

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 607 061**

51 Int. Cl.:

C12N 9/42 (2006.01)

C12P 19/02 (2006.01)

C12P 19/14 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **02.05.2012** **E 12166458 (5)**

97 Fecha y número de publicación de la concesión europea: **14.09.2016** **EP 2660319**

54 Título: **Endoglucanasas con propiedades mejoradas**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
29.03.2017

73 Titular/es:

**CLARIANT PRODUKTE (DEUTSCHLAND) GMBH
(100.0%)
Brüningstrasse 50
65929 Frankfurt am Main, DE**

72 Inventor/es:

**REISINGER, CHRISTOPH, DR.;
GEBHARD, GABI, DR.;
CLAREN, JÖRG, DR.;
MITROVIC, ALEKSANDRA;
UNTERSTRASSER, ISABEL y
FLICKER, KARLHEINZ**

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 607 061 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

Endoglucanasas con propiedades mejoradas

Campo de la invención y técnica anterior

La celulosa es un componente principal del material vegetal. Es la base de la integridad estructural de las plantas y frecuentemente se encuentra en una matriz de lignocelulosa compuesta de celulosa, hemicelulosas y lignina. Las aplicaciones que emplean celulosa se aprovechan de sus propiedades estructurales (fibras, productos textiles, papel, etc.) o de su naturaleza de hidrato de carbono, produciendo D-glucosa, celobiosa y/u oligómeros de celulosa.

Las lignocelulosas están fácilmente disponibles de la agricultura y silvicultura que incluyen corrientes de subproductos de cereales, maíz, caña de azúcar, remolacha, madera, etc. Las plantas que están optimizadas para su contenido y rendimiento de lignocelulosa ("cultivos de energía") contribuirán probablemente como un recurso importante en el futuro cercano.

Las celulasas comprenden una clase estructural y funcionalmente diversa de glucohidrolasas que actúan sobre la celulosa. Las celulasas se encuentran en bacterias, arqueas, hongos y plantas. Teniendo en común la actividad de escisión hidrolítica de los enlaces glucosídicos presentes en polímeros u oligómeros de celulosa, difieren en la especificidad de sustrato, modo de acción y parámetros de enzima que incluyen capacidad de procesamiento, pH y temperatura óptimos. La mayoría de las celulasas actúan sobre enlaces β -1,4 entre dos restos de glucosa. Sin embargo, otros enlaces encontrados en las lignocelulosas también pueden ser hidrolizados. Las celulasas pueden subdividirse por su modo de acción en endo- y exoenzimas. Las endoglucanasas introducen escisiones aleatorias en el polímero de celulosa, reduciendo así el grado de polimerización. Las exoenzimas, como las celobiohidrolasas, funcionan en un modo de acción sucesivo liberando celobiosa (D-glucosa- β -1,4-D-glucopiranosido) del extremo reductor o no reductor del polímero.

La base de datos CAZY [Cantarel BL, Coutinho PM, Rancurei C, Bernard T, Lombard V, Henrissat B (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics, *Nucleic Acids Res* 37:D233-238 PMID: 18838391] mantiene, entre otras cosas, una colección de glucohidrolasas conocidas que incluyen enzimas degradadoras de celulosa (es decir, celulasas). En esta base de datos las enzimas se clasifican en diferentes clases de GH según elementos estructurales. Varias clases de GH incluyen endoglucanasas, en particular las clases GH5, GH7, GH9, GH12, GH16, GH45, GH48, GH61 y GH74. A pesar de la alta diversidad dentro de algunas de las clases de GH, los miembros de una clase de GH frecuentemente tienen parámetros físicos y enzimáticos similares. Esto permite hacer afirmaciones generales como especificidad de sustrato, intervalo de pH, estabilidad o eficiencia catalítica para los miembros de una cierta clase de GH.

Los microorganismos degradadores de celulosa frecuentemente producen y secretan una mezcla compleja de celulasas. Por ejemplo, en el secretoma de *Trichoderma reesei* se han identificado 7 endoglucanasas que pertenecen a 6 clases diferentes de GH (Cel5A, Cel7B, Cel12A, Cel45A, Cel61A, Cel61B, Cel74A). Las diferentes endoglucanasas muestran un espectro de propiedades (Karlsson J, Siika-aho M, Tenkanen M, Tjerneld F. Enzymatic properties of the low molecular mass endoglucanases Cel12A (EG III) and Cel45A (EG V) of *Trichoderma reesei*. *J Biotechnol*. 2002 Oct 9;99(1):63-78. PubMed PMID: 12; Karlsson J, Momcilovic D, Wittgren B, Schulein M, Tjerneld F, Brinkmalm G. Enzymatic degradation of carboxymethyl cellulose hydrolyzed by the endoglucanases Cel5A, Cel7B, and Cel45A from *Humicola insolens* and Cel7B, Cel12A and Cel45Acore from *Trichoderma reesei*. *Biopolymers*. 2002 Jan;63(1):32-40. PubMed PMID:11754346). Las dos endoglucanasas predominantes, EGI (Cel7B, GH7) y EGI (Cel5A) se consideran las enzimas más activas de las mismas.

La actividad sinérgica de las enzimas celulolíticas permite la descomposición eficiente de sustratos complejos (B. Henrissat, H. Driguez, C. Viet & M. Schulein: Synergism of Cellulases from *Trichoderma reesei* in the Degradation of Cellulose; *Nature Biotechnology* 3, 722 - 726 (1985) doi:10.1038/nbt0885-722) y excluye el reemplazo de un componente de una clase estructural por una enzima de un segundo pliegue, cuando al mismo tiempo la eficiencia hidrolítica requiere mantenerse al máximo nivel (no equivalencia de diferentes EGs). No siempre es posible un simple reemplazo por otra enzima de clase GH. En términos generales, los miembros de endoglucanasas de la familia GH5 (que incluye EGs de bacterias termófilas) muestran termoestabilidad más alta en comparación con endoglucanasas de la familia GH7; sin embargo, la aplicación de una proteína termoestable de la familia GH7 frecuentemente es ventajosa para obtener altas tasas de hidrólisis.

Se informaron muchas aplicaciones de endoglucanasas como parte de mezclas enzimáticas complejas como actividades enzimáticas individuales. Las celulasas son importantes para hacer biocombustibles derivados de celulosa. Después del corte y opcionalmente pretratamiento químico y/o físico, las lignocelulosas se incuban con celulasas para liberar monómeros de azúcar que son adicionalmente procesados. Las condiciones de proceso necesitan ser adaptadas para optimizar velocidades de hidrólisis, rendimientos y/o estabilidad. Frecuentemente se prefieren temperaturas más altas en estos procesos, pero requieren enzimas más termoestables. Los procesos de sacarificación e hidrólisis simultáneos (SSF) requieren enzimas celulolíticas que son activas bajo condiciones fermentativas. El bioprocésamiento consolidado (CBP) requiere adicionalmente la combinación de propiedades de la enzima, con el fin de hacer en una sola etapa la producción, sacarificación y fermentación de la enzima.

Otras aplicaciones de endoglucanasas están dirigidas solamente a una hidrólisis o modificación parcial de las fibras de celulosa (modificación de fibra, biopulido, lavado a la piedra biológico, etc.). Por lo tanto, es necesario que las endoglucanasas usadas trabajen y/o sean estables a temperaturas elevadas, condiciones extremas de pH (por ejemplo, alcalino, ácido) y químicas (por ejemplo, lavandería, detergentes, proteasas, disolventes, etc.). El daño a la fibra debe ser minimizado para estas aplicaciones. Las endoglucanasas también pueden ayudar en la separación de fracciones no celulósicas del material de fibra en procesos de formación de pulpa (producción de pulpa y papel) o mejorar las propiedades reológicas de las corrientes de proceso. La estabilidad del detergente y la resistencia de la proteasa pueden considerarse un producto de la estabilidad elevada de la estructura de enzima, una propiedad que también está relacionada con una elevada estabilidad térmica. Las endoglucanasas también tienen aplicación en el procesamiento de alimentos y piensos (cervecías, producción de vino, recuperación de aceite de la torta de la prensa, horneado, preparación de masa). Frecuentemente la esterilización o pasteurización requiere temperaturas más altas. Para acortar los tiempos de procesamiento, puede ser ventajosa la estabilidad operacional de la endoglucanasa.

Las proteínas endoglucanasa I (Cel7B) derivadas de hongos del género *Trichoderma* (anamorfo *Hypocrea*) muestran un alto grado de identidad y se consideran mesófilas. Los miembros más estables de las endoglucanasas de la familia 7 de GH son las enzimas nativas de *Humicola insolens* (Cel7B) y *Fusarium oxysporum* (*eg1*) (documento US5912157). Según dicho informe, EGI no presenta actividad por encima de 60 °C. Así, existe la necesidad en el campo de la provisión de endoglucanasas más termoestables de la familia 7 de GH.

Se informó que algunas endoglucanasas pueden ser inactivadas térmicamente a temperaturas más altas (Dominguez JM, Acebal C, Jimenez J, de la Mata I, Macarron R, Castillon MP. Mechanisms of thermoinactivation of endoglucanase I from *Trichoderma reesei* QM 9414. *Biochem J.* 1992 Oct 15;287 (Pt 2):583-8). Los autores de dicho estudio también intentaron la reactivación de endoglucanasa termoinactivada, pero esto requiere condiciones agresivas que implican urea 8 M y agentes adicionales. Se mostraron efectos descritos como repliegue productivo sobre otras proteínas diferentes de las endoglucanasas [Zhang N. Suen WC, Windsor W, Xiao L, Madison V. Zaks A. Improving tolerance of *Candida antarctica* lipase B towards irreversible thermal inactivation through directed evolution. *Protein Eng.* 2003 Aug;16(8):599-605], pero a conocimiento de los inventores, no para endoglucanasas, en particular endoglucanasas de GH7. Se cree en la materia que las endoglucanasas termoinactivadas son de poco uso en la descomposición industrial de celulosa. Por otra parte, la elevada termoestabilidad frecuentemente es deseable para las endoglucanasas, en particular para enzimas de origen fúngico. Hasta la fecha solo se informaron algunas mejoras para las endoglucanasas de GH12 y GH45. Se han informado endoglucanasas termoestables de los pliegues estructurales de GH5 y GH48. Dichas endoglucanasas difieren sustancialmente con respecto a sus propiedades cinéticas y preferencia de sustrato de las endoglucanasas de la clase GH7.

En resumen, existe la necesidad de endoglucanasas procesables, particularmente de la familia GH7, con perfiles de temperatura superiores. Además sería deseable lograr buena productividad de su huésped de expresión. La necesidad es adicionalmente apoyada por el hecho de que muchos procesos de relevancia industrial operan bajo condiciones agresivas y a temperaturas elevadas. Un problema a resolver por la presente invención es la provisión de endoglucanasas mejoradas, particularmente de endoglucanasas con propiedades térmicas mejoradas. Problemas adicionales tratados y resueltos por la presente invención serán evidentes a partir de las siguientes secciones.

Sumario de la invención

La invención se refiere a proteínas (polipéptidos) de endoglucanasa termoestables. Las soluciones proporcionadas son:

1. Una proteína que tiene actividad de endoglucanasa que pertenece a la clase GH7 y que muestra termoestabilización activa (reivindicado por la invención hasta el punto que está cubierto por las reivindicaciones).

2. Una proteína que tiene actividad de endoglucanasa que comprende una secuencia de aminoácidos que tiene al menos el 96 %, preferentemente al menos el 97 %, más preferentemente al menos el 98 %, incluso más preferentemente al menos el 99 %, tal como al menos el 99,5 % de identidad con SEQ ID NO: 2, en las que las proteínas de endoglucanasa de la invención muestran al menos el 90 % de actividad de conversión del sustrato residual a 60 °C cuando la incubación se hace durante una hora.

Un aspecto adicional de la invención son ácidos nucleicos que codifican dichos polipéptidos, y construcciones de expresión que comprenden estos polinucleótidos en un esqueleto de vector contenido en un organismo. Otro aspecto relacionado con la invención es la aplicación de las proteínas de la invención para el procesamiento de lignocelulosa y materiales de celulosa. En particular, la sacarificación de materia prima de lignocelulosa en procesos consolidados, parcialmente consolidados o no consolidados, o en el procesamiento de alimento, pienso, fibra de celulosa, o aplicaciones de limpieza.

La invención también se refiere a organismos de producción/expresión para la producción de las proteínas de la invención, y a procedimientos para el cultivo de tales organismos con el fin de la producción de proteínas. Los organismos se seleccionan de organismos que incluyen microorganismos (hongos, bacterias o arqueas) o plantas.

Breve descripción de las figuras

5 Figura 1: El gel de SDS muestra la expresión de una proteína de SEQ ID NO: 8 - una variante de SEQ ID NO: 2 - secretada en el sobrenadante. La banda de la proteína expresada es visible entre 75 y 100 kDa.

10 Figura 2: Plásmido de expresión de *Trichoderma reesei*. La secuencia de ADN que codifica el gen de endoglucanasa madura se clona en fusión con la secuencia del péptido señal *TrCBHI* bajo el control del promotor de *TrCBHI*. El casete de expresión escindible *Swal/SbfI* contiene un casete de resistencia a higromicina para la selección de transformantes.

