

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 502 743**

51 Int. Cl.:

**C12N 9/34** (2006.01)

**C12N 15/56** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **08.04.2008** **E 12157549 (2)**

97 Fecha y número de publicación de la concesión europea: **11.06.2014** **EP 2481796**

54 Título: **Variantes de glucoamilasa**

30 Prioridad:

**09.10.2007 WO PCT/US2007/021683**

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

**06.10.2014**

73 Titular/es:

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

72 Inventor/es:

**AEHLE, WOLFGANG;  
BOTT, RICHARD R.;  
NIKOLAEV, IGOR;  
SCHEFFERS, MARTIJN;  
VAN SOLINGEN, PIET y  
VROEMEN, CASPER**

74 Agente/Representante:

**RIZZO, Sergio**

**ES 2 502 743 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**

Variantes de glucoamilasa

**CAMPO DE LA INVENCION**

5 [0001] La presente invención hace referencia a variantes de glucoamilasa. En concreto, la invención hace referencia a variantes en el dominio de unión a almidón (SBD por sus siglas en inglés) de una glucoamilasa. La invención también hace referencia a variantes que presentan propiedades alteradas (p. ej., termoestabilidad mejorada y/o actividad específica aumentada) comparadas con una glucoamilasa precursora correspondiente. La presente invención también proporciona composiciones de enzima que comprenden glucoamilasas variantes; construcciones de ADN que comprenden polinucleótidos que codifican las variantes y métodos para producir las  
10 variantes de glucoamilasa en células huésped.

**ANTECEDENTES DE LA INVENCION**

15 [0002] Las enzimas de glucoamilasa (glucano 1,4- $\alpha$ -glucohidrolasa, EC 3.2.1.3) son carbohidrasas hidrolizantes de almidón que actúan en exo, que catalizan la eliminación de unidades de glucosa sucesivas de los extremos no reductores de moléculas relacionadas con oligosacáridos o polisacáridos o almidón. Las glucoamilasas pueden hidrolizar tanto los enlaces glucosídicos lineales como ramificados de almidón (p.ej., amilosa y amilopectina).

20 [0003] Las glucoamilasas se producen mediante numerosas cepas de bacterias, hongos, levaduras y plantas. Especialmente interesante, y comercialmente importante, las glucoamilasas son enzimas fúngicas que se producen de forma extracelular, por ejemplo, a partir de cepas de *Aspergillus* (Svensson *et al.* (1983) *Carlsberg Res. Commun.* 48:529 - 544; Boel *et al.*, (1984) *EMBO J.* 3:1097 - 1102; Hayashida *et al.*, (1989) *Agric. Biol. Chem.* 53:923 - 929; patente estadounidense 5.024.941; patente estadounidense 4.794.175 y WO 88/09795); *Talaromyces* (patente estadounidense 4.247.637; patente estadounidense 6.255.084 y patente estadounidense 6.620.924); *Rhizopus* (Ashikari *et al.*, (1986) *Agric. Biol. Chem.* 50:957 - 964; Ashikari *et al.*, (1989) *App. Microbiol. and Biotech.* 32:129 - 133 y patente estadounidense 4.863.864); *Humicola* (WO 05/052148 y patente estadounidense 4.618.579) y *Mucor* (Houghton-Larsen *et al.*, (2003) *Appl. Microbiol. Biotechnol.* 62:210 - 217). Muchos de los genes que codifican estas enzimas han sido clonados y expresados en células bacterianas, fúngicas y/o levadura.

30 [0004] Comercialmente, las glucoamilasas son enzimas muy importantes y se han utilizado en una amplia variedad de aplicaciones que requieren la hidrólisis de almidón (p. ej., para la producción de glucosa y otros monosacáridos a partir de almidón). Las glucoamilasas se utilizan para producir edulcorantes de maíz con alto contenido de fructosa, que comprenden más del 50 % del mercado de edulcorantes en Estados Unidos. En general, las glucoamilasas pueden utilizarse, y normalmente se utilizan, con alfa-amilasas en procesos hidrolizantes de almidón para hidrolizar almidón en dextrinas y posteriormente glucosa. La glucosa puede entonces convertirse en fructosa por parte de otras enzimas (p. ej., isomerasas de glucosa); cristalizada; o  
35 utilizada en fermentaciones con el fin de producir numerosos productos finales (p. ej., etanol, ácido cítrico, ácido láctico, succinato, intermediarios de ácido ascórbico, ácido glutámico, glicerol y 1,3-propanediol). El etanol producido mediante el uso de glucoamilasas en la fermentación de almidón y/o material que contiene celulosa puede utilizarse como una fuente de combustible o para consumo alcohólico.

40 [0005] Aunque las glucoamilasas se han utilizado con éxito en aplicaciones comerciales durante muchos años, todavía existe la necesidad de nuevas glucoamilasas con propiedades modificadas, tal como actividad específica mejorada y termoestabilidad aumentada.

45 [0006] WO 2003/029449 expone variantes de la glucoamilasa *Aspergillus niger* G1 que presenta termoestabilidad y/o actividad específica modificada. WO 2007/057018 expone las variantes de glucoamilasa *Talaromyces* and *Aspergillus*. WO 2006/060062 expone secuencias para la glucoamilasa *Trichoderma reesei* y para usarse en la conversión de almidón, pienso animal y fermentación alcohólica.

**SUMARIO DE LA INVENCION**

[0007] En un aspecto, la presente invención proporciona variantes de glucoamilasa que comprenden un dominio catalítico y un dominio de unión a almidón (SBD). El dominio catalítico comprende al menos un 85 % de identidad de secuencia con la secuencia de aminoácidos

SVDDFISTETPIALNLLCNVGPDGCRAFGTSAGAVIASPSTIDPDYYYMWTR  
 DSALVFKNLIDRFTETYDAGLQRRIEQYITAQVTLQGLSNPSGSLADGSGLGEP  
 KFELTLKPFTGNWGRPQRDGPALRAIALIGYSKWLINNNYQSTVSNVIWPIVR  
 NDLNVAQYWNQTGFDLWEEVNGSSFFTVANQHRALVEGATLAATLGQSG  
 SAYSSVAPQVLCFLQRFWVSSGGYVDSNINTNEGRTGKDVNSVLTSIHTFDPN  
 LGCDAGTFQPCSDKALSNLKVVVDSFRSIYGVNKGIPAGAAVAIGRYAEDVY  
 YNGNPWYLATFAAAEQLYDAIYVWKKTGSIPTVATSLAFFQELVPGVTAGTY  
 SSSSSTFTNIINAVSTYADGFLSEAAKYVPADGSLAEQFDRNSGTPLSALHLTW  
 SYASFLTATARRAGIVPPSWANSSASTIPSTC (SEQ ID NO:3)

y el SBD comprende la sustitución de aminoácidos N563I, N563C, N563E, N563A, N563K, N563L, N563Q, N563T o N563V en la secuencia de aminoácidos

N563L, N563Q, N563T or N563V9H, A539M or A539S in the amino acid sequence  
 SVDDFISTETPIALNLLCNVGPDGCRAFGTSAGAVIASPSTIDPDYYYMWTR  
 DSALVFKNLIDRFTETYDAGLQRRIEQYITAQVTLQGLSNPSGSLADGSGLGEP  
 KFELTLKPFTGNWGRPQRDGPALRAIALIGYSKWLINNNYQSTVSNVIWPIVR  
 NDLNVAQYWNQTGFDLWEEVNGSSFFTVANQHRALVEGATLAATLGQSG  
 SAYSSVAPQVLCFLQRFWVSSGGYVDSNINTNEGRTGKDVNSVLTSIHTFDPN  
 LGCDAGTFQPCSDKALSNLKVVVDSFRSIYGVNKGIPAGAAVAIGRYAEDVY  
 YNGNPWYLATFAAAEQLYDAIYVWKKTGSIPTVATSLAFFQELVPGVTAGTY  
 SSSSSTFTNIINAVSTYADGFLSEAAKYVPADGSLAEQFDRNSGTPLSALHLTW  
 SYASFLTATARRAGIVPPSWANSSASTIPSTCSGASVVGYSRPTATSFPPSQTP  
 KPGVPSGTPYTPLCATPTSVAVTFHELVSTQFGQTVKVAGNAAALGNWSTS  
 AAVALDAVNYADNHPLWIGTVNLEAGDVVEYKYINVGDGQSVTWESDPNH  
 TYTVPVAVACVTQVVKEDTWQS (SEQ ID NO: 2),

5 o una posición equivalente en un SBD de glucoamilasa precursora como se determina mediante la alineación de secuencia. Las variantes de glucoamilasa de la invención muestran termoestabilidad aumentada (medida como el % de actividad residual tras la incubación durante 1 hora a 64 grados centígrados en tampón NaAc pH 4,5) y/o actividad específica aumentada comparada con la glucoamilasa precursora.

10 **[0008]** En algunas formas de realización, las variantes de glucoamilasa presentan un dominio catalítico con al menos un 90 % o al menos un 95 % de identidad de secuencia con la secuencia de aminoácidos de SEQ ID NO:3.

**[0009]** En algunas formas de realización, la glucoamilasa precursora comprende un SBD que presenta al menos un 95 % de identidad de secuencia de aminoácidos con un SBD elegido de:

CATPTSVAVTFHELVSTQFGQTVKVAGNAAALGNWSTSAAVALDAVNYAD  
 NHPLWIGTVNLEAGDVVEYKYINVGDGQSVTWESDPNH  
 TYTVPVAVACVTQVVKEDTWQS (SEQ ID NO: 11);

CADASEVYVTFNERVSTAWGETIKVVGNPALGNWDTSKAVTLSASGYKSN  
DPLWSITVPIKATGSAVQYKYIKVGTNGKITWESDPNRSITLQTASSAGKCAA  
QTVNDSWR (SEQ ID NO:385);

CTPPSEVTLTFNALVDTAFGQNIYLVGSIPELGSWDPANALLMSAKSWTSGNP  
VWTLSSISLPAGTSFEYKFIRKDDGSSDVVWESDPNRSYNVPKDCGANTATVN  
SWWR (SEQ ID NO:386);

CTTPTSAVTFDEIVSTSYGETIYLAGSIPELGNWSTASAIPLRADAYTNSNPL  
WYVTVNLPPGTSFEYKFFKNQTDGTIVWEDDPNRSYTPAYCGQTTAILDDS  
WQ (SEQ ID NO:387);

CTTPTAVAVTFDLTATTTYGENIYLVGSISQLGDWETSDGIALSADKYTSSDPL  
WYVTVTLPAGESFEYKFIRIESDDSVWESDPNREYTPQACGTSTATVTDT  
WR (SEQ ID NO:388);

5 o

CTTPTAVAVTFDLTATTTYGENIYLVGSISQLGDWDTSDGIALSADKYTSSNP  
LWYVTVTLPAGESFEYKFIRIESDDSVWESDPNREYTPQACGESTATVTDT  
WR (SEQ ID NO:389).

10 **[0010]** En algunas formas de realización, las variantes de glucoamilasa comprenden además una o más sustituciones en las posiciones correspondientes a la posición 493, 494, 495, 501, 502, 503, 508, 511, 517, 518, 519, 520, 525, 527, 531, 533, 535, 536, 537, 538, 539, 540, 545, 546, 547, 549, 551, 561, 567, 569, 577, 579 o 583 de SEQ ID NO:2 o en una posición equivalente en la glucoamilasa precursora como se determina mediante la alineación de secuencia;

15 **[0011]** En algunas formas de realización, la variante de glucoamilasa de la reivindicación comprende una o más sustituciones de aminoácidos en una posición correspondiente a una posición de residuo: 10, 14, 15, 23, 42, 45, 46, 59, 60, 61, 67, 68, 72, 73, 97, 98, 99, 102, 108, 110, 113, 114, 122, 124, 125, 133, 140, 144, 145, 147, 152, 153, 164, 175, 182, 204, 205, 214, 216, 219, 228, 229, 230, 231, 236, 239, 240, 241, 242, 244, 263, 264, 265, 268, 269, 276, 284, 291, 300, 301, 303, 310, 311, 313, 316, 338, 342, 344, 346, 349, 359, 361, 364, 379, 382, 390, 391, 393, 394, 408, 410, 415, 417, 418, 430, 431, 433, 436, 442, 443, 444, 448 y 451 de SEQ ID NO:2 o SEQ ID NO:3.

**[0012]** En algunas formas de realización, la glucoamilasa precursora se obtiene de una *Trichoderma spp.*

20 **[0013]** La presente invención también proporciona polinucleótidos que codifican las variantes de glucoamilasa como se define en la reivindicación 1 y células huésped que comprenden dichos polinucleótidos.

25 **[0014]** También se proporcionan composiciones de enzima que comprenden variantes de glucoamilasa según se definen en la reivindicación 1, en la que la composición es adecuada para utilizarse en un proceso de conversión de almidón, una fórmula de pienso animal o un proceso de fermentación alcohólica que contiene almidón. En algunas formas de realización, la composición adecuada para utilizarse en un proceso de conversión de almidón comprende además un alfa-amilasa.

30 **[0015]** En otro aspecto, la presente invención proporciona un método para producir una glucoamilasa variante en una célula huésped, que comprende la transformación de una célula huésped con una construcción de ADN que comprende un polinucleótido que codifica una variante de glucoamilasa según se define en la reivindicación 1; el cultivo de la célula huésped en las condiciones adecuadas para la expresión y producción de la variante de glucoamilasa y la producción de la variante.

En algunas formas de realización, el método comprende además la recuperación de la glucoamilasa del cultivo.

**BREVE DESCRIPCIÓN DE LOS DIBUJOS**

**[0016]**

La fig. 1A ilustra una glucoamilasa *Trichoderma reesei* (TrGA) que presenta 632 aminoácidos (SEQ ID NO: 1) El péptido señal se subraya, la región catalítica (SEQ ID NO:3) que comienza con los residuos de aminoácidos SVDDFI (SEQ ID NO:12) y que presenta 453 residuos de aminoácidos está en negrita; la región de enlace está en cursiva y el dominio de unión a almidón (SBD) es tanto en cursiva como subrayado. La proteína madura que incluye el dominio catalítico (SEQ ID NO:3), la región de enlace (SEQ ID NO:10) y el dominio de unión a almidón (SEQ ID NO:11) se representan mediante SEQ ID NO:2. Con respecto a la numeración de SBD de la molécula de glucoamilasa TrGA, se hace referencia en la presente exposición tanto a) las posiciones 491 a 599 en SEQ ID NO:2 de la TrGA madura y/o posiciones 1 a 9 en SEQ ID NO:11 que representa la secuencia SBD aislada de la TrGA madura. Con respecto a la numeración del dominio catalítico de la molécula TrGA se hace referencia a la SEQ ID NO:2 y SEQ ID NO:3.

La fig. 1B ilustra el ADNc (SEQ ID NO:4) que codifica la TrGA. La fig. 1C ilustra los dominios de TrGA de proteína madura y precursora.

La fig. 2 ilustra el plásmido de destino pDONR-TrGA que incluye el ADNc (SEQ ID NO:4) de la TrGA.

La fig. 3 ilustra el plásmido PTTT-Dest.

La fig. 4 ilustra el vector de expresión final pTTT-TrGA.

Las figs. 5A-5B ilustran una comparación de alineación de los dominios catalíticos de las glucoamilasas precursoras que incluyen glucoamilasas derivadas de *Aspergillus awamori* (AaGA) (SEQ ID NO:5); *Aspergillus niger* (AnGA) (SEQ ID NO:6); *Aspergillus oryzae* (AoGA) (SEQ ID NO:7); *Trichoderma reesei* (TrGA) (SEQ ID NO:3); *Humicola grisea* (HgGA) (SEQ ID NO:8); e *Hypocrea vinosa* (HvGA) (SEQ ID NO:9). Los aminoácidos idénticos se indican con un asterisco (\*).

La fig. 5C ilustra una secuencia de proteína madura de glucoamilasa *Talaromyces* (TeGA) (SEQ ID NO:384).

Las figs. 5D-5E ilustran una alineación que compara el dominio de unión a almidón (SBD) de glucoamilasas precursoras entre las que se incluyen *Trichoderma reesei* (SEQ ID NO: 11), *Humicola grisea* (HgGA) (SEQ ID NO:385), *Thermomyces lanuginosus* (ThGA) (SEQ ID NO:386), *Talaromyces emersonii* (TeGA) (SEQ ID NO:387), *Aspergillus niger* (AnGA) (SEQ ID NO:388); y *Aspergillus awamori* (AaGA) (SEQ ID NO:389).

La fig. 6 es una comparación de las estructuras tridimensionales de glucoamilasa *Trichoderma* (negra) (SEQ ID NO:2) y la glucoamilasa *Aspergillus awamori* (gris) vistas desde el lateral. El lateral se mide en referencia al centro activo. Por ejemplo, en las figuras 6-8 la entrada al centro activo se define como "la parte superior" de la molécula.

La fig. 7 es una composición de las estructuras tridimensionales de glucoamilasa *Trichoderma* (negra) y *Aspergillus awamori* (gris) vistas desde la parte superior.

La fig. 8 es una alineación de las estructuras tridimensionales de TrGA (negra) y *A. niger* GA (gris) vistas desde el lado que muestra el lugar de unión 1 y 2.

La fig. 9 es un modelo de la unión de acarbosa a la estructura de cristal de TrGA mostrada en la figura 6.

**DESCRIPCIÓN DETALLADA DE LA INVENCION****I. Definiciones**

**[0017]** A menos que se defina lo contrario, todos los términos técnicos y científicos utilizados en el presente documento tienen el mismo significado según lo entiende comúnmente alguien especializado en la técnica a la que esta invención pertenece. Singleton, *et al.*, *DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY*, 2D ED., John Wiley and Sons, New York (1994), y Hale & Markham, *THE HARPER COLLINS DICTIONARY OF BIOLOGY*, Harper Perennial, N.Y. (1991) proporcionan al experto el significado general de muchos de los términos empleados en el presente documento. Aún así, determinados términos se definen a continuación para una mayor claridad y facilidad de referencia.

**[0018]** El término "glucoamilasa (EC 3.2.1.3)" se refiere en el presente documento a una enzima que cataliza la liberación de D-glucosa desde los extremos no reductores de almidón y polisacáridos y oligosacáridos relacionados.

**[0019]** El término "precursor" o "secuencia precursora" hace referencia a una secuencia natural o que se da de forma natural o una secuencia de referencia que presenta una secuencia y/o identidad estructural con TrGA (SEQ ID NOS: 1 y/o 2).

**[0020]** El término "TrGA" hace referencia a una secuencia de glucoamilasa *Trichoderma reesei* precursora que presenta la secuencia de proteína madura ilustrada en SEQ ID NO:2 que incluye el dominio catalítico que tiene la secuencia ilustrada en SEQ ID NO:3. El aislamiento, clonación y expresión de la TrGA se describen en WO 2006/060062 y la publicación de patente estadounidense nº 2006/0094080 publicada el 4 de mayo, 2006 que se incorporan al presente documento como referencia. La TrGA también se considera una secuencia de glucoamilasa precursora. En algunas formas de realización, la secuencia precursora hace referencia a una TrGA que es el punto inicial para la ingeniería de las proteínas.

**[0021]** La frase "forma madura de una proteína o polipéptido" hace referencia a la forma funcional final de la proteína o polipéptido. Con el fin de ejemplificarlo, una forma madura de TrGA incluye el dominio catalítico, la región de enlace y el dominio de unión a almidón que presenta la secuencia de aminoácidos de SEQ ID NO:2.

**[0022]** El término "homólogos de glucoamilasa *Trichoderma*" hace referencia a glucoamilasas que presentan al menos un 50 % de identidad de secuencia, al menos un 60 % de identidad de secuencia, al menos un 70 % de identidad de secuencia, al menos un 80 % de identidad de secuencia con la secuencia TrGA (SEQ ID NO:2) y que retienen las características funcionales de una glucoamilasa.

**[0023]** El término una "secuencia homóloga" se refiere en el presente documento a una secuencia de ácido nucleico o de polipéptidos que presenta al menos un 100 %, al menos un 99 %, al menos un 98 %, al menos un 97 %, al menos un 96 %, al menos un 95 %, al menos un 94 %, al menos un 93 %, al menos un 92 %, al menos un 91 %, al menos un 90 %, al menos un 88 %, al menos un 85 %, al menos un 80 %, al menos un 75 %, al menos un 70 %, al menos un 65 %, al menos un 60 %, al menos un 55 %, al menos un 50 % o al menos un 45 % de identidad de secuencia con una secuencia de ácido nucleico o una secuencia de polipéptidos cuando se alinea de forma óptima para compararla, en la que la función de la secuencia de ácido nucleico candidata o secuencia de polipéptidos es básicamente la misma que la secuencia de ácido nucleico o secuencia de polipéptidos con la que se compara dicha secuencia homóloga candidata. En algunas formas de realización, las secuencias homólogas presentan entre un 85 % y un 100 % de identidad de secuencia, mientras que en otras formas de realización existe entre un 90 % y 100 % de identidad de secuencia y en otras formas de realización, existe un 95 % y un 100 % de identidad de secuencia. En algunas formas de realización, la secuencia homóloga candidata (p. ej., la secuencia de referencia) o precursora se compara con la secuencia de ácido nucleico TrGA o la secuencia de proteína madura. La identidad de secuencia puede medirse sobre toda la longitud de la secuencia precursora u homóloga.

**[0024]** Los términos "variante de glucoamilasa", "variante con SBD" y "variante de TrGA" se emplean el presente documento en referencia a glucoamilasas que son similares a una secuencia de glucoamilasa precursora o de referencia (p. ej., homólogos de glucoamilasa TrGA o *Trichoderma*) pero presentan al menos una sustitución, delección o inserción en la secuencia de aminoácidos del SBD que hace que sean diferentes en la secuencia de una glucoamilasa precursora o de referencia.

**[0025]** El término "dominio catalítico" hace referencia en el presente documento a una región estructural de un polipéptido que contiene el centro activo para la hidrólisis del sustrato.

**[0026]** El término "enlazador" se refiere a una secuencia de aminoácidos que presenta generalmente entre 3 y 40 residuos de aminoácidos que une de forma covalente una secuencia de aminoácidos que comprende un dominio de unión a almidón con una secuencia de aminoácidos que comprende un dominio catalítico.

**[0027]** El término "dominio de unión a almidón (SBD)" se refiere a una secuencia de aminoácidos que se une preferentemente a un sustrato de almidón.

**[0028]** Los términos "secuencia mutante" y "gen mutante" se utilizan en el presente documento indistintamente y hacen referencia a una secuencia de polinucleótidos que presenta una alteración en al menos un codón que se da en la secuencia original de una célula huésped. El producto de expresión de la secuencia mutante es una proteína variante con una secuencia de aminoácidos modificada en relación con la precursora. El producto de expresión puede tener una capacidad funcional modificada (p. ej., actividad enzimática mejorada).

**[0029]** El término "propiedad" o equivalentes gramaticales del mismo en el contexto de un polipéptido se refiere en el presente documento a cualquier característica o atributo de un polipéptido que puede elegirse o detectarse. Entre estas propiedades se incluyen, sin carácter limitativo, estabilidad oxidativa, especificidad de sustrato, actividad catalítica, estabilidad térmica, perfil de actividad de pH, resistencia a degradación proteolítica, KM, K<sub>cat</sub>, relación K<sub>cat</sub>/KM, plegamiento de proteínas, capacidad de unirse a un sustrato y capacidad de ser secretado.

**[0030]** El término "propiedad" o equivalentes gramaticales del mismo en el contexto de un ácido nucleico se refiere en el presente documento a cualquier característica o atributo de un ácido nucleico que pueda elegirse o detectarse. Entre estas propiedades se incluyen, sin carácter limitativo, una propiedad que afecta a la transcripción de genes (p. ej., fuerza del promotor o reconocimiento del promotor), una propiedad que afecta al procesamiento de ARN (p. ej., corte y empalme de ARN y estabilidad de ARN), una propiedad que afecta a la traducción (p. ej., regulación, unión de ARNm a proteínas ribosómicas).

**[0031]** El término "actividad específica" se define como la actividad por mg de proteína de glucoamilasa activa. La actividad se determina utilizando el análisis de etanol como se describe en el presente documento. Una variante identificada con un índice de rendimiento (PI por sus siglas en inglés) > 1,0 en comparación con el PI de la TrGA precursora se considera que presenta actividad específica aumentada. El PI se calcula a partir de las

actividades específicas (actividad/mg enzima) de la glucoamilasa variante y precursora (WT o tipo silvestre) Es el cociente "actividad específica variante/actividad específica WT".

5 **[0032]** Los términos "estable térmicamente" y "termoestable" hacen referencia a variantes de glucoamilasa de la presente invención que retienen una cantidad específica de actividad enzimática tras la exposición a temperaturas identificadas sobre un periodo de tiempo determinado en condiciones que se mantienen durante la hidrólisis de sustratos de almidón, por ejemplo, mientras se exponen a temperaturas modificadas.

**[0033]** El término "estabilidad mejorada" en el contexto de una propiedad tal como termoestabilidad se refiere a una mayor actividad hidrolítica de almidón retenido con el tiempo en comparación con otra glucoamilasa de referencia (p. ej., glucoamilasa precursora).

10 **[0034]** El término "estabilidad disminuida" en el contexto de una propiedad tal como termoestabilidad se refiere a una menor actividad hidrolítica de almidón retenido con el tiempo en comparación con otras glucoamilasas, variantes y/o glucoamilasas de tipo silvestre.

15 **[0035]** Los términos "activo" y "biológicamente activo" se refieren a una actividad biológica asociada con una proteína específica. Por consiguiente, la actividad biológica de una proteína determinada se refiere a cualquier actividad biológica típicamente atribuida a aquella proteína por los expertos en la técnica. Por ejemplo, una actividad enzimática asociada a una glucoamilasa es hidrolítica y, por tanto, una glucoamilasa activa presenta actividad hidrolítica.

20 **[0036]** Los términos "polinucleótidos" y "ácido nucleico", utilizados indistintamente en el presente documento, se refieren a una forma polimérica de nucleótidos de cualquier longitud, ya sean ribonucleótidos o desoxirribonucleótidos. Entre estos términos se incluyen, aunque sin carácter limitativo, ADN de cadena simple, doble o triple, ADN genómico, ADNc, ARN, híbrido ADN-ARN o un polímero que comprenda bases de pirimidina y purina u otras bases de nucleótidos naturales, modificadas de forma química, bioquímica, no naturales o derivatizadas.

25 **[0037]** Los términos "construcción de ADN, "ADN transformante" y "vector de expresión" se usan indistintamente para referirse al ADN con el fin de introducir secuencias en una célula huésped u organismo. El ADN puede generarse *in vitro* mediante PCR o cualquier otra técnica adecuada conocida por los expertos en la técnica. La construcción de ADN, el ADN transformante o el casete de expresión recombinante pueden incorporarse en un plásmido, cromosoma, ADN mitocondrial, ADN plastidial, virus o fragmento de ácido nucleico. Normalmente, la parte del casete de expresión recombinante de un vector de expresión, construcción de ADN o ADN transformante incluye, entre otras secuencias, una secuencia de ácido nucleico que ha de transcribirse y un promotor. En algunas formas de realización, los vectores de expresión tienen la capacidad de incorporar y expresar fragmentos de ADN heterólogos en una célula huésped.

30

35 **[0038]** El término "vector" se refiere en el presente documento a una construcción de polinucleótidos diseñada para introducir ácidos nucleicos en uno o más tipos de células. Entre los vectores se incluyen vectores de clonación, vectores de expresión, vectores transportadores, plásmidos, casetes y similares.

**[0039]** Tal y como se emplea en el presente documento en el contexto de introducción de una secuencia de ácido nucleico en una célula, el término "introducido" hace referencia a cualquier método adecuado para transferir la secuencia de ácido nucleico en la célula. Tales métodos para la introducción incluyen sin carácter limitativo fusión de protoplasto, transfección, transformación, conjugación y transducción.

40 **[0040]** Los términos "transformado" y "transformado de forma estable" hacen referencia en el presente documento a una célula que tiene una secuencia de polinucleótidos no natural (heteróloga) integrada en su genoma o como un plásmido episomal que se mantiene durante al menos dos generaciones.

45 **[0041]** Los términos "marcador seleccionable" y "marcador selectivo" hacen referencia en el presente documento a un ácido nucleico (p. ej., un gen) capaz de expresión en células huésped que permite la facilidad de selección de aquellos huéspedes que contienen el vector. Normalmente, los marcadores seleccionables son genes que conceden resistencia antimicrobiana o una ventaja metabólica en la célula huésped con el fin de permitir que las células que contienen el ADN exógeno se distingan de las células que no han recibido ninguna secuencia exógena durante la transformación.

50 **[0042]** El término "promotor" hace referencia en el presente documento a una secuencia de ácido nucleico que funciona para dirigir la transcripción de un gen aguas abajo. El promotor, junto con otras secuencias de ácido nucleico reguladoras de traducción y transcripción (también denominadas "secuencias de control"), es necesario para expresar un gen determinado. En general, entre las secuencias reguladoras de traducción y transcripción se incluyen, sin carácter limitativo, secuencias promotoras, lugares de unión ribosómicos, secuencias de transcripción de inicio y parada, secuencias de traducción de inicio y de parada y secuencias potenciadoras o

activadoras.

- 5 **[0043]** Un ácido nucleico está "unido de forma operativa" cuando se coloca en una relación funcional con otra secuencia de ácido nucleico. Por ejemplo, ADN que codifica un líder secretor (es decir, un péptido señal), está unido de forma operativa a un ADN para un polipéptido si se expresa como una preproteína que participa en la secreción del polipéptido. Generalmente, "unido de forma operativa" indica que las secuencias de ADN que están unidas son contiguas y, en el caso de un líder secretor, contiguas y en fase de lectura.
- [0044]** El término "gen" se refiere en el presente documento a un polinucleótido (p. ej., un segmento de ADN) que codifica un polipéptido e incluye regiones que preceden y siguen a las regiones codificantes así como secuencias intermedias (intrones) entre segmentos codificantes individuales (exones).
- 10 **[0045]** El término "genes homólogos" se refiere en el presente documento a un par de genes de diferentes especies pero normalmente relacionadas, que se corresponden una a la otra y que son idénticas o muy similares una a la otra. El término abarca genes que se separan mediante especiación (es decir, el desarrollo de nuevas especies) (p. ej., genes ortólogos), así como genes que se han separado mediante duplicación genética (p. ej., genes parálogos).
- 15 **[0046]** Los términos "ortólogo" y "genes ortólogos" se refieren en el presente documento a genes en diferentes especies que han evolucionado de un gen ancestral común (es decir, un gen homólogo) mediante especiación. Normalmente, los ortólogos conservan la misma función durante el transcurso de la evolución. La identificación de ortólogos es útil a la hora de predecir de forma fiable la función del gen en genomas secuenciados recientemente.
- 20 **[0047]** Los términos "parálogo" y "genes parálogos" se refieren en el presente documento a genes que están relacionados por duplicación dentro de un genoma. Mientras que los ortólogos conservan la misma función durante el transcurso de la evolución, los parálogos desarrollan nuevas funciones, aunque algunas funciones están relacionadas normalmente con la original. Entre los ejemplos de genes parálogos se incluyen, sin carácter limitativo, genes que codifican tripsina, quimotripsina, elastasa y trombina, que son todas serina proteasas y se dan de forma conjunta en las mismas especies.
- 25 **[0048]** El término "homología" se refiere en el presente documento a identidad o similaridad de secuencia, prefiriéndose la identidad. Esta homología se determina mediante las técnicas convencionales conocidas en la técnica (véase p. ej., Smith y Waterman, *Adv. Appl. Math.*, 2:482 [1981]; Needleman y Wunsch, *J. Mol. Biol.*, 48:443 [1970]; Pearson and Lipman, *Proc. Natl. Acad. Sci. EEUU* 85:2444 [1988]; programas como GAP, BESTFIT, FASTA y TFASTA en Wisconsin Genetics Software Package (Genetics Computer Group, Madison, WI); y Devereux *et al.*, *Nucl. Acid Res.*, 12:387-395 [1984]).
- 30 **[0049]** El "porcentaje (%) de identidad de secuencia de ácido nucleico" o "porcentaje (%) de identidad de secuencia de aminoácidos" se define como el porcentaje de residuos de nucleótidos o residuos de aminoácidos en una secuencia candidata que son idénticos a los residuos de nucleótidos o residuos de aminoácidos de la secuencia inicial (es decir, TrGA).
- 35 **[0050]** Las secuencias homólogas se determinan mediante métodos conocidos de alineación de secuencias. Un método de alineación utilizado normalmente es BLAST descrito por Altschul *et al.*, (Altschul *et al.*, (1990) *J. Mol. Biol.*, 215:403-410.; y Karlin *et al.*, (1993) *Proc. Natl. Acad. Sci. EEUU* 90:5873-5787). Un programa BLAST especialmente útil es el programa WU-BLAST-2 (Véase, Altschul *et al.*, (1996) *Meth. Enzymol.*, 266:460-480). WU-BLAST-2 utiliza diferentes parámetros de búsqueda, la mayoría de los cuales están establecidos a los valores por defecto. Los parámetros regulables están configurados con los siguientes valores: periodo de solapamiento = 1, fracción de solapamiento = 0,125, umbral de palabras (T) = 11. Los parámetros HSP S y HSP S2 son valores dinámicos y los establece el propio programa dependiendo de la composición de la secuencia específica y la composición de la base de datos específica contra la que se busca la secuencia de interés. Sin embargo, los valores pueden regularse con el fin de aumentar la sensibilidad. Un valor de % de identidad de secuencia de aminoácidos se determina por el número de residuos idénticos que coinciden dividido por el número total de residuos de la secuencia "más larga" en la región alineada. La secuencia "más larga" es la que presenta los residuos más reales en la región alineada (se ignoran los huecos introducidos por WU-Blast-2 con el fin de maximizar la puntuación de alineación).
- 40 **[0051]** Otros métodos pueden ser útiles para alinear secuencias. Un ejemplo de un algoritmo útil es PILEUP. PILEUP crea una alineación de secuencias múltiple de un grupo de secuencias relacionadas utilizando alineaciones por pares progresivas. También puede trazar un árbol que muestre las relaciones de agrupamiento utilizadas para crear la alineación. PILEUP utiliza una simplificación del método de alineación progresiva de Feng y Doolittle (Feng y Doolittle, *J. Mol. Evol.*, 35:351-360 [1987]). El método es similar al descrito por Higgins y Sharp (Higgins y Sharp, *CABIOS* 5:151-153 [1989]). Parámetros útiles de PILEUP incluyen un peso del hueco por defecto de 3,00, un peso de la longitud del hueco por defecto de 0,10 y huecos de extremo ponderados.
- 50 **[0051]** Otros métodos pueden ser útiles para alinear secuencias. Un ejemplo de un algoritmo útil es PILEUP. PILEUP crea una alineación de secuencias múltiple de un grupo de secuencias relacionadas utilizando alineaciones por pares progresivas. También puede trazar un árbol que muestre las relaciones de agrupamiento utilizadas para crear la alineación. PILEUP utiliza una simplificación del método de alineación progresiva de Feng y Doolittle (Feng y Doolittle, *J. Mol. Evol.*, 35:351-360 [1987]). El método es similar al descrito por Higgins y Sharp (Higgins y Sharp, *CABIOS* 5:151-153 [1989]). Parámetros útiles de PILEUP incluyen un peso del hueco por defecto de 3,00, un peso de la longitud del hueco por defecto de 0,10 y huecos de extremo ponderados.
- 55



**[0052]** El término "alineación óptima" hace referencia a la alineación que proporciona el mayor porcentaje de puntuación de identidad.

**[0053]** Una "posición equivalente" se refiere a una alineación entre dos secuencias en la que la alineación es óptima. Por ejemplo, utilizando las figs. 5D y 5E, la posición 491 en TrGA (SEQ ID NO:2) es C491; la posición equivalente para *Aspergillus niger* es la posición C509; y la posición equivalente para *Aspergillus awamori* es la posición Q538. Véase la figura 8 para una alineación de ejemplo de la secuencia tridimensional.

**[0054]** El término "hibridación" se refiere en el presente documento al proceso por el que una hebra de ácido nucleico se une con una hebra complementaria a través del apareamiento de bases, como se conoce en la técnica.

**[0055]** Se considera que una secuencia de ácido nucleico es "hibridable de forma selectiva" en referencia a una secuencia de ácido nucleico si dos secuencias se hibridan una a la otra de forma específica bajo condiciones de lavado e hibridación con una astringencia de moderada a alta. Las condiciones de hibridación se basan en la temperatura de fusión ( $T_m$ ) de la sonda o complejo que une ácido nucleico. Por ejemplo, la "astringencia máxima" normalmente se da a aproximadamente  $T_m - 5$  °C (5° por debajo de la  $T_m$  de la sonda); "astringencia alta" a aproximadamente 5-10 °C por debajo de la  $T_m$ ; "astringencia intermedia" a aproximadamente 10-20 °C por debajo de la  $T_m$  de la sonda y "baja astringencia" a aproximadamente 20-25°C por debajo de la  $T_m$ . De forma funcional, pueden utilizarse las condiciones de astringencia máxima para identificar secuencias que tengan identidad precisa o casi precisa con la sonda de hibridación; mientras que se puede utilizar una hibridación de astringencia intermedia o baja para identificar o detectar homólogos de secuencia de polinucleótidos.

**[0056]** Las condiciones de hibridación de astringencia moderada y alta son ya conocidas en la técnica. Un ejemplo de condiciones astringentes altas incluye hibridación a aproximadamente 42 °C en formamida 50 %, 5X SSC, 5X solución Denhardt, 0,5 % SDS y 100 µg/ml de ADN portador desnaturalizado seguido de dos lavados en 2X SSC y 0,5 % SDS a temperatura ambiente y dos veces adicionales en 0,1X SSC y 0,5 % SDS a 42 °C. Un ejemplo de condiciones astringentes moderadas incluye una incubación durante la noche a 37 °C en una solución que comprende formamida 20 %, 5X SSC (150 mM NaCl, 15 mM citrato trisódico), 50 mM fosfato de sodio (pH 7,6), 5 x solución Denhardt, sulfato de dextrano 10 % y 20 mg/ml ADN de esperma de salmón cortado desnaturalizado seguido del lavado de los filtros en 1x SSC a aproximadamente 37 - 50 °C. Aquellos expertos en la técnica saben cómo regular la temperatura, la fuerza iónica, etc. según sea necesario adaptarse a factores tales como la longitud de sonda y similares.

