SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
No. 12-048841- -00000-0000
Fecha: 2012-03-22 16:25:59 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 498
Act. 411 PRESENTACION Folios: 499

Industria y Comercio
SUPERINTENDENCIA

DELEGATURA DE PROPIEDAD INDUSTRIAL

División de Nuevas Creaciones

SOLICITUD

PATENTE DE INVENCIÓN

21. **EXPEDIENTE No.** _____

54. **TÍTULO** NUEVAS ENZIMAS DE HIDROLASA DE GLUCOLISIS Y SUS APLICACIONES

51. **CLASIFICACIÓN INTERNACIONAL** _____

71. **SOLICITANTE** DANISCO US INC

DOMICILIO 925 Page Mill Road, Palo Alto, California Estados Unidos de América

74. **APODERADO** JERONIMO CASTILLO BONILLA

22. **BOGOTÁ, D. C.,** _____

(FORMA P-10)

115198-Imprenta Nacional de Colombia-2011

Industria y Comercio
SUPERINTENDENCIA

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
No. 12-048841- -00000-0000
Fecha: 2012-03-22 16:25:59 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 498
Act. 411 PRESENTACION Folios: 499

FORMULARIO ÚNICO DE SOLICITUD DE PATENTES
PCT
ENTRADA FASE NACIONAL

SOLICITUD DE REGISTRO DE:

☒ Patente de Invención ☐ Patente Modelo de Utilidad

☒ Capítulo I ☐ Capitulo II

SOLICITANTES	Nombre: **DANISCO US INC** Dirección: 925 Page Mill Road, Palo Alto, California 94304, Estados Unidos de América Domicilio: Estadounidense Lugar de Constitución: Delaware Teléfono: Fax: E - mail:	**IDENTIFICACIÓN** ☐ C.C. ☐ NIT ☐ C.E. ☐ Otro Cual: _____ Número: _____
REPRESENTANTE O APODERADO	Nombre: **JERÓNIMO CASTILLO BONILLA** Dirección: Calle 93A Nº 14-17, Ofc 311 Teléfono: 2363493 Fax: 2182351 E-mail: info@pcastillopineda.com	**IDENTIFICACIÓN** ☒ C.C. ☐ NIT ☐ C.E. ☐ Otro Número: 79462057 TP: 104167
INVENTORES	Nombre: BOWER, Benjamín Dirección: 925 Page Mill Road, Palo Alto, California 94304, Estados Unidos de America. Nacionalidad o Domicilio: Estadounidense	

PCT
Solicitud Internacional No.: PCT/US2010/049849 Fecha: Septiembre 22, 2010
Publicación Internacional No.: WO 2011/038019 A2 Fecha: Marzo 31, 2011

Titulo:

NUEVAS ENZIMAS DE HIDROLASA DE GLUCÓLISIS Y SUS APLICACIONES.

Clasificación Internacional:

Prioridad
☒ SI ☐ NO

País de Origen	Fecha	Número de Solicitud
US	Septiembre 23, 2009	61/245.273
US	Diciembre 23, 2009	61/289.886

Comprobante de pago No: **Fecha:**

Anexos

- ☒ Comprobante de pago de la tasa de presentación de la solicitud.
- ☒ Documento que acredita la existencia y representación legal de la persona jurídica peticionaria, en cuanto no estén inscritas en alguna Cámara de Comercio Colombiana.
- ☒ Poderes, si fuere el caso.
- ☒ Documento de cesión del inventor al solicitante o a su causante.
- ☒ Traducción de la solicitud internacional al castellano (Resumen, descripción, reivindicaciones, dibujos)
- ☐ Copia del examen preliminar internacional efectuado por la Autoridad Internacional de Examen preliminar (IPEA)
- ☐ Traducción de examen preliminar internacional efectuado por la IPEA
- ☐ De ser el caso, copia del contrato de acceso.
- ☐ De ser el caso, documento que acredite la licencia o autorización de uso de conocimientos tradicionales de las comunidades indígenas.
- ☐ De ser el caso, certificado de depósito del material biológico
- ☒ Arte final 12x12.
- ☐ De ser el caso, información sobre otras solicitudes de patente o títulos obtenidos en el extranjero por el mismo titular o su causante, relacionadas parcial o totalmente con la invención de esta solicitud.
- ☐ De ser el caso, modificaciones efectuadas a la solicitud internacional.
- ☐ Otros.

11. FIGURA CARACTERÍSTICA:

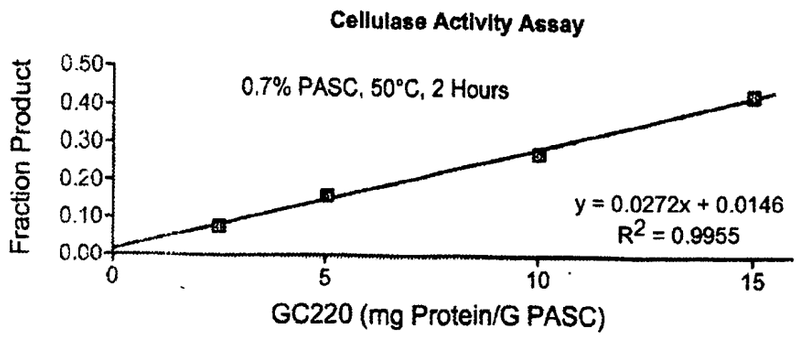


FIG. 25

12. Solicito la concesión de la patente

NOMBRE: JERÓNIMO CASTILLO B.

FIRMA:

C.C. 79.462.057 T.P. 104167

Instrucciones para el diligenciamiento del presente formulario, ver reverso de esta página

(Continuación Inventores)

Nombre : HIS, Megan Yee
Dirección : 925 Page Mill Road, Palo Alto, California 94304, Estados Unidos de América
Nacionalidad : estadounidense

Nombre : KAPER, Thijs.
Dirección : 925 Page Mill Road, Palo Alto, California 94304, Estados Unidos de América
Nacionalidad : holandesa

Nombre : KELEMEN, Bradley R.
Dirección : 925 Page Mill Road, Palo Alto, California 94304, Estados Unidos de América
Nacionalidad : estadounidense

Nombre : LANTZ, Suzanne E
Dirección : 925 Page Mill Road, Palo Alto, California 94304, Estados Unidos de América
Nacionalidad : estadounidense

Nombre : LARENAS, Edmund A
Dirección : 925 Page Mill Road, Palo Alto, California 94304, Estados Unidos de América
Nacionalidad : estadounidense

Nombre : MITCHINSON, Colin
Dirección : 925 Page Mill Road, Palo Alto, California 94304, Estados Unidos de América
Nacionalidad : británica

Nombre : KIM, Steven
Dirección : 925 Page Mill Road, Palo Alto, California 94304, Estados Unidos de América
Nacionalidad : estadounidense

Nombre : HITZ, William D.
Dirección : 404 Hillside Road, Wilmington, Delaware 19807, Estados Unidos de América
Nacionalidad : estadounidense

Nombre : EMPTAGE, Mark
Dirección : 2822 Lardon Drive, Wilmington, Delaware 19810, Estados Unidos de América
Nacionalidad : estadounidense

Nombre : WING, Keith Dumont
Dirección : 122 Yardley Lane, Wilmington, Delaware 19810, Estados Unidos de América
Nacionalidad : estadounidense

State of California

SECRETARY OF STATE

Requested for use in Colombia.
Not for use within the United States of America.
The purpose of the Apostille is to certify the authenticity of the signature of the official signing the document, the capacity in which the official signing the document has acted, and, where appropriate, the identity of the seal or stamp

APOSTILLE

(Convention de La Haye du 5 octobre 1961)

1. *Country: United States of America*
   This public document

2. *has been signed by Carol A. See*

3. *acting in the capacity of Notary Public, State of California*

4. *bears the seal/stamp of Carol A. See, Notary Public.*
   State of California

CERTIFIED

5 *At Sacramento California*

6. *the 1st day of September 2009*

7. *by Deputy Secretary of State, State of California*

8. *No. 723031*

9. *Seal/Stamp:*

10. *Signature*

Debra Bowen
Secretary of State

BY _____

NP-24 A (REV. 1-07)

OSP 06 99602

FOR PATENTS AND TRADE MARKS IN
COLOMBIA

Poder General

1 Applicant's name (1)

2 Applicant's address domiciliado en (2)

nombra al Dr. JERONIMO CASTILLO BONILLA y/o Dra. XIMENA BONILLA PEREIRA y/o Dr. ALVARO CASTILLO GRAU Calle 93A N° 14-17 Oficina 311 en Bogotá, Colombia, como sus representantes con poder especial, amplio y suficiente, para recabar de las oficinas y autoridades Colombianas, administrativas, judiciales o de policía, en primera o segunda instancia, la obtención de Patentes de Invención, Registro de Marcas, de Modelos y de Dibujos Industriales, de Nombres Comerciales y Enseñas, Lemas Comerciales y Denominaciones de Origen, lo mismo que traspasos, extensiones, renovaciones, cambios de nombre, registro de licencias de explotación y registro de licencias de uso referentes a tales actuaciones, cancelación de Registros de marca, a cuyo efecto les faculta para dar ante dichas autoridades todos los pasos necesarios al objeto indicado, elevar solicitudes, formular descripciones, enmiendas, protestas, declaraciones, apelaciones y reclamos, solicitar la aplicación de medidas cautelares; solicitar la nulidad de Patentes de Invención, modelos de utilidad, diseños industriales, así como la de registros de marcas ante la autoridad nacional competente; aceptar en nuestro nombre y representación la Cesión de derechos de invención y de patente de invención que se nos haga por cualquier persona, abonar todos los impuestos, cuotas y pagos determinados por la ley, recibir todos los documentos y valores dando el descargo respectivo, llenar cualesquiera otros requisitos, desistir y tomar en fin todas las medidas que creyeren conducentes al resguardo de nuestros intereses y en caso de presentarse oposiciones y/o observaciones para que intervengan como demandantes o como demandados, con todas las demás facultades necesarias, así como iniciar y tramitar procesos jurisdiccionales y/o administrativos relativos a competencia desleal. Este poder comprende además las facultades de recibir, desistir, transigir, comprometer, sustituir, revocar sustituciones, conciliar, renunciar, recibir notificaciones y nombrar apoderados judiciales o extrajudiciales.

3 Date
4 Applicant's signature

Dado y firmado en (3)

General Power of Attorney

(1) Danisco US Inc.

domiciled in (2)
925 Page Mill Road, Palo Alto
California 94304
U.S.A.

hereby appoint, JERONIMO CASTILLO BONILLA and/or XIMENA BONILLA PEREIRA and/or ALVARO CASTILLO GRAU, Calle 93A N° 14-17 Office 311 en Bogotá, Colombia, as our representatives with special, ample and sufficient power to obtain from the offices and Colombian authorities, administrative, judicial and police, in first or in second instance, Privileged Patents of invention, Registration of Trade Marks, of Models and Industrial Drawings, of Commercial Names and ensigns, Commercial slogans and Origin Denominations as well as assignments, extensions, renewals, changes of name, registration of license of exploitation and of license of use referred to said actions, Cancellation Actions against trademark Registrations, who are empowered to take before said authorities all of the necessary steps for the purposes indicated, to file applications, to formulate descriptions, corrections, protests, declarations, appeals, and claims, apply for precautionary measures; apply for the nullity of Patent Inventions, Utility Models, Industrial Designs, as well as for Trademark Registrations before the competent national authority; to accept on our behalf and representation the assignment of rights of inventions and titles of invention that may be done to us by any person, to pay all taxes, quotas and payments prescribed by law, to receive all the documents and proceeds giving the respective acquittance or receipt, to fulfill any other requisites to desist and take in fact any other measures that may be deemed conducive to the safeguard of our interests, and in case that oppositions and/or observations are presented to intervene as complainant and defendant, with all necessary legal faculties, as well as to file jurisdictional processes and/or administrative related to disloyal competition.
This power of attorney includes the faculties to receive, desist, compromise, transact substitute it in all or in part, to revoke substitutions, to conciliate, to renounce, to receive notifications and to appoint extra-judicial or judicial attorneys.

Given and signed in (3)
Palo Alto, California, U.S.A.
27th day of August, 2009

4)Signature (s)/
X Soon Lee
Soonhee Jang

DECLARACION NOTARIAL
(Individual)

A los días del mes de
compareció y/o comparecieron

personalmente ante mi, Notario Público

quien (es) doy fe de conocer y me consta que es (son) la(s) persona(s) descrita(s) y firmante(s) del documento que precede, declarando ante mi que otorga(n) el mismo para los fines y propósitos que en él se expresan.

Notario Público,

.....

NOTARIAL CERTIFICATE
(Individual)

On this day of of
personally appeared before me

a Notary Public

to me known, and known to me to be the person(s) described in and who executed the foregoing document, and he (they) duly acknowledged to me that he (they) executed same for the uses and purposes therein expresed.

.....
Notary Public.

(Sociedades)

El suscrito Notario Certifica: que la firma que antecede y dice

es auténtica; que el signatario desempeña actualmente el cargo de

de,

que tiene facultad para otorgar este poder; y que dicha compañía esta domiciliada en

y actualmente existe y esta organizada de acuerdo a las leyes de .

Todos los hechos anteriores le constan por haber examinado los documentos respectivos y de todo ello da fe.

Dado y firmado en

El Notario Público,

.....

(Con legalización Consular)

(Corporations)

The undersigned Notary Public certifies: that the preceding signature reading
Soonhee Jang

is authentic; that the signer at present fills the post of Vice President, IP Strategy & Chief IP Counsel

of Danisco US Inc.

and that she is empowered to grant this power; and that said company is domiciled in
Palo Alto, California, U.S.A.

and actually exists and is organized in accordance with the laws of Delaware, USA .

All the foregoing facts are known to the Notary due to her having examined the respective documents to which she, the Notary attests..

Given and signed in Palo Alto, California, USA
Carol A. See, Notary Public

Carol A. See
.....
Notary Public.

(With Consular legalization)

State of California
Secretary of State

APOSTILLE (Convention de La Haye du 5 octobre 1961)			
1. Country: Pays / País:	United States of America		
This public document Le présent acte public / El presente documento público			
2. has been signed by a été signé par ha sido firmado por	*Aida M. Martin*		
3. acting in the capacity of agissant en qualité de quien actúa en calidad de	*Notary Public, State of California*		
4. bears the seal / stamp of est revêtu du sceau / timbre de y está revestido del sello / timbre de	*Aida M. Martin, Notary Public, State of California*		
Certified Attesté / Certificado			
5. at à / en	Sacramento, California	**6. the** le / el día	*9th day of March 2012*
7. by par / por	Secretary of State, State of California		
8. N° sous n° bajo el número	*36097*		
9. Seal / stamp: Sceau / timbre: Sello / timbre:		**10. Signature:** Signature: Firma:	

This Apostille is the trilingual model Apostille Certificate as suggested by the Permanent Bureau and developed in response to the 2009 Special Commission on the practical operation of the Hague Apostille Convention.
This Apostille only certifies the authenticity of the signature and the capacity of the person who has signed the public document, and, where appropriate, the identity of the seal or stamp which the public document bears.
This Apostille does not certify the content of the document for which it was issued.
This Apostille is not valid for use anywhere within the United States of America, its territories or possessions.
To verify the issuance of this Apostille, see: www.sos.ca.gov/business/notary/apostille-search/.

Cette apostille est le modèle d'Apostille trilingue tel que suggéré par le Bureau Permanent et élaboré en réponse à la Commission spéciale de 2009 sur le fonctionnement pratique de la Convention de La Haye Apostille.
Cette Apostille atteste uniquement la véracité de la signature, la qualité en laquelle le signataire de l'acte a agi et, le cas échéant, l'identité du sceau ou timbre dont cet acte public est revêtu.
Cette Apostille ne certifie pas le contenu de l'acte pour lequel elle a été émise.
L'utilisation de cette Apostille n'est pas valable en / au États-Unis d'Amérique, ses territoires où possessions.
Cette Apostille peut être vérifiée à l'adresse suivante: www.sos.ca.gov/business/notary/apostille-search/.

Esta apostilla es el modelo trilingüe Certificado de Apostilla según lo sugerido por la Oficina Permanente y desarrollado en respuesta a la Comisión especial de 2009 sobre el funcionamiento práctico del Convenio de La Haya sobre Apostilla
Esta Apostilla certifica únicamente la autenticidad de la firma, la calidad en que el signatario del documento haya actuado y, en su caso, la identidad del sello o timbre del que el documento público esté revestido.
Esta Apostilla no certifica el contenido del documento para el cual se expidió.
No es válido el uso de esta Apostilla en Estados Unidos de América, sus territorios o posesiones.
Esta Apostilla se puede verificar en la dirección siguiente: www.sos.ca.gov/business/notary/apostille-search/.

I hereby certify that the attached document is a true and correct copy of the recorded Worldwide Assignment of rights from E.I. Du Pont De Nemours and Company to Danisco US Inc. for priority applications USSN 61/289,886 and USSN 61/245,273.

IN WITNESS WHEREOF I undersign as follows:

Date: March 5, 2012

Fabesha Greene

STATE OF CALIFORNIA
COUNTY OF SANTA CLARA

On March 5, 2012, before me, Aida M. Martin , Notary Public, personally appeared, Fabesha Greene, who proved to me on the basis of satisfactory evidence to be the person(s) whose name(s) is/are subscribed to the within instrument and acknowledged to me that he/she/they executed the same in his/her/their authorized capacity(ies), and that by his/her/their signature(s) on the instrument the person(s), or the entity upon behalf of which the person(s) acted, executed the instrument. I certify under PENALTY OF PERJURY under the laws of the State of California that the foregoing paragraph is true and correct.

WITNESS my hand and official seal.

Signature

(Seal)

AIDA M. MARTIN
Commission # 1871469
Notary Public - California
Santa Clara County
My Comm. Expires Dec 13, 2013

THE STATE OF

31376_CO_ASSIGN

UNITED STATES PATENT AND TRADEMARK OFFICE

Facsimile Transmission

To:	**Name:**	DANISCO US INC.
	Company:	925 PAGE MILL ROAD
	Fax Number:	16508456504
	Voice Phone:	
From:	**Name:**	ASSIGNMENT SERVICES BRANCH
	Voice Phone:	571-272-3350

37 C.F.R. 1.6 sets forth the types of correspondence that can be communicated to the Patent and Trademark Office via facsimile transmissions. Applicants are advised to use the certificate of facsimile transmission procedures when submitting a reply to a non-final or final Office action by facsimile (37 CFR 1.8(a)).

Fax Notes:

Pg#	Description
1	Cover Page
2	100.TXT
4	Document 1, Batch 2183254

USPTO ASSIGNMENT SYSTEM PROCESSING

Date and time of transmission: Tuesday, September 21, 2010 9:31:44 PM
Number of pages including this cover sheet: 04

8

UNITED STATES PATENT AND TRADEMARK OFFICE

UNDER SECRETARY OF COMMERCE FOR INTELLECTUAL PROPERTY AND
DIRECTOR OF THE UNITED STATES PATENT AND TRADEMARK OFFICE

501297140A

SEPTEMBER 21, 2010

PTAS

DANISCO US INC.
925 PAGE MILL ROAD
WENFANG CHEN
PALO ALTO, CA 94304

UNITED STATES PATENT AND TRADEMARK OFFICE
NOTICE OF RECORDATION OF ASSIGNMENT DOCUMENT

THE ENCLOSED DOCUMENT HAS BEEN RECORDED BY THE ASSIGNMENT DIVISION OF THE U.S. PATENT AND TRADEMARK OFFICE. A COMPLETE MICROFILM COPY IS AVAILABLE AT THE ASSIGNMENT SEARCH ROOM ON THE REEL AND FRAME NUMBER REFERENCED BELOW.

PLEASE REVIEW ALL INFORMATION CONTAINED ON THIS NOTICE. THE INFORMATION CONTAINED ON THIS RECORDATION NOTICE REFLECTS THE DATA PRESENT IN THE PATENT AND TRADEMARK ASSIGNMENT SYSTEM. IF YOU SHOULD FIND ANY ERRORS OR HAVE QUESTIONS CONCERNING THIS NOTICE, YOU MAY CONTACT THE EMPLOYEE WHOSE NAME APPEARS ON THIS NOTICE AT 571-272-3350. PLEASE SEND REQUEST FOR CORRECTION TO: U.S. PATENT AND TRADEMARK OFFICE, MAIL STOP: ASSIGNMENT SERVICES BRANCH, P.O. BOX 1450, ALEXANDRIA, VA 22313.

RECORDATION DATE: 09/21/2010 REEL/FRAME: 025021/0683
NUMBER OF PAGES: 3

BRIEF: ASSIGNMENT OF ASSIGNOR'S INTEREST (SEE DOCUMENT FOR DETAILS).
DOCKET NUMBER: 31376USP2 -USP3

ASSIGNOR:
E.I. DU PONT DE NEMOURS AND COMPANY DOC DATE: 09/20/2010

ASSIGNEE:
DANISCO US INC.
925 PAGE MILL ROAD
PALO ALTO, CALIFORNIA 94304

SERIAL NUMBER: 61245273 FILING DATE: 09/23/2009
PATENT NUMBER: ISSUE DATE:
TITLE: NOVEL GLYCOSYL HYDROLASE ENZYMES AND USES THEREOF

SERIAL NUMBER: 61289886 FILING DATE: 12/23/2009
PATENT NUMBER: ISSUE DATE:
TITLE: NOVEL GLYCOSYL HYDROLASE ENZYMES AND USES THEREOF

9

025021/0683 PAGE 2

ASSIGNMENT SERVICES BRANCH
PUBLIC RECORDS DIVISION

PATENT ASSIGNMENT

Electronic Version v1.1
Stylesheet Version v1.1

09/21/2010
501297140

SUBMISSION TYPE:	NEW ASSIGNMENT
NATURE OF CONVEYANCE:	ASSIGNMENT

CONVEYING PARTY DATA

Name	Execution Date
E.I. du Pont de Nemours and Company	09/20/2010

RECEIVING PARTY DATA

Name:	Danisco US Inc.
Street Address:	925 Page Mill Road
City:	Palo Alto
State/Country:	CALIFORNIA
Postal Code:	94304

PROPERTY NUMBERS Total: 2

Property Type	Number
Application Number:	61245273
Application Number:	61289886

CH $80.00 61245273

CORRESPONDENCE DATA

Fax Number: (650)845-6504
Correspondence will be sent via US Mail when the fax attempt is unsuccessful.
Phone: 650-846-7500
Email: aida.martin@danisco.com
Correspondent Name: Danisco US Inc.
Address Line 1: 925 Page Mill Road
Address Line 2: Wenfang Chen
Address Line 4: Palo Alto, CALIFORNIA 94304

ATTORNEY DOCKET NUMBER:	31376USP2 -USP3
NAME OF SUBMITTER:	Wenfang Chen

Total Attachments: 2
source=DuPont Assignment#page1.tif
source=DuPont Assignment#page2.tif

ASSIGNMENT

E. I. du Pont de Nemours and Company ("DuPont") hereby declares that WILLIAM D. HITZ, MARK EMPTAGE and KEITH DUMONT WING are under an obligation to assign and have assigned to DuPont all their right title and an interest in an invention entitled NOVEL GLYCOSYL HYDROLASE ENZYMES AND USES THEREOF which is disclosed in United States Patent Application No. 61/245,273, filed on September 23, 2009 and United States Patent Application No. 61/289,886, filed on December 23, 2009 and which are collectively identified as Case Number IDN 31376; and for valuable consideration, the receipt and adequacy of which is hereby acknowledged and in fulfillment of pre-existing obligation of assignment, E. I. du Pont de Nemours and Company hereby:

I. Assigns and transfers unto **Danisco US Inc.,** a corporation organized and existing under and by virtue of the laws of the State of Delaware, having its principal place of business at 925 Page Mill Road, Palo Alto, California 94304-1013,: (1) the entire interest throughout the world in and to the aforesaid Applications for Letters Patent including any priority rights derived from the aforesaid Applications for Letters Patent by virtue of the U.S. patent law, including all divisions, continuations, continuations-in-part, substitute applications, renewals, reissues, re-examinations and all foreign counterparts thereof; or the International Convention for the Protection of Industrial Property for any and all member countries of the aforesaid International Convention or other treaty or understanding;; (2) the entire interest in and to any and all inventions disclosed in the aforesaid Applications for inventions disclosed in the aforesaid Applications for Letters Patent; (3) any and all patent applications corresponding to the aforesaid U.S. Applications for Letters Patent; and (4) the sole right: (i) to file any additional applications; (ii) file such applications under the aforesaid International Convention or other treaty or understanding; (iii) have said patents granted in its name or ours; and (iv) enforce said patents and to sue for and recover profits and damages for any and all infringements thereof whether past or future; and

II. Agree, whenever requested, to communicate to said assignees, their successors, assigns, and legal representatives, any facts known to us respecting said rights, to testify in any legal proceeding respecting said rights, to execute all applications, papers or instruments necessary or required by said assignees, their successors, assigns and legal representatives, to carry into effect any of the provisions of this instrument, and generally to do everything possible to aid said assignees, their successors, assigns, and legal representatives to obtain and enforce proper patent protection for said inventions in any and all countries.

E. I. du Pont de Nemours and Company

By Roger A. Bowman

Name ROGER A. BOWMAN

Title Assistant Secretary - Patent Board

Date September 17, 2010

RECEIVED BY:

Danisco US Inc

By _____

Name Michael V. Arbige

Title Exec V.P. Technology

Date September 20, 2010

State of California
Secretary of State

APOSTILLE (Convention de La Haye du 5 octobre 1961)			
1. Country: Pays / País:	United States of America		
This public document Le présent acte public / El presente documento público			
2. has been signed by a été signé par ha sido firmado por	*Aida M. Martin*		
3. acting in the capacity of agissant en qualité de quien actúa en calidad de	*Notary Public, State of California*		
4. bears the seal / stamp of est revêtu du sceau / timbre de y está revestido del sello / timbre de	*Aida M. Martin, Notary Public, State of California*		
Certified Attesté / Certificado			
5. at à / en	Sacramento, California	**6. the** le / el día	*9th day of March 2012*
7. by par / por	Secretary of State, State of California		
8. N° sous n° bajo el número	*92293*		
9. Seal / stamp: Sceau / timbre: Sello / timbre:		**10. Signature:** Signature: Firma:	

This Apostille is the trilingual model Apostille Certificate as suggested by the Permanent Bureau and developed in response to the 2009 Special Commission on the practical operation of the Hague Apostille Convention.
This Apostille only certifies the authenticity of the signature and the capacity of the person who has signed the public document, and, where appropriate, the identity of the seal or stamp which the public document bears.
This Apostille does not certify the content of the document for which it was issued.
This Apostille is not valid for use anywhere within the United States of America, its territories or possessions.
To verify the issuance of this Apostille, see: www.sos.ca.gov/business/notary/apostille-search/.

Cette apostille est le modèle d'Apostille trilingue tel que suggéré par le Bureau Permanent et élaboré en réponse à la Commission spéciale de 2009 sur le fonctionnement pratique de la Convention de La Haye Apostille.
Cette Apostille atteste uniquement la véracité de la signature, la qualité en laquelle le signataire de l'acte a agi et, le cas échéant, l'identité du sceau ou timbre dont cet acte public est revêtu.
Cette Apostille ne certifie pas le contenu de l'acte pour lequel elle a été émise.
L'utilisation de cette Apostille n'est pas valable en / au États-Unis d'Amérique, ses territoires où possessions.
Cette Apostille peut être vérifiée à l'adresse suivante: www.sos.ca.gov/business/notary/apostille-search/.

Esta apostilla es el modelo trilingüe Certificado de Apostilla según lo sugerido por la Oficina Permanente y desarrollado en respuesta a la Comisión especial de 2009 sobre el funcionamiento práctico del Convenio de La Haya sobre Apostilla
Esta Apostilla certifica únicamente la autenticidad de la firma, la calidad en que el signatario del documento haya actuado y, en su caso, la identidad del sello o timbre del que el documento público esté revestido.
Esta Apostilla no certifica el contenido del documento para el cual se expidió.
No es válido el uso de esta Apostilla en Estados Unidos de América, sus territorios o posesiones.
Esta Apostilla se puede verificar en la dirección siguiente: www.sos.ca.gov/business/notary/apostille-search/.

I hereby certify that the attached document is a true and correct copy of the recorded Worldwide Assignment of rights for priority application USSN 61/289,886.

IN WITNESS WHEREOF I undersign as follows:

Date: March 5, 2012

Fabesha Greene

STATE OF CALIFORNIA
COUNTY OF SANTA CLARA

On March 5, 2012 before me, Aida M. Martin, Notary Public, personally appeared, Fabesha Greene, who proved to me on the basis of satisfactory evidence to be the person(s) whose name(s) is/are subscribed to the within instrument and acknowledged to me that he/she/they executed the same in his/her/their authorized capacity(ies), and that by his/her/their signature(s) on the instrument the person(s), or the entity upon behalf of which the person(s) acted, executed the instrument. I certify under PENALTY OF PERJURY under the laws of the State of California that the foregoing paragraph is true and correct.

WITNESS my hand and official seal.

Signature

(Seal)

AIDA M. MARTIN
Commission # 1871469
Notary Public - California
Santa Clara County
My Comm. Expires Dec 13, 2013

THE OF STATE

31376_CO_assn886

15

UNITED STATES PATENT AND TRADEMARK OFFICE

Facsimile Transmission

To:	**Name:**	DANISCO US INC.
	Company:	925 PAGE MILL ROAD
	Fax Number:	16508456504
	Voice Phone:	
From:	**Name:**	ASSIGNMENT SERVICES BRANCH
	Voice Phone:	571-272-3350

37 C.F.R. 1.6 sets forth the types of correspondence that can be communicated to the Patent and Trademark Office via facsimile transmissions. Applicants are advised to use the certificate of facsimile transmission procedures when submitting a reply to a non-final or final Office action by facsimile (37 CFR 1.8(a)).

Fax Notes:

Pg#	Description
1	Cover Page
2	170.TXT
4	Document 1, Batch 2174675

USPTO ASSIGNMENT SYSTEM PROCESSING

Date and time of transmission: Wednesday, September 15, 2010 4:38:16 PM
Number of pages including this cover sheet: 05

16

UNITED STATES PATENT AND TRADEMARK OFFICE

UNDER SECRETARY OF COMMERCE FOR INTELLECTUAL PROPERTY AND
DIRECTOR OF THE UNITED STATES PATENT AND TRADEMARK OFFICE

501289666A

SEPTEMBER 15, 2010

PTAS

DANISCO US INC.
925 PAGE MILL ROAD
WENFANG CHEN
PALO ALTO, CA 94304

UNITED STATES PATENT AND TRADEMARK OFFICE
NOTICE OF RECORDATION OF ASSIGNMENT DOCUMENT

THE ENCLOSED DOCUMENT HAS BEEN RECORDED BY THE ASSIGNMENT DIVISION OF THE U.S. PATENT AND TRADEMARK OFFICE. A COMPLETE MICROFILM COPY IS AVAILABLE AT THE ASSIGNMENT SEARCH ROOM ON THE REEL AND FRAME NUMBER REFERENCED BELOW.

PLEASE REVIEW ALL INFORMATION CONTAINED ON THIS NOTICE. THE INFORMATION CONTAINED ON THIS RECORDATION NOTICE REFLECTS THE DATA PRESENT IN THE PATENT AND TRADEMARK ASSIGNMENT SYSTEM. IF YOU SHOULD FIND ANY ERRORS OR HAVE QUESTIONS CONCERNING THIS NOTICE, YOU MAY CONTACT THE EMPLOYEE WHOSE NAME APPEARS ON THIS NOTICE AT 571-272-3350. PLEASE SEND REQUEST FOR CORRECTION TO: U.S. PATENT AND TRADEMARK OFFICE, MAIL STOP: ASSIGNMENT SERVICES BRANCH, P.O. BOX 1450, ALEXANDRIA, VA 22313.

RECORDATION DATE: 09/14/2010 REEL/FRAME: 024983/0402
NUMBER OF PAGES: 6

BRIEF: ASSIGNMENT OF ASSIGNOR'S INTEREST (SEE DOCUMENT FOR DETAILS).
DOCKET NUMBER: 31376USP3

ASSIGNOR:
BOWER, BENJAMIN S DOC DATE: 09/10/2010

ASSIGNOR:
HSI, MEGAN YEE DOC DATE: 09/03/2010

ASSIGNOR:
KAPER, THIJS DOC DATE: 09/03/2010

ASSIGNOR:
KELEMAN, BRADLEY R. DOC DATE: 09/03/2010

ASSIGNOR:
LANTZ, SUZANNE E. DOC DATE: 09/10/2010

ASSIGNOR:
LARENAS, EDMUND A DOC DATE: 09/08/2010

17

024983/0402 PAGE 2

ASSIGNOR:
MITCHINSON, COLIN DOC DATE: 09/10/2010

ASSIGNEE:
DANISCO US INC.
925 PAGE MILL ROAD
PALO ALTO, CALIFORNIA 94304

SERIAL NUMBER: 61289886 FILING DATE: 12/23/2009
PATENT NUMBER: ISSUE DATE:
TITLE: NOVEL GLYCOSYL HYDROLASE ENZYMES AND USES THEREOF

ASSIGNMENT SERVICES BRANCH
PUBLIC RECORDS DIVISION

PATENT ASSIGNMENT

Electronic Version v1.1 09/14/2010
Stylesheet Version v1.1 501289666

SUBMISSION TYPE:	NEW ASSIGNMENT
NATURE OF CONVEYANCE:	ASSIGNMENT

CONVEYING PARTY DATA

Name	Execution Date
Benjamin S Bower	09/10/2010
Megan Yee Hsi	09/03/2010
Thijs Kaper	09/03/2010
Bradley R. Keleman	09/03/2010
Suzanne E. Lantz	09/10/2010
Edmund A Larenas	09/08/2010
Colin Mitchinson	09/10/2010

RECEIVING PARTY DATA

Name:	Danisco US Inc.
Street Address:	925 Page Mill Road
City:	Palo Alto
State/Country:	CALIFORNIA
Postal Code:	94304

PROPERTY NUMBERS Total: 1

Property Type	Number
Application Number:	61289886

CH $40.00 61289886

CORRESPONDENCE DATA

Fax Number: (650)845-6504
Correspondence will be sent via US Mail when the fax attempt is unsuccessful.
Phone: 650-846-7500
Email: aida.martin@danisco.com
Correspondent Name: Danisco US Inc.
Address Line 1: 925 Page Mill Road
Address Line 2: Wenfang Chen
Address Line 4: Palo Alto, CALIFORNIA 94304

ATTORNEY DOCKET NUMBER:	31376USP3

NAME OF SUBMITTER:	Wenfang Chen

Total Attachments: 4
source=31376USP3-Assignment#page1.tif
source=31376USP3-Assignment#page2.tif
source=31376USP3-Assignment#page3.tif
source=31376USP3-Assignment#page4.tif

Assignment

WHEREAS, **Benjamin S. Bower**, residing at 8479 Mayhews Landing Road, Newark, CA 94560, **Megan Yee Hsi**, residing at 80 Descanso Drive, Unit 2101, San Jose, CA 95134, **Thijs Kaper,** residing at 670 Myrtle Street, Half Moon Bay, Ca 94019, **Bradley R. Keleman**, residing at 658 12th Avenue, Menlo Park, CA 94040, **Suzanne E. Lantz**, residing at 2072 Eaton Avenue, San Carlos, CA 94070, **Edmund A. Larenas**, residing at 301 Nevada. Moss Beach, CA 94038, and **Colin Mitchinson**, residing at 381 Myrtle Street, Half Moon Bay, CA 94019, have invented a new and useful invention entitled **Novel Glycosyl Hydrolase Enzymes and Uses Thereof** and an application for United States Letters Patent was filed on December 23, 2009, United States Provisional Serial No. 61/289,886 and

WHEREAS, **Danisco US Inc.,** a corporation organized and existing under and by virtue of the laws of the State of Delaware, having its principal place of business at 925 Page Mill Road, Palo Alto, California 94304-1013, is desirous of acquiring an interest in and to said invention, and in the Letters Patents to be obtained therefor;

NOW, THEREFORE, be it known by all whom it may concern:

That for good and valuable consideration the receipt of which is hereby acknowledged, the said inventors have and do hereby sell, assign, transfer and set over unto the said Danisco US Inc., its successors and assigns, the full and exclusive right, title and interest, including all rights under the Paris Convention for the Protection of Industrial Property, in and to said invention, and in and to any and all Letters Patents to be granted and issued therefor, not only for, to and in the United States of America, its territories and possessions, but for, to and in all other countries, and it has been and is hereby authorized and requested that the appropriate government agencies issue said Letters Patents to said Danisco US Inc., . in accordance with this Assignment.

Said Inventor/Inventors hereby covenant/covenants and agree/agrees to cooperate with Danisco US Inc. to enable said Danisco US Inc. to enjoy to the fullest extent the right, title and interest herein conveyed in the United States and foreign countries. Such cooperation by said Inventors shall include prompt production of pertinent facts and documents, giving of testimony, execution of petitions, oaths, specifications, declarations or other papers and other assistance, all to the extent deemed necessary or desirable by said Danisco US Inc. for (a) perfecting the right, title and interest herein conveyed, (b) prosecuting any of said applications, (c) filing and prosecuting applications for reissuance of any said patents, (d) interference or other priority proceedings involving said invention, and (e) legal proceedings involving said invention and any

12

applications therefor and any patents granted thereon, including without limitation, opposition proceedings, cancellation proceedings, priority contests, public use proceedings, infringement actions and court actions, provided, however, that the expense incurred by said Inventors in providing such cooperation shall be paid for by said Danisco US Inc.

The terms and covenants of this Assignment shall inure to the benefit of said Danisco US Inc., its successors, assigns and other legal representatives, and shall be binding upon said Inventors, his/her respective heirs, legal representatives and assigns.

Said Inventor hereby warrants and represents that he/she has not entered and will not enter into any assignment, contract or understanding in conflict herewith.

IN WITNESS WHEREOF I undersign as follows:

Signed at Palo Alto, this 10 the day of Sept. 2010.

By: Benjamin S. Bower
Benjamin S. Bower

WITNESSED BY:

Signature: Aida Martin

Printed Name of Witness: Aida Martin

IN WITNESS WHEREOF I undersign as follows:

Signed at Palo Alto, this 3 the day of Sept 2010.

By: Megan Yee Hsi
Megan Yee Hsi

WITNESSED BY:

Signature: Aida Martin

Printed Name of Witness: Aida Martin

IN WITNESS WHEREOF I undersign as follows:

Signed at Palo Alto, this 3rd the day of Sept 2010.

By: _____
Thijs Kaper

WITNESSED BY:

Signature: _____

Printed Name of Witness: Aida Martin

IN WITNESS WHEREOF I undersign as follows:

Signed at Palo Alto, this 3 the day of September 2010.

By: _____
Bradley R. Keleman

WITNESSED BY:

Signature: _____

Printed Name of Witness: Aida Martin

IN WITNESS WHEREOF I undersign as follows:

Signed at Palo Alto, this 10 the day of Sept. 2010.

By: _____
Suzanne E. Lantz

WITNESSED BY:

Signature: _____

Printed Name of Witness: Aida Martin

IN WITNESS WHEREOF I undersign as follows:

Signed at Palo Alto, this 8 the day of September 2010.

By: _____
Edmund A. Larenas

WITNESSED BY:

Signature: _____

Printed Name of Witness: Aida Martin

IN WITNESS WHEREOF I undersign as follows:

Signed at Palo Alto, this 10 the day of Sept, 2010.

By: _____
Colin Mitchinson

WITNESSED BY:

Signature: _____

Printed Name of Witness: Aida Martin

State of California
Secretary of State

APOSTILLE (Convention de La Haye du 5 octobre 1961)			
1. Country: Pays / País:	United States of America		
This public document Le présent acte public / El presente documento público			
2. has been signed by a été signé par ha sido firmado por	*Aida M. Martin*		
3. acting in the capacity of agissant en qualité de quien actúa en calidad de	*Notary Public, State of California*		
4. bears the seal / stamp of est revêtu du sceau / timbre de y está revestido del sello / timbre de	*Aida M. Martin, Notary Public, State of California*		
Certified Attesté / Certificado			
5. at à / en	Sacramento, California	**6. the** le / el día	*9th day of March 2012*
7. by par / por	Secretary of State, State of California		
8. N° sous n° bajo el número	*36097*		
9. Seal / stamp: Sceau / timbre: Sello / timbre:		**10. Signature:** Signature: Firma:	

This Apostille is the trilingual model Apostille Certificate as suggested by the Permanent Bureau and developed in response to the 2009 Special Commission on the practical operation of the Hague Apostille Convention.
This Apostille only certifies the authenticity of the signature and the capacity of the person who has signed the public document, and, where appropriate, the identity of the seal or stamp which the public document bears.
This Apostille does not certify the content of the document for which it was issued.
This Apostille is not valid for use anywhere within the United States of America, its territories or possessions.
To verify the issuance of this Apostille, see: www.sos.ca.gov/business/notary/apostille-search/.

Cette apostille est le modèle d'Apostille trilingue tel que suggéré par le Bureau Permanent et élaboré en réponse à la Commission spéciale de 2009 sur le fonctionnement pratique de la Convention de La Haye Apostille.
Cette Apostille atteste uniquement la véracité de la signature, la qualité en laquelle le signataire de l'acte a agi et, le cas échéant, l'identité du sceau ou timbre dont cet acte public est revêtu.
Cette Apostille ne certifie pas le contenu de l'acte pour lequel elle a été émise.
L'utilisation de cette Apostille n'est pas valable en / au États-Unis d'Amérique, ses territoires où possessions.
Cette Apostille peut être vérifiée à l'adresse suivante: www.sos.ca.gov/business/notary/apostille-search/.

Esta apostilla es el modelo trilingüe Certificado de Apostilla según lo sugerido por la Oficina Permanente y desarrollado en respuesta a la Comisión especial de 2009 sobre el funcionamiento práctico del Convenio de La Haya sobre Apostilla
Esta Apostilla certifica únicamente la autenticidad de la firma, la calidad en que el signatario del documento haya actuado y, en su caso, la identidad del sello o timbre del que el documento público esté revestido.
Esta Apostilla no certifica el contenido del documento para el cual se expidió.
No es válido el uso de esta Apostilla en Estados Unidos de América, sus territorios o posesiones.
Esta Apostilla se puede verificar en la dirección siguiente: www.sos.ca.gov/business/notary/apostille-search/.

I hereby certify that the attached document is a true and correct copy of the recorded Worldwide Assignment of rights for priority application USSN 61/245,273.

IN WITNESS WHEREOF I undersign as follows:

Date: March 5, 2012

Fabesha Greene

STATE OF CALIFORNIA
COUNTY OF SANTA CLARA

On March 5, 2012 before me, Aida M. Martin, Notary Public, personally appeared, Fabesha Greene, who proved to me on the basis of satisfactory evidence to be the person(s) whose name(s) is/are subscribed to the within instrument and acknowledged to me that he/she/they executed the same in his/her/their authorized capacity(ies), and that by his/her/their signature(s) on the instrument the person(s), or the entity upon behalf of which the person(s) acted, executed the instrument. I certify under PENALTY OF PERJURY under the laws of the State of California that the foregoing paragraph is true and correct.

WITNESS my hand and official seal.

Signature

(Seal)

AIDA M. MARTIN
Commission # 1871469
Notary Public - California
Santa Clara County
My Comm. Expires Dec 13, 2013

UNITED STATES PATENT AND TRADEMARK OFFICE

Facsimile Transmission

To:	**Name:**	DANISCO US INC.
	Company:	925 PAGE MILL ROAD
	Fax Number:	16508456504
	Voice Phone:	
From:	**Name:**	ASSIGNMENT SERVICES BRANCH
	Voice Phone:	571-272-3350

37 C.F.R. 1.6 sets forth the types of correspondence that can be communicated to the Patent and Trademark Office via facsimile transmissions. Applicants are advised to use the certificate of facsimile transmission procedures when submitting a reply to a non-final or final Office action by facsimile (37 CFR 1.8(a)).

Fax Notes:

Pg#	Description
1	Cover Page
2	140.TXT
4	Document 1, Batch 2174671

USPTO ASSIGNMENT SYSTEM PROCESSING

Date and time of transmission: Wednesday, September 15, 2010 4:37:34 PM
Number of pages including this cover sheet: 05

27

UNITED STATES PATENT AND TRADEMARK OFFICE

UNDER SECRETARY OF COMMERCE FOR INTELLECTUAL PROPERTY AND
DIRECTOR OF THE UNITED STATES PATENT AND TRADEMARK OFFICE

501289662A

SEPTEMBER 15, 2010

PTAS

DANISCO US INC.
925 PAGE MILL ROAD
WENFANG CHEN
PALO ALTO, CA 94304

UNITED STATES PATENT AND TRADEMARK OFFICE
NOTICE OF RECORDATION OF ASSIGNMENT DOCUMENT

THE ENCLOSED DOCUMENT HAS BEEN RECORDED BY THE ASSIGNMENT DIVISION OF THE U.S. PATENT AND TRADEMARK OFFICE. A COMPLETE MICROFILM COPY IS AVAILABLE AT THE ASSIGNMENT SEARCH ROOM ON THE REEL AND FRAME NUMBER REFERENCED BELOW.

PLEASE REVIEW ALL INFORMATION CONTAINED ON THIS NOTICE. THE INFORMATION CONTAINED ON THIS RECORDATION NOTICE REFLECTS THE DATA PRESENT IN THE PATENT AND TRADEMARK ASSIGNMENT SYSTEM. IF YOU SHOULD FIND ANY ERRORS OR HAVE QUESTIONS CONCERNING THIS NOTICE, YOU MAY CONTACT THE EMPLOYEE WHOSE NAME APPEARS ON THIS NOTICE AT 571-272-3350. PLEASE SEND REQUEST FOR CORRECTION TO: U.S. PATENT AND TRADEMARK OFFICE, MAIL STOP: ASSIGNMENT SERVICES BRANCH, P.O. BOX 1450, ALEXANDRIA, VA 22313.

RECORDATION DATE: 09/14/2010 REEL/FRAME: 024983/0380
NUMBER OF PAGES: 6

BRIEF: ASSIGNMENT OF ASSIGNOR'S INTEREST (SEE DOCUMENT FOR DETAILS).
DOCKET NUMBER: 31376USP2

ASSIGNOR:
BOWER, BENJAMIN S DOC DATE: 09/10/2010

ASSIGNOR:
HSI, MEGAN YEE DOC DATE: 09/03/2010

ASSIGNOR:
KAPER, THIJS DOC DATE: 09/03/2010

ASSIGNOR:
KELEMEN, BRADLEY R. DOC DATE: 09/03/2010

ASSIGNOR:
LANTZ, SUZANNE E. DOC DATE: 09/10/2010

ASSIGNOR:
LARENAS, EDMUND A DOC DATE: 09/08/2010

28

024983/0380 PAGE 2

ASSIGNOR:
MITCHINSON, COLIN DOC DATE: 09/10/2010

ASSIGNEE:
DANISCO US INC.
925 PAGE MILL ROAD
PALO ALTO, CALIFORNIA 94304

SERIAL NUMBER: 61245273 FILING DATE: 09/23/2009
PATENT NUMBER: ISSUE DATE:
TITLE: NOVEL GLYCOSYL HYDROLASE ENZYMES AND USES THEREOF

ASSIGNMENT SERVICES BRANCH
PUBLIC RECORDS DIVISION

PATENT ASSIGNMENT

Electronic Version v1.1
Stylesheet Version v1.1

09/14/2010
501289662

SUBMISSION TYPE:	NEW ASSIGNMENT
NATURE OF CONVEYANCE:	ASSIGNMENT

CONVEYING PARTY DATA

Name	Execution Date
Benjamin S Bower	09/10/2010
Megan Yee Hsi	09/03/2010
Thijs Kaper	09/03/2010
Bradley R. Kelemen	09/03/2010
Suzanne E. Lantz	09/10/2010
Edmund A Larenas	09/08/2010
Colin Mitchinson	09/10/2010

RECEIVING PARTY DATA

Name:	Danisco US Inc.
Street Address:	925 Page Mill Road
City:	Palo Alto
State/Country:	CALIFORNIA
Postal Code:	94304

PROPERTY NUMBERS Total: 1

Property Type	Number
Application Number:	61245273

CH $40.00 61245273

CORRESPONDENCE DATA

Fax Number: (650)845-6504
Correspondence will be sent via US Mail when the fax attempt is unsuccessful.
Phone: 650-846-7500
Email: aida.martin@danisco.com
Correspondent Name: Danisco US Inc.
Address Line 1: 925 Page Mill Road
Address Line 2: Wenfang Chen
Address Line 4: Palo Alto, CALIFORNIA 94304

ATTORNEY DOCKET NUMBER:	31376USP2

NAME OF SUBMITTER:	Wenfang Chen

Total Attachments: 4
source=31376USP2-Assignment#page1.tif
source=31376USP2-Assignment#page2.tif
source=31376USP2-Assignment#page3.tif
source=31376USP2-Assignment#page4.tif

Assignment

WHEREAS, **Benjamin S. Bower**, residing at 8479 Mayhews Landing Road, Newark, CA 94560, **Megan Yee Hsi**, residing at 80 Descanso Drive, Unit 2101, San Jose, CA 95134, **Thijs Kaper,** residing at 670 Myrtle Street, Half Moon Bay, Ca 94019, **Bradley R. Keleman**, residing at 658 12th Avenue, Menlo Park, CA 94040, **Suzanne E. Lantz**, residing at 2072 Eaton Avenue, San Carlos, CA 94070, **Edmund A. Larenas**, residing at 301 Nevada. Moss Beach, CA 94038, and **Colin Mitchinson**, residing at 381 Myrtle Street, Half Moon Bay, CA 94019, have invented a new and useful invention entitled **Novel Glycosyl Hydrolase Enzymes and Uses Thereof** and an application for United States Letters Patent was filed on September 23, 2009, United States Provisional Serial No. 61/245,273 and

WHEREAS, **Danisco US Inc.,** a corporation organized and existing under and by virtue of the laws of the State of Delaware, having its principal place of business at 925 Page Mill Road, Palo Alto, California 94304-1013, is desirous of acquiring an interest in and to said invention, and in the Letters Patents to be obtained therefor;

NOW, THEREFORE, be it known by all whom it may concern:

That for good and valuable consideration the receipt of which is hereby acknowledged, the said inventors have and do hereby sell, assign, transfer and set over unto the said Danisco US Inc., its successors and assigns, the full and exclusive right, title and interest, including all rights under the Paris Convention for the Protection of Industrial Property, in and to said invention, and in and to any and all Letters Patents to be granted and issued therefor, not only for, to and in the United States of America, its territories and possessions, but for, to and in all other countries, and it has been and is hereby authorized and requested that the appropriate government agencies issue said Letters Patents to said Danisco US Inc., . in accordance with this Assignment.

Said Inventor/Inventors hereby covenant/covenants and agree/agrees to cooperate with Danisco US Inc. to enable said Danisco US Inc. to enjoy to the fullest extent the right, title and interest herein conveyed in the United States and foreign countries. Such cooperation by said Inventors shall include prompt production of pertinent facts and documents, giving of testimony, execution of petitions, oaths, specifications, declarations or other papers and other assistance, all to the extent deemed necessary or desirable by said Danisco US Inc. for (a) perfecting the right, title and interest herein conveyed, (b) prosecuting any of said applications, (c) filing and prosecuting applications for reissuance of any said patents, (d) interference or other priority proceedings involving said invention, and (e) legal proceedings involving said invention and any

applications therefor and any patents granted thereon, including without limitation, opposition proceedings, cancellation proceedings, priority contests, public use proceedings, infringement actions and court actions, provided, however, that the expense incurred by said Inventors in providing such cooperation shall be paid for by said Danisco US Inc.

The terms and covenants of this Assignment shall inure to the benefit of said Danisco US Inc., its successors, assigns and other legal representatives, and shall be binding upon said Inventors, his/her respective heirs, legal representatives and assigns.

Said Inventor hereby warrants and represents that he/she has not entered and will not enter into any assignment, contract or understanding in conflict herewith.

IN WITNESS WHEREOF I undersign as follows:

Signed at Palo Alto, this 10 the day of Sept 2010.

By: _____
Benjamin S. Bower

WITNESSED BY:

Signature: _____

Printed Name of Witness: Aida Martin

IN WITNESS WHEREOF I undersign as follows:

Signed at Palo Alto, this 3 the day of Sept 2010.

By: _____
Megan Yee Hsi

WITNESSED BY:

Signature: _____

Printed Name of Witness: Aida Martin

IN WITNESS WHEREOF I undersign as follows:

Signed at Palo Alto, this 3 the day of Sept 2010.

By: _____
Thijs Kaper

WITNESSED BY:

Signature: _____

Printed Name of Witness: Aida Martin

IN WITNESS WHEREOF I undersign as follows:

Signed at Palo Alto, this 3 the day of September 2010.

By: _____
Bradley R. Keleman

WITNESSED BY:

Signature: _____

Printed Name of Witness: Aida Martin

IN WITNESS WHEREOF I undersign as follows:

Signed at Palo Alto, this 10 the day of Sept. 2010.

By: _____
Suzanne E. Lantz

WITNESSED BY:

Signature: _____

Printed Name of Witness: Aida Martin

IN WITNESS WHEREOF I undersign as follows:

Signed at Palo Alto, this 8 the day of September 2010.

By: _____
Edmund A. Larenas

WITNESSED BY:

Signature: _____

Printed Name of Witness: Aida Martin

IN WITNESS WHEREOF I undersign as follows:

Signed at Palo Alto, this 10 the day of Sept. 2010.

By: _____
Colin Mitchinson

WITNESSED BY:

Signature: _____

Printed Name of Witness: Aida Martin

NOVEL GLYCOSYL HYDROLASE ENZYMES AND USES THEREOF

1. CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority to United States Provisional Applications 61/245,273, filed September 23, 2009, and 61/289,886, filed December 23, 2009, the disclosures of which are incorporated herein by reference.

2. TECHNICAL FIELD

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

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 the1970s, 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.

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 cellooligosaccharides (*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.

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.

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.

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.

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

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.

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.

In one aspect, the invention of 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).

The enzyme blends/compositions of the invention 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.

The enzyme blends/compositions of the invention can comprise a number of different enzymes, including, *e.g.,* cellulases and/or hemicellulases. For example, the enzymes blends/compositions of the invention can be used to hydrolyze biomass or a suitable feedstock. The enzyme blends/compositions of the invention 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 of the invention 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 hemicelluose 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 of the invention 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 of the invention comprises a mixture of a xylanase, a xylosidase, a

cellobiohydrolase, an arabinofuranosidase, and a β-glucosidase. The enzyme blend/composition of the invention is suitably one that is non-naturally occurring.

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-occuring 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 hemicelluloytic 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 hemicelluloytic 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 endogeneous 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.

The enzyme blend/composition of the invention 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.

The enzyme blend/composition of the invention described herein is, in another example, a cellulase compostion. 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.*

For example, an enzyme blend/composition of the invention 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 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.*

For example, an enzyme blend/composition of the invention 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 celluase enzyme(s) of (e) is desirably able to 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.*

In further examples, an enzyme blend/composition of the invention 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

41

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.*

An enzyme blend/composition of the invention 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

AC

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.*

Moreover, an enzyme blend/composition of the invention 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.*

An enzyme blend/composition of the invention 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, wheras 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.*

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.

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.

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 hemicellolytic 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.

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.

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 expressable 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.*

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.

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

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.

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.

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.

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).

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.

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.

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.

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

98

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.

All publically available information as of the filing date, including, *e.g.,* publications, patents, patent applications, GenBank sequences, and ATCC deposits cited herein are hereby expressly incorporated by reference.

5. BRIEF DESCRIPTION OF THE FIGURES AND TABLES

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 "Fv51A" 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

49

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-μL 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

51

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).

SC

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 21A: *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/Pegl1-Fv43D plasmid.

Figure 46: TOPO Blunt/Pegl1-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).

[0100] Figure 62: Coding sequence for a construct comprising a CBH1 signal sequence (underlined) upstream of genomic DNA encoding mature Fo43A (SEQ ID NO:49).

[0101] 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).

[0102] 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 wtih flanking residues predicted to be involved in substrate binding.

[0103] 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

[0104] Enzymes have traditionally been classified by substrate specificity and reaction products. In the pre-genomic era, function was regarded as the most amenable (and

50

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.

[0105] 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.).

[0106] 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).

[0107] 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.

[0108] The enzymes of the disclosure belong, *inter alia*, to the glycosyl hydrolase families 3, 10, 11, 30, 39, 43, and/or 51.

[0109] 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.

[0110] 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.

[0111] 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 stearothermophylus* 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.

[0112] 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.

[0113] 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.

[0114] 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).

[0115] 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 $(\alpha/\beta)_8$ 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-b-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.

1.1.1. 6.1 Polypeptides of the Disclosure

[0116] 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.

[0117] 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.

[0118] 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.

[0119] 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.

[0120] 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 of the invention, for example, one that is non-naturally occurring.

[0121] **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 of the invention comprises a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the mature Fv3A sequence as shown in Figure 1B. The Fv3A polypeptide of the invention preferably has β-xylosidase activity.

[0122] Accordingly an Fv3A polypeptide of the invention 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.

[0123] **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*-nitophenyl-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 of the invention 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 of the invention comprises a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the mature Pf43A sequence as shown in Figure 2B. The Pf43A polypeptide of the invention preferably has β-xylosidase activity.

[0124] Accordingly a Pf43A polypeptide of the invention 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.

[0125] **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-nitophenyl-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 of the invention preferably has β-xylosidase activity.

[0126] Accoringly, an Fv43E polypeptide of the invention 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.

[0127] **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*-nitophenyl-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 or 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 of the invention preferably has β-xylosidase activity.

[0128] Accordingly, an Fv39A polypeptide of the invention 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.

[0129] **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-nitophenyl-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 of the invention preferably has β-xylosidase activity.

[0130] Accordingly an Fv43A polypeptide of the invention 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.

[0131] **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-nitophenyl-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 of the present invention preferably has β-xylosidase activity, L-α-arabinofuranosidase activity, or both β-xylosidase and L-α-arabinofuranosidase activities.

[0132] Accordingly, an Fv43B polypeptide of the invention 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.

[0133] **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*-nitophenyl-□α-L-arabinofuranoside. It was shown to catalyze the release of arabinose from branched arabino-xylo 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 of the invention preferably has β-xylosidase activity, L-α-arabinofuranosidase activity, or both β-xylosidase and L-α-arabinofuranosidase activities.

[0134] Accordingly, a Pa51A polypeptide of the invention 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.

[0135] **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*-nitophenyl-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 of the invention preferably has β-xylosidase activity.

[0136] Accordingly a Gz43A polypeptide of the invention 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.

[0137] **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*-nitophenyl-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 of the invention preferably has β-xylosidase activity.

[0138] Accordingly an Fo43A polypeptide of the invention 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.

[0139] **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*-nitophenyl-□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 of the invention preferably has L-α-arabinofuranosidase activity.

[0140] Accordingly an Af43A polypeptide of the invention 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.

[0141] **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 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 of the invention preferably has L-α-arabinofuranosidase activity.

[0142] Accordingly a Pf51A polypeptide of the invention 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.

[0143] **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 of the invention preferably has xylanase activity.

[0144] 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%,

92

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 of the invention preferably has xylanase activity.

[0145] **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*-nitophenyl-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 of the invention preferably has β-xylosidase activity.

[0146] Accordingly an Fv43D polypeptide of the invention 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.

[0147] **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 of the invention preferably has β-xylosidase activity.

[0148] Accordingly a Pf43B polypeptide of the invention 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.

[0149] **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 of the invention preferably has L-α-arabinofuranosidase activity.

[0150] Accordingly an Fv51A polypeptide of the invention 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.

[0151] **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 oberservation 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 Crystalogr. 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 polypetpide of the invention preferably has xylanase activity.

[0152] **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 175contiguous 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 of the invention preferably has xylanase activity.

[0153] **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*-nitophenyl-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

176

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 of the invention preferably has β-xylosidase activity.

[0154] Accordingly a *Trichoderma reesei* Bxl1 polypeptide of the invention 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.

[0155] **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 of the invention preferably has β-glucosidase activity.

[0156] Accordingly, the present disclosure provides a number of isolated, synthetic, or recombinant hemicelluloytic polypeptides or variants as described below:

[0157] (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

[0158] (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

[0159] (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

[0160] (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

[0161] (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

[0162] (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

[0163] (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

78

[0164] (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

[0165] (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

[0166] (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

[0167] (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

[0168] (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

[0169] (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.

[0170] 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

179

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.

[0171] Additionally the instant disclosure provides host cells that are recombiantly 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).

1.1.1.1.1. **6.1.1 Fusion proteins**

[0172] 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.

[0173] 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).

[0174] 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

80

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.

[0175] 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).

[0176] 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.

[0177] 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.

1.1.2. 6.2 Nucleic Acids and Host Cells

[0178] The present disclosure provides nucleic acids encoding a polypeptide of the disclosure, for example one described in Section 6.1 above.

[0179] 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).

[0180] 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.

[0181] 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.

[0182] For example, the disclosure provides an isolated nucleic acid molecule, wherein the nucleic acid molecule encodes:

[0183] (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

[0184] (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

[0185] (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

82

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

[0186] (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

[0187] (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

[0188] (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

[0189] (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

[0190] (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

[0191] (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

[0192] (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

[0193] (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

83

[0194] (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

[0195] (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.

[0196] The instant disclosure also provides:

[0197] (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

[0198] (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

[0199] (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

[0200] (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

[0201] (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

[0202] (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

[0203] (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

[0204] (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

[0205] (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

[0206] (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

[0207] (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

[0208] (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

[0209] (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.

[0210] The disclosure also provides expression cassettes and/or vectors comprising the above-described nucleic acids.

[0211] 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,

85

endoglucanase, or β-glucosidase promoter. A particulary suitable promoter can be, for example, a *T. reesei* cellobiohydrolase, endoglucanase, or β-glucosidase promoter. For example, the promoter is a cellobiohydrolase I (*cbh*1) promoter. Non-limiting examples of promoters include a *cbh1, cbh2, egl1, egl2, egl3, egl4, egl5, pki1, gpd1, xyn1*, or *xyn2* promoter. Additional non-limiting examples of promoters include a *T. reesei cbh1, cbh2, egl1, egl2, egl3, egl4, egl5, pki1, gpd1, xyn1*, or *xyn2* promoter.

[0212] 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.

[0213] 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*.

[0214] 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*.

[0215] 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, Chrysoporium, Coprinus, Coriolus, Corynascus, Chaetomium, Cryptococcus, Filobasidium, Fusarium, Gibberella, Humicola, Hypocrea, Magnaporthe, Mucor, Myceliophthora, Mucor, Neocallimastix, Neurospora, Paecilomyces, Penicillium, Phanerochaete, Phlebia, Piromyces, Pleurotus,Scytaldium, Schizophyllum, Sporotrichum, Talaromyces, Thermoascus, Thielavia, Tolypocladium, Trametes,* and *Trichoderma*. Suitable cells can also include cells of various anamorph and teleomorph forms fo these filamentous fungal genera.

[0216] 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 radiate, Pleurotus eryngii, Talaromyces flavus, Thielavia terrestris, Trametes villosa, Trametes versicolor, Trichoderma harzianum, Trichoderma koningii, Trichoderma longibrachiatum, Trichoderma reesei, and Trichoderma viride.

[0217] 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.

[0218] 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 recombiantly express one or more of *Trichoderma reesei* Xyn2, *Trichoderma reesei* Xyn3, *Trichoderma reesei* Bxl1 and/or *Trichoderma reesei* Bgl1.

81

[0219] 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.

[0220] 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.

[0221] 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.

1.1.3. 6.3 Enzyme Blends for Saccharification

[0222] 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.

1.1.3.1.1. 6.3.1 Background

[0223] 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

88

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.

[0224] 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.

[0225] 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.

[0226] 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.

[0227] 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 of the present invention 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.

[0228] 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) a composition made by combining component enzymes, whether in the form of a fermentation broth or partially or completely isolated or purified;

(2) a composition produced by an organism modified to express one or more component enzymes; in certain embodiments, the organism used to express one or more component enzymes can be modified to delete one or more genes; in certain other embodiments, the organism used to express one or more component enzymes can further comprise proteins affecting xylan hydrolysis, hemicellulose hydrolysis, and/or cellulose hydrolysis;

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

(4) an enzyme mixture produced *in situ*, *e.g.,* during a saccharification or fermentation reaction;

(5) a composition produced in accordance with any or all of the above (1)-(4).

[0229] The term "fermentation broth" as used herein refers to an enzyme preparation produced by fermentation that undergoes no or minimal recovery and/or purification subsequent to fermentation. For example, microbial cultures are grown to saturation, incubated under carbon-limiting conditions to allow protein synthesis (*e.g.,* expression of enzymes). Then, once the enzyme(s) are secreted into the cell culture media, the fermentation broths can be used. The fermentation broths of the disclosure can contain unfractionated or fractionated contents of the fermentation materials derived at the end of the fermentation. For example, the fermentation broths of the invention are unfractionated and comprise the spent culture medium and cell debris present after the microbial cells (*e.g.,* filamentous fungal cells) undergo a fermentation process. The fermentation broth can suitably contain the spent cell culture media, extracellular enzymes, and live or killed microbial cells. Alternatively, the fermentation broths can be fractionated to remove the microbial cells. In those cases, the fermentation broths can, for example, comprise the spent cell culture media and the extracellular enzymes.

[0230] 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.

1.1.3.1.2. **6.3.2 <u>Biomass</u>**

[0231] 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

90

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.

[0232] 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.

[0233] 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.

1.1.3.1.3. **6.3.3. Pretreatment**

[0234] 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 susceptable to enzymes and thus more amenable to hydrolysis by the enzyme(s) and/or enzyme blends/compositions of the disclosure.

[0235] 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.

[0236] 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.

[0237] 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.

[0238] 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.

[0239] Further pretreatment methods can involve the use of hydrogen peroxide H_2O_2. See Gould, 1984, Biotech, and Bioengr. 26:46-52.

[0240] 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.

[0241] 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.

[0242] 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.

1.1.3.1.4. **6.3.4 Exemplary Enzyme Blends**

[0243] 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.

[0244] 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.

[0245] 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.

[0246] **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.

[0247] **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.

[0248] **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

99

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.

[0249] **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.

[0250] **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

95

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.

[0251] The present disclosure provides a first enzyme blend/composition for lignocellulose saccharification comprising:

(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) 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) 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) 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) 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).

[0252] The present disclosure provides a second enzyme blend/composition for lignocellulose saccharification comprising:

(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.%

96

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) 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) 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) 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) 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).

[0253] The present disclosure provides a third enzyme blend/composition for lignocellulose saccharification comprising:

(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) 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.

[0254] The present disclosure provides a fourth enzyme blend/composition for lignocellulose saccharification comprising:

(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) 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) 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.

[0255] The present disclosure provides a fifth enzyme blend/composition for lignocellulose saccharification comprising:

(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) 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.

[0256] The present disclosure provides a sixth enzyme blend/composition for lignocellulose saccharification comprising:

(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) 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) 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) 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.

[0257] 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.

[0258] 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.

[0259] 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.

[0260] 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) 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.

1.1.3.1.5. **6.3.5. <u>Cellulases</u>**

[0261] 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.

[0262] 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*

100

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.*

[0263] 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.

(a) **6.3.5.1** **<u>β-Glucosidase</u>**

[0264] 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.

[0265] 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.

[0266] 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)).

[0267] 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.

[0268] 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.).

[0269] Moreover, the β-glucosidase can be a component of a whole cellulase, as described in Section 6.3.5.4 below.

[0270] 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 μmoL of Nitrophenol liberated from 4-nitrophenyl-β-D-glucopyranoside in 10 min at 50°C (122°F) and pH 4.8.

(b) **6.3.5.2** **Endoglucanases**

[0271] 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 endoglucanse can be produced by expressing an endogenous or exogenous endoglucanase gene. The endoglucanase can be, in some circumstances, overexpressed or underexpressed.

[0272] 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.

[0273] 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 *Acidothermus cellulolyticus* El 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 carotovara* (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.

(c) **6.3.5.3** **Cellobiohydrolases**

[0274] 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

102

expressing an endogeneous or exogeneous cellobiohydrolase gene. The cellobiohydrolase can be, in some circumstances, overexpressed or under expressed.

[0275] 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.

[0276] 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 CBHI.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).

(d) **6.3.5.4** **Whole Cellulases**

[0277] 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 cellooligosaccharides (*e.g.,* cellobiose).

[0278] 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.

[0279] 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.

[0280] 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.

[0281] 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).

[0282] 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.

[0283] 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

[0284] 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.

[0285] 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).

[0286] 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.

1.1.3.1.6. **6.3.6 Xylanases**

[0287] 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.

[0288] 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).

[0289] 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.

1.1.3.1.7. **6.3.7 β-Xylosidases**

[0290] The enzyme blends/compositions of the disclosure, for example, can suitablycomprise 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 [0115] 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.

[0291] 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).

[0292] 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.

1.1.3.1.8. **6.3.8 L-α-Arabinofuranosidases**

[0293] 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.

[0294] 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,

107

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 thermoviolacus, 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).

[0295] 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.

1.1.3.1.9. **6.3.9 Accessory proteins**

[0296] 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.

108

1.1.4. 6.4 Further Applications

[0297] 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.

1.1.4.1.1. 6.4.1 Wood, Paper and Pulp Treatments

[0298] 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.

[0299] 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.

[0300] 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.

[0301] 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.

1.1.4.1.2. 6.4.2 Treating Fibers and Textiles

[0302] 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.

[0303] 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 cellulosics 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).

1.1.4.1.3. 6.4.3 Treating Foods and Food Processing

[0304] 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.

[0305] 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.

[0306] 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.

[0307] 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.

1.1.4.1.4. **6.4.4 Animal Feeds and Food or Feed or Food Additives**

[0308] 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.

[0309] 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.

[0310] 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.

[0311] 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.

[0312] 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.

[0313] 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.

1.1.4.1.5. **6.4.5 Waste Treatment**

[0314] 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.

112

1.1.4.1.6. **6.4.6 Detergent, Disinfectant and Cleaning Compositions**

[0315] 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.

[0316] 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.

[0317] 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.

[0318] 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 anti-redeposition agents, and corrosion inhibitors.

[0319] 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.

[0320] 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 hemicelluose and optionally cellose with an enzyme blend/composition as described herein. The enzyme blend/compositon used in such a processs of the invention 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 of the invention 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/compositon used in such a process of the invention 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 hemicellolose in the biomass material. The enzyme blend/composition used in such a process of the invention 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 hemicellolose 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.

[0321] A suitable process of the invention 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 of the invention 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.

The process of the invention optionally further comprises recovering monosaccharides.

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

7. EXAMPLE 1: REPRESENTATIVE EXPERIMENTAL METHODS

1.1.5. 7.1 Materials and Methods

[0322] The following assays/methods were used in Example 1 and subsequent examples. Any deviations from the protocols provided below are indicated.

1.1.5.1.1. **7.1.1 Preparation of hemicellulose from plant tissues**

[0323] 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 lyophylization. Yield of recovered carbohydrate ranged from 6 to 23% of the starting plant material depending on the tissue and the preparation.

1.1.5.1.2. **7.1.2 Dilute ammonia pretreatment of biomass substrates**

[0324] Corncob and switchgrass were pretreated prior to enzymatic hydrolysis according to the methods and processing ranges described in WO06110901A (unless otherwise noted).

1.1.5.1.3. **7.1.3 Compositional analysis of biomass**

[0325] 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.

1.1.5.1.4. **7.1.4 Preparation of crude oligomers from ammonia pretreated corncob**

[0326] 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

115

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.

1.1.5.1.5. **7.1.5 Total protein assays**

[0327] 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.

[0328] 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)

[0329] 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.

[0330] 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

[0331] 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.

[0332] The total protein of purified samples was determined by A280 (*see, e.g.,* Pace et al., Protein Science, 1995, 4:2411).

[0333] 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).

[0334] 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.

[0335] In other cases, Coomassie Plus- the Better Bradford Assay (Thermo Scientific, Rockford, IL product #23238) was used according to manufacturer recommendation.

1.1.5.1.6. **7.1.6 Synthetic substrate (para-nitrophenyl substrate) activity assays**

[0336] 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 µL) was dispensed into Costar 96 well plates (Cat no. 9017). 20 µL 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 µL 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.

1.1.5.1.7. **7.1.7 Cob saccharification assay**

[0337] Typically, Corncob saccharification in a microtiter plate format was performed in accordance with the following procedures. The biomass substrate, dilute ammonia

[0332] The total protein of purified samples was determined by A280 (*see, e.g.,* Pace et al., Protein Science, 1995, 4:2411).

[0333] 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).

[0334] 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.

[0335] In other cases, Coomassie Plus- the Better Bradford Assay (Thermo Scientific, Rockford, IL product #23238) was used according to manufacturer recommendation.

1.1.5.1.6. **7.1.6 Synthetic substrate (para-nitrophenyl substrate) activity assays**

[0336] 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 µL) was dispensed into Costar 96 well plates (Cat no. 9017). 20 µL 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 µL 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.

1.1.5.1.7. **7.1.7 Cob saccharification assay**

[0337] Typically, Corncob saccharification in a microtiter plate format was performed in accordance with the following procedures. The biomass substrate, dilute ammonia

pretreated corncob, 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 corncob 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 corncob 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, pH10.0 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.

[0338] This describes a typical method that was used in multiple Examples herein. In certain Examples, corncob saccharification was measured using a modified protocol. The modifications are described with the individual examples.

1.1.5.1.8. **7.1.8 Sugar analysis by HPLC**

[0339] 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 H_2SO_4 run at 0.6 mL/min. Column temperature was 50^0C 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.

[0340] 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.

[0341] 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 H_2SO_4 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.

1.1.5.1.9. **7.1.9 Protein Analysis by HPLC**

[0342] 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 gycosylation 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.

1.1.5.1.10. **7.1.10 Cellulase Activity Assay Using Calcofluor White**

[0343] 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.

[0344] 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 sample - FI buffer)/(FI zero enzyme - FI buffer),
wherein FP is fraction product, and FI = fluorescence units.

1.1.5.1.11. **7.1.11 Cultivation of *Fusarium verticillioides* and purification of major hemicellulase activities detected in the extracellular protein**

[0345] 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.

[0346] 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.

[0347] 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^{0}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 de-starched corn pericarp in 1 L base media (15 g KH_2PO_4, 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.

[0348] 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).

[0349] 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 GH Health Sciences Akta system using Unicorn software. Proteins were eluted with a 0-0.5 M NaCl gradient in MES buffer.

[0350] 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. 100 µL of 1 N sodium carbonate stop solution was added and absorbance at 405 nm was read.

[0351] 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.

1.1.5.1.12. **7.1.12 Protein Identification by MS-MS**

[0352] 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.

[0353] For the identification of GH54 LARF, 200 µL of 1 N HCl (EM Industries, Hawthorne, NY) was added to 250 µL of "Fraction B3" (0.35 mg protein) from the purification procedure described above and the sample iced for 10 min. Next, 200 µL 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 µL of 8 M urea (MP Biomedicals, Solon, OH) was added to the pellet, then 4 µL of 0.2 M DTT (Sigma-Aldrich, St. Louis, MO). Sample was incubated at 52°C for 30 min and then 4 µL 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 µL of 0.1% n-octyl B-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 µg Trypsin (Promega Corp., Madison, WI) was added, and to the other half, 4 µg 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 u, 300A, 0.2 x 150 mm, Michrom Bioresources, Auburn, CA) was run at a flow rate of 200 µL/min. The injection volumes were 50 µL, 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 H_2O (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 μg, 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.

1.1.5.1.13. **7.1.13 Induction of *Fusarium verticillioides* Enzymes by Dilute Aqueous Ammonia Pretreated Switchgrass**

[0354] 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.

[0355] 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 Christakapoulos media (0.1% KH_2PO_4, 0.03% $CaCl_2$, 0.03% $MgSO_4$ x 7 H_2O, 2.61% Na_2HPO_4 x 7 H_20, 0.134% NaH_2PO_4 x 1 H_2O, 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.

[0356] 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.

1.1.5.1.14. **7.1.14 Size exclusion chromatography fractionation of the *Fusarium verticillioides* culture broth**

[0357] 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).

1.1.5.1.15. **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**

[0358] 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.

1.1.5.1.16. **7.1.16 Purification of Xylanases**

[0359] Xylanase purification was conducted in two stages:

1.1.5.1.18. **7.1.18 Purification of Fv51A**

[0366] 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 Stemdim 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*

1.1.6. 8.1 *Fusarium verticillioides* genes

[0367] 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.

[0368] 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.

[0369] 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)

RCS

[0360] Stage 1: Hydrophobic interaction chromatography (HIC). $(NH_4)_2SO_4$ 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 $(NH4)_2SO_4$. Stepwise Elution: 45% B; 0% B; water+10% glycerol.

[0361] 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.

[0362] 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).

[0363] 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.

[0364] The purity of the eluted protein was assessed using 10% Bis-Tris NUPAGE® SDS-PAGE (Novex), and determined to be greater than 95% pure.

1.1.5.1.17. **7.1.17 Purification of Fv43D**

[0365] 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.

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'-CACCATGTGGAAACTCCTCGTCAGC-3') (SEQ ID NO:76)

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

[0370] 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.

[0371] 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.

[0372] 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 (Δ*cbh1*, Δ*cbh2*, Δ*eg1*, Δ*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 µL/well of Glycine Minimal media (6.0 g/L glycine; 4.7 g/L $(NH_4)_2SO_4$; 5.0 g/L KH_2PO_4; 1.0 g/L $MgSO_4 \cdot 7H_2O$; 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 $FeSO_4 \cdot 7H_2O$; 16 g/L $ZnSO_4 \cdot 7H_2O$; 3.2 g/L $CuSO_4 \cdot 5H_2O$; 1.4 g/L $MnSO_4 \cdot H_2O$; 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.

1.1.7. 8.2 Genes from other species

[0373] 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.

[0374] 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)

[0375] 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'-CCGCGGCCGCACCATGGTTTCTTTCTCCTACCTGCTGCTG-3') (SEQ ID NO:82)
A.fumi-Q4WG11-R:
(5'-CCGGCGCGCCCTTACTAGTAGACAGTGATGGAAGCAGATCCG-3') (SEQ ID NO:83)
AfuXyn5:
A.fumi-Q4WFZ8-F: (5'-CCGCGGCCGCACCATGATCTCCATTTCCTCGCTCAGCT-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)

[0376] 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

[0377] 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

1.1.8. 10.1 Saccharification performance of expressed proteins

[0378] 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)

[0379] 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, Cg51B, 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.

[0380] 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.

[0381] 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 corncoob 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)

[0382] 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.

[0383] 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)

[0384] 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.

[0385] 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)

[0386] 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.

[0387] 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.

[0388] 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)

[0389] 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.

1.1.9. 10.2 Activity of candidate endo-xylanases with birchwood xylan

[0390] 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 µL) to the microtiter plate and the plates incubated at 50°C for 10 min. Meanwhile, 100 µL of DNS solution was added to PCR tubes. Following the 10-min incubation, 60 µL 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 µmole of xylose reducing sugar equivalents per minute under the conditions of the assay. Results are shown in Table 3.

1.1.10. 10.3 Enzyme hydrolysis of arabinoxylan oligomers from saccharified corncob

[0391] 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 momomer 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.

[0392] 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.

[0393] 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 B-XYLOSIDASES FOR EFFECTIVE CORNCOB HYDROLYSIS

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.

The Bio-gel P2 fractions containing oligomers of greater than DP2 (5-10 mg) were lyophilized then dissolved in 0.7 mL of a D_2O 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 ^{13}C 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.

[0394] The arabinose containing oligomers as determined by NMR, are dominated by one or more branched structures in which arabinose is linked β- 1→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→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→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.

[0395] 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→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.

[0396] 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

[0397] 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

[0398] Pretreated corncob as described in Example 1 was used as water slurry adjusted to about pH 5 with H_2SO_4 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.

[0399] 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.

[0400] 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.

[0401] 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.

[0402] 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.

[0403] 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).

[0404] 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

[0405] 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

[0406] 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.

[0407] 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 µL 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 µL of water, centrifuged and filtered before HPLC analysis for monomer sugars as described in Example 1. 200 µL of each hemicellulose suspension was diluted with 200 µL of 0.8 N H_2SO_4, 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.

[0408] 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→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*

[0409] 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.

[0410] The construction of the expression cassettes for these different genes and the transformation of *T. reesei* strain are described below.

11.1.1. 15.1 Construction of the β-glucosidase expression cassette

[0411] 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'–CTGTCGCGGGACCACTGGGCAAGACTCCGCAGGGCGGTCG-3') (SEQ ID NO:90)
Reverse Primer SK942:
(5'– CCTACGCTACCGACAGAGTG-3') (SEQ ID NO:91)

[0412] 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)

11.1.2. 15.2 <u>Construction of the endoxylanase expression cassette</u>

[0413] 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)

[0414] 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.

[0415] Forward Primer SK745: (5' – GAGTTGTGAAGTCGGTAATCC-3') (SEQ ID NO:96)

[0416] Reverse Primer SK822: (5' – CACGAAGAGCGGCGATTC-3') (SEQ ID NO:97)

11.1.3. 15.3 Construction of the β-xylosidase *Fv3A* expression cassette

[0417] The *F. verticilloides* β-xylosidase *fv3A* gene was amplified from a *F. verticilloides* genomic DNA sample using the primers MH124 and MH125.

[0418] 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)

[0419] 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)

11.1.4. 15.4 Construction of the β-xylosidase Fv43D expression cassette

[0420] 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'–CATGCGATCGCGACGTTTTGGTCAGGTCG-3') (SEQ ID NO:105)
Reverse Primer SK1321: (5'-GACAGAAACTTGAGCTGCATGGTGTGGGACAACAAGAAGG-3') (SEQ ID NO:106)

[0421] 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.

11.1.5. 15.5 Construction of the α-arabinofuranosidase expression cassette

[0422] 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 *egl1* 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)

[0423] 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)

11.1.6. 15.6 Co-Transformation of *T. reesei* with the β-glucosidase and endoxylanase expression cassettes

[0424] A *Trichoderma reesei* mutant strain, derived from RL-P37 (Sheir-Neiss, G *et al.* Appl. Microbiol. Biotechnol. 1984, 20:46-53) and selected for high cellulase production was co-transformed with the β-glucosidase expression cassette (*cbh1* promoter, *T.reesei* β-glucosidase1 gene, *cbh1* terminator, and *amdS* marker), and the endoxylanase expression cassette (*cbh1* promoter, *T.reesei xyn3*, and *cbh1* terminator) using PEG-mediated

41

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.

11.1.7. 15.7 Co-transformation of *T. reesei* strain #229 with two β-xylosidase and α-arabinofuranosidase expression cassettes

[0425] *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_2O bring to post-sterile addition:	980 mL
50% Glucose	20 mL

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

In 750 mL deionized H2O, dissolve successively:

$Na_3Citrate*2H_2O$	125.00 g
KH_2PO_4 (Anhydrous)	250.00 g
NH_4NO_3 (Anhydrous)	100 g
$MgSO_4*7H_2O$	10.00 g
$CaCl_2*2H_2O$	5.00 g
Vogels Trace Element Solution	5.0 mL
Vogels Biotin Solution	2.5 mL
[0426] With deionized H_2O,	bring to 1 L

[0427] 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

[0428] The saccharification performance of enzyme compositions produced by *T. reesei* integrated expression strains on dilute ammonia pretreated corncob was evaluated. *T.*

142

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.

[0429] 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.

[0430] A chromatographic comparison of the enzyme composition produced by three different integrated strains is shown in Figure 47D.

11.1.8. 16.1 HPLC Analysis

[0431] 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.

[0432] 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

[0433] 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.

[0434] 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 H_2SO_4 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

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

[0435] 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.

[0436] 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.

[0437] 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.

[0438] 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

[0439] 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.

[0440] 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).

[0441] 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.

[0442] 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).

[0443] 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

[0444] 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.

[0445] 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.

[0446] 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

[0447] 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% NH_3 (weight / weight DM, added as NH_4OH). 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.

[0448] 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

[0449] 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%.

[0450] 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.

147

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

[0451] 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

[0452] 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

[0453] 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.

[0454] 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.

25. SPECIFIC EMBODIMENTS AND INCORPORATION BY REFERENCE

[0455] All publications, patents, patent applications and other documents cited in this application are hereby incorporated by reference in their entireties for all purposes to the same extent as if each individual publication, patent, patent application or other document were individually indicated to be incorporated by reference for all purposes.

[0456] While various specific embodiments have been illustrated and described, it will be appreciated that various changes can be made without departing from the spirit and scope of the disclosure(s).

WHAT IS CLAIMED IS:

1. A non-naturally occurring composition comprising:

   a. a first polypeptide or a variant thereof having β-xylosidase activity, L-α-arabinofuranosidase activity, or both β-xylosidase activity and L-α-arabinofuranosidase activity; and

   b. a second polypeptide or a variant thereof having xylanase activity.

2. The composition of claim 1, wherein the first polypeptide or variant thereof is an Fv43D, Fv3A, Pf43A, Fv43E, Fv39A, Fv43A, Fv43B, Pa51A, Fo43A, Gz43A, *Trichoderma reesei* Bxl1, Fv51A, Af43A, or Pf51A polypeptide.

3. The composition of claim 1 or 2, wherein the first polypeptide or variant thereof has β-xylosidase activity.

4. The composition of claim 3, wherein the first polypeptide or variant thereof having β-xylosidase activity is an Fv43D, Fv3A, Pf43A, Fv43E, Fv39A, Fv43A, Fv43B, Pa51A, Fo43A, Gz43A, or *Trichoderma reesei* Bxl1 polypeptide.

5. The composition of any one of claims 1-4, further comprising a third polypeptide or a variant thereof having L-α-arabinofuranosidase activity.

6. The composition of claim 5, wherein the third polypeptide or variant thereof is an Fv51A, Af43A, Fv43B, Pa51A or Pf51A polypeptide.

7. The composition of any one of claims 1-4, further comprising a third polypeptide having β-xylosidase activity.

8. The composition of claim 7, wherein the first polypeptide is an Fv3A polypeptide or an Fv43A polypeptide, and the third polypeptide is an Fv43D, Pa51A, Gz43A, *Trichoderma reesei* Bxl1, Pf43A, Fv43E, Fv39A, Fo43A, or Fv43B polypeptide.

9. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Fv43D polypeptide has at least 90% sequence identity to 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.

10. The composition of any one of claims 2, 4, 6 and 8, wherein, if present, the Fv3A polypeptide has at least 90% 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.

11. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Pf43A polypeptide has at least 90% 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.

12. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Fv43E polypeptide has at least 90% 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.

13. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Fv39A polypeptide has at least 90% 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 145-439 of SEQ ID NO:8.

14. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Fv43A polypeptide has at least 90% 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, or (vi) 321-449 of SEQ ID NO:10.

15. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Fv43B polypeptide has at least 90% 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.

16. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Pa51A polypeptide has at least 90% 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.

17. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Fo43A polypeptide has at least 90% 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.

18. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Gz43A polypeptide has at least 90% 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.

20. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the *Trichoderma reesei* Bxl1 polypeptide has at least 90% sequence identity to the amino acid sequence of SEQ ID NO:44.

21. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Fv51A polypeptide has at least 90% 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.

22. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Af43A polypeptide has at least 90% sequence identity to SEQ ID NO:20, or to residues (i) 15-558, or (ii) 15-295 of SEQ ID NO:20.

23. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Pf51A polypeptide has at least 90% 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.

24. The composition of any one of claims 2, 4, 6, or 8, wherein, if present, the Fv43D polypeptide is encoded by a nucleic acid having at least 90% sequence identity to SEQ ID NO:27, or by a nucleic acid that is capable of hybridizing, under high stringency conditions, to SEQ ID NO:27, or to a fragment thereof.

151

25. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Fv3A polypeptide is encoded by a nucleic acid having at least 90% sequence identity to SEQ ID NO:1, or by a nucleic acid that is capable of hybridizing, under high stringency conditions, to SEQ ID NO:1, or to a fragment thereof.

26. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Pf43A polypeptide is encoded by a nucleic acid having at least 90% sequence identity to SEQ ID NO:3, or by a nucleic acid that is capable of hybridizing, under high stringency conditions, to SEQ ID NO:3, or to a fragment thereof.

27. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Fv43E polypeptide is encoded by a nucleic acid having at least 90% sequence identity to SEQ ID NO:5, or by a nucleic acid that is capable of hybridizing, under high stringency conditions, to SEQ ID NO:5, or to a fragment thereof.

28. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Fv39A polypeptide is encoded by a nucleic acid having at least 90% sequence identity to SEQ ID NO:7, or by a nucleic acid that is capable of hybridizing, under high stringency conditions, to SEQ ID NO:7, or to a fragment thereof.

29. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Fv43A polypeptide is encoded by a nucleic acid having at least 90% sequence identity to SEQ ID NO:9, or by a nucleic acid that is capable of hybridizing, under high stringency conditions, to SEQ ID NO:9, or to a fragment thereof.

30. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Fv43B polypeptide is encoded by a nucleic acid having at least 90% sequence identity to SEQ ID NO:11, or by a nucleic acid that is capable of hybridizing, under high stringency conditions, to SEQ ID NO:11, or to a fragment thereof.

31. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Pa51A polypeptide is encoded by a nucleic acid having at least 90% sequence identity to SEQ ID NO:13, or by a nucleic acid that is capable of hybridizing, under high stringency conditions, to SEQ ID NO:13, or to a fragment thereof.

32. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Fo43A polypeptide is encoded by a nucleic acid having at least 90% sequence identity to SEQ ID NO:17, or by a nucleic acid that is capable of hybridizing, under high stringency conditions, to SEQ ID NO:17, or to a fragment thereof.

33. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Gz43A polypeptide is encoded by a nucleic acid having at least 90% sequence identity to SEQ ID NO:15, or by a nucleic acid that is capable of hybridizing, under high stringency conditions, to SEQ ID NO:15, or to a fragment thereof.

34. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Fv51A polypeptide is encoded by a nucleic acid having at least 90% sequence identity to SEQ ID

NO:31, or by a nucleic acid that is capable of hybridizing, under high stringency conditions, to SEQ ID NO:31, or to a fragment thereof.

35. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Af43A polypeptide is encoded by a nucleic acid having at least 90% sequence identity to SEQ ID NO:19, or by a nucleic acid that is capable of hybridizing, under high stringency conditions, to SEQ ID NO:19, or to a fragment thereof.

36. The composition of any one of claims 2, 4, 6, and 8, wherein, if present, the Pf51A polypeptide is encoded by a nucleic acid having at least 90% sequence identity to SEQ ID NO:21, or by a nucleic acid that is capable of hybridizing, under high stringency conditions, to SEQ ID NO:21, or to a fragment thereof.

37. The composition of any one of claims 1-36, wherein the second polypeptide is a *Trichoderma* xylanase or an *Aspergillus* xylanase.

38. The composition of claim 37, wherein the *Trichoderma* xylanase is a *Trichoderma reesei* Xyn3 polypeptide or a *Trichoderma reesei* Xyn2 polypeptide.

39. The composition of claim 38, wherein, if present, the *Trichoderma reesei* Xyn3 polypeptide has at least 90% sequence identity to SEQ ID NO:42, or to residues 17-347 of SEQ ID NO:42.

40. The composition of claim 38, wherein, if present, the *Trichoderma reesei* Xyn2 polypeptide has at least 90% sequence identity to SEQ ID NO:43, or to residues 33-222 of SEQ ID NO:43.

41. The composition of claim 38, wherein, if present, the *Trichoderma reesei* Xyn3 polypeptide is encoded by a nucleic acid having at least 90% sequence identity to SEQ ID NO:41, or by a nucleic acid that is capable of hybridizing under high stringency conditions to SEQ ID NO:41, or to a fragment thereof.

42. The composition of claim 37, wherein the *Aspergillus* xylanase is an AfuXyn2 polypeptide or an AfuXyn5 polypeptide.

43. The composition of claim 42, wherein, if present, the AfuXyn2 polypeptide has at least 90% sequence identity to SEQ ID NO:24, or to residues 19-228 of SEQ ID NO:24.

44. The composition of claim 42, wherein, if present, the AfuXyn5 polypeptide has at least 90% sequence identity to SEQ ID NO:26, or to residues 20-313 of SEQ ID NO:26.

45. The composition of claim 42, wherein, if present, the AfuXyn2 polypeptide is encoded by a nucleic acid having at least 90% sequence identity to SEQ ID NO:23, or by a nucleic acid that is capable of hybridizing under high stringency to SEQ ID NO:23, or to a fragment thereof.

46. The composition of claim 42, wherein, if present, the AfuXyn5 polypeptide is encoded by a nucleic acid having at least 90% sequence identity to SEQ ID NO:25, or by a nucleic acid that is capable of hybridizing under stringent conditions to a complement of SEQ ID NO:25, or to a fragment thereof.

47. The composition of any one of claims 1-46, further comprising one or more polypeptides having cellulase activity.

48. The composition of claim 47, wherein the one or more polypeptides having cellulase activity are independently components of a whole cellulase.

49. The composition of claim 48, wherein the whole cellulase is a β-glucosidase-enriched whole cellulase.

50. The composition of claim 48 or 49, wherein the whole cellulase is a filamentous fungal whole cellulase.

51. The composition of any one of claims 48-50, wherein the filamentous fungal whole cellulase is a preparation from an *Acremonium, Aspergillus, Emericella, Fusarium, Humicola, Mucor, Myceliophthora, Neurospora, Penicillium, Scytalidium, Thielavia, Chrysosporium, Phanerochaete, Tolypocladium,* or *Trichoderma* species.

52. The composition of any one of claims 48-50, wherein the whole cellulase is a *Chrysosporium lucknowense, Phanerochaete chrysosporium, Trichoderma reesei, Trichoderma harzianum, Trichoderma koningii, Trichoderma longibrachiatum, Trichoderma viride,* or *Pencillium funiculosum* whole cellulase.

53. The composition of claim 47, wherein the one or more polypeptides having cellulase activity are selected from a β-glucosidase, an endoglucanase, and a cellobiohydrolase.

54. The composition of claim 49 or 53, wherein the β-glucosidase is a *Trichoderma reesei* Bgl1 polypeptide.

55. The composition of claim 54, wherein the *Trichoderma reesei* Bgl1 polypeptide has at least 90% sequence identity to SEQ ID NO:45, or to residues 32-744 of SEQ ID NO:45.

56. The composition of any one of claims 49 and 53-55, wherein the β-glucosidase polypeptide constitutes up to 50 wt.% of the total weight of cellulase enzymes in the composition.

57. The composition of any one of claims 49 and 53-56, wherein the β-glucosidase polypeptide constitutes 2 wt.% to 50 wt.% of the total weight of proteins in the composition.

58. The composition of any one of claims 47-57, wherein the one or more polypeptides having cellulase activity constitute 30 wt.% to 80 wt.% of the total weight of proteins in the composition.

59. The composition of claim 58, wherein the one or more polypeptides having cellulase activity together are capable of achieving at least 0.00005 fraction product per mg protein per gram of phosphoric acid swollen cellulose (PASC) as determined by a calcofluor assay.

60. The composition of any one of claims 1-59, further comprising one or more accessory proteins.

61. The composition of claim 60, wherein the one or more accessory proteins are selected from the group consisting of mannanases, galactanases, arabinases, ligninases, amyloases, glucuronidases, proteases, esterases, lipases, xyloglucanases, glycoside hydrolase Family 61 polypeptides, CIP1, CIP2, swollenin, and expansins.

62. The composition of claim 60, wherein the one or more accessory proteins are selected from the group consisting of CIP1-like proteins, CIP2-like proteins, cellobiose dehydrogenases, and manganese peroxidases.

63. The composition of any one of claims 60-62, wherein the one or more accessory proteins constitute 1 wt.% to 50 wt.% of the total weight of proteins in the composition.

64. The composition of any one of claims 1-63, capable of converting 60% or more of the hemicellulose xylan from a biomass into xylose.

65. The composition of claim 64, capable of converting 70% or more of the hemicellulose xylan from a biomass into xylose.

66. The composition of claim 65, capable of converting 80% or more of the hemicellulose xylan from a biomass into xylose.

67. The composition of any one of claims 64-66, wherein the biomass is corncob, corn stover, corn fiber, switchgrass, sorghum, paper, pulp, or sugarcane bagasse.

68. The composition of any one of claims 64-66, wherein the biomass is *Miscanthus.*

69. The composition of any one of claims 1-68, further comprising a biomass.

70. The composition of claim 69, wherein the combined weight of polypeptides having xylanase activity is 1 g to 40 g per 1 kg of hemicellulose in the biomass.

71. The composition of claim 69, wherein the combined weight of polypeptides having xylanase activity is 0.5 g to 40 g per 1 kg of hemicellulose in the biomass.

72. The composition of claim 71, wherein the combined weight of polypetpides having xylanase activity is 0.5 g to 20 g per 1 kg of hemicellulose in the biomass.

73. The composition of any one of claims 70-72, wherein the combined weight of polypeptides having xylanase activity is 2 g to 20 g per 1 kg of hemicellulose in the biomass.

74. The composition of any one of claims 69-73, wherein the combined weight of polypeptides having β-xylosidase activity is 1 g to 50 g per 1 kg of hemicellulose in the biomass.

75. The composition of any one of claims 69-73, wherein the combined weight of polypeptides having β-xylosidase activity is 0.5 g to 50 g per 1 kg of hemicellulose in the biomass.

76. The composition of claim 75, wherein the combined weight of polypeptides having β-xylosidase activity is 0.5 g to 40 g per 1 kg of hemicellulose in the biomass.

155

77. The composition of any one of claims 74-76, wherein the combined weight of polypeptides having β-xylosidase activity is 2 g to 40 g per 1 kg of hemicellulose in the biomass.

78. The composition of any one of claims 69-77, wherein the combined weight of polypeptides having L-α-arabinofuranosidase activity, if present, is 0.5 g to 20 g per 1 kg of hemicellulose in the biomass.

79. The composition of any one of claims 69-77, wherein the combined weight of polypetpides having L-α-arabinofuranosidase activity, if present, is 0.2 g to 20 g per 1 kg of hemicellulose in the biomass.

80. The composition of any one of claims 69-79, wherein the combined weight of polypeptides having cellulase activity, if present, is 1 g to 100 g per 1 kg of cellulose in the biomass.

81. A fermentation broth comprising the composition of any one of claims 1-80.

82. The fermentation broth of claim 81, which is the fermentation broth of a filamentous fungus.

83. The fermentation broth of claim 82, wherein the filamentous fungus is a *Trichoderma, Humicola, Fusarium, Aspergillus, Neurospora, Penicillium, Cephalosporium, Achlya, Podospora, Endothia, Mucor, Cochliobolus, Pyricularia, Phanerochaete,* or *Chrysosporium.*

84. The fermentation broth of claim 82, wherein the filamentous fungus is a *Aspergillus awamori, Aspergillus fumigatus, Aspergillus foetidus, Aspergillus japonicus, Aspergillus nidulans, Aspergillus niger, Aspergillus oryzae, Chrysosporium lucknowense, Fusarium bactridioides, Fusarium cerealis, Fusarium crookwellense, Fusarium culmorum, Fusarium graminearum, Fusarium graminum, Fusarium heterosporum, Fusarium negundi, Fusarium oxysporum, Fusarium reticulatum, Fusarium roseum, Fusarium sambucinum, Fusarium sarcochroum, Fusarium sporotrichioides, Fusarium sulphureum, Fusarium torulosum, Fusarium trichothecioides, Fusarium venenatum, Bjerkandera adusta, Ceriporiopsis aneirina, Ceriporiopsis aneirina, Ceriporiopsis caregiea, Ceriporiopsis gilvescens, Ceriporiopsis pannocinta, Ceriporiopsis rivulosa, Ceriporiopsis subrufa, Ceriporiopsis subvermispora, Coprinus cinereus, Coriolus hirsutus, Humicola insolens, Humicola lanuginosa, Mucor miehei, Myceliophthora thermophila, Neurospora crassa, Neurospora intermedia, Penicillium purpurogenum, Penicillium canescens, Penicillium solitum, Penicillium funiculosum, Phanerochaete chrysosporium, Phlebia radiate, Pleurotus eryngii, Talaromyces flavus, Thielavia terrestris, Trametes villosa, Trametes versicolor, Trichoderma harzianum, Trichoderma koningii, Trichoderma longibrachiatum, Trichoderma reesei,* or *Trichoderma viride.*

85. The fermentation broth of claim 82, wherein the filamentous fungus is an *Aspergillus,* or a *Fusarium.*

86. The fermentation broth of claim 82, wherein the filamentous fungus is a *Trichoderma reesei.*

150

87. The fermentation broth of claim 82, wherein the filamentous fungus is a *Penicillium funiculosum.*

88. The fermentation broth of any one of claims 81-87, which is a cell-free fermentation broth.

89. A method of converting biomass to sugars, comprising contacting the biomass with the composition of any one of claims 1-80, or with the fermentation broth of any one of claims 81-88.

90. A saccharification process comprising treating a material comprising hemicellulose with the composition of any one of claims 1-80 or with the fermentation broth of any one of claims 81-89.

91. The process of claim 90, wherein the material comprising hemicellulose is corncob, corn stover, corn fiber, switchgrass, sorghum, paper, pulp, or sugarcane bagasse.

92. The process of claim 90, wherein the material comprising hemicellulose is *Miscanthus.*

93. The process of any one of claims 90-92, which yields at least 60% xylose from hemicellulose xylan of the material comprising hemicellulose.

94. The process of any one of claims 90-93, further comprising recovering monosaccharides.

95. An isolated, synthetic, or recombiant nucleic acid encoding a polypeptide having at least 90% sequence identity to the amino acid sequence of SEQ ID NO:28, or to the amino acid sequence corresponding to residues (i) 20-341, (ii) 21-350, (iii) 107-341, or (iv) 107-350 of SEQ ID NO:28, which has β-xylosidase activity.

96. An isolated, synthetic, or recombinant nucleic acid encoding a polypeptide having at least 90% sequence identity to the amino acid sequence of SEQ ID NO:4, or to the amino acid sequence corresponding 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, which has β-xylosidase activity.

97. An isolated, synthetic, or recombinant nucleic acid encoding a polypeptide having at least 90% sequence identity to the amino acid sequence of SEQ ID NO:6, or to the amino acid sequence corresponding to residues (i) 19-530, (ii) 29-530, (iii) 19-300, or (iv) 29-300 of SEQ ID NO:6, which has β-xylosidase activity.

98. An isolated, synthetic, or recombinant nucleic acid encoding a polypeptide having at least 90% sequence identity to the amino acid sequence of SEQ ID NO:8, or to the amino acid sequence corresponding to residues (i) 20-439, (ii) 20-291, (iii) 145-291, or (iv) 145-439 of SEQ ID NO:8, which has β-xylosidase activity.

99. An isolated, synthetic, or recombinant nucleic acid encoding a polypeptide having at least 90% sequence identity to the amino acid sequence of SEQ ID NO:10, or to the amino acid sequence corresponding to residues (i) 23-449, (ii) 23-302, (iii) 23-320, (iv) 23-448, (v)

157

303-448, (vi) 303-449, (vii) 321-448, or (viii) 321-449 of SEQ ID NO:10, which has β-xylosidase activity.

100. An isolated, synthetic, or recombinant nucleic acid encoding a polypeptide having at least 90% sequence identity to the amino acid sequence of SEQ ID NO:12, or to the amino acid sequence corresponding to residues (i) 17-574, (ii) 27-574, (iii) 17-303, or (iv) 27-303 of SEQ ID NO:12, which has β-xylosidase and/or L-α-arabinofuranosidase activities.

101. An isolated, synthetic, or recombinant nucleic acid encoding a polypeptide having at least 90% sequence identity to the amino acid sequence of SEQ ID NO:14, or to the amino acid sequence corresponding to residues (i) 21-676, (ii) 21-652, (iii) 469-652, or (iv) 469-676 of SEQ ID NO:14, which has β-xylosidase and/or L-α-arabinofuranosidase activities.

102. An isolated, synthetic, or recombinant nucleic acid encoding a polypeptide having at least 90% sequence identity to the amino acid sequence of SEQ ID NO:16, or to the amino acid sequence corresponding to residues (i) 19-340, (ii) 53-340, (iii) 19-383, or (iv) 53-383 of SEQ ID NO:16, which has β-xylosidase activity.

103. An isolated, synthetic, or recombinant nucleic acid encoding a polypeptide having at least 90%% sequence identity to the amino acid sequence of SEQ ID NO:18, or to the amino acid sequence corresponding to residues (i) 21-341, (ii) 107-341, (iii) 21-348, or (iv) 107-348 of SEQ ID NO:18, which has β-xylosidase activity.

104. An isolated, synthetic, or recombinant nucleic acid encoding a polypeptide having at least 90% sequence identity to the amino acid sequence of SEQ ID NO:20, or to the amino acid sequence corresponding to residues (i) 15-558, or (ii) 15-295 of SEQ ID NO:20, which has L-α-arabinofuranosidase activity.

105. An isolated, synthetic, or recombinant nucleic acid encoding a polypeptide having at least 90% sequence identity to the amino acid sequence of SEQ ID NO:22, or to the amino acid sequence corresponding to residues (i) 21-632, (ii) 461-632, (iii) 21-642, or (iv) 461-642 of SEQ ID NO:22, which has L-α-arabinofuranosidase activity.

106. An isolated, synthetic, or recombinant nucleic acid encoding a polypeptide having at least 90% sequence identity to the amino acid sequence of SEQ ID NO:2, or to the amino acid sequence corresponding to residues (i) 24-766, (ii) 73-321, (iii) 73-394, (iv) 395-622, (v) 24-622, or (iv) 73-622 of SEQ ID NO:2, which has β-xylosidase activity.

107. An isolated, synthetic, or recombinant nucleic acid encoding a polypeptide having at least 90% sequence identity to the amino acid sequence of SEQ ID NO:32, or to the amino acid sequence corresponding to residues (i) 21-660, (ii) 21-645, (iii) 450-645, or (iv) 450-660 of SEQ ID NO:32, which has L-α-arabinofuranosidase activity.

108. An isolated polypeptide, wherein said polypeptide has at least 90% sequence identity to the amino acid sequence of SEQ ID NO:28, or to the amino acid sequence corresponding to residues (i) 20-341, (ii) 21-350, (iii) 107-341, or (iv) 107-350 of SEQ ID NO:28, which has β-xylosidase activity.

109. An isolated polypeptide, wherein said polypeptide has at least 90% sequence identity to the amino acid sequence of SEQ ID NO:4, or to the amino acid sequence corresponding 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, which has β-xylosidase activity.

110. An isolated polypeptide, wherein said polypeptide has at least 90% sequence identity to the amino acid sequence of SEQ ID NO:6, or to the amino acid sequence corresponding to residues (i) 19-530, (ii) 29-530, (iii) 19-300, or (iv) 29-300 of SEQ ID NO:6, which has β-xylosidase activity.

111. An isolated polypeptide, wherein said polypeptide has at least 90% sequence identity to the amino acid sequence of SEQ ID NO:8, or to the amino acid sequence corresponding to residues (i) 20-439, (ii) 20-291, (iii) 145-291, or (iv) 145-439 of SEQ ID NO:8, which has β-xylosidase activity.

112. An isolated polypeptide, wherein said polypeptide has at least 90% sequence identity to the amino acid sequence of SEQ ID NO:10, or to the amino acid sequence corresponding 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, which has β-xylosidase activity.

113. An isolated polypeptide, wherein said polypeptide has at least 90% sequence identity to the amino acid sequence of SEQ ID NO:12, or to the amino acid sequence corresponding to residues (i) 17-574, (ii) 27-574, (iii) 17-303, or (iv) 27-303 of SEQ ID NO:12, which has β-xylosidase and/or L-α-arabinofuranosidase activities.

114. An isolated polypeptide, wherein said polypeptide has at least 90% sequence identity to the amino acid sequence of SEQ ID NO:14, or to the amino acid sequence corresponding to residues (i) 21-676, (ii) 21-652, (iii) 469-652, or (iv) 469-676 of SEQ ID NO:14, which has β-xylosidase and/or L-α-arabinofuranosidase activities.

115. An isolated polypeptide, wherein said polypeptide has at least 90% sequence identity to the amino acid sequence of SEQ ID NO:16, or to the amino acid sequence corresponding to residues (i) 19-340, (ii) 53-340, (iii) 19-383, or (iv) 53-383 of SEQ ID NO:16, which has β-xylosidase activity.

116. An isolated polypeptide, wherein said polypeptide has at least 90% sequence identity to the amino acid sequence of SEQ ID NO:18, or to the amino acid sequence corresponding to residues (i) 21-341, (ii) 107-341, (iii) 21-348, or (iv) 107-348 of SEQ ID NO:18, which has β-xylosidase activity.

117. An isolated polypeptide, wherein said polypeptide has at least 90% sequence identity to the amino acid sequence of SEQ ID NO:20, or to the amino acid sequence corresponding to residues (i) 15-558, or (ii) 15-295 of SEQ ID NO:20, which has L-α-arabinofuranosidase activity.

118. An isolated polypeptide, wherein said polypeptide has at least 90% sequence identity to the amino acid sequence of SEQ ID NO:22, or to the amino acid sequence corresponding to residues (i) 21-632, (ii) 461-632, (iii) 21-642, or (iv) 461-642 of SEQ ID NO:22, which has L-α-arabinofuranosidase activity.

119. An isolated polypeptide, wherein said polypeptide has at least 90% sequence identity to the amino acid sequence of SEQ ID NO:2, or to the amino acid sequence corresponding to residues (i) 24-766, (ii) 73-321, (iii) 73-394, (iv) 395-622, (v) 24-622, or (iv) 73-622 of SEQ ID NO:2, which has β-xylosidase activity.

120. An isolated polypeptide, wherein said polypeptide has at least 90% sequence identity to the amino acid sequence of SEQ ID NO:32, or to the amino acid sequence corresponding to residues (i) 21-660, (ii) 21-645, (iii) 450-645, or (iv) 450-660 of SEQ ID NO:32, which has L-α-arabinofuranosidase activity.

121. A host cell engineered to recombinantly express a polypeptide of any one of claims 108-120.

122. The host cell of claim 121, which is a cell of a filamentous fungus.

123. The host cell of claim 122, which is a cell of a *Trichoderma, Humicola, Fusarium, Aspergillus, Neurospora, Penicillium, Cephalosporium, Achlya, Podospora, Endothia, Mucor, Cochliobolus, Pyricularia, Phanerochaete,* or *Chrysosporium.*

124. The host cell of claim 122, which is a cell of a *Aspergillus awamori, Aspergillus fumigatus, Aspergillus foetidus, Aspergillus japonicus, Aspergillus nidulans, Aspergillus niger, Aspergillus oryzae, Chrysosporium lucknowense, Fusarium bactridioides, Fusarium cerealis, Fusarium crookwellense, Fusarium culmorum, Fusarium graminearum, Fusarium graminum, Fusarium heterosporum, Fusarium negundi, Fusarium oxysporum, Fusarium reticulatum, Fusarium roseum, Fusarium sambucinum, Fusarium sarcochroum, Fusarium sporotrichioides, Fusarium sulphureum, Fusarium torulosum, Fusarium trichothecioides, Fusarium venenatum, Bjerkandera adusta, Ceriporiopsis aneirina, Ceriporiopsis aneirina, Ceriporiopsis caregiea, Ceriporiopsis gilvescens, Ceriporiopsis pannocinta, Ceriporiopsis rivulosa, Ceriporiopsis subrufa, Ceriporiopsis subvermispora, Coprinus cinereus, Coriolus hirsutus, Humicola insolens, Humicola lanuginosa, Mucor miehei, Myceliophthora thermophila, Neurospora crassa, Neurospora intermedia, Penicillium purpurogenum, Penicillium canescens, Penicillium solitum, Penicillium funiculosum, Phanerochaete chrysosporium, Phlebia radiate, Pleurotus eryngii, Talaromyces flavus, Thielavia terrestris, Trametes villosa, Trametes versicolor, Trichoderma harzianum, Trichoderma koningii, Trichoderma longibrachiatum, Trichoderma reesei,* or *Trichoderma viride.*

125. The host cell of claim 122 or 123, which is a cell of a *Trichoderma spp.*, a *Pencillium,* an *Aspergillus,* or a *Fusarium.*

126. The host cell of claim 125, wherein:

(a) the *Trichoderma spp.* is *Trichoderma reesei;* or

(b) the *Pencillium* is *Pencillium funiculosum;* or

(c) the *Aspergillus* is *Aspergillus oryzae* or *Aspergillus nidulans;* or

(d) the *Fusarium* is *Fusarium verticillioides* or *Fusarium oxysporum.*

127. The host cell of claim 121, which is a cell of a bacterium.

128. The host cell of claim 127, wherein the bacterium is a *Streptomyces,* a *Thermomonospora,* a *Bacillus,* or a *Cellulomonas.*

129. A method of producing the composition of any one of claims 1-80 or the fermentation broth of any one of claims 81-88.

130. A process of saccharification comprising treating a biomass material comprising hemicellulose with the composition of any one of claims 1-80 or the fermentation broth of any one of claims 81-88.

131. The process of claim 130, wherein the biomass material further comprises cellulose.

132. The process of claim 130 or 131, wherein the amount of polypeptides having xylanase activity is 1 g to 40 g per kg of hemicellulose in the biomass material.

133. The process of claim 130 or 131, wherein the amount of polypeptides having xylanase activity is 0.5 g to 40 g per kg of hemicellulose in the biomass material.

135. The process of any one of claims 130-133, wherein the amount of polypeptides having β-xylosidase activity is 1 g to 50 g per kg of hemicellulose in the biomass material.

136. The process of any one of claims 130-133, wherein the amount of polypeptide having β-xylosidase activity is 0.5 g to 50 g per kg of hemicellulose in the biomass material.

137. The process of any one of claims 130-136, wherein the amount of polypeptides having L-α-arabinofuranosidase activity is 0.5 g to 20 g per kg of hemicellulose in the biomass material.

138. The process of any one of claims 130-136, wherein the amount of polypeptides having L-α-arabinofuranosidase activity is 0.2 g to 20 g per kg of hemicellulose in the biomass material.

139. The process of any one of claims 130-138, wherein the amount of polypeptides having cellulase activity is 1 g to 100 g per kg of cellulose in the biomass material.

140. The process of any one of claims 130-139, wherein the amount of polypeptides having β-glucosidase activity is 50% or less of the total weight of polypeptides having cellulase activity.

ABSTRACT

The present disclosure is generally directed to glycosyl hydrolase enzymes, compositions comprising such enzymes, and methods of using the enzymes and compositions, for example for the saccharification of cellulosic and hemicellulosic materials into sugars.