Figura 3: Variantes de endoglucanasa que muestran una elevada estabilidad a la temperatura ([6] y [3]) y variantes con elevada estabilidad a la temperatura y termoestabilización activa [1], [4], [5], [7], [8], [9] y [10] en comparación con una proteína GH7 nativa [2].

15 Figura 4: Calibración de porciones de 200 μ l de disolución de 4-metilumbeliferona alcalina para la lectura de fluorescencia en un lector de placas Tecan Infinite M200. Se preparó una disolución 10 mM disolviendo 440 mg de 4-metilumbeliferona (Sigma Aldrich Cat. N.º 69580) en 250 ml de disolución 0,5 M de carbonato de sodio. Se prepararon diluciones sucesivas en carbonato sódico 0,5 M. La intensidad de fluorescencia se midió a 360 nm/454 nm con una ganancia de 50.

Figura 5: Estabilización térmica de SEQ ID NO: 13 en comparación con SEQ ID NO: 4.

20 Figura 6: Determinación de las semividas a 70 °C (ejemplo 7) para SEQ ID NO: 14 en comparación con SEQ ID NO: 4.

Definiciones

“**Termoestabilidad**” es un término usado para describir una propiedad intrínseca de una proteína particular con actividad de endoglucanasa según la presente invención.

25 “**Termoestabilización activa**” es un término usado para describir una propiedad intrínseca de una proteína particular con actividad de endoglucanasa según la presente invención.

Determinación de la termoestabilidad y/o termoestabilización activa: Se determinan la termoestabilidad y la termoestabilización activa del siguiente modo.

30 1.) La enzima se expresa en *Pichia pastoris* como se describe en Ejemplo 2. Opcionalmente se purifica la enzima.

2.) Ajuste de la concentración de enzima.

35 Se prepara una disolución de la enzima de una concentración apropiada diluyendo la enzima purificada o sobrenadante de cultivo de *Pichia pastoris* en tampón acetato sódico (50 mM, pH 5) a una concentración de trabajo aplicable. Para la determinación de la concentración de trabajo aplicable, se prepara una dilución sucesiva de la enzima obtenida en la etapa 1) anterior en el tampón acetato sódico y se prueban alícuotas de 10 μ l en el gradiente de temperatura como se describe en el Ejemplo 4. Una concentración de trabajo aplicable se define como una concentración que produce una señal de fluorescencia entre 5.000 y 15.000 en un lector de placas Tecan Infinite M200 a una ganancia de 50, o una concentración equivalente de 4-metilumbeliferona 5,4 μ M a 19 μ M después de la incubación como se describe en el Ejemplo 4.

40 3.) Determinación de la capacidad de conversión de sustrato como se describe en el Ejemplo 4, con la excepción de que la alícuota de 10 μ l del sobrenadante de cultivo se reemplaza por la alícuota de 10 μ l de la disolución de enzima a la concentración de trabajo aplicable como se define en la etapa 2).

45 4.) Normalización de la medición por división de todas las lecturas de unidades relativas de fluorescencia (urf) entre la máxima lectura de urf dentro del gradiente de temperatura para obtener una conversión relativa de sustrato para cada proteína probada a cada temperatura probada.

5.) Representación de la concentración relativa de sustrato frente a las temperaturas de reacción probadas.

6.) Determinación de la estabilidad a la temperatura como se describe en (a), o determinación de la termoestabilización activa como se describe en (b), del siguiente modo.

a. **Determinación de la estabilidad a la temperatura:** Una proteína se caracteriza como estable a la temperatura si la conversión relativa de sustrato a 60 °C es 0,5 o más, preferentemente 0,7 o más y más preferentemente 0,9 o más, tal como 0,95 o más.

5 **Determinación de la termoestabilización activa:** Análisis de la gráfica obtenida en la etapa 5) para la presencia de una meseta a una conversión relativa de sustrato que es menor que el nivel máximo (que es 1), pero que es al menos tan alta como 0,15.

Una meseta se define como un nivel de la conversión relativa de sustrato que permanece esencialmente sin cambio dentro de un intervalo de temperatura de al menos 5 °C, preferentemente de 70 a 75 °C (es decir, dentro de +/- 0,1 alrededor del valor promedio dentro de dicho intervalo de temperatura).

10 **b.** Las variantes que no muestran termoestabilización activa tienen una conversión relativa de sustrato entre 0 y menos de 0,15, normalmente alrededor de 0,1. Sin desear ceñirse a teoría particular alguna, se cree que la conversión relativa de sustrato medida de normalmente aproximadamente 0,1 (en lugar de 0,0, como era de esperar para una enzima inactiva a una temperatura dada), se debe a las rampas de temperatura finita en el ciclador térmico y/o durante la manipulación de las mezclas de muestra.

15 **Propiedades térmicas** es un término usado generalmente para referirse a las propiedades de una enzima a temperaturas más altas (por ejemplo, 60 °C o más). El término puede incluir uno o ambos de "estabilidad a la temperatura", como se ha definido anteriormente, y "termoestabilización activa", como se ha descrito anteriormente.

20 **Actividad de endoglucanasa** en el contexto de esta invención se define como la aceleración catalítica de la rotura de enlaces glucosídicos β -1,4 mediante ataque nucleófilo por una molécula polar como agua o moléculas orgánicas con sus funciones hidroxilo o mercapto o amino, por una proteína. La definición también incluye la escisión de moléculas sintéticas que tienen una molécula no de hidrato de carbono unida a glucosa, celobiosa o lactosa, por medio del enlace glucosídico β -1,4. Reacciones de ejemplo catalizadas por endoglucanasas se enumeran por Brenda Database (<http://www.brenda-enzymes.info> (Edición 01.2012 (enero de 2012)); Enzyme data and metabolic information: BRENDA, a resource for research in biology, biochemistry, and medicine Schomburg, I., Hofmann, O., Baensch, C., Chang, A., Schomburg, D. Gene Funct. Dis. 3-4, 109-18 (2000))

25 **Actividad residual** se define como la actividad enzimática que se recupera después de la incubación de la enzima durante un tiempo definido a una temperatura definida (elevada) en comparación con la actividad sin la etapa de incubación. En el Ejemplo 4 se da el protocolo para la determinación de la actividad residual.

30 **Alineamiento de secuencias con SEQ ID NO: 2:** El alineamiento por pares de cualquier segunda secuencia de endoglucanasa GH7 con la secuencia de origen (SEQ ID NO: 2) se hace usando el algoritmo ClustalW (Larkin M.A., Blackshields G., Brown N.P., Chenna R., McGettigan P.A., McWilliam H., Valentin F., Wallace I.M., Wilm A., Lopez R., Thompson J.D., Gibson T.J. y Higgins D.G. (2007), ClustalW and ClustalX versión 2. Bioinformatics 2007 23(21): 2947-2948). El alineamiento por pares mostrará los números de posición para SEQ ID NO: 2. Dichos números pueden usarse para referencia, por ejemplo, cuando se dice que, por ejemplo, el resto correspondiente a la posición N.º 2 de SEQ ID NO: 2 está mutado en la segunda endoglucanasa de GH7. Como convención para la numeración de aminoácidos y designación de variantes de proteína para la descripción de variantes de proteína, el aminoácido dentro de la secuencia de proteína de origen SEQ ID NO: 2 se denomina el número de posición 1 o S1 o serina 1. La numeración de todos los aminoácidos será según su posición en la secuencia de origen dada en SEQ ID NO: 2 con respecto a este número de posición 1.

40 **Identidad de secuencia:** Para la determinación de la identidad de secuencia se usa el software AlignX de VectorNT Package comercializado por Life Technology Corporation, usando los ajustes estándares (Penalización por abertura de hueco 10, Penalización por extensión de hueco 0,1).

45 Las **variantes de proteína** son polipéptidos cuya secuencia de aminoácidos difiere en una o más posiciones de su proteína de origen, por lo que las diferencias podrían ser sustituciones de uno o más resto(s) de aminoácido con otros, deleciones de uno o varios resto(s) de aminoácido, o inserción de resto(s) de aminoácido adicionales o estiramientos de resto(s) de de aminoácido en la secuencia de origen. Las proteínas pueden modificarse en posiciones definidas por introducción de mutaciones puntuales en los ácidos nucleicos codificantes. El término secuencia de proteína modificada en el presente documento siempre se refiere a proteínas que resultan de la transcripción y traducción, además de la modificación postraduccional opcional y procesos de translocación de los ácidos nucleicos modificados correspondientemente, tanto *in vitro* como por un huésped de expresión adecuado. Los métodos para la generación de tales variantes de proteína son muy conocidos en la técnica y así no están limitados, ejemplos incluyen mutagénesis al azar o dirigida al sitio, mutagénesis de saturación de sitio, ensamblaje de fragmentos basado en PCR, barajado de ADN, recombinación homologa *in vitro* o *in vivo*, y métodos de síntesis de genes basados en síntesis química de ADN.

55 La **nomenclatura de aminoácidos**, péptidos, nucleótidos y ácidos nucleicos se hace según la IUPAC. Generalmente, los aminoácidos se nombran en este documento según el código de una letra.

Los **intercambios** de aminoácidos individuales se describen nombrando el código de una sola letra del aminoácido original, seguido de su número de posición y el código de una sola letra del aminoácido de reemplazo, es decir, el cambio de glutamina en la posición uno a una leucina en esta posición se describe como "Q1L". Para delecciones de posiciones individuales de la secuencia, el símbolo del aminoácido de reemplazo se sustituye por la abreviatura de tres letras "del", así, la delección de alanina en la posición 3 se denominaría "A3del". Los aminoácidos adicionales insertados reciben el número de la posición precedente extendida por una letra minúscula en orden alfabético con respecto a su distancia a su punto de inserción. Así, la inserción de dos triptófanos después de la posición 3 se denomina "3aW, 3bW". La introducción de codones no traducidos TAA, TGA y TAG en la secuencia de ácido nucleico se indica "*" en la secuencia de aminoácidos, así la introducción de un codón de terminación en la posición 4 de la secuencia de aminoácidos se denomina "G4". Las mutaciones múltiples se separan por un signo más o una barra diagonal o una coma. Por ejemplo, dos mutaciones en las posiciones 20 y 21 que sustituyen glicina y serina con alanina y ácido glutámico, respectivamente, se indican "A20G+E21S" o "A20G/E21S", "A20G, E21S". Cuando un resto de aminoácido en una posición dada se sustituye con dos o más restos de aminoácido alternativos, estos restos se separan por una coma o una barra diagonal. Por ejemplo, la sustitución de alanina en la posición 30 con tanto glicina como ácido glutámico se indica como "A20G,E" o "A20G/E" o "A20G, A20E". Cuando en el presente documento se identifica una posición adecuada para modificación sin ninguna modificación específica que se sugiera, debe entenderse que cualquier resto de aminoácido puede ser sustituido con el residuo de aminoácido presente en la posición. Así, por ejemplo, cuando se menciona una modificación de una alanina en la posición 20, pero no se especifica, debe entenderse que la alanina puede ser delecionada o sustituida por cualquier otro resto de aminoácido (es decir, uno cualquiera de R, N, D, C, Q, E, G, H, I, L, K, M, F, P, S, T, W, Y y V).

Los términos "**mutación similar**" o "**sustitución similar**" se refieren a una mutación de aminoácido en la que un resto de aminoácido en una primera mutación (con respecto a la secuencia de origen, tal como por ejemplo SEQ ID NO: 2) es reemplazado nuevamente por una segunda mutación, y con lo cual el resto de aminoácido producido por la segunda mutación tiene propiedades similares al resto de aminoácido que había sido producido por la primera mutación. En este contexto, similar significa un aminoácido que tiene propiedades químicas similares. Si, por ejemplo, una primera mutación en una posición específica produce a una sustitución de un resto de aminoácido no alifático (por ejemplo, Ser) con un resto de aminoácido alifático (por ejemplo, Leu), entonces una sustitución en la misma posición con un aminoácido alifático diferente por medio de una segunda mutación (por ejemplo, Ile o Val) se denomina una mutación similar. Propiedades químicas adicionales incluyen tamaño del resto hidrofobia, polaridad, carga, valor de pK, y similares. Así, una mutación similar puede incluir sustitución tal como básico por básico, ácido por ácido, polar por polar, etc. Es probable que los conjuntos de aminoácidos así derivados estén conservados por razones estructurales. Estos conjuntos pueden describirse en forma de un diagrama de Venn (Livingstone CD. y Barton GJ. (1993) "Protein sequence 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). Pueden hacerse sustituciones similares, por ejemplo, según la siguiente agrupación de aminoácidos: Hidrófobo: F W Y H K M I L V A G; Aromáticos: F W Y H; Alifáticos: I L V; Polares: W Y H K R E D C S T N; Cargados H K R E D; Cargados positivamente: H K R; Cargados negativamente: E D.

