

Industria y Comercio
SUPERINTENDENCIA

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
INVENTARIO

BT

Delegatura de Propiedad Industrial
División de Nuevas Creaciones

PGT

SOLICITUD
PATENTE DE INVENCION

10-48157

21. **EXPEDIENTE No.**

54. **TÍTULO** SECUENCIAS SEÑAL Y CHAPERONAS Co-EXPRESADAS PARA MEJORAR
LA PRODUCCION DE PROTEINA EN UNA CELULA HOSPEDERA.

51. **CLASIFICACIÓN INTERNACIONAL**

71. **SOLICITANTE** DANISCO US INC.

DOMICILIO PALO ALTO, CALIFORNIA, ESTADOS UNIDOS DE AMERICA

74. **APODERADO** JERONIMO CASTILLO BONILLA.

22. **BOGOTÁ, D.C.,**

(FORMA P-10)

2b
01-01-2010

PETITORIO



SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 10-048157- -00000-0000

Fecha: 2010-04-23 17:13:15 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 142

FORMULARIO ÚNICO DE SOLICITUD DE PATENTE

PCT

ENTRADA EN FASE NACIONAL

1

SOLICITUD DE :

☒ Patente de Invención.

☐ Patente de Modelo de Utilidad,

☒ Capítulo I.

☐ Capítulo II

2

SOLICITANTE
(71)

Nombre: DANISCO US INC

Dirección: 925 Page Mill Road, Palo Alto, California,
94304-1013, Estados Unidos de America

Nacionalidad o Domicilio: estadounidense

Lugar de Constitución: Delaware

Teléfono:

Fax:

E-mail:

IDENTIFICACIÓN

C.C. ☐ NIT ☐

C.E. ☐ Otro ☐

Cual _____

Número _____

3

REPRESENTANTE
O APODERADO

Nombre: JERÓNIMO CASTILLO BONILLA

Dirección: Calle 93 A N° 14-17, Oficina 311
Bogotá, D.C.

Teléfono: 2 363493 Fax: 2 182351

E-mail: info@pcastillopineda.com

IDENTIFICACIÓN

C.C. ☒ , NIT ☐

C.E. ☐ Otro ☐

Cual _____

Número 79462057 TP 104167

4

INVENTOR (ES)
(72)

Nombre: BAO, Kai

Dirección: 925 Page Mill Road, Palo Alto, California, Estados Unidos de
America

Nacionalidad o Domicilio estadounidense

5

PCT (86)

Solicitud Internacional No. PCT/US2008/081721

Fecha Octubre 30, 2008

Publicación Internacional No. WO 2009/058956 A1

Fecha Mayo 7, 2009

6

Título (54) SECUENCIAS SEÑAL Y CHAPERONAS CO-EXPRESADAS PARA MEJORAR
LA PRODUCCION DE PROTEINA EN UNA CÉLULA HOSPEDERA

7

Clasificación Internacional (51) C12N 15/62, 15/80, C12P 21/02

8

Prioridad

SI ☒ NO ☐

(33) País de Origen

US

(32) Fecha

Nov. 1, 2007

(31) Número de Solicitud

60/984.430

9

Comprobante de pago No.

Fecha

ANEXOS

- ☒ Comprobante de pago de la tasa de representación de la solicitud.
- ☒ Documento que acredite la existencia y representación legal cuando el solicitante sea persona jurídica.
- ☒ Poderes, si fuere el caso.
- ☐ Documentos de cesión del inventor al solicitante o a su causante.
- ☐ Declaraciones.
- ☒ Copia de la solicitud internacional. (Resumen, descripción, reivindicaciones, dibujos).
- ☒ 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 del 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 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. Notificación de la Oficina Internacional sobre el cambio de Nombre de Danisco US Inc., Genencor Division por el de DANISCO US INC. y la Publicación Internacional.

11

FIGURA CARACTERÍSTICA:

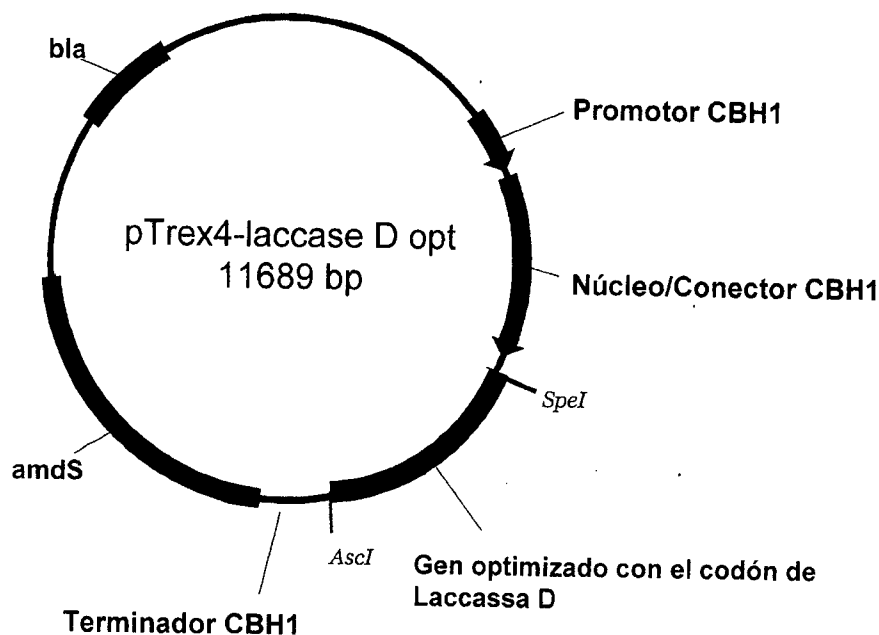


FIG. 1

12

Solicito la concesión de la patente.

NOMBRE: JERÓNIMO CASTILLO BONILLA

FIRMA:

C.C. 79462057

T.P. 104167

(continuación Inventores)

- Nombre : WANG, Huaming
- Dirección : 925 Page Mill Road, Palo Alto, California, Estados Unidos
de América
Nacionalidad : estadounidense

CORRECTED VERSION

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF THE RECORDING
OF A CHANGE(PCT Rule 92bis.1 and
Administrative Instructions, Section 422)

To:

BACSI, Steven, G.
Danisco US Inc.
925 page Mill Road
Palo Alto, CA 94304
ETATS-UNIS D'AMERIQUE

Date of mailing (day:month:year) 04 March 2010 (04.03.2010)	IMPORTANT NOTIFICATION
Applicant's or agent's file reference 30993-WO	
International application No. PCT/US2008/081721	
International filing date (day:month:year) 30 October 2008 (30.10.2008)	

1. The following indications appeared on record concerning:		
☒ the applicant ☐ the inventor ☐ the agent ☐ the common representative		
Name and Address DANISCO US, INC. Genencor Division 925 Page Mill Road Palo Alto, CA 94304-1013 United States of America	State of Nationality US	State of Residence US
	Telephone No. 650-846-7615	
	Facsimile No. 650-621-8115	
	E-mail address	
2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:		
☐ the person ☒ the name ☐ the address ☐ the nationality ☐ the residence		
Name and Address DANISCO US INC. 925 Page Mill Road Palo Alto, CA 94304-1013 United States of America	State of Nationality US	State of Residence US
	Telephone No. 650-846-7615	
	Facsimile No. 650-621-8115	
	E-mail address ☐ Notifications by e-mail authorized	
3. Further observations, if necessary:		
4. A copy of this notification has been sent to:		
☒ the receiving Office ☐ the International Preliminary Examining Authority		
☐ the International Searching Authority ☒ the designated Offices concerned		
☐ the Authority(ies) specified for supplementary search ☐ the elected Offices concerned		
☐ other:		

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland	Authorized officer Porohovski Youri
Facsimile No. +41 22 338 87 20	e-mail pt03.pct@wipo.int
Form PCT/IB/306 (January 2009)	Telephone No. +41 22 338 74 03

I/IB/306X150

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
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:*

CERTIFIED

10. Signature



Secretary of State

BY

[Signature]
Secretary of State



LAW OFFICES
PEDRO CASTILLO PINEDA &
ASSOCIATES
P.O. BOX 48362
Bogotá, Colombia / South America

IMPORTANT NOTE
The lodging of this GENERAL POWER with the Patent Office obviates the need of presenting special powers with each individual trade mark application.

FOR PATENTS AND TRADE MARKS IN
COLOMBIA

Poder General -

General Power of Attorney
(1) Danisco US Inc.

1 Applicant's (1)
name

2 Applicant's domiciliado en (2)
address

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

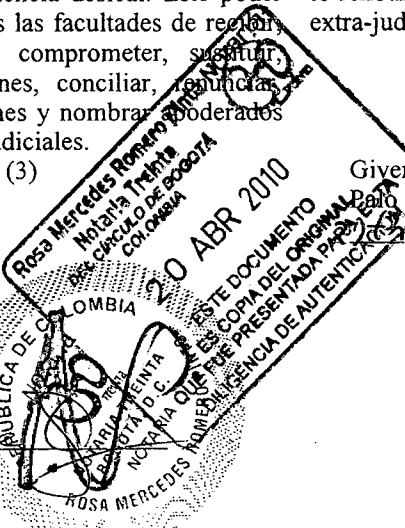
nombra al Dr. JERONIMO CASTILLO BONILLA y/o Dra. XIMENA BONILLA PEREIRA y/o Dr. ALVARO CASTILLO GRAU Calle 93A N° 14-17 Oficina 311 en Bogotá, Colombia, como sus representantes con poder especial, amplio y suficiente, para recabar de las oficinas y autoridades Colombianas, administrativas, judiciales 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.

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.

3 Date
4 Applicant's signature

Dado y firmado en (3)
Given and signed in (3)
Palo Alto, California, U.S.A.
20 day of August, 2009 ,

4)Signature (s)
X 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

Notary Public.

(with Consular legalization)



IN WITNESS WHEREOF I undersign as follows:

Date: August 27, 2009

Soonhee Jang
Soonhee Jang,
Vice President, IP Strategy &
Chief IP Counsel

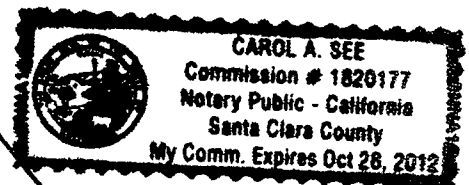
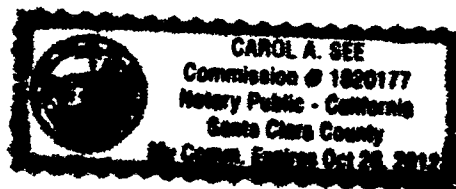
STATE OF CALIFORNIA
COUNTY OF SANTA CLARA

On 8/27/09, before me, Carol A. See, Notary Public, personally appeared, Soonhee Jang, 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 Carol A. See

(Seal)



(12) SOLICITUD INTERNACIONAL PUBLICADA SEGÚN EL TRATADO DE COOPERACIÓN INTERNACIONAL EN MATERIA DE PATENTES (PCT)

(19) Organización Mundial de Propiedad Intelectual
Oficina Internacional



(43) Fecha de Publicación Internacional
7 de mayo de 2009 (07.05.2009)

PCT

(10) Número de Publicación Internacional
WO 2009/058956 A1

(51) Clasificación Internacional de Patentes:

C12N 9/02 (2006.01) C12N15/62 (2006.01)
C12N 9/34 (2006.01) C12N15/80 (2006.01)
C12N 9/42 (2006.01) C12P 21/02 (2006.01)
C12N 9/58 (2006.01)

(74) Agente: KUNG, Viola, T.; Howrey Lip, 2941 Fairview
Park Drive, Box 7, Falls Church, VA 22042 (US).

(81) Países designados (a menos que se indique lo contrario para cada clase de protección disponible):
AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FT, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(21) No. Solicitud Internacional: PCT/US2008/081721

(22) Fecha de Presentación Internacional:
30 de octubre de 2008 (30.10.2008)

(25) Idioma de Presentación: Inglés

(26) Idioma de Publicación: Inglés

(30) Datos relativos a la Prioridad:
60/984,430 1 de noviembre de 2007 (01.11.2007) US

(71) Solicitante (para todos los países destinatarios con excepción de US): **DANISCO US, INC.** [US/US];
Genencor Division, 925 Page Mill Road, Palo Alto, CA 94304-1013 (US).

(72) Inventores, e

(75) Inventores/Solicitantes (para US únicamente): **BAO, Kai** [US/US]; Danisco Us, Inc., Genencor Division, 925 Page Mill Road, Palo Alto, CA 94304-1013 (US).
WANG, Huaming [US/US]; Danisco Us, Inc., Genencor Division, 925 Page Mill Road, Palo Alto, CA 94304-1013 (US).

(84) Países designados (a menos que se indique lo contrario para cada clase de protección disponible):
ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasiática (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), Europea (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FT, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Publicado:

— Con informe internacional de búsqueda

[Continúa en la siguiente página]

(54) Título: SECUENCIAS SEÑAL Y CHAPERONAS CO-EXPRESADAS PARA MEJORAR LA PRODUCCIÓN DE PROTEÍNA EN UNA CÉLULA HOSPEDERA

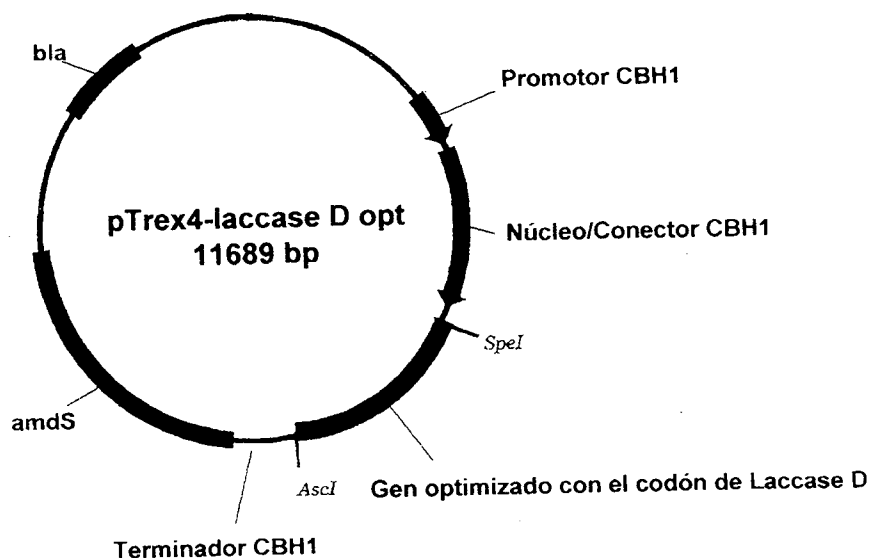


FIG. 1

(57) Extracto: La invención provee métodos y composiciones para la producción incrementada de proteína. El método comprende los pasos de: (a) introducir en una célula hospedera una primera secuencia de ácido nucleico que incluye una secuencia señal conectada operacionalmente a una secuencia de proteína deseada; (b) expresar la primera secuencia de secuencia señal conectada operacionalmente a una secuencia de proteína deseada; (c) co-expresar una segunda secuencia de ácido nucleico que codifica una chaperona o foldasa seleccionada del grupo conformado por *bip1*, *ero1*, *pdi1*, *tig1*, *prp1*, *ppi1*, *ppi2*, *prp3*, *prp4*, calnexina, y *lhs1*; and (d) recolectar la proteína deseada segregada de la célula hospedera. La primera secuencia de ácido nucleico incluye opcionalmente una secuencia enzimática entre la secuencia señal y la secuencia de la proteína deseada.

(12) SOLICITUD INTERNACIONAL PUBLICADA SEGÚN EL TRATADO DE COOPERACIÓN INTERNACIONAL EN MATERIA DE PATENTES (PCT)

(19) Organización Mundial de Propiedad Intelectual
Oficina Internacional



(43) Fecha de Publicación Internacional
7 de mayo de 2009 (07.05.2009)

PCT

(10) Número de Publicación Internacional
WO 2009/058956 A1

(51) Clasificación Internacional de Patentes:

C12N 9/02 (2006.01) C12N15/62 (2006.01)
C12N 9/34 (2006.01) C12N15/80 (2006.01)
C12N 9/42 (2006.01) C12P 21/02 (2006.01)
C12N 9/58 (2006.01)

(74) Agente: KUNG, Viola, T.; Howrey Lip, 2941 Fairview Park Drive, Box 7, Falls Church, VA 22042 (US).

(81) Países designados (a menos que se indique lo contrario para cada clase de protección disponible): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FT, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(21) No. Solicitud Internacional: PCT/US2008/081721

(22) Fecha de Presentación Internacional:
30 de octubre de 2008 (30.10.2008)

(25) Idioma de Presentación: Inglés

(26) Idioma de Publicación: Inglés

(30) Datos relativos a la Prioridad:
60/984,430 1 de noviembre de 2007 (01.11.2007) US

(71) Solicitante (para todos los países destinatarios con excepción de US): **DANISCO US, INC.** [US/US]; Genencor Division, 925 Page Mill Road, Palo Alto, CA 94304-1013 (US).

(72) Inventores, e

(75) Inventores/Solicitantes (para US únicamente): **BAO, Kai** [US/US]; Danisco Us, Inc., Genencor Division, 925 Page Mill Road, Palo Alto, CA 94304-1013 (US). **WANG, Huaming** [US/US]; Danisco Us, Inc., Genencor Division, 925 Page Mill Road, Palo Alto, CA 94304-1013 (US).

(84) Países designados (a menos que se indique lo contrario para cada clase de protección disponible): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasiática (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), Europea (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FT, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Publicado:

— Con informe internacional de búsqueda

[Continúa en la siguiente página]

(54) Título: SECUENCIAS SEÑAL Y CHAPERONAS CO-EXPRESADAS PARA MEJORAR LA PRODUCCIÓN DE PROTEÍNA EN UNA CÉLULA HOSPEDERA

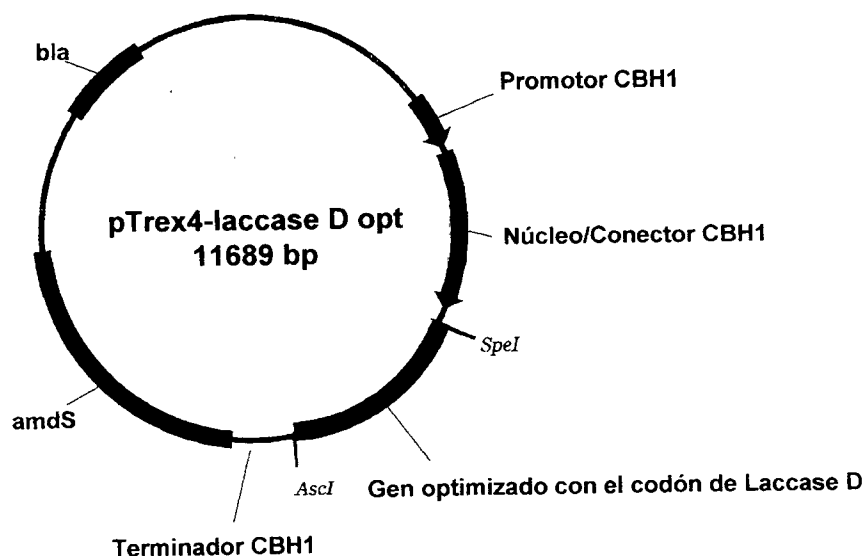


FIG. 1

(57) Extracto: La invención provee métodos y composiciones para la producción incrementada de proteína. El método comprende los pasos de: (a) introducir en una célula hospedera una primera secuencia de ácido nucleico que incluye una secuencia señal conectada operacionalmente a una secuencia de proteína deseada; (b) expresar la primera secuencia de ácido nucleico; (c) co-expresar una segunda secuencia de ácido nucleico que codifica una chaperona o foldasa seleccionada del grupo conformado por *bip1*, *ero1*, *pdi1*, *tig1*, *prp1*, *ppi1*, *ppi2*, *prp3*, *prp4*, calnexina, y *lhs1*; and (d) recolectar la proteína deseada segregada de la célula hospedera. La primera secuencia de ácido nucleico incluye opcionalmente una secuencia enzimática entre la secuencia señal y la secuencia de la proteína deseada.

WO 2009/058956 A1



— Con parte de la descripción del listado de secuencias publicada separadamente en formato electrónico y disponible a solicitud de La Oficina Internacional.

SECUENCIAS SEÑAL Y CHAPERONAS CO-EXPRESADAS PARA MEJORAR LA PRODUCCIÓN DE PROTEÍNA EN UNA CÉLULA HOSPEDERA

CAMPO DE LA INVENCION

Esta invención provee métodos y composiciones para producción mejorada de proteínas. En algunas incorporaciones, los métodos provistos en esta especificación involucran el uso de una secuencia señal conectada operacionalmente a una proteína. En algunas incorporaciones, la secuencia señal conectada operacionalmente a una proteína se expresa en combinación con por lo menos una chaperona en una célula hospedera. En algunas incorporaciones, la proteína se expresa en una célula fúngica filamentosa. En incorporaciones adicionales, los métodos de la presente invención involucran fusión de una proteína al dominio catalítico de una enzima, tal como una glucoamilasa o una CBH1. Algunas incorporaciones proveen combinaciones de una secuencia señal, una o más de una chaperona, chaperonina y/o foldasa, y/o fusión de la proteína a una proteína catalítica o dominio.

ARTE PREVIO DE LA INVENCION

Células hospederas tales como levadura, hongos filamentosos y bacterias se han empleado hace mucho tiempo para expresar y segregar proteína extraña. Normalmente, la producción de estas proteínas extrañas en levadura, hongos filamentosos y bacterias implica la expresión y purificación parcial o completa de la proteína de la célula hospedera o el medio de cultivo en el cual se desarrollan las células. Aunque algunas proteínas requieren purificación del medio intracelular de las células hospederas, la purificación puede simplificarse en gran medida si las proteínas se segregan de la célula en el medio de cultivo.

La secreción de proteína extracelular es un aspecto importante y complicado de la producción de proteínas en varios sistemas de expresión celular. Uno de los factores asociados con la secreción de proteínas es el plegamiento apropiado de la proteína. Muchas proteínas pueden desplegarse irreversiblemente y replegarse *in vitro* en concentraciones de dilución, cuando toda la información requerida para especificar una estructura de proteína plegada compacta está

presente en la secuencia de aminoácidos de las proteínas. No obstante, el plegamiento de proteínas *in vivo* ocurre en un medio concentrado de numerosas proteínas, en el cual reacciones de agregación intermolecular compiten con el proceso de plegamiento intramolecular. Estas complicaciones son más significativas en sistemas de expresión eucariotas que en sistemas procariotas.

El primer paso en la vía secretora eucariota es la translocación del polipéptido nascente a través de la membrana de retículo endoplasmático (ER – *endoplasmic reticulum*) en forma extendida. El plegamiento correcto y ensamblado de un polipéptido ocurre en el ER a través de la vía secretora. Sin embargo, en muchos casos, aunque las proteínas se sobre-expresan en gran medida, se segregan deficientemente. Por supuesto, en muchos casos las señales de secreción que deben facilitar tal expresión no parecen realizarla. La expresión de proteínas deseadas se complica aún más por la interacción de otras proteínas. Estos factores son incluso más significativos cuando se intenta la expresión de una proteína, obtenida de una especie, género o familia de organismos, en otra especie, género o familia. Por ejemplo, normalmente las proteínas de Basidiomycetes (v.g., laccasa) se expresan en forma deficiente en hospederos Ascomycetes tales como *Trichoderma*. Por supuesto, a pesar de una gran cantidad de trabajo en el área de sistemas fúngicos de expresión, subsiste una necesidad de una expresión extracelular mejorada de proteínas deseadas.

SUMARIO DE LA INVENCION

La invención provee métodos y composiciones para una producción mejorada de proteína. Los métodos implican el uso de una secuencia señal conectada operativamente a una proteína deseada, la cual se expresa en combinación con por lo menos una chaperona en una célula hospedera. En algunas incorporaciones, la proteína se expresa en una célula fúngica filamentosa. En incorporaciones adicionales, los métodos de la presente invención implican la fusión de una proteína deseada al dominio catalítico de una proteína hospedera, tal como una glucoamilasa o una CBH1.

En algunas incorporaciones, la presente invención provee métodos y composiciones para aumentar la producción de proteínas en hospederos fúngicos

filamentosos (v.g., Ascomycetes), mediante el uso de una señal secretora en combinación con la expresión de una proteína chaperona obtenida del mismo organismo que la proteína. En algunas incorporaciones, la proteína es una proteína distinta de Ascomycete, que se fusiona a la señal secretora de una proteína hospedera de Ascomycetes. En algunas incorporaciones adicionales, por lo menos una proteína chaperona encuentra aplicación en aumentar la expresión de proteínas fusionadas al dominio catalítico de una proteína de Ascomycetes.

Algunas incorporaciones proveen métodos para producir por lo menos una proteína en una célula hospedera de Ascomycetes, introduciendo en una célula hospedera un polinucleótido que comprende una proteína deseada conectada operacionalmente a una secuencia señal del mismo filum, género o especie que el hospedero; co-expresando una chaperona del mismo filum, género o especie que la proteína; cultivando la célula hospedera bajo condiciones apropiadas de cultivo para la expresión y producción de la proteína; y produciendo la proteína. El método incluye opcionalmente recuperar la proteína producida. Algunas incorporaciones incluyen fusionar la proteína al dominio catalítico de una enzima de Ascomycetes. Otras incorporaciones incluyen fusionar la proteína a una enzima de longitud total de Ascomycetes. En algunas incorporaciones, la célula hospedera de Ascomycetes es *Trichoderma*. En algunas incorporaciones, la chaperona es por lo menos una de las siguientes, BIP1, ERO1, PDI1, TIG1, PRP1, PPI1, PPI2, PRP3, PRP4, CALNEXINA y LHS1.

La elección de la proteína no es limitante, y puede incluir cualquiera de las siguientes proteínas de cualquier género, especie, y/o familia: laccasas, glucoamilasas, α -amilasas, enzimas que hidrolizan almidón granular, celulasas, lipasas, xilanasas, cutinasas, hemicelulasas, proteasas, oxidasas, laccasas y combinaciones de las mismas. Algunas incorporaciones incluyen secuencias señal de genes NSP24 o CBH1. En algunas incorporaciones, el gen de la chaperona es bip1. Incorporaciones del método pueden incluir asimismo un promotor de Ascomycetes. En algunas incorporaciones, la célula hospedera y la secuencia señal es del mismo hospedero Ascomycetes. En algunas incorporaciones, el promotor es el promotor CBH1 de *Trichoderma*. En algunas incorporaciones, la proteína es una proteína de Basidiomycetes. En algunas incorporaciones, la célula hospedera es una célula hospedera de Ascomycetes. En algunas incorporaciones,

la célula hospedera es una célula hospedera de Basidiomycetes y la proteína es una proteína de Ascomycetes.

Algunas incorporaciones adicionales proveen métodos para producir por lo menos una proteína en una célula hospedera de Ascomycetes, introduciendo en una célula hospedera de Ascomycetes un polinucleótido que comprende una proteína deseada fusionada al dominio catalítico de una enzima de Ascomycetes, en donde la proteína deseada es una proteína de Basidiomycetes; co-expresando una chaperona de Ascomycetes; cultivando la célula hospedera de Ascomycetes bajo condiciones apropiadas de cultivo para la expresión y producción de la proteína; y produciendo la proteína. En algunas incorporaciones, la proteína producida se recupera. En algunas incorporaciones, la proteína se conecta operacionalmente a una secuencia señal de Ascomycetes.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

Figura 1 muestra una representación esquemática del plásmido de expresión pTrex4-laccase D opt de *Trichoderma*. La secuencia polinucleotídica se muestra como SEQ ID NO: 1.

Figura 2 muestra una representación esquemática del plásmido de expresión pTrex2g-Bip1 de *Trichoderma*. La secuencia polinucleotídica se muestra como SEQ ID NO: 2.

Figura 3 muestra una representación esquemática del plásmido de expresión pTrex2g-Pdi1 de *Trichoderma*. La secuencia polinucleotídica se muestra como SEQ ID NO: 3.

Figura 4 muestra una representación esquemática de la secuencia *Ero1* usada en el plásmido de expresión pTrex2g-Ero1 de *Trichoderma*. La secuencia polinucleotídica se muestra como SEQ ID NO: 4.

Figura 5 muestra una representación esquemática del plásmido de expresión pTrGA-laccase D opt de *Trichoderma*. La secuencia polinucleotídica se muestra como SEQ ID NO: 5.

Figura 6 muestra una representación esquemática del plásmido de expresión pKB408 de *Trichoderma*. La secuencia polinucleotídica se muestra como SEQ ID NO: 6.

Figura 7 muestra una representación esquemática del plásmido de expresión pKB410 de *Trichoderma*. La secuencia polinucleotídica se muestra como SEQ ID NO: 7.

Figuras 8-1 a 8-4 muestran la SEQ ID NO:8 del marco de lectura abierto (ORF – *Open Reading Frame*) de NSP24 de *T. reesei*. El péptido señal son los primeros 20 aminoácidos (SEQ ID NO: 9).

Figuras 9-1 y 9-2 muestran la (SEQ ID NO: 10) del ORF de CBH1 de *T. reesei*. La secuencia señal comienza en el par de bases 210 y termina en el par de bases 260 (SEQ ID NO: 11). El núcleo catalítico comienza en el par de bases 261 al par de bases 1698 (SEQ ID NO: 12), incluyendo el intrón 1 (desde el par de bases 671 a 737) y el intrón 2 (desde el par de bases 1435 a 1497). La secuencia conectora comienza en el par de bases 1699 y termina en el par de bases 1770 (SEQ ID NO: 13). La secuencia de la proteína CBH1 se muestra como SEQ ID NO: 14.

Figura 10 ilustra la mejora de producción de laccasa por fusión del gen que codifica la lacasa de *C. unicolor* a la glucoamilasa de *Trichoderma* de longitud total. La cepa #8-2 es la fusión de laccasa CBH1. Las cepas 1066-9, 1066-13, y 1066-15 son fusión de laccasa TrGA.

Figura 11 ilustra la mejora de producción de laccasa por fusión del gen, que codifica la laccasa de *C. unicolor*, a la secuencia señal CBH1 o NSP24 en matraces de agitación. El eje Y muestra la actividad de la laccasa como unidades/ml. El eje X muestra las cepas (fusión de CBH1 sola, o con secuencia señal).

Figura 12 ilustra la mejora de producción de laccasa por fusión del gen, que codifica la laccasa *C. unicolor*, a la secuencia señal CBH1 o NSP24 en

fermentadores. El eje Y muestra la actividad de la laccasa como unidades/ml. El eje X muestra el tiempo de fermentación como horas.

Figura 13 ilustra la mejora de producción de laccasa proporcionada por la secuencia señal CBH1 más la expresión de la chaperona BIP1. El eje Y muestra la actividad de la laccasa como unidades/ml. El eje X muestra el tiempo de fermentación como horas.

Figura 14 ilustra la mejora de producción de laccasa mediante co-expresión de chaperonas con *C. unicolor* en matraces de agitación en 3, 4, y 5 días. El eje Y muestra la actividad de la laccasa como unidades/ml. El eje X muestra las cepas (KB410-13, o con co-expresión de bip).

Figura 15 ilustra la mejora de producción de laccasa por fusión del gen, que codifica la laccasa *C. unicolor*, a la secuencia señal CBH1, dominio catalítico y conector y co-expresión con la chaperona Bip1, pdi1 o ero1. El eje Y muestra la actividad de la laccasa como unidades/ml. El eje X muestra las cepas.

DESCRIPCIÓN DETALLADA DE LA INVENCION

A menos que se indique lo contrario, la práctica de la presente invención involucra técnicas convencionales usadas comúnmente en biología molecular, ingeniería de proteínas, técnicas de ADN recombinante, microbiología, biología celular, cultivo de células, biología transgénica, inmunología, y purificación de proteínas, las cuales están dentro de la habilidad del arte. Tales técnicas son conocidas por los expertos en el arte y se describen en numerosos textos y trabajos de referencia. Todas las patentes, solicitudes de patente, artículos y publicaciones mencionadas en esta especificación, tanto *supra* como *infra*, se incorporan aquí expresamente por referencia.

Definiciones

El término "Ascomycetes" se refiere a una clase de hongos pertenecientes al filum Ascomycota. Miembros de este filum se distinguen por la presencia de ascas (es decir, células especializadas en forma de sacos que contienen

ascosporas).

El término "Basidiomycetes" se refiere a una clase de hongos pertenecientes al filum Basidiomycota. Miembros de este filum se caracterizan por la producción de basidiosporas, (es decir, esporas sexuales que se localizan en áreas externas especializadas de células terminales claviformes conocidas como basidios).

"Proteasa" significa una proteína o dominio polipeptídico de una proteína o polipéptido que tiene la capacidad de catalizar la escisión de enlaces peptídicos en una o más de varias posiciones de un esqueleto proteico (v.g. E.C. 3.4). Las proteasas son obtenibles de microorganismos (v.g. hongos o bacterias), plantas, y/o animales.

Una "protease ácida" se refiere a una proteasa que tiene la capacidad de hidrolizar proteínas bajo condiciones ácidas.

Tal como se emplea en esta especificación, el término "chaperona" o "chaperonas moleculares" facilitan el plegamiento de la proteína blindando regiones desplegadas de proteínas circundantes y no mejoran la tasa de plegamiento de proteínas. Esto puede incluir proteínas y sus homólogos que ayudan al plegamiento y glicosilación de las proteínas secretoras en el retículo endoplasmático (ER). Las chaperonas pueden ser residentes en el ER. Chaperonas ejemplares incluyen Bip (GRP78), GRP94 y Lhs1p de levadura y aquellas que ayudan a la proteína secretora a plegarse uniéndose a regiones hidrofóbicas expuestas en los estados desplegados y evitando interacciones desfavorables. Las chaperonas incluyen asimismo proteínas que están involucradas en la translocación de proteínas a través de la membrana del ER.

Tal como se emplea en esta especificación, "chaperoninas" son proteínas que contribuyen al plegamiento de la proteína al estado nativo (estado activo) utilizando ATP. A menudo las subunidades de proteína se ensamblan entre sí para formar ensambles anulares grandes. Por ejemplo, las chaperoninas actúan uniendo proteínas no nativas en sus cavidades centrales y luego, después de unir ATP, liberan la proteína del sustrato en una cavidad ahora encapsulada para

plegarse productivamente.

"Proteínas foldasas" significa proteínas que catalizan pasos en el plegamiento de proteínas para aumentar la tasa de plegamiento de proteínas. Por ejemplo, pueden ayudar en la formación de puentes bisulfuro y en la formación de la conformación correcta de cadenas peptídicas adyacentes a residuos de prolina. Foldasas ejemplares incluyen la proteína bisulfuro isomerasa (pdi - *protein disulphide isomerase*) y sus homólogos, y prolil-peptidil cis-trans isomerasa y sus homólogos.

Tal como se emplea en esta especificación, "proteasa de la familia NSP24" significa una enzima que tiene actividad proteasa en su forma nativa o tipo silvestre, perteneciente a la familia de las proteasas NSP24. Las proteasas NSP24 son proteasas ácidas, tales como proteasas fúngicas ácidas. Las proteasas NSP24 tienen por lo menos 85%, por lo menos 90%, por lo menos 93%, por lo menos 95%, por lo menos 96%, por lo menos 97%, por lo menos 98% y por lo menos 99% de identidad de secuencia con la secuencia de aminoácidos de SEQ ID NO: 8 y sus fragmentos biológicamente activos.

Tal como se emplea en esta especificación, el término "una proteína deseada" significa una proteína de interés. Una proteína deseada y una proteína de interés se utilizan indistintamente en esta solicitud. En algunas incorporaciones, la proteína deseada es una proteína industrial comercialmente importante. Se pretende que el término abarque proteínas que son codificadas por genes de ocurrencia natural, genes mutados y/o genes sintéticos. La proteína deseada puede ser una proteína nativa a la célula hospedera, o no nativa (heteróloga) a la célula hospedera.

Tal como se emplea en esta especificación, "derivada" significa una proteína que se deriva de una proteína precursora o progenitora (v.g., la proteína nativa) mediante adición de uno o más aminoácidos a cualquiera, o ambos, de los extremos C- y N-terminales, sustitución de uno o más aminoácidos en uno, o en un número de sitios diferentes en la secuencia de aminoácidos, delección de uno o más aminoácidos en cualquiera, o ambos extremos de la proteína, o en uno o más sitios en la secuencia de aminoácidos, o inserción de uno o más aminoácidos en

uno o más sitios en la secuencia de aminoácidos.

El término "recombinante" se refiere a un polinucleótido o polipéptido que no ocurre naturalmente en una célula hospedera. Una molécula recombinante puede contener dos o más secuencias que ocurren naturalmente, las cuales se conectan entre sí en una forma que no ocurre naturalmente.

Los términos "péptidos," "proteínas," y "polipéptidos" se usan indistintamente en esta especificación.

Tal como se emplea en esta especificación, "porcentaje (%) de identidad de secuencia" con respecto a secuencias nucleotídicas o de aminoácidos se define como el porcentaje de residuos de aminoácido o nucleótidos en una secuencia candidata que son idénticos con los residuos de aminoácido o nucleótidos en una secuencia de interés (v.g., una secuencia de péptido señal NSP24), después de alinear las secuencias e introducir espacios, si es necesario, para lograr el porcentaje máximo de identidad de secuencia, y no considerando sustituciones conservadoras de algún tipo como parte de la identidad de secuencia.

Tal como se emplea en esta especificación, el término " α -amilasa (v.g., E.C. clase 3.2.1.1)" se refiere a enzimas que catalizan la hidrólisis de ligamientos α -1,4-glucosídicos. Estas enzimas se han descrito también como aquellas que efectúan la exo- o endohidrólisis de ligamientos 1,4- α -D-glucosídicos en polisacáridos que contienen unidades de D-glucosa 1,4- α -ligadas. Otro término empleado para describir estas enzimas es "glucogenasa." Enzimas ejemplares incluyen α -1,4-glucan 4-glucanohidrasa glucanohidrolasa.

Tal como se emplea en esta especificación, el término "glucoamilasa" se refiere a la clase amiloglucosidasa de enzimas (v.g., EC.3.2.1.3, glucoamilasa, 1,4- α -D-glucan glucohidrolasa). Éstas son enzimas exo-actantes, las cuales liberan residuos glucosilo de los extremos no reductores de moléculas de amilosa y amilopectina. La enzima hidroliza igualmente los ligamientos α -1,6 y α -1,3 aunque a una velocidad mucho más lenta que los ligamientos α -1,4.

El término "promotor" significa una secuencia reguladora involucrada en

unir ARN polimerasa para iniciar la transcripción de un gen.

Un "promotor heterólogo", tal como se emplea en esta especificación, se refiere a un promotor que se ha puesto en asociación con un gen o ácido nucleico purificado, pero que no está asociado naturalmente con ese gen o ácido nucleico purificado.

Una "preparación purificada" y una "preparación sustancialmente pura" de un polipéptido, tal como se emplea en esta especificación, significan un polipéptido que se ha separado de células, otras proteínas, lípidos o ácidos nucleicos, con los cuales ocurre naturalmente.

"Homólogo," tal como se emplea en esta especificación, se refiere a la similaridad de secuencia entre dos o más moléculas de polipéptido o entre dos o más moléculas de ácido nucleico. Cuando una posición en las secuencias que se están comparando está ocupada por la misma base o subunidad monomérica de aminoácido, (v.g., si una posición en cada una de dos moléculas de ADN está ocupada por adenina), entonces las moléculas son homólogas en esa posición. El porcentaje de homología entre dos secuencias es una función del número de posiciones coincidentes u homólogas, compartidas por las dos secuencias, dividido por el número de posiciones comparadas x 100. Por ejemplo, si 6 de 10, de las posiciones en dos secuencias son coincidentes u homólogas, entonces las dos secuencias son 60% homólogas. A modo de ejemplo, Las secuencias de ADN ATTGCC y TATGGC comparten 50% de homología. Generalmente se hace una comparación cuando dos secuencias están alineadas para dar la máxima homología. El término "% de homología" se utiliza indistintamente en esta especificación con el término "% de identidad" en esta especificación y se refiere al nivel identidad de secuencia de aminoácidos o ácido nucleico entre las secuencias de ácido nucleico o secuencias de aminoácidos, cuando se alinean usando un programa de alinación de secuencias.

Tal como se emplea en esta especificación, el término "vector" se refiere a una secuencia polinucleotídica diseñada para introducir ácidos nucleicos en uno o más tipos de células. Los vectores incluyen vectores de clonación, vectores de expresión, vectores lanzadera, plásmidos, partículas de bacteriófago, casetes y

similares.

Tal como se emplea en esta especificación, "vector de expresión" significa un constructo de ADN que incluye una secuencia de ADN que se conecta operacionalmente a una secuencia de control apropiada, capaz de afectar la expresión del ADN en un hospedero apropiado.

El término "expresión" significa el proceso mediante el cual se produce un polipéptido basado en la secuencia de ácido nucleico de un gen.

El término "co-expresión" significa que por lo menos dos diferentes genes se expresan en una célula. Éstos pueden ser genes exógenos, o genes endógenos. Pueden integrarse o expresarse de los mismos o diferentes plásmidos, y pueden expresarse del mismo o diferente promotor.

Tal como se emplea en esta especificación, "conectado(a) operacionalmente" significa que una región reguladora, tal como un promotor, terminador, señal de secreción o región potenciadora se une o conecta a un gen estructural y controla la expresión de ese gen. Una secuencia señal se conecta operacionalmente a una proteína si dirige la proteína a través del sistema de secreción de una célula hospedera.

Tal como se emplea en esta especificación, "microorganismo" se refiere a una bacteria, un hongo, un virus, un protozoo, y demás microbios u organismos microscópicos.

El término "hongos filamentosos" se refiere a todas las formas filamentosas de la subdivisión Eumycotina, como se conoce en el arte. Estos hongos se caracterizan por un micelio vegetativo con una pared celular compuesta de quitina, celulosa, y otros polisacáridos complejos. Los hongos filamentosos de la presente invención son morfológica, fisiológica y genéticamente distintos de las levaduras. El crecimiento vegetativo por hongos filamentosos se produce por alargamiento de la hifa, y el catabolismo del carbono es obligatoriamente aeróbico.

Tal como se emplea en esta especificación, el término "*Trichoderma*" y

"*Trichoderma sp.*" se refieren a cualquier género fúngico clasificado previa o actualmente como *Trichoderma*.

Tal como se emplea en esta especificación el término "cultivar" se refiere a hacer crecer una población de células microbianas bajo condiciones apropiadas en un medio líquido, semi-sólido o sólido. En algunas incorporaciones, el cultivo se conduce en un tanque o reactor, como se conoce en el arte. En algunas incorporaciones, el cultivo resulta en la bioconversión fermentadora de un sustrato de almidón, tal como un sustrato que comprende almidón granular, en un producto final.

"Fermentación" se refiere a la descomposición enzimática y anaeróbica de sustancias orgánicas por microorganismos para producir compuestos orgánicos más simples. Aunque, con frecuencia, la fermentación ocurre bajo condiciones anaeróbicas, no se pretende que el término se limite únicamente a condiciones anaeróbicas estrictas, ya que la fermentación también ocurre en la presencia de oxígeno.

El término "introducido(a)", en el contexto de insertar una secuencia de ácido nucleico en una célula, significa "transfección," "transformación" o "transducción," e incluye referencia a la incorporación de una secuencia de ácido nucleico en una célula eucariota o procariota, en donde la secuencia de ácido nucleico se incorpora en el genoma de la célula (v.g., cromosoma, plásmido, plástido, o ADN mitocondrial), se convierte en un replicón autónomo, o se expresa transitoriamente (v.g., ARNm transfectado).

Tal como se emplea en esta especificación, los términos "transformado(a)," "transformado(a) de manera estable" y "transgénico(a)", utilizados en referencia a una célula, significan que la célula tiene una secuencia de ácido nucleico no nativa integrada en su genoma o como un plásmido episómico que se mantiene a través de múltiples generaciones.

Tal como se emplea en esta especificación, el término "heterólogo", utilizado en referencia a un polipéptido o un polinucleótido que codifica una proteína deseada, significa un polipéptido o polinucleótido que no ocurre

naturalmente en una célula hospedera.

El término "homólogo" o "endógeno" con referencia a un polipéptido o un polinucleótido que codifica una proteína deseada se refiere a un polipéptido o un polinucleótido que ocurre naturalmente en, o se expresa naturalmente por, la célula hospedera.

El término "sobre-expresión" significa el proceso de expresar un polipéptido en una célula hospedera en un nivel que es mayor que el producido por una célula hospedera tipo silvestre. En algunas incorporaciones, por lo menos un polinucleótido se introduce en la célula hospedera. En algunas incorporaciones adicionales, el término se refiere a la expresión de un polipéptido homólogo en una concentración que es mayor que la expresión del mismo polipéptido homólogo, expresado por una célula tipo silvestre.

Tal como se describe en esta especificación, un aspecto de la invención presenta un ácido nucleico "sustancialmente puro" que comprende una secuencia nucleotídica que codifica un péptido señal NSP24 o un péptido señal CBH1 conectado operacionalmente a una proteína, y/o equivalentes de tales ácidos nucleicos. En estas incorporaciones, el ácido nucleico se aísla de otros ácidos nucleicos y/o constituyentes de la célula.