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

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 *Fusarium verticillioides*
28.	Amino acid	Protein sequence of Fv43D
29.	Nucleotide	Nucleotide sequence for Pf43B, a GH43 family enzyme from *Penicillium funiculosum*
30.	Amino acid	Protein sequence of Pf43B
31.	Nucleotide	Nucleotide sequence for Fv51A, a GH51 family enzyme *Fusarium verticilloides*
32.	Amino acid	Protein sequence of Fv51A
33.	Nucleotide	Nucleotide sequence for Cg51B, a GH51 family enzyme from *Chaetomium globosum*
34.	Amino acid	Protein sequence of Cg51B
35.	Nucleotide	Nucleotide sequence for Fv43C, a GH43 family enzyme from *Fusarium verticillioides*
36.	Amino acid	Fv43C protein sequence
37.	Nucleotide	Nucleotide sequence for Fv30A, a GH30 family enzyme from *Fusarium verticillioides*
38.	Amino acid	Fv30A protein sequence
39.	Nucleotide	Nucleotide sequence for Fv43F, a GH43 family enzyme from *Fusarium verticillioides*
40.	Amino acid	Fv43F protein sequence
41.	Nucleotide	Nucleotide sequence for Xyn3, a GH10 family xylanase from *Trichoderma reesei*
42.	Amino acid	Xyn3 protein sequence
43.	Amino acid	Protein sequence of Xyn2, a GH11 xylanase from *Trichoderma reesei*
44.	Amino acid	Protein sequence of Bxl1, a GH3 β-xylosidase from *Trichoderma reesei*
45.	Amino acid	Protein sequence of Bgl1, a GH3 β-glucosidase from *Trichoderma reesei*
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)	*Magnaporthe grisea*
At51A	XP_001212060 (NCBI)	*Aspergillus terreus*
Pt51A	XP_001939835 (NCBI)	*Pyrenophora tritici-repentis*
Ss51A	XP_001595513 (NCBI)	*Sclerotinia sclerotiorum*
Vd51A	EEU33949 (NCBI)	*Verticilium dahliae*
AfuXyn3	Q4WLV2 (Uniprot)	*Aspergillus fumigatus*
PfuXyn1	Q8J0K5 (Uniprot)	*Penicillium funiculosum*
SspXyn1	CAA67143 (NCBI)	*Streptomyces 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	25.642	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	0.057	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	0.187	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	0.985	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

SUBSTITUTE SHEET (RULE 26)

151

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	69.85	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

SUBSTITUTE SHEET (RULE 26)

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

SUBSTITUTE SHEET (RULE 26)

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

SUBSTITUTE SHEET (RULE 26)

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	Ara
A	**5.6%**	**73.2%**	**69.4%**	5.8%	76.2%	71.8%	2.6%	84.8%	71.6%	2.1%	92.6%	92.3%
B	**3.2%**	**62.6%**	**42.6%**	4.8%	72.0%	64.6%	59.8%	83.1%	70.3%	2.0%	95.9%	93.6%
C	**1.8%**	**73.3%**	**68.3%**	14.3%	70.6%	67.0%	2.5%	98.5%	81.1%	1.8%	96.1%	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	180	47.5	14	49.3%	67.4%	57.1%
2	130	145	40	21	49.4%	78.3%	61.8%
3	90	180	47.5	14	53.5%	74.4%	62.5%
4	130	145	52	23	51.7%	78.0%	63.0%
5	90	180	47.5	14	53.5%	76.5%	63.4%
6	90	180	47.5	14	47.0%	67.5%	55.8%
7	90	180	47.5	14	51.0%	71.4%	59.8%
8	170	180	47.5	14	50.0%	64.4%	56.2%
9	90	180	45.5	28	60.9%	84.0%	70.8%
10	90	180	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

SUBSTITUTE SHEET (RULE 26)

	g/L			% Conversion		
Sample Name pH / Temp	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

SUBSTITUTE SHEET (RULE 26)

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

atgctgctcaatcttcaggtcgctgccagcgctttgtcgctttctcttttaggtggattggctgaggctg
ctacgccatataccttccggactgtaccaaaggacctttgagcaagaatggaatctgcgatacttcgtt
atctccagctaaaagagcggctgctctagttgctgctctgacgcccgaagagaaggtgggcaatctggtc
aggtaaaatataccccccccccataatcactattcggagattggagctgacttaacgcagcaatgcaactg
gtgcaccaagaatcggacttccaaggtacaactggtggaacgaagcccttcatggcctcgctggatctcc
aggtggtcgctttgccgacactcctccctacgacgcggccacatcatttcccatgcctcttctcatggcc
gctgctttcgacgatgatctgatccacgatatcggcaacgtcgtcggcaccgaagcgcgtgcgttcacta
acggcggttggcgcggagtcgacttctggacacccaacgtcaaccctttaaagatcctcgctggggtcg
tggctccgaaactccaggtgaagatgcccttcatgtcagccggtatgctcgctatatcgtcaggggtctc
gaaggcgataaggagcaacgacgtattgttgctacctgcaagcactatgctggaaacgactttgaggact
ggggaggcttcacgcgtcacgactttgatgccaagattactcctcaggacttggctgagtactacgtcag
gcctttccaggagtgcacccgtgatgcaaaggttggttccatcatgtcgcctacaatgccgtgaacggc
attcccgcatgcgcaaactcgtatctgcaggagacgatcctcagagggcactggaactggacgcgcgata
acaactggatcactagtgattgtggcgccatgcaggatatctggcagaatcacaagtatgtcaagaccaa
cgctgaaggtgcccaggtagcttttgagaacggcatggattctagctgcgagtatactactaccagcgat
gtctccgattcgtacaagcaaggcctcttgactgagaagctcatggatcgttcgttgaagcgccttttcg
aagggcttgttcatactggtttctttgacggtgccaaagcgcaatggaactcgctcagttttgcggatgt
caacaccaaggaagctcaggatcttgcactcagatctgctgtggagggtgctgttcttcttaagaatgac
ggcactttgcctctgaagctcaagaagaaggatagtgttgcaatgatcggattctgggccaacgatactt
ccaagctgcagggtggttacagtggacgtgctccgttcctccacagcccgctttatgcagctgagaagct
tggtcttgacaccaacgtggcttggggtccgacactgcagaacagctcatctcatgataactggaccacc
aatgctgttgctgcggcgaagaagtctgattacattctctactttggtggtcttgacgcctctgctgctg
gcgaggacagagatcgtgagaaccttgactggcctgagagccagctgacccttcttcagaagctctctag
tctcggcaagccactggttgttatccagcttggtgatcaagtcgatgacaccgctcttttgaagaacaag
aagattaacagtattctttgggtcaattaccctggtcaggatggcggcactgcagtcatggacctgctca
ctggacgaaagagtcctgctggccgactacccgtcacgcaatatcccagtaaatacactgagcagattgg
catgactgacatggacctcagacctaccaagtcgttgccagggagaacttatcgctggtactcaactcca
gttcttccctacggctttggcctccactacaccaagttccaagccaagttcaagtccaacaagttgacgt
ttgacatccagaagcttctcaagggctgcagtgctcaatactccgatacttgcgcgctgccccccatcca
agttagtgtcaagaacaccggccgcattacctccgactttgtctctctggtctttatcaagagtgaagtt
ggacctaagccttaccctctcaagaccttgcggcttatggtcgcttgcatgatgtcgcgccttcatcga
cgaaggatatctcactggagtggacgttggataacattgcgcgacggggagagaatggtgatttggttgt
ttatcctgggacttacactctgttgctggatgagcctacgcaagccaagatccaggttacgctgactgga
aagaaggctattttggataagtggcctcaagaccccaagtctgcgtaa

FIG. 1A

SEQ ID NO:2
Protein sequence of Fv3A

mllnlqvaasalslsllqqlaeaatpytlpdctkgplsknqicdtslspakraaalvaaltpeekvgnlv
snATGAPRIGLPRYNWWNEALHGLAGSPGGRFADTPPYDAATSFPMPLLMAAAFDDDLIHDIGNVVGTEA
RAFTNGGWRGVDFWTPNVNPFKDPRWGRGSETPGEDALHVSRYARYIVRGLEGDKEQRRIVATCKHYAGN
DFEDWGGFTRHDFDAKITPQDLAEYYVRPFQECTRDAKVGSIMCAYNAVNGIPACANSYLQETILRGHWN
WTRDNNWITSDCGAMQDIWQNHKYVKTNAEGAQVAFENGMDssceytttsdvsdsykqglliteklmdrsl
krlfeglvhtgffdgakaqwnslsfadvntkeaqdlalrsavegAVLLKNDGTLPLKLKKKDSVAMIGFW
ANDTSKLQGGYSGRAPFLHSPLYAAEKLGLDTNVAWGPTLQNSSSHDNWTTNAVAAAKKSDYILYFGGLD
ASAAGEDRDRENLDWPESQLTLLQKLSSLGKPLVVIQLGDQVDDTALLKNKKINSILWVNYPGQDGGTAV
MDLLTGRKSPAGRLPVTQYPSKYTEQIGMTDMDLRPTKSLPGRTYRWYSTPVLPYGFGLHYTkfqakfks
nkltfdiqkllkgcsaqysdtcalppiqvsvkntgritsdfvslvfiksevgpkpyplktlaaygrlhdv
apsstkdislewtldniarrgengdlvvypgtytllldeptqakiqvtltgkkaildkwpqdpksa

FIG. 1B

175

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

atgcttcagcgatttgcttatattttaccactggctctattgagtgttggagtgaaagccgacaacccct
ttgtgcagagcatctacaccgctgatccggcaccgatggtatacaatgaccgcgtttatgtcttcatgga
ccatgacaacaccggagctacctactacaacatgacagactggcatctgttctcgtcagcagatatggcg
aattggcaagatcatggcattccaatgagcctggccaatttcacctgggccaacgcgaatgcgtgggccc
cgcaagtcatccctcgcaacggccaattctacttttatgctcctgtccgacacaacgatggttctatggc
tatcggtgtgggagtgagcagcaccatcacaggtccataccatgatgctatcggcaaaccgctagtagag
aacaacgagattgatcccaccgtgttcatcgacgatgacggtcaggcatacctgtactgggaaatccag
acctgtggtacgtcaaattgaaccaagatatgatatcgtacagcgggagccctactcagattccactcac
cacggctggatttggtactcgaacgggcaatgctcaacggccgaccacttttgaagaagctccatgggta
tacaaacgcaacggcatctactatatcgcctatgcagccgattgttgttctgaggatattcgctactcca
cgggaaccagtgccactggtccgtggacttatcgaggcgtcatcatgccgacccaaggtagcagcttcac
caatcacgagggtattatcgacttccagaacaactcctacttttttctatcacaacggcgctcttcccggc
ggaggcggctaccaacgatctgtatgtgtggagcaattcaaatacaatgcagatggaaccattccgacga
tcgaaatgaccaccgccggtccagctcaaattgggactctcaacccttacgtgcgacaggaagccgaaac
ggcggcatggtcttcaggcatcactacggaggtttgtagcgaaggcggaattgacgtcgggtttatcaac
aatggcgattacatcaaagttaaaggcgtagctttcggttcaggagcccattctttctcagcgcgggttg
cttctgcaaatagcggcggcactattgcaatacacctcggaagcacaactggtacgctcgtgggcacttg
tactgtccccagcactggcggttggcagacttggactaccgttacctgttctgtcagtggcgcatctggg
acccaggatgtgtattttgttttcggtggtagcggaacaggatacctgttcaactttgattattggcagt
tcgcataa

FIG. 2A

SEQ ID NO:4
Protein sequence of Pf43A

mlqrfayilplallsvqvka**dnpfvqsiytadpapmvyndrvyvfmdhdntgatyynmtdwhlfssadma**
nwqdhgipmslanftwananawapqvipmngqfyfyapvrhndgsmaigvgvsstitgpyhdaigkplve
nneidptvfidddgqaylywgnpdlwyvklnqdmisysgsptqipllttagfgtrtgnaqrpttfeeapwv
ykrngiyyiayaadccsedirystgtsatgpwtyrgvimptqgssftnhegiidfqnnsyffyhngalpg
gggyqrsvcveqfkynadgti*ptiemttagpaqigtlnpyvrq*EAETAAWSSGITTEVCSEGGIDVGFIN
NGDYIKVKGVAFGSGAHSFSARVASANSGGTIAIHLGSTTGTLVGTCTVPSTGGWQTWTTVTCSVSGASG
TQDVYFVFGGSGTGYLFNFDYWQFa

FIG. 2B

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

atgaaggtatactggctcgtggcgtgggccacttctttgacgccggcactggctggcttgattggacacc
gtcgcgccaccaccttcaacaatcctatcatctactcagactttccagataacgatgtattcctcggtcc
agataactactactacttctctgcttccaacttccacttcagcccaggagcacccgttttgaagtctaaa
gatctgctaaactgggatctcatcggccattcaattccccgcctgaactttggcgacggctatgatcttc
ctcctggctcacgttattaccgtggaggtacttgggcatcatcctcagatacagaaagagcaatggaca
gtggtactggatcggctgcatcaacttctggcagacctgggtatacactgcctcatcgccggaaggtcca
tggtacaacaagggaaacttcggtgataacaattgctactacgacaatggcatactgatcgatgacgatg
ataccatgtatgtcgtatacggttccggtgaggtcaaagtatctcaactatctcaggacggattcagcca
ggtcaaatctcaggtagttttcaagaacactgatattggggtccaagacttggagggtaaccgcatgtac
aagatcaacgggctctactatatcctaaacgatagcccaagtggcagtcagacctggatttggaagtcga
aatcaccctgggcccttatgagtctaaggtcctcgccgacaaagtcaccccgcctatctctggtggtaa
ctcgccgcatcagggtagtctcataaagactcccaatggtggctggtacttcatgtcattcacttgggcc
tatcctgccggccgtcttccggttcttgcaccgattacgtggggtagcgatggtttccccattcttgtca
agggtgctaatggcggatgggatcatcttacccaacacttcctggcacggatggtgtgacaaagaattg
gacaaggactgatacctccgcggaacctcacttgctccgtcctgggagtggaaccataatccggacgtc
aactccttcactgtcaacaacggcctgactctccgcactgctagcattacgaaggatatttaccaggcga
ggaacacgctatctcaccgaactcatggtgatcatccaacaggaatagtgaagattgatttctctccgat
gaaggacggccgaccgggccgggctttcagcgtttcgagaccaaagtgcatacatcggtattcatcgagat
aacggaaagttcacaatcgctacgaagcatgggatgaatatggatgagtggaacggaacaacaacagacc
tgggacaaataaaagccacagctaatgtgccttctggaaggaccaagatctggctgagacttcaacttga
taccaacccagcaggaactggcaacactatcttttcttacagttgggatggagtcaagtatgaaacactg
ggtcccaacttcaaactgtacaatggttgggcattctttattgcttaccgattcggcatcttcaacttcg
ccgagacggctttaggaggctcgatcaaggttgagtctttcacagctgcatag

FIG. 3A

SEQ ID NO:6
Protein sequence of Fv43E

mkvywlvawatsltpalaglighrrattfnnpiiysdfpdndvflgpdnyyyfsasnfhfspgapvlksk
dllnwdlighsiprlnfgdgydlppgsryyrggtwasslryrksngqwywigcinfwqtwvytasspegp
wynkgnfgdnncyydngilidddtmyvvygsgevkvsqlsqdgfsqvksqvvfkntdigvqdlegnrmy
kinglyyilndspsgsqtwiwkskspwgpyeskvladkvtppisggnsphqgsliktpnggwyfmsftwa
ypagrlpvlapitwgsdgfpilvkganggwgssyptlpgtdgvtknwtrtdtfrgtslapswewnhnpdv
nsftvnngltlrtasitkdiyqarntlshrthgdhptgivkidfspmkdgdraglsafrdqsayigihrd
ngkftiatkhqmnmdewngtttdlgqikatanvpsgrtkiwlrlqldtnpagtgntifsyswdgvkyetl
gpnfklyngwaffiayrfgifnfaetalggsikvesftaa

FIG. 3B

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

atgcactacgctaccctcaccactttggtgctggctctgaccaccaacgtcgctgcacagcaaggcacag
caactgtcgacctctccaaaaatcatggaccggcgaaggcccttggttcaggcttcatatacggctggcc
tgacaacggaacaagcgtcgacacctccataccagatttcttggtaactgacatcaaattcaactcaaac
cgcggcggtggcgcccaaatcccatcactgggttgggccagaggtggctatgaaggatacctcggccgct
tcaactcaaccttatccaactatcgcaccacgcgcaagtataacgctgactttatcttgttgcctcatga
cctctgggtgcggatggcgggcagggttcaaactccccgtttcctggcgacaatggcaattggactgag
atggagttattctggaatcagcttgtgtctgacttgaaggctcataatatgctggaaggtcttgtgattg
atgtttggaatgagcctgatattgatatcttttgggatcgcccgtggtcgcagtttcttgagtattacaa
tcgcgcgaccaaactacttcggtgagtctactactgatccatacgtatttacagtgagctgactggtcga
attagaaaaacacttcccaaaactcttctcagtggcccagccatggcacattctcccattctgtccgatg
ataaatggcatacctggcttcaatcagtagcgggtaacaagacagtccctgatatttactcctggcatca
gattggcgcttgggaacgtgagccggacagcactatccccgactttaccaccttgcgggcgcaatatggc
gttcccgagaagccaattgacgtcaatgagtacgctgcacgcgatgagcaaaatccagccaactccgtct
actacctctctcaactagagcgtcataaccttagaggtcttcgcgcaaactggggtagcggatctgacct
ccacaactggatgggcaacttgatttacagcactaccggtacctcggaggggacttactaccctaatggt
gaatggcaggcttacaagtactatgcggccatggcagggcagagacttgtgaccaaagcatcgtcggact
tgaagtttgatgtctttgccactaagcaaggccgtaagattaagattatagccggcacgaggaccgttca
agcaaagtataacatcaaaatcagcggtttggaagtagcaggacttcctaagatgggtacggtaaaggtc
cggacttatcggttcgactgggctgggccgaatggaaaggttgacgggcctgttgatttgggggagaaga
agtatacttattcggccaatacggtgagcagcccctctacttga

FIG. 4A

SEQ ID NO:8
Protein sequence of Fv39A

mhyatlttlvlalttnvaaqggtatvdlsknhgpakalgsgfiygwpdngtsvdtsipdflvtdikfnsn
rgggaqipslgwarggyegylgrfnstlsnyrttrkynadfillphdlwgadgggqgsnspfpgdngnwte
melf**wnqlvsdlkahnmleglvidvwnepdidifwdrpwsqfleyynratkllrktlpktllsgpamahs**
pilsddkwhtwlqsvagnktvpdiyswhqigawerepdstipdfttlraqygvpekpidvneyaardeqn
pansvyylsqlerhnlrglranwgsgsdlhnwmgnliysttgtsegtyypngewqaykyyaamagqrlvt
kassdlkfdvfatkqgrkikiiagtrtvqakynikisglevaglpkmgtvkvrtyrfdwagpngkvdgpv
dlgekkytysantvsspst

FIG. 4B

SEQ ID NO:9
Nucleotide sequence for Fv43A, a GH43 family enzyme from *Fusarium verticillioides*

atgtggctgacctccccattgctgttcgccagcaccctcctgggcctcactggcgttgctctagcagaca
accccatcgtccaagacatctacaccgcagacccagcaccaatggtctacaatggccgcgtctacctctt
cacaggccatgacaacgacggctctaccgacttcaacatgacagactggcgtctcttctcgtcagcagac
atggtcaactggcagcaccatggtgtccccatgagcttaaagaccttcagctgggccaacagcagagcct
gggctggtcaagtcgttgcccgaaacggaaagttttacttctatgttcctgtccgtaatgccaagacggg
tggaatggctattggtgtcggtgttagtaccaacatccttgggccctacactgatgcccttggaaagcca
ttggtcgagaacaatgagatcgacccaactgtctacatcgacactgatggccaggcctatctctactggg
gcaaccctggattgtactacgtcaagctcaaccaagacatgctctcctacagtggtagcatcaacaaagt
atcgctcacaacagctggattcggcagccgcccgaacaacgcgcagcgtcctactactttcgaggaagga
ccgtggctgtacaagcgtggaaatctctactacatgatctacgcagccaactgctgttccgaggacattc
gctactcaactggacccagcgccactggaccttggacttaccgcggtgtcgtgatgaacaaggcgggtcg
aagcttcaccaaccatcctggcatcatcgactttgagaacaactcgtacttctttaccacaatggcgct
cttgatggaggtagcggttatactcggtctgtggctgtcgagagcttcaagtatggttcggacggtctga
tccccgagatcaagatgactacgcaaggcccagcgcagctcaagtctctgaacccatatgtcaagcagga
ggccgagactatcgcctggtctgagggtatcgagactgaggtctgcagcgaaggtggtctcaacgttgct
ttcatcgacaatggtgactacatcaaggtcaagggagtcgactttggcagcaccggtgcaaagacgttca
gcgcccgtgttgcttccaacagcagcggaggcaagattgagcttcgacttggtagcaagaccggtaagtt
ggttggtacctgcacggtaacgactacgggaaactggcagacttataagactgtggattgccccgtcagt
ggtgctactggtacgagcgatctattctttgtcttcacgggctctgggtctggctctctgttcaacttca
actggtggcagtttagctaa

FIG. 5A

SEQ ID NO:10
Protein sequence of Fv43A

mwltspllfastllqltqvala**dnpivqdiytadpapmvyngrvylftghdndgstdfnmtdwrlfssad**
mvnwqhhgvpmslktfswansrawagqvvarngkfyfyvpvrnaktggmaigvgvstnilgpytdalgkp
lvenneidptvyidtdgqaylywgnpglyyvklnqdmlsysgsinkvslttagfgsrpnnaqrpttfeeg
pwlykrgnlyymiyaanccsedirystgpsatgpwtyrgvvmnkagrsftnhpgiidfennsyffyhnga
ldggsgytrsvavesfkygsdg*lipeikmttggpaqlksl*NPYVKQEAETIAWSEGIETEVCSEGGLNVA
FIDNGDYIKVKGVDFGSTGAKTFSARVASNSSGGKIELRLGSKTGKLVGTCTVTTTGNWQTYKTVDCPVS
GAIGTSDLFFVFTGSGSGSLFNFNWWQFs

FIG. 5B

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

atgcgcttctcttggctattgtgccccttctagcgatgggaagtgctcttcctgaaacgaagacggatg
tttcgacatacaccaaccctgtccttccaggatggcactcggatccatcgtgtatccagaaagatggcct
ctttctctgcgtcacttcaacattcatctccttcccaggtcttcccgtctatgcctcaagggatctagtc
aactggcgtctcatcagccatgtctggaaccgcgagaaacagttgcctggcattagctggaagacggcag
gacagcaacagggaatgtatgcaccaaccattcgataccacaagggaacatactacgtcatctgcgaata
cctgggcgttggagatattattggtgtcatcttcaagaccaccaatccgtgggacgagagtagctggagt
gaccctgttaccttcaagccaaatcacatcgaccccgatctgttctgggatgatgacggaaaggtttatt
gtgctacccatggcatcactctgcaggagattgatttggaaactggagagcttagcccggagcttaatat
ctggaacggcacaggaggtgtatggcctgagggtccccatatctacaagcgcgacggttactactatctc
atgattgccgagggtggaactgccgaagaccacgctatcacaatcgctcgggcccgcaagatcaccggcc
cctatgaagcctacaataacaacccaatcttgaccaaccgcgggacatctgagtacttccagactgtcgg
tcacggtgatctgttccaagataccaagggcaactggtggggtctttgtcttgctactcgcatcacagca
cagggagtttcacccatgggccgtgaagctgttttgttcaatggcacatggaacaagggcgaatggccca
agttgcaaccagtacgaggtcgcatgcctggaaacctcctcccaaagccgacgcgaaacgttcccggaga
tgggcccttcaacgctgacccagacaactacaacttgaagaagactaagaagatccctcctcactttgtg
caccatagagtcccaagagacggtgccttctctttgtcttccaagggtctgcacatcgtgcctagtcgaa
acaacgttaccggtagtgtgttgccaggagatgagattgagctatcaggacagcgaggtctagctttcat
cggacgccgccaaactcacactctgttcaaatatagtgttgatatcgacttcaagcccaagtccgatgat
caggaagctggaatcaccgttttccgcacgcagttcgaccatatcgatcttggcattgttcgtcttccta
caaaccaaggcagcaacaagaaatctaagcttgccttccgattccgggccacaggagctcagaatgttcc
tgcaccgaaggtagtaccggtccccgatggctgggagaagggcgtaatcagtctacatatcgaggcagcc
aacgcgacgcactacaaccttggagcttcgagccacagaggcaagactctcgacatcgcgacagcatcag
caagtcttgtgagtggaggcacgggttcatttgttggtagtttgcttggaccttatgctacctgcaacgg
caaaggatctggagtggaatgtcccaagggaggtgatgtctatgtgacccaatggacttataagcccgtg
gcacaagagattgatcatggtgtttttgtgaaatcagaattgtag

FIG. 6A

SEQ ID NO:12
Protein sequence of Fv43B

mrfswllcpllamqsalpetktdvst**ytnpvlpgwhsdpsciqkdglflcvtstfisfpglpvyasrdlv**
nwrlishvwnrekqlpgiswktagqqqgmyaptiryhkgtyyviceylgvgdiigvifkttnpwdesswS
dpvtfkpnhidpdlfwdddgkvycathgitlqeidletgelspelniwngtggvwpegphiykrdgyyyl
miaeggtaedhaitiararkitgpyeaynnnpiltnrgtseyfqtvghgdlfqdtkgnwwglclatrita
qgvspmgreavlfngtwnkgewpklqpvrgrmpgnllpkptrnvpgdgpfnadpdnynlkktkkipphfv
hhrvprdgafslsskglhivpsrnnvtgsvlpgdeielsgqrglafigrrqthtlfkysvdidfkpksdd
qeagitvfrtqfdhidlgivrlptnqgsnkksklafrfratgaqnvpapkvvpvpdgwekgvislhieaa
nathynlgasshrgktldiatasaslvsggtgsfvgsllgpyatcngkgsgvecpkggdvyv
tqwtykpvaqeidhgvfvksel

FIG. 6B

180

SEQ ID NO:13
Nucleotide sequence for Pa51A, a GH51 family enzyme from *Podospora anserina*

atgatccacctcaagccagccctcgcggcgttgttggcgctgtcgacgcaatgtgtggctattgatttgtttgtcaagtc
ttcgggggggaataagacgactgatatcatgtatggtcttatgcacgaggtatgtgttttgcgagatctccctttgttt
ttgcgcactgctgacatggagactgcaaacaggatatcaacaactccggcgacggcggcatctacgccgagctaatctcc
aaccgcgcgttccaagggagtgagaagttcccctccaacctcgacaactggagccccgtcggtggcgctacccttaccct
tcagaagcttgccaagccccttttcctctgcgttgccttactccgtcaatgttgccaaccccaaggagggcaagggcaagg
gcaaggacaccaaggggaagaaggttggcttggccaatgctgggttttggggtatggatgtcaagaggcagaagtacact
ggtagcttccacgttactggtgagtacaagggtgactttgaggttagcttgcgcagcgcgattaccggggagacctttgg
caagaaggtggtgaagggtgggagtaagaaggggaagtggaccgagaaggagtttgagttggtgcctttcaaggatgcgc
ccaacagcaacaacacctttgttgtgcagtgggatgccgaggtatgtgcttctttgatattggctgagatagaagttggg
ttgacatgatgtggtgcagggcgcaaaggacggatctttggatctcaacttgatcagcttgttccctccgacattcaagg
gaaggaagaatgggctgagaattgatcttgcgcagacgatggttgagctcaagccggtaagtcctctctagtcagaaaag
tagagcctttgttaacgcttgacagaccttcttgcgcttccccggtggcaacatgctcgagggtaacaccttggacactt
ggtggaagtggtacgagaccattggccctctgaaggatcgcccgggcatggctggtgtctgggagtaccagcaaaccctt
ggcttgggtctggtcgagtacatggagtgggccgatgacatgaacttggagcccagtatgtgatcccattttctggagtg
acttctcttgctaacgtatccacagttgtcggtgtcttcgctggtcttgccctcgatggctcgttcgttcccgaatccga
gatgggatgggtcatccaacaggctctcgacgaaatcgagttcctcactggcgatgctaagaccaccaaatggggtgccg
tccgcgcgaagcttggtcaccccaagccttggaaggtcaagtgggttgagatcggtaacgaggattggcttgccggacgc
cctgctggcttcgagtcgtacatcaactaccgcttccccatgatgatgaaggccttcaacgaaaagtaccccgacatcaa
gatcatcgcctcgccctccatcttcgacaacatgacaatccccgcgggtgctgccggtgatcaccacccgtacctgactc
ccgatgagttcgttgagcgattcgccaagttcgataacttgagcaaggataacgtgacgctcatcggcgaggctgcgtcg
acgcatcctaacggtggtatcgcttgggagggagatctcatgcccttgccttggtggggcggcagtgttgctgaggctat
cttcttgatcagcactgagagaaacggtgacaagatcatcggtgctacttacgcgcctggtcttcgcagcttggaccgct
ggcaatggagcatgacctgggtgcagcatgccgccgacccggcccctcaccactcgctcgaccagttggtatgtctggaga
atcctcgcccaccacatcatccgtgagacgctcccggtcgatgccccggccggcaagcccaactttgaccctctgttcta
cgttgccggaaagagcgagagtggcaccggtatcttcaaggctgccgtctacaactcgactgaatcgatcccggtgtcgt
tgaagtttgatggtctcaacgagggagcggttgccaacttgacggtgcttactgggccggaggatccgtatggatacaac
gacccccttcactggtatcaatgttgtcaaggagaagaccaccttcatcaaggccggaaagggcggcaagttcaccttcac
cctgccgggcttgagtgttgctgtgttggagacggccgacgcggtcaagggtggcaagggaaagggcaagggcaagggaa
agggtaactga

FIG. 7A

SUBSTITUTE SHEET (RULE 26)

181

SEQ ID NO:14
Protein sequence of Pa51A

mihlkpalaallalstgcvaidlfvkssggnkttdimyglmhedinnsgdggiyaelisnrafqgsekfp
snldnwspvggatlllqklakplssalpysvnvanpkegkgkgkdtkgkkvglanagfwgmdvkrqkytg
sfhvtgeykgdfevslrsaitgetfgkkvvkggskkgkwtekefelvpfkdapnsnntfvvqwdaegakd
gsldlnlislfpptfkgrknglridlaqtmvelkptflrfpggnmlegntldtwwkwyetigplkdrpgm
agvweyqqtlglglveymewaddmnlepivgvfaglaldgsfvpesemgwviqqaldeiefltgdaktk
wgavraklghpkpwkvkwveignedwlagrpagfesyinyrfpmmmkafnekypdikiiaspsifdnmti
pagaagdhhpyltpdefverfakfdnlskdnvtligeaasthpnggia**wegdlmplpwwggsvaeaifli**
sterngdkiigatyapglrsldrwqwsmtwvqhaadpalttrstswyvwrilahhiiretlpvdapagkp
nfdplfyvagksesgtgifkaavynstesipvslkfdglnegavanltvltgpedpygyndpftginvvk
ekttfikagkggkftftlpglsvavletadavkggkgkgkgkgkgn

FIG. 7B

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

atgaagtccaagttgttattcccactcctctctttcgttggtcaaagtcttgccaccaacgacgactgtcctctcatcactagtag
atggactgcggatccttcggctcatgtctttaacgacaccttgtggctctacccgtctcatgacatcgatgctggatttgagaatg
atcctgatggaggccagtacgccatgagagattaccatgtctactctatcgacaagatctacggttccctgccggtcgatcacggt
acggccctgtcagtggaggatgtcccctgggcctctcgacagatgtgggctcctgacgctgcccacaagaacggcaaatactacct
atacttccctgccaaagacaaggatgatatcttcagaatcggcgttgctgtctcaccaaccccggcggaccattcgtccccgaca
agagttggatccctcacactttcagcatcgaccccgccagtttcgtcgatgatgatgacagagcctacttggcatggggtggtatc
atgggtggccagcttcaacgatggcaggataagaacaagtacaacgaatctggcactgagccaggaaacggcaccgctgccttgag
ccctcagattgccaagctgagcaaggacatgcacactctggcagagaagcctcgcgacatgctcattcttgaccccaagactggca
agccgctcctttctgaggatgaagaccgacgcttcttcgaaggaccctggattcacaagcgcaacaagatttactacctcacctac
tctactggcacaacccactatcttgtctatgcgacttcaaagacccccctatggtcctttacacctaccagggcagaattctggagcc
agttgatggctggactactcactctagtatcgtcaagtaccagggtcagtggtggctattttatcacgatgccaagacatctggca
aggactatcttcgccaggtaaaggctaagaagatttggtacgatagcaaaggaaagatcttgacaaagaagccttga

FIG. 8A

SUBSTITUTE SHEET (RULE 26)

187

SEQ ID NO:16
Protein sequence of Gz43A

mkskllfpllsfvqqslatnddcplitsrwtadpsahvfndtlwlypshdid**agfendpdggqyamrdyh**
vysidkiygslpvdhgtalsvedvpwasrqmwapdaahkngkyylyfpakdkddifrigvavsptpggpf
vpdkswiphtfsidpasfvddddraylawggimggqlqrwqdknkynesgtepgngtaalspqiaklskd
mhtlaekprdmlildpktgkpllsededrrffegpwihkrnkiyyltystgtthylvyatsktpygpyty
qgrilepvdgwtthssivkyqgqwwlfyhdaktsgkdylrqvkakkiwydskgkiltkkp

FIG. 8B

SEQ ID NO:17
Nucleotide sequence for Fo43A, a GH43 family enzyme from *Fusarium oxysporum*

atgcagctcaagtttctgtcttcagcattgctgttctctctgaccagcaaatgcgctgcgcaagacacta
atgacattcctccccctgatcaccgacctctggtccgcagatccctcggctcatgttttcgaaggcaagct
ctgggtttacccatctcacgacatcgaagccaatgttgtcaacggcacaggaggcgctcaatacgccatg
agggattaccatacctactccatgaagagcatctatggtaaagatcccgttgtcgaccacggcgtcgctc
tctcagtcgatgacgttcccctgggcgaagcagcaaatgtgggctcctgacgcagctcataagaacggcaa
atattatctgtacttccccgccaaggacaaggatgagatcttcagaattggagttgctgtctccaacaag
cccagcggtcctttcaaggccgacaagagctggatccctggcacgtacagtatcgatcctgctagctacg
tcgacactgataacgaggcctacctcatctggggcggtatctgggcggccagctccaagcctggcagga
taaaaagaactttaacgagtcgtggattggagacaaggctgctcctaacggcaccaatgccctatctcct
cagatcgccaagctaagcaaggacatgcacaagatcaccgaaacaccccgcgatctcgtcattctcgccc
ccgagacaggcaagcctcttcaggctgaggacaacaagcgacgattcttcgagggcccttggatccacaa
gcgcggcaagctttactacctcatgtactccaccggtgatacccacttccttgtctacgctacttccaag
aacatctacggtccttatacctaccggggcaagattcttgatcctgttgatgggtggactactcatggaa
gtattgttgagtataagggacagtggtggcttttctttgctgatgcgcatacgtctggtaaggattacct
tcgacaggtgaaggcgaggaagatctggtatgacaagaacggcaagatcttgcttcaccgtccttag

FIG. 9A

SEQ ID NO:18
Protein sequence of Fo43A

mqlkflssallfsltskcaaqdtndipplitdlwsadpsahvfegklwvypshdieanvvngtggaqyam
rdyhtysmksiygkdpvvdhgvalsvddvpwakqqm**wapdaahkngkyylyfpakdkdeifrigvavsnk**
psgpfkadkswipgtysidpasyvdtdneayliwggiwggqlqawqdkknfneswigdkaapngtnalsp
qiaklskdmhkitetprdlvilapetgkplqaednkrrffegpwihkrgklyylmystgdthflvyatsk
niygpytyrgkildpvdgwtthgsiveykgqwwlffadahtsgkdylrqvkarkiwydkngkillhrp

FIG. 9B

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

atggcagctccaagtttatcctaccccacaggtatccaatcgtataccaatcctctcttccctggttggc
actccgatcccagctgtgcctacgtagcggagcaagacacctttttctgcgtgacgtccactttcattgc
cttccccggtcttcctctttatgcaagccgagatctgcagaactggaaactggcaagcaatattttcaat
cggcccagccagatccctgatcttcgcgtcacggatggacagcagtcgggtatctatgcgcccactctgc
gctatcatgagggccagttctacttgatcgtttcgtacctgggcccgcagactaagggcttgctgttcac
ctcgtctgatccgtacgacgatgccgcgtggagcgatccgctcgaattcgcggtacatggcatcgacccg
gatatcttctgggatcacgacgggacggtctatgtcacgtccgccgaggaccagatgattaagcagtaca
cactcgatctgaagacgggggcgattggcccggttgactacctctggaacggcaccggaggagtctggcc
cgagggcccgcacatttacaagagagacggatactactacctcatgatcgcagagggaggtaccgagctc
ggccactcggagaccatggcgcgatctagaacccggacaggtccctgggagccataccgcacaatccgc
tcttgtcgaacaagggcacctcggagtacttccagactgtgggccatgcggacttgttccaggatgggaa
cggcaactggtgggccgtggcgttgagcacccgatcagggcctgcatggaagaactatcccatgggtcgg
gagacggtgctcgccccgccgcttgggagaagggtgagtggcctgtcattcagcctgtgagaggccaaa
tgcaggggccgtttccaccaccaaataagcgagttcctcgcggcgagggcggatggatcaagcaacccga
caaagtggatttcaggcccggatcgaagataccggcgcacttccagtactggcgatatcccaagacagag
gattttaccgtctcccctcggggccacccgaatactcttcggctcacaccctcctttacaacctcaccg
gaactgcggacttcaagccggatgatggcctgtcgcttgttatgcgcaaacagaccgacaccttgttcac
gtacactgtggacgtgtcttttgaccccaaggttgccgatgaagaggcgggtgtgactgttttccttacc
cagcagcagcacatcgatcttggtattgtccttctccagacaaccgaggggctgtcgttgtccttccggt
tccgcgtggaaggccgcggtaactacgaaggtcctcttccagaagccaccgtgcctgttcccaaggaatg
gtgtggacagaccatccggcttgagattcaggccgtgagtgacaccgagtatgtctttgcggctgccccg
gctcggcaccctgcacagaggcaaatcatcagccgcgccaactcgttgattgtcagtggtgatacgggac
ggtttactggctcgcttgttggcgtgtatgccacgtcgaacgggggtgccggatccacgcccgcatatat
cagcagatggagatacgaaggacggggccagatgattgattttggtcgagtggtcccgagctactga

FIG. 10A

SEQ ID NO:20
Protein sequence of Af43A

maapslsyptgiqs**ytnplfpgwhsdpscayvaeqdtffcvtstfiafpglplyasrdlqnwklasnifn**
rpsqipdlrvtdgqqsgiyaptlryhegqfylivsylgpqtkgllftssdpyddaawsdplefavhgidp
difwdhdgtvyvtsaedqmikqytldlktgaigpvdylwngtggvwpegphiykrdgyyylmiaeggtel
ghsetmarsrtrtgpwepyphnpllsnkgtseyfqtvghadlfqdgngnwwavalstrsgpawknypmgr
etvlapaawekgewpviqpvrgqmqgpfpppnkrvprgeggwikqpdkvdfrpgskipahfqywrypkte
dftvsprghpntlrltpsfynltgtadfkpddglslvmrkqtdtlftytvdvsfdpkvadeeagvtvflt
qqqhidlgivllqtteglslsfrfrvegrgnyegplpeatvpvpkewcgqtirleiqavsdteyvfaaap
arhpaqrqiisranslivsgdtgrftgslvgvyatsnggagstpayisrwryegrgqmidfgrvvpsy

FIG. 10B

SEQ ID NO:21
Nucleotide sequence for Pf51A, a GH51 family enzyme from *Penicillium funiculosum*

atgggaaagatgtggcattcgatcttggttgtgttgggcttattgtctgtcgggcatgccatcactatcaacgtgtcccaaagtgg
cggcaataagaccagtcctttgcaatatggtctgatgttcgaggtaatccttctcttataccacatataaaagttgcgtcatttct
aagacaagtcaaggacataaatcacggcggtgatggcggtctgtatgcagagcttgttcgaaaccgagcattccaaggtagcaccg
tctatccagcaaacctcgatggatacgactcggtcaatggagcaatcctagcgcttcagaatttgacaaaccctctatcaccctcc
atgcctagctctctcaacgtcgccaaggggtccaacaatggaagcatcggtttcgcaaatgaaggctggtggggatagaagtcaa
gccgcaaagatacgcgggctcattctacgtccaggggactatcaaggagatttcgacatctctcttcagtcgaaattgacacaag
aagtcttcgcaacggcaaaagtcaggtcctcgggcaaacacgaggactgggttcaatacaagtacgagttggtgcccaaaaaggca
gcatcaaacaccaataacactctgaccattacttttgactcaaaggtatgttaaattttgggtttagttcgatgtctggcaattgt
cttacgagaaacgtagggattgaaagacggatccttgaacttcaacttgatcagcctatttcccccaacttacaacaatcggccca
atggcctaagaatcgacctggttgaagctatggctgaactagaggggtaagctcttacaaatcaactttatctttacgaagacta
atgtgaaaacttagaaatttctgcggtttccaggcggtagcgatgtggaaggtgtacaagctccttactggtataagtggaatgaa
acggtaggagatctcaaggaccgttatagtaggcccagtgcatggacgtacgaagaaagcaatggaattggcttgattgagtacat
gaattggtgtgatgacatggggcttgagccgagtgagtgtattccattcagcgtcaaatccagtgttctaatcatacacatcagtt
cttgccgtatgggatggacattacctttcgaacgaagtgatatcggaaaacgatttgcagccatatatcgacgacaccctcaacca
actggaattcctgatgggtgccccagatacgccatatggtagttggcgtgcgtctctgggctatccgaagccgtggacgattaact
acgtcgagattggaaacgaagacaatctatacgggggactagaaacatacatcgcctaccggtttcaggcatattacgacgctata
acagctaaatatccccatatgacggtcatggaatctttgacggagatgcctggtccggcggccgctgcaagcgattaccatcaata
ttctactcctgatgggtttgtttcccagttcaactactttgatcagatgccagtcactaatagaacactgaacggtatgaaaaccc
ccccttttttaaatatgcttttaatggtattaaccatctttcataggagagattgcaaccgtttatccaaataatcctagtaattc
ggtggcctggggaagcccattcccctttgtatccttggtggattgggtccgttgcagaagctgttttcctaattggtgaagagagga
attcgccaaagataatcggtgctagctacgtacggaattctacttttcgagattttaacattggataagaaggactaacctcaata
caggctccaatgttcagaaatatcaacaattggcagtggtctccaacactcatcgcttttgacgctgactcgtcgcgtacaagtcg
ttcaacaagctggcatgtgatcaaggtatgctaatttttcctcctcattcaaacccgcagatgtgagctaactttccgaagcttctc
tcgacaaacaaaatcacgcaaaatttacccacgacttggagtggcggtgacataggtccattatactgggtagctggacgaaacga
caatacaggatcgaacatattcaaggccgctgtttacaacagcacctcagacgtccctgtcaccgttcaatttgcaggatgcaacg
caaagagcgcaaatttgaccatcttgtcatccgacgatccgaacgcatcgaactaccctgggggggcccgaagttgtgaagactgag
atccagtctgtcactgcaaatgctcatggagcatttgagttcagtctcccgaacctaagtgtggctgttctcaaaacggagtaa

FIG. 11A

SUBSTITUTE SHEET (RULE 26)

SEQ ID NO:22
Protein sequence of Pf51A

mgkmwhsilvvlgllsvqhaitinvsqsggnktsplqyglmfedinhggdgglyaelvrnrafqgstvyp
anldgydsvngailalqnltnplspsmpsslnvakgsnngsigfanegwwgievkpqryagsfyvqgdyq
gdfdislqskltqevfatakvrssgkhedwvqykyelvpkkaasntnntltitfdskglkdgslnfnlis
lfppqynnrpnglridlveamaelegkflrfpggsdvegvqapywykwnetvgdlkdrysrpsawtyees
ngiglieymnwcddmglepilavwdghylsnevisendlqpyiddtlnqleflmgapdtpygswraslgy
pkpwtinyveignednlyqgletyiayrfqayydaitakyphmtvmesltempgpaaaasdyhqystpdg
fvsqfnyfdqmpvtnrtlngeiatvypnnpsnsvawgspfplyp**wwigsvaeavfligeernspkiigas**
yapmfrninnwqwsptliafdadssrtsrstswhviklistnkitqnlpttwsggdigplywvagrndnt
gsnifkaavynstsdvpvtvqfagcnaksanltilssddpnasnypggpevvkteiqsvtanahgafefs
lpnlsvavlkte

FIG. 11B

SEQ ID NO:23
Nucleotide sequence for AfuXyn2, a GH11 family enzyme from *Aspergillus fumigatus*

atggtttctttctcctacctgctgctggcgtgctccgccattggagctctggctgcccccgtcgaacccg
agaccacctcgttcaatgagactgctcttcatgagttcgctgagcgcgccggcaccccaagctccaccgg
ctggaacaacggctactactactccttctggactgatggcggcggcgacgtgacctacaccaatggcgcc
ggtggctcgtactccgtcaactggaggaacgtgggcaactttgtcggtggaaagggctggaaccctggaa
gcgctaggtaccgagctttgtcaacgtcggatgtgcagacctgtggctgacagaagtagaaccatcaact
acggaggcagcttcaaccccagcggcaatggctacctggctgtctacggctggaccaccaaccccttgat
tgagtactacgttgttgagtcgtatggtacatacaaccccggcagcggcggtaccttcaggggcactgtc
aacaccgacggtggcacttacaacatctacacggccgttcgctacaatgctccctccatcgaaggcacca
agaccttcacccagtactggtctgtgcgcacctccaagcgtaccggcggcactgtcaccatggccaacca
cttcaacgcctggagcagactgggcatgaacctgggaactcacaactaccagattgtcgccactgagggt
taccagagcagcggatctgcttccatcactgtctactag

FIG. 12A

SEQ ID NO:24
Protein sequence of AfuXyn2

mvsfsylllacsaigalaapvepettsfnetalhefaeragtpsst**gwnngyyysfwtdgggdvtytnga**
ggsysvnwrnvgnfvggkgwnpgsartinyggsfnpsgngylavygwttnplieyyvvesygtynpgsgg
tfrgtvntdggtyniytavrynapsiegtktftqywsvrtskrtggtvtmanhfnawsrlgmnlgthnyq
ivategyqssgsasitvy

FIG. 12B

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

atgatctccatttcctcgctcagctttggactcgccgctatcgccggcgcatatgctcttccgagtgaca
aatccgtcagcttagcggaacgtcagacgatcacgaccagccagacaggcacaaacaatggctactacta
ttccttctggaccaacggtgccggatcagtgcaatatacaaatggtgctggtggcgaatatagtgtgacg
tgggcgaaccagaacggtggtgactttacctgtgggaagggctggaatccagggagtgaccagtaggcaa
cgcccgagaactatagaagaggacgcaaagaaagcactaaactctctactagtgacattaccttctctgg
cagcttcaatccttccggaaatgcttacctgtccgtgtatggatggactaccaacccccctagtcgaatac
tacatcctcgagaactatggcagttacaatcctggctcgggcatgacgcacaagggcaccgtcaccagcg
atggatccacctacgacatctatgagcaccaacaggtcaaccagccttcgatcgtcggcacggccaccctt
caaccaatactggtccatccgccaaaacaagcgatccagcggcacagtcaccaccgcgaatcacttcaaag
gcctgggctagtctggggatgaacctgggtacccataactatcagattgtttccactgagggatatgaga
gcagcggtacctcgaccatcactgtctcgtctggtggttcttcttctggtggaagtggtggcagctcgtc
tactacttcctcaggcagctcccctactggtggctccggcagtgtaagtcttcttccatatggttgtggc
tttatgtgtattctgactgtgatagtgctctgctttgtggggccagtgcggtggaattggctggtctggt
cctacttgctgctcttcgggcacttgccaggtttcgaactcgtactactcccagtgcttgtagtaccttc
ttgcagggttatatccaagtga

FIG. 13A

SEQ ID NO:26
Protein sequence of AfuXyn5

MISISSLSFGLAAIAGAYALPSDKSVSLAERQTITTSQT**GTNNGYYYSFWTNGAGSVQYTNGAGGEYSVT**
WANQNGGDFTCGKGWNPGSDHDITFSGSFNPSGNAYLSVYGWTTNPLVEYYILENYGSYNPGSGMTHKGT
VTSDGSTYDIYEHQQVNQPSIVGTATFNQYWSIRQNKRSSGTVTTANHFKAWASLGMNLGTHNYQIVSTE
GYESSGTSTITV*SSGGSSSGGSGGSSSTTSSGSSPTGGSGSCSA***LWGQCGGIGWSGPTCCSSGTCQVSNS**
YYSQCL

FIG. 13B

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

atgcagctcaagtttctgtcttcagcattgttgctgtctttgaccggcaattgcgctgcgcaagacacta
atgatatccctcctctgatcaccgacctctggtctgcggatccctcggctcatgttttcgagggcaaact
ctgggtttacccatctcacgacatcgaagccaatgtcgtcaacggcaccggaggcgctcagtacgccatg
agagattatcacacctattccatgaagaccatctatggaaaagatcccgttatcgaccatggcgtcgctc
tgtcagtcgatgatgtcccatgggccaagcagcaaatgtgggctcctgacgcagcttacaagaacggcaa
atattatctctacttccccgccaaggataaagatgagatcttcagaattggagttgctgtctccaacaag
cccagcggtcctttcaaggccgacaagagctggatccccggtacttacagtatcgatcctgctagctatg
tcgacactaatggcgaggcatacctcatctggggcggtatctggggcggccagcttcaggcctggcagga
tcacaagacctttaatgagtcgtggctcggcgacaaagctgctcccaacggcaccaacgccctatctcct
cagatcgccaagctaagcaaggacatgcacaagatcaccgagacaccccgcgatctcgtcatcctggccc
ccgagacaggcaagcccctcaagcagaggacaataagcgacgatttttcgaggggccctgggttcacaa
gcgcggcaagctgtactacctcatgtactctaccggcgacacgcacttcctcgtctacgcgacttccaag
aacatctacggtccttatacctatcagggcaagattctcgaccctgttgatgggtggactacgcatggaa
gtattgttgagtacaagggacagtggtggttgttctttgcggatgcgcatacttctggaaaggattatct
gagacaggttaaggcgaggaagatctggtatgacaaggatggcaagattttgcttactcgtcctaagatt
tag

FIG. 14A

SEQ ID NO:28
Protein sequence of Fv43D

mqlkflssalllsltqncaaqdtndipplitdlwsadpsahvfegklwvypshdieanvvngtggaqyam
rdyhtysmktiygkdpvidhgvalsvddvpwakqqm**wapdaaykngkyylyfpakdkdeifrigvavsnk**
psgpfkadkswipgtysidpasyvdtngeayliwggiwggqlqawqdhktfneswlgdkaapngtnalsp
qiaklskdmhkitetprdlvilapetgkplqaednkrrffegpwvhkrgklyylmystgdthflvyatsk
niygpytyqgkildpvdgwtthgsiveykgqwwlffadahtsgkdylrqvkarkiwydkdgkilltrpki

FIG. 14B

SEQ ID NO:29
Nucleotide sequence for Pf43B, GH43 family enzyme from *Penicillium funiculosum*

atgagtcgcagcatccttccgtacgcctctgttttcgccctcctgggcggggctatcgccgaaccgtttt
tggttctcaatagcgattttcccgatcccagtctcatagagacatccagcggatactatgcattcggtac
caccggaaacggagtcaatgcgcaggttgcttcttcaccagactttaatacctggactttgctttccggc
acagatgccctcccgggaccatttccgtcatgggtagcttcgtctccacaaatctgggcgccagatgttt
tggttaaggtatgttcttatggaataacagttttaggagtaggtcagccaggatattgacaaaattataa
taggccgatggtacctatgtcatgtacttttcggcatctgctgcgagtgactcgggcaaacactgcgttg
gtgccgcaactgcgacctcaccggaaggaccttacaccccggtcgatagcgctgttgcctgtccattaga
ccagggaggagctattgatgccaatggatttattgacaccgacggcactatatacgttgtatacaaaatt
gatggaaacagtctagacggtgatggaaccacacatcctacccccatcatgcttcaacaaatggaggcag
acggaacaaccccaaccggcagcccaatccaactcattgaccgatccgacctcgacggacctttgatcga
ggctcctagtttgctcctctccaatggaatctactacctcagtttctcttccaactactacaacactaat
tactacgacacttcatacgcctatgcctcgtcgattactggtccttggaccaaacaatctgcgccttatg
caccttgttggttactggaaccgagactagcaatgacggcgcattgagcgcccctggtggtgccgattt
ctccgtcgatggcaccaagatgttgttccacgcaaacctcaatggacaagatatctcgggcggacgcgcc
ttatttgctgcgtcaattactgaggccagcgatgtggttacattgcagtag

FIG. 15A

SEQ ID NO:30
Protein sequence of Pf43B

msrsilpyasvfalIqgaiaepflvlnsdfpdpslietssgyyafgttgngvnaqvasspdfntwtllsg
tdalpgpfpswvasspqiwapdvlvkadgtyvmyfsasaasdsgkhcvgaatatspegpytpvdsavacp
ldqggaidangfidtdgtiyvvykidgnsldgdgtthptpimlqqmeadgttptgspiqlidrsdldgpl
ieapslllsngiyylsfssnyyntnyydtsyayassitgpwtkqsapyapllvtgtetsndgalsapgga
dfsvdgtkmlfhanlngqdisggralfaasiteasdvvtlq

FIG. 15B

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

atggttcgcttcagttcaatcctagcggctgcggcttgcttcgtggctgttgagtcagtcaacatcaaggtcgacag
caagggcggaaacgctactagcggtcaccaatatggcttccttcacgaggttggtattgacacaccactggcgatga
ttgggatgctaacttggagctaggatatcaacaattccggtgatggtggcatctacgctgagctcatccgcaatcgt
gctttccagtacagcaagaaataccctgtttctctatctggctggagacccatcaacgatgctaagctctccctcaa
ccgtctcgacactcctctctccgacgctctccccgtttccatgaacgtgaagcctggaaagggcaaggccaaggaga
ttggtttcctcaacgagggttactggggaatggatgtcaagaagcaaaagtacactggctctttctgggttaagggc
gcttacaagggccactttacagcttctttgcgatctaaccttaccgacgatgtctttggcagcgtcaaggtcaagtc
caaggccaacaagaagcagtgggttgagcatgagtttgtgcttactcctaacaagaatgcccctaacagcaacaaca
cttttgctatcacctacgatcccaaggtgagtaacaatcaaaactgggacgtgatgtatactgacaatttgtagggc
gctgatggagctcttgacttcaacctcattagcttgttccctcccacctacaagggccgcaagaacggtcttcgagt
tgatcttgccgaggctctcgaaggtctccacccccgtaaggtttaccgtctcacgtgtatcgtgaacagtcgctgact
tgtagaaaagagcctgctgcgcttccccggtggtaacatgctcgagggcaacaccaacaagacctggtgggactgga
aggataccctcggacctctccgcaaccgtcctggtttcgagggtgtctggaactaccagcagacccatggtcttgga
atcttggagtacctccagtgggctgaggacatgaaccttgaaatcagtaggttctataaaattcagtgacggttatg
tgcatgctaacagatttcagttgtcggtgtctacgctggcctctccctcgacggctccgtcacccccaaggaccaac
tccagcccctcatcgacgacgcgctcgacgagatcgaattcatccgaggtcccgtcacttcaaagtggggaaagaag
cgcgctgagctcggccaccccaagcctttcagactctcctacgttgaagtcggaaacgaggactggctcgctggtta
tcccactggctggaactcttacaaggagtaccgcttccccatgttcctcgaggctatcaagaaagctcaccccgatc
tcaccgtcatctcctctggtgcttctattgaccccgttggtaagaaggatgctggtttcgatattcctgctcctgga
atcggtgactaccaccctaccgcgagcctgatgttcttgttgaggagttcaacctgtttgataacaataagtatgg
tcacatcattggtgaggttgcttctacccaccccaacggtggaactggctggagtggtaaccttatgccttacccct
ggtggatctctggtgttggcgaggccgtcgctctctgcggttatgagcgcaacgccgatcgtattcccggaacattc
tacgctcctatcctcaagaacgagaaccgttggcagtgggctatcaccatgatccaattcgccgccgactccgccat
gaccacccgctccaccagctggtatgtctggtcactcttcgcaggccaccccatgacccatactctccccaccaccg
ccgacttcgaccccctctactacgtcgctggtaagaacgaggacaagggaactcttatctggaagggtgctgcgtat
aacaccaccaagggtgctgacgttcccgtgtctctgtccttcaagggtgtcaagcccggtgctcaagctgagcttac
tcttctgaccaacaaggagaaggatccttttgcgttcaatgatcctcacaagggcaacaatgttgttgatactaaga
agactgttctcaaggccgatggaaagggtgctttcaacttcaagcttcctaacctgagcgtcgctgttcttgagacc
ctcaagaagggaaagccttactctagctag

FIG. 16A

SUBSTITUTE SHEET (RULE 26)

SEQ ID NO:32
Protein sequence of Fv51A

mvrfssilaaaacfvavesvnikvdskggnatsghqygflhedinnsgdggiyaelirnrafqyskkypv
slsgwrpindaklslnrldtplsdalpvsmnvkpgkgkakeigflnegywgmdvkkqkytgsfwvkgayk
ghftaslrsnltddvfgsvkvkskankkqwvehefvltpnknapnsnntfaitydpkgadgaldfnlisl
fpptykgrknglrvdlaealeglhpsllrfpggnmlegntnktwwdwkdtlgplrnrpgfegvwnyqqth
glgileylqwaedmnleiivgvyaglsldgsvtpkdqlqpliddaldeiefirgpvtskwgkkraelghp
kpfrlsyvevgnedwlagyptgwnsykeyrfpmfleaikkahpdltvissgasidpvgkkdagfdipapg
igdyhpyrepdvlveefnlfdnnkyghii**gevasthpnggtgwsgnlmpypwwisgvgeavalcgyerna**
dripgtfyapilknenrwqwaitmiqfaadsamttrstswyvwslfaghpmthtlpttadfdplyyvagk
nedkgtliwkgaaynttkgadvpvslsfkgvkpgaqaeltlltnkekdpfafndphkgnnvvdtkktvlk
adgkgafnfklpnlsvavletlkkgkpyss

FIG. 16B

SEQ ID NO:33
Nucleotide sequence for Cg51B, a GH51 family enzyme from *Chaetomium globosum*

atggcgcccctttcgcttcgggccctctcgctgctcgcgctcacaggagccgcagccgcggtgaccctat
cggtcgcgaactctggcggtaatgatacgtctccgtacatgtatggcatcatgttcgaggacatcaatca
gagcggtgacggcgggctgtaagttctgtcgcggcttcccctgacaagcttgcatgatgcttaactaaag
tccttaggtacgccgagctgattcgcaaccgagccttccataatagctccctccaggcctggaccgccgt
ggggacagcactctcgaggtcgtaacctctgcaccgttatcggatgccctgcctcgctcggtcaaggtc
acgagtggaaagggcaaggcgggcttgaagaatgccggctactggggaatggacgtccagaagaccgaca
agtatagcggcagcttctactcgtacggcgcctacgacggaaagtttacccctctctctggtgtcggacat
cacaaatgagaccctggccaccaccaagatcaagtccaggtcggtggagcatgcctggaccgagcacaag
ttcgagcttctcccgaccaagagcgcggcgaacagcaacaacagcttcgtgctggagttccgcccctgcc
accagacggagctccagttcaacctcatcagcttgttcccgccgacgtataagaacaggcccaacggcat
gcgccgagagctcatggagaagctcgcagacctcaagcccagtttccttcggattccaggaggcaacaac
ctgtaagtgcttccggcgaaactagcagtagttgcctgagagacactaatctcagcgaacaacagcgagg
gcaactatgctggcaactactggaactggtcaagcacacttggcccgctgaccgaccggcccggtcgtga
cggcgtgtggacgtacgccaacacggacggcatcgggctggtcgagtacatgcactgggccgaggacctc
gacgtggaggttgtgctggcggtcgccgcaggcctgtacctgaacggcgatgtggtcccggaggaggagc
tgcacgtcttcgtggaggatgcgctgaacgagctcgagttcctcatgggcgacgtctcgacccccttgggg
cgcgcgccgcgctaagctcggctaccccaagccgtggaacatcaagttcgtcgaggtcggcaacgaggac
aacctgtgggcggcctcgactcgtacaagagctaccggctgaagactttctacgacgccatcaaggcga
agtaccccgacatctccatcttttcgtcgaccgacgagtttgtgtacaaggagtcgggccaggactacca
caagtacaccccggccggactactccgtgtcccagttcgacctgtttgacaactgggccgacggccaccc
atcatcatcggagagtgagtgaacggcgacccccacctcccctaacgcgggatcgcgagctgatagatc
accccaggtatgcgaccatccagaacaacacgggcaagctcgaggacacggactgggacgcgcccaagaa
caagtggtccaactggatcggctccgtcgccgaggccgtcttcatcctcggagccgagcgcaacggcgac
cgggtctgggccaccacctttgcgccgatcctccagaacctcaacagctaccaatgggctgtaagtacat
acatacataccgcacccccaaccccaaccccccccaaagcgcacctccacccacccacccaaacacaccac
aactacctagctaacccgccacacaaacaaacagcccgacctaatctccttcaccgccaacccggccgac
accacgcccagcgtctcgtacccgatcatccagctgctcgcctcgcaccgcatcacgcacaccctccccg
tcagcagcgccgacgccttcggcccggcctactgggtggccggtcgcggcgccgacgacggctcgtacat
cctcaaggcggccgtgtacaacagcacgggggtgcggatgtaccggtgagggtgcagtttgaggcgggg
ggtggtggtggtggtggtggtggtggtggtggtggtgatgggaaggggaagggtaaagggaagg
gaggggagggtggtgagggtgtgaagaagggtgaccgcgcgcagttgaccgtgttgacggcgccggaggg
gccctgggcgcataatacgccggagaataaggggcggtcaagacgacagtgacgacgttgaaggccggg
aggggtggggtgtttgagtttagtctgccggatttgtcggtggcggtgttggtggtggaggggagaagt
ga

FIG. 17A

SEQ ID NO:34
Protein sequence of Cg51B

maplslralsllaltqaaaavtlsvansggndtspymygimfedinqsgdgglyaelirnrafhnsslqa
wtavgdstlevvtsaplsdalprsvkvtsgkgkaglknagywgmdvqktdkysgsfysygaydgkftlsl
vsditnetlattkiksrsvehawtehkfellptksaansnnsfvlefrpchqtelqfnlislfpptyknr
pngmrrelmekladlkpsflripggnnlegnyagnywnwsstlgpltdrpgrdgvwtyantdgiglveym
hwaedldvevvlavaaglylngdvvpeeelhvfvedalneleflmgdvstpwgarraklgypkpwnikfv
evgnednlwggldsyksyrlktfydaikakypdisifsstdefvykesgqdyhkytrpdysvsqfdlfdn
wadghpiii**geyatiqnntgkledtdwdapknkwsnwigsvaeavfilgaerngdrvwgttfapilqnln**
syqwapdlisftanpadttpsvsypiiqllashrithtlpvssadafgpaywvagrgaddgsyilkaavy
nstggadvpvrvqfeaggggggggggggggdgkgkgkgkggeggegvkkgdraqltvltapegpwahnt
penkgavkttvttlkagrggvfefslpdlsvavlvvegek

FIG. 17B

SEQ ID NO:35
Nucleotide sequence for Fv43C, a GH43 family enzyme from *Fusarium verticillioides*

atgcgtcttctatcgtttcccagccatctcctcgtggccttcctaaccctcaaagaggcttcatccctcg
ccctcagcaaacgggatagccctgtcctccccggcctctgggcggaccccaacatcgccatcgtcgacaa
gacatactacatcttccctaccaccgacggtttcgaaggctggggcggcaacgtcttctactggtggaaa
tcaaaagatctcgtatcatggacaaagagcgacaagccattccttactctcaatggtacgaatggcaacg
ttccctgggctacaggtaatgcctgggctcctgctttcgctgctcgcggaggcaagtattacttctacca
tagtgggaataatccctctgtgagtgatgggcataagagtattggtgcggcggtggctgatcatcctgag
gggccgtggaaggcacaggataagccgatgatcaagggaacttctgatgaggagattgtcagcaaccagg
ctatcgatcccgctgcctttgaagaccctgagactggaaagtggtatatctactgggggaaacggtgtccc
cattgtcgcagagctcaacgacgacatggtctctctcaaagcaggctggcacaaaatcacaggtcttcag
aatttccgcgagggtcttttcgtcaactatcgcgatggaacatatcatctgacatactctatcgacgata
cgggctcagagaactatcgcgttgggtacgctacggcggataaccccattggaccttggacatatcgtgg
tgttcttctggagaaggacgaatcgaagggcattcttgctacgggacataactccatcatcaacattcct
ggaacggatgagtggtatatcgcgtatcatcgcttccatattcccgatggaaatgggtataatagggaga
ctacgattgatagggtacccatcgacaaggatacgggtttgtttggaaaggttacgccgactttgcagag
tgttgatcctaggcctttgtag

FIG. 18A

SEQ ID NO:36
Protein sequence for Fv43C

mrllsfpshllvafltlkeasslalsk**rdspvlpglwadpniaivdktyyifpttdgfegwggnvfywwk**
skdlvswtksdkpfltlngtngnvpwatgnawapafaarggkyyfyhsgnnpsvsdghksigaavadhpe
gpwkaqdkpmikgtsdeeivsnqaidpaafedpetgkwyiywgngvpivaelnddmvslkagwhkitglq
nfreglfvnyrdgtyhltysiddtgsenyrvgyatadnpigpwtyrgvllekdeskgilatghnsiinip
gtdewyiayhrfhipdgngynrettidrvpidkdtglfgkvtptlqsvdprpl

FIG. 18B

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

atgctcttctcgctcgttcttcctaccttgcctttcaagccagcctggcgctcggcgatacatccgtta
ctgtcgacaccagccagaaactccaggtcatcgatggctttggtgtctcagaagcctacggccacgccaa
acaattccaaaacctcggtcctggaccacagaaagagggcctcgatcttctcttcaacactacaaccggc
gcaggcttatccatcatccgaaacaagatcggctgcgacgcctccaactccatcaccagcaccaacaccg
acaacccagataagcaggctgtttaccattttgacggcgatgatgatggtcaggtatggtttagcaaaca
ggccatgagctatggtgtagatactatctacgctaatgcttggtctgcgcctgtatacatgaagtcagcc
cagagtatgggccgtctctgcggtacacctggtgtgtcgtgctcctctggagattggagacatcgttacg
ttgagatgatagctgagtacctctcctactacaagcaggctggcatcccagtgtcgcacgttggattcct
caatgagggtgacggctcggactttatgctctcaactgccgaacaggctgcagatgtcattcctcttcta
cacagcgctttgcagtccaagggccttggcgatatcaagatgacgtgctgtgataacatcggttggaagt
cacagatggactataccgccaagctggctgagcttgaggtggagaagtatctatctgtcatcacatccca
cgagtactccagcagccccaaccagcctatgaacactacattgccaacctggatgtccgagggagctgcc
aatgaccaggcatttgccacagcgtggtacgtcaacggcggttccaacgaaggtttcacatgggcagtca
agatcgcacaaggcatcgtcaatgccgacctctcagcgtatatctactgggagggcgttgagaccaacaa
caaggggtctctatctcacgtcatcgacacggacggtaccaagtttaccatatcctcgattctctgggcc
attgctcactggtcgcgccatattcgccctggtgcgcatagactttcgacttcaggtgttgtgcaagata
cgattgttggtgcgtttgagaacgttgatggcagtgtcgtcatggtgctcaccaactctggcactgctgc
tcagactgtggacctgggtgtttcgggaagtagcttctcaacagctcaggctttcacttcggatgctgag
gcgcagatggtcgataccaaggtgactctgtccgacggtcgtgtcaaggttacggtcccggtgcacggtg
tcgtcactgtgaagctcacaacagcaaaaagctccaaaccggtctcaactgctgtttctgcgcaatctgc
ccccactccaactagtgttaagcacaccttgactcaccagaagacttcttcaacaacactctcgaccgcc
aaggccccaacctccactcagactacctctgtagttgagtcagccaaggcggtgaaataccctgtccccc
ctgtagcatccaagggatcctcgaagagtgctcccaagaagggtaccaagaagaccactacgaagaaggg
ctcccaccaatcgcacaaggcgcatagtgctactcatcgtcgatgccgccatggaagttaccgtcgtggc
cactgcaccaactaa

FIG. 19A

SEQ ID NO:38
Protein sequence of Fv30A

mlfslvlptlafqaslalgdtsvtvdtsqklqvidgfgvseayghakqfqnlgpgpqkegldllfntttg
aglsiirnkigcdasnsitstntdnpdkqavyhfdgdddgqsaqsmgrlcgtpgvscssgdwrhryvemi
aeylsyykqagipvshvgflnegdgsdfmlstaeqaadvipllhsalqskglgdikmtccdnigwksqmd
ytaklaelevekylsvitsheyssspnqpmnttlptwmsegaandqafatawyvnggsnegftwavkiaq
givnadlsayiywegvetnnkgslshvidtdgtkftissilwaiahwsrhirpgahrlstsgvvqdtivg
afenvdgsvvmvltnsgtaaqtvdlgvsgssfstaqaftsdaeaqmvdtkvtlsdgrvkvtvpvhgvvtv
klttaksskpvstavsaqsaptptsvkhtlthqktssttlstakaptstqttsvvesakavkypvppvas
kgssksapkkgtkktttkkgshqshkahsathrrcrhgsyrrghctn

FIG. 19B

SEQ ID NO:39
Nucleotide sequence for Fv43F, a GH43 family enzyme from *Fusarium verticillioides*

atgtggaaactcctcgtcagcggtcttgtcgccgtcgcgtccctcagcggcgtgaacgctgcttatccta
accctggtcccgtcaccggcgatactcgtgttcacgaccctacggttgtcaagactcccagcggtggata
cttgctggctcatactggcgataacgtttcgctcaagacttcttctgatcgaactgcttggaaggatgca
ggtgctgttttccccaacggtgcgccttggactacgcagtacaccaagggcgacaagaacctctgggccc
ctgatatctcctaccacaacggccagtactatctgtactactccgcctcttccttcggtcagcgtacctc
tgccatttttctcgctaccagcaagaccggtgcatccggctcgtggaccaaccaaggcgtcgtcgtcgag
tccaacaacaacaacgactacaatgccattgacggaaatctctttgtcgactctgatggaaaatggtggc
tctccttcggctctttctggtccggcatcaagctcatccaactcgaccccaagaccggcaagcgcaccgg
ctcaagcatgtactccctcgccaaacgcgacgcctccgtcgaaggcgccgtcgaggctccgttcatcacc
aaacgcggaagcacctactacctctgggtgtcgttcgacaagtgttgccagggcgctgctagcacgtacc
gtgtcatggttggacggtcgagcagcattactggtccttatgttgacaaggctggtaagcagatgatgtc
tggtggaggaacggagattatggctagtcacggatctattcatggaccgggacataatgctgttttcact
gataacgatgcggacgttcttgtctatcattactacgataacgctggcacagcgctgttgggcatcaact
tgctcagatatgacaatggctggcctgttgcttattag

FIG. 20A

SEQ ID NO:40
Protein sequence for Fv43F

mwkllvsglvavaslsgvnaaypnpgpvtgdtrvhdptvvktpsggyllahtgdnvslktssdrtawkda
gavfpngapwttqytkgdknlwapdisyhngqyylyysassfgqrtsaiflatsktgasgswtnqgvvve
snnnndynaidgnlfvdsdgkwwlsfgsfwsgikliqldpktgkrtgssmyslakrdasvegaveapfit
krgstyylwvsfdkccqgaastyrvmvgrsssitgpyvdkagkqmmsgggteimashgsihgpghnavft
dndadvlvyhyydnagtallginllrydngwpvay

FIG. 20B

SEQ ID NO:41
Nucleotide sequence for Xyn3, a GH10 xylanase from *Trichoderma reesei*

atgaaagcaaacgtcatcttgtgcctcctggccccctggtcgccgctctccccaccgaaaccatccacc
tcgaccccgagctcgccgctctccgcgccaacctcaccgagcgaacagccgacctctgggaccgccaagc
ctctcaaagcatcgaccagctcatcaagagaaaaggcaagctctactttggcaccgccaccgaccgcggc
ctcctccaacgggaaaagaacgcggccatcatccaggcagacctcggccaggtgacgccggagaacagca
tgaagtggcagtcgctcgagaacaaccaaggccagctgaactggggagacgccgactatctcgtcaactt
tgcccagcaaaacggcaagtcgatacgcggccacactctgatctggcactcgcagctgcctgcgtgggtg
aacaatatcaacaacgcggatactctgcggcaagtcatccgcacccatgtctctactgtggttgggcggt
acaagggcaagattcgtgcttgggtgagttttgaacaccacatgccccttttcttagtccgctcctcctc
ctcttggaacttctcacagttatagccgtatacaacattcgacaggaaatttaggatgacaactactgac
tgacttgtgtgtgtgatggcgataggacgtggtcaatgaaatcttcaacgaggatggaacgctgcgctct
tcagtcttttccaggctcctcggcgaggagtttgtctcgattgccttcgtgctgctcgagatgctgacc
cttctgcccgtctttacatcaacgactacaatctcgaccgcgccaactatggcaaggtcaacgggttgaa
gacttacgtctccaagtggatctctcaaggagttcccattgacggtattggtgagccacgacccctaaat
gtccccattagagtctctttctagagccaaggcttgaagccattcagggactgacacgagagccttctc
tacaggaagccagtcccatctcagcggcggcggaggctctggtacgctgggtgcgctccagcagctggca
acggtacccgtcaccgagctggccattaccgagctggacattcaggggcaccgacgacggattacaccc
aagttgttcaagcatgcctgagcgtctccaagtgcgtcggcatcaccgtgtggggcatcagtgacaaggt
aagttgcttccctgtctgtgcttatcaactgtaagcagcaacaactgatgctgtctgtctttacctagg
actcgtggcgtgccagcaccaaccctcttctgtttgacgcaaacttcaaccccaagccggcatataacag
cattgttggcatcttacaatag

FIG. 21A

SEQ ID NO:42
Protein sequence for Xyn3

mkanvilcllaplvaaiptetihldpelaalranltertadlwdrqasq**sidqlikrkgklyfgtatdrg**
llqreknaaiiqadlgqvtpensmkwqslennqgqlnwgdadylvnfaqqngksirghtliwhsqlpawv
nninnadtlrqvirthvstvvgrykgkirawdvvneifnedgtlrssvfsrllgeefvsiafraardadp
sarlyindynldranygkvnglktyvskwisqgvpidgigsqshlsggggsgtlgalqqlatvpvtelai
teldiqgapttdytqvvqaclsvskcvgitvwgisdkdswrastnpllfdanfnpkpaynsivgilq

FIG. 21B

SEQ ID NO:43
Protein sequence of Xyn2, a GH11 family xylanase from *Trichoderma reesei*

mvsftsllaasppsrascrpaaevesvavekrqtiqpgt**gynngyfysywndghggvtytngpggqfsvn wsnsgnfvggkgwqpgtknkvinfsgsynpngnsylsvygwsrnplieyyivenfgtynpstgatklgev tsdgsvydiyrtqrvnqpsiigtatfyqywsvrrnhrssgsvntanhfnawaqqgltlgtmdyqivaveg yfssgsasitvs**

FIG. 22

SEQ ID NO:44
Protein sequence of Bxl1, a GH3 family β-xylosidase from *Trichoderma reesei*

mvnnaallaalsallptalaqnnqtyanysaqgqpdlypetlatltlsfpdcehgplknnlvcdssagyv eraqalislftleelilntqn**sgpgvprlglpnyqvwnealhgldranfatkggqfewatsfpmpiltta alnrtlihqiadiistqarafsnsgrygldvyapnvngfrsplwgrgqetpgedafflssaytyeyitgi qqggvdpehlkvaatvkhfagydlenwnnqsrlgfdaiitqqdlseyytpqflaaaryaksrslmcaynsv ngvpscansfflqtllreswgfpewgyvssdcdavynvfnphdyasnqssaaasslragtd**idcgqtypw hlnesfvagevsrgeiersvtrlyanlvrlgyfdkknqyrslgwkdvvktdawnisyeaavegivllknd **gtlplskkvrsialigpwanattqmqgnyygpapylispleaakkagyhvnfelgteiagnsttgfakai aaakksdaiiylggidntieqegadrtdiawpgnqldlikqlsevgkplvvlqmgggqvdssslksnkkv nslvwggypgqsggvalfdilsgkrapagrlvttqypaeyvhqfpqndmnlrpdgksnpgqtyiwytgkp vyefgsglfyttfketlashpkslkfntssilsaphpgytyseqipvftfeaniknsgktespytamlfv** rtsnagpapypnkwlvgfdrladikpghssklsipipvsalarvdshgnrivypgkyelalntdesvkle felvgeevtienwpleeqqikdatpda

FIG. 23

SEQ ID NO:45
Protein sequence of Bgl1, a GH3 family β-glucosidase from *Trichoderma reesei*

mryrtaaalalatqpfaradshstsgasaeavvppagtpwgtaydkakaalaklnlqdkvgivsgv**gwng gpcvgntspaskisypslclqdgplgvrystgstaftpgvqaastwdvnlirergqfigeevkasgihvi lgpvagplgktpqggrnwegfgvdpyltgiamgqtingiqsvgvqatakhyilneqelnretissnpddr tlhelytwpfadavqanvasvmcsynkvnttwacedqytlqtvlkdqlgfpgyvmtdwnaqhttvqsans gldmsmpgtdfngnnrlwgpaltnavnsnqvptsrvddmvtrilaawyltgqdqagypsfnisrnvqgnh** ktnvraiardgivllkn**danilplkkpasiavvgsaaiignharnspscndkgcddgalgmgwgsgavny pyfvapydaintrassqgtqvtlsntdntssgasaargkdvaivfitadsgegyitvegnagdrnnldpw hngnalvqavagansnvivvvhsvgaiileqilalpqvkavvwaglpsqesgnalvdvlwgdvspsgklv ytiakspndyntrivsqgsdsfseglfidykhfddanitpryefgyglsyt**kfnysrlsvlstaksgpat gavvpggpsdlfqnvatvtvdiansgqvtgaevaqlyityppssaprtppkqlrgfaklnltpgqsgtatf nirrrdlsywdtasqkwvvpsgsfgisvgassrdirltstlsva

FIG. 24

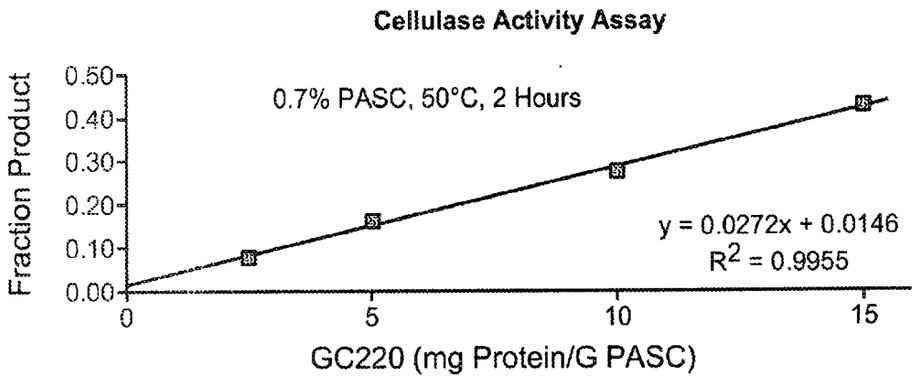

Cellulase Activity Assay
0.7% PASC, 50°C, 2 Hours
y = 0.0272x + 0.0146
R2 = 0.9955
Fraction Product
0.50
0.40
0.30
0.20
0.10
0.00
0
5
10
15
GC220 (mg Protein/G PASC)

FIG. 25

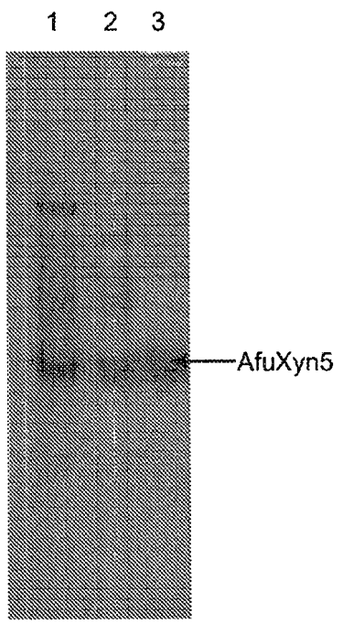

1 2 3
AfuXyn5

FIG. 27

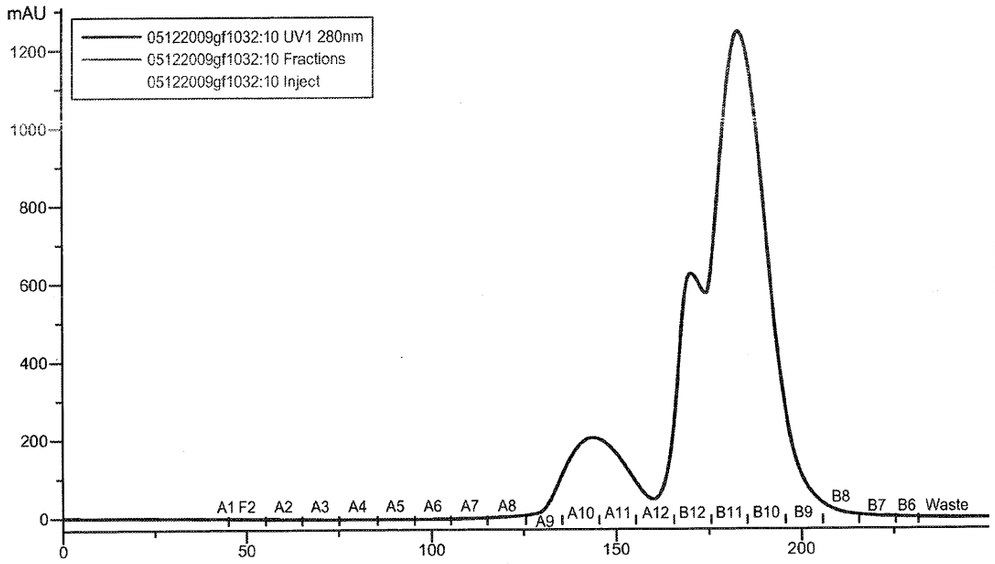


FIG. 26

SUBSTITUTE SHEET (RULE 26)

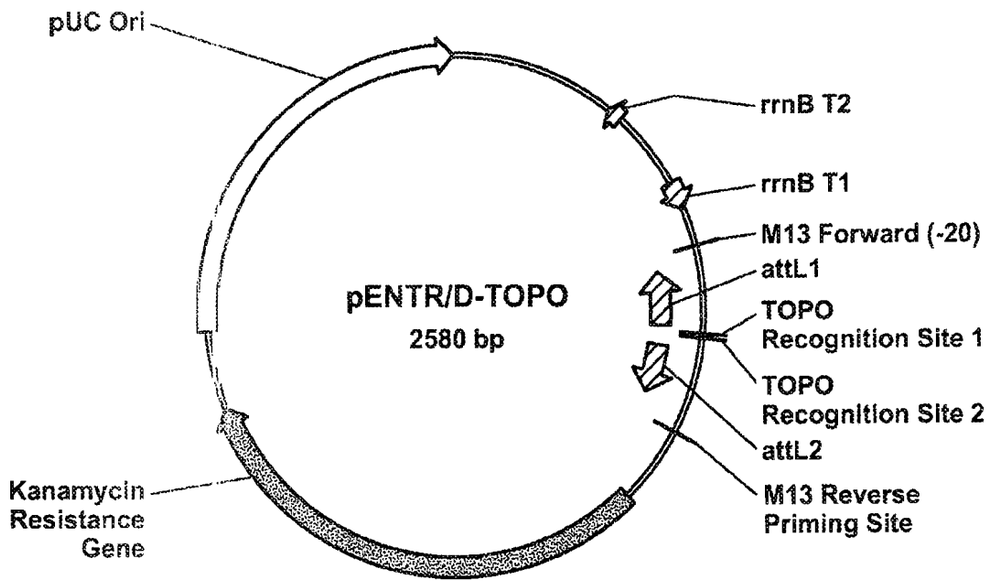

pUC Ori
rrnB T2
rrnB T1
M13 Forward (-20)
attL1
TOPO
Recognition Site 1
TOPO
Recognition Site 2
attL2
M13 Reverse
Priming Site
Kanamycin
Resistance
Gene
pENTR/D-TOPO
2580 bp

FIG. 28

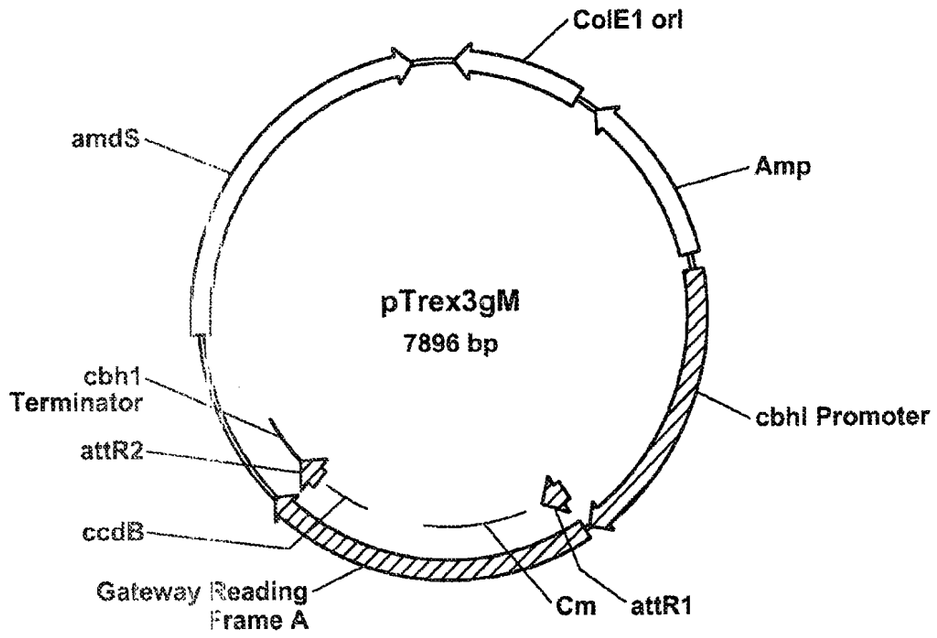
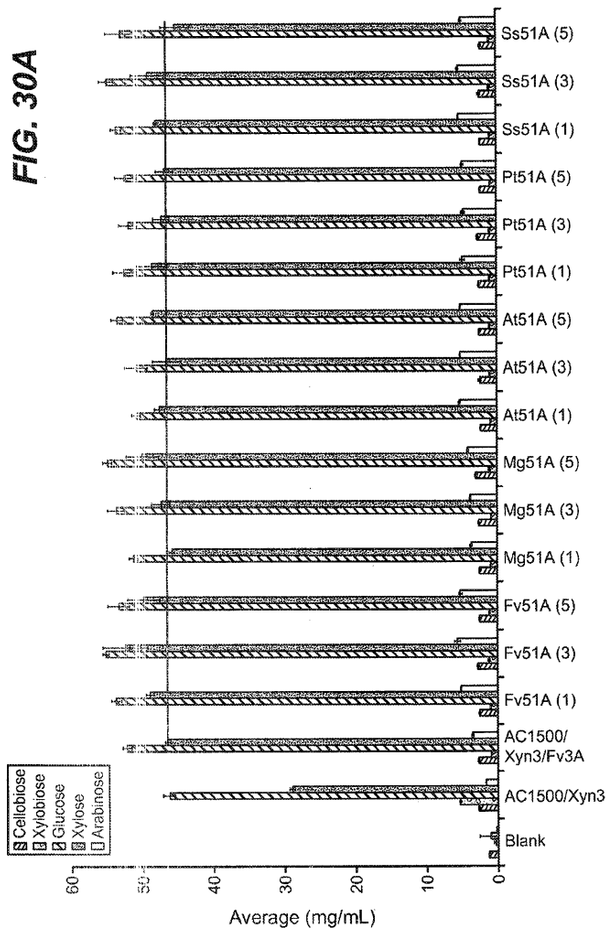

ColE1 ori
amdS
Amp
pTrex3gM
7896 bp
cbh1
Terminator
attR2
cbhI Promoter
ccdB
Gateway Reading
Frame A
Cm
attR1

FIG. 29

FIG. 30A


Ss51A (5)
Ss51A (3)
Ss51A (1)
Pt51A (5)
Pt51A (3)
Pt51A (1)
At51A (5)
At51A (3)
At51A (1)
Mg51A (5)
Mg51A (3)
Mg51A (1)
Fv51A (5)
Fv51A (3)
Fv51A (1)
AC1500/ Xyn3/Fv3A
AC1500/Xyn3
Blank
Cellobiose
Xylobiose
Glucose
Xylose
Arabinose
60
50
40
30
20
10
0
Average (mg/mL)

200

FIG. 30B

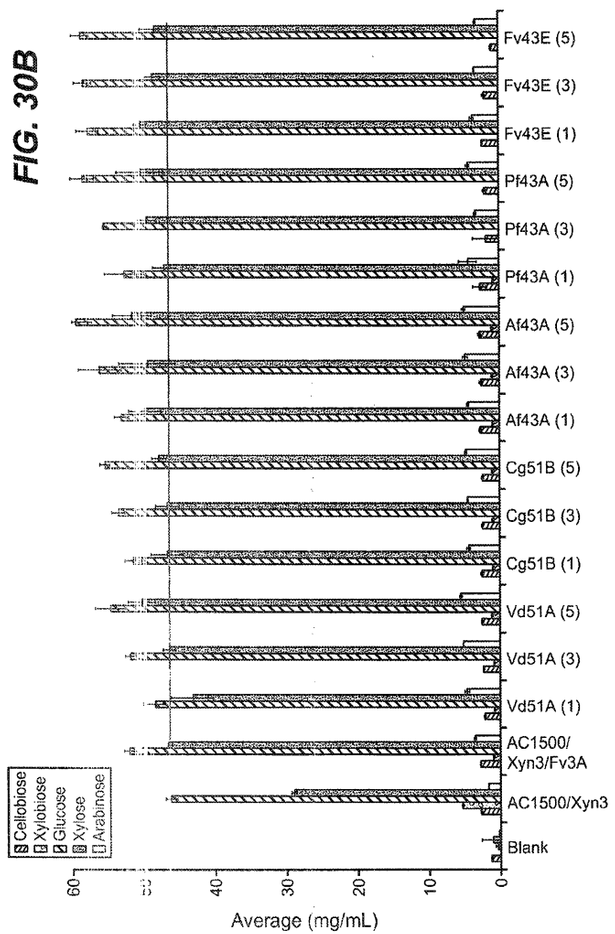

Fv43E (5)
Fv43E (3)
Fv43E (1)
Pf43A (5)
Pf43A (3)
Pf43A (1)
Af43A (5)
Af43A (3)
Af43A (1)
Cg51B (5)
Cg51B (3)
Cg51B (1)
Vd51A (5)
Vd51A (3)
Vd51A (1)
AC1500/ Xyn3/Fv3A
AC1500/Xyn3
Blank
Cellobiose
Xylobiose
Glucose
Xylose
Arabinose
60
50
40
30
20
10
0
Average (mg/mL)

FIG. 31

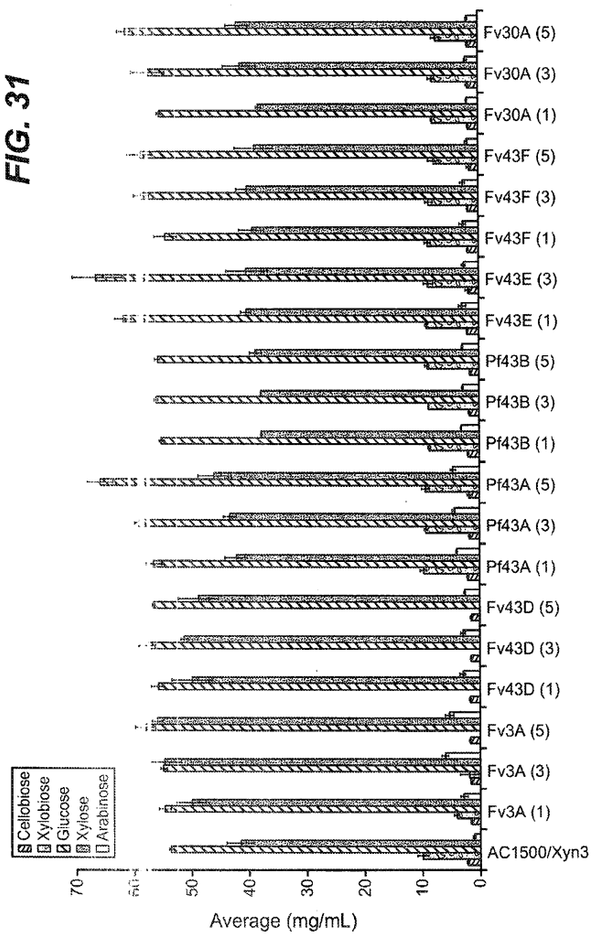

Fv30A (5)
Fv30A (3)
Fv30A (1)
Fv43F (5)
Fv43F (3)
Fv43F (1)
Fv43E (3)
Fv43E (1)
Pf43B (5)
Pf43B (3)
Pf43B (1)
Pf43A (5)
Pf43A (3)
Pf43A (1)
Fv43D (5)
Fv43D (3)
Fv43D (1)
Fv3A (5)
Fv3A (3)
Fv3A (1)
AC1500/Xyn3
Cellobiose
Xylobiose
Glucose
Xylose
Arabinose
70
50
40
30
20
10
0
Average (mg/mL)

FIG. 32

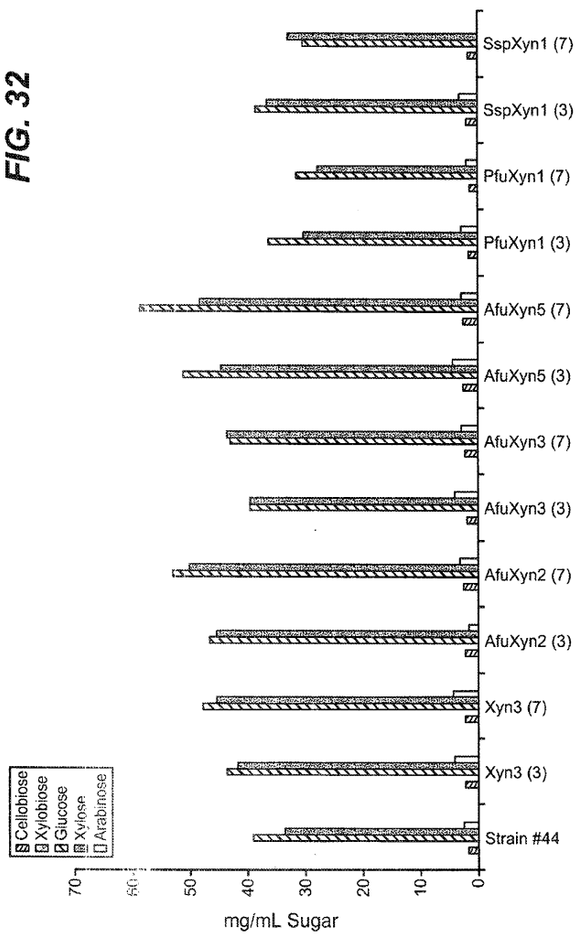

SspXyn1 (7)
SspXyn1 (3)
PfuXyn1 (7)
PfuXyn1 (3)
AfuXyn5 (7)
AfuXyn5 (3)
AfuXyn3 (7)
AfuXyn3 (3)
AfuXyn2 (7)
AfuXyn2 (3)
Xyn3 (7)
Xyn3 (3)
Strain #44
Cellobiose
Xylobiose
Glucose
Xylose
Arabinose
70
50
40
30
20
10
0
mg/mL Sugar

FIG. 33

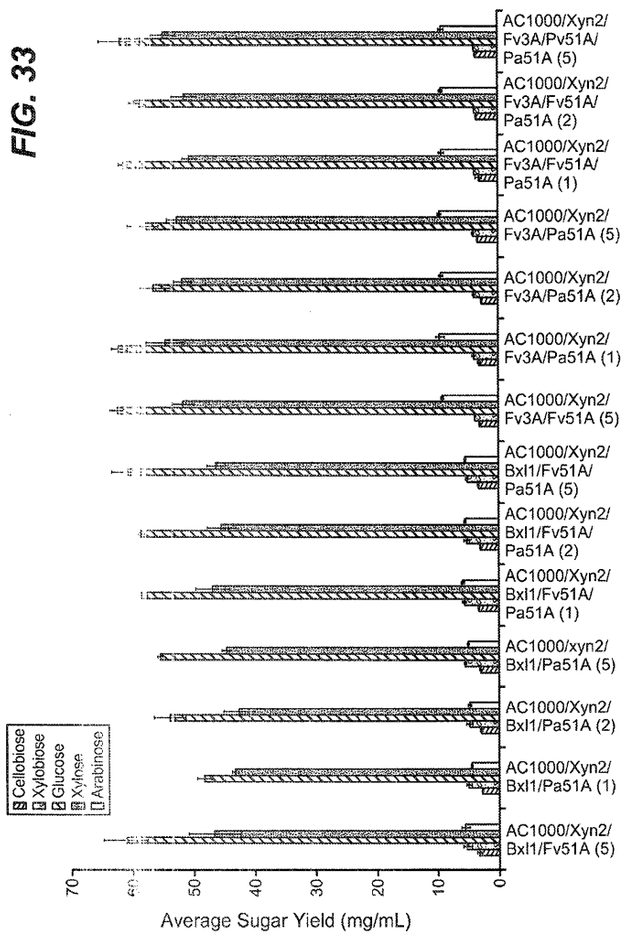
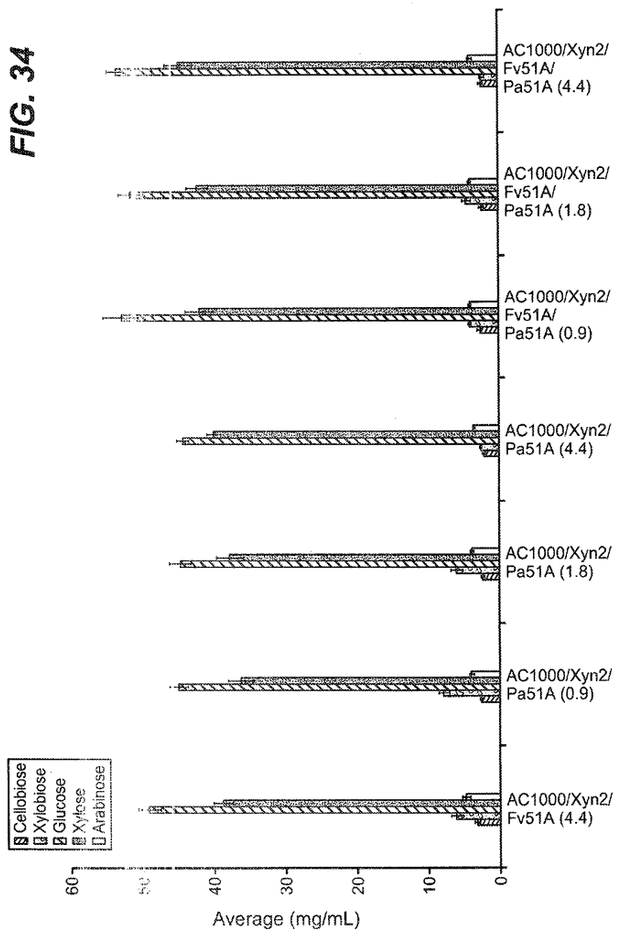

AC1000/Xyn2/ Fv3A/Pv51A/ Pa51A (5)
AC1000/Xyn2/ Fv3A/Fv51A/ Pa51A (2)
AC1000/Xyn2/ Fv3A/Fv51A/ Pa51A (1)
AC1000/Xyn2/ Fv3A/Pa51A (5)
AC1000/Xyn2/ Fv3A/Pa51A (2)
AC1000/Xyn2/ Fv3A/Pa51A (1)
AC1000/Xyn2/ Fv3A/Fv51A (5)
AC1000/Xyn2/ Bxl1/Fv51A/ Pa51A (5)
AC1000/Xyn2/ Bxl1/Fv51A/ Pa51A (2)
AC1000/Xyn2/ Bxl1/Fv51A/ Pa51A (1)
AC1000/xyn2/ Bxl1/Pa51A (5)
AC1000/Xyn2/ Bxl1/Pa51A (2)
AC1000/Xyn2/ Bxl1/Pa51A (1)
AC1000/Xyn2/ Bxl1/Fv51A (5)
Cellobiose
Xylobiose
Glucose
Xylose
Arabinose
70
60
50
40
30
20
10
0
Average Sugar Yield (mg/mL)

FIG. 34


AC1000/Xyn2/
Fv51A/
Pa51A (4.4)
AC1000/Xyn2/
Fv51A/
Pa51A (1.8)
AC1000/Xyn2/
Fv51A/
Pa51A (0.9)
AC1000/Xyn2/
Pa51A (4.4)
AC1000/Xyn2/
Pa51A (1.8)
AC1000/Xyn2/
Pa51A (0.9)
AC1000/Xyn2/
Fv51A (4.4)
Cellobiose
Xylobiose
Glucose
Xylose
Arabinose
60
50
40
30
20
10
0
Average (mg/mL)

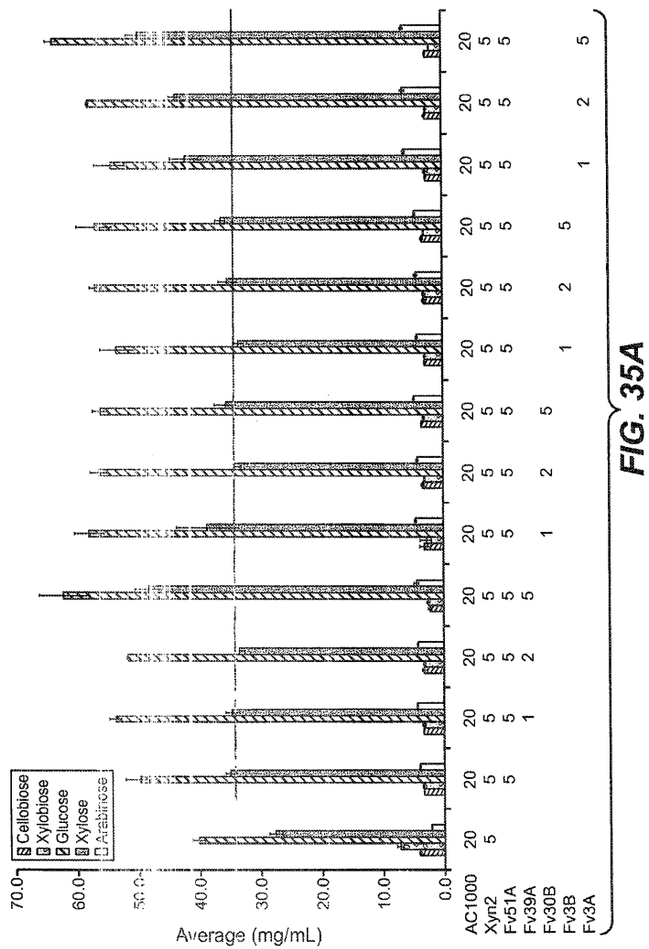

Cellobiose
Xylobiose
Glucose
Xylose
Arabinose
70.0
60.0
40.0
30.0
20.0
10.0
0.0
Average (mg/mL)
AC1000 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20
Xyn2 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
Fv51A 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
Fv39A 1 2 5 1 2 5 1 2 5 1 2 5
Fv30B 1 2 5 1 2 5
Fv3B 1 2 5
Fv3A 1 2 5

FIG. 35A

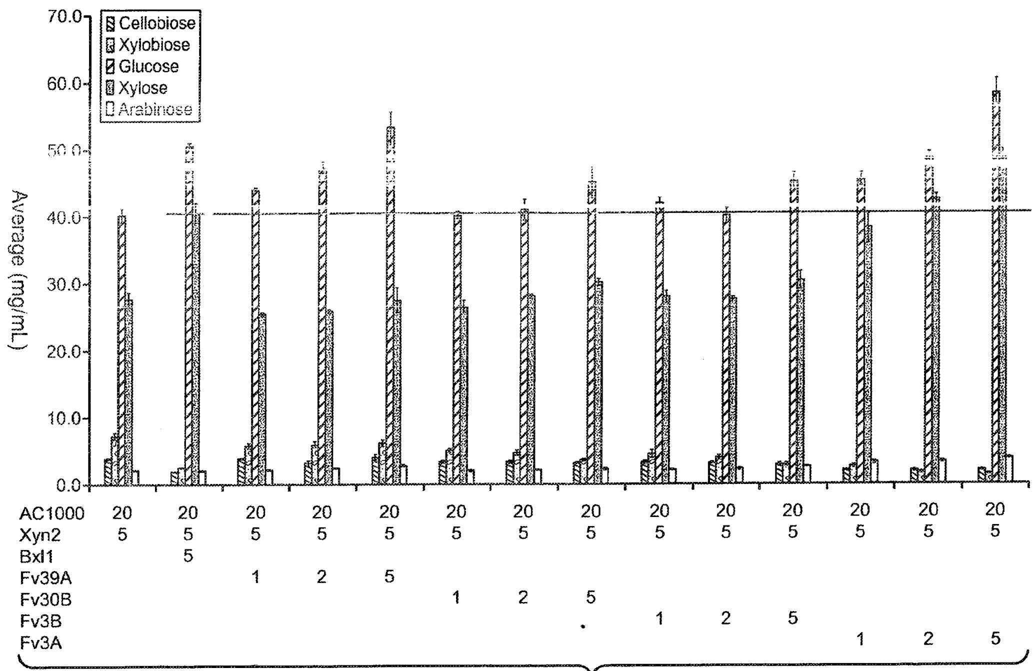

Cellobiose
Xylobiose
Glucose
Xylose
Arabinose
Average (mg/mL)
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
AC1000 20 20 20 20 20 20 20 20 20 20 20 20 20 20
Xyn2 5 5 5 5 5 5 5 5 5 5 5 5 5 5
Bxl1 5
Fv39A 1 2 5
Fv30B 1 2 5
Fv3B 1 2 5
Fv3A 1 2 5

FIG. 35B

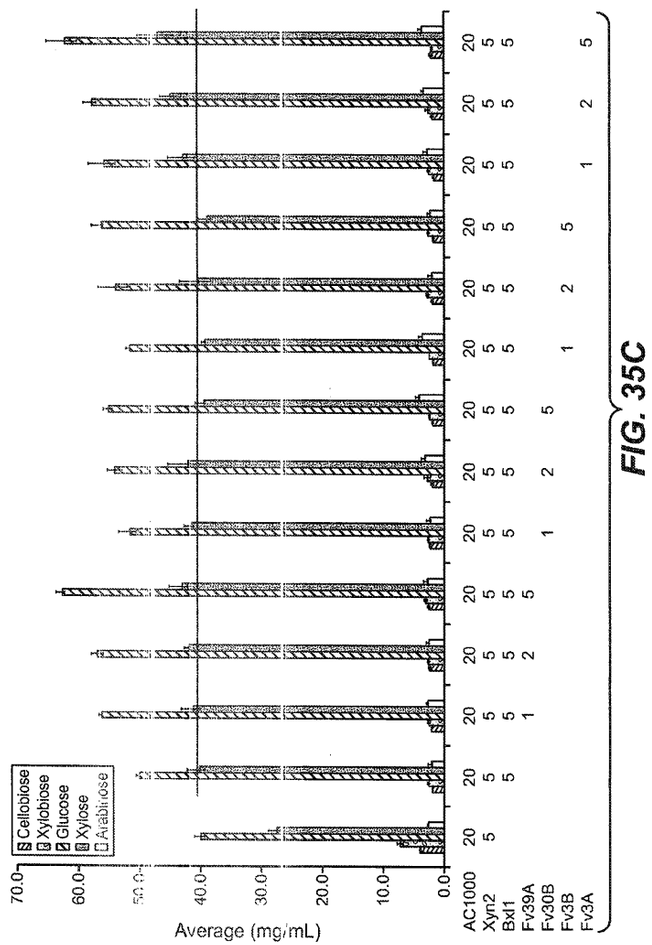

Cellobiose
Xylobiose
Glucose
Xylose
Arabinose
70.0
60.0
40.0
30.0
20.0
10.0
0.0
Average (mg/mL)
AC1000 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20
Xyn2 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
Bxl1 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
Fv39A 1 2 5 1 2 5 1 2 5
Fv30B 1 2 5 1 2 5
Fv3B 1 2 5 1 2 5
Fv3A 1 2 5 1 2 5

FIG. 35C

208

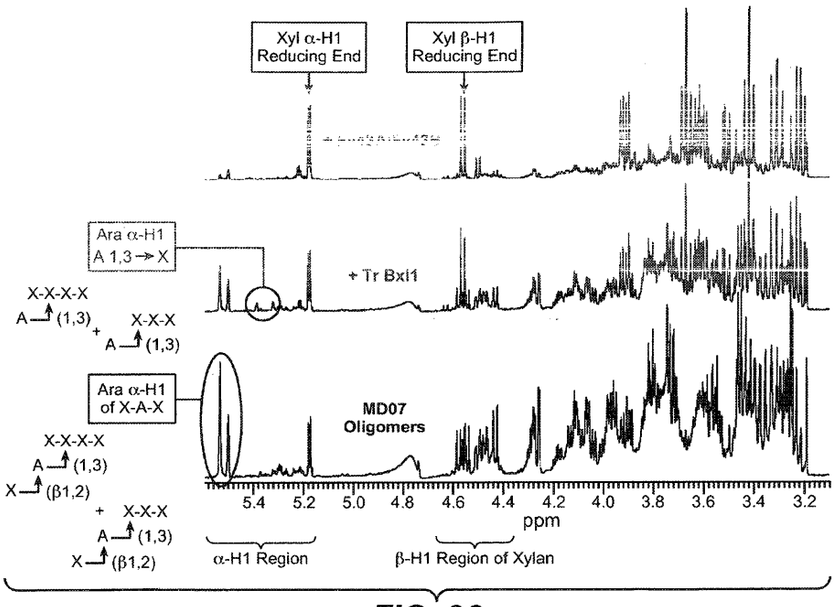


FIG. 36

SUBSTITUTE SHEET (RULE 26)

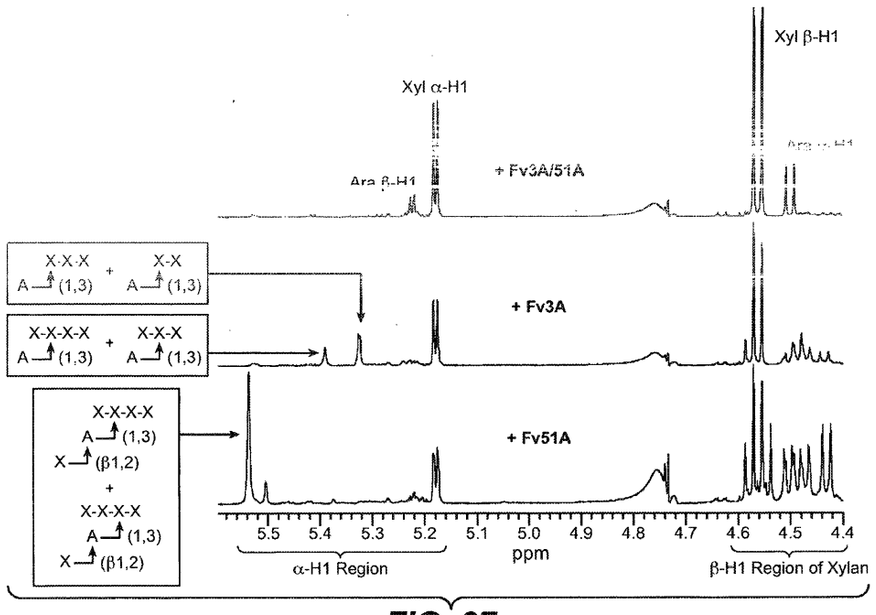


FIG. 37

SUBSTITUTE SHEET (RULE 26)

```
gi|reesei|Bxl1  -MVNNAALLAALSALIPTALAQNNQTYANYSAQGQPDLYPETLATLTLSFPDCEHGPLKN
gi|fvert| Fv3A  MLLNLQVAASALSLSLLGGLAEAATPYT----------------------LPDCTKGPLSK
                 ::*  .  :***  *  .**:   .*:                        :*** :***.:
gi|reesei|      NLVCDSSAGYVERAQALISLFTLEELILNTQNSGPGVPRLGLPNYQVWNEALHGLDRAN-
gi|fvert|       NGICDTSLSPAKRAAALVAALTPEEKVGNLVSNATGAPRIGLPRYNWWNEALHGLAGSPG
                * :**:* . .:** **:: :* ** : *  ....*.**:***.*: ********  :
gi|reesei|      ---FATKGGQFEWATSFPMPILTTAALNRTLIHQIADIISTQARAFSNSGRYGLDVYAPNV
gi|fvert|       GRFADTP-PYDAATSFPMPLLMAAAFDDDLIHDIGNVVGTEARAFTNGGWRGVDFWTPNV
                  ** .   :: *******:* :**::  ***:*.:::.*:****:*.*  *:*.::***
gi|reesei|      NGFRSPLWGRGQETPGEDAFFLSSAYTYEYITGIQGGVDPEHLKVAATVKHFAGYDLENW
gi|fvert|       NPFKDPRWGRGSETPGEDALHVS-RYARYIVRGLEG--DKEQRRIVATCKHYAGNDFEDW
                * *:.* ****.*******:.:*  *:   : *::*  * *: ::.** **:** *:*:*
gi|reesei|      NNQSRLGFDAIITQQDLSEYYTPQFLAAARYAKSRSLMCAYNSVNGVPSCANSFFLQTLL
gi|fvert|       GGFTRHDFDAKITPQDLAEYYVRPFQECTRDAKVGSIMCAYNAVNGIPACANSYLQETIL
                .. :* .*** ** ***:***.  *  .:* **  *:*****:***:*:****:: :*:*
gi|reesei|      RESWGFP-EWGYVSSDCDAVYNVFNPHDYASNQSSAAASSLRAGTDIDCGQTYPWHLNES
gi|fvert|       RGHWNWTRDNNWITSDCGAMQDIWQNHKYVKTNAEGAQVAFENGMDSSCEYTTTSDVSDS
                *  *.:. : .::***.*: :::: *.*...::..*  ::. * * .*  * . .:.:*
gi|reesei|      FVAGEVSRGEIERSVTRLYANLVRLGYFD-KKNQYRSLGWKDVVKTDAWNISYEAAVEGI
gi|fvert|       YKQGLLTEKLMDRSLKRLFEGLVHTGFFDGAKAQWNSLSFADVNTKEAQDLALRSAVEGA
                :  * ::.  ::**:.**: .**: *:**  * *:.**.: ** ..:* ::: .:****
gi|reesei|      VLLKNDGTLPLSKKVR-SIALIGPWANATTQMQGNYYGPAPYLISPLEAAKKAGYHVNFE
gi|fvert|       VLLKNDGTLPLKLKKKDSVAMIGFWANDTSKLQGGYSGRAPFLHSPLYAAEKLGLDTNVA
                **********. * : *:*:** *** *:::**.* * **:* *** **:* * ..*.
gi|reesei|      LGTEIAGNSTTG--FAKAIAAAKKSDAIIYLGGIDNTIEQEGADRTDIAWPGNQLDLIKQ
gi|fvert|       WGPTLQNSSSHDNWTTNAVAAAKKSDYILYFGGLDASAAGEDRDRENLDWPESQLTLLQK
                 *. : ..*: .   ::*:******* *:*:**:* :   *. ** :: ** .** *:::
gi|reesei|      LSEVGKPLVVLQMGGGQVDSSSLKSNKKVNSLVWGGYPGQSGGVALFDILSGKRAPAGRL
gi|fvert|       LSSLGKPLVVIQLG-DQVDDTALLKNKKINSILWVNYPGQDGGTAVMDLLTGRKSPAGRL
                **.:******:*:* .***.::* .***:**::* .****.**.*::*:*:*:::*****
gi|reesei|      VTTQYPAEYVHQFPQNDMNLRPDGKSNPGQTYIWYTGKPVYEFGSGLFYTTFKETLASHP
gi|fvert|       PVTQYPSKYTEQIGMTDMDLRPT-KSLPGRTYRWYS-TPVLPYGFGLHYTKFQAKFKSN-
                 .****::*..*:  .**:***  ** **:** **: .**  :* **.**.*: .: *:
gi|reesei|      KSLKFNTSSILSAPHPGYTYSEQIPVFTFEANIKNSGKTESPYTAMLFVRTSNAGPAPYP
gi|fvert|       -KLTFDIQKLLKG--CSAQYSDTCALPPIQVSVKNTGRITSDFVSLVFIKS-EVGPKPYP
                 .*.*: ..:*..   .  **:  .: .::..:**:*:  * :.:::*::: :.** ***
gi|reesei|      NKWLVGFDRLADIKPGHSSKLSIPIPVSALARVDSHGNRIVYPGKYELALNTDESVKLEF
gi|fvert|       LKTLAAYGRLHDVAPSSTKDISLEWTLDNIARRGENGDLVVYPGTYTLLLDEPTQAKIQV
                 * *..:.** *: *. :..:*:  .:. :** ..:*: :****.* * *:   ..*::.
gi|reesei|      ELVGEEVTIENWPLEEQQIKDATPDA
gi|fvert|       TLTGKKAILDKWPQDPKSA-------
```

FIG. 38

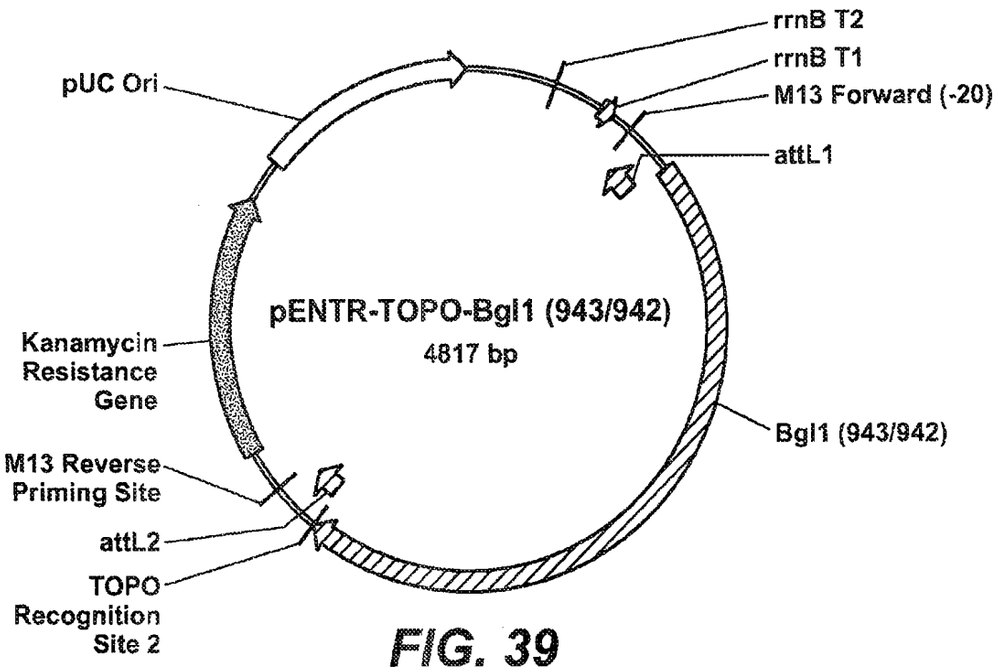

rrnB T2
rrnB T1
M13 Forward (-20)
attL1
pUC Ori
pENTR-TOPO-Bgl1 (943/942)
4817 bp
Kanamycin Resistance Gene
Bgl1 (943/942)
M13 Reverse Priming Site
attL2
TOPO Recognition Site 2

FIG. 39

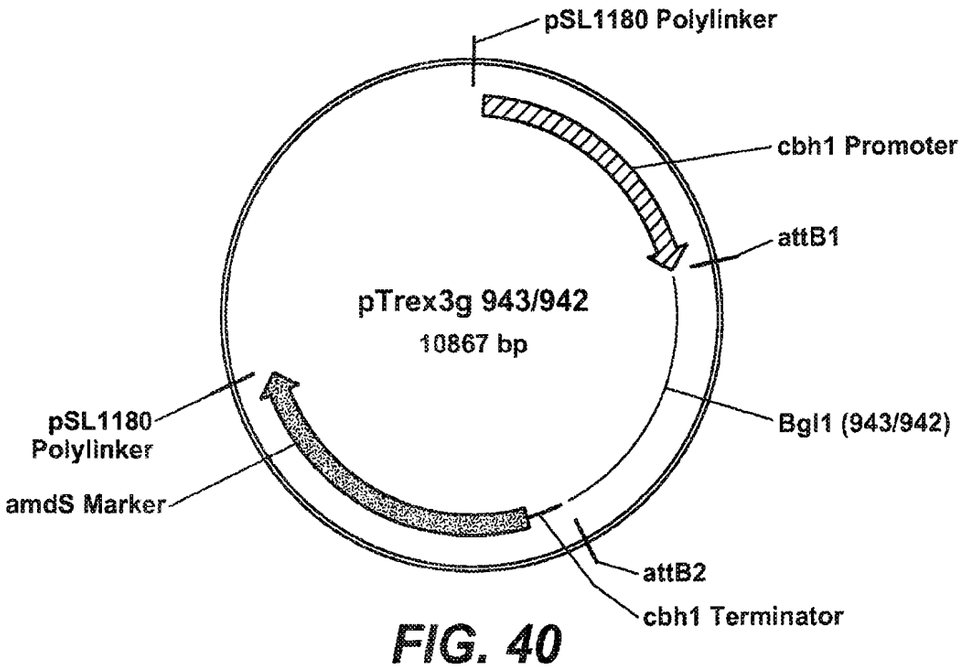

pSL1180 Polylinker
cbh1 Promoter
attB1
pTrex3g 943/942
10867 bp
Bgl1 (943/942)
pSL1180 Polylinker
amdS Marker
attB2
cbh1 Terminator

FIG. 40

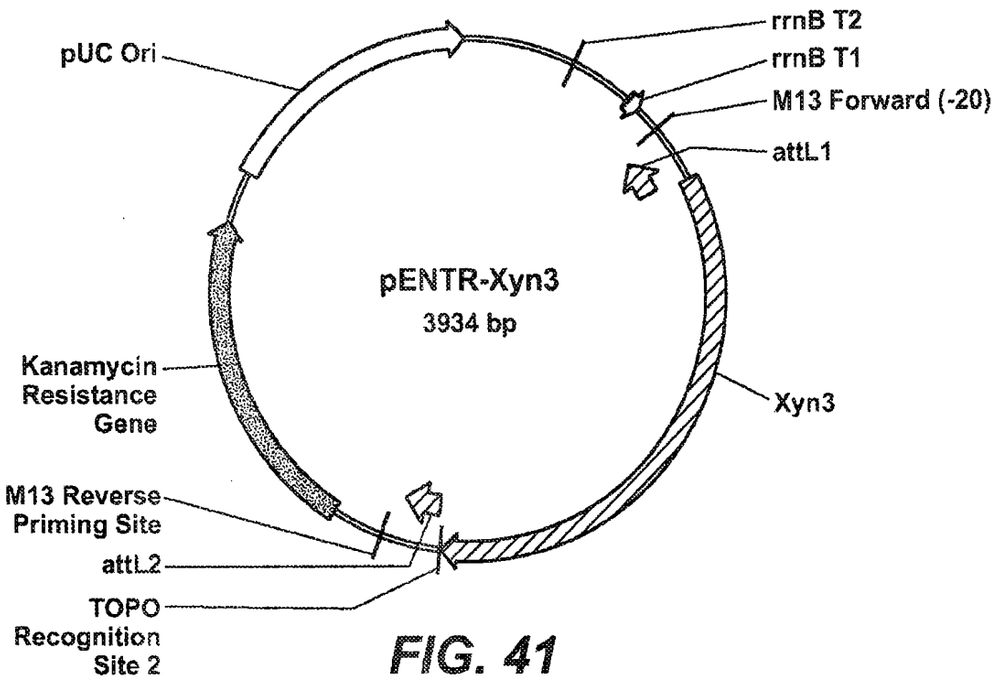

pUC Ori
rrnB T2
rrnB T1
M13 Forward (-20)
attL1
pENTR-Xyn3
3934 bp
Kanamycin Resistance Gene
Xyn3
M13 Reverse Priming Site
attL2
TOPO Recognition Site 2

FIG. 41

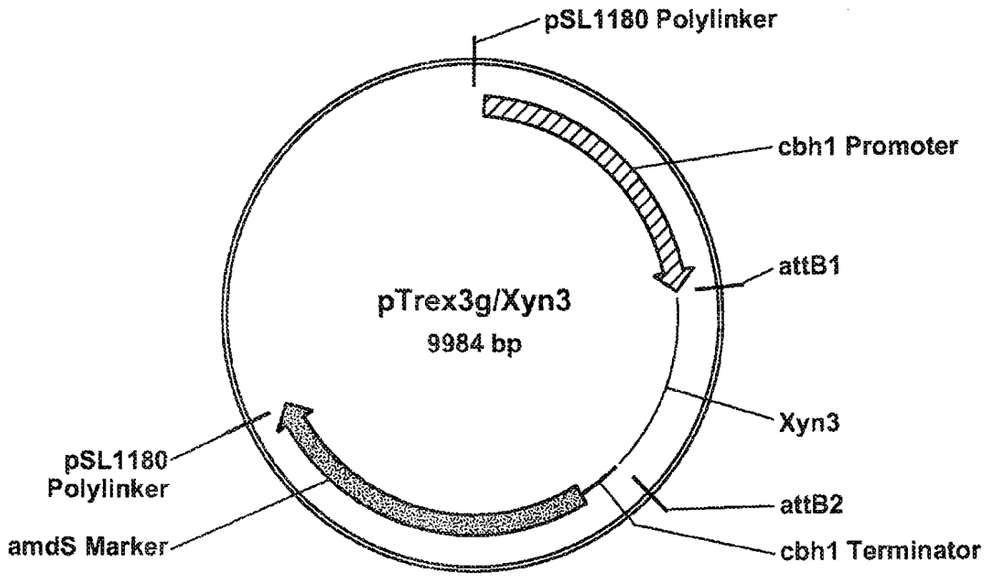

pSL1180 Polylinker
cbh1 Promoter
attB1
pTrex3g/Xyn3
9984 bp
Xyn3
pSL1180 Polylinker
attB2
amdS Marker
cbh1 Terminator

FIG. 42

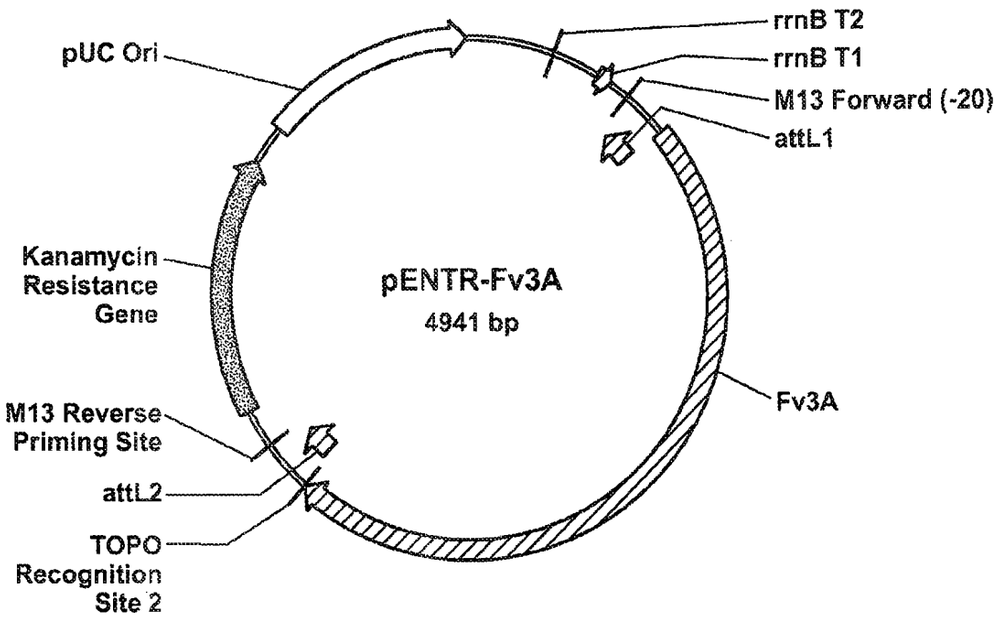

pUC Ori
rrnB T2
rrnB T1
M13 Forward (-20)
attL1
Kanamycin Resistance Gene
pENTR-Fv3A
4941 bp
Fv3A
M13 Reverse Priming Site
attL2
TOPO Recognition Site 2

FIG. 43

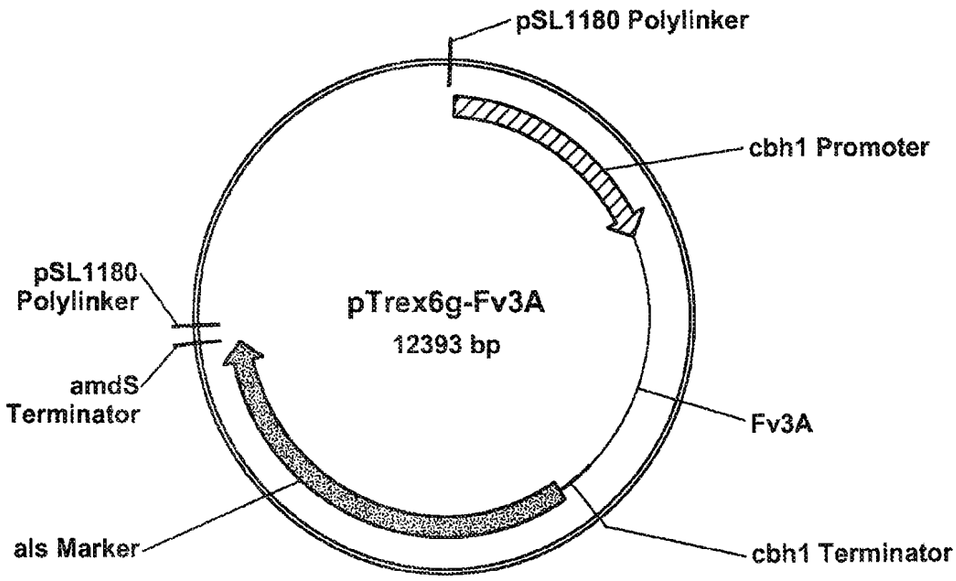

pSL1180 Polylinker
cbh1 Promoter
pSL1180 Polylinker
pTrex6g-Fv3A
12393 bp
amdS Terminator
Fv3A
als Marker
cbh1 Terminator