Una **construcción de expresión** en el presente documento se define como una secuencia de ADN que comprende todos los elementos de secuencia requeridos para establecer la expresión de un marco de lectura abierto (ORF) comprendido en la célula huésped, que incluye secuencias para la iniciación de la transcripción (promotores), terminación y regulación, sitios para la iniciación de la traducción, regiones para la replicación estable o integración en el genoma del huésped y un marcador genético de selección. El marco de lectura abierto consiste opcionalmente en una fusión de un ácido nucleico que codifica la proteína diana con elementos adicionales, especialmente señales de secreción, un dominio de unión de celulosa, TAGs para potenciar el nivel de expresión o para facilitar la purificación o aislamiento del caldo de fermentación. Así, la configuración funcional puede estar ya establecida o alcanzarse por medio de evento de disposición (integración, etc.) en la célula huésped. En una realización preferida, la construcción de expresión contiene un promotor funcionalmente unido al marco de lectura abierto, seguido de una secuencia de terminación opcional. Promotores preferidos son promotores de fuerza media a alta, funcionales en los huéspedes seleccionados bajo las condiciones de fermentación. Para una ilustración, ejemplos de promotores preferidos se dan del siguiente modo:

- Bacterias (por ejemplo, *Escherichia coli*): lac, tac, trp, tet, T3 T7, CP7, CP21, araBAD
- Levadura (por ejemplo, *Pichia*, *Saccharomyces*): AOXI, AOXII, FMDH, GAP, TEF, PFK1, FBA1, PGK1, ADH1, ADH2, TDH3
- Hongos (por ejemplo, *Trichoderma*): CBHI, CBHII, EGI, PGK, BGL, XYL1, XYL2

Ejemplos adicionales de promotores adecuados para expresión heteróloga se informan en la bibliografía. Otras partes de la construcción de expresión son elementos genéticos requeridos para una herencia estable de los ácidos nucleicos introducidos, y marcadores de selección que incluyen elementos genéticos que se refieren a resistencia a antibiótico o que complementan las auxotrofías definidas de la cepa huésped.

La secuencia de todos los ácidos nucleicos de la invención, o de los ácidos nucleicos que codifican los polipéptidos/proteínas de la invención, puede ajustarse hacia un uso de codón óptimo en el huésped de expresión

seleccionado. Los ácidos nucleicos que tienen tal uso de codón optimizado/óptimo para el huésped de expresión particular también son parte de la presente invención. Un **huésped** de producción se usa en el presente documento como sinónimo de huésped de expresión, y significa un organismo que, tras el cultivo, produce la proteína de la presente invención. En una realización, la proteína de la presente invención no es secretada por el huésped de producción; sin embargo, en una realización preferida, es secretada al medio circundante. Tales organismos se seleccionan preferentemente del reino de las bacterias, arqueas, levaduras, hongos y/o plantas. Un huésped de expresión preferido es *Pichia pastoris*.

“**Bacterias**” debe referirse en el presente documento a organismos procariotas. En una realización preferida, las bacterias son eubacterias, e incluso más preferentemente se seleccionan de entre los géneros *Escherichia*, *Bacillus*, *Klebsiella*, *Streptomyces*, *Lactococcus* y *Lactobacillus*, en particular *Escherichia coli*, *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus amyloliquefaciens*, *Bacillus megaterium*, *Klebsiella planticola*, *Streptomyces lividans*, *Lactococcus lactis*, *Lactobacillus brevis*.

“**Levadura**” debe referirse en el presente documento a organismos eucariotas inferiores que muestran un estado vegetativo unicelular en su ciclo de vida. Éstas incluyen especialmente organismos de la clase *Saccharomycetes*, en particular del género *Saccharomyces*, *Pachysolen*, *Pichia*, *Candida*, *Yarrowina*, *Debaromyces*, *Klyveromyces*, *Zygosaccharomyces*.

“**Hongos filamentosos**” u “**hongos**” debe referirse en el presente documento a todos los organismos eucariotas inferiores que muestran un crecimiento de hifa durante al menos un estado en su ciclo de vida. Éstos incluyen especialmente organismos del filo Ascomycota y Basidiomycota, en particular de los géneros *Trichoderma*, *Talaromyces*, *Aspergillus*, *Penicillium*, *Chrysosporium*, *Phanerochaete*, *Thermoascus*, *Agaricus*, *Pleutrus*, *Irpex*.

“**Planta**” debe referirse en el presente documento a todos los organismos eucariotas que pertenecen al reino de las plantas. En una realización preferida, el huésped de expresión se selecciona de plantas del género *Zea*, *Triticum*, *Hordeum*, *Secale*, *Miscanthus*, *Saccharum*, *Solanum*, *Ipomea*, *Manihot*, *Helianthus*, *Camellia*, *Aspalathus*, *Eucalyptus*, *Beta*, *Fagus*, miembros de la familia de *Pinaceae*, *Betulaceae*, *Malvaceae*, *Cupressaceae*, *Rosaceae*, *Arecaceae*.

Por **formulación de enzima** se indica cualquier composición líquida o sólida que contiene la enzima como una fracción. Componentes adicionales comprenden preferentemente agua, polioles, azúcares, detergentes, agentes de tamponamiento, agentes reductores, sales inorgánicas, vehículos sólidos, agentes conservantes, especialmente con actividad antibacteriana o antifúngica, colorantes, fragancias y/o perfumes.

Usos de las endoglucanasas, tales como particularmente de la endoglucanasa de la presente invención (ejemplos no limitantes): hidrólisis de materias primas de lignocelulosa para la generación de azúcares monoméricos, diméricos u oligoméricos; producción de pulpa y papel; aplicaciones textiles para la mejora o el procesamiento general de fibras, hilos o tela vaquera; aplicaciones de limpieza para aplicaciones de uso industrial o doméstico; liberación de nutrientes, aumento del rendimiento de producción o mejora de las propiedades de la masa en el campo de los alimentos y piensos.

Descripción detallada de la invención

La invención se refiere a endoglucanasas GH7 con propiedades superiores, en tanto que estén cubiertas por las reivindicaciones. Más particularmente, la invención se refiere a proteínas (polipéptidos) de endoglucanasa termoestables, en tanto que estén cubiertas por las reivindicaciones. Las soluciones proporcionadas y reivindicadas, en tanto que estén cubiertas por las reivindicaciones, son:

1. Una proteína que tiene actividad de endoglucanasa que pertenece a la clase GH7 y que muestra termoestabilización activa.

2. Una proteína que tiene actividad de endoglucanasa que comprende una secuencia de aminoácidos que tiene al menos el 96 %, preferentemente al menos el 97 %, más preferentemente al menos el 98 %, incluso más preferentemente al menos el 99 %, tal como al menos el 99,5 % de identidad con SEQ. ID NO: 2, en la que la proteína muestra al menos el 90 % de capacidad de conversión de sustrato residual a la temperatura 60 °C cuando la incubación se hace durante una hora.

Estas dos realizaciones se describen en detalle a continuación.

La estabilidad a la temperatura se define anteriormente. En el Ejemplo 4 se da un ejemplo de la determinación de la termoestabilidad. Las endoglucanasas de la clase GH7 se enumeran en la Tabla 1 (EC 3.2.1.4). A menos que se excluya por restricciones particulares de identidad de secuencia en una reivindicación particular, la invención se refiere a variantes de todas las endoglucanasas de la clase GH7, comprendidas las variantes de las que se muestran en la Tabla 1.

Tabla 1: Endoglucanasas conocidas de la clase GH7

	Nombre de proteína	Organismo	GenBank	PDB/3D
1	celulasa III-A (fragmento de péptido)	<i>Acremonium cellulolyticus</i>		
2	endo- β -1,4-glucanasa (EglB;AN3418.2)	<i>Aspergillus nidulans</i> FGSC A4	EAA63386.1	
3	endo- β -1,4-glucanasa (CelB)	<i>Aspergillus oryzae</i> KBN616	BAA22589.1	
4	endo- β -1,4-glucanasa (CelB;A0090010000314)	<i>Aspergillus oryzae</i> RIB40	AEB00821.1	
5	endo- β -1,4-glucanasa I	<i>Aspergillus terreus</i> MS-31	ADR78837.1	
6	endo- β -1,3-1,4-glucanasa (Bgl7A)	<i>Bispora</i> sp. MEY-1 / CGMCC 2500	ACT53749.1	
7	EG I (fragmento de péptido) (Cel7C)	<i>Chrysosporium lucknowense</i>		
8	endo- β -1,4-glucanasa (CLhgEG1)	Simbionte de <i>Coptotermes lacteus</i> WH2002	BAC07551.1	
9	endo- β -1,4-glucanasa (CLhgEG2)	Simbionte de <i>Coptotermes lacteus</i> WH2002	BAC07552.1	
10	endo- β -1,4-glucanasa (EglB)	<i>Emericella nidulans</i>	AAM54071.1	
11	endo- β -1,4-glucanasa I (EG I;Eg1;Fof7) (Cel7B)	<i>Fusarium oxysporum</i>	AAA65586.1	1OWW[A,B,C,D]
12	endoglucanasa 3 (HmEG3) (fragmento)	<i>Holomastigotoides mirabile</i>	BAB64565.1	
13	endoglucanasa 2 (HmEG2)	<i>Holomastigotoides mirabile</i>	BAB64564.1	
14	endoglucanasa 1 (HmEG1)	<i>Holomastigotoides mirabile</i>	BAB64563.1	
15	endo- β -1,4-glucanasa I (Egl1;EG-I)	<i>Humicola grisea</i> var. <i>thermoidea</i>	BAA09786.1	
16	endoglucanasa I (EG I;EG1) (Cel7B)	<i>Humicola insolens</i>	AAE25068.1	1 A39[A]
17	endo- β -1,4-glucanasa I (EG1;Egl1;EG-I) (Cel7B)	<i>Hypocrea jecorina</i>	AAA34212.1	1EG1[A ,C]
18	endoglucanasa I (Egl)	<i>Hypocrea jecorina</i> M5	ADM08177.1	
19	endo- β -1,4-glucanasa (Egl1)	<i>Hypocrea jecorina</i> PTCC 5142	AAX28897.1	
20	endoglucanasa I	<i>Hypocrea pseudokoningii</i>	ABM90986.1	
21	endoglucanasa I (Eg1)	<i>Hypocrea pseudokoningii</i> 3.3002	AEQ29501.1	
22	endoglucanasa I (Eg1)	<i>Hypocrea rufa</i>	AEO17039.1	
23	endo- β -1,4-glucanasa I (EG1;BglI)	<i>Hypocrea rufa</i> AS 3.3711	AAQ21382.1	
24	endoglucanasa I	<i>Hypocrea rufa</i> HK-75		
25	endo- β -1,4-glucanasa (Egl1 ;MG02532.4)	<i>Magnaporthe grisea</i> 70-15	XP_366456.1	
26	endoglucanasa	<i>Myceliophthora thermophila</i> CBS 117.65	AAE25067.1	
27	endoglucanasa I (Egl1) (Cel7B)	<i>Penicillium decumbens</i> 114-2	ABY56790.1	
28	endoglucanasa I (Egl1)	<i>Penicillium decumbens</i> L-06	ACJ 15337.1	
29	endoglucanasa I (Egl1;Eg1)	<i>Penicillium oxalicum</i>	ACS32299.1	

	Nombre de proteína	Organismo	GenBank	PDB/3D
30	endoglucanasa (Cel7B)	<i>Penicillium purpurogenum</i>	AEL78899.1	
31	endoglucanasa (Bgl7C7)	<i>Penicillium sp. C7</i>	AEG74551.1	
32	endo- β -1,4-glucanasa (Egl) (fragmento de péptidos) (Cel7B)	<i>Penicillium verruculosum</i>		
33	endoglucanasa 3 (PgEG3)	<i>Pseudotriconympha grassii</i>	BAB64562.1	
34	endoglucanasa 2 (PgEG2)	<i>Pseudotriconympha grassii</i>	BAB64561.1	
35	endoglucanasa 1 (PgEG1h)	<i>Pseudotriconympha grassii</i>	BAB64553.1	
36	Egl1 (fragmento)	<i>Trichoderma asperellum T203</i>	AAS37698.1	
37	endoglucanasa I	<i>Trichoderma longibrachiatum 3.1029</i>	AEI71804.1	
38	endoglucanasa I (Egl1)	<i>Trichoderma longibrachiatum 36MS</i>	AEC03714.1	
39	endo- β -1,4-glucanasa I (Egl1;EglI;TICel7A) (Cel7A)	<i>Trichoderma longibrachiatum CECT 2606</i>	1920181A	
40	endoglucanasa I	<i>Trichoderma longibrachiatum FU05</i>	ACZ34302.1	
41	endoglucanasa I (Egl;EglI)	<i>Trichoderma sp. SSL</i>	ACH68455.1	
42	endo- β -1,4-glucanasa (RsSymEG1;SM2038B11)	protista simbiótico sin cultivar de <i>Reticulitermes speratus</i>	BAF57296.1	

A continuación se describirá en detalle el primer y segundo aspecto.

Primer aspecto: Una proteína que tiene actividad de endoglucanasa que pertenece a la clase GH7 y que muestra termoestabilización activa.

5 En el primer aspecto de la invención, las proteínas tienen actividad de endoglucanasa y propiedades térmicas superiores. Las propiedades térmicas superiores se definen como una estabilidad a la temperatura que se manifiesta en una actividad de conversión de sustrato relativa superior al 90 % (tal como superior al 95 %) tras la incubación a temperaturas de 60 °C o más altas, y termoestabilización activa. La termoestabilización activa se describe a continuación.

10 Los inventores de la presente invención han descubierto sorprendentemente que las proteínas que muestran termoestabilización activa también muestran estabilidad a la temperatura.

