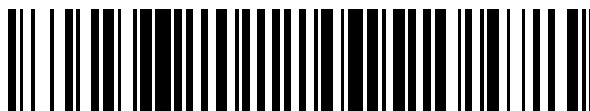


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 501 190**

51 Int. Cl.:

C12N 9/24 (2006.01)
C12N 1/20 (2006.01)
C12N 15/75 (2006.01)
C12P 21/06 (2006.01)
C12P 19/04 (2006.01)
C07H 21/04 (2006.01)
C08B 1/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **10.09.2004 E 04821802 (8)**

97 Fecha y número de publicación de la concesión europea: **18.06.2014 EP 1668125**

54 Título: **Enzimas modificadas, métodos para producir enzimas modificadas y usos de las mismas**

30 Prioridad:

15.09.2003 US 503251 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

01.10.2014

73 Titular/es:

**DANISCO US INC. (100.0%)
925 Page Mill Road
Palo Alto, California 94304, US**

72 Inventor/es:

**CLARKSON, KATHLEEN A. y
FENEL, FRED**

74 Agente/Representante:

RIZZO, Sergio

ES 2 501 190 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Enzimas modificadas, métodos para producir enzimas modificadas y usos de las mismas

CAMPO

5 [0001] Esta publicación afecta a las enzimas modificadas con una estabilidad aumentada en entornos duros industriales, como una temperatura y/o un pH aumentado.

ANTECEDENTES

10 [0002] Las xilanasas se han descubierto en al menos cien organismos diferentes. Las xilanasas son glicosil hidrolasas que hidrolizan cadenas de xilopiranosido con uniones β -1,4. Dentro de la clasificación basada en secuencia de las familias glicosil hidrolasas establecidas por Henrissat y Bairoch (1993), la mayoría de xilanasas se encuentran en las familias 10 y 11. Las características comunes para los miembros de la familia 11 incluyen una alta homología genética, un tamaño de aproximadamente 20 kDa y un mecanismo catalítico de doble desplazamiento (Tenkanen et al., 1992; Wakarchuk et al., 1994). Las familias se han agrupado ahora, basándose en similitudes de estructura, en Clanes (Henrissat and Davies, 1995). La familia 11 de glicosil hidrolasas, que son principalmente xilanasas, residen en el Clan C junto con la familia 12 de enzimas, conocidas todas por ser celulasas.

15 [0003] Las xilanasas pueden utilizarse a menudo para aplicaciones importantes como el blanqueamiento de la pulpa, la modificación de fibras textiles y en la alimentación animal (p.ej., las xilanasas pueden ayudar la digestión de un animal, Prade, 1996). Las xilanasas son útiles para la producción de comida para seres humanos también. Por ejemplo, la xilanasas mejora las propiedades de la masa de pan y la calidad del pan. Las xilanasas también pueden ayudar en el proceso de elaboración de cerveza mejorando la filtrabilidad de las cervezas que contienen xilano. Las xilanasas pueden emplearse en la descomposición de la materia vegetativa incluyendo la eliminación/el uso de residuos agrícolas y de residuos que resultan del procesamiento de los productos agrícolas, incluyendo la producción de fueles u otros productos/materiales de base biológica a partir de la biomasa.

20 [0004] Sin embargo, a menudo las condiciones extremas en estas aplicaciones, como una temperatura y/o pH altos, etc, hacen que las xilanasas rindan con menos efectividad que bajo condiciones normales. Durante el blanqueo de pulpa, por ejemplo, el material que viene de una fase de lavado alcalino puede tener una temperatura alta, a veces mayor a 80°C y un pH alto, como un pH más alto de 10. Ya que la mayoría de xilanasas no funcionan bien bajo esas condiciones, la pulpa debe enfriarse y el pH alcalino neutralizarse antes de que la xilanasas normal pueda funcionar. Tomando algunos de estos pasos en cuenta, el proceso puede volverse más caro ya que debe alterarse para adaptar la xilanasas.

25 [0005] En otro ejemplo, las xilanasas también son útiles en las aplicaciones de alimentos para animales. Ahí, las enzimas pueden enfrentarse a condiciones de temperaturas altas durante un periodo corto de tiempo (p.ej. -0,5 – 5 min a 95°C o más) durante la preparación de la alimentación. La inactivación de la enzima puede ocurrir bajo estas condiciones de temperatura y las enzimas resultan inservibles cuando se necesitan a una temperatura más baja como, por ejemplo, ~37 °C.

30 [0006] Se han descubierto xilanasas con cualidades mejoradas. Varias xilanasas termoestables, alcalofílicas y acidofílicas se han descubierto y clonado a partir de organismos termófilos (Bodie et al., 1995; Fukunaga et al., 1998). Sin embargo, suele resultar difícil producir las enzimas en cantidades económicamente eficientes. *T. reesei*, por otra parte, produce xilanasas, que no son termoestables como las xilanasas de los organismos termófilos. *T. reesei* es conocido por producir xilanasas diferentes de las que las xilanasas I y II (XynI y XynII, respectivamente) son las mejor caracterizadas (Tenkanen et al., 1992). XynI tiene un tamaño de 19 kDa, un pI de 5,5 y un pH entre 3 y 4. XynII tiene un tamaño de 20 kDa, un pI de 9,0 y un pH óptimo de 5,0-5,5 (Törrönen y Rouvinen, 1995). Estas xilanasas presentan un perfil de pH, especificidad y actividad específica favorables en una cantidad de aplicaciones y puede producirse de manera económica en procesos de producción a gran escala.

35 [0007] Se han realizado esfuerzos para producir una xilanasas con cualidades favorables. Por ejemplo, algunos han intentado mejorar la estabilidad de la xilanasas *Bacillus circulans* añadiendo puentes de disulfuro que unen el N-término de la proteína al C-término y la parte N-terminal de la hélice- α a la hebra- β cercana (Wakarchuk et al., 1994). Además, Campbell et al. (1995) modificó la xilanasas *Bacillus circulans* mediante enlaces inter e intra moleculares de disulfuro con tal de aumentar la termoestabilidad. De manera similar, la estabilidad de la *T. reesei* xilanasas II ha sido mejorada al cambiar la región N-terminal a una parte respectiva de una xilanasas termófila (Sung et al., 1998). Además de la termoestabilidad mejorada, el intervalo de actividad de la enzima se amplió

para incluir un pH alcalino. Las mutaciones de punto único también se han utilizado para aumentar la estabilidad de la xilanasa *Bacillus pumilus* (Arase *et al.*, 1993).

[0008] Al comparar las estructuras de las enzimas termófilas y mesófilas se ha obtenido mucha información (Vogt *et al.*, 1997). El análisis estructural de las xilanasas termófilas también ha dado información sobre los factores que influyen la termoestabilidad de la xilanasas (Gruber *et al.*, 1998; Harris *et al.*, 1997).

[0009] Kimura T. et al. (Bioscience, Biotechnology, Biochemistry, Japan Society for Bioscience and Agrochemistry, TOKYO, JAPÓN, vol. 64, núm. 6, 1 Junio 2000, páginas 1230-123) publica la purificación, caracterización y clonación molecular de una xilanasa acidófila a partir de *Penicillium sp.40*.

[0010] WO02/38746 publica la secuencia de una xilanasa alcalófila a partir de *Aspergillus kawachii*.

[0011] Sin embargo, actualmente, existe una necesidad de enzimas, especialmente xilanasas con propiedades mejoradas en condiciones industriales.

RESUMEN DE LA INVENCION

[0012] La presente publicación hace referencia a enzimas modificadas. Específicamente, la invención hace referencia a enzimas modificadas con un funcionamiento mejorado en condiciones extremas de pH y temperatura.

[0013] En un primer aspecto, la publicación describe una xilanasa modificada comprendiendo un polipéptido con una secuencia de aminoácido según se establece en la SEQ ID NO: 1, donde la secuencia tiene al menos un residuo de aminoácido sustituido en una posición seleccionada del grupo que consiste en: 2, 5, 7, 10, 11, 16, 19, 22, 26, 28, 29, 30, 34, 36, 38, 57, 58, 61, 63, 65, 67, 92, 93, 97, 105, 108, 110, 111, 113, 132, 143, 144, 147, 149, 151, 153, 157, 160, 162, 165, 169, 180, 184, 186, 188, 190 y +191. Preferiblemente, la sustitución se selecciona del grupo que consiste en: 2, 22, 28, 58, 65, 92, 93, 97, 105, 108, 144, 162, 180, 186 y +191. Preferiblemente, la xilanasa modificada tiene al menos una sustitución seleccionada del grupo que consiste en: H22K, S65C, N92C, F93W, N97R, V108H, H144C, H144K, F180Q y S186C. Además, preferiblemente, la xilanasa modificada muestra una termofilicidad, alcalofilicidad o una combinación de las mismas mejorada, en comparación con un tipo silvestre de xilanasas.

[0014] En un segundo aspecto, la publicación describe una enzima modificada, la enzima modificada comprendiendo una secuencia de aminoácido, la secuencia de aminoácido siendo homóloga a la secuencia establecida en SEQ ID NO:1, la secuencia de aminoácido teniendo al menos un residuo de aminoácido sustituido en una posición equivalente a una posición seleccionada del grupo que consiste en: 2, 5, 7, 10, 11, 16, 19, 22, 26, 28, 29, 30, 34, 36, 38, 57, 58, 61, 63, 65, 67, 92, 93, 97, 105, 108, 110, 111, 113, 132, 143, 144, 147, 149, 151, 153, 157, 160, 162, 165, 169, 180, 184, 186, 188, 190 y +191. En un modo de realización preferido, la secuencia de aminoácido tiene al menos un residuo de aminoácido sustituido en una posición equivalente a la posición seleccionada del grupo que consiste en: 2, 22, 28, 58, 65, 92, 93, 97, 105, 108, 144, 162, 180, 186 y +191. En un modo de realización preferido, la secuencia de aminoácido tiene al menos un residuo de aminoácido seleccionado del grupo que consiste en: H22K, S65C, N92C, F93W, N97R, V108H, H144C, H144K, F180Q y S186C.

[0015] En un modo de realización preferido, la enzima modificada es una glicosilhidrolasa del Clan C comprendiendo una secuencia de aminoácido, la secuencia de aminoácido siendo homóloga a la secuencia establecida en SEQ ID NO:1 la secuencia de aminoácido teniendo al menos un residuo de aminoácido sustituido en una posición equivalente a la posición seleccionada del grupo que consiste en: 2, 5, 7, 10, 11, 16, 19, 22, 26, 28, 29, 30, 34, 36, 38, 57, 58, 61, 63, 65, 67, 92, 93, 97, 105, 108, 110, 111, 113, 132, 143, 144, 147, 149, 151, 153, 157, 160, 162, 165, 169, 180, 184, 186, 188, 190 y +191. En un modo de realización preferido, la secuencia de aminoácido tiene al menos un residuo de aminoácido sustituido en una posición equivalente a una posición seleccionada del grupo que consiste en: 2, 22, 28, 58, 65, 92, 93, 97, 105, 108, 144, 162, 180, 186 y +191. En un modo de realización preferido, la secuencia de aminoácido tiene al menos un residuo de aminoácido sustituido seleccionado del grupo que consiste en: H22K, S65C, N92C, F93W, N97R, V108H, H144C, H144K, F180Q y S186C. Las enzimas modificadas preferidas se publican aquí.

[0016] En un modo de realización preferido, la enzima modificada es una xilanasas de familia 11 que comprende una secuencia de aminoácido, la secuencia de aminoácido siendo homóloga a la secuencia presentada en SEQ ID NO:1, la secuencia de aminoácido teniendo al menos un residuo de aminoácido sustituido en una posición equivalente a una posición seleccionada del grupo que consiste en: 2, 5, 7, 10, 11, 16, 19, 22, 26, 28, 29, 30, 34, 36, 38, 57, 58, 61, 63, 65, 67, 92, 93, 97, 105, 108, 110, 111, 113, 132, 143, 144, 147, 149, 151, 153, 157, 160, 162, 165, 169, 180, 184, 186, 188, 190 y +191. En un modo de realización preferido, la secuencia de aminoácido tiene al menos un residuo de aminoácido sustituido en una posición equivalente a una posición seleccionada del grupo que consiste en: 2, 22, 28, 58, 65, 92, 93, 97, 105, 108, 144, 162, 180, 186 y +191. En un modo de realización preferido, la secuencia de aminoácido tiene al menos un residuo de aminoácido sustituido

seleccionado del grupo que consiste en: H22K, S65C, N92C, F93W, N97R, V108H, H144C, H144K, F180Q y S186C. Las enzimas modificadas preferidas de la familia 11 se publican aquí.

5 **[0017]** En otro modo de realización preferido, la enzima modificada es una celulasa de familia 12 comprendiendo una secuencia de aminoácido, la secuencia de aminoácido siendo homóloga a la secuencia presentada en SEQ ID NO:1, la secuencia de aminoácido teniendo al menos un residuo de aminoácido sustituido en una posición equivalente a una posición seleccionada del grupo que consiste en: 2, 5, 7, 10, 11, 16, 19, 22, 26, 28, 29, 30, 34, 36, 38, 57, 58, 61, 63, 65, 67, 92, 93, 97, 105, 108, 110, 111, 113, 132, 143, 144, 147, 149, 151, 153, 157, 160, 162, 165, 169, 180, 184, 186, 188, 190 y +191. En un modo de realización preferido, la secuencia de aminoácido tiene al menos un residuo de aminoácido sustituido en una posición equivalente a una posición seleccionada del grupo que consiste en: 2, 22, 28, 58, 65, 92, 93, 97, 105, 108, 144, 162, 180, 186 y +191. En un modo de realización preferido, la secuencia de aminoácido tiene al menos un residuo de aminoácido sustituido seleccionado del grupo que consiste en: H22K, S65C, N92C, F93W, N97R, V108H, H144C, H144K, F180Q y S186C, donde la posición es una posición equivalente, según se define aquí. Las enzimas modificadas preferidas de familia 12 son las publicadas aquí.

15 **[0018]** En un modo de realización preferido, la celulasa de familia 12 es celulasa *Trichoderma* EGIII según se presenta en SEQ ID NO:3, la modificación comprende al menos un aminoácido seleccionado del grupo que consiste en: 2, 13, 28, 34, 77, 80, 86, 122, 123, 134, 137, 140, 164, 174, 183, 209, 215 y 218, la numeración de posición siendo con respecto a SEQ ID NO:3. En un modo de realización preferido, la sustitución es al menos una mutación seleccionada del grupo que consiste en T2C, N13H, S28K, T34C, S77C, P80R, S86C, G122C, 20 K123W, Q134H, Q134K, Q134R, V137H, G140C, N164C, N164K, N174C, K183H, N209C, A215D y N218C, la numeración de posición siendo con respecto a SEQ ID NO:3.

[0019] Los modos de realización del primer y segundo aspecto, según se publica arriba, también proporcionan ácidos nucleicos codificando cualquiera de las enzimas modificadas, según se presenta arriba, así como complementos. En otro modo de realización preferido, la publicación proporciona composiciones que comprenden al menos una enzima modificada, según se publica aquí y otro ingrediente. En otro modo de realización preferido, la publicación describe vectores que comprenden una enzima modificada, según se publica aquí, las células comprendiendo la enzima modificada y métodos de expresar la enzima modificada.

30 **[0020]** En un tercer aspecto, la publicación describe un método para modificar una enzima que comprende modificar un primer sitio en la enzima para que el primer sitio pueda unirse a un segundo sitio en la enzima. En un modo de realización preferido, el primer sitio es un lazo o secuencia adyacente a una lámina-β. En un modo de realización preferido, el segundo sitio se sitúa en una lámina-β.

35 **[0021]** En un modo de realización preferido, la enzima modificada es una xilanasa. Por ejemplo, en un modo de realización favorito, la publicación describe una xilanasa modificada, donde la xilanasa está modificada por al menos uno de los siguientes métodos: (i) modificando una secuencia N-terminal para que la secuencia N-terminal se una mediante un puente de disulfuro a una hebra-β adyacente; (ii) modificando una secuencia C-terminal para que la secuencia C-terminal se una a una hebra-β adyacente; (iii) modificando una hélice-α o secuencia adyacente a una hélice-α, para que la hélice-α, o una secuencia adyacente a la hélice-α se una más firmemente al cuerpo de la proteína; (iv) modificando una secuencia adyacente a la hebra-β para que la secuencia adyacente a la hebra-β pueda unirse más firmemente a una secuencia adyacente. Por ejemplo, en un modo de realización preferido, la modificación puede ocurrir en una hebra-β cerca del cordón.

BREVE DESCRIPCIÓN DE LAS FIGURAS

[0022]

45 La figura 1 muestra una alineación de aminoácido entre la familia 11 de xilanasas. La numeración de aminoácido se compara con la *T.Reesei* xilanasa II, según se indica en la parte superior de las secuencias. Los residuos comunes al menos a un 75% de la familia 11 de xilanasas están subrayados. Los siguientes están alineados (por abreviatura) en la figura: XYN2_TRIRE Endo-1,4-betaxilanasa 2 precursor (EC 3.2.1.8) (Xilanasa 2) (1,4-beta-D-xilano xilanohidrolasa 2) - *Trichoderma reesei* (*Hypocrea jecorina*) >sp|P36217|; XYN1_TRIRE Endo-1,4-beta-xilanasa 1 precursor (EC 3.2.1.8) (Xylanase 1) (1,4-beta-Dxylan xylanohidrolase 1) - *Trichoderma reesei* (*Hypocrea jecorina*) >sp|P36218|; XYN2_BACST Endo-1,4-betaxilanasa precursor (EC 3.2.1.8) (Xilanasa) (1,4-beta-D-xilano xilanohidrolasa) - *Bacillus stearothermophilus* >sp|P457031|; XYN1_HUMIN Endo-1,4-beta-xilanasa 1 precursor (EC 3.2.1.8) (Xilanasa 1) (1,4-beta-D-xylan xilanohidrolasa 1) - *Humicola insolens* >sp|P55334|; XYN1_ASPAW Endo-1,4-beta-xilanasa I precursor (EC 3.2.1.8)(Xilanasa I) (1,4-beta-D-xilano xilanohidrolasa I) - *Aspergillus awamori* >sp|P55328|; XYNA_BACST Endo-1,4-beta-xilanasa A precursor (EC 3.2.1.8) (Xilanasa A) (1,4-beta-D-xylan >sp|P45705|).

ES 2 501 190 T3

- La figura 2 muestra una alineación de aminoácido de la familia 12 de celulasas con Xynll. Las siguientes se alinean (por abreviatura) en la figura: 1ENX *XylanaseII Trichoderma reesei*, y miembros de la familia cel12 Q8NJY2 *Aspergillus awamori*, Q8NJJ3 *Humicola grisea*, Q8NJJ4 *Trichoderma viride*, Q8NJJ5 *Hypocrea koningii*, Q8NJJ6 *Hypocrea schweinitzii*, Q8NJJ7 *Stachybotrys echinata*, Q8NJJ8 *Bionectria ochroleuca*, Q8NJJ9 *Bionectria ochroleuca*, Q8NJJ0 *Bionectria ochroleuca*, Q8NJJ1 *Bionectria ochroleuca*, Q8NJJ2 *Fusarium solani* (subsp. *Cucurbitae*), Q8NJJ3 *Fusarium solani* (subsp. *cucurbitae*), Q8NJJ4 *Fusarium equiseti* (*Fusarium scirpi*), Q8NJJ5 *Emericella desertorum*, Q8NJJ6 *Chaetomium brasiliense*, Q9KIH1 *Streptomyces* sp. 11AG8. En la figura, las dos flechas indican la posición de los puentes de disulfuro (secuencia de señal no eliminada).
- La figura 3 muestra la secuencia nucleótida de *Trichoderma reesei* oligonucleótidos utilizados en la mutagénesis de la xilanasa, con los cambios de codón subrayados.
- La figura 4 muestra un gráfico que compara la actividad con respecto a la temperatura del tipo-silvestre Xynll con las xilanasas mutadas Y2 e Y5. Las xilanasas mutadas tienen las siguientes mutaciones: K58R y un ácido aspártico añadido a la serina C-terminal en la posición 190 (+191D) (=Y2); T2C, T28C, K58R +191D, (=Y5). La figura ejemplifica que un puente salino, sólo, no aumenta la termofilicidad ni la estabilidad térmica. En su lugar, la introducción de un puente de disulfuro aumenta la estabilidad y temperatura dependiendo de la actividad. La actividad se mide según *Bailey et al.*, 1992.
- La figura 5 muestra un gráfico comparando la actividad con respecto al pH del tipo silvestre de Xynll con la xilanasa mutada Y5 con las siguientes mutaciones: T2C, T28C, K58R con un ácido aspártico añadido a la serina C-terminal posición 190 (+191D). La actividad se mide según *Bailey et al.*, 1992.
- La figura 6 muestra un gráfico que compara la actividad con respecto a la temperatura del tipo silvestre de Xynll con la xilanasa mutada Y5 con las siguientes mutaciones: T2C, T28C, K58R con un ácido aspártico añadido a la serina C-terminal posición 190 (+191D). La actividad se mide según *Bailey et al.*, 1992.
- La figura 7 muestra un gráfico comparando la actividad residual con pH 5,0, con la inactivación con pH 8 con respecto a la temperatura del tipo silvestre de Xynll xilanasa con la xilanasa mutada Y5 teniendo las siguientes mutaciones: T2C, T28C, K58R con un ácido aspártico añadido a la serina C-terminal posición 190 (+191D). La actividad se mide según *Bailey et al.*, 1992.
- La figura 8 muestra un gráfico comparando la actividad residual en pH 5,3, con la inactivación en pH 8 con respecto a la temperatura de la xilanasa mutada Y5 con una Xynll xilanasa (SS105/162) teniendo las siguientes mutaciones adicionales Q162C y L105C. La actividad se mide según *Bailey et al.*, 1992.
- La figura 9 muestra un gráfico comparando la actividad residual en pH 5, con la inactivación en pH 9 con respecto a la temperatura de la xilanasa mutada Y5 con una Xynll xilanasa (P9) teniendo las siguientes mutaciones adicionales: F93W, N97R y H144K. La actividad se mide según *Bailey et al.*, 1992.
- La figura 10 muestra un gráfico comparando la actividad residual en pH 5, con inactivación en pH 5 con respecto a la temperatura de la xilanasa mutada Y5 con una Xynll xilanasa (P12) teniendo las siguientes mutaciones adicionales H144C y N92C. La actividad se mide según *Bailey et al.*, 1992.
- La figura 11 muestra un gráfico comparando la actividad residual en pH 5, con la inactivación en pH 9 con respecto a una temperatura de la xilanasa Y5 con una Xynll xilanasa (P12) teniendo las siguientes mutaciones adicionales H144C y N92C. La actividad se mide según *Bailey et al.*, 1992.
- La figura 12 muestra un gráfico comparando la actividad residual en pH 5,2 con la inactivación en pH 8 respecto a la temperatura de la xilanasa mutada Y5 con una xilanasa Xynll (P15) teniendo las siguientes mutaciones adicionales: F180Q, H144C y N92C. La actividad se mide según *Bailey et al.*, 1992.
- La figura 13 muestra un gráfico comparando la actividad residual en pH 5, con la inactivación en pH 8 con respecto a la temperatura de la xilanasa mutada Y5 con una xilanasa Xynll (P21) teniendo las siguientes mutaciones adicionales: H22K, F180Q, H144C y N92C. La actividad se mide según *Bailey et al.*, 1992.
- La figura 14 muestra un gráfico comparando la actividad residual en pH 5,17, con la inactivación en pH 7,8 con respecto a la temperatura de la xilanasa mutada Y5 con una Xynll xilanasa (P20) teniendo las siguientes mutaciones adicionales: H22K y F180Q. La actividad se mide según *Bailey et al.*, 1992.
- La figura 15 muestra un gráfico comparando la actividad en pH 8, con respecto a la temperatura de la xilanasa mutada Y5 con una xilanasa Xynll (J17) teniendo las siguientes mutaciones adicionales: V108H. La actividad se mide según *Bailey et al.*, 1992.

La figura 16 muestra un gráfico comparando la actividad en pH 8, con respecto a la temperatura de la xilanas mutada Y5 con una xilanas XynII (J21) teniendo las siguientes mutaciones adicionales: S65C y S186C (J21 en el gráfico). La actividad se mide según *Bailey et al.*, 1992.

5 La figura 17 muestra una alineación estructural de *Trichoderma reesei* xilanasell (XynII, PDB 1 ENX, en azul;) y *Trichoderma reesei* endoglucanasell (Cal12A, PDB 1H8V, en rojo).

La figura 18 establece un aminoácido nucleótido de secuencia de XynII.

La figura 19 establece un aminoácido nucleótido de secuencia de EGIII.

La figura 20 establece un aminoácido nucleótido de secuencia de XynII.

Descripción detallada de los modos de realización preferidos

10 **[0023]** A no ser que se defina lo contrario aquí, todos los términos técnicos y científicos utilizados aquí tienen el mismo significado que se suele entender por alguien con un conocimiento normal de la técnica. Singleton. *et al.*,
 15 **DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY**, 2D ED., John Wiley and Sons, New York (1994), y Hale & Marham, **THE HARPER COLLINS DICTIONARY OF BIOLOGY**, Harper Perennial, NY (1991) proporcionan un diccionario general de muchos términos utilizados en esta publicación. Pese a que cualquier
 20 método y material similar o equivalente a aquellos aquí descritos pueden utilizarse en la práctica o en los ensayos, se describen los métodos y materiales preferidos. Los intervalos numéricos incluyen los números que definen el intervalo. A no ser que se indique lo contrario, los ácidos nucleicos se escriben de izquierda a derecha en la orientación 5' a 3'; las secuencias de aminoácidos se escriben de izquierda a derecha en la orientación de amino a carboxi, respectivamente. Los practicantes se dirigen particularmente a Sambrook *et al.*, 1989, y Ausubel FM *et al.*, 1993, para definiciones y términos de la técnica. Debe entenderse que esta publicación no está limitada a la metodología particular, protocolos y reactivos descritos, ya que pueden cambiar.

[0024] Los títulos aquí proporcionados no son limitaciones de los varios aspectos o modos de realización que pueden tenerse como referencia a la especificación en su totalidad. En consecuencia, los términos definidos inmediatamente debajo son definidos con más detalle con referencia a la especificación en su totalidad.

25 **[0025]** Según se utiliza aquí, el término "polipéptido" hace referencia a un compuesto elaborado de una única cadena de residuos de aminoácidos unidos por enlaces péptidos. El término "proteína" aquí puede ser sinónimo del término "polipéptido" o puede hacer referencia, además, a un complejo de dos o más polipéptidos.

30 **[0026]** Según se utiliza aquí, el término "expresión" hace referencia al proceso por el que un polipéptido se produce basándose en la secuencia de ácido nucleico del gen. El proceso incluye tanto la transcripción como la traducción.

[0027] Según se utiliza aquí, el término "gen" se refiere al segmento de ADN involucrado en producir una cadena de polipéptido, que puede incluir o no regiones que preceden o siguen a la región de codificación.

35 **[0028]** Según se utiliza aquí, con referencia a una numeración de posición, el término "equivalente" hace referencia a posiciones según se determina mediante las alineaciones de secuencia y estructura con *Trichoderma reesei* xilanas II (xynII) como una secuencia de referencia o una estructura de referencia, según se proporciona aquí (véase, por ejemplo, la figura 2 para una alineación de secuencia múltiple y *Trichoderma reesei* xilanasall con otras secuencias y la figura 17 para una alineación estructural de *Trichoderma reesei* Xyn II con *Trichoderma reesei* endoglucanasall). La numeración de posición deberá ser con respecto a *Trichoderma reesei* xynII, según se establece en SEQ ID NO:1. El sistema de numeración, pese a que puede utilizar una secuencia específica como un punto de referencia base, también es aplicable a todas las secuencias homólogas relevantes. La homología de secuencia entre proteínas puede determinarse utilizando programas de alineación reconocidos y según se describe aquí utilizando técnicas de hibridación aquí descritas.

45 **[0029]** Según se utiliza aquí, el término "adyacente" hace referencia a una proximidad cercana lineal y/o espacial cercana entre los residuos de aminoácido o regiones o áreas de una proteína. Por ejemplo, un primer residuo o una primera región o una primera área que es adyacente a un segundo residuo o una segunda región o una segunda área (en un sentido lineal), respectivamente, deberá tener preferiblemente aproximadamente 7, preferiblemente aproximadamente 5, preferiblemente aproximadamente 2 residuos de aminoácido que intervienen entre ellos. De manera alternativa, por ejemplo, cuando un primer conjunto de residuos o una primera región o un primer área es adyacente a un segundo conjunto de residuos o una segunda región o un segundo área, entonces el primer conjunto de residuos o primera región o primer área puede ser proximal (en espacio, según se muestra, por ejemplo, mediante la estructura terciaria de una proteína) al segundo conjunto de residuos

o segunda región o segunda área. Alguien especializado en la técnica, cuando fuera posible, sabría cómo solucionar la estructura terciaria de una proteína.

- 5 **[0030]** Según se utiliza aquí, con referencia a las posiciones de secuencia, la designación “+” seguida por un entero significa que un polipéptido ha sido modificado para incluir aminoácidos adicionales en la posición putativa, según se especifica mediante el entero. Por ejemplo, la designación +191 significa que un polipéptido que normalmente tiene 190 aminoácidos en la secuencia de aminoácido tiene un aminoácido añadido.
- [0031]** Según se utiliza aquí, el término “molécula de ácido nucleico” incluye moléculas de ARN, ADN y ADNc. Se entenderá que como resultado de la degeneración del código genético, pueden producirse una multitud de secuencias nucleótidas que codifican una proteína dada, como las proteínas mutantes aquí descritas.
- 10 **[0032]** Según se utiliza aquí, el término “puente de disulfuro” o “enlace de disulfuro” hace referencia al enlace formado entre los átomos de sulfuro y los residuos de cisteína en un polipéptido o una proteína. En esta publicación, un puente de disulfuro o enlace de disulfuro puede no ocurrir de manera natural y ser introducido mediante mutación de punto.
- 15 **[0033]** Según se utiliza aquí, el término “puente salino” hace referencia al enlace formado entre residuos cargados de manera opuesta, los aminoácidos en un polipéptido o proteína. En esta publicación, un puente salino puede no ocurrir de manera natural e introducirse mediante mutación de punto.
- [0034]** Según se utiliza aquí, una “enzima” hace referencia a una proteína o polipéptido que cataliza una reacción química.
- [0035]** Según se utiliza aquí, el término “actividad” hace referencia a una actividad biológica asociada a una proteína particular, como una actividad enzimática asociada a una proteasa. La actividad biológica hace referencia a cualquier actividad que normalmente le atribuiría a esa proteína alguien especialista en la técnica.
- 20 **[0036]** Según se utiliza aquí, el término “xilanasas” hace referencia a glicosil hidrolasas que hidrolizan cadenas de xilopiranosida con uniones β -1,4.
- [0037]** Según se utiliza aquí, el término “XynI” hace referencia a *Trichoderma reesei* xilanasas, xilanasas I. XynI tiene el tamaño de 19 kDa, un pI de 5,5 y un pH óptimo entre 3 y 4.
- 25 **[0038]** Según se utiliza aquí, “XynII” hace referencia a *Trichoderma reesei* xilanasas, xilanasas II. XynII tiene un tamaño de 20 kDa, un pI de 9,0 y un pH óptimo entre 5 y 5,5.
- [0039]** Según se utiliza aquí, “xilopiranosida” hace referencia a un polímero de xilosa con uniones β -1,4, incluyendo polímeros sustituidos de xilosa, es decir, polímeros de xilofiranosas ramificados con uniones β -D-1,4, altamente sustituidos con grupos acetilo, arabinosil y uronil (véase, por ejemplo, Biely, P. (1985) *Microbial Xylanolytic Systems*. Trends Biotechnol., 3, 286-290.).
- 30 **[0040]** Según se utiliza aquí, el término “glicosil hidrolasa” hace referencia a una enzima que hidroliza el enlace glicosídico entre dos o más carbohidratos o entre un carbohidrato y una fracción no carbohidrato. La hidrólisis enzimática del enlace glicosídico se produce mediante la catálisis de ácido general y requiere dos residuos críticos: un donante protón y un nucleófilo/base. La nomenclatura enzimática IUB-MB de las glicosil hidrolasas se basa en la especificidad del sustrato y ocasionalmente en el mecanismo molecular.
- 35 **[0041]** Según se utiliza aquí, el término “hidrolizar” hace referencia a una enzima que cataliza una reacción donde un enlace químico se adhiere de manera enzimática con la adición de una molécula de agua.
- [0042]** Según se utiliza aquí, “hidrólisis” hace referencia al proceso de reacción por la que un enlace químico se adhiere con la adición de una molécula de agua.
- 40 **[0043]** Según se utiliza aquí, “Clan C” hace referencia a agrupaciones de familias que comparten un pliegue tridimensional y una maquinaria catalítica idéntica (véase, por ejemplo, Henrissat, B. y Bairoch, A., (1996) *Biochem. J.*, 316, 695-696).
- [0044]** Según se utiliza aquí, “familia 11” hace referencia a una familia de enzimas según se establece por Henrissat y Bairoch (1993) *Biochem J.*, 293, 781-788 (véase también, Henrissat y Davies (1997) *Current Opinion in Structural Biol.* 1997, &:637-644). Las características comunes para los miembros de la familia 11 incluyen una alta homología genética, un tamaño de aproximadamente 20 kDa y un mecanismo catalítico de doble desplazamiento (véase Tenkanen *et al.*, 1992; Wakarchuk *et al.*, 1994). La estructura de la familia 11 de xilanasas incluye dos grandes láminas- β elaboradas con hebras- β y hélices- α . La familia 11 de xilanasas

incluyen lo siguiente: *Aspergillus niger* XynA, *Aspergillus kawachii* XynC, *Aspergillus tubigensis* XynA, *Bacillus circulans* XynA, *Bacillus pumilus* XynA, *Bacillus subtilis* XynA, *Neocallimastix patriciarum* XynA, *Streptomyces lividans* XynB, *Streptomyces lividans* XynC, *Streptomyces thermoviolaceus* XynII, *Thermomonospora fusca* XynA, *Trichoderma harzianum* Xyn, *Trichoderma reesei* XynI, *Trichoderma reesei* XynII, *Trichoderma viride* Xyn.

5 **[0045]** Según se utiliza aquí, la “familia 12” hace referencia a una familia de enzimas establecida por Henrissat y Bairoch (1993) en la que glicosil hidrolasas conocidas se clasificaron en familias basadas en similitudes de las secuencias de aminoácido. Hasta la fecha, todas las enzimas de la familia 12 son celulasas. Las enzimas de la familia 12 hidrolizan el enlace glucosídico con uniones β -1,4 en la celulosa mediante una reacción de doble desplazamiento y un intermediario de glicosil-enzima que resulta en una retención de la configuración anomérica del producto. Los estudios estructurales de los miembros de familia 12 revelan una estructura compacta tipo sándwich- β que se curva para crear un extenso sitio de enlace en el sustrato sobre la cara cóncava de la lámina- β .

[0046] Según se utiliza aquí, el término “proteasa” hace referencia a una enzima que se degrada hidrolizando al menos algunos de sus enlaces péptidos.

15 **[0047]** Según se utiliza aquí, “enlace péptido” hace referencia al enlace químico entre el grupo carbonil de un aminoácido y el grupo amino de otro aminoácido.

[0048] Según se utiliza aquí, “tipo silvestre” hace referencia a una secuencia o una proteína que es nativa u ocurre de manera natural.

20 **[0049]** Según se utiliza aquí, “mutación de punto” hace referencia a un cambio en un único nucleótido de ADN, especialmente donde ese cambio resulta en un cambio en una proteína.

[0050] Según se utiliza aquí, “mutante” hace referencia a una versión de un organismo o proteína donde la versión es otra diferente al tipo silvestre. El cambio puede verse afectado por métodos bien conocidos por aquellos especialistas en la técnica, por ejemplo, mediante mutación de punto en la que la proteína resultante puede referirse como una mutante.

25 **[0051]** Según se utiliza aquí, “mutagénesis” hace referencia al proceso que produce un cambio de tipo silvestre a mutante.

30 **[0052]** Según se utiliza aquí, “sustituido” y “modificado” se utilizan de manera intercambiable y se refieren a una secuencia, como una secuencia de aminoácido que comprende un polipéptido, que incluye la supresión, inserción, reemplazo o interrupción de una secuencia que ocurre de manera natural. A menudo en el contexto de la invención, una secuencia sustituida se refiere, por ejemplo, al reemplazo de una secuencia que ocurre de manera natural.

[0053] Según se utiliza aquí, “enzima modificada” hace referencia a una enzima que incluye la supresión, inserción, reemplazo o interrupción de una secuencia que ocurre de manera natural.

35 **[0054]** Según se utiliza aquí, “hebras- β ” hace referencia a esa parte de una secuencia de aminoácido que forma una secuencia lineal que ocurre en una lámina- β .

[0055] Según se utiliza aquí, “láminas- β ” hace referencia a una estructura de tipo lámina que resulta cuando los aminoácidos de hidrógeno se enlazan entre ellos para formar una estructura tipo lámina.

40 **[0056]** Según se utiliza aquí, “hélice- α ” hace referencia a la estructura que resulta cuando una única cadena de polipéptido gira sobre sí misma regularmente para hacer un cilindro rígido en el que cada enlace péptido es hidrógeno regular unido a otros enlaces péptidos en la cadena cercana.

[0057] Según se utiliza aquí, “pulgar” hace referencia a un lazo entre las hebras- β B7 y B8 en XynI y en XynII (véase, por ejemplo, en Torronen, A. y Rouvinen, J.; Biochemistry 1995, 34, 847-856).

45 **[0058]** Según se utiliza aquí, “cordón” hace referencia a un lazo entre las hebras- β B7 y B8 que hacen un pulgar y una parte del lazo entre las hebras- β B6a y B9 que cruzan la grieta en un lado (véase, por ejemplo, Torronen, A. y Rouvinen, J.; Biochemistry 1995, 34, 847-856).

[0059] Según se utiliza aquí, “alcalino” hace referencia al estado o cualidad de ser básico.

[0060] Según se utiliza aquí, “alcalofilo” hace referencia a la cualidad de ser más robusto en una atmósfera alcalina que un miembro no alcalofilo. Por ejemplo, un organismo alcalofilo hace referencia a un organismo que

sobrevive o prospera bajo condiciones alcalinas donde un organismo normal no puede, y una proteína alcalofila es una cuya actividad es activa o más robusta bajo condiciones alcalinas donde una proteína normal sería menos activa.

[0061] Según se utiliza aquí, "ácido" hace referencia al estado o cualidad de ser ácido.

5 **[0062]** Según se utiliza aquí, "acidófilo" hace referencia a la cualidad de ser más robusto en una atmósfera ácida que un miembro no acidófilo. Por ejemplo, un organismo acidófilo hace referencia a un organismo que sobrevive o prospera bajo condiciones ácidas donde un organismo normal no puede, y una proteína acidófila es una cuya actividad es activa o más robusta bajo condiciones ácidas donde una proteína normal sería menos activa.

10 **[0063]** Según se utiliza aquí, "termoestable" hace referencia a la cualidad de ser estable en una atmósfera que implica temperatura. Por ejemplo, un organismo termoestable es uno que es más estable bajo condiciones de temperatura específicas que un organismo no termoestable.

[0064] Según se utiliza aquí "termoestabilidad" hace referencia a la cualidad de ser termoestable.

15 **[0065]** Según se utiliza aquí "termófilo" hace referencia a la cualidad de ser más robusto en una atmósfera caliente que un miembro no termófilo. Por ejemplo, un organismo termófilo hace referencia a un organismo que sobrevive o prospera bajo condiciones calientes donde un organismo normal no puede, y una proteína termófila es una cuya actividad es activa o más robusta bajo condiciones calientes donde una proteína normal sería menos activa.

20 **[0066]** Según se utiliza aquí, "mesófilo" hace referencia a la cualidad de ser más robusto en una atmósfera normal que un miembro no mesófilo. Por ejemplo, un organismo mesófilo hace referencia a un organismo que sobrevive o prospera bajo condiciones normales donde otro organismo no puede, y una proteína mesófila es una cuya actividad es activa o más robusta bajo condiciones normales donde otra proteína sería menos activa.

[0067] Según se utiliza aquí, "oligonucleótidos" hace referencia a una secuencia nucleótida corta que puede utilizarse, por ejemplo, como un cebador en una reacción utilizada para crear proteínas mutantes.

25 **[0068]** Según se utiliza aquí "codón" hace referencia a una secuencia de tres nucleótidos en una molécula de ADN o de ARNm que representa la instrucción para incorporar un aminoácido específico en una cadena de polipéptido.

[0069] Según se utiliza aquí "Y5" hace referencia a una xilanasas mutante según se publica, por ejemplo, en la publicación número WO 01/27252.

[0070] Según se utiliza aquí, las siguientes designaciones hacen referencia a los siguientes mutantes:

30 "P2" = N97R + H144K / Y5
 "P3" = F93W + H144K en Y5
 "P8" = F180Q en Y5
 "P9" = N97R in F93W + H144K en Y5
 "P12" = H144C + N92C en Y5
 35 "P15" = F180Q en H144C + N92C en Y5
 "P16" = N97R en H144C + N92C en Y5
 "P18" = H22K en Y5
 "P20" = H22K + F 180Q in Y5
 "P21" = H22K + F180Q + H144C + N92C en Y5
 40 "J17" = V108H en Y5
 "J21" = S65C + S186C en Y5

donde la numeración de posición es con respecto a XynII.

[0071] La presente publicación hace referencia a enzimas modificadas con un funcionamiento mejorado en condiciones extremas, como la temperatura y el pH.

45 **[0072]** En un primer aspecto, la publicación describe una xilanasas modificada comprendiendo un polipéptido con una secuencia de aminoácido según se presenta en SEQ ID NO:1, donde la secuencia tiene al menos un residuo de aminoácido sustituido en una posición seleccionada del grupo que consiste en: 2, 5, 10, 11, 16, 19, 22, 26, 28, 29, 30, 34, 36, 38, 57, 58, 61, 63, 65, 67, 92, 93, 97, 105, 108, 110, 111, 113, 132, 143, 144, 147, 149, 151, 153, 157, 160, 162, 165, 169, 180, 184, 186, 188, 190 y +191, donde la numeración de posición es con respecto a
 50 SEQ ID NO:1. Preferiblemente, la sustitución se selecciona del grupo que consiste en: 2, 22, 28, 58, 65, 92, 93,

97, 105, 108, 144, 162, 180, 186 y +191. Preferiblemente, la xilanasa modificada tiene al menos una sustitución seleccionada del grupo que consiste en H22K, S65C, N92C, F93W, N97R, V108H, H144C, H144K, F180Q y S186C. Además, preferiblemente, la xilanasa modificada muestra una termofilia, alcalofilia o una combinación de estas mejoradas, en comparación con una xilanasa de tipo silvestre.

5 **[0073]** En un segundo aspecto, la publicación describe una enzima modificada, la enzima modificada comprendiendo una secuencia de aminoácido, la secuencia de aminoácido siendo homóloga a la secuencia presentada en SEQ ID NO:1, la secuencia de aminoácido teniendo al menos un residuo de aminoácido sustituido en una posición equivalente a la posición seleccionada del grupo que consiste en: 2, 5, 7, 10, 11, 16, 19, 22, 26, 28, 29, 30, 34, 36, 38, 44, 57, 58, 61, 63, 65, 67, 92, 93, 97, 105, 108, 110, 111, 113, 132, 143, 144, 147, 149,
10 151, 153, 157, 160, 162, 165, 169, 180, 184, 186, 188, 190 y +191, donde la numeración de posición es con respecto a SEQ ID NO:1. En un modo de realización preferido, la secuencia de aminoácido tiene al menos un residuo de aminoácido en una posición equivalente a una posición seleccionada del grupo que consiste en: 2, 22, 28, 58, 65, 92, 93, 97, 105, 108, 144, 162, 180, 186 y +191. En un modo de realización preferido, la secuencia de aminoácido tiene al menos un residuo de aminoácido sustituido en una posición equivalente a una posición seleccionada del grupo que consiste en: 2, 22, 28, 58, 65, 92, 93, 97, 105, 108, 144, 162, 180, 186 y +191. En un modo de realización preferido, la secuencia de aminoácido tiene al menos un residuo de aminoácido sustituido seleccionado del grupo que consiste en: H22K, S65C, N92C, F93W, N97R, V108H, H144C, H144K, F180Q y S186C.

20 **[0074]** En un modo de realización preferido, la enzima modificada es una glicosil hidrolasa del Clan C comprendiendo una secuencia de aminoácido, la secuencia de aminoácido siendo homóloga a la secuencia presentada en SEQ ID NO:1, la secuencia de aminoácido teniendo al menos un residuo de aminoácido sustituido en una posición equivalente a una posición seleccionada del grupo que consiste en 2, 5, 7, 10, 11, 16, 19, 22, 26, 28, 29, 30, 34, 36, 38, , 57, 58, 61, 63, 65, 67, 92, 93, 97, 105, 110, 108, 110, 111, 113, 132, 143, 144, 147, 149, 151, 153, 157, 160, 162, 165, 169, 180, 184, 186, 188, 190 y +191. En un modo de realización preferido, la secuencia de aminoácido tiene al menos un residuo de aminoácido en una posición equivalente a una posición seleccionada del grupo que consiste en: 2, 22, 28, 58, 65, 92, 93, 97, 105, 108, 144, 162, 180, 186 y +191. En un modo de realización preferido, la secuencia de aminoácido tiene al menos un residuo de aminoácido sustituido seleccionado del grupo que consiste en: H22K, S65C, N92C, F93W, N97R, V108H, H144C, H144K, F180Q y S186C. Las enzimas modificadas preferidas son según se publican aquí.

30 **[0075]** En un modo de realización preferido, la enzima modificada es una xilanasa de familia 11 que comprende una secuencia de aminoácido, la secuencia de aminoácido siendo homóloga a la secuencia presentada en SEQ ID NO:1, la secuencia de aminoácido teniendo al menos un residuo de aminoácido sustituido en una posición equivalente a una posición seleccionada del grupo que consiste en: 2, 5, 7, 10, 11, 16, 19, 22, 26, 28, 29, 30, 34, 36, 38, 57, 58, 61, 63, 65, 67, 92, 93, 97, 105, 108, 110, 111, 113, 132, 143, 144, 147, 149, 151, 153, 157, 160,
35 162, 165, 169, 180, 184, 186, 188, 190 y +191. En un modo de realización preferido, la secuencia de aminoácido tiene al menos un residuo de aminoácido sustituido en una posición equivalente a una posición seleccionada del grupo que consiste en: 2, 22, 28, 58, 65, 92, 93, 97, 105, 108, 144, 162, 180, 186 y +191. En un modo de realización preferido, la secuencia de aminoácido tiene al menos un residuo de aminoácido sustituido seleccionado del grupo que consiste en: H22K, S65C, N92C, F93W, N97R, V108H, H144C, H144K, F180Q y S186C. Las enzimas de familia 11 modificadas preferidas son según se publican aquí.

45 **[0076]** En otro modo de realización preferido, la enzima modificada es una celulasa de familia 12 comprendiendo una secuencia de aminoácido, la secuencia de aminoácido siendo homóloga a la secuencia presentada en SEQ ID NO:1, la secuencia de aminoácido teniendo al menos un residuo de aminoácido sustituido en una posición equivalente a una posición seleccionada del grupo que consiste en: 2, 5, 7, 10, 11, 16, 19, 22, 26, 28, 29, 30, 34, 36, 38, 57, 58, 61, 63, 65, 67, 92, 93, 97, 105, 108, 110, 111, 113, 132, 143, 144, 147, 149, 151, 153, 157, 160, 162, 165, 169, 180, 184, 186, 188, 190 y +191. En un modo de realización preferido, la secuencia de aminoácido tiene al menos un residuo de aminoácido sustituido en una posición equivalente a una posición seleccionada del grupo que consiste en: 2, 22, 28, 58, 65, 92, 93, 97, 105, 108, 144, 162, 180, 186 y +191. En un modo de realización preferido, la secuencia de aminoácido tiene al menos un residuo de aminoácido sustituido seleccionado del grupo que consiste en: H22K, S65C, N92C, F93W, N97R, V108H, H144C, H144K, F180Q y S186C. Las enzimas de familia 12 modificadas preferidas son según se publicada aquí.

50 **[0077]** En un modo de realización preferido, la celulasa de familia 12 es *Trichoderma* EGIII celulasa según se presenta en SEQ ID NO:3, la modificación comprende al menos un aminoácido seleccionado del grupo que consiste en: 2, 13, 28, 34, 77, 80, 86, 122, 123, 134, 137, 140, 164, 174, 183, 209, 215 y 218, la numeración de posición siendo con respecto a SEQ ID NO:3. En un modo de realización preferido, la sustitución es al menos una mutación seleccionada del grupo que consiste en T2C, N13H, S28K, T34C, S77C, P80R, S86C, G122C, K123W, Q134H, Q134K, Q134R, V137H, G140C, N164C, N164K, N174C, K183H, N209C, A215D y N218C, , la numeración de posición siendo con respecto a SEQ ID NO:3.

[0078] XynII muestra una homología de aminoácido significativa con otros miembros de familia 11, aproximadamente 20-90%, así como una similitud estructural general. La homología, según se utiliza aquí, puede estar determinada por alguien especializado en la técnica; específicamente, se contemplan las homologías de al menos 20%, preferiblemente 30% o más, preferiblemente 40% o más, preferiblemente 50% o más, preferiblemente 60% o más, preferiblemente 70% o más, preferiblemente 80% o más, preferiblemente 90% o más, preferiblemente 95% o más y preferiblemente 97% o más (según se calcula en el nivel de aminoácido y el nivel nucleótido y según se utiliza aquí). Existen similitudes estructurales entre las enzimas de familia 11 y de familia 12. Las proteínas beta tienen dos láminas beta apiladas y una hélice alfa envasada contra una de las láminas beta formando una estructura de rollo llamada "beta-jelly". (véase Stirk, H.J., Woolfson, D.N., Hutchison, E.G. y Thornton, J.M. (1992) *Depicting topology and handedness in jellyroll structures*. FEBS Letters 308 p1-3).

[0079] Basándose en esta similitud estructural, ambas familias de enzimas han sido asignadas a una "súper familia" llamada Clan C (véase Sandgren, M. et. al., J. Mol. Bio. (2001) 308, 295-310.)).

[0080] Pese a que la homología de secuencia entre las familias 11 y 12 es baja, la similitud estructural general de las dos familias es destacable según se ve al comparar las figuras 2 y 16. La longitud de los lazos que conectan las dos láminas-beta comprende las mayores diferencias estructurales entre las familias (Sandgren et. al., J. Mol., Biol., 2001). Actualmente, ninguna enzima de familia 11 es conocida por contener puentes de disulfuro N terminal mientras que muchas celulasas de familia 12, en general parecen contener un puente de disulfuro cerca del N-término (p.ej., entre los residuos 4 y 32 en *T. reesei* Cel 12A). Ese puente de disulfuro en las enzimas de familia 12 se sitúa cerca de la posición donde un disulfuro se introdujo en la VARIANTEe Trichoderma (Y5), pese a estar lejos del N-término (véase, por ejemplo, la publicación WO 01/27252). La importancia de una restricción que establezca la región N-terminal de las enzimas de familia 11 se examinó en *Trichoderma reesei* xylanase II (XynII). Al insertar un puente de disulfuro no natural entre los residuos (TSC y T28C), se consiguió un aumento en T_m de 11°C. En estas dos familias de estructura similar, la familia 11 y la familia 12, los puentes de disulfuro N-terminal juegan roles similares en cuanto a la estabilidad. Esto se ha demostrado reemplazando la cisteína en la posición 32 con una alanina en Cel12A resultando en una disminución significativa en T_m de 18,5°C. De manera interesante, la magnitud del cambio en la estabilidad por añadir un disulfuro N-terminal no natural en XynII es comparable a la de extraer uno natural de Cel 12A (véase la tabla A).

Tabla A

Enzima	Delta Tm	Tm (grados C)
WT Cel12A		54,4
C32A	-18,5	35,9
WT xynII		58,6
Y5	+10,7	69,3

La tabla A muestra las temperaturas de fundición, T_m del tipo silvestre de Cel12A comparado con la VARIANTEe con la sustitución en la posición 32 y el tipo silvestre XynII comparado con la VARIANTEe Y5 de esta enzima.

[0081] Las estructuras tridimensionales de los puentes de disulfuro N-terminal de las tres estructuras públicamente conocidas para las glicosil hidrolasas de familia 12 (*Trichoderma reesei*- PDB 1H8V, *Aspergillus niger*- PDB 1KS5, *Streptomyces lividans*-PDB 2NLR), muestran un cambio en la posición del puente de disulfuro en comparación con el puente de disulfuro no natural en los sitios 2 y 28 en xilanasa Y5. La Tabla B muestra la posición del puente de disulfuro en una xilanasa Y5 ("PDB 1ENX" siendo una xilanasa XynII de tipo silvestre) y en las tres estructuras de familia conocidas 12. Las posiciones estructurales de las mutaciones en 2 y 28 de xilanasa Y5 pueden traducirse en los residuos correspondientes en las estructuras Cel 12. En cada caso, el disulfuro no nativo de Y5 está más cercano al N-término; y para la estructura de *A. niger* (PDB 1KS5) un disulfuro podría diseñarse para utilizar el residuo N-terminal por sí mismo (en residuos Q1C, V35C, de acuerdo con la numeración de *A. niger*). En lugar de estar limitado por la secuencia natural, los datos de rayos x podrían utilizarse para diseñar extensiones y truncamientos del N-término para facilitar disulfuros no nativos que se adhieren específicamente a los nuevos residuos N-terminal.

Tabla B

Código	WTN-terminal S-S posición	Sitio correspondiente a 2-28 en xynII	Donde(de acuerdo con la estructura) podría una S-S insertarse en la N-terminal
PDB 1ENX	No	-	
Y5	C2-C28	T2-T28	T2C-T28C
PDB 1H8V	C4-C32	T2-T34	T2C-T34C
PDB 1KS5	C4-C32	T2-Y34	Q1C-V35C
PDB 2NLR	C5-C31	T3-T33	T3C-T33C

5 **[0082]** Un gran número de secuencias de familia 12 (Tabla C) son conocidas por poder estar estabilizadas potencialmente a través de un puente de disulfuro N-terminal, particularmente aquellas moléculas donde un puente de disulfuro no nativo podría introducirse o un disulfuro nativo podría moverse más cerca de un N-término. La Tabla C enumera una serie de secuencias donde la extracción predecible se la secuencia de señal produce secuencias de proteína madura muy similares a aquellas de la conocida familia 12 de estructuras. La Tabla C también enumera la distancia entre las dos cisteínas N-terminal (26-28 aminoácidos) similares al enlace de disulfuro de Y5. En las predicciones de división de sitio, unas secuencias de señal se eliminan en teoría mediante parámetros conocidos declarados (véase, por ejemplo, "*Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites*". Henrik Nielsen, Jacob Engelbrecht, Søren Brunak and Gunnar von Heijne, Protein Engineering 10, 1-6 (1997)).

10 **[0083]** Un gran grupo de secuencias de estructuras tridimensionales desconocidas en la Tabla C caen dentro del grupo estructuralmente similar de las enzimas de familia 12, que tienen de manera similar un residuo de cisteína en la N-terminal en el sitio 5 +/- 2 residuos, formando un puente de disulfuro con residuo 32 +/- 7, para que la primera hebra beta o hebras de la lámina beta puedan unirse a la lámina beta adyacente. Todas estas secuencias podrían tratarse de la manera descrita en el debate en torno a la Tabla B para mejorar la estabilidad.

Tabla C

ID	Secuencia	Eucariota/Gram-/Gram+	Sitio de división pronosticada	Número de cisteína adecuada (1ª en enlace ss)	aa's hacia la 2ª cisteína en el enlace ss
Q8NJY2	Endoglucanasa{GEN:CEL12B} Aspergillus awamori (var. kawachi)	Eu	16-17	6	28
Q8NJY4	Endoglucanasa{GEN:CEL12A} - Trichoderma viride	Eu	16-17	4	28
Q8NJY5	Endoglucanasa{GEN:CEL12A} - Hypocrea koningii	Eu	16-17	4	28
Q8NJY6	Endoglucanasa{GEN:CEL12A} - Hypocrea schweinitzii	Eu	16-17	4	28
Q8NJY7	Endoglucanasa{GEN:CEL12A} - Stachybotrys echinata	Eu	16-17	4	28
Q8NJY8	Endoglucanasa{GEN:CEL12D} - Bionectria ochroleuca	Eu	17-18	4	28
Q8NJY9	Endoglucanasa{GEN:CEL12C} - Bionectria ochroleuca	Eu	17-18	3	28
Q8NJZ1	Endoglucanasa{GEN:CEL12A} - Bionectria ochroleuca	Eu	18-19	4	28
Q8NJZ4	Endoglucanasa{GEN:CEL12A} - Fusarium equiseti (Fusarium scirpi)	Eu	17-18	4	28
Q9KIH1	Cellulasa 12A{GEN:CEL12A} - Streptomyces sp. 1AG8	Gram+	31-32	5	26

20 **[0084]** La tabla D enumera además un número de secuencias de enzimas de familia 12 con una secuencia de señal sin descomponerse. Todas tienen cisteínas 30-39 aminoácidos aparte, y tras la extracción de la secuencia de señal (la extracción puede realizarse como en la tabla C) son capaces estructuralmente de formar un puente de disulfuro en la N-terminal (según se muestra en las estructuras públicamente conocidas, véase la tabla B). El sitio de mutación propuesto tiene correlación con el sitio correspondiente del puente de disulfuro entre los sitios 2-28 del mutante Y5. Las secuencias de glicosil hidrolasa se alinearon utilizando el programa MOE (de Chemical Computing Corp) utilizando métodos de casación de secuencia estándar.

25

Tabla D

Código de secuencia	Enzima	Especies	Mutaciones
Tr O94218	Cel 12	Aspergillus aculeatus	D22C/ G52C
Sp P22669	Cel 12	Aspergillus aculeatus	Q20C/ T52C
Sp Q12679	Cel 12	Aspergillus awamori	T18C/ Y50C
Tr O13454	Cel 12	Aspergillus oryzae	E18C/ Y50C
Sp P16630	Cel 12	Erwina carotovora	A32C/ I68C
Tr O31030	Cel 12	Pectobacterium carotovora	A32C/ V68C
Tr Q9V2TO	Cel 12	Pyrococcus furiosus	P57C/ T96C
Tr O33897	Cel 12	Rhodothermus marinus	E40C / E70C
Tr Q9RJY3	Cel 12	Streptomyces coelicolor	T43C/ T73C
Tr O08468	Cel 12	Streptomyces halstedii	L40C/ T70C
Tr Q59963	Cel 12	Streptomyces rochei	T40C/ T70C
Tr Q9KIH1	Cel 12	Streptomyces sp. 11AG8	Q34C/ N64C
Tr Q60032	Cel 12	Thermotoga maritima	V2C/ K38C
Tr Q60033	Cel 12	Thermotoga maritima	V20C/ K56C
Tr O08428	Cel 12	Thermotoga neopolitana	V2C/ R38C
Tr P96492	Cel 12	Thermotoga neopolitana	V20C / K56C
AF435072	Cel12A	Aspergillus Kawachi	Q20C / T52C
AF434180	Cel12A	Chaetium brasiliense	S28C/ Y61C
AF434181	Cel12A	Emericella desertorum	D30C/ G63C
AF434182	Cel12A	Fusarium equiseti	D19C/ H51C
AF434183	Cel12A	Nectria ipomoeae	Q25C / T58C
AF434184	Cel12B	Nectria ipomoeae	T32C / T65C
AF435063	Cel12A	Bionectria ochroleuca	T20C / Y52C
AF435064	Cel12B	Bionectria ochroleuca	T34C / T66C
AF435065	Cel12C	Bionectria ochroleuca	A18C / T50C
AF435066	Cel12D	Bionectria ochroleuca	S19C / Y51C
AF435071	Cel12A	Humicola grisea	S34C / Y67C
AF435068	Cel12A	Hypochrea schweinitzii	T18C / T50C
AF435067	Cel12A	Stachybotrys echinata	S 18C / Y50C

- 5 **[0085]** No sólo muestra la región N-terminal una alta similitud estructural entre familias 11 y 12; ambas familias muestran una estructura similar a una mano, la de una “mano derecha parcialmente cerrada” según se describe en Törrönen et al. 1997. Las dos láminas- β forman “dedos” y un par retorcido de una lámina- β y la hélice- α forma una “palma”. El lazo largo entre las hebras- β B7 y B8 hace el “pulgar” y una parte del lazo entre las hebras- β B6b (residuos 95-102 en xynII y 125-131 en Cel12A) y B9 forma un “cordón”, que cruza la grieta en un lado (Törrönen A. y Rouvinen, J. Biochem. 1995, 34, 847-0856). El efecto estabilizador de insertar sustituciones de agentes
- 10 especiales que le dan rigidez entre la hebra beta B6b y el lazo adyacente y/o el “cordón” se ve en la mutación en los sitios 92, 93, 144 (N92C-H144C, al menos una de las siguientes mutaciones N97R, F93W + H144K (XynII), y puede de manera similar ser introducido en los sitios correspondientes en la familia 12.
- 15 La tabla E muestra la numeración de una selección de sitios equivalentes estructuralmente entre xynII y Cel12A. La alta similitud estructural entre las dos familias permite un gran número de sustituciones similares (véase Sandgren et. al., Mol., Biol., 2001 para la comparación estructural).

Tabla E

Ejemplos de ubicaciones equivalentes	
XynII	Cel12A
T2C	T2C
T28C	T34C
N92C	G122C
H144C, K	N164C, K
F93W	K123W
Q162H	K183H

- 20 **[0086]** Las enzimas modificadas pueden comprender una o más mutaciones además de aquellas presentadas arriba. Otras mutaciones como las supresiones, inserciones, sustituciones, transversiones, transiciones e inversiones, en una o más ubicaciones, también pueden incluirse. De manera similar, la enzima modificada puede no tener al menos una de las sustituciones arriba expuestas.

[0087] La enzima modificada también puede comprender una sustitución conservadora que puede ocurrir como una sustitución de comparación (p.ej., básico por básico, ácido por ácido, polar por polar, etc.). Las sustituciones no conservadoras también pueden ocurrir, es decir, de una clase de residuo a otra o alternativamente involucrando la inclusión de aminoácidos antinaturales como ornitina, ornitina de ácido diaminobutírico, ornitina norleucina, pirlalanina, tienilalanina, naftil alanina y fenilglicina.

[0088] Las secuencias también pueden tener supresiones, inserciones o sustituciones de residuos de aminoácidos que producen un cambio silencioso y resultan en una sustancia equivalente funcionalmente. Las sustituciones de aminoácido deliberadas pueden estar hechas sobre la base de similitud en las propiedades en aminoácido (como la polaridad, carga, solubilidad, hidrofobicidad, hidrofiliidad, y/o la naturaleza anfipática de los residuos) y por lo tanto resulta útil agrupar aminoácidos juntos en grupos funcionales. Los aminoácidos pueden agruparse basándose en las propiedades sólo de su cadena lateral. Sin embargo, resulta más útil incluir los datos de mutación también. Es probable que se conserven los conjuntos de aminoácidos derivados así por razones estructurales. Estos conjuntos pueden describirse en la forma de un diagrama Venn (Livingstone C.D. and Barton G.J. (1993) "*Protein secuencia alignments: a strategy for the hierarchical analysis of residue conservation*" Comput.Appl Biosci. 9: 745-756) (Taylor W.R. (1986) "*The classification of amino acid conservation*" J.Theor.Biol. 119; 205-218). Las sustituciones conservadoras pueden llevarse a cabo, por ejemplo de acuerdo con la tabla abajo que describe un diagrama Venn generalmente aceptado para agrupar aminoácidos.

Conjunto		Subconjunto	
Hidrofóbico	F W Y H K M I L V A G C	Aromático	F W Y H
		Alifático	I L V
Polar	W Y H K R E D C S T N Q	Cargado	H K R E D
		Cargado positivamente	H K R
		Cargado negativamente	E D
Pequeño	V C A G S P T N D	Diminuto	A G S

[0089] Las secuencias de aminoácido VARIANTEs también pueden incluir grupos espaciadores adecuados insertados entre cualquier dos residuos aminoácidos de la secuencia incluyendo los grupos alquil como los grupos metil, etil o propil además de los espaciadores aminoácidos como los residuos de alanina-β o glicina. Una forma adicional de variación incluye la presencia de uno o más residuos de aminoácido en forma peptóide.

[0090] Las comparaciones de homología pueden llevarse a cabo con la vista, o más generalmente, con la ayuda de programas de comparación de secuencias fácilmente disponibles. Estos programas de ordenador disponibles comercialmente pueden calcular el % de homología entre dos o más secuencias. El % de homología puede calcularse en secuencias contiguas, es decir, una secuencia se alinea con la otra secuencia y cada aminoácido en una secuencia se compara directamente con el aminoácido correspondiente en la otra secuencia un residuo cada vez. Esto se llama un alineamiento "sin espacios". Normalmente, dichos alineamientos sin espacios se llevan a cabo sólo con un número relativamente corto de residuos.

[0091] Pese a que este es un método muy simple y consistente, falla al no tener en cuenta que, por ejemplo, en un par de secuencias de otra manera idénticas, una inserción o eliminación provocará que los siguientes residuos de aminoácido se pongan fuera del alineamiento, potencialmente resultando así en una gran reducción en % de homología cuando se lleva a cabo un alineamiento global. En consecuencia, la mayoría de métodos de comparación de secuencia están diseñados para producir alineamientos óptimos que tienen en cuenta posibles inserciones y eliminaciones sin penalizar indebidamente la puntuación de homología total. Esto se consigue insertando "espacios" en la alineación de secuencia para intentar maximizar la homología local.

[0092] Sin embargo, estos métodos más complejos asignan "penalizaciones de espacio" a cada espacio que ocurre en la alineación para que, para el mismo número de aminoácidos idénticos, una alineación de secuencia con el menor número de espacios posible - reflejando una afinidad mayor entre las dos secuencias comparadas - conseguirá una puntuación más alta que uno con muchos espacios. Suelen utilizarse "costes de espacios afines" que cargan un coste relativamente alto por la existencia de un espacio y una penalización más pequeña por cada residuo posterior en el espacio. Este es el sistema de puntuación de espacio más comúnmente utilizado. Las altas penalizaciones de espacio producirán por supuesto alineaciones optimizadas con menos espacios. La mayoría de programas de alineación permiten modificar las penalizaciones de espacio. Sin embargo, es preferible utilizar los valores por defecto cuando se utiliza dicho software para comparaciones de secuencia. Por ejemplo cuando se utiliza el paquete *GCG Wisconsin Bestfit* la penalización del espacio por defecto para las secuencias de aminoácidos es -12 por un espacio y -4 por cada extensión.

[0093] El cálculo de máximo % de homología por lo tanto requiere en primer lugar la producción de un alineamiento óptimo, teniendo en cuenta las penalizaciones de espacio. Un programa de ordenador adecuado para llevar a cabo dicho alineamiento es el paquete *GCG Wisconsin Bestfit* (Devereux et al 1984 Nuc. Acids Research 12 p387). Ejemplos de otro software que puede ejecutar comparaciones de secuencia incluyen, sin carácter limitativo, el paquete BLAST (véase Ausubel et al., 1999 *Short Protocols in Molecular Biology*, 4ª Ed – capítulo 18), FASTA (Altschul et al., 1990 J. Mol. Biol. 403-410) y la suite GENWORKS de herramientas de comparación. Tanto BLAST como FASTA están disponibles para búsquedas offline y online (véase Ausubel et al., 1999, *Short Protocols in Molecular Biology*, páginas 7-58 a 7-60). Sin embargo, para algunas aplicaciones, es preferible utilizar el programa *GCG Bestfit*. *BLAST 2 Secuencias* también está disponible para comparar una secuencia nucleótida y una proteína (véase *FEMS Microbiol Lett* 1999 174(2): 247-50; *FEMS Microbiol Lett* 1999 177(1): 187-8 y tatiana@ncbi.nlm.nih.gov).

[0094] Pese a que el % de homología final puede medirse en términos de identidad, el proceso de alineación por sí mismo no está comúnmente basado en una comparación de par todo-o-nada. En su lugar, se utiliza generalmente una matriz de puntuación de similitud a escala que asigna puntuaciones a cada comparación del tipo par basándose en una similitud química o en la distancia evolutiva. Un ejemplo de dicha matriz comúnmente utilizada es la matriz BLOSUM62 – la matriz por defecto para la suite de programas BLAST. Los programas *GCG Wisconsin* generalmente utilizan valores por defecto públicos o una tabla de comparación de símbolo personalizada si se suministra (véase el manual de usuario para detalles adicionales). Para algunas aplicaciones, es preferible utilizar los valores de defecto públicos para el paquete GCG, o en el caso de otro software, la matriz de defecto, como BLOSUM62.

[0095] De manera alternativa, el porcentaje de homologías puede calcularse utilizando la función de alineamiento múltiple en ADNSIS™ (Hitachi Software), basada en un algoritmo, análogo a CLUSTAL (Higgins DG & Sharp PM (1988), Gen 73(1), 237-244).

[0096] Una vez que el software ha producido una alineación óptima, es posible calcular el % de homología, preferiblemente un % de identidad de secuencia. El software normalmente hace esto como parte de la comparación de secuencia y genera un resultado numérico.

[0097] Los modos de realización del primer y segundo aspecto, según se publican arriba, proporcionan un ácido nucleico que codifica cualquiera de las enzimas modificadas, según se establece arriba, así como los complementos de la misma. En otro modo de realización preferido, la publicación proporciona composiciones que comprenden al menos una enzima modificada, según se publica aquí y otro ingrediente. En otro modo de realización preferido, la publicación proporciona vectores comprendiendo una enzima modificada, según se publica aquí, las células comprendiendo la enzima modificada y métodos de expresar la enzima modificada.

[0098] Alguien especializado en la técnica tendrá en cuenta la relación entre la secuencia de ácido nucleico y la secuencia de polipéptido, en particular, el código genético y la degeneración de este código, y será capaz de construir dichas enzimas modificadas sin dificultad. Por ejemplo, alguien especializado en la técnica tendrá en cuenta que para cada sustitución de aminoácido en la secuencia de enzima modificada puede haber uno o más codones que codifican el aminoácido sustituto. En consecuencia, será evidente que, dependiendo de la degeneración del código genético con respecto a ese residuo de aminoácido particular, una o más secuencias de ácido nucleico de enzima modificada pueden generarse correspondiendo con esa secuencia de polipéptido de enzima modificada.

[0099] Las mutaciones en la secuencia de aminoácido y la secuencia de ácido nucleico pueden realizarse mediante cualquier número de técnicas, según se conoce en el campo. En los modos de realización particularmente preferidos, las mutaciones se introducen en secuencias parentales mediante una PCR (reacción en cadena de la polimerasa) utilizando cebadores apropiados, como se ilustra en los Ejemplos. Las enzimas parentales pueden modificarse en el nivel de aminoácido o en el nivel de ácido nucleico para generar las secuencias de enzima modificada aquí descritas. Por lo tanto, un modo de realización preferido proporciona la generación de enzimas modificadas introduciendo uno o más cambios de codón correspondiente en la secuencia nucleótida que codifica una enzima modificada.

[0100] Se apreciará que los cambios de codón arriba podrían realizarse en cualquier secuencia de ácido nucleico de enzima modificada. Por ejemplo, los cambios de secuencia pueden realizarse en cualquiera de las secuencias homólogas aquí descritas.

[101] La enzima modificada puede comprender la enzima “completa”, es decir, en su longitud total como ocurre en la naturaleza (o como mutada), o puede comprender una forma truncada de la misma. La enzima modificada derivada de esta puede por lo tanto ser truncada o ser de “longitud-completa”: El truncamiento puede estar en el

extremo N-terminal o en el extremo C-terminal. La enzima modificada puede carecer de una o más partes, como sub-secuencias, secuencias de señal, dominios y fracciones, tanto activas como no.

[0102] Una secuencia nucleótida que codifica una enzima que tiene las propiedades específicas según se definen aquí o una enzima que es adecuada para la modificación, como una enzima modificada, puede ser identificada y/o aislada y/o purificada a partir de cualquier célula u organismo que produce dicha enzima. Varios métodos son conocidos en la técnica para identificar y/o aislar y/o purificar las secuencias nucleótidas. Como modo de ejemplo, las técnicas de amplificación PCR para preparar más de una secuencia pueden utilizarse una vez que una secuencia adecuada se haya identificado y/o aislado y/o purificado.

[0103] Como modo de ejemplo adicional, un banco de datos ADNc y/o ADN genómico puede construirse utilizando ADN cromosómico o ARN mensajero a partir del organismo que produce la enzima. Si la secuencia de aminoácido de la enzima o una parte de la secuencia de aminoácido de la enzima se conoce, las sondas etiquetadas de oligonucleótidos pueden estar sintetizadas y utilizadas para identificar los clones de codificación-enzima a partir del banco de datos genómico preparado a partir del organismo. De manera alternativa, una sonda etiquetada de oligonucleótido que contiene secuencias homólogas a otro gen de enzima conocido podría utilizarse para identificar clones de codificación-enzima. En el último caso, se utilizan condiciones de lavado y de hibridación menos exigentes.

[0104] De manera alternativa, los clones de codificación-enzima podrían identificarse insertando fragmentos de ADN genómico en un vector de expresión, como un plásmido, transformando las bacterias de enzima-negativa con el banco de datos de ADN genómico resultante y después revistiendo la bacteria transformada sobre placas de agar que contienen un sustrato para la enzima permitiendo así identificar los clones que expresan la enzima.

[0105] En una alternativa adicional, la secuencia nucleótida que codifica la enzima modificada puede prepararse de manera sintética estableciendo los métodos estándar, p.ej., el método fosforamidita descrito por Beucage S.L. et al., (1981) *Tetrahedron Letters* 22, p 1859-1869 o el método descrito por Matthes et al., (1984) *EMBO J.* 3, p 801-805. En el método fosforamidita, los oligonucleótidos se sintetizan, p.ej., en un sintetizador de ADN automático, purificado, recocado, ligado y clonado en vectores apropiados.

[0106] La secuencia nucleótida puede ser de origen mixto genómico y sintético, de origen mixto sintético y ADNc o de origen mixto genómico y ADNc, preparado al ligar fragmentos de origen sintético, genómico o de ADNc de acuerdo con técnicas estándar. Cada fragmento ligado corresponde con varias partes de la secuencia nucleótida entera. La secuencia de ADN también puede prepararse mediante la reacción en cadena de la polimerasa (PCR) utilizando cebadores específicos, por ejemplo los descritos en US 4,683,202 o en Saiki R K et al., (*Science* (1988) 239, pp 487-491).

[0107] Las secuencias nucleótidas aquí descritas y adecuadas para el uso en los métodos y composiciones aquí descritos pueden incluir dentro de ellas nucleótidos modificados o sintéticos. Un número de diferentes tipos de modificación a los oligonucleótidos son conocidos en la técnica. Estos incluyen los ejes centrales de metilfosfonato y fosforotioato y/o la adición de acridina o cadenas de polilisina en los extremos 3' y/o 5' de la molécula. A los efectos de este documento, debe entenderse que las secuencias nucleótidas aquí descritas pueden ser modificadas por cualquier método disponible en la técnica. Dichas modificaciones pueden llevarse a cabo para aumentar la actividad *in vivo* o la esperanza de vida de las secuencias nucleótidas.

[0108] Un modo de realización preferido proporciona secuencias nucleótidas y el uso de secuencias nucleótidas que son complementarias a las secuencias aquí presentadas, o cualquier derivado, fragmento o derivado de la misma. Si la secuencia es complementaria a un fragmento de la misma entonces esa secuencia puede utilizarse como sonda para identificar secuencias de codificación similares en otros organismos, etc.

[0109] Los polinucleótidos que no son 100% homólogos a las secuencias de enzima modificada pueden obtenerse de muchas maneras. Otras VARIANTES de las secuencias aquí descritas pueden obtenerse por ejemplo investigando bancos de datos de ADN elaborados a partir de un intervalo de individuos, por ejemplo individuos de diferentes poblaciones. Además, otros homólogos pueden obtenerse y tales homólogos y fragmentos de los mismos en general serán capaces de hibridarse de manera selectiva a las secuencias mostradas en la lista de secuencia aquí. Dichas secuencias pueden obtenerse investigando los bancos de datos de ADNc elaborados de bancos de datos de ADN genómico a partir de otras especies e investigando dichos bancos de datos con sondeos que comprenden toda o parte de cualquiera de las secuencias en los listados de secuencia adheridos bajo condiciones de exigencia mediana a alta. Consideraciones similares se aplican para obtener especies homólogas y VARIANTES alélicas de las secuencias nucleótidas o polipéptidas aquí descritas.

[0110] Las VARIANTES y variedades/especies homólogas también pueden obtenerse utilizando una PCR degenerada que utilizará cebadores diseñados para apuntar a las secuencias dentro de las VARIANTES y

homólogas codificando secuencias de aminoácido conservadas. Los cebadores utilizados en la PCR degenerada contendrán una o más posiciones degeneradas y se utilizará en condiciones de exigencia menor a aquellas utilizadas para las secuencias de clonación con cebadores de secuencia única contra las secuencias conocidas. Las secuencias conservadas pueden pronosticarse, por ejemplo, alineando las secuencias de aminoácido a partir de algunas VARIANTEes/homólogos. Las alineaciones de secuencia pueden realizarse utilizando un software de ordenador conocido en la técnica aquí descrita.

[0111] De manera alternativa, dichos polinucleótidos pueden obtenerse mediante mutagénesis dirigida al sitio de secuencias caracterizadas, según se proporcionan aquí. Esto puede resultar útil cuando, por ejemplo, los cambios de secuencia de codón se requieren para optimizar las preferencias de codón para una célula anfitriona particular en la que las secuencias polinucleótidas se expresan. Otros cambios de secuencia pueden desearse con tal de introducir sitios de reconocimiento de restricción de enzima, o para alterar la propiedad o función de los polipéptidos codificados por los polinucleótidos.

[0112] Los polinucleótidos pueden utilizarse para producir un cebador, p.ej. un cebador PCR, un cebador para una reacción de amplificación alternativa, una sonda p.ej. etiquetada con una etiqueta de identificación mediante un medio convencional utilizando etiquetas radioactivas o no radioactivas o los polinucleótidos pueden clonarse en vectores. Dichos cebadores, sondas y otros fragmentos serán al menos 15, preferiblemente al menos 20, por ejemplo al menos 25, 30 o 40 nucleótidos en longitud y también están abarcados por el término polinucleótidos.

[0113] Los polinucleótidos como los polinucleótidos de aDN y sondas pueden producirse de manera recombinante o sintética o mediante cualquier medio disponible para aquellos especialistas en la técnica. También pueden clonarse mediante técnicas estándar. En general, los cebadores se producirán mediante medios sintéticos, incluyendo una fabricación gradual de la secuencia de ácido nucleico deseada con un nucleótido cada vez. Las técnicas para conseguir esto utilizando técnicas automatizadas están fácilmente disponibles en la técnica.

[0114] Los polinucleótidos más largos se producirán generalmente utilizando medios recombinantes, por ejemplo utilizando técnicas de clonación PCR (reacción en cadena de la polimerasa). Los cebadores pueden diseñarse para contener sitios de reconocimiento de restricción de enzima adecuados para que el ADN amplificado pueda clonarse en un vector de clonación adecuado. Preferiblemente, las secuencias VARIANTEes son al menos tan biológicamente activas como las secuencias aquí presentadas.

[0115] Un modo de realización preferido incluye secuencias que son complementarias a la enzima modificada o secuencias que son capaces de hibridar tanto en secuencias nucleótidas de las enzimas modificadas (incluyendo secuencias complementarias de aquellas aquí presentadas), así como secuencias nucleótidas que son complementarias a secuencias que pueden hibridar en secuencias nucleótidas de las enzimas modificadas (incluyendo secuencias complementarias de aquellas aquí presentadas). Un modo de realización preferido proporciona secuencias polinucleótidas que son capaces de hibridar en secuencias nucleótidas aquí presentadas bajo condiciones de exigencia intermedia a máxima.

[0116] Un modo de realización preferido incluye secuencias nucleótidas que pueden hibridar en la secuencia nucleótida del ácido nucleico de enzima modificada, o el complemento del mismo, bajo condiciones exigentes (p.ej., 50°C y 0,2xSSC). Más preferiblemente, las secuencias nucleótidas pueden hibridar en la secuencia nucleótida de la enzima modificada, o el complemento de la misma, bajo condiciones de exigencia alta (p.ej., 65°C y 0,1xSSC).

[0117] Puede resultar deseable mutar la secuencia con tal de preparar una enzima modificada. Por consiguiente, un mutante puede prepararse a partir de las enzimas modificadas aquí proporcionadas. Las mutaciones pueden introducirse utilizando oligonucleótidos sintéticos. Estos oligonucleótidos contienen secuencias nucleótidas acompañando los sitios de mutación deseados. Un método adecuado se publica en Morinaga et al., (*Biotechnology* (1984) 2, p646-649). Otro método de introducir mutaciones en secuencias nucleótidas de codificación-enzima se describe en Nelson and Long (*Analytical Biochemistry* (1989), 180, p 147-151). Un método adicional se describe en Sarkar and Sommer (*Biotechniques* (1990), 8, p404-407 - "*The megaprimer method of site directed mutagenesis*"). Otros métodos para mutar la secuencia se emplean y publican aquí.

[0118] En un modo de realización preferido, la secuencia para su uso en los métodos y composiciones aquí descritos es una secuencia recombinante – es decir, una secuencia que ha sido preparada utilizando técnicas de ADN recombinante. Dichas técnicas se explican, por ejemplo, en la bibliografía, por ejemplo, J. Sambrook, E. F. Fritsch, y T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual, Second Edition, Books 1-3*, Cold Spring Harbor Laboratory Press.

[0119] Otro modo de realización proporciona composiciones y formulaciones que comprenden enzimas modificadas. Las composiciones incluyen la enzima modificada junto con otro componente.

[0120] Otro modo de realización proporciona vectores que comprenden la enzima modificada, células que comprenden la enzima modificada y métodos para expresar la enzima modificada. La secuencia nucleótida para su uso en métodos y composiciones aquí descritas pueden incorporarse en un vector replicable recombinante. El vector puede utilizarse para replicar y expresar la secuencia nucleótida, en forma de enzima, en y/o desde una célula anfitriona compatible. La expresión puede controlarse utilizando secuencias de control, p.ej., secuencias reguladoras. La enzima producida por una célula recombinante anfitriona mediante la expresión de la secuencia nucleótida puede segregarse o puede contenerse de manera intra celular dependiendo de la secuencia y/o del vector utilizado. Las secuencias de codificación pueden diseñarse con secuencias de señal que dirigen la secreción de las secuencias de codificación de sustancia a través de una membrana de célula procariota o eucariota particular. Los polinucleótidos pueden incorporarse en un vector replicable recombinante. El vector puede utilizarse para replicar el ácido nucleico en una célula anfitriona compatible. El vector comprendiendo la secuencia polinucleótida puede transformarse en una célula anfitriona adecuada. Las anfitrionas adecuadas pueden incluir células bacterianas, de levadura, de insecto y fúngicas.

[0121] Las enzimas modificadas y sus polinucleótidos pueden expresarse introduciendo un polinucleótido en un vector replicable, introduciendo el vector en una célula anfitriona compatible y cultivando la célula anfitriona bajo condiciones que provocan la réplica del vector. El vector puede recuperarse de la célula anfitriona.

[0122] El ácido nucleico de enzima modificada puede estar unido de manera operativa a elementos reguladores de transcripción y traducción activos en una célula anfitriona de interés. El ácido nucleico de enzima modificada también puede codificar una proteína de fusión que comprende secuencias de señal como, por ejemplo, aquellas derivadas del gen de glucoamilasa a partir de *Schwanniomyces occidentalis*, gen del tipo acoplamiento al factor- α a partir de *Saccharomyces cerevisiae* y la amilasa-TAKA a partir de *Aspergillus oryzae*. De manera alternativa, el ácido nucleico de enzima modificada puede codificar una proteína de fusión que comprende un dominio vinculante a la membrana.

[0123] La enzima modificada puede expresarse en los niveles deseados en un organismo anfitrión utilizando un vector de expresión. Un vector de expresión comprendiendo un ácido nucleico de enzima modificada puede ser cualquier vector capaz de expresar el gen que codifica el ácido nucleico de enzima modificada en el organismo anfitrión seleccionado y la elección del vector dependerá de la célula anfitriona en la que se va a introducir. Así, el vector puede ser un vector que se replica de manera autónoma, es decir un vector que existe en una entidad episomal, cuya replicación es independiente a la replicación cromosomal, como, por ejemplo, un plásmido, un bacteriófago o un elemento episomal, un minicromosoma o un cromosoma artificial. De manera alternativa, el vector puede ser uno que, cuando se introduce en una célula anfitriona, se integra en el genoma de célula anfitriona, se integra en un genoma de célula anfitriona y se replica junto con el cromosoma.

[0124] El vector de expresión típicamente incluye los componentes de un vector de clonación, como, por ejemplo, un elemento que permite la replicación autónoma del vector en el organismo anfitrión seleccionado y uno o más marcadores detectables fenotípicamente con el objetivo de la selección. El vector de expresión normalmente comprende secuencias nucleótidas de control codificando un promotor, operador, sitio vinculante de ribosoma, señal de iniciación de traducción y opcionalmente, un gen represor o uno o más genes de activación. De manera adicional, el vector de expresión puede comprender una secuencia que codifica una secuencia de aminoácido capaz de dirigir la enzima modificada a un orgánulo celular anfitrión como un peroxisoma o a un compartimento de célula anfitriona particular. Dicha secuencia dirigida incluye, sin carácter limitativo, la secuencia SKL. Para la expresión bajo la dirección de secuencias de control, la secuencia de ácido nucleico la enzima modificada está enlazada de manera operable a las secuencias de control de una manera apropiada respecto a la expresión.

[125] Preferiblemente, un polinucleótido en un vector está enlazado de manera operable a una secuencia de control que es capaz de proporcionar la expresión de la secuencia de codificación mediante la célula anfitriona, es decir, el vector es un vector de expresión. Las secuencias de control pueden modificarse, por ejemplo, mediante la adición de elementos reguladores transcripcionales adicionales para hacer que el nivel de transcripción dirigido por las secuencias de control sea más receptivo a los moduladores transcripcionales. Las secuencias de control pueden comprender en particular promotores.

[0126] En el vector, la secuencia de ácido nucleico de codificación para la enzima modificada está combinada de manera operable con una secuencia promotora adecuada. El promotor puede ser cualquier secuencia de ADN con una actividad de transcripción en el organismo anfitrión de elección y puede derivarse de los genes que son homólogos o heterólogos al organismo anfitrión. Los ejemplos de promotores adecuados para dirigir la transcripción de la secuencia nucleótida modificada, como los ácidos nucleicos de enzima modificada, en un

anfitrión bacteria incluyen el promotor del *lac* operon de *E. coli*, los promotores *dagA* del gen agarasa *Streptomyces coelicolor*, los promotores del *Bacillus licheniformis* α -amilasa gen (*amyL*), el promotor *aprE* de *Bacillus subtilis*, los promotores del *Bacillus stearothermophilus* gen amilasa maltogénica (*amyM*), los promotores del *Bacillus amyloliquefaciens* α -gen amilasa (*amyQ*), los promotores de los *Bacillus subtilis* *xylA* y *xylB* genes y un promotor derivado de un *Lactococcus* sp.-promotor derivado incluyendo el promotor P170. Cuando el gen que codifica la enzima modificada se expresa en una especie bacteria como *E. coli*, un promotor adecuado puede seleccionarse, por ejemplo, de un promotor bacteriófago incluyendo un promotor T7 un promotor lambda fago. Para la transcripción en una especie fúngica, ejemplos de promotores útiles son aquellos derivados de los genes que codifican la *Aspergillus oryzae* TAKA amilasa, *Rhizomucor miehei* proteinasa aspártica, *Aspergillus niger* neutral α -amilasa, *A. niger* ácido estable α -amilasa, *A. niger glucoamylase*, *Rhizomucor miehei lipase*, *Aspergillus oryzae* proteasa alcalina, *Aspergillus oryzae* isomerasa fosfato triosa o *Aspergillus nidulans* acetamidasa. Los ejemplos de promotores adecuados para la expresión en especies de levadura incluyen, sin carácter limitativo, los promotores Gal 1 y Gal 10 de *Saccharomyces cerevisiae* y los promotores *Pichia pastoris* AOX1 o AOX2.

[0127] Ejemplos de organismos anfitriones bacteriales son especies bacteriales positivas gram como *Bacillaceae* incluyendo *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus lentus*, *Bacillus brevis*, *Bacillus stearothermophilus*, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus coagulans*, *Bacillus lautus*, *Bacillus megaterium* y *Bacillus thuringiensis*, especies *Streptomyces* como *Streptomyces murinus*, especies bacteriales de ácido láctico incluyendo *Lactococcus* spp. como *Lactococcus lactis*, *Lactobacillus* spp. incluyendo *Lactobacillus reuteri*, *Leuconostoc* spp., *Pediococcus* spp. y *Streptococcus* spp. De manera alternativa, las VARIANTEEs de especies bacteriales de gram negativo que pertenecen a *Enterobacteriaceae* incluyendo *E. coli*, o *Pseudomonadaceae* pueden seleccionarse como el organismo anfitrión. Un organismo anfitrión de levadura adecuado puede seleccionarse de las especies de levaduras relevantes en la biotecnología como, sin carácter limitativo, las especies de levadura como *Pichia* sp., *Hansenula* sp o *Kluyveromyces*, especies *Yarrowinia* o una especie de *Saccharomyces* incluyendo *Saccharomyces cerevisiae* o una especie que pertenece a *Schizosaccharomyce* como por ejemplo, especie *S. Pombe*. Preferiblemente una variedad de especies de levadura metilotrófica *Pichia pastoris* se utiliza como el organismo anfitrión. Preferiblemente el organismo anfitrión es una especie *Hansenula*. Los organismos anfitriones adecuados entre los hongos filamentosos incluyen especies de *Aspergillus*, p.ej. *Aspergillus niger*, *Aspergillus oryzae*, *Aspergillus tubigensis*, *Aspergillus awamori* o *Aspergillus nidulans*. De manera alternativa, variedades de una especie *Fusarium*, p.ej. *Fusarium oxysporum* o de una especie *Rhizomucor* como *Rhizomucor miehei* puede utilizarse como organismo anfitrión. Otras VARIANTEEs adecuadas incluyen especies *Thermomyces* y *Mucor*.

[0128] Las células anfitrionas que comprenden polinucleótidos pueden utilizarse para expresar polipéptidos, como las enzimas modificadas aquí publicadas, fragmentos, homólogos, VARIANTEEs y derivados de los mismos. Las células anfitrionas pueden cultivarse bajo condiciones adecuadas que permiten la expresión de las proteínas. La expresión de los polipéptidos puede ser constitutiva para que se produzcan continuamente, o inducible, requiriendo un estímulo para iniciar la expresión. En el caso de expresión inducible, la producción de proteína puede iniciarse cuando sea requerida, por ejemplo, mediante adición de una sustancia inductora al medio de cultivo, por ejemplo dexametasona o IPTG. Los polipéptidos pueden extraerse de células anfitrionas mediante una variedad de técnicas conocidas en el campo, incluyendo la lisis enzimática, química y/u osmótica y la disrupción física. Los polipéptidos también pueden producirse de manera recombinante en un sistema libre de células *in vitro*, como el sistema de reticulocitos de conejo TnT™ (Promega).

[0129] En un tercer aspecto, la publicación describe un método para modificar una enzima comprendiendo modificar un primera sitio en la parte de enzima de una región definida estructuralmente para que el primer sitio pueda vincularse al segundo sitio. En un modo de realización preferido, el primer sitio está en un lazo o secuencia adyacente a la lámina- β . En un modo de realización preferido, el segundo sitio se sitúa en una lámina- β . En un modo de realización preferido, la enzima modificada es una xilanaso o Clan C.

[0130] En un modo de realización preferido, la publicación describe una xilanaso modificada o un método para modificar una xilanaso (o enzima modificada), de acuerdo con al menos uno de los siguientes: (i) modificar la secuencia N-terminal para que la región N-terminal esté vinculada mediante un puente de disulfuro a una hebra- β adyacente (véase Gruber, *et al.*, 1998 en *T. reesei* XynII los aminoácidos 1-4 y 24-30 respectivamente); (ii) modificar la C-terminal (en *T. reesei* XynII aminoácidos 183-190, véase Gruber, *et al.*, 1998) para que se vincule a una hebra- β adyacente; (iii) modificar una hélice- α de la enzima para que pueda vincularse de manera más firme al cuerpo de la proteína; (iv) modificar al menos un lazo adyacente para que vincule la hebra beta adyacente B6a (en *T. reesei* XynII aminoácidos 91-94, Gruber, *et al.*, 1998) o (v) modificar el residuo equivalente a XynII, según se presenta arriba.

[0031] Según otro modo de realización (por los ejemplos) la mutagénesis puede utilizarse para crear puentes de disulfuro, puentes salinos y mutaciones de punto separadas en diferentes regiones. Por ejemplo, la enzima

puede modificarse para crear al menos un puente de disulfuro, para que al menos un puente de disulfuro pueda: 1) estabilizar la región N-terminal o vincular la hebra beta N-terminal a la lámina beta adyacente (posiciones 2-28, 5-19, 7-16, 10-29 en Xynll, o una posición equivalente, según se publica aquí); 2) estabilizar la región de hélice alfa (posiciones 105-162, 57-153, 110-151, 111-151, en Xynll, o una posición equivalente según se publica aquí); 3) estabilizar la región C-terminal (posiciones 63-188, 61-190, 36-186 or 34-188 en Xynll, o una posición equivalente según se publica aquí); o 4) estabilizar el lazo vinculándolo a la hebra beta como B6b (92-144, 113-143 en Xynll o una posición equivalente a la publicada aquí) y/o 5) estabilizar la lámina beta (posiciones 26-38, 61-149, 63-147, 65-186, 67-184 en Xynll, o una posición equivalente, según se proporciona aquí).

[0132] Los puentes salinos pueden crearse en diferentes sitios de la enzima (p.ej., posiciones 22, 180, 58 or +191D en Xynll, o una posición equivalente, según se proporciona aquí) y mutaciones de punto único pueden introducirse en diferentes sitios de la molécula (p.ej., posiciones 108, 26, 30, 67, 93, 97, 132, 157, 160, 165, 169 or 186 en Xynll, o una posición equivalente, según se proporciona aquí) aumentando así la termoestabilidad y/o termofilia y o alcalofilia la proteína. En cuanto al Y5 mutante, el término-C puede vincularse de manera más firme al cuerpo de la enzima añadiendo como un cambio recombinante un aminoácido (p.ej., ácido aspártico o ácido glutámico) que entonces puede formar un puente salino desde el término-C hasta el cuerpo de la enzima. Si resulta apropiado, un reemplazo de aminoácido adecuado puede llevarse a cabo en el cuerpo de la proteína, para permitir la formación de un puente salino o para estabilizar la enzima en la parte C-terminal mediante la hélice- α o una región cercana a la hélice- α .

[0133] Pueden crearse mutantes adicionales. La estructura de la hebra beta N-terminal A1 o el lazo N-terminal en las enzimas de familia 11 y 12 se describen como la hebra beta, una parte de la lámina beta A anterior a/hasta una estructura de curva beta que lleva hasta la hebra beta B1 o el lazo N-terminal anterior a la primera hebra beta de la lámina beta. (véase, Törrönen et al., Biochemistry 1995, 34, 847-856; Sandgren, et. al., J. Mol. Bio. (2001) 308, 295-310; Gruber, et al., 1998). La hebra beta B1 de la región N-terminal se describe como la parte de hebra beta de la lámina beta B anterior a/hasta una estructura de curva beta que lleva a la hebra beta B2 o el lazo anterior a la primera hebra beta o la lámina beta. La región de hebra beta A1 está vinculada preferiblemente a la hebra beta A2 o a cualquier otra región adyacente (Xynll o un equivalente del mismo). La región de hebra beta B1 está vinculada preferiblemente a la hebra beta B2 o a cualquier otra región adyacente (Xynll o un equivalente del mismo). En Xynll A1 comprende residuos 1-4, A2 residuos 25-30, B1 residuos 6-10 y B2 residuos 13-19.

[0134] La estructura de la hebra beta C-terminal A4 o el lazo C-terminal en las enzimas de familia 11 y 12 es la parte de hebra beta de la lámina beta A entre las hebras beta A3 y A5 o el lazo que sigue a la lámina beta A4 (véase Törrönen et al., Biochemistry 1995, 34, 847-856; Sandgren, et. al., J. Mol. Bio. (2001) 308,295-310; Gruber, et al., 1998). La región de hebra beta A4 está vinculada preferiblemente a la hebra beta A3 o A5, o a cualquier otra región adyacente. En Xynll A4 es residuos 183-190, A3 es residuos 33-39 y A5 es residuos 61-69. El cordón de la familia 11 y 12 se describe como el lazo que conecta las hebras beta B6b y B9. La hebra beta de la familia 11 y 12 B6b se describe como la hebra beta anterior al cordón (Törrönen et al., Biochemistry 1995, 34, 847-856; Sandgren, et. al., J. Mol. Bio. (2001) 308, 295-310; Gruber, et al., 1998). La hebra beta B6b puede estar vinculada al cordón o al lazo entre las hebras beta A6 y B7, o a cualquier otra región adyacente. En Xynll, B6b es residuos 90-94 y B9 es residuos 103-110, el cordón es 95-102, la hebra beta A6 es residuos 148-152, la hebra beta B7 es residuos 134-142 y el lazo entre las hebras beta A6 y B7 es residuos 143-147.

[0135] La hélice de las enzimas de familia 11 y 12 se describe como la región que sigue a la hebra beta A6 y forma una estructura helicoidal paralela a la hebra beta B9 (Törrönen et al., Biochemistry 1995, 34, 847-856; Sandgren, et. al., J. Mol. Bio. (2001) 308, 295-310; Gruber, et al., 1998). La hélice de las enzimas de familia 11 y 12 está vinculada preferiblemente a la hebra beta B9 o a cualquier otra región adyacente. En Xynll la hélice es residuos 153-162, la hebra beta A6 es residuos 148-152 y la hebra beta B9 es residuos 103-110.

EJEMPLOS

EJEMPLO 1.

Plásmidos utilizados para la expresión de xilanas II y patrón de mutagénesis

[0136] El marco abierto de lectura que codifica el producto de gen *Trichoderma reesei* XYNII se amplió mediante la reacción en cadena de la polimerasa (PCR) a partir del banco de datos ADNc *T. reesei*. XYNII ADNc se clonó en pKktac (VTT, Espoo, Finlandia) o de manera alternativa en pALK143 (ROAL, Rajamäki, Finlandia).

EJEMPLO 2.

Mutagénesis dirigida al sitio para la generación de un mutante de xilanas II

[0137] Los vectores de expresión que contienen ADNc-codificando xilanasas II según se describe en el Ejemplo 1 se utilizaron como patrón en la mutagénesis dirigida al sitio de tipo gradual en consecutivas ampliaciones PCR. Los cebadores oligonucleótidos sintéticos que contienen codones alterados para las mutaciones X-Y se utilizaron para la inserción de la alteración deseada en la secuencia de aminoácido primaria de xilanasas II nativa. Mediante este enfoque los residuos de los sitios 92, 93 y 144 de los mutantes de enzima de tipo silvestre se generaron para vincular el lazo N143- S146 de xynII a la hebra-β cercana. De manera adicional, la mutagénesis se llevó a cabo para generar las mutaciones en los sitios 22, 65, 97 y 108 en la secuencia primaria de xilanasas. Las secuencias oligonucleótidas utilizadas en la mutagénesis se muestran en la Figura 3. La PCR se llevó a cabo según se ha descrito en la mutagénesis dirigida al sitio *Quick Change Site-directed mutagenesis* (Stratagene, La Jolla, Ca, EEUU) de acuerdo con los procedimientos estándar de PCR. *PfuTurbo* (Stratagene) se utilizó como ADN polimerasa para amplificar el ADN plásmido. El ADN plásmido de la amplificación PCR de mutagénesis dirigida al sitio se transformó en *E. coli* XL-1 azul y las células de bacteria transformadas se propagaron entonces en LB, con ampicilina 100 ug/ml por selección de ADN plásmido y amplificación del ADN mutado. Los plásmidos se aislaron y secuenciaron para confirmar que contienen las mutaciones deseadas. El ADN plásmido mutado que codifica las VARIANTEes mutantes se sobreexpresa en *E. coli* para examinar la influencia de la mutagénesis en las propiedades enzimáticas de los mutantes Y5 *T. reesei* xilanasas.

EJEMPLO 3.

Producción de los productos de gen XYNII modificados en la variedad *E. coli* y ensayo para la actividad de xilanasas

[0138] Las variedades de *E. coli* sobreexpresando las VARIANTEes mutadas de xilanasas II se cultivaron sobre placas complementadas con 1% xilano de abedul (Sigma, Steinheim, Alemania) acopladas con Rhemazol Brilliant Blue. Rhemazol Brilliant Blue acoplado al xilano se utilizó para detectar la actividad de xilanasas que era fácilmente visualizada por una característica formación de halo debido a la desaparición del color azul alrededor de las colonias de bacteria que expresan la actividad de la xilanasas (Biely *et al.*, 1985).

[0139] Los genes de xilanasas mutadas (véase arriba; Ejemplo 2) se expresaron en *E. coli* a +37°C en frascos de agitación en un medio de cultivo LB. Los cultivos de células expresando las VARIANTEes de enzima se centrifugaron y el sedimento celular se separó del sobrenadante que alberga la enzima que se secretó de las células hacia el medio de cultivo. El ensayo de la actividad de enzima xilanasas se llevó a cabo de acuerdo con métodos estándar. El medio de crecimiento que contiene los mutantes de xilanasas secretados se incubaron durante 10 min en 1% xilano de abedul (Sigma) a 50°C en 50mM de buffer citrato-fosfato (ph 5,0 -t) y 50 mM Tris-HCl en pH 7-9. (Bailey *et al.*, 1992). Si fue necesario, el medio de crecimiento inactivado por calor se utilizó para diluir las muestras. La actividad enzimática de las VARIANTEes mutantes se examinó en comparación con el tipo silvestre y la enzima de mutación Y5 en condiciones variadas (véase, para Bailey *et al.*, 1992).

EJEMPLO 4.

Determinación de la estabilidad dependiente de la temperatura y la actividad dependiente del pH de los mutantes de xilanasas II

Actividad como una función de temperatura;

[0140] La actividad xilanasas de las VARIANTEes mutantes se determinó a temperaturas variadas y se seleccionaron los valores pH (véase las figuras aquí). Los mutantes se incubaron durante 10 min con 1% xilano de abedul (Sigma) en 50mM buffer citrato-fosfato (ph 4,5-7) o 50 mM Tris-HCl en pH 7-9. La cantidad relativa de azúcares reductores liberados se detectó con el ensayo de método DNS según se describe en el ejemplo 3.

Actividad residual

[0141] Las VARIANTEes mutantes se incubaron durante 10 minutos a temperaturas variadas sin sustrato. Después de la inactivación, las muestras se enfriaron con hielo y la actividad residual se determinó mediante un método DNS según se describe en el ejemplo 3.

Actividad dependiente del pH

[0142] La actividad de xilanasas dependiente del pH se determinó detectando la actividad de enzima en un pH variado que oscila entre XX – YY durante 10 min en 1% xilano de abedul a temperaturas seleccionadas (véase las imágenes) en 50mM buffer citrato-fosfato (ph 4,5-7) y 50 mM Tris-HCl con pH 7,5-9. Esto estuvo seguido por un ensayo DNS según se describe en el ejemplo 3.

EJEMPLO 5.

Preparación y muestreo de xilanasas mutantes para propiedades mejoradas

5 **[0143]** Las xilanasas mutantes se prepararon con sustituciones en una o más sustituciones de diferentes regiones de la molécula. Las sustituciones eran mutaciones de punto separado en contacto con otras mutaciones de punto separado o estaban preparadas para actuar sobre un elemento estructural descubierto de manera común tanto en las enzimas de familia 11 como de familia 12. Los ensayos de enzima se llevaron a cabo según se perfila en los ejemplos. Los ejemplos de sustituciones "estructurales" se publican aquí y se muestran en los ejemplos.

10 **[0144]** El puente de disulfuro puede situarse entre los sitios 2 y 28 (T2c, T28C). La figura 4 muestra la importancia de la región N-terminal al sustituir los residuos del wt por una VARIANTEe más termófila. De manera similar, la eliminación del puente de disulfuro nativo (residuos C4 y C32, numeración Cel12A) de T.reesei EGIII afecta ampliamente la estabilidad de la enzima, según se muestra en las figuras proporcionadas y las tablas aquí (véase, especialmente, la Tabla A).

15 **[0145]** La región de la lámina beta común para ambas familias 11 y 12 nombrada hebra beta B6b (según Gruber et al), se muestra importante para la estabilidad, especialmente en condiciones álcali. Este efecto se observa en las sustituciones (al comparar con la VARIANTEe Y5) como estabilidad mejorada en pH 9 vs pH5 para P12, según se muestra en las figuras (véase, por ejemplo, la figura 9, la figura 10 y la figura 11).

20 **[0146]** La importancia de la región está claramente demostrada por un conjunto diferente de mutaciones (aunque en la misma región) afectando a la misma hebra beta. Cuando los sitios 93, 97 y 144 se sustituyen (F93W, N97R, H144K, P9 en el gráfico), puede observarse en la figura 9 un efecto similar en la estabilización de la enzima que aquel cuando se substituyen los sitios 92 y 144 (N92C, H144C= P12 en el gráfico.)

[0147] Un ejemplo de las características mejoradas de sustituciones separadas en los sitios 22 y 180 se muestra abajo. La VARIANTEe que contiene las sustituciones H22K y F180Q (P20 en la figura 14) muestra una estabilidad térmica mejorada sobre Y5 con pH 7,8.

25 **[0148]** Además, la región C-terminal es importante para la estabilidad. En la sustitución S65C, S186C (J21 en el gráfico) la enzima muestra una actividad mejorada con respecto a la temperatura con pH 8.

30 **[0149]** Alguien especializado en la técnica apreciaría fácilmente que la presente publicación está bien adaptada para realizar los objetos y obtener los fines y ventajas mencionados, así como aquellos inherentes a ella. Los complejos moleculares y los métodos, procedimientos, tratamientos, moléculas, compuestos específicos aquí descritos son presentemente representativos de modos de realización preferidos, son ejemplares y no pretenden ser limitativos. Será fácilmente aparente para alguien especializado en la técnica que pueden realizarse sustituciones y modificaciones variadas.

[0150] Todas las patentes y publicaciones mencionadas en la especificación son indicativas de los niveles de aquellos especializados en la técnica.

35 **[0151]** La publicación descrita aquí de manera ilustrativa y adecuada puede practicarse en la ausencia de cualquier elemento o elementos, limitación o limitaciones que no se haya publicado aquí específicamente. Los términos y expresiones que se han empleado se utilizan como términos de descripción y no de limitación y no existe intención de que al usar dichos términos y expresiones se excluyan equivalentes de las características mostradas y descritas o partes de las mismas, sino que se reconoce que son posibles varias modificaciones. Así, debe entenderse que pese a que la presente publicación ha sido publicada específicamente con modos de realización preferidos y características opcionales, la modificación y variación de los conceptos aquí publicados puede llevarse a cabo por aquellos especialistas en la técnica y que dichas modificaciones y variaciones están bajo consideración.

LISTADO DE SECUENCIA.

[0152]

- 45 <110> Genencor International, Inc
 .
 <120> Enzimas modificadas, Métodos para producir enzimas modificadas y sus usos
- 50 <130> SMK/FP6360325
 <140> EP 04821802.8
 <141> 2004-09-10

<150> PCT/US2004/029575
<151> 2004-09-10
5
<150> US 60/503,251
<151> 2003-09-15
10
<160> 51
<170> FastSEQ para Windows Versión 4.0
15
<210> 1
<211> 223
<212> PRT
<213> Trichoderma reesei
<400> 1

ES 2 501 190 T3

Met Val Ser Phe Thr Ser Leu Leu Ala Gly Val Ala Ala Ile Ser Gly
 1 5 10 15
 Val Leu Ala Ala Pro Ala Ala Glu Val Glu Ser Val Ala Val Glu Lys
 20 25 30
 Arg Gln Thr Ile Gln Pro Gly Thr Gly Tyr Asn Asn Gly Tyr Phe Tyr
 35 40 45
 Ser Tyr Trp Asn Asp Gly His Gly Gly Val Thr Tyr Thr Asn Gly Pro
 50 55 60
 Gly Gly Gln Phe Ser Val Asn Trp Ser Asn Ser Gly Asn Phe Val Gly
 65 70 75 80
 Gly Lys Gly Trp Gln Pro Gly Thr Lys Asn Lys Val Ile Asn Phe Ser
 85 90 95
 Gly Ser Tyr Asn Pro Asn Gly Asn Ser Tyr Leu Ser Val Tyr Gly Trp
 100 105 110
 Ser Arg Asn Pro Leu Ile Glu Tyr Tyr Ile Val Glu Asn Phe Gly Thr
 115 120 125
 Tyr Asn Pro Ser Thr Gly Ala Thr Lys Leu Gly Glu Val Thr Ser Asp
 130 135 140
 Gly Ser Val Tyr Asp Ile Tyr Arg Thr Gln Arg Val Asn Gln Pro Ser
 145 150 155 160
 Ile Ile Gly Thr Ala Thr Phe Tyr Gln Tyr Trp Ser Val Arg Arg Asn
 165 170 175
 His Arg Ser Ser Gly Ser Val Asn Thr Ala Asn His Phe Asn Ala Trp
 180 185 190
 Ala Gln Gln Gly Leu Thr Leu Gly Thr Met Asp Tyr Gln Ile Val Ala
 195 200 205
 Val Glu Gly Tyr Phe Ser Ser Gly Ser Ala Ser Ile Thr Val Ser
 210 215 220

<210> 2
 <211> 781
 <212> ADN
 5 <213> Trichoderma reesei

<400> 2

atggtctcct tcacctccct cctcgccggc gtcgccgcca tctcgggctg cttggccgct 60
 cccgccgccc aggtcgaatc cgtggctgtg gagaagcgcc agacgattca gcccgccacg 120
 ggctacaaca acggctactt ctactcgtac tggaacgatg gccacggcgg cgtgacgtac 180
 accaatggtc cccgcccggc gttctccgtc aactggtcca actcgggcaa ctttgtcggc 240
 ggcaagggat ggcagcccgg caccaagaac aagtaagact acctactctt accccctttg 300
 accaacacag cacaacacaa tacaacacat gtgactacca atcatggaat cggatcctaac 360
 agctgtgttt tcaaaaaaaaa gggatcatcaa cttctcgggc agtacaacc ccaacggcaa 420
 cagctacctc tccgtgtacg gctgggtccc caaccctctg atcgagtact acatcgctcga 480
 gaactttggc acctacaacc cgtccacggg cgcaccaag ctgggagagg tcacctccga 540
 cggcagcgtc tacgacatct accgcacgca gcgcgtcaac cagccgtcca tcatcggcac 600
 cgccaccttt taccagtact ggtccgctcc cgcgaaccac cgctcgagcg gctccgtcaa 660
 cacggcgaac cacttcaacg cgtgggctca gcaaggcctg acgctcggga cgatggatta 720
 ccagattgtt gccgtggagg gttactttag ctctggctct gcttccatca ccgtcagcta 780
 a 781

<210> 3
 10 <211> 234
 <212> PRT
 <213> Trichoderma reesei

<400> 3

15

ES 2 501 190 T3

Met Lys Phe Leu Gln Val Leu Pro Ala Leu Ile Pro Ala Ala Leu Ala
 1 5 10 15
 Gln Thr Ser Cys Asp Gln Trp Ala Thr Phe Thr Gly Asn Gly Tyr Thr
 20 25 30
 Val Ser Asn Asn Leu Trp Gly Ala Ser Ala Gly Ser Gly Phe Gly Cys
 35 40 45
 Val Thr Ala Val Ser Leu Ser Gly Gly Ala Ser Trp His Ala Asp Trp
 50 55 60
 Gln Trp Ser Gly Gly Gln Asn Asn Val Lys Ser Tyr Gln Asn Ser Gln
 65 70 75 80
 Ile Ala Ile Pro Gln Lys Arg Thr Val Asn Ser Ile Ser Ser Met Pro
 85 90 95
 Thr Thr Ala Ser Trp Ser Tyr Ser Gly Ser Asn Ile Arg Ala Asn Val
 100 105 110
 Ala Tyr Asp Leu Phe Thr Ala Ala Asn Pro Asn His Val Thr Tyr Ser
 115 120 125
 Gly Asp Tyr Glu Leu Met Ile Trp Leu Gly Lys Tyr Gly Asp Ile Gly
 130 135 140
 Pro Ile Gly Ser Ser Gln Gly Thr Val Asn Val Gly Gly Gln Ser Trp
 145 150 155 160
 Thr Leu Tyr Tyr Gly Tyr Asn Gly Ala Met Gln Val Tyr Ser Phe Val
 165 170 175
 Ala Gln Thr Asn Thr Thr Asn Tyr Ser Gly Asp Val Lys Asn Phe Phe
 180 185 190
 Asn Tyr Leu Arg Asp Asn Lys Gly Tyr Asn Ala Ala Gly Gln Tyr Val
 195 200 205
 Leu Ser Tyr Gln Phe Gly Thr Glu Pro Phe Thr Gly Ser Gly Thr Leu
 210 215 220
 Asn Val Ala Ser Trp Thr Ala Ser Ile Asn
 225 230

<210> 4

<211> 826

<212> ADN

5 <213> Trichoderma reesei

<400> 4

atgaagtcc ttcaagtct cctgccctc ataccggccg cctggccca aaccagctgt 60
 gaccagtggg caaccttcac tggcaacggc tacacagtca gcaacaacct ttggggagca 120
 tcagccggct ctggatttgg ctgcgtgacg gcggtatcgc tcagcggcgg gcctcctgg 180
 cacgcagact ggcagtggtc cggcggccag aacaacgtca agtcgtacca gaactctcag 240
 attgccattc cccagaagag gaccgtcaac agcatcagca gcatgcccac cactgccagc 300
 tggagctaca gcgggagcaa catccgcgct aatgttgcgt atgacttgtt caccgcagcc 360
 aaccggaatc atgtcaagta ctcgggagac tacgaactca tgatctggta agccataaga 420
 agtgaccctc cttgatagtt tcgactaaca acatgtcttg aggcttgga aatacggcga 480
 tattgggccc attgggtcct cacagggaac agtcaacgtc ggtggccaga gctggacgct 540
 ctactatggc tacaacggag ccatgcaagt ctattccttt gtggcccaga ccaacactac 600
 caactacagc ggagatgtca agaacttctt caattatctc cgagacaata aaggatacaa 660
 cgctgcaggc caatatgttc ttagtaagtc accctcactg tgactgggct gagtttggttg 720
 caacgtttgc taacaaaacc ttcgtatagg ctaccaattt ggtaccgagc ccttcacggg 780
 cagtggaact ctgaacgtcg catcctggac cgcactatc aactaa 826

10 <210> 5

<211> 222

<212> PRT

<213> Trichoderma reesei

<400> 5

ES 2 501 190 T3

```

Met Val Ser Phe Thr Ser Leu Leu Ala Ala Ser Pro Pro Ser Arg Ala
 1          5          10          15
Ser Cys Arg Pro Ala Ala Glu Val Glu Ser Val Ala Val Glu Lys Arg
      20          25          30
Gln Thr Ile Gln Pro Gly Thr Gly Tyr Asn Asn Gly Tyr Phe Tyr Ser
      35          40          45
Tyr Trp Asn Asp Gly His Gly Gly Val Thr Tyr Thr Asn Gly Pro Gly
      50          55          60
Gly Gln Phe Ser Val Asn Trp Ser Asn Ser Gly Asn Phe Val Gly Gly
      65          70          75          80
Lys Gly Trp Gln Pro Gly Thr Lys Asn Lys Val Ile Asn Phe Ser Gly
      85          90          95
Ser Tyr Asn Pro Asn Gly Asn Ser Tyr Leu Ser Val Tyr Gly Trp Ser
      100          105          110
Arg Asn Pro Leu Ile Glu Tyr Tyr Ile Val Glu Asn Phe Gly Thr Tyr
      115          120          125
Asn Pro Ser Thr Gly Ala Thr Lys Leu Gly Glu Val Thr Ser Asp Gly
      130          135          140
Ser Val Tyr Asp Ile Tyr Arg Thr Gln Arg Val Asn Gln Pro Ser Ile
      145          150          155          160
Ile Gly Thr Ala Thr Phe Tyr Gln Tyr Trp Ser Val Arg Arg Asn His
      165          170          175
Arg Ser Ser Gly Ser Val Asn Thr Ala Asn His Phe Asn Ala Trp Ala
      180          185          190
Gln Gln Gly Leu Thr Leu Gly Thr Met Asp Tyr Gln Ile Val Ala Val
      195          200          205
Glu Gly Tyr Phe Ser Ser Gly Ser Ala Ser Ile Thr Val Ser
      210          215          220

```

<210> 6

<211> 227

<212> PRT

5 <213> Humicola insolens

<400> 6

```

Met Val Ser Leu Lys Ser Val Leu Ala Ala Ala Thr Ala Val Ser Ser
 1          5          10          15
Ala Ile Ala Ala Pro Phe Asp Phe Val Pro Arg Asp Asn Ser Thr Ala
      20          25          30
Leu Gln Ala Arg Gln Val Thr Pro Asn Ala Glu Gly Trp His Asn Gly
      35          40          45

```

ES 2 501 190 T3

Tyr Phe Tyr Ser Trp Trp Ser Asp Gly Gly Gly Gln Val Gln Tyr Thr
 50 55 60
 Asn Leu Glu Gly Ser Arg Tyr Gln Val Arg Trp Arg Asn Thr Gly Asn
 65 70 75 80
 Phe Val Gly Gly Lys Gly Trp Asn Pro Gly Thr Gly Arg Thr Ile Asn
 85 90 95
 Tyr Gly Gly Tyr Phe Asn Pro Gln Gly Asn Gly Tyr Leu Ala Val Tyr
 100 105 110
 Gly Trp Thr Arg Asn Pro Leu Val Glu Tyr Tyr Val Ile Glu Ser Tyr
 115 120 125
 Gly Thr Tyr Asn Pro Gly Ser Gln Ala Gln Tyr Lys Gly Thr Phe Tyr
 130 135 140
 Thr Asp Gly Asp Gln Tyr Asp Ile Phe Val Ser Thr Arg Tyr Asn Gln
 145 150 155 160
 Pro Ser Ile Asp Gly Thr Arg Thr Phe Gln Tyr Trp Ser Ile Arg
 165 170 175
 Lys Asn Lys Arg Val Gly Gly Ser Val Asn Met Gln Asn His Phe Asn
 180 185 190
 Ala Trp Gln Gln His Gly Met Pro Leu Gly Gln His Tyr Tyr Gln Val
 195 200 205
 Val Ala Thr Glu Gly Tyr Gln Ser Ser Gly Glu Ser Asp Ile Tyr Val
 210 215 220
 Gln Thr His
 225

<210> 7
 <211> 210
 5 <212> PRT
 <213> Bacillus stearothermophilus

<400> 7

Met Lys Leu Lys Lys Lys Met Leu Thr Leu Leu Leu Thr Ala Ser Met
 1 5 10 15
 Ser Phe Gly Leu Phe Gly Ala Thr Ser Ser Ala Ala Thr Asp Tyr Trp
 20 25 30
 Gln Tyr Trp Thr Asp Gly Gly Gly Met Val Asn Ala Val Asn Gly Pro
 35 40 45
 Gly Gly Asn Tyr Ser Val Thr Trp Gln Asn Thr Gly Asn Phe Val Val
 50 55 60
 Gly Lys Gly Trp Thr Val Gly Ser Pro Asn Arg Val Ile Asn Tyr Asn
 65 70 75 80
 Ala Gly Ile Trp Glu Pro Ser Gly Asn Gly Tyr Leu Thr Leu Tyr Gly
 85 90 95
 Trp Thr Arg Asn Ala Leu Ile Glu Tyr Tyr Val Val Asp Ser Trp Gly
 100 105 110
 Thr Tyr Arg Pro Thr Gly Asn Tyr Lys Gly Thr Val Asn Ser Asp Gly
 115 120 125
 Gly Thr Tyr Asp Ile Tyr Thr Thr Met Arg Tyr Asn Ala Pro Ser Ile
 130 135 140
 Asp Gly Thr Gln Thr Phe Gln Gln Phe Trp Ser Val Arg Gln Ser Lys
 145 150 155 160
 Arg Pro Thr Gly Ser Asn Val Ser Ile Thr Phe Ser Asn His Val Asn
 165 170 175
 Ala Trp Arg Ser Lys Gly Met Asn Leu Gly Ser Ser Trp Ala Tyr Gln
 180 185 190
 Val Leu Ala Thr Glu Gly Tyr Gln Ser Ser Gly Arg Ser Asn Val Thr
 195 200 205
 Val Trp
 210

ES 2 501 190 T3

<210> 8

<211> 229

<212> PRT

<213> *Trichoderma reesei*

5

<400> 8

```

Met Val Ala Phe Ser Ser Leu Ile Cys Ala Leu Thr Ser Ile Ala Ser
 1          5          10          15
Thr Leu Ala Met Pro Thr Gly Leu Glu Pro Glu Ser Ser Val Asn Val
 20          25          30
Thr Glu Arg Gly Met Tyr Asp Phe Val Leu Gly Ala His Asn Asp His
 35          40          45
Arg Arg Arg Ala Ser Ile Asn Tyr Asp Gln Asn Tyr Gln Thr Gly Gly
 50          55          60
Gln Val Ser Tyr Ser Pro Ser Asn Thr Gly Phe Ser Val Asn Trp Asn
 65          70          75          80
Thr Gln Asp Asp Phe Val Val Gly Val Gly Trp Thr Thr Gly Ser Ser
 85          90          95
Ala Pro Ile Asn Phe Gly Gly Ser Phe Ser Val Asn Ser Gly Thr Gly
 100         105         110
Leu Leu Ser Val Tyr Gly Trp Ser Thr Asn Pro Leu Val Glu Tyr Tyr
 115         120         125
Ile Met Glu Asp Asn His Asn Tyr Pro Ala Gln Gly Thr Val Lys Gly
 130         135         140
Thr Val Thr Ser Asp Gly Ala Thr Tyr Thr Ile Trp Glu Asn Thr Arg
 145         150         155         160
Val Asn Glu Pro Ser Ile Gln Gly Thr Ala Thr Phe Asn Gln Tyr Ile
 165         170         175
Ser Val Arg Asn Ser Pro Arg Thr Ser Gly Thr Val Thr Val Gln Asn
 180         185         190
His Phe Asn Ala Trp Ala Ser Leu Gly Leu His Leu Gly Gln Met Asn
 195         200         205
Tyr Gln Val Val Ala Val Glu Gly Trp Gly Gly Ser Gly Ser Ala Ser
 210         215         220
Gln Ser Val Ser Asn
 225
    
```

<210> 9

<211> 211

<212> PRT

<213> *Aspergillus awamori*

10

<400> 9

ES 2 501 190 T3

```

Met Lys Val Thr Ala Ala Phe Ala Gly Leu Leu Val Thr Ala Phe Ala
 1          5          10          15
Ala Pro Val Pro Glu Pro Val Leu Val Ser Arg Ser Ala Gly Ile Asn
 20          25          30
Tyr Val Gln Asn Tyr Asn Gly Asn Leu Gly Asp Phe Thr Tyr Asp Glu
 35          40          45
Ser Ala Gly Thr Phe Ser Met Tyr Trp Glu Asp Gly Val Ser Ser Asp
 50          55          60
Phe Val Val Gly Leu Gly Trp Thr Thr Gly Ser Ser Asn Ala Ile Thr
 65          70          75
Tyr Ser Ala Glu Tyr Ser Ala Ser Gly Ser Ser Ser Tyr Leu Ala Val
 85          90          95
Tyr Gly Trp Val Asn Tyr Pro Gln Ala Glu Tyr Tyr Ile Val Glu Asp
100          105          110
Tyr Gly Asp Tyr Asn Pro Cys Ser Ser Ala Thr Ser Leu Gly Thr Val
115          120          125
Tyr Ser Asp Gly Ser Thr Tyr Gln Val Cys Thr Asp Thr Arg Thr Asn
130          135          140
Glu Pro Ser Ile Thr Gly Thr Ser Thr Phe Thr Gln Tyr Phe Ser Val
145          150          155          160
Arg Glu Ser Thr Arg Thr Ser Gly Thr Val Thr Val Ala Asn His Phe
165          170          175
Asn Phe Trp Ala Gln His Gly Phe Gly Asn Ser Asp Phe Asn Tyr Gln
180          185          190
Val Met Ala Val Glu Ala Trp Ser Gly Ala Gly Ser Ala Ser Val Thr
195          200          205
Ile Ser Ser
210

```

- <210> 10
- <211> 330
- 5 <212> PRT
- <213> Bacillus stearothermophilus
- <400> 10

ES 2 501 190 T3

Met Cys Ser Ser Ile Pro Ser Leu Arg Glu Val Phe Ala Asn Asp Phe
 1 5 10 15
 Arg Ile Gly Ala Ala Val Asn Pro Val Thr Leu Glu Ala Gln Gln Ser
 20 25 30
 Leu Leu Ile Arg His Val Asn Ser Leu Thr Ala Glu Asn His Met Lys
 35 40 45
 Phe Glu His Leu Gln Pro Glu Glu Gly Arg Phe Thr Phe Asp Ile Ala
 50 55 60
 Ile Lys Ser Ser Thr Ser Pro Phe Ser Ser His Gly Val Arg Gly His
 65 70 75 80
 Thr Leu Val Trp His Asn Gln Thr Pro Ser Trp Val Phe Gln Asp Ser
 85 90 95
 Gln Gly His Phe Val Gly Arg Asp Val Leu Leu Glu Arg Met Lys Ser
 100 105 110
 His Ile Ser Thr Val Val Gln Arg Tyr Lys Gly Lys Val Tyr Cys Trp
 115 120 125
 Asp Val Ile Asn Glu Ala Val Ala Asp Glu Gly Ser Glu Trp Leu Arg
 130 135 140
 Ser Ser Thr Trp Arg Gln Ile Ile Gly Asp Asp Phe Ile Gln Gln Ala
 145 150 155 160
 Phe Leu Tyr Ala His Glu Ala Asp Pro Glu Ala Leu Leu Phe Tyr Asn
 165 170 175
 Asp Tyr Asn Glu Cys Phe Pro Glu Lys Arg Glu Lys Ile Tyr Thr Leu
 180 185 190
 Val Lys Ser Leu Arg Asp Lys Gly Ile Pro Ile His Gly Ile Gly Met
 195 200 205
 Gln Ala His Trp Ser Leu Asn Arg Pro Thr Leu Asp Glu Ile Arg Ala
 210 215 220
 Ala Ile Glu Arg Tyr Ala Ser Leu Gly Val Ile Leu His Ile Thr Glu
 225 230 235 240
 Leu Asp Ile Ser Met Phe Glu Phe Asp Asp His Arg Lys Asp Leu Ala
 245 250 255
 Ala Pro Thr Asn Glu Met Val Glu Arg Gln Ala Glu Arg Tyr Glu Gln
 260 265 270
 Ile Phe Ser Leu Phe Lys Glu Tyr Arg Asp Val Ile Gln Asn Val Thr
 275 280 285
 Phe Trp Gly Ile Ala Asp Asp His Thr Trp Leu Asp His Phe Pro Val
 290 295 300
 Gln Gly Arg Lys Asn Trp Pro Leu Leu Phe Asp Glu Gln His Asn Pro
 305 310 315 320
 Lys Pro Ala Phe Trp Arg Val Val Asn Ile
 325 330

<210> 11

<211> 190

<212> PRT

5 <213> Trichoderma reesei

<400> 11

ES 2 501 190 T3

Gln Thr Ile Gln Pro Gly Thr Gly Tyr Asn Asn Gly Tyr Phe Tyr Ser
 1 5 10 15
 Tyr Trp Asn Asp Gly His Gly Gly Val Thr Tyr Thr Asn Gly Pro Gly
 20 25 30
 Gly Gln Phe Ser Val Asn Trp Ser Gly Asn Phe Val Gly Gly
 35 40 45
 Lys Gly Trp Gln Pro Gly Thr Lys Asn Lys Val Ile Asn Phe Ser Gly
 50 55 60
 Ser Tyr Asn Pro Asn Gly Asn Ser Tyr Leu Ser Val Tyr Gly Trp Ser
 65 70 75 80
 Arg Asn Pro Leu Ile Glu Tyr Tyr Ile Val Glu Asn Phe Gly Thr Tyr
 85 90 95
 Asn Pro Ser Thr Gly Ala Thr Lys Leu Gly Glu Val Thr Ser Asp Gly
 100 105 110
 Ser Val Tyr Asp Ile Tyr Arg Thr Gln Arg Val Asn Gln Pro Ser Ile
 115 120 125
 Ile Gly Thr Ala Thr Phe Tyr Gln Tyr Trp Ser Val Arg Arg Asn His
 130 135 140
 Arg Ser Ser Gly Ser Val Asn Thr Ala Asn His Phe Asn Ala Trp Ala
 145 150 155 160
 Gln Gln Gly Leu Thr Leu Gly Thr Met Asp Tyr Gln Ile Val Ala Val
 165 170 175
 Glu Gly Tyr Phe Ser Ser Gly Ser Ala Ser Ile Thr Val Ser
 180 185 190

<210> 12
 <211> 237
 <212> PRT

5 <213> Aspergillus awamori

<400> 12

Met Lys Ala Phe His Leu Leu Ala Ala Leu Ser Gly Ala Ala Val Ala
 1 5 10 15
 Gln Gln Ala Gln Leu Cys Asp Gln Tyr Ala Thr Tyr Thr Gly Gly Val
 20 25 30
 Tyr Thr Ile Asn Asn Asn Leu Trp Gly Lys Asp Ala Gly Ser Gly Ser
 35 40 45
 Gln Cys Thr Thr Val Asn Ser Ala Ser Ser Ala Gly Thr Ser Trp Ser
 50 55 60
 Thr Lys Trp Asn Trp Ser Gly Gly Glu Asn Ser Val Lys Ser Tyr Ala
 65 70 75 80
 Asn Ser Gly Leu Ser Phe Asn Lys Lys Leu Val Ser Gln Ile Ser His
 85 90 95
 Ile Pro Thr Ala Ala Arg Trp Ser Tyr Asp Asn Thr Cys Ile Arg Arg
 100 105 110
 Gly Arg Ala Tyr Asp Leu Phe Thr Ala Ala Asp Ile Asn His Val Thr
 115 120 125
 Trp Ser Gly Asp Tyr Glu Leu Met Ile Trp Leu Ala Arg Tyr Gly Gly
 130 135 140
 Val Gln Pro Leu Gly Ser Gln Ile Ala Thr Ala Thr Val Glu Gly Gln
 145 150 155 160
 Thr Trp Glu Leu Trp Tyr Gly Val Asn Gly Ala Gln Lys Thr Tyr Ser
 165 170 175
 Phe Val Ala Ala Asn Pro Ile Thr Ser Phe Gln Gly Asp Ile Asn Asp
 180 185 190
 Phe Phe Lys Tyr Leu Thr Gln Asn His Gly Phe Pro Ala Ser Ser Gln
 195 200 205
 Tyr Leu Ile Thr Leu Gln Phe Gly Thr Glu Pro Phe Thr Gly Gly Pro
 210 215 220
 Ala Thr Leu Asn Val Ala Asp Trp Ser Ala Ser Val Gln
 225 230 235

10 <210> 13
 <211> 233

ES 2 501 190 T3

<212> PRT

<213> *Trichoderma viride*

<400> 13

```

Met Lys Phe Leu Gln Ile Ala Pro Thr Leu Leu Pro Val Ala Leu Ala
 1      5      10      15
Gln Ser Ser Cys Ser Gln Tyr Ala Thr Phe Ser Gly Gly Asn Tyr Ala
 20      25      30
Leu Ser Asn Asn Leu Trp Gly Gln Ser Ala Gly Ser Gly Ser Gly Cys
 35      40      45
Ile Thr Asp Val Ser Leu Gly Gly Ser Ala Val Trp Ser Thr Thr Trp
 50      55      60
Asp Trp Ser Gly Gly Gln Ser Asn Val Lys Gly Tyr Pro Asn Ile Ala
 65      70      75      80
Leu Asn Ile Pro Asn Lys Arg Leu Val Ser Ser Ile Ser Ser Met Pro
 85      90      95
Thr Thr Ala Gln Trp Ser Tyr Ser Gly Ser Ser Ile Arg Ala Asp Val
 100     105     110
Ala Tyr Asp Leu Phe Thr Ala Ser Asn Pro Asn His Val Thr Tyr Ser
 115     120     125
Gly Asp Tyr Glu Leu Met Ile Trp Leu Gly Lys Tyr Gly Asp Ile Gln
 130     135     140
Pro Ile Gly Ser Ser Gln Gly Thr Val Asn Val Gly Gly Thr Ser Trp
 145     150     155     160
Asn Leu Trp Tyr Gly Pro Asn Gly Ser Met Gln Val Tyr Ser Phe Val
 165     170     175
Ala Pro Gly Asn Leu Thr Asn Trp Ser Gly Asp Val Lys Asn Phe Tyr
 180     185     190
Thr Tyr Leu Gln Asn Asn Lys Gly Tyr Pro Ala Ser Ser Gln Tyr Val
 195     200     205
Leu Ser Tyr Gln Phe Gly Thr Glu Ala Phe Thr Gly Ser Gly Thr Leu
 210     215     220
Asn Asn Thr Trp Thr Ala Ser Ile Asn
 225     230

```

5

<210> 14

<211> 234

<212> PRT

<213> *Hypocrea koningii*

10

<400> 14

ES 2 501 190 T3

Met	Lys	Leu	Ile	His	Val	Leu	Pro	Ala	Leu	Ile	Pro	Ala	Ala	Leu	Ala
1				5					10					15	
Gln	Thr	Ser	Cys	Asp	Gln	Tyr	Ala	Val	Phe	Thr	Gly	Ser	Asp	Tyr	Thr
			20					25					30		
Val	Ser	Asn	Asn	Leu	Trp	Gly	Gln	Ser	Ala	Gly	Ser	Gly	Phe	Gly	Cys
		35					40					45			
Val	Thr	Ala	Glu	Ser	Leu	Ser	Gly	Ser	Ala	Ser	Trp	His	Ala	Asp	Trp
	50					55					60				
Gln	Trp	Ser	Gly	Gly	Gln	Asn	Asn	Val	Lys	Ser	Tyr	Gln	Asn	Ser	Gln
65					70					75					80
Ile	Pro	Ile	Pro	Gln	Lys	Arg	Thr	Val	Asn	Ser	Ile	Ser	Ser	Met	Pro
				85					90					95	
Thr	Thr	Ala	Ser	Trp	Ser	Tyr	Thr	Gly	Ser	Asp	Ile	Arg	Ala	Asn	Val
				100				105						110	
Ala	Tyr	Asp	Leu	Phe	Thr	Ala	Ala	Asn	Pro	Asn	His	Val	Thr	Tyr	Ser
		115						120				125			
Gly	Asp	Tyr	Glu	Leu	Met	Ile	Trp	Leu	Gly	Arg	Tyr	Gly	Asp	Ile	Gly
	130					135					140				
Pro	Ile	Gly	Ser	Ser	Gln	Gly	Thr	Val	Asn	Val	Gly	Gly	Gln	Ser	Trp
145					150					155					160
Thr	Leu	Tyr	Tyr	Gly	Tyr	Asn	Gly	Ala	Met	Gln	Val	Tyr	Ser	Phe	Val
				165					170					175	
Ala	Gln	Thr	Asn	Thr	Thr	Ser	Tyr	Ser	Gly	Asp	Val	Lys	Asn	Phe	Phe
			180					185					190		
Asn	Tyr	Leu	Arg	Asp	Asn	Lys	Gly	Tyr	Asn	Ala	Ala	Gly	Gln	Tyr	Val
		195					200					205			
Leu	Ser	Tyr	Gln	Phe	Gly	Thr	Glu	Pro	Phe	Thr	Gly	Ser	Gly	Thr	Leu
	210					215					220				
Asn	Val	Ala	Ser	Trp	Thr	Ala	Ser	Ile	Asn						
225					230										

- <210> 15
- <211> 234
- 5 <212> PRT
- <213> Hypocrea schweinitzii
- <400> 15

ES 2 501 190 T3

Met Lys Phe Leu Gln Val Leu Pro Ala Ile Leu Pro Ala Ala Leu Ala
 1 5 10 15
 Gln Thr Ser Cys Asp Gln Tyr Ala Thr Phe Ser Gly Asn Gly Tyr Ile
 20 25 30
 Val Ser Asn Asn Leu Trp Gly Ala Ser Ala Gly Ser Gly Phe Gly Cys
 35 40 45
 Val Thr Ser Val Ser Leu Asn Gly Ala Ala Ser Trp His Ala Asp Trp
 50 55 60
 Gln Trp Ser Gly Gly Gln Asn Asn Val Lys Ser Tyr Gln Asn Val Gln
 65 70 75 80
 Ile Asn Ile Pro Gln Lys Arg Thr Val Asn Ser Ile Gly Ser Met Pro
 85 90 95
 Thr Thr Ala Ser Trp Ser Tyr Ser Gly Ser Asp Ile Arg Ala Asn Val
 100 105 110
 Ala Tyr Asp Leu Phe Thr Ala Ala Asn Pro Asn His Val Thr Tyr Ser
 115 120 125
 Gly Asp Tyr Glu Leu Met Ile Trp Leu Gly Lys Tyr Gly Asp Ile Gly
 130 135 140
 Pro Ile Gly Ser Ser Gln Gly Thr Val Asn Val Gly Gly Gln Thr Trp
 145 150 155 160
 Thr Leu Tyr Tyr Gly Tyr Asn Gly Ala Met Gln Val Tyr Ser Phe Val
 165 170 175
 Ala Gln Ser Asn Thr Thr Ser Tyr Ser Gly Asp Val Lys Asn Phe Phe
 180 185 190
 Asn Tyr Leu Arg Asp Asn Lys Gly Tyr Asn Ala Gly Gly Gln Tyr Val
 195 200 205
 Leu Ser Tyr Gln Phe Gly Thr Glu Pro Phe Thr Gly Ser Gly Thr Leu
 210 215 220
 Asn Val Ala Ser Trp Thr Ala Ser Ile Asn
 225 230

<210> 16

<211> 237

<212> PRT

5 <213> Stachybotrys echinata

<400> 16

ES 2 501 190 T3

Met Lys Val Ala Ala Leu Leu Val Ala Leu Ser Pro Leu Ala Phe Ala
 1 5 10 15
 Gln Ser Leu Cys Asp Gln Tyr Ser Tyr Tyr Ser Ser Asn Gly Tyr Glu
 20 25 30
 Phe Asn Asn Asn Met Trp Gly Arg Asn Ser Gly Gln Gly Asn Gln Cys
 35 40 45
 Thr Tyr Val Asp Tyr Ser Ser Pro Asn Gly Val Gly Trp Arg Val Asn
 50 55 60
 Trp Asn Trp Ser Gly Gly Asp Asn Asn Val Lys Ser Tyr Pro Tyr Ser
 65 70 75 80
 Gly Arg Gln Leu Pro Thr Lys Arg Ile Val Ser Trp Ile Gly Ser Leu
 85 90 95
 Pro Thr Thr Val Ser Trp Asn Tyr Gln Gly Asn Asn Leu Arg Ala Asn
 100 105 110
 Val Ala Tyr Asp Leu Phe Thr Ala Ala Asn Pro Asn His Pro Asn Ser
 115 120 125
 Ser Gly Asp Tyr Glu Leu Met Ile Trp Leu Gly Arg Leu Gly Asn Val
 130 135 140
 Tyr Pro Ile Gly Asn Gln Val Ala Thr Val Asn Ile Ala Gly Gln Gln
 145 150 155 160
 Trp Asn Leu Tyr Tyr Gly Tyr Asn Gly Ala Met Gln Val Tyr Ser Phe
 165 170 175
 Val Ser Pro Asn Gln Leu Asn Tyr Phe Ser Gly Asn Val Lys Asp Phe
 180 185 190
 Phe Thr Tyr Leu Gln Tyr Asn Arg Ala Tyr Pro Ala Asp Ser Gln Tyr
 195 200 205
 Leu Ile Thr Tyr Gln Phe Gly Thr Glu Pro Phe Thr Gly Gln Asn Ala
 210 215 220
 Val Phe Thr Val Ser Asn Trp Ser Ala Gln Gln Asn Asn
 225 230 235

<210> 17

<211> 238

<212> PRT

5 <213> Fusarium equiseti

<400> 17

ES 2 501 190 T3

Met Lys Ser Thr Leu Leu Leu Ala Gly Ala Phe Ala Pro Leu Ala Phe
 1 5 10 15
 Ala Lys Asp Leu Cys Glu Gln Tyr Gly Tyr Leu Ser Ser Asp Gly Tyr
 20 25 30
 Ser Leu Asn Asn Val Trp Gly Lys Asp Ser Gly Thr Gly Asp Gln
 35 40 45
 Cys Thr His Val Asn Trp Asn Asn Ala Asn Gly Ala Gly Trp Asp Val
 50 55 60
 Glu Trp Asn Trp Ser Gly Gly Lys Asp Asn Val Lys Ser Tyr Pro Asn
 65 70 75 80
 Ser Ala Leu Leu Ile Gly Glu Asp Lys Lys Thr Ile Ser Ser Ile Thr
 85 90 95
 Asn Met Gln Ser Thr Ala Glu Trp Lys Tyr Ser Gly Asp Asn Leu Arg
 100 105 110
 Ala Asp Val Ala Tyr Asp Leu Phe Thr Ala Ala Asp Pro Asn His Glu
 115 120 125
 Thr Ser Ser Gly Glu Tyr Glu Leu Met Val Trp Leu Ala Arg Ile Gly
 130 135 140
 Gly Val Gln Pro Ile Gly Ser Leu Gln Thr Ser Val Thr Ile Glu Gly
 145 150 155 160
 His Thr Trp Glu Leu Trp Val Gly Met Asn Gly Ser Met Lys Val Phe
 165 170 175
 Ser Phe Val Ala Pro Thr Pro Val Asn Asn Phe Asn Ala Asp Ile Lys
 180 185 190
 Gln Phe Trp Asp Tyr Leu Thr Lys Ser Gln Asn Phe Pro Ala Asp Asn
 195 200 205
 Gln Tyr Leu Leu Thr Phe Gln Phe Gly Thr Glu Pro Phe Thr Gly Asp
 210 215 220
 Asn Ala Lys Phe Thr Val Thr Asn Phe Asn Ala His Leu Lys
 225 230 235

<210> 18

<211> 237

5 <212> PRT

<213> Bionectria ochroleuca

<400> 18

ES 2 501 190 T3

Met Lys Thr Gly Ile Ala Tyr Leu Ala Ala Val Leu Pro Leu Ala Met
 1 5 10 15
 Ala Glu Ser Leu Cys Asp Gln Tyr Ala Tyr Leu Ser Arg Asp Gly Tyr
 20 25 30
 Asn Phe Asn Asn Asn Glu Trp Gly Ala Ala Thr Gly Thr Gly Asp Gln
 35 40 45
 Cys Thr Tyr Val Asp Ser Thr Ser Ser Gly Gly Val Ser Trp His Ser
 50 55 60
 Asp Trp Thr Asn Ser Gly Ser Glu Ser Glu Ile Lys Ser Tyr Pro Tyr
 65 70 75 80
 Ser Gly Leu Asp Leu Pro Glu Lys Lys Ile Val Thr Ser Ile Gly Ser
 85 90 95
 Ile Ser Thr Gly Ala Glu Trp Ser Tyr Ser Gly Ser Asn Ile Arg Ala
 100 105 110
 Asp Val Ala Tyr Asp Ile Phe Thr Ala Ala Asp Pro Asn His Ala Thr
 115 120 125
 Ser Ser Gly Asp Tyr Glu Val Met Ile Trp Leu Ala Asn Leu Gly Gly
 130 135 140
 Leu Thr Pro Ile Gly Ser Pro Ile Gly Thr Val Lys Ala Ala Gly Arg
 145 150 155 160
 Asp Trp Glu Leu Trp Asp Gly Tyr Asn Gly Ala Met Arg Val Tyr Ser
 165 170 175
 Phe Val Ala Pro Ser Gln Leu Asn Ser Phe Asp Gly Glu Ile Met Asp
 180 185 190
 Phe Phe Tyr Val Val Lys Asp Met Arg Gly Phe Pro Ala Asp Ser Gln
 195 200 205
 His Leu Leu Thr Val Gln Phe Gly Thr Glu Pro Ile Ser Gly Ser Gly
 210 215 220
 Ala Lys Phe Ser Val Ser His Trp Ser Ala Lys Leu Gly
 225 230 235

<210> 19

<211> 236

<212> PRT

5 <213> Bionectria ochroleuca

<400> 19

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

ES 2 501 190 T3

```

                85                90                95
Asn Gly Gly Ala Asp Trp Asp Tyr Ser Gly Ser Asn Ile Arg Ala Asn
      100                105                110
Val Ala Tyr Asp Ile Phe Thr Ser Ala Asp Pro Asn His Val Thr Ser
      115                120                125
Ser Gly Asp Tyr Glu Leu Met Ile Trp Leu Gly Lys Leu Gly Asp Ile
      130                135                140
Tyr Pro Ile Gly Asn Ser Ile Gly Arg Val Lys Ala Ala Asn Arg Glu
      145                150                155
Trp Asp Leu His Val Gly Tyr Asn Gly Ala Met Lys Val Phe Ser Phe
      165                170                175
Val Ala Pro Ser Pro Val Thr Arg Phe Asp Gly Asn Ile Met Asp Phe
      180                185                190
Phe Tyr Val Met Arg Asp Met Gln Gly Tyr Pro Met Asp Lys Gln Tyr
      195                200                205
Leu Leu Thr Leu Gln Phe Gly Thr Glu Pro Phe Thr Gly Ser Asn Ala
      210                215                220
Lys Phe Ser Cys Trp Tyr Phe Gly Ala Lys Ile Lys
      225                230                235

```

<210> 20

<211> 240

<212> PRT

5 <213> Bionectria ochroleuca

<400> 20

```

Met Lys Ala Asn Ile Val Ile Leu Ser Leu Phe Ala Pro Leu Ala Ala
 1      5      10      15
Val Ala Gln Thr Leu Cys Gly Gln Tyr Ser Ser Asn Thr Gln Gly Gly
 20      25      30
Tyr Ile Phe Asn Asn Asn Met Trp Gly Met Gly Ser Gly Ser Gly Ser
 35      40      45
Gln Cys Thr Tyr Val Asp Lys Val Trp Ala Glu Gly Val Ala Trp His
 50      55      60
Thr Asp Trp Ser Trp Ser Gly Gly Asp Asn Asn Val Lys Ser Tyr Pro
 65      70      75      80
Tyr Ser Gly Arg Glu Leu Gly Thr Lys Arg Ile Val Ser Ser Ile Lys
 85      90      95
Ser Ile Ser Ser Gly Ala Asp Trp Asp Tyr Thr Gly Ser Asn Leu Arg
 100     105     110
Ala Asn Ala Ala Tyr Asp Ile Phe Thr Ser Ala Asn Pro Asn His Ala
 115     120     125
Thr Ser Ser Gly Asp Tyr Glu Val Met Ile Trp Leu Gly Arg Tyr Gly
 130     135     140
Gly Val Tyr Pro Ile Gly Asn Ser Ile Gly Thr Val Arg Ala Ala Gly
 145     150     155     160
Arg Asp Trp Ala Leu His Ile Gly Tyr Asn Gly Ala Met Lys Val Phe
 165     170     175
Ser Phe Val Ala Ala Asn Pro Val Thr Arg Phe Asp Gly Glu Ile Met
 180     185     190
Asp Phe Phe Tyr Leu Leu Arg Asp Met Gln Gly Tyr Pro Met Thr Ser
 195     200     205
Gln Tyr Leu Leu Thr Leu Gln Phe Gly Thr Glu Pro Phe Thr Gly Ser
 210     215     220
Gly Ala Lys Phe Asn Cys Trp Tyr Phe Gly Ala Thr Leu Ser Tyr Trp
 225     230     235     240

```

10 <210> 21

ES 2 501 190 T3

<211> 254
 <212> PRT
 <213> Humicola grisea

5 <400> 21

Met	Leu	Lys	Ser	Ala	Leu	Leu	Leu	Gly	Ala	Ala	Ala	Val	Ser	Val	Gln
1				5					10					15	
Ser	Ala	Ser	Ile	Pro	Thr	Ile	Pro	Ala	Asn	Leu	Glu	Pro	Arg	Gln	Ile
			20					25					30		
Arg	Ser	Leu	Cys	Glu	Leu	Tyr	Gly	Tyr	Trp	Ser	Gly	Asn	Gly	Tyr	Glu
		35					40					45			
Leu	Leu	Asn	Asn	Leu	Trp	Gly	Lys	Asp	Thr	Ala	Thr	Ser	Gly	Trp	Gln
	50					55					60				
Cys	Thr	Tyr	Leu	Asp	Gly	Thr	Asn	Asn	Gly	Gly	Ile	Gln	Trp	Asn	Thr
65					70					75					80
Ala	Trp	Glu	Trp	Gln	Gly	Ala	Pro	Asp	Asn	Val	Lys	Asn	Tyr	Pro	Tyr
				85					90					95	
Val	Gly	Lys	Gln	Ile	Gln	Arg	Gly	Arg	Lys	Ile	Ser	Asp	Ile	Asn	Ser
			100					105					110		
Met	Arg	Thr	Ser	Val	Ser	Trp	Thr	Tyr	Asp	Arg	Thr	Asp	Leu	Arg	Ala
		115						120				125			
Asn	Val	Ala	Tyr	Asp	Val	Phe	Thr	Ala	Arg	Asp	Pro	Asp	His	Pro	Asn
	130					135					140				
Trp	Gly	Gly	Asp	Tyr	Glu	Leu	Met	Ile	Trp	Leu	Ala	Arg	Tyr	Gly	Gly
145					150					155					160
Ile	Tyr	Pro	Ile	Gly	Thr	Phe	His	Ser	Gln	Val	Asn	Leu	Ala	Gly	Arg
				165					170					175	
Thr	Trp	Asp	Leu	Trp	Thr	Gly	Tyr	Asn	Gly	Asn	Met	Arg	Val	Tyr	Ser
			180					185					190		
Phe	Leu	Pro	Pro	Ser	Gly	Asp	Ile	Arg	Asp	Phe	Ser	Cys	Asp	Ile	Lys
		195						200				205			
Asp	Phe	Phe	Asn	Tyr	Leu	Glu	Arg	Asn	His	Gly	Tyr	Pro	Ala	Arg	Glu
	210					215					220				
Gln	Asn	Leu	Ile	Val	Tyr	Gln	Val	Gly	Thr	Glu	Cys	Phe	Thr	Gly	Gly
225					230					235					240
Pro	Ala	Arg	Phe	Thr	Cys	Arg	Asp	Phe	Arg	Ala	Asp	Leu	Trp		
				245					250						

<210> 22
 <211> 247
 <212> PRT
 <213> Chaetomium brasiliense

10

<400> 22

Met	Lys	Leu	Thr	Leu	Val	Leu	Phe	Val	Ser	Ser	Leu	Ala	Ala	Ala	Thr
1				5					10					15	
Pro	Leu	Gly	Trp	Arg	Glu	Arg	Arg	Gln	Gln	Val	Ser	Leu	Cys	Gly	Gln
			20					25					30		
Ser	Ser	Ser	Trp	Ser	Gly	Asn	Gly	Tyr	Gln	Leu	Asn	Asn	Asn	Leu	Trp
		35					40					45			
Gly	Gln	Ser	Arg	Ala	Thr	Ser	Gly	Ser	Gln	Cys	Thr	Tyr	Leu	Asp	Ser
	50					55					60				
Ser	Ser	Asn	Ser	Gly	Ile	His	Trp	His	Thr	Thr	Trp	Thr	Trp	Glu	Gly
65					70						75				80
Gly	Glu	Gly	Glu	Val	Lys	Ser	Tyr	Ala	Tyr	Ser	Gly	Arg	Gln	Val	Ser
				85						90				95	
Thr	Gly	Leu	Thr	Ile	Ala	Ser	Ile	Asp	Ser	Met	Gln	Thr	Ser	Val	Ser
			100					105					110		
Trp	Glu	Tyr	Asn	Thr	Thr	Asp	Ile	Gln	Ala	Asn	Val	Ala	Tyr	Asp	Ile
		115						120				125			
Phe	Thr	Ala	Glu	Asp	Pro	Asp	His	Glu	His	Ser	Ser	Gly	Asp	Tyr	Glu
		130					135				140				
Val	Met	Ile	Trp	Leu	Ala	Arg	Tyr	Asn	Asn	Val	Ser	Pro	Ile	Gly	Ser
145					150					155					160

ES 2 501 190 T3

Ser Val Ala Thr Ala Thr Val Gly Gly Asp Thr Trp Asp Leu Phe Ala
 165 170 175
 Gly Ala Asn Gly Asp Met Glu Val Tyr Ser Phe Val Ala Glu Asn Thr
 180 185 190
 Met Asn Ser Phe Ser Gly Asp Val Lys Asp Phe Phe Asp Tyr Leu Glu
 195 200 205
 Gln Asn Val Gly Phe Pro Val Asp Asp Gln Tyr Leu Leu Val Phe Glu
 210 215 220
 Leu Gly Ser Glu Ala Phe Thr Gly Gly Pro Ala Thr Leu Ser Val Ser
 225 230 235 240
 Gln Phe Ser Ala Asn Ile Ala
 245

<210> 23

<211> 357

<212> PRT

5 <213> Bionectria ochroleuca

<400> 23

Met Lys Ser Ile Ile Ser Phe Phe Gly Leu Ala Thr Leu Val Ala Ala
 1 5 10 15
 Ala Pro Ser Gln Asn Pro Thr Arg Thr Gln Pro Leu Glu Lys Arg Ala
 20 25 30
 Thr Thr Leu Cys Gly Gln Trp Asp Ser Val Glu Thr Gly Tyr Thr
 35 40 45
 Ile Tyr Asn Asn Leu Trp Gly Gln Asp Asn Gly Ser Gly Ser Gln Cys
 50 55 60
 Leu Thr Val Glu Gly Val Thr Asp Gly Leu Ala Ala Trp Ser Ser Thr
 65 70 75 80
 Trp Ser Trp Ser Gly Gly Ser Ser Ser Val Lys Ser Tyr Ser Asn Ala
 85 90 95
 Val Leu Ser Ala Glu Ala Ala Arg Ile Ser Ala Ile Ser Ser Ile Pro
 100 105 110
 Ser Lys Trp Glu Trp Ser Tyr Thr Gly Thr Asp Ile Val Ala Asn Val
 115 120 125
 Ala Tyr Asp Leu Phe Ser Asn Thr Asp Cys Gly Asp Thr Pro Glu Tyr
 130 135 140
 Glu Ile Met Ile Trp Leu Ser Ala Leu Gly Gly Ala Gly Pro Ile Ser
 145 150 155 160
 Ser Thr Gly Ser Ser Ile Ala Thr Val Thr Ile Ala Gly Ala Ser Trp
 165 170 175
 Asn Leu Trp Gln Gly Gln Asn Asn Gln Met Thr Val Phe Ser Phe Val
 180 185 190
 Ala Glu Ser Asp Gln Lys Ser Phe Ser Gly Asp Leu Asn Asp Phe Ile
 195 200 205
 Gln Tyr Leu Val Asp Ser Gln Gly Tyr Ser Gly Ser Gln Cys Leu Tyr
 210 215 220
 Ser Ile Gly Ala Gly Thr Glu Pro Phe Thr Gly Thr Asp Ala Glu Phe
 225 230 235 240
 Ile Thr Thr Gly Tyr Ser Val Ser Val Ser Ala Gly Asp Ser Gly Ser
 245 250 255
 Asp Glu Thr Thr Thr Ser Ser Gln Ala Gln Ser Ser Thr Val Glu Thr
 260 265 270
 Ser Thr Ala Thr Gln Pro Gln Ser Ser Ser Thr Val Val Pro Thr Val
 275 280 285
 Thr Leu Ser Gln Pro Ser Asn Glu Ser Thr Thr Thr Pro Val Gln Ser
 290 295 300
 Gln Pro Ser Ser Val Glu Thr Thr Pro Thr Ala Gln Pro Gln Ser Ser
 305 310 315 320
 Ser Val Gln Thr Thr Thr Ala Gln Ala Gln Pro Thr Pro Glu Arg
 325 330 335

 Ala Ala Pro Asp Ala Gly Ser Ala Glu Leu Leu Ser Ser Ala Thr Met
 340 345 350
 His Leu Asp Arg Arg
 355

ES 2 501 190 T3

Met Lys Leu Leu Ala Leu Ser Leu Val Ser Leu Ala Ser Ala Ala Ser
 1 5 10 15
 Ala Ala Ser Ile Leu Ser Asn Thr Phe Thr Arg Arg Ser Asp Phe Cys
 20 25 30
 Gly Gln Trp Asp Thr Ala Thr Val Gly Asn Phe Ile Val Tyr Asn Asn
 35 40 45
 Leu Trp Gly Gln Asp Asn Ala Asp Ser Gly Ser Gln Cys Thr Gly Val
 50 55 60
 Asp Ser Ala Asn Gly Asn Ser Ile Ser Trp His Thr Thr Trp Ser Trp
 65 70 75 80
 Ser Gly Gly Ser Ser Ser Val Lys Ser Tyr Ala Asn Ala Ala Tyr Gln
 85 90 95
 Phe Thr Ser Thr Lys Leu Asn Ser Leu Ser Ser Ile Pro Thr Ser Trp
 100 105 110
 Lys Trp Gln Tyr Ser Thr Thr Asp Ile Val Ala Asn Val Ala Tyr Asp
 115 120 125
 Leu Phe Thr Ser Ser Ser Ala Gly Gly Asp Ser Glu Tyr Glu Ile Met
 130 135 140
 Ile Trp Leu Ala Ala Leu Gly Gly Ala Gly Pro Ile Ser Ser Thr Gly
 145 150 155 160
 Ser Ser Ile Ala Thr Val Thr Leu Gly Gly Val Thr Trp Ser Leu Tyr
 165 170 175
 Ser Gly Pro Asn Gly Ser Met Gln Val Tyr Ser Phe Val Ala Ser Ser
 180 185 190
 Thr Thr Glu Ser Phe Ser Ala Asp Leu Met Asp Phe Ile Asn Tyr Leu
 195 200 205
 Ala Glu Asn Gln Gly Leu Ser Ser Ser Gln Tyr Leu Thr His Val Gln
 210 215 220
 Ala Gly Thr Glu Pro Phe Thr Gly Thr Asp Ala Thr Leu Thr Val Ser
 225 230 235 240
 Ser Tyr Ser Val Ser Val Ser
 245

5

<210> 24
 <211> 247
 <212> PRT
 <213> Emericella desertorum

10

<400> 24

ES 2 501 190 T3

Met Lys Leu Leu Ala Leu Ser Leu Val Ser Leu Ala Ser Ala Ala Ser
 1 5 10 15
 Ala Ala Ser Ile Leu Ser Asn Thr Phe Thr Arg Arg Ser Asp Phe Cys
 20 25 30
 Gly Gln Trp Asp Thr Ala Thr Val Gly Asn Phe Ile Val Tyr Asn Asn
 35 40 45
 Leu Trp Gly Gln Asp Asn Ala Asp Ser Gly Ser Gln Cys Thr Gly Val
 50 55 60
 Asp Ser Ala Asn Gly Asn Ser Ile Ser Trp His Thr Thr Trp Ser Trp
 65 70 75 80
 Ser Gly Gly Ser Ser Ser Val Lys Ser Tyr Ala Asn Ala Ala Tyr Gln
 85 90 95
 Phe Thr Ser Thr Lys Leu Asn Ser Leu Ser Ser Ile Pro Thr Ser Trp
 100 105 110
 Lys Trp Gln Tyr Ser Thr Thr Asp Ile Val Ala Asn Val Ala Tyr Asp
 115 120 125
 Leu Phe Thr Ser Ser Ser Ala Gly Gly Asp Ser Glu Tyr Glu Ile Met
 130 135 140
 Ile Trp Leu Ala Ala Leu Gly Gly Ala Gly Pro Ile Ser Ser Thr Gly
 145 150 155 160
 Ser Ser Ile Ala Thr Val Thr Leu Gly Gly Val Thr Trp Ser Leu Tyr
 165 170 175
 Ser Gly Pro Asn Gly Ser Met Gln Val Tyr Ser Phe Val Ala Ser Ser
 180 185 190
 Thr Thr Glu Ser Phe Ser Ala Asp Leu Met Asp Phe Ile Asn Tyr Leu
 195 200 205
 Ala Glu Asn Gln Gly Leu Ser Ser Ser Gln Tyr Leu Thr His Val Gln
 210 215 220
 Ala Gly Thr Glu Pro Phe Thr Gly Thr Asp Ala Thr Leu Thr Val Ser
 225 230 235 240
 Ser Tyr Ser Val Ser Val Ser
 245

<210> 25

<211> 244

5

<212> PRT

<213> *Fusarium solani*

<400> 25

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

ES 2 501 190 T3

```

      85          90          95
Thr Leu Gly Glu Val Ser Ser Ile Pro Ser Ala Trp Asp Trp Thr Tyr
      100          105          110
Thr Gly Asn Gly Ile Ile Ala Asn Val Ala Tyr Asp Leu Phe Thr Ser
      115          120          125
Ser Thr Glu Ser Gly Asp Ala Glu Tyr Glu Phe Met Ile Trp Leu Ser
      130          135          140
Ala Leu Gly Gly Ala Gly Pro Ile Ser Asn Asp Gly Ser Pro Val Ala
      145          150          155
Thr Val Glu Leu Ala Gly Thr Ser Trp Lys Leu Tyr Gln Gly Lys Asn
      165          170          175
Asn Gln Met Thr Val Phe Ser Phe Val Ala Glu Ser Asp Val Asn Asn
      180          185          190
Phe Cys Gly Asp Leu Ala Asp Phe Thr Asp Tyr Leu Val Asp Asn His
      195          200          205
Gly Val Ser Ser Ser Gln Ile Leu Gln Ser Val Gly Ala Gly Thr Glu
      210          215          220
Pro Phe Glu Gly Thr Asn Ala Val Phe Thr Thr Asn Asn Tyr His Ala
      225          230          235          240
Asp Val Glu Tyr

```

<210> 26

<211> 250

<212> PRT

5

<213> Fusarium solani

<400> 26

ES 2 501 190 T3

Met Lys Phe Phe Gly Val Val Ser Ala Phe Leu Ala Ala Thr Ala Val
 1 5 10 15
 Ala Thr Pro Thr Thr Pro Thr Glu Thr Ile Glu Lys Arg Asp Thr Thr
 20 25 30
 Trp Cys Asp Ala Phe Gly Ser Leu Ala Thr Ser Gly Tyr Thr Val Tyr
 35 40 45
 His Asn Asn Trp Gly Lys Gly Asp Ala Thr Ser Gly Ser Gln Cys Thr
 50 55 60
 Thr Phe Thr Ser Val Ser Asn Asn Asn Phe Val Trp Ser Thr Ser Trp
 65 70 75 80
 Thr Trp Ala Gly Gly Ala Gly Lys Val Lys Ser Tyr Ser Asn Val Ala
 85 90 95
 Leu Glu Lys Ile Asn Lys Lys Ile Ser Asp Ile Lys Ser Val Ser Thr
 100 105 110
 Arg Trp Ile Trp Arg Tyr Thr Gly Thr Lys Met Ile Ala Asn Val Ser
 115 120 125
 Tyr Asp Leu Trp Phe Ala Pro Thr Ala Ser Ser Asn Asn Ala Tyr Glu
 130 135 140
 Ile Met Ile Trp Val Gly Ala Tyr Gly Gly Ala Leu Pro Ile Ser Thr
 145 150 155 160
 Pro Gly Lys Gly Val Ile Asp Arg Pro Thr Leu Ala Gly Ile Pro Trp
 165 170 175
 Asp Val Tyr Lys Gly Pro Asn Gly Asp Val Thr Val Ile Ser Phe Val
 180 185 190
 Ala Ser Ser Asn Gln Gly Asn Phe Gln Ala Asp Leu Lys Glu Phe Leu
 195 200 205
 Asn Tyr Leu Thr Ser Lys Gln Gly Leu Pro Ser Asn Tyr Val Ala Thr
 210 215 220
 Ser Phe Gln Ala Gly Thr Glu Pro Phe Glu Gly Thr Asn Ala Val Leu
 225 230 235 240
 Lys Thr Ser Ala Tyr Thr Ile Ser Val Asn
 245 250

<210> 27

<211> 371

5

<212> PRT

<213> Streptomyces sp. 11AG8

<400> 27

ES 2 501 190 T3

Met Arg Ser His Pro Arg Ser Ala Thr Met Thr Val Leu Val Val Leu
 1 5 10 15
 Ala Ser Leu Gly Ala Leu Leu Thr Ala Ala Ala Pro Ala Gln Ala Asn
 20 25 30
 Gln Gln Ile Cys Asp Arg Tyr Gly Thr Thr Thr Ile Gln Asp Arg Tyr
 35 40 45
 Val Val Gln Asn Asn Arg Trp Gly Thr Ser Ala Thr Gln Cys Ile Asn
 50 55 60
 Val Thr Gly Asn Gly Phe Glu Ile Thr Gln Ala Asp Gly Ser Val Pro
 65 70 75 80
 Thr Asn Gly Ala Pro Lys Ser Tyr Pro Ser Val Tyr Asp Gly Cys His
 85 90 95
 Tyr Gly Asn Cys Ala Pro Arg Thr Thr Leu Pro Met Arg Ile Ser Ser
 100 105 110
 Ile Gly Ser Ala Pro Ser Ser Val Ser Tyr Arg Tyr Thr Gly Asn Gly
 115 120 125
 Val Tyr Asn Ala Ala Tyr Asp Ile Trp Leu Asp Pro Thr Pro Arg Thr
 130 135 140
 Asn Gly Val Asn Arg Thr Glu Ile Met Ile Trp Phe Asn Arg Val Gly
 145 150 155 160
 Pro Val Gln Pro Ile Gly Ser Pro Val Gly Thr Ala His Val Gly Gly
 165 170 175
 Arg Ser Trp Glu Val Trp Thr Gly Ser Asn Gly Ser Asn Asp Val Ile
 180 185 190
 Ser Phe Leu Ala Pro Ser Ala Ile Ser Ser Trp Ser Phe Asp Val Lys
 195 200 205
 Asp Phe Val Asp Gln Ala Val Ser His Gly Leu Ala Thr Pro Asp Trp
 210 215 220
 Tyr Leu Thr Ser Ile Gln Ala Gly Phe Glu Pro Trp Glu Gly Gly Thr
 225 230 235 240
 Gly Leu Ala Val Asn Ser Phe Ser Ser Ala Val Asn Ala Gly Gly Gly
 245 250 255
 Asn Gly Gly Thr Pro Gly Thr Pro Ala Ala Cys Gln Val Ser Tyr Ser
 260 265 270
 Thr His Thr Trp Pro Gly Gly Phe Thr Val Asp Thr Thr Ile Thr Asn
 275 280 285
 Thr Gly Ser Thr Pro Val Asp Gly Trp Glu Leu Asp Phe Thr Leu Pro
 290 295 300
 Ala Gly His Thr Val Thr Ser Val Trp Asn Ala Leu Ile Ser Pro Ala
 305 310 315 320
 Ser Gly Ala Val Thr Ala Arg Ser Thr Gly Ser Asn Gly Arg Ile Ala
 325 330 335
 Ala Asn Gly Gly Thr Gln Ser Phe Gly Phe Gln Gly Thr Ser Ser Gly
 340 345 350
 Ala Gly Phe Thr Ala Pro Ala Gly Ala Arg Leu Asn Gly Thr Ser Cys
 355 360 365
 Thr Val Arg
 370

<210> 28

<211> 221

<212> PRT

5 <213> conSecuencia Artificial

<220>

<223> consensus secuencia

<221> VARIANTE

<222> (1)...(221)

10 <223> Xaa = cualquier Amino Acid

<400> 28

ES 2 501 190 T3

Cys Xaa Gln Tyr Xaa Xaa Xaa Xaa Xaa Xaa Gly Tyr Xaa Xaa Xaa Asn
 1 5 10 15
 Asn Xaa Trp Gly Xaa Xaa Xaa Xaa Xaa Ser Gly Xaa Gln Cys Thr Xaa
 20 25 30
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Trp Xaa Xaa Xaa Trp
 35 40 45
 Xaa Trp Ser Gly Gly Xaa Xaa Xaa Val Lys Ser Tyr Xaa Xaa Xaa Xaa
 50 55 60
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ser Xaa Ile Xaa Ser Xaa
 65 70 75 80
 Xaa Xaa Xaa Xaa Xaa Trp Xaa Tyr Xaa Gly Xaa Xaa Xaa Xaa Ala Asn
 85 90 95
 Val Ala Tyr Asp Leu Phe Thr Xaa Xaa Xaa Pro Xaa His Xaa Xaa Xaa
 100 105 110
 Xaa Gly Xaa Tyr Glu Xaa Met Ile Trp Leu Xaa Xaa Xaa Gly Gly Xaa
 115 120 125
 Xaa Pro Ile Gly Ser Xaa Xaa Xaa Xaa Val Xaa Xaa Xaa Xaa Xaa Xaa
 130 135 140
 Gly Xaa Xaa Trp Xaa Leu Xaa Xaa Gly Xaa Asn Gly Xaa Met Xaa Val
 145 150 155 160
 Xaa Ser Phe Val Ala Xaa Ser Ser Ser Ser Ser Ser Phe Xaa Gly Asp
 165 170 175
 Xaa Xaa Xaa Phe Xaa Xaa Tyr Leu Xaa Xaa Xaa Xaa Gly Xaa Pro Xaa
 180 185 190
 Xaa Xaa Gln Tyr Leu Xaa Xaa Xaa Gln Xaa Gly Thr Glu Pro Phe Thr
 195 200 205
 Gly Xaa Xaa Ala Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ala
 210 215 220

<210> 29

<211> 25

5 <212> ADN

<213> Secuencia Artificial

<220>

<223> oligonucleotido sintético

<400> 29

10 gaacgatggc aagggcgggc tgacg 25

<210> 30

<211> 27

<212> ADN

<213> Secuencia Artificial

15 <220>

<223> oligonucleotido sintético

<400> 30

cttctcgggc tgctacaacc caaacgg 27

<210> 31

20 <211> 26

<212> ADN

<213> Secuencia Artificial
 <220>
 <223> oligonucleotido sintético
 <400> 31
 5 acatcgtcga gtgtttggc acctac 26
 <210> 32
 <211> 29
 <212> ADN
 <213> Secuencia Artificial
 10 <220>
 <223> oligonucleotido sintético
 <400> 32
 catcgtcgag aactggggca cctacaacc 29
 <210> 33
 15 <211> 21
 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> oligonucleotido sintético
 20 <400> 33
 ggcacctacc gaccgtccac g 21
 <210> 34
 <211> 25
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> oligonucleotido sintético
 <400> 34
 caagctgggc gagcacacct cggac 25
 30 <210> 35
 <211> 21
 <212> ADN
 <213> Secuencia Artificial
 <220>

<223> oligonucleotido sintético
 <400> 35
 cgccgcaact gtcgctcgag c 21
 <210> 36
 5 <211> 29
 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> oligonucleotido sintético
 10 <400> 36
 gtggagggtt accaaagctc tgctctgc 29
 <210> 37
 <211> 27
 <212> ADN
 15 <213> Secuencia Artificial
 <220>
 <223> oligonucleotido sintético
 <400> 37
 tctggctctg cttgcatcac cgtcagc 27
 20 <210> 38
 <211> 27
 <212> ADN
 <213> Secuencia Artificial
 <220>
 25 <223> oligonucleotido sintético
 <400> 38
 gagaagcgcc agtgcattca gcccggc 27
 <210> 39
 <211> 27
 30 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> oligonucleotido sintético
 <400> 39

gtgacgtact gcaatggtcc cggcggg 27
 <210> 40
 <211> 33
 <212> ADN
 5 <213> Secuencia Artificial
 <220>
 <223> oligonucleotido sintético
 <400> 40
 ggcaccaaga acaggtcat caacttctcg ggc 33
 10 <210> 41
 <211> 33
 <212> ADN
 <213> Secuencia Artificial
 <220>
 15 <223> oligonucleotido sintético
 <400> 41
 tccatcaccg tcagcgatta aagggggctc ttc 33
 <210> 42
 <211> 32
 20 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> oligonucleotido sintético
 <400> 42
 25 cccagacgat tcagtgcggc acgggctaca ac 32
 <210> 43
 <211> 32
 <212> ADN
 <213> Secuencia Artificial
 30 <220>
 <223> oligonucleotido sintético
 <400> 43
 ctctactcg tactggtgcg atggccacgg cg 32
 <210> 44

<211> 32
 <212> ADN
 <213> Secuencia Artificial
 <220>
 5 <223> oligonucleotido sintético
 <400> 44
 cgattcagcc cggctgcggc tacaacaacg gc 32
 <210> 45
 <211> 35
 10 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> oligonucleotido sintético
 <400> 45
 15 caacggctac ttctactgct actggaacga tggcc 35
 <210> 46
 <211> 34
 <212> ADN
 <213> Secuencia Artificial
 20 <220>
 <223> oligonucleotido sintético
 <400> 46
 ccggcaggg ctactgcaac ggctacttct actc 34
 <210> 47
 25 <211> 31
 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> oligonucleotido sintético
 30 <400> 47
 ggctgtgacgt acacctgagg tcccggcggg c 31
 <210> 48
 <211> 27
 <212> ADN

ES 2 501 190 T3

<213> Secuencia Artificial
 <220>
 <223> oligonucleotido sintético
 <400> 48
 5 ggcgccacca agtgcggcga ggtcacc 27
 <210> 49
 <211> 28
 <212> ADN
 <213> Secuencia Artificial
 10 <220>
 <223> oligonucleotido sintético
 <400> 49
 gcgtgggctc agtgcggcct gacgctcg 28
 <210> 50
 15 <211> 752
 <212> ADN
 <213> Trichoderma reesei
 <400> 50
 atggttgect tttccagcct catctgcgct ctcaccagca tgcgccagta tctggcgatg 60
 cccacaggcc tcgagcctga gagcagtgtc aacgtcacag agcgtggcat gtacgacttt 120
 gttcttggag ctcacaaatga tcacgcgct cgtgctagca tcaactacga ccaaaactac 180
 caaactggcg gacaagtcag ctattcgcct tccaactcag gcttctcagt gaactggaac 240
 actcaagatg actttgttgt gggcgttggg tggacgactg gatcttctgc gtaggaggac 300
 tcctcatcat tctgcacttt gaaagcatct tctgacccaa agcttctctt agtcccatca 360
 actttggcgg ctcttttagt gtcaacagcg gaactggcct gctttccgtc tatggctgga 420
 gcaccaaccc actggttgag tactacatca tggaggacaa ccacaactac ccagcacagg 480
 gtaccgtcaa gggaaaccgtc accagcgacg gagccactta caccatctgg gagaataccc 540
 gtgtcaacga gccttccatc cagggcacag cgaccttcaa ccagtacatt tccgtgcgga 600
 actcgcgccag gaccagcggg actgttactg tgcagaacca cttcaatgct tgggcctcgc 660
 ttggcctgca ccttgggcag atgaactacc aggttgtcgc tgtcgaaggc tggggtggtg 720
 gtggttctgc ctcacagagt gtcagcaact ag 752
 20 <210> 51
 <211> 248
 <212> PRT
 <213> Trichoderma reesei
 <400> 51

ES 2 501 190 T3

Met Val Ala Phe Ser Ser Leu Ile Cys Ala Leu Thr Ser Ile Ala Ser
 1 5 10 15
 Thr Leu Ala Met Pro Thr Gly Leu Glu Pro Glu Ser Ser Val Asn Val
 20 25 30
 Thr Glu Arg Gly Met Tyr Asp Phe Val Leu Gly Ala His Asn Asp His
 35 40 45
 Arg Arg Arg Ala Ser Ile Asn Tyr Asp Gln Asn Tyr Gln Thr Gly Gly
 50 55 60
 Gln Val Ser Tyr Ser Pro Ser Asn Thr Gly Phe Ser Val Asn Trp Asn
 65 70 75 80
 Thr Gln Asp Asp Phe Val Val Gly Val Gly Trp Thr Thr Gly Ser Ser
 85 90 95
 Ala Glu Asp Ser Ser Ser Phe Cys Thr Leu Lys Ala Ser Ser Asp Gln
 100 105 110
 Lys Leu Leu Leu Val Pro Ser Thr Leu Ala Ala Leu Leu Val Ser Thr
 115 120 125
 Ala Glu Leu Ala Cys Phe Pro Ser Met Ala Gly Ala Pro Thr His Trp
 130 135 140
 Leu Ser Thr Thr Ser Trp Arg Thr Thr Thr Thr Thr Gln His Arg Val
 145 150 155 160
 Pro Ser Arg Glu Pro Ser Pro Ala Thr Glu Pro Leu Thr Pro Ser Gly
 165 170 175
 Arg Ile Pro Val Ser Thr Ser Leu Pro Ser Arg Ala Gln Arg Pro Ser
 180 185 190
 Thr Ser Thr Phe Pro Cys Gly Thr Arg Pro Gly Pro Ala Glu Leu Leu
 195 200 205
 Leu Cys Arg Thr Thr Ser Met Leu Gly Pro Arg Leu Ala Cys Thr Leu
 210 215 220
 Gly Arg Thr Thr Arg Leu Ser Leu Ser Lys Ala Gly Val Val Val Val
 225 230 235 240
 Leu Pro His Arg Val Ser Ala Thr
 245

Reivindicaciones

1. Un ácido nucleico, codificando una xilanasas modificada comprendiendo un polipéptido con una secuencia de aminoácidos según se presenta en SEQ ID NO:1;

donde la secuencia tiene las mutaciones T2C, T28C, K58R y +191D ;

5 y donde la secuencia tiene mutaciones adicionales seleccionadas del grupo que consiste en:

- (i) F93W, N97R y H144K,
(ii) H144C y N92C,
(iii) F180Q, H144C y N92C,
10 (iv) V108H,
(v) S65C y S 186C,
(vi) H22K, F180Q, H144C y N92C;

donde la posición de los aminoácidos sustituidos está numerada desde el aminoácido después de la señal y la prosequencia.

2. Una xilanasas modificada comprendiendo un polipéptido con una secuencia de aminoácido según se
15 presenta en SEQ ID NO:1;

donde la secuencia tiene las mutaciones T2C, T28C, K58R y +191D;

y donde la secuencia comporta las mutaciones adicionales seleccionadas del grupo que consiste en:

- (i) F93W, N97R y H144K,
20 (ii) H144C y N92C,
(iii) F180Q, H144C y N92C,
(iv) V108H,
(v) S65C y S 186C,
(vi) H22K, F180Q, H144C y N92C;

25 donde la posición del aminoácido sustituido está numerada a partir del aminoácido después de la señal y la prosequencia.

Figura 1

```

      10      20      30      40      50
XYN2_TRIRE ---MVSFTSLLAASPP-SRASC RPAAEV---ESVAVEKRQTIQ-----P
XYN1_HUMIN ---MVSLKSVLAAATAVSSAIAAPFDVPRDNSTALQARQVTP-----N
XYN2_BACST -----MKLKKKMLTLLLTASMSFGLF-----G
XYN1_TRIRE ---MVAFSSLICALTSIASTLAMPTGLEPESSVNVYTERGMYDFVLGAHND
XYN1_ASPAW -----MKVTAAFAGLLVTAFAAPVPEPVLVS-----
XYN2_BACST MCSSIPSLREVFANDFRIGAAVNPVTLAQQLLIRHVNSLTAENHMKFE

      60      70      80      90     100
XYN2_TRIRE GTGYNNGYFYFYSYWNDDGHGGVYTYTNGPGGQFSVNWS--NSG-NFVGGKGWQ
XYN1_HUMIN AEGWHNGYFYFYSWWSDDGGGQVQYTNLEGSRYQVRWR--NTG-NFVGGKGWN
XYN2_BACST ATSSAATDYWQYWTDDGGGMVNAVNGPGGNYSVTWQ--NTG-NFVVGKGT
XYN1_TRIRE HRRRASINYDQNYQTG-GQVSYSPSNTG-FSVNWN--TQD-DFVVGVGWT
XYN1_ASPAW --RSAGINYVQNYNGNLGDFTYDESAGT-FSMYWEDGVSS-DFVVGGLWT
XYN2_BACST HLQPEEGRFTFDIAIKSSTSPFSSHGVRGHTLVVHNQTPSWVFQDSQGHF

      110     120     130     140     150
XYN2_TRIRE PGTKNKVINFS-GSYNPNGNSYLSVYGWGRNPLIEYYIVENF---GTYNP
XYN1_HUMIN PGT-GRTINYG-GYFNPQNGYLAVYGWTRNPLVEYYVIESY---GTYNP
XYN2_BACST VGSPPNRVINYNAGIWEPSGNGYLTLYGWRNALIEYYVVDSW---GTYRP
XYN1_TRIRE TGS-SAPINFGGSFVSNSGTGLLSVYGWSTNPLVEYYIMED-----NHNY
XYN1_ASPAW TGS-SNAITYSAEYSASGSSSYLAVYGWVNYPQAEYYIVEDY---GDYNP
XYN2_BACST VGRDVLLERMKSHISTVVQRYKGVYCWVINEAVADEGSEWLRSSSTWRQ

      160     170     180     190     200
XYN2_TRIRE STGATKLGCVTSDGSVYDIYRTQRVNQPSIIGTATFYQYWSVRRNHRSSG
XYN1_HUMIN GSQAQYKGTFFYTDGDQYDIFVSTRYNQPSIDGTRTFQYQYWSIRKNKRVGG
XYN2_BACST T--GNFKGTVNSDGGTYDIYTTMRYNAPSIDGTQTFQYQYWSVRSKRPTG
XYN1_TRIRE PAQGTVKGTVTSIDGATYTIWENTRVNEPSIQGTATFNQYISVRNSPRTSG
XYN1_ASPAW CSSATSLGTVYSDGSTYQVCTDTRTNEPSITGTSTFTQYFVRESRTSG
XYN2_BACST IIGDDFIQQAFLYAHEADPEALLFYNDYNECFPEKREKIYTLVKSLRDKG

      210     220     230     240     250
XYN2_TRIRE S----VNTANHFNAWA-QQGLTLGTMD-YQIVAVEGYFSSGSASITVS--
XYN1_HUMIN S----VNMQNHFNWQ-QHGMLGQHY-YQVVATEGYQSSGESDIYVQTH
XYN2_BACST SNV-SITFSNHVNAWR-SKGMNLGSSWAYQVLATEGYQSSGRSNVTVW--
XYN1_TRIRE T----VTVQNHFNWQ-SLGLHLGQMN-YQVVAVEGWGGSGSASQSVSN-
XYN1_ASPAW T----VTVANHFNFWA-QHGFNSDPN-YQVMAVEAWSGAGSASVTISS-
XYN2_BACST IPIHGIGMQAHWSLNRPTLDEIRAAIERYASLGVILHITELDISMFEFDD

      260     270     280     290     300
XYN2_TRIRE -----

```

Figura 1

Figura 1

```
XYN1_HUMIN -----  
XYNA_BACST -----  
XYN1_TRIRE -----  
XYN1_ASPAW -----  
XYN2_BACST HRKDIAAPTNE MVERQAERYEQIFSLFKEYRDVIQNVTFWGIADDHTWLD
```

```
          310      320      330  
          |        |        |  
XYN2_TRIRE -----  
XYN1_HUMIN -----  
XYNA_BACST -----  
XYN1_TRIRE -----  
XYN1_ASPAW -----  
XYN2_BACST HFPVQGRKNWPLLEFDEQHNPKPAFWRVVNI
```

Figura 2

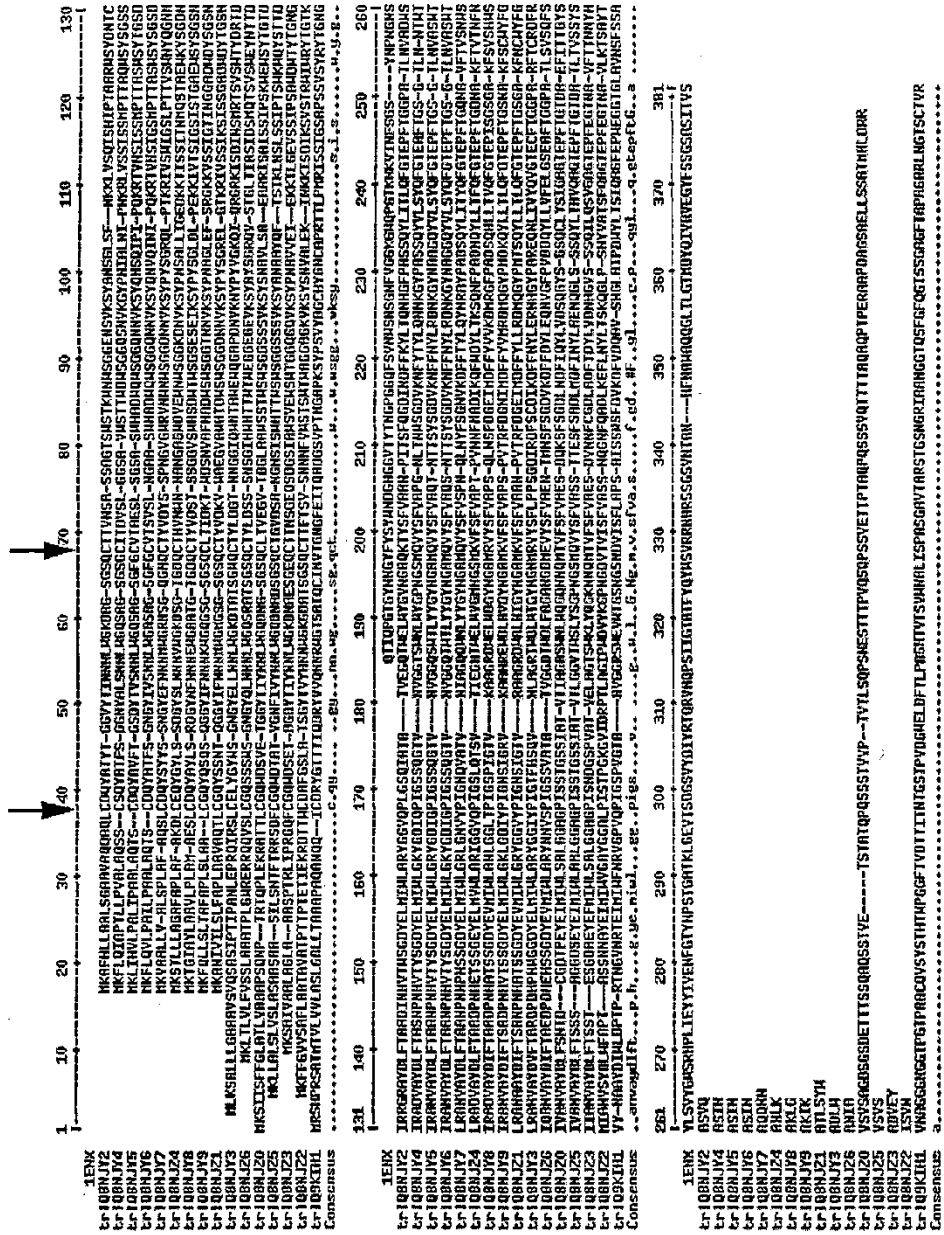


Figura 3

H22K	5'- GAACGATGGCAAGGGCGGCGTGACG -3'
S65C	5'- CTTCTCGGGCTGCTACAACCCAAACGG -3'
N92C	5'- ACATCGTCGAGTGT ^{TTTT} GGCACCTAC -3'
F93W	5'- CATCGTCGAGAACTGGGGGCACCTACAACC -3'
N97R	5'- GGCACCTACCGACCGTCCACG -3'
V108H	5'- CAAGCTGGGCGAGCACACCTCCGAC -3'
H144C	5'- CGCCGCAACTGT ^{CGCT} CGAGC -3'
F180Q	5'- GTGGAGGGTTACCAAAGCTCTGGCTCTGC -3'
S186C	5'- TCTGGCTCTGCTTGCATCACCGTCAGC -3'
T2C	5'-GAGAAGCGCCAGTGCATTAGCCCGGC-3'
T28C	5'-GTGACGTACTGCAATGGTCCC ^{GGCGGG} -3'
K58R	5'-GGCACCAAGAACAGGGT ^{CATCAACTTCTCGGGC} -3'
191D	5'-TCCATCACCGTCAGCGATTAAAGGGGGCTCTTC-3'
P5C	5'-CCCAGACGATTAGTGC ^{GGCACGGGCTACAAC} -3'
N19C	5'-CTTCTACTCGTACTGGTGC ^{GATGGCCACGGCG} -3'
T7C	5'-CGATTAGCCCGGCTGC ^{GGCTACAACAACGGC} -3'
S16C	5'-CAACGGTACTTCTACTGCTACTGGAACGATGGCC-3'
N10C	5'-CCGGCACGGGCTACTGCAACGGTACTTCTACTC-3'
N29C	5'-GGCGTGACGTACACCTGC ^{GGTCCC} GGCGGGC-3'
L105C	5'-GGCGCCACCAAGTGC ^{GGCGAGGTCACC} -3'
Q162C	5'-GCGTGGGCTCAGTGC ^{GGCCTGACGCTCG} -3'

Figura 4

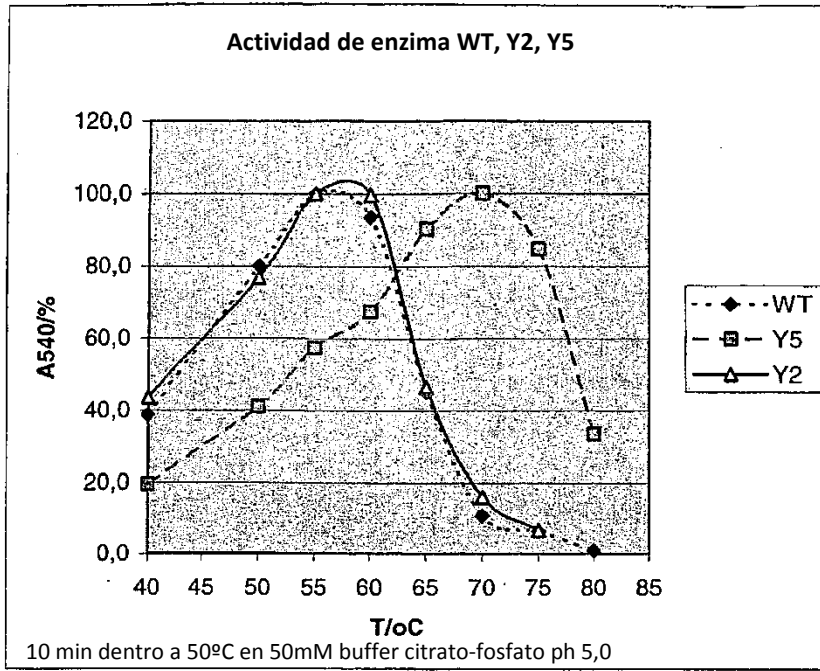


Figura 5

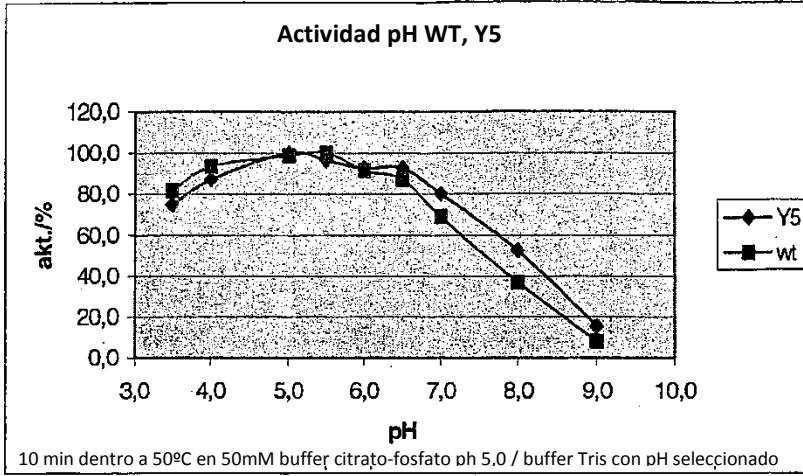


Figura 6

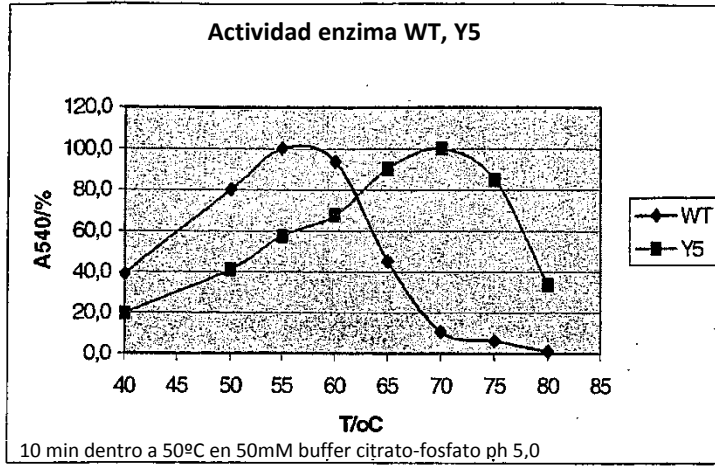


Figura 7

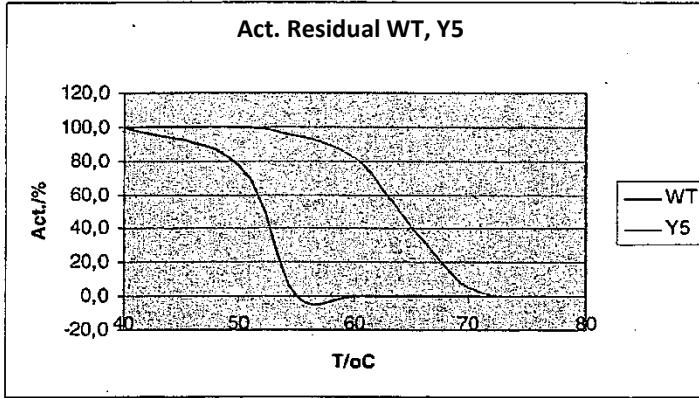


Figura 8

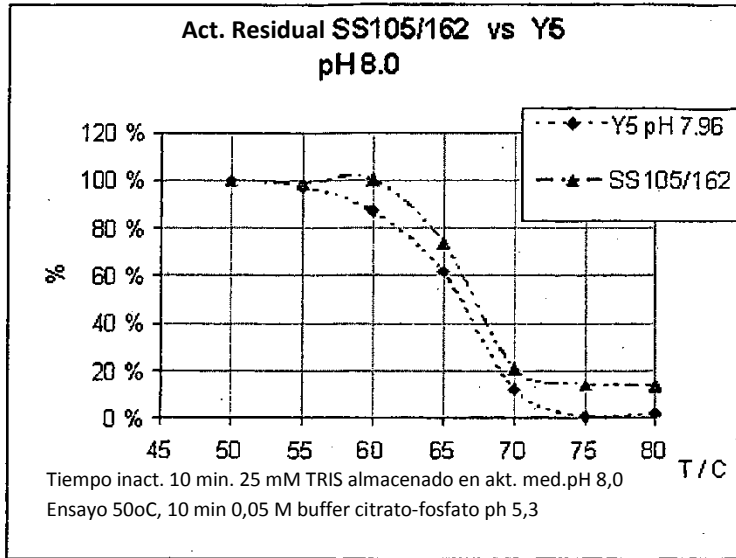


Figura 9

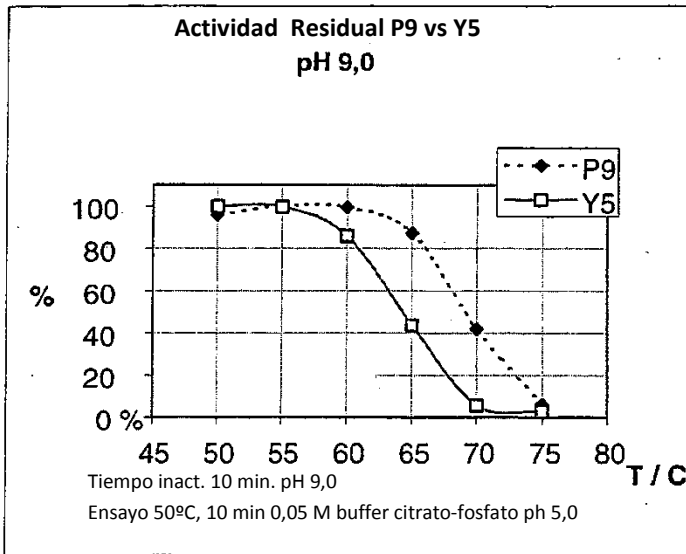


Figura 10

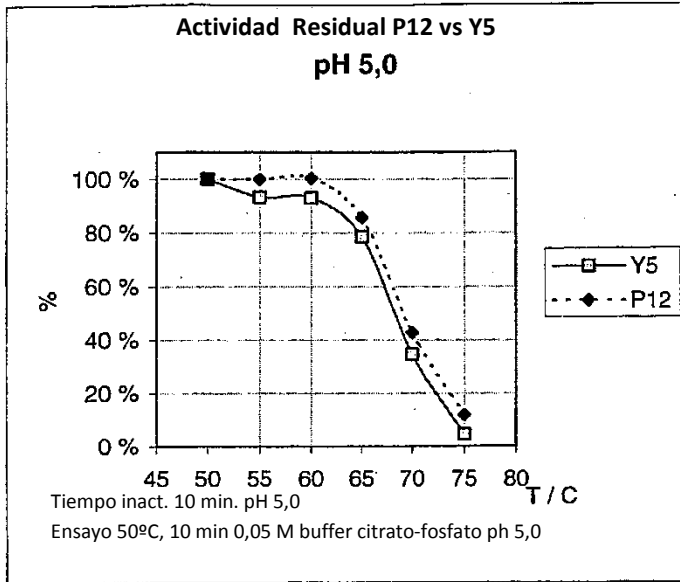


Figura 11

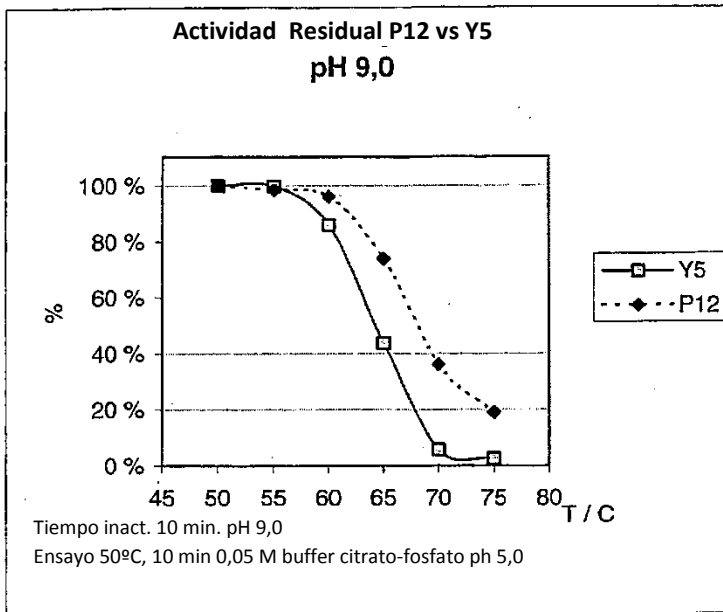


Figura 12

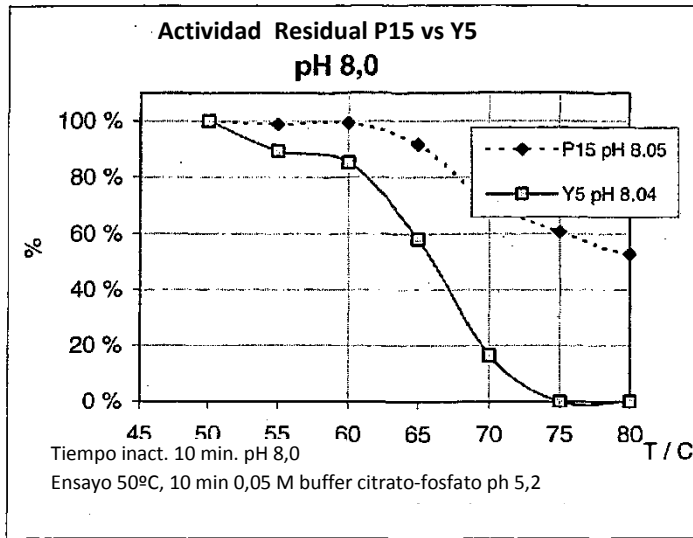


Figura 13

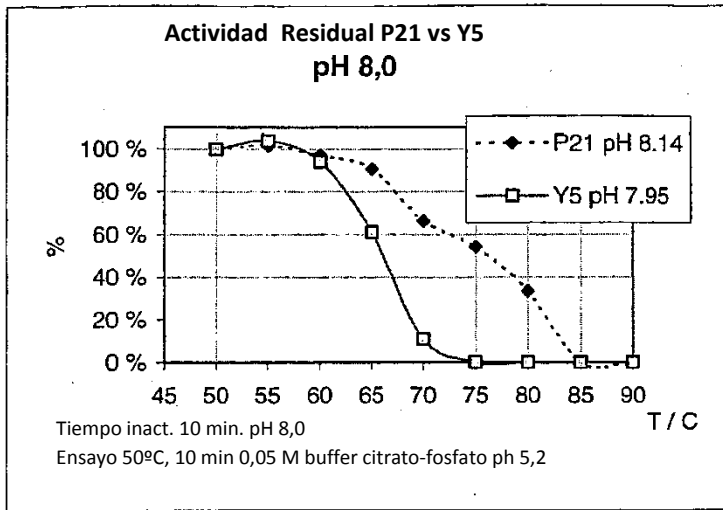


Figura 14

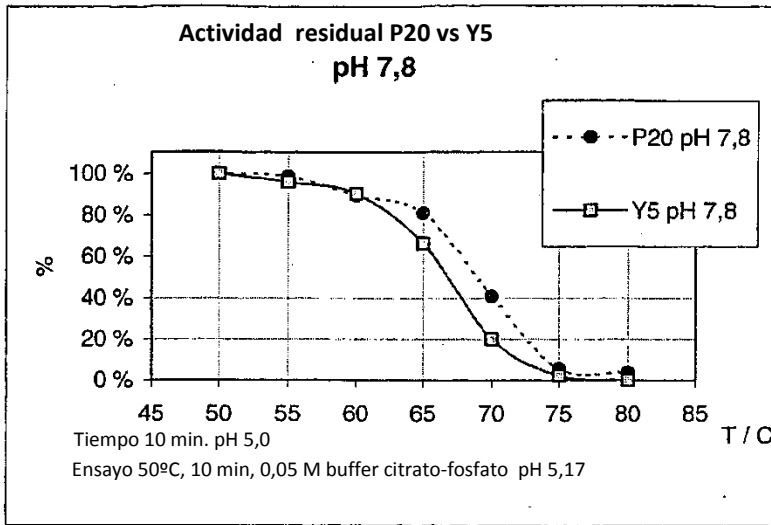


Figura 15

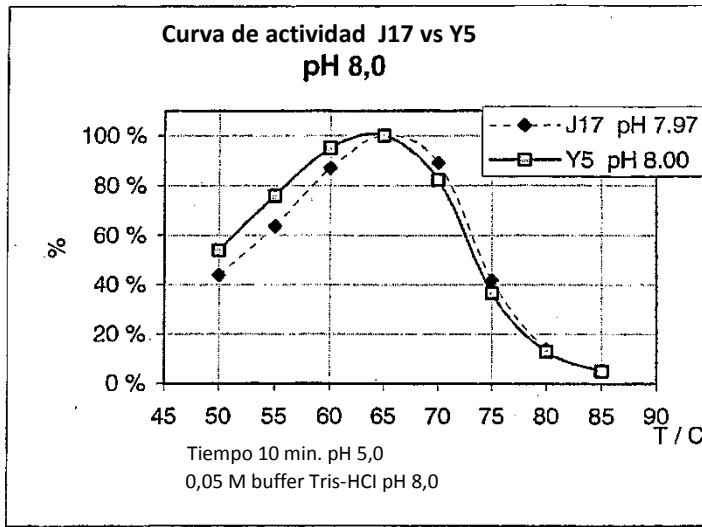


Figura 16

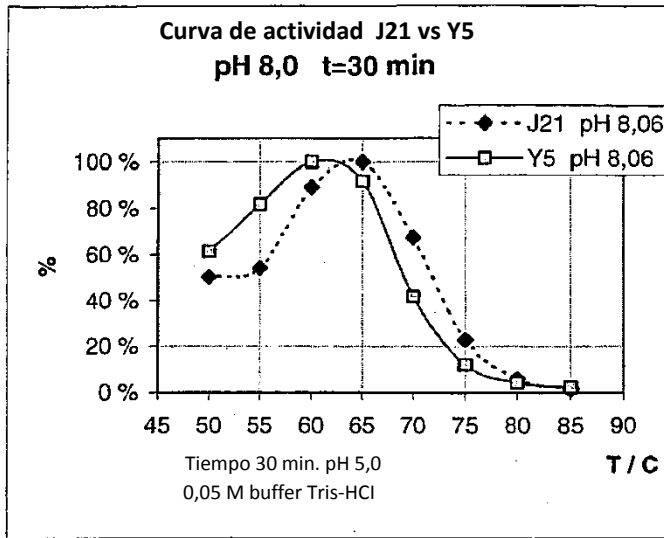


Figura 17

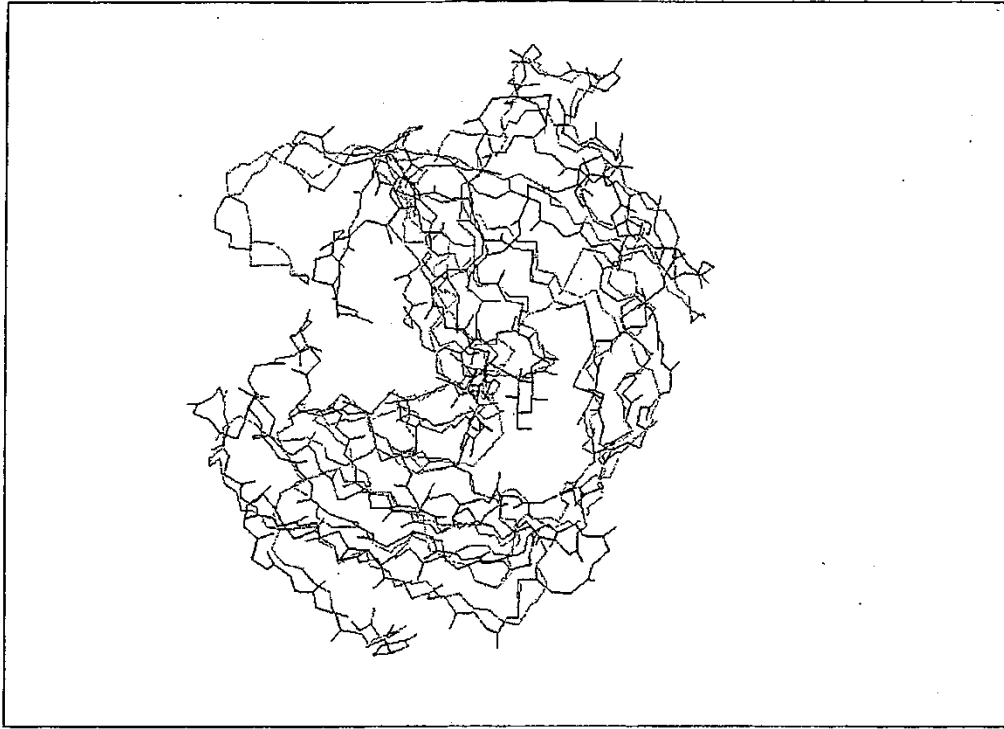


Figura 18

Proteína trichoderma reesei Xyl II (xilanaso alto pl)
 la secuencia total, incluyendo la señal y la pro secuencia

MVSTLSLLAGVAISGVLAAAPAAEVESVAVEKRTIQPFGTGYNNGYFYSYWNDDHGGVVTYNGPQQFSDVNNWSNSGNFVVG
 GKGWQPGTKNKVINFSGSYNPNNGNSYLSVYGWSRNPLEYIYIVENFGTYNPFSTGATKLGVEVTSDDGSVYDIYRTQRVNPQS
 IIGTATFYQYWSVRRNHRSSGSVNTANHFNAWAQQQLTLGTMDYQIVAVEGYFSSGSASITVS (SEQ. ID NO: 1)

Gen trichoderma reesei Xyl II (xilanaso alto pl)
 ADN desde que empieza el codón hasta que para el codón (incluye un único intrón)

ATGGTCTCCTCACCTCCCTCCGCGGGTCCGCGCCATCTCGGGCGTCTGGGGCGTCTCCGCGCGGAGGTGGAATC
 CGTGGCTGTGGAGAAGCGCCAGACGATTTCAGCCGGCACGGGCTACAACAACGGCTACTTCTACTCGTACTGGAACGATG
 GCCACGGCGGCGTGACGTACACCAATGGTCCCGGGCGGAGTTCTCCGTCAACTGGTCCAACCTCGGCAACTTTGTCTGGC
 GGCAAGGGATGGCAGCCCGGCCCAAGAACAAGTAAGACTACCTACTCTTACCCCTTTGACCAACACAGCACACACAAA
 TACAACACATGTGACTACCAATCATGGAATCGGATCTAACAGCTGTGTTTTCAAAAAAAGGGTCACTCAACTTCTCGGGC
 AGCTACAACCCCAACGGCAACAGCTACCTCTCCGTGTACGGTGGTCCCGCAACCCCTGATCGAGTACTACATCGTCGA
 GAACTTGGACCTACAAACCGGTCCACGGGCGCCACCAAGCTGGCGAGGTACCTCCGACGGCAGCGTCTACGACATTT
 ACCGCACGCGGGGTCAACCAGCCGTCCATCGGACCGCCACCTTTTACCAGTACTGGTCCGTCCGCGCAACCCAC
 CGCTCGAGCGGGTCCGTCAACACGGGGAACCACTTCAACGGGTGGGCTCAGCAAGGCCTGACGCTCGGGACGATGGATTA
 CCAGATTGTTGCCGTGGAGGGTTACTTTAGCTCTGGCTCTCCATCACCGTCAAGCTAA (SEQ. ID NO: 2)

Figura 19

Proteína trichoderma reesei EGL III (endoglucanasa III)
 secuencia completa, incluyendo la secuencia de señal

MKFLQVLPAALAQTSQDQWATFTNGYTVSNLWLGASAGSGFGCVTAVSLSGGASWHADWQWSSGGQNNVKSQNSQ
 IAIQKRTVNSISSMPTTASWSYSGSNIRANVAYDLFTAANPNHVTYSGDYELMIWLKYGDIGPIGSSQGTVNVGGQSW
 TLYYGYNGAMQVYSFVAQTNTTNYSGDVKNFFNYLRDNKGYNAAGQYVLSYQFGTEPFTGSGTLNVAASWTASIN
 (SEQ. ID NO: 3)

Gen trichoderma reesei EG II (endoglucanasa III)

ADN desde que empieza el codón hasta que para el codón (incluye dos intrones)

ATGAAGTTCCTTCAAGTCCTCCCTGCCCCTCATACCGGCCCGCCCTGGCCCAAACAGCTGTGACCCAGTGGGCAACCTTCAC
 TGGCAACGGCTACACAGTCAGCAACAACCTTTGGGGAGCATCAGCCGGCTCTGGATTTGGCTGCGTGACGGCGGTATCGC
 TCAGCGGGGGCCCTCTGGCACGCGAGCTGCCAGTGGTCCGGCGGCCAGAACACGTCAGTCCGTACCAGAACTCTCAG
 ATTGCCATCCCCAGAAGAGGACCGTCAACAGCATCAGCAGCATGCCACCCTGCCAGCTGGAGCTACAGCGGGAGCAA
 CATCCGCGCTAATGTTGGGTATGACTTGTTCACCGCAGCCAACCCGAATCATGTCACGTAATCGGGAGACTACGAACTCA
 TGATCTGGTAAGCCATAAGAAGTGACCCTCCTTGATAGTTTCGACTAACAAACATGCTTGAGGCTTGGCAAATACGGCGGA
 TATTGGCCCGATTGGTCCCTCACAGGGAACAGTCAACGTCGGTGGCCAGAGCTGGACGCTCTACTATGGCTACAACGGGAG
 CCATGCAAGTCTATTCCCTTGTGGCCAGACCAACTACCAACTACAGGGAGATGTCAGAACTTCTTCAATTATCTC
 CGAGACATAAAGGATACAACGCTGCAGGGCCAATATGTTCTAGTAAGTCAACCTCAGTGGCTGAGTTTGTG
 CAACGTTTGCTAACAAACCTTCGTATAGGCTACCAATTTGGTACCGGCCCTTCACGGGCAGTGGAACTCTGAACGTCG
 CATCCTGGACCGCATCTCAACTAA (SEQ. ID NO: 4)