FIG. 44

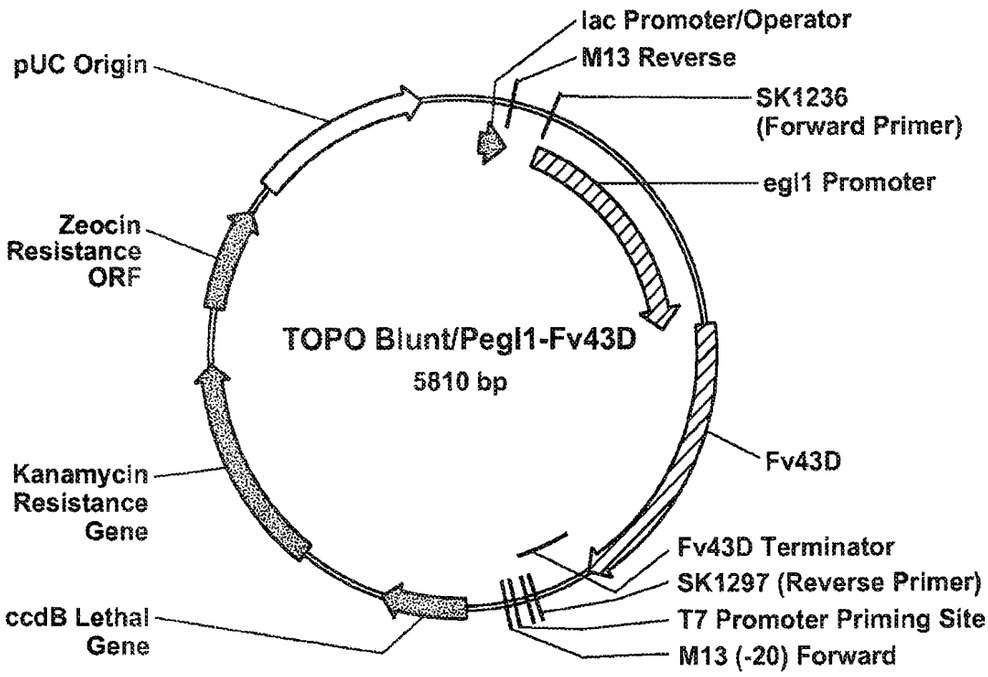

lac Promoter/Operator
M13 Reverse
SK1236
(Forward Primer)
egl1 Promoter
pUC Origin
Zeocin
Resistance
ORF
TOPO Blunt/Pegl1-Fv43D
5810 bp
Fv43D
Kanamycin
Resistance
Gene
Fv43D Terminator
SK1297 (Reverse Primer)
T7 Promoter Priming Site
ccdB Lethal
Gene
M13 (-20) Forward

FIG. 45

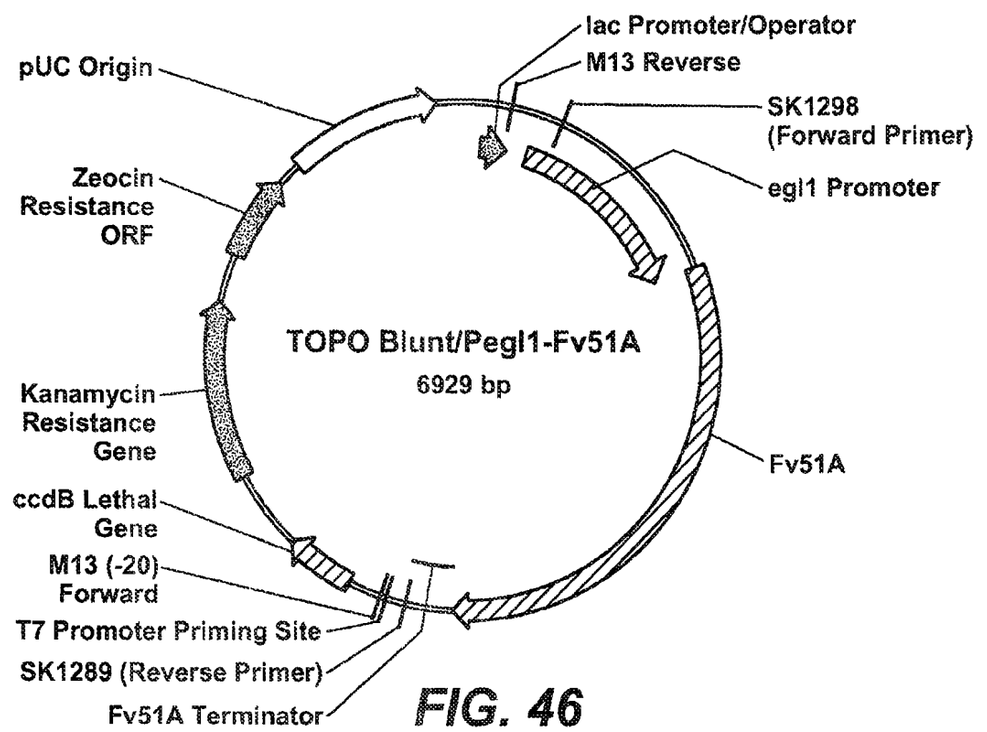

lac Promoter/Operator
M13 Reverse
SK1298
(Forward Primer)
pUC Origin
egl1 Promoter
Zeocin
Resistance
ORF
TOPO Blunt/Pegl1-Fv51A
6929 bp
Kanamycin
Resistance
Gene
Fv51A
ccdB Lethal
Gene
M13 (-20)
Forward
T7 Promoter Priming Site
SK1289 (Reverse Primer)
Fv51A Terminator

FIG. 46

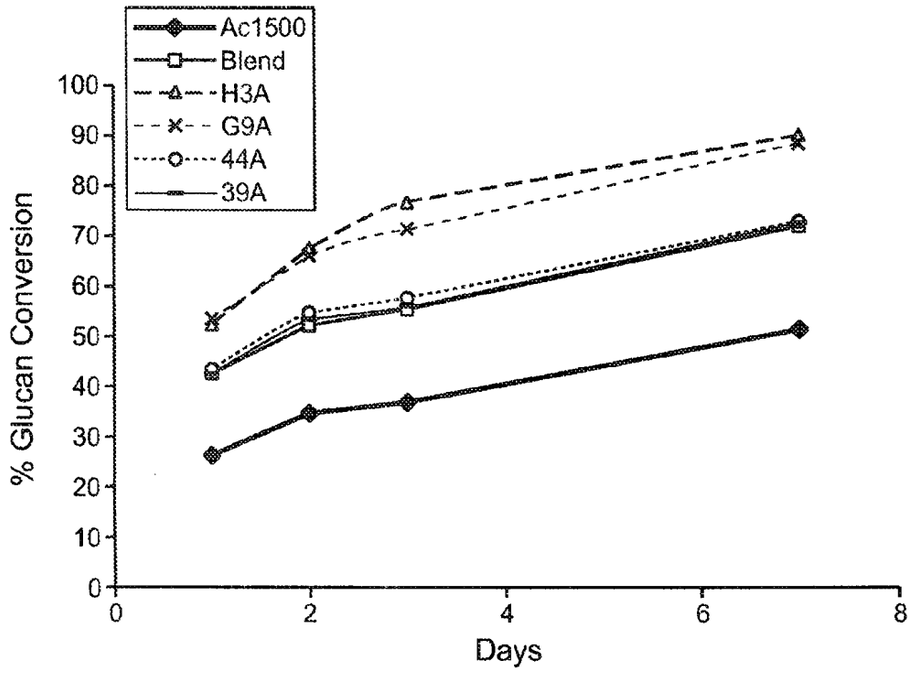


FIG. 47A

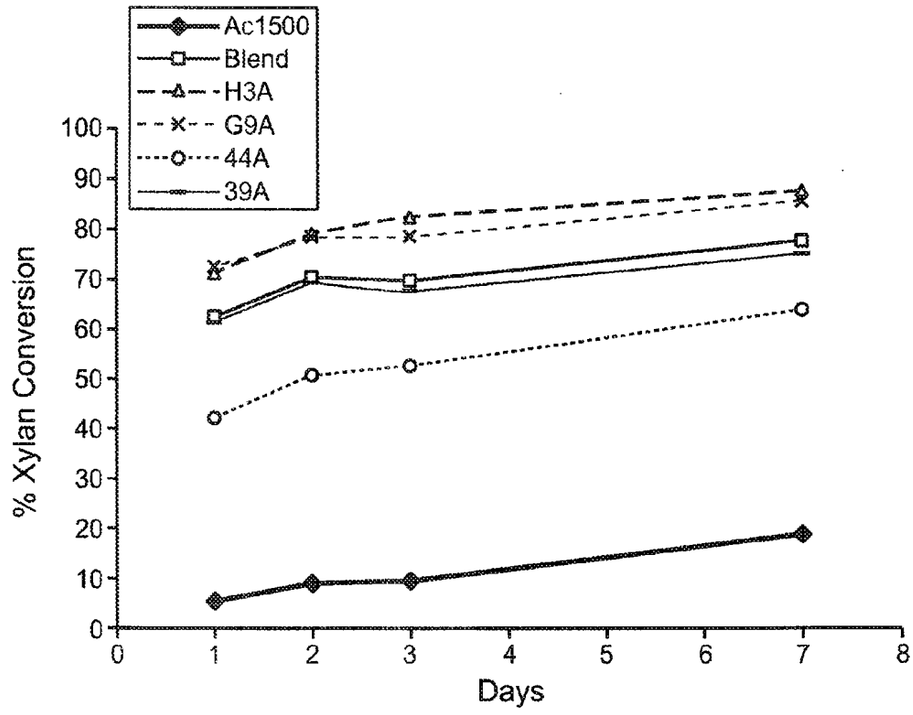


FIG. 47B

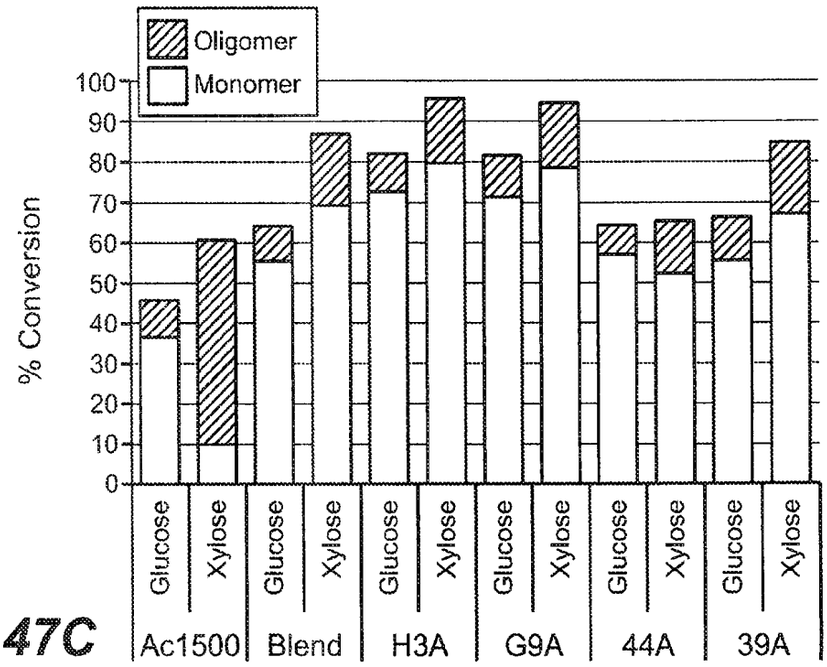


FIG. 47C

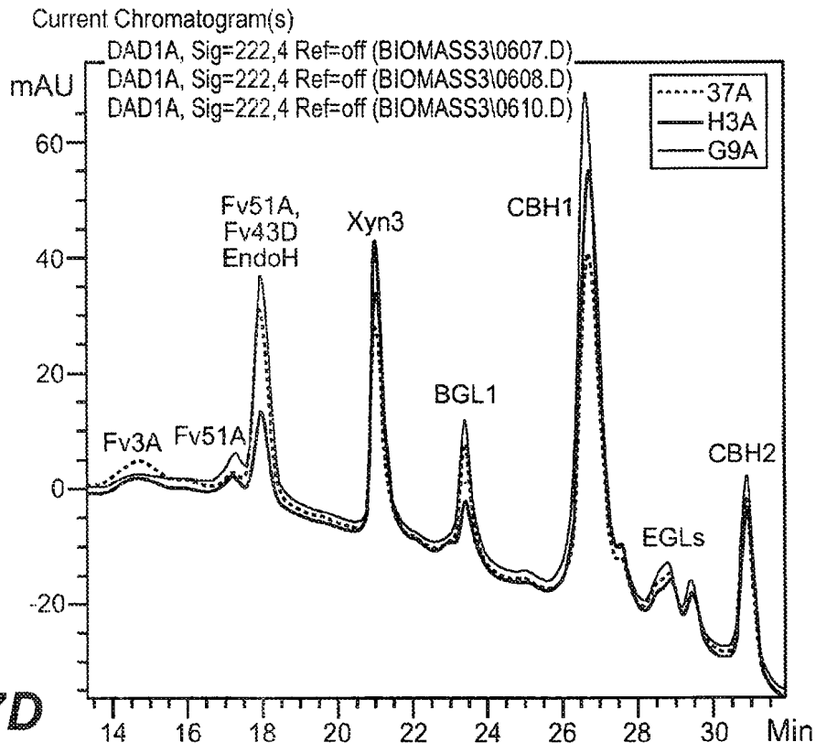


FIG. 47D

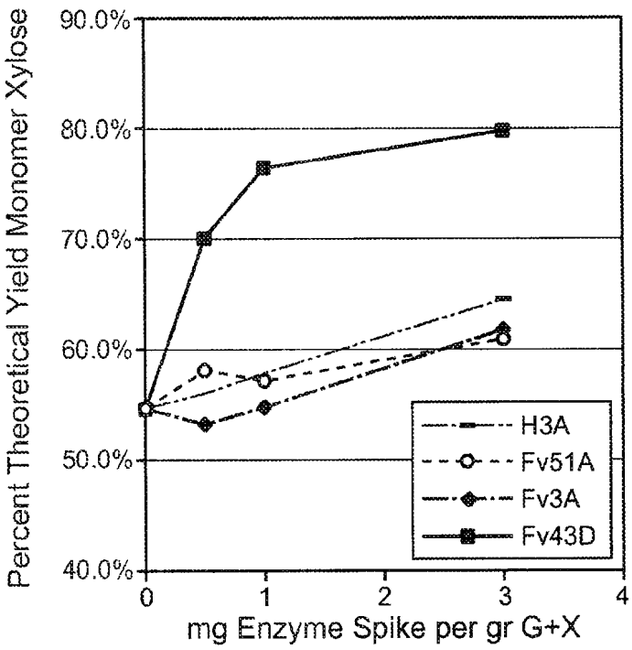


FIG. 48A

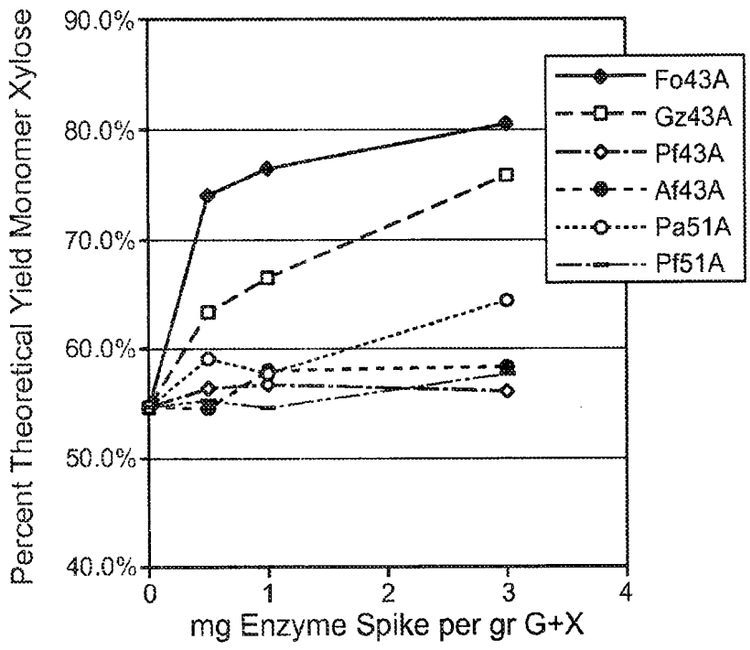


FIG. 48B

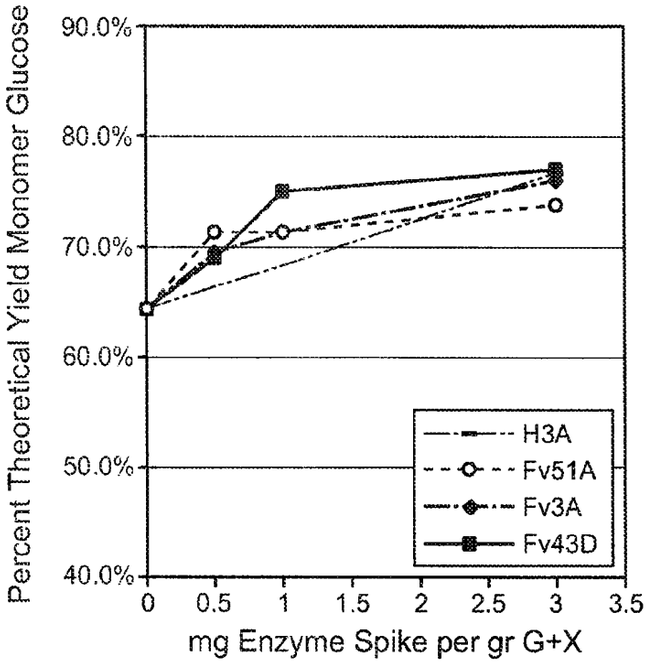


FIG. 49A

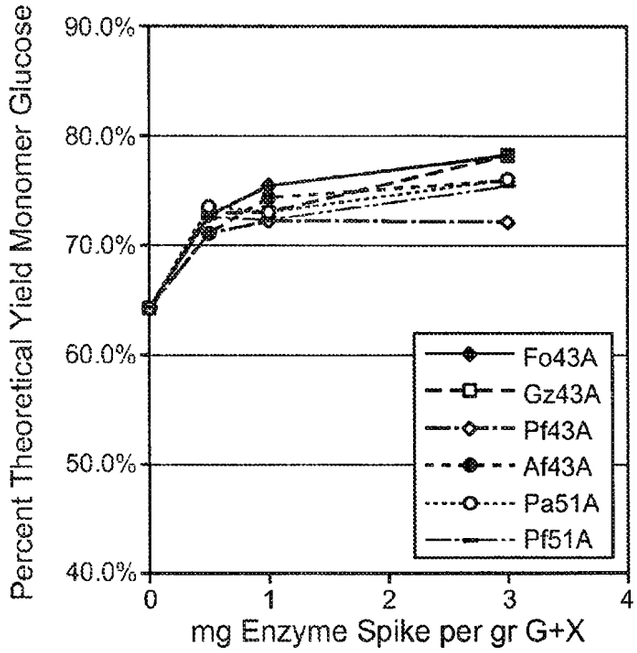


FIG. 49B

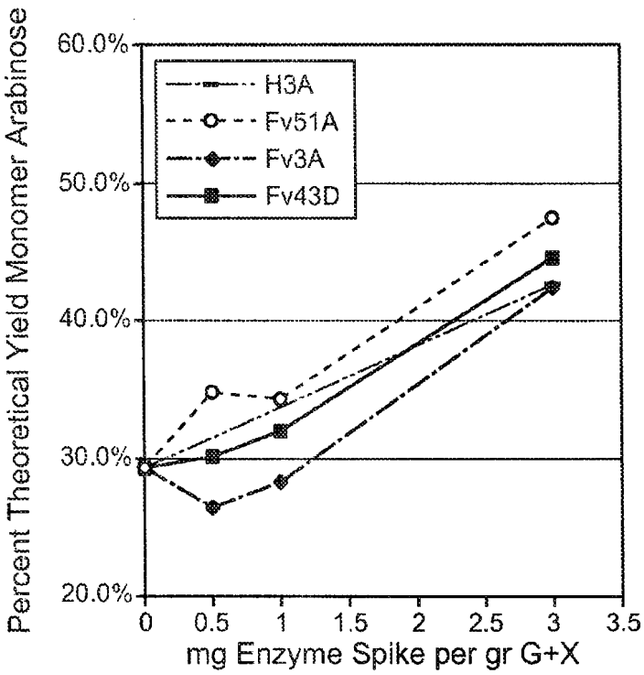


FIG. 50A

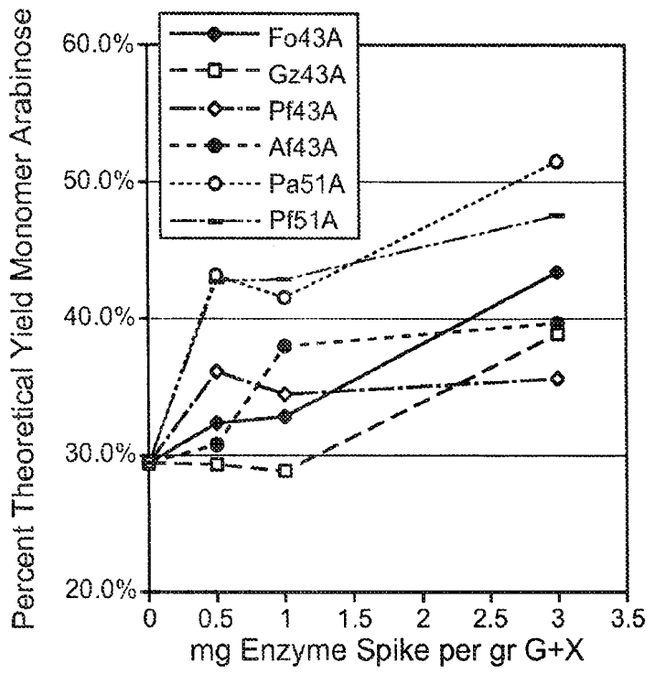


FIG. 50B

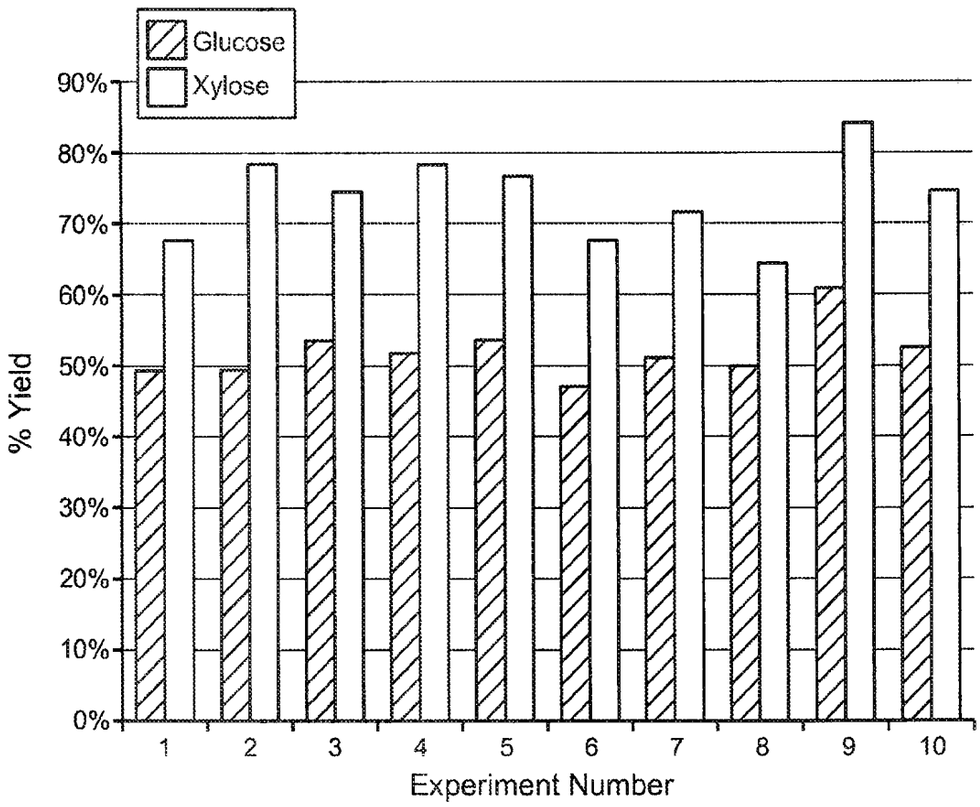

Glucose
Xylose
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
% Yield
1
2
3
4
5
6
7
8
9
10
Experiment Number

FIG. 51

```
                    *        20         *        40         *        60
Fv39A       : MHYATLTTLVLALTTNVAAQQGTATVDLSKNHGPAKALGSGFIYGWPDNGTSVDTSIPDF : 60
Q9ZFM2_XYN  : --------------MKVVNVPSNGREKFKKNW--KFCVGTGRL-GLALQKEYLDH--LKL : 41
P36906_XYN  : --------------MIKVRVPDFSDKKFSDRW--RYCVGTGRL-GLALQKEYIET--LKY : 41
                    *        80         *       100         *       120
Fv39A       : LVTDIKFNSNRGGGAQIPSLGWARGGYEG----------YLGR-FNSTLSNYRITRKYNA : 109
Q9ZFM2_XYN  : VQEKIGFRYIRGHGLLSDDVGIYREVEIDGEMKPFYNFTYIDRIVDSYLA--LNIRPF-I : 98
P36906_XYN  : VKENIDFKYIRGHGLLCDDVGIYREDVVGDEVKPFYNFTYIDRIFDSFLE--IGIRPF-V : 98
                    *       140         *       160         *       180
Fv39A       : DFILLPHDLWGADGG----QGSNSPFPGDNGNWTEM-ELFWNQLVSDLKAHNMLEGLVID : 164
Q9ZFM2_XYN  : EFGFMPKALASGDQTVEYWKGNVTP-PKDYNKWRDLIVAVVSHFIERYGIEEVRT-WLFE : 156
P36906_XYN  : EIGFMPKKLASGTQTVEYWEGNVTP-PKDYEKWSDLVKAVLHHFISRYGIEEVLK-WPFE : 156
                  A     *       200         *       220         *       240
Fv39A       : VWNEPDIDIFW-DRPWSQFLEYYNRATKLLRKTLPKTLLSGPAMAHSPILSDDKWHTWLQ : 223
Q9ZFM2_XYN  : VWNEPNLVNFWKDANKQEYFKLYEVTARAVKSVDPHLQVGGPAICGGS----DEWITDFL : 212
P36906_XYN  : IWNEPNLKEFWKDADEKEYFKLYKVTAKAIKEVNENLKVGGPAICGGA----DYWIEDFL : 212
                    *       260         *       280         *       300
Fv39A       : SVAGNKTVP-DIYSWHQIGAW--------------EREP-DSTIPDFTTLRA---QYGVP : 264
Q9ZFM2_XYN  : HFCAERRVPVDFVSRHAYTSKAPHKKTFEYYYQELELEPPEDMLEQFKTVRALIRQSPFF : 272
P36906_XYN  : NFCYEENVPVDFVSRHATTSK-QGEYTPHLIYQ--EIMPSEYMLNEFKTVREIIKNSHFP : 269
                  N *       320         *       340         *       360
Fv39A       : EKPIDVNEYAARDEQNPANSVYYLSQLERHNLRGLRANWGSGSDLHNWMGNLIYSTTGTS : 324
Q9ZFM2_XYN  : HLPLHITEYNT--SYSPINPVHDTALNAAYIARILSEGGDYVDSFSYWTFSDVFEEMDVP : 330
P36906_XYN  : NLPFHITEYNT--SYSPQNPVHDTPFNAAYIARILSEGGDYVDSFSYWTFSDVFEERDVP : 327
                    *       380         *       400         *       420
Fv39A       : EGTYY-----------PNGEWQAYKYYAAMAGQRL----------------------VT : 350
Q9ZFM2_XYN  : KALFHGGFGLVALHSIPKPTFHAFTFFNALGDELLYRDGEMIVTRRKDGSIAAVLWNLVM : 390
P36906_XYN  : RSQFHGGFGLVALNMIPKPTFYTFKFFNAMGEEMLYRDEHMLVTRRDDGSVALIAWNEVM : 387
                    *       440         *       460         *       480
Fv39A       : KASSDLKFDVFATKQGRKIKIIAGTRTVQAKYNIKISGLEVAGLPKMGTVKVRTYRFDWA : 410
Q9ZFM2_XYN  : EKGEGLTKEVQLVIPVSFSAVFIKRQIVNEQYGNAWRVWKQMGRPRFPSRQAVETLPS-A : 449
P36906_XYN  : DKTENPDEDYEVEIPVRFRDVFIKRQLIDEEHGNPWGTWIHMGRPRYPSKEQVNTLREVA : 447
                    *       500         *       520         *
Fv39A       : GP----------NGKVDGPVDLGEKKYT-YSANTVSSPST--------------- : 439
Q9ZFM2_XYN  : QPHVMTEQRRATDGVIHLSIVLSKNEVTLIEIEQVRDETSTYVGLDDGEITSYSS : 504
P36906_XYN  : KPEIMTSQPVANDGYLNLKFKLGKNAVVLYELTERIDESSTYIGLDDSKINGY-- : 500
```

FIG. 52

222

```
Fv43D ---MQLKFLSSALLLSLTGNCAAQDTNDIPPLITDLWSADPSAHVFEGKLWVYPSHDIEA
Fo43A ---MQLKFLSSALLFSLTSKCAAQDTNDIPPLITDLWSADPSAHVFEGKLWVYPSHDIEA
Gz43A ---MKSKLLFP--LLSFVG--QSLATNDDCPLITSRWTADPSAHVFNDTLWLYPSHDIDA
Pf43A --MLQRFAYILPLALLSVG--VKADN----PFVQSIYTADPAPMVYNDRVYVFMDHDNTG
Fv43A -MWLTSPLLFASTLLGLTGVALADN-----PIVQDIYTADPAPMVYNGRVYLFTGHDNDG
Fv43B --MRFSWLLCPLLAMGSALPETKTDVSTYTNPVLPGWHSDPSC-IQKDGLFLCVTSTFIS
Af43A --------------MAAPSLSYPTGIQSYTNPLFPGWHSDPSCAYVAEQDTFFCVTSTFI
Pf43B ----MSRSILPYASVFALLGGAIAEP----FLVLNSDFPDPSLIETSSGYYAFGTTGNGV
Fv43E MKVYWLVAWATSLTPALAGLIGHRRATTFNNPIIYSDFPDNDVFLGPDNYYYFSASNFHF

Fv43D NVVNGTGGAQYAMRDYHTYSMKTIYGKDPVIDHGVALSVDDVPWAKQQMWAPDAAYK--N
Fo43A NVVNGTGGAQYAMRDYHTYSMKSIYGKDPVVDHGVALSVDDVPWAKQQMWAPDAAHK--N
Gz43A GFENDPDGGQYAMRDYHVYSIDKIYGSLP-VDHGTALSVEDVPWASRQMWAPDAAHK--N
Pf43A -------ATYYNMTDWHLFSSADMANWQD---HGIPMSLANFTWANANAWAPQVIPR--N
Fv43A -------STDFNMTDWRLFSSADMVNWQH---HGVPMSLKTFSWANSRAWAGQVVAR--N
Fv43B FP----GLPVYASRDLVNWRLISHVWNRE---KQLPGISWKTAGQQQGMYAPTIRYH--K
Af43A AFP---GLPLYASRDLQNWKLASNIFNRP---SQIPDLR-VTDGQQSGIYAPTLRYH--E
Pf43B N------AQVASSPDFNTWTLLSGT-------DALPGPFPSWVASSPQIWAPDVLVKA-D
Fv43E SP----GAPVLKSKDLLNWDLIGHSIPRLNFGDGYDLPPGSRYYRG-GTWASSLRYRKSN

Fv43D GKYYLYFPAK-DK-DEIFRIGVAVSNKPSGPFK---ADK-SWIPGTYSIDPASYVDTNGE
Fo43A GKYYLYFPAK-DK-DEIFRIGVAVSNKPSGPFK---ADK-SWIPGTYSIDPASYVDTDNE
Gz43A GKYYLYFPAK-DK-DDIFRIGVAVSPTPGGPFV---PDK-SWIPHTFSIDPASFVDDDDR
Pf43A GQFYFYAPVR-HN-DGSMAIGVGVSSTITGPYH---DAIGKPLVENNEIDPTVFIDDDGQ
Fv43A GKFYFYVPVRNAK-TGGMAIGVGVSTNILGPYT---DALGKPLVENNEIDPTVYIDTDGQ
Fv43B GTYYVICEYLGVG-DIIGVIFKTTNPWDESSWS---DPV---TFKPNHIDPDLFWDDDGK
Af43A GQFYLIVSYLGP--QTKGLLFTSSDPYDDAAWS---DPL---EFAVHGIDPDIFWDHDGT
Pf43B GTYVMYFSASAASDSGKHCVGAATATSPEGPYTPVDSAVACPLDQGGAIDANGFIDTDGT
Fv43E GQWYWIGCIN------FWQTWVYTASSPEGPWY----NKGNFGDNNCYYDNGILIDDDDT

Fv43D AYLIWGGI-WGGQLQAWQDHKTFNESWLGDKAAPNGTNALSPQIAKLSKDMHKITETPRD
Fo43A AYLIWGGI-WGGQLQAWQDKKNFNESWIGDKAAPNGTNALSPQIAKLSKDMHKITETPRD
Gz43A AYLAWGGI-MGGQLQRWQDKNKYNES--GTEPG-NGTAALSPQIAKLSKDMHTLAEKPRD
Pf43A AYLYWG--------------------------------NPDLWYYVKLNQDMISYSGSPTQ
Fv43A AYLYWG--------------------------------NPGLYYVKLNQDMLSYSGSINK
Fv43B VYCATHG----ITLQEIDLETGELSPELNIWNGTGGVWPEGPHIYKRDGYYYLMIAEGGT
Af43A VYVTSAED-QMIKQYTLDLKTGAIGPVDYLWNGTGGVWPEGPHIYKRDGYYYLMIAEGGT
Pf43B IYVVYKID-------------------------GNSLDGDGTTHPTPIMLQQMEADGT
Fv43E MYVVYGSGEVKVSQLSQDGFSQVKSQVVFKNTDIGVQDLEGNRMYKING-------LYYI
```

FIG. 53A

```
Fv43D LVILAPETGKPLQAEDNKRRFFEGP-----WVHKRGKLYYLMYSTG----------
Fo43A LVILAPETGKPLQAEDNKRRFFEGP-----WIHKRGKLYYLMYSTG----------
Gz43A MLILDPKTGKPLLSEDEDRRFFEGP-----WIHKRNKIYYLTYSTG----------
Pf43A IPLTTAGFGTRTGNAQRPTTFEEAP-----WVYKRNGIYYIAYAAD----------
Fv43A VSLTTAGFGSRPNNAQRPTTFEEGP-----WLYKRGNLYYMIYAAN----------
Fv43B ----AEDHAITIARARKITGPYEAYNNNPILTNRGTSEYFQTVGHGDLFQDTKGNWWGLC
Af43A ----ELGHSETMARSRTRTGPWEPYPHNPLLSNKGTSEYFQTVGHADLFQDGNGNWWAVA
Pf43B --TPTGSPIQLIDRSDLDGPLIEAP-----SLLLSNGIYYLSFSSN----------
Fv43E LNDSPSGSQTWIWKSKSPWGPYESKVLADKVTPPISGGNSPHQGSLIKTPNGGWY-----

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 LATRITAQGVSPMGREAVLFNGTWNKGEWPKLQPVRGRMPGNLLPKPTRN------VPGD
Af43A LSTRSGPAWKNYPMGRETVLAPAAWEKGEWPVIQPVRGQMQG-PFPPPNKR----VPRGE
Pf43B YYNTNYYDTSYAYASSITGPWT----------KQSAPYAPLLVTGT--------ETSND
Fv43E FMSFTWAYPAGRLPVLAPITWG----------SDGFPILVKGANGGWGSSYPTLPGT

Fv43D SIVEYKGQWWLFFAD-AHTSGKDYLRQVKARKIWYDKDG-----KILLTRPKI-------
Fo43A SIVEYKGQWWLFFAD-AHTSGKDYLRQVKARKIWYDKNG-----KILLHRP---------
Gz43A SIVKYQGQWWLFYHD-AKTSGKDYLRQVKAKKIWYDSKG-----KILTKKP---------
Pf43A GIIDFQNNSYFFYHNGALPGGGGYQRSVCVEQFKYNADG-----TIPTIEMTTAG-----
Fv43A GIIDFENNSYFFYHNGALDGGSGYTRSVAVESFKYGSDG-----LIPEIKMTTQG-----
Fv43B GPFNADPDNYNLKKTKKIPPHFVHHRVPRDGAFSLSSKG----LHIVPSRNNVTGSVLPG
Af43A GGWIKQPDKVDFRPGSKIPAHFQYWRYPKTEDFTVSPRGHPNTLRLTPSFYNLTG-----
Pf43B GALSAPGGADFSVDGTKMLFHANLNGQDISGGRALFAAS------ITEASDVVTLQ----
Fv43E DGVTKNWTRTDTFRGTSLAPSWEWNHNPDVNSFTVNNGLTLRTASITKDIYQARN-----

Fv43D ------------------------------------------------------------
Fo43A ------------------------------------------------------------
Gz43A ------------------------------------------------------------
Pf43A ----------------PAQIGTLNPYVRQEAETAAWSSGITTEVCSEGGIDVGFINNG
Fv43A ----------------PAQLKSLNPYVKQEAETIAWSEGIETEVCSEGGLNVAFIDNG
Fv43B DEIELSGQRGLAFIGRRQTHTLFKYSVDIDFKPKSDDQEAGITVFRTQFDHIDLGIVRLP
Af43A -TADFKPDDGLSLVMRKQTDTLFTYTVDVSFDPKVADEEAGVTVFLTQQQHIDLGIVLLQ
Pf43B ------------------------------------------------------------
Fv43E ----------TLSHRTHGDHPTGIVKIDFSPMKDGDRAGLSAFRDQSAYIGIHRDNGK
```

FIG. 53B

```
Fv43D ------------------------------------------------------------
Fo43A ------------------------------------------------------------
Gz43A ------------------------------------------------------------
Pf43A DYIK----------VKGVAFGS-GAESFSARVASANSGGTIAIHLGSTTGTLVGTCTV
Fv43A DYIK----------VKGVDFGSTGAKTFSARVASNSSGGKIELRLGSKTGKLVGTCTV
Fv43B TNQGSNKKSKLAFRFRATGAQNVPAPK---VVPVPDGWEKGVISLHIEAANATHYNLGAS
Af43A TTEG----LSLSFRFRVEGRGNYEGPLPEATVPVPKEWCGQTIRLEIQAVSDTEYVFAAA
Pf43B ------------------------------------------------------------
Fv43E FTIAT----KHGMNMDEWNGTTTDLGQIKATANVPSGRTKIWLRLQLDTNPAGTGNTIFS

Fv43D ------------------------------------------------------------
Fo43A ------------------------------------------------------------
Gz43A ------------------------------------------------------------
Pf43A PSTGGWQTWTTVTCSVSGASGTQ--------DVYFVFGGSGTGYLFN------FDYWQFA
Fv43A TTTGNWQTYKTVDCPVSGATGTS--------DLFFVFTGSGSGSLFN------FNWWQFS
Fv43B --SHRGKTLDIATASASLVSGGTGSFVGSLLGPYATCNGKGSGVECPKGGDVYVTQWTYK
Af43A PARHPAQRQIISRANSLIVSGDTGRFTGSLVGVYATSNG-GAGSTP-----AYISRWRYE
Pf43B ------------------------------------------------------------
Fv43E YSWDGVKYETLGPNFKLYNG----------WAFFIAYRFGIFNFAETALGGSIKVESFT

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

FIG. 53C

```
Pa51A   MIHLKPALAALLALSTQCVAIDLFVKSSGGNKTTDIMYGLMHEDINNSGDGGIYAELISN
Fv51A   MVRFSSILAAAACF-VAVESVNIKVDSKGGNATSGHQYGFLHEDINNSGDGGIYAELIRN
Pf51A   MGKMWHSILVVLGLLSVGHAITINVSQSGGNKTSPLQYGLMFEDINHGGDGGLYAELVRN

Pa51A   RAFQGSEKFPSNLDNWSPVGGATLTLQKLAKPLSSALPYSVNVANPKEGKGKGKDTKGKK
Fv51A   RAFQYSKKYPVSLSGWRPINDAKLSLNRLDTPLSDALPVSMNVK---PGKGK-----AKE
Pf51A   RAFQGSTVYPANLDGYDSVNGAILALQNLTNPLSPSMPSSLNVA-----KGS----NNGS

Pa51A   VGLANAGFWGMDVKRQKYTGSFHVTGEYKGDFEVSLRSAITGETFGKKVVKGGSKKGKWT
Fv51A   IGFLNEGYWGMDVKKQKYTGSFWVKGAYKGHFTASLRSNLTDDVFGSVKVKSKANKKQWV
Pf51A   IGFANEGWWGIEVKPQRYAGSFYVQGDYQGDFDISLQSKLTQEVFATAKVRSSGKHEDWV

Pa51A   EKEFELVPFKDAPNSNNTFVVQWDAEGAKDGSLDLNLISLFPPTFKGRKNGLRIDLAQTM
Fv51A   EHEFVLTPNKNAPNSNNTFAITYDPKGA-DGALDFNLISLFPPTYKGRKNGLRVDLAEAL
Pf51A   QYKYELVPKKAASNTNNTLIITFDSKGLKDGSLNFNLISLFPPTYNNRPNGLRIDLVEAM

Pa51A   VELKPTFLRFPGGNMLEGNTLDTWWKWYETIGPLKDRPGMAGVWEYQQTLGLGLVEYMEW
Fv51A   EGLHPSLLRFPGGNMLEGNTNKTWWDWKDTLGPLRNRPGFEGVWNYQQTHGLGILEYLQW
Pf51A   AELEGKFLRFPGGSDVEGVQAPYWYKWNETVGDLKDRYSRPSAWTYEESNGIGLIEYMNW

Pa51A   ADDMNLEPIVGVFAGLALDGSFVPESEMGWVIQQALDEIEFLTGDAKTTKWGAVRAKLGH
Fv51A   AEDMNLEIIVGVYAGLSLDGSVTPKDQLQPLIDDALDEIEFIRG-PVTSKWGKKRAELGH
Pf51A   CDDMGLEPILAVWDGHYLSNEVISENDLQPYIDDTLNQLEFLMG-APDTPYGSWRASLGY

Pa51A   PKPWKVKWVEIGNEDWLAGRPAGFESYINYRFPMMMKAFNEKYPDIKIIASPSIFD----
Fv51A   PKPFRLSYVEVGNEDWLAGYPTGWNSYKEYRFPMFLEAIKKAHPDLTVISSGASIDPVGK
Pf51A   PKPWTINYVEIGNEDNLYG---GLETYIAYRFQAYYDAITAKYPHMTVMESLTEMPG---

Pa51A   ---NMTIPAGAAGDHHPYLIPDEFVERFAKFDNLSKDNVTLIGEAASTHPNG---GIAWE
Fv51A   KDAGFDIPAPGIGDYHPYREPDVLVEEFNLFDNNKYG--HIIGEVASTHPNG---GTGWS
Pf51A   -------PAAAASDYHQYSIPDGFVSQFNYFDQMPVTNRTLNGEIATVYPNNPSNSVAWG

Pa51A   GDLMPLPWWGGSVAEAIFLISTERNGDKIIGATYAPGLRSLDRWQWSMTWVQHAADPALT
Fv51A   GNLMPYPWWISGVGEAVALCGYERNADRIPGTFYAPILKNENRWQWAITMIQFAADSAMT
Pf51A   SPFPLYPWWIGSVAEAVFLIGEERNSPKIIGASYAPMFRNINNWQWSPTLIAFDADSSRT

Pa51A   TRSTSWYVWRILAHHIIRETLPVDAPAGKPNFDPLFYVAGKSES-GTGIFKAAVYNSTES
Fv51A   TRSTSWYVWSLFAGHPMTHTLPTTA-----DFDPLYYVAGKNEDKGTLIWKGAAYNTTKG
Pf51A   SRSTSWHVIKLLSTNKITQNLPTTWSGG--DIGPLYWVAGRNDNTGSNIFKAAVYNSTSD

Pa51A   --IPVSLKFDGLNEGAVANLTVLTGPE-DPYGYNDPFTGINVVKEKTTFIKAGKGGKFTF
Fv51A   ADVPVSLSFKGVKPGAQAELTLLTNKEKDPFAFNDPHKGNNVVDTKKTVLKADGKGAFNF
Pf51A   --VPVTVQFAGCN-AKSANLTILSSDD--PNASNYPG-GPEVVKTEIQSVTANAHGAFEF

Pa51A   TLPGLSVAVLETADAVKGGKGKGKGKGKGN
Fv51A   KLPNLSVAVLETLK---------KGKPYSS
Pf51A   SLPNLSVAVLKTE----------------
```

FIG. 54

```
                 *         20          *         40          *
xyn3    : -MKANVILC--LLAPLVAALPTETIHLDPELAALRANLTERTADLWDRQA :  47
P56588  : ----------------------------------------------QA :   2
P23360  : MVRPTILLTSLLLAPFAAASPI----------------------LEERQA :  28

                60          *         80          *        100
xyn3    : SQSIDQLIKRKGKLYFGTATDRGLLQRE-KNAAIIQADLGQVTPENSMKW :  96
P56588  : SVSIDAKFKAHGKKYLGTIGDQYTLTKNTKNPAIIKADFGQLTPENSMKW :  52
P23360  : AQSVDQLIKARGKVYFGVATDQNRLTTG-KNAAIIQADFGQVTPENSMKW :  77

                *         120         *         140         *
xyn3    : QSLENNQGQLNWGDADYLVNFAQQNGKSIRGHTLIWHSQLPAWVNNINNA : 146
P56588  : DATEPNRGQFTFSGSDYLVNFAQSNGKLIRGHTLVWHSQLPGWVSSITDK : 102
P23360  : DATEPSQGNFNFAGADYLVNWAQQNGKLIRGHTLVWHSQLPSWVSSITDK : 127

               160          *        180          *        200
xyn3    : DTLRQVIRTHVSTVVGRYKGKIRAWDVVNEIFNEDGTLRSSVFSRLLGEE : 196
P56588  : NTLISVLKNHITTVMTRYKGKIYAWDVLNEIFNEDGSLRNSVFYNVIGED : 152
P23360  : NTLTNVMKNHITTLMTRYKGKIRAWDVVNEAFNEDGSLRQTVFLNVIGED : 177

                *         220         *         240         *
xyn3    : FVSIAFRAARDADPSARLYINDYNLDRANYGKVNGLKTYVSKWISQGVPI : 246
P56588  : YVRIAFETARSVDPNAKLYINDYNLDSAGYSKVNGMVSHVKKWLAAGIPI : 202
P23360  : YIPIAFQTARAADPNAKLYINDYNLDSASYPKTQAIVNRVKQWRAAGVPI : 227

               260          *        280      N  *        300
xyn3    : DGIGSQSHLSGGGGSGTLGALQQLATVPVTELAITELDIQGAPTTDYTQV : 296
P56588  : DGIGSQTHLGAGAGSAVAGALNALASAGTKEIAITELDIAGASSTDYVNV : 252
P23360  : DGIGSQTHLSAGQGAGVLQALPLLASAGTPEVAITELDVAGASPTDYVNV : 277

                *         320         *         340         *
xyn3    : VQACLSVSKCVGITVWGISDKDSWRASTNPLLFDANFNPKPAYNSIVGIL : 346
P56588  : VNACLNQAKCVGITVWGVADPDSWRSSSSPLLFDGNYNPKAAYNAIANAL : 302
P23360  : VNACLNVQSCVGITVWGVADPDSWRASTTPLLFDGNFNPKPAYNAIVQDL : 327

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

FIG. 55A

207

```
  1  MVSFSYLLLACSAIG-ALAAPVEPETTSFN  AfuXyn2
  1  MISISSLSFGLAAIAGAYALPSD-KSVSLA  AfuXyn5
  1  MVSFTSLLAASPPSRASCRPAAEVESVAVE  Xyn2

 30  ETALHEFAERAGTPSSTGWNNGYYYSFWTD  AfuXyn2
 30  ER--------QTITTSQTGTNNGYYYSFWTN  AfuXyn5
 31  KR--------QTIQPG-TGYNNGYFYSYWND  Xyn2

 60  GGGDVTYTNGAGGSYSVNWRN--VGNFVGG  AfuXyn2
 53  GAGSVQYTNGAGGEYSVTWANQNGGDFTCG  AfuXyn5
 53  GHGGVTYTNGPGGQFSVNWSN--SGNFVGG  Xyn2

 88  KGWNPGSA-RTINYGGSFNPSGNGYLAVYG  AfuXyn2
 83  KGWNPGSD-HDITFSGSFNPSGNAYLSVYG  AfuXyn5
 81  KGWQPGTKNKVINFSGSYNPNGNSYLSVYG  Xyn2
                N
117  WTTNPLIEYYVVESYGTYNPGSGGTFRGTV  AfuXyn2
112  WTTNPLVEYYILENYGSYNPGSGMTHKGTV  AfuXyn5
111  WSRNPLIEYYIVENFGTYNPSTGATKLGEV  Xyn2

147  NTDGGTYNIYTAVRYNAPSIEGTKTFTQYW  AfuXyn2
142  TSDGSTYDIYEHQQVNQPSIVGTATFNQYW  AfuXyn5
141  TSDGSVYDIYRTQRVNQPSIIGTATFYQYW  Xyn2

177  SVRTSKRTGGTVTMANHFNAWSRLGMNLGT  AfuXyn2
172  SIRQNKRSSGTVTTANHFKAWASLGMNLGT  AfuXyn5
171  SVRRNHRSSGSVNTANHFNAWAQQGLTLGT  Xyn2
                A
207  HNYQIVATEGYQSSGSASITVY          AfuXyn2
202  HNYQIVSTEGYESSGTSTITVSSGGSSSGG  AfuXyn5
201  MDYQIVAVEGYFSSGSASITVS          Xyn2

228                                  AfuXyn2
232  SGGSSSTTSSGSSPTGGSGSCSALWGQCGG  AfuXyn5
222                                  Xyn2

228                                  AfuXyn2
262  IG-WSGPTCCSSGTCQVSNSYYSQCL      AfuXyn5
222                                  Xyn2
```

FIG. 55B

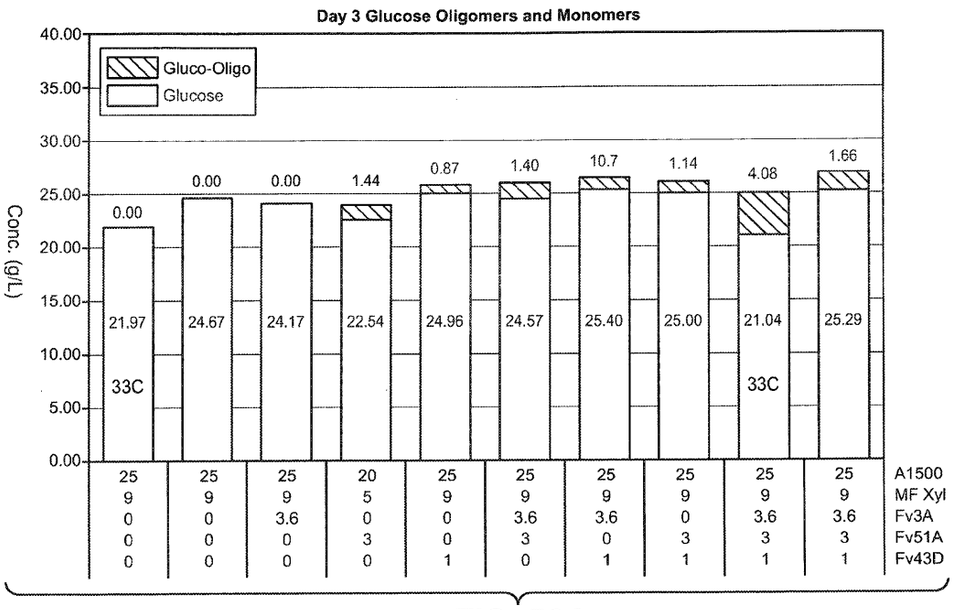


FIG. 56A

SUBSTITUTE SHEET (RULE 26)

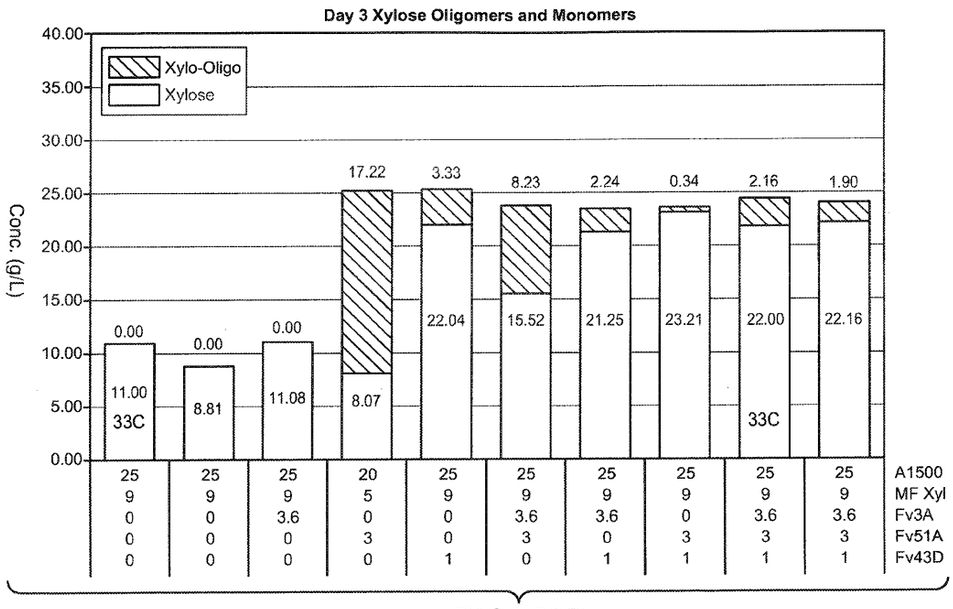


FIG. 56B

SUBSTITUTE SHEET (RULE 26)

230

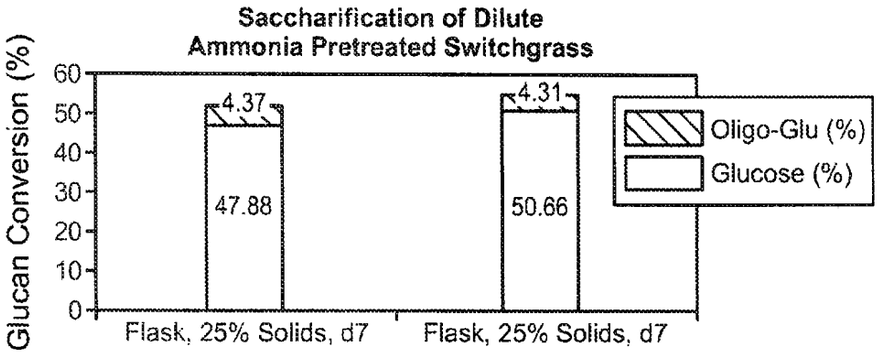


FIG. 57A

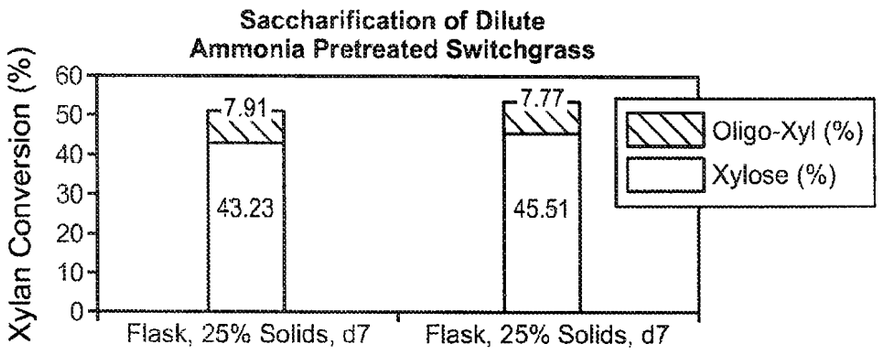


FIG. 57B

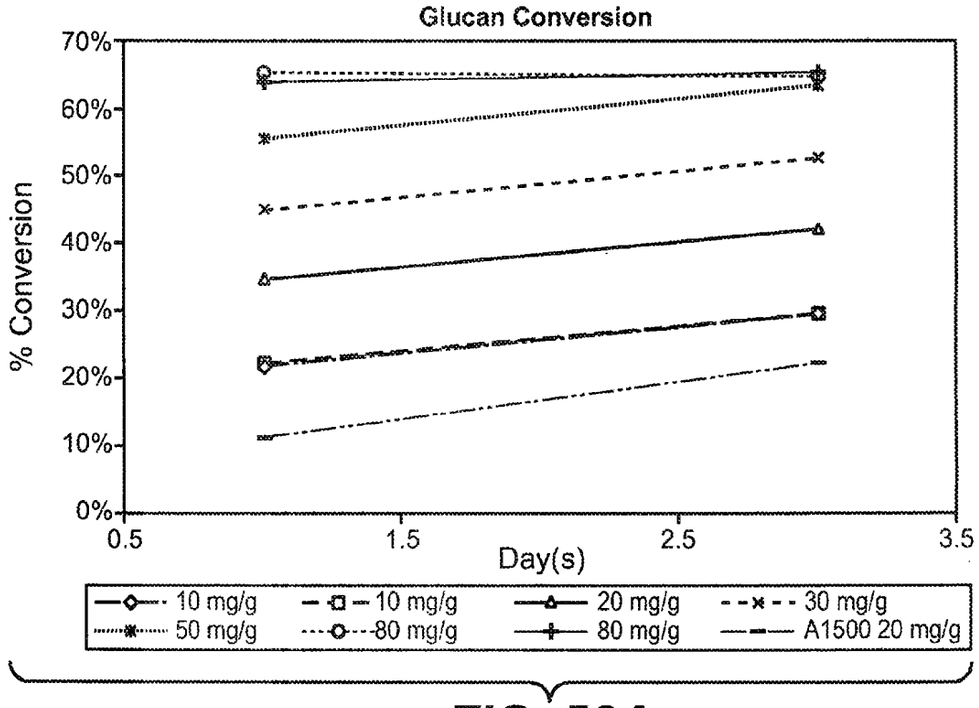


FIG. 58A

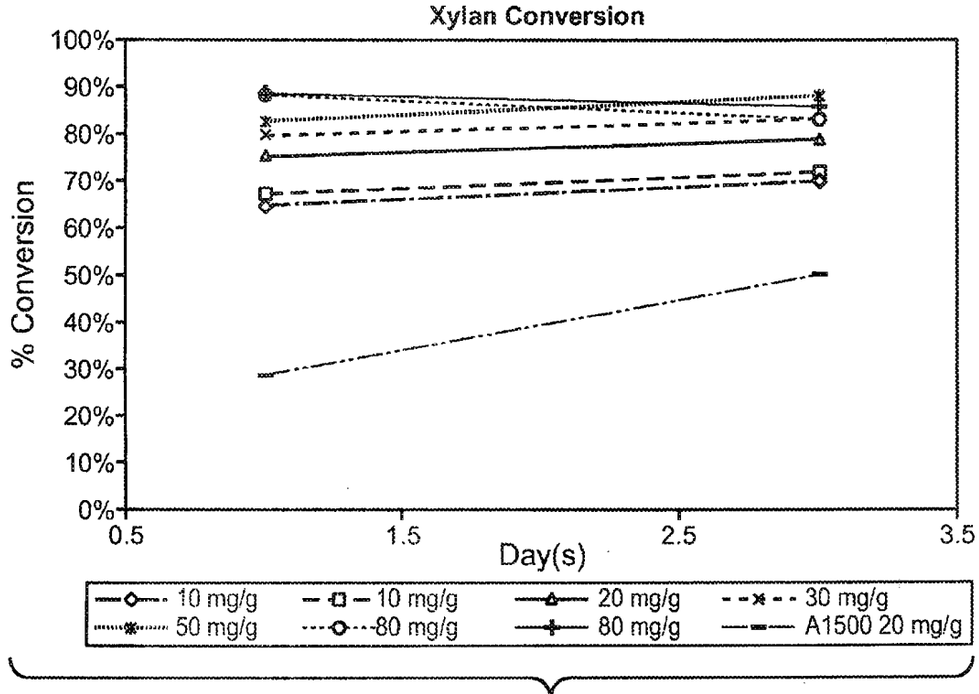


FIG. 58B

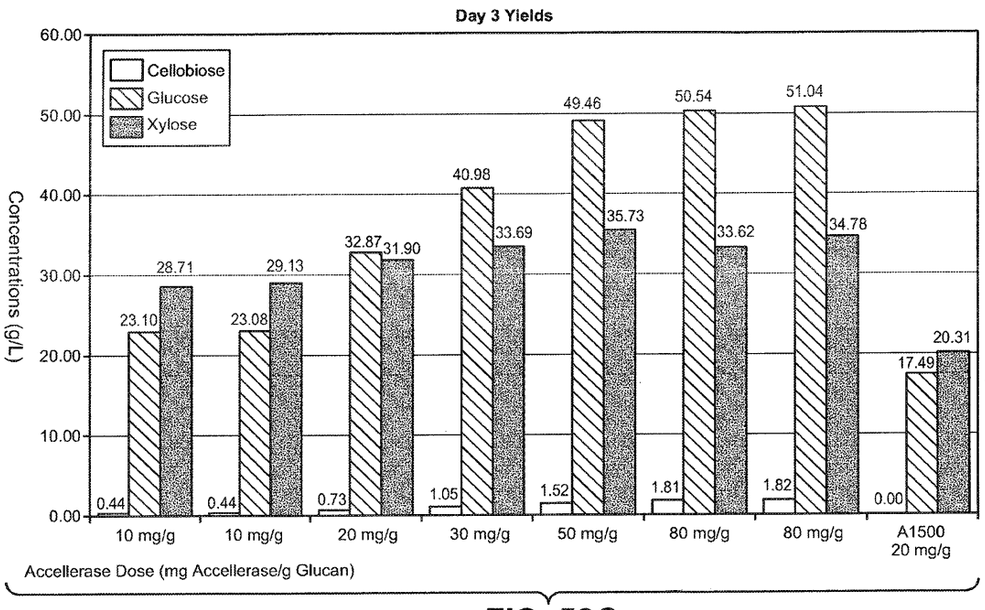


FIG. 58C

SUBSTITUTE SHEET (RULE 26)

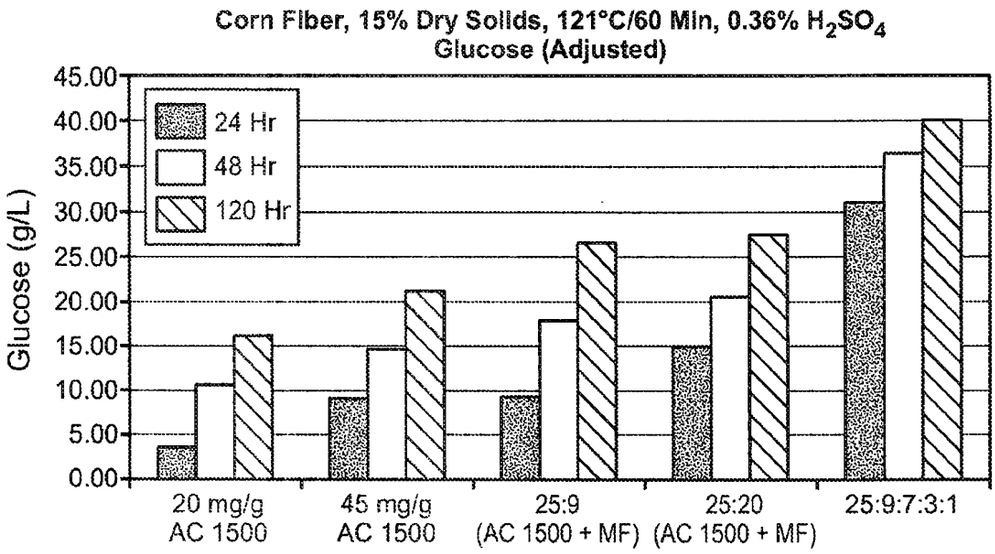


FIG. 59A

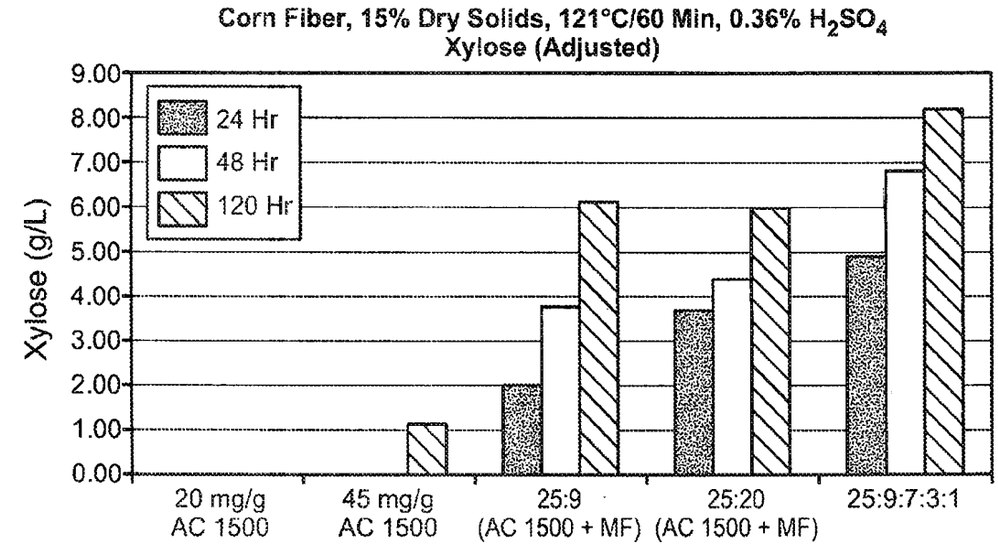


FIG. 59B

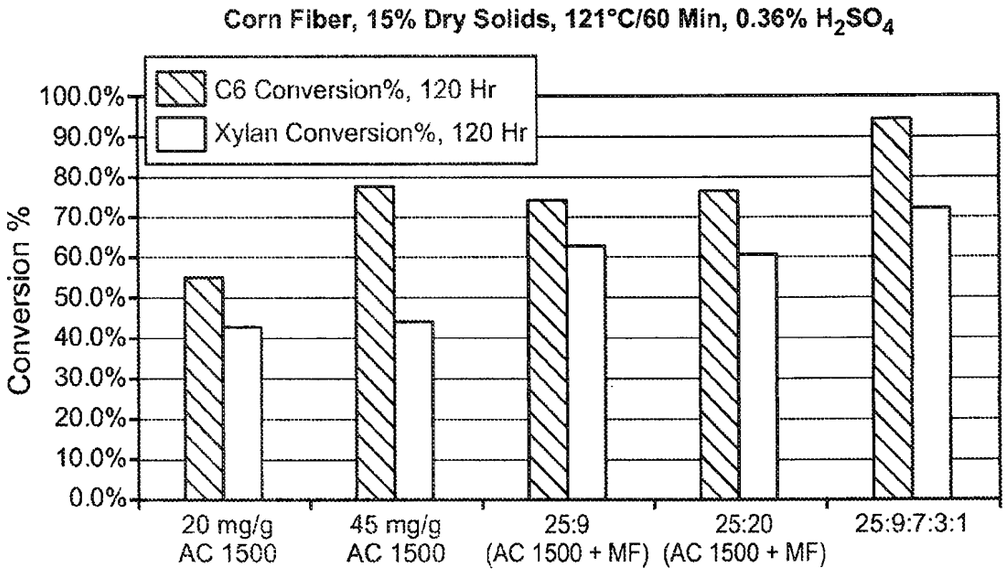

Corn Fiber, 15% Dry Solids, 121°C/60 Min, 0.36% H_2SO_4
C6 Conversion%, 120 Hr
Xylan Conversion%, 120 Hr
Conversion %
100.0%
90.0%
80.0%
70.0%
60.0%
50.0%
40.0%
30.0%
20.0%
10.0%
0.0%
20 mg/g AC 1500
45 mg/g AC 1500
25:9 (AC 1500 + MF)
25:20 (AC 1500 + MF)
25:9:7:3:1

FIG. 59C

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

atgatccacctcaagccagccctcgcggcgttgttggcgctgtcgacgcaatgtgtggctattgatttgtttgtcaagtcttcggg
ggggaataagacgactgatatcatgtatggtcttatgcacgaggatatcaacaactccggcgacggcggcatctacgccgagctaa
tctccaaccgcgcgttccaagggagtgagaagttcccctccaacctcgacaactggagccccgtcggtggcgctacccttaccctt
cagaagcttgccaagccccttcctctgcgttgccttactccgtcaatgttgccaaccccaaggagggcaagggcaagggcaagga
caccaagggggaagaaggttggcttggccaatgctgggttttggggtatggatgtcaagaggcagaagtacactggtagcttccacg
ttactggtgagtacaagggtgactttgaggttagcttgcgcagcgcgattaccggggagacctttggcaagaaggtggtgaagggt
gggagtaagaaggggaagtggaccgagaaggagtttgagttggtgcctttcaaggatgcgcccaacagcaacaacacctttgttgt
gcagtgggatgccgagggcgcaaaggacggatctttggatctcaacttgatcagcttgttccctccgacattcaagggaaggaaga
atgggctgagaattgatcttgcgcagacgatggttgagctcaagccgaccttcttgcgcttccccggtggcaacatgctcgagggt
aacaccttggacacttggtggaagtggtacgagaccattggccctctgaaggatcgcccgggcatggctggtgtctgggagtacca
gcaaaccctggcttgggtctggtcgagtacatggagtgggccgatgacatgaacttggagcccattgtcggtgtcttcgctggtc
ttgccctcgatggctcgttcgttcccgaatccgagatgggatgggtcatccaacaggctctcgacgaaatcgagttcctcactggc
gatgctaagaccaccaaatggggtgccgtccgcgcgaagcttggtcaccccaagccttggaaggtcaagtgggttgagatcggtaa
cgaggattggcttgccggacgccctgctggcttcgagtcgtacatcaactaccgcttccccatgatgatgaaggccttcaacgaaa
agtaccccgacatcaagatcatcgcctcgcctccatcttcgacaacatgacaatcccgcgggtgctgccggtgatcaccacccg
tacctgactcccgatgagttcgttgagcgattcgccaagttcgataacttgagcaaggataacgtgacgctcatcggcgaggctgc
gtcgacgcatcctaacggtggtatcgcttgggagggagatctcatgcccttgccttggtggggcggcagtgttgctgaggctatct
tcttgatcagcactgagagaaacggtgacaagatcatcggtgctacttacgcgcctggtcttcgcagcttggaccgctggcaatgg
agcatgacctgggtgcagcatgccgccgacccggccctcaccactcgctcgaccagttggtatgtctggagaatcctcgcccacca
catcatccgtgagacgctcccggtcgatgccccggccggcaagcccaactttgaccctctgttctacgttgccggaaagagcgaga
gtggcaccggtatcttcaaggctgccgtctacaactcgactgaatcgatcccggtgtcgttgaagtttgatggtctcaacgaggga
gcggttgccaacttgacggtgcttactgggccggaggatccgtatggatacaacgacccttcactggtatcaatgttgtcaagga
gaagaccaccttcatcaaggccggaaagggcggcaagttcaccttcaccctgccgggcttgagtgttgctgtgttggagacggccg
acgcggtcaagggtggcaagggaaagggcaagggcaagggaaagggtaactga

FIG. 60A

SUBSTITUTE SHEET (RULE 26)

736

SEQ ID NO:47
Codon optimized cDNA for Pa51A, a GH51 family enzyme from *Podospora anserina*

atgatccacctcaagcccgccctcgccgccctcctcgccctcagcacccaatgcgtcgccatcgacctct
tcgtcaagagcagcggcggcaacaagaccaccgacatcatgtacggcctcatgcacgaggacatcaacaa
cagcggcgacggcggcatctacgccgagctgatcagcaaccgcgccttccagggcagcgagaagttcccc
agcaacctcgacaactggtccccgtcggcggcgccaccctcaccctccagaagctcgccaagcccctgt
cctctgccctccctactccgtcaacgtcgccaaccccaaggagggtaagggtaagggcaaggacaccaa
gggcaagaaggtcggcctcgccaacgccggcttttggggcatggacgtcaagcgccagaaatacaccggc
agcttccacgtcaccggcgagtacaagggcgacttcgaggtcagcctccgcagcgccattaccggcgaga
ccttcggcaagaaggtcgtcaagggcggcagcaagaagggcaagtggaccgagaaggagttcgagctggt
ccccttcaaggacgcccccaacagcaacaacaccttcgtcgtccagtgggacgccgagggcgccaaggac
ggcagcctcgacctcaacctcatcagcctcttcccgcccaccttcaagggccgcaagaacggcctccgca
tcgacctcgcccagaccatggtcgagctgaagcccaccttcctccgctttcccggcggcaacatgctcga
gggcaacaccctcgacacctggtggaagtggtacgagaccatcggcccccctgaaggaccgccctggcatg
gccggcgtctgggagtaccagcagacgctgggcctcggcctggtcgagtacatggagtgggccgacgaca
tgaacctcgagcccatcgtcggcgtctttgctggcctggccctggatggcagctttgtccccgagagcga
gatgggctgggtcatccagcaggctctcgatgagatcgagttcctcaccggcgacgccaagaccaccaag
tggggcgccgtccgcgccaagctcggccaccctaagccctggaaggtcaaatgggtcgagatcggcaacg
aggactggctcgccggccgacctgccggcttcgagagctacatcaactaccgcttccccatgatgatgaa
ggccttcaacgagaaataccccgacatcaagatcattgccagcccctccatcttcgacaacatgaccatt
ccagccggtgctgccggtgaccaccaccctacctcaccccgacgaatttgtcgagcgcttcgccaagt
tcgacaacctcagcaaggacaacgtcaccctcattggcgaggccgccagcacccaccccaacggcggcat
tgcctgggagggcgacctcatgccccctgccctggtggggcggcagcgtcgccgaggccatcttcctcatc
agcaccgagcgcaacggcgacaagatcatcggcgccacctacgccccctggcctccgatctctcgaccgct
ggcagtggagcatgacctgggtccagcacgccgccgaccctgccctcaccaccccgcagcaccagctggta
cgtctggcgcatcctcgcccaccacatcattcgcgagaccctccccgtcgacgcccccgccggcaagccc
aacttcgaccccctcttctacgtcgctggcaagtcggagagcggcaccggcatcttcaaggccgccgtct
acaacagcaccgagagcatccccgtcagcctcaagttcgacggcctcaacgagggcgccgtcgccaacct
caccgtcctcaccggccccgaggaccccctacggctacaacgacccctttcaccggcatcaacgtcgtcaag
gaaaagaccaccttcatcaaggccggcaagggcggcaagttcacctttaccctccccggcctctctgtcg
ccgtcctcgagaccgccgacgccgtgaagggtggcaagggaaagggaaagggcaagggtaagggtaacta
a

FIG. 60B

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

atgtatcggaagttggccgtcatctcggccttcttggccacagctcgtgctaccaacgacgactgtcctc
tcatcactagtagaaggactgcggatccttcggctcatgtctttaacgacaccttgtggctctacccgtc
tcatgacatcgatgctggatttgagaatgatcctgatggaggccagtacgccatgagagattaccatgtc
tactctatcgacaagatctacggttccctgccggtcgatcacggtacggccctgtcagtggaggatgtcc
cctgggcctctcgacagatgtgggctcctgacgctgcccacaagaacggcaaatactacctatacttccc
tgccaaagacaaggatgatatcttcagaatcggcgttgctgtctcaccaacccccggcggaccattcgtc
cccgacaagagttggatccctcacactttcagcatcgaccccgccagtttcgtcgatgatgatgacagag
cctacttggcatggggtggtatcatgggtggccagcttcaacgatggcaggataagaacaagtacaacga
atctggcactgagccaggaaacggcaccgctgccttgagccctcagattgccaagctgagcaaggacatg
cacactctggcagagaagcctcgcgacatgctcattcttgaccccaagactggcaagccgctcctttctg
aggatgaagaccgacgcttcttcgaaggaccctggattcacaagcgcaacaagatttactacctcaccta
ctctactggcacaacccactatcttgtctatgcgacttcaaagacccccctatggtccttacacctaccag
ggcagaattctggagccagttgatggctggactactcactctagtatcgtcaagtaccagggtcagtggt
ggctattttatcacgatgccaagacatctggcaaggactatcttcgccaggtaaaggctaagaagatttg
gtacgatagcaaaggaaagatcttgacaaagaagccttga

FIG. 61

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

atgtatcggaagttggccgtcatctcggccttcttggccacagctcgtgctcaagacactaatgacattc
ctccctgatcaccgacctctggtccgcagatccctcggctcatgttttcgaaggcaagctctgggtttta
cccatctcacgacatcgaagccaatgttgtcaacggcacaggaggcgctcaatacgccatgagggattac
catacctactccatgaagagcatctatggtaaagatcccgttgtcgaccacggcgtcgctctctcagtcg
atgacgttccctgggcgaagcagcaaatgtgggctcctgacgcagctcataagaacggcaaatattatct
gtacttccccgccaaggacaaggatgagatcttcagaattggagttgctgtctccaacaagcccagcggt
cctttcaaggccgacaagagctggatccctggcacgtacagtatcgatcctgctagctacgtcgacactg
ataacgaggcctacctcatctgggggcggtatctgggggcggccagctccaagcctggcaggataaaaagaa
ctttaacgagtcgtggattgcagacaaggctgctcctaacggcaccaatgccctatctcctcagatcgcc
aagctaagcaaggacatgcacaagatcaccgaaacaccccgcgatctcgtcattctcgccccccgagacag
gcaagcctcttcaggctgaggcacaacaagcgacgattcttcgagggcccttggatccacaagcgcggcaa
gctttactacctcatgtactccaccggtgataccacttccttgtctacgctacttccaagaacatctac
ggtccttatacctaccggggcaagattcttgatcctgttgatgggtggactactcatggaagtattgttg
agtataagggacagtggtggctttttctttgctgatgcgcatacgtctggtaaggattaccttcgacaggt
gaaggcgaggaagatctggtatgacaagaacggcaagatcttgcttcaccgtccttag

FIG. 62

SEQ ID NO:50
Nucleotide sequence for Pf51A, a GH51 family enzyme from *Penicillium funiculosum*

atgtaccggaagctcgccgtcatcagcgccttcctggcgactgctcgcgccatcaccatcaacgtcagcc
agagcggcggcaacaagaccagcccgctccagtacggcctcatgttcgaggacatcaaccacggcggcga
cggcggcctctacgccgagctggtccggaaccgggccttccagggcagcaccgtctacccggccaacctc
gacggctacgactcggtgaacggcgcgattctcgcgctccagaacctcaccaacccgctcagcccgagca
tgccctcgtcgctgaacgtcgccaagggctcgaacaacggcagcatcggcttcgccaacgaggggtggtg
gggcatcgaggtcaagccgcagcggtacgccggcagcttctacgtccagggcgactaccagggcgacttc
gacatcagcctccagagcaagctcacccaggaggtcttcgcgacggcgaaggtccggtcgagcggcaagc
acgaggactgggtccagtacaagtacgagctggtcccgaagaaggccgccagcaacaccaacaacaccct
caccatcaccttcgacagcaagggcctcaaggacggcagcctcaacttcaacctcatcagcctcttcccg
ccgacctacaacaaccggccgaacggcctccggatcgacctcgtcgaggccatggcggagctggagggca
agttcctccgcttccccggccgctcggacgtggagggcgtccaggccccgtactggtacaagtggaacga
gaccgtcggcgacctcaaggaccgctactcgcgcccgagcgcctggacctacgaggagagcaacggcatc
ggcctcatcgagtacatgaactggtgcgacgacatgggcctcgagccgatcctcgccgtctgggacggcc
actacctcagcaacgaggtcatcagcgagaacgacctccagccgtacatcgacgacaccctcaaccagct
cgagttcctcatgggcgccccggacactccctacgggtcttggagggctagcctcggctacccgaagccg
tggaccatcaactacgtcgagatcggcaacgaggacaacctctacggcggcctcgagacctacatcgcct
accggttccaggcctactacgacgccatcaccgccaagtacccgcacatgaccgtcatggagagcctcac
cgagatgcccggccccgctgccgcggcgtcggactaccaccagtactcgacgcccgacggcttcgtcagc
cagttcaactacttcgaccagatgccggtcaccaaccgcacgctgaacggcgagatcgccaccgtctacc
ccaacaacccgagcaactcggtggcgtggggcagcccgttcccgctctacccgtggtggatcgggtccgt
ggctgaggccgtcttcctcatcggcgaggagcggaacagcccgaagatcatcggcgccagctacgccccc
atgttccgcaacattaacaactggcagtggagcccgaccctgatcgccttcgacgccgacagcagccgga
cgtcgcgctctacttcctggcacgtcatcaagctcctcagcaccaacaagatcacccagaacctgcccac
gacgtggtctgggggggacatcggcccgctctactgggtcgccggccggaacgacaacaccggcagcaac
atcttcaaggccgccgtctacaacagcaccagcgacgtcccggtcaccgtccagttcgccggctgcaacg
ccaagagcgccaacctcaccatcctctcgtcggacgaccccaacgccagcaactacccgggcggccccga
ggtcgtcaagaccgagatccagagcgtcaccgccaacgcccacggcgccttcgagttcagcctcccgaac
ctgtcggtggctgtgctgaagacggagtag

FIG. 63

```
Fv3A    MLLNLQVAASALSLSLLGGLAEAATPYT--------------------LPDCTKGPLSK
Bxl1    -MVNNAALLAALSALLPTALAQNNQTYANYSAQGQPDLYPETLATLTLSFPDCEHGPLKN
Bgl1    MRY-RTAAALALATG------PFARADSHSTSG---ASAEAVVP----------------

Fv3A    NGICDTSLSPAKRAAALVAALTPEEKVGNLVSNATGAPRIGLPRYNWWNEALHGLAGSPG
Bxl1    NLVCDSSAGYVERAQALISLFTLEELILNTQNSGPGVPRLGLPNYQVWNEALHGLDRAN-
Bgl1    -PAGTPWGTAYDKAKAALAKLNLQDKVGIVSGVGWNGGPCVGNTSPASKISYPSLCLQDG

Fv3A    GRFADTP-PYDAATSFPMPLLMAAAFDDDLIHDIGNVVGTEARAFTNGGWRGVDFWTPNV
Bxl1    --FATKGGQFEWATSFPMPILTTAALNRTLIHQIADIISTQARAFSNSGRYGLDVYAPNV
Bgl1    ---PLGVRYSTGSTAFTPGVQAASTWDVNLIRERGQFIGEEVKASGIHVILGPVAGPLGK

Fv3A    NPFKDPRWGRGSETPGEDALHVS-RYARYIVRGLEG--DKEQRRIVATCKHYAGNDFEDW
Bxl1    NGFRSPLWGRGQETPGEDAFFLSSAYTYEYITGIQGGVDPEHLKVAATVKHFAGYDLENW
Bgl1    TPQGGRNWEGF---GVDP--YLTGIAMGQTINGIQS------VGVQATAKHYILNEQELN

Fv3A    GGFTR---------HDFDAKITPQDLAEYYVRPFQECTRDAKVGSIMCAYNAVNGIPACA
Bxl1    NNQSR---------LGFDAIITQQDLSEYYTPQFLAAARYAKSRSLMCAYNSVNGVPSCA
Bgl1    R-------------ETISSNPDDRTLHELYTWPFADAVQAN-VASVMCSYNKVNTTWACE

Fv3A    NSYLQETILRGHWNWTRDNNWITSDCGAMQDIWQNHKYVKTNAEGAQVAFENGMDSSCEY
Bxl1    NSFFLQTLLRESWGFP-EWGYVSSDCDAVYNVFNPHDYASNQSSAAASSLRAGTDIDCGQ
Bgl1    DQYTLQTVLKDQLGFP---GYVMTDWNAQHTTVQSANSGLDMSMPG-TDFNGNNRLWGPA

Fv3A    TTTSDVSDSYKQGLLTEKLMDRSLKRLFEGLVHTGFFDGAKAQWN---------------
Bxl1    TYPWHLNESFVAGEVSRGEIERSVTRLYANLVRLGYFD-KKNQYR---------------
Bgl1    LTNAVNSNQVPTSRVDDMVTRILAAWYLTGQDQAGYPSFNIS-----------------

Fv3A    -SLSFADVNTKEAQDLALRSAVEGAVLLKNDG-TLPLKLKKKDSVAMIGFWAN-------
Bxl1    -SLGWKDVVKTDAWNISYEAAVEGIVLLKNDG-TLPLSKKVR-SIALIGPWAN-------
Bgl1    ------RNVQGNHKTNVRAIARDGIVLLKNDANILPLKKPASIAVVGSAAIIGNHARNSP

Fv3A    ----------DTSKLQGGYSGRAPFLHSPLYAAEKLGLDTNVAWGPTLQNSSSHDNWTTN
Bxl1    ----------ATTQMQGNYYGPAPYLISPLEAAKKAGYHVNFELGTEIAGNSTTG--FAK
Bgl1    SCNDKGCDDGALGMGWGSGAVNYPYFVAPYDAINTRASSQGT--QVTLSNTDNTS---SG

Fv3A    AVAAAKKSDYILYFGGLDAS----AAGEDRDRENLDWPESQLTLLQKLSSLGKPLVVIQL
Bxl1    AIAAAKKSDAIIYLGGIDNT----IEQEGADRTDIAWPGNQLDLIKQLSEVGKPLVVLQM
Bgl1    ASAARGKDVAIVFITADSGEGYITVEGNAGDRNNLDPWHNGNALVQAVAGANSNVIVVVH

Fv3A    G-DQVDDTALLKNKKINSILWVNYPGQDGGTAVMDLLTGRKSPAGRLPVTQYP-------
Bxl1    GGQQVDSSSLKSNKKVNSLVWGGYPGQSGGVALFDILSGKRAPAGRLVTTQYP-------
Bgl1    SVGAIILEQILALPQVKAVVWAGLPSQESGNALVDVLWGDVSPSGKLVYTIAKSPNDYNT
```

FIG. 64A

```
Fv3A    --------------SKYTEQIGMTDMDLRPT-KSLPGRTYRWYS-TPVLPYGFGLHYTKF
Bxl1    --------------AEYVHQFPQNDMNLRPDGKSNPGQTYIWYTGKPVYEFGSGLFYTTF
Bgl1    RIVSGGS-------DSFSEGLFIDYKHFDDAN------------ITPRYEFGYGLSYTKF

Fv3A    QAKFKSN--KLTFDIQKLLKG--CSAQYS-------------------------------
Bxl1    KETLASEPKSLKFNTSSILSAPHPGYTYS-------------------------------
Bgl1    NYSRLSVLSTAKS--GPATGAVVP------------------------------------

Fv3A    ----------------------------------------DTCALPPIQVSVKNTGR
Bxl1    ----------------------------------------EQIPVFTFEANIKNSGK
Bgl1    ------------------------------------GGPSDLFQNVATVTVDIANSGQ

Fv3A    ITSDFVSLVFIKS-EVGPKPYPLKTLAAYGRLHDVAPSSTKDISLEWTLDNIARRGENGD
Bxl1    TESPYTAMLFVRTSNAGPAPYPNKWLVGFDRLADIKPGHSSKLSIPIPVSALARVDSHGN
Bgl1    VTGAEVAQLYITYPSSAPR-TPPKQLRGFAKLN-LTPGQSGTATFNIRRRDLSYWDTASQ

Fv3A    LVVYPGTYTLLLDEPTQAKIQVTLTGKKAILDKWPQDPKSA-------
Bxl1    RIVYPGKYELALNTDESVKLEFELVGEEVTIENWPLEEQQIKDATPDA
Bgl1    KWVVPSGSFGISVGASSRDIRLTSTLSVA--------------------
```

FIG. 64B

```
Fv30B.pro MNPLSLGLAALSLLGYVGVNFVAAFPTDSNSGSEVLISVNGHVKHQELDGFGASQAFQRAEDILGKDGLSKEGTQHVLDL   80
Fv30A.pro ML-FSLVLPTLA--------FQASLALGDTS-----VTVDTSQKLQVIDGFGVSEAYGHAKQF---QNLGPGPQKEGLDL   63

Fv30B.pro LFSKDIGAGFSILRNGIGSSNSSDKNFMNSIEPFSPGSPGAKPHYVWDGYDSGQLTVAQEAFKRGLKFLYGDAWSAPGYM  160
Fv30A.pro LFNTTTGAGLSIIRNKIGCDAS------NSITSTNTDNPDKQAVYHFDGDDDGQ--------------------------  111

Fv30B.pro KTNHDENNGGYLCGVTGAACASGDWKQAYADYLLQWVEFYRKSGVKVTNLGFLNEPQFAAPYAGMLSNGTQAADFIRVLG  240
Fv30A.pro ----SAQSMGRLCGTPGVSCSSGDWRHRYVEMIAEYLSYYKQAGIPVSHVGFLNEGD-GSDF--MLSTAEQAADVIPLLH  184

Fv30B.pro KTIRKRGIHDLTIACCDGEGWDLQ---EDMMAGLTAGPDPAINYLSVVTGHGYVSPPNHPLSTTKKTWLTEWADLTGQFT  317
Fv30A.pro SALQSKGLGDIKMTCCDNIGWKSQMDYTAKLAELEVE-----KYLSVITSHEYSSSPNQPMNTTLPTWMSEGAANDQAFA  259

Fv30B.pro PYTFYNNSGQGEGMTWAGRIQTALVDANVSGFLYWIGAENSTTNS--ALINMIGDKVIPSKRFWAFASFSRFARPGARRI  395
Fv30A.pro T-AWYVNGGSNEGFTWAVKIAQGIVNADLSAYIYWEGVETNNKGSLSHVIDTDGTKFTISSILWAIAHWSRHIRPGAHRL  338

Fv30B.pro EATSSVPLVTVSSFLNTDGTVATQVLNNDTVAHSVQLVVSGTG-----------------------------------  438
Fv30A.pro STSGVVQDTIVGAFENVDGSVVMVLTNSGTAAQTVDLGVSGSSFSTAQAFTSDAEAQMVDTKVTLSDGRVKVTVPVHGVV  418

Fv30B.pro ----------------------RNPHSLKPFLTDNSNDLTALKHLKATGKGSFQTTIPPRSLVSFVT-------------  483
Fv30A.pro TVKLTTAKSSKPVSTAVSAQSAPTPTSVKHTLTHQKTSSTTLSTAKAPTSTQTTSVVESAKAVKYPVPPVASKGSSKSAP  498

Fv30B.pro -------------------------------------DF                                           485
Fv30A.pro KKGTKKTTTKKGSHQSHKAHSATHRRCRHGSYRRGHCTN                                           537
```

FIG. 65

NUEVAS ENZIMAS DE HIDROLASA DE GLUCÓLISIS Y SUS APLICACIONES

REFERENCIA CRUZADA CON APLICACIONES RELACIONADAS

Esta solicitud reclama prioridad para los usos provisionales 61/245.273 de Estados Unidos, presentado el 23 de septiembre de 2009, y 61/289.886, presentado el 23 de diciembre de 2009, cuyas divulgaciones son incorporadas aquí para referencia.

2. CAMPO TÉCNICO

La presente divulgación se relaciona generalmente con enzimas de hidrolasa de glucólisis, las composiciones que comprenden dichas enzimas y los métodos de uso de las enzimas y las composiciones, por ejemplo para la sacarificación o la conversión de materiales celulósicos y hemicelulosicos en azúcares.

3. ANTECEDENTES

La bioconversión de la biomasa lignocelulósica renovable en un azúcar fermentable que se fermenta posteriormente para producir alcohol (p. ej. etanol) como una alternativa a los combustibles líquidos ha atraído la atención acelerada de investigadores desde los años 70, cuando ocurrió la crisis del petróleo debido a que la OPEP disminuyó la producción de petróleo (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). El etanol ha sido ampliamente utilizado como una mezcla del 10% de la gasolina en los E.E.U.U. o como combustible sólo para vehículos en Brasil durante las dos últimas décadas. La importancia del bioetanol como combustible aumentará paralelamente con los precios disparados del petróleo y el agotamiento gradual de sus fuentes. Además, los azúcares fermentables se utilizan cada vez más para producir plásticos, polímeros y otro productos de base biológica y se espera que esta industria se amplíe substancialmente en los próximos años. Así, la demanda por los abundantes azúcares fermentables de bajo costo, que pueden utilizarse como materia prima en lugar de las materias primas derivadas del petróleo, continúa creciendo.

Principalmente entre las materias primas renovables útiles están la celulosa y la hemicelulosa (xilanos), que pueden convertirse en azúcares fermentables. La conversión enzimática de estos polisacáridos en azúcares solubles, por ejemplo

glucosa, xilosa, arabinosa, galactosa, manosa y/o otras hexosas y pentosas, ocurre debido a las acciones combinadas de varias enzimas. Por ejemplo, endo-1,4-β-glucanases (EG) y los exo-cellobiohydrolases (CBH) catalizan la hidrólisis de la celulosa insoluble en los celooligosacáridos (*p. ej.* siendo la celobiosa el producto principal), mientras que las β-glucosidasas (BGL) catalizan la conversión de los oligosacáridas en glucosa. Xilanasas junto con otras proteínas accesorias (ejemplos no limitantes de los cuales se incluyen L-α-arabinofuranosidases, feruloyl y los esterases acetylxylan, glucuronidases, y los β-xilosidasas) catalizan la hidrólisis de hemicelulosas.

Las membranas celulares de las plantas se componen de una mezcla heterogénea de polisacáridos complejos que interactúan a través de medios covalentes y no covalentes. Los polisacáridos complejos de membranas celulares más altas de plantas incluyen, por ejemplo, la celulosa (glucan β-1,4) que constituye generalmente el 35-50% del carbono encontrado en los componentes de la pared de las células. Los polímeros de la celulosa auto asociados a través de la vinculación del hidrógeno, las interacciones de van der Waals y las interacciones hidrofóbicas para formar microfibrils semi-cristalinos de la celulosa. Estas microfibrilas también incluyen las regiones no cristalinas, conocidas generalmente como celulosa amorfa. Las microfibrilas de la celulosa se encajan en una matriz formada por hemicelulosas (incluyendo, por ejemplo, xilanos , arabinosas y manosas), las pectinas (p. ej. galacturonasas y galactanos), y varios otros glucanos β-1,3 y β-1,4. Estos polímeros de la matriz se substituyen a menudo por ejemplo con residuos de arabinosa, galactosa y/o xilosa para rendir arabinoxylans, arabinogalactans, galactomannans y xyloglucans altamente complejos. La matriz de la hemicelulosa, a su vez es rodeada por la lignina polifenólica.

La complejidad de la matriz la hace difícil de degradar por los microorganismos mientras que los componentes de la lignina y de la hemicelulosa deban ser analizados antes de que las enzimas puedan actuar en las microfibrilas de la celulosa de base. Normalmente se requiere un consorcio de diversas actividades enzimáticas para analizar los polímeros de la pared celular para liberar los monosacáridos constitutivos. Para la sacarificación de las membranas celulares de la planta, la lignina debe estar permeabilizada y la hemicelulosa interrumpida para permitir que las enzimas de degradación celulosa puedan actuar en su sustrato.

Sin importar el tipo de materia prima celulósica, el costo y la eficacia hidrolítica de enzimas son los factores importantes que restringen la comercialización de los procesos de la bioconversión de la biomasa. Los costos de producción de enzimas microbianas producidas están conectados fuertemente con la productividad de la cepa productora de enzima y la producción final de la actividad en el caldo de fermentación. La eficacia hidrolítica de un complejo de multienzimas depende tanto de las características de enzimas individuales, como de las sinergias entre ellas y de su proporción en la mezcla del multienzimas .

Existe una necesidad en el arte de identificar la enzima y/o las mezclas/las composiciones enzimáticas que son capaces de convertir la planta y/o otros materiales celulósicos o hemicelulosicos en los azúcares fermentables con eficacia y producción mejoradas.

4. RESUMEN

La divulgación proporciona ciertos polipéptidos de hidrolasa de glucólisis que tienen actividad hemicelulolítica, incluyendo, p. ej., *xilanasas* (endoxylanasas*p. ej.,*), xilosidasas (β-xylosidase*s.g.,*), arabinofuranosidasas (*e.g.,* L-α-arabinofuranosidasas), los ácidos nucléicos que codifican estos polipéptidos y métodos para hacer y usar los polipéptidos y/o los ácidos nucléicos. La divulgación se basa, en parte, en el descubrimiento de las nuevas enzimas y las variantes que tienen xilanasa, β-xilosidasa, y/o las actividades de L-α-arabinofuranosidasa. La divulgación también se basa en la identificación de las mezclas de la enzima (o de las composiciones) que catalizan eficientemente la hidrólisis de materiales celulósicos y hemicelulosicos. Para fines de esta divulgación, una enzima puede definirse como un polipéptido que tiene la actividad enzimática particular o como esa enzima. Por ejemplo, una xilanasa se puede referir como polipéptido que tiene la actividad de la xilanasa o como una enzima xilanasa y un β-xilosidasa se puede referir como un polipéptido que tiene la actividad del β-xilosidasa o enzima del β-xilosidasa.

Las enzimas y/o las mezclas de enzimas/ composiciones de la divulgación pueden utilizarse para producir los azúcares a partir de la biomasa. Los azúcares así producidos pueden ser utilizados por microorganismos para la producción de etanol o se pueden utilizar para producir otros bioproductos para varios usos industriales. Por lo tanto, la divulgación también proporciona usos industriales (p. ej. procesos de

245

sacarificación en la producción del etanol) usando las enzimas y/o mezclas de enzimas /composiciones aquí descritas. Las enzimas y/o mezclas de enzimas /composiciones de la divulgación actual pueden utilizarse para disminuir costos de la enzima en la producción del biocombustible.

En un aspecto, la invención de la divulgación hace referencia a las enzimas (incluidas sus variantes), o a enzimas y/o mezclas de enzimas /composiciones que son útiles para hidrolizar los componentes principales de una biomasa lignocelulósica (incluyendo, p. ej. la celulosa, hemicelulosa, y lignina) o cualquier material que comprenda la celulosa y/o la hemicelulosa. Tal biomasa y/o material lignocelulósico que comprenda celulosa y/o hemicelulosa incluyen, por ejemplo, semillas, granos, tubérculos, desechos de plantas o subproductos de procesador de alimentos o transformación industrial (p. ej. tallos), maíz (incluyendo, mazorcas, rastrojos y similares), hierbas (hierba india, tal como Sorghastrum nutans; o, pasto mijo, *p. ej. especies Panicum* como Panicum virgatum), madera (incluyendo, virutas de madera , desechos procesados), papel, pulpa, papel reciclado (p. ej. papel periódico).

Las mezclas /composiciones de la enzima de la invención se pueden utilizar para hidrolizar la celulosa que comprende una cadena linear de fracciones de glucosa de β-1,4-vinculada, o hemicelulosa, de una estructura compleja que varía de planta en planta.

Las mezclas/composiciones de la enzima de la invención pueden comprender un número de diversas enzimas, incluyendo, p. ej. celulasas y/o hemicelulasas. Por ejemplo, las mezclas /composiciones de la enzima de la invención se pueden utilizar para hidrolizar la biomasa o una materia prima conveniente. Las mezclas /composiciones de la enzima de la invención preferiblemente comprenden las mezclas de enzimas, seleccionadas a partir de, p. ej. xilanasas, xilosidasas, celobiohidrolasas, arabinofuranosidasas, y/o de otras enzimas que puedan digerir la hemicelulosa en azúcares de monómero. Una mezcla/composición de la enzima de la invención puede constar de una mezcla de dos o más, tres o más, o de cuatro o más enzimas seleccionadas de una o más xilanasas, de una o más xilosidasas, de una o más celobiohidrolasas, de una o más arabinofuranosidasas y de una o más otras enzimas capaces de convertir la hemiceluosa en azúcares de monómero. Las otras enzimas que pueden asimilar hemicelulosa en azúcares de monómero incluyen, sin limitación, una

296

celulasa, una hemicelulasa o una composición que comprende una celulasa o una hemicelulasa. Una mezcla/composición de la enzima de la invención puede comprender una mezcla dos o más, tres o más, o de cuatro o más enzimas seleccionadas de una xilanasa, de una xilosidasa, de una celobiohidrolasa, de una arabinofuranosidasa y por lo menos otra enzima capaz de convertir la hemicelulosa en azúcares de monómero. Un ejemplo no limitador de una mezcla/composición de la enzima de la invención comprende una mezcla de una xilanasa, de una xilosidasa, de una celobiohidrolasa, de una arabinofuranosidasa y de una ββ- glucosidasa. La mezcla/composición de la enzima de la invención es convenientemente una que está ocurriendo de manera no natural.

Según se utiliza aquí el término "composición que ocurre naturalmente" hace referencia a una composición producida por una fuente natural, que comprende uno o más componentes o actividades enzimáticos, en donde cada uno de estos componentes o actividades se encuentra en la proporción y nivel producidos por la fuente natural que se encuentra en la naturaleza, sin ser tocada ni modificada por el ser humano. En consecuencia, una composición natural es, por ejemplo, una que es producida por un organismo sin modificaciones con respecto a las enzimas celulolíticas o hemiceluloíticas, de tal manera que la relación o los niveles de las enzimas componentes permanecen inalterados respecto a aquellas producidas por el organismo original en su entorno nativo. Por otro lado, una "composición no natural", se refiere a una composición producida por: (1) combinación de componentes celulolíticas o enzimas hemiceluloiticas ya sea en una relación natural o no natural, *es decir,* alterada, proporción; o (2) modificación de un organismo para expresar, sobreexpresar o subexpresar una o más enzimas endógenas o exógenas; o (3) modificar un organismo tal que al menos una enzima endógena sea eliminada. Una "composición no natural" puede referirse a una composición producida por un organismo que se producen naturalmente y sin modificar, pero cultivada en un medio artificial o el medio ambiente que es diferente del entorno nativo del organismo, de tal forma que difieren de las cantidades o proporciones de peso de enzimas particulares en la composición de las existentes en una composición hecha por un organismo nativo cultivado en su hábitat natural.

La mezcla/composición de la enzima de la invención descrita aquí es, por ejemplo, un caldo de fermentación. El caldo de fermentación puede ser uno de un hongo filamentoso, incluyendo, *p. ej.* un *Trichoderma*, *Humicola*, *Fusarium*, *Aspergillus*,

247

neuroespora, penicillium, Cephalosporium, Achlya, Podospora, Endothia, Mucor, Cochliobolus, Pyricularia, o *Chrysosporium.* Un hongo ejemplar de *Trichoderma* spp. es un reesei de *Trichoderma.* Un hongo ejemplar *del penicillium* spp. es un funiculosum del penicillium . El caldo de fermentación puede ser, por ejemplo, un caldo sin células de fermentación o un caldo de célula completa.

La mezcla/composición de la enzima de la invención aquí descrita es, en otro ejemplo, una composición de la celulasa. La composición de la celulasa es una composición fungicida filamentosa de la celulasa, incluyendo, por ejemplo, una *Trichoderma,* tal como un reesei de *Trichoderma, composición de la celulasa.* La composición de celulasa puede ser producida por hongos filamentosos, por ejemplo, una *Trichoderma,* como un *Trichoderma reesei.*

Por ejemplo, una mezcla/ composición de la enzima de la invención puede incluir (a) una o más enzima(s) de xilanasa donde al menos una de dicha una o más enzima(s) de xilanasa es un Trichoderma reesei Xyn2, un Trichoderma reesei 0 C 9 Xyn3, un AfuXyn2 o un AfuXyn5;(b) una o más - enzima(s) xilosidasa, en el que al menos una de dicha una o más D 0 9 - enzima(s) de xilosidasa es un grupo 1 0D 9 - xilosidasa o un grupo 2 0D 9 - xilosidasa, en el grupo 1 0D 9 - xilosidasa es un Fv3A o un Fv43A y el grupo 2 0D 9 - xilosidasa es un Pf43A, un Fv43D, un Fv39A, un Fv43E, un Fo43A, un Fv43B, un Pa51A, un Gz43A o un Trichoderma reesei C 0 9 Bxl1; (c) una o más L-- enzima(s) arabinofuranosidasa, donde al menos una de dicha una o más L-- arabinofuranosidasa es un Af43A, un Fv43B, un Pf51A, un Pa51A o un Fv5 (d) una o más enzimas de celulasa; y, opcionalmente, (e) uno o más otros componentes. La mezcla/composición de la enzima es convenientemente una que está ocurriendo de manera no natural. Una o más enzima(s) de celulasa de (d) es deseablemente capaz de alcanzar por lo menos 0.00005 productos de la fracción por proteína de magnesio por gramo de la celulosa hinchada del ácido fosfórico (PASC) según lo determinado por un análisis del calcofluor. En un ejemplo no-limitador, el peso combinado de enzimas de xilanasa en la composición puede representar o constituir 5 wt.% a 45 wt.% (p. ej. 5 wt. % a 25 wt.%, 5 wt.% a 15 wt.%, 10 wt.% a 15 wt.%) del peso de la proteína combinada o total en la composición, mientras que el peso combinado de β- enzima de xilosidasa puede representar o constituir 2 wt.% a 50 wt.% (p. ej. 2 wt.% a 30 wt.%, 5 wt.% a 25 wt. %, 5 wt.% a 10 wt.%) de las proteínas totales en la composición, mientras que el peso combinado de las enzimas de L-α– arabinofuranosidasa puede representar o constituir 2

240

wt.% a 50 wt.% (p. ej. 2 wt.% a 30 wt.%, 2 wt.% a 20 wt.%, 5 wt.% a 15 wt.%, 5 wt.% a 10 wt.%) del peso de la proteína combinada o total en la composición, mientras que el peso combinado de enzimas de celulasa puede representar o constituir 30 wt.% a 80 wt. % (p. ej. 40 wt.% a 70 wt.%, 50 wt.% a 60 wt.%) del peso de la proteína combinada o total en la composición. La mezcla/composición de la enzima como se describe aquí es, por ejemplo, una composición de caldo de fermentación. El caldo de fermentación puede ser por ejemplo uno de un hongo filamentoso, incluyendo, sin limitación un *Trichoderma, Humicola, Fusarium, Aspergillus, neuroespora, penicillium, Cephalosporium, Achlya, Podospora, Endothia, Mucor, Cochliobolus, Pyricularia, o Chrysosporium*. Un hongo ejemplar de *Trichoderma* spp. es un reesei de *Trichoderma.* Un hongo ejemplar *del penicillium* spp. es un funiculosum del penicillium . El caldo de fermentación puede ser, por ejemplo, un caldo sin células de fermentación o un caldo de célula completa. La mezcla/composición de la enzima según lo descrito aquí puede también ser una composición de celulasa, por ejemplo, una composición de celulasa fungicida filamentosa. La composición de celulasa puede ser producida por hongos filamentosos, por ejemplo, una Trichoderma, como un Trichoderma reesei.

Por ejemplo, una mezcla/composición de la enzima de la invención puede comprender (a) una o más enzimas de xilanasa donde por lo menos una de dicha unas o más enzimas de xilanasa es un reesei *Xyn2 de* Trichoderma, un reesei *Xyn3 de* Trichoderma, un AfuXyn2, o un AfuXyn5; (b) uno o ambos grupos 1 β- enzimas de xilosidasa: Fv3A y Fv43A; (c) uno o más del grupo 2 β- enzimas de xilosidasa seleccionadas de Pf43A, de un Fv43D, de un Fv39A, de un Fv43E, de un Fo43A, de un Fv43B, de un Pa51A, de un Gz43A, y/o de un reesei *Bxl1 de* Trichoderma; (d) una o más enzimas de celulasa; y opcionalmente (e) uno o más otros componentes. Una o más enzima(s) de celulasa de (d) es deseablemente capaz de alcanzar por lo menos 0.00005 productos de la fracción por proteína de magnesio por gramo de la celulosa hinchada del ácido fosfórico (PASC) según lo determinado por un análisis del calcofluor. La mezcla/composición de la enzima es convenientemente una que está ocurriendo de manera no natural. En un ejemplo no-limitador, el peso combinado de enzimas de xilanasa en la composición puede representar o constituir 5 wt.% a 45 wt.% (p. ej. 5 wt. % a 25 wt.%, 5 wt.% a 15 wt.%, 10 wt.% a 15 wt.%) del peso de la proteína combinada o total en la composición, mientras que el peso combinado del grupo 1 β- enzima de xiilosidasa puede representar o constituir 2 wt.% a 50 wt.% (p. ej. 2 wt.% a 30 wt.%, 5

wt.% a 25 wt.%, 5 wt.% a 10 wt.%) de las proteínas totales en la composición, mientras que el peso combinado del grupo 2 β las enzimas de xilosidasa puede representar o constituir 2 wt.% a 50 wt.% (p. ej. 2 wt.% a 30 wt.%, , 5 wt.% a 25 wt.%, 5 wt.% a 10 wt. %) del peso de la proteína combinada o total en la composición, mientras que el peso combinado de enzimas de celulasa puede representar o constituir 30 wt.% a 80 wt.% (p. ej. 40 wt.% a 70 wt.%, 50 wt.% a 60 wt.%) del peso de la proteína combinada o total en la composición. La relación entre el peso del grupo 1 β- enzimas de xilosidasa y el peso del grupo 2 β- enzimas de xilosidasa puede ser, por ejemplo, 1:10 a 10:1 (p. ej. 1:8 a 8:1, 1:6 a 6:1, 1:4 a 4:1, 1:2 a 2:1, o a 1:1). La mezcla/composición de la enzima puede comprender más componentes adicionales, que pueden ser proteínas accesorias u otra proteína/componentes sin proteínas. Los componentes adicionales pueden constituir, por ejemplo, 1 wt.% a 50 wt.%, 1 wt.% a 10 wt.%, 2 wt.% a 5 wt.%, 5 wt.% a 10 wt.%, o 5 wt.% a 20 wt.% del peso total de proteínas en la composición. La mezcla/composición de la enzima como se describe aquí es, por ejemplo, una composición de caldo de fermentación. El caldo de fermentación puede ser por ejemplo uno de un hongo filamentoso, incluyendo, sin limitación un Trichoderma, Humicola, Fusarium, Aspergillus, neuroespora, penicillium, Cephalosporium, Achlya, Podospora, Endothia, Mucor, Cochliobolus, Pyricularia, o Chrysosporium. Un hongo ejemplar de *Trichoderma* spp. es un reesei de *Trichoderma.* Un hongo ejemplar *del penicillium* spp. es un funiculosum del penicillium . El caldo de fermentación puede ser, por ejemplo, un caldo sin células de fermentación o un caldo de célula completa. La mezcla/composición de la enzima según lo descrito aquí puede también ser una composición de celulasa, por ejemplo, una composición de celulasa fungicida filamentosa. La composición de celulasa puede ser producida por hongos filamentosos, por ejemplo, una Trichoderma, como un Trichoderma reesei.

En otros ejemplos, una mezcla/ composición de la enzima de la invención puede incluir (a) una o más enzima(s) de xilanasa donde al menos una de dicha una o más enzima(s) de xilanasa es un Trichoderma reesei Xyn2, un Trichoderma reesei Xyn3, un AfuXyn2 o un AfuXyn5;(b) una o más - enzima(s) xilosidasa, en el que al menos una de dicha una o más - enzima(s) de xilosidasa es un grupo 1 - enzima xilosidasa Fv3A o un Fv43A o un grupo 2 - enzima xilosidasa Pf43A, un Fv43D, un Fv39A, un Fv43E, un Fo43A, un Fv43B, un Pa51A, un Gz43A y/o un Trichoderma reesei Bxl1; (c) una o más L-- enzima(s) arabinofuranosidasa, donde al menos una de dicha una o más enzima L- -

250

arabinofuranosidasa es un Af43A, un Fv43B, un Pf51A, un Pa51A o un Fv5 (d) una o más enzimas de glucosidasa; y, opcionalmente, (e) uno o más otros componentes. La mezcla/composición de la enzima es convenientemente una que está ocurriendo de manera no natural. En un ejemplo no-limitador, el peso combinado de enzimas de xilanasa en la composición puede representar o constituir 5 wt.% a 45 wt.% (p. ej. 5 wt. % a 25 wt.%, 5 wt.% a 15 wt.%, 10 wt.% a 15 wt.%) del peso de la proteína combinada o total en la composición, mientras que el peso combinado de la enzima de xiilosidasa puede representar o constituir 2 wt.% a 50 wt.% (p. ej. 2 wt.% a 30 wt.%, 5 wt.% a 25 wt. %, 5 wt.% a 10 wt.%) de las proteínas totales en la composición, mientras que el peso combinado de las enzimas L- - arabinofuranosidasa pueden representar o constituir 2 wt.% a 50 wt.% (p. ej. 2 wt.% a 30 wt.%, , 5 wt.% a 25 wt.%, 5 wt.% a 10 wt.%) del peso de la proteína combinada o total en la composición, mientras que el peso combinado de enzimas de glucosidasa pueden representar o constituir 2 wt.% a 50 wt.% (p. ej. hasta 50 wt.%, 2 wt.% a 10 wt.%, o 3 wt.% a 80 wt.%) del peso de la proteína combinada o total en la composición. La mezcla/composición de la enzima puede comprender además componentes adicionales, que pueden ser proteínas accesorias u otra proteína/componentes sin proteínas. Los componentes adicionales pueden constituir, por ejemplo, 1 wt.% a 50 wt.%, 1 wt.% a 10 wt.%, 2 wt.% a 5 wt.%, 5 wt.% a 10 wt.%, o 5 wt.% a 20 wt.% del peso total de proteínas en la composición. La mezcla/composición de la enzima como se describe aquí es, por ejemplo, una composición de caldo de fermentación. El caldo de fermentación puede ser por ejemplo uno de un hongo filamentoso, incluyendo, sin limitación un Trichoderma, Humicola, Fusarium, Aspergillus, neuroespora, penicillium, Cephalosporium, Achlya, Podospora, Endothia, Mucor, Cochliobolus, Pyricularia, o Chrysosporium. Un hongo ejemplar de *Trichoderma* spp. es un reesei de *Trichoderma*. Un hongo ejemplar *del penicillium* spp. es un funiculosum del penicillium . El caldo de fermentación puede ser, por ejemplo, un caldo sin células de fermentación o un caldo de célula completa. La mezcla/composición de la enzima según lo descrito aquí puede también ser una composición de celulasa, por ejemplo, una composición de celulasa fungicida filamentosa. La composición de celulasa puede ser producida por hongos filamentosos, por ejemplo, una Trichoderma, como un Trichoderma reesei.

Por ejemplo, una mezcla/composición de la enzima de la invención puede comprender también (a) una o más enzimas de xilanasa donde por lo menos una de dicha una o más enzimas de xilanasa es un Xyn2 Trichoderma reesei *un* Xyn3

Trichoderma reesei, un AfuXyn2, o un AfuXyn5; (b) uno o ambos grupos 1 β- enzimas de xilosidasa: Fv3A y Fv43A; (c) uno o más del grupo 2 - enzimas de xilosidasa seleccionadas de Pf43A, de un Fv43D, de un Fv39A, de un Fv43E, de un Fo43A, de un Fv43B, de un Pa51A, de un Gz43A, y/o de un Trichoderma reesei Bxl1; y (d) una o más enzimas de glucosidasa; y opcionalmente (e) uno o más otros componentes. La mezcla/composición de la enzima es convenientemente una que está ocurriendo de manera no natural. En un ejemplo no-limitador, el peso combinado de enzimas de xilanasa constituye un 5 wt.% a 45 wt.% (p. ej. 5 wt.% a 25 wt.%, 5 wt.% a 15 wt.%, 10 wt.% a 15 wt.%) del peso de la proteína combinada o total en la composición, mientras que el peso combinado del grupo 1 - enzima de xiilosidasa constituye 2 wt.% a 50 wt.% (p. ej. 2 wt.% a 30 wt.%, 5 wt.% a 25 wt.%, 5 wt.% a 10 wt.%) del peso total de las proteínas en la composición, mientras que el peso combinado del grupo 2 las enzimas de xilosidasa puede representar o constituir 2 wt.% a 50 wt.% (p. ej. 2 wt.% a 30 wt.%, , 5 wt.% a 25 wt.%, 5 wt.% a 10 wt.%) del peso total de la proteína en la composición, mientras que el peso combinado de enzimas de glucosidasa constituye el 2 wt.% a 50 wt.% (p. ej. 2 wt.% a 30 wt.%, 5 wt.% a 25 wt.%, 5 wt.% a 10 wt.%) del peso total de las proteína en la composición. La relación entre el peso del grupo 1 - enzimas de xilosidasa y el peso del grupo 2 - enzimas de xilosidasa puede ser, por ejemplo, 1:10 a 10:1 p. ej.emplo 1:8 a 8:1, 1:6 a 6:1, 1:4 a 4:1, 1:2 a 2:1, o a 1:1. La mezcla/composición de la enzima puede comprender además componentes adicionales, que pueden ser proteínas accesorias u otra proteína/componentes sin proteínas. Los componentes adicionales pueden constituir, por ejemplo, 1 wt.% a 50 wt.%, 1 wt.% a 10 wt.%, 2 wt.% a 5 wt.%, 5 wt.% a 10 wt.%, o 5 wt.% a 20 wt.% del peso total de proteínas en la composición. La mezcla/composición de la enzima como se describe aquí es, por ejemplo, una composición de caldo de fermentación. El caldo de fermentación puede ser por ejemplo uno de un hongo filamentoso, incluyendo, sin limitación un Trichoderma, Humicola, Fusarium, Aspergillus, neuroespora, penicillium, Cephalosporium, Achlya, Podospora, Endothia, Mucor, Cochliobolus, Pyricularia, o Chrysosporium. Un hongo ejemplar de *Trichoderma* spp. es un Trichoderma reesei. Un hongo ejemplar *del penicillium* spp. es un funiculosum del penicillium . El caldo de fermentación puede ser, por ejemplo, un caldo sin células de fermentación o un caldo de célula completa. La mezcla/composición de la enzima según lo descrito aquí puede también ser una composición de celulasa, por ejemplo, una composición de celulasa fungicida

152

filamentosa. La composición de celulasa puede ser producida por hongos filamentosos, por ejemplo, una Trichoderma.

Por ejemplo, una mezcla/ composición de la enzima de la invención puede incluir (a) una o más enzima(s) de xilanasa donde al menos una de dicha una o más enzima(s) de xilanasa es un Trichoderma reesei Xyn2, un Trichoderma reesei Xyn3, un AfuXyn2 o un AfuXyn5;(b) una o más - enzima(s) xilosidasa, en el que al menos una de dicha una o más - enzima(s) de xilosidasa es un grupo 1 - xilosidasa o un grupo 2 - xilosidasa, donde en el grupo 1 - xilosidasa puede ser un Fv3A o un Fv43A y el grupo 2 - xilosidasa puede ser un Pf43A, un Fv43D, un Fv39A, un Fv43E, un Fo43A, un Fv43B, un Pa51A, un Gz43A o un Trichoderma reesei Bxl1; (c) una o más enzimas L--arabinofuranosidasa, donde al menos una de dicha una o más enzimas L--arabinofuranosidasa es un Af43A, un Fv43B, un Pf51A, un Pa51A o un Fv5; y opcionalmente (d) una o más de otros componentes. La mezcla/composición de la enzima es convenientemente una que está ocurriendo de manera no natural. En un ejemplo no-limitador, el peso combinado de las enzimas de xilanasa constituye 5 wt.% a 45 wt.% (p. ej. 5 wt.% a 25 wt.%, 5 wt.% a 15 wt.%, 10 wt.% a 15 wt.%) del peso total de la proteína en la composición, mientras que el peso combinado de la β- enzima de xilosidasa constituye 2 wt.% a 50 wt.% (p. ej. 2 wt.% a 30 wt.%, 5 wt.% a 25 wt.%, 5 wt. % a 10 wt.%) del peso total de la proteína en la composición, mientras que el peso combinado de las enzimas α- L - arabinofuranosidasa constituye un 2 wt.% a 50 wt.% (p. ej. 2 wt.% a 30 wt.%, 5 wt.% a 25 wt.%, 5 wt.% a 10 wt.%) del peso total de la proteína en la composición. La mezcla/composición de la enzima puede comprender además componentes adicionales, que pueden ser proteínas accesorias u otra proteína/componentes sin proteínas. Los componentes adicionales pueden constituir, por ejemplo, 1 wt.% a 50 wt.%, 1 wt.% a 10 wt.%, 2 wt.% a 5 wt.%, 5 wt.% a 10 wt.%, o 5 wt.% a 20 wt.% del peso total de proteínas en la composición. La mezcla/composición de la enzima como se describe aquí es, por ejemplo, una composición de caldo de fermentación. El caldo de fermentación puede ser por ejemplo uno de un hongo filamentoso, incluyendo, sin limitación un Trichoderma, Humicola, Fusarium, Aspergillus, neuroespora, penicillium, Cephalosporium, Achlya, Podospora, Endothia, Mucor, Cochliobolus, Pyricularia, o Chrysosporium. Un hongo ejemplar de Trichoderma spp. es un Trichoderma reesei. Un hongo ejemplar *del penicillium* spp. es un funiculosum del penicillium . El caldo de fermentación puede ser, por ejemplo, un caldo sin células de fermentación o un caldo de célula completa. La mezcla/composición de la enzima

según lo descrito aquí puede también ser una composición de celulasa, por ejemplo, una composición de celulasa fungicida filamentosa. La composición de celulasa puede ser producida por hongos filamentosos, tales como una Trichoderma.

Una mezcla/composición de la enzima de la invención también puede ser una que comprenda (a) una o más enzimas de xilanasa donde por lo menos una de dicha una o más enzimas de xilanasa es un Trichoderma reesei Xyn2 *un* Trichoderma reesei Xyn3 , un AfuXyn2, o un AfuXyn5; (b) uno o ambos grupos 1 β- enzimas de xilosidasa: Fv3A y Fv43A; (c) uno o más del grupo 2 β- enzimas de xilosidasa seleccionadas de Pf43A, de un Fv43D, de un Fv39A, de un Fv43E, de un Fo43A, de un Fv43B, de un Pa51A, de un Gz43A, y/o de un Trichoderma reesei *Bxl1* ; y opcionalmente (d) uno o más de otros componentes. La mezcla/composición de la enzima es convenientemente una que está ocurriendo de manera no natural. En un ejemplo no-limitador, el peso combinado de las enzimas de xilanasa puede constituir 5 wt.% a 45 wt.% (p. ej. 5 wt.% a 25 wt.%, 5 wt.% a 15 wt.%, 10 wt.% a 15 wt.%) del peso total de la proteína en la composición, mientras que el peso combinado del grupo 1 la β- enzima de xilosidasa constituye 2 wt.% a 50 wt.% (p. ej. 2 wt.% a 30 wt.%, 5 wt.% a 25 wt.%, 5 wt.% a 10 wt. %) del peso total de la proteína en la composición, mientras que el peso combinado del grupo 2 de las enzimas β- xilosidasa puede constituir un 2 wt.% a 50 wt.% (p. ej. 2 wt.% a 30 wt.%, 5 wt.% a 25 wt.%, 5 wt.% a 10 wt.%) del peso total de la proteína en la composición. La relación entre el peso del grupo 1 - enzimas de xilosidasa y el peso del grupo 2 - enzimas de xilosidasa puede ser, por ejemplo, 1:10 a 10:1 p. ej.emplo 1:8 a 8:1, 1:6 a 6:1, 1:4 a 4:1, 1:2 a 2:1, o a 1:1. La mezcla/composición de la enzima puede comprender además componentes adicionales, que pueden ser proteínas accesorias u otra proteína/componentes sin proteínas. Los componentes adicionales pueden constituir, por ejemplo, 1 wt.% a 50 wt.%, 1 wt.% a 10 wt.%, 2 wt.% a 5 wt.%, 5 wt.% a 10 wt.%, o 5 wt.% a 20 wt.% del peso total de proteínas en la composición. La mezcla/composición de la enzima como se describe aquí es, por ejemplo, una composición de caldo de fermentación. El caldo de fermentación puede ser por ejemplo uno de un hongo filamentoso, incluyendo, sin limitación un Trichoderma, Humicola, Fusarium, Aspergillus, neuroespora, penicillium, Cephalosporium, Achlya, Podospora, Endothia, Mucor, Cochliobolus, Pyricularia, o Chrysosporium. Un hongo ejemplar de Trichoderma spp. es un Trichoderma reesei. Un hongo ejemplar *del penicillium* spp. es un funiculosum del penicillium . El caldo de fermentación puede ser, por ejemplo, un

caldo sin células de fermentación o un caldo de célula completa. La mezcla/composición de la enzima según lo descrito aquí puede también ser una composición de celulasa, por ejemplo, una composición de celulasa fungicida filamentosa. La composición de celulasa puede ser producida por hongos filamentosos, tales como una Trichoderma.