Esto se mostró por el siguiente ejemplo. Los inventores generaron una endoglucanasa GH7, que es una variante particular de SEQ ID NO: 4 (es decir, la dada por SEQ ID NO: 2) del siguiente modo. Se obtuvo un ácido nucleico que codifica un polipéptido con SEQ ID NO: 2 por mutagénesis al azar (PCR propensa a error, como se describe en el Ejemplo 1). Los métodos de mutagénesis al azar son muy conocidos en la técnica. Además, ahora que los inventores han desvelado aquí la conveniencia de un polipéptido codificado por SEQ ID NO: 2, el experto puede preparar directamente un ácido nucleico que codifica esta proteína. Los métodos para esto incluyen, por ejemplo, síntesis de genes o mutagénesis dirigida al sitio, a partir de un ácido nucleico con un alto grado de identidad de secuencia (por ejemplo, superior al 90 %) con SEQ ID NO: 4, e introducción de mutaciones por mutaciones dirigida al sitio (en uno o varias etapas) para obtener el ácido nucleico que codifica la proteína de SEQ ID NO: 2. Una secuencia inicial de la que puede obtenerse el ácido nucleico que codifica SEQ ID NO: 2 por mutagénesis es Cel7B de *Hypocrea pseudokoningii*, dada aquí como SEQ ID NO. 4 (número de acceso de Gene Bank ABM90986).

Los inventores de la presente invención caracterizaron la termoestabilidad de la proteína que tiene SEQ ID NO: 2. Como puede apreciarse en la Figura 3, esta proteína resuelve el problema técnico que subyace a la presente invención, es decir, tiene termoestabilidad más alta que su proteína de origen (SEQ ID NO: 4). Esto es evidente, por ejemplo, del hecho de que la conversión relativa de sustrato está casi próxima a su máximo, por ejemplo, a 60 °C, mientras que la conversión relativa de sustrato de la proteína que tiene SEQ ID NO: 4 está a un nivel muy bajo a dicha temperatura (véase la Figura 3).

30 Sorprendentemente, los inventores han encontrado que incluso a temperaturas más altas, por ejemplo en el intervalo de 68 a 76 °C (incluyendo 70 a 74 °C), la conversión relativa de sustrato no desciende significativamente al aumentar

la temperatura. Esto contrasta fuertemente con las propiedades de la proteína de origen que tiene SEQ ID NO: 4 que, en una gráfica frente a temperaturas crecientes, muestra una disminución de la conversión relativa de sustrato, llegando la disminución hasta niveles de fondo sin meseta intermedia. Se cree que la proteína que tiene SEQ ID NO: 4, cuando se expone a temperaturas más altas, por ejemplo 60 °C o más, tal como 70 °C o más, no está presente en su estado activo. Sin desear ceñirse a teoría particular alguna, se cree que este efecto es debido al desdoblamiento térmico (o plegamiento de conformaciones inactivas) de la proteína. Sin desear ceñirse a teoría particular alguna, el efecto de la pérdida de actividad a temperaturas altas se denominará a continuación desdoblamiento térmico. El desdoblamiento térmico es un fenómeno muy conocido de las proteínas de casi cualquier tipo, particularmente enzimas, a temperaturas más altas. Así, el desdoblamiento térmico observado para la proteína que tiene SEQ ID NO: 4 está en línea con las expectativas de un experto. La proteína de SEQ ID NO: 4 no es parte de la invención.

En marcado contraste, la proteína de este aspecto de la invención muestra una fase de meseta a temperaturas más altas, por ejemplo, en el intervalo de 68 a 76 °C (incluyendo 70 a 74 °C). Esta meseta es más baja que la conversión relativa de sustrato máxima, pero más alta que la conversión relativa de sustrato de fondo. Sin desear ceñirse a teoría particular alguna, los inventores de la presente invención llegan a la conclusión de que la proteína de la invención está presente a estas temperaturas más altas en un estado que es diferente del estado plegado a temperaturas más bajas (por ejemplo, 46 °C), pero todavía esta proteína es enzimáticamente activa. Así, puede asumirse que a temperaturas más altas esta proteína de la invención se vuelve a plegar activamente, es decir, se vuelve a plegar para obtener un estado activo adicional (y así permitir la conversión relativa de sustrato observada a temperaturas más altas). Por lo tanto, los inventores han llamado a esta propiedad, que también se define arriba en la sección de definiciones, "termoestabilización activa".

La proteína de la invención resuelve así el problema técnico que subyace a la presente invención al ser estable a la temperatura. Además, basándose en la divulgación de la presente invención, se da al experto en la técnica una guía para la identificación de proteínas adicionales según este primer aspecto de la invención.

Tales proteínas adicionales pueden encontrarse del siguiente modo. En primer lugar, puede introducirse cualquier tipo de mutaciones (incluyendo delección, inserción o sustitución de uno o varios restos de aminoácido, y siendo aleatoria o dirigida) en cualquier endoglucanasa de la familia GH7, particularmente en una cualquiera de las indicadas en la Tabla 1, para obtener una proteína mutante, o una biblioteca de las mismas. La proteína mutante así obtenida o la biblioteca de las mismas puede cribarse para termoestabilización activa como se ha definido anteriormente. Las proteínas conocidas dadas en la Tabla 1 anterior no son parte de la invención, pero cualquier mutante de las mismas que muestre termoestabilización activa está incluido en la invención.

Y, lo que es más importante, todas las enzimas del primer aspecto de la presente invención, es decir, las que muestran termoestabilización activa como se ha definido anteriormente, muestran estabilidad a la temperatura como se ha definido anteriormente. Así, la termoestabilización activa en un primer aspecto es una solución al problema que subyace a la presente invención. Si una proteína dada está o no dentro del primer aspecto de la invención puede probarse fiablemente por el ensayo de termoestabilización activa dado anteriormente.

Segundo aspecto: Una proteína que tiene actividad de endoglucanasa que comprende una secuencia de aminoácidos que tiene al menos el 96 %, preferentemente al menos el 97 %, más preferentemente al menos el 98 %, incluso más preferentemente al menos el 99 %, tal como al menos el 99,5 % de identidad con SEQ ID NO: 2 en la que la proteína muestra al menos el 90 % de actividad de conversión de sustrato residual a la temperatura 60 °C cuando la incubación se hace durante una hora.

Al buscar una segunda solución al problema que subyace a la presente invención, los inventores emprendieron un proyecto de mutagénesis a partir de la proteína de SEQ ID NO: 2. Así, los inventores han introducido mutaciones, tales como mutaciones puntuales, en la proteína que tiene la SEQ ID NO: 2 (es decir, modificando el ácido nucleico subyacente, como se describe más abajo). Los inventores han encontrado que muchos de tales mutantes también muestran estabilidad a la temperatura y así resuelven el problema subyacente en un segundo aspecto. En la Figura 3 se dan ejemplos de las soluciones.

Así, *en un segundo aspecto*, la invención se refiere a una proteína que tiene actividad de endoglucanasa que comprende una secuencia de aminoácidos que tiene al menos el 96 %, preferentemente al menos el 97 %, más preferentemente al menos el 98 %, incluso más preferentemente al menos el 99 %, tal como al menos el 99,5 % de identidad con SEQ ID NO: 2 en la que la proteína muestra al menos el 90 % de actividad de conversión de sustrato residual a la temperatura 60 °C cuando la incubación se hace durante una hora. Esta proteína puede normalmente pertenecer a la clase GH7.

Particularmente, la presente invención también proporciona mutantes específicos de la proteína con la secuencia de SEQ ID NO: 2. Así, la secuencia dada en SEQ ID NO: 2 se modifica en una o más posiciones (preferentemente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18 o 19) posiciones. Tal modificación puede consistir en sustitución, delección, inserción y similares. En una realización particular de la misma, la modificación consiste en una sustitución. En una realización incluso más específica, la modificación consiste en una sustitución de una cualquiera o más de las posiciones específicas de SEQ ID NO: 2 que están individualizadas en la columna más a la izquierda de las Tablas 2, 3, 4.

Aunque tal modificación en cualquiera de estas posiciones dadas puede en principio ser una sustitución con cualquier resto de aminoácido, se prefiere que la sustitución sea una sustitución con un resto de aminoácido dado en el número de carril 4 de una cualquiera o más de las Tablas 2, 3 o 4, o con un resto de aminoácido similar al mismo (mutación similar como se ha definido anteriormente). Así, podrían introducirse mutaciones similares como se ha definido anteriormente en lugar de las enumeradas. El Ejemplo 5 muestra algunos de tales mutantes. Los métodos para la introducción de mutaciones son conocidos en la técnica. Puede tomarse como guía a modo de ejemplo el Ejemplo 1. En otras palabras, una realización preferida de la invención se refiere a posiciones preferidas de mutagénesis de endoglucanasas de la clase GH7. Una lista de intercambios preferidos se da en la Tabla 2, carril 2. En otra realización preferida, las mutaciones preferidas se seleccionan del listado de la Tabla 3, carril 2. En otra realización preferida de la invención, las mutaciones preferidas se seleccionan de la Tabla 4, carril 2. También es posible combinar dos o tres de estas realizaciones preferidas, por ejemplo, pueden combinarse uno o más intercambios preferidos dados en la Tabla 2 con uno o más intercambios preferidos dados en la Tabla 3 y/o Tabla 4.

Tabla 2: Intercambios preferidos de aminoácidos con respecto a SEQ ID NO. 2

Posición de SEQ ID NO: 2	Intercambios alternativos de aminoácidos para SEQ ID NO: 2				Posición de SEQ ID NO: 2	Intercambios alternativos de aminoácidos para SEQ ID NO: 2			
	Aminoácido en la posición en SEQ ID NO: 2	Aminoácidos más preferidos en endoglucanasas GH7 por el alineamiento ClustalW para SEQ ID NO: 2	Aminoácidos más preferidos en endoglucanasas GH7 por el alineamiento ClustalW para SEQ ID NO: 2	Aminoácidos más preferidos en endoglucanasas GH7 por el alineamiento ClustalW para SEQ ID NO: 2		Aminoácido en la posición en SEQ ID NO: 2	Aminoácidos más preferidos en endoglucanasas GH7 por el alineamiento ClustalW para SEQ ID NO: 2	Aminoácidos más preferidos en endoglucanasas GH7 por el alineamiento ClustalW para SEQ ID NO: 2	Aminoácidos más preferidos en endoglucanasas GH7 por el alineamiento ClustalW para SEQ ID NO: 2
	Carril 1	Carril 2	Carril 3	Carril 4		Carril 1	Carril 2	Carril 3	Carril 4
2	L	L,Q	L,Q	Q	217	A	A,Y	A	
8	T	T,C	T		231	R	R,G,H,K	R,H,K	H,K
16	T	T,C	T		233	G	G,N	G	
19	K	K,E	K		235	P	P,S	P	
23	S	S,H	S,H	H	242	G	G,D	G	
30	N	N,D	N		249	P	P,R	P,R	R
32	Y	Y,S	Y,S	S	261	G	G,C	G	
41	W	W,R	W		263	P	P,T	P,T	T
42	I	I,M	I,M	M	271	T	T,K	T,K	K
48	IM	N,Y	N		277	N	N,D,E	N,D,E	D,E
55	G	G,C	G,C	C	290	T	T,E	T,E	E
64	E	E,H,K	E,H,K	H,K	293	S	S,T	S	
65	A	A,D	A,D	D	299	T	T,A	T,E	E
67	G	G,C	G,C	C	312	E	E,D,S	E,D,S	D,S
68	S	S,C,G	S,C,G	C,G	317	I	I,V	I	
75	G	G,C	G		322	N	N,W	N	
86	N	N,S	N,S	S	323	D	D,N	D	
88	S	S,C,D,F,T	S,D	D	325	S	S,T	S	
93	N	N,H,R,Y	N,H,R	H,R	327	Y	Y,F	Y	
104	T	T,S	T,S	S	328	M	M,K	M	

Intercambios alternativos de aminoácidos para SEQ ID NO: 2	Aminoácidos más preferidos en endoglucanasas GH7 por el alineamiento ClustalW para SEQ ID NO: 2	Aminoácido en la posición en SEQ ID NO: 2	Posición de SEQ ID NO: 2
E,S	D,E,S	D	335
	N	N	352
E	H,E	H	357
	Y	Y	3E0
L	P,L	P	379
	P	P,del	382
	P	P,del	383
	S	S,L	390
I	T,	T,I	391
T	A,T	A	392
T	S,T	S	398
T	I,T	I	405
	H	Y,H	431
G	S,G	S	432
Y	D,Y	D	434
	H	H,Y	448

Intercambios alternativos de aminoácidos para SEQ ID NO: 2	Aminoácidos más preferidos en endoglucanasas GH7 por el alineamiento ClustalW para SEQ ID NO: 2	Aminoácido en la posición en SEQ ID NO: 2	Posición de SEQ ID NO: 2
	S	S	107
	K,E	K	118
D	A,D	A	137
	A,G	A	144
	S,A	S	145
E	Q,E	Q	150
	E,K	E	153
	G,S	G	164
	Q,L	Q	179
D,E	T,D,E	T	185
K	Q,K	Q	191
	F,S	F	193
F	L,F	L	201
M	L,M,Y	L	210
	P,L,S	P	212
	N,T	N	216

Tabla 3: Intercambios preferidos de aminoácidos con respecto a SEQ ID NO. 2

Intercambios alternativos de aminoácidos para SEQ ID NO: 2	Aminoácidos más preferidos en endoglucanasas GH7 por el alineamiento ClustalW para SEQ ID NO: 2	Aminoácido en la posición en SEQ ID NO: 2	Posición de SEQ ID NO: 2
Carril 4	Carril 2	Carril 1	
H	S,H	S	23
	N	N	30
	W	W	41