**[0057]** El término "recombinante" en el presente documento incluye referencia a una célula o vector que ha sido modificado mediante la introducción de una secuencia de ácido nucleico homóloga o heteróloga o que la célula se ha derivado de una célula modificada de ese modo. Por lo tanto, por ejemplo, las células recombinantes expresan genes que no se hallan en forma idéntica dentro de la forma natural (no recombinante) de la célula o expresan genes naturales que se han expresado de otro modo de forma anormal, subexpresados o no expresados en general como resultado de una intervención humana deliberada.

**[0058]** En algunas formas de realización de la presente exposición, las secuencias de ADN mutado se generan con mutagénesis por saturación del lugar en al menos un codón. En otra forma de realización, la mutagénesis por saturación del lugar se lleva a cabo en dos o más codones. En otra forma de realización adicional, las secuencias de ADN mutante presentan más del 50 %, más del 55 %, más del 60 %, más del 65 %, más del 70 %, más del 75 %, más del 80 %, más del 85 %, más del 90 %, más del 95 %, más del 98 % o más del 99 % de homología con la secuencia precursora. En formas de realización alternativas, el ADN mutante se genera *in vivo* utilizando cualquier procedimiento mutagénico tal como, por ejemplo, radiación, nitrosoguanidina y similares. A continuación, la secuencia de ADN deseada se aísla y se utiliza en los métodos proporcionados en el presente documento.

**[0059]** El término "proteína heteróloga" se refiere en el presente documento a una proteína o polipéptido que no se da de forma natural en una célula huésped.

**[0060]** El término "proteína homóloga" se refiere en el presente documento a una proteína o polipéptido natural o que se da de forma natural en una célula e incluye proteínas naturales que están sobreexpresadas en la célula ya sea mediante tecnología de ADN recombinante o de forma natural.

**[0061]** Una enzima está "sobreexpresada" en una célula huésped si la enzima se expresa en una célula con un nivel mayor que el nivel al que se expresa en una célula de tipo silvestre correspondiente.

**[0062]** Los términos "proteína" y "polipéptido" se utilizan indistintamente en el presente documento.

En la presente exposición y reivindicaciones, se utilizan los códigos de una letra y tres letras convencionales para residuos de aminoácidos. Los códigos de tres letras para aminoácidos se definen según la Comisión sobre

nomenclatura bioquímica IUPAC-IUB (JCBN). También se entiende que un polipéptido puede codificarse mediante más de una secuencia de nucleótidos debido a la degeneración del código genético.

5 **[0063]** Se describen variantes de la presente exposición mediante la siguiente nomenclatura: [residuo aminoácido original/posición/residuo aminoácido sustituido]. Por ejemplo, la sustitución de leucina por arginina en la posición 76 se representa como R76L. Cuando se sustituye más de un aminoácido en una determinada posición, la sustitución se representa como 1) Q172C, Q172D o Q172R; 2) Q172C, D, o R o c) Q172C/D/R. Cuando se identifica en la misma una posición adecuada para la sustitución sin un aminoácido específico sugerido, se ha de entender que cualquier residuo de aminoácido puede sustituirse por el residuo de aminoácido presente en la posición. Cuando una glucoamilasa variante contiene una delección en comparación con otras glucoamilasas la delección se indica con "\*\*". Por ejemplo, una delección en la posición R76 se representa como R76\*. Una delección de dos o más aminoácidos consecutivos se indica por ejemplo como (76-78)\*.

**[0064]** Una "prosecuencia" es una secuencia de aminoácidos entre la secuencia señal y la proteína madura que es necesaria para la secreción de la proteína. La escisión de la prosecuencia tendrá como resultado una proteína activa madura.

15 **[0065]** El término "secuencia señal" o "péptido señal" se refiere a cualquier secuencia de nucleótidos y/o aminoácidos que puedan participar en la secreción de las formas precursoras o maduras de la proteína. Esta definición de secuencia señal es una funcional, que pretende incluir todas las secuencias de aminoácidos codificadas por la parte N-terminal del gen de la proteína, que participan en la realización de la secreción de proteína. Normalmente, aunque no de forma universal, están ligados a la parte N-terminal de una proteína o a la parte N-terminal de una proteína precursora. La secuencia señal puede ser endógena o exógena. La secuencia señal puede estar asociada normalmente a la proteína (p. ej., glucoamilasa) o puede ser de un gen que codifica otra proteína secretada.

20 **[0066]** El término forma "precursora" de una proteína o péptido se refiere a la forma madura de la proteína que tiene una prosecuencia unida de forma operativa al carbonilo o amino-terminal de la proteína. El precursor puede también tener una secuencia "señal" unida de forma operativa al amino-terminal de la prosecuencia. El precursor puede también presentar polipéptidos adicionales que estén involucrados en la actividad traduccional posterior (p. ej., polipéptidos escindidos de este para dejar la forma madura de una proteína o péptido).

**[0067]** Los términos "cepa huésped" o "célula huésped" se refieren a un huésped adecuado para un vector de expresión que comprende ADN según la presente invención.

30 **[0068]** Los términos "derivado de" y "obtenido de" se refieren no solo a una glucoamilasa producida o que se puede producir mediante una cepa del organismo en cuestión, sino también a una glucoamilasa codificada por una secuencia de ADN aislada de tal cepa y producida en un organismo huésped que contiene dicha secuencia de ADN. Además, el término se refiere a una glucoamilasa que está codificada por una secuencia de ADN de origen sintético y/o ADNc y que tiene las características de identificación de la glucoamilasa en cuestión.

35 **[0069]** Un "derivado" dentro del alcance de esta definición generalmente conserva la característica actividad hidrolizante observada en la forma de tipo silvestre, natural o precursora hasta el punto de que el derivado es útil para fines similares como la forma precursora, natural o de tipo silvestre. Los derivados funcionales de glucoamilasas abarcan péptidos o fragmentos de péptidos que se dan de forma natural, se producen de forma sintética o recombinante que presentan las características generales de las glucoamilasas de la presente invención.

40 **[0070]** Los términos "aislado" o "purificado" se refieren a un material que se ha extraído de su medio original (p. ej., el medio natural si se da de forma natural). En algunas formas de realización, una proteína aislada presenta más de un 10 % de pureza, preferiblemente más de un 20 % de pureza e incluso más preferiblemente más de un 30 % de pureza, según se determina mediante SDS-PAGE. Aspectos adicionales de la presente exposición incluyen la proteína en forma altamente purificada (es decir, más de un 40 % de pureza, más de un 60 % de pureza, más de un 80 % de pureza, más de un 90 % de pureza, más de un 95 % de pureza, más de un 97 % de pureza y más de un 99 % de pureza) según se determina mediante SDS-PAGE.

45 **[0071]** El término "mutagénesis combinatoria" hace referencia en el presente documento a métodos en los que se generan las bibliotecas de variantes de una secuencia inicial. En estas bibliotecas, las variantes contienen una o más mutaciones elegidas de un conjunto de mutaciones predefinidas. Además, los métodos proporcionan medios para introducir mutaciones aleatorias que no eran elementos del conjunto de mutaciones predefinido. En algunas formas de realización, los métodos incluyen aquellos descritos en la patente estadounidense 6.582.914. En formas de realización alternativas, los métodos de mutagénesis combinatoria incluyen kits disponibles comercialmente (p. ej., Quik-Change® Multisite, Stratagene, San Diego, CA).

55 **[0072]** El término "biblioteca de mutantes" hace referencia en el presente documento a una población de células

que son idénticas en la mayoría de su genoma pero incluyen diferentes homólogos de uno o más genes. Tales bibliotecas pueden utilizarse, por ejemplo, para identificar genes u operones con rasgos mejorados.

**[0073]** El término "contenido de sólidos secos (DS o ds por sus siglas en inglés)" hace referencia en el presente documento a los sólidos totales de una suspensión en % sobre una base de peso en seco.

5 **[0074]** El término "propiedad diana" hace referencia en el presente documento a la propiedad del gen inicial que ha de modificarse. No se pretende que la presente invención esté limitada por cualquier propiedad diana específica. Sin embargo, en algunas formas de realización, la propiedad diana es la estabilidad de un producto génico (p. ej., resistencia a la desnaturalización, proteólisis u otros factores de degradación), mientras que en  
10 otras formas de realización, el nivel de producción en un huésped de producción se modifica. De hecho, se contempla que cualquier propiedad de un gen inicial será útil en la presente invención. Pueden aparecer otras definiciones de los términos en toda la memoria.

**[0075]** Antes de describir las formas de realización de ejemplo con más detalle, ha de entenderse que la presente invención no está limitada a las formas de realización específicas descritas, ya que tales pueden, obviamente, variar. También ha de entenderse que la terminología empleada en el presente documento tiene el  
15 fin de describir formas de realización específicas únicamente y no pretenden poseer carácter limitativo.

**[0076]** En los casos en los que se presente una gama de valores, ha de entenderse que también se indica de forma específica cada valor intermedio, a la décima parte de la unidad del límite inferior a menos que el contexto indique claramente lo contrario, entre los límites superior e inferior de cada gama. Cada gama menor entre  
20 cualquier valor indicado o valor intermedio en una gama indicada y cualquier otro valor indicado o intermedio en esa gama indicada se incluye dentro de la invención. Los límites superior e inferior de estas gamas menores pueden incluirse o excluirse de forma independiente en la gama, y cada gama en la que cualquier límite, ambos o ninguno se incluyan en las gamas menores también se incluye dentro de la invención, sujeto a cualquier límite excluido de forma específica en la gama indicada. En los casos en los que la gama indicada incluya uno o ambos límites, las gamas que excluyan alguno o ambos de esos límites incluidos también se incluyen en la invención.

25 **[0077]** Cabe destacar que tal y como se utilizan en el presente documento y en las reivindicaciones adjuntas, las formas singulares "un/a" y "el/la" incluyen referentes plurales a menos que el contexto indique claramente lo contrario. Por lo tanto, por ejemplo, la referencia a "un gen" incluye una pluralidad de tales agentes candidatos y la referencia a "la célula" incluye referencia a una o más células y equivalentes de estos conocidos por los expertos en la técnica, etc.

30 **[0078]** Las publicaciones detalladas en el presente documento se presentan únicamente para su exposición anterior a la fecha de presentación de la presente solicitud. Nada de lo que aparece en el presente documento ha de entenderse como una admisión de que la presente invención no tiene derecho a preceder a tal publicación en virtud de la invención anterior.

## II. Formas de realización

### 35 **Glucoamilasas precursoras:**

**[0079]** En algunas formas de realización, la presente invención proporciona una variante de glucoamilasa de una glucoamilasa precursora como se define en las reivindicaciones. La glucoamilasa precursora comprende un dominio catalítico y un dominio de unión a almidón. La glucoamilasa precursora comprende una secuencia que presenta una secuencia y/o identidad estructural con TrGA (SEQ ID NO:2).

40 **[0080]** En algunas formas de realización de la presente exposición, la glucoamilasa precursora comprende un dominio catalítico que presenta una secuencia de aminoácidos con al menos un 50 % de identidad de secuencia de aminoácidos con una o más de las secuencias de aminoácidos ilustradas en SEQ ID NO:1, 2, 3, 5, 6, 7, 8 o 384 incluyendo al menos un 60 %, al menos un 70 %, al menos un 80 %, al menos un 90 %, al menos un 95 % y  
45 al menos un 99 % de identidad de secuencia con SEQ ID NO:1, 2, 3, 5, 6, 7, 8 y/o 384. En otras formas de realización, la glucoamilasa precursora presenta al menos un 80 % de identidad de secuencia, al menos un 85 % de identidad de secuencia, al menos un 90 % de identidad de secuencia, al menos un 95 % de identidad de secuencia, al menos un 97 % de identidad de secuencia, al menos un 98 % de identidad de secuencia y al menos un 99 % de identidad de secuencia con el dominio catalítico de la secuencia de aminoácidos TrGA SEQ ID NO:3.

50 **[0081]** En algunas formas de realización, la glucoamilasa de la presente exposición comprende un dominio de unión a almidón que presenta una identidad estructural con SEQ ID NO:11. En algunas formas de realización, la glucoamilasa precursora comprende un dominio de unión a almidón que presenta una secuencia de aminoácidos con al menos un 30 % de identidad de secuencia, al menos un 40 % de identidad de secuencia, al menos un 50 % de identidad de secuencia, al menos un 60 % de identidad de secuencia, al menos un 70 % de identidad de

secuencia, al menos un 80 % de identidad de secuencia, al menos un 85 % de identidad de secuencia, al menos un 90 % de identidad de secuencia, al menos un 95 % de identidad de secuencia, al menos un 97 % de identidad de secuencia, al menos un 98 % de identidad de secuencia y al menos un 99 % de identidad de secuencia con el SBD de la secuencia de aminoácidos TrGA SEQ ID NO:11.

5 **[0082]** La glucoamilasa precursora puede codificarse mediante una secuencia de ADN que hibrida bajo condiciones astringentes, altas o medias con un ADN que codifica una glucoamilasa que comprende una de las secuencias de aminoácidos de SEQ ID NO:1, 2, y/o 11. En algunas formas de realización, la glucoamilasa codificada presenta al menos un 50 % de identidad de secuencia, al menos un 60 % de identidad de secuencia, al menos un 70 % de identidad de secuencia, al menos un 80 % de identidad de secuencia, al menos un 85 % de identidad de secuencia, al menos un 90 % de identidad de secuencia, al menos un 95 % de identidad de secuencia, al menos un 97 % de identidad de secuencia, al menos un 98 % de identidad de secuencia y al menos un 99 % de identidad de secuencia con la secuencia de aminoácidos SEQ ID NO:1, 2 y/o 11. En algunas formas de realización, la glucoamilasa precursora es una secuencia natural o que se da de forma natural en una célula huésped. En algunas formas de realización, la glucoamilasa precursora es una variante que se da de forma natural. En algunas formas de realización, la glucoamilasa precursora es una secuencia de referencia que es una variante que se ha diseñado o es una glucoamilasa híbrida.

10 **[0083]** La estructura predicha y las secuencias conocidas de glucoamilasas se conservan entre especies fúngicas (Coutinho *et al.*, 1994 *Protein Eng.*, 7:393 - 400 y Coutinho *et al.*, 1994, *Protein Eng.*, 7: 749-760). En algunas formas de realización, la glucoamilasa precursora es una glucoamilasa fúngica filamentosa. En algunas formas de realización, la glucoamilasa precursora se obtiene de una cepa de *Trichoderma* (p. ej., *T. reesei*, *T. longibrachiatum*, *T. strictipellis*, *T. asperellum*, *T. konilangbra* y *T. haz-ianum*), una cepa de *Aspergillus* (p. ej., *A. niger*, *A. nidulans*, *A. kawachi*, *A. awamori* y *A. oryzae*), una cepa de *Talaromyces* (p. ej., *T. emersonii*, *T. thermophilus*, y *T. duponti*), una cepa de *Hypocrea* (p. ej., *H. gelatinosa*, *H. orientalis*, *H. vinosa*, y *H. citrina*), una cepa de *Fusarium* (p. ej., *F. oxysporum*, *F. roseum*, y *F. venenatum*), una cepa de *Neurospora* (p. ej., *N. crassa*) una cepa de *Humicola* (p. ej., *H. grisea*, *H. insolens* y *H. lanuginosa*), una cepa de *Penicillium* (p. ej., *P. notatum* o *P. chrysogenum*), o una cepa de *Saccharomycopsis* (p. ej., *S. fibuligera*). En algunas formas de realización, la glucoamilasa precursora comprende la secuencia de aminoácidos de aquellas secuencias ilustradas en las figuras 5A-E.

20 **[0084]** En algunas formas de realización, la glucoamilasa precursora puede ser una glucoamilasa bacteriana. Por ejemplo, la glucoamilasa puede obtenerse de cepas bacterianas gram positivas tales como *Bacillus* (p. ej., *B. alkalophilus*, *B. amyloliquefaciens*, *B. lentus*, *B. licheniformis*, *B. stearothermophilus*, *B. subtilis* y *B. thuringiensis*) o una cepa de *Streptomyces* (p. ej., *S. lividans*).

25 **[0085]** En algunas formas de realización, la glucoamilasa precursora de la presente exposición tendrá al menos un 50 % de identidad de secuencia, al menos un 60 % de identidad de secuencia, al menos un 70 % de identidad de secuencia, al menos un 80 % de identidad de secuencia, al menos un 85 % de identidad de secuencia, al menos un 88 % de identidad de secuencia, al menos un 90 % de identidad de secuencia, al menos un 93 % de identidad de secuencia, al menos un 95 % de identidad de secuencia, al menos un 96 % de identidad de secuencia, al menos un 97 % de identidad de secuencia, al menos un 98 % de identidad de secuencia, y al menos un 99 % de identidad de secuencia con la secuencia de aminoácidos TrGA de SEQ ID NO: 2. En algunas formas de realización, la glucoamilasa precursora también identidad estructural con SEQ ID NO:2.

30 **[0086]** En formas de realización adicionales, se obtendrá un homólogo de glucoamilasa *Trichoderma* a partir de una cepa de *Trichoderma* o *Hypocrea*. Algunos homólogos de glucoamilasa *Trichoderma* se describen en la patente estadounidense pub. nº 2006/0094080 y se hace referencia de forma específica a secuencias de aminoácidos detallados en SEQ ID NOs: 17 - 22 y 43 - 47 de dicha referencia.

35 **[0087]** En algunas formas de realización, la glucoamilasa precursora es TrGA que comprende la secuencia de aminoácidos SEQ ID NO:2 o un homólogo de glucoamilasa *Trichoderma* que presenta al menos un 50 %, al menos un 60 %, al menos un 70 %, al menos un 80 %, al menos un 85 %, al menos un 88 %, al menos un 90 %, al menos un 93 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 % de identidad de secuencia con la secuencia de TrGA. En algunas formas de realización, la glucoamilasa precursora presenta identidad estructural con la secuencia de TrGA (SEQ ID NO:2) y comprende un SBD con identidad estructural con el SBD de TrGA (SEQ ID NO:11).

40 **[0088]** Una glucoamilasa precursora puede aislarse y/o identificarse mediante técnicas de ADN recombinante convencionales. Se puede utilizar cualquier técnica convencional conocida por los expertos. Por ejemplo, se pueden utilizar sondas y/o cebadores específicos para áreas conservadas de la glucoamilasa con el fin de identificar homólogos en células bacterianas o fúngicas (el dominio catalítico, el centro activo, etc.). De forma alternativa, se puede utilizar una PCR con cebadores degenerados para identificar homólogos en células bacterianas o fúngicas. En algunos casos, se pueden analizar secuencias conocidas, tales como en una base de datos, para la secuencia y/o identidad estructural con una de las glucoamilasas conocidas, entre las que se

incluyen SEQ ID NO:2 o un dominio de unión a almidón conocido, entre el que se incluye SEQ ID NO:11. También se pueden utilizar análisis funcionales para identificar la actividad de glucoamilasa en una célula bacteriana o fúngica. Las proteínas con actividad de glucoamilasa pueden aislarse e invertirse con el fin de aislar la secuencia de ADN correspondiente. El experto en la técnica conoce tales métodos.

##### 5 **Homología estructural de la glucoamilasa:**

10 **[0089]** El dogma central de la biología molecular es que la secuencia de ADN que codifica un gen para una enzima específica, determina la secuencia de aminoácidos de la proteína, secuencia que a su vez determina el plegamiento tridimensional de la enzima. Este plegamiento reúne los residuos dispares que crean un centro catalítico y una superficie de unión al sustrato y tiene como resultado una alta especificidad y actividad de las enzimas en cuestión.

15 **[0090]** Las glucoamilasas consisten en hasta tres dominios estructurales distintos, un dominio catalítico de aproximadamente 450 residuos que está conservado de forma estructural en glucoamilasas, generalmente seguido de una región de enlace consistente en entre 30 y 80 residuos que están conectados a un dominio de unión a almidón de aproximadamente 100 residuos. La estructura de la glucoamilasa *Trichoderma reesei* con las tres regiones intactas se determinó con la resolución Angstrom 1,8 en el presente documento (véase la tabla 8 y el ejemplo 11). Al utilizar las coordenadas (véase la tabla 8) la estructura catalítica se alineó con las coordenadas del dominio catalítico de la cepa *Aspergillus awamori* X100 que se determinó previamente (Aleshin, A.E., Hoffman, C., Firsov, L.M., and Honzatko, R.B. 1994 "Refined crystal structures of glucoamylase from *Aspergillus awamori* var. X100". *J Mol Biol* 238: 575-591.). La estructura de cristal *Aspergillus awamori* incluía únicamente el dominio catalítico. Como se ve en las figuras 6 y 7, la estructura de los dominios catalíticos se superpone muy estrechamente y es posible identificar residuos equivalentes basándose en esta superposición estructural. Los inventores creen que las glucoamilasas comparten la estructura básica representada en las figuras 6 y 7.

25 **[0091]** Las figuras 6 y 7 son comparaciones de las estructuras tridimensionales de la glucoamilasa *Trichoderma* (negra) de SEQ ID NO:1 (véase la figura 1 para la secuencia de aminoácidos) y de *Aspergillus awamori* (gris) vistas desde el lateral y parte superior, respectivamente. El lateral se define en relación con el centro activo de la molécula que está en la "parte superior". En la vista lateral, puede verse la relación entre el dominio catalítico y la región de enlace y el dominio de unión a almidón. Las glucoamilasas mostradas en el presente documento y de hecho glucoamilasas conocidas a día de hoy comparten esta homología estructural, especialmente en el dominio catalítico. La conservación de la estructura de la molécula de glucoamilasa se corresponde con la conservación de la actividad y un mecanismo conservado de acción para todas las glucoamilasas. Teniendo en cuenta esta alta homología, las variantes específicas del lugar de la glucoamilasa *Trichoderma* que tienen como resultado una función modificada tendrían consecuencias estructurales y, por tanto, funcionales similares en otras glucoamilasas. Por lo tanto, la información dada a conocer sobre qué variantes tendrían beneficios aconsejables puede aplicarse a otras glucoamilasas.

35 **[0092]** Se produjo una estructura de cristal adicional utilizando las coordenadas de la tabla 8 para el dominio de unión a almidón. El SBD para TrGA se alineó con el SBD para *A. niger*. Como se muestra en la figura 8, la estructura de los SBD de *A. niger* y TrGA se superponen estrechamente. Los inventores creen que, aunque todos los dominios de unión a almidón comparten al menos parte de la estructura básica representada en la figura 8, algunos SBD se parecen más estructuralmente que otros. Por ejemplo, el SBD de TrGA puede clasificarse como dentro de la familia del módulo de unión de carbohidrato 20 dentro de la base de datos CAZY (cazy.org). La base de datos CAZY describe las familias de módulos de unión de carbohidratos y catalíticos relacionados estructuralmente (o dominios funcionales) de enzimas que degradan, modifican o crean enlaces glicosídicos. Dada una homología estructural alta, las variantes específicas del lugar del SBD de TrGA que tienen como resultado una función modificada también tendrían consecuencias estructurales y, por tanto, funcionales similares en otras glucoamilasas que tengan SBD con estructura similar a la del SBD de TrGA, especialmente aquellos clasificados dentro de la familia de módulo 20 de unión de carbohidratos. Por lo tanto, la información dada a conocer sobre qué variantes tendrían beneficios aconsejables puede aplicarse a otros SBD con similaridad estructural.

50 **[0093]** La identidad estructural determina si los residuos de aminoácidos son equivalentes. La identidad estructural es un equivalente topológico directo cuando las dos estructuras (estructuras de aminoácidos y tridimensionales) se alinean. Una posición de residuos (aminoácidos) de una glucoamilasa es equivalente a un residuo de glucoamilasa *T. reesei* si es tanto homóloga (es decir, se corresponde en posición ya sea en estructura primaria o terciaria) como análoga con un residuo específico o parte de ese residuo en la glucoamilasa *T. reesei* (que tenga la misma capacidad funcional o similar para combinar, reaccionar o interactuar de forma química).

55 **[0094]** Con el fin de establecer la identidad con la estructura primaria, la secuencia de aminoácidos de una glucoamilasa *Trichoderma reesei* puede compararse directamente con la secuencia primaria de glucoamilasa y específicamente con un conjunto de residuos que se sabe son invariables en glucoamilasas por cuya secuencia

son conocidos. Por ejemplo, las figuras 5A y B en el presente documento muestran los residuos conservados entre los dominios catalíticos de glucoamilasa. Las figuras 5D y E muestran una alineación de dominios de unión a almidón a partir de glucoamilasas. Una vez se han alineado los residuos conservados, y permitiendo las inserciones y deleciones necesarias con el fin de mantener la alineación (es decir, evitando la eliminación de residuos conservados a través de la deleción e inserción arbitraria), se definen los residuos equivalentes a aminoácidos específicos en la secuencia primaria de glucoamilasa *Trichoderma reesei*. La alineación de residuos conservados puede conservar el 100 % de tales residuos. Sin embargo, la alineación de más del 75 % o tan solo un 40 % de residuos conservados también es adecuada para definir residuos equivalentes. Además, la identidad estructural puede utilizarse junto con la identidad de secuencia con los residuos equivalentes de identidad.

[0095] Por ejemplo, en las figuras 5A y 5B, se alinean glucoamilasas de seis organismos para proporcionar la máxima cantidad de homología entre las secuencias de aminoácidos. Una comparación de estas secuencias muestra que existe un número de residuos conservados contenidos en cada secuencia como se indica mediante un asterisco. Por lo tanto, se pueden utilizar estos residuos conservados para definir los correspondientes residuos de aminoácidos equivalentes de la glucoamilasa *Trichoderma reesei* en otras glucoamilasas tal como la glucoamilasa de *Aspergillus niger*. Las figuras 5D y 5E muestran los SBD de seis organismos alineados para identificar los residuos equivalentes entre ellos.

[0096] La identidad estructural supone la identificación de residuos equivalentes entre las dos estructuras. "Los residuos equivalentes" pueden definirse mediante la determinación de homología en el nivel de la estructura terciaria (identidad estructural) para una enzima cuya estructura terciaria ha sido determinada mediante técnicas NMR y/o cristalografía de rayos X. Para la cristalografía de rayos X, los residuos equivalentes se definen como aquellos para cuyas coordenadas atómicas de dos o más de los átomos de la cadena principal de un residuo de aminoácido específico de la glucoamilasa *Trichoderma reesei* (N en N, CA en CA, C en C y O en O) están dentro de 0,13 nm y preferiblemente 0,1 nm tras la alineación. La alineación se consigue después de que el mejor modelo se haya orientado y colocado para proporcionar la máxima superposición de coordenadas atómicas de átomos de proteína no hidrógenos de la glucoamilasa en cuestión con la glucoamilasa *Trichoderma reesei*. El mejor modelo es el modelo cristalográfico que proporciona el factor R más bajo para los datos de difracción experimental con la resolución más alta disponible.

$$\text{Factor } R = \frac{\sum_h |F_o(h)| - |F_c(h)|}{\sum_h |F_o(h)|}$$

[0097] Los residuos equivalentes que son análogos funcionalmente a un residuo específico de la glucoamilasa *Trichoderma reesei* se definen como aquellos aminoácidos de la enzima que pueden adoptar una conformación de tal forma que bien cambian, modifican o contribuyen a la estructura de proteína, unión de sustrato o catálisis de forma definida y atribuida a un residuo específico de la glucoamilasa *Trichoderma reesei*. Además, se trata de aquellos residuos de la enzima (para los que se ha obtenido una estructura terciaria mediante cristalografía de rayos X) que ocupan una posición análoga hasta el punto de que, aunque los átomos de la cadena principal del residuo determinado puedan no cumplir los criterios de equivalencia sobre la base de ocupar una posición homóloga, las coordenadas atómicas de al menos dos de los átomos de la cadena lateral del residuo quedan con 0,13 nm de los correspondientes átomos de la cadena literal de la glucoamilasa *Trichoderma reesei*. Las coordenadas de la estructura tridimensional de la glucoamilasa *Trichoderma reesei* están establecidos en la tabla 8 y pueden utilizarse como se detalla anteriormente con el fin de determinar residuos equivalentes en el nivel de estructura terciaria.

#### Variantes de SBD

[0098] Las variantes según la presente exposición incluyen al menos una sustitución, deleción o inserción en la secuencia de aminoácidos del SBD de una glucoamilasa precursora. En algunas formas de realización, las variantes de SBD tendrán al menos un 20 %, al menos un 40%, al menos un 50%, al menos un 60%, al menos un 70%, al menos un 80%, al menos un 85%, al menos un 90%, al menos un 95%, el menos un 97 % y también al menos un 100 % de la actividad de glucoamilasa en comparación con la actividad de glucoamilasa de TrGA (SEQ ID NO:2) cuando se compara con las mismas condiciones básicamente.

[0099] En algunas formas de realización, las variantes de SBD según la presente exposición comprenderán una sustitución, deleción o inserción en al menos una posición de aminoácido del SBD de la TrGA precursora (SEQ ID NO:2) o en una posición equivalente en la secuencia de otra glucoamilasa precursora. En algunas formas de realización, la glucoamilasa precursora presenta al menos un 50 %, al menos un 60 %, al menos un 70 %, al menos un 80 %, al menos un 90 % de identidad de secuencia con la secuencia de TrGA, que incluye sin carácter limitativo, al menos un 93 % de identidad de secuencia, al menos un 95 %, al menos un 97 % y al menos un 99 % de identidad de secuencia con el dominio catalítico de TrGA (SEQ ID NO:3). En algunas formas de realización, la glucoamilasa precursora presenta al menos un 50 %, al menos un 60 %, al menos un 70 %, al menos un 80 %, al menos un 90 % de identidad de secuencia con la secuencia TrGA que incluye sin carácter limitativo, al menos

un 93 % de secuencia de identidad, al menos un 95 %, al menos un 97 % y al menos un 99 % de identidad de secuencia con la proteína madura de TrGA (SEQ ID NO:2). En algunas formas de realización, la glucoamilasa precursora presenta identidad estructural con la secuencia de TrGA

5 **[0100]** En algunas formas de realización, la variante con SBD comprenderá una sustitución, delección o inserción y preferiblemente una sustitución en la región del bucle 1 (aa 560-570) y/o la región del bucle 2 (aa 523--527) de la secuencia correspondiente a SEQ ID NO:2. En concreto, las regiones que incluyen los residuos de aminoácidos 558-562 y/o los residuos de aminoácidos 570-578 son regiones de sustitución.

10 **[0101]** En otras formas de realización, la variante según la presente exposición comprenderá una sustitución, delección o inserción en al menos una posición de aminoácido de un fragmento del SBD de la TrGA precursora, en el que el fragmento comprende el dominio catalítico y al menos parte del SBD de la secuencia de TrGA (SEQ ID NO:3) o en una posición equivalente en un fragmento que comprende el dominio catalítico de una glucoamilasa precursora, dominio catalítico que presenta al menos un 50 %, al menos un 60%, al menos un 70%, al menos un 80% de identidad de secuencia, al menos un 90%, al menos un 95%, al menos un 97% y al menos un 99% de identidad de secuencia con el fragmento de la secuencia de TrGA. En algunas formas de realización, el fragmento de SBD comprenderá al menos 40, 50, 60, 70, 80, 90, 100, y/o 109 residuos de aminoácidos del SBD (SEQ ID NO:11). En algunas formas de realización, cuando la glucoamilasa precursora incluye un dominio catalítico, la región de enlace y el dominio de unión a almidón, el fragmento puede incluir parte de la región de enlace. En algunas formas de realización, la variante comprenderá una sustitución, delección o inserción en la secuencia de aminoácidos de un fragmento de la secuencia TrGA (SEQ ID NO:2). En algunas formas de realización, la variante tendrá identidad estructural con la secuencia de TrGA (SEQ ID NO: 2).

**[0102]** En algunas formas de realización, la variante de glucoamilasa incluirá al menos una sustitución en la secuencia de aminoácidos del SBD de una precursora. En formas de realización adicionales, la variante tendrá más de una sustitución (p. ej., dos, tres o cuatro sustituciones).

25 **[0103]** Aunque las variantes pueden estar en cualquier posición en el dominio de unión a almidón de la secuencia de proteína madura (SEQ ID NO: 2), en algunas formas de realización, una variante de glucoamilasa comprende una o más sustituciones en las siguientes posiciones en la secuencia de aminoácidos establecida en la SEQ ID NO:2: 493, 494, 495, 501, 502, 503, 508, 511, 517, 518, 519, 520, 525, 527, 531, 533, 535, 536, 537, 538, 539, 540, 545, 546, 547, 549, 551, 561, 563, 567, 569, 577, 579, y 583 y/o en una posición equivalente en una glucoamilasa precursora. En algunas formas de realización, la variante de glucoamilasa comprende una o más sustituciones de aminoácidos correspondientes a la posición, 493, 494, 495, 502, 503, 508, 511, 518, 519, 520, 527, 531, 535, 536, 537, 539, 563, y 577 de SEQ ID NO:2 o una posición equivalente en una glucoamilasa precursora.

30 **[0104]** En algunas formas de realización, la variante incluirá al menos una sustitución en una posición equivalente a una posición establecida en SEQ ID NO:2 y especialmente en una posición correspondiente a T493, P494, T495, H502, E503, Q508, Q511, N518, A519, A520, T527, V531, A535, V536, N537, A539, N563, y N577 de SEQ ID NO:2 o una posición equivalente en una glucoamilasa precursora.

35 **[0105]** En algunas formas de realización, la glucoamilasa precursora tendrá al menos un 50 % de identidad de secuencia, al menos un 80 %, al menos un 85%, al menos un 90%, al menos un 95%, al menos un 96%, al menos un 97%, al menos un 98% y al menos un 99% de identidad de secuencia con SEQ ID NO: 2. En algunas formas de realización, la glucoamilasa precursora es un homólogo de glucoamilasa *Trichoderma*.

40 **[0106]** En formas de realización adicionales, la variante con SBD de una precursora de glucoamilasa comprende al menos una de las siguientes sustituciones en las siguientes posiciones en una secuencia de aminoácidos establecida en SEQ ID NO:2: T493C/M/N/Q/Y; P494H/I/M/N/Q/W; T495M/P/R; H502A/M/SN; E503C/D/H/S/W; Q508N/P/Y; Q511C/G/H/I/K/T/N; N518S/P/T; A519E/I/K/R/T/V/Y; A520C/E/L/P/Q/R/W; T527AN; V531A/L/N/R/S/T; A535E/F/G/KL/N/P/R/S/T/V/W/Y; V536C/E/I/L/M/Q/R/S; A539E/H/M/R/S/W; N563A/C/E/I/K/L/Q/T/V; N577A/K/P/R/V y/o una sustitución en una posición equivalente en un homólogo de glucoamilasa precursora.

45 **[0107]** En algunas formas de realización, la variante de glucoamilasa aislada comprende un dominio catalítico que presenta al menos un 85 %, al menos un 90%, al menos un 95%, al menos un 97 %, al menos un 99% de identidad de secuencia de aminoácidos con SEQ ID NO:3 y un SBD que comprende una o más sustituciones de aminoácidos con una posición correspondiente a la posición: 3, 4, 5, 11, 12, 13, 18, 21, 27, 28, 29, 30, 35, 37, 41, 43, 45, 46, 47, 48, 49, 50, 55, 56, 57, 59, 61, 71, 73, 77, 79, 87, 89, y 93 de SEQ ID NO: 11 o un SBD que comprende una o más sustituciones de aminoácidos en una posición equivalente a SEQ ID NO: 11 de un SBD de glucoamilasa precursora.

55 **Variantes con SBD que incluyen además una sustitución, delección o inserción en el dominio catalítico:**

- [0108]** En algunas formas de realización, la variante de glucoamilasa tendrá una sustitución, delección o inserción en un dominio catalítico además del SBD. En algunas formas de realización, la sustitución, delección o inserción del dominio catalítico incluirá cualquiera de los residuos de aminoácidos ilustrados en la secuencia de SEQ ID NO:3. En algunas formas de realización, la variante comprende una de las variantes del dominio catalítico descritas en el presente documento. En otras formas de realización, la variante comprende una de las variantes del dominio catalítico descritas en la solicitud de PCT PCT/US07/21683, presentada el 9 de octubre, 2007. En algunas formas de realización, la variación del dominio catalítico incluirá más de una sustitución en el dominio catalítico (p. ej., dos, tres o cuatro sustituciones) en comparación con un correspondiente dominio catalítico de glucoamilasa precursora.
- [0109]** En algunas formas de realización, la variante de glucoamilasa con al menos una variación en el SBD también comprende una sustitución, delección o inserción en al menos una posición de aminoácidos en una posición correspondiente a las regiones de aminoácidos no conservados del dominio catalítico como se ilustra en las figuras 5A y 5B (p. ej., posiciones de aminoácidos correspondientes a aquellas posiciones que no están indicadas con "\*" en las figuras 5A y 5B). En algunas formas de realización, la variante con SBD también comprende una sustitución en al menos una posición de aminoácidos en una posición correspondiente a las regiones de aminoácidos no conservados como se ilustra en las figuras 5A y 5B.
- [0110]** En algunas formas de realización, la variante con SBD incluida en la presente exposición comprende una o más sustituciones en las siguientes posiciones en la secuencia de aminoácidos establecida en SEQ ID NO: 2 o 3. 10, 14, 15, 23, 42, 45, 46, 59, 60, 61, 67, 68, 72, 73, 97, 98, 99, 102, 108, 110, 113, 114, 122, 124, 125, 133, 140, 144, 145, 147, 152, 153, 164, 175, 182, 204, 205, 214, 216, 219, 228, 229, 230, 231, 236, 239, 240, 241, 242, 244, 263, 264, 265, 268, 269, 276, 284, 291, 300, 301, 303, 310, 311, 313, 316, 338, 342, 344, 346, 349, 359, 361, 364, 375, 379, 382, 390, 391, 393, 394, 408, 410, 415, 417, 418, 430, 431, 433, 436, 442, 443, 444, 448 y 451, y/o en una posición equivalente en una glucoamilasa precursora. En algunas formas de realización, la glucoamilasa precursora presentará al menos un 50 %, al menos un 60 %, al menos un 70 %, al menos un 80 %, al menos un 90 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 % y al menos un 99 % de identidad de secuencia con SEQ ID NO:2 o SEQ ID NO:3. En algunas formas de realización, la glucoamilasa precursora será un homólogo de glucoamilasa *Trichoderma*. En algunas formas de realización, la variante tendrá propiedades modificadas en comparación con una glucoamilasa precursora. En algunas formas de realización, la glucoamilasa precursora tendrá identidad estructural con la glucoamilasa de SEQ ID NO:2 o 3.
- [0111]** En algunas formas de realización, la variante de glucoamilasa con SBD también comprende una o más sustituciones en las siguientes posiciones en el dominio catalítico en la secuencia de aminoácidos establecida en la SEQ ID NO:2 o 3: T10, L14, N15, P23, T42, P45, D46, F59, K60, N61, T67, E68, A72, G73, S97, L98, A99, S102, K108, E110, L113, K114, R122, Q124, R125, I133, K140, N 144, N 145, Y 147, S 152, N 153, N164, F175, N182, A204, T205, S214, V216, Q219, W228, V229, S230, S231, D236, I239, N240, T241, N242, G244, N263, L264, G265, A268, G269, D276, V284, S291, P300, A301, A303, Y310, A311, D313, Y316, V338, T342, S344, T346, A349, V359, G361, A364, T375, N379, S382, S390, E391, A393, K394, R408, S410, S415, L417, H418, T430, A431, R433, I436, A442, N443, S444, T448 y S451 y/o una posición equivalente en una precursora.
- [0112]** En algunas formas de realización, la variante de glucoamilasa con SBD incluida en el presente documento también comprende una o más sustituciones en las siguientes posiciones en el dominio catalítico en la secuencia de aminoácidos establecida en la SEQ ID NO:2 o 3: 10, 14, 15, 23, 59, 60, 61, 65, 67, 68, 72, 73, 97, 98, 99, 102, 110, 113, 133, 140, 144, 145, 147, 152, 153, 164, 182, 204, 205, 214, 216, 219, 228, 229, 230, 231, 236, 239, 241, 242, 263, 264, 265, 268, 269, 276, 284, 291, 300, 301, 303, 311, 338, 342, 344, 346, 349, 359, 361, 364, 375, 379, 382, 390, 391, 393, 394, 410, 417, 418, 430, 431, 433, 442, 444, 448, y 451 y/o una posición equivalente en una glucoamilasa precursora. En algunas formas de realización, la variante presenta una o más sustituciones correspondientes a una de las siguientes posiciones: 61, 67, 72, 97, 102, 133, 205, 219, 228, 230, 231, 239, 263, 268, 291, 342, 394, 430, 431 y 451 de SEQ ID NO:2 y/o 3. En algunas formas de realización, la sustitución en estas posiciones se elige de entre: N61I, T67M, A72Y, S97N, S102M, I133T, T205Q, Q219S, W228H, W228M, S230F, S230G, S230N, S230R, S231L, I239V, I239Y, N263P, A268C, A268G, A268K, S291A, T342V, K394S, T430K, A431 Q y S451K de SEQ ID NO:2 y/o 3. En algunas formas de realización, la variante presenta una o más sustituciones correspondientes a una de las siguientes posiciones: 72, 133, 219, 228, 230, 231, 239, 263, 268 y 451 de SEQ ID NO:2 y/o 3. En algunas formas de realización, la sustitución en estas posiciones se elige de entre: A72Y, I133T, Q219S, W228H, W228M, S230R, S230F, S230G, S231L, I239V, N263P, A268C, A268G y S451K de SEQ ID NO:2 y/o 3.
- [0113]** En formas de realización adicionales, la variante con SBD de una precursora de glucoamilasa comprende también al menos una de las siguientes sustituciones en las siguientes posiciones del dominio catalítico en una secuencia de aminoácidos establecida en SEQ ID NO:2 o 3: T10D/F/G/K/L/M/P/R/S; L14E/H; N15D/N; P23A/G; F59A/G; K60F/H; N61D/I/L/Q/V/W; R65A/C/G/I/K/M/S/V/Y; T67C/I/K/M/T; E68I/M/W; A72E/G/L/M/Q/R/W/Y; G73C/L/W; S97F/M/N/P/R/S/V/W/Y; L98H/M; A99C/L/M/N/P; S102A/C/I/L/M/N/RN/W/Y; E110Q/S/W; L113E/N; K114C/D/E/L/M/Q/S/T/V; I113K/R/S/T; K 140A/E/F/H/K/L/M/N/Q/R/S/V/W/Y; N144C/D/E/I/K; N145A/C/E/I/K/L/M/Q/R/V/W/Y; Y147A/M/R; S152H/M; N153C/D/K/L/W/Y; N164A/G; N182C/E/K/P/R;



A204C/D/G/I/M/Q/T; T205A/D/H/I/K/M/N/P/Q/S/V/W/Y; S214P/T; V216C/G/H/K/N/Y; Q219D/G/H/N/P/S;  
 W228A/F/G/H/I/L/M/Q/S/T/V/Y; V229E/I/M/N/Q; S230C/D/E/F/G/H/I/K/L/N/P/Q/R/T/V/Y;  
 S231C/D/F/L/M/N/Q/R/SN/Y; D236F/G/L/M/N/P/S/T/V; I239M/Q/SN/W/Y; T241C/E/H/L/M/P/S/TN;  
 N242C/F/H/M/T/V/W; N263H/K/P; L264A/C/E/F/L/S; G265E/G/H/I/K/R/T; A268C/D/E/F/G/I/K/L/P/R/T/W;  
 5 G269E; D276S; V284R/T/V/Y/A/E/F/H/K/N/P/W; P300K/R; A301E/K/L/P/S/W;  
 A303C/D/F/H/I/K/L/N/R/T/V/W/Y; A311N/P/Q/S/Y; V338P/Q/S/Y; T342N/V; S344A/T; T346G/H/M/N/P/Q/Y;  
 A349L/I/K/M/N/Q/R/W; G361H/L/R; A364M/W; T375C/D/E/H/V/W/Y; N379A/C/D/G/I/M/P/S; S382A/N/P/V/W;  
 S390A/Y; E391A/E/I/K/L/M/Q/R/V/W/Y; A393E/G/H/I/K/L/M/N/Q/R/ST/V/W/Y; K394A/1-UK/L/M/Q/R/S/TN/W;  
 S410E/H/N; L417A/D/E/F/G/I/K/Q/R/S/TN/W/Y; H418E/M; T430A/E/F/G/H/I/K/M/N/Q/R/V;  
 10 A431C/E/H/I/L/M/Q/R/S/W/Y; R433A/C/E/F/G/K/L/M/N/SN/W/Y; I436E/F/G/H/K/P/R/S/TN/Y;  
 S444M/N/P/Q/R/T/V/W; T448F/G/I/P/Q/TN; y S451 E/H/K/L/Q/T y/o una sustitución en una posición equivalente  
 en una glucoamilasa precursora.

**[0114]** En otras formas de realización, la variante con SBD de una precursora de glucoamilasa también  
 15 comprende una o más sustituciones en las siguientes posiciones en la secuencia de aminoácidos del dominio  
 catalítico establecida en SEQ ID NO:2 o 3: 10, 15, 23, 42, 59, 60, 61, 68, 72, 73, 97, 98, 99, 102, 114, 133, 140,  
 144, 147, 152, 153, 164, 182, 204, 205, 214, 216, 228, 229, 230, 231, 236, 241, 242, 263, 264, 265, 268, 269,  
 276, 284, 291, 300, 301, 303, 311, 338, 342, 344, 346, 349, 359, 361, 364, 375, 379, 382, 390, 391, 393, 394,  
 410, 417, 430, 431, 433, 436, 442, 443, 444, 448 y 451 o una posición equivalente en una glucoamilasa  
 precursora. En algunas formas de realización, la variante presenta una o más sustituciones correspondientes a  
 20 una de las siguientes posiciones: 10, 42, 68, 73, 97, 114, 153, 229, 231, 236, 264, 291, 301, 344, 361, 364, 375,  
 417 y 433 de SEQ ID NO:2 y/o 3. En algunas formas de realización, la sustitución en estas posiciones se elige de  
 entre: T10S, T42V, E68C, E68M, G73F, G73W, K114M, K114T, N153A, N153S, N153V, W228V, D236R, G361  
 D, G361E, G361P, G361 Y, A364D, A364E, A364F, A364G, A364K, A365L, A365R, R433C, R433G, R433L,  
 R433N, R433S, R433V y 1436H de SEQ ID NO:2 y/o 3. En algunas formas de realización, la variante presenta  
 25 una o más sustituciones correspondientes a una de las siguientes posiciones: 42, 68, 73, 114, 153, 236, 361 y  
 364 de SEQ ID NO:2 y/o 3. En algunas formas de realización, la sustitución en estas posiciones se elige de  
 entre: T42V, E68M, G73F, G73W, K114T, N153S, N153V, D236R, G361D, A364F y A364L de SEQ ID NO:2 y/o  
 3.

**[0115]** En algunas formas de realización, la variante con SBD presenta una o más sustituciones  
 30 correspondientes a una de las siguientes posiciones: 228, 230, 231, 268, 291, 417, 433 y 451 de SEQ ID NO:2 o  
 3. En algunas formas de realización, la sustitución en estas posiciones se elige de entre: W228H, W228M,  
 S230F, S230G, S230R, S231L, A268C, A268G, S291 A, L417R, R433Y y S451K de SEQ ID NO:2 y/o 3.

**Variantes de glucoamilasa con SBD híbridas o quiméricas:**

**[0116]** Las variantes de glucoamilasa de la presente exposición pueden también incluir glucoamilasas híbridas o  
 35 quiméricas con, por ejemplo, un dominio de unión a almidón (SBD) de una glucoamilasa y un dominio catalítico y  
 enlazador de otra. Por ejemplo, una glucoamilasa híbrida puede formarse mediante el intercambio del SBD de  
 AnGA con el SBD de TrGA, realizando un híbrido con el SBD de AnGA y el dominio catalítico y enlazador de  
 TrGA. De forma alternativa, el SBD y el enlazador de AnGA pueden intercambiarse por el SBD y el enlazador de  
 TrGA. En una forma de realización, una variante de glucoamilasa de SBD de acuerdo con la presente exposición  
 40 comprenderá un dominio catalítico de una glucoamilasa *Aspergillus*, una glucoamilasa *Humicola*, una  
 glucoamilasa *Thermomyces* o una glucoamilasa *Talaromyces* y un SBD que comprende una o más sustituciones  
 de aminoácidos en una posición correspondiente a la posición: 3, 4, 5, 11, 12, 13, 18, 21, 27, 28, 29, 30, 35, 37,  
 41, 43, 45, 46, 47, 48, 49, 50, 55, 56, 57, 59, 61, 71, 73, 77, 79, 87, 89 y 93 de SEQ ID NO:11. En algunas  
 45 formas de realización, el dominio catalítico puede comprender cualquiera de las siguientes secuencias, SEQ ID  
 NO:5; SEQ ID NO:6; SEQ ID NO:7, SEQ ID NO:8 o SEQ ID NO:9.

**Propiedades modificadas:**

**[0117]** Las variantes de glucoamilasa incluidas en la presente exposición pueden también tener al menos una  
 propiedad modificada (p. ej., propiedad mejorada) en comparación con una glucoamilasa precursora y en  
 50 concreto en comparación con la TrGA. En algunas formas de realización, se elige al menos una propiedad  
 modificada de entre estabilidad de PH aumentada, estabilidad térmica aumentada y/o actividad específica  
 aumentada. En algunas formas de realización, la estabilidad de pH aumentada está a un pH 7,0 - 8,5. En formas  
 de realización adicionales, la estabilidad de pH aumentada presenta niveles de pH inferiores a 7,0, inferiores a  
 6,5, inferiores a 6,0 (tal como pH entre 6,5 y 6,0, entre 6,0 y 5,5, entre 6,0 y 5,0 y entre 6,0 y 4,5).

**[0118]** Las variantes de glucoamilasa de la presente exposición pueden también proporcionar índices más altos  
 55 de hidrólisis de almidón con concentraciones de sustrato bajas en comparación con la glucoamilasa precursora.  
 La variante puede tener una Vmax superior o Km inferior que una glucoamilasa precursora cuando se realiza el  
 análisis con las mismas condiciones. Por ejemplo, la glucoamilasa variante puede tener una Vmax superior con  
 una gama de temperatura de 25 °C a 70 °C (p. ej., a 25 °C a 35 °C; 30 °C - 35 °C; 40 °C a 50 °C; de 50 °C a 55

°C y de 55°C a 62°C). La constante Michaelis-Menten, los valores Km y Vmax pueden determinarse fácilmente utilizando procedimientos conocidos convencionales.

**Variantes con SBD con termoestabilidad aumentada:**

[0119] En algunas formas de realización, las glucoamilasas variantes incluidas en la presente exposición tendrán estabilidad térmica modificada en comparación con una precursora (tipo silvestre). La termoestabilidad se mide como el % de actividad residual tras la incubación durante 1 hora a 64 grados centígrados en un tampón NaAc pH 4,5. Con estas condiciones, TrGA presenta una actividad residual de entre aproximadamente 15 % y 44 % debido a la variación diaria en comparación con la actividad inicial antes de la incubación. Por lo tanto, en algunas formas de realización, las variantes con termoestabilidad aumentada presentan una actividad residual que está entre al menos un 1 % y al menos un 50 % superior a la de la precursora (tras incubación durante 1 hora a 64 grados centígrados en tampón NaAc pH 4,5) incluidos 2 %, 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 %, 10 %, 11 %, 12 %, 13 %, 14 %, 15 %, 16 %, 17 %, 18 %, 19 %, 20 %, 21 %, 22 %, 23 %, 24 %, 25 %, 26 %, 27 %, 28 %, 29 %, 30 %, 31 %, 32 %, 33 %, 34 %, 35 %, 36 %, 37 %, 38 %, 39 %, 40 %, 41 %, 42 %, 43 %, 44 %, 45 %, 46 %, 47 %, 48 %, 49 %, y 50 % en comparación con la actividad inicial antes de la incubación. Por ejemplo, cuando la actividad residual precursora es un 15 %, una variante con estabilidad térmica aumentada puede tener una actividad residual de entre aproximadamente un 15 % y aproximadamente un 75 %. En algunas formas de realización, las variantes de la presente exposición tendrán una termoestabilidad mejorada tal como retener al menos 50 %, 60 %, 70 %, 75 %, 80 %, 85 %, 90 %, 92 %, 95 %, 96 %, 97 %, 98 %, 99 % de actividad enzimática tras la exposición a temperaturas modificadas durante un periodo de tiempo determinado, por ejemplo, al menos 60 minutos, 120 minutos, 180 minutos, 240 minutos, 300 minutos, etc. En algunas formas de realización, la variante ha aumentado la estabilidad térmica en comparación con la glucoamilasa precursora con las temperaturas elegidas en el rango que va desde aproximadamente 40 a aproximadamente 90 °C, desde aproximadamente 40 a aproximadamente 85 °C, desde aproximadamente 45 a aproximadamente 80 °C, desde aproximadamente 50 a aproximadamente 75 °C, desde aproximadamente 60 a aproximadamente 75 °C y también desde aproximadamente 60 a 70 °C. En algunas formas de realización, una variante con SBD presenta termoestabilidad aumentada en comparación con una glucoamilasa precursora a temperaturas por encima de 70 °C, por encima de 75°C, por encima de 80°C y por encima de 85°C con una gama de pH de 4,0 a 6,0. En algunas formas de realización, la termoestabilidad se determina tal y como se describe en los ejemplos. En algunas formas de realización, la variante presenta estabilidad aumentada con temperaturas inferiores en comparación con la glucoamilasa precursora tal como en la gama de 20 a 50 °C incluyendo 35 a 45 y 30°C a 40°C.

[0120] En algunas formas de realización, las variantes de glucoamilasa expuestas en el presente documento con termoestabilidad aumentada incluyen aquellas con una o más sustituciones en las siguientes posiciones en la secuencia de aminoácidos establecida en SEQ ID NO: 2: 493, 495, 503, 508, 511, 518, 519, 520, 527, 531, 535, 536, 537, 539, 563 y 577 o una posición equivalente en una glucoamilasa precursora. En algunas formas de realización, la glucoamilasa precursora será un homólogo de glucoamilasa *Trichoderma* y en formas de realización adicionales, la glucoamilasa precursora tendrá al menos un 90 %, al menos un 95 % y al menos un 98 % de identidad de secuencia con SEQ ID NO:2. En algunas formas de realización, la sustitución se elegirá de entre T493I, T495K/R/S, E503A/C/S/T/V, Q508H/R/S/T, Q511A/D/H/N/S, N518S, A519E/K/R/T/V/Y, A520C/L/P, T527A/V, V531L, A535D/K/N/P/R, VS36I/R, N537W, A539E/H/M/R/S, N563A/C/E/I/K/L/Q/T/V y N577A/K/P/R/V.

[0121] En formas de realización adicionales, las variantes de glucoamilasa expuestas en el presente documento con termoestabilidad aumentada incluyen las variantes con SBD descritas en el párrafo anterior e incluyen de forma adicional una o más deleciones, sustituciones o inserciones y especialmente sustituciones en las siguientes posiciones en la secuencia de aminoácidos establecida en SEQ ID NO:2 o SEQ ID NO:3. T10, N15, P23, T42, F59, K60, N61, E68, A72, G73, S97, L98, A99, S102, K114, I133, K 140, N 144, Y 147, S152, N153, N164, N182, A204, T205, S214, V216, W228, V229, S230, S231, D236, T241, N242, N263, L264, G265, A268, G269, D276, V284, S291, P300, A301, A303, A311, V338, T342, S344, T346, A349, V359, G361, A364, T375, N379, S382, S390, E391, A393, K394, S410, L417, T430, A431, R433, 1436, A442, N443, S444, T448 y S451 y/o una posición equivalente en una glucoamilasa precursora. En algunas formas de realización, la glucoamilasa precursora será un homólogo de glucoamilasa *Trichoderma* y en formas de realización adicionales, la glucoamilasa tendrá al menos un 50 %, al menos un 60 %, al menos un 70 %, al menos un 80 %, al menos un 90 %, al menos un 95 % y al menos un 98 % de identidad de secuencia con SEQ ID NO:2 o SEQ ID NO:3. En algunas formas de realización, la glucoamilasa precursora también presentará identidad estructural con SEQ ID NO:2 y/o SEQ ID NO:3. En algunas formas de realización, la variante con termoestabilidad aumentada cuenta con una sustitución en al menos una de las posiciones: 10, 42, 68, 73, 97, 153, 229, 231, 236, 264, 291, 301, 344, 361, 364, 375, y/o 417 de SEQ ID NO:2 y/o SEQ ID NO:3. En algunas formas de realización, la variante con termoestabilidad aumentada cuenta con una sustitución en al menos una de las posiciones: 42, 68, 73, 153, 236, 344, 361, 364, y/o 365 de SEQ ID NO:2 y/o SEQ ID NO:3.

**Variantes con SBD con actividad específica aumentada**

**[0122]** En algunas formas de realización, una glucoamilasa variante expuesta en el presente documento tendrá una actividad específica modificada en comparación con una glucoamilasa de tipo silvestre o precursora. En algunas formas de realización, la actividad específica modificada es una actividad específica aumentada. La actividad específica aumentada puede definirse como un índice de rendimiento aumentado (PI) mayor o igual que aproximadamente 1,0 incluyendo mayor o igual que aproximadamente 1,1; 1,2; 1,3; 1,4; 1,5; 1,6; 1,7; 1,8; 1,9; 2; 2,3; 2,5; 3,0; 3,3; 3,5; 4,0; 4,3; 4,5 e incluso mayor o igual que 5,0. En algunas formas de realización, la variante cuenta con una actividad específica al menos 1 vez más alta que la glucoamilasa precursora, incluyendo al menos 1,1 veces, 1,2 veces, 1,3 veces, 1,4 veces, 1,5 veces, 1,6 veces, 1,7 veces, 1,8 veces, 1,9 veces, 2 veces, 2,2 veces, 2,5 veces, 2,7 veces, 2,9 veces, 3 veces, 4 veces y 5 veces.

**[0123]** En algunas formas de realización, las variantes de glucoamilasa expuestas en el presente documento con actividad específica aumentada incluyen aquellas con una o más sustituciones en las siguientes posiciones en la secuencia de aminoácidos establecida en SEQ ID NO: 2. 493, 494, 495, 502, 503, 508, 511, 518, 519, 520, 531, 535, 536 y 539 y/o una posición equivalente en una glucoamilasa precursora. En algunas formas de realización, las variantes incluyen una o más sustituciones en el SBD. En algunas formas de realización, las variantes de la presente exposición con actividad específica mejorada incluyen una sustitución correspondiente a la posición 495, 519, 520, 535 y 539 de SEQ ID NO:2. En algunas formas de realización, las variantes de la presente exposición con actividad específica mejorada incluyen una sustitución elegida de entre: T493C, T493M, T493N, T493Q, T493Y, P494H, P494I, P494M, P494N, P494Q, P494W, T495M, T495P, T495R, H502A, H502M, H502S, H502V, E503C, E503D, E503H, E503S, E503W, Q508N, Q508P, Q508Y, Q511C, Q511G, Q511H, Q511I, Q511K, Q511T, Q511V, N518P, N518T, A519I, A520C, A520E, A520L, A520P, A520Q, A520R, A520W, V531A, V531L, V531N, VS31R, V531S, V531T, A535E, A535F, A535G, A535K, A535L, A535N, A535P, A535R, A535S, A535T, A535V, A535W, A535Y, V536C, V536E, V536I, V536L, V536M, V536Q, V536S, A539E, A539M, A539R, A539S y A539W y/o una posición equivalente en una glucoamilasa precursora. En algunas formas de realización, las variantes de la presente exposición con una actividad específica mejorada incluyen una sustitución correspondiente a la posición T495M, A519I, A520C/L/P, A535R y A539R. En algunas formas de realización, la glucoamilasa precursora comprenderá una secuencia con al menos un 80 %, al menos un 85 %, al menos un 90 %, al menos un 95 % o 97 % de identidad de secuencia con la secuencia de SEQ ID NO:2.

**[0124]** En formas de realización adicionales, las variantes de glucoamilasa de la presente exposición con actividad específica aumentada incluyen las variantes con SBD definidas en el párrafo anterior e incluyen de forma adicional una o más deleciones, sustituciones o inserciones y especialmente sustituciones en las siguientes posiciones en la secuencia de aminoácidos establecida en SEQ ID NO:2 y/o SEQ ID NO:3. T10, L14, N15, P23, F59, K60, N61, T67, E68, A72, G73, S97, L98, A99, S102, E110, L113, I133, K140, N144, N145, Y147, S152, N153, N164, N182, A204, T205, S214, V216, Q219, W228, V229, S230, S231, D236, I239, T241, N242, N263, L264, G265, A268, G269, D276, V284, S291, P300, A301, A311, V338, T342, S344, T346, A349, V359, G361, A364, T375, N379, S382, S390, E391, A393, K394, S410, S415, L417, H418, T430, A431, R433, A442, S444, T448 y/o S451 y/o una posición equivalente en una glucoamilasa precursora. En algunas formas de realización, las variantes de la presente exposición con actividad específica mejorada incluyen una sustitución en las siguientes posiciones en la secuencia de aminoácidos establecida en SEQ ID NO:2 o SEQ ID NO:3: 61, 67, 72, 97, 102, 133, 205, 219, 228, 230, 231, 239, 263, 268, 291, 342, 394, 430, 431 y 451 y/o una posición equivalente en una glucoamilasa precursora. En algunas formas de realización, la glucoamilasa precursora comprenderá una secuencia con al menos un 50 %, al menos un 60 %, al menos un 70 %, al menos un 80 %, al menos un 90 % o 95 % de identidad de secuencia con la secuencia de SEQ ID NO:2 o SEQ ID NO:3. En algunas formas de realización, la glucoamilasa precursora también presentará identidad estructural con SEQ ID NO:2 y/o SEQ ID NO:3. En algunas formas de realización, la variante con actividad específica aumentada cuenta con una sustitución en al menos una de las posiciones: 72, 133, 219, 228, 230, 231, 239, 263, 268 y 451 de SEQ ID NO:2 y/o SEQ ID NO:3.

**Variantes con SBD tanto con termoestabilidad aumentada como con actividad específica aumentada**

**[0125]** En algunos aspectos, la presente exposición hace referencia a una glucoamilasa variante tanto con estabilidad térmica modificada como con actividad específica modificada en comparación con una de tipo silvestre o precursora. En algunas formas de realización, la actividad específica modificada es una actividad específica aumentada y la termoestabilidad modificada es una termoestabilidad aumentada en comparación con la glucoamilasa precursora.

**[0126]** En algunas formas de realización, las variantes de glucoamilasa de la presente exposición con termoestabilidad aumentada y actividad específica aumentada incluyen aquellas con una o más sustituciones, inserciones o deleciones y específicamente sustituciones en las siguientes posiciones en la secuencia de aminoácidos establecida en SEQ ID NO:2: 493, 495, 503, 508, 511, 518, 519, 520, 527, 531, 535, 536, 537, 539, 563 y 577 y/o una posición equivalente en una glucoamilasa precursora.

**[0127]** En algunas formas de realización, las variantes de glucoamilasa de la presente exposición con termoestabilidad aumentada y actividad específica aumentada incluyen aquellas con una o más sustituciones,

inserciones o deleciones y específicamente sustituciones en las siguientes posiciones en la secuencia de aminoácidos establecida en SEQ ID NO:2: 495, 503, 511, 520, 531, 535, 536 y 539 y/o una posición equivalente en una glucoamilasa precursora. En formas de realización adicionales, la variante de glucoamilasa expuesta en el presente documento con termoestabilidad aumentada y actividad específica aumentada incluye las siguientes sustituciones T495R, E503C/S, Q511H, A520C/L/P, V531L, A535K/N/P/R, V536I y A539E/M/R/S.

**[0128]** En algunas formas de realización, la glucoamilasa precursora será un homólogo de glucoamilasa *Trichoderma* y en formas de realización adicionales, la glucoamilasa precursora tendrá al menos un 90 %, al menos un 95 % y al menos un 98 % de identidad de secuencia con SEQ ID NO:2. En algunas formas de realización, la variante con termoestabilidad aumentada y actividad específica aumentada cuenta con una sustitución en al menos una de las posiciones: 503, 511, 519, 531, 535, 539, 563 y 577. En algunas formas de realización, las sustituciones se eligen de entre: T4931, T495K, T495R, T495S, E503A, E503C, E503S, E503T, E503V, Q508H, Q508R, Q508S, Q508T, Q511A, Q511D, Q511H, Q511N, Q511S, N518S, A519E, A519K, A519R, A519T, A519V, A519Y, A520C, A520L, A520P, T527A, T527V, V531L, A535D, A535K, A535N, A535P, A535R, V536I, V536R, N537W, A539E, A539H, A539M, A539R, A539S, N563A, N563C, N563E, N563I, N563K, N563L, N563Q, N563T, N563V, N577A, N577K, N577P, N577R y N577V de SEQ ID NO:2.

**[0129]** En formas de realización adicionales, las variantes de glucoamilasa de la presente exposición con actividad específica aumentada y termoestabilidad aumentada incluyen las variantes con SBD definidas en el párrafo anterior e incluyen de forma adicional una o más deleciones, sustituciones o inserciones y especialmente sustituciones en las siguientes posiciones en la secuencia de aminoácidos establecida en SEQ ID NO:2 y/o SEQ ID NO:3. T10, N15, F59, N61, E68, A72, G73, S97, A99, S102, I133, K140, N153, N182, A204, T205, S214, W228, V229, S230, S231, D236, T241, N242, L264, G265, A268, D276, V284, S291, P300, A301, A303, A311, V338, S344, T346, V359, G361, A364, T375, N379, S382, E391, A393, K394, S410, L417, T430, A431, R433, S444, T448 y/o S451 y/o una posición equivalente en una glucoamilasa precursora. En algunas formas de realización, la glucoamilasa precursora será un homólogo de glucoamilasa *Trichoderma* y en formas de realización adicionales, la glucoamilasa tendrá al menos un 50 %, al menos un 60 %, al menos un 70 %, al menos un 80 %, al menos un 90 %, al menos un 95 % y al menos un 98 % de identidad de secuencia con SEQ ID NO:2 o SEQ ID NO:3. En algunas formas de realización, la glucoamilasa precursora también presentará identidad estructural con SEQ ID NO:2 y/o SEQ ID NO:3. En algunas formas de realización, la variante con termoestabilidad y actividad específica aumentadas cuenta con una sustitución en al menos una de las posiciones: 228, 230, 231, 268, 291, 417, 433 y 451 de SEQ ID NO:2 y/o SEQ ID NO:3.

### **Polinucleótidos:**

**[0130]** La presente exposición también hace referencia a polinucleótidos aislados que codifican una glucoamilasa variante con SBD de la presente exposición. Los polinucleótidos pueden prepararse mediante técnicas consolidadas conocidas en la técnica. Los polinucleótidos pueden prepararse de forma sintética, tal como mediante un sintetizador de ADN automático. La secuencia de ADN puede ser de origen sintético y genómico combinados (o ADNc) preparada mediante la ligación conjunta de fragmentos. Los polinucleótidos pueden también prepararse mediante la reacción en cadena de la polimerasa (PCR por sus siglas en inglés) utilizando cebadores específicos. En general, se hace referencia a Minshull J., *et al.*, (2004), *Methods* 32(4):416 - 427). También determinadas empresas sintetizan actualmente ADN tal como Geneart AG, Regensburg, Alemania.

**[0131]** En algunas formas de realización, un polinucleótido aislado comprende una secuencia de nucleótidos (i) con al menos un 70 % de identidad con SEQ ID NO:4 o (ii) capaz de hibridar con una sonda derivada de una secuencia de nucleótidos establecida en SEQ ID NO:4, con condiciones de astringencia intermedia a alta, o (iii) complementario a una secuencia de nucleótidos con al menos un 90 % de identidad de secuencia con la secuencia establecida en SEQ ID NO:4. Sondas útiles según la presente exposición puede incluir al menos 50, 100, 150, 200, 250, 300 o más nucleótidos contiguos de SEQ ID NO:4.

**[0132]** La presente exposición proporciona además polinucleótidos aislados que codifican glucoamilasas variantes con una o más sustituciones de aminoácidos correspondientes a las posiciones 3, 4, 5, 12, 13, 18, 21, 28, 29, 30, 37, 41, 45, 46, 47, 49, 73 y 87 de SEQ ID NO:11 o una posición equivalente en un SBD de una glucoamilasa precursora. En otras formas de realización, la glucoamilasa precursora presentará una secuencia de aminoácidos que comprende al menos un 50 % de identidad de secuencia, al menos un 60 % de identidad de secuencia, al menos un 70 % de identidad de secuencia, al menos un 80 % de identidad de secuencia, al menos un 85 % de identidad de secuencia, al menos un 90 % de identidad de secuencia, al menos un 95 % de identidad de secuencia, al menos un 97 % de identidad de secuencia, al menos un 98 % de secuencia de aminoácidos y al menos un 99 % de identidad secuencia de aminoácidos con SEQ ID NO:11.

**[0133]** La presente exposición también proporciona fragmentos (es decir, partes) del ADN que codifica las glucoamilasas variantes proporcionadas en el presente documento. Estos fragmentos son útiles a la hora de obtener fragmentos de ADN de longitud parcial capaces de ser usados para aislar o identificar polinucleótidos

que codifican enzimas de glucoamilasa maduras descritas en el presente documento a partir de células fúngicas filamentosas (p. ej., *Trichoderma*, *Aspergillus*, *Fusarium*, *Penicillium*, y *Humicola*) o un segmento de estos con actividad de glucoamilasa. En algunas formas de realización, los fragmentos de ADN pueden comprender al menos 50, 100, 150, 200, 250 300 o más nucleótidos contiguos. En algunas formas de realización, partes del ADN proporcionado en SEQ ID NO:11 son útiles a la hora de obtener la glucoamilasa precursora y especialmente homólogos de la glucoamilasa *Trichoderma* de otras especies, tal como hongos filamentosos que codifican una glucoamilasa.

### **Construcciones de ADN y vectores**

**[0134]** Según algunas formas de realización de la presente exposición, una construcción de ADN que comprende un polinucleótido según se describe anteriormente y que codifica una glucoamilasa variante incluida en la presente exposición y unida de forma operativa a una secuencia promotora se junta para transferirse a una célula huésped.

**[0135]** La construcción de ADN puede introducirse en una célula huésped utilizando un vector. El vector puede ser cualquier vector que, cuando se introduce en una célula huésped, pueda integrarse en un genoma de célula huésped y se duplica. En algunas formas de realización, el vector se transforma de forma estable y/o integra en una célula huésped. Entre los vectores se incluyen vectores de clonación, vectores de expresión, vectores transportadores, partículas fágicas, plásmidos, casetes y similares. En algunas formas de realización, el vector es un vector de expresión que comprende secuencias reguladoras unidas de forma operativa a la secuencia que codifica la glucoamilasa.

**[0136]** Ejemplos de vectores de expresión y/o integración se presentan en Sambrook *et al.*, (1989) *supra*, y Ausubel (1987) *supra*, y van den Hondel *et al.* (1991) en Bennett y Lasure (Eds.) *MORE GENE MANIPULATIONS IN FUNGI*, Academic Press pp. 396-428 y la patente estadounidense No. 5,874,276. También se hace referencia al catálogo de cepas del Centro de Almacenamiento genético fúngico (FGSC, < www.fgsc.net>) para un listado de vectores. Entre los vectores especialmente útiles se incluyen vectores obtenidos de por ejemplo Invitrogen y Promega.

**[0137]** Plásmidos adecuados para utilizarse en células bacterianas incluyen pBR322 y pUC19 que permiten la replicación en *E.coli* y pE194 por ejemplo que permite la replicación en *Bacillus*. Otros vectores específicos adecuados para utilizarse en células huésped *E. coli* incluyen vectores tales como pFB6, pBR322, pUC18, pUC100, pDONR™201, pDONR™221, pENTR™, pGEM@3Z y pGEM@4Z.

**[0138]** Vectores específicos adecuados para utilizarse en células fúngicas incluyen pRAX, un vector de expresión con fin general útil en *Aspergillus*, pRAX con un promotor *glaA*, y en *Hypocrea/Trichoderma* incluye pTrex3g con un promotor *cbh1*.

**[0139]** En algunas formas de realización, el promotor muestra actividad transcripcional en una célula huésped fúngica o bacteriana y puede derivarse de genes que codifican proteínas ya sean homólogas o heterólogas a la célula huésped. El promotor puede ser un promotor híbrido, truncado y/o mutante. Los promotores mencionados anteriormente son conocidos en la técnica. Ejemplos de promotores útiles en células fúngicas y especialmente en células fúngicas filamentosas tal como *Trichoderma* o *Aspergillus* incluyen tales promotores de ejemplo como los promotores *T. reesei cbh1*, *cbh2*, *egl1*, *egl2*, *eg5*, *xln1* y *xln2*. Otros ejemplos de promotores útiles incluyen promotores de genes de glucoamilasa *A. awamori* y *A. niger* (*glaA*) (Véase, Nunberg *et al.*, (1984) *Mol. Cell Biol.* 4:2306-2315 y Boel *et al.*, (1984) *EMBO J.* 3:1581-1585), promotor de TAKA amilasa *A. oryzae*, el promotor TPI (triosa fosfato isomerasa) de *S. cerevisiae*, el promotor de genes acetamidasa *Aspergillus nidulans* y genes lipasa *Rhizomucor miehei*. Ejemplos de promotores adecuados útiles en células bacterianas incluyen aquellos obtenidos del operón *lac E. coli*; gen alfa-amilasa *Bacillus licheniformis* (*amyL*), gen amilasa *B. stearothermophilus* (*amyS*); genes *xyIA* y *xyIB Bacillus subtilis*, el gen beta-lactamasa y el promotor *tac*. En algunas formas de realización, el promotor es uno que es natural de la célula huésped. Por ejemplo, cuando el huésped es *T. reesei*, el promotor es un promotor natural de *T. reesei*. En otras formas de realización, el promotor es uno que es heterólogo a la célula huésped fúngica. En algunas formas de realización, el promotor será el promotor de la glucoamilasa precursora (p. ej., el promotor de TrGA).

**[0140]** En algunas formas de realización, la construcción de ADN incluye ácidos nucleicos que codifican una secuencia señal, es decir, una secuencia de aminoácidos enlazada al amino terminal del polipéptido que dirige al polipéptido codificado a la trayectoria secretora de la célula. El extremo 5' de la secuencia codificante de la secuencia de ácido nucleico puede incluir de forma natural una región codificante de péptido señal que está enlazada de forma natural en el marco de lectura de traducción con el segmento de la secuencia codificante de glucoamilasa que codifica la glucoamilasa secretada o el extremo 5' de la secuencia codificante de la secuencia de ácido nucleico puede incluir un péptido señal que es ajeno a la secuencia codificante. En algunas formas de realización, la construcción de ADN incluye una secuencia señal que está asociada de forma natural a un gen de glucoamilasa precursora a partir del cual se obtiene la glucoamilasa variante. En algunas formas de realización,

la secuencia señal será la secuencia representada en SEQ ID NO:1 o una secuencia con al menos un 90 %, al menos un 94 % y al menos un 98 % de identidad de secuencia con la misma. Secuencias señales eficaces pueden incluir las secuencias señales obtenidas de otras enzimas fúngicas filamentosas, tal como de *Trichoderma* (glucoamilasa *T. reesei*), *Humicola* (celulasa *H. insolens* o glucoamilasa *H. grisea*), *Aspergillus* (glucoamilasa *A. niger* y TAKA amilasa *A. oryzae*) y *Rhizopus*.