Las enzimas, mezcla/composiciones de enzimas de la divulgación pueden utilizarse en el sector alimenticio, *e.g.,* para hornear, para procesar frutas y verduras, en la subdivisión de desechos agrícolas, en la fabricación de alimentos para animales, en la producción de la pulpa y papel, en la fabricación del textiles, o en agentes de limpieza industriales o para el hogar. Las enzimas, y las enzimas en las mezclas / composiciones de enzimas de la divulgación son, por ejemplo, producidas cada una independientemente por un microorganismo, *p. ej.* por hongos o bacterias.

Las enzimas, mezcla/composiciones de enzimas de la divulgación pueden utilizarse también como enzimas o composiciones comerciales para digerir la lignocellulosa de cualquier fuente conveniente, incluyendo todas las fuentes biológicas, tales como biomasas de las plantas, p. ej. maíz , granos, hierbas (hierba india , tal como Sorghastrum nutans; o, mijo, *p. ej. especies Panicum como Panicum* virgatum), o, las maderas o subproductos para procesar madera, *p. ej.* en el proceso de madera, pulpa y/o la industria de papel, en la fabricación del textiles, en el hogar y agentes de limpieza industriales, y/o en el proceso de desechos de biomasa.

En otro aspecto, la divulgación proporciona ácidos nucleicos aislados, sintéticos o recombinantes que tienen el cerca de 70%, por ejemplo, por lo menos cerca del 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%, o el total (100%) de identidades de la secuencia a una secuencia del ácido nucleico del 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, o 50, sobre una región de por lo menos cerca de 10, *p. ej.* por lo menos cerca de 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, o 2000, residuos. Relativamente, la divulgación proporciona ácidos nucleicos aislados, sintéticos, o recombinados capaces de hibridación, bajo condiciones altamente

estrictas , a un complemento del 97%, 98%, 99%, o total (100%) de identidades de secuencia a una secuencia del ácido nucleic de la 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, o 50, o a un fragmento de eso. El fragmento, por ejemplo, puede ser de por lo menos 150 residuos contiguos en longitud, por ejemplo, por lo menos 200, 250, o 300 residuos contiguos en longitud. La divulgación actual proporciona los ácidos nucleicos que codifican un polipéptido que tiene actividad hemicelulolitica. La actividad hemicelolitica ejemplar incluye, sin limitación, xilanasa, β-xilosidasa, y/o la actividad de L-α-arabinofuranosidasa. Los polipéptidos ejemplares que tienen actividad hemicelulolitica incluyen, sin limitación, una xilanasa, una β-xilosidasa, y/o una L-α- arabinofuranosidasa.

La divulgación además proporciona ácidos nucleicos aislados, sintéticos, o recombinados que codifican una enzima de la divulgación, incluyendo un polipéptido que comprende la secuencia del aminoácido cuyo de Secuencia de Identificación Número (SEQ ID NO) es: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 43, o 44, o una subsecuencia de ello (*p. ej.* un dominio catalítico ("CD" por su sigla en inglés) o módulo vinculante del carbohidrato ("CBM" por su sigla en inglés), o una variante apropiada de eso. En algunas realizaciones, un ácido nucleico de la divulgación codifica la porción madura de una proteína del aminoácido con 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, o 44, que es operable opcionalmente unido a una secuencia heteróloga de la señal, *p. ej.* la secuencia de la señal de Trichoderma reesei *CBHI* . El ácido nucleico codifica deseablemente un polipéptido que tiene actividad hemicelulolítica, p. ej. la xilanasa , la β-xilosidasa y/ola actividad de L-α-arabinofuranosidasa. El ácido nucleico de la divulgación codifica una hemicelulasa, por ejemplo, una xilanasa, una β-xilosidasa, y/o una L-α- arabinofuranosidasa , o una variante conveniente de eso. Otros ácidos nucleicos de la divulgación se describen en la siguiente sección 6.2.

La divulgación proporciona además casetes de expresión que comprenden un ácido nucleico de la divulgación o una subsecuencia del mismo. Por ejemplo, el ácido nucleico comprende al menos 70%, *p. ej.,* al menos alrededor del 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% de identidad de secuencia con una secuencia de ácido nucleico de SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 3941, 46, 47, 48, 49 o 50, en una región de por lo menos unos 10

256

residuos, *p. ej.,* al menos unos 10, 20, 30, 40. 50, 75, 90, 100, 150, 200, 250, 300, 350, 400 o 500 residuos. En otro ejemplo, el ácido nucleico codifica un polipéptido de 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, o 44, en el cual el ácido nucleico está unido operativa y opcionalmente a un promotor. Puede ser el promotor, *p. ej.,* un fungicida, viral, bacteriano, mamífero o planta promotora. El promotor puede ser un promotor constitutivo o inducible. Un promotor ejemplar adecuado se expresa en hongos filamentosos, por ejemplo, Trichoderma reesei. Un promotor adecuado puede derivarse de un hongos filamentoso, *p. ej., Trichoderma reesei,* p. ej., una celobiohidrolasa I ("*cbh1*") promotor del gen de *Trichoderma reesei.*

La divulgación proporciona adicionalmente una celda recombinada diseñada para expresar un ácido nucleico de la divulgación o un casete de expresión de la divulgación. Por ejemplo, el ácido nucleico comprende al menos 70%, *p. ej.,* al menos alrededor del 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% de identidad de secuencia a una secuencia de ácido nucleico de SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 3941, 46, 47, 48, 49 o 50, en una región de por lo menos unos 10 residuos, *p. ej.,* al menos unos 15, 20, 25, 50, 75, 100, 150, 200, 300, 400 o 500 residuos nucleótidos. El ácido nucleico puede codificar un polipéptido de 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, o 44, en el cual el ácido nucleico está unido operativa y opcionalmente a un promotor. El casete de la expresión puede comprender el ácido nucleico con al menos 70%, p. ej., al menos alrededor del 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% de identidad de secuencia a una 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 o 50, en una región de por lo menos unos 10 residuos nucleótidos. Por ejemplo, el casete de expresión puede incluir la codificación del ácido nucleico de un polipéptido de 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, o 44, en el cual el ácido nucleico opcional y operativamente está vinculado a un promotor. La célula recombinante preferiblemente es una célula bacteriana recombinante, una célula mamífera recombinante, una célula fungicida recombinante, una célula recombinante de levadura, una célula recombinante del insecto, una célula de alga recombinante, o una célula de planta recombinante. Por ejemplo, la célula recombinante es una célula de hongo filamentosa recombinante, como un *Trichoderma*, *Humicola*, *Fusarium*, *Aspergillus*, *Neurospora*, *Penicillium*, *fue*, *Achlya*, *PodosporaEndothia*, *Mucor*, *Cochliobolus*, *Pyricularia,* o *Chrysosporium* celular.

257

La divulgación proporciona plantas transgénicas que comprenden un ácido nucleico de la divulgación o un casete de expresión de la divulgación.

La divulgación proporciona polipéptidos sintéticos, aislados o recombinantes que comprenden una secuencia de aminoácidos que tienen al menos 80%, p. ej., al menos alrededor del 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% o total identidad de secuencia (100%) con un polipéptido de SEQ ID NO: 2, 4, 6, 810, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 43 y 44, en una región de por lo menos unos 10, por ejemplo, al menos de 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, o 350 residuos, o sobre el polipéptido inmaduro de longitud total, el polipéptido maduro de longitud total, el CD de longitud total o la longitud total de CBM. Los polipéptidos ejemplares incluyen fragmentos de por lo menos cerca de 10, por ejemplo, por lo menos cerca de 15, 20, 25, 30, 35, 40, 45, 50, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600 residuos en longitud. En realizaciones específicas, los fragmentos comprenden un CD y/o un CBM. Donde un fragmento comprende un CD y un CBM de una enzima de la divulgación, el fragmento incluye opcionalmente un vinculante que separa el CD y el CBM. El vinculador puede ser un enlazador nativo o un enlazador heterólogo. En ciertas realizaciones, los polipéptidos de la divulgación tienen una o más actividades de hemiculasa. Los Polipéptidos o secuencias de péptidos de la divulgación incluyen secuencias codificadas por los ácidos nucleicos de la divulgación. Los polipéptidos ejemplares se describen en la sección 6.1.

La divulgación proporciona además una proteína quimérica o fusión que comprende al menos un dominio de un polipéptido de la divulgación (*por ejemplo,* el CD, el CBM o ambos). Al menos un dominio puede ser vinculado operativamente a una segunda secuencia de aminoácidos, *p. ej.,* una señal de secuencia péptida.

Inversamente, la divulgación proporciona una proteína quimérica o de fusión que comprende una secuencia de señal de un polipéptido de la divulgación vinculado operativamente a una segunda secuencia, *p. ej.,* codificación de la secuencia de aminoácidos de un polipéptido heterólogo que no está naturalmente asociado a la secuencia de señal. En consecuencia, la divulgación proporciona un polipéptido recombinante compuesto por residuos 1 a 13, de 1 a 14, 1 a 15, 1 a 16, 1 a 17, 1 a 18, 1 a 19, 1 a 20, 1 a 21, 1 al 22, 1 a 23, 1 a 24, 1 a 25, 1 a 26, 1 a 27, 1 a 28, 1 a 28, 1 a 30,

258

1 a 31, 1 a 32, 1 a 33, 1 a 34, 1 a 35, 1 a 36, 1 a 37, 1 a 38, o 1 a 40 de un polipéptido de *p. ej.,* 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 o 44. Otros polipéptidos quiméricos o de la fusión ejemplares se describen en la sección 6.1.1.

La divulgación también proporciona métodos para producir un polipéptido recombinante que comprende: (a) una célula de cultivo diseñada para expresar un polipéptido de la divulgación; y (b) recuperando el polipéptido. La recuperación del polipéptido incluye, por ejemplo, la *recuperación del caldo de fermentación que comprende el polipéptido* . En ciertas divulgaciones, la recuperación del polipéptido puede incluir más pasos de purificación.

La divulgación proporciona métodos para hidrolizar, romper o interrumpir una celooligosacarida, un oligómero de arabinoxilano o una composición que comprende un compuesto de glucano o celulosa que comprende entrar en contacto con la composición con una enzima, la mezcla y composición de enzima de la divulgación en condiciones adecuadas, donde la enzima o la mezcla/composición de la enzima hidroliza, rompe o interrumpe la celooligosacarida, arabinoxilano oligómero o la composición que comprende el glucano o la celulosa.

La divulgación proporciona "mezclas" o composiciones de enzimas (también denominadas aquí "mezcla/composición de enzima"), que comprende un polipéptido de la divulgación o un polipéptido codificado por un ácido nucleico de la divulgación. En algunas realizaciones, el polipéptido del acceso tiene unas o más actividades seleccionadas de xilanasa, de β-xilosidasa y de las actividades de α-L-arabinofuranosidasa. En ciertas realizaciones, las mezclas/composiciones de enzimas se utilizan o son útiles para la despolimerización de polímeros celulósicos y hemicelulosicos para metabolizar las fracciones del carbón. Las mezclas de enzima de la divulgación pueden ser en forma de *composición, por ejemplo,* como un producto de fabricación. La composición puede ser *por ejemplo,* una formulación y puede tomar la forma física de *p. ej.,* de un líquido o de un sólido. En divulgaciones ejemplares, una mezcla y composición de la enzima de la divulgación incluye una celulasa que comprende al menos tres tipos diferentes de enzimas seleccionados de (1) una endoglucanasa, (2) una celobiohydrolasa y (3) una β-glucosidasa; o por lo menos tres diferentes actividades enzimáticas seleccionadas de (1) una actividad endoglucanasa

que cataliza la segmentación de β 1,4 interno, vínculos de materiales celulósicos o hemicelulosicos, resultando en una glucooligosaccharides menor, (2) una actividad celobiohidrolasa que cataliza la segmentación y liberación, de manera "exo", de las unidades de celobiosa (*por ejemplo,* disacárida de glucosa -glucosa β-1,4), y (3) una actividad β-glucosidasa que cataliza la liberación de monómeros de glucosa de una celooligosacaridas corta (*p. ej.* celobiosa). Las mezclas /las composiciones ejemplares de la divulgación se describen en la sección 6.3.4. a continuación.

En otro aspecto, la divulgación proporciona métodos para el procesamiento de un material de biomasa que comprende lignocelulosa que comprende contactar una composición que comprende una celulosa y/o azúcares fermentables con un polipéptido de la divulgación o un polipéptido codificado por un ácido nucleico de la divulgación, o una mezcla/composición de enzima (*por ejemplo,* un producto de fabricación) de la divulgación. Material de biomasa adecuada que comprende lignocelulosa puede derivarse de un cultivo agrícola, de un subproducto de un alimento o producción de alimentos, de un producto de residuos lignocelulósicos, un residuo de una planta, o un papel de desecho o un producto de residuos de papel. Los polipéptidos de la divulgación pueden tener una o más actividades enzimáticas seleccionadas de la celulasa, endoglucanasa, celobiohidrolasa, β-glucosidasa, xilanasa, mananasa, β-xilosidasa, arabinofuranosidasa y otras actividades de la hemicelulasa. Los residuos adecuados de la planta pueden constar de grano, semillas, tallos, hojas, cascos, cascarillas, mazorcas, rastrojo, paja, pastos, madera, astillas de madera, pasta de madera y aserrín. Las hierbas pueden ser, *p. ej.,* pasto India o mijo. Las hierbas también pueden ser, por ejemplo, *Miscanthus.* Los residuos de papel puede ser, *p. ej.,* papel desechado o de fotocopias, papel para imprimir, papel de cuadernos, papel de Bloc de notas, papel para la máquina de escribir, periódicos, revistas, cartón y diversos materiales de empaque hechos de papel. También pueden ser los residuos de papel, por ejemplo, la pulpa.

La divulgación proporciona composiciones (incluyendo enzimas, mezclas y composiciones de enzimas, *p. ej.,* productos de fabricación de la divulgación) que comprende una mezcla de enzimas hidrolizadas de hemicelulosa y celulosa y por lo menos un material de biomasa. Opcionalmente, el material de biomasa comprende un material lignocelulósico derivado de un cultivo agrícola. Alternativamente, el material de la biomasa es un subproducto de la producción de un alimento o de comidas. El material adecuado de biomasa también puede ser un producto de residuos lignocelulósicos, un

260

residuo de plantas, residuos de papel o productos de papel de desecho, o comprende un residuo de la planta. Los residuos de la planta pueden constar de granos, semillas, tallos, hojas, cascos, cascarillas, mazorcas, rastrojo, paja, pastos, madera, astillas de madera, pasta de madera o aserrín. Los pasos ejemplares incluyen, sin limitación, pasto indio o mijo. Los pastos ejemplares también pueden incluir *Miscanthus.* Los residuos ejemplares de papel incluyen, sin limitación, papel desechado o de fotocopias, papel para imprimir, papel de cuadernos, papel de Bloc de notas, papel para la máquina de escribir, periódicos, revistas, cartón y diversos materiales de empaque hechos de papel. Los desechos ejemplares de papel también pueden incluir, por ejemplo pulpa.

Toda la información pública disponible en la fecha de presentación, incluyendo, p. ej. publicaciones, patentes, usos de patente, secuencias de GenBank, y depósitos ATCC citados aquí expresamente son incorporados por referencia.

5. BREVE DESCRIPCIÓN DE LAS FIGURAS Y TABLAS

Tablas 1A-1B: La Tabla 1A proporciona un resumen de los identificadores de secuencia utilizados en la presente divulgación para enzimas de hidrolasa de glicólisis; la Tabla 1B proporciona números de adhesión para las enzimas de hidrolasa glicólisis adicionales mencionadas en los ejemplos.

Tabla 2: Actividad de las proteínas expresadas sobre sustratos sintéticos pNPA y pNPX (como se define en la sección 7.1.6. a continuación) en términos relativos a la actividad Quad para eliminar antecedentes del huésped *Trichoderma reesei.* El Quad para eliminar antecedentes del huésped (o "XQuad") se define como la actividad de la(s) proteína(s) expresada en la cepa eliminada Quad dividida por la actividad de la cepa eliminada Quad de antecedentes, sin la proteína(s) expresada. Por ejemplo, un valor de >1 indica que la proteína expresada tiene una actividad mayor que la de los antecedentes.

Tabla 3: Actividad de xilanasa de candidato purificado de endo-xilanasas con xilano de madera de abedul, incubado a 50°C, pH 5.

Tabla 4: Porcentaje de conversión de oligómeros de arabinoxilano de mazorca a productos de monómero basados en el total de azúcar disponible según lo determinado por hidrólisis ácida. Ver el ejemplo 4.

761

Tabla 5: El experimento resulta (según lo descrito en el ejemplo 7) definiendo el nivel de la actividad de la hemicelulasa para hidrolizar la mazorca tratada con azúcares del monómero. La columna titulada "corrida #" indica la orden experimental seleccionada al azar. La columna titulada "ensayo #" indica la orden experimental estándar del diseño. La columna titulada "Quad" indica la fracción de la proteína total que es del cultivo sobrenadante para un crecimiento de la cepa suprimdia *Quad T. reesei*. La columna titulada "Xyn3" indica la fracción de la proteína total que es *Trichoderma reesei* Xyn3. La columna titulada "Fv43D" indica la fracción de la proteína total que es Fv43D. La columna titulada "Fv51A" indica la fracción de la proteína total que es Fv51A. La columna titulada "Fv43A" indica la fracción de la proteína total que es Fv43A. La columna titulada "Fv43B" indica la fracción de la proteína total que es Fv43B. La columna titulada "cargua (carbo ug/mg)" indica la proteína cargada en la reacción de la sacarificación en unidades de microgramos de proteína por miligramo del carbohidrato. Las columnas tituladas "Xyl mg/mL, Glu mg/mL, Arab mg/mL, y G+X+A mg/mL" indican la concentración de xilosa, glucosa, arabinosa y de la combinación de esos tres productos de azúcares detectados al final de la reacción de la sacarificación. Las columnas tituladas "Xyl %theor, Glu %theor y Arab %theor" indican el porcentaje de producción teórica de xilosa, glucosa y arabinosa alcanzado al final de la reacción de la sacarificación.

Tabla 6: Las relaciones calculadas de los siete componentes enzimáticos para la producción máxima predicha de glucosa, xilosa y arabinosa de la hidrólisis de la mazorca. La columna "cargua total de mg/gr carb" indica la dosis total de enzimas en la que se calculan las predicciones. Las filas tituladas "Total mg/ml G + X + A, % rendimiento de glucosa, % de rendimiento de xilosa y % de rendimiento de arabinosa" indican la respuesta que se ha calculado para el óptimo. La columna titulada "datos de r2 apropiado para el modelo (incluye ambas cargas)" indica el parámetro estadístico r al cuadrado para el modelo de ajuste a los datos presentados en la Tabla 5. Las columnas tituladas "fracción Accellerase, sup de fración del Quad, fracción purificada Trichoderma reesei Xyn3, fracción purificada Fv43D, fracción purificada Fv51A, fracción purificada Fv43A y fracción purificada Fv43B" indican la fracción de ese componente que el modelo equipado con los datos en la tabla 5 calcula para ser óptima.

Tabla 7: El refinamiento de las cargas de enzima para la conversión hidrolítica máxima de la mazorca usando mezclas incluyendo las enzimas Fv3A y Fv43D en 1.06

g, 14% de sólidos secos de reacciones pretratadas de la mazorca. Las condiciones para la sacarificación fueron según lo descrito en el ejemplo 7. Las columnas marcadas con nombres de la enzima indican el magnesio de cada una de las enzimas mencionadas por gramo del glucano o xilano utilizado.

Tabla 8: El refinamiento de las cargas de la enzima para la conversión máxima de la mazorca usando las mezclas que contenían Fv51A como la único α-L-arabinofuranosidasa en 1.06 gr, 14% de sólidos secos de reacciones pretratadas de la mazorca. Las condiciones utilizadas para la sacarificación fueron según lo descrito en el ejemplo 7. Las columnas marcadas con nombres de la enzima indican el magnesio de cada una de las enzimas mencionadas por gramo de glucano o xilano utilizado. Las columnas marcadas con los carbohidratos indican el magnesio por mL de cada producto de carbohidrato producido basado en las medidas hechas con cromatografía de exclusión de tamaño. La columna >dp2 incluye todos los oligomeros mayores a un disacárido.

Tabla 9: Las producciones de azúcar obtenidas de las mezclas A, B, C de los hemicelulasas purificadas probadas en 1.06 g, 14% de sólidos secos de reacciones de penetración de la mazorca. Las condiciones utilizadas para la sacarificación fueron según lo descrito en el ejemplo 7. Mezcla A: 6 mg de *Trichoderma reesei* Xyn3, 4 mg Fv3A, 1 mg Fv51A por gramo de xilano. Mezcla B: 6 mg *Trichoderma reesei* Xyn3, Fv43D, 1 mg Fv43D, de 3 mg Fv43A, de 3 mg Fv43B por gramo xilano. Mezcla C: 6 mg de *Trichoderma reesei* Xyn3, 4 mg Fv3A, 1 mg Fv43D, 1 mg Fv51A por gramo de xilano. Las columnas marcadas con los azúcares del monómero indican los % de la producción de cada uno del monómero mencionado.

Tabla 10: La producción de azúcar por el tratamiento de las preparaciones de la hemicelulosa hechas de la mazorca, sorgo, mijo y el bagazo de la caña de azúcar por mezclas A, B, C de hemicelulasa. Las reacciones se corrieron a la escala de 100 -µL en 50 mM pH 5.0 de solución amortiguadora/tampón de acetato del sodio para 6 h a 48°C según lo descrito en el ejemplo 8. Se muestra el % de producción para cada monómero de azúcar .

Tabla 11: Concentración de la mayoría de enzimas expresadas por varias cepas de expresión integrada de *T. reesei* (designadas H3A, 39A, 69A, A10A, G6A, 102, 44A, 11A, G9A) según lo determinado por el porcentaje del área integrada HPLC.

Tabla 12: La lista de los parámetros de pretratamiento del mijo y los resultados de la sacarificación de diferentes tratamientos.

Tabla 13: Los resultados de la sacarificación de la pulpa de madera dura con la composición de la enzima producidos por una cepa integrada *T. reesei* , H3A, en reacciones con diversas cantidades de sólidos, enzimas y tiempo de incubación.

Tabla 14: La sacarificación de la pulpa de madera dura con la composición de la enzima producida por una cepa integrada H3A sobre una gama de temperaturas y de pH.

Se predijeron las secuencias de la señal enumeradas a continuación y en las figuras. Las predicciones fueron hechas con el algoritmo SignalP (*disponible en* http://www.cbs.dtu.dk). Las predicciones del dominio fueron hechas con base en una o más bases de datos Pfam, SMART, o NCBI.

Figuras 1A-1B: Figura 1A: Secuencia nucleotídica Fv3A (SEQ ID NO.: 1). Figura 1B: Secuencia del aminoácido Fv3A (SEQ ID NO.: 2). SEQ. ID. NO.: 2 es la secuencia del Fv3A inmaduro. Fv3A tiene una secuencia predicha de la señal que corresponde a las posiciones 1 a 23 del SEQ ID NO.: 2 (subrayado); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a las posiciones 24 a 766 del SEQ. ID. NO.: 2. El dominio conservado predicho aparece en negrilla.

Figuras 2A-2B: Figura 2A: Secuencia nucleotídica Pf43A (SEQ IDNO.: 3). Figura 2B: Secuencia del aminoácido Pf43A (SEQ ID NO.: 4). SEQ ID NO.: 4 es la secuencia del Pf43A inmaduro. Pf43A tiene una secuencia predicha de la señal que corresponde a las posiciones 1 a 20 del SEQ. ID. NO.: 4 (subrayado); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a las posiciones 21 a 445 del SEQ ID NO.: 4. El dominio conservado predicho se encuentra en negrilla, el módulo vinculante del carbohidrato predicho ("CBM" por su sigla en inglés) aparece en letras mayúsculas y el enlazador predicho que separa el CD y el CBM aparece en itálica.

Figuras 3A-3B: Figura 3A: Secuencia nucleotídica Fv43E (SEQ ID NO.: 5). Figura 3B: Secuencia del aminoácido Fv43E (SEQ ID NO.: 6). SEQ ID NO.: 6 es la secuencia del Fv43E inmaduro. Fv43E tiene una secuencia predicha de la señal que

corresponde a las posiciones 1 a 18 del SEQ ID NO.: 6 (subrayado); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a las posiciones 19 a 530 del SEQ ID NO.: 6. El dominio conservado predicho aparece en negrilla.

Figuras 4A-4B: Figura 4A: Secuencia nucleotídica Fv39A (SEQ ID NO.: 7). Figura 4B: Secuencia del aminoácido Fv39A (SEQ. ID. NO.: 8). SEQ ID NO.: 8 es la secuencia del Fv39A inmaduro. Fv39A tiene una secuencia predicha de la señal que corresponde a las posiciones 1 a 19 del SEQ ID NO.: 8 (subrayado); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a las posiciones 20 a 439 del SEQ ID NO.: 8. El dominio conservado predicho aparece en negrilla.

Figuras 5A-5B: Figura 5A: Secuencia nucleotídica Fv43A (SEQ ID NO.: 9). Figura 5B: Secuencia del aminoácido Fv43A (SEQ. ID. NO.: 10). SEQ. ID. NO.: 10 es la secuencia del Fv43A inmaduro. Fv43A tiene una secuencia predicha de la señal que corresponde a las posiciones 1 a 22 del SEQ ID NO.: 10 (subrayado); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a las posiciones 23 a 449 del SEQ. ID. NO.: 10. El dominio conservado predicho se encuentra en negrilla, el CBM predicho aparece en letras mayúsculas y el enlazador predicho que separa el CD y el CBM aparece en itálica.

Figuras 6A-6B: Figura 6A: Secuencia nucleotídica Fv43B (SEQ. ID. NO.: 11). Figura 6B: Secuencia del aminoácido Fv43B (SEQ ID NO.: 12). SEQ. ID. NO.: 12 es la secuencia del Fv43B inmaduro. Fv43B tiene una secuencia predicha de la señal que corresponde a las posiciones 1 a 16 del SEQ. ID. NO.: 12 (subrayado); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a las posiciones 17 a 574 del SEQ. ID. NO.: 12. El dominio conservado predicho aparece en negrilla.

Figuras 7A-7B: Figura 7A: Secuencia nucleotídica Pa51A (SEQ. ID. NO.: 13). Figura 7B: Secuencia del aminoácido Pa51A (SEQ. ID. NO.: 14). SEQ. ID. NO.: 14 es la secuencia del Pa51A inmaduro. Pa51A tiene una secuencia predicha de la señal que corresponde a las posiciones 1 a 20 del SEQ. ID. NO.: 14 (subrayado); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a las posiciones 21 a 676 del SEQ. ID. NO.: 14. El dominio

conservado α-L-arabinofuranosidasa predicho aparece en negrilla. Para fines de la expresión, el ADN genómico fue un codón optimizado para la expresión en *T. reesei* (véase la figura 60B).

Figuras 8A-8B: Figura 8A: Secuencia nucleotídica Gz43A (SEQ. ID. NO.: 15). Figura 8B: Secuencia del aminoácido Gz43A (SEQ ID NO.:16). SEQ. ID. NO.: 16 es la secuencia del Gz43A inmaduro. Gz43A tiene una secuencia predicha de la señal que corresponde a las posiciones 1 a 18 del SEQ ID NO.: 16 (subrayado); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a las posiciones 19 a 340 del SEQ ID NO.: 16. El dominio conservado predicho aparece en negrilla. Para fines de la expresión, la secuencia de señal predicha Gz43A fue substituida por la secuencia de señal *T. reesei* CBH1 (myrklavisaflatara (SEQ. ID. NO.: 51)) en *T. reesei* (véase el cuadro 61).

Figuras 9A-9B: Figura 9A: Secuencia nucleotídica Fo43A (SEQ ID NO.: 17). Figura 9B: Secuencia del aminoácido Fo43A (SEQ ID NO.: 18). SEQ ID NO.: 18 es la secuencia del Fo43A inmaduro. Fo43A tiene una secuencia predicha de la señal que corresponde a las posiciones 1 a 20 del SEQ ID NO.: 18 (subrayado); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a las posiciones 21 a 348 del SEQ ID NO.: 18. El dominio conservado predicho aparece en negrilla. Para fines de la expresión, la secuencia de señal predicha Fo43A fue substituida por la secuencia de señal T. reesei CBH1 (myrklavisaflatara (SEQ. ID. NO.: 51) (véase Figura 62).

Figuras 10A-10B: Figura 10A: Secuencia nucleotídica Af43A (SEQ ID NO.: 19). Figura 10B: Secuencia del aminoácido Af43A (SEQ. ID. NO.: 20). SEQ ID NO.: 20 es la secuencia del Af43A inmaduro. El dominio conservado predicho aparece en negrilla.

Figuras 11A-11B: Figura 11A: Secuencia nucleotídica Pf51A (SEQ ID NO.:21). Figura 11B: Secuencia del aminoácido Pf51A (SEQ ID NO.: 22). SEQ ID NO.: 22 es la secuencia del Pf51A inmaduro. Pf51A tiene una secuencia predicha de la señal que corresponde a las posiciones 1 a 20 del SEQ ID NO.: 22 (subrayado); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a las posiciones 21 a 642 del SEQ ID NO.: 22. El dominio conservado α-L-arabinofuranosidasa predicho aparece en negrilla. Para fines de

expresión, la secuencia de señal predicha Pf51A fue reemplazada por un codón optimizado la señal de secuencia *T. reesei* CBH1 (myrklavisaflatara (SEQ ID NO: 51)) (subrayado) y la secuencia de nucleótidos Pf51A fue un codón optimizado por la expresión en *T. reesei* (véase Figura 63).

Figuras 12A-12B: Figura 12A: Secuencia nucleotídica AfuXyn2 (SEQ. ID. NO.: 23). Figura 12B: Secuencia del aminoácido AfuXyn2 (SEQ ID NO.:24). SEQ ID NO.: 24 es la secuencia de AfuXyn2 inmaduro. AfuXyn2 tiene una secuencia de señal predicha que corresponde a las posiciones 1 a 18 de SEQ ID NO.: 24 (subrayado); la segmentación de la secuencia de señal se predice para rendir una proteína madura con una secuencia que corresponde a las posiciones 19 a 228 del SEQ ID NO.: 24. El dominio conservado predicho GH11 aparece en negrilla.

Figuras 13A-13B: Figura 13A: Secuencia nucleotídica AfuXyn2 (SEQ ID NO.: 25). Figura 13B: Secuencia del aminoácido AfuXyn2 (SEQ ID NO.:26). SEQ ID NO.: 26 es la secuencia de AfuXyn5 inmaduro. AfuXyn5 tiene una secuencia de señal predicha que corresponde a las posiciones 1 a 19 de SEQ ID NO.: 26 (subrayado); la segmentación de la secuencia de señal se predice para rendir una proteína madura con una secuencia que corresponde a las posiciones 20 a 313 del SEQ ID NO.: 26. El dominio conservado predicho GH11 aparece en negrilla.

Figuras 14A-14B: Figura 14A: Secuencia nucleotídica Fv43D (SEQ ID NO.: 27). Figura 14B: Secuencia del aminoácido Fv43D (SEQ ID NO.: 28). SEQ. ID. NO.: 28 es la secuencia del Fv43D inmaduro. Fv43D tiene una secuencia predicha de la señal que corresponde a las posiciones 1 a 20 del SEQ ID NO.: 28 (subrayado); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a las posiciones 21 a 350 del SEQ ID NO.: 28. El dominio conservado predicho aparece en negrilla.

Figuras 15A-15B: Figura 15A: Secuencia nucleotídica Pf43B (SEQ ID NO.: 29). Figura 15B: Secuencia del aminoácido Pf43B (SEQ ID NO.: 30). SEQ ID NO.: 30 es la secuencia del Pf43B inmaduro. Pf43B tiene una secuencia predicha de la señal que corresponde a las posiciones 1 a 20 del SEQ ID NO.: 30 (subrayado); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a las posiciones 21 a 321 del SEQ ID NO.: 30. El dominio conservado predicho aparece en negrilla.

Figuras 16A-16B: Figura 16A: Secuencia nucleotídica Fv51A (SEQ ID NO.: 31). Figura 16B: Secuencia del aminoácido Fv51A (SEQ ID NO.: 32). SEQ ID NO.: 32 es la secuencia del Fv51A inmaduro. Fv51A tiene una secuencia predicha de la señal que corresponde a las posiciones 1 a 19 del SEQ ID NO.: 32 (subrayado); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a las posiciones 20 a 660 del SEQ ID NO.: 32. El dominio conservado α-L-arabinofuranosidasa predicho aparece en negrilla.

Figuras 17A-17B: Figura 17A: Secuencia nucleotídica Cg51B (SEQ ID NO.:33). Figura 17B: Secuencia del aminoácido Cgf51B (SEQ. ID. NO.: 34). SEQ ID NO.: 34 es la secuencia del Cg51B inmaduro. Cg51B tiene una secuencia predicha de la señal que corresponde a las posiciones 1 a 20 del SEQ ID NO.: 34 (subrayado); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a las posiciones 21 a 670 del SEQ ID NO.: 34. El dominio conservado predicho aparece en negrilla.

Figuras 18A-18B: Figura 18A: Secuencia nucleotídica Fv43C (SEQ ID NO.: 35). Figura 18B: Secuencia del aminoácido Fv43C (SEQ ID NO.: 36). SEQ. ID. NO.: 36 es la secuencia del Fv43C inmaduro. Fv43C tiene una secuencia predicha de la señal que corresponde a las posiciones 1 a 22 del SEQ ID NO.: 36 (subrayado); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a las posiciones 23 a 333 del SEQ ID NO.: 36. El dominio conservado predicho aparece en negrilla.

Figuras 19A-19B: Figura 19A: Secuencia nucleotídica Fv30A (SEQ ID NO.: 37). Figura 19B: Secuencia del aminoácido Fv30A (SEQ ID NO.: 38). SEQ ID NO.: 38 es la secuencia del Fv30A inmaduro. Fv30A tiene una secuencia predicha de la señal que corresponde a las posiciones 1 a 19 del SEQ ID NO.: 38 (subrayado); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a las posiciones 20 a 537 del SEQ ID NO.: 38.

Figuras 20A-20B: Figura 20A: Secuencia nucleotídica Fv43F (SEQ ID NO.: 39). Figura 20B: Secuencia del aminoácido Fv43F (SEQ ID NO.: 40). SEQ ID NO.: 40 es la secuencia del Fv43F inmaduro. Fv43F tiene una secuencia predicha de la señal que corresponde a las posiciones 1 a 20 del SEQ ID NO.: 40 (subrayado); la segmentación

de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a las posiciones 21 a 315 del SEQ ID NO.: 40.

Figuras 21A-21B: Figura 21A: Secuencia *núcleotídica Trichoderma reesei* Xyn3 (SEQ ID NO.: 41). Figura 21B: *Secuencia del aminoácido Trichoderma reesei* Xyn3 (SEQ ID NO.: 42). SEQ ID NO.: 42 es la secuencia de Trichoderma reesei *Xyn3* inmaduro . *Trichoderma reesei Xyn3* tiene una secuencia de señal predicha que corresponde a las posiciones 1 a 16 de SEQ ID NO.:42 (subrayado); la segmentación de la secuencia de señal se predice para rendir una proteína madura con una secuencia que corresponde a las posiciones 17 a 347 del SEQ. ID. NO.: 42. El dominio conservado predicho aparece en negrilla.

Figura 22: Secuencia del aminoácido Trichoderma reesei Xyn2 (SEQ ID NO.: 43). Se subraya la secuencia de la señal. El dominio conservado predicho aparece en negrilla. La secuencia de codificación se puede encontrar en Törrönen *y otros.* Biotecnología, 1992, 10:1461 - 65.

Figura 23: Secuencia del aminoácido Trichoderma reesei Bxl1 (SEQ ID NO.: 44). Se subraya la secuencia de la señal. El dominio conservado predicho aparece en negrilla. La secuencia de codificación se encuentra en Margolles-Clark *y otros.* Appl. Environ. Microbiología. 1996, 62 (10): 3840-46.

Figura 24: Secuencia del aminoácido Trichoderma reesei Bgl1 (SEQ ID NO.: 45). Se subraya la secuencia de la señal. El dominio conservado predicho aparece en negrilla. La secuencia de codificación se encuentra en Barnett *y otros.* Bio-Technology, 1991, 9 (6): 562-567.

Figura 25: Análisis de la actividad de la celulasa usando hidrólisis PASC con detección de calcofluor.

Figura 26: Perfil de elución de la xilanasa.

Figura 27: Detección de SDS-PAGE de la separación de dos etapas de AfuXyn5. Carril 1: Muestra sin procesar; Carril 2: Eluye de la columna de fenilo; Carril 3: Eluye de la columna de GF.

Figura 28: plásmido D/pENTR-TOPO .

Figura 29: pTrex3gM.

Figuras 30A-30B: Funcionamiento de diferentes enzimas en el sustrato de la mazorca. Las barras de error representan los errores experimentales asociados con ensayos de mazorcas, por triplicado. Los números entre paréntesis a lo largo del eje x representan las dosis de la enzima en el magnesio de proteína por g de celulosa.

Figura 31: El funcionamiento de diferentes mezclas/composiciones de enzimas en el sustrato de la mazorca. Barras de error representan los errores experimentales asociados con ensayos de mazorcas, por triplicado. Los números entre paréntesis a lo largo del eje x representan las dosis de enzima en el magnesio de proteína por g de celulosa.

Figura 32: El funcionamiento de diferentes mezclas/composiciones de enzimas en el sustrato de la mazorca. Las barras de error representan los errores experimentales asociados con ensayos de mazorcas, por triplicado. Los números entre paréntesis a lo largo del eje x representan las dosis de enzima en el magnesio de proteína por g de celulosa.

Figura 33: El funcionamiento de diferentes mezclas/composiciones de enzimas en el sustrato de la mazorca. Las barras de error representan los errores experimentales asociados con ensayos de mazorcas, por triplicado. Los números entre paréntesis a lo largo del eje x representan las dosis de enzima en el magnesio de proteína por g de celulosa.

Figura 34: El funcionamiento de diferentes mezclas/composiciones de enzimas en el sustrato de la mazorca. Las barras de error representan los errores experimentales asociados con ensayos de mazorcas, por triplicado. Los números entre paréntesis a lo largo del eje x representan las dosis de enzima en el magnesio de proteína por g de celulosa.

Figuras 35A-35 C: El funcionamiento de diferentes mezclas/composiciones de enzimas en el sustrato de la mazorca. Las barras de error representan los errores experimentales asociados con ensayos de mazorcas, por triplicado. Los números a lo largo del eje x representan las dosis de enzima en el magnesio de proteína por g de celulosa.

210

Figura 36: La región NMR del protón Anomeric de oligomeros arabinoxilano cortos demuestra segmentación por Fv43A más Fv43B.

Figura 37: La región NMR del protón Anomeric de oligomeros arabinoxilan cortos demuestra la segmentación de la xilosa enlazada β-1,2 de la arabinosa por Fv3A.

Figura 38: Alineación entre las secuencias del aminoácido de Trichoderma reesei β-xylosidase *de* y Fv3A.

Figura 39: pENTR-TOPO-Bgl1 (943/942) plásmido.

Figura 40: pTrex3g vector de expresión 943/942 Bgl1.

Figura 41: *pENTR Trichoderma reesei* Xyn3 plásmido.

Figura 42: pTrex3g/*Trichoderma reesei* Xyn3 vector de expresión .

Figura 43: Fv3A pENTR plásmido.

Figura 44: vector de expresión pTrex6g/Fv3A .

Figura 45: TOPO Blunt/Pegl1-Fv43D plasmido.

Figura 45: TOPO Blunt/Pegl1-Fv51A plasmido.

Figuras 47A-47D: Glucano (Figura 47A) y conversiones de xilano (Figura 47B) en monómeros de azúcares de caldos secretados de fermentación de la enzima *T. reesei* cepas de expresión integrada. En la muestra de 3 días se analizó el grado de conversión del glucano y del xilano al monómero y a los productos solubles del oligomero (figura 47C). La figura 47D demuestra una comparación cromatográfica del producto de la enzima a partir de tres cepas de expresión integrada de *T. reesei* En el ejemplo 1 se describen las condiciones experimentales. Las proporciones de las proteínas difieren en los transformantes y pueden ser cuantificadas como un porcentaje del área total pico integrada. El "EGLs" marca el área sumada de los picos de las endoglucanasas. EndoH fue agregado a la muestra de proteína en cantidades pequeñas como reactivo para el análisis del HPLC.

Figuras 48A-48B: La sacarificación creciente de la producción del monómero de xilosa en respuesta a la adición de hemicelulosa en la composición de la enzima

2011

producida por una cepa integrada a los 7 mg totales de proteína por gramo de glucano + el xilano en el amoníaco de la mazorca pretratada. Figura 48A: Las hemicelulosas de los verticillioides Fusarium o *del componente* en la composición de la enzima producido por la cepa integrada. Figura 48B: Hemicelulosas de otros hongos.

Figuras 49A-extensión: La sacarificación creciente en la producción del monómero de glucosa en respuesta a la adición de hemicelulosa en la composición de la enzima producida por una cepa integrada a los 7 mg totales de proteína por gramo de glucano + el xilano en el amoníaco de la mazorca pretratada. Figura 49A: Las hemicelulosas de los verticillioides Fusarium *del componente* en la composición de la enzima producido por la cepa integrada. Figura 49B: Hemicelulosas de otros hongos.

Figuras 50A-50B: La sacarificación creciente en la producción del monómero de arabinosa en respuesta a la adición de hemicelulosa en la composición de la enzima producida por una cepa integrada a los 7 mg totales de proteína por gramo de glucano + el xilano en el amoníaco de la mazorca pretratada. Figura 50A: Las hemicelulosas de los verticillioides Fusarium *del componente* en la composición de la enzima producido por una cepa integrada. Figura 50B: Hemicelulosas de otros hongos.

Figura 51: Una presentación gráfica del funcionamiento de la sacarificación a través de condiciones de pretratamiento. El eje x corresponde a los resultados experimentales que se enumeran en la tabla 12. Los rendimientos se calculan según las cantidades teóricas de glucano o xilano disponible en el mijo crudo. Todos los rendimientos se basan en azúcares monoméricos liberados tras 3 días de sacarificación con el cóctel de la enzima.

Figura 52: Alineación de la secuencia del aminoácido de un número de GH39 β-xilosidasas. Los residuos subrayados en negrilla son el residuo general catalítico predicho de base ácida (marcado con "A" encima de la alineación) y residuo nucleófilo catalítico (marcado con "N" encima de la alineación). Los residuos subrayados en letra normal en las dos secuencias inferiores están dentro de 4 angstroms del sustrato en los sitios activos de las respectivas estructuras 3D (pdb: 1uhv y 2bs9, respectivamente). Los residuos subrayados en la secuencia de Fv39A se predicen que están dentro de 4 angstroms de un sustrato vinculado en el sitio activo.

Figura 53: Alineación de la secuencia del aminoácido de un número GH43 de la familia de las hidrolasas. Los residuos del aminoácido altamente conservados entre los miembros de la familia se muestran subrayados y en negrilla.

Figura 53: Alineación de la secuencia del aminoácido de un número GH51 de la familia de las enzimas. Los residuos del aminoácido altamente conservados entre los miembros de la familia aparecen subrayados y en negrilla.

Figuras 55A-55B: Alineaciones de la secuencia del aminoácido de un número GH10 y GH11 de la familia de las endoxilanasas. Figura 55A; Alineación de GH10 de la familia de las xilanasas. Los residuos subrayados en negrilla son los residuos nucleófilos catalíticos (marcados con "N" por encima de la alineación). Figura 55B; Alineación de GH11 de la familia de las xilanasas. Los residuos subrayados en negrilla son el residuo necleófilo catalítico y los residuos generales de base ácida (marcados con "A" respectivamente encima de la alineación).

Figuras 56A-56B: Sacarificación de mijo de amoníaco diluido pretratado con diversas composiciones y mezclas de enzima ; Figura 56 A: conversión de glucano; Figura 56B: conversión de xilano. Los números debajo de las cifras sobre el eje x se refieren a la cantidad total de mg de cada proteína en una mezcla/composición dada por g de glucano o xilano, tal como se describe en el ejemplo 13.

Figuras 57A-57B: Sacarificación del mijo con amoníaco diluido pretratado con una composición de enzima producido por una cepa integrada H3A; Figura 57A: conversión de glucano; Figura 57B: conversión de xilano. En el ejemplo 14 se describen las condiciones experimentales.

Figuras 58A-58 C: La sacarificación del bagazo de caña de azúcar ampliado por vapor con una composición de la enzima produjo una cepa integrada H3A en diversas dosis de la enzima; Figura 58A: conversión del glucano; Figura 58B: conversión del xilano; Figura 58C: conversiones de 3 días del glucano y del xilano En el ejemplo 17 se describen las condiciones experimentales.

Figuras 59A-59C: Sacarificación de ácido diluido pretratado con fibra de maíz con diferentes enzimas o mezclas de enzimas; Figura 59A: conversión de glucano; Figura 59B: conversión de xilano: Figura 59C: conversiones de 5 días de glucano y xilano. El azúcar ajustada (glucosa o xilosa) reflejó el azúcar producida del paso

enzimático menos los niveles de azúcar iniciales. Las proporciones que se muestran en el eje x representan la dosis de enzimas en mg de proteína total por gramo de celulosa. En el ejemplo 18 se describen las condiciones experimentales.

Figuras 60A-60B: Figura 60A: cADN deducido para Pa51A (SEQ ID NO.: 46). Figura 60B: Codón cADN optimizado para Pa51A (SEQ ID NO.: 47).

Figura 61: Secuencia de codificación para una construcción que abarca una secuencia de la señal CBH1 (subrayada) antes del ADN genómico que codifica un Gz43A maduro (SEQ ID NO.: 48).

Figura 62: Codificación de la secuencia para una construcción que comprende una señal de secuencia CBH1 (subrayado) arriba del ADN genómico codificando un Fo43A maduro (SEQ ID NO: 49).

Figura 63: Secuencia de codificación de codón optimizado para una construcción que comprende una secuencia de señal CBH1 (subrayada) antes del ADN genómico que codifica a Pf51A maduro (SEQ ID NO.: 50).

Figura 64: Alineación de la secuencia del aminoácido de un número GH43 de la familia de las hidrolasas. Los residuos del aminoácido altamente conservados entre los miembros de la familia aparecen subrayados y en negrilla. La caja marca el residuo catalítico predicho con residuos predichos que flanquean estar implicados en la vinculación del sustrato.

Figura 65: Alineación de la secuencia del aminoácido de dos Fusarium representativos GH30 de la familia de las hidrolasa . Los residuos del aminoácido altamente conservados entre los miembros de la familia aparecen subrayados y en negrilla.

6. DESCRIPCIÓN DETALLADA

Las enzimas han sido clasificadas tradicionalmente por la especificidad del sustrato y por la reacción del sustrato. En la era pre-genomica, la función era considerada como la base más susceptible (y quizás más útil) para comparar enzimas y por muchos años se han desarrollado ensayos para diversas actividades enzimáticas, resultando en el esquema de clasificación familiar EC. Las celulasas y otras hidrolasas de glicólisis, que actúan sobre enlaces glucosídicos entre dos fracciones del

carbohidrato (o una mitad del carbohidrato y del no-carbohidrato, como ocurre en los derivados del nitrofenol-glucósido), están designados, bajo este esquema de la clasificación, como EC 3.2.1., donde el número final indica el tipo exacto de enlace de segmentación. Por ejemplo, según este esquema una celulasa endo que actúa (1,4-β-endoglucanase) se designa EC 3.2.1.4.

Con el advenimiento de los proyectos generalizados de la secuencia del genoma, los datos de secuencia han facilitado el análisis y comparación de genes y de proteínas relacionadas. Además, un número creciente de las enzimas capaces de actuar en las fracciones del carbohidrato (es decir, carbohidrasas) se han cristalizado, solucionando sus estructuras tridimensionales. Tales análisis han identificado discretas familias de enzimas con secuencia relacionada, que contienen pliegues tridimensionales conservados que se pueden predecir basados en su secuencia del aminoácido. Además, se ha demostrado que las enzimas con igual o similar pliegues tridimensionales exhiben la misma o similar estereospecificidad de la hidrólisis, aun cuando catalizando diferentes reacciones (Henrissat *y otros.*, 1998, FEBS Lett 425 (2): 352-4; Coutinho y Henrissat, 1999, en genética, bioquímica y ecología de la degradación de la celulosa. T. Kimura. Tokyo, Uni Publishers Co: 15-23.).

Estos hallazgos forman la base de una clasificación basada en secuencia de los módulos de carbohidrasa, que está disponible bajo la forma de base de datos de Internet, el servidor Carbohydrate-Active enZYme (CAZy), disponible en http://afmb.cnrs-mrs.fr/CAZY/index.html (enzimas activas de carbohidrato: una base de datos de enfoque integrado. Veáse Cantarel *y otros.*, 2009, Nucleic Acids Res. 37 (edición de la base de datos): D233-38).

CAZy define cuatro clases importantes de carbohidrasas distinguibles por el tipo de reacción catalizada: Hidrolasas de Glicólisis (GH), Glicosiltransferasas (GT), Liasas del polisacárido (PL) y Esterasas del carbohidrato (CE). Las enzimas de la divulgación son hidrolasas de glicólisis. Las GH son un grupo de enzimas que hidrolizan el enlace glucosídico entre dos o más carbohidratos, o entre un carbohidrato y una fracción del no-carbohidrato. Un sistema de clasificación para las hidrolasas de glicólisis, agrupado por semejanza de la secuencia, ha conducido a la definición de más de 85 familias diferentes. Esta clasificación está disponible en el sitio Web CAZy.

Las enzimas de la divulgación pertenecen, *entre otras cosas*, a las familias de hidrolasas de glicólisis 3, 10, 11, 30, 39, 43, y/o 51.

Las enzimas de la familia 3 ("GH3") de hidrolasas de glicólisis incluyen, p. ej. β-glucosidasa (EC: 3.2.1.21); β-xylosidasa (EC: 3.2.1.37); N-acetilo β-glucosaminidasa (EC: 3.2.1.52); glucano β-1,3-glucosidasa (EC: 3.2.1.58); celodextrinasa (EC: 3.2.1.74); exo-1,3-1,4-glucanasa (EC: 3.2.1); y β-galactosidase (EC 3.2.1.23). Por ejemplo, las enzimas GH3 pueden ser las que tienen β-glucosidasa, β-xylosidasa, β-glucosaminidasa de N-acetilo, glucano β-1,3-glucosidasa, el celodextrinasa, exo-1,3-1,4-glucanasa, y/o actividad de β-galactosidasa. Generalmente, las enzimas GH3 son proteínas globulares y pueden consistir en dos o más subdominios. Un residuo catalítico se ha identificado como un residuo del aspartato que, en β-glucosidasas, está localizado en el tercer terminal-N del péptido y se coloca dentro del fragmento del aminoácido SDW (Li *y otros.* 2001, Biochem. J. 355:835 - 840). La secuencia correspondiente en Bgl1 de *T. reesei* es T266D267W268 (contando desde la metionina en el punto de inicio), donde el residuo catalítico aspartato es D267. La secuencia del hidroxilo/ aspartato también se conserva en las β-xilosidasas GH3 probadas. Por ejemplo, la secuencia correspondiente en *T. reesei* Bxl1 es S310D311 y la secuencia correspondiente en Fv3A es S290D291.

Las enzimas de la familia de hidrolasas de glicólisis 39 ("GH39" de la hidrolasa) tienen actividad α-L-iduronidasa (EC: 3.2.1.76) o β-xilosidasa (EC: 3.2.1.37). La estructura tridimensional de dos β-xilosidases GH39, del saccharolyticum Thermoanaerobacterium *(Uniprot Accession No.* P36906) y *Geobacillus* stearothermophilus (Uniprot Accession No. Q9ZFM2), se han solucionado (véase Yang *y otros.* J. Mol. Biol. 2004, 335 (1): 155-65 y Czjzek *y otros.*, J. Mol. Biol. 2005, 353 (4): 838-46). Las regiones más altamente conservadas en estas enzimas se encuentran en sus secciones de N-terminal, que tienen un pliegue de barril TIM clásico 8 (α/β) con los dos ácidos glutámicos de sitio activo clave situados en los extremos de la terminal c de β-cepas 4 (ácido/base) y 7 (nucleófilo). Residuos de Fv39A E168 y E272 se predijeron para funcionar como catalizador de base ácida y nucleófilo, respectivamente, basado en un alineamiento de secuencia de la citada β-xilosidasas de GH39 de *Thermoanaerobacterium saccharolyticum* y *Geobacilo stearothermophilus* con Fv39A.

Las enzimas de la familia de hidrolasa de glicólisis 43 ("GH43") incluyen, *e.g.*, L-α-arabinofuranosidasa (EC 3.2.1.55); β-xilosidase (EC 3.2.1.37); endo-arabinanasa

276

(EC 3.2.1.99); y/o β-galactosidasa de galactano 1.3 (EC 3.2.1.145). Por ejemplo, las enzimas GH43 pueden tener la actividad L-α-arabinofuranosidasa, actividad de β-xylosidasa, actividad de endo-arabinanasa, y/o actividad de β-galactosidasa de galactano 1.3. Las enzimas de la familia GH43 muestran una estructura como de un impulsor de cinco hélices-β. La estructura de hélice se basa en una repetición de cinco de helices compuestas de cuatro hojas de cepas β. La base general catalítica, un aspartato, el ácido general catalítico, un glutamato, y un aspartato que modula el pKa de la base general fueron identificados a través de la estructura cristalina de *Cellvibrio* japonicus CjAbn43A, y confirmados por mutagénesis dirigida al sitio (véase Nurizzo *y otros.* Nat. Struct. Biol. 2002, 9 (9) 665-8). Los residuos catalíticos se organizan en tres bloques conservados extendidos ampliamente a través de la secuencia del aminoácido (Pons *et al* Proteínas: Structure, Function and Bioinformatics, 2004, 54:424 - 432). Entre la familia GH43 se probaron las enzimas para actividades útiles en la hidrólisis de la biomasa, los residuos catalíticos predichos se muestran como residuos en negrilla y subrayados en las secuencias de la Figura 53. La estructura cristalina de la xilosidasa del Geobacilo stearothermophylus (Brux *y otros.* J. Mol. Bio., 2006, 359:97 - 109) sugiere varios residuos adicionales que puedan ser importantes para el sustrato que vincula esta enzima. Como las enzimas de la familia GH43 probadas para la hidrólisis de la biomasa tenían distintas preferencias del sustrato, estos residuos no se conservan completamente en las secuencias alineadas en la Figura 53. Sin embargo entre las xilosidasas probadas, varios residuos conservados que contribuyen al enlace de sustrato, a través de interacciones hidrofóbicas o enlaces de hidrógeno, se conservan y se caracterizan por un subrayado sencillo en la Figura 53.

Las enzimas de la familia de hidrolasas de glicólisis 51 ("GH51") tienen L-α-arabinofuranosidasa (EC 3.2.1.55) y/o actividad de endoglucanasa (EC 3.2.1.4). La estructura cristalina de alta resolución de un GH51 L-α-arabinofuranosidase de *Geobacilo* stearothermophilus T-6 muestra que la enzima es un hexamer, con cada monómero organizado en dos dominios: los 8 barriles (β/α) y un sandwich de 12 cadenas β con topologia jelly-roll (véase Hövel *y otros.* EMBO J. 2003, 22 (19): 4922-4932). Se puede esperar que los residuos catalíticos sean ácidos y conservados a través de las secuencias de enzimas en la familia. Cuando las secuencias del aminoácido de Fv51A, de Pf51A, y de Pa51A se alinean con las enzimas GH51 de una secuencia más diversa, 8 residuos ácidos permanecen conservados. Esos aparecen en negrita y subrayados en la Figura 54.

Las enzimas de la familia hidrolasa de glicólisis 10 ("GH10") también tienen una estructura de 8 barriles (β/α). Se hidrolizan de manera endo con un mecanismo de retención que utiliza por lo menos un residuo catalítico ácido en un proceso de catálisis generalmente ácido/base (Pell *y otros.*, J. Biol. Chem., 2004, 279 (10): 9597-9605). Las estructuras de cristal de las xilanasas de GH10 de *Penicillium simplicissimum* (Uniprot P56588) y *Thermoascus aurantiacus* (Uniprot P23360) complejas con sustratos en los sitios activos han sido resueltos (véase Schmidt *et al.*. Bioquímica., 1999, 38:2403 - 2412; y Leggio bajo *y otros.* FEBS Lett. 2001, 509: 303-308). *Los residuos TRICHODERMA reesei* Xyn3 importantes para la catálisis y enlace del sustrato pueden derivarse de una alineación con las secuencias de las citado xilanasas GH10 de *Penicillium simplicissimum* y *Thermoascus aurantiacus* (Figura 55A). *El residuo TRICHODERMA reesei* Xyn3 E282 está previsto para ser el residuo nucleófilo catalítico, mientras que los residuos E91, N92, K95, Q97, S98, H128, W132, Q135, N175, E176, Y219, Q252, H254, W312 y/o W320 se predijeron para participar en enlace de sustrato y/o catálisis.

Las enzimas de la familia de hidrolasas de glicólisis 11 ("GH11") tienen una β estructura de jelly-roll. Se hidrolizan de manera endo con un mecanismo de retención que utiliza por lo menos un residuo catalítico ácido en un proceso de catálisis generalmente ácido/base. Varios otros residuos extendidos a lo largo de su estructura pueden contribuir a la estabilización de las unidades de xilosa en el sustrato vecinos de la pareja de monómeros de xilosa que se segmenta por hidrólisis. Se probaron tres endoxilanasas de la familia de GH11 y sus secuencias están alineadas en Figura 55B. E118 (o E86 en maduro *T. reesei* Xyn2) y E209 (o E177 en maduro *T. reesei* Xyn2) han sido identificados como nucleófilo catalítico y residuos base general/ácido *Trichoderma reesei Xyn2* , respectivamente (véase Havukainen *et al.* Biochem., 1996, 35:9617-24).

Las enzimas de la familia de hidrolasas de glicólisis 30 ("GH30") oestán conservando las enzimas que tienen glucosilceramidasa (EC 3.2.1.45); β-1,6-glucanasa (EC 3.2.1.75); β-xylosidasa (EC 3.2.1.37); actividad de β-glucosidasa (3.2.1.21). La primera estructura de cristal de GH30 fue la enfermedad humana de Gaucher β-glucocerebrosidasa resuelto por Grabowski, Gatt y Horowitz (Crit Rev Biochem Mol Biol 1990; 25(6) 385-414). GH30 tiene un pliegue de barril (α/β) TIM_8 con los dos ácidos glutamicos clave activos situados en los extremos del terminal C de las cadenas β 4 (ácido/base) y 7 (nucleofilo) (Henrissat B, y otros. Proc Natl Acad Sci U S A, 92 (15):

208

7090-4, 1995; Jordan et al., Applied Microbiol Biotechnol, 86:1647, 2010). El glutamato 162 de Fv30A se conserva en 14 de las 14 proteínas alineadas GH30 (13 proteínas bacterianas y una endo-b-xilanasa de la accesión de los hongos *Biospora número* ADG62369) y el glutamato 250 de Fv30A se conserva en 10 de los mismos 14, es un aspartato en otros tres y un no ácido en uno. Hay otros residuos ácidos moderadamente conservados pero no hay otros tan extensamente conservados.

6.1 Polipéptidos de la divulgación

La divulgación proporciona polipéptidos sintéticos, aislados o recombinantes que comprenden una secuencia de aminoácidos que tienen al menos 80%, p. ej., al menos alrededor del 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% o total identidad de secuencia (100%) a un polipéptido de SEQ ID NO: 2, 4, 6, 810, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 43 y 44, en una región de por lo menos unos 10, por ejemplo, al menos de 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, o 350 residuos, o sobre la longitud total del polipéptido inmaduro, el polipéptido maduro de longitud total, la longitud total del dominio conservado y/o la longitud total de CBM. El dominio conservado puede ser un dominio catalítico predicho ("CD"). Los polipéptidos ejemplares también incluyen fragmentos de por lo menos cerca de 10, p. ej., por lo menos cerca de 15, 20, 25, 30, 35, 40, 45, 50, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600 residuos en longitud. Los fragmentos pueden incluir un dominio conservado y/o un CBM. Donde un fragmento comprende un domino conservado y un CBM de una enzima, el fragmento incluye opcionalmente un vinculante que separa los dos. El vinculante puede ser un vinculante nativo o un vinculante heterólogo. Se contempla que los polipéptidos de la divulgación pueden ser codificados por una secuencia de ácido nucleico teniendo por lo menos alrededor del 85%, 86%, 87%, 88%, 89%, o 90% de identidad de secuencia a SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 2123, 25, 27, 29, 31, 33, 35, 37, 39 o 41, o por una secuencia de ácido nucleico capaz de hibridizar bajo altas condiciones de rigor a un complemento de SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, o 41 o a un fragmento de lo mismo. Otros ácidos nucleicos ejemplares de la divulgación se describen en la siguiente Sección 6.2.

Los polipéptidos de la divulgación incluyen proteínas con una secuencia de aminoácidos de al menos el 85%, *p. ej.,* al menos el 86%, 87%, 88%, 89%, 90%, 91%,

92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% de identidad de secuencia para por lo menos 50 residuos de aminoácidos contiguos de las secuencias de hidrolasa de glicólisis de secuencias de SEQ ID NO: 2, 4, 6, 8, 1012, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 43, 44 y 45. Por ejemplo, un polipéptido de la divulgación puede incluir secuencias de aminoácidos que tienen al menos el 85%, *p. ej.,* al menos el 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% de identidad de secuencia a por lo menos 10, *p. ej.,* por lo menos 11,12, 13, 14, 15, 20, 25, 30, 40, 50, 75, 100, 125, 150, 175, 200, 250, 300 o 350 residuos de aminoácidos contiguos de las secuencias de hidrolasas de glicólisis de 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, o 45. La secuencia contigua del aminoácido corresponde al dominio conservado y/o el CBM y/o la secuencia de la señal.

Cualquiera de las secuencias de aminoácidos descritas aquí pueden ser producidas conjuntamente o en conjunción con al menos 1, p. ej., al menos 2, 3, 5, 10, o 20 aminoácidos heterólogos flanqueando cada uno de los extremos C y/o terminal N de la secuencia de aminoácido especificado y/o eliminaciones de al menos 1, p. ej., por lo menos 2, 3, 5, 10 o 20 aminoácidos de los extremos C y/o terminal N de una enzima de la divulgación.

Otras variaciones también están dentro del alcance de esta divulgación. Por ejemplo, unos o más residuos del aminoácido pueden modificarse para aumentar o disminuir el pI de una enzima. El cambio del valor del pI puede alcanzarse quitando un residuo del glutamato o substituyéndolo con otro residuo del aminoácido.

La divulgación específicamente proporciona un polipéptido Fv3A, un Pf43A, un Fv43E, un Fv39A, un Fv43A, un Fv43B, un Pa51A, un Gz43A, un Fo43A, un Af43A, un Pf51A, un AfuXyn2, un AfuXyn5, un Fv43D, Pf43B, Fv43B, Fv51A, un Trichoderma reesei Xyn3, un Trichoderma reesei Xyn2, un Trichoderma reesei Bxl1, y/o un Trichoderma reesei Bgl1. Una combinación de una o más de estas enzimas está convenientemente presente en la mezcla/composición de la enzima, por ejemplo, una que esté ocurriendo naturalmente.

Fv3A: La secuencia de aminoácidos de Fv3A (SEQ ID NO: 2) se muestra en las Figuras 1B, 38 y 64. SEQ ID NO.: 2 es la secuencia del Fv3A inmaduro. Fv3A tiene una secuencia predicha de la señal que corresponde a los residuos 1 a 23 del SEQ ID

280

NO.: 2 (subrayado); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a los residuos 24 a 766 del SEQ ID NO.: 2. Los dominios conservados predichos están en negrilla en la Figura 1B. Se demostró que Fv3A tiene una actividad de β-xiosidasa, por ejemplo, en un ensayo enzimático usando *p*- nitrofeniol-β-xilopiranosida, xilobiosa, xilo-oligomeros lineales mezclados, oligomeros arabinoxilanos ramificados de la hemicelulosa, o mazorcas pretratadas con amoníaco diluído como sustratos. El residuo catalítico previsto es D291, mientras que los residuos flanqueados, S290 y C292, se predijeron que participen en el sustrato vinculante (Figura 64). E175 y E213 se conservan a través de otras enzimas GH3 y se predicen que tienen funciones catalíticas. Según lo utilizado aquí, "un polipéptido de Fv3A" se refiere a un polipéptido y/o a una variante de eso que comprende una secuencia que tiene por lo menos el 85%, *p. ej.,* por lo menos identidad de la secuencia del 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o del 100% por lo menos a 50, *p. ej., por lo menos los residuos contiguos del aminoácido* de 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, o 700 de 24 a 766 del NO. SEQ de la identificación: 2. Un polipéptido de Fv3A es preferiblemente inalterado con respecto a Fv3A nativo en los residuos D291, S290, C292, E175, y E213. Un polipéptido Fv3A preferentemente es inalterado en al menos el 70%, 75%, 80%, 85%, 90%, 95%, 98% o 99% de los residuos de aminoácidos que se conservan entre Fv3A, *Trichoderma reesei* Bxl1 y/o *Trichoderma reesei* Bgl1, como se muestra en la alineación de la Figura 64. Un polipéptido Fv3A comprende adecuadamente todo el dominio conservado predicho de nativos Fv3A como se muestra en la Figura 1B. Un polipéptido Fv3A ejemplar de la invención comprende una secuencia que tiene al menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con la secuencia de Fv3A madura, como se muestra en la Figura 1B. El polipéptido Fv3A de la invención preferentemente tiene actividad β-xilosidasa.

En consecuencia un polipéptido Fv3A de la invención comprende adecuadamente una secuencia de aminoácidos con al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos de SEQ ID NO: 2, o residuos (i) 24-766, (ii) 73-321, (iii) 73-394, (iv) 395-622, (v) 24-622, o (vi) 73-622 de SEQ ID NO: 2. El polipéptido convenientemente tiene actividad β-xilosidasa.

Pf43A: La secuencia de aminoácidos de Pf43A (SEQ ID NO: 2) se muestra en las Figuras 2B y 53. SEQ ID NO.: 4 es la secuencia del Pf43A inmaduro. Pf43A tiene una secuencia predicha de la señal que corresponde a los residuos 1 a 20 del SEQ ID NO.: 4 (subrayado en la Figura 2B); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a los residuos 21 a 445 del SEQ ID NO.: 4. El dominio conservado predicho está en negrilla, el CBM predicho está en mayúsculas y el vinculante previsto separando el CD y el CBM está en cursiva en la Figura 2B. Pf43A ha demostrado tener actividad β-xilosidasa, en, por ejemplo, un ensayo enzimático utilizando *p*- nitofenilo-β-xylopiranosida, xilobiosa, mezcla lineal xilo-oligómeros o mazorca pretratada con amoníaco como sustratos. Los residuos catalíticos predichos incluyen D32 o D60, D145, y E206. La región del terminal C subrayada en la Figura 53 es el CBM predicho. Según lo utilizado aquí, "un polipéptido de Pf43A" se refiere a un polipéptido y/o a una variante de eso que comprende una secuencia que tiene por lo menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de la identidad de secuencia a por lo menos a 50, 75, 100, 125, 150, 175, 200, 250, 300, 350 o 400, de residuos contiguos de aminoácidos entre los residuos 21 a 445 de SEQ ID NO.: 4. Un polipéptido Pf43A preferentemente es inalterable en comparación con los Pf43A nativos en residuos D32 o D60, D145 y E206. Un Pf43A es preferiblemente inalterado en por lo menos el 70%, 80%, 90%, 95%, 98%, o 99% de los residuos del aminoácido que se encuentran conservados en la familia de proteínas incluyendo Pf43A y 1, 2, 3, 4, 5, 6, 7, o todos los 8 de otras secuencias del aminoácido en la alineación de la Figura 53. Un polipéptido Pf43A de la invención convenientemente comprende dos o más o todos los siguientes dominios: (1) el polipéptido CBM, (2) el dominio conservado predicho y (3) el vinculante Pf43A que se muestra en la Figura 2B. Un polipéptido Pf43A ejemplar de la invención comprende una secuencia que tiene al menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con la secuencia Pf43A madura, como se muestra en la Figura 2B. El polipéptido Pf43A de la invención preferentemente tiene actividad β-xilosidasa.

En consecuencia un polipéptido Pf43A de la invención comprende adecuadamente una secuencia de aminoácidos con al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de la identidad de secuencia con la secuencia de aminoácidos de SEQ ID NO: 4, o residuos (i) 21-445, (ii) 21-301, (iii) 21-323, (iv) 21-

444, (v) 302-444, (vi) 324-445, (vii) 324-444, o (viii) 324-445 de SEQ ID NO: 4. El polipéptido convenientemente tiene actividad β-xilosidasa.

Fv43E: La secuencia de aminoácidos de Fvf43E (SEQ ID NO.: 6) se muestra en las Figuras 3B y 53. SEQ ID NO.:6 es la secuencia del Fv3E inmaduro. Fv43E tiene una secuencia predicha de la señal que corresponde a los residuos 1 a 18 del SEQ ID NO.: 6 (subrayado en la Figura 3B); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a los residuos 19 a 530 de SEQ ID NO.: 6. Los dominios conservados predichos aparecen en negrilla en la Figura 3B. Se demostró que Fv3E tiene una actividad de β-xiosidasa, por ejemplo, en un ensayo enzimático usando 4- nitrofeniol-β-xilopiranosida, xilobiosa y xilo-oligomeros lineales mezclados, o mazorcas pretratadas con amoníaco diluído como substratos. Los residuos catalíticos predichos incluyen D40 o D71, D155 y E241. Según lo utilizado aquí, "un polipéptido de Fv43A" se refiere a un polipéptido y/o a una variante de eso que comprende una secuencia que tiene por lo menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de la identidad de secuencia a por lo menos a 50, 75, 100, 125, 150, 175, 200, 250, 300, 350 o 400, 450 o 500 de residuos contiguos de aminoácidos entre los residuos 19 a 530 de SEQ ID NO.: 6. Un polipéptido de Fv43E es preferiblemente inalterado con respecto al Fv43E nativo en los residuos D40 o D71, D155, y E241. Un polipéptido Fv43E es preferiblemente inalterado en por lo menos el 70%, 80%, 90%, 95%, 98%, o 99% de los residuos del aminoácido que se encuentran conservados en la familia de enzimas incluyendo Fv43E y 1, 2, 3, 4, 5, 6, 7, o todos las 8 otras secuencias del aminoácido en la alineación de la Figura 53. Un polipéptido Fv43A comprende adecuadamente todo el dominio conservado predicho de nativos Fv43E como se muestra en la Figura 3B. Un polipéptido Fv43A ejemplar comprende una secuencia que tiene al menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de la identidad con la secuencia de Fv43E madura, como se muestra en la Figura 3B. El polipéptido Fv43A de la invención preferentemente tiene actividad β-xilosidasa.

En consecuencia un polipéptido Fv43E de la invención comprende adecuadamente una secuencia de aminoácidos con al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos de SEQ ID NO:6, o residuos (i) 19-530, (ii) 29-530, (iii) 19-300 o (iv) 29-300 de SEQ ID NO: 6. El polipéptido convenientemente tiene actividad β-xilosidasa.

Fv39A: La secuencia de aminoácidos de Fv39A (SEQ ID NO: 8) se muestra en las Figuras 4B y 52. SEQ ID NO.: 8 es la secuencia del Fv39A inmaduro. Fv39A tiene una secuencia predicha de la señal que corresponde a los residuos 1 a 19 de SEQ ID NO.: 8 (subrayado en la Figura 4B); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a los residuos 20 a 439 del SEQ ID NO.: 8. Los dominios conservados predichos aparecen en negrilla en la Figura 4B. Se demostró que Fv39A tiene una actividad de β-xiosidasa, por ejemplo, en un ensayo enzimático usando p- nitrofenil-β-xilopiranosida, xilobiosa o xilo-oligomeros lineales mezclados, como sustratos. Residuos de Fv39A E168 y E272 están predichos para funcionar como catalizador de base ácida y nucleófilo, respectivamente, basados en un alineamiento de secuencia mencionado en las xilosidasas GH39 de *Thermoanaerobacterium saccharolyticum* (Uniprot adhesión Nº P36906) y *Geobacilo stearothermophilus* (Uniprot adhesión Nº Q9ZFM2) con Fv39A. Como se utiliza en el presente documento, "un polipéptido Fv39A" se refiere a un polipéptido y/o a una variante de lo mismo que comprende una secuencia que tiene al menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad de secuencia con por lo menos 50, 75, 100, 125, 150, 175, 200, 250, 300, 350 o 400 residuos contiguos de aminoácidos entre residuos 20 a 439 de SEQ ID NO: 8. Un polipéptido de Fv39A está preferiblemente inalterado con respecto a Fv39A nativo en los residuos E168 y E272. Un polipéptido Fv39A está preferentemente inalterado en al menos el 70%, 80%, 90%, 95%, 98% o 99% de los residuos de aminoácidos que se conservan entre una familia o enzimas, incluyendo Fv39A y xilosidasas de *Thermoanaerobacterium saccharolyticum* y *Geobcilos stearothermophilus* (véase el anterior párrafo). Un polipéptido Fv39A comprende adecuadamente todo el dominio conservado predicho de Fv3A nativos como se muestra en la Figura 4B. Un polipéptido Fv39A ejemplar comprende una secuencia que tiene al menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con la secuencia de Fv3A madura, como se muestra en la Figura 4B. El polipéptido Fv39A de la invención preferentemente tiene actividad β-xilosidasa.

En consecuencia un polipéptido Fv39A de la invención comprende adecuadamente una secuencia de aminoácidos con al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos de SEQ ID NO: 8, o residuos (i) 20-439, (ii) 20-291, (iii) 145-291, (iv) 145-439 de SEQ ID NO: 8. El polipéptido convenientemente tiene actividad β-xilosidasa.

Fv43A: La secuencia de aminoácidos de Fv43A (SEQ ID NO: 10) se muestra en las Figuras 5B y 53. SEQ ID NO.:10 es la secuencia del Fv3A inmaduro. Fv43A tiene una secuencia predicha de la señal que corresponde a los residuos 1 a 22 del SEQ ID NO.: 10 (subrayado en la Figura 5B); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a los residuos 23 a 449 de SEQ ID NO.:10. En la Figura 5B el dominio conservado predicho se encuentra en negrilla, el CBM predicho aparece en letras mayúsculas y el vinculante predicho que separa el CD y el CBM aparece en itálica. Se demostró que Fv43A tiene actividad β-xilosidasa, por ejemplo, un ensayo enzimático utilizando 4-nitophenil-β-D-xilopiranosida, xilobiosa, xilo-oligómeros mixtos, lineales, oligómeros de arabinoxilano ramificados de hemicelulosa y/o xilo-oligómeros lineales como sustratos. Los residuos catalíticos predichos incluyen D34 o D62, D148 y E209. Según lo utilizado aquí, "un polipéptido de Fv43A" se refiere a un polipéptido y/o a una variante de eso que comprende una secuencia que tiene por lo menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de la identidad de secuencia con por lo menos 50, 75, 100, 125, 150, 175, 200, 250, 300, 350 o 400 residuos contiguos de aminoácidos entre los residuos 23 a 449 de SEQ ID NO.:10. Un polipéptido Fv43E está preferiblemente inalterado al compararse con Fv43A nativo, en los residuos D34 o D62, D148 y E209, un polipéptido Fv43A preferiblemente no alterado en por lo menos el 70%, 80%, 90%, 95%, 98%, o 99% de los residuos del aminoácido que se encuentran conservados en la familia de enzimas incluyendo Fv43A y 1, 2, 3, 4, 5, 6, 7, o todas las 9 otras secuencias del aminoácido en la alineación de la Figura 53. Un polipéptido Fv43A comprende adecuadamente todo el dominio conservado predicho de CBM del Fv3A nativo y/o todo el dominio conservado predicho de Fv43A nativo, y/o el enlazador como se muestra en la Figura 5B. Un polipéptido Fv43A ejemplar comprende una secuencia que tiene al menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con la secuencia de Fv43A madura, como se muestra en la Figura 5B. El polipéptido Fv45A de la invención preferentemente tiene actividad β-xilosidasa.

En consecuencia un polipéptido Fv43A de la invención comprende adecuadamente una secuencia de aminoácidos con al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos de SEQ ID NO: 10, o residuos (i) 23-449, (ii) 23-302, (iii) 23-320, (iv) 23-

448, (v) 303-448 (vii) 321-448 o (viii) 321-449 de SEQ ID NO:10. El polipéptido convenientemente tiene actividad β-xilosidasa.

Fv43B: La secuencia de aminoácidos de Fv43B (SEQ ID NO.: 12) se muestra en las Figuras 6B y 53. SEQ ID NO.:12 es la secuencia del Fv43B inmaduro. Fv43B tiene una secuencia predicha de la señal que corresponde a los residuos 1 a 16 de SEQ ID NO.: 12 (subrayado en la Figura 6B); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a los residuos 17 a 574 del SEQ ID NO.:12. El dominio conservado predicho está en negrilla en la figura 6B. Fv43B demostró tener actividades de L-α-arabinofuranosidasa y β-xilosidasa, por ejemplo, un primer ensayo enzimático utilizando 4-nitofenilo-β-D-xilopiranosida y *p*- nitrofenilo-α-L-arabinofuranosida como sustratos. Se mostró, en un segundo ensayo enzimático, para impulsar la liberación de arabinosa ramificada arabino-xilooligomeros y para catalizar la liberación incrementada de xilosa de mezclas de oligómero en presencia de otras enzimas de xilosidasa. Los residuos catalíticos predichos incluyen D38 o D68, D151 y E236. Según lo utilizado aquí, "un polipéptido de Fv43B" se refiere a un polipéptido y/o a una variante de eso que comprende una secuencia que tiene por lo menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de la identidad de secuencia a por lo menos a 50, 75, 100, 125, 150, 175, 200, 250, 300, 350 o 400, 450, 500 o 550 de residuos contiguos de aminoácidos entre los residuos 17 a 574 de SEQ ID NO.:12. Un polipéptido de Fv43B está preferiblemente inalterado con respecto a Fv43B nativo en los residuos D38 o D68, D151 y E236. Un polipéptido Fv43B está preferiblemente inalterado en por lo menos el 70%, 80%, 90%, 95%, 98%, o 99% de los residuos del aminoácido que se encuentran conservados en una familia de enzimas incluyendo Fv43B y 1, 2, 3, 4, 5, 6, 7, 8 o todas las 9 otras secuencias del aminoácido en la alineación de la Figura 53. Un polipéptido Fv43B comprende adecuadamente el dominio entero conservado predicho de Fv43B nativo como se muestra en las Figuras 6B y 53. Un ejemplar polipéptido Fv43B comprende una secuencia que tiene al menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con la secuencia de Fv43B madura, como se muestra en la Figura 6B. El polipéptido Fv43B de la presente invención preferentemente tiene actividad β-xilosidasa, actividad de L-α-arabinofuranosidasa, o tanto actividades de β-xilosidasa como de L-α-arabinofuranosidasa.

En consecuencia un polipéptido Fv43B de la invención comprende adecuadamente una secuencia de aminoácidos con al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos de SEQ ID NO: 12, o residuos (i) 17-574, (ii) 27-574, (iii) 17-303, (iv) 27-303 de SEQ ID NO: 12. El polipéptido tiene convenientemente actividad de β-xilosidasa, actividad de L-α-arabinofuranosidasa, o las dos actividades β-xilosidasa y L-α-arabinofuranosidasa.

Pa51A: La secuencia de aminoácidos de Pa51A (SEQ ID NO: 14) se muestra en las Figuras 7B y 54. SEQ ID NO.: 14 es la secuencia del Pa51A inmaduro. Pa51A tiene una secuencia predicha de la señal que corresponde a los residuos 1 a 20 del SEQ ID NO.:14 (subrayado en la Figura 7B); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a los residuos 21 a 676 del SEQ ID NO.:14. El dominio conservado predicho L-α-arabinofuranosidasa está en negrilla en la Figura 7B. Pa51A demostró tener actividades de y β-xilosidasa y de L-α-arabinofuranosidasa, por ejemplo, en un ensayo enzimático utilizando sustratos artificiales de p-nitrofenilo-β-D-xilopiranosida y p- nitofenilo-α-L-arabinofuranosida. Se demostró para catalizar la liberación de arabinosa de los arabino-xilooligomeros ramificados y para catalizar la liberación incrementada de xilosa de las mezclas de oligómero en presencia de otras enzimas de xilosidasa. Los residuos ácidos conservados incluyen E43, D50, E257, E296, E340, E370, E485, y E493. Según lo utilizado aquí, "un polipéptido de Pa51A" se refiere a un polipéptido y/o a una variante de eso que comprende una secuencia que tiene por lo menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de la identidad de secuencia a por lo menos a 50, 75, 100, 125, 150, 175, 200, 250, 300, 350 o 400, 450, 500, 550, 600 o 650 de residuos contiguos de aminoácidos entre los residuos 21 a 676 de SEQ ID NO.: 14. Un polipéptido de Pa51A está preferiblemente inalterado, cuando se compara con Pa51A nativo, en los residuos E43, D50, E257, E296, E340, E370, E485, y E493. Un polipéptido Pa51A está preferentemente inalterado en al menos el 70%, 80%, 90%, 95%, 98% o 99% de los residuos de aminoácidos que se conservan entre un grupo de enzimas incluyendo Pa51A, Fv51A y Pf51A, como se muestra en la alineación de la Figura 54. Un polipéptido Pa51A comprende adecuadamente un dominio conservado predicho de Pa51A nativo como se muestra en la figura 7B. Un polipéptido ejemplar Pa51A comprende una secuencia que tiene al menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de

identidad con la secuencia de Pa51A madura, como se muestra en la Figura 7B. El polipéptido Pa51A de la invención preferentemente tiene actividad β-xilosidasa, actividad de L-α-arabinofuranosidasa, o las dos actividades de β-xilosidasa y de L-α-arabinofuranosidasa.

En consecuencia un polipéptido Pa51A de la invención comprende adecuadamente una secuencia de aminoácidos con al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos de SEQ ID NO: 14, o residuos (i) 21-676, (ii) 21-652, (iii) 469-652 o (iv) 469-676 de SEQ ID NO: 14. El polipéptido tiene convenientemente actividad de β-xilosidasa, actividad de L-α-arabinofuranosidasa, o las dos actividades β-xilosidasa y L-α-arabinofuranosidasa.

Gz43A: La secuencia de aminoácido de Gz43A (SEQ ID NO: 16) se muestra en las Figuras 8B y 53. SEQ ID NO.: 16 es la secuencia de Gz43A inmaduro. Gz43A tiene una secuencia predicha de la señal que corresponde a los residuos 1 a 18 del SEQ ID NO.: 16 (subrayado en la Figura 8B); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a los residuos 19 a 340 del SEQ ID NO.: 16. Los dominios conservados predichos aparecen en negrilla en la Figura 8B. Se demostró que Gz43A tiene una actividad de β-xiosidasa, por ejemplo, en un ensayo enzimático usando p- nitofenil-β-xilopiranosida, xilobiosa y/o xilo-oligomeros lineales mezclados como substratos. Los residuos catalíticos predichos incluyen D33 o D68 y E243. Según lo utilizado aquí, "un polipéptido de Gz43A" se refiere a un polipéptido y/o a una variante de eso que comprende una secuencia que tiene por lo menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de identidad de secuencia con por lo menos 50, 75, 100, 125, 150, 175, 200, 250 o 300 residuos contiguos de aminoácidos entre los residuos 19 a 340 de SEQ ID NO.: 16. Un polipéptido de Gz43A está preferiblemente inalterado con respecto a Gz43A nativo en los residuos D33 o D68, D154 y E243. Un polipéptido Gz43A está preferentemente inalterado en por lo menos el 70%, 80%, 90%, 95%, 98%, o 99% de los residuos del aminoácido que se encuentran conservados en un grupo de enzimas incluyendo Gz43A y 1, 2, 3, 4, 5, 6, 7, 8 o todas las 9 otras secuencias del aminoácido en la alineación de la Figura 53. Un polipéptido Gz43A comprende adecuadamente todo el dominio conservado predicho de nativos Gz43A como se muestra en la Figura 8B. Un polipéptido Gz43A ejemplar comprende una

288

secuencia que tiene al menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con la secuencia de Gz43A madura, como se muestra en la Figura 8B. El polipéptido Gz43A de la invención preferentemente tiene actividad β-xilosidasa.

En consecuencia un polipéptido Gz43A de la invención comprende adecuadamente una secuencia de aminoácidos con al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos de SEQ ID NO: 16, o residuos (i) 19-340, (ii) 53-340, (iii) 19-383 o (iv) 53-383 de SEQ ID NO: 16. El polipéptido convenientemente tiene actividad β-xilosidasa.

Fo43A: La secuencia de aminoácido de Fo43A (SEQ ID NO:18) se muestra en las Figuras 9B y 53. SEQ ID NO.: 18 es la secuencia del Fo3A inmaduro. Fo3A tiene una secuencia predicha de la señal que corresponde a los residuos 1 a 20 del SEQ ID NO.: 18 (subrayado en la Figura 9B); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a los residuos 21 a 348 del SEQ ID NO.: 18. Los dominios conservados predichos aparecen en negrilla en la Figura 9B. Se demostró que Fo43A tiene una actividad de β-xiosidasa, por ejemplo, en un ensayo enzimático usando p- nitofenilo-β-xilopiranosida, xilobiosa y/o xilo-oligomeros lineales mezclados como substratos. Los residuos catalíticos predichos incluyen D37 o D72, D159 y E251. Según lo utilizado aquí, "un polipéptido de Fo43A" se refiere a un polipéptido y/o a una variante de eso que comprende una secuencia que tiene por lo menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de la identidad de secuencia con por lo menos a 50, 75, 100, 125, 150, 175, 200, 250 o 300 de residuos contiguos de aminoácidos entre los residuos 18 a 344 de SEQ ID NO.: 18. Un polipéptido de Fo43B está preferiblemente inalterado con respecto a Fo43B nativo en los residuos D37 o D72 y E251. Un polipéptido Fo43A está preferentemente inalterado en por lo menos el 70%, 80%, 90%, 95%, 98%, o 99% de los residuos del aminoácido que se encuentran conservados en un grupo de enzimas incluyendo Fo43A y 1, 2, 3, 4, 5, 6, 7, 8 o todas las otras 9 secuencias del aminoácido en la alineación de la Figura 53. Un polipéptido Fo43A comprende adecuadamente el dominio conservado predicho de nativos Fo43A como se muestra en la Figura 9B. Un polipéptido Fo43A ejemplar comprende una secuencia que tiene al menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con la secuencia de Fo43A madura, como se muestra

289

en la Figura 9B. El polipéptido Fo43A de la invención preferentemente tiene actividad β-xilosidasa.

En consecuencia un polipéptido Fo43A de la invención comprende adecuadamente una secuencia de aminoácidos con al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos de SEQ ID NO: 18, o residuos (i) 21-341, (ii) 107-341, (iii) 21-348 o (iv) 107-348 de SEQ ID NO: 18. El polipéptido convenientemente tiene actividad β-xilosidasa.

Af43A: La secuencia del aminoácido Af43A (SEQ ID NO: 20) se muestra en las Figuras 10B y 53. SEQ ID NO.: 20 es la secuencia del Af43A no maduro. El dominio conservado predicho está en negrilla en la Figura 10B. Se demostró que Af43A tiene actividad de L-α-arabinofuranosidasa, por ejemplo, en un análisis enzimático usando *p*-nitofenilo- α-L-arabinofuranosida como sustrato. Se demostró que Af43A cataliza la liberación de arabinosa del sistema de oligómeros liberados de la hemicelulosa vía la acción de la endoxilanasa. Los residuos catalíticos predichos incluyen D26 o D58 y E227. Según lo utilizado aquí, "un polipéptido de Af43A" se refiere a un polipéptido y/o a una variante de eso que comprende una secuencia que tiene por lo menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de identidad de secuencia con por lo menos a 50, 75, 100, 125, 150, 175, 200, 250 o 300 de residuos contiguos de aminoácidos del SEQ ID NO.: 20. Un polipéptido de Af43A está preferiblemente inalterado con respecto a Af43A nativo en los residuos D26, D58 y E227. Un polipéptido Af43A está preferentemente inalterado en por lo menos el 70%, 80%, 90%, 95%, 98%, o 99% de los residuos del aminoácido que se encuentran conservados en un grupo de enzimas incluyendo Af43A y 1, 2, 3, 4, 5, 6, 7, 8 o todas las otras 9 secuencias del aminoácido en la alineación de la Figura 53. Un polipéptido de Af43A abarca convenientemente el dominio conservado predicho de Af43A nativo según las indicaciones de la Figura 10B. Un polipéptido ejemplar de Af43A comprende una secuencia que tiene por lo menos identidad de secuencia del 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o del 100% del SEQ ID NO.: 20. El polipéptido de Af43A de la invención tiene preferiblemente actividad de L-α-arabinofuranosidasa.

zao

En consecuencia un polipéptido Af43A de la invención comprende adecuadamente una secuencia de aminoácidos con al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos de SEQ ID NO.: 20, o residuos (i) 15-558, (ii) 15-295 de SEQ ID NO.: 20. El polipéptido tiene convenientemente actividad de L-α-arabinofuranosidasa.

Pf51A: La secuencia de aminoácidos de Pf51A (SEQ ID NO.: 22) se muestra en las Figuras 11B y 54. SEQ ID NO.: 22 es la secuencia del Pf51A inmaduro. Pf51A tiene una secuencia predicha de la señal que corresponde a los residuos 1 a 20 del SEQ ID NO.: 22 (subrayado en la Figura 11B); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a los residuos 21 a642 del SEQ ID NO.:22. El dominio conservado predicho L-α-arabinofuranosidasa está en negrilla en la Figura 11B. Pf51A demostró tener actividades de y β-xilosidasa y de L-α-arabinofuranosidasa, por ejemplo, en un ensayo enzimático utilizando como sustratos 4-nitrofenil -α-L-arabinofuranosida. Se demostró que Pf51A cataliza la liberación de arabinosa del sistema de oligomeros liberados de la hemicelulosa vía la acción de la endoxilanasa. Los residuos ácidos conservados predichos incluyen E43, D50, E248, E287, E331, E360, E472, y E480. Según lo utilizado aquí, "un polipéptido de Pf51A" se refiere a un polipéptido y/o a una variante de eso que comprende una secuencia que tiene por lo menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de la identidad de secuencia a por lo menos a 50, 75, 100, 125, 150, 175, 200, 250, 300, 350 o 400, 450, 500, 550 o 600 de residuos contiguos de aminoácidos entre los residuos 21 a 642 del SEQ ID NO.: 22. Un polipéptido de Pf51A es preferiblemente inalterado, cuando se compara con Pf51A nativo, en los residuos E43, D50, E248, E287, E331, E360, E472, y E480. Un polipéptido Pf51A es preferentemente inalterado en al menos el 70%, 80%, 90%, 95%, 98% o 99% de los residuos de aminoácidos que se conservan entre un grupo de enzimas incluyendo Pf51A, Pa51A y Fv51A, como se muestra en la alineación de Figura 54. Un polipéptido Pa51A comprende adecuadamente un dominio conservado predicho de Pa51A nativo como se muestra en la figura 11B. Un polipéptido ejemplar Pa51A comprende una secuencia que tiene al menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con la secuencia de Pa51A madura, como se muestra en la Figura 11B. El polipéptido Pa51A de la invención preferentemente tiene actividad de L-α-arabinofuranosidasa.

291

En consecuencia un polipéptido Pf51A de la invención comprende adecuadamente una secuencia de aminoácidos con al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos de SEQ ID NO.: 22, o con residuos (i) 21-632, (ii) 461-632, (iii) 21-642 o (iv) 461-642 del SEQ ID NO.: 22. El polipéptido tiene actividad de L-α-arabinofuranosidasa.

AfuXyn2: La secuencia del aminoácido de AfuXyn2 (SEQ ID NO.: 24) se muestra en las Figuras 12B y 55B. SEQ ID NO.: 24 es la secuencia de AfuXyn2 inmaduro. AfuXyn2 tiene una secuencia de señal predicha que corresponde a los residuos 1 a 18 del SEQ ID NO.: 24 (subrayado en la Figura 12B); la segmentación de la secuencia de señal se predice para rendir una proteína madura con una secuencia que corresponde a los residuos 19 a 228 del SEQ ID NO.: 24. El dominio conservado GH11 predicho está en negrilla en la Figura 12B. Se demostró que AfuXyn2 tiene actividad indirecta de la endoxilanasa observando su capacidad de catalizar la producción creciente del monomero de xilosa en presencia de la xilobiosidasa cuando las enzimas actúan en biomasa pretratada o en la hemicelulosa aislada. Los residuos catalíticos conservados incluyen E124, E129, y E215. Según lo utilizado aquí, "un polipéptido de AfuXyn2" se refiere a un polipéptido y/o a una variante de eso que comprende una secuencia que tiene por lo menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de la identidad de secuencia con por lo menos a 50, 75, 100, 125, 150, 175 o 200 de residuos contiguos de aminoácidos entre los residuos 19 a 228 del SEQ ID NO.: 24. Un polipéptido de AfuXyn2 está preferiblemente inalterado con respecto a AfuXyn2 nativo en los residuos E124, E129 y E215. Un polipéptido AfuXyn2 está preferentemente inalterado en al menos el 70%, 80%, 90%, 95%, 98% o 99% de los residuos de aminoácidos que se conservan entre AfuXyn2, AfuXyn5 y *Trichoderma reesei* Xyn2, como se muestra en la alineación de la Figura 55B. Un polipéptido AfuXyn2 comprende adecuadamente todo el dominio conservado predicho de AfuXyn2 nativo que se muestra en la Figura 12B. Un polipéptido AfuXyn2 ejemplar comprende una secuencia que tiene al menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con la secuencia de AfuXyn2 madura que se muestra en la Figura 12B. El polipéptido AfuXyn2 de la invención preferentemente tiene actividad xilanasa.

AfuXyn5: La secuencia del aminoácido de AfuXyn2 (SEQ ID NO.: 26) se muestra en las Figuras 13B y 55B. El SEQ ID NO.: 26 es la secuencia de AfuXyn5 inmaduro. AfuXyn5 tiene una secuencia de señal predicha que corresponde a los residuos 1 a 19 del SEQ ID NO.: 26 (subrayado en la Figura 13B); la segmentación de la secuencia de señal se predice para rendir una proteína madura con una secuencia que corresponde a los residuos 20 a 313 del SEQ ID NO.: 26. Los dominios conservados GH11 predichos están en negrilla en la Figura 13B. Se demostró que AfuXyn5 tiene actividad indirecta de endoxilanasa observando su capacidad de catalizar la producción creciente del monomero de xilosa en presencia de la xilobiosidasa cuando las enzimas actúan en biomasa pretratada o en la hemicelulosa aislada. Los residuos catalíticos conservados incluyen E119, E124 y E210. El CBM predicho está cerca del extremo del terminal C, caracterizado por los residuos hidrofóbicos numerosos y sigue la serina larga, serie treonina rica en aminoácidos. La región se muestra subrayada en la Figura 55B. Según lo utilizado aquí, "un polipéptido de AfuXyn5" se refiere a un polipéptido y/o a una variante de eso que comprende una secuencia que tiene por lo menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de identidad de secuencia con por lo menos 50, 75, 100, 125, 150, 175, 200, 250 o 275 de residuos contiguos de aminoácidos entre los residuos 20 a 313 del SEQ ID NO.: 26. Un polipéptido de AfuXyn5 es preferiblemente inalterado con respecto a AfuXyn5 nativo en los residuos E119, E120 y E210. Un polipéptido AfuXyn5 es preferentemente inalterado en al menos el 70%, 80%, 90%, 95%, 98% o 99% de los residuos de aminoácidos que se conservan entre AfuXyn5, AfuXyn2 y *Trichoderma reesei* Xyn2, como se muestra en la alineación de la Figura 55B. Un polipéptido AfuXyn5 comprende adecuadamente todo el dominio conservado predicho CBM de AfuXyn5 nativo y/o todo el dominio conservado predicho AfuXyn5 nativo (subrayado) que se muestra en la Figura 13B. Un polipéptido AfuXyn5 ejemplar comprende una secuencia que tiene al menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con la secuencia de AfuXyn5 madura que se muestra en la Figura 13B. El polipéptido AfuXyn5 de la invención preferentemente tiene actividad de xilanasa.

Fv43D: La secuencia de aminoácido de Fv43D (SEQ ID NO.:28) se muestra en las Figuras 14B y 53. El SEQ ID NO.: 28 es la secuencia del Fv43D inmaduro. Fv43D tiene una secuencia predicha de la señal que corresponde a los residuos 1 a 20 del SEQ ID NO.: 28 (subrayado en la Figura 14B); la segmentación de la secuencia de la

señal se predice para rendir una proteína madura con una secuencia que corresponde a los residuos 21 a 350 del SEQ ID NO.: 28. Los dominios conservados predichos aparecen en negrilla en la Figura 14B. Se demostró que Fv43D tiene una actividad de β-xiosidasa, por ejemplo, en un ensayo enzimático usando p- nitofenilo-β-xilopiranosida, xilobiosa y/o xilo-oligomeros lineales mezclados como sustratos. Los residuos catalíticos predichos incluyen D37 o D72, D159 y E251. Según lo utilizado aquí, "un polipéptido de Fv43D" se refiere a un polipéptido y/o a una variante de eso que comprende una secuencia que tiene por lo menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de la identidad de secuencia con por lo menos 50, 75, 100, 125, 150, 175, 200, 250, 300 o 320 de residuos contiguos de aminoácidos entre los residuos 21 a 350 del SEQ ID NO.: 28. Un polipéptido de Fv43D está preferiblemente inalterado con respecto a Fv43D nativo en los residuos D37 o D72 y E251. Un polipéptido Fv43D está preferentemente inalterado en por lo menos el 70%, 80%, 90%, 95%, 98%, o 99% de los residuos del aminoácido que se encuentran conservados en un grupo de enzimas incluyendo Fv43D y 1, 2, 3, 4, 5, 6, 7, 8 o todas las otras 9 secuencias del aminoácido en la alineación de la Figura 53. Un polipéptido Fv43D comprende adecuadamente todo el dominio conservado predicho de Fv43D nativos como se muestra en la Figura 14B. Un polipéptido Fv43D ejemplar comprende una secuencia que tiene al menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con la secuencia de Fv43D madura, como se muestra en la Figura 14B. El polipéptido Fv43D de la invención preferentemente tiene actividad β-xilosidasa.

En consecuencia un polipéptido Fv43D de la invención comprende adecuadamente una secuencia de aminoácidos con al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos del SEQ ID NO.: 28, o con residuos (i) 20-341, (ii) 21-350, (iii) 107-341 o (iv) 107-350 del SEQ ID NO.: 28. El polipéptido convenientemente tiene actividad β-xilosidasa.

Pf43B: La secuencia de aminoácido de Pf43B (SEQ ID NO.: 30) se muestra en las Figuras 15B y 53. El SEQ ID NO.: 30 es la secuencia del Pf43B inmaduro. Pf43B tiene una secuencia predicha de la señal que corresponde a los residuos 1 a 20 del SEQ ID NO.:30 (subrayado en la Figura 15B); la segmentación de la secuencia de la señal se predice para rendir una proteína madura con una secuencia que corresponde a

los residuos 21 a 321 del SEQ ID NO.:30. El dominio conservado predicho está en negrilla en la Figura 15B. Los residuos ácidos conservados dentro del dominio conservado incluyen D32, D61, D148 y E212. Se demostró que Pf43B tiene actividad β-xilosidasa, por ejemplo, en un ensayo enzimático utilizando p-nitrofenil-β-D-xilopiranosida, xilobiosa, y/o xilo-oligómeros mixtos, lineales como sustratos. Según lo utilizado aquí, "un polipéptido de Pf43B" se refiere a un polipéptido y/o a una variante de eso que comprende una secuencia que tiene por lo menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de identidad de secuencia con por lo menos 50, 75, 100, 125, 150, 175, 200, 250 o 280 de residuos contiguos de aminoácidos entre los residuos 21 a 321 del SEQ ID NO.: 30. Un polipéptido de Pf43B está preferiblemente inalterado con respecto a Pf43B nativo en los residuos D32 o D61, D148 y E212. Un polipéptido Pf43B está preferentemente inalterado en por lo menos el 70%, 80%, 90%, 95%, 98%, o 99% de los residuos del aminoácido que se encuentran conservados en un grupo de enzimas incluyendo Pf43B y 1, 2, 3, 4, 5, 6, 7, 8 o todas las otras 9 secuencias del aminoácido en la alineación de la Figura 53. Un polipéptido Pf43B comprende adecuadamente el dominio conservado predicho de Pf43B nativo como se muestra en la Figura 15B. Un polipéptido Pf43B ejemplar comprende una secuencia que tiene al menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con la secuencia de Pf43B madura, como se muestra en la Figura 15B. El polipéptido Pf43B de la invención preferentemente tiene actividad β-xilosidasa.

En consecuencia un polipéptido Pf43B de la invención comprende adecuadamente una secuencia de aminoácidos con al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos de la SEQ ID NO.: 30. El polipéptido convenientemente tiene actividad β-xilosidasa.

Fv51A: La secuencia de aminoácidos de Fv51A (SEQ ID NO: 32) se muestra en las Figuras 16B y 54. La SEQ ID NO.: 32 es la secuencia del Fv51A inmaduro. Fv3A tiene una secuencia predicha de señal que corresponde a los residuos 1 a 19 de la SEQ ID NO.: 32 (subrayados en la Figura 16B); la segmentación de la secuencia de señal se predice para rendir una proteína madura con una secuencia que corresponde a los residuos 20 a 660 de la SEQ ID NO.: 32. El dominio conservado predicho L-α-arabinofuranosidasa está en negrilla en la Figura 16B. Fv51A demostró tener

actividades de L-α-arabinofuranosidasa, por ejemplo, en un ensayo enzimático utilizando como sustratos 4-nitrofenilo -α-L-arabinofuranosida. Se demostró que Fv51A cataliza la liberación de arabinosa del sistema de oligomeros liberados de la hemicelulosa vía la acción de la endoxilanasa. Los residuos conservados incluyen E42, D49, E247, E286, E330, E359, E479 y E487. Según lo utilizado aquí, "un polipéptido de Fv51A " se refiere a un polipéptido y/o a una variante de eso que comprende una secuencia que tiene por lo menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de la identidad de secuencia con por lo menos a 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600 o 625 de residuos contiguos de aminoácidos entre los residuos 20 a 660 de la SEQ ID NO.: 32. Un polipéptido de Fv51A es preferiblemente inalterado con respecto a Fv51A nativo en los residuos E42, D49, E247, E286, E359, E479 y E487. Un polipéptido Fv51A está preferentemente inalterado en al menos el 70%, 80%, 90%, 95%, 98% o 99% de los residuos de aminoácidos que se conservan entre Fv51A , Pa51A y Pf51A, como se muestra en la alineación de Figura 54. Un polipéptido Fv51A comprende adecuadamente un dominio conservado predicho de Fv51A nativo como se muestra en la Figura 16B. Un polipéptido ejemplar Fv51A comprende una secuencia que tiene al menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con la secuencia de Pa51A madura, como se muestra en la Figura 16B. El polipéptido Fv51A de la invención preferentemente tiene actividad de L-α-arabinofuranosidasa.

En consecuencia un polipéptido Fv51A de la invención comprende adecuadamente una secuencia de aminoácidos con al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos de SEQ ID NO: 32, o con residuos (i) 21-660, (ii) 21-645, (iii) 450-645 o (iv) 450-660 de la SEQ ID NO:32. El polipéptido tiene convenientemente actividad de L-α-arabinofuranosidasa.

Xyn3: La secuencia del aminoácido de Trichoderma reesei Xyn3 (SEQ ID NO: 42) se muestra en la Figura 21B. La SEQ ID NO: 42 es la secuencia de Trichoderma reesei Xyn3 inmaduro. Trichoderma reesei Xyn3 tiene una secuencia de señal predicha que corresponde a los residuos 1 a 16 de la SEQ ID NO.: 42 (subrayado en la Figura 21B); la segmentación de la secuencia de señal se predice para rendir una proteína madura con una secuencia que corresponde a los residuos 17 a 347 de la SEQ ID NO.:

42. El dominio conservado predicho está en negrilla en la Figura 21B. Se demostró que Trichoderma reesei Xyn3 tiene actividad indirecta de la endoxilanasa observando su capacidad de catalizar la producción creciente del monomero xilosa en presencia de la xilobiosidasa cuando las enzimas actúan en biomasa pretratada o en la hemicelulosa aislada. Los residuos catalíticos conservados incluyen E91, E176, E180, E195, y E282, según lo determinado por la alineación con otra enzima de la familia GH10, el delta Xys1 del Streptomyces halstedii *(Canals* y otros, ., 2003, Act Crystalogr. D Biol. 59:1447 - 53), que tiene 33% de identidad de la secuencia con Trichoderma reesei Xyn3 . Según lo utilizado aquí, "un polipéptido de Trichoderma reesei Xyn3" se refiere a un polipéptido y/o a una variante de eso que comprende una secuencia que tiene por lo menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de identidad de secuencia con por lo menos 50, 75, 100, 125, 150, 175, 200, 250 o 300 de residuos contiguos de aminoácidos entre los residuos 17 a 347 de la SEQ ID NO.: 42. Un polipéptido de Trichoderma reesei *Xyn3 de* está preferiblemente inalterado, con respecto al Trichoderma reesei *Xyn3 nativo, en los residuos E91, E176, E180, E195 y E282. Un polipéptido Trichoderma reesei* Xyn3 está preferentemente inalterado en al menos el 70%, 80%, 90%, 95%, 98% o 99% de los residuos de aminoácidos que se conservan entre *Trichoderma reesei* Xyn3 y Xys1 delta. Un polipéptido *Trichoderma reesei* Xyn3 comprende adecuadamente todo el dominio conservado predicho del *Trichoderma reesei Xyn3 nativo* que se muestra en la Figura 21B. Un polipéptido ejemplar *Trichoderma reesei* Xyn3 comprende una secuencia al menos del 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con el Trichoderma reesei Xyn3 maduro que se muestra en la Figura 21B. El polipéptido *Trichoderma reesei* Xyn3 de la invención preferentemente tiene actividad xilanasa.

Xyn2: La secuencia del aminoácido de Trichoderma reesei Xyn2 (SEQ ID NO: 42) se muestra en las Figuras 22 y 55B. La SEQ ID NO: 43 es la secuencia de Trichoderma reesei Xyn2 inmaduro.El Trichoderma reesei Xyn2 tiene una secuencia predicha del prepropeptido que corresponde a los residuos 1 a 33 de la SEQ ID NO.: 43 (subrayado en la Figura 22); la segmentación de la secuencia predicha de la señal entre las posiciones 16 y 17 se predice para rendir un propeptido, que es procesado por una proteasa como un kexin entre las posiciones 32 y 33, generando la proteína madura con una secuencia que corresponde a los residuos 33 a 222 de la SEQ ID NO.: 43. El dominio conservado predicho aparece en negrilla en la Figura 22. *Se demostró que*

Trichoderma reesei Xyn2 tiene actividad indirecta de la endoxilanasa observando su capacidad de catalizar una producción creciente del monómero xilosa en presencia de la xilobiosidasa cuando las enzimas actúan en biomasa pretratada o en la hemicelulosa aislada. Los residuos ácidos conservados incluyen E118, E123, y E209. Según lo utilizado aquí, "un polipéptido de Trichoderma reesei Xyn2" se refiere a un polipéptido y/o a una variante de eso que comprende una secuencia que tiene por lo menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de identidad de secuencia con por lo menos 50, 75, 100, 125, 150 o 175 de residuos contiguos de aminoácidos entre los residuos 33 a 222 de la SEQ ID NO.: 43. Un polipéptido de Trichoderma reesei Xyn2 está preferentemente inalterado, con respecto al Trichoderma reesei Xyn2 nativo, en los residuos E118, E123 y E209. Un polipéptido Trichoderma reesei Xyn2 está preferentemente inalterado en al menos el 70%, 80%, 90%, 95%, 98% o 99% de los residuos de aminoácidos que se conservan entre *Trichoderma reesei* Xyn2, AfuXyn2 y AfuXyn5, como se muestra en la alineación de Figura 55B. Un polipéptido *Trichoderma reesei* Xyn2 comprende adecuadamente todo el dominio conservado predicho de *Trichoderma reesei Xyn2* nativo que se muestra en la Figura 22. Un polipéptido *ejemplar Trichoderma reesei* Xyn2 comprende una secuencia teniendo al menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con la secuencia madura del *Trichoderma reesei* Xyn2 que se muestra en la Figura 22. El polipéptido *Trichoderma reesei* Xyn2 de la invención preferentemente tiene actividad xilanasa.

Bxl1: La secuencia de aminoácidos de Trichoderma reesei Bxl1 (SEQ ID NO: 44) se muestra en las Figuras 23 y 64. La SEQ ID NO.: 42 es la secuencia de Trichoderma reesei *Bxl1 inmaduro . Trichoderma reesei Bxl1 tiene una secuencia de señal predicha que corresponde a los residuos 1 a 18 de la SEQ ID NO.: 44 (subrayado en la Figura 23); la segmentación de la secuencia de señal se predice para rendir una proteína madura con una secuencia que corresponde a los residuos 19 a 797 de la SEQ ID NO.: 44. Los dominios conservados predichos aparecen en negrilla en la Figura 23. Se demostró que Trichoderma reesei Bxl1 tiene actividad β-xilosidasa, por ejemplo, en un ensayo enzimático utilizando p-nitofenilo-β--xilopiranosida, xilobiosa, y/o xilo-oligómeros mixtos, lineales como sustratos. Los residuos ácidos conservados incluyen E193, E234 y D310. Según lo utilizado aquí, "un polipéptido de Trichoderma reesei Bxl1" se refiere a un polipéptido y/o a una variante de eso que comprende una secuencia que tiene por lo menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%,*

96%, 97%, 98%, 99%, o 100% de identidad de secuencia con por lo menos 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700 o 750 de residuos contiguos de aminoácidos entre los residuos 17 a 797 de la SEQ ID NO.: 44. Un polipéptido de Trichoderma reesei Bxl1 está preferiblemente inalterado, con respecto al Trichoderma reesei Bxl1 nativo, en los residuos E193, E234 y D310. Un polipéptido Trichoderma reesei Bxl1 preferentemente está inalterado en al menos el 70%, 80%, 90%, 95%, 98% o 99% de los residuos de aminoácidos que se conservan entre Trichoderma reesei Bxl1, Fv3A y Trichoderma reesei Bxl1, como se muestra en la alineación de la Figura 64. Un polipéptido Trichoderma reesei Bxl1 comprende adecuadamente todos los dominios conservados predichos de *Trichoderma reesei Bxl1* nativo que se muestra en la Figura 23. Un polipéptido *ejemplar Trichoderma reesei* Bxl1 comprende una secuencia teniendo al menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con la secuencia madura del *Trichoderma reesei* Bxl1 que se muestra en la Figura 23. El polipéptido *Trichoderma reesei* Bxl1 de la invención preferentemente tiene actividad β-xilanasa.

En consecuencia un polipéptido Trichoderma reesei Bxl1 de la invención comprende adecuadamente una secuencia de aminoácidos con al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos de la SEQ ID NO.: 44. El polipéptido convenientemente tiene actividad β-xilosidasa.

Bgl1: La secuencia de aminoácidos de Trichoderma reesei Bgl1 (SEQ ID NO: 45) se muestra en las Figuras 24 y 64.Trichoderma reesei Bgl1 tiene una secuencia de señal predicha que corresponde a los residuos 1 a 19 de la SEQ ID NO.: 45 (subrayado en la Figura 24); la segmentación de la secuencia de señal se predice para rendir una proteína madura con una secuencia que corresponde a los residuos 20 a 744 de la SEQ ID NO.: 45. El dominio conservado predicho aparece en negrilla en la Figura 24. *El Trichoderma reesei* Bgl1 ha sido demostrado para tener actividad de β-glucosidasa por la observación de una capacidad de catalizar la hidrólisis de para-nitrofenil-β-D-glucopiranosida para producir para nitrofenol y una capacidad de catalizar la hidrólisis de la celobiosa. Los residuos ácidos conservados incluyen D164, E197 y D267. Según lo utilizado aquí, "un polipéptido de Trichoderma reesei Bgl1" se refiere a un polipéptido y/o a una variante de eso que comprende una secuencia que tiene por lo menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100%

299

de identidad de secuencia con por lo menos 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750 o 780 de residuos contiguos de aminoácidos entre los residuos 20 a 744 de la SEQ ID NO.: 45 Un polipéptido de Trichoderma reesei Bgl1 está preferentemente inalterado, con respecto al Bgl1 nativo, en los residuos D164, E197 y D267. Un polipéptido Trichoderma reesei Bgl1 preferentemente está inalterado en al menos el 70%, 80%, 90%, 95%, 98% o 99% de los residuos de aminoácidos que se conservan entre Trichoderma reesei Bgl1, Fv3A y Trichoderma reesei Bxl1, como se muestra en la alineación de la Figura 64. Un polipéptido *Trichoderma reesei* Bgl1 comprende adecuadamente todos los dominios conservados predichos de *Trichoderma reesei Bgl1* nativo que se muestra en la Figura 24. Un polipéptido *ejemplar Trichoderma reesei* Bgl1 comprende una secuencia teniendo al menos el 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con la secuencia madura del *Trichoderma reesei* Bgl1 que se muestra en la Figura 24. El polipéptido *Trichoderma reesei* Bgl1 de la invención preferentemente tiene actividad β-glucosidasa.

Por consiguiente, el actual acceso proporciona un número de polipéptidos hemiceluloiticos aislados, sintéticos, o recombinados o de variantes como las descritas a continuación:

(1) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos correspondientes a las posiciones (i) 24 a 766 de SEQ ID NO: 2; (ii) 73 a 321 de SEQ ID NO: 2; (iii) 73 a 394 de SEQ ID NO: 2; (iv) 395 a 622 de SEQ ID NO: 2; (v) 24 a 622 de SEQ ID NO: 2; o (iv) 73 a 622 de SEQ ID NO: 2; el polipéptido preferentemente tiene actividad de β- xilosidasa; o

(2) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos correspondientes a las posiciones (i) 21 a 445 de SEQ ID NO: 4; (ii) 21 a 301 de SEQ ID NO: 4; (iii) 21 a 323 de SEQ ID NO: 4; (iv) 21 a 444 de SEQ ID NO: 4; (v) 302 a 444 de SEQ ID NO: 4; (vi) 302 a 445 de SEQ ID NO: 4; (vii) 324 a 444 de SEQ ID NO: 4; o (viii) 324 a 445 de SEQ ID NO: 4; el polipéptido preferentemente tiene actividad de β-xilosidasa; o

500

(1) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos correspondientes a las posiciones (i) 19 a 530 de SEQ ID NO: 6; (ii) 29 a 530 de SEQ ID NO: 6; (iii) 19 a 300 de SEQ ID NO:6; (iv) 29 a 300 de SEQ ID NO: 6; el polipéptido preferentemente tiene actividad de β- xilosidasa; o

(1) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos correspondientes a las posiciones (i) 20 a 439 de SEQ ID NO: 8; (ii) 20 a 291 de SEQ ID NO: 8; (iii)145 a 291 de SEQ ID NO: 8; o (iv) 145 a 439 de SEQ ID NO: 8; el polipéptido preferentemente tiene actividad de - xilosidasa; o

(2) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos correspondientes a las posiciones (i) 23 a 449 de SEQ ID NO: 10; (ii) 23 a 302 de SEQ ID NO:10; (iii) 23 a 320 de SEQ ID NO: 10; (iv) 23 a 448 de SEQ ID NO:10; (v) 303 a 448 de SEQ ID NO:10; (vi) 303 a 449 de SEQ ID NO:10; (vii) 321 a 448 de SEQ ID NO:10; o (viii) 321 a 449 de SEQ ID NO:10; el polipéptido preferentemente tiene actividad de β- xilosidasa; o

(1) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos correspondientes a las posiciones (i) 17 a 574 de SEQ ID NO: 12; (ii) 27 a 574 de SEQ ID NO: 12; (iii) 17 a 303 de SEQ ID NO:12; (iv) 27 a 303 de SEQ ID NO: 12;el polipéptido preferentemente tiene actividad de - xilosidasa y L-arabinofuranosidasa; o

(7) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos correspondientes a las posiciones (i) 21 a 676 de SEQ ID NO: 14; (ii) 21 a 652 de SEQ ID NO: 14; (iii) 469 a 652 de SEQ ID NO: 14; o (iv) 469 a 676 de SEQ ID NO: 14; el polipéptido tiene preferentemente las dos actividades β- xilosidasa y L-α-arabinofuranosidasa; o

(8) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos correspondientes a las posiciones (i) 19 a 340 de SEQ ID NO: 16; (ii) 53 a 340 de SEQ

ID NO: 16; (iii) 19 a 383 de SEQ ID NO: 16; o (iv) 53 a 383 de SEQ ID NO: 16; el polipéptido preferentemente tiene actividad de β- xilosidasa; o

(9) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos correspondientes a las posiciones (i) 21 a 341 de SEQ ID NO: 18; (ii) 107 a 341 de SEQ ID NO: 18; (iii) 21 a 348 de SEQ ID NO: 18; o (iv) 107 a 348 de SEQ ID NO: 18; el polipéptido preferentemente tiene actividad de β- xilosidasa; o

(10) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos correspondientes a las posiciones (i) 15 a 558 de SEQ ID NO: 20; o (ii) 15 a 295 de SEQ ID NO: 20; el polipéptido tiene preferencia por la actividad L-α- arabinofuranosidasa; o

(11) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos correspondientes a las posiciones (i) 21 a 632 de SEQ ID NO: 22; (ii) 461 a 632 de SEQ ID NO: 22; (iii) 21 a 642 de SEQ ID NO: 22; o (iv) 461 a 642 de SEQ ID NO: 22; el polipéptido tiene actividad preferente L-α- arabinofuranosidasa ; o

(12) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia para la secuencia de aminoácidos correspondientes a las posiciones (i) 20 a 341 de SEQ ID NO: 28; (ii) 21 a 350 de SEQ ID NO: 28; (iii) 107 a 341 de SEQ ID NO: 28; o (iv) 107 a 350 de SEQ ID NO: 28; el polipéptido tiene actividad de β- xilosidasa; o

(13) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia para la secuencia de aminoácidos correspondientes a las posiciones (i) 21 a 660 de SEQ ID NO: 32; (ii) 21 a 645 de SEQ ID NO: 32; (iii) 450 a 645 de SEQ ID NO: 32; o (iv) 450 a 660 de SEQ ID NO: 32; el polipéptido tiene actividad preferente de L-α- arabinofuranosidasa.

La divulgación actual también proporciona composiciones (*p. ej.*, composiciones de celulasa, o mezclas de enzima/composiciones) o caldos de fermentación enriquecidas con uno o más de los polipéptidos anteriormente descritos. La mezcla/composición de la enzima es así una composición que ocurre de manera no

natural. La composición de la celulasa puede ser por ejemplo una composición fungicida filamentosa de la celulasa, como por ejemplo, una composición Trichoderma de la celulasa. El caldo de fermentación puede ser un caldo de fermentación de hongos filamentosos, por ejemplo, un *Trichoderma, Humicola, Fusarium, Aspergillus, Neurospora, Penicillium, Cefalosporium, Achlya, Podospora, Endothia, Mucor, Cochliobolus, Pyricularia,* o *caldo de fermentación Chrysosporium* . En particular, el caldo de fermentación puede ser, por ejemplo, uno de los *Trichoderma spp.* como un *Trichoderma reesei,* o *Penicillium spp,* como un *Penicillium funiculosum.* El caldo de la fermentación puede también convenientemente ser un caldo de fermentación sin células.

Además la divulgación instantánea proporciona células huésped que son recombinantes diseñadas para expresar un polipéptido previamente descrito. Las células huésped pueden ser, por ejemplo, células huésped fungicidas filamentosas, tales como *Trichoderma, Humicola, Fusarium, Aspergillus, Neuroespora, Penicillium, Cephalosporium, Achlya, Podospora, Endothia, Mucor, cochliobolus, Pyricularia,* o células de *Chrysosporium* . En particular, las células huésped pueden ser, por ejemplo, *una célula Trichoderma spp.* (tal como una célula de Trichoderma reesei), o *una célula del penicillium* (tal como una célula de Penicillium funiculosum), *una célula Aspergillus* (tal como célula de los *Aspergillus nidulans*), o una célula de *Fusarium* (tal como Fusarium verticilloides o célula de *Fusarium oxysporum*).

6.1.1 **Proteínas de fusión**

La presente divulgación también proporciona una proteína de fusión que incluye un dominio de una proteína de la presente divulgación conectado a uno o más segmentos de fusión, que son normalmente heterólogos a la proteína (es decir, derivado de una fuente diferente a la de la proteína de la divulgación). Los segmentos convenientes de la fusión incluyen, sin limitación, los segmentos que pueden realzar la estabilidad de una proteína, proporcionar otra actividad biológica deseable y/o facilitar la purificación de la proteína (*p- ej.,* por cromatografía de afinidad). Un segmento de fusión adecuado puede ser un dominio de cualquier tamaño que tiene la función deseada (por ejemplo, imparte mayor estabilidad, solubilidad, acción o actividad biológica; y/o simplifica la purificación de la proteína). Los segmentos de la fusión se pueden ensamblar a los términos amino y/o carboxyl de los dominios de una proteína de la presente divulgación. Los segmentos de la fusión pueden ser susceptibles a la

segmentación. Puede haber una cierta ventaja en tener esta susceptibilidad, por ejemplo, puede permitir la recuperación directa de la proteína en cuestión. Las proteínas de la fusión son producidas preferiblemente cultivando una célula recombinada transfectada con un ácido nucleico de la fusión que codifica una proteína, que incluye un segmento de la fusión unido al extremo terminal carboxilo o amino, o los segmentos de fusión unidos a los extremos terminales carboxilo y amino, de una proteína, o un dominio de eso.

En consecuencia, las proteínas de la presente divulgación también incluyen productos de expresión de fusiones de genes (por ejemplo, una forma sobreexpresada, soluble y activa de una proteína recombinante), de genes mutagenizados (por ejemplo, genes con modificaciones codón para mejorar la transcripción genética y traducción) y de genes truncados (por ejemplo, genes donde se han eliminado la secuencias de señal o sustituidos con una secuencia de señal heteróloga).