ES 2 607 061 T3

Posición de SEQ ID NO: 2	Aminoácido en la posición en SEQ ID NO: 2	Aminoácidos preferidos en endoglucanasas GH7 por el alineamiento ClustalW para SEQ ID NO: 2	Aminoácidos más preferidos en endoglucanasas GH7 por el alineamiento ClustalW para SEQ ID NO: 2	intercambios alternativos de aminoácidos para SEQ ID NO: 2
55	G	G,C	G,C	C
64	E	E,H,K	E,H,K	H,K
65	A	A,D	A,D	D
67	G	G,C	G,C	C
118	K	K,E	K	
137	A	A,D	A,D	D
144	A	A,G	A	
150	Q	Q,E	Q,E	E
164	G	G,S	G	
179	Q	Q,L	Q	
185	T	T,D,E	T,D,E	D,E
191	Q	Q,K	Q,K	K
201	L	L,F	L,F	F
212	P	P,L,S	P	
216	N	N,T	N	
231	R	R,G,H,K	R,H,K	H,K
242	G	G,D	G	
249	P	P,R	P,R	R
261	G	G,C	G	
263	P	P,T	P,T	T
271	T	T,K	T,K	K
277	N	N,D,E	N,D,E	D,E
290	T	T,E	T,E	E
299	T	T,A	T,E	E
312	E	E,D,S	E,D,S	D,S
323	D	D,N	D	
325	S	S,T	S	
328	M	M,K	M	
335	D	D,E,S	D,E,S	E,S

Intercambios alternativos de aminoácidos para SEQ ID NO: 2	Aminoácidos más preferidos en endoglucanasas GH7 por el alineamiento ClustalW para SEQ ID NO: 2	Aminoácidos preferidos en endoglucanasas GH7 por el alineamiento ClustalW para SEQ ID NO: 2	Aminoácido en la posición en SEQ ID NO: 2	Posición de SEQ ID NO: 2
E	H,E	H,E	H	357
L	P,L	P,L	P	379
	S	S,L	S	390
I	T,I	T,I	T	391
T	I,T	I,T	I	405
G	S,G	S,G	S	432
Y	D,Y	D,Y	D	434

Tabla 4: Intercambios preferidos de aminoácidos con respecto a SEQ ID NO. 2

Intercambios alternativos de aminoácidos para SEQ ID NO: 2	Aminoácidos más preferidos en endoglucanasas GH7 por el alineamiento ClustalW para SEQ ID NO: 2	Aminoácidos preferidos en endoglucanasas GH7 por el alineamiento ClustalW para SEQ ID NO: 2	Aminoácido en la posición en SEQ ID NO: 2	Posición de SEQ ID NO: 2
Carril 4	Carril 3	Carril 2	Carril 1	
Q	L,Q	L,Q	L	2
	T	T,C	T	8
C	T,C	T,C	T	16
	K	K,E	K	19
S	Y,S	Y,S	Y	32
M	I,M	I,M	I	42
	N	N,Y	N	48
G	S,G	S,G	S	68
	G	G,C	G	75
5	N,S	N,S	N	86
C	S,C	S,C,D,F,T	S	88
	N	N,H,R,Y	N	93

Posición de SEQ ID NO: 2	Aminoácido en la posición en SEQ ID NO: 2	Aminoácidos preferidos en endoglucanasas GH7 por el alineamiento ClustalW para SEQ ID NO: 2	Aminoácidos más preferidos en endoglucanasas GH7 por el alineamiento ClustalW para SEQ ID NO: 2	intercambios alternativos de aminoácidos para SEQ ID NO: 2
104	T	T,S	T,S	S
107	S	S,T	S	
145	S	S,A	S	
153	E	E,K	E	
193	F	F,S	F	
210	L	L,M,Y	L	
217	A	A,Y	A	
233	G	G,N	G	
235	P	P,S	P	
293	S	S,T	S	
317	I	I,V	I	
322	N	N,W	N	
327	Y	Y,F	Y	
352	N	N,V,W	N,V	V
360	Y	Y,F	Y	
382	P	P,del	P	
383	P	P,del	P	
392	A	A,T	A,T	T
398	S	S,T	S,T	T
431	Y	Y,H	H	
448	H	H,Y	H	

- El primer y segundo aspecto de la invención (reivindicados en tanto que estén cubiertos por las reivindicaciones), aunque son soluciones diferentes para el mismo problema, no son necesariamente mutuamente excluyentes. Así, la invención se refiere a proteínas que cumplen con las condiciones tanto del primer aspecto como del segundo aspecto anteriores. Es importante notar que el primer aspecto y el segundo aspecto son dos soluciones alternativas al problema de proporcionar enzimas GH7 con estabilidad a la temperatura mejorada. Estas soluciones son independientes (aunque para algunos ejemplos se solapan) y así no necesitan ser necesariamente combinadas. Por ejemplo, la proteína identificada como [6] en la Figura 3 muestra estabilidad a la temperatura en comparación con la proteína que tiene SEQ ID NO: 4 ([2] en la Figura 3), sin embargo no muestra termoestabilización activa.
- 5 Así, la invención proporciona diferentes variantes de enzima según el primer aspecto anterior y/o según el segundo aspecto anterior. Puede probarse fácilmente si cualquier enzima dada muestra las propiedades térmicas deseadas (estabilidad a la temperatura y/o termoestabilización activa) por la prueba denominada "Determinación de la termoestabilidad y/o termoestabilización activa" anterior.
- 10

Como se da en detalle anteriormente en la sección de definiciones, además de individualizado por los ejemplos de más abajo, ahora se resume brevemente cómo pueden obtenerse las mutaciones deseadas:

- Alineamiento por pares de cualquier secuencia de endoglucanasa GH7 con SEQ ID NO: 2 usando el algoritmo ClustalW
- 5
- Identificación de posiciones correspondientes (carril 1) en la secuencia diana de endoglucanasa GH7
 - Modificación de posiciones correspondientes de la secuencia diana de endoglucanasa GH7 según los intercambios preferidos propuestos dados en el carril 2, o preferentemente en el carril 3
 - Expresión de la secuencia modificada y prueba de la proteína expresada para propiedades térmicas mejoradas
- 10
- Se cree que las enzimas termoestables de la invención también presentan reducción de la formación de aglomerados a temperaturas más altas, y así precipitación reducida. La anulación de tales precipitados es particularmente ventajosa en presencia de granates, tela vaquera o materiales tejidos, además de para la aplicación en reactores de membrana, reduciendo las características de ensuciamiento de la membrana.
- Las proteínas de fusión que comprenden cualquier proteína de la invención también son parte de la invención.
- 15
- Otra aspecto relacionado con la invención, pero no reivindicado por la invención, está relacionado con la producción de las proteínas de la invención por expresión heteróloga en un huésped de producción, también llamado huésped de expresión. Los métodos para la expresión heteróloga comprenden la transferencia de un ácido nucleico que codifica la proteína de la invención (construcción de expresión) en el huésped de producción por transformación, transfección, cruzamiento o métodos equivalentes con respecto a la transferencia de ácido nucleico (ADN o ARN).
- 20
- Los métodos para la transformación dentro del significado de la presente invención no están particularmente limitados. Se han informado ejemplos para una variedad de especies e incluyen electroporación, transformación en protoplastos, transformación química y transferencia mediante partículas balísticas, microinyección, infección viral, entrecruzamiento, o el uso de cepas o líneas de células competentes naturales. Un huésped de producción preferido segrega conjuntamente la endoglucanasa de la invención con otras celulasas, hemicelulasas o pectinasas en el caldo de cultivo. Así, se prefiere que la secuencia codificante de la construcción de expresión codifique la endoglucanasa de la invención precedida por una señal para la secreción de la cepa huésped particular. Tales señales son muy conocidas en la técnica; por ejemplo, en las eubacterias se denominan péptidos señal. Sin desear ceñirse a teoría particular, estas señales tienen en común la capacidad de dirigir la secreción de una proteína, normalmente en modo de co-traducción. Un huésped de expresión preferido es *Trichoderma reesei*.
- 25
- Un aspecto adicional de la invención es la aplicación de las proteínas endoglucanasas anteriormente descritas. Esto incluye las aplicaciones de las preparaciones de proteína purificada, parcialmente purificada o cruda, como tales, o en una formulación de enzima, además de la aplicación de células completas u organismos que expresan la proteína diana. Los campos de aplicación de las endoglucanasas pueden encontrarse en el capítulo Campo de la invención. Como se estableció ahí, es altamente deseable la aplicación de proteínas termoestables. Una aplicación preferida de la endoglucanasa está en el campo de la conversión enzimática de lignocelulosa.
- 30
- 35

Visión general de las secuencias desveladas en el presente documento, en las que SEQ ID NOs 5, 8, 13 y 14 no son parte de la invención

SEQ ID NO	Tipo	Función	Fuente
1	ADN	Secuencia de ADN que codifica SEQ ID NO: 2 (endoglucanasa GH7 de la invención) - uso de codón adaptado para <i>Pichia pastoris</i>	Artificial
2	Proteína	Endoglucanasa GH7 de la invención-secuencia de proteína madura	Artificial
3	ADN	Secuencia de ADN que codifica SEQ ID NO: 4 (ADN de endoglucanasa GH7 Cel7B) - uso de codón adaptado	Artificial
4	Proteína	Endoglucanasa I (Cel7B) - secuencia de proteína madura	<i>Hypocrea pseudokoningii</i> -ABM90986
5	Proteína	Endoglucanasa I (Cel7B) - secuencia de proteína madura	Artificial
6	Proteína	Endoglucanasa GH7 de la invención - secuencia de proteína madura	Artificial
7	Proteína	Endoglucanasa GH7 de la invención - secuencia de proteína madura	Artificial

SEQ ID NO	Tipo	Función	Fuente
8	Proteína	Endoglucanasa I (Cel7B) - variante de la secuencia de proteína madura	Artificial
9	Proteína	Endonucleasa GH7 de la invención - secuencia de proteína madura	Artificial
10	Proteína	Endonucleasa GH7 de la invención - secuencia de proteína madura	Artificial
11	Proteína	Endonucleasa GH7 de la invención - secuencia de proteína madura	Artificial
12	Proteína	Endonucleasa GH7 de la invención - secuencia de proteína madura	Artificial
13	Proteína	Endoglucanasa I (Cel7B) - variante de la secuencia de proteína madura	Artificial
14	Proteína	Endoglucanasa I (Cel7B) - variante de la secuencia de proteína madura	Artificial
15	ADN	Secuencia de ADN que codifica la expresión de SEQ ID NO. 2 con TAG del extremo N 6x histidina (<i>cursiva</i>) con péptido señal SP _{mfa} (<u>subrayado</u>) en <i>Pichia pastoris</i>	Artificial
16	ADN	Secuencia de ADN que codifica SEQ ID NO. 2 en fusión con el péptido señal CBHI (<u>subrayado</u>) para la expresión en <i>Trichoderma reesei</i>	Artificial

Secuencias desveladas en el presente documento (NO: 1 -16)

SEQ ID NO: 1

TCTCTGCAGCCAGGAACCTTCTACTCCAGAGGTGCACCCAAAGCTGACCACCTACAAGTGTACCACCTCTGGTGGT
 TGTGTGCTCAGAACACCTATGTTGTTCTGGACTGGAACCTACAGATGGATCCACGACGCCAACTACAACCTTTGT
 ACCGTGAACGGTGGTGTCAACACTACTCTGTGTCCAGACGAGGCTACTGGTAGCAAGAAGCTGCTTCATCGAGGGT
 GTTGACTACGCTGCTTCTGGTGTACTGCCAATGGTCTACCTTGACCCTGAACCAGTACATGCCATCTTCTCT
 GCGGTTACACTTCTGTGTCGCCAAGACTGTACTTGTGGGTCCAGACGGTAAGTACGTTATGCTGAAGCTGAAC
 GGACAGGAGCTGTCTTTGACGTTGACCTGTCTGCTTTGCCATGTGGAGAGAACGCTTCTCTGTACCCTGTCTCAG
 ATGGACGAGAACGGTGGAGCTAACAGTACAACACCCGCGGTGCTAACTACGGTCTGGTTACTGTGACGCCAG
 TGTCCAGTTCAGACTTGGAGAAAACGGAACCTGAACACTTCTGGCCAGGGATTCTGCTGTAACGAGATGGACATC
 TTGGAGGAAAACCTAGAGCTAACGCTCTGACCCACACTCTTGTAAATGCTACCGCTTGTGACTCTGTGGTTGC
 GGTTTAACCACATACCGCTCGGGTTACCCAAACTACTTTGGCCAGGTGGCACTGTTGACACCTCGAAGCCATT
 ACCATCATCACCCAGTTCAACACCGACAACGGTCTCCATCTGGTAACCTGGTGTGATCACCAGAAAGTACAGA
 CAGAACGGCGTTGACATCCCATCTGCTAAACCAGGTGGCGACACCATTTGCTCTTGTCCATCTGCCTCTACTTAC
 GGTGGATTGGCTACCATGGGAAAGGCTCTGTCCGAGGGAATGGTGTGATCTTCTCGATCTGGAACGACAACCTCG
 CAGTACATGAACGGCTGGACTCTGGTGTGCTGGTCCATGTTCTTCTACCGAGGGCAACCCATCTAACATCCTG
 GCTAACAAACCTGGTACTCAGTGGTGTACTCGAACATTAGATGGGGCGACATGGTTCTACCACCAACTCTACC
 GGTGGTAACCCACCACCACCTGCATCTTCTACCACCTTCTCGACCGCCAGAAGATCGTCTACCTCCTCTTCT
 TCTCCATCTTGTATCCAGACTCACTGGGGTCAGTGTGGTGGTATTGGCTACACCGGCTGTAAGACCTGTACCTCT
 GGAACCACTTGCCAGTACAGCAACGACTACTACTCTCAGTGCCTGTGA

