

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 517 524**

51 Int. Cl.:

**C12N 15/62** (2006.01)

**C12N 15/80** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **21.06.2007** **E 13151041 (4)**

97 Fecha y número de publicación de la concesión europea: **23.07.2014** **EP 2607486**

54 Título: **Regiones de escisión KEX2 de proteínas de fusión recombinantes**

30 Prioridad:

**11.07.2006 US 484814**

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

**03.11.2014**

73 Titular/es:

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

72 Inventor/es:

**WANG, HUAMING y**  
**WARD, MICHAEL**

74 Agente/Representante:

**RIZZO, Sergio**

**ES 2 517 524 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**

Regiones de escisión KEX2 de proteínas de fusión recombinantes

**CAMPO DE LA INVENCION**

5 **[0001]** La presente invención hace referencia a una secreción y escisión aumentada de proteínas deseadas, como proteínas de anticuerpos funcionales y enzimas industriales a partir de hongos filamentosos. La invención revela constructos de ADN de fusión, vectores y polipéptidos de fusión que incorporan regiones KEX2 para la escisión de proteínas y métodos de producción de las proteínas deseadas.

**ANTECEDENTES**

10 **[0002]** Durante la secreción de proteínas en una célula fúngica, determinadas proteínas se escinden mediante KEX2, un miembro de la familia KEX2 o familia "kexina" de serina peptidasa (EC 3.4.21.61). KEX2 es una endopeptidasa calcio-dependiente altamente específica que escinde el enlace peptídico que es inmediatamente C-terminal con respecto a un par de aminoácidos básicos (es decir, el "sitio KEX2") en un sustrato de proteínas durante la secreción de esa proteína. Las proteínas KEX2 contienen generalmente un residuo de cisteína cerca del residuo de histidina de su sitio activo y son inhibidas mediante *p*-mercuribenzoato. El miembro fundador de este grupo, la peptidasa KEX2 de *S. cerevisiae* (Fuller et al., 1989, Proc. Natl. Acad. Sci. USA 86:1434-1438), escinde los precursores de toxina killer y feromona factor  $\alpha$  durante su secreción.

15 **[0003]** Se ha informado de la producción de polipéptidos de fusión en varios organismos que incluyen *E. coli*, levadura y hongos filamentosos. Por ejemplo, se ha producido quimosina bovina en *Aspergillus niger* como una fusión a glucoamilasa de longitud completa (GAI) (Ward et al., (1990) Bio/technology 8:435 - 440; USP 6.265.204 y USP 6.590.078); se ha producido interleucina 6 humana (hIL6) en *Aspergillus nidulans* como una fusión a glucoamilasa de *A. niger* de longitud completa (GAI) (Contreras et al., (1991) Biotechnology 9:378 - 381); lisozima de clara de huevo de gallina (Jeenes et al., (1993) FEMS Microbiol. Lett. 107:267 - 273) y lactoferrina humana (Ward et al., (1995) Bio/Technology 13:498 -503) se han producido en *Aspergillus niger* como una fusión a los residuos 1 - 498 de glucoamilasa; y se ha producido quimosina bovina en *Aspergillus niger* como una fusión con alfa-amilasa nativa de longitud completa (Korman et al., (1990) Curr. Genet. 17: 203-212) y en *Aspergillus oryzae* como una fusión con formas truncadas de glucoamilasa de *A. oryzae* (Tsuchiya et al., (1994) Biosci. Biotech. Biochem. 58: 895 - 899). Se hace referencia también a Shoemaker et al., 1981 Bio/Technology 1: 691 -696; Nunberg et al., (1984) Mol. Cell. Biol. 4:2306 - 2315 y Boel et al., (1984) EMBO J. 3:1097 - 1102. En algunas de estas proteínas de fusión, se inserta un sitio de reconocimiento de proteasa KEX2 (Lys-Arg) entre una glucoamilasa y una proteína deseada (p.ej., Contreras et al., 1991 y Ward et al., 1995). Los inventores de la presente invención han descubierto que la secreción de proteína y/o escisión de proteína puede mejorarse en una proteína de fusión cuando el sitio de reconocimiento KEX2 ha sido manipulado para incluir una presecuencia de sitio KEX2 de aminoácidos.

20 **[0004]** La literatura específica de interés incluye: Ward et al., (2004) Appl. Environ. Microbiol. 70:2567-2576; Goller et al.,(1998) Appl. Environ. Microbiol. 64:3202-3208; La Grange et al., (1996) Appl. Environ. Microbiol. 62:1036-1044; Bergquist et al., (2002) Biochem. Biotechnol. 100:165-176; Spencer et al., (1998) Eur. J. Biochem. 258:107-112; Jalving et al., (2000) Appl. Environ. Microbiol. 66:363 - 368); Brenner y Fuller (1992) Proc. Natl. Acad. Sci. 89:922-926; Durand et al., (1999) Appl. Microbiol. Biotechnol. 52: 208-214; Ahn et al., (2004) Appl. Microbiol. Biotechnol. 64:833-839; Gouka et al., (1997) Appl Microbiol Biotechnol. 47:1-11 Broekhuijsen et al., (1993) J. Biotechnol. 31:135-145; MacKenzie et al., (1998) J. Biotechnol. 63:137 - 146 y las solicitudes de patente publicadas 20040018573 y 20050158825. También USP 4.816.567 y USP 6.331.415 revelan procesos para producir moléculas de inmunoglobulina en células huésped recombinantes. WO 01/18218 describe el uso de la señal de secreción de xilanas II para aumentar la expresión y secreción de una proteína de interés. WO 90/10075 describe la expresión de polipéptidos de fusión en levadura. Mikosch et al. (Journal of Biotechnology 52 (1996) 97-106) revela la secreción de inhibidor de proteinasa de la mucosa humana activa mediante *Aspergillus niger* tras el procesamiento similar a KEX2 de una proteína de fusión de inhibidor-glucoamilasa. WO 03/089614 revela la secreción de inmunoglobulina tras el procesamiento similar a KEX2 de una proteína de fusión de inmunoglobulina-glucoamilasa.

25 **[0005]** Aunque hay numerosos métodos disponibles para la producción de enzimas industriales y proteínas terapéuticas, continúa existiendo una necesidad de métodos alternativos para la producción de proteínas y específicamente para la producción de proteínas terapéuticas, como producción de anticuerpos, lo que resultará en un tiempo de aumento de escala relativamente rápido y altos niveles de proteína producida con riesgo limitado de contaminación mediante agentes virales u otros agentes extraños. La presente invención responde a esta necesidad.

30 **SUMARIO DE LA INVENCION**

**[0006]** Se proporcionan un constructo de ADN de fusión, vectores, un polipéptido de fusión, una célula que comprende el constructo de ADN de fusión, y métodos para mejorar la secreción y/o escisión de una proteína deseada desde una célula. Más específicamente, una región KEX2 abarcada por la invención ha sido incluida en un polipéptido de fusión para proporcionar la escisión de una proteína deseada desde el polipéptido de fusión. Por lo tanto, la invención pertenece a una región KEX2 para la escisión de proteínas. Por ello, en un primer aspecto, la presente invención proporciona un constructo de ADN de fusión que codifica un polipéptido de fusión que comprende en enlace operable desde el extremo 5' de dicho constructo,

un promotor;

una primera molécula de ADN que codifica una secuencia señal funcional como una secuencia de secreción en una célula fúngica filamentosa;

una segunda molécula de ADN que codifica una proteína portadora, donde la proteína portadora es un polipéptido fúngico secretado de manera natural o una porción funcional del mismo;

una tercera molécula de ADN que codifica una región KEX2, donde la región KEX2 es VAVYKR; y

una cuarta molécula de ADN que codifica una proteína deseada.

**[0007]** En otros aspectos, la presente invención proporciona un vector que comprende el constructo de ADN de fusión de la presente invención y una célula huésped que comprende el vector.

**[0008]** En otro aspecto adicional, la presente invención proporciona un proceso para producir una proteína deseada en una célula fúngica filamentosa que comprende:

(a) obtener una célula huésped fúngica filamentosa que comprende un constructo de ADN de fusión de la presente invención;

(b) cultivar la célula huésped en condiciones adecuadas que permitan la expresión y producción de la proteína deseada; y

(c) recuperar la proteína deseada.

**[0009]** En otro aspecto adicional, la presente invención proporciona un proceso para aumentar la producción de un anticuerpo a partir de una célula fúngica filamentosa que comprende obtener una célula fúngica filamentosa que comprende un constructo de ADN de fusión de la presente invención, cultivar la célula fúngica en condiciones adecuadas para la expresión del polipéptido de fusión y permitir la secreción del polipéptido de fusión, donde la secreción de la proteína deseada es aumentada en comparación con la secreción de un polipéptido de fusión equivalente que no incluye una presecuencia KEX2.

**[0010]** En otro aspecto adicional, la presente invención proporciona un polipéptido de fusión que comprende a partir de un extremo amino-terminal de dicho polipéptido de fusión una primera secuencia de aminoácidos que comprende:

(a) una secuencia señal funcional como secuencia secretora en una célula fúngica filamentosa;

(b) una segunda secuencia de aminoácidos que comprende una proteína portadora, donde la proteína portadora es un polipéptido fúngico secretado de manera natural;

(c) una tercera secuencia de aminoácidos que comprende una región KEX2, donde dicha región es VAVYKR; y

(d) una cuarta secuencia de aminoácidos que comprende una proteína deseada.

### **BREVE DESCRIPCIÓN DE LAS FIGURAS**

**[0011]** Determinados aspectos de la descripción detallada a continuación se entienden mejor cuando se leen conjuntamente con los dibujos que acompañan. Se enfatiza que, según la práctica común, las diversas características de los dibujos no se encuentran a escala. Por el contrario, las dimensiones de las diversas características se aumentan o reducen de manera arbitraria para mayor claridad. En los dibujos se encuentran incluidas las siguientes figuras:

La Fig. 1 ilustra de manera esquemática un modo de realización de un polipéptido de fusión según la invención, que incluye una proteína portadora, una región KEX2 y una proteína deseada, donde la proteína portadora se ilustra como un enlazador/núcleo de celobiohidrolasa I (CBH1, en inglés), que comprende el dominio catalítico y parte de la región enlazadora de la proteína CBH y la proteína deseada se ilustra como una cadena pesada o cadena ligera de anticuerpo.

La Fig. 2 representa un mapa del plásmido pTrex4-her2 cadena ligera DNA2.0 usado para la expresión de un polipéptido de fusión. El plásmido incluye un promotor de *cbh1* de *Trichoderma reesei*; un polipéptido que codifica una secuencia señal de CBH1 y proteína portadora; una región KEX2 insertada inmediatamente tras el sitio Spel, un polinucleótido que codifica la proteína deseada ilustrado como una cadena ligera de anticuerpo (trastuzumab); un terminador de celobiohidrolasa (*cbh1*) de *Trichoderma reesei*; un marcador de acetamidasa de *amdS Aspergillus nidulans*.

Las Fig. 3A - E proporcionan la secuencia de nucleótidos (SEQ ID NO: 103) (10885 pb) del plásmido pTrex4 -her2 cadena ligera DNA2.0 de la Fig. 2.

La Fig. 4 muestra un análisis Western blot de sobrenadante de células *Trichoderma reesei* cultivadas que comprenden secuencias de región KEX2 como se describe en mayor detalle en los ejemplos 1, 2, 3 y 4. La vía 1 representa un marcador de peso molecular (Véase Blue Plus 2, Invitrogen). Las vías 2 y 3 representan una variante GGGKR (SEQ ID NO: 5); la vía 4 representa una variante GGGKRGGG (SEQ ID NO: 7); la vía 5 representa una variante VAVEKR (SEQ ID NO: 9) de la región KEX2 abarcadas por la invención; y las vías 6 y 7 representan una variante KRGGG (SEQ ID NO: 2).

La Fig. 5 muestra un análisis Western blot de sobrenadante de células *Trichoderma reesei* cultivadas que comprenden regiones KEX2 abarcadas por la invención como se describe en mayor detalle en el ejemplo 5. La vía 1 representa un marcador de peso molecular como se describe arriba. Las vías 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 y 13 representan de manera correspondiente las regiones KEX2 VAVEKR (SEQ ID NO: 9), VAVWKR (SEQ ID NO: 25), VAVGKR (SEQ ID NO: 26), VAVRKR (SEQ ID NO: 27), VAVTKR (SEQ ID NO: 28), VAVVKR (SEQ ID NO: 29), VAVAKR (SEQ ID NO: 30), VAVLKR (SEQ ID NO: 31), VAVDKR (SEQ ID NO: 32), VAVNKR (SEQ ID NO: 33), VAVYKR (SEQ ID NO: 34), VAVHKR (SEQ ID NO: 35).

La Fig. 6 muestra un Western blot de sobrenadantes de células *Trichoderma reesei* cultivadas que contienen secuencias de región KEX2 como se describe en mayor detalle en los ejemplos 5, 6 y 7. Las vías 1 y 10 representan un marcador de peso molecular, como se describe arriba. Las vías 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15 y 16 representan de manera correspondiente regiones KEX2 AAVEKR (SEQ ID NO: 38), GAVEKR (SEQ ID NO: 37), MAVEKR (SEQ ID NO: 36), LAVEKR (SEQ ID NO: 39), WAVEKR (SEQ ID NO: 40), KAVEKR (SEQ ID NO: 41), PAVEKR (SEQ ID NO: 42), DAVEKR (SEQ ID NO: 51), VAVEKR (SEQ ID NO: 9), HAVEKR (SEQ ID NO: 52), QAVEKR (SEQ ID NO: 47), SAVEKR (SEQ ID NO: 46), NVISKR (SEQ ID NO: 22), y SDVTKR (SEQ ID NO: 24).

La Fig. 7 muestra un análisis Western blot de sobrenadantes de células *Trichoderma reesei* cultivadas que contienen secuencias de región KEX2 como se describe en mayor medida en el ejemplo 5. La vía 1 representa un marcador de peso molecular, como se describe arriba. Las vías 2, 3, 4, 5, 6, 7, 8, 9, 10 y 11 representan de manera correspondiente regiones KEX2 VAVEKR (SEQ ID NO: 9), VGVEKR (SEQ ID NO: 56), VTVEKR (SEQ ID NO: 40), VEVEKR (SEQ ID NO: 55), VPVEKR (SEQ ID NO: 62), VWVEKR (SEQ ID NO: 67), VKVEKR (SEQ ID NO: 58), VRVEKR (SEQ ID NO: 63), VVVEKR (SEQ ID NO: 66), y VIVEKR (SEQ ID NO: 57).

La Fig. 8 muestra un análisis Western blot de sobrenadantes de células *Trichoderma reesei* cultivadas que contienen secuencias de región KEX2 como se describe en mayor medida en el ejemplo 5. Las vías 1 - 11 representan de manera correspondiente regiones KEX2 VADEKR (SEQ ID NO: 70), VAAEKR (SEQ ID NO: 69), VAFEKR (SEQ ID NO: 72), VAGEKR (SEQ ID NO: 73), VAIEKR (SEQ ID NO: 74), VANEKR (SEQ ID NO: 76), VALKEKR (SEQ ID NO: 75), VASEKR (SEQ ID NO: 79), VAREKR (SEQ ID NO: 78) y VAPEKR (SEQ ID NO: 83).

La Fig. 9 muestra un gel de SDS-PAGE de sobrenadantes de células *A. niger* cultivadas que contienen una región KEX 2 VAVEKR (SEQ ID NO: 9) como se describe en mayor detalle en el ejemplo 8. La vía 1 representa un marcador de peso molecular, Marker 12 MW estándar (Invitrogen). Las vías 2, 3 y 4 representan 3 transformantes y corresponden respectivamente a transformantes A10, A11 y A12.

## **DEFINICIONES**

**[0012]** A menos que se defina lo contrario, todos los términos técnicos y científicos usados aquí tienen el mismo significado que el entendido comúnmente por un experto en la técnica a la que pertenece esta invención. Singleton, et al., DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY, 2D ED., John Wiley and Sons, Nueva 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 aquí usados. Aun así, se definen a continuación determinados términos para una mayor claridad y facilidad de referencia.

**[0013]** El término "recombinante" cuando se usa en referencia a una célula, ácido nucleico, proteína o vector, indica que la célula, ácido nucleico, proteína o vector ha sido modificado mediante la introducción de una proteína o ácido nucleico heterólogos o la alteración de una proteína o ácido nucleico nativos, o que la célula se deriva de una célula modificada de este modo. Por ello, por ejemplo, las células recombinantes expresan ácidos nucleicos o polipéptidos que no se encuentran en la forma nativa (no recombinante) de la célula o expresan genes nativos que de otra manera se expresarían de manera anormal, serían subexpresados o sobreexpresados o no se expresarían.

**[0014]** Un "gen" se refiere a un segmento de ADN que participa en la producción de un polipéptido e incluye regiones que preceden y siguen a las regiones codificantes, p.ej., el promotor y terminador, así como secuencias que intervienen (intrones) entre segmentos de codificación individuales (exones).

**[0015]** El término "ácido nucleico" abarca ADN, ARN, monocatenario o bicatenario y modificaciones químicas de los mismos. Los términos "ácido nucleico" y "polinucleótido" pueden usarse de manera intercambiable aquí.

Puesto que el código genético es degenerado, puede usarse más de un codón para codificar un aminoácido particular y la presente invención abarca polinucleótidos, que codifican una secuencia de aminoácidos concreta.

5 **[0016]** El término "constructo de ADN" significa una secuencia de ADN que se enlaza de manera operable a una secuencia de control adecuada capaz de efectuar la expresión de una proteína en un huésped adecuado. Dichas secuencias de control pueden incluir un promotor para efectuar la transcripción, una secuencia de operador opcional para controlar la transcripción, una secuencia de codificación de sitios de enlace de ribosoma adecuados en el ARNm, potenciadores y secuencias que controlan la terminación de la transcripción y traducción.

10 **[0017]** El término "constructo de ADN de fusión" o "ácido nucleico de fusión" hace referencia a un ácido nucleico que comprende desde 5' a 3' varias secuencias de polinucleótidos (p.ej., una molécula de ADN que codifica una secuencia señal, una molécula de ADN que codifica una proteína portadora, una molécula de ADN que codifica un sitio KEX2 y una molécula de ADN que codifica una proteína deseada) enlazadas de manera operable y que codifican un polipéptido de fusión.

15 **[0018]** Un "vector" se refiere a una secuencia de polinucleótidos diseñada para introducir ácidos nucleicos en uno o más tipos celulares. Los vectores incluyen vectores de clonación, vectores de expresión, vectores transportadores, plásmidos, partículas de fago, casetes y similares.