Las hidrolasas de glicólisis que utilizan los substratos insolubles son a menudo enzimas modulares. Normalmente comprenden módulos catalíticos anexados a uno o más dominios no catalíticos de enlace de carbohidratos (CBM). En la naturaleza, los CBM se piensan para promover la interacción de la hidrolasa de glicólisis con su polisacárido del substrato objetivo. Por lo tanto, la divulgación proporciona enzimas quiméricas habiendo alterado la especificidad de sustrato; incluyendo, por ejemplo, enzimas quiméricas con múltiples sustratos como resultado de CBM heterólogos "empalmados". Los CBM heterólogos de las enzimas quiméricas de la divulgación se pueden también diseñar para que sean modulares, de tal forma que estén anexados a un módulo catalítico o a un dominio catalítico (un "CD", *p. ej.*, en un sitio activo), que pueden además ser heterólogos u homólogos a la hidrolasa de glicólisis.

Por lo tanto, la divulgación proporciona péptidos y polipéptidos conformando o comprendiendo, módulos CBM/CD, que pueden ser emparejados homólogamente o unidos para formar pares quiméricos CBM/CD (heterólogos). Por lo tanto, estos péptidos y polipéptidos quiméricos pueden utilizarse para mejorar o alterar el rendimiento de una enzima de interés. En consecuencia, la divulgación proporciona enzimas quiméricas que comprenden, por ejemplo, al menos un CBM de una enzima de 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, o 45. Un polipéptido de la divulgación, por ejemplo, incluye una secuencia de aminoácidos que

comprende el CD y/o CBM de la secuencia de hidrolasa de glicólisis de 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, o 45. El polipéptido de la divulgación así puede ser convenientemente una proteína de fusión que comprende los dominios funcionales de dos o más proteínas diferentes (*por ejemplo,* un CBM de una proteína vinculada a un CD desde otra proteína).

Los polipéptidos de la divulgación convenientemente pueden ser obtenidos o utilizados "sustancialmente" puros. Por ejemplo, un polipéptido de la divulgación constituye por lo menos cerca del 80 wt.% *p. ej.,* por lo menos cerca de 85 wt.%, 90 wt. %, 91 wt.%, 92 wt.%, 93 wt.%, 94 wt.%, 95 wt.%, 96 wt.%, 97 wt.%, 98 wt.%, o 99 wt.%) de la proteína total en una composición dada, que también incluye otros ingredientes tales como un tampón o solución.

También, los polipéptidos de la divulgación pueden convenientemente ser obtenidos y/o utilizarse en caldos de cultivo recombinantes (*por ejemplo,* un caldo de cultivo de hongos filamentosos). Los caldos de cultivo recombinantes pueden ocurrir de manera no natural; por ejemplo, el caldo de cultivo puede presentarse por una célula recombinante que está diseñada para expresar un polipéptido heterólogo de la divulgación, o por una célula recombinante diseñada para expresar un polipéptido endógeno de la divulgación en mayor o menor cantidad que los niveles de expresión endógena (por ejemplo, en una cantidad que es 1-, 2 -3, 4, 5- o más veces mayor o menor que los niveles de expresión endógena). Además, los polipéptidos de la divulgación convenientemente pueden obtenerse o utilizarse como caldos de cultivo recombinante producidos por cepas de células "integradas" huésped diseñadas para expresar una pluralidad de polipéptidos de la divulgación en proporciones deseadas. Las relaciones deseadas ejemplares se describen, por ejemplo, en la Sección 6.3.4 a continuación.

6.2 Ácidos nucleicos y células huésped

La presente divulgación proporciona ácidos nucleicos codificando a un polipéptido de la divulgación, por ejemplo uno descrito en la Sección 6.1 anterior.

En otro aspecto, la divulgación proporciona ácidos nucleicos aislados, sintéticos o recombinantes que constan de *cerca del 70%, por ejemplo, por lo menos cerca del 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%, o el total (100%) de identidades de la secuencia con una secuencia del ácido nucleico de la 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, o 50, sobre una región de por lo menos cerca de 10, por ej. *por lo menos cerca de 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, o 2000, fracciones.* La presente divulgación también proporciona los ácidos nucleicos codificando al menos un polipéptido teniendo una actividad hemicelulolitica (*por ejemplo,* una xilanasa, β-xilosidasa y actividad de L-α-arabinofuranosidasa).

Los ácidos nucleicos de la divulgación también incluyen ácidos nucleicos aislados, sintéticos o recombinantes, codificando una enzima con la secuencia de 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, o 45 y su subsecuencias (*por ejemplo,* un dominio conservado o dominios de enlace de carbohidratos ("CBM") y sus variantes. Un ácido nucleico de la divulgación por ejemplo, puede codificar la parte madura de una proteína de 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, o 45.

La divulgación específicamente proporciona un ácido nucleico codificando un Fv3A, un Pf43A, un Fv43E, un Fv39A, un Fv43A, un Fv43B, un Pa51A, un Gz43A, un Fo43A, un Af43A, un Pf51A, un AfuXyn2, un AfuXyn5, un Fv43D, un Pf43B, un Fv43B, un Fv51A, un Trichoderma reesei Xyn3, un Trichoderma reesei Xyn2, un Trichoderma reesei Bxl1, o un polipéptido Trichoderma reesei Bgl1.

Por ejemplo, la divulgación proporciona una molécula de ácido nucleico aislada, donde la molécula de ácido nucleico codifica:

(1) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos correspondientes a las posiciones (i) 24 a 766 de SEQ ID NO: 2; (ii) 73 a 321 de SEQ ID NO: 2; (iii) 73 a 394 de SEQ ID NO: 2; (iv) 395 a 622 de SEQ ID NO: 2; (v) 24 a 622 de SEQ ID NO: 2; o (iv) 73 a 622 de SEQ ID NO: 2; o

306

(2) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos correspondientes a las posiciones (i) 21 a 445 de SEQ ID NO: 4; (ii) 21 a 301 de SEQ ID NO: 4; (iii) 21 a 323 de SEQ ID NO: 4; (iv) 21 a 444 de SEQ ID NO: 4; (v) 302 a 444 de SEQ ID NO: 4; (vi) 302 a 445 de SEQ ID NO: 4; (vii) 324 a 444 de SEQ ID NO: 4; o (viii) 324 a 445 de SEQ ID NO: 4; o

(3) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos correspondientes a las posiciones (i) 19 a 530 de SEQ ID NO: 6; (ii) 29 a 530 de SEQ ID NO: 6; (iii) 19 a 300 de SEQ ID NO:6; o (iv) 29 a 300 de SEQ ID NO: 6; o

(4) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos correspondientes a las posiciones (i) 20 a 439 de SEQ ID NO: 8; (ii) 20 a 291 de SEQ ID NO: 8; (iii)145 a 291 de SEQ ID NO: 8; o (iv) 145 a 439 de SEQ ID NO: 8; o

(5) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos correspondientes a las posiciones (i) 23 a 449 de SEQ ID NO: 10; (ii) 23 a 302 de SEQ ID NO:10; (iii) 23 a 320 de SEQ ID NO: 10; (iv) 23 a 448 de SEQ ID NO:10; (v) 303 a 448 de SEQ ID NO:10; (vi) 303 a 449 de SEQ ID NO:10; (vii) 321 a 448 de SEQ ID NO:10; o (viii) 321 a 449 de SEQ ID NO:10; o

(6) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos correspondientes a las posiciones (i) 17 a 574 de SEQ ID NO: 12; (ii) 27 a 574 de SEQ ID NO: 12; (iii) 17 a 303 de SEQ ID NO:12; o (iv) 27 a 303 de SEQ ID NO: 12; o

(7) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos correspondientes a las posiciones (i) 21 a 676 de SEQ ID NO: 14; (ii) 21 a 652 de SEQ ID NO: 14; (iii) 469 a 652 de SEQ ID NO: 14; o (iv) 469 a 676 de SEQ ID NO: 14; o

(8) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos

309

correspondientes a las posiciones (i) 19 a 340 de SEQ ID NO: 16; (ii) 53 a 340 de SEQ ID NO: 16; (iii) 19 a 383 de SEQ ID NO: 16; o (iv) 53 a 383 de SEQ ID NO: 16; o

(9) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos correspondientes a las posiciones (i) 21 a 341 de SEQ ID NO: 18; (ii) 107 a 341 de SEQ ID NO: 18; (iii) 21 a 348 de SEQ ID NO: 18; o (iv) 107 a 348 de SEQ ID NO: 18; o

(10) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos correspondientes a las posiciones (i) 15 a 558 de SEQ ID NO: 20; o (ii) 15 a 295 de SEQ ID NO: 20; o

(11) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos correspondientes a las posiciones (i) 21 a 632 de SEQ ID NO: 22; (ii) 461 a 632 de SEQ ID NO: 22; (iii) 21 a 642 de SEQ ID NO: 22; o (iv) 461 a 642 de SEQ ID NO: 22; o

(12) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia para la secuencia de aminoácidos correspondientes a las posiciones (i) 20 a 341 de SEQ ID NO: 28; (ii) 21 a 350 de SEQ ID NO: 28; (iii) 107 a 341 de SEQ ID NO: 28; o (iv) 107 a 350 de SEQ ID NO: 28; o

(13) un polipéptido compuesto por una secuencia de aminoácidos con al menos el 90%, 95%, 98%, 99% o 100% de identidad de secuencia para la secuencia de aminoácidos correspondientes a las posiciones (i) 21 a 660 de SEQ ID NO: 32; (ii) 21 a 645 de SEQ ID NO: 32; (iii) 450 a 645 de SEQ ID NO: 32; o (iv) 450 a 660 de SEQ ID NO: 32.

La divulgación instantánea también proporciona:

(1) un ácido nucleico que tiene por lo menos el 90% (por *ejemplo,* al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más) de identidad de secuencia SEQ ID NO: 1, o un ácido nucleico que es capaz de hibridizar bajo condiciones rigurosas con un complemento de SEQ ID NO: 1, o un fragmento; o

(2) un ácido nucleico que tiene por lo menos el 90% (por *ejemplo,* al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más) de identidad de secuencia con SEQ ID NO: 3, o un ácido nucleico que es capaz de hibridizar en condiciones rigurosas con un complemento de SEQ ID NO: 3, o con un fragmento; o

(3) un ácido nucleico que tiene por lo menos el 90% (por *ejemplo,* al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más) identidad de secuencia con SEQ ID NO: 5, o un ácido nucleico que es capaz de hibridizar en condiciones rigurosas con un complemento de SEQ ID NO: 5, o con un fragmento; o

(4) un ácido nucleico que tiene al menos el 90% (por *ejemplo,* al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más) identidad de secuencia con SEQ ID NO: 7, o un ácido nucleico capaz de hibridizar en condiciones rigurosas con un complemento de SEQ ID NO: 7, o con un fragmento; o

(5) un ácido nucleico que tiene al menos el 90% (por *ejemplo,* al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más) identidad de secuencia con SEQ ID NO: 9, o un ácido nucleico capaz de hibridizar en condiciones rigurosas con un complemento de SEQ ID NO: 9, o con un fragmento; o

(6) un ácido nucleico que tiene al menos el 90% (por *ejemplo,* al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más) identidad de secuencia con SEQ ID NO: 11 o un ácido nucleico capaz de hibridizar en condiciones rigurosas con un complemento de SEQ ID NO: 11, o con un fragmento; o

(7) un ácido nucleico que tiene al menos el 90% (por *ejemplo,* al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más) identidad de secuencia con SEQ ID NO: 13, o un ácido nucleico capaz de hibirizar en condiciones rigurosas con un complemento de SEQ ID NO: 13, o con un fragmento; o

(8) un ácido nucleico que tiene al menos el 90% (*por ejemplo,* al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más) identidad de secuencia con SEQ ID NO: 15, o un ácido nucleico capaz de hibridizar en condiciones rigurosas con un complemento de SEQ ID NO: 15, o con un fragmento; o

(9) un ácido nucleico que tiene al menos el 90% (por *ejemplo,* al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más) identidad de secuencia con SEQ ID

NO: 17 o un ácido nucleico capaz de hibridizar en condiciones rigurosas con un complemento de SEQ ID NO: 17, o con un fragmento; o

(10) un ácido nucleico que tiene al menos el 90% (por *ejemplo,* al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más) identidad de secuencia con SEQ ID NO: 19, o un ácido nucleico capaz de hibridizar en condiciones rigurosas con un complemento de SEQ ID NO: 19, o con un fragmento; o

(11) un ácido nucleico que tiene al menos el 90% (*por ejemplo,* al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más) identidad de secuencia con SEQ ID NO: 21, o un ácido nucleico capaz de hibridizar en condiciones rigurosas con un complemento de SEQ ID NO: 21, o con un fragmento; o

(12) un ácido nucleico que tiene al menos el 90% (*por ejemplo,* al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más) identidad de secuencia con SEQ ID NO: 27 o un ácido nucleico capaz de hibridizar en condiciones rigurosas con un complemento de SEQ ID NO: 27, o con un fragmento; o

(13) un ácido nucleico que tiene al menos el 90% (*por ejemplo,* al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más) de identidad de secuencia con SEQ ID NO: 31, o un ácido nucleico capaz de hibridizar en condiciones rigurosas con un complemento de SEQ ID NO: 31, o con un fragmento de ello.

La divulgación también proporciona casetes de expresión y/o vectores de los ácidos nucleicos descritos anteriormente.

Convenientemente, el ácido nucleico que codifica una enzima de la divulgación está unida operativamente a un promotor. En concreto, donde se desea la expresión recombinante en un huésped de hongo filamentoso, el promotor puede ser un promotor de hongo filamentoso. Los ácidos nucleicos pueden estar, por ejemplo, bajo el control de promotores heterólogos. Los ácidos nucleicos se pueden también expresar bajo el control de promotores constitutivos o inducibles. Los ejemplos de los promotores que pueden ser utilizados incluyen, pero no se limitan a un promotor del celulasa, promotor del xianasa, el promotor 1818 (identificado previamente como proteína altamente expresada mapeando EST la *Trichoderma*). Por ejemplo, el promotor puede convenientemente ser una celobiohidrolasa, una endoglucanasa, o un promotor del β-glucosidasa. Un promotor particularmente adecuado puede ser, por ejemplo, un

promotor *T. reesei* celobiohidrolasa, endoglucanasa o β-glucosidasa. Por ejemplo, el promotor es un promotor de la celobiohidrolasa I (*cbh*1). Los ejemplos no limitantes de promotores incluyen un promotor *cbh1, cbh2, egl1, egl2, egl3, egl4, egl5, pki1, gpd1, xyn1* o *xyn2*. Los ejemplos no-limitadores adicionales de promotores incluyen un promotor *T. reesei cbh1, cbh2, egl1, egl2, egl3, egl4, egl5, pki1, gpd1, xyn1* o *xyn2* .

La actual divulgación proporciona células huésped diseñadas para expresar una o más de las enzimas de la divulgación. Las células huésped convenientes incluyen las células de cualquier microorganismo (células *p. ej.*, de una bacteria, un protist, un alga, un hongo (*p. ej.*, una levadura o un hongo filamentoso), u otro microbio), y son preferentemente células de una bacteria, de levadura o de un hongo filamentoso.

Las células huésped adecuadas de géneros bacterianos incluyen, pero no se limita a las células de Escherichia, Bacillus, Lactobacillus, Pseudomonas y Streptomyces. Las células convenientes de especies bacterianas incluyen, pero no se limitan a las células *de Escherichia coli*, *bacilo subtilis*, *bacilo licheniformis*, *lactobacilo*brevis, aeruginosa de los Pseudomonas , *lividans* de los Streptomyces .

Las células huésped convenientes de los géneros de levadura incluyen, pero no se limitan a las células *del Saccharomyces*, *Schizosaccharomyces*, *Candida* , *Hansenula*, *Pichia*, *Kluyveromyces* y *Ph affia*. Las células convenientes de las especies de la levadura incluyen, pero no se limitan a *células del Saccharomyces*cerevisiae, Schizosaccharomyces pombe, *Candida albicans*, Hansenula polymorpha, *Pichia pastoris*, *P.* canadensis, Kluyveromyces marxianus y Phaffia rhodozyma.

Las células huésped convenientes de hongos filamentosos incluyen todas las formas filamentosas de la subdivisión *Eumycotina*. Las células convenientes de géneros fungicidas filamentosos incluyen, pero no se limitan a las células de *Acremonium, Aspergillus, Aureobasidium, Bjerkandera, Ceriporiopsis, Chrysoporium, Coprinus, Coriolus, Corynascus, Chaetomium, Cryptococcus, Filobasidium, Fusarium,Gibberella, Humicola, Hypocrea, Magnaporthe, Mucor, Myceliophthora, Mucor, Neocallimastix, Neuroespora, Paecilomyces, Penicillium, Phanerochaete, Phlebia, Piromyces, Pleurotus, Scytaldium, Schizophyllum, Sporotrichum, Talaromyces, Thermoascus, Thielavia, Tolypocladium, Trametes,* y *Trichoderma.* Las células convenientes pueden también incluir las células de varias formas anamorfas y teleomorfas para estos géneros fungicidas filamentosos.

311

Células convenientes de las especies fungicidas filamentosas incluyen, pero no están limitadas a las células de *Aspergillus awamori, Aspergillus fumigatus, Aspergillus foetidus, Aspergillus japonicus, Aspergillus nidulans, Aspergillus niger, Aspergillus oryzae, Chrysosporium lucknowense, Fusarium bactridioides, Fusarium cerealis, Fusarium crookwellense, Fusarium culmorum, Fusarium graminearum, Fusarium graminum, Fusarium heterosporum, Fusarium negundi, Fusarium oxysporum, Fusarium reticulatum, Fusarium roseum, Fusarium sambucinum, Fusarium sarcochroum, Fusarium sporotrichioides, Fusarium sulphureum, Fusarium torulosum, Fusarium trichothecioides, Fusarium venenatum, Bjerkandera adusta, Ceriporiopsis aneirina, Ceriporiopsis aneirina, Ceriporiopsis caregiea, Ceriporiopsis gilvescens, Ceriporiopsis pannocinta, Ceriporiopsis rivulosa, Ceriporiopsis subrufa, Ceriporiopsis subvermispora, Coprinus cinereus, Coriolus hirsutus, Humicola insolens, Humicola lanuginosa, Mucor miehei, Myceliophthora thermophila, Neurospora crassa, Neurospora intermedia, Penicillium purpurogenum, Penicillium canescens, Penicillium solitum, Penicillium funiculosum, Phanerochaete chrysosporium, Phlebia radiate, Pleurotus eryngii, Talaromyces flavus, Thielavia terrestris, Trametes villosa, Trametes versicolor, Trichoderma harzianum, Trichoderma koningii, Trichoderma longibrachiatum, Trichoderma reesei,* y *Trichoderma viride.*

La divulgación adicionalmente proporciona una célula recombinante que está diseñada para expresar uno o más, dos o más, tres o más, cuatro o más, cinco o más de un Fv3A, Pf43A, Fv43E, Fv39A, Fv43A, Fv43B, Pa51A, Gz43A, Fo43A, Af43A, Pf51A, AfuXyn2, AfuXyn5, Fv43D, un Pf43B y un polipéptido Fv51A. La célula huésped recombinante es, por ejemplo, una célula huésped recombinante de Trichoderma reesei . En un ejemplo particular, la divulgación proporciona un hongo recombinado, tal como una Trichoderma reesei recombinada, que se manipula para expresar uno o más, dos o más, tres o más, cuatro o más, o cinco o más de un polipéptido Fv3A, Pf43A, Fv43E, Fv39A, Fv43A, Fv43B, Pa51A, Gz43A, Fo43A, Af43A, Pf51A, AfuXyn2, AfuXyn5, Fv43D, Pf43B, y de un de Fv51A. La divulgación proporciona una célula huésped recombinante Trichoderma reesei diseñada para expresar 1, 2, 3, 4, 5, o más de un polipéptido Fv3A, Pf43A, Fv43E, Fv39A, Fv43A, Fv43B, Pa51A, Gz43A, Fo43A, Af43A, Pf51A, AfuXyn2, AfuXyn5, Fv43D, un Pf43B y un Fv51A.

La divulgación proporciona una célula huésped, por ejemplo, una célula fúngica recombinante u hongos filamentosos recombinante, diseñada para expresar

31C

recombinantemente al menos una xilanasa, al menos una β-xilosidasa y una L-α-arabinofuranosidasa. La divulgación también proporciona una célula huésped recombinante, *p. ej.,* una célula huésped fungicida recombinante o un hongo filamentoso como un *Trichoderma reeseirecombinante*, que está diseñado para expresar 1, 2, 3, 4, 5 o más de un polipéptido Fv3A, Pf43A, Fv43E, Fv39A, Fv43A, Fv43B, Pa51A, Gz43A, Fo43A, Af43A, Pf51A, AfuXyn2, AfuXyn5, Fv43D, Pf43B y un Fv51A, además de uno o más de *Trichoderma reesei* Xyn2 *Trichoderma reesei* Xyn3, *Trichoderma reesei Bxl1* y/o *Trichoderma reesei Bgl1* . La célula huésped recombinante es, por ejemplo, una célula huésped recombinante de Trichoderma reesei . El hongo recombinante es, por ejemplo, un Trichoderma reesei*recombinante* .La divulgación también proporciona una célula huésped Trichoderma reesei o un hongo Trichoderma reesei recombinante, que está diseñado para expresar recombinantemente 1, 2, 3, 4, 5 o más de un polipéptido Fv3A, Pf43A, Fv43E, Fv39A, Fv43A, Fv43B, Pa51A, Gz43A, Fo43A, Af43A, Pf51A, AfuXyn2, AfuXyn5, Fv43D, Pf43B y un Fv51A, además de uno o más de expresar recombinantemente uno o más de Trichoderma reesei Xyn2 Trichoderma reesei Xyn3, Trichoderma reesei Bxl1 y/o Trichoderma reesei Bgl1 .

La presente divulgación también proporciona una célula huésped recombinante por ejemplo, una célula fúngica recombinante o un organismo recombinante, por ejemplo, un hongo filamentoso, como un Trichoderma reesei recombinante que está diseñado para expresar recombinantemente polipéptidos Trichoderma reesei Xyn3, Trichoderma reesei Bgl1, Fv3A, Fv43D y Fv51A. Por ejemplo, la célula huésped recombinante adecuada es una célula huésped de *de* Trichoderma reesei. El hongo recombinante es convenientemente una Trichoderma reesei recombinante. La divulgación proporciona, por ejemplo, una célula huésped *Trichoderma reesei* diseñada para *expresar recombinadamente polipéptidos Trichoderma reesei* Xyn3 *Trichoderma reesei* Bgl1, Fv3A, Fv43D y Fv51A.

Además la divulgación proporciona una célula huésped recombinante u hongo recombinante diseñado para expresar una mezcla de enzima compuesta por enzimas adecuadas en proporciones adecuadas para la sacarificación. La célula huésped recombinante es, por ejemplo, una célula huésped fúngica. El hongo recombinante es, por ejemplo, una Trichoderma reesei*recombinante* . Los cocientes/cantidades ejemplares de la enzima presentes en mezclas convenientes de la enzima se describen en la Sección 6.3.4 a continuación.

La divulgación proporciona además plantas transgénicas que comprenden un ácido nucleico de la divulgación o un casete de expresión de la divulgación. La planta transgénica puede ser, por ejemplo, una planta de cereal, de maíz, de patata, de tomate, de trigo, de gárgola, de canola, de soja, de arroz, de cebada, o una planta de tabaco.

6.3 Mezclas de enzimas **para** sacarificación

La actual divulgación proporciona una composición que comprende una mezcla/composición de enzima capaz de romper el material lignoceluloso. Tal mezcla/composición multienzimática comprende al menos un polipéptido de la presente divulgación, en combinación con uno o más polipéptidos adicionales de la presente divulgación, o con una o más enzimas de otros microorganismos, plantas u organismos. Se contemplan combinaciones de enzima sinérgica y métodos relacionados. La divulgación incluye métodos para identificar las proporciones óptimas de las enzimas incluidas en las mezclas/composiciones para degradar un determinado material lignocelulósicos. Estos métodos incluyen *p. ej., pruebas* para identificar la mezcla/composición de enzima optimas y cocientes para la conversión eficiente de un substrato lignocelulósico dado a sus azúcares constitutivos. Los siguientes ejemplos incluyen ensayos que pueden utilizarse para identificar las proporciones óptimas y mezclas y composiciones de enzimas a los cuales se degradan los materiales lignocelulósicos.

6.3.1 Antecedentes

Las paredes celulares de las plantas superiores se componen de una variedad de componentes de polímeros de carbohidratos (CP). Estos CP interactúan en medios covalentes y no-covalentes, proporcionando la integridad estructural requerida para formar membranas celulares rígidas y para resistir la presión de turgencia de las plantas. El CP principal encontrado en plantas es la celulosa, que forma la espina dorsal estructural de la pared de la célula. Durante la biosíntesis de la celulosa, las cadenas de poly-β-1,4-D-glucosa se auto asocian con la vinculación del hidrógeno e interacciones hidrofóbicas para formar los microfibrils de celulosa, que más adelante se auto asocian para formar fibrillas más grandes. Las microfibrillas de celulosa son a menudo estructuralmente irregulares y contienen regiones de cristalinidad variante. El grado de cristalinidad de las fibrillas de celulosa depende de que tan firmemente está ordenada la vinculación del hidrógeno entre sus cadenas de componentes de la celulosa. Las áreas

con menos vinculación ordenada y por lo tanto cadenas de glucosa más accesibles, se refieren como regiones amorfas.

El modelo general de despolimerización de celulosa en glucosa implica un mínimo de tres actividades enzimáticas distintas. Las endoglucanasas segmentan las cadenas de la celulosa internamente a cadenas más cortas en un proceso que aumenta el número de los extremos accesibles, que son más susceptibles a la actividad de la exoglucanasa que las cadenas intactas de la celulosa. Estas exoglucanasas (p. ej., celobiohidrolasas) son específicos para reducir extremos o extremos no reductores, liberando, en la mayoría de los casos, la celobiosa, el dímero de la glucosa. La celobiosa acumulada entonces esta sujeta a segmentación por celobiasas (*por ejemplo,* β-1,4-glucosidasas) a la glucosa.

La celulosa contiene solamente anhidro-glucosa. En cambio, la hemicelulosa contiene un número de diversos monómeros de azúcar. Por ejemplo, aparte de la glucosa, los monómeros de azúcar en la hemicelulosa pueden también incluir xilosa, manosa, galactosa, ramnosa y arabinosa. Las Hemicelulosas contienen sobre todo azúcares D-pentosa y de vez en cuando pequeñas cantidades de L-azúcares. La Xilosa está típicamente presente en la cantidad más grande, pero el ácido manuronico y el ácido galacturonico también tienden a estar presentes. Las Hemicelulosas incluyen xilano, glucuronoxilano, arabinoxilano, glucomanano y xiloglucano.

Las enzimas y composiciones multienzimáticas de la divulgación son útiles para la sacarificación de materiales de hemicelulosa, incluyendo, *p. ej.,* xilano, arabinoxilano y sustratos que contengan xilano o arabinoxilano. El Arabinoxilano es un polisacárido integrado por la xilosa y la arabinosa, donde los residuos de L-α - arabinofuranosa se unen como puntos de bifurcación a un β- (1.4) - espina dorsal polimérica ligada a la xilosa.

La mayoría de las fuentes de la biomasa son algo complejas, conteniendo celulosa, hemicelulosa, pectina, lignina, proteína y ceniza, entre otros componentes. En consecuencia, en ciertos aspectos, la actual divulgación proporciona mezclas y composiciones de enzimas que contienen enzimas que imparten un rango o una variedad de características del sustrato al trabajar juntos para degradar la biomasa en azúcares fermentables de la manera más eficiente. Un ejemplo de una mezcla/composición de la multi-enzimas de la presente invención es una mezcla de

315

celobiohidrolasas, xilanasas, endoglucanases, β-glucosidasas, β-xilosidasas y, opcionalmente proteínas accesorias. La mezcla/composición de la enzima es así una composición que ocurre convenientemente de manera no natural.

Por consiguiente, la divulgación proporciona mezclas y composiciones de enzima (incluidos los productos de fabricación) que comprende una mezcla de xilano hidrolizante, hemicelulosa o enzimas hidrolizantes de celulosa, que incluyen al menos uno, varios o todos de una celulasa, incluyendo una glucanasa; una celobiohidrolasa; una L-α-arabinofuranosidasa; una xilanasa; una β-glucosidasa y una β-xilosidasa. Preferentemente cada una de las mezclas y composiciones de enzima de la divulgación comprende al menos una enzima de la divulgación. La mezcla/composición de la enzima también proporciona mezclas/composiciones que ocurren convenientemente de manera no natural. Como se utiliza en el presente documento, el término "mezclas/composiciones de enzima" se refiere a:

(1) una composición hecha combinando las enzimas componentes, bien sea en forma de caldo de fermentación o aislada parcial o totalmente o purificada;

(2) una composición producida por un organismo modificado para expresar una o más enzimas componentes; en ciertas realizaciones el organismo utilizado para expresar una o más de las enzimas componentes puede ser modificada para eliminar uno o más genes; en algunas otras realizaciones, el organismo utilizado para expresar uno o más componentes de enzimas pueden incluir más proteínas que afectan la hidrólisis de xilano, hidrólisis de hemicelulosa y/o hidrólisis de la celulosa;

(3) una composición combinando las enzimas componentes simultáneamente, por separado, o secuencialmente durante una reacción de sacarificación o fermentación; y

(4) una mezcla de la enzima producida in situ, *p. ej.,* durante una reacción de la sacarificación o de la fermentación;

(5) una composición producida de acuerdo con alguna o todas las anteriores (1)-(4).

El término "caldo de fermentación" como es usado aquí se refiere a la preparación de una enzima producida por la fermentación que sufre ninguna o mínima recuperación y/o purificación posterior a la fermentación. Por ejemplo, los cultivos microbianos se cultivan a saturación, incubados bajo condiciones limitantes de carbono para permitir la síntesis de proteínas (*p. ej.,* expresión de enzimas). Entonces, una vez que las enzimas se secreten en los medios de cultivo de la célula, se pueden utilizar los caldos de la fermentación. Los caldos de la fermentación del acceso pueden incluir el contenido sin fraccionar o fraccionado de los materiales de la fermentación derivados al final de la fermentación. Por ejemplo, los caldos de la fermentación de la invención no son fraccionados y comprenden el medio de cultivo gastado y el ripio de la célula presente después de las células microbianas (p. ej., células fungicidas filamentosas) experimenta un proceso de fermentación. El caldo de la fermentación puede contener convenientemente los medios de cultivo gastados de la célula, las enzimas extracelulares y células microbianas vivas o muertas. Alternativamente, los caldos de la fermentación se pueden fraccionar para retirar las células microbianas. En esos casos, los caldos de fermentación pueden, por ejemplo, comprenden los medios de cultivo celular gastado y las enzimas extracelulares.

Cualquiera de las enzimas específicamente descritas pueden combinarse con una o más de las enzimas descritas en este documento o con cualquiera otras enzimas disponibles y adecuadas, para producir una adecuada mezcla/composición multienzimática. La divulgación no está restringida o limitada a las combinaciones específicas de los ejemplares enumerados a continuación.

6.3.2 Biomasa

La divulgación proporciona métodos y procesos de sacarificación de la biomasa, utilizando enzimas, mezclas y composiciones de enzima de la divulgación. El término "biomasa", en este documento, se refiere a cualquier composición de celulosa y hemicelulosa (opcionalmente también lignina en materiales de biomasa lignocelulósicos). Tal como se utiliza aquí, la biomasa incluye sin limitación, semillas , granos, tubérculos, desechos de plantas o subproductos de procesador de alimentos o transformación industrial (p. ej. tallos), maíz (incluyendo, mazorcas, rastrojos y similares), hierbas (hierba india, tal como Sorghastrum nutans; o, pasto mijo, por ej. especies Panicum como Panicum virgatum), madera (incluyendo, virutas de madera, desechos procesados), papel, pulpa, papel reciclado (por ej. papel periódico y

similares). Otros materiales de la biomasa incluyen, sin limitación, patatas, soja (p. ej. canola) cebada, centeno, avena, trigo, remolachas y bagazo de la caña de azúcar.

En otro aspecto, la divulgación proporciona métodos de sacarificación que comprende contactar una composición que consta de biomasa, por ejemplo un material que comprende xilano, hemicelulosa, celulosa y/o azúcar fermentada con un polipéptido de la divulgación, o un polipéptido codificado por un ácido nucleico de la divulgación, o cualquiera de las mezclas/composiciones de enzima o productos de fabricación de la divulgación.

La biomasa sacarificada (p ej., el material lignocelulósico transformado por las enzimas de la divulgación) puede convertirse en una serie de productos de base biológica vía procesos tales como, por ejemplo, fermentación microbiana y/o productos químicos de síntesis. Según se utiliza en el presente documento, "fermentación microbiana" se refiere a un proceso de cultivo y cosecha de fermentación de microorganismos bajo las condiciones adecuadas. El microorganismo de la fermentación puede ser cualquier microorganismo conveniente para el uso en un proceso de fermentación deseado para la producción de productos de base biológica. Los microorganismos convenientes de la fermentación incluyen, sin limitación, hongos filamentosos, levadura y bacterias. La biomasa sacarificada puede ser convertida, por ejemplo, en un combustible (*p. ej.,* un biocombustible como un bioetanol, biobutanol, biometanol, biopropanol, biodiesel, jet fuel o similares) *a través la fermentación y/o síntesis química* . La biomasa sacarificada también puede convertirse, por ejemplo, en un producto químico (p. ej., ácido ascórbico, isopreno, propanediol 1.3), lípidos, aminoácidos, proteínas y enzimas, *vía la fermentación* y/o síntesis química.

6.3.3. Pretratamiento

Antes de la sacarificación, la biomasa (por *ejemplo,* materiales lignocelulósicos) es preferentemente sometida a uno o más pasos de pretratamiento para procesar el xilano, la hemicelulosa, la celulosa y el material de lignina más accesible o susceptible a las enzimas y así más susceptible a la hidrólisis por la enzima(s) y/o mezclas/composiciones de enzimas de la divulgación.

En una realización ejemplar, el pretratamiento conlleva someter un material de biomasa a un catalizador que comprende una solución diluida de un ácido fuerte y una

sal de metal en un reactor. El material de biomasa puede, *p. ej.,* ser una materia prima o un material seco. Este tratamiento previo puede bajar la energía de activación, o la temperatura de la hidrólisis de la celulosa, permitiendo en última instancia producciones más altas de azúcares fermentables. Véase, *p. ej.,* Patente de los Estados Unidos Nos. 6.660.506; 6.423.145.

Otro método ejemplar del tratamiento previo a conlleva hidrolizar la biomasa sometiendo el material de la biomasa a una primera hidrólisis en un medio acuoso a la temperatura y presión elegidas para efectuar principalmente la despolimerización de la hemicelulosa sin lograr una despolimerización significativa de celulosa en la glucosa. Este paso rinde una mezcla en la cual la fase acuosa líquida contiene monosacáridos disueltos resultando de la despolimerización de la hemicelulosa y una fase sólida que contiene la celulosa y la lignina. La mezcla está entonces sujeta a un segundo paso de la hidrólisis bajo condiciones que permitan que una porción importante de la celulosa sea despolimerizada, rindiendo una fase líquida acuoso que contiene productos de celulosa de despolimerización disuelta/soluble. Véase, *p.ej.,* Patente de los E.E.U.U. No. 5.536.325.

El método ejemplar adicional implica procesar un material de biomasa por unas o más etapas de hidrólisis del ácido diluido usando cerca de 0.4% a cerca de 2% de un ácido fuerte; seguido por el tratamiento del componente lignocelulósico sólido sin reaccionar del material hidrolizado ácido con la deslignificación alcalina. Véase, *p. ej.,* Patente de los E.E.U.U. No. 6.409.841.

Otro método de pretratamiento ejemplar comprende biomasa prehidrolizante (*por ejemplo, materiales de* lignocelulósicos) en un reactor de prehidrolisis; añadir un líquido ácido a los materiales lignocelulósicos sólidos para hacer una mezcla; calentar la mezcla a temperatura de reacción; mantener la temperatura de reacción por un período de tiempo suficiente para fraccionar el material lignocelulósico en una parte soluble que contengan al menos un 20% de la lignina de los materiales lignocelulósicos y una fracción sólida que contienen celulosa; separar la parte solubilizada de la fracción sólida y eliminando la parte solubilizada mientras se llega a la de temperatura de reacción; y recuperando la parte solubilizada. La celulosa en la fracción sólida se hace más favorable para la digestión enzimática. Véase, *p.ej.,* Patente de los E.E.U.U. No. 5.705.369.

Otros métodos de pretratamiento pueden implicar el uso de peróxido de hidrógeno H2O2. Véase Gould, 1984, Biotech y Bioengr. 26:46-52.

El Pretratamiento puede incluir también contactar un material de biomasa con cantidades estequiométricas de hidróxido de sodio y el hidróxido de amónico en una concentración muy baja. Véase a Teixeira *y otros.*, 1999, Appl. Biochem.and Biotech. 77-79: 19-34.

El Pretratamiento también puede incluir contactar una lignocelulosa con un producto químico (*p. ej.,* una base, como el carbonato de sodio o hidróxido de potasio) a un pH entre 9 y 14, a temperatura, presión y pH moderados. Véase la publicación PCT WO2004/081185.

El amoníaco se utiliza, por ejemplo, en un método de pretratamiento preferido. Dicho método de pretratamiento consiste en someter un material de biomasa a una baja concentración de amoniaco bajo condiciones de sólidos altos. Véase, *p. ej.,* Patente de los E.E.U.U. No. 20070031918 y publicación PCT WO 06110901.

6.3.4 Mezclas de enzimas ejemplares

La actual divulgación proporciona mezclas/composiciones de enzimas que comprenden una o más enzimas de la divulgación. Una o más enzimas de las mezclas/composiciones de la enzima se pueden producir por una célula huésped recombinante o por un organismo recombinante. Las mezclas/composiciones de la enzima son composiciones convenientes que no ocurren de manera natural.

Una mezcla/composición de la enzima de la divulgación puede constar adecuadamente de un primer polipéptido con actividad β-xilosidasa y comprender además de 1, 2, 3 o 4 de un segundo polipéptido con actividad β-xilosidasa, uno o más polipéptidos con actividad L-α-arabinofuranosidasa, uno o más polipéptidos con actividad xilanasa y uno o más polipéptidos con actividad celulasa. El primer polipéptido que tiene actividad de β-xilosidasa es, por ejemplo, un Fv3A, un Pf43A, un Fv43E, un Fv43A, un Fv43B, un Pa51A, un Gz43A, un Fo43A, o un polipéptido de Fv39A. El segundo polipéptido que tiene actividad de β-xilosidasa, si está presente, es, por ejemplo, diferente del primer polipéptido que tiene actividad de β-xilosidasa y es convenientemente un Fv3A, un Pf43A, un Fv43E, un Fv43A, un Fv43B, un Pa51A, un Gz43A, un Fv43D, un Fo43A, un Fv39A, o un polipéptido de Trichoderma reesei *Bxl1* .

320

Cada uno o más polipéptidos con actividad de L-α-arabinofuranosidasa, si presente, es, por ejemplo, un Af43A, un Pf51A, un Pa51A, un Fv43B, o un polipéptido de Fv51A. Cada uno o más polipéptidos con actividad de xilanasa es, por ejemplo, un Trichoderma reesei *Xyn, un Trichoderma reesei* Xyn2, un AfuXyn2 o un AfuXyn5 . Cada uno o más polipéptidos con actividad de cellulasa, si está presente, es, por ejemplo, una endoglucanasa, por ejemplo, una Trichoderma reesei *EG1* o EG2, una celobiohidrolasa, por ejemplo, una Trichoderma reesei *CBH1 o CBH2, o una β-glucosidasa, por ejemplo, una Trichoderma reesei* Bgl1.

Otra mezcla/composición de la enzima de la divulgación puede constar adecuadamente de un primer polipéptido con actividad L-α-arabinofuranosidasa y comprender además 1, 2, 3 o 4 de un segundo polipéptido con actividad L-α-arabinofuranosidasa, uno o más polipéptidos con actividad de β-xilosidasa, uno o más polipéptidos con actividad xilanasa y uno o más polipéptidos con actividad celulasa. La primera L-α-arabinofuranosidasa es un Af43A, un Pf51A, o un polipéptido de Fv51A. La segunda L-α-arabinofuranosidasa es diferente de la primera L-α-arabinofuranosidasa, y es, por ejemplo, un Af43A, un Pf51A, un Pa51A, un Fv43B, o un polipéptido de Fv51A. Cada una o más β-xilosidasas es, por ejemplo, un Fv3A, un Pf43A, un Fv43E, un Fv43A, un Fv43B, un Pa51A, un Fv43D, un Gz43A, un Fo43A, un Fv39A, o un polipéptido de Trichoderma reesei *Bxl1* . En ciertas realizaciones cada uno o más xilanasas es una Trichoderma reesei Xyn3, una Trichoderma reesei Xyn2, un AfuXyn2 o un polipéptido AfuXyn5 . En ciertas realizaciones cada uno o más celulasas es independientemente una endoglucanasa, por ejemplo, una Trichoderma reesei EG1 o EG2, un polipéptido, una celobiohidrolasa, por ejemplo, un polipéptido Trichoderma reesei CBH1 o CBH2, o una β-glucosidasa, por ejemplo, un polipéptido Trichoderma reesei Bgl1.

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

321

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.

Xilanasas: La xilanasa constituye convenientemente cerca de 0.05 wt.% hasta cerca del 75 wt.% de las enzimas en una mezcla/composición de la enzima de la divulgación (*es decir,* el porcentaje de xilanasas es un porcentaje del peso relativo al peso de todas las proteínas en la composición) o de una base relativa del peso (*es decir,* donde la xilanasa es un porcentaje del peso relativo al peso combinado de xilanasas, de β-xylosidasas, de celulasas, de L-α-arabinofuranosidasas y de proteínas accesorias). La relación de cualquier par de proteínas respecto a los demás puede calcularse fácilmente en las mezclas y composiciones de enzima de la divulgación. Se contemplan mezclas/composiciones que comprenden enzimas en cualquier relación de peso derivable de los porcentajes del peso divulgados aquí. El contenido de xilanasa puede ser en un rango donde el límite más bajo es 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.%, o 50 wt.% del peso total de enzimas en la mezcla/composición de la enzima y el límite superior es 10 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 35 wt.%, 40 wt.%, 50 wt.%, 55 wt.%, 60 wt.%, 65 wt.%, 70 wt.%, o 75 wt. % del peso total de enzimas en la mezcla/composición de la enzima. Una única o más xilanasas en una mezcla/composición de la enzima puede representar, por ejemplo, 5 wt.% a 20 wt.%, 10 wt.% a 15 wt.%, o 15 wt.% a 25 wt.% de las enzimas totales en la mezcla/composición de la enzima. Las xilanasas ejemplares adecuadas para ser incluidas en las mezclas/composiciones de la enzima de la divulgación se describen en la Sección 6.3.6 a continuación.

L-α-arabinofuranosidasas: La L-α-arabinofuranosidasa constituye convenientemente cerca del 0.05 wt.% hasta cerca del 75 wt.% del total de enzimas en una mezcla/composición de la enzima dada (*es decir,* el porcentaje de L-α-

322

arabinofuranosidasas es un porcentaje del peso relativo al peso de todas las proteínas en la mezcla/composición) o de una base relativa del peso (*es decir,* donde el porcentaje de la L-α-arabinofuranosidasa es un porcentaje del peso relativo al peso combinado de xilanasas, de β-xylosidasas, de celulasas, de L-α-arabinofuranosidasas y de proteínas accesorias). La relación de cualquier par de proteínas respecto a las demás puede calcularse fácilmente con base en la divulgación. Se contemplan mezclas/composiciones que comprenden enzimas en cualquier relación de peso derivable de los porcentajes del peso divulgados aquí. El contenido de L-α-arabinofuranosidasas puede ser en un rango donde el límite inferior es 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.%, o 50 wt.% del peso total de enzimas en la mezcla/composición de la enzima y el límite superior es 10 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 35 wt.%, 40 wt.%, 50 wt.%, 55 wt.%, 60 wt.%, 65 wt.%, 70 wt.%, o 75 wt.% del peso total de enzimas en la mezcla/composición de la enzima. Por ejemplo, la única o más L-α-arabinofuranosidasa(s) puede representar convenientemente 2 wt.% a 25 wt.%; 5 wt.% a 20 wt.%; o 5 wt.% a 10 wt.% del peso total de enzimas en la mezcla/composición. Las L-α-arabinofuranosidasas ejemplares adecuadas para ser incluidas en las mezclas/composiciones de la enzima de la divulgación se describen en la Sección 6.3.8 a continuación.

β-Xylosidasas: La β-xylosidasa constituye convenientemente cerca de 0.05 wt.% a cerca de 75 wt.% del peso total de enzimas en una mezcla/composición de la enzima. La relación de cualquier par de proteínas respecto a las demás puede calcularse fácilmente con base en esta divulgación. Se contemplan mezclas/composiciones que comprenden enzimas en cualquier relación de peso derivable de los porcentajes del peso divulgados aquí. El contenido de La β-xylosidasa puede ser en un rango donde el límite inferior es cercano a 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.%, o 50 wt.% del peso total de enzimas en la mezcla/composición de la enzima y el límite superior es cercano a 10 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 35 wt.%, 40 wt.%, 50 wt.%, 55 wt.%, 60 wt.%, 65 wt.%, 70 wt.%, o 75 wt.% del peso total de enzimas en la mezcla/composición de la enzima. Por ejemplo, las β-xilosidasas pueden representar cerca del 0.05 wt.% a cerca del 75 wt.% del peso total de enzimas en la mezcla/composición. También, las β-xilosidasas pueden representar el 0.05 wt.% a cerca del 70 wt.%, cerca de 1 wt.% a cerca de 65 wt.%, cerca

de 1 wt.% a cerca de 60 wt.%, cerca de 2 wt.% a cerca de 55 % peso, cerca de 3 wt.% a cerca de 50 wt.%, cerca de 4 wt.% a cerca de 45 wt.%, o cerca de 5 wt.% a cerca de 40 wt.% del peso total de enzimas en la mezcla/composición. En todavía otro ejemplo, la β-xilosidasa(s) representa adecuadamente 2 wt.% a 30 wt.%; 10 wt.% a 20 wt.%; o 5 wt.% a 10 wt.% del peso total de enzimas en la mezcla/composición. La β-xilosidasa(s) ejemplar adecuada se describe en la Sección 6.3.7 a continuación.

Celulasas: La celulasa constituye convenientemente cerca del 0.05 wt.% a cerca del 90 wt.% del peso total de enzimas en una mezcla/composición de la enzima. La relación de cualquier par de proteínas respecto a las demás puede calcularse fácilmente con base en esta divulgación. Se contemplan mezclas/composiciones que comprenden enzimas en cualquier relación de peso derivable de los porcentajes del peso divulgados aquí. El contenido de celulasa puede ser un rango donde el límite inferior es cerca del 0.05 wt.%, 5 wt.%, 10 wt.%, 20 wt.%, 30 wt.%, 40 wt.%, 50 wt.%, 60 wt.%, 70 wt.% del peso total de enzimas en la mezcla/composición de la enzima y el limite superior es cerca del 20 wt.%, 30 wt.%, 40 wt.%, 50 wt.%, 60 wt.%, 70 wt.%, 80 wt.%, or 90 wt.% del peso total de enzimas en la mezcla/composición de la enzima. Por ejemplo, la celulasa representa convenientemente 30 wt.% a 80 wt.%; 50 wt.% a 70 wt. %; o 40 wt.% a 60 wt.% del peso total de enzimas en la mezcla/composición. Las celulasas ejemplares adecuadas se describen en la Sección 6.3.5 a continuación. Los componentes de celulasa en una mezcla/composición de la enzima de la divulgación son convenientemente capaces de lograr por lo menos 0.005 de fracción de producto por mg de proteína por gramo de ácido fosfórico de celulosa hinchada (PASC) según lo determinado por un ensayo de calcofluor. Por ejemplo, los componentes de celulasa en una mezcla/composición de la divulgación son capaces de lograr un rango de productos de fracción por mg de proteína por gramo de PASC, donde el límite inferior del rango es cercano a 0.005, 0.01, 0.015, 0.02, 0,03, 0,04, 0.05, 0,06, 0,075 o 0,1 y donde el limite superior del rango es, 0,03 0,04 0,05, 0,06, 0,075, 0,1, 0,2, 0,3, 0,4, 0,5, 0,6 o 0,7. Los componentes de celulasa en una mezcla y composición de la divulgación pueden, por ejemplo, lograr 0,00005 – 0.0001,-0.001 0.0005, 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 o 0.015 0,075 fracción de producto por mg de proteínas por gramo PASC determinado por un ensayo de calcofluor. La celulasa puede ser, por ejemplo, una celulasa completa. La celulasa puede también, por ejemplo, ser convenientemente enriquecida con una β-glucosidasa.

Proteínas accesorias: La mezcla/composición de la enzima puede convenientemente comprender además una o más proteínas accesorias. El contenido proteínico accesorio de una mezcla/composición de la enzima puede extenderse entre 0 wt.% a cerca de 60 wt.% del peso total de proteínas en una mezcla/composición de la enzima. La relación de cualquier par de proteínas respecto a las demás puede determinarse fácilmente con base en esta divulgación. Se contemplan mezclas/composiciones que comprenden enzimas en cualquier relación de peso derivable de los porcentajes del peso divulgados aquí. El contenido de la proteína accesoria puede estar en el rango donde el límite inferior es 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.% del peso total de proteínas en la mezcla/composición de la enzima y el límite superior es 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.% del peso total de proteínas en la mezcla/composición de la enzima. Por ejemplo, las proteínas accesorias pueden representar convenientemente 0 wt.% a 2 wt. %, 5 wt.% a 10 wt.%, 20 wt.% a 50 wt.%, o 2 wt.% a 5 wt.% en la mezcla/composición de la enzima. Las proteínas accesorias ejemplares adecuados para ser incluidas en las mezclas/composiciones de la enzima de la divulgación se describen en la Sección 6.3.9 a continuación.

La divulgación actual prevé una primera mezcla/composición de la enzima para la sacarificación lignocelulosa, compuesta por:

(1) cerca del 30 wt.% a cerca del 80 wt.% (*p. ej.,* 30 wt.% a 80 wt.%, 35 wt.% a 75 wt.%, 40 wt.% a 70 wt.%, 40 wt.% a 60 wt.%, 50 wt.% a 70 wt.%, etc.) de celulasa, *p.ej.,* celulasa entera o β-glucosidasa enriquecida con celulasa entera;

(2) cerca de 3 wt.% a cerca de 50 wt.% (*p.ej.,* 5 wt.% a 40 wt.%, 10 wt.% a 30 wt.%, 5 wt.% a 20 wt.%, 10 wt.% a 15 wt.%, 15 wt.% a 25 wt.%, etc.) xilanasa, *p.ej.,* una Trichoderma reesei *Xyn2* , una Trichoderma reesei *Xyn3, un AfuXyn2, un AfuXyn5, o una mezcla de dos o más de las enzimas precedentes;*

(3) cerca de 2 wt.% a cerca de 40 wt.% (*p.ej.,* 2 wt.% a 35 wt.%, 5 wt.% a 30 wt.%, 10 wt.% a 25 wt.%, 2 wt.% a 30 wt.%, 10 wt.% a 20 wt.%, 5 wt.% a 10 wt.%, etc.) β-xilosidasa, *p.ej.,* un Fv3A, un Pf43A, un Fv43D, un

Fv39A, un Fv43E, un Fv43A, un Fv43B, un Pa51A, un Fo43A, un Gz43A, una Trichoderma reesei *Bxl1, o una mezcla de dos o más de las enzimas precedentes;*

(4) cerca de 2 wt.% a cerca de 40 wt.% (*p.ej.,* 2 wt.% a 35 wt.%, 5 wt.% a 30 wt.%, 10 wt.% a 25 wt.%, 2 wt.% a 25 wt.%, 5 wt.% a 20 wt.%, 5 wt.% a 10 wt.%, etc) L-α-arabinofuranosidasa, *p.ej.,* un Af43A, un Fv43B, un Pa51A, un Pf51A, un Fv51A, o una mezcla de dos o más de las enzimas precedentes; y

(5) cerca de 0 wt.% a cerca de 50 wt.% (2 wt.% a 40 wt.%, 5 wt.% a 30 wt.%, 10 wt.% a 25 wt.%, 0 wt.% a 2 wt.%, 5 wt.% a 10 wt.%, 20 wt.% a 50 wt. %, 2 wt.% a 5 wt.%, etc) de proteínas accesorias.

La divulgación actual prevé una segunda mezcla/composición de la enzima para la sacarificación lignocelulosa compuesta por:

(1) cerca del 30 wt.% a cerca del 80 wt.% (*p. ej.,* 30 wt.% a 80 wt.%, 35 wt.% a 75 wt.%, 40 wt.% a 70 wt.%, 40 wt.% a 60 wt.%, 50 wt.% a 70 wt.%, etc.) la celulasa, *p.ej.,* celulasa entera o celulasa entera enriquecida β-glucosidasa, o cerca de 2 wt.% a cerca de 10 wt.% (*p.ej.,* 2 wt.% a 8 wt. %, 4 wt.% a 6 wt.%, 2 wt.% a 4 wt.%, 6 wt.% a 8 wt.%, 8 wt.% a 10 wt.%, 2 wt.% a 10 wt.%, etc) β-glucosidasa, *p.ej.,* una Trichoderma reesei *Bgl1 ;*

(2) cerca de 3 wt.% a cerca de 50 wt.% (*p.ej.,* 5 wt.% a 40 wt.%, 10 wt.% a 30 wt.%, 5 wt.% a 20 wt.%, 10 wt.% a 15 wt.%, 15 wt.% a 25 wt.%, etc.) xilanasa, p.ej., una Trichoderma reesei Xyn2 , una Trichoderma reesei Xyn3, un AfuXyn2, un AfuXyn5, o una mezcla de dos o más de las enzimas precedentes;

(3) cerca de 2 wt.% a cerca de 40 wt.% (*p.ej.,* 2 wt.% a 35 wt.%, 5 wt.% a 30 wt.%, 10 wt.% a 25 wt.%, 2 wt.% a 30 wt.%, 10 wt.% a 20 wt.%, 5 wt.% a 10 wt.%, etc.) de por lo menos dos β-xylosidasas, donde por lo menos una β-xylosidasa se selecciona del grupo 1 y por lo menos una β-xylosidasa se selecciona del grupo 2;

donde:

Grupo 1: un Fv3A, un Fv43A, o una mezcla de eso;

Grupo 2: un Fv43D, un Pa51A, un Gz43A, un *Trichoderma reesei* Bxl1, un Pf43A, un Fv43E, un Fv39A, un Fo43A, un Fv43B, o una mezcla de dos o más de las enzimas precedentes;

326

(4) cerca de 2 wt.% a cerca de 25 wt.% (*p.ej.,* 2 wt.% a 25 wt.%, 5 wt.% a 20 wt.%, 5 wt.% a 10 wt.%, etc) L-α-arabinofuranosidasa(s), *p.ej.,* un Af43A, un Fv43B, un Pa51A, un Pf51A, un Fv51A, o una mezcla de dos o más de las enzimas precedentes; y

(5) cerca de 0 wt.% a cerca de 50 wt.% (2 wt.% a 40 wt.%, 5 wt.% a 30 wt.%, 10 wt.% a 25 wt.%, 0 wt.% a 2 wt.%, 5 wt.% a 10 wt.%, 20 wt.% a 50 wt. %, 2 wt.% a 5 wt.%, etc) de proteínas accesorias.

La divulgación actual prevé una tercera mezcla/composición de la enzima para la sacarificación lignocelulosa compuesta por:

(2) cerca de 3 wt.% a cerca de 50 wt.% (*p.ej.,* 5 wt.% a 40 wt.%, 10 wt.% a 30 wt.%, 5 wt.% a 20 wt.%, 10 wt.% a 15 wt.%, 15 wt.% a 25 wt.%, etc.) xilanasa, p.ej., una Trichoderma reesei Xyn2 , una Trichoderma reesei Xyn3, un AfuXyn2, un AfuXyn5, o una mezcla de dos o más de las enzimas precedentes; y

(2) cerca de 2 wt.% a cerca de 40 wt.% (*p.ej.,* 2 wt.% a 35 wt.%, 5 wt.% a 30 wt.%, 10 wt.% a 25 wt.%, 2 wt.% a 30 wt.%, 10 wt.% a 20 wt.%, 5 wt.% a 10 wt.%, etc.) de por lo menos dos β-xylosidasas, donde por lo menos una β-xylosidasa se selecciona del grupo 1 y por lo menos una β-xylosidasa se selecciona del grupo 2;

donde:

Grupo 1: un Fv3A, un Fv43A, o una mezcla de eso;

Grupo 2: un Fv43D, un Pa51A, un Gz43A, un *Trichoderma reesei* Bxl1, un Pf43A, un Fv43E, un Fv39A, un Fo43A, un Fv43B, o una mezcla de dos o más de las enzimas precedentes;

La divulgación actual prevé una cuarta mezcla/composición de la enzima para la sacarificación lignocelulosa compuesta por:

(1) cerca de 5 wt.% a cerca de 25 del wt.% de xilanasa (*p.ej.,* 2 wt.% a 25 wt. %, 5 wt.% a 20 wt.%, 5 wt.% a 10 wt.%, etc), *p.ej.,* una Trichoderma reesei *Xyn2 , una* Trichoderma reesei *Xyn3, un AfuXyn2, un AfuXyn5, o una mezcla de dos o más de las enzimas precedentes;*

(2) cerca de 2 wt.% a cerca de 30 wt.% (*p.ej.,* 2 wt.% a 30 wt.%, 10 wt.% a 20 wt.%, 5 wt.% a 10 wt.%, etc.) β-xilosidasa(s), p.ej., un Fv3A, un Pf43A, un Fv43D, un Fv39A, un Fv43E, un Fv43A, un Fv43B, un Pa51A, un

327

Gz43A, un Fo43A una Trichoderma reesei Bxl1, o una mezcla de dos o más de las enzimas precedentes; y

(3) cerca de 2 wt.% a cerca de 50 del wt.% β-glucosidasa (*p.ej.,* 2 wt.% a 5 wt.%, 5 wt.% a 45 wt.%, 10 wt.% a 40 wt.%, 15 wt.% a 30 wt.%, 10 wt.% a 25 wt.%, 5 wt.% a 20 wt.%, 15 wt.% a 40 wt.%, etc), *p.ej.,* una Trichoderma reesei *Bgl1.*

La divulgación actual prevé una quinta mezcla/composición de la enzima para la sacarificación lignocelulosa compuesta por:

(1)cerca de 5 wt.% a cerca de 25 del wt.% de xilanasa (p.ej., 2 wt.% a 25 wt.%, 5 wt.% a 20 wt.%, 5 wt.% a 10 wt.%, etc), p.ej., una Trichoderma reesei Xyn2 , una Trichoderma reesei Xyn3, un AfuXyn2, un AfuXyn5, o una mezcla de dos o más de las enzimas precedentes; y

(2) cerca de 2 wt.% a cerca de 30 wt.% (*p.ej.,* 2 wt.% a 30 wt.%, 10 wt.% a 20 wt.%, 5 wt.% a 10 wt.%, etc.) β-xilosidasa(s), p.ej., un Fv3A, un Pf43A, un Fv43D, un Fv39A, un Fv43E, un Fv43A, un Fv43B, un Pa51A, un Gz43A, un Fo43A una Trichoderma reesei Bxl1, o una mezcla de dos o más de las enzimas precedentes;

La divulgación actual prevé una sexta mezcla/composición de la enzima para la sacarificación lignocelulosa compuesta por:

(1) cerca de 2 wt.% a cerca de 50 del wt.% β-glucosidasa (p.ej., 2 wt.% a 5 wt.%, 5 wt.% a 45 wt.%, 10 wt.% a 40 wt.%, 15 wt.% a 30 wt.%, 10 wt.% a 25 wt.%, 5 wt.% a 20 wt.%, 15 wt.% a 40 wt.%, etc), p.ej., una Bgl1.

(2) cerca de 3 wt.% a cerca de 50 wt.% (*p.ej.,* 5 wt.% a 40 wt.%, 10 wt.% a 30 wt.%, 5 wt.% a 20 wt.%, 10 wt.% a 15 wt.%, 15 wt.% a 25 wt.%, etc.) xilanasa, p.ej., una Trichoderma reesei Xyn2 , una Trichoderma reesei Xyn3, un AfuXyn2, un AfuXyn5, o una mezcla de dos o más de las enzimas precedentes;

(3) cerca de 2 wt.% a 40 wt.% (*p.ej.,* 2 wt.% a 35 wt.%, 5 wt.% a 30 wt.%, 10 wt.% a 25 wt.%, 2 wt.% a 30 wt.%, 10 wt.% a 20 wt.%, 5 wt.% a 10 wt.%, etc.) β-xilosidasa(s), p.ej., un Fv3A, un Pf43A, un Fv43D, un Fv39A, un Fv43E, un Fv43A, un Fv43B, un Pa51A, un Gz43A, un Fo43A, una

328

Trichoderma reesei Bxl1, o una mezcla de dos o más de las enzimas precedentes; y

(4) cerca de 2 wt.% a cerca de 25 wt.% (*p.ej.,* 2 wt.% a 25 wt.%, 5 wt.% a 20 wt.%, 5 wt.% a 10 wt.%, etc) L-α-arabinofuranosidasa(s), *p.ej.,* un Af43A, un Fv43B, un Pa51A, un Pf51A, un Fv51A, o una mezcla de dos o más de las enzimas precedentes;

La sexta mezcla/composición de la enzima para la sacarificación de la lignocelulosa anterior puede, por ejemplo, comprender cerca de 2 wt.% a cerca de 40 wt.% por lo menos dos β-xilosidasas, donde por lo menos una β-xilosidasa es seleccionada del grupo 1 y por lo menos una β-xilosidasa es seleccionada del grupo 2; donde:

Grupo 1: un Fv3A, un Fv43A, o una mezcla de eso;

Grupo 2: un Fv43D, un Pa51A, un Gz43A, un *Trichoderma reesei* Bxl1, un Pf43A, un Fv43E, un Fv39A, un Fo43A, un Fv43B, o una mezcla de dos o más de las enzimas precedentes;

Donde una mezcla/composición de la enzima de la divulgación contiene β-xilosidasas del grupo 1 y del grupo 2, la relación de las β-xilosidasas del grupo 1 al grupo 2 es preferiblemente 1:10 a 10:1. Por ejemplo, la proporción es convenientemente 1:2 a 2:1, 2:5 a 5:2, 3:8 a 8:3, 1:4 a 4:1, 1:5 a 5:1, 1:7 a 7:1, o cualquier rango entre cualquier par los extremos anteriores (, *por ejemplo,* 1:10 a 2:1, *4:1 a 2:5, 3:8 a 5:1,* etc.). Un ejemplo concreto de una proporción adecuada es de aproximadamente 1:1.

Donde una enzima de la mezcla/composición de la divulgación contiene un Fv43A como una β-xilosidasa, la mezcla/composición puede contener además Fv43B como una L-α-arabinofuranosidasa.

Una mezcla/composición de la enzima de la divulgación es, por ejemplo, convenientemente parte de una mezcla de reacción de sacarificación con biomasa, además de los componentes de la mezcla/composición de la enzima. Por ejemplo, la mezcla de reacción de la sacarificación puede caracterizarse por los 1, 2, 3 o 4 de las siguientes características:

(1) el peso total de xilanasa por kilogramo de hemicelulasa en dicha mezcla de reacción de la sacarificación está en un rango en el cual el límite inferior es cerca de 0.5 g, 1 g, 2 g, 3 g, 4 g, 5 g, 7 g, o 10 g, y el límite superior es independientemente cerca de 5 g, 7 g, 10 g, 15 g, 20 g, 30 g, o 40 g; por ejemplo, el peso total de xilanasa por kilogramo de hemicelulasa en la mezcla de reacción puede ser 0.5 g a 40 g, 0.5g a 30 g, 0.5 g a 20 g, 0.5 a 10 g, 0.5 a 5 g, 1 g a 40 g, 2 g a 40 g, 3 g a 40 g, 5 g a 40 g, 7 g a 30 g, 10 g a 30 g, 5 g a 20 g, o 5 g a 30 g.

(ii) el peso total de β-xilosidasa por kilogramo de hemicelulasa en dicha mezcla de reacción de la sacarificación está en un rango donde el límite inferior es cerca de 0.5 g, 1 g, 2 g, 3 g, 4 g, 5 g, 7 g, o 10 g y el límite superior es independientemente cerca de 5 g, 7 g, 10 g, 15 g, 20 g, 30 g, 40 g, o 50 g; por ejemplo, el peso total de β-xilosidasa por kilogramo de hemicelulasa en la mezcla de reacción puede ser 0.5 g a 40 g, 0.5 a 50 g, 0.5 g a 30 g, 0.5 g a 20 g, 0.5 g a 10 g, 0.5 g a 5g, 1 g a 40 g, 2 g a 40 g, 3 g a 40 g, 5 g a 40 g, 7 g a 30 g, 10 g, a 30 g, 5 g a 30 g, 5 g a 20 g.

(iii) el peso total de L-α-arabinofuranosidasa por kilogramo de hemicelulasa en dicha mezcla de reacción de la sacarificación está en un rango en el cual el límite inferior es cerca de 0.2 g, 0.5 g, 1 g, 1.5 g, 2 g, 2.5 g, 3 g, 4 g, o 5 g y el límite superior es independientemente cerca de 2 g, 3 g, 4 g, 5 g, 7 g, 10 g, 15 g, o 20 g; por ejemplo, el peso total de L-α-arabinofuranosidasa por kilogramo de hemicelulasa en la mezcla de reacción puede ser 0.2 g a 20 g, 0.5 g a 20 g, 1 g a 20 g, 2 g a 20 g, 2.5 g a 20 g, 3 g a 15 g, 4 g a 20 g, 5 g a 15 g, 5 g a 10 g, 5 g a 20 g, o 2.5 g a 15 g.

(iv) el peso total de celulasa por kilogramo de celulasa en dicha mezcla de reacción de la sacarificación está en un rango donde el límite inferior es cerca de 1 g, 3 g, 5 g, 7 g, 10 g, 12 g, 15 g, 18 g, o 20g, y el límite superior es independientemente cerca de 10 g, 15 g, 18 g, 20 g, 25 g, 30g, 50 g, 75 g, o 10 por ejemplo, el peso total de celulasa por el kilogramo del celulasa en la mezcla de reacción dicha puede ser 1 g a 100 g, 3 g a 100 g, 5 g a 100 g, 7 g a 100 g, 12 g a 100 g, 15 g a 100 g, 18 g a 100 g, 3 g a 75 g, 5 g a 50 g, 7 g a 75 g, 10 g a 75 g, 10 g a 50 g,

330

12 g a 75 g, 12 g a 50 g, 15 g a 75 g, 15 g a 50 g, 18 g a 30 g, 18 g a 75 g.

6.3.5. Celulasas

La mezcla/composiciones de la enzima de la divulgación puede comprender una o más celulasas. Las celulasas son enzimas que hidrolizan la celulosa (β-1,4-glucano o enlaces de β D-glucosidico) resultado la formación de la glucosa, celobiosa, celooligosacaridas y similares. Las celulasas han sido tradicionalmente divididas en tres clases principales: endoglucanasas (EC 3.2.1.4) ("EG"), exoglucanasas o celobiohidrolasas (EC 3.2.1.91) ("CBH") y β-glucosidasas (β -D-glucósido glucohidrolasa; EC 3.2.1.21) ("BG") (Knowles *et al.*, 1987, Trends in biotechnology 5(9):255-261; Shulein, 1988, Methods in Enzimology, 160:234-242). Las endoglucanasas actúan principalmente en las partes amorfas de la fibra de la celulosa, mientras que las celobiohidrolasas pueden también degradar la celulosa cristalina.

Para utilizar las celulasas de acuerdo con los métodos y las composiciones de la divulgación pueden obtenerse, o producirse recombinantemente, *entre otras cosas*, de uno o más de los siguientes organismos: *Crinipellis scapella*, *Macrophomina phaseolina*, *Myceliophthora thermophila*, *Sordaria fimicola*, *Volutella coletotricoides*, *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.* (p.ej., Trichoderma reesei) y *Cylindrocarpon sp.*

331

Por ejemplo, una celulasa, para utilizarse en el método y/o la composición de la divulgación, es una celulasa entera y/o es capaz de alcanzar por lo menos 0.1 (p. ej. 0.1 a 0.4) productos de la fracción según lo determinado por el ensayo calcofluor descrito en la Sección 7.1.10 a continuación.

6.3.5.1 β-Glucosidasa

Las mezclas/composiciones de la enzima de la divulgación pueden comprender una o más β-glucosidasas. El término β-glucosidasa según lo utilizado aquí se refiere a una glucohidrolasa β-D-glucósida clasificada como EC 3.2.1.21, y/o a los miembros de ciertas familias GH, incluyendo, sin limitación, los miembros de las familias 1, 3, 9 o 48 GH, que catalizan la hidrólisis de la celobiosa para liberar β-D-glucosa.

La β-glucosidasa adecuada se puede obtener de un número de microorganismos, por medios recombinantes, o se puede comprar de fuentes comerciales. Los ejemplos de β- glucosidasas de microorganismos incluyen, sin limitación, unos de bacterias y hongos. Por ejemplo, una β-glucosidasa de la actual divulgación se obtiene convenientemente de un hongo filamentoso.

Las β-glucosidasas se pueden obtener, o producir recombinantemente , *entre otras cosas, del Aspergillus aculeatus* (Kawaguchi *y otros* 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), o *Trichoderma reesei* (*p.ej.,* β-glucosidasa 1 (Patente de los EE.UU. No. 6,022,725), β-glucosidasa 3 (Patente de los EE.UU. No.6,982,159), β- glucosidasa 4 (Patente de los EE.UU. No. 7,045,332), β-glucosidasa 5 (Patente de los EE.UU No. 7,005,289), β-glucosidasa 6 (Publicación de los EE.UU. No. 20060258554), β-glucosidasa 7 (Publicación de los EE.UU. No. 20060258554)).

La β-glucosidasa puede ser producida expresando un gen endógeno o exógeno que codifica una β-glucosidasa. Por ejemplo, la β-glucosidasa se puede secretar en el espacio extracelular *p. ej.,* por los organismos Gram-positivos (p. ej., Bacilos o los Actinomycetes), o los eukaryotic huésped (*p.ej., Trichoderma*, *Aspergillus*,

332

Saccharomyces, o *Pichia*). La β-glucosidasa puede estar, en algunas circunstancias, sobre o sub expresada .

La β-glucosidasa se puede también obtener de fuentes comerciales. Los ejemplos de la preparación comercial de β-glucosidasa conveniente para utilizarse en la presente divulgación incluyen, por ejemplo, *Trichoderma reesei* β-glucosidasa en Accellerase® BG (Danisco US Inc., Genencor); NOVOZYM™ 188 (una β-glucosidasa de *Aspergillus niger*); *Agrobacterium sp.* β-glucosidasa y *Thermatoga maritima* β-glucosidasa de Megazyme (Megazyme International Ireland Ltd., Irlanda).

Además, la β-glucosidasa puede ser un componente de una celulasa entera, como se describe en la Sección 6.3.5.4 a continuación.

La actividad de la β-glucosidasa puede determinarse por un número de medios convenientes sabidos en técnica, tal como el ensayo descrito por Chen *y otros.*, en Biochimica et Biophysica Acta 1992, 121:54-60, donde 1 pNPG denota 1 μmoL de nitrofenol liberado de 4βnitrofenilo- - D-glucopiranosida en 10 min a 50°C (122°F) y pH 4.8.

6.3.5.2 Endoglucanasas

Las mezclas/composiciones de la enzima de la divulgación pueden comprender una o más endoglucanasas. Cualquier endoglucanasa (EC 3.2.1.4) puede utilizarse en los métodos y composiciones de la presente divulgación. Una endoglucansa puede ser producida expresando un gen endógeno o exógeno de la endoglucanasa. La endoglucanasa puede estar, en algunas circunstancias, sobre o sub expresada .

Por ejemplo, Trichoderma reesei EG1 (Penttila et al, Gene 1986, 63:103-112) y/o EG2 (Saloheimo et al., Gene 1988, 63:11-21) se utilizan adecuadamente en los métodos y las composiciones de la presente divulgación.

Una endoglucanasa *Thielavia terrestris* termoestable (Kvesitadaze *et al*, Applied Biochem. Biotech. 1995, 50:137 - 143), en otro ejemplo, se utiliza en los métodos y las composiciones de la presente divulgación. Además, un *Trichoderma reesei* EG3 (Okada *et al.* Appl. Environ. Microbiol. 1988, 64:555 - 563), EG. 4 (Saloheimo *y otros.* Eur. J. Biochem. 1997, 249:584 - 591), EG. 5 (Saloheimo *y otros.*

333

Molecular Microbiology 1994, 13:219-228), EG6 (publicación de la Patente de los EE.UU. No. 20070213249), o EG7 (publicación de la Patente de los EE.UU. No. 20090170181), un *Acidothermus cellulolyticus* El endoglucanasa (Patente de Estados Unidos No. 5.536.655), un *Humicola insolens* endoglucanasa V (EGV) (Protein Data Bank entrada 4ENG), un *Staphylotrichum coccosporum* endoglucanasa (publicación de la Patente de los EE.UU. No. 20070111278), un *Aspergillus aculeatus* endoglucanasa F1-CMC (Ooi *et al* Nucleic Acid Res. 1990, 18:5884), un *Aspergillus kawachii* IFO 4308 endoglucanase CMCase-1 (Sakamoto *et al* Curr. Genet. 1995, 27:435-439), un *Erwinia carotovara* (Saarilahti *et al* Gene 1990, 90:9-14); or también se puede utilizar un *Acremonium thermophilum* ALKO4245 endoglucanasa (publicación de la Patente de los EE.UU. No. 20070148732). Las endoglucanasas convenientes adicionales se describen en, *p. ej.,* WO 91/17243, WO 91/17244, WO 91/10732, de la Patente de los EE.UU. No. 6.001.639.

6.3.5.3 Celobiohidrolasas

Cualquier celobiohidrolasa (EC 3.2.1.91) ("CBH") puede utilizarse opcionalmente en los métodos y mezclas/composiciones de la presente divulgación. La celobiohidrolasa puede ser producida expresando un gen endogeno o exogeno de la celobiohidrolasa. La celobiohidrolasa puede estar, en algunas circunstancias, sobre o sub expresada .

Por ejemplo, Trichoderma reesei *CBHI (Shoemaker* et al. Bio/Technology 1983, 1:691-696) y/o CBHII (Teeri *et al.* Bio/technology 1983, 1:696 - 699) se pueden utilizar convenientemente en los métodos y mezclas/composiciones de la presente divulgación.

Los CBH adecuados pueden seleccionarse de un *Agaricus bisporus* CBH1 (Swiss Prot Accession No. Q92400), un *Aspergillus aculeatus* CBH1 (Swiss Prot Accession No. O59843), un *Aspergillus nidulans* CBHA (GenBank Accession No. AF420019) o CBHB (GenBank Accession No. AF420020), un *Aspergillus niger* CBHA (GenBank Accession No. AF156268) o CBHB (GenBank Accession No. AF156269), un *Claviceps purpurea* CBH1 (Swiss Prot Accession No. O00082), un *Cochliobolus carbonarum* CBH1 (Swiss Prot Accession No. Q00328), un *Cryphonectria parasitica* CBH1 (Swiss Prot Accession No. Q00548), a *Fusarium oxysporum* CBH1 (Cel7A) (Swiss Prot Accession No. P46238), una *Humicola grisea* CBH1.2 (GenBank Accession No.

U50594), una *Humicola grisea* var. *thermoidea* CBH1 (GenBank Accession No. D63515) una CBHI.2 (GenBank Accession No. AF123441), o un exo1 (GenBank Accession No. AB003105), una *Melanocarpus albomyces* Cel7B (GenBank Accession No. AJ515705), una *Neurospora crassa* CBHI (GenBank Accession No. X77778), una *Penicillium funiculosum* CBHI (Cel7A) (publicación de la Patente de los EE.UU. No. 20070148730), una *Penicillium janthinellum* CBHI (GenBank Accession No. S56178), una *Phanerochaete chrysosporium* CBH (GenBank Accession No. M22220), o un CBHI-2 (Cel7D) (GenBank Accession No. L22656), una *Talaromyces emersonii* CBH1A (GenBank Accession No. AF439935), una *Trichoderma viride* CBH1 (GenBank Accession No. X53931), o una *Volvariella volvacea* V14 CBH1 (GenBank Accession No. AF156693).

6.3.5.4 Celulasas enteras

Una mezcla/composición de la enzima de la divulgación puede además comprender una celulasa entera. Según lo utilizado aquí, una "celulasa entera" se refiere a una composición con contenido de celulasa de ocurrencia natural o no, que comprende por lo menos 3 tipos diferentes de enzimas: (1) una endoglucanasa, (2) una celobiohidrolasa y (3) una β-glucosidasa, o comprende por lo menos 3 actividades enzimáticas diferentes: (1) una actividad de endoglucanasa, que cataliza la segmentación interna de vínculos β 1,4, resultando en una glucooligosacarida más corta, (2) una actividad celobiohidrolasa, que cataliza un tipo "exo" de liberación de unidades de celobiosa (disacárido de glucosa-glocusa β-1,4) y (3) una actividad β-glucosidasa, que cataliza la liberación del monómero de glucosa de la celooligosacarida corta (*por ejemplo,* celobiosa).

Una composición que contiene celulasa que ocurre naturalmente " es una producida por una fuente que ocurre naturalmente, que comprende uno o más componentes o actividades de tipo de celobiohidrolasa, uno o más de tipo endoglucanasa y uno o más de tipo β-glucosidasa, en donde cada uno de estos componentes o actividades se encuentra en la relación y nivel producidos por la naturaleza, sin ser tocados por el hombre. En consecuencia, una composición que contenga celulosa que ocurre naturalmente es, por ejemplo, una que es producida por

un organismo inmodificado con respecto a las enzimas celulolíticas o hemiceluloíticas, de tal manera que la relación o los niveles de las enzimas componentes permanecen inalterados respecto a aquellas producidas por el organismo original en su entorno nativo. Por otro lado, una "composición que contiene celulosa que no ocurre naturalmente", hace referencia a una composición producida por: (1) enzimas celulolíticas de componente de combinación bien sea en una relación de ocurrencia natural o no natural , *es decir,* alterada, proporción; o (2) modificando un organismo para sobreexpresar, o subexpresar una o más enzimas celulolíticas; o (3) modificando un organismo tal que al menos una enzima celulolítica es eliminada. Una composición "que contiene celulasa de ocurrencia no natural " puede referirse también a una composición resultante de ajustar las condiciones de cultivo para un organismo que ocurre naturalmente, de tal forma que el organismo de ocurrencia natural crece bajo condición no nativa y produce enzimas de nivel o proporción alterada. Por consiguiente, en algunas realizaciones, la preparación celulosa entera de la presente divulgación puede tener un o más EG y/o CBHs y/o β-glucosidasas suprimidas y/o sobreexpresadas.

En la divulgación presente, una preparación de celulasa entera puede ser desde cualquier microorganismo que sea capaz de hidrolizar un material celulósico. En algunas realizaciones, la preparación de la celulasa entera es una celulasa entera de hongo filamentoso. Por ejemplo, la preparación de celulasa entera puede ser desde una especie de *Acremonium, Aspergillus, Emericella, Fusarium, Humicola, Mucor, Myceliophthora, Neurospora, Penicillium, Scytalidium, Thielavia, Tolypocladium,* o *Trichoderma* . Por ejemplo, la preparación de la celulasa entera es un *Aspergillus aculeatus, Aspergillus awamori, Aspergillus foetidus, Aspergillus japonicus, Aspergillus nidulans, Aspergillus niger,* o *Aspergillus oryzae* de celulasa entera. Además, la preparación de celulasa entera puede ser una preparación de celulasa entera de *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*o preparación de *Fusarium venenatum* . La preparación de la celulasa entera también puede ser una preparación de celulasa entera de *Humicola insolens, Humicola lanuginosa, Mucor miehei, Myceliophthora thermophila, Neurospora crassa, Penicillium purpurogenum, Penicillium brasilianum, Scytalidium thermophilum*o *Thielavia terrestris.* Además, la preparación de

celulasa entera puede ser un *Trichoderma harzianum, Trichoderma koningii, Trichoderma longibrachiatum, Trichoderma reesei* (, *por ejemplo,* 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), o una preparación de celulosa entera de *Tricoderma viride* (*p. ej.,* ATCC 32098 y 32086).

En particular, la preparación de celulasa entera puede ser convenientemente una preparación de celulasa entera de Trichoderma reesei RutC30, que está disponible en la American Type Culture Collection como Trichoderma reesei ATCC 56765. Por ejemplo, la preparación de la celulosa entera puede convenientemente también ser una celulasa entera de Penicillium funiculosum, que está disponible en American Type Culture Collection como Penicillium funiculosum *ATCC No.* : 10446.

La preparación de la celulasa entera puede también obtenerse de fuentes comerciales. Los ejemplos de las preparaciones comerciales de celulasa convenientes para el uso en métodos y composiciones de la presente divulgación incluyen, por ejemplo, CELLUCLAST™ y Cellic™ (Novozymes A/S) y LAMINEX™ BG, INDIAGE™ 44L, PRIMAFAST™ 100, PRIMAFAST™ 200, SPEZYME™ CP, ACCELLERASE ®1000 Y ACCELLERASE ® 1500 (DANISCO US. Inc., Genencor).

Las preparaciones convenientes de la celulasa entera se pueden hacer usando cualquier método de cultivo del microorganismo conocido en la técnica, especialmente fermentación, resultado en la expresión de enzimas capaces de hidrolizar un material celulósico. Según lo utilizado aquí, "fermentación" se refiere a agitar el cultivo del termo, fermentación a pequeña o gran escala, tales como fermentación continua, de lote, alimentación por lotes o fermentación en estado sólido en laboratorio o fermentación industrial realizada en un medio apropiado y bajo condiciones que permiten que la celulasa y las enzimas de interés sean expresadas y/o aisladas.

Generalmente, el microorganismo se cultiva en un medio de cultivo de célula conveniente para la producción de enzimas capaces de hidrolizar un material celulósico. El cultivo ocurre en un medio nutriente conveniente que comprende fuentes de carbón y nitrógeno y sales inorgánicas, usando procedimientos y variaciones conocidas en la técnica. Los medios de cultivo convenientes, los rangos de temperaturas y otras condiciones para el crecimiento y producción de celulasa son conocidas en la técnica.

Como un ejemplo no-limitador, un rango típico de temperaturas para la producción de celulasas por Trichoderma reesei *es de 24°C a 28°C*

La preparación de celulasa entera puede utilizarse en la medida en que se produce por la fermentación con ninguna o mínima recuperación y/o purificación. Por ejemplo, una vez que las celulasas son secretados en el medio de cultivo celular, el medio de cultivo de células que contienen las celulasas puede utilizarse directamente. La preparación de la celulasa entera puede comprender el contenido no fraccionado del material de fermentación, incluyendo el medio de cultivo celular gastado, enzimas extracelulares y células. Por otra parte, la preparación de celulasa entera puede también estar sujeta a un mayor procesamiento en un número de pasos rutinarios, p. ej. precipitación centrifugación, cromatografía de afinidad, filtración o similares. Por ejemplo, la preparación de celulasa entera se puede concentrar y luego utilizar sin purificación adicional. La preparación de celulasa entera por ejemplo, puede formularse para comprender ciertos agentes químicos que disminuyen la viabilidad celular o mata las células después de la fermentación. Por ejemplo, las células pueden ser aisladas o permeabilizadas utilizando métodos conocidos en el arte.

La actividad de endoglucanasa de la preparación de la celulasa entera se puede determinar usando la celulosa carboximetil (CMC) como substrato. Un ensayo conveniente mide la producción de reducir los extremos creados por la mezcla de la enzima que actúa en CMC en donde 1 unidad es la cantidad de enzima que libera 1 µmoL del producto/min (Ghose, T. K., puro y Appl. Chem. 1987, 59, pp. 257-268).

La celulasa entera puede ser una celulasa β-glucosidasa-enriquecida. La celulasa entera β-glucosidasa enriquecida generalmente comprende una β-glucosidasa y una preparación de celulasa entera. Las composiciones de celulasas enteras β-glucosidasa-enriquecidas se pueden producir por medios recombinantes. Por ejemplo, dicha preparación de celulasa entera puede lograrse expresando una β-glucosidasa en un microorganismo capaz de producir una celulasa entera. La composición de celulasa entera β-glucosidasa enriquecida puede también, por ejemplo, comprender una preparación de celulasa entera y una β-glucosidasa. Por ejemplo, la composición de celulasa entera β-glucosidasa-enriquecida de la celulasa puede comprender convenientemente por lo menos 5 wt.%, 7 wt.%, 10 wt.%, 15 wt.% o 20 wt.%, y hasta 25

wt.%, 30 wt.%, 35 wt.%, 40 o 50 wt.% β-glucosidasa basada en el peso total de proteínas en esa mezcla/composición.

6.3.6 Xilanasas

Las mezclas/composiciones de enzima de la divulgación, por ejemplo, pueden, comprenden una o más xilanasas grupo A, que pueden ser un Trichoderma reesei Xyn2, un Trichoderma reesei Xyn3, un AfuXyn2 o un AfuXyn5. *Los polipéptidos adecuados Trichoderma reesei* Xyn2, *Trichoderma reesei* Xyn3, AfuXyn2 o AfuXyn5 se describen en la Sección 6.1 anterior.

Las mezclas/composiciones de la enzima de la divulgación, opcionalmente, comprenden una o más xilanasas además de o en lugar de una o más xilanasas del grupo A. Cualquier xilanasa (EC 3.2.1.8) se puede utilizar como una o más xilanasas adicionales. Las xilanasas convenientes incluyen, *p. ej.,* una *xilanasa Caldocellum saccharolyticum* (Luthi *et al* 1990, Appl. Environ. Microbiol. 56 (9): 2677-2683), una xilanasa Thermatoga maritima (Winterhalter y Liebel, 1995, Appl. Environ. Microbiol. 61 (5): 1810-1815), un *Thermatoga SP.* Xilanasa de la cepa FJSS-B.1 (Simpson *y otros.* 1991, Biochem. J. 277, 413-417), una *xilanasa Bacillus circulans* (BcX) (Patente de los EE.UU No. 5,405,769), una *xilanasa Aspergillus niger* (Kinoshita *et al.* 1995, Journal of Fermentation and Bioengineering 79(5):422-428), a *xilanasa Streptomyces lividans* (Shareck *et al.* 1991, Gene 107:75 - 82; Morosoli *y otros.* 1986 Biochem. J. 239:587 - 592; Kluepfel *y otros.* 1990, Biochem. J. 287:45 - 50), *una xilanasa Bacilo subtilis* (Bernier *y otros.* 1983, Gene 26(1):59–65), a *Cellulomonas fimi* xilanase (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* xilanasa (Dominguez et al., 1995, Nature Structural Biology 2:569-576), a *Bacillus pumilus* xilanasa (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 xilanasa (Zappe *et al.* 1990, Nucleic Acids Res. 18(8):2179), o una xilanasa *Trichoderma harzianum* (Rose *et al.* 1987, J. Mol. Biol.194 (4): 755-756).

La xilanasa puede ser producida expresando un gen endógeno o exógeno que codifica una xilanasa. La xilanasa, en algunas circunstancias puede estar sobre o sub expresada .

6.3.7 β-Xilosidasas

Las mezclas/composiciones de la divulgación, por ejemplo, pueden convenientemente comprender una o más β-xilosidasas. Por ejemplo, la β-xilosidasa es una enzima β-xilosidasa del grupo 1 (*p.ej.,* una Fv3A o una Fv43A) o una enzima β-xilosidasa del grupo 2 (*p.ej.,* una Pf43A, una Fv43D, una Fv39A, una Fv43E, una Fo43A, una Fv43B, una Pa51A, una Gz43A, o una Trichoderma reesei *Bxl1*). Estos polipéptidos se describen en la Sección 0 anterior **(sic)**. Por ejemplo, una mezcla/composición de la enzima del acceso puede comprender convenientemente una o más β-xilosidasas del grupo 1 y una o más β-xilosidasas del grupo 2.

Las mezclas/composiciones de la enzima de la divulgación opcionalmente pueden constar de una o más β-xilosidasas, además de o en lugar de las β-xilosidasas del grupo 1 y/o grupo 2 anteriores. Cualquier β-xilosidasa (EC 3.2.1.37) puede utilizarse como β-xilosidasas adicionales. Las β-xilosidasas convenientes incluyen, por ejemplo, un Talaromyces emersonii *Bxl1 (Reen* y otros 2003, Biochem Biophys Res Commun. 305(3):579-85), a *Geobacillus stearothermophilus* β-xilosidasas (Shallom *et al.* 2005, Biochemistry 44:387-397), unas β-xilosidasas *Scytalidium thermophilum* (Zanoelo *et al.* 2004, J. Ind. Microbiol. Biotechnol. 31:170-176), unas β-xilosidasas *Trichoderma lignorum* (Schmidt, 1998, Methods Enzymol. 160:662-671), unas β-xilosidasas *Aspergillus awamori* (Kurakake *et al.* 2005, Biochim. Biophys. Acta 1726:272-279), unas β-xilosidasas *Aspergillus versicolor* (Andrade *et al.* 2004, Process Biochem. 39:1931-1938), una β-xilosidasas *Streptomyces sp.* (Pinphanichakarn *et al.* 2004, World J. Microbiol. Biotechnol. 20:727-733), β-xilosidasas *Thermotoga maritima* (Xue and Shao, 2004, Biotechnol. Lett. 26:1511-1515), una *Trichoderma sp.* SY β-xilosidasas (Kim *et al.* 2004, J. Microbiol. Biotechnol. 14:643-645), una β-xilosidasas *Aspergillus niger* (Oguntimein and Reilly, 1980, Biotechnol. Bioeng. 22:1143-1154), o una β-xilosidasas *Penicillium wortmanni* (Matsuo *et al.* 1987, Agric. Biol. Chem. 51:2367 - 2379).

La β-xilosidasa puede ser producida expresando un gen endógeno o exógeno que codifica una β-xilosidasa. La β-xilosidasas, en algunas circunstancias pueden estar sobre o sub expresadas.

6.3.8 L-α-Arabinofuranosidasas

Las mezclas/composiciones de la enzima de la divulgación, por ejemplo, pueden convenientemente comprender una o más L-α-Arabinofuranosidasas. La L-α-

340

arabinofuranosidasa es, por ejemplo, un polipéptido Af43A, Fv43B, Pf51A, Pa51A, o Fv51A. Los polipéptidos de Af43A, Fv43B, Pf51A, Pa51A y Fv51A se describen en la sección 6.1 anterior.

Las mezclas/composiciones de la enzima de la divulgación opcionalmente comprenden una o más L-α-Arabinofuranosidasas además o en lugar de las anteriormente mencionadas L-α-Arabinofuranosidasas. Las L-α-arabinofuranosidasas (EC 3.2.1.55) de cualquier organismo conveniente se puede utilizar como L-α-arabinofuranosidasas adicionales. Las L-α-arabinofuranosidasas conveniente incluyen, *p.ej.,* una L-α-arabinofuranosidasas del *Aspergillus oryzae* (Numan y 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 *y otros.*, 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 thermoviolacus, 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).

341

La L-α-arabinofuranosidasa puede ser producida expresando un gen endógeno o exógeno que codifica un L-α-arabinofuranosidasa. La L-α-arabinofuranosidasa, en algunas circunstancias puede estar sobre o sub expresada .

6.3.9 Proteínas accesorias

Las mezclas/composiciones de la enzima de la divulgación, por ejemplo, pueden convenientemente además comprender una o más proteínas accesorias. Ejemplos de proteínas accesorias incluyen, sin limitación, mananasas (, *por ejemplo,* endomananasas, exomananasas y β-manosidasas), galactanasas (*p. ej.,* endo - y exo-galactanasas), arabinasas (, *por ejemplo,* endo-arabinasas y exo-arabinasas), ligninasas, amilasas, glucuronidasas, proteasas, esterasas (*p. ej.,* ácido ferúlico esterasas, esterasas de xilano de acetilo, esterasas ácidos hidroxicinámico o esterasas de metilo pectina), lipasas, polipéptidos de glucósido hidrolasa familia 61, xiloglucanasas, 6p21.2, CIP2, swollenin, expansins y proteínas de interrupción de celulosa . Los ejemplos de proteínas accesorias también incluyen proteínas como CIP1, proteínas como CIP2, las de hidrogenasas de la celobiosa y las peroxidasas del manganeso. En realizaciones particulares, las proteínas que interrumpen la celulosa son módulos de enlace de celulosa.

6.4 Otros usos

Además de sacarificación de la biomasa, las enzimas o mezclas /composiciones de enzimas de la divulgación pueden ser utilizadas en procesos industriales, agrícolas, de alimentos y piensos, así como alimentos y procesos para procesar suplementos alimenticios. Los usos ejemplares son descritos a continuación.

6.4.1 Tratamientos de pulpa, papel y madera

Las enzimas, mezclas/composiciones de enzimas y los métodos de la divulgación pueden utilizarse en madera, productos de madera, residuos de madera o subproducto, papel, productos de papel, papel o pulpa de madera, pulpa de Kraft, o madera o papel reciclado de tratamiento o proceso industrial. Estos procesos incluyen, p. ej., los tratamientos de la madera, la pulpa de madera, residuos de papel, papel, o pulpa, o el destintado de la madera o del papel. Las enzimas, mezclas/composiciones de la enzima de la divulgación pueden, por ejemplo, utilizarse para tratar/pretratar la pulpa de papel o el papel reciclado o la pulpa de papel y similares reciclados. Las enzimas, mezclas/composiciones de enzima de la divulgación pueden utilizarse para

aumentar el "brillo" del papel cuando se incluyen en el papel, pulpa, papel reciclado o tratamiento/pretratamiento de la pulpa de papel. Se puede apreciar que cuanto mayor sea el grado de papel, mayor será el brillo; el brillo puede afectar la capacidad de escaneo del equipo óptico de escaneo. Como tal, las enzimas, las mezclas/composiciones de enzimas y los métodos y procesos pueden utilizarse para fabricar papeles de alto grado, "brillantes", incluidos papeles de inyección de tinta, láser y de calidad de impresión fotográfica.

Las enzimas, las mezclas/composiciones de enzimas de la divulgación pueden utilizarse para procesar o tratar otros materiales celulósicos, incluyendo, p. ej. las fibras de la madera, del algodón, del cáñamo, de la linaza o del lino.

En consecuencia, la revelación proporciona madera, pulpa de madera, papel, pulpa de papel, desechos de papel o madera o procesos de tratamiento del papel reciclado mediante una enzima, la mezcla/composición de la enzima de la divulgación.

Las enzimas, mezclas/composiciones de la enzima de la divulgación pueden utilizarse para destintar el papel de desecho impreso, tales como periódicos o para destintar papel de desecho impreso sin contacto, *p. ej.,* papel de impresiones xerográficas y papel impreso en láser y mezclas de papeles de desecho impresos de contacto y no contacto, como se describe en la Patente de los EE.UU. No. 6.767.728 o 6.426.200; Neo, j. Wood Chem. Tech. 1986, 6 (2): 147. Pueden también ser utilizados para producir la xilosa de una pulpa de la madera dura del papel en un proceso que implica extraer xilano contenido en la pulpa en una fase líquida, exponiendo el xilano contenido en la fase líquida obtenida a las condiciones suficientes para hidrolizar el xilano a xilosa y recuperando la xilosa. El paso de extracción, por ejemplo, puede incluir por lo menos un tratamiento de una suspensión acuosa de la pulpa o de un material soluble en álcali por una enzima o de una mezcla/composición de la enzima (véase, la Patente de los E.E.U.U. No. 6.512.110). Las enzimas, mezclas/composiciones de la enzima de la divulgación se pueden utilizar para disolver la pulpa de fibras celulósicas tales como productos de papel reciclados hechos de fibra de la madera dura, una mezcla de la fibra de la madera blanda, papel usado, *p.ej.,* de sobres sin imprimir, de sobres destintados, del papel oficio sin imprimir, del papel oficio destintado y similares, según lo descrito p. ej., en la Patente de los E.E.U.U. No. 6.254.722.

6.4.2 Tratamiento de fibras y textiles

La revelación proporciona métodos de tratamiento para fibras y tejidos utilizando una o más enzimas, mezclas/composiciones de enzima de la divulgación. Las enzimas, mezclas/composiciones de la enzima se pueden utilizar en cualquier fibra o método para tratar telas, conocidas en la técnica. Véase, *p. ej.,* Patente de los Estados Unidos Nos. 6.261.828; 6.077.316; 6.024.766; 6.021.536; 6.017.751; 5.980.581; publicación de la Patente de los EE.UU. No.20020142438 A1. Por ejemplo, las enzimas, mezclas/composiciones de la enzima de la divulgación pueden utilizarse para desencolar la fibra y/o la tela. La sensación y aspecto de una tela se pueden, por ejemplo, mejorar por un método que comprende poner en contacto la tela con una enzima o la mezcla/composición de la enzima de la divulgación en una solución. Opcionalmente, la tela se trata con la solución bajo presión. Las enzimas, mezclas/composición de la enzima de la divulgación se pueden utilizar también para quitar manchas.

Las enzimas, mezclas/composiciones de la enzima de la divulgación pueden utilizarse para tratar otros materiales celulósicos, incluyendo fibras (por ejemplo, fibras de algodón, cáñamo, linaza, o lino), telas cosidas y sin coser, por ejemplo, tejidos, hilados, mezclillas, lanas y toallas, hecha de algodón, mezclas de algodón o celulosicas naturales o artificiales o sus mezclas. Los procesos que tratan el textil se pueden utilizar conjuntamente con otros tratamientos textiles, *p. ej.,* pulir y/o blanquear. El pulimento, por ejemplo, es el retiro de material no celulosico de la fibra del algodón, *p. ej.,* la cutícula (principalmente consiste en ceras) y la pared principal de la célula (principalmente consiste en pectina, proteína y xiloglucano).

6.4.3 Tratamiento y procesamiento de alimentos

Las enzimas, mezclas/composiciones de la enzima de la divulgación tienen numerosas aplicaciones en la industria del proceso de alimentos. Por ejemplo, pueden ser utilizados para mejorar la extracción de aceite de plantas ricas en aceite *por ejemplo,* semillas ricas en aceite. Las enzimas, mezclas /composiciones de la enzima de la divulgación pueden utilizarse para extraer el aceite de soja de las sojas, el aceite de oliva de aceitunas, el aceite de canola de la canola, o el aceite de girasol de las semillas de girasol.

Las enzimas, mezclas/composiciones de la enzima de la revelación también pueden utilizarse para separar los componentes de materiales de células vegetales. Por ejemplo, pueden ser utilizados para separar las células vegetales en componentes. Las enzimas, mezclas/composiciones de la enzima de la divulgación pueden también utilizarse para separar cosechas fracciones de proteínas, aceite y caparazones. El proceso de separación puede realizarse usando métodos conocidos.

Las enzimas, mezclas/composiciones de la enzima de la divulgación pueden, además de las aplicaciones anteriores, ser utilizadas para aumentar la producción en la preparación de jugos de frutas o vegetales, extractos, jarabes y similares. Pueden también ser utilizadas en el tratamiento enzimático de varios materiales derivados de las paredes de las células vegetales o de materiales de desecho p. ej., cereales, granos, producción de vino o jugo, o residuos agrícolas por ejemplo, los caparazones vegetales, las vainas de la haba, pulpa de remolacha, pulpa de aceitunas, pulpa de patata y similares. Además, pueden ser utilizadas para modificar la consistencia y/o el aspecto de frutas o verduras procesadas. Además, pueden ser utilizadas para tratar el material vegetal como para facilitar el proceso del material de la planta (incluyendo alimentos), purificación o extracción de los componentes de la planta. Las enzimas, mezclas/composiciones de la enzima de la divulgación pueden utilizarse para mejorar el valor de la alimentación, para disminuir la capacidad de enlace del agua, para mejorar la degradabilidad en las plantas de agua de desecho y/o para mejorar la conversión del material de la planta en el ensilado y similares.

Las enzimas, mezclas/composición de la enzima de la divulgación se pueden utilizar también para hornear. En algunas realizaciones se utilizan para crear las pastas no pegajosas para su fácil manejo en las máquinas y para reducir los tamaños de las galletas. Puede utilizarse también para hidrolizar las arabinoxilanas para evitar una rápida rehidratación del producto cocido que puede conducir a la pérdida del efecto crujiente y reducir su vida útil . Por ejemplo, se utilizan como aditivos en el proceso de la masa.

6.4.4 Piensos y alimentos o piensos o aditivos alimentarios

La divulgación proporciona métodos para el tratamiento de piensos y alimentos y aditivos de alimentos o piensos (suplementos) utilizando enzimas, mezclas y composiciones de enzimas de la divulgación. Animales incluyendo mamíferos (p. ej.,

seres humanos), pájaros, pescados y similares. La divulgación proporciona piensos y alimentos y aditivos (suplementos) que comprenden enzimas, mezclas y composiciones de enzimas de la divulgación. Al tratar los piensos, los alimentos y los aditivos utilizando las enzimas de la divulgación, puede ayudar en la disponibilidad de nutrientes, p. ej., almidón, proteínas y similares, en el pienso o en el aditivo (suplementos). Al descomponer las proteínas de difícil digestión o al desenmascarar directa o indirectamente el almidón (u otros nutrientes), las enzimas, las mezclas/composiciones de la enzima pueden hacer que los nutrientes sean más accesibles a otras enzimas endógenas o exógenas. Pueden también causar simplemente la liberación de nutrientes y azúcares listos para ser digeridos y de fácil absorción.

Cuando se agrega a la alimentación animal, enzimas, mezclas/composiciones de enzima de la divulgación mejora la descomposición en vivo del material de la pared celular vegetal, en parte porque reduce la viscosidad intestinal (ver, por ejemplo, Bedford et al., actas del primer Simposio de enzimas en la Alimentación Animal, 1993, pp. 73-77), donde se logra una mejor utilización de los nutrientes de la planta por parte del animal. Así, usando las enzimas, las mezclas/composiciones de la divulgación en alimentos, la tasa de crecimiento y/o la relación de conversión de la alimentación del animal puede mejorarse (*es decir,* el peso del alimento injerido con respecto al aumento de peso).

El aditivo del alimento animal de la revelación puede ser una enzima granulada que se puede mezclar fácilmente con componentes de la alimentación. Alternativamente, los aditivos de la alimentación de la divulgación pueden formar un componente de una premezcla. El producto de enzima granulado de la divulgación puede ser con recubrimiento o no . El tamaño de partícula de la enzima granulada puede ser compatible con la de la alimentación y/o de los componentes de la premezcla. Esto proporciona un medio seguro y conveniente de incorporar las enzimas en la alimentación. Alternativamente, el aditivo del pienso de la divulgación puede ser una composición líquida estabilizada. Esto puede ser una mezcla acuosa o una base líquida de aceite. Véase, *p. ej.,* Patente de los E.E.U.U. No. 6.245.546.

Una enzima, mezcla/composición de la enzima de la divulgación puede ser proveída expresando las enzimas directamente en las cosechas de alimentación transgénica (*p. ej.,* plantas transgénicas, semillas y similares), tales como granos,

cereales, maíz, soja, canola, lupino y similares. Según lo discutido anteriormente, la divulgación proporciona plantas transgénicas, partes de plantas y células de plantas que comprenden una secuencia de ácido nucléico que codifica un polipéptido de la divulgación. El ácido nucleico se expresa de tal forma que la enzima de la revelación es producida en cantidades recuperables. La xilanasa se puede recuperar de cualquier planta o parte de la planta. Alternativamente, la planta o la parte de la planta que contiene el polipéptido recombinante puede ser utilizada como tal para mejorar la calidad de un alimento o una alimentación, *p. ej.*, mejorando el valor nutricional, el sabor agradable y las propiedades reologicas, o para destruir un factor antinutritivo.

La divulgación proporciona métodos para eliminar oligosacáridos de los piensos antes del consumo animal sujeto al uso de una enzima, la mezcla/composición de la enzima de la divulgación. En este proceso el pienso se forma para tener un valor metabolizable creciente de energía. Además de las enzimas, las mezclas/composiciones de la enzima de la divulgación, se pueden usar galactosidasas, celulasas y combinaciones de las mismas.

La divulgación proporciona métodos para utilizar una enzima, una mezcla/composición de la enzima de la divulgación como un suplemento nutricional en las dietas de los animales mediante la preparación de un suplemento nutricional que contiene una enzima recombinante de la divulgación y la administración del suplemento nutricional a un animal para aumentar la utilización de la hemicelulasa contenida en los alimentos de ingesta animal.

6.4.5 Tratamiento de desechos

Las enzimas, las mezclas/composiciones de la enzima de la divulgación, pueden utilizarse en gran variedad de aplicaciones industriales, p. ej., en el tratamiento de desperdicios o desechos. Por ejemplo, en un aspecto, la divulgación proporciona un proceso de digestión de residuos sólidos utilizando enzimas, mezclas y composiciones de enzima de la divulgación. Los métodos pueden constar de reducción de la masa y volumen sustancialmente de residuos sólidos sin tratar. Los residuos sólidos pueden tratarse con un proceso digestivo enzimático en presencia de una solución enzimática (incluyendo las enzimas, mezclas y composiciones de enzima de la divulgación) a una temperatura controlada. Esto da lugar a una reacción sin fermentación bacteriana apreciable de microorganismos agregados. Los residuos sólidos que se convierten en

397

un desperdicio licuado y cualquier desecho residual sólido. Los residuos licuados resultantes se pueden separar de cualquier dicho desperdicio residual solidificado. Véase, *p. ej.,* Patente de los E.E.U.U. No. 5.709.796.

6.4.6 Detergente, Desinfectante y Composiciones de limpieza

La divulgación ofrece composiciones de detergentes, desinfectantes o limpiadores (limpieza o depuración) que comprenden una o más enzimas, mezclas y composiciones de enzima de la divulgación y métodos para fabricar y utilizar estas composiciones. La divulgación incorpora todos los métodos conocidos para fabricar y utilizar composiciones de detergentes, desinfectantes o limpiadores. Véase, *p. ej.,* Patente de los Estados Unidos Nos. 6,413,928; 6,399,561; 6,365,561; 6,380,147.

En divulgaciones específicas, las composiciones de los detergentes o limpiadores pueden ser una composición acuosa de una o dos partes, una composición líquida no acuosa, una composición sólida o en forma granular o de partícula, una tableta comprimida, un gel y/o una pasta y también en forma líquida. Las enzimas, mezclas/composiciones de la enzima de la divulgación pueden utilizarse también como un producto aditivo detergente, desinfectante o limpiador en forma sólida o líquida. Tales productos aditivos se pretenden para suplir o para potencializar el funcionamiento de composiciones de detergentes convencionales y se pueden agregar en cualquier etapa del proceso de limpieza.

La presente divulgación ofrece composiciones de limpieza incluyendo composiciones de detergentes para la limpieza de superficies duras, composiciones de detergentes para la limpieza de telas, composiciones para lavavajillas, composiciones para limpieza oral, composiciones de limpieza para prótesis dentales y soluciones de limpieza para lentes de contacto.

Cuando las enzimas de la divulgación son componentes de composiciones adecuadas para su uso en el método de lavado de las lavadoras, las composiciones pueden incluir, además de una enzima, la mezcla/composición de la enzima de la revelación, un compuesto surfactante y generador. Además pueden constar de uno o más componentes de detergentes, *p. ej.,* compuestos poliméricos orgánicos, agentes blanqueadores, enzimas adicionales, supresores de espuma, dispersantes, dispersantes

de cal, agentes de suspensión de suciedad y agentes anti-redeposiciones e inhibidores de corrosión.

Las composiciones de lavandería de la divulgación también contienen agentes suavizantes, como componentes adicionales de detergentes. Tales composiciones que contienen carbohidrasa ofrecen limpieza de telas, eliminación de manchas, mantenimiento de blancura, suavizante, apariencia del color, inhibición de desteñido y desinfección cuando se formula como composiciones para detergentes de lavandería.

La divulgación además proporciona un proceso de sacarificación de un material de biomasa que comprende hemicelulosa. Tal material de biomasa puede incluir opcionalmente celulosa. Los materiales ejemplares de la biomasa incluyen, sin limitación, mazorca, mijo, sorgo y/o bagazo. En consecuencia la divulgación proporciona un proceso de sacarificación, que comprende el tratamiento de un material de biomasa que comprende hemiceluosa y, opcionalmente, celosa con una mezcla/composición enzimática tal como se describe en este documento. La mezcla/composición de la enzima utilizada en tales procesos de la invención incluye 0.5 g a 40 g (*p. ej.,* 0.5 g a 20 g, 0.5 g a 30 g, 0.5 g a 40 g, 0.5 g a 15 g, 0.5 g a 10 g, 0.5 g a 5 g, 0.5 g a 7 g, etc.) de polipéptidos que tienen actividad de xilanasa por el kilogramo de hemicelulosa en el material de la biomasa. La mezcla/composición de la enzima utilizada en tal proceso de la invención puede también incluir 1 g a 40 g (*p. ej.,* 2 g a 20 g, a 3 g a 7 g, 1 g a 5 g, o 2 g a 5 g, etc.) de los polipéptidos que tienen actividad de xilanasa por kilogramo de hemicelulosa en el material de la biomasa. La mezcla/composición de la enzima utilizada en tal proceso puede incluir 0.5 g a 50 g (*p. ej.,* 0.5 g a 50 g, 0.5 g a 45 g, 0.5 g a 40 g, 0.5 g a 30 g, 0.5 g a 25 g, 0.5 g a 20 g, 0.5 g a 15 g, 0.5 g a 10 g, etc.) de polipéptidos que tienen β actividad de la xilosidasa por kilogramo de hemicelulosa en el material de la biomasa. La mezcla/composición de la enzima utilizada en tal proceso de la invención puede también incluir 1 g a 50 g (*p. ej.,* 2 g a 40 g, 4 g a 20 g, 4g a 10 g, 2 g a 10g, 3 g a 7 g, etc.) de los polipéptidos que tienen actividad de xilanasa por kilogramo de hemicelulosa en el material de la biomasa. La mezcla/composición de la enzima utilizada en tal proceso puede incluir 0.2 g to 20 g (*p. ej.,* 0.2 g a 18 g, 0.2 g a 15 g, 0.3 g a 10 g, 0.2 g a 8 g, 0.2 g a 5 g, etc.) de polipéptidos con actividad L-α-arabinofuranosidasa por kg de hemicelolosa en el material de la biomasa . La mezcla/composición de la enzima utilizada en tal proceso de la invención puede incluir 0.5 g to 20 g (*p.ej.,* 1 g a 10 g, 1 g a 5 g, 2 g a 6 g, 0.5 g a 4 g, o 1 g a 3 g, etc.) de

polipéptidos con actividad L-α-arabinofuranosidasa por kg of hemicelolosa en el material de la biomasa . La mezcla/composición de la enzima también puede incluir 1 g a 100 g (*por ejemplo,* 1 g a 100 g, 2 g, 80 g, 3 g a 50 g, 5 g a 40 g, 2 g, 20 g, 10 g a 30 g o 12 g 18 g, etc.) de polipéptidos con actividad celulasa por kg de celulosa en el material de la biomasa. Opcionalmente, la cantidad de polipéptidos que tienen β actividad de la glucosidasa puede constituir hasta el 50% del peso total de polipéptidos que tienen actividad del celulasa.

El proceso conveniente de la invención preferiblemente produce entre el 60% y 90% de xilosa del xilano de hemicelulosa del material tratado de la biomasa. Los materiales convenientes de la biomasa incluyen uno o más, por ejemplo, de mazorca, mijo, sorgo y/o bagazo. Como tal, un proceso de la invención rinde preferiblemente por lo menos el 70% (p. ej., por lo menos 75%, por lo menos del 80%) de xilosa del xilano de hemicelulosa de uno o más de estos materiales de la biomasa. Por ejemplo, el proceso rinde entre 60% y 90% de la xilosa del xilano de la hemicelulosa de un material de la biomasa que comprende hemicelulosa, incluyendo, sin limitación, mazorca, mijo, sorgo y/o bagazo.

El proceso de la invención comprende opcionalmente además la recuperación de los monosacáridos.

La invención además es ilustrada por los siguientes ejemplos. Los ejemplos se proporcionan para fines de ilustración únicamente. Por ningún motivo deben ser interpretados como limitación del alcance o del contenido de la invención.

7. EJEMPLO 1: MÉTODOS EXPERIMENTALES REPRESENTATIVOS

7.1 Materiales y Métodos

Los siguientes ensayos y métodos fueron utilizados en el Ejemplo 1 y en ejemplos posteriores. Cualquier desviación de los protocolos proporcionados se indica a continuación.

7.1.1 Preparación de la hemicelulosa a partir de tejidos vegetales

Las preparaciones de la hemicelulosa utilizaron una modificación del procedimiento de NaOH/sonicación descrito por Erbringerova *y otros.* (Carbohydrate Polymers 1998, 37:231). El material vegetal seco se molió para pasar por una pantalla de 1 mm y 10 g

de este material fue suspendido en 250 mL de 5% (wt/v) de NaOH. La suspensión fue calentada a 80°C sin revolver por 30 minutos después sonicada por 15 minutos a temperatura ambiente usando un sonicador de sonda ajustado en valores altos. La suspensión se regresó a 80°C durante 30 minutos adicional después se permitió que se enfriara a temperatura ambiente. Se retiraron los sólidos de la suspensión por centrifugación a 3000 x g por 15 minutos y el sobrenadante resultante fue decantado en 1 L del etanol que fue enfriado con hielo. Después de 30 minutos el precipitado resultante fue recuperado primero decantando el líquido transparente por encima del precipitado, luego se filtró sin permitir que el precipitado se secara completamente. El producto de la filtración primero fue lavado con 200 mL de frío, 80% de etanol fue removido del filtro sin permitir que se secara. El producto de la filtración fue redisuelto en 200 mL de agua y el pH de la solución fue ajustado a 5.5 con el ácido acético. El carbohidrato extraído fue reprecipitado agregando 1 L de etanol al hielo y el precipitado resultante fue recuperado de nuevo por la filtración anteriormente descrita. El producto de la filtración se congeló y el solvente y agua restantes fueron removidos por medio de liofilización. La producción del carbohidrato recuperado osciló entre 6 y 23% del material de planta inicial dependiendo del tejido y de la preparación.

7.1.2 Pretratamiento del amoníaco diluido de los sustratos de la biomasa

Se pretrataron la mazorca y el mijo antes de la hidrólisis enzimática según los métodos y rangos de procesamiento descritos en WO06110901A (a menos que se indique lo contrario).

7.1.3 Análisis de composición de la biomasa

Los 2 pasos de los métodos de hidrólisis de ácido descritos en la "Determinación de carbohidratos y lignina estructurales en la biomasa" (el laboratorio nacional de Energía Renovable, Golden, CO 2008, disponible en http://www.nrel.gov/biomass/pdfs/42618.pdf) fueron utilizado para medir la composición de los substratos de la biomasa. Los resultados enzimáticos de la hidrólisis se divulgan en términos de conversión de porcentajes con respecto a la producción teórica del contenido del glucano y xilano del sustrato.

7.1.4 Preparación de oligómeros no refinados a partir de mazorcas pretratadas en amoniaco

Se prepararon los oligómeros sin refinar para detectar hemicelulasas utilizando mazorcas con el siguiente procedimiento. Las mazorcas molidas con martillo (~1/4 diámetro promedio) más 6% (w/w) de amoníaco caliente a 145°C con inyección directa de vapor en un reactor agitador a presión. Después de 20 minutos el exceso de amoníaco es sacado del reactor a un vacío final de ~0.1 bar. La mazorca pretratada con amoníaco luego se coloca en un reactor agitador estéril para la sacarificación de la enzima. Se agrega suficiente agua para obtener una carga final total de sólidos del 25% (w/w) después de realizar todas las adiciones. El pH del reactor se mantuvo en pH 5.3 con el ácido sulfúrico de 4 N y la temperatura controlada de 47°C. Se agregó Spezyme ®^{CP}, Xilanasa Multifect ® (Danisco los US Inc., Genencor) y Novo 188 (Novozymes, Dinamarca) a las cargas de 20, 10 y 5 mg/g de celulosa, respectivamente y se permitió sacarificar la mazorca pretratada con azúcares y oligómeros durante 116 H. El material entonces fue enfriado a 33°C y el pH se ajustó a 5.8 con NaOH 4 N. La glucosa y la xilosa se fermentaron en etanol agregando un cultivo de semilla de una cepa recombinante Zymomonas mobilis (volumen total del 10%, No. de accesión del ATCC. PTA-1798) según lo descrito en los E.E.U.U. 7.354.755. Se siguió el progreso de la fermentación hasta que toda la glucosa y el ~95% de la xilosa se consumió. Una alícuota de 0,5 l de caldo de fermentación se aclaró por centrifugación (21.000 x g) durante 20 minutos seguido por filtración del sobrenadante a través de una unidad de filtrado de 0,2 micras (Nalgene). Se retiró el etanol del caldo de fermentación filtrado en un rotovap que se mantuvo a 35°C en vacío. El volumen total del licor final fue reducido en ~4x utilizando el último procedimiento.

7.1.5 Ensayos de la proteína total

Se emplearon diferentes métodos para determinar las proteínas totales dependiendo de la naturaleza de la muestra de proteína (es decir, purificada, caldo de fermentación, producto comercial, etc.). El ensayo de la proteína BCA es un ejemplo de un ensayo colorimétrico que mide la concentración de la proteína de proteína con un espectrofotómetro.

Reactivos: BCA Protein Assay Kit (Pierce Chemical, Product #23227), 50 mM solución amortiguadora de Acetato de Sodio pH 5.0, 15% ácido trichloroacetico (TCA),

0.1 N NaOH, BSA solución de caldo, Reactivo A, Reactivo B (del Kit de ensayo de proteína)

Procedimiento: Se prepararon diluciones de enzima en tubos de ensayo con solución amortiguadora de acetato de sodio de 50 mM. La solución diluida de enzima (0,1 mL) fue agregada a los tubos centrífugos de Eppendorf de 2 mL que contenían 1 mL 15% TCA. Los tubos eran de vórtice y se colocaron en hielo durante 10 Min. Las muestras entonces fueron centrifugadas a 14.000 RPM durante 6 Min. El sobrenadante fue vertido, el gránulo resuspendido en 1 mL 0,1 N NaOH y los tubos se sometieron al vórtice hasta que el gránulo fue disuelto. Las soluciones estándar de BSA se prepararon desde una solución de caldo de 2 mg/mL. La solución de trabajo de BCA fue preparada mezclando 0,5 mL de reactivo B con 25 mL de reactivo A. Se agregó la proteína resuspendida (0,1 mL cada una) en 3 tubos centrífugos . Se agregó la solución de trabajo de dos mL Pierce BCA a cada una de las muestras y a los tubos estándar Eppendorf BSA diluidas en serie . Todos los tubos fueron incubados al baño maría a 37°C durante 30 minutos. Las muestras fueron luego enfriadas a temperatura ambiente (15 min) y se medió la absorbancia a 562 nm en un espectrofotómetro.

Cálculos: Los valores medios para cada absorbencia estándar de la proteína BSA fueron calculados y trazados, la absorbencia en el eje x y la concentración (mg/ml) en el eje y. Se utilizó una curva lineal de ajuste y la ecuación de la línea se calculó mediante la fórmula: $y = mx + b$

La concentración sin refinar de las muestras de enzima se calculó mediante la sustitución de la absorbancia por el valor x. La concentración de la proteína total fue calculada multiplicándola por el factor de dilusión.

La proteína total de las muestras purificadas fue determinada por A280 (*ver, p. ej.,* Pace y otros, Protein Science, 1995, 4:2411).

Algunas muestras de proteínas se miden utilizando el método Biuret modificado por Weichselbaum y Gornall utilizando seroalbúmina bovina como un calibrador (Biuret modificado) (Weichselbaum, Amer. J. Clin. Path. 1960, 16:40; Gornall *y otros.* , J. Biol. Chem. 1949, 177:752).

El contenido proteínico total de los productos de la fermentación también es medido algunas veces como nitrógeno total por la combustión, la captura y la medida del

353

nitrógeno liberado, por Kjeldahl (laboratorios rtech, www.rtechlabs.com) o interno por el método DUMAS (TruSpec CN, www.leco.com) (SADER, *y otros.* Archivos de la ciencia veterinaria, 9 (2): 73-79, 2004). Para muestras con proteínas complejas, p. ej. los caldos de la fermentación, se utilizó un contenido N 16% promedio y el factor de conversión de 6.25 para la relación nitrógeno a proteína . En algunos casos, la proteína precipitable total fue medida para retirar el nitrógeno sin proteínas que interfería. Se utilizó una concentración final de TCA de 12,5% y el gránulo TCA que contenía proteína fue resuspendido en 0,1 M NaOH.

En otros casos se utilizó, Coomassie más el mejor ensayo de Bradford (Thermo Scientific, Rockford, IL product #23238) siguiendo la recomendación del fabricante.

7.1.6 Ensayos de la actividad de sustrato sintético (sustrato para-nitrofenilo)

La proteína activa de la expresión de genes clonados *T. reesei* fue confirmada con ensayos de sustrato modelo. Las actividades de la celulasa y hemicelulasa en los sustratos sintéticos, tales como 4-nitrofenilo α-L-arabinofuranosida (pNPA, Sigma N3641) y 4-nitrofenilo β-D-glucopyranosida (pNPG, Sigma N7006) y 4-nitrofenilo β-D-xylopyranosida (pNPX, N2132 de Sigma) se midieron de la siguiente forma: La solución del substrato fue preparada disolviendo 30mg del substrato sintético en 100 mL 50 mL de solución amortiguadora de acetato de sodio , pH 4.8. El carbonato de sodio (1 M) fue preparado para mitigar la reacción. La solución del substrato (µL 100) fue dispensada en las placas Costar 96 (Cat No. 9017). El µL 20 de la muestra de la enzima fue dispensado en una microplaca con múltiples agujeros. La microplaca fue incubada a 50°C durante 10 mínimos utilizando un calentador Thermomixer R y un agitador enfriador (Eppendorf). Se agregó µL 50 de carbonato de sodio de 1 M a cada uno de los agujeros para mitigar la reacción. La absorbencia en longitud de onda de 400 nanómetros fue leída con el espectrofotómetro de Microplaca SpectraMax 340C384 (dispositivos moleculares). Las unidades por mL fueron determinadas usando una curva estándar *p*- nitrofenol. Un huésped Trichoderma de eliminación Quad., desde el cual los genes *cbh1, cbh2, egl1* y *egl2* fueron eliminados (véase WO 05/001036), fue analizado con las muestras de enzima a manera de control de la actividad de las enzimas expresadas en los antecedentes.

7.1.7 Ensayo de sacarificación de la mazorca

Normalmente se realiza la sacarificación de la mazorca en un formato de placa de microtitulación de conformidad con los siguientes procedimientos. El sustrato de la biomasa, la mazorca pretratada en amoníaco diluído, fue diluido en agua y el pH-ajustado con el ácido sulfúrico para crear un pH 5, 7% de celulasa líquida que fue utilizada directamente en el ensayo. Las enzimas probadas incluyen: productos comerciales de celulasa, por ejemplo, Accellerase® 1000, Accellerasa ® 1500 (Danisco US Inc., Genencor), caldos de fermentación T. reesei y enzimas purificadas. Se cargaron las enzimas con base en la proteína total de mg por gramo de celulosa (según lo determinado por análisis de composición) en el substrato de la mazorca. Se diluyeron las enzimas en 50 milímetros de acetato de sodio pH 5.0 para obtener la concentración deseada de carga en el volumen requerido. Cuarenta microlitros de la solución de la enzima fueron agregados a 70 mg de mazorca pretratada diluida en amoníaco a 7% de celulosa por placa (equivalente a 4.5% de celulosa final). La placa del ensayo fue incubada a temperatura ambiente por 10 Min. Las placas del ensayo fueron cubiertas con los selladores de placa de aluminio y las placas fueron incubadas a 50°C, 200 RPM, por tres días. Al final del período de la incubación, la reacción de la sacarificación fue mitigada agregando 100 µL de 100 mM de solución amortiguadora/tampón de glicina, pH10.0 por placa de agujeros y la placa fue centrifugada por 5 minutos a 3.000 RPM. 10 Ml de sobrenadante se añadieron a 200 µL de agua MilliQ en una placa de 96 agujeros HPLC y los azúcares solubles se midieron por HPLC.

Esto describe un método típico utilizado en los múltiples ejemplos del presente documento. En ciertos ejemplos, la sacarificación de mazorca fue medida utilizando un protocolo modificado. Las modificaciones se describen con los ejemplos individuales.

7.1.8 Análisis de azúcar por HPLC

Normalmente, las muestras de hidrólisis de sacarificación de la mazorca se prepararon por centrifugación para despejar el material insoluble, filtración a través de un filtro de nylon 0.22 µm (Spin-X filtro de tubo centrífugo, Corning Incorporated, Corning, NY) y dilución a una concentración apropiada de azúcares solubles con agua destilada. Los azúcares de monómero se determinaron en un Shodex SH1011 de SH-G de azúcar, 8x300mm con una columna de guardia 6 x 50 mm SH-1011P (www.shodex.net). El solvente era 0.01 N H_2SO_4 operando a 0.6 mL/min. La temperatura de la columna fue de 50°C y la detección fue hecha usando un detector índice de

refracción. Los estándares externos de la glucosa, xilosa y de la arabinosa fueron operados con cada conjunto de muestra. Ciertos ejemplos aquí utilizan un protocolo para alcanzar el mismo fin con un conjunto algo modificado de protocolos. Las modificaciones específicas a los protocolos se describen con los ejemplos individuales.

Los azúcares oligoméricos fueron separados por cromatografía de exclusión de tamaño usando una columna de 7.5 mm x 60 cm Tosoh Biosep G2000PW (www.tosohbioscience.de). El solvente era agua destilada en 0.6 mL/min y la columna fue operada a temperatura ambiente. Los seis estándares de azúcar carbono utilizados para la calibración por tamaño fueron: stachyose, raffinose, celobiosa y glucosa; y 5 los azúcares de carbón fueron: xilohexosa, xilopentosa, xilotetrosa, xilotriosa, xilobiosa y xilosa. Los Xilo-oligomeros fueron obtenidos de Megazyme (www.megazyme.com). La detección se realizó por el índice de refracción y cuando se reportan los resultados cuantitativo están como unidades de área máxima o áreas máximas relativas por porcentaje.

Los azúcares solubles totales fueron determinados por hidrólisis de ácido de las muestras centrifugadas y clarificadas por filtro. La muestra clarificada fue diluida 1:1 con 0,8 n H_2SO 4 y la solución resultante se trató en autoclave en un frasco tapado por un ciclo total de tiempo de 1 h a 121°C. Los resultados se registran sin corrección de la pérdida de monómero de azúcar durante la hidrólisis.

7.1.9 <u>Análisis de proteína por HPLC</u>

Para separar y cuantificar las enzimas contenidas en el caldo de las fermentaciones 14L de las cepas integradas de la expresión, se realizaron cromatografías líquidas (LC por su sigla en inglés) y espectroscopia de masa (MS). Las muestras de la enzima primero fueron tratadas con una glicosidasa recombinante expresando glicosidasa endoH de S. plicatus (e.g., NEB P0702L). EndoH fue utilizado en una proporción de 0.01-0.03 µg proteína endoH por µg proteína total de la muestra e incubada durante 3 h a 37°C, pH 4.5-6.0 para retirar enzimáticamente la gicosilación vinculada N antes del análisis HPLC. Aproximadamente 50 µg de proteína fue inyectado para la cromatografía de interacción hidrofóbica utilizando un sistema Agilent 1100 HPLC con una columna de HIC-fenilo y se realizó un gradiente de sal de alto a bajo durante más de 35 min sobre las muestras del caldo de fermentación concentrado. El gradiente fue alcanzado usando la solución amortiguadora A alta en sal: El Sulfato de amonio de 4 m que contiene 20 mM de fosfato

350

de potasio pH 6,75 y solución amortiguadora B baja en sal: 20 milímetros de fosfato de potasio pH 6.75. Los picos fueron detectados con la luz UV a 222 nanómetros y las fracciones fueron recogidas e identificadas por espectroscopia másica.

7.1.10 Ensayo de la actividad de la celulasa utilizando blanco de Calcofluor

La actividad de la celulasa fue medida en PASC utilizando un método de detección blanco de calcofluor (Appl. Biochem. Biotechnol. 161:313 - 317). Todos los productos químicos utilizados fueron de grado analítico. Avicel PH-101 fue comprado de FMC Biopolímero (Philadelphia, PA). El blanco de Calcofluor fue comprado a Sigma (St. Louis, MO). La celulosa hinchada de ácido fosfórico (PASC) fue preparada de Avicel PH-101 usando un protocolo adaptado de Walseth, TAPPI 1971, 35:228 y Wood, Biochem. J. 1971, 121:353 - 362. En definitiva, el Avicel fue solubilizado en ácido fosfórico concentrado después precipitado usando agua desionizada fría. Después de que la celulosa fuera recogida y lavada con más agua para neutralizar el pH, fue diluida hasta sólidos del 1% en la solución amortiguadora del acetato del sodio de 50 milímetros, pH 5.0.

Todas las dilusiones de la enzima fueron realizadas en la solución amortiguadora del acetato del sodio de 50 milímetros, pH 5.0. La celulasa GC220 (Danisco los US Inc., Genencor) fue diluida 2.5, 5, 10 y 15 mg proteína/g PASC, para producir una curva lineal de calibración. Las muestras que se probarán fueron diluidas para quedar en la gama de la curva de calibración, *es decir.* para obtener una respuesta de 0.1 a 0.4 productos de la fracción. Los150 µL de frío 1% PASC se agregaron a 20 µL de la solución de la enzima en 96 placas microtítuladas. La placa fue cubierta e incubada para 2 h a 50°C, 200 RPM en una incubadora Innova/agitador. La reacción fue mitigada con100 µL de 50 µg/mL Calcofluor en 100 milímetros de Glicina, pH 10. La fluorescencia fue leída en un lector microplaca de la fluorescencia (SpectraMax M5 por dispositivos moleculares) a una longitud de onda de excitación EX = 365 nanómetros y longitud de onda de la emisión Em = 435 nanómetros. El resultado (demostrado en la Figura 25) se expresa como el producto de la fracción según la ecuación:

Punto de congelación = 1 - (muestra FI)/(Enzima cero FI - solución amortiguadora FI), donde el FP es producto de la fracción y FI = unidades de fluorescencia.

7.1.11 Cultivo de Fusarium verticillioides y purificación de las principales actividades de la hemicelulasa detectadas en la proteína extracelular

La fuente *Fusarium verticillioides tipo silvestre* fue tal como se describe en la Tabla 1 de Fuchs *et al*, Fungal Genetics and Biology 2004, 41:852-863. El hongo fue cultivado utilizando fibra de grano de maíz desalmidonada mediante una modificación del método descrito en Li *et al*, Applied Biochemistry and Biotechnology 2005, (121-124): 321-334.

Se retiró el almidón de la fracción del pericarpio de maíz seco de un fraccionamiento de maíz seco en molino mezclando 200 g del pericarpio de maíz en 3 L de agua del grifo y calentándolo a 80°C con 5 mL (500 mg) de amilasa a alta temperatura Liquozyme SC DS (Novozymes, Dinamarca). La mezcla fue agitada de vez en cuando con una espátula y manteniéndola a 80°C por 30 minutos y luego se enfrió a temperatura ambiente por cerca de 2 horas. La mezcla resultante fue decantada sobre una criba de malla de 20 para desecar parcialmente. Los sólidos en la criba fueron lavados más a fondo con 4 L de agua del grifo después secados al aire durante la noche antes de un secado final en un horno a 60°C. El material secado fue molido pasando por una criba de 1 milímetro en un molino de cuchillas antes de utilizarse.

El cultivo de F. verticillioides se mantuvo en agar de dextrosa de patata (Sigma P6685) y se inoculó en un medio rico de crecimiento de dextrosa de patata de 24 g/L con micelio de la placa. El cultivo de crecimiento fue incubado por 3 días a 30°C agitándose a 130 RPM. Después de 3 días de crecimiento los 150 ml de la célula y los medios gastados fueron utilizados para inocular un medio de inducción del pericarpio del maíz. El medio de inducción fue de 80 g de pericarpio de maíz desalmidonado en 1 L de base (15 g KH_2PO_4, 5g de sulfatos de amonio, 20 g de extracto de levadura, 0.5 g de sulfato de magnesio y 1 g de Tween 80 en pH4.8). El cultivo fue mantenido a 30°C con agitación a 130 RPM por 7 días.

La proteína extracelular fue separada del grano gastado y de la masa fungicida de la célula por centrifugación en 2.000 x G por 20 Min. El sobrenadante clarificado parcialmente pasó primero por el filtro de 0,45 µm y, a continuación, a través de un filtro µm 0,22. El líquido filtrado clarificado fue almacenado a 4°C hasta ser utilizado. El filtrado se concentró aproximadamente 14 veces con Amicon ultra-4 filtros centrífugos con un límite de 30 kD (Millipore).

Las actividades L-α-arabinofuranosidasa (LARF) y β-xylosidasa (BXL) fueron purificadas del sobrenadante fungicida, primero intercambiando las proteínas en 50 mM pH 6.3 (N-Morfolino) de la sustancia amortiguadora ácido-NaOH etanosulfonico (MES), luego aplicando la proteína sobre la columna de intercambio de anion amina GE Health Science Q Sepharose (20 ml de volumen de lecho equilibrado con la sustancia amortiguadora MES). El sistema de purificación de proteína utilizado fue un sistema GH Health Sciences Akta que utiliza software Unicorn. Las proteínas se eluyen con un gradiente de 0-0,5 M NaCl en el tampón MES.

La actividad en fracciones de la columna se superviso usando 4 βnitrofenilo - D-xylopyranosida (para la actividad BXL) o 4 αnitrofenilo - L-arabinofuranosida (para la actividad LARF). En ambos casos, se agregaron 2 µL de la fracción de la columna a 95 µL de 50 milímetros del tampón de acetato del sodio y 5 µL del sustrato del p-nitrofenilo, mezclado e incubado 2 minutos a 23°C. 100 µL de 1 N solución de parada del carbonato de sodio de fue agregada y se leyó la absorbencia a 405 nanómetros.

Se recogieron las fracciones activas de la enzima del sistema del intercambio de aniones, concentrándose en la filtración de Amicon Centriprep y aplicado la columna de intercambio catiónico del sulfopropyl GE Health Sciences SP-Sepharose Fast Flow (lecho equilibrado de 20 ml con el tampón de acetato del sodio) utilizando el mismo sistema de la purificación de la proteína descrito anteriormente. Las proteínas se eluyen con un gradiente de 0-1 M NaCl en el tampón de acetato de sodio y la actividad en las fracciones fue supervisada como se describió anteriormente. Las fracciones activas luego fueron recogidas, reunidas y concentradas en caso de necesidad. Las proteínas en fracciones activas fueron identificadas por MS-MS.

7.1.12 Identificación de proteínas mediante MS-MS

La identificación de un conjunto representativo de *familia de* Fusarium verticillioides GH54 L-α-arabinofuranosidasa (LARF), familia GH51 LARF, dos familias GH3 β-xylosidasas y una familia GH30 β-xilosidasa (BXL) se describe a continuación.

Para la identificación de GH54 LARF, 200 µl de 1 N HCl (EM Industries, Hawthorne, NY) se agregaron a 250 µl de la "Fracción B3" (0,35 mg de proteínas) del procedimiento de purificación descrito anteriormente y la muestra se enfrió durante 10 minutos. Luego se agregaron 200 µl de 50% TCA (Sigma-Aldrich, St. Louis, MO) y las

muestras se enfriaron otros 10 min para precipitar las proteínas. Las muestras fueron centrifugadas por 2 minutos a 16.000 rcf y se desecharon los sobrenadantes. La proteína granulada fue lavada con 90% de acetona fría (J.T. Baker, Phillipsburg, NJ) y, a continuación, se centrifugó por 1 min a 16.000 rcf. Se descartó el sobrenadante y el granulado se dejó secar al aire. En primer lugar, 30 µl de urea de 8 M (MP Biomedicals, Solon, OH) fue añadido al granulo, luego 4 µl de 0,2 M DTT (Sigma-Aldrich, St. Louis, MO). La muestra fue incubada a 52°C durante 30 minutos y luego se agregaron µL 4 de yodoacetamida de 0.44 M (Sigma-Aldrich, St. Agregaron a Louis, MES) y la solución fue incubada a temperatura ambiente en la oscuridad por 30 Min. Luego se agregó lentamente 120 µl de 0,1% de agua n-octil B-D-glucopiranosida (Sigma-Aldrich, St. Louis, MO) y a continuación se dividió la muestra en dos partes iguales. A una mitad de la muestra se le agregó 7,5 µg tripsina (Promega Corp., Madison, WI) y a la otra mitad se agregó 4 µg de AspN (Roche Applied Science, Indianapolis, IN). Ambas muestras fueron incubadas a 37°C durante 1 H. Las muestras fueron mitigadas con 10% de ácido Trifluoroacetico (Thermoscientific, Waltham, mA) para un volumen final de 0.1% TFA. Ambas muestras fueron corridas en Thermofinnigan (San Jose, CA) un espectómetro de masas atrapa iones en una ionización por electrospray (ESI) LCQ-Deca . Una columna Vydac C18 (5 u, 300A, 0,2 x 150 mm, recursos biológicos Michrom, Auburn, CA) se corrió a un caudal de 200 µl/min. Los volúmenes de inyección fueron 50 µL y se filtraron con un cartucho de captura en línea (péptido CapTrap, Michrom recursos biológicos, Auburn, CA) antes de cargarlos en la columna. Se realizó la separación de los procesos con el siguiente gradiente (Solvente A: 0,1% ácido trifluoroacético en H_2O (J.T. Baker, Phillipsburg, NJ), Solvente B: 0.08% ácido trifluoroacetico en acetonitrilo (J.T. Panadero, Phillipsburg, NJ)). Para la identificación de GH51 LARF, las dos β-xilosidasas GH3 y β-xilosidasa GH30, se utilizó el método descrito anteriormente, excepto que las muestras fueron digeridas con el Chymotrypsin (7.5 µg, Sigma-Aldrich, St. Louis, MES) además del Trypsin y AspN. Se utilizó el motor de búsqueda TurboSequest en Bioworks 3.31 (Thermo, San Jose) para identificar las proteínas de la base de datos de proteínas de los verticillioides *de* Fusarium. La base de datos de la proteína fue descargada del instituto BROAD http://www.broadinstitute.org/annotation/genome/fusarium_verticillioides/MultiDownloads.html

300

7.1.13 Inducción de enzimas *Fusarium verticillioides* por medio de dilución de mijo pretratado en amoniaco acuoso

Para la inducción de enzimas *Fusarium verticillioides* en respuesta a la dilución de mijo pretratado en amoníaco acuoso, se molió el mijo a < 1 mm de diámetro, luego se pretrató con 6% de base de amoníaco de peso seco a 50% inicial de materia seca inicial a una temperatura máxima de 16°C durante 90 minutos. El material resultante fue de 43,3% de sólidos y contiene 0,336% de amoníaco residual

El cultivo de F. verticillioides tipo salvaje se mantuvo en agar de dextrosa de patata (Sigma P6685) y se inoculó un medio rico de crecimiento de dextrosa de la patata de 24 g/L con micelio de la placa. El cultivo de crecimiento fue incubado durante 3 días a 24°C con agitación a 140 rpm, tiempo durante el cual se torno turbio con células *Fusarium* . Después del crecimiento de 3 días, 4 frascos cada uno de 100 ml de base de medio Christakapoulos (0.1% KH_2PO_4, 0.03% $CaCl_2$, 0.03% $MgSO_4$ x 7 H_2O, 2.61% Na_2HPO_4 x 7 H_20, 0.134% NaH_2PO_4 x 1 H_2O, 1.0% de fosfato de anhidro del amonio) fueron inoculados con 4 ml de la suspensión de la célula resultante. A la suspensión se le agregaron dos gramos de materia seca de dilución de mijo pretratado en amoniaco acuoso como la fuente única del carbón.

Después de agregar el mijo pretratado, se ajustó el pH a 6.5 y el frasco se agitó en espiral a 180 rpm a 23 °C durante 168 horas. Antes de la cromatografía de exclusión por tamaño, se analizaron muestras en diferentes momentos del ensayo utilizando el ensayo Bradford para proteínas, actividades de p-nitrofenil arabinofuranosidasa, xilosidasa p-nitrofenilo y p-nitrofenilo glucosidasa y geles de SDS-PAGE para la inducción de enzimas. El mijo pretratado intacto, 2% glucosa y ninguna fuente de carbono se incluyeron como controles paralelos y se encontró que tenían niveles mucho más bajos de inducción de enzima en el caso de la fuente de carbono diluido de mijo pretratado en amoníaco acuoso.

7.1.14 Fraccionamiento cromatográfico de exclusión de tamaño del caldo de cultivo de *Fusarium verticillioides*

Los medios de cultivo de los Fusarium verticillioides que contiene las enzimas expresadas de la inducción de 168 h descrita anteriormente fueron centrifugados a 3.500 x g para retirar ripios y células. El sobrenadante fue filtrado a través de un filtro de µ0,4 m dando un líquido claro amarillo profundo. El líquido fue concentrado (~4X) dos veces en un concentrador (UFC901024) grande de corte Amicon ultra 10K MW Se

361

reservó una muestra concentrada de 1,7 mL f para la cromatografía de exclusión de tamaño (SEC) y se realizó un ensayo para el contenido de proteína según el ensayo BCA como se describió anteriormente. La muestra concentrada de la proteína fue cargada sobre una columna Superdex 16/60 SEC equilibrada con un tampón pH 5.0 de 50 milímetros de acetato del sodio. La columna se eluyó en 1 mL/min a 4°C. Se monitoreó aa absorbencia de la muestra a 214 nanómetros y a 280 nanómetros. Después de la elución del volumen vacío, se recogieron fracciones de 1 mL en una placa de microtitulación profunda de polipropileno. Las fracciones con proteínas fueron agregadas en una placa aparte de microtitulación profunda y se midió la concentración de proteína total de todas las fracciones por medio del ensayo BCA y las actividades de arabinosidasa del p-nitrofenilo, xilosidasa del p-nitrofenilo y glucosidase del p-nitrofenilo según lo descrito en el ejemplo 1 (ensayo de la actividad del substrato Sintético).

7.1.15 Purificación de una serie de GH43, GH51 y homólogos de GH3 (sección 8.1) de la fermentación de *T. reesei* de cromatografía de intercambio catiónico

Se llevó a cabo la producción de enzimas del frasco de agitación de cultivos *Trichoderma reesei* de tipo silvestre como se describe en WO 2005/001036. Las preparaciones enzimáticas extracelulares fueron concentradas por filtración Amicon Centriprep en caso de necesidad, y aplicadas a la columna de intercambio catiónico del sulfopropyl GE Health Sciences SP-Sepharose Fast Flow (lecho de 20 ml equilibrado con tampón de acetato de sodio) utilizando un sistema de purificación de proteínas sistema GE Health Sciences Akta utilizando el software Unicorn. Las proteínas se eluyen con un gradiente de 0-1 M NaCl en el tampón de acetato de sodio y la actividad en las fracciones fue supervisada como se describió anteriormente. Las fracciones activas luego fueron recogidas, reunidas y concentradas en caso de necesidad. La actividad en fracciones de la columna se supervisaron utilizando 4 nitrofenilo - D-xylopyranosida (para la actividad BXL) o 4 nitrofenilo - L-arabinofuranosida (para la actividad LARF). En ambos casos, se agregaron 2 µL de la fracción de la columna a 95 µL de 50 milímetros pH 5.0 del tampón de acetato de sodio y 5 µL del sustrato del p-nitrofenilo, mezclado e incubado 2 minutos a 23°C. Se agregaron 100 microlitros de solución de parada de carbonato de sodio al 1 N y se midió la absorbencia a 405 nanómetros. En algunos casos el paso de la purificación de la proteína del intercambio de aniones fue útil antes del intercambio catiónico. En este caso, las fracciones de la enzima fueron intercambiadas en 50 milímetros pH 6.3 (N-Morfolino) tampón ácido-

NaOH ethanesulphonic (MES) y luego se aplicó la muestra de la proteína a la columna de intercambio cuaternario de aniones de la amina de GE HealthSciences Q-Sepharose Fast Flow (volumen de lecho de 20 ml equilibrado con el tampón MES). Se utilizó el sistema de purificación de la proteína Akta según lo descrito para el paso del intercambio catiónico. Las proteínas se eluyeron con un gradiente de 0-0,5 M NaCl en el tampón MES. Las enzimas activas entonces fueron intercambiadas en 50 milímetros, tampón de acetato del sodio del pH 5.0 antes de la cromatografía del intercambio catiónico.

7.1.16 Purificación de Xilanasas

La purificación de las Xianasas fue conducida en dos etapas:

Etapa 1: Cromatografía de interacción hidrofóbica (HIC). Se agregó (NH4) 2SO 4 al sobrenadante de la fermentación para lograr una concentración final de 1 M. La columna de separación utilizada fue una Hiprep fenil (highsub), 16/10, 20 mL. *Tampón A*: 20 milímetros de fosfato del sodio, pH 6.0. *Tampón B*: Tampón A + 1 M $(NH4)_2SO_4$. Elución escalonada: 45% B; 0% B; agua + glicerol al 10%.

Etapa 2: Filtración de gel (GF). Se recogió el eluido de 0% B de la columna HIC y se cargó en una columna HiLoad 26/60 Superdex 75 prep grade (320 ml, GE Healthcare). La fase móvil era de 20 milímetros de fosfato del sodio, pH 6.8, que contiene NaCl de 0.15 M. La Figura 26 muestra el perfil de la elución. La Figura 27 muestra la detección SDS-PAGE de la separación de dos etapas de AfuXyn 5.

Para purificar la proteína T reesei XYN3 se utilizó La visión aplicada de biosistemas BIOCAD® y el explorador Amersham Pharmacia Biotech AKTA . Aproximadamente 150 ml de Xilanasa 3, de un concentrado ultrafiltrado, fueron cargados en una columna de desalación Sephadex G-25M (volumen total ~525 ml) equilibrada en 10 milímetros en tampón TES, pH 6.8. 300-400 ml de esta muestra desalada fueron cargados en una columna del intercambio aniónico (resina cuaternaria de amina de alta densidad; Applied Biosystems Inc.). Luego, la proteína unida fue eluida utilizando un gradiente de sal entre 0-1 M de cloruro de sodio utilizando volúmenes de 8 columnas de 25 milímetros de tampón TES. La elución de la proteína se produjo entre 0-250 milímetros de cloruro de sodio y fue detectada utilizando 10% Bis-Tris NUPAGE ® SDS-PAGE (Novex).

363

Las muestras eluidas que contienen 3 xilanasa se concentraron a 10 mL con concentradores de membrana Vivaspin 5.000 MWCO (límite de peso molecular) (Vivascience; GE Healthcare). El concentrado luego fue cargado en una HIgh Load 26/60 Superdex 200 (identificación no 17-1071-01; GE Healthcare). La columna fue equilibrada con 25 milímetros de tampón TES, pH 6.8, conteniendo 100 milímetros de cloruro de sodio y la proteína unida fue eluida en volúmenes de 8 columnas del tampón TES con cloruro de sodio.

La pureza de la proteína eluida fue determinada utilizando 10% Bis-Tris NUPAGE® SDS-PAGE (Novex) y determinada a tener más del 95% de pureza.

7.1.17 Purificación de Fv43D

Al concentrado de la ultrafiltración (UFC) de los Fusarium verticillioides 43D se le intercambió el tampón y se dializó contra el tampón de acetato del sodio de 50 milímetros, pH 4.0, durante la noche. El material dializado fue pasado por una columna HiTrap de 1 ml preempacada con resina de flujo rápido de sefarosa de sulfopropilo (GE Healthcare) diseñada para utilizarse con jeringuilla. El Fv43D purificado fue eluido con 50 milímetros de tampón de acetato de sodio, pH 4.0, con 250 milímetros de cloruro de sodio. El tampón de acetato de UFC y del sodio fue empujado a través de la columna usando una jeringuilla de 5 ml conectada al conectador GE Healthcare. El Fv43D purificado fue dializado durante la noche contra el tampón de acetato de sodio de 50 milímetros, pH 4.0. La proteína purificada fue ensayada por SDS-PAGE, HPLC y con una espectroscopia de la masa para demostrar homogeneidad.

7.1.18 Purificación de Fv51A

Al concentrado de la ultrafiltración (UFC) de los Fusarium verticillioides 51A se le intercambió el tampón y se dializó contra el tampón de acetato de sodio de 50 milímetros, pH 5.0, durante la noche. El material dializado fue pasado a través de una columna RESOURCE 15 de 6 ml preempacada con medios de sulfonato de metílio (GE Healthcare). El UFC fue cargado en 1 mL/min contra el tampón de acetato de sodio de 50 milímetros, pH 5.0 y eluido en 5 mL/min contra el tampón de acetato de sodio de 50 milímetros, pH 5.0, usando un gradiente de cloruro de sodio de 0 a 250 milímetros. Las fracciones eluidas fueron recogidas y ensayadas por SDS-PAGE. Las fracciones con Fv51A purificado fueron concentradas usando un concentrador de 10.000 MWCO Vivaspin de Sartorius Stemdim Biotech. El Fv51A purificado fue dializado durante la

noche contra el tampón de acetato de sodio de 50 milímetros, pH 5.0. La proteína purificada fue ensayada por SDS-PAGE, HPLC y con una espectroscopia de la masa para demostrar homogeneidad. El sistema AKTA Explorer 100 AKTA de GE Healthcare fue utilizado para la purificación de Fv51A.

8. EJEMPLO 2: EXPRESIÓN DE LOS GENES INDIVIDUALES DE LA HEMICELULASA DE VARIAS ESPECIES EN TRICHODERMA REESEI

8.1 *Genes Fusarium verticillioides*

La secuencia Fv51A se obtuvo mediante la búsqueda del genoma de Fusarium verticillioides en la base de datos del Broad Institute (http://www.broadinstitute.org/) para homólogos de GH51 arabinofuranosidasa.

Los siguientes genes de *Fusarium verticillioides* se expresaron en *Trichoderma reesei*: Fv3A, Fv43A, Fv43B, Fv43D, Fv51A, Fv3B, Fv43C, Fv39A, Fv43E, Fv30A, Fv30B, y Fv43F. La secuencia de Fv3A fue obtenida buscando homólogos de la β-xilosidasa GH3 en el genoma de los Fusarium verticillioides . La secuencia anotada careció de una secuencia y se utilizó el programa Augustus de predicción de señal del gen (http://augustus.gobics.de/) para identificar la secuencia anteriores que contuvieron una secuencia de la señal. Las secuencias para Fv39A, Fv43A, Fv43B, Fv43D, Fv43E, y Fv30A fueron obtenidas buscando el genoma de los Fusarium verticillioides para homologs de la β-xilosidasa GH39, GH30, y GH43.

Los marcos abiertos de lectura de los genes de la hemicelulasa de interés fueron ampliados por PCR utilizando ADN genómico puro/exraido de los Fusarium verticillioides como la plantilla. El PCR thermocycler utilizado fue ADN Engine Tetrad 2 Peltier Thermal Cycler (BioRad Laboratories). La polimerasa de ADN utilizada fue Polimerasa PfuUltra II Fusion HS ADN (Stratagene). Los cebadores utilizados para ampliar los marcos abiertos de lectura fueron los siguientes:

Fv3A:

Cebador directo MH124 (5 ' - CACCCATGCTGCTCAATCTTCAG-3') (SEQ ID NO.: 52)

Cebador inverso MH125 (5 ' - TTACGCAGACTTGGGGTCTTGAG-3') (SEQ ID NO.: 53)

Fv43A:

Cebador directo MH075 (5 ' - CACCATGTGGCTGACCTCCCCATT-3') (SEQ ID NO.: 54)

365

Cebador inverso MH076 (5 ' - TTAGCTAAACTGCCACCAGTTGAAGTTG -3 ') (SEQ ID NO.: 55)

Fv43B:

Cebador directo MH077 (5'-CACCATGCGCTTCTCTTGGCTATTGT-3') (SEQ ID NO.: 56)

Cebador inverso MH078 (5 ' - CTACAATTCTGATTTCACAAAAACACC-3') (SEQ ID NO: 57)

Fv43D:

Cebador directo MH0815 (5'-CACCATGCAGCTCAAGTTTCTG-3') (SEQ ID NO.: 58)

Cebador inverso MH082 (5'-CTAAATCTTAGGACGAGTAAGC-3') (SEQ ID NO.: 59)

Fv51A:

Cebador directo SK1159 (5'-CACCATGGTTCGCTTCAGTTCAATCCTAG-3') (SEQ ID NO.: 60)

Cebador inverso SK1160 (5'–CTAGCTAGAGTAAGGCTTTCC-3') (SEQ ID NO: 61)

Fv39A:

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

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

Fv3B:

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

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

Fv30A:

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

Cebador inverso MH113 (5'-TTA GTT GGT GCA GTG GCC ACG-3') (SEQ ID NO.: 67)

Fv30B:

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

Cebador inverso MH115 (5'-CAG CCC TCA TAG TCG TCT TCT TC-3') (SEQ ID NO.: 69)

Fv43C:

Cebador directo MH079 (5'- CAC CAT GCG TCT TCT ATC GTT TCC -3') (SEQ ID NO.: 70)

Cebador inverso MH080 (5 '-CTA CAA AGG CCT AGG ATC AA -3') (SEQ ID NO.: 71)

Fv39A:

Cebador directo MH116 (5'-CAC CAT GCA CTA CGC TAC CCT CAC CAC-3') (SEQ ID NO.: 72)

Cebador inverso MH117 (5' - TCA AGT AGA GGG GCT GCT CAC C-3') (SEQ ID NO.: 73)

Fv43E:
Cebador directo MH147 (5'-CAC CAT GAA GGT ATA CTG GCT CGT GG-3') (SEQ ID No.: 74)
Ceador inverso MH148 (5'-CTA TGC AGC TGT GAA AGA CTC AAC C-3') (SEQ ID NO.: 75)

Fv43F:
Cebador directo MH149 (5'-CACCATGTGGAAACTCCTCGTCAGC-3') (SEQ ID NO.: 76)
Cebador inverso MH150 (5'-CTA ATA AGC AAC AGG CCA GCC ATT G-3') (SEQ ID NO.: 77)

Los cebadores directos incluyeron cuatro nucleotides adicionales (las secuencias - CACC) en el extremo 5' para facilitar la reproducción direccional en pENTR/D-TOPO (Invitrogen, Carlsbad, CA) (Figura 28). Las condiciones de PCR para ampliar los marcos abiertos de la lectura fueron las siguientes (a excepción de Fv51A): Paso 1: 94°C por 2 Min. Paso 2: 94°C por 30 sec. Paso 3: 57°C por 30 sec. Paso 4: 72°C por 30-45 sec. Los pasos 2, 3 y 4 fueron repetidos para 29 ciclos adicionales. Paso 5: 72°C por 2 Min. Para Fv51A, se utilizaron las siguientes condiciones: Paso 1: 94°C por 2 Min. Paso 2: 94°C por 30 sec. Paso 3: 56°C por 30 seg. Paso 4: 72°C por 30 seg. Los pasos 2, 3 y 4 fueron repetidos para 25 ciclos adicionales. Paso 5: 4°C hold.

Los productos de PCR de los marcos abiertos de lectura de la hemicelulasa correspondiente fueron purificados mediante un Kit de purificación Qiaquick PCR (Qiagen, Valencia, CA). Los productos purificados de PCR fueron clonados en el vector pENTR/D-TOPO, transformados en TOP10 *químicamente competente células E. coli* (Invitrogen, Carlsbad, CA) y plateado en las placas LA con 50 ppm kanamycin. La ADN Plasmid fue obtenida de los transformantes *E. coli* utilizando un kit de preparación QIAspin plasmid (Qiagen). Los datos de la secuencia para la ADN insertada en el vector pENTR/D-TOPO fueron obtenidos utilizando cebadores M13 directos e inversos (Sequetech, Mountain View, CA). Un vector pENTR/D-TOPO con la secuencia correcta de ADN del marco abierto de lectura de la hemicelulasa correspondiente fue recombinado con el vector de destino pTrex3gM (WO 05/001036, Figura 29) utilizando la mezcla de reacción clonase LR (Invitrogen, Carlsbad, CA) según las instrucciones del fabricante. El producto de la reacción clonase LR fue transformado posteriormente en *células* TOP10 *químicamente competentes E. coli* que luego fueron plateadas en el LA que contenía 50ppm de carbenicillina. Los vectores de la expresión resultante eran plasmids pTrex3gM que contenían los marcos abiertos de lectura correspondientes de hemicelulasa que resultaron de la recombinación entre los sitios attR1 y attR2 de

pTrex3gM y los sitios attL1 y attL2 de pENTR/D-TOPO; y el marcador de selección del Aspergillus nidulans acetamidase (*amdS*). El ADN de los vectores de la expresión que contienen los marcos abiertos de lectura correspondientes a la hemicelulasa fue aislado utilizando un kit Qiagen miniprep utilizado para la transformación biolistica de las esporas del Trichoderma reesei.

La transformación Biolistica del Trichoderma reesei con el vector de expresión de pTrex3gM que contenía el marco abierto de lectura correspondiente a la hemicelulasa fue realizada utilizando el siguiente protocolo Transformación de la cepa de *celulasa quad delete Trichoderma reesei* (Δ*cbh1*, Δ*cbh2*, Δ*eg1*, Δ*Eg2*) por bombardeo de helio se logró mediante un sistema Biolistic ® PDS-1000/He Particle Delivery de Bio-Rad (Hercules, CA) siguiendo las instrucciones del fabricante (véase publicaciones de patentes WO 05/001036 y Estados Unidos 2006/0003408). Los transformantes fueron transferidos a placas de selección acetamida (véase la publicación de la patente WO 2009114380). Los transformants estables fueron inoculados en placas filtro de microtitulación (Corning), conteniendo 200 µL/pocillo de medios mínimos del Glycine (6.0 g/l de glicina; 4.7 g/l $(NH_4)_2SO_4$; 5.0 g/l KH_2PO_4; 1.0 g/l $MgSO_4 \cdot 7H_2O$; 33.0 g/l PIPPS; pH 5.5) con la adición post estéril de ~2% de la mezcla de glucosa/soforosa (E.E.U.U. 7.713.725) del como la fuente de los elementos de traza del carbón, 10 ml/l de 100 g/l de $CaCl_2$, 2.5 ml/l de *T. reesei* (400X): 175 g/l de ácido cítrico anhidro; 200 g/l $FeSO_4 \cdot 7H_2O$; 16 g/l $ZnSO_4 \cdot 7H_2O$; 3.2 g/l $CuSO_4 \cdot 5H_2O$; 1.4 g/l $MnSO_4 \cdot H_2O$; 0.8 g/l H_3Bo_3). Los transformantes fueron cultivados en un cultivo líquido durante 5 días en una cámara rica $_o$2 alojada en una incubadora a 28°C. Se obtuvieron las muestras sobrenadante de la placa filtro de microplacas en un colector de vacío. Las muestras de sobrenadante se corrieron en geles 4-12% NuPAGE (Invitrogen, Carlsbad, CA) de acuerdo a las instrucciones del fabricante para comprobar si hay expresión. El gel fue manchado con la mancha simplemente Azul (Invitrogen, Carlsbad, CA). Purificación de las enzimas GH43, GH51 y GH3 de fermentaciones *T. reesei* fueron realizadas por cromatografía del intercambio catiónico según lo descrito en el ejemplo 1.

8.2 Genes de otras especies

Las secuencias de Pf43A y Pf51A fueron obtenidas ordenando del genoma del penicillium funiculosum . Las preguntas usadas para las búsquedas del genoma *P. funiculosum* fueron otros homólogos fungicidas GH43 y GH51 disponibles en los genomas públicos. Un programa de predicción de genes llamado Augustus

(http://augustus.gobics.de/) se utilizó para comprobar las secuencias intrón junto con el codón de inicio. La secuencia de Fv43D fue utilizada para consultar los genomas *Gibberella zeae* (*Fusarium graminearum*) y *Fusarium oxysporum* disponibles en la base de datos del Instituto Broad y recuperar secuencias de Gz43A y Fo43A respectivamente. Los genes para Gz43A y Fo43A fueron sintetizados por GeneArt (Geneart GmbH, Regensburg, Alemania) con la secuencia de señal CBH1 en lugar de la secuencia nativa de la señal. Ninguno de los dos genes contiene intrones. El gen Pf51A fue optimizado con codón y sintetizado por GeneArt con la secuencia de señal CBH1 en lugar de la secuencia original de la señal.

Pf43A y Pf51A fueron clonados y expresados en *Trichoderma reesei* utilizando los siguientes cebadores:

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)

Tres genes fueron clonados de *Aspergillus fumigatus* para la expresión en *Trichoderma reesei*: GH11 xilanasa 2 (AfuXyn2), GH11 xyilanasa 5 (AfuXyn5) y GH43 Af43A. Los cebadores para los cebadores directos de la reproducción AfuXyn2, de AfuXyn5, y Af43A se demuestran a continuación:

AfuXyn2:

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

A.fumi-Q4WG11-R:

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

AfuXyn5:

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

A.fumi-Q4WF28-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)

Los métodos utilizados para la clonación y expresión de todos estos genes fueron similares al procedimiento descrito para clonar genes de *de Fusarium* . Los genes adicionales incluyendo ésos enumerados en la tabla 1B, fueron reproducidos de manera similar.

9. EJEMPLO 3: PRUEBA PARA LA ACTIVIDAD DE LAS NUEVAS HEMICELULASAS EN SUBSTRATOS SINTÉTICOS

Las actividades de Fv3A, Fv43A, Fv43B, Fv43D y un número de otras proteínas, por ejemplo, la Tabla 2, fueron probadas sobre sustratos sintéticos pNPX y pNPA tal como se describe en el ensayo de actividad de sustrato sintético en el ejemplo 1. *T. reesei* Bxl1 fue utilizado en 0,7 g/l. Las otras muestras de la enzima y el control huésped Quad-Supr, se agregaron por volumen de cultivos de crecimiento (placas de microtitulación o escala de agitación del frasco). Por lo tanto, la actividad absoluta no se puede comparar a través de muestras sino es una indicación de una proteína expresada activa con la actividad relativa del pNPX y del pNPA demostrada en la tabla 2. La actividad sobre los sustratos para-nitrofenilos no se utiliza como un predictor de rendimiento en la sacarificación de la biomasa.

10. EJEMPLO 4: HIDRÓLISIS DE MAZORCA PRETRATADA POR PREPARATIVOS DE CELULASA Y HEMICELULOSA

10.1 Rendimiento de la sacarificación de proteínas expresadas

Se evaluó el desempeño de la sacarificación de proteínas expresadas como adiciones a una mezcla de enzimas con una deficiencia de L-α-arabinofuranosidasa . Se evaluaron candidatos de L-α-arabinofuranosidas en un ensayo de sacarificación de 4 días de la mazorca por adición a una mezcla de enzima de Accellerase® 1500 /*T. reesei* Xyn3/Fv3A. La criba se llevó a cabo como se describe en el Ensayo de Sacarificación de la Mazorca (ejemplo 1) con las siguientes enzimas y cantidades/concentraciones:

- Accellerase® 1500, TP (nitrógeno Total) 54,2 mg/mL
- *Trichoderma reesei* Xyn3, TP (purificado) 2,9 mg/mL

370

- Fv3A, 3.2 mg/ml TP (purificados)
- Fv51A, 7.8 mg/ml TP (purificados)
- 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)

La proteína total (TP) de muestras purificadas se determina por A280, a menos que se indique en forma diferente. La proteína total de muestras sin purificar fue determinada por BCA según instrucciones del fabricante, a menos que se indique en forma diferente. Accellerase® 1500 fue añadida a 20 mg/g de proteínas/celulasa; *Trichoderma reesei* Xyn3 fue agregado a 5 mg/g de proteínas/celulasa; se agregó Fv3A a 5 mg/g de proteína/celulasa. Fv51A, Mg51A, At51A, Pt51A, Ss51A, Vd51A, Cg51B, Af43A, Pf43A, o Fv43E fueron agregados en 1, 3, y 5 mg de celulasa de proteina/g. Después de 4 días de incubación a 50°C, 200 RPM, la placa del ensayo fue mitigado y analizado por HPLC para los azúcares solubles.

Se encontró que las enzimas mejoran el rendimiento de glucosa o xilosa o arabinosa o una concentración reducida de celobiosa o xilobiosa en esta mezcla de enzimas. Los resultados se demuestran en las figuras 30A y 30B.

El rendimiento de la sacarificación de proteínas expresadas también fue evaluado como adiciones a la mezcla de enzimas con una deficiencia de L-α-arabinofuranosidasa y β-xilosidasa. Los candidatos B-xilosidasas fueron evaluados en un ensayo de sacarificaciónde la mazorca de 3 días adicionando una mezcla de enzimas Accellerase® 1500 /*T. reesei* Xyn3. La criba fue conducida según lo descrito en el ensayo de la sacarificación de la mazorca (ejemplo 1) con las siguientes enzimas y cantidades/concentraciones :

- Accellerase® 1500, TP (nitrógeno Total) 54,2 mg/mL
- *Trichoderma reesei* Xyn3, TP (purificado) 2,9 mg/mL

- Fv3A, 3.2 mg/ml TP (purificados)
- Fv43D, 6.8 mg/ml (purificados)
- Pf43A, 2 mg/mL TP (TCA/BCA)
- Pf43B, 2.7 mg/ml del TP (BCA)
- Fv43E, 1.37 mg/mL TP (TCA/BCA)
- Fv43F, 2.8 mg/ml TP (BCA)
- Fv30A, 2.7 mg/ml del TP (BCA)

Accellerase® 1500 fue añadida a 17.9 mg/g de proteínas/celulasa; Trichoderma reesei Xyn3 fue agregada a 5 mg/g de proteínas/celulasa; se agregó Fv3A a 5 mg/g de proteína/celulasa. Fv3A, Fv43D, Pf43A, Pf43B, Fv43E, Fv43F, o Fv30A fueron agregados en 1, 3, y 5 mg de celulosa de proteína/g (Fv43E fue agregado solamente en 1 y 3 mg/g). Después de 3 días de incubación a 50°C, 200 RPM, la placa del ensayo se inactivó y se analizo en HPLC para los azúcares solubles. Los resultados se muestran en la Figura 31. Se encontró que las enzimas mejoran el rendimiento de glucosa o xilosa o arabinosa o la concentración reducida de celobiosa o xilobiosa en esta mezcla de enzimas.

Se evaluó el desempeño de la sacarificación de proteínas expresadas también como adiciones a una mezcla de enzimas con una deficiencia de xilanasa. Durante la construcción de laexpresión integrada cepas *Trichoderma reesei* (descritas en el ejemplo 9 a continuación), una cepa *T. reesei* (# 44) fue aislada que sobre expresó las proteínas Bgl1, Fv3A, Fv51A, Fv43D pero no sobre expresó la endo-xilanasa. Esta cepa fue utilizada como fondo al que se le agregaron xilanasas para medir de rendimiento. Los candidatos endo-xilanasa fueron evaluados en un ensayo de 3 días de sacarificación de la mazorca agregando a la enzima los productos de la cepa # 44. La criba fue conducida según lo descrito en el ensayo de la sacarificación de la mazorca (ejemplo 1) con las siguientes enzimas y cantidades/concentraciones :

- Enzima Cepa #44 78.6 mg/ml TP (Biuret modificado)
- *Trichoderma reesei* Xyn3,TP (purificado) 2,9 mg/mL
- AfuXyn2 3.3 mg/ml TP (purificado)
- AfuXyn3 5.8 mg/ml TP (purificado)
- AfuXyn5 14.8 mg/ml TP (purificado)

372

- PfuXyn1 1.9 mg/ml TP (purificado)
- SspXyn1 1.2 mg/ml TP (purificado)

La composición de enzima producida por la Cepa # 44 fue añadida a la celulosa de 20 mg/g de proteínas; se agregaron las candidatas de enzimas de xilanasa a 3 y 7 mg de celulosa de proteína/g. Después de 3 días de incubación a 50°C, 200 RPM, la placa del ensayo se inactivó y se analizo en HPLC para los azúcares solubles. Se encontró que las enzimas mejoran el rendimiento de xilosa, o glucosa o arabinosa o que reduce la concentración de celobiosa o xilobiosa en esta mezcla de enzimas. Los resultados se muestran en la Figura 32.

Se evaluó y se comparó el desempeño de la sacarificación de proteínas expresadas Fv51A y Pa51A en las mezclas de enzimas con una deficiencia de L-α-arabinofuranosidasa. Se evaluaron Fv51A y Pa51A en un ensayo de 3 días de sacarificación de la mazorca agregando a una mezcla de enzima de Accellerase® 1000 /*T. reesei* Xyn2/Bxl1 o Fv3A. La criba se llevó a cabo como se describe en el Ensayo de Sacarificación de la Mazorca (ejemplo 1) con las siguientes enzimas y cantidades/concentraciones:

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

Accellerase® 1000 fue añadida a 20 mg celulasa de proteína/g; Trichoderma reesei Xyn2 fue agregado a 5 mg de celulosa de proteínas/g; se agregó Trichoderma reesei Bxl1 o Fv3A a 5 mg de celulosa de proteína/g. Fv51A fue agregado en 5 mg de celulosa de protein/g. Pa51A fue agregado en 1, 2, o 5 mg de celulosa de proteína/g. Después de 3 días de incubación a 50°C, 200 RPM, la placa del ensayo se inactivó y se analizo en HPLC para los azúcares solubles. Se encontró que las combinaciones de enzimas mejoran el rendimiento de xilosa, o glucosa o arabinosa o que reduce la concentración de celobiosa o xilobiosa en esta mezcla de enzimas. Los resultados se muestran en la Figura 33.

Fv51A y Pa51A también fueron evaluados en un ensayo de la sacarificación de la mazorca de 3 días agregando a una mezcla de enzima de Accellerase ® 1000/ *Trichodermadel reesei Xyn2*. Los Fv51A (29 mg/mL TP) y Pa51A (29 mg/mL) purificados fueron utilizados en esta parte del estudio. Accellerase® 1000 fue agregada a 17.5 mg de celulosa de proteína/g; *Trichoderma reesei* Xyn2 fue agregado a 4.4 mg de celulosa de proteína/g. Fv51A fue agregado a 4.4 mg de celulosa de proteína/g. Pa51A fue agregado a 0.9, 1.8 y 4.4 mg de celulosa de proteína/g. Después de 3 días de incubación a 50°C, 200 RPM, la placa del ensayo se inactivó y se analizaron los azucares solubles en HPLC. Se encontró que las combinaciones de enzimas mejoran el rendimiento de xilosa, o glucosa o arabinosa en esta mezcla de enzimas. Los resultados se muestran en la Figura 34.

El rendimiento de sacarificación de proteínas expresadas también fue evaluado como adiciones a la mezcla de enzima con deficiencia de β-xilosidasa. Se evaluaron los candidatos B-xilosidasa en un ensayo de 3 días de sacarificación de la mazorca agregando a la mezcla de enzimas de Accellerase® 1000 / *Trichoderma reesei* Xyn2/Fv51A. La criba se llevó a cabo como se describe en el ensayo de sacarificación de la mazorca (ejemplo 1) con las siguientes cantidades y concentraciones y enzimas:

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

Accellerase® 1000 fue añadida a 20 mg celulasa de proteína/g; Trichoderma reesei Xyn2 fue agregada a 5 mg de celulosa de proteínas/g; se agregó Fv51A a 5 mg de celulosa de proteína/g. Fv39A, Fv30B, Fv3A, o Fv3B fueron agregados en 1, 2, o 5 mg de celulosa de proteina/g. Después de 3 días de incubación a 50°C, 200 RPM, la placa del ensayo se inactivó y se analizaron los azucares solubles en HPLC. Se encontró que las combinaciones de enzimas mejoraron el rendimiento de glucosa o xilosa, o

arabinosa o que redujeron la concentración de celobiosa o en esta mezcla de enzima, con o sin β-xilosidasa . Los resultados se muestran en la Figura 33.

10.2 Actividad del candidato endo-xilanasas con xilano de madera

La actividad del candidato endo-xilanasas fue evaluado utilizando madera xilano como un sustrato mediante el siguiente ensayo. 90 Ml de una solución de 1% (peso/vol) xileno de madera (Sigma X 0502) fue añadido a los pocillos de una placa de microplacas de 96 pocillos y pre-incubado a 50 ° C durante 10 minutos. Las dilusiones enzimaticas y los estándares del xiliosa fueron agregados (10 µL) a la placa de microplaca y a las placas incubadas a 50°C por 10 Min. Mientras tanto, 100 µL100 de la solución del DNS fue agregada a los tubos de PCR. Después de la incubación de 10 minutos, los 60 µL de la reacción enzimática fueron transferidos a los tubos de PCR que contenían la solución del DNS. Los tubos fueron incubados en un thermocycler a 95°C durante 5 minutos y después enfriados a 4°C. Cien microlitros de la mezcla de reacción fueron transferidos a una placa de 96 pocillos y se midió la absorbencia a 540 nanómetro. Se generó una curva estándar de xilosa y se utilizó para calcular la actividad. Una unidad de xilanasa se define como la cantidad de enzima necesaria para generar 1 µmole de xilosa reduciendo el equivalente de azúcar por minuto en las condiciones del ensayo. Resultados se muestran en la tabla 3.

10.3 Hidrólisis de la enzima de oligómeros de arabinoxilano de la mazorca sacarificada.

En este estudio, la hidrólisis enzimática de los oligómeros de arabinoxilano permanece después de la digestión de amoníaco diluido pretratado de la mazorca con celulasa y preparativos se monitorearon las preparaciones de hemicelulasa. La preparación de oligomeros crudos se describe en el ejemplo 1. Las azúcares totales del oligómero fueron determinadas por HPLC (véase el ejemplo 1) después de la hidrólisis de ácido de oligomeros sin refinar con2% (v/v) de ácido sulfúrico en un frasco sellado a 121°C por 30 Min. Se corrigió un pequeña cantidad de degradación del azúcar en sus concentraciones según lo determinado por las muestras de control de las mezclas conocidas del azúcar tratadas con el mismo procedimiento. La concentración de azúcares totales en la preparación sin refinar del oligomero determinada por este método fue de 45 g/l de glucosa, 168 g/l de xilosa y 46 g/l de arabinosa. Al explicar el azúcar del monómero presente en oligomeros sin refinar antes de la hidrólisis de ácido, el 86% de la glucosa, el 90% del xilosa y el 43% de la arabinosa estaban presentes en

375

forma oligomerica. Varias β-xilosidasas, arabinofuranosidasas y mezclas de eso fueron probados para la conversión creciente del monomero de la arabinosa en la preparación de oligomeros sin refinar. La preparación del oligomero sin refinar se diluyo 20 veces a 12 g/l oligomeros en un tampón de acetato de sodio de 50 milímetros, pH 5.0 y mantenido a 50°C en un bloque calentador en tubos tapados Eppendorf de 1.5 ml. Las enzimas fueron agregadas en las concentraciones finales de 0.06-0.09 g/l e incubadas durante 24 h hasta su terminación. Las muestras entonces fueron retiradas para el análisis del HPLC de los azúcares del momomero según lo descrito en el ejemplo 1. Los resultados se enumeran en la tabla 4 como % de la conversión al monomero a base de azúcar en el azúcar total según lo determinado por hidrólisis de ácido.

Para obtener las producciones más altas (44-71%) de arabinosa para oligomeros arabinoxilano restantes en la mezcla sin refinar del oligomero, los datos en la tabla 4 demuestran que las combinaciones binarias de Fv3A+Fv51A, de Fv3A+Fv43B, y de Fv43A+Fv43B proporcionan los mejores resultados. De las familias de sequencias y de la actividad en los substratos artificiales, se deduce que Fv3A es una β-xilosidasa y Fv51A es una L-α-arabinofuranosidasa.

Se sabe que los azúcares de la arabinosa en arabinoxilano de la mazorca están ligadas con frecuencia a la xilosa en las 2 y 3 posiciones del carbón del azúcar de la arabinosa. Así la actividad de Fv3A es probable que hidrolice el vínculo del xil (1-2)ara que entonces hace disponible el vínculo del ara (1-3) xil a la hidrólisis por la L-α-arabinofuranosidasa. De las pruebas de β-xylosidasas solamente Fv3A y Fv43A parecían tener esta actividad. También en esta selección sólo Fv43B parecía tener actividad de L-α-arabinofuranosidase entre los miembros de la familia 43 de *Fusarium verticillioides*. Los resultados se muestran en la Tabla 4.

1. **EJEMPLO 5: GAMA DEL SUSTRATO DE B-XYLOSIDASAS PARA LA HIDRÓLISIS EFICAZ DE LA MAZORCA**

En este ejemplo, se determinó la gama del sustrato de 3 β-iylosidasas y su relación para una conversión eficaz de xilooligomeros de la mazorca a los azúcares del monomero. La preparación del hidrolizado de mazorca que contiene azúcares oligoméricos el ensayo de los azúcares de monómero se realizó tal como se describe en el ejemplo 1. Se determinaron los espectros NMR del protón de azúcares oligomericos con el grado de polimerización (DP) mayor que 2 según lo separado por cromatografía

de la exclusión Bio-Gel P2 (Figura 36 y Figura 37). Los espectros de oligómeros antes del tratamiento de la enzima tiene la etiqueta "Oligómeros MD07" en el panel inferior de la Figura 36 y los espectros del mismo oligo que contengan fracciones después tratamientos enzimáticos se encuentran en los grupos restantes de la Figura 36 y Figura 37 etiquetados con la enzima de tratamiento.

Las fracciones de Bio-gel P2 que contienen oligómeros superiores a DP2 (5-10 mg) fueron liofilizadas luego disueltas en 0,7 mL de una solución D_2O que contiene 0,5 mM 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) como patrón interno. Las alícuotas de 0.55 ml fueron utilizadas para las muestras NMR para optimizar la supresión del pico residual del agua. Se utilizaron versiones estándares de la secuencia de pulso de correlación Varian 2D, con la optimización de la anchura espectral de ^{13}C para los experimentos heteronucleares. Los espectros fueron adquiridos en una Varian Unity Inova utilizando una criosonda de alto sensibilidad funcionando a 500 MHz. La aclaración de la estructura fue hecha identificando las correlaciones que caracterizan el sistema de giro para los residuos individuales de azucar y luego identificar las correlaciones inter-glucosídicas.

La arabinosa que contiene oligomeros según lo determinado NMR, son dominados por uno o más estructuras ramificadas en las cuales la arabinosa es β- 1→3 unida a un residuo de la xilosa en un fragmento del polímero. El residuo de la arabinosa en ramificación resultante es sustituido aún más por residuo de xilosa β-1 unida →2 a la arabinosa. Sin esta segunda sustitución, hay muy poca arabinosa presente en los xilo-oligomeros restantes. *T. reesei* Bxl1 no es muy eficaz para segmentar la xilosa más alejada 1→2 a la unión arabinosa como lo demuestran las señales restantes en 5.5 a 5.55 ppm en los espectros. La combinación de la Fv43A que es eficaz en oligómeros de xilosa de cadena más largas y la Fv43B L-α-arabinofuranosidasa retira la mayoría pero no todas las especies ramificadas con señales en el rango de 5,5 ppm y a diferencia del tratamiento con *T. reesei* Bxl1, no deja ninguna señal en 5,35 que se atribuible a la arabinosa ramificada restante respectoal oligómero de xilosa.

La región de protones anoméricos de los espectros de los xllo-oligómeros restantes después del tratamiento con el Fv51A solos o en combinación con el Fv3A muestran resultados diferentes. La L-α- arabinofuranosidasa por sí sola no es eficaz para reducir la complejidad de las señales en la región. El Fv3A β- xilosidasa retira esencialmente toda la xilosa suyacente en la ramificación de la arabinosa dejando en su

377

mayoría estructuras unidas deoligómero de xilosa 13 arabinosa Añadir L-α-arabinofuranosidasa a la β- xilosidasa resulta en una conversión bastante completa a azúcares de monómero como lo demuestra el aumento de la señal procedente de los protones de anomérico α y β de azúcares reductores.

Por lo tanto puede concluirse que la efectividad aumento de una β-xilosidasa sobre otra es probable que sea debido a la gama del sustrato en términos de la complejidad estructural del aglicon permitido como sustrato.

12. EJEMPLO 6: COMPARACIÓN DE SECUENCIAS Y RESIDUOS CRÍTICOS EN LA DETERMINACIÓN DE LA GAMA DEL SUSTRATO EN GH3 B-XILOSIDASAS

El *T. reesei* β-xilosidasa (Bxl1) es 61% similar a la *F. verticillioides* ortholog (Fv3A) y comparte un 42% de identidad de secuencia como se muestra en la alineación de la Figura 38. Fv3A tiene gama mucho más amplia del sustrato que Bxl1, que es probable que se atribuya a los aminoácidos no-conservados entre las dos proteínas.

13. EJEMPLO 1: LA DETERMINACIÓN DE LA ACTIVIDAD DEFINIDA DE LA HEMICELULASA PARA LA HIDRÓLISIS DE LA MAZORCA A LOS MONÓMEROS DEL AZÚCAR

La mazorca pretratada según lo descrito en el ejemplo 1 fue utilizado como mezcla del agua ajustada a un pH 5 con H_2SO_4 en agua a sólidos/10g de 18.6 gr de mazorca seca de la mezcla total. 0.78 g de la mezcla fueron agregados a frascos de cristal de 4 ml y suficiente pH 5.0, 50 milímetros de tampón de acetato del sodio fue agregado para dar un peso total de la reacción de 1.06 g después de las adiciones deseadas de la enzima. Las hemicelulasas fueron agregados como las preparaciones purificadas descritas en el ejemplo 1. El sobrenadante de la cepa quad deleted *T. reesei* (Quad en las Tablas 5 y 6) es el concentrado de proteínas de fondo expresadas por la cepa *T. reesei* eliminada en 4 actividades principales de celulasa (descritas en WO 05/001036). Accellerase® 1000 es una mezcla completa de celulasa con alta actividad de β-glucosidasa. Los frascos fueron incubados en 230 RPM en una coctelera orbital a 48°C durante 72 h luego se agregaron 2 ml de agua. Se tomó una sub-muestra y luego se diluyó aún más, se centrifugo y filtró para el análisis HPLC de monómeros de azúcares como se describe en el ejemplo 1. Los resultados experimentales que definen cantidades útiles de actividad definida de la hemicelulasa para hidrolizar la mazorca pretratado en monómeros de azúcares se muestran en la tabla 5.

Los resultados desde el punto de vista de diseño de experimentos (DoE) se ajustaron a un modelo de superficie y se utilizaron para determinar la mejor relación entre los componentes de 7 enzimas en cuanto al mejor rendimiento de glucosa, xilosa y arabinosa en las dos concentraciones de proteínas totales probadas. Resultados de las relaciones de los componentes de las 7 enzimas se muestran en la tabla 6.

Otra exploración de las relaciones (tabla 7) fue realizada incluyendo Fv3A y nuevamente incluyendo Fv43D y manteniendo esa actividad constante a bajo nivel. La reacción instalada y las condiciones de la reacción fueron idénticas a aquellas descritas para el experimento completo de DoE.

Las reacciones (correr los números) 20, 21, y 22 contienen solamente la mezcla entera de la enzima celulasa y a los 21 mg/g del glucano monomerizado cerca del 48% de la glucosa presente en la mazorca y el 24% de la xilosa. Agregar la endoxilanasa, *Trichoderma reesei* Xyn3 permitió la disminución de la carga total de proteína celulasa, conservando sobre el mismo rendimiento de monómero de glucosa y aumentando el rendimiento de monómero de xilosa a alrededor del 40% (corrida # 1). Todas las combinaciones que dieron las producciones de la arabinosa por encima del 40% requirieron la combinación de Fv43D, Fv43A y Fv43B o Fv51A o de Fv3A y Fv51A o Fv43B. Esas combinaciones también tendieron a tener la liberación más alta de xilosa para el monómero de azúcar.

Otra serie de reacciones encaminadas a perfeccionar la mezcla requerida de hemicelulasas se ejecutó manteniendo constantes la carga de toda la celulasa y la carga de la endoxilanasa. La preparación de la celulosa completa Accellerase® 1000 se mantuvo constante 12 mg/g de glucano y la *T. reesei* Xyn3 endoxilanasa purificada se mantuvo a 6 mg/g de xilano. Fv51A fue la única L-α-arabinofuranosidase en la mezcla, en diferentes dosis. Otras condiciones de reacción permanecieron iguales pero el hidrolizado fue analizado por cromatografía de exclusión de tamaño como se describe en el ejemplo 1. La cuantificación de azúcares individuales fue realizada por área pico solamente y los resultados se muestran en la tabla 8.

Todas las combinaciones mostraron cantidades similares de glucosa y casi la misma cantidad de xilosa total soluble. El grado de reducción al monómero de xilosa varió por tratamiento. Sin actividad agregada de convertir oligómeros a monómeros cerca del 50% de la xilosa soluble seguía siendo oligomerica excepto que por lo menos se

agregara una β-xilosidasa. Fv3A a 2 mg Fv3A proteina/g xilano reducido a >DP2 oligómerosa 3.6 mg/ml. La adición 2 mg Fv51A proteina/g xilano a 2 mg Fv3A proteina/g xilano disminuyó áun más los oligomers >DP2 a 1.5 mg/ml. Unos 2 mg de xilano de Fv51A/g parecen ser suficientes para reducir los oligómeros >DP2 a un mínimo cuando los requerido 2 mg/g Fv3A estaba presente (Tabla 8).

Una mezcla de 6 mg/g xilano *T. reesei* Xyn3, Fv3A de 1 o 2 mg/g y 1 o 2 mg/g Fv51A es una carga adecuada para reducir toda la mazorca arabinoxilano en monómeros de azúcares. La adición de Fv43D a las mezclas de ayuda llevando la xilobiosa u otros oligomeros DP2 al monómero. La hidrólisis de la Arabinosa al monómero no fue medida en este experimento.

14. EJEMPLO 8: LA EFICACIA DE LAS HEMICELULASAS EN LA PRODUCCIÓN DE MONÓMEROS DE AZÚCAR DE LA MAZORCA

En este ejemplo se demuestra la eficacia de un conjunto de actividades de hemicelulasa purificada en la producción de monómeros de azúcares de arabinosa y xilosa cuando actúan solos en amoníaco diluido de mazorca pretratada. Tres mezclas (mezclas A, B, y C) de hemicelulasas purificadas fueron preparadas y utilizadas para hidrolizar la hemicellulose de la mazorca pretratada en 1 g total, 14% de las reacciones de sólidos preparados como en el ejemplo 1 y corridas bajo las condiciones descritas en el ejemplo 1. Los monómeros de azúcares fueron analizados por HPLC como descrito en el ejemplo 1 después de 72 h de reacción y las cantidades obtenidas se muestran en la Tabla 9.

- **Mezcla A:** 6 mg de Trichoderma reesei Xyn3, 4 mg Fv3A, 1 mg Fv51A/ por gramo de xilano.
- **Mezcla B:** 6 mg Trichoderma reesei Xyn3, Fv43D, 3 mg Fv43A, de 3 mg Fv43B por gramo xilano.
- **Mezcla C:** 6 mg de Trichoderma reesei Xyn3, 3 mg Fv3A, 1 mg Fv43D, 1 mg Fv51A por gramo de xilano.

En este experimento los conjuntos hemicelulasa definidos cedieron ligeramente menos monómero de azúcar que lo visto en anteriores experimentos que incluía actividades para solubilizar la celulosa. Las producciones seguían siendo mayores que ésos considerados solo con la adición de endoxilanasa a las

380

preparaciones enteras celulasa. Las actividades de la hemicellulase son eficaces para llevar el xilano a monómero.

El mismo conjunto de mezclas se utilizó para las preparaciones de la hemicelulosa fabricada a partir de la mazorca, rastrojo total del grano del sorgo, mijo y bagazo de caña de azúcar, utilizando los procedimientos de los métodos generales de conformidad con el ejemplo 1. Se realizaron las suspensiones comunes de cada preparación de la hemicelulosa a 100 mg/ml en 50 milímetros pH 5.0 en un tampón de acetato de sodio y el pH fue comprobado. Cada uno de ellos fue diluido a una preparación de 10 mg / mL con tampón de acetato de más de 50 mM. Las alícuotas de cada mezcla de enzima fueron agregadas a 100 µL de los 10 mg/ml de suspensión y las reacciones se corrieron dos veces, incubado a 48°C durante 6 h con agitación. Las reacciones fueron diluidas con 100 µL de agua, centrifugadas y filtradas antes del análisis HPLC para monómero de azúcares como se describe en el ejemplo 1. 200 µL de cada suspensión de hemicelulosa fue diluido con 200 µL de 0,8 N H_2SO4, autosegmentandose en 121°C durante 30 minutos en ciclo líquido luego filtradas y azúcares analizados por HPLC tal como se describe en el ejemplo 1. Los resultados demostrados en la Tabla 10 se divulgan como el monómero promedio de azúcar por la mezcla de enzima como porcentaje del azúcar hidrolizable ácida presente en la reacción.

Mezclas A y C realizada bien en hemicelulosa de mazorca y como se ve en otros experimentos sobre mazorca pretratada. Las mezclas que contengan Fv3A aumentan la conversión de arabinosa a monómero y dan una ligera ventaja en la conversión de xilosa a monómero. Todas las 3 mezclas funcionan bien en hemicelulosa purificada de otras monocotiledóneas. La mezcla de una endoxilanasa, uno o dos β-xilosidasas, uno de los cuales tiene especificidad de sustrato más allá de dos o tres unidades de xilosa enlazadas β 1→4 y una l α-arabinofuranosidasa da como resultado una mezcla de hemicelulasa eficaz contra hemicelulosa monocot.

15. EJEMPLO 9: CONSTRUCCIÓN DE LA CEPA INTEGRADA DE LA EXPRESIÓN TRICHODERMA REESEI

Se construyó una cepa expresión integrada Trichoderma reesei que co-expresó cinco genes: *T. reesei β-glucosidase* gen *bgl1*, *T. reesei* endoxylanase gen

xyn3, f. verticillioides β-xylosidase gen *fv3Af. verticillioides β-xylosidase* gen *fv43D*y *f. verticillioides α-arabinofuranosidase* gen *fv51A.*

La construcción de los casetes de expresión para estos genes diferentes y la transformación de la cepa *T. reesei* se describen a continuación.

15.1 Construcción del casete de expresión β- glucosidasa

La parte terminal N del gen original de T. reesei β-glucosidase bgl1 fue codón optimizado por ADN 2.0 (Menlo Park, Estados Unidos). Esta parte sintetizada compuesta por las primeras 447 bases de la región de codificación. Este fragmento fue ampliado PCR mediante cebos SK943 y SK941. La región restante del gen original *bgl1* fue ampliado po PCR de una muestra genómica de ADN extraída de la cepa *T. reesei RL-P37* , usando cebos SK940 y SK942. Estos dos fragmentos de PCR del gen *bgl1* fueron fundidos juntos en una reacción de fusión PCR, usando cebos SK943 y SK942:

Cebador directo SK943:
(5' - CACCATGAGATATAGAACAGCTGCCGCT-3') (SEQ ID NO:88)
Cebador inverso SK941:
(5' - CGACCGCCCTGCGGAGTCTTGCCCAGTGGTCCCGCGACAG-3') (SEQ ID. NO.: 89))
Cebador directo SK940:
(5' - CTGTCGCGGGACCACTGGGCAAGACTCCGCAGGGCGGTCG-3') (SEQ IDDNO.: 90)
Cebador inverso SK942:
(5' - CCTACGCTACCGACAGAGTG-3') (SEQ ID NO.: 91)

Los fragmentos resultantes de la fusión PCR fueron clonados en Gateway ® vector pENTR ™/D-TOPO ® y transformados en E. coli One Shot ® TOP10 células químicamente competente (Invitrogen) resultando en el vector intermedio, pENTRY-943/942 (figura 39). Se determinó la secuencia nucleotida del ADN insertado. El vector de pENTRY-943/942 con la secuencia correcta de *bgl1* fue recombinado con pTrex3g usando un protocolo de reacción de LR clonase ® esbozado por Invitrogen. La mezcla de reacción clonase LR se transformó en *e. coli One Shot ® TOP10* células químicamente competentes (Invitrogen), resultando en el vector de expresión final, pTrex3g 943/942 (figura 40). El vector contiene también el gen *Aspergillus nidulans amdS* que codifica la acetamidasa como un marcador seleccionable para la

transformación de *T. reesei*. El casete de expresión era PCR amplificado con cebos SK745 y SK771 para generar el producto para la transformación de la cepa, usando el método de electroporación descrito en WO 08153712.

Cebador directo SK771: (5' – GTCTAGACTGGAAACGCAAC -3') (SEQ ID NO:94)
Cebador inverso SK745: (5' – GAGTTGTGAAGTCGGTAATCC -3') (SEQ ID NO:95)

15.2 Construcción de la expresión endoxilanasa

El gen original T. reesei endoxilanasa xyn3 fue PCR amplificada de una muestra de ADN genómica extraído T. reesei, usando cebos xyn3F 2 y xyn3R-2.

Cebador directo xyn3F-2:
(5'–CACCATGAAAGCAAACGTCATCTTGTGCCTCCTGG-3') (SEQ ID NO:94)
Cebador inverso (xyn3R-2):
(5'-CTATTGTAAGATGCCAACAATGCTGTTATATGCCGGCTTGGGG-3') (SEQ ID NO:95)

Los fragmentos resultantes PCR fueron clonados en Gateway ® vector pENTR ™/D-TOPO ® y transformados en E. coli One Shot ® TOP10 células químicamente competentes (Invitrogen) resultando en el vector intermedio, pENTR/Xyn3 (Figura 41). Se determinó la secuencia nucleotida del ADN insertado. El vector pENTR/Xyn3 con la secuencia correcta de *xyn3* fue recombinado con pTrex3g usando un protocolo de reacción de LR clonase ® esbozado por Invitrogen. La mezcla de reacción LR clonase se transformó en *e. coli One Shot ® TOP10* células químicamente competentes (Invitrogen), resultando en el vector de expresión final, pTrex3g/Xyn3 (Figura 42). El vector contiene también el gen *Aspergillus nidulans amdS* que codifica la acetamidasa como un marcador seleccionable para la transformación de *T. reesei*. El casete de expresión era PCR amplificado con cebos SK745 y SK822 para generar el producto para la transformación de la cepa, usando el método de electroporación.

Cebador directo SK745: (5' – GAGTTGTGAAGTCGGTAATCC-3') (SEQ ID NO:96)
Cebador inverso SK822: (5' – CACGAAGAGCGGCGATTC-3') (SEQ ID NO:97)

15.3 Construcción del *casete de expresión de β-xilosidasa* Fv3A

El gen F. verticilloides β-xilosidasa fv3A fue amplificado de una muestra de ADN genómico de F. verticilloides utilizando los cebos MH124 y MH125.

Cebador directo MH124: (5' – CAC CCA TGC TGC TCA ATC TTC AG -3') (SEQ ID NO:98)
Cebador inverso MH125: (5' - TTA CGC AGA CTT GGG GTC TTG AG -3') (SEQ ID NO:99)

Los fragmentos resultantes PCR fueron clonados en Gateway ® vector pENTR ™/D-TOPO ® y transformados en E. coli One Shot ® TOP10 células químicamente competentes (Invitrogen) resultando en el vector intermedio, pENTR -Fv3A (Figura 43). Se determinó la secuencia nucleotida del ADN insertado. El vector pENTR-Fv3A3 con la secuencia correcta *fv3A* fue recombinado con pTrex6g usando un protocolo de reacción de LR clonase ® esbozado por Invitrogen. La mezcla de reacción LR clonase se transformó en *e. coli One Shot ® TOP10* células químicamente competentes (Invitrogen), resultando en el vector de expresión final, pTrex6g/Fv3A (Figura 44). El vector contiene también un mutante resistente de etilo chlorimuron del gen origianl *T.reesei* acetolactate sintasa (*als*), designado *alsR*, que se utiliza junto con su promotor orignal y terminador como un marcador seleccionable para la transformación de *T. reesei* (WO2008/039370 A1). El casete de expresión era PCR amplificado con cebos SK1334, SK1335 y SK1299 para generar el producto para la transformación de T. reesei , usando el método de electroporación, (véase p.ej.,WO2008153712 A2) .

Cebador directo SK1334: (5' – GCTTGAGTGTATCGTGTAAG -3') (SEQ ID NO:100)
Cebador directo SK1335: (5' – GCAACGGCAAAGCCCCACTTC -3') (SEQ ID NO:101)
Cebador inverso SK1299: (5' - GTAGCGGCCGCCTCATCTCATCTCATCCATCC -3') (SEQ ID NO:102)

15.4 <u>Construcción del casete de expresión de β-xilosidasa Fv43D</u>

Para la construcción el casete de expresión F verticilloides β-xilosidasa Fv43D, el producto del gen fv43D fue amplificado de F.verticilloides ADN genómico cebadores SK1322 y SK1297. La región del gen promotor de la endoglucanasa *bgl1* fue ampliado por PCR de una cepa RL-P37 genomica de ADN extraída de *T. reesei* , usando cebos SK1236 y SK1321. Estos dos fragmentos de ADN de PCR amplificado fueron fundidos juntos en una reacción de fusión PCR, usando cebos SK1236 y SK1297: El fragmento PCR resultante de la fusión fue clonado en pCR-Blunt II-vector TOPO (Invitrogen) para dar el plásmido TOPO Blunt Pegl1-Fv43D (figura 45) y las céclulas *e. coli One Shot ® TOP10*

químicamente competentes (Invitrogen) fueron transformadas mediante este plásmido. Se extrajo el ADN Plásmido de varios clones *E.coli* clones y confirmados por digestión de restricción.

Cebador directo SK1322: (5'–CACCATGCAGCTCAAGTTTCTGTC-3') (SEQ ID NO:103)
Cebador inverso SK1297: (5'–GGTTACTAGTCAACTGCCCGTTCTGTAGCGAG-3') (SEQ ID NO:104)
Cebador directo SK1236: (5'–CATGCGATCGCGACGTTTTGGTCAGGTCG-3') (SEQ ID NO:105)
Cebador inverso SK1321: (5'-GACAGAAACTTGAGCTGCATGGTGTGGGACAACAAGAAGG-3') (SEQ ID NO:106)

El casete de expresión fue PCR amplificado con cebos SK1236 y SK1297 de TOPO Blunt/Pegl1-Fv43D para generar el producto para la transformación de T. reesei, usando el método de electroporación descrito en WO2008153712A2.

15.5 Construcción del casete de expresión de α-arabinofuranosidasa

Para la construcción del casete de expresión del gen F verticilloides α-arabinofuranosidase fv51A, el producto del gen fv51A fue amplificado del ADN genómico F.verticilloides utilizando los cebos SK1159 y SK1289. La región del gen promotor de la endoglucanasa *egl1* fue ampliado por PCR de una muestra de ADN genómico T reesei extraído de la cepa RL-P37 , *usando cebos SK1236 y SK1262. Estos dos fragmentos de ADN de PCR amplificado fueron fundidos juntos en una reacción de fusión PCR, usando cebos SK1236 y SK1289: El fragmento PCR resultante de la fusión fue clonado en pCR-Blunt II-vector TOPO (Invitrogen) para dar el plásmido TOPO Blunt Pegl1-Fv51A (Figura 46) y las células e. coli One Shot ® TOP10 químicamente competentes (Invitrogen) fueron transformadas mediante este plásmido.*

Cebador directo SK1159: (5'–CACCATGGTTCGCTTCAGTTCAATCCTAG-3') (SEQ ID NO:107)
Cebador inverso SK1289: (5'-GTGGCTAGAAGATATCCAACAC-3') (SEQ ID NO:108)

Cebador directo SK1236: (5'–CATGCGATCGCGACGTTTTGGTCAGGTCG-3') (SEQ ID NO:109)

Cebador inverso SK1262:

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

El casete de expresión fue amplificado por PCR con cebos SK12985 y SK1289 para generar el producto para la transformación de T reesei, usando el método de electroporación.

Cebador directo SK1298: (5'-GTAGTTATGCGCATGCTAGAC-3') (SEQ ID NO:111)

Cebador inverso SK1289: (5'-GTGGCTAGAAGATATCCAACAC-3') (SEQ ID NO:112)

15.6 Co-Transformación de *T. reesei* con los βcasetes de expresión - glucosidasa y endoxilanasa

Una cepa mutante *Trichoderma reesei*, derivada de RL-P37 (Sheir-Neiss, G *et al* Appl. Microbiol. Biotechnol. 1984, 20:46 - 53) y seleccionada para alta producción del cellulasa fue co - transformada con el βcasete de expresión glucosidasa (promotor de *cbh1* , *T.reesei* β- gen terminador cbn1 , y *marcador del amdS*), y el casete de la expresión de endoxylanasa (promotor de *cbh1* , de *T.reesei xyn3* y de *cbh1*terminador) usando la transformación PEG-mediada (Penttila, M *y otros.* Gene 1987, 61 (2): 155-64). Los numerosos transformantes fueron aislados y examinados para βla producción de glucosidasa y endoxilanasa. Un *transformante llamado cepa T. reesei #229* fue utilizada para la transformación con los otros casetes de expresión.

15,7 Co transformación de la cepa *T. reesei* #229 con dos casetes de expresión β- xilosidasa y α-arabinofuranosidas

La cepa T. reesei # 229 fue co-transformada con el casete de expresión β-xilosidasa fv3A (cbh1 promotor, gen fv3A terminador cbh1 y marcador alsR), el casete de expresión β-xilosidasa fv43D (egl1 promotor, gen fv43D, terminador original fv43D) y el casete de la expresión de α-arabinofuranosidasa (promotor egl , gen fv51A, terminador original fv51A) mediante electroporación. Se seleccionaron transformantes en placas de agar Vogels conteniendo chlorimuron etílico (80 ppm). El agar de Vogels fue preparado como sigue, por litro.

solución común de50 x Vogels (abajo)	20 ml

Agar BBL20 g
con H2O desionizado llevan la adición post-estéril: 980 ml
50% Glucosa20 ml

solución común de Vogels x 50 (WO 2005/001036), por litro:

En 750 ml H2O desionizada, disolver sucesivamente:

$Na_3Citrate*2H_2O$	125.00 g
KH_2PO_4	(anhidro) 250.00 g
NH_4NO_3	(anhidro) 100 g
$MgSO_4*7H_2O$	10.00 g
$CaCl_2*2H_2O$	5.00 g
Solución Vogels Trace Element	5,0 mL
Solución Vogels Biotin2.5 ml	
con H2O desionizada,	llevar a 1 L

Los numerosos transformantes fueron aislados y examinados para la producción de L-α-arabinofuranosidasa y β-xilosidasa. También se examinaron transformantes para rendimiento de conversión de biomasa de acuerdo con el ensayo de sacarificación de la mazorca descrito en el ejemplo 1. Los ejemplos de *T. reesei* de cepas de expresión integradas descritas son H3A, 39A, A10A, 11A y G9A, que expresan todos los genes de *T. reesei* beta-glucosidasa 1, *T. reesei* Xyn3, Fv3A, Fv51A y Fv43D, en diferentes proporciones (tabla 11). Los ejemplos de *T. el reesei* de cepas de expresión integradas descritas aquí también incluye 44A, 69A, G6A y 102, y cada uno incluye la mayor parte de los genes para *T. reesei beta-glucosidasa 1,* T. reesei XYN3, Fv3A, Fv51A, y Fv43D, expresado en diversos proporciones. La cepa 44A faltante sobreexpresada *T. reesei* XYN3; la cepa 69A faltante Fv51A (confirmado por Western Biot, no se muestra); las cepas G6A y 102 faltantes Fv3A (tabla 11), según lo determinado por el análisis de la proteína de HPLC (ejemplo 1).

16. EJEMPLO 10: FUNCIONAMIENTO DE LA SACARIFICACIÓN DE *T. REESEI* CEPAS DE EXPRESIÓN INTEGRADA SOBRE MAZORCA PRETRATADA CON AMONIACO

Se evaluó el rendimiento de la sacarificación de composiciones de enzima producido por *T. reesei* con cepas de expresión integrada en mazorca pretratada en amoníaco diluido . *Se generaron muestras de la enzima T. reesei* como concentrados de ultrafiltración (UFC) o como centrato. Para la generación de la UFC, se obtuvieron caldos de fermentación T. reesei (escala 14 L) después de la separación de células por centrifugación, concentrado con membrana de ultrafiltración a través de un corte de membrana de peso molecular de Millipore 10 kD . Luego se ajusto el pH a 4.8. El caldo

de separación de célula fue pulido por filtración, usando la filtración de FW6 Buchner. En cada muestra de la enzima se probó la concentración de la proteína total usando el método modificado del Biuret.

El rendimiento de sacarificación se evaluó en viales o frascos de agitar. En cada preparación enzimática se probó el funcionamiento de la sacarificación en una carga de 20% sólidos secos (DS) de la mazorca pretratada diluida en amoniaco (véase, WO2006/110901). Todas las reacciones de la sacarificación entonces fueron tituladas con ácido sulfúrico a pH 5.0 y el azide de sodio fue agregado a una concentración final de 0.01% (peso/volumen), para controlar de contaminación microbiana. Cada reacción de la sacarificación fue dosificada con 20 mg de la preparación enzimática de la proteína total (TP) por g de glucano o xilano de sustrato, según fuera apropiado. Accellerase® 1500 (Ac1500) y la cepa integrada UFC fueron dosificadas en 20 mg de proteína/g glucano total Se preparó una mezcla de enzima en una relación de 25:9:4:3:1 Accellerase® 1500:Xyn3:Fv3A:Fv51A: FV43D. En la mezcla, Accellerase® 1500 fue dosificado en 20 mg proteína/g glucano total y las hemicelulasas fueron dosificadas por gramo de xilano (4.3 mg de Xyn3/g de xilano, 1,7 mg Fv3A/g de xilano, 1,4 mg Fv51A/g de xilano, 0,5 mg Fv43D/g de xilano). Cada reacción de la sacarificación fue incubada a 50°C en un agitador de rotación fijado a 200 rpm, luego se tomaron muestras y se diluyó 10 x (v/v), antes de determinar la concentración monomérica del azúcar usando el análisis HPLC (detallado en la sección 16.1 bajo la sección "del análisis monomerica del HPLC") después de 1, 2, 3 y 7 días de la sacarificación (Figuras 47A, 47B). El 3 día de la sacarificación, cada reacción también fue muestreada por peso (w/w) para la concentración oligomerica del azúcar por el análisis HPLC (sección 16.1, Figura 47C). Estos resultados muestran que las composiciones enzimáticas producidas por cepas integradas H3A y G9A proporcionan una mejor glucosa y rendimientos de xilosa que Accellerase ®[1500 de] mezclas de enzima creadas de las enzimas expresadas individualmente.

La Figura 47D muestra una comparación cromatografíca de la composición de la enzima producida por tres diferentes cepas integradas.

16.1 <u>Análisis HPLC</u>

<u>Análisis monomérico HPLC de azúcar</u>: Cada muestra fue analizada por HPLC mediante una columna de exclusión de ion BioRad Aminex HPX-87 H (300 x 7,8 mm).

Todo el día 1, 2, 3 y 7 muestras fueron diluidas 10 x volumétricamente con ácido sulfúrico de 5 mM, filtradas a través de un filtro de 0,2 µm antes de inyectar en HPLC y se corrieron bajo las especificaciones de fabricación.

Análisis HPLC de azúcar monomérica (Hidrólisis ácida): Las muestras de la sacarificación del día 3 fueron diluidas 10 x por peso con agua Milli-Q, luego se agregó el ácido sulfúrico a la concentración final de 4% (w/w). Una xilosa y glucosa estándar (estándar de recuperación de azúcar – "SRS") de concentración conocida también se preparó y se midió el análisis HPLC de azúcar monomérica como se indicó anteriormente y el análisis HPLC de azúcar de oligómero. Las muestras HPLC de Oligomero y fueron autosegmentadas a 121°C por 15 minutos y filtradas a través de un filtro de 0.2 µm antes de la inyección en la columna de la exclusión del ion de BioRad Aminex HPX-87H del HPLC (300 milímetros x 7.8 milímetros), y se corrió según las especificaciones del fabricante. Se determinó la concentración del azúcar del Oligomero multiplicando el porcentaje retenido de la concentración de xilosa y glucosa del estándar de azúcar recuperado(SRS) después de hidrólisis de ácido, por las concentraciones de azúcar del HPLC de la hidrólisis de ácido post de cada muestra, luego restando la concentración monomeric del azúcar determinada en la sección anterior "Análisis HPLC del azúcar monomerica".

17. EJEMPLO 11: IDENTIFICACIÓN DE CLASES DE ENZIMA FILOGENÉTICAMENTE AMPLIAS QUE PUEDEN MEJORAR EL RENDIMIENTO DE SACARIFICACIÓN DE LA CEPA DE EXPRESIÓN INTEGRADA DE *T. REESEI*

En este ejemplo, se identificaron las actividades de la enzima que limitan la eficacia de las composiciones de enzima producida por una cepa de expresión integrada de *T. reesei* y se ejemplificaron las enzimas de diferentes especies que pueden compensar las actividades de limitación o lo que se podría sustituir por componentes de cepa integrada.

0.95g de la mazorca pretratada según lo descrito en el ejemplo 1 fue agregado a los frascos del cristal de 20 ml. Se agregó suficiente pH 5.0, 50 mM de tampón de acetato del sodio y 1 N H_2SO_4 para dar un peso total de la reacción de 3.00 g en mazorca seca de solidos/100g de la mezcla total, adiciones post enzima de pH 5.0. A la composición de la enzima T. reesei producido por cepa integrada H3A (figuras 48, 49A y 50A) se agregó como concentrado ultrafiltrado *(es decir,* libre de células) a 7 mg de proteína total por g de glucano y xilano combinado en la materia prima a todos los

frascos. Además, las hemicelulasas del candidato fueron agregados como las preparaciones ultrafiltradas a partir de 14 L cultivos de fermentación expresadas por la cepa *T. reesei Quad delete* (descrita en WO 05/001036). Las hemicelulasas candidatas fueron agregados en la proteína enzimática a 0, 0.5, 1.0 o 3.0mg por g de glucano y de xilano combinados. Las hemicelulasas incluyeron los componentes de la composición enzimatica producida por la cepa integrada misma, incluyendo:

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

Así como las enzimas de diversos hongos

- *Fusarium oxysporum* Fo43A
- *Gibberella zeae* Gz43A
- *Penicillium brasilianum* Pf43A
- *Aspergillus fumigatus* Af43A
- *Podospora anserina* Pa51A
- *Penicillium brasilianum* Pf51A

Los frascos fueron incubados a 180 RPM en una coctelera orbital a 48°C durante 72 h. Luego se agregaron 12 ml de agua. Se tomó una sub-muestra y luego se diluyó aún más, se centrifugo y filtró para el análisis HPLC de monómeros de azúcares como se describe en el ejemplo 1.

Los resultados que se muestran en la Figura 48A y 48B demostraron que la adición de las actividades de β-xylosidase, como Fusarium verticillioides Fv43D y Fusarium oxysporum Fo43A notablemente mejoraron la liberación del monómero de xilosa relativo a 7mg de la proteína total de la cepa integrada de T. reesei por g de glucano y xilano combinados. Incluso se observaron mejoras significativas en las adiciones bajas de 0.5 y 1.0 mg/g.

Los resultados demostrados en la figura 49A y 49B demuestran que la adición de todas las hemicelulasas conduce a un aumento significativo en glucosa del monomero, incluso el >10% en las cargas de hemicelulasa de 1-3 mg por g del Glucan y del Xylan combinados.

Los resultados que se muestran en la figura 50A y 50B demostraron que además de la adición de enzimas GH51 como Fusarium verticillioides Fv51A y especialmente Podospora anserine Pa51A y Penicillium brasilianum Pfu51A, condujeron a un aumento en el nivel de arabinosa de monómero.

18. EJEMPLO 12: SACARIFICACIÓN DE MIJO VARIADAMENTE PRETRATADO CON PREPARACIONES DE CELULASA Y HEMICELULASA

Se evaluó el rendimiento de sacarificación de celulasas expresadas y hemicelulasas en mijo sin procesar pretratado. Se evaluó una gama de condiciones para el pretratamiento del mijo con amoníaco diluido para el desempeño de la sacarificación con una enzima cóctel compuesta de enzimas aquí descritas. Las condiciones de pretratamiento varían y eficacia del pretratamiento afecta el rendimiento de hidrólisis enzimática.

El pretratamiento del mijo sin procesar fue interpretado en tubos de acero inoxidable sellados, 6 x ½ pulg., inmersos en un baño de arena caliente. Una mezcla liquida de mijo sin procesar, agua e hidróxido de amonio (~ 28% solución) fue mezclada hasta el porcentaje deseado de sólidos y porcentaje de amoníaco y luego se cargó en un tubo de pretratamiento. Los tubos se mantuvieron a la temperatura deseada (+/- 2°C) por el tiempo deseado y luego se inactivaron en el agua helada durante aproximadamente 1 minuto antes de traerse a temperatura ambiente. La mezcla pretratada fue quitada de los tubos y permitida secarse durante la noche en la capilla (>90% de sólidos alcanzables).

Los sólidos pretratados secos fueron sacarificados al 10% de sólidos, pH 5, 50°C, 200 rpm usando Accellerase ® 1500, 3 Xyn, Fv3A, Fv51A y Fv43D (25, 9, 7, 3, 1 mg total proteína/g glucano o xilano respectivamente). Se utilizó un volumen total del hidrolisato de 5 ml en frascos de centelleo de 20 ml.

Las producciones de Glucano y de xilano se basaron en glucosa monomérica y xilosa liberada comparadas con el glucano o xilano disponible en la biomasa sn procesar. Las concentraciones Monomeric del azúcar fueron medidas por HPLC (columna de BioRad Aminex HPX 87-H).

El tiempo del pretratamiento, la temperatura, el porcentaje de sólidos y el porcentaje de NH_3, fue variado sobre una amplia gama a fin de optimizar los resultados

de sacarificación. Cada condición del pretratamiento tuvo producciones de glucano y xilano mejor que ~50% que se considera como un ejecutor fuerte. Los parámetros del pretratamiento que tuvieron fuertes producciones se enumeran en la tabla 12 junto con sus respectivos glucanos, xilanos y porcentajes totales de las producciones. Las conversiones del Glucano y o Xilano se basan en las azúcares monomérico liberadas durante la sacarificación con respecto al glucano o al xilano teóricamente disponible en el mijo sin procesar. La conversión total es glucosa y xilosa solamente (Figura 51).

19. EJEMPLO 13: SACARIFICACIÓN DEL MIJO PRETRATADO CON PREPARACIONES DE CELULASA Y HEMICELULASA

Además de los ejemplos anteriores, se probó el rendimiento de la sacarificación de las mezclas de enzimas y composiciones de enzimas producidas por una cepa integrada en varios sustratos, pretratamientos y condiciones. Estos experimentos demuestran la gama del funcionamiento usando la mezcla de la enzima o un producto de cepa integrado. Demuestran buen rendimiento en una amplia gama de sustratos y pretratamientos, pH y temperaturas.

El mijo pretratado en amoníaco diluido fue preparado de acuerdo a los métodos y rangos de procesos en WO06110901A: Mijo (38,7% glucano, xilano de 22,2%, 2,5% arabinan, 23,2% lignina) fue molido con martillo para pasar a través de una pantalla de 1 mm y, a continuación, tratado a 160oc durante 90 minutos con 6% NH3 (peso / peso DM, agregado como NH 4OH). Este sustrato pretratado fue tratada con mezclas de enzima que contiene Accellerase® 1500, Multifect® xilanasa (ambos productos comerciales de Danisco A/S, División de Genencor, Palo Alto, CA), Fv3A, Fv51A y Fv43D en una masa de reacción total de 50 g de sólidos de 15%. La proteína total (TP) de los productos comerciales fue determinada por el ensayo de Biuret. Las otras enzimas fueron concentrados de ultra filtración (UFC) siguiendo la expresión en cepas de celulasa quad-deleted *T. reesei*, con TP determinada por análisis de nitrógeno Total de proteínas precipitables TCA. Todas las reacciones fueron dosificadas con Accellerase® 1500 a 25 mg glucano TP/g y se agregaron Multifect ® xilanasa (MF Xyl) a 9 mg xilano TP/g y Fv3A, Fv51A y Fv43D como se indica en la figura 56 A- 56B a 3,6 mg TP/g xilano, 3.0 mg xilano TP/g y 1,0 mg xilano TP/g, respectivamente. Todas las enzimas fueron dosificadas relativas al contenido de carbohidrato del mijo antes del pretratamiento. Las reacciones de sacarificación se llevaron a cabo en 47 ° y 33 ° C a pH 5.3 durante tres días.

Los resultados a 33 ° C demostraron que la adición de Fv3A, Fv51A y Fv43D incrementaba la conversión de glucano (figura 56A) y se duplicó mucho más la conversión de xilano (figura 56B). Los resultados a 47 ° C demostraron que la adición de Fv3A, Fv51A y Fv43D dieron algunas conversiones de mayor glucano (figura 56 A) y más del doble de la conversión de xilano (figura 56B). Las adiciones de solo Fv51A o Fv43D, aumentó ampliamente la conversión de xilano, especialmente a xylo-oligómeros o xilosa monómeros, respectivamente. La adición solo de Fv3A aumentó los rendimientos de la xilosa, pero en combinación con Fv51A produjo un gran aumento en la conversión de xilano a monómero.

20. EJEMPLO 14: SACARIFICACIÓN DEL MIJO PRETRATADO POR UNA *CEPA T.* REESEI INTEGRADA

El rendimiento de sacarificación de una composición de enzima producida por una cepa integrada *T. reesei* (H3A) fue evaluada en el mijo pretratado diluido en amoníaco de acuerdo con los métodos y rangos de procesos en WO06110901A: El mijo (37% glucano,21% xilano 5% arabinan, 18% lignina) fue molido con martillo para pasar a través de una pantalla de 1 mm y, a continuación, fue pretratado a 160°C durante 90 minutos con 10.0% NH3 (peso / peso DM, agregado como NH 4OH). En frascos Erlenmeyer de vidrio duplicados de 500 mL, 50 g de mezcla líquida pretratada en 25% de sólidos fue sacarificado a 48°C, pH 5.3 durante 7 días, con sobrenadante de la cepa integrada dosificada en 14mg TP/g de hidratos de carbono (glucano más xilano). El TP de H3A fue determinado por análisis total del nitrógeno de la proteína precipitable TCA. Todas las enzimas fueron dosificadas relativas al contenido de carbohidrato inicial del mijo antes del pretratamiento.

Al final de los 7 días, los altos niveles de conversión de glucano (52–55%, Figura 57A) y conversión de xilano (51% –53%, (figura 57B) fueron medidos por HPLC. Estos resultados demuestran que la composición de la enzima producida por la cepa integrada puede sacarificar mijo pretratado en amoníaco diluído en sólidos altos (materia seca del 25%).

21. EJEMPLO 15: SACARIFICACIÓN DE LA PULPA DE LA MADERA DURA POR UNA CEPA *INTEGRADA T.*REESEI

El rendimiento de sacarificación de una composición de enzima producida por una cepa *integrada de* T reesei H3A fue evaluado en pulpa sin blanquear de madera

industrial (derivado del proceso Kraft y delignificación de oxígeno, Smurfit Kappa Cellulose Du Pin, Biganos, Francia) con la siguiente composición: Glucano el 75.1%, Xilano el 19.1%, lignina soluble en el ácido el 2.2%. Los estudios enzimáticos de la sacarificación fueron realizados usando el método de comprobación estándar de NREL LAP-009 "Enzymatic Saccharification of Lignocellulosic Biomass" (http://www.nrel.gov/biomass/pdfs/42629.pdf), salvo que la carga de la celulosa fue diferente (variando de 9.3-20%) y se utilizó una masa total de 100 g. La condición experimental fue de 200 RPM y pH 5.0, 50°C. La enzima fue dosificada en 20mg de glucano TP/g (basado en materia seca final) al principio del experimento. Las muestras fueron tomadas en intervalos sincronizados si eran licuados y luego analizadas por HPLC para la concentración del azúcar. La glucosa, xilosa y la concentración de celobiosa fueron determinados utilizando un sistema de Aguas HPLC (Alliance system, Waters Corp., Milford, mA). La columna HPLC usada para el análisis del azúcar fue de BioRad (columna de exclusión del ion de Aminex HPX-87H (300 milímetros x 7.8 milímetros), BioRad Inc., Hércules, CA). Todas las muestras fueron diluidas 10 x con 5 mM de ácido sulfúrico, filtradas a través de un filtro de 0,2 µm antes de inyectarse en HPLC. Como se indicó (tabla 13) se llevaron a cabo dos experimentos en modo "lote alimentado": La pulpa pretratada de la madera dura llevada a un contenido inicial de la materia seca de 7.0% en Tiempo 0 y el resto del sustrato fue agregado en momentos discretos en cuatro porciones iguales durante las primeras 24 horas para traer los sólidos secos finales cargando hasta 20%.

Los resultados mostraron altos niveles de conversión glucano y xilano a monómeros (hasta un 89% y el 90%, respectivamente) con la composición de enzima producida por la cepa integrada H3A. Conversiones aumentaron con mayores cargas de enzima y tiempos más lagos de sacarificación. La conversión fue más baja cuando los sólidos estaban al 20% y no al 15%, pero esta conversión más baja al 20% podría ser atenuada parcialmente usando un proceso de alimentación por lote.

22. EJEMPLO 16: SACARIFICACIÓN DE LA PULPA DE LA MADERA DURA POR CEPA *INTEGRADA DE T. REESEI* EN UNA GAMA DE TEMPERATURA Y PH

El rendimiento de la sacarificación de una composición de enzima producida por una cepa integrada *T. reesei* (H3A) sobre pulpa sin blanquear de madera industrial (como se utilizó en el ejemplo anterior) en 7% de carga de celulosa y 20 mg TP/g

glucano de sobrenadante de cepa integrada (basado en la materia seca final y en la composición del sustrato pretratado) se probó a temperaturas de 45° C - 60 ° C y pH de 4.65-5.4 (tampón en citrato de sodio de 0,1 M). Los resultados después de la sacarificación de 2 días se muestran en la TABLA 14 y muestran buenas conversiones del glucano y del xilano sobre la gama entera de las condiciones probadas. Estos resultados sugirieron las condiciones óptimas para la sacarificación de este sustrato como pH 4.9 y 50°C pero las buenas conversiones se ven incluso en pH 5.0, 60°C. En un experimento de seguimiento en 50°C, incluyendo pH inferior (y de otra manera sin modificar las condiciones experimentales) se vieron buenas sacarificaciones en pH 3.8, pH 4.0 y pH 4.25, con titulaciones de glucosa & xilosa de 45,1 & 9,7 g/L; 50.0 & 11.4 g/L; y 57,2 & 13.2 g/L, respectivamente.

23. EJEMPLO 17: LA SACARIFICACIÓN DEL BAGAZO DE CAÑA DE AZÚCAR AMPLIADO POR VAPOR CON UNA COMPOSICIÓN DE LA ENZIMA PRODUCIDA POR UNA CEPA INTEGRADA EN DIVERSAS DOSIS DE LA ENZIMA

El rendimiento de sacarificación de una composición de enzima producida por una cepa integrada *T. reesei* (H3A) del bagazo de caña de azúcar ampliado por vapor fue evaluado en 7% de carga de celulosa. El bagazo fue pretratado con inyección de vapor en un reactor de StakeTech a 210 psig, 200°C con un tiempo de residencia de 4 minutos. El material pretratado tenía la siguiente composición: Glucano 40.9%, Xilano 20.8%, lignina 27%. El sobrenadante de la cepa integrada fue dosificado en 10, 20, 30, 50 u 80mg de glucano TP/g (basado en materia seca final y en la composición del sustrato pretratado). La sacarificación fue realizada en un volumen de reacción de 5 ml a 50°C, pH 5 durante 3 días. Los resultados (figura 58A-C) mostraron que la cepa integrada producto superó el Accellerase® 1500 (20 mg de proteína/g de glucano) en la conversión del glucano a glucosa (figura 58A y figura 58C) y, especialmente, la conversión de xilano en xilosa (figura 58B y figura 58C) incluso a la mitad de la dosis de Accellerase ® 1500. La conversión del Glucan, especialmente en el día 1, aumentó perceptiblemente mientras que la dosis de la proteína total de H3A aumentó (figura 58A y figura 58C), mientras que la conversión del xilano era más rápida, con poco incremento del día 1 al día 3, y muy alto, incluso en la dosis más baja de la proteína total de H3A (figura 58B y figura 58C).

24. EJEMPLO 18: LA SACARIFICACIÓN DE LA FIBRA DE MAÍZ PRETRATADA CON ÁCIDO DILUIDO CON LAS VARIAS ENZIMAS

El rendimiento de la sacarificación de mezclas de enzima en fibra de maíz diluida con ácido sulfúrico se evaluó en un agitador de 250 mL. En un experimento típico, la fibra de maíz (composición inicial de 38% C6 azúcares y azúcares de 27% C5) se ajustó al 15% DS (sólidos secos) y se agregó ácido sulfúrico 0,36% (w/w%). La mezcla de la fibra del maíz luego fue autosegmentada a 121°C durante 60 Min. La mezcla luego fue ajustada a pH 5.0 utilizando NaOH de 6 N. El contenido de azúcar de la muestra pretratada fue de 21 g/l de glucosa y xilosa de 12 g/l. Las enzimas se añadieron al sustrato pretratado como sigue (como se indica en la figura 59A, 59B, 59C); Accellerase® 1500 (AC 1500): glucano TP/g de 20 mg y 45 mg TP/g glucano; Accellerase® 1500 +Multifect ® xilanasa (MF): (25 mg/g glucano + xilano de mg 9 TP/g) y (25 mg TP/g glucano + xilano de 20 mg TP/g); Accellerase® 1500 + Xyn 3++ Fv3A + Fv51A + Fv43D: (25 mg TP/g glucano xilano TP/g 9 mg + xilano TP/g 7 mg + xilano TP/g 3 mg + xilano TP/g 1 mg) (toda la enzima "mezclada"). La proteína total (TP) de los productos comerciales (Accellerase® 1500 y Multifect® xilanasa: Danisco los US Inc., Genencor) fueron determinados por el análisis de Biuret. Las otras enzimas fueron concentrados de ultra filtración (UFC) siguiendo la expresión en cepas de celulasa quad-deleted *T. reesei*, (como se describió anteriormente) se determinaron sus TP con análisis de nitrógeno Total de proteínas precipitables TCA. Todas las enzimas fueron dosificadas relátivas al contenido de carbohidrato inicial de la fibra de maíz antes del pretratamiento. La sacarificación enzimática fue realizada en 200 RPM y 50°C durante 24 h, 48 h y 120 H. Las muestras fueron retiradas en diversos intervalos del tiempo y analizadas para la formación de las azúcares de la glucosa y xilosa por HPLC. El azúcar ajustada (glucosa o xilosa) reflejó el azúcar producida del paso enzimático menos los azúcares iniciales.

Los resultados muestran que la mezcla de enzima completa superó el (Accellerase® 1500 + Multifect® C 1 8 xilanasa) ("1500 AC + MF") en la conversión de glucano y en la conversión de xilano, incluso cuando ambos fueron dosificados en la misma proteína total. La mezcla completa de enzima dio casi una conversión completa de glucano de este sustrato después de la sacarificación de 5 días.

25. REALIZACIONES ESPECÍFICAS Y LA INCORPORACIÓN POR REFERENCIA

Todas las publicaciones, patentes, solicitudes de patentes y otros documentos citados en esta aplicación se incorporan por referencia en su totalidad para todos los efectos en la misma medida como si cada publicación individual, patente, solicitud de patente u otro documento fueron individualmente indicados para ser incorporados por referencia para todos los efectos.

Aunque diversas realizaciones específicas han sido ilustradas y descritas, se apreciará que se pueden realizar varios cambios sin apartarse del espíritu y alcance de la divulgación/divulgaciones.LO QUE SE REIVINDICA ES:

1. Una composición que ocurre de manera no natural que comprende:

   a. un primer polipéptido o una variante de ello teniendo una actividad - xilosidasa, L-- arabinofuranosidasa, o las dos actividades xilosidasa y L - arabinofuranosidasa; y

   b. un segundo polipéptido o una variante de ello con una actividad de xilanasa.

2. La composición de la reivindicación 1, donde el primer polipéptido o variante de ello es un Fv43D, un Fv3A, un Pf43A, un Fv43E, un Fv39A, un Fv43A, un Fv43B, un Pa51A, un Fo43A, un Gz43A, un Trichoderma reesei *Bxl1*, un polipéptido de Fv51A, de Af43A, o de Pf51A.

3. La composición de la reivindicación 1 ó 2, donde el primer polipéptido o variante de ello tiene β- actividad de xilosidasa.

4. La composición de la reivindicación 3, donde el primer polipéptido o variante de ello con actividad de xilosidasa es un Fv43D, un Fv3A, un Pf43A, un Fv43E, un Fv39A, un Fv43A, un Fv43B, un Pa51A, un Fo43A, un Gz43A, o un polipéptidoTrichoderma reesei Bxl1.

5. La composición de cualquiera de las reivindicaciones 1-4, comprende más aún un tercer polipéptido o una variante de ello que tiene actividad de lα-arabinofuranosidasa.

6. La composición de la reivindicación 5, donde el tercer polipéptido o variante de ello es un polipéptido de Fv51A, de Af43A, de Fv43B, de Pa51A o de Pf51A.

7. La composición de cualquiera de las reivindicaciones 1-4, comprende más aún un tercer polipéptido que tiene actividad de β xilosidasa.

8. La composición de la reivindicación 7, donde el primer polipéptido es un polipéptido de Fv3A o un polipéptido de Fv43A, y el tercer polipéptido es un Fv43D, un Pa51A, un Gz43A, un Trichoderma reesei *Bxl1* , un polipéptido de Pf43A, de Fv43E, de Fv39A, de Fo43A, o de Fv43B.

9. La composición de cualquier de las reivindicaciones 2, 4, 6, y 8, donde, si presente, el polipéptido de Fv43D tiene por lo menos secuencia de identidad del 90% al SEQ ID NO.: 28, o a los residuos (i) 20-341, (ii) 21-350, (iii) 107-341, o (iv) 107-350 de SEQ ID NO.: 28.

10. La composición de cualquier de las reivindicaciones 2, 4, 6, y 8, donde, si presente, el polipéptido de Fv43D tiene por lo menos secuencia de identidad del 90% con la secuencia del aminoácido SEQ ID NO.: 2, o a los residuos (i) 24-766, (ii) 73-394, (iii) 73-394, (iv) 395-622, (v) 24- 622 o (vi) 73-622 de SEQ ID NO.: 2.

11. La composición de cualquiera de las reivindicaciones 2, 4, 6, y 8, donde, si presente, el polipéptido Pf43A tiene por lo menos secuencia de identidad del 90% con la secuencia del aminoácido SEQ ID NO.: 4, o con los residuos (i) 24-445, (ii) 21-301, (iii) 21-323, (iv) 21-444, (v) 302-444 (vi) 302-445, (vii) 324-444 o (viii) 324-445 de SEQ ID NO.:4.

12. La composición de cualquiera de las reivindicaciones 2, 4, 6, y 8, donde, si presente, el polipéptido de Fv43E tiene por lo menos secuencia de identidad del 90% con la secuencia del aminoácido SEQ ID NO.: 6, o con los residuos (i) 19-530, (ii) 29-530, (iii) 19-300 o (iv) 29-300 de SEQ ID NO.: 6.

13. La composición de cualquiera de las reivindicaciones 2, 4, 6, y 8, donde, si presente, el polipéptido de Fv39A tiene por lo menos secuencia de identidad del 90% con la secuencia del aminoácido SEQ ID NO.: 8, o con los residuos (i) 20-439, (ii) 20-291, (iii) 145-291 o 145-439 de SEQ ID NO.: 8.

14. La composición de cualquiera de las reivindicaciones 2, 4, 6, y 8, donde, si presente, el polipéptido de Fv43A tiene por lo menos secuencia de identidad del 90% con la secuencia del aminoácido SEQ ID NO.: 10, o con los residuos (i) 23-449, (ii) 23-302, (iii) 23-320, (iv) 23-448, (v) 303-448 o (vi) 321-449 de SEQ ID NO.: 10.

15. La composición de cualquiera de las reivindicaciones 2, 4, 6, y 8, donde, si presente, el polipéptido de Fv43B tiene por lo menos secuencia de identidad del 90% con la secuencia del aminoácido SEQ ID NO.:12, o con los residuos (i) 17-574, (ii) 27-574, (iii) 17-303 o (iv) 27-303 de SEQ ID NO.: 12.

16. La composición de cualquiera de las reivindicaciones 2, 4, 6, y 8, donde, si presente, el polipéptido de Pa51A tiene por lo menos secuencia de identidad del 90%

con la secuencia del aminoácido SEQ ID NO.: 14, o con los residuos (i) 21-676, (ii) 21-652, (iii) 469-652 o (iv) 469-676 de SEQ ID NO.: 14.

17. La composición de cualquiera de las reivindicaciones 2, 4, 6, y 8, donde, si presente, el polipéptido de Fo43A tiene por lo menos secuencia de identidad del 90% al SEQ ID NO.: 18, o con los residuos (i) 21-341, (ii) 107-341, (iii) 21-348, o (iv) 107-348 de SEQ ID NO.: 18.

18. La composición de cualquiera de las reivindicaciones 2, 4, 6, y 8, donde, si presente, el polipéptido de Gz43A tiene por lo menos secuencia de identidad del 90% con la secuencia del aminoácido SEQ ID NO.:16, o con los residuos (i) 19-340, (ii) 53-340, (iii) 19-383 o (iv) 53-383 de SEQ ID NO.:16.

20. La composición de cualquiera de las reivindicaciones 2, 4, 6, y 8, donde, si presente, el polipéptido de Trichoderma reesei Bxl1 tiene por lo menos secuencia de identidad del 90% con la secuencia del aminoácido SEQ ID NO.:44.

21. La composición de cualquiera de las reivindicaciones 2, 4, 6, y 8, donde, si presente, el polipéptido de Fv51A tiene por lo menos secuencia de identidad del 90% con la secuencia del aminoácido SEQ ID NO.: 32, o con los residuos (i) 21-660, (ii) 21-645, (iii) 450-645 o (iv) 450-660 de SEQ ID NO.: 32.

22. La composición de cualquiera de las reivindicaciones 2, 4, 6, y 8, donde, si presente, el polipéptido de Af43A tiene por lo menos secuencia de identidad del 90% al SEQ ID NO.: 20, o con los residuos (i) 15-558 o (ii) 15-2950 de SEQ ID NO.: 20.

23. La composición de cualquiera de las reivindicaciones 2, 4, 6, y 8, donde, si presente, el polipéptido de Pa51A tiene por lo menos secuencia de identidad del 90% con la secuencia del aminoácido SEQ ID NO.:22, o con los residuos (i) 21-632, (ii) 461-632, (iii) 21-642 o (iv) 461-642 de SEQ ID NO.::22.

24. La composición de cualquiera de las reclamaciones 2, 4, 6 y 8, donde, si presente, el polipéptido Fv43D está codificado por un ácido nucleico que tiene al menos secuencia de identidad del 90% con SEQ ID NO: 27 por un ácido nucleico capaz de hibridizar, bajo condiciones de alto rigor, con SEQ ID NO: 27, o a un fragmento de la misma.

25. La composición de cualquiera de las reivindicaciones 2, 4, 6 y 8, donde, si presente, el polipéptido Fv43D está codificado por un ácido nucleico que tiene al menos secuencia de identidad del 90% con SEQ ID NO: 1 con un ácido nucleico capaz de

400

hibridizar, bajo condiciones de alto rigor, con SEQ ID NO: 1, o a un fragmento de la misma.

26. La composición de cualquiera de las reivindicaciones 2, 4, 6 y 8, donde, si presente, el polipéptido Pf43A está codificado por un ácido nucleico que tiene al menos secuencia de identidad del 90% con SEQ ID NO: 3, por un ácido nucleico capaz de hibridizar, bajo condiciones de alto rigor, con SEQ ID NO: 3, o a un fragmento de la misma.

27. La composición de cualquiera de las reivindicaciones 2, 4, 6 y 8, donde, si presente, el polipéptido Fv43E está codificado por un ácido nucleico que tiene al menos secuencia de identidad del 90% con SEQ ID NO: 5 por un ácido nucleico capaz de hibridizar, bajo condiciones de alto rigor, con SEQ ID NO: 5, o con un fragmento de la misma.

28. La composición de cualquiera de las reivindicaciones 2, 4, 6 y 8, donde, si presente, el polipéptido Fv39A está codificado por un ácido nucleico que tiene al menos secuencia de identidad del 90% con SEQ ID NO: 7 con un ácido nucleico capaz de hibridizar, bajo condiciones de alto rigor, con SEQ ID NO: 7, o a un fragmento de la misma.

29. La composición de cualquiera de las reivindicaciones 2, 4, 6 y 8, donde, si presente, el polipéptido Fv43A está codificado por un ácido nucleico que tiene al menos secuencia de identidad del 90% con SEQ ID NO: 9 con un ácido nucleico capaz de hibridizar, bajo condiciones de alto rigor, con SEQ ID NO:9, o a un fragmento de la misma.

30. La composición de cualquiera de las reivindicaciones 2, 4, 6 y 8, donde, si presente, el polipéptido Fv43B está codificado por un ácido nucleico que tiene al menos secuencia de identidad del 90% con SEQ ID NO: 11 por un ácido nucleico capaz de hibridizar, bajo condiciones de alto rigor, con SEQ ID NO: 11, o con un fragmento de la misma.

31 La composición de cualquiera de las reivindicaciones 2, 4, 6 y 8, donde, si presente, el polipéptido Pa51A está codificado por un ácido nucleico que tiene al menos secuencia de identidad del 90% con SEQ ID NO:13, o por un ácido nucleico capaz de hibridizar, bajo condiciones de alto rigor, con SEQ ID NO: 13, o con un fragmento de la misma.

401

32. La composición de cualquiera de las reivindicaciones 2, 4, 6 y 8, donde, si presente, el polipéptido Fo43A está codificado por un ácido nucleico que tiene al menos secuencia de identidad del 90% con SEQ ID NO: 17, o con un ácido nucleico capaz de hibridizar, bajo condiciones de alto rigor, con SEQ ID NO:17, o a un fragmento de la misma.

33. La composición de cualquiera de las reivindicaciones 2, 4, 6 y 8, donde, si presente, el polipéptido Gz43A está codificado por un ácido nucleico que tiene al menos secuencia de identidad del 90% con SEQ ID NO:15, o con un ácido nucleico capaz de hibridizar, bajo condiciones de alto rigor, con SEQ ID NO:15, o a un fragmento de la misma.

34. La composición de cualquiera de las reivindicaciones 2, 4, 6 y 8, donde, si presente, el polipéptido Fv51A está codificado por un ácido nucleico que tiene al menos secuencia de identidad del 90% con SEQ ID NO:31, o con un ácido nucleico capaz de hibridizar, bajo condiciones de alto rigor, con SEQ ID NO:13, o a un fragmento de la misma.

35. La composición de cualquiera de las reivindicaciones 2, 4, 6 y 8, donde, si presente, el polipéptido Af43A está codificado por un ácido nucleico que tiene al menos secuencia de identidad del 90% con SEQ ID NO: 19 por un ácido nucleico capaz de hibridizar, bajo condiciones de alto rigor, con SEQ ID NO: 19, o con un fragmento de la misma.

36. La composición de cualquiera de las reivindicaciones 2, 4, 6 y 8, donde, si presente, el polipéptido Pf51A está codificado por un ácido nucleico que tiene al menos secuencia de identidad del 90% con SEQ ID NO:21, o por un ácido nucleico capaz de hibridizar, bajo condiciones de alto rigor, con SEQ ID NO: 21, o a un fragmento de la misma.

37. La composición de cualquiera de las reivindicaciones 1-36, donde el segundo polipéptido es un xilanasa de *Trichoderma* o *una xilanasa del Aspergillus* .

38. La composición de la reivindicación 37, donde la xilanasa de *Trichoderma* es un polipéptido de Trichoderma reesei *Xyn3* o un polipéptido de Trichoderma reesei *Xyn2* .

39. La composición de la reivindicación 38, donde, si presente, el polipéptido Trichoderma reesei Xyn3 tiene por lo menos secuencia de identidad del 90% al SEQ ID NO.: 42, o con los residuos 17-347 de SEQ ID NO.: 42.

402

40. La composición de la reivindicación 38, donde, si presente, el polipéptido de Trichoderma reesei Xyn2 tiene por lo menos secuencia de identidad del 90% con SEQ ID NO.: 43, o con los residuos 33-222 de SEQ ID NO.:43.

41. La composición de la reivindicación 38, donde, si presente, el polipéptido Trichoderma reesei Xyn3 está codificado por un ácido nucleico que tiene al menos secuencia de identidad del 90% con SEQ ID NO: 41, o con un ácido nucleico capaz de hibridizar, bajo condiciones de alto rigor, con SEQ ID NO:41, o a un fragmento de la misma.

42. La composición de la reivindicación 37, en donde *la xilanasa del Aspergillus* es un polipéptido AfuXyn2 o un polipéptido AfuXyn5.

43. La composición de la reivindicación 42, donde, si presente, el polipéptido de AfuXyn2 tiene por lo menos secuencia de identidad del 90% al SEQ ID NO.: 24, o con los residuos 19-228 de SEQ ID NO.: 24.

44. La composición de de la reivindicación 42, donde, si presente, el polipéptido de AfuXyn5 tiene por lo menos secuencia de identidad del 90% al SEQ ID NO.: 26, o con los residuos 20-313 de SEQ ID NO.: 26.

45. La composición de la reivindicación 42, donde, si presente, el polipéptido AfuXyn2 está codificado por un ácido nucleico que tiene al menos secuencia de identidad del 90% con SEQ ID NO: 23, o por un ácido nucleico capaz de hibridizar, bajo condiciones de alto rigor, con SEQ ID NO: 23, o con un fragmento de la misma.

46. La composición de la reivindicación 42 , donde, si presente, el polipéptido AfuXyn5 está codificado por un ácido nucleico que tiene al menos secuencia de identidad del 90% con SEQ ID NO: 25, o por un ácido nucleico capaz de hibridizar, bajo condiciones de alto rigor, con un complemento de SEQ ID NO: 25, o con un fragmento de la misma.

47. La composición de cualquiera reivindicación 1-46, que comprende además unos o más polipéptidos que tienen actividad del celulasa.

48. La composición de la reivindicación 47, donde los unos o más polipéptidos que tienen actividad del celulasa son componentes independientes de una celulasa entero.

49. La composición de la reivindicación 48, donde la celulasa entera es una β-celulaa entera de glucosidasa enriquecida.

50. La composición de la reivindicación 48 o 49, donde la celulasa entera es una celulasa entera fungicida filamentosa.

51. La composición de cualquiera de las reivindicaciones 48-50, donde el celulasa entera fungicida filamentosa es una preparación de unas especies *Acremonium, Aspergillus, Emericella, Fusarium, Humicola, Mucor, Myceliophthora, Neurospora, Penicillium, Scytalidium, Thielavia, Chrysosporium, Phanerochaete,Tolypocladium,* o *Trichoderma* .

52. La composición de cualquiera de las reivindicaciones 48-50, donde la celulasa entera es una celulasa entera *Chrysosporium lucknowense, Phanerochaete chrysosporium, Trichoderma reesei, Trichoderma harzianum, Trichoderma koningii, Trichoderma longibrachiatum, Trichoderma viride,* o *Pencillium brasilianum* .

53. La composición de la reivindicación 47, donde uno o más polipéptidos que tienen actividad de celulasa se seleccionan de una β- glucosidasa, una endoglucanasa y de una celobiohidrolasa.

54. La composición de la reivindicación 49 o 53, donde βla glucosidasa es un polipéptido Trichoderma reesei Bgl1

55. La composición de la reivindicación 54, donde, el polipéptido de Trichoderma reesei Bgl1 tiene por lo menos secuencia de identidad del 90% con SEQ ID NO.: 45, o con los residuos 33-744 de SEQ ID NO.:45.

56. La composición de cualquiera de las reivindicaciones 49 y 53-55, donde el β- el polipéptido de glucosidasa constituye hasta 50 wt.% del peso total de enzimas de la celulasa en la composición.

57. La composición de cualquiera de las reivindicaciones 49 y 53-56, donde β- el polipéptido de glucosidasa constituye 2 wt.% a 50 wt.% del peso total de proteínas en la composición.

58. La composición de cualquiera de las reivindicaciones 47-57, donde uno o más polipéptidos que tienen actividad de celulasa constituyen entre 30 wt.% a 80 wt.% del peso total de proteínas en la composición.

59. La composición de la reivindicación 58, donde uno o más polipéptidos que tienen actividad de celulasa juntos son capaces de alcanzar por lo menos 0.00005 producto de

404

la fracción por mg de proteína por gramo de la celulosa hinchada de ácido fosfórico (PASC) según lo determinado por un ensayo calcofluor.

60. La composición de cualesquiera reivindicaciones 1-59, comprende además una o más proteínas accesorias.

61. La composición de la reivindicación 60, donde una o más proteínas accesorias se seleccionan del grupo que consta de mananasas, galactanasas, arabinasas, ligninasas, amiloasas, glucuronidasas, proteasas, esterasas, lipasas, xiloglucanasas, polipéptidos de la familia 61 de hidrolasa glucósida, CIP1, CIP2, del swollenin y expansinas.

62. La composición de la reivindicación 60, en donde una o más proteínas accesorias se seleccionan del grupo que consiste en las proteínas como CIP1, CIP2, dehidrogenasas de celobiosa y peroxidasas de manganeso.

63. La composición de cualquiera de las reivindicaciones 60-62, donde una o más proteínas accesorias constituyen 1 wt.% a 50 wt.% del peso total de proteínas en la composición.

64. La composición de cualquiera de las reivindicaciones 1-63, capaz de convertir el 60% o más del xilano de la hemicelulosa de una biomasa en xilosa.

65. La composición de la reivindicación 64, capaz de convertir el 70% o más del xilano de la hemicelulosa de una biomasa en xilosa.

66. La composición de la reivindicación 65, capaz de convertir el 80% o más del xilano de la hemicelulosa de una biomasa en xilosa.

67. La composición de cualquiera de las reivindicaciones 64-66, donde la biomasa es mazorca, rastrojo del maíz, fibra del maíz, mijo, sorgo, papel, pulpa, o bagazo de la caña de azúcar.

68. La composición de cualquiera de las reivindicaciones 64-66, donde la biomasa es *Miscanthus.*

69. La composición de cualquiera de las reivindicaciones 1-68, además comprende una biomasa.

70. La composición de la reivindicación 69, donde el peso combinado de los polipéptidos que tienen actividad de xilanasa es 1 g a 40 g por 1 kilogramo de hemicelulosa en la biomasa.

405

71. La composición de la reivindicación 69, donde el peso combinado de los polipéptidos que tienen actividad de xilanasa es 0.5 g a 40 g por 1 kilogramo de hemicelulosa en la biomasa.

72. La composición de la reivindicación 71, donde el peso combinado de los polipéptidos que tienen actividad de xilanasa es 0.5 g a 20 g por 1 kilogramo de hemicelulosa en la biomasa.

73. La composición de cualquiera de las reivindicaciones 70-72, donde el peso combinado de los polipéptidos que tienen actividad de xilanasa es 2 g a 20 g por 1 kilogramo de hemicelulosa en la biomasa.

74. La composición de cualquiera de las reivindicaciones 69 - 73, donde el peso combinado de los polipéptidos que tienen actividad de xilosidasa es 1 g a 50 g por 1 kilogramo de hemicelulosa en la biomasa.

75. La composición de cualquiera de las reivindicaciones 69-73, donde el peso combinado de los polipéptidos que tienen actividad de xilosidasa es 0.5 g a 50 g por 1 kilogramo de hemicelulosa en la biomasa.

76. La composición de la reivindicación 75, donde el peso combinado de los polipéptidos que tienen actividad de xilosidasa es 0.5 g a 40 g por 1 kilogramo de hemicelulosa en la biomasa.

77. La composición de cualquiera de las reivindicaciones 74-76, donde el peso combinado de los polipéptidos que tienen actividad de xilosidasa es 2 g a 40 g por 1 kilogramo de hemicelulosa en la biomasa.

78. La composición de cualquiera de las reivindicaciones 69-77, donde el peso combinado de los polipéptidos que tienen actividad de L-arabinofuranosidasa, si presente, está entre 0-5 g a 20 g por 1 kilogramo de hemicelulosa en la biomasa.

79. La composición de cualquiera de las reivindicaciones 69-77, donde el peso combinado de los polipéptidos que tienen actividad de L-arabinofuranosidasa, si presente, es 0.2 g a 20 g por 1 kilogramo de hemicelulosa en la biomasa.

80. La composición de cualquiera de las reivindicaciones 69-77, donde el peso combinado de los polipéptidos que tienen actividad de celulasa, si presente, está entre 1 g a 100 g por 1 kilogramo de celulosa en la biomasa.

406

81 Un caldo de fermentación que comprende la composición de cualquiera de las reivindicaciones 1-80.

82. El caldo de fermentación de la reivindicación 81, que es el caldo de fermentación de un hongo filamentoso.

83. El caldo de fermentación de la reivindicación 82, donde el hongo filamentoso, es un *Trichoderma, Humicola, Fusarium, Aspergillus, Neurospora, Penicillium, Cefalosporium, Achlya, Podospora, Endothia, Mucor, Cochliobolus, Pyricularia,* Phanerochaete o *Chrysosporium* .

84. El caldo de fermentación de la reivindicación 82 donde el hongo filamentoso es un *Aspergillus awamori, Aspergillus fumigatus, Aspergillus foetidus, Aspergillus japonicus, Aspergillus nidulans, Aspergillus niger, Aspergillus oryzae, Chrysosporium lucknowense, Fusarium bactridioides, Fusarium cerealis, Fusarium crookwellense, Fusarium culmorum, Fusarium graminearum, Fusarium graminum, Fusarium heterosporum, Fusarium negundi, Fusarium oxysporum, Fusarium reticulatum, Fusarium roseum, Fusarium sambucinum, Fusarium sarcochroum, Fusarium sporotrichioides, Fusarium sulphureum, Fusarium torulosum, Fusarium trichothecioides, Fusarium venenatum, Bjerkandera adusta, Ceriporiopsis aneirina, Ceriporiopsis aneirina, Ceriporiopsis caregiea, Ceriporiopsis gilvescens, Ceriporiopsis pannocinta, Ceriporiopsis rivulosa, Ceriporiopsis subrufa, Ceriporiopsis subvermispora, Coprinus cinereus, Coriolus hirsutus, Humicola insolens, Humicola lanuginosa, Mucor miehei, Myceliophthora thermophila, Neurospora crassa, Neurospora intermedia, Penicillium purpurogenum, Penicillium canescens, Penicillium solitum, Penicillium funiculosum, Phanerochaete chrysosporium, Phlebia radiate, Pleurotus eryngii, Talaromyces flavus, Thielavia terrestris, Trametes villosa, Trametes versicolor, Trichoderma harzianum, Trichoderma koningii, Trichoderma longibrachiatum, Trichoderma reesei,* o *Trichoderma viride.*

85. El caldo de fermentación de la reivindicación 82, donde el hongo filamentoso es un Aspergillus , o un *Fusarium.*

86. El caldo de fermentación de la reivindicación 82, donde el hongo filamentoso es un Trichoderma reesei .

87. El caldo de fermentación de la reivindicación 82, donde el hongo filamentoso es un funiculosum de penicillium .

88. El caldo de fermentación de cualquiera de las reivindicaciones 81-87, que es un caldo sin células de fermentación.

89. Un método para convertir la biomasa en azúcares, que comprende entrar en contacto con la biomasa con la composición de cualquiera de las reivindicaciones 1-80, o con el caldo de fermentación de cualquiera de las reivindicaciones 81-88.

90. Un proceso de sacarificación que comprende el tratamiento de un material compuesto por hemicelulosa con la composición de cualquiera de las reivindicaciones 1-80 o con el caldo de fermentación de cualquiera de las reivindicaciones 81-89.

91. La proceso de la reivindicación 90, donde el material que comprende la hemicelulosa es mazorca, rastrojo del maíz, fibra del maíz, mijo, sorgo, papel, pulpa, o bagazo de la caña de azúcar.

92. El proceso de la reivindicación 90, donde el material que comprende la hemicelulosa es *Miscanthus.*

93. El proceso de cualquiera de las reivindicaciones 90-92, con producción de por lo menos 60% de xilosa del xilano de hemicelulosa del material que comprende la hemicelulosa.

94. El proceso de cualquiera de las reivindicaciones 90-93, además comprende la recuperación de los monosacáridos.

95. Un ácido nucléico aislado, sintético o recombinante que codifica un polipéptido que tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.: 28, o a la secuencia del aminoácido que corresponde a los residuos (i) 20-341, (ii) 21-350, (iii) 107-341 o (iv) 107-350 de SEQ ID NO.: 28, que tiene β actividad de xilosidasa.

96. Un ácido nucléico aislado, sintético o recombinante que codifica un polipéptido que tiene por lo menos identidad de la secuencia del 90% a la secuencia del aminoácido de SEQ ID NO.:4, o a la secuencia del aminoácido que corresponde a los residuos (i) 21-445, (ii) 21-301, (iii) 21-323, (iv) 21-444, (v) 302-444, (vi) 302-445, (vii) 324-444, o (viii) 324-445 de la SEQ ID NO.: 4, que tiene β- actividad del xilosidasa.

97. Un ácido nucléico aislado, sintético o recombinante que codifica un polipéptido que tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.: 6, o a la secuencia del aminoácido que corresponde a los

408

residuos (i) 19-530, (ii) 29-530, (iii) 19-300 o (iv) 29-300 de SEQ ID NO.:6, que tiene β actividad de xilosidasa.

98. Un ácido nucléico aislado, sintético o recombinante que codifica un polipéptido que tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.: 8, o a la secuencia del aminoácido que corresponde a los residuos (i) 20-439, (ii) 20-291, (iii) 145-291 o (iv) 145-439 de SEQ ID NO.: 8, que tiene β actividad de xilosidasa.

99. Un ácido nucléico aislado, sintético o recombinante que codifica un polipéptido que tiene por lo menos identidad de la secuencia del 90% a la secuencia del aminoácido de SEQ ID NO.:10, o a la secuencia del aminoácido que corresponde a los residuos (i)23-449 (ii) 23-302, (iii) 23-320, (iv) 23-448, (v) 303-448, (vi) 303-449, (vii) 321-448, u (viii) 321-449 de la SEQ ID NO.: 10, que tiene β- actividad del xilosidasa.

100. Un ácido nucléico aislado, sintético o recombinante que codifica un polipéptido que tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.: 12, o a la secuencia del aminoácido que corresponde a los residuos (i) 17-574, (ii) 27-574, (iii) 17-303 o (iv) 27-303 de SEQ ID NO.:12, que tiene actividad de xilosidasa y/o de L-arabinofuranosidasa.

101. Un ácido nucléico aislado, sintético o recombinante que codifica un polipéptido que tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.: 14, o a la secuencia del aminoácido que corresponde a los residuos (i) 21-676, (ii) 21-652, (iii) 469-652 o (iv) 469-676 de SEQ ID NO.:14, que tiene actividades de xilosidasa y/o L-arabinofuranosidasa.

102. Un ácido nucléico aislado, sintético o recombinante que codifica un polipéptido que tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.: 16, o a la secuencia del aminoácido que corresponde a los residuos (i) 19-340, (ii) 53-340, (iii) 19-383 o (iv) 53-383 de SEQ ID NO.:16, que tiene actividad de xilosidasa.

103. Un ácido nucléico aislado, sintético o recombinante que codifica un polipéptido que tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.: 18, o a la secuencia del aminoácido que corresponde a los residuos (i) 21-341, (ii) 107-341, (iii) 21-348 o (iv) 107-348 de SEQ ID NO.: 18, que tiene actividad de xilosidasa.

104. Un ácido nucléico aislado, sintético o recombinante que codifica un polipéptido que tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.:20, o a la secuencia del aminoácido que corresponde a los residuos (i) 15-558 o (ii) 15-295, de SEQ ID NO.:20, que tiene actividad de L-arabinofuranosidasa.

105. Un ácido nucléico aislado, sintético o recombinante que codifica un polipéptido que tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.: 22, o a la secuencia del aminoácido que corresponde a los residuos (i) 21-632, (ii) 461-632, (iii) 21-642 o (iv) 461-642 de SEQ ID NO.:22, que tiene actividad de L-arabinofuranosidasa.

106. Un ácido nucléico aislado, sintético o recombinante que codifica un polipéptido que tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.: 2, o a la secuencia del aminoácido que corresponde a los residuos (i) 24-766, (ii) 73-321, (iii) 73-394 (iv) 395-622, (v) 24-622 o (vi) 73-622 de SEQ ID NO.:2, que tiene β actividad de xilosidasa.

107. Un ácido nucléico aislado, sintético o recombinante que codifica un polipéptido que tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.: 32, o a la secuencia del aminoácido que corresponde a los residuos (i) 21-660, (ii) 21-645, (iii) 450-645 o (iv) 450-660 de SEQ ID NO.:32, que tiene actividad de L-arabinofuranosidasa.

108. Un polipéptido aislado, donde dicho polipéptido que tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.: 28, o a la secuencia del aminoácido que corresponde a los residuos (i) 20-341, (ii) 21-350, (iii) 107-341 o (iv) 107-350 de SEQ ID NO.: 28, que tiene β actividad de xilosidasa.

109. Un polipéptido aislado, donde dicho polipéptido que tiene por lo menos identidad de la secuencia del 90% a la secuencia del aminoácido de SEQ ID NO.:4, o a la secuencia del aminoácido que corresponde a los residuos (i) 21-445, (ii) 21-301, (iii) 21-323, (iv) 21-444, (v) 302-444, (vi) 302-445, (vii) 324-444, o (viii) 324-445 de la SEQ ID NO.: 4, que tiene β- actividad del xilosidasa.

110. Un polipéptido aislado, donde dicho polipéptido que tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.: 6, o a la

secuencia del aminoácido que corresponde a los residuos (i) 19-530, (ii) 29-530, (iii) 19-300 o (iv) 29-300 de SEQ ID NO.: 6, que tiene β actividad de xilosidasa.

111. Un polipéptido aislado, donde dicho polipéptido que tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.: 8, o a la secuencia del aminoácido que corresponde a los residuos (i) 20-439, (ii) 20-291, (iii) 145-291 o (iv) 145-439 de SEQ ID NO.: 8, que tiene β actividad de xilosidasa.

112. Un polipéptido aislado, donde dicho polipéptido que tiene por lo menos identidad de la secuencia del 90% a la secuencia del aminoácido de SEQ ID NO.:10, o a la secuencia del aminoácido que corresponde a los residuos (i) 23-449, (ii) 23-302, (iii) 23-320, (iv) 23-448, (v) 303-448, (vi) 303-449, (vii) 321-448, o (viii)321-449 de la SEQ ID NO.: 10, que tiene β- actividad del xilosidasa.

113. Un polipéptido aislado, donde dicho polipéptido que tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.: 12, o a la secuencia del aminoácido que corresponde a los residuos (i) 17-574, (ii) 27-574, (iii) 17-303 o (iv) 27-303 de SEQ ID NO.: 12, que tiene actividad de xilosidasa y/o de L-arabinofuranosidasa.

114. Un polipéptido aislado, donde dicho polipéptido que tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.: 14, o a la secuencia del aminoácido que corresponde a los residuos (i) 21-676, (ii) 21-652, (iii) 469-652 o (iv) 469-676 de SEQ ID NO.: 14, que tiene actividad de xilosidasa y/o L-arabinofuranosidasa.

115. Un polipéptido aislado, donde dicho polipéptido que tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.: 16, o a la secuencia del aminoácido que corresponde a los residuos (i) 19-340, (ii) 53-340, (iii) 19-383 o (iv) 53-383 de SEQ ID NO.:16, que tiene actividad de xilosidasa.

116. Un polipéptido aislado, donde dicho polipéptido que tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.:18, o a la secuencia del aminoácido que corresponde a los residuos (i) 21-341, (ii) 107-341, (iii) 21-348 o (iv) 107-348 de SEQ ID NO.:18, que tiene actividad de xilosidasa.

117. Un polipéptido aislado, donde dicho polipéptido tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.:20, o a la

41

secuencia del aminoácido que corresponde a los residuos (i) 15-558 o (ii) 15-295, de SEQ ID NO.:20, que tiene actividad de L-arabinofuranosidasa.

118. Un polipéptido aislado, donde dicho polipéptido tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.: 22, o con la secuencia del aminoácido que corresponde a los residuos (i) 21-632, (ii) 461-632, (iii) 21-642 o (iv) 461-642 de SEQ ID NO.:22, que tiene actividad de L-arabinofuranosidasa.

119. Un polipéptido aislado, donde dicho polipéptido tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.: 2, o con la secuencia del aminoácido que corresponde a los residuos (i) 24-766, (ii) 73-321, (iii) 73-394 (iv) 395-622, (v) 24-622 o (vi) 73-622 de SEQ ID NO.: 2, que tiene β actividad de xilosidasa.

120. Un polipéptido aislado, donde dicho polipéptido tiene por lo menos identidad de la secuencia del 90% con la secuencia del aminoácido del SEQ ID NO.: 32, o con la secuencia del aminoácido que corresponde a los residuos (i) 21-660, (ii) 21-645, (iii) 450-645 o (iv) 450-660 de SEQ ID NO.:32, que tiene actividad de L-arabinofuranosidasa.

121. Una célula huésped dirigida para expresar un polipéptido recombinante de cualquiera de las reivindicaciones 108-120.

122. La célula huésped de la reivindicación 121, que es una célula de un hongo filamentoso.

123. La célula huésped de la reivindicación 122, que es una célula de un *Trichoderma, Humicola, Fusarium, Aspergillus, neuroespora, penicillium, Cephalosporium, Achlya, Podospora, Endothia, Mucor, Cochliobolus, Pyricularia, Phanerochaete,* o *Chrysosporium.*

124. La célula huésped de la reivindicación 122 que es una célula de un *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*

42

adusta, Ceriporiopsis aneirina, Ceriporiopsis aneirina, Ceriporiopsis caregiea, Ceriporiopsis gilvescens, Ceriporiopsis pannocinta, Ceriporiopsis rivulosa, Ceriporiopsis subrufa, Ceriporiopsis subvermispora, Coprinus cinereus, Coriolus hirsutus, Humicola insolens, Humicola lanuginosa, Mucor miehei, Myceliophthora thermophila, Neurospora crassa, Neurospora intermedia, Penicillium purpurogenum, Penicillium canescens, Penicillium solitum, Penicillium funiculosum, Phanerochaete chrysosporium, Phlebia radiate, Pleurotus eryngii, Talaromyces flavus, Thielavia terrestris, Trametes villosa, Trametes versicolor, Trichoderma harzianum, Trichoderma koningii, Trichoderma longibrachiatum, Trichoderma reesei, o *Trichoderma viride.*

125. La célula huésped de la reivindicación 122 o 123, que es una célula de un *Trichoderma spp.,* un *Pencillium,* un Aspergillus , o un *Fusarium.*

126. La célula huésped de la reivindicación 125, donde:

(a) la Trichoderma spp. es Trichoderma reesei; o

(b) el *Pencillium* es Pencillium funiculosum *;* o

(c) la Aspergillus es Aspergillus oryzae o Aspergillus nidulans; o

(d) la Fusarium es Fusarium verticillioides o Fusarium oxysporum.

127. La célula huésped de la reivindicación 121, que es una célula de una bacteria.

128. La célula huésped de la reivindicación 127, donde la bacteria es un *Streptomyces,* un *Thermomonospora,* un bacilo , o una Celulomonas .

129. Un método de producir la composición de cualquiera de las reivindicaciones 1-80 o el caldo de fermentación de cualquier de las reivindicaciones 81-88.

130. Un proceso de sacarificación que comprende el tratamiento de un material de biomasa compuesto por hemicelulosa con la composición de cualquiera de las reivindicaciones 1-80 o con el caldo de fermentación de cualquiera de las reivindicaciones 81-89.

131. El proceso de la reivindicación 130, donde el material de la biomasa además comprende celulosa.

132. El proceso de la reivindicación 130 o 131, donde la cantidad de polipéptidos que tienen actividad del xilanasa es 1 g a 40 g por kilogramo de hemicelulosa en el material de la biomasa.

133. El proceso de la reivindicación 130 o 131, donde la cantidad de polipéptidos que tienen actividad de xilanasa es 0.5 g a 40 g por kilogramo de hemicelulosa en el material de la biomasa.

135. El proceso de cualquiera de las reivindicaciones 130 o 131, donde la cantidad de polipéptidos que tienen actividad de xilosidasa es 1 g a 50 g por kilogramo de hemicelulosa en el material de la biomasa.

136. El proceso de cualquiera de las reivindicaciones 130 o 131, donde la cantidad de polipéptidos que tienen actividad del xilosidasa es 0.5 g a 50 g por kilogramo de hemicelulosa en el material de la biomasa.

137. El proceso de cualquiera de las reivindicaciones 130 o 136, donde la cantidad de polipéptidos que tienen actividad de L-arabinofuranosidasa es 0.5 g a 20 g por kilogramo de hemicelulosa en el material de la biomasa.

138. El proceso de cualquiera de las reivindicaciones 130 o 136, donde la cantidad de polipéptidos que tienen actividad de L-arabinofuranosidasa es 0.2 g a 20 g por kilogramo de hemicelulosa en el material de la biomasa.

139. El proceso de cualquiera de las reivindicaciones 130 o 138, donde la cantidad de polipéptidos que tienen actividad de celulasa es 1 g a 100 g por kilogramo de celulosa en el material de la biomasa.

140. El proceso de cualquiera de las reivindicaciones 130-139, en donde la cantidad de polipéptidos que tienen β- la actividad de la glucosidasa es el 50% o menos del peso total de polipéptidos que tienen actividad del celulasa.

RESUMEN

La presente divulgación está dirigida generalmente a enzimas de hidrolasa de glucólisis, cuyas composiciones comprenden dichas enzimas y métodos de uso de las enzimas y de las composiciones, por ejemplo para la sacarificación o conversión de materiales celulósicos y hemicelulosicos en azúcares.

415

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

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 *Fusarium verticillioides*
28.	Amino acid	Protein sequence of Fv43D
29.	Nucleotide	Nucleotide sequence for Pf43B, a GH43 family enzyme from *Penicillium funiculosum*
30.	Amino acid	Protein sequence of Pf43B
31.	Nucleotide	Nucleotide sequence for Fv51A, a GH51 family enzyme *Fusarium verticilloides*
32.	Amino acid	Protein sequence of Fv51A
33.	Nucleotide	Nucleotide sequence for Cg51B, a GH51 family enzyme from *Chaetomium globosum*
34.	Amino acid	Protein sequence of Cg51B
35.	Nucleotide	Nucleotide sequence for Fv43C, a GH43 family enzyme from *Fusarium verticillioides*
36.	Amino acid	Fv43C protein sequence
37.	Nucleotide	Nucleotide sequence for Fv30A, a GH30 family enzyme from *Fusarium verticillioides*
38.	Amino acid	Fv30A protein sequence
39.	Nucleotide	Nucleotide sequence for Fv43F, a GH43 family enzyme from *Fusarium verticillioides*
40.	Amino acid	Fv43F protein sequence
41.	Nucleotide	Nucleotide sequence for Xyn3, a GH10 family xylanase from *Trichoderma reesei*
42.	Amino acid	Xyn3 protein sequence
43.	Amino acid	Protein sequence of Xyn2, a GH11 xylanase from *Trichoderma reesei*
44.	Amino acid	Protein sequence of Bxl1, a GH3 β-xylosidase from *Trichoderma reesei*
45.	Amino acid	Protein sequence of Bgl1, a GH3 β-glucosidase from *Trichoderma reesei*
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

417

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

QB

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

SUBSTITUTE SHEET (RULE 26)

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	25.642	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	0.057	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	0.187	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	0.985	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	69.85	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

SUBSTITUTE SHEET (RULE 26)

			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

SUBSTITUTE SHEET (RULE 26)

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

SUBSTITUTE SHEET (RULE 26)

SUBSTITUTE SHEET (RULE 26)

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	Ara
A	**5.6%**	**73.2%**	**69.4%**	5.8%	76.2%	71.8%	2.6%	84.8%	71.6%	2.1%	92.6%	92.3%
B	**3.2%**	**62.6%**	**42.6%**	4.8%	72.0%	64.6%	59.8%	83.1%	70.3%	2.0%	95.9%	93.6%
C	**1.8%**	**73.3%**	**68.3%**	14.3%	70.6%	67.0%	2.5%	98.5%	81.1%	1.8%	96.1%	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	180	47.5	14	49.3%	67.4%	57.1%
2	130	145	40	21	49.4%	78.3%	61.8%
3	90	180	47.5	14	53.5%	74.4%	62.5%
4	130	145	52	23	51.7%	78.0%	63.0%
5	90	180	47.5	14	53.5%	76.5%	63.4%
6	90	180	47.5	14	47.0%	67.5%	55.8%
7	90	180	47.5	14	51.0%	71.4%	59.8%
8	170	180	47.5	14	50.0%	64.4%	56.2%
9	90	180	45.5	28	60.9%	84.0%	70.8%
10	90	180	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

SUBSTITUTE SHEET (RULE 26)

SUBSTITUTE SHEET (RULE 26)

	g/L			% Conversion		
Sample Name pH / Temp	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*

atgctgctcaatcttcaggtcgctgccagcgctttgtcgctttctcttttaggtggattggctgaggctg
ctacgccatataccctttccggactgtaccaaaggacctttgagcaagaatggaatctgcgatacttcgtt
atctccagctaaaagagcggctgctctagttgctgctctgacgcccgaagagaaggtgggcaatctggtc
aggtaaaatatacccccccccataatcactattcggagattggagctgacttaacgcagcaatgcaactg
gtgcaccaagaatcggacttccaaggtacaactggtggaacgaagcccttcatggcctcgctggatctcc
aggtggtcgctttgccgacactcctccctacgacgcggccacatcatttcccatgcctcttctcatggcc
gctgctttcgacgatgatctgatccacgatatcggcaacgtcgtcggcaccgaagcgcgtgcgttcacta
acggcggttggcgcggagtcgacttctggacacccaacgtcaaccctttaaagatcctcgctggggtcg
tggctccgaaactccaggtgaagatgcccttcatgtcagccggtatgctcgctatatcgtcagggtctc
gaaggcgataaggagcaacgacgtattgttgctacctgcaagcactatgctggaaacgactttgaggact
ggggaggcttcacgcgtcacgactttgatgccaagattactcctcaggacttggctgagtactacgtcag
gcctttccaggagtgcacccgtgatgcaaaggttggttccatcatgtgcgcctacaatgccgtgaacggc
attcccgcatgcgcaaactcgtatctgcaggagacgatcctcagagggcactggaactggacgcgcgata
acaactggatcactagtgattgtggcgccatgcaggatatctggcagaatcacaagtatgtcaagaccaa
cgctgaaggtgcccaggtagcttttgagaacggcatggattctagctgcgagtatactactaccagcgat
gtctccgattcgtacaagcaaggcctcttgactgagaagctcatggatcgttcgttgaagcgccttttcg
aagggcttgttcatactggtttctttgacggtgccaaagcgcaatggaactcgctcagttttgcggatgt
caacaccaaggaagctcaggatcttgcactcagatctgctgtggagggtgctgttcttcttaagaatgac
ggcactttgcctctgaagctcaagaagaaggatagtgttgcaatgatcggattctgggccaacgatactt
ccaagctgcagggtggttacagtggacgtgctccgttcctccacagcccgctttatgcagctgagaagct
tggtcttgacaccaacgtggcttggggtccgacactgcagaacagctcatctcatgataactggaccacc
aatgctgttgctgcggcgaagaagtctgattacattctctactttggtggtcttgacgcctctgctgctg
gcgaggacagagatcgtgagaaccttgactggcctgagagccagctgacccttcttcagaagctctctag
tctcggcaagccactggttgttatccagcttggtgatcaagtcgatgacaccgctcttttgaagaacaag
aagattaacagtattctttgggtcaattaccctggtcaggatggcggcactgcagtcatggacctgctca
ctggacgaaagagtcctgctggccgactacccgtcacgcaatatcccagtaaatacactgagcagattgg
catgactgacatggacctcagacctaccaagtcgttgccagggagaacttatcgctggtactcaactcca
gtttcttccctacggctttggcctccactacaccaagttccaagccaagttcaagtccaacaagttgacgt
ttgacatccagaagcttctcaagggctgcagtgctcaatactccgatacttgcgcgctgccccccatcca
agttagtgtcaagaacaccggccgcattacctccgactttgtctctctggtctttatcaagagtgaagtt
ggacctaagccttaccctctcaagaccttgcggcttatggtcgcttgcatgatgtcgcgccttcatcga
cgaaggatatctcactggagtggacgttggataacattgcgccgacggggagagaatggtgatttggttgt
ttatcctgggacttacactctgttgctggatgagcctacgcaagccaagatccaggttacgctgactgga
aagaaggctattttggataagtggcctcaagaccccaagtctgcgtaa

FIG. 1A

SEQ ID NO:2

Protein sequence of Fv3A

mllnlqvaasalslsllgglaeaatpytlpdctkgplskngicdtslspakraaalvaaltpeekvgnlv
sn**ATGAPRIGLPRYNWWNEALHGLAGSPGGRFADTPPYDAATSFPMPLLMAAAFDDDLIHDIGNVVGTEA
RAFTNGGWRGVDFWTPNVNPFKDPRWGRGSETPGEDALHVSRYARYIVRGLEGDKEQRRIVATCKHYAGN
DFEDWGGFTRHDFDAKITPQDLAEYYVRPFQECTRDAKVGSIMCAYNAVNGIPACANSYLQETILRGHWN
WTRDNNWITSDCGAMQDIWQNHKYVKTNAEGAQVAFENGMD**ssceytttsdvsdsykqqlliteklmdrsl
krlfeglvhtgffdgakaqwnslsfadvntkeaqdlalrsaveg**AVLLKNDGTLPLKLKKKDSVAMIGFW
ANDTSKLQGGYSGRAPFLHSPLYAAEKLGLDTNVAWGPTLQNSSSHDNWTTNAVAAAKKSDYILYFGGLD
ASAAGEDRDRENLDWPESQLTLLQKLSSLGKPLVVIQLGDQVDDTALLKNKKINSILWVNYPGQDGGTAV
MDLLTGRKSPAGRLPVTQYPSKYTEQIGMTDMDLRPTKSLPGRTYRWYSTPVLPYGFGLHYTkfqakfks
nkltfdiqkllkgcsaqysdtcalppiqvsvkntgritsdfvslvfiksevgpkpyplktlaaygrlhdv
apsstkdislewtldniarrgengdlvvypgtytlllldeptqakiqvtltgkkaildkwpqdpksa

FIG. 1B

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

atgcttcagcgatttgcttatattttaccactggctctattgagtgttggagtgaaagccgacaaccctt
ttgtgcagagcatctacaccgctgatccggcaccgatggtatacaatgaccgcgtttatgtcttcatgga
ccatgacaacaccggagctacctactacaacatgacagactggcatctgttctcgtcagcagatatggcg
aattggcaagatcatggcattccaatgagcctggccaatttcacctgggccaacgcgaatgcgtgggccc
cgcaagtcatccctcgcaacggccaattctacttttatgctcctgtccgacacaacgatggttctatggc
tatcggtgtgggagtgagcagcaccatcacaggtccataccatgatgctatcggcaaaccgctagtagag
aacaacgagattgatcccaccgtgttcatcgacgatgacggtcaggcatacctgtactgggggaaatccag
acctgtggtacgtcaaattgaaccaagatatgatatcgtacagcgggagccctactcagattccactcac
cacggctggatttggtactcgaacgggcaatgctcaacggccgaccacttttgaagaagctccatgggta
tacaaacgcaacggcatctactatatcgcctatgcagccgattgttgttctgaggatattcgctactcca
cgggaaccagtgccactggtccgtggacttatcgaggcgtcatcatgccgacccaaggtagcagcttcac
caatcacgagggtattatcgacttccagaacaactcctacttttctatcacaacggcgctcttcccggc
ggaggcggctaccaacgatctgtatgtgtggagcaattcaaatacaatgcagatggaaccattccgacga
tcgaaatgaccaccgccggtccagctcaaattgggactctcaacccttacgtgcgacaggaagccgaaac
ggcggcatggtcttcaggcatcactacggaggtttgtagcgaaggcggaattgacgtcgggtttatcaac
aatggcgattacatcaaagttaaaggcgtagctttcggttcaggagcccattctttctcagcgcgggttg
cttctgcaaatagcggcggcactattgcaatacacctcggaagcacaactggtacgctcgtgggcacttg
tactgtccccagcactggcggttggcagacttggactaccgttacctgttctgtcagtggcgcatctggg
acccaggatgtgtattttgttttcggtggtagcggaacaggatacctgttcaactttgattattggcagt
tcgcataa

FIG. 2A

SEQ ID NO:4
Protein sequence of Pf43A

mlqrfayilplallsvgvkadnpfvqsiytadpapmvyndrvyvfmdhdntgatyynmtdwhlfssadma
nwqdhgipmslanftwananawapqvipmngqfyfyapvrhndgsmaigvgvsstitgpyhdaigkplve
nneidptvfidddgqaylywgnpdlwyvklnqdmisysgsptqiplttagfgtrtgnaqrpttfeeapwv
ykrngiyyiayaadccsedirystgtsatgpwtyrgvimptqgssftnhegiidfqnnsyffyhngalpg
gggyqrsvcveqfkynadgtiptiemttagpaqigtlnpyvrqEAETAAWSSGITTEVCSEGGIDVGFIN
NGDYIKVKGVAFGSGAHSFSARVASANSGGTIAIHLGSTTGTLVGTCTVPSTGGWQTWTTVTCSVSGASG
TQDVYFVFGGSGTGYLFNFDYWQFa

FIG. 2B

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

atgaaggtatactggctcgtggcgtgggccacttctttgacgccggcactggctggcttgattggacacc
gtcgcgccaccaccttcaacaatcctatcatctactcagactttccagataacgatgtattcctcggtcc
agataactactactacttctctgcttccaacttccacttcagcccaggagcacccgttttgaagtctaaa
gatctgctaaactgggatctcatcggccattcaattcccgcctgaactttggcgacggctatgatcttc
ctcctggctcacgttattaccgtggaggtacttgggcatcatccctcagatacagaaagagcaatggaca
gtggtactggatcggctgcatcaacttctggcagacctgggtatacactgcctcatcgccggaaggtcca
tggtacaacaagggaaacttcggtgataacaattgctactacgacaatggcatactgatcgatgacgatg
ataccatgtatgtcgtatacggttccggtgaggtcaaagtatctcaactatctcaggacggattcagcca
ggtcaaatctcaggtagttttcaagaacactgatattggggtccaagacttggagggtaaccgcatgtac
aagatcaacgggctctactatatcctaaacgatagcccaagtggcagtcagacctggatttggaagtcga
aatcaccctggggcccttatgagtctaaggtcctcgccgacaaagtcaccccgcctatctctggtggtaa
ctcgccgcatcagggtagtctcataaagactcccaatggtggctggtacttcatgtcattcacttgggcc
tatcctgccggccgtcttccggttcttgcaccgattacgtggggtagcgatggtttccccattcttgtca
agggtgctaatggcggatggggatcatcttacccaacacttcctggcacggatggtgtgacaaagaattg
gacaaggactgatacctttccgcggaacctcacttgctccgtcctgggagtggaaccataatccggacgtc
aactccttcactgtcaacaacggcctgactctccgcactgctagcattacgaaggatatttaccaggcga
ggaacacgctatctcaccgaactcatggtgatcatccaacaggaatagtgaagattgatttctctccgat
gaaggacggcgaccgggccgggctttcagcgtttcgagaccaaagtgcatacatcggtattcatcgagat
aacggaaagttcacaatcgctacgaagcatgggatgaatatggatgagtggaacggaacaacaacagacc
tgggacaaataaaagccacagctaatgtgccttctggaaggaccaagatctggctgagacttcaacttga
taccaacccagcaggaactggcaacactatcttttcttacagttgggatggagtcaagtatgaaacactg
ggtcccaacttcaaactgtacaatggttgggcattctttattgcttaccgattcggcatcttcaacttcg
ccgagacggctttaggaggctcgatcaaggttgagtctttcacagctgcatag

FIG. 3A

SEQ ID NO:6
Protein sequence of Fv43E

mkvywlvawatsltpalaglighrrattfnnpiiysdfpdndvflgpdnyyyfsasnfhfspgapvlksk
dllnwdlighsiprlnfgdgydlppgsryyrggtwasslryrksngqwywigcinfwqtwvytasspegp
wynkgnfgdnncyydngilidddtmyvvygsgevkvsqlsqdgfsqvksqvvfkntdigvqdlegnrmy
kinglyyilndspsgsqtwiwkskspwgpyeskvladkvtppisggnsphqgsliktpnggwyfmsftwa
ypagrlpvlapitwgsdgfpilvkganggwgssyptlpgtdgvtknwtrtdtfrgtslapswewnhnpdv
nsftvnngltlrtasitkdiyqarntlshrthgdhptgivkidfspmkdgdraglsafrdqsayigihrd
ngkftiatkhqmnmdewngtttdlgqikatanvpsgrtkiwlrlqldtnpagtgntifsyswdgvkyetl
gpnfklyngwaffiayrfgifnfaetalggsikvesftaa

FIG. 3B

430

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

atgcactacgctaccctcaccactttggtgctggctctgaccaccaacgtcgctgcacagcaaggcacag
caactgtcgacctctccaaaaatcatggaccggcgaaggcccttggttcaggcttcatatacggctggcc
tgacaacggaacaagcgtcgacacctccataccagatttcttggtaactgacatcaaattcaactcaaac
cgcggcggtggcgcccaaatcccatcactgggttgggccagaggtggctatgaaggatacctcggccgct
tcaactcaaccttatccaactatcgcaccacgcgcaagtataacgctgactttatcttgttgcctcatga
cctctgggtgcggatggcgggcagggttcaaactccccgtttcctggcgacaatggcaattggactgag
atggagttattctggaatcagcttgtgtctgacttgaaggctcataatatgctggaaggtcttgtgattg
atgtttggaatgagcctgatattgatatcttttgggatcgcccgtggtcgcagtttcttgagtattacaa
tcgcgcgaccaaactacttcggtgagtctactactgatccatacgtatttacagtgagctgactggtcga
attagaaaaacacttcccaaaactcttctcagtggcccagccatggcacattctcccattctgtccgatg
ataaatggcatacctggcttcaatcagtagcgggtaacaagacagtccctgatatttactcctggcatca
gattggcgcttgggaacgtgagccggacagcactatccccgactttaccaccttgcgggcgcaatatggc
gttcccgagaagccaattgacgtcaatgagtacgctgcacgcgatgagcaaaatccagccaactccgtct
actacctctctcaactagagcgtcataaccttagaggtcttcgcgcaaactggggtagcggatctgacct
ccacaactggatgggcaacttgatttacagcactaccggtacctcggaggggacttactaccctaatggt
gaatggcaggcttacaagtactatgcggccatggcagggcagagacttgtgaccaaagcatcgtcggact
tgaagtttgatgtctttgccactaagcaaggccgtaagattaagattatagccggcacgaggaccgttca
agcaaagtataacatcaaaatcagcggtttggaagtagcaggacttcctaagatgggtacggtaaaggtc
cggacttatcggttcgactgggctgggccgaatggaaaggttgacgggcctgttgatttgggggagaaga
agtatacttattcggccaatacggtgagcagcccctctacttga

FIG. 4A

SEQ ID NO:8
Protein sequence of Fv39A

mhyatlttlvlalttnvaaqggtatvdlsknhgpakalgsgfiygwpdngtsvdtsipdflvtdikfnsn
rgggaqipslgwarggyegylgrfnstlsnyrttrkynadfillphdlwgadgggqgsnspfpgdngnwte
melf**wnqlvsdlkahnmleglvidvwnepdidifwdrpwsqfleyynratkllrktlpktllsgpamahs**
pilsddkwhtwlqsvagnktvpdiyswhqigawerepdstipdfttlraqygvpekpidvneyaardeqn
pansvyylsqlerhnlrglranwgsgsdlhnwmgnliysttgtsegtyypngewqaykyyaamagqrlvt
kassdlkfdvfatkqgrkikiiagtrtvqakynikisglevaglpkmgtvkvrtyrfdwagpngkvdgpv
dlgekkytysantvsspst

FIG. 4B

431

SEQ ID NO:9
Nucleotide sequence for Fv43A, a GH43 family enzyme from *Fusarium verticillioides*

atgtggctgacctccccattgctgttcgccagcaccctcctgggcctcactggcgttgctctagcagaca
accccatcgtccaagacatctacaccgcagacccagcaccaatggtctacaatggccgcgtctacctctt
cacaggccatgacaacgacggctctaccgacttcaacatgacagactggcgtctcttctcgtcagcagac
atggtcaactggcagcaccatggtgtccccatgagcttaaagaccttcagctgggccaacagcagagcct
gggctggtcaagtcgttgcccgaaacggaaagttttacttctatgttcctgtccgtaatgccaagacggg
tggaatggctattggtgtcggtgttagtaccaacatccttgggccctacactgatgcccttggaaagcca
ttggtcgagaacaatgagatcgacccaactgtctacatcgacactgatggccaggcctatctctactggg
gcaaccctggattgtactacgtcaagctcaaccaagacatgctctcctacagtggtagcatcaacaaagt
atcgctcacaacagctggattcggcagccgcccgaacaacgcgcagcgtcctactactttcgaggaagga
ccgtggctgtacaagcgtggaaatctctactacatgatctacgcagccaactgctgttccgaggacattc
gctactcaactggacccagcgccactggaccttggacttaccgcggtgtcgtgatgaacaaggcgggtcg
aagcttcaccaaccatcctggcatcatcgactttgagaacaactcgtacttcttttaccacaatggcgct
cttgatggaggtagcggttatactcggtctgtggctgtcgagagcttcaagtatggttcggacggtctga
tccccgagatcaagatgactacgcaaggcccagcgcagctcaagtctctgaacccatatgtcaagcagga
ggccgagactatcgcctggtctgagggtatcgagactgaggtctgcagcgaaggtggtctcaacgttgct
ttcatcgacaatggtgactacatcaaggtcaagggagtcgactttggcagcaccggtgcaaagacgttca
gcgcccgtgttgcttccaacagcagcggaggcaagattgagcttcgacttggtagcaagaccggtaagtt
ggttggtacctgcacggtaacgactacgggaaactggcagacttataagactgtggattgccccgtcagt
ggtgctactggtacgagcgatctattctttgtcttcacgggctctgggtctggctctctgttcaacttca
actggtggcagtttagctaa

FIG. 5A

SEQ ID NO:10
Protein sequence of Fv43A

mwltspllfastllqltgvala**dnpivqdiytadpapmvyngrvylftghdndgstdfnmtdwrlfssad**
mvnwqhhgvpmslktfswansrawagqvvarngkfyfyvpvrnaktggmaigvgvstnilgpytdalgkp
lvenneidptvyidtdgqaylywgnpglyyvklnqdmlsysgsinkvslttagfgsrpnnaqrpttfeeg
pwlykrgnlyymiyaanccsedirystgpsatgpwtyrgvvmnkagrsftnhpgiidfennsyffyhnga
ldggsgytrsvavesfkygsdg*lipeikmttggpaqlksl*NPYVKQEAETIAWSEGIETEVCSEGGLNVA
FIDNGDYIKVKGVDFGSTGAKTFSARVASNSSGGKIELRLGSKTGKLVGTCTVTTTGNWQTYKTVDCPVS
GATGTSDLFFVFTGSGSGSLFNFNWWQFs

FIG. 5B

432

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

atgcgcttctcttggctattgtgccccttctagcgatgggaagtgctcttcctgaaacgaagacggatg
tttcgacatacaccaaccctgtccttccaggatggcactcggatccatcgtgtatccagaaagatggcct
ctttctctgcgtcacttcaacattcatctccttcccaggtcttcccgtctatgcctcaagggatctagtc
aactggcgtctcatcagccatgtctggaaccgcgagaaacagttgcctggcattagctggaagacggcag
gacagcaacagggaatgtatgcaccaaccattcgataccacaagggaacatactacgtcatctgcgaata
cctgggcgttggagatattattggtgtcatcttcaagaccaccaatccgtgggacgagagtagctggagt
gaccctgttaccttcaagccaaatcacatcgaccccgatctgttctgggatgatgacggaaaggtttatt
gtgctacccatggcatcactctgcaggagattgatttggaaactggagagcttagcccggagcttaatat
ctggaacggcacaggaggtgtatggcctgagggtccccatatctacaagcgcgacggttactactatctc
atgattgccgagggtggaactgccgaagaccacgctatcacaatcgctcgggcccgcaagatcaccggcc
cctatgaagcctacaataacaacccaatcttgaccaaccgcgggacatctgagtacttccagactgtcgg
tcacggtgatctgttccaagataccaagggcaactggtggggtctttgtcttgctactcgcatcacagca
cagggagtttcacccatgggccgtgaagctgttttgttcaatggcacatggaacaagggcgaatggccca
agttgcaaccagtacgaggtcgcatgcctggaaacctcctcccaaagccgacgcgaaacgttcccggaga
tgggcccttcaacgctgacccagacaactacaacttgaagaagactaagaagatccctcctcactttgtg
caccatagagtcccaagagacggtgccttctctttgtcttccaagggtctgcacatcgtgcctagtcgaa
acaacgttaccggtagtgtgttgccaggagatgagattgagctatcaggacagcgaggtctagctttcat
cggacgccgccaaactcacactctgttcaaatatagtgttgatatcgacttcaagcccaagtccgatgat
caggaagctggaatcaccgttttccgcacgcagttcgaccatatcgatcttggcattgttcgtcttccta
caaaccaaggcagcaacaagaaatctaagcttgccttccgattccgggccacaggagctcagaatgttcc
tgcaccgaaggtagtaccggtccccgatggctgggagaagggcgtaatcagtctacatatcgaggcagcc
aacgcgacgcactacaaccttggagcttcgagccacagaggcaagactctcgacatcgcgacagcatcag
caagtcttgtgagtggaggcacgggttcatttgttggtagtttgcttggaccttatgctacctgcaacgg
caaaggatctggagtggaatgtcccaagggaggtgatgtctatgtgacccaatggacttataagcccgtg
gcacaagagattgatcatggtgtttttgtgaaatcagaattgtag

FIG. 6A

SEQ ID NO:12
Protein sequence of Fv43B

mrfswllcpllamqsalpetktdvst**ytnpvlpgwhsdpsciqkdglflcvtstfisfpglpvyasrdlv**
nwrlishvwnrekqlpgiswktagqqqgmyaptiryhkgtyyviceylgvgdiigvifkttnpwdesswss
dpvtfkpnhidpdlfwdddgkvycathgitlqeidletgelspelniwngtggvwpegphiykrdgyyyl
miaeggtaedhaitiararkitgpyeaynnnpiltnrgtseyfqtvghgdlfqdtkgnwwglclatrita
qgvspmgreavlfngtwnkgewpklqpvrgrmpgnllpkptrnvpgdgpfnadpdnynlkktkkipphfv
hhrvprdgafslsskglhivpsrnnvtgsvlpgdeielsgqrglafigrrqthtlfkysvdidfkpksdd
qeagitvfrtqfdhidlgivrlptnqgsnkksklafrfratgaqnvpapkvvpvpdgwekgvislhieaa
nathynlgasshrgktldiatasaslvsggtgsfvgsllgpyatcngkgsgvecpkggdvyv
tqwtykpvaqeidhgvfvksel

FIG. 6B

SUBSTITUTE SHEET (RULE 26)

SEQ ID NO:13
Nucleotide sequence for Pa51A, a GH51 family enzyme from *Podospora anserina*

atgatccacctcaagccagccctcgcggcgttgttggcgctgtcgacgcaatgtgtggctattgatttgtttgtcaagtc
ttcgggggggaataagacgactgatatcatgtatggtcttatgcacgaggtatgtgttttgcgagatctccctttgtttt
ttgcgcactgctgacatggagactgcaaacaggatatcaacaactccggcgacggcggcatctacgccgagctaatctcc
aaccgcgcgttccaagggagtgagaagttcccctccaacctcgacaactggagccccgtcggtggcgctacccttaccct
tcagaagcttgccaagccccttcctctgcgttgccttactccgtcaatgttgccaaccccaaggagggcaagggcaagg
gcaaggacaccaagggggaagaaggttggcttggccaatgctgggttttggggtatggatgtcaagaggcagaagtacact
ggtagcttccacgttactggtgagtacaagggtgactttgaggttagcttgcgcagcgcgattaccggggagacctttgg
caagaaggtggtgaagggtgggagtaagaaggggaagtggaccgagaaggagtttgagttggtgcctttcaaggatgcgc
ccaacagcaacaacacctttgttgtgcagtgggatgccgaggtatgtgcttctttgatattggctgagatagaagttggg
ttgacatgatgtggtgcagggcgcaaaggacggatctttggatctcaacttgatcagcttgttccctccgacattcaagg
gaaggaagaatgggctgagaattgatcttgcgcagacgatggttgagctcaagccggtaagtcctctctagtcagaaaag
tagagcctttgttaacgcttgacagaccttcttgcgcttccccggtggcaacatgctcgagggtaacaccttggacactt
ggtggaagtggtacgagaccattggccctctgaaggatcgcccgggcatggctggtgtctgggagtaccagcaaacccct
ggcttgggtctggtcgagtacatggagtgggccgatgacatgaacttggagcccagtatgtgatcccattttctggagtg
acttctcttgctaacgtatccacagttgtcggtgtcttcgctggtcttgccctcgatggctcgttcgttcccgaatccga
gatgggatgggtcatccaacaggctctcgacgaaatcgagttcctcactggcgatgctaagaccaccaaatggggtgccg
tccgcgcgaagcttggtcaccccaagccttggaaggtcaagtgggttgagatcggtaacgaggattggcttgccggacgc
cctgctggcttcgagtcgtacatcaactaccgcttccccatgatgatgaaggccttcaacgaaaagtaccccgacatcaa
gatcatcgcctcgccctccatcttcgacaacatgacaatccccgcgggtgctgccggtgatcaccacccgtacctgactc
ccgatgagttcgttgagcgattcgccaagttcgataacttgagcaaggataacgtgacgctcatcggcgaggctgcgtcg
acgcatcctaacggtggtatcgcttgggagggagatctcatgcccttgccttggtggggcggcagtgttgctgaggctat
cttcttgatcagcactgagagaaacggtgacaagatcatcggtgctacttacgcgcctggtcttcgcagcttggaccgct
ggcaatggagcatgacctgggtgcagcatgccgccgaccccggcctcaccactcgctcgaccagttggtatgtctggaga
atcctcgcccaccacatcatccgtgagacgctcccggtcgatgccccggccggcaagcccaactttgaccctctgttcta
cgttgccggaaagagcgagagtggcaccggtatcttcaaggctgccgtctacaactcgactgaatcgatcccggtgtcgt
tgaagtttgatggtctcaacgagggagcggttgccaacttgacggtgcttactgggccggaggatccgtatggatacaac
gacccccttcactggtatcaatgttgtcaaggagaagaccaccttcatcaaggccggaaagggcggcaagttcaccttcac
cctgccgggcttgagtgttgctgtgttggagacggccgacgcggtcaagggtggcaagggaaagggcaagggcaagggaa
agggtaactga

FIG. 7A

434

SEQ ID NO:14
Protein sequence of Pa51A

mihlkpalaallalstqcvaidlfvkssggnkttdimyglmhedinnsgdqgiyaelisnrafqgsekfp
snldnwspvqgatltlqklakplssalpysvnvanpkegkgkgkdtkgkkvglanagfwgmdvkrqkytg
sfhvtgeykgdfevslrsaitgetfgkkvvkggskkgkwtekefelvpfkdapnsnntfvvqwdaegakd
gsldlnlislfpptfkgrknglridlaqtmvelkptflrfpggnmleqntldtwwkwyetigplkdrpgm
agvweyqqtlglglveymewaddmnlepivgvfaglaldgsfvpesemgwviqqaldeiefltgdakttk
wgavraklghpkpwkvkwveignedwlagrpagfesyinyrfpmwmmkafnekypdikiiaspsifdnmti
pagaagdhhpyltpdefverfakfdnlskdnvtligeaasthpnggia**wegdlmplpwwggsvaeaifli**
sterngdkiigatyapglrsldrwqwsmtwvqhaadpalttrstswyvwrilahhiiretlpvdapagkp
nfdplfyvagksesgtgifkaavynstesipvslkfdglnegavanltvltgpedpygyndpftginvvk
ekttfikagkggkftftlpglsvavletadavkggkgkgkgkgkgn

FIG. 7B

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

atgaagtccaagttgttattcccactcctctctttcgttggtcaaagtcttgccaccaacgacgactgtcctctcatcactagtag
atggactgcggatccttcggctcatgtctttaacgacaccttgtggctctacccgtctcatgacatcgatgctggatttgagaatg
atcctgatggaggccagtacgccatgagagattaccatgtctactctatcgacaagatctacggttccctgccggtcgatcacggt
acggccctgtcagtggaggatgtcccctgggcctctcgacagatgtgggctcctgacgctgcccacaagaacggcaaatactacct
atacttccctgccaaagacaaggatgatatcttcaqaatcqqcgttgctgtctcaccaaccccqqcqqaccattcgtccccgaca
agagttggatccctcacactttcagcatcgaccccgccagtttcgtcgatgatgatgacagagcctacttggcatggggtggtatc
atgggtggccagcttcaacgatggcaggataagaacaagtacaacgaatctggcactgagccaggaaacggcaccgctgccttgag
ccctcagattgccaagctgagcaaggacatgcacactctggcagagaagcctcgcgacatgctcattcttgaccccaagactggca
aqccqctcctttctqaqqatqaaqaccqacgcttcttccqaaqqaccctqqattcacaaqcqcaacaaqatttactacctcacctac
tctactggcacaacccactatcttgtctatgcgacttcaaagacccctatggtccttacacctaccagggcagaattctggagcc
agttgatggctggactactcactctagtatcgtcaagtaccagggtcagtggtggctattttatcacgatgccaagacatctggca
aggactatcttcgccaggtaaaggctaagaagatttggtacgatagcaaaggaaagatcttgacaaagaagccttga

FIG. 8A

435

SEQ ID NO:16
Protein sequence of Gz43A

mkskllfpllsfvqqslatnddcplitsrwtadpsahvfndtlwlypshdid**agfendpdggqyamrdyh**
vysidkiygslpvdhgtalsvedvpwasrqmwapdaahkngkyylyfpakdkddifrigvavsptpggpf
vpdkswiphtfsidpasfvdddraylawggimggqlqrwqdknkynesgtepgngtaalspqiaklskd
mhtlaekprdmlildpktgkpllsededrrffegpwihkrnkiyyltystgtthylvyatsktpygpyty
qgrilepvdgwtthssivkyqgqwwlfyhdaktsgkdylrqvkakkiwydskgkiltkkp

FIG. 8B

SEQ ID NO:17
Nucleotide sequence for Fo43A, a GH43 family enzyme from *Fusarium oxysporum*

atgcagctcaagtttctgtcttcagcattgctgttctctctgaccagcaaatgcgctgcgcaagacacta
atgacattcctccctgatcaccgacctctggtccgcagatccctcggctcatgttttcgaaggcaagct
ctgggtttacccatctcacgacatcgaagccaatgttgtcaacggcacaggaggcgctcaatacgccatg
agggattaccatacctactccatgaagagcatctatggtaaagatcccgttgtcgaccacggcgtcgctc
tctcagtcgatgacgttccctgggcgaagcagcaaatgtgggctcctgacgcagctcataagaacggcaa
atattatctgtacttcccgccaaggacaaggatgagatcttcagaattggagttgctgtctccaacaag
cccagcggtcctttcaaggccgacaagagctggatccctggcacgtacagtatcgatcctgctagctacg
tcgacactgataacgaggcctacctcatctggggcggtatctgggcggccagctccaagcctggcagga
taaaaagaactttaacgagtcgtggattggagacaaggctgctcctaacggcaccaatgccctatctcct
cagatcgccaagctaagcaaggacatgcacaagatcaccgaaacaccccgcgatctcgtcattctcgccc
ccgagacaggcaagcctcttcaggctgaggacaacaagcgacgattcttcgagggcccttggatccacaa
gcgcggcaagctttactacctcatgtactccaccggtgatacccacttccttgtctacgctacttccaag
aacatctacggtcctatacctaccggggcaagattcttgatcctgttgatgggtggactactcatggaa
gtattgttgagtataagggacagtggtggcttttctttgctgatgcgcatacgtctggtaaggattacct
tcgacaggtgaaggcgaggaagatctggtatgacaagaacggcaagatcttgcttcaccgtccttag

FIG. 9A

SEQ ID NO:18
Protein sequence of Fo43A

mqlkflssallfsltskcaaqdtndipplitdlwsadpsahvfegklwvypshdieanvvngtggaqyam
rdyhtysmksiygkdpvvdhgvalsvddvpwakqqm**wapdaahkngkyylyfpakdkdeifrigvavsnk**
psgpfkadkswipgtysidpasyvdtdneayliwggiwggqlqawqdkknfneswigdkaapngtnalsp
qiaklskdmhkitetprdlvilapetgkplqaednkrrffegpwihkrgklyylmystgdthflvyatsk
niygpytyrgkildpvdgwtthgsiveykgqwwlffadahtsgkdylrqvkarkiwydkngkillhrp

FIG. 9B

436

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

atggcagctccaagtttatcctaccccacaggtatccaatcgtataccaatcctctcttccctggttggc
actccgatcccagctgtgcctacgtagcggagcaagacacctttttctgcgtgacgtccactttcattgc
cttccccggtcttcctctttatgcaagccgagatctgcagaactggaaactggcaagcaatattttcaat
cggcccagccagatccctgatcttcgcgtcacggatggacagcagtcgggtatctatgcgcccactctgc
gctatcatgagggccagttctacttgatcgtttcgtacctgggcccgcagactaagggcttgctgttcac
ctcgtctgatccgtacgacgatgccgcgtggagcgatccgctcgaattcgcggtacatggcatcgacccg
gatatcttctgggatcacgacgggacggtctatgtcacgtccgccgaggaccagatgattaagcagtaca
cactcgatctgaagacgggggcgattggcccggttgactacctctggaacggcaccggaggagtctggcc
cgagggcccgcacatttacaagagagacggatactactacctcatgatcgcagagggaggtaccgagctc
ggccactcggagaccatggcgcgatctagaacccggacaggtccctgggagccataccgcacaatccgc
tcttgtcgaacaagggcacctcggagtacttccagactgtgggccatgcggacttgttccaggatgggaa
cggcaactggtgggccgtggcgttgagcacccgatcagggcctgcatggaagaactatcccatgggtcgg
gagacggtgctcgcccccgccgcttgggagaagggtgagtggcctgtcattcagcctgtgagaggccaaa
tgcaggggccgtttccaccaccaaataagcgagttcctcgcggcgagggcggatggatcaagcaacccga
caaagtggatttcaggcccggatcgaagataccggcgcacttccagtactggcgatatcccaagacagag
gattttaccgtctcccctcggggccacccgaatactcttcggctcacaccctcctttacaacctcaccg
gaactgcggacttcaagccggatgatggcctgtcgcttgttatgcgcaaacagaccgacaccttgttcac
gtacactgtggacgtgtcttttgaccccaaggttgccgatgaagaggcgggtgtgactgttttccttacc
cagcagcagcacatcgatcttggtattgtccttctccagacaaccgaggggctgtcgttgtccttccggt
tccgcgtggaaggccgcggtaactacgaaggtcctcttccagaagccaccgtgcctgttcccaaggaatg
gtgtggacagaccatccggcttgagattcaggccgtgagtgacaccgagtatgtctttgcggctgccccg
gctcggcaccctgcacagaggcaaatcatcagccgcgccaactcgttgattgtcagtggtgatacgggac
ggtttactggctcgcttgttggcgtgtatgccacgtcgaacgggggtgccggatccacgcccgcatatat
cagcagatggagatacgaaggacggggccagatgattgattttggtcgagtggtcccgagctactga

FIG. 10A

SEQ ID NO:20
Protein sequence of Af43A

maapslsyptgiqs**ytnplfpgwhsdpscayvaeqdtffcvtstfiafpglplyasrdlqnwklasnifn**
rpsqipdlrvtdgqqsgiyaptlryhegqfylivsylgpqtkgllftssdpyddaawsdplefavhgidp
difwdhdgtvyvtsaedqmikqytldlktgaigpvdylwngtggvwpegphiykrdgyyylmiaeggtel
ghsetmarsrtrtgpwepyphnpllsnkgtseyfqtvghadlfqdgngnwwavalstrsgpawknypmgr
etvlapaawekgewpviqpvrgqmqgpfpppnkrvprgeggwikqpdkvdfrpgskipahfqywrypkte
dftvsprghpntlrltpsfynltgtadfkpddglslvmrkqtdtlftytvdvsfdpkvadeeagvtvflt
qqqhidlgivllqtteglslsfrfrvegrgnyegplpeatvpvpkewcgqtirleiqavsdteyvfaaap
arhpaqrqiisranslivsgdtgrftgslvgvyatsnggagstpayisrwryegrgqmidfgrvvpsy

FIG. 10B

SUBSTITUTE SHEET (RULE 26)

SEQ ID NO:21
Nucleotide sequence for Pf51A, a GH51 family enzyme from *Penicillium funiculosum*

atgggaaagatgtggcattcgatcttggttgtgttgggcttattgtctgtcgggcatgccatcactatcaacgtgtcccaaagtgg
cggcaataagaccagtcctttgcaatatggtctgatgttcgaggtaatccttctcttataccacatataaaagttgcgtcatttct
aagacaagtcaaggacataaatcacggcggtgatggcggtctgtatgcagagcttgttcgaaaccgagcattccaaggtagcaccg
tctatccagcaaacctcgatggatacgactcggtcaatggagcaatcctagcgcttcagaatttgacaaaccctctatcaccctcc
atgcctagctctctcaacgtcgccaaggggtccaacaatggaagcatcggtttcgcaaatgaaggctggtgggggatagaagtcaa
gccgcaaagatacgcgggctcattctacgtccaggggactatcaaggagatttcgacatctctcttcagtcgaaattgacacaag
aagtcttcgcaacggcaaaagtcaggtcctcgggcaaacacgaggactgggttcaatacaagtacgagttggtgcccaaaaaggca
gcatcaaacaccaataacactctgaccattacttttgactcaaaggtatgttaaattttgggtttagttcgatgtctggcaattgt
cttacgagaaacgtagggattgaaagacggatccttgaacttcaacttgatcagcctatttcccccaacttacaacaatcggccca
atggcctaagaatcgacctggttgaagctatggctgaactagaggggtaagctcttacaaatcaactttatctttacgaagacta
atgtgaaaacttagaaatttctgcggtttccaggcggtagcgatgtggaaggtgtacaagctccttactggtataagtggaatgaa
acggtaggagatctcaaggaccgttatagtaggcccagtgcatggacgtacgaagaaagcaatggaattggcttgattgagtacat
gaattggtgtgatgacatgggcttgagccgagtgagtgtattccattcagcgtcaaatccagtgttctaatcatacacatcagtt
cttgccgtatgggatggacattacctttcgaacgaagtgatatcggaaaacgatttgcagccatatatcgacgacaccctcaacca
actggaattcctgatgggtgccccagatacgccatatggtagttggcgtgcgtctctgggctatccgaagccgtggacgattaact
acgtcgagattggaaacgaagacaatctatacgggggactagaaacatacatcgcctaccggtttcaggcatattacgacgctata
acagctaaatatccccatatgacggtcatggaatctttgacggagatgcctggtccggcggccgctgcaagcgattaccatcaata
ttctactcctgatgggtttgtttcccagttcaactactttgatcagatgccagtcactaatagaacactgaacggtatgaaaaccc
ccccttttttaaatatgcttttaatggtattaaccatctttcataggagagattgcaaccgtttatccaaataatcctagtaattc
ggtggcctggggaagcccattcccctttgtatcctggtggattgggtccgttgcagaagctgttttcctaattggtgaagagagga
attcgccaaagataatcggtgctagctacgtacggaattctacttttcgagattttaacattggataagaaggactaacctcaata
caggctccaatgttcagaaatatcaacaattggcagtggtctccaacactcatcgcttttgacgctgactcgtcgcgtacaagtcg
ttcaacaagctggcatgtgatcaaggtatgctaattttcctcctcattcaaacccgcagatgtgagctaactttccgaagcttctc
tcgacaaacaaaatcacgcaaaatttacccacgacttggagtggcggtgacataggtccattatactgggtagctggacgaaacga
caatacaggatcgaacatattcaaggccgctgtttacaacagcacctcagacgtccctgtcaccgttcaatttgcaggatgcaacg
caaagagcgcaaatttgaccatcttgtcatccgacgatccgaacgcatcgaactaccctgggggcccgaagttgtgaagactgag
atccagtctgtcactgcaaatgctcatggagcatttgagttcagtctcccgaacctaagtgtggctgttctcaaaacggagtaa

FIG. 11A

438

SEQ ID NO:22
Protein sequence of Pf51A

mgkmwhsilvvlgllsvghaitinvsqsggnktsplqyglmfedinhggdgglyaelvrnrafqgstvyp
anldgydsvnqailalqnltnplspsmpsslnvakgsnngsigfanegwwgievkpqryagsfyvqgdyq
gdfdislqskltqevfatakvrssgkhedwvqykyelvpkkaasntnntltitfdskglkdgslnfnlis
lfpptynnrpnglridlveamaelegkflrfpggsdvegvqapywykwnetvgdlkdrysrpsawtyees
ngiglieymnwcddmglepilavwdqhylsnevisendlqpyiddtlnqleflmgapdtpygswraslgy
pkpwtinyveignednlyggletyiayrfqayydaitakyphmtvmesltempgpaaaasdyhqystpdg
fvsqfnyfdqmpvtnrtlngeiatvypnnpsnsvawgspfplyp**wwigsvaeavfligeernspkiigas**
yapmfrninnwqwsptliafdadssrtsrstswhvikllstnkitqnlpttwsggdigplywvagrndnt
gsnifkaavynstsdvpvtvqfagcnaksanltilssddpnasnypggpevvkteiqsvtanahgafefs
lpnlsvavlkte

FIG. 11B

SEQ ID NO:23
Nucleotide sequence for AfuXyn2, a GH11 family enzyme from *Aspergillus fumigatus*

atggtttctttctcctacctgctgctggcgtgctccgccattggagctctggctgcccccgtcgaacccg
agaccacctcgttcaatgagactgctcttcatgagttcgctgagcgcgccggcaccccaagctccaccgg
ctggaacaacggctactactactccttctggactgatggcggcggcgacgtgacctacaccaatgqcgcc
ggtggctcgtactccgtcaactggaggaacgtgggcaactttgtcggtggaaagggctggaaccctggaa
gcgctaggtaccgagctttgtcaacgtcggatgtgcagacctgtggctgacagaagtagaaccatcaact
acggaggcagcttcaaccccagcggcaatggctacctggctgtctacggctggaccaccaaccccttgat
tgagtactacgttgttgagtcgtatggtacatacaaccccggcagcggcggtaccttcaggggcactgtc
aacaccgacggtggcacttacaacatctacacggccgttcgctacaatgctccctccatcgaaggcacca
agaccttcacccagtactggtctgtgcgcacctccaagcgtaccggcggcactgtcaccatggccaacca
cttcaacgcctggagcagactgggcatgaacctgggaactcacaactaccagattgtcgccactgagggt
taccagagcagcggatctgcttccatcactgtctactag

FIG. 12A

SEQ ID NO:24
Protein sequence of AfuXyn2

mvsfsvlllacsaigalaapvepettsfnetalhefaeragtpsst**gwnngyyysfwtdgggdvtytnga**
ggsysvnwrnvgnfvggkgwnpgsartinyggsfnpsgngylavygwttnplieyyvvesygtynpgsgg
tfrgtvntdggtyniytavrynapsiegtktftqywsvrtskrtggtvtmanhfnawsrlgmnlgthnyq
ivategyqssgsasitvy

FIG. 12B

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

atgatctccatttcctcgctcagctttggactcgccgctatcgccggcgcatatgctcttccgagtgaca
aatccgtcagcttagcggaacgtcagacgatcacgaccagccagacaggcacaaacaatggctactacta
ttccttctggaccaacggtgccggatcagtgcaatatacaaatggtgctggtggcgaatatagtgtgacg
tgggcgaaccagaacggtggtgactttacctgtgggaagggctggaatccagggagtgaccagtaggcaa
cgcccgagaactatagaagaggacgcaaagaaagcactaaactctctactagtgacattaccttctctgg
cagcttcaatccttccggaaatgcttacctgtccgtgtatggatggactaccaaccccctagtcgaatac
tacatcctcgagaactatggcagttacaatcctggctcgggcatgacgcacaagggcaccgtcaccagcg
atggatccacctacgacatctatgagcaccaacaggtcaaccagccttcgatcgtcggcacggccaccttt
caaccaatactggtccatccgccaaaacaagcgatccagcggcacagtcaccaccgcgaatcacttcaag
gcctgggctagtctggggatgaacctgggtacccataactatcagattgtttccactgagggatatgaga
gcagcggtacctcgaccatcactgtctcgtctggtggttcttcttctggtggaagtggtggcagctcgtc
tactacttcctcaggcagctcccctactggtggctccggcagtgtaagtcttcttccatatggttgtggc
tttatgtgtattctgactgtgatagtgctctgctttgtggggccagtgcggtggaattggctggtctggt
cctacttgctgctcttcgggcacttgccaggtttcgaactcgtactactcccagtgcttgtagtaccttc
ttgcagggttatatccaagtga

FIG. 13A

SEQ ID NO:26
Protein sequence of AfuXyn5

MISISSLSFGLAAIAGAYALPSDKSVSLAERQTITTSQT**GTNNGYYYSFWTNGAGSVQYTNGAGGEYSVT**
WANQNGGDFTCGKGWNPGSDHDITFSGSFNPSGNAYLSVYGWTTNPLVEYYILENYGSYNPGSGMTHKGT
VTSDGSTYDIYEHQQVNQPSIVGTATFNQYWSIRQNKRSSGTVTTANHFKAWASLGMNLGTHNYQIVSTE
GYESSGTSTITV*SSGGSSSGGSGGSSSTTSSGSSPTGGSGSCSA***LWGQCGGIGWSGPTCCSSGTCQVSNS**
YYSQCL

FIG. 13B

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

atgcagctcaagtttctgtcttcagcattgttgctgtctttgaccggcaattgcgctgcgcaagacacta
atgatatccctcctctgatcaccgacctctggtctgcggatccctcggctcatgttttcgagggcaaact
ctgggtttacccatctcacgacatcgaagccaatgtcgtcaacggcaccggaggcgctcagtacgccatg
agagattatcacacctattccatgaagaccatctatggaaaagatcccgttatcgaccatggcgtcgctc
tgtcagtcgatgatgtcccatgggccaagcagcaaatgtgggctcctgacgcagcttacaagaacggcaa
atattatctctacttccccgccaaggataaagatgagatcttcagaattggagttgctgtctccaacaag
cccagcggtcctttcaaggccgacaagagctggatccccggtacttacagtatcgatcctgctagctatg
tcgacactaatggcgaggcatacctcatctggggcggtatctggggcggccagcttcaggcctggcagga
tcacaagacctttaatgagtcgtggctcggcgacaaagctgctcccaacggcaccaacgccctatctcct
cagatcgccaagctaagcaaggacatgcacaagatcaccgagacaccccgcgatctcgtcatcctggccc
ccgagacaggcaagcccctcaagcagaggacaataagcgacgatttttcgagggcccctgggttcacaa
gcgcggcaagctgtactacctcatgtactctaccggcgacacgcacttcctcgtctacgcgacttccaag
aacatctacggtccttatacctatcagggcaagattctcgaccctgttgatgggtggactacgcatggaa
gtattgttgagtacaagggacagtggtggttgttctttgcggatgcgcatacttctggaaaggattatct
gagacaggttaaggcgaggaagatctggtatgacaaggatggcaagattttgcttactcgtcctaagatt
tag

FIG. 14A

SEQ ID NO:28
Protein sequence of Fv43D

mqlkflssalllsltqncaaqdtndipplitdlwsadpsahvfegklwvypshdieanvvngtggaqyam
rdyhtysmktiygkdpvidhgvalsvddvpwakqqm**wapdaaykngkyylyfpakdkdeifrigvavsnk**
psgpfkadkswipgtysidpasyvdtngeayliwggiwggqlqawqdhktfneswlgdkaapngtnalsp
qiaklskdmhkitetprdlvilapetgkplqaednkrrffegpwvhkrgklyylmystgdthflvyatsk
niygpytyqgkildpvdgwtthgsiveykgqwwlffadahtsgkdylrqvkarkiwydkdgkilltrpki

FIG. 14B

SEQ ID NO:29
Nucleotide sequence for Pf43B, GH43 family enzyme from *Penicillium funiculosum*

atgagtcgcagcatccttccgtacgcctctgttttcgccctcctgggcggggctatcgccgaaccgtttt
tggttctcaatagcgattttcccgatcccagtctcatagagacatccagcggatactatgcattcggtac
caccggaaacggagtcaatgcgcaggttgcttcttcaccagactttaatacctggactttgctttccggc
acagatgccctcccgggaccatttccgtcatgggtagcttcgtctccacaaatctgggcgccagatgttt
tggttaaggtatgttcttatggaataacagttttaggagtaggtcagccaggatattgacaaaattataa
taggccgatggtacctatgtcatgtacttttcggcatctgctgcgagtgactcgggcaaacactgcgttg
gtgccgcaactgcgacctcaccggaaggaccttacaccccggtcgatagcgctgttgcctgtccattaga
ccagggaggagctattgatgccaatggatttattgacaccgacggcactatatacgttgtatacaaaatt
gatggaaacagtctagacggtgatggaaccacacatcctacccccatcatgcttcaacaaatggaggcag
acggaacaaccccaaccggcagcccaatccaactcattgaccgatccgacctcgacggacctttgatcga
ggctcctagtttgctcctctccaatggaatctactacctcagtttctcttccaactactacaacactaat
tactacgacacttcatacgcctatgcctcgtcgattactggtccttggaccaaacaatctgcgccttatg
caccttgttggttactggaaccgagactagcaatgacggcgcattgagcgcccctggtggtgccgatttt
ctccgtcgatggcaccaagatgttgttccacgcaaacctcaatggacaagatatctcgggcggacgcgcc
ttatttgctgcgtcaattactgaggccagcgatgtggttacattgcagtag

FIG. 15A

SEQ ID NO:30
Protein sequence of Pf43B

msrsilpyasvfallqgaiaepfl**vlnsdfpdpslietssgyyafgttgngvnaqvasspdfntwtllsg**
tdalpgpfpswvasspqiwapdvlvkadgtyvmyfsasaasdsgkhcvgaatatspegpytpvdsavacp
ldqggaidangfidtdgtiyvvykidgnsldgdgtthptpimlqqmeadgttptgspiqlidrsdldgpl
ieapslllsngiyylsfssnyyntnyydtsyayassitgpwtkqsapyapllvtgtetsndgalsapgga
dfsvdgtkmlfhanlngqdisggralfaasiteasdvvtlq

FIG. 15B

442

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

atggttcgcttcagttcaatcctagcggctgcggcttgcttcgtggctgttgagtcagtcaacatcaaggtcgacag
caagggcggaaacgctactagcggtcaccaatatggcttccttcacgaggttggtattgacacaccactggcgatga
ttgggatgctaacttggagctaggatatcaacaattccggtgatggtggcatctacgctgagctcatccgcaatcgt
gctttccagtacagcaagaaatacccctgtttctctatctggctggagacccatcaacgatgctaagctctccctcaa
ccgtctcgacactcctctctccgacgctctccccgtttccatgaacgtgaagcctggaaagggcaaggccaaggaga
ttggtttcctcaacgagggttactggggaatggatgtcaagaagcaaaagtacactggctctttctgggttaagggc
gcttacaagggccactttacagcttctttgcgatctaaccttaccgacgatgtctttggcagcgtcaaggtcaagtc
caaggccaacaagaagcagtgggttgagcatgagtttgtgcttactcctaacaagaatgcccctaacagcaacaaca
cttttgctatcacctacgatcccaaggtgagtaacaatcaaaactgggacgtgatgtatactgacaatttgtagggc
gctgatggagctcttgacttcaacctcattagcttgttccctcccacctacaaggccgcaagaacggtcttcgagt
tgatcttgccgaggctctcgaaggtctccacccccgtaaggtttaccgtctcacgtgtatcgtgaacagtcgctgact
tgtagaaaagagcctgctgcgcttccccggtggtaacatgctcgagggcaacaccaacaagacctggtgggactgga
aggataccctcggacctctccgcaaccgtcctggtttcgagggtgtctggaactaccagcagacccatggtcttgga
atcttggagtacctccagtgggctgaggacatgaaccttgaaatcagtaggttctataaaattcagtgacggttatg
tgcatgctaacagatttcagttgtcggtgtctacgctggcctctccctcgacggctccgtcaccccccaaggaccaac
tccagcccctcatcgacgacgcgctcgacgagatcgaattcatccgaggtcccgtcacttcaaagtggggaaagaag
cgcgctgagctcggccaccccaagcctttcagactctcctacgttgaagtcggaaacgaggactggctcgctggtta
tcccactggctggaactcttacaaggagtaccgcttccccatgttcctcgaggctatcaagaaagctcaccccgatc
tcaccgtcatctcctctggtgcttctattgaccccgttggtaagaaggatgctggtttcgatattcctgctcctgga
atcggtgactaccaccctaccgcgagcctgatgttcttgttgaggagttcaacctgtttgataacaataagtatgg
tcacatcattggtgaggttgcttctacccaccccaacggtggaactggctggagtggtaaccttatgccttaccct
ggtggatctctggtgttggcgagccgtcgctctctgcggttatgagcgcaacgccgatcgtattcccggaacattc
tacgctcctatcctcaagaacgagaaccgttggcagtgggctatcaccatgatccaattcgccgccgactccgccat
gaccacccgctccaccagctggtatgtctggtcactcttcgcaggccaccccatgacccatactctccccaccaccg
ccgacttcgaccccctctactacgtcgctggtaagaacgaggacaagggaactcttatctggaagggtgctgcgtat
aacaccaccaagggtgctgacgttcccgtgtctctgtccttcaagggtgtcaagcccggtgctcaagctgagcttac
tcttctgaccaacaaggagaagcgatccttttgcgttcaatgatcctcacaagggcaacaatgttgttgatactaaga
agactgttctcaaggccgatggaaagggtgctttcaacttcaagcttcctaacctgagcgtcgctgttcttgagacc
ctcaagaagggaaagccttactctagctag

FIG. 16A

443

SEQ ID NO:32
Protein sequence of Fv51A

mvrfssilaaaacfvavesvnikvdskggnatsghqygflhedinnsgdggiyaelirnrafqyskkypv
slsgwrpindaklslnrldtplsdalpvsmnvkpgkgkakeigflnegywgmdvkkqkytgsfwvkgayk
ghftaslrsnltddvfgsvkvkskankkqwvehefvltpnknapnsnntfaitydpkgadgaldfnlisl
fpptykgrknglrvdlaealeglhpsllrfpggnmlegntnktwwdwkdtlgplrnrpgfegvwnyqqth
glgileylqwaedmnleiivgvyaglsldgsvtpkdqlqpliddaldeiefirgpvtskwgkkraelghp
kpfrlsyvevgnedwlagyptgwnsykeyrfpmfleaikkahpdltvissgasidpvgkkdagfdipapg
igdyhpyrepdvlveefnlfdnnkyghii**gevasthpnggtgwsgnlmpypwwisgvgeavalcgyerna**
dripgtfyapilknenrwqwaitmiqfaadsamttrstswyvwslfaghpmthtlpttadfdplyyvagk
nedkgtliwkgaayntttkgadvpvslsfkgvkpgaqaeltlltnkekdpfafndphkgnnvvdtkktvlk
adgkgafnfklpnlsvavletlkkgkpyss

FIG. 16B

SEQ ID NO:33
Nucleotide sequence for Cg51B, a GH51 family enzyme from *Chaetomium globosum*

atggcgcccctttcgcttcgggccctctcgctgctcgcgctcacaggagccgcagccgcggtgaccctat
cggtcgcgaactctggcggtaatgatacgtctccgtacatgtatggcatcatgttcgaggacatcaatca
gagcggtgacggcgggctgtaagttctgtcgcggcttcccctgacaagcttgcatgatgcttaactaaag
tccttaggtacgccgagctgattcgcaaccgagccttccataatagctcccctccaggcctggaccgccgt
ggggacagcactctcgaggtcgtaacctctgcaccgttatcggatgccctgcctcgctcggtcaaggtc
acgagtggaaagggcaaggcgggcttgaagaatgccggctactggggaatggacgtccagaagaccgaca
agtatagcggcagcttctactcgtacggcgcctacgacggaaagtttaccctctctctggtgtcggacat
cacaaatgagaccctggccaccaccaagatcaagtccaggtcggtggagcatgcctggaccgagcacaag
ttcgagcttctcccgaccaagagcgcggcgaacagcaacaacagcttcgtgctggagttccgcccctgcc
accagacggagctccagttcaacctcatcagcttgttcccgccgacgtataagaacaggcccaacggcat
gcgccgagagctcatggagaagctcgcagacctcaagcccagtttccttcggattccaggaggcaacaac
ctgtaagtgcttccggcgaaactagcagtagttgcctgagagacactaatctcagcgaacaacagcgagg
gcaactatgctggcaactactggaactggtcaagcacacttggcccgctgaccgaccggcccggtcgtga
cggcgtgtggacgtacgccaacacggacggcatcgggctggtcgagtacatgcactgggccgaggacctc
gacgtggaggttgtgctggcggtcgccgcaggcctgtacctgaacggcgatgtggtcccggaggaggagc
tgcacgtcttcgtggaggatgcgctgaacgagctcgagttcctcatgggcgacgtctcgacccccttgggg
cgcgcgccgccgctaagctcggctaccccaagccgtggaacatcaagttcgtcgaggtcggcaacgaggac
aacctgtgggcggcctcgactcgtacaagagctaccggctgaagactttctacgacgccatcaaggcga
agtaccccgacatctccatcttttcgtcgaccgacgagtttgtgtacaaggagtcgggccaggactacca
caagtacacccggccggactactccgtgtcccagttcgacctgtttgacaactgggccgacggccacccc
atcatcatcggagagtgagtgaacggcgacccccacctcccctaacgcgggatcgcgagctgatagatc
accccaggtatgcgaccatccagaacaacacgggcaagctcgaggacacggactgggacgcgcccaagaa
caagtggtccaactggatcggctccgtcgccgaggccgtcttcatcctcggagccgagcgcaacggcgac
cgggtctgggcaccacctttgcgccgatcctccagaacctcaacagctaccaatgggctgtaagtacat
acatacataccgcacccccaaccccaaccccccccaaagcgcacctccacccacccacccaaacacaccac
aactacctagctaacccgccacacaaacaaacagcccgacctaatctccttcaccgccaacccggccgac
accacgcccagcgtctcgtacccgatcatccagctgctcgcctcgcaccgcatcacgcacaccctccccg
tcagcagcgccgacgccttcggcccggcctactgggtggccggtcgcggcgccgacgacggctcgtacat
cctcaaggcggccgtgtacaacagcacgggggggtgcggatgtaccggtgagggtgcagtttgaggcgggg
ggtggtggtggtggtggtggtggtggtggtggtggtggtgatgggaaggggaagggtaaagggaaggg
gaggggagggtggtgagggtgtgaagaagggtgaccgcgcgcagttgaccgtgttgacggcgccggaggg
gccctgggcgcataatacgccggagaataaggggcggtcaagacgacagtgacgacgttgaaggccggg
agggtgggggtgtttgagtttagtctgccggatttgtcggtggcggtgttggtggtggaggggagaagt
ga

FIG. 17A

SEQ ID NO:34
Protein sequence of Cg51B

maplslralsllaltgaaaavtlsvansggndtspymygimfedinqsgdgglyaelirnrafhnsslqa
wtavgdstlevvtsaplsdalprsvkvtsgkgkaglknagywgmdvqktdkysgsfysygaydgkftlsl
vsditnetlattkiksrsvehawtehkfellptksaansnnsfvlefrpchqtelqfnlislfpptyknr
pngmrrelmekladlkpsflripggnnlegnyagnywnwsstlgpltdrpgrdgvwtyantdgiglveym
hwaedldvevvlavaaglylngdvvpeeelhvfvedalneleflmgdvstpwgarraklgypkpwnikfv
evgnednlwggldsyksyrlktfydaikakypdisifsstdefvykesgqdyhkytrpdysvsqfdlfdn
wadghpiii**geyatiqnntgkledtdwdapknkwsnwigsvaeavfilgaerngdrvwgttfapilqnln**
syqwapdlisftanpadttpsvsypiiqllashrithtlpvssadafgpaywvagrgaddgsyilkaavy
nstggadvpvrvqfeaggggggggggggggdgkgkgkgkggeggegvkkgdraqltvltapegpwahnt
penkgavkttvttlkagrggvfefslpdlsvavlvvegek

FIG. 17B

SEQ ID NO:35
Nucleotide sequence for Fv43C, a GH43 family enzyme from *Fusarium verticillioides*

atgcgtcttctatcgtttcccagccatctcctcgtggccttcctaaccctcaaagaggcttcatccctcg
ccctcagcaaacgggatagccctgtcctccccggcctctgggcggaccccaacatcgccatcgtcgacaa
gacatactacatcttccctaccaccgacggtttcgaaggctggggcggcaacgtcttctactggtggaaa
tcaaaagatctcgtatcatggacaaagagcgacaagccattccttactctcaatggtacgaatggcaacg
ttccctgggctacaggtaatgcctgggctcctgctttcgctgctcgcggaggcaagtattacttctacca
tagtgggaataatccctctgtgagtgatgggcataagagtattggtgcggcggtggctgatcatcctgag
gggccgtggaaggcacaggataagccgatgatcaagggaacttctgatgaggagattgtcagcaaccagg
ctatcgatcccgctgcctttgaagaccctgagactggaaagtggtatatctactggggaaacggtgtccc
cattgtcgcagagctcaacgacgacatggtctctctcaaagcaggctggcacaaaatcacaggtcttcag
aatttccgcgagggtcttttcgtcaactatcgcgatggaacatatcatctgacatactctatcgacgata
cgggctcagagaactatcgcgttgggtacgctacggcggataaccccattggaccttggacatatcgtgg
tgttcttctggagaaggacgaatcgaagggcattcttgctacgggacataactccatcatcaacattcct
ggaacggatgagtggtatatcgcgtatcatcgcttccatattcccgatggaaatgggtataatagggaga
ctacgattgatagggtacccatcgacaaggatacgggtttgtttggaaaggttacgccgactttgcagag
tgttgatcctaggcctttgtag

FIG. 18A

SEQ ID NO:36
Protein sequence for Fv43C

mrllsfpshllvafltlkeasslalsk**rdspvlpglwadpniaivdktyyifpttdgfegwggnvfywwk**
skdlvswtksdkpfltlngtngnvpwatgnawapafaarggkyyfyhsgnnpsvsdghksigaavadhpe
gpwkaqdkpmikgtsdeeivsnqaidpaafedpetgkwyiywgngvpivaelnddmvslkagwhkitglq
nfreglfvnyrdgtyhltysiddtgsenyrvgyatadnpigpwtyrgvllekdeskgilatghnsiinip
gtdewyiayhrfhipdgngynrettidrvpidkdtglfgkvtptlqsvdprpl

FIG. 18B

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

atgctcttctcgctcgttcttcctaccttgcctttcaagccagcctggcgctcggcgatacatccgtta
ctgtcgacaccagccagaaactccaggtcatcgatggctttggtgtctcagaagcctacggccacgccaa
acaattccaaaacctcggtcctggaccacagaaagagggcctcgatcttctcttcaacactacaaccggc
gcaggcttatccatcatccgaaacaagatcggctggacgcctccaactccatcaccagcaccaacaccg
acaacccagataagcaggctgtttaccattttgacggcgatgatgatggtcaggtatggtttagcaaaca
ggccatgagctatggtgtagatactatctacgctaatgcttggtctgcgcctgtatacatgaagtcagcc
cagagtatgggccgtctctgcggtacacctggtgtgtcgtgctcctctggagattggagacatcgttacg
ttgagatgatagctgagtacctctcctactacaagcaggctggcatcccagtgtcgcacgttggattcct
caatgagggtgacggctcggactttatgctctcaactgccgaacaggctgcagatgtcattcctcttcta
cacagcgctttgcagtccaagggccttggcgatatcaagatgacgtgctgtgataacatcggttggaagt
cacagatggactataccgccaagctggctgagcttgaggtggagaagtatctatctgtcatcacatccca
cgagtactccagcagccccaaccagcctatgaacactacattgccaacctggatgtccgagggagctgcc
aatgaccaggcatttgccacagcgtggtacgtcaacggcggttccaacgaaggtttcacatgggcagtca
agatcgcacaaggcatcgtcaatgccgacctctcagcgtatatctactgggagggcgttgagaccaacaa
caagggtctctatctcacgtcatcgacacggacggtaccaagtttaccatatcctcgattctctgggcc
attgctcactggtcgcgccatattcgccctggtgcgcatagactttcgacttcaggtgttgtgcaagata
cgattgttggtgcgtttgagaacgttgatggcagtgtcgtcatggtgctcaccaactctggcactgctgc
tcagactgtggacctgggtgtttcgggaagtagcttctcaacagctcaggctttcacttcggatgctgag
gcgcagatggtcgataccaaggtgactctgtccgacggtcgtgtcaaggttacggtcccggtgcacggtg
tcgtcactgtgaagctcacaacagcaaaaagctccaaaccggtctcaactgctgtttctgcgcaatctgc
ccccactccaactagtgttaagcacaccttgactcaccagaagacttcttcaacaacactctcgaccgcc
aaggccccaacctccactcagactacctctgtagttgagtcagccaaggcggtgaaataccctgtccccc
ctgtagcatccaagggatcctcgaagagtgctcccaagaagggtaccaagaagaccactacgaagaaggg
ctcccaccaatcgcacaaggcgcatagtgctactcatcgtcgatgccgccatggaagttaccgtcgtggc
cactgcaccaactaa

FIG. 19A

SEQ ID NO:38
Protein sequence of Fv30A

mlfslvlptlafqaslalgdtsvtvdtsqklqvidgfgvseayghakqfqnlgpgpqkegldllfntttg
aglsiirnkigcdasnsitstntdnpdkqavyhfdgdddgqsaqsmgrlcgtpgvscssgdwrhryvemi
aeylsyykqagipvshvgflnegdgsdfmlstaeqaadvipllhsalqskglgdikmtccdnigwksqmd
ytaklaelevekylsvitsheyssspnqpmnttlptwmsegaandqafatawyvnggsnegftwavkiaq
givnadlsayiywegvetnnkgslshvidtdgtkftissilwaiahwsrhirpgahrlstsgvvqdtivg
afenvdgsvvmvltnsgtaaqtvdlgvsgssfstaqaftsdaeaqmvdtkvtlsdgrvkvtvpvhgvvtv
klttaksskpvstavsaqsaptptsvkhtlthqktssttlstakaptstqttsvvesakavkypvppvas
kgssksapkkgtkktttkkgshqshkahsathrrcrhgsyrrghctn

FIG. 19B

SEQ ID NO:39
Nucleotide sequence for Fv43F, a GH43 family enzyme from *Fusarium verticillioides*

atgtggaaactcctcgtcagcggtcttgtcgccgtcgcgtccctcagcggcgtgaacgctgcttatccta
accctggtcccgtcaccggcgatactcgtgttcacgaccctacggttgtcaagactcccagcggtggata
cttgctggctcatactggcgataacgtttcgctcaagacttcttctgatcgaactgcttggaaggatgca
ggtgctgttttccccaacggtgcgccttggactacgcagtacaccaagggcgacaagaacctctgggccc
ctgatatctcctaccacaacggccagtactatctgtactactccgcctcttccttcggtcagcgtacctc
tgccatttttctcgctaccagcaagaccggtgcatccggctcgtggaccaaccaaggcgtcgtcgtcgag
tccaacaacaacaacgactacaatgccattgacggaaatctctttgtcgactctgatggaaaatggtggc
tctccttcggctctttctggtccggcatcaagctcatccaactcgaccccaagaccggcaagcgcaccgg
ctcaagcatgtactccctcgccaaacgcgacgcctccgtcgaaggcgccgtcgaggctccgttcatcacc
aaacgcggaagcacctactacctctgggtgtcgttcgacaagtgttgccagggcgctgctagcacgtacc
gtgtcatggttggacggtcgagcagcattactggtccttatgttgacaaggctggtaagcagatgatgtc
tggtggaggaacggagattatggctagtcacggatctattcatggaccgggacataatgctgttttcact
gataacgatgcggacgttcttgtctatcattactacgataacgctggcacagcgctgttgggcatcaact
tgctcagatatgacaatggctggcctgttgcttattag

FIG. 20A

SEQ ID NO:40
Protein sequence for Fv43F

mwkllvsglvavaslsgvnaaypnpgpvtgdtrvhdptvvktpsggyllahtgdnvslktssdrtawkda
gavfpngapwttqytkgdknlwapdisyhngqyylyysassfgqrtsaiflatsktgasgswtnqgvvve
snnnndynaidgnlfvdsdgkwwlsfgsfwsgikliqldpktgkrtgssmyslakrdasvegaveapfit
krgstyylwvsfdkccqgaastyrvmvgrsssitgpyvdkagkqmmsgggteimashgsihgpghnavft
dndadvlvyhyydnagtallginllrydngwpvay

FIG. 20B

447

SEQ ID NO:41
Nucleotide sequence for Xyn3, a GH10 xylanase from *Trichoderma reesei*

atgaaagcaaacgtcatcttgtgcctcctggccccctggtcgccgctctccccaccgaaaccatccacc
tcgaccccgagctcgccgctctccgcgccaacctcaccgagcgaacagccgacctctgggaccgccaagc
ctctcaaagcatcgaccagctcatcaagagaaaaggcaagctctactttggcaccgccaccgaccgcggc
ctcctccaacgggaaaagaacgcggccatcatccaggcagacctcggccaggtgacgccggagaacagca
tgaagtggcagtcgctcgagaacaaccaaggccagctgaactggggagacgccgactatctcgtcaactt
tgcccagcaaaacggcaagtcgatacgcggccacactctgatctggcactcgcagctgcctgcgtgggtg
aacaatatcaacaacgcggatactctgcggcaagtcatccgcacccatgtctctactgtggttgggcggt
acaagggcaagattcgtgcttgggtgagttttgaacaccacatgccccttttcttagtccgctcctcctc
ctcttggaacttctcacagttatagccgtatacaacattcgacaggaaatttaggatgacaactactgac
tgacttgtgtgtgtgatggcgataggacgtggtcaatgaaatcttcaacgaggatggaacgctgcgctct
tcagtcttttccaggctcctcggcgaggagtttgtctcgattgcctttcgtgctgctcgagatgctgacc
cttctgcccgtctttacatcaacgactacaatctcgaccgcgccaactatggcaaggtcaacgggttgaa
gacttacgtctccaagtggatctctcaaggagttcccattgacggtattggtgagccacgacccctaaat
gtcccccattagagtctctttctagagccaaggcttgaagccattcagggactgacacgagagccttctc
tacaggaagccagtcccatctcagcggcggcggaggctctggtacgctgggtgcgctccagcagctggca
acggtacccgtcaccgagctggccattaccgagctggacattcaggggcaccgacgacggattacaccc
aagttgttcaagcatgcctgagcgtctccaagtgcgtcggcatcaccgtgtggggcatcagtgacaaggt
aagttgcttccccctgtctgtgcttatcaactgtaagcagcaacaactgatgctgtctgtctttacctagg
actcgtggcgtgccagcaccaaccctcttctgtttgacgcaaacttcaaccccaagccggcatataacag
cattgttggcatcttacaatag

FIG. 21A

SEQ ID NO:42
Protein sequence for Xyn3

mkanvilcllaplvaalptetihldpelaalranltertadlwdrqasq**sidqlikrkgklyfgtatdrg**
llqreknaaiiqadlgqvtpensmkwqslennqgqlnwgdadylvnfaqqngksirghtliwhsqlpawv
nninnadtlrqvirthvstvvgrykgkirawdvvneifnedgtlrssvfsrllgeefvsiafraardadp
sarlyindynldranygkvnglktyvskwisqgvpidgigsqshlsggggsgtlgalqqlatvpvtelai
teldiqgapttdytqvvqaclsvskcvgitvwgisdkdswrastnpllfdanfnpkpaynsivgilq

FIG. 21B

SEQ ID NO:43
Protein sequence of Xyn2, a GH11 family xylanase from *Trichoderma reesei*

mvsftsllaasppsrascrpaaevesvavekrqtiqpgtgynngyfysywndghggvtytngpggqfsvn
wsnsgnfvggkgwqpgtknkvinfsgsynpngnsylsvygwsrnplieyyivenfgtynpstgatklgev
tsdgsvydiyrtqrvnqpsiigtatfyqywsvrrnhrssgsvntanhfnawaqqgltlgtmdyqivaveg
yfssgsasitvs

FIG. 22

SEQ ID NO:44
Protein sequence of Bxl1, a GH3 family β-xylosidase from *Trichoderma reesei*

mvnnaallaalsallptalaqnnqtyanysaqgqpdlypetlatltlsfpdcehgplknnlvcdssagyv
eraqalislftleelilntqnsgpgvprlglpnyqvwnealhgldranfatkggqfewatsfpmpiltta
alnrtlihqiadiistqarafsnsgrygldvyapnvngfrsplwgrgqetpgedafflssaytyeyitgi
qqgvdpehlkvaatvkhfagydlenwnnqsrlgfdaiitqqdlseyytpqflaaaryaksrslmcaynsv
ngvpscansfflqtllreswgfpewgyvssdcdavynvfnphdyasnqssaaasslragtdidcgqtypw
hlnesfvagevsrgeiersvtrlyanlvrlgyfdkknqyrslgwkdvvktdawnisyeaavegivllknd
gtlplskkvrsialigpwanattqmqgnyygpapylispleaakkagyhvnfelgteiagnsttgfakai
aaakksdaiiylggidntieqegadrtdiawpgnqldlikqlsevgkplvvlqmgggqvdssslksnkkv
nslvwggypgqsggvalfdilsgkrapagrlvttqypaeyvhqfpqndmnlrpdgksnpgqtyiwytgkp
vyefgsglfyttfketlashpkslkfntssilsaphpgytyseqipvftfeaniknsgktespytamlfv
rtsnagpapypnkwlvgfdrladikpghssklsipipvsalarvdshgnrivypgkyelalntdesvkle
felvgeevtienwpleeqqikdatpda

FIG. 23

SEQ ID NO:45
Protein sequence of Bgl1, a GH3 family β-glucosidase from *Trichoderma reesei*

mryrtaaalalatqpfaradshstsgasaeavvppagtpwgtaydkakaalaklnlqdkvgivsgvgwng
gpcvgntspaskisypslclqdgplgvrystgstaftpgvqaastwdvnlirergqfigeevkasgihvi
lgpvagplgktpqggrnwegfgvdpyltgiamgqtingiqsvgvqatakhyilneqelnretissnpddr
tlhelytwpfadavqanvasvmcsynkvnttwacedqytlqtvlkdqlgfpgyvmtdwnaqhttvqsans
gldmsmpgtdfngnnrlwgpaltnavnsnqvptsrvddmvtrilaawyltgqdqagypsfnisrnvqgnh
ktnvraiardgivllkndanilplkkpasiavvgsaaiignharnspscndkgcddgalgmgwgsgavny
pyfvapydaintrassqgtqvtlsntdntssgasaargkdvaivfitadsgegyitvegnagdrnnldpw
hngnalvqavaganснvivvvhsvgaiileqilalpqvkavvwaglpsqesgnalvdvlwgdvspsgklv
ytiakspndyntrivsggsdsfseglfidykhfddanitpryefgyglsytkfnysrlsvlstaksgpat
gavvpggpsdlfqnvatvtvdiansgqvtgaevaqlyityppssaprtppkqlrgfaklnltpgqsgtatf
nirrrdlsywdtasqkwvvpsgsfgisvgassrdirltstlsva

FIG. 24

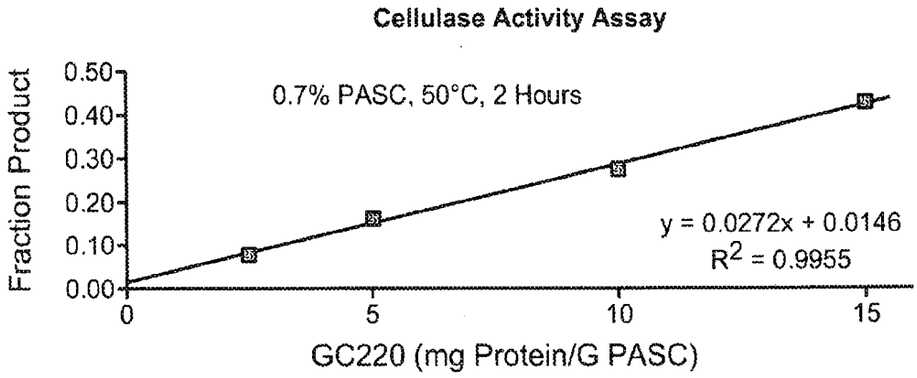


FIG. 25

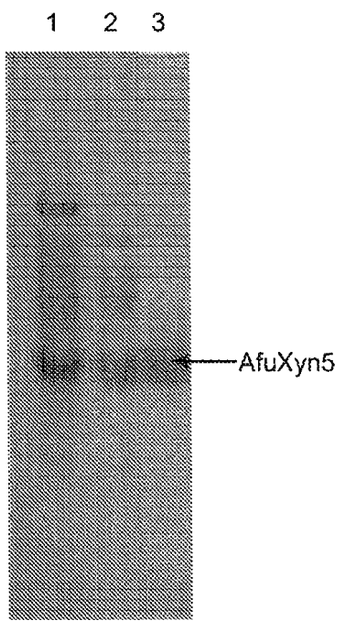


FIG. 27

450

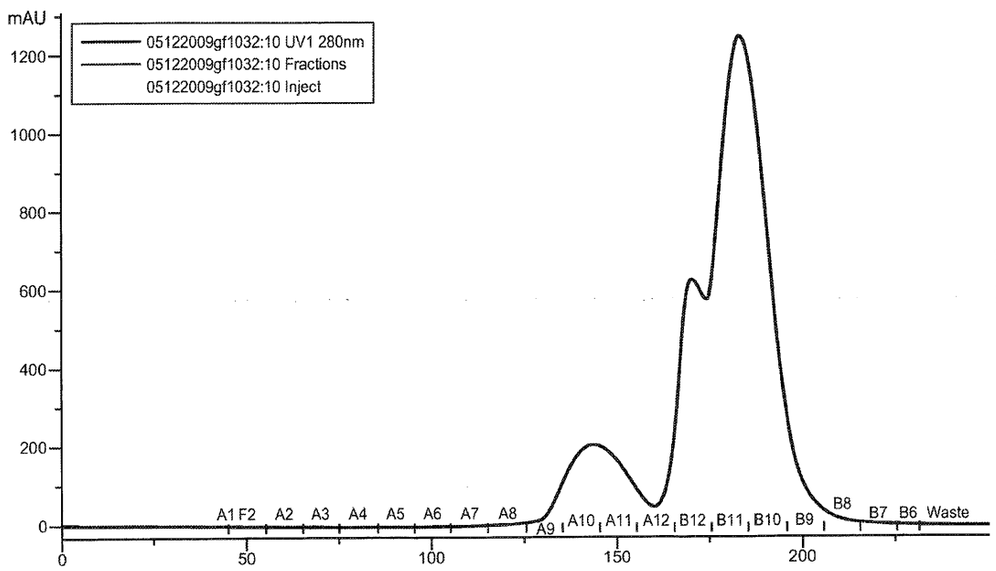


FIG. 26

SUBSTITUTE SHEET (RULE 26)

451

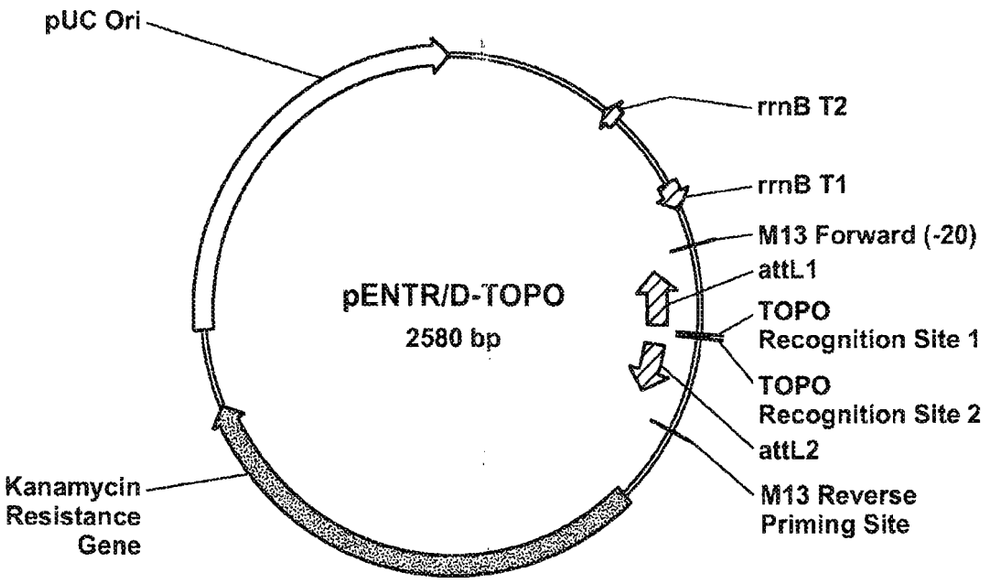

pUC Ori
rrnB T2
rrnB T1
M13 Forward (-20)
attL1
TOPO
Recognition Site 1
TOPO
Recognition Site 2
attL2
M13 Reverse
Priming Site
Kanamycin
Resistance
Gene
pENTR/D-TOPO
2580 bp

FIG. 28

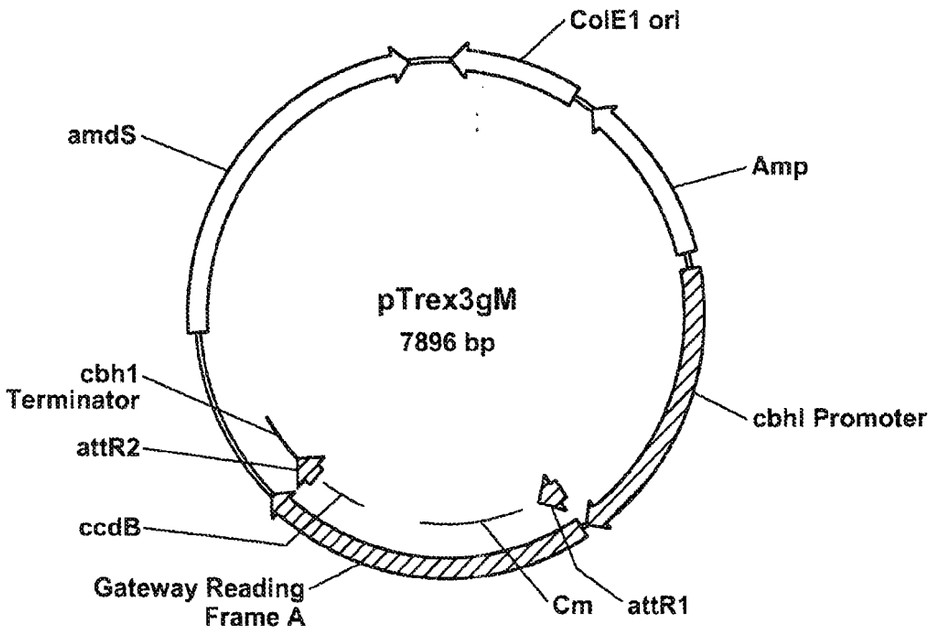

ColE1 ori
amdS
Amp
pTrex3gM
7896 bp
cbh1
Terminator
attR2
cbhI Promoter
ccdB
Gateway Reading
Frame A
Cm
attR1

FIG. 29

FIG. 30A

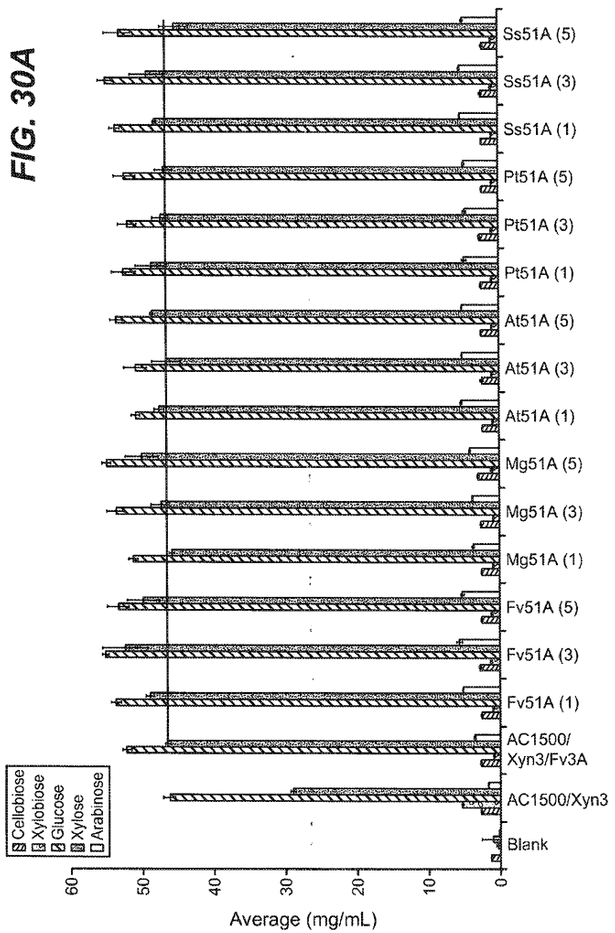

Ss51A (5)
Ss51A (3)
Ss51A (1)
Pt51A (5)
Pt51A (3)
Pt51A (1)
At51A (5)
At51A (3)
At51A (1)
Mg51A (5)
Mg51A (3)
Mg51A (1)
Fv51A (5)
Fv51A (3)
Fv51A (1)
AC1500/ Xyn3/Fv3A
AC1500/Xyn3
Blank
Cellobiose
Xylobiose
Glucose
Xylose
Arabinose
60
50
40
30
20
10
0
Average (mg/mL)

FIG. 30B

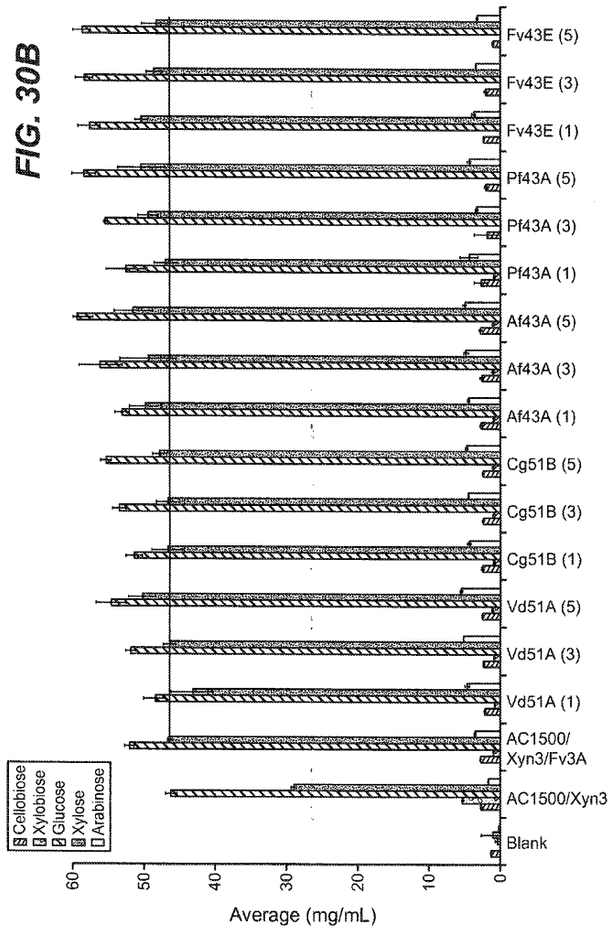

Fv43E (5)
Fv43E (3)
Fv43E (1)
Pf43A (5)
Pf43A (3)
Pf43A (1)
Af43A (5)
Af43A (3)
Af43A (1)
Cg51B (5)
Cg51B (3)
Cg51B (1)
Vd51A (5)
Vd51A (3)
Vd51A (1)
AC1500/ Xyn3/Fv3A
AC1500/Xyn3
Blank
Cellobiose
Xylobiose
Glucose
Xylose
Arabinose
0
10
20
30
40
50
60
Average (mg/mL)

FIG. 31

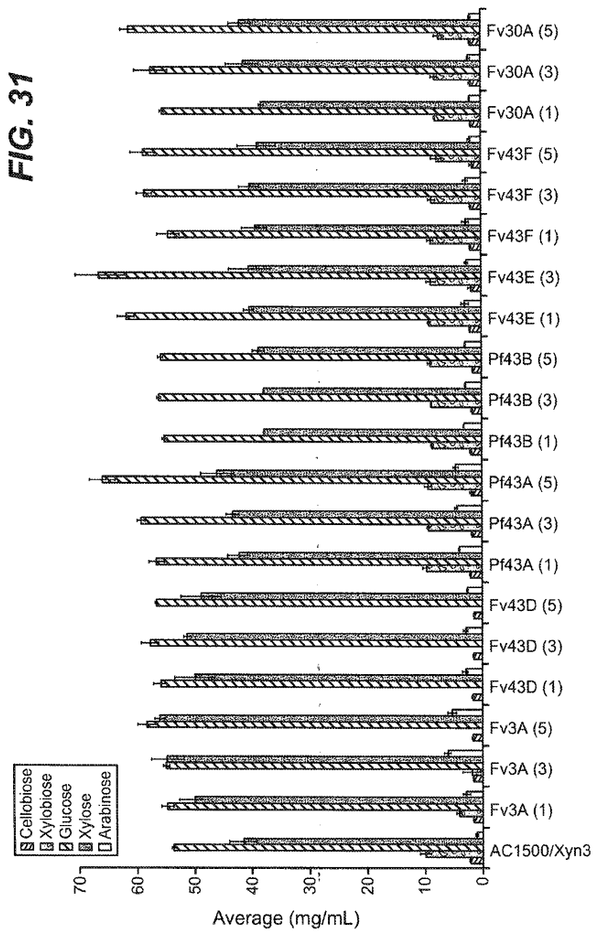

Fv30A (5)
Fv30A (3)
Fv30A (1)
Fv43F (5)
Fv43F (3)
Fv43F (1)
Fv43E (3)
Fv43E (1)
Pf43B (5)
Pf43B (3)
Pf43B (1)
Pf43A (5)
Pf43A (3)
Pf43A (1)
Fv43D (5)
Fv43D (3)
Fv43D (1)
Fv3A (5)
Fv3A (3)
Fv3A (1)
AC1500/Xyn3
Cellobiose
Xylobiose
Glucose
Xylose
Arabinose
70
60
50
40
30
20
10
0
Average (mg/mL)

FIG. 32

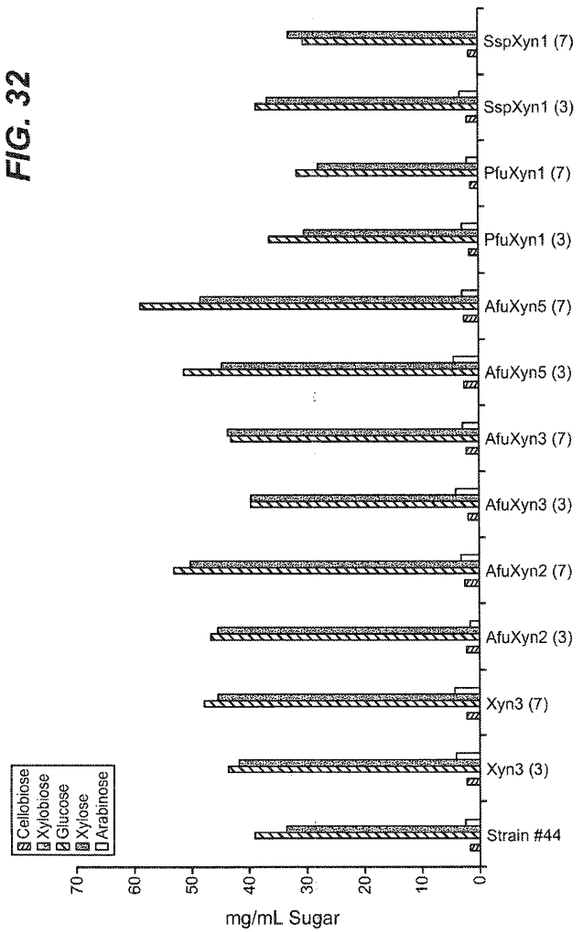
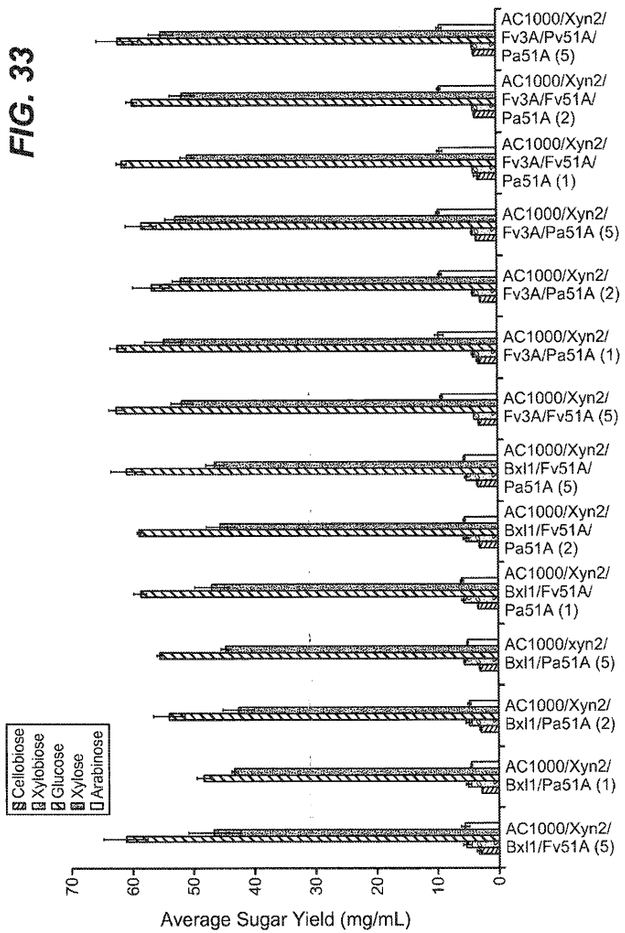
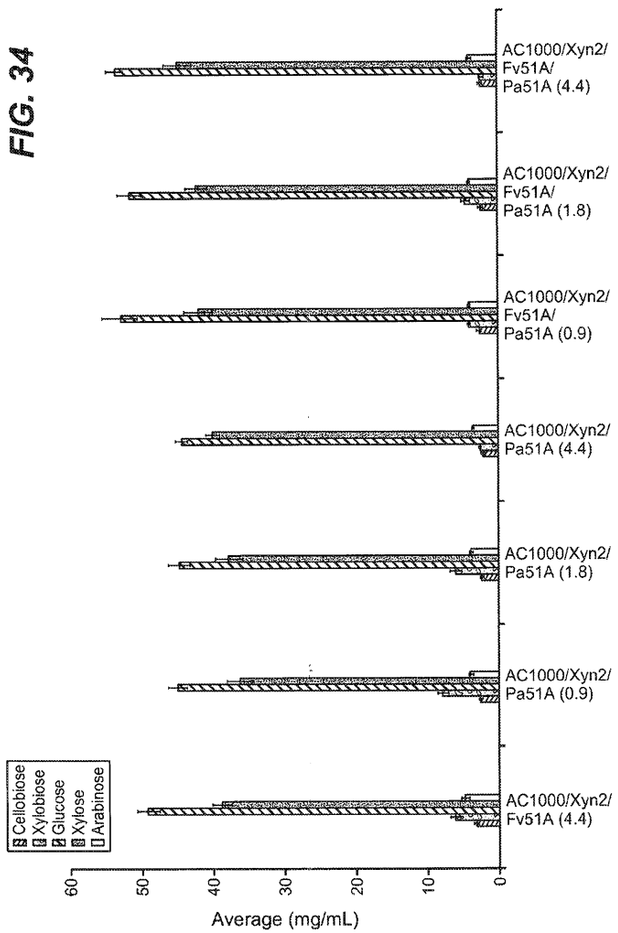

SspXyn1 (7)
SspXyn1 (3)
PfuXyn1 (7)
PfuXyn1 (3)
AfuXyn5 (7)
AfuXyn5 (3)
AfuXyn3 (7)
AfuXyn3 (3)
AfuXyn2 (7)
AfuXyn2 (3)
Xyn3 (7)
Xyn3 (3)
Strain #44
Cellobiose
Xylobiose
Glucose
Xylose
Arabinose
70
60
50
40
30
20
10
0
mg/mL Sugar

FIG. 33


AC1000/Xyn2/ Fv3A/Pv51A/ Pa51A (5)
AC1000/Xyn2/ Fv3A/Fv51A/ Pa51A (2)
AC1000/Xyn2/ Fv3A/Fv51A/ Pa51A (1)
AC1000/Xyn2/ Fv3A/Pa51A (5)
AC1000/Xyn2/ Fv3A/Pa51A (2)
AC1000/Xyn2/ Fv3A/Pa51A (1)
AC1000/Xyn2/ Fv3A/Fv51A (5)
AC1000/Xyn2/ Bxl1/Fv51A/ Pa51A (5)
AC1000/Xyn2/ Bxl1/Fv51A/ Pa51A (2)
AC1000/Xyn2/ Bxl1/Fv51A/ Pa51A (1)
AC1000/xyn2/ Bxl1/Pa51A (5)
AC1000/Xyn2/ Bxl1/Pa51A (2)
AC1000/Xyn2/ Bxl1/Pa51A (1)
AC1000/Xyn2/ Bxl1/Fv51A (5)
Cellobiose
Xylobiose
Glucose
Xylose
Arabinose
70
60
50
40
30
20
10
0
Average Sugar Yield (mg/mL)

458

FIG. 34


AC1000/Xyn2/ Fv51A/ Pa51A (4.4)
AC1000/Xyn2/ Fv51A/ Pa51A (1.8)
AC1000/Xyn2/ Fv51A/ Pa51A (0.9)
AC1000/Xyn2/ Pa51A (4.4)
AC1000/Xyn2/ Pa51A (1.8)
AC1000/Xyn2/ Pa51A (0.9)
AC1000/Xyn2/ Fv51A (4.4)
Cellobiose
Xylobiose
Glucose
Xylose
Arabinose
60
50
40
30
20
10
0
Average (mg/mL)

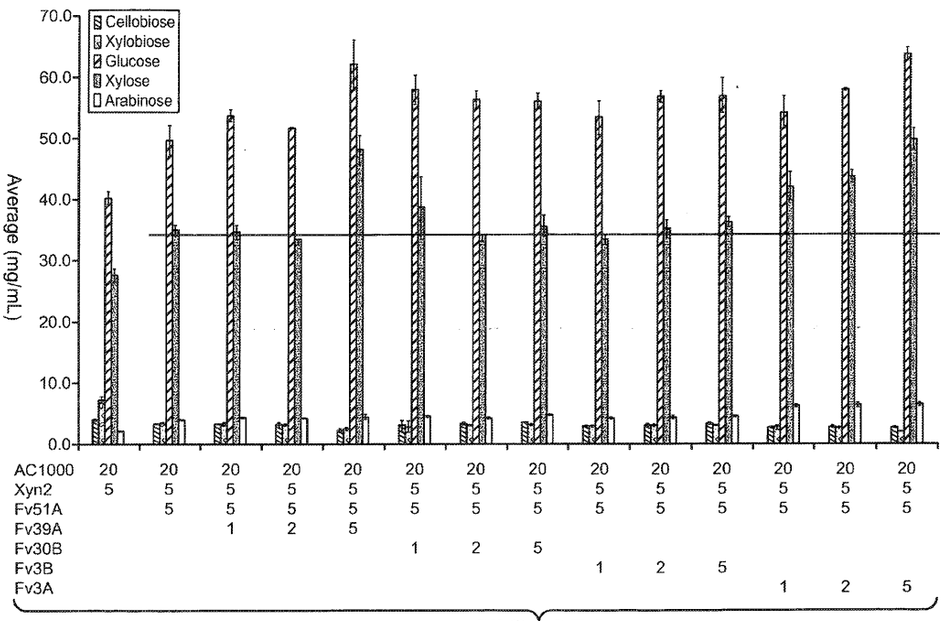


FIG. 35A

SUBSTITUTE SHEET (RULE 26)