SEQ ID NO: 2

SLQPGTSTPEVHPKLTYYKCTTSGGCVAQNTYVVLWDWNYRWHIDANYNSCTVNGGVNTLLCPDEATGSKNCFIEG
 VDYAASGVTANGSTLTLNQYMPSSSGGYTSVSPRLYLLGPDGKYVMLKLNQELSFVDVLSALPCGENASLYLSQ
 MDENGGANQYNTAGANYGSGYCDAPVQVTRNGTLNLSGQGFCCNEMDILEGNSRANALTPHSCNATAACDSAGC
 GFNPYRSGYPNYFGPGGTVDTSKPFTIITQFNTDNGSPSGNLVSI TRKYRQNGVDIPSAKPGGDTISSCPSASTY
 GGLATMGKALSEGMVLIFS IWN DNSQYMNWLDSDAGPCSSTEGNSNILANNP GTHVVSNI RWDIGSTTNST
 GGNPPPPASSTTFSTARRSSTSSSSPSCIQTHWQCGGIGYTGCKTCTSGTTCQYSNDYYSQCL*

ES 2 607 061 T3

SEQ ID NO: 3

TCTCAGCAGCCAGGAACCTTCTACTCCAGAGGTGCACCCAAAGCTGACCACCTACAAGTGTACCACCTCTGGTGGT
TGTTGTGCTCAGGACACCTCTGTGTTCTGGACTGGAACCTACAGATGGATGCACGACGCCAACTACAACCTCTTGT
ACCGTGAACGGTGGTGTCAACACTACTCTGTGTCCAGACGAGGCTACTTGTGGCAAGAAGCTGCTTCATCGAGGGT
GTTGACTACGCTGCTTCTGGTGTACTGCCTCTGGTTCTACCTTGACCCTGAACCAGTACATGCCATCTTCTCT
GGCGGTTACTCTTCTGTGTCGCCAAGACTGTACTTGTGGGTCCAGACGGTGAAGTACGTTATGCTGAAGCTGAAC
GGACAGGAGCTGTCTTTTACGTTGACCTGTCTGCTTTGCCATGTGGAGAGAACCGGTTCTCTGTACCTGTCTCAG
ATGGACGAGAACGGTGGAGCTAACCCAGTACAACACCCGCCGGTGTAACTACGGTCTGGTTACTGTGACGCCAG
TGTCCAGTTCAGACTTGGAGAAACCGAACCTTCTGACCCTTCTGTCGATACAGAGATGGACATC
TTGGAGGGAAACTCTAGAGCTAACGCTCTGACCCACACTCTTGTACTGCTACCGCTTGTGACTCTGCTGGTTGC
GGTTTTAACCCATACGGCTCGGGTTACCCAAACTACTTTGGCCAGGTGACACTGTTGACACCTCGAAGCCATTC
ACCATCATCACCCAGTTCAACACCGACAACGGTCTCCATCTGGTAACCTGGTGTGATCACCAGAAAGTACAGA
CAGAACGGCGTTGACATCCCATCTGCTAAACAGGTGGGACACCATTTCTGCTTGTCCATCTGCCTCTGCTTAC
GGTGGATTGGCTACCATGGGAAAGGCTCTGTCTCTGGAATGGTGTGATCTTCTCGATCTGGAACGACAACCTG
CAGTACATGAACCTGGCTGACTCTGGTTCTGCTGGTCCATGTTCTTCTACCGAGGGCAACCCATCTAACACTCTG
GCTAACCAACCCIGGTACTCAGTGGTGTACTCGAACATTAGATGGGGCGACATGGTTCTACCACCACTTACC
GGTGGTAACCCACCACCACCCTGCATCTTCTACCACCTTCTCGACCACCAGAAGATCGTCTACCACCTCTTCT
TCTCCATCTTGTACCCAGACTCACTGGGGTCAAGTGTGGTGGTATGGCTACACCGGCTGTAAGACCTGTACCTCT
GGAACCACTTGGCAGTACGGCAACGACTACTACTCTCAGTGCCTGTGA

SEQ ID NO: 4

SQQPGTSTPEVHPKLTTYKCTTSSGGCVAQDTSVVLWDWNYRWMHDANYNSCTVNGGVNTTLLCPDEATCGKNCFIG
VDYAASGVTASGSLTLTLNQYMPSSSGGYSSVSPRLLYLLGPDGEYVMLKLNQQLSFDVDLSALPCGENGLYLSQ
MDENGANQYNTAGANYGSGYCDACQCPVQTRNGTLNTSGQGFCCNEMDILEGNSRANALTPHSCATAACDSAGC
GFNPYSGYPNYFGPGDVTDSKPFITITQFNTDNGSPSGNLVSI TRKYRQNGVDIPSAKPGGDTISSCPSASAY
GGLATMGKALSSGMVLIFSIWINDNSQYMNWLDSSGAGPCSSTEGNP SNILANNPGTHVVYSNIRWGDIGSTTNST
GGNPPPPASSTTFSTTRRSSTSSSPSCTQTHWQCGGIGYTGCKTCTSGTTCQYGN DYYSQCL*

SEQ ID NO: 5

SQQPGTSTPEVHPKLTTYKCTTSSGGCVAQDTSVVLWDWNYRWMHDANYNSCTVNGGVNTTLLCPDEATCGKNCFIG
VDYAASGVTASGSLTLTLNQYMPSSSGGYSSVSPRLLYLLGPDGEYVMLKLNQQLSFDVDLSALPCGENGLYLSQ
MDKNGANQYNTAGANYGSGYCDACQCPVQTRNGTLNTSGQGFCCNEMDILEGNSRANALTPHSCATAACDSAGC
GFNPYSGYPNYFGPGDVTDSKPFITITQFNTDNGSPSGNLVSI TRKYRQNGVDIPSAKPGGDTISSCPSASAY
GGLATMGKALSSGMVLIFSIWINDNSQYMNWLDSSGAGPCSSTEGNP SNILANNPGTHVVYSNIRWGDIGSTTNST
GGNPPPPASSTTFSTTRRSSTSSSPSCTQTHWQCGGIGYTGCKTCTSGTTCQYGN DYYSQCL*

SEQ ID NO: 6

SLQPGTSTPEVHPKLTTYKCTTSSGGCVAQNTSVVLWDWNYRWMHDANYNSCTVNGGVNTTLLCPDEATGGKNCFIG
VDYAASGVTASGSLTLTLNQYMPSSSGGYSSVSPRLLYLLGPDGKYVMLKLNQQLSFDVDLSALPCGENASLYLSQ
MDENGANQYNTAGANYGSGYCDACQCPVQTRNGTLNTSGQGFCCNEMDILEGNSRANALTPHSCNATAACDSAGC
GFNPYSGYPNYFGPGDVTDSKPFITITQFNTDNGSPSGNLVSI TRKYRQNGVDIPSAKPGGDTISSCPSASTY
GGLATMGKALSSGMVLIFSIWINDNSQYMNWLDSSGAGPCSSTEGNP SNILANNPGTHVVYSNIRWGDIGSTTNST
GGNPPPPASSTTFSTTRRSSTSSSPSCIQTHWQCGGIGYTGCKTCTSGTTCQYSNDYYSQCL*

SEQ ID NO: 7

SLQPGTSTPEVHPKLTTYKCTTSSGGCVAQNTYVVLWDWNYRWIHDANYNSCTVNGGVNTTLLCPDEATGSKNCFIG
VDYAASGVTANGSLTLTLNQYMPSSSGGYTSVSPRLLYLLGPDGKYVMLKLNQQLSFDVDLSALPCGENASLYLSQ
MDENGANQYNTAGANYGSGYCDACQCPVQTRNGTLNTSGQGFCCNEMDILEGNSRANALTPHSCNATAACDSAGC
GFNPYSGYPNYFGPGDVTDSKPFITITQFNTDNGSPSGNLVSI TRKYRQNGVDIPSAKPGGDTISSCPSASTY
GGLATMGKALSSGMVLIFSIWINDNSQYMNWLDSSGAGPCSSTEGNP SNILANNPGTHVVYSNIRWGDIGSTTNST
GGNPPPPASSTTFSTTRRSSTSSSPSCIQTHWQCGGIGYTGCKTCTSGTTCQYSNDYYSQCL*

ES 2 607 061 T3

SEQ ID NO: 8

SQQPGTSTPEVHPKLTTYKCTTSGGCVAQNTSVVLDWNYRWMHDANYNSCTVNGGVNTTLCPEATCGKNCFIG
VDYAASGVTASGSTLTILNQYMPSSSGGYSSVSPRLLYLLGPDGKYVMLKLNQELSFVDVLSALPCGENASLYLSQ
MDENGGANQYNTAGANYGSGYCDACQCPVQVQWRNGTLNTSGQGFCNEMDILEGNSRANALTPHSCNATAACDSAGC
GFNPYKSGYPNYFPGPQVDTVDTSKPFTIIITQFNTDNGSPSGNLVSI TRKYRQNGVDIPSAKPGGDTISSCPASAY
GGLATMGKALSSGMVLIFSIWINDNSQYMNWLDSSGAGPCSSSTEGNPSNILANNPGTHVVYSNIRWGDIGSTTNST
GGNPPPPASSTTFSTTRRSSTTSSSPSCTQTHWGQCGGIGYTGCKTCTSGTTCQYGN DYYSQCL*

SEQ ID NO: 9

SLQPGTSTPEVHPKLTTYKCTTSGGCVAQNTSVVLDWNYRWMHDANYNSCTVNGGVNTTLCPEATCCKNCFIG
VDYAASGVTASGSTLTILNQYMPSSSGGYSSVSPRLLYLLGPDGKYVMLKLNQELSFVDVLSALPCGENASLYLSQ
MDENGGANQYNTAGANYGSGYCDACQCPVQVQWRNGTLNTSGKGFCNEMDILEGNSRANALTPHSCNATAACDSAGC
GFNPYKSGYPNYFPGPQVDTVDTSKPFTIIITQFNTDNGSPSGNLVSI TRKYRQNGVDIPSAKPGGDTISSCPASAY
GGLATMGKALSSGMVLIFSIWINDNSQYMNWLDSSGAGPCSSSTEGNPSNILANNPGTHVVYSNIRWGDIGSTTNST
GGNPPPPASSTTFSTTRRSSTTSSSPSCTQTHWGQCGGIGYTGCKTCTSGTTCQYSNDYYSQCL*

SEQ ID NO: 10

SQQPGTSTPEVHPKLTTYKCTTSGGCVAQNTSVVLDWNYRWMHDANYNSCTVNGGVNTTLCPEATCGKNCFIG
VDYAASGVTASGSTLTILNQYMPSSSGGYSSVSPRLLYLLGPDGKYVMLKLNQELSFVDVLSALPCGENASLYLSQ
MDENGGANQYNTAGANYGSGYCDACQCPVQVQWRNGTLNTSGQGFCNEMDILEGNSRANALTPHSCNATAACDSAGC
GFNPYKSGYPNYFPGPQVDTVDTSKPFTIIITQFNTDNGSPSGNLVSI TRKYRQNGVDIPSAKPGGDTISSCPASAY
GGLATMGKALSEGVLIFSIWINDNSQYMNWLDSSGAGPCSSSTEGNPSNILANNPGTHVVYSNIRWGDIGSTTNST
GGNPPPPASSTTFSTTRRSSTTSSSPSCTQTHWGQCGGIGYTGCKTCTSGTTCQYGN DYYSQCL*

SEQ ID NO: 11

SQQPGTSTPEVHPKLTTYKCTTSGGCVAQNTSVVLDWNYRWMHDANYNSCTVNGGVNTTLCPEATCGKNCFIG
VDYAASGVTASGSTLTILNQYMPSSSGGYSSVSPRLLYLLGPDGKYVMLKLNQELSFVDVLSALPCGENASLYLSQ
MDENGGANQYNTAGANYGSGYCDACQCPVQVQWRNGTLNTSGQGFCNEMDILEGNSRANALTPHSCNATAACDSAGC
GFNPYKSGYPNYFPGPQVDTVDTSKPFTIIITQFNTDNGSPSGNLVSI TRKYRQNGVDIPSAKPGGDTISSCPASAY
GGLATMGKALS DGMVLIFSIWINDNSQYMNWLDSSGAGPCSSSTEGNPSNILANNPGTHVVYSNIRWGDIGSTTNST
GGNPPPPASSTTFSTTRRSSTTSSSPSCTQTHWGQCGGIGYTGCKTCTSGTTCQYGN DYYSQCL*