El término "equivalente" se refiere a secuencias nucleotídicas que codifican polipéptidos funcionalmente equivalentes. Secuencias nucleotídicas equivalentes abarcan secuencias que difieren en una o más sustituciones, adiciones y/o deleciones nucleotídicas, tales como variantes alélicas. Por ejemplo, en algunas incorporaciones, debido a la degeneración del código genético, secuencias nucleotídicas equivalentes incluyen secuencias que difieren de la secuencia nucleotídica de SEQ ID NO:8, pero que resultan en la producción de polipéptidos que son funcionalmente equivalentes a la secuencia polipeptídica codificada por SEQ ID NO:8.

Esta invención provee un método para producir una proteína deseada. El método comprende los pasos de: (a) introducir en una célula hospedera una primera secuencia de ácido nucleico que comprende una secuencia señal

conectada operacionalmente a una secuencia de la proteína deseada; (b) expresar la primera secuencia de ácido nucleico; (c) co-expresar una segunda secuencia de ácido nucleico que codifica una chaperona o foldasa seleccionada del grupo conformado por *bip1*, *ero1*, *pdi1*, *tig1*, *prp1*, *ppi1*, *ppi2*, *prp3*, *prp4*, *calnexina*, y *lhs1*; y (d) recolectar la proteína deseada segregada de la célula hospedera.

En una incorporación, la primera secuencia de ácido nucleico comprende además una secuencia enzimática entre la secuencia señal y la secuencia de la proteína deseada. Por ejemplo, la secuencia enzimática se obtiene de una enzima glucoamilasa o CBH1. En una incorporación, la secuencia enzimática es una secuencia enzimática de longitud total que comprende un dominio catalítico, un conector, y un dominio de unión. En otra incorporación, la secuencia enzimática comprende una secuencia de dominio catalítico, la cual se conecta a la secuencia de la proteína deseada mediante un conector. En algunas incorporaciones, la enzima es una proteína hospedera que se expresa y/o se segrega altamente en su hospedero natural.

La primera secuencia de ácido nucleico comprende además un promotor corriente arriba de una secuencia señal. En una incorporación, el promotor es nativo a la célula hospedera y no está naturalmente asociado con la secuencia de la proteína deseada.

La segunda secuencia de ácido nucleico se conecta operacionalmente a un promotor. En una incorporación, el promotor es nativo a la célula hospedera y no está naturalmente asociado con la segunda secuencia de ácido nucleico.

Expresión Incrementada de Proteínas

La presente invención provee un método para la producción de una proteína deseada en una célula hospedera. La producción de proteína se aumenta por la inclusión de una señal secretora (v.g., el péptido señal NSP24 o el péptido señal CBH1) en combinación con la co-expresión de una proteína chaperona, chaperonina, y/o foldasa. En algunas incorporaciones, la señal secretora es de una proteína hospedera de Ascomycetes. En alguna incorporación, la proteína

deseada se fusiona al dominio catalítico de una enzima.

La presente invención ofrece ventajas significativas, especialmente en vista del hecho de que puede ser difícil producir grandes cantidades de proteínas de otras familias de hongos en hospederos Ascomycetes. Por supuesto, los expertos en el arte saben que, a menudo, es difícil producir cualquier proteína fúngica heteróloga en hospederos fúngicos o bacterianos. La presente invención provee métodos y composiciones adecuados para la producción de cualquier proteína apropiada en un hospedero fúngico o bacteriano conveniente. En algunas incorporaciones, el hospedero fúngico es un Ascomycetes y la proteína es una proteína de Basidiomycetes, mientras que en otras incorporaciones, el hospedero fúngico es un Basidiomycetes y la proteína es una proteína de Ascomycetes.

En algunas incorporaciones, la presente invención provee métodos para aumentar la expresión y/o secreción de una proteína en un hospedero usando un péptido señal hospedero en combinación con la co-expresión de una o más chaperonas o foldasas del mismo organismo como la fuente de la proteína. Por consiguiente, en algunas incorporaciones, una proteína heteróloga de Ascomycetes se expresa en un hospedero Basidiomycetes usando un péptido señal de hospedero Basidiomycetes y una chaperona de Ascomycetes. En algunas incorporaciones alternativas, una proteína heteróloga de Basidiomycetes se expresa en un hospedero Ascomycetes usando un péptido señal de Ascomycetes y una chaperona de Ascomycetes o Basidiomycetes. En algunas incorporaciones, el hospedero Ascomycetes es un miembro del género *Trichoderma*. En algunas incorporaciones, el *Trichoderma* es *Trichoderma reesei*, incluyendo varias cepas de *T. reesei*. En algunas incorporaciones alternativas, el Basidiomycetes es un miembro del género *Cerrena*, incluyendo, pero no limitado a, *C. unicolor*.

En algunas incorporaciones de la presente invención, la expresión y/o secreción de una proteína deseada se aumentan fusionando la proteína a una enzima hospedera en combinación con la co-expresión exógena de una o más chaperonas del mismo organismo que la proteína deseada. La co-expresión se lleva a cabo o bien mediante el mismo plásmido, o bien mediante plásmidos separados.

En incorporaciones todavía adicionales, la expresión y/o secreción de una proteína deseada se incrementa conectando la proteína al dominio catalítico de una enzima hospedera, en combinación con conectando operacionalmente la proteína a una secuencia señal hospedera, y co-expresión exógena de una o más chaperonas, chaperoninas, y/o foldasas, preferiblemente del mismo organismo que la proteína.

Se contempla que elementos enumerados en varias incorporaciones provistas en esta especificación encontrarán aplicación en cualquier combinación apropiada. Por consiguiente, no se pretende que las incorporaciones se limiten a las enumeraciones específicas provistas en esta especificación, ya que aspectos de las varias incorporaciones hallan aplicación en combinación unas con otras.

Péptidos Señal

El péptido señal específico usado en la presente invención no es crítico, siempre y cuando el péptido señal sea operacional en el hospedero. Se provee un "péptido señal operacional" si el péptido señal incrementa la secreción de una proteína cuando se conecta operacionalmente a la proteína en una célula hospedera. En algunas incorporaciones, el péptido señal se obtiene de una proteína fuertemente segregada y/o es un péptido señal fuerte. Un "péptido señal fuerte" resulta cuando la proteína natural es fuertemente segregada por su hospedero natural. En algunas incorporaciones, el péptido señal se obtiene de un organismo dentro del mismo filum que la célula hospedera. Por supuesto, en algunas incorporaciones, esto es ventajoso. En algunas incorporaciones, el péptido señal y la célula hospedera son del mismo género, mientras que en algunas incorporaciones adicionales, el péptido señal y la célula hospedera son de la especie. Por ejemplo, en algunas incorporaciones, la célula hospedera es una célula hospedera de Ascomycetes y el péptido señal se obtiene de Ascomycetes. En algunas incorporaciones, la célula hospedera es una *Trichoderma* y el péptido señal es de una *Trichoderma*. En algunas incorporaciones, la célula hospedera es *T. reesei* y el péptido señal se obtiene de *T. reesei*. En algunas incorporaciones, el péptido señal es un péptido señal fuerte. En algunas incorporaciones alternativas, la célula hospedera es una célula hospedera de Basidiomycetes y el péptido señal se obtiene de Basidiomycetes. Algunos ejemplos de péptidos señal que

encuentran aplicación en la presente invención incluyen, pero no se limitan a los péptido señal CBH1 y NSP24. Aunque los péptidos señal pueden funcionar en otros miembros de un filum tal como Ascomycetes, en algunas incorporaciones, péptidos señal encuentran aplicación óptima cuando se emplean en el género del cual se obtuvieron (*es decir*, para facilitar una secreción fuerte).

Tal como se emplea en esta especificación, una "proteína fuertemente segregada" es cualquier proteína que forma una cantidad significativa de la proteína total segregada de la célula. La proteína total segregada de la célula se denomina también "proteína extracelular." Por ejemplo, una proteína fuertemente segregada incluye por lo menos alrededor de 2% de la proteína extracelular, por lo menos alrededor de 3%, por lo menos alrededor de 4%, por lo menos alrededor de 5%, por lo menos alrededor de 6%, por lo menos alrededor de 7%, por lo menos alrededor de 8%, por lo menos alrededor de 9%, por lo menos alrededor de 10%, por lo menos alrededor de 15%, por lo menos alrededor de 20%, por lo menos alrededor de 30%, por lo menos alrededor de 40%, por lo menos alrededor de 50%, por lo menos alrededor de 60%, por lo menos alrededor de 70%, por lo menos alrededor de 75%, por lo menos alrededor de 80%, por lo menos alrededor de 85%, por lo menos alrededor de 90%, por lo menos alrededor de 95%, o por lo menos alrededor de 99%. En algunas incorporaciones, la proteína fuertemente segregada comprende por lo menos alrededor de 5% de la proteína extracelular en el sobrenadante de cultivo.

Péptidos Señal CBH1, Conectores y Dominios Catalíticos

Trichoderma reesei produce varias enzimas celulasas, incluyendo celobiohidrolasa I (CBH1), que se pliegan en dos dominios separados (*es decir*, los dominios catalítico y de unión) que están separados por una región conectora extendida. Polipéptidos extraños se han segregado en *T. reesei* como fusiones con el dominio catalítico más región conectora de CBH1 (Véase *v.g.*, Nyysönen *et al.*, Bio/Technol. 11:591-595 [1993]). *T. longibrachiatum* produce asimismo una CBH1 que encuentra aplicación en fusiones, así como en el aislamiento de un péptido señal y/o un conector. Los conectores encuentran aplicación en conectar un dominio catalítico de una enzima y el polipéptido deseado. Cualquier conector apropiado encuentra aplicación en la presente invención, siempre que forme un

distanciador semi-rígido, extendido, entre dominios plegados independientemente. Tales regiones conectoras se encuentran en varias proteínas, especialmente hidrolasas (v.g., celulasas fúngicas y bacterianas y hemicelulasas; Véase v.g., Libby *et al.*, Protein Engineering, Design and Selection (1994) vol. 7, 1109-1114).

Como se muestra en Figura 9, para CBH1 (SEQ ID NO: 10), la secuencia señal comienza en el par de bases 210 y termina en el par de bases 260 (SEQ ID NO: 11). El núcleo catalítico comienza en el par de bases 261 al par de bases 1698 (SEQ ID NO: 12), incluyendo el intrón 1 (desde el par de bases 671 a 737) y el intrón 2 (desde el par de bases 1435 a 1497). La secuencia conectora comienza en el par de bases 1699 y termina en el par de bases 1770 (SEQ ID NO: 13). El dominio de unión de la celulosa comienza en el par de bases 1771 al par de bases 1878. La información de la secuencia y el dominio para CBH1 puede hallarse a través del sitio web de la organización ExPASy y se designa uniprot/P62694. Homólogos de CBH1 se han identificado en un número de otras especies *Trichoderma* así como otros hongos filamentosos y encuentran aplicación en la presente invención como apropiados.

Péptidos Señal NSP24 y Polinucleótidos

El gen NSP24 se aisló y secuenció de *T. reesei* (Véase v.g., la Patente de los Estados Unidos No. 7,429,476, la cual se incorpora en esta especificación por referencia en su totalidad). El secuenciamiento de este gen identificó una secuencia que codifica un marco de lectura abierto de 407 aminoácidos (SEQ ID NO: 8), como se muestra en Figura 8. Un péptido señal se identificó como los primeros 20 aminoácidos (MQTFGAFLVSFLAASGLAAA; SEQ ID NO: 9) de SEQ ID NO: 8. Homólogos NSP24 se han identificado en un número de otras especies *Trichoderma* así como otros hongos filamentosos y encuentran aplicación en la presente invención como apropiados. En algunas incorporaciones, la secuencia señal NSP24 se usa en un organismo Ascomycetes. En algunas incorporaciones, la secuencia se usa en *Trichoderma spp.*, y, en algunas incorporaciones aún más particularmente, en *T. reesei*.

Por consiguiente, la presente invención provee péptidos señal proteasa de la familia NSP24 que encuentran aplicación en segregar una proteína. En algunas

incorporaciones, el péptido señal NSP24 se designa "péptido señal proteasa aspártica NSP24."

Polinucleótidos de la Invención

La presente invención provee varios polinucleótidos, incluyendo pero no limitados a polinucleótidos que codifican proteínas deseadas, péptidos señal, dominios catalíticos, conectores, chaperonas, chaperoninas y foldasas. En algunas incorporaciones, los polinucleótidos comprenden por lo menos dos de los anteriores. En todavía otras incorporaciones, los polinucleótidos de la presente invención comprenden por lo menos tres de los anteriores.

En algunas incorporaciones, los polinucleótidos codifican proteínas que comprenden por lo menos una sustitución de aminoácido tal como una "sustitución conservadora de aminoácido" usando L-aminoácidos, en donde un aminoácido se reemplaza por otro aminoácido biológicamente similar. Sustituciones conservadoras de aminoácido son aquellas que preservan la carga general, hidrofobicidad/hidrofiliidad, y/o el volumen estérico ("**steric bulk**") del aminoácido que se esta sustituyendo. Ejemplos de sustituciones conservadoras son aquellas entre los siguientes grupos: Gly/Ala, Val/Ile/Leu, Lys/Arg, Asn/Gln, Glu/Asp, Ser/Cys/Thr, y Phe/Trp/Tyr. En algunas incorporaciones, "proteínas derivadas" encuentran aplicación en la presente invención. En algunas de estas incorporaciones, las proteínas derivadas difieren sólo alrededor de 1 a alrededor de 10 residuos de aminoácido, tal como alrededor de 6 a alrededor de 10, sólo alrededor de 5, sólo alrededor de 4, alrededor de 3, alrededor de 2, o incluso 1 residuo de aminoácido, comparadas con la secuencia de la proteína "progenitora". La Tabla 1 presenta sustituciones conservadoras ejemplares de aminoácido, reconocidas en el arte. En incorporaciones adicionales, la sustitución involucra una o más sustituciones no conservadoras de aminoácido, deleciones, o inserciones que no suprimen la actividad del péptido señal.

Tabla 1. Sustituciones Conservadoras de Aminoácido

Para el Aminoácido	Código de una Letra	Reemplazar con Cualquiera de los Siguietes
Alanina	A	D-Ala, Gly, beta-Ala, L-Cys, D-Cys
Arginina	R	D-Arg, Lys, D-Lys, homo-Arg, D-homo-Arg, Met, Ile, D-Met, D-Ile, Orn, D-Orn
Asparagina	N	D-Asn, Asp, D-Asp, Glu, D-Glu, Gin, D-Gln
Ácido Aspártico	D	D-Asp, D-Asn, Asn, Glu, D-Glu, Gin, D-Gln
Cisteína	C	D-Cys, S-Me-Cys, Met, D-Met, Thr, D-Thr
Glutamina	Q	D-Gln, Asn, D-Asn, Glu, D-Glu, Asp, D-Asp
Ácido Glutámico	E	D-Glu, D-Asp, Asp, Asn, D-Asn, Gin, D-Gln
Glicina	G	Ala, D-Ala, Pro, D-Pro, b-Ala, Acp
Isoleucina	I	D-Ile, Val, D-Val, Leu, D-Leu, Met, D-Met
Leucina	L	D-Leu, Val, D-Val, Leu, D-Leu, Met, D-Met
Lisina	K	D-Lys, Arg, D-Arg, homo-Arg, D-homo-Arg, Met, D-Met, Ile, D-Ile, Orn, D-Orn
Metionina	M	D-Met, S-Me-Cys, Ile, D-Ile, Leu, D-Leu, Val, D-Val
Fenilalanina	F	D-Phe, Tyr, D-Thr, L-Dopa, His, D-His, Trp, D-Trp, Trans-3,4, or 5-fenilprolina, cis-3, 4, o 5-fenilprolina
Prolina	P	D-Pro, ácido L-l-tioazolidin-4-carboxílico, ácido D- o L-1-oxazolidin-4-carboxílico
Serina	S	D-Ser, Thr, D-Thr, allo-Thr, Met, D-Met, Met(O), D-Met(O), L-Cys, D-Cys
Treonina	T	D-Thr, Ser, D-Ser, allo-Thr, Met, D-Met, Met(O), D-Met(O), Val, D-Val
Tirosina	Y	D-Tyr, Phe, D-Phe, L-Dopa, His, D-His
Valina	V	D-Val, Leu, D-Leu, Ile, D-Ile, Met, D-Met

En algunas incorporaciones, los polinucleótidos de la invención son secuencias nativas. En algunas incorporaciones, las secuencias nativas se aíslan de la naturaleza, mientras que en otras incorporaciones aquellas se producen por medios recombinantes o sintéticos. El término "secuencia nativa" abarca específicamente formas truncadas o segregadas, que ocurren naturalmente, (v.g.,

fragmentos biológicamente activos), y formas variantes, que ocurren naturalmente, de las secuencias nativas.

Debido a la degeneración del código genético, más de un codón puede usarse para un aminoácido particular. Por lo tanto, en algunas incorporaciones, se emplean diferentes secuencias de ADN para codificar cualquiera de los polipéptidos tales como el péptido señal, la proteína, el dominio catalítico, y/o las chaperonas. Por supuesto, se pretende que la presente invención abarque diferentes secuencias polinucleotídicas que codifican el mismo polipéptido.

Un ácido nucleico es hibridable a otra secuencia de ácido nucleico si una forma de una sola hebra del ácido nucleico puede empalmarse al otro ácido nucleico bajo condiciones apropiadas de temperatura y fuerza iónica de la solución. Las condiciones de hibridación y lavado son bien conocidas en el arte para hibridación bajo condiciones de baja, media, alta y muy alta restricción. En general, la hibridación involucra una sonda nucleotídica y una secuencia de ADN homóloga que forman híbridos de doble hebra con emparejamiento extensivo de bases de polinucleótidos complementarios. En algunas incorporaciones, el filtro con la sonda y la secuencia homóloga se lavan en 2x cloruro de sodio/citrato de sodio (SSC), 0.5% SDS a alrededor de 60°C (restricción media), 65°C (restricción media/alta), 70°C (restricción alta) y alrededor de 75°C (restricción muy alta) (Véase v.g., Current Protocols in Molecular Biology, John Wiley & Sons, New York, 1989, 6.3.1 - 6.3.6, incorporado en esta especificación por referencia);

La presente invención comprende variaciones alélicas, mutantes naturales, mutantes inducidos, proteínas codificadas por ADN que se hibrida, bajo condiciones de alta o baja restricción, a un ácido nucleico que codifica una laccasa, una secuencia señal de NSP24, una secuencia señal de CBH1, dominios catalíticos, chaperonas, chaperoninas y foldasas. Ácidos nucleicos y polipéptidos de la presente invención incluyen aquellos que difieren de las secuencias divulgadas en esta especificación en virtud de errores de secuenciamiento en las secuencias divulgadas.

La "homología de secuencias de ADN" se determina por el grado de identidad entre dos secuencias de ADN. La homología o "identidad porcentual" se determina a menudo para secuencias polipeptídicas y/o secuencias nucleotídicas

usando programas de computador. Métodos para llevar a cabo el alineamiento de secuencias y determinar identidad de secuencia son bien conocidos por el artesano experto, pueden efectuarse sin experimentación excesiva, y cálculos de valores de identidad pueden obtenerse con certeza. Se dispone de un número de algoritmos, y son conocidos por los expertos en el arte, para alinear secuencias y determinar la identidad de secuencia. Programas computarizados que utilizan estos algoritmos se encuentran igualmente disponibles y son conocidos por los expertos en el arte, incluyendo, pero no limitados a: software ALIGN o Megalign (DNASTAR), o WU-BLAST-2, GAP, BESTFIT, BLAST, FASTA, TFASTA, y CLUSTAL. Los expertos en el arte saben cómo determinar parámetros apropiados para medir el alineamiento, incluyendo algoritmos necesarios para lograr el alineamiento máximo en toda la longitud de las secuencias que se están comparando. La identidad de secuencia puede determinarse usando los parámetros predefinidos ("**the default parameters**") establecidos por el programa. En algunas incorporaciones, la identidad de secuencia se determina por el algoritmo de búsqueda de homología de Smith-Waterman (Smith Waterman, Meth. Mol. Biol., 70:173-187 [1997]) como se implementa en el programa MSPRCH (Oxford Molecular) usando una búsqueda de brechas ("*gap*") con los siguientes parámetros de búsqueda: penalización por brecha abierta 12, y penalización por extensión de brecha 1.

Comparaciones de aminoácidos pareados pueden llevarse a cabo usando el programa GAP del paquete de software para análisis de secuencias GCG de Genetics Computer Group, Inc. (Madison, WI), empleando la matriz de sustitución de aminoácidos blosum62, con una ponderación de brecha 12 y una ponderación de longitud 2. Con respecto al alineamiento óptimo de dos secuencias de aminoácidos, el segmento contiguo de la secuencia de aminoácidos variante puede tener residuos adicionales de aminoácido o residuos delecionados de aminoácido con respecto a la secuencia aminoácidos de referencia. El segmento contiguo usado para comparación con la secuencia de aminoácidos de referencia incluirá por lo menos alrededor de 20 residuos contiguos de aminoácido, y puede ser alrededor de 30, alrededor de 40, alrededor de 50, o más residuos de aminoácido. En algunas incorporaciones, correcciones para una mayor identidad de secuencia, asociadas con la inclusión de brechas en la secuencia de aminoácidos de la derivada se hacen asignando penalizaciones por brecha.

En algunas incorporaciones, la proteína, el péptido señal, el dominio catalítico de la enzima, la chaperona, la chaperonina, y/o foldasa abarcadas por la invención se derivan de una bacteria o un hongo, tal como un hongo filamentoso. Hongos filamentosos ejemplares incluyen *Aspergillus spp.* y *Trichoderma spp.* Una *Trichoderma spp.* ejemplar es *T. reesei*. No obstante, en algunas incorporaciones, el péptido señal y/o el ADN que codifica el péptido señal provistos por la presente invención se derivan de otro género o especie de hongos, incluyendo pero no limitados a *Absidia spp.*; *Acremonium spp.*; *Agaricus spp.*; *Anaeromyces spp.*; *Aspergillus spp.*, incluyendo, pero no limitados a *A. aculeatus*, *A. awamori*, *A. flavus*, *A. foetidus*, *A. fumaricus*, *A. fumigatus*, *A. nidulans*, *A. niger*, *A. oryzae*, *A. terreus* y *A. versicolor*; *Aeurobasidium spp.*; *Cerrena spp.*; *Cephalosporium spp.*; *Cephalosporium spp.*; *Chaetomium spp.*; *Coprinus spp.*; *Dactyllum spp.*; *Dactylium spp.*; *Fusarium spp.*, incluyendo *F. conglomerans*, *F. decemcellulare*, *F. javanicum*, *F. lini*, *F. oxysporum* y *F. solani*; *Gliocladium spp.*; *Humicola spp.*, incluyendo *H. insolens* y *H. lanuginosa*; *Mucor spp.*; *Neurospora spp.*, incluyendo *N. crassa* y *N. sitophila*; *Neocallimastix spp.*; *Orpinomyces spp.*; *Penicillium spp.*; *Phanerochaete spp.*; *Phlebia spp.*; *Piromyces spp.*; *Rhizopus spp.*; *Schizophyllum spp.*; *Stachybotrys spp.*; *Trametes spp.*; *Trichoderma spp.*, incluyendo *T. reesei*, *T. reesei (longibrachiatum)* y *T. viride*; y *Zygorhynchus spp.*

Fusión de Dominio Catalítico

Fusionar una proteína deseada a una enzima permite a menudo una expresión y/o secreción incrementadas de la proteína deseada. En general, la secuencia enzimática se da corriente arriba de la secuencia de proteína deseada en el constructo. Por ejemplo, la enzima se obtiene de una enzima glucoamilasa o CBH1. En una incorporación, la secuencia enzimática es una secuencia enzimática de longitud total que comprende un dominio catalítico, un conector, y un dominio de unión. En otra incorporación, la secuencia enzimática comprende una secuencia de dominio catalítico, la cual se conecta a la secuencia de la proteína deseada por un conector o una porción del conector. En algunas incorporaciones, la enzima es una proteína hospedera que está altamente expresada y/o segregada en su hospedero natural. Por ejemplo, cuando la célula hospedera es un célula hospedera *Trichoderma*, la enzima es de una proteína de

Trichoderma. No obstante, hay que entender que muchas proteínas fúngicas filamentosas encuentran aplicación en la fusión a proteínas y que pueden utilizarse con éxito en otros hospederos fúngicos filamentosos.

Chaperonas, Chaperoninas y Foldasas

La chaperona, chaperonina, y/o foldasa específica, usada en los métodos y polinucleótidos incluidos en la invención no es crítica. Más aún, cuando se describen los usos de chaperona, chaperonina, y/o foldasa en esta especificación, se usan indistintamente en un método. Por ejemplo, cuando se describe un método que utiliza una chaperona, hay que entender que podrían emplearse una foldasa y/o chaperonina en lugar, o en adición, de la chaperona enumerada. Chaperona, chaperonina, y/o foldasa apropiadas para esta invención son aquellas que son activas en una célula hospedera y actúan para aumentar la expresión de la proteína deseada.

En algunas incorporaciones, la chaperona, chaperonina, y/o foldasa es del mismo filum de organismos que la proteína, y puede ser del mismo género, y puede ser también del mismo género o especie. En algunas incorporaciones, la chaperona, chaperonina, y/o foldasa es de un Basidiomycete y la proteína es una proteína de Basidiomycetes. En algunas incorporaciones, la chaperona, chaperonina, y/o foldasa se usan en combinación. En algunas incorporaciones, pueden emplearse fragmentos de chaperona, chaperonina, y/o foldasa que tienen sustancialmente la misma función que la chaperona, chaperonina, y/o foldasa de longitud total. Chaperona, chaperonina, y/o foldasa ejemplares incluyen las divulgadas en la solicitud de Patente de los Estados Unidos 60/919,332 y WO 2008/115596, las cuales se incorporan en esta especificación por referencia en su totalidad. Chaperona, chaperonina, y/o foldasa ejemplares incluyen, pero no se limitan a: BIP1, CLX1, ERO1, LHS1, PRP3, PRP4, PRP1, TIG1, PDI1, PPI1, PPI2, SCJ1, ERV2, EDEM, y SIL1. La Tabla 2 presenta un número de las secuencias para chaperona, chaperonina, y/o foldasa utilizables en la invención.

Tabla 2: Secuencias Ejemplares de Ácido Nucleico y Polipeptídicas de Proteínas que Mejoran la Secreción		
Proteína	Secuencia Ejemplar de Ácido Nucleico	Secuencia Ejemplar Polipeptídica
BIP1	SEQIDNO:15	SEQ ID NO: 30
CLX1	SEQIDNO:16	SEQ ID NO: 31
ERO1	SEQIDNO: 17	SEQ ID NO: 32
LHS1	SEQIDNO: 18	SEQ ID NO: 33
PRP3	SEQIDNO: 19	SEQ ID NO: 34
PRP4	SEQ ID NO: 20	SEQ ID NO: 35
PRP1	SEQIDNO: 21	SEQ ID NO: 36
TIG1	SEQ ID NO: 22	SEQ ID NO: 37
PDI1	SEQ ID NO: 23	SEQ ID NO: 38
PPI1	SEQ ID NO: 24	SEQ ID NO: 39
PPI2	SEQ ID NO: 25	SEQ ID NO: 40
SCJ1	SEQ ID NO: 26	SEQ ID NO: 41
ERV2	SEQ ID NO: 27	SEQ ID NO: 42
EDEM	SEQ ID NO: 28	SEQ ID NO: 43
SIL1	SEQ ID NO: 29	SEQ ID NO: 44

Biología Molecular - Promotores y Vectores de Expresión

La presente invención utiliza técnicas rutinarias en el campo de la genética recombinante, bien conocidas por los expertos en el arte. En algunas incorporaciones, la presente invención provee genes heterólogos que comprenden secuencias promotoras de genes (v.g., de hongos filamentosos), que típicamente se clonan en vectores intermedios antes de la transformación en células hospederas (v.g., células de *Trichoderma reesei*) para replicación y/o expresión. Estos vectores intermedios son típicamente vectores procariotas (v.g., plásmidos, o vectores lanzadera).

En general, la expresión de una proteína deseada se lleva a cabo bajo cualquier promotor apropiado. En una incorporación, un promotor no nativo a un hospedero se conecta operacionalmente a un polinucleótido que codifica una proteína deseada que es nativa o no nativa a un hospedero. En otra incorporación,

un promotor nativo a un hospedero se conecta operacionalmente un polinucleótido que codifica una proteína deseada, la cual es nativa o no nativa a un hospedero. En algunas incorporaciones, la proteína deseada se expresa bajo un promotor heterólogo, el cual no se asocia naturalmente con el gen de la proteína deseada. Aunque en algunas otras incorporaciones, la proteína deseada se expresa bajo un promotor constitutivo o inducible. En algunas incorporaciones, la proteína deseada se expresa en un sistema de expresión de *Trichoderma* con un promotor de celulasa (v.g., el promotor CBH1).

Tal como se emplea en esta especificación, el término "promotor" se refiere a una secuencia de ácido nucleico que funciona para dirigir la transcripción de un gen corriente abajo. Un promotor puede incluir secuencias de ácido nucleico necesarias cerca del sitio de iniciación de la transcripción, tal como, en el caso de un promotor tipo polimerasa II, un elemento TATA. El promotor, junto con otras secuencias reguladoras transcripcionales y translacionales, denominadas colectivamente "secuencias reguladoras", controla la expresión de un gen. En general, las secuencias reguladoras incluyen, pero no se limitan a, secuencias promotoras, sitios de unión ribosómica, secuencias de inicio y parada transcripcionales, secuencias de inicio y parada translacionales, y secuencias potenciadoras o activadoras. Las secuencias reguladoras son generalmente apropiadas para, y reconocidas por, el hospedero en el cual se está expresando el gen corriente abajo. En algunas incorporaciones, el promotor usado es del mismo filum que la célula hospedera y, en otra incorporación, el promotor es del mismo género que la célula hospedera y, en algunas incorporaciones, del mismo género y especie que la célula hospedera.

Un "promotor constitutivo" es un promotor que es activo bajo la mayoría de las condiciones ambientales y de desarrollo. Un "promotor inducible" o "reprimible" es un promotor que es activo bajo regulación ambiental o de desarrollo. En algunas incorporaciones, los promotores son inducibles o reprimibles debido a cambios en factores ambientales incluyendo, pero no limitados a, disponibilidad de carbono, nitrógeno u otros nutrientes, temperatura, pH, osmolaridad, la presencia de metal(es) pesado(s), la concentración de inhibidor(es), estrés, o una combinación de los anteriores, como se conoce en el arte. En algunas otras incorporaciones, los promotores son inducibles o reprimibles por factores

metabólicos, tales como el nivel de ciertas fuentes de carbono, el nivel de ciertas fuentes de energía, el nivel de ciertos catabolitos, o una combinación de los anteriores, como se conoce en el arte.

Ejemplos apropiados no limitantes de promotores incluyen *cbh1*, *cbh2*, *egl1*, *egl2*, *egl3*, *egl4*, *egl5*, *xyn1*, y *xyn2*, el promotor del gen reprimible de la fosfatasa ácida (*phoA*) de *P. chrysogenum* (Véase, Graessle *et al.*, Appl. Environ. Microbiol., 63:753-756 [1997]), el promotor PCK1 reprimible por la glucosa (Véase, Leuker *et al.*, Gene 192:235-240 [1997]), el promotor MRP1 inducible por la maltosa, reprimible por la glucosa (Véase, Munro *et al.*, Mol. Microbiol., 39 1414-1426 [2001]), el promotor MET3 reprimible por la metionina (Véase, Liu *et al.*, Eukary. Cell 5:638-649 [2006]), el promotor *pKi*, y el promotor *cpc1*.

En algunas incorporaciones de la presente invención, el promotor en el constructo del gen reportero es un promotor sensible a la temperatura. En algunas incorporaciones, la actividad del promotor sensible a la temperatura se reprime por temperatura elevada. En algunas incorporaciones, el promotor es un promotor reprimido por catabolito. En algunas incorporaciones, el promotor es reprimido por cambios en la osmolaridad. En algunas incorporaciones, el promotor es inducible o reprimible por los niveles de polisacáridos, disacáridos, o monosacáridos presentes en el medio de cultivo.

Un ejemplo de un promotor inducible que encuentra aplicación en la presente invención es el promotor *cbh1* de *T. reesei*, cuya secuencia nucleotídica se halla depositada en GenBank bajo el Número de Acceso D86235. Otros promotores ejemplares incluyen promotores involucrados en la regulación de genes que codifican enzimas celulasa, incluyendo, pero no limitados a, *cbh2*, *egl1*, *egl2*, *egl3*, *egl5*, *xyn1* y *xyn2*.

En algunas incorporaciones de la presente invención, con el fin de obtener altos niveles de expresión de un gen clonado, el gen heterólogo se posiciona ventajosamente alrededor de la misma distancia del promotor que en el gen que ocurre naturalmente. Sin embargo, como se conoce en el arte, cierta variación en la distancia puede acomodarse sin pérdida de la función promotora.

En algunas incorporaciones, un promotor natural modificado por reemplazo, sustitución, adición o eliminación de uno o más nucleótidos encuentra aplicación en la presente invención, siempre y cuando las modificaciones no cambien la función del promotor. Por supuesto, se pretende que la presente invención abarque y no se restrinja por tales alteraciones al promotor.

El vector/construido de expresión contiene normalmente una unidad de transcripción o casete de expresión que contiene la totalidad de los elementos adicionales requeridos para la expresión de la secuencia heteróloga. Por consiguiente, un típico casete de expresión contiene un promotor conectado operacionalmente a la secuencia heteróloga de ácido nucleico y a las señales requeridas para una poliadenilación eficiente del transcripto, sitios de unión al ribosoma, y terminación de la traducción. Elementos adicionales en el casete pueden incluir potenciadores y, si se emplea ADN genómico como el gen estructural, intrones con sitios donantes y aceptores de empalme funcionales, péptidos líder de secreción, secuencias líder, conectores, y sitios de escisión.

La práctica de la presente invención no está restringida por la elección del promotor en el constructo genético. Como se indica más arriba, promotores ejemplares son los promotores *chb1*, *cbh2*, *eg1*, *eg2*, *eg3*, *eg5*, *xln1* y *xln2* de *Trichoderma reesei*. Promotores adicionales que encuentran aplicación en la presente invención incluyen los de los genes de glucoamilasa de *A. awamori* y *A. niger* (*glaA*) (Véase, Nunberg *et al.*, Mol. Cell Biol., 4:2306-2315 [1984]) y el promotor de acetamidasa de *A. nidulans*. Un promotor ejemplar para vectores utilizados en *Bacillus subtilis* es el promotor *AprE*; un promotor ejemplar usado en *E. coli* es el promotor *Lac*, un promotor ejemplar usado en *Saccharomyces cerevisiae* es *PGK1*, un promotor ejemplar usado en *Aspergillus niger* es *glaA*, y un promotor ejemplar para *Trichoderma reesei* es *cbh1*. Sin embargo, no se pretende que la presente invención se limite a estas células específicas ni a estos promotores específicos, ya que otras células y promotores encuentran aplicación en varias incorporaciones.

En algunas incorporaciones, además de una secuencia promotora, el casete de expresión incluye igualmente una región de terminación de transcripción corriente abajo del gen estructural para proporcionar una terminación eficiente. En

algunas incorporaciones, la región de terminación se obtiene del mismo gen que la secuencia promotora, mientras que en otras incorporaciones, se obtiene de genes diferentes.

Aunque en la presente invención encuentra aplicación cualquier terminador fúngico funcional apropiado, algunos terminadores ejemplares incluyen, pero no se limitan, al terminador del gen *trpC* de *Aspergillus nidulans* (véase, Yelton *et al.*, Proc. Natl. Acad. Sci. USA 81:1470-1474 (1984); Mullaney *et al.*, (Molecular Genetics and Genomics [MGG] 199:37-45 (1985)), los genes de glucoamilasa de *Aspergillus awamori* o *Aspergillus niger* (véase, Nunberg *et al.*, Mol. Cell. Biol., 4:2306 (1984); Boel *et al.*, EMBO J., 3:1581-1585 (1984)), el gen de la amilasa TAKA de *Aspergillus oryzae*, el gen de la carboxil proteasa de *Mucor miehei* (Publicación de Patente EP No. 0 215 594) y el gen CBH1 de *Trichoderma reesei*.

No se pretende que el vector de expresión utilizado para transportar la información genética dentro de la célula hospedera esté limitado a algún vector particular. Se contempla que cualquiera de los vectores convencionales empleados para expresión en células procariotas o eucariotas encontrará aplicación en la presente invención. Vectores bacterianos estándar de expresión incluyen, pero no se limitan a bacteriófagos λ y M13, así como plásmidos tales como los plásmidos a base de pBR322, pSKF, pET23D, y sistemas de expresión de fusión tales como MBP, GST, y LacZ. En algunas incorporaciones, se añaden identificaciones de epítipo a las proteínas recombinantes para proporcionar métodos convenientes de aislamiento (v.g., c-myc). Ejemplos de vectores apropiados de expresión y/o integración son bien conocidos por los expertos en el arte (véase, v.g., Bennett y Lasure (eds.) More Gene Manipulations in Fungi, Academic Press páginas 70-76 y páginas 396-428 (1991); Patente de los Estados Unidos No. 5,874,276. Varios proveedores comerciales (v.g., Promega, Invitrogen, etc.) suministran vectores útiles, como se conocen por los expertos en el arte. Algunos vectores específicos útiles incluyen, pero no se limitan a pBR322, pUC18, pUC100, pDON™201, pENTR™, pGEN®3Z y pGEN®4Z. No obstante, se pretende que la presente invención incluya otros factores de expresión que sirven funciones equivalentes y que son, o se hacen, conocidos en el arte. Por consiguiente, una amplia variedad de combinaciones de hospedero/factor de expresión encuentran aplicación en expresar las secuencias de ADN de la

presente invención. En algunas incorporaciones, vectores útiles de expresión comprenden segmentos de secuencias de ADN cromosómico, no cromosómico y/o sintético (v.g., varios derivados conocidos de SV40) y plásmidos bacterianos conocidos (v.g., plásmidos de *E. coli* incluyendo col E1, pCR1, pBR322, pMb9, pUC 19, pSL1180 y sus derivados), plásmidos de rango de hospedero más amplio (v.g., RP4), ADN de bacteriófago (v.g., los numerosos derivados del bacteriófago lambda, tales como NM989, y otros bacteriófagos de ADN, tales como M13, bacteriófagos filamentosos de ADN de una sola hebra), y plásmidos de levadura (V.g., el plásmido 2.mu o derivados del mismo).

En algunas incorporaciones, un vector de expresión incluye un marcador seleccionable. Ejemplos de marcadores seleccionables incluyen los que confieren resistencia antimicrobiana. Marcadores nutricionales encuentran también aplicación en la presente invención, incluyendo aquellos marcadores conocidos en el arte como *amdS*, *argB* y *pyr4*. Marcadores útiles para la transformación de *Trichoderma* son conocidos en el arte (Véase, v.g., Finkelstein, en *Biotechnology of Filamentous Fungi*, Finkelstein *et al.*, (eds.), Butterworth-Heinemann, Boston MA, capítulo 6 (1992)). En algunas incorporaciones, los vectores de expresión incluyen asimismo un replicón, un gen que codifica la resistencia a los antibióticos para permitir la selección de bacterias que albergan plásmidos recombinantes, y/o sitios únicos de restricción en regiones no esenciales del plásmido para permitir la inserción de secuencias heterólogas. Se pretende que cualquier gen apropiado de resistencia a los antibióticos encontrará aplicación en la presente invención. En algunas incorporaciones, en las cuales *T. reesei* es la célula hospedera, las secuencias procariotas se seleccionan preferiblemente de tal modo que no interfieran la replicación o la integración del ADN en *T. reesei*.

En algunas incorporaciones, un vector de expresión incluye un gen reportero solo o, de manera opcional, como una fusión con la proteína de interés. Ejemplos de genes reporteros incluyen, pero no se limitan a, reporteros fluorescentes, reporteros detectables por color (v.g., (3-galactosidasa), y reporteros biotinilados. En algunas incorporaciones, cuando la molécula reportera se expresa, se emplea para identificar si el péptido señal es activo en una célula hospedera. Si el péptido señal es activo, la molécula reportera se segrega de la célula. En algunas incorporaciones, inicialmente el péptido señal se conecta

operacionalmente al reportero, con el fin de identificar la secreción de una célula hospedera particular. Métodos alternativos, tales como los que usan anticuerpos específicos para la proteína de interés y/o el péptido señal, encuentran igualmente aplicación en determinar si la proteína de interés se segrega o no.

En algunas incorporaciones, los métodos de transformación de la presente invención resultan en la integración estable de la totalidad o parte del vector de transformación en el genoma de una célula hospedera, tal como una célula hospedera fúngica filamentosa. No obstante, se contempla también la transformación resultante en el mantenimiento de un factor de transformación extracromosómico auto-replicante.

Muchos métodos de transfección estándar encuentran aplicación en la presente invención para producir estirpes celulares fúngicas filamentosas y bacterianas (v.g., *Aspergillus* o *Trichoderma*), que expresan grandes cantidades de las proteínas. Métodos para la introducción de constructos de ADN en cepas productoras de celulasa de *Trichoderma* son bien conocidos por los expertos en el arte (véase v.g., Lorito *et al.*, Curr. Genet, 24:349-356 [1993]; Goldman *et al.*, Curr. Genet., 17:169-174 [1990]; Penttilä *et al.*, Gene 6: 155-164 [1987]; Patente de los Estados Unidos No. 6.022,725; Patente de los Estados Unidos No. 6,268,328; Nevalainen *et al.*, "The Molecular Biology of *Trichoderma* and its Application to the Expression of Both Homologous and Heterologous Genes" en Molecular Industrial Mycology, Leong and Berka (eds.), Marcel Dekker Inc., NY [1992] páginas 129-148; Yelton *et al.*, Proc. Natl. Acad. Sci. USA 81: 1470-1474 [1984]; Bajar *et al.*, Proc. Natl. Acad. Sci. USA 88: 8202-8212 [1991]; Fernandez-Abalos *et al.*, Microbiol., 149:1623 - 1632 [2003]; y Brigidi *et al.*, FEMS Microbiol. Lett., 55:135-138 [1990]).

No obstante, cualquiera de los procedimientos bien conocidos para introducir secuencias nucleotídicas extrañas en células hospederas encuentran aplicación en la presente invención. Estos métodos incluyen, pero no se limitan a, el uso de transfección de fosfato de calcio, polibreno, fusión de protoplasto, electroporación, biolística, liposomas, microinyección, vectores plasmídicos, vectores víricos y cualquiera de otros métodos bien conocidos para introducir ADN genómico clonado, ADNc, ADN sintético u otro material genético extraño en una célula hospedera, como es bien conocido por los expertos en el arte. Igualmente

útil es el método de transfección mediada por *Agrobacterium* (véase v.g., la Patente de los Estados Unidos No. 6,255,115). Es únicamente necesario que el procedimiento de ingeniería genética particular empleado sea capaz de introducir exitosamente por lo menos un gen en una célula hospedera que pueda expresar el gen. En algunas incorporaciones, la invención provee métodos para producir una proteína, que comprenden los pasos de introducir en una célula hospedera un polinucleótido que contiene un péptido señal NSP24 conectado a un ácido nucleico que codifica una proteína, cultivar la célula hospedera bajo condiciones apropiadas de cultivo para la expresión y producción de la proteína, y producir la citada proteína. En algunas incorporaciones, la proteína se segrega de la célula hospedera. En algunas incorporaciones alternativas, la presente invención provee métodos para producir una proteína, que comprenden los pasos de introducir en una célula hospedera un polinucleótido que contiene un péptido señal CBH1 conectado operacionalmente a un ácido nucleico que codifica una proteína, cultivar la célula hospedera bajo condiciones apropiadas de cultivo para la expresión y producción de la proteína, y producir la citada proteína. En algunas incorporaciones, la proteína se segrega de la célula hospedera.

Una vez el vector de expresión se introduce en las células hospederas, las células transfectadas o transformadas se cultivan bajo condiciones que favorezcan la expresión de genes bajo control de las secuencias promotoras del gen. En algunas incorporaciones, se cultivan grandes lotes de células transformadas. En algunas incorporaciones, el producto (*es decir*, la proteína) se cosecha de las células y/o se recupera del cultivo empleando técnicas estándar.

Por consiguiente, la invención en esta especificación permite la expresión y secreción mejoradas de polipéptidos deseados cuya secreción se intensifica por secuencias de péptido señal, secuencias de ADN de fusión, y varios constructos heterólogos, así como expresión de chaperonas, chaperoninas y/o foldasas. La invención provee asimismo procesos para expresar y segregar altos niveles de tales polipéptidos deseados.

Proteínas Deseadas

El término "proteína deseada" significa cualquier proteína de interés. La

proteína deseada puede ser una proteína nativa a una célula hospedera, o no nativa (heteróloga) a una célula hospedera. En algunas incorporaciones, la proteína deseada es una proteína fúngica. En algunas incorporaciones, el hospedero es un hospedero Ascomycete y la proteína es cualquier proteína distinta de una proteína de Ascomycetes. En algunas incorporaciones, el hospedero es un hospedero Basidiomycete y la proteína es cualquier proteína distinta de una proteína de Basidiomycete. En algunas incorporaciones, la proteína es cualquier proteína distinta de una proteína de *Trichoderma*. En algunas otras incorporaciones, la proteína es cualquier proteína distinta de una proteína de *Aspergillus*.

No se pretende que la presente invención se limite a cualquier tipo particular de proteína. Por supuesto, se pretende que la presente invención comprenda cualquier proteína de interés. Algunos ejemplos no limitantes de proteínas deseadas incluyen, pero no se limitan a glucoamilasas, α -amilasas, enzimas que hidrolizan almidón granular, celulasas, lipasas, xilanasas, cutinasas, hemicelulasas, proteasas, oxidasas, laccasas y combinaciones de las mismas.

En algunas incorporaciones, la glucoamilasa es una glucoamilasa tipo silvestre, obtenida de una fuente fúngica filamentosa, tal como una cepa de *Aspergillus*, *Trichoderma* o *Rhizopus*. No obstante, en otras incorporaciones, la glucoamilasa es una glucoamilasa modificada por técnicas de ingeniería genética (v.g., una variante de una glucoamilasa de *Aspergillus niger*). En algunas otras incorporaciones, las composiciones de la presente invención comprenden igualmente por lo menos una proteasa y por lo menos una α -amilasa. En algunas incorporaciones, la α -amilasa se obtiene de una fuente bacteriana (v.g., *Bacillus spp.*), o de una fuente fúngica (v.g., an *Aspergillus spp.*). En algunas incorporaciones, las composiciones incluyen asimismo enzimas, por lo menos una proteasa, y/o por lo menos una glucoamilasa, y/o por lo menos una α -amilasa. En algunas incorporaciones, la proteína es laccasa, tal como laccasa obtenida de Basidiomycetes, y en algunas incorporaciones, del género *Cerrena*, tal como *C. unicolor*. Fuentes comerciales de estas enzimas son conocidas y están disponibles de, por ejemplo, Genencor International, Inc. y Novozymes A/S.

Laccasas y Enzimas Relacionadas con Laccasa

En una incorporación preferida, laccasas y enzimas relacionadas con laccasa son proteínas deseadas. No se pretende que la presente invención se limite a alguna laccasa particular, ya que se incluye cualquier enzima laccasa en la clasificación de enzimas (EC 1.10.3.2). En algunas incorporaciones, las enzimas laccasas se obtienen de origen vegetal o microbiano. En algunas incorporaciones, las enzimas laccasas microbianas se derivan de bacterias u hongos (incluyendo hongos filamentosos y levaduras). Aunque no se pretende que la presente invención se limite a laccasas específicas, ejemplos apropiados incluyen laccases derivables de *Aspergillus*, *Neurospora* (v.g., *N. crassa*), *Podospora*, *Botrytis*, *Collybia*, *Cerrena*, *Stachybotrys*, *Panus*, (v.g., *Panus rudis*), *Thielavia*, *Fomes*, *Lentinus*, *Pleurotus*, *Trametes* (v.g., *T. villosa* y *T. versicolor*), *Rhizoctonia* (v.g., *R. solani*), *Coprinus* (v.g., *C. plicatilis* y *C. cinereus*), *Psatyrella*, *Myceliophthora* (v.g., *M. thermophila*), *Schytalidium*, *Phlebia* (v.g., *P. radita*; véase, v.g., WO 92/01046), *Coriolus* (v.g., *C. hirsutus*; véase, v.g., JP 2-238885), *Spongipellis*, *Polyporus*, *Ceriporiopsis subvermisporea*, *Ganoderma tsunodae* y *Trichoderma*.

En algunas incorporaciones, las laccasas incluyen laccasa *Cerrena* A1, B1 y D2 de la cepa CBS115.075, la laccasa *Cerrena* A2, B2, C, D1, y E de la cepa CBS154.29, la enzima laccasa *Cerrena* B3 de la cepa ATCC20013 (véase, v.g., Publicación de los Estados Unidos No. 2008/0196173, incorporada en esta especificación por referencia en su totalidad). Versiones adicionalmente optimizadas de estas laccasas encuentran también aplicación en la presente invención.

En otras incorporaciones, las laccasas incluyen la proteína madura de la laccasa *Cerrena* D expresada en *Trichoderma*; cuya secuencia de aminoácidos se muestra en la siguiente forma (SEQ ID NO: 45).

AIGPVADLHIVNKDLAPDGVQRPTVLAGGTFPGTLITGQKGNFQLNVIDDLTDDRM
 LTPTSIHWHGFFQKGTAWADGPAFVTQCPIADNSFLYDFDVPDQAGTFWYHSHLST
 QYCDGLRGAFVVYDPNDPHKDLVDVDDGGTVITLADWYHVLAQTVVGAATPDSTLI
 NGLGRSQTGPADAELAVISVEHNKRYRFRLLVSISCDPNFTFSVDGHNMTVIEVDGVN
 TRPLTVDSIQIFAGQRYSFVLNANQPEDNYWIRAMPNIGRNTTTLDGKNAAILRYKNA
 SVEEPKTVGGPAQSPLNEADLRPLVPAPVPGNAVPGGADINHRLNLTFSNGLFSINNA
 SFTNPSVPALLQILSGAQNAQDLLPTGSYIGLELGKVVELVIPPLAVGGPHPFHLHGHN
 FWVVRSGSDEYNFDDAILRDVVSIGAGTDEVTIRFVTDNPGPWFLHCHIDWHLEAG
 LAIVFAEGINQTAAANPTPQAWDELCPKYNGLSASQKVKPKKGTAI

Células Hospederas

La presente invención provee células hospederas transformadas con constructos de ADN y un vector, tal como se describe en esta especificación. En algunas incorporaciones, la presente invención posibilita células hospederas transformadas con constructos de ADN, que codifican una proteína deseada y conectada operacionalmente al péptido señal NSP24 o CBH1, tal como se describe en esta especificación. En algunas incorporaciones, la invención proporciona constructos de ADN, que codifican por lo menos una proteína deseada tal como proteasa, laccasa, α -amilasa, glucoamilasa, xilanas, y celulosa, en donde los constructos se introducen en una célula hospedera. En algunas incorporaciones, la presente invención posibilita la expresión de genes de proteínas y/o la sobre-expresión de genes de proteínas bajo el control de promotores génicos funcionales en células hospederas bacterianas y/o fúngicas.

Se pretende que cualesquiera células hospederas son útiles con la presente invención. No se pretende que la presente invención se limite a alguna célula hospedera particular. En algunas incorporaciones, la célula hospedera es una célula en la cual el péptido señal tiene actividad en segregar la proteína de interés. Por ejemplo, células hospederas, para las cuales encuentra aplicación un péptido señal de *T. reesei*, incluyen, pero no se limitan a, células fúngicas y bacterianas. Células hospederas incluyen células fúngicas filamentosas, incluyendo pero no limitadas a *Trichoderma* spp. (v.g., *T. viride* y *T. reesei*, el morfo asexual de *Hypocrea jecorina*, previamente clasificada como *T. longibrachiatum*), *Penicillium* spp., *Humicola* spp. (v.g., *H. insolens* y *H. grisea*),

Aspergillus spp. (v.g., *A. niger*, *A. nidulans*, *A. oryzae*, y *A. awamori*), *Fusarium* spp. (v.g., *F. gramineum*), *Neurospora* spp., *Hypocrea* spp. y *Mucor* spp. Células hospederas alternativas incluyen, pero no se limitan a, *Bacillus* spp (v.g., *B. subtilis*, *B. licheniformis*, *B. lentus*, *B. stearothermophilus* y *B. brevis*) y *Streptomyces* spp. (v.g., *S. coelicolor* y *S. lividans*).

En el arte se conocen muchos métodos para identificar si una proteína se segrega en una célula hospedera o permanece en el citoplasma. Se pretende que cualquier método apropiado encontrará aplicación en identificar células hospederas en las cuales la secuencia señal es activa.

Expresión de Proteína

Proteínas deseadas de la presente invención se producen cultivando células transformadas con un vector tal como un vector de expresión que contiene genes cuya secreción se intensifica por la secuencia del péptido señal NSP24 o CBH1, foldasas, chaperoninas, y/o chaperonas. La presente invención es particularmente útil para mejorar la producción intracelular y/o extracelular de proteínas. Como saben los expertos en el arte, las condiciones óptimas para la producción de las proteínas variarán con la elección de la célula hospedera y la proteína que se va a expresar. Tales condiciones se determinan fácilmente por los expertos en el arte.

En algunas incorporaciones, la proteína de interés se aísla o recupera y purifica después de la expresión. Varios métodos para el aislamiento y purificación de proteínas son conocidos por los expertos en el arte. Cualquier método apropiado encuentra aplicación en la presente invención. Por ejemplo, métodos de purificación estándar, que encuentran aplicación en la presente invención, incluyen, pero no se limitan a técnicas electroforéticas, moleculares, inmunológicas y cromatográficas, incluyendo intercambio iónico, hidrofóbicas, afinidad, y cromatografía HPLC de fase inversa, y cromatoenfocado. Por ejemplo, en algunas incorporaciones, la proteína de interés se purifica usando una columna de anticuerpo estándar que contiene anticuerpos dirigidos contra la proteína de interés. Técnicas de ultrafiltración y diafiltración, en conjunción con concentración de proteínas, encuentran igualmente aplicación en algunas incorporaciones. Como

es conocido por los expertos en el arte, el grado de purificación necesario varía dependiendo del uso de la proteína de interés. Por supuesto, en algunas incorporaciones, no es necesaria purificación.

En algunas incorporaciones, proteínas de interés producidas por células hospederas transformadas, como se proveen por la presente invención, se recuperan del medio de cultivo por procedimientos convencionales conocidos por los expertos en el arte. Estos métodos incluyen, pero no se limitan a separar las células hospederas del medio por centrifugación o filtración. En algunas incorporaciones, las células se disgregan, y se retira el sobrenadante de la fracción celular y los desechos. En algunas incorporaciones, los componentes proteináceos del sobrenadante o filtrado se precipitan por medio de una sal (v.g., sulfato de amonio) luego de la clarificación. Las proteínas precipitadas se solubilizan entonces y, en algunas incorporaciones, se purifican por cualquier método apropiado, incluyendo procedimientos cromatográficos (v.g., cromatografía de intercambio iónico, cromatografía de filtración en gel, cromatografía de afinidad, y otros procedimientos reconocidos en el arte).

En algunas incorporaciones adicionales, anticuerpos dirigidos contra los péptidos y proteínas producidos usando la presente invención se generan inmunizando un animal (v.g., un conejo o ratón), y recuperando anticuerpos anti-proteína y/o péptido señal NSP24, usando cualquier método apropiado conocido en el arte. En algunas incorporaciones adicionales, se producen anticuerpos monoclonales utilizando cualquier método apropiado conocido en el arte.

En algunas incorporaciones, análisis conocidos por las personas expertas en el arte encuentran aplicación en la presente invención, incluyendo, pero no limitados a, los descritos en WO 99/34011 y en la Patente de los Estados Unidos No. 6,605,458, las cuales se incorporan ambas por referencia en esta especificación en su totalidad.

Fusiones

En algunas incorporaciones, la proteína deseada se produce como una proteína de fusión. En algunas incorporaciones adicionales, la proteína deseada

se fusión a una proteína que se segrega eficientemente por un hongo filamentoso, y se fusiona al dominio catalítico de una enzima del mismo filum, género, y/o especies que la célula hospedera utilizada para expresión de la proteína de fusión. En algunas incorporaciones, la proteína deseada se fusiona a un polipéptido CBH1, o una porción del mismo. En algunas incorporaciones adicionales, la proteína deseada se fusiona a un polipéptido CBH1, o una porción del mismo, que se altera para minimizar o eliminar la actividad catalítica. En incorporaciones todavía adicionales, la proteína deseada se fusiona a un polipéptido de glucoamilasa de *Trichoderma*, o una porción del mismo. En algunas incorporaciones adicionales, la proteína deseada se fusiona a una glucoamilasa de *Trichoderma*, o una porción de la misma, que se altera para minimizar o eliminar actividad catalítica. En algunas incorporaciones adicionales, la proteína deseada se fusiona a un polipéptido para intensificar la secreción, facilitar la purificación subsiguiente y/o mejorar la estabilidad.

En general, el primero, segundo, y/o tercer polinucleótido en el hospedero de expresión de la presente invención se insertan genéticamente o se integran en la estructura genómica ("**the genomic makeup**") del hospedero de expresión (v.g., se integra en el cromosoma del hospedero de expresión). No obstante, en algunas incorporaciones, es extracromosómico (v.g., existe como un vector replicante en el hospedero de expresión). En algunas incorporaciones adicionales, el polinucleótido extracromosómico se expresa bajo condiciones apropiadas de selección para un marcador de selección que está presente en el vector).

Análisis de Nivel de Secreción

Tal como se describe en esta especificación, el nivel de secreción de un polipéptido deseado en el hospedero de expresión se determina empleando cualquier método apropiado. Por ejemplo, en algunas incorporaciones, el nivel de secreción se basa en varios factores (v.g., condiciones de crecimiento del hospedero), etc. Sin embargo, en algunas incorporaciones, el nivel de secreción del polipéptido deseado, expresado en el hospedero, es más alto que el nivel de secreción del polipéptido deseado, expresado sin la presencia de una proteína potenciadora de la secreción. En algunas incorporaciones, el nivel de secreción deseado de un polipéptido (v.g., laccasa de *Cerrena unicolor* en un hospedero de

expresión tal como *T. reesei*) es por lo menos alrededor de 1 mg/litro, alrededor de 2 mg/litro, alrededor de 3 mg/litro, alrededor de 4 mg/litro, o alrededor de 5 mg/litro cuando el hospedero es se cultiva en modo de fermentación en lotes en un matraz de agitación, o por lo menos alrededor de 50 mg/litro, alrededor de 100 mg/litro, alrededor de 150 mg/litro, alrededor de 200 mg/litro, alrededor de 250 mg/litro, alrededor de 500 mg/litro, alrededor de 1000 mg/litro, alrededor de 2.000 mg/litro, alrededor de 5.000 mg/litro, alrededor de 10.000 mg/litro o alrededor de 20.000 mg/litro cuando el hospedero se cultiva en un ambiente fermentador con pH controlado, tasa de alimentación, etc. (v.g., fermentación alimentada ("**fed-batch fermentation**")).

Por ejemplo, con el fin de evaluar la expresión y/o secreción de un polipéptido segregable, se llevan a cabo análisis en el nivel de proteína, el nivel de ARN, y/o mediante el uso de bioanálisis funcionales apropiados para la actividad y/o producción del polipéptido segregable. Análisis ejemplares empleados para analizar la expresión y/o secreción del polipéptido segregable incluyen, pero no se limitan a, Northern blotting, dot blotting (análisis de ADN o ARN), RT-PCR (reacción en cadena de polimerasa de transcriptasa inversa), o hibridación *in situ*, usando una sonda apropiadamente etiquetada (basada en la secuencia codificante del ácido nucleico), Southern blotting convencional y auto-radiografía.

En algunas incorporaciones, la producción, expresión y/o secreción de un polipéptido segregable se mide directamente en una muestra. En algunas incorporaciones, las mediciones se hacen usando análisis para detección de la actividad enzimática, expresión y/o producción. En algunas incorporaciones, la expresión de proteína se evalúa por métodos inmunológicos (v.g., tinción inmunohistoquímica de células y/o secciones de tejido, o inmunoanálisis de medio de cultivo de tejidos por métodos Western blotting o ELISA). Tales inmunoanálisis encuentran aplicación en evaluar cualitativa y/o cuantitativamente la expresión del polipéptido segregable. Estos métodos son conocidos por los expertos en el arte. Por supuesto, existen numerosos kits y reactivos disponibles en el comercio para uso en tales métodos.

En algunas incorporaciones, la presente invención provee asimismo extractos (v.g., sólidos o sobrenadantes) obtenidos del medio de cultivo empleado

para cultivar el hospedero de expresión. En algunas incorporaciones, el sobrenadante no contiene una cantidad sustancial del hospedero de expresión, mientras que en algunas incorporaciones alternativas, el sobrenadante no contiene ninguna cantidad del hospedero de expresión.

Cultivo de Células

Como se conoce en el arte, las células hospederas y las células transformadas de la presente invención pueden cultivarse en medios nutrientes convencionales. Sin embargo, en algunas incorporaciones, el medio de cultivo para células hospederas transformadas se modifica, según sea apropiado, para activar promotores y seleccionar transformantes. Las condiciones específicas de cultivo, tales como temperatura, pH y similares, son típicamente aquellas que se usan para la célula hospedera seleccionada para la expresión, y serán evidentes a los expertos en el arte. El medio y las condiciones de cultivo para células hospederas son conocidos por los expertos en el arte. Se anota que, en cultivo, transformantes estables de células hospederas fúngicas, tales como células de *Trichoderma*, son generalmente distinguibles de transformantes inestables por su tasa de crecimiento más rápida o la formación de colonias circulares con un perfil suavizado, más que desigual, en medio de cultivo sólido.

Composiciones

En algunas incorporaciones, la presente invención provee composiciones y métodos para expresar proteínas deseadas usando la secuencia señal NSP24 o CBH1, constructos y vectores. En algunas incorporaciones, la presente invención provee composiciones que contienen enzimas, incluyendo, pero no limitadas a, laccasas, glucoamilasas, α -amilasas, enzimas hidrolizantes de almidón granular, celulasas, lipasas, fosfolipasas, xilanasas, cutinasas, hemicelulasas, oxidasas, peroxidasas, proteasas, fitasas, queratinasas, pululanasas, glucoamilasas, pectinasas, oxidorreductasas, reductasas, perhidrolasas, fenol oxidasas, lipoxigenasas, ligninasas, tannanasas, pululanasas, pentosanasas, β -glucanasas, arabinosidasas, hialuronidasas, condroitinasas, mannanasas, esterases, acil transferasas, y combinaciones de las mismas.

Aplicaciones

Las proteínas deseadas producidas por la presente invención encuentran aplicación en cualesquiera aplicaciones apropiadas para esa proteína. Ejemplos de aplicaciones para proteínas tales como enzimas incluyen, pero no se limitan a comidas para animales para mejorar la ingesta y eficiencia de los alimentos (v.g., proteasas), hidrolisados de proteína dietética (v.g., para individuos con sistemas digestivos deteriorados), tratamiento de cuero, tratamiento de fibras proteínicas (v.g., lana y seda), limpieza, procesamiento de proteínas (v.g., para retirar péptidos amargos, mejorar el sabor de los alimentos, y/o para producir queso y/o cacao en polvo), productos de higiene personal (v.g., composiciones para el cabello), edulcorantes (v.g., producción de jarabes con alto contenido de maltosa o fructosa), fermentación y bioetanol (v.g., α -amilasas y glucoamilasas empleadas para tratar granos para fermentación para producir bioetanol). Ejemplos de aplicaciones para laccasas incluyen, pero no se limitan a, blanqueamiento de pulpa y papel, blanqueamiento de textiles, tratamiento de papel de desecho, destintado de papel de desecho, polimerización de compuestos aromáticos o proteínas, polimerización mediada por radicales y reacciones de entrecruzamiento (v.g., pinturas, recubrimientos, biomateriales), la activación de tinturas, y para acoplar compuestos orgánicos. De igual forma, las laccasas encuentran aplicación en composición de limpieza, incluyendo, pero sin limitación, detergentes para lavandería y otros detergentes.

EJEMPLOS

Los siguientes ejemplos se ofrecen para ilustrar, mas no para limitar la invención reivindicada. Será evidente para los expertos en el arte que pueden practicarse muchas modificaciones, tanto a los materiales como a los métodos, sin apartarse del ámbito de la invención.

En la divulgación experimental que sigue, se aplican las siguientes abreviaturas: M (molar); μ M (micromolar); N (normal); mol (moles); mmol (milimoles); μ mol (micromoles); nmol (nanomoles); g (gramos); mg (miligramos); kg (kilogramos); μ g y μ g (microgramos); L (litros); ml (mililitros); μ l y μ l (microlitros); cm (centímetros); mm (milímetros); μ m (micrómetros); nm (nanómetros); °C

(grados Centígrados); h y hr (horas); min (minutos); sec (segundos); msec (milisegundos); V (voltaje); xg (veces la gravedad); °F (grados Fahrenheit); amdS (acetamidasa, un marcador selectivo obtenido de *A. nidulans*); lccD (laccasa); BioRad (BioRad Laboratories, Hercules, CA); Difco (Difco Laboratories, Detroit, MI); Calbiochem (marca Calbiochem de propiedad de EMD Chemicals Inc., San Diego, CA); Sigma (Sigma Chemical Co., St. Louis, MO); Spectronic (Spectronic Devices, Ltd., Bedfordshire, Reino Unido); Advanced Kinetics (Advanced Kinetics and Technology Solutions, Suiza).

La mayoría de los vectores de expresión en los ejemplos se produjeron basándose en el esqueleto del plásmido pSL1180, cuya secuencia se provee en la base de datos GENBANK®, bajo el identificador U13865. Los marcadores tales como el marcador amdS, chaperonas o foldasas, laccasa (1ccD), las secuencias señal, fusiones TrGA y terminadores se agregaron usando el policonector y/o métodos PCR conocidos en el arte.

Los sitios en los plásmidos se identifican como sigue: cbh1 - celobiohidrolasa; Tcbh1 - el terminador de cbh1; TrGA - *Trichoderma* glucoamilasa; 1ccD - laccasa D; marcador amdS marcador seleccionable para autotrofismo; pSL1180 - el esqueleto del plásmido; laccase D opt - una versión optimizada del gen de la laccasa D que se construye con utilización de codones optimizados para la expresión en el hospedero (*Trichoderma*); Pcpc-1 - un promotor del gen de control-1 de la vía cruzada de *Neurospora crassa*; bla - el gen de la β -lactamasa (es decir, un marcador selectivo de *E. coli*); y HphR - el gen de la resistencia a la higromicina (un marcador selectivo de *E. coli*).

Para construir los plásmidos de expresión, se diseñaron cebadores y se emplearon en la reacción PCR Herculasa (Stratagene) que contenía la plantilla de ADN.

EJEMPLO 1. Construcción del Vector de Expresión pTrex4-laccaseD opt

Este Ejemplo describe los pasos implicados en la construcción del vector de expresión pTrex4-laccaseD opt. El plásmido se produjo para expresar el gen de la laccasa D optimizado por codón de *C. unicolor* usando el promotor CBH1 y la

secuencia señal CBH1. Este vector de expresión contenía el gen optimizado por codón de la laccasa D, fusionado al núcleo/conector de la CBH1 (celobiohidrolasa) expresado del promotor CBH1. Figura 1 ofrece una representación esquemática del plásmido de expresión *Trichoderma*. La secuencia del plásmido pTrex4-laccaseD opt se muestra como SEQ ID NO: 1. Los siguientes segmentos de ADN se ensamblaron en la construcción de pTrex4-laccase D opt (véase, Figura 1). Un fragmento del ADN genómico de *T. reesei*, que representa el promotor CBH1 y la secuencia señal CBH1 y el núcleo/conector CBH1 se insertó en el vector pSL1180 del plásmido. Se insertó una copia optimizada por codón del gen de la laccasa D (*laccase D opt*) de *C. unicolor*, de modo que quedara conectado operacionalmente a la CBH1 en su región conectora. Un terminador CBH1 de *T. reesei* se conectó operacionalmente al gen de la laccasa D. El gen *amdS* se agregó como un marcador autotrópico seleccionable. El gen *bla* (que codifica la beta-lactamasa, un marcador selectivo obtenido de *E. coli*) está presente en el vector pSL1180.

EJEMPLO 2. Construcción del Vector de Expresión pTrex2g-Bip1

El plásmido pTrex2g/Bip1 se produjo para expresar la chaperona *bip1* de *T. reesei*. Figura 2 ilustra la representación esquemática del plásmido de expresión pTrex2g-Bip1 de *Trichoderma*. La secuencia del plásmido se proporciona como SEQ ID NO: 2. Los siguientes segmentos de ADN se ensamblaron en la construcción de pTrex2g-Bip1. UN fragmento de 2267 bp de *T. reesei bip1* se insertó en el vector del plásmido pSL1180, conectado operacionalmente al promotor *Ppki* (piruvato cinasa de *T. reesei*). El terminador *cbh1* de *Trichoderma* se conectó operacionalmente al gen *bip1*. El marcador seleccionable HphR de *E. coli* se incluyó para selección y se conectó operacionalmente al promotor *Pcpc-1* (gen de control-1 de vía cruzada de *Neurospora crassa*) y el terminador *trpC* (gen C de la síntesis de triptófano de *A. nidulans*).

EJEMPLO 3. Construcción del Vector de Expresión pTrex2g-Pdi1

El plásmido pTrex2g-Pdi1 se produjo para expresar la chaperona *pdi1* en la misma forma que el pTrex2g-Bip1 (véase, Ejemplo 2), excepto que el gen de la chaperona *pdi1* de *T. reesei* (2465 bp) se insertó en lugar del gen de la chaperona *bip1*. Figura 3 ilustra la representación esquemática del plásmido pTrex2g-Pdi1 de

expresión de *Trichoderma*; la secuencia del plásmido se proporciona como SEQ ID NO: 3.

EJEMPLO 4. Construcción del Vector de Expresión pTrex2g-Ero1

El plásmido pTrex2g-Ero1 se produjo para expresar la chaperona *ero1* en la misma forma que el pTrex2g-Bip1 (véase, Ejemplo 2), excepto que el gen de la chaperona *ero1* de *T. reesei* (2465 bp) se insertó en lugar del gen de la chaperona *bip1*. Figura 4 ilustra la representación esquemática del *ero1* en el plásmido pTrex2g-Ero1 de expresión de *Trichoderma*. La secuencia de *ero1* se proporciona como SEQ ID NO: 4.

EJEMPLO 5. Construcción del Vector de Expresión pTrGA-laccaseD opt

El plásmido pTrGA-laccaseD opt se produjo análogamente al de Ejemplo 1, excepto que pTrGA-laccase D opt expresa una fusión de la glucoamilasa de longitud total de *T. reesei* y la laccasa D de *C. unicolor* con codones optimizados. Figura 5 ilustra la representación esquemática del plásmido pTrGA-laccaseD opt de expresión de *Trichoderma*; la secuencia polinucleotídica se muestra como SEQ ID NO: 5.

EJEMPLO 6. Construcción del Vector de Expresión pKB408

El plásmido pKB408 se produjo para expresar la laccasa D opt de *C. unicolor* fusionado operacionalmente al péptido señal NSP-24 de *T. reesei*. El plásmido se construyó análogamente al mostrado en Figura 1, excepto que los constructos de laccasa D se conectaron operacionalmente al péptido señal NSP-24, el cual se insertó en lugar de la laccasa D opt conectada a la secuencia señal CBH1, dominio catalítico y conector. Figura 6 ilustra la representación esquemática del plásmido pKB408 de expresión de *Trichoderma*; la secuencia polinucleotídica se muestra como SEQ ID NO: 6.

EJEMPLO 7. Construcción del Vector de Expresión pKB410.

El plásmido pKB410 se produjo como se describe en el Ejemplo 6, excepto

que se utilizó la secuencia señal CHB1 de *T. reesei* en lugar de la secuencia señal NSP-24. Figura 7 ilustra la representación esquemática del plásmido pKB410 de expresión de *Trichoderma*; la secuencia polinucleotídica se muestra como SEQ ID NO: 7.

EJEMPLO 8. Transformación de *T. reesei* y Análisis de Expresión

En este ejemplo, la cepa *T. reesei* recombinante estable, derivada de RL-P37 (véase, Sheir-Neiss and Montenecourt, Appl. Microbiol. Biotechnol., 20:46-53 (1984)) y delecionada para los genes *cbh1*, *cbh2*, *egl1*, y *egl2* descrita por Bower *et al* (véase, Bower *et al.*, *Carbohydrases From Trichoderma reesei and Other Micro-organisms*, Royal Society of Chemistry, Cambridge, páginas 327-334 (1998)) se usó para transformar los plásmidos de los Ejemplos 1-14 solos o en varias combinaciones. Se emplearon métodos biolísticos y de electroporación para transformar los plásmidos, como se describe a continuación.

Transformación Biolística

El plásmido de expresión se confirmó mediante secuenciamiento de ADN y se transformó biolísticamente en una cepa *Trichoderma*. La transformación de la cepa *Trichoderma* por el método de transformación biolística se llevó a cabo utilizando un Sistema de Bombardeo de Partículas Biolistic® PDS-1000/he (Bio-Rad) siguiendo las instrucciones del fabricante (véase, WO 05/001036 y la Solicitud de Patente de los Estados Unidos No. 2006/0003408). Los transformantes se seleccionaron y transfirieron en medio mínimo con placas de acetamida (MMA) y se cultivaron durante 4 días a 28-30°C. Un pequeño bloque de una colonia individual que incluía esporas y micelio se transfirió en 30 mls de caldo definido de lactosa NREL (pH 6.2), que contenía 1 mM de cobre. Los cultivos se desarrollaron durante 5 días a 28°C. Los caldos de cultivo se centrifugaron, y los sobrenadantes se analizaron utilizando el análisis ABTS, como se describe a continuación, para detectar la actividad laccasa.

Electroporación

Se llevó a cabo electroporación, como se describe en la Solicitud de

Patente de los Estados Unidos No: 60/931,072, incorporada en esta especificación por referencia en su totalidad. Una cepa *T. reesei* se cultivó y se puso a esporular en placas de agar de dextrosa de papa (Difco) durante alrededor de 10-20 días. Las esporas se lavaron de la superficie de las placas con agua y se purificaron por filtración a través de Miracloth (Calbiochem). Se recolectaron las esporas mediante centrifugación (3000 xg, 12 min), se lavaron una vez con agua helada y una vez con 1.1M sorbitol helado. El sedimento de las esporas se re-suspendió en un pequeño volumen de 1.1 M sorbitol frío, mezclado con alrededor de 8 µg de un fragmento de ADN purificado en gel, aislado de ADN de plásmido (pKB408 y pKB410, Figuras 6 y 7) por 100 µl de suspensión de esporas. La mezcla (100 µl) se puso en una cubeta de electroporación (gap 1mm) y se sometió a un pulso eléctrico usando los siguientes parámetros de electroporación: voltaje 6000-20000 V/cm, capacitancia = 25 µF, resistencia = 500. Después de la electroporación, las esporas se diluyeron alrededor de 100 veces en una mezcla 5:1 de 1.1 M sorbitol y YEPD (1% de extracto de levadura, 2% Bacto-peptona, 2% glucosa, pH 5.5), se pusieron en matraces de agitación y se incubaron durante 16-18 horas en un agitador orbital (28°C y 200 rpm). Las esporas se recogieron nuevamente una vez mediante centrifugación, se re-suspendieron en alrededor de 10 veces el volumen del sedimento 1.1 M sorbitol y se esparcieron sobre dos placas de Petri de 15 cm que contenían medio modificado con *amdS* (acetamida 0.6 g/l, cloruro de cesio 1.68 g/l, glucosa 20 g/l, fosfato potásico monobásico 15 g/l, heptahidrato de sulfato de magnesio 0.6 g/l, dihidrato de cloruro de calcio 0.6 g/l, sulfato de hierro (II) 5mg/l, sulfato de zinc 1.4 mg/l, cloruro de cobalto (II) 1mg/l, sulfato de manganeso (II) 1.6 mg/l, agar 20 g/l y pH 4.25). Los transformantes aparecieron alrededor de 1 semana de incubación a 28-30°C.

El análisis ABTS se llevó a cabo en la siguiente forma: Se preparó una solución patrón concentrada de ABTS que contenía 4.5 mM ABTS en agua (ABTS; Sigma Cat# A-1888). Se preparó amortiguador que contenía 0.1 M acetato de sodio pH 5.0. Luego, se agregaron 1.5 ml de amortiguador y 0.2 ml de solución patrón concentrada de ABTS a cubetas (10x4x45mm, No./REF67.742) y se mezclaron bien. Se preparó una cubeta extra como blanco. A continuación, se agregaron a las mezclas 50 µl de cada muestra de enzima a examinarse (usando varias diluciones).

Se midió la actividad ABTS en una máquina Genesys2 (Spectronic) usando una configuración de programa de análisis cinético ABTS: (Advanced Kinetics) en la siguiente forma: longitud de onda 420nm, tiempo de intervalo (sec) 2.0, tiempo total de ejecución (sec) 14.0, factor 1.000, límite inferior -000000.00, límite superior 999999.00, y el orden de reacción fue primero.

El procedimiento implicaba agregar 1.5 mL de NaOAc (120mM amortiguador NaOAc, pH 5.0), luego agregar 0.2 mL de 4.5mM ABTS a la cubeta, después el blanco a la cubeta, agregar 0.05 mL de la muestra de la enzima a la cubeta, mezclar rápidamente y bien y, finalmente, medir el cambio de absorción a 420nm, cada 2 segundos durante 14 segundos. Una unidad ABTS se define como un cambio de A420 por minuto (dada ninguna solución a la muestra). Cálculo de ABTS U/mL: (cambio en A420/min * factor de dilución).

EJEMPLO 9. Análisis de la Expresión del Gen de Fusión de Laccasa/Glucoamilasa en Transformantes de *T. reesei*

El medio de cultivo de los transformantes obtenidos y cultivados como se describe en el Ejemplo 8 se separó del micelio mediante centrifugación (16000 xg, 10 min), y se analizó la actividad ABTS de los sobrenadantes. Los resultados se muestran en Figura 10. La Tabla 3 presenta las cepas descritas en Figura 10. Figura 10 ilustra la mejora de la producción de laccasa por fusión del gen, que codifica laccasa *C. unicolor*, a la glucoamilasa de *Trichoderma* de longitud total. Los resultados mostraron que la expresión de laccasa mejoró 24-29% cuando se fusionó a la glucoamilasa de *Trichoderma*, que cuando se fusionó a CBH1.

Tabla 3. Cepas Utilizadas en Figura 10

Número de Identificación de la Cepa	Tipo de Cepa
#8-2	Fusión de laccasa CBH1
1066-9	Fusión de laccasa TrGA
1066-13	Fusión de laccasa TrGA
1066-15	Fusión de laccasa TrGA

EJEMPLO 10. Análisis de Producción de Laccasa Usando Secuencias Señal NSP24 y CBH1

Cuando la secuencia señal CBH1 de *T. reesei* se conectó operacionalmente al gen de la laccasa, se mejoró la expresión 4-5 veces sobre la cepa #8-2 de fusión CBH1 inicial sola en matraces de agitación, y 5-6 veces en un fermentador de 14 litros, como se muestra por los resultados presentados en las Figuras 11 (matraces de agitación) y 12 (fermentador). Cuando se empleó la secuencia señal NSP-24 de *T. reesei*, la expresión mejoró 3-4 veces en matraces de agitación y 4-5 veces en un fermentador de 14 litros. Tres clones se analizaron en los matraces de agitación con respecto a la secuencia señal CBH1 (#7, #10, y #13) y dos clones se analizaron con respecto a la secuencia señal NSP24 (#7 y #25), y la expresión se analizó a los 3 días (primera barra), a los 4 días (segunda barra) y a los 5 días (tercera barra). Un solo clon de cada una se analizó en los fermentadores de 14 litros, como se muestra por los resultados en Figura 12. En esta Figura, el diamante indica la secuencia señal NSP24 conectada operacionalmente a la laccasa D, el cuadrado indica la secuencia señal CBH1 conectada operacionalmente a la laccasa D, y el triángulo indica la fusión CBH1 sola.

EJEMPLO 11. Análisis de Producción de Laccasa Usando la Secuencia Señal CBH1 y Co-Expresión de *bip1* en un Fermentador

El plásmido de la secuencia señal CBH1 (conectada operacionalmente a laccasa) se co-transformó con el plásmido Bip1 de *T. reesei*, y se analizó la expresión. Los resultados se muestran en Figura 13. En Figura 13, los diamantes indican los datos obtenidos para la secuencia señal CHB1 (conectada operacionalmente a laccasa) más BIP1, mientras que los cuadrados indican los datos obtenidos para la secuencia señal CBH1 (conectada operacionalmente a laccasa) sola. Figura 13 ilustra la mejora de la producción de laccasa proporcionada por la secuencia señal CBH1 más la expresión de la chaperona BIP1, que aumentó la expresión significativamente, en más de 15% en fermentadores.

EJEMPLO 12. Análisis de Producción de Laccasa Usando la Secuencia Señal CBH1 y Co-Expresión de *bip1* en un Matraz de Agitación

El plásmido de la secuencia señal CBH1 (conectada operacionalmente a laccasa) se co-transformó con el plásmido *bip1* de *T. reesei*, se cultivó, y se analizó la expresión de la laccasa utilizando el análisis ABTS. Los resultados se presentan en Figura 14. Se analizaron 5 diferentes clones durante 3 días (primera barra), durante 4 días (segunda barra) y durante 5 días (tercera barra). KB410-13 fue un control que tenía el plásmido de secuencia señal CBH1 solo. Los otros 4 clones fueron KB410-13 con uno de los co-transformantes *bip1*: E32, E9, E16, y E10. Figura 14 ilustra la mejora de la producción de laccasa mediante co-expresión de chaperonas con *C. unicolor* en matraces de agitación. La co-expresión con *bip1* aumentó la expresión significativamente (de 14-41%) en matraces de agitación.

EJEMPLO 13. Análisis de Producción de Laccasa Usando Fusión de CBH1-laccasa D y Co-Expresión de una Variedad de Chaperonas

El plásmido de expresión que tenía una secuencia señal CBH1, dominio catalítico y conector conectados operacionalmente a laccasa se co-transformó con una variedad de plásmidos de chaperona de *T. reesei* (BIP1, PDI1, y ERO1). La célula transformada resultante se desarrolló en cultivo, y se analizó la expresión de laccasa. Figura 15 ilustra la mejora de la producción de laccasa por fusión del gen que codifica la laccasa de *C. unicolor* a la secuencia señal CBH1, dominio catalítico y conector, y la co-expresión con las chaperonas *bip1*, *pdi1* y *ero1*.

Todas las cepas tenían la secuencia señal CBH1, dominio catalítico y conector conectados a laccasa D. Las cepas 1B1, 1B12 y 1B19 tenían el casete de expresión *bip1*; había tres transformantes independientes, con diferencia en los números de copias del plásmido *bip1* y ubicación de integración. Las cepas 3B2 y 3B8 tenían el casete de expresión *pdi1*; ellas son dos transformantes independientes, con diferencia en los números de copias del plásmido *pdi1* y ubicación de integración. Las cepas 9B6 y 9B7 tenían el casete de expresión *ero1*; ellas son dos transformantes independientes, con diferencia en los números de copias del plásmido *ero1*, y la ubicación de integración puede ser diferente. #8-

2 es la cepa de control que no tiene casete de expresión de la chaperona.

Los resultados de Figura 15 indican que el aumento más alto en expresión se obtuvo con la co-expresión con la chaperona *Bip1*.

EJEMPLO 14. Análisis de Producción de Laccasa Usando la Secuencia Señal CBH1 y Co-Expresión de una Variedad de Chaperonas

El plásmido de la secuencia señal CBH1 (es decir, conectada operacionalmente a laccasa) se co-transformó con una variedad de plásmidos de la chaperona de *T. reesei* (*bip1*, *lhs1*, *pdi1*, *ppi1*, *ppi2*, *tig1*, *prp1*, y *ero1*), solos o en combinación. Los cultivos se desarrollaron en matraces de agitación, como se conoce en el arte, y se examinó la expresión de laccasa usando el análisis ABTS. Los clones se analizaron en triplicado. Los datos presentados en la Tabla 4 muestran que, agregando más de una chaperona, no se aumentó la expresión de laccasa por encima de la de *bip1* sola. Los datos en la Tabla 4 presentan tres muestras independientes purificadas de esporas (o clones) de la misma cepa.

Tabla 4: Expresión de Laccasa en la Presencia de Chaperonas					
Co-transformación de KB413-32A con Diferentes Chaperonas					
Cada cepa tiene 3 repeticiones: -A,-B,-C					
	Muestras	Chaperonas	4 días Caldo SF	6 días	
1	KB413-32A-A	bip1 únicamente	4.52	6.32	
2	KB413-32A-B	bip1 únicamente	4.26	6.35	
3	KB413-32A-C	bip1 únicamente	4.28	6.13	
4	KB414-1-A	bip1, ero1	3.88	5.89	
5	KB414-1-B	bip1, ero1	3.78	5.93	
6	KB414-1-C	bip1, ero1	3.76	5.59	
7	KB415-2-A	bip1, lhs1, blanca	3.8	5.93	
8	KB415-2-B	bip1, lhs1, blanca	3.72	5.92	
9	KB415-2-C	bip1, lhs1, blanca	3.78	6.06	
10	KB415-3-A	bip1, lhs1, gris	4.38	6.32	
11	KB415-3-B	bip1, lhs1, gris	4.3	6.66	
12	KB415-3-C	bip1, lhs1, gris	3.98	6.15	

13	KB416-3-A	bip1, pdi1	4.18	6.58	
14	KB416-3-B	bip1, pdi1	5.26	7.12	
15	KB416-3-C	bip1, pdi1	4.22	6.06	
16	KB417-3-A	bip1, ppi1	4.32	6.23	
17	KB417-3-B	bip1, ppi1	3.96	6.32	
18	KB417-3-C	bip1, ppi1	4.18	6.88	
19	KB418-2-A	bip1, ppi2	4.24	6.59	
20	KB418-2-B	bip1, ppi2	3.96	5.69	
21	KB418-2-C	bip1, ppi2	4.04	5.92	
22	KB419-1-A	bip1, tigA	4.66	5.98	
23	KB419-1-B	bip1, tigA	5.26	7.25	
24	KB419-1-C	bip1, tigA	4.18	6.05	
25	KB413-prp2-A	bip1, prpA	3.96	5.63	
26	KB413-prp2-B	bip1, prpA	3.9	5.59	
27	KB413-prp2-C	bip1, prpA	3.92	5.86	
28	KB414-1-A	bip1, ero1	4.2	6.01	
29	KB414-1-B	bip1, ero1	3.88	5.69	
30	KB414-1-C	bip1, ero1	3.92	5.88	

La invención y la manera y el proceso realizarla y usarla se han descrito ahora en términos completos, claros, concisos y exactos, con el fin de permitir su realización y utilización a cualquier persona experta en el arte al cual pertenece la invención. Hay que entender que lo anterior describe incorporaciones preferidas de la presente invención y que pueden introducirse modificaciones a ellas sin apartarse del ámbito de la presente invención, como se expone en las reivindicaciones. Para destacar en particular y reivindicar claramente el asunto considerado como invención, las siguientes reivindicaciones concluyen esta especificación.

LO QUE SE REIVINDICA ES:

1. Un método para producir una proteína deseada, que comprende los pasos de:
 - (a) introducir en una célula hospedera una primera secuencia de ácido nucleico, que comprende una secuencia señal conectada operacionalmente a una secuencia de la proteína deseada;
 - (b) expresar la primera secuencia de ácido nucleico;
 - (c) co-expresar una segunda secuencia de ácido nucleico que codifica una chaperona o foldasa seleccionada del grupo conformado por *bip1*, *ero1*, *pdi1*, *tig1*, *prp1*, *ppi1*, *ppi2*, *prp3*, *prp4*, calnexina, y *lhs1*; y
 - (d) recolectar la proteína deseada segregada de la célula hospedera.
2. El método de acuerdo con la Reivindicación 1, en donde la primera secuencia de ácido nucleico comprende además una secuencia enzimática entre la secuencia señal y la secuencia de la proteína deseada.
3. El método de acuerdo con la Reivindicación 2, en donde la secuencia enzimática se obtiene de una enzima glucoamilasa o CBH1.
4. El método de acuerdo con la Reivindicación 2, en donde la secuencia enzimática comprende una secuencia enzimática de longitud total.
5. El método de acuerdo con la Reivindicación 2, en donde la secuencia enzimática comprende una secuencia de dominio núcleo catalítico.
6. El método de acuerdo con la Reivindicación 5, en donde la primera secuencia de ácido nucleico comprende además una secuencia conectora entre la secuencia de dominio núcleo catalítico y la secuencia de la proteína deseada.
7. El método de acuerdo con la Reivindicación 1, en donde la proteína deseada es una laccasa.
8. El método de acuerdo con la Reivindicación 7, en donde la citada laccasa se deriva de un hongo filamentoso o levadura.

9. El método de acuerdo con la Reivindicación 8, en donde la citada laccasa se deriva de *Aspergillus*, *Neurospora*, *Podospora*, *Botrytis*, *Collybia*, *Cerrena*, *Stachybotrys*, *Panus*, *Thielavia*, *Fomes*, *Lentinus*, *Pleurotus*, *Trametes*, *Rhizoctonia*, *Coprinus*, *Psatyrella*, *Myceliophthora*, *Schytalidium*, *Phlebia*, *Coriolus*, *Spongipellis*, *Polyporus*, *Ceriporiopsis subvermispora*, *Ganoderma tsunodae*, o *Trichoderma*.
10. El método de acuerdo con la Reivindicación 9, en donde la citada laccasa se deriva de la laccasa A1, A2, B1, B2, B3, C, D1, D2, o E de *Cerrena*.
11. El método de acuerdo con la Reivindicación 9, en donde la citada laccasa se deriva de la proteína madura de la laccasa D de *Cerrena*.
12. El método de acuerdo con la Reivindicación 1, en donde la secuencia señal codifica un péptido señal celobiohidrolasa I o un péptido señal NSP24.
13. El método de acuerdo con la Reivindicación 1, en donde el hospedero es un hongo filamentoso.
14. El método de acuerdo con la Reivindicación 13, en donde el hospedero es ascomycetes.
15. El método de acuerdo con la Reivindicación 14, en donde el hospedero es *Trichoderma*.
16. El método de acuerdo con la Reivindicación 1, en donde la primera secuencia de ácido nucleico comprende además un promotor corriente arriba de una secuencia señal.
17. El método de acuerdo con la Reivindicación 16, en donde el promotor es nativo a célula hospedera y no está asociado naturalmente con la secuencia de la proteína deseada.
18. El método de acuerdo con la Reivindicación 1, en donde la chaperona es BIP1.
19. El método de acuerdo con la Reivindicación 1, en donde la segunda secuencia de ácido nucleico se conecta operacionalmente a un promotor.

20. El método de acuerdo con la Reivindicación 19, en donde el promotor es nativo a la célula hospedera y no está asociado naturalmente con la segunda secuencia de ácido nucleico.

21. El método de acuerdo con la Reivindicación 2, en donde la proteína deseada es una laccasa y la laccasa se produce como una proteína de fusión con la enzima.

1 / 16

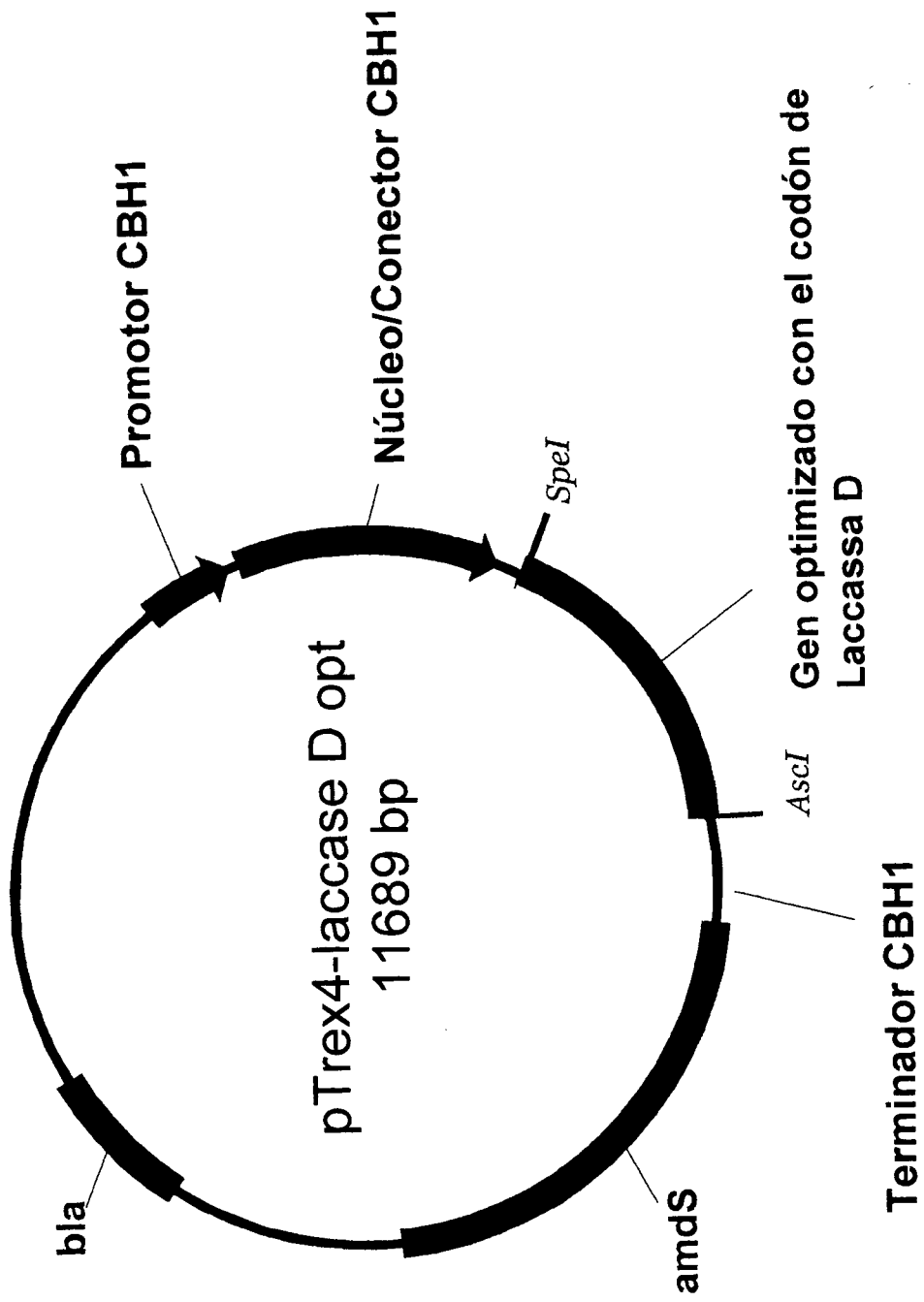


FIG. 1

2 / 16

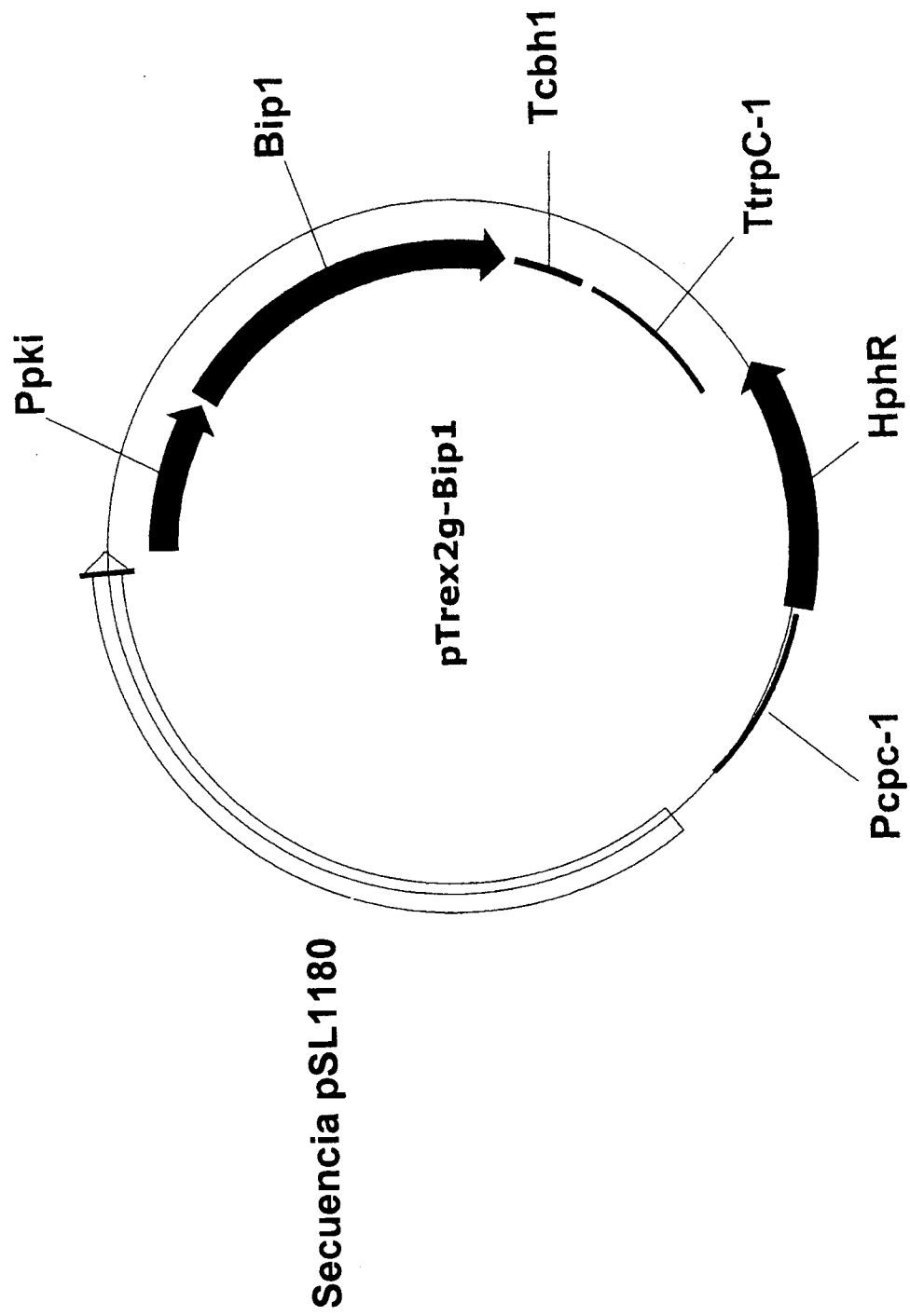
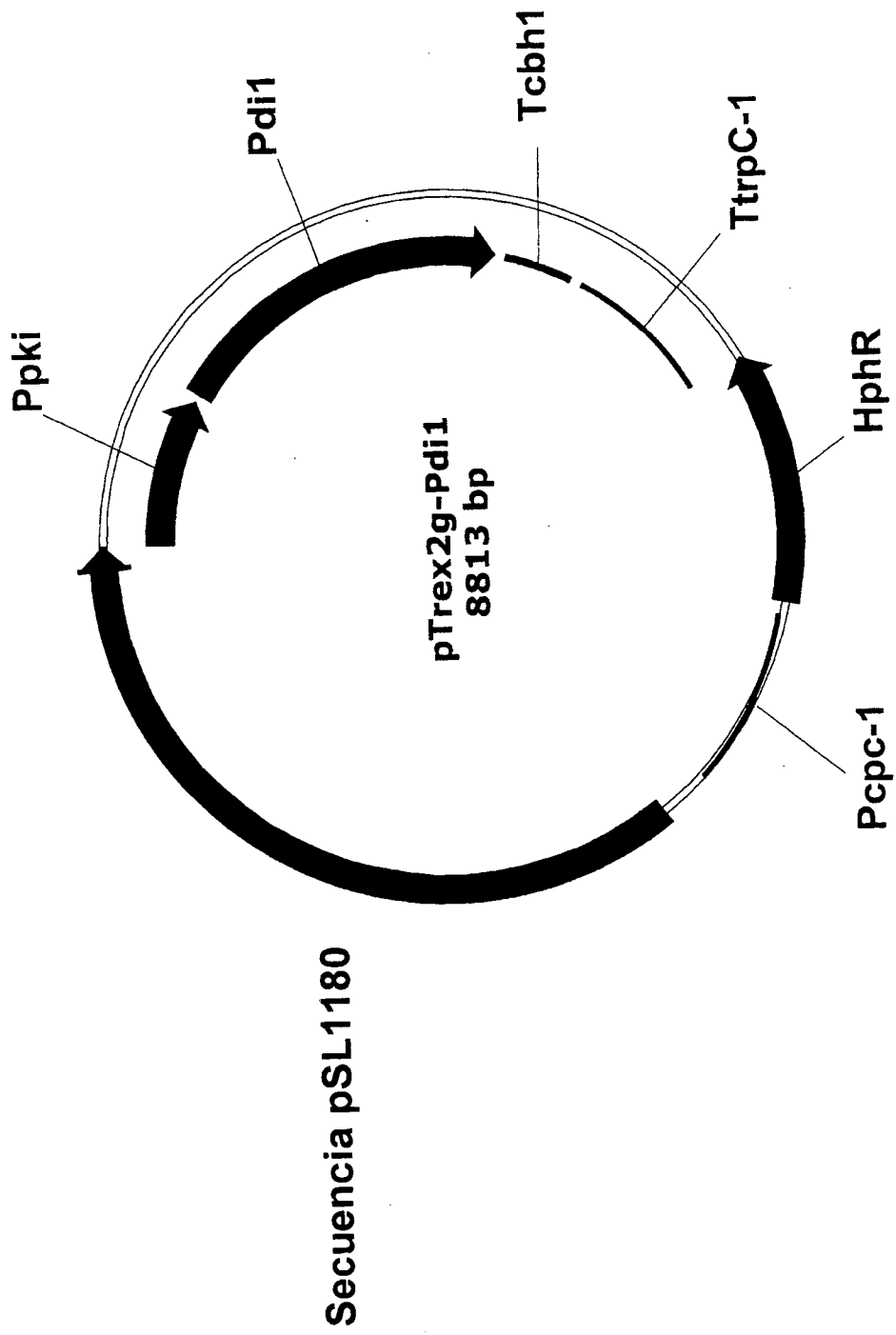


FIG. 2

3 / 16



4 / 16

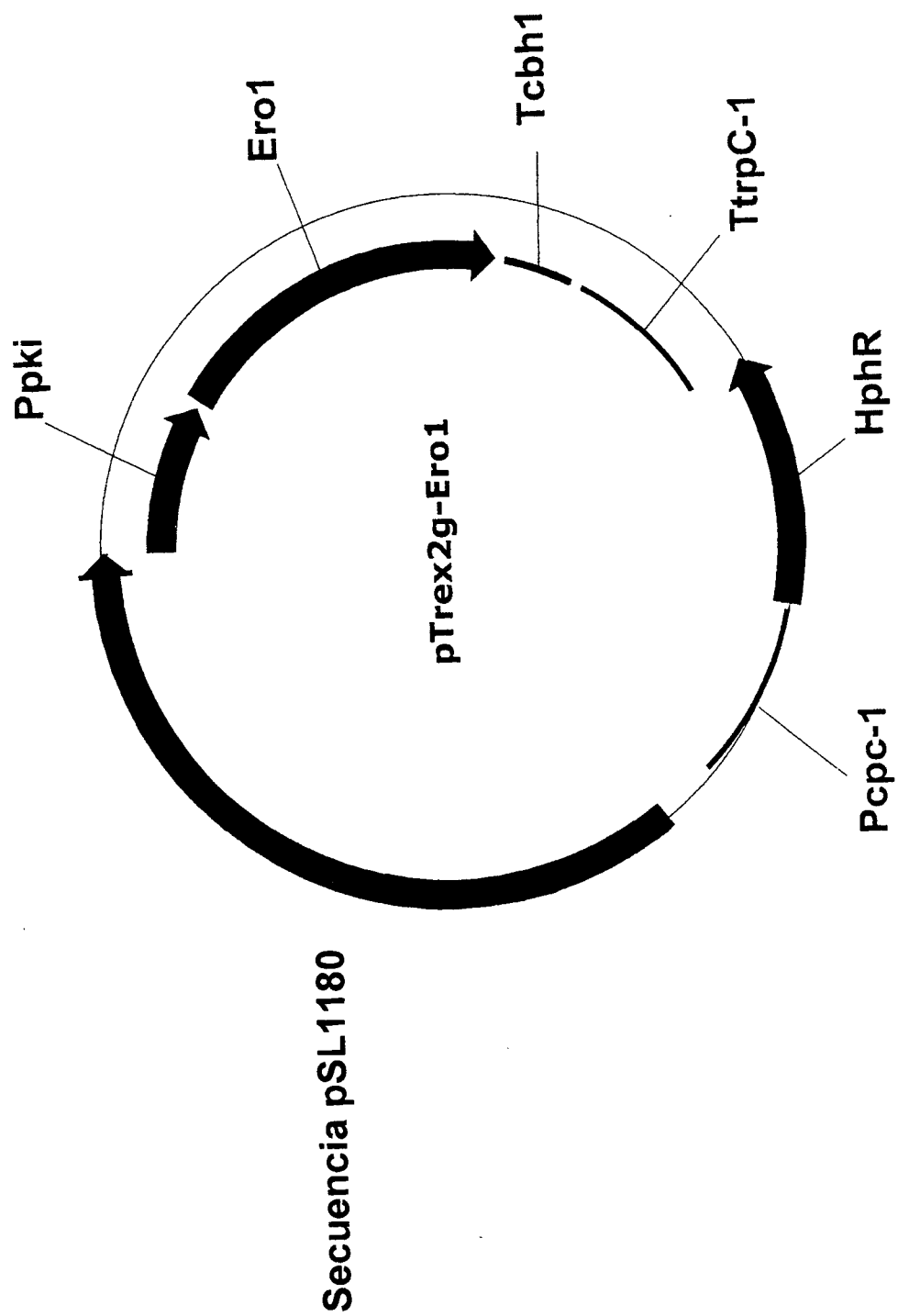


FIG. 4

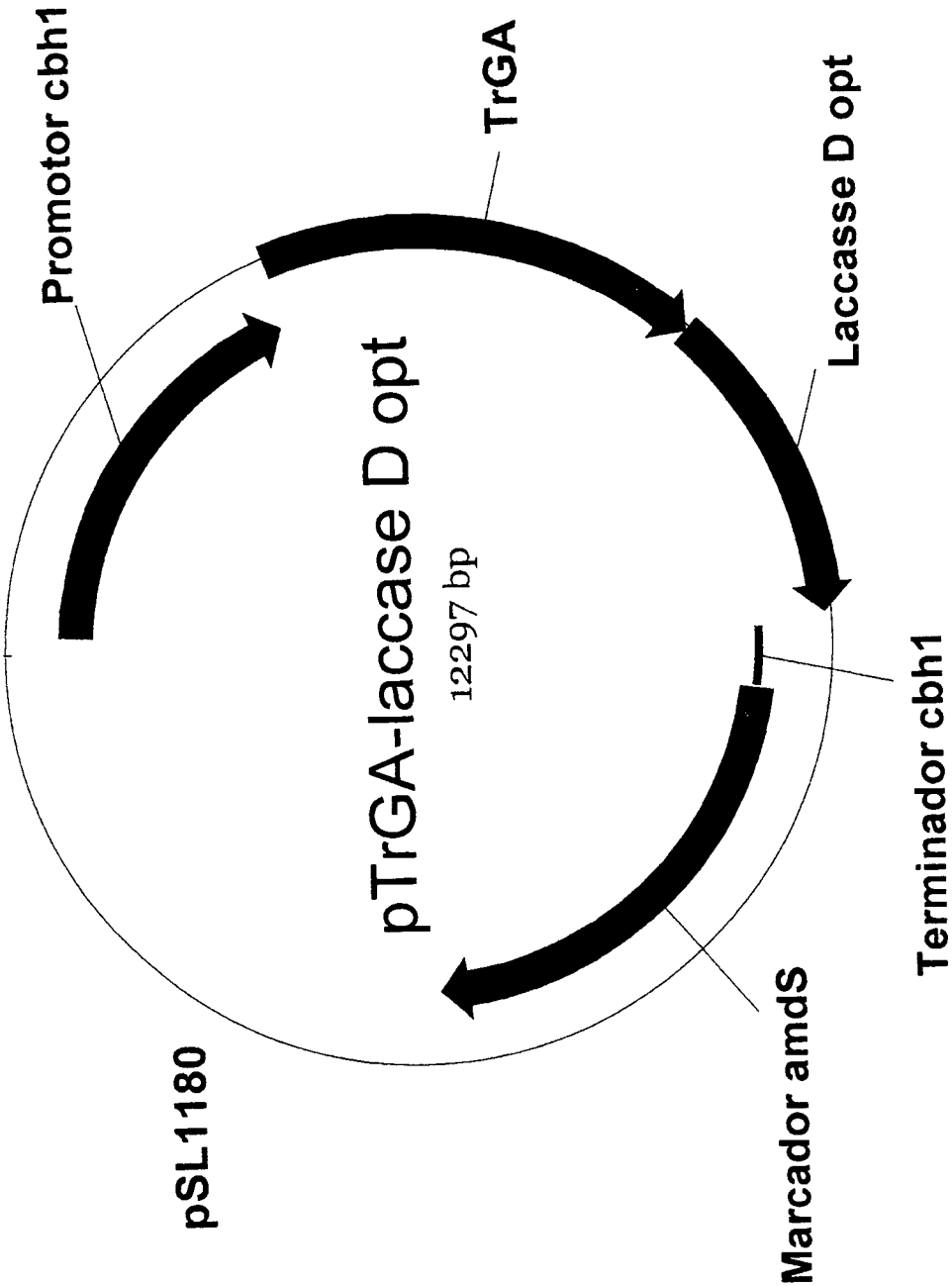


FIG. 5

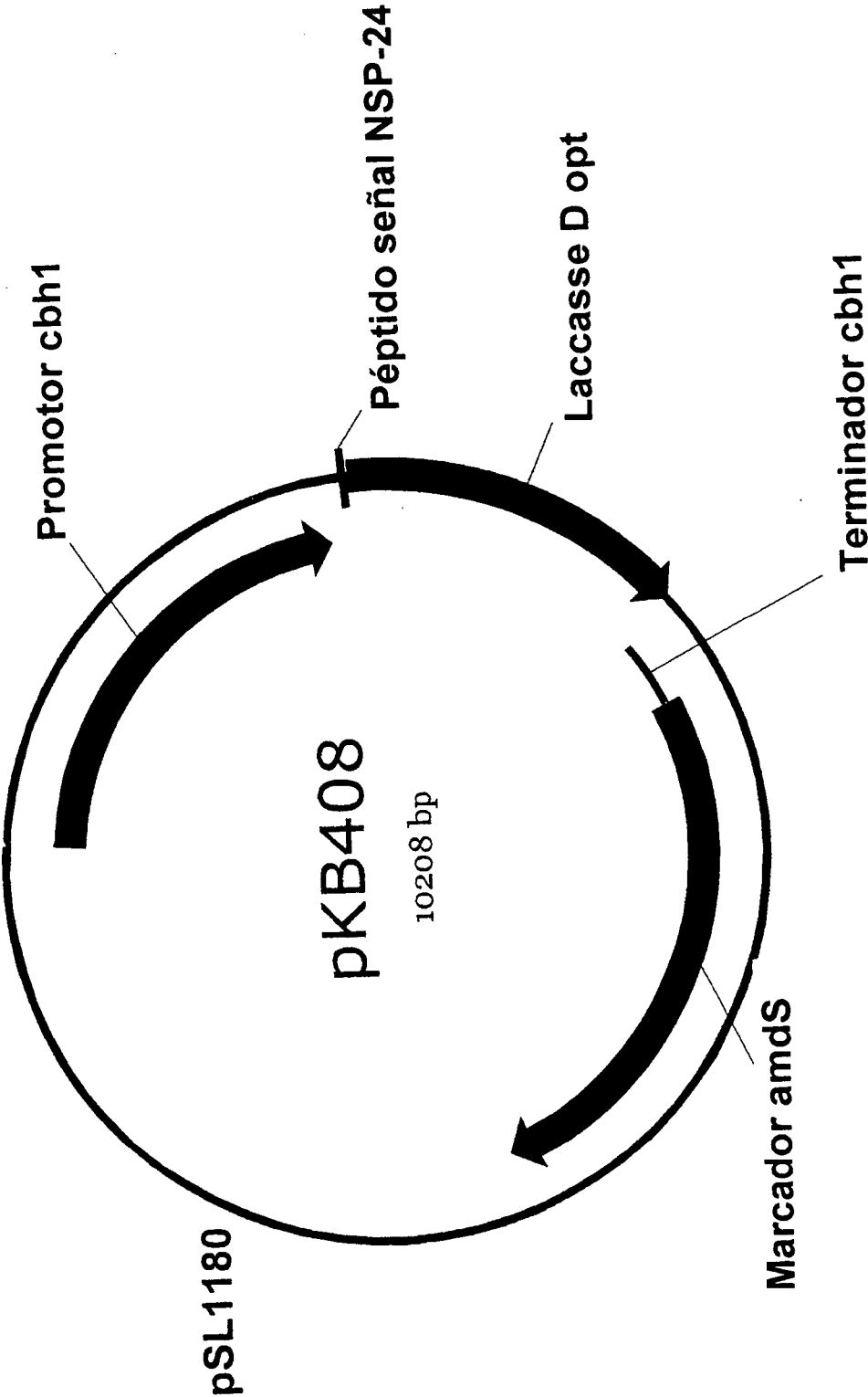


FIG. 6

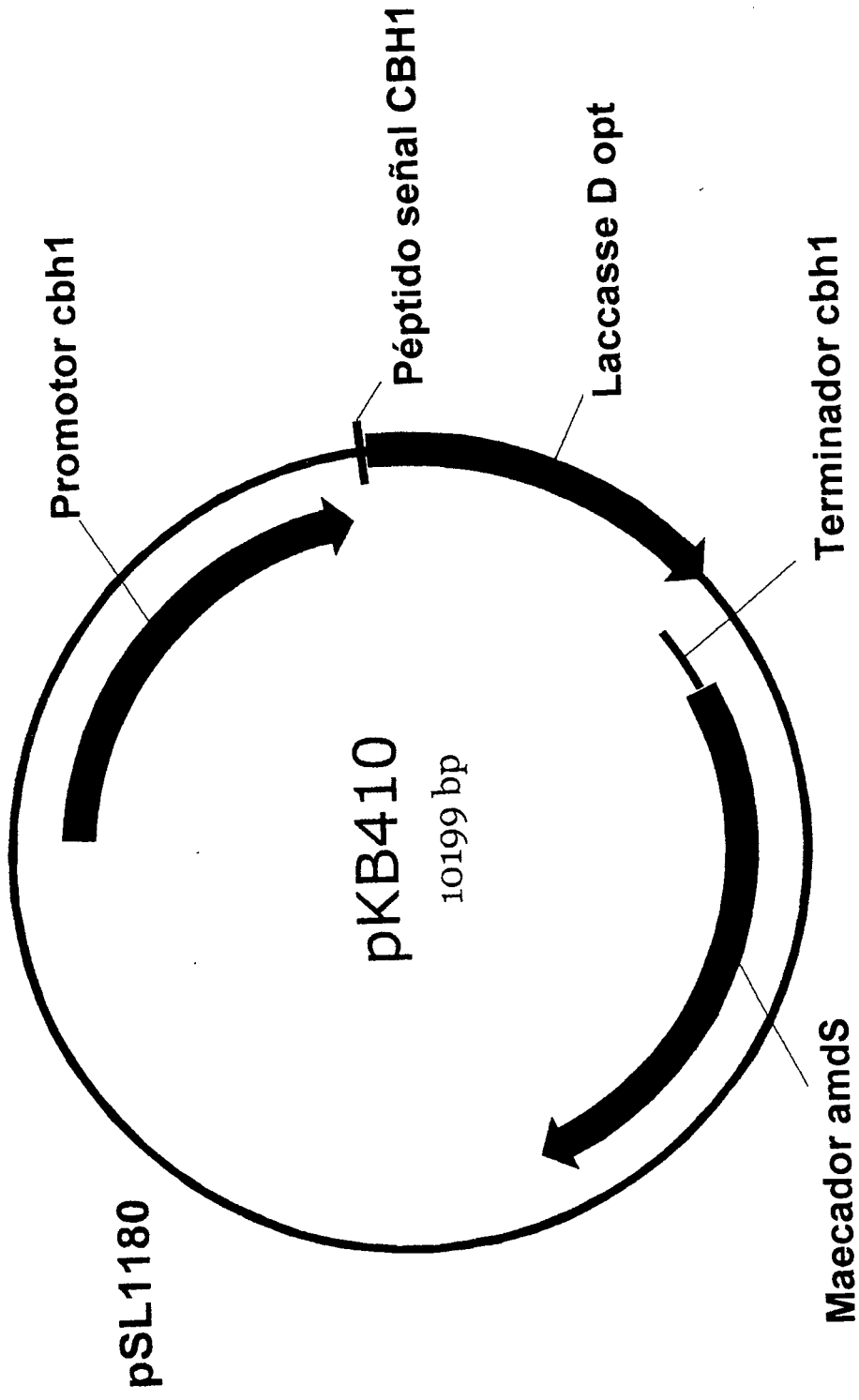


FIG. 7

8 / 16

Secuencia de ácido nucleico NSP24 (SEQ ID NO: 8)

CTGCAGCCACTTGCAGTCCCGTGAATTCCTCACGGTGAATGTAGGCCTTTTGTAGGG
TAGGAATTGTCACTCAAGCACCCCAACCTCCATTACGCCTCCCCCATAGAGTTCCC
AATCAGTGAGTCATGGCACTGTTCTCAAATAGATTGGGGAGAAGTTGACTTCCGCCC
AGAGCTGAAGGTCGCACAACCGCATGATATAGGGTCGGCAACGGCAAAAAGCACGT
GGCTCACCGAAAAGCAAGATGTTTGCGATCTAACATCCAGGAACCTGGATACATCCA
TCATCACGCACGACCACTTTGATCTGCTGGTAAACTCGTATTCGCCCTAAACCGAAG
TGCGTGGTAAATCTACACGTGGGCCCCCTTTCGGTATACTGCGTGTGTCTTCTCTAGG
TGCCATTCTTTTCCCTTCCCTCTAGTGTTGAATTGTTTGTGTTGGAGTCCGAGCTGTA
ACTACCTCTGAATCTCTGGAGAATGGTGGACTAACGACTACCGTGCACCTGCATCAT
GTATATAATAGTGATCCTGAGAAGGGGGGTTTGGAGCAATGTGGGACTTTGATGGTC
ATCAAACAAAGAACGAAGACGCCTCTTTTGCAAAGTTTGTTCGGCTACGGTGAAG
AACTGGATACTTGTGTGTCTTCTGTGTATTTTGTGGCAACAAGAGGCCAGAGACA
ATCTATTCAAACACCAAGCTTGCTCTTTTGAGCTACAAGAACCTGTGGGGTATATAT
CTAGAGTTGTGAAGTCGGTAATCCCGCTGTATAGTAATACGAGTCGCATCTAAATAC
TCCGAAGCTGCTGCGAACCCGGAGAATCGAGATGTGCTGGAAAGCTTCTAGCGAGCG
GCTAAATTAGCATGAAAGGCTATGAGAAATTCTGGAGACGGCTTGTGTAATCATGGC
GTTCCATTCTTCGACAAGCAAAGCGTTCCGTCGCAGTAGCAGGCACTCATTCCCGAA
AAAACTCGGAGATTCCTAAGTAGCGATGGAACCGGAATAATATAATAGGCAATACAT
TGAGTTGCCCTCGACGGTTGCAATGCAGGGGTACTGAGCTTGGACATAACTGTTCCGT
ACCCACCTCTTCTCAACCTTTGGCGTTTCCCTGATTTCAGCGTACCCGTACAAGTCG
TAATCACTATTAACCCAGACTGACCGGACGTGTTTGGCCCTTCATTTGGAGAAATAA
TGTCATTGCGATGTGTAATTTGCCCTGCTTGACCGACTGGGGCTGTTTGAAGCCCGAA
TGTAAGGATTGTTATCCGAACCTCTGCTCGTAGAGGCATGTTGTGAATCTGTGTTCGGC
AGGACACGCCTCGAAGGTTACGGCAAGGGAAACCACCGATAGCAGTGTCTAGTAGC
AACCTGTAAAGCCGCAATGCAGCATCACTGGAAAATACAAACCAATGGCTAAAAGTA
CATAAGTTAATGCCTAAAGAAGTCATATACCAGCGGCTAATAATTGTACAATCAAGT
GGCTAAACGTACCGTAATTTGCCAACGGCTTGTGGGGTTGCAGAAGCAACGGCAAAG
CCCCACTTCCCCACGTTTGTCTTCACTCAGTCCAATCTCAGCTGGTGATCCCCCA
ATTGGGTCGCTTGTGTTTCCGGTGAAGTGAAAGAAGACAGAGGTAAGAATGTCTGA
CTCGGAGCGTTTTGCATACAACCAAGGGCAGTGATGGAAGACAGTGAAATGTTGACA
TTCAAGGAGTATTTAGCCAGGGATGCTTGAGTGTATCGTGTAAGGAGGTTTGTCTGC
CGATACGACGAATACTGTATAGTCACTTCTGATGAAGTGGTCCATATTGAAATGTAA
GTCGGCACTGAACAGGCAAAAGATTGAGTTGAAACTGCCTAAGATCTCGGGCCCTCG
GGCCTTCGGCCTTTGGGTGTACATGTTTGTGCTCCGGGCAAATGCAAAGTGTGGTAG
GATCGAACACACTGCTGCCTTTACCAAGCAGCTGAGGGTATGTGATAGGCAAATGTT
CAGGGGCCACTGCATGGTTTTCGAATAGAAAGAGAAGCTTAGCCAAGAACAATAGCCG
ATAAAGATAGCCTCATTAACGGAATGAGCTAGTAGGCAAAGTCAGCGAATGTGTAT
ATATAAAGGTTTCGAGGTCCGTGCCTCCCTCATGCTCTCCCCATCTACTCATCAACTC
AGATCCTCCAGGAGACTTGTACACCATCTTTTGAGGCACAGAAACCCAATAGTCAAC
CATACAAGTTTGTACAAAAAGCAGGCTCCGCGGCCGCCCCCTTCACCATGCAGAC
CTTTGGAGCTTTTCTCGTTTCCCTTCCTCGCCGCCAGCGGCCTGGCCGCGGCCCTCCC
CACCGAGGGTCAGAAGACGGCTTCCGTCGAGGTCCAGTACAACAAGAACTACGTCCC
CCACGGCCCTACTGCTCTCTTCAAGGCCAAGAGAAAGTATGGCGCTCCCATCAGCGA

FIG. 8-1

9 / 16

CAACCTGAAGTCTCTCGTGGCTGCCAGGCAGGCCAAGCAGGCTCTCGCCAAGCGCCA
GACCGGCTCGGCGCCCAACCACCCAGTGACAGCGCCGATTTCGGAGTACATCACCTC
CGTCTCCATCGGCACTCCGGCTCAGGTCTCCCCCTGGACTTTGACACCGGCTCCTC
CGACCTGTGGGTCTTTAGCTCCGAGACGCCCAAGTCTTCGGCCACCGGCCACGCCAT
CTACACGCCCTCCAAGTCGTCCACCTCCAAGAAGGTGTCTGGCGCCAGCTGGTCCAT
CAGCTACGGCGACGGCAGCAGCTCCAGCGGCGATGTCTACACCGACAAGGTCACCAT
CGGAGGCTTCAGCGTCAACACCCAGGGCGTCGAGTCTGCCACCCGCGTGTCCACCGA
GTTTCGTCCAGGACACGGTCATCTCTGGCCTCGTCGGCCTTGCCCTTTGACAGCGGCAA
CCAGGTCAGGCCGCACCCGCAGAAGACGTGGTTCTCCAACGCCGCCAGCAGCCTGGC
TGAGCCCCCTTTTCACTGCCGACCTGAGGCACGGACAGAGTAAGTAGACACTCACTGG
AATTCGTTCTTTTCCCGATCATCATGAAAGCAAGTAGACTGACTGAACCAAACAAC
AGACGGCAGCTACAACCTTTGGCTACATCGACACCAGCGTCGCCAAGGGCCCCGTTGC
CTACACCCCCGTTGACAACAGCCAGGGCTTCTGGGAGTTCACTGCCTCGGGCTACTC
TGTTCGGCGGGCGGCAAGCTCAACCGCAACTCCATCGACGGCATTGCCGACACCGGCAC
CACCTGCTCCTCCTCGACGACAACGTCTCGATGCCTACTACGCCAACGTCCAGTC
GGCCCAGTACGACAACCAGCAGGAGGTGTCTCTTCGACTGCGACGAGGACCTCCC
TTCGTTTCAGCTTCGGTGTGGAAGCTCCACCATCACCATCCCTGGCGATCTGCTGAA
CCTGACTCCCCCTCGAGGAGGGCAGCTCCACCTGCTTCGGTGGCCTCCAGAGCAGCTC
CGGCATTGGCATCAACATCTTTGGTGACGTTCGCCCTCAAGGCTGCCCTGGTTGTCTT
TGACCTCGGCAACGAGCGCCTGGGCTGGGCTCAGAAATAAAAGGGTGGGCGCGCCGA
CCCAGCTTTCTTGTAACAAGTGGTGATCGCGCCAGCTCCGTGCGAAAGCCTGACGCA
CCGGTAGATTCTTGTTGAGCCCGTATCATGACGGCGGGGAGCTACATGGCCCCGG
GTGATTTATTTTTTTTGTATCTACTTCTGACCCCTTTTCAAATATACGGTCAACTCAT
CTTTCACTGGAGATGCGGCCTGCTTGGTATTGCGATGTTGTCAGCTTGGCAAATTGT
GGCTTTCGAAAACACAAAACGATTTCCTTAGTAGCCATGCATTTTAAGATAACGGAAT
AGAAGAAAGAGGAAATTAAAAAACAACATCCCGTTTCATAACCCGTA
GAATCGCCGCTCTTCGTGTATCCAGTACCAGTTTATTTTGAATAGCTCGCCCGCTG
GAGAGCATCCTGAATGCAAGTAACAACCGTAGAGGCTGACACGGCAGGTGTTGCTAG
GGAGCGTCGTGTTCTACAAGGCCAGACGTCTTCGCGGTTGATATATATGTATGTTTG
ACTGCAGGCTGCTCAGCGACGACAGTCAAGTTCGCCCTCGCTGCTTGTGCAATAATC
GCAGTGGGGAAGCCACACCGTGACTCCCATCTTTTCAGTAAAGCTCTGTTGGTGTTTA
TCAGCAATACACGTAATTTAACTCGTTAGCATGGGGCTGATAGCTTAATTACCGTT
TACCAGTGCCATGGTTCTGTCAGCTTTCTTGGCCCGTAAATTCGGCGAAGCCAGCC
AATCACCAGCTAGGCACCAGCTAAACCCATATAATTAGTCTCTTATCAACACCATCCG
CTCCCCCGGGATCAATGAGGAGAATGAGGGGGATGCGGGGCTAAAGAAGCCTACATA
ACCCTCATGCCAACTCCCAGTTTACACTCGTCGAGCCAACATCCTGACTATAAGCTA
ACACAGAATGCCTCAATCCTGGGAAGAACTGGCCGCTGATAAGCGCGCCCGCCTCGC
AAAAACCATCCCTGATGAATGGAAAGTCCAGACGCTGCCTGCGGAAGACAGCGTTAT
TGATTTCCCAAAGAAATCGGGGATCCTTTTCAGAGGCCGAAGTGAAGATCACAGAGGC
CTCCGCTGCAGATCTTGTGTCCAAGCTGGCGGCCGGAGAGTTGACCTCGGTGGAAGT
TACGCTAGCATTTCTGTAAACGGGCAGCAATCGCCCAGCAGTTAGTAGGGTCCCCCTCT
ACCTCTCAGGGAGATGTAACAACGCCACCTTATGGGACTATCAAGCTGACGCTGGCT
TCTGTGCAGACAACTGCGCCACGAGTTCTTCCCTGACGCCGCTCTCGCGCAGGCA
AGGGAACCTCGATGAATACTACGCAAAGCACAAAGAGACCCGTTGGTCCACTCCATGGC
CTCCCCATCTCTCTCAAAGACCAGCTTCGAGTCAAGGTACACCGTTGCCCTTAAGTC
GTTAGATGTCCCTTTTGTGTCAGCTAACATATGCCACCAGGGCTACGAAACATCAATG
GGCTACATCTCATGGCTAAACAAGTACGACGAAGGGGACTCGGTTCTGACAACCATG
CTCCGCAAAGCCGGTGCCGTCTTCTACGTCAAGACCTCTGTCCCGCAGACCCTGATG

FIG. 8-2

10 / 16

GTCTGCGAGACAGTCAACAACATCATCGGGCGCACCGTCAACCCACGCAACAAGAAC
TGGTCGTGCGGCGGCAGTTCTGGTGGTGAGGGTGCGATCGTTGGGATTCGTGGTGGC
GTCATCGGTGTAGGAACGGATATCGGTGGCTCGATTTCGAGTGCCGGCCGCGTTCAAC
TTCCTGTACGGTCTAAGGCCGAGTCATGGGCGGCTGCCGTATGCAAAGATGGCGAAC
AGCATGGAGGGTCAGGAGACGGTGCACAGCGTTGTCTGGGCCGATTACGCACCTCTGTT
GAGGGTGAGTCCTTCGCCTCTTCTTCTTTTCTTCTGCTCTATACCAGGCCTCCACTGT
CCTCCTTTCTTGCTTTTTTATACTATATACGAGACCGGCAGTCACTGATGAAGTATGT
TAGACCTCCGCCTCTTCACCAAATCCGTCTCGGTTCAGGAGCCATGGAAATACGACT
CCAAGGTCATCCCCATGCCCTGGCGCCAGTCCGAGTCGGACATTATTGCCTCCAAGA
TCAAGAACGGCGGGCTCAATATCGGCTACTACAACCTTCGACGGCAATGTCCTTCCAC
ACCCTCCTATCCTGCGCGGCGTGGAAACCACCGTCGCCGCACTCGCCAAAGCCGGTC
ACACCGTGACCCCGTGGACGCCATAACAAGCACGATTTTCGCCACGATCTCATCTCCC
ATATCTACGCGGCTGACGGCAGCGCCGACGTAATGCGCGATATCAGTGCATCCGGCG
AGCCGGCGATTCCAAATATCAAAGACCTACTGAACCCGAACATCAAAGCTGTTAACA
TGAACGAGCTCTGGGACACGCATCTCCAGAAGTGGAATTACCAGATGGAGTACCTTG
AGAAATGGCGGGAGGCTGAAGAAAAGGCCGGGAAGGAAGTGGACGCCATCATCGCGC
CGATTACGCCTACCGCTGCGGTACGGCATGACCAGTTCGGGTACTATGGGTATGCCT
CTGTGATCAACCTGCTGGATTTCACGAGCGTGGTTGTTCCGGTTACCTTTGCGGATA
AGAACATCGATAAGAAGAATGAGAGTTTCAAGGCGGTTAGTGAGCTTGATGCCCTCG
TGCAGGAAGAGTATGATCCGGAGGCGTACCATGGGGCACCGGTTGCAGTGCAGGTTA
TCGGACGGAGACTCAGTGAAGAGAGGACGTTGGCGATTGCAGAGGAAGTGGGGAAGT
TGCTGGGAAATGTGGTGACTCCATAGCTAATAAGTGTGAGATAGCAATTTGCACAAG
AAATCAATACCAGCAACTGTAAATAAGCGCTGAAGTGACCATGCCATGCTACGAAAG
AGCAGAAAAAACCTGCCGTAGAACCAGAGATATGACACGCTTCCATCTCTCAAA
GGAAGAATCCCTTCAGGGTTGCGTTTCCAGTCTAGACACGTATAACGGCACAAGTGT
CTCTCACCAAATGGGTTATATCTCAAATGTGATCTAAGGATGGAAAGCCAGAATAT
CGATCGCGCGCAGATCCATATATAGGGCCCCGGGTTATAATTACCTCAGGTCGACGTC
CCATGGCCATTTCGAATTTCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTGTTA
TCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGG
TGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCCGCTTTCCA
GTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGG
CGGTTTGCCTATTGGGCGCTCTTCCGCTTCTCTCGCTCACTGACTCGCTGCGCTCGGT
CGTTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCAC
AGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAG
GAACCGTAATAAGGCCGCTTGCTGGCGTTTTTTCATAGGCTCCGCCCCCTGACGA
GCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAG
ATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCC
GCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGCGCTTTCTCATAG
CTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGT
GCACGAACCCCCGTTTCAGCCCGACCGCTGCGCCTTATCCGGTAACCTATCGTCTTGA
GTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGAT
TAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA
CGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT
CGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGG
TTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCC
TTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAACTCACGTTAAGGGAT
TTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATG
AAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATG

FIG. 8-3

11 / 16

CTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGC
CTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAG
TGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAA
CCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCAT
CCAGTCTATTAATTGTTGCCGGAAGCTAGAGTAAGTAGTTCCGCCAGTTAATAGTTT
GCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTAT
GGCTTCATTTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTT
GTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGC
CGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCC
ATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATA
GTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAAATACCGCGCC
ACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAAC
CTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGTAACCCACTCGTGCACCCAA
CTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG
GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACT
CTTCCTTTTTCATATTTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATA
CATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCGCGCACATTTCCCCG
AAAAGTGCCACCTGACGTCTAAGAAACCATTTATTCATGACATTAACCTATAAAAA
TAGGCGTATCACGAGGGCCCTTTCTGTCTCGCGCGTTTCGGTGATGACGGTGAAAACCT
CTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAG
CAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAA
CTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATAAAATTGTAAACGTTAA
TATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCATTTTTTAACCAATA
GGCCGAAATCGGCAAAATCCCTTATAAATCAAAGAATAGCCCGAGATAGGGTTGAG
TGTTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAA
AGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCAAATC
AAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCC
CCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA
AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAAGCGGTCACGCTGCGCGTAAC
CACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTACTATGGTTGCTTTGA
CGTATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGC
CATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCG
CTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAAGTTGGGTAACG
CCAGGGTTTTCCAGTCACGACGTTGTAAAACGACGGCCAGTGCCCAAGCTTACTAG
TACTTCTCGAGCTCTGTACATGTCCGGTCGCGACGTACGCGTATCGATGGCGCCAGC
TGCAGGCGGCGCGC

FIG. 8-4

13 / 16

Ser Glu Ala Leu Thr Pro His Pro Cys Thr Thr Val Gly Gln Glu Ile Cys Gly Asp Gly Cys Gly Thr Tyr Ser Asp Asn Arg
 1170 TATGGCGGCACTTGGGATCCCGATGGCTGGAGTGGAAACCCATACCGCTGGCAACACACAGCTTCTACGGCCCTGGCTCAAGCTTTTACC
 Tyr Gly Thr Cys Asp Pro Asp Gly Cys Asp Trp Asn Pro Tyr Arg Leu Gly Asn Thr Ser Phe Tyr Gly Pro Gly Ser Ser Phe Thr
 1260 CTCGATACCAACCAAGAAATTGACCGTTGTACCCAGTTCGAGACGTCGGGTGCCATCAACCGATATGTCCAGAAATGGCGTCACTTTC
 Leu Asp Thr Thr Lys Lys Leu Thr Val Val Thr Gln Phe Glu Thr Ser Gly Ala Ile Asn Arg Tyr Tyr Val Gln Asn Gly Val Thr Phe
 1350 CAGCAGCCCAACGCCGAGCTTGGTAGTTACTCTGGCAACGAGCTCAACGATGATTACTGCACAGCTGAGGAGGCAGAAATTCGGCGGATCC
 Gln Gln Pro Asn Ala Glu Leu Gly Ser Tyr Ser Gly Asn Glu Leu Asn Asp Tyr Cys Thr Ala Glu Glu Ala Glu Phe Gly Gly Ser
 1440 TCTTTCTCAGACAAAGGCGGCGCTGACTCAGTTCAGAAAGGCTACCTCTGGCGGCATGGTTCTGGTCATGAGTCTGTGGGATGATGTGAGT
 Ser Phe Ser Asp Lys Gly Gly Leu Thr Gln Phe Lys Ala Thr Ser Gly Gly Met Val Leu Val Met Ser Leu Trp Asp Asp
 1530 TTGATGGACAAACATGCGCGTTGACAAAGAGTCAAGCAGCTGACTGAGATGTTACAGTACTACGCCCAACATGCTGTGGCTGGACTCCAC
 Tyr Tyr Ala Asn Met Leu Trp Leu Asp Ser Thr
 1620 TACCCGACAAACGAGACCTCTCCACACCCGGTGGCGGTGGGAAAGCTGCTCCACAGCTCCGGTGTCCCTGCTCAGGTCGAAATCTCAG
 Tyr Pro Thr Asn Glu Thr Ser Ser Thr Pro Gly Ala Val Arg Gly Ser Cys Ser Thr Ser Ser Gly Val Pro Ala Gln Val Glu Ser Gln
 1710 TCTCCCAACGCCAAGGTACCTTCTCAACATCAAGTTCGGACCCATTGGCAGCACCCGGCAACCCTAGCGCGGCAACCCCTCCCGGGGA
 Ser Pro Asn Ala Lys Val Thr Phe Ser Asn Ile Lys Phe Gly Pro Ile Gly Ser Thr Gly Asn Pro Ser Gly Gly Asn Pro Pro Gly Gly
 1800 AACCGTGGCACCCACCCCGCCCGCCAGCCACTACCACTGGAAGCTCTCCGGACCTACCCAGTCTCACTACGCGCCAGTGGCGGGT
 Asn Arg Gly Thr Thr Thr Arg Arg Pro Ala Thr Thr Gly Ser Ser Pro Gly Pro Thr Gln Ser His Tyr Gly Gln Cys Gly Gly
 1890 ATTGGCTACAGCGGCCCCACGGTCTGGCCAGCGGCACAACTTGCCAGGTCTGAAACCCCTACTACTCTCAGTGCCTGTAAAGCTCCGTG
 Ile Gly Tyr Ser Gly Pro Thr Val Cys Ala Ser Gly Thr Thr Cys Gln Val Leu Asn Pro Tyr Tyr Ser Gln Cys Leu
 1980 CGAAAGCCTGACGCACCGGTAGATTCTTGGTGAGCCCGGTATCATGACGGCGGGGAGCTACATGGCCCCGGGTGATTATTTTGT
 ATCTACTCTGACCCCTTTCAAATATACGGTCAACTCATCTTCACTGGAGATGGCGCTGCTGGTATTGGGATGTTCTCAGCTTGGCA
 2070 AATTGTGGCTTTCGAAAAACACAAACGATTCTCTAGTAGCCATGCATTTTAAAGATAACGGAATAGAGAAAGAGGAAATTAATAAAAAA
 2160 AAAAAACAAACATCCCGTTCAATACCCGTAGAAATCCCGCTCTCTCGTGTATCCCGAGTACCA
 2221

FIG. 9-2

14 / 16

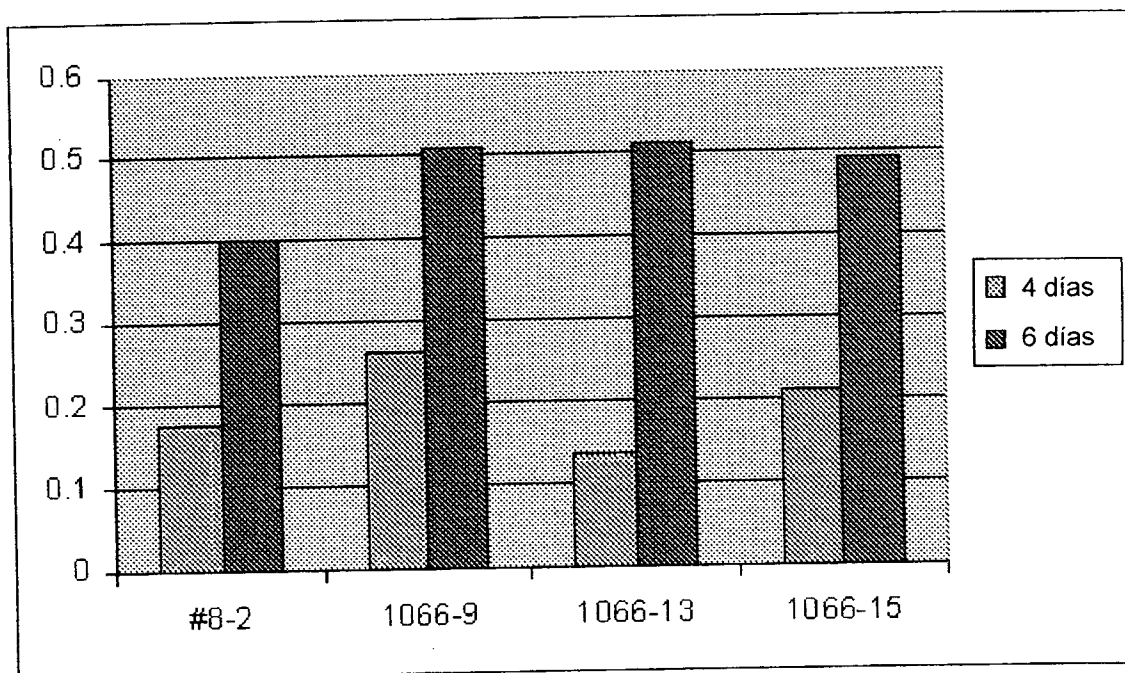


FIG. 10

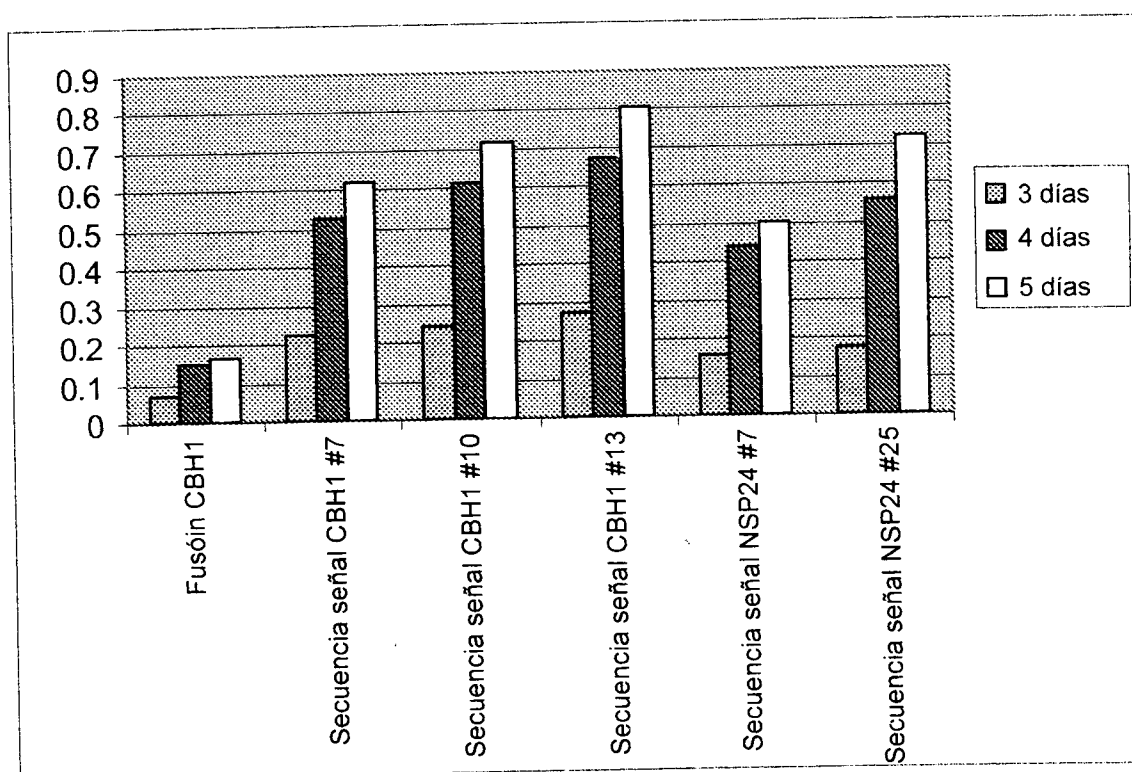


FIG. 11

15 / 16

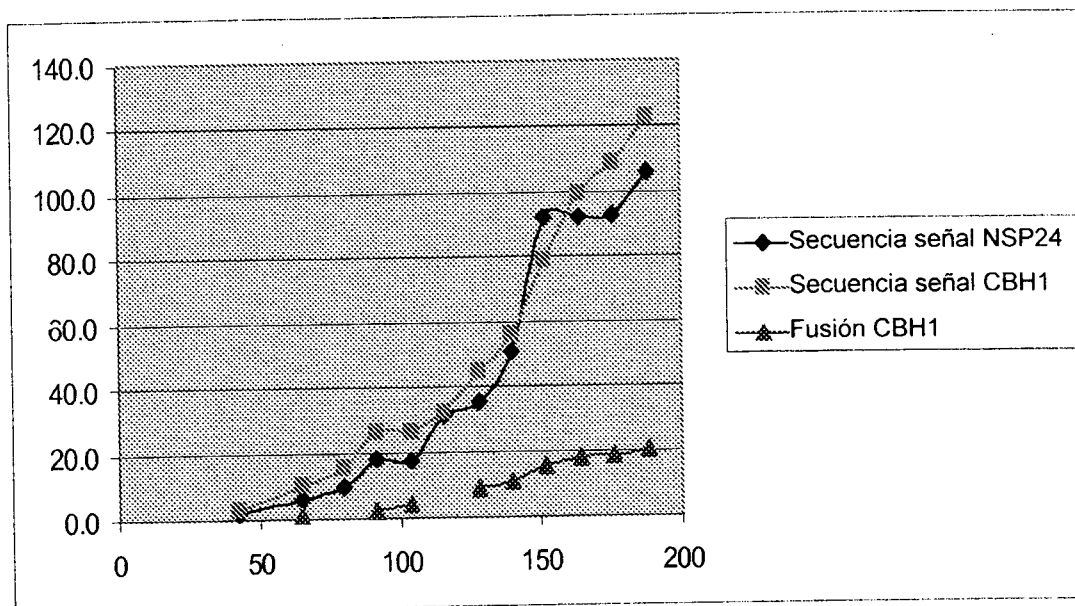


FIG. 12

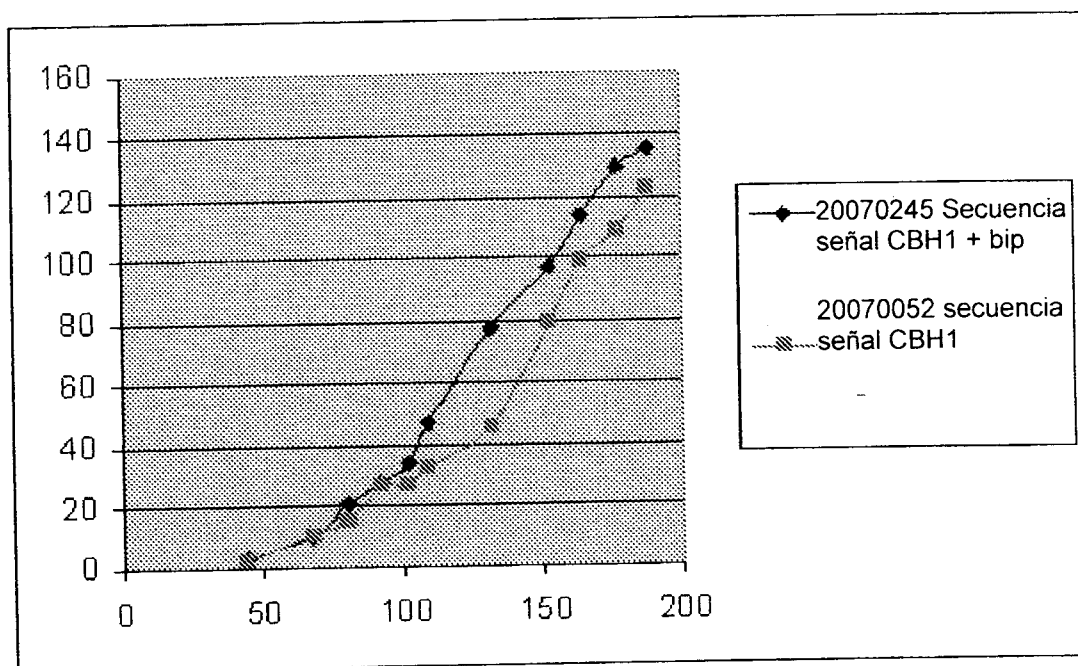


FIG. 13

16 / 16

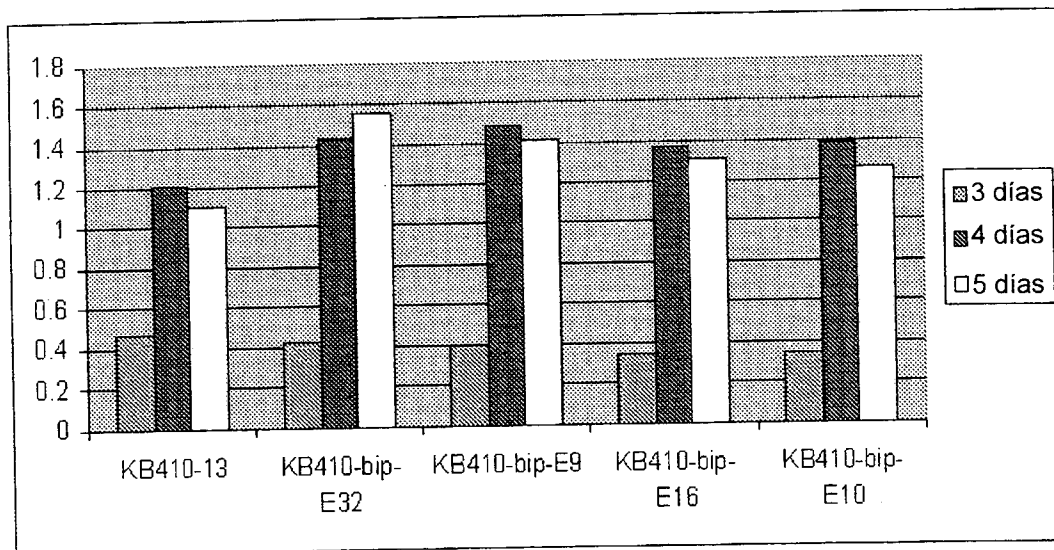


FIG. 14

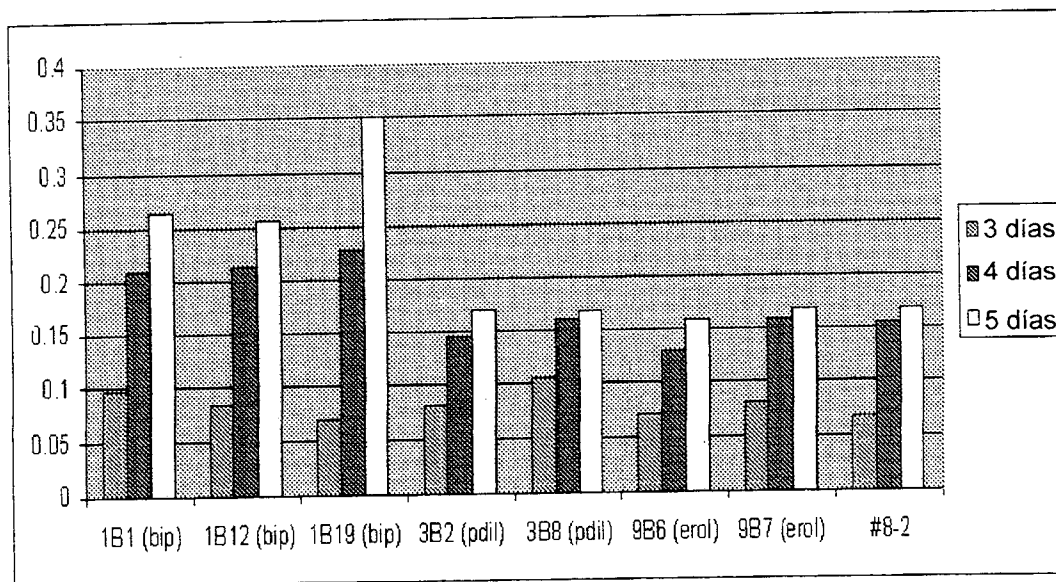


FIG. 15

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
7 May 2009 (07.05.2009)

PCT

(10) International Publication Number
WO 2009/058956 A1

(51) International Patent Classification:

C12N 9/02 (2006.01) C12N 15/62 (2006.01)
C12N 9/34 (2006.01) C12N 15/80 (2006.01)
C12N 9/42 (2006.01) C12P 21/02 (2006.01)
C12N 9/58 (2006.01)

(21) International Application Number:

PCT/US2008/081721

(22) International Filing Date: 30 October 2008 (30.10.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/984,430 1 November 2007 (01.11.2007) US

(71) Applicant (for all designated States except US):
DANISCO US, INC. [US/US]; Genencor Division,
925 Page Mill Road, Palo Alto, CA 94304-1013 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): BAO, Kai [US/US];
Danisco Us, Inc., Genencor Division, 925 Page Mill Road,
Palo Alto, CA 94304-1013 (US). WANG, Huaming
[US/US]; Danisco Us, Inc., Genencor Division, 925 Page
Mill Road, Palo Alto, CA 94304-1013 (US).

(74) Agent: KUNG, Viola, T.; Howrey LLP, 2941 Fairview
Park Drive, Box 7, Falls Church, VA 22042 (US).

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

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

Published:

— with international search report

[Continued on next page]

(54) Title: SIGNAL SEQUENCES AND CO-EXPRESSED CHAPERONES FOR IMPROVING PROTEIN PRODUCTION IN
A HOST CELL

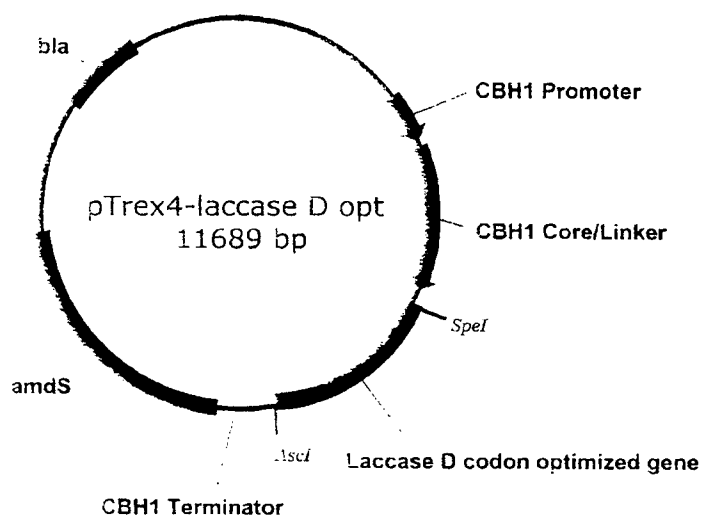


FIG. 1

(57) Abstract: The invention provides methods and compositions for improved protein production. The method comprises the steps of: (a) introducing into a host cell a first nucleic acid sequence comprising a signal sequence operably linked to a desired protein sequence; (b) expressing the first nucleic acid sequence; (c) co-expressing a second nucleic acid sequence encoding a chaperone or foldase selected from the group consisting of *bip1*, *ero1*, *pdi1*, *tig1*, *prp1*, *ppi1*, *ppi2*, *prp3*, *prp4*, calnexin, and *lht1*; and (d) collecting the desired protein secreted from the host cell. The first nucleic acid sequence optionally comprises an enzyme sequence between the signal sequence and the desired protein sequence.

WO 2009/058956 A1



— with sequence listing part of description published separately in electronic form and available upon request from the International Bureau

SIGNAL SEQUENCES AND CO-EXPRESSED CHAPERONES FOR IMPROVING PROTEIN PRODUCTION IN A HOST CELL

FIELD OF THE INVENTION

This invention provides methods and compositions for improved protein production. In some embodiments, the methods provided herein involve the use of a signal sequence operably linked to a protein. In some embodiments, the signal sequence operably linked to a protein is expressed in combination with at least one chaperone in a host cell. In some embodiments, the protein is expressed in a filamentous fungal cell. In further embodiments, the methods of the present invention involve fusion of a protein to the catalytic domain of an enzyme, such as a glucoamylase or a CBH1. Some embodiments provide combinations of a signal sequence, one or more of a chaperone, chaperonin, and/or foldase, and/or fusion of the protein to a catalytic protein or domain.

BACKGROUND OF THE INVENTION

Host cells such as yeast, filamentous fungi and bacteria have long been used to express and secrete foreign protein. Typically, production of these foreign or proteins in yeast, filamentous fungi and bacteria involves the expression and partial or complete purification of the protein from the host cell or the culture medium in which the cells are grown. While some proteins require purification from the intracellular milieu of the host cells, purification can be greatly simplified if the proteins are secreted from the cell into the culture media.

Extracellular protein secretion is a complicated and important aspect of protein production in various cell expression systems. One of the factors associated with protein secretion is proper protein folding. Many proteins can be reversibly unfolded and refolded *in vitro* at dilute concentrations, as all of the information required to specify a compact folded protein structure is present in the amino acid sequence of proteins. However, protein folding *in vivo* occurs in a concentrated milieu of numerous proteins in which intermolecular aggregation reactions compete with the intramolecular folding process. These complications are more significant in eukaryotic expression systems than in prokaryotic systems.

The first step in the eukaryotic secretory pathway is translocation of the nascent polypeptide across the endoplasmic reticulum (ER) membrane in extended form. Correct folding and assembly of a polypeptide occurs in the ER through the secretory pathway. However, in many cases, although the proteins are greatly overexpressed, they are poorly

secreted. Indeed, in many cases the secretion signals that should facilitate such expression do not appear to accomplish this. The expression of desired proteins is further complicated by the interaction of other proteins. These factors are even more significant when expression of a protein obtained from one species, genus or family of organisms is attempted in another species, genus or family. For example, Basidiomycetes proteins (*e.g.*, laccase) typically express poorly in Ascomycetes hosts such as *Trichoderma*. Indeed, despite much work in the area of fungal expression systems, there remains a need for improved extracellular expression of desired proteins.

SUMMARY OF THE INVENTION

The invention provides methods and compositions for improved protein production. The methods involve the use of a signal sequence operably linked to a desired protein, which is expressed in combination with at least one chaperone in a host cell. In some embodiments, the protein is expressed in a filamentous fungal cell. In further embodiments, the methods of the present invention involve fusion of a desired protein to the catalytic domain of a host protein, such as a glucoamylase or a CBH1.

In some embodiments, the present invention provides methods and compositions to increase the production of proteins in filamentous fungal hosts (*e.g.*, Ascomycetes), through the use of a secretory signal in combination with expression of a chaperone protein obtained from the same organism as the protein. In some embodiments, the protein is a non-Ascomycete protein that is fused to the secretory signal from an Ascomycetes host protein. In some additional embodiments, at least one chaperone protein finds use in increasing the expression of proteins fused to the catalytic domain of an Ascomycetes protein.

Some embodiments provide methods for producing at least one protein in an Ascomycetes host cell, by introducing into a host cell a polynucleotide comprising a desired protein operably linked to signal sequence from the same phylum, genus and/or species as the host; co-expressing a chaperone from the same phylum, genus and/or species as the protein; culturing the host cell under suitable culture conditions for the expression and production of the protein; and producing the protein. The method optionally includes recovering the produced protein. Some embodiments include fusing the protein to the catalytic domain of an enzyme from Ascomycetes. Other embodiments include fusing the protein to a full-length enzyme from Ascomycetes. In some embodiments, the Ascomycetes host cell is *Trichoderma*. In some embodiments, the chaperone is at least one of the following, BIP1, ERO1, PDI1, TIG1, PRP1, PPI1, PPI2, PRP3, PRP4, CALNEXIN, and LHS1.

The choice of protein is not limiting, and can include any of the following proteins from any genus, species, and/or family: laccases, glucoamylases, alpha amylases, granular starch hydrolyzing enzymes, cellulases, lipases, xylanases, cutinases, hemicellulases, proteases, oxidases, laccases and combinations thereof. Some embodiments include signal sequences from NSP24 or CBH1 genes. In some embodiments, the chaperone gene is *bip1*. Embodiments of the method can also include an Ascomycetes promoter. In some embodiments, the host cell and the signal sequence is from the same Ascomycetes host. In some embodiments, the promoter is the CBH1 promoter from *Trichoderma*. In some embodiments, the protein is a Basidiomycetes protein. In some embodiments, the host cell is an Ascomycetes host cell. In some embodiments, the host cell is a Basidiomycetes host cell and the protein is an Ascomycetes protein.

Some further embodiments provide methods for producing at least one protein in an Ascomycetes host cell, by introducing into an Ascomycetes host cell a polynucleotide comprising a desired protein fused to the catalytic domain of an enzyme from Ascomycetes, wherein the desired protein is a Basidiomycetes protein; co-expressing an Ascomycetes chaperone; culturing the Ascomycetes host cell under suitable culture conditions for the expression and production of the protein; and producing the protein. In some embodiments, the produced protein is recovered. In some embodiments, the protein is operably linked to an Ascomycetes signal sequence.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the schematic of the *Trichoderma* expression plasmid pTrex4-laccaseD opt. The polynucleotide sequence is shown as SEQ ID NO: 1.

Figure 2 shows the schematic of the *Trichoderma* expression plasmid pTrex2g-Bip1. The polynucleotide sequence is shown as SEQ ID NO: 2.

Figure 3 shows the schematic of the *Trichoderma* expression plasmid pTrex2g-Pdi1. The polynucleotide sequence is shown as SEQ ID NO: 3.

Figure 4 shows the schematic of the *Ero1* sequence used in the *Trichoderma* expression plasmid pTrex2g-Ero1. The polynucleotide sequence is shown as SEQ ID NO: 4.

Figure 5 shows the schematic of the *Trichoderma* expression plasmid pTrGA-laccaseD opt. The polynucleotide sequence is shown as SEQ ID NO: 5.

Figure 6 shows the schematic of the *Trichoderma* expression plasmid pKB408. The polynucleotide sequence is shown as SEQ ID NO: 6.

Figure 7 shows the schematic of the *Trichoderma* expression plasmid pKB410. The polynucleotide sequence is shown as SEQ ID NO: 7.

Figures 8-1 to 8-4 show the *T. reesei* NSP24 Open Reading frame (ORF) SEQ ID NO:8. The signal peptide is the first 20 amino acids (SEQ ID NO: 9).

Figures 9-1 and 9-2 show the *T. reesei* CBH1 ORF (SEQ ID NO: 10). The signal sequence begins at base pair 210 and ends at base pair 260 (SEQ ID NO: 11). The catalytic core begins at base pair 261 through base pair 1698 (SEQ ID NO: 12), including intron 1 (from base pair 671 to 737) and intron 2 (from base pair 1435 to 1497). The linker sequence begins at base pair 1699 and ends at base pair 1770 (SEQ ID NO: 13). The CBH1 protein sequence is shown as SEQ ID NO: 14.

Figure 10 illustrates the improvement of laccase production by fusion of the gene encoding *C. unicolor* laccase to the full-length *Trichoderma* glucoamylase. Strain #8-2 is CBH1 laccase fusion. Strain 1066-9, 1066-13, and 1066-15 are TrGA laccase fusion.

Figure 11 illustrates the improvement of laccase production by fusion of the gene encoding *C. unicolor* laccase to the CBH1 or NSP24 signal sequence in shake flasks. Y axis shows the laccase activity as units/ml. X axis shows the strains (CBH1 fusion alone, or with signal sequence).

Figure 12 illustrates the improvement of laccase production by fusion of the gene encoding *C. unicolor* laccase to the CBH1 or NSP24 signal sequence in fermentors. Y axis shows the laccase activity as units/ml. X axis shows the fermentation time as hours.

Figure 13 illustrates the improvement of laccase production provided by the CBH1 signal sequence plus BIP1 chaperone expression. Y axis shows the laccase activity as units/ml. X axis shows the fermentation time as hours.

Figure 14 illustrates the improvement of laccase production by co-expression of chaperones with *C. unicolor* in shake flasks at 3, 4, and 5 days. Y axis shows the laccase activity as units/ml. X axis shows the strains (KB410-13, or with co-expression of bip).

Figure 15 illustrates the improvement of laccase production by fusion of the gene encoding *C. unicolor* laccase to the CBH1 signal sequence, catalytic domain and linker and co-expression with Bip1, pdi1 or ero1 chaperone. Y axis shows the laccase activity as units/ml. X axis shows the strains.

DETAILED DESCRIPTION OF THE INVENTION

Unless otherwise indicated, the practice of the present invention involves conventional techniques commonly used in molecular biology, protein engineering, recombinant DNA techniques, microbiology, cell biology, cell culture, transgenic biology, immunology, and protein purification, which are within the skill of the art. Such techniques are known to those of skill in the art and are described in numerous texts and reference works. All patents, patent applications, articles and publications mentioned herein, both *supra* and *infra*, are hereby expressly incorporated herein by reference.

Definitions

The term "Ascomycetes" refers to a class of fungi belonging to the phylum Ascomycota. Members of this phylum are distinguished by the presence of asci (*i.e.*, specialized sac-like cells that contain ascospores).

The term "Basidiomycetes" refers to a class of fungi belonging to the phylum Basidiomycota. Members of this phylum are characterized by the production of basidiospores, (*i.e.*, sexual spores that are located on external areas of specialized club-shaped end cells referred to as basidia).

"Protease" means a protein or polypeptide domain of a protein or polypeptide that has the ability to catalyze cleavage of peptide bonds at one or more of various positions of a protein backbone (*e.g.* E.C. 3.4). Proteases are obtainable from microorganisms (*e.g.* a fungi or bacteria), plants, and/ or animals.

An "acid protease" refers to a protease having the ability to hydrolyze proteins under acidic conditions.

As used herein, the term "chaperone" or "molecular chaperones" facilitate protein folding by shielding unfolded regions from surrounding proteins and do not enhance the rate of protein folding. This can include proteins and their homologs that assist the folding and glycosylation of the secretory proteins in the endoplasmic reticulum (ER). Chaperones may be resident in the ER. Exemplary chaperones include Bip (GRP78), GRP94 and yeast Lhs1p and those help the secretory protein to fold by binding to exposed hydrophobic regions in the unfolded states and preventing unfavorable interactions. Chaperones also include proteins that are involved in translocation of proteins through the ER membrane.

As used herein, "chaperonins" are proteins that assist protein folding to the native state (active state) utilizing ATP. Often the protein subunits are assembled together to form a large ring assemblies. For example, chaperonins act by binding nonnative proteins in their

central cavities and then, upon binding ATP, release the substrate protein into a now-encapsulated cavity to fold productively.

“Foldase proteins” means proteins that catalyze steps in protein folding to increase the rate of protein folding. For example, they can assist in formation of disulphide bridges and formation of the right conformation of peptide chains adjacent to proline residues. Exemplary foldases include protein disulphide isomerase (pdi) and its homologs and prolyl-peptidyl cis-trans isomerase and its homologs.

As used herein, “NSP24 family protease” means an enzyme having protease activity in its native or wild type form that belonging to the family of NSP24 proteases. NSP24 proteases are acid proteases, such as acid fungal proteases. The NSP24 proteases have at least 85%, at least 90%, at least 93%, at least 95%, at least 96%, at least 97%, at least 98% and at least 99% sequence identity to the amino acid sequence of SEQ ID NO: 8 and biologically active fragments thereof.

As used herein, the term “a desired protein” means a protein of interest. A desired protein and a protein of interest are used interchangeably in this application. In some embodiments, the desired protein is a commercially important industrial protein. It is intended that the term encompass proteins that are encoded by naturally occurring genes, mutated genes and/or synthetic genes. The desired protein can be a protein native to the host cell, or non-native (heterologous) to the host cell.

As used herein, “derivative” means a protein which is derived from a precursor or parent protein (*e.g.*, the native protein) by addition of one or more amino acids to either or both the C- and N-terminal end(s), substitution of one or more amino acids at one or a number of different sites in the amino acid sequence, deletion of one or more amino acids at either or both ends of the protein or at one or more sites in the amino acid sequence, or insertion of one or more amino acids at one or more sites in the amino acid sequence.

The term “recombinant” refers to a polynucleotide or polypeptide that does not naturally occur in a host cell. A recombinant molecule may contain two or more naturally occurring sequences that are linked together in a way that does not occur naturally.

The terms “peptides,” “proteins,” and “polypeptides” are used interchangeably herein.

As used herein, “percent (%) sequence identity” with respect to amino acid or nucleotide sequences is defined as the percentage of amino acid residues or nucleotides in a candidate sequence that are identical with the amino acid residues or nucleotides in a sequence of interest (*e.g.* a NSP24 signal peptide sequence), after aligning the sequences and

introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity.

As used herein, the term "alpha-amylase (*e.g.*, E.C. class 3.2.1.1)" refers to enzymes that catalyze the hydrolysis of alpha-1,4-glucosidic linkages. These enzymes have also been described as those effecting the exo or endohydrolysis of 1,4- α -D-glucosidic linkages in polysaccharides containing 1,4- α -linked D-glucose units. Another term used to describe these enzymes is "glycogenase." Exemplary enzymes include alpha-1,4-glucan 4-glucanohydrase glucanohydrolase.

As used herein, the term "glucoamylase" refers to the amyloglucosidase class of enzymes (*e.g.*, EC.3.2.1.3, glucoamylase, 1, 4-alpha-D-glucan glucohydrolase). These are exo-acting enzymes, which release glucosyl residues from the non-reducing ends of amylose and amylopectin molecules. The enzyme also hydrolyzes alpha-1, 6 and alpha -1,3 linkages although at much slower rate than alpha-1, 4 linkages.

The term "promoter" means a regulatory sequence involved in binding RNA polymerase to initiate transcription of a gene.

A "heterologous promoter" as used herein refers to a promoter that has been placed in association with a gene or purified nucleic acid, but which is not naturally associated with that gene or purified nucleic acid.

A "purified preparation" and "substantially pure preparation" of a polypeptide, as used herein, mean a polypeptide that has been separated from cells, other proteins, lipids or nucleic acids with which it naturally occurs.

"Homologous," as used herein, refers to the sequence similarity between two or more polypeptide molecules or between two or more nucleic acid molecules. When a position in the sequences being compared is occupied by the same base or amino acid monomer subunit, (*e.g.*, if a position in each of two DNA molecules is occupied by adenine), then the molecules are homologous at that position. The percent of homology between two sequences is a function of the number of matching or homologous positions shared by the two sequences divided by the number of positions compared x 100. For example, if 6 of 10, of the positions in two sequences are matched or homologous then the two sequences are 60% homologous. By way of example, the DNA sequences ATTGCC and TATGGC share 50% homology. Generally, a comparison is made when two sequences are aligned to give maximum homology. The term "% homology" is used interchangeably herein with the term "% identity" herein and refers to the level of nucleic acid or amino acid sequence identity

between the nucleic acid sequences or amino acid sequences, when aligned using a sequence alignment program.

As used herein, the term "vector" refers to a polynucleotide sequence designed to introduce nucleic acids into one or more cell types. Vectors include cloning vectors, expression vectors, shuttle vectors, plasmids, phage particles, cassettes and the like.

As used herein, "expression vector" means a DNA construct including a DNA sequence which is operably linked to a suitable control sequence capable of affecting the expression of the DNA in a suitable host.

The term "expression" means the process by which a polypeptide is produced based on the nucleic acid sequence of a gene.

The term "co-expression" means that at least two different genes are expressed in one cell. They can be exogenous genes, or endogenous genes. They can be integrated or expressed from the same or different plasmids, and they can be expressed from the same or different promoter.

As used herein, "operably linked" means that a regulatory region, such as a promoter, terminator, secretion signal or enhancer region is attached to or linked to a structural gene and controls the expression of that gene. A signal sequence is operably linked to a protein if it directs the protein through the secretion system of a host cell.

As used herein, "microorganism" refers to a bacterium, a fungus, a virus, a protozoan, and other microbes or microscopic organisms.

The term "filamentous fungi" refers to all filamentous forms of the subdivision Eumycotina, as known in the art. These fungi are characterized by a vegetative mycelium with a cell wall composed of chitin, cellulose, and other complex polysaccharides. The filamentous fungi of the present invention are morphologically, physiologically, and genetically distinct from yeasts. Vegetative growth by filamentous fungi is by hyphal elongation and carbon catabolism is obligatory aerobic.

As used herein, the term "*Trichoderma*" and "*Trichoderma sp.*" refer to any fungal genus previously or currently classified as *Trichoderma*.

As used herein the term "culturing" refers to growing a population of microbial cells under suitable conditions in a liquid, semi-solid or solid medium. In some embodiments, culturing is conducted in a vessel or reactor, as known in the art. In some embodiments, culturing results in the fermentative bioconversion of a starch substrate, such as a substrate comprising granular starch, to an end-product.

“Fermentation” refers to the enzymatic and anaerobic breakdown of organic substances by microorganisms to produce simpler organic compounds. While fermentation often occurs under anaerobic conditions, it is not intended that the term be solely limited to strict anaerobic conditions, as fermentation also occurs in the presence of oxygen.

The term “introduced” in the context of inserting a nucleic acid sequence into a cell, means “transfection,” “transformation” or “transduction,” and includes reference to the incorporation of a nucleic acid sequence into a eukaryotic or prokaryotic cell wherein the nucleic acid sequence is either incorporated into the genome of the cell (*e.g.*, chromosome, plasmid, plastid, or mitochondrial DNA), converted into an autonomous replicon, or transiently expressed (*e.g.*, transfected mRNA).

As used herein, the terms “transformed,” “stably transformed” and “transgenic” used in reference to a cell means the cell has a non-native nucleic acid sequence integrated into its genome or as an episomal plasmid that is maintained through multiple generations.

As used herein, the term “heterologous” used in reference to a polypeptide or a polynucleotide encoding a desired protein means a polypeptide or polynucleotide that does not naturally occur in a host cell.

The term “homologous” or “endogenous” with reference to a polypeptide or a polynucleotide encoding a desired protein refers to a polypeptide or a polynucleotide that occurs naturally in or is naturally expressed by the host cell.

The term “overexpression” means the process of expressing a polypeptide in a host cell at a level that is greater than that produced by a wild-type host cell. In some embodiments, at least one polynucleotide is introduced into the host cell. In some further embodiments, the term refers to the expression of a homologous polypeptide at a concentration that is greater than that expression of the same homologous polypeptide expressed by a wild-type cell.

As described herein, one aspect of the invention features a “substantially pure” nucleic acid that comprises a nucleotide sequence encoding an NSP24 signal peptide or CBH1 signal peptide operably linked to a protein, and/or equivalents of such nucleic acids. In these embodiments, the nucleic acid is isolated from other nucleic acids and/or cell constituents.

The term “equivalent” refers to nucleotide sequences encoding functionally equivalent polypeptides. Equivalent nucleotide sequences encompass sequences that differ by one or more nucleotide substitutions, additions and/or deletions, such as allelic variants. For example in some embodiments, due to the degeneracy of the genetic code equivalent nucleotide sequences include sequences that differ from the nucleotide sequence of SEQ ID

NO:8, but that result in the production of polypeptides that are functionally equivalent to the polypeptide sequence encoded by SEQ ID NO:8.

This invention provides a method for producing a desired protein. The method comprises the steps of: (a) introducing into a host cell a first nucleic acid sequence comprising a signal sequence operably linked to a desired protein sequence; (b) expressing the first nucleic acid sequence; (c) co-expressing a second nucleic acid sequence encoding a chaperone or foldase selected from the group consisting of *bip1*, *ero1*, *pdi1*, *tig1*, *prp1*, *ppi1*, *ppi2*, *prp3*, *prp4*, *calnexin*, and *lhs1*; and (d) collecting the desired protein secreted from the host cell.

In one embodiment, the first nucleic acid sequence further comprises an enzyme sequence between the signal sequence and the desired protein sequence. For example, the enzyme sequence is obtained from a glucoamylase or from a CBH1 enzyme. In one embodiment, the enzyme sequence is a full-length enzyme sequence comprising a catalytic domain, a linker, and a binding domain. In another embodiment, the enzyme sequence comprises a catalytic domain sequence, which is linked to the desired protein sequence by a linker. In some embodiments, the enzyme is a host protein that is highly expressed and/or secreted in its natural host.

The first nucleic acid sequence further comprises a promoter upstream to a signal sequence. In one embodiment, the promoter is native to the host cell and is not naturally associated with the desired protein sequence.

The second nucleic acid sequence is operably linked to a promoter. In one embodiment, the promoter is native to the host cell and is not naturally associated with the second nucleic acid sequence.

Increased Expression of Proteins

The present invention provides a method for the production of a desired protein in a host cell. The protein production is increased by inclusion of a secretory signal (e.g. NSP24 signal peptide or CBH1 signal peptide) in combination with co-expression of a chaperone, chaperonin, and/or foldase protein. In some embodiments, the secretory signal is from an Ascomycetes host protein. In some embodiment, the desired protein is fused to the catalytic domain of an enzyme.

The present invention provides significant advantages, especially in view of the fact that it can be difficult to produce large amounts of proteins from other fungi families in Ascomycete hosts. Indeed, those skilled in the art know that it is often difficult to produce

any heterologous fungal protein in fungal or bacterial hosts. The present invention provides methods and compositions suitable for the production of any suitable protein in a suitable fungal or bacterial host. In some embodiments, the fungal host is an Ascomycetes and the protein is a Basidiomycetes protein, while in other embodiments, the fungal host is a Basidiomycetes and the protein is an Ascomycetes protein.

In some embodiments, the present invention provides methods for increasing expression and/or secretion of a protein in a host using a host signal peptide in combination with co-expression of one or more chaperones or foldases from the same organism as the source of the protein. Thus, in some embodiments, a heterologous Ascomycetes protein is expressed in a Basidiomycetes host using a Basidiomycetes host signal peptide and an Ascomycetes chaperone. In some alternative embodiments, a heterologous Basidiomycetes protein is expressed in an Ascomycetes host using an Ascomycetes signal peptide and an Ascomycetes or Basidiomycetes chaperone. In some embodiments, the Ascomycetes host is a member of the *Trichoderma* genus. In some embodiments, the *Trichoderma* is *Trichoderma reesei*, including various strains of *T. reesei*. In some alternative embodiments, the Basidiomycetes is a member of the genus *Cerrena*, including but not limited to *C. unicolor*.

In some embodiments of the present invention, expression and/or secretion of a desired protein is increased by fusing the protein to a host enzyme in combination with exogenous co-expression of one or more chaperones from the same organism as the desired protein. Co-expression is accomplished either via the same plasmid, or via separate plasmids.

In yet additional embodiments, expression and/or secretion of a desired protein is increased by linking the protein to a the catalytic domain of a host enzyme, in combination with operably linking the protein to a host signal sequence, and exogenous co-expression of one or more chaperones, chaperonins, and/or foldases, preferably from the same organism as the protein.

It is contemplated that elements recited in various embodiments provided herein will find use in any suitable combination. Thus, it is not intended that the embodiments be limited to the specific recitations provided herein, as aspects of the various embodiments find use in combination with each other.

Signal Peptides

The specific signal peptide used in the present invention is not critical, as long as the signal peptide is operable in the host. An "operable signal peptide" is provided when the signal peptide increases secretion of a protein when operably linked to the protein in a host

cell. In some embodiments, the signal peptide is obtained from a strongly secreted protein and/or is a strong signal peptide. A "strong signal peptide" results when the natural protein is strongly secreted by its natural host. In some embodiments, the signal peptide is obtained from an organism within the same phylum as the host cell. Indeed, in some embodiments, this is advantageous. In some embodiments, the signal peptide and the host cell are of the same genus, while in some additional embodiments, the signal peptide and the host cell are of the species. For example, in some embodiments, the host cell is an Ascomycetes host cell and the signal peptide is obtained from Ascomycetes. In some embodiments, the host cell is a *Trichoderma* and the signal peptide is from a *Trichoderma*. In some embodiments, the host cell is *T. reesei* and the signal peptide is obtained from *T. reesei*. In some embodiments, the signal peptide is a strong signal peptide. In some alternative embodiments, the host cell is a Basidiomycetes host cell and the signal peptide is obtained from Basidiomycetes. Some examples of signal peptides that find use in the present invention include, but are not limited to CBHI and NSP24 signal peptides. While the signal peptides can work in other members of a phylum such as Ascomycetes, in some embodiments, signal peptides find optimum use when used in the genus from which it was obtained (*i.e.*, to provide strong secretion).

As used herein, a "strongly secreted protein" is any protein that forms a significant amount of the total protein secreted from the cell. The total protein secreted from the cell is also referred to as "extracellular protein." For example, a strongly secreted protein includes at least about 2% of the extracellular protein, at least about 3%, at least about 4%, at least about 5%, at least about 6%, at least about 7%, at least about 8%, at least about 9%, at least about 10%, at least about 15%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or at least about 99%. In some embodiments, the strongly secreted protein comprises at least about 5% of the extracellular protein in the culture supernatant.

CBHI Signal Peptides, Linkers, and Catalytic Domains

Trichoderma reesei produces several cellulase enzymes, including cellobiohydrolase I (CBHI), which are folded into two separate domains (*i.e.*, catalytic and binding domains) that are separated by an extended linker region. Foreign polypeptides have been secreted in *T. reesei* as fusions with the catalytic domain plus linker region of CBHI (See *e.g.*, Nyysönen *et al.*, Bio/Technol. 11:591-595 [1993]). *T. longibrachiatum* also produces a CBHI that finds use in fusions, as well as in the isolation of a signal peptide and/or a linker. Linkers find use

in connecting a catalytic domain of an enzyme and the desired polypeptide. Any suitable linker finds use in the present invention, as long as it forms an extended, semi-rigid spacer between independently folded domains. Such linker regions are found in several proteins, especially hydrolases (*e.g.*, bacterial and fungal cellulases and hemicellulases; *See e.g.*, Libby *et al.*, Protein Engineering, Design and Selection (1994) vol. 7, 1109-1114).

As shown in Figure 9, for CBHI (SEQ ID NO: 10), the signal sequence begins at base pair 210 and ends at base pair 260 (SEQ ID NO: 11). The catalytic core begins at base pair 261 through base pair 1698 (SEQ ID NO: 12), including intron 1 (from base pair 671 to 737) and intron 2 (from base pair 1435 to 1497). The linker sequence begins at base pair 1699 and ends at base pair 1770 (SEQ ID NO: 13). The cellulose binding domain begins at base pair 1771 through base pair 1878. The sequence and domain information for CBHI can be found via the expasy organization website and is designated uniprot/P62694. CBHI homologs have been identified in a number of other *Trichoderma species* as well as other filamentous fungi and find use in the present invention as appropriate.

NSP24 Signal Peptides and Polynucleotides

The NSP24 gene was isolated and sequenced from *T. reesei* (*See e.g.*, U.S. Patent No. 7,429,476, which is incorporated herein by reference in its entirety). Sequencing of this gene identified a sequence encoding a 407 amino acid open reading frame (SEQ ID NO: 8), as shown in Figure 8. A signal peptide was identified as the first 20 amino acids (MQTFGAFLVSFLAASGLAAA; SEQ ID NO: 9) of SEQ ID NO: 8. NSP24 homologs have been identified in a number of other *Trichoderma species* as well as other filamentous fungi and find use in the present invention as appropriate. In some embodiments, the NSP24 signal sequence is used in an Ascomycetes organism. In some embodiments, the sequence is used in *Trichoderma spp.*, and in some even more particularly embodiments, in *T. reesei*.

Thus, the present invention provides NSP24 family protease signal peptides that find use in secreting a protein. In some embodiments, the NSP24 signal peptide is designated "NSP24 aspartic protease signal peptide."

Polynucleotides of the Invention

The present invention provides various polynucleotides, including but not limited to polynucleotides encoding desired proteins, signal peptides, catalytic domains, linkers, chaperones, chaperonins and foldases. In some embodiments, polynucleotides comprise at

least two of the above. In yet other embodiments, the polynucleotides of the present invention comprise at least three of the above.

In some embodiments, the polynucleotides encode proteins that comprise at least one amino acid substitution such as a "conservative amino acid substitution" using L-amino acids, wherein one amino acid is replaced by another biologically similar amino acid. Conservative amino acid substitutions are those that preserve the general charge, hydrophobicity/hydrophilicity, and/or steric bulk of the amino acid being substituted. Examples of conservative substitutions are those between the following groups: Gly/Ala, Val/Ile/Leu, Lys/Arg, Asn/Gln, Glu/Asp, Ser/Cys/Thr, and Phe/Trp/Tyr. In some embodiments, "derivative proteins" find use in the present invention. In some of these embodiments, the derivative proteins differ by as few as about 1 to about 10 amino acid residues, such as about 6 to about 10, as few as about 5, as few as about 4, about 3, about 2, or even 1 amino acid residue, compared to the "parent" protein sequence. Table 1 provides exemplary conservative amino acid substitutions recognized in the art. In additional embodiments, substitution involves one or more non-conservative amino acid substitutions, deletions, or insertions that do not abolish the signal peptide activity.

Table 1. Conservative Amino Acid Replacements

For Amino Acid	One Letter Code	Replace with Any Of the Following
Alanine	A	D-Ala, Gly, beta-Ala, L-Cys, D-Cys
Arginine	R	D-Arg, Lys, D-Lys, homo-Arg, D-homo-Arg, Met, Ile, D-Met, D-Ile, Orn, D-Orn
Asparagine	N	D-Asn, Asp, D-Asp, Glu, D-Glu, Gln, D-Gln
Aspartic Acid	D	D-Asp, D-Asn, Asn, Glu, D-Glu, Gln, D-Gln
Cysteine	C	D-Cys, S-Me-Cys, Met, D-Met, Thr, D-Thr
Glutamine	Q	D-Gln, Asn, D-Asn, Glu, D-Glu, Asp, D-Asp
Glutamic Acid	E	D-Glu, D-Asp, Asp, Asn, D-Asn, Gln, D-Gln
Glycine	G	Ala, D-Ala, Pro, D-Pro, b-Ala, Acp

For Amino Acid	One Letter Code	Replace with Any Of the Following
Isoleucine	I	D-Ile, Val, D-Val, Leu, D-Leu, Met, D-Met
Leucine	L	D-Leu, Val, D-Val, Leu, D-Leu, Met, D-Met
Lysine	K	D-Lys, Arg, D-Arg, homo-Arg, D-homo-Arg, Met, D-Met, Ile, D-Ile, Orn, D-Orn
Methionine	M	D-Met, S-Me-Cys, Ile, D-Ile, Leu, D-Leu, Val, D-Val
Phenylalanine	F	D-Phe, Tyr, D-Thr, L-Dopa, His, D-His, Trp, D-Trp, Trans-3,4, or 5-phenylproline, cis-3, 4, or 5-phenylproline
Proline	P	D-Pro, L-I-thioazolidine-4-carboxylic acid, D-or L-1-oxazolidine-4-carboxylic acid
Serine	S	D-Ser, Thr, D-Thr, allo-Thr, Met, D-Met, Met(O), D-Met(O), L-Cys, D-Cys
Threonine	T	D-Thr, Ser, D-Ser, allo-Thr, Met, D-Met, Met(O), D-Met(O), Val, D-Val
Tyrosine	Y	D-Tyr, Phe, D-Phe, L-Dopa, His, D-His
Valine	V	D-Val, Leu, D-Leu, Ile, D-Ile, Met, D-Met

In some embodiments, the polynucleotides of the invention are native sequences. In some embodiments, the native sequences are isolated from nature, while in other embodiments they are produced by recombinant or synthetic means. The term "native sequence" specifically encompasses naturally-occurring truncated or secreted forms (*e.g.*, biologically active fragments), and naturally-occurring variant forms of the native sequences.

Because of the degeneracy of the genetic code, more than one codon may be used to code for a particular amino acid. Therefore, in some embodiments, different DNA sequences are used to encode any of the polypeptides such as the signal peptide, the protein, the catalytic domain, and/or the chaperones. Indeed, it is intended that the present invention encompass different polynucleotide sequences that which encode the same polypeptide.

A nucleic acid is hybridizable to another nucleic acid sequence when a single stranded form of the nucleic acid can anneal to the other nucleic acid under appropriate conditions of temperature and solution ionic strength. Hybridization and washing conditions are well known in the art for hybridization under low, medium, high and very high stringency conditions. In general, hybridization involves a nucleotide probe and a homologous DNA sequence that form stable double stranded hybrids by extensive base-pairing of complementary polynucleotides. In some embodiments, the filter with the probe and homologous sequence are washed in 2x sodium chloride/sodium citrate (SSC), 0.5% SDS at about 60°C (medium stringency), 65°C (medium/high stringency), 70°C (high stringency) and about 75°C (very high stringency) (*See e.g.*, Current Protocols in Molecular Biology, John Wiley & Sons, New York, 1989, 6.3.1 - 6.3.6, hereby incorporated by reference);

The present invention encompasses allelic variations, natural mutants, induced mutants, proteins encoded by DNA that hybridizes under high or low stringency conditions to a nucleic acid which encodes a laccase, a signal sequence of NSP24, a signal sequence of CBHI, catalytic domains, chaperones, chaperonins and foldases. Nucleic acids and polypeptides of the present invention include those that differ from the sequences disclosed herein by virtue of sequencing errors in the disclosed sequences.

"Homology of DNA sequences" is determined by the degree of identity between two DNA sequences. Homology or "percent identity" is often determined for polypeptide sequences and/or nucleotides sequences using computer programs. Methods for performing sequence alignment and determining sequence identity are well-known to the skilled artisan, may be performed without undue experimentation, and calculations of identity values are obtainable with definiteness. A number of algorithms are available and known to those of skill in the art, for aligning sequences and determining sequence identity. Computerized programs using these algorithms are also available and well-known to those in the art, including, but are not limited to: ALIGN or Megalign (DNASTAR) software, or WU-BLAST-2, GAP, BESTFIT, BLAST, FASTA, TFASTA, and CLUSTAL. Those skilled in the art know how to determine appropriate parameters for measuring alignment, including algorithms needed to achieve maximal alignment over the length of the sequences being compared. The sequence identity can be determined using the default parameters determined by the program. In some embodiments, sequence identity is determined by the Smith-Waterman homology search algorithm (Smith Waterman, Meth. Mol. Biol., 70:173-187 [1997]) as implemented in MSPRCH program (Oxford Molecular) using an affine gap search with the following search parameters: gap open penalty of 12, and gap extension penalty of 1.

Paired amino acid comparisons can be carried out using the GAP program of the GCG sequence analysis software package of Genetics Computer Group, Inc. (Madison, WI), employing the blosum62 amino acid substitution matrix, with a gap weight of 12 and a length weight of 2. With respect to optimal alignment of two amino acid sequences, the contiguous segment of the variant amino acid sequence may have additional amino acid residues or deleted amino acid residues with respect to the reference amino acid sequence. The contiguous segment used for comparison to the reference amino acid sequence will include at least about 20 contiguous amino acid residues, and may be about 30, about 40, about 50, or more amino acid residues. In some embodiments, corrections for increased sequence identity associated with inclusion of gaps in the derivative's amino acid sequence are made by assigning gap penalties.

In some embodiments, the protein, signal peptide, enzyme catalytic domain, chaperone, chaperonin, and/or foldase encompassed by the invention is derived from a bacterium or a fungus, such as a filamentous fungus. Exemplary filamentous fungi include *Aspergillus* spp. and *Trichoderma* spp. One exemplary *Trichoderma* spp. is *T. reesei*. However, in some embodiments, the signal peptide and/or DNA encoding the signal peptide provided by the present invention is derived from another genus or species of fungi, including but not limited to *Absidia* spp.; *Acremonium* spp.; *Agaricus* spp.; *Anaeromyces* spp.; *Aspergillus* spp., including, but not limited to *A. aculeatus*, *A. awamori*, *A. flavus*, *A. foetidus*, *A. fumigatus*, *A. nidulans*, *A. niger*, *A. oryzae*, *A. terreus* and *A. versicolor*; *Aeurobasidium* spp.; *Cerrena* spp.; *Cephalosporium* spp.; *Cephalosporium* spp.; *Chaetomium* spp.; *Coprinus* spp.; *Dactyllum* spp.; *Dactylium* spp.; *Fusarium* spp., including *F. conglomerans*, *F. decemcellulare*, *F. javanicum*, *F. lini*, *F. oxysporum* and *F. solani*; *Gliocladium* spp.; *Humicola* spp., including *H. insolens* and *H. lanuginosa*; *Mucor* spp.; *Neurospora* spp., including *N. crassa* and *N. sitophila*; *Neocallimastix* spp.; *Orpinomyces* spp.; *Penicillium* spp.; *Phanerochaete* spp.; *Phlebia* spp.; *Piromyces* spp.; *Rhizopus* spp.; *Schizophyllum* spp.; *Stachybotrys* spp.; *Trametes* spp.; *Trichoderma* spp., including *T. reesei*, *T. reesei* (*longibrachiatum*) and *T. viride*; and *Zygorhynchus* spp.

Catalytic Domain Fusion

Fusing a desired protein to an enzyme often allows for increased expression and/or secretion of the desired protein. In general, the enzyme sequence is upstream to the desired protein sequence in the construct. For example, the enzyme is obtained from a glucoamylase or from a CBH1 enzyme. In one embodiment, the enzyme sequence is a full-length enzyme

sequence comprising a catalytic domain, a linker, and a binding domain. In another embodiment, the enzyme sequence comprises a catalytic domain sequence, which is linked to the desired protein sequence by a linker or a portion of the linker. In some embodiments, the enzyme is a host protein that is highly expressed and/or secreted in its natural host. For example, when the host cell is a *Trichoderma* host cell, the enzyme is from a *Trichoderma* protein. However, it is to be understood that many filamentous fungal proteins find use in fusion to proteins and can be used in other filamentous fungal hosts with success.

Chaperones, Chaperonins and Foldases

The specific chaperone, chaperonin, and/or foldase used in the methods and polynucleotides included in the invention is not critical. Further, when describing the uses of chaperone, chaperonin, and/or foldase herein, they are used interchangeably in a method. For example, when describing a method using a chaperone, it is to be understood that a foldase and/or chaperonin could be used in place of or in addition to the recited chaperone. Chaperone, chaperonin, and/or foldase suitable for this invention are those that are active in a host cell and act to increase expression of the desired protein.

In some embodiments, the chaperone, chaperonin, and/or foldase is from the same phylum of organisms as the protein, and can be from the same genus, and can also be from the same genus and species. In some embodiments, the chaperone, chaperonin, and/or foldase is from a Basidiomycete and the protein is a basidiomycetes protein. In some embodiments, the chaperone, chaperonin, and/or foldase are used in combination. In some embodiments, fragments of chaperone, chaperonin, and/or foldase having substantially the same function as the full-length chaperone, chaperonin, and/or foldase can be used. Exemplary chaperone, chaperonin, and/or foldase include those disclosed in U.S. patent application 60/919,332 and WO 2008/115596, which are incorporated herein by reference in their entirety. Exemplary chaperone, chaperonin, and/or foldase include, but are not limited to: BIP1, CLX1, ERO1, LHS1, PRP3, PRP4, PRP1, TIG1, PDI1, PPI1, PPI2, SCJ1, ERV2, EDEM, and SIL1. Table 2 provides a number of the sequences for chaperone, chaperonin, and/or foldase usable in the invention.

Table 2: Exemplary Nucleic Acid and Polypeptide Sequences of Secretion-Enhancing Proteins		
Protein	Exemplary Nucleotide Acid Sequence	Exemplary Polypeptide Sequence
BIP1	SEQ ID NO: 15	SEQ ID NO: 30
CLX1	SEQ ID NO: 16	SEQ ID NO: 31
ERO1	SEQ ID NO: 17	SEQ ID NO: 32
LHS1	SEQ ID NO: 18	SEQ ID NO: 33
PRP3	SEQ ID NO: 19	SEQ ID NO: 34
PRP4	SEQ ID NO: 20	SEQ ID NO: 35
PRP1	SEQ ID NO: 21	SEQ ID NO: 36
TIG1	SEQ ID NO: 22	SEQ ID NO: 37
PDI1	SEQ ID NO: 23	SEQ ID NO: 38
PPI1	SEQ ID NO: 24	SEQ ID NO: 39
PPI2	SEQ ID NO: 25	SEQ ID NO: 40
SCJ1	SEQ ID NO: 26	SEQ ID NO: 41
ERV2	SEQ ID NO: 27	SEQ ID NO: 42
EDEM	SEQ ID NO: 28	SEQ ID NO: 43
SIL1	SEQ ID NO: 29	SEQ ID NO: 44

Molecular Biology – Promoters and Expression Vectors

The present invention utilizes routine techniques in the field of recombinant genetics, well-known to those of skill in the art. In some embodiments, the present invention provides heterologous genes comprising gene promoter sequences (*e.g.*, from, filamentous fungi) that are typically cloned into intermediate vectors before transformation into host cells (*e.g.*, *Trichoderma reesei* cells) for replication and/or expression. These intermediate vectors are typically prokaryotic vectors (*e.g.*, plasmids, or shuttle vectors).

In general, the expression of a desired protein is accomplished under any suitable promoter. In one embodiment, a promoter non-native to a host is operably linked to a polynucleotide encoding a desired protein that is either native or non-native to a host. In another embodiment, a promoter native to a host is operably linked to a polynucleotide encoding a desired protein that is either native or non-native to a host. In some embodiments,

the desired protein is expressed under a heterologous promoter, which is not naturally associated with the desired protein gene. While in some other embodiments, the desired protein is expressed under a constitutive or inducible promoter. In some embodiments, the desired protein is expressed in a *Trichoderma* expression system with a cellulase promoter (e.g., the *cbh1* promoter).

As used herein, the term "promoter" refers to a nucleic acid sequence that functions to direct transcription of a downstream gene. A promoter can include necessary nucleic acid sequences near the start site of transcription, such as, in the case of a polymerase II type promoter, a TATA element. The promoter together with other transcriptional and translational regulatory nucleic acid sequences, collectively referred to as "regulatory sequences" controls the expression of a gene. In general, the regulatory sequences include, but are not limited to, promoter sequences, ribosomal binding sites, transcriptional start and stop sequences, translational start and stop sequences, and enhancer or activator sequences. The regulatory sequences are generally appropriate for and recognized by the host in which the downstream gene is being expressed. In some embodiments, the promoter used is from the same phylum as the host cell, and in other embodiment the promoter is from the same genus as the host cell, and in some embodiments from the same genus and species as the host cell.

A "constitutive promoter" is a promoter that is active under most environmental and developmental conditions. An "inducible" or "repressible promoter" is a promoter that is active under environmental or developmental regulation. In some embodiments, promoters are inducible or repressible due to changes in environmental factors including, but not limited to, carbon, nitrogen or other nutrient availability, temperature, pH, osmolarity, the presence of heavy metal(s), the concentration of inhibitor(s), stress, or a combination of the foregoing, as is known in the art. In some other embodiments, promoters are inducible or repressible by metabolic factors, such as the level of certain carbon sources, the level of certain energy sources, the level of certain catabolites, or a combination of the foregoing, as is known in the art.

Suitable non-limiting examples of promoters include *cbh1*, *cbh2*, *egl1*, *egl2*, *egl3*, *egl4*, *egl5*, *xyn1*, and *xyn2*, repressible acid phosphatase gene (*phoA*) promoter of *P. chrysogenum* (See, Graessle *et al.*, Appl. Environ. Microbiol., 63:753-756 [1997]), glucose-repressible PCK1 promoter (See, Leuker *et al.*, Gene 192 :235-240 [1997]), maltose-inducible, glucose-repressible MRP1 promoter (See, Munro *et al.*, Mol. Microbiol., 39 1414-

1426 [2001]), methionine-repressible MET3 promoter (*See, Liu et al., Eukary. Cell* 5:638-649 [2006]), pKi promoter, and cpcl promoter.

In some embodiments of the present invention, the promoter in the reporter gene construct is a temperature-sensitive promoter. In some embodiments, the activity of the temperature-sensitive promoter is repressed by elevated temperature. In some embodiments, the promoter is a catabolite-repressed promoter. In some embodiments, the promoter is repressed by changes in osmolarity. In some embodiments, the promoter is inducible or repressible by the levels of polysaccharides, disaccharides, or monosaccharides present in the culture medium.

An example of an inducible promoter that finds use in the present invention is the cbh1 promoter of *T. reesei*, the nucleotide sequence of which is deposited in GenBank under Accession Number D86235. Other exemplary promoters include promoters involved in the regulation of genes encoding cellulase enzymes, including, but not limited to, cbh2, egl1, egl2, egl3, egl5, xyn1 and xyn2.

In some embodiments of the present invention, in order to obtain high levels of expression of a cloned gene, the heterologous gene is advantageously positioned about the same distance from the promoter as in the naturally occurring gene. However, as is known in the art, some variation in this distance can be accommodated without loss of promoter function.

In some embodiments, a natural promoter modified by replacement, substitution, addition or elimination of one or more nucleotides finds use in the present invention, as long as the modifications do not change the function of the promoter. Indeed, it is intended that the present invention encompasses and is not constrained by such alterations to the promoter.

The expression vector/construct typically contains a transcription unit or expression cassette that contains all of the additional elements required for the expression of the heterologous sequence. Thus, a typical expression cassette contains a promoter operably linked to the heterologous nucleic acid sequence and signals required for efficient polyadenylation of the transcript, ribosome binding sites, and translation termination. Additional elements within the cassette may include enhancers and, if genomic DNA is used as the structural gene, introns with functional splice donor and acceptor sites, secretion leader peptides, leader sequences, linkers, and cleavage sites.

The practice of the present invention is not constrained by the choice of promoter in the genetic construct. As indicated above, exemplary promoters are the *Trichoderma reesei* cbh1, cbh2, eg1, eg2, eg3, eg5, xln1 and xln2 promoters. Additional promoters that find use

in the present invention include those from *A. awamori* and *A. niger* glucoamylase genes (*glaA*) (See, Nunberg *et al.*, Mol. Cell Biol., 4:2306–2315 [1984]) and the promoter from *A. nidulans* acetamidase. An exemplary promoter for vectors used in *Bacillus subtilis* is the *AprE* promoter; an exemplary promoter used in *E. coli* is the *Lac* promoter, an exemplary promoter used in *Saccharomyces cerevisiae* is PGK1, an exemplary promoter used in *Aspergillus niger* is *glaA*, and an exemplary promoter for *Trichoderma reesei* is *cbhI*. However, it is not intended that the present invention be limited to these specific cells nor these specific promoters, as other cells and promoters find use in various embodiments.

In some embodiments, in addition to a promoter sequence, the expression cassette also contains a transcription termination region downstream of the structural gene to provide for efficient termination. In some embodiments, the termination region is obtained from the same gene as the promoter sequence, while in other embodiments, it is obtained from different genes.

Although any suitable functional fungal terminator finds use in the present invention, some exemplary terminators include, but are not limited to the terminator from *Aspergillus nidulans* *trpC* gene (See, Yelton *et al.*, Proc. Natl. Acad. Sci. USA 81:1470-1474 (1984); Mullaney *et al.*, (Molecular Genetics and Genomics [MGG] 199:37-45 (1985)), the *Aspergillus awamori* or *Aspergillus niger* glucoamylase genes (See, Nunberg *et al.*, Mol. Cell Biol., 4:2306 (1984); Boel *et al.*, EMBO J., 3:1581-1585 (1984)), the *Aspergillus oryzae* TAKA amylase gene, the *Mucor miehei* carboxyl protease gene (EP Pat. Publ. No. 0 215 594) and the *Trichoderma reesei* CBH1 gene.

It is not intended that the expression vector used to transport the genetic information into the host cell be limited to any particular vector. It is contemplated that any of the conventional vectors used for expression in eukaryotic or prokaryotic cells will find use in the present invention. Standard bacterial expression vectors include, but are not limited to bacteriophages λ and M13, as well as plasmids such as pBR322-based plasmids, pSKF, pET23D, and fusion expression systems such as MBP, GST, and LacZ. In some embodiments, epitope tags are added to recombinant proteins to provide convenient methods of isolation (*e.g.*, c-myc). Examples of suitable expression and/or integration vectors are well-known to those in the art (See *e.g.*, Bennett and Lasure (eds.) More Gene Manipulations in Fungi, Academic Press pp. 70–76 and pp. 396–428 (1991); US Pat. No. 5,874,276. Various commercial vendors (*e.g.*, Promega, Invitrogen, etc.) provide useful vectors, as known to those of skill in the art. Some specific useful vectors include, but are not limited to pBR322, pUC18, pUC100, pDON™201, pENTR™, pGEN®3Z and pGEN®4Z. However, it is

intended that the present invention encompass other expression vectors which serve equivalent functions and which are, or become, known in the art. Thus, a wide variety of host/expression vector combinations find use in expressing the DNA sequences of the present invention. In some embodiments, useful expression vectors comprise segments of chromosomal, non-chromosomal and/or synthetic DNA sequences (e.g., various known derivatives of SV40) and known bacterial plasmids (e.g., plasmids from *E. coli* including col E1, pCR1, pBR322, pMb9, pUC 19, pSL1180 and their derivatives), wider host range plasmids (e.g., RP4), phage DNAs (e.g., the numerous derivatives of phage lambda, such as NM989, and other DNA phages, such as M13, and filamentous single stranded DNA phages), and yeast plasmids (e.g., the 2.μ plasmid or derivatives thereof).

In some embodiments, an expression vector includes a selectable marker. Examples of selectable markers include those that confer antimicrobial resistance. Nutritional markers also find use in the present invention, including those markers known in the art as *amdS*, *argB* and *pyr4*. Markers useful for the transformation of *Trichoderma* are known in the art (See e.g., Finkelstein, in *Biotechnology of Filamentous Fungi*, Finkelstein *et al.*, (eds.), Butterworth-Heinemann, Boston MA, chapter 6 (1992)). In some embodiments, the expression vectors also include a replicon, a gene encoding antibiotic resistance to permit selection of bacteria that harbor recombinant plasmids, and/or unique restriction sites in nonessential regions of the plasmid to allow insertion of heterologous sequences. It is intended that any suitable antibiotic resistance gene will find use in the present invention. In some embodiments in which *T. reesei* is the host cell, the prokaryotic sequences are preferably chosen such that they do not interfere with the replication or integration of the DNA in *T. reesei*.

In some embodiments, an expression vector includes a reporter gene alone or, optionally as a fusion with the protein of interest. Examples of reporter genes include but are not limited to, fluorescent reporters, color detectable reporters (e.g., β-galactosidase), and biotinylated reporters. In some embodiments, when the reporter molecule is expressed, it is used to identify whether the signal peptide is active in a host cell. If the signal peptide is active, the reporter molecule is secreted from the cell. In some embodiments, the signal peptide is initially operably linked to the reporter, in order to identify secretion from a particular host cell. Alternative methods such as those using antibodies specific to the protein of interest and/or the signal peptide also find use in determining whether or not the protein of interest is secreted.

In some embodiments, the methods of transformation of the present invention result in the stable integration of all or part of the transformation vector into the genome of a host cell,

such as a filamentous fungal host cell. However, transformation resulting in the maintenance of a self-replicating extra-chromosomal transformation vector is also contemplated.

Many standard transfection methods find use in the present invention to produce bacterial and filamentous fungal (*e.g.*, *Aspergillus* or *Trichoderma*) cell lines that express large quantities of the proteins. Methods for the introduction of DNA constructs into cellulase-producing strains of *Trichoderma* are well-known to those of skill in the art (*See e.g.*, Lorito *et al.*, *Curr. Genet.*, 24:349-356 [1993]; Goldman *et al.*, *Curr. Genet.*, 17:169-174 [1990]; Penttilä *et al.*, *Gene* 6: 155-164 [1987]; US Pat. No. 6,022,725; US Pat. No. 6,268,328; Nevalainen *et al.*, "The Molecular Biology of *Trichoderma* and its Application to the Expression of Both Homologous and Heterologous Genes" in *Molecular Industrial Mycology*, Leong and Berka (eds.), Marcel Dekker Inc., NY [1992] pp 129-148; Yelton *et al.*, *Proc. Natl. Acad. Sci. USA* 81: 1470-1474 [1984]; Bajar *et al.*, *Proc. Natl. Acad. Sci. USA* 88: 8202-8212 [1991]; Fernandez-Abalos *et al.*, *Microbiol.*, 149:1623 – 1632 [2003]; and Brigidi *et al.*, *FEMS Microbiol. Lett.*, 55:135-138 [1990]).

However, any of the well-known procedures for introducing foreign nucleotide sequences into host cells find use in the present invention. These methods include, but are not limited to the use of calcium phosphate transfection, polybrene, protoplast fusion, electroporation, biolistics, liposomes, microinjection, plasmid vectors, viral vectors and any of the other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA or other foreign genetic material into a host cell, as well-known to those of skill in the art. Also of use is the *Agrobacterium*-mediated transfection method (*See e.g.*, U.S. Patent No. 6,255,115). It is only necessary that the particular genetic engineering procedure used be capable of successfully introducing at least one gene into a host cell that is capable of expressing the gene. In some embodiments, the invention provides methods for producing a protein, comprising the steps of introducing into a host cell a polynucleotide comprising an NSP24 signal peptide linked to a nucleic acid encoding a protein, culturing the host cell under suitable culture conditions for the expression and production of the protein, and producing said protein. In some embodiments, the protein is secreted from the host cell. In some alternative embodiments, the present invention provides methods for producing a protein, comprising the steps of introducing into a host cell a polynucleotide comprising an CBH1 signal peptide operably linked to a nucleic acid encoding a protein, culturing the host cell under suitable culture conditions for the expression and production of the protein, and producing said protein. In some embodiments, the protein is secreted from the host cell.

After the expression vector is introduced into the host cells, the transfected or transformed cells are cultured under conditions favoring expression of genes under control of the gene promoter sequences. In some embodiments, large batches of transformed cells are cultured. In some embodiments, the product (*i.e.*, the protein) is harvested from the cells and/or recovered from the culture using standard techniques.

Thus, the invention herein provides for the expression and enhanced secretion of desired polypeptides whose secretion is enhanced by signal peptide sequences, fusion DNA sequences, and various heterologous constructs as well as expression of chaperones, chaperonins and/or foldases. The invention also provides processes for expressing and secreting high levels of such desired polypeptides.

Desired Proteins

The term "desired protein" means any protein of interest. The desired protein can be a protein native to a host cell, or non-native (heterologous) to a host cell. In some embodiments, the desired protein is a fungal protein. In some embodiments, the host is an Ascomycete host and the protein is any protein other than an Ascomycetes protein. In some embodiments, the host is a Basidiomycete host and the protein is any protein other than a Basidiomycete protein. In some embodiments, the protein is any protein other than a *Trichoderma* protein. In some other embodiments, the protein is any protein other than an *Aspergillus* protein.

It is not intended that the present invention be limited to any particular type of protein. Indeed, it is intended that the present invention encompass any protein of interest. Some non-limiting examples of desired proteins include, but are not limited to glucoamylases, alpha amylases, granular starch hydrolyzing enzymes, cellulases, lipases, xylanases, cutinases, hemicellulases, proteases, oxidases, laccases and combinations thereof.

In some embodiments, the glucoamylase is a wild type glucoamylase obtained from a filamentous fungal source, such as a strain of *Aspergillus*, *Trichoderma* or *Rhizopus*. However, in other embodiments, the glucoamylase is a protein engineered glucoamylase (*e.g.*, a variant of an *Aspergillus niger* glucoamylase). In some other embodiments, compositions of the present invention also comprise at least one protease and at least one alpha amylase. In some embodiments, the alpha amylase is obtained from a bacterial source (*e.g.*, *Bacillus spp.*), or from a fungal source (*e.g.*, an *Aspergillus spp.*). In some embodiments, the compositions also include at least one protease, and/or at least one glucoamylase, and/or at least one alpha amylase enzymes. In some embodiments, the protein is laccase, such as

laccase obtained from Basidiomycetes, and in some embodiments, from the genus *Cerrena*, such as *C. unicolor*. Commercial sources of these enzymes are known and available from, for example Genencor International, Inc. and Novozymes A/S.

Laccase and Laccase Related Enzymes

In one preferred embodiment, laccases and laccase-related enzymes are desired proteins. It is not intended that the present invention be limited to any particular laccase, as any laccase enzyme within the enzyme classification (EC 1.10.3.2) is encompassed. In some embodiments, the laccase enzymes are obtained from microbial or plant origin. In some embodiments, the microbial laccase enzymes are derived from bacteria or fungi (including filamentous fungi and yeasts). Although it is not intended that the present invention be limited to specific laccases, suitable examples include laccases derivable from *Aspergillus*, *Neurospora* (e.g. *N. crassa*), *Podospora*, *Botrytis*, *Collybia*, *Cerrena*, *Stachybotrys*, *Panus*, (e.g., *Panus rudis*), *Thieilava*, *Fomes*, *Lentinus*, *Pleurotus*, *Trametes* (e.g., *T. villosa* and *T. versicolor*), *Rhizoctonia* (e.g. *R. solani*), *Coprinus* (e.g. *C. plicatilis* and *C. cinereus*), *Psatyrella*, *Myceliophthora* (e.g., *M. thermonhila*), *Schytalidium*, *Phlebia* (e.g. *P. radita*; See e.g., WO 92/01046), *Coriolus* (e.g. *C. hirsutus*; See e.g., JP 2--238885), *Spongipellis*, *Polyporus*, *Ceriporiopsis subvermisporea*, *Ganoderma tsunodae* and *Trichoderma*.

In some embodiments, laccases include *Cerrena* laccase A1, B1 and D2 from CBS115.075 strain, *Cerrena* laccase A2, B2, C, D1, and E from CBS154.29 strain, *Cerrena* laccase B3 enzyme from ATCC20013 strain (see e.g., US Publication No. 2008/0196173, incorporated herein by reference in its entirety). Further optimized versions of these laccases also find use in the present invention.

In another embodiment, laccases include the mature protein of *Cerrena* laccase D expressed in *Trichoderma*; the amino acid sequence of which is shown as follows (SEQ ID NO: 45).

AIGPVADLHIVNKDLAPDGVQRPTVLGGTFPGTLITGQKGNFQLNVIDDLTDDRM
LTPTSIHWHGFFQKGTAWADGPAFVTQCPIADNSFLYDFDVPDQAGTFWYHSHLST
QYCDGLRGAFVVYDPNDPHKDLYDVDDGGTVITLADWYHVLAQTVVGAATPDSTLI
NGLGRSQTGPADAELAVISVEHNKRYRFRLVSISCDPNFTFSVDGHNMTVIEVDGVN
TRPLTVDSIQIFAGQRYSFVLNANQPEDNYWIRAMPNIGRNTTTLDGKNAAILRYKNA
SVEEPKTVGGPAQSPLNEADLRPLVPAPVPGNAVPGGADINHRLNLTFSNGLFSINNA
SFTNPSVPALLQILSGAQNAQDLLPTGSYIGLELGKVVELVIPPLAVGGPHPFHLHGHN

FWVVRSGSDEYNFDDAILRDVVSIGAGTDEVTIRFVTDNPGPWFLHCHIDWHLEAG
LAIVFAEGINQTAAANPTPQAWDELCPKYNGLSASQKVKKPGGTAI

Host Cells

The present invention provides host cells transformed with DNA constructs and vector as described herein. In some embodiments, the present invention provides for host cells transformed with DNA constructs encoding a desired protein and operably linked to the NSP24 or CBHI signal peptide as described herein. In some embodiments, the invention provides DNA constructs that encode at least one desired protein such as protease, laccase, alpha amylase, glucoamylase, xylanase, and cellulose, wherein the constructs are introduced into a host cell. In some embodiments, the present invention provides for the expression of protein genes and/or overexpression of protein genes under control of gene promoters functional in bacterial and/or fungal host cells.

It is intended that any suitable host cell are useful with the present invention. It is not intended that the present invention be limited to any particular host cell. In some embodiments, the host cell is a cell in which the signal peptide has activity in secreting the protein of interest. For example, host cells for which a *T. reesei* signal peptide find use include, but are not limited to, fungal and bacterial cells. Host cells include filamentous fungal cells, including but not limited to *Trichoderma* spp. (e.g., *T. viride* and *T. reesei*, the asexual morph of *Hypocrea jecorina*, previously classified as *T. longibrachiatum*), *Penicillium* spp., *Humicola* spp. (e.g., *H. insolens* and *H. grisea*), *Aspergillus* spp. (e.g., *A. niger*, *A. nidulans*, *A. oryzae*, and *A. awamori*), *Fusarium* spp. (e.g., *F. gramineum*), *Neurospora* spp., *Hypocrea* spp. and *Mucor* spp. Alternative host cells include, but are not limited to *Bacillus* spp. (e.g., *B. subtilis*, *B. licheniformis*, *B. lentus*, *B. stearothermophilus* and *B. brevis*) and *Streptomyces* spp. (e.g., *S. coelicolor* and *S. lividans*).

Many methods are known in the art for identifying whether a protein is secreted in a host cell or remains in the cytoplasm. It is intended that any suitable method will find use in identifying host cells in which the signal sequence is active.

Protein Expression

Desired proteins of the present invention are produced by culturing cells transformed with a vector such as an expression vector containing genes whose secretion is enhanced by the NSP24 or CBH1 signal peptide sequence, foldases, chaperonins, and/or chaperones. The present invention is particularly useful for enhancing the intracellular and/or extracellular

production of proteins. As those of skill in the art know, optimal conditions for the production of the proteins will vary with the choice of the host cell and protein to be expressed. Such conditions are easily determined by those of skill in the art.

In some embodiments, the protein of interest is isolated or recovered and purified after expression. Various methods for protein isolation and purification are known to those of skill in the art. Any suitable method finds use in the present invention. For example, standard purification methods that find use in the present invention include, but are not limited to electrophoretic, molecular, immunological and chromatographic techniques, including ion exchange, hydrophobic, affinity, and reverse-phase HPLC chromatography, and chromatofocusing. For example, in some embodiments, the protein of interest is purified using a standard antibody column comprising antibodies directed against the protein of interest. Ultrafiltration and diafiltration techniques, in conjunction with protein concentration, also find use in some embodiments. As known to those of skill in the art, the degree of purification necessary varies depending on the use of the protein of interest. Indeed, in some embodiments, no purification is necessary.

In some embodiments, proteins of interest produced by transformed host cells, as provided by the present invention, are recovered from the culture medium by conventional procedures known to those of skill in the art. These methods include, but are not limited to separating the host cells from the medium by centrifugation or filtration. In some embodiments, the cells are disrupted and the supernatant is removed from the cellular fraction and debris. In some embodiments, the proteinaceous components of the supernatant or filtrate are precipitated by means of a salt (*e.g.*, ammonium sulfate) after clarification. The precipitated proteins are then solubilized and in some embodiments, are purified by any suitable method, including chromatographic procedures (*e.g.*, ion exchange chromatography, gel filtration chromatography, affinity chromatography, and other art-recognized procedures).

In some further embodiments, antibodies directed against the peptides and proteins produced using the present invention are generated by immunizing an animal (*e.g.*, a rabbit or mouse), and recovering anti-protein and/or NSP24 signal peptide antibodies using any suitable method known in the art. In some additional embodiments, monoclonal antibodies are produced using any suitable method known in the art.

In some embodiments, assays known to those of skill in the art find use in the present invention, including, but not limited to those described in WO 99/34011 and U.S. Patent No. 6,605,458, both of which are incorporated by reference herein in their entirety.

Fusions

In some embodiments, the desired protein is produced as a fusion protein. In some further embodiments, the desired protein is fused to a protein that is efficiently secreted by a filamentous fungus, and fused to an enzyme catalytic domain from the same phylum, genus, and/or species as the host cell used for expression of the fusion protein. In some embodiments, the desired protein is fused to a CBHI polypeptide, or portion thereof. In some additional embodiments, the desired protein is fused to a CBHI polypeptide, or portion thereof, that is altered to minimize or eliminate catalytic activity. In some still further embodiments, the desired protein is fused to a *Trichoderma* glucoamylase polypeptide, or portion thereof. In some additional embodiments, the desired protein is fused to a *Trichoderma* glucoamylase, or portion thereof, that is altered to minimize or eliminate catalytic activity. In some further embodiments, the desired protein is fused to a polypeptide to enhance secretion, facilitate subsequent purification and/or enhance stability.

In general, the first, second, and/or third polynucleotide in the expression host of the present invention is either genetically inserted or integrated into the genomic makeup of the expression host (e.g., it is integrated into the chromosome of the expression host). However, in some embodiments, it is extrachromosomal (e.g., it exists as a replicating vector within the expression host). In some further embodiments, the extrachromosomal polynucleotide is expressed under suitable selection conditions for a selection marker that is present on the vector).

Secretion Level Assays

As described herein, the secretion level of a desired polypeptide in the expression host is determined using any suitable method. For example, in some embodiments, the secretion level is based on various factors (e.g., growth conditions of the host), etc. However, in some embodiments, the secretion level of the desired polypeptide expressed in the host is higher than the secretion level of the desired polypeptide expressed without the presence of a secretion enhancing protein. In some embodiments, the secretion level of a desired polypeptide (e.g., laccase from *Cerrena unicolor* in an expression host such as *T. reesei*) is at least about 1 mg/liter, about 2 mg/liter, about 3 mg/liter, about 4 mg/liter, or about 5 mg/liter when the host is grown in batch fermentation mode in a shake flask, or at least about 50 mg/liter, about 100 mg/liter, about 150 mg/liter, about 200 mg/liter, about 250 mg/liter, about 500 mg/liter, about 1000 mg/liter, about 2000 mg/liter, about 5000 mg/liter, about 10,000

mg/liter or about 20,000 mg/liter when the host is grown in a fermenter environment with controlled pH, feed-rate, etc. (e.g., fed-batch fermentation).

For example, in order to evaluate the expression and/or secretion of a secretable polypeptide, assays are carried out at the protein level, the RNA level, and/or through the use of functional bioassays suitable for the secretable polypeptide activity and/or production. Exemplary assays employed to analyze the expression and/or secretion of secretable polypeptide include but are not limited to, Northern blotting, dot blotting (DNA or RNA analysis), RT-PCR (reverse transcriptase polymerase chain reaction), or *in situ* hybridization, using an appropriately labeled probe (based on the nucleic acid coding sequence), conventional Southern blotting and autoradiography.

In some embodiments, the production, expression and/or secretion of a secretable polypeptide is directly measured in a sample. In some embodiments, the measurements are made using assays for enzyme activity, expression and/or production. In some embodiments, protein expression is evaluated by immunological methods (e.g., immunohistochemical staining of cells and/or tissue sections, or immunoassays of tissue culture medium by Western blotting or ELISA methods). Such immunoassays find use in qualitatively and/or quantitatively evaluating the expression of secretable polypeptide. These methods are known to those of skill in the art. Indeed, there are numerous commercially available kits and reagents for use in such methods.

In some embodiments, the present invention also provides extracts (e.g., solids or supernatants) obtained from the culture medium used to grow the expression host. In some embodiments, the supernatant does not contain substantial amount of the expression host, while in some alternative embodiments, the supernatant does not contain any amount of the expression host.

Cell Culture

As known in the art, the host cells and transformed cells of the present invention can be cultured in conventional nutrient media. However, in some embodiments, the culture media for transformed host cells is modified as appropriate, for activating promoters and selecting transformants. The specific culture conditions, such as temperature, pH and the like, are typically those that are used for the host cell selected for expression, and will be apparent to those skilled in the art. Culture media and conditions for host cells are known to those of skill in the art. It is noted that in culture, stable transformants of fungal host cells, such as *Trichoderma* cells are generally distinguishable from unstable transformants by their faster

growth rate or the formation of circular colonies with a smooth, rather than ragged outline on solid culture medium.

Compositions

In some embodiments, the present invention provides compositions and methods for expressing desired proteins using the NSP24 or CBH1 signal sequence, constructs and vectors. In some embodiments, the present invention provides compositions that include enzymes, including, but not limited to laccases, glucoamylases, alpha amylases, granular starch hydrolyzing enzymes, cellulases, lipases, phospholipases, xylanases, cutinases, hemicellulases, oxidases, peroxidases, proteases, phytases, keratinases, pullulanases, glucoamylases, pectinases, oxidoreductases, reductases, perhydrolases, phenol oxidases, lipoxigenases, ligninases, tannanases, pullulanases, pentosanases, beta-glucanases, arabinosidases, hyaluronidases, chondroitinases, mannanases, esterases, acyl transferases, and combinations thereof.

Applications

The desired proteins produced by the present invention find use in any applications appropriate for that protein. Examples of applications for proteins such as enzymes include, but are not limited to animal feeds for improvement of feed intake and feed efficiency (*e.g.*, proteases), dietary protein hydrolysates (*e.g.*, for individuals with impaired digestive systems), leather treatment, treatment of protein fibers (*e.g.*, wool and silk), cleaning, protein processing (*e.g.*, to remove bitter peptides, enhance the flavor of food, and/or to produce cheese and/or cocoa), personal care products (*e.g.*, hair compositions), sweeteners (*e.g.*, production of high maltose or high fructose syrups), fermentation and bioethanol (*e.g.*, alpha amylases and glucoamylases used to treat grains for fermentation to produce bioethanol). Examples of applications for laccases include, but are not limited to bleaching of pulp and paper, textile bleaching, treatment of waste water, de-inking of waste paper, polymerization of aromatic compounds or proteins, radical-mediated polymerization and cross-linking reactions (*e.g.*, paints, coatings, biomaterials), the activation of dyes, and to couple organic compounds. The laccases also find use in cleaning composition, including but not limited to laundry and other detergents.

EXAMPLES

The following examples are offered to illustrate, but not to limit the claimed invention. It will be apparent to those skilled in the art that many modifications, both to materials and methods, may be practiced without departing from the scope of the invention.

In the experimental disclosure which follows, the following abbreviations apply: M (Molar); μ M (micromolar); N (Normal); mol (moles); mmol (millimoles); μ mol (micromoles); nmol (nanomoles); g (grams); mg (milligrams); kg (kilograms); μ g and μ g (micrograms); L (liters); ml (milliliters); μ l and μ l (microliters); cm (centimeters); mm (millimeters); μ m (micrometers); nm (nanometers); $^{\circ}$ C (degrees Centigrade); h and hr (hours); min (minutes); sec (seconds); msec (milliseconds); V (voltage); xg (times gravity); $^{\circ}$ F (degrees Fahrenheit); amdS (acetamidase, a selective marker obtained from *A. nidulans*); lccD (laccase); BioRad (BioRad Laboratories, Hercules, CA); Difco (Difco Laboratories, Detroit, MI); Calbiochem (Calbiochem brand owned by EMD Chemicals Inc., San Diego, CA); Sigma (Sigma Chemical Co., St. Louis, MO); Spectronic (Spectronic Devices, Ltd., Bedfordshire, UK); Advanced Kinetics (Advanced Kinetics and Technology Solutions, Switzerland); .

Most of the expression vectors in the examples were produced based on the pSL1180 plasmid backbone, the sequence of which is provided in the GENBANK® database, under the identifier U13865. The markers such as the amdS marker, chaperones or foldases, laccase (lccD), the signal sequences, TrGA fusions and terminators were added using the polylinker and/or PCR methods as known in the art.

The sites on the plasmids are identified as follows: cbh1 – cellobiohydrolase; Tcbh1 – the terminator from cbh1; TrGA – *Trichoderma* glucoamylase; lccD – laccase D; amdS – marker selectable marker for autotrophism; pSL1180 – the plasmid backbone; laccase D opt – an optimized version of the laccase D gene that is constructed with codon usage optimized for expression in the host (*Trichoderma*); Pcpc-1 – a promoter from the cross pathway control-1 gene from *Neurospora crassa*; bla – β -lactamase gene (*i.e.*, a selective marker from *E. coli*); and HphR – the hygromycin-resistance gene (a selective marker from *E. coli*).

To construct the expression plasmids, primers were designed and used in the Herculase PCR reaction (Stratagene) containing the DNA template.

EXAMPLE 1. Construction of Expression Vector pTrex4-laccaseD opt

This Example describes the steps involved in the construction of the expression vector pTrex4-laccaseD opt. The plasmid was produced to express the codon optimized laccase D

gene from *C. unicolor* using the CBH1 promoter and CBH1 signal sequence. This expression vector contained the laccase D codon optimized gene fused to the CBH1 (cellobiohydrolase) core/linker and expressed from the CBH1 promoter. Figure 1 provides a schematic of the *Trichoderma* expression plasmid. The sequence of the pTrex4-laccaseD opt plasmid is shown as SEQ ID NO: 1. The following segments of DNA were assembled in the construction of pTrex4-laccase D opt (See, Figure 1). A fragment of *T. reesei* genomic DNA representing the CBH1 promoter and the CBH1 signal sequence and CBH1 core/linker was inserted into the plasmid pSL1180 vector. A codon optimized copy of the *C. unicolor* laccase D (*laccase D opt*) gene was inserted, such that it was operably linked to the CBH1 at its linker region. A CBH1 terminator from *T. reesei* was operably linked to the laccase D gene. The *amdS* gene was added as a selectable autotrophic marker. The *bla* gene (encoding beta-lactamase, a selective marker obtained from *E. coli*) is present in the pSL1180 vector.

EXAMPLE 2. Construction of Expression Vector pTrex2g-Bip1

The pTrex2g/Bip1 plasmid was produced to express the bip1 chaperone from *T. reesei*. Figure 2 provides the schematic of the *Trichoderma* expression plasmid pTrex2g-Bip1; The sequence of the plasmid is provided as SEQ ID NO: 2. The following segments of DNA were assembled in the construction of pTrex2g-Bip1. A 2267 bp fragment of *T. reesei* *bip1* was inserted into the plasmid pSL1180 vector operably linked to the *Ppki* promoter (pyruvate kinase from *T. reesei*), The *Trichoderma cbh1* terminator was operably linked to the *bip1* gene. The HphR selectable marker from *E. coli* was included for selection and was operably linked to the *Pcpc-1* promoter (cross pathway control-1 gene from *Neurospora crassa*) and the *trpC* terminator (tryptophan synthesis gene C from *A. nidulans*).

EXAMPLE 3. Construction of Expression Vector pTrex2g-Pdi1

The pTrex2g-Pdi1 plasmid was produced to express the chaperone *pdi1* in the same way as the pTrex2g-Bip1 (See, Example 2), except that the *T. reesei pdi1* chaperone gene (2465 bp) was inserted in place of the *bip1* chaperone gene. Figure 3 provides the schematic of the *Trichoderma* expression plasmid pTrex2g-Pdi1; the sequence of the plasmid is provided as SEQ ID NO: 3.

EXAMPLE 4. Construction of Expression Vector pTrex2g-Ero1

The pTrex2g-Ero1 plasmid was produced to express the chaperone *ero1* in the same way as the pTrex2g-Bip1 (See, Example 2), except that the *T. reesei ero1* chaperone gene

(2465 bp) was inserted in place of the *bip1* chaperone gene. Figure 4 provides the schematic of the *ero1* in the *Trichoderma* expression plasmid pTrex2g-Ero1. The sequence of *ero1* is provided as SEQ ID NO: 4.

EXAMPLE 5. Construction of Expression Vector pTrGA-laccaseD opt

The pTrGA-laccaseD opt plasmid was produced similarly to that in Example 1, except that pTrGA-laccase D opt expresses a fusion of the full-length glucoamylase from *T. reesei* and *C. unicolor* laccase D with optimized codons. Figure 5 provides the schematic of the *Trichoderma* expression plasmid pTrGA-laccaseD opt; the polynucleotide sequence is shown as SEQ ID NO: 5.

EXAMPLE 6. Construction of Expression Vector pKB408

The pKB408 plasmid was produced to express *C. unicolor* laccase D opt operably fused to the *T. reesei* NSP-24 signal peptide. The plasmid was constructed similarly to that shown in Figure 1 except that the laccase D constructs were operably linked to the NSP-24 signal peptide, which was inserted in place of the laccase D opt linked to the CBH1 signal sequence, catalytic domain and linker. Figure 6 provides the schematic of the *Trichoderma* expression plasmid pKB408; the polynucleotide sequence is shown as SEQ ID NO: 6.

EXAMPLE 7. Construction of Expression Vector pKB410.

The pKB410 plasmid was produced as described in Example 6, except the *T. reesei* CHB1 signal sequence was used instead of the NSP-24 signal sequence. Figure 7 provides the schematic of the *Trichoderma* expression plasmid pKB410; the polynucleotide sequence is shown as SEQ ID NO: 7.

EXAMPLE 8. Transformation of *T. reesei* and Analysis of Expression

In this example, the stable recombinant *T. reesei* strain derived from RL-P37 (See, Sheir-Neiss and Montenecourt, Appl. Microbiol. Biotechnol., 20:46-53 (1984)) and deleted for the *cbh1*, *cbh2*, *egl1*, and *egl2* genes described by Bower *et al* (See, Bower *et al.*, Carbohydrases From *Trichoderma reesei* and Other Micro-organisms, Royal Society of Chemistry, Cambridge, pp. 327-334 (1998)) was used for transforming the plasmids from Examples 1-14 alone or in various combinations. Biolistic and electroporation methods were used to transform the plasmids, as described below.

Biolistic Transformation

The expression plasmid was confirmed by DNA sequencing and transformed biolistically into a *Trichoderma* strain. Transformation of the *Trichoderma* strain by the biolistic transformation method was accomplished using a Biolistic® PDS-1000/he Particle Delivery System (Bio-Rad) following the manufacturer's instructions (See, WO 05/001036 and US Pat. Appl. Publ. No. 2006/0003408). Transformants were selected and transferred onto minimal media with acetamide (MMA) plates and grown for 4 days at 28-30°C. A small plug of a single colony including spores and mycelium was transferred into 30 mls of NREL lactose defined broth (pH 6.2) containing 1 mM copper. The cultures were grown for 5 days at 28°C. Culture broths were centrifuged and supernatants were analyzed using the ABTS assay as described below for laccase activity.

Electroporation

Electroporation was performed as described in US Patent application No: 60/931,072, herein incorporated by reference in its entirety. A *T. reesei* strain was grown and sporulated on Potato Dextrose Agar plates (Difco) for about 10-20 days. The spores were washed from the surface of the plates with water and purified by filtration through Miracloth (Calbiochem). The spores were collected by centrifugation (3000 xg, 12 min), washed once with ice-cold water and once with ice-cold 1.1M sorbitol. The spore pellet was re-suspended in a small volume of cold 1.1 M sorbitol, mixed with about 8 µg of gel-purified DNA fragment isolated from plasmid DNA (pKB408 and pKB410, Figures 6 and 7) per 100 µl of spore suspension. The mixture (100 µl) was placed into an electroporation cuvette (1mm gap) and subjected to an electric pulse using the following electroporation parameters: voltage 6000-20000 V/cm, capacitance = 25 µF, resistance = 50Ω. After electroporation, the spores were diluted about 100-fold into 5:1 mixture of 1.1 M sorbitol and YEPD (1% yeast extract, 2% Bacto-peptone, 2% glucose, pH 5.5), placed in shake flasks and incubated for 16-18 hours in an orbital shaker (28°C and 200 rpm). The spores were once again collected by centrifugation, re-suspended in about 10-fold of pellet volume of 1.1 M sorbitol and plated onto two 15 cm Petri plates containing *amdS* modified medium (acetamide 0.6 g/l, cesium chloride 1.68 g/l, glucose 20 g/l, potassium dihydrogen phosphate 15 g/l, magnesium sulfate heptahydrate 0.6 g/l, calcium chloride dihydrate 0.6 g/l, iron (II) sulfate 5mg/l, zinc sulfate 1.4 mg/l, cobalt (II) chloride 1mg/l, manganese (II) sulfate 1.6 mg/l, agar 20 g/l and pH 4.25). Transformants appeared at about 1 week of incubation at 28-30°C.

The ABTS assay was performed as follows: An ABTS stock solution was prepared containing 4.5 mM ABTS in water (ABTS; Sigma Cat# A-1888). Buffer was prepared containing 0.1 M sodium acetate pH 5.0. Then, 1.5 ml of buffer and 0.2 ml of ABTS stock solution were added to cuvettes (10x4x45mm, No./REF67.742) and mixed well. One extra cuvette was prepared as a blank. Then, 50 ul of each enzyme sample to be tested (using various dilutions) were added to the mixtures.

The ABTS activity was measured in a Genesys2 machine (Spectronic) using an ABTS kinetic assay program set up: (Advanced Kinetics) as follows: wave length 420nm, interval time (Sec) 2.0, total run time (sec) 14.0, factor 1.000, low limit -000000.00, high limit 999999.00, and the reaction order was first.

The procedure involved adding 1.5 mL of NaOAc (120mM NaOAc Buffer pH 5.0), then add 0.2 mL of 4.5mM ABTS to the cuvette, then to blank the cuvette, adding 0.05 mL of the enzyme sample to the cuvette, mixing quickly and well and, finally, measuring the change of absorption at 420nm, every 2 seconds for 14 seconds. One ABTS unit is defined as change of A420 per minute (given no dilution to the sample). Calculation of ABTS U/mL: (change in $\Delta 420/\text{min}$ * dilution factor).

EXAMPLE 9. Analysis of Laccase/Glucoamylase Fusion Gene Expression in *T. reesei* Transformants

The culture medium of the transformants obtained and cultivated as described in Example 8 was separated from mycelium by centrifugation (16000 xg, 10 min) and ABTS activity from the supernatants were analyzed. The results are shown in Figure 10. Table 3 provides the strains described in Figure 10. Figure 10 illustrates the improvement of laccase production by fusion of the gene encoding *C. unicolor* laccase to the full-length *Trichoderma* glucoamylase. The results showed that expression of laccase improved 24-29% when fused to the *Trichoderma* glucoamylase, than fused to CBH1.

Table 3. Strains Used in Figure 10

Strain Identification Number	Strain Type
#8-2	CBH1 laccase fusion
1066-9	TrGA laccase fusion
1066-13	TrGA laccase fusion
1066-15	TrGA laccase fusion

EXAMPLE 10. Analysis of Laccase Production Using NSP24 and CBH1 Signal Sequences

When the *T. reesei* CBH1 signal sequence was operably linked to the laccase gene, expression was improved 4-5 folds over initial CBH1 fusion strain #8-2 alone in shake flasks and 5-6 folds in a 14 liter fermentor as shown by the results provided in Figures 11 (shake flasks) and 12 (fermentor). When the *T. reesei* NSP-24 signal sequence was used, the expression improved 3-4 folds in shake flasks and 4-5 folds in a 14 liter fermentor. Three clones were analyzed in the shake flasks for the CBH1 signal sequence (#7, #10, and #13) and two clones were analyzed for the NSP24 signal sequence (#7 and #25) and the expression was analyzed at 3 days (first bar), 4 days (second bar) and 5 days (third bar). A single clone of each was analyzed in the 14 liter fermenters, as shown by the results in Figure 12. In this Figure, the diamond indicates the NSP24 signal sequence operably linked to the laccase D, the square indicates the CBH1 signal sequence operably linked to the laccase D and the triangle indicates the CBH1 fusion alone.

EXAMPLE 11. Analysis of Laccase Production Using CBH1 Signal Sequence and Co-Expression of *bip1* in a Fermenter

The CBH1 signal sequence plasmid (operably linked to laccase) was co-transformed with the *T. reesei* Bip1 plasmid and expression analyzed. The results are shown in Figure 13. In Figure 13, diamonds indicate the data obtained for the CHB1 signal sequence (operably linked to laccase) plus BIP1, while the squares indicate the data obtained for the CBH1 signal sequence (operably linked to laccase) alone. Figure 13 illustrates the improvement of laccase production provided by the CBH1 signal sequence plus BIP1 chaperone expression, which increased expression significantly, by more than 15% in fermentors.

EXAMPLE 12. Analysis of Laccase Production Using CBH1 Signal Sequence and Co-Expression of *bip1* in a Shake Flask

The CBH1 signal sequence plasmid (operably linked to laccase) was co-transformed with the *T. reesei* *bip1* plasmid, grown in and laccase expression analyzed using the ABTS assay. The results are presented in Figure 14. Five different clones were analyzed for 3 days (first bar) 4 days (second bar) and 5 days (third bar). KB410-13 was a control having CBH1 signal sequence plasmid alone. The other 4 clones were KB410-13 with one of the *bip1* co-transformants: E32, E9, E16, and E10. Figure 14 illustrates the improvement of laccase

production by co-expression of chaperones with *C. unicolor* in shake flasks. The co-expression with *bip1* increased expression significantly (from 14-41%) in shake flasks.

EXAMPLE 13. Analysis of Laccase Production Using CBH1-laccase D Fusion and Co-Expression of a Variety of Chaperones

The expression plasmid having a CBH1 signal sequence, catalytic domain and linker operably linked to laccase was co-transformed with a variety of *T. reesei* chaperone plasmids (*BIP1*, *PDI1*, and *ERO1*). The resultant transformed cell was grown in culture and laccase expression analyzed. Figure 15 illustrates the improvement of laccase production by fusion of the gene encoding *C. unicolor* laccase to the CBH1 signal sequence, catalytic domain and linker and co-expression with *bip1*, *pdi1* and *erol* chaperones.

All strains had CBH1 signal sequence, catalytic domain and linker linked to laccase D. Strains 1B1, 1B12 and 1B19 had *bip1* expression cassette; they were three independent transformants, with difference in the *bip1* plasmid copy numbers and location of integration. Strains 3B2 and 3B8 had *pdi1* expression cassette; they are two independent transformants, with difference in the *pdi1* plasmid copy numbers and location of integration. Strains 9B6 and 9B7 had *erol* expression cassette; they are two independent transformants, with difference in the *erol* plasmid copy numbers and location of integration may be different. #8-2 is the control strain which has no chaperone expression cassette.

The results of Figure 15 indicate that the highest increase in expression was obtained with the co-expression with the *bip1* chaperone.

EXAMPLE 14. Analysis of Laccase Production Using CBH1 Signal Sequence and Co-Expression of a Variety of Chaperones

The CBH1 signal sequence plasmid (*i.e.*, operably linked to laccase) was co-transformed with a variety of *T. reesei* chaperone plasmids (*bip1*, *lhs1*, *pdi1*, *ppi1*, *ppi2*, *tig1*, *prp1*, and *erol*), either alone or in combination. The cultures were grown in shake flasks as known in the art and laccase expression analyzed using the ABTS assay. The clones were analyzed in triplicate. The data provided in Table 4 show that adding more than one chaperone did not increase expression of laccase above that of *bip1* alone. The data in Table 4 show three independent spore-purified samples (or clones) from the same strain.

Table 4: Expression of Laccase in the Presence of Chaperones				
Co-transformation of KB413-32A with Different Chaperones				
Each Strain has 3 repeats: -A,-B,-C				
	Samples	Chaperones	4 days SF broth	6 days
1	KB413-32A-A	bip1 only	4.52	6.32
2	KB413-32A-B	bip1 only	4.26	6.35
3	KB413-32A-C	bip1 only	4.28	6.13
4	KB414-1-A	bip1, ero1	3.88	5.89
5	KB414-1-B	bip1, ero1	3.78	5.93
6	KB414-1-C	bip1, ero1	3.76	5.59
7	KB415-2-A	bip1, lhs1, white	3.8	5.93
8	KB415-2-B	bip1, lhs1, white	3.72	5.92
9	KB415-2-C	bip1, lhs1, white	3.78	6.06
10	KB415-3-A	bip1, lhs1, gray	4.38	6.32
11	KB415-3-B	bip1, lhs1, gray	4.3	6.66
12	KB415-3-C	bip1, lhs1, gray	3.98	6.15
13	KB416-3-A	bip1, pdi1	4.18	6.58
14	KB416-3-B	bip1, pdi1	5.26	7.12
15	KB416-3-C	bip1, pdi1	4.22	6.06
16	KB417-3-A	bip1, ppi1	4.32	6.23
17	KB417-3-B	bip1, ppi1	3.96	6.32
18	KB417-3-C	bip1, ppi1	4.18	6.88
19	KB418-2-A	bip1, ppi2	4.24	6.59
20	KB418-2-B	bip1, ppi2	3.96	5.69
21	KB418-2-C	bip1, ppi2	4.04	5.92
22	KB419-1-A	bip1, tigA	4.66	5.98
23	KB419-1-B	bip1, tigA	5.26	7.25
24	KB419-1-C	bip1, tigA	4.18	6.05
25	KB413-prp2-A	bip1, prpA	3.96	5.63
26	KB413-prp2-B	bip1, prpA	3.9	5.59
27	KB413-prp2-C	bip1, prpA	3.92	5.86
28	KB414-1-A	bip1, ero1	4.2	6.01

Table 4: Expression of Laccase in the Presence of Chaperones					
Co-transformation of KB413-32A with Different Chaperones					
Each Strain has 3 repeats: -A,-B,-C					
	Samples	Chaperones	4 days SF broth	6 days	
29	KB414-1-B	bip1, ero1	3.88	5.69	
30	KB414-1-C	bip1, ero1	3.92	5.88	

The invention, and the manner and process of making and using it, are now described in such full, clear, concise and exact terms as to enable any person skilled in the art to which it pertains, to make and use the same. It is to be understood that the foregoing describes preferred embodiments of the present invention and that modifications may be made therein without departing from the scope of the present invention as set forth in the claims. To particularly point out and distinctly claim the subject matter regarded as invention, the following claims conclude this specification.

WHAT IS CLAIMED IS:

1. A method for producing a desired protein, comprising the steps of:
 - (a) introducing into a host cell a first nucleic acid sequence comprising a signal sequence operably linked to a desired protein sequence;
 - (b) expressing the first nucleic acid sequence;
 - (c) co-expressing a second nucleic acid sequence encoding a chaperone or foldase selected from the group consisting of *bip1*, *ero1*, *pdi1*, *tig1*, *prp1*, *ppi1*, *ppi2*, *prp3*, *prp4*, calnexin, and *lhs1*; and
 - (d) collecting the desired protein secreted from the host cell.
2. The method according to Claim 1, wherein the first nucleic acid sequence further comprises an enzyme sequence between the signal sequence and the desired protein sequence.
3. The method according to Claim 2, wherein the enzyme sequence is obtained from a glucoamylase or from a CBH1 enzyme.
4. The method according to Claim 2, wherein the enzyme sequence comprises a full-length enzyme sequence.
5. The method according to Claim 2, wherein the enzyme sequence comprises a catalytic core domain sequence.
6. The method according to Claim 5, wherein the first nucleic acid sequence further comprises a linker sequence between the catalytic core domain sequence and the desired protein sequence.
7. The method according to Claim 1, wherein the desired protein is a laccase.
8. The method according to Claim 7, wherein said laccase is derived from a filamentous fungus or yeast.
9. The method according to Claim 8, wherein said laccase is derived from *Aspergillus*, *Neurospora*, *Podospora*, *Botrytis*, *Collybia*, *Cerrena*, *Stachybotrys*, *Panus*, *Thielavia*, *Fomes*, *Lentinus*, *Pleurotus*, *Trametes*, *Rhizoctonia*, *Coprinus*, *Psatyrella*, *Myceliophthora*,

Schytalidium, *Phlebia*, *Coriolus*, *Spongipellis*, *Polyporus*, *Ceriporiopsis subvermispora*, *Ganoderma tsunodae*, or *Trichoderma*.

10. The method according to Claim 9, wherein said laccase is derived from *Cerrena* laccase A1, A2, B1, B2, B3, C, D1, D2, or E.
11. The method according to Claim 9, wherein said laccase is derived from the mature protein of *Cerrena* laccase D.
12. The method according to Claim 1, wherein the signal sequence encodes Cellobiohydrolase I signal peptide or NSP24 signal peptide.
13. The method according to Claim 1, wherein the host is a filamentous fungus.
14. The method according to Claim 13, wherein the host is ascomycetes.
15. The method according to Claim 14, wherein the host is *Trichoderma*.
16. The method according to Claim 1, wherein the first nucleic acid sequence further comprises a promoter upstream to a signal sequence.
17. The method according to Claim 16, wherein the promoter is native to the host cell and is not naturally associated with the desired protein sequence.
18. The method according to Claim 1, wherein the chaperon is BIP1.
19. The method according to Claim 1, wherein the second nucleic acid sequence is operably linked to a promoter.
20. The method according to Claim 19, wherein the promoter is native to the host cell and is not naturally associated with the second nucleic acid sequence.
21. The method according to Claim 2, wherein the desired protein is a laccase and the laccase is produced as a fusion protein with the enzyme.

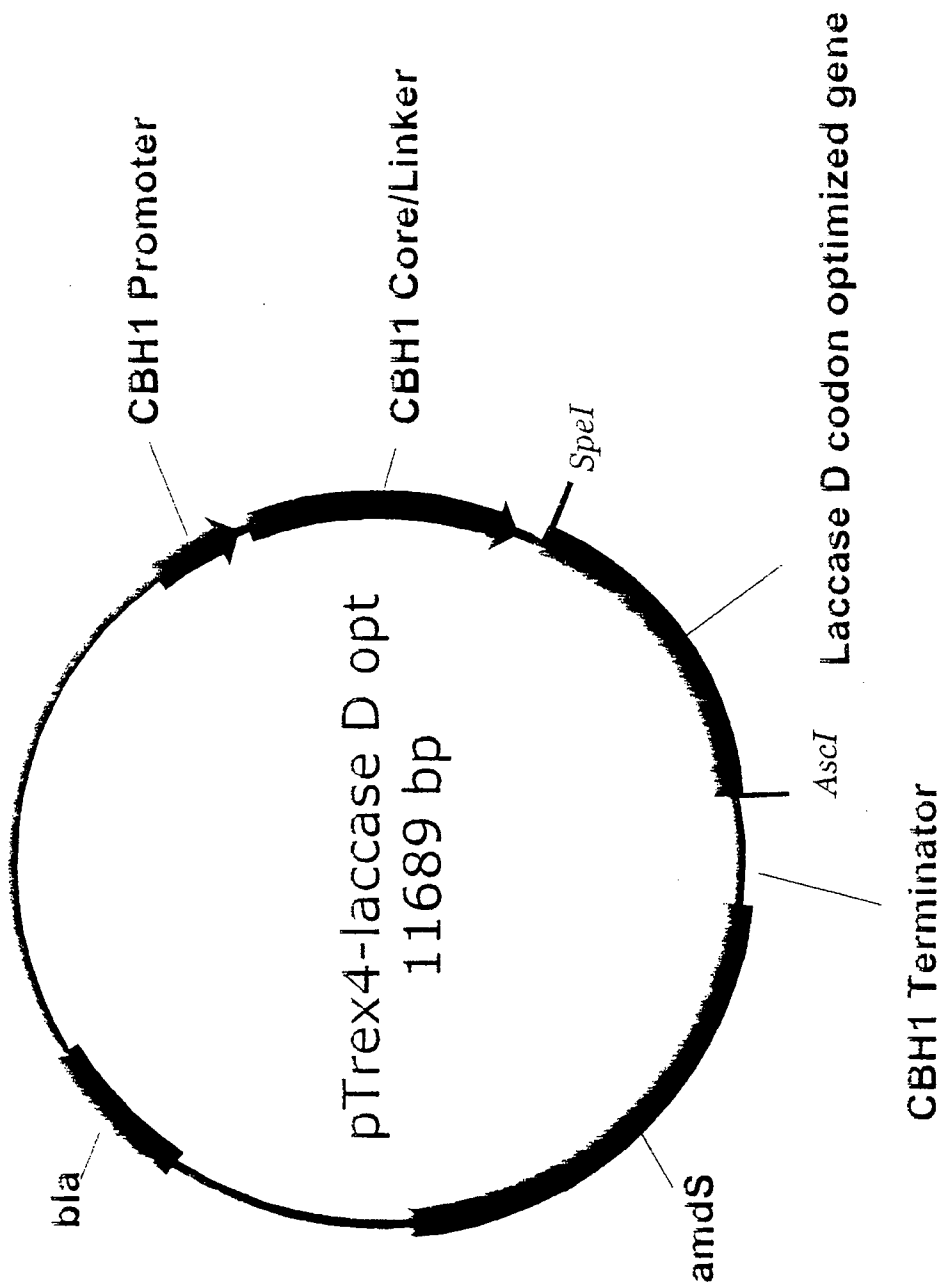


FIG. 1

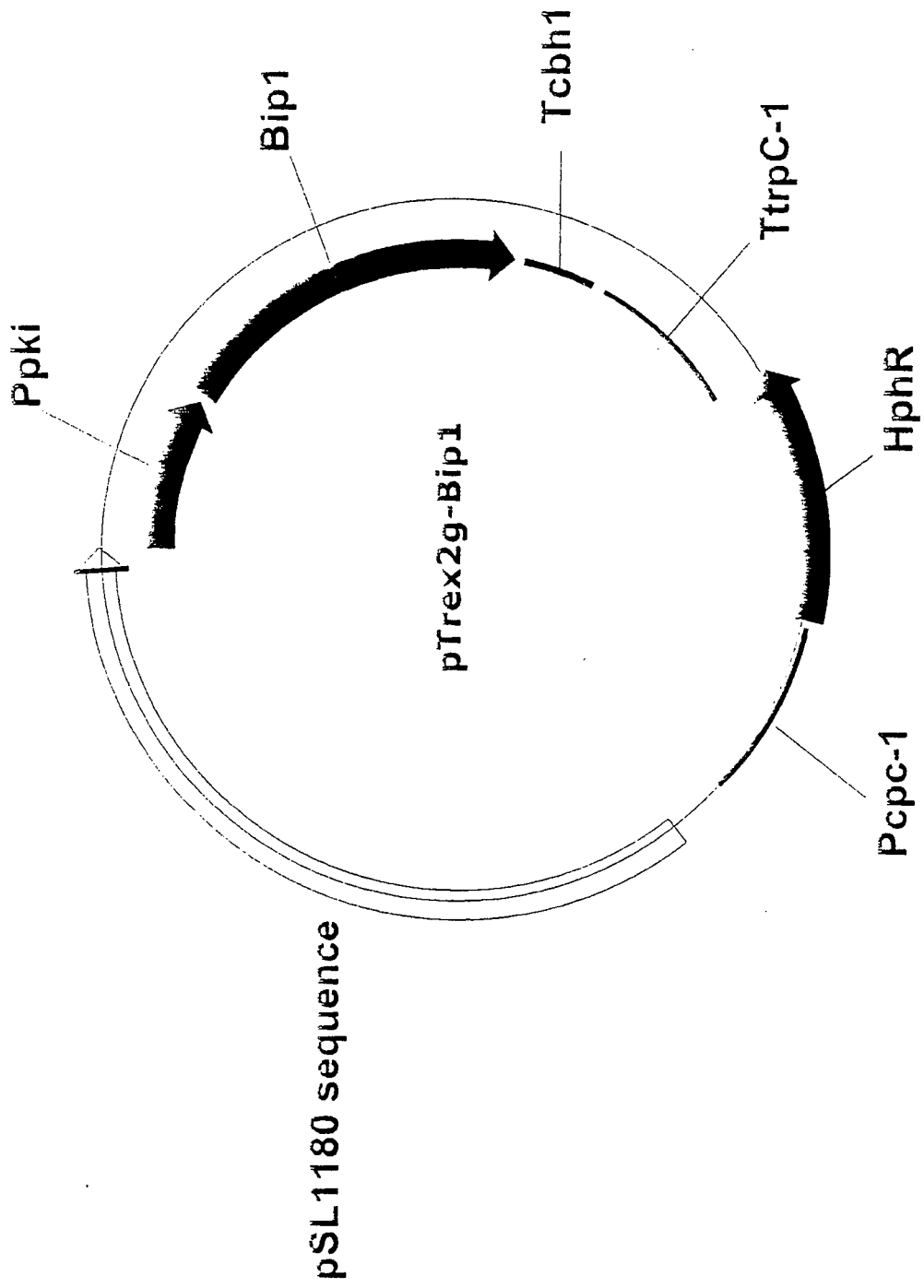


FIG. 2

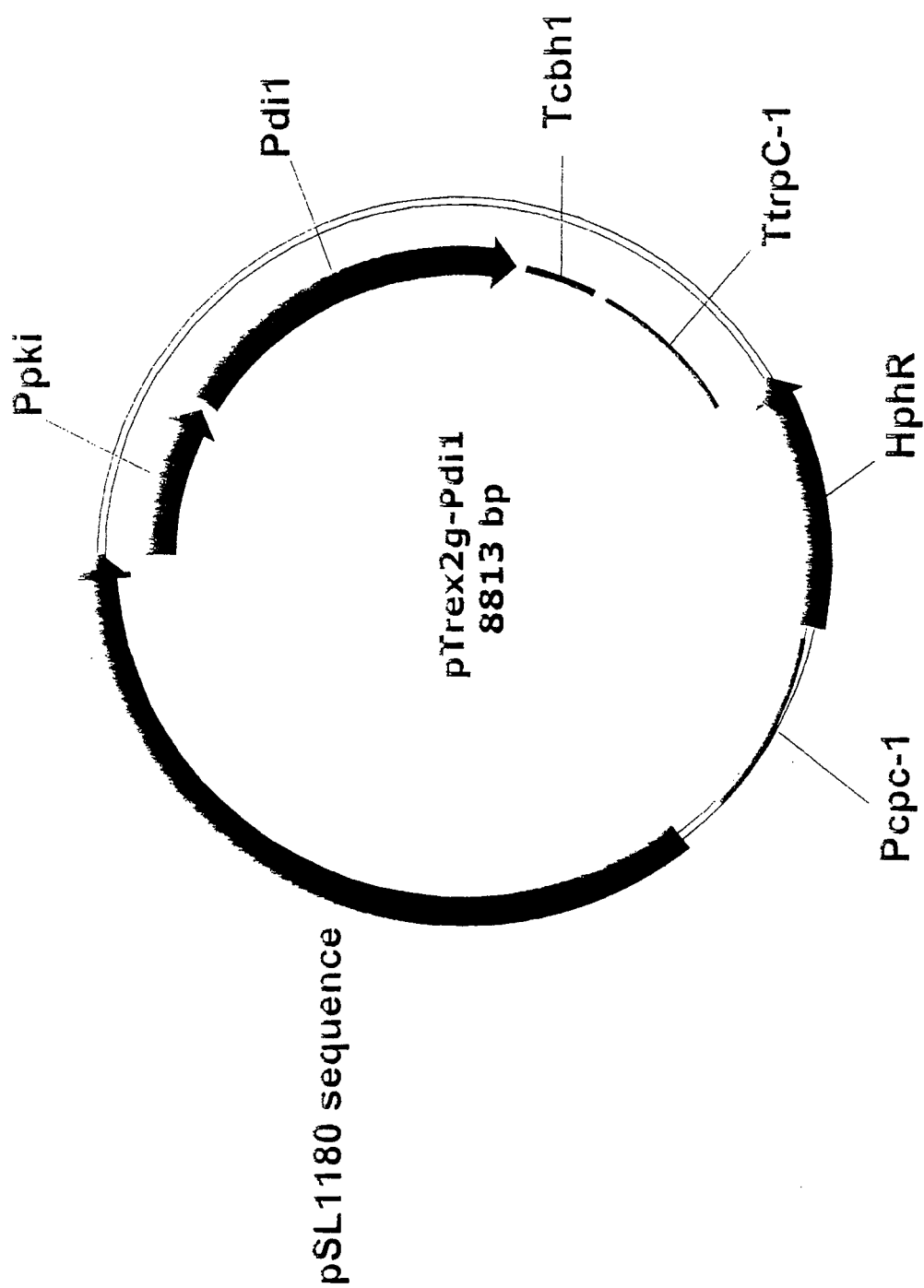


FIG. 3

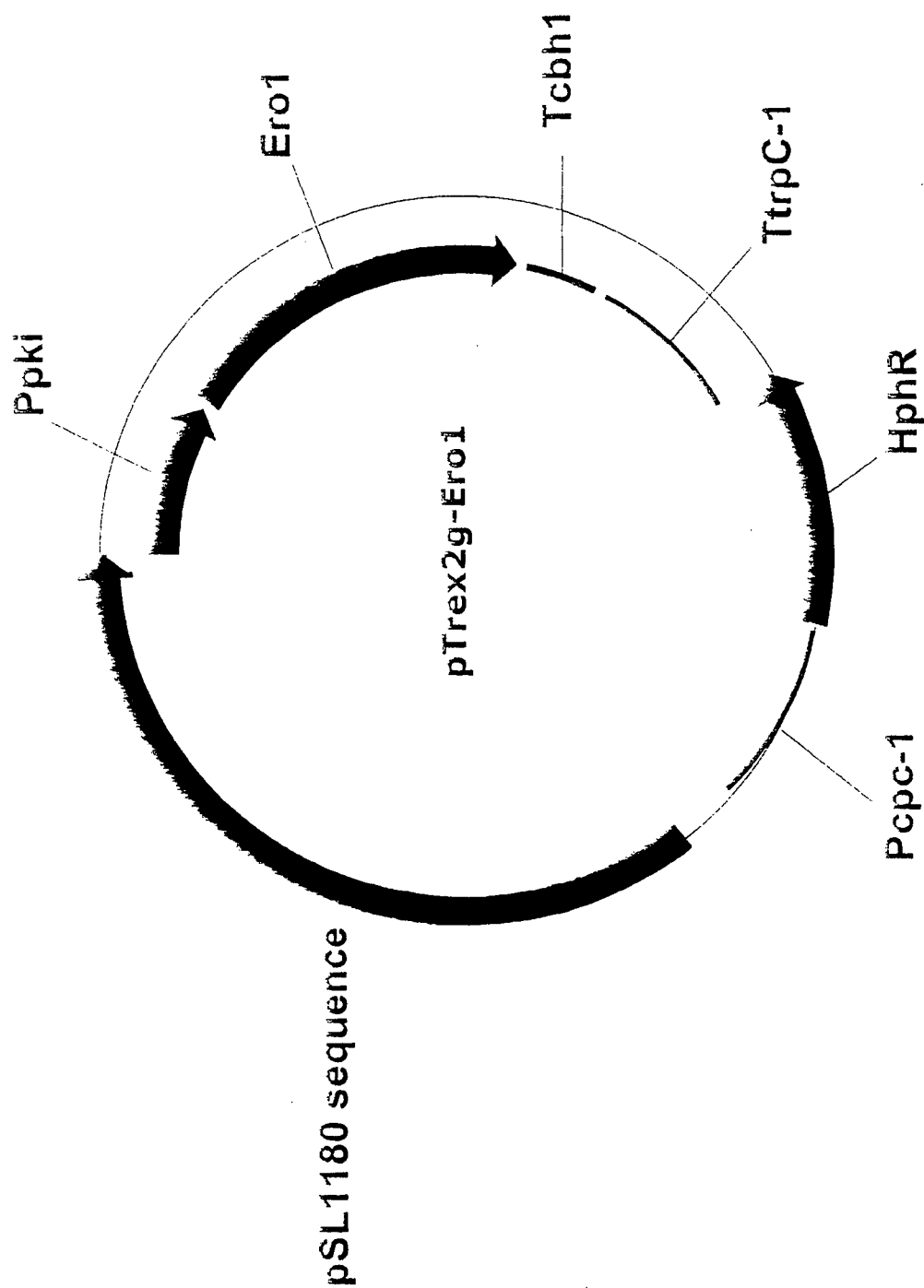


FIG. 4

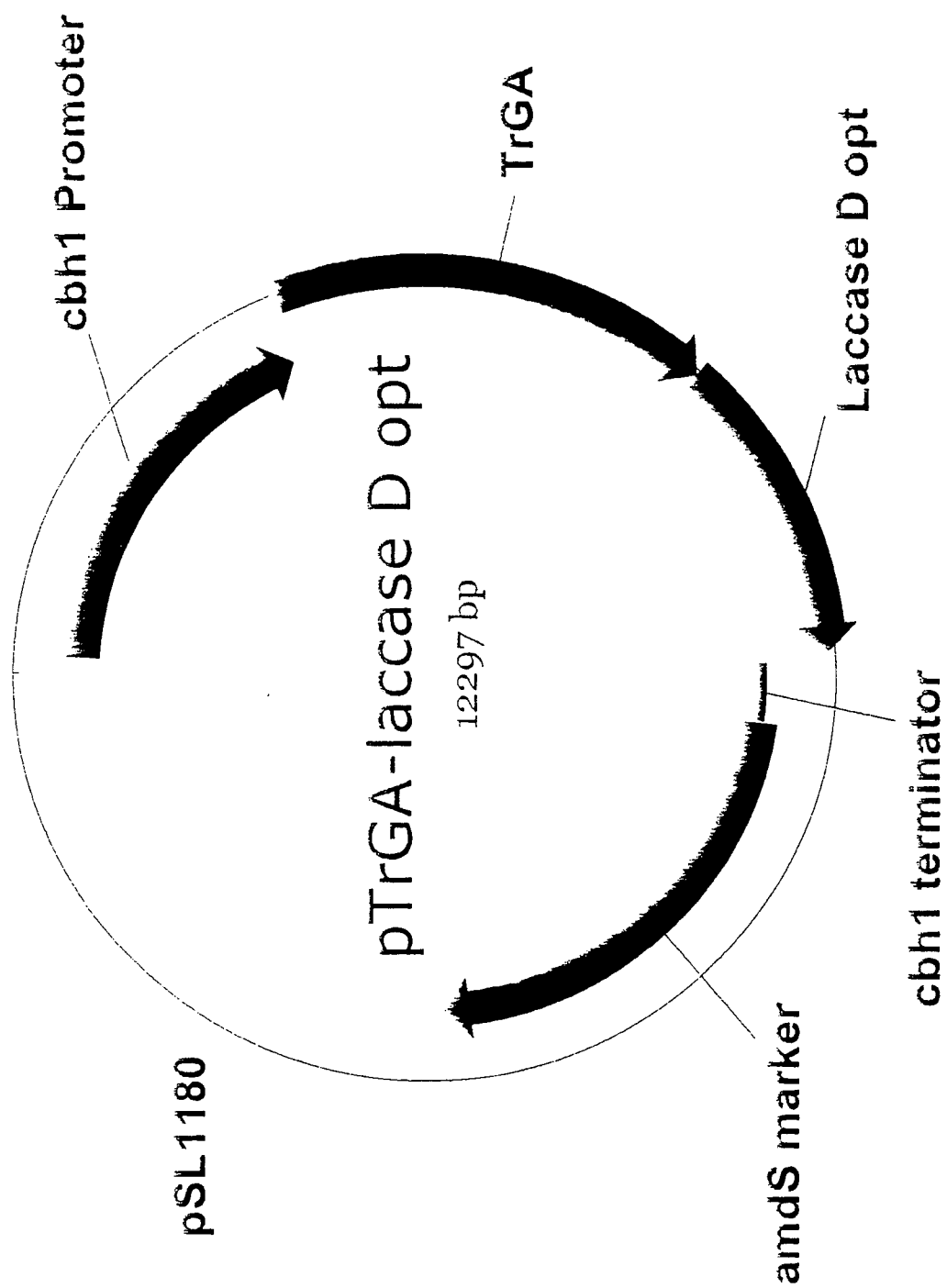


FIG. 5

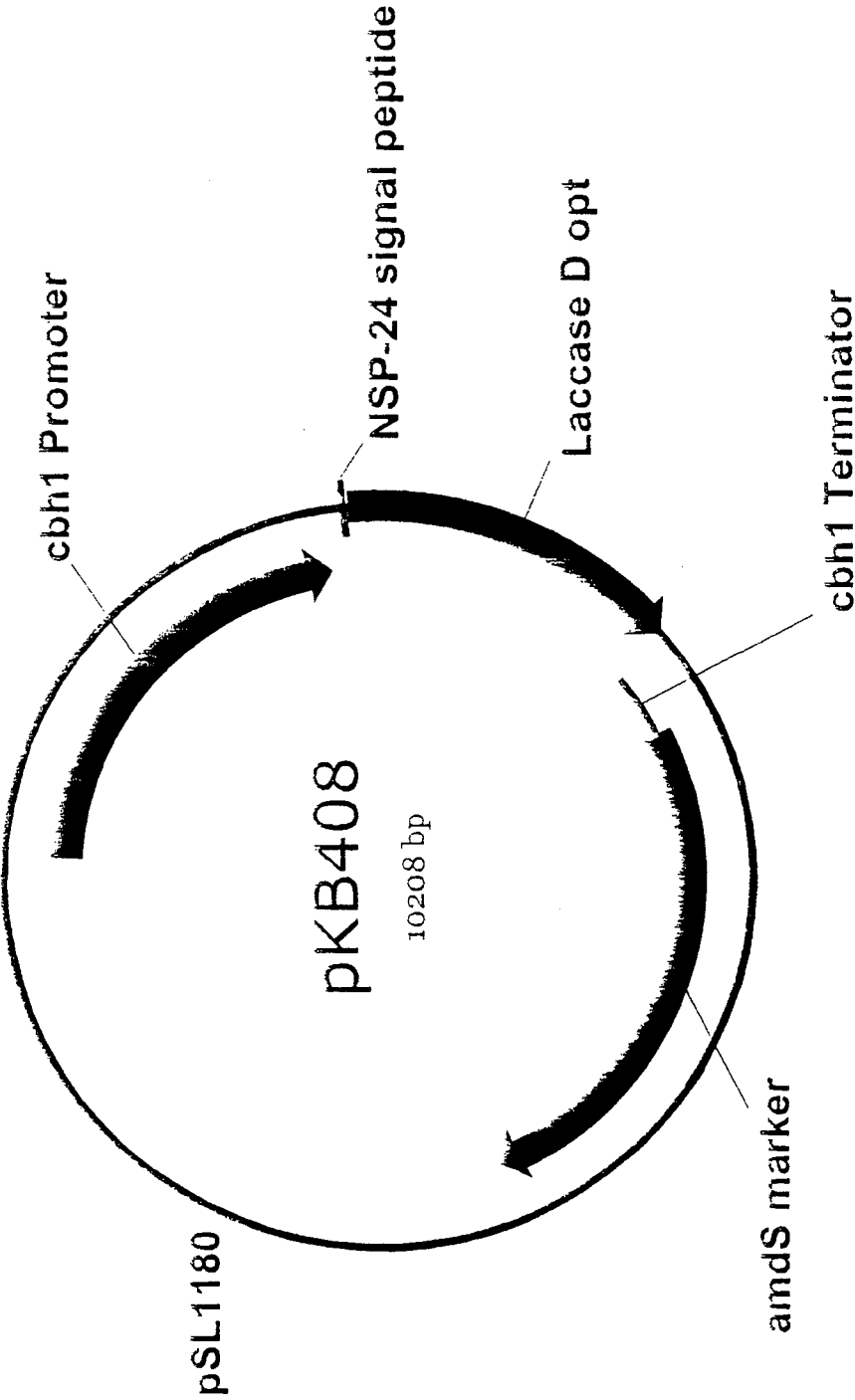


FIG. 6

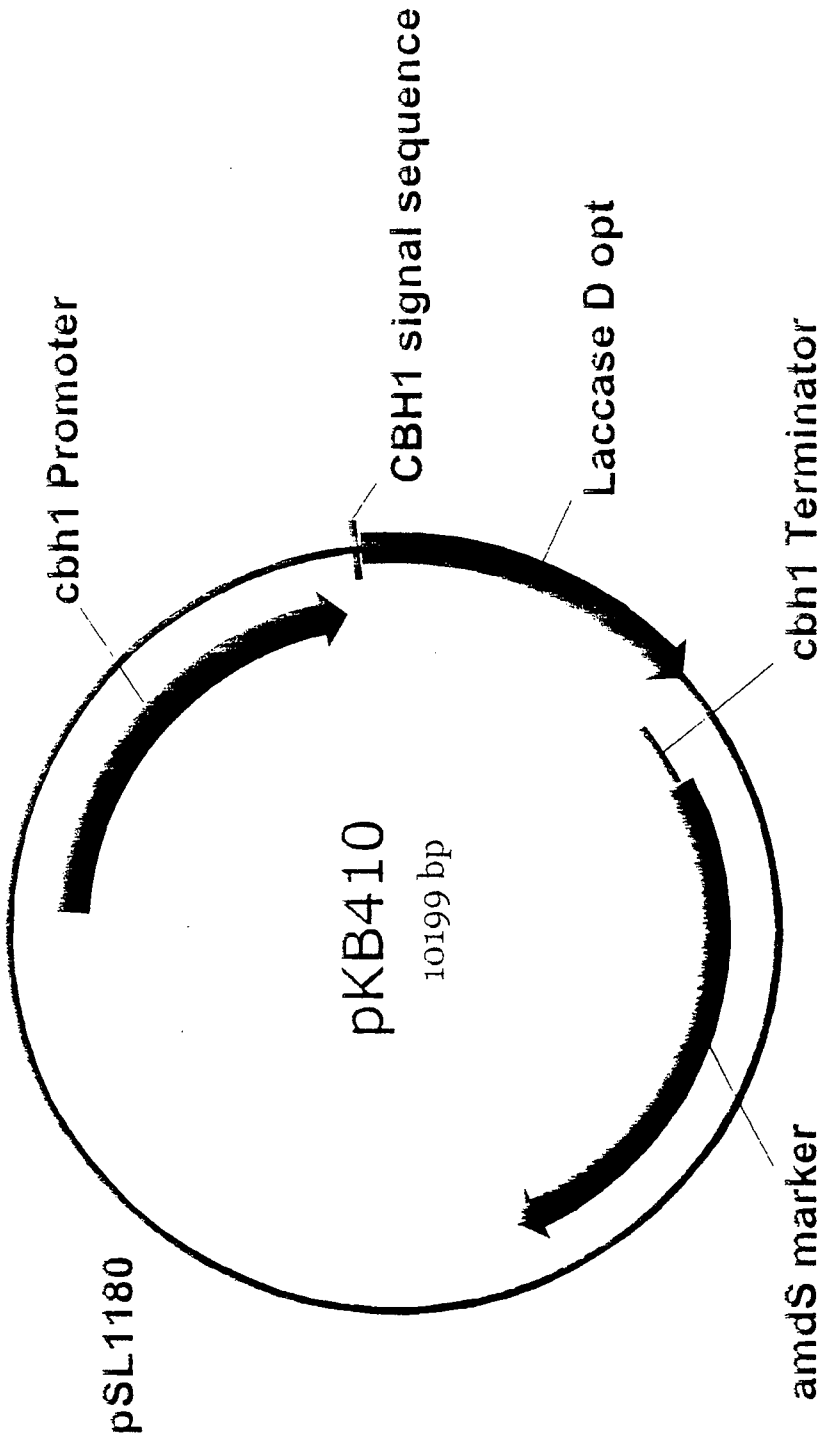


FIG. 7

8/16

NSP24 nucleic acid sequence (SEQ ID NO: 8):