**[0019]** Un "vector de expresión" se refiere a un vector que tiene la capacidad de incorporar y expresar un fragmento de ADN heterólogo en una célula extraña. Muchos vectores de expresión eucarióticos o procarióticos se encuentran disponibles en el mercado.

20 **[0020]** Un "promotor" es una secuencia reguladora que participa en la unión de ARN polimerasa para iniciar la transcripción de un gen.

25 **[0021]** El término "secuencia señal" hace referencia a una secuencia de aminoácidos en el extremo amino de una proteína que dirige la proteína al sistema de secreción para la secreción a partir de una célula. La secuencia señal se escinde de la proteína antes de la secreción de la proteína. En algunos casos, una secuencia señal puede denominarse también "péptido señal" o "péptido líder". La definición de una secuencia señal es funcional. La forma madura de la proteína extracelular carece de la secuencia señal que se separa durante el proceso de secreción.

30 **[0022]** "Bajo control transcripcional" es un término bien comprendido en la técnica que indica que la transcripción de una secuencia de polinucleótidos, normalmente una secuencia de ADN, depende de si está unida de manera operable a un elemento que contribuye a la iniciación de o promueve la transcripción.

**[0023]** "Bajo control traduccional" es un término bien comprendido en la técnica que indica un proceso regulador que se produce tras la formación de ARNm.

35 **[0024]** El término "unido de manera operable" hace referencia a la yuxtaposición donde los elementos se encuentran en una disposición que les permite estar relacionados funcionalmente. Por ejemplo, un promotor está unido de manera operable a una secuencia codificante si controla la transcripción de la secuencia.

40 **[0025]** El término "marcador selectivo" hace referencia a una proteína capaz de expresión en un huésped que permite la facilidad de selección de aquellos huéspedes que contienen un vector o ácido nucleico introducido. Los ejemplos de marcadores seleccionables incluyen, sin carácter limitativo, antimicrobianos (p.ej., higromicina, bleomicina o cloranfenicol) y/o genes que confieren una ventaja metabólica, como una ventaja nutricional sobre la célula huésped.

**[0026]** Los términos "proteína" y "polipéptido" se usan aquí de manera intercambiable. Se usa aquí el código convencional de una letra o tres letras para los residuos de aminoácidos.

**[0027]** El término "proteína portadora" hace referencia a una secuencia de polipéptido o parte funcional de la misma de un polipéptido fúngico secretado de manera natural.

45 **[0028]** El término "proteína de anticuerpo" se usa de manera intercambiable con inmunoglobulinas (Igs), hace referencia a una proteína que contiene uno o más polipéptidos que se unen específicamente a un antígeno. Se incluyen mediante este término anticuerpos de cualquier isotipo, fragmentos de anticuerpos que conservan la unión específica al antígeno, incluyendo, sin carácter limitativo, anticuerpos Fab, Fv, scFv, Fd, Fab', Fv, F(ab')<sub>2</sub>, fragmentos de anticuerpos que conservan la unión específica al antígeno, anticuerpos monoclonales, anticuerpos quiméricos, anticuerpos humanizados, anticuerpos de una sola cadena, anticuerpos híbridos bifuncionales (es decir, biespecíficos) y proteínas de fusión que comprenden una parte de unión al antígeno de un anticuerpo y una proteína no de anticuerpo.

50

- 5 **[0029]** La forma monomérica de un anticuerpo comprende cuatro cadenas de polipéptido de dos tipos diferentes, una pesada y una ligera. Se reconocen los diferentes tipos de cadenas pesada y ligera. Las cadenas ligeras se dividen estructuralmente en dos dominios, una región variable (VL) y una región constante (CL). La cadena pesada también se divide en dominios estructurales diferentes. Por ejemplo, la cadena pesada comprende, desde el extremo amino-terminal, una región variable (VH), una región constante (CH1), una región bisagra, una segunda región constante (CH2) y una tercera región constante (CH3).
- 10 **[0030]** El término "polipéptido de fusión equivalente" hace referencia a un polipéptido de fusión que tiene una secuencia de aminoácidos idéntica comparada con un polipéptido de fusión de referencia, excepto por una presecuencia de sitio KEX2. Un primer polipéptido de fusión con una primera presecuencia de sitio KEX2 es equivalente a un segundo polipéptido de fusión que tiene una presecuencia de sitio KEX2 diferente si los polipéptidos tienen secuencias de aminoácidos idénticas, excepto por la diferencia en la presecuencia de sitio KEX2.
- 15 **[0031]** El término "derivado" abarca los términos "originado a partir de", "obtenido" u "obtenible a partir de" y "aislado a partir de".
- 20 **[0032]** "Cepa huésped" o "célula huésped" significa un huésped adecuado para un vector de expresión o constructo de ADN que comprende un polinucleótido que codifica un polipéptido y especialmente un polipéptido de fusión recombinante abarcado por la invención. En modos de realización específicos, las cepas huésped pueden ser una célula fúngica filamentosa. El término "célula huésped" incluye tanto células como protoplastos.
- 25 **[0033]** El término "hongos filamentosos" hace referencia a todas las formas filamentosas de la subdivisión Eumycotina (Véase, Alexopoulos, C. J. (1962), INTRODUCTORY MYCOLOGY, Wiley, Nueva York). Estos hongos se caracterizan por un micelio vegetativo con una pared celular compuesta por quitina, glucanos y otros polisacáridos complejos. Los hongos filamentosos de la presente invención son distintos de forma morfológica, fisiológica y genética de las levaduras. El crecimiento vegetativo por los hongos filamentosos se produce por elongación de las hifas y el catabolismo de carbono es obligatoriamente aeróbico.
- 30 **[0034]** El término "cultivar" se refiere a desarrollar una población de células microbianas en condiciones adecuadas en un medio sólido o líquido.
- 35 **[0035]** El término "heterólogo" en relación con un polinucleótido o polipéptido hace referencia a un polinucleótido o polipéptido que no se produce de manera natural en una célula huésped. En algunos modos de realización, la proteína es una proteína industrial importante desde un punto de vista comercial y en algunos modos de realización, la proteína heteróloga es una proteína terapéutica. Se desea que el término abarque proteínas que son codificadas por los genes de origen natural, genes mutados y/o genes sintéticos.
- 40 **[0036]** El término "homólogo" en relación con un polinucleótido o proteína hace referencia a un polinucleótido o proteína que se produce de manera natural en la célula huésped.
- 45 **[0037]** Los términos "recuperado", "aislado" y "separado" según su uso aquí hacen referencia a una proteína, célula, ácido nucleico o aminoácido que se extrae de al menos un componente con el que se asocia de manera natural.
- 50 **[0038]** Según su uso aquí, los términos "transformado", "transformado de manera estable" y "transgénico" usado en relación con una célula significa que la célula tiene una secuencia de ácido nucleico no nativa (p.ej., heteróloga) o copia adicional de una secuencia de ácido nucleico nativa (p.ej., homóloga) integrada en su genoma o tiene un plásmido episomal que se mantiene a lo largo de múltiples generaciones.
- [0039]** Según su uso aquí, el término "expresión" se refiere al proceso por el cual se produce un polipéptido basándose en la secuencia de ácido nucleico de un gen. El proceso incluye tanto transcripción como traducción.
- [0040]** El término proteína "glicosilada" significa una proteína que tiene moléculas de oligosacáridos añadidas a un determinado residuo de aminoácido en la proteína.
- [0041]** El término proteína "no glicosilada" es una proteína que no tiene moléculas de oligosacáridos unidas a la proteína.
- [0042]** El término "introducido" en el contexto de insertar una secuencia de ácido nucleico en una célula, significa "transfección" o "transformación" o "transducción" e incluye referencia a la incorporación de una secuencia de ácido nucleico en una célula eucariótica o procariótica donde la secuencia de ácido nucleico puede incorporarse al genoma de la célula (p.ej., ADN mitocondrial, cromosoma, plásmido o plástido), convertido en un replicón autónomo, o expresado de manera transitoria (p.ej., ARNm transfectado).

**[0043]** El término "KEX2" hace referencia a una endopeptidasa calcio-dependiente que tiene una actividad definida como EC 3.4.21.61, según la Nomenclatura de Enzimas de la IUBMB. KEX2 escinde una unión peptídica (el sitio de escisión KEX2) que es inmediatamente C-terminal con respecto a un par de aminoácidos básicos durante la secreción de proteína.

5 **[0044]** El término "región KEX2" hace referencia a una región de ocho a cuatro residuos de aminoácidos contiguos que se sitúa entre el extremo C-terminal de una proteína portadora y el extremo N-terminal de una proteína deseada en un polipéptido de fusión. La región KEX2 consta de un sitio KEX2 y una presecuencia de sitio KEX2.

10 **[0045]** El término "sitio KEX2" hace referencia a motivos de escisión KEX2 de dos aminoácidos en una proteína. Un sitio KEX2 contiene dos aminoácidos básicos contiguos (p.ej., lisina, histidina y/o arginina) en cualquier orden, (p.ej., KK, RR, KR o RK).

15 **[0046]** El término "presecuencia de sitio KEX2" hace referencia a los de dos a seis aminoácidos contiguos [(X)<sub>n</sub>, donde n es de 2 a 6] que preceden inmediatamente al sitio KEX2 (es decir, inmediatamente N-terminal con respecto al sitio). Por ejemplo, si una región KEX2 se define como VAVEKR, el motivo "KR" es el sitio KEX2 de la región; n es 4 y el motivo "VAVE" corresponde a la presecuencia de sitio KEX2 de la región.

**[0047]** El término "variante" se refiere a una región de una proteína que contiene uno o más aminoácidos diferentes en comparación con una proteína de referencia.

20 **[0048]** El término "proteína secretada" hace referencia a una región de un polipéptido que es liberada de una célula durante la secreción de proteínas. En algunos modos de realización, la proteína secretada es la proteína que se libera o escinde de un polipéptido de fusión recombinante de la invención.

**[0049]** El término "secreción" se refiere al movimiento selectivo de una proteína a través de una membrana en una célula huésped al espacio extracelular o medio circundante.

25 **[0050]** Los términos "determinar", "medir", "evaluar", "valorar" y "ensayar" se usan aquí de manera intercambiable para referirse a cualquier forma de medición, e incluyen la determinación de si un elemento está presente o no. El término incluye tanto determinaciones cuantitativas como cualitativas. La evaluación puede ser relativa o absoluta.

**[0051]** La "evaluación de la presencia de" incluye determinar la cantidad de algo presente, así como determinar si está presente o ausente.

**[0052]** Pueden aparecer otras definiciones de términos a lo largo de la especificación.

30 **DESCRIPCIÓN DETALLADA**

35 **[0053]** Antes de describir los modos de realización de ejemplo en mayor detalle, se debe entender que esta invención no se limita a los modos de realización concretos descritos, ya que, por supuesto, estos pueden variar. También se entenderá que la terminología utilizada aquí sigue el fin exclusivo de describir modos de realización concretos, y no pretende ser limitativa, ya que el alcance de la presente invención quedará limitado exclusivamente por las reivindicaciones adjuntas.

40 **[0054]** Cuando se proporciona un intervalo de valores, se entiende que cada valor intermedio, hasta la décima de la unidad del límite inferior a menos que el contexto indique claramente lo contrario, entre los límites superior e inferior del intervalo también se revela de manera específica. Cada intervalo menor entre cualquier valor mencionado o valor intermedio en un intervalo mencionado y cualquier otro valor mencionado o intermedio en ese intervalo mencionado queda abarcado en la invención. Los límites superior e inferior de estos intervalos más pequeños pueden incluirse o excluirse de manera independiente en el intervalo, y cada intervalo en el que alguno, ninguno o ambos límites están incluidos en los intervalos más pequeños también queda abarcado en la invención, sujeto a algún límite excluido de manera específica en el intervalo mencionado. Cuando el intervalo mencionado incluye uno o ambos límites, los intervalos que excluyen alguno o ambos de aquellos límites incluidos se incluyen también en la invención.

45 **[0055]** Aunque cualquier método y material similar o equivalente a aquellos descritos aquí puede usarse en la práctica o ensayos de la presente invención, se describen a continuación los métodos y materiales preferidos y de ejemplo.

50 **[0056]** Debe observarse que según su uso aquí y en las reivindicaciones adjuntas, las formas singulares "un", "una", "el" y "la" incluyen los referentes plurales a menos que el contexto indique claramente lo contrario. Por tanto, por ejemplo, la referencia a "un gen" incluye una pluralidad de dichos agentes candidatos y la referencia a

"la célula" incluye la referencia a una o más células y equivalentes de los mismos conocidos por aquellos con experiencia en la técnica, etc.

5 **[0057]** Las publicaciones analizadas aquí se proporcionan únicamente para su divulgación antes de la fecha de presentación de la presente solicitud. Nada de la presente memoria debe interpretarse como una admisión de que la presente invención no tiene derecho a preceder dicha publicación en virtud de invención anterior. Además, las fechas de la publicación proporcionadas pueden ser diferentes a las fechas de publicación reales, lo que debería confirmarse de manera independiente.

**POLIPÉPTIDOS DE FUSIÓN -**

10 **[0058]** Como se ha apuntado arriba, el polipéptido de fusión en cuestión comprende: a) una secuencia señal; b) una proteína portadora; c) una región KEX2 que comprende: i) un sitio KEX2 y ii) una presecuencia del sitio KEX2 inmediatamente N-terminal con respecto al sitio KEX2; y d) una proteína deseada.

15 **[0059]** La Fig. 1 ilustra un polipéptido de fusión referido de la invención. Las diversas partes del polipéptido mencionado (es decir, "secuencia señal", proteína portadora, "región KEX2" y "proteína deseada") son etiquetadas así únicamente para mayor claridad y conveniencia. Se reconoce que el polipéptido de fusión en cuestión también puede denominarse "proteína" o "proteína precursora" puesto que contiene generalmente una región N-terminal que se separa durante la secreción y una región C-terminal que es secretada.

Secuencias señal y proteínas portadoras -

20 **[0060]** La secuencia señal de un polipéptido de fusión en cuestión puede ser cualquier secuencia señal que facilite la secreción de proteínas desde una célula fúngica filamentosa. En determinados modos de realización, el polipéptido de fusión referido puede comprender una secuencia señal para una proteína que se conoce que es altamente secretada a partir de la célula filamentosa en la que se va a producir la proteína de fusión. La secuencia señal empleada puede ser endógena o no endógena a la célula en la que el polipéptido de fusión va a producirse. En modos de realización concretos, la secuencia señal puede comprender un "portador" que contiene la secuencia señal en su extremo N-terminal, donde el portador es al menos una parte N-terminal de una proteína que es endógena a la célula y secretada eficientemente por la célula.

30 **[0061]** Las secuencias señal y portadores adecuados son conocidos en la técnica (véase, p.ej., Ward et al, Bio/Technology 1990 8:435-440; y Paloheimo et al, Applied and Environmental Microbiology 2003 69: 7073-7082). Los ejemplos de secuencias señal y proteínas portadoras adecuadas incluyen aquellos de celobiohidrolasa I, celobiohidrolasa II, endoglucanasas I, II y III,  $\alpha$ -amilasa, aspartil proteasas, glucoamilasa, fitasa, mananasa,  $\alpha$  y  $\beta$  glucosidasas, quimosina bovina, interferón humano y activador tisular del plasminógeno humano y secuencias señal eucarióticas consenso sintéticas como aquellas descritas por Gwynne et al., (1987) Bio/Technology 5:713-719.

35 **[0062]** En algunos modos de realización, si se emplea *Trichoderma* (p.ej. *T. reesei*) como célula huésped, la secuencia señal o portador de mananasa I de *T.reesei* (Man5A, o MANI), celobiohidrolasa II de *T. reesei* (Cel6A o CBHII), endoglucanasa I (Cel7b o EGI), endoglucanasa II (Cel5a o EGII), endoglucanasa III (Cel12A o EGIII), xilanasas I o II (XynIIa o XynIIb) o celobiohidrolasa I de *T. reesei* (Cel7a o CBHI) pueden emplearse en el polipéptido de fusión.

40 **[0063]** En otros modos de realización, si se usa un *Aspergillus* (p.ej. *A. niger*) como célula huésped, la secuencia señal o portador de glucoamilasa de *A. niger* (GlaA) o alfa-amilasa puede emplearse en el polipéptido de fusión. Las glucoamilasas de *Aspergillus niger* y *Aspergillus awamori* tienen secuencias de aminoácidos idénticas. Se reconocen generalmente dos formas de la enzima en los sobrenadantes de cultivos. GAI es la forma de longitud completa (residuos de aminoácidos 1 - 616) y GAII es un fragmento proteolítico natural que comprende los residuos de aminoácidos 1 - 512. Se conoce que GAI se pliega en dos dominios separados unidos por una región enlazadora extendida. Los dos dominios son el dominio catalítico de 471 residuos (aminoácidos 1 - 471) y el dominio de enlace al almidón de 108 residuos (aminoácidos 509 - 616), siendo la región enlazadora entre los dos dominios de 36 residuos (aminoácidos 472-508). GAII carece del dominio de enlace al almidón. Se hace referencia a Libby et al., (1994) Protein Engineering 7:1109 - 1114. En algunos modos de realización, la glucoamilasa que se usa como proteína portadora y que incluye una secuencia señal presentará más de un 95%, 96%, 97%, 98% y 99% de identidad de secuencia con un dominio catalítico de una glucoamilasa de *Aspergillus* o *Trichoderma*. El término "dominio catalítico" hace referencia a una región o parte estructural de la secuencia de aminoácidos de una proteína que posee la actividad catalítica de la proteína.

50 **[0064]** En determinados modos de realización, la secuencia señal y la proteína portadora se obtienen del mismo gen. En algunos modos de realización, la secuencia señal y la proteína portadora se obtienen de genes diferentes.



**[0065]** La proteína portadora puede incluir todo o parte de la secuencia madura de un polipéptido secretado. En algunos modos de realización, se usan polipéptidos secretados de longitud completa. Sin embargo, pueden emplearse partes funcionales de polipéptidos secretados. Según su uso aquí, "parte" de un polipéptido secretado o equivalentes gramaticales significa un polipéptido secretado truncado que conserva su capacidad para plégarse en una configuración normal, aunque truncada.

**[0066]** En algunos casos, el truncamiento del polipéptido secretado significa que la proteína funcional conserva una función biológica. En algunos modos de realización, se usa el dominio catalítico del polipéptido secretado, aunque podrían usarse otros dominios funcionales, por ejemplo el dominio de unión al sustrato. En un modo de realización, cuando se usa glucoamilasa como la proteína portadora (es decir, glucoamilasa de *Aspergillus niger*), las partes funcionales preferidas conservan el dominio catalítico de la enzima e incluyen los aminoácidos 1 - 471 (véase, WO 03089614, p.ej., ejemplo 10). En otro modo de realización, cuando se usa CBH I como la proteína portadora (es decir, CBH I de *Trichoderma reesei*) las partes funcionales preferidas conservan el dominio catalítico de la enzima. Se hace referencia a la SEQ ID NO: de la figura 2 de WO 05093073, donde se revela la secuencia que codifica una secuencia señal de CBH1 de *Trichoderma reesei*, dominio catalítico de CBH1 de *T. reesei* (también denominado dominio de núcleo o núcleo catalítico) y enlazador de CBH1 de *T. reesei*. En algunos modos de realización, una proteína portadora de CBH1 y que incluye una secuencia señal tendrá un 95%, 96%, 97%, 98% y 99% de identidad de secuencia con la SEQ ID NO:1 de la figura 2 de WO 05093073.

**[0067]** En general, si la proteína portadora es una proteína truncada, es una proteína truncada de manera C-terminal (es decir, contiene un extremo N-terminal intacto). De manera alternativa, la proteína portadora puede truncarse de manera N-terminal, o truncarse opcionalmente en ambos extremos para dejar una parte funcional. Generalmente, dichas partes de una proteína secretada que comprende una proteína portadora comprenden más del 50%, más del 70%, más del 80% y más del 90% de la proteína secretada y preferiblemente la parte N-terminal de la proteína secretada. En algunos modos de realización, la proteína portadora incluirá una región enlazadora además del dominio catalítico. En los constructos de fusión de los ejemplos aquí contenidos, parte de la región enlazadora de la proteína de CBHI se usó en la proteína portadora.

**[0068]** Según su uso aquí, la primera secuencia de aminoácidos que comprende una secuencia señal funcional como secuencia secretora es codificada por una primera molécula de ADN. La segunda secuencia de aminoácidos que comprende la proteína portadora es codificada por una segunda secuencia de ADN. Sin embargo, como se describe arriba la secuencia señal y la proteína portadora pueden obtenerse del mismo gen.

#### Región KEX2 -

**[0069]** La región KEX2 comprende un sitio KEX2 ( $B_1B_2$ ) y una presecuencia de sitio KEX2 ( $(X)_{n=2-6}$ ) inmediatamente N-terminal con respecto a dicho sitio KEX2. En algunos modos de realización, la región KEX2 proporciona medios para la escisión (separación) en el extremo amino-terminal de la proteína deseada desde el polipéptido de fusión *in vivo*. La región KEX2 de un polipéptido de fusión de los abarcados por la invención no es una región de origen natural entre la proteína portadora y la proteína deseada.

**[0070]** El sitio de escisión KEX2, que se produce en el extremo C-terminal de la región KEX2, puede escindirse mediante una proteasa fúngica filamentosa nativa (p.ej., una proteasa tipo KEXB de *Aspergillus* nativa o proteasa de KEX2 de *Trichoderma* nativa). La proteína deseada se escinde de un polipéptido de fusión según la invención inmediatamente aguas abajo del sitio KEX2.

**[0071]** El sitio KEX2 contiene la secuencia de aminoácidos "B<sub>1</sub>B<sub>2</sub>" donde B<sub>1</sub> y B<sub>2</sub>, son independientemente, aminoácidos básicos. Preferiblemente, el sitio KEX2 incluye cualquiera de KK, KR, RK o RR y más preferiblemente es KR.

**[0072]** La presecuencia de sitio KEX2 comprende una secuencia de aminoácidos  $(X)_{n=2-6}$  donde X es cualquier aminoácido y n es de 2 a 6 y preferiblemente 4. La región KEX2 según se define aquí no se encuentra de manera natural en la proteína portadora en el extremo C-terminal de la proteína portadora, que comprende el polipéptido de fusión según la invención. En algunos modos de realización, la presecuencia de sitio KEX2 es una secuencia de aminoácidos que es diferente de los residuos de aminoácido  $(X)_{n=2-6}$  contiguos de origen natural en el extremo C-terminal de la proteína portadora. Sin embargo, los residuos de aminoácido  $(X)_{n=2-6}$  contiguos pueden encontrarse en otras partes de la proteína portadora y pueden unirse con un sitio KEX2 ( $B_1B_2$ ) pero la región KEX2 no se acoplará al extremo N-terminal de la proteína deseada.

**[0073]** Según se describe aquí, cuando la presecuencia de sitio KEX2 se define como  $X_4X_3X_2X_1B_1B_2$ ,

- a)  $X_1$ ,  $X_2$  y  $X_3$  no son G;
- b)  $X_1$  no es S, si  $X_2$  y  $X_3$  son G,  $X_4$  es A, o  $X_3$  es S;

- c)  $X_4$  no es T, si  $X_3$  es A y  $X_2$  es S; o  
d)  $X_1$  no es D.

**[0074]** Como se describe aquí, la región KEX2 es  $X_4X_3X_2X_1B_1B_2$  donde  $B_1B_2$  es KR y

- 5 a)  $X_1$ ,  $X_2$  y  $X_3$  no son G;  
a)  $X_1$  no es S, si  $X_2$  y  $X_3$  son G,  $X_4$  es A, o  $X_3$  es S;  
b)  $X_4$  no es T, si  $X_3$  es A y  $X_2$  es S; o  
c) d)  $X_1$  no es D.

10 **[0075]** Según se describe aquí, la presecuencia de sitio KEX2 se define como  $X_4X_3X_2X_1$  donde,

- a)  $X_4$  es V, S, N, L, o K;  
a)  $X_3$  es A, V, D, W, E o P;  
b)  $X_2$  es V, I, L o F; y  
c)  $X_1$  es E, S, T o Y.

15

**[0076]** Según se describe aquí, la presecuencia de sitio KEX2 se define como  $X_4X_3X_2X_1$  donde,

- a)  $X_4$  es V, N, o L;  
a)  $X_3$  es A, V, D, W, E o P;  
b)  $X_2$  es V, I, L o F; y  
c)  $X_1$  es E o Y.

20

**[0077]** Según se describe aquí, la región KEX2  $X_4X_3X_2X_1KR$  puede seleccionarse entre el grupo de  $X_4$  es V;  $X_3$  es A;  $X_2$  es V;  $X_1$  es E o Y y combinaciones de los mismos.

**[0078]** Como se describe aquí, la presecuencia de sitio KEX2 se selecciona entre el grupo compuesto por VAVE (SEQ ID NO: 84); NVIS (SEQ ID NO: 85); SDVT (SEQ ID NO: 86); VAVY (SEQ ID NO: 87); LAVE (SEQ ID NO: 88); KAVE (SEQ ID NO: 89); VAIE (SEQ ID NO: 90); VALE (SEQ ID NO: 91); VAFE (SEQ ID NO: 92); WVVE (SEQ ID NO: 93); VEVE (SEQ ID NO: 94); y VPVE (SEQ ID NO: 95).

25

**[0079]** Como se describe aquí, la presecuencia de sitio KEX2 no es KSRS (SEQ ID NO: 109); SRIS (SEQ ID NO: 111); GGGS (SEQ ID NO: 110); TSTY (SEQ ID NO: 96); ASIS (SEQ ID NO: 97); ATAS (SEQ ID NO: 98); TASQ (SEQ ID NO: 99); TASL (SEQ ID NO: 100), SVIS (SEQ ID NO: 101); NVIS (SEQ ID NO: 85); GGG; TSRD (SEQ ID NO: 102); SPMD (SEQ ID NO: 106); DLGE (SEQ ID NO: 107); o TPTA (SEQ ID NO: 108).

30

**[0080]** En algunos modos de realización, la presecuencia de sitio KEX2 proporciona una escisión y/o secreción mejoradas de una proteína deseada a partir de una célula huésped en comparación con la escisión y/o secreción de la proteína deseada a partir de un polipéptido de fusión equivalente que carece de una presecuencia de sitio KEX2.

35

**[0081]** En algunos modos de realización, la presecuencia de sitio KEX2 es una presecuencia de sitio KEX2 optimizada. Una presecuencia de KEX2 optimizada es una presecuencia de KEX2 abarcada por la invención pero que proporciona una escisión o secreción mayor o más eficiente desde una célula huésped en comparación con otras variantes de presecuencias de sitio KEX2.

**[0082]** En algunos modos de realización, el polipéptido de fusión abarcado por la invención incluirá una presecuencia de KEX2 optimizada como la presecuencia de KEX2. La presecuencia de KEX2 optimizada puede emplearse con cualquier secuencia señal, cualquier región portadora de una proteína secretada, cualquier sitio KEX2 o cualquier proteína deseada. La región KEX2 mencionada que contiene una presecuencia de sitio KEX2 optimizada puede no ser de origen natural. En determinados modos de realización, la región KEX2 referida que contiene una presecuencia de sitio KEX2 optimizada no se encuentra en ninguna proteína que es secretada desde una célula fúngica filamentosa.

40

45

#### Proteínas deseadas -

**[0083]** La proteína deseada (o la proteína portadora) puede ser cualquier parte de una proteína que puede ser secretada desde una célula fúngica filamentosa. Tales proteínas incluyen las denominadas enzimas industriales, proteínas terapéuticas, hormonas, proteínas estructurales, proteínas plasmáticas, aditivos alimentarios y comestibles y similares. La proteína deseada puede ser una proteína heteróloga u homóloga y puede incluir polipéptidos híbridos que comprenden una combinación de polipéptidos completos o parciales cada uno de los

50

5 cuales puede ser homólogo o heterólogo con respecto al huésped de expresión fúngico. La proteína secretada deseada puede derivarse de fuentes bacterianas (p.ej., especies *Bacillus* y *Pseudomonas*), fúngicas (p.ej., especies *Aspergillus*, *Trichoderma*, *Humicola*, o *Mucor*), virales (p.ej., Hepatitis A o B o Adenovirus), mamíferas (p.ej., seres humanos o ratones) y vegetales. Las proteínas deseadas incluyen variaciones alélicas de proteínas de origen natural así como variaciones obtenidas por ingeniería genética.

10 **[0084]** En un modo de realización, la proteína deseada puede ser una enzima como una carbohidrasa, como una  $\alpha$ -amilasa hidrolizante de almidón, una  $\alpha$ -amilasa alcalina, una  $\beta$ -amilasa, una celulasa, una dextranasa, una  $\alpha$ -glucosidasa, una  $\alpha$ -galactosidasa, una glucoamilasa, una hemicelulasa, una pentosanasa, una xilanas, una invertasa, una lactasa, una naringanasa, una pectinasa o una pullulanasa; una proteasa como una proteasa ácida, una proteasa alcalina, bromelaína, ficina, una proteasa neutra, papaína, pepsina, una peptidasa, cuajo, renina, quimosina, subtilisina, termolisina, una proteinasa aspártica, o tripsina; una enzima hidrolizante de almidón granular, como una glucoamilasa o una alfa-amilasa; una lipasa o esterasa, como una trigliceridasa, una fosfolipasa, una esterasa pregástrica, una fosfatasa, una fitasa, una amidasa, una iminoacilasa, una glutaminasa, una lisozima, o una penicilina acilasa; una isomerasa como glucosa isomerasa; una enzima oxidante de fenol, p.ej., una lacasa; oxidorreductasas, p.ej., una oxidasa de aminoácidos, una catalasa, una cloroperoxidasa, una glucosa oxidasa, una hidroxisteroide deshidrogenasa o una peroxidasa; una liasa como acetolactata descarboxilasa; una  $\beta$ -descarboxilasa aspártica, una fumarasa o una histadasa; una transferasa como ciclodextrina glicosiltransferasa o una aciltransferasa; o una ligasa, por ejemplo. En modos de realización específicos, la proteína puede ser una aminopeptidasa, una carboxipeptidasa, una quitinasa, una glucoamilasa, una alfa-amilasa, una cutinasa, una fitasa, una desoxirribonucleasa, una  $\alpha$ -galactosidasa, una  $\beta$ -galactosidasa, una  $\beta$ -glucosidasa, una lacasa, una manosidasa, una mutanasa, una enzima pectinolítica, una polifenoloxidas, ribonucleasa o transglutaminasa.

25 **[0085]** En otros modos de realización, la proteína deseada puede ser una proteína terapéutica (es decir, una proteína que tiene una actividad biológica terapéutica). Los ejemplos de proteínas terapéuticas adecuadas incluyen: eritropoyetina, citocinas como interferón- $\alpha$ , interferón- $\beta$ , interferón- $\gamma$ , interferón- $\omega$  y granulocito-CSF, GM-CSF, factores de coagulación como factor VIII, factor IX y proteína C humana, antitrombina III, trombina, cadena  $\alpha$  de receptor de IgE soluble, inmunoglobulina, como inmunoglobulina G (IgG), fragmentos de IgG, fusiones de IgG, IgM o IgA; interleucinas, urocinasa, quimasa e inhibidor de tripsina urea, proteína de unión a IGF, factor de crecimiento epidérmico, factor de liberación de hormona del crecimiento, proteína de fusión de anexina V, angioestatina, factor de crecimiento endotelial vascular 2, factor inhibidor de progenitor mielóide 1, osteoprotegerina,  $\alpha$ -1-antitripsina,  $\alpha$ -fetoproteínas, ADNasa II, *kringle* 3 del plasminógeno humano, glucocerebrosidasa, proteína 1 de unión a TNF, hormona estimulante del folículo, Ig-antígeno 4 asociado al linfocito T citotóxico, activador transmembrana y modulador del calcio y ligando de ciclofilina, fusión de Fc y receptor de TNF soluble, proteína 1 similar al glucagón y agonista del receptor de IL-2.

35 **[0086]** En algunos modos de realización preferidos, la proteína deseada es una inmunoglobulina de cualquier clase, G, A, M, E o D. (Véase, USP 4.816.567 y las referencias citadas en ella para el análisis de la estructura de inmunoglobulinas). En otros modos de realización preferidos, las proteínas de anticuerpo como los anticuerpos monoclonales que incluyen cadenas ligeras o pesadas y fragmentos de las mismas. En modos de realización adicionales, los anticuerpos humanizados resultan de especial interés como proteína deseada (p.ej., trastuzumab (herceptin)). Algunos ejemplos específicos de fragmentos de anticuerpo monoclonal preferidos son formas truncadas de la cadena pesada para separar parte de la región constante como fragmentos Fab en los que la cadena pesada (Fd) carece de la región bisagra y los dominios CH2 y CH3; fragmentos Fab' en los que la cadena pesada incluye la región bisagra pero carece de los dominios CH2 y CH3; y fragmentos F(ab')<sub>2</sub> que incluyen la parte Fab conectada mediante la región bisagra. (Verma et al., (1998) J. Immunological Methods 216:165 - 181 y Pennell y Eldin (1998) Res. Immunol. 149:599 - 603). También resultan de interés los anticuerpos monocatenarios (ScFv) y anticuerpos de un solo dominio (p.ej., anticuerpos de camélidos).

**[0087]** En algunos modos de realización especialmente preferidos, un polipéptido de fusión según la invención comprenderá en unión operable una secuencia señal, una proteína portadora, una región KEX2 y una proteína deseada como se indica a continuación:

#### 50 **VECTORES Y CONSTRUCTOS DE ADN DE FUSIÓN -**

55 **[0088]** En algunos modos de realización, la invención proporciona un constructo de ADN de fusión que codifica un polipéptido de fusión como se ha revelado arriba, que comprende en unión operativa desde el extremo 5' de dicho constructo, un promotor; una primera molécula de ADN que codifica una secuencia señal; una segunda molécula de ADN que codifica una proteína portadora; una tercera molécula de ADN que codifica una región KEX2, dicha región KEX2 que comprende un sitio KEX2 y una presecuencia de sitio KEX2 inmediatamente 5' con respecto al sitio KEX2; y una cuarta molécula de ADN que codifica una proteína deseada. Puesto que se conoce el código genético, el diseño y producción de estos ácidos nucleicos se enmarca perfectamente en las habilidades del experto, dada la descripción del polipéptido de fusión en cuestión. En determinados modos de realización, los ácidos nucleicos pueden ser de codones optimizados para la expresión del polipéptido de fusión

en una célula huésped concreta. Puesto que las tablas de uso de codones están disponibles para numerosas especies de hongos filamentosos, el diseño y producción de ácidos nucleicos de codones optimizados que codifican un polipéptido de fusión mencionado se enmarca dentro de la habilidad del experto en la técnica.

#### Promotores -

- 5 **[0089]** Los ejemplos de promotores adecuados para dirigir la transcripción del ácido nucleico mencionado en una célula huésped fúngica filamentosa son promotores obtenidos de los genes para TAKA amilasa de *Aspergillus oryzae*, proteinasa aspártica de *Rhizomucor miehei*, alfa-amilasa neutra de *Aspergillus niger*, alfa-amilasa estable en ácido de *Aspergillus niger* (Korman et al (1990) Curr. Genet 17:203-212; Gines et al., (1989) Gene 79: 107 - 117), glucoamilasa de *Aspergillus niger* o *Aspergillus awamori* (glaA) (Nunberg et al., (1984) Mol. Cell Biol. 4:2306 - 2315; Boel E. et al., (1984) EMBO J. 3: 1581 -1585), lipasa de *Rhizomucor miehei*, proteasa alcalina de *Aspergillus oryzae*, triosa fosfato isomerasa de *Aspergillus oryzae*, acetamidasa de *Aspergillus nidulans* (Hyner et al., (1983) Mol. Cell. Biol. 3:1430 - 1439), amiloglucosidasa de *Fusarium venenatum*, proteasas similares a la tripsina de *Fusarium oxysporum* (WO 96/00787), celobiohidrolasa I de *Trichoderma reesei* (Shoemaker et al. (1984) EPA EPO 0137280), celobiohidrolasa II de *Trichoderma reesei*, endoglucanasa I de *Trichoderma reesei*, endoglucanasa II de *Trichoderma reesei*, endoglucanasa III de *Trichoderma reesei*, endoglucanasa IV de *Trichoderma reesei*, endoglucanasa V de *Trichoderma reesei*, xilanasa I de *Trichoderma reesei*, xilanasa II de *Trichoderma reesei*, beta-xilosidasa de *Trichoderma reesei*, así como el promotor NA2-tpi (un híbrido de los promotores de los genes para la alfa-amilasa neutra de *Aspergillus niger* y triosa fosfato isomerasa de *Aspergillus oryzae*); y promotores híbridos, truncados y mutantes de los mismos. Se hace referencia también a Yelton et al., (1984) Proc. Natl. Acad. Sci. USA 81:1470 - 1474; Mullaney et al., (1985) Mol. Gen. Genet. 199:37 - 45; Lockington et al., (1986) Gene 33: 137 - 149; Macknight et al., (1986) Cell 46: 143 - 147; Hynes et al., (1983) Mol. Cell Biol. 3: 1430 - 1439. Los promotores eucarióticos superiores como promotor temprano SV40 (Barclay et al (1983) Molecular and Cellular Biology 3:2117 - 2130) también pueden ser útiles. Los promotores pueden ser promotores constitutivos o inducibles. Algunos promotores preferidos incluyen una celobiohidrolasa I o II de *Trichoderma reesei*, una endoglucanasa I, II o III de *Trichoderma reesei* y una xilanasa II de *Trichoderma reesei*.

#### Vectores -

- [0090]** Un polinucleótido en cuestión puede estar presente en un vector, por ejemplo, un vector fágico, plasmídico, viral o retroviral. En determinados modos de realización, el vector puede ser un vector de expresión para expresar un polipéptido de fusión en cuestión en una célula fúngica filamentosa.
- 30 **[0091]** Los vectores para la expresión de proteínas recombinantes son bien conocidas en la técnica (Ausubel, et al, Short Protocols in Molecular Biology, 3ª ed., Wiley & Sons, 1995; Sambrook, et al., Molecular Cloning: A Laboratory Manual, Segunda Edición, (1989) Cold Spring Harbor, N.Y.).
- [0092]** Un constructo de ADN de fusión según la invención puede construirse usando técnicas conocidas como se describe generalmente por ejemplo en la publicación EPO 0 215 594.
- 35 **[0093]** Los fragmentos de polinucleótido natural o sintético que codifican la proteína deseada (p.ej., una inmunoglobulina) puede incorporarse en vectores o constructos de ácido nucleico heterólogos, capaces de introducirse y replicar en una célula fúngica filamentosa.
- [0094]** Una vez que se fabrica un constructo de ADN o más específicamente un constructo de ADN de fusión abarcado por la invención puede incorporarse a una variedad de vectores como se conoce en la técnica. Aunque el constructo de ADN incluirá preferiblemente una secuencia promotora, en algunos modos de realización el vector incluirá otras secuencias reguladoras funcionales en el huésped a transformar, como sitios de unión de ribosomas, secuencias de iniciación y terminación de la transcripción, secuencias terminadoras, señales de poliadenilación, potenciadores y/o activadores. En algunos modos de realización, un polinucleótido que codifica la proteína deseada y región KEX2 se insertará en un vector que comprende un promotor, secuencia señal y proteína portadora en un sitio de endonucleasa de restricción apropiado mediante procedimientos estándares. Dichos procedimientos y procedimientos de subclonación relacionados se considera que recaen dentro del alcance del conocimiento de aquellos con experiencia en la técnica.
- 40 **[0095]** Las secuencias de terminación que son reconocidas por el huésped de expresión para terminar la transcripción pueden unirse de manera operable al extremo 3' del constructo de ADN de fusión que codifica la proteína de fusión a expresar. Aquellos con experiencia en la técnica son conocedores de diversas secuencias terminadoras que pueden usarse con hongos filamentosos. Los ejemplos no limitativos incluyen el terminador del gen trpC de *Aspergillus nidulans*, (Yelton M. et al., (1984) Proc. Natl. Acad. Sci. USA 81:1470 - 1474) o el terminador de los genes de glucoamilasa de *Aspergillus niger* (Nunberg et al. (1984) Mol. Cell. Biol. 4:2306 - 2353) o el terminador del gen de celobiohidrolasa I de *Trichoderma reesei*.
- 55 **[0096]** Las secuencias de poliadenilación son secuencias de ADN que cuando se transcriben son reconocidas

por el huésped de expresión para añadir residuos de poliadenosina a ARNm transcrito. Los ejemplos incluyen secuencias de poliadenilación del gen *trpC* de *A. nidulans*, (Yelton M. et al., (1984) Proc. Natl. Acad. Sci. USA 81:1470 - 1474); de gen de glucoamilasa de *A. niger* (Nunberg et al. (1984) Mol. Cell. Biol. 4:2306 - 2315); el gen de alfa-amilasa de *A. niger* o *A. oryzae* y el gen de carboxil proteasa de *Rhizomucor miehei*. Cualquier secuencia de poliadenilación fúngica es probable que sea funcional en la presente invención.

**[0097]** En modos de realización adicionales, el constructo de ADN de fusión o el vector que comprende el constructo de ADN de fusión contendrá un gen marcador seleccionable para permitir la selección de células huésped transformadas. Los genes marcadores de selección son conocidos en la técnica y variarán con la célula huésped usada. Los ejemplos de marcadores seleccionables incluyen, sin carácter limitativo, aquellos que confieren resistencia antimicrobiana (p.ej., higromicina, bleomicina, cloramfenicol y fleomicina). Los genes que confieren una ventaja metabólica, como marcadores selectivos nutricionales, también serán útiles en la invención. Algunos de estos marcadores incluyen *amdS*. También pueden usarse como marcadores de selección genes que codifican secuencias que complementan un defecto auxotrófico (p.ej., complementación de *pyr4* de un *A. nidulans*, *A. awamori* o *Trichoderma reesei* deficiente en *pyr4* y complementación de *argB* de una cepa deficiente en *argB*). Se hace referencia a Kelley et al., (1985) EMBO J. 4: 475 - 479; Penttila et al., (1987) Gene 61:155 -164 y Kinghorn et al (1992) Applied Molecular Genetics of Filamentous Fungi, Blackie Academic and Professional, Chapman and Hall, Londres.

#### Células huésped -

**[0098]** Se proporciona también una célula huésped que comprende un constructo de ADN de fusión según la invención. En modos de realización determinados, las células huésped pueden ser una célula huésped fúngica filamentosa. En algunos modos de realización, las células pueden ser células fúngicas filamentosas de una cepa que tiene un historial de uso para la producción de proteínas que tienen un status GRAS, es decir, un *Generally Recognized as Safe* (reconocido generalmente como seguro) por la FDA.

**[0099]** En modos de realización específicos, la célula fúngica referida puede ser una célula de las siguientes especies: *Trichoderma*, (p.ej., *Trichoderma reesei* (clasificada anteriormente como *T. longibrachiatum* y actualmente también conocida como *Hypocrea jecorina*), *Trichoderma viride*, *Trichoderma koningii*, y *Trichoderma harzianum*); *Penicillium* sp.: *Humicola* sp. (p.ej., *Humicola insolens* y *Humicola grisea*); *Chrysosporium* sp. (p.ej., *C. lucknowense*); *Gliocladium* sp.; *Aspergillus* sp. (p.ej., *Aspergillus oryzae*, *Aspergillus niger*, *Aspergillus nidulans*, *Aspergillus kawachi*, *Aspergillus aculeatus*, *Aspergillus japonicus*, *Aspergillus sojae*, y *Aspergillus awamori*), *Fusarium* sp.; *Mucor* sp.; *Neurospora* sp.; *Hypocrea* sp.; o *Emericella* sp. (Véase también, Innis et al., (1985) Sci. 228:21-26), entre otros. En algunos modos de realización, las células fúngicas referidas pueden ser cepas de *Aspergillus oryzae*, ATCC 11490, *Aspergillus niger* que incluyen ATCC 22342, ATCC 44733, ATCC 14331, NRRL 3112, y cepas derivadas de las mismas. En algunos modos de realización, las células fúngicas referidas pueden ser cepas de *Trichoderma* que incluyen equivalentes funcionales de RL-P37 (Sheir-Neiss et al. (1984) Appl. Microbiol. Biotechnology 20:46 -53). Las cepas huésped de *Trichoderma* útiles incluyen; NRRL 15709, ATCC 13631, ATCC 26921 (QM 9414) ATCC 32098, ATCC 32086 y ATCC 56765 (RUTC-30).

**[0100]** En algunos modos de realización, una célula huésped puede ser una donde los genes nativos se hayan eliminado o inactivado. En algunos modos de realización, las células huésped preferidas presentan genes de proteasa inactivados (p.ej., aspartil proteasa) y se hace referencia a Berka et al. (1990) Gene 86:153-162 y USP 6.509.171. En algunos modos de realización, las células huésped preferidas tienen genes de celulasa inactivados (p.ej., *cbh1*, *cbh2* y *egl1*, y *egl2*) y se hace referencia a la cepa con delección cuádruple de *T. reesei* revelada en WO 05/001036.

**[0101]** El constructo de ADN de fusión descrito arriba puede estar presente en el genoma nuclear de la célula huésped o puede estar presente en un plásmido que replica en la célula huésped, por ejemplo.

#### Transformación

**[0102]** La introducción de un constructo de ADN o vector en una célula huésped incluye técnicas como la transformación; electroporación, microinyección nuclear; transducción; transfección, (p.ej., transfección mediada por DEAE-Dextrina y mediada por lipofección); incubación con precipitado de ADN y fosfato cálcico; bombardeo a alta velocidad con microproyectiles recubiertos con ADN; y fusión de protoplastos. Las técnicas de transformación generales son conocidas en el sector (Véase, p.ej., Ausubel et al., (1987), *supra*, capítulo 9; y Sambrook (1989) *supra*, y Campbell et al., (1989) Curr. Genet. 16:53-56). Se hace referencia también a WO 05/001036; USP 6.022.725; USP 6.103.490; USP 6.268.328; [y solicitudes de patente estadounidenses publicadas 20060041113, 20060040353, 20060040353 y 20050208623].

**[0103]** La expresión de proteínas introducidas de manera recombinante en *Trichoderma* se describe en USP 6.022.725; USP 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 and Berka, Marcel Dekker Inc., NY (1992) págs. 129 - 148). Se hace referencia también a Cao et al., (2000) Protein Sci. 9:991 - 1001; Yelton et al., (1984) Proc. Natl. Acad. Sci. 81:1470 - 1471; USP 6,590,078; y Berka, et al., (1991) en: Applications of Enzyme Biotechnology, Eds. Kelly and Baldwin, Plenum Press, NY) para la transformación de cepas de *Aspergillus*.

**[0104]** En un modo de realización, la preparación de *Trichoderma sp.* para la transformación implica la preparación de protoplastos a partir de micelos fúngicos. (Véase, Penttila et al., (1987) Gene 61:155 - 164). En algunos modos de realización, los micelios se obtienen a partir de esporas vegetativas germinadas.

**[0105]** Generalmente, las células se cultivan en un medio estándar que contiene nutrientes y sales fisiológicas (véase, p.ej., Pourquie, J. et al., BIOCHEMISTRY AND GENETICS OF CELLULOSE DEGRADATION, eds. Aubert, J. P. et al., Academic Press, págs. 71-86, 1988 y Ilmen, M. et al., (1997) Appl. Environ. Microbiol. 63:1298-1306). También son útiles en la presente invención los medios preparados comunes en el mercado (p.ej., caldo de extracto de levadura-malta (YM), caldo Luria Bertani (LB) y caldo Sabouraud dextrosa (SD)). Las condiciones de cultivo preferidas para un hongo filamentoso dado son conocidas en la técnica y pueden encontrarse en la literatura científica y/o a partir de la fuente de los hongos como la American Type Culture Collection (ATCC) y Fungal Genetics Stock Center.

**[0106]** En algunos modos de realización, cuando una inmunoglobulina es la proteína deseada, las células que expresan inmunoglobulina se cultivarán en las condiciones empleadas normalmente para cultivar la línea celular parental. Generalmente, las células se cultivarán en un medio estándar que contiene nutrientes y sales fisiológicas como las descritas por Ilmen et al., (1997) supra. Las condiciones de cultivo también serán estándares (p.ej., incubación a 25 - 30°C en matraces de agitación sobre un agitador rotatorio) hasta que se alcanzan niveles deseados de expresión de inmunoglobulina.

#### **MÉTODOS DE PRODUCCIÓN DE PROTEÍNAS**

**[0107]** Los métodos de producción de una proteína deseada en una célula fúngica filamentosa también se encuentran abarcados por la invención. En algunos modos de realización, estos métodos incluyen, obtener una célula huésped filamentosa que comprende un constructo de ADN de fusión o vector según la invención y cultivar la célula huésped filamentosa en condiciones adecuadas que permitan la expresión y secreción de la proteína deseada. Aunque un cultivo de células huésped (es decir, una composición que contiene las células huésped mencionadas y medio de cultivo) puede contener la proteína secretada del polipéptido de fusión descrito arriba, en algunos modos de realización, la proteína deseada se recupera del medio de cultivo. En algunos modos de realización, la proteína deseada es purificada. La proteína puede recuperarse a partir del medio de cultivo mediante cualquier método conveniente.

**[0108]** En algunos modos de realización, una célula fúngica en cuestión puede cultivarse en condiciones de fermentación continua o por lotes. Una fermentación por lotes clásica es un sistema cerrado, donde la composición del medio se establece al comienzo de la fermentación y no se somete a alteraciones artificiales durante la fermentación. Por tanto, al comienzo de la fermentación el medio es inoculado con el organismo u organismos deseados. En este método, se permite que ocurra la fermentación sin la adición de componentes al sistema. Normalmente, una fermentación por lotes se califica como un "lote" con respecto a la adición de la fuente de carbono y a menudo se realizan intentos de controlar factores como el pH y concentración de oxígeno. Las composiciones de biomasa y metabolitos del sistema de lotes cambian constantemente hasta el momento en el que para la fermentación. En los cultivos por lotes, las células progresan a través de una fase de latencia estática hasta una fase logarítmica de alto crecimiento y finalmente hasta una fase estacionaria donde el índice de crecimiento es reducido o interrumpido. Si no se tratan, las células en la fase estacionaria finalmente morirán. En general, las células en fase logarítmica son responsables de la mayoría de la producción del producto final.

**[0109]** Una variación en el sistema de lotes estándar es el sistema de "fermentación por lote alimentado", que también es útil con la presente invención. En esta variación de un sistema por lotes típico, el sustrato se añade en incrementos a medida que avanza la fermentación. Los sistemas por lote alimentado son útiles cuando la presión catabólica es apta para inhibir el metabolismo de las células y cuando es deseable tener cantidades limitadas de sustrato en el medio. La medición de la concentración de sustrato real en los sistemas por lote alimentado es difícil y, por tanto, se estima basándose en los cambios de factores medibles como pH, oxígeno disuelto y la presión parcial de gases residuales como CO<sub>2</sub>. La fermentación por lotes y por lote alimentado son comunes y conocidas en la técnica.

**[0110]** La fermentación continua es un sistema abierto donde un medio de fermentación definido se añade continuamente a un biorreactor y se extrae una cantidad equivalente de medio condicionado simultáneamente para su procesamiento. La fermentación continua generalmente mantiene los cultivos a una densidad elevada constante donde las células se encuentran principalmente en un crecimiento en fase logarítmica.

**[0111]** La fermentación continua permite la modulación de un factor o cualquier número de factores que afecten al crecimiento celular y/o concentración de producto final. Por ejemplo, en un modo de realización, un nutriente limitante como la fuente de carbono o la fuente de nitrógeno se mantiene a un índice fijo y se permite moderar el resto de parámetros. En otros sistemas, pueden alterarse de forma continua una variedad de factores que afectan al crecimiento mientras que la concentración celular, medida por la turbidez del medio, se mantiene constante. Los sistemas continuos se esfuerzan por mantener las condiciones de crecimiento estables. Por tanto, la pérdida celular debido a la extracción de medio debe equilibrarse con el índice de crecimiento celular en la fermentación. Se conocen métodos de modulación de nutrientes y factores de crecimiento para procesos de fermentación continuos así como las técnicas para maximizar el índice de formación de producto.

## 10 **Expresión y secreción -**

**[0112]** La producción de una proteína deseada en una célula fúngica filamentosa que comprende un constructo de ADN de fusión que codifica un polipéptido de fusión resulta en la secreción de la proteína deseada del polipéptido de fusión. Durante el proceso de secreción en hongos, pueden unirse cadenas de azúcar a una proteína que va a ser secretada para producir una proteína glicosilada. En la presente invención, la producción de la proteína deseada, (p.ej., un anticuerpo), puede incluir proteína glicosilada o no glicosilada.

**[0113]** En algunos modos de realización, la proteína secretada del referido polipéptido de fusión generalmente está presente en el medio de cultivo de la célula fúngica filamentosa en una cantidad que es mayor a la cantidad de la proteína secretada deseada de un polipéptido de fusión equivalente que carece de la presecuencia de sitio KEX2, producido por una célula fúngica filamentosa equivalente (es decir, el mismo tipo celular, cultivado en las mismas condiciones). Un cultivo de las células en cuestión que producen una proteína deseada a partir de un polipéptido de fusión según la invención puede contener más del 5%, más del 10%, más del 20%, más del 40%, más del 60%, más del 80%, más del 100%, más del 150%, más del 200%, más del 300%, más del 500% y más del 1000% de proteína deseada en el medio de cultivo, en comparación con un cultivo celular equivalente que expresa una proteína por lo demás equivalente que no contiene una presecuencia de sitio KEX2 según queda abarcado por la invención.

**[0114]** En algunos modos de realización, el nivel de expresión y secreción de una proteína deseada (p.ej., un anticuerpo de longitud completa) será mayor de 0,5 g/L. De manera rutinaria, puede recuperarse más de 1,0 g/L de la proteína deseada a partir del medio de cultivo. Pueden alcanzarse niveles reproducibles de más de 1,5, 2,0 y 3,0 g/L. En algunos modos de realización, el nivel de expresión y secreción de la proteína deseada será mayor de 10 g/L e incluso mayor de 20 g/L.

**[0115]** En algunos modos de realización de la invención, la escisión de la proteína deseada del polipéptido de fusión recombinante será mayor que la escisión de la misma proteína deseada de un polipéptido de fusión recombinante equivalente que carece de la presecuencia de sitio KEX2. En algunos modos de realización, la presecuencia de sitio KEX2 puede resultar en una proteína de fusión que se escinde con una eficiencia de al menos un 80%, al menos un 85%, al menos un 90%, al menos un 95%, al menos un 97%, al menos un 98%, al menos un 99% o al menos un 100%, donde la eficiencia del 100% resulta en una proteína de secreción deseada totalmente escindida del polipéptido de fusión.

**[0116]** En determinados modos de realización, la eficiencia de escisión de proteínas puede calcularse determinando la cantidad de escisión que se ha producido, p.ej., determinado la cantidad de proteína escindida frente a la cantidad de proteína no escindida. En un modo de realización, la cantidad de escisión de proteína puede calcularse determinando la proporción de la cantidad de proteína escindida en el medio de cultivo con respecto a la cantidad de proteína de fusión no escindida en el medio de cultivo por volumen del cultivo celular.

**[0117]** Un polipéptido de fusión que contiene una presecuencia de sitio KEX2 o una presecuencia de sitio KEX2 optimizada, en algunos modos de realización, puede resultar en un polipéptido de fusión que es escindido a una eficiencia de al menos un 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97%, 98%, 99% o 100%, donde el 100% de eficiencia es una proteína deseada completamente escindida.

**[0118]** En otros modos de realización, la eficiencia de secreción de un polipéptido de fusión en cuestión puede calcularse determinando la cantidad de parte secretada de ese polipéptido de fusión en el medio de crecimiento de una célula que secreta esa proteína. Esta determinación puede ser cuantitativa, cualitativa, relativa o absoluta. En un modo de realización, la cantidad de proteína secretada en el medio de cultivo de una célula que secreta una fusión en cuestión puede ser al menos un 10%, al menos un 30%, al menos un 50%, al menos un 70%, al menos un 90%, al menos dos veces, al menos cinco veces o al menos diez veces superior a la cantidad de la proteína secretada por una célula que produce un polipéptido de fusión equivalente que no contiene una presecuencia de KEX2 optimizada.

**[0119]** En algunos modos de realización, el aumento en la secreción y/o escisión puede medirse frente a una región KEX2 estándar definida como GGGB<sub>1</sub>B<sub>2</sub>, donde B<sub>1</sub>B<sub>2</sub> es KK, KR, RK o RR y preferiblemente KR. En un

modo de realización, la cantidad de proteína secretada o proteína deseada en el medio de cultivo de una célula que secreta una fusión en cuestión puede ser al menos un 10%, al menos un 30%, al menos un 50%, al menos un 70%, al menos un 80%, al menos un 90%, al menos un 100%, al menos 2 veces, al menos 3 veces, al menos 5 veces y al menos 10 veces mayor a la cantidad de la proteína secretada o proteína deseada secretada por un polipéptido de fusión equivalente en un huésped equivalente en las mismas condiciones esencialmente.

#### **MÉTODOS DE DETECCIÓN -**

**[0120]** También se proporcionan métodos de detección para identificar presecuencias de sitio KEX2 optimizadas. Estos métodos pueden incluir: a) alterar una presecuencia de sitio KEX2 de un polipéptido de fusión parental para producir un polipéptido de prueba y b) evaluar la secreción del polipéptido de fusión de prueba por una célula fúngica filamentosa. En determinados modos de realización, la secreción y/o escisión de la proteína deseada a partir de polipéptidos de fusión de prueba se compara con la secreción y/o escisión de la proteína de fusión parental. En modos de realización concretos, el método incluye evaluar la cantidad de una proteína secretada del polipéptido de fusión en un medio de cultivo en relación con la cantidad de una parte secretada del polipéptido de fusión en una célula, por volumen de cultivo, o evaluar la cantidad de una proteína secretada de un polipéptido de fusión recombinante en un medio de cultivo. En otros modos de realización, el método incluye evaluar la cantidad de una proteína secretada (proteína deseada) liberada o escindida de un polipéptido de fusión en un medio de cultivo en relación con la cantidad de la proteína secretada que permanece en la forma del polipéptido de fusión (p.ej., unida a la proteína portadora).

**[0121]** En estos ensayos de detección, la proteína de fusión parental tiene una secuencia de aminoácidos que se ilustra de manera esquemática en la Fig. 1, donde X es cualquier aminoácido. En determinados modos de realización, la proteína de fusión parental y la proteína de fusión de prueba pueden ser idénticas excepto por la presecuencia de sitio KEX2. Una proteína de fusión recombinante parental y una proteína de fusión recombinante de prueba pueden diferir en uno, dos, tres o cuatro aminoácidos en la presecuencia de sitio KEX2. Una alteración puede ser una sustitución, inserción o delección de aminoácido, y si hay dos o tres alteraciones, las alteraciones pueden darse en aminoácidos contiguos, aminoácidos no contiguos o una combinación de aminoácidos contiguos y no contiguos.

**[0122]** En un modo de realización, la presecuencia de sitio KEX2 de un polipéptido de fusión parental puede alterarse para producir una pluralidad de polipéptidos de fusión de prueba diferentes que contienen cada uno presecuencias de sitio KEX2 diferentes y evaluar después la secreción y/o escisión de los polipéptidos de fusión de prueba y los polipéptidos de fusión parentales mediante una célula fúngica filamentosa.

**[0123]** Estos métodos pueden llevarse a cabo usando protocolos que son conocidos generalmente (véase, p.ej., Ward et al (1990) Bio/Technology 8:435-440 y Spencer (1998) Eur. J. Biochem 258: 107-112, entre otros), en los que se introduce un vector en una célula, se cultiva la célula, y se somete a ensayo el cultivo celular para detectar la presencia de la proteína celular. En un modo de realización, se altera un ácido nucleico recombinante que codifica una fusión parental (la estructura de la cual se muestra en la Fig. 1) para producir un ácido nucleico que codifica un polipéptido de prueba, y se usan los dos ácidos nucleicos para transformar células fúngicas filamentosas idénticas (que pueden ser cualquiera de las células huésped arriba enumeradas). Las dos líneas celulares se cultivan en condiciones idénticas, y se evalúa la eficiencia de secreción y/o escisión de la parte secretada de la proteína. La secuencia señal, proteína secretada, el sitio KEX2 y la presecuencia de sitio KEX2 de la proteína de fusión parental pueden ser cualquier secuencia señal, proteína secretada, sitio KEX2 o presecuencia de sitio KEX2 conocidos, incluyendo aquellos enumerados arriba.

**[0124]** Como se ha apuntado arriba, la eficiencia de la secreción o escisión de proteína puede evaluarse de muchas formas diferentes, por ejemplo, comparando las cantidades normalizadas o absolutas de parte secretada en el medio de cultivo entre los diferentes cultivos, o comparando la cantidad de la parte secretada de la proteína con la cantidad de la parte no secretada de la proteína. Esta evaluación puede ser cuantitativa, cualitativa, relativa o absoluta, por ejemplo.

**[0125]** Una presecuencia de sitio KEX2 optimizada puede identificarse probando una pluralidad de diferentes proteínas de fusión de prueba según los métodos anteriores; y determinando cuales de las diferentes proteínas de fusión de prueba es secretada y/o escindida más eficientemente; donde la presecuencia de sitio KEX2 optimizada es la presecuencia de sitio KEX2 de la proteína de fusión recombinante de prueba que es secretada más eficientemente.

**[0126]** Un cultivo de células que contiene al menos un 10%, al menos un 20%, al menos un 30%, al menos un 50%, al menos un 70% y al menos un 95% más de proteína deseada o secretada que un cultivo de control indica que la presecuencia de sitio KEX2 aumenta la secreción y/o escisión de proteína a partir de esas células.

#### **EJEMPLOS**



[0127] Los siguientes ejemplos se plantean para proporcionar a aquellos con experiencia en la técnica una divulgación y descripción completa de cómo realizar y utilizar la presente invención, y no pretenden limitar el alcance de lo que los inventores consideran su invención ni pretenden representar que los experimentos a continuación son todos o los únicos experimentos llevados a cabo. Se han realizado esfuerzos por garantizar la exactitud con respecto a los números utilizados (p.ej., cantidades, temperatura, etc.) pero deben tenerse en cuenta algunos errores y desviaciones experimentales. A menos que se indique lo contrario, las partes son partes en peso, el peso molecular es el peso molecular medio en peso, la temperatura se da en grados centígrados (°C), la presión es atmosférica o cercana a la atmosférica y se aplican las siguientes abreviaturas, M (Molar);  $\mu$ M (micromolar); N (Normal); mol (moles); mmol (milimoles);  $\mu$ mol (micromoles); nmol (nanomoles); g (gramos); mg (miligramos) kg (kilogramos);  $\mu$ g (microgramos); L (litros); ml (mililitros); h (horas); min (minutos); PAGE (electroforesis en gel de poliacrilamida); kDa (kilodaltons); y pb (pares de bases). Se usan los siguientes ensayos y métodos en los ejemplos que se proporcionan a continuación:

#### A. Construcción del vector pTrex4:

[0128] Se clonó ADN sintético en un vector de expresión de *Trichoderma* (pTrex4) para generar plásmidos de expresión apropiados para su uso en los ejemplos descritos a continuación.

[0129] pTrex4 es una versión modificada de pTrex2 y derivada de un vector de expresión pTrex3g. La construcción de pTrex3g se describe en detalle en el Ejemplo 6 de WO 05/001036. En resumen, el pTrex3g se basa en el vector de *E.coli* pSL1180 (Pharmacia, Inc., Piscataway, NJ) que es un vector basado en el plásmido pUC 118 con un sitio de clonación múltiple extendido que contiene 64 secuencias de reconocimiento de enzimas de restricción hexaméricas. Se diseñó como un vector de destino de Gateway (Hartley, J.L. et al., (2000) Genome Research 10:1788 - 1795) para permitir la inserción usando Gateway Technology (Invitrogen) de cualquier marco de lectura abierto deseado entre las regiones de promotor y terminador del gen *cbh1* de *T. reesei*.

[0130] Los detalles del pTrex4-her2 cadena ligera DNA2.0 son los siguientes (Fig. 2 y Fig. 3): El plásmido es de 10885 kb de tamaño (SEQ ID NO: 103). En la región poliligadora de pSL 1180 se encuentran insertados los siguientes segmentos de ADN:

Un segmento de 2,2 pb de ADN de la región promotora del *cbh1* de *T. reesei*;  
la secuencia de ADN de la secuencia señal de *cbh1* de *T. reesei* (subrayada); dominio catalítico, enlazador (cursiva) (1570 bases) (SEQ ID NO: 104)

ATGTATCGGAAGTTGGCCGTCATCTCGGCCTTCTTGGCCACAGCTCGTGCTCAGTCGGCCTG  
CACTCTCCAATCGGAGACTCACCCGCCTCTGACATGGCAGAAATGCTCGTCTGGTGGCACTT  
GCACTCAACAGACAGGCTCCGTGGTCATCGACGCCAACTGGCGCTGGACTCACGCTACGAA  
CAGCAGCACGAACTGCTACGATGGCAACA CTTGGAGCTCGACCCTATGTCCTGACAACGAG  
ACCTGCGCGAAGA AACTGCTGTCTGGACGGTGCCGCCTACGCGTCCACGTACGGAGTTACCA  
CGAGCGGTAACAGCCTCTCCATTGGCTTTGTACCCAGTCTGCGCAGAAGAACGTTGGCGCT  
CGCCTTTACCTTATGGCGAGCGACACGACCTACCAGGAATTCACCCTGCTTGGCAACGAGTT  
CTCTTTTCGATGTTGATGTTTCGCAGCTGCCGTAAGTGACTTACCATGAACCCCTGACGTATCT

TCTTGTGGGCTCCCAGCTGACTGGCCAATTTAAGGTGCGGCTTGAACGGAGCTCTCTACTTC  
 GTGTCCATGGACGCGGATGGTGGCGTGAGCAAGTATCCCACCAACACCGCTGGCGCCAAGT  
 ACGGCACGGGGTACTGTGACAGCCAGTGTCCCCGCGATCTGAAGTTCATCAATGGCCAGGC  
 CAACGTTGAGGGCTGGGAGCCGTCATCCAACAACGCAAACACGGGCATTGGAGGACACGG  
 AAGCTGCTGCTCTGAGATGGATATCTGGGAGGCCAACTCCA TCTCCGAGGCTCTTACCCCC  
 ACCCTTGCACGACTGTGCGCCAGGAGATCTGCGAGGGTGTGATGGGTGCGGCGGAACTTACTC  
 CGATAACAGATATGGCGGCACTTGCATCCCGATGGCTGCGACTGGAACCCATACCGCCTG  
 GGCAACACCAGCTTCTACGGCCCTGGCTCAAGCTTTACCCTCGATACCACCAAGAAATTGAC  
 CGTTGTCAACCAGTTCGAGACGTGCGGTGCCATCAACCGATACTATGTCCAGAATGGCGTCA  
 CTTTCCAGCAGCCCAACGCCGAGCTTGGTAGTTACTCTGGCAACGAGCTCAACGATGATTAC  
 TGCACAGCTGAGGAGGCAGAATTCGGCGGATCCTCTTTCTCAGACAAGGGCGGCCTGACTC  
 AGTTCAAGAAGGCTACCTCTGGCGGCATGGTCTGGTTCATGAGTCTGTGGGATGATGTGAGT  
 TTGATGGACAAACATGCGCGTTGACAAAGAGTCAAGCAGCTGACTGAGATGTTACAGTACT  
 ACGCCAACATGCTGTGGCTGGACTCCACCTACCCGACAAACGAGACCTCCTCCACACCCCGT  
 GCCGTGCGCGGAAGCTGCTCCACCAGCTCCGGTGTCCCTGCTCAGGTCGAATCTCAGTCTCC  
 CAACGCCAAGGTCACCTTCTCCAACATCAAGTTCGGACCCATTGGCAGCACCCGGCAACCCTA  
 GCGGCGGCAACCCTCCCGCGGAAACCCGCCTGGCACCACCACCACCCGCGCCAGCCACTA  
 CCACTGGAAGCTCTCCCGGACCTACTAGT

[0131] La secuencia de aminoácidos de la secuencia señal de *cbh1* de *T. reesei*; dominio catalítico; enlazador (480 aminoácidos) se representa a continuación (SEQ ID NO: 105)

5

MYRKLAVISAFLATARAQSACTLQSEHPPLTWQKCSSGGTCTQQTGSVVIDANWRWT  
 HATNSSTNCYDGNTWSSTLCPDNETCAKNCCLDGAAYASTYGVTTSGNSLSIGFVTQSAQKNV  
 GARLYLMASDTTYQEFTLLGNEFSFDVDVSQLPCGLNGALYFVSMDADGGVSKYPTNTAGAKY  
 GTGYCDSQCPRDLKFINGQANVEGWEPSSNNANTGIGGHGSCSEMDIWEANSISEALTPHPCTT  
 VGQEICEGDGCGGTYSNRYGGTCDPDGCDWNPYRLGNTSFYGPSSFTLDTTKKLTVVVTQFE  
 TSGAINRYYVQNGVTFQQPNAELGSYSGNELNDDYCTAEAEFGGSSFSKGLTQFKKATSGG  
 MVLVMSLWDDYYANMLWLDSTYPTNETSSTPGA VRGSCSTSSGVPAQVESQSPNAKVTFNSNIK  
 FGPIGSTGNPSGGNPPGGNPPGTTTTRRPATTTGSSPGPTS

[0132] El plásmido también contiene un terminador de *cbh1*, un marcador seleccionable de *amdS* de *A. nidulans* y nucleótidos que codifican la cadena ligera del anticuerpo.

10 **B. Transformación biolística de *T. reesei*:**

[133] En todos los ejemplos a continuación, la transformación se llevó a cabo en un derivado de la cepa de *T. reesei* (WO 05/001036) con delección cuádruple ( $\Delta cbh1$ ,  $\Delta cbh2$ ,  $\Delta eg1$ , y  $\Delta eg2$ ) derivada de manera original de RL-P37 (Sheir-Neiss et al., (1984) Appl. Microbiol. Biotechnol. 20:46 - 53; USP 4.797.361) con el vector pTrex4 apropiado usando el protocolo explicado a continuación.

15 [0134] Se preparó una suspensión de esporas (aproximadamente  $5 \times 10^8$  esporas/ml) de la cepa *Trichoderma*. Se extendieron 100ul - 200ul de suspensión de esporas en el centro de las placas de medio mínimo con acetamida. El medio mínimo con acetamida tenía la siguiente composición: 0,6 g/L acetamida; 1,68 g/L CsCl; 20 g/L glucosa;

20 g/L KH<sub>2</sub>PO<sub>4</sub>; 0,6 g/L CaCl<sub>2</sub>.2H<sub>2</sub>O; 1 ml/L 1000X solución de oligoelementos; 20 g/L agar Noble; pH 5,5. La solución de oligoelementos 1000X contenía 5,0 g/l FeSO<sub>4</sub>.7H<sub>2</sub>O, 1,6 g/l MnSO<sub>4</sub>.H<sub>2</sub>O, 1,4 g/l ZnSO<sub>4</sub>.7H<sub>2</sub>O y 1,0 g/l CoCl<sub>2</sub>.6H<sub>2</sub>O. Se dejó secar la suspensión de esporas sobre la superficie del medio mínimo con acetamida.

5 [0135] La transformación de la cepa *Trichoderma* por el método de transformación biolística se logró usando un sistema de liberación de partículas Biolistic® PDS-1000/He de Bio-Rad (Hercules, CA) siguiendo las instrucciones del fabricante (véase, WO 05/001036 y US 2006/0003408).

**C. Transformación de *Aspergillus* -**

10 [0136] El protocolo de transformación de *Aspergillus* era una modificación del método Campbell (Campbell et al. (1989). Curr. Genet. 16:53-56). También se revelan detalles del método de transformación de *Aspergillus niger* en WO 03089614 y la publicación de patente estadounidense 20050153399. Se sometieron a ensayo los transformantes para la producción de proteínas en gel SDS y Western blot para seleccionar los transformantes en función de la cantidad de proteína producida.

**D. Fermentación de cepas *T. reesei* y *Aspergillus niger* transformadas con el vector de expresión:**

15 [0137] En general, se siguió el protocolo de fermentación descrito en Foreman et al., (2003) J. Biol. Chem 278:31988 -31997.

E. El **Medio Proflo** contiene: 30 g/L α-lactosa; 6,5 g/L (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>; 2g/L KH<sub>2</sub>PO<sub>4</sub>; 0,3g/L MgSO<sub>4</sub>.7H<sub>2</sub>O; 0,2g/L CaCl<sub>2</sub>; 1ml/L solución salina de oligoelementos 1000X; 2ml/ L 10% Tween 80; 22,25g/L harina de semilla de algodón Proflo (Traders Protein, Memphis, TN); 0,72g/L CaCO<sub>3</sub>.

20 F. El **Medio definido** contiene: 5g/L (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>; 33g/L tampón PIPPS; 9g/L casaminoácidos; 4,5g/L KH<sub>2</sub>PO<sub>4</sub>; 1g/L CaCl<sub>2</sub>; 1g/L MgSO<sub>4</sub>.7H<sub>2</sub>O; 5ml/L antiespumante Mazu DF60-P (Mazur Chemicals, Gurnee, IL); 1 ml/L solución de oligoelementos 1000X. Tras colocar en autoclave, se añadieron 40 ml de 40% lactosa.

G. La **solución de oligoelementos 1000X** contiene: 5,0 g/l FeSO<sub>4</sub>.7H<sub>2</sub>O, 1,6 g/l MnSO<sub>4</sub>.H<sub>2</sub>O, 1,4 g/l ZnSO<sub>4</sub>.7H<sub>2</sub>O y 1,0 g/l CoCl<sub>2</sub>.6H<sub>2</sub>O

H. El análisis de proteínas se realizó mediante análisis de Western blot y gel SDS estándar.

25

**Ejemplo 1**

**Construcción de una cepa de expresión de la cadena ligera de trastuzumab que contiene un sitio de escisión KEX2 KRGGG (SEQ ID NO: 2)**

30 [0138] El ADN (SEQ ID NO: 1) que codifica la cadena ligera de trastuzumab según la secuencia de aminoácidos publicada de anticuerpo 4D5-8 (Carter et al, Proc. Natl. Acad. Sci. 1992 89: 4285-4289) fue sintetizado por DNA2.0 Inc. (1455 Adams Drive, Menlo Park, CA94025).

ACTAGTAAACGCGGTGGCGGTGATATTCAAATGACACAATCTCCTTCTTCTCTGTCA  
 GCCTCAGTGGGCGACCGTGTGACGATTACTTGCCGCGCCTCTCAGGACGTTAACACT  
 GCCGTTCGCATGGTACCAGCAGAAGCCAGGCAAGGCGCCCAAGCTTCTGATTTACAG  
 CGTTTCGTTTCTGTACTCTGGCGTGCCATCCCGCTTCTCTGGCAGCCGAAGCGGCAC  
 GGATTTACCCTGACCATTTTCGTCCTGCAGCCCGAGGATTTGCCACGTATTACTG  
 CCAGCAGCACTACCACTCCACCCACCTTTGGCCAAGGAACGAGAGTCGAAATCA  
 CTCGCACGGTTCGCTGCCCTTCAGTCTTCATCTTCCCCCCCAGCGACGAACAGCTGA  
 AGTCTGGTACGGCCAGCGTCGTTTGCTTGCTTAATAACTTCTATCCGCGAGAGGCGA  
 AGGTCCAATGGAAGGTTGATAACGTTCTGCAGTCCGGCAATTTCGAGGAGAGCGTG  
 ACCGAGCAGGATTCAAAGGATAGCACCTACTCACTCAGCAGCACCCCTGACGTTGTC  
 CAAGGCCGATTACGAGAAGCATAAGTTGTATGCATGCGAGGTCACCCACCAGGGAC  
 TGTC AAGCCCAGTTACCAAGTCGTTCAATCGAGGCGAGTGCTAAGGCGCGCC (SEQ  
 ID NO: 1).

**[0139]** La cadena ligera codificada por el ADN contiene un sitio de escisión KEX2 KRGGG (SEQ ID NO:2) en su extremo N-terminal. Los sitios de restricción SpeI y AclI se incluyeron para fines de clonación. El ADN sintético se clonó en vector de expresión de *Trichoderma* (pTrex4) para generar un plásmido de expresión llamado pTrex4-her2 cadena ligera DNA2.0 (Fig. 2). El plásmido resultante codifica una proteína de fusión que contiene una región enlazadora/núcleo de CBHI de *Trichoderma* y la cadena ligera de anticuerpo, separados por un sitio KEX2. El plásmido fue digerido con la enzima de restricción XbaI y transformado biolísticamente en una cepa de *Trichoderma reesei* derivada de la cepa con delección cuádruple descrita en WO 05001036, ejemplo 5). Se obtuvieron más de 20 transformantes y se transfirieron a nuevas placas. Se seleccionaron veinte transformantes estables para su cultivo en medio Proflo durante 2 días a 30°C. Se transfirieron 5 ml del cultivo de dos días de antigüedad de Proflo a 50 ml de medio definido. Los cultivos se desarrollaron durante 5 días a 28°C. Los caldos de cultivo fueron centrifugados y se usaron los sobrenadantes para el análisis de proteínas. Los datos del Western blot (Fig. 4) indicaron que más del 90% de la proteína de fusión estaba escindida en la mejor cepa productora de cadena ligera (transformante 1010-18, variante KRGGG). Sin embargo, GGG continúa en el extremo N-terminal de la cadena ligera de anticuerpo escindido, lo que no es deseable. También se detectó una banda de aproximadamente 50 kd en el Western blot, lo que puede resultar de la dimerización de dos moléculas de cadena ligera.

### Ejemplo 2

#### Construcción de una cepa de expresión de la cadena ligera de trastuzumab que contiene el sitio de escisión KEX2 GGGKR (SEQ ID NO:5)

**[0140]** Se sintetizaron dos cebadores (GGACTAGTGGTGGCGGTAAACGCGATATTCAAATGACACAATCT C; SEQ ID NO:3 y AAGGCGCGCCTTAGCACTCGCCTCGATTG; SEQ ID NO:4) mediante Invitrogen (1600 Faraday Avenue, Carlsbad, CA 92008) y se usaron para amplificar ADN de la cadena ligera de trastuzumab.

**[0141]** El fragmento de PCR resultante codifica la cadena ligera de anticuerpo que contiene un sitio KEX2 de secuencia GGGKR (SEQ ID NO:5) en su extremo N-terminal. El fragmento de PCR fue digerido con enzimas de restricción SpeI y AclI y clonado al vector de expresión pTrex4 para generar un plásmido llamado pTrex4-GGGKR-her2 DNA2.0. Se analizó la fidelidad del fragmento de PCR mediante secuenciación de ADN. El plásmido fue digerido con la enzima de restricción XbaI y transformado biolísticamente usando técnicas estándares en la cepa *T. reesei* descrita arriba. Se obtuvieron más de 20 transformantes y se transfirieron a nuevas placas. Se seleccionaron un total de 21 transformantes estables para su cultivo en medio Proflo durante 2 días a 30°C. Se transfirieron 5 ml de cultivo de 2 días de antigüedad de Proflo a 50 ml de medio definido. Los cultivos se mantuvieron durante 5 días a 28°C. Los caldos de cultivo fueron centrifugados y se usaron los sobrenadantes. El análisis Western blot indicó que más del 95% de la proteína del transformante 1010-B5 (variante GGGKR) y el transformante 1010-B6 (variante GGGKR) era una proteína de fusión no escindida (Fig. 4). También se detectó una banda de aproximadamente 150 kd en el Western blot. Esto puede resultar de la dimerización de las dos moléculas de fusión de cadena ligera-núcleo de CBHI.

### Ejemplo 3

#### Construcción de una cepa de expresión de la cadena ligera de trastuzumab que contiene un sitio de escisión KEX2 GGGKRGGG (SEQ ID NO: 7)

**[0142]** Se sintetizaron dos oligos, GGACTAGTGGCGGTGGCAAACGCGGTGGCGGTGATATTC (SEQ ID NO. 6) y AAGGCGCGCCTTAGCACTCGCCTCGATTG (SEQ ID NO. 4), mediante Invitrogen y se usaron para amplificar ADN de la cadena ligera. El fragmento de PCR resultante codifica la cadena ligera y la secuencia GGGKRGGG (SEQ ID NO:7) para la escisión de KEX2. El fragmento de PCR fue digerido con enzimas de restricción SpeI y AclI y clonado al vector de expresión pTrex4 para generar un plásmido llamado pTrex4-GGGKRGGG-her2 cadena ligera DNA2.0. Se analizó la fidelidad del fragmento de PCR mediante secuenciación de ADN. El plásmido fue digerido con enzima de restricción XbaI y transformado biolísticamente en la cepa *T. reesei* como se describe arriba. Se obtuvieron más de 10 transformantes y se transfirieron a nuevas placas. Se seleccionaron 3 transformantes estables para su cultivo en medio Proflo durante 2 días a 30°C. Se transfirieron 5 ml del cultivo de dos días de antigüedad de Proflo a 50 ml de medio definido. Los cultivos se mantuvieron durante 5 días a 28°C. Los caldos de cultivo fueron centrifugados y se usaron los sobrenadantes para el análisis de proteínas. Los datos del gel Western indicaron que, en transformante 1011-1 (variante GGGKRGGG), más del 90% de la proteína de fusión fue escindida (Fig. 4). Sin embargo, GGG continuó en el extremo N-terminal de la cadena ligera de anticuerpo escindido, lo que no es deseable.

### Ejemplo 4

**[0143]** Construcción de una cepa de expresión de la cadena ligera de trastuzumab que contiene una región KEX2 VAVEKR (SEQ ID NO: 9)

5 **[0143]** Una región KEX 2 VAVEKR (SEQ ID NO: 9) se encuentra de forma natural en la prorrregión de la xilanasas de alto punto isoeléctrico (pI) de *T. reesei*, Xyn 2 (Torronen et al., (1992) *Biotechnol.* 10:1461 - 1465). Para construir un polipéptido de fusión según la invención, se sintetizaron dos oligos, GGACTAGTGTCCCGTTGAGAAACGCGATATTCAAATGACACAAT CTCC (SEQ ID NO. 8) y AAGGCGCGCCTTAGCACTCGCCTCGATTG (SEQ ID NO. 4), mediante Invitrogen y se usaron para amplificar ADN de la cadena ligera.

10 **[0144]** El fragmento de PCR resultante codifica la cadena ligera y la secuencia VAVEKR (SEQ ID NO:9) para la escisión de KEX2. El fragmento de PCR fue digerido con enzimas de restricción SpeI y AclI y clonado al vector de expresión pTrex4 para generar un plásmido llamado pTrex4-VAVE-her2 cadena ligera DNA2.0. Se analizó la fidelidad del fragmento de PCR mediante secuenciación de ADN. El plásmido fue digerido con enzima de restricción XbaI y transformado biolísticamente en la cepa *T.reesei* como se describe arriba. Se obtuvieron más de 20 transformantes y se transfirieron a nuevas placas. Se seleccionaron 6 transformantes estables para su cultivo en medio Proflo durante 2 días a 30°C. Se transfirieron 5 ml del cultivo de dos días de antigüedad de Proflo a 50 ml de medio definido. Los cultivos se mantuvieron durante 5 días a 28°C. Los caldos de cultivo fueron centrifugados y se usaron los sobrenadantes para el análisis de proteínas. Los datos del gel Western indicaron que, en transformante 1012-2 (variante VAVEKR, SEQ ID NO: 9) más del 95% de las proteínas de fusión fueron escindidas (Fig. 4).

#### Ejemplo 5

20 **Construcción de una cepa de expresión de la cadena ligera de trastuzumab que contiene variantes de la región KEX2 VAVEKR (SEQ ID NO: 9)**

**[0145]** Se sintetizó ADN (SEQ ID NO: 10) que codifica la cadena ligera del anticuerpo trastuzumab mediante Genent (Josef- Engert-Strasse 11, 93053 Regensburg, Alemania).

ACTAGTAAGCGGGCGGGCGGGGAGGTCCAGCTCGTTCGAGAGCGGGCGGGCCTCGT  
 CCAGCCCGGGCGGCAGCCTCCGCCTCAGCTGCGCCGCCAGCGGCTTCAACATCAAGG  
 ACACCTACATCCACTGGGTCCGCCAGGCCCGGCAAGGGCCTCGAGTGGGTGCGC  
 CGCATCTACCCACCAACGGCTACACCCGCTACGCCGACAGCGTCAAGGGCCGCTT  
 CACCATCAGCGCCGACACCAGCAAGAACACCGCCTACCTCCAGATGAACAGCCTCC  
 GCGCCGAGGACACCGCCGTCTACTACTGCAGCCGCTGGGGCGGGGACGGCTTCTAC  
 GCCATGGACTACTGGGGCCAGGGCACCTCGTCACGGTCTCCAGCGCCAGCACCAA  
 GGGCCCAAGCGTCTTCCCTCGCCCCAGCAGCAAGAGCACCAGCGGGCGGCACCG  
 CCGCCCTCGGCTGCCTCGTCAAGGACTACTTCCCCGAGCCCGTCACTGTCAGCTGGA  
 ACAGCGGGCGTCTCACCAGCGGGTCCACACCTTCCCCGCCGTCCTCCAGAGCAGC  
 GGCCTCTACAGCCTCAGCAGCGTCGTCACCGTCCCCAGCAGCAGCCTCGGCACCCA  
 GACCTACATCTGCAACGTCAACCACAAGCCCAGCAACACCAAGGTGACAAAGCGCG  
 TCGAGCCCAAGAGCTGCGACAAGACCCACACCTGCCCCCCTGCCCCGCCCCGAG  
 CTGCTCGGCGGGCCCTCCGTCTTCTCTTCCCCCAGCCCAAGGACACCCTCATG  
 ATCAGCCGCACCCCGAGGTACCTGCGTTCGTCGATGTCAGCCACGAGGACCC  
 CGAGGTCAAGTTCAACTGGTACGTGACGGCGTCGAGGTCCACAACGCCAAGACCA  
 AGCCCCGCGAGGAGCAGTACAACAGCACCTACCGCGTCGTCAGCGTCTGACCGTC  
 CTCCACCAGGACTGGCTCAACGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAGGC  
 CCTCCCCGCCCCATCGAAAAGACCATCAG  
 CAAGGCCAAGGGCCAGCCCCGCGAGCCCCAGGTCTACACCCTCCCCCAGCCGCG  
 AGGAGATGACCAAGAACCAGGTCTCCCTCACCTGCCTGGTCAAGGGCTTCTACCC  
 AGCGACATCGCCGTCGAGTGGGAGAGCAACGGCCAGCCCGAGAACAACACTACAAGA  
 CCACCCCCCGTCTCTGACAGCGACGGCAGCTTCTTCTCTACAGCAAGCTCACCG  
 TCGACAAGAGCCGCTGGCAGCAGGGCAACGTCTTTAGCTGCAGCGTCATGCACGAG  
 GCCCTCCACAACCACTACACCCAGAAGAGCCTCAGCCTCAGCCCCGGAAGTAAGG  
 CGCG

(SEQ ID NO: 10)

**[0146]** Este ADN codifica KRGGG (SEQ ID NO: 2) y la cadena ligera del anticuerpo humano. Se incluyeron dos sitios de restricción SpeI y AclI para fines de clonación. La secuencia de nucleótidos fue mutada para eliminar un sitio *kex2* interno mediante mutagénesis dirigida (Stratagene, 11011 North Torrey Pines Road, La Jolla, CA 92037) y los dos cebadores usados para la mutagénesis en la reacción de PCR fueron TCGAGATCACCCGCACCGTCGCG GCGCCAAG (SEQ ID NO: 11) y CGACGGTGCGGGTGATCTCGACCTTGGTGCCCTGG CCG (SEQ ID NO: 12). El ADN que codifica la cadena ligera resultante contenía dos nucleótidos sustituidos en la secuencia de ADN que cambiaban el aminoácido K a T. Se sintetizaron dos oligos, GGACTAGTGTGCCGTTGAGAAACGCGACATCCAGATGACCCAGAGC (SEQ ID NO: 13) y CTAAAGGGAACAAAAGCTGGAGC (SEQ ID NO: 14), mediante Invitrogen y se usaron para amplificar ADN de la cadena ligera. El fragmento de PCR resultante codifica la cadena ligera y VAVEKR (SEQ ID NO: 9). El fragmento de PCR fue digerido con enzimas de restricción SpeI y AclI y clonado al vector de expresión pTrex4 para generar un plásmido llamado pTrex4-VAVE-her2 cadena ligera Geneart (KR-TR). Se analizó la fidelidad del fragmento de PCR mediante secuenciación de ADN. El plásmido fue digerido con enzima de restricción XbaI y cotransformado biolísticamente en la cepa *T.reesei* con el plásmido de expresión de la cadena pesada. Se obtuvieron más de 40 transformantes y se transfirieron a nuevas placas. Se seleccionaron más de 20 transformantes estables para su cultivo en medio Proflo durante 2 días a 30°C. Se transfirieron 5 ml del cultivo de dos días de antigüedad de Proflo a 50 ml de medio definido. Los cultivos se mantuvieron durante 4 días a 28°C.

Los caldos de cultivo fueron centrifugados y se usaron los sobrenadantes para el análisis de proteínas. Los datos de Western blot indicaron que en la variante VAVEKR (transformante 17-43), más del 90% de la proteína de fusión fue escindida (Fig. 5).

5 **[0147]** Para generar cambios de aminoácido en el residuo de glutamina de la presecuencia de sitio KEX2 de VAVEKR (SEQ ID NO: 9), se sintetizó un cebador degenerado (GGACTAGTGTGCGCCGTTNNSAAACGCGACATCC AGATGACCCAGAG (SEQ ID NO:15) y se usó en una reacción de PCR con cebador inverso (SEQ ID NO: 14) para amplificar ADN para generar un conjunto de fragmentos de PCR. Los fragmentos de PCR mezclados se clonaron en vector de expresión de *Trichoderma* p(Trex4). Se secuenciaron 13 clones y se produjeron 7 variantes (tabla 1). Los 7 plásmidos se transformaron biolísticamente en la cepa *T. reesei*. Se obtuvieron más de 40 transformantes para cada variante y se transfirieron a nuevas placas. Para el primer grupo de tres variantes (VAVWKR (SEQ ID NO: 25), VAVGKR (SEQ ID NO: 26) y VAVRKR (SEQ ID NO: 27)), se seleccionaron 15 transformantes estables para cada variante. Para el segundo conjunto de cuatro variantes (VAVTKR (SEQ ID NO: 28), VAVVKR (SEQ ID NO: 29), VAVAKR (SEQ ID NO: 30) y VAVLKR (SEQ ID NO: 31)), se seleccionaron 11 transformantes estables para cada variante. Los transformantes seleccionados se cultivaron en medio Proflo durante 2 días a 28°C. Se transfirieron 5 ml de cultivo de dos días de antigüedad de Proflo a 50 ml de medio definido. Los cultivos se mantuvieron durante 4 días a 28°C. Los caldos de cultivo fueron centrifugados y se usaron los sobrenadantes para el análisis de proteínas.

20 **[0148]** Se sintetizó un nuevo cebador (GGACTAGTGTGCGCCGTTNACAAACGCGACATCCAGATGAC CCAGAG SEQ ID NO: 16) y se usó en una reacción de PCR con cebador inverso (SEQ ID NO: 14) para amplificar ADN para generar fragmentos de PCR con secuencias múltiples. Los fragmentos de PCR mezclados se clonaron en vector de expresión de *Trichoderma* p(Trex4). Se secuenciaron 10 clones y se produjeron 4 variantes más (VAVDKR (SEQ ID NO: 32), VAVNKR (SEQ ID NO: 33), VAVYKR (SEQ ID NO: 34) y VAVHKR (SEQ ID NO: 35)). Se transformaron los plásmidos biolísticamente en la cepa *Trichoderma* descrita arriba. Se obtuvieron más de 40 transformantes para cada variante y se transfirieron a nuevas placas. Se seleccionaron 10 transformantes estables de cada variante y se cultivaron en medio Proflo durante 2 días a 28°C. Se transfirieron 5 ml del cultivo de dos días de antigüedad de Proflo a 50 ml de medio definido. Los cultivos se mantuvieron durante 4 días a 28°C. Los caldos de cultivo fueron centrifugados y se usaron los sobrenadantes para el análisis de proteínas.

30 **[0149]** Se seleccionó un transformante (el mejor transformante productor de cadena ligera) de cada variante en el residuo de glutamina para compararse. El análisis Western indicó que la variante VAVYKR (SEQ ID NO: 34) produjo más cadena ligera que cualquier otra variante. Las variantes VAVTKR (SEQ ID NO: 28) y VAVDKR (SEQ ID NO: 32) tienen más proteína de fusión lo que indica una escisión menos eficiente. (Fig. 5).

35 **[0150]** Para generar cambios de aminoácido en el primer residuo de valina del sitio de presecuencia de KEX2 (VAVEKR, SEQ ID NO: 9), se sintetizó un cebador degenerado (GGACTAGTNNSGCCGTCGAGAAGCGCGACATCCAGATGACCCAG AG; SEQ ID NO: 17) que se usó en una reacción de PCR con cebador inverso CTAAAGGGAACAAAAGCTGGAGC (SEQ ID NO:14) para amplificar ADN para generar fragmento de PCR con múltiples secuencias. Los fragmentos de PCR mezclados se clonaron en vector de expresión de *Trichoderma* (pTrex4). Se secuenciaron 30 clones y se produjeron 13 variantes (MAVEKR (SEQ ID NO: 36), GAVEKR (SEQ ID NO: 37), AAVEKR (SEQ ID NO:38), LAVEKR (SEQ ID NO: 39), WAVEKR (SEQ ID NO: 40), KAVEKR (SEQ ID NO: 41), PAVEKR (SEQ ID NO: 42), RAVEKR (SEQ ID NO: 43), NAVEKR (SEQ ID NO: 44), TAVEKR (SEQ ID NO: 45), SAVEKR (SEQ ID NO: 46), QAVEKR (SEQ ID NO: 47) y EAVEKR (SEQ ID NO: 48)). Se diseñó un nuevo cebador, se sintetizó (GGACTAGTNWCGCCGTCGAGAAGCGCGACATCCAGATGACCCAGAG SEQ ID NO: 18) y se usó en una reacción de PCR con cebador inverso CTAAAGGGAACAAAAGCTGGAGC (SEQ ID NO: 14) para amplificar ADN para generar fragmento de PCR con múltiples secuencias. Los fragmentos de PCR mezclados se clonaron en vector de expresión de *Trichoderma* (pTrex4). Se secuenciaron 19 clones y se produjeron 5 variantes más (YAVEKR (SEQ ID NO: 49), FAVEKR (SEQ ID NO: 50), DAVEKR (SEQ ID NO: 51), HAVEKR (SEQ ID NO: 52) y IAVEKR (SEQ ID NO: 53)). Los plásmidos que contenían las siguientes 11 variantes (MAVEKR (SEQ ID NO: 36), GAVEKR (SEQ ID NO: 37), AAVEKR (SEQ ID NO: 38), LAVEKR (SEQ ID NO:39), WAVEKR (SEQ ID NO: 40), KAVEKR (SEQ ID NO: 41), PAVEKR (SEQ ID NO: 42), HAVEKR (SEQ ID NO: 52), DAVEKR (SEQ ID NO: 51), SAVEKR (SEQ ID NO: 46) y QAVEKR (SEQ ID NO: 47)) se transformaron biolísticamente en la cepa *T. reesei*.

55 **[0151]** Se obtuvieron más de 20 transformantes para cada variante y se transfirieron a nuevas placas. Se seleccionaron más de 8 transformantes estables de cada variante y se cultivaron en medio Proflo durante 2 días a 28°C. Se transfirieron 5 ml del cultivo de dos días de antigüedad de Proflo a 50 ml de medio definido. Los cultivos se mantuvieron durante 4 días a 28°C. Los caldos de cultivo fueron centrifugados y se analizaron los sobrenadantes mediante SDS-PAGE de proteínas. Se seleccionó un transformante (el mejor transformante productor) de cada variante. El análisis Western indicó (Fig. 6) que todas las variantes producían cadena ligera. Todas mostraron menos del 95% de escisión excepto LAVEKR (SEQ ID NO: 39). Esta variante mostró escisión de KEX2 más eficiente que la variante VAVEKR (SEQ ID NO: 9).

**[0152]** Para generar cambios de aminoácido en el residuo de alanina de la región KEX2 (VAVEKR, SEQ ID NO: 9)), se sintetizó un cebador degenerado (GGACTAGTGTCCNNSGTTGAGAAAGGCGACATCCAGATGACCCA GAGC; SEQ ID NO: 19) que se usó en una reacción de PCR con cebador inverso (SEQ ID NO:14) para amplificar ADN para generar fragmento de PCR con múltiples secuencias. Los fragmentos de PCR mezclados se clonaron en vector de expresión de *Trichoderma* (pTrex4). Se secuenciaron 96 clones y se produjeron 15 variantes (VDVEKR (SEQ ID NO: 54), VEVEKR (SEQ ID NO: 55), VGVEKR (SEQ ID NO: 56), VIVEKR (SEQ ID NO: 57), VKVEKR (SEQ ID NO: 58), VLVEKR (SEQ ID NO: 59), VMVEKR (SEQ ID NO: 60), VNVEKR (SEQ ID NO: 61), VPVEKR (SEQ ID NO:62), VRVEKR (SEQ ID NO: 63), VSVEKR (SEQ ID NO: 64), VTVEKR (SEQ ID NO: 65), VVVEKR (SEQ ID NO: 66), VWVEKR (SEQ ID NO: 67) y VYVEKR (SEQ ID NO: 68)). Se transformaron 5 plásmidos biolísticamente en la cepa *T. reesei*. Se obtuvieron más de 20 transformantes de cada variante y se transfirieron a nuevas placas. Para este primer conjunto de 5 variantes (VGVEKR (SEQ ID NO: 56), VTVEKR (SEQ ID NO: 65), VWVEKR (SEQ ID NO: 67), VEVEKR (SEQ ID NO: 55) y VPVEKR (SEQ ID NO: 62)), se seleccionaron 10 transformantes estables. Para este segundo conjunto de 4 variantes (VKVEKR (SEQ ID NO: 58), VRVEKR (SEQ ID NO: 63), VVVEKR (SEQ ID NO: 66) y VIVEKR (SEQ ID NO: 57)), se seleccionaron 10 transformantes estables. Los transformantes seleccionados se cultivaron en medio Proflo durante 2 días a 28°C. Se transfirieron 5 ml del cultivo de dos días de antigüedad de Proflo a 50 ml de medio definido. Los cultivos se mantuvieron durante 4 días a 28°C. Los caldos de cultivo fueron centrifugados y se analizaron los sobrenadantes mediante gel de SDS de proteínas. Se seleccionó un transformante (el mejor transformante productor) de cada variante para compararse (tabla 1). El análisis Western (Fig. 7) indicó que solo la cadena ligera libre podía detectarse en las tres variantes: VGVEKR (SEQ ID NO: 56); VEVEKR (SEQ ID NO: 55) y VWVEKR (SEQ ID NO: 67). La variante VPVEKR (SEQ ID NO: 62) produjo menos cadena ligera libre y alguna fusión ligera-CBHI no escindida.

**[0153]** Para generar cambios de aminoácido en el segundo residuo de valina del sitio KEX2 (VAVEKR, SEQ ID NO: 9), se sintetizó un cebador degenerado (GGACTAGTGTCCNNSGAGAAACGCGACATCCAGATGACCCAG AG; SEQ ID NO:20) que se usó en una reacción de PCR con cebador inverso (SEQ ID NO:14) para amplificar ADN para generar fragmento de PCR con múltiples secuencias. Los fragmentos de PCR mezclados se clonaron en vector de expresión de *Trichoderma* (pTrex4). Se secuenciaron 36 clones y se produjeron 15 variantes (VAAEKR (SEQ ID NO: 69), VADEKR (SEQ ID NO: 70), VAAEKR (SEQ ID NO: 71), VAFEKR (SEQ ID NO: 72), VAGEKR (SEQ ID NO: 73), VAIEKR (SEQ ID NO: 74), VALEKR (SEQ ID NO: 75), VANEKR (SEQ ID NO: 76), VAQEKR (SEQ ID NO: 77), VAREKR (SEQ ID NO: 78), VASEKR (SEQ ID NO: 79), VATEKR (SEQ ID NO: 80), VAWEKR (SEQ ID NO: 81), VAYEKR (SEQ ID NO: 82) y VAPEKR (SEQ ID NO: 83)). Se transformaron plásmidos biolísticamente en la cepa *T. reesei*. Se obtuvieron más de 20 transformantes para cada variante y se transfirieron a nuevas placas. Para el primer conjunto de 8 variantes (VAAEKR (SEQ ID NO: 69), VADEKR (SEQ ID NO: 70), VAAEKR (SEQ ID NO: 71), VAFEKR (SEQ ID NO: 72), VAGEKR (SEQ ID NO: 73), VANEKR (SEQ ID NO: 76), VALEKR (SEQ ID NO: 75) y VAIEKR (SEQ ID NO: 74)), se seleccionaron 10 transformantes estables. Para el segundo conjunto de 2 variantes (VASEKR (SEQ ID NO: 79) y VAREKR (SEQ ID NO: 78), se seleccionaron 8 transformantes estables. Se seleccionaron solo 4 transformantes para la variante VAPEKR (SEQ ID NO: 83). Los transformantes seleccionados se cultivaron en medio Proflo durante 2 días a 28°C. Se transfirieron 5 ml del cultivo de dos días de antigüedad de Proflo a 50 ml de medio definido. Los cultivos se mantuvieron durante 4 días a 28°C. Los caldos de cultivo fueron centrifugados y se analizaron los sobrenadantes. Se seleccionó un transformante (el mejor transformante productor) de cada variante para compararse (Tabla 1). El análisis Western (Fig. 8) indicó que VAIEKR (SEQ ID NO: 74) y VALEKR (SEQ ID NO: 75) generaban una escisión completa del polipéptido de fusión ya que no se observó una banda de fusión en el gel. El Western blot (Fig. 8) indicó que VAFEKR (SEQ ID NO: 72) produjo la mayor cantidad de cadena ligera de anticuerpo a pesar de que la escisión no fue del 100%.

Tabla 1

MAVEKR (SEQ ID NO: 36)	VKVEKR (SEQ ID NO: 58)	VAAEKR (SEQ ID NO: 69)	VAVWKR (SEQ ID NO: 25)
GAVEKR (SEQ ID NO: 37)	VRVEKR (SEQ ID NO: 63)	VADEKR (SEQ ID NO: 70)	VAVGKR (SEQ ID NO:26)
AAVEKR (SEQ ID NO: 38)	VVVEKR (SEQ ID NO: 66)	VAAEKR (SEQ ID NO: 71)	VAVRKR (SEQ ID NO:27)
LAVEKR (SEQ ID NO: 39)	VIVEKR (SEQ ID NO: 57)	VAFEKR (SEQ ID NO: 72)	VAVTKR (SEQ ID NO:28)
WAVEKR (SEQ ID NO: 40)	VEVE (SEQ ID NO: 55)	VAGEKR (SEQ ID NO: 73)	VAVVKR (SEQ ID NO: 29)
KAVEKR (SEQ ID NO: 41)	VGVEKR (SEQ ID NO: 56)	VAIEKR (SEQ ID NO: 74)	VAVAKR (SEQ ID NO: 30)



PAVEKR (SEQ ID NO: 42)	VPVEKR (SEQ ID NO: 62)	VALEKR (SEQ ID NO: 75)	VAVLKR (SEQ ID NO: 31)
SAVEKR (SEQ ID NO: 46)	VTVEKR (SEQ ID NO: 65)	VANEKR (SEQ ID NO: 76)	VAVDKR (SEQ ID NO: 32)
QAVEKR (SEQ ID NO: 47)	VWVEKR (SEQ ID NO: 67)	VASEKR (SEQ ID NO: 79)	VAVNKR (SEQ ID NO: 33)
DAVEKR (SEQ ID NO: 51)		VAREKR (SEQ ID NO: 78)	VAVYKR (SEQ ID NO: 34)
HAVEKR (SEQ ID NO: 52)		VAPEKR (SEQ ID NO: 83)	VAVHKR (SEQ ID NO: 35)

5

## 10 Ejemplo 6

### Construcción de una cepa de expresión de la cadena ligera de trastuzumab que contiene la región KEX2 NVISKR (SEQ ID NO: 22)

[0154] Una región KEX2 NVISKR se encuentra de manera natural en la prosequencia de la glucoamilasa de *A. niger* (glaA). Para construir un polipéptido de fusión, se sintetizó un oligo, GGACTAGTAACGTCATCAGCAAGCGCGACATCCAGATGACCCAGAGC (SEQ ID NO: 21) mediante Invitrogen y se usó para amplificar ADN de cadena ligera con cebador inverso (SEQ ID NO: 14). El fragmento de PCR resultante codifica la cadena ligera y la secuencia NVISKR (SEQ ID NO: 22) para la escisión de KEX2. El fragmento de PCR fue digerido con enzimas de restricción SpeI y AclI y clonado al vector de expresión pTrex4 para generar un plásmido llamado pTrex4-NVIS-her2 cadena ligera geneart (KR-TR). Se analizó la fidelidad del fragmento de PCR mediante secuenciación de ADN. Se transformó el plásmido biolísticamente en la cepa *Trichoderma reesei*. Se obtuvieron más de 20 transformantes y se transfirieron a nuevas placas. Se seleccionaron 10 transformantes estables para su cultivo en medio Proflo durante 2 días a 30°C. Se transfirieron 5 ml del cultivo de dos días de antigüedad de Proflo a 50 ml de medio definido. Los cultivos se mantuvieron durante 5 días a 28°C. Los caldos de cultivo fueron centrifugados y se usaron los sobrenadantes para el análisis de proteínas. El análisis Western indicó que más del 95% de las proteínas de fusión fueron escindidas (Fig. 6).

15

20

25

## Ejemplo 7

### Construcción de una cepa de expresión de la cadena ligera de trastuzumab que contiene la región KEX2 SDVTKR (SEQ ID NO: 24)

[0155] Se sintetizó un oligo, GGACTAGTAGCGACGTCACCAAGCGCGACATCCAGATGACCCAGAGC (SEQ ID NO: 23) mediante Invitrogen y se usó para amplificar ADN de cadena ligera con cebador inverso (SEQ ID NO: 14). El fragmento de PCR resultante codifica la cadena ligera y la secuencia SDVTKR (SEQ ID NO: 24) para la escisión de kex2. El fragmento de PCR fue digerido con enzimas de restricción SpeI y AclI y clonado al vector de expresión pTrex4 para generar un plásmido denominado pTrex4-SDVT-her2 cadena ligera geneart (KR-TR). Se analizó la fidelidad del fragmento de PCR mediante secuenciación de ADN. Se transformó el plásmido biolísticamente en la cepa *Trichoderma reesei*. Se obtuvieron más de 20 transformantes y se transfirieron a nuevas placas. Se seleccionaron 10 transformantes estables para su cultivo en medio Proflo durante 2 días a 30°C. Se transfirieron 5 ml del cultivo de dos días de antigüedad de Proflo a 50 ml de medio definido. Los cultivos se mantuvieron durante 5 días a 28°C. Los caldos de cultivo fueron centrifugados y se usaron los sobrenadantes para el análisis de proteínas. El análisis Western indicó que más del 50% de las proteínas de fusión fueron escindidas (Fig. 6).

30

35

40

## Ejemplo 8

### Construcción de una cepa de expresión de la cadena ligera de trastuzumab que contiene una región KEX2 VAVEKR (SEQ ID NO: 9) en *Aspergillus niger*

[0156] El plásmido pTrex4-VAVE-her2 cadena ligera geneart (KR-TR) del Ejemplo 5 se digirió con SpeI y AclI. El extremo del fragmento de ADN del sitio de corte de AclI generó extremo romo mediante ADN polimerasa de T4. El fragmento fue asilado en un gel de agarosa 1,2% y ligado a plásmido de expresión de *A. niger* (pSLGAMpR2-BBI según se divulga en la publicación de patente estadounidense nº 2005 0153399) que se cortó con NheI y BstEII con el extremo BstEII generado romo con ADN polimerasa de T4. El nuevo plásmido, llamado

45

pSLGAMpR2-VAVE-her2 LC geneart se transformó en la cepa de *A. niger* dgr246:Δamy5;pyr- que se deriva de la cepa dgr246:ΔGAP:pyr- divulgada en la publicación de patente estadounidense 20050153399. La diferencia consiste en que el nivel de proteína de α-amilasa se ve enormemente reducido en este plásmido debido a una mutación.

- 5 [0157] El dgr246ΔGAP:pyr2- se deriva de la cepa dgr246 P2 que tiene el gen *pepA* eliminado, es *pyrG* negativo y se ha sometido a diversas rondas de mutagénesis y cribado o selección para la producción mejorada de un producto génico heterólogo (Ward, M. et al., 1993, Appl. Microbiol. Biotech. 39:738-743 y referencias allí contenidas). Para crear cepa dgr246ΔGAP:pyr2-, se eliminó el gen *glaA* (glucoamilasa) en la cepa dgr246 P2 usando exactamente el mismo plásmido de delección (pΔGAM NB-Pyr) y procedimiento presentado por Fowler, T. et al (1990) Curr. Genet. 18:537-545. Brevemente, la delección se logró por transformación con un fragmento de ADN lineal que tenía secuencias flanqueantes de *glaA* en cualquier extremo y con parte de la región promotora y codificante del gen *glaA* sustituida por el gen *pyrG* de *Aspergillus nidulans* como marcador seleccionable. Los transformantes en los que el fragmento lineal que contiene las secuencias flanqueantes de *glaA* y el gen *pyrG* se habían integrado en el locus de *glaA* cromosómico se identificaron mediante análisis Southern blot. Este cambio se había producido en la cepa transformada dgr246ΔGAP. Se colocaron en placas esporas de este transformante en medio que contenía ácido fluoroorótico y se obtuvieron mutantes resistentes espontáneos como describe van Hartingsveldt, W. et al. (1987) Mol. Gen. Genet. 206:71-75. Uno de estos, dgr246ΔGAP:pyr2-, demostró ser una cepa auxótrofa para uridina que podía complementarse mediante transformación con plásmidos que portaban un gen *pyrG* de tipo salvaje.
- 10
- 15
- 20 [0158] Se obtuvieron más de 20 transformantes y se transfirieron a nuevas placas. Se cultivaron 17 transformantes en medio Promosoy durante 5 días a 28°C. Los caldos de cultivo fueron centrifugados y se usaron los sobrenadantes para el análisis SDS PAGE de proteínas y análisis Western. Los datos indicaron que todos los transformantes produjeron cadena ligera de anticuerpo. El transformante #A12 produjo la mayor cantidad de cadena ligera de anticuerpo, y el 60-70% de la proteína de fusión fue escindida (Fig. 9).
- 25 [0159] La descripción que precede simplemente ilustra los principios de los modos de realización de ejemplo.
- [0160] Además, todos los ejemplos y lenguaje condicional aquí mencionados pretenden principalmente ayudar al lector en la comprensión de los principios de la invención y los conceptos aportados por los inventores para el avance de la técnica, y deben interpretarse como no limitativos a dichos ejemplos y condiciones mencionados de manera específica.
- 30 [0161] Además, todas las afirmaciones aquí contenidas que recogen principios, aspectos y modos de realización de la invención, así como ejemplos específicos de la misma, pretenden abarcar equivalentes tanto estructurales como funcionales de las mismas. Además, se pretende que dichos equivalente incluyan equivalentes actualmente conocidos así como equivalentes que se desarrollen en el futuro, es decir, cualquier elemento desarrollado que lleve a cabo la misma función, independientemente de la estructura. El alcance de la presente invención, por ello, no pretende limitarse a los modos de realización de ejemplo mostrados y descritos en la presente memoria.
- 35

40

45

**Reivindicaciones**

1. Un constructo de ADN de fusión que codifica un polipéptido de fusión que comprende en unión operable desde el extremo 5' de dicho constructo, un promotor;
- 5 una primera molécula de ADN que codifica una secuencia señal funcional como una secuencia secretora en una célula fúngica filamentosa;
- una segunda molécula de ADN que codifica una proteína portadora, donde la proteína portadora es un polipéptido fúngico secretado de manera natural o una porción funcional del mismo;
- 10 una tercera molécula de ADN que codifica una región KEX2, donde la región KEX2 es VAVYKR; y
- una cuarta molécula de ADN que codifica una proteína deseada.
  
2. El constructo de ADN de fusión de la reivindicación 1, donde:
  - 15 (a) la primera molécula de ADN y la segunda molécula de ADN codifican una secuencia señal de CBH1 de *Trichoderma* y proteína portadora o una secuencia señal de endoglucanasa de *Trichoderma* y proteína portadora; o
  - (b) la primera molécula de ADN y la segunda molécula de ADN codifican una secuencia señal de glucoamilasa y proteína portadora o una secuencia señal de alfa-amilasa y proteína portadora.
  
- 20 3. El constructo de ADN de fusión de la reivindicación 1, donde la proteína deseada es una enzima o la proteína deseada es una proteína terapéutica.
  
4. El constructo de ADN de fusión de la reivindicación 3, donde la proteína terapéutica es un anticuerpo o es un anticuerpo monoclonal de cadena pesada o cadena ligera.
  
- 25 5. El constructo de ADN de fusión de la reivindicación 1, donde:
  - (a) la primera molécula de ADN y segunda molécula de ADN codifican una secuencia señal de CBH1 y proteína portadora y la cuarta molécula de ADN codifica una cadena ligera de anticuerpo o fragmento de la misma; o
  - 30 (b) la primera molécula de ADN y segunda molécula de ADN codifican una secuencia señal de glucoamilasa y proteína portadora y la cuarta molécula de ADN codifica una cadena ligera de anticuerpo o fragmento de la misma.
  
6. La proteína de fusión codificada por el constructo de ADN de fusión de cualquiera de las reivindicaciones precedentes.
  
- 35 7. Un vector que comprende el constructo de ADN de fusión de cualquiera de las reivindicaciones de la 1 a la 5.
  
8. Una célula huésped que comprende el vector de la reivindicación 7.
  
9. La célula huésped de la reivindicación 8, donde la célula huésped es una célula huésped de *Trichoderma*.
  
10. La célula huésped de la reivindicación 9, donde la célula de *Trichoderma* es una célula de *T. reesei*.
  
- 40 11. Un proceso para producir una proteína deseada en una célula fúngica filamentosa que comprende:
  - (a) obtener una célula huésped fúngica filamentosa que comprende un constructo de ADN de fusión según cualquiera de las reivindicaciones de la 1 a la 5;
  - (b) cultivar la célula huésped en condiciones adecuadas que permitan la expresión y producción de la proteína deseada; y
  - 45 (c) recuperar la proteína deseada.
  
12. Un proceso para escindir una proteína deseada de un polipéptido de fusión recombinante que comprende expresar en una célula fúngica filamentosa un polipéptido de fusión codificado por un constructo de ADN de fusión según cualquiera de las reivindicaciones de la 1 a la 5, donde dicha región KEX2 proporciona un sitio de escisión de proteína y obtener una proteína deseada que se escinde del polipéptido de fusión expresado.
- 50 13. El proceso según la reivindicación 12, donde la escisión de la proteína deseada se aumenta en comparación con la escisión de dicha proteína deseada de un polipéptido de fusión equivalente que carece de la presecuencia de sitio KEX2.

- 5 **14.** Un proceso para aumentar la producción de un anticuerpo a partir de una célula fúngica filamentosa que comprende obtener una célula fúngica filamentosa que comprende un constructo de ADN de fusión de cualquiera de las reivindicaciones de la 1 a la 5, cultivar la célula fúngica en condiciones adecuadas para la expresión del polipéptido de fusión y permitir la secreción del polipéptido de fusión, donde la secreción de la proteína deseada se aumenta en comparación con la secreción de un polipéptido de fusión equivalente que no incluye una presecuencia KEX2.
- 15.** El proceso de la reivindicación 14, donde la secreción de la proteína deseada es aumentada en al menos un 30% en comparación con la secreción de la proteína deseada a partir del polipéptido de fusión equivalente.
- 16.** El proceso de cualquiera de las reivindicaciones de la 11 a la 15, donde:
- 10 (a) la proteína deseada es una proteína terapéutica, una inmunoglobulina, un anticuerpo monoclonal; y/o  
(b) un anticuerpo monoclonal de cadena ligera o cadena pesada o fragmento del mismo.
- 15 **17.** Un polipéptido de fusión que comprende a partir de un extremo amino terminal de dicho polipéptido de fusión una primera secuencia de aminoácidos que comprende:
- 20 (a) una secuencia señal funcional como secuencia secretora en una célula fúngica filamentosa;  
(b) una segunda secuencia de aminoácidos que comprende una proteína portadora, donde la proteína portadora es un polipéptido fúngico secretado de manera natural;  
(c) una tercera secuencia de aminoácidos que comprende una región KEX2, donde dicha región es VAVYKR; y  
(d) una cuarta secuencia de aminoácidos que comprende una proteína deseada.

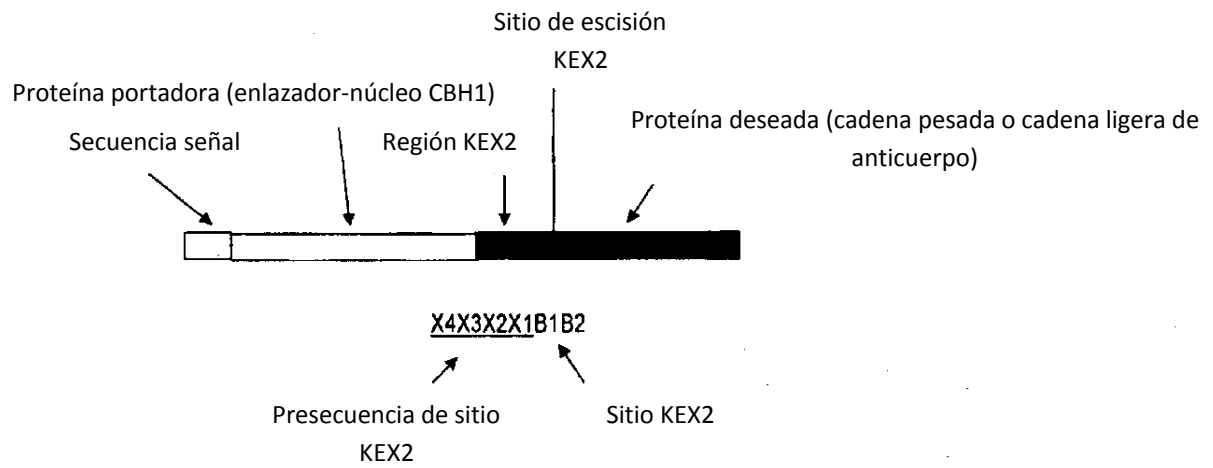


FIG. 1

Fig. 2

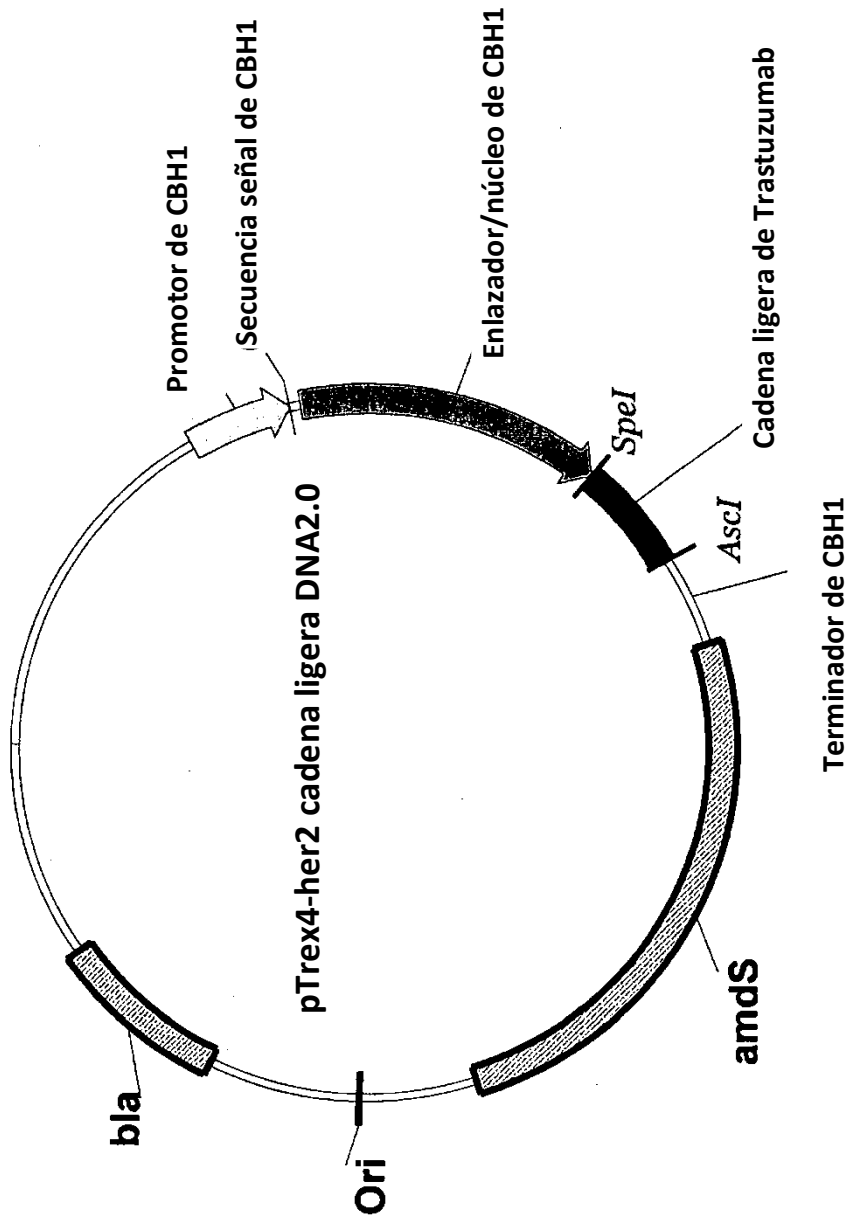


FIG. 3A

Secuencia de ADN de pTrex4-her2 cadena ligera DNA2.0 (10885 bases) (SEQ ID NO: 103)

AAGCTTAATACTTCTCGAGCTCTGTACATGTCCGGTCGCGACGTACGCGTATCGATG  
 GCGCCAGCTGCAGGCGGCCGCTGCAGCCACTTGCAGTCCCGTGGAATTCTCACGGT  
 GAATGTAGGCCTTTTGTAGGGTAGGAATTGTCACTCAAGCACCCCAACCTCCATTA  
 CGCCTCCCCATAGAGTTCCTCAATCAGTGAGTCATGGCACTGTTCTCAAATAGATTG  
 GGGAGAAGTTGACTTCCGCCCAGAGCTGAAGGTGCGACAACCGCATGATATAGGGTC  
 GGCAACGGCAAAAAAGCACGTGGCTCACCGAAAAGCAAGATGTTTGCGATCTAACAT  
 CCAGGAACCTGGATACATCCATCATCACGCACGACCACTTTTGATCTGCTGGTAACT  
 CGTATTCGCCCTAAACCGAAGTGACGTGGTAAATCTACACGTGGGCCCTTTCGGTA  
 TACTGCGTGTGTCTTCTCTAGGTGCCATTCTTTTCCCTTCCTCTAGTGTGAAATTGT  
 TTGTGTTGGAGTCCGAGCTGTAACCTCTGAATCTCTGGAGAATGGTGGACTAAC  
 GACTACCGTGCACCTGCATCATGTATATAATAGTGATCCTGAGAAGGGGGGTTTGGGA  
 GCAATGTGGGACTTTGATGGTCATCAAACAAAGAACGAAGACGCTCTTTTGCAAAG  
 TTTTGTTCGGCTACGGTGAAGAACTGGATACTTGTGTGTCTTCTGTGTATTTTTG  
 TGGAACAAGAGGCCAGAGACAATCTATTCAAACACCAAGCTTGCTCTTTTGAGCTA  
 CAAGAACCTGTGGGGTATATATCTAGAGTTGTGAAGTCGGTAATCCCGCTGTATAGT  
 AATACGAGTCGCATCTAAATACTCCGAAGCTGCTGCGAACCCGGAGAATCGAGATGT  
 GCTGGAAAGCTTCTAGCGAGCGGCTAAATTAGCATGAAAGGCTATGAGAAATCTGG  
 AGACGGCTTGTGAAATCATGGCGTTCATTCTTCGACAAGCAAAGCGTTCCGTCGCA  
 GTAGCAGGCACTCATTCCCGAAAAAACTCGGAGATTCTAAGTAGCGATGGAACCGG  
 AATAATATAATAGGCAATACATTGAGTTGCCCTCGACGGTTGCAATGCAGGGGTACTG  
 AGCTTGACATAACTGTTCCGTACCCACCTCTTCTCAACCTTTGGCGTTCCTCTGA  
 TTCAGCGTACCCGTACAAGTCGTAATCACTATTAACCCAGACTGACCGGACGTGTTT  
 TGCCCTTCATTTGGAGAAATAATGTCATTGCGATGTGTAATTTGCCTGCTTGACCGA  
 CTGGGGCTGTTTCGAAGCCCGAATGTAGGATTGTTATCCGAACCTCTGCTCGTAGAGGC  
 ATGTTGTGAATCTGTGTGCGGGCAGGACACGCCTCGAAGGTTACGGCAAGGGAAACC  
 ACCGATAGCAGTGTCTAGTAGCAACCTGTAAAGCCGCAATGCAGCATCACTGGAAAA  
 TACAAACCAATGGCTAAAAGTACATAAGTTAATGCCTAAAGAAGTCATATAACCAGCG  
 GCTAATAATGTACAATCAAGTGGCTAAACGTACCGTAATTTGCCAACGGCTTGTGG  
 GGTTCAGAAGCAACGGCAAAGCCCCACTTCCCACGTTTGTTCCTTCACTCAGTCC  
 AATCTCAGCTGGTGATCCCCAATGGGTGCTTGTGTTTCCGGTGAAGTGAAAGA  
 AGACAGAGGTAAGAATGTCTGACTCGGAGCGTTTTGCATACAACCAAGGGCAGTGAT  
 GGAAGACAGTGAAATGTTGACATTCAAGGAGTATTTAGCCAGGGATGCTTGAGTGTA  
 TCGTGTAAGGAGGTTTGTCTGCCGATACGACGAATACTGTATAGTCACTTCTGATGA  
 AGTGGTCCATATTGAAATGTAAGTCGGCACTGAACAGGCAAAGATTGAGTTGAAAC  
 TGCCTAAGATCTCGGGCCCTCGGGCCTTGGGTGTACATGTTTGTGCTCC  
 GGGCAAATGCAAAGTGTGGTAGGATCGAACACACTGCTGCCTTTACCAAGCAGCTGA  
 GGGTATGTGATAGGCAAATGTTTCAGGGCCACTGCATGGTTTTCGAATAGAAAGAGAA  
 GCTTAGCCAAGAACAAATAGCCGATAAAGATAGCCTCATTAAACGGAATGAGCTAGTA  
 GGCAAAGTCAGCGAATGTGTATATATAAAGGTTTCGAGGTCGGTGCCTCCCTCATGCT  
 CTCCCCATCTACTCATCAACTCAGATCCTCCAGGAGACTGTACACCATCTTTTGAG  
 GCACAGAAACCAATAGTCAACCGGACTGCGCATCATGTATCGGAAGTTGGCCGT  
 CATCTCGGCCCTTCTTGCCACAGCTCGTGCTCAGTCGGCCTGCACTCTCCAATCGGA  
 GACTCACCCGCTCTGACATGGCAG

FIG. 3B

AAATGCTCGTCTGGTGGCACTTGCACTCAACAGACAGGCTCCGTGGTCATCGACGCC  
 AACTGGCGCTGGACTCACGCTACGAACAGCAGCACGAACTGCTACGATGGCAACACT  
 TGGAGCTCGACCCTATGTCTTGACAACGAGACCTGCGCGAAGAACTGCTGTCTGGAC  
 GGTGCCGCCTACGCGTCCACGTACGGAGTTACCACGAGCGGTAAACAGCCTCTCCATT  
 GGCTTTGTCACCAGTCTGCGCAGAAGAACGTTGGCGCTCGCCTTTACