

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
No. 11-125473-00000-0000
Fecha: 2011-09-26 14:15:31 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYI Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 223

Now
26/9

Industria y Comercio

SUPERINTENDENCIA

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
INVENTARIO

DELEGATURA DE PROPIEDAD INDUSTRIAL

División de Nuevas Creaciones

PCT
SOLICITUD

PATENTE DE INVENCION

11-125473

21. EXPEDIENTE No.

54. TÍTULO

POLYPEPTIDOS QUE TIENEN ACTIVIDAD DE CELULASA

51. CLASIFICACIÓN INTERNACIONAL

C12N 15/63

71. SOLICITANTE

CALIFORNIA INSTITUTE OF TECHNOLOGY

DOMICILIO

PASADENA, ESTADO DE CALIFORNIA, E.U.A.

74. APODERADO

LUIS FELIPE CASTILLO

22. BOGOTÁ, D.C.,

26 SET. 2011

Pol., L.
06-11-2011
Ni xcc

PETITORIO



SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-125473-00000-0000

Fecha: 2011-09-26 14:15:31
Tra. 11 PATENTEIYI
Act. 411 PRESENTACION

Dep. 2020 DIR.NUEVASCR
Eve: 378 FASENACIONALI
Folios: 223

Act. 411 PRESENTACION Folios: 214

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: CALIFORNIA INSTITUTE OF TECHNOLOGY

Dirección: California Institute of Technology,
1200 E. California Blvd., MS 201-85

Nacionalidad o domicilio: Pasadena, Estado de California E.U.A.

IDENTIFICACIÓN

C.C. ☐ NIT ☐

C.E. ☐ Otro ☐

Cual _____
Número _____

3

REPRESENTANTE
O APODERADO

Nombre: LUIS FELIPE CASTILLO GIBSON

Dirección: Carrera 13 No. 37-43 Piso 12

Teléfono: 285 74 60 Fax: 285 92 67
E-mail info@castillograu.com

IDENTIFICACIÓN

C.C. ☒ NIT ☐

C.E. ☐ OTRO ☐

Cual _____

Número 79.397.879 de TP 68.995 del
Bogotá C.S.J.

4

INVENTOR(ES)
(72)

Nombre: Frances H. ARNOLD - Pete HEINZELMAN

Dirección 4341 Encinas Dr. - 44 North Oak Avenue

Nacionalidad o Domicilio: La Canada, CA. - Pasadena, CA. (E.U.A.)

5

PCT (86)

Solicitud Internacional No. PCT/US2010/030133 Fecha: 06/04/2010

Publicación Internacional No. WO 2010/118058 A2 Fecha: 14/10/2010

6

Título (54) POLIPÉPTIDOS QUE TIENEN ACTIVIDAD DE CELULASA

7

Clasificación Internacional (51) C12N 15/63

8

Prioridad

Si

☒

No

☐

(33) País de origen
E.U.A.

(32) Fecha

06.04.2009 -
13.05.2009

(31) Número de Solicitud

61/166,993 - 61/177,882

9

Comprobante de pago N°:

94094

Fecha:

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

Recibido: 11/09/2011 12:00

10

ANEXOS

- ☒ Comprobante de pago de la tasa de presentación de la solicitud
- ☐ Documento que acredita la existencia y representación legal cuando el solicitante sea persona jurídica.
- ☐ Poderes, si fuere el caso
- ☐ Certificado de existencia y representación legal cuando el solicitante sea persona jurídica
- ☐ Documento de Cesión del inventor al solicitante o a su acusante
- ☐ 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 efectuado por la IPEA.
- ☐ Traducción del examen preliminar 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 de material biológico.
- ☒ Arte final 12 x 12
- ☐ De ser el caso, información sobre otras solicitudes de patente o títulos obtenidos en el extranjero por el mismo titular o su causante, relacionadas parcial o totalmente con la invención de esta solicitud.
- ☐ Otros

11

FIGURA CARACTERISTICA

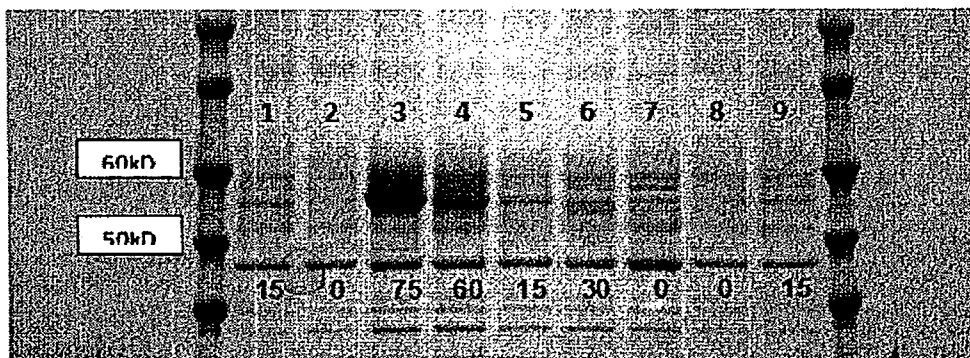


FIGURA 1A

12

Solicito la concesión de patente,

NOMBRE: LUIS FELIPE CASTILLO GIBSONFIRMA: C.C. 79397-879
de BogotáT.P. 68.995
del C.S.J.

POLIPÉPTIDOS QUE TIENEN ACTIVIDAD DE CELULASA

REFERENCIA CRUZADA CON SOLICITUDES RELACIONADAS

[0001] La solicitud reivindica prioridad según 35 U.S.C. §119 a la Solicitud provisional estadounidense Nos. de serie 61/166.993, radicada el 6 de abril de 2009 y 61/177.882, radicada el 13 de mayo de 2009, cuyas revelaciones se incorporan a la presente mediante referencia.

DECLARACIÓN SOBRE INVESTIGACIÓN PATROCINADA POR GOBIERNO FEDERAL

[0002] El gobierno de los Estados Unidos tiene ciertos derechos sobre esta invención de acuerdo con la Subvención No. GM068664 otorgada por los National Institutes of Health y la Subvención No. DAAD19-03-0D-0004 otorgada por ARO - US Army Robert Morris Acquisition Center.

CAMPO TÉCNICO

[0003] La revelación se relaciona con la ingeniería y el diseño biomolecular y con proteínas y ácidos nucleicos obtenidas por ingeniería.

ANTECEDENTES

[0004] El desempeño de las mezclas de celulasa en procesos de conversión de biomasa depende de muchas propiedades de la enzima que incluyen estabilidad, inhibición del producto, sinergia entre diferentes componentes de la celulasa, enlace productivo frente a adsorción no productiva y dependencia del pH, además del estado físico y la composición del sustrato de celulosa. Dada la naturaleza multivariada de la hidrólisis de la celulosa, resulta deseable tener diversas celulasas de entre las cuales escoger con el fin de optimizar las formulaciones de la enzima para diferentes aplicaciones y materias primas.

RESUMEN

[0005] La revelación proporciona polipéptidos recombinantes que tienen actividad de celulasa y un aumento de la termoestabilidad y la actividad en comparación con la proteína de tipo silvestre. La revelación prevé y demuestra que las quimeras CBHII y las enzimas nativas que tienen una mutación Cys a Ser en el extremo del C terminal (p.ej., cerca del aminoácido 310-315 dependiendo de la secuencia de la proteína nativa, ver, p.ej., SEQ ID NO: 2 y 4) hidroliza más celulosa sólida que la enzima nativa en pruebas de hidrólisis a largo plazo.

[0006] La revelación proporciona un polipéptido recombinante que comprende una sustitución C→S en la región C terminal en un motivo que comprende la secuencia GECDG (SEQ ID NO: 2 de 312-316), en donde la variante comprende un aumento de la termoestabilidad y la actividad de celulasa en comparación con una celobiohidrolasa de tipo silvestre. Por ejemplo, la revelación prevé polipéptidos que tienen un aumento de la termoestabilidad y la actividad de celulasa que comprenden una secuencia que es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 2 que comprende una C314S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 4 que comprende una C311S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 12 que

comprende una C310S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 13 que comprende una C312S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 14 que comprende una C314S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 15 que comprende una C315S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 16 que comprende una C313S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 17 que comprende una C311S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 19 que comprende una C313S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 21 que comprende una C312S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 22 que comprende una C311S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 64 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 65 que comprende una C407S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 66 que comprende una C394S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 67 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 68 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 69 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 70 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 71 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 72 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 73 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 74 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 75 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 76 que comprende una C407S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 77 que comprende una C394S; o es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 78 que comprende una C412S, en donde los anteriores polipéptidos tienen actividad de celulasa y mejor termoestabilidad en comparación con su proteína progenitora correspondiente (de tipo silvestre) que carece de una mutación Cys→Ser.

[0007] La revelación también proporciona polipéptidos sustancialmente purificados que o bien se producen por técnica recombinante, por síntesis o de otra forma se generan no naturalmente en donde el polipéptido comprende una secuencia como la que se estipula más adelante que tiene de 1-10, 10-20 o 20-30 sustituciones conservadoras de aminoácido excepto en la posición que se identifica a continuación en donde está presente una sustitución C→S: SEQ ID NO: 2 que comprende una C314S; SEQ ID NO: 4 que comprende una C311S; SEQ ID NO: 12 que comprende una C310S; SEQ ID NO: 13 que comprende una C312S; SEQ ID NO: 14 que comprende una C314S; SEQ ID NO: 15 que comprende una C315S; SEQ ID NO: 16 que comprende una C313S; SEQ ID NO: 17 que comprende una C311S; SEQ ID NO: 19 que comprende una C313S; SEQ ID NO: 21 que comprende una C312S; SEQ ID NO: 22 que comprende una C311S; SEQ ID NO: 64 que comprende una C400S; SEQ ID NO: 65 que comprende una C407S; SEQ ID NO: 66 que comprende una C394S; SEQ ID NO: 67 que comprende una C400S; SEQ ID NO: 68 que comprende una C400S; SEQ ID NO: 69 que comprende una C400S; SEQ ID NO: 70 que comprende una C400S; SEQ ID NO: 71 que

6

comprende una C400S; SEQ ID NO: 72 que comprende una C400S; SEQ ID NO: 73 que comprende una C400S; SEQ ID NO: 74 que comprende una C400S; SEQ ID NO: 75 que comprende una C400S; SEQ ID NO: 76 que comprende una C407S; SEQ ID NO: 77 que comprende una C394S; o la SEQ ID NO: 78 que comprende una C412S.

[0008] La revelación proporciona un polipéptido recombinante que comprende una secuencia seleccionada de entre el grupo compuesto por: (a) un polipéptido que tiene una identidad de 85% o más con la SEQ ID NO: 2, que tiene una Ser en la posición 314 y en donde el polipéptido tiene actividad de celulasa; (b) un polipéptido que tiene una identidad de por lo menos 70% o más con la SEQ ID NO: 4, que tiene una Ser en la posición 311 y en donde el polipéptido tiene actividad de celulasa; (c) un polipéptido que tiene una identidad del 70% o más con una secuencia seleccionada de entre el grupo compuesto por: (i) SEQ ID NO: 12 y que tiene una Ser en la posición 310, (ii) SEQ ID NO: 13 y que tiene una Ser en la posición 312, (iii) SEQ ID NO: 14 y que tiene una Ser en la posición 314, (iv) SEQ ID NO: 15 y que tiene una Ser en la posición 315, (v) SEQ ID NO: 16 y que tiene una Ser en la posición 313, (vi) SEQ ID NO: 17 y que tiene una Ser en la posición 311, (vii) SEQ ID NO: 19 y que tiene una Ser en la posición 313, (viii) SEQ ID NO: 21 y que tiene una Ser en la posición 312 y (ix) SEQ ID NO: 22 y que tiene una Ser en la posición 311 y en donde cada uno de los anteriores polipéptidos tiene actividad de celulasa; y (d) un polipéptido quimérico que comprende por lo menos dos dominios de dos polipéptidos progenitores de celobiohidrolasa diferentes, en donde los dominios comprenden de N- a C- terminal: (segmento 1)-(segmento 2)-(segmento 3)-(segmento 4)-(segmento 5)-(segmento 6)-(segmento 7)-(segmento 8); en donde: el segmento 1 comprende una secuencia que es por lo menos 50-100% idéntica al residuo de aminoácidos de cerca de 1 a cerca de x_1 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 2 comprende una secuencia que es por lo menos 50-100% idéntica al residuo de aminoácidos x_1 a cerca de x_2 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 3 comprende una secuencia que es por lo menos 50-100% idéntica al residuo de aminoácidos x_2 a cerca de x_3 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 4 comprende una secuencia que es por lo menos 50-100% idéntica al residuo de aminoácidos x_3 a cerca de x_4 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 5 comprende una secuencia que es por lo menos 50-100% idéntica a cerca del residuos de aminoácidos x_4 a cerca de x_5 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 6 comprende una secuencia que es por lo menos 50-100% idéntica al residuo de aminoácidos x_5 a cerca de x_6 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 7 comprende una secuencia que es por lo menos 50-100% idéntica al residuo de aminoácidos x_6 a cerca de x_7 de la SEQ ID NO: 2 ("1") o la SEQ ID NO: 4 ("2"); y el segmento 8 comprende una secuencia que es por lo menos 50-100% idéntica al residuo de aminoácidos x_7 a cerca de x_8 de la SEQ ID NO: 2 ("1") o la SEQ ID NO: 4 ("2"); en donde x_1 es el residuo 43, 44, 45, 46 o 47 de la SEQ ID NO: 2 o el residuo 42, 43, 44, 45 o 46 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_2 es el residuo 70, 71, 72, 73 o 74 de la SEQ ID NO: 2 o el residuo 68, 69, 70, 71, 72, 73 o 74 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_3 es el residuo 113, 114, 115, 116, 117 o 118 de la SEQ ID NO: 2 o el residuo 110, 111, 112, 113, 114, 115 o 116 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_4 es el residuo

153, 154, 155, 156 o 157 de la SEQ ID NO: 2 o el residuo 149, 150, 151, 152, 153, 154, 155 o 156 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_5 es el residuo 220, 221, 222, 223 o 224 de la SEQ ID NO: 2 o el residuo 216, 217, 218, 219, 220, 221, 222 o 223 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_6 es el residuo 256, 257, 258, 259, 260 o 261 de la SEQ ID NO: 2 o el residuo 253, 254, 255, 256, 257, 258, 259 o 260 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_7 es el residuo 312, 313, 314, 315 o 316 de la SEQ ID NO: 2 o el residuo 309, 310, 311, 312, 313, 314, 315 o 318 de la SEQ ID NO: 4; y x_8 es un residuo de aminoácidos correspondiente al C terminal del polipéptido que tiene la secuencia de la SEQ ID NO: 2, SEQ ID NO: 4 o SEQ ID NO: 6, en donde el polipéptido quimérico comprende una Ser en la posición 314 de la SEQ ID NO: 2 o en la posición 311 de la SEQ ID NO: 4 y en donde el polipéptido quimérico tiene actividad de celulasa y mejor termoestabilidad y/o estabilidad de pH en comparación con un polipéptido CBH II que comprende la SEQ ID NO: 2, 4 o 6. En una modalidad del polipéptido recombinante el segmento 1 comprende un residuo de aminoácidos de cerca de 1 a cerca de x_1 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y tiene 1-10 sustituciones conservadoras de aminoácido; el segmento 2 va de cerca del residuo de aminoácidos x_1 a cerca de x_2 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y tiene cerca de 1-10 sustituciones conservadoras de aminoácido; el segmento 3 va de cerca del residuo de aminoácidos x_2 a cerca de x_3 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y tiene cerca de 1-10 sustituciones conservadoras de aminoácido; el segmento 4 va de cerca del residuo de aminoácidos x_3 a cerca de x_4 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y tiene cerca de 1-10 sustituciones conservadoras de aminoácido; el segmento 5 va de cerca del residuo de aminoácidos x_4 a cerca de x_5 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y tiene cerca de 1-10 sustituciones conservadoras de aminoácido; el segmento 6 va de cerca del residuo de aminoácidos x_5 a cerca de x_6 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y tiene cerca de 1-10 sustituciones conservadoras de aminoácido; el segmento 7 va de cerca del residuo de aminoácidos x_6 a cerca de x_7 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y tiene cerca de 1-10 sustituciones conservadoras de aminoácido; y segmento 8 va de cerca del residuo de aminoácidos x_7 a cerca de x_8 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y tiene cerca de 1-10 sustituciones conservadoras de aminoácido excepto en la posición 314 de la SEQ ID NO: 2, la posición 311 o la SEQ ID NO: 4 o 313 de la SEQ ID NO: 6. En otra modalidad, el polipéptido quimérico comprende una secuencia que es por lo menos 80%, 90%, 95%, 98% o 99% idéntica a una secuencia seleccionada de entre el grupo compuesto por la SEQ ID NO: 12-62 y 63.

[0009] La revelación también proporciona un polipéptido recombinante que consta de una secuencia como la estipulada en la SEQ ID NO: 12-62 o 63.

[0010] La revelación también proporciona un polinucleótido que codifica cualquiera de los polipéptidos arriba descritos, vectores que contienen el polinucleótido y células hospederas que contienen el polinucleótido o vector.

X

[0011] La revelación también proporciona una preparación enzimática que comprende un polipéptido de la revelación en forma sustancialmente purificada o como parte de un lisado de células.

[0012] La revelación también proporciona un método de tratamiento de una biomasa que comprende celulosa, el cual método comprende poner en contacto la biomasa con un polipéptido o preparación enzimática de la revelación.

BREVE DESCRIPCIÓN DE LAS FIGURAS

[0013] **Figura 1A-B** Gel SDS-PAGE de gen progenitor candidato CBH II sobrenadantes del cultivo de expresión de levadura. (A) Carriles de gel (Izquierda a derecha): 1-*H. jecorina*, 2-Vector vacío, 3-*H. insolens*, 4-*C. thermophilum*, 5-*H. jecorina* (duplicado), 6-*P. chrysosporium*, 7-*T. emersonii*, 8-Vector vacío (duplicado), 9-*H. jecorina* (triplicado). Los números en la parte inferior del gel representan concentración de azúcar reductor ($\mu\text{g/mL}$) presente en reacción después de prueba de hidrólisis PASC de 2 horas a 50°C . La comparación subsiguiente SDS-PAGE con estándar BSA permitió hacer la estimación del nivel de expresión de *H. insolens* de 5-10 mg/L. (B) Muestra de análisis en gel SDS-PAGE de la expresión del CBH II en sobrenadantes de cultivo de *S. cerevisiae*. Las bandas del CBH II aparecen justo por debajo del estándar de peso molecular de 60 kDa. Carriles, de izquierda a derecha, 1-H. jeco de tipo silvestre, 2-H. jeco B7P3, 3-H. jeco C311S, 4-C. ther de tipo silvestre, 5-H. inso de tipo silvestre, 6-H. inso B7P3, 7-H. inso C314S. Los números denotan μg equivalentes de glucosa/mL de volumen de reacción por mL del equivalente del sobrenadante del cultivo de expresión de SDCAA producido durante 100 minutos de incubación con PASC (1 mg/mL) a 50°C en acetato de sodio 50 mM, pH 4,8. Los valores para los carriles 1-4 han sido divididos por 2 para corregir por el doble del volumen de sobrenadante del cultivo concentrado que se está cargando en donde la omisión de esta corrección haría que los valores de actividad específica para las enzimas del *H. insolens* parecieran artificialmente bajos.

[0014] La **Figura 2A-C** muestra ilustraciones de las fronteras de bloques de la biblioteca de la quimera CBH II. (A) diagrama de la cinta de dominio catalítico del CBH II de *H. insolens* con los bloques distinguidos por colores. la enzima CBH II hace complejos con inhibidor de la glucosidasa de la isofagomina celobio-derivado. (B) representación lineal del dominio catalítico del *H. insolens* que muestra elementos de la estructura secundaria, enlaces disulfuro y divisiones en bloques denotadas por flechas negras. (C) Mapa de contacto de la cadena lateral que denota contactos (átomos pesados de la cadena lateral dentro de $4.5\text{\AA}$) que se pueden fragmentar en el momento de la recombinación. La mayoría de los contactos fragmentados tiene lugar entre bloques consecutivos.

[0015] La **Figura 3** muestra un número de contactos fragmentados (E) y número de mutaciones del progenitor más cercano (m) para juego de muestras de quimeras 23 segregadas/activas y 15 no segregadas/no activas.

[0016] La **Figura 4** muestra actividad específica, normalizada a pH 5,0, como una función del pH para las enzimas CBH II progenitoras y tres quimeras termoestables. Los datos que se presentan son promedios de dos réplicas, en donde las barras de error para HJPlus y *H. jeco* denotan valores para dos estudios independientes. 16 horas de reacción, 300 μg enzima/g

PASC, 50°C, amortiguador de citrato de sodio 12,5 mM/fosfato de sodio 12,5 mM al pH que se muestra.

[0017] La **Figura 5** muestra los resultados de la prueba de hidrólisis de la celulosa a largo plazo (μg equivalente de azúcar reductor glucosa/ μg de enzima CBH II) para progenitores y quimeras termoestables a través de un intervalo de temperaturas. Las barras de error indican errores estándar para tres réplicas de HJPlus y las enzimas CBH II de *H. insolens*. 40 horas de reacción, 100 μg de enzima/g de PASC, acetato de sodio 50 mM, pH 4,8.

[0018] La **Figura 6** muestra actividades residuales normalizadas para las quimeras del conjunto de validación después de 12 horas de incubación a 63°C. Las actividades residuales para las enzimas CBH II en sobrenadantes del cultivo concentrado determinadas en una prueba de 2 horas con PASC como sustrato, 50°C, amortiguador de acetato de sodio 25 mM, pH 4,8.

[0019] **Figura 7** Mapa para vector de expresión del progenitor y la quimera de la enzima CBH II Yep352/PGK91-1-ss. El vector ilustrado contiene el gen cel6a de *H. jecorina* de tipo silvestre (enzima CBH II). Tanto para las enzimas CBH II quiméricas como para las progenitoras, la secuencia de aminoácidos CBD/conectora que sigue al sitio ss Lys-Arg Kex2 es:

ASCSSVWGQCGGQNWSGPTCCASGSTCVYSNDYYSQCCLPGAASSSSSTRAASTTSRVSP
TSRSSATPPPGSTTTTRVPPVGS GTATYS (SEQ ID NO: 8).

[0020] La **Figura 8** muestra los valores observados y predichos de T_{50} para progenitores de CBH II y 51 quimeras de CBH II. La línea denota ecuación del modelo de regresión lineal (parámetros en la Tabla 7). Los valores de T_{50} de la CBH II progenitora se ilustran como cuadrados.

[0021] La **Figura 9A-C** muestra actividades específicas de la CBH II hacia Avicel como una función de la temperatura. (a) Actividades específicas de la CBH II progenitora y quimera. (b) Actividades específicas de CBH II progenitora, C311S mutante y quimera de sustitución de bloque único B7P3. Se llevaron a cabo reacciones durante 16 horas en acetato de sodio 50 mM, pH 4,8 con una concentración de Avicel de 15 mg/mL. (c) Actividades de CBH II progenitora, mutante de único punto y quimera de sustitución de único bloque (μg /glucosa/mL reacción) hacia Avicel como una función de la temperatura. Se llevaron a cabo reacciones durante 150 minutos en acetato de sodio 50 mM, pH 4,8 con una concentración de Avicel de 15 mg/mL. La CBH II en sobrenadantes de cultivo de levadura se dosificó para alcanzar concentraciones más o menos equivalentes del producto azúcar reductor a 55°C. Los datos que se presentan son promedios de dos réplicas independientes en donde las barras de error indican los valores de actividad por duplicado para cada punto de temperatura.

[0022] La **Figura 10** muestra alineación de múltiples secuencias ClustalW para bloque 7 de progenitor 1, *H. insolens* y progenitor 3, *C. thermophilum*. Las flechas denotan residuos cambiados en mutantes de inversión.

[0023] La **Figura 11** muestra valores de T_{50} para mutantes puntuales de la quimera 21111331. Los valores se muestran como promedio de dos duplicados independientes, las barras de error indican valores duplicados de T_{50} para cada mutante puntual. Se llevó a cabo la inactivación durante 10 minutos a la temperatura que se estaba examinando en amortiguador

de acetato de sodio 50 mM, pH 4,8. La actividad residual se determinó por incubación con 1 g/L de celulosa hinchada en ácido fosfórico (PASC) en el anterior amortiguador durante 100 minutos a 50°C.

[0024] La **Figura 12** muestra valores de T_{50} para las CBH II progenitoras de *H. insolens* y *H. jecorina*, mutantes Ser de punto único y quimeras de sustitución en bloque B7P3. Los valores se muestran como promedios de tres réplicas independientes y las barras de error indican una desviación estándar para cada CBH II. La inactivación se llevó a cabo durante 10 minutos a la temperatura que se estaba examinando en amortiguador de acetato de sodio 50 mM, pH 4,8. La actividad residual se determinó por incubación con 1 g/L de celulosa hinchada en ácido fosfórico (PASC) en el anterior amortiguador durante 100 minutos a 50°C.

[0025] La **Figura 13** muestra los valores de T_{50} para quimeras de CBH II 31311112, 13231111 y el dominio catalítico de la CBH II de tipo silvestre de *P. chrysosporium* (fusionado al CBM de *H. jecorina*) y segregadas de forma heteróloga de *S. cerevisiae*. Los valores se muestran como dos réplicas independientes con las barras de error indicando los valores para cada estudio. La inactivación se llevó a cabo durante 10 minutos a la temperatura examinada, en amortiguador de acetato de sodio 50 mM, pH 4,8. La actividad residual se determinó por incubación con 1 g/L de celulosa hinchada en ácido fosfórico (PASC) en el anterior amortiguador durante 100 minutos a 50°C.

[0026] La **Figura 14A-D** muestra interfaces del bloque de recombinación de la CBH II. (a) Los sitios inter bloques en donde los pares de residuos no parentales son posibles (esferas conectadas) suelen estar expuestos a la superficie, permitiendo potencialmente que el solvente tamice las interacciones. (b) Una interfaz de ejemplo (B5-B6) ilustra conservación de la columna vertebral (caricaturas para *H. jecorina* y *H. insolens* alineadas), residuos variables sobre la superficie y la comparativamente rara posibilidad de un par hidrofóbico enterrado novedoso en los residuos 173 y 253 (flecha). (c) Los Bloques 1-4 de *H. jecorina* (caricatura negra) se corresponden con los bloques emparentados de *H. insolens* (caricatura con código de color) sin desviaciones grandes, aunque se observa desplazamiento asociado con unión al sustrato (flecha) en parte de B3 (amarillo). (d) Los bloques emparentados 5-8 también son similares, aunque el indel en la unión B6, B7 (flecha) requerirá un cambio conformacional.

[0027] La **Figura 15** muestra un análisis estructural de la mutación C314S y su efecto estabilizador. (a) Las posiciones de hidrógeno para la estructura del *H. insolens* de alta resolución (1.0cn) fueron añadidas con REDUCE.1 (b) La geometría reconfigurada de la estructura análoga de la serina se modeló en PyMOL (<http://www.pymol.org>). La optimización de la cadena lateral en la plataforma de modelado SHARPEN2 (con una función de energía de Rosetta de todos los átomos) también sugirió que tanto la Cys314 como la Ser314 donarían enlaces de hidrógeno al carbonilo de Pro339 y aceptarían enlaces de hidrógeno de la amida de Gly316. La superior capacidad de enlace de hidrógeno de la serina puede desempeñar un papel en la mayor estabilidad de las variantes que contienen serina. Otra posible explicación es la complementariedad geométrica. Específicamente, la posición Cys de 1.0cn muestra evidencia de tensión conformacional en que la cadena lateral está notoriamente doblada (es decir, el ángulo diedro inadecuado de N-C-C α -C β está a 6° de la posición estándar), lo cual aumenta la

distancia desde el Pro carbonilo. Los números que aparecen en la figura no precedidos por letras denotan distancias de enlace de hidrógeno (Å).

DESCRIPCIÓN DETALLADA

[0028] Tal como se usan aquí y en las reivindicaciones que se anexan, las formas singulares "un", "y" y "el" incluyen referentes plurales a menos que el contexto claramente dicte otra cosa. Así pues, por ejemplo, la referencia a "un dominio" incluye una pluralidad de tales dominios y la referencia a "la proteína" incluye referencia a una o más proteínas y así sucesivamente.

[0029] También, el uso de la conjunción "o" significa "y/o" a menos que se indique lo contrario. De manera similar, "comprende" "comprenden" "que comprende" "incluyen" "incluye", y "que incluyen" son intercambiables y no tienen por objeto ser limitantes.

[0030] También se entenderá que en donde las descripciones de diversas modalidades usan el término "que comprende", la persona medianamente versada en la técnica entenderá que en algunos casos específicos, una modalidad puede ser descrita como alternativa usando expresiones como "compuesto/a esencialmente por" o "compuesto/a por."

[0031] Aunque en la práctica de los métodos y composiciones reveladas se pueden usar métodos y materiales similares o equivalentes a los aquí descritos, los métodos, dispositivos y materiales ejemplares se describen aquí.

[0032] A menos que se defina otra cosa, todos los términos científicos y técnicos utilizados aquí tienen el significado que entiende la persona medianamente versada en la materia técnica a la cual pertenece esta revelación. Así las cosas, tal como se usa a lo largo de la solicitud en cuestión, los siguientes términos tendrán los siguientes significados.

[0033] En estudios recientes se ha demostrado el superior desempeño de celulasas de hongos termofílicos con respecto a sus contrapartes mesofílicas en procesos de conversión de biomasa a escala de laboratorio, en donde una mejora de la estabilidad conduce a la retención de la actividad durante periodos más prolongados a temperaturas tanto moderadas como elevadas. Las celulasas de los hongos resultan atractivas porque son altamente activas y se pueden expresar en hongos hospederos como *Hipocrea jecorina* (*Trichoderma reesei* anamórfica) en niveles de hasta 40 g/L en el sobrenadante. Por desgracia, el conjunto de celulasas termoestables documentadas de los hongos es reducido. En el caso de las enzimas celobiohidrolasas procesivas de clase II (CBH II), menos de 10 secuencias de genes termoestables naturales están anotadas en la base de datos CAZy.

[0034] Como se describe en mayor detalle en este documento, con el uso de generación y análisis recursivos de polipéptidos quiméricos, dominios estabilizadores particulares y en últimas de aminoácidos específicos, se identificaron la termoestabilidad impartida y la actividad mejorada.

[0035] Como se describirá en mayor detalle más adelante, la invención está basada, por lo menos en parte, en la generación y expresión de enzimas novedosas que catalizan la hidrólisis de la celulosa. En una modalidad, se proporcionan polipéptidos novedosos modificados por ingeniería genética para que hidrolicen la celulosa a altas temperaturas. Dichos polipéptidos incluyen variantes de la celobiohidrolasa que han sido alterados para que incluyan sustituciones de aminoácidos en residuos específicos. Aunque estas variantes serán descritas

en mayor detalle más adelante, se entiende que los polipéptidos de la revelación pueden contener uno o más aminoácidos modificados. La presencia de aminoácidos modificados puede resultar provechosa, por ejemplo, (a) aumentando la vida media, (b) la termoestabilidad y (c) el recambio de sustrato de un polipéptido. El o los aminoácidos son modificados, por ejemplo, de forma co-traslacional o post-traslacional durante la producción recombinante (p.ej., glucosilación ligada al N en motivos N--X--S/T durante la expresión en células de mamíferos) o modificados por medios sintéticos. Según esto, una proteína, enzima, polinucleótido, gen o célula "mutante", "variante" o "modificada", significa una proteína, enzima, polinucleótido, gen o célula, que han sido alteradas o derivadas o que de alguna manera son diferentes o han cambiado, con respecto a una proteína, enzima, polinucleótido, gen o célula progenitora. Una proteína o enzima mutante o modificada suele ser, aunque no necesariamente es así, expresada a partir de un polinucleótido o gen mutante.

[0036] Una "mutación" significa cualquier proceso o mecanismo que se traduce en una proteína, enzima, polinucleótido, gen o célula mutante. Esto incluye cualquier mutación en la cual una proteína, enzima, polinucleótido o secuencia genética resulta alterada y cualquier cambio detectable en una célula que surja de tal mutación. Lo típico es que una mutación se presente en un polinucleótido o secuencia genética debido a mutaciones puntuales, deleciones o inserciones de residuos de nucleótidos únicos o múltiples. Una mutación incluye alteraciones de polinucleótidos que surgen dentro de una región codificadora de proteína de un gen así como alteraciones en regiones por fuera de una secuencia codificadora de proteína, como, pero sin limitaciones, secuencias reguladoras o promotoras. Una mutación en un gen puede ser "silente", es decir, no reflejada en una alteración de aminoácido en el momento de la expresión, lo cual conduce a una variante "conservadora de secuencia" del gen. Esto por lo general surge cuando un aminoácido corresponde a más de un codón.

[0037] Algunos ejemplos no limitantes de un aminoácido modificado incluyen un aminoácido glucosilado, un aminoácido sulfatado, un aminoácido prenilado (p.ej., farnesilado, geranilgeranilado), un aminoácido acetilado, un aminoácido acilado, un aminoácido pegilado, un aminoácido biotinilado, un aminoácido carboxilado, un aminoácido fosforilado y similares. La bibliografía está plena de referencias adecuadas para orientar a la persona versada en la técnica en la modificación de los aminoácidos. Algunos ejemplos de protocolos se encuentran en Walker (1998) Protein Protocols en CD-ROM (Humana Press, Towata, N.J.).

[0038] Los métodos recombinantes para producir y aislar los polipéptidos modificados de celobiohidrolasa de la revelación se describen aquí. Además de la producción recombinante, los polipéptidos se pueden producir mediante por síntesis directa de los péptidos usando técnicas de fase sólida (p.ej., Stewart et al. (1969) Solid Phase Peptide Synthesis (WH Freeman Co, San Francisco); y Merrifield (1963) J. Am. Chem. Soc. 85: 2149-2154). La síntesis de los péptidos se puede llevar a cabo usando técnicas manuales o automatizadas. La síntesis automatizada se puede lograr, por ejemplo, usando el Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer, Foster City, Calif.) de conformidad con las instrucciones proporcionadas por el fabricante.

[0039] Los términos “celobiohidrolasa II” o “enzima CBH II” significan una enzima de las 6 proteínas de la familia de la celulasa, que están ampliamente distribuidas en bacterias y hongos. Las enzimas están involucradas en la hidrólisis de la celulosa.

[0040] Por “actividad de celulasa” se entiende una enzima que es capaz de hidrolizar la celulosa. La celulasa se refiere a una clase de enzimas producidas por hongos, bacterias y protozoarios que catalizan la hidrólisis de la celulosa. Sin embargo, también existen celulasas producidas por otros tipos de organismos como plantas y animales. El número EC para este grupo de enzimas es EC 3.2.1.4. Existen cinco tipos generales de celulasas con base en el tipo de reacción catalizada: endo-celulasa; exo-celulasa, dentro de esta categoría hay dos tipos principales de exo-celulasas (o celobiohidrolasas, abreviadas como CBH) – un tipo funciona de forma procesiva a partir del extremo reductor y un tipo funciona de forma procesiva a partir del extremo no reductor de la celulosa; hidrolasas celobiasa o beta-glucosidasa; celulasas oxidativas; y fosforilasas de celulosa que despolimerizan la celulosa usando fosfatos en vez de agua. La mayoría de las celulasas de los hongos tiene dos dominios: un dominio catalítico y un dominio de unión a la celulosa, que están conectados por un conector flexible. En modalidades específicas de la revelación la actividad de celulasa es una actividad de CBH. Las secuencias aquí descritas incluyen, en algunos casos, tanto el dominio de unión a la celulosa como el dominio catalítico o solamente el dominio catalítico. En tales casos en donde solamente se da la secuencia del dominio catalítico, se reconocerá que un dominio de unión a la celulosa (CBD) como el proporcionado en la SEQ ID NO: 8, puede estar eslabonado funcionalmente (bien sea como parte de la secuencia codificadora o fusionado más tarde) al dominio catalítico ya sea directamente o a través de un conector.

[0041] Una “proteína” o “polipéptido”, términos que se usan de forma intercambiable aquí, comprende una o más cadenas de bloques fundamentales químicos llamados aminoácidos que están conectados entre sí por enlaces químicos llamados enlaces peptídicos. Una “enzima” significa cualquier sustancia, de preferencia compuesta en todo o en su mayor parte por proteína, que cataliza o promueve, más o menos específicamente, una o más reacciones químicas o bioquímicas. Una proteína, enzima, polinucleótido, gen o célula “nativa” o “de tipo silvestre”, significa una proteína, enzima, polinucleótido, gen o célula que se encuentra en estado natural.

[0042] Una “secuencia de aminoácidos” es un polímero de aminoácidos (una proteína, polipéptido, etc.) o una sarta de caracteres que representa un polímero de aminoácidos, dependiendo del contexto. Los términos “proteína”, “polipéptido”, y “péptido” se usan de forma intercambiable en este documento. “Aminoácido” es una molécula que tiene la estructura en la cual un átomo de carbono central está enlazado a un átomo de hidrógeno, un grupo de ácido carboxílico (cuyo átomo de carbono se menciona aquí como un “átomo de carbono carboxilo”), un grupo amino (cuyo átomo de nitrógeno se menciona aquí como un “átomo de nitrógeno amino”) y un grupo de cadena lateral, R. Cuando se incorpora dentro de un péptido, polipéptido o proteína, un aminoácido pierde uno o más átomos de sus grupos carboxílicos aminoácidos en la reacción de deshidratación que enlaza un aminoácido a otro. Como resultado, cuando se incorpora dentro de una proteína, un aminoácido se conoce como un “residuo de aminoácido.”



[0043] Una secuencia particular de aminoácidos de una proteína dada (es decir, la "estructura primaria" del polipéptido cuando se escribe desde el amino-terminal hacia el carboxi-terminal) está determinada por la secuencia de nucleótidos de la porción codificadora de un ARNm, el cual a su vez es especificado por información genética, típicamente ADN genómico (incluido ADN de los organelos, *p.ej.*, ADN mitocondrial o del cloroplasto). Así, determinar la secuencia de un gen ayuda a predecir la secuencia primaria de un polipéptido correspondiente y más particularmente el papel o la actividad que desempeñan el polipéptido o las proteínas codificadas por ese gen o secuencias de polinucleótidos.

[0044] "Sustitución conservadora de aminoácidos" o, más sencillamente, "variaciones conservadoras" de una secuencia en particular quiere decir el reemplazo de un aminoácido o serie de aminoácidos, con secuencias de aminoácidos esencialmente idénticas. La persona versada en la técnica reconocerá que las sustituciones, deleciones o adiciones individuales que alteren, adicionen o eliminen un solo aminoácido o un porcentaje de aminoácidos en una secuencia codificada dan lugar a "variaciones conservadoras" en las cuales las alteraciones se traducen en la delección de un aminoácido, la adición de un aminoácido o la sustitución de un aminoácido por un aminoácido químicamente similar.

[0045] Las tablas de sustitución conservadora que proporcionan aminoácidos funcionalmente similares son bien conocidas del estado de la técnica. Por ejemplo, un grupo de sustitución conservadora incluye Alanina (A), Serina (S) y Treonina (T). Otro grupo de sustitución conservadora incluye Ácido aspártico (D) y Ácido glutámico (E). Otro grupo de sustitución conservadora incluye Asparagina (N) y Glutamina (Q). Otro grupo de sustitución conservadora incluye Arginina (R) y Lisina (K). Otro grupo de sustitución conservadora incluye Isoleucina, (I) Leucina (L), Metionina (M) y Valina (V). Otro grupo de sustitución conservadora incluye Fenilalanina (F), Tirosina (Y) y Triptófano (W).

[0046] Así pues, las "sustituciones conservadoras de aminoácido" de una secuencia de polipéptidos registrada (*p.ej.*, SEQ ID Nos: 2, 4, 6 y 12-78) incluyen sustituciones de un porcentaje, típicamente menos de 10%, de los aminoácidos de la secuencia del polipéptido, con un aminoácido seleccionado conservadoramente del mismo grupo de sustitución conservadora. Según esto, una variación sustituida conservadoramente de un polipéptido de la revelación puede contener 100, 75, 50, 25 o 10 sustituciones con una variación sustituida conservadoramente del mismo grupo de sustitución conservadora.

[0047] Se entiende que la adición de secuencias que no alteran la actividad codificada de una molécula de ácido nucleico, como la adición de una secuencia no funcional o no codificadora, es una variación conservadora del ácido nucleico básico. La "actividad" de una enzima es una medida de su capacidad de catalizar una reacción, es decir, una "función" y puede expresarse como la tasa a la cual se produce el fruto de la reacción. Por ejemplo, la actividad enzimática puede ser representada como la cantidad de producto producida por unidad de tiempo o por unidad de enzima (*p.ej.*, concentración o peso) o en términos de las constantes de afinidad o de disociación. Como se usa de forma intercambiable en este documento una "actividad de celobiohidrolasa o actividad de celulasa", "actividad biológica de celobiohidrolasa o celulasa" o "actividad funcional de celobiohidrolasa o celulasa", hace referencia a una actividad ejercida por una proteína, polipéptido que tiene actividad de celulasa

14

y en modalidades específicas actividad de celobiohidrolasa sobre un sustrato de celulosa, según se determine *in vivo* o *in vitro*, de acuerdo con técnicas estándar.

[0048] La persona versada en la técnica apreciará que muchas variaciones conservadoras de los constructos de ácido nucleico que se revelan arrojan un constructo funcionalmente idéntico. Por ejemplo, como se dijo arriba, debido a la degeneración del código genético, las "sustituciones silentes" (es decir, las sustituciones en una secuencia de ácidos nucleicos que no se traduce en una alteración de un polipéptido codificado) son una característica implícita de todas las secuencias de ácidos nucleicos que codifican un aminoácido. De manera similar, las "sustituciones conservadoras de aminoácido" en que uno o unos pocos aminoácidos en una secuencia de aminoácidos son sustituidos por diferentes aminoácidos con propiedades muy similares, también se identifican fácilmente como muy similares a un constructo revelado. Dichas variaciones conservadoras de cada secuencia revelada son una característica de los polipéptidos previstos aquí.

[0049] Las "variantes conservadoras" son proteínas o enzimas en las cuales un residuo de aminoácidos dado ha sido cambiado sin alterar la conformación y función generales de la proteína o enzima, incluidos, pero sin limitaciones, el reemplazo de un aminoácido por uno que tiene propiedades similares, incluidos carácter polar o no polar, tamaño, forma y carga. Los aminoácidos diferentes de los indicados como conservados pueden diferir en una proteína o enzima de manera que el porcentaje de similitud en la proteína o secuencia de aminoácidos entre dos proteínas cualesquiera de función similar puede variar y puede ser, por ejemplo, por lo menos de 30%, por lo menos de 50%, por lo menos de 70%, por lo menos de 80% o por lo menos de 90%, según se determine de acuerdo con un esquema de alineación. Tal como se menciona aquí, "similitud de la secuencia" significa el grado hasta el cual las secuencias de nucleótidos o proteínas están emparentadas. El grado de similitud entre dos secuencias puede estar basado en porcentaje de identidad y/o conservación de la secuencia. "Identidad de la secuencia" aquí significa el grado hasta el cual dos nucleótido o secuencias de aminoácidos son invariables. "Alineación de la secuencia" significa el proceso de alinear dos o más secuencias para alcanzar niveles máximos de identidad (y, en el caso de secuencias de aminoácidos, conservación) con el propósito de valorar el grado de similitud. Hay numerosos métodos para alinear secuencias y valorar similitud/identidad conocidos de la técnica como, por ejemplo, el Método de cúmulo, en donde la similitud se basa en el algoritmo MEGALIGN, así como BLASTN, BLASTP y FASTA (Lipman y Pearson, 1985; Pearson y Lipman, 1988). Cuando se usan todos estos programas, los escenarios preferidos son aquellos que se traducen en la máxima similitud de la secuencia.

[0050] Las modificaciones no conservadoras de un polipéptido en particular son aquellas que sustituyen cualquier aminoácido no caracterizado como una sustitución conservadora. Por ejemplo, cualquier sustitución que atravesase las fronteras de los seis grupos explicados arriba. Estos incluyen sustituciones de aminoácidos básicos o ácidos por aminoácidos neutros, (p.ej., Asp, Glu, Asn o Gln por Val, Ile, Leu o Met), aminoácidos aromáticos por aminoácidos básicos o ácidos (p.ej., Phe, Tyr o Trp para Asp, Asn, Glu o Gln) o cualquier otra sustitución que no reemplace un aminoácido por un aminoácido similar. Las cadenas laterales básicas incluyen lisina (K), arginina (R), histidina (H); cadenas laterales ácidas incluyen ácido aspártico (D),

ácido glutámico (E); cadenas laterales polares no cargadas incluyen glicina (G), asparagina(N), glutamina (Q), serina (S), treonina (T), tirosina (Y), cisteína (C); cadenas laterales no polares incluyen alanina (A), valina (V), leucina (L), isoleucina (I), prolina (P), fenilalanina (F), metionina (M), triptófano (W); las cadenas laterales beta-ramificadas incluyen treonina (T), valina (V), isoleucina (I); las cadenas laterales aromáticas incluyen tirosina (Y), fenilalanina (F), triptófano (W), histidina (H).

[0051] Una proteína, enzima, polinucleótido, gen o célula "progenitora" es cualquier proteína, enzima, polinucleótido, gen o célula, a partir de las cuales cualquier otra proteína, enzima, polinucleótido, gen o célula, es derivado o hecho, usando cualquier método, herramienta o técnica y si o no el progenitor es en sí mismo nativo o mutante. Un polinucleótido o gen progenitor codifica una proteína o enzima progenitora.

[0052] Además de proporcionar variantes de polipéptidos CBH II, se proporcionan los polipéptidos quiméricos que comprenden: 1) una variante de dominio aislada de una primera hebra parental y modificada para incluir una sustitución de aminoácido; y 2) un dominio aislado de una segunda hebra parental bien sea no modificada o modificada para incluir una nueva actividad o una actividad que complementa el dominio. Aquí se revelan métodos para obtener por ingeniería genética un polipéptido quimérico de la revelación.

[0053] La revelación proporciona variantes, mutantes y quimeras de celulasa y celobiohidrolasa (CBH) II que tienen termoestabilidad aumentada en comparación con una proteína de tipo silvestre o parental, en donde la proteína de tipo silvestre está compuesta por la SEQ ID NO: 2, 4 o 6. La variante comprende una Serina en la región C terminal en un motivo que comprende la secuencia GEXDG, en donde X es C, A o G (SEQ ID NO: 107), en donde X está sustituido con Serina, la variante que comprende actividad de celulasa y en donde el polipéptido tiene termoestabilidad aumentada en comparación con una celulasa de tipo silvestre que carece de una serina en la secuencia GEXDG (SEQ ID NO: 107). En una modalidad, las variantes comprenden por lo menos una mutación de una Cys→Ser en el motivo GECDG (ver, p.ej., SEQ ID NO: 2 de aminoácido 312-316) encontrada en la mayoría de las proteínas de celulasa y celobiohidrolasa II (como se describe en mayor detalle más adelante) y pueden comprender mutaciones adicionales que mejoran la termoestabilidad o la actividad. La identidad entre las celulasas puede ser bastante baja. La sustitución de serina arriba descrita es aplicable a cualquier celulasa que tiene el motivo de la SEQ ID NO: 107 (p.ej., en donde el polipéptido tiene una identidad de por lo menos 60% o más con la SEQ ID NO: 2 o 4).

[0054] Por ejemplo, la revelación proporciona polipéptidos que tienen un aumento de la termoestabilidad y la actividad de celulasa que comprenden una secuencia que es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 2 que comprende una C314S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 4 que comprende una C311S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 12 que comprende una C310S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 13 que comprende una C312S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 14 que comprende una C314S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 15 que comprende una C315S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 16 que comprende una C313S; es por lo menos 85%, 90%, 95%,

16

98%, 99% idéntica a la SEQ ID NO: 17 que comprende una C311S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 19 que comprende una C313S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 21 que comprende una C312S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 22 que comprende una C311S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 64 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 65 que comprende una C407S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 66 que comprende una C394S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 67 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 68 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 69 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 70 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 71 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 72 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 73 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 74 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 75 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 76 que comprende una C407S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 77 que comprende una C394S; o es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 78 que comprende una C412S, en donde los anteriores polipéptidos tienen actividad de celulasa y mejor termoestabilidad en comparación con su proteína progenitora correspondiente (de tipo silvestre) que carece de una mutación Cys→Ser.

[0055] En otra modalidad, la revelación proporciona los polipéptidos arriba descritos; sin embargo, estos además comprenden por lo menos una mutación de adición que puede ser determinada por alineación una SEQ ID NO: 64, en donde la SEQ ID NO: 64 comprende una Pro en la posición 413 o una Ser o Thr en la posición 231 o una Ser o Thr en la posición 305 o una Gln o Asn en la posición 410 o una Glu en la posición 82 o cualquier combinación de las anteriores. Se pueden identificar sustituciones similares por alineación de la secuencia de la secuencia de aminoácidos de la SEQ ID NO: 64 con las de la SEQ ID Nos: 2, 4, 6, 12-63 y 65-78.

[0056] La revelación también proporciona polipéptidos sustancialmente purificados que son producidos bien sea por técnica recombinante, sintéticos o generados de otra forma no natural en donde el polipéptido comprende una secuencia como la que se estipula más adelante que tiene de 1-10, 10-20 o 20-30 sustituciones conservadoras de aminoácido excepto en la posición que se identifica a continuación en donde está presente una sustitución C→S:

SEQ ID NO: 2 que comprende una C314S;
 SEQ ID NO: 4 que comprende una C311S;
 SEQ ID NO: 12 que comprende una C310S;
 SEQ ID NO: 13 que comprende una C312S;
 SEQ ID NO: 14 que comprende una C314S;
 SEQ ID NO: 15 que comprende una C315S;
 SEQ ID NO: 16 que comprende una C313S;
 SEQ ID NO: 17 que comprende una C311S;

SEQ ID NO: 19 que comprende una C313S;
 SEQ ID NO: 21 que comprende una C312S;
 SEQ ID NO: 22 que comprende una C311S;
 SEQ ID NO: 64 que comprende una C400S;
 SEQ ID NO: 65 que comprende una C407S;
 SEQ ID NO: 66 que comprende una C394S;
 SEQ ID NO: 67 que comprende una C400S;
 SEQ ID NO: 68 que comprende una C400S;
 SEQ ID NO: 69 que comprende una C400S;
 SEQ ID NO: 70 que comprende una C400S;
 SEQ ID NO: 71 que comprende una C400S;
 SEQ ID NO: 72 que comprende una C400S;
 SEQ ID NO: 73 que comprende una C400S;
 SEQ ID NO: 74 que comprende una C400S;
 SEQ ID NO: 75 que comprende una C400S;
 SEQ ID NO: 76 que comprende una C407S;
 SEQ ID NO: 77 que comprende una C394S; o
 SEQ ID NO: 78 que comprende una C412S.

[0057] "Polipéptido aislado" hace referencia a un polipéptido que está separado de otros contaminantes que naturalmente lo acompañan, *p.ej.*, proteína, lípidos y polinucleótidos. El término abarca polipéptidos que han sido removidos o purificados de su ambiente o sistema de expresión natural (*p.ej.*, célula hospedera o síntesis *in vitro*).

[0058] "Polipéptido sustancialmente puro" hace referencia a una composición en la cual la especie de polipéptido es la especie predominante presente (es decir, con base molar o en peso es más abundante que cualquier otra especie macromolecular individual en la composición) y es generalmente una composición sustancialmente purificada cuando la especie en cuestión comprende por lo menos cerca de 50 por ciento de la especie macromolecular presente por mol o % en peso. Generalmente, una composición de polipéptido sustancialmente puro comprenderá cerca de 60 % o más, cerca de 70% o más, cerca de 80% o más, cerca de 90% o más, cerca de 95% o más y cerca de 98% o más de toda la especie macromolecular por mol o % en peso presente en la composición. En algunas modalidades, la especie en cuestión está purificada hasta la homogeneidad esencial (es decir, no es posible detectar una especie contaminante en la composición por los métodos convencionales de detección) en donde la composición consiste esencialmente en una única especie macromolecular. La especie de solvente, moléculas pequeñas (<500 Daltons) y especies de iones elementales no se consideran como especies macromoleculares.

[0059] "Secuencia de referencia" hace referencia a una secuencia definida que se usa como base de comparación con la secuencia. Una secuencia de referencia puede ser un subconjunto de una secuencia más grande como, por ejemplo, un segmento de un gen de longitud total o la secuencia del polipéptido. Generalmente, una secuencia de referencia puede tener por lo menos 20 nucleótidos o residuos de aminoácidos de longitud, por lo menos 25 nucleótidos o residuos de longitud, por lo menos 50 nucleótidos o residuos de longitud o la longitud completa del ácido nucleico o polipéptido. Dado que cada uno de dos polinucleótidos o polipéptidos puede (1) comprender una secuencia (es decir, una porción de la secuencia completa) que es similar entre las dos secuencias y (2) además puede comprender una secuencia que es divergente entre las dos secuencias, las comparaciones de secuencia entre dos (o más) polinucleótidos o polipéptidos típicamente se suelen realizar comparando

secuencias de los dos polinucleótidos o polipéptidos durante una "ventana de comparación" para identificar y comparar regiones locales de similitud de la secuencia.

[0060] "Identidad de la secuencia" significa que dos secuencias de aminoácidos son sustancialmente idénticas (es decir, aminoácido a aminoácido) sobre una ventana de comparación. El término "similitud de la secuencia" hace referencia a aminoácidos similares que comparten las mismas características biofísicas. El término "porcentaje de identidad de la secuencia" o "porcentaje de similitud de la secuencia" se calcula comparando dos secuencias óptimamente alineadas sobre una ventana de comparación, determinando el número de posiciones en las cuales los residuos idénticos (o residuos similares) se presentan en las dos secuencias del polipéptido para arrojar el número de posiciones pareadas, dividiendo el número de posiciones pareadas por el número total de posiciones en la ventana de comparación (es decir, el tamaño de la ventana) y multiplicando el resultado por 100 para arrojar el porcentaje de identidad de la secuencia (o el porcentaje de similitud de la secuencia). En lo que respecta a una secuencia de polinucleótidos, los términos identidad de la secuencia y similitud de la secuencia tienen significado comparable al descrito para las secuencias de proteínas, en donde el término "porcentaje de identidad de la secuencia" indica que dos secuencias de polinucleótidos son idénticas (nucleótido por nucleótido) sobre una ventana de comparación. En tal sentido, también se puede calcular un porcentaje de identidad de las secuencias de polinucleótidos (o un porcentaje de similitud de las secuencias de polinucleótidos, *p.ej.*, para sustituciones silentes u otras sustituciones, con base en el algoritmo de análisis). La correspondencia máxima se puede determinar usando uno de los algoritmos de secuencia aquí descritos (u otros algoritmos disponibles para la persona medianamente versada en la materia técnica) o por inspección visual.

[0061] Igual que se aplica a los polipéptidos, el término identidad sustancial o similitud sustancial significa que dos secuencias de péptidos, cuando están óptimamente alineadas, como cuando se usan los programas BLAST, GAP o BESTFIT con pesos de brecha por defecto o por inspección visual, comparten la identidad de la secuencia o la similitud de la secuencia. De manera similar, igual que se aplica en el contexto de dos ácidos nucleicos, el término identidad sustancial o similitud sustancial significa que las dos secuencias de ácidos nucleicos, cuando se encuentran óptimamente alineadas, como cuando se usan los programas BLAST, GAP o BESTFIT usando pesos de brecha por defecto (que se describen en otro aparte de este documento) o por inspección visual, comparten la identidad de la secuencia o la similitud de la secuencia.

[0062] Un ejemplo de un algoritmo que es apropiado para determinar el porcentaje de identidad de la secuencia o de similitud de la secuencia es el algoritmo FASTA, que está descrito en Pearson, W. R. & Lipman, D. J., (1988) Proc. Natl. Acad. Sci. USA 85:2444. Ver también, W. R. Pearson, (1996) Methods Enzymology 266:227-258. Los parámetros preferidos empleados en una alineación con FASTA de las secuencias de ADN para calcular el porcentaje de identidad o el porcentaje de similitud son optimizados, BL50 Matriz 15: -5, k-tuple=2; penalidad de unión=40, optimización=28; penalidad de brecha -12, penalidad de longitud brecha=-2; y ancho=16.

[0063] Otro ejemplo de un algoritmo útil es el PILEUP. El PILEUP crea una alineación de múltiples secuencias a partir de un grupo de secuencias relacionadas usando alineaciones progresivas, pareadas para mostrar relación y porcentaje de identidad de la secuencia o porcentaje de similitud de la secuencia. También traza un árbol o dendograma que muestra las relaciones de aglomeración que se usan para crear la alineación. En el PILEUP se usa una simplificación del método de alineación progresiva de Feng & Doolittle, (1987) J. Mol. Evol. 35:351-360. El método empleado es similar al método descrito por Higgins & Sharp, CABIOS 5:151-153, 1989. El programa puede alinear hasta 300 secuencias, cada una con una longitud máxima de 5,000 nucleótidos o aminoácidos. El procedimiento de alineación múltiple comienza con la alineación pareada de las dos secuencias más similares, produciendo un aglomerado de dos secuencias alineadas. Este aglomerado entonces se alinea con la siguiente secuencia o aglomerado de secuencias alineadas más emparentados. Dos aglomerados de secuencias se alinean por una simple extensión de la alineación pareada de dos secuencias individuales. La alineación final se logra por una serie de alineaciones progresivas, pareadas. El programa se corre designando secuencias específicas y sus coordenados de aminoácido o nucleótido para la comparación de las regiones de secuencia y designando los parámetros del programa. Con el uso del PILEUP, una secuencia de referencia se pone en comparación con otras secuencias de prueba para determinar la relación de porcentaje de identidad de la secuencia (o el porcentaje de similitud de la secuencia) usando los siguientes parámetros: peso de la brecha por defecto (3.00), peso de la longitud de la brecha por defecto (0.10) y brechas terminales ponderadas. El PILEUP se puede obtener en el paquete de software de análisis de secuencias GCG, *p.ej.*, versión 7.0 (Devereaux *et al.*, (1984) Nuc. Acids Res. 12:387-395).

[0064] Otro ejemplo de un algoritmo que es apropiado para múltiples alineaciones de ADN y secuencias de aminoácidos es el programa CLUSTALW (Thompson, J. D. *et al.*, (1994) Nuc. Acids Res. 22:4673-4680). El CLUSTALW realiza múltiples comparaciones pareadas entre grupos de secuencias y las ensambla en una alineación múltiple con base en la identidad de la secuencia. Las penalidades de Brecha abierta y extensión de Brecha fueron de 10 y 0,05 respectivamente. Para las alineaciones de aminoácidos se puede usar el algoritmo BLOSUM como una matriz de peso de proteína (Henikoff y Henikoff, (1992) Proc. Natl. Acad. Sci. USA 89:10915-10919).

[0065] Como se mencionó arriba, los miembros de la familia de la celobiohidrolasa y celulasa se pueden identificar por alineación de la secuencia y una sustitución en el motivo GECDG (ver, *p.ej.*, SEQ ID NO: 2 del aminoácido 312-316). Entonces el polipéptido modificado se puede someter a examen para determinar su actividad como se describe más adelante a diversas temperaturas y condiciones para identificar aquellas modificaciones que introducen una actividad favorable. Algunas secuencias a manera de ejemplo se pueden encontrar en los siguientes números de acceso al GenBank, cuyas secuencias se incorporan a la presente mediante referencia.

29

P07987
GUX2_TRIRE
Q9HEY8
Q9HEY8_TRIRE
Q7LSP2
Q7LSP2_TRIKO
Q6UJX9
Q6UJX9_TRIVI
A3QVU7
A3QVU7_TRIVI
1HCL5
Q1HCL5_TRIKO
Q66PN1
Q66PN1_9HYPO
B5TWC7
B5TWC7_9HYPO
Q9C1S9
GUX6_HUMIN
Q2GMP2
Q2GMP2_CHAGB
A7E6G7
A7E6G7_SCLS1
Q0UPA5
Q0UPA5_PHANO
A6S7A6
A6S7A6_BOTFB
P49075
GUX3_AGABI
Q02321
Q02321_PHACH
Q9C1R4
Q9C1R4_LENED
Q96VU2
Q96VU2_LENED
B2ABX7
B2ABX7_PODAN
A4RPH6
A4RPH6_MAGGR

Q6E5B1
Q6E5B1_9AGAR
B7X9Z0
B7X9Z0_COPCI
A8NEJ3
A8NEJ3_COPC7
Q96V98
Q96V98_ORPSP
Q7Z7X6
Q7Z7X6_PIREQ
Q870B2
Q870B2_9FUNG
Q874E1
Q874E1_ORPSP
A9FHT2
A9FHT2_SORCS
B0FEV9
B0FEV9_9FUNG
Q6EY63
Q6EY63_9FUNG
Q6EH22
Q6EH22_NEOFR
B6EA50
B6EA50_NEOPA
B0FEV4
B0FEV4_NEOPA
6E1Y8
Q6E1Y8_NEOFR
Q9UW10
Q9UW10_9FUNG
Q12646
Q12646_NEOPA
Q6A4K7
Q6A4K7_9FUNG
Q9UW11
Q9UW11_9FUNG
Q9P8Q8
Q9P8Q8_9FUNG
B0FEV8
B0FEV8_9FUNG

Todavía en otras modalidades, la familia de la variante del polipéptido de celulasa que tiene mejor termoestabilidad incluye aquellos integrantes que aparecen en la siguiente Tabla que tienen una sustitución C→S, G→S o A→S. Además, los polipéptidos que tienen 85%, 90%, 95%, 98% o 99% de identidad de la secuencia con cualquiera de las siguientes secuencias que tienen las sustituciones identificadas en la tabla que aparece a continuación, que tienen actividad de celulasa y termoestabilidad también están abarcadas por la revelación.

La alineación del marco de aminoácidos encierra la CBH II Cys311 de *H. jecorina* para las secuencias de proteínas que tienen la más alta identidad con la CBH II de *H. jecorina*. Se muestran los residuos en la posición equivalente de 311 denotados con negrita y subrayados. Las secuencias para las CBH II recombinantes de *H. insolens* y *P. chrys* estudiadas en este trabajo se denotan como *H. inso* y *P. chrys*. Se excluyeron 54 de las 250 secuencias más idénticas debido a redundancia (es decir, mutantes puntuales para estudios estructurales o isoformas con identidad >95%). Se identifica el número de acceso para la celulasa y la secuencia correspondiente se incorpora a la presente mediante referencia como si se copiara directamente del número de acceso. Las secuencias asociadas con el número de accesos se

✓

denominan SEQ ID NO: 79-106. Un reemplazo de del residuo en **negrita y subrayado** (p.ej., C, A o G) con S. El número entre paréntesis que sigue a la secuencia identifica la SEQ ID NO:)

<i>H.jeco</i>	---T---G---D---S---L---LDSFVWVKPGGE <u>CDG</u> ---T---S-----	(4)
XP_001903170	---T---G---L---D---I---EDAFVWIKPGGE <u>CDG</u> ---T---S-----	(79)
XP_001226029	---T---G---H---D---L---LDAFVWIKPGGE <u>CDG</u> ---T---S-----	(80)
XP_360146	---T---G---S---E---L---ADAFVWIKPGGE <u>CDG</u> ---V---S-----	(81)
<i>H.inso</i>	---T---G---H---Q---Y---VDAFVWVKPGGE <u>CDG</u> ---T---S-----	(2)
XP_001598803	---T---G---D---A---L---EDAFVWVKPGGE <u>ADG</u> ---T---S-----	(82)
XP_001796781	---T---D---D---P---L---LDAYVWVKPGGE <u>G</u> DG---T---S-----	(83)
AAA50608	---T---G---S---S---L---IDAIWVVKPGGE <u>CDG</u> ---T---S-----	(84)
AAK28357	---T---G---S---S---L---IDSIVWVKPGGE <u>CDG</u> ---T---S-----	(85)
BAH59082	---T---G---S---P---L---IDSIVWVKPGGE <u>CDG</u> ---T---S-----	(86)
AAT64008	---T---G---S---S---L---IDAIWVVKPGGE <u>CDG</u> ---T---T-----	(87)
<i>P.crys</i>	---T---G---S---Q---F---IDSIVWVKPGGE <u>CDG</u> ---T---S-----	(12)
BAH59083	---T---P---S---S---L---IDSIVWVKPGGE <u>ADG</u> ---T---S-----	(88)
XP_001833045	---T---P---S---S---A---IDAIWVVKPGGE <u>ADG</u> ---T---S-----	(89)
XP_002391276	---T---G---S---S---L---IDSIVWVKPGGE-----	(90)
AAD51055	---P---D---S---SKP-L-LDAYMWIKTPGE <u>ADG</u> ---S---S-----	(91)
ABY52798	---S---G---Y---P---L---LDAFMWLKTPGE <u>ADG</u> ---S---A-----	(92)
AAF34679	---P---D---A---SMP-L-LDAYMWLKTPGE <u>ADG</u> ---S---A-----	(93)
ABY52797	---P---S---K---P---L---LDAYMWIKTPGE <u>ADG</u> ---S---S-----	(94)
AAR08200	---PNP-G---M---P---L---LDAYMWLKTPGE <u>ADG</u> ---S---S-----	(95)
AAB92678	---P---N---P---GSMPL-LDAYMWIKTPGE <u>ADG</u> ---S---S-----	(96)
ABY52799	---S---P---DPEKFP---L-LDAYFWLKPPGE <u>ADG</u> ---S---D-----	(97)
AAC60491	---T---G---D---A---N---IDAYLWVKPPGE <u>ADG</u> -----	(98)
AAC09068	---V---K---M---P---L---LDAYMWLKTPGE <u>ADG</u> ---S---D-----	(99)
ZP_04371095	---T---G---D---A---A---VDAFLWIKPPGE <u>ADG</u> ---C---A-----	(100)
ZP_03818362	---T---G---D---S---Q---IDAFWLWVKIVGE <u>ADG</u> -----	(101)
ZP_03817628	---T---G---D---P---Q---IDAFWLWVKIPGE <u>ADG</u> -----	(102)
ZP_04331392	---T---G---N---P---L---IDAFIWTKLPG <u>ADG</u> -----	(103)
2BOE-X	---T---G---D---P---M---IDAFWLWIKLPGE <u>ADG</u> -----	(104)
ZP_04608509	---T---G---D---S---A---IAAYLWVKLPGE <u>ADG</u> -----	(105)
P26414	---T---G---D---P---A---IDAFWLWIKPPGE <u>ADG</u> -----	(106)

[0066] Para propósitos de la revelación, un polipéptido de la revelación exhibe mejor termoestabilidad con respecto a un polipéptido progenitor correspondiente si tiene una T₅₀ que es por lo menos cerca de 4°C o por lo menos cerca de 9°C más alta que la de la celulasa progenitora o por ejemplo una celobiohidrolasa que tiene una T₅₀ de cerca de 4°C a cerca de 30°C más alta o cualquier cantidad entre ellas o una T₅₀ de cerca de 9°C a cerca de 30°C más alta o cualquier cantidad entre ellas, cuando se compara con la de la celobiohidrolasa progenitora. La T₅₀ es la temperatura a la cual la enzima modificada o natural retiene 50% de su actividad residual después de una preincubación de 15 minutos y se determina por la prueba detallada en los Ejemplos que se dan a continuación o como se conoce en el estado de la técnica.

[0067] Las celobiohidrolasas o celulasas modificadas de la revelación pueden tener una T₅₀ que es de cerca de 4°C a cerca de 30°C más alta que la de una celobiohidrolasa progenitora correspondiente (p.ej., SEQ ID NO: 2, 4 o 6) o cualquier intervalo entre ellas, cerca de 5°C a cerca de 20°C más alta o cualquier intervalo entre ellas, cerca de 8°C a cerca de 15°C más alta o cualquier intervalo entre ellas o de cerca de 9°C a cerca de 15°C más alta o cualquier intervalo entre ellas. por ejemplo, la celulasa modificada puede tener una T₅₀ que es por lo

menos cerca de 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 o 30°C más alta que la de la celobiohidrolasa progenitora correspondiente.

[0068] Las variantes que se identifican aquí también se pueden usar para generar celobiohidrolasas quiméricas. Por ejemplo, previamente se ha usado SCHEMA para crear familias de cientos de quimeras activas de β -lactamasa y enzimas del citocromo P450. El SCHEMA usa datos de estructura proteínica para definir fronteras de "bloques" contiguos de aminoácidos que minimizan $\langle E \rangle$, el número promedio de la biblioteca de contactos de aminoácidos de cadena lateral que se rompen cuando los bloques se intercambian entre diferentes progenitores. Se ha demostrado que la probabilidad de que una quimera de β -lactamasa esté plegada y activa guardaba una relación inversa con el valor de E para esa secuencia. El algoritmo RASPP (la recombinación como el problema del camino más corto) se usó para identificar las fronteras de los bloques que minimizaron $\langle E \rangle$ con respecto al número promedio de mutaciones de la biblioteca, $\langle m \rangle$. Más de 20% de las ~500 quimeras singulares caracterizadas a partir de una colección de β -lactamasas estaba compuesto por 8 bloques de 3 progenitores ($3^8 = 6.561$ posibles secuencias) que eran catalíticamente activos. Un abordaje parecido produjo una familia de quimeras del citocromo P450 de 3 progenitores, 8 bloques que contienen más de 2.300 enzimas novedosas, catalíticamente activas. Las quimeras de estas dos colecciones se caracterizaron por elevados números de mutaciones, 66 y 72 aminoácidos en promedio desde el progenitor más cercano, respectivamente. Así las cosas, SCHEMA/RASPP permitió diseñar familias de quimeras que tienen significativa diversidad de secuencia y una fracción apreciable de miembros funcionales.

[0069] También se ha demostrado que las termoestabilidades de las quimeras SCHEMA se pueden predecir con base en datos de estabilidad de la secuencia a partir de una pequeña muestra de secuencias. El modelo de regresión lineal de los datos de inactivación térmica para 184 quimeras del citocromo P450 mostró que los bloques SCHEMA hicieron contribuciones aditivas a la termoestabilidad. Se predijo que más de 300 quimeras serían termoestables por este modelo y las 44 que se examinaron fueron más estables que el progenitor más estable. Se estimó que se podrían usar menos de 35 mediciones de termoestabilidad para predecir las quimeras más termoestables de todas. Además, las quimeras termoestables de P450 desplegaron perfiles singulares de actividad y especificidad, lo cual demostró que la quimeragénesis puede conducir a propiedades útiles adicionales de la enzima. Aquí la recombinación SCHEMA de las enzimas CBH II puede generar celulasas quiméricas que son activas sobre celulosa hinchada en ácido fosfórico (PASC) a altas temperaturas, durante periodos prolongados de tiempo y amplios intervalos de pH.

[0070] Usando los métodos aquí descritos se generaron numerosos polipéptidos quiméricos que tienen actividad de celobiohidrolasa con mejores características en comparación con las proteínas CBH II progenitoras de tipo silvestre.

[0071] Una familia diversa de novedosas enzimas CBH II se construyó permutando bloques de secuencia de tres enzimas CBH II micóticas. Veintitrés de 48 secuencias quiméricas muestreadas de este conjunto fueron segregadas en forma activa por *S. cerevisiae* y cinco tienen vidas medias a 63°C más prolongadas que la del progenitor más estable. Dado que este conjunto de muestra de 48 integrantes representa menos de 1% de las posibles 6561

secuencias, predecimos que esta colección de quimeras ya contiene cientos de enzimas CBH II termoestable activas, un número que eclipsa las aproximadamente veinte enzimas CBH II micóticas de la base de datos CAZy.

[0072] El abordaje de usar los datos de estabilidad de la secuencia del conjunto de muestras para identificar bloques que contribuyen positivamente a la termoestabilidad de la quimera se validó al hallar que las 10 quimeras catalíticamente activas del segundo conjunto de validación CBH II fueron más termoestables que el progenitor más estable, una CBH II naturalmente termoestable del hongo termofílico, *H. insolens*. Esta revelación hasta el momento ha generado un total de 33 nuevas enzimas CBH II que se expresan en forma catalíticamente activa en *S. cerevisiae*, 15 de las cuales son más termoestables que el progenitor más estable a partir del cual fueron construidas. Estas 15 enzimas termoestables son diversas en secuencia, difiriendo entre sí y con sus homólogos naturales más cercanos en hasta 94 y 58 posiciones de aminoácido, respectivamente.

[0073] El análisis de las termoestabilidades de las quimeras de CBH II en los conjuntos combinados de muestra y de validación indica que los cuatro bloques de termoestabilización identificados, B1P1, B6P3, B7P3 y B8P2, hacen contribuciones acumulativas a la estabilidad térmica cuando están presentes en la misma quimera. Cuatro de los cinco juego de muestras de quimeras que son más termoestables que la CBH II del *H. insolens* contienen bien sea dos o tres de estos bloques estabilizadores (Tabla 1). Los diez miembros activos del conjunto de validación, todos los cuales son más estables que la enzima del *H. insolens*, contienen por lo menos dos bloques estabilizadores, cinco de las seis quimeras más termoestables en este grupo conteniendo ya sean tres o cuatro bloques estabilizadores.

[0074] La revelación demuestra que los bloques estabilizadores pueden recombinarse para crear novedosas celulasas activas altamente estables. El modelo de regresión de la estabilidad predice que la biblioteca de CBH II SCHEMA contiene 2.026 quimeras que son más estables que la enzima progenitora más estable. Estas quimeras son diversas y distintas de las celulasas nativas: difieren de sus progenitoras por entre 8 y 72 mutaciones (un promedio de 50) y entre sí por un promedio de 63 mutaciones. Un total de 33 genes de este conjunto fueron sintetizados y expresados en *S. cerevisiae*: se encontró que todas estas CBH II quiméricas fueron más estables que la celulasas progenitora más estable del hongo termofílico *H. insolens*, medida bien sea por su vida media de inactivación a 63°C o por la T_{50} . Al reducir la complejidad de la secuencia haciendo quimeras de apenas ocho bloques se logró la generación de un modelo de estabilidad de secuencia y la identificación de un único bloque de secuencia altamente estabilizador. Al poner a prueba solamente diez sustituciones de aminoácidos en este bloque se identificó una única, sustitución altamente estabilizadora. El considerable efecto estabilizador de la sustitución C313S (con referencia a SEQ ID NO: 6; C314S, SEQ ID NO: 2 y C311S, SEQ ID NO: 4) observado en todas las quimeras y en las enzimas CBH II nativas de *P. chrysosporium*, *H. insolens* y *H. jecorina* sugiere que la mutación de cualquier residuo en esta posición a una Ser puede estabilizar cualquier familia de celulasas en la cual se introduce.

[0075] La minimización del número de contactos fragmentados luego de la recombinación (Figura 2C) permite que los bloques se aproximen como unidades desacopladas que hacen contribuciones independientes a la estabilidad de la proteína completa, lo cual conlleva

contribuciones acumulativas o aun aditivas a la termoestabilidad de la quimera. Por esta razón la recombinación SCHEMA de la enzima CBH II tuvo el efecto de minimizar esos contactos fragmentados: en tanto que hay 303 contactos totales interbloques definidos en la estructura cristalina de la CBH II progenitora del *H. insolens*, el diseño de la biblioteca SCHEMA de la CBH II se traduce en apenas 33 contactos fragmentados potenciales. Dado que las enzimas progenitoras de CBH II no muestran subdominios estructurales obvios y solamente cuatro de los ocho bloques (1, 5, 7 y 8) se parecen a unidades o módulos estructurales compactos, el reducido número de contactos fragmentados demuestra que el algoritmo SCHEMA/RASPP es efectivo en los casos en los cuales el número de bloques parece mayor que el número de subdivisiones estructurales. Como se había observado antes con las quimeras de β -lactamasa y del citocromo P450, los valores bajos de *E* fueron predictivos de plegamiento y actividad de quimera. Aunque no se usa aquí, esta relación debe ser valiosa para diseñar conjuntos de muestras de quimeras que contienen una elevada fracción de miembros activos.

[0076] La revelación también usó la quimera para determinar si se podría mejorar la estabilidad del pH en las enzimas CBH II. En tanto que la actividad específica de la CBH II de *H. jecorina* muestra una aguda declinación a medida que aumenta el pH por encima del valor óptimo de 5, la HJPlus, creada sustituyendo bloques estabilizadores sobre la enzima CBH II más relevante industrialmente de *H. jecorina*, retiene significativamente más actividad en estos pH más altos (Figura 4). Las quimeras termoestables 11113132 y 13311332 y también los progenitores de la celulasa CHB II de *H. insolens* y *C. thermophilum*, tienen perfiles de pH/actividad aún más amplios que la HJPlus. El estrecho perfil de pH/actividad de la CBH II de *H. jecorina* ha sido atribuido a la desprotonación de varios pares carboxilo-carboxilato, los cuales desestabilizan la proteína por encima de un pH ~6. La sustitución del progenitor 3 en el bloque 7 de la HJPlus cambia el aspartato 277 por histidina, eliminando el par carboxilo-carboxilato entre D277 y D316 (del bloque 8). Al reemplazar D277 por la histidina de carga positiva se puede prevenir la repulsión desestabilizadora de la carga a un pH no ácido, permitiendo que la HJPlus retenga la actividad a un pH más alto que la CBH II de *H. jecorina*. Los perfiles aún más amplios de pH/actividad de las dos quimeras termoestables restantes y de las enzimas CBH II progenitoras de *H. insolens* y *C. thermophilum* pueden deberse a la ausencia de residuos ácidos en posiciones correspondientes al par carboxilo-carboxilato E57-E119 de HJPlus y de la CBH II de *H. jecorina*.

[0077] La HJPlus exhibe tanto una actividad específica relativamente alta y como una elevada termoestabilidad. La Figura 5 muestra que estas propiedades conducen a un buen desempeño en experimentos de hidrólisis a largo plazo: la HJPlus hidrolizó la celulosa a temperaturas 7-15°C más altas que las enzimas CBH II progenitoras y también tuvo un aumento significativo de la a largo plazo con respecto a todos los progenitores a su temperatura óptima, superando a la CBH II de la *H. jecorina* en un factor de 1,7. Dado que la actividad específica de la quimera HJPlus es menor de la de la CBH II progenitora de *H. jecorina*, este aumento de la actividad a largo plazo se puede atribuir a la capacidad de la HJPlus termoestable a retener la actividad a temperaturas óptimas de hidrólisis durante tiempos de reacción más largos.

b
N

[0078] Las otras dos quimeras termoestables compartieron el amplio intervalo de temperatura de la HJPlus. Esta observación respalda una correlación positiva entre el $t_{1/2}$ a alta temperatura y máxima temperatura operativa y sugiere que muchas de las quimeras termoestables entre las 6.561 secuencias de quimera de CBH II también serán capaces de degradar la celulosa a altas temperaturas. Aunque esta capacidad de hidrolizar el sustrato amorfo PASC a altas temperaturas augura la potencial utilidad de quimeras termoestables de hongos de CBH II, los estudios con más sustratos cristalinos más dificultosos y sustratos que contienen lignina permitirán hacer una valoración más completa de la relevancia de esta novedosa familia de enzimas CBH II en aplicaciones de degradación de biomasa.

[0079] La mayoría de procesos de conversión de biomasa usa mezclas de celulasas de los hongos (principalmente CBH II, celobiohidrolasa clase I (CBH I), endoglucanasas y α -glucosidasa) para alcanzar niveles elevados de hidrólisis de la celulosa. La generación de un grupo diverso de quimeras termoestables de la enzima CBH II es el primer paso hacia la construcción de un inventario de celulasas estables, altamente activas a partir de las cuales se pueden formular y optimizar mezclas de enzimas para aplicaciones y materias primas específicas.

[0080] "Segmento peptídico" hace referencia a una porción o fragmento de un polipéptido o proteína más grande. Un segmento peptídico no tiene que tener actividad funcional por sí solo, aunque en algunos casos, un segmento peptídico puede corresponder a un dominio de un polipéptido en donde el dominio tiene su propia actividad biológica. Un segmento peptídico asociado a la estabilidad es un segmento peptídico que se encuentra en un polipéptido que promueve la estabilidad, la función o el plegamiento en comparación con a un polipéptido emparentado que carece del segmento peptídico. Un segmento peptídico asociado a la desestabilización es un segmento peptídico que se identifica como causante de pérdida de estabilidad, función o plegamiento cuando está presente en un polipéptido.

[0081] Los términos "fusionado", "vinculado de forma operativa", y "asociado de forma operativa" se usan de forma intercambiable en este documento para referirse ampliamente a un acople químico o físico de dos dominios o segmentos peptídicos por lo demás diferentes, en donde cada dominio o segmento peptídico cuando está vinculado de forma operativa puede proporcionar un polipéptido funcional que tiene una actividad deseada. Los dominios o segmentos peptídicos pueden estar conectados por medio de conectores peptídicos de manera que son funcionales o pueden fusionarse a través de otros intermediarios o enlaces químicos. Por ejemplo, dos dominios pueden ser parte de la misma secuencia codificadora, en donde los polinucleótidos se encuentran en el marco de manera que el polinucleótido cuando se transcribe codifica un único ARNm que cuando se traduce comprende los dos dominios como un único polipéptido. Como alternativa, los dos dominios pueden expresarse por separado como polipéptidos individuales y fusionarse entre sí usando métodos químicos. Típicamente, los dominios codificadores se vincularán "en el marco" ya sea directamente o separados por un péptido conector y codificados por un único polinucleótido. En el estado de la técnica se conocen varias secuencias codificadoras para conectores peptídicos y péptidos.

[0082] "Polinucleótido" o "secuencia de ácidos nucleicos" hace referencia a una forma polimérica de nucleótidos. En algunos casos un polinucleótido hace referencia a una secuencia

29

que no está inmediatamente contigua con ninguna de las secuencias codificadoras con las cuales es inmediatamente contigua (una en el extremo 5' y una en el extremo 3') en el genoma natural del organismo del cual se deriva. El término por consiguiente incluye, por ejemplo, un ADN recombinante que se incorpora dentro de un vector; dentro de un plásmido o virus que se duplica de forma autónoma; o dentro del ADN genómico de un procariote o eucariote o que existe como molécula separada (*p.ej.*, un ADNc) independiente de otras secuencias. Los nucleótidos de la revelación pueden ser ribonucleótidos, desoxirribonucleótidos o formas modificadas de cualquier nucleótido. El término polinucleótido como se usa en este contexto hace referencia, entre otros, ADN de cadena simple y doble, ADN que es una mezcla de regiones de cadena simple y doble, ARN de cadena simple y doble y ARN que es mezcla de regiones de cadena simple y doble, moléculas híbridas que comprenden ADN y ARN que puede ser de cadena única o, más típicamente, de cadena doble o una mezcla de regiones de cadena simple y doble. El término polinucleótido abarca ADN o ARN genómico (dependiendo del organismo, es decir, ARN del genoma de los virus), así como ARNm codificado por el ADN genómico y el ADNc.

[0083] Los términos "segmento de ácido nucleico", "segmento de oligonucleótido" o "segmento de polinucleótido" hacen referencia a una porción de una molécula de polinucleótido más grande. El segmento de polinucleótido no tiene que corresponder a un dominio funcional codificado de una proteína; sin embargo, en algunos casos el segmento codificará un dominio funcional de una proteína. Un segmento de polinucleótido puede tener cerca de 6 nucleótidos o más de longitud (*p.ej.*, 6-20, 20-50, 50-100, 100-200, 200-300, 300-400 o más nucleótidos de longitud). Un segmento peptídico asociado a la estabilidad puede ser codificado por un segmento de polinucleótido asociado a la estabilidad, en donde el segmento peptídico promueve estabilidad, función o plegamiento en comparación con un polipéptido que carece del segmento peptídico.

[0084] El término "quimera" hace referencia a una combinación de por lo menos dos segmentos de por lo menos dos diferentes proteínas progenitoras. Como lo apreciará la persona versada en la técnica, los segmentos en realidad no tienen que venir de cada uno de los progenitores, por cuanto es la secuencia particular la que es relevante y no los ácidos nucleicos físicos mismos. Por ejemplo, una celobiohidrolasa micótica quimérica de clase II (celulasas CBH II) tendrá por lo menos dos segmentos de dos diferentes polipéptidos progenitores CBH II. Los dos segmentos están conectados para dar lugar a un nuevo polipéptido que tiene actividad de celulasa. En otras palabras, una proteína no será una quimera si tiene la secuencia idéntica de cualquiera de los progenitores de longitud completa. Un polipéptido quimérico puede comprender más de dos segmentos de dos diferente proteínas progenitoras. Por ejemplo, puede haber 2, 3, 4, 5-10, 10-20 o más progenitores para cada quimera final o biblioteca de quimeras. El segmento de cada polipéptido progenitor puede ser muy corto o muy largo, los segmentos pueden variar en longitud de aminoácidos contiguos de 1 a 90%, 95%, 98% o 99% de la longitud completa de la proteína. En una modalidad, la longitud mínima es 10 aminoácidos. En una modalidad, un solo punto de cruce se define para dos progenitores. La localización del cruce define en dónde se va a detener un segmento de aminoácidos del progenitor y en dónde va a comenzar el siguiente segmento de aminoácidos

X
N

del progenitor. Así pues, una quimera simple solamente tendrá una localización del cruce en la cual el segmento que precede a la localización del cruce pertenecerá a un progenitor y el segmento que sigue a la localización del cruce pertenecerá al segundo progenitor. En una modalidad, la quimera tiene más de una localización de cruce. Por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11-30 o más localizaciones de cruce. Más adelante se examinan los nombres y definiciones de estas localizaciones de cruce. En una modalidad en la cual existen dos localizaciones de cruce y dos progenitores, habrá un primer segmento contiguo de un primer progenitor, seguido por un segundo segmento contiguo de un segundo progenitor, seguido por un tercer segmento contiguo del primer progenitor. Contiguo significa que no hay nada de significancia que interrumpa los segmentos. Estos segmentos contiguos están conectados formando una secuencia contigua de aminoácidos. Por ejemplo, una quimera CBH II de *Humicola insolens* (en lo sucesivo "1") y *H. jecori* (en lo sucesivo "2"), con dos cruces en 100 y 150, podría tener los primeros 100 aminoácidos de 1, seguidos por los siguientes 50 de 2, seguidos por el resto de los aminoácidos de 1, todos conectados en una cadena de aminoácidos contiguos. Como alternativa, la quimera CBH II podría tener los primeros 100 aminoácidos de 2, los siguientes 50 de 1 y el resto seguido por 2. Como lo apreciará la persona versada en la técnica, existen variantes de las quimeras así como las secuencias exactas. Así las cosas, no se necesita que el 100% de cada segmento esté presente en la quimera final si es una quimera variante. La cantidad que se puede alterar, ya sea a través de residuos adicionales o remoción o alteración de residuos se definirá a medida que se define el término variante. Por supuesto, como lo entenderá la persona versada en la técnica, la anterior discusión se aplica no solamente a los aminoácidos sino también a los ácidos nucleicos que codifican a los aminoácidos.

[0085] La revelación describe además de variantes específicas, variantes que pueden usarse para generar quimeras de CBH II. Se usó una biblioteca dirigida de recombinación SCHEMA para generar enzimas de celobiohidrolasa con base en un miembro especialmente bien estudiado de esta diversa familia de enzimas y de modo más particular las enzimas celobiohidrolasas II: El *H. insolens* es el progenitor "1" (SEQ ID NO: 2), el *H. jecorina* es el progenitor "2" (SEQ ID NO: 4) y el *C. thermophilum* es el progenitor "3" (SEQ ID NO: 6). SCHEMA es un método computarizado para predecir cuáles fragmentos de proteínas homólogas se pueden recombinar sin afectar la integridad estructural de la proteína (ver, *p.ej.*, Meyer *et al.*, (2003) Protein Sci., 12:1686-1693). Este abordaje computarizado identificó siete puntos de recombinación en las proteínas progenitoras de CBH II, lo cual permitió la formación de una biblioteca de polipéptidos quimeras de CBH II, en donde cada polipéptido comprende ocho segmentos. Las quimeras de mayor estabilidad son identificable determinando la contribución sumatoria de cada segmento a la estabilidad general, bien sea usando la regresión lineal de los datos de estabilidad de la secuencia o basados en el análisis de consenso de las MSAs de proteínas plegadas frente a no plegadas. La recombinación SCHEMA asegura que las quimeras retienen la función biológica y exhiben alta diversidad de secuencia porque conservan importantes residuos funcionales al tiempo que intercambian los tolerantes.

[0086] Así las cosas, como se ilustra en varias modalidades aquí, la revelación proporciona polipéptidos CBH II que comprenden una quimera de dominios parentales de los cuales una hebra parental o la secuencia codificadora quimérica resultante se pueden modificar para

cl
N

comprender una sustitución C→S arriba descrita. En algunas modalidades, el polipéptido comprende una quimera que tiene una pluralidad de dominios de N- a C terminal de diferentes proteínas parentales CBH II: (segmento 1)-(segmento 2)-(segmento 3)-(segmento 4)-(segmento 5)-(segmento 6)-(segmento 7)-(segmento 8);

en donde el segmento 1 comprende residuos de aminoácidos de cerca de 1 a cerca de x_1 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 2 va de cerca del residuo de aminoácidos x_1 a cerca de x_2 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 3 va de cerca del residuo de aminoácidos x_2 a cerca de x_3 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 4 va de cerca del residuo de aminoácidos x_3 a cerca de x_4 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 5 va de cerca del residuo de aminoácidos x_4 a cerca de x_5 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 6 va de cerca del residuo de aminoácidos x_5 a cerca de x_6 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 7 va de cerca del residuo de aminoácidos x_6 a cerca de x_7 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); y segmento 8 va de cerca del residuo de aminoácidos x_7 a cerca de x_8 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3");

en donde: x_1 es el residuo 43, 44, 45, 46 o 47 de la SEQ ID NO: 2 o el residuo 42, 43, 44, 45 o 46 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_2 es el residuo 70, 71, 72, 73 o 74 de la SEQ ID NO: 2 o el residuo 68, 69, 70, 71, 72, 73 o 74 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_3 es el residuo 113, 114, 115, 116, 117 o 118 de la SEQ ID NO: 2 o el residuo 110, 111, 112, 113, 114, 115 o 116 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_4 es el residuo 153, 154, 155, 156 o 157 de la SEQ ID NO: 2 o el residuo 149, 150, 151, 152, 153, 154, 155 o 156 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_5 es el residuo 220, 221, 222, 223 o 224 de la SEQ ID NO: 2 o el residuo 216, 217, 218, 219, 220, 221, 222 o 223 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_6 es el residuo 256, 257, 258, 259, 260 o 261 de la SEQ ID NO: 2 o el residuo 253, 254, 255, 256, 257, 258, 259 o 260 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_7 es el residuo 312, 313, 314, 315 o 316 de la SEQ ID NO: 2 o el residuo 309, 310, 311, 312, 313, 314, 315 o 318 de la SEQ ID NO: 4 o la SEQ ID NO: 6; y x_8 es un residuo de aminoácidos correspondiente al C terminal del polipéptido que tiene la secuencia de la SEQ ID NO: 2, SEQ ID NO: 4 o la SEQ ID NO: 6.

[0087] Usando las anteriores referencias de dominio se generaron varias estructuras quiméricas como las que aparecen en la Tabla 1.

1.588 secuencias de quimeras CBH II con valores de T_{50} predichos mayores que el valor de T_{50} de 64,8°C para la CBH II progenitora del *H. insolens*

29

1,588 CBH II chimera sequences with T₅₀ values predicted to be greater than the measured T₅₀ value of 64.8 C for the *H. insolens* parent CBH II.

31313232	13132231	13212231	21113231	22112331	33211132	22223232	32123131
31323333	11221233	13331133	21133232	21222133	33211131	31213132	11221333
11212133	33123232	13232232	22221133	32333132	12311232	22223231	11211231
31313231	33123231	13232231	21133231	32333131	12321333	22322232	11231232
31333232	21311333	22121133	23211232	22132332	33231132	31213131	11231231
11232133	21331333	33223232	23221333	22132331	12311231	31312132	33123332
31333231	32213332	23111232	23211231	33313332	33231131	22322231	33123331
21323112	32213331	23121333	23231232	33313331	12331232	31312131	31321232
21323111	32312332	33223231	11311333	33333332	12331231	31233132	31321231
32113332	21211133	33322232	23231231	33333331	23122232	31233131	33122132
32113331	32312331	23111231	21122133	33213132	23122231	31332132	33122131
31223133	13321112	33322231	11331333	33213131	11113232	31332131	12123232
21111133	32233332	23131232	11211133	33312132	11123333	23321133	12123231
31322133	13321111	23131231	11231133	33312131	11113231	23222232	32121332
32133332	32233331	11323112	33113132	12313232	11133232	23222231	32121331
32133331	32332332	11323111	33113131	12323333	12221133	11213232	21323133
21131133	21231133	11111133	33133132	33233132	11133231	11223333	13122232
32112132	32332331	11131133	33133131	12313231	31111332	11213231	13122231
32112131	32121232	32221232	31321133	33233131	31111331	11312232	22113132
32132132	32121231	32221231	31222232	33332132	31131332	11322333	22113131
32132131	13313133	22313232	31222231	33332131	13211232	11312231	22133132
21321312	32212132	22323333	12123133	12333232	31131331	11233232	22133131
33112332	32212131	22313231	32111132	12333231	13221333	11233231	23113332
21321311	13333133	22333232	13113232	32311332	13211231	11332232	23113331
33112331	32232132	22333231	13123333	32311331	13231232	11332231	23133332
11223233	32232131	31122232	32111131	32331332	22212332	31211332	23133331
33132332	33212332	31122231	13113231	32331331	13231231	31211331	12212332
11322233	33212331	11321312	32131132	12223133	22212331	31231332	21311232
33132331	22123133	11321311	13133232	12322133	11122133	31231331	12212331
31211232	33232332	22223133	32131131	31113332	22232332	12112332	21321333
31221333	33232331	22322133	13133231	31113331	22232331	31323232	21311231
31211231	23113232	23213232	33111332	31133332	33321232	12112331	12232332
21313333	23123333	23223333	33111331	32211132	33321231	11222133	21331232
31231232	23113231	23213231	33131332	13213232	23323133	31323231	12232331
31231231	23133232	23312232	11321233	31133331	31213332	12132332	21331231
21221112	23133231	23322333	33131331	13223333	31213331	12132331	23112132
21333333	11121112	23312231	13122133	32211131	31312332	32123332	23112131
21221111	12311133	23233232	12111232	13213231	31312331	32123331	23132132
12112232	11121111	11313333	12121333	13312232	31233332	21121133	23132131
12122333	12212232	23233231	12111231	13322333	31233331	32122132	32223332

3

12112231	12222333	23332232	12131232	13312231	31332332	32122131	22111332
12132232	12212231	23332231	12131231	32231132	31332331	33122332	32223331
12132231	21321233	11221112	32313332	13233232	31121232	33122331	32322332
21213133	12331133	11333333	32313331	32231131	31121231	31221232	22111331
21312133	12232232	11221111	21311133	13233231	22321133	31221231	21221133
13323112	12232231	23222133	21212232	13332232	22222232	21313232	32322331
13323111	13311333	11213133	21222333	13332231	31212132	21323333	22131332
21233133	23122133	11312133	32333332	33211332	22222231	21313231	22131331
21332133	13331333	11233133	21212231	33211331	31212131	21333232	13323133
32123232	11113133	11332133	32333331	33231332	31232132	21333231	32222132
32123231	11133133	22211232	21331133	33231331	31232131	12122232	32222131
13111133	32223232	22221333	21232232	22122232	23311232	12122231	33311132
13131133	22111232	22211231	21232231	31112132	23321333	22113332	33311131
12321112	22121333	22231232	32213132	22122231	23311231	22113331	33222332
12321111	32223231	22231231	32213131	31112131	23331232	21223133	33222331
33122232	32322232	31323133	32312132	31132132	23331231	21322133	33331132
33122231	22111231	21112232	32312131	33323232	21123133	22133332	33331131
21211333	32322231	21122333	32233132	31132131	23221133	22133331	23123232
13321312	22131232	21112231	32233131	13222133	11311133	13121133	23123231
13321311	22131231	21132232	32332132	33323231	11212232	22112132	12321133
21231333	13211133	21132231	32332131	12211232	11222333	22112131	12222232
11123112	13231133	32113132	33213332	12221333	11212231	22132132	12222231
12313133	11323312	32113131	33213331	12211231	11331133	22132131	13311232
11123111	11323311	11211333	33312332	12231232	11232232	33313132	13321333
21323233	33321133	32133132	33312331	12231231	11232231	33313131	13311231
12333133	33222232	32133131	33233332	23121133	31223232	23112332	22213332
13313333	33222231	11231333	33233331	11112232	21111232	23112331	13331232
32212332	11111333	33113332	33332332	11122333	21121333	33333132	22213331
32212331	11131333	33113331	33332331	11112231	31223231	33333131	22312332
13221112	11223112	33133332	33121232	11132232	31322232	23132332	13331231
13333333	11223111	11323233	33121231	32321232	21111231	23132331	22312331
13221111	11322112	33133331	33212132	11132231	31322231	21211232	11123133
32232332	11322111	31311232	33212131	32321231	21131232	21221333	22233332
32232331	13321233	31321333	12213232	31123232	21131231	21211231	22233331
22113232	22213232	31311231	12223333	31123231	32122332	21231232	22332332
22123333	22223333	31331232	12213231	22323133	32122331	21231231	22332331
22113231	22213231	31331231	12312232	33221232	13123133	12323133	22121232
22133232	22312232	21321112	12322333	33221231	33111132	32311132	31111132
22133231	22322333	33112132	33232132	23313232	33111131	13313232	22121231
13213133	22312231	21321111	12312231	23323333	33131132	13323333	31111131
13312133	22233232	33112131	33232131	23313231	33131131	32311131	31131132
13233133	22233231	12113232	12233232	23333232	21213232	13313231	31131131
13332133	22332232	12123333	12233231	23333231	21223333	32222332	13221133

16

11121312	22332231	12113231	12332232	11321112	21213231	32222331	22212132
12311333	31121133	33132132	12332231	11321111	21312232	32331132	22212131
11121311	11221312	33132131	32211332	23223133	21322333	13333232	22232132
22122133	11221311	12133232	32211331	23322133	21312231	32331131	22232131
12331333	22222133	12133231	23123133	11313133	21233232	13333231	23212332
33323133	23311133	32111332	32231332	11333133	21233231	33311332	23212331
23112232	23212232	32111331	32231331	22311232	21332232	33311331	23232332
23122333	23222333	31221133	32323232	22321333	21332231	33331332	23232331
23112231	23212231	32131332	12222133	22311231	12121133	22123232	11111232
11113333	23331133	21313133	32323231	31212332	13111232	33331331	11121333
23132232	11213333	32131331	31112332	31212331	13121333	31113132	11111231
23132231	23232232	21333133	31112331	22331232	13111231	22123231	11131232
11133333	11312333	12122133	13311133	22331231	32313132	31113131	11131231
12211133	23232231	13112232	31132332	31232332	32313131	31133132	31313332
21221233	11233333	13122333	13212232	31232331	13131232	31133131	31313331
12231133	11332333	13112231	31132331	21113232	22112332	13223133	31333332
13211333	11121233	13132232	13222333	21123333	13131231	13322133	31333331
22131131	22331331	13111331	33323132	33321132	22121132	11311132	11321132
33223332	21113332	13131332	33323131	33321131	22121131	11311131	11321131
23111332	21113331	13131331	23122332	11111332	23121332	11222332	11323132
33223331	21133332	21212132	23122331	11111331	23121331	11222331	11323131
33322332	22211132	21212131	11113332	11131332	11111132	11331132	11321332
23111331	21133331	21232132	11113331	11131331	11111131	11331131	11321331
33322331	22211131	21232131	11133332	13321232	11131132	21121332	11221132
23131332	22231132	12313332	12211132	13321231	11131131	21121331	11221131
23131331	22231131	12313331	11133331	22223332	22323332	13123132	21321132
33222132	23211332	12333332	12211131	22223331	22323331	13123131	21321131
33222131	23211331	12333331	21221232	22322332	22223132	21223332	13321132
12223232	23231332	33121132	21221231	22322331	22223131	21223331	13321131
12223231	23231331	33121131	12231132	31121132	22322132	21322332	11121132
12322232	21112132	12213132	12231131	31121131	22322131	21322331	11121131
12322231	21112131	12213131	13211332	22222132	23223332	12121132	11323332
32221332	21132132	12312132	13211331	22222131	23223331	12121131	11323331
32221331	23323232	12312131	13231332	23311132	23322332	13121332	11223132
22313332	21132131	21223232	13231331	23311131	23322331	13121331	11223131
22313331	23323231	21223231	11112132	23222332	11313332	21222132	11322132
22333332	11323133	21322232	11112131	23222331	11313331	21222131	11322131
22333331	22321232	12233132	11132132	23331132	11333332	12323332	11221332
31122332	31311132	21322231	32321132	11213332	11333331	12323331	11221331
31122331	22321231	12233131	13323232	23331131	23222132	12223132	21323132
13321133	31311131	12332132	11132131	11213331	23222131	12223131	21323131
13222232	31222332	12332131	32321131	11312332	11213132	12322132	21321332
13222231	31222331	13213332	13323231	11312331	11213131	12322131	21321331

2
3

22213132	31331132	13213331	33321332	11233332	11312132	13223332	21221132
22213131	31331131	13312332	33321331	11233331	11312131	13223331	21221131
22312132	12113132	13312331	31123132	11332332	11233132	13322332	13323132
22312131	12113131	13233332	31123131	11332331	11233131	13322331	13323131
22233132	21123232	13233331	33221132	11121232	11332132	13222132	12321132
22233131	21123231	13332332	33221131	11121231	11332131	13222131	12321131
22332132	12133132	13332331	23313132	31323332	22221332	12221332	13321332
22332131	12133131	13121232	12321232	11212132	22221331	12221331	13321331
23213332	13113332	13121231	23313131	31323331	31323132	23121132	11123132
23213331	13113331	32323132	12321231	11212131	31323131	23121131	11123131
23312332	13133332	32323131	23333132	11232132	21122332	11122332	13221132
23312331	13133331	22122332	23333131	11232131	21122331	11122331	13221131
23233332	23221232	22122331	11123232	31223132	11211332	22323132	11121332
23233331	23221231	33323332	11123231	21111132	11211331	22323131	11121331
23332332	11311232	13212132	31121332	31223131	11231332	23323332	23321132
23332331	11321333	33323331	31121331	31322132	11231331	23323331	23321131
23121232	11311231	13212131	13221232	21111131	11323232	23223132	11223332
23121231	11331232	13232132	13221231	31322131	11323231	23223131	11223331
23212132	11331231	13232131	22311132	21131132	31321332	23322132	11322332
23212131	13112132	12211332	22311131	21131131	31321331	23322131	11322331
23232132	13112131	12211331	22222332	11223232	12123332	11313132	11222132
23232131	13132132	12231332	22222331	11223231	12123331	11313131	11222131
22211332	13132131	12231331	22331132	11322232	31221132	11333132	21121132
22211331	12111332	33223132	22331131	11322231	31221131	11333131	21121131
11121133	12111331	23111132	23311332	31221332	21313132	22321332	21323332
22231332	11221133	33223131	23311331	31221331	21313131	22321331	21323331
22231331	12131332	33322132	23331332	21313332	21333132	21123332	21223132
22323232	12131331	23111131	23331331	21313331	21333131	21123331	21223131
31313132	33123132	33322131	21113132	21333332	12122132	22221132	21322132
22323231	33123131	12323232	21113131	21333331	12122131	22221131	21322131
31313131	21212332	12323231	21133132	12122332	13122332	23221332	13121132
21112332	21212331	23131132	21133131	12122331	13122331	23221331	13121131
21112331	21232332	23131131	23211132	21213132	11221232	11311332	21221332
31333132	21232331	11112332	23211131	21213131	11221231	11311331	21221331
31333131	32121132	11112331	23231132	21312132	21311332	21122132	12323132
21132332	13123232	11132332	23231131	21312131	21311331	21122131	12323131
21132331	32121131	32321332	11212332	21233132	21331332	11331332	13323332
23223232	13123231	11132331	11212331	21233131	21331331	11331331	13323331
23223231	33121332	32321331	11232332	21332132	21211132	11211132	13223132
23322232	33121331	31123332	11232331	21332131	21211131	11211131	13223131
23322231	12213332	31123331	31223332	13111132	21231132	11231132	13322132
11313232	12213331	32221132	21111332	13111131	21231131	11231131	13322131
11323333	12312332	13223232	31223331	13131132	13313132	31321132	12321332

33

11313231	12312331	32221131	31322332	13131131	13313131	31321131	12321331
11333232	12233332	13223231	21111331	21211332	13333132	12123132	11123332
11333231	12233331	13322232	31322331	21211331	13333131	12123131	11123331
31311332	12332332	13322231	21131332	21231332	22123132	13123332	12221132
31311331	12332331	22313132	21131331	21231331	22123131	13123331	12221131
31331332	23113132	22313131	31222132	12313132	23123332	11321232	13221332
31331331	12121232	22333132	31222131	12313131	23123331	11321231	13221331
12113332	23113131	33221332	23321232	21323232	12311132	13122132	11122132
12113331	12121231	22333131	23321231	21323231	12311131	13122131	11122131
11223133	23133132	33221331	13113132	12333132	12222332	12121332	23323132
11322133	23133131	23313332	13113131	12333131	21321232	12121331	23323131
12133332	32323332	23313331	13133132	13313332	12222331	21311132	22321132
12133331	12212132	31122132	13133131	13313331	21321231	21311131	22321131
22221232	32323331	23333332	11321133	13333332	12331132	21222332	23321332
31211132	12212131	31122131	11222232	13333331	12331131	21222331	23321331
22221231	21321133	23333331	11222231	22123332	13311332	21331132	21123132
31211131	21222232	23213132	21213332	22123331	13311331	21331131	21123131
31231132	21222231	12221232	21213331	13213132	23122132	12223332	23221132
31231131	12232132	23213131	21312332	13213131	23122131	12223331	23221131
12112132	12232131	23312132	21312331	13312132	13331332	12322332	
12112131	13212332	12221231	21233332	13312131	13331331	12322331	
21122232	13212331	23312131	21233331	13233132	11113132	23123132	
21122231	13232332	23233132	21332332	13233131	11113131	23123131	
12132132	13232331	23233131	12111132	13332132	11133132	12222132	
12132131	32223132	23332132	21332331	13332131	11133131	12222131	
13112332	22111132	23332131	12111131	12311332	22121332	13311132	
13112331	32223131	22311332	21121232	12311331	22121331	13311131	
13132332	32322132	22311331	21121231	22122132	13211132	13222332	
13132331	22111131	11122232	12131132	22122131	13211131	13222331	
32123132	32322131	22331332	12131131	12331332	13231132	13331132	
11211232	22131132	11122231	13111332	12331331	13231131	13331131	

De acuerdo con la anterior Tabla, cada dígito hace referencia a un dominio de un polipéptido quimérico CBH II. El número denota la hebra parental del dominio del cual se deriva. Por ejemplo, un polipéptido quimérico CBH II que tiene la secuencia 12111131, indica que el polipéptido comprende una secuencia del N-terminal al C terminal de: aminoácidos de cerca de 1 a x_1 de la SEQ ID NO: 2 ("1") vinculada a los aminoácidos de cerca de x_1 a x_2 de la SEQ ID NO: 4 ("2") vinculada a los aminoácidos de cerca de x_2 a cerca de x_3 de la SEQ ID NO: 2 vinculada a los aminoácidos de cerca de x_3 a cerca de x_4 de la SEQ ID NO: 2 vinculada a los aminoácidos de cerca de x_4 a cerca de x_5 de la SEQ ID NO: 2 vinculada a los aminoácidos de cerca de x_5 a cerca de x_6 de la SEQ ID NO: 2 vinculada a los aminoácidos de cerca de x_6 a x_7 de la SEQ ID NO: 6 ("3") vinculada a los aminoácidos de cerca de x_7 a x_8 (p.ej., el C terminal) de la SEQ ID NO: 2.

[0088] En algunas modalidades, el polipéptido CBH II tiene una estructura de segmento quimérica seleccionada de entre el grupo compuesto por: 11113132, 21333331, 21311131, 22232132, 33133132, 33213332, 13333232, 12133333, 13231111, 11313121, 11332333, 12213111, 23311333, 13111313, 31311112, 23231222, 33123313, 22212231, 21223122, 21131311, 23233133, 31212111 y 32333113.

[0089] En algunas modalidades el polipéptido tiene mejor termoestabilidad comparado con un polipéptido de tipo silvestre de la SEQ ID NO: 2, 4 o 6. La actividad del polipéptido se puede medir con cualquiera o una combinación de sustratos como se describe en los ejemplos. como será aparente para la persona medianamente versada en la técnica, se pueden examinar y usar otros compuestos dentro de la clase de los compuestos ejemplificados por los que se discuten en los ejemplos.

[0090] En algunas modalidades, el polipéptido puede tener varios cambios de la secuencia de aminoácidos con respecto a una secuencia de referencia. Los cambios pueden ser una sustitución, delección o inserción de uno o más aminoácidos. Cuando el cambio es una sustitución, el cambio puede ser una sustitución conservadora, una conservadora no conservadora o una combinación de sustituciones conservadoras y no conservadoras. Por ejemplo, la quimera puede comprender una sustitución C→S en C314 de la SEQ ID NO: 2 o C311 de la SEQ ID NO: 4.

[0091] Así pues, en algunas modalidades, los polipéptidos pueden comprender una estructura general del N-terminal al C terminal: (segmento 1)-(segmento 2)-(segmento 3)-(segmento 4)-(segmento 5)-(segmento 6)-(segmento 7)-(segmento 8),

en donde el segmento 1 comprende residuos de aminoácidos de cerca de 1 a cerca de x_1 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y que tienen 1-10 sustituciones conservadoras de aminoácido; el segmento 2 va de cerca del residuo de aminoácidos x_1 a cerca de x_2 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y tiene cerca de 1-10 sustituciones conservadoras de aminoácido; el segmento 3 va de cerca del residuo de aminoácidos x_2 a cerca de x_3 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y tiene cerca de 1-10 sustituciones conservadoras de aminoácido; el segmento 4 va de cerca del residuo de aminoácidos x_3 a cerca de x_4 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y tiene cerca de 1-10 sustituciones conservadoras de aminoácido; el segmento 5 va de cerca del residuo de aminoácidos x_4 a cerca de x_5 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y tiene cerca de 1-10 sustituciones conservadoras de aminoácido; el segmento 6 va de cerca del residuo de aminoácidos x_5 a cerca de x_6 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y tiene cerca de 1-10 sustituciones conservadoras de aminoácido; el segmento 7 va de cerca del residuo de aminoácidos x_6 a cerca de x_7 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y tiene cerca de 1-10 sustituciones conservadoras de aminoácido; y el segmento 8 va de cerca del residuo de aminoácidos x_7 a cerca de x_8 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y tiene cerca de 1-10 sustituciones conservadoras de aminoácido;

[0092] en donde x_1 es el residuo 43, 44, 45, 46 o 47 de la SEQ ID NO: 2 o el residuo 42, 43, 44, 45 o 46 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_2 es el residuo 70, 71, 72, 73 o 74 de la SEQ ID NO: 2 o el residuo 68, 69, 70, 71, 72, 73 o 74 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_3 es el residuo 113, 114, 115, 116, 117 o 118 de la SEQ ID NO: 2 o el residuo 110, 111, 112, 113, 114, 115 o 116 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_4 es el residuo 153, 154, 155, 156 o 157 de la SEQ ID NO: 2 o el residuo 149, 150, 151, 152, 153, 154, 155 o 156 de la SEQ ID

NO: 4 o la SEQ ID NO: 6; x_5 es el residuo 220, 221, 222, 223 o 224 de la SEQ ID NO: 2 o el residuo 216, 217, 218, 219, 220, 221, 222 o 223 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_6 es el residuo 256, 257, 258, 259, 260 o 261 de la SEQ ID NO: 2 o el residuo 253, 254, 255, 256, 257, 258, 259 o 260 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_7 es el residuo 312, 313, 314, 315 o 316 de la SEQ ID NO: 2 o el residuo 309, 310, 311, 312, 313, 314, 315 o 318 de la SEQ ID NO: 4 o la SEQ ID NO: 6; y x_8 es un residuo de aminoácidos correspondiente al C terminal del polipéptido y tiene la secuencia de la SEQ ID NO: 2, SEQ ID NO: 4 o la SEQ ID NO: 6 y en donde la quimera tiene un algoritmo como el estipulado en la Tabla 1 y en donde la quimera comprende una sustitución C→S correspondiente a una C314 de la SEQ ID NO: 2 o C311 de la SEQ ID NO: 4.

[0093] En algunas modalidades, el número de sustituciones puede ser 2, 3, 4, 5, 6, 8, 9 o 10 o más sustituciones de aminoácidos (p.ej., 10-20, 21-30, 31-40 y similares sustituciones de aminoácidos).

[0094] En algunas modalidades, los polipéptidos funcionales CBH II pueden tener actividad de celulasa junto con termoestabilidad aumentada, como para un sustrato definido examinado en los ejemplos y también tener un nivel de identidad de secuencia de aminoácidos con la celobiohidrolasa de referencia o segmentos de la misma. La enzima o segmento de referencia puede ser el de una enzima de tipo silvestre (p.ej., natural) o una enzima obtenida por ingeniería genética. Así pues, en algunas modalidades, los polipéptidos de la revelación pueden comprender una estructura general del N-terminal al C terminal:

en donde el segmento 1 comprende una secuencia que es por lo menos 50-100% idéntica al residuo de aminoácidos de cerca de 1 a cerca de x_1 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 2 comprende una secuencia que es por lo menos 50-100% idéntica al residuo de aminoácidos x_1 a cerca de x_2 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 3 comprende una secuencia que es por lo menos 50-100% idéntica al residuo de aminoácidos x_2 a cerca de x_3 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 4 comprende una secuencia que es por lo menos 50-100% idéntica al residuo de aminoácidos x_3 a cerca de x_4 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 5 comprende una secuencia que es por lo menos 50-100% idéntica a cerca del residuos de aminoácidos x_4 a cerca de x_5 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 6 comprende una secuencia que es por lo menos 50-100% idéntica al residuo de aminoácidos x_5 a cerca de x_6 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 7 comprende una secuencia que es por lo menos 50-100% idéntica al residuo de aminoácidos x_6 a cerca de x_7 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); y el segmento 8 comprende una secuencia que es por lo menos 50-100% idéntica al residuo de aminoácidos x_7 a cerca de x_8 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3");

en donde x_1 es el residuo 43, 44, 45, 46 o 47 de la SEQ ID NO: 2 o el residuo 42, 43, 44, 45 o 46 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_2 es el residuo 70, 71, 72, 73 o 74 de la SEQ ID NO: 2 o el residuo 68, 69, 70, 71, 72, 73 o 74 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_3 es el residuo 113, 114, 115, 116, 117 o 118 de la SEQ ID NO: 2 o el residuo 110, 111, 112,

36

113, 114, 115 o 116 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_4 es el residuo 153, 154, 155, 156 o 157 de la SEQ ID NO: 2 o el residuo 149, 150, 151, 152, 153, 154, 155 o 156 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_5 es el residuo 220, 221, 222, 223 o 224 de la SEQ ID NO: 2 o el residuo 216, 217, 218, 219, 220, 221, 222 o 223 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_6 es el residuo 256, 257, 258, 259, 260 o 261 de la SEQ ID NO: 2 o el residuo 253, 254, 255, 256, 257, 258, 259 o 260 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_7 es el residuo 312, 313, 314, 315 o 316 de la SEQ ID NO: 2 o el residuo 309, 310, 311, 312, 313, 314, 315 o 318 de la SEQ ID NO: 4 o la SEQ ID NO: 6; y x_8 es un residuo de aminoácidos correspondiente al C terminal del polipéptido que tiene la secuencia de la SEQ ID NO: 2, SEQ ID NO: 4 o la SEQ ID NO: 6 y en donde la quimera tiene un algoritmo como el estipulado en la Tabla 1 y en donde la quimera comprende una sustitución C→S correspondiente a C314 de la SEQ ID NO: 2 o C311 de la SEQ ID NO: 4.

[0095] En algunas modalidades, cada segmento del polipéptido quimérico puede tener por lo menos 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98% o 99% o más de identidad de secuencia en comparación con el segmento de referencia indicado para cada uno de los (segmento 1), (segmento 2), (segmento 3), (segmento 4)-(segmento 5), (segmento 6), (segmento 7) y (segmento 8) de la SEQ ID NO: 2, SEQ ID NO: 4 o la SEQ ID NO: 6.

[0096] En algunas modalidades, las variantes del polipéptido pueden tener mejor termoestabilidad en comparación con la actividad enzimática del polipéptido de tipo silvestre de la SEQ ID NO: 2, 4 o 6 y en donde la quimera comprende una sustitución C→S correspondiente a C314 de la SEQ ID NO: 2 o C311 de la SEQ ID NO: 4.

[0097] Las enzimas quiméricas aquí descritas pueden prepararse de varias formas, como lisados, extractos crudos o preparaciones aisladas. Los polipéptidos pueden disolverse en soluciones apropiadas; formulados como polvos, como un polvo de acetona (con o sin estabilizadores); o prepararse como liofilizados. En algunas modalidades, el polipéptido puede ser un polipéptido aislado.

[0098] En algunas modalidades, los polipéptidos puede estar en forma de filas. Las enzimas pueden estar en forma soluble, por ejemplo, como soluciones en los pozos de placas de microtitulación o inmovilizados sobre un sustrato. El sustrato puede ser un sustrato sólido o un sustrato poroso (p.ej., membrana), el cual puede estar compuesto por polímeros orgánicos como poliestireno, polietileno, polipropileno, polifluoroetileno, polietilenooxi y poliacrilamida, así como copolímeros e injertos de los mismos. Un soporte sólido también puede ser inorgánico, como vidrio, bióxido de silicio, vidrio de poro controlado (CPG), bióxido de silicio de fase inversa o metal, como oro o platino. La configuración de un sustrato puede estar en forma de perlas, esferas, partículas, gránulos, un gel, una membrana o una superficie. Las superficies pueden ser planas, sustancialmente planas o no planas. Los soportes sólidos pueden ser porosos o no porosos y pueden tener características higroscópicas o no higroscópicas. Un soporte sólido puede estar en la forma de un pozo, depresión u otro recipiente, vaso, característica o localización. Una pluralidad de soportes puede ser configurada sobre una fila en diversas localizaciones, con acceso para la adición robótica de reactivos o métodos y/o instrumentos de detección.

7
3

[0099] La revelación también proporciona polinucleótidos que codifican los polipéptidos CBH II de ingeniería genética aquí revelados. Los polinucleótidos pueden estar vinculados operativamente a una o más secuencias heterólogas reguladoras o de control que controlan la expresión genética para crear un polinucleótido recombinante capaz de expresar el polipéptido. Los constructos de expresión que contienen un polinucleótido heterólogo que codifica la quimera CBH II se pueden introducir en células hospederas adecuadas para expresar el polipéptido.

[00100] Dado el conocimiento de las secuencias específicas de las enzimas quiméricas CBH II (p.ej., la estructura del segmento de la CBH II quimérica), las secuencias de polinucleótidos serán aparentes a partir de la secuencia de aminoácidos de las enzimas quiméricas CBH II obtenidas por ingeniería genética para la persona versada en la técnica. El conocimiento de los codones correspondientes a varios aminoácidos aunado al conocimiento de la secuencia de aminoácidos de los polipéptidos le permite a la persona medianamente versada en la técnica hacer diferentes polinucleótidos que codifican los polipéptidos de la revelación. Así pues, la revelación contempla todas y cada una de las posibles variaciones de los polinucleótidos que podrían hacerse seleccionando combinaciones con base en posibles elecciones de codones y todas esas variaciones deben considerarse como específicamente reveladas para cualquiera de los polipéptidos aquí descritos.

[00101] En algunas modalidades, los polinucleótidos codifican los polipéptidos aquí descritos pero tienen cerca de 80% o más de identidad de secuencia, cerca de 85% o más de identidad de secuencia, cerca de 90% o más de identidad de secuencia, cerca de 91% o más de identidad de secuencia, cerca de 92% o más de identidad de secuencia, cerca de 93% o más de identidad de secuencia, cerca de 94% o más de identidad de secuencia, cerca de 95% o más de identidad de secuencia, cerca de 96% o más de identidad de secuencia, cerca de 97% o más de identidad de secuencia, cerca de 98% o más de identidad de secuencia o cerca de 99% o más de identidad de secuencia en el nivel de los nucleótidos con un polinucleótido de referencia que codifica la variante CBH II de polipéptidos quiméricos y que tienen una sustitución C→S arriba descrita (p.ej., en donde el polipéptido o quimera comprende una sustitución C→S correspondiente a C314 de la SEQ ID NO: 2 o C311 de la SEQ ID NO: 4).

[00102] En algunas modalidades, los polinucleótidos aislados que codifican los polipéptidos se pueden manipular de diversas formas para dar lugar a la expresión del polipéptido. La manipulación del polinucleótido aislado antes de su inserción dentro de un vector puede ser deseable o necesaria dependiendo del vector de expresión. Las técnicas para modificar los polinucleótidos y las secuencias de ácidos nucleicos utilizando métodos de ADN recombinante son bien conocidas del estado de la técnica. Se encuentra orientación en Sambrook *et al.*, 2001, *Molecular Cloning: A Laboratory Manual*, 3ª Ed., Cold Spring Harbor Laboratory Press; y *Current Protocols in Molecular Biology*, Ausubel, F. ed., Greene Pub. Associates, 1998, actualizados a 2007.

[00103] En algunas modalidades, los polinucleótidos están operativamente vinculados a secuencias de control para la expresión de los polinucleótidos y/o polipéptidos. En algunas modalidades, la secuencia de control puede ser una secuencia promotora adecuada, la cual se puede obtener de genes que codifican polipéptidos extracelulares o intracelulares, bien sea

32

homólogos o heterólogos a la célula hospedera. En el caso de las células hospederas bacterianas, los promotores apropiados para dirigir la transcripción de los constructos de ácidos nucleicos de la presente revelación, incluyen los promotores obtenidos de los genes lac operón de *E. coli* operon, xylA y xylB del *Bacillus subtilis*, genes de utilización de xilosa del *Bacillus megatarium* (p.ej., Rygus *et al.*, (1991) Appl. Microbiol. Biotechnol. 35:594-599; Meinhardt *et al.*, (1989) Appl. Microbiol. Biotechnol. 30:343-350), gen de la beta-lactamasa de los procariotes (Villa-Kamaroff *et al.*, (1978) Proc. Natl Acad. Sci. USA 75: 3727-3731), así como el promotor tac (DeBoer *et al.*, (1983) Proc. Natl Acad. Sci. USA 80: 21-25). Varios promotores apropiados se describen en "Useful proteins from recombinant bacteria" en Scientific American, 1980, 242:74-94; y en Sambrook *et al.*, *supra*.

[00104] En algunas modalidades, la secuencia de control también puede ser una secuencia apropiada de terminación de transcripción, una secuencia reconocida por una célula hospedera para terminar la transcripción. La secuencia terminadora está vinculada de forma operativa al 3' terminal de la secuencia de ácidos nucleicos que codifica el polipéptido. Se puede usar cualquier terminador que sea funcional en la célula hospedera de elección.

[00105] En algunas modalidades, la secuencia de control también puede ser una secuencia líder adecuada, una región no traducida de un ARNm que es importante para la traducción por la célula hospedera. La secuencia líder está vinculada de forma operativa al 5' terminal de la secuencia de ácidos nucleicos que codifica el polipéptido. Se puede usar cualquier secuencia líder que sea funcional en la célula hospedera de elección.

[00106] En algunas modalidades, la secuencia de control también puede ser una región codificadora de un péptido de señal que codifica una secuencia de aminoácidos vinculada al amino terminal de un polipéptido y dirige el polipéptido codificado dentro de la vía secretoria de la célula. El extremo 5' de la secuencia codificadora de la secuencia de ácidos nucleicos puede de forma inherente contener una región codificadora de un péptido de señal vinculada naturalmente en marco de lectura de la traducción con el segmento de la región codificadora que codifica el polipéptido secretado. Como alternativa, el extremo 5' de la secuencia codificadora puede contener una región codificadora de un péptido de señal que es ajena a la secuencia codificadora. La región codificadora ajena de un péptido de señal se puede requerir cuando la secuencia codificadora no contiene naturalmente una región codificadora de un péptido de señal. La región codificadora de un péptido de señales efectiva para células hospederas bacterianas puede ser la región codificadora de un péptido de señales obtenida de los genes de la amilasa maltogénica del *Bacillus* NCIB 11837, alfa-amilasa del *Bacillus stearothermophilus*, beta-lactamasa de *Bacillus licheniformis subtilisin*, *Bacillus licheniformis*, proteasas naturales de *Bacillus stearothermophilus* (nprT, nprS, nprM) y prsA de *Bacillus subtilis*. Otros péptidos de señales son descritos por Simonen y Palva, (1993) Microbiol Rev 57: 109-137.

[00107] La revelación va dirigida además a un vector de expresión recombinante que comprende un polinucleótido que codifica la variante CBH II obtenida por ingeniería genética o polipéptido quimérico y una o más regiones reguladoras de expresión como un promotor y un terminador, un origen de replicación, etc., dependiendo del tipo de huéspedes dentro de los cuales se van a introducir. Al crear el vector de expresión, la secuencia codificadora se localiza

39

en el vector de manera que la secuencia codificadora está vinculada de forma operativa con las secuencias apropiadas de control para expresión.

[00108] El vector de expresión recombinante puede ser cualquier vector (*p.ej.*, un plásmido o virus), el cual puede ser sometido de forma conveniente a procedimientos recombinantes de ADN y producir la expresión de las secuencias de polinucleótidos. La elección del vector típicamente dependerá de la compatibilidad del vector con la célula hospedera dentro de la cual se va a introducir el vector. Los vectores pueden ser plásmidos lineales o circulares cerrados.

[00109] El vector de expresión puede ser un vector que se replica de forma autónoma, es decir, un vector que existe como una entidad extracromosómica, cuya replicación es independiente de la replicación cromosómica, *p.ej.*, un plásmido, un elemento extracromosómico, un minicromosoma o un cromosoma artificial. El vector puede contener cualquier medio para asegurar la auto-replicación. Como alternativa, el vector puede ser tal que, cuando se introduce dentro de la célula hospedera, se integra dentro del genoma y se replica al tiempo con el o los cromosomas dentro de los cuales se ha integrado. Además, se pueden usar un vector o plásmido único o dos o más vectores o plásmidos que juntos contienen el ADN total que se va a introducir dentro del genoma de la célula hospedera o un transposón.

[00110] En algunas modalidades, el vector de expresión de la revelación contiene uno o más marcadores seleccionables, lo cual facilita la selección de las células transformadas. Un marcador seleccionable es un gen cuyo producto da lugar a resistencia biocida o viral, resistencia a metales pesados, prototrofia a auxotrofos y similares. Algunos ejemplos de marcadores seleccionables bacterianos son los genes o marcadores del *Bacillus subtilis* o el *Bacillus licheniformis*, los cuales confieren resistencia a los antibióticos como ampicilina, kanamicina, cloramfenicol (Ejemplo 1) o tetraciclina. Otros marcadores útiles serán aparentes para la persona medianamente versada en la técnica.

[00111] En otra modalidad, la revelación proporciona una célula hospedera que comprende un polinucleótido que codifica la variante CBH II o el polipéptido quimérico, el cual polinucleótido está operativamente vinculado a una o más secuencias de control para la expresión del polipéptido en la célula hospedera. Las células hospederas que se pueden usar para expresar los polipéptidos codificados por los vectores de expresión de la revelación son bien conocidas del estado de la técnica e incluyen, pero sin limitaciones, células bacterianas, como *E. coli* y *Bacillus megaterium*; células eucarióticas, como células de levadura, células CHO y similares, células de insectos como la *Drosophila* S2 y la *Spodoptera* Sf9; células animales como CHO, COS, BHK, 293 y células del melanoma de Bowes; y células de plantas. Otras células hospederas adecuadas serán aparentes para la persona medianamente versada en la técnica. Los medios de cultivo y condiciones de crecimiento adecuadas para las células hospederas arriba descritas son bien conocidos del estado de la técnica.

[00112] La variante CBH II o los polipéptidos quiméricos de la revelación se pueden obtener usando métodos bien conocidos del estado de la técnica. Los polinucleótidos pueden ser sintetizados por técnicas recombinantes, como las previstas en Sambrook *et al.*, 2001, Molecular Cloning: A Laboratory Manual, 3ª Ed., Cold Spring Harbor Laboratory Press; y Current Protocols in Molecular Biology, Ausubel, F. ed., Greene Pub. Associates, 1998,

40

actualizados a 2007. Los polinucleótidos que codifican las enzimas o los iniciadores para la amplificación también se pueden preparar por métodos estándar de fase sólida, de acuerdo con métodos sintéticos conocidos, por ejemplo usando método de la fosforamitida descrito por Beaucage *et al.*, (1981) Tet Lett 22:1859-69 o el método descrito por Matthes *et al.*, (1984) EMBO J. 3:801-05, *p.ej.*, como es la práctica típica de los métodos sintéticos automatizados. Además, esencialmente cualquier ácido nucleico se puede obtener de cualquiera de entre una variedad de fuentes comerciales, como The Midland Certified Reagent Company, Midland, TX, The Great American Gen Company, Ramona, CA, ExpressGen Inc. Chicago, IL, Operon Technologies Inc., Alameda, CA y muchos otros.

[00113] Las enzimas obtenidas por ingeniería genética expresadas en una célula hospedera se pueden recuperar de las células o del medio de cultivo usando una o más de las técnicas bien conocidas para la purificación de proteínas, incluidas, entre otras, tratamiento de lisozimas, sonicación, filtración, separación por adición de sal, ultra-centrifugación, cromatografía y separación por afinidad (*p.ej.*, anticuerpos ligados al sustrato). Las soluciones apropiadas para el lisado y la extracción de proteínas de alta eficiencia de las bacteria, como la *E. coli*, están disponibles en el comercio con el nombre comercial CellLytic BTM de Sigma-Aldrich de St. Louis MO.

[00114] Las técnicas de cromatografía para el aislamiento de los polipéptidos incluyen, entre otras, cromatografía de fase inversa, cromatografía líquida de alta resolución, cromatografía de intercambio iónico, electroforesis en gel y cromatografía de afinidad. Las condiciones para purificar una enzima en particular dependerán, en parte, de factores como carga neta, hidrofobicidad, hidrofiliidad, peso molecular, forma molecular, etc., y serán aparentes para la persona medianamente versada en la técnica.

[00115] En los ejemplos de este documento se dan las descripciones SCHEMA dirigidas a la recombinación y síntesis de polipéptidos quiméricos, así como en Otey *et al.*, (2006), PLoS Biol. 4(5):e112; Meyer *et al.*, (2003) Protein Sci., 12:1686-1693; Solicitud estadounidense de patente No. 12/024.515, radicada en febrero 1, 2008; y Solicitud estadounidense de patente No. 12/027,885, radicada en febrero 7, 2008; las cuales referencias se incorporan a la presente en su integridad mediante referencia.

[00116] Como se dijo arriba, el polipéptido se puede usar en diversas aplicaciones, como, entre otras, generación de biocombustibles, degradación de celulosa y similares.

[00117] Los siguientes ejemplos tienen por objeto explicar un poco más, pero no limitar la anterior revelación o las reivindicaciones que se anexan.

EJEMPLOS

[00118] **Construcción del plásmido de expresión de la CBH II.** Los genes parentales y quiméricos que codifican las enzimas CBH II fueron clonados dentro del vector de expresión de levadura YEp352/PGK91-1- α ss (Figura 6). Las secuencias de ADN que codifican los dominios catalíticos de CBH II parentales y quiméricas fueron diseñados con sesgo de codón de *S. cerevisiae* usando el software GeneDesigner (DNA2.0) y sintetizados por ADN2.0. Los genes del dominio catalítico de la CBH II fueron digeridos con XhoI y KpnI, ligado dentro del vector entre los sitios XhoI y KpnI y transformados dentro de *E. coli* XL-1 Blue (Stratagene). Los genes de la CBH II fueron secuenciados usando iniciadores: CBH2L (5'-

GCTGAACGTGTCATCGGTTAC-3' (SEQ ID NO: 9) y RSQ3080 (5'-GCAACACCTGGCAATTCCTTACC-3' (SEQ ID NO: 10)). Los constructos *C terminal* His₆ progenitor y quimera de la CBH II se hicieron amplificando el gen de la CHB II con iniciador delantero *CBH2LPCR* (5'-GCTGAACGTGTCATCGTTACTTAG-3' (SEQ ID NO: 11)) e iniciadores reversos complementarios al el gen de la CHB II adecuado con His₆ codones de remate y detención. Los productos de la PCR fueron ligados, transformados y secuenciados como se dijo arriba.

[00119] **Expresión de la enzima CBH II en *S. cerevisiae*.** La cepa del *S. cerevisiae* YDR483W BY4742 (*Mata his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 ΔKRE2*, ATCC No. 4014317) se convirtió en competente usando el Kit de Transformación EZ Levadura II (Zymo Research), transformado con ADN de plásmido y sembrado en agar de uracilo sintético. Las colonias se recogieron en cultivos de la noche a la mañana en 5 mL de medio sintético de casaminoácidos dextrosa (SDCAA) (20 g/L dextrosa, 6,7 g/L base de nitrógeno de levadura Difco, 5 g/L casaminoácidos Bacto, 5,4 g/L Na₂HPO₄, 8,56 g/L NaH₂PO₄·H₂O) suplementados con 20 μg/mL de triptófano y cultivados durante la noche a 30°C, 250 rpm. Los cultivos de 5 mL se expandieron en frascos Tunair de 40 mL de SDCAA en 250 mL (Shelton Scientific) y se agitaron a 30°C, 250 rpm durante 48 horas. Los cultivos se centrifugaron y los sobrenadantes se concentraron a 500 μL, usando una celda de ultrafiltración Amicon adaptada con membrana PES de 30-kDa, para usar en pruebas de t_{1/2}. Los sobrenadantes concentrados se llevaron a 1 mM fenilmetilsulfonilfluoruro y 0,02% de NaN₃. Las proteínas CBH II rotuladas con His₆ fueron purificadas usando columnas giratorias Ni-NTA (Qiagen) de acuerdo con el protocolo del fabricante y las proteínas se intercambiaron dentro de acetato de sodio 50 mM, pH 4,8, usando columnas desalinizadoras Zeba-Spin (Pierce). La concentración de proteína purificada se determinó usando reactivo de proteínas Pierce Coomassie Plus con BSA como estándar. El análisis SDS-PAGE se llevó a cabo cargando ya sean 20 μL de sobrenadante del cultivo concentrado o aproximadamente 5 μg de enzima CBH II purificada sobre un gel de 7,5% Tris-HCl (Biorad) y coloreando con tinción segura SimplyBlue (Invitrogen). Los sobrenadantes o las proteínas purificadas CBH II se trataron con EndoH (New England Biolabs) durante 1 h a 37°C siguiendo las instrucciones del fabricante. la actividad de la enzima CBH II en sobrenadantes concentrados de cultivo de levadura se midió añadiendo 37,5 μL del sobrenadante del cultivo concentrado a 37,5 μL de PASC e incubando durante 2 h a 50°C. Los equivalentes de azúcar reductor formados se determinaron por medio de la prueba de Nelson-Somogyi como se describe más adelante.

[00120] **Vida media, actividad específica, actividad de pH e hidrólisis a largo plazo de las mediciones de celulosa.** Se preparó la celulosa hinchada en ácido fosfórico (PASC). Para intensificar la actividad de la enzima CBH II sobre el sustrato, la PASC se preincubó a una concentración de 10 g/L con 10 mg/mL de endoglucanasa de *A. niger* (Sigma) en 50mM de acetato de sodio, pH 4,8 durante 1 h a 37°C. La endoglucanasa se inactivó calentando hasta 95°C durante 15 minutos, la PASC se lavó dos veces con 50mM amortiguador de acetato y se resuspendió a 10 g/L en agua desionizada.

[00121] Los t_{1/2}s de la enzima CBH II se midieron añadiendo sobrenadante concentrado de cultivo de expresión de CBH II a acetato de sodio 50 mM, pH 4,8 a una concentración que da

un A_{520} de 0,5 medida por la prueba de azúcares reductores de Nelson-Somogyi después de la incubación con PASC tratado como se describe más adelante. 37,5 μ L las mezclas de enzima CBH II/amortiguador se inactivaron en un baño de agua a 63°C. Después de la inactivación, se añadieron 37,5 μ L PASC tratada con endoglucanasa y la hidrólisis se llevó a cabo durante 2 h a 50°C. Los sobrenadantes de la reacción se filtraron a través de placas Multiscreen HTS (Millipore). Los valores $\log(A_{520})$ de la prueba de Nelson-Somogyi, obtenidos usando un lector de microplacas SpectraMax (Molecular Devices) corregidos para los antecedentes de absorbancia, se trazaron frente al tiempo y las vidas medias de la enzima CBH II obtenida por regresión lineal usando Microsoft Excel.

[00122] Para las mediciones de la actividad específica, la enzima CBH II purificada se añadió a PASC para dar un volumen final de reacción de 75 μ L de acetato de sodio 25 mM, pH 4,8, con 5 g/L de PASC y una concentración de la enzima CBH II de 3 mg enzima/g PASC. La incubación procedió durante 2 h en un baño de agua a 50°C y se determinó la concentración de azúcar reductor. Para las mediciones de perfil de pH/actividad, se añadió enzima CBH II purificada a una concentración de 300 μ g/g PASC en un volumen de reacción de 75 μ L. Las reacciones fueron amortiguadas con citrato de sodio 12,5 mM/fosfato de sodio 12,5 mM, durante 16 h a 50°C y se determinó el azúcar reductor. La hidrólisis a largo plazo de las mediciones de celulosa se llevó a cabo con volúmenes de 300 μ L de 1 g/L de PASC tratado en acetato de sodio 100 mM, pH 4,8, NaCl 20 mM. Se añadió enzima CBH II purificada a 100 μ g/g de PASC y las reacciones se realizaron en baños de agua durante 40 h antes de la determinación de azúcar reductor.

[00123] Se sintetizaron cinco genes progenitores candidatos que codifican las enzimas CBH II con sesgo de codón de *S. cerevisiae*. Los cinco contenían secuencias codificadoras N-terminal idénticas, en donde los residuos 1-89 corresponden al módulo fijador de celulosa (CBM), una región conectora flexible y los cinco residuos N-terminal del dominio catalítico de *H. jecorina*. Dos de las enzimas CBH II candidatas, de *Humicola insolens* y *Chaetomium termophilum*, fueron segregadas de *S. cerevisiae* en niveles mucho más altos que las otros tres, de *Hipocrea jecorina*, *Phanerochaete chrysosporium* y *Talaromyces emersonii* (Figura 1). Dado que las bandas en el gel SDS-PAGE de los tres progenitores candidatos débilmente expresados fueron difíciles de discernir, se llevaron a cabo pruebas de actividad en las cuales los sobrenadantes del cultivo concentrado se incubaron con celulosa hinchada en ácido fosfórico (PASC) para confirmar la presencia de celulasas activa. Los valores para el azúcar reductor formado, que se presentan en la Figura 1, confirmaron la presencia de CBH II activa en sobrenadantes concentrados de cultivo de *S. cerevisiae* para todas las enzimas excepto CBH II de *T. emersonii*. Se eligieron las secuencias de *H. insolens* y *C. termophilum* para recombinarlas con la enzima CBH II de hongo más relevante industrialmente, del *H. jecorina*. La identidad respectiva de las secuencias de los dominios catalíticos son 64% (1:2), 66% (2:3) y 82% (1:3), en donde el *H. insolens* es el progenitor 1, *H. jecorina* es el progenitor 2 y *C. termophilum* es el progenitor 3. Estos dominios catalíticos respectivos contienen 360, 358 y 359 residuos de aminoácidos.

[00124] La expresión de la proteína heteróloga en el hongo filamentoso *H. jecorina*, el organismo que se usa con mayor frecuencia para producir celulasas para aplicaciones

23

industriales, es mucho más ardua que en el *Saccharomyces cerevisiae*. La secreción observada de CBH II de *H. jecorina* del *S. cerevisiae* motivó la elección de este huésped heterólogo. Para hiperglucosilación, de la que se ha informado que reduce la actividad de las celulasas recombinantes, los genes recombinantes de la CBH II se expresaron en una cepa *dkRE2* de *S. cerevisiae* deficiente en glucosilación. Se espera que esta cepa conecte oligómeros de manosa más pequeños a los sitios de glucosilación ligados al N y ligados al O que las cepas de tipo silvestre, las cuales imitan más de cerca la glucosilación de la enzima CBH II nativa producida por la *H. jecorina*. El análisis en gel SDS-PAGE de las proteínas CBH II, con y sin tratamiento EndoH para remover estructuras altas en manosa, mostró que el tratamiento EndoH no aumentó la movilidad electroforética de las enzimas segregadas por esta cepa, lo cual confirmó la ausencia de los grupos funcionales de manosa ramificada que las cepas de tipo silvestre de *S. cerevisiae* conectan a los sitios de glucosilación en las proteínas recombinantes.

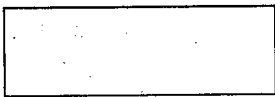
[00125] La estructura de alta resolución de *H. insolens* (pdb entrada 1ocn) se usó como plantilla para el SCHEMA para identificar los contactos que podrían romperse tras la recombinación. El RASPP devolvió cuatro bibliotecas candidatas, cada una con <E> por debajo de 15. Todas las bibliotecas candidatas tienen <E> más bajo que las bibliotecas quimeras construidas previamente, lo cual sugiere que se podría obtener una fracción aceptable de quimeras activas plegadas para un <m> relativamente alto. La diversidad de la secuencia de la quimera se maximizó seleccionando las fronteras de los bloques que conducen a la <m> = 50 más grande. Los bloques para este diseño se ilustran en la Figura 2B y se detallan en la Tabla 2.

Tabla 2: Alineación ClustalW de múltiples secuencias para los dominios catalíticos de la enzima CBH II progenitora. Los bloques 2, 4, 6 y 8 se denotan por casillas y sombreado gris. Los bloques 1, 3, 5 y 7 no están sombreados. (H. inso: SEQ ID NO: 2; H. Jeco: SEQ ID NO: 4 y C. Ther: SEQ ID NO: 6).

H.inso
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDNRNVTVDTLVQ 60
C.ther
GNPFSGVQLWANTYYSSEVHTLAIPSL-
PELAAKAAKVAEVPVSFQWLDNRNVTVDTLFSG 59
H.jeco
GNPFVGVTPWANAYYASEVSSLAIPSLT-GAMATAAAAVAKVPSFMWLD-
LDKTPLMEQ 58
**** * * * * * . * * * . * * . * * * * * . * .

H.inso
TLSEIREANQAGANPQYAAQIVVYDLPDRDCAAAASNGEWAIANNGVNNYKAYINRIREI 120
C.ther
TLAEIRAANQRGANPPYAGIFVVYDLPDRDCAAAASNGEWSIANNGANNYKRYIDRIEL 119
H.jeco
TLADIRTANKNGN--
YAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDTIRQI 116
* * * * * . * * * . * * * * * * * * * . * * . * * * . * * .

H.inso
LISFSDVRTILVIEPDSLNMVTNMNVPKCSGAASTYRELTIIYALKQLDLPVAMYMDAG 180
C.ther
LIQYSDIRTILVIEPDSLNMVTNMNVQKCSNAASTYKELTVYALKQLNLPVAMYMDAG 179
H.jeco
VVEYSDIRTLLVIEPDSLNLVTNLGTPKCANAQSAYLECINYAVTQLNLPNVAMYLDAG 176
* * * * * . * * * . * * * * * * * * * . * * . * * * . * * .



H.inso F- 360
C.ther F- 359
H.jeco FL 358

[00126] El dominio catalítico de la CBH II del *H. insolens* tiene una estructura de barril α/β en la cual las ocho hélices definen el perímetro del barril y siete hojas β paralelas del sitio activo (Figura 2A). Dos asas extendidas forman un techo sobre el sitio activo, creando un túnel a través del cual el sustrato de las cadenas de celulosa pasa durante la hidrólisis. Cinco de las siete fronteras de bloque caen entre los elementos de la estructura secundaria, en tanto que el bloque 4 comienza y termina en la mitad de hélices α consecutivas (Figuras 2A, 2B). La mayoría de los contactos de cadena lateral interbloques tiene lugar entre bloques que son adyacentes en la estructura primaria (Figura 2C).

[00127] Un juego de muestras de 48 genes de quimera se diseñó como tres juegos de 16 quimeras que tienen cinco bloques de un progenitor y tres bloques bien sea de uno o de los dos progenitores restantes (Tabla 3); las secuencias se seleccionaron para igualar la representación de cada progenitor en cada posición de bloque. Los genes correspondientes fueron sintetizados y expresados.

Tabla 3: Secuencias del juego de muestras de las quimeras de la enzima CBH II.

Inactiva	Activa
13121211	11332333
12122221	21131311
33332321	31212111

65

33321331 22232132
 21322232 33213332
 21112113 23233133
 31121121 13231111
 32312222 12213111
 23223223 31311112
 31313323 11113132
 32121222 13111313
 12121113 21311131
 22133222 11313121
 33222333 21223122
 11131231 22212231
 11112321 23231222
 12111212 32333113
 31222212 12133333
 22322312 13333232
 12222213 33123313
 12221122 21333331
 22212323 23311333
 23222321 33133132
 32333223
 33331213

[00128] Veintitrés de los 48 sobrenadantes del cultivo concentrado del conjunto de muestras de *S. cerevisiae* evidenciaron actividad hidrolítica hacia la PASC. Estos resultados hacen pensar que miles de las 6.561 posibles secuencias de quimeras de CBH II (ver p.ej., la Tabla 1) codifican enzimas activas. Las 23 quimeras activas CBH II del juego de muestras muestran considerable diversidad de secuencia, y difieren del progenitor más cercano a la secuencia y entre sí en por lo menos 23 y 36 sustituciones de aminoácidos y hasta 54 y 123, respectivamente. Su promedio de nivel de mutación $\langle m \rangle$ es de 36.

[00129] Como Meyer *et al.*, encontraron correlaciones entre E , m y la probabilidad de que una quimera esté plegada y activa, se hizo el análisis de si existían correlaciones similares para el conjunto de muestras de quimeras de CBH II. La cantidad de actividad de la enzima CBH II en sobrenadantes concentrados de cultivo de expresión, medido determinando la actividad sobre la PASC, se correlacionó con la intensidad de las bandas del CBH II en los geles SDS-PAGE (Figura 1). Igual que sucedió con la CBH II progenitora de *H. jecorina*, se podía detectar actividad para algunas quimeras de CBH II con bandas de gel indetectables. No hubo observaciones de quimeras de CBH II que presentaran bandas de gel pero que carecieran de actividad. La probabilidad de que una quimera CBH II fuese segregadas en forma activa fue inversamente proporcional tanto a E como a m (Figura 3).

[00130] Las vidas medias de inactivación térmica ($t_{1/2}$) se midieron a 63°C para los sobrenadantes del cultivo concentrado las enzimas CBH II progenitoras y quiméricas activas. Las vidas medias de la CBH II progenitora de *H. insolens*, *H. jecorina* y *C. thermophilum* fueron de 95, 2 y 25 minutos, respectivamente (Tabla 1). El juego de muestras de quimeras activas evidenció un amplio intervalo de vidas medias, de menos de 1 minuto a más de 3.000. cinco de las 23 quimeras activas tuvieron vidas medias mayores que las del progenitor más termoestable, CBH II del *H. insolens*.

[00131] Con la intención de construir un modelo predictivo cuantitativo para la vida media de una quimera CBH II, se usaron cinco diferentes algoritmos de modelación de datos de

46

regresión lineal (Tabla 4). Cada algoritmo se usó para construir un modelo relacionando las composiciones de bloques de cada quimera CBH II del juego de muestras y los progenitores con el $\log(t_{1/2})$. Estos modelos produjeron valores de peso de termoestabilidad que cuantificaron la contribución de un bloque al $\log(t_{1/2})$. Para los cinco algoritmos de modelado, este proceso se repitió 1.000 veces, omitiendo dos secuencias seleccionadas al azar de cada cálculo, de manera que cada algoritmo produjo 1.000 valores de peso para cada uno de los 24 bloques. La media y la desviación estándar (DE) se calcularon para cada peso de termoestabilidad del bloque. La exactitud predictiva de cada algoritmo modelo se valoró midiendo qué tan bien predecía cada modelo los $t_{1/2}$ s de las dos secuencias omitidas. La correlación entre los valores medidos y predichos para las 1.000 iteraciones del algoritmo es la puntuación de validación cruzada del algoritmo modelo. Para los cinco modelos, las puntuaciones de validación cruzada (X-val) fueron menores o iguales a 0,57 (Tabla 4), lo cual indica que el modelo de regresión lineal no podría aplicarse a este pequeño conjunto de datos de $t_{1/2}$ de 23 quimeras para la predicción cuantitativa de la vida media de la quimera CBH II.

Tabla 4: Valores de validación cruzada valores para la aplicación de 5 algoritmos de regresión lineal las puntuaciones de estabilidad del bloque de la enzima quimera CBH II. Abreviaturas del algoritmo: regresión de la cresta (RR), regresión parcial del mínimo cuadrado (PLSR), regresión de la máquina del vector de soporte (SVMR), programación lineal de la regresión de la máquina del vector de soporte (LPSVMR) y programación lineal de la amplificación de la regresión (LPBoostR).

Método	Cresta	PLS	SVMR	LSVM	LPBoost
X-val	0.56	0.55	0.50	0.42	0.43

[00132] El modelo de regresión lineal se usó para clasificar cualitativamente los bloques como estabilizador, desestabilizador o neutro. EL impacto de cada bloque sobre la termoestabilidad de la quimera se caracterizó usando un sistema de calificación que da cuenta de la contribución a la termoestabilidad determinada por cada uno de los algoritmos de regresión. Para cada algoritmo, los bloques con un valor de peso de termoestabilidad más de 1 DE por encima del valor neutro recibieron una calificación de “+1”, los bloques dentro de 1 DE del neutro se calificaron con cero y los bloques 1 o más DE por debajo del neutro se calificaron con “-1”. Se obtuvo una “puntuación de estabilidad” para cada bloque sumando las calificaciones de estabilidad de 1, 0, -1 de cada uno de los cinco modelos. En la Tabla 5 se resumen las puntuaciones para cada bloque. Se identificó a los Bloque 1/progenitor 1 (B1P1), B6P3, B7P3 y B8P2 como los que tuvieron el máximo efecto estabilizador, y se encontró que B1P3, B2P1, B3P2, B6P2, B7P1, B7P2 y B8P3 fueron los bloques desestabilizadores más fuertes de todos.

Tabla 5: Resultados de la clasificación cualitativa de los bloques generada por cinco algoritmos de regresión lineal¹ para el juego de muestras de las enzimas quimeras de CBH II. La puntuación de +1 denota un bloque con peso de termoestabilidad (métrica sin dimensiones para la contribución de un bloque a la termoestabilidad de la quimera) mayor de una desviación estándar por encima del valor neutro (estabilizador), la puntuación de 0 denota un bloque con peso dentro de una desviación estándar del valor neutro y -1 denota un bloque con un peso de más de una desviación estándar por debajo del valor neutro (desestabilizador).

27

Block	Ridge	PLS	SVMR	LSVM	LPBoost	Sum
B1P1	1	0	1	1	0	3
B1P2	0	0	0	-1	0	-1
B1P3	-1	0	-1	-1	-1	-4
B2P1	-1	0	0	-1	-1	-3
B2P2	1	0	0	0	0	1
B2P3	1	0	0	0	0	1
B3P1	1	0	1	0	0	2
B3P2	-1	0	-1	-1	-1	-4
B3P3	1	0	1	0	0	2
B4P1	0	0	0	0	0	0
B4P2	0	0	0	0	0	0
B4P3	0	0	0	-1	0	-1
B5P1	0	0	0	0	0	0
B5P2	0	0	0	0	-1	-1
B5P3	-1	0	0	-1	0	-2
B6P1	1	0	0	-1	-1	-1
B6P2	-1	0	-1	-1	-1	-4
B6P3	1	1	1	1	1	5
B7P1	-1	0	-1	-1	-1	-4
B7P2	-1	0	-1	-1	-1	-4
B7P3	1	0	1	1	1	4
B8P1	1	0	1	-1	0	1
B8P2	1	0	1	1	0	3
B8P3	-1	0	-1	-1	-1	-4

[00133] Se sintetizó un segundo juego de genes que codifica las quimeras de las enzimas CBH II con el fin de validar los bloques estabilizadores predichos e identificar celulasas más termoestables que el progenitor más estable. Las 24 quimeras incluidas en este conjunto de validación (Tabla 6) estaban desprovistas de los siete bloques de los que se predecía que eran los más desestabilizadores y enriquecidas con los cuatro bloques más estabilizadores, en donde la representación se sesgó hacia puntuaciones de estabilidad más altas. Además, la quimera “HJPlus” 12222332 se construyó sustituyendo los bloques más estabilizadores predichos dentro de la enzima CBH II de *H. jecorina* (progenitor 2).

Tabla 6: Secuencias del conjunto de validación de 24 quimeras de las enzimas CBH II, nueve de las cuales se expresaron de forma activa.

Inactiva	Activa
12122132	12111131
12132332	12132331
12122331	12131331
12112132	12332331
13122332	13332331
13111132	13331332
13111332	13311331
13322332	13311332
22122132	22311331
22322132	
22311332	
23111332	
23321131	
23321332	
23321331	

cb
X

[00134] Los sobrenadantes concentrados de los cultivos de expresión de *S. cerevisiae* para nueve de las 24 quimeras del conjunto de validación, así como la quimera HJPlus, mostraron actividad hacia PASC (Tabla 6). De las 15 quimeras para las cuales no se detectó actividad, nueve contenían el bloque B4P2. De las 16 quimeras que contenían B4P2 en el juego de muestras inicial, solamente una mostró actividad hacia PASC. Sumadas a los dos juegos de quimera sets y HJPlus, solamente dos de las 26 quimeras que exhibían B4P2 fueron activas, lo cual indicó que este bloque en particular es altamente nocivo para la expresión de la celulasa activa en *S. cerevisiae*.

[00135] Se evaluaron las estabilidades de las 10 enzimas quiméricas funcionales CBH II del conjunto de validación. Dado que las enzimas estables ya tenían vidas medias de más de 50 horas, se usó la actividad hidrolítica residual hacia la PASC después de una inactivación térmica de 12 horas a 63°C como la métrica para la evaluación preliminar. Esta incubación de 12 horas produjo una disminución cuantificable de la actividad del conjunto de la quimera más termoestable de la muestra, 11113132, e inactivó por completo la CBH II progenitora termoestable del *H. insolens*. Las diez quimeras funcionales del conjunto de validación retuvieron una mayor fracción de sus actividades que el progenitor más estable, la CBH II del *H. insolens*.

[00136] Se analizaron las actividades de las quimeras termoestables seleccionadas usando enzimas purificadas. Las enzimas CBH II progenitoras y tres quimeras termoestables, la quimera más termoestable del juego de muestras 11113132, la quimera más termoestable del conjunto de validación 13311332 y la quimera HJPlus 12223332, se expresaron con rótulos His₆ en el C terminal y se purificaron. Para minimizar la inactivación térmica de las enzimas CBH II durante la prueba de actividad, usamos una incubación más corta, de dos horas, con el sustrato PASC a 50°C, pH 4,8. Como se muestra en la Tabla 3, el actividades específicas de las CBH II progenitora y quimérica estuvieron dentro de un factor de cuatro de la enzima CBH II progenitora más activa, de *H. jecorina*. La actividad específica de HJPlus fue mayor que la de todas las demás enzimas CBH II examinadas, excepto la CBH II de *H. jecorina*.

[00137] La dependencia del pH de la actividad de celulasa también es importante, por cuanto un perfil de pH/actividad amplio permitiría el uso de una quimera CBH II en un intervalo más amplio de condiciones potenciales de hidrólisis de la celulosa. Se ha observado que la CBH II de *H. jecorina* tiene actividad óptima en el intervalo de pH de 4 a 6, con una actividad marcadamente reducida por fuera de estos valores.¹⁶ La Figura 4 muestra que las enzimas CBH II de *H. insolens* y *C. thermophilum* y las tres quimeras termoestables purificadas de la CBH II tienen perfiles de pH/actividad que considerablemente más amplios que el de la CBH II de *H. jecorina*. Aunque Liu *et al.*, informan de un pH óptimo de 4 para la CBH II de *C. thermophilum*, el pH óptimo de la enzima recombinante aquí estuvo cerca de 7. La CBH II nativa de *H. insolens* tiene un amplio perfil de pH/actividad, con máxima actividad alrededor del pH 9 y aproximadamente 60% de esta actividad máxima a pH 4. Se observó un perfil amplio similar para la enzima recombinante. La quimera HJPlus tiene un perfil de pH/actividad mucho más amplio que la CBH II de *H. jecorina*, que muestra una dependencia del pH similar a las otras dos las enzimas CBH II progenitoras.

49

[00138] Lograr actividad a alta temperatura y retención de la actividad en periodos prolongados de tiempo son dos motivaciones primarias para obtener por ingeniería genética enzimas CBH II altamente estables. El desempeño de quimeras termoestables de CBH II en la hidrólisis de la celulosa se examinó a lo largo de un intervalo de temperaturas durante 40 horas. Como se muestra en la Figura 5, las tres quimeras termoestables fueron activas sobre PASC a temperaturas más altas que las enzimas CBH II progenitoras. Las quimeras retuvieron su actividad a 70°C, en tanto que la CBH II de *H. jecorina* no hidrolizó la PASC por encima de 57°C y la enzima estable de *H. insolens* no mostró hidrólisis por encima de 63°C. La actividad de HJPlus en las pruebas de hidrólisis de la celulosa a largo plazo superó la de todos los progenitores a sus respectivas temperaturas óptimas.

[00139] La biblioteca de CBH II tiene menos interrupciones potenciales por varias razones. Además de la más alta identidad de las secuencias de la CBH II progenitora, la topología en barril del pliegue de la CBH II limita el número de contactos de largo rango que pueden romperse por la recombinación. Los contactos entre bloques (átomos pesados dentro de 4,5 Å) comprenden solamente 27% (503/1831) del total en un mapa de contactos derivado de la estructura 1ocn del *H. insolens*. Cuando solamente hay contactos para los cuales son posibles los novedosos pares de residuos en las quimeras, el total de inter bloques se reduce a 23% (68/294). Además, la mayor parte de estas interacciones se dan entre residuos sobre la superficie de la proteína y la posibilidad de tamizaje del solvente reduce aún más las probabilidades de interacciones drásticas de interrupción residuo-residuo (Figura 14a). Una excepción, una interacción enterrada entre las posiciones 176 y 256, se ilustra en la Figura 14b. En este sitio, las quimeras tienen B6P2 y bien sea los pares Met173:Trp253 B5P1 o B5P3 (aminoácidos más grandes que los pares parentales Met176:Phe256 o Leu173:Trp253). No obstante, la inspección de los modelos cristalográficos parentales, se consideró que un choque estérico en esta posición es improbable debido al desplazamiento en la porción de la columna vertebral de la proteína que dispone Trp253 y la flexibilidad intrínseca de las cadenas laterales Met. Cabe anotar que una quimera caracterizada sigue este patrón (13333232) y es más estable que las progenitoras (67°C), de conformidad con el caso del modelo de regresión (68°C).

[00140] Otro mecanismo por el cual podría surgir el acoplamiento, la divergencia estructural de los bloques, no depende de la presencia de pares novedosos de residuos en las interfases de los bloques. Antes bien, como las secuencias parentales divergen, las estructuras intrínsecas de los bloques pueden divergir, entorpeciendo los trasplantes modulares de los bloques. En el caso de la biblioteca de CBH II, la alta identidad de los pares de progenitores de los valores de la secuencia (82%, 66% y 64%) sugiere que solo son probables las desviaciones menores de la estructura (<1 Å RMSD). Esta posibilidad se puede evaluar comparando las estructuras cristalográficas de las CBH II de *H. insolens* y *H. jecorina* (la CBH II de *C. thermophilum* carece de una estructura cristalina pero es 82% idéntica a la de *H. insolens*). La alineación de los bloques a partir de las estructuras para cada progenitor (1ocn y 1cb2), genera bajos valores de carbono alfa RMSD (0,5, 0,5, 0,6, 0,5, 0,3, 0,7, 0,3 y 0,4 Å RMSD). Los bloques de *H. jecorina* superpuestos sobre los de *H. insolens* se ilustran en la Figura complementaria 5c. Se revisaron los efectos dependientes del contexto a una recombinación

5

estructural *silico*, empalmando cada bloque alineado sobre la estructura opuesta del huésped. Es posible construir modelos estructurales que no colisionen (carbonos alfa > 3 Å de separación) para todas las quimeras de sustitución de bloques únicos (p.ej., 11112111 o 22122222), con excepción de una colisión (2,65 Å) cuando se usa B7P2 (11111121) debido a la inserción de Asn entre los bloques 6 y 7 (Figura 14D).

[00141] Se llevaron a cabo otros experimentos para determinar las contribuciones de varios bloques/segmentos a la estabilidad de la quimera y la mejor termoestabilidad y/o estabilidad del pH. Los genes parentales y quiméricos que codifican las enzimas CBH II se clonaron dentro del vector de expresión de levadura YEp352/PGK91-1-*ass* y expresión en medio sintético de dextrosa casaminoácidos (SDCAA). Para las pruebas de actividad Avicel, los sobrenadantes de cultivo de levadura peptona dextrosa (YPD) se llevaron a 1 mM fenilmetilsulfonilfluoruro y 0,02% NaN_3 y se usaron sin concentrar. La actividad de la enzima CBH II en sobrenadantes concentrados de cultivo de levadura SDCAA se midió añadiendo diluciones de sobrenadante del cultivo concentrado a 37,5 μL de PASC y 225 μL de acetato de sodio 50 mM, pH 4,8 e incubando durante 2 h a 50°C. Los equivalentes de azúcar reductor formados se determinaron por medio de la prueba de Nelson-Somogyi.

[00142] Los valores de T_{50} de la enzima CBH II se midieron añadiendo sobrenadante concentrado de cultivo de expresión de CBH II SDCAA a acetato de sodio 50 mM, pH 4,8 a una concentración que dio un A_{520} de 0,5 medido en la prueba de azúcares reductores de Nelson-Somogyi después de incubación con PASC tratada con endoglucanasa. 200 μL de la mezcla de enzima CBH II/amortiguador se incubaron en un baño de agua a la temperatura de interés durante 10 minutos. Después de la incubación se añadieron 37,5 μL de PASC tratada con endoglucanasa y 62,5 μL de acetato de sodio 50 mM y la hidrólisis se llevó a cabo durante 2 h a 50°C. La temperatura de incubación a la cual la enzima perdió la mitad de su actividad se determinó por interpolación lineal de los valores A_{520} de la prueba de Nelson-Somogyi trazados frente a la temperatura.

[00143] Para las mediciones de hidrólisis Avicel PH101 (Fluka) a largo plazo, se incubaron 0,3 μg de CBH II purificada con 3 mg de Avicel en 270 μL de acetato de sodio 50 mM, pH 4,8, en tubos de PCR puestos en un baño de agua durante 16 horas. Los tubos se enfriaron en un baño de agua a temperatura ambiente durante 10 minutos, se centrifugaron a 1000 g durante 10 minutos y se extrajeron los sobrenadantes para el análisis de azúcar reductor.

[00144] Para la estimación de actividad de CBH II en sobrenadantes de cultivo de expresión YPD, se añadieron volúmenes del sobrenadante que oscilaron de 2 mL a 40 mL a 800 μL de 33 mg/mL de Avicel suspendido en acetato de sodio 50 mM, pH 4,8 en tubos cónicos. Se dejó que las CBH II se unieran al Avicel a 4°C durante una hora, se centrifugaron a 2000 g durante 2 minutos y se lavaron dos veces con acetato de sodio 50 mM, pH 4,8. Después de la segunda lavada, el Avicel ligado a la CBH II se resuspendió en 2,75 mL de amortiguador de acetato de sodio, se dividió en alícuotas de 270 μL y se incubó a 50°C durante 2,5 horas. La centrifugación y el análisis de azúcar reductor del sobrenadante se llevaron a cabo como se explicó arriba.

[00145] El paquete de regresión lineal en Matemática se usó para adaptar los datos de T_{50} de la quimera CBH II a un modelo aditivo de bloques de 17 parámetros, y también se usó para el análisis de validación cruzada. Los efectos del bloque se informan con respecto a un estado

de progenitor 1 (CBH II de *H. insolens*) de referencia en donde 16 parámetros representan la sustitución de cada uno de los 8 bloques de los progenitores 2 y 3.

[00146] Los valores de T_{50} , definidos aquí como la temperatura a la cual una enzima pierde 50% de su actividad durante una incubación de diez minutos, se determinaron para las tres celobiohidrolasas progenitoras, 33 quimeras activas de CBH II de experimentos previos y 18 quimeras adicionales que el modelo de estabilidad cualitativa predijo que estarían entre las más termoestable, es decir, que no contienen ninguno de los 7 bloques desestabilizadores predichos y ya sean 3 o 4 de los 4 bloques estabilizadores predichos. Las 51 secuencias de quimera se enumeran en la Tabla 8. La repetición del cultivo y la concentración de todas las quimeras termoestables predichas previamente clasificadas como no segregadas permitió obtener cantidades suficientes de CBH II 12112132, 13111132 y 13322332 para la determinación de T_{50} . El juego completo de los valores de T_{50} para las quimeras y CBH II progenitoras se da en la Tabla 8. Las secuencias de aminoácidos para todas estas CBH II aparecen en la Tabla 7. Todas las 31 quimeras termoestables predichas examinadas tienen valores de T_{50} más de dos grados más alta que la de la enzima progenitora más termoestable (64,8°C). En la Tabla también se identifica el residuo Cys en el bloque/dominio 7 que puede mutarse a una Ser para dar una termoestabilidad aumentada. Según esto, la revelación proporciona polipéptidos de cualquiera de las siguientes secuencias en donde la Cys subrayada/en cursiva/en negrita está sustituido por un residuo de Ser y en donde el polipéptido que se traduce tiene mejor termoestabilidad que una enzima de tipo silvestre.

TABLA 7: Secuencias de aminoácidos para los dominios catalíticos de la CBH II progenitora y quimérica que se muestran en la Tabla 8. La Tabla también incluye el dominio catalítico para la CBH II de *P. chrysosporium*. Todas las CBH II recombinantes comparten el CBM N-terminal y el conector de la CBH II nativa de *H. jecorina*, CSSVWGQCGGQNWSGPTCCASGSTCVYSNDYYSQCLPGAASSSSSTRAASTTSRVSPPTSRSSSATPPPGSTTTRVPPVSGGTATYS (SEQ ID NO:8).

Progenitor 1 (H. insolens) (SEQ ID NO: 2)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDRNVTVDTLVQTLSEIREANQAGANPQYAAQIVVYDLPDRDCAAAASNGEWAIANNGVNKYKAYINRIREILISFSDVRTILVIEPDSLNMVTNMNVKCSGAASTYRELTIYALKQLDLPHVAMYMDAGHAGWLGWPANIQPAAELFAKIYEDAGKPRAVRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHYIEAFRPLLEARGFPAQFIVDQGRSGKQPTGQKEWGHWCNAIGTGFGMRPTANTGHQYVDAFVWVKPGGECDGTSDTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF

Progenitor 2 (H. jecorina) (SEQ ID NO: 4)
GNPFVGVTPWANAYYASEVSSLAIPSLTGAMATAAAVAKVPSFMWLDTLDKTPLMEQTLADIRANKNGGNYAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDTIRQIVVEYSDIRTLLVIEPDSLNLVTNLGTPKCANASAYLECIYAVTQLNLPNVAMYLDAGHAGWLGWPANQDPAAQLFANVYKNASSPRALRGLATNVANYNGWNITSPPSYTQGNVYNEKLYIHAIGPLLANHGWSNAFFITDQGRSGKQPTGQQQWGDWCNVIGTGFGIRPSANTGDSLLDSFVWVKPGGECDGTS DSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

Progenitor 3 (C. thermophilum) (SEQ ID NO: 6)
GNPFSGVQLWANTYYSEVHTLAIPSLSPELAAKAAKVAEPVSFQWLDRNVTVDTLFSGTLAEIRANQRGANPPYAGIFVVYDLPDRDCAAAASNGEWSIANNGANNKYKRYIDRIRELLIQYSDIRTLVIEPDSLNMVTNMNVQKCSNAASTYKELTVYALKQLNLPNVAMYMDAGHAGWLGWPANIQPAAELFAQIYRDAGRPAAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQGFDAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDAFVWVKPGGESDGTSDTSAARYDYHCGLSDALTPAPEAGQWFQAYFEQLLINANPPF

P. chrysosporium CBH II (SEQ ID NO: 12)

2
5

NNPWTGFQIFLSPYYANEVAAAAKQITDPTLSSKAASVANIPTFTWLDSVAKIPDLGTYLASASA
LGKSTGTKQLVQIVYDLPDRDCAAKASNGEFSIANNGQANYENYIDQIVAQIQQFPDVRVAVI
EPDSLNLVTLNLNVQKCANAKTTYLACVNYALTNLAKVGVYMYMDAGHAGWLGWSPANLSPAA
QLFTQVWQNAGKSPFIKGLATNVANYNALQAASPDPTQGNPNYDEIHYINALAPLLQQAGWD
ATFIVDQGRSGVQNIQQWGDWCNIKGAGFGTRPTTNTGSQFIDSIVVWKPGGECDGTSNSS
SPRYDSTCSLPDAAQPAPEAGTWFAQYFQTLVSAANPPL

32333113 (SEQ ID NO: 13)
GNPFSGVQLWANTYYSSEVHTLAIPSLSPELAAKAAKVAEVPSFMWLDTLDKTPLMEQTLADIR
TANKNGANPPYAGIFVYDLPDRDCAAAASNGEWSIANNGANNYKRYIDRIRELLIQYSDIRTL
VIEPDSLNMVTNMNVQKCSNAASTYKELTVYALKQLNLPHVAMYMDAGHAGWLGWSPANIQP
AAELFAQIYRDAGRPAAVRGLATNVANYNAWSVSSPPPYTSPNPYDEKHIEAFRPLLEARG
FPAQFIVDQGRSGKQPTGQKEWGHWCNAIGTGFGMRPTANTGHQYVDAFVWVKPGGECDG
TSDTSAARYDYHCGLSDALTPAPEAGQWFQAYFEQLLINANPPF

13111313 (SEQ ID NO: 14)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDRNVTVDTLFSGTLA
EIRANQRGANPQYAAQIVYDLPDRDCAAAASNGEWAIANNGVNNYKAYINRIREILISFSDVR
TILVIEPDSLNMVTNMNVKCSGAASTYRELTIALKQLDLPHVAMYMDAGHAGWLGWSPANI
QPAAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPYDEKHIEAFAPLLRNQ
GFDAQFIVDQGRSGKQPTGQKEWGHWCNAIGTGFGMRPTANTGHQYVDAFVWVKPGGECD
GTSDTSAARYDYHCGLSDALTPAPEAGQWFQAYFEQLLINANPPF

11313121 (SEQ ID NO: 15)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDRNVTVDTLVQTLSE
EIREANQAGANPPYAGIFVYDLPDRDCAAAASNGEWSIANNGANNYKRYIDRIREILISFSDVR
TILVIEPDSLNMVTNMNVKCSGAASTYKELTVYALKQLNLPHVAMYMDAGHAGWLGWSPANI
QPAAELFAQIYRDAGRPAAVRGLATNVANYNAWSVSSPPPYTSPNPYDEKHIEAFRPLLEA
RGFPNAFFITDQGRSGKQPTGQQQWGDWCNVIGTGFGIRPSANTGDSLLDSFVWVKPGGEC
DGTSDDTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF

21131311 (SEQ ID NO: 16)
GNPFVGVTPWANAYYASEVSSLAIPSLTGAMATAAAAVAKVPSFQWLDRNVTVDTLVQTLSEI
REANQAGANPQYAAQIVYDLPDRDCAAAASNGEWAIANNGVNNYKAYINRIRELLIQYSDIRTI
LVIEPDSLNMVTNMNVQKCSNAASTYRELTIALKQLDLPHVAMYMDAGHAGWLGWSPANIQP
AAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPYDEKHIEAFAPLLRNQGF
DAQFIVDQGRSGKQPTGQKEWGHWCNAIGTGFGMRPTANTGHQYVDAFVWVKPGGECDGT
SDTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF

31212111 (SEQ ID NO: 17)
GNPFSGVQLWANTYYSSEVHTLAIPSLSPELAAKAAKVAEVPSFQWLDRNVTVDTLVQTLSEI
REANQAGGNYAGQFVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDRIREILISFSDVRTILV
EPDSLNMVTNMNVKCSGAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWSPANQDPA
AQLFANVYKNASSPRALRGLATNVANYNAWSVSSPPPYTSPNPYDEKHIEAFRPLLEARGF
PAQFIVDQGRSGKQPTGQKEWGHWCNAIGTGFGMRPTANTGHQYVDAFVWVKPGGECDGT
SDTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF

23233133 (SEQ ID NO: 18)
GNPFVGVTPWANAYYASEVSSLAIPSLTGAMATAAAAVAKVPSFQWLDRNVTVDTLFSGTLAEI
RAANQRGGNYAGQFVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDRIRELLIQYSDIRTLV
IEPDSLNMVTNMNVQKCSNAASTYKELTVYALKQLNLPHVAMYMDAGHAGWLGWSPANIQPA
AELFAQIYRDAGRPAAVRGLATNVANYNAWSVSSPPPYTSPNPYDEKHIEAFRPLLEARGF
PAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGTS
D TSAARYDYHCGLSDALTPAPEAGQWFQAYFEQLLINANPPF

31311112 (SEQ ID NO: 19)
GNPFSGVQLWANTYYSSEVHTLAIPSLSPELAAKAAKVAEVPSFQWLDRNVTVDTLVQTLSE
IREANQAGANPPYAGIFVYDLPDRDCAAAASNGEWSIANNGANNYKRYIDRIREILISFSDVRTI
LVIEPDSLNMVTNMNVKCSGAASTYRELTIALKQLDLPHVAMYMDAGHAGWLGWSPANIQP
AAELFAKIYEDAGKPRAVRGLATNVANYNAWSVSSPPPYTSPNPYDEKHIEAFRPLLEARGF
P AQFIVDQGRSGKQPTGQKEWGHWCNAIGTGFGMRPTANTGHQYVDAFVWVKPGGECDG
TSDSSAPRFD SHCALPDALQPAPQAGAWFAQYFVQLLTNANPSFL

3

22212231 (SEQ ID NO: 20)
GNPFVGVTPWANAYYASEVSSLAIPSLTGAMATAAAAVAKVPSFMWLDTLDKTPLMEQTLADI
RTANKNGGNYAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDRIREILISFSDVRTILVI
EPDSLANMVTNMNVPKCSGAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQDPA
AQLFANVYKNASSPRALRGLATNVANYNGWNITSPPSYTGQNAVYNEKLYIHAIGPLLANHGW
S_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGT
SDTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF

13231111 (SEQ ID NO: 21)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDNRNVTVDTLFSGTLA
EIRAANQRGGNYAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDRIRELLIQYSDIRTI
LVIEPDSLANMVTNMNVQKCSNAASTYRELTIALKQLDLPHVAMYMDAGHAGWLGWPANIQP
AAELFAKIYEDAGKPRAVRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHYIEAFRPLLEARGF
PAQFIVDQGRSGKQPTGQKEWGHWCNAIGTGFGMRPTANTGHQYVDAFVWVKPGGEC^uDGT
SDTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF

12213111 (SEQ ID NO: 22)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDTLDKTPLMEQTLADI
RTANKNGGNYAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDRIREILISFSDVRTILVI
EPDSLANMVTNMNVPKCSGAASTYKELTVYALKQLNLPHVAMYMDAGHAGWLGWPANIQPA
AELFAQIYRDAGRPAAVRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHYIEAFRPLLEARGF
PAQFIVDQGRSGKQPTGQKEWGHWCNAIGTGFGMRPTANTGHQYVDAFVWVKPGGEC^uDGT
SDTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF

12133333 (SEQ ID NO: 23)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDTLDKTPLMEQTLADI
RTANKNGANPQYAAQIVVYDLPDRDCAAAASNGEWAIANNGVNNYKAYINRIRELLIQYSDIRTI
LVIEPDSLANMVTNMNVQKCSNAASTYKELTVYALKQLNLPHVAMYMDAGHAGWLGWPANIQ
PAAELFAQIYRDAGRPAAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQG
FDAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGT
SDTSAARYDYHCGLSDALTPAPEAGQWFQAYFEQLLINANPPF

33133132 (SEQ ID NO: 24)
GNPFSGVQLWANTYYSEVHTLAIPSLPELAAKAAKVAEVPVSFQWLDNRNVTVDTLFSGTLAEI
RAANQRGANPQYAAQIVVYDLPDRDCAAAASNGEWAIANNGVNNYKAYINRIRELLIQYSDIRTI
LVIEPDSLANMVTNMNVQKCSNAASTYKELTVYALKQLNLPHVAMYMDAGHAGWLGWPANIQ
PAAELFAQIYRDAGRPAAVRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHYIEAFRPLLEAR
GFPKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTS^uSDSAPRFD^uSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

11332333 (SEQ ID NO: 25)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDNRNVTVDTLVQTLS
EIREANQAGANPPYAGIFVVYDLPDRDCAAAASNGEWSIANNGANNYKRYIDRIRELLIQYSDIR
TILVIEPDSLANMVTNMNVQKCSNAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQ
DPAAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQ
GFDKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTS^uSDTSAARYDYHCGLSDALTPAPEAGQWFQAYFEQLLINANPPF

23311333 (SEQ ID NO: 26)
GNPFVGVTPWANAYYASEVSSLAIPSLTGAMATAAAAVAKVPSFQWLDNRNVTVDTLFSGTLAEI
RAANQRGANPPYAGIFVVYDLPDRDCAAAASNGEWSIANNGANNYKRYIDRIREILISFSDVRTI
LVIEPDSLANMVTNMNVPKCSGAASTYRELTIALKQLDLPHVAMYMDAGHAGWLGWPANIQP
AAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQGF
DAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGT
D^uTS^uAARYDYHCGLSDALTPAPEAGQWFQAYFEQLLINANPPF

33213332 (SEQ ID NO: 27)
GNPFSGVQLWANTYYSEVHTLAIPSLPELAAKAAKVAEVPVSFQWLDNRNVTVDTLFSGTLAEI
RAANQRGGNYAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDRIREILISFSDVRTILVI
EPDSLANMVTNMNVPKCSGAASTYKELTVYALKQLNLPHVAMYMDAGHAGWLGWPANIQPA
AELFAQIYRDAGRPAAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQGF^uD

AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGTSD
SSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13333232 (SEQ ID NO: 28)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDNRNVTVDTLFSGTLA
EIRANQQRGANPPYAGIFVYDLPDRDCAAAASNGEWSIANNGANNYKRYIDRIRELLIQYSDIR
TILVIEPDSLANMVTNMNVQKCSNAASTYKELTVYALKQLNLPVAMYMDAGHAGWLGWPANI
QPAAELFAQIYRDAGRPAAVRGLATNVANYNGWNITSPPSYTQGNVYNEKLYIHAIGPLLANH
GWS_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGES
DGTSDSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

22232132 (SEQ ID NO: 29)
GNPFVGVTPWANAYYASEVSSLAIPSLTGAMATAAAAVAKVPSFMWLDTLDKTPLMEQTLADI
RTANKNGGNYAGQFVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDRIRELLIQYSDIRTI
EPDSLANTMVTNMNVQKCSNAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQDPA
AQLFANVYKNASSPRALRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHIEAFRPLLEARGF
PAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGT
DSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

11113132 (SEQ ID NO: 30)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDNRNVTVDTLVQTL
EIREANQAGANPQYAAQIVYDLPDRDCAAAASNGEWAIANNGVNNYKAYINRIREILISFSDVR
TILVIEPDSLANTMVTNMNVKCSGAASTYKELTVYALKQLNLPVAMYMDAGHAGWLGWPANI
QPAAELFAQIYRDAGRPAAVRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHIEAFRPLLEA
RGFPAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTSDSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

21333331 (SEQ ID NO: 31)
GNPFVGVTPWANAYYASEVSSLAIPSLTGAMATAAAAVAKVPSFQWLDNRNVTVDTLVQTLSEI
REANQAGANPPYAGIFVYDLPDRDCAAAASNGEWSIANNGANNYKRYIDRIRELLIQYSDIRTI
LVIEPDSLANTMVTNMNVQKCSNAASTYKELTVYALKQLNLPVAMYMDAGHAGWLGWPANI
PAAELFAQIYRDAGRPAAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHIEAFAPLLRNQ
FPAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGT
SDTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF

21311131 (SEQ ID NO: 32)
GNPFVGVTPWANAYYASEVSSLAIPSLTGAMATAAAAVAKVPSFQWLDNRNVTVDTLVQTLSEI
REANQAGANPPYAGIFVYDLPDRDCAAAASNGEWSIANNGANNYKRYIDRIREILISFSDVRTI
LVIEPDSLANTMVTNMNVKCSGAASTYRELTIALKQLDLPHVAMYMDAGHAGWLGWPANI
AAELFAKIYEDAGKPRVRLATNVANYNAWSVSSPPPYTSPNPNYDEKHIEAFRPLLEARGF
PAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGT
DTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF

12332331 (SEQ ID NO: 33)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDTLDKTPLMEQTLADI
RTANKNGANPPYAGIFVYDLPDRDCAAAASNGEWSIANNGANNYKRYIDRIRELLIQYSDIRTI
LVIEPDSLANTMVTNMNVQKCSNAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQD
PAAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHIEAFAPLLRNQ
GFDPAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTSDTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF

13112332 (SEQ ID NO: 34)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDNRNVTVDTLFSGTLA
EIRANQQRGANPQYAAQIVYDLPDRDCAAAASNGEWAIANNGVNNYKAYINRIREILISFSDVR
TILVIEPDSLANTMVTNMNVKCSGAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQ
DPAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHIEAFAPLLRNQ
GFDPAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTSDSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

22311331 (SEQ ID NO: 35)
GNPFVGVTPWANAYYASEVSSLAIPSLTGAMATAAAAVAKVPSFMWLDTLDKTPLMEQTLADI
RTANKNGANPPYAGIFVYDLPDRDCAAAASNGEWSIANNGANNYKRYIDRIREILISFSDVRTI

57

LVIEPDSLANMVTNMNVPKCSGAASTYRELTIIYALKQLDLPHVAMYMDAGHAGWLGWPANIQP
AAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQGF
DAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGS
DTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF

12111332 (SEQ ID NO:36)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPSPFMWLDLTDKTPLMEQTLA
DIRTANKNGANPQYAAQIVVYDLPDRDCAAAASNGEWAIANNGVNYYKAYINRIREILISFSDVR
TILVIEPDSLANMVTNMNVPKCSGAASTYRELTIIYALKQLDLPHVAMYMDAGHAGWLGWPANI
QPAAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQ
GFD_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTS DSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

12112332 (SEQ ID NO:37)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPSPFMWLDLTDKTPLMEQTLA
DIRTANKNGANPQYAAQIVVYDLPDRDCAAAASNGEWAIANNGVNYYKAYINRIREILISFSDVR
TILVIEPDSLANMVTNMNVPKCSGAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQ
DPAAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQ
GFD_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTS DSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

12131331 (SEQ ID NO:38)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPSPFMWLDLTDKTPLMEQTLADI
RTANKNGANPQYAAQIVVYDLPDRDCAAAASNGEWAIANNGVNYYKAYINRIRELLIQYSDIRTI
LVIEPDSLANMVTNMNVQKCSNAASTYRELTIIYALKQLDLPHVAMYMDAGHAGWLGWPANIQP
AAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQGF
DAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGS
DTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF

12131332 (SEQ ID NO: 39)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPSPFMWLDLTDKTPLMEQTLADI
RTANKNGANPQYAAQIVVYDLPDRDCAAAASNGEWAIANNGVNYYKAYINRIRELLIQYSDIRTI
LVIEPDSLANMVTNMNVQKCSNAASTYRELTIIYALKQLDLPHVAMYMDAGHAGWLGWPANIQP
AAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQGF
DAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGS
DSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

12332332 (SEQ ID NO: 40)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPSPFMWLDLTDKTPLMEQTLADI
RTANKNGANPPYAGIFVVYDLPDRDCAAAASNGEWSIANNGANNYKRYIDRIRELLIQYSDIRTI
LVIEPDSLANMVTNMNVQKCSNAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQD
PAAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQ
GFD A KFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTS DSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

12111131 (SEQ ID NO: 41)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPSPFMWLDLTDKTPLMEQTLADI
RTANKNGANPQYAAQIVVYDLPDRDCAAAASNGEWAIANNGVNYYKAYINRIREILISFSDVRTI
LVIEPDSLANMVTNMNVPKCSGAASTYRELTIIYALKQLDLPHVAMYMDAGHAGWLGWPANIQP
AAELFAKIYEDAGKPRAVRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHYIEAFRPLLEARGF
PAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGS
DTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF

12311332 (SEQ ID NO: 42)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPSPFMWLDLTDKTPLMEQTLADI
RTANKNGANPPYAGIFVVYDLPDRDCAAAASNGEWSIANNGANNYKRYIDRIREILISFSDVRTI
LVIEPDSLANMVTNMNVPKCSGAASTYRELTIIYALKQLDLPHVAMYMDAGHAGWLGWPANIQP
AAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQGF
DAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGS
DSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13332331 (SEQ ID NO: 43)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDRNVTVDTLFSGTLA
EIRAANQRGANPPYAGIFVVDLPDRDCAAAASNGEWSIANNGANNYKRYIDRIRELLIQYSDIR
TILVIEPDSLANTMNMNVQKCSNAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPNQ
DPAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQ
GFDAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTSDTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF

12132331 (SEQ ID NO: 44)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFMWLDTLDKTPLMEQTLADI
RTANKNGANPQYAAQIVVDLPDRDCAAAASNGEWAIANNGVNNYKAYINRIRELLIQYSDIRTI
LVIEPDSLANTMNMNVQKCSNAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPNQD
PAAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQ
GFDAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTSDTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF

12132332 (SEQ ID NO: 45)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFMWLDTLDKTPLMEQTLADI
RTANKNGANPQYAAQIVVDLPDRDCAAAASNGEWAIANNGVNNYKAYINRIRELLIQYSDIRTI
LVIEPDSLANTMNMNVQKCSNAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPNQD
PAAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQ
GFDAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTSDSSAPRFDHSCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13332332 (SEQ ID NO: 46)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDRNVTVDTLFSGTLA
EIRAANQRGANPPYAGIFVVDLPDRDCAAAASNGEWSIANNGANNYKRYIDRIRELLIQYSDIR
TILVIEPDSLANTMNMNVQKCSNAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPNQ
DPAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQ
GFDAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTSDSSAPRFDHSCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

12112132 (SEQ ID NO: 47)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFMWLDTLDKTPLMEQTLADI
RTANKNGANPQYAAQIVVDLPDRDCAAAASNGEWAIANNGVNNYKAYINRIREILISFSDVRTI
LVIEPDSLANTMNMNVKCSGAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPNQD
PAAQLFANVYKNASSPRALRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHYIEAFRPLLEAR
GFPAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTSDSSAPRFDHSCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13322332 (SEQ ID NO: 48)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDRNVTVDTLFSGTLA
EIRAANQRGANPPYAGIFVVDLPDRDCAAAASNGEWSIANNGANNYKRYIDTIRQIVVEYSDIR
TLLVIEPDSLANTLVTNLGTPKCANAQSAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPNQ
DPAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQ
GFDAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTSDSSAPRFDHSCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13131332 (SEQ ID NO: 49)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDRNVTVDTLFSGTLA
EIRAANQRGANPQYAAQIVVDLPDRDCAAAASNGEWAIANNGVNNYKAYINRIRELLIQYSDIR
TILVIEPDSLANTMNMNVQKCSNAASTYRELTIALKQLDLPHVAMYMDAGHAGWLGWPANI
QPAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQ
GFDAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTSDSSAPRFDHSCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

12331332 (SEQ ID NO: 50)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFMWLDTLDKTPLMEQTLADI
RTANKNGANPPYAGIFVVDLPDRDCAAAASNGEWSIANNGANNYKRYIDRIRELLIQYSDIRTI
LVIEPDSLANTMNMNVQKCSNAASTYRELTIALKQLDLPHVAMYMDAGHAGWLGWPANIQP
AAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQGF

7

DAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGS
DSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13312332 (SEQ ID NO:51)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDRNVTVDTLFSGTLA
EIRAANQRGANPPYAGIFVVDLPDRDCAAAASNGEWSIANNGANNYKRYIDRIEILISFSDVR
TILVIEPDSL ANMVTNMNVPKCSGAASAYLECIN YAVTQLNLPNVAMYLDAGHAGWLGWPANQ
DPAAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNP NYDEKHYIEAFAPLLRNQ
GFD AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTS DSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

11113332 (SEQ ID NO:52)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDRNVTVDTLVQTL S
EIREANQAGANPQYAAQIVVDLPDRDCAAAASNGEWAIANNGVN NYKAYINRIEILISFSDVR
TILVIEPDSL ANMVTNMNVPKCSGAASTYKELTVYALKQLNLP HVAMYMDAGHAGWLGWPANI
QPAAELFAQIYRDAGRPA AVRGLATNVANYNAWSIASPPSYTSPNP NYDEKHYIEAFAPLLRNQ
GFD AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTS DSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13113132 (SEQ ID NO:53)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDRNVTVDTLFSGTLA
EIRAANQRGANPQYAAQIVVDLPDRDCAAAASNGEWAIANNGVN NYKAYINRIEILISFSDVR
TILVIEPDSL ANMVTNMNVPKCSGAASTYKELTVYALKQLNLP HVAMYMDAGHAGWLGWPANI
QPAAELFAQIYRDAGRPA AVRGLATNVANYNAWSVSSPPPYTSPNP NYDEKHYIEAFRPLLEA
RGFPAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTS DSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

11112132 (SEQ ID NO:54)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDRNVTVDTLVQTL S
EIREANQAGANPQYAAQIVVDLPDRDCAAAASNGEWAIANNGVN NYKAYINRIEILISFSDVR
TILVIEPDSL ANMVTNMNVPKCSGAASAYLECIN YAVTQLNLPNVAMYLDAGHAGWLGWPANQ
DPAAQLFANVYKNASSPRALRGLATNVANYNAWSVSSPPPYTSPNP NYDEKHYIEAFRPLLEA
RGFP_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGES
DGTSDSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

12113132 (SEQ ID NO:55)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFMWLDTLDKTPLMEQTLADI
RTANKNGANPQYAAQIVVDLPDRDCAAAASNGEWAIANNGVN NYKAYINRIEILISFSDVRTI
LVIEPDSL ANMVTNMNVPKCSGAASTYKELTVYALKQLNLP HVAMYMDAGHAGWLGWPANIQ
PAAELFAQIYRDAGRPA AVRGLATNVANYNAWSVSSPPPYTSPNP NYDEKHYIEAFRPLLEAR
GFPAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTS DSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13132332 (SEQ ID NO:56)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDRNVTVDTLFSGTLA
EIRAANQRGANPQYAAQIVVDLPDRDCAAAASNGEWAIANNGVN NYKAYINRIELLIQYSDIR
TILVIEPDSL ANMVTNMNVPKCSNAASAYLECIN YAVTQLNLPNVAMYLDAGHAGWLGWPANQ
DPAAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNP NYDEKHYIEAFAPLLRNQ
GFD AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTS DSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

11111132 (SEQ ID NO:57)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDRNVTVDTLVQTL S
EIREANQAGANPQYAAQIVVDLPDRDCAAAASNGEWAIANNGVN NYKAYINRIEILISFSDVR
TILVIEPDSL ANMVTNMNVPKCSGAASTYRELT IYALKQLDLPHVAMYMDAGHAGWLGWPANI
QPAAELFAKIYEDAGKPRAVRGLATNVANYNAWSVSSPPPYTSPNP NYDEKHYIEAFRPLLEAR
GFPAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTS DSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13331332 (SEQ ID NO:58)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDRNVTVDTLFSGTLA
EIRAANQRGANPPYAGIFVVDLPDRDCAAAASNGEWSIANNGANNYKRYIDRIELLIQYSDIR

13

TILVIEPDSLANMVTNMNVQKCSNAASTYRELTIIYALKQLDLPHVAMYMDAGHAGWLGWPANI
QPAAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQ
GFDAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTSDSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13111132 (SEQ ID NO:59)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDRNVTVDTLFSGTLA
EIRAANQRGANPQYAAQIVVYDLPDRDCAAAASNGEWAIANNGVNNYKAYINRIREILISFSDVR
TILVIEPDSLANMVTNMNVKCSGAASTYRELTIIYALKQLDLPHVAMYMDAGHAGWLGWPANI
QPAAELFAKIYEDAGKPRAVRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHYIEAFRPLLEAR
GFP_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTSDSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

12222132 (SEQ ID NO: 60)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDRNVTVDTLFSGTLA
RTANKNGGNYAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDTIRQIVVEYSDIRTLLV
IEPDSLNLVTNLGTPKCANAAQSAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQDPAA
QLFANVYKNASSPRALRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHYIEAFRPLLEAR
GFP_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTSDSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

12222332 (SEQ ID NO: 61)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDRNVTVDTLFSGTLA
RTANKNGGNYAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDTIRQIVVEYSDIRTLLV
IEPDSLNLVTNLGTPKCANAAQSAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQDPAA
QLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQGFD
AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTSDSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13311332 (SEQ ID NO: 62)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDRNVTVDTLFSGTLA
EIRAANQRGANPPYAGIFVVYDLPDRDCAAAASNGEWSIANNGANNYKRYIDRIREILISFSDVR
TILVIEPDSLANMVTNMNVKCSGAASTYRELTIIYALKQLDLPHVAMYMDAGHAGWLGWPANI
QPAAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQ
GFDAKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTSDSSAPRFD SHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13311331 (SEQ ID NO: 63)
GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDRNVTVDTLFSGTLA
EIRAANQRGANPPYAGIFVVYDLPDRDCAAAASNGEWSIANNGANNYKRYIDRIREILISFSDVR
TILVIEPDSLANMVTNMNVKCSGAASTYRELTIIYALKQLDLPHVAMYMDAGHAGWLGWPANI
QPAAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQ
GFD_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD
GTSDTTAARYDYHCGLEDALKAPEAGQWFNEYFIQLLRNANPPF

TABLA 8: Dos valores independientes duplicados de T₅₀ (°C) para las CBH II progenitoras, 23 quimeras de CBH II del juego de muestras original y quimeras termoestables predichas de CBH II. Las 18 quimeras sintetizadas para este trabajo van precedidas por un asterisco.

Juego de muestras de Quimeras & Progenitores				Quimeras termoestables predichas			
Secuencia	T ₅₀ (1)	T ₅₀ (2)	Media T ₅₀	Secuencia	T ₅₀ (1)	T ₅₀ (2)	Media T ₅₀
32333113	52	51	51,5	12332331	66,5	67	66,8
13111313	56	53,5	54,8	*13112332	67	67	67
11313121	55	55,5	55,3	22311331	68	68	68
21131311	57,5	57	57,3	*12111332	68	68	68
31212111	59	58	58,5	*12112332	68,5	67,5	68
Progenitor 2	60	58	59	12131331	68,5	69	68,8
23233133	61	61	61	*12131332	70	67,5	68,8
31311112	60	62	61	*12332332	69	69	69

59

22212231	63	61	62	12111131	70	68,5	69,3
13231111	63	6,5	63,3	12311332	70	69	69,5
12213111	63	63,5	63,3	13332331	70	69	69,5
Progenitor 3	63,5	64,5	64	12132331	70,5	69	69,8
12133333	64	64	64	*12132332	70,5	69	69,8
Progenitor 1	64	65,5	64,8	*13332332	69,5	70	69,8
33133132	65	66	65	12112132	71	68,5	69,8
11332333	64,5	66	65,3	13322332	71	68,5	69,8
23311333	65	66	65,5	*13131332	70	70	70
33213332	66	66	66	*12331332	71	69	70
13333232	67,5	67	67,3	*13312332	70	70	70
22232132	68	68	68	*11113332	69,5	70,5	70
11113132	71,5	71	71,3	*13113132	70,5	69,5	70
21333331	73,5	75,5	74,5	*11112132	70,5	70	70,3
21311131	75,5	75,5	75	*12113132	70,5	70,5	70,5
				*13132332	69,5	71,5	70,5
				*11111132	71	70,5	70,8
				13331332	72	70	71
				*13111132	72	69,5	71,3
				*12222132	72,5	70	71,3
				12222332	72	69,5	71,3
				13311332	71	71,5	71,7
				13311331	73,5	72,5	73

[00147] La aplicación de la regresión lineal a los datos de estabilidad de la secuencia dio lugar a un modelo de diez parámetros que adapta los valores observados de T_{50} con $R^2 = 0,88$ (Figura 8). Se hizo un mejor estimativo de la capacidad predictiva del modelo de regresión por fuera del conjunto de entrenamiento, una validación cruzada de once veces que se traduce en un R^2 de 0,57, en donde la remoción de dos valores atípicos, (11313121 y 22222222) aumenta el R^2 de la validación cruzada a 0,76. El modelo de regresión usa el progenitor más estable 1 (*H. insolens*) como T_{50} del estado de referencia e incluye nueve términos adicionales que tienen valores de $p \leq 0,1$. Los parámetros del modelo (Tabla 9) muestran que un único bloque, el bloque 7 del progenitor 3 (B7P3), es de lejos el contribuyente más fuerte a la termoestabilidad de la quimera con respecto a la CBH II de *H. insolens*. Este bloque de la CBH II de *C. termophilum* contribuye aproximadamente 8,5°C a la estabilidad de las quimeras que lo contienen. Se encontró que dos de los 8 bloques restantes con valores de $p \leq 0,1$ hacen contribuciones más pequeñas a la estabilidad, de 1,2°C y 2,7°C, en tanto que los otros seis disminuyen la estabilidad.

TABLE 9: Parámetros y valores de p del modelo de regresión lineal de T_{50} . Los valores del parámetro con $p \leq 0,1$, usados para calcular la línea de adaptación de la regresión de la Figura 1, aparecen en negrita. Los efectos de bloque se informan con respecto a un estado de referencia de un progenitor 1 (CBH II *H. insolens*) en donde 16 parámetros representan la sustitución de cada uno de los 8 bloques de los progenitores 2 y 3.

Bloque	Parámetro Valor	Valor de p
Progenitor1	62,8	0,00
B12	-0,9	0,35
B13	-3,5	0,00

B22	-1,7	0,06
B23	-1,1	0,25
B32	0,5	0,68
B33	1,2	0,10
B42	2,7	0,05
B43	0,0	0,99
B52	-1,3	0,10
B53	-0,6	0,50
B62	-3,5	0,02
B63	-0,7	0,37
B72	-3,8	0,05
B73	8,5	0,00
B82	0,0	1,00
B83	-5,6	0,00

[00148] La alineación de las secuencias B7P1 y B7P3 (Figura 10) muestra que el bloque 7 difiere en 10 de 56 posiciones de aminoácido en las enzimas de *H. insolens* y *C. termophilum*. En los antecedentes de la quimera con el más alto valor de T_{50} , 21311131, cada residuo en B7P3 (segmento 7 del progenitor 3 (SEQ ID NO: 6)) se mutó individualmente al residuo correspondiente en B7P1 (segmento 7 del progenitor 1 (SEQ ID NO: 2)) y se determinaron los valores de T_{50} para cada una de las mutantes puntuales. Una mutación, S313C, alteró marcadamente la termoestabilidad de la quimera: Esta única mutación redujo el T_{50} de 21311131 en aproximadamente 10°C (Figura 11).

[00149] Para estudiar el efecto de la mutación inversa en diferentes antecedentes, se construyeron los genes para las CBH II progenitoras de *H. insolens* y *H. jecorina* que codifican la sustitución C313S (C314S en *H. insolens* y C311S en *H. jecorina*), se expresaron y se determinaron los valores de T_{50} de las enzimas. También se cuantificaron las estabilidades de las quimeras 11111131 y 2222232, en las cuales el B7P3 estabilizador es sustituido dentro de las enzimas de tipo silvestre de *H. insolens* y *H. jecorina*. Tanto la sustitución del bloque B7P3 como la mutación puntual Cys-Ser estabilizaron marcadamente las CBH II progenitoras; el efecto más grande fue un aumento de ~8°C en el T_{50} para la CBH II de *H. jecorina* que contenía la sustitución C311S (Figura 12). la mutación Cys-Ser también se examinó en dos quimeras, 31311112 y 13231111, que no contenían B7P3 así como en una CBH II homóloga (de *Phanerochaete chrysosporium*) la cual no estaba en el juego de recombinación del progenitor. El dominio catalítico de la CBH II de *P. chrysosporium* es solamente 55-56% idéntico a los dominios catalíticos de la CBH II progenitora. Todas estas enzimas fueron estabilizadas por la sustitución Cys-Ser; la CBH II de *P. chrysosporium* se estabilizó notoriamente en 10°C (Figura 13).

[00150] Ocho de las quimeras termoestables de CBH II y las enzimas progenitoras que contienen la mutación equivalente C313S fueron rotuladas con His6 y purificadas de manera que sus actividades específicas pudieran ser determinadas. Como se muestra en la Tabla 10A, las actividades específicas, medidas sobre celulosa amorfa (PASC) a 50°C, para estas quimeras y las enzimas nativas que contienen la mutación Cys-Ser son similares a las de los

progenitores de tipo silvestre. Así pues, el aumento de la termoestabilidad no se obtiene a expensas de la actividad específica.

TABLA 10A: Valores de actividad específica (μg equivalente de azúcar reductor glucosa/ $(\mu\text{g}$ la enzima CBH II $\times \text{min} \times 10^2$)) para CBH II nativas, mutantes puntuales y quiméricas termoestables seleccionadas. Las barras de error muestran los errores estándar, en donde el error estándar se define como desviación estándar/ $\sqrt{n}$, para tres réplicas. Reacción de 2 horas, 3 mg de enzima/g PASC, 50°C, 25 mM acetato de sodio, pH 4,8.

Enzima CBH II	Actividad específica
	μg Azúcar reductor/ $(\mu\text{g}$ Enzima $\times \text{min}) \times 10^2$
<i>Humicola insolens</i> (Progenitor 1)	5,3 +/- 0,5
<i>Hipocrea jecorina</i> (Progenitor 2)	8,4 +/- 0,4
<i>Chaetomium thermophilum</i> (Progenitor 3)	4,8 +/- 0,3
<i>Phanerochaete chrysosporium</i>	7,7 +/- 0,3
<i>Humicola insolens</i> C314S	5,3 +/- 0,9
<i>Hipocrea jecorina</i> C311S	7,8 +/- 0,5
<i>Phanerochaete chrysosporium</i> C311S	8,5 +/- 0,1
HJPlus (Quimera 12222332)	9,6 +/- 0,8
Quimera 13111132	8,5 +/- 0,3
Quimera 2222232	7,7 +/- 0,3
Quimera 13311332	6,8 +/- 0,6
Quimera 13311331	6,2 +/- 0,3
Quimera 11111131	6,1 +/- 0,9
Quimera 13112332	5,6 +/- 0,4
Quimera 21311131	5,5 +/- 0,3
Quimera 11113132	5,3 +/- 0,5
Quimera 21333331	3,8 +/- 0,4

Tabla 10B: Actividad total en sobrenadantes de medio de cultivo de expresión tanto sintético (SDCAA) como enriquecido (YPD) para CBH II de *H. jecorina* y *H. insolens* de tipo silvestre, mutante puntual C313S y sustitución de bloque B7P3. Los valores presentados son μg de glucosa/mL de actividad de celulasa por mL de sobrenadante de cultivo de expresión de equivalente de CBH II añadido a la prueba de actividad de celulasa. Para los cultivos SDCAA, se usaron sobrenadantes concentrados de cultivo SDCAA y se midió la actividad hacia la celulosa hinchada en ácido fosfórico (1 mg/mL) a 50°C durante 100 minutos en acetato de sodio 50 mM, pH 4,8. El sobrenadante YPD de CBH II se concentró mediante enlace a Avicel y se midió la actividad hacia el Avicel (15 mg/mL) a 55°C durante 150 minutos en acetato de sodio 50 mM, pH 4,8.

CBH II	SDCAA (1)	SDCAA (2)	SDCAA Mean	YPD (1)	YPD (2)	YPD Mean
<i>H. jecorina</i>	19	17	18	0.4	0.4	0.4
<i>H. jeco</i> C311S	50	43	47	6.1	5.6	5.9
<i>H. jeco</i> B7P3	35	33	34	3.9	3.6	3.8
<i>H. insolens</i>	73	83	78	6.2	6.0	6.1
<i>H. inso</i> C314S	100	97	98	8.8	8.0	8.4
<i>H. inso</i> B7P3	39	42	40	4.4	4.1	4.2

[00151] Estas mismas ocho quimeras termoestables (T_{50} 2-10°C más alta que el progenitor más estable) se examinaron entonces para determinar la actividad sobre celulosa cristalina durante una incubación de 16 horas a lo largo de una gama de temperaturas, incluidas aquellas a las cuales las enzimas progenitoras evidencian poca o ninguna actividad. La Figura 9a muestra que 7 de 8 quimeras termoestables examinadas mostraron actividad máxima hacia

2
6

Avicel a 60-65°C, y todas las 8 quimeras retuvieron su actividad a 70°C, la más alta temperatura examinada. en contraposición, las tres CBH II progenitoras muestran máxima actividad a 50°C y son o bien completa o casi completamente inactivas a 70°C. Además, las siete quimeras con temperaturas aumentadas de actividad óptima hidrolizan significativamente más Avicel que cualquiera de las enzimas CBH II progenitoras. Como se muestra en la Figura 9b, se observan comportamientos similares para las progenitoras de *H. insolens* y *H. jecorina* que contienen la mutación puntual Cys-Ser. La mutación puntual Cys-Ser también aumentó la hidrólisis de Avicel y la máxima temperatura operativa para la CBH II de *P. chrysosporium*. La sustitución del bloque P3B7, que se hizo en las progenitoras de *H. insolens* y *H. jecorina*, aumentó tanto la temperatura operativa como la hidrólisis de la CBH II del *H. insolens* pero, a pesar de aumentar la máxima temperatura operativa, no mejoró la hidrólisis global de la celulosa por la enzima de *H. jecorina*.

[00152] Se observó una baja (< 1 mg/L) secreción de CBH II de tipo silvestre de *H. jecorina* del huésped heterólogo de expresión *S. cerevisiae*. La mutación C311S en la enzima CBH II de tipo silvestre de *H. jecorina* aumenta marcadamente la actividad total de la CBH II segregada (Tabla 11). En medio sintético (SDCAA), las sustituciones C311S y B7P3 aumentan la actividad total segregada de la CBH II de *H. jecorina* por un factor de dos, en tanto que en medio enriquecido (YPD) el aumento de la actividad es de diez veces. En cuanto a la CBH II progenitora de *H. insolens*, la cual se expresa en niveles mucho más altos que las otras dos CBH II progenitoras, la mutación C314S aumentó la actividad segregada en un factor de ~1,5, en tanto que la sustitución del bloque B7P3 la redujo. Dado que los *H. insolens* y *H. jecorina* de tipo silvestre y las mutantes Cys-Ser tienen actividades específicas similares (Tabla 10), el aumento de la actividad total segregada de celulasa es el resultado de una mejor secreción de la enzima funcional. Se ha observado una correlación entre la secreción de la proteína heteróloga de *S. cerevisiae* y la estabilidad de la proteína, lo cual sugiere que el aumento de la secreción de las CBH II mutantes Cys-Ser podría reflejar sus estabilidades más altas.

TABLA 11: Valores de actividad específica (µg de equivalente de azúcar reductor glucosa/(µg de enzima CBH II x min x 10²)) para CBH II nativas, mutantes puntuales y quiméricas termoestables seleccionadas. Las barras de error muestran los errores estándar, en donde el error estándar se define como desviación estándar/sqrt (n), para tres réplicas. Reacción de 2 horas, 3 mg de enzima/g PASC, 50°C, 25 mM acetato de sodio, pH 4,8.

Enzima CBH II	Actividad específica µg Azúcar reductor/(µg Enzima x min) x 10 ²
<i>Humicola insolens</i> (Progenitor 1)	5,3 +/- 0,5
<i>Hipocrea jecorina</i> (Progenitor 2)	8,4 +/- 0,4
<i>Chaetomium termophilum</i> (Progenitor 3)	4,8 +/- 0,3
<i>Phanerochaete chrysosporium</i>	7,7 +/- 0,3
<i>Humicola insolens</i> C314S	5,3 +/- 0,9
<i>Hipocrea jecorina</i> C311S	7,8 +/- 0,5
<i>Phanerochaete chrysosporium</i> C311S	8,5 +/- 0,1
HJPlus (Quimera 12222332)	9,6 +/- 0,8
Quimera 13111132	8,5 +/- 0,3
Quimera 22222232	7,7 +/- 0,3

63

Quimera 13311332	6,8 +/- 0,6
Quimera 13311331	6,2 +/- 0,3
Quimera 11111131	6,1 +/- 0,9
Quimera 13112332	5,6 +/- 0,4
Quimera 21311131	5,5 +/- 0,3
Quimera 11113132	5,3 +/- 0,5
Quimera 21333331	3,8 +/- 0,4

[00153] Para modelar la mutación Cys-Ser se usó la estructura cristalina de alta resolución 1ocn de la CBH II de *H. insolens*. Primero, la red de enlaces de hidrógeno se optimizó con REDUCE. Se predijo que Cys314 formaría un enlace de hidrógeno con el carbonilo de Pro 339. Para confirmar esta predicción, se optimizó el empaque de la cadena lateral usando la plataforma de modelado SHARPEN. Se predice que Ser314 hace las interacciones similares a Cys314, lo cual se traduce en enlaces de hidrógeno más fuertes y una geometría más favorable (Figura 14).

[00154] Varios efectos podrían explicar por qué la mutación Cys-Ser estabiliza un amplio intervalo de CBH II, incluidas las CBH II nativas y las quimeras. Cys y Ser son similares (aunque no isoestéricos) y estos dos aminoácidos dominan la alineación de las secuencias en esta posición en comparación con otras alternativas. Los socios del enlace de hidrógeno para este residuo son elementos de la columna vertebral (la amida de Gly316 y el carbonilo de Pro339) y por consiguiente es menos probable que sean dependientes de variaciones de terceros aminoácidos. Además, las cadenas laterales vecinas inmediatas de este bolsillo (Asn283, Pro339, Phe345) se conservan entre las cuatro células CBH II nativas estudiadas.

[00155] La estructura cristalina de alta resolución (1,3 Å) de *H. insolens* (pdb entrada 1ocn6) muestra que Cys314 forma parte de una red de enlaces de hidrógeno (Figura 15). El aumento de la capacidad de enlace de hidrógeno de la Ser con respecto a Cys podría sugerir un papel para interacciones más fuertes de los enlaces de hidrógeno en la estabilización. La estructura cristalina también sugiere que Ser puede ser preferida por razones estéricas. Específicamente, cuando la cadena lateral de la Cys se reconstruye con ángulos canónicos del enlace, se remueve una angulación de 6° y la Cys es empujada más cerca del carbonilo de Pro339, creando una interacción estérica desfavorable.

[00156] Una alineación de las 196 secuencias de proteínas comparte la máxima identidad con la CBH II de *H. jecorina*. Cincuenta y cuatro de las 250 secuencias más idénticas se excluyeron de la alineación debido a redundancia (es decir, mutantes puntuales para estudios estructurales o isoformas con identidad >95%). Existe un sesgo en favor de Ser311: 158 secuencias tienen Ser, 20 tienen Ala, 10 tienen Cys, 5 tienen una delección y 3 tienen Gly. Sin embargo, existen otras 42 posiciones en donde la elección más frecuente tiene lugar con más de dos veces la frecuencia del aminoácido de *H. jecorina*.

[00157] El gran efecto estabilizador de la mutación Cys-Ser suscita la posibilidad que la Ser en esta posición es un indicador global de termoestabilidad de la celulasa nativa. Sin embargo, la T₅₀ de 64,8°C para la CBH II de *H. insolens*, que tiene Cys en esta posición, es mayor que la

64

de la CBH II de *C. thermophilum* (64,0°C), lo cual indica que la Ser no es el único determinante de estabilidad.

[00158] La termoestabilidad no es la única propiedad de interés de las celulasas industriales. También son importantes la actividad específica, los cambios en la unión a la celulosa y los efectos sobre la expresión y la inhibición del producto. Las quimeras y los datos que se muestran aquí demuestran que la recombinación da como resultado quimeras de CBH II cuya mejor termoestabilidad no sucede a expensas de la actividad específica medida en la prueba de hidrólisis de la celulosa a corto plazo (*p.ej.*, 2 horas). Se hicieron observaciones parecidas con las CBH II que contienen la mutación termoestabilizadora Cys-Ser. En pruebas de hidrólisis a largo plazo, varias de las quimeras de CBH II y las tres CBH II con la mutación Cys-Ser hidrolizaron más celulosa que las CBH II nativas. Este desempeño superior quizá sea el resultado de tener una actividad específica comparable a la de las CBH II progenitoras junto con mayor termoestabilidad que permite a la enzima continuar funcionando más tiempo a temperaturas elevadas. Dado que estos exámenes se llevaron a cabo con cantidades iguales de enzimas progenitoras, quimeras y mutantes Cys-Ser purificadas, las mejoras observadas en la hidrólisis a alta temperatura no son el resultado de un aumento de la secreción del huésped de expresión *S. cerevisiae*. Las quimeras termoestables y las mutantes Cys-Ser por consiguiente pueden resultar siendo útiles componentes de las formulaciones de la enzima para la degradación la celulosa.

[00159] Aunque se han descrito e ilustrado diversas modalidades específicas, se apreciará que se pueden introducir varios cambios sin apartarse del espíritu y el alcance de la invención.

5

Lo que se reivindica es:

1. Un polipéptido recombinante que comprende una Serina en la región C terminal en un motivo que comprende la secuencia GEXDG, en donde X es C, A o G (SEQ ID NO: 107), en donde X está sustituido por Serina, en donde la variante comprende actividad de celulasa y en donde el polipéptido tiene termoestabilidad aumentada en comparación con una celulasa de tipo silvestre que carece de una Serina en la secuencia GEXDG (SEQ ID NO: 107).
2. El polipéptido recombinante de la reivindicación 1, el cual polipéptido comprende una secuencia que es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 2 que comprende una C314S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 4 que comprende una C311S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 12 que comprende una C310S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 13 que comprende una C312S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 14 que comprende una C314S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 15 que comprende una C315S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 16 que comprende una C313S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 17 que comprende una C311S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 19 que comprende una C313S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 21 que comprende una C312S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 22 que comprende una C311S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 64 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 65 que comprende una C407S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 66 que comprende una C394S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 67 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 68 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 69 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 70 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 71 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 72 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 73 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 74 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 75 que comprende una C400S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 76 que comprende una C407S; es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 77 que comprende una C394S; o es por lo menos 85%, 90%, 95%, 98%, 99% idéntica a la SEQ ID NO: 78 que comprende una C412S, en donde los anteriores polipéptidos tienen actividad de celulasa y mejor termoestabilidad en comparación con su proteína progenitora correspondiente (de tipo silvestre) que carece de la mutación Cys→Ser.

66

3. Un polipéptido recombinante que comprende una secuencia como la que se estipula más adelante que tienen de 1-10, 10-20 o 20-30 sustituciones conservadoras de aminoácido excepto en la posición que se identifica a continuación en donde está presente una sustitución C→S: SEQ ID NO: 2 que comprende una C314S; SEQ ID NO: 4 que comprende una C311S; SEQ ID NO: 12 que comprende una C310S; SEQ ID NO: 13 que comprende una C312S; SEQ ID NO: 14 que comprende una C314S; SEQ ID NO: 15 que comprende una C315S; SEQ ID NO: 16 que comprende una C313S; SEQ ID NO: 17 que comprende una C311S; SEQ ID NO: 19 que comprende una C313S; SEQ ID NO: 21 que comprende una C312S; SEQ ID NO: 22 que comprende una C311S; SEQ ID NO: 64 que comprende una C400S; SEQ ID NO: 65 que comprende una C407S; SEQ ID NO: 66 que comprende una C394S; SEQ ID NO: 67 que comprende una C400S; SEQ ID NO: 68 que comprende una C400S; SEQ ID NO: 69 que comprende una C400S; SEQ ID NO: 70 que comprende una C400S; SEQ ID NO: 71 que comprende una C400S; SEQ ID NO: 72 que comprende una C400S; SEQ ID NO: 73 que comprende una C400S; SEQ ID NO: 74 que comprende una C400S; SEQ ID NO: 75 que comprende una C400S; SEQ ID NO: 76 que comprende una C407S; SEQ ID NO: 77 que comprende una C394S; o la SEQ ID NO: 78 que comprende una C412S.

4. El polipéptido recombinante de la reivindicación 1, que comprende una secuencia seleccionada de entre el grupo compuesto por:

(a) un polipéptido que tiene una identidad de 85% o más con la SEQ ID NO: 2, que tiene una Ser en la posición 314 y en donde el polipéptido tiene actividad de celulasa;

(b) un polipéptido que tiene una identidad de por lo menos 70% o más con la SEQ ID NO: 4, que tiene una Ser en la posición 311 y en donde el polipéptido tiene actividad de celulasa;

(c) un polipéptido que tiene una identidad del 70% o más con una secuencia seleccionada de entre el grupo compuesto por: (i) SEQ ID NO: 12 y que tiene una Ser en la posición 310, (ii) SEQ ID NO: 13 y que tiene una Ser en la posición 312, (iii) SEQ ID NO: 14 y que tiene una Ser en la posición 314, (iv) SEQ ID NO: 15 y que tiene una Ser en la posición 315, (v) SEQ ID NO: 16 y que tiene una Ser en la posición 313, (vi) SEQ ID NO: 17 y que tiene una Ser en la posición 311, (vii) SEQ ID NO: 19 y que tiene una Ser en la posición 313, (viii) SEQ ID NO: 21 y que tiene una Ser en la posición 312 y (ix) SEQ ID NO: 22 y que tiene una Ser en la posición 311 y en donde cada uno de los anteriores polipéptidos tienen actividad de celulasa;

(d) un polipéptido quimérico que comprende por lo menos dos dominios de dos polipéptidos progenitores de celobiohidrolasa diferentes, en donde los dominios comprenden de N- a C- terminal: (segmento 1)-(segmento 2)-(segmento 3)-(segmento 4)-(segmento 5)-(segmento 6)-(segmento 7)-(segmento 8);

en

donde:

el segmento 1 comprende una secuencia que es por lo menos 50-100% idéntica al residuo de aminoácidos de cerca de 1 a cerca de x_1 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 2 comprende una secuencia que es por lo menos 50-100% idéntica al residuo de aminoácidos x_1 a cerca de x_2 de la SEQ ID NO: 2 ("1"), la SEQ

1
2

ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 3 comprende una secuencia que es por lo menos 50-100% idéntica al residuo de aminoácidos x_2 a cerca de x_3 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 4 comprende una secuencia que es por lo menos 50-100% idéntica al residuo de aminoácidos x_3 a cerca de x_4 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 5 comprende una secuencia que es por lo menos 50-100% idéntica a cerca del residuos de aminoácidos x_4 a cerca de x_5 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 6 comprende una secuencia que es por lo menos 50-100% idéntica al residuo de aminoácidos x_5 a cerca de x_6 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); el segmento 7 comprende una secuencia que es por lo menos 50-100% idéntica al residuo de aminoácidos x_6 a cerca de x_7 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3"); y segmento 8 comprende una secuencia que es por lo menos 50-100% idéntica al residuo de aminoácidos x_7 a cerca de x_8 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3");

en donde x_1 es el residuo 43, 44, 45, 46 o 47 de la SEQ ID NO: 2 o el residuo 42, 43, 44, 45 o 46 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_2 es el residuo 70, 71, 72, 73 o 74 de la SEQ ID NO: 2 o el residuo 68, 69, 70, 71, 72, 73 o 74 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_3 es el residuo 113, 114, 115, 116, 117 o 118 de la SEQ ID NO: 2 o el residuo 110, 111, 112, 113, 114, 115 o 116 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_4 es el residuo 153, 154, 155, 156 o 157 de la SEQ ID NO: 2 o el residuo 149, 150, 151, 152, 153, 154, 155 o 156 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_5 es el residuo 220, 221, 222, 223 o 224 de la SEQ ID NO: 2 o el residuo 216, 217, 218, 219, 220, 221, 222 o 223 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_6 es el residuo 256, 257, 258, 259, 260 o 261 de la SEQ ID NO: 2 o el residuo 253, 254, 255, 256, 257, 258, 259 o 260 de la SEQ ID NO: 4 o la SEQ ID NO: 6; x_7 es el residuo 312, 313, 314, 315 o 316 de la SEQ ID NO: 2 o el residuo 309, 310, 311, 312, 313, 314, 315 o 318 de la SEQ ID NO: 4 o la SEQ ID NO: 6; y x_8 es un residuo de aminoácidos correspondiente al C terminal del polipéptido que tiene la secuencia de la SEQ ID NO: 2, SEQ ID NO: 4 o la SEQ ID NO: 6,

en donde el polipéptido quimérico comprende una Ser en la posición 314 de la SEQ ID NO: 2 o posición 311 de la SEQ ID NO: 4 y en donde el polipéptido quimérico tiene actividad de celulasa y mejor termoestabilidad y/o estabilidad de pH en comparación con un polipéptido CBH II que comprende la SEQ ID NO: 2, 4 o 6.

5. El polipéptido recombinante de la reivindicación 4, en donde el segmento 1 comprende residuos de aminoácidos de cerca de 1 a cerca de x_1 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y que tienen 1-10 sustituciones conservadoras de aminoácido; el segmento 2 va de cerca del residuo de aminoácidos x_1 a cerca del x_2 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y tiene cerca de 1-10 sustituciones conservadoras de aminoácido; el segmento 3 va de cerca del residuo de aminoácidos x_2 a cerca del x_3 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y tiene cerca de 1-10 sustituciones conservadoras de aminoácido; el segmento 4 va de cerca del residuo de aminoácidos x_3 a cerca del x_4 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la

6

SEQ ID NO: 6 ("3") y tiene cerca de 1-10 sustituciones conservadoras de aminoácido; el segmento 5 va de cerca del residuo de aminoácidos x_4 a cerca del x_5 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y tiene cerca de 1-10 sustituciones conservadoras de aminoácido; el segmento 6 va de cerca del residuo de aminoácidos x_5 a cerca del x_6 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y tiene cerca de 1-10 sustituciones conservadoras de aminoácido; el segmento 7 va de cerca del residuo de aminoácidos x_6 a cerca del x_7 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y tiene cerca de 1-10 sustituciones conservadoras de aminoácido; y segmento 8 va de cerca del residuo de aminoácidos x_7 a cerca del x_8 de la SEQ ID NO: 2 ("1"), la SEQ ID NO: 4 ("2") o la SEQ ID NO: 6 ("3") y tiene cerca de 1-10 sustituciones conservadoras de aminoácido excepto en la posición 314 de la SEQ ID NO: 2, posición 311 o la SEQ ID NO: 4 o 313 de la SEQ ID NO: 6.

6. El polipéptido recombinante de la reivindicación 4, en donde el polipéptido quimérico comprende una secuencia que es por lo menos 80%, 90%, 95%, 98% o 99% idéntica a una secuencia seleccionada de entre el grupo compuesto por la SEQ ID NO: 12-62 y 63.
7. Un polipéptido recombinante de la reivindicación 1, el cual polipéptido comprende una secuencia que tiene por lo menos 85% de identidad con las SEQ ID NO: 79-105 o 106 y tiene una Serina en X en la secuencia de la SEQ ID NO: 107.
8. Un polinucleótido que codifica un polipéptido de las reivindicaciones 1, 2, 3, 4 o 7.
9. Un vector que comprende un polinucleótido de la reivindicación 8.
10. Una célula hospedera que comprende el polinucleótido de la reivindicación 8.
11. Una célula hospedera que comprende el vector de la reivindicación 9.
12. Una preparación enzimática que comprende un polipéptido de la reivindicación 1, 2, 3, 4 o 7.
13. Una preparación enzimática que comprende un polipéptido producido por una célula hospedera de la reivindicación 10.
14. Un método para el tratamiento de biomasa que comprende celulosa, el cual método comprende poner en contacto la biomasa con un polipéptido de la reivindicación 1, 2, 3, 4 o 7.
15. Un método para el tratamiento de biomasa que comprende celulosa, el cual método comprende poner en contacto la biomasa con una preparación enzimática de la reivindicación 12.

69

RESUMEN

La presente revelación se relaciona con polipéptidos quiméricos de fusión CBH II, ácidos nucleicos que codifican los polipéptidos y células hospederas para producir los polipéptidos.

70

1/11

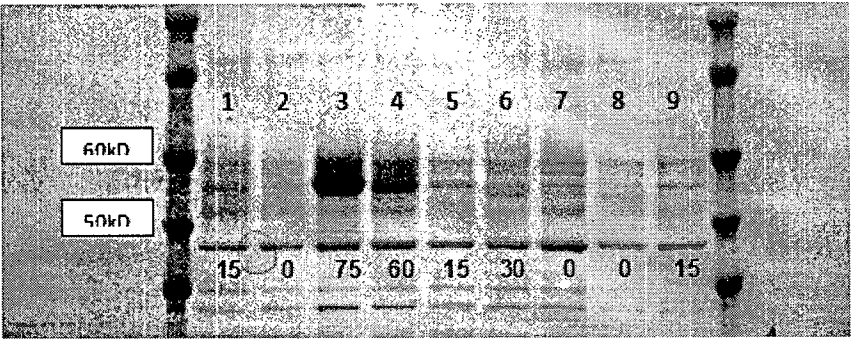


FIGURA 1A

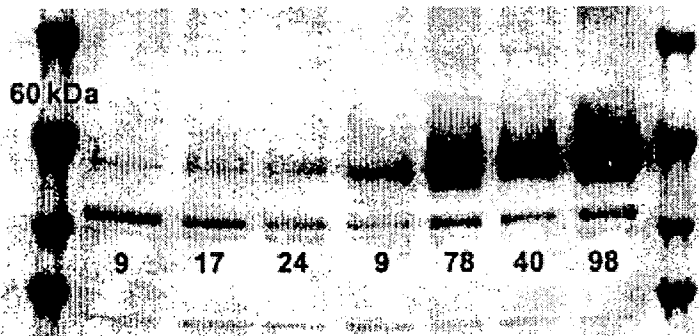


FIGURA 1B

2/11

21

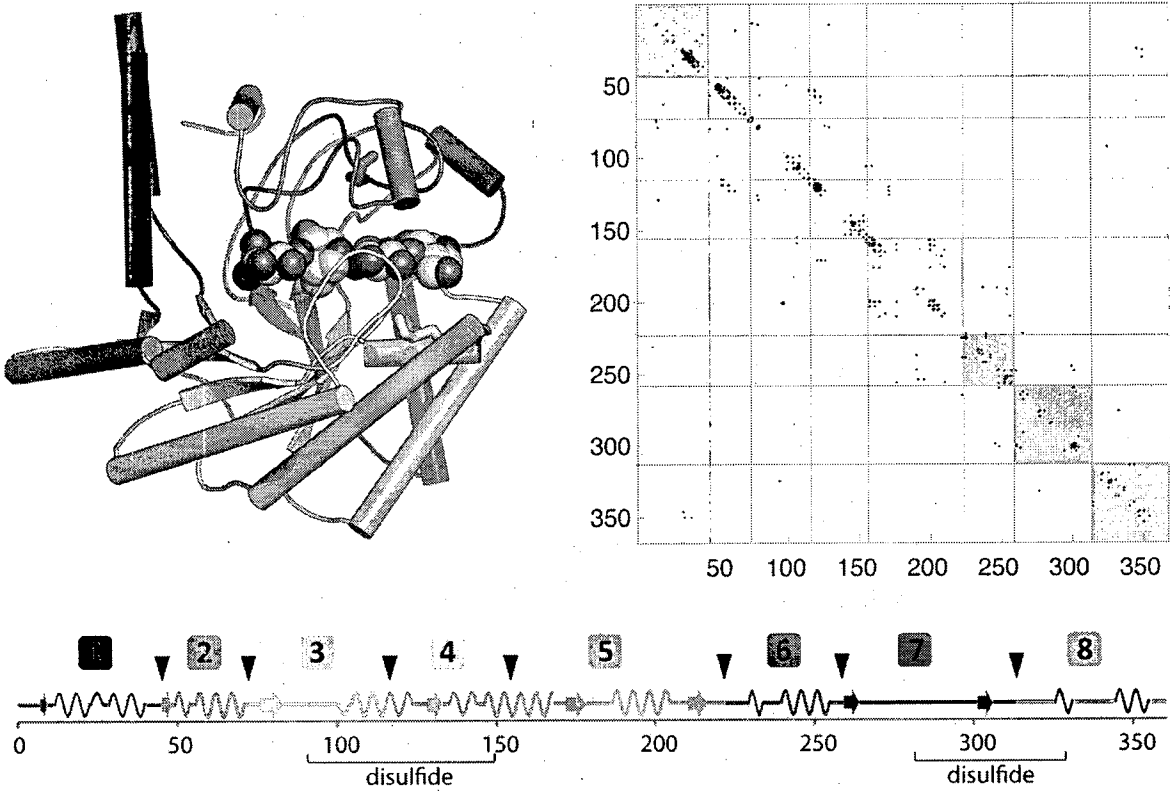


FIGURA 2

2
2

3/11

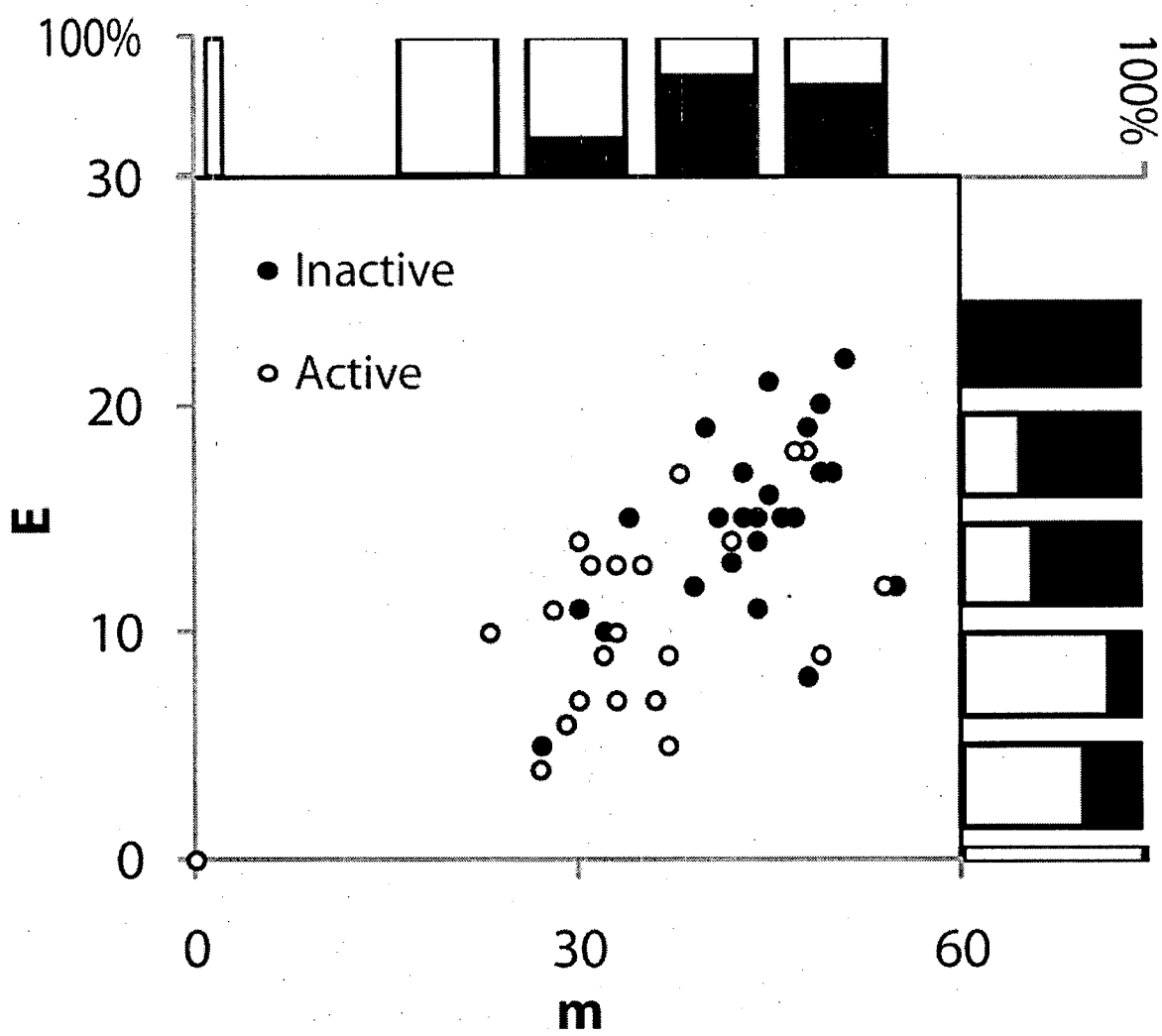


FIGURA 3

4/11

23

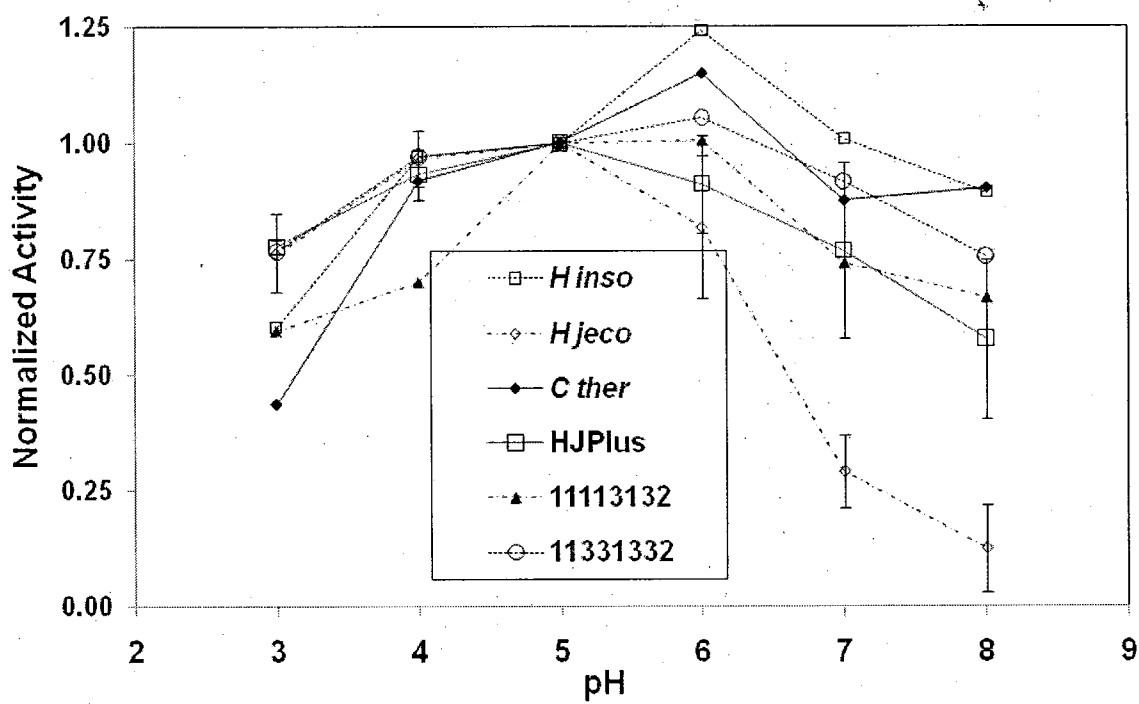


FIGURA 4

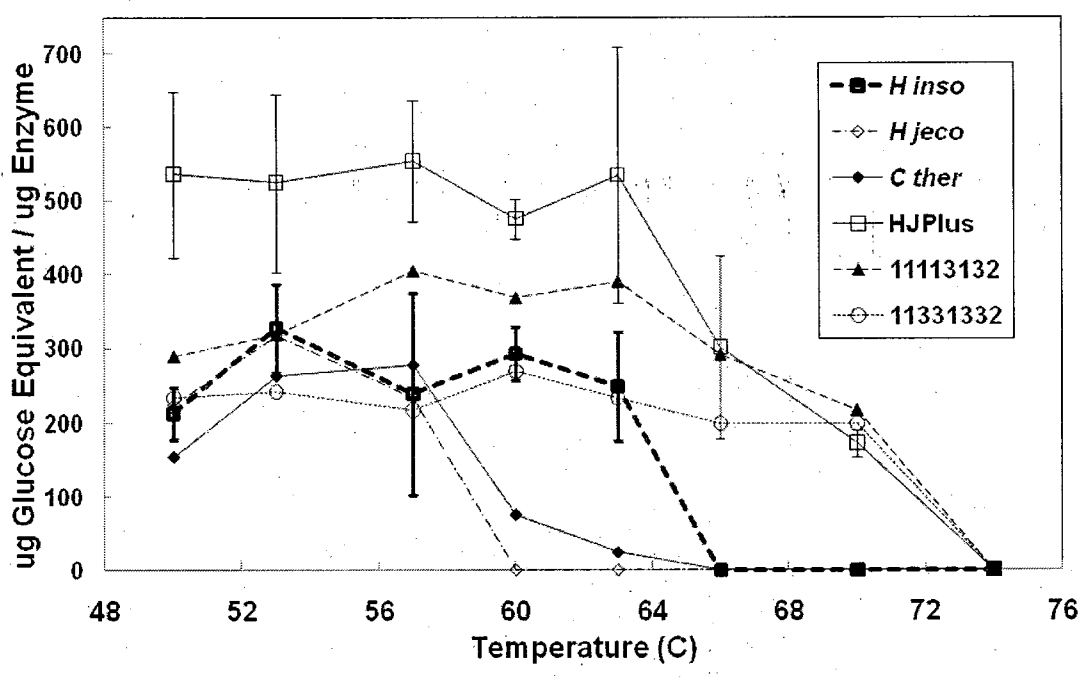


FIGURA 5

74

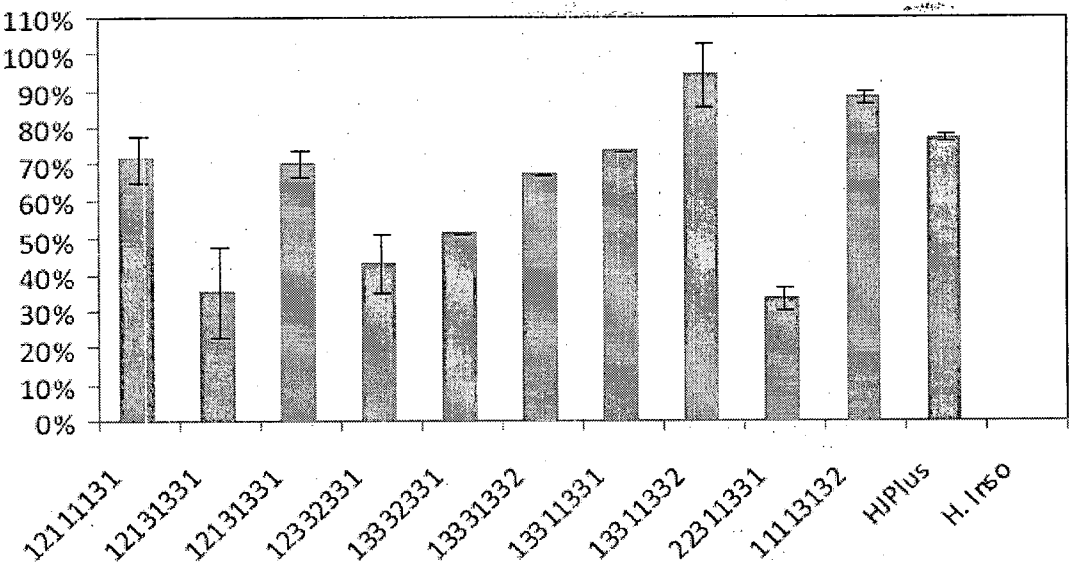


FIGURA 6

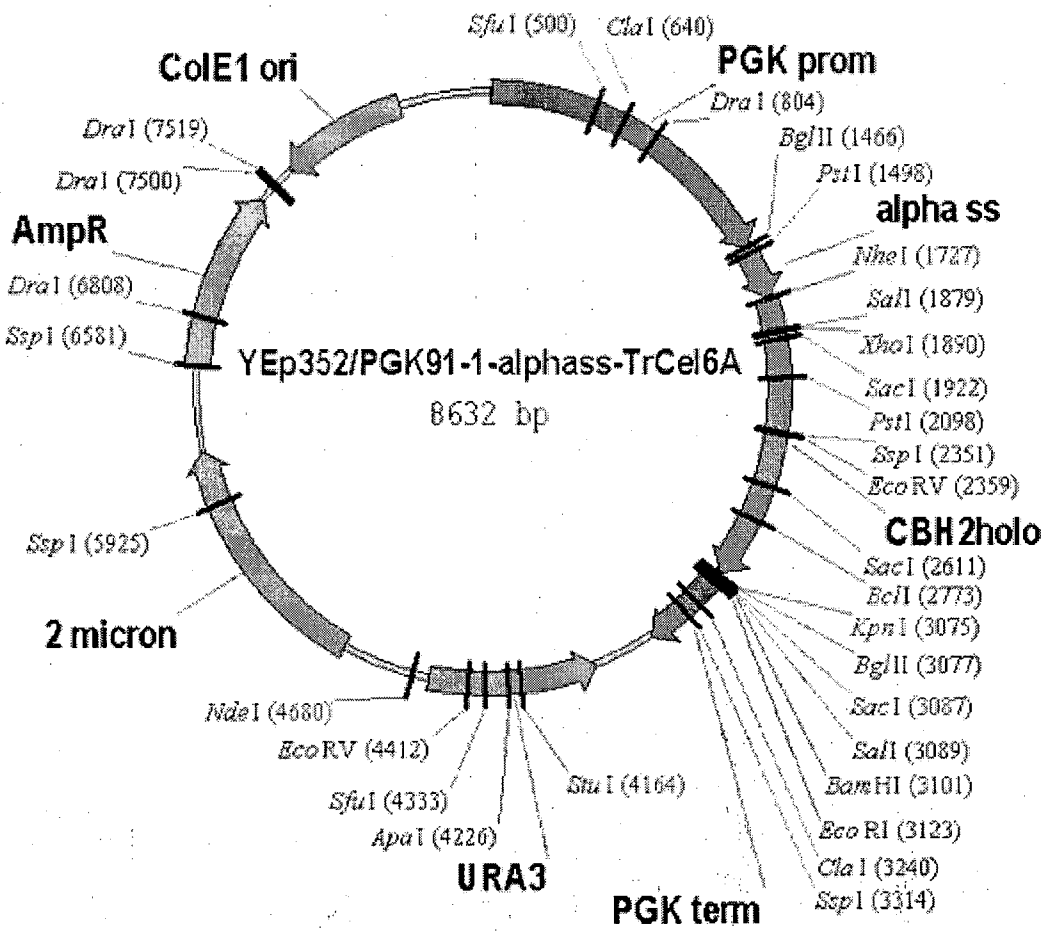


FIGURA 7

25

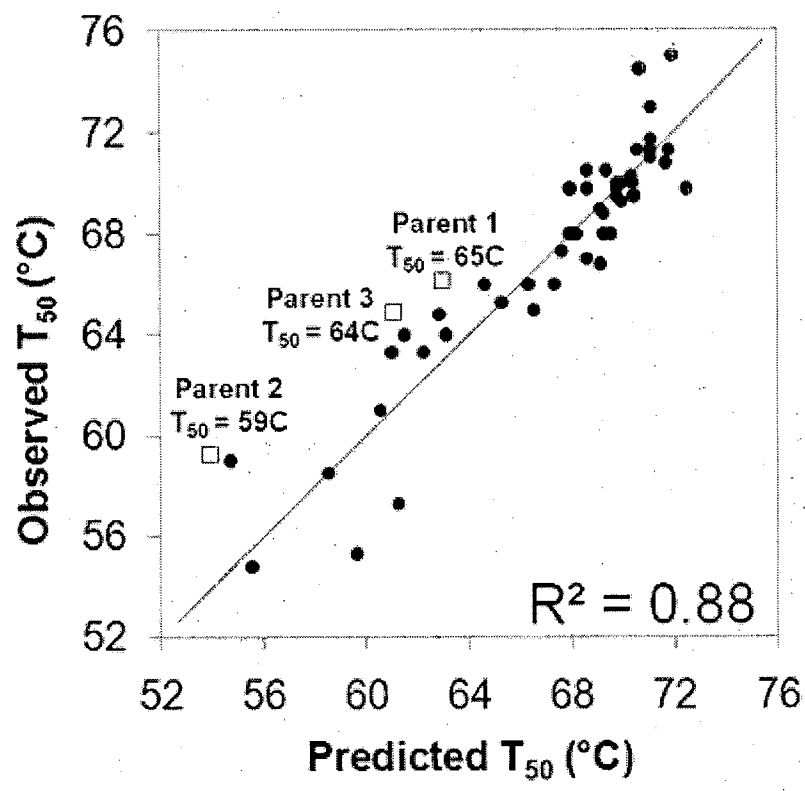


FIGURA 8

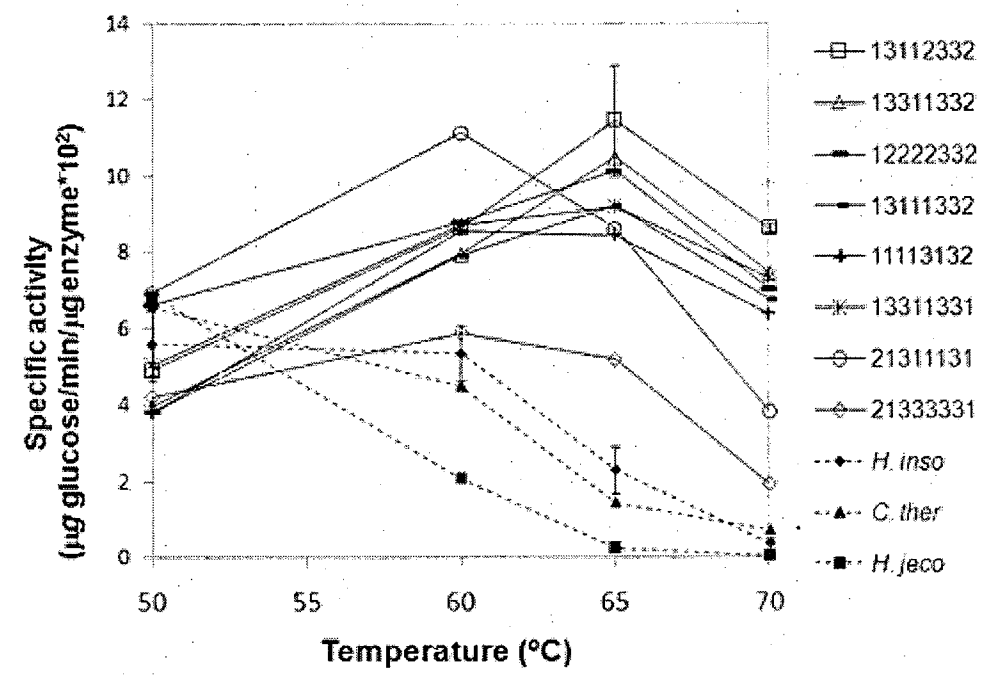


FIGURA 9A

26

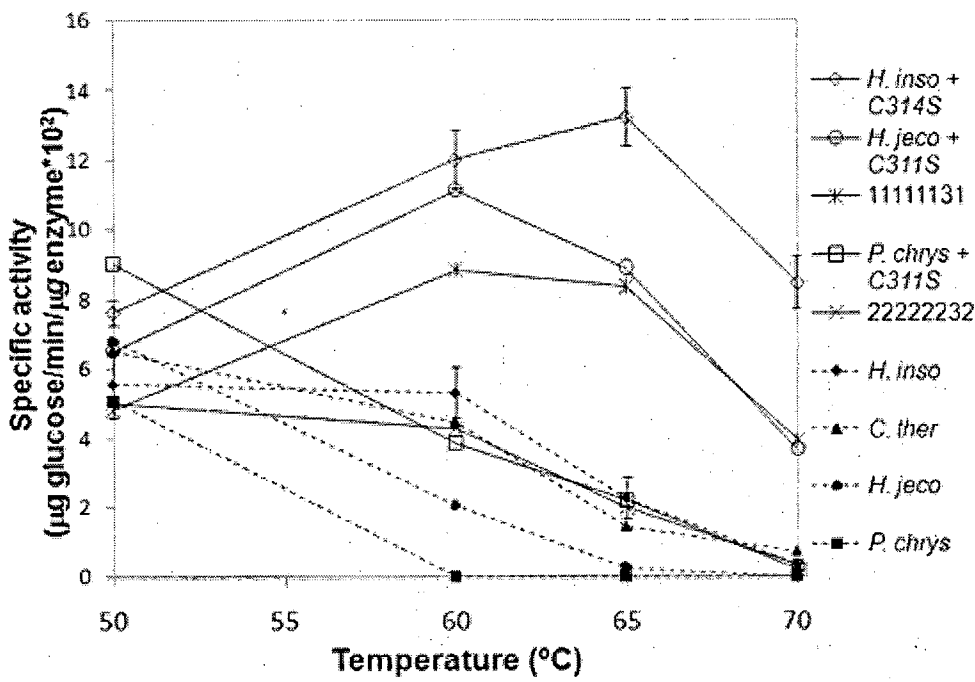


FIGURA 9B

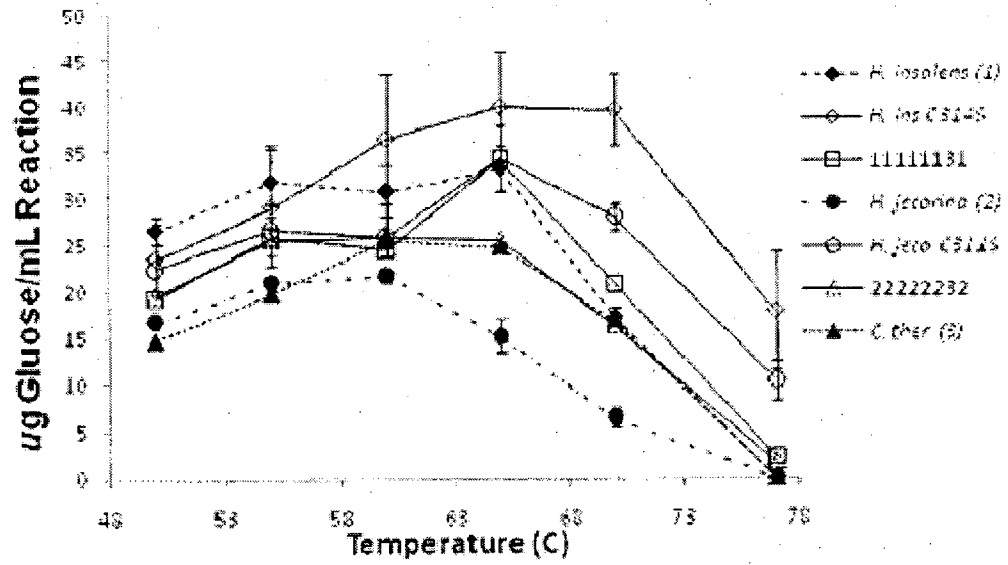


FIGURA 9C



FIGURA 10

7/7

8/11

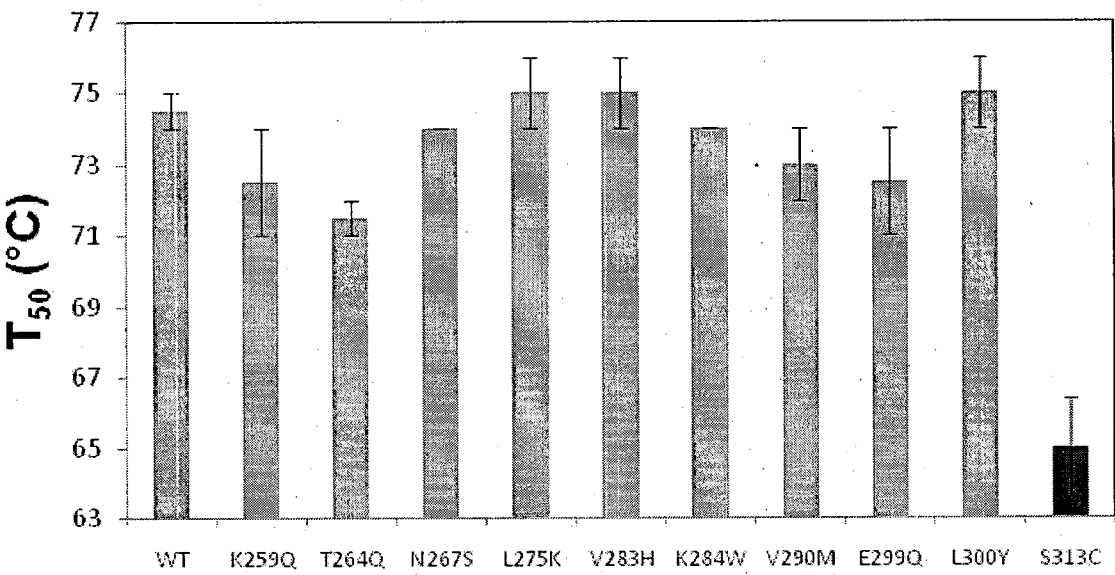


FIGURA 11

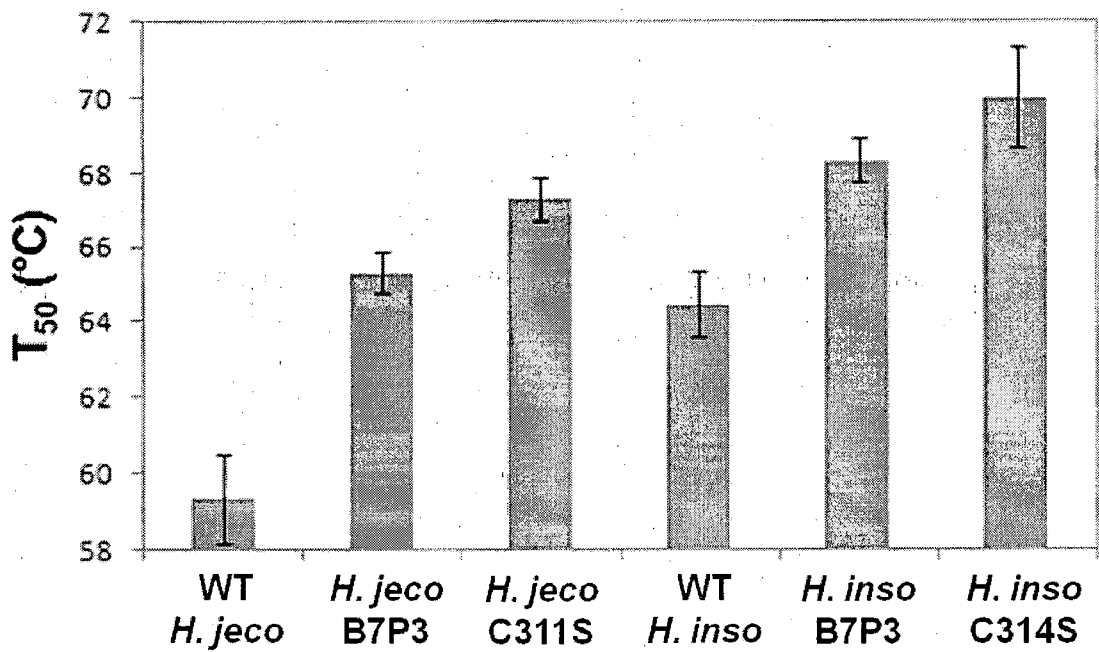


FIGURA 12

28

9/11

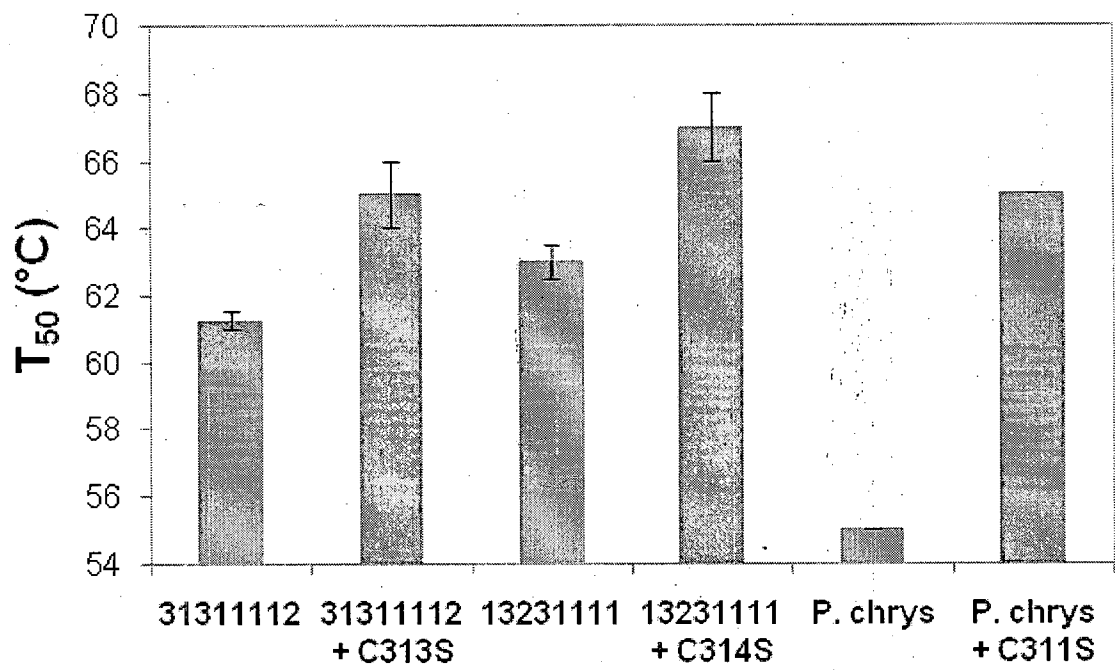


FIGURA 13

10/11

79

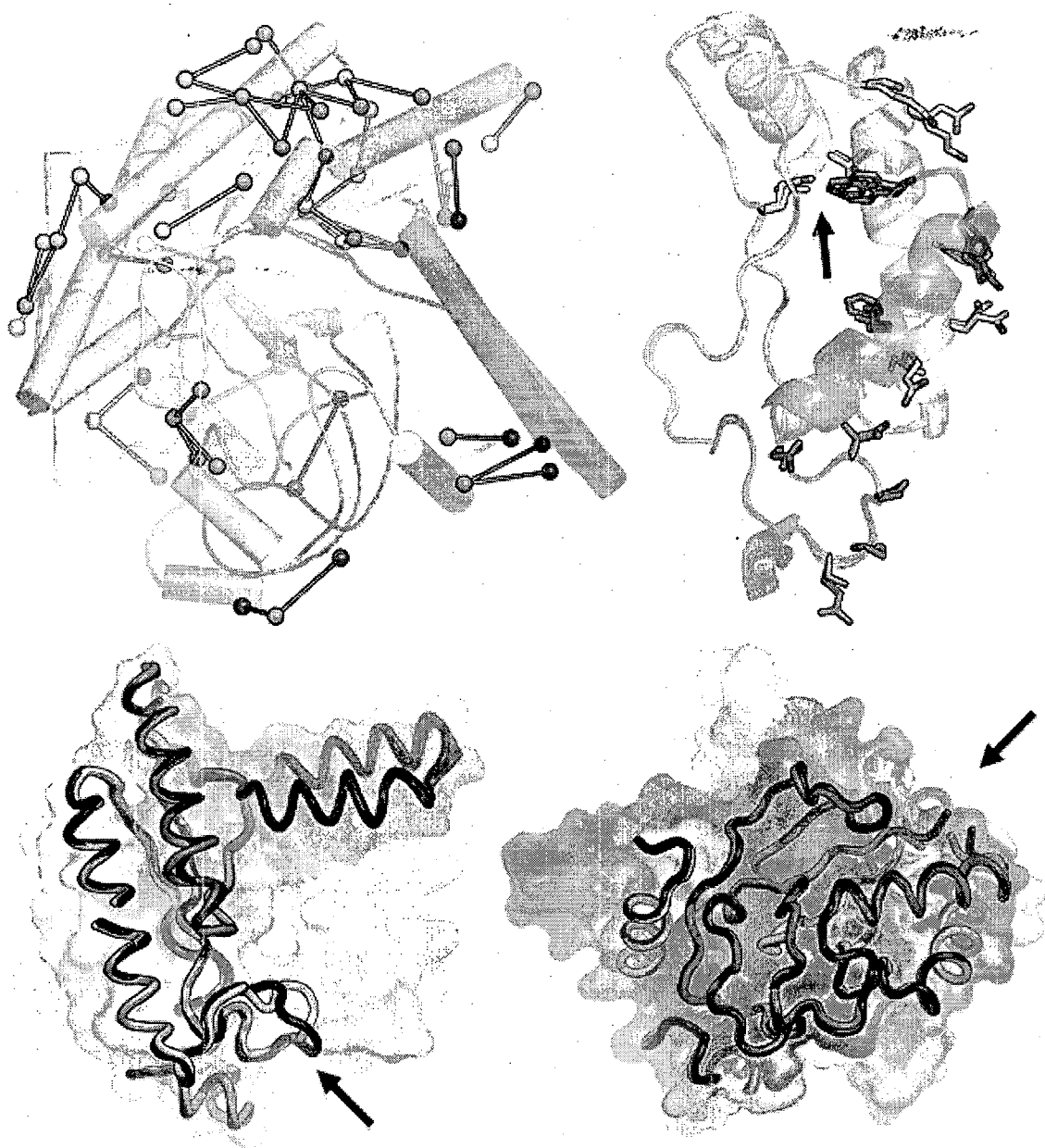


FIGURA 14A-D

80

11/11

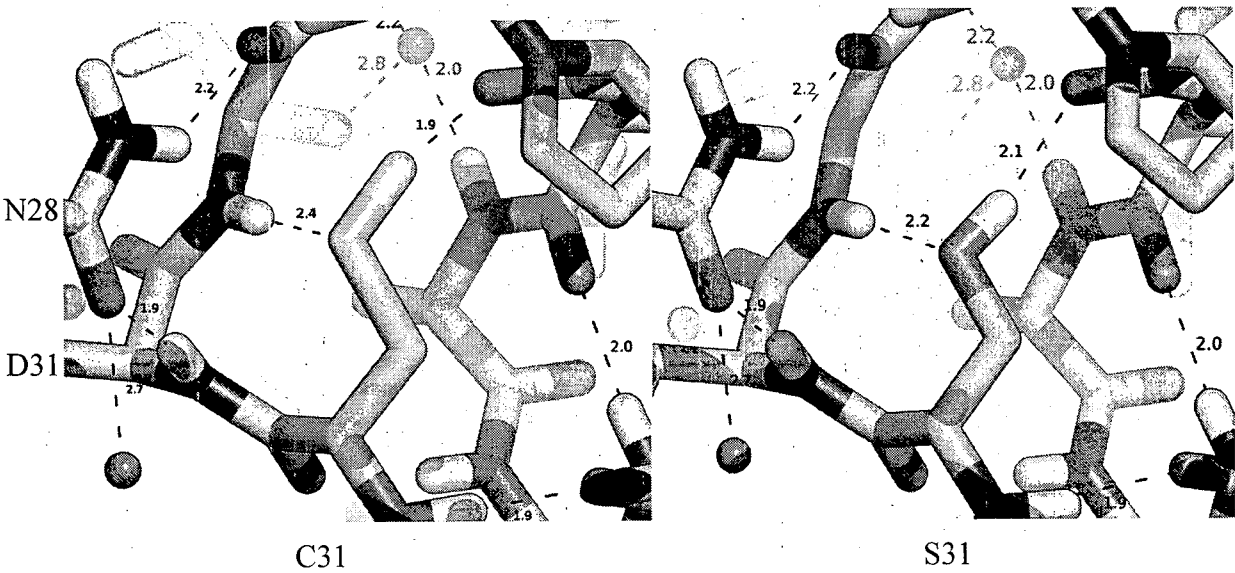


FIGURA 15

46

LISTADO DE SECUENCIAS

<110> California Institute of Technology
Arnold, Frances H.
Heinselman, Pete

<120> POLIPÉPTIDOS QUE TIENEN ACTIVIDAD DE CELULASA

<130> 00016-047WO1

<140> Todavía desconocida

<141> 2010-04-06

<150> US 61/166,993

<151> 2009-04-06

<150> US 61/177,882

<151> 2009-05-13

<160> 107

<170> PatentIn versión 3.5

<210> 1

<211> 1083

<212> ADN

<213> Humicola insolens

<220>

<221> CDS

<222> (1)..(1083)

<400> 1

ggt aac ccc ttt gaa ggt gtt cag ctg tgg gct aat aac tat tat aga 48
Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

tct gag gta cat aca ctg gcc att ccg caa att aca gac ccc gcg ttg 96
Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

cgt gcc gca gct agt gct gtg gct gag gtg cca agt ttt caa tgg ctg 144
Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45

gac aga aat gta aca gtg gat act ttg ttg gta cag act ttg tca gaa 192
Asp Arg Asn Val Thr Val Asp Thr Leu Leu Val Gln Thr Leu Ser Glu
50 55 60

atc cgt gag gcc aat caa gca ggt gct aat ccc caa tat gca gcg caa 240
Ile Arg Glu Ala Asn Gln Ala Gly Ala Asn Pro Gln Tyr Ala Ala Gln
65 70 75 80

atc gtg gtc tat gat ctg ccc gat aga gac tgt gca gct gcc gcc tgg 288
Ile Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser
85 90 95

aat ggt gaa tgg gca ata gcg aac aac ggt gta aac aat tac aaa gct 336
Asn Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala
100 105 110

2
6

tac att aat aga att aga gag ata ttg ata agt ttt tgg gac gtt aga 384
Tyr Ile Asn Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg
115 120 125

acg ata tta gtc att gag cca gat agt cta gct aat atg gtc aca aat 432
Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn
130 135 140

atg aat gtc ccg aag tgt tcc ggt gca gcc agc act tat agg gaa tta 480
Met Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Thr Tyr Arg Glu Leu
145 150 155 160

acc ata tat gca ctg aag caa ttg gat ctg cct cat gtc gct atg tac 528
Thr Ile Tyr Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr
165 170 175

atg gat gcc ggc cac gct gga tgg tta ggc tgg ccg gca aac att cag 576
Met Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln
180 185 190

cca gcc gca gaa ttg ttt gcc aaa att tac gaa gat gct gga aag cct 624
Pro Ala Ala Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro
195 200 205

aga gca gtg aga ggt ctt gca act aat gtt gct aat tac aat gca tgg 672
Arg Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp
210 215 220

tca gtt tca tcc cct cca cca tac aca agt cca aat cca aac tac gat 720
Ser Val Ser Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp
225 230 235 240

gaa aag cat tat atc gaa gca ttc aga ccc tta tta gaa gcc cgt ggt 768
Glu Lys His Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly
245 250 255

ttc cca gcc caa ttt ata gtg gat cag gga aga tca ggt aag caa cca 816
Phe Pro Ala Gln Phe Ile Val Asp Gln Gly Arg Ser Gly Lys Gln Pro
260 265 270

act ggc caa aag gag tgg ggg cat tgg tgt aat gct att ggc aca gga 864
Thr Gly Gln Lys Glu Trp Gly His Trp Cys Asn Ala Ile Gly Thr Gly
275 280 285

ttt ggt atg aga cct act gct aat acc ggt cac cag tat gtg gat gct 912
Phe Gly Met Arg Pro Thr Ala Asn Thr Gly His Gln Tyr Val Asp Ala
290 295 300

ttt gtt tgg gtt aaa ccg ggc ggt gaa tgc gac ggg acc agc gat act 960
Phe Val Trp Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp Thr
305 310 315 320

acg gcg gcc aga tat gat tat cat tgt ggt ctg gaa gat gca tta aaa 1008
Thr Ala Ala Arg Tyr Asp Tyr His Cys Gly Leu Glu Asp Ala Leu Lys
325 330 335

cca gct cct gaa gcc ggc cag tgg ttc aac gaa tac ttc att caa ttg 1056
Pro Ala Pro Glu Ala Gly Gln Trp Phe Asn Glu Tyr Phe Ile Gln Leu
340 345 350

ctt agg aac gct aac ccg ccc ttt taa 1083

Leu Arg Asn Ala Asn Pro Pro Phe
355 360

<210> 2
<211> 360
<212> PRT
<213> Humicola insolens

<400> 2

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45

Asp Arg Asn Val Thr Val Asp Thr Leu Leu Val Gln Thr Leu Ser Glu
50 55 60

Ile Arg Glu Ala Asn Gln Ala Gly Ala Asn Pro Gln Tyr Ala Ala Gln
65 70 75 80

Ile Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser
85 90 95

Asn Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala
100 105 110

Tyr Ile Asn Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg
115 120 125

Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn
130 135 140

Met Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Thr Tyr Arg Glu Leu
145 150 155 160

Thr Ile Tyr Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr
165 170 175

Met Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln
180 185 190

Pro Ala Ala Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro
195 200 205

Arg Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp
210 215 220

Ser Val Ser Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp
225 230 235 240

Glu Lys His Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly
245 250 255

Phe Pro Ala Gln Phe Ile Val Asp Gln Gly Arg Ser Gly Lys Gln Pro
260 265 270

Thr Gly Gln Lys Glu Trp Gly His Trp Cys Asn Ala Ile Gly Thr Gly

84

275 280 285

Phe Gly Met Arg Pro Thr Ala Asn Thr Gly His Gln Tyr Val Asp Ala
290 295 300

Phe Val Trp Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp Thr
305 310 315 320

Thr Ala Ala Arg Tyr Asp Tyr His Cys Gly Leu Glu Asp Ala Leu Lys
325 330 335

Pro Ala Pro Glu Ala Gly Gln Trp Phe Asn Glu Tyr Phe Ile Gln Leu
340 345 350

Leu Arg Asn Ala Asn Pro Pro Phe
355 360

<210> 3
<211> 1077
<212> ADN
<213> Hipocrea jecorina

<220>
<221> CDS
<222> (1)..(1077)

<400> 3
ggt aat cca ttc gtt ggg gtg aca ccc tgg gcg aac gcc tat tat gct 48
Gly Asn Pro Phe Val Gly Val Thr Pro Trp Ala Asn Ala Tyr Tyr Ala
1 5 10 15

tct gag gtt tca tcc cta gct att ccc tct tta aca ggt gca atg gct 96
Ser Glu Val Ser Ser Leu Ala Ile Pro Ser Leu Thr Gly Ala Met Ala
20 25 30

aca gcc gcc gct gcc gtt gca aag gtc cct tcc ttc atg tgg ctg gat 144
Thr Ala Ala Ala Ala Val Ala Lys Val Pro Ser Phe Met Trp Leu Asp
35 40 45

act ttg gac aaa acc ccc tta atg gaa caa acg ttg gct gat ata cgt 192
Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile Arg
50 55 60

act gcg aat aaa aac ggc ggc aat tat gct gga caa ttt gtg gtt tat 240
Thr Ala Asn Lys Asn Gly Gly Asn Tyr Ala Gly Gln Phe Val Val Tyr
65 70 75 80

gac ctg ccg gat aga gat tgt gct gca cta gcg agc aac ggg gag tac 288
Asp Leu Pro Asp Arg Asp Cys Ala Ala Leu Ala Ser Asn Gly Glu Tyr
85 90 95

agc att gcg gat ggc ggt gtc gca aag tac aaa aac tat ata gat act 336
Ser Ile Ala Asp Gly Gly Val Ala Lys Tyr Lys Asn Tyr Ile Asp Thr
100 105 110

atc agg caa ata gtt gtc gaa tac agt gat att cgt acg ctg ctt gta 384
Ile Arg Gln Ile Val Val Glu Tyr Ser Asp Ile Arg Thr Leu Leu Val
115 120 125

atc gaa ccc gat tcc tta gcg aac ttg gta aca aat cta ggt act ccg 432
Ile Glu Pro Asp Ser Leu Ala Asn Leu Val Thr Asn Leu Gly Thr Pro
130 135 140

45

aag tgt gcg aac gcg cag agt gct tat ctt gag tgc atc aat tat gca 480
Lys Cys Ala Asn Ala Gln Ser Ala Tyr Leu Glu Cys Ile Asn Tyr Ala
145 150 155 160

gtc acc cag ttg aat ttg cca aac gtt gca atg tat ctt gat gct ggt 528
Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu Asp Ala Gly
165 170 175

cat gcc ggg tgg ttg ggt tgg cca gca aat cag gat ccc gct gcg cag 576
His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp Pro Ala Ala Gln
180 185 190

ctg ttt gca aat gtt tac aaa aat gcc tca agt cct aga gcg ctg agg 624
Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro Arg Ala Leu Arg
195 200 205

ggt ctt gca aca aat gtt gct aat tac aac gga tgg aat att acc tca 672
Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Gly Trp Asn Ile Thr Ser
210 215 220

ccc cca tca tac aca caa gga aat gct gtt tac aat gaa aaa ctt tat 720
Pro Pro Ser Tyr Thr Gln Gly Asn Ala Val Tyr Asn Glu Lys Leu Tyr
225 230 235 240

att cat gcc att ggt cca ctg ctg gct aat cac gga tgg agt aat gcc 768
Ile His Ala Ile Gly Pro Leu Leu Ala Asn His Gly Trp Ser Asn Ala
245 250 255

ttt ttc att aca gat caa ggg aga agt ggt aaa caa cct act gga caa 816
Phe Phe Ile Thr Asp Gln Gly Arg Ser Gly Lys Gln Pro Thr Gly Gln
260 265 270

caa caa tgg ggt gac tgg tgt aat gtt atc ggt act ggg ttt ggc atc 864
Gln Gln Trp Gly Asp Trp Cys Asn Val Ile Gly Thr Gly Phe Gly Ile
275 280 285

aga cca tca gcg aat acg ggt gat tca ttg ttg gac tca ttt gtt tgg 912
Arg Pro Ser Ala Asn Thr Gly Asp Ser Leu Leu Asp Ser Phe Val Trp
290 295 300

gtt aaa ccc ggg ggt gaa tgt gat gga acg agt gat tct tct gct cca 960
Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp Ser Ser Ala Pro
305 310 315 320

agg ttc gat tct cat tgc gca tta cca gat gct ttg cag cca gca cct 1008
Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro Ala Pro
325 330 335

caa gca gga gct tgg ttc caa gct tat ttt gta caa tta ctg act aac 1056
Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu Thr Asn
340 345 350

gcc aat cct agt ttt cta taa 1077
Ala Asn Pro Ser Phe Leu
355

<210> 4
<211> 358
<212> PRT
<213> Hipocrea jecorina

<400> 4

Gly Asn Pro Phe Val Gly Val Thr Pro Trp Ala Asn Ala Tyr Tyr Ala
1 5 10 15

Ser Glu Val Ser Ser Leu Ala Ile Pro Ser Leu Thr Gly Ala Met Ala
20 25 30

Thr Ala Ala Ala Val Ala Lys Val Pro Ser Phe Met Trp Leu Asp
35 40 45

Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile Arg
50 55 60

Thr Ala Asn Lys Asn Gly Gly Asn Tyr Ala Gly Gln Phe Val Val Tyr
65 70 75 80

Asp Leu Pro Asp Arg Asp Cys Ala Ala Leu Ala Ser Asn Gly Glu Tyr
85 90 95

Ser Ile Ala Asp Gly Gly Val Ala Lys Tyr Lys Asn Tyr Ile Asp Thr
100 105 110

Ile Arg Gln Ile Val Val Glu Tyr Ser Asp Ile Arg Thr Leu Leu Val
115 120 125

Ile Glu Pro Asp Ser Leu Ala Asn Leu Val Thr Asn Leu Gly Thr Pro
130 135 140

Lys Cys Ala Asn Ala Gln Ser Ala Tyr Leu Glu Cys Ile Asn Tyr Ala
145 150 155 160

Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu Asp Ala Gly
165 170 175

His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp Pro Ala Ala Gln
180 185 190

Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro Arg Ala Leu Arg
195 200 205

Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Gly Trp Asn Ile Thr Ser
210 215 220

Pro Pro Ser Tyr Thr Gln Gly Asn Ala Val Tyr Asn Glu Lys Leu Tyr
225 230 235 240

Ile His Ala Ile Gly Pro Leu Leu Ala Asn His Gly Trp Ser Asn Ala
245 250 255

Phe Phe Ile Thr Asp Gln Gly Arg Ser Gly Lys Gln Pro Thr Gly Gln
260 265 270

Gln Gln Trp Gly Asp Trp Cys Asn Val Ile Gly Thr Gly Phe Gly Ile
275 280 285

Arg Pro Ser Ala Asn Thr Gly Asp Ser Leu Leu Asp Ser Phe Val Trp
290 295 300

Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp Ser Ser Ala Pro
305 310 315 320

Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro Ala Pro
325 330 335

Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu Thr Asn
340 345 350

Ala Asn Pro Ser Phe Leu
355

<210> 5
<211> 1077
<212> ADN
<213> Chaetomium termophilium

<220>
<221> CDS
<222> (1)..(1077)

<400> 5
ggt aac cct ttc agt ggt gtg cag tta tgg gct aat act tac tat tct 48
Gly Asn Pro Phe Ser Gly Val Gln Leu Trp Ala Asn Thr Tyr Tyr Ser
1 5 10 15

tca gaa gtc cac acc tta gct atc cca agc tta agt cca gaa tta gcg 96
Ser Glu Val His Thr Leu Ala Ile Pro Ser Leu Ser Pro Glu Leu Ala
20 25 30

gct aag gcg gcg aaa gta gct gaa gtg cca tca ttc caa tgg tta gat 144
Ala Lys Ala Ala Lys Val Ala Glu Val Pro Ser Phe Gln Trp Leu Asp
35 40 45

aga aac gtg act gtg gat act ctg ttt tct ggt aca ctt gct gag ata 192
Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Glu Ile
50 55 60

agg gcg gct aac caa agg gga gct aat cca cca tat gct ggc atc ttt 240
Arg Ala Ala Asn Gln Arg Gly Ala Asn Pro Pro Tyr Ala Gly Ile Phe
65 70 75 80

gtg gtt tat gac ctt cct gat aga gat tgt gct gcc gct gca agc aat 288
Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn
85 90 95

ggt gaa tgg agt ata gct aac aac ggt gct aac aac tat aag aga tat 336
Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg Tyr
100 105 110

atc gat aga att aga gaa ttg ttg att cag tac tca gat atc agg aca 384
Ile Asp Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg Thr
115 120 125

att ttg gtt att gaa cca gac agt cta gca aat atg gtt act aac atg 432
Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

aac gta caa aaa tgt tct aac gca gca tct acg tat aaa gaa ctg act 480
Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Thr Tyr Lys Glu Leu Thr
145 150 155 160

gtg tat gca ttg aaa cag ttg aac ttg cca cac gta gcc atg tat atg 528
Val Tyr Ala Leu Lys Gln Leu Asn Leu Pro His Val Ala Met Tyr Met
165 170 175

48

gat gca ggt cac gcc ggc tgg tta ggc tgg ccc gct aat ata cag cct 576
Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro
180 185 190

gcc gca gaa tta ttc gcg caa ata tac aga gac gct gga cgt ccg gct 624
Ala Ala Glu Leu Phe Ala Gln Ile Tyr Arg Asp Ala Gly Arg Pro Ala
195 200 205

gcg gtc agg ggt ctt gcc act aac gtt gca aat tac aac gct tgg tca 672
Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220

ata gcg agt cct cca tcg tac aca agc cct aac cca aac tac gat gag 720
Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu
225 230 235 240

aag cat tac ata gaa gca ttt gct cct ttg ctt cgt aac caa ggt ttt 768
Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly Phe
245 250 255

gat gca aag ttt atc gtc gat acc gga aga aac ggc aag cag ccg aca 816
Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr
260 265 270

ggg cag cta gaa tgg ggg cac tgg tgc aat gtc aag ggt acg ggt ttc 864
Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe
275 280 285

ggt gtt aga ccc acg gct aac act ggg cat gag ttg gtt gat gca ttc 912
Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe
290 295 300

gtt tgg gta aaa ccc gga gga gag tca gac ggt act tct gat act agt 960
Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Thr Ser
305 310 315 320

gct gcc aga tac gat tac cac tgt ggc ctt tct gat gct ttg aca cca 1008
Ala Ala Arg Tyr Asp Tyr His Cys Gly Leu Ser Asp Ala Leu Thr Pro
325 330 335

gcc cct gaa gcc ggg caa tgg ttc cag gcc tac ttc gaa caa cta ttg 1056
Ala Pro Glu Ala Gly Gln Trp Phe Gln Ala Tyr Phe Glu Gln Leu Leu
340 345 350

att aac gca aac cca cca tag 1077
Ile Asn Ala Asn Pro Pro
355

<210> 6
<211> 358
<212> PRT
<213> Chaetomium termophilium

<400> 6

Gly Asn Pro Phe Ser Gly Val Gln Leu Trp Ala Asn Thr Tyr Tyr Ser
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Ser Leu Ser Pro Glu Leu Ala
20 25 30

89

Ala Lys Ala Ala Lys Val Ala Glu Val Pro Ser Phe Gln Trp Leu Asp
35 40 45

Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Glu Ile
50 55 60

Arg Ala Ala Asn Gln Arg Gly Ala Asn Pro Pro Tyr Ala Gly Ile Phe
65 70 75 80

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn
85 90 95

Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg Tyr
100 105 110

Ile Asp Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg Thr
115 120 125

Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Thr Tyr Lys Glu Leu Thr
145 150 155 160

Val Tyr Ala Leu Lys Gln Leu Asn Leu Pro His Val Ala Met Tyr Met
165 170 175

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro
180 185 190

Ala Ala Glu Leu Phe Ala Gln Ile Tyr Arg Asp Ala Gly Arg Pro Ala
195 200 205

Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220

Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu
225 230 235 240

Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly Phe
245 250 255

Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr
260 265 270

Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe
275 280 285

Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe
290 295 300

Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Thr Ser
305 310 315 320

Ala Ala Arg Tyr Asp Tyr His Cys Gly Leu Ser Asp Ala Leu Thr Pro
325 330 335

Ala Pro Glu Ala Gly Gln Trp Phe Gln Ala Tyr Phe Glu Gln Leu Leu
340 345 350

Ile Asn Ala Asn Pro Pro
355

90

<210> 7
<211> 267
<212> ADN
<213> Secuencia artificial

<220>
<223> Conector CBD

<220>
<221> CDS
<222> (1)..(267)

<400> 7
gct agc tgc tca agc gtc tgg ggc caa tgt ggt ggc cag aat tgg tcg 48
Ala Ser Cys Ser Ser Val Trp Gly Gln Cys Gly Gly Gln Asn Trp Ser
1 5 10 15

ggt ccg act tgc tgt gct tcc gga agc aca tgc gtc tac tcc aac gac 96
Gly Pro Thr Cys Cys Ala Ser Gly Ser Thr Cys Val Tyr Ser Asn Asp
20 25 30

tat tac tcc cag tgt ctt ccc ggc gct gca agc tca agc tcg tcc acg 144
Tyr Tyr Ser Gln Cys Leu Pro Gly Ala Ala Ser Ser Ser Ser Thr
35 40 45

cgc gcc gcg tcg acg act tct cga gta tcc ccc aca aca tcc cgg tcg 192
Arg Ala Ala Ser Thr Thr Ser Arg Val Ser Pro Thr Thr Ser Arg Ser
50 55 60

agc tcc gcg acg cct cca cct ggt tct act act acc aga gta cct cca 240
Ser Ser Ala Thr Pro Pro Pro Gly Ser Thr Thr Thr Arg Val Pro Pro
65 70 75 80

gtc gga tcg gga acc gct acg tat tca 267
Val Gly Ser Gly Thr Ala Thr Tyr Ser
85

<210> 8
<211> 89
<212> PRT
<213> Secuencia artificial

<220>
<223> Constructo sintético

<400> 8
Ala Ser Cys Ser Ser Val Trp Gly Gln Cys Gly Gly Gln Asn Trp Ser
1 5 10 15

Gly Pro Thr Cys Cys Ala Ser Gly Ser Thr Cys Val Tyr Ser Asn Asp
20 25 30

Tyr Tyr Ser Gln Cys Leu Pro Gly Ala Ala Ser Ser Ser Ser Ser Thr
35 40 45

Arg Ala Ala Ser Thr Thr Ser Arg Val Ser Pro Thr Thr Ser Arg Ser
50 55 60

Ser Ser Ala Thr Pro Pro Pro Gly Ser Thr Thr Thr Arg Val Pro Pro
65 70 75 80

91

Val Gly Ser Gly Thr Ala Thr Tyr Ser
85

<210> 9
<211> 21
<212> ADN
<213> Secuencia artificial

<220>
<223> Iniciador de oligonucleótido CBH2L

<400> 9
gctgaacgtg tcatcggtta c 21

<210> 10
<211> 23
<212> ADN
<213> Secuencia artificial

<220>
<223> Iniciador de oligonucleótido RSQ3080

<400> 10
gcaacacctg gcaattcctt acc 23

<210> 11
<211> 24
<212> ADN
<213> Secuencia artificial

<220>
<223> Iniciador de oligonucleótido CBH2LPCR

<400> 11
gctgaacgtg tcatcggttac ttag 24

<210> 12
<211> 356
<212> PRT
<213> Phanerochaete chrysosporium

<400> 12

Asn Asn Pro Trp Thr Gly Phe Gln Ile Phe Leu Ser Pro Tyr Tyr Ala
1 5 10 15

Asn Glu Val Ala Ala Ala Lys Gln Ile Thr Asp Pro Thr Leu Ser
20 25 30

Ser Lys Ala Ala Ser Val Ala Asn Ile Pro Thr Phe Thr Trp Leu Asp
35 40 45

Ser Val Ala Lys Ile Pro Asp Leu Gly Thr Tyr Leu Ala Ser Ala Ser
50 55 60

Ala Leu Gly Lys Ser Thr Gly Thr Lys Gln Leu Val Gln Ile Val Ile
65 70 75 80

Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Lys Ala Ser Asn Gly Glu
85 90 95

92

Phe Ser Ile Ala Asn Asn Gly Gln Ala Asn Tyr Glu Asn Tyr Ile Asp
100 105 110

Gln Ile Val Ala Gln Ile Gln Gln Phe Pro Asp Val Arg Val Val Ala
115 120 125

Val Ile Glu Pro Asp Ser Leu Ala Asn Leu Val Thr Asn Leu Asn Val
130 135 140

Gln Lys Cys Ala Asn Ala Lys Thr Thr Tyr Leu Ala Cys Val Asn Tyr
145 150 155 160

Ala Leu Thr Asn Leu Ala Lys Val Gly Val Tyr Met Tyr Met Asp Ala
165 170 175

Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Leu Ser Pro Ala Ala
180 185 190

Gln Leu Phe Thr Gln Val Trp Gln Asn Ala Gly Lys Ser Pro Phe Ile
195 200 205

Lys Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Leu Gln Ala Ala
210 215 220

Ser Pro Asp Pro Ile Thr Gln Gly Asn Pro Asn Tyr Asp Glu Ile His
225 230 235 240

Tyr Ile Asn Ala Leu Ala Pro Leu Leu Gln Gln Ala Gly Trp Asp Ala
245 250 255

Thr Phe Ile Val Asp Gln Gly Arg Ser Gly Val Gln Asn Ile Arg Gln
260 265 270

Gln Trp Gly Asp Trp Cys Asn Ile Lys Gly Ala Gly Phe Gly Thr Arg
275 280 285

Pro Thr Thr Asn Thr Gly Ser Gln Phe Ile Asp Ser Ile Val Trp Val
290 295 300

Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asn Ser Ser Ser Pro Arg
305 310 315 320

Tyr Asp Ser Thr Cys Ser Leu Pro Asp Ala Ala Gln Pro Ala Pro Glu
325 330 335

Ala Gly Thr Trp Phe Gln Ala Tyr Phe Gln Thr Leu Val Ser Ala Ala
340 345 350

Asn Pro Pro Leu
355

<210> 13

<211> 358

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 13

Gly Asn Pro Phe Ser Gly Val Gln Leu Trp Ala Asn Thr Tyr Tyr Ser

93

1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Ser Leu Ser Pro Glu Leu Ala
20 25 30

Ala Lys Ala Ala Lys Val Ala Glu Val Pro Ser Phe Met Trp Leu Asp
35 40 45

Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile Arg
50 55 60

Thr Ala Asn Lys Asn Gly Ala Asn Pro Pro Tyr Ala Gly Ile Phe Val
65 70 75 80

Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn Gly
85 90 95

Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg Tyr Ile
100 105 110

Asp Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg Thr Ile
115 120 125

Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met Asn
130 135 140

Val Gln Lys Cys Ser Asn Ala Ala Ser Thr Tyr Lys Glu Leu Thr Val
145 150 155 160

Tyr Ala Leu Lys Gln Leu Asn Leu Pro His Val Ala Met Tyr Met Asp
165 170 175

Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro Ala
180 185 190

Ala Glu Leu Phe Ala Gln Ile Tyr Arg Asp Ala Gly Arg Pro Ala Ala
195 200 205

Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser Val
210 215 220

Ser Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu Lys
225 230 235 240

His Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly Phe Pro
245 250 255

Ala Gln Phe Ile Val Asp Gln Gly Arg Ser Gly Lys Gln Pro Thr Gly
260 265 270

Gln Lys Glu Trp Gly His Trp Cys Asn Ala Ile Gly Thr Gly Phe Gly
275 280 285

Met Arg Pro Thr Ala Asn Thr Gly His Gln Tyr Val Asp Ala Phe Val
290 295 300

Trp Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp Thr Ser Ala
305 310 315 320

Ala Arg Tyr Asp Tyr His Cys Gly Leu Ser Asp Ala Leu Thr Pro Ala
325 330 335

94

Pro Glu Ala Gly Gln Trp Phe Gln Ala Tyr Phe Glu Gln Leu Leu Ile
340 345 350

Asn Ala Asn Pro Pro Phe
355

<210> 14
<211> 360
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido quimérico CBH II

<400> 14

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45

Asp Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Glu
50 55 60

Ile Arg Ala Ala Asn Gln Arg Gly Ala Asn Pro Gln Tyr Ala Ala Gln
65 70 75 80

Ile Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser
85 90 95

Asn Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala
100 105 110

Tyr Ile Asn Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg
115 120 125

Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn
130 135 140

Met Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Thr Tyr Arg Glu Leu
145 150 155 160

Thr Ile Tyr Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr
165 170 175

Met Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln
180 185 190

Pro Ala Ala Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro
195 200 205

Arg Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp
210 215 220

Ser Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp
225 230 235 240

Glu Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly

245 250 255

Phe Asp Ala Gln Phe Ile Val Asp Gln Gly Arg Ser Gly Lys Gln Pro
260 265 270

Thr Gly Gln Lys Glu Trp Gly His Trp Cys Asn Ala Ile Gly Thr Gly
275 280 285

Phe Gly Met Arg Pro Thr Ala Asn Thr Gly His Gln Tyr Val Asp Ala
290 295 300

Phe Val Trp Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp Thr
305 310 315 320

Ser Ala Ala Arg Tyr Asp Tyr His Cys Gly Leu Ser Asp Ala Leu Thr
325 330 335

Pro Ala Pro Glu Ala Gly Gln Trp Phe Gln Ala Tyr Phe Glu Gln Leu
340 345 350

Leu Ile Asn Ala Asn Pro Pro Phe
355 360

<210> 15

<211> 361

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 15

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45

Asp Arg Asn Val Thr Val Asp Thr Leu Leu Val Gln Thr Leu Ser Glu
50 55 60

Ile Arg Glu Ala Asn Gln Ala Gly Ala Asn Pro Pro Tyr Ala Gly Ile
65 70 75 80

Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser
85 90 95

Asn Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg
100 105 110

Tyr Ile Asp Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg
115 120 125

Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn
130 135 140

Met Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Thr Tyr Lys Glu Leu
145 150 155 160

96

Thr Val Tyr Ala Leu Lys Gln Leu Asn Leu Pro His Val Ala Met Tyr
165 170 175

Met Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln
180 185 190

Pro Ala Ala Glu Leu Phe Ala Gln Ile Tyr Arg Asp Ala Gly Arg Pro
195 200 205

Ala Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp
210 215 220

Ser Val Ser Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp
225 230 235 240

Glu Lys His Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly
245 250 255

Phe Pro Asn Ala Phe Phe Ile Thr Asp Gln Gly Arg Ser Gly Lys Gln
260 265 270

Pro Thr Gly Gln Gln Gln Trp Gly Asp Trp Cys Asn Val Ile Gly Thr
275 280 285

Gly Phe Gly Ile Arg Pro Ser Ala Asn Thr Gly Asp Ser Leu Leu Asp
290 295 300

Ser Phe Val Trp Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp
305 310 315 320

Thr Thr Ala Ala Arg Tyr Asp Tyr His Cys Gly Leu Glu Asp Ala Leu
325 330 335

Lys Pro Ala Pro Glu Ala Gly Gln Trp Phe Asn Glu Tyr Phe Ile Gln
340 345 350

Leu Leu Arg Asn Ala Asn Pro Pro Phe
355 360

<210> 16
<211> 359
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido quimérico CBH II

<400> 16

Gly Asn Pro Phe Val Gly Val Thr Pro Trp Ala Asn Ala Tyr Tyr Ala
1 5 10 15

Ser Glu Val Ser Ser Leu Ala Ile Pro Ser Leu Thr Gly Ala Met Ala
20 25 30

Thr Ala Ala Ala Val Ala Lys Val Pro Ser Phe Gln Trp Leu Asp
35 40 45

Arg Asn Val Thr Val Asp Thr Leu Leu Val Gln Thr Leu Ser Glu Ile
50 55 60

99

Arg Glu Ala Asn Gln Ala Gly Ala Asn Pro Gln Tyr Ala Ala Gln Ile
65 70 75 80

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ser Asn
85 90 95

Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala Tyr
100 105 110

Ile Asn Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg Thr
115 120 125

Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Thr Tyr Arg Glu Leu Thr
145 150 155 160

Ile Tyr Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr Met
165 170 175

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro
180 185 190

Ala Ala Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro Arg
195 200 205

Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220

Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu
225 230 235 240

Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly Phe
245 250 255

Asp Ala Gln Phe Ile Val Asp Gln Gly Arg Ser Gly Lys Gln Pro Thr
260 265 270

Gly Gln Lys Glu Trp Gly His Trp Cys Asn Ala Ile Gly Thr Gly Phe
275 280 285

Gly Met Arg Pro Thr Ala Asn Thr Gly His Gln Tyr Val Asp Ala Phe
290 295 300

Val Trp Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp Thr Thr
305 310 315 320

Ala Ala Arg Tyr Asp Tyr His Cys Gly Leu Glu Asp Ala Leu Lys Pro
325 330 335

Ala Pro Glu Ala Gly Gln Trp Phe Asn Glu Tyr Phe Ile Gln Leu Leu
340 345 350

Arg Asn Ala Asn Pro Pro Phe
355

<210> 17

<211> 357

<212> PRT

<213> Secuencia artificial

98

<220>

<223> Polipéptido quimérico CBH II

<400> 17

Gly Asn Pro Phe Ser Gly Val Gln Leu Trp Ala Asn Thr Tyr Tyr Ser
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Ser Leu Ser Pro Glu Leu Ala
20 25 30

Ala Lys Ala Ala Lys Val Ala Glu Val Pro Ser Phe Gln Trp Leu Asp
35 40 45

Arg Asn Val Thr Val Asp Thr Leu Leu Val Gln Thr Leu Ser Glu Ile
50 55 60

Arg Glu Ala Asn Gln Ala Gly Gly Asn Tyr Ala Gly Gln Phe Val Val
65 70 75 80

Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Leu Ala Ser Asn Gly Glu
85 90 95

Tyr Ser Ile Ala Asp Gly Gly Val Ala Lys Tyr Lys Asn Tyr Ile Asp
100 105 110

Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg Thr Ile Leu
115 120 125

Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met Asn Val
130 135 140

Pro Lys Cys Ser Gly Ala Ala Ser Ala Tyr Leu Glu Cys Ile Asn Tyr
145 150 155 160

Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu Asp Ala
165 170 175

Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp Pro Ala Ala
180 185 190

Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro Arg Ala Leu
195 200 205

Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser Val Ser
210 215 220

Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu Lys His
225 230 235 240

Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly Phe Pro Ala
245 250 255

Gln Phe Ile Val Asp Gln Gly Arg Ser Gly Lys Gln Pro Thr Gly Gln
260 265 270

Lys Glu Trp Gly His Trp Cys Asn Ala Ile Gly Thr Gly Phe Gly Met
275 280 285

Arg Pro Thr Ala Asn Thr Gly His Gln Tyr Val Asp Ala Phe Val Trp
290 295 300

90

Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp Thr Thr Ala Ala
305 310 315 320

Arg Tyr Asp Tyr His Cys Gly Leu Glu Asp Ala Leu Lys Pro Ala Pro
325 330 335

Glu Ala Gly Gln Trp Phe Asn Glu Tyr Phe Ile Gln Leu Leu Arg Asn
340 345 350

Ala Asn Pro Pro Phe
355

<210> 18

<211> 357

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 18

Gly Asn Pro Phe Val Gly Val Thr Pro Trp Ala Asn Ala Tyr Tyr Ala
1 5 10 15

Ser Glu Val Ser Ser Leu Ala Ile Pro Ser Leu Thr Gly Ala Met Ala
20 25 30

Thr Ala Ala Ala Val Ala Lys Val Pro Ser Phe Gln Trp Leu Asp
35 40 45

Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Glu Ile
50 55 60

Arg Ala Ala Asn Gln Arg Gly Gly Asn Tyr Ala Gly Gln Phe Val Val
65 70 75 80

Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Leu Ala Ser Asn Gly Glu
85 90 95

Tyr Ser Ile Ala Asp Gly Gly Val Ala Lys Tyr Lys Asn Tyr Ile Asp
100 105 110

Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg Thr Ile Leu
115 120 125

Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met Asn Val
130 135 140

Gln Lys Cys Ser Asn Ala Ala Ser Thr Tyr Lys Glu Leu Thr Val Tyr
145 150 155 160

Ala Leu Lys Gln Leu Asn Leu Pro His Val Ala Met Tyr Met Asp Ala
165 170 175

Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro Ala Ala
180 185 190

Glu Leu Phe Ala Gln Ile Tyr Arg Asp Ala Gly Arg Pro Ala Ala Val
195 200 205

Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser Val Ser

91

210 215 220

Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu Lys His
225 230 235 240

Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly Phe Pro Ala
245 250 255

Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr Gly Gln
260 265 270

Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe Gly Val
275 280 285

Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe Val Trp
290 295 300

Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Thr Ser Ala Ala
305 310 315 320

Arg Tyr Asp Tyr His Cys Gly Leu Ser Asp Ala Leu Thr Pro Ala Pro
325 330 335

Glu Ala Gly Gln Trp Phe Gln Ala Tyr Phe Glu Gln Leu Leu Ile Asn
340 345 350

Ala Asn Pro Pro Phe
355

<210> 19

<211> 360

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 19

Gly Asn Pro Phe Ser Gly Val Gln Leu Trp Ala Asn Thr Tyr Tyr Ser
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Ser Leu Ser Pro Glu Leu Ala
20 25 30

Ala Lys Ala Ala Lys Val Ala Glu Val Pro Ser Phe Gln Trp Leu Asp
35 40 45

Arg Asn Val Thr Val Asp Thr Leu Leu Val Gln Thr Leu Ser Glu Ile
50 55 60

Arg Glu Ala Asn Gln Ala Gly Ala Asn Pro Pro Tyr Ala Gly Ile Phe
65 70 75 80

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ser Asn
85 90 95

Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg Tyr
100 105 110

Ile Asp Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg Thr
115 120 125

92

Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Thr Tyr Arg Glu Leu Thr
145 150 155 160

Ile Tyr Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr Met
165 170 175

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro
180 185 190

Ala Ala Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro Arg
195 200 205

Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220

Val Ser Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu
225 230 235 240

Lys His Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly Phe
245 250 255

Pro Ala Gln Phe Ile Val Asp Gln Gly Arg Ser Gly Lys Gln Pro Thr
260 265 270

Gly Gln Lys Glu Trp Gly His Trp Cys Asn Ala Ile Gly Thr Gly Phe
275 280 285

Gly Met Arg Pro Thr Ala Asn Thr Gly His Gln Tyr Val Asp Ala Phe
290 295 300

Val Trp Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp Ser Ser
305 310 315 320

Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro
325 330 335

Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu
340 345 350

Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 20

<211> 356

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 20

Gly Asn Pro Phe Val Gly Val Thr Pro Trp Ala Asn Ala Tyr Tyr Ala
1 5 10 15

Ser Glu Val Ser Ser Leu Ala Ile Pro Ser Leu Thr Gly Ala Met Ala
20 25 30

93

Thr Ala Ala Ala Ala Val Ala Lys Val Pro Ser Phe Met Trp Leu Asp
35 40 45

Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile Arg
50 55 60

Thr Ala Asn Lys Asn Gly Gly Asn Tyr Ala Gly Gln Phe Val Val Tyr
65 70 75 80

Asp Leu Pro Asp Arg Asp Cys Ala Ala Leu Ala Ser Asn Gly Glu Tyr
85 90 95

Ser Ile Ala Asp Gly Gly Val Ala Lys Tyr Lys Asn Tyr Ile Asp Arg
100 105 110

Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg Thr Ile Leu Val
115 120 125

Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met Asn Val Pro
130 135 140

Lys Cys Ser Gly Ala Ala Ser Ala Tyr Leu Glu Cys Ile Asn Tyr Ala
145 150 155 160

Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu Asp Ala Gly
165 170 175

His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp Pro Ala Ala Gln
180 185 190

Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro Arg Ala Leu Arg
195 200 205

Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Gly Trp Asn Ile Thr Ser
210 215 220

Pro Pro Ser Tyr Thr Gln Gly Asn Ala Val Tyr Asn Glu Lys Leu Tyr
225 230 235 240

Ile His Ala Ile Gly Pro Leu Leu Ala Asn His Gly Trp Ser Ala Lys
245 250 255

Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr Gly Gln Leu
260 265 270

Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe Gly Val Arg
275 280 285

Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe Val Trp Val
290 295 300

Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Thr Thr Ala Ala Arg
305 310 315 320

Tyr Asp Tyr His Cys Gly Leu Glu Asp Ala Leu Lys Pro Ala Pro Glu
325 330 335

Ala Gly Gln Trp Phe Asn Glu Tyr Phe Ile Gln Leu Leu Arg Asn Ala
340 345 350

Asn Pro Pro Phe
355

94

<210> 21
 <211> 358
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Polipéptido quimérico CBH II

<400> 21

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
 1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
 20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
 35 40 45

Asp Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Glu
 50 55 60

Ile Arg Ala Ala Asn Gln Arg Gly Gly Asn Tyr Ala Gly Gln Phe Val
 65 70 75 80

Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Leu Ala Ser Asn Gly
 85 90 95

Glu Tyr Ser Ile Ala Asp Gly Gly Val Ala Lys Tyr Lys Asn Tyr Ile
 100 105 110

Asp Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg Thr Ile
 115 120 125

Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met Asn
 130 135 140

Val Gln Lys Cys Ser Asn Ala Ala Ser Thr Tyr Arg Glu Leu Thr Ile
 145 150 155 160

Tyr Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr Met Asp
 165 170 175

Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro Ala
 180 185 190

Ala Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro Arg Ala
 195 200 205

Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser Val
 210 215 220

Ser Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu Lys
 225 230 235 240

His Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly Phe Pro
 245 250 255

Ala Gln Phe Ile Val Asp Gln Gly Arg Ser Gly Lys Gln Pro Thr Gly
 260 265 270

95

Gln Lys Glu Trp Gly His Trp Cys Asn Ala Ile Gly Thr Gly Phe Gly
275 280 285

Met Arg Pro Thr Ala Asn Thr Gly His Gln Tyr Val Asp Ala Phe Val
290 295 300

Trp Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp Thr Thr Ala
305 310 315 320

Ala Arg Tyr Asp Tyr His Cys Gly Leu Glu Asp Ala Leu Lys Pro Ala
325 330 335

Pro Glu Ala Gly Gln Trp Phe Asn Glu Tyr Phe Ile Gln Leu Leu Arg
340 345 350

Asn Ala Asn Pro Pro Phe
355

<210> 22

<211> 357

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 22

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Met Trp Leu
35 40 45

Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile
50 55 60

Arg Thr Ala Asn Lys Asn Gly Gly Asn Tyr Ala Gly Gln Phe Val Val
65 70 75 80

Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Leu Ala Ser Asn Gly Glu
85 90 95

Tyr Ser Ile Ala Asp Gly Gly Val Ala Lys Tyr Lys Asn Tyr Ile Asp
100 105 110

Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg Thr Ile Leu
115 120 125

Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met Asn Val
130 135 140

Pro Lys Cys Ser Gly Ala Ala Ser Thr Tyr Lys Glu Leu Thr Val Tyr
145 150 155 160

Ala Leu Lys Gln Leu Asn Leu Pro His Val Ala Met Tyr Met Asp Ala
165 170 175

Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro Ala Ala

96

180 185 190
Glu Leu Phe Ala Gln Ile Tyr Arg Asp Ala Gly Arg Pro Ala Ala Val
195 200 205
Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser Val Ser
210 215 220
Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu Lys His
225 230 235 240
Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly Phe Pro Ala
245 250 255
Gln Phe Ile Val Asp Gln Gly Arg Ser Gly Lys Gln Pro Thr Gly Gln
260 265 270
Lys Glu Trp Gly His Trp Cys Asn Ala Ile Gly Thr Gly Phe Gly Met
275 280 285
Arg Pro Thr Ala Asn Thr Gly His Gln Tyr Val Asp Ala Phe Val Trp
290 295 300
Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp Thr Thr Ala Ala
305 310 315 320
Arg Tyr Asp Tyr His Cys Gly Leu Glu Asp Ala Leu Lys Pro Ala Pro
325 330 335
Glu Ala Gly Gln Trp Phe Asn Glu Tyr Phe Ile Gln Leu Leu Arg Asn
340 345 350
Ala Asn Pro Pro Phe
355
<210> 23
<211> 359
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido quimérico CBH II

<400> 23
Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15
Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30
Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Met Trp Leu
35 40 45
Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile
50 55 60
Arg Thr Ala Asn Lys Asn Gly Ala Asn Pro Gln Tyr Ala Ala Gln Ile
65 70 75 80
Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn
85 90 95

97

Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala Tyr
100 105 110

Ile Asn Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg Thr
115 120 125

Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Thr Tyr Lys Glu Leu Thr
145 150 155 160

Val Tyr Ala Leu Lys Gln Leu Asn Leu Pro His Val Ala Met Tyr Met
165 170 175

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro
180 185 190

Ala Ala Glu Leu Phe Ala Gln Ile Tyr Arg Asp Ala Gly Arg Pro Ala
195 200 205

Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220

Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu
225 230 235 240

Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly Phe
245 250 255

Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr
260 265 270

Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe
275 280 285

Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe
290 295 300

Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Thr Ser
305 310 315 320

Ala Ala Arg Tyr Asp Tyr His Cys Gly Leu Ser Asp Ala Leu Thr Pro
325 330 335

Ala Pro Glu Ala Gly Gln Trp Phe Gln Ala Tyr Phe Glu Gln Leu Leu
340 345 350

Ile Asn Ala Asn Pro Pro Phe
355

<210> 24

<211> 360

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 24

Gly Asn Pro Phe Ser Gly Val Gln Leu Trp Ala Asn Thr Tyr Tyr Ser
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Ser Leu Ser Pro Glu Leu Ala
20 25 30

Ala Lys Ala Ala Lys Val Ala Glu Val Pro Ser Phe Gln Trp Leu Asp
35 40 45

Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Glu Ile
50 55 60

Arg Ala Ala Asn Gln Arg Gly Ala Asn Pro Gln Tyr Ala Ala Gln Ile
65 70 75 80

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ser Asn
85 90 95

Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala Tyr
100 105 110

Ile Asn Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg Thr
115 120 125

Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Thr Tyr Lys Glu Leu Thr
145 150 155 160

Val Tyr Ala Leu Lys Gln Leu Asn Leu Pro His Val Ala Met Tyr Met
165 170 175

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro
180 185 190

Ala Ala Glu Leu Phe Ala Gln Ile Tyr Arg Asp Ala Gly Arg Pro Ala
195 200 205

Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220

Val Ser Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu
225 230 235 240

Lys His Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly Phe
245 250 255

Pro Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr
260 265 270

Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe
275 280 285

Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe
290 295 300

Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser Ser
305 310 315 320

Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro
325 330 335

99

Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu
340 345 350

Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 25

<211> 360

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 25

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45

Asp Arg Asn Val Thr Val Asp Thr Leu Leu Val Gln Thr Leu Ser Glu
50 55 60

Ile Arg Glu Ala Asn Gln Ala Gly Ala Asn Pro Pro Tyr Ala Gly Ile
65 70 75 80

Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser
85 90 95

Asn Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg
100 105 110

Tyr Ile Asp Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg
115 120 125

Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn
130 135 140

Met Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Ala Tyr Leu Glu Cys
145 150 155 160

Ile Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr
165 170 175

Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp
180 185 190

Pro Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro
195 200 205

Arg Ala Leu Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp
210 215 220

Ser Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp
225 230 235 240

100

Glu Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly
245 250 255

Phe Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro
260 265 270

Thr Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly
275 280 285

Phe Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala
290 295 300

Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Thr
305 310 315 320

Ser Ala Ala Arg Tyr Asp Tyr His Cys Gly Leu Ser Asp Ala Leu Thr
325 330 335

Pro Ala Pro Glu Ala Gly Gln Trp Phe Gln Ala Tyr Phe Glu Gln Leu
340 345 350

Leu Ile Asn Ala Asn Pro Pro Phe
355 360

<210> 26

<211> 359

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 26

Gly Asn Pro Phe Val Gly Val Thr Pro Trp Ala Asn Ala Tyr Tyr Ala
1 5 10 15

Ser Glu Val Ser Ser Leu Ala Ile Pro Ser Leu Thr Gly Ala Met Ala
20 25 30

Thr Ala Ala Ala Val Ala Lys Val Pro Ser Phe Gln Trp Leu Asp
35 40 45

Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Glu Ile
50 55 60

Arg Ala Ala Asn Gln Arg Gly Ala Asn Pro Pro Tyr Ala Gly Ile Phe
65 70 75 80

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn
85 90 95

Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg Tyr
100 105 110

Ile Asp Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg Thr
115 120 125

Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Thr Tyr Arg Glu Leu Thr

101

145 150 155 160

Ile Tyr Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr Met
165 170 175

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro
180 185 190

Ala Ala Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro Arg
195 200 205

Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220

Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu
225 230 235 240

Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly Phe
245 250 255

Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr
260 265 270

Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe
275 280 285

Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe
290 295 300

Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Thr Ser
305 310 315 320

Ala Ala Arg Tyr Asp Tyr His Cys Gly Leu Ser Asp Ala Leu Thr Pro
325 330 335

Ala Pro Glu Ala Gly Gln Trp Phe Gln Ala Tyr Phe Glu Gln Leu Leu
340 345 350

Ile Asn Ala Asn Pro Pro Phe
355

<210> 27

<211> 358

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 27

Gly Asn Pro Phe Ser Gly Val Gln Leu Trp Ala Asn Thr Tyr Tyr Ser
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Ser Leu Ser Pro Glu Leu Ala
20 25 30

Ala Lys Ala Ala Lys Val Ala Glu Val Pro Ser Phe Gln Trp Leu Asp
35 40 45

Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Glu Ile
50 55 60

102

Arg Ala Ala Asn Gln Arg Gly Gly Asn Tyr Ala Gly Gln Phe Val Val
65 70 75 80

Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Leu Ala Ser Asn Gly Glu
85 90 95

Tyr Ser Ile Ala Asp Gly Gly Val Ala Lys Tyr Lys Asn Tyr Ile Asp
100 105 110

Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg Thr Ile Leu
115 120 125

Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met Asn Val
130 135 140

Pro Lys Cys Ser Gly Ala Ala Ser Thr Tyr Lys Glu Leu Thr Val Tyr
145 150 155 160

Ala Leu Lys Gln Leu Asn Leu Pro His Val Ala Met Tyr Met Asp Ala
165 170 175

Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro Ala Ala
180 185 190

Glu Leu Phe Ala Gln Ile Tyr Arg Asp Ala Gly Arg Pro Ala Ala Val
195 200 205

Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser Ile Ala
210 215 220

Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu Lys His
225 230 235 240

Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly Phe Asp Ala
245 250 255

Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr Gly Gln
260 265 270

Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe Gly Val
275 280 285

Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe Val Trp
290 295 300

Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser Ser Ala Pro
305 310 315 320

Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro Ala Pro
325 330 335

Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu Thr Asn
340 345 350

Ala Asn Pro Ser Phe Leu
355

<210> 28

<211> 361

<212> PRT

<213> Secuencia artificial

103

<220>

<223> Polipéptido quimérico CBH II

<400> 28

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45

Asp Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Glu
50 55 60

Ile Arg Ala Ala Asn Gln Arg Gly Ala Asn Pro Pro Tyr Ala Gly Ile
65 70 75 80

Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser
85 90 95

Asn Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg
100 105 110

Tyr Ile Asp Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg
115 120 125

Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn
130 135 140

Met Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Thr Tyr Lys Glu Leu
145 150 155 160

Thr Val Tyr Ala Leu Lys Gln Leu Asn Leu Pro His Val Ala Met Tyr
165 170 175

Met Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln
180 185 190

Pro Ala Ala Glu Leu Phe Ala Gln Ile Tyr Arg Asp Ala Gly Arg Pro
195 200 205

Ala Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Gly Trp
210 215 220

Asn Ile Thr Ser Pro Pro Ser Tyr Thr Gln Gly Asn Ala Val Tyr Asn
225 230 235 240

Glu Lys Leu Tyr Ile His Ala Ile Gly Pro Leu Leu Ala Asn His Gly
245 250 255

Trp Ser Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro
260 265 270

Thr Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly
275 280 285

Phe Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala
290 295 300

104

Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser
305 310 315 320

Ser Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln
325 330 335

Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu
340 345 350

Leu Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 29

<211> 357

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 29

Gly Asn Pro Phe Val Gly Val Thr Pro Trp Ala Asn Ala Tyr Tyr Ala
1 5 10 15

Ser Glu Val Ser Ser Leu Ala Ile Pro Ser Leu Thr Gly Ala Met Ala
20 25 30

Thr Ala Ala Ala Val Ala Lys Val Pro Ser Phe Met Trp Leu Asp
35 40 45

Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile Arg
50 55 60

Thr Ala Asn Lys Asn Gly Gly Asn Tyr Ala Gly Gln Phe Val Val Tyr
65 70 75 80

Asp Leu Pro Asp Arg Asp Cys Ala Ala Leu Ala Ser Asn Gly Glu Tyr
85 90 95

Ser Ile Ala Asp Gly Gly Val Ala Lys Tyr Lys Asn Tyr Ile Asp Arg
100 105 110

Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg Thr Ile Leu Val
115 120 125

Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met Asn Val Gln
130 135 140

Lys Cys Ser Asn Ala Ala Ser Ala Tyr Leu Glu Cys Ile Asn Tyr Ala
145 150 155 160

Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu Asp Ala Gly
165 170 175

His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp Pro Ala Ala Gln
180 185 190

Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro Arg Ala Leu Arg
195 200 205

105

Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser Val Ser Ser
210 215 220

Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu Lys His Tyr
225 230 235 240

Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly Phe Pro Ala Lys
245 250 255

Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr Gly Gln Leu
260 265 270

Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe Gly Val Arg
275 280 285

Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe Val Trp Val
290 295 300

Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser Ser Ala Pro Arg
305 310 315 320

Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro Ala Pro Gln
325 330 335

Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu Thr Asn Ala
340 345 350

Asn Pro Ser Phe Leu
355

<210> 30

<211> 361

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 30

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45

Asp Arg Asn Val Thr Val Asp Thr Leu Leu Val Gln Thr Leu Ser Glu
50 55 60

Ile Arg Glu Ala Asn Gln Ala Gly Ala Asn Pro Gln Tyr Ala Ala Gln
65 70 75 80

Ile Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser
85 90 95

Asn Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala
100 105 110

Tyr Ile Asn Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg

106

115 120 125

Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn
130 135 140

Met Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Thr Tyr Lys Glu Leu
145 150 155 160

Thr Val Tyr Ala Leu Lys Gln Leu Asn Leu Pro His Val Ala Met Tyr
165 170 175

Met Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln
180 185 190

Pro Ala Ala Glu Leu Phe Ala Gln Ile Tyr Arg Asp Ala Gly Arg Pro
195 200 205

Ala Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp
210 215 220

Ser Val Ser Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp
225 230 235 240

Glu Lys His Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly
245 250 255

Phe Pro Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro
260 265 270

Thr Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly
275 280 285

Phe Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala
290 295 300

Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser
305 310 315 320

Ser Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln
325 330 335

Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu
340 345 350

Leu Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 31

<211> 359

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 31

Gly Asn Pro Phe Val Gly Val Thr Pro Trp Ala Asn Ala Tyr Tyr Ala
1 5 10 15

Ser Glu Val Ser Ser Leu Ala Ile Pro Ser Leu Thr Gly Ala Met Ala
20 25 30

107

Thr Ala Ala Ala Val Ala Lys Val Pro Ser Phe Gln Trp Leu Asp
35 40 45

Arg Asn Val Thr Val Asp Thr Leu Leu Val Gln Thr Leu Ser Glu Ile
50 55 60

Arg Glu Ala Asn Gln Ala Gly Ala Asn Pro Pro Tyr Ala Gly Ile Phe
65 70 75 80

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ser Asn
85 90 95

Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg Tyr
100 105 110

Ile Asp Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg Thr
115 120 125

Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Thr Tyr Lys Glu Leu Thr
145 150 155 160

Val Tyr Ala Leu Lys Gln Leu Asn Leu Pro His Val Ala Met Tyr Met
165 170 175

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro
180 185 190

Ala Ala Glu Leu Phe Ala Gln Ile Tyr Arg Asp Ala Gly Arg Pro Ala
195 200 205

Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220

Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu
225 230 235 240

Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly Phe
245 250 255

Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr
260 265 270

Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe
275 280 285

Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe
290 295 300

Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Thr Thr
305 310 315 320

Ala Ala Arg Tyr Asp Tyr His Cys Gly Leu Glu Asp Ala Leu Lys Pro
325 330 335

Ala Pro Glu Ala Gly Gln Trp Phe Asn Glu Tyr Phe Ile Gln Leu Leu
340 345 350

Arg Asn Ala Asn Pro Pro Phe

108

355

<210> 32

<211> 359

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 32

Gly Asn Pro Phe Val Gly Val Thr Pro Trp Ala Asn Ala Tyr Tyr Ala
1 5 10 15

Ser Glu Val Ser Ser Leu Ala Ile Pro Ser Leu Thr Gly Ala Met Ala
20 25 30

Thr Ala Ala Ala Ala Val Ala Lys Val Pro Ser Phe Gln Trp Leu Asp
35 40 45

Arg Asn Val Thr Val Asp Thr Leu Leu Val Gln Thr Leu Ser Glu Ile
50 55 60

Arg Glu Ala Asn Gln Ala Gly Ala Asn Pro Pro Tyr Ala Gly Ile Phe
65 70 75 80

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn
85 90 95

Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg Tyr
100 105 110

Ile Asp Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg Thr
115 120 125

Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Thr Tyr Arg Glu Leu Thr
145 150 155 160

Ile Tyr Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr Met
165 170 175

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro
180 185 190

Ala Ala Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro Arg
195 200 205

Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220

Val Ser Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu
225 230 235 240

Lys His Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly Phe
245 250 255

Pro Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr
260 265 270

109

Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe
275 280 285

Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe
290 295 300

Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Thr Thr
305 310 315 320

Ala Ala Arg Tyr Asp Tyr His Cys Gly Leu Glu Asp Ala Leu Lys Pro
325 330 335

Ala Pro Glu Ala Gly Gln Trp Phe Asn Glu Tyr Phe Ile Gln Leu Leu
340 345 350

Arg Asn Ala Asn Pro Pro Phe
355

<210> 33

<211> 359

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 33

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Met Trp Leu
35 40 45

Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile
50 55 60

Arg Thr Ala Asn Lys Asn Gly Ala Asn Pro Pro Tyr Ala Gly Ile Phe
65 70 75 80

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn
85 90 95

Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg Tyr
100 105 110

Ile Asp Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg Thr
115 120 125

Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Ala Tyr Leu Glu Cys Ile
145 150 155 160

Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu
165 170 175

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp Pro
180 185 190

Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro Arg
195 200 205

Ala Leu Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220

Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu
225 230 235 240

Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly Phe
245 250 255

Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr
260 265 270

Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe
275 280 285

Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe
290 295 300

Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Thr Thr
305 310 315 320

Ala Ala Arg Tyr Asp Tyr His Cys Gly Leu Glu Asp Ala Leu Lys Pro
325 330 335

Ala Pro Glu Ala Gly Gln Trp Phe Asn Glu Tyr Phe Ile Gln Leu Leu
340 345 350

Arg Asn Ala Asn Pro Pro Phe
355

<210> 34

<211> 361

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 34

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45

Asp Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Glu
50 55 60

Ile Arg Ala Ala Asn Gln Arg Gly Ala Asn Pro Gln Tyr Ala Ala Gln
65 70 75 80

Ile Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ser

85 90 95

Asn Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala
100 105 110

Tyr Ile Asn Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg
115 120 125

Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn
130 135 140

Met Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Ala Tyr Leu Glu Cys
145 150 155 160

Ile Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr
165 170 175

Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp
180 185 190

Pro Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro
195 200 205

Arg Ala Leu Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp
210 215 220

Ser Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp
225 230 235 240

Glu Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly
245 250 255

Phe Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro
260 265 270

Thr Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly
275 280 285

Phe Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala
290 295 300

Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser
305 310 315 320

Ser Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln
325 330 335

Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu
340 345 350

Leu Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 35

<211> 358

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 35

Gly Asn Pro Phe Val Gly Val Thr Pro Trp Ala Asn Ala Tyr Tyr Ala
1 5 10 15

Ser Glu Val Ser Ser Leu Ala Ile Pro Ser Leu Thr Gly Ala Met Ala
20 25 30

Thr Ala Ala Ala Ala Val Ala Lys Val Pro Ser Phe Met Trp Leu Asp
35 40 45

Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile Arg
50 55 60

Thr Ala Asn Lys Asn Gly Ala Asn Pro Pro Tyr Ala Gly Ile Phe Val
65 70 75 80

Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn Gly
85 90 95

Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg Tyr Ile
100 105 110

Asp Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg Thr Ile
115 120 125

Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met Asn
130 135 140

Val Pro Lys Cys Ser Gly Ala Ala Ser Thr Tyr Arg Glu Leu Thr Ile
145 150 155 160

Tyr Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr Met Asp
165 170 175

Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro Ala
180 185 190

Ala Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro Arg Ala
195 200 205

Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser Ile
210 215 220

Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu Lys
225 230 235 240

His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly Phe Asp
245 250 255

Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr Gly
260 265 270

Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe Gly
275 280 285

Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe Val
290 295 300

Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Thr Thr Ala
305 310 315 320

Ala Arg Tyr Asp Tyr His Cys Gly Leu Glu Asp Ala Leu Lys Pro Ala

325 330 335

Pro Glu Ala Gly Gln Trp Phe Asn Glu Tyr Phe Ile Gln Leu Leu Arg
340 345 350

Asn Ala Asn Pro Pro Phe
355

<210> 36

<211> 360

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 36

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Met Trp Leu
35 40 45

Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile
50 55 60

Arg Thr Ala Asn Lys Asn Gly Ala Asn Pro Gln Tyr Ala Ala Gln Ile
65 70 75 80

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn
85 90 95

Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala Tyr
100 105 110

Ile Asn Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg Thr
115 120 125

Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Thr Tyr Arg Glu Leu Thr
145 150 155 160

Ile Tyr Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr Met
165 170 175

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro
180 185 190

Ala Ala Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro Arg
195 200 205

Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220

Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu
225 230 235 240

114

Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly Phe
245 250 255

Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr
260 265 270

Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe
275 280 285

Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe
290 295 300

Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser Ser
305 310 315 320

Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro
325 330 335

Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu
340 345 350

Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 37

<211> 360

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 37

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Met Trp Leu
35 40 45

Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile
50 55 60

Arg Thr Ala Asn Lys Asn Gly Ala Asn Pro Gln Tyr Ala Ala Gln Ile
65 70 75 80

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn
85 90 95

Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala Tyr
100 105 110

Ile Asn Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg Thr
115 120 125

Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

115

Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Ala Tyr Leu Glu Cys Ile
145 150 155 160

Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu
165 170 175

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp Pro
180 185 190

Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro Arg
195 200 205

Ala Leu Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220

Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu
225 230 235 240

Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly Phe
245 250 255

Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr
260 265 270

Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe
275 280 285

Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe
290 295 300

Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser Ser
305 310 315 320

Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro
325 330 335

Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu
340 345 350

Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 38
<211> 359
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido quimérico CBH II

<400> 38

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Met Trp Leu
35 40 45

Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile

50 55 60

Arg Thr Ala Asn Lys Asn Gly Ala Asn Pro Gln Tyr Ala Ala Gln Ile
65 70 75 80

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn
85 90 95

Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala Tyr
100 105 110

Ile Asn Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg Thr
115 120 125

Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Thr Tyr Arg Glu Leu Thr
145 150 155 160

Ile Tyr Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr Met
165 170 175

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro
180 185 190

Ala Ala Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro Arg
195 200 205

Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220

Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu
225 230 235 240

Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly Phe
245 250 255

Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr
260 265 270

Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe
275 280 285

Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe
290 295 300

Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Thr Thr
305 310 315 320

Ala Ala Arg Tyr Asp Tyr His Cys Gly Leu Glu Asp Ala Leu Lys Pro
325 330 335

Ala Pro Glu Ala Gly Gln Trp Phe Asn Glu Tyr Phe Ile Gln Leu Leu
340 345 350

Arg Asn Ala Asn Pro Pro Phe
355

<210> 39

<211> 360

<212> PRT

112

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 39

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Met Trp Leu
35 40 45

Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile
50 55 60

Arg Thr Ala Asn Lys Asn Gly Ala Asn Pro Gln Tyr Ala Ala Gln Ile
65 70 75 80

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn
85 90 95

Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala Tyr
100 105 110

Ile Asn Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg Thr
115 120 125

Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Thr Tyr Arg Glu Leu Thr
145 150 155 160

Ile Tyr Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr Met
165 170 175

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro
180 185 190

Ala Ala Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro Arg
195 200 205

Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220

Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu
225 230 235 240

Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly Phe
245 250 255

Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr
260 265 270

Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe
275 280 285

Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe

118

290 295 300

Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser Ser
305 310 315 320

Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro
325 330 335

Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu
340 345 350

Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 40

<211> 360

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 40

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Met Trp Leu
35 40 45

Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile
50 55 60

Arg Thr Ala Asn Lys Asn Gly Ala Asn Pro Pro Tyr Ala Gly Ile Phe
65 70 75 80

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn
85 90 95

Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg Tyr
100 105 110

Ile Asp Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg Thr
115 120 125

Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Ala Tyr Leu Glu Cys Ile
145 150 155 160

Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu
165 170 175

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp Pro
180 185 190

Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro Arg
195 200 205

Ala Leu Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220

Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu
225 230 235 240

Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly Phe
245 250 255

Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr
260 265 270

Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe
275 280 285

Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe
290 295 300

Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser Ser
305 310 315 320

Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro
325 330 335

Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu
340 345 350

Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 41

<211> 359

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 41

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Met Trp Leu
35 40 45

Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile
50 55 60

Arg Thr Ala Asn Lys Asn Gly Ala Asn Pro Gln Tyr Ala Ala Gln Ile
65 70 75 80

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ser Asn
85 90 95

Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala Tyr
100 105 110

120

Ile Asn Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg Thr
115 120 125

Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Thr Tyr Arg Glu Leu Thr
145 150 155 160

Ile Tyr Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr Met
165 170 175

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro
180 185 190

Ala Ala Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro Arg
195 200 205

Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220

Val Ser Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu
225 230 235 240

Lys His Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly Phe
245 250 255

Pro Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr
260 265 270

Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe
275 280 285

Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe
290 295 300

Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Thr Thr
305 310 315 320

Ala Ala Arg Tyr Asp Tyr His Cys Gly Leu Glu Asp Ala Leu Lys Pro
325 330 335

Ala Pro Glu Ala Gly Gln Trp Phe Asn Glu Tyr Phe Ile Gln Leu Leu
340 345 350

Arg Asn Ala Asn Pro Pro Phe
355

<210> 42

<211> 360

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 42

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu

20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Met Trp Leu
35 40 45

Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile
50 55 60

Arg Thr Ala Asn Lys Asn Gly Ala Asn Pro Pro Tyr Ala Gly Ile Phe
65 70 75 80

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn
85 90 95

Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg Tyr
100 105 110

Ile Asp Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg Thr
115 120 125

Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Thr Tyr Arg Glu Leu Thr
145 150 155 160

Ile Tyr Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr Met
165 170 175

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro
180 185 190

Ala Ala Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro Arg
195 200 205

Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220

Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu
225 230 235 240

Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly Phe
245 250 255

Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr
260 265 270

Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe
275 280 285

Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe
290 295 300

Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser Ser
305 310 315 320

Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro
325 330 335

Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu
340 345 350

122

Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 43

<211> 360

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 43

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45

Asp Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Glu
50 55 60

Ile Arg Ala Ala Asn Gln Arg Gly Ala Asn Pro Pro Tyr Ala Gly Ile
65 70 75 80

Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser
85 90 95

Asn Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg
100 105 110

Tyr Ile Asp Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg
115 120 125

Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn
130 135 140

Met Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Ala Tyr Leu Glu Cys
145 150 155 160

Ile Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr
165 170 175

Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp
180 185 190

Pro Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro
195 200 205

Arg Ala Leu Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp
210 215 220

Ser Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp
225 230 235 240

Glu Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly
245 250 255

Phe Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro

123

260 265 270

Thr Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly
275 280 285

Phe Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala
290 295 300

Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Thr
305 310 315 320

Thr Ala Ala Arg Tyr Asp Tyr His Cys Gly Leu Glu Asp Ala Leu Lys
325 330 335

Pro Ala Pro Glu Ala Gly Gln Trp Phe Asn Glu Tyr Phe Ile Gln Leu
340 345 350

Leu Arg Asn Ala Asn Pro Pro Phe
355 360

<210> 44

<211> 359

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 44

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Met Trp Leu
35 40 45

Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile
50 55 60

Arg Thr Ala Asn Lys Asn Gly Ala Asn Pro Gln Tyr Ala Ala Gln Ile
65 70 75 80

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn
85 90 95

Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala Tyr
100 105 110

Ile Asn Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg Thr
115 120 125

Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Ala Tyr Leu Glu Cys Ile
145 150 155 160

Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu
165 170 175

124

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp Pro
180 185 190

Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro Arg
195 200 205

Ala Leu Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220

Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu
225 230 235 240

Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly Phe
245 250 255

Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr
260 265 270

Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe
275 280 285

Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe
290 295 300

Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Thr Thr
305 310 315 320

Ala Ala Arg Tyr Asp Tyr His Cys Gly Leu Glu Asp Ala Leu Lys Pro
325 330 335

Ala Pro Glu Ala Gly Gln Trp Phe Asn Glu Tyr Phe Ile Gln Leu Leu
340 345 350

Arg Asn Ala Asn Pro Pro Phe
355

<210> 45

<211> 360

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 45

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Met Trp Leu
35 40 45

Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile
50 55 60

Arg Thr Ala Asn Lys Asn Gly Ala Asn Pro Gln Tyr Ala Ala Gln Ile
65 70 75 80

125

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ser Asn
85 90 95

Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala Tyr
100 105 110

Ile Asn Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg Thr
115 120 125

Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Ala Tyr Leu Glu Cys Ile
145 150 155 160

Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu
165 170 175

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp Pro
180 185 190

Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro Arg
195 200 205

Ala Leu Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220

Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu
225 230 235 240

Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly Phe
245 250 255

Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr
260 265 270

Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe
275 280 285

Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe
290 295 300

Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser Ser
305 310 315 320

Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro
325 330 335

Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu
340 345 350

Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 46

<211> 361

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

126

<400> 46

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45

Asp Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Glu
50 55 60

Ile Arg Ala Ala Asn Gln Arg Gly Ala Asn Pro Pro Tyr Ala Gly Ile
65 70 75 80

Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser
85 90 95

Asn Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg
100 105 110

Tyr Ile Asp Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg
115 120 125

Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn
130 135 140

Met Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Ala Tyr Leu Glu Cys
145 150 155 160

Ile Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr
165 170 175

Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp
180 185 190

Pro Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro
195 200 205

Arg Ala Leu Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp
210 215 220

Ser Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp
225 230 235 240

Glu Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly
245 250 255

Phe Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro
260 265 270

Thr Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly
275 280 285

Phe Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala
290 295 300

Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser
305 310 315 320

Ser Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln
325 330 335

Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu
340 345 350

Leu Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 47

<211> 360

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 47

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Met Trp Leu
35 40 45

Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile
50 55 60

Arg Thr Ala Asn Lys Asn Gly Ala Asn Pro Gln Tyr Ala Ala Gln Ile
65 70 75 80

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn
85 90 95

Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala Tyr
100 105 110

Ile Asn Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg Thr
115 120 125

Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Ala Tyr Leu Glu Cys Ile
145 150 155 160

Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu
165 170 175

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp Pro
180 185 190

Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro Arg
195 200 205

Ala Leu Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220

Val Ser Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu

128

225 230 235 240

Lys His Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly Phe
245 250 255

Pro Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr
260 265 270

Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe
275 280 285

Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe
290 295 300

Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser Ser
305 310 315 320

Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro
325 330 335

Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu
340 345 350

Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 48

<211> 361

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 48

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45

Asp Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Glu
50 55 60

Ile Arg Ala Ala Asn Gln Arg Gly Ala Asn Pro Pro Tyr Ala Gly Ile
65 70 75 80

Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser
85 90 95

Asn Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg
100 105 110

Tyr Ile Asp Thr Ile Arg Gln Ile Val Val Glu Tyr Ser Asp Ile Arg
115 120 125

Thr Leu Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Leu Val Thr Asn
130 135 140

129

Leu Gly Thr Pro Lys Cys Ala Asn Ala Gln Ser Ala Tyr Leu Glu Cys
145 150 155 160

Ile Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr
165 170 175

Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp
180 185 190

Pro Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro
195 200 205

Arg Ala Leu Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp
210 215 220

Ser Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp
225 230 235 240

Glu Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly
245 250 255

Phe Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro
260 265 270

Thr Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly
275 280 285

Phe Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala
290 295 300

Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser
305 310 315 320

Ser Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln
325 330 335

Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu
340 345 350

Leu Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 49
<211> 361
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido quimérico CBH II

<400> 49

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45

130

Asp Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Glu
50 55 60

Ile Arg Ala Ala Asn Gln Arg Gly Ala Asn Pro Gln Tyr Ala Ala Gln
65 70 75 80

Ile Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser
85 90 95

Asn Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala
100 105 110

Tyr Ile Asn Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg
115 120 125

Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn
130 135 140

Met Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Thr Tyr Arg Glu Leu
145 150 155 160

Thr Ile Tyr Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr
165 170 175

Met Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln
180 185 190

Pro Ala Ala Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro
195 200 205

Arg Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp
210 215 220

Ser Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp
225 230 235 240

Glu Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly
245 250 255

Phe Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro
260 265 270

Thr Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly
275 280 285

Phe Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala
290 295 300

Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser
305 310 315 320

Ser Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln
325 330 335

Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu
340 345 350

Leu Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 50
<211> 360

131

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 50

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Met Trp Leu
35 40 45

Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile
50 55 60

Arg Thr Ala Asn Lys Asn Gly Ala Asn Pro Pro Tyr Ala Gly Ile Phe
65 70 75 80

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn
85 90 95

Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg Tyr
100 105 110

Ile Asp Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg Thr
115 120 125

Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Thr Tyr Arg Glu Leu Thr
145 150 155 160

Ile Tyr Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr Met
165 170 175

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro
180 185 190

Ala Ala Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro Arg
195 200 205

Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220

Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu
225 230 235 240

Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly Phe
245 250 255

Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr
260 265 270

Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe
275 280 285

Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe
290 295 300

Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser Ser
305 310 315 320

Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro
325 330 335

Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu
340 345 350

Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 51

<211> 361

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 51

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45

Asp Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Glu
50 55 60

Ile Arg Ala Ala Asn Gln Arg Gly Ala Asn Pro Pro Tyr Ala Gly Ile
65 70 75 80

Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser
85 90 95

Asn Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg
100 105 110

Tyr Ile Asp Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg
115 120 125

Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn
130 135 140

Met Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Ala Tyr Leu Glu Cys
145 150 155 160

Ile Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr
165 170 175

Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp
180 185 190

Pro Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro

133

195 200 205

Arg Ala Leu Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp
210 215 220

Ser Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp
225 230 235 240

Glu Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly
245 250 255

Phe Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro
260 265 270

Thr Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly
275 280 285

Phe Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala
290 295 300

Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser
305 310 315 320

Ser Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln
325 330 335

Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu
340 345 350

Leu Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 52

<211> 361

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 52

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45

Asp Arg Asn Val Thr Val Asp Thr Leu Leu Val Gln Thr Leu Ser Glu
50 55 60

Ile Arg Glu Ala Asn Gln Ala Gly Ala Asn Pro Gln Tyr Ala Ala Gln
65 70 75 80

Ile Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser
85 90 95

Asn Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala
100 105 110

134

Tyr Ile Asn Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg
115 120 125

Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn
130 135 140

Met Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Thr Tyr Lys Glu Leu
145 150 155 160

Thr Val Tyr Ala Leu Lys Gln Leu Asn Leu Pro His Val Ala Met Tyr
165 170 175

Met Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln
180 185 190

Pro Ala Ala Glu Leu Phe Ala Gln Ile Tyr Arg Asp Ala Gly Arg Pro
195 200 205

Ala Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp
210 215 220

Ser Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp
225 230 235 240

Glu Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly
245 250 255

Phe Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro
260 265 270

Thr Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly
275 280 285

Phe Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala
290 295 300

Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser
305 310 315 320

Ser Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln
325 330 335

Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu
340 345 350

Leu Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 53

<211> 361

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 53

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

135

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45

Asp Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Glu
50 55 60

Ile Arg Ala Ala Asn Gln Arg Gly Ala Asn Pro Gln Tyr Ala Ala Gln
65 70 75 80

Ile Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser
85 90 95

Asn Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala
100 105 110

Tyr Ile Asn Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg
115 120 125

Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn
130 135 140

Met Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Thr Tyr Lys Glu Leu
145 150 155 160

Thr Val Tyr Ala Leu Lys Gln Leu Asn Leu Pro His Val Ala Met Tyr
165 170 175

Met Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln
180 185 190

Pro Ala Ala Glu Leu Phe Ala Gln Ile Tyr Arg Asp Ala Gly Arg Pro
195 200 205

Ala Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp
210 215 220

Ser Val Ser Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp
225 230 235 240

Glu Lys His Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly
245 250 255

Phe Pro Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro
260 265 270

Thr Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly
275 280 285

Phe Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala
290 295 300

Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser
305 310 315 320

Ser Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln
325 330 335

Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu
340 345 350

136

Leu Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 54
<211> 361
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido quimérico CBH II

<400> 54

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45

Asp Arg Asn Val Thr Val Asp Thr Leu Leu Val Gln Thr Leu Ser Glu
50 55 60

Ile Arg Glu Ala Asn Gln Ala Gly Ala Asn Pro Gln Tyr Ala Ala Gln
65 70 75 80

Ile Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser
85 90 95

Asn Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala
100 105 110

Tyr Ile Asn Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg
115 120 125

Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn
130 135 140

Met Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Ala Tyr Leu Glu Cys
145 150 155 160

Ile Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr
165 170 175

Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp
180 185 190

Pro Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro
195 200 205

Arg Ala Leu Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp
210 215 220

Ser Val Ser Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp
225 230 235 240

Glu Lys His Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly
245 250 255

132

Phe Pro Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro
260 265 270

Thr Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly
275 280 285

Phe Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala
290 295 300

Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser
305 310 315 320

Ser Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln
325 330 335

Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu
340 345 350

Leu Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 55

<211> 360

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 55

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Met Trp Leu
35 40 45

Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile
50 55 60

Arg Thr Ala Asn Lys Asn Gly Ala Asn Pro Gln Tyr Ala Ala Gln Ile
65 70 75 80

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn
85 90 95

Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala Tyr
100 105 110

Ile Asn Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg Thr
115 120 125

Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
130 135 140

Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Thr Tyr Lys Glu Leu Thr
145 150 155 160

Val Tyr Ala Leu Lys Gln Leu Asn Leu Pro His Val Ala Met Tyr Met

38

165 170 175
Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro
180 185 190
Ala Ala Glu Leu Phe Ala Gln Ile Tyr Arg Asp Ala Gly Arg Pro Ala
195 200 205
Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser
210 215 220
Val Ser Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu
225 230 235 240
Lys His Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly Phe
245 250 255
Pro Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr
260 265 270
Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe
275 280 285
Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe
290 295 300
Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser Ser
305 310 315 320
Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro
325 330 335
Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu
340 345 350
Thr Asn Ala Asn Pro Ser Phe Leu
355 360
<210> 56
<211> 361
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido quimérico CBH II

<400> 56
Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15
Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30
Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45
Asp Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Glu
50 55 60
Ile Arg Ala Ala Asn Gln Arg Gly Ala Asn Pro Gln Tyr Ala Ala Gln
65 70 75 80

139

Ile Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser
85 90 95

Asn Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala
100 105 110

Tyr Ile Asn Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg
115 120 125

Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn
130 135 140

Met Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Ala Tyr Leu Glu Cys
145 150 155 160

Ile Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr
165 170 175

Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp
180 185 190

Pro Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro
195 200 205

Arg Ala Leu Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp
210 215 220

Ser Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp
225 230 235 240

Glu Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly
245 250 255

Phe Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro
260 265 270

Thr Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly
275 280 285

Phe Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala
290 295 300

Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser
305 310 315 320

Ser Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln
325 330 335

Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu
340 345 350

Leu Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 57

<211> 361

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

140

<400> 57

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45

Asp Arg Asn Val Thr Val Asp Thr Leu Leu Val Gln Thr Leu Ser Glu
50 55 60

Ile Arg Glu Ala Asn Gln Ala Gly Ala Asn Pro Gln Tyr Ala Ala Gln
65 70 75 80

Ile Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser
85 90 95

Asn Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala
100 105 110

Tyr Ile Asn Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg
115 120 125

Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn
130 135 140

Met Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Thr Tyr Arg Glu Leu
145 150 155 160

Thr Ile Tyr Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr
165 170 175

Met Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln
180 185 190

Pro Ala Ala Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro
195 200 205

Arg Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp
210 215 220

Ser Val Ser Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp
225 230 235 240

Glu Lys His Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly
245 250 255

Phe Pro Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro
260 265 270

Thr Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly
275 280 285

Phe Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala
290 295 300

Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser
305 310 315 320

1241

Ser Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln
325 330 335

Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu
340 345 350

Leu Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 58

<211> 361

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 58

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45

Asp Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Glu
50 55 60

Ile Arg Ala Ala Asn Gln Arg Gly Ala Asn Pro Pro Tyr Ala Gly Ile
65 70 75 80

Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser
85 90 95

Asn Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg
100 105 110

Tyr Ile Asp Arg Ile Arg Glu Leu Leu Ile Gln Tyr Ser Asp Ile Arg
115 120 125

Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn
130 135 140

Met Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Thr Tyr Arg Glu Leu
145 150 155 160

Thr Ile Tyr Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr
165 170 175

Met Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln
180 185 190

Pro Ala Ala Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro
195 200 205

Arg Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp
210 215 220

142

Ser Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp
225 230 235 240

Glu Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly
245 250 255

Phe Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro
260 265 270

Thr Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly
275 280 285

Phe Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala
290 295 300

Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser
305 310 315 320

Ser Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln
325 330 335

Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu
340 345 350

Leu Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 59
<211> 361
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido quimérico CBH II

<400> 59

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45

Asp Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Glu
50 55 60

Ile Arg Ala Ala Asn Gln Arg Gly Ala Asn Pro Gln Tyr Ala Ala Gln
65 70 75 80

Ile Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ser
85 90 95

Asn Gly Glu Trp Ala Ile Ala Asn Asn Gly Val Asn Asn Tyr Lys Ala
100 105 110

Tyr Ile Asn Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg
115 120 125

Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn

143

130 135 140

Met Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Thr Tyr Arg Glu Leu
145 150 155 160

Thr Ile Tyr Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr
165 170 175

Met Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln
180 185 190

Pro Ala Ala Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro
195 200 205

Arg Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp
210 215 220

Ser Val Ser Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp
225 230 235 240

Glu Lys His Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly
245 250 255

Phe Pro Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro
260 265 270

Thr Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly
275 280 285

Phe Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala
290 295 300

Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser
305 310 315 320

Ser Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln
325 330 335

Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu
340 345 350

Leu Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 60

<211> 358

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 60

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Met Trp Leu
35 40 45

144

Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile
50 55 60

Arg Thr Ala Asn Lys Asn Gly Gly Asn Tyr Ala Gly Gln Phe Val Val
65 70 75 80

Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Leu Ala Ser Asn Gly Glu
85 90 95

Tyr Ser Ile Ala Asp Gly Gly Val Ala Lys Tyr Lys Asn Tyr Ile Asp
100 105 110

Thr Ile Arg Gln Ile Val Val Glu Tyr Ser Asp Ile Arg Thr Leu Leu
115 120 125

Val Ile Glu Pro Asp Ser Leu Ala Asn Leu Val Thr Asn Leu Gly Thr
130 135 140

Pro Lys Cys Ala Asn Ala Gln Ser Ala Tyr Leu Glu Cys Ile Asn Tyr
145 150 155 160

Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu Asp Ala
165 170 175

Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp Pro Ala Ala
180 185 190

Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro Arg Ala Leu
195 200 205

Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser Val Ser
210 215 220

Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu Lys His
225 230 235 240

Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly Phe Pro Ala
245 250 255

Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr Gly Gln
260 265 270

Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe Gly Val
275 280 285

Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe Val Trp
290 295 300

Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser Ser Ala Pro
305 310 315 320

Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro Ala Pro
325 330 335

Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu Thr Asn
340 345 350

Ala Asn Pro Ser Phe Leu
355

<210> 61

145

<211> 358
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido quimérico CBH II

<400> 61

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Met Trp Leu
35 40 45

Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile
50 55 60

Arg Thr Ala Asn Lys Asn Gly Gly Asn Tyr Ala Gly Gln Phe Val Val
65 70 75 80

Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Leu Ala Ser Asn Gly Glu
85 90 95

Tyr Ser Ile Ala Asp Gly Gly Val Ala Lys Tyr Lys Asn Tyr Ile Asp
100 105 110

Thr Ile Arg Gln Ile Val Val Glu Tyr Ser Asp Ile Arg Thr Leu Leu
115 120 125

Val Ile Glu Pro Asp Ser Leu Ala Asn Leu Val Thr Asn Leu Gly Thr
130 135 140

Pro Lys Cys Ala Asn Ala Gln Ser Ala Tyr Leu Glu Cys Ile Asn Tyr
145 150 155 160

Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu Asp Ala
165 170 175

Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp Pro Ala Ala
180 185 190

Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro Arg Ala Leu
195 200 205

Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser Ile Ala
210 215 220

Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu Lys His
225 230 235 240

Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly Phe Asp Ala
245 250 255

Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr Gly Gln
260 265 270

Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly Phe Gly Val
275 280 285

146

Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala Phe Val Trp
290 295 300

Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser Ser Ala Pro
305 310 315 320

Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro Ala Pro
325 330 335

Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu Thr Asn
340 345 350

Ala Asn Pro Ser Phe Leu
355

<210> 62

<211> 361

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 62

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45

Asp Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Glu
50 55 60

Ile Arg Ala Ala Asn Gln Arg Gly Ala Asn Pro Pro Tyr Ala Gly Ile
65 70 75 80

Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser
85 90 95

Asn Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg
100 105 110

Tyr Ile Asp Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg
115 120 125

Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn
130 135 140

Met Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Thr Tyr Arg Glu Leu
145 150 155 160

Thr Ile Tyr Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr
165 170 175

Met Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln
180 185 190

147

Pro Ala Ala Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro
195 200 205

Arg Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp
210 215 220

Ser Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp
225 230 235 240

Glu Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly
245 250 255

Phe Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro
260 265 270

Thr Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly
275 280 285

Phe Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala
290 295 300

Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Ser
305 310 315 320

Ser Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln
325 330 335

Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu
340 345 350

Leu Thr Asn Ala Asn Pro Ser Phe Leu
355 360

<210> 63

<211> 360

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido quimérico CBH II

<400> 63

Gly Asn Pro Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg
1 5 10 15

Ser Glu Val His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu
20 25 30

Arg Ala Ala Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu
35 40 45

Asp Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Glu
50 55 60

Ile Arg Ala Ala Asn Gln Arg Gly Ala Asn Pro Pro Tyr Ala Gly Ile
65 70 75 80

Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser
85 90 95

Asn Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Arg

148

100 105 110
Tyr Ile Asp Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg
115 120 125
Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn
130 135 140
Met Asn Val Pro Lys Cys Ser Gly Ala Ala Ser Thr Tyr Arg Glu Leu
145 150 155 160
Thr Ile Tyr Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr
165 170 175
Met Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln
180 185 190
Pro Ala Ala Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro
195 200 205
Arg Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp
210 215 220
Ser Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp
225 230 235 240
Glu Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg Asn Gln Gly
245 250 255
Phe Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly Lys Gln Pro
260 265 270
Thr Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val Lys Gly Thr Gly
275 280 285
Phe Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu Leu Val Asp Ala
290 295 300
Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Thr
305 310 315 320
Thr Ala Ala Arg Tyr Asp Tyr His Cys Gly Leu Glu Asp Ala Leu Lys
325 330 335
Pro Ala Pro Glu Ala Gly Gln Trp Phe Asn Glu Tyr Phe Ile Gln Leu
340 345 350
Leu Arg Asn Ala Asn Pro Pro Phe
355 360
<210> 64
<211> 447
<212> PRT
<213> Trichoderma reesei

<400> 64
Gln Ala Cys Ser Ser Val Trp Gly Gln Cys Gly Gly Gln Asn Trp Ser
1 5 10 15
Gly Pro Thr Cys Cys Ala Ser Gly Ser Thr Cys Val Tyr Ser Asn Asp
20 25 30

149

Tyr Tyr Ser Gln Cys Leu Pro Gly Ala Ala Ser Ser Ser Ser Ser Thr
35 40 45

Arg Ala Ala Ser Thr Thr Ser Arg Val Ser Pro Thr Thr Ser Arg Ser
50 55 60

Ser Ser Ala Thr Pro Pro Pro Gly Ser Thr Thr Thr Arg Val Pro Pro
65 70 75 80

Val Gly Ser Gly Thr Ala Thr Tyr Ser Gly Asn Pro Phe Val Gly Val
85 90 95

Thr Pro Trp Ala Asn Ala Tyr Tyr Ala Ser Glu Val Ser Ser Leu Ala
100 105 110

Ile Pro Ser Leu Thr Gly Ala Met Ala Thr Ala Ala Ala Val Ala
115 120 125

Lys Val Pro Ser Phe Met Trp Leu Asp Thr Leu Asp Lys Thr Pro Leu
130 135 140

Met Glu Gln Thr Leu Ala Asp Ile Arg Thr Ala Asn Lys Asn Gly Gly
145 150 155 160

Asn Tyr Ala Gly Gln Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys
165 170 175

Ala Ala Leu Ala Ser Asn Gly Glu Tyr Ser Ile Ala Asp Gly Gly Val
180 185 190

Ala Lys Tyr Lys Asn Tyr Ile Asp Thr Ile Arg Gln Ile Val Val Glu
195 200 205

Tyr Ser Asp Ile Arg Thr Leu Leu Val Ile Glu Pro Asp Ser Leu Ala
210 215 220

Asn Leu Val Thr Asn Leu Gly Thr Pro Lys Cys Ala Asn Ala Gln Ser
225 230 235 240

Ala Tyr Leu Glu Cys Ile Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro
245 250 255

Asn Val Ala Met Tyr Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp
260 265 270

Pro Ala Asn Gln Asp Pro Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys
275 280 285

Asn Ala Ser Ser Pro Arg Ala Leu Arg Gly Leu Ala Thr Asn Val Ala
290 295 300

Asn Tyr Asn Gly Trp Asn Ile Thr Ser Pro Pro Ser Tyr Thr Gln Gly
305 310 315 320

Asn Ala Val Tyr Asn Glu Lys Leu Tyr Ile His Ala Ile Gly Pro Leu
325 330 335

Leu Ala Asn His Gly Trp Ser Asn Ala Phe Phe Ile Thr Asp Gln Gly
340 345 350

Arg Ser Gly Lys Gln Pro Thr Gly Gln Gln Gln Trp Gly Asp Trp Cys

150

355 360 365

Asn Val Ile Gly Thr Gly Phe Gly Ile Arg Pro Ser Ala Asn Thr Gly
370 375 380

Asp Ser Leu Leu Asp Ser Phe Val Trp Val Lys Pro Gly Gly Glu Cys
385 390 395 400

Asp Gly Thr Ser Asp Ser Ser Ala Pro Arg Phe Asp Ser His Cys Ala
405 410 415

Leu Pro Asp Ala Leu Gln Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln
420 425 430

Ala Tyr Phe Val Gln Leu Leu Thr Asn Ala Asn Pro Ser Phe Leu
435 440 445

<210> 65
<211> 453
<212> PRT
<213> Humicola insolens

<400> 65

Ala Ser Cys Ala Pro Thr Trp Gly Gln Cys Gly Gly Ile Gly Phe Asn
1 5 10 15

Gly Pro Thr Cys Cys Gln Ser Gly Ser Thr Cys Val Lys Gln Asn Asp
20 25 30

Trp Tyr Ser Gln Cys Leu Pro Gly Ser Gln Val Thr Thr Thr Ser Thr
35 40 45

Thr Ser Thr Ser Ser Ser Thr Thr Ser Arg Ala Thr Ser Thr Thr
50 55 60

Ser Thr Gly Gly Val Thr Ser Ile Thr Thr Ala Pro Thr Arg Thr Val
65 70 75 80

Thr Ile Pro Gly Gly Ala Thr Thr Thr Ala Ser Tyr Asn Gly Asn Pro
85 90 95

Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg Ser Glu Val
100 105 110

His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu Arg Ala Ala
115 120 125

Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu Asp Arg Asn
130 135 140

Val Thr Val Asp Thr Leu Leu Val Glu Thr Leu Ser Glu Ile Arg Ala
145 150 155 160

Ala Asn Gln Ala Gly Ala Asn Pro Pro Tyr Ala Ala Gln Ile Val Val
165 170 175

Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn Gly Glu
180 185 190

Trp Ala Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Gly Tyr Ile Asn
195 200 205

151

Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg Thr Ile Leu
210 215 220

Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met Asn Val
225 230 235 240

Ala Lys Cys Ser Gly Ala Ala Ser Thr Tyr Arg Glu Leu Thr Ile Tyr
245 250 255

Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr Met Asp Ala
260 265 270

Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro Ala Ala
275 280 285

Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro Arg Ala Val
290 295 300

Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser Ile Ser
305 310 315 320

Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu Lys His
325 330 335

Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly Phe Pro Ala
340 345 350

Gln Phe Ile Val Asp Gln Gly Arg Ser Gly Lys Gln Pro Thr Gly Gln
355 360 365

Lys Glu Trp Gly His Trp Cys Asn Ala Ile Gly Thr Gly Phe Gly Met
370 375 380

Arg Pro Thr Ala Asn Thr Gly His Gln Tyr Val Asp Ala Phe Val Trp
385 390 395 400

Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp Thr Thr Ala Ala
405 410 415

Arg Tyr Asp Tyr His Cys Gly Leu Glu Asp Ala Leu Lys Pro Ala Pro
420 425 430

Glu Ala Gly Gln Trp Phe Gln Ala Tyr Phe Glu Gln Leu Leu Arg Asn
435 440 445

Ala Asn Pro Pro Phe
450

<210> 66
<211> 440
<212> PRT
<213> Phanerochaete chrysosporium

<400> 66

Ala Ser Ser Glu Trp Gly Gln Cys Gly Gly Ile Gly Trp Thr Gly Pro
1 5 10 15

Thr Thr Cys Val Ser Gly Thr Thr Cys Thr Val Leu Asn Pro Tyr Tyr
20 25 30

152

Ser Gln Cys Leu Pro Gly Ser Ala Val Thr Thr Thr Ser Val Ile Thr
35 40 45

Ser His Ser Ser Ser Val Ser Ser Val Ser Ser His Ser Gly Ser Ser
50 55 60

Thr Ser Thr Ser Ser Pro Thr Gly Pro Thr Gly Thr Asn Pro Pro Pro
65 70 75 80

Pro Pro Ser Ala Asn Asn Pro Trp Thr Gly Phe Gln Ile Phe Leu Ser
85 90 95

Pro Tyr Tyr Ala Asn Glu Val Ala Ala Ala Ala Lys Gln Ile Thr Asp
100 105 110

Pro Thr Leu Ser Ser Lys Ala Ala Ser Val Ala Asn Ile Pro Thr Phe
115 120 125

Thr Trp Leu Asp Ser Val Ala Lys Ile Pro Asp Leu Gly Thr Tyr Leu
130 135 140

Ala Ser Ala Ser Ala Leu Gly Lys Ser Thr Gly Thr Lys Gln Leu Val
145 150 155 160

Gln Ile Val Ile Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Lys Ala
165 170 175

Ser Asn Gly Glu Phe Ser Ile Ala Asn Asn Gly Gln Ala Asn Tyr Glu
180 185 190

Asn Tyr Ile Asp Gln Ile Val Ala Gln Ile Gln Gln Phe Pro Asp Val
195 200 205

Arg Val Val Ala Val Ile Glu Pro Asp Ser Leu Ala Asn Leu Val Thr
210 215 220

Asn Leu Asn Val Gln Lys Cys Ala Asn Ala Lys Thr Thr Tyr Leu Ala
225 230 235 240

Cys Val Asn Tyr Ala Leu Thr Asn Leu Ala Lys Val Gly Val Tyr Met
245 250 255

Tyr Met Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Leu
260 265 270

Ser Pro Ala Ala Gln Leu Phe Thr Gln Val Trp Gln Asn Ala Gly Lys
275 280 285

Ser Pro Phe Ile Lys Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala
290 295 300

Leu Gln Ala Ala Ser Pro Asp Pro Ile Thr Gln Gly Asn Pro Asn Tyr
305 310 315 320

Asp Glu Ile His Tyr Ile Asn Ala Leu Ala Pro Leu Leu Gln Gln Ala
325 330 335

Gly Trp Asp Ala Thr Phe Ile Val Asp Gln Gly Arg Ser Gly Val Gln
340 345 350

Asn Ile Arg Gln Gln Trp Gly Asp Trp Cys Asn Ile Lys Gly Ala Gly
355 360 365

Phe Gly Thr Arg Pro Thr Thr Asn Thr Gly Ser Gln Phe Ile Asp Ser
370 375 380

Ile Val Trp Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asn Ser
385 390 395 400

Ser Ser Pro Arg Tyr Asp Ser Thr Cys Ser Leu Pro Asp Ala Ala Gln
405 410 415

Pro Ala Pro Glu Ala Gly Thr Trp Phe Gln Ala Tyr Phe Gln Thr Leu
420 425 430

Val Ser Ala Ala Asn Pro Pro Leu
435 440

<210> 67

<211> 447

<212> PRT

<213> Trichoderma reesei

<400> 67

Gln Ala Cys Ser Ser Val Trp Gly Gln Cys Gly Gly Gln Asn Trp Ser
1 5 10 15

Gly Pro Thr Cys Cys Ala Ser Gly Ser Thr Cys Val Tyr Ser Asn Asp
20 25 30

Tyr Tyr Ser Gln Cys Leu Pro Gly Ala Ala Ser Ser Ser Ser Thr
35 40 45

Arg Ala Ala Ser Thr Thr Ser Arg Val Ser Pro Thr Thr Ser Arg Ser
50 55 60

Ser Ser Ala Thr Pro Pro Pro Gly Ser Thr Thr Thr Arg Val Pro Pro
65 70 75 80

Val Gly Ser Gly Thr Ala Thr Tyr Ser Gly Asn Pro Phe Val Gly Val
85 90 95

Thr Pro Trp Ala Asn Ala Tyr Tyr Ala Ser Glu Val Ser Ser Leu Ala
100 105 110

Ile Pro Ser Leu Thr Gly Ala Met Ala Thr Ala Ala Ala Val Ala
115 120 125

Lys Val Pro Ser Phe Met Trp Leu Asp Thr Leu Asp Lys Thr Pro Leu
130 135 140

Met Glu Gln Thr Leu Ala Asp Ile Arg Thr Ala Asn Lys Asn Gly Gly
145 150 155 160

Asn Tyr Ala Gly Gln Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys
165 170 175

Ala Ala Leu Ala Ser Asn Gly Glu Tyr Ser Ile Ala Asp Gly Gly Val
180 185 190

Ala Lys Tyr Lys Asn Tyr Ile Asp Thr Ile Arg Gln Ile Val Val Glu
195 200 205

154

Tyr Ser Asp Ile Arg Thr Leu Leu Val Ile Glu Pro Asp Ser Leu Ala
210 215 220

Asn Leu Val Thr Asn Leu Gly Thr Pro Lys Cys Ala Asn Ala Gln Ser
225 230 235 240

Ala Tyr Leu Glu Cys Ile Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro
245 250 255

Asn Val Ala Met Tyr Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp
260 265 270

Pro Ala Asn Gln Asp Pro Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys
275 280 285

Asn Ala Ser Ser Pro Arg Ala Leu Arg Gly Leu Ala Thr Asn Val Ala
290 295 300

Asn Tyr Asn Gly Trp Asn Ile Thr Ser Pro Pro Ser Tyr Thr Gln Gly
305 310 315 320

Asn Ala Val Tyr Asn Glu Lys Leu Tyr Ile His Ala Ile Gly Pro Leu
325 330 335

Leu Ala Asn His Gly Trp Ser Asn Ala Phe Phe Ile Thr Asp Gln Gly
340 345 350

Arg Ser Gly Lys Gln Pro Thr Gly Gln Gln Gln Trp Gly Asp Trp Cys
355 360 365

Asn Val Ile Gly Thr Gly Phe Gly Ile Arg Pro Ser Ala Asn Thr Gly
370 375 380

Asp Ser Leu Leu Asp Ser Phe Val Trp Val Lys Pro Gly Gly Glu Cys
385 390 395 400

Asp Gly Thr Ser Asp Ser Ser Ala Pro Arg Phe Asp Pro His Cys Ala
405 410 415

Leu Pro Asp Ala Leu Gln Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln
420 425 430

Ala Tyr Phe Val Gln Leu Leu Thr Asn Ala Asn Pro Ser Phe Leu
435 440 445

<210> 68

<211> 447

<212> PRT

<213> Trichoderma reesei

<400> 68

Gln Ala Cys Ser Ser Val Trp Gly Gln Cys Gly Gly Gln Asn Trp Ser
1 5 10 15

Gly Pro Thr Cys Cys Ala Ser Gly Ser Thr Cys Val Tyr Ser Asn Asp
20 25 30

Tyr Tyr Ser Gln Cys Leu Pro Gly Ala Ala Ser Ser Ser Ser Ser Thr
35 40 45

Arg Ala Ala Ser Thr Thr Ser Arg Val Ser Pro Thr Thr Ser Arg Ser

155

50 55 60

Ser Ser Ala Thr Pro Pro Pro Gly Ser Thr Thr Thr Arg Val Pro Pro
65 70 75 80

Val Glu Ser Gly Thr Ala Thr Tyr Ser Gly Asn Pro Phe Val Gly Val
85 90 95

Thr Pro Trp Ala Asn Ala Tyr Tyr Ala Ser Glu Val Ser Ser Leu Ala
100 105 110

Ile Pro Ser Leu Thr Gly Ala Met Ala Thr Ala Ala Ala Val Ala
115 120 125

Lys Val Pro Ser Phe Met Trp Leu Asp Thr Leu Asp Lys Thr Pro Leu
130 135 140

Met Glu Gln Thr Leu Ala Asp Ile Arg Thr Ala Asn Lys Asn Gly Gly
145 150 155 160

Asn Tyr Ala Gly Gln Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys
165 170 175

Ala Ala Leu Ala Ser Asn Gly Glu Tyr Ser Ile Ala Asp Gly Gly Val
180 185 190

Ala Lys Tyr Lys Asn Tyr Ile Asp Thr Ile Arg Gln Ile Val Val Glu
195 200 205

Tyr Ser Asp Ile Arg Thr Leu Leu Val Ile Glu Pro Asp Ser Leu Ala
210 215 220

Asn Leu Val Thr Asn Leu Ser Thr Pro Lys Cys Ala Asn Ala Gln Ser
225 230 235 240

Ala Tyr Leu Glu Cys Ile Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro
245 250 255

Asn Val Ala Met Tyr Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp
260 265 270

Pro Ala Asn Gln Asp Pro Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys
275 280 285

Asn Ala Ser Ser Pro Arg Ala Leu Arg Gly Leu Ala Thr Asn Val Ala
290 295 300

Ser Tyr Asn Gly Trp Asn Ile Thr Ser Pro Pro Ser Tyr Thr Gln Gly
305 310 315 320

Asn Ala Val Tyr Asn Glu Lys Leu Tyr Ile His Ala Ile Gly Pro Leu
325 330 335

Leu Ala Asn His Gly Trp Ser Asn Ala Phe Phe Ile Thr Asp Gln Gly
340 345 350

Arg Ser Gly Lys Gln Pro Thr Gly Gln Gln Gln Trp Gly Asp Trp Cys
355 360 365

Asn Val Ile Gly Thr Gly Phe Gly Ile Arg Pro Ser Ala Asn Thr Gly
370 375 380

156

Asp Ser Leu Leu Asp Ser Phe Val Trp Val Lys Pro Gly Gly Glu Cys
385 390 395 400

Asp Gly Thr Ser Asp Ser Ser Ala Pro Gln Phe Asp Pro His Cys Ala
405 410 415

Leu Pro Asp Ala Leu Gln Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln
420 425 430

Ala Tyr Phe Val Gln Leu Leu Thr Asn Ala Asn Pro Ser Phe Leu
435 440 445

<210> 69

<211> 447

<212> PRT

<213> Trichoderma reesei

<400> 69

Gln Ala Cys Ser Ser Val Trp Gly Gln Cys Gly Gly Gln Asn Trp Ser
1 5 10 15

Gly Pro Thr Cys Cys Ala Ser Gly Ser Thr Cys Val Tyr Ser Asn Asp
20 25 30

Tyr Tyr Ser Gln Cys Leu Pro Gly Ala Ala Ser Ser Ser Ser Ser Thr
35 40 45

Arg Ala Ala Ser Thr Thr Ser Arg Val Ser Pro Thr Thr Ser Arg Ser
50 55 60

Ser Ser Ala Thr Pro Pro Pro Gly Ser Thr Thr Thr Arg Val Pro Pro
65 70 75 80

Val Gly Ser Gly Thr Ala Thr Tyr Ser Gly Asn Pro Phe Val Gly Val
85 90 95

Thr Pro Trp Ala Asn Ala Tyr Tyr Ala Ser Glu Val Ser Ser Leu Ala
100 105 110

Ile Pro Ser Leu Thr Gly Ala Met Ala Thr Ala Ala Ala Val Ala
115 120 125

Lys Val Pro Ser Phe Met Trp Leu Asp Thr Leu Asp Lys Thr Pro Leu
130 135 140

Met Glu Gln Thr Leu Ala Asp Ile Arg Thr Ala Asn Lys Asn Gly Gly
145 150 155 160

Asn Tyr Ala Gly Gln Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys
165 170 175

Ala Ala Leu Ala Ser Asn Gly Glu Tyr Ser Ile Ala Asp Gly Gly Val
180 185 190

Ala Lys Tyr Lys Asn Tyr Ile Asp Thr Ile Arg Gln Ile Val Val Glu
195 200 205

Tyr Ser Asp Ile Arg Thr Leu Leu Val Ile Glu Pro Asp Ser Leu Ala
210 215 220

Asn Leu Val Thr Asn Leu Ser Thr Pro Lys Cys Ala Asn Ala Gln Ser

157

225 230 235 240

Ala Tyr Leu Glu Cys Ile Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro
245 250 255

Asn Val Ala Met Tyr Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp
260 265 270

Pro Ala Asn Gln Asp Pro Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys
275 280 285

Asn Ala Ser Ser Pro Arg Ala Leu Arg Gly Leu Ala Thr Asn Val Ala
290 295 300

Asn Tyr Asn Gly Trp Asn Ile Thr Ser Pro Pro Ser Tyr Thr Gln Gly
305 310 315 320

Asn Ala Val Tyr Asn Glu Lys Leu Tyr Ile His Ala Ile Gly Pro Leu
325 330 335

Leu Ala Asn His Gly Trp Ser Asn Ala Phe Phe Ile Thr Asp Gln Gly
340 345 350

Arg Ser Gly Lys Gln Pro Thr Gly Gln Gln Gln Trp Gly Asp Trp Cys
355 360 365

Asn Val Ile Gly Thr Gly Phe Gly Ile Arg Pro Ser Ala Asn Thr Gly
370 375 380

Asp Ser Leu Leu Asp Ser Phe Val Trp Val Lys Pro Gly Gly Glu Cys
385 390 395 400

Asp Gly Thr Ser Asp Ser Ser Ala Pro Arg Phe Asp Pro His Cys Ala
405 410 415

Leu Pro Asp Ala Leu Gln Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln
420 425 430

Ala Tyr Phe Val Gln Leu Leu Thr Asn Ala Asn Pro Ser Phe Leu
435 440 445

<210> 70

<211> 447

<212> PRT

<213> Trichoderma reesei

<400> 70

Gln Ala Cys Ser Ser Val Trp Gly Gln Cys Gly Gly Gln Asn Trp Ser
1 5 10 15

Gly Pro Thr Cys Cys Ala Ser Gly Ser Thr Cys Val Tyr Ser Asn Asp
20 25 30

Tyr Tyr Ser Gln Cys Leu Pro Gly Ala Ala Ser Ser Ser Ser Thr
35 40 45

Arg Ala Ala Ser Thr Thr Ser Arg Val Ser Pro Thr Thr Ser Arg Ser
50 55 60

Ser Ser Ala Thr Pro Pro Pro Gly Ser Thr Thr Thr Arg Val Pro Pro
65 70 75 80

15

Val Gly Ser Gly Thr Ala Thr Tyr Ser Gly Asn Pro Phe Val Gly Val
85 90 95

Thr Pro Trp Ala Asn Ala Tyr Tyr Ala Ser Glu Val Ser Ser Leu Ala
100 105 110

Ile Pro Ser Leu Thr Gly Ala Met Ala Thr Ala Ala Ala Val Ala
115 120 125

Lys Val Pro Ser Phe Met Trp Leu Asp Thr Leu Asp Lys Thr Pro Leu
130 135 140

Met Glu Gln Thr Leu Ala Asp Ile Arg Thr Ala Asn Lys Asn Gly Gly
145 150 155 160

Asn Tyr Ala Gly Gln Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys
165 170 175

Ala Ala Leu Ala Ser Asn Gly Glu Tyr Ser Ile Ala Asp Gly Gly Val
180 185 190

Ala Lys Tyr Lys Asn Tyr Ile Asp Thr Ile Arg Gln Ile Val Val Glu
195 200 205

Tyr Ser Asp Ile Arg Thr Leu Leu Val Ile Glu Pro Asp Ser Leu Ala
210 215 220

Asn Leu Val Thr Asn Leu Gly Thr Pro Lys Cys Ala Asn Ala Gln Ser
225 230 235 240

Ala Tyr Leu Glu Cys Ile Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro
245 250 255

Asn Val Ala Met Tyr Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp
260 265 270

Pro Ala Asn Gln Asp Pro Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys
275 280 285

Asn Ala Ser Ser Pro Arg Ala Leu Arg Gly Leu Ala Thr Asn Val Ala
290 295 300

Ser Tyr Asn Gly Trp Asn Ile Thr Ser Pro Pro Ser Tyr Thr Gln Gly
305 310 315 320

Asn Ala Val Tyr Asn Glu Lys Leu Tyr Ile His Ala Ile Gly Pro Leu
325 330 335

Leu Ala Asn His Gly Trp Ser Asn Ala Phe Phe Ile Thr Asp Gln Gly
340 345 350

Arg Ser Gly Lys Gln Pro Thr Gly Gln Gln Gln Trp Gly Asp Trp Cys
355 360 365

Asn Val Ile Gly Thr Gly Phe Gly Ile Arg Pro Ser Ala Asn Thr Gly
370 375 380

Asp Ser Leu Leu Asp Ser Phe Val Trp Val Lys Pro Gly Gly Glu Cys
385 390 395 400

Asp Gly Thr Ser Asp Ser Ser Ala Pro Arg Phe Asp Pro His Cys Ala

15

405 410 415

Leu Pro Asp Ala Leu Gln Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln
420 425 430

Ala Tyr Phe Val Gln Leu Leu Thr Asn Ala Asn Pro Ser Phe Leu
435 440 445

<210> 71

<211> 447

<212> PRT

<213> Trichoderma reesei

<400> 71

Gln Ala Cys Ser Ser Val Trp Gly Gln Cys Gly Gly Gln Asn Trp Ser
1 5 10 15

Gly Pro Thr Cys Cys Ala Ser Gly Ser Thr Cys Val Tyr Ser Asn Asp
20 25 30

Tyr Tyr Ser Gln Cys Leu Pro Gly Ala Ala Ser Ser Ser Ser Ser Thr
35 40 45

Arg Ala Ala Ser Thr Thr Ser Arg Val Ser Pro Thr Thr Ser Arg Ser
50 55 60

Ser Ser Ala Thr Pro Pro Pro Gly Ser Thr Thr Thr Arg Val Pro Pro
65 70 75 80

Val Gly Ser Gly Thr Ala Thr Tyr Ser Gly Asn Pro Phe Val Gly Val
85 90 95

Thr Pro Trp Ala Asn Ala Tyr Tyr Ala Ser Glu Val Ser Ser Leu Ala
100 105 110

Ile Pro Ser Leu Thr Gly Ala Met Ala Thr Ala Ala Ala Val Ala
115 120 125

Lys Val Pro Ser Phe Met Trp Leu Asp Thr Leu Asp Lys Thr Pro Leu
130 135 140

Met Glu Gln Thr Leu Ala Asp Ile Arg Thr Ala Asn Lys Asn Gly Gly
145 150 155 160

Asn Tyr Ala Gly Gln Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys
165 170 175

Ala Ala Leu Ala Ser Asn Gly Glu Tyr Ser Ile Ala Asp Gly Gly Val
180 185 190

Ala Lys Tyr Lys Asn Tyr Ile Asp Thr Ile Arg Gln Ile Val Val Glu
195 200 205

Tyr Ser Asp Ile Arg Thr Leu Leu Val Ile Glu Pro Asp Ser Leu Ala
210 215 220

Asn Leu Val Thr Asn Leu Gly Thr Pro Lys Cys Ala Asn Ala Gln Ser
225 230 235 240

Ala Tyr Leu Glu Cys Ile Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro
245 250 255

160

Asn Val Ala Met Tyr Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp
260 265 270

Pro Ala Asn Gln Asp Pro Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys
275 280 285

Asn Ala Ser Ser Pro Arg Ala Leu Arg Gly Leu Ala Thr Asn Val Ala
290 295 300

Asn Tyr Asn Gly Trp Asn Ile Thr Ser Pro Pro Ser Tyr Thr Gln Gly
305 310 315 320

Asn Ala Val Tyr Asn Glu Lys Leu Tyr Ile His Ala Ile Gly Pro Leu
325 330 335

Leu Ala Asn His Gly Trp Ser Asn Ala Phe Phe Ile Thr Asp Gln Gly
340 345 350

Arg Ser Gly Lys Gln Pro Thr Gly Gln Gln Gln Trp Gly Asp Trp Cys
355 360 365

Asn Val Ile Gly Thr Gly Phe Gly Ile Arg Pro Ser Ala Asn Thr Gly
370 375 380

Asp Ser Leu Leu Asp Ser Phe Val Trp Val Lys Pro Gly Gly Glu Cys
385 390 395 400

Asp Gly Thr Ser Asp Ser Ser Ala Pro Gln Phe Asp Pro His Cys Ala
405 410 415

Leu Pro Asp Ala Leu Gln Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln
420 425 430

Ala Tyr Phe Val Gln Leu Leu Thr Asn Ala Asn Pro Ser Phe Leu
435 440 445

<210> 72
<211> 447
<212> PRT
<213> Trichoderma reesei

<400> 72

Gln Ala Cys Ser Ser Val Trp Gly Gln Cys Gly Gly Gln Asn Trp Ser
1 5 10 15

Gly Pro Thr Cys Cys Ala Ser Gly Ser Thr Cys Val Tyr Ser Asn Asp
20 25 30

Tyr Tyr Ser Gln Cys Leu Pro Gly Ala Ala Ser Ser Ser Ser Thr
35 40 45

Arg Ala Ala Ser Thr Thr Ser Arg Val Ser Pro Thr Thr Ser Arg Ser
50 55 60

Ser Ser Ala Thr Pro Pro Pro Gly Ser Thr Thr Thr Arg Val Pro Pro
65 70 75 80

Val Gly Ser Gly Thr Ala Thr Tyr Ser Gly Asn Pro Phe Val Gly Val
85 90 95

Thr Pro Trp Ala Asn Ala Tyr Tyr Ala Ser Glu Val Ser Ser Leu Ala
100 105 110

Ile Pro Ser Leu Thr Gly Ala Met Ala Thr Ala Ala Ala Val Ala
115 120 125

Lys Val Pro Ser Phe Met Trp Leu Asp Thr Leu Asp Lys Thr Pro Leu
130 135 140

Met Glu Gln Thr Leu Ala Asp Ile Arg Thr Ala Asn Lys Asn Gly Gly
145 150 155 160

Asn Tyr Ala Gly Gln Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys
165 170 175

Ala Ala Leu Ala Ser Asn Gly Glu Tyr Ser Ile Ala Asp Gly Gly Val
180 185 190

Ala Lys Tyr Lys Asn Tyr Ile Asp Thr Ile Arg Gln Ile Val Val Glu
195 200 205

Tyr Ser Asp Ile Arg Thr Leu Leu Val Ile Glu Pro Asp Ser Leu Ala
210 215 220

Asn Leu Val Thr Asn Leu Ser Thr Pro Lys Cys Ala Asn Ala Gln Ser
225 230 235 240

Ala Tyr Leu Glu Cys Ile Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro
245 250 255

Asn Val Ala Met Tyr Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp
260 265 270

Pro Ala Asn Gln Asp Pro Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys
275 280 285

Asn Ala Ser Ser Pro Arg Ala Leu Arg Gly Leu Ala Thr Asn Val Ala
290 295 300

Ser Tyr Asn Gly Trp Asn Ile Thr Ser Pro Pro Ser Tyr Thr Gln Gly
305 310 315 320

Asn Ala Val Tyr Asn Glu Lys Leu Tyr Ile His Ala Ile Gly Pro Leu
325 330 335

Leu Ala Asn His Gly Trp Ser Asn Ala Phe Phe Ile Thr Asp Gln Gly
340 345 350

Arg Ser Gly Lys Gln Pro Thr Gly Gln Gln Gln Trp Gly Asp Trp Cys
355 360 365

Asn Val Ile Gly Thr Gly Phe Gly Ile Arg Pro Ser Ala Asn Thr Gly
370 375 380

Asp Ser Leu Leu Asp Ser Phe Val Trp Val Lys Pro Gly Gly Glu Cys
385 390 395 400

Asp Gly Thr Ser Asp Ser Ser Ala Pro Arg Phe Asp Pro His Cys Ala
405 410 415

Leu Pro Asp Ala Leu Gln Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln
420 425 430

162

Ala Tyr Phe Val Gln Leu Leu Thr Asn Ala Asn Pro Ser Phe Leu
435 440 445

<210> 73

<211> 447

<212> PRT

<213> Trichoderma reesei

<400> 73

Gln Ala Cys Ser Ser Val Trp Gly Gln Cys Gly Gly Gln Asn Trp Ser
1 5 10 15

Gly Pro Thr Cys Cys Ala Ser Gly Ser Thr Cys Val Tyr Ser Asn Asp
20 25 30

Tyr Tyr Ser Gln Cys Leu Pro Gly Ala Ala Ser Ser Ser Ser Ser Thr
35 40 45

Arg Ala Ala Ser Thr Thr Ser Arg Val Ser Pro Thr Thr Ser Arg Ser
50 55 60

Ser Ser Ala Thr Pro Pro Pro Gly Ser Thr Thr Thr Arg Val Pro Pro
65 70 75 80

Val Gly Ser Gly Thr Ala Thr Tyr Ser Gly Asn Pro Phe Val Gly Val
85 90 95

Thr Pro Trp Ala Asn Ala Tyr Tyr Ala Ser Glu Val Ser Ser Leu Ala
100 105 110

Ile Pro Ser Leu Thr Gly Ala Met Ala Thr Ala Ala Ala Val Ala
115 120 125

Lys Val Pro Ser Phe Met Trp Leu Asp Thr Leu Asp Lys Thr Pro Leu
130 135 140

Met Glu Gln Thr Leu Ala Asp Ile Arg Thr Ala Asn Lys Asn Gly Gly
145 150 155 160

Asn Tyr Ala Gly Gln Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys
165 170 175

Ala Ala Leu Ala Ser Asn Gly Glu Tyr Ser Ile Ala Asp Gly Gly Val
180 185 190

Ala Lys Tyr Lys Asn Tyr Ile Asp Thr Ile Arg Gln Ile Val Val Glu
195 200 205

Tyr Ser Asp Ile Arg Thr Leu Leu Val Ile Glu Pro Asp Ser Leu Ala
210 215 220

Asn Leu Val Thr Asn Leu Ser Thr Pro Lys Cys Ala Asn Ala Gln Ser
225 230 235 240

Ala Tyr Leu Glu Cys Ile Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro
245 250 255

Asn Val Ala Met Tyr Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp
260 265 270

3
16

Pro Ala Asn Gln Asp Pro Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys
275 280 285

Asn Ala Ser Ser Pro Arg Ala Leu Arg Gly Leu Ala Thr Asn Val Ala
290 295 300

Asn Tyr Asn Gly Trp Asn Ile Thr Ser Pro Pro Ser Tyr Thr Gln Gly
305 310 315 320

Asn Ala Val Tyr Asn Glu Lys Leu Tyr Ile His Ala Ile Gly Pro Leu
325 330 335

Leu Ala Asn His Gly Trp Ser Asn Ala Phe Phe Ile Thr Asp Gln Gly
340 345 350

Arg Ser Gly Lys Gln Pro Thr Gly Gln Gln Gln Trp Gly Asp Trp Cys
355 360 365

Asn Val Ile Gly Thr Gly Phe Gly Ile Arg Pro Ser Ala Asn Thr Gly
370 375 380

Asp Ser Leu Leu Asp Ser Phe Val Trp Val Lys Pro Gly Gly Glu Cys
385 390 395 400

Asp Gly Thr Ser Asp Ser Ser Ala Pro Gln Phe Asp Pro His Cys Ala
405 410 415

Leu Pro Asp Ala Leu Gln Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln
420 425 430

Ala Tyr Phe Val Gln Leu Leu Thr Asn Ala Asn Pro Ser Phe Leu
435 440 445

<210> 74
<211> 447
<212> PRT
<213> Trichoderma reesei

<400> 74

Gln Ala Cys Ser Ser Val Trp Gly Gln Cys Gly Gly Gln Asn Trp Ser
1 5 10 15

Gly Pro Thr Cys Cys Ala Ser Gly Ser Thr Cys Val Tyr Ser Asn Asp
20 25 30

Tyr Tyr Ser Gln Cys Leu Pro Gly Ala Ala Ser Ser Ser Ser Thr
35 40 45

Arg Ala Ala Ser Thr Thr Ser Arg Val Ser Pro Thr Thr Ser Arg Ser
50 55 60

Ser Ser Ala Thr Pro Pro Pro Gly Ser Thr Thr Thr Arg Val Pro Pro
65 70 75 80

Val Gly Ser Gly Thr Ala Thr Tyr Ser Gly Asn Pro Phe Val Gly Val
85 90 95

Thr Pro Trp Ala Asn Ala Tyr Tyr Ala Ser Glu Val Ser Ser Leu Ala
100 105 110

Ile Pro Ser Leu Thr Gly Ala Met Ala Thr Ala Ala Ala Val Ala

164

115 120 125
Lys Val Pro Ser Phe Met Trp Leu Asp Thr Leu Asp Lys Thr Pro Leu
130 135 140
Met Glu Gln Thr Leu Ala Asp Ile Arg Thr Ala Asn Lys Asn Gly Gly
145 150 155 160
Asn Tyr Ala Gly Gln Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys
165 170 175
Ala Ala Leu Ala Ser Asn Gly Glu Tyr Ser Ile Ala Asp Gly Gly Val
180 185 190
Ala Lys Tyr Lys Asn Tyr Ile Asp Thr Ile Arg Gln Ile Val Val Glu
195 200 205
Tyr Ser Asp Ile Arg Thr Leu Leu Val Ile Glu Pro Asp Ser Leu Ala
210 215 220
Asn Leu Val Thr Asn Leu Gly Thr Pro Lys Cys Ala Asn Ala Gln Ser
225 230 235 240
Ala Tyr Leu Glu Cys Ile Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro
245 250 255
Asn Val Ala Met Tyr Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp
260 265 270
Pro Ala Asn Gln Asp Pro Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys
275 280 285
Asn Ala Ser Ser Pro Arg Ala Leu Arg Gly Leu Ala Thr Asn Val Ala
290 295 300
Ser Tyr Asn Gly Trp Asn Ile Thr Ser Pro Pro Ser Tyr Thr Gln Gly
305 310 315 320
Asn Ala Val Tyr Asn Glu Lys Leu Tyr Ile His Ala Ile Gly Pro Leu
325 330 335
Leu Ala Asn His Gly Trp Ser Asn Ala Phe Phe Ile Thr Asp Gln Gly
340 345 350
Arg Ser Gly Lys Gln Pro Thr Gly Gln Gln Gln Trp Gly Asp Trp Cys
355 360 365
Asn Val Ile Gly Thr Gly Phe Gly Ile Arg Pro Ser Ala Asn Thr Gly
370 375 380
Asp Ser Leu Leu Asp Ser Phe Val Trp Val Lys Pro Gly Gly Glu Cys
385 390 395 400
Asp Gly Thr Ser Asp Ser Ser Ala Pro Gln Phe Asp Pro His Cys Ala
405 410 415
Leu Pro Asp Ala Leu Gln Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln
420 425 430
Ala Tyr Phe Val Gln Leu Leu Thr Asn Ala Asn Pro Ser Phe Leu
435 440 445

165

<210> 75
<211> 447
<212> PRT
<213> Trichoderma reesei

<400> 75

Gln Ala Cys Ser Ser Val Trp Gly Gln Cys Gly Gly Gln Asn Trp Ser
1 5 10 15

Gly Pro Thr Cys Cys Ala Ser Gly Ser Thr Cys Val Tyr Ser Asn Asp
20 25 30

Tyr Tyr Ser Gln Cys Leu Pro Gly Ala Ala Ser Ser Ser Ser Ser Thr
35 40 45

Arg Ala Ala Ser Thr Thr Ser Arg Val Ser Pro Thr Thr Ser Arg Ser
50 55 60

Ser Ser Ala Thr Pro Pro Gly Ser Thr Thr Thr Arg Val Pro Pro
65 70 75 80

Val Gly Ser Gly Thr Ala Thr Tyr Ser Gly Asn Pro Phe Val Gly Val
85 90 95

Thr Pro Trp Ala Asn Ala Tyr Tyr Ala Ser Glu Val Ser Ser Leu Ala
100 105 110

Ile Pro Ser Leu Thr Gly Ala Met Ala Thr Ala Ala Ala Val Ala
115 120 125

Lys Val Pro Ser Phe Met Trp Leu Asp Thr Leu Asp Lys Thr Pro Leu
130 135 140

Met Glu Gln Thr Leu Ala Asp Ile Arg Thr Ala Asn Lys Asn Gly Gly
145 150 155 160

Asn Tyr Ala Gly Gln Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys
165 170 175

Ala Ala Leu Ala Ser Asn Gly Glu Tyr Ser Ile Ala Asp Gly Gly Val
180 185 190

Ala Lys Tyr Lys Asn Tyr Ile Asp Thr Ile Arg Gln Ile Val Val Glu
195 200 205

Tyr Ser Asp Ile Arg Thr Leu Leu Val Ile Glu Pro Asp Ser Leu Ala
210 215 220

Asn Leu Val Thr Asn Leu Ser Thr Pro Lys Cys Ala Asn Ala Gln Ser
225 230 235 240

Ala Tyr Leu Glu Cys Ile Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro
245 250 255

Asn Val Ala Met Tyr Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp
260 265 270

Pro Ala Asn Gln Asp Pro Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys
275 280 285

Asn Ala Ser Ser Pro Arg Ala Leu Arg Gly Leu Ala Thr Asn Val Ala

290 295 300

Ser Tyr Asn Gly Trp Asn Ile Thr Ser Pro Pro Ser Tyr Thr Gln Gly
305 310 315 320

Asn Ala Val Tyr Asn Glu Lys Leu Tyr Ile His Ala Ile Gly Pro Leu
325 330 335

Leu Ala Asn His Gly Trp Ser Asn Ala Phe Phe Ile Thr Asp Gln Gly
340 345 350

Arg Ser Gly Lys Gln Pro Thr Gly Gln Gln Gln Trp Gly Asp Trp Cys
355 360 365

Asn Val Ile Gly Thr Gly Phe Gly Ile Arg Pro Ser Ala Asn Thr Gly
370 375 380

Asp Ser Leu Leu Asp Ser Phe Val Trp Val Lys Pro Gly Gly Glu Cys
385 390 395 400

Asp Gly Thr Ser Asp Ser Ser Ala Pro Gln Phe Asp Pro His Cys Ala
405 410 415

Leu Pro Asp Ala Leu Gln Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln
420 425 430

Ala Tyr Phe Val Gln Leu Leu Thr Asn Ala Asn Pro Ser Phe Leu
435 440 445

<210> 76

<211> 453

<212> PRT

<213> Humicola insolens

<400> 76

Ala Ser Cys Ala Pro Thr Trp Gly Gln Cys Gly Gly Ile Gly Phe Asn
1 5 10 15

Gly Pro Thr Cys Cys Gln Ser Gly Ser Thr Cys Val Lys Gln Asn Asp
20 25 30

Trp Tyr Ser Gln Cys Leu Pro Gly Ser Gln Val Thr Thr Thr Ser Thr
35 40 45

Thr Ser Thr Ser Ser Ser Ser Thr Thr Ser Arg Ala Thr Ser Thr Thr
50 55 60

Ser Thr Gly Gly Val Thr Ser Ile Thr Thr Ala Pro Thr Arg Thr Val
65 70 75 80

Thr Ile Pro Gly Gly Ala Thr Thr Thr Ala Ser Tyr Asn Gly Asn Pro
85 90 95

Phe Glu Gly Val Gln Leu Trp Ala Asn Asn Tyr Tyr Arg Ser Glu Val
100 105 110

His Thr Leu Ala Ile Pro Gln Ile Thr Asp Pro Ala Leu Arg Ala Ala
115 120 125

Ala Ser Ala Val Ala Glu Val Pro Ser Phe Gln Trp Leu Asp Arg Asn
130 135 140

167

Val Thr Val Asp Thr Leu Leu Val Glu Thr Leu Ser Glu Ile Arg Ala
145 150 155 160

Ala Asn Gln Ala Gly Ala Asn Pro Pro Tyr Ala Ala Gln Ile Val Val
165 170 175

Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn Gly Glu
180 185 190

Trp Ala Ile Ala Asn Asn Gly Ala Asn Asn Tyr Lys Gly Tyr Ile Asn
195 200 205

Arg Ile Arg Glu Ile Leu Ile Ser Phe Ser Asp Val Arg Thr Ile Leu
210 215 220

Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met Asn Val
225 230 235 240

Ala Lys Cys Ser Gly Ala Ala Ser Thr Tyr Arg Glu Leu Thr Ile Tyr
245 250 255

Ala Leu Lys Gln Leu Asp Leu Pro His Val Ala Met Tyr Met Asp Ala
260 265 270

Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro Ala Ala
275 280 285

Glu Leu Phe Ala Lys Ile Tyr Glu Asp Ala Gly Lys Pro Arg Ala Val
290 295 300

Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser Ile Ser
305 310 315 320

Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu Lys His
325 330 335

Tyr Ile Glu Ala Phe Arg Pro Leu Leu Glu Ala Arg Gly Phe Pro Ala
340 345 350

Gln Phe Ile Val Asp Gln Gly Arg Ser Gly Lys Gln Pro Thr Gly Gln
355 360 365

Lys Glu Trp Gly His Trp Cys Asn Ala Ile Gly Thr Gly Phe Gly Met
370 375 380

Arg Pro Thr Ala Asn Thr Gly His Gln Tyr Val Asp Ala Phe Val Trp
385 390 395 400

Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp Thr Thr Ala Ala
405 410 415

Arg Tyr Asp Pro His Cys Gly Leu Glu Asp Ala Leu Lys Pro Ala Pro
420 425 430

Glu Ala Gly Gln Trp Phe Gln Ala Tyr Phe Glu Gln Leu Leu Arg Asn
435 440 445

Ala Asn Pro Pro Phe
450

<210> 77

168

<211> 440
<212> PRT
<213> Phanerochaete chrysosporium

<400> 77

Ala Ser Ser Glu Trp Gly Gln Cys Gly Gly Ile Gly Trp Thr Gly Pro
1 5 10 15

Thr Thr Cys Val Ser Gly Thr Thr Cys Thr Val Leu Asn Pro Tyr Tyr
20 25 30

Ser Gln Cys Leu Pro Gly Ser Ala Val Thr Thr Thr Ser Val Ile Thr
35 40 45

Ser His Ser Ser Ser Val Ser Ser Val Ser Ser His Ser Gly Ser Ser
50 55 60

Thr Ser Thr Ser Ser Pro Thr Gly Pro Thr Gly Thr Asn Pro Pro Pro
65 70 75 80

Pro Pro Ser Ala Asn Asn Pro Trp Thr Gly Phe Gln Ile Phe Leu Ser
85 90 95

Pro Tyr Tyr Ala Asn Glu Val Ala Ala Ala Ala Lys Gln Ile Thr Asp
100 105 110

Pro Thr Leu Ser Ser Lys Ala Ala Ser Val Ala Asn Ile Pro Thr Phe
115 120 125

Thr Trp Leu Asp Ser Val Ala Lys Ile Pro Asp Leu Gly Thr Tyr Leu
130 135 140

Ala Ser Ala Ser Ala Leu Gly Lys Ser Thr Gly Thr Lys Gln Leu Val
145 150 155 160

Gln Ile Val Ile Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Lys Ala
165 170 175

Ser Asn Gly Glu Phe Ser Ile Ala Asn Asn Gly Gln Ala Asn Tyr Glu
180 185 190

Asn Tyr Ile Asp Gln Ile Val Ala Gln Ile Gln Gln Phe Pro Asp Val
195 200 205

Arg Val Val Ala Val Ile Glu Pro Asp Ser Leu Ala Asn Leu Val Thr
210 215 220

Asn Leu Asn Val Gln Lys Cys Ala Asn Ala Lys Thr Thr Tyr Leu Ala
225 230 235 240

Cys Val Asn Tyr Ala Leu Thr Asn Leu Ala Lys Val Gly Val Tyr Met
245 250 255

Tyr Met Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Leu
260 265 270

Ser Pro Ala Ala Gln Leu Phe Thr Gln Val Trp Gln Asn Ala Gly Lys
275 280 285

Ser Pro Phe Ile Lys Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala
290 295 300

169

Leu Gln Ala Ala Ser Pro Asp Pro Ile Thr Gln Gly Asn Pro Asn Tyr
305 310 315 320

Asp Glu Ile His Tyr Ile Asn Ala Leu Ala Pro Leu Leu Gln Gln Ala
325 330 335

Gly Trp Asp Ala Thr Phe Ile Val Asp Gln Gly Arg Ser Gly Val Gln
340 345 350

Asn Ile Arg Gln Gln Trp Gly Asp Trp Cys Asn Ile Lys Gly Ala Gly
355 360 365

Phe Gly Thr Arg Pro Thr Thr Asn Thr Gly Ser Gln Phe Ile Asp Ser
370 375 380

Ile Val Trp Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asn Ser
385 390 395 400

Ser Ser Pro Arg Tyr Asp Pro Thr Cys Ser Leu Pro Asp Ala Ala Gln
405 410 415

Pro Ala Pro Glu Ala Gly Thr Trp Phe Gln Ala Tyr Phe Gln Thr Leu
420 425 430

Val Ser Ala Ala Asn Pro Pro Leu
435 440

<210> 78

<211> 459

<212> PRT

<213> Trichoderma reesei

<400> 78

Asp Tyr Lys Asp Asp Asp Asp Lys Glu Phe Leu Glu Ala Ser Cys Ser
1 5 10 15

Ser Val Trp Gly Gln Cys Gly Gly Gln Asn Trp Ser Gly Pro Thr Cys
20 25 30

Cys Ala Ser Gly Ser Thr Cys Val Tyr Ser Asn Asp Tyr Tyr Ser Gln
35 40 45

Cys Leu Pro Gly Ala Ala Ser Ser Ser Ser Ser Thr Arg Ala Ala Ser
50 55 60

Thr Thr Ser Arg Val Ser Pro Thr Thr Ser Arg Ser Ser Ser Ala Thr
65 70 75 80

Pro Pro Pro Gly Ser Thr Thr Thr Arg Val Pro Pro Val Gly Ser Gly
85 90 95

Thr Ala Thr Tyr Ser Gly Asn Pro Phe Val Gly Val Thr Pro Trp Ala
100 105 110

Asn Ala Tyr Tyr Ala Ser Glu Val Ser Ser Leu Ala Ile Pro Ser Leu
115 120 125

Thr Gly Ala Met Ala Thr Ala Ala Ala Val Ala Lys Val Pro Ser
130 135 140

Phe Met Trp Leu Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr
145 150 155 160

Leu Ala Asp Ile Arg Thr Ala Asn Lys Asn Gly Gly Asn Tyr Ala Gly
165 170 175

Gln Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Leu Ala
180 185 190

Ser Asn Gly Glu Tyr Ser Ile Ala Asp Gly Gly Val Ala Lys Tyr Lys
195 200 205

Asn Tyr Ile Asp Thr Ile Arg Gln Ile Val Val Glu Tyr Ser Asp Ile
210 215 220

Arg Thr Leu Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Leu Val Thr
225 230 235 240

Asn Leu Gly Thr Pro Lys Cys Ala Asn Ala Gln Ser Ala Tyr Leu Glu
245 250 255

Cys Ile Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met
260 265 270

Tyr Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln
275 280 285

Asp Pro Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser
290 295 300

Pro Arg Ala Leu Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Gly
305 310 315 320

Trp Asn Ile Thr Ser Pro Pro Ser Tyr Thr Gln Gly Asn Ala Val Tyr
325 330 335

Asn Glu Lys Leu Tyr Ile His Ala Ile Gly Pro Leu Leu Ala Asn His
340 345 350

Gly Trp Ser Asn Ala Phe Phe Ile Thr Asp Gln Gly Arg Ser Gly Lys
355 360 365

Gln Pro Thr Gly Gln Gln Gln Trp Gly Asp Trp Cys Asn Val Ile Gly
370 375 380

Thr Gly Phe Gly Ile Arg Pro Ser Ala Asn Thr Gly Asp Ser Leu Leu
385 390 395 400

Asp Ser Phe Val Trp Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser
405 410 415

Asp Ser Ser Ala Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala
420 425 430

Leu Gln Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val
435 440 445

Gln Leu Leu Thr Asn Ala Asn Pro Ser Phe Leu
450 455

<210> 79

<211> 484

<212> PRT

<213> Podospora anserina

<400> 79

Met Ala Lys Arg Leu Leu Leu Thr Ala Ala Leu Ala Ala Thr Thr Leu
1 5 10 15

Ala Ala Pro Val Ile Glu Glu Arg Gln Asn Cys Gly Ser Val Trp Ser
20 25 30

Gln Cys Gly Gly Gln Gly Trp Thr Gly Ala Thr Cys Cys Ala Ser Gly
35 40 45

Ser Thr Cys Val Ala Gln Asn Gln Trp Tyr Ser Gln Cys Leu Pro Gly
50 55 60

Ser Gln Val Thr Thr Thr Ala Gln Ala Pro Ser Ser Thr Arg Thr Thr
65 70 75 80

Thr Ser Ser Ser Ser Arg Pro Thr Ser Ser Ser Ile Ser Thr Ser Ala
85 90 95

Val Asn Val Pro Thr Thr Thr Thr Ser Ala Gly Ala Ser Val Thr Val
100 105 110

Pro Pro Gly Gly Gly Ala Ser Ser Thr Ala Ser Tyr Ser Gly Asn Pro
115 120 125

Phe Leu Gly Val Gln Gln Trp Ala Asn Ser Tyr Tyr Ser Ser Glu Val
130 135 140

His Thr Leu Ala Ile Pro Ser Leu Thr Gly Pro Met Ala Thr Lys Ala
145 150 155 160

Ala Ala Val Ala Lys Val Pro Ser Phe Gln Trp Met Asp Arg Asn Val
165 170 175

Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Asp Ile Arg Ala Ala
180 185 190

Asn Arg Ala Gly Ala Asn Pro Pro Tyr Ala Gly Ile Phe Val Val Tyr
195 200 205

Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn Gly Glu Trp
210 215 220

Ala Ile Ala Asp Gly Gly Ala Ala Lys Tyr Lys Ala Tyr Ile Asp Arg
225 230 235 240

Ile Arg His His Leu Val Gln Tyr Ser Asp Ile Arg Thr Ile Leu Val
245 250 255

Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met Asn Val Pro
260 265 270

Lys Cys Gln Gly Ala Ala Asn Thr Tyr Lys Glu Leu Thr Val Tyr Ala
275 280 285

Leu Lys Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu Asp Ala Gly
290 295 300

His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gly Pro Ala Ala Glu
305 310 315 320

Leu Phe Ala Gly Ile Tyr Lys Asp Ala Gly Arg Pro Thr Ser Leu Arg
325 330 335

Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Gly Trp Ser Leu Ser Ser
340 345 350

Ala Pro Ser Tyr Thr Thr Pro Asn Pro Asn Phe Asp Glu Lys Arg Phe
355 360 365

Val Gln Ala Phe Ser Pro Leu Leu Thr Ala Ala Gly Phe Pro Ala His
370 375 380

Phe Ile Thr Asp Thr Gly Arg Ser Gly Lys Gln Pro Thr Gly Gln Leu
385 390 395 400

Glu Trp Gly His Trp Cys Asn Ala Ile Gly Thr Gly Phe Gly Pro Arg
405 410 415

Pro Thr Thr Asp Thr Gly Leu Asp Ile Glu Asp Ala Phe Val Trp Ile
420 425 430

Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp Thr Thr Ala Ala Arg
435 440 445

Tyr Asp His His Cys Gly Phe Ala Asp Ala Leu Lys Pro Ala Pro Glu
450 455 460

Ala Gly Gln Trp Phe Gln Ala Tyr Phe Glu Gln Leu Leu Thr Asn Ala
465 470 475 480

Asn Pro Pro Phe

<210> 80

<211> 485

<212> PRT

<213> Chaetomium Globosum

<400> 80

Met Ala Ala Lys Leu Phe Leu Ala Ala Ala Leu Ala Ala Thr Ala Leu
1 5 10 15

Ala Ala Pro Val Val Glu Glu Arg Gln Asn Cys Ala Thr Leu Trp Gly
20 25 30

Gln Cys Gly Gly Asn Gly Trp Asn Gly Ala Thr Cys Cys Ala Ser Gly
35 40 45

Ser Thr Cys Thr Lys Gln Asn Asp Trp Tyr Ser Gln Cys Leu Pro Gly
50 55 60

Gly Ala Val Thr Thr Pro Gly Thr Thr Thr Lys Pro Thr Ser Thr Ser
65 70 75 80

Thr Ser Thr Ser Thr Ser Ser Arg Ser Thr Ser Thr Ser Gln Gly Gly
85 90 95

Gly Val Ser Ser Ser Thr Ser Ser Pro Pro Val Val Thr Asn Pro Pro

173

100 105 110
Thr Ser Ile Pro Gly Gly Ala Ser Ser Thr Ala Ser Tyr Thr Gly Asn
115 120 125
Pro Phe Ser Gly Val Gln Met Trp Ala Asn Asp Tyr Tyr Arg Ser Glu
130 135 140
Val His Thr Leu Ala Met Pro Ser Leu Thr Gly Ala Met Ala Thr Lys
145 150 155 160
Ala Ala Lys Val Ala Glu Val Pro Ser Tyr Gln Trp Met Asp Arg Asn
165 170 175
Val Thr Val Asp Thr Leu Phe Ser Gly Thr Leu Ala Gln Ile Arg Ala
180 185 190
Ala Asn Gln Ala Gly Ala Ser Pro Pro Tyr Ala Gly Ile Phe Val Val
195 200 205
Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn Gly Glu
210 215 220
Trp Ser Ile Ala Asn Gly Gly Ala Ala Asn Tyr Lys Ala Tyr Ile Lys
225 230 235 240
Arg Ile Arg Glu Leu Ile Ile Gln Tyr Ser Asp Ile Arg Met Leu Leu
245 250 255
Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met Gly Val
260 265 270
Ala Lys Cys Ala Gly Ala Ala Ser Thr Tyr Lys Glu Leu Thr Ile His
275 280 285
Ala Leu Lys Glu Leu Asn Leu Pro Asn Val Ala Met Tyr Leu Asp Ala
290 295 300
Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro Ala Ala
305 310 315 320
Asp Leu Phe Ala Thr Leu Tyr Lys Asp Ala Gly Arg Pro Ala Ala Val
325 330 335
Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser Val Ser
340 345 350
Ser Ala Pro Ala Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu Lys His
355 360 365
Tyr Val Glu Ala Phe Ser Pro Leu Leu Thr Ala Ala Gly Phe Pro Ala
370 375 380
His Phe Ile Thr Asp Thr Gly Arg Ser Gly Lys Gln Pro Thr Gly Gln
385 390 395 400
Leu Glu Trp Gly His Trp Cys Asn Ala Val Gly Thr Gly Phe Gly Gln
405 410 415
Arg Pro Ser Ala Asn Thr Gly His Asp Leu Leu Asp Ala Phe Val Trp
420 425 430

174

Ile Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp Thr Thr Ala Ala
435 440 445

Arg Tyr Asp His Asn Cys Gly Leu Ala Asp Ala Leu Lys Pro Ala Pro
450 455 460

Glu Ala Gly Gln Trp Phe Gln Ala Tyr Phe Glu Gln Leu Leu Thr Asn
465 470 475 480

Ala Asn Pro Pro Phe
485

<210> 81

<211> 487

<212> PRT

<213> Magnaporthe grisea

<400> 81

Met Ala Ser Lys Leu Phe Leu Ala Ala Ala Leu Leu Gln Gly Ala Leu
1 5 10 15

Ser Ser Pro Leu Ala Val Glu Glu Arg Gln Ala Cys Ala Ala Gln Trp
20 25 30

Gly Gln Cys Gly Gly Gln Asp Tyr Thr Gly Pro Thr Cys Cys Gln Ser
35 40 45

Gly Ser Thr Cys Val Val Ser Asn Gln Trp Tyr Ser Gln Cys Leu Pro
50 55 60

Gly Ser Ser Asn Pro Thr Thr Thr Ser Arg Thr Ser Thr Ser Ser Ser
65 70 75 80

Ser Ser Thr Ser Arg Thr Ser Ser Ser Thr Ser Arg Pro Pro Ser Ser
85 90 95

Val Pro Thr Thr Pro Thr Ser Val Pro Pro Thr Ile Thr Thr Thr Pro
100 105 110

Thr Thr Thr Pro Thr Gly Gly Ser Gly Pro Gly Thr Thr Ala Ser Phe
115 120 125

Thr Gly Asn Pro Phe Ala Gly Val Asn Leu Phe Pro Asn Lys Phe Tyr
130 135 140

Ser Ser Glu Val His Thr Leu Ala Ile Pro Ser Leu Thr Gly Ser Leu
145 150 155 160

Val Ala Lys Ala Ser Ala Val Ala Gln Val Pro Ser Phe Gln Trp Leu
165 170 175

Asp Ile Ala Ala Lys Val Glu Thr Leu Met Pro Gly Ala Leu Ala Asp
180 185 190

Val Arg Ala Ala Asn Ala Ala Gly Gly Asn Tyr Ala Ala Gln Leu Val
195 200 205

Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn Gly
210 215 220

Glu Phe Ser Ile Ala Asp Gly Gly Val Val Lys Tyr Lys Ala Tyr Ile

175

225 230 235 240

Asp Ala Ile Arg Lys Gln Leu Leu Ala Tyr Ser Asp Val Arg Thr Ile
245 250 255

Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met Gly
260 265 270

Val Pro Lys Cys Ala Gly Ala Lys Asp Ala Tyr Leu Glu Cys Thr Ile
275 280 285

Tyr Ala Val Lys Gln Leu Asn Leu Pro His Val Ala Met Tyr Leu Asp
290 295 300

Gly Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Leu Gln Pro Ala
305 310 315 320

Ala Asp Leu Phe Gly Lys Leu Tyr Ala Asp Ala Gly Lys Pro Ser Gln
325 330 335

Leu Arg Gly Met Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Asp Leu
340 345 350

Thr Thr Ala Pro Ser Tyr Thr Thr Pro Asn Pro Asn Phe Asp Glu Lys
355 360 365

Lys Tyr Ile Ser Ala Phe Ala Pro Leu Leu Ala Ala Lys Gly Trp Ser
370 375 380

Ala His Phe Ile Ile Asp Gln Gly Arg Ser Gly Lys Gln Pro Thr Gly
385 390 395 400

Gln Lys Glu Trp Gly His Trp Cys Asn Gln Gln Gly Val Gly Phe Gly
405 410 415

Arg Arg Pro Ser Ala Asn Thr Gly Ser Glu Leu Ala Asp Ala Phe Val
420 425 430

Trp Ile Lys Pro Gly Gly Glu Cys Asp Gly Val Ser Asp Pro Thr Ala
435 440 445

Pro Arg Phe Asp His Phe Cys Gly Thr Asp Tyr Gly Ala Met Ser Asp
450 455 460

Ala Pro Gln Ala Gly Gln Trp Phe Gln Lys Tyr Phe Glu Met Leu Leu
465 470 475 480

Thr Asn Ala Asn Pro Pro Leu
485

<210> 82

<211> 421

<212> PRT

<213> Sclerotinia sclerotiorum

<400> 82

Met Gly Leu Lys Asn Val Leu Leu Ala Ala Ala Val Ala Pro Thr
1 5 10 15

Val Tyr Ala Gln Gly Ala Gly Tyr Ser Gln Cys Gly Gly Gln Gly Trp
20 25 30

176

Ser Gly Ala Thr Thr Cys Val Ser Gly Phe Thr Cys Thr Tyr Thr Asn
35 40 45

Glu Tyr Tyr Ser Gln Cys Leu Pro Gly Ser Gly Gly Ala Ser Ser
50 55 60

Ser Arg Pro Thr Thr Thr Ala Pro Thr Thr Ile Val Thr Ser Thr Lys
65 70 75 80

Ala Ser Thr Thr Thr Gly Ser Ser Ala Thr Thr Thr Ala Ala Pro Ala
85 90 95

Ala Gly Asn Pro Phe Val Gly Lys Ala Leu Tyr Val Asn Pro Tyr Tyr
100 105 110

Ala Ser Glu Ile Ser Ala Ser Ala Ile Pro Ser Leu Thr Gly Ala Met
115 120 125

Ala Thr Lys Ala Ala Ala Val Ala Lys Val Pro Thr Phe Phe Trp Leu
130 135 140

Asp Thr Ala Asp Lys Val Pro Thr Met Gly Thr Tyr Leu Ser Asn Ile
145 150 155 160

Arg Ala Leu Asn Lys Ala Gly Ala Asn Pro Pro Val Ala Gly Thr Phe
165 170 175

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn
180 185 190

Gly Glu Tyr Ser Ile Ala Asn Asn Gly Val Ala Asn Tyr Lys Ala Tyr
195 200 205

Ile Asp Ser Ile Val Thr Ile Leu Lys Asn Tyr Ser Asp Thr Ser Val
210 215 220

Ile Leu Ile Ile Val Asp Leu Pro Asn Val Ser Met Tyr Leu Asp Ala
225 230 235 240

Gly His Ala Gly Trp Leu Gly Trp Ser Ala Asn Ile Gly Pro Ala Ala
245 250 255

Gln Leu Phe Gly Gln Val Tyr Lys Ala Ala Gly Ser Pro Ser Gln Val
260 265 270

Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Thr Ser Ser
275 280 285

Ser Cys Pro Ser Tyr Thr Ser Gly Asp Ser Asn Cys Asn Glu Lys Leu
290 295 300

Tyr Ile Asn Ala Leu Ala Pro Leu Leu Thr Ala Gln Gly Phe Pro Ala
305 310 315 320

His Phe Ile Met Asp Thr Gly Arg Asn Gly Val Gln Pro Thr Ala Gln
325 330 335

Gln Ala Trp Gly Asp Trp Cys Asn Val Ile Gly Thr Gly Phe Gly Val
340 345 350

Arg Pro Thr Thr Asn Thr Gly Asp Ala Leu Glu Asp Ala Phe Val Trp

355 360 365

Val Lys Pro Gly Gly Glu Ala Asp Gly Thr Ser Asn Thr Thr Ala Ala
370 375 380

Arg Tyr Asp Phe His Cys Gly Leu Ser Asp Ala Leu Gln Pro Ala Pro
385 390 395 400

Glu Ala Gly Thr Trp Phe Gln Ala Tyr Phe Ala Gln Leu Leu Thr Asn
405 410 415

Ala Asn Pro Ser Phe
420

<210> 83

<211> 449

<212> PRT

<213> Phaeosphaeria nodorum

<400> 83

Met Lys Ala Val Ser Phe Leu Ala Val Ala Ala Leu Ala Pro Ala Ile
1 5 10 15

Lys Ala Gln Ala Ser Leu Tyr Gln Gln Cys Gly Gly Thr Gly Phe Ser
20 25 30

Gly Ser Thr Thr Cys Val Ser Gly Ala Tyr Cys Ser Lys Val Asn Asp
35 40 45

Ser Ala Thr Ser Ala Ala Pro Ala Pro Thr Thr Phe Lys Thr Ser Lys
50 55 60

Thr Val Gly Ser Pro Ala Thr Gly Ser Ser Thr Thr Gly Ser Ser Ala
65 70 75 80

Thr Gly Thr Ala Ser Pro Gly Asp Gly Ser Asn Pro Leu Lys Gly Lys
85 90 95

Asn Phe Tyr Ala Asn Ser Tyr Tyr Ala Ser Glu Ile Asn Asn Leu Ala
100 105 110

Ala Pro Ser Leu Val Ala Ala Gly Asn Ala Ala Leu Ala Ala Lys Ala
115 120 125

Ser Asn Val Ala Lys Val Gly Thr Phe Tyr Trp Leu Asp Val Arg Ala
130 135 140

Lys Val Pro Ile Ile Ser Thr Phe Ala Lys Asp Val Gln Lys Arg Asn
145 150 155 160

Ala Ala Gly Ala Asn Glu Val Leu Pro Leu Val Val Tyr Asp Leu Pro
165 170 175

Glu Arg Asp Cys Ala Ala Leu Ala Ser Asn Gly Glu Leu Ser Leu Ala
180 185 190

Asn Asn Gly Thr Ala Leu Tyr Gln Glu Tyr Ile Asp Met Ile Ala Ala
195 200 205

Gln Ile Lys Gln Phe Pro Asp Val Thr Phe Leu Leu Val Val Glu Pro
210 215 220

178

Asp Ser Leu Ala Asn Leu Val Thr Asn Leu Asn Val Ala Lys Cys Ala
225 230 235 240

Asn Ala Ala Thr Ala Tyr Lys Thr Leu Thr Ala Tyr Ala Ile Lys Thr
245 250 255

Leu Asn Leu Lys Asn Val Ile Met Tyr Leu Asp Ala Gly His Ala Gly
260 265 270

Trp Leu Gly Trp Thr Ala Asn Ile Glu Pro Ala Ala Glu Leu Phe Gly
275 280 285

Ala Leu Tyr Lys Ser Ala Gly Ser Pro Ala Ala Val Arg Gly Leu Val
290 295 300

Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser Ile Ala Thr Cys Pro Ser
305 310 315 320

Tyr Thr Gln Gly Asn Thr Asn Cys Asp Glu Lys Arg Tyr Val Asn Ala
325 330 335

Leu Ala Pro Leu Leu Val Lys Asn Gly Phe Pro Ala His Phe Leu Thr
340 345 350

Asp Thr Gly Arg Asn Gly Val Gln Pro Thr Lys Gln Gln Ala Trp Gly
355 360 365

Asp Trp Cys Asn Val Ile Gly Thr Gly Phe Gly Ile Arg Pro Ser Ser
370 375 380

Thr Thr Asp Asp Pro Leu Leu Asp Ala Tyr Val Trp Val Lys Pro Gly
385 390 395 400

Gly Glu Gly Asp Gly Thr Ser Asp Thr Ser Ala Val Arg Tyr Asp Ala
405 410 415

His Cys Gly Tyr Ala Asp Ala Leu Lys Pro Ala Pro Glu Ala Gly Ser
420 425 430

Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu Ser Asn Ala Ser Pro Ala
435 440 445

Phe

<210> 84

<211> 438

<212> PRT

<213> Agaricus bisporus

<400> 84

Met Phe Lys Phe Ala Ala Leu Leu Ala Leu Ala Ser Leu Val Pro Gly
1 5 10 15

Phe Val Gln Ala Gln Ser Pro Val Trp Gly Gln Cys Gly Gly Asn Gly
20 25 30

Trp Thr Gly Pro Thr Thr Cys Ala Ser Gly Ser Thr Cys Val Lys Gln
35 40 45

Asn Asp Phe Tyr Ser Gln Cys Leu Pro Asn Asn Gln Ala Pro Pro Ser
50 55 60

Thr Thr Thr Gln Pro Gly Thr Thr Pro Pro Ala Thr Thr Thr Ser Gly
65 70 75 80

Gly Thr Gly Pro Thr Ser Gly Ala Gly Asn Pro Tyr Thr Gly Lys Thr
85 90 95

Val Trp Leu Ser Pro Phe Tyr Ala Asp Glu Val Ala Gln Ala Ala Ala
100 105 110

Asp Ile Ser Asn Pro Ser Leu Ala Thr Lys Ala Ala Ser Val Ala Lys
115 120 125

Ile Pro Thr Phe Thr Trp Phe Asp Thr Val Ala Lys Val Pro Asp Leu
130 135 140

Gly Gly Tyr Leu Ala Asp Ala Gln Ser Lys Asn Gln Leu Val Gln Ile
145 150 155 160

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Leu Ala Ser Asn
165 170 175

Gly Glu Phe Ser Leu Ala Asn Asp Gly Leu Asn Lys Tyr Lys Asn Tyr
180 185 190

Val Asp Gln Ile Ala Ala Gln Ile Lys Gln Phe Pro Asp Val Ser Val
195 200 205

Val Ala Val Ile Glu Pro Asp Ser Leu Ala Asn Leu Val Thr Asn Leu
210 215 220

Asn Val Gln Lys Cys Ala Asn Ala Gln Ser Ala Tyr Lys Glu Gly Val
225 230 235 240

Ile Tyr Ala Ile Gln Lys Leu Asp Ala Val Gly Val Thr Met Tyr Ile
245 250 255

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Leu Ser Pro
260 265 270

Ala Ala Gln Leu Phe Ala Gln Ile Tyr Arg Asp Ala Gly Ser Pro Arg
275 280 285

Asn Leu Arg Gly Ile Ala Thr Asn Val Ala Asn Phe Asn Ala Leu Arg
290 295 300

Ala Ser Ser Pro Asp Pro Ile Thr Gln Gly Asn Ser Asn Tyr Asp Glu
305 310 315 320

Ile His Tyr Ile Glu Ala Leu Ala Pro Met Leu Ser Asn Ala Gly Phe
325 330 335

Pro Ala His Phe Ile Val Asp Gln Gly Arg Ser Gly Val Gln Asn Ile
340 345 350

Arg Asp Gln Trp Gly Asp Trp Cys Asn Val Lys Gly Ala Gly Phe Gly
355 360 365

Gln Arg Pro Thr Thr Asn Thr Gly Ser Ser Leu Ile Asp Ala Ile Val
370 375 380

180

Trp Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp Ser Ser Ser
385 390 395 400

Pro Arg Phe Asp Ser His Cys Ser Leu Ser Asp Ala His Gln Pro Ala
405 410 415

Pro Glu Ala Gly Thr Trp Phe Gln Ala Tyr Phe Glu Thr Leu Val Ala
420 425 430

Asn Ala Asn Pro Ala Leu
435

<210> 85

<211> 443

<212> PRT

<213> Lentinula edodes

<400> 85

Met Lys Ile Thr Ser Thr Gly Leu Leu Ala Leu Ser Ser Leu Leu Pro
1 5 10 15

Phe Ala Leu Gly Gln Ser Gln Leu Tyr Ala Gln Cys Gly Gly Ile Gly
20 25 30

Trp Ser Gly Ala Thr Thr Cys Val Ser Gly Ala Thr Cys Thr Val Val
35 40 45

Asn Ala Tyr Tyr Ser Gln Cys Leu Pro Gly Ser Ala Ser Ala Pro Pro
50 55 60

Thr Ser Thr Ser Ser Ile Gly Thr Gly Thr Thr Thr Ser Ser Ala Pro
65 70 75 80

Gly Ser Thr Gly Thr Thr Thr Pro Ala Ala Gly Asn Pro Phe Thr Glu
85 90 95

Gln Ile Tyr Leu Ser Pro Tyr Tyr Ala Asn Glu Ile Ala Ala Ala Val
100 105 110

Thr Gln Ile Ser Asp Pro Thr Thr Ala Ala Ala Ala Ala Lys Val Ala
115 120 125

Asn Ile Pro Thr Phe Ile Trp Leu Asp Gln Val Ala Lys Val Pro Asp
130 135 140

Leu Gly Thr Tyr Leu Ala Asp Ala Ser Ala Lys Gln Lys Ser Glu Gly
145 150 155 160

Lys Asn Tyr Leu Val Gln Ile Val Val Tyr Asp Leu Pro Asp Arg Asp
165 170 175

Cys Ala Ala Leu Ala Ser Asn Gly Glu Phe Thr Ile Ala Asp Asn Gly
180 185 190

Glu Ala Asn Tyr His Asp Tyr Ile Asp Gln Ile Val Ala Gln Ile Lys
195 200 205

Gln Tyr Pro Asp Val His Val Val Ala Val Ile Glu Pro Asp Ser Leu
210 215 220

181

Ala Asn Leu Val Thr Asn Leu Ser Val Ala Lys Cys Ala Asn Ala Gln
225 230 235 240

Thr Thr Tyr Leu Glu Cys Val Thr Tyr Ala Met Gln Gln Leu Ser Ala
245 250 255

Val Gly Val Thr Met Tyr Leu Asp Ala Gly His Ala Gly Trp Leu Gly
260 265 270

Trp Pro Ala Asn Leu Ser Pro Ala Ala Gln Leu Phe Thr Ser Leu Tyr
275 280 285

Ser Asn Ala Gly Ser Pro Ser Gly Val Arg Gly Leu Ala Thr Asn Val
290 295 300

Ala Asn Tyr Asn Ala Leu Val Ala Thr Thr Pro Asp Pro Ile Thr Gln
305 310 315 320

Gly Asp Pro Asn Tyr Asp Glu Met Leu Tyr Ile Glu Ala Leu Ala Pro
325 330 335

Leu Leu Gly Ser Phe Pro Ala His Phe Ile Val Asp Gln Gly Arg Ser
340 345 350

Gly Val Gln Asp Ile Arg Gln Gln Trp Gly Asp Trp Cys Asn Val Leu
355 360 365

Gly Ala Gly Phe Gly Thr Gln Pro Thr Thr Asn Thr Gly Ser Ser Leu
370 375 380

Ile Asp Ser Ile Val Trp Val Lys Pro Gly Gly Glu Cys Asp Gly Thr
385 390 395 400

Ser Asn Thr Ser Ser Pro Arg Tyr Asp Ala His Cys Gly Leu Pro Asp
405 410 415

Ala Thr Pro Asn Ala Pro Glu Ala Gly Thr Trp Phe Gln Ala Tyr Phe
420 425 430

Glu Thr Leu Val Glu Lys Ala Asn Pro Pro Leu
435 440

<210> 86

<211> 460

<212> PRT

<213> Coniophora puteana

<400> 86

Met Phe Lys Phe Ala Ala Leu Ser Ala Phe Val Ala Leu Val Pro Leu
1 5 10 15

Leu Val Asn Ala Gln Val Ala Ala Tyr Gly Gln Cys Gly Gly Gln Asp
20 25 30

Trp Thr Gly Ala Thr Ala Cys Ala Ser Gly Thr Ala Cys Thr Lys Val
35 40 45

Asn Asp Tyr Tyr Tyr Gln Cys Leu Pro Gly Ser Ser Gly Ser Ser Val
50 55 60

Ser Gly Gly Ser Gly Ser Gly Ser Thr Ser Ala Pro Ser Pro Thr Ser

182

65 70 75 80

Thr Val Pro Thr Ser Thr Ser Ser Ala Ser Thr Ala Pro Ser Ser Thr
85 90 95

Ser Thr Ser Ser Ala Ala Ser Ser Asp Asn Pro Tyr Thr Gly Tyr Gln
100 105 110

Ile Phe Leu Asn Pro Glu Tyr Ala Ser Glu Val Gln Ala Ala Ile Pro
115 120 125

Ser Ile Thr Asp Ser Ala Val Ala Ala Lys Ala Leu Lys Val Ala Glu
130 135 140

Val Pro Val Phe Phe Trp Leu Asp Gln Val Ala Lys Val Pro Asp Leu
145 150 155 160

Glu Thr Tyr Leu Ala Ala Ala Asp Lys Gln Gly Lys Ser Ser Gly Gln
165 170 175

Lys Gln Leu Leu Gln Ile Val Val Tyr Asp Leu Pro Asp Arg Asp Cys
180 185 190

Ala Ala Asn Ala Ser Asn Gly Glu Phe Ser Ile Ser Asp Asp Gly Gln
195 200 205

Ala Lys Tyr Glu Asn Tyr Ile Asp Gln Ile Val Ala Ile Val Lys Lys
210 215 220

Tyr Pro Asp Val Arg Val Val Ala Val Val Glu Pro Asp Ser Met Gly
225 230 235 240

Asn Leu Val Thr Asn Met Asp Leu Pro Lys Cys Ser Ala Ala Ala Pro
245 250 255

Thr Tyr Lys Thr Cys Ile Asn Tyr Ala Ile Ala Gln Leu Ser Ser Ala
260 265 270

Gly Val Tyr Met Tyr Val Asp Ala Gly His Ala Gly Trp Leu Gly Trp
275 280 285

Pro Asn Asn Leu Ala Pro Ala Ala Gln Leu Phe Gly Glu Leu Tyr Glu
290 295 300

Thr Ser Gly Lys Ser Ala Tyr Phe Arg Gly Leu Ala Thr Asn Val Ala
305 310 315 320

Asn Tyr Asn Ala Leu Asn Thr Ser Ser Pro Asp Pro Cys Thr Gln Asn
325 330 335

Ala Pro Asn Tyr Asp Glu Met Leu Tyr Ile Asn Ala Leu Ser Pro Leu
340 345 350

Leu Gln Gln Gln Gly Phe Ser Ala Gln Phe Ile Val Asp Gln Gly Arg
355 360 365

Ser Gly Val Gln Asn Ile Arg Asn Ala Trp Gly Asp Trp Cys Asn Ile
370 375 380

Lys Gly Ala Gly Phe Gly Ile Arg Pro Thr Thr Asp Thr Gly Ser Pro
385 390 395 400

183

Leu Ile Asp Ser Ile Val Trp Val Lys Pro Gly Gly Glu Cys Asp Gly
405 410 415

Thr Ser Asn Ser Ser Ala Pro Arg Tyr Asp Ser Thr Cys Ser Leu Ser
420 425 430

Asp Ser Leu Gln Pro Ala Pro Glu Ala Gly Thr Trp Phe Gln Gln Tyr
435 440 445

Phe Glu Ala Leu Val Thr Asn Ala Val Pro Ser Leu
450 455 460

<210> 87
<211> 442
<212> PRT
<213> Volvariella volvacea

<400> 87

Met Ser Arg Phe Ser Ala Leu Thr Ala Leu Leu Ser Leu Pro Leu
1 5 10 15

Leu Ala Ile Ala Gln Ser Pro Leu Tyr Gly Gln Cys Gly Gly Asn Gly
20 25 30

Trp Thr Gly Pro Lys Thr Cys Val Ser Gly Ala Thr Cys Thr Val Ile
35 40 45

Asn Asp Trp Tyr Trp Gln Cys Leu Pro Gly Asn Gly Pro Thr Ser Ser
50 55 60

Ser Pro Thr Ser Thr Pro Thr Thr Thr Thr Thr Gly Gly Pro Gln
65 70 75 80

Pro Thr Val Pro Ala Ala Gly Asn Pro Tyr Thr Gly Tyr Glu Ile Tyr
85 90 95

Leu Ser Pro Tyr Tyr Ala Ala Glu Ala Gln Ala Ala Ala Gln Ile
100 105 110

Ser Asp Ala Thr Gln Lys Ala Lys Ala Leu Lys Val Ala Gln Ile Pro
115 120 125

Thr Phe Thr Trp Phe Asp Val Ile Ala Lys Thr Ser Thr Leu Gly Asp
130 135 140

Tyr Leu Ala Glu Ala Ser Ala Leu Gly Lys Ser Ser Gly Lys Lys Tyr
145 150 155 160

Leu Val Gln Ile Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala
165 170 175

Leu Ala Ser Asn Gly Glu Phe Ser Ile Ala Asn Asn Gly Leu Asn Asn
180 185 190

Tyr Lys Gly Tyr Ile Asp Gln Leu Val Ala Gln Ile Lys Lys Tyr Pro
195 200 205

Asp Val Arg Val Val Ala Val Ile Glu Pro Asp Ser Leu Ala Asn Leu
210 215 220

Val Thr Asn Leu Asn Val Ser Lys Cys Ala Asn Ala Gln Thr Ala Tyr

184

225 230 235 240

Lys Ala Gly Val Thr Tyr Ala Leu Gln Gln Leu Asn Ser Val Gly Val
245 250 255

Tyr Met Tyr Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala
260 265 270

Asn Leu Asn Pro Ala Ala Gln Leu Phe Ser Gln Leu Tyr Arg Asp Ala
275 280 285

Gly Ser Pro Gln Tyr Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr
290 295 300

Asn Ala Leu Ser Ala Ser Ser Pro Asp Pro Val Thr Gln Gly Asn Pro
305 310 315 320

Asn Tyr Asp Glu Leu His Tyr Ile Asn Ala Leu Ala Pro Ala Leu Gln
325 330 335

Ser Gly Gly Phe Pro Ala His Phe Ile Val Asp Gln Gly Arg Ser Gly
340 345 350

Val Gln Asn Ile Arg Gln Gln Trp Gly Asp Trp Cys Asn Val Lys Gly
355 360 365

Ala Gly Phe Gly Gln Arg Pro Thr Leu Ser Thr Gly Ser Ser Leu Ile
370 375 380

Asp Ala Ile Val Trp Ile Lys Pro Gly Gly Glu Cys Asp Gly Thr Thr
385 390 395 400

Asn Thr Ser Ser Pro Arg Tyr Asp Ser His Cys Gly Leu Ser Asp Ala
405 410 415

Thr Pro Asn Ala Pro Glu Ala Gly Gln Trp Phe Gln Ala Tyr Phe Glu
420 425 430

Thr Leu Val Arg Asn Ala Ser Pro Pro Leu
435 440

<210> 88
<211> 389
<212> PRT
<213> Coniophora puteana

<400> 88

Met Phe Ser Pro Val Val Leu Gly Ala Leu Ala Ala Leu Leu Pro Ala
1 5 10 15

Ala Val Gln Ala Met Pro Ala Ser Thr Gln Ala Arg Ala Ala Asp Ala
20 25 30

Thr Ala Asn Pro Tyr Thr Gly Tyr Thr Ile Phe Lys Asn Pro Glu Tyr
35 40 45

Val Ala Glu Val Gln Ala Ala Val Gln Gln Ile Ser Asp Ser Ser Leu
50 55 60

Ala Ser Ala Ala Ala Gly Val Glu Asp Val Pro Val Phe Phe Trp Leu
65 70 75 80

Asp Gln Val Ala Lys Val Pro Asn Leu Thr Thr Tyr Leu Ala Ala Ala
85 90 95

Asp Ala Glu Ala Lys Ser Ser Gly Ser Gln Gln Leu Phe Gln Ile Val
100 105 110

Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn Gly
115 120 125

Glu Phe Ser Ile Ser Asp Asn Gly Gln Ala Asn Tyr Glu Asn Tyr Ile
130 135 140

Asp Gln Ile Val Ala Ser Ile Lys Gln Tyr Pro Asp Val Arg Val Val
145 150 155 160

Ala Val Val Glu Pro Asp Ser Met Ala Asn Leu Val Thr Asn Leu Ser
165 170 175

Val Gln Lys Cys Ala Asp Ala Glu Ser Thr Tyr Lys Thr Cys Val Ala
180 185 190

Tyr Ala Ile Glu Gln Leu Ala Thr Val Gly Val Tyr Met Tyr Leu Asp
195 200 205

Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Leu Ser Pro Ala
210 215 220

Ala Glu Leu Phe Ala Gln Met Tyr Ser Thr Thr Gly Ser Ser Pro Tyr
225 230 235 240

Phe Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ser Leu Thr Thr
245 250 255

Asp Ser Pro Asp Pro Ile Thr Ser Gly Asp Ser Asn Tyr Asp Glu Leu
260 265 270

Leu Tyr Ile Glu Ala Leu Ser Pro Leu Leu Val Asp Asn Gly Phe Pro
275 280 285

Ala Gln Phe Ile Val Glu Gln Ala Arg Ser Gly Val Gln Asn Ile Arg
290 295 300

Ser Ala Trp Gly Asp Trp Cys Asn Val Lys Gly Ala Gly Phe Gly Leu
305 310 315 320

Arg Pro Ser Thr Asp Thr Pro Ser Ser Leu Ile Asp Ser Ile Val Trp
325 330 335

Val Lys Pro Gly Gly Glu Ala Asp Gly Thr Ser Asn Ser Ser Ala Ala
340 345 350

Arg Tyr Asp Tyr His Cys Ser Leu Ser Asp Ala Leu Gln Pro Ala Pro
355 360 365

Glu Ala Gly Thr Trp Phe Gln Thr Tyr Phe Glu Asp Leu Val Ser Gly
370 375 380

Ala Asn Pro Ala Phe
385

<210> 89

186

<211> 454
 <212> PRT
 <213> Coprinopsis cinerea okayama

<400> 89

Met Leu Lys Gly Ser Lys Phe Phe Ala Leu Ser Leu Ala Leu Leu Pro
 1 5 10 15

Ala Leu Val Gln Ala Gln Arg Pro Leu Tyr Ala Gln Cys Gly Gly Thr
 20 25 30

Gly Trp Thr Gly Glu Thr Thr Cys Val Ser Gly Ala Val Cys Glu Val
 35 40 45

Ile Asn Gln Trp Tyr His Gln Cys Leu Pro Gly Ser Asn Gln Pro Gln
 50 55 60

Pro Pro Val Thr Thr Gln Pro Pro Val Val Val Pro Thr Thr Ser Gln
 65 70 75 80

Pro Pro Val Val Val Pro Thr Asn Pro Pro Gly Gly Thr Pro Val Pro
 85 90 95

Ser Thr Gly Asn Pro Phe Glu Gly Tyr Asp Ile Tyr Leu Ser Pro Tyr
 100 105 110

Tyr Ala Glu Glu Val Glu Ala Ala Ala Ala Met Ile Asp Asp Pro Val
 115 120 125

Leu Lys Ala Lys Ala Leu Lys Val Lys Glu Ile Pro Thr Phe Ile Trp
 130 135 140

Phe Asp Val Val Arg Lys Thr Pro Asp Leu Gly Arg Tyr Leu Ala Asp
 145 150 155 160

Ala Thr Ala Ile Gln Gln Arg Thr Gly Arg Lys Gln Leu Val Gln Ile
 165 170 175

Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn
 180 185 190

Gly Glu Phe Ser Leu Ala Asp Gly Gly Met Glu Lys Tyr Lys Asp Tyr
 195 200 205

Val Asp Arg Leu Ala Ser Glu Ile Arg Lys Tyr Pro Asp Val Arg Ile
 210 215 220

Val Ala Val Ile Glu Pro Asp Ser Leu Ala Asn Met Val Thr Asn Met
 225 230 235 240

Asn Val Ala Lys Cys Arg Gly Ala Glu Ala Ala Tyr Lys Glu Gly Val
 245 250 255

Ile Tyr Ala Leu Arg Gln Leu Ser Ala Leu Gly Val Tyr Ser Tyr Val
 260 265 270

Asp Ala Gly His Ala Gly Trp Leu Gly Trp Asn Ala Asn Leu Ala Pro
 275 280 285

Ser Ala Arg Leu Phe Ala Gln Ile Tyr Lys Asp Ala Gly Arg Ser Ala
 290 295 300

187

Phe Ile Arg Gly Leu Ala Thr Asn Val Ser Asn Tyr Asn Ala Leu Ser
305 310 315 320

Ala Thr Thr Arg Asp Pro Val Thr Gln Gly Asn Asp Asn Tyr Asp Glu
325 330 335

Leu Arg Phe Ile Asn Ala Leu Ala Pro Leu Leu Arg Asn Glu Gly Trp
340 345 350

Asp Ala Lys Phe Ile Val Asp Gln Gly Arg Ser Gly Val Gln Asn Ile
355 360 365

Arg Gln Glu Trp Gly Asn Trp Cys Asn Val Tyr Gly Ala Gly Phe Gly
370 375 380

Met Arg Pro Thr Leu Asn Thr Pro Ser Ser Ala Ile Asp Ala Ile Val
385 390 395 400

Trp Ile Lys Pro Gly Gly Glu Ala Asp Gly Thr Ser Asp Thr Ser Ala
405 410 415

Pro Arg Tyr Asp Thr His Cys Gly Lys Ser Asp Ser His Lys Pro Ala
420 425 430

Pro Glu Ala Gly Thr Trp Phe Gln Glu Tyr Phe Val Asn Leu Val Lys
435 440 445

Asn Ala Asn Pro Pro Leu
450

<210> 90
<211> 361
<212> PRT
<213> Moniliophthora perniciosa

<400> 90

Ile Pro Gly Ser Asp Pro Gly Asn Pro Gly Pro Thr Ser Ser Ser Thr
1 5 10 15

Leu Ser Ser Thr Ala Ala Pro Pro Thr Asn Thr Gln Ser Pro Val Glu
20 25 30

Asp Asn Pro Tyr Thr Gly Tyr Thr Ile Tyr Leu Ser Pro Tyr Tyr Ala
35 40 45

Asp Glu Ile Asp Ala Ala Ala Ala Lys Ile Thr Asp Pro Thr Leu Lys
50 55 60

Val Gln Ala Leu Lys Val Lys Glu Ile Pro Thr Phe Ile Trp Phe Asp
65 70 75 80

Thr Thr Ala Lys Leu Ser Thr Leu Glu Pro Tyr Leu Lys Asp Ala Ser
85 90 95

Ala Lys Gly Lys Ala Glu Gly Lys Lys Tyr Leu Leu Gln Ile Val Val
100 105 110

Tyr Thr Leu Pro Glu Arg Asp Cys Ala Ala Leu Ala Ser Asn Gly Glu
115 120 125

Leu Ser Ile Asp Asn Gly Gly Glu Val Lys Ser Arg Glu Tyr Ile Asp
130 135 140

Thr Met Val Ala Thr Ile Lys Lys Tyr Pro Asp Val Arg Val Val Ala
145 150 155 160

Val Val Glu Pro Asp Ser Leu Ala Asn Leu Val Thr Asn Leu Asn Val
165 170 175

Gln Lys Cys Ser Lys Ala Gln Thr Ile Tyr Lys Thr Ser Thr Gln Tyr
180 185 190

Ala Leu Lys Gln Leu Asp Thr Ala Gly Val Tyr Met Tyr Leu Asp Ala
195 200 205

Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Leu Thr Pro Thr Ala
210 215 220

Gln Leu Phe Gln Gln Val Trp Gln Asp Ala Gly Ser Pro Lys Phe Val
225 230 235 240

Arg Gly Leu Ala Thr Asn Val Ala Asn Phe Asn Ala Leu Arg Ala Ala
245 250 255

Ser Pro Asp Pro Val Thr Ser Gln Asn Pro Asn Tyr Asp Glu Ile His
260 265 270

Tyr Ile Glu Gly Arg Ala Gly Gln Gln Asn Leu Arg Lys Glu Trp Gly
275 280 285

Asp Trp Cys Asn Val Lys Gly Ala Gly Phe Gly Thr Arg Pro Thr Thr
290 295 300

Asn Thr Gly Ser Ser Leu Ile Asp Ser Ile Val Trp Val Lys Pro Gly
305 310 315 320

Gly Glu Ser Ala Arg Phe Asp Ala Lys Cys Val Ser Ala Ser Ser His
325 330 335

Val Pro Ala Pro Glu Ala Gly Thr Trp Phe Gln Glu Tyr Phe Glu Ala
340 345 350

Leu Val Arg Asn Ala Asn Pro Ala Leu
355 360

<210> 91

<211> 460

<212> PRT

<213> *Piromyces rhizinflatus*

<400> 91

Met Lys Phe Ser Thr Leu Ile Gly Thr Leu Phe Ala Thr Gly Ala Leu
1 5 10 15

Ala Ser Ser Cys His Arg Asp Tyr Pro Cys Cys Asn Asp Cys Asn Val
20 25 30

Val Tyr Gln Asp Trp Glu Arg Asp Trp Gly Val Leu Asn Gly Gln Glu
35 40 45

Trp Cys Phe Ile Asp Lys Asn Arg Cys Asn Gly Gly Gly Tyr Cys Lys

189

50 55 60

Phe Glu Ser Leu Gly Tyr Pro Cys Cys Asn Gly Cys Asp Val Tyr Tyr
65 70 75 80

Thr Asp Asn Asp Gly Arg Trp Gly Val Glu Asn Gly Asn Trp Cys Gly
85 90 95

Ile Arg Asp Asp Lys Cys Asn Gly Tyr Gln Gln Pro Arg Thr Thr Thr
100 105 110

Thr Thr Arg Thr Thr Thr Arg Thr Thr Thr Thr Gln Arg Pro Val Gln
115 120 125

Thr Asn Val Ser Asp Asn Phe Phe Glu Asn Thr Leu Tyr Ser Asn Phe
130 135 140

Lys Phe Gln Gly Glu Val Gln Ser Ser Ile Gln Lys Leu Ser Gly Asp
145 150 155 160

Met Ala Lys Lys Ala Glu Lys Val Lys Tyr Val Pro Thr Ala Val Trp
165 170 175

Leu Ala Trp Glu Gly Ala Pro Arg Glu Val Pro Gln Tyr Leu Asp Asp
180 185 190

Ala Gly Ser Lys Thr Val Val Phe Val Leu Tyr Met Ile Pro Thr Arg
195 200 205

Asp Cys Asn Ala Asn Ala Ser Val Gly Gly Ser Ala Thr Leu Glu Lys
210 215 220

Tyr Lys Gly Tyr Ile Asp Asn Ile Tyr Asn Thr Phe Asn Gln Tyr Pro
225 230 235 240

Asn Ser Lys Ile Val Met Ile Leu Glu Pro Asp Thr Ile Gly Asn Leu
245 250 255

Val Thr Ala Asn Asn Ala Asn Cys Met Asn Val Gln Asn Leu His Lys
260 265 270

Gln Gly Leu Ala Tyr Ala Ile Ser Lys Phe Gly Thr Gln Lys Asn Val
275 280 285

Arg Val Tyr Leu Asp Ala Ala His Gly Ala Trp Leu Ser Ser His Ala
290 295 300

Asp Lys Thr Ala Gln Val Ile Lys Glu Ile Leu Asn Asn Ala Gly Ser
305 310 315 320

Gly Lys Leu Arg Gly Ile Thr Thr Asn Val Ser Asn Tyr Gln Thr Val
325 330 335

Asn Asp Glu Tyr Ser Tyr Gln Met Arg Leu Asn Ser Ala Leu Gln Asn
340 345 350

Leu Gly Val Arg Asp Leu His Tyr Ile Ile Asp Thr Ser Arg Asn Gly
355 360 365

Ala Asn Ile Ala Gln Gln Phe Asn Gln Ser Gly Thr Trp Cys Asn Phe
370 375 380

Lys Gly Ala Gly Leu Gly Ala Arg Pro Gln Ala Asn Pro Asp Ser Ser
385 390 395 400

Lys Pro Leu Leu Asp Ala Tyr Met Trp Ile Lys Thr Pro Gly Glu Ala
405 410 415

Asp Gly Ser Ser Ser Gly Ser Arg Ala Asp Pro Val Cys Gly Arg Trp
420 425 430

Asp Ser Leu Gln Gly Ala Pro Asp Ala Gly Ser Trp Phe His Asp Tyr
435 440 445

Phe Val Met Leu Leu Gln Asn Ala Asn Pro Pro Phe
450 455 460

<210> 92

<211> 518

<212> PRT

<213> Piromyces rhizinflatus

<400> 92

Met Lys Phe Ile Val Cys Ala Ser Ile Leu Ser Leu Gly Leu Phe Lys
1 5 10 15

Phe Ala Asp Ala Ala Cys Ala Gly Pro Tyr Ala Gln Cys Gly Gly Asn
20 25 30

Asn Phe Asn Gly Glu Asn Cys Cys Gln Ser Gly Tyr Lys Cys Val Ala
35 40 45

Ile Asn Glu Trp Tyr Ser Gln Cys Gln Glu Gly Ala Ala Glu Pro Glu
50 55 60

Pro Val Pro Gln Ser Ser Ala Ala Asp Asp Gln Trp Asn Asn Asn Asn
65 70 75 80

Gln Trp Asn Asn Asn Gln Gln Asn Asn Gln Trp Asn Asn Pro Trp Asp
85 90 95

Asn Asn Asn Asn Gln Trp Asn Asn Asn Gln Trp Asn Asn Asn Gln Gln
100 105 110

Asn Asn Gln Trp Asn Asn Asn Gln Gln Gln Asn Asn Gln Trp Asp Asn
115 120 125

Asn Asn Gln Trp Asn Asn Asn Gln Trp Asn Asn Asn Gln Gln Gln Asn
130 135 140

Asn Gln Gln Gln Asn Asn Gln Gln Gln Asn Asn Gln Trp Asn Asn Asn
145 150 155 160

Asn Gln Trp Asn Asn Asn Gln Gln Gln Asn Asn Gln Gln Gln Asn Asn
165 170 175

Gln Ala Pro Ala Gln Ser Asn Gly Gly Gly Ala Ser Gly Ser Ser Gln
180 185 190

Asn Phe Phe Thr Asn Glu Ile Tyr Ala Asn Pro Arg Phe Ile Glu Glu
195 200 205

Ile Asp Ser Ser Ile Pro Lys Leu Thr Pro Glu Leu Ala Ala Lys Ala

191

210 215 220

Glu Lys Val Lys Gln Val Pro Thr Ala Val Trp Leu Ala Trp Asp Gly
225 230 235 240

Ala Pro Gly Glu Val Glu Gly His Leu Lys Ala Ala Gly Ser Lys Thr
245 250 255

Val Val Phe Ile Leu Tyr Met Ile Pro Thr Arg Asp Cys Asn Ser Asn
260 265 270

Ala Ser Ala Gly Gly Ala Ser Asp Leu Asn Lys Tyr Lys Gly Tyr Val
275 280 285

Asp Asp Ile Ala Gly Thr Ile Lys Ser His Pro Glu Ser Lys Val Val
290 295 300

Met Ile Val Glu Pro Asp Thr Leu Gly Asn Leu Val Thr Gly Ser Ser
305 310 315 320

Glu Ala Cys Lys Asn Val His Ser Leu His Lys Ser Ala Leu Ser Tyr
325 330 335

Ala Val Asn Val Phe Gly Ala Met Ser Asn Val Ser Val Tyr Leu Asp
340 345 350

Ala Ala His Gly Lys Trp Leu Gly Gly Ser Thr Asp Lys Val Ala Ser
355 360 365

Val Leu Lys Glu Ile Leu Asp Asn Ala Pro Asn Gly Lys Ile Arg Gly
370 375 380

Leu Ser Thr Asn Val Ser Asn Tyr Gln Pro Leu Ser Ser Glu Ala Gln
385 390 395 400

Tyr His Gln Gly Leu Ser Ser Ala Leu Ala Ala Val Gly Tyr Pro Gly
405 410 415

Met Lys Phe Val Val Asp Thr Gly Arg Asn Gly Val Asp Val Ser Ser
420 425 430

Thr Phe Ser Ile Asn Glu Thr Trp Cys Asn Phe Val Gly Thr Gly Phe
435 440 445

Gly Glu Arg Pro Gln Gly Asn Pro Ser Gly Tyr Pro Leu Leu Asp Ala
450 455 460

Phe Met Trp Leu Lys Thr Pro Gly Glu Ala Asp Gly Ser Ala Thr Gly
465 470 475 480

Ser Arg Ala Asp Pro Val Cys Ala Arg Gln Asp Ser Leu Gln Gly Ala
485 490 495

Pro Asp Ala Gly Gln Trp Phe His Glu Tyr Phe Val Gln Leu Leu Glu
500 505 510

Asn Ala Lys Pro Gly Phe
515

<210> 93
<211> 376
<212> PRT

192

<213> Piromyces rhizinflatus

<400> 93

Glu Thr Leu Pro Gln Gln Ser Asn Ser Ala Lys Thr Leu Pro Gln Gln
1 5 10 15

Ser Asp Ser Ala Lys Thr Ile Pro Gln Pro Thr Ser Ala Glu Ser Gln
20 25 30

Thr Ser Lys Thr Leu Pro Gln Thr Gly Gly Ser Gly Asn Gly Ser Ser
35 40 45

Gln Asn Phe Phe Leu Asn Glu Ile Tyr Ala Asn Pro Lys Phe Ile Glu
50 55 60

Glu Val Glu Asp Ser Ile Glu Lys Leu Thr Pro Glu Leu Gln Ala Lys
65 70 75 80

Ala Glu Lys Val Lys Asp Val Pro Thr Ala Val Trp Leu Ala Trp Asp
85 90 95

Gly Ser Pro Gly Glu Val Glu Gly His Leu Val Ala Ala Gly Ser Lys
100 105 110

Thr Val Val Phe Leu Leu Tyr Met Ile Pro Thr Arg Asp Cys Asn Ser
115 120 125

Asn Ala Ser Ala Gly Gly Ala Ala Ser Leu Asp Lys Tyr Lys Gly Tyr
130 135 140

Ile Asp Asp Ile Ser Asn Thr Ile Lys Ser His Pro Glu Ser Lys Val
145 150 155 160

Val Met Val Val Glu Pro Asp Thr Leu Gly Asn Leu Val Thr Gly Asn
165 170 175

Ser Glu Ala Cys Lys Asn Val His Thr Leu His Lys Asn Ala Leu Ser
180 185 190

Tyr Ala Val Asp Val Phe Gly Ala Met Ser Asn Val Ser Val Tyr Leu
195 200 205

Asp Ala Ala His Gly Met Trp Leu Gly Pro His Thr Asp Lys Val Ala
210 215 220

Ser Val Ile Lys Glu Ile Leu Asn Asn Ala Pro Asn Gly Lys Ile Arg
225 230 235 240

Gly Leu Ser Thr Asn Val Ser Asn Tyr Gln Pro Val Ser Ser Glu Tyr
245 250 255

Gln Tyr His Gln Lys Leu Ala Ala Ser Leu Ala Ala Val Gly Val Asn
260 265 270

Asp Val His Phe Ile Val Asp Thr Gly Arg Ser Gly Val Asp Val Thr
275 280 285

Glu Thr Phe Ser Lys Gln Gln Thr Trp Cys Asn Phe Ile Gly Ala Gly
290 295 300

Leu Gly Pro Arg Pro Gln Gly Asn Pro Asp Ala Ser Met Pro Leu Leu

193

305 310 315 320

Asp Ala Tyr Met Trp Leu Lys Thr Pro Gly Glu Ala Asp Gly Ser Ala
325 330 335

Val Gly Asp Arg Ala Asp Pro Val Cys Ser His Glu Asp Ser Leu Gln
340 345 350

Val Ala Pro Asp Ala Gly Gln Trp Phe His Asp Tyr Phe Val Leu Leu
355 360 365

Leu Lys Asn Ala Asn Pro Pro Phe
370 375

<210> 94
<211> 463
<212> PRT
<213> Piromyces rhizinflatus

<400> 94

Met Lys Phe Leu Ser Leu Ile Gly Thr Leu Phe Ala Thr Gly Ala Leu
1 5 10 15

Ala Ser Gln Cys His Pro Asn Trp Pro Cys Cys Leu Asn Cys Asp Val
20 25 30

Val Tyr Gln Asp Gly Glu Gly Asp Trp Gly Val Leu Asn Asn Asp Trp
35 40 45

Cys Phe Ile Ser Lys Ser Arg Cys Gly Asn Asn Asn Asn Asn Asn Asn
50 55 60

Gly Tyr Cys Lys Phe Gln Ser Leu Gly Tyr Pro Cys Cys Ser Gly Cys
65 70 75 80

Gln Val Val Tyr Thr Asp Gly Asp Gly Asn Trp Gly Val Glu Asn Gly
85 90 95

Asn Trp Cys Gly Ile Arg Asp Glu Gln Cys Gly Gly Gly Asn Asn Trp
100 105 110

Gln Gln Pro Thr Thr Thr Arg Thr Thr Thr Arg Thr Thr Thr Arg Thr
115 120 125

Gln Asn Asn Arg Pro Thr Ser Asp Asn Phe Phe Asp Asn Thr Leu Tyr
130 135 140

Ser Asn Tyr Lys Phe Gln Asp Glu Val Gln Ser Ser Ile Asn Lys Leu
145 150 155 160

Ser Gly Gln Met Ala Glu Lys Ala Lys Lys Val Lys Tyr Val Pro Thr
165 170 175

Ala Ala Trp Leu Ala Trp Ser Gly Ala Pro Asp Glu Val Pro Arg Tyr
180 185 190

Leu Gln Glu Ala Gly Ser Asp Thr Val Val Phe Val Leu Tyr Met Ile
195 200 205

Pro Thr Arg Asp Cys Asn Ala Asn Ala Ser Ala Gly Gly Ser Ala Asn
210 215 220

194

Leu Asp Thr Tyr Lys Gly Tyr Val Asn Lys Ile Tyr Asn Thr Ile Asn
225 230 235 240

Gln Tyr Pro Asn Ser Arg Ile Val Met Ile Leu Glu Pro Asp Thr Ile
245 250 255

Gly Asn Leu Val Thr Ala Asn Asn Gln Asn Cys Gln Asn Val Gln Asn
260 265 270

Leu His Lys Asn Ala Leu Ala Tyr Ala Ile Ser Lys Phe Gly Thr Gln
275 280 285

Ser Asn Val Ser Val Tyr Leu Asp Ala Ala His Gly Ala Trp Leu Ser
290 295 300

Pro His Ala Asp Lys Thr Ala Gln Val Ile Lys Glu Ile Leu Ser Lys
305 310 315 320

Ala Gly Asn Gly Lys Ile Arg Gly Ile Ser Thr Asn Val Ser Asn Tyr
325 330 335

Gln Thr Ile Asp Asp Glu Tyr Asn Tyr His Gln Arg Leu Asn Ser Ala
340 345 350

Leu Gln Asn Leu Gly Val Ser Asn Met His Phe Ile Val Asp Thr Ser
355 360 365

Arg Asn Gly Ala Asn Ile Ala Ser Gln Phe Asn Gln Ser Gly Thr Trp
370 375 380

Cys Asn Phe Lys Gly Ala Gly Leu Gly Gln Arg Pro Lys Gly Ser Pro
385 390 395 400

Asp Pro Ser Lys Pro Leu Leu Asp Ala Tyr Met Trp Ile Lys Thr Pro
405 410 415

Gly Glu Ala Asp Gly Ser Ser Ser Gly Ala Arg Ala Asp Pro Val Cys
420 425 430

Gly Arg Trp Asp Ser Leu Gln Gly Ala Pro Asp Ala Gly Ser Trp Phe
435 440 445

His Asp Tyr Phe Val Met Leu Leu Gln Asn Ala Asn Pro Ser Phe
450 455 460

<210> 95
<211> 510
<212> PRT
<213> Neocallimastix frontalis

<400> 95

Met Lys Phe Ile Val Phe Ala Ser Ile Leu Ser Ser Gly Ile Ile Lys
1 5 10 15

Leu Ser Asn Ala Ala Cys Gly Gly Pro Tyr Ala Gln Cys Gly Gly Ser
20 25 30

Gly Phe Ser Gly Glu Ala Cys Cys Gln Asp Gly Tyr Lys Cys Val Ala
35 40 45

195

Met Asn Glu Trp Tyr Ser Gln Cys Gln Ala Gly Ser Asp Ala Pro Ala
50 55 60

Asn Ser Ala Ala Pro Val Asn Ser Ala Val Gly Asn Asp Asn Asn Asn
65 70 75 80

Asn Asn Asn Gln Trp Asn Asn Gln Trp Asn Asn Asn Gly Trp Asn Trp
85 90 95

Gly Asn Gly Asn Asn Glu Ala Gln Asn Pro Trp Asn Asn Asn Gly Trp
100 105 110

Ser Trp Glu Gly Gly Asn Asn Gly Asn Gln Asp Gln Asn Gln Trp Asp
115 120 125

Asn Asn Gly Trp Pro Trp Gly Gly Asn Asn Gly Asn Gln Gly Pro Val
130 135 140

Gln Val Asn Ile Gly Glu Asn Asn Asn Gln Asn Gln Asn Pro Ala Asn
145 150 155 160

Asp Ala Pro Ala Pro Pro Ala Gln Gly Gln Ala Pro Ala Pro Ala
165 170 175

Pro Ala Ala Ala Ala Gly Gly Ser Gly Ser Ser Gln Asn Phe Phe Gln
180 185 190

Asn Glu Ile Tyr Ala Asn Pro Lys Phe Ile Glu Glu Val Asp Ser Ser
195 200 205

Ile Ala Lys Leu Asp Gly Glu Leu Lys Ala Lys Ala Glu Lys Val Lys
210 215 220

Ser Val Pro Thr Ala Val Trp Leu Ala Trp Asp Gly Ala Pro Gly Glu
225 230 235 240

Val Ala Gln His Leu Glu Ala Ala Gly Ser Lys Thr Val Val Phe Ile
245 250 255

Met Tyr Met Ile Pro Thr Arg Asp Cys Asn Ala Asn Ala Ser Ala Gly
260 265 270

Gly Ala Ser Asn Leu Gln Thr Tyr Lys Gly Tyr Val Asp Ser Ile Ser
275 280 285

Asn Thr Ile Lys Lys Tyr Pro Asn Ser Lys Val Val Met Ile Leu Glu
290 295 300

Pro Asp Thr Leu Gly Asn Leu Val Thr Ala Asn Ser Glu Asn Cys Lys
305 310 315 320

Asn Val His Gln Leu His Lys Asp Ala Leu Ser Tyr Gly Val Asn Val
325 330 335

Phe Gly Ser Met Ser Asn Val Ser Val Tyr Leu Asp Ala Ala His Gly
340 345 350

Ala Trp Leu Gly Asp Ser Thr Asp Lys Val Ala Ala Val Val Lys Glu
355 360 365

Ile Leu Ser Asn Ala Pro Asn Gly Lys Ile Arg Gly Leu Ser Thr Asn
370 375 380

196

Ile Ser Asn Tyr Gln Pro Val Asp Ser Glu Tyr Gly Tyr His Gln Lys
385 390 395 400

Leu Ala Ser Ala Leu Ser Ala Ala Gly Tyr Pro Asp Met His Phe Val
405 410 415

Val Asp Thr Gly Arg Asp Gly Val Ala Ile Ser Ser Gly Thr Trp Cys
420 425 430

Asn Leu Ile Gly Thr Gly Phe Gly Glu Arg Pro Lys Gly Asn Pro Asn
435 440 445

Pro Gly Met Pro Leu Leu Asp Ala Tyr Met Trp Leu Lys Thr Pro Gly
450 455 460

Glu Ala Asp Gly Ser Ser Thr Gly Ala Arg Ala Asp Pro Val Cys Ala
465 470 475 480

Lys Ser Asp Ser Leu Pro Gly Ala Pro Asp Ala Gly Gln Trp Phe His
485 490 495

Asp Tyr Phe Val Gln Leu Leu Lys Asn Ala Lys Pro Ala Phe
500 505 510

<210> 96

<211> 459

<212> PRT

<213> Orpinomyces sp. PC-2

<400> 96

Met Lys Phe Ser Thr Val Leu Ala Thr Leu Phe Ala Thr Gly Ala Leu
1 5 10 15

Ala Ser Glu Cys His Trp Gln Tyr Pro Cys Cys Lys Asp Cys Thr Val
20 25 30

Tyr Tyr Thr Asp Thr Glu Gly Lys Trp Gly Val Leu Asn Asn Asp Trp
35 40 45

Cys Met Ile Asp Asn Arg Arg Cys Ser Ser Asn Asn Asn Asn Cys Ser
50 55 60

Ser Ser Ile Thr Ser Gln Gly Tyr Pro Cys Cys Ser Asn Asn Asn Cys
65 70 75 80

Lys Val Glu Tyr Thr Asp Asn Asp Gly Lys Trp Gly Val Glu Asn Asn
85 90 95

Asn Trp Cys Gly Ile Ser Asn Ser Cys Gly Gly Gly Gln Gln Gln Gln
100 105 110

Pro Thr Gln Pro Thr Gln Pro Thr Gln Pro Gln Gln Pro Thr Gln Pro
115 120 125

Ser Ser Asp Asn Phe Phe Glu Asn Glu Ile Tyr Ser Asn Tyr Lys Phe
130 135 140

Gln Gly Glu Val Asp Ile Ser Ile Lys Lys Leu Asn Gly Asp Leu Lys
145 150 155 160

197

Ala Lys Ala Glu Lys Val Lys Tyr Val Pro Thr Ala Val Trp Leu Ala
165 170 175

Trp Asp Gly Ala Pro Gln Glu Val Pro Arg Tyr Leu Gln Glu Ala Gly
180 185 190

Asn Lys Thr Val Val Phe Val Leu Tyr Met Ile Pro Thr Arg Asp Cys
195 200 205

Gly Ala Asn Ala Ser Ala Gly Gly Ser Ala Thr Ile Asp Lys Tyr Lys
210 215 220

Gly Tyr Ile Asn Asn Ile Tyr Asn Thr Ser Asn Gln Tyr Lys Asn Ser
225 230 235 240

Lys Ile Val Met Ile Leu Glu Pro Asp Thr Ile Gly Asn Leu Val Thr
245 250 255

Asn Asn Asn Asp Asn Cys Arg Asn Val Arg Asn Met His Lys Gln Ala
260 265 270

Leu Ser Tyr Ala Ile Ser Lys Phe Gly Thr Gln Ser His Val Lys Val
275 280 285

Tyr Leu Asp Ala Ala His Gly Ala Trp Leu Asn Gln Tyr Ala Asp Gln
290 295 300

Thr Ala Asn Val Ile Lys Glu Ile Leu Asn Asn Ala Gly Ser Gly Lys
305 310 315 320

Leu Arg Gly Ile Ser Thr Asn Val Ser Asn Tyr Gln Ser Ile Glu Ser
325 330 335

Glu Tyr Lys Tyr His Gln Asn Leu Asn Arg Ala Leu Glu Ser Lys Gly
340 345 350

Val Arg Gly Leu Lys Phe Ile Val Asp Thr Ser Arg Asn Gly Ala Asn
355 360 365

Val Glu Gly Ala Phe Asn Ala Ser Gly Thr Trp Cys Asn Phe Lys Gly
370 375 380

Ala Gly Leu Gly Gln Arg Pro Lys Gly Asn Pro Asn Pro Gly Ser Met
385 390 395 400

Pro Leu Leu Asp Ala Tyr Met Trp Ile Lys Thr Pro Gly Glu Ala Asp
405 410 415

Gly Ser Ser Gln Gly Ser Arg Ala Asp Pro Val Cys Ala Arg Gly Asp
420 425 430

Ser Leu Gln Gly Ala Pro Asp Ala Gly Ser Trp Phe His Glu Tyr Phe
435 440 445

Thr Met Leu Ile Gln Asn Ala Asn Pro Pro Phe
450 455

<210> 97

<211> 345

<212> PRT

<213> Piromyces rhizinflatus

198

<400> 97

Thr Ala Val Trp Leu Ala Trp Asp Gly Ala Pro Gly Glu Val Glu Gly
1 5 10 15

His Leu Lys Asn Ala Gly Asp Lys Thr Ile Val Phe Ile Leu Tyr Met
20 25 30

Ile Pro Thr Arg Asp Cys Asn Ser Lys Ala Ser Ala Gly Gly Ala Ser
35 40 45

Asn Leu Lys Lys Tyr Gln Gly Tyr Val Asp Ser Ile Ala Asn Thr Ile
50 55 60

Ser Lys His Pro Glu Ser Lys Val Val Met Val Ile Glu Pro Asp Thr
65 70 75 80

Leu Gly Asn Leu Ile Thr Gly Glu Thr Glu Glu Cys Lys Thr Val His
85 90 95

Thr Leu His Lys Asp Ala Leu Ala Tyr Ala Val Asn Val Phe Gly Asp
100 105 110

Met Ser Asn Val Ser Ala Tyr Leu Asp Ala Ala His Gly Lys Trp Leu
115 120 125

Gly Trp Ala Ala Gly Lys Thr Ala Ala Val Ile Lys Glu Ile Leu Asp
130 135 140

Asn Ala Pro Asn Gly Asn Ile Arg Gly Phe Ser Thr Asn Val Ser Asn
145 150 155 160

Tyr Gln Pro Ile Glu Ser Glu Tyr Glu Tyr His Glu Lys Leu Asn Ala
165 170 175

Ala Leu Glu Asp Leu Gly Ile Thr Gly Lys Lys Phe Ile Val Asp Thr
180 185 190

Gly Arg Ser Gly Val Asp Val Thr Glu Glu Phe Asn Leu Asn Gln Thr
195 200 205

Trp Cys Asn Leu Ile Tyr Ala Gly Leu Gly Glu Pro Ser Arg Gly Ser
210 215 220

Pro Asp Pro Glu Lys Phe Pro Leu Leu Asp Ala Tyr Phe Trp Leu Lys
225 230 235 240

Pro Pro Gly Glu Ala Asp Gly Ser Asp Thr Gly Ser Arg Ala Asp Pro
245 250 255

Val Cys Gly Arg Glu Asp Ser Phe Pro Gly Ala Pro Asp Ala Gly Ser
260 265 270

Trp Phe Ser Glu Tyr Phe Ala Ser Met Leu Glu Lys Ser Pro Phe Tyr
275 280 285

Gly Glu Gly Ile Glu Glu Glu Pro Glu Glu Pro Glu Glu Pro Leu Cys
290 295 300

Ser Gln Lys Phe Leu Asp Gln Gly Tyr Gln Cys Cys Ser Gln Cys Gly
305 310 315 320

199

Thr Ile Tyr Tyr Val Asp Asp Ala Gly Asn Trp Gly Val Glu Asn Asn
325 330 335

Glu Trp Cys Gly Leu Pro Glu Asp Cys
340 345

<210> 98
<211> 457
<212> PRT
<213> Micromonospora cellulolyticum

<400> 98

Val Ala Ile Leu Ser Ala Arg Arg Arg Ser Ala Ala Ile Ser Val Thr
1 5 10 15

Ala Val Ala Gly Leu Ala Ala Ala Gly Val Leu Arg Val Gly Gly Val
20 25 30

Ala Gly Thr Val Ser Gly Ser Leu Tyr Arg Asp Pro Ser Ser Ala Val
35 40 45

Val Arg Trp Val Ala Ala Asn Pro Gly Asp Phe Arg Ala Ala Val Ile
50 55 60

Arg Glu Lys Ile Ala Ser Gln Pro Gln Ala Arg Trp Tyr Ala Asn Phe
65 70 75 80

Asn Pro Ser Thr Ile Gln Ser Glu Val Ser Ala Phe Ile Gly Ala Ala
85 90 95

Asn Ser Ala Gln Gln Ile Pro Val Leu Ser Val Tyr Glu Ile Thr Asn
100 105 110

Arg Asp Cys Gly Gly Ala His Ala Gly Gly Ala Pro Asp Leu Asn Gln
115 120 125

Tyr Gln Thr Trp Val Ser Asn Phe Ala Arg Gly Leu Gly Asn Gln Thr
130 135 140

Val Leu Ile Ile Leu Glu Thr Asp Ser Leu Ala Leu Gln Thr Cys Leu
145 150 155 160

Ser Thr Ser Glu Leu Asn Ala Arg Asn Gln Ala Leu Ser Thr Ala Thr
165 170 175

Gln Thr Ile Lys Ser Ala Asn Pro Asn Ala Lys Val Tyr Leu Asp Gly
180 185 190

Gly His Ser Thr Trp Asn Ser Ala Asn Asp Thr Ala Asn Arg Leu Arg
195 200 205

Ala Ala Gly Val Gln Tyr Ala Asp Gly Phe Phe Thr Asn Val Ser Asn
210 215 220

Phe Asn Pro Thr Ser Ser Glu Ala Asn Phe Gly Arg Ala Val Ile Ser
225 230 235 240

Ala Leu Asn Gly Met Gly Ile Ser Gly Lys Arg Gln Val Ile Asp Thr
245 250 255

Ser Arg Asn Gly Gly Ala Ala Gly Asp Trp Cys Ala Asp Asp Asn Thr

260 265 270

Asp Arg Arg Ile Gly Gln Tyr Pro Thr Thr Asn Thr Gly Asp Ala Asn
275 280 285

Ile Asp Ala Tyr Leu Trp Val Lys Pro Pro Gly Glu Ala Asp Gly Cys
290 295 300

Ala Thr Arg Gly Ser Phe Gln Pro Asp Leu Ala Phe Ser Leu Ala Asn
305 310 315 320

Gly Val Pro Asn Pro Pro Thr Thr Ala Pro Pro Thr Thr Asn Arg Ala
325 330 335

Asp Asp Arg Pro Pro Thr Thr Ala Pro Pro Thr Thr Asp Thr Pro Thr
340 345 350

Thr Ala Pro Pro Thr Thr Pro Pro Pro Ala Gly Asn Gly Leu Ser Ala
355 360 365

Ser Val Ala Ile Thr Gln Trp Asn Gly Gly Phe Thr Ala Ser Val Asn
370 375 380

Val Thr Ala Gly Ser Ala Ile Asn Gly Trp Thr Val Thr Val Ala Leu
385 390 395 400

Pro Gly Gly Ala Ala Ile Thr Gly Thr Trp Asn Ala Gln Ala Ser Gly
405 410 415

Thr Ser Gly Thr Val Arg Phe Thr Asn Val Gly Tyr Asn Gly Gln Val
420 425 430

Gly Ala Gly Gln Thr Thr Asn Phe Gly Phe Gln Gly Thr Gly Thr Gly
435 440 445

Gln Gly Ala Thr Ala Thr Cys Ala Ala
450 455

<210> 99

<211> 193

<212> PRT

<213> Orpinomyces sp. PC-2

<400> 99

Thr Leu His Lys Asn Ala Leu Ser Tyr Ala Val Asn Val Phe Gly Ser
1 5 10 15

Met Lys Asn Val Ser Val Tyr Leu Asp Ala Ala His Gly Met Trp Leu
20 25 30

Ser Ala Val Ala Asp Lys Thr Ala Ala Val Ile Lys Glu Val Leu Asp
35 40 45

Asn Ala Pro Asn Gly Lys Ile Arg Gly Leu Ser Thr Asn Ile Ser Asn
50 55 60

Tyr Gln Pro Val Tyr Ser Glu Tyr Lys Tyr His Glu Lys Leu Ser Ala
65 70 75 80

Glu Leu Glu Lys Leu Gly Val Ser Asp Ile His Phe Ile Val Asp Thr
85 90 95

201

Gly Arg Asn Gly Val Asp Ile Thr Glu Thr Phe Ser Lys Thr Gln Thr
100 105 110

Trp Cys Asn Phe Val Gly Thr Gly Phe Gly Glu Arg Pro Gln Gly Asn
115 120 125

Pro Asp Pro Val Lys Met Pro Leu Leu Asp Ala Tyr Met Trp Leu Lys
130 135 140

Thr Pro Gly Glu Ala Asp Gly Ser Asp Thr Gly Ser Arg Ala Asp Pro
145 150 155 160

Val Cys Ala Arg Glu Asp Ser Leu Pro Gly Ser Pro Asp Ala Gly Gln
165 170 175

Trp Phe His Asp Tyr Phe Val Gln Leu Leu Glu Asn Ala Asn Pro Ala
180 185 190

Phe

<210> 100

<211> 458

<212> PRT

<213> Catenulispora acidiphila

<400> 100

Met Val Ala Thr Gln Leu Gly Gly Val Ala Met Ala Gly Thr Ile Ala
1 5 10 15

Ser Gly Thr Gln Phe Tyr Ala Asp Pro Asn Ser Gln Val Val Lys Trp
20 25 30

Asp Ala Ala Asn Pro Gly Asp Ala Arg Glu Pro Ala Ile Ala Ser Gln
35 40 45

Ile Ala Ser Val Ser Gln Gly Ile Trp Phe Ser Asn Tyr Arg Pro Ser
50 55 60

Thr Val Gln Ser Asp Val Ser Ala Val Thr Thr Ala Ala Ala Ala
65 70 75 80

Gly Lys Thr Pro Val Leu Val Val Tyr Glu Ile Pro Asn Arg Asp Cys
85 90 95

Gly Gly Ala Ser Ala Gly Gly Ala Pro Asp Ile Ser Ser Tyr Glu Asn
100 105 110

Tyr Ile Gln Ser Phe Ala Asn Gly Leu Gly Ser His Gln Val Ile Val
115 120 125

Ile Leu Glu Pro Asp Ser Leu Ala Leu Gln Thr Cys Leu Ser Ser Gln
130 135 140

Gln Ala Thr Asp Arg Asp Asn Ala Ile Ala Phe Ala Gly Ala His Leu
145 150 155 160

Lys Ser Ala Asp Pro Ala Ala Lys Val Tyr Leu Asp Ala Gly His Ser
165 170 175

202

Ser Trp Asn Ser Pro Ser Ala Gln Ala Ala Ala Leu Asn Ala Ala Gly
180 185 190

Val Lys Thr Ser Ser Asp Gly Ile Phe Ser Asn Val Ser Asn Phe Gln
195 200 205

Thr Thr Ala Ser Glu Val Ser Tyr Asp Lys Gln Val Leu Ala Ala Leu
210 215 220

Gly Ser Pro Ser Asn Leu His Ile Val Val Asp Thr Ser Arg Asn Gly
225 230 235 240

Asn Gly Pro Ala Gly Ser Ala Trp Cys Asp Pro Ser Gly Arg Ala Leu
245 250 255

Gly Gln Ala Pro Thr Ala Asn Thr Gly Asp Ala Ala Val Asp Ala Phe
260 265 270

Leu Trp Ile Lys Pro Pro Gly Glu Ala Asp Gly Cys Ala Asp Ala Ala
275 280 285

Gly Thr Phe Asp Pro Ala Leu Ala Tyr Ala Leu Ile Thr Asn Gly Gly
290 295 300

Gly Pro Pro Pro Thr Ser Pro Ser Ser Thr Pro Ser Thr Thr Pro Ser
305 310 315 320

Thr Thr Pro Ser Thr Thr Pro Ser Thr Thr Pro Ser Thr Thr Pro Ser
325 330 335

Thr Thr Pro Ser Thr Pro Pro Ser Thr Thr Pro Ser Ser Pro Pro Ser
340 345 350

Ser Ser Pro Ala Gly Cys Gln Val Thr Tyr Thr Arg Thr Asn Glu Trp
355 360 365

Ala Gly Gly Phe Thr Ala Asn Val Ser Ile Thr Ser Ser Lys Ala Leu
370 375 380

Ser Ser Trp Thr Val Gly Phe Thr Tyr Gly Gly Asp Gln Gln Ile Thr
385 390 395 400

Asn Ser Trp Asn Gly Asn His Thr Gln Ser Gly Arg Asn Val Thr Leu
405 410 415

Thr Ser Leu Ser Tyr Asn Gly Ser Ile Gly Ala Gly Gln Thr Leu Thr
420 425 430

Gly Val Gly Val Gln Gly Thr Trp Thr Ser Ser Asp Ala Ala Pro Ser
435 440 445

Ala Phe Thr Leu Asn Gly Val Ala Cys His
450 455

<210> 101

<211> 422

<212> PRT

<213> Actinosynnema mirum

<400> 101

Met Ala Gly Ala Ala Val Leu Val Thr Gly Gly Gly Gln Thr Ala Thr

203

1 5 10 15

Ala Ala Asp Ser Ala Phe Tyr Thr Asp Pro Gly Ser Ser Ser Ala Arg
20 25 30

Trp Val Ala Ala Asn Pro Asn Asp Ser Arg Ala Ala Val Ile Arg Asp
35 40 45

Arg Val Ala Ser Val Pro Gln Ala Lys Trp Phe Thr Thr Thr Asn Thr
50 55 60

Ser Thr Val Arg Ser Glu Val Ser Ala Phe Val Gly Ala Ala Ala Ser
65 70 75 80

Ala Gly Lys Ile Pro Ile Leu Val Val Tyr Asn Ile Pro Asn Arg Asp
85 90 95

Cys Gly Gly Ala Ser Gly Gly Gly Ala Pro Ser His Gln Ala Tyr Arg
100 105 110

Ala Trp Val Asp Glu Val Ala Ala Gly Leu Gly Gly Arg Pro Ala Ser
115 120 125

Ile Ile Leu Glu Pro Asp Val Leu Pro Ile Met Ser Asn Cys Gln Ser
130 135 140

Ala Asp Gln Gln Asn Gln Thr Lys Ala Ser Met Ser Tyr Ala Gly Arg
145 150 155 160

Lys Leu Lys Ser Gly Ser Gly Gln Ala Lys Val Tyr Phe Asp Ile Gly
165 170 175

Asn Ser Asp Trp Leu Ala Pro Ala Glu Ala Ala Asn Arg Leu Arg Gly
180 185 190

Ala Asp Val Ser Gly Ser Ser Asp Gly Ile Ala Ser Asn Val Ser Asn
195 200 205

Tyr Arg Ala Thr Gln Ala Glu Val Ser Tyr Thr Lys Ala Ile Leu Asn
210 215 220

Ala Leu Gly Asp Gly Arg Leu Lys Ala Val Ile Asp Thr Ser Arg Asn
225 230 235 240

Gly Asn Gly Pro Leu Gly Ser Glu Trp Cys Asp Pro Pro Gly Arg Ala
245 250 255

Ile Gly Thr Pro Ser Thr Lys Asn Thr Gly Asp Ser Gln Ile Asp Ala
260 265 270

Phe Leu Trp Val Lys Ile Val Gly Glu Ala Asp Gly Cys Ile Ala Ser
275 280 285

Ala Gly Gln Phe Val Pro Gln Arg Ala Tyr Asp Leu Ala Val Ala Ala
290 295 300

Gly Pro Val Pro Thr Thr Thr Thr Thr Pro Gly Gly Asn Pro Gly
305 310 315 320

Gly Gly Cys Ala Val Thr His Arg Val Val Ser Gln Trp Asn Gly Gly
325 330 335

204

Phe Thr Gly Glu Val Val Val Glu Asn Arg Gly Pro Ala Ile Ser Ser
340 345 350

Trp Thr Leu Glu Phe Ser Ala Pro Gly Val Thr Val Thr Gln Gly Trp
355 360 365

Asn Gly Thr Trp Thr Asp Thr Gly Asp Gly Val Arg Val Val Asn Thr
370 375 380

Ala Trp Asn Gly Ala Leu Ala Ser Gly Gly Arg Val Thr Ala Gly Tyr
385 390 395 400

Asn Ala Asn Tyr Gly Gly Gly Ala Pro Pro Phe Ser Ser Pro Thr Leu
405 410 415

Asn Gly Ala Ala Cys Ser
420

<210> 102

<211> 444

<212> PRT

<213> Actinosynnema mirum

<400> 102

Met Ala Ala Gly Ala Leu Ser Ser Ala Leu Val Ala Ala Ala Thr Ala
1 5 10 15

Ile Ala Thr Gly Thr Ala Ser Pro Ala Ala Val Ala Ala Asp Ser Glu
20 25 30

Phe Tyr Ser Asp Pro Ala Thr Ser Ala Ala Arg Trp Val Ala Ala Asn
35 40 45

Pro Asn Asp Ser Arg Ala Ala Val Ile Arg Asp Arg Val Ala Ser Val
50 55 60

Pro Gln Ala Lys Trp Phe Thr Thr Thr Asn Thr Ser Thr Ile Arg Ala
65 70 75 80

Glu Val Asp Ala His Thr Ser Ala Ala Ala Ser Ala Gly Lys Thr Pro
85 90 95

Ile Leu Val Val Tyr Asn Ile Pro Asn Arg Asp Cys Gly Gly Ala Ser
100 105 110

Gly Gly Gly Ala Pro Ser His Gly Ala Tyr Arg Gln Trp Val Asp Gln
115 120 125

Phe Ala Ala Gly Leu Ala Gly Arg Pro Ala Ala Ile Ile Leu Glu Pro
130 135 140

Asp Val Leu Pro Ile Met Ser Thr Cys Gln Ser Ala Ser Gln Gln Ala
145 150 155 160

Glu Thr Arg Ala Ser Met Ala Tyr Ala Gly Lys Ala Leu Lys Ala Ala
165 170 175

Ser Ser Gln Ala Lys Val Tyr Phe Asp Ile Gly His Ser Ala Trp Leu
180 185 190

Thr Pro Ala Glu Ala Ala Asn Arg Leu Arg Ala Ala Glu Val Ser Thr

205

195 200 205
Ser Ala Asp Gly Ile Ala Thr Asn Val Ser Asn Tyr Arg Arg Thr Ala
210 215 220
Asp Glu Val Ala Phe Ala Lys Ala Thr Leu Asn Ala Leu Gly Asp Gly
225 230 235 240
Arg Leu Lys Ala Val Val Asp Thr Ser Arg Asn Gly Asn Gly Pro Leu
245 250 255
Gly Ser Glu Trp Cys Asp Pro Pro Gly Arg Ala Ile Gly Thr Pro Ser
260 265 270
Thr Arg Asn Thr Gly Asp Pro Gln Ile Asp Ala Phe Leu Trp Val Lys
275 280 285
Ile Pro Gly Glu Ala Asp Gly Cys Ile Ala Gly Ala Gly Gln Phe Val
290 295 300
Pro Gln Arg Ala Tyr Asp Met Ala Val Ala Ala Gly Pro Ala Pro Thr
305 310 315 320
Thr Thr Thr Thr Thr Thr Thr Thr Thr Arg Val Thr Thr Thr Thr Thr
325 330 335
Thr Pro Pro Pro Asn Gly Ala Ala Cys Val Val Arg His Arg Val Val
340 345 350
Ser Ser Trp Ser Gly Gly His Thr Gly Glu Val Val Ile Glu Asn Arg
355 360 365
Gly Pro Ala Leu Gln Asn Trp Thr Leu Glu Phe Ser Ala Pro Gly Val
370 375 380
Ala Val Ser Gln Gly Trp Asn Gly Thr Trp Thr Asp Leu Gly Asp Thr
385 390 395 400
Val Arg Val Thr Ser Ala Ser Trp Asn Gly Gly Ile Ala Thr Gly Gly
405 410 415
Thr Ala Thr Thr Gly Tyr Ser Ala Ser Phe Ser Gly Gly Thr Pro Pro
420 425 430
Phe Thr Ser Pro Val Leu Asn Gly Thr Ala Cys Ala
435 440
<210> 103
<211> 426
<212> PRT
<213> Nocardiosis dassonvillei

<400> 103
Met Ser Ser Val Ser Ala Leu Ala Leu Gly Thr Ala Leu Leu Ala Thr
1 5 10 15
Ala Pro Ala Ser Ala Ala Asp Ser Glu Phe Tyr Val Asn Pro Asn Thr
20 25 30
Ser Ala Ala Val Trp Val Glu Glu Asn Pro Asn Asp Pro Arg Ala Asp
35 40 45

206

Val Ile Arg Asp Arg Ile Ala Ser Val Ala Gln Ala Thr Trp Phe Thr
50 55 60

Gln Tyr Asn Pro Ala Glu Val Arg Asp Asp Val Asp Ala Val Val Ser
65 70 75 80

Ala Ala Asp Ala Gln Gly Gln Thr Pro Ile Leu Val Val Tyr Asn Ile
85 90 95

Pro Gly Arg Asp Cys Gly Asn His Ser Gly Gly Gly Ala Pro Ser His
100 105 110

Asp Ala Tyr Arg Ala Trp Val Asp Glu Val Ala Ala Gly Leu Glu Gly
115 120 125

Arg Ser Ala Thr Ile Val Leu Glu Pro Asp Ala Leu Pro Leu Val Ser
130 135 140

Gly Cys Ser Asp Pro Ser Glu Leu Leu Asp Ser Met Ala Tyr Ala Gly
145 150 155 160

Lys Ala Leu Met Glu Gly Ser Ser Glu Ala Arg Val Tyr Phe Asp Ile
165 170 175

Gly Asn Ser Ala Trp Leu Asp Pro Gln Glu Ala Ala Gly Leu Leu Asn
180 185 190

Gly Ala Asp Val Ala Asn Ser Ala His Gly Val Ala Thr Asn Thr Ser
195 200 205

Asn Tyr Asn Trp Thr His Asp Glu Val Ala Phe Ala Glu Ala Val Ile
210 215 220

Ala Ala Thr Gly Val Pro Gly Leu Gly Ala Val Ile Asp Thr Ser Arg
225 230 235 240

Asn Gly Asn Gly Pro Ala Pro Gln Asn Glu Trp Cys Asp Pro Pro Gly
245 250 255

Arg Met Ile Gly Arg Pro Ser Thr Thr Asp Thr Gly Asn Pro Leu Ile
260 265 270

Asp Ala Phe Ile Trp Thr Lys Leu Pro Gly Glu Ala Asp Gly Cys Ile
275 280 285

Ala Pro Ala Gly Gln Phe Val Pro Gln Ala Ala Tyr Asp Met Ala Val
290 295 300

Asn Ala Pro Glu Tyr Pro Thr Asp Pro Gly Glu Pro Thr Asp Pro Glu
305 310 315 320

Glu Pro Thr Asp Pro Pro Glu Gly Glu Gly Cys Thr Ala Asp Tyr Arg
325 330 335

Val Val Ser Glu Trp Gly Asn Gly Phe Gln Ala Ala Val Thr Val Thr
340 345 350

Ala Glu Asp Ser Leu Ser Gly Trp Thr Val Thr Trp Thr Tyr Ala Asp
355 360 365

Gly Gln Arg Phe Ser Gln Gly Trp Asn Ala Glu Phe Ser Ser Ser Gly

207

370 375 380

Ser Arg Val Thr Ala Ser Asp Leu Gly Trp Asn Gly Thr Leu Ser Ala
385 390 395 400

Gly Gly Ser Thr Glu Phe Gly Phe Thr Gly Thr His Gly Gly Ser Asn
405 410 415

Gly Val Pro Glu Val Thr Cys Ser Ala Ala
420 425

<210> 104

<211> 286

<212> PRT

<213> Termobifida fusca

<400> 104

Asn Asp Ser Pro Phe Tyr Val Asn Pro Asn Met Ser Ser Ala Glu Trp
1 5 10 15

Val Arg Asn Asn Pro Asn Asp Pro Arg Thr Pro Val Ile Arg Asp Arg
20 25 30

Ile Ala Ser Val Pro Gln Gly Thr Trp Phe Ala His His Asn Pro Gly
35 40 45

Gln Ile Thr Gly Gln Val Asp Ala Leu Met Ser Ala Ala Gln Ala Ala
50 55 60

Gly Lys Ile Pro Ile Leu Val Val Ser Asn Ala Pro Gly Arg Asp Cys
65 70 75 80

Gly Asn His Ser Ser Gly Gly Ala Pro Ser His Ser Ala Tyr Arg Ser
85 90 95

Trp Ile Asp Glu Phe Ala Ala Gly Leu Lys Asn Arg Pro Ala Tyr Ile
100 105 110

Ile Val Glu Pro Asp Leu Ile Ser Leu Met Ser Ser Cys Met Gln His
115 120 125

Val Gln Gln Glu Val Leu Glu Thr Met Ala Tyr Ala Gly Lys Ala Leu
130 135 140

Lys Ala Gly Ser Ser Gln Ala Arg Ile Tyr Phe Asp Ala Gly His Ser
145 150 155 160

Ala Trp His Ser Pro Ala Gln Met Ala Ser Trp Leu Gln Gln Ala Asp
165 170 175

Ile Ser Asn Ser Ala His Gly Ile Ala Thr Asn Thr Ser Asn Tyr Arg
180 185 190

Trp Thr Ala Asp Glu Val Ala Tyr Ala Lys Ala Val Leu Ser Ala Ile
195 200 205

Gly Asn Pro Ser Leu Arg Ala Val Ile Asp Thr Ser Arg Asn Gly Asn
210 215 220

Gly Pro Ala Gly Asn Glu Trp Cys Asp Pro Ser Gly Arg Ala Ile Gly
225 230 235 240

208

Thr Pro Ser Thr Thr Asn Thr Gly Asp Pro Met Ile Asp Ala Phe Leu
245 250 255

Trp Ile Lys Leu Pro Gly Glu Ala Asp Gly Cys Ile Ala Gly Ala Gly
260 265 270

Gln Phe Val Pro Gln Ala Ala Tyr Glu Met Ala Ile Ala Ala
275 280 285

<210> 105

<211> 448

<212> PRT

<213> Micromonospora sp. ATCC 39149

<400> 105

Met Ala Ala Ser Ala Leu Thr Ala Ala Val Ala Val Ser Ile Leu Gly
1 5 10 15

Gly Thr Pro Ala Ser Ala Ala Asp Ser Ala Phe Tyr Val Asp Pro Gln
20 25 30

Ala Ser Ala Ala Arg Trp Val Ala Ala Asn Pro Gly Asp Trp Arg Ala
35 40 45

Ala Val Ile Arg Asp Arg Ile Ala Ala Val Pro Gln Gly Arg Trp Phe
50 55 60

Thr Thr Thr Asn Thr Ser Thr Val Arg Ser Glu Val Asp Gln Phe Val
65 70 75 80

Gly Ala Ala Ala Ala Ala Gly Lys Val Pro Ile Met Val Val Tyr Asn
85 90 95

Ile Pro Asn Arg Asp Cys Ser Gly Ala Ser Gly Gly Gly Ala Pro Ser
100 105 110

His Thr Ala Tyr Arg Gln Trp Val Asp Gln Val Ala Ala Gly Leu Ala
115 120 125

Gly Arg Pro Ala Thr Ile Val Leu Glu Pro Asp Val Leu Pro Ile Met
130 135 140

Thr Asn Cys Gln Asn Ala Ser Gln Gln Ala Glu Thr Arg Ala Ser Met
145 150 155 160

Ala Tyr Ala Gly Lys Lys Leu Lys Ser Gly Ser Ala Gln Ala Lys Val
165 170 175

Tyr Phe Asp Ala Gly Asn Ser Ala Trp Leu Ala Pro Ala Glu Ile Ala
180 185 190

Ser Arg Leu Asn Gly Ala Asp Ile Ala Asn Ser Ala Asp Gly Ile Ser
195 200 205

Leu Asn Val Ser Asn Tyr Arg Thr Thr Ala Glu Ser Val Ser Tyr Ala
210 215 220

Lys Gln Val Ile Ala Ala Thr Gly Val Ser Arg Leu Lys Ala Val Ile
225 230 235 240

209

Asp Thr Ser Arg Asn Gly Asn Gly Pro Leu Gly Ser Glu Trp Cys Asp
245 250 255

Pro Pro Gly Arg Ala Ile Gly Thr Pro Ser Thr Thr Ala Thr Gly Asp
260 265 270

Ser Ala Ile Ala Ala Tyr Leu Trp Val Lys Leu Pro Gly Glu Ala Asp
275 280 285

Gly Cys Ile Ala Pro Ala Gly Gln Phe Val Pro Gln Arg Ala Tyr Asp
290 295 300

Leu Ala Ile Ala Ala Gly Pro Val Pro Thr Thr Ala Pro Pro Thr Thr
305 310 315 320

Ala Pro Pro Thr Thr Ala Pro Pro Thr Thr Ala Pro Pro Thr Thr Ala
325 330 335

Pro Pro Thr Thr Pro Pro Asn Gly Ala Cys Lys Val Thr Phe Thr Pro
340 345 350

Asn Thr Trp Ser Gly Gly Phe Thr Ala Glu Leu Arg Val Thr Asn Gly
355 360 365

Gly Ser Ala Leu Asn Gly Trp Ser Leu Ser Phe Gly Phe Gly Ser Gly
370 375 380

Ser Gly Val Arg Leu Thr Ser Gly Trp Asn Gly Glu Trp Ser Gln Asn
385 390 395 400

Gly Asp Val Phe Leu Val Arg Asn Ala Ala Trp Asn Gly Asn Leu Pro
405 410 415

Ala Gly Gly Thr Leu Ser Val Gly Phe Gln Gly Thr Phe Ser Gly Ala
420 425 430

Ser Leu Pro Thr Ala Val Gly Phe Thr Leu Asn Gly Ser Arg Cys Asn
435 440 445

<210> 106
<211> 456
<212> PRT
<213> Termobispora bispora

<400> 106

Met Ser Arg Ile Arg Arg Phe Leu Ala Thr Ala Leu Ala Ala Ala Thr
1 5 10 15

Ala Gly Val Gly Ala Ile Val Thr Ala Ile Ala Ser Ala Gly Pro Ala
20 25 30

His Ala Tyr Asp Ser Pro Phe Tyr Val Asp Pro Gln Ser Asn Ala Ala
35 40 45

Lys Trp Val Ala Ala Asn Pro Asn Asp Pro Arg Thr Pro Val Ile Arg
50 55 60

Asp Arg Ile Ala Ala Val Pro Thr Gly Arg Trp Phe Ala Asn Tyr Asn
65 70 75 80

Pro Ser Thr Val Arg Ala Glu Val Asp Ala Tyr Val Gly Ala Ala Ala

210

85 90 95

Ala Ala Gly Lys Ile Pro Ile Met Val Val Tyr Ala Met Pro Asn Arg
100 105 110Asp Cys Gly Gly Pro Ser Ala Gly Gly Ala Pro Asn His Thr Ala Tyr
115 120 125Arg Ala Trp Ile Asp Glu Ile Ala Ala Gly Leu Arg Asn Arg Pro Ala
130 135 140Val Ile Ile Leu Glu Pro Asp Ala Leu Pro Ile Met Thr Asn Cys Met
145 150 155 160Ser Pro Ser Glu Gln Ala Glu Val Gln Ala Ser Ala Val Gly Ala Gly
165 170 175Lys Lys Phe Lys Ala Ala Ser Ser Gln Ala Lys Val Tyr Phe Asp Ala
180 185 190Gly His Asp Ala Trp Val Pro Ala Asp Glu Met Ala Ser Arg Leu Arg
195 200 205Gly Ala Asp Ile Ala Asn Ser Ala Asp Gly Ile Ala Leu Asn Val Ser
210 215 220Asn Tyr Arg Tyr Thr Ser Gly Leu Ile Ser Tyr Ala Lys Ser Val Leu
225 230 235 240Ser Ala Ile Gly Ala Ser His Leu Arg Ala Val Ile Asp Thr Ser Arg
245 250 255Asn Gly Asn Gly Pro Leu Gly Ser Glu Trp Cys Asp Pro Pro Gly Arg
260 265 270Ala Thr Gly Thr Trp Ser Thr Thr Asp Thr Gly Asp Pro Ala Ile Asp
275 280 285Ala Phe Leu Trp Ile Lys Pro Pro Gly Glu Ala Asp Gly Cys Ile Ala
290 295 300Thr Pro Gly Val Phe Val Pro Asp Arg Ala Tyr Glu Leu Ala Met Asn
305 310 315 320Ala Ala Pro Pro Thr Tyr Ser Pro Ser Pro Thr Pro Ser Thr Pro Ser
325 330 335Pro Ser Pro Ser Gln Ser Asp Pro Gly Ser Pro Ser Pro Ser Pro Ser
340 345 350Gln Pro Pro Ala Gly Arg Ala Cys Glu Ala Thr Tyr Ala Leu Val Asn
355 360 365Gln Trp Pro Gly Gly Phe Gln Ala Glu Val Thr Val Lys Asn Thr Gly
370 375 380Ser Ser Pro Ile Asn Gly Trp Thr Val Gln Trp Thr Leu Pro Ser Gly
385 390 395 400Gln Ser Ile Thr Gln Leu Trp Asn Gly Asp Leu Ser Thr Ser Gly Ser
405 410 415

211

Asn Val Thr Val Arg Asn Val Ser Trp Asn Gly Asn Val Pro Ala Gly
420 425 430

Gly Ser Thr Ser Phe Gly Phe Leu Gly Ser Gly Thr Gly Gln Leu Ser
435 440 445

Ser Ser Ile Thr Cys Ser Ala Ser
450 455

<210> 107

<211> 5

<212> PRT

<213> Secuencia artificial

<220>

<223> Secuencia de consenso

<220>

<221> CARACTERÍSTICA_MISC

<222> (3)..(3)

<223> Xaa es A, G o C

<400> 107

Gly Glu Xaa Asp Gly
1 5

2/2

 Industria y Comercio SUPERINTENDENCIA	EXTRACTO PARA PUBLICACIÓN ☒ Patente de Invención ☐ Patente modelo de utilidad ☐ PCT	
(21) No. De solicitud	(51) int. Cl: C12N 15/63	
(22) Fecha de solicitud	(71) Solicitante (s) CALIFORNIA INSTITUTE OF TECHNOLOGY	
(30) Prioridad	(72) Interventor (es) Frances H. ARNOLD - Pete HEINZELMAN	
(31) No. Prioridad 61/166,993 – 61/177,882		
(32) Fecha 06.04.2009 - 13.05.2009		
(33) Pais E.U.A.	(74) Apoderado (s) LUIS FELIPE CASTILLO GIBSON	
(54) Título POLIPÉPTIDOS QUE TIENEN ACTIVIDAD DE CELULASA		
Las casillas identificadas con 85, 86 y 87 sólo aplican para PCT		
(85) Fecha límite inicio Fase Nacional		
(86) Datos relativos a la presentación de la solicitud PCT Fecha de presentación de la solicitud Fecha: 06/04/2010 No. de solicitud PCT: PCT/US2010/030133	(87) Publicación internacional Fecha Fecha: 14/10/2010 No. Publicación WO 2010/118058 A2	
Resumen reivindicación (es): 1. Un polipéptido recombinante que comprende una Serina en la región C terminal en un motivo que comprende la secuencia GEXDG, en donde X es C, A o G (SEQ ID NO: 107), en donde X está sustituido por Serina, en donde la variante comprende actividad de celulasa y en donde el polipéptido tiene termoestabilidad aumentada en comparación con una celulasa de tipo silvestre que carece de una Serina en la secuencia GEXDG (SEQ ID NO: 107).		

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800.176.089-2

-/-

Industria y Comercio
SUPERINTENDENCIA

RECIBO DE CAJA

No. 11 - 0094094

Bogotá D.C., Septiembre 26 de 2011 - 13:49:58

RECIBIDO DE : CASTILLO GRAU & ASOCIADOS LTDA.

NI 800.098.704

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	VR. PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	392736831	26/09/2011	2.894.000.00

*** Conceptos Pagados ***

CANT.	RENTISTICO	CONCEPTO	Vr. UNIDITARIO	Vr. CONCEPTO
1	50005-01-01 SOLICITUDES	1 TRAMITES DE SOL. DE PATENTE DE INVENCIÓN	571.000.00	571.000.00
				<u>\$571.000.00</u>

SON: **QUINIENTOS SETENTA Y UN MIL PESOS MONEDA CORRIENTE**

Responsable

Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-125473- -00000-0000

Fecha: 2011-09-26 14:15:31	Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTE IYII	Eve: 378 FASENACIONALI
Act. 411 PRESENTACION	Folios: 223
Act. 411 PRESENTACION	Folios: 214

Sede Centro: Carrera 13 No. 27 - 00 Pisos 3,4,5 y 10 Bogotá, D.C.- Colombia

Web: www.sic.gov.co e-mai: info@sic.gov.co Conmutador: (571) 5870000 Fax: (571) 5870284 Linea: 018000-910165 Call Center: (571) 6513240

Septiembre 26 de 2011 - 13:50:01

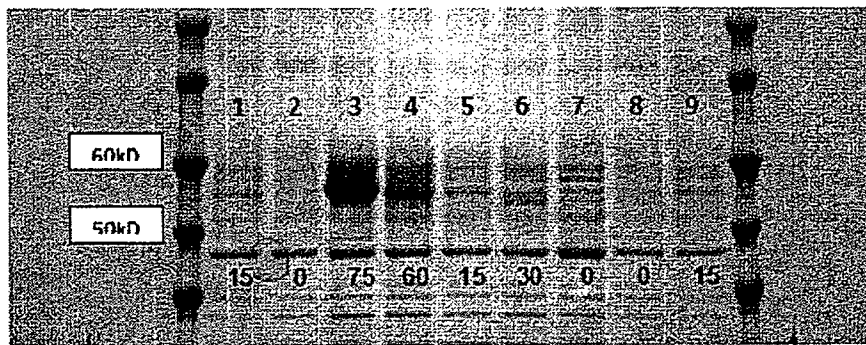


FIGURA 1A

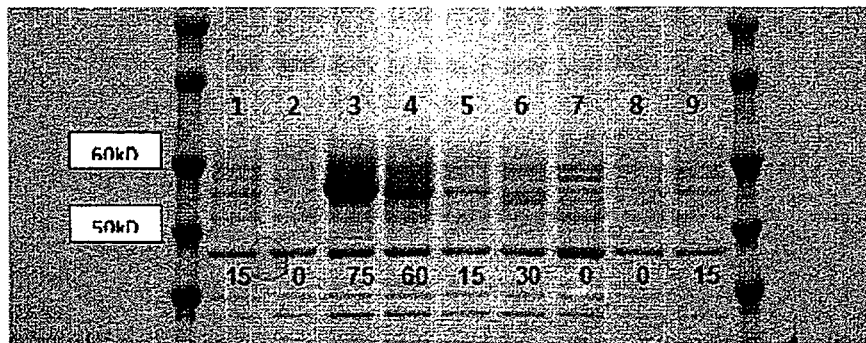


FIGURA 1A



No. 11-125473- -00000-0000

Fecha: 2011-09-26 14:15:31

Dep. 2020 DIR.NUEVASCR

Tra. 11 PATENTEIYII

Eve: 378 FASENACIONALI

Act. 411 PRESENTACION

Folios: 223

Act. 411 PRESENTACION

Folios: 214

PATENTE DE INVENCION ☐

MODELO DE

☐	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 ☐