SEQ ID NO: 12

SQQPGTSTPEVHPKLTTYKCTTSGGCVAQNTSVVLDWNYRWMHDANYNSCTVNGGVNTTLCPEATCGKNCFIG
VDYAASGVTASGSTLTILNQYMPSSSGGYSSVSPRLLYLLGPDGKYVMLKLNQELSFVDVLSALPCGENASLYLSQ
MDENGGANQYNTAGANYGSGYCDACQCPVQVQWRNGTLNTSGQGFCNEMDILEGNSRANALTPHSCNATAACDSAGC
GFNPYKSGYPNYFPGPQVDTVDTSKPFTIIITQFNTDNGSPSGNLVSI TRKYRQNGVDIPSAKPGGDTISSCPASAY
GGLATMGKALS DGMVLIFSIWINDNSQYMNWLDSSGAGPCSSSTEGNPSNILANNPGTHVVYSNIRWGDIGSTTNST
GGNPPPPASSTTFSTTRRSSTTSSSPSCTQTHWGQCGGIGYTGCKTCTSGTTCQYGN DYYSQCL*

SEQ ID NO: 13

SQQPGTSTPEVHPKLTTYKCTTSGGCVAQNTSVVLDWNYRWHHDANYNSCTVNGGVNTTLCPEATCSKNCFIG
VDYAASGVTANGSTLTILNQYMPSSSGGYSSVSPRLLYLLGPDGEYVMLKLNQELSFVDVLSALPCGENGSLYLSQ
MDENGGANQYNTAGANYGSGYCDACQCPVQVQWRNGTLNTSGQGFCNEMDILEGNSRANALTPHSCNATAACDSAGC
GFNPYKSGYPNYFPGPQVDTVDTSKPFTIIITQFNTDNGSPSGNLVSI TRKYRQNGVDIPSAKPGGDTISSCPASAY
GGLATMGKALSSGMVLIFSIWINDNSQYMNWLDSSGAGPCSSSTEGNPSNILANNPGTHVVYSNIRWGDIGSTTNST
GGNPPPPASSTTFSTTRRSSTTSSSPSCTQTHWGQCGGIGYTGCKTCTSGTTCQYGN DYYSQCL*

SEQ ID NO: 14

SQQPGTSTPEVHPKLTTYKCTTSGGCVAQNTSVVLDWNYRWHHDANYNSCTVNGGVNTTLCPEATCGKNCFIG
VDYAASGVTANGSTLTILNQYMPSSSGGYSSVSPRLLYLLGPDGEYVMLKLNQELSFVDVLSALPCGENGSLYLSQ
MDENGGANQYNTAGANYGSGYCDACQCPVQVQWRNGTLNTSGQGFCNEMDILEGNSRANALTPHSCNATAACDSAGC
GFNPYKSGYPNYFPGPQVDTVDTSKPFTIIITQFNTDNGSPSGNLVSI TRKYRQNGVDIPSAKPGGDTISSCPASAY
GGLATMGKALSSGMVLIFSIWINDNSQYMNWLDSSGAGPCSSSTEGNPSNILANNPGTHVVYSNIRWGDIGSTTNST
GGNPPPPASSTTFSTTRRSSTTSSSPSCTQTHWGQCGGIGYTGCKTCTSGTTCQYGN DYYSQCL*

SEQ ID NO: 15

atgagatttccttcaatttttactgcagttttatctgcagcatcctccgcattagctgctccagtcaacactaca
acagaagatgaaacggcacaattccggctgaagctgtcatcggttacttagatttagaaggggatttcgatggt
gctgttttgccattttccaacagcacaataacgggttattggttataaatactactattgccagcattgctgct
aaagaagaaggggtatctttggataaacgtgagggcgaagcatgccaccaccaccaccactcctcgggtct
ctgcagccaggaacttctactccagaggtgcacccaaagctgaccacctacaagtgtaccacctctggtggtgt
ggtgctcagaacacctatggtgtctgactggaactacagatggatccacgacgccaactacaactctgtacc
gtgaacgggtggtgtcaacactactctgtgtccagacgaggetactggtagcaagaactgcttcacogaggggtgt
gactacgctgcttctggtgttactgccaatggttctacctgaccctgaaccagtagatgccatcttctctggc
ggttacacttctgtgctcgcaagactgtactgttgggtccagacggtaagtacggttatgctgaagctgaacgga
caggagctgtctttgacgttgacctgtctgctttgcatgtggagagaacgcttctctgtacctgtctcagatg
gacgagaacgggtggagctaacagtagacaacaccgcccgtgctaactacggttctggttactgtgacgcccagtgt
ccagttcagacttggagaaacggaacctgaacacttctggccagggattctgctgtaacagatggacatcttg
gagggaaactctagagctaacgctctgaccccacactctgtaatgctaccgcttgtgactctgctggttgccgt
tttaaccataccgctcgggttacccaaactacttggcccaggtggcactgtgacacctcgaagccattcacc
atcatcaccagttcaacaccgacaacgggttctccatctggtaacctggtgtcgatcaccagaaagtacagacag
aacggcgttgacatcccatctgctaaaccaggtggcgacaccatttcgctctgtccatctgcctctacttacggt
ggattggctaccatgggaaaggtctgtccgagggaaatggtgctgatcttctcgatctggaacgacaactcgcag
tacatgaactggctggactctggtgatgctggtccatgttcttctaccgagggcaacctcacaacatcctggct
aacaacctggtactcacgtggtgactcgaacattagatggggcgacattggttctaccaccaactctaccggt
ggtaacccaccaccaccactgcactcttctaccacettctcgaccgcccagaagatcgtctacctctcttctct
ccatctgtatccagactcactggggcagtggtggtattggctacaccggtgtaagacctgtacctctgga
accacttgccagtagcaacgactactactctcagtgccctgtga

SEQ ID NO: 16:

atgtatcgggaagttggcctcatctcggccttcttggccacagcaacgggcttctctgcaaccgggtaccagcacc
cccgaggtccatcccaagttgacaacctacaagtgtaaacctccgggggggtgcgtggccagaacacctatgtg
gtccttgactggaactaccgctggatccacgacgcaactacaactcgtgacccgtcaacggcggcgtcaacacc
acgctctgcctgacgagggcaccggtagcaagaactgcttcatcgagggcgtcgactacgccgctcgggcgtc
acggccaatggcagcaccctcaccctgaaccagtagatgccagcagctctggcggctacactagcgtctctct
cggctgtatctcctgggtccagacggtaagtacgtgatgctgaagctcaacggccaggagctgagcttcgacgtc
gacctctctgctctgccgtgtggagagaacgcctcgtctacctgtctcagatggacgagaacgggggcccac
cagtagaacacggccgggtgccaactacgggagcggctactgcgatgctcagtgcccctcagacatggaggaac
ggcaccctcaacactagcggccagggcttctgctgcaacgagatggatctcctggagggcaactcgagggcgaat
gccttgacccctcactcttgcaatgccacggcctgcgactctgccggttgcggcttcaacccctatcgcagcggc
taccxaaactacttcggccccggaggcaccggttgacacctccaagccattcaccatcatcaccagttcaacacg
gacaacggctcgcctcgggcaaccttgtgagcatcaccgcaagtagacacaaaacggcgtcgacatccccagc
gccaaacccggcggcgacaccatctcgtcctgcccctcaccctcaacttacggcggcctcgcaccatgggcaag
gccctgagcggagggcattggtcctcctcagcatttggaacgacaacagccagtagaactggctcgacagc
ggcgtgcccggccctgcagcagcaccgagggcaacctccaacatcctggccaacaaccccggtagcagcgtc
gtctactccaacatccgctggggagacattgggtctactacgaactcagctggtggtccgccccgcctgcgtcc
agcacgagcttttcgactgcccgaggagctcgacgtcctcgagcagcccagctgcatccagactcactggggg
cagtgccgtggcattgggtacaccgggtgcaagacgtgcacgtcgggcaactcagtgccagtagcaacgactac
tactcgaatgcctttaa

Ejemplos

5 Ejemplo 1: Generación de bibliotecas y variantes específicas

Se produjo una biblioteca basada en SEQ ID NO: 3 (biblioteca “N7”) usando SEQ ID NO: 3 como molde por PCR propensa a error usando polimerasa Taq siguiendo el protocolo de la bibliografía (Joyce et al), usando condiciones de PCR del siguiente modo: 2 min a 95 °C, 30 ciclos de (1 min a 95 °C, 1 min a 56 °C, 1 min a 72 °C), 5 min a 72 °C. Todos los productos obtenidos de las PCRs se purificaron con el kit QIAquick PCR Purification (Qiagen, Hilden, Alemania).

10 Se prepararon variantes específicas de SEQ ID NO: 3 por un protocolo de PCR modificado usando cebadores que contenían la secuencia de nucleótidos mutada (Ho, S. N. et al. Gene; 1989; 77; 51-9).

Ejemplo 2: Expresión en *Pichia pastoris*

Construcción de un casete de expresión lineal (LEC) - Se construyeron LECs (Liu Z, et al. Chembiochem. 2008 Jan 4; 9 (1):58-61) con marcador Zeocin y el promotor GAP por un protocolo de PCR modificado.

5 Transformación y cultivo de *Pichia pastoris* - Se prepararon células competentes y se transformaron como se describe (Lin-Cereghino, J., et al. BioTechniques. 2005, 38, 44-48). Los transformantes se seleccionaron sobre placas de agar YPD que contenían Zeocin 100 mg/l, y se llevaron a placas de pocillos profundos (DWP) (BMD 5 % 250 ml/pocillo) por medio de un robot (QPix2, Genetix). Las DWPs inoculadas se cultivaron durante 60 h a 28 °C, 80 % de humedad y 280 rpm.

Ejemplo 3: Expresión en *Trichoderma reesei*

10 **Construcción del vector de expresión de *Trichoderma reesei***

Se transformó ADN del plásmido pV7 linealizado digerido con SbfI/SwaI (Figura 2) en *Trichoderma reesei* SCF41, esencialmente como se describe por Penttilä et al 1997. La selección de transformantes se hizo en placas de medio de Mandel's Andreotti que contenían higromicina como agente selectivo (100 mg/l). Los transformantes se verificaron por PCR.

15 **Ejemplo 4: Determinación de la capacidad de conversión de sustrato a diferentes temperaturas para indicación de la termoestabilidad de SEQ ID NO: 2 - Variantes usando 4-metilumbeliferil-β-D-celobiosido (4-MUC)**

20 Para una comparación precisa de la estabilidad térmica, se incubaron 10 µl de los sobrenadantes de cultivo de *Pichia pastoris* que contenían las variantes de endoglucanasa secretadas con 90 µl de 4-MUC 100 µM (disuelto en tampón acetato sódico (50 mM, pH 5.0)), en el gradiente de temperatura de un ciclador térmico de gradiente de Eppendorf. Se incubaron 24 mezclas de reacción a un gradiente de temperatura que iba de 45 °C a 65 °C y de 55 °C a 75 °C (cada reacción se mantuvo a un nivel de temperatura constante único) durante una hora. La actividad enzimática a la temperatura respectiva pudo determinarse después de la adición de 100 µl de carbonato de sodio 1 M a cada reacción, y medición de la intensidad de fluorescencia a 360 nm/454 nm en un lector de placas Tecan Infinite M200. Para comparación de la termoestabilidad, las cuentas de fluorescencia de cada punto de temperatura, se determinó la actividad enzimática relativa dividiendo entre el recuento máximo de una serie (normalización a 1). El perfil de temperatura para cualquier enzima dada se generó representando la actividad enzimática relativa sobre el intervalo de temperatura medido.

Ejemplo 5: Termoestabilización activa de algunas variantes de endoglucanasa

30 Este ejemplo describe ejemplos del sorprendente efecto de la *termoestabilización activa*. En este ejemplo se usaron proteínas (sobrenadante de cultivo) (Tabla 5 más adelante, en la que SEQ ID NOs 5, 8, 13 y 14 no son parte de la invención) expresadas en *Pichia pastoris*.

35 La Figura 3 muestra las propiedades determinadas de las proteínas de la invención: las proteínas designadas como [1], [4], [5], [7], [8], [9] y [10] muestran termoestabilización activa y estabilidad a la temperatura, mientras que las proteínas designadas [3] y [6] muestran estabilidad a la temperatura, pero no termoestabilización activa.