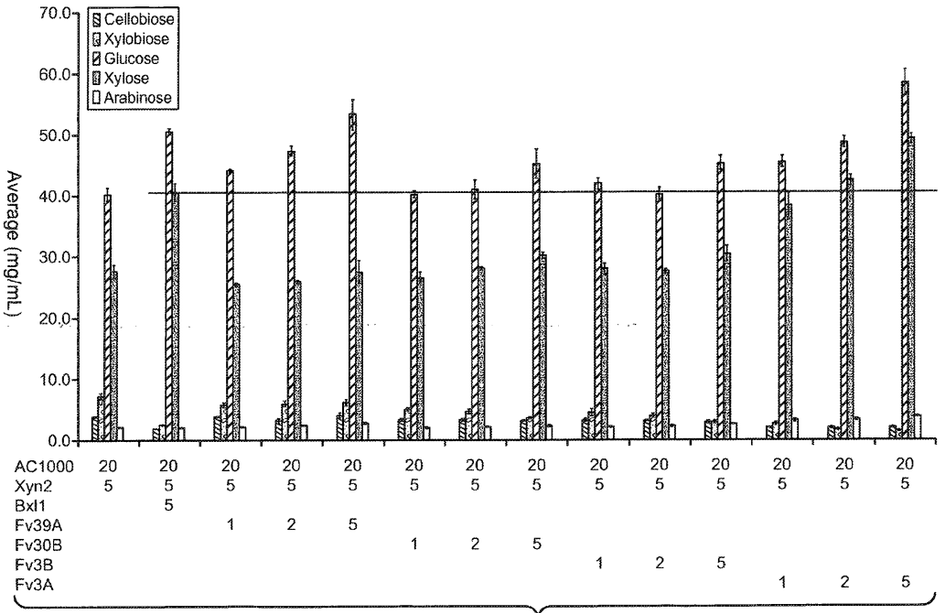


FIG. 35B

SUBSTITUTE SHEET (RULE 26)

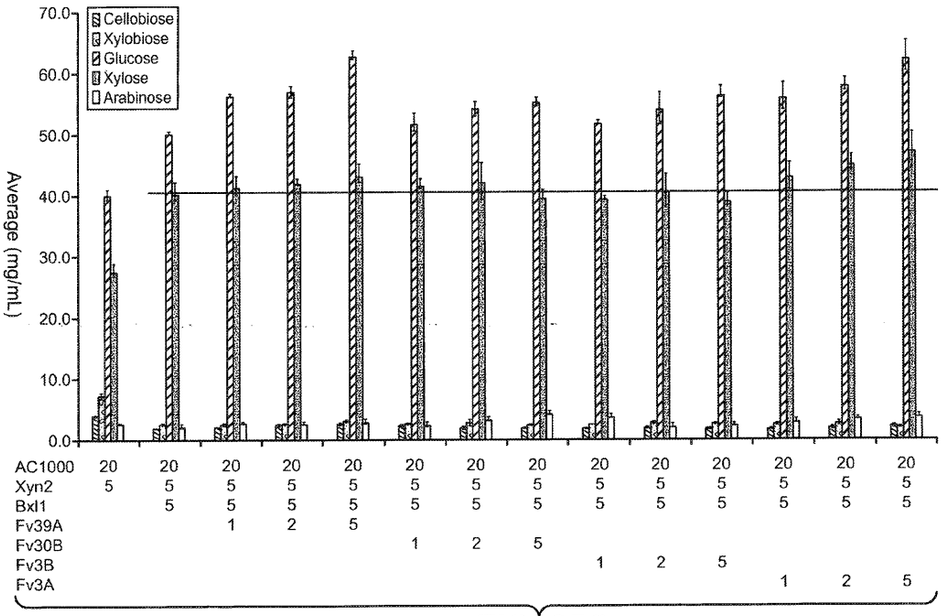


FIG. 35C

SUBSTITUTE SHEET (RULE 26)

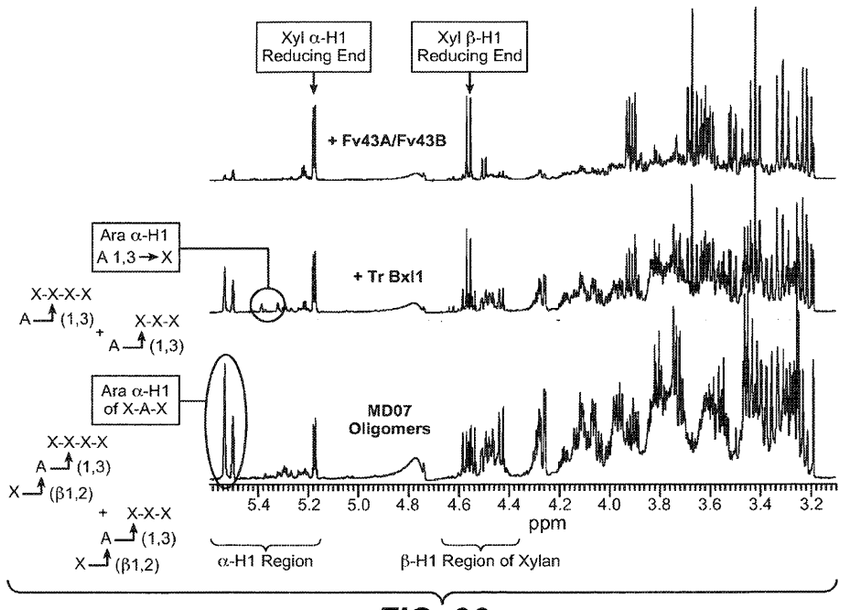


FIG. 36

SUBSTITUTE SHEET (RULE 26)

465

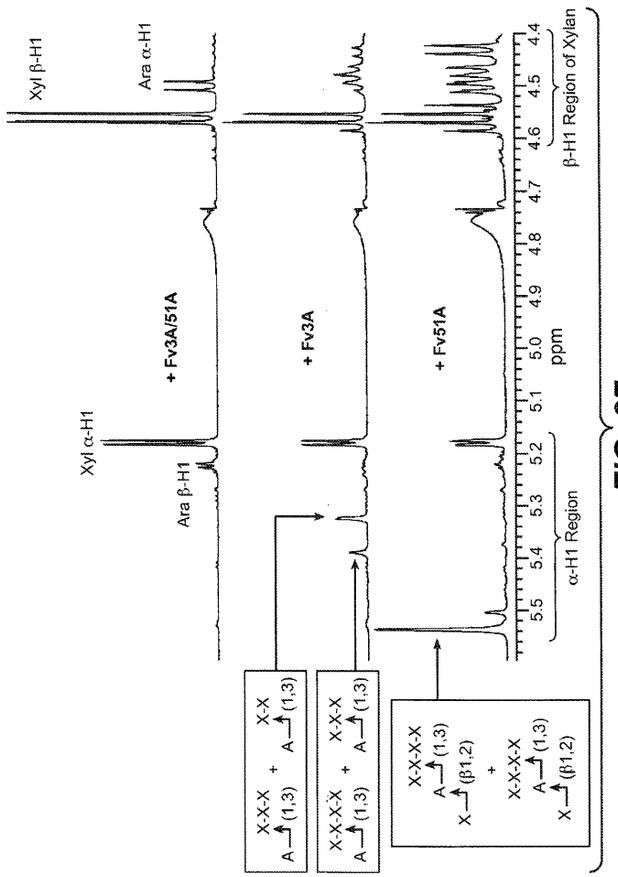

Xyl β-H1
Ara α-H1
Xyl α-H1
Ara β-H1
+ Fv3A/51A
+ Fv3A
+ Fv51A
X-X-X
A (1,3)
X-X
A (1,3)
X-X-X-X
A (1,3)
X-X-X
A (1,3)
X-X-X-X
A (1,3)
X (β1,2)
+
X-X-X-X
A (1,3)
X (β1,2)
5.5 5.4 5.3 5.2 5.1 5.0 4.9 4.8 4.7 4.6 4.5 4.4
ppm
α-H1 Region
β-H1 Region of Xylan

FIG. 37

```
gi|reesei|Bxl1  -MVNNAALLAALSALLPTALAQNNQTYANYSAQGQPDLYPETLATLTLSFPDCEHGPLKN
gi|fvert| Fv3A  MLLNLQVAASALSLSLLGGLAEAATPYT----------------------LPDCTKGPLSK
                 ::*  .  :***  *  .**:   .*:                      :*** :***.:
gi|reesei|      NLVCDSSAGYVERAQALISLFTLEELILNTQNSGPGVPRLGLPNYQVWNEALHGLDRAN-
gi|fvert|       NGICDTSLSPAKRAAALVAALTPEEKVGNLVSNATGAPRIGLPRYNWWNEALHGLAGSPG
                * :**:* . .:** **:: :* ** : *  ....*.**:***.*: ********  :
gi|reesei|      --FATKGGQFEWATSFPMPILTTAALNRTLIHQIADIISTQARAFSNSGRYGLDVYAPNV
gi|fvert|       GRFADTP-PYDAATSFPMPLLMAAAFDDDLIHDIGNVVGTEARAFTNGGWRGVDFWTPNV
                  ** .   :: *******:* :**::  ***:*.:::.*:****:*.*  *:*.::***
gi|reesei|      NGFRSPLWGRGQETPGEDAFFLSSAYTYEYITGIQGGVDPEHLKVAATVKHFAGYDLENW
gi|fvert|       NPFKDPRWGRGSETPGEDALHVS-RYARYIVRGLEG--DKEQRRIVATCKHYAGNDFEDW
                * *:.* ****.*******:.:*  *:   : *::*  * *: ::.** **:** *:*:*
gi|reesei|      NNQSRLGFDAIITQQDLSEYYTPQFLAAARYAKSRSLMCAYNSVNGVPSCANSFFLQTLL
gi|fvert|       GGFTRHDFDAKITPQDLAEYYVRPFQECTRDAKVGSIMCAYNAVNGIPACANSYLQETIL
                .. :* .*** ** ***:***.  *  .:* **  *:*****:***:*:****:: :*:*
gi|reesei|      RESWGFP-EWGYVSSDCDAVYNVFNPHDYASNQSSAAASSLRAGTDIDCGQTYPWHLNES
gi|fvert|       RGHWNWTRDNNWITSDCGAMQDIWQNHKYVKTNAEGAQVAFENGMDSSCEYTDTSDVSDS
                *  *.:. : .:::***.*: :::: *.*...::..*  ::. * * .*  * . .:.:*
gi|reesei|      FVAGEVSRGEIERSVTRLYANLVRLGYFD-KKNQYRSLGWKDVVKTDAWNISYEAAVEGI
gi|fvert|       YKQGLLTEKLMDRSLKRLFEGLVHTGFFDGAKAQWNSLSFADVNTKEAQDLALRSAVEGA
                :  * ::.  ::**:.**: .**: *:**  * *:.**.: ** ..:* ::: .:****
gi|reesei|      VLLKNDGTLPLSKKVR-SIALIGPWANATTQMQGNYYGPAPYLISPLEAAKKAGYHVNFE
gi|fvert|       VLLKNDGTLPLKLKKKDSVAMIGFWANDTSKLQGGYSGRAPFLHSPLYAAEKLGLDTNVA
                ***********. * : *:*:** *** *:::**.* * **:* *** **:* * ..*.
gi|reesei|      LGTEIAGNSTTG--FAKAIAAAKKSDAIIYLGGIDNTIEQEGADRTDIAWPGNQLDLIKQ
gi|fvert|       WGPTLQNSSSHDNWTTNAVAAAKKSDYILYFGGLDASAAGEDRDRENLDWPESQLTLLQK
                 *. : ..*: .   ::*:******* *:*:**:* :   *. ** :: ** .** *:::
gi|reesei|      LSEVGKPLVVLQMGGGQVDSSSLKSNKKVNSLVWGGYPGQSGGVALFDILSGKRAPAGRL
gi|fvert|       LSSLGKPLVVIQLG-DQVDDTALLKNKKINSILWVNYPGQDGGTAVMDLLTGRKSPAGRL
                **.:******:*:* .***.::* .***:**::* .****.**.*::*:*:*:::*****
gi|reesei|      VTTQYPAEYVHQFPQNDMNLRPDGKSNPGQTYIWYTGKPVYEFGSGLFYTTFKETLASHP
gi|fvert|       PVTQYPSKYTEQIGMTDMDLRPT-KSLPGRTYRWYS-TPVLPYGFGLHYTKFQAKFKSN-
                 .****::*..*:  .**:***  ** **:** **: .**  :* **.**.*: .: *:
gi|reesei|      KSLKFNTSSILSAPHPGYTYSEQIPVFTFEANIKNSGKTESPYTAMLFVRTSNAGPAPYP
gi|fvert|       -KLTFDIQKLLKG--CSAQYSDTCALPPIQVSVKNTGRITSDFVSLVFIKS-EVGPKPYP
                .*.*: ..:*..   .  **:  .: .::..:**:*:  * :.:::*::: :.** ***
gi|reesei|      NKWLVGFDRLADIKPGHSSKLSIPIPVSALARVDSHGNRIVYPGKYELALNTDESVKLEF
gi|fvert|       LKTLAAYGRLHDVAPSSTKDISLEWTLDNIARRGENGDLVVYPGTYTLLLDEPTQAKIQV
                 * *..:.** *: *. :..:*:  .:. :** ..:*: :****.* * *:   ..*::.
gi|reesei|      ELVGEEVTIENWPLEEQQIKDATPDA
gi|fvert|       TLTGKKAILDKWPQDPKSA--------
```

FIG. 38

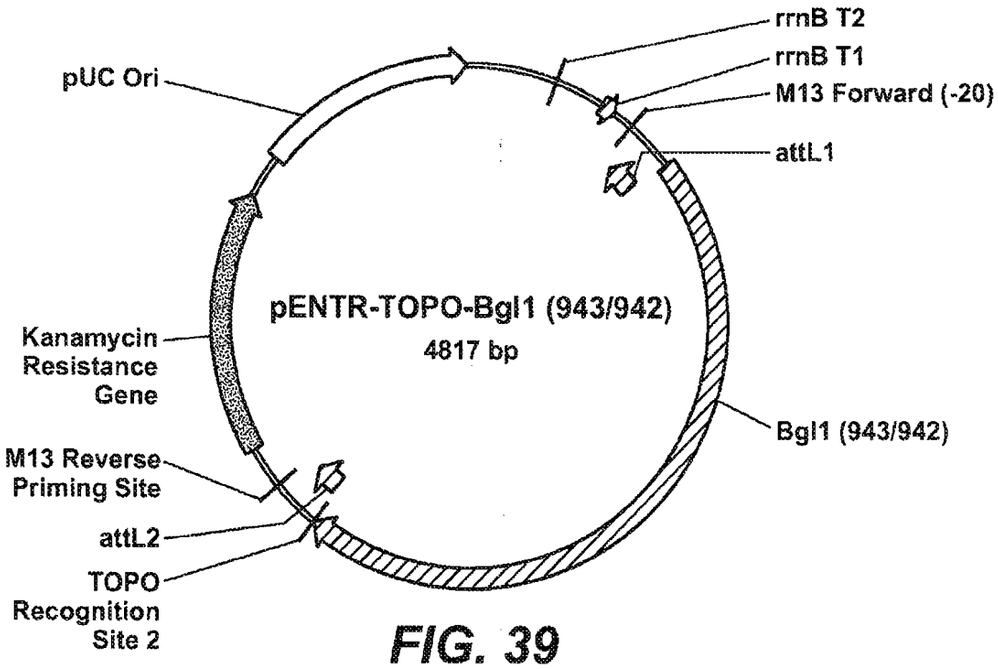

rrnB T2
rrnB T1
M13 Forward (-20)
attL1
pUC Ori
pENTR-TOPO-Bgl1 (943/942)
4817 bp
Kanamycin Resistance Gene
Bgl1 (943/942)
M13 Reverse Priming Site
attL2
TOPO Recognition Site 2

FIG. 39

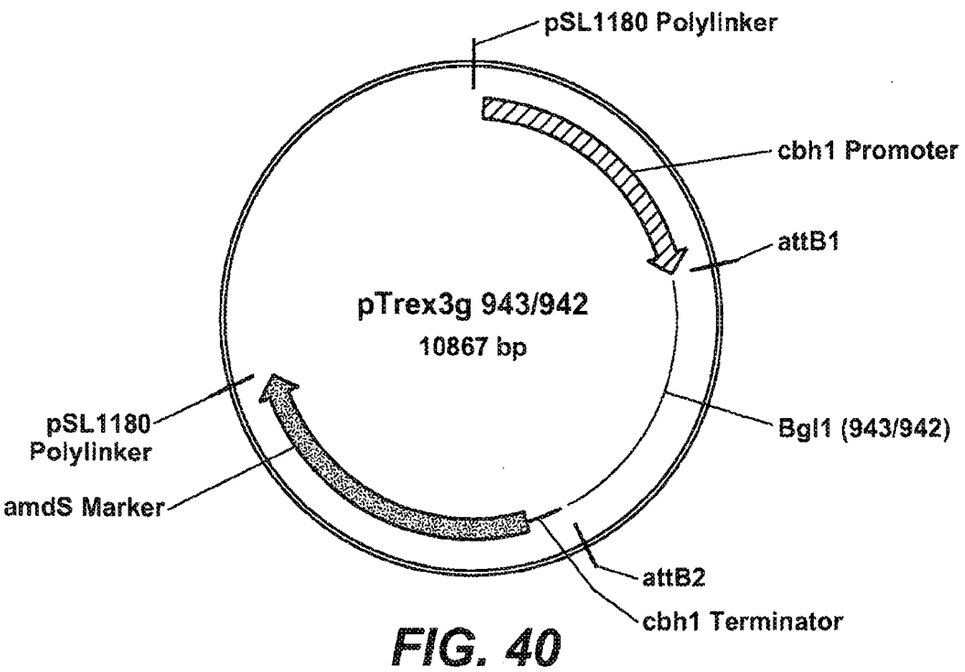

pSL1180 Polylinker
cbh1 Promoter
attB1
pTrex3g 943/942
10867 bp
Bgl1 (943/942)
pSL1180 Polylinker
amdS Marker
attB2
cbh1 Terminator

FIG. 40

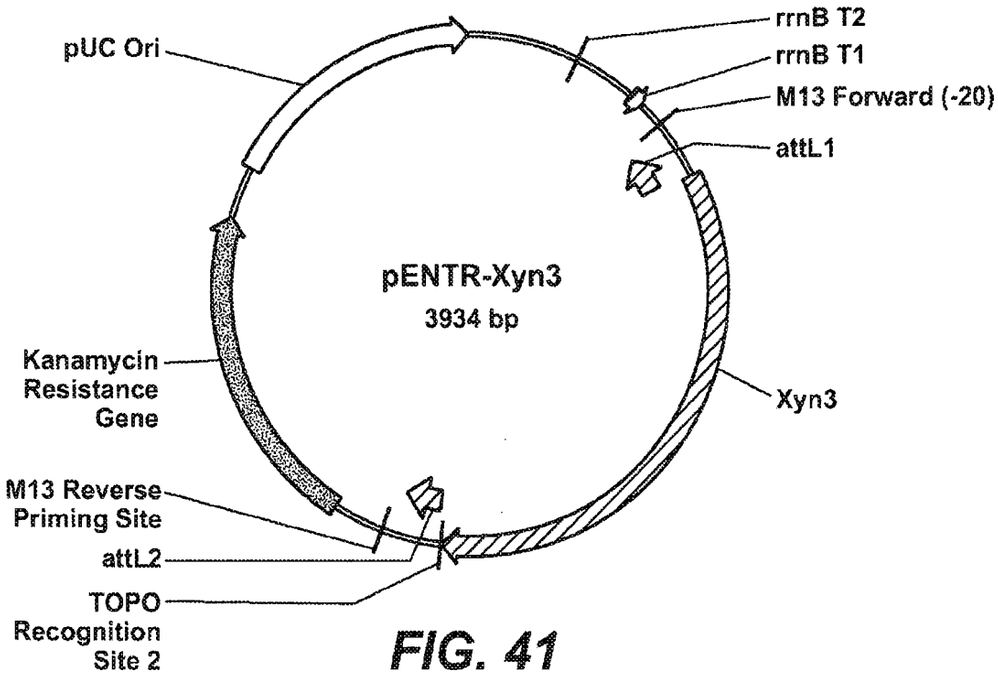

pUC Ori
rrnB T2
rrnB T1
M13 Forward (-20)
attL1
pENTR-Xyn3
3934 bp
Kanamycin Resistance Gene
Xyn3
M13 Reverse Priming Site
attL2
TOPO Recognition Site 2

FIG. 41

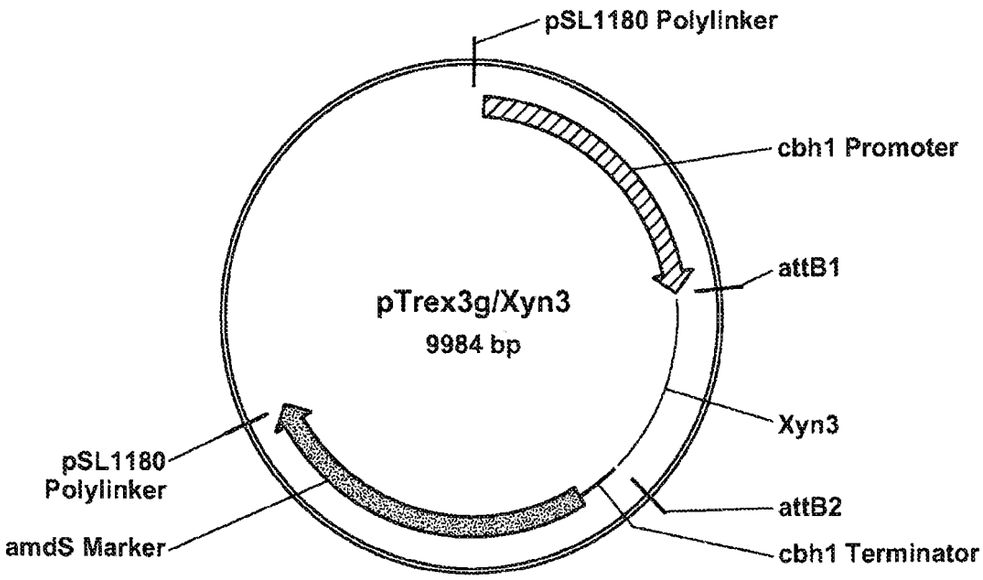

pSL1180 Polylinker
cbh1 Promoter
attB1
pTrex3g/Xyn3
9984 bp
Xyn3
pSL1180 Polylinker
attB2
amdS Marker
cbh1 Terminator

FIG. 42

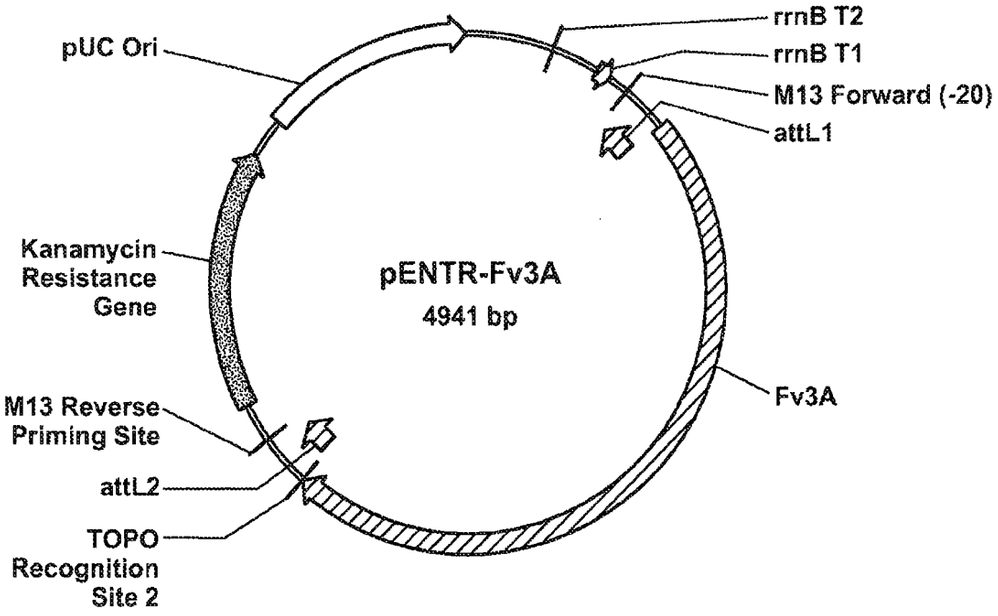

pUC Ori
rrnB T2
rrnB T1
M13 Forward (-20)
attL1
Kanamycin Resistance Gene
pENTR-Fv3A
4941 bp
Fv3A
M13 Reverse Priming Site
attL2
TOPO Recognition Site 2

FIG. 43

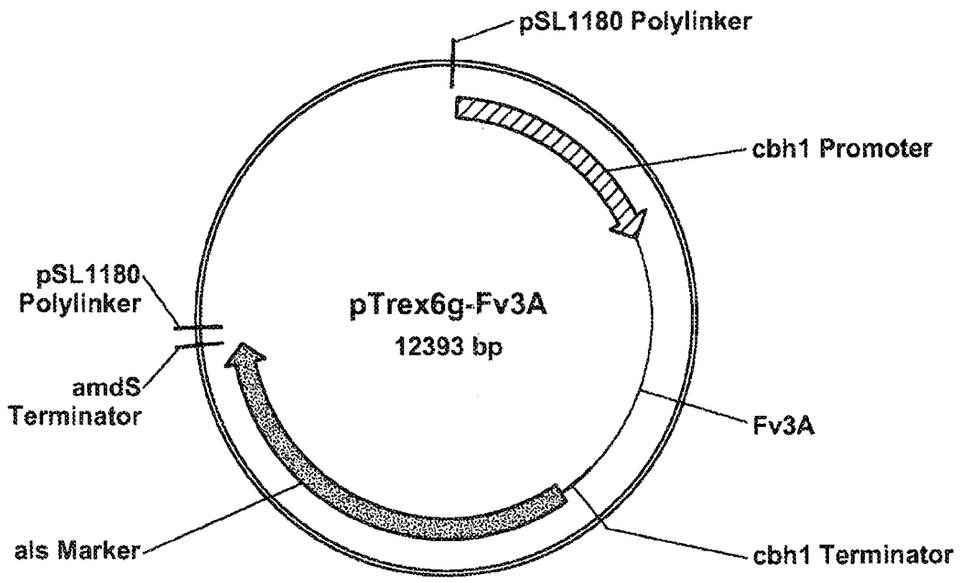

pSL1180 Polylinker
cbh1 Promoter
pSL1180 Polylinker
pTrex6g-Fv3A
12393 bp
amdS Terminator
Fv3A
als Marker
cbh1 Terminator

FIG. 44

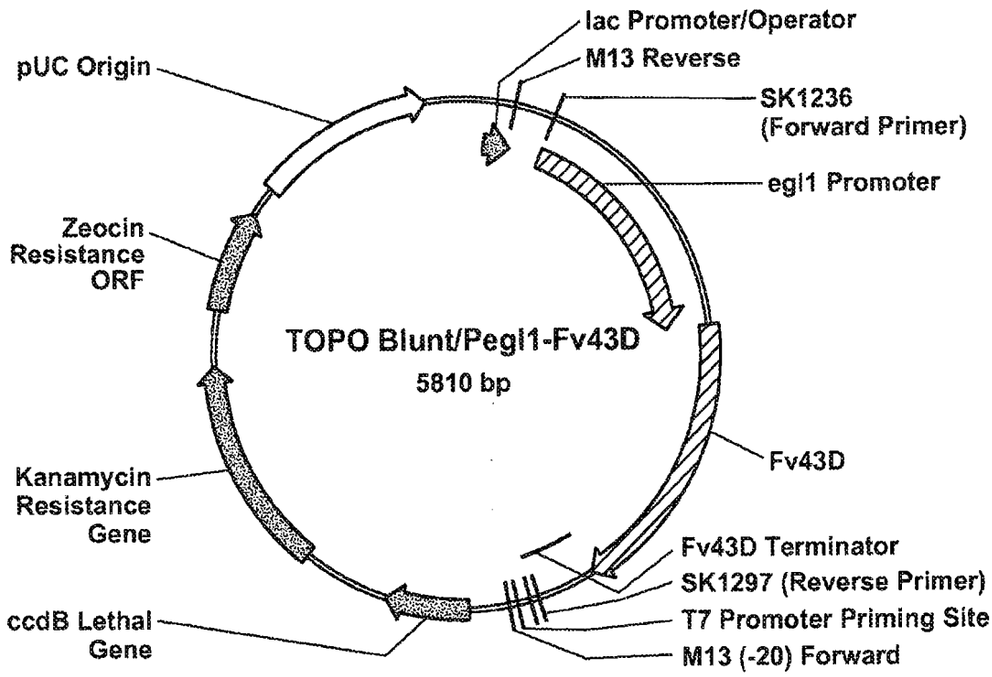

lac Promoter/Operator
M13 Reverse
SK1236
(Forward Primer)
egl1 Promoter
pUC Origin
Zeocin
Resistance
ORF
TOPO Blunt/Pegl1-Fv43D
5810 bp
Fv43D
Kanamycin
Resistance
Gene
Fv43D Terminator
SK1297 (Reverse Primer)
T7 Promoter Priming Site
ccdB Lethal
Gene
M13 (-20) Forward

FIG. 45

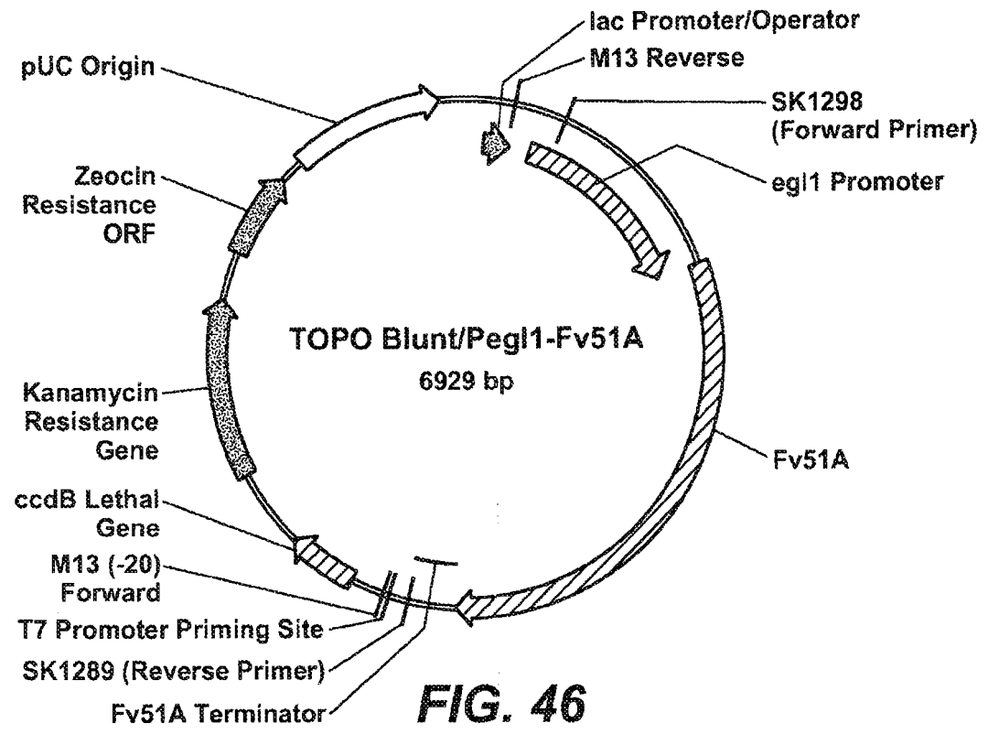

lac Promoter/Operator
M13 Reverse
SK1298
(Forward Primer)
egl1 Promoter
pUC Origin
Zeocin
Resistance
ORF
TOPO Blunt/Pegl1-Fv51A
6929 bp
Fv51A
Kanamycin
Resistance
Gene
ccdB Lethal
Gene
M13 (-20)
Forward
T7 Promoter Priming Site
SK1289 (Reverse Primer)
Fv51A Terminator

FIG. 46

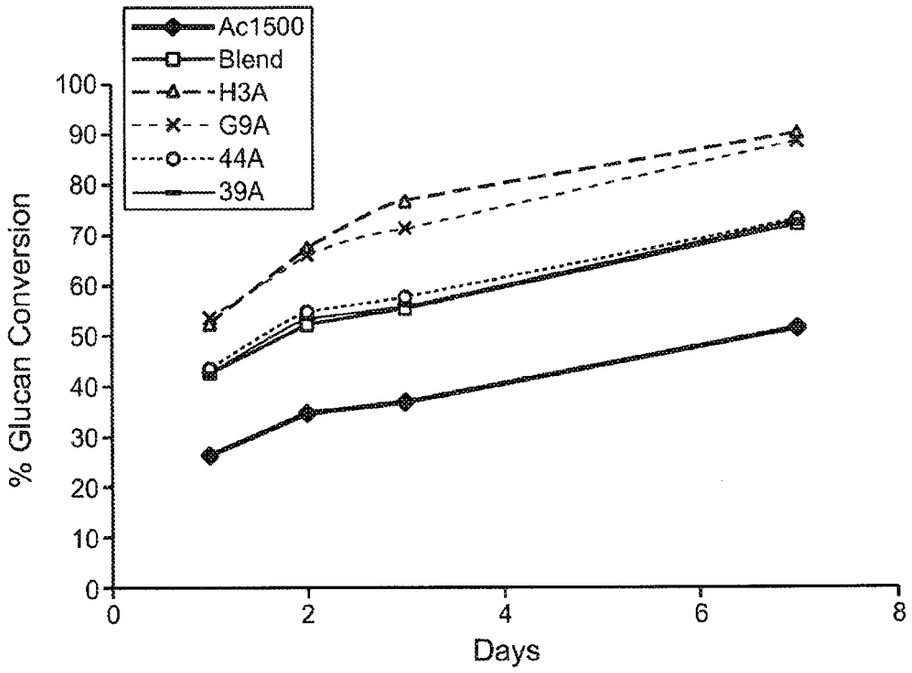


FIG. 47A

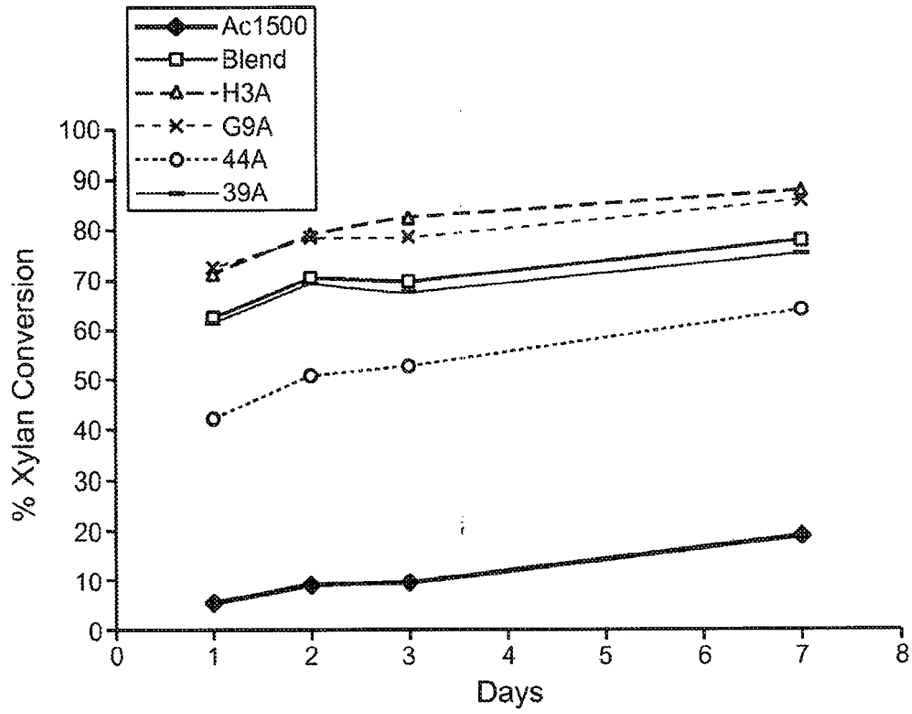


FIG. 47B

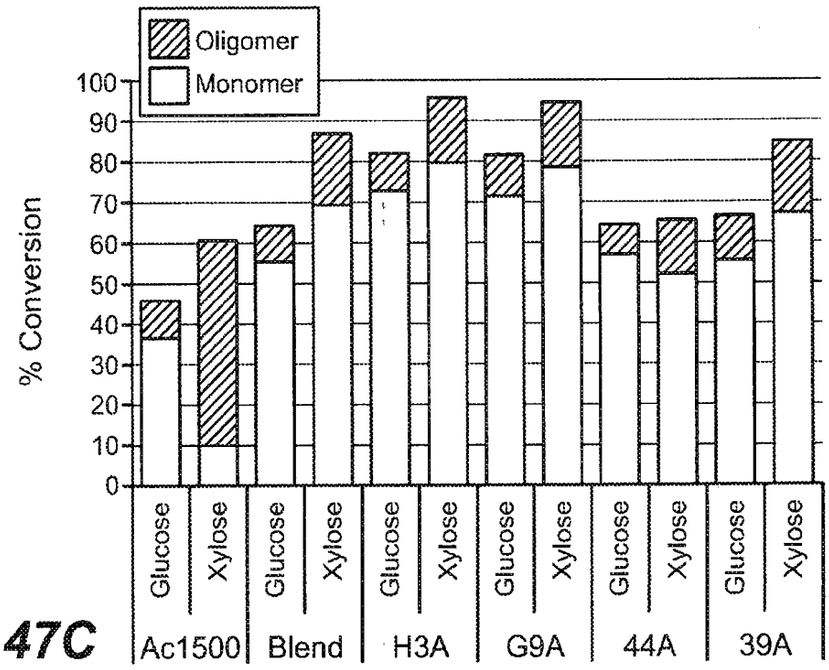


FIG. 47C

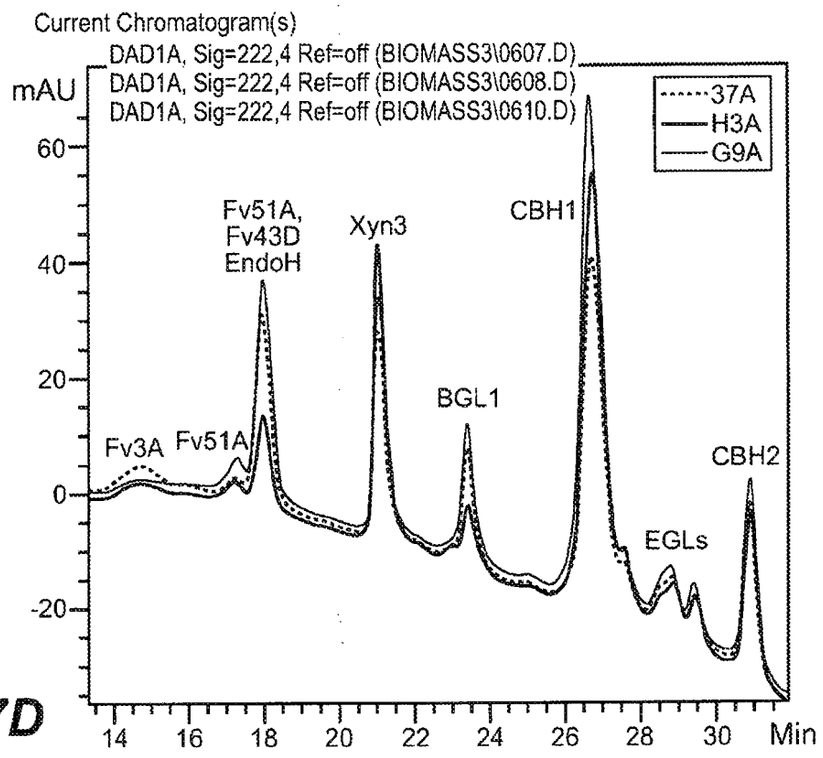


FIG. 47D

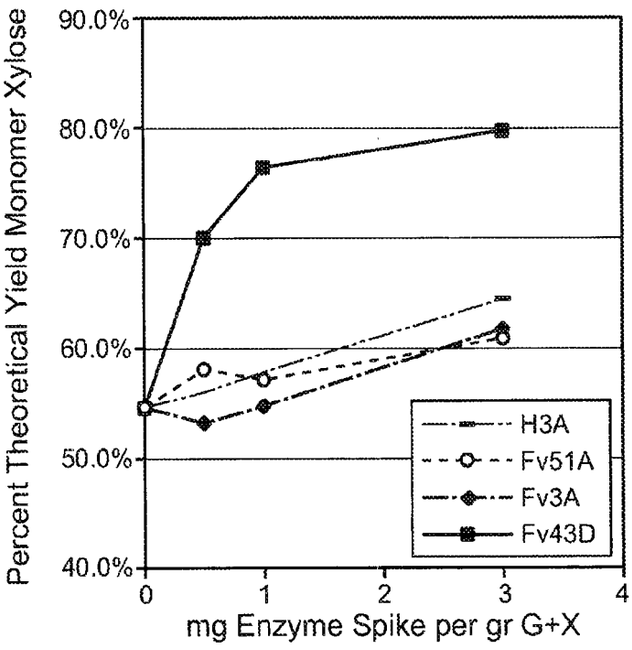


FIG. 48A

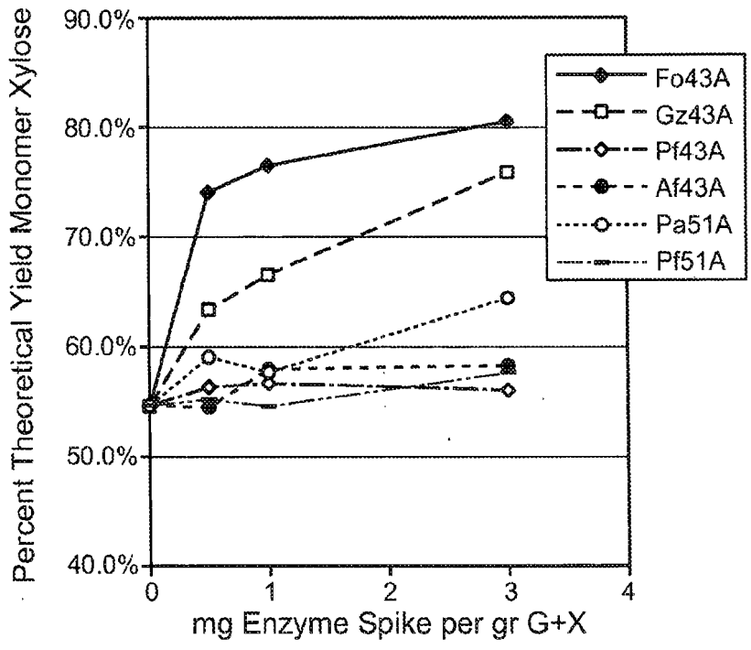


FIG. 48B

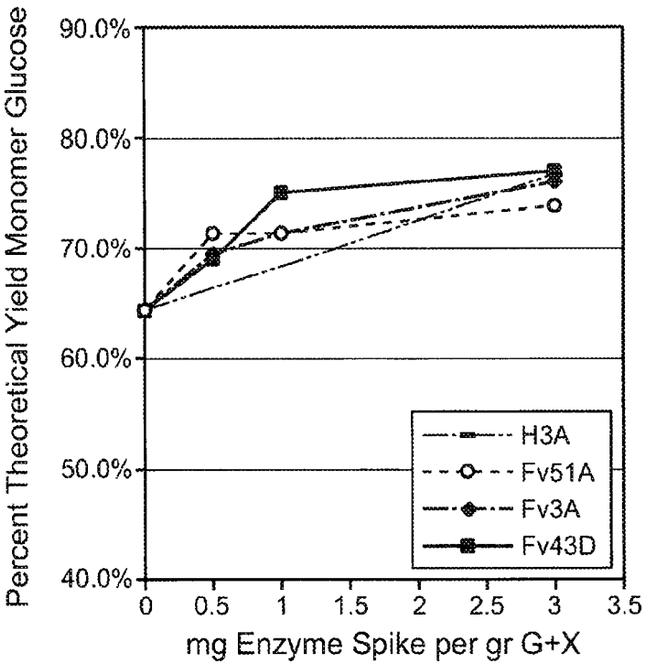


FIG. 49A

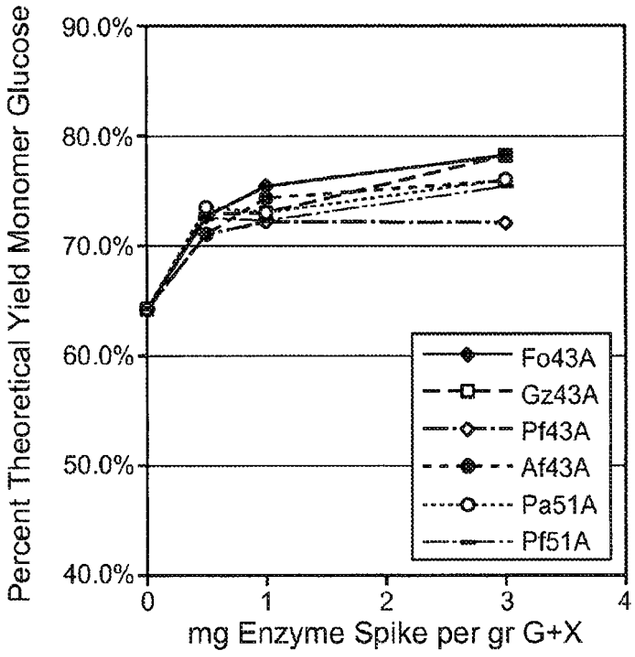


FIG. 49B

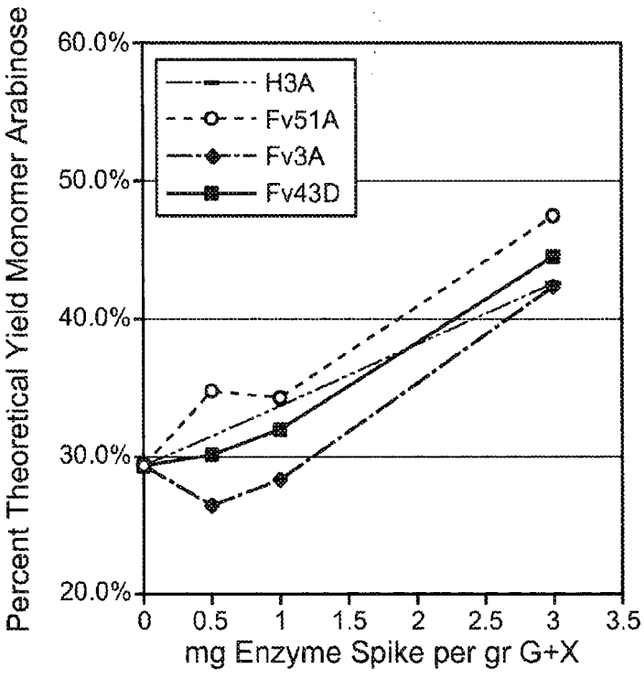


FIG. 50A

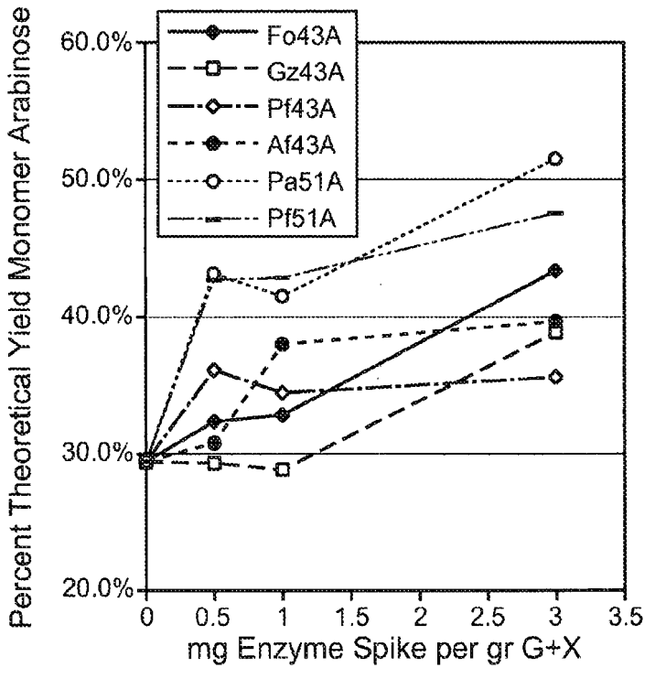


FIG. 50B

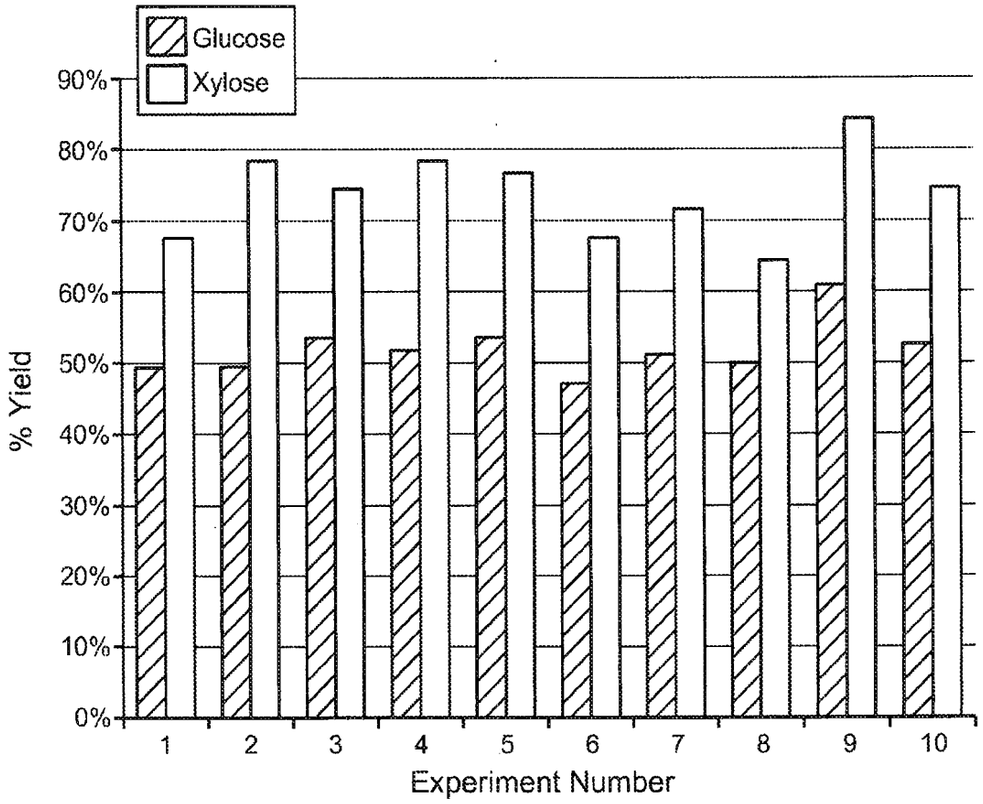

Glucose
Xylose
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
% Yield
1
2
3
4
5
6
7
8
9
10
Experiment Number

FIG. 51

```
                    *         20          *         40          *         60
Fv39A       :  MHYATLTTLVLALTTNVAAQQGTATVDLSKNHGPAKALGSGFIYGWPDNGTSVDTSIPDF : 60
Q9ZFM2_XYN  :  --------------MKVVNVPSNGREKFKKNW--KFCVGTGRL-GLALQKEYLDH--LKL : 41
P36906_XYN  :  --------------MIKVRVPDFSDKKFSDRW--RYCVGTGRL-GLALQKEYIET--LKY : 41
                    *         80          *         100         *         120
Fv39A       :  LVTDIKFNSNRGGGAQIPSLGWARGGYEG----------YLGR-FNSTLSNYRTTRKYNA : 109
Q9ZFM2_XYN  :  VQEKIGFRYIRGHGLLSDDVGIYREVEIDGEMKPFYNFTYIDRIVDSYLA--LNIRPF-I : 98
P36906_XYN  :  VKENIDFKYIRGHGLLCDDVGIYREDVVGDEVKPFYNFTYIDRIFDSFLE--IGIRPF-V : 98
                    *         140         *         160         *         180
Fv39A       :  DFILLPHDLWGADGG----QGSNSPFPGDNGNWTEM-ELFWNQLVSDLKAHNMLEGLVID : 164
Q9ZFM2_XYN  :  EFGFMPKALASGDQTVEYWKGNVTP-PKDYNKWRDLIVAVVSHFIERYGIEEVRT-WLFE : 156
P36906_XYN  :  EIGFMPKKLASGTQTVEYWEGNVTP-PKDYEKWSDLVKAVLHHFISRYGIEEVLK-WPFE : 156
                  A     *         200         *         220         *         240
Fv39A       :  VWNEPDIDIEW-DRPWSQFLEYYNRATKLLRKTLPKTLLSGFAMAHSPILSDDKWHTWLQ : 223
Q9ZFM2_XYN  :  VWNEPNLVNEWKDANKQEYFKLYEVTARAVKSVDPHLQVGGPAICGGS----DEWITDFL : 212
P36906_XYN  :  IWNEPNLKEEWKDADEKEYFKLYKVTAKAIKEVNENLKVGGPAICGGA----DYWIEDFL : 212
                    *         260         *         280         *         300
Fv39A       :  SVAGNKTVP-DIYSWHQIGAW--------------EREP-DSTIPDFTTLRA---QYGVP : 264
Q9ZFM2_XYN  :  HFCAERRVPVDFVSRHAYTSKAPHKKTFEYYYQELELEPPEDMLEQFKTVRALIRQSPFP : 272
P36906_XYN  :  NFCYEENVPVDFVSRHATTSK-QGEYTPHLIYQ--EIMPSEYMLNEFKTVREIIKNSHFP : 269
                  N *         320         *         340         *         360
Fv39A       :  EKPIDVNEYAARDEQNPANSVYYLSQLERHNLRGLRANWGSGSDLHNWMGNLIYSTTGTS : 324
Q9ZFM2_XYN  :  HLPLHITEYNT--SYSPINPVHDTALNAAYIARILSEGGDYVDSFSYWTFSDVFEEMDVP : 330
P36906_XYN  :  NLPFHITEYNT--SYSPQNPVHDTPFNAAYIARILSEGGDYVDSFSYWTFSDVFEERDVP : 327
                    *         380         *         400         *         420
Fv39A       :  EGTYY----------PNGEWQAYKYYAAMAGQRL----------------------VT : 350
Q9ZFM2_XYN  :  KALFHGGFGLVALHSIPKPTFHAFTFFNALGDELLYRDGEMIVTRRKDGSIAAVLWNLVM : 390
P36906_XYN  :  RSQFHGGFGLVALNMIPKPTFYTFKFFNAMGEEMLYRDEHMLVTRRDDGSVALIAWNEVM : 387
                    *         440         *         460         *         480
Fv39A       :  KASSDLKFDVFATKQGRKIKIIAGTRTVQAKYNIKISGLEVAGLPKMGTVKVRTYRFDWA : 410
Q9ZFM2_XYN  :  EKGEGLTKEVQLVIPVSFSAVFIKRQIVNEQYGNAWRVWKQMGRPRFPSRQAVETLPS-A : 449
P36906_XYN  :  DKTENPDEDYEVEIPVRFRDVFIKRQLIDEEHGNPWGTWIHMGRPRYPSKEQVNTLREVA : 447
                    *         500         *         520         *
Fv39A       :  GP----------NGKVDGPVDLGEKKYT-YSANTVSSPST--------------- : 439
Q9ZFM2_XYN  :  QPHVMTEQRRATDGVIHLSIVLSKNEVTLIEIEQVRDETSTYVGLDDGEITSYSS : 504
P36906_XYN  :  KPEIMTSQPVANDGYLNLKFKLGKNAVVLYELTERIDESSTYIGLDDSKINGY-- : 500
```

FIG. 52

```
Fv43D ---MQLKFLSSALLLSLTGNCAAQDTNDIPPLITDLWSADPSAHVFEGKLWVYPSHDIEA
Fo43A ---MQLKFLSSALLFSLTSKCAAQDTNDIPPLITDLWSADPSAHVFEGKLWVYPSHDIEA
Gz43A ---MKSKLLFP--LLSFVG--QSLATNDDCPLITSRWTADPSAHVFNDTLWLYPSHDIDA
Pf43A --MLQRFAYILPLALLSVG--VKADN----PFVQSIYTADPAPMVYNDRVYVFMDHDNTG
Fv43A -MWLTSPLLFASTLLGLTGVALADN-----PIVQDIYTADPAPMVYNGRVYLFTGHDNDG
Fv43B --MRFSWLLCPLLAMGSALPETKTDVSTYTNPVLPGWHSDPSC-IQKDGLFLCVTSTFIS
Af43A --------------MAAPSLSYPTGIQSYTNPLFPGWHSDPSCAYVAEQDTFFCVTSTFI
Pf43B ----MSRSILPYASVFALLGGAIAEP----FLVLNSDFPDPSLIETSSGYYAFGTTGNGV
Fv43E MKVYWLVAWATSLTPALAGLIGHRRATTFNNPIIYSDFPDNDVFLGPDNYYYFSASNFHF

Fv43D NVVNGTGGAQYAMRDYHTYSMKTIYGKDPVIDHGVALSVDDVPWAKQQMWAPDAAYK--N
Fo43A NVVNGTGGAQYAMRDYHTYSMKSIYGKDPVVDHGVALSVDDVPWAKQQMWAPDAAHK--N
Gz43A GFENDPDGGQYAMRDYHVYSIDKIYGSLP-VDHGTALSVEDVPWASRQMWAPDAAHK--N
Pf43A -------ATYYNMTDWHLFSSADMANWQD---HGIPMSLANFTWANANAWAPQVIPR--N
Fv43A -------STDFNMTDWRLFSSADMVNWQH---HGVPMSLKTFSWANSRAWAGQVVAR--N
Fv43B FP----GLPVYASRDLVNWRLISHVWNRE---KQLPGISWKTAGQQQGMYAPTIRYH--K
Af43A AFP---GLPLYASRDLQNWKLASNIFNRP---SQIPDLR-VTDGQQSGIYAPTLRYH--E
Pf43B N------AQVASSPDFNTWTLLSGT-------DALPGPFPSWVASSPQIWAPDVLVKA-D
Fv43E SP----GAPVLKSKDLLNWDLIGHSIPRLNFGDGYDLPPGSRYYRG-GTWASSLRYRKSN

Fv43D GKYYLYFPAK-DK-DEIFRIGVAVSNKPSGPFK---ADK-SWIPGTYSIDPASYVDTNGE
Fo43A GKYYLYFPAK-DK-DEIFRIGVAVSNKPSGPFK---ADK-SWIPGTYSIDPASYVDTDNE
Gz43A GKYYLYFPAK-DK-DDIFRIGVAVSPTPGGPFV---PDK-SWIPHTFSIDPASFVDDDDR
Pf43A GQFYFYAPVR-HN-DGSMAIGVGVSSTITGPYH---DAIGKPLVENNEIDPTVFIDDDGQ
Fv43A GKFYFYVPVRNAK-TGGMAIGVGVSTNILGPYT---DALGKPLVENNEIDPTVYIDTDGQ
Fv43B GTYYVICEYLGVG-DIIGVIFKTTNPWDESSWS---DPV---TFKPNHIDPDLFWDDDGK
Af43A GQFYLIVSYLGP--QTKGLLFTSSDPYDDAAWS---DPL---EFAVHGIDPDIFWDHDGT
Pf43B GTYVMYFSASAASDSGKHCVGAATATSPEGPYTPVDSAVACPLDQGGAIDANGFIDTDGT
Fv43E GQWYWIGCIN------FWQTWVYTASSPEGPWY----NKGNFGDNNCYYDNGILIDDDDT

Fv43D AYLIWGGI-WGGQLQAWQDHKTFNESWLGDKAAPNGTNALSPQIAKLSKDMHKITETPRD
Fo43A AYLIWGGI-WGGQLQAWQDKKNFNESWIGDKAAPNGTNALSPQIAKLSKDMHKITETPRD
Gz43A AYLAWGGI-MGGQLQRWQDKNKYNES--GTEPG-NGTAALSPQIAKLSKDMHTLAEKPRD
Pf43A AYLYWG----------------------------------NPDLWYVKLNQDMISYSGSPTQ
Fv43A AYLYWG----------------------------------NPGLYYVKLNQDMLSYSGSINK
Fv43B VYCATHG----ITLQEIDLETGELSPELNIWNGTGGVWPEGPHIYKRDGYYYLMIAEGGT
Af43A VYVTSAED-QMIKQYTLDLKTGAIGPVDYLWNGTGGVWPEGPHIYKRDGYYYLMIAEGGT
Pf43B IYVVYKID--------------------------GNSLDGDGTTHPTPIMLQQMEADGT
Fv43E MYVVYGSGEVKVSQLSQDGFSQVKSQVVFKNTDIGVQDLEGNRMYKING-------LYYI
```

FIG. 53A

```
Fv43D LVILAPETGKPLQAEDNKRRFFEGP-----WVHKRGKLYYLMYSTG--------------
Fo43A LVILAPETGKPLQAEDNKRRFFEGP-----WIHKRGKLYYLMYSTG--------------
Gz43A MLILDPKTGKPLLSEDEDRRFFEGP-----WIHKRNKIYYLTYSTG--------------
Pf43A IPLTTAGFGTRTGNAQRPTTFEEAP-----WVYKRNGIYYIAYAAD--------------
Fv43A VSLTTAGFGSRPNNAQRPTTFEEGP-----WLYKRGNLYYMIYAAN--------------
Fv43B ----AEDHAITIARARKITGPYEAYNNNPILTNRGTSEYFQTVGHGDLFQDTKGNWWGLC
Af43A ----ELGHSETMARSRTRTGPWEPYPHNPLLSNKGTSEYFQTVGHADLFQDGNGNWWAVA
Pf43B --TPTGSPIQLIDRSDLDGPLIEAP-----SLLLSNGIYYLSFSSN--------------
Fv43E LNDSPSGSQTWIWKSKSPWGPYESKVLADKVTPPISGGNSPHQGSLIKTPNGGWY-----

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 LATRITAQGVSPMGREAVLFNGTWNKGEWPKLQPVRGRMPGNLLPKPTRN------VPGD
Af43A LSTRSGPAWKNYPMGRETVLAPAAWEKGEWPVIQPVRGQMQG-PFPPPNKR----VPRGE
Pf43B YYNTNYYDTSYAYASSITGPWT-----------KQSAPYAPLLVTGT--------ETSND
Fv43E FMSFTWAYPAGRLPVLAPITWG-------------SDGFPILVKGANGGWGSSYPTLPGT

Fv43D SIVEYKGQWWLFFAD-AHTSGKDYLRQVKARKIWYDKDG-----KILLTRPKI-------
Fo43A SIVEYKGQWWLFFAD-AHTSGKDYLRQVKARKIWYDKNG-----KILLHRP---------
Gz43A SIVKYQGQWWLFYHD-AKTSGKDYLRQVKAKKIWYDSKG-----KILTKKP---------
Pf43A GIIDFQNNSYFFYHNGALPGGGGYQRSVCVEQFKYNADG-----TIPTIEMTTAG-----
Fv43A GIIDFENNSYFFYHNGALDGGSGYTRSVAVESFKYGSDG-----LIPEIKMTTQG-----
Fv43B GPFNADPDNYNLKKTKKIPPHFVHHRVPRDGAFSLSSKG----LHIVPSRNNVTGSVLPG
Af43A GGWIKQPDKVDFRPGSKIPAHFQYWRYPKTEDFTVSPRGHPNTLRLTPSFYNLTG-----
Pf43B GALSAPGGADFSVDGTKMLFHANLNGQDISGGRALFAAS------ITEASDVVTLQ----
Fv43E DGVTKNWTRTDTFRGTSLAPSWEWNHNPDVNSFTVNNGLTLRTASITKDIYQARN-----

Fv43D ------------------------------------------------------------
Fo43A ------------------------------------------------------------
Gz43A ------------------------------------------------------------
Pf43A ------------------PAQIGTLNPYVRQEAETAAWSSGITTEVCSEGGIDVGFINNG
Fv43A ------------------PAQLKSLNPYVKQEAETIAWSEGIETEVCSEGGLNVAFIDNG
Fv43B DEIELSGQRGLAFIGRRQTHTLFKYSVDIDFKPKSDDQEAGITVFRTQFDHIDLGIVRLP
Af43A -TADFKPDDGLSLVMRKQTDTLFTYTVDVSFDPKVADEEAGVTVFLTQQQHIDLGIVLLQ
Pf43B ------------------------------------------------------------
Fv43E ------------TLSHRTHGDHPTGIVKIDFSPMKDGDRAGLSAFRDQSAYIGIHRDNGK
```

FIG. 53B

```
Fv43D ------------------------------------------------------------
Fo43A ------------------------------------------------------------
Gz43A ------------------------------------------------------------
Pf43A DYIK----------VKGVAFGS-GAHSFSARVASANSGGTIAIHLGSTTGTLVGTCTV
Fv43A DYIK----------VKGVDFGSTGAKTFSARVASNSSGGKIELRLGSKTGKLVGTCTV
Fv43B TNQGSNKKSKLAFRFRATGAQNVPAPK---VVPVPDGWEKGVISLHIEAANATHYNLGAS
Af43A TTEG----LSLSFRFRVEGRGNYEGPLPEATVPVPKEWCGQTIRLEIQAVSDTEYVFAAA
Pf43B ------------------------------------------------------------
Fv43E FTIAT----KHGMNMDEWNGTTTDLGQIKATANVPSGRTKIWLRLQLDTNPAGTGNTIFS

Fv43D ------------------------------------------------------------
Fo43A ------------------------------------------------------------
Gz43A ------------------------------------------------------------
Pf43A PSTGGWQTWTTVTCSVSGASGTQ--------DVYFVFGGSGTGYLFN------FDYWQFA
Fv43A TTTGNWQTYKTVDCPVSGATGTS--------DLFFVFTGSGSGSLFN------FNWWQFS
Fv43B --SHRGKTLDIATASASLVSGGTGSFVGSLLGPYATCNGKGSGVECPKGGDVYVTQWTYK
Af43A PARHPAQRQIISRANSLIVSGDTGRFTGSLVGVYATSNG-GAGSTP-----AYISRWRYE
Pf43B ------------------------------------------------------------
Fv43E YSWDGVKYETLGPNFKLYNG----------WAFFIAYRFGIFNFAETALGGSIKVESFT

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

FIG. 53C

```
Pa51A   MIHLKPALAALLALSTQCVAIDLFVKSSGGNKTTDIMYGLMHEDINNSGDGGIYAELISN
Fv51A   MVRFSSILAAAACF-VAVESVNIKVDSKGGNATSGHQYGFLHEDINNSGDGGIYAELIRN
Pf51A   MGKMWHSILVVLGLLSVGHAITINVSQSGGNKTSPLQYGLMFEDINHGGDGGLYAELVRN

Pa51A   RAFQGSEKFPSNLDNWSPVGGATLTLQKLAKPLSSALPYSVNVANPKEGKGKGKDTKGKK
Fv51A   RAFQYSKKYPVSLSGWRPINDAKLSLNRLDTPLSDALPVSMNVK---PGKGK-----AKE
Pf51A   RAFQGSTVYPANLDGYDSVNGAILALQNLTNPLSPSMPSSLNVA-----KGS----NNGS

Pa51A   VGLANAGFWGMDVKRQKYTGSFHVTGEYKGDFEVSLRSAITGETFGKKVVKGGSKKGKWT
Fv51A   IGFLNEGYWGMDVKKQKYTGSFWVKGAYKGHFTASLRSNLTDDVFGSVKVKSKANKKQWV
Pf51A   IGFANEGWWGIEVKPQRYAGSFYVQGDYQGDFDISLQSKLTQEVFATAKVRSSGKHEDWV

Pa51A   EKEFELVPFKDAPNSNNTFVVQWDAEGAKDGSLDLNLISLFPPTFKGRKNGLRIDLAQTM
Fv51A   EHEFVLTPNKNAPNSNNTFAITYDPKGA-DGALDFNLISLFPPTYKGRKNGLRVDLAEAL
Pf51A   QYKYELVPKKAASNTNNTLTITFDSKGLKDGSLNFNLISLFPPTYNNRPNGLRIDLVEAM

Pa51A   VELKPTFLRFPGGNMLEGNTLDTWWKWYETIGPLKDRPGMAGVWEYQQTLGLGLVEYMEW
Fv51A   EGLHPSLLRFPGGNMLEGNTNKTWWDWKDTLGPLRNRPGFEGVWNYQQTHGLGILEYLQW
Pf51A   AELEGKFLRFPGGSDVEGVQAPYWYKWNETVGDLKDRYSRPSAWTYEESNGIGLIEYMNW

Pa51A   ADDMNLEPIVGVFAGLALDGSFVPESEMGWVIQQALDEIEFLTGDAKTTKWGAVRAKLGH
Fv51A   AEDMNLEIIVGVYAGLSLDGSVTPKDQLQPLIDDALDEIEFIRG-PVTSKWGKKRAELGH
Pf51A   CDDMGLEPILAVWDGHYLSNEVISENDLQPYIDDTLNQLEFLMG-APDTPYGSWRASLGY

Pa51A   PKPWKVKWVEIGNEDWLAGRPAGFESYINYRFPMMMKAFNEKYPDIKIIASPSIFD----
Fv51A   PKPFRLSYVEVGNEDWLAGYPTGWNSYKEYRFPMFLEAIKKAHPDLTVISSGASIDPVGK
Pf51A   PKPWTINYVEIGNEDNLYG---GLETYIAYRFQAYYDAITAKYPHMTVMESLTEMPG---

Pa51A   ---NMTIPAGAAGDHHPYLTPDEFVERFAKFDNLSKDNVTLIGEAASTHPNG---GIAWE
Fv51A   KDAGFDIPAPGIGDYHPYREPDVLVEEFNLFDNNKYG--HIIGEVASTHPNG---GTGWS
Pf51A   -------PAAAASDYHQYSTPDGFVSQFNYFDQMPVTNRTLNGEIATVYPNNPSNSVAWG

Pa51A   GDLMPLPWWGGSVAEAIFLISTERNGDKIIGATYAPGLRSLDRWQWSMTWVQHAADPALT
Fv51A   GNLMPYPWWISGVGEAVALCGYERNADRIPGTFYAPILKNENRWQWAITMIQFAADSAMT
Pf51A   SPFPLYPWWIGSVAEAVFLIGEERNSPKIIGASYAPMFRNINNWQWSPTLIAFDADSSRT

Pa51A   TRSTSWYVWRILAHHIIRETLPVDAPAGKPNFDPLFYVAGKSES-GTGIFKAAVYNSTES
Fv51A   TRSTSWYVWSLFAGHPMTHTLPTTA-----DFDPLYYVAGKNEDKGTLIWKGAAYNTTKG
Pf51A   SRSTSWHVIKLLSTNKITQNLPTTWSGG--DIGPLYWVAGRNDNTGSNIFKAAVYNSTSD

Pa51A   --IPVSLKFDGLNEGAVANLTVLTGPE-DFYGYNDPFTGINVVKEKTTFIKAGKGGKFTF
Fv51A   ADVPVSLSFKGVKPGAQAELTLLTNKEKDPFAFNDPHKGNNVVDTKKTVLKADGKGAFNF
Pf51A   --VPVTVQFAGCN-AKSANLTILSSDD--PNASNYPG-GPEVVKTEIQSVTANAHGAFEF

Pa51A   TLPGLSVAVLETADAVKGGKGKGKGKGKGN
Fv51A   KLPNLSVAVLETLK---------KGKPYSS
Pf51A   SLPNLSVAVLKTE----------------
```

FIG. 54

```
                 *         20          *         40          *
xyn3    : -MKANVILC--LLAPLVAALPTETIHLDPELAALRANLTERTADLWDRQA :  47
P56588  : ----------------------------------------------QA :   2
P23360  : MVRPTILLTSLLLAPFAAASPI----------------------LEERQA :  28

                60          *         80          *         100
xyn3    : SQSIDQLIKRKGKLYFGTATDRGLLQRE-KNAAIIQADLGQVTPENSMKW :  96
P56588  : SVSIDAKFKAHGKKYLGTIGDQYTLTKNTKNPAIIKADFGQLTPENSMKW :  52
P23360  : AQSVDQLIKARGKVYFGVATDQNRLTTG-KNAAIIQADFGQVTPENSMKW :  77

                  *        120          *        140          *
xyn3    : QSLENNQGQLNWGDADYLVNFAQQNGKSIRGHTLIWHSQLPAWVNNINNA : 146
P56588  : DATEPNRGQFTFSGSDYLVNFAQSNGKLIRGHTLVWHSQLPGWVSSITDK : 102
P23360  : DATEPSQGNFNFAGADYLVNWAQQNGKLIRGHTLVWHSQLPSWVSSITDK : 127

                160          *        180          *        200
xyn3    : DTLRQVIRTHVSTVVGRYKGKIRAWDVVNEIFNEDGTLRSSVFSRLLGEE : 196
P56588  : NTLISVLKNHITTVMTRYKGKIYAWDVLNEIFNEDGSLRNSVFYNVIGED : 152
P23360  : NTLTNVMKNHITTLMTRYKGKIRAWDVVNEAFNEDGSLRQIVFLNVIGED : 177

                  *        220          *        240          *
xyn3    : FVSIAFRAARDADPSARLYINDYNLDRANYGKVNGLKTYVSKWISQGVPI : 246
P56588  : YVRIAFETARSVDPNAKLYINDYNLDSAGYSKVNGMVSHVKKWLAAGIPI : 202
P23360  : YIPIAFQTARAADPNAKLYINDYNLDSASYPKTQAIVNRVKQWRAAGVPI : 227

                160          *        180     N    *        300
xyn3    : DGIGSQSHLSGGGGSGTLGALQQLATVPVTELAITELDIQGAPTTDYTQV : 296
P56588  : DGIGSQTHLGAGAGSAVAGALNALASAGTKEIAITELDIAGASSTDYVNV : 252
P23360  : DGIGSQTHLSAGQGAGVLQALPLLASAGTPEVAITELDVAGASPTDYVNV : 277

                  *        320          *        340          *
xyn3    : VQACLSVSKCVGITVWGISDKDSWRASTNPLLFDANFNPKPAYNSIVGIL : 346
P56588  : VNACLNQAKCVGITVWGVADPDSWRSSSSPLLFDGNYNPKAAYNAIANAL : 302
P23360  : VNACLNVQSCVGITVWGVADPDSWRASTTPLLFDGNFNPKPAYNAIVQDL : 327


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

FIG. 55A

```
  1  MVSFSYLLLACSAIG-ALAAPVEPETTSFN  AfuXyn2
  1  MISISSLSFGLAAIAGAYALPSD-KSVSLA  AfuXyn5
  1  MVSFTSLLAASPPSRASCRPAAEVESVAVE  Xyn2

 30  ETALHEFAERAGTPSSTGWNNGYYYSFWTD  AfuXyn2
 30  ER--------QTITTSQTGTNNGYYYSFWTN  AfuXyn5
 31  KR--------QTIQPG-TGYNNGYFYSYWND  Xyn2

 60  GGGDVTYTNGAGGSYSVNWRN--VGNFVGG  AfuXyn2
 53  GAGSVQYTNGAGGEYSVTWANQNGGDFTCG  AfuXyn5
 53  GHGGVTYTNGPGGQFSVNWSN--SGNFVGG  Xyn2

 88  KGWNPGSA-RTINYGGSFNPSGNGYLAVYG  AfuXyn2
 83  KGWNPGSD-HDITFSGSFNPSGNAYLSVYG  AfuXyn5
 81  KGWQPGTKNKVINFSGSYNPNGNSYLSVYG  Xyn2
               N
117  WTTNPLIEYYVVESYGTYNPGSGGTFRGTV  AfuXyn2
112  WTTNPLVEYYILENYGSYNPGSGMTHKGTV  AfuXyn5
111  WSRNPLIEYYIVENFGTYNPSTGATKLGEV  Xyn2

147  NTDGGTYNIYTAVRYNAPSIEGTKTFTQYW  AfuXyn2
142  TSDGSTYDIYEHQQVNQPSIVGTATFNQYW  AfuXyn5
141  TSDGSVYDIYRTQRVNQPSIIGTATFYQYW  Xyn2

177  SVRTSKRTGGTVTMANHFNAWSRLGMNLGT  AfuXyn2
172  SIRQNKRSSGTVTTANHFKAWASLGMNLGT  AfuXyn5
171  SVRRNHRSSGSVNTANHFNAWAQQGLTLGT  Xyn2
               A
207  HNYQIVATEGYQSSGSASITVY          AfuXyn2
202  HNYQIVSTEGYESSGTSTITVSSGGSSSGG  AfuXyn5
201  MDYQIVAVEGYFSSGSASITVS          Xyn2

228                                  AfuXyn2
232  SGGSSSTTSSGSSPTGGSGSCSALWGQCGG  AfuXyn5
222                                  Xyn2

228                                  AfuXyn2
262  IG-WSGPTCCSSGTCQVSNSYYSQCL      AfuXyn5
222                                  Xyn2
```

FIG. 55B

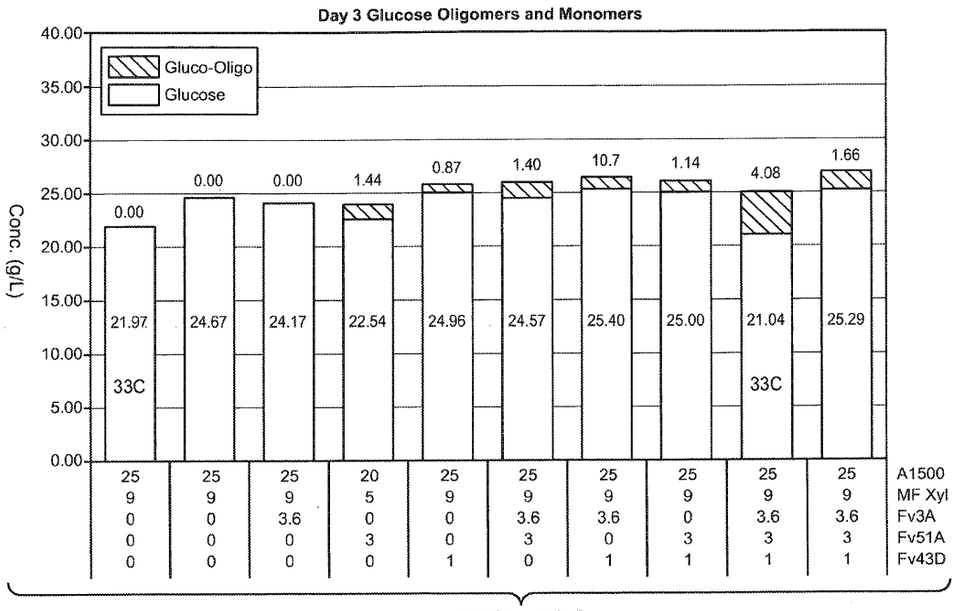

Day 3 Glucose Oligomers and Monomers
Gluco-Oligo
Glucose
Conc. (g/L)
40.00
35.00
30.00
25.00
20.00
15.00
10.00
5.00
0.00
0.00
0.00
0.00
1.44
0.87
1.40
10.7
1.14
4.08
1.66
21.97
24.67
24.17
22.54
24.96
24.57
25.40
25.00
21.04
25.29
33C
33C
25 25 25 20 25 25 25 25 25 25 A1500
9 9 9 5 9 9 9 9 9 9 MF Xyl
0 0 3.6 0 0 3.6 3.6 0 3.6 3.6 Fv3A
0 0 0 3 0 3 0 3 3 3 Fv51A
0 0 0 0 1 0 1 1 1 1 Fv43D

FIG. 56A

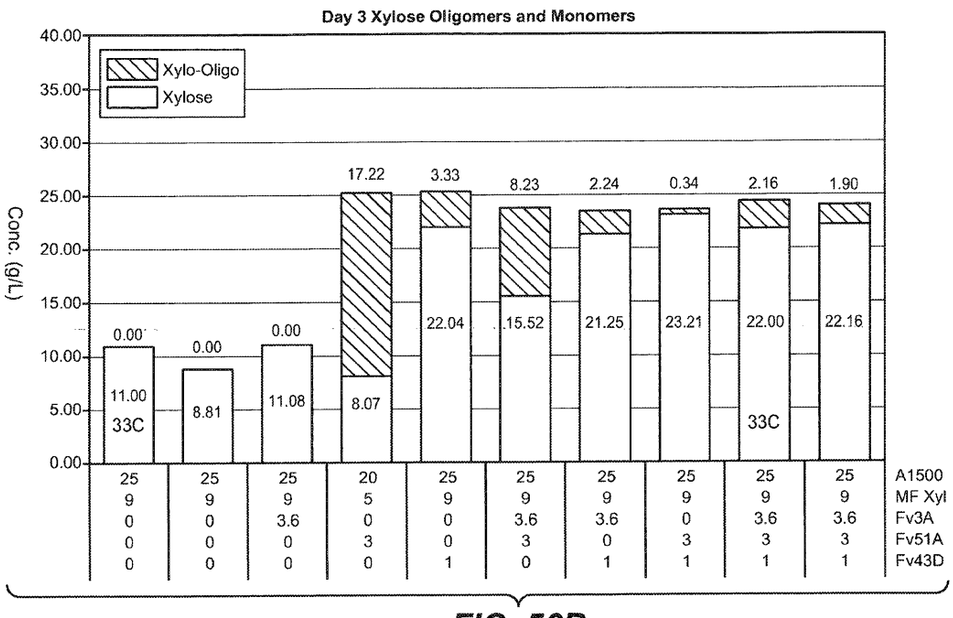

Day 3 Xylose Oligomers and Monomers
Xylo-Oligo
Xylose
Conc. (g/L)
40.00
35.00
30.00
25.00
20.00
15.00
10.00
5.00
0.00
0.00
0.00
0.00
17.22
3.33
8.23
2.24
0.34
2.16
1.90
11.00
33C
8.81
11.08
8.07
22.04
15.52
21.25
23.21
22.00
33C
22.16
25 9 0 0 0 A1500
25 9 0 0 0 MF Xyl
25 9 3.6 0 0 Fv3A
20 5 0 3 0 Fv51A
25 9 0 0 1 Fv43D
25 9 3.6 3 0
25 9 3.6 0 1
25 9 0 3 1
25 9 3.6 3 1
25 9 3.6 3 1

FIG. 56B

484

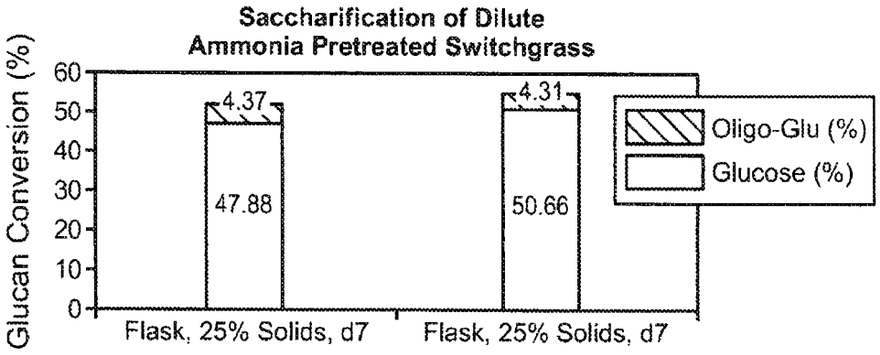


FIG. 57A

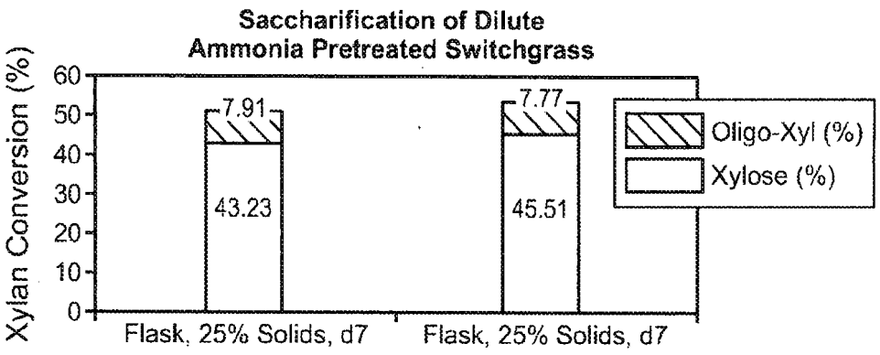


FIG. 57B

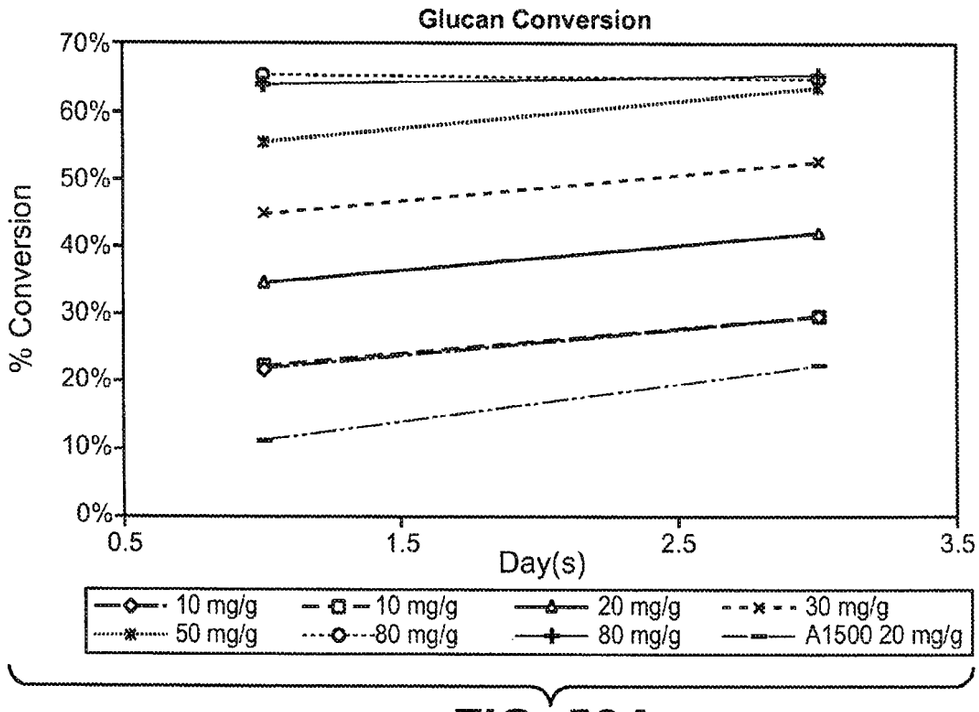


FIG. 58A

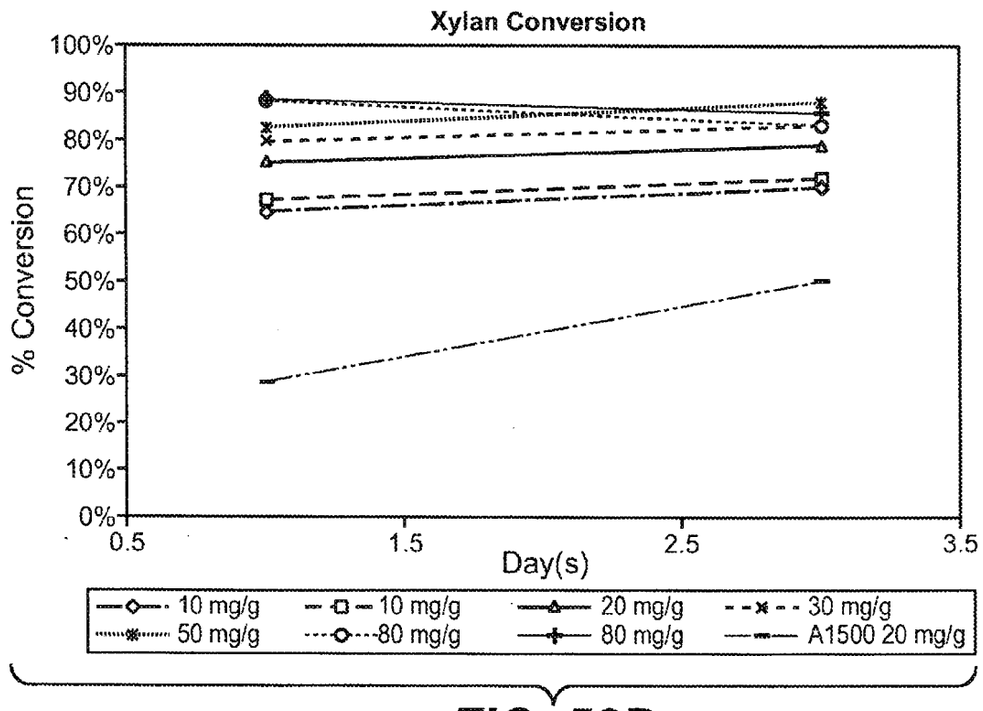


FIG. 58B

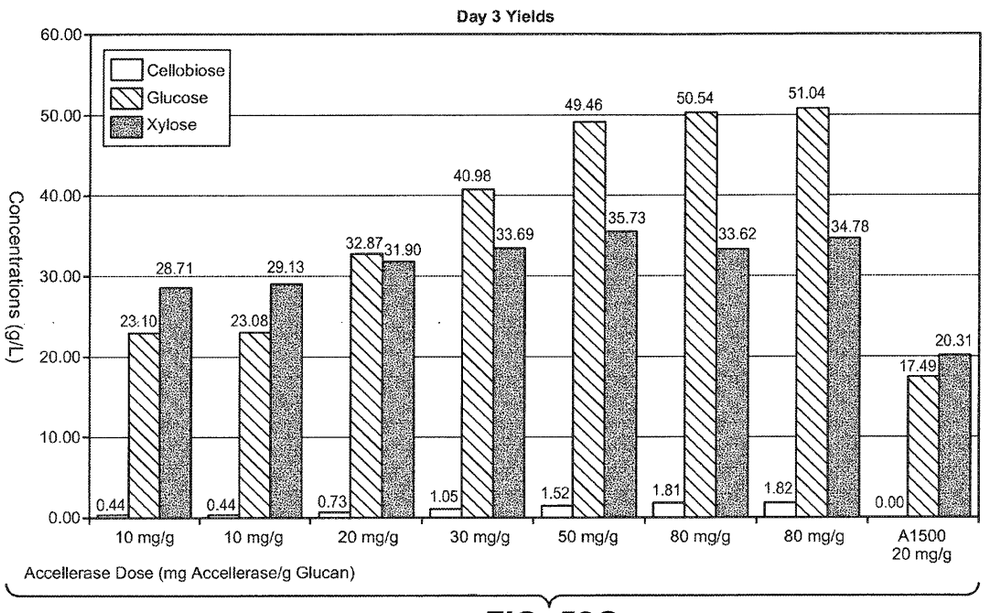

Day 3 Yields
Cellobiose
Glucose
Xylose
Concentrations (g/L)
60.00
50.00
40.00
30.00
20.00
10.00
0.00
0.44
23.10
28.71
0.44
23.08
29.13
0.73
32.87
31.90
1.05
40.98
33.69
1.52
49.46
35.73
1.81
50.54
33.62
1.82
51.04
34.78
0.00
17.49
20.31
10 mg/g
10 mg/g
20 mg/g
30 mg/g
50 mg/g
80 mg/g
80 mg/g
A1500
20 mg/g
Accellerase Dose (mg Accellerase/g Glucan)

FIG. 58C

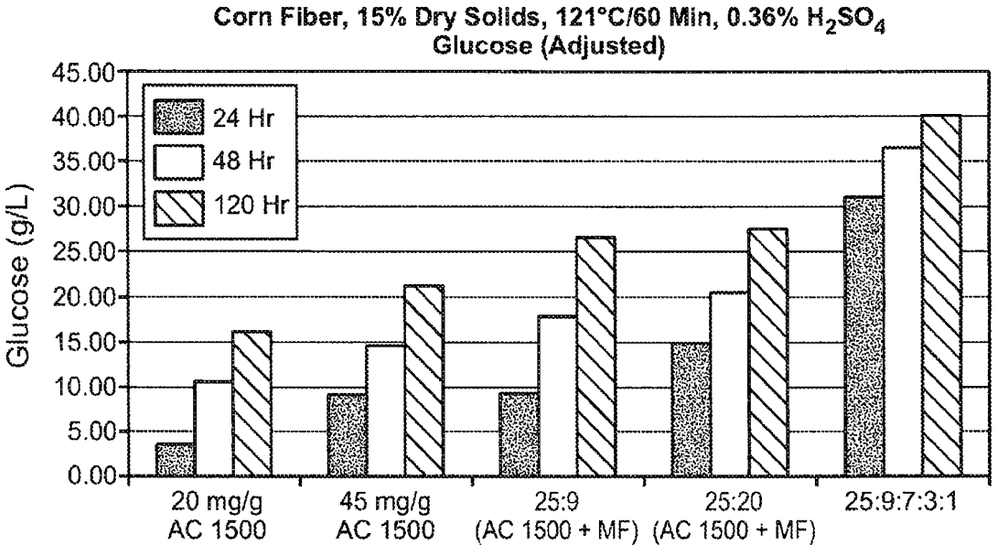


FIG. 59A

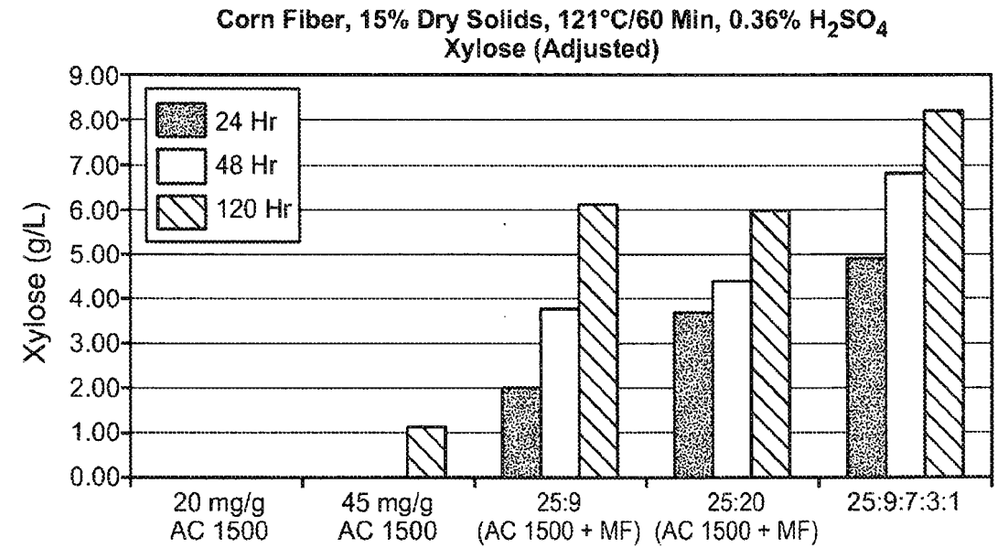


FIG. 59B

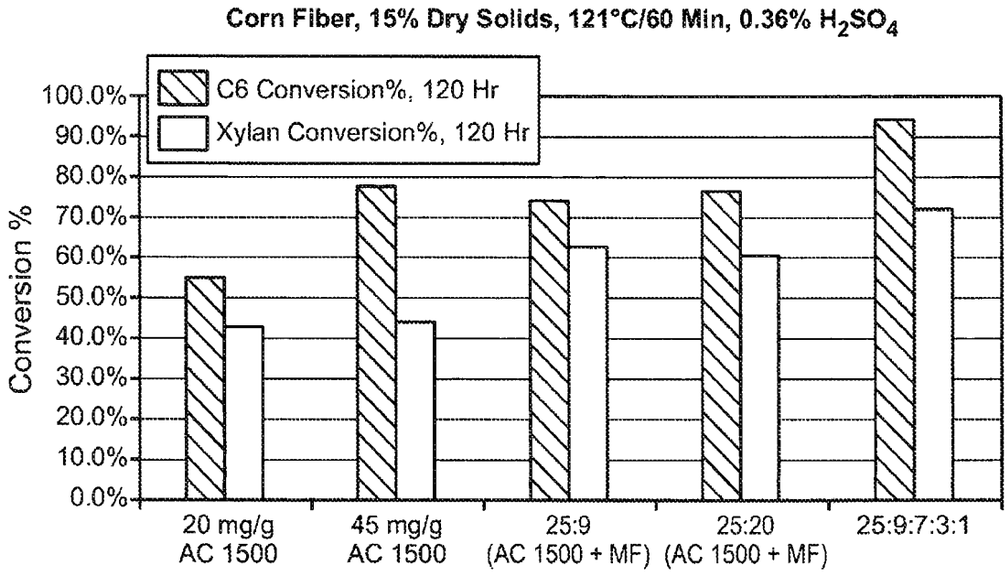

Corn Fiber, 15% Dry Solids, 121°C/60 Min, 0.36% H_2SO_4
C6 Conversion%, 120 Hr
Xylan Conversion%, 120 Hr
Conversion %
100.0%
90.0%
80.0%
70.0%
60.0%
50.0%
40.0%
30.0%
20.0%
10.0%
0.0%
20 mg/g AC 1500
45 mg/g AC 1500
25:9 (AC 1500 + MF)
25:20 (AC 1500 + MF)
25:9:7:3:1

FIG. 59C

SUBSTITUTE SHEET (RULE 26)

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

atgatccacctcaagccagccctcgcggcgttgttggcgctgtcgacgcaatgtgtggctattgatttgtttgtcaagtcttcggg
ggggaataagacgactgatatcatgtatggtcttatgcacgaggatatcaacaactccggcgacggcggcatctacgccgagctaa
tctccaaccgcgcgcgttccaagggagtgagaagttcccctccaacctcgacaactggagccccgtcggtggcgctacccttaccctt
cagaagcttgccaagccccttcctctgcgttgccttactccgtcaatgttgccaaccccaaggagggcaagggcaagggcaagga
caccaagggaagaaggttggcttggccaatgctgggttttggggtatggatgtcaagaggcagaagtacactggtagcttccacg
ttactggtgagtacaagggtgactttgaggttagcttgcgcagcgcgattaccggggagacctttggcaagaaggtggtgaagggt
gggagtaagaaggggaagtggaccgagaaggagtttgagttggtgcctttcaaggatgcgcccaacagcaacaacacctttgttgt
gcagtgggatgccgagggcgcaaaggacggatctttggatctcaacttgatcagcttgttccctccgacattcaagggaaggaaga
atgggctgagaattgatcttgcgcagacgatggttgagctcaagccgaccttcttgcgcttccccggtggcaacatgctcgagggt
aacaccttggacacttggtggaagtggtacgagaccattggccctctgaaggatcgcccgggcatggctggtgtctgggagtacca
gcaaacccttggcttgggtctggtcgagtacatggagtgggccgatgacatgaacttggagcccattgtcggtgtcttcgctggtc
ttgccctcgatggctcgttcgttcccgaatccgagatgggatgggtcatccaacaggctctcgacgaaatcgagttcctcactggc
gatgctaagaccaccaaatggggtgccgtccgcgcgaagcttggtcaccccaagccttggaaggtcaagtgggttgagatcggtaa
cgaggattggcttgccggacgccctgctggcttcgagtcgtacatcaactaccgcttcccatgatgatgaaggccttcaacgaaa
agtaccccgacatcaagatcatcgcctcgccctccatcttcgacaacatgacaatccccgcgggtgctgccggtgatcaccacccg
tacctgactcccgatgagttcgttgagcgattcgccaagttcgataacttgagcaaggataacgtgacgctcatcggcgaggctgc
gtcgacgcatcctaacggtggtatcgcttgggagggagatctcatgcccttgccttggtggggcggcagtgttgctgaggctatct
tcttgatcagcactgagagaaacggtgacaagatcatcggtgctacttacgcgcctggtcttcgcagcttggaccgctggcaatgg
agcatgacctgggtgcagcatgccgccgacccggccctcaccactcgctcgaccagttggtatgtctggagaatcctcgcccacca
catcatccgtgagacgctcccggtcgatgcccggccggcaagcccaactttgaccctctgttctacgttgccggaaagagcgaga
gtggcaccggtatcttcaaggctgccgtctacaactcgactgaatcgatcccggtgtcgttgaagtttgatggtctcaacgaggga
gcggttgccaacttgacggtgcttactgggccggaggatccgtatggatacaacgacccccttcactggtatcaatgttgtcaagga
gaagaccaccttcatcaaggccggaaagggcggcaagttcaccttcaccctgccgggcttgagtgttgctgtgttggagacggccg
acgcggtcaagggtggcaagggaaagggcaagggcaagggaaagggtaactga

FIG. 60A

490

SEQ ID NO:47
Codon optimized cDNA for Pa51A, a GH51 family enzyme from *Podospora anserina*

atgatccacctcaagcccgccctcgccgccctcctcgccctcagcacccaatgcgtcgccatcgacctct
tcgtcaagagcagcggcggcaacaagaccaccgacatcatgtacggcctcatgcacgaggacatcaacaa
cagcggcgacggcggcatctacgccgagctgatcagcaaccgcgccttccagggcagcgagaagttcccc
agcaacctcgacaactggtccccgtcggcggcgccaccctcaccctccagaagctcgccaagcccctgt
cctctgccctcccctactccgtcaacgtcgccaaccccaaggagggtaagggtaagggcaaggacaccaa
gggcaagaaggtcggcctcgccaacgccggcttttggggcatggacgtcaagcgccagaaatacaccggc
agcttccacgtcaccggcgagtacaagggcgacttcgaggtcagcctccgcagcgccattaccggcgaga
ccttcggcaagaaggtcgtcaagggcggcagcaagaagggcaagtggaccgagaaggagttcgagctggt
ccccttcaaggacgcccccaacagcaacaacaccttcgtcgtccagtgggacgccgagggcgccaaggac
ggcagcctcgacctcaacctcatcagcctcttcccgcccaccttcaagggccgcaagaacggcctccgca
tcgacctcgcccagaccatggtcgagctgaagcccaccttcctccgctttcccggcggcaacatgctcga
gggcaacaccctcgacacctggtggaagtggtacgagaccatcggccccctgaaggaccgccctggcatg
gccggcgtctgggagtaccagcagacgctgggcctcggcctggtcgagtacatggagtgggccgacgaca
tgaacctcgagcccatcgtcggcgtctttgctggcctggccctggatggcagctttgtccccgagagcga
gatgggctgggtcatccagcaggctctcgatgagatcgagttcctcaccggcgacgccaagaccaccaag
tggggcgccgtccgcgccaagctcggccaccctaagccctggaaggtcaaatgggtcgagatcggcaacg
aggactggctcgccggccgacctgccggcttcgagagctacatcaactaccgcttccccatgatgatgaa
ggccttcaacgagaaataccccgacatcaagatcattgccagcccctccatcttcgacaacatgaccatt
ccagccggtgctgccggtgaccaccaccctacctcaccccgacgaatttgtcgagcgcttcgccaagt
tcgacaacctcagcaaggacaacgtcaccctcattggcgaggccgccagcaccccaccccaacggcggcat
tgcctgggagggcgacctcatgccctgccctggtggggcggcagcgtcgccgaggccatcttcctcatc
agcaccgagcgcaacggcgacaagatcatcggcgccacctacgccctggcctccgatctctcgaccgct
ggcagtggagcatgacctgggtccagcacgccgccgaccctgccctcaccaccccgcagcaccagctggta
cgtctggcgcatcctcgcccaccacatcattcgcgagaccctccccgtcgacgcccccgccggcaagccc
aacttcgacccccctcttctacgtcgctggcaagtcggagagcggcaccggcatcttcaaggccgccgtct
acaacagcaccgagagcatccccgtcagcctcaagttcgacggcctcaacgagggcgccgtcgccaacct
caccgtcctcaccggcccgaggacccctacggctacaacgacccccttcaccggcatcaacgtcgtcaag
gaaaagaccaccttcatcaaggccggcaagggcggcaagttcacctttaccctccccggcctctctgtcg
ccgtcctcgagaccgccgacgccgtgaagggtggcaagggaaagggaaagggcaagggtaagggtaacta
a

FIG. 60B

491

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

atgtatcggaagttggccgtcatctcggccttcttggccacagctcgtgctaccaacgacgactgtcctc
tcatcactagtagatggactgcggatccttcggctcatgtctttaacgacaccttgtggctctacccgtc
tcatgacatcgatgctggatttgagaatgatcctgatggaggccagtacgccatgagagattaccatgtc
tactctatcgacaagatctacggttccctgccggtcgatcacggtacggccctgtcagtggaggatgtcc
cctgggcctctcgacagatgtgggctcctgacgctgcccacaagaacggcaaatactacctatacttccc
tgccaaagacaaggatgatatcttcagaatcggcgttgctgtctcaccaacccccggcggaccattcgtc
cccgacaagagttggatccctcacactttcagcatcgaccccgccagtttcgtcgatgatgatgacagag
cctacttggcatggggtggtatcatgggtggccagcttcaacgatggcaggataagaacaagtacaacga
atctggcactgagccaggaaacggcaccgctgccttgagccctcagattgccaagctgagcaaggacatg
cacactctggcagagaagcctcgcgacatgctcattcttgaccccaagactggcaagccgctcctttctg
aggatgaagaccgacgcttcttcgaaggaccctggattcacaagcgcaacaagatttactacctcaccta
ctctactggcacaacccactatcttgtctatgcgacttcaaagacccccctatggtccttacacctaccag
ggcagaattctggagccagttgatggctggactactcactctagtatcgtcaagtaccagggtcagtggt
ggctattttatcacgatgccaagacatctggcaaggactatcttcgccaggtaaaggctaagaagatttg
gtacgatagcaaaggaaagatcttgacaaagaagccttga

FIG. 61

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

atgtatcggaagttggccgtcatctcggccttcttggccacagctcgtgctcaagacactaatgacattc
ctcccctgatcaccgacctctggtccgcagatcctcggctcatgttttcgaaggcaagctctgggtttta
cccatctcacgacatcgaagccaatgttgtcaacggcacaggaggcgctcaatacgccatgagggattac
catacctactccatgaagagcatctatggtaaagatcccgttgtcgaccacggcgtcgctctctcagtcg
atgacgttccctgggcgaagcagcaaatgtgggctcctgacgcagctcataagaacggcaaatattatct
gtacttccccgccaaggacaaggatgagatcttcagaattggagttgctgtctccaacaagcccagcggt
cctttcaaggccgacaagagctggatccctggcacgtacagtatcgatcctgctagctacgtcgacactg
ataacgaggcctacctcatctgggggcggtatctgggggcggccagctccaagcctggcaggataaaaagaa
ctttaacgagtcgtggattggagacaaggctgctcctaacggcaccaatgccctatctcctcagatcgcc
aagctaagcaaggacatgcacaagatcaccgaaacaccccgcgatctcgtcattctcgccccgagacag
gcaagcctcttcaggctgaggacaacaagcgacgattcttcgagggcccttggatccacaagcgcggcaa
gctttactacctcatgtactccaccggtgataccacttccttgtctacgctacttccaagaacatctac
ggtccttatacctaccggggcaagattcttgatcctgttgatgggtggactactcatggaagtattgttg
agtataagggacagtggtggcttttctttgctgatgcgcatacgtctggtaaggattaccttcgacaggt
gaaggcgaggaagatctggtatgacaagaacggcaagatcttgcttcaccgtccttag

FIG. 62

SEQ ID NO:50
Nucleotide sequence for Pf51A, a GH51 family enzyme from *Penicillium funiculosum*

atgtaccggaagctcgccgtgatcagcgccttcctggcgactgctcgcgccatcaccatcaacgtcagcc
agagcggcggcaacaagaccagcccgctccagtacggcctcatgttcgaggacatcaaccacggcggcga
cggcggcctctacgccgagctggtccggaaccgggccttccagggcagcaccgtctacccggccaacctc
gacggctacgactcggtgaacggcgcgattctcgcgctccagaacctcaccaacccgctcagcccgagca
tgccctcgtcgctgaacgtcgccaagggctcgaacaacggcagcatcggcttcgccaacgaggggtggtg
gggcatcgaggtcaagccgcagcggtacgccggcagcttctacgtccagggcgactaccagggcgacttc
gacatcagcctccagagcaagctcacccaggaggtcttcgcgacggcgaaggtccggtcgagcggcaagc
acgaggactgggtccagtacaagtacgagctggtcccgaagaaggccgccagcaacaccaacaacaccct
caccatcaccttcgacagcaagggcctcaaggacggcagcctcaacttcaacctcatcagcctcttcccg
ccgacctacaacaaccggccgaacggcctccggatcgacctcgtcgaggccatggcggagctggagggca
agttcctccgcttccccggcggctcggacgtggagggcgtccaggccccgtactggtacaagtggaacga
gaccgtcggcgacctcaaggaccgctactcgcgcccgagcgcctggacctacgaggagagcaacggcatc
ggcctcatcgagtacatgaactggtgcgacgacatgggcctcgagccgatcctcgccgtctgggacggcc
actacctcagcaacgaggtcatcagcgagaacgacctccagccgtacatcgacgacaccctcaaccagct
cgagttcctcatgggcgccccggacactccctacgggtcttggagggctagcctcggctacccgaagccg
tggaccatcaactacgtcgagatcggcaacgaggacaacctctacggcggcctcgagacctacatcgcct
accggttccaggcctactacgacgccatcaccgccaagtacccgcacatgaccgtcatggagagcctcac
cgagatgcccggccccgctgccgcggcgtcggactaccaccagtactcgacgcccgacggcttcgtcagc
cagttcaactacttcgaccagatgccggtcaccaaccgcacgctgaacggcgagatcgccaccgtctacc
ccaacaacccgagcaactcggtggcgtggggcagcccgttcccgctctacccgtggtggatcgggtccgt
ggctgaggccgtcttcctcatcggcgaggagcggaacagcccgaagatcatcggcgccagctacgccccc
atgttccgcaacattaacaactggcagtggagcccgaccctgatcgccttcgacgccgacagcagccgga
cgtcgcgctctacttcctggcacgtcatcaagctcctcagcaccaacaagatcacccagaacctgcccac
gacgtggtctgggggggacatcggcccgctctactgggtcgccggccggaacgacaacaccggcagcaac
atcttcaaggccgccgtctacaacagcaccagcgacgtcccggtcaccgtccagttcgccggctgcaacg
ccaagagcgccaacctcaccatcctctcgtcggacgaccccaacgccagcaactacccgggcggccccga
ggtcgtcaagaccgagatccagagcgtcacgccaacgcccacggcgccttcgagttcagcctcccgaac
ctgtcggtggctgtgctgaagacggagtag

FIG. 63

```
Fv3A    MLLNLQVAASALSLSLLGGLAEAATPYT----------LPDCTKGPLSK
Bxl1    -MVNNAALLAALSALLPTALAQNNQTYANYSAQGQPDLYPETLATLTLSFPDCEHGPLKN
Bgl1    MRY-RTAAALALATG------PFARADSHSTSG---ASAEAVVP----------

Fv3A    NGICDTSLSPAKRAAALVAALTPEEKVGNLVSNATGAPRIGLPRYNWWNEALHGLAGSPG
Bxl1    NLVCDSSAGYVERAQALISLFTLEELILNTQNSGPGVPRLGLPNYQVWNEALHGLDRAN-
Bgl1    -PAGTPWGTAYDKAKAALAKLNLQDKVGIVSGVGWNGGPCVGNTSPASKISYPSLCLQDG

Fv3A    GRFADTP-PYDAATSFPMPLLMAAAFDDDLIHDIGNVVGTEARAFTNGGWRGVDFWTPNV
Bxl1    --FATKGGQFEWATSFPMPILTTAALNRTLIHQIADIISTQARAFSNSGRYGLDVYAPNV
Bgl1    ---PLGVRYSTGSTAFTPGVQAASTWDVNLIRERGQFIGEEVKASGIHVILGPVAGPLGK

Fv3A    NPFKDPRWGRGSETPGEDALHVS-RYARYIVRGLEG--DKEQRRIVATCKHYAGNDFEDW
Bxl1    NGFRSPLWGRGQETPGEDAFFLSSAYTYEYITGIQGGVDPEHLKVAATVKHFAGYDLENW
Bgl1    TPQGGRNWEGF---GVDP--YLTGIAMGQTINGIQS------VGVQATAKHYILNEQELN

Fv3A    GGFTR---------HDFDAKITPQDLAEYYVRPFQECTRDAKVGSIMCAYNAVNGIPACA
Bxl1    NNQSR---------LGFDAIITQQDLSEYYTPQFLAAARYAKSRSLMCAYNSVNGVPSCA
Bgl1    R-------------ETISSNPDDRTLHELYTWPFADAVQAN-VASVMCSYNKVNTTWACE

Fv3A    NSYLQETILRGHWNWTRDNNWITSDCGAMQDIWQNHKYVKTNAEGAQVAFENGMDSSCEY
Bxl1    NSFFLQTLLRESWGFP-EWGYVSSDCDAVYNVFNPHDYASNQSSAAASSLRAGTDIDCGQ
Bgl1    DQYTLQTVLKDQLGFP---GYVMTDWNAQHTTVQSANSGLDMSMPG-TDFNGNNRLWGPA

Fv3A    TTTSDVSDSYKQGLLTEKLMDRSLKRLFEGLVHTGFFDGAKAQWN----------
Bxl1    TYPWHLNESFVAGEVSRGEIERSVTRLYANLVRLGYFD-KKNQYR----------
Bgl1    LTNAVNSNQVPTSRVDDMVTRILAAWYLTGQDQAGYPSFNIS----------

Fv3A    -SLSFADVNTKEAQDLALRSAVEGAVLLKNDG-TLPLKLKKKDSVAMIGFWAN-------
Bxl1    -SLGWKDVVKTDAWNISYEAAVEGIVLLKNDG-TLPLSKKVR-SIALIGPWAN-------
Bgl1    ------RNVQGNHKTNVRAIARDGIVLLKNDANILPLKKPASIAVVGSAAIIGNHARNSP

Fv3A    ----------DTSKLQGGYSGRAPFLHSPLYAAEKLGLDTNVAWGPTLQNSSSHDNWTTN
Bxl1    ----------ATTQMQGNYYGPAPYLISPLEAAKKAGYHVNFELGTEIAGNSTTG--FAK
Bgl1    SCNDKGCDDGALGMGWGSGAVNYPYFVAPYDAINTRASSQGT--QVTLSNTDNTS---SG

Fv3A    AVAAAKKSDYILYFGGLDAS----AAGEDRDRENLDWPESQLTLLQKLSSLGKPLVVIQL
Bxl1    AIAAAKKSDAIIYLGGIDNT----IEQEGADRTDIAWPGNQLDLIKQLSEVGKPLVVLQM
Bgl1    ASAARGKDVAIVFITADSGEGYITVEGNAGDRNNLDPWHNGNALVQAVAGANSNVIVVVH

Fv3A    G-DQVDDTALLKNKKINSILWVNYPGQDGGTAVMDLLTGRKSPAGRLPVTQYP-------
Bxl1    GGGQVDSSSLKSNKKVNSLVWGGYPGQSGGVALFDILSGKRAPAGRLVTTQYP-------
Bgl1    SVGAIILEQILALPQVKAVVWAGLPSQESGNALVDVLWGDVSPSGKLVYTIAKSPNDYNT
```

FIG. 64A

```
Fv3A    --------------SKYTEQIGMTDMDLRPT-KSLPGRTYRWYS-TPVLPYGFGLHYTKF
Bxl1    --------------AEYVHQFPQNDMNLRPDGKSNPGQTYIWYTGKPVYEFGSGLFYTTF
Bgl1    RIVSGGS-------DSFSEGLFIDYKHFDDAN------------ITPRYEFGYGLSYTKF

Fv3A    QAKFKSN--KLTFDIQKLLKG--CSAQYS-------------------------------
Bxl1    KETLASHPKSLKFNTSSILSAPHPGYTYS-------------------------------
Bgl1    NYSRLSVLSTAKS--GPATGAVVP------------------------------------

Fv3A    ----------------------------------------DTCALPPIQVSVKNTGR
Bxl1    ----------------------------------------EQIPVFTFEANIKNSGK
Bgl1    ------------------------------------GGPSDLFQNVATVTVDIANSGQ

Fv3A    ITSDFVSLVFIKS-EVGPKPYPLKTLAAYGRLHDVAPSSTKDISLEWTLDNIARRGENGD
Bxl1    TESPYTAMLFVRTSNAGPAPYPNKWLVGFDRLADIKPGHSSKLSIPIPVSALARVDSHGN
Bgl1    VTGAEVAQLYITYPSSAPR-TPPKQLRGFAKLN-LTPGQSGTATFNIRRRDLSYWDTASQ

Fv3A    LVVYPGTYTLLLDEPTQAKIQVTLTGKKAILDKWPQDPKSA-------
Bxl1    RIVYPGKYELALNTDESVKLEFELVGEEVTIENWPLEEQQIKDATPDA
Bgl1    KWVVPSGSFGISVGASSRDIRLTSTLSVA-------------------
```

FIG. 64B

SUBSTITUTE SHEET (RULE 26)

```
Fv30B.pro MNPLSLGLAALSLLGYVGVNFVAAFPTDSNSGSEVLISVNGHVKHQELDGFGASQAFQRAEDILGKDGLSKEGTQHVLDL   80
Fv30A.pro ML-FSLVLPTLA--------FQASLALGDTS-----VTVDTSQKLQVIDGFGVSEAYGHAKQF---QNLGPGPQKEGLDL   63

Fv30B.pro LFSKDIGAGFSILRNGIGSSNSSDKNFMNSIEPFSPGSPGAKPHYVWDGYDSGQLTVAQEAFKRGLKFLYGDAWSAPGYM  160
Fv30A.pro LFNTTTGAGLSIIRNKIGCDAS------NSITSTNTDNPDKQAVYHFDGDDDGQ--------------------------  111

Fv30B.pro KTNHDENNGGYLCGVTGAACASGDWKQAYADYLLQWVEFYRKSGVKVTNLGFLNEPQFAAPYAGMLSNGTQAADFIRVLG  240
Fv30A.pro ----SAQSMGRLCGTPGVSCSSGDWRHRYVEMIAEYLSYYKQAGIPVSHVGFLNEGD-GSDF--MLSTAEQAADVIPLLH  184

Fv30B.pro KTIRKRGIHDLTIACCDGEGWDLQ---EDMMAGLTAGPDPAINYLSVVTGHGYVSPPNHPLSTTKKTWLTEWADLTGQFT  317
Fv30A.pro SALQSKGLGDIKMTCCDNIGWKSQMDYTAKLAELEVE-----KYLSVITSHEYSSSPNQPMNTTLPTWMSEGAANDQAFA  259

Fv30B.pro PYTFYNNSGQGEGMTWAGRIQTALVDANVSGFLYWIGAENSTTNS--ALINMIGDKVIPSKRFWAFASFSRFARPGARRI  395
Fv30A.pro T-AWYVNGGSNEGFTWAVKIAQGIVNADLSAYIYWEGVETNNKGSLSHVIDTDGTKFTISSILWAIAHWSRHIRPGAHRL  338

Fv30B.pro EATSSVPLVTVSSFLNTDGTVATQVLNNDTVAHSVQLVVSGTG-----------------------------------  438
Fv30A.pro STSGVVQDTIVGAFENVDGSVVMVLTNSGTAAQTVDLGVSGSSFSTAQAFTSDAEAQMVDTKVTLSDGRVKVTVPVHGVV  418

Fv30B.pro ----------------------RNPHSLKPFLTDNSNDLTALKHLKATGKGSFQTTIPPRSLVSFVT--------------  483
Fv30A.pro TVKLTTAKSSKPVSTAVSAQSAPTPTSVKHTLTHQKTSSTTLSTAKAPTSTQTTSVVESAKAVKYPVPPVASKGSSKSAP  498

Fv30B.pro -------------------------------------DF                                           485
Fv30A.pro KKGTKKTTTKKGSHQSHKAHSATHRRCRHGSYRRGHCTN                                           537
```

FIG. 65

496

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
NIT : 800.176.089-2
- / -

Industria y Comercio
SUPERINTENDENCIA

RECIBO DE CAJA

No. 12 - 0028656

Bogotá D.C., Marzo 22 de 2012 - 15:59:44

RECIBIDO DE : CASTILLO PINEDA SAS NI 860.046.502

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	VR.PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	425045891	22/03/2012	7.235.000.00

*** Conceptos Pagados ***

CANT.	RENTISTICO	CONCEPTO	Vr.UNDITARIO	Vr.CONCEPTO
1	50005-01-01 SOLICITUDES	1 TRAMITES DE SOL. DE PATENTE DE INVENCION	590.000.00	590.000.00
				$590.000.00

SON: **QUINIENTOS NOVENTA MIL PESOS MONEDA CORRIENTE***

Responsable: _____
-*******-

Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

No. 12-048841- -00000-0000

Fecha: 2012-03-22 16:25:59 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 498

Act. 411 PRESENTACION Folios: 499

497
~~498~~

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
NIT : 800.176.089-2
- / -

Industria y Comercio
SUPERINTENDENCIA

RECIBO DE CAJA

No. 12 - 0028663

Bogotá D.C., Marzo 22 de 2012 - 15:59:44

RECIBIDO DE : CASTILLO PINEDA SAS NI 860.046.502

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	VR.PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	425045891	22/03/2012	7.235.000.00

*** Conceptos Pagados ***

CANT.	RENTISTICO	CONCEPTO	Vr.UNDITARIO	Vr.CONCEPTO
130	50005-01-01 SOLICITUDES	1607 REIVINDICACION UNITARIA ADICIONAL A LAS 10 INICICIALES	30.000.00	3.900.000.00
				$3.900.000.00

SON: **TRES MILLONES NOVECIENTOS MIL PESOS MONEDA CORRIENTE*

Responsable: _____
-*****-

Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

No. 12-048841- -00000-0000

Fecha: 2012-03-22 16:25:59 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 498

Act. 411 PRESENTACION Folios: 499

498
~~499~~

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

No. 12-048841- -00000-0000

Fecha: 2012-03-22 16:25:59 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 498

Act. 411 PRESENTACION Folios: 499

Industria y Comercio
SUPERINTENDENCIA

PATENTE DE INVENCIÓN ☐ MODELO DE UTILIDAD ☐ Art 33 Decisión 486/00

- ☐ Indicación que se solicita una patente.
- ☐ Datos de identificación del solicitante o de la persona que presenta la solicitud
- ☐ Descripción de la invención
- ☐ Dibujos de ser estos pertinentes
- ☐ Comprobante de pago de las tasas establecidas (De ser el caso formato de descuento)

Completa ☐ Incompleta ☐

PATENTE DE INVENCIÓN PCT ☐ MODELO DE UTILIDAD PCT ☐ Art.33 Decisión 486/00, Circular Única

- ☐ Indicación que se solicita una PCT
- ☐ Copia de la solicitud en español, tal como fue presentada inicialmente (capítulo descriptivo, reivindicatorio, resumen)
- ☐ Dibujos de ser estos pertinentes
- ☐ Comprobante de pago de las tasas establecidas (de ser el caso formato de descuento)

Completa ☐ Incompleta ☐

DISEÑO INDUSTRIAL ☐ (Art. 119 Decisión 486/00)

- ☐ Indicación que se solicita Diseño industrial
- ☐ Datos de identificación del solicitante o de la persona que presenta la solicitud
- ☐ Representación gráfica y fotográfica del Diseño industrial o muestra del material que incorpora el diseño
- ☐ Comprobante de pago de las tasas establecidas

Completa ☐ Incompleta ☐

ESQUEMA DE TRAZADO ☐ (Art. 92 Decisión 486/00)

- ☐ Indicación que se solicita un esquema de trazado
- ☐ Datos de identificación del solicitante o de la persona que presenta la solicitud
- ☐ Representación gráfica de un esquema de trazado
- ☐ Comprobante de pago de las tasas establecidas

Completa ☐ Incompleta ☐