CTGCAGCCACTTGCAGTCCCGTGGGAATCTCACGGTGAATGTAGGCCTTTTGTAGGG
TAGGAATTGTCACCTCAAGCACCCCAACCTCCATTACGCCCTCCCCCATAGAGTTCCC
AATCAGTGAGTCATGGCACTGTTCTCAAAATAGATTGGGGAGAAGTTGACTTCCGCCC
AGAGCTGAAGGTCGCACAACCGCATGATATAGGGTTCGGCAACGGCAAAAAAGCACGT
GGCTCACCGAAAAGCAAGATGTTTGCGATCTAACATCCAGGAACCTGGATACATCCA
TCATCACGCACGACCACCTTTGATCTGCTGGTAAACTCGTATTCGCCCTAAACCGAAG
TGCGTGGTAAATCTACACGTGGGCCCCCTTTCGGTATACTGCGTGTGTCTTCTCTAGG
TGCCATTCTTTTCCCTTCTCTAGTGTGAATTGTTTGTGTTGGAGTCCGAGCTGTA
ACTACCTCTGAATCTCTGGAGAATGGTGGACTAACGACTACCGTGCACCTGCATCAT
GTATATAATAGTGATCCTGAGAAGGGGGGTTTGGAGCAATGTGGGACTTTGATGGTC
ATCAAACAAAGAACGAAGACGCCTCTTTTGCAAAGTTTGTTCGGCTACGGTGAAG
AACTGGATACTTGTGTGTCTTCTGTGTATTTTGTGGCAACAAGAGGCCAGAGACA
ATCTATTCAAACACCAAGCTTGCTCTTTTGGAGCTACAAGAACCTGTGGGGTATATAT
CTAGAGTTGTGAAGTCGGTAATCCCGCTGTATAGTAATACGAGTCGCATCTAAATAC
TCCGAAGCTGCTGCGAACCCGGAGAATCGAGATGTGCTGGAAAGCTTCTAGCGAGCG
GCTAAATTAGCATGAAAGGCTATGAGAAATCTGGAGACGGCTTGTGAATCATGGC
GTTCCATTCTTCGACAAGCAAAGCGTTCCGTCGCAGTAGCAGGCACTCATTCCCGAA
AAAACCTCGGAGATTCTTAAGTAGCGATGGAACCGGAATAATATAATAGGCAATACAT
TGAGTTGCCTCGACGGTTGCAATGCAGGGGTACTGAGCTTGGACATAACTGTTCCGT
ACCCACCTCTTCTCAACCTTTGGCGTTTCCCTGATTACAGCGTACCCGTACAAGTCG
TAATCACTATTAACCCAGACTGACCGGACGTGTTTTCCTTCATTGAGAGAAATAA
TGTCATTGCGATGTGTAATTTGCCTGCTTGACCGACTGGGGCTGTTTGAAGCCCGAA
TGTAGGATGTTATCCGAACCTGCTCGTAGAGGCATGTTGTGAATCTGTGTGCGGC
AGGACACGCCCTCGAAGGTTACGGCAAGGGAAACCACCGATAGCAGTGTCTAGTAGC
AACCTGTAAAGCCGCAATGCAGCATCACTGGAAAATACAAACCAATGGCTAAAAGTA
CATAAGTTAATGCCTAAAGAAGTCATATACCAGCGGCTAATAATTGTACAATCAAGT
GGCTAAACGTACCGTAATTTGCCAACGGCTTGTGGGGTTGCAGAAGCAACGGCAAAAG
CCCCACTTCCCCACGTTTGTCTTCACTCAGTCCAATCTCAGCTGGTGATCCCCCA
ATTGGGTCGCTTGTGTTCGGGTGAAGTGAAAGAAGACAGAGGTAAGAATGTCTGA
CTCGGAGCGTTTTGCATACAACCAAGGGCAGTGATGGAAGACAGTGAAATGTTGACA
TTCAAGGAGTATTTAGCCAGGGATGCTTGAGTGTATCGTGTAAGGAGGTTTGTCTGC
CGATACGACGAATACTGTATAGTCACTTCTGATGAAGTGGTCCATATTGAAATGTAA
GTCGGCACTGAACAGGCAAAAGATTGAGTTGAAACTGCCTAAGATCTCGGGCCCTCG
GGCCTTCGGCCTTTGGGTGTACATGTTTGTGCTCCGGGCAAATGCAAAGTGTGGTAG
GATCGAACACACTGCTGCCTTTACCAAGCAGCTGAGGGTATGTGATAGGCAAATGTT
CAGGGGCCACTGCATGGTTTCGAATAGAAAGAGAAGCTTAGCCAAGAACAATAGCCG
ATAAAGATAGCCTCATTAAACGGAATGAGCTAGTAGGCAAAGTCAGCGAATGTGTAT
ATATAAAGGTTTCGAGGTCCGTGCCTCCCTCATGCTCTCCCATCTACTCATCAACTC
AGATCCTCCAGGAGACTTGTAACCATCTTTTGGAGCACAGAAACCAATAGTCAAC
CATCACAAGTTTGTACAAAAAAGCAGGCTCCGCGGCCGCCCCCTTACCATGCAGAC
CTTTGGAGCTTTTCTCGTTTCTTCTCGCCGCCAGCGGCCTGGCCGCGGCCCTCCC
CACCGAGGGTCAGAAGACGGCTTCCGTCGAGGTCCAGTACAACAAGAACTACGTCCC
CCACGGCCCTACTGCTCTCTTCAAGGCCAAGAGAAAGTATGGCGCTCCCATCAGCGA

FIG. 8-1

SUBSTITUTE SHEET (RULE 26)

9/16

CAACCTGAAGTCTCTCGTGGCTGCCAGGCAGGCCAAGCAGGCTCTCGCCAAGCGCCA
GACCGGCTCGGCGCCCAACCACCCAGTGACAGCGCCGATTCGGAGTACATCACCTC
CGTCTCCATCGGCACTCCGGCTCAGGTCTCCCCCTGGACTTTTGACACCGGCTCCTC
CGACCTGTGGGTCTTTAGCTCCGAGACGCCCAAGTCTTCGGCCACCGGCCACGCCAT
CTACACGCCCTCCAAGTCGTCCACCTCCAAGAAGGTGTCTGGCGCCAGCTGGTCCAT
CAGCTACGGCGACGGCAGCAGCTCCAGCGCGATGTCTACACCGACAAGGTCACCAT
CGGAGGCTTCAGCGTCAACACCCAGGGCGTCGAGTCTGCCACCCGCGTGTCCACCGA
GTTTCGTCCAGGACACGGTCATCTCTGGCCTCGTCGGCCTTGCCCTTTGACAGCGGCAA
CCAGGTCAGGCCGACCCGCGAGAAGACGTGGTTCTCCAACGCCGCCAGCAGCCTGGC
TGAGCCCCCTTTTCACTGCCGACCTGAGGCACGGACAGAGTAAGTAGACACTCACTGG
AATTCGTTCCCTTTCCCGATCATCATGAAAGCAAGTAGACTGACTGAACCAACAACCT
AGACGGCAGCTACAACCTTTGGCTACATCGACACCAGCGTCGCCAAGGGCCCCGTTGC
CTACACCCCCGTTGACAACAGCCAGGGCTTCTGGGAGTTCACTGCCTCGGGCTACTC
TGTCGGCGGGCGCAAGCTCAACCGCAACTCCATCGACGGCATTGCCGACACCGGCAC
CACCCTGCTCCTCCTCGACGACAACGTCTGTCGATGCCCTACTACGCCAACGTCCAGTC
GGCCCAGTACGACAACCAGCAGGAGGGTGTCTCTTCGACTGCGACGAGGACCTCCC
TTCGTTTCAGCTTCGGTGTGGAAGCTCCACCATCACCATCCCTGGCGATCTGCTGAA
CCTGACTCCCCCTCGAGGAGGGCAGCTCCACCTGCTTCGGTGGCCTCCAGAGCAGCTC
CGGCATTGGCATCAACATCTTTGGTGACGTTGCCCTCAAGGCTGCCCTGGTTGTCTT
TGACCTCGGCAACGAGCGCCTGGGCTGGGCTCAGAAATAAAAGGGTGGGCGCGCCGA
CCCAGCTTTCTTGTACAAAGTGGTGATCGCGCCAGCTCCGTGCGAAAGCCTGACGCA
CCGGTAGATTCTTGGTGAGCCCGTATCATGACGGCGGCGGGAGCTACATGGCCCCGG
GTGATTTATTTTTTTTGTATCTACTTCTGACCCTTTTCAAATATACGGTCAACTCAT
CTTTCCTGAGATGCGGCCTGCTTGGTATTGCGATGTTGTGTCAGCTTGGCAAATTGT
GGCTTTCGAAAACACAAAACGATTCCTTAGTAGCCATGCATTTTAAGATAACGGAAT
AGAAGAAAGAGGAAATTAAAAAACAACATCCCCGTTTCATAACCCGTA
GAATCGCCGCTCTTCGTGTATCCAGTACCAGTTTATTTTGAATAGCTCGCCCCGTG
GAGAGCATCCTGAATGCAAGTAACAACCGTAGAGGCTGACACCGGCAGGTGTTGCTAG
GGAGCGTCGTGTTCTACAAGGCCAGACGTCTTCGCGGTTGATATATATGTATGTTTG
ACTGCAGGCTGCTCAGCGACGACAGTCAAGTTCGCCCTCGCTGCTTGTGCAATAATC
GCAGTGGGGAAGCCACACCGTGACTCCCATCTTTCAGTAAAGCTCTGTTGGTGTTTA
TCAGCAATACAGTAATTTAACTCGTTAGCATGGGGCTGATAGCTTAATTACCGTT
TACCAGTGCCATGGTTCTGCAGCTTTCCTTGGCCCGTAAAATTCGGCGAAGCCAGCC
AATCACCAGCTAGGCACCAGCTAAACCCTATAATTAGTCTCTTATCAACACCATCCG
CTCCCCCGGGATCAATGAGGAGAATGAGGGGGATGCGGGGCTAAAGAAGCCTACATA
ACCCTCATGCCAACTCCAGTTTACACTCGTCGAGCCAACATCCTGACTATAAGCTA
ACACAGAATGCCTCAATCCTGGGAAGAAGTGGCCGCTGATAAGCGCGCCCGCCTCGC
AAAAACCATCCCTGATGAATGGAAAGTCCAGACGCTGCCTGCGGAAGACAGCGTTAT
TGATTTCCCAAAGAAATCGGGGATCCTTTTCAGAGGCCGAAGTGAAGATCACAGAGGC
CTCCGCTGCAGATCTTGTGTCCAAGCTGGCGGCCGAGAGTTGACCTCGGTGGAAGT
TACGCTAGCATCTGTAAACGGGCAGCAATCGCCCAGCAGTTAGTAGGGTCCCCCTCT
ACCTCTCAGGGAGATGTAACAACGCCACCTTATGGGACTATCAAGCTGACGCTGGCT
TCTGTGCAGACAACTGCGCCACGAGTTCTTCCCTGACGCCGCTCTCGCGCAGGCA
AGGGAACTCGATGAATACTACGCAAAGCACAAAGAGACCCGTTGGTCCACTCCATGGC
CTCCCCATCTCTCTCAAAGACCAGCTTCGAGTCAAGGTACACCGTTGCCCTAAGTC
GTTAGATGTCCCTTTTGTGTCAGCTAACATATGCCACCAGGGCTACGAAACATCAATG
GGCTACATCTCATGGCTAAACAAGTACGACGAAGGGGACTCGGTTCTGACAACCATG
CTCCGCAAAGCCGGTGCCGTCTTCTACGTCAAGACCTCTGTCCCGCAGACCCTGATG

FIG. 8-2

SUBSTITUTE SHEET (RULE 26)

10/16

GTCTGCGAGACAGTCAACAACATCATCGGGCGCACCGTCAACCCACGCAACAAGAAC
TGGTCGTGCGGCGGCAGTTCTGGTGGTGAGGGTGCGATCGTTGGGATTTCGTGGTGGC
GTCATCGGTGTAGGAACGGATATCGGTGGCTCGATTTCGAGTGCCGGCCGCGTTCAAC
TTCCTGTACGGTCTAAGGCCGAGTCATGGGCGGCTGCCGTATGCAAAGATGGCGAAC
AGCATGGAGGGTCAGGAGACGGTGCACAGCGTTGTCGGGCCGATTACGCACTCTGTT
GAGGGTGAGTCCTTCGCCTCTTCCTTCTTTTCTGCTCTATAACCAGGCCTCCACTGT
CCTCCTTTCTTGCTTTTATACTATATACGAGACCGGCAGTCACTGATGAAGTATGT
TAGACCTCCGCCTCTTCACCAAATCCGTCCTCGGTTCAGGAGCCATGGAAATACGACT
CCAAGGTCATCCCCATGCCCTGGCGCCAGTCCGAGTCGGACATTATTCCTCCAAAGA
TCAAGAACGGCGGGCTCAATATCGGCTACTACAACCTTCGACGGCAATGTCTTCCAC
ACCTCCTATCCTGCGCGGCGTGGAACCACCGTCGCCGCACTCGCCAAAGCCGGTC
ACACCGTGACCCCGTGACGCCATACAAGCACGATTTTCGGCCACGATCTCATCTCCC
ATATCTACGCGGCTGACGGCAGCGCCGACGTAATGCGCGATATCAGTGCATCCGGCG
AGCCGGCGATTCCAAATATCAAAGACCTACTGAACCCGAACATCAAAGCTGTAAACA
TGAACGAGCTCTGGGACACGCATCTCCAGAAGTGGAATTACCAGATGGAGTACCTTG
AGAAATGGCGGGAGGCTGAAGAAAAGGCCGGGAAGGAAGTGGACGCCATCATCGCGC
CGATTACGCCTACCGCTGCGGTACGGCATGACCAGTTCCGGTACTATGGGTATGCCT
CTGTGATCAACCTGCTGGATTTCACGAGCGTGTTGTTCCGGTTACCTTTGCGGATA
AGAACATCGATAAGAAGAATGAGAGTTTCAAGGCGGTTAGTGAGCTTGATGCCCTCG
TGCAGGAAGAGTATGATCCGGAGGCGTACCATGGGGCACCGGTTGCAGTGCAGGTTA
TCGGACGGGAGACTCAGTGAAGAGAGGACGTTGGCGATTGCAGAGGAAGTGGGGAAGT
TGCTGGGAAATGTGGTGACTCCATAGCTAATAAGTGTGATAGCAATTTGCACAAG
AAATCAATACCAGCAACTGTAAATAAGCGCTGAAGTGACCATGCCATGCTACGAAAG
AGCAGAAAAAACCTGCCGTAGAACCGAAGAGATATGACACGCTTCCATCTCTCAA
GGAAGAATCCCTTCAGGGTTGCGTTTCCAGTCTAGACACGTATAACGGCACAAGTGT
CTCTACCAAAATGGGTTATATCTCAAATGTGATCTAAGGATGGAAAGCCAGAATAT
CGATCGCGCGCAGATCCATATATAGGGCCCGGTTATAATTACCTCAGGTCGACGTC
CCATGGCCATTTCGAATTTCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTGTTA
TCCGCTCACAAATCCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGG
TGCCTAATGAGTGAGCTAACTCACATTAATTCGCTTGCGCTCACTGCCCCGTTTCCA
GTCCGGAAACCTGTCTGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGG
CGGTTTGCGTATTGGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGT
CGTTCCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCAC
AGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAG
GAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGA
GCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAG
ATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCC
GCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGCGCTTCTCATAG
CTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGTTTCGCTCCAAGCTGGGCTGTGT
GCACGAACCCCCGTTACGCCCCGACCGCTGCGCCTTATCCGGTAACCTATCGTCTTGA
GTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGAT
TAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA
CGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT
CGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGG
TTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCC
TTTGATCTTTTCTACGGGCTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGAT
TTTGCTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAATG
AAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATG

FIG. 8-3

SUBSTITUTE SHEET (RULE 26)

11/16

CTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGC
CTGACTCCCCGTCGTGTAGATAAATACTACGATACGGGAGGGCTTACCATCTGGCCCCAG
TGCTGCAATGATAACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAA
CCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCCTCCAT
CCAGTCTATTAATTGTTGCGGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTT
GCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCGTTTGGTAT
GGCTTCATTTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTT
GTGCAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGC
CGCAGTGTTTACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTTCATGCC
ATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATA
GTGTATGCGGCGACCGAGTTGCTCTTGCCCCGGCGTCAATACGGGATAATACCGCGCC
ACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACCT
CTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAA
CTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG
GCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACT
CTTCCTTTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGACGGGATA
CATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCGC
AAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAA
TAGGCGTATCACGAGGCCCTTTTCGTCTCGCGCGTTTCGGTGATGACGGTGAAAACCT
CTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAG
CAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGCGGGGTGTCGGGGCTGGCTTAA
CTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATAAAAATTGTAAACGTTAA
TATTTTGTTAAAATTCGCGTTAAATTTTTTGTTAAATCAGCTCATTTTTTTAACCAATA
GGCCGAAATCGGCAAAATCCCTTATAAATCAAAAAGAAATAGCCCGAGATAGGGTTGAG
TGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAA
AGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCAAATC
AAGTTTTTTGGGGTTCGAGGTGCCGTAAAGCACTAAATCGGAACCTAAAGGGAGCCC
CCGATTTAGAGCTTGACGGGGAAGCCGGCGAACGTGGCGAGAAAGGAAGGAAGAA
AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGATGCGGTCACGCTGCGCGTAAC
CACCACACCCGCGCGCTTAATGCGCCGCTACAGGGCGGTACTATGGTTGCTTTGA
CGTATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGC
CATTGCGCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCG
CTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAAGTTGGGTAAACG
CCAGGGTTTTCCAGTCACGACGTTGTAAAACGACGGCCAGTGCCCAAGCTTACTAG
TACTTCTCGAGCTCTGTACATGTCCGGTCGCGACGTACGCGTATCGATGGCGCCAGC
TGCAGGCGGCGCG

FIG. 8-4

SUBSTITUTE SHEET (RULE 26)

CBHI MAP

90	AA GCTTAGCCAAAGAACAAATAGCCCGATAAAGATAGCCCTCATTAACGGGAATGAGCTAGTGGCAAAAGTCAGCGAATGTGTATATATAAAGG
180	TTCCGAGGTCCTCGTCCCTCATGCTCTCCCATCTACTCATCAACTCAGATCCCTCCAGGAGACTTGTACACCATNTTTTGAGGACACAGA
270	AACCCAAATAGTCAACCGGAGCTGGCATCATGTATCGGAAGTTGGCCGTCATCTCGGCCCTTCTTGGCCACAGCTCGTGTCTCAGTCGGCCT
360	Met Tyr Arg Lys Leu Ala Val Ile Ser Ala Phe Leu Ala Thr Ala Arg Ala Gln Ser Ala GCACTCTCCAATCGGAGACTCACCCCGCTCTGACATGGCAGAAAATGCTCGTCTGGTGGCACTTGCACCTCAACAGACAGGCTCCGTGGTCA
450	Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr Trp Gln Lys Cys Ser Ser Gly Gly Thr Cys Thr Gln Gln Thr Gly Ser Val Val TCGACGCCCAACTGGCGCTGGACTCACGCTACGAACAGCAGCACGAACCTGCTACGATGGCAACACTTGGAGCTCGACCCCTATGTCTCTGACA
540	Ile Asp Ala Asn Trp Arg Trp Thr His Ala Thr Asn Ser Ser Thr Asn Cys Tyr Asp Gly Asn Thr Trp Ser Ser Thr Leu Cys Pro Asp ACGAGACCTGGCGGAAGAACTGCTGTCTGGACGGTGCGCGCTACGCGTCCACGTACGGAGTTACCCACGAGCGGTAACAGCCCTCTCCATTG
630	Asn Glu Thr Cys Ala Lys Asn Cys Cys Leu Asp Gly Ala Ala Tyr Ala Ser Thr Tyr Gly Val Thr Thr Ser Gly Asn Ser Leu Ser Ile GCTTTGTACCCAGTCTGGCGCAGAAGAACGTTGGCGCTCGCCCTTACCTTATGGCGAGCGACACGACCTACCAGGAATTCACCCCTGCTTG
720	Gly Phe Val Thr Gln Ser Ala Gln Lys Asn Val Gly Ala Arg Leu Tyr Leu Met Ala Ser Asp Thr Thr Tyr Gln Glu Phe Thr Leu Leu GCAACGAGTCTCTTTCGATGTTGATGTTTCGCAGCTGCCGTAAGTACTTACCATGAACCCCTGACGTATCTTCTTGTGGGCTCCCAGC
810	Gly Asn Glu Phe Ser Phe Asp Val Asp Val Ser Gln Leu Pro TGACTGGCCCAATTTAAGGTGGCGCTTGAACGGAGCTCTCTACTTCTCGTGTCCATGGACGCGGATGGTGGCGTGAGCAAGTATCCCAAC
900	ACCCTGGCGCCCAAGTACGGCACGGGTACTGTGACAGCCAGTGTCCCCCGGATCTGAAGTTCATCAATGGCCAGGCCAACGTTGAGGGG
990	Thr Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp Ser Gln Cys Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln Ala Asn Val Glu Gly TGGGAGCCGTCATCCCAACAACGCAACACGGGCATTTGAGGACACGGAAGCTGCTCTGAGATGGATATCTGGGAGGCCCAACTCCATC
1080	Trp Glu Pro Ser Ser Asn Asn Ala Asn Thr Gly Ile Gly Gly His Gly Ser Cys Ser Glu Met Asp Ile Trp Glu Ala Asn Ser Ile TCCGAGGCTCTTACCCCCCACCCTTGACGACTGTGGCCAGGAGATCTGCGAGGGTGATGGTGGCGGGAACCTTACTCCGATAACAGA

FIG. 9-1

Ser Glu Ala Leu Thr Pro His Pro Cys Thr Thr Val Gly Gln Glu Ile Cys Glu Gly Asp Gly Cys Gly Thr Tyr Ser Asp Asn Arg
 TATGGCGGCACTTGGATCCCGATGGCTGGACTTGAACCCATACCGCTGGGCAACACCAAGCTTCTACGGCCCTGGCTCAAGCTTTACC
 1170

Tyr Gly Gly Thr Cys Asp Pro Asp Gly Cys Asp Tyr Asn Pro Tyr Arg Leu Gly Asn Thr Ser Phe Tyr Gly Pro Gly Ser Phe Thr
 CTCGATACCAACAAGAAATTGACCGTTTGTACCCAGTTCGAGACGTCCGGTGCCATCAACCGATACTATGTCCAGAAATGGCGTCACTTTC
 1260

Leu Asp Thr Thr Lys Lys Leu Thr Val Val Thr Gln Phe Glu Thr Ser Gly Ala Ile Asn Arg Tyr Tyr Val Gln Asn Gly Val Thr Phe
 CAGCAGCCCCAAGCGGAGCTTGGTAGTTACTTCTGGCAACGAGCTCAACGATGATTACTGCACAGCTGAGGAGGCAGAAATTCGGCGGATCC
 1350

Gln Gln Pro Asn Ala Glu Leu Gly Ser Tyr Ser Gly Asn Glu Leu Asn Asp Tyr Cys Thr Ala Glu Ala Glu Phe Gly Gly Ser
 TCTTCTCAGACAAGGGCGGCTGACTCAGTTCAGAGGCTACCTCTGGCGCATGGTCTGTGTCATGAGTCTGTGGGATGATGTGAGT
 1440

Ser Phe Ser Asp Lys Gly Gly Leu Thr Gln Phe Lys Lys Ala Thr Ser Gly Met Val Leu Val Met Ser Leu Trp Asp Asp
 TTGATGGACAAACATGCGCGTTGACAAAAGAGTCAAGCAGCTGACTGAGATGTTACAGTACTACGCCAACATGCTGTGGCTGGACTCCACC
 1530

TACCCGACAAACGAGACCTCCTCCACACCCCGTGCCGTGCGGGAAGCTGCTCCACCAGCTCCGGTGTCCCTGCTCAGGTGGAATCTCAG
 Tyr Tyr Ala Asn Met Leu Trp Leu Asp Ser Thr
 1620

Tyr Pro Thr Asn Glu Thr Ser Ser Thr Pro Gly Ala Val Arg Gly Ser Cys Ser Thr Ser Ser Gly Val Pro Ala Gln Val Glu Ser Gln
 TCTCCCAACGCCAAGGTCACCTTCTCCCAACATCAAGTTCGGACCCATGGCAGCACCAGGCAACCCCTAGCGGGCGGCAACCCCTCCCGGCGGA
 1710

Ser Pro Asn Ala Lys Val Thr Phe Ser Asn Ile Lys Phe Gly Pro Ile Gly Ser Thr Gly Asn Pro Ser Gly Gly Asn Pro Pro Gly Gly
 AACCGTGGCACCACCACCCCGCGCGCCAGCCACTACCACTGGGAAGCTCTCCCGGACCTACCCAGTCTCACTACGCGCCAGTGGCGCGGT
 1800

Asn Arg Gly Thr Thr Thr Arg Arg Pro Ala Thr Thr Thr Gly Ser Ser Pro Gly Pro Thr Gln Ser His Tyr Gly Gln Cys Gly Gly
 ATTGGCTACAGCGGCGCCACGGTCTGCGCCAGCGGCACAACTTGCCAGGTCCCTGAACCCCTTACTACTCTCAGTGCCTGTAAAGCTCCCGTG
 1890

Ile Gly Tyr Ser Gly Pro Thr Val Cys Ala Ser Gly Thr Thr Cys Gln Val Leu Asn Pro Tyr Tyr Ser Gln Cys Leu
 CGAAAGCCTGACGCACCGGTAGATTCTTGGTGAGCCCGTATCATGACGGCGCGGAGCTACATGGCCCCGGGTGATTTATTTTGT
 1980

ATCTACTTCTGACCCCTTTTCAAATATACGGTCAACTCATCTTCACTGGAGATGCGGCCCTGCTTGGTATTTGGGATGTTGTGTCAGCTTGGCA
 2070

AATGTGGCTTTCGAAAAACACAAAACGATTCTTAGTAGCCATGTCATTTTAAGATAACGGAAATAGAAAGAGGAAATTAATAAAAAA
 2160

AAAAAACAAACATCCCGTTTCATAACCCGTAGAAATCGCCGCTCTTCTGTATCCAGTACCA
 2221

FIG. 9-2

14/16

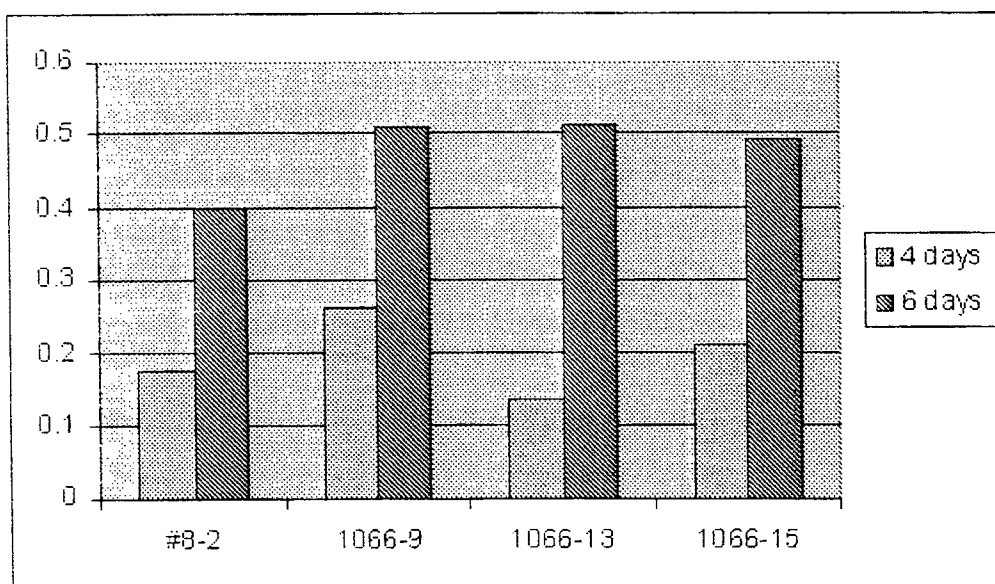


FIG. 10

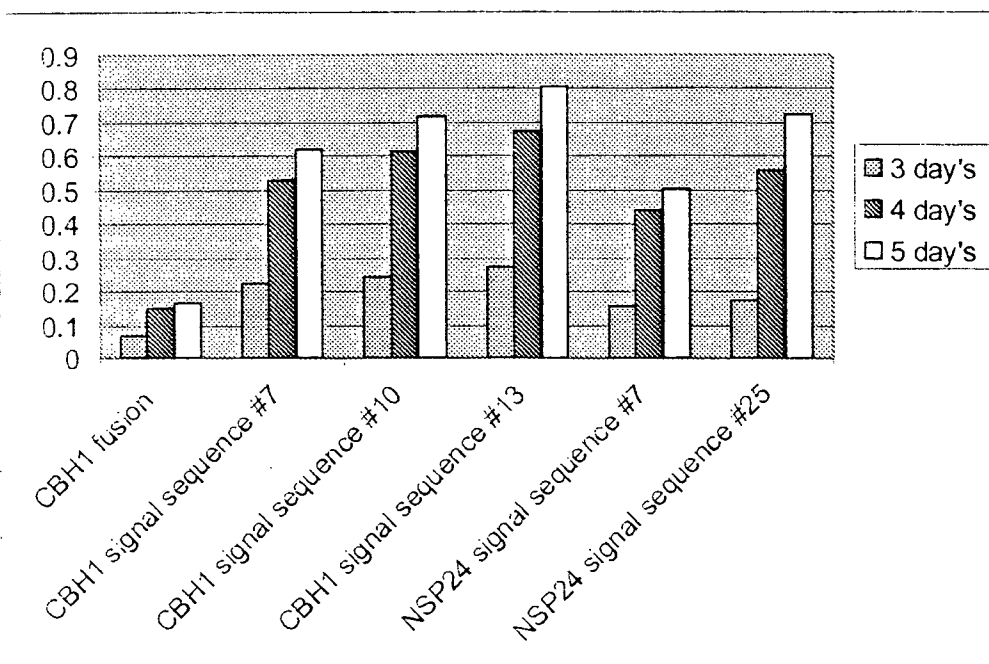


FIG. 11

SUBSTITUTE SHEET (RULE 26)

15/16

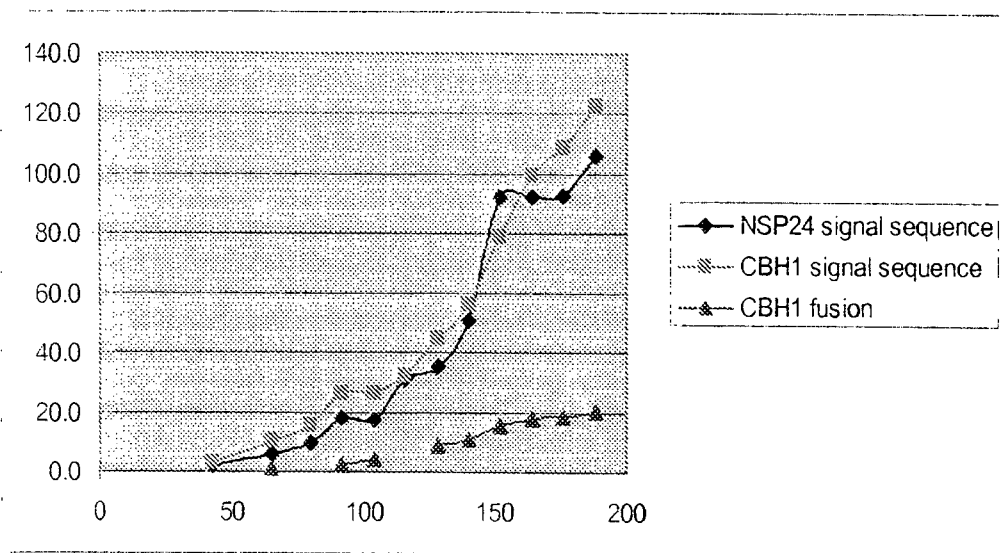


FIG. 12

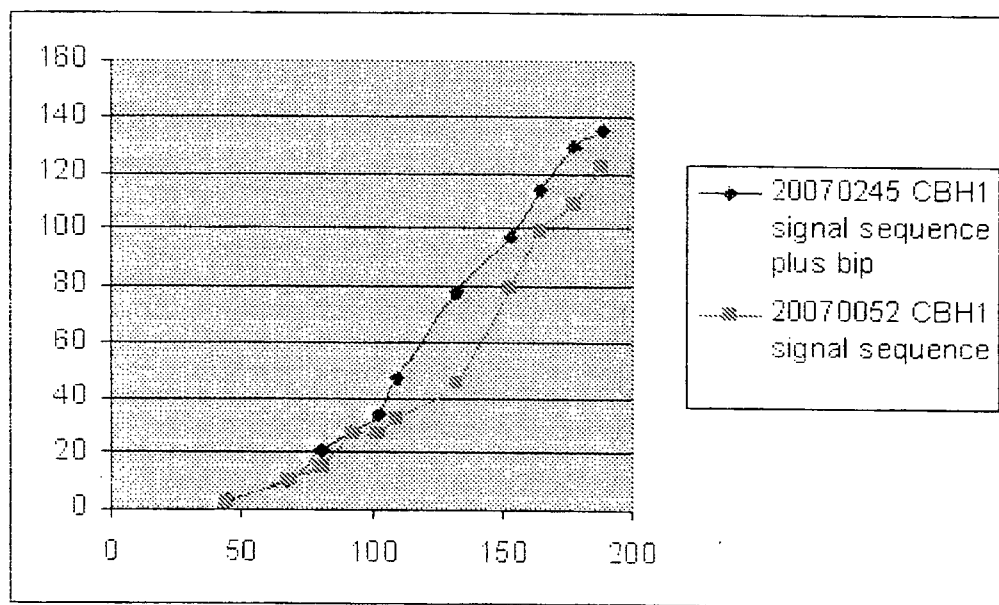


FIG. 13

SUBSTITUTE SHEET (RULE 26)

16/16

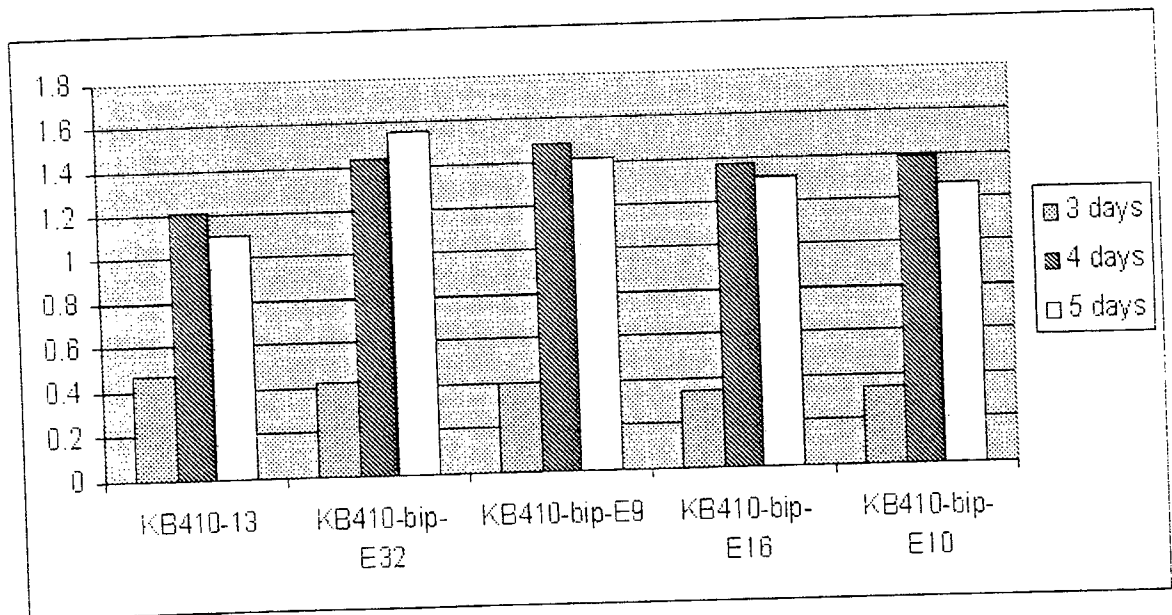


FIG. 14

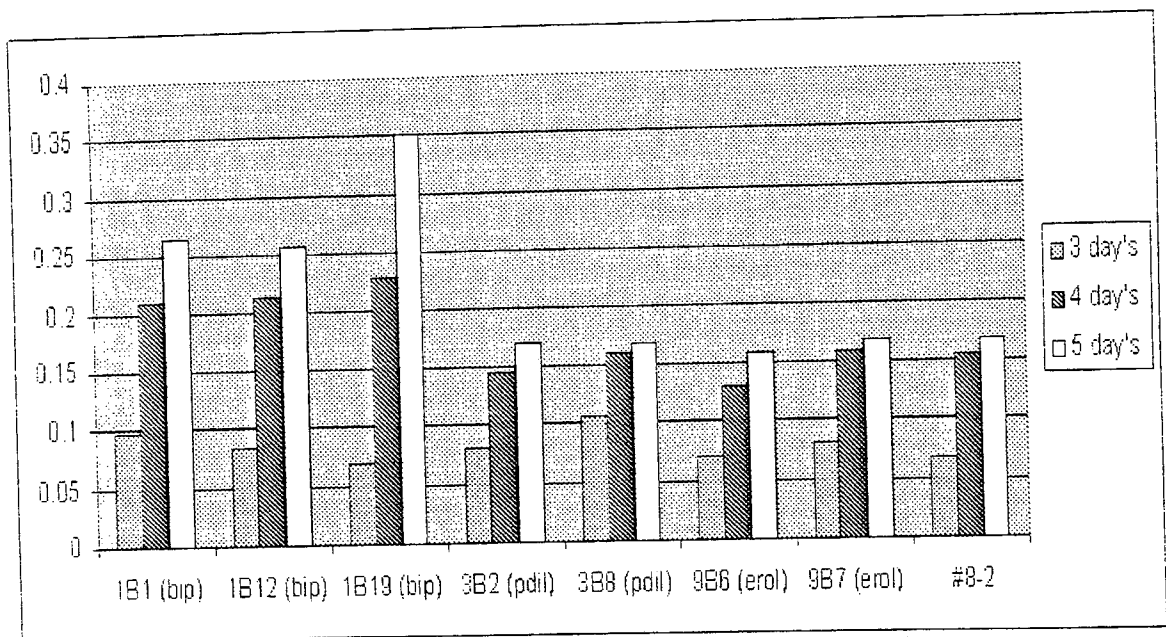


FIG. 15

SUBSTITUTE SHEET (RULE 26)

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 10-048157-00000-0000

Fecha: 2010-04-23 17:13:15 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 142

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800176089-2

RECIBO OFICIAL DE CAJA : 10 - 31,625
FECHA : ABRIL 23 DE 2010

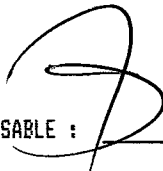
***** CONSIGNACION *****

DEPOSITANTE	TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PGO	Vr. PAGO
CASTILLO PINEDA	CONSIGNACION	BANCO DE BOGOTA	062754387	105315335	23/04/2010	6,528,000.00

***** CONCEPTO *****

CANT.	RENTISTICO	CONCEPTO	TOTAL CONCEPTO
1	50005-01-01 SOLICITUDES	1 TRAMITES DE SOL. DE PATENTE DE IN	554,000.00
		TOTAL :	\$ 554,000.00

SON: QUINIENTOS CINCUENTA Y CUATRO MIL PESOS

RESPONSABLE : 

RECIBO DE CAJA APLICADO AL EXPEDIENTE No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 10-048157- -00000-0000

Fecha: 2010-04-23 17:13:15 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 142

NIT : 8.000.000

CERTIFICACION DE EXCEDENTE : 10 - 31,628
FECHA : ABRIL 23 DE 2010

***** CONSIGNACION *****

DEPOSITANTE	TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PGO	Vr. PAGO
CASTILLO PINEDA	CONSIGNACION	BANCO DE BOGOTA	062754387	105315335	23/04/2010	6,528,000.00

***** CONCEPTO *****

CANT.	RENTISTICO	CONCEPTO	TOTAL CONCEPTO
1	50005-01-09	OTROS SERVICIOS DE PROPIEDAD I	
99		OTROS SERVCIOS ADMINISTRATIVOS	286,000.00
		TOTAL :	\$ 286,000.00

SON: DOSCIENTOS OCHENTA Y SEIS MIL PESOS

RESPONSABLE :

RECIBO DE CAJA APLICADO AL EXPEDIENTE No. _____

ESTA CERTIFICACION DE EXCEDENTE DE PAGO SOLO PODRA SER UTILIZADA
COMO PARTE DE LA CANCELACION DE UNA TARIFA VIGENTE Y NO PODRA SER
SUSTITUIDA POR NINGUN RECIBO.



No. 10-048157-00000-0000

S

Fecha: 2010-04-23 17:13:15 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 142

PATENTE DE INVENCION ☐

MODELO ☐

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 INVENCION 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 ☐