Tabla 5: Proteínas probadas en el Ejemplo 5, 6 y 7 Designación	SEQ ID NO:	Figura N.º	Mutaciones con respecto a SEQ ID NO 2
EGL de <i>Hypocrea pseudokoningii</i>	SEQ ID NO: 4	[2]	L2Q, N30D, Y32S, I42M, G67C, S68G, N86S, T104S, K118E, A144G, N216T, R231 G, G242D, T299A, E312S, D335S, A392T, S398T, I405T, S432G
Variante de EGL de <i>Hypocrea pseudokoningii</i> (E153K con respecto a SEQ ID NO: 4)	SEQ ID NO: 5	[3]	L2Q, N30D, Y32S, I42M, G67C, S68G, N86S, T104S, K118E, A144G, E153K, N216T, R231G, G242D, T299A, E312S, D335S, A392T, S398T, I405T, S432G
Variante de EGL de la invención; derivada por mutagénesis de la secuencia de EGL de <i>Hypocrea pseudokoningii</i>	SEQ ID NO: 2	[1]	
Variante 1 de ejemplo de SEQ ID NO: 2	SEQ ID NO: 6	[4]	Y32S, I42M, S68G, N86S, T104S, R231G, E312S, D335S, A392T

Tabla 5: Proteínas probadas en el Ejemplo 5, 6 y 7 Designación	SEQ ID NO:	Figura N.º	Mutaciones con respecto a SEQ ID NO 2
Variante 2 de ejemplo de SEQ ID NO: 2	SEQ ID NO: 7	[5]	R231G, E312S, D335S, A392T
Variante de EGI de <i>Hypocrea pseudokoningii</i> (con respecto a SEQ ID NO: 4)	SEQ ID NO: 8	[6]	,L2Q, Y32S, I42M, G67C, S68G, N86S, T104S, R231G, T2 99A, E312S, D335S, A392T, S398T, I405T, S432G
Variante 3 de ejemplo de SEQ ID NO: 2	SEQ ID NO: 9	[7]	, Y32S, I42M, G67C, S68C, N86S, T104S, Q191 K, R231 G, T299A, E312S, D335S, A392T, S398T, I405T
Variante 4 de ejemplo de SEQ ID NO: 2	SEQ ID NO: 10	[8]	, L2Q, Y32S, I42M, G67C, S68G, N86S, T104S, R231K, T2 99A, D335S, A392T, S398T, I405T, S432G
Variante 5 de ejemplo de SEQ ID NO: 2	SEQ ID NO: 11	[9]	, L2Q, Y32S, I42M, G67C, S68G, N86S, T104S, R231G, T2 99A, E312D, D335E, A392T, S398T, I405T, S432G
Variante 6 de ejemplo de SEQ ID NO: 2	SEQ ID NO: 12	[10]	, L2Q, Y32S, 142M, G67C, S68G, N86S, T104S, R231K, T2 99A, E312D, D335S, A392T, S398T, I405T, S432G
Variante 7 de ejemplo de SEQ ID NO: 2	SEQ ID NO: 13		, L2Q, N30D, Y32S, G67C, T104S, K118E, A144G, N216T, R231G, G242D, T299A, E312S, D335S, A392T, S398T, I4 05T, S432G
Variante 8 de ejemplo de SEQ ID NO: 2	SEQ ID NO: 14		, L2Q, N30D, Y32S, G67C, S68G, T104S, K118E, A144G, N 216T, R231G, G242D, T299A, E312S, D335S, A392T, S398T, I4 05T, S432G

Ejemplo 6: Determinación de la liberación de azúcar reductor en paja

5 Se determinó la liberación de azúcar reductor en paja aplicando paja de trigo pretratada con ácido con una materia seca del 2,5 %. Se añadieron las siguientes enzimas a la mezcla de reacción: celobiohidrolasa I (12,5 mg/l), beta-glucosidasa (40 CBU/mg de celobiohidrolasa I) y la variante de endoglucanasa GH7 probada (12,5 mg/l). La hidrólisis de la paja se incubó a 60 °C mediante agitación continua durante 48 h.

Ejemplo 7: Determinación del perfil de temperatura de las variantes de SEQ ID NO: 2

10 Para el ensayo de actividad de MUL (4-metilumbeliferil-β-D-lactopiranosido), se mezclaron 10 µl del cultivo sobrenadante con 90 µl de MUL 100 µM en tampón acetato de Na 25 mM a pH 4,8. Las placas se sellaron y se incubaron por 2 h, con agitación de 300 rpm, a 45 °C y 59 °C cada vez (para repetir el cribado también a 65 °C). La reacción se inactivó añadiendo 100 µl de Na₂CO₃ por pocillo. La excitación se realizó a 365 nm y la fluorescencia se midió a 450 nm. Los resultados se muestran en la Figura 5.

Ejemplo 8: Determinación del perfil de temperatura de las variantes de SEQ ID NO: 2

15 Se determinaron las semividas de las enzimas midiendo la actividad residual usando el ensayo de MUL descrito en el Ejemplo 7, después incubar los sobrenadantes de expresión de cultivos de *Pichia pastoris* a 70 °C durante 0 a 7min en un baño de agua. Las muestras se pusieron sobre hielo después del periodo de incubación preciso antes de establecer el ensayo de actividad.

REIVINDICACIONES

- 5 1. Una proteína que tiene actividad de endoglucanasa que comprende una secuencia de aminoácidos que tiene al menos el 96 %, preferentemente al menos el 97 %, más preferentemente al menos el 98 %, incluso más preferentemente al menos el 99 %, tal como al menos el 99,5 % de identidad con SEQ ID NO: 2, en la que la proteína muestra al menos el 90 % de capacidad de conversión de sustrato residual a la temperatura 60 °C cuando la incubación se hace durante una hora.
2. La proteína según la reivindicación 1, que es idéntica a SEQ ID NO: 6; SEQ ID NO: 7; SEQ ID NO: 9; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 12; o SEQ ID NO: 2; o una proteína que tiene al menos una mutación con respecto a SEQ ID NO: 2.
- 10 3. El polipéptido según la reivindicación 2, en el que la al menos una mutación es una seleccionada entre los cambios preferidos, más preferidos de cambios alternativos mostrados en la Tabla 2.
4. El polipéptido según la reivindicación 2, en el que la al menos una mutación es una seleccionada entre los cambios preferidos, más preferidos de cambios alternativos mostrados en la Tabla 3.
- 15 5. El polipéptido según la reivindicación 2, en el que la al menos una mutación es una seleccionada entre los cambios preferidos, más preferidos de cambios alternativos mostrados en la Tabla 4.
6. Un ácido nucleico que codifica una proteína de cualquiera de las reivindicaciones precedentes.
7. Un vector de expresión que comprende el ácido nucleico de la reivindicación 6.
8. Un microorganismo que contiene la construcción de vector de la reivindicación 7.
- 20 9. Una mezcla que contiene la proteína según una cualquiera de las reivindicaciones 1-5 y una o más enzima(s), adicional(es) seleccionada(s) preferentemente de una o más de celulasas, hemicelulasas y pectinasas
10. Uso de la proteína de una cualquiera de las reivindicaciones 1-5 o de la mezcla de la reivindicación 9 para la sacarificación de lignocelulosa.

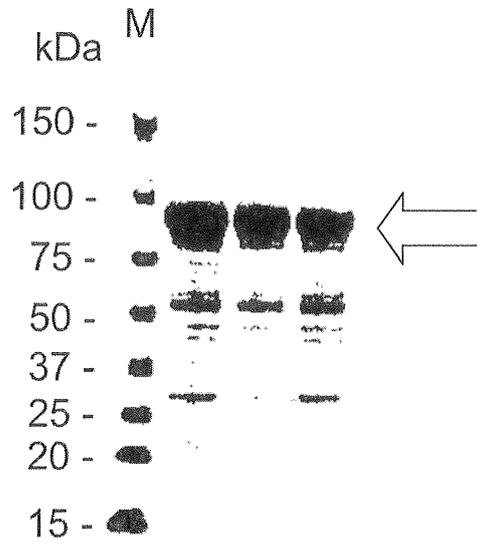


Figura 1

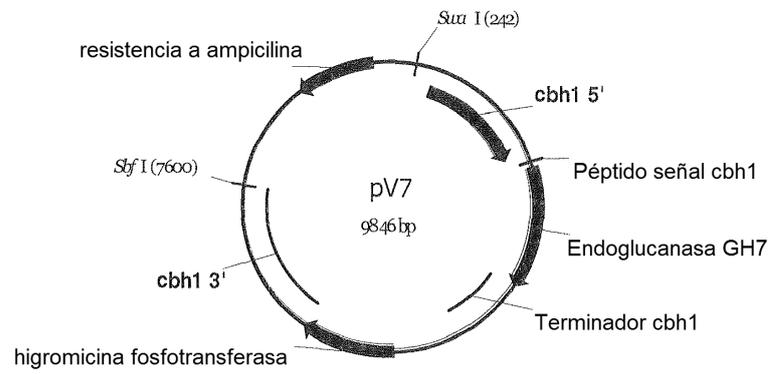


Figura 2

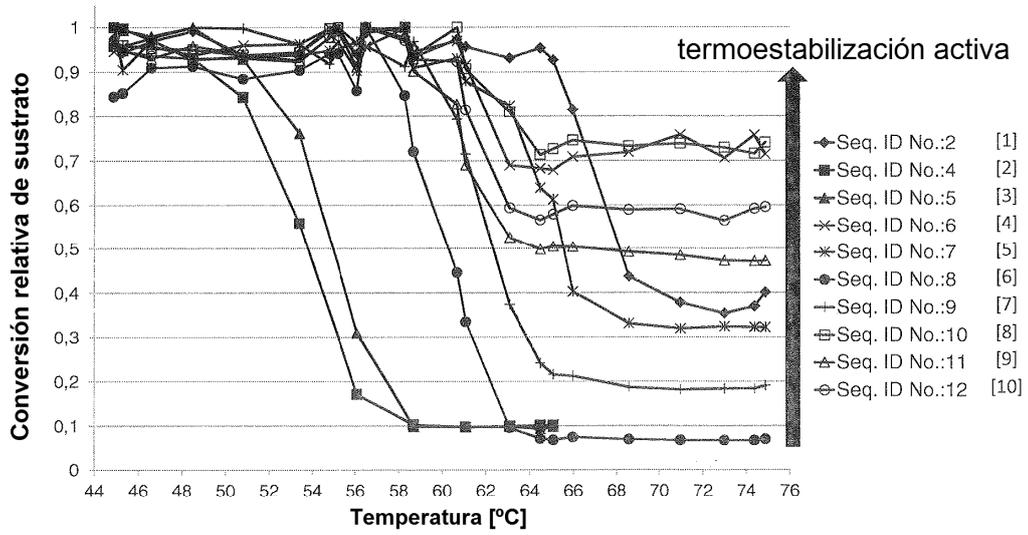


Figura 3

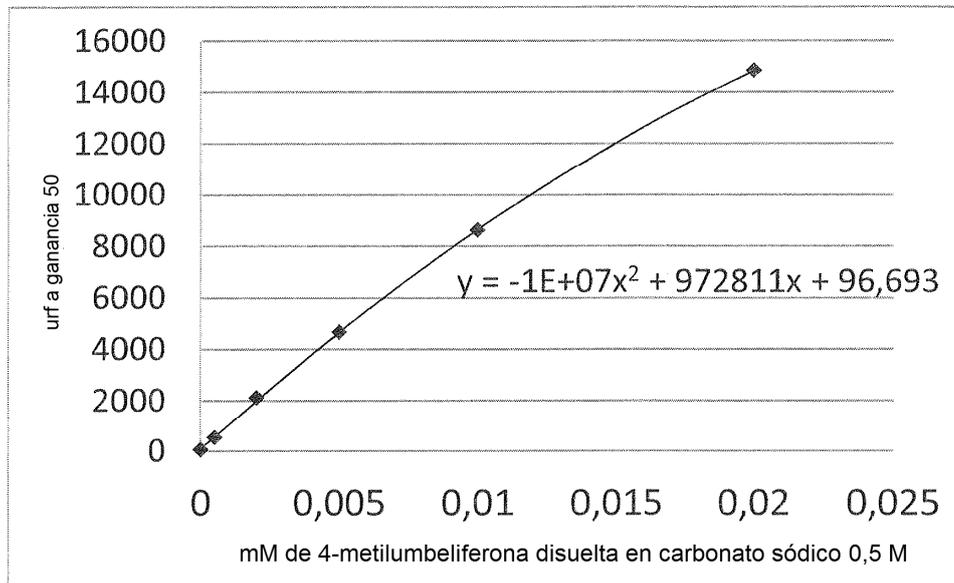


Figura 4

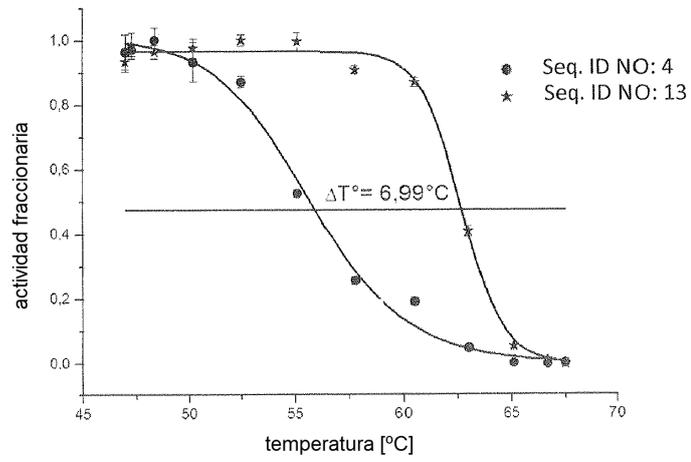


Figura 5

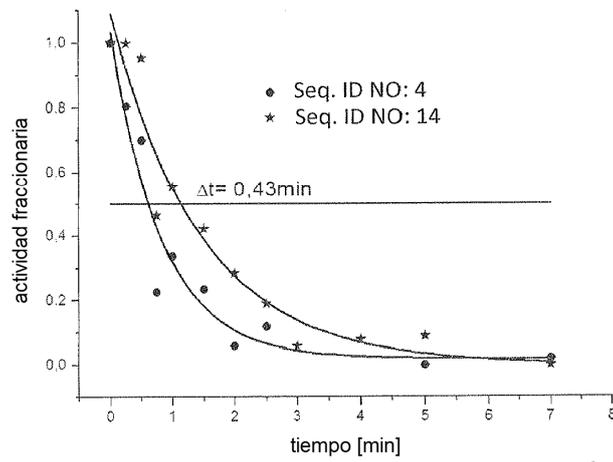


Figura 6