**[0141]** En formas de realización adicionales, una construcción de ADN o vector que comprende una secuencia señal y una secuencia promotora que han de introducirse en una célula huésped se derivan de la misma fuente. En algunas formas de realización, se puede utilizar la secuencia señal de glucoamilasa natural de un homólogo de glucoamilasa *Trichoderma*, tal como una secuencia señal de una cepa de *Hypocrea*.

**[0142]** En algunas formas de realización, el vector de expresión también incluye una secuencia de terminación. Puede utilizarse cualquier secuencia de terminación funcional en la célula huésped. En algunas formas de realización, la secuencia de terminación y la secuencia promotora se derivan de la misma fuente. En otra forma de realización, la secuencia de terminación es homóloga a la célula huésped. Entre las secuencias de terminación útiles se incluyen secuencias de terminación obtenidas a partir de los genes de *Trichoderma reesei* *cbh1*; glucoamilasa *A. niger* o *A. awamori* (Nunberg et al. (1984) *supra*, y Boel et al., (1984) *supra*), antranilato sintasa *Aspergillus nidulans*, Taka amilasa *Aspergillus oryzae*, o *A. nidulans trpC* (Punt et al., (1987) *Gene* 56:117 - 124).

**[0143]** En algunas formas de realización, el vector de expresión incluye un marcador seleccionable. Entre los ejemplos de marcadores seleccionables se incluyen unos que confieren resistencia antimicrobiana (p. ej., higromicina y fleomicina). Los marcadores selectivos nutricionales también son útiles incluyendo aquellos marcadores conocidos en la técnica como *amds* (acetamidasa), *argB* (ornitina carbamoiltransferasa) y *pyrG* (orotidina-5' fosfato descarboxilasa). Los marcadores útiles en sistemas de vectores para la transformación de *Trichoderma* son conocidos en la técnica (Véase, p. ej., Finkelstein, capítulo 6 en *BIOTECHNOLOGY OF FILAMENTOUS FUNGI*, Finkelstein et al. Eds. Butterworth-Heinemann, Boston, MA (1992); Kinghorn et al. (1992) *APPLIED MOLECULAR GENETICS OF FILAMENTOUS FUNGI*, Blackie Academic and Professional, Chapman and Hall, London; Berges y Barreau (1991) *Curr. Genet.* 19:359-365; y van Hartingsveldt et al., (1987) *Mol. Gen. Genet.* 206:71 - 75). En algunas formas de realización, el marcador selectivo es el gen *amdS*, que codifica la enzima acetamidasa, lo que permite que las células transformadas crezcan en acetamida como una fuente de nitrógeno. El uso del gen *amdS* *A. nidulans* como marcador selectivo se describe en Kelley et al., (1985) *EMBO J.* 4:475 - 479 y Penttilä et al., (1987) *Gene* 61:155-164.

**[0144]** Los métodos utilizados para ligar la construcción de ADN que comprende una secuencia de ácido nucleico que codifica una glucoamilasa variante, un promotor, una terminación y otras secuencias y para insertarlos en un vector adecuado son conocidos en la técnica. El enlace se consigue generalmente mediante la ligación en los lugares de restricción convenientes. Si existen estos lugares, los enlazadores de oligonucleótidos sintéticos se utilizan según la práctica convencional. (Véase, Sambrook (1989) *supra*, y Bennett y Lasure, *MORE GENE MANIPULATIONS IN FUNGI*, Academic Press, San Diego (1991) pp 70-76.). De forma adicional, los vectores pueden construirse utilizando técnicas de recombinación conocidas (p. ej., Invitrogen Life Technologies, Gateway Technology).

#### **Células huésped:**

**[0145]** Algunas formas de realización de la presente exposición incluyen células huésped que comprenden un polinucleótido que codifica una glucoamilasa variante de la presente exposición. En algunas formas de realización, las células huésped se eligen de células bacterianas, fúngicas, vegetales y de levadura. El término célula huésped incluye tanto las células, la progenie de las células y los protoplastos creados a partir de las células que se utilizan para producir una glucoamilasa variante según la presente exposición.

**[0146]** En algunas formas de realización, las células huésped son células fúngicas. En algunas formas de realización, las células huésped son células huésped fúngicas filamentosas. El término "hongos filamentosos" hace referencia a las formas filamentosas de la subdivisión Eumycotina (Véase Alexopoulos, C. J. (1962), *INTRODUCTORY MYCOLOGY*, Wiley, New York). Estos hongos se caracterizan por un micelio vegetativo con una pared celular compuesta de quitina, celulosa y otros polisacáridos complejos. Los hongos filamentosos de la presente exposición son distintos morfológica, fisiológica y genéticamente de las levaduras. El crecimiento vegetativo por hongos filamentosos se realiza mediante elongación de las hifas y el catabolismo de carbono es aerobio obligado. En la presente invención, la célula precursora fúngica filamentosa puede ser una célula de una especie de entre, sin carácter limitativo, *Trichoderma*, (p. ej., *Trichoderma reesei*), la mutación asexual de *Hypocrea jecorina*, clasificada anteriormente como *T. longibrachiatum*, *Trichoderma viride*, *Trichoderma koningii*, *Trichoderma harzianum*) (Sheir-Neirs et al., (1984) *Appl. Microbiol. Biotechnol* 20:46 - 53; ATCC No. 56765 y ATCC No. 26921); *Penicillium* sp., *Humicola* sp. (p. ej., *H. insolens*, *H. lanuginosa* y *H. grisea*); *Chrysosporium* sp. (p. ej., *C. lucknowense*), *Gliocladium* sp., *Aspergillus* sp. (p. ej., *A. oryzae*, *A. niger*, *A. sojae*, *A. japonicus*, *A. nidulans*, y *A. awamori*) (Ward et al., (1993) *Appl. Microbiol. Biotechnol.* 39:738 - 743 y Goedegebuur et al.,

(2002) Genet 41:89 -98), *Fusarium sp.*, (p. ej., *F. roseum*, *F. graminum* *F. cerealis*, *F. oxysporum* y *F. venenatum*), *Neurospora sp.*, (*N. crassa*), *Hypocrea sp.*, *Mucor sp.*, (*M. miehei*), *Rhizopus sp.* y *Emericella sp.* (Véase también, Innis *et al.*, (1985) *Sci.* 228:21 -26). Los términos "*Trichoderma*" o "*Trichoderma sp.*" o "*Trichoderma spp.*" hacen referencia a cualquier género fúngico clasificado previamente o actualmente como *Trichoderma*.

5 [0147] En algunas formas de realización, las células huésped serán células bacterianas gram positivas. Entre los ejemplos sin carácter limitativo se incluyen cepas de *Streptomyces*, (p. ej., *S. lividans*, *S. coelicolor* y *S. griseus*) y *Bacillus*. Tal y como se utiliza en el presente documento, "el género *Bacillus*" incluye todas las especies dentro del género "*Bacillus*", según lo conocen los expertos en la técnica, entre las que se incluyen sin carácter limitativo *B. subtilis*, *B. licheniformis*, *B. lentus*, *B. brevis*, *B. stearothermophilus*, *B. alkalophilus*, *B. amyloliquefaciens*, *B. clausii*, *B. halodurans*, *B. megaterium*, *B. coagulans*, *B. circulans*, *B. lautus*, y *B. thuringiensis*. Se reconoce que el género *Bacillus* continúa experimentando reorganización taxonómica. Por lo tanto, se pretende que el género incluya especies que han sido reclasificadas, incluyendo sin carácter limitativo tales organismos como *B. stearothermophilus*, que ahora se denomina "*Geobacillus stearothermophilus*".

10 [0148] En algunas formas de realización, la célula huésped es una cepa bacteriana gram negativa, tal como *E. coli* o *Pseudomonas sp.* En otras formas de realización, las células huésped pueden ser células de levadura tal como *Saccharomyces sp.*, *Schizosaccharomyces sp.*, *Pichia sp.* o *Candida sp.*

15 [0149] En otras formas de realización, la célula huésped será una célula huésped diseñada genéticamente en la que los genes naturales se han desactivado, por ejemplo, mediante delección en células bacterianas o fúngicas. Cuando se desee obtener una célula huésped fúngica con uno o más genes desactivados, se pueden utilizar métodos conocidos (p. ej., los métodos expuestos en la patente estadounidense N.º. 5.246.853, patente estadounidense N.º 5.475.101 y WO 92/06209). La desactivación génica puede conseguirse mediante una delección completa o parcial, mediante desactivación por inserción o mediante cualquier otro medio que proporcione un gen no funcional para el fin destinado (de forma que se impida la expresión del gen a partir de una proteína funcional). En algunas formas de realización, cuando la célula huésped es una célula *Trichoderma* y especialmente una célula huésped *T. reesei*, los genes *cbh1*, *cbh2*, *egl1* y *egl2* se desactivarán y podrán eliminarse. En algunas formas de realización, las células huésped *Trichoderma reesei* tienen codificación génica con eliminación cuádruple para proteínas tales como aquellas establecidas y descritas en la patente estadounidense 5.847.276 y WO 05/001036. En otras formas de realización, la célula huésped es una cepa carente de proteasa o proteasa negativa.

### 30 **Transformación de células huésped:**

[0150] La introducción de una construcción de ADN o vector en una célula huésped incluye técnicas tales como transformación, electroporación, microinyección nuclear, transducción, transfección (p. ej., transfección mediada por lipofección y mediada por DEAE-Dextrina); incubación con precipitado de ADN fosfato calcio; bombardeo de alta velocidad con microproyectiles revestidos de ADN; y fusión de protoplastos. Las técnicas de transformación generales se conocen en la técnica (Véase, p. ej., Ausubel *et al.*, (1987), *supra*, capítulo 9; y Sambrook (1989) *supra*, y Campbell *et al.*, (1989) *Curr. Genet.* 16:53-56).

[0151] Los métodos de transformación para *Bacillus* se exponen en numerosas referencias entre las que se incluyen Anagnostopoulos C y J. Spizizen (1961) *J. Bacteriol.* 81:741 - 746 y WO 02/14490.

40 [0152] Los métodos de transformación para *Aspergillus* se describen en Yelton *et al.* (1984) *Proc. Natl. Acad. Sci.* EEUU 81:1470-1474; Berka *et al.*, (1991) en *Applications of Enzyme Biotechnology*, Eds. Kelly and Baldwin, Plenum Press (NY); Cao *et al.*, (2000) *Sci.* 9:991 - 1001; Campbell *et al.*, (1989) *Curr. Genet.* 16:53-56 y EP 238 023. La expresión de proteína heteróloga en *Trichoderma* se describe en la patente estadounidense 6.022.725; patente estadounidense 6.268.328; Harkki *et al.* (1991); *Enzyme Microb. Technol.* 13:227-233; Harkki *et al.*, (1989) *Bio Technol.* 7:596-603; EP 244.234; EP 215.594; y Nevalainen *et al.*, "The Molecular Biology of *Trichoderma* and its Application to the Expression of Both Homologous and Heterologous Genes", en *MOLECULAR INDUSTRIAL MYCOLOGY*, Eds. Leong y Berka, Marcel Dekker Inc., NY (1992) pp. 129 - 148). También se hace referencia a WO96/00787 y Bajar *et al.* (1991) *Proc. Natl. Acad. Sci.* EEUU 88-8202 - 28212 para la transformación de cepas *Fusarium*.

45 [0153] En una forma de realización específica, la preparación de *Trichoderma sp.* para la transformación implica la preparación de protoplastos a partir de micelios fúngicos (Véase, Campbell *et al.* (1989) *Curr. Genet.* 16:53-56; Pentilla *et al.*, (1987) *Gene* 61:155 - 164). La transformación mediada por *Agrobacterium tumefaciens* de hongos filamentosos es conocida. (Véase, de Groot *et al.*, (1998) *Nat. Biotechnol.* 16:839 - 842). También se hace referencia a la patente estadounidense 6.022.725 y a la patente estadounidense 6.268.328 para la transformación de procedimientos utilizados con huéspedes fúngicos filamentosos.

50 [0154] En algunas formas de realización, los transformantes estables genéticamente se construyen con sistemas de vectores por los que el ácido nucleico que codifica la glucoamilasa variante se integra de forma

estable en un cromosoma de la cepa huésped. A continuación, se purifican los transformantes con técnicas conocidas.

5 **[0155]** En algunas formas de realización adicionales, las células huésped son células vegetales, tal como células de una planta monocotiledónea (p. ej., maíz, trigo y sorgo) o células de una planta dicotiledónea (p. ej., soja). Se conocen los métodos para realizar construcciones de ADN útiles para la transformación de plantas y métodos para la transformación vegetal. Algunos de estos métodos incluyen transferencia génica mediada por *Agrobacterium tumefaciens* bombardeo de microproyectiles, transformación de protoplastos mediada por PEG, electroporación y similares. Se hace referencia a la patente estadounidense 6.803.499, la patente estadounidense 6.777.589; Fromm *et al.* (1990) *Biotechnol.* 8:833-839; Potrykus *et al.* (1985) *Mol.Gen. Genet.* 199:169 - 177).

#### **Producción de proteínas:**

15 **[0156]** En algunas formas de realización, la presente exposición se dirige a métodos para producir las variantes de glucoamilasa en una célula huésped. En algunas formas de realización, el método comprende la transformación de una célula huésped con un vector que comprende un polinucleótido que codifica una glucoamilasa variante según la presente exposición, cultivando de forma opcional la célula huésped según condiciones adecuadas para la expresión y producción de la variante de glucoamilasa y recuperando de forma opcional la variante. Entre las condiciones adecuadas de cultivo se incluyen sin carácter limitativo el cultivo con agitación del matraz, fermentaciones a pequeña o gran escala (incluyendo fermentaciones continuas, por lotes y discontinuas) en fermentadores industriales o de laboratorio, con el medio adecuado que contenga sales fisiológicas y nutrientes (Véase, p. ej., Pourquie, J. *et al.*, *BIOCHEMISTRY AND GENETICS OF CELLULOSE DEGRADATION*, eds. Aubert, J. P. *et al.*, Academic Press, pp. 71-86, 1988 e Ilmen, M. *et al.*, (1997) *Appl. Environ. Microbiol.* 63:1298-1306). Los medios comunes preparados comercialmente (p. ej., caldo de extracto de malta levadura, caldo Luria Bertani y el caldo Sabouraud Dextrosa) son útiles en la presente exposición. Las condiciones de cultivo para células fúngicas y filamentosas son conocidas en la técnica y pueden hallarse en la literatura científica y/o desde la fuente de los hongos tal como American Type Culture Collection y Fungal Genetics Stock Center. En los casos en los que una secuencia codificante de glucoamilasa esté bajo el control de un promotor inducible, el agente inductor (p. ej., un azúcar, una sal metálica o antimicrobiano) se añade al medio con una concentración eficaz para inducir la expresión de glucoamilasa.

30 **[0157]** En algunas formas de realización, los ensayos se llevan a cabo para evaluar la expresión de una variante de glucoamilasa mediante una línea celular que ha sido transformada con un polinucleótido que codifica una variante incluida en la presente exposición. Los ensayos pueden llevarse a cabo en el nivel de proteína, el nivel de ARN y/o mediante el uso de bioanálisis funcionales específicos de la actividad y/o producción de glucoamilasa. Algunos de estos ensayos incluyen la prueba de Northern blot, dot blot (análisis de ADN o ARN), RT-PCR (reacción en cadena de la polimerasa con transcriptasa inversa), hibridación *in situ* utilizando una sonda marcada de forma adecuada (según la secuencia codificante de ácido nucleico) y la prueba de Southern blot convencional y autorradiografía.

40 **[0158]** Además, la producción y/o expresión de una glucoamilasa variante de la presente exposición puede medirse en una muestra directamente, por ejemplo, mediante ensayos que miden directamente los azúcares reductores tal como glucosa en el medio de cultivo y mediante ensayos que miden la actividad, expresión y/o producción de glucoamilasa. En concreto, la actividad de glucoamilasa puede analizarse mediante el método del ácido 3,5-dinitrosalicílico (DNS) (Véase, Goto *et al.*, (1994) *Biosci. Biotechnol. Biochem.* 58:49 - 54). En formas de realización adicionales, la expresión de proteína se evalúa mediante métodos inmunológicos, tal como la coloración inmunohistoquímica de células, secciones de tejido o inmunoanálisis de un medio de cultivo de tejido (p. ej., mediante Western blot o ELISA). Tales inmunoanálisis pueden utilizarse para evaluar de forma cualitativa y cuantitativa la expresión de una glucoamilasa. Los expertos en la técnica conocen los detalles de tales métodos y muchos reactivos para practicar tales métodos están disponibles comercialmente.

**[0159]** Las glucoamilasas variantes con SBD de la presente exposición pueden recuperarse o purificarse del medio de cultivo mediante una variedad de procedimientos conocidos en la técnica entre los que se incluyen la centrifugación, la filtración, la extracción, la precipitación y similares.

#### 50 **Composiciones:**

**[0160]** Las variantes de glucoamilasa pueden utilizarse en composiciones de enzima entre las que se incluyen sin carácter limitativo composiciones para la hidrólisis y sacarificación de almidón, composiciones de limpieza y detergente (p. ej., detergentes para la colada, detergentes para lavar los platos y composiciones de limpieza de superficies duras), composiciones para la fermentación del alcohol, composiciones para pienso animal y composiciones de bebidas. Además, las variantes pueden utilizarse en aplicaciones de cocción, tal como la producción de pan y pasteles, destilación, atención sanitaria, textil, procepción de conversión de residuos medioambientales, procesamiento de biopulpeo y aplicaciones de conversión de biomasa.



[0161] En algunas formas de realización, se obtendrá una composición de enzima que incluye una variante con SBD incluida en la presente exposición en un medio de cultivo o se recuperará y purificará del medio de cultivo. En algunas formas de realización, la variante con SBD se utilizará junto con cualquier de las siguientes enzimas o una combinación de estas - alfa-amilasas, proteasas, pululanasa, isoamilasas, celulasas, hemicelulasas, xilanasas, ciclodextrina glicos transferasas, lipasas, fitasas, laccasas, oxidasas, esterases, cutinasas, xilanasas, enzima hidrolizante de almidón granular y otras glucoamilasas.

[0162] En algunas composiciones, la variante con SBD de la presente exposición se combinará con alfa-amilasa, tal como alfa-amilasas fúngicas (p. ej., *Aspergillus sp.*) o alfa-amilasas bacterianas (p. ej., *Bacillus sp.* tal como *B. stearothermophilus*, *B. amyloliquefaciens* y *B. licheniformis*) y variantes e híbridos de estas. En algunas formas de realización, el alfa-amilasa es un alfa-amilasa estable ácida. En algunas formas de realización, la alfa-amilasa es una enzima hidrolizante de almidón granular (GSHE). En algunas formas de realización, la alfa-amilasa es alfa-amilasa *Aspergillus kawachi* (AKAA), véase US 7.037.704. Las alfa-amilasas disponibles comercialmente contempladas para utilizarse en las composiciones de la presente exposición son conocidas e incluyen GZYME G997, SPEZYME FRED, SPEZYME XTRA, STARGEN (Danisco US, Inc, Genencor Division), TERMAMYL 120-L y SUPRA (Novozymes, Biotech.) y VIRIDIUM (Diversa).

[0163] En algunas formas de realización, las proteasas son proteasas fúngicas ácidas. En una forma de realización adicional, las proteasas fúngicas ácidas son de *Trichoderma* (p. ej., NSP-24, véase también US 2006/015342, publicada el 7/13/2006, SEQ ID NO:10).

[0164] En otras formas de realización, las variantes de glucoamilasa de la presente exposición pueden combinarse con otras glucoamilasas. En algunas formas de realización, las glucoamilasas de la presente exposición se combinarán con una o más glucoamilasas derivadas de cepas de *Aspergillus* o variantes de esta, tal como *A. oryzae*, *A. niger*, *A. kawachi*, y *A. awamori*; glucoamilasas derivadas de cepas de *Humicola* o variantes de esta, especialmente *H. grisea*, tal como la glucoamilasa con al menos un 90 %, 93 %, 95 %, 96 %, 97 %, 98 % y 99% de identidad de secuencia con SEQ ID NO:3 expuesta en WO 05/052148; glucoamilasas derivadas de cepas de *Talaromyces* o variantes de esta, especialmente *T. emersonii*; glucoamilasas derivadas de cepas de *Athelia* y especialmente *A. rolfsii*; glucoamilasas derivadas de cepas de *Penicillium*, especialmente *P. chrysogenum*.

#### Usos:

[0165] En concreto, las variantes con SBD pueden utilizarse para los procesos de conversión de almidón, y especialmente en la producción de dextrosa para jarabes de fructosa, azúcares especializados y en alcohol (p. ej., etanol y butanol) y la producción de otros productos finales (p. ej., ácido orgánico, ácido ascórbico y aminoácidos) a partir de la fermentación de sustratos que contienen almidón (G.M.A van Beynum *et al.*, Eds. (1985) *STARCH CONVERSION TECHNOLOGY*, Marcel Dekker Inc. NY). Las dextrinas producidas utilizando composiciones de variante de glucoamilasa de la presente exposición pueden tener como resultado una producción de glucosa de al menos un 80 %, al menos un 85 %, al menos un 90 % y al menos un 95 %. La producción de alcohol a partir de la fermentación de sustratos de almidón con glucoamilasas de la presente exposición puede incluir la producción de carburante alcohólico y alcohol de boca. En algunas formas de realización, la producción de alcohol será mayor cuando la variante de glucoamilasa se utilice con las mismas condiciones que la glucoamilasa precursora. En algunas formas de realización, la producción de alcohol será entre aproximadamente un 0,5 % y 2,5 % mayor, incluyendo sin carácter limitativo 0,6 %; 0,7 %; 0,8 %; 0,9 %; 1,0 %; 1,1 %; 1,2 %; 1,3 %; 1,4 %; 1,5 %; 1,6 %; 1,7 %; 1,8 %; 1,9 %; 2,0 %; 2,1 %; 2,2 %; 2,3 % y 2,4 % más de alcohol que la glucoamilasa precursora con las mismas condiciones básicamente.

[0166] En algunas formas de realización, las glucoamilasas variantes con SBD de la presente exposición serán útiles en la hidrólisis de almidón de diferentes sustratos con base vegetal, que se utilizan para la producción de alcohol. En algunas formas de realización, los sustratos con base vegetal incluirán maíz, trigo, cebada, centeno, sorgo, arroz, caña de azúcar, patatas y combinaciones de estos. En algunas formas de realización, el sustrato con base vegetal se fraccionará en material vegetal, por ejemplo, un grano de cereal, tal como maíz, que se fracciona en componentes tal como fibra, germen, proteína y almidón (endospermo) (patente estadounidense 6.254.914 y patente estadounidense 6.899.910). Los métodos para las fermentaciones de alcohol se describen en *THE ALCOHOL TEXTBOOK, A REFERENCE FOR THE BEVERAGE, FUEL AND INDUSTRIAL ALCOHOL INDUSTRIES*, 3ª Ed., Eds K.A. Jacques *et al.*, 1999, Nottingham University Press, Reino Unido. En determinadas formas de realización, el alcohol será etanol. En concreto, los procesos de producción de fermentación de alcohol se caracterizan como procesos de molienda húmeda o seca. En algunas formas de realización, la glucoamilasa variante se utilizará en un proceso de fermentación de molienda húmeda y en otras formas de realización la glucoamilasa variante será útil en un proceso de molienda seca.

[0167] La molienda de grano húmeda implica un número de etapas básicas, que incluyen generalmente: triturado, cocción, licuefacción, sacarificación, fermentación y separación de líquidos y sólidos con el fin de producir alcohol y otros subproductos. Se tritura material vegetal y especialmente granos de cereales integrales,

tal como maíz, trigo o centeno. En algunos casos, el grano se puede fraccionar primero en partes componentes. El material vegetal triturado puede molerse para obtener una partícula fina o gruesa. El material vegetal triturado se mezcla con líquido (p. ej., agua y/o finos residuos de la destilería) en un tanque de suspensión. La suspensión está sujeta a altas temperaturas (p. ej., de 90 °C a 105 °C o superiores) en un cocedor a chorro junto con enzimas de licuefacción (p. ej., alfa-amilasas) para solubilizar e hidrolizar el almidón en el grano a dextrinas. La mezcla se enfría y se trata además con enzimas de sacarificación, tal como glucoamilasas incluidas en la presente exposición, para producir glucosa. A continuación, la mezcla que contiene glucosa puede fermentarse durante aproximadamente 24 a 120 horas en presencia de microorganismos de fermentación, tal como microorganismos que producen etanol y especialmente levadura (*Saccharomyces spp*). Los sólidos en la mezcla se separan de la fase líquida y se obtiene alcohol tal como etanol y subproductos útiles tal como granos de destilación.

**[0168]** En algunas formas de realización, la etapa de sacarificación y la etapa de fermentación se combinan y se hace referencia al proceso como sacarificación y fermentación simultáneas o sacarificación, propagación de levadura y fermentación simultáneas.

**[0169]** En otras formas de realización, la glucoamilasa variante con SBD se utiliza en un proceso para la hidrólisis de almidón en el que la temperatura del proceso está entre 30 °C y 75 °C en algunas formas de realización, entre 40 °C y 65 °C. En algunas formas de realización, la glucoamilasa variante se utiliza en un proceso para la hidrólisis de almidón con un pH de entre pH 3,0 y pH 6,5. Los procesos de fermentación en algunas formas de realización incluyen la molienda de un grano de cereal o grano fraccionado y la combinación del grano de cereal molido con líquido con el fin de formar una suspensión que a continuación se mezcla en un único recipiente con una glucoamilasa variante según la presente exposición y de forma opcional otras enzimas tal como, sin carácter limitativo, alfa-amilasas, otras glucoamilasas, fitasas, proteasas, pululanases, isoamilasas u otras enzimas con actividad hidrolizante de almidón granular y levadura con el fin de producir etanol y otros subproductos (véase, p. ej., la patente estadounidense 4.514.496, WO 04/081193 y WO 04/080923).

**[0170]** En algunas formas de realización, la presente exposición pertenece a un método para sacarificar una solución de almidón líquida, que comprende una etapa de sacarificación enzimática con una glucoamilasa variante de la presente exposición.

**[0171]** La presente exposición también presenta una composición de pienso animal o formulación que comprende al menos una glucoamilasa variante con SBD incluida en la presente exposición. Los métodos para usar una enzima de glucoamilasa variante con SBD en la producción de piensos que comprenden almidón se presentan en WO 03/049550, presentada el 13 de diciembre, 2002. En resumen, la variante de glucoamilasa con SBD se mezcla con un pienso que comprende almidón. La variante de glucoamilasa con SBD es capaz de degradar almidón resistente para ser usado por el animal.

## **EXPERIMENTAL**

**[0172]** En la exposición y la sección experimental que se presenta a continuación, se aplican las siguientes abreviaturas: GA (glucoamilasa); GAU (unidad glucoamilasa); % p (porcentaje en peso); °C (grados centígrados); rpm (revoluciones por minuto); H<sub>2</sub>O (agua); dH<sub>2</sub>O (agua desmineralizada), dH<sub>2</sub>O (agua desmineralizada, filtración Milli-Q) aa o AA (aminoácido); pb (pares de bases); kb (pares de kilobases); kD (kilodaltons); g o gm (gramos); µg (microgramos); mg (miligramos); µL (microlitros); ml y mL (mililitros); mm (milímetros), µm (micrómetro); M (molar); mM (milimolar); µM (micromolar); U (unidades); V (voltios); PM (peso molecular); s (segundo/segundos); min(s) o m(s) (minuto/minutos); hr(s) o h(s) (hora/horas); OD (oxígeno disuelto); ABS (absorbancia); EtOH (etanol); SF (suero fisiológico); m/v (masa/volumen) y PMT (placa microtitulación); N (Normal); DP1 (monosacáridos); DP2 (disacáridos); DP>3 (oligosacáridos, azúcares con un grado de polimerización superior a 3); ppm (partes por millón).

**[0173]** Se utilizaron los siguientes análisis y métodos en los ejemplos presentados a continuación. Sin embargo, cabe destacar que pueden utilizarse diferentes métodos para proporcionar variantes de una molécula precursora y la invención no está limitada a los métodos utilizados en los ejemplos. Se pretende que se pueda utilizar cualquier medio adecuado para realizar variantes y para la selección de variantes.

### **Análisis de la actividad glucoamilasa pNPG para placas de microtitulación de 96 pocillos:**

**[0174]** Las soluciones de reactivos fueron: tampón NaAc (200 mM de tampón de acetato de sodio pH 4,5); sustrato (50 mM p-nitro-fenil- $\alpha$ -D- glucopiranosido (Sigma N-1377) en tampón NaAc (0,3 g/20 ml) y solución de cese de reacción (800 mM tampón glicina-NaOH pH 10). Se colocaron 30 µl de sobrenadante filtrado en una PMT nueva con base plana de 96 pocillos. Se añadieron a cada pocillo 50 µl de tampón NaAc y 120 µl de sustrato y se incubaron durante 30 minutos a 50 °C (incubadora/agitador HT iEMS de Thermolab systems). Se finalizó la reacción añadiendo 100 µl de solución de cese de reacción. La absorbancia se midió a 405 nm en un lector de PMT (Spectramax 384 plus de Molecular Devices) y la actividad se calculó utilizando un coeficiente de

extinción molar de 0,011  $\mu\text{M}/\text{cm}$ .

#### **Análisis de termoestabilidad:**

- 5 **[0175]** Se añadió sobrenadante crudo (8  $\mu\text{l}$ ) a 280  $\mu\text{l}$  50 mM de tampón NaAc pH 4,5. La muestra diluida se dividió equitativamente sobre 2 PMT. Una PMT (placa inicial) se incubó durante 1 hr a 4  $^{\circ}\text{C}$  y la otra PMT (placa residual) se incubó a 64  $^{\circ}\text{C}$  (incubadora/agitador HT iEMS de Thermolab systems) durante 1 hr. La placa residual se enfrió durante 10 min en hielo. Se midió la actividad de ambas placas utilizando el análisis de detección de etanol descrito a continuación. Se añadieron 60  $\mu\text{l}$  de la placa inicial o placa residual a 120  $\mu\text{l}$  4 % almidón de maíz soluble y se incubaron durante 2 horas a 32  $^{\circ}\text{C}$  900 rpm (incubadora/agitador HT iEMS de Thermolab systems).
- 10 **[0176]** Se calculó la termoestabilidad como % de actividad residual como se muestra a continuación:

$$\frac{\text{ABS (340) residual} - \text{vacío}}{\text{ABS (340) inicial} - \text{vacío}} \times 100\%$$

- 15 **[0177]** Se analizó el material sobrenadante crudo para detectar la glucosa restante en el medio de cultivo tras el periodo de crecimiento. No se calculó la termoestabilidad si se hallaba glucosa restante en el medio de cultivo. Según la actividad residual del WT y mutante, se calculó el índice de rendimiento (PI) para detectar la termoestabilidad. El PI de una variante es el cociente "actividad residual variante/actividad residual WT". El PI de WT es 1,0 y una variante con un PI > 1,0 tiene una actividad específica que es mayor que WT.

#### **Análisis de datos y cálculo del índice de rendimiento del análisis de aplicación de etanol**

- 20 **[0178]** Se midieron los niveles de proteína utilizando un instrumento de electroforesis de microfluidos (Caliper Life Sciences, Hopkinton, MA, EEUU). Se prepararon las muestras de proteína y el chip de microfluidos según las instrucciones del fabricante (LabChip® HT Protein Express, P/N 760301). Se prepararon los sobrenadantes de cultivo y se almacenaron en placas de microtitulación de 96 pocillos a -20  $^{\circ}\text{C}$  hasta que se utilizaron, momento en el que se descongelaron mediante el calentamiento en un incubadora a 37  $^{\circ}\text{C}$  durante 30 minutos. Después de una breve agitación, se transfirieron 2  $\mu\text{l}$  de cada muestra de cultivo a una placa de PCR de 96 pocillos (Bio-Rad, Hercules, CA, EEUU) con muestras de 7  $\mu\text{l}$  de solución amortiguadora (Caliper) seguido del calentamiento de la placa a 90  $^{\circ}\text{C}$  durante 5 minutos sobre una calentador de placa controlado de forma termoestática. Se dejó que la placa se enfriara antes de añadir 40  $\mu\text{l}$  de agua a cada muestra. La placa se colocó en el instrumento junto con un ejemplo de proteína suministrado y calibrado por el fabricante. Cuando las proteínas pasan por un punto focal en el chip, la señal fluorescente se registró y se determinó la concentración de proteína mediante la cuantificación de la señal en relación a la señal generada por el conjunto calibrado de ejemplos de proteína.
- 25 **[0179]** Después de la determinación de proteína Caliper, se procesaron los datos de la siguiente forma. Se comprobaron las escaleras de calibración para la exactitud del patrón pico. Si la escalera de calibración que estaba asociada a la secuencia no era suficiente, se reemplazaba con una escalera de calibración de una secuencia adyacente. Para la detección del pico, se empleó la configuración por defecto de la opción para hallar el pico global del software caliper. El pico de interés se eligió a 75 kDa +/- 10 %.
- 30 **[0180]** El resultado se exportó a un programa de hojas de cálculo y el área pico se relacionó con la actividad correspondiente (ABS340-medición en blanco) en el análisis de aplicación de etanol. Con los números de actividad y el área de 12 muestras de tipo silvestre, se realizó una línea de calibración utilizando la ecuación "Cinética de enzimas" del programa Grafit Version 5 (Erithacus Software, Horley, Reino Unido) junto con una función de ajuste no lineal. Se empleó la configuración por defecto para calcular los parámetros de Vmax y Km.
- 35 **[0181]** Según estos dos parámetros, se calculó una curva de calibración Michaelis-Menten y la actividad específica de cada variante se calculó según la curva de calibración.
- 40 **[0181]** Según la actividad residual del WT y variantes, se calculó el índice de rendimiento (PI) para detectar la actividad específica. El PI de una variante se calculó como el cociente "actividad específica variante/actividad específica WT". Utilizando este cociente, el PI de WT es 1,0 y una variante con un PI > 1,0 tiene una actividad específica que es mayor que WT.
- 45

#### **Análisis de actividad de hexoquinasa:**

- 50 **[0182]** Cóctel de hexoquinasa: 10-15 minutos antes de usarse, se añadieron 90 ml de agua a un recipiente BoatIL de glucosa HK R1 (kit de análisis de glucosa IL (HK), Instrumento de Laboratorio nº 182507-40) y se mezcló gradualmente. Se añadieron 100  $\mu\text{l}$  de cóctel de hexoquinasa a 85  $\mu\text{l}$  de dH<sub>2</sub>O. Se añadieron 15  $\mu\text{l}$  de muestra a las mezclas y se incubaron durante 10 minutos en la oscuridad a temperatura ambiente. Se leyó la absorbancia a 340 nm en un lector de PMT. Las concentraciones de glucosa se calcularon según una curva estándar de glucosa (0 - 2 mg/ml).

**Condiciones de análisis del análisis de detección de etanol:**

[0183] 8 % de solución madre: Se suspendieron 8 g de almidón de maíz soluble (Sigma nº S4180) en 40 ml dH<sub>2</sub>O a temperatura ambiente. Se añadieron 50 ml de dH<sub>2</sub>O hirviendo a la suspensión en un matraz de 250 ml y se coció durante 5 minutos. La solución de almidón se enfrió a 25 °C y se reguló el volumen a 100 ml con dH<sub>2</sub>O.

5 [0184] Se diluyeron 5 µl de sobrenadante crudo con 175 µl 50 mM de tampón NaAc pH 4,5 en una PMT de 96 pocillos de base plana. Se añadieron 60 µl de esta solución a 120 µl 14% almidón de maíz soluble y se incubaron durante 2 horas a 32 °C 900 rpm (incubadora/agitador HT iEMS de Thermolabsystems) Se finalizó la reacción añadiendo 90 µl de solución de cese de reacción enfriada a 4 °C. La muestra se colocó en hielo. El almidón se centrifugó a 1118 xg a 15 °C durante 5 minutos (SIGMA 6K15) y se utilizaron 15 µl de sobrenadante en el  
10 análisis de actividad de Hexoquinasa descrito anteriormente con el fin de determinar el contenido de glucosa. Solución de cese de reacción (800 mM tampón glicina-NaOH, pH 10). 4 % (m/v) solución de trabajo de almidón soluble: la solución madre se diluyó (1:1) con 100 mM tampón acetato de sodio.

[0185] Se analizó el material sobrenadante crudo para detectar la glucosa restante en el medio de cultivo tras el periodo de crecimiento. La cantidad de glucosa producida por la glucoamilasa no se calculó si se halló glucosa  
15 restante en el medio de cultivo.

**EJEMPLOS**

[0186] Los siguientes ejemplos se presentan con el fin de demostrar e ilustrar además determinadas formas de realización y aspectos de la presente invención y no han de considerarse como limitadores del alcance de la misma.

20 **Ejemplo 1:****Construcción de bibliotecas de evaluación del lugar TrGA (SEL por sus siglas en inglés) en el vector pTTT para la expresión en *Trichoderma reesei***

[0187] Se clonó una secuencia de ADNc *Trichoderma reesei* (SEQ ID NO: 4) en pDONR™201 mediante la reacción de recombinación Gateway® BP (Invitrogen, Carlsbad, CA, EEUU) teniendo como resultado el vector de entrada pDONR-TrGA (Fig. 2). La secuencia de ADNc (SEQ ID NO:4) codifica el péptido señal TrGA, la prosequencia y la proteína madura, incluyendo el dominio catalítico, la región de enlace y el dominio de unión a almidón (SEQ ID NO:1). SEQ ID NO:4 y SEQ ID NO:1 se muestran en la figura 1B y 1A. La figura 1C ilustra los dominios de TrGA de proteína madura y precursora.

[0188] Con el fin de expresar la proteína TrGA en *Trichoderma reesei*, la secuencia codificante TrGA (SEQ ID NO:4) se clonó en el vector de destino compatible pTTT-Dest Gateway (Fig. 3) mediante la reacción de recombinación GATEWAY® LR. El vector de expresión contenía las regiones terminadoras y promotoras derivadas de *cbhl T. reesei* que permitían una fuerte expresión inducible de un gen de interés. El vector también contenía el marcador selectivo *amdS Aspergillus nidulans* que permitía el crecimiento de transformantes en acetamida como una única fuente de nitrógeno. El vector de expresión también contenía regiones de telómero *T. reesei* que permitían el mantenimiento plásmido no cromosómico en una célula fúngica. En el plásmido pTTT-Dest de destino, las regiones de promotor y terminador de *cbhl* se separaron mediante el gen de resistencia cloranfenicol, CmR y el gen letal *E. coli*, *ccdB*, flanqueado por lugares de recombinación específicos basados en bacteriófago lambda *attR1*, *attR2*. Esta configuración permitía la selección directa de recombinantes que contenían el gen TrGA bajo el control de los elementos reguladores *cbhl* en la orientación adecuada mediante la reacción de recombinación GATEWAY® LR. El vector de expresión final pTTT-TrGA se muestra en la figura 4.

[0189] Los SEL de TrGA se crearon utilizando el vector de entrada pDONR-TrGA (fig. 2) como plantilla y los cebadores enumerados en la tabla 1. Todos los cebadores utilizados en los experimentos de la mutagénesis contenían el triplete NNS (N = A,C,T,G y S=C o G) en la posición que se alinea con el codón de la secuencia TrGA diseñada para mutarse (SEQ ID NO:1), lo que permite una incorporación aleatoria de nucleótidos en la posición preseleccionada. La construcción de cada SEL empezó con dos amplificaciones por PCR independientes en el vector de entrada pDONR-TrGA : una utilizando el Gateway F (pDONR201 - FW) y un cebador de mutagénesis específico R (tabla 2) y la otra, el cebador Gateway R (pDONR201 - RV) y un cebador de mutagénesis específico F (tabla 2). La polimerasa de ADN PHUSION de alta fidelidad (Finnzymes OY, Espoo, Finlandia) se utilizó en una reacción de amplificación por PCR que incluía cebadores 0,2 µM. Las reacciones se llevaron a cabo durante 25 ciclos según el protocolo proporcionado por Finnzymes. 1 µl alícuotas de los fragmentos de PCR obtenidos se utilizaron como plantillas para una reacción por PCR de fusión posterior junto con los cebadores Gateway FW y Gateway RV (Invitrogen). La amplificación por PCR, después de 22 ciclos, produjo una población de fragmentos de ADN de TrGA lineal de longitud completa mutados de forma aleatoria en la posición de codón específica. Los fragmentos se flanquearon por los lugares de recombinación *attL1*, *attL2* específicos de Gateway en ambos extremos. Los fragmentos de ADN se purificaron con un kit de limpieza  
55

CHARGESWITCH® PCR (Invitrogen, Carlsbad EEUU) y a continuación se recombinaron con 100 ng del vector de destino pTTT (fig. 3) utilizando la mezcla de enzimas LR CLONASE™ II según el protocolo proporcionado por Invitrogen. Los productos de recombinación que se generaron se transformaron en *E.coli* Max Efficiency DH5a, según describe el proveedor (Invitrogen). Las construcciones de expresión finales pTTT-TrGA con mutaciones en la posición deseada se eligieron mediante el recubrimiento de bacterias en placas de agar 2xYT (16 g/L Bacto Tryptone (Difco), 10 g/L Bacto Yeast Extract (Difco), 5 g/L NaCl, 16 g/L Bacto Agar (Difco) con 100 µg/ml ampicilina.

**[0190]** Se cultivaron 96 colonias únicas de cada biblioteca durante 24 hrs a 37 °C en una PMT con 200 µl de medio 2xYT con 100 µg/ml ampicilina. Los cultivos se utilizaron directamente para amplificar fragmentos por PCR incluidos en la región en la que se introducía la mutación específica. Los productos por PCR específica obtenidos se secuenciaron utilizando un analizador de secuencias ABI3100 (Applied Biosystems). Cada biblioteca contenía de 15 a 19 variantes de TrGA diferentes en el vector de expresión final. Estas variantes se transformaron de forma individual en *T. reesei*, como se describe a continuación. Las bibliotecas se enumeran desde 1 a 182 en referencia al residuo de aminoácido específico en la secuencia de TrGA que se ha mutado de forma aleatoria.

**Tabla 1A y B - Cebadores utilizados para generar SEL de TrGA para el dominio catalítico (Tabla 1A) y el enlazador y el dominio de unión a almidón (Tabla 1B) (D = Directo; I = Inverso)**

Tabla 1A – cebadores dominio catalítico		
Posición AA	D/I	Secuencia ADN 5' a 3'
pDON R201-	D	TCGCGTTAACGCTAGCATGGATCTC (SEQ ID NO:13)
pDON R201-	I	TCGCGTTAACGCTAGCATGGATCTC (SEQ ID NO:14)
4	D	CGTCACCAAGAGGTCTGTTGACNNSSTTCATCAGCACCGAGACGCC (SEQ ID NO:15)
4	I	GTCAACAGACCTCTTGGTGACGTCG (SEQ ID NO:16)
5	D	CACCAAGAGGTCTGTTGACGACNNSATCAGCACCGAGACGCCTATTGC (SEQ ID NO:17)
5	I	GTCGTCAACAGACCTCTTGGTGAC (SEQ ID NO:18)
10	D	TGACGACTTCATCAGCACCGAGNNSCCTATTGCACTG (SEQ ID NO:19)
10	I	CTCGGTGCTGATGAAGTCGTC (SEQ 10 NO:20)
12	D	TCATCAGCACCGAGACGCCTNNSGCACTGAACAATCTTCTTTGCA (SEQ ID NO:21)
12	I	AGGCGTCTCGGTGCTGATGAAGTCG (SEQ ID NO:22)
14	D	CAGCACCGAGACGCCTATTGCANNSAACAATCTTCTT (SEQ ID NO:23)
14	I	TGCAATAGGCGTCTCGGTGCT (SEQ ID NO:24)
15	D	CACCGAGACGCCTATTGCACTGNNSAATCTTCTTTGC (SEQ ID NO:25)
15	I	CAGTGCAATAGGCGTCTCGGT (SEQ ID NO:26)
23	D	CAATCTTCTTTGCAATGTTGGTNNSGATGGATGCCGT (SEQ ID NO:27)
23	I	ACCAACATTGCAAAGAAGATTG (SEQ ID NO:28)
24	D	TTCTTTGCAATGTTGGTCCTNNSGGATGCCGTGCATTCGGCACAT (SEQ ID NO:29)
24	I	AGGACCAACATTGCAAAGAAGATTG (SEQ ID NO:30)
29	D	GTCCTGATGGATGCCGTGCANNSGGCACATCAGCTGGTGCGGTGA (SEQ ID NO:31)

ES 2 502 743 T3

		NO:31)
29	I	TGCACGGCATCCATCAGGACCAACA (SEQ ID NO:32)
42	D	TGCGGTGATTGCATCTCCAGCNNSATTGACCCGGAC (SEQ ID NO:33)
42	I	GCTGGGAGATGCAATCACCGCA (SEQ ID NO:34)
43	D	TGATTGCATCTCCAGCACANNSGACCCGGACTACTATTACATGT (SEQ ID NO:35)
43	I	TGTGCTGGGAGATGCAATCACCGCA (SEQ ID NO:36)
44	D	TTGCATCTCCAGCACAAATTNNSCCGGACTACTATTACATGTGGA (SEQ ID NO:37)
44	I	AATTGTGCTGGGAGATGCAATCACCC (SEQ ID NO:38)
45	D	CATCTCCAGCACAAATTGACNNSGACTACTATTACATGTGGACGC (SEQ ID NO:39)
45	I	GTCAATTGTGCTGGGAGATGCAATC (SEQ ID NO:40)
46	D	CTCCAGCACAAATTGACCCGNNSTACTATTACATGTGGACGCGAGA SEQ ID NO:41)
46	I	CGGGTCAATTGTGCTGGGAGATGCA (SEQ ID NO:42)
47	D	CCAGCACAAATTGACCCGGACNNSTATTACATGTGGACGCGAGATA (SEQ ID NO:43)
47	I	TCCGGGTCAATTGTGCTGGGAGAT (SEQ ID NO:44)
49	D	CAATTGACCCGGACTACTATNNSATGTGGACGCGAGATAGCGCTC (SEQ ID NO:45)
49	I	ATAGTAGTCCGGGTCAATTGTGCTG (SEQ ID NO:46)
51	D	ACCCGGACTACTATTACATGNNSACGCGAGATAGCGCTCTTGTCT (SEQ ID NO:47)
51	I	CATGTAATAGTAGTCCGGGTCAATT (SEQ ID NO:48)
59	D	ACGCGAGATAGCGCTCTTGTGNNSAAGAACCTCATC (SEQ ID NO:49)
59	I	GACAAGAGCGCTATCTCGCGT (SEQ ID NO:50)
60	D	GCGAGATAGCGCTCTTGTCTTCNNSAACCTCATCGAC (SEQ ID NO:51)
60	I	GAAGACAAGAGCGCTATCTCG (SEQ ID NO:52)
61	D	AGATAGCGCTCTTGTCTTCAAGNNSCTCATCGACCGC (SEQ ID NO:53)
61	I	CTTGAAGACAAGAGCGCTATC (SEQ ID NO:54)
65	D	TGTCTTCAAGAACCTCATCGACNNSTCACCGAAACG (SEQ ID NO:55)
65	I	GTCGATGAGGTTCTTGAAGAC (SEQ ID NO:56)
67	D	CAAGAACCTCATCGACCGCTTCNNSGAAACGTACGAT (SEQ ID NO:57)
67	I	GAAGCGGTCGATGAGGTTCTT (SEQ ID NO:58)
68	D	GAACCTCATCGACCGCTTACCNNNSACGTACGATGCG (SEQ ID NO:59)

ES 2 502 743 T3

68	I	GGTGAAGCGGTCGATGAGGTT (SEQ ID NO:60)
70	D	TCGACCGCTTCACCGAAACGNNSGATGCGGGCCTGCAGCGCCGCA (SEQ ID NO:61)
70	I	CGTTTCGGTGAAGCGGTCGATGAGG (SEQ ID NO:62)
72	D	CCGCTTCACCGAAACGTACGATNNSGGCCTGCAGCGC (SEQ ID NO:63)
72	I	ATCGTACGTTTCGGTGAAGCGG (SEQ ID NO:64)
73	D	CTTCACCGAAACGTACGATGCGNNSCTGCAGCGCCGC (SEQ ID NO:65)
73	I	CGCATCGTACGTTTCGGTGAA (SEQ ID NO:66)
75	D	AAACGTACGATGCGGGCCTGNNSCGCCGCATCGAGCAGTACATTA (SEQ ID NO:67)
75	I	CAGGCCCGCATCGTACGTTTCGGTG (SEQ ID NO:68)
76	D	CGTACGATGCGGGCCTGCAGNNSCGCATCGAGCAGTACATTACTG (SEQ ID NO:69)
76	I	CTGCAGGCCCGCATCGTACGTTTCG (SEQ ID NO:70)
94	D	CTCTCCAGGGCCTCTCTAACNNSTCGGGCTCCCTCGCGGACGGCT (SEQ ID NO:71)
94	I	GTTAGAGAGGCCCTGGAGAGTGACC (SEQ ID NO:72)
97	D	GGGCCTCTCTAACCCCTCGGGCNSCTCGCGGACGGC (SEQ ID NO:73)
97	I	GCCCGAGGGGTTAGAGAGGCC (SEQ ID NO:74)
98	D	CTCTCTAACCCCTCGGGCTCCNNSGCGGACGGCTCT (SEQ ID NO:75)
98	I	GGAGCCCGAGGGGTTAGAGAG (SEQ ID NO:76)
99	D	CTCTAACCCCTCGGGCTCCCTCNNSGACGGCTCTGGT (SEQ ID NO:77)
99	I	GAGGGAGCCCGAGGGGTTAGA (SEQ ID NO:78)
100	D	ACCCCTCGGGCTCCCTCGCGNNSGGCTCTGGTCTCGGCGAGCCCA (SEQ ID NO:79)
100	I	CGCGAGGGAGCCCGAGGGGTTAGAG (SEQ ID NO:80)
102	D	CTCGGGCTCCCTCGCGGACGGCNSGGTCTCGGCGAG (SEQ ID NO:81)
102	I	CGCCGTCCGCGAGGGAGCCCGA (SEQ ID NO:82)
110	D	TGGTCTCGGCGAGCCCAAGTTTNNSTTGACCCTGAAG (SEQ ID NO:83)
110	I	AAACTTGGGCTCGCCGAGACCA (SEQ ID NO:84)
111	D	TCTCGGCGAGCCCAAGTTTGAGNNSACCCTGAAGCCT (SEQ ID NO:85)
111	I	CTCAAACCTGGGCTCGCCGAG (SEQ ID NO:86)
113	D	CGAGCCCAAGTTTGAGTTGACCNNSAAGCCTTTCACC (SEQ ID NO:87)
113	I	GGTCAACTCAAACCTGGGCTC (SEQ ID NO:88)

ES 2 502 743 T3

114	D	CCAAGTTTGAGTTGACCCTGNNSCCTTTCACCGGCAACTGGGGTC (SEQ ID NO:89)
114	I	CAGGGTCAACTCAAACCTGGGCTCG (SEQ ID NO:90)
116	D	TTGAGTTGACCCTGAAGCCTNNSACCGGCAACTGGGGTCGACCGCA (SEQ ID NO:91)
116	I	AGGCTTCAGGGTCAACTCAAACCTG (SEQ ID NO:92)
119	D	CCCTGAAGCCTTTCACCGGCNNSTGGGGTCGACCGCAGCGGGATG (SEQ ID NO:93)
119	I	GCCGGTGAAAGGCTTCAGGGTCAAC (SEQ ID NO:94)
122	D	CTTTCACCGGCAACTGGGGTNNSCCGCAGCGGGATGGCCCAGCTC (SEQ ID NO:95)
122	I	ACCCAGTTGCCGGTGAAAGGCTTC (SEQ ID NO:96)
125	D	GCAACTGGGGTCGACCGCAGNNSGATGGCCCAGCTCTGCGAGCCA (SEQ ID NO:97)
125	I	CTGCGGTGACCCAGTTGCCGGTG (SEQ ID NO:98)
133	D	GGATGGCCCAGCTCTGCGAGCCNNSGCCTTGATTGGA (SEQ ID NO:99)
133	I	GGCTCGCAGAGCTGGGCCATCC (SEQ ID NO:100)
137	D	TGCGAGCCATTGCCTTGATTNNSTACTCAAAGTGGCTCATCAACA (SEQ ID NO:101)
137	I	AATCAAGGCAATGGCTCGCAGAGCT (SEQ ID NO:102)
140	D	CATTGCCTTGATTGGATACTCANNSTGGCTCATCAAC (SEQ ID NO:103)
140	I	TGAGTATCCAATCAAGGCAATG (SEQ ID NO:104)
144	D	TGGATACTCAAAGTGGCTCATCNNSAACAACATCAG (SEQ ID NO:105)
144	I	GATGAGCCACTTTGAGTATCC (SEQ ID NO:106)
145	D	ATACTCAAAGTGGCTCATCAACNNSAATCAGTCG (SEQ ID NO:107)
145	I	GTTGATGAGCCACTTTGAGTA (SEQ ID NO:108)
146	D	CAAAGTGGCTCATCAACAACNNSTATCAGTCGACTGTGTCCAACG (SEQ ID NO:109)
146	I	GTTGTTGATGAGCCACTTTGAGTAT (SEQ ID NO:110)
147	D	AAAGTGGCTCATCAACAACAACNNSCAGTCGACTGTG (SEQ ID NO:111)
147	I	GTTGTTGTTGATGAGCCACTT (SEQ ID NO:112)
148	D	GGCTCATCAACAACAACATNNSTCGACTGTGTCCAACGTCATCT (SEQ ID NO:113)
148	I	ATAGTTGTTGTTGATGAGCCACTTT (SEQ ID NO:114)
152	D	CAACAACATCAGTCGACTGTGNNSAACGTCATCTGG (SEQ ID NO:115)
152	I	CACAGTCGACTGATAGTTGTT (SEQ ID NO:116)



ES 2 502 743 T3

153	D	CAACTATCAGTCGACTGTGTCCNNSGTCATCTGGCCT (SEQ ID NO:117)
153	I	GGACACAGTCGACTGATAGTT (SEQ ID NO:118)
164	D	GCCTATTGTGCGCAACGACCTCNNSTATGTTGCCAGT (SEQ ID NO:119)
164	I	GAGGTCGTTGCGCACAATAGG (SEQ ID NO:120)
169	D	ACCTCAACTATGTTGCCAGNNSTGGAACCAAACCGGCTTTGACC (SEQ ID NO:121)
169	I	CTGGGCAACATAGTTGAGGTCGTTG (SEQ ID NO:122)
172	D	ATGTTGCCAGTACTGGAACNNSACCGGCTTTGACCTCTGGGAAG(SEQ ID NO:123)
172	I	GTTCCAGTACTGGGCAACATAGTTG (SEQ ID NO:124)
175	D	AGTACTGGAACCAAACCGGCNNSGACCTCTGGGAAGAAGTCAATG (SEQ ID NO:125)
175	I	GCCGGTTTGGTTCCAGTACTGGGCA (SEQ ID NO:126)
178	D	ACCAAACCGGCTTTGACCTCNNNSGAAGAAGTCAATGGGAGCTCAT(SEQ ID NO:127)
178	I	GAGGTCAAAGCCGGTTTGGTTCCAG (SEQ ID NO:128)
180	D	CCGGCTTTGACCTCTGGGAANNSGTCAATGGGAGCTCATTCTTTA (SEQ ID NO:129)
180	I	TTCCAGAGGTCAAAGCCGGTTTGG (SEQ ID NO:130)
181	D	GCTTTGACCTCTGGGAAGAANNSAATGGGAGCTCATTCTTTACTG (SEQ ID NO:131)
181	I	TTCTTCCCAGAGGTCAAAGCCGGTT (SEQ ID NO:132)
182	D	CTTTGACCTCTGGGAAGAAGTCNNSGGGAGCTCATTC (SEQ ID NO:133)
182	I	GACTTCTTCCCAGAGGTCAAAG (SEQ ID NO:134)
204	D	TGTCGAGGGCGCCACTCTTGCTNNSACTCTTGCCAG (SEQ ID NO: 135)
204	I	AGCAAGAGTGGCGCCCTCGAC (SEQ ID NO:136)
205	D	CGAGGGCGCCACTCTTGCTGCCNNSCTTGCCAGTCG (SEQ ID NO:137)
205	I	GGCAGCAAGAGTGGCGCCCTC (SEQ ID NO:138)
208	D	CTCTTGCTGCCACTCTTGGCANNSTCGGAAGCGCTTATTCATCTG(SEQ ID NO:139)
208	I	GCCAAGAGTGGCAGCAAGAGTGGCG (SEQ ID NO:140)
211	D	CCACTCTTGCCAGTCGGGANNSGCTTATTCATCTGTTGCTCCCC(SEQ ID NO:141)
211	I	TCCCGACTGGCCAAGAGTGGCAGCA (SEQ ID NO:142)
214	D	TGGCCAGTCGGGAAGCGCTTATNNSCTGTTGCTCCC (SEQ ID NO:143)
214	I	ATAAGCGCTTCCCGACTGGCC (SEQ ID NO:144)

ES 2 502 743 T3

216	D	GTCGGGAAGCGCTTATTCATCTNNSGCTCCCCAGGTT (SEQ ID NO:145)
216	I	AGATGAATAAGCGCTTCCCGA (SEQ ID NO:146)
219	D	CGCTTATTCATCTGTTGCTCCCNNSGTTTTGTGCTTT (SEQ ID NO:147)
219	I	GGGAGCAACAGATGAATAAGC (SEQ ID NO:148)
228	D	TGTGCTTTCTCCAACGATTCCNNSGTGTCGTCTGGTGGATACGTCG(SEQ ID NO:149)
228	I	GAATCGTTGGAGAAAGCACAAAACCT (SEQ ID NO:150)
229	D	GTGCTTTCTCCAACGATTCTGGNNSTCGTCTGGTGGGA (SEQ ID NO: 151)
229	I	CCAGAATCGTTGGAGAAAGCA (SEQ ID NO: 152)
230	D	(SEQ ID NO:153)
230	I	CACCCAGAATCGTTGGAGAAA (SEQ ID NO:154)
231	D	TCTCCAACGATTCTGGGTGTCGNNSGGTGGATACGTC (SEQ ID NO:155)
231	I	CGACACCCAGAATCGTTGGAGA (SEQ ID NO:156)
236	D	GGTGTCGTCTGGTGGATACGTCNNSTCCAACATCAACAC (SEQ ID NO:157)
236	I	GACGTATCCACCAGACGACAC (SEQ ID NO:158)
239	D	TGGTGGATACGTCGACTCCAACNNSAACACCAACGAG (SEQ ID NO:159)
239	I	GTTGGAGTCGACGTATCCACC (SEQ ID NO: 160)
240	D	TGGATACGTCGACTCCAACATCNNSACCAACGAGGGCA (SEQ ID NO:161)
240	I	GATGTTGGAGTCGACGTATCCA (SEQ ID NO:162)
241	D	ATACGTCGACTCCAACATCAACNNSAACGAGGGCAGGAC(SEQID NO:163)
241	I	GTTGATGTTGGAGTCGACGTA (SEQ ID NO:164)
242	D	TCGACTCCAACATCAACACCNNSGAGGGCAGGACTGGCAAGGATG (SEQ ID NO:165)
242	I	GGTGTTGATGTTGGAGTCGACGTAT(SEQ ID NO:166)
243	D	ACTCCAACATCAACACCAACNNSGGCAGGACTGGCAAGGATGTCA (SEQ ID NO: 167)
243	I	GTTGGTGTGATGTTGGAGTCGACG (SEQ ID NO: 168)
244	D	CTCCAACATCAACACCAACGAGNNSAGGACTGGCAAG (SEQ ID NO:169)
244	I	CTCGTTGGTGTGATGTTGGAGT (SEQ ID NO:170)
245	D	ACATCAACACCAACGAGGGCNNSACTGGCAAGGATGTCAACTCCG (SEQ ID NO:171)
245	I	GCCCTCGTTGGTGTGATGTTGGAGT (SEQ ID NO:172)
263	D	TTCCATCCACACCTTCGATCCCNNSCTGGCTGTGAC (SEQ ID NO:173)
263	I	GGGATCGAAGGTGTGGATGGA (SEQ ID NO:174)

ES 2 502 743 T3

264	D	CATCCACACCTTCGATCCCAACNNSGGCTGTGACGCA (SEQ ID NO:175)
264	I	GTTGGGATCGAAGGTGTGGAT (SEQ ID NO: 176)
265	D	CCACACCTTCGATCCCAACCTTNNSTGTGACGCAGGC (SEQ ID NO: 177)
265	I	AAGGTTGGGATCGAAGGTGTG (SEQ ID NO:178)
268	D	CGATCCCAACCTTGGCTGTGACNNSGGCACCTTCCAGC (SEQ ID NO:179)
268	I	GTCACAGCCAAGGTTGGGATC (SEQ ID NO:180)
269	D	TCCCAACCTTGGCTGTGACGCANNSACCTTCCAGCCA (SEQ ID NO:181)
269	I	TGCGTCACAGCCAAGGTTGGG (SEQ ID NO:182)
276	D	AGGCACCTTCCAGCCATGCAGTNNSAAAGCGCTCTCC (SEQ ID NO:183)
276	I	ACTGCATGGCTGGAAGGTGCC (SEQ ID NO:184)
284	D	CAAAGCGCTCTCCAACCTCAAGNNSGTTGTGACTCCT (SEQ ID NO:185)
284	I	CTTGAGGTTGGAGAGCGCTTT (SEQ ID NO:186)
291	D	GGTTGTTGTGCGACTCCTTCCGCNNSATCTACGGCGTG (SEQ ID NO:187)
291	I	GCGGAAGGAGTCGACAACAAC (SEQ ID NO:188)
292	D	TTGTCGACTCCTTCCGCTCCNNSTACGGCGTGAACAAGGGCATT(CSEQ ID NO:189)
292	I	GGAGCGGAAGGAGTCGACAACAACC (SEQ ID NO:190)
294	D	ACTCCTTCCGCTCCATCTACNNSGTGAACAAGGGCATTCTGCCG(SEQ ID NO:191)
294	I	GTAGATGGAGCGGAAGGAGTCGACA (SEQ ID NO:192)
297	D	GCTCCATCTACGGCGTGAACNNSGGCATTCTGCCGGTGTGCCG(SEQ ID NO:193)
297	I	GTTACGCCGTAGATGGAGCGGAAG (SEQ ID NO:194)
300	D	CTACGGCGTGAACAAGGGCATTNNSGCCGGTGTGCCG (SEQ ID NO:195)
300	I	AATGCCCTTGTTACGCCGTA (SEQ ID NO:196)
301	D	CGGCGTGAACAAGGGCATTCTNNSGGTGTGCCGTC (SEQ ID NO:197)
301	I	AGGAATGCCCTTGTTACGCC (SEQ ID NO:198)
303	D	GAACAAGGGCATTCTGCCGGTNNSGCCGTCGCCATT (SEQ ID NO:199)
303	I	ACCGGCAGGAATGCCCTTGTT (SEQ ID NO:200)
309	D	GTGCTGCCGTCGCCATTGGCENNSTATGCAGAGGATGTGTACTACA(SEQ ID NO:201)
309	I	GCCAATGGCGACGGCAGCACCGGCA (SEQ ID NO:202)
310	D	CTGCCGTGCCATTGGCCGGNNSGCAGAGGATGTGTACTACAACG (SEQ ID NO:203)

ES 2 502 743 T3

310	I	CCGGCCAATGGCGACGGCAGCACCG (SEQ ID NO:204)
311	D	TGCCGTGCGCCATTGGCCGGTATNNSGAGGATGTGTAC (SEQ ID NO:205)
311	I	ATACCGGCCAATGGCGACGGC (SEQ ID NO:206)
313	D	CCATTGGCCGGTATGCAGAGNNSGTGTACTACAACGGCAACCCTT(SEQ ID NO:207)
313	D	CCATTGGCCGGTATGCAGAGNNSGTGTACTACAACGGCAACCCTT(SEQ ID NO:208)
313	I	CTCTGCATACCGGCCAATGGCGACG (SEQ ID NO:209)
313	I	CTCTGCATACCGGCCAATGGCGACG (SEQ ID NO:210)
314	D	TTGGCCGGTATGCAGAGGATNNSTACTACAACGGCAACCCTTGGT(SEQ ID NO:211)
314	I	ATCCTCTGCATACCGGCCAATGGCG (SEQ ID NO:212)
315	D	GCCGGTATGCAGAGGATGTGNNSACAACGGCAACCCTTGGTATC (SEQ ID NO:213)
315	I	CACATCCTCTGCATACCGGCCAATG (SEQ ID NO:214)
316	D	GGTATGCAGAGGATGTGTACNNSAACGGCAACCCTTGGTATCTTG(SEQ ID NO:215)
316	I	GTACACATCCTCTGCATACCGGCCAAT(SEQ ID NO:216)
317	D	ATGCAGAGGATGTGTACTACNNSGGCAACCCTTGGTATCTTGCTA(SEQ ID NO:217)
317	D	ATGCAGAGGATGTGTACTACNNSGGCAACCCTTGGTATCTTGCTA(SEQ ID NO:218)
317	I	GTAGTACACATCCTCTGCATACCGGC (SEQ ID NO:219)
317	I	GTAGTACACATCCTCTGCATACCGGC (SEQ ID NO:220)
321	D	TGTACTACAACGGCAACCCTNNSTATCTTGCTACATTTGCTGCTG(SEQ ID NO:221)
321	D	TGTACTACAACGGCAACCCTNNSTATCTTGCTACATTTGCTGCTG (SEQ ID NO:222)
321	I	AGGGTTGCCGTTGTAGTACACATCC (SEQ ID NO:223)
321	I	AGGGTTGCCGTTGTAGTACACATCC (SEQ ID NO:224)
338	D	GCAGCTGTACGATGCCATCTACNNSTGGAAGAAGACG (SEQ ID NO:225)
338	I	GTAGATGGCATCGTACAGCTG (SEQ ID NO:226)
340	D	ACGATGCCATCTACGTCTGGNNSAAGACGGGCTCCATCACGGTGA(SEQ ID NO:227)
340	I	CCAGACGTAGATGGCATCGTACAGC (SEQ ID NO:228)
341	D	ATGCCATCTACGTCTGGAAGNNSACGGGCTCCATCACGGTGACCG (SEQ ID NO:229)

ES 2 502 743 T3

341	I	CTTCCAGACGTAGATGGCATCGTACAGC (SEQ ID NO:230)
342	D	ATGCCATCTACGTCTGGAAGAAGNNSGGCTCCATCACG (SEQ ID NO:231)
342	I	CTTCTTCCAGACGTAGATGGC (SEQ ID NO:232)
344	D	CTACGTCTGGAAGAAGACGGGCNNSATCACGGTGACC (SEQ ID NO:233)
344	I	GCCCGTCTTCTTCCAGACGTAG (SEQ ID NO:234)
346	D	CTGGAAGAAGACGGGCTCCATCNNSGTGACCGCCACCTC (SEQ ID NO:235)
346	I	GATGGAGCCCGTCTTCTTCCA (SEQ ID NO:236)
349	D	GACGGGCTCCATCACGGTGACCNNSACCTCCCTGGCC (SEQ ID NO:237)
349	I	GGTCACCGTGATGGAGCCCGT (SEQ ID NO:238)
350	D	GCTCCATCACGGTGACCGCCNNSCTCCCTGGCCTTCTTCCAGGAGC (SEQ ID NO:239)
350	I	GGCGGTCACCGTGATGGAGCCCGTC (SEQ ID NO:240)
356	D	CCACCTCCCTGGCCTTCTCNNSGAGCTTGTTCTGGCGTGACGG (SEQ ID NO:241)
356	I	GAAGAAGGCCAGGGAGGTGGCGGTC (SEQ ID NO:242)
359	D	CCTGGCCTTCTTCCAGGAGCTTNSCCTGGCGTGACG (SEQ ID NO:243)
359	I	AAGCTCCTGGAAGAAGGCCAG (SEQ ID NO:244)
361	D	CTTCTTCCAGGAGCTTGTTCTNNSGTGACGGCCGGG (SEQ ID NO:245)
361	I	AGGAACAAGCTCCTGGAAGAA (SEQ ID NO:246)
363	D	AGGAGCTTGTTCTGGCGTGNNSGCCGGGACCTACTCCAGCAGCT (SEQ ID NO:247)
363	I	CACGCCAGGAACAAGCTCCTGGAAG (SEQ ID NO:248)
364	D	GGAGCTTGTTCTGGCGTGACGNNSGGGACCTACTCC (SEQ ID NO:249)
364	I	CGTCACGCCAGGAACAAGCTC (SEQ ID NO:250)
368	D	GCGTGACGGCCGGGACCTACNNSAGCAGCTCTTCGACCTTTACCA (SEQ ID NO:251)
368	I	GTAGGTCCCGGCCGTCACGCCAGGA (SEQ ID NO:252)
369	D	TGACGGCCGGGACCTACTCCNNSAGCTCTTCGACCTTTACCAACA (SEQ ID NO:253)
369	I	GGAGTAGGTCCCGGCCGTCACGCCA (SEQ ID NO:254)
375	D	CTCCAGCAGCTCTTCGACCTTTNNSAACATCATCAACG (SEQ ID NO:255)
375	I	AAAGGTCGAAGAGCTGCTGGA (SEQ ID NO:256)
376	D	GCAGCTCTTCGACCTTTACNNSATCATCAACGCCGTCTCGACAT (SEQ ID NO:257)
376	I	GGTAAAGGTCGAAGAGCTGCTGGAG (SEQ ID NO:258)

ES 2 502 743 T3

379	D	TTCGACCTTTACCAACATCATCNNSGCCGTCTCGACA (SEQ ID NO:259)
379	I	GATGATGTTGGTAAAGGTCGA (SEQ IN NO:260)
382	D	TACCAACATCATCAACGCCGTCNNSACATACGCCGAT (SEQ ID NO:261)
382	I	GACGGCGTTGATGATGTTGGT (SEQ ID NO:262)
390	D	GACATACGCCGATGGCTTCCTCNNSGAGGCTGCCAAG (SEQ ID NO:263)
390	I	GAGGAAGCCATCGGCGTATGT (SEQ ID NO:264)
391	D	ATACGCCGATGGCTTCCTCAGCNNSGCTGCCAAGTAC (SEQ ID NO:265)
391	I	GCTGAGGAAGCCATCGGCGTA (SEQ ID NO:266)
393	D	CGATGGCTTCCTCAGCGAGGCTNNSAAGTACGTCCCC (SEQ ID NO:267)
393	I	AGCCTCGCTGAGGAAGCCATC (SEQ ID NO:268)
394	D	TGGCTTCCTCAGCGAGGCTGCCNNSTACGTCCCCGCC (SEQ ID NO:269)
394	I	GGCAGCCTCGCTGAGGAAGCC (SEQ ID NO:270)
395	D	TCCTCAGCGAGGCTGCCAAGNNSGTCCCCGCCGACGGTTCGCTGG (SEQ ID NO:271)
395	I	CTTGGCAGCCTCGCTGAGGAAGCCA (SEQ ID NO:272)
398	D	AGGCTGCCAAGTACGTCCCCNNSGACGGTTCGCTGGCCGAGCAGTT (SEQ ID NO:273)
398	I	GGGGACGTACTIONTGGCAGCCTCGCTG (SEQ ID NO:274)
401	D	AGTACGTCCCCGCCGACGGTNNSCTGGCCGAGCAGTTTGACCGCA (SEQ ID NO:275)
401	I	ACCGTCGGCGGGGACGTACTIONTGGCAG (SEQ ID NO:276)
408	D	CGCTGGCCGAGCAGTTTGACNNSAACAGCGGCACTCCGCTGTCTG(SEQ ID NO:277)
408	I	GTCAAAGTCTCGGCCAGCGAACCG (SEQ ID NO:278)
409	D	TGGCCGAGCAGTTTGACCGCNNSAGCGGCACTCCGCTGTCTGCGC (SEQ ID NO:279)
409	I	GCGGTCAAAGTCTCGGCCAGCGAA (SEQ ID NO:280)
410	D	GGCCGAGCAGTTTGACCGCAACNNSGGCACTCCGCTG (SEQ ID NO:281)
410	I	GTTGCGGTCAAAGTCTCGGC (SEQ ID NO:282)
412	D	AGTTTGACCGCAACAGCGGCNNSCCGCTGTCTGCGCTTCACCTGA(SEQ ID NO:283)
412	I	GCCGCTGTTGCGGTCAAAGTCTCG (SEQ ID NO:284)
415	D	GCAACAGCGGCACTCCGCTGNNSGCGCTTCACCTGACGTGGTCGT (SEQ ID NO:285)
415	I	CAGCGGAGTGCCGCTGTTGCGGTCA (SEQ ID NO:286)

ES 2 502 743 T3

417	D	CAGCGGCACTCCGCTGTCTGCGNNSCACCTGACGTGGT (SEQ ID NO:287)
417	I	CGCAGACAGCGGAGTGCCGCT (SEQ ID NO:288)
418	D	GCACTCCGCTGTCTGCGCTTNSCTGACGTGGTCGTACGCCTCGT(SEQ ID NO:289)
418	I	AAGCGCAGACAGCGGAGTGCCGCTG (SEQ ID NO:290)
421	D	TGTCTGCGCTTACCTGACGNNSTCGTACGCCTCGTTCTTGACAG(SEQ ID NO:291)
421	I	CGTCAGGTGAAGCGCAGACAGCGGA (SEQ ID NO:292)
430	D	GTACGCCTCGTTCTTGACAGCCNNSGCCCGTCGGGCT (SEQ ID NO:293)
430	I	GGCTGTCAAGAACGAGGCGTA (SEQ ID NO:294)
431	D	CGCCTCGTTCTTGACAGCCACGNNSCGTCGGGCTGGC (SEQ ID NO:295)
431	I	CGTGGCTGTCAAGAACGAGGC (SEQ ID NO:296)
433	D	TCTTGACAGCCACGGCCCGTNSGCTGGCATCGTGCCCCCTCGT(SEQ ID NO:297)
433	I	ACGGGCCGTGGCTGTCAAGAACGAG (SEQ ID NO:298)
436	D	CCACGGCCCGTCGGGCTGGCNNSGTGCCCCCTCGTGGGCCAACA (SEQ ID NO:299)
436	I	GCCAGCCCGACGGGCCGTGGCTGTC (SEQ ID NO:300)
442	D	TGGCATCGTGCCCCCTCGTGGNNSAACAGCAGCGCT (SEQ ID NO:301)
442	I	CCACGAGGGGGGCACGATGCC (SEQ ID NO:302)
443	D	CATCGTGCCCCCTCGTGGGCCNNSAGCAGCGCTAGC (SEQ ID NO:303)
443	I	GGCCCACGAGGGGGGCACGAT (SEQ ID NO:304)
444	D	CGTGCCCCCTCGTGGGCCAACNNSAGCGCTAGCACG (SEQ ID NO:305)
444	I	GTTGGCCCACGAGGGGGGCAC (SEQ ID NO:306)
448	D	GTGGGCCAACAGCAGCGCTAGCNNSATCCCCTCGACG (SEQ ID NO:307)

Tabla 1B: Enlazador y cebadores SBD

Posición AA	D/I	Secuencia ADN 5' a 3'
451	D	GCAGCGCTAGCACGATCCCNNSACGTGCTCCGGCGCGTCCGTGG (SEQ ID NO:308)
451	I	GGGGATCGTGCTAGCGCTGCTGTTG (SEQ ID NO:309)
493	D	CTACACGCCCTGCCCTGCGCGNNSCCAACCTCCGTG (SEQ ID NO:310)
493	I	CGCGCAGGGCAGGGGGCGTGTA (SEQ ID NO:311)
494	D	CACGCCCTGCCCTGCGCGACCNNSACCTCCGTGGCC (SEQ ID NO:312)

ES 2 502 743 T3

494	I	GGTCGCGCAGGGCAGGGGCGT (SEQ ID NO:313)
495	D	GCCCCTGCCCTGCGCGACCCCANNSTCCGTGGCCGTC (SEQ ID NO:314)
495	I	TGGGGTCGCGCAGGGCAGGGG (SEQ ID NO:315)
501	D	CCCAACCTCCGTGGCCGTCACCNNSCACGAGCTCGTGT (SEQ ID NO:316)
501	I	GGTGACGGCCACGGAGGTTGG (SEQ ID NO:317)
502	D	AACCTCCGTGGCCGTCACCTTCNNSGAGCTCGTGTCG (SEQ ID NO:318)
502	I	GAAGGTGACGGCCACGGAGGT (SEQ ID NO:319)
503	D	CTCCGTGGCCGTCACCTTCCACNNSCTCGTGTGACACA (SEQ ID NO:320)
503	I	GTGGAAGGTGACGGCCACGGA (SEQ ID NO:321)
508	D	CTTCCACGAGCTCGTGTGACANNSTTTGGCCAGACG (SEQ ID NO:322)
508	I	TGTCGACACGAGCTCGTGAA (SEQ ID NO:323)
511	D	GCTCGTGTGACACAGTTTGGCNNSACGGTCAAGGTG (SEQ ID NO:324)
511	I	GCCAACTGTGTGACACGAG (SEQ ID NO:325)
514	D	CACAGTTTGGCCAGACGGTCNNSGTGGCGGGCAACGCCGCGGCC (SEQ ID NO:326)
514	I	GACCGTCTGGCCAACTGTGTGAC (SEQ ID NO:327)
517	D	TGGCCAGACGGTCAAGGTGGCGNNSAACGCCGCGGCCCTGGG (SEQ ID NO:328)
517	I	CGCCACCTTGACCGTCTGGCCAACTG (SEQ ID NO:329)
518	D	CCAGACGGTCAAGGTGGCGGGCNNSGCCGCGGCCCTGGGCAACT (SEQ ID NO:330)
518	I	GCCCGCCACCTTGACCGTCTGGCCAAA (SEQ ID NO:331)
519	D	GACGGTCAAGGTGGCGGGCAACNNSGCCGCGGCCCTGGGCAACT (SEQ ID NO:332)
519	I	GTTGCCCGCCACCTTGACCGTCTGGCC (SEQ ID NO:333)
520	D	GGTCAAGGTGGCGGGCAACGCCNNSGCCCTGGGCAACTGGA (SEQ ID NO:334)
520	I	GCGTGGCCCGCCACCTTGACCGTCTG (SEQ ID NO:335)
525	D	CAACGCCGCGGCCCTGGGCAACNNSAGCACGAGCGCCGCG (SEQ ID NO:336)
525	I	GTTGCCAGGGCCGCGCGTTGCCCGC (SEQ ID NO:337)
527	D	CGCGGCCCTGGGCAACTGGAGCNNSAGCGCCGCGTGGCTC (SEQ ID NO:338)
527	I	GCTCCAGTTGCCAGGGCCGCGCGTT (SEQ ID NO:339)
531	D	CAACTGGAGCACGAGCGCCGCCNNSGCTCTGGACGCCGTCA (SEQ ID NO:340)
531	I	GGCGGCGCTCGTGCTCCAGTTGCCAG (SEQ ID NO:341)
533	D	GAGCACGAGCGCCGCGTGGCTNNSGACGCCGTCAACTATGC (SEQ ID NO:342)
533	I	AGCCACGGCGCGCTCGTGCTCCAGTT (SEQ ID NO:343)



ES 2 502 743 T3

535	D	GAGCGCCGCGCTGGCTCTGGACNSGTCAACTATGCCGATA (SEQ ID NO:344)
535	I	GTCCAGAGCCACGGCGGCGCTCGTGCT (SEQ ID NO:345)
536	D	CGCCGCGCTGGCTCTGGACGCCNNSAACTATGCCGATAACC (SEQ ID NO:346)
536	I	GGCGTCCAGAGCCACGGCGGCGCTCGT (SEQ ID NO:347)
537	D	CGCCGTGGCTCTGGACGCCGTCNNSTATGCCGATAAC (SEQ ID NO:348)
537	D	CGCCGTGGCTCTGGACGCCGTCNNSTATGCCGATAACCACCCC (SEQ ID NO:349)
537	I	GACGGCGTCCAGAGCCACGGCGGCGCT (SEQ ID NO:350)
537	I	GACGGCGTCCAGAGCCACGGCGGCGCT (SEQ ID NO:351)
538	D	CGTGGCTCTGGACGCCGTCAACNNSGCCGATAACCACCCCC (SEQ ID NO:352)
538	I	GTTGACGGCGTCCAGAGCCACGGCGGCG (SEQ ID NO:353)
539	D	GGCTCTGGACGCCGTCAACTATNNSGATAACCACCCCCTGT (SEQ ID NO:354)
539	I	ATAGTTGACGGCGTCCAGAGCCACGGC (SEQ ID NO:355)
540	D	TCTGGACGCCGTCAACTATGCCNNSAACCCCCCCTGTGGATT (SEQ ID NO:356)
540	I	GGCATAGTTGACGGCGTCCAGAGCCAC (SEQ ID NO:357)
541	D	GGACGCCGTCAACTATGCCGATNNSCACCCCCTGTGGATTGGG (SEQ ID NO:358)
541	I	ATCGGCATAGTTGACGGCGTCCAGAGC (SEQ ID NO:359)
545	D	CTATGCCGATAACCACCCCCTGNNSATTGGGACGGTCAACCTC (SEQ ID NO:360)
545	I	CAGGGGGTGGTTATCGGCATAGTTGAC (SEQ ID NO:361)
546	D	TGCCGATAACCACCCCCTGTGGNNSGGGACGGTCAACCTCGAG (SEQ ID NO:362)
546	I	CCACAGGGGGTGGTTATCGGCATAGTT (SEQ ID NO:363)
547	D	CGATAACCACCCCCTGTGGATTNNSACGGTCAACCTCGAGGCT (SEQ ID NO:364)
547	I	AATCCACAGGGGGTGGTTATCGGCATA (SEQ ID NO:365)
549	D	CCACCCCCTGTGGATTGGGACGNNSAACCTCGAGGCTGGAGAC (SEQ ID NO:366)
549	I	CGTCCCAATCCACAGGGGGTGGTTATC (SEQ ID NO:367)
551	D	CCTGTGGATTGGGACGGTCAACNNSGAGGCTGGAGACGTCGTG (SEQ ID NO:368)
551	I	GTTGACCGTCCCAATCCACAGGGGGTG (SEQ ID NO:369)
561	D	TGGAGACGTCGTGGAGTACAAGNNSATCAATGTGGGCCAAGAT (SEQ ID NO:370)
561	I	CTTGTACTCCACGACGTCTCCAGCCTC (SEQ ID NO:371)
563	D	CGTCGTGGAGTACAAGTACATCNNSGTGGGCCAAGATGGCTCC (SEQ ID NO:372)
563	I	GATGTACTTGTACTCCACGACGTCTCC (SEQ ID NO:373)
567	D	CAAGTACATCAATGTGGGCCAANNSGGCTCCGTGACCTGGGAG (SEQ ID NO:374)
567	I	TTGGCCACATTGATGTACTTGTACTC (SEQ ID NO:375)

569	D	CATCAATGTGGGCCAAGATGGCNSGTGACCTGGGAGAGTGAT (SEQ ID NO:376)
569	I	GCCATCTTGGCCACATTGATGTACTTG (SEQ ID NO:377)
577	D	CGTGACCTGGGAGAGTGATCCCNNSCACACTTACACGGTTCCT (SEQ ID NO:378)
577	I	GGGATCACTCTCCCAGGTCACGGAGCC (SEQ ID NO:379)
579	D	CTGGGAGAGTGATCCCAACCACNNSTACACGGTTCCTGCGGTG (SEQ ID NO:380)
579	I	GTGGTTGGGATCACTCTCCCAGGTCAC (SEQ ID NO:381)
583	D	TCCCAACCACACTTACACGGTTNNSGCGGTGGCTTGTGTGACG (SEQ ID NO:382)
583	I	AACCGTGTAAGTGTGGTTGGGATCACT (SEQ ID NO:383)

**Ejemplo 2:****Transformación de SEL de TrGA en *Trichoderma reesei***

[0191] Los SEL se transformaron en *T. reesei* utilizando el método PEG-protoplasto (Véase p. ej., Penttilä *et al.* (1987) Gene 61:155-164). Los clones *E. coli* del SMM confirmado por análisis de secuencia se cultivaron durante la noche a 37 °C en placas de microtitulación de pocillos hondos (Greiner Art. N° 780271) con 1,200 µl del medio 2xYT con ampicilina (100 µg/ml) y kanamicina (50 µg/ml). Los ADN plasmídicos se aislaron de los cultivos utilizando el kit CHEMAGIC® Plasmid Mini (Chemagen - Biopolymer Technologie AG, Baesweiler, Alemania) y se transformaron de forma individual en una cepa huésped *T. reesei* derivada de RL-P37 con 4 deleciones de gen ( $\Delta$ cbh1,  $\Delta$ cbh2,  $\Delta$ egl1,  $\Delta$ egl2, es decir, "con eliminación cuádruple"; véase la patente estadounidense 5.847.276, WO 92/06184 y WO 05/001036) utilizando el método PEG-Protoplasto con las siguientes modificaciones.

[0192] Para la preparación de protoplasto, se cultivaron esporas durante 16-24 horas a 24 °C en un medio mínimo (MM) de *Trichoderma*. ( 20 g/L glucosa, 15 g/L KH<sub>2</sub>PO<sub>4</sub>, pH 4,5; 5g/L (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0,6 g/L MgSO<sub>4</sub>x<sub>7</sub>H<sub>2</sub>O, 0,6 g/L CaCl<sub>2</sub>x<sub>2</sub>H<sub>2</sub>O, 1 ml de 1000X solución de oligoelementos de *T. reesei* {5 g/L FeSO<sub>4</sub>x<sub>7</sub>H<sub>2</sub>O, 1,4 g/L ZnSO<sub>4</sub>x<sub>7</sub>H<sub>2</sub>O, 1,6 g/L MnSO<sub>4</sub>x<sub>2</sub>H<sub>2</sub>O, 3,7 g/L CoCl<sub>2</sub>x<sub>6</sub>H<sub>2</sub>O}) con agitación a 150 rpm. La germinación de esporas se recolectó mediante centrifugación y se trató con 15 mg/ml de solución β-D-glucanasa-G (Interspex - Art.No. 0439-1) para lisar las paredes de las células fúngicas. Se llevó a cabo una preparación adicional de protoplastos mediante un método convencional, como lo describe Penttilä *et al.* (1987 *supra*). 2) El método de transformación se redujo 10 veces. En general, las mezclas de transformación que contenían hasta 600 ng de ADN y 1-5x 10<sup>5</sup> protoplastos en un volumen total de 25 µl se trataron con 200 ml de 25 % solución PEG, diluido con 2 volúmenes de 1,2 M solución sorbitol, mezclado con 3 % MM agarosa superior selectiva con acetamida (el mismo medio mínimo que se ha mencionado anteriormente pero (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> se sustituyó con 20 mM acetamida) y se vertió en 2 % agarosa selectiva con acetamida bien en placas de microtitulación de 24 pocillos o en una bandeja Q de 20x20 cm dividida en 48 pocillos. Las placas se incubaron a 28 °C durante un periodo de 5 a 8 días. Las esporas de la población total de transformantes regeneradas en cada pocillo individual se recolectaron de las placas utilizando una solución de 0,85 % NaCl, 0,015 % Tween 80. Las suspensiones de esporas se utilizaron para inocular fermentaciones en PMT de 96 pocillos. En el caso de PMT de 24 pocillos, se introdujo una etapa adicional de recubrimiento sobre una PMT de 24 pocillos nueva con MM acetamida selectiva con el fin de enriquecer el número de esporas.

**Ejemplo 3:****Fermentación de transformantes de *T. reesei* que expresan variantes de TrGA en un formato PMT**

[0193] Los transformantes se fermentaron en placas de microtitulación de filtrado y se utilizaron para los análisis los sobrenadantes de cultivo obtenidos con variantes de proteína de TrGA expresadas. En resumen, las placas de filtrado de 96 pocillos (Corning Art. N° 3505) con 200 µl de medio LD-GSM (5,0 g/L (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 33 g/L 1,4-Piperazinabis(ácido propanosulfónico), pH 5,5, 9,0 g/L Casamino ácidos, 1,0 g/L KH<sub>2</sub>PO<sub>4</sub>, 1,0 g/L CaCl<sub>2</sub>x<sub>2</sub>H<sub>2</sub>O, 1,0 g/L MgSO<sub>4</sub>x<sub>7</sub>H<sub>2</sub>O, 2,5 ml/L de 1000X oligoelementos *T. reesei* 20 g/L glucosa, 10 g/L soforosa) se inocularon por cuadruplicado con suspensiones de esporas de transformantes *T. reesei* que expresan variantes de TrGA (más de 104 esporas por pocillo). Las placas se incubaron a 28 °C con agitación a 230 rpm y 80 % de humedad durante 6 días. Los sobrenadantes de cultivo se recolectaron mediante filtración al vacío. Los sobrenadantes se utilizaron en diferentes análisis para la detección de variantes con propiedades mejoradas.

**Ejemplo 4:**

**Preparación de las muestras de caldo integrales a partir de transformantes que producen GA**

5 [0194] Los transformantes que producen TrGA se cultivaron inicialmente en matraces de agitación de 250 ml con 30 ml de medio Proflo. El medio Proflo contenía: 30 g/L  $\alpha$ -lactosa, 6,5 g/L  $(\text{NH}_4)_2\text{SO}_4$ , 2 g/L  $\text{KH}_2\text{PO}_4$ , 0,3 g/L  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ , 0,2 g/L  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ , 1 ml/L 1000X solución salina de oligoelementos como se ha mencionado anteriormente, 2 ml/L 10% Tween 80, 22,5 g/L harina de semilla de algodón ProFlo (Traders protein, Memphis, TN), 0,72 g/L  $\text{CaCO}_3$ . Tras dos días de cultivo a 28 °C y 140 rpm, 10 % del cultivo Proflo se transfirió a un matraz de agitación de 250 ml con 30 ml de Medio definido de lactosa. La composición del medio definido de lactosa era como se muestra a continuación: 5 g/L  $(\text{NH}_4)_2\text{SO}_4$ , 33 g/L tampón 1,4-Piperazinabis (ácido propanosulfónico), pH 5,5, 9 g/L casamino ácidos, 4,5 g/L  $\text{KH}_2\text{PO}_4$ , 1,0 g/L  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ , 5 ml/L Mazu DF60-P antiespumante (Mazur Chemicals, IL), 1ml/L of 1000X solución de oligoelementos. 40 ml/L de 40% (p/v) solución lactosa se añadió al medio tras la esterilización. Los matraces de agitación con el medio definido de lactosa se incubaron a 28 °C, 140 rpm durante 4 - 5 días.

15 [0195] Se extrajo el micelio de las muestras de cultivo mediante centrifugación y se analizó el sobrenadante para detectar el contenido de proteína total (BCA Protein Assay Kit, Pierce Cat. No.23225) y la actividad GA, tal y como se describe anteriormente en la sección experimental.

20 [0196] El perfil proteínico de las muestras de caldo integrales se determinó mediante electroforesis SDS PAGE. Las muestras del sobrenadante de cultivo se mezclaron con el mismo volumen de 2X tampón de carga de muestra con agente reductor y se separaron en NUPAGE® Novex 10% Bis-Tris Gel con MES SDS tampón de migración (Invitrogen, Carlsbad, CA, USA). Las bandas de polipéptidos se visualizaron en el gel SDS con SIMPLYBLUE SafeStain (Invitrogen, Carlsbad, CA, USA).

[0197] El ejemplo 5-7 proporciona variantes en el dominio catalítico con propiedades mejoradas. El ejemplo 8-10 proporciona variantes en el dominio de unión a almidón con propiedades mejoradas.

**Ejemplo 5: Variantes en el dominio catalítico con estabilidad térmica mejorada**

25 [0198] La molécula TrGA precursora presentaba una actividad residual entre 15 y 44 % (variación diaria) según las condiciones descritas. El índice de rendimiento se calculó según la termoestabilidad de TrGA WT del mismo lote. Los índices de rendimiento son los cocientes  $P1 = (\text{actividad residual variante})/(\text{actividad residual TrGA WT})$ . Al utilizar este cociente, un índice de rendimiento  $>1$  indica una estabilidad mejorada. Las variantes que presentaban un índice de rendimiento de estabilidad térmica de más de 1,0 se muestran en la siguiente tabla 2.

Tabla 2- Detección estabilidad térmica

Variante	PI estabilidad térmica
T010F	1.11
T010G	1.13
T010M	1.12
T010Q	1.06
T010R	1.06
T010S	1.24
T042V	1.31
F059A	1.03
F059G	1.07
F059L	1.06
F059M	1.12
F059Q	1.10
F059V	1.03
N061V	1.06
E068C	1.23
E068F	1.14
E068G	1.15
E068I	1.19
E068K	1.01
E068M	1.29
E068N	1.18
E068Q	1.15
E068W	1.09
A072E	1.07
A072Q	1.02
G073F	1.44
G073M	1.01
G073N	1.10
G073W	1.36
S097E	1.08
S097G	1.03
S097T	1.17
S097V	1.03
L098C	1.07
A099C	1.07
A099F	1.02
A099I	1.14

A099L	1.09
A099Q	1.07
A099R	1.02
A099S	1.11
A099T	1.02
S102A	1.02
K114C	1.19
K114D	1.17
K114E	1.16
K114L	1.10
K114M	1.21
K114Q	1.25
I133V	1.21
K140D	1.09
K140Q	1.06
K140S	1.05
K140W	1.04
N144A	1.11
N144F	1.06
S152C	1.04
S152G	1.09
S152I	1.09
S152N	1.12
N153A	1.28
N153D	1.06
N153E	1.29
N153F	1.16
N153H	1.01
N153L	1.06
N153M	1.27
N153S	1.31
N153V	1.34
N153W	1.19
N182R	1.02
A204D	1.02
A204M	1.08
T205C	1.02
T205D	1.06
T205N	1.09
T205P	1.17

ES 2 502 743 T3

T205S	1.04
T205V	1.06
T205Y	1.07
S214C	1.02
S214E	1.03
S214N	1.07
S214P	1.04
S214Q	1.13
S214T	1.06
S214V	1.13
S214W	1.03
S214Y	1.04
V216I	1.13
W228A	1.01
W228F	1.12
W228G	1.06
W228H	1.05
W228I	1.06
W228L	1.14
W228M	1.04
W228Q	1.15
W228S	1.15
W228T	1.06
W228V	1.21
W228Y	1.10
V229A	1.16
V229D	1.18
V229E	1.16
V229F	1.17
V229G	1.18
V229H	1.13
V229I	1.21
V229L	1.20
V229M	1.11
V229N	1.07
V229P	1.13
V229Q	1.10
V229R	1.07
V229S	1.12
V229T	1.11
V229W	1.07
V229Y	1.08
S230C	1.09
S230D	1.08
S230E	1.11
S230F	1.08
S230G	1.09
S230H	1.03
S230K	1.02
S230L	1.02
S230M	1.08
S230N	1.16
S230P	1.12
S230Q	1.20

S230R	1.11
S230T	1.07
S230V	1.11
S230Y	1.05
S231A	1.06
S231C	1.07
S231D	1.18
S231E	1.10
S231F	1.14
S231K	1.09
S231L	1.15
S231M	1.09
S231N	1.13
S231Q	1.18
S231R	1.13
S231T	1.16
S231V	1.21
S231W	1.06
S231Y	1.10
D236A	1.10
D236C	1.16
D236E	1.06
D236F	1.11
D236G	1.07
D236H	1.16
D236I	1.14
D236K	1.13
D236L	1.08
D236M	1.15
D236N	1.15
D236P	1.06
D236R	1.29
D236T	1.16
D236V	1.18
D236Y	1.07
T241V	1.05
N242F	1.03
N242H	1.01
N263A	1.18
N263C	1.14
N263L	1.02
N263M	1.10
N263R	1.04
N263T	1.07
L264A	1.03
L264D	1.20
L264I	1.14
L264K	1.24
L264M	1.15
L264P	1.02
L264R	1.02
L264V	1.04
L264Y	1.15
G265D	1.08

ES 2 502 743 T3

G265E	1.03
G265F	1.08
G265H	1.09
G265I	1.09
G265K	1.05
G265L	1.07
G265N	1.15
G265P	1.07
G265Q	1.15
G265R	1.12
G265S	1.13
G265T	1.17
G265V	1.15
G265W	1.11
G265Y	1.10
A268C	1.16
A268D	1.20
A268E	1.10
A268F	1.07
A268G	1.11
A268I	1.12
A268K	1.11
A268L	1.11
A268M	1.18
A268N	1.10
A268P	1.13
A268R	1.15
A268S	1.09
A268T	1.14
A268W	1.05
G269K	1.01
G269N	1.08
G269P	1.08
G269Q	1.02
G269R	1.06
D276E	1.03
D276Q	1.03
D276S	1.02
D276V	1.17
V284C	1.09
V284E	1.14
V284G	1.13
V284I	1.10
V284Q	1.08
V284R	1.09
V284S	1.06
V284T	1.08
V284W	1.08
V284Y	1.12
S291A	1.21
S291D	1.02
S291E	1.02
S291F	1.24
S291G	1.05

S291H	1.25
S291I	1.10
S291K	1.13
S291M	1.21
S291N	1.08
S291P	1.11
S291Q	1.08
S291R	1.07
S291T	1.21
S291V	1.06
S291W	1.15
P300A	1.10
P300N	1.06
P300Q	1.05
P300R	1.02
P300T	1.07
P300V	1.02
P300W	1.11
P300Y	1.04
A301A	1.13
A301K	1.11
A301L	1.07
A301P	1.22
A301R	1.21
A301S	1.12
A301T	1.14
A301V	1.07
A301W	1.14
A301Y	1.14
A303E	1.02
A303I	1.02
A303L	1.03
A303Q	1.01
A311C	1.05
A311E	1.06
A311G	1.11
A311H	1.09
A311I	1.04
A311K	1.13
A311L	1.07
A311R	1.07
A311S	1.01
A311Y	1.02
V338H	1.12
V338I	1.21
V338L	1.17
V338M	1.13
V338N	1.23
V338P	1.18
V338Q	1.20
V338S	1.19
V338Y	1.10
T342C	1.06
T342I	1.04

ES 2 502 743 T3

T342L	1.17
T342P	1.08
S344A	1.05
S344C	1.16
S344D	1.02
S344F	1.10
S344K	1.12
S344M	1.26
S344N	1.16
S344P	1.20
S344Q	1.22
S344R	1.22
S344T	1.19
S344V	1.20
S344W	1.03
T346D	1.03
T346L	1.02
T346M	1.07
T346N	1.15
T346P	1.13
T346Q	1.12
T346S	1.13
T346V	1.09
T346W	1.06
T346Y	1.07
A349D	1.15
A349E	1.13
A349F	1.09
A349G	1.12
A349H	1.11
A349K	1.14
A349L	1.16
A349M	1.07
A349N	1.09
A349P	1.03
A349Q	1.09
A349R	1.04
A349T	1.09
A349V	1.08
A349W	1.04
A349Y	1.04
V359Q	1.02
V359R	1.13
V359Y	1.01
G361A	1.13
G361C	1.16
G361D	1.35
G361E	1.24
G361F	1.20
G361H	1.11
G361I	1.20
G361K	1.16
G361L	1.22
G361M	1.28

G361P	1.27
G361R	1.19
G361S	1.22
G361T	1.16
G361V	1.15
G361W	1.22
G361Y	1.26
A364C	1.18
A364D	1.25
A364E	1.27
A364F	1.33
A364G	1.25
A364K	1.27
A364L	1.31
A364M	1.21
A364Q	1.19
A364R	1.28
A364S	1.23
A364T	1.23
A364V	1.23
A364W	1.23
T375A	1.17
T375E	1.03
T375F	1.06
T375G	1.05
T375H	1.13
T375K	1.10
T375L	1.03
T375M	1.17
T375N	1.20
T375P	1.15
T375V	1.16
T375W	1.11
T375Y	1.17
N379A	1.11
N379C	1.03
N379D	1.06
N379F	1.07
N379G	1.09
N379H	1.07
N379I	1.04
N379K	1.04
N379L	1.04
N379M	1.08
N379P	1.17
N379Q	1.11
N379R	1.07
N379T	1.09
N379V	1.13
N379W	1.09
N379Y	1.06
S382A	1.06
S382D	1.04
S382E	1.05

ES 2 502 743 T3

S382G	1.06
S382I	1.07
S382K	1.03
S382N	1.07
S382P	1.10
S382R	1.05
S382T	1.02
S382Y	1.04
S390M	1.06
S390Q	1.02
S390R	1.02
E391L	1.04
E391R	1.02
E391S	1.02
E391W	1.04
E391Y	1.02
A393D	1.11
A393E	1.03
A393F	1.09
A393G	1.12
A393H	1.05
A393I	1.06
A393K	1.05
A393L	1.17
A393M	1.07
A393N	1.11
A393O	1.02
A393R	1.03
A393S	1.13
A393T	1.08
A393V	1.09
A393W	1.10
A393Y	1.12
K394A	1.08
K394C	1.06
K394E	1.07
K394F	1.09
K394G	1.05
K394H	1.08
K394L	1.08
K394M	1.06
K394Q	1.07
K394R	1.06
K394T	1.03
K394V	1.02
S410E	1.05
S410H	1.11
S410I	1.06
S410K	1.11
S410L	1.05
S410M	1.02
S410N	1.10
S410Q	1.15
S410R	1.18

S410T	1.04
S410V	1.13
L417I	1.04
L417K	1.20
L417M	1.05
L417Q	1.04
L417R	1.20
L417V	1.07
L417Y	1.01
T430A	1.05
T430E	1.02
T430F	1.06
T430H	1.10
T430I	1.04
T430K	1.08
T430M	1.17
T430N	1.13
T430Q	1.05
T430R	1.13
T430S	1.17
T430V	1.05
A431I	1.03
A431N	1.03
A431P	1.08
A431R	1.08
A431V	1.03
R433A	1.13
R433C	1.24
R433E	1.22
R433F	1.17
R433G	1.23
R433K	1.12
R433L	1.23
R433M	1.10
R433N	1.28
R433S	1.23
R433V	1.28
R433W	1.16
R433Y	1.18
I436E	1.07
I436F	1.02
I436G	1.09
I436H	1.20
I436K	1.14
I436P	1.15
I436R	1.16
I436S	1.17
I436T	1.18
I436V	1.12
I436Y	1.05
A442N	1.06
A442R	1.04
A442T	1.09
S444E	1.04



S444K	1.07
S444M	1.05
S444Q	1.04
T448A	1.02
T448E	1.09
T448F	1.12
T448I	1.12
T448L	1.09
T448M	1.11
T448Q	1.17
T448R	1.09

T448S	1.07
T448V	1.13
T448W	1.01
T448Y	1.08
S451E	1.04
S451H	1.18
S451K	1.08
S451L	1.01
S451Q	1.06
S451T	1.02

**Ejemplo 6: Variantes en el dominio catalítico con actividad específica mejorada (AE) en un análisis de detección de etanol**

- 5 [0199] Se analizaron las variantes en un análisis de detección de etanol utilizando los análisis descritos anteriormente. La tabla 3 muestra los resultados del análisis de detección para las variantes con un índice de rendimiento (PI) > 1,0 en comparación con el PI de la TrGA precursora. El PI de la actividad específica es el cociente "actividad específica-variante/actividad específica-WT". Al utilizar esto, el PI de la actividad específica para TrGA de tipo silvestre es 1,0 y una variante con un PI > 1,0 tiene una actividad específica que es mayor que
- 10 la TrGA precursora. La actividad específica se determinó en este ejemplo como la actividad medida mediante análisis de detección de etanol dividida por los resultados obtenidos en el análisis Caliper descrito anteriormente.

**Tabla 3- Detección etanol**

Variante	PI Actividad Específica
T010D	1.09
T010F	1.06
T010G	1.12
T010K	1.05
T010L	1.07
T010M	1.05
T010P	1.05
T010R	1.08
T010S	1.09
L014E	1.03
L014H	1.05
N015D	1.02
P023A	1.16
P023F	1.13
P023N	1.05
F059A	1.17
F059F	1.05
F059G	1.05
K060F	1.06
K060H	1.03
N061D	1.05
N061I	1.21
N061L	1.18
N061Q	1.08
N061V	1.11

N061W	1.02
R065A	1.17
R065C	1.08
R065G	1.08
R065I	1.11
R065K	1.09
R065M	1.07
R065S	1.12
R065V	1.14
R065Y	1.01
T067C	1.14
T067I	1.13
T067K	1.05
T067M	1.22
E068I	1.06
E068M	1.10
E068W	1.03
A072E	1.11
A072G	1.02
A072L	1.03
A072M	1.11
A072Q	1.10
A072R	1.10
A072W	1.06
A072Y	1.30
G073C	1.02
G073L	1.07
G073W	1.03
S097F	1.11

ES 2 502 743 T3

S097M	1.11
S097N	1.23
S097P	1.18
S097R	1.07
S097V	1.12
S097W	1.09
S097Y	1.18
L098H	1.04
L098M	1.09
A099C	1.07
A099L	1.01
A099M	1.03
A099N	1.11
A099P	1.08
S102A	1.20
S102C	1.04
S102I	1.04
S102L	1.05
S102M	1.25
S102N	1.19
S102R	1.21
S102V	1.07
S102W	1.06
S102Y	1.10
E110Q	1.02
E110S	1.07
E110W	1.11
L113E	1.15
L113N	1.08
I133K	1.04
I133R	1.16
I133S	1.08
I133T	1.29
K140A	1.04
K140E	1.04
K140F	1.03
K140H	1.14
K140L	1.10
K140M	1.11
K140N	1.15
K140Q	1.08
K140R	1.12
K140S	1.13
K140V	1.15
K140W	1.07
K140Y	1.06
N144C	1.05
N144D	1.15
N144E	1.16
N144I	1.13
N144K	1.05
N145A	1.07
N145C	1.09
N145E	1.03

N145I	1.20
N145K	1.05
N145L	1.03
N145M	1.07
N145Q	1.14
N145R	1.11
N145V	1.12
N145W	1.14
N145Y	1.05
Y147A	1.02
Y147M	1.02
Y147R	1.12
S152H	1.08
S152M	1.10
N153C	1.09
N153D	1.20
N153K	1.13
N153L	1.12
N153W	1.07
N153Y	1.13
N164A	1.02
N164G	1.03
N182C	1.12
N182E	1.13
N182K	1.07
N182P	1.01
N182R	1.03
A204C	1.04
A204D	1.09
A204G	1.02
A204I	1.06
A204M	1.09
A204Q	1.09
A204T	1.05
T205A	1.03
T205D	1.03
T205H	1.03
T205I	1.05
T205K	1.09
T205M	1.05
T205N	1.09
T205P	1.17
T205Q	1.25
T205S	1.10
T205V	1.06
T205W	1.05
T205Y	1.18
S214P	1.08
S214T	1.07
V216C	1.08
V216G	1.05
V216H	1.03
V216K	1.02
V216N	1.13

ES 2 502 743 T3

V216Y	1.09
Q219D	1.05
Q219G	1.06
Q219H	1.03
Q219N	1.08
Q219P	1.16
Q219S	1.29
W228A	1.20
W228F	1.22
W228G	1.17
W228H	1.33
W228I	1.18
W228L	1.12
W228M	1.35
W228T	1.19
V229E	1.01
V229I	1.02
V229M	1.03
V229N	1.01
V229Q	1.02
S230C	1.23
S230D	1.13
S230E	1.10
S230F	1.63
S230G	1.77
S230H	1.05
S230I	1.18
S230K	1.04
S230L	1.20
S230N	1.23
S230P	1.13
S230Q	1.20
S230R	1.84
S230T	1.11
S230V	1.12
S230Y	1.08
S231C	1.13
S231D	1.08
S231F	1.17
S231L	1.29
S231M	1.08
S231N	1.04
S231Q	1.05
S231R	1.02
S231V	1.07
S231Y	1.07
D236F	1.14
D236G	1.05
D236L	1.11
D236M	1.07
D236N	1.03
D236P	1.06
D236S	1.03
D236T	1.14

D236V	1.04
I239M	1.04
I239Q	1.08
I239S	1.11
I239V	1.52
I239W	1.02
I239Y	1.25
T241C	1.07
T241E	1.03
T241H	1.10
T241L	1.04
T241M	1.05
T241P	1.02
T241S	1.08
T241V	1.05
N242C	1.08
N242F	1.06
N242M	1.04
N242T	1.08
N242V	1.03
N242W	1.05
N263H	1.05
N263K	1.02
N263P	1.40
L264A	1.04
L264C	1.08
L264E	1.16
L264F	1.03
L264S	1.05
G265E	1.10
G265H	1.12
G265I	1.06
G265K	1.03
G265R	1.06
G265T	1.10
A268C	1.50
A268D	1.14
A268E	1.18
A268F	1.15
A268G	1.35
A268I	1.15
A268K	1.23
A268L	1.06
A268P	1.08
A268R	1.14
A268T	1.18
A268W	1.05
G269E	1.01
D276S	1.01
V284R	1.06
V284T	1.05
V284Y	1.07
S291A	1.26
S291E	1.09

ES 2 502 743 T3

S291F	1.13
S291H	1.13
S291K	1.07
S291N	1.04
S291P	1.12
S291W	1.04
P300K	1.10
P300P	1.12
P300R	1.11
A301A	1.01
A301E	1.09
A301K	1.09
A301L	1.05
A301P	1.02
A301S	1.03
A301W	1.04
A303C	1.06
A303D	1.04
A303F	1.09
A303H	1.05
A303J	1.09
A303K	1.02
A303L	1.05
A303N	1.04
A303R	1.10
A303T	1.11
A303V	1.04
A303W	1.15
A303Y	1.07
A311N	1.04
A311P	1.09
A311Q	1.19
A311S	1.01
A311Y	1.06
V338P	1.04
V338Q	1.12
V338S	1.14
V338Y	1.05
T342N	1.06
T342V	1.23
S344A	1.05
S344T	1.01
T346G	1.14
T346H	1.07
T346M	1.06
T346N	1.09
T346P	1.07
T346Q	1.04
T346Y	1.06
A349L	1.02
V359I	1.14
V359K	1.08
V359M	1.09
V359N	1.02

V359Q	1.14
V359R	1.15
V359W	1.04
G361H	1.11
G361L	1.04
G361R	1.04
A364M	1.05
A364W	1.07
T375C	1.12
T375D	1.01
T375E	1.02
T375H	1.05
T375V	1.04
T375W	1.02
T375Y	1.05
N379A	1.05
N379C	1.10
N379D	1.05
N379G	1.07
N379I	1.01
N379M	1.04
N379P	1.06
N379S	1.01
S382A	1.02
S382N	1.02
S382P	1.10
S382S	1.09
S382V	1.10
S382W	1.10
S390A	1.05
S390Y	1.03
E391A	1.17
E391E	1.10
E391I	1.13
E391K	1.18
E391L	1.18
E391M	1.05
E391Q	1.04
E391R	1.08
E391V	1.05
E391W	1.12
E391Y	1.08
A393E	1.05
A393G	1.14
A393H	1.10
A393I	1.07
A393K	1.09
A393L	1.12
A393M	1.07
A393N	1.18
A393Q	1.02
A393R	1.09
A393S	1.10
A393T	1.13

A393V	1.16
A393W	1.04
A393Y	1.03
K394A	1.11
K394H	1.11
K394K	1.08
K394L	1.01
K394M	1.14
K394Q	1.09
K394R	1.19
K394S	1.22
K394T	1.08
K394V	1.05
K394W	1.13
S410E	1.01
S410H	1.06
S410N	1.04
L417A	1.12
L417D	1.19
L417E	1.10
L417F	1.08
L417G	1.19
L417I	1.10
L417K	1.02
L417Q	1.04
L417R	1.30
L417S	1.05
L417T	1.10
L417V	1.21
L417W	1.05
L417Y	1.10
H418E	1.01
H418M	1.12
T430A	1.19
T430E	1.15
T430F	1.09
T430G	1.16
T430H	1.15
T430I	1.06
T430K	1.24
T430M	1.16
T430N	1.07
T430Q	1.15
T430R	1.04
T430V	1.09
A431C	1.04
A431E	1.08
A431H	1.11
A431I	1.20
A431L	1.21
A431M	1.12
A431Q	1.22
A431R	1.11
A431S	1.09

A431W	1.04
A431Y	1.13
R433A	1.09
R433M	1.17
R433W	1.06
R433Y	1.22
A442A	1.13
S444K	1.03
S444M	1.13
S444N	1.04
S444P	1.08
S444Q	1.08
S444R	1.03
S444T	1.15
S444V	1.14
S444W	1.16
T448F	1.02
T448G	1.08
T448I	1.10
T448P	1.08
T448Q	1.04
T448V	1.04
T451K	1.29

#### Ejemplo 7: Variantes de actividad específica y de termoestabilidad combinadas en el dominio catalítico

[0200] La tabla 4 muestra las variantes que presentaban un índice de rendimiento (PI) de 1,0 o mejor en comparación con la precursora para ambas propiedades: actividad específica y termoestabilidad. Estas incluían los siguientes lugares. 10, 15, 59, 61, 68, 72, 73, 97, 99, 102, 133, 140, 153, 182, 204, 205, 214, 228, 229, 230, 231, 236, 241, 242, 264, 265, 268, 275, 284, 291, 300, 301, 303, 311, 338, 344, 346, 359, 361, 364, 375, 370, 382, 391, 393, 394, 410, 417, 430, 431, 433, 444, 448 y 451. Los lugares que muestran la actividad específica y termoestabilidad combinadas más altas incluían: 228, 230, 231, 268, 291, 417, 433 y 451.

**Tabla 4: Variantes combinadas**

Variante	PI de Act. Específica	PI de Estabi. Térmica
T010F	1.06	1.11
T010G	1.12	1.13
T010M	1.05	1.12
T010R	1.08	1.06
T010S	1.09	1.24
N015N	1.06	1.06
F059A	1.17	1.03
F059G	1.05	1.07
N061V	1.11	1.06
E068I	1.06	1.19
E068M	1.10	1.29
E068W	1.03	1.09
A072E	1.11	1.07
A072Q	1.10	1.02
G073W	1.03	1.36
S097V	1.12	1.03
A099C	1.07	1.07
A099L	1.01	1.09
S102A	1.20	1.02
K140Q	1.08	1.06
K140S	1.13	1.05
K140W	1.07	1.04
N153D	1.20	1.06
N153L	1.12	1.06
N153W	1.07	1.19
N182R	1.03	1.02
A204D	1.09	1.02
A204M	1.09	1.08
T205D	1.03	1.06
T205N	1.09	1.09
T205P	1.17	1.17
T205S	1.10	1.04
T205V	1.06	1.06
T205Y	1.18	1.07
S214P	1.08	1.04
S214T	1.07	1.06
W228A	1.20	1.01
W228F	1.22	1.12

W228G	1.17	1.06
W228H	1.33	1.05
W228I	1.18	1.06
W228L	1.12	1.14
W228M	1.35	1.04
W228T	1.19	1.06
V229E	1.01	1.16
V229I	1.02	1.21
V229M	1.03	1.11
V229N	1.01	1.07
V229Q	1.02	1.10
S230C	1.23	1.09
S230D	1.13	1.08
S230E	1.10	1.11
S230F	1.63	1.08
S230G	1.77	1.09
S230H	1.05	1.03
S230K	1.04	1.02
S230L	1.20	1.02
S230N	1.23	1.16
S230P	1.13	1.12
S230Q	1.20	1.20
S230R	1.84	1.11
S230T	1.11	1.07
S230V	1.12	1.11
S230Y	1.08	1.05
S231C	1.13	1.07
S231D	1.08	1.18
S231F	1.17	1.14
S231L	1.29	1.15
S231M	1.08	1.09
S231N	1.04	1.13
S231Q	1.05	1.18
S231R	1.02	1.13
S231V	1.07	1.21
S231Y	1.07	1.10
D236F	1.14	1.11
D236G	1.05	1.07
D236L	1.11	1.08
D236M	1.07	1.15
D236N	1.03	1.15
D236P	1.06	1.06

ES 2 502 743 T3

D236T	1.14	1.16
D236V	1.04	1.18
T241V	1.05	1.05
N242F	1.06	1.03
L264A	1.04	1.03
G265E	1.10	1.03
G265H	1.12	1.09
G265I	1.06	1.09
G265K	1.03	1.05
G265R	1.06	1.12
G265T	1.10	1.17
A268C	1.50	1.16
A268D	1.14	1.20
A268E	1.18	1.10
A268F	1.15	1.07
A268G	1.35	1.11
A268I	1.15	1.12
A268K	1.23	1.11
A268L	1.06	1.11
A268P	1.08	1.13
A268R	1.14	1.15
A268T	1.18	1.14
A268W	1.05	1.05
D276S	1.01	1.02
V284R	1.06	1.09
V284T	1.05	1.08
V284Y	1.07	1.12
S291A	1.26	1.21
S291E	1.09	1.02
S291F	1.13	1.24
S291H	1.13	1.25
S291K	1.07	1.13
S291N	1.04	1.08
S291P	1.12	1.11
S291W	1.04	1.15
P300R	1.11	1.02
A301K	1.09	1.11
A301L	1.05	1.07
A301P	1.02	1.22
A301S	1.03	1.12
A301W	1.04	1.14
A303I	1.09	1.02
A303L	1.05	1.03
A311S	1.01	1.01
A311Y	1.06	1.02
V338P	1.04	1.18
V338Q	1.12	1.20
V338S	1.14	1.19
V338Y	1.05	1.10
S344A	1.05	1.05
S344T	1.01	1.19
T346M	1.06	1.07
T346N	1.09	1.15
T346P	1.07	1.13

T346Q	1.04	1.12
T346Y	1.06	1.07
A349L	1.02	1.16
V359Q	1.14	1.02
V359R	1.15	1.13
G361H	1.11	1.11
G361L	1.04	1.22
G361R	1.04	1.19
A364M	1.05	1.21
A364W	1.07	1.23
T375E	1.02	1.03
T375H	1.05	1.13
T375V	1.04	1.16
T375W	1.02	1.11
T375Y	1.05	1.17
N379A	1.05	1.11
N379C	1.10	1.03
N379D	1.05	1.06
N379G	1.07	1.09
N379I	1.01	1.04
N379M	1.04	1.08
N379P	1.06	1.17
S382A	1.02	1.06
S382N	1.02	1.07
S382P	1.10	1.10
E391L	1.18	1.04
E391R	1.08	1.02
E391W	1.12	1.04
E391Y	1.08	1.02
A393E	1.05	1.03
A393G	1.14	1.12
A393H	1.10	1.05
A393I	1.07	1.06
A393K	1.09	1.05
A393L	1.12	1.17
A393M	1.07	1.07
A393N	1.18	1.11
A393Q	1.02	1.02
A393R	1.09	1.03
A393S	1.10	1.13
A393T	1.13	1.08
A393V	1.16	1.09
A393W	1.04	1.10
A393Y	1.03	1.12
K394A	1.11	1.08
K394H	1.11	1.08
K394L	1.01	1.08
K394M	1.14	1.06
K394Q	1.09	1.07
K394R	1.19	1.06
K394T	1.08	1.03
K394V	1.05	1.02
S410E	1.01	1.05
S410H	1.06	1.11

S410N	1.04	1.10
L417I	1.10	1.04
L417K	1.02	1.20
L417Q	1.04	1.04
L417R	1.30	1.20
L417V	1.21	1.07
L417Y	1.10	1.01
T430A	1.19	1.05
T430E	1.15	1.02
T430F	1.09	1.06
T430H	1.15	1.10
T430I	1.06	1.04
T430K	1.24	1.08
T430M	1.16	1.17
T430N	1.07	1.13
T430Q	1.15	1.05
T430R	1.04	1.13

T430V	1.09	1.05
A431I	1.20	1.03
A431R	1.11	1.08
R433A	1.09	1.13
R433M	1.17	1.10
R433W	1.06	1.16
R433Y	1.22	1.18
S444K	1.03	1.07
S444M	1.13	1.05
S444Q	1.08	1.04
T448F	1.02	1.12
T448I	1.10	1.12
T448Q	1.04	1.17
T448V	1.04	1.13
S451K	1.29	1.08

**Ejemplo 8: Variantes con dominio de unión a almidón con estabilidad térmica mejorada**

5 [0201] La molécula TrGA precursora presentaba una actividad residual entre 15 y 44 % (variación diaria) según las condiciones descritas. El índice de rendimiento se calculó según la termoestabilidad de TrGA WT del mismo lote. Los índices de rendimiento son los cocientes  $P1 = (\text{actividad residual variante})/(\text{actividad residual TrGA WT})$ . Al utilizar este cociente, un índice de rendimiento  $>1$  indica una estabilidad mejorada. Las variantes que presentaban un índice de rendimiento de estabilidad térmica de más del 1,0 se muestran en la siguiente tabla 5.

Tabla 5 - Detección estabilidad térmica – SBD

Variante	PI de estabilidad térmica
T493I	1,15
T495K	1,20
T495R	1,10
T495S	1,23
E503A	1,43
E503C	1,39
E503S	1,02
E503T	1,04
E503V	1,68
Q508H	1,19
Q508R	1,29
Q508S	1,13
QS11A	1,11
Q511D	1,12
Q511H	1,33
QS11N	1,14
Q511S	1,15
N518S	1,17
A519E	1,11
A519K	1,33
A519R	1,32



A519T	1,06
A519V	1,07
A519Y	1,29
A520C	1,10
A520L	1,06
A520P	1,06
T527A	1,08
T527V	1,12
V531L	1,21
A535D	1,12
A535K	1,24
A535N	1,37
A535P	1,59
A535R	1,26
V536I	1,02
V536R	1,10
N537W	1,16
A539E	1,32
A539H	1,17
A539M	1,05
A539R	1,36
A539S	1,30
N563A	1,14
N563C	1,49
N563E	1,44
N563I	1,65
N563K	1,77
N563L	1,60
N563Q	1,29
N563T	1,31
N563V	1,53
N577A	1,19
N577K	1,23
N577P	1,41
N577R	1,31
N577V	1,12

**Ejemplo 9: Variantes con dominio de unión a almidón con actividad específica mejorada (AE) en un análisis de detección de etanol**

5 **[0202]** Se analizaron las variantes en un análisis de detección de etanol utilizando los análisis descritos anteriormente. La tabla 6 muestra los resultados del análisis de detección para las variantes con un índice de rendimiento (PI) > 1,0 en comparación con el PI de la TrGA precursora. El PI de las mutaciones mostró un índice de rendimiento de la actividad específica para TrGA de tipo silvestre es 1,0 y una variante con una amplia variedad de sustituciones en estos lugares tuvo como resultado termoestabilidad aumentada. Estos lugares

incluían 539.

**Tabla 6 Detección de Etanol SBD**

Variante	PI Actividad Específica
T493C	1,05
T493M	1,19
T493N	1,07
T493Q	1,03
T493Y	1,07
P494H	1,03
P494I	1,12
P494M	1,12
P494N	1,16
P494Q	1,02
P494W	1,07
T495M	1,49
T495P	1,04
T495R	1,15
H502A	1,16
H502M	1,13
H502S	1,15
H502V	1,10
E503C	1,05
E503D	1,06
E503H	1,01
E503S	1,10
E503W	1,04
Q508N	1,11
Q508P	1,07
Q508Y	1,09
Q511C	1,07
Q511G	1,06
Q511H	1,05
Q511I	1,10
Q511K	1,09
Q511T	1,04
Q511V	1,04
N518P	1,13
N518T	1,02
A519I	1,21
A520C	1,38
A520E	1,16

A520L	1,46
A520P	1,50
A520Q	1,05
A520R	1,06
A520W	1,07
V531A	1,02
V531L	1,04
V531N	1,19
V531R	1,06
V531S	1,08
V531T	1,19
A535E	1,19
A535F	1,06
A535G	1,02
A535K	1,07
A535L	1,02
A535N	1,04
A535P	1,14
A535R	1,22
A535S	1,06
A535T	1,04
A535V	1,04
A535W	1,09
A535Y	1,13
V536C	1,09
V536E	1,09
V536I	1,04
V536L	1,07
V536M	1,20
V536Q	1,02
V536S	1,05
A539E	1,08
A539M	1,03
A539S	1,02
A539W	1,06
A539R	1,22

**Ejemplo 10: Variantes de glucoamilasa con AE y termoestabilidad aumentadas**

**[0203]** Las variantes en los ejemplos 8 y 9 se analizaron para detectar actividad específica aumentada y estabilidad térmica aumentada. La tabla 7 muestra las variantes con un índice de rendimiento (PI) > 1,0 en comparación con el PI de la TrGA precursora para ambas propiedades.

Tabla 7: Variantes con SBD con propiedades modificadas combinadas

Variante	PI Actividad Específica	PI Estabilidad Térmica
T495R	1,15	1,10
E503C	1,05	1,39
E503S	1,10	1,02
Q511H	1,05	1,33
A520C	1,38	1,10
A520L	1,46	1,06
A520P	1,50	1,06
V531L	1,04	1,21
A535K	1,07	1,24
A535N	1,04	1,37
A535P	1,14	1,59
A535R	1,22	1,26
V536I	1,04	1,02
A539E	1,08	1,32
A539M	1,03	1,05
A539R	1,22	1,36
A539S	1,02	1,30

Referencia **Ejemplo 11: Estructura de cristal de TrGA**

5 **[0204]** La estructura tridimensional completa de la glucoamilasa *Trichoderma reesei* (*Hypocrea jecorina*) (TrGA) se determinó con la resolución 1,9 Å. La tabla 8 muestra las coordenadas para la estructura de cristal de la glucoamilasa *Trichoderma*. TrGA se cristalizó de forma intacta con 599 residuos y todas las modificaciones traduccionales posteriores que normalmente tendrían lugar en el huésped natural. La estructura de cristal se produjo y se analizó de la siguiente manera:

10 **[0205]** Expresión proteínica y purificación - El gen codificante de GA *H. jecorina* se clonó y se expresó según los protocolos descritos en la publicación de solicitud de patente estadounidense nº: US 2006/0094080 A1, Dunn-Coleman *et al.* Mayo 4/2006).

15 **[0206]** El material de proteína TrGA usado para los experimentos de cristalización se purificó inicialmente en una etapa mediante cromatografía de intercambio de aniones como se muestra a continuación: se prepararon sobrenadantes de cultivo concentrados de TrGA expresada consistentes en 180 mg/ml de proteína total mediante la dilución de la muestra 1:10 en un tampón de 25 mM Tris-HCl pH 8,0. Se empleó una columna FF Sefarosa Q HIPREP 16/10 (GE Healthcare) para la purificación de intercambio aniónico. La columna HIPREP se equilibró con 4 volúmenes de columna (VC) de tampón de inicio (25 mM Tris-HCl, pH 8,0) seguido de una aplicación de 10 ml de la muestra proteínica diluida. Se aplicó un gradiente lineal 8 VC de 0 a 140 mM NaCl en el tampón de migración (25 mM Tris-HCl, pH 8,0) con el fin de eluir la proteína unida. Se eluyó la TrGA unida de la columna sefarosa Q HIPREP con una concentración de sal de aproximadamente 80 mM NaCl. Las fracciones con proteína TrGA pura se acumularon y concentraron a 50 mg/ml utilizando un tubo de concentración centrífuga VIVASPIN (Viva Science) con un límite de peso molecular (LPM) de 10 kD. Se cambió el tampón del material de TrGA purificado y concentrado utilizando una columna desaladora DG-10 (Bio-Rad) equilibrada con 50 mM tampón acetato de sodio, pH 4,3. Las concentraciones de proteína se determinaron mediante la medición de la absorbancia a 280 nm. El contenido proteico TrGA concentrado y purificado inicialmente se almacenó a -20 °C.

25 **[0207]** Se introdujeron dos etapas de purificación adicionales, una purificación mediante intercambio de aniones adicional y una purificación por exclusión de tamaño, con el fin de mejorar la tendencia del material proteínico TrGA a formar cristales. Estas dos etapas de purificación adicionales se llevaron a cabo de la siguiente forma: En la primera etapa de purificación por intercambio de aniones se utilizó una columna MONOQ 10 ml (GE Healthcare). Una muestra de 1 ml de material TrGA congelado y purificado inicialmente (50 mg proteína) se descongeló y el tampón se cambió a 20 mM Tris-HCL, pH 8,0, mediante dilución repetida de la muestra a 6 ml en el nuevo tampón, seguido de una concentración de la muestra de nuevo a 0,5 ml utilizando un tubo de

concentración de 6 ml 5 kD LPM. La muestra TrGA se diluyó tras la última etapa de concentración en agua destilada hasta que se alcanzó una conductividad de la muestra proteínica que correspondía a la conductividad del tampón inicial de la purificación de aniones, es decir, 25 mM Tris-HCl, pH 8,0. La columna MONOQ se equilibró primero con 4 volúmenes de columna (VC) de tampón de inicio, seguido de la aplicación de la muestra proteínica diluida a la columna. Se eluyó la proteína unida de la columna MONOQ mediante dos gradientes diferentes. En el primero, se aplicó un gradiente pH lineal 4 VC donde el pH del tampón de inicio se disminuyó de 8,0 a 6,0. En el segundo gradiente, se aplicó un gradiente de sal largo 8 CV en el que la concentración de sal se aumentó de 0 a 350 mM NaCl en el tampón de migración (25 mM Tris-HCl, pH 6,0). Se halló que la TrGA unida se eluía de la columna durante el segundo gradiente de sal con una concentración de NaCl aproximada de 150 mM. Las fracciones con TrGA se acumularon y concentraron a 2 ml utilizando un tubo de concentración VIVASPIN 6 ml 5 kD LPM. A continuación, la muestra TrGA concentrada se aplicó a una columna de exclusión Superdex 200 16/60 de tamaño (GE Helthcare) equilibrada con 4 VC de 20 Tris-cl, pH 8,0 y 50 mM NaCl, que también se utilizó como tampón de migración. Las fracciones del principal pico de elución tras la purificación por exclusión de tamaño se acumularon y concentraron a una concentración de proteína aproximada de 7,5 mg/ml utilizando un tubo de concentración VIVASPIN de 6 ml 5kD LPM.

**[0208]** Cristalización de proteína - La muestra proteínica que se utilizó para hallar las condiciones de cristalización de TrGA iniciales fue una muestra del material TrGA que se purificó una vez mediante purificación por intercambio de aniones y a partir de ahí se almacenó a -20 °C. La muestra proteínica TrGA se descongeló y se diluyó con 50 mM tampón acetato de sodio, pH 4,3 a aproximadamente 12 mg/ml antes de los experimentos de cristalización iniciales. Se utilizó el conjunto de datos de rayos X ortorrómbicos para resolver la estructura TrGA mediante reemplazo molecular (RM) y el conjunto de datos ortorrómbicos de alta resolución se utilizó para el modelo de estructura TrGA del grupo espacial ortorrómbico final. Se halló que los cristales TrGA ortorrómbicos se cultivaban en solución consistente en 25 % PEG 3350 0,20 acetato de amonio, 0,10 M Bis-Tris pH 5,5 (solución de reserva), utilizando el método de difusión por vapor con gotas colgantes (McPherson 1982), a 20 °C. Las gotas de cristalización se prepararon mediante la mezcla de iguales cantidades de solución proteínica (12 mg/ml) y la solución de reserva para un volumen final de 10 µl. Se halló que los cristales TrGA pertenecían al grupo espacial ortorrómbico P212121 con las dimensiones celulares aproximadas:  $a = 52,2 \text{ \AA}$ ,  $b = 99,2 \text{ \AA}$ ,  $c = 121,2 \text{ \AA}$ , y presentan una  $V_m$  calculada de 2,3 (Matthews 1968) con una molécula en la unidad asimétrica.

**[0209]** Recogida de datos de rayos X - Los dos conjuntos de datos TrGA ortorrómbicos se recogieron de cristales únicos fijados en tubos capilares sellados, a temperatura ambiente. El conjunto de datos de rayos X TrGA ortorrómbicos de baja resolución iniciales, utilizado para solucionar la estructura mediante métodos de reemplazo molecular (RM), se recogió en una fuente de rayos X doméstica, un detector de placa de imagen Raxis IV++ MSC/Rigaku (Molecular Structures Corp., The Woodlands, Texas) con espejos de enfoque con una radiación Cu K $\alpha$  de un generador de ánodos giratorio Rigaku RU200. Este conjunto de datos se procesó, se extrajo y se realizó la media utilizando el software d\*trek suministrado por MSC/Rigaku. El conjunto de datos monoclinicos centrados C se recogió de un único cristal TrGA congelado a 100 K equilibrado en un agente crioprotector compuesto de 25 % PEG 3350, 15% Glicerol 50 mM CaCl $_2$  y 0,1 M Bis-Tris pH 5,5 como crioprotector, fijado en bucles de fibra rayón y sumergido congelado en nitrógeno líquido antes de su transporte al sincrotrón. El conjunto de datos ortorrómbicos de alta resolución (1,9 Å) y el conjunto de datos monoclinicos centrados se recogieron ambos en una fuente de sincrotrón, línea de haz 911:5 en MAX LAB en Lund, Suecia. Ambos datos de conjuntos recogidos en una fuente de sincrotrón se procesaron con MOSFLM y se organizaron con el programa SCALA incluido en el paquete de programas CCP4 (Collaborative Computational Project Number 4 1994). El procesamiento de datos posterior se llevó a cabo utilizando el paquete de programas CCP4 (Collaborative Computational Project Number 4 1994), a menos que se indique lo contrario. Un conjunto del 5 % de las reflexiones de cada conjunto de datos se deja apartado y se utiliza para controlar el libre R (Bruger *et al.* (1992) 355:472-475).

**[0210]** Solución de estructura - La estructura TrGA se solucionó inicialmente mediante RM con el programa de reemplazo automático MOLREP (Collaborative Computational Project Number 4 1994), incluido en el paquete de programas CCP4, utilizando el conjunto de datos ortorrómbicos de resolución baja iniciales y utilizando las coordenadas de la variante *Aspergillus awamori* GA X 100 (pdb entrada 1 GLM (Aleshin *et al.* (1994) J Mol Biol 238: 575-591) como modelo de investigación. Se editó el modelo de investigación *A. awamori* GA para eliminar todas las fracciones de glicosilación fijadas a la molécula de proteína como glicosilaciones N- y O- y todas las moléculas de solventes antes de llevar a cabo los experimentos RM. Todas las reflexiones entre 36,8 y 2,8 Å de resolución del conjunto de datos de baja resolución iniciales se utilizaron para la solución RM. El programa RM halló una única solución de función de rotación, con un máximo de 11,1  $\sigma$  por encima de lo ya registrado, el siguiente valor más alto fue 3,8  $\sigma$  por encima de lo ya registrado. La solución de función de traducción proporcionó un factor R de 48,7 % y tuvo un factor de contraste de 17,4. La solución MR se refinó durante 10 ciclos de refinamiento de cuadrados mínimos limitado utilizando el programa Refmac 5,0 (Murshudov *et al.* (1997) Acta Crystallogr., D53: 240-255). Esto redujo el factor R cristalográfico a 31,1 % mientras que el valor libre R cayó desde 42,2 % a 41,1 %.

**[0211]** Refinamiento y encaje modelo - El modelo de solución RM refinado se utilizó para calcular un mapa de

densidad inicial a partir del conjunto de datos TrGA ortorrómbicos de baja resolución. La densidad de electrones para un puente disulfuro entre los residuos 19 y 26 de TrGA, un puente disulfuro no presente en la variante *A. awamori* X 100 modelo de estructura, puede identificarse fácilmente en este mapa de densidad de electrones. Esto se tomó como una indicación de que el mapa de densidad de electrones tenía la suficiente calidad para utilizarse para construir un modelo de estructura de TrGA a partir de su secuencia de aminoácidos. El modelo de estructura TrGA inicial, según el conjunto de datos de baja resolución, se refinó con ciclos alternos de construcción de modelo mediante Coot (Emsley, *et al.* Acta Crystallogr D Biol Crystallogr (2004) 60:2126-2132), y la máxima probabilidad de refinamiento utilizando Refmac 5,0.

[0212] La resolución del modelo de estructura TrGA inicial se extendió a la resolución del conjunto de datos ortorrómbicos de alta resolución (1,9 Å) mediante el refinamiento del modelo de estructura TrGA inicial contra el conjunto de datos de alta resolución durante 10 ciclos de refinamiento limitado utilizando el programa Refmac 5,0. La mayoría de las moléculas de agua en los modelos de estructura se ubicaron de forma automática mediante el uso de protocolos de recogida de agua en los programas de refinamiento y, a continuación, de forma manual se eligieron o descartaron mediante inspección visual. Todas las comparaciones estructurales se realizaron con Coot (Emsley *et al. supra*) u O (Jones *et al.* (1991) Acta Crystallogr. A47: 110-119), y las figuras se prepararon con PyMOL (DeLano *et al.* (2002) The PyMOL Molecular Graphics system, Palo Alto, CA EEUU: DeLano Scientific).

[0213] A partir de estos resultados, puede verse que el segmento principal catalítico TrGA seguía la misma topología de 6 barriles ( $\alpha/\alpha$ ) descrita por Aleshin *et al.* (*supra*) para AaGA, consistente en un doble barril de hélices alfa con el c-terminal de la hélice externa hacia el N-terminal de la hélice interna. Fue posible identificar diferencias clave en la densidad de electrones tal como el puente disulfuro entre los residuos 19 y 26 y en la inserción (residuos 257-260) en relación con AaGA. El segmento que comprende 80-100 también experimentó una reconstrucción de modelo extensiva. Se identificó un lugar de glicosilación principal en Asn 171 que tenía hasta cuatro fracciones de glucósidos fijadas. Se identificó un lugar de glicosilación similar en AaGa. De forma adicional, el núcleo catalítico con tres cis péptidos entre los residuos 22-23, 44-45 y 122-123 se conservó entre TrGA y AaGA. En general, existía una variación rms de 0,535 Å entre 409 de 453 átomos Ca cuando se compararon las coordenadas de los núcleos catalíticos de TrGA y AaGA.

#### Referencia **Ejemplo 12: Homología entre TrGA y *Aspergillus awamori* GA**

[0214] La estructura de cristal del TrGA identificada en el ejemplo 11 se superpuso sobre la estructura de cristal identificada previamente de la *Aspergillus awamori* GA (AaGA) La estructura de cristal AaGA se obtuvo a partir de la base de datos proteínica (BDP) y la forma de AaGA que se cristalizó fue la forma con solo un dominio catalítico. La estructura de la glucoamilasa *Trichoderma reesei* con las tres regiones intactas se determinó con la resolución Angstrom 1,8 en el presente documento (véase la tabla 8 y el ejemplo 11). Al utilizar las coordenadas (véase la tabla 8) la estructura se alineó con las coordenadas del dominio catalítico de la cepa *Aspergillus awamori* X100 que se determinó previamente (Aleshin, (1994) *J Mol Biol* 238:575-591 y el BDP). Como puede observarse en las figuras 6 y 7, la estructura del dominio catalítico se superpuso bastante cerca y permitió la identificación de los residuos equivalentes según esta superposición estructural.

#### *Dominio catalítico de TrGA*

[0215] Según el presente análisis, se identificaron los lugares que podían mutarse en el dominio catalítico de TrGA y tener como resultado una estabilidad y/o actividad específica aumentada. Estos lugares incluyen 108, 124, 175 y 316 en el lugar activo. También se identificaron variantes por pares específicos Y47W/Y315F e Y47F/Y315W. Se identificaron otros lugares 143, D44, P45, D46, R122, R125, V181, E242, Y310, D313, V314, N317, R408 y N409. Debido a la alta homología estructural se espera que las variantes beneficiosas halladas en los lugares en *Trichoderma reesei* GA tengan consecuencias similares en *Aspergillus awamori* y otras glucoamilasas homólogas.

#### **Región de enlace y dominio de unión a almidón de TrGA**

[0216] El enlazador TrGA, residuos 454-490, se define como el segmento que abarca la región entre dos puentes disulfuros, uno entre los residuos 222 y 453 y otro entre los residuos 491 y 587. Nueve de los residuos en el enlazador son prolina. A partir de la estructura de cristal, el enlazador se extiende desde la parte de atrás de la molécula en un amplio arco seguido de un giro abrupto tras el residuo 477 lisina en la superficie cerca de la superficie de unión al sustrato. El enlazador se extiende como una bobina aleatoria que se fija mediante interacciones de las cadenas laterales de Tyr 452, Pro 465, Phe 470, Gln 474, Pro 475, Lys 477, Val 480 y Tyr 786 a regiones en la superficie del dominio catalítico.

[0217] El dominio de unión a almidón está compuesto de un sándwich beta de dos hojas beta torcidas, atadas a un extremo mediante un puente disulfuro entre Cys 491 y Cys 587 y al otro extremo, con una serie de bucles que comprenden un lugar de unión para almidón conectado mediante bucles alargados. La estructura del SBD de

TrGA es bastante similar a la estructura media del SBD de AnGA determinada mediante NMR (Sorimachi, K, *et al. Structure* (1997) 5(5): p. 647-661) y el SBD de beta amilasa de *Bacillus cereus* (Mikami, B., *et al. Biochemistry* (1999) 38(22): p. 7050-61). La figura 8 muestra una alineación de las estructuras de cristal de TrGA y *A. niger* con SBD. Cuando se alinea con uno o ambos de estos SBD, destaca un bucle como bastante variable en relación con los residuos 537-543 (en *A. niger* el bucle es 554-560 y en *B. cereus* el bucle es 462-465). En la estructura NMR de beta ciclodextrina, un análogo de almidón complejo con el SBD de *A. niger* GA (Sorimachi, *et al* 1997 *supra*), el bucle cambia de forma sustancial tras la unión a ciclodextrina. Por lo tanto, el bucle se designa como bucle flexible. Este bucle flexible forma parte del "lugar de unión 2" (véase la figura 8 para este lugar de unión en TrGA). Se identificó también un segundo lugar de unión en AnGA (lugar de unión 1), un lugar primario que comparte similitudes con otras proteínas que unen carbohidratos. En general, la conservación de residuos e incluso configuraciones laterales en el lugar de unión 1 de estos SBD es muy alta. Las figuras demuestran las interacciones en estos lugares de unión entre el SBD y el dominio catalítico que sirven para unirse al almidón.

**[0218]** Si se toma de forma conjunta, parece existir un patrón común para las interacciones entre el enlazador y el SBD con el dominio catalítico (DC). La interacción se presenta en forma de cadena lateral de anclaje que interactúa con el área de superficie del dominio cercano. En general, el residuo de anclaje se halla en el segmento enlazador. En el caso de interacciones entre el DC y SBD, los residuos de anclaje pueden aportarse desde cualquier dominio como en el caso de los residuos Ile 43 y Phe 29 que vienen del DC o el residuo 592, que viene del SBD.

#### Referencia Ejemplo 13: **Modelo de la unión de acarbosa a TrGA**

**[0219]** Se determinó la estructura de cristal de la TrGA complejada con el inhibidor acarbosa. Los cristales del complejo se obtuvieron mediante la impregnación de cristales de TrGA naturales precultivados en acarbosa. Tras la impregnación durante 3 días de los cristales, se fijaron en un tubo capilar de vidrio sellado y se recogió mediante difracción de rayos X con un detector de placa de imagen Rigaku Raxis IV++ con una resolución de 2,0 Å. Se ajustaron las coordenadas para un mapa de densidad de electrones de diferencia. El modelo se refinó a un factor R de 0,154 con un libre R de 0,201 para un total de 41276 reflexiones que representan todos los datos recogidos entre 27 y 2,0 Å de resolución. El modelo de la estructura refinada resultante se muestra en la figura 9.

**[0220]** Según el conocimiento de que la presencia del SBD tiene un impacto en la hidrólisis de almidón insoluble, se entiende que debería existir una interacción del SBD con moléculas de almidón mayores. Por lo tanto, se comparó la estructura de TrGA con estructuras conocidas de un DC unido a acarbosa de AaGA y un SBD de *A. niger* complejado con beta-ciclodextrina. Esto mostró que la beta-ciclodextrina unida al lugar de unión 2 estaba cerca de la ubicación del sustrato como indicaba la ubicación de acarbosa unida al DC de *A. awamori*. Por lo tanto, las coordenadas de acarbosa del modelo de estructura de la AaGA (pdb entry 1GAI, Aleshin, *et al. 1994 supra*) se alinearon en el centro activo de TrGA. Además, se alineó la estructura del SBD de AnGA unida a ciclodextrina (pdb entry 1AC0: Sorimachi, *et al* 1997 *supra*). A partir de esto, se realizó un modelo para acarbosa que se une a TrGA (véase la figura 9). El modelo mostró que el SBD localizaría el DC de TrGA cerca del almidón afectado, y también impediría que la enzima se dispersara lejos del sustrato mientras se liberaba el producto del centro activo tras la hidrólisis. El SBD de TrGA se uniría al almidón por el lugar 1 y favorecería la localización donde un fragmento afectado podría unirse al lugar 2 dentro de un extremo suelto que señala al lugar catalítico (el lado activo para el dominio catalítico). Este modelo muestra cómo la función propuesta de la enzima se aporta mediante la estructura del SBD y enlazador. Las cadenas laterales de aminoácidos implicadas en la interacción específica entre el DC, el enlazador y el SBD son específicas para *Trichoderma reesei* GA, sin embargo, en otras glucoamilasas, cambios de secuencia complementarios permitirían interacciones generales similares y la juxtaposición de dominios.

**[0221]** Según el presente modelo, se identificaron los lugares en los que podrían realizarse sustituciones en el SBD de TrGA y tener como resultado una estabilidad y/o actividad específica aumentada. Por lo tanto, dos bucles que forman parte del lugar de unión 1 son candidatos probables de modificaciones con el fin de aumentar o disminuir la unión a la molécula de almidón más grande. Estos son el bucle 1 (aa 560-570) y el bucle 2 (aa 523-527). Puesto que los dos residuos Trp (triptófano) en los aminoácidos 525 y 572 están probablemente implicados de forma directa en la unión a almidón, no serían tan propicios al cambio. Sin embargo, los residuos subyacentes, incluyendo 516-518 serían propicios, al igual que los residuos subyacentes 558-562. El bucle de los residuos 570-578 es también un buen candidato para modificaciones. Los residuos 534-541 son parte del lugar de unión 2 que interactúa con el lugar catalítico en el DC. Por lo tanto, estos son buenos candidatos para modificaciones que pueden aumentar o disminuir la actividad específica.

**[0222]** Debido a la alta homología estructural del SBD de TrGA, se espera que las variantes beneficiosas halladas en los lugares en *Trichoderma reesei* GA tengan consecuencias similares en *Aspergillus awamori* y otras glucoamilasas homólogas. Por lo tanto, la estructura del SBD de TrGA proporciona una base para el diseño de esta y enzimas relacionadas con propiedades modificadas en comparación con una glucoamilasa precursora. Estas propiedades modificadas pueden ser beneficiosas para procesos de generación de combustibles según materias primas de almidón.

**REIVINDICACIONES**

1. Una variante de glucoamilasa aislada que comprende un dominio catalítico y un dominio de unión a almidón (SBD por sus siglas en inglés),

5

(a) comprendiendo dicho dominio catalítico al menos un 85 % de identidad de secuencia con la secuencia de aminoácidos

SVDDFISTETPIALNNLLCNVGPDGCRAFGTSAGAVIASPSTIDPDYYYMWTR  
 DSALVFKNLIDRFTETDYDAGLQRRIEQYITAQVTLQGLSNPSGSLADGSGLGE  
 PKFELTLKPFTGNWGRPQRDGPALRAIALIGYSKWLINNNYQSTVSNVIWPIV  
 RNDLNYVAQYWNQTGFDLWEEVNGSSFFTANQHRALVEGATLAATLGQS  
 GSA YSSVAPQVLCFLQRFWVSSGGYVDSNINTNEGRTGKDVNSVLTSIHFD  
 NLGCDAGTFQPCSDKALS NLKVVVDSFRSIYGVNKGIPAGAAVAIGRYAEDV  
 YYNGNPWYLATFAAAEQLYDAIYVWKKTG SITVTATSLAFFQELVPGVTAG  
 TYSSSSSTFTNIINAVSTYADGFLSEAAKYVPADGSLAEQFDRNSGTPLSALHL  
 TWSYASFLTATARRAGIVPPSWANSSASTIPSTC (SEQ ID NO:3),

y

(b) comprendiendo dicho SBD la sustitución de aminoácidos N563I, N563C, N563E, N563A, N563K, N563L, N563Q, N563T o N563V en la secuencia de aminoácidos:

SVDDFISTETPIALNNLLCNVGPDGCRAFGTSAGAVIASPSTIDPDYYYMWTR  
 DSALVFKNLIDRFTETDYDAGLQRRIEQYITAQVTLQGLSNPSGSLADGSGLGE  
 PKFELTLKPFTGNWGRPQRDGPALRAIALIGYSKWLINNNYQSTVSNVIWPIV  
 RNDLNYVAQYWNQTGFDLWEEVNGSSFFTANQHRALVEGATLAATLGQS  
 GSAYSSVAPQVLCFLQRFWVSSGGYVDSNINTNEGRTGKDVNSVLTSIHFD  
 NLGCDAGTFQPCSDKALS NLKVVVDSFRSIYGVNKGIPAGAAVAIGRYAEDV  
 YYNGNPWYLATFAAAEQLYDAIYVWKKTG SITVTATSLAFFQELVPGVTAG  
 TYSSSSSTFTNIINAVSTYADGFLSEAAKYVPADGSLAEQFDRNSGTPLSALHL  
 TWSYASFLTATARRAGIVPPSWANSSASTIPSTCSGASVVGYSRPTATSFPPS  
 QTPKPGVPSGTPYTPLPCATPTSVAVTFHELVSTQFGQTVKVAGNAAALGNW  
 STSAAVALDAVNYADNHPLWIGTVNLEAGDVVEYKYINVGQDGSVTWESDP  
 NHTYTVPAVACVTQVVKEDTWQS (SEQ ID NO: 2);

10

o en una posición equivalente en un SBD de glucoamilasa precursora como se determina mediante la alineación de secuencia;

presentando dicha variante de glucoamilasa una termoestabilidad aumentada comparada con dicha glucoamilasa precursora, en la que la termoestabilidad se mide como el % de actividad residual tras la incubación durante 1 hora a 64 grados centígrados en un tampón de NaAc pH 4,5.

15

2. Variante de glucoamilasa aislada de acuerdo con la reivindicación 1, en la que el dominio catalítico presenta al menos un 90 % o al menos un 95 % de identidad de secuencia con la secuencia de aminoácidos de SEQ ID NO:3.



3. Variante de glucoamilasa de acuerdo con la reivindicación 1, en la que la glucoamilasa precursora comprende un SBD que presenta al menos un 95 % de identidad de secuencia de aminoácidos con un SBD elegido de entre:

CATPTSVAVTFHELVTQFGQTVKVAGNAAALGNWSTSAVALDAV  
 NYADNHPLWIGTVNLEAGDVVEYKYINVVGQDGSVTWESDPNHTYTVPAVA  
 CVTQVVKEDTWQS (SEQ ID NO: 11);

CADASEVYVTFNERVSTAWGETIKVVGPNPALGNWDTSKAVTLSAS  
 GYKSNLPLWSITVPIKATGSAVQYKYIKVGTNGKITWESDPNRSITLQTASSA  
 GKCAAQTVNDSWR (SEQ ID NO:385);

CTPPSEVTLTFNALVDTAFGQNIYLVGSIPELGSWDPANALLMSAKSW  
 TSGNPVWTLISLPAAGTSFEYKFKRKDDGSSDVVWESDPNRSYNVPKDCGANT  
 ATVNSWWR (SEQ ID NO:386);

CTTPTSVAVTFDEIVSTSYGETIYLAGSIPELGNWSTASAIPLRADAYTN  
 SNPLWYVTVNLPPGTSFEYKFFKNQTDGTIVWEDDPNRSYTVPAYCGQTTAI  
 LDDSWQ (SEQ ID NO:387);

CTTPTAVAVTFDLTATTTYGENIYLVGSISQLGDWETS DGIALSADKYT  
 SSDPLWYVTVTL PAGESFEYKFIRESDDSV EWESDPNREYTV PQACGTSTAT  
 VTDTWR (SEQ ID NO:388);

CTTPTAVAVTFDLTATTTYGENIYLVGSISQLGDWDTSDGIALSADKY  
 TSSNPLWYVTVTL PAGESFEYKFIRESDDSV EWESDPNREYTV PQACGESTA  
 TVTDTWR (SEQ ID NO:389).

4. Variante de glucoamilasa de acuerdo con la reivindicación 1, que comprende además una o más sustituciones en una posición correspondiente a la posición 493, 494, 495, 501, 502, 503, 508, 511, 517, 518, 519, 520, 525, 527, 531, 533, 535, 536, 537, 538, 539, 540, 545, 546, 547, 549, 551, 561, 567, 569, 577, 579 o 583 de SEQ ID NO:2 o en una posición equivalente en dicha glucoamilasa precursora como se determina mediante la alineación de secuencia.
5. Variante de glucoamilasa aislada de acuerdo con la reivindicación 1, que comprende además una o más sustituciones de aminoácidos en una posición correspondiente a la posición de residuo: 10, 14, 15, 23, 42, 45, 46, 59, 60, 61, 67, 68, 72,73,97,98,99, 102, 108, 110, 113, 114, 122, 124, 125, 133, 140, 144, 145, 147, 152, 153, 164, 175, 182, 204, 205, 214, 216, 219, 228, 229, 230, 231, 236, 239, 240, 241, 242, 244, 263, 264, 265, 268, 269, 276, 284, 291, 300, 301, 303, 310, 311, 313, 316, 338, 342, 344, 346, 349, 359, 361, 364, 379, 382, 390, 391, 393, 394, 408, 410, 415, 417, 418, 430, 431, 433, 436, 442, 443, 444, 448 y 451 de SEQ ID NO:2 o SEQ ID NO:3.
6. Variante de glucoamilasa de acuerdo con la reivindicación 1, en la que la glucoamilasa precursora se obtiene a partir de una *Trichoderma spp.*
7. Un polinucleótido que codifica la variante de acuerdo con la reivindicación 1.
8. Una célula huésped que comprende un polinucleótido de acuerdo con la reivindicación 7.
9. Una composición de enzima que comprende la variante de glucoamilasa según la reivindicación 1, en la que la composición es adecuada para utilizarse en un proceso de conversión de almidón, una formulación

de pienso animal o un proceso de fermentación alcohólica.

- 5
10. Una composición de enzima adecuada para utilizarse en un proceso de conversión de almidón de acuerdo con la reivindicación 9 que comprende además un alfa-amilasa.
  11. Un método para producir una variante de glucoamilasa en una célula huésped que comprende la transformación de una célula huésped con una construcción de ADN que comprende un polinucleótido de acuerdo con la reivindicación 7; el cultivo de la célula huésped en las condiciones adecuadas para la expresión y producción de dicha variante de glucoamilasa y la producción de dicha variante.
  12. Método de acuerdo con la reivindicación 11, que comprende además la recuperación de la variante de glucoamilasa de dicho cultivo.

## Dibujos

FIG. 1A

(A) Proteína precursora TrGA (632 aminoácidos) (SEQ ID NO:1)

```

1  MHVLSTAVLL  GSVAVQKVLG  RPGSSGLSDV  TKRSVDDFIS  TETPIALNNL
51  LCNVGPDGCR  AFGTSAGAVI  ASPSTIDPDY  YYMWTRDSAL  VFKNLIDRFT
101 ETYDAGLQRR  IEQYITAQVT  LQGLSNPSGS  LADGSGLGEP  KFELTLKPFT
151 GNWGRPQRDG  PALRAIALIG  YSKWLNINNY  QSTVSNVIWP  IVRNDLNYVA
201 QYWNQTGFDL  WEEVNGSSFF  TVANQHRALV  EGATLAATLG  QSGSAYSSVA
251 PQVLCFLQRF  WVSSGGYVDS  NINTNEGRTG  KDVNSVLTSI  HTFDPNLGCD
301 AGTFQPCSDK  ALSNLKVVVD  SFRSIYGVNK  GIPAGAAVAI  GRYAEDVYYN
351 GNPWYLATFA  AAEQLYDAIY  VWKRTGSITY  TATSLAFFQE  LVPGVTAGTY
401 SSSSSTFTNI  INAVSTYADG  FLSEAAKYVP  ADGSLAEQFD  RNSGTPLSAL
451 HLTWSYASFL  TATARRAGIV  PPSWANSSAS  TIPSTCSGAS  VVGSYSRPTA
501 TSFPPSQTPK  PGVPSGTPYT  PLPCATPTSV  AVTFHELVST  QFGQTVKVGAG
551 NAAALGNWST  SAAVALDAVN  YADNHPLWIG  TVNLEAGDVV  EYKYINVGQD
601 GSVTWESDPN  HTYTVPAVAC  VTQVVKEDTW  QS

```

(B) Secuencia codificante ADN de TrGA (1899 pb) (SEQ ID NO:4)

```

1  ATGCACGTCC  TGTCGACTGC  GGTGCTGCTC  GGCTCCGTTG  CCGTTCAAAA
51  GGTCCTGGGA  AGACCAGGAT  CAAGCGGTCT  GTCCGACGTC  ACCAAGAGGT
101 CTGTTGCAATG  CTTTCATCAGC  ACCGAGACGC  CTATTGCACT  GAACAATCTT
151 CTTTGC AATG  TTGGTCCTGA  TGGATGCCGT  GCATTCGGCA  CATCAGCTGG
201 TGCGGTGATT  GCATCTCCCA  GCACAATTGA  CCCGGACTAC  TATTACATGT
251 GGACGCGAGA  TAGCGCTCTT  GTCTTCAAGA  ACCTCATCGA  CCGCTTCAAC
301 GAAACGTACG  ATGCGGGCCT  GCAGCGCCGC  ATCGAGCAGT  ACATTACTGC
351 CCAGGTCACT  CTCCAGGGCC  TCTCTAACC  CTCGGGCTCC  CTCGCGGACG
401 GCTCTGGTCT  CGGCGAGCCC  AAGTTTGAGT  TGACCCTGAA  GCCTTTCACC
451 GGCAACTGGG  GTCGACCGCA  GCGGGATGGC  CCAGCTCTGC  GAGCCATTGC
501 CTTGATTGGA  TACTCAAAGT  GGCTCATCAA  CAACAATAT  CAGTCGACTG
551 TGTCCAACGT  CATCTGGCCT  ATTGTGCGCA  ACGACCTCAA  CTATGTTGCC
601 CAGTACTGGA  ACCAAACCGG  CTTTGACCTC  TGGGAAGAAG  TCAATGGGAG
651 CTCATTCTTT  ACTGTTGCCA  ACCAGCACCG  AGCACTTGTC  GAGGGCGCCA
701 CTCTTGCTGC  CACTCTTGGC  CAGTCGGGAA  GCGCTTATTC  ATCTGTTGCT
751 CCCAGGTTT  TGTGCTTTCT  CCAACGATTC  TGGGTGTCGT  CTGGTGGATA
801 CGTCGACTCC  AACATCAACA  CCAACGAGGG  CAGGACTGGC  AAGGATGTCA
851 ACTCCGTCC  GACTTCCATC  CACACCTTCG  ATCCCAACCT  TGGCTGTGAC
901 GCAGGCACCT  TCCAGCCATG  CAGTGACAAA  GCGCTCTCCA  ACCTCAAGGT
951 TGTGTCGAC  TCCTTCCGCT  CCATCTACGG  CGTGAACAAG  GGCATTCCTG
1001 CCGGTGCTGC  CGTCGCCATT  GGCCGGTATG  CAGAGGATGT  GTACTACAAC
1051 GGCAACCCTT  GGTATCTTGC  TACATTTGCT  GCTGCCGAGC  AGCTGTACGA
1101 TGCCATCTAC  GTCTGGAAGA  AGACGGGCTC  CATCACGGTG  ACCGCCACCT
1151 CCCTGGCCTT  CTTCCAGGAG  CTTGTTCTCG  GCGTGACGGC  CGGGACCTAC
1201 TCCAGCAGCT  CTTCCAGCTT  TACCAACATC  ATCAACGCCG  TCTCGACATA
1251 CGCCGATGGC  TTCCTCAGCG  AGGCTGCCAA  GTACGTCCCC  GCCGACGGTT
1301 CGCTGGCCGA  GCAGTTTGAC  CGCAACAGCG  GCACTCCGCT  GTCTGCGCTT
1351 CACCTGACGT  GGTCGTACGC  CTCGTTCTTG  ACAGCCACGG  CCCGTCGGGC
1401 TGGCATCGTG  CCCCCCTCGT  GGGCCAACAG  CAGCGCTAGC  ACGATCCCCT
1451 CGACGTGCTC  CGGCGCGTCC  GTGGTCCGAT  CCTACTCGCG  TCCCACCGCC

```

FIG. 1B

```

1501 ACGTCATTCC CTCCGTCGCA GACGCCCAAG CCTGGCGTGC CTTCCGGTAC
1551 TCCCTACACG CCCCTGCCCT GCGCGACCCC AACCTCCGTG GCCGTCACTT
1601 TCCACGAGCT CGTGTCGACA CAGTTTGGCC AGACGGTCAA GGTGGCGGGC
1651 AACGCCGCGG CCCTGGGCAA CTGGAGCACG AGCGCCGCCG TGGCTCTGGA
1701 CGCCGTCAAC TATGCCGATA ACCACCCCCT GTGGATTGGG ACGGTCAACC
1751 TCGAGGCTGG AGACGTCGTG GAGTACAAGT ACATCAATGT GGGCCAAGAT
1801 GGCTCCGTGA CCTGGGAGAG TGATCCCAAC CACTTACA CGGTTCTGTC
1851 GGTGGCTTGT GTGACGCAGG TTGTCAAGGA GGACACCTGG CAGTCGTAA
    
```

FIG. 1C

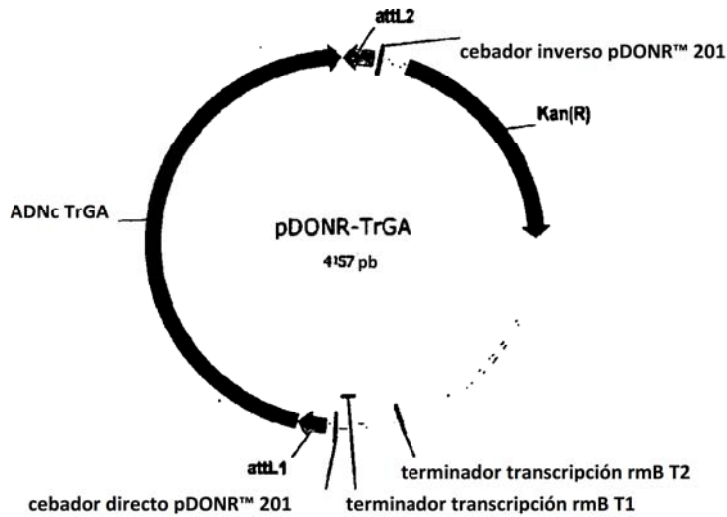
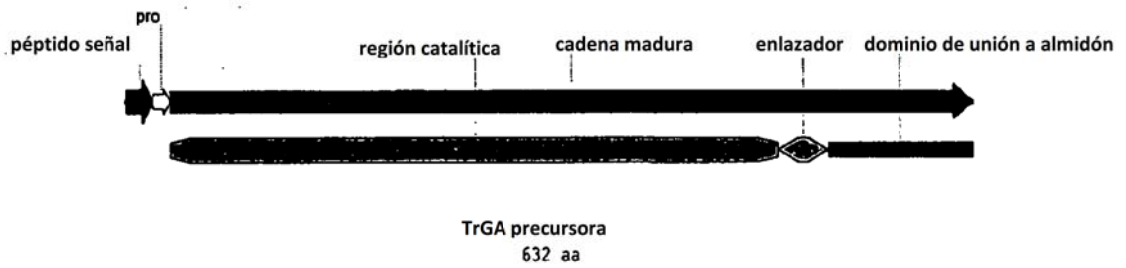


Figura 2

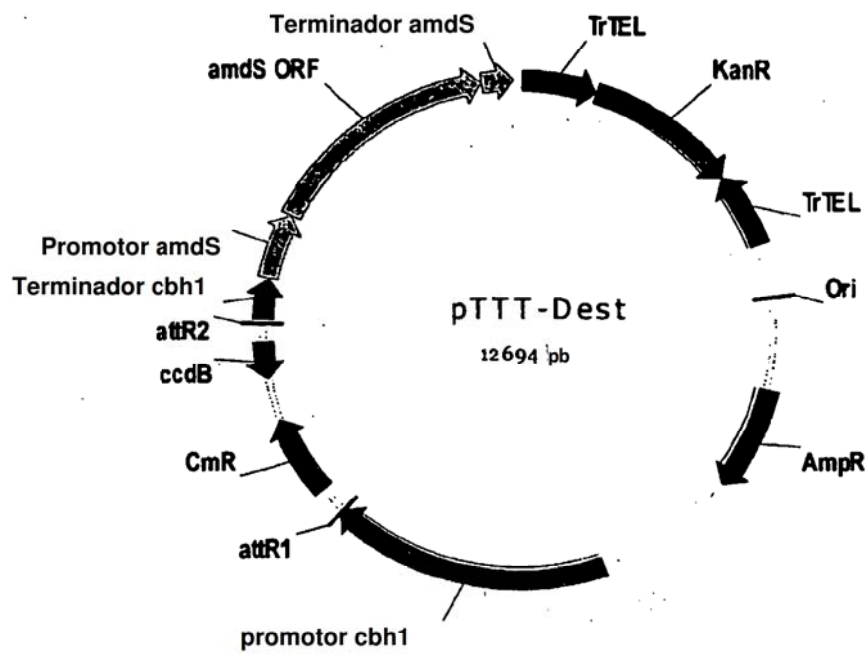


Figura 3

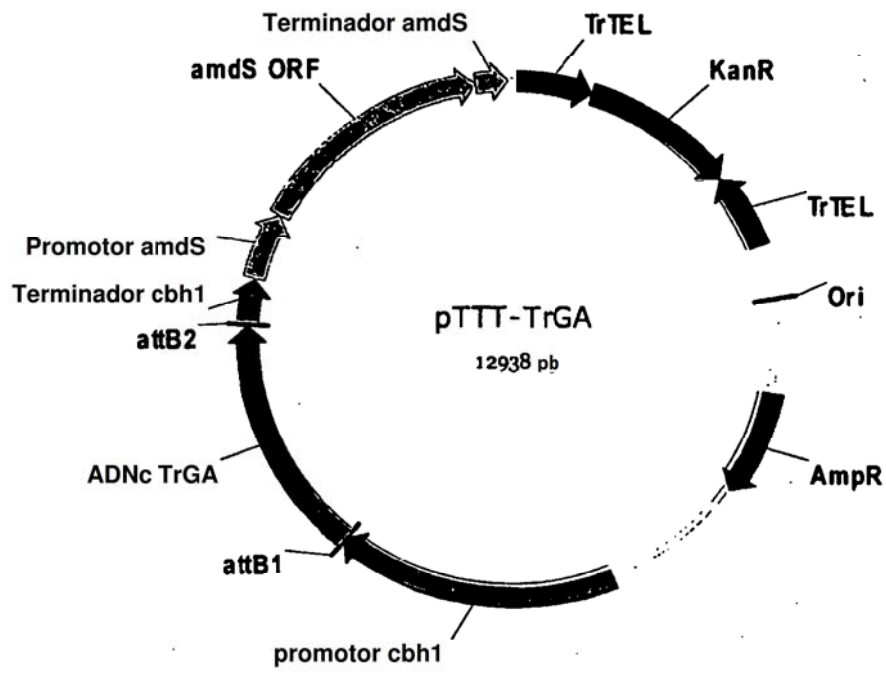


Figura 4

FIG. 5A (página 1 de 2)

AaGA	(1)	-ATLDSWLSNEATVARTAILNNGADGAWVSGADSGIVVASPSTDNDPDYF
AnGA	(1)	-ATLDSWLSNEATVARTAILNNGADGAWVSGADSGIVVASPSTDNDPDYF
AoGA	(1)	QSDLNAFIEAQTPIAKQGYLNNIGADGKLVGAAAGIVYASPSKSNPDYF
HgGA	(1)	-AAVDTFINTEKPIAWNKLLANIGPNGKAAPGAAAGVVIASPSRTDPPYF
HvGA	(1)	--SVDDFINTQTPIALNNLLCNVGPDGCRAFGTSAAGAVIASPSTTDPDYF
TrGA	(1)	--SVDDFISTETPIALNNLLCNVGPDGCRAFGTSAAGAVIASPSTIDPDYF
		* * * * *
AaGA	(50)	YTWRDSSLVLIKTLVDFRNGDTE-LLSTIENYISSQAIVQGISNPSGDL
AnGA	(50)	YTWRDSSLVLIKTLVDFRNGDTE-LLSTIENYISSQAIVQGISNPSGDL
AoGA	(51)	YTWRDAGLTMEEYIEQFIGGDAT-LESTIQNYVDSQANEQAVSNPSGGL
HgGA	(50)	FTWTPDAALVLTGIIESLGHNYNT-----TLQQVSNPSGTF
HvGA	(49)	YMWTRDSALVFNKIVDRFTQYDAGLQRRIEQYISAQVTLQGISNPSGSL
TrGA	(49)	YMWTRDSALVFNKILDRFTETDAGLQRRIEQYITAQVTLQGLSNPSGSL
		* * * * *
AaGA	(99)	SSGG-LGEPKFNVDDETAYTGSWGRPQRDGPALRATAMIGFRQWLLDNGYT
AnGA	(99)	SSGAGLGEKPFNVDDETAYTGSWGRPQRDGPALRATAMIGFGQWLLDNGYT
AoGA	(100)	SDGSLAEKPFYINISQFTDQSWGRPQRDGPALRASALIAYGNLSISDKQ
HgGA	(86)	ADGSLGAEKFNVDLTAFTGEWGRPQRDGPPLRAIALIQYAKWLIANGYK
HvGA	(99)	SDGSLGEPKFEFELTSLQFTGNWGRPQRDGPALRAIALIGYSKWLINNNYQ
TrGA	(99)	ADGSLGEPKFEFELTLKPFTEGNWGRPQRDGPALRAIALIGYSKWLINNNYQ
		* * * * *
AaGA	(148)	SAATEIVWPLVRNDLSYVAQYWNQTYDLWEEVNGSSFFTIQVHRALVE
AnGA	(149)	STATDIVWPLVRNDLSYVAQYWNQTYDLWEEVNGSSFFTIQVHRALVE
AoGA	(150)	SVVKANIWPIYQNDLSYVQYWNQTYGFDLWEEVQGSFFTVAVQHKALVE
HgGA	(136)	STAKSVVWPVVKNDLAYTAQYWNQTYGFDLWEEVPGSSFFTIASSHRALTE
HvGA	(149)	STVSNIIWPIVRNDLNYVAQYWNQTYGFDLWEEVNGSSFFTVANQHRALVE
TrGA	(149)	STVSNVIWPIVRNDLNYVAQYWNQTYGFDLWEEVNGSSFFTVANQHRALVE
		* * * * *
AaGA	(198)	GSAFATAVGSSCSWCDSQAPQILCYLQSFWTG--EYILANFDSS--RSGK
AnGA	(199)	GSAFATAVGSSCSWCDSQAPEILCYLQSFWTG--SFILANFDSS--RSGK
AoGA	(200)	GDFAKALGEECQACS-VAPQILCHLQDFWNG--SAVLSNLPTNG-RSGL
HgGA	(186)	GAYLAAQLDTECPPTTVAQVLCFQAFWNSKGNVYVSTSTAGEYRSGK
HvGA	(199)	GATLAATLGQSGSTYSSVAPQILCFQRFWVS--GGYIDSNINTNEGRTGK
TrGA	(199)	GATLAATLGQSGSAYSSVAPQVLCFQRFWVSSGGYVDSNINTNEGRTGK
		* * * * *
AaGA	(244)	DTNTLLGSIHTFDPEAGCDDSTFQPCSPRALANHKEVVDSFRSIYTLNDG
AnGA	(245)	DANTLLGSIHTFDPEAACDDSTFQPCSPRALANHKEVVDSFRSIYTLNDG
AoGA	(246)	DTNSLLGSIHTFDPAACDDTTFQPCSSRALSNHKLVDVDSFRSVYGINNG
HgGA	(236)	DANSILASIHNFDPAGCDNLTFQPCSERALANHKAYVDSFRNLYAINKG
HvGA	(248)	DANLLASIHNFDPPLGCDASTFQPCSDKALSNLKVVVDSFRSIYGVNKG
TrGA	(249)	DVNSVLTSIHTFDPNLGCDAFTFQPCSDKALSNLKVVVDSFRSIYGVNKG
		* * * * *
AaGA	(294)	LSDSEAVAVGRYPKDSYYNGNPWFLECTLAAAQLYDALYQWDKQGSLEIT
AnGA	(295)	LSDSEAVAVGRYPEDTYNGNPWFLECTLAAAQLYDALYQWDKQGSLEVT
AoGA	(296)	RGAGKAAAVGPYAEDTYQGGNPWYLTTLVAAELLYDALYQWDKQGVNVT
HgGA	(286)	IAQKAAVAVGRYSEDVYNGNPWYLANFAAAQLYDAIYVWVKQGSITVT
HvGA	(298)	IPAGSAVAIGRYPEDVYFNGNPWYLATFAAAQLYDSVYVWVKKTSITVT
TrGA	(299)	IPAGAAVAIGRYAEDVYNGNPWYLATFAAAQLYDAIYVWVKKTSITVT
		* * * * *

FIG. 5B (página 2 de 2)

AaGA (344) DVSLDFFQALYSDAATGTYSKSSSSSTYSIVDAVKTFADGFVSIVETHAAS  
 AnGA (345) DVSLDFFKALYSDAATGTYSKSSSSSTYSIVDAVKTFADGFVSIVETHAAS  
 AoGA (346) ETSLPFFKDLSSNVTTGSYAKSSSAYESLTSVAVKTYADGFISVVQEYTPD  
 HgGA (336) SVSLPFFRDLVSSVSTGTYSKSSSTFTNIVNAVKAYADGFIEVAAKYTPS  
 HvGA (348) STSSAFFQELVPGVAAGTYSSSQSTFTSIINAISTYADGFLSEAAKYVPA  
 TrGA (349) ATSLAFFQELVPGVTAGTYSSSSSTFTNLIINAVSTYADGFLSEAAKYVPA  
 \* \*\* \* \* \* \* \* \*

AaGA (394) NGSLSEQYDKSDGDELSARDLTWSYAALLTANRRNSVMPPSWGETSAS-  
 AnGA (395) NGSMSEQYDKSDGEQLSARDLTWSYAALLTANRRNSVVPASWGETSAS-  
 AoGA (396) GGALAEQYSRDQGT PVSASDLTWSYAFLSAVGRRNGTVPASWGSSTAN-  
 HgGA (386) NGALAEQYDRNTGKPD SAADLTWSYSAFLSAIDRRAGLVPPSWRASVAKS  
 HvGA (398) DGSLAEQFDRNTGTPLSAVHLTWSYASF LTAAARRAGVVPPSWASSGAN-  
 TrGA (399) DGSLAEQFDRNSGTPLSALHLTWSYASFLTATARRAGIVPPSWANSSAS-  
 \* \*\* \* \* \* \* \* \* \*

AaGA (443) SVPGTC  
 AnGA (444) SVPGTC  
 AoGA (445) AVPSQC  
 HgGA (436) QLPSTC  
 HvGA (447) TVPSSC  
 TrGA (448) TIPSTC  
 \* \*

FIG. 5C

gsldsflatetpialqgvlnnigpnagadvagasagivvaspsrsdpdyfyswtrdaaltakylvdafia  
 gnkdleqtiqeyisaqaqvqtisnpsgdltstgggkgepkfnvnetaftgppwgrpqrdgpalrataliaya  
 nylidngqastadeiiwpiqvndlsyvtqywnsstfdlweevegssffttavqhralvegnaatrlnh  
 tcpncvsqapqvlcflqsywtgsyvlanfgsggrsgkdvnslgsihtfdpaggcddstfqpcsarala  
 nhkvvtdsfrsvyavnsgiaegsavavgrypedvygggnpwyLataaaqlydaiyqwnkigsisid  
 vsLAffqdiypsaavgtyngssthndiisavqtyadgylsiiekytpsdgslteqfirsrdgtplsag  
 ltwsyaslltaarrqsivpaswgessassvpavcsatsatgppystatntawpssgsgpstttsvpctt  
 ptsvavtfdeivsttygetiylagsipelgnwspssaiplradaytssnplwyvtlnlpagtsfeykff  
 kktdgtiweddpnrstypaycggttailddswq



Trichoderma reesei	posición	491	492	483	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	
	Aminoácido	C	A	T	P	T	S	V	A	V	T	F	H	E	L	V	S	T	Q	F	G	Q	T	V	K	V	A	G	N	A	A	
Humicola grisea	posición	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	
	Aminoácido	C	A	D	A	S	E	V	Y	V	T	F	N	E	R	V	S	T	A	W	G	E	T	I	K	V	V	G	N	V	P	
Thielavia terrestris	posición	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	
	Aminoácido	C	S	T	P	T	A	V	A	V	T	F	N	E	R	V	T	T	Q	W	G	Q	T	I	K	V	V	G	D	A	A	
Thermomyces lanuginosus	posición	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	
	Aminoácido	C	T	P	P	S	E	V	T	L	T	F	N	A	L	V	D	T	A	F	G	Q	N	I	Y	L	V	G	S	I	P	
Talaromyces emersonii	posición	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	
	Aminoácido	C	T	T	P	T	S	V	A	V	T	F	D	E	I	V	S	T	S	Y	G	E	T	I	Y	L	A	G	S	I	P	
Aspergillus niger	posición	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	
	Aminoácido	C	T	T	P	T	A	V	A	V	T	F	D	L	T	A	T	T	Y	G	E	N	I	Y	L	V	G	S	I	S		
Aspergillus awamori	posición	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	
	Aminoácido	C	T	T	P	T	A	V	A	V	T	F	D	L	T	A	T	T	Y	G	E	N	I	Y	L	V	G	S	I	S		
		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

FIG. 5D

Trichoderma reesei	posición	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	
	Aminoácido	A	L	G	N	W	S	T	S	A	A	V	A	L	D	A	V	N	Y	A	D	N	H	P	L	W	I	G	T	V	N	
Humicola grisea	posición	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	
	Aminoácido	A	L	G	N	W	D	T	S	K	A	V	T	L	S	A	S	G	Y	K	S	N	D	P	L	W	S	I	T	V	P	
Thielavia terrestris	posición	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	
	Aminoácido	A	L	G	G	W	D	T	S	K	A	V	P	L	S	A	A	G	Y	T	A	S	D	P	L	W	S	G	T	V	D	
Thermomyces lanuginosus	posición	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	
	Aminoácido	E	L	G	S	W	D	P	A	N	A	L	L	M	S	A	K	S	W	T	S	G	N	P	V	W	T	L	S	I	S	
Talaromyces emersonii	posición	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	
	Aminoácido	E	L	G	N	W	S	T	A	S	A	I	P	L	R	A	D	A	Y	T	N	S	N	P	L	W	Y	V	T	V	N	
Aspergillus niger	posición	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	
	Aminoácido	Q	L	G	D	W	E	T	S	D	G	I	A	L	S	A	D	K	Y	T	S	S	D	P	L	W	Y	V	T	V	T	
Aspergillus awamori	posición	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	
	Aminoácido	Q	L	G	D	W	D	T	S	D	G	I	A	L	S	A	D	K	Y	T	S	S	N	P	L	W	Y	V	T	V	T	
		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

FIG. 5D (continúa)

Trichoderma reesei	posición	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581
	Aminoácido	L	E	A	G	D	V	V	E	Y	K	Y	I	N	V	G	Q	D	G	S	V	T	-	W	E	S	D	P	N	H		
Humicola grisea	posición	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581		
	Aminoácido	I	K	A	T	G	S	A	V	Q	Y	K	Y	I	K	V	G	T	N	G	K	I	T	-	W	E	S	D	P	N	R	
Thielavia terrestris	posición	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576			
	Aminoácido	L	P	A	-	G	L	A	V	Q	Y	K	Y	I	N	V	A	A	D	G	G	V	T	-	W	E	A	D	P	N	H	
Thermomyces lanuginosus	posición	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563		
	Aminoácido	L	P	A	-	G	T	S	F	E	Y	K	F	I	R	K	D	D	G	S	S	D	V	V	W	E	S	D	P	N	R	
Talaromyces emersonii	posición	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569			
	Aminoácido	L	P	P	-	G	T	S	F	E	Y	K	F	F	K	N	Q	T	D	G	T	I	V	-	W	E	D	D	P	N	R	
Aspergillus niger	posición	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596			
	Aminoácido	L	P	A	-	G	E	S	F	E	Y	K	F	I	R	I	E	S	D	S	V	E	-	W	E	S	D	P	N	R		
Aspergillus awamori	posición	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595			
	Aminoácido	L	P	A	-	G	E	S	F	E	Y	K	F	I	R	I	E	S	D	S	V	E	-	W	E	S	D	P	N	R		

FIG. 5E



FIG. 6

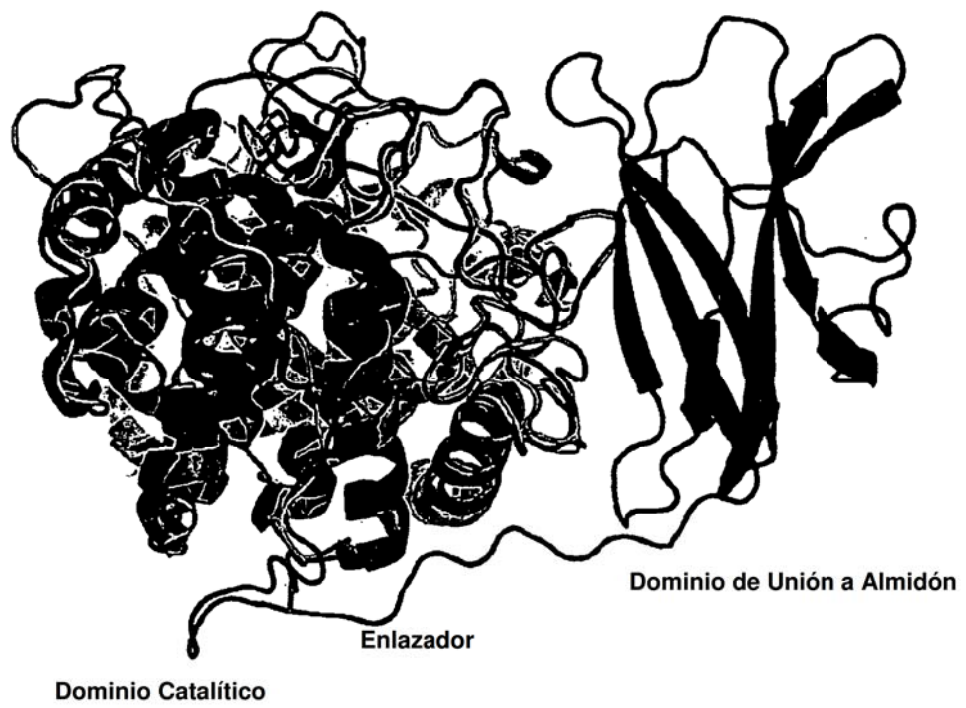


FIG. 7

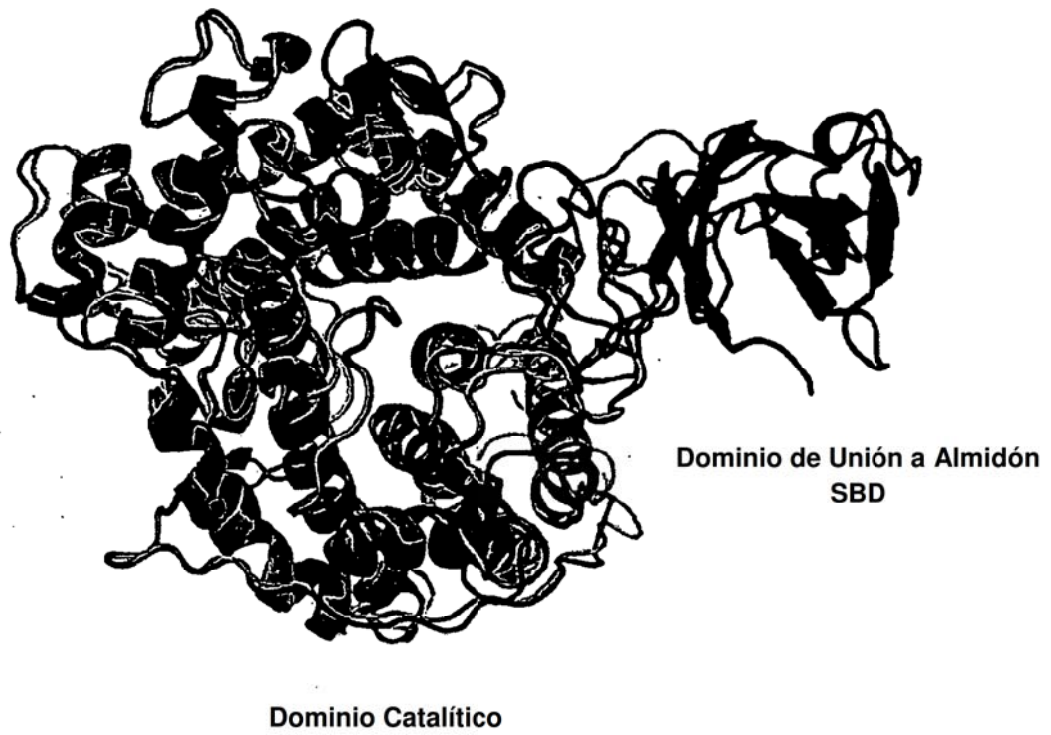


FIG. 8

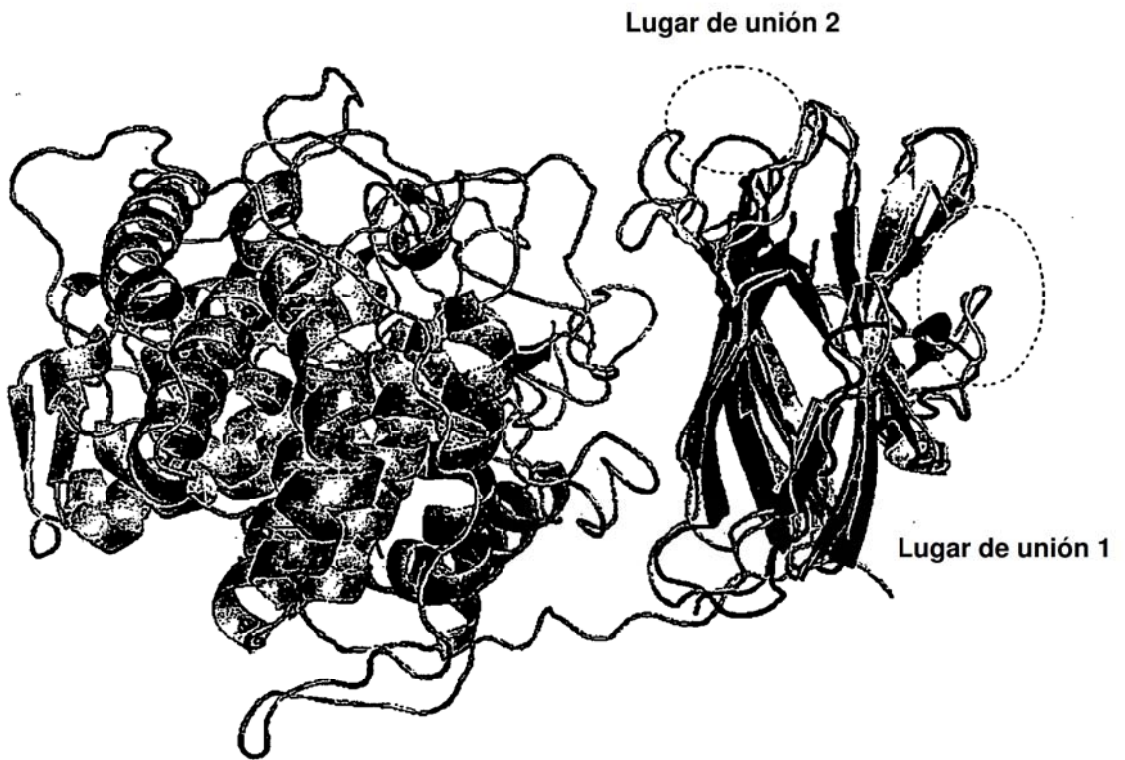


FIG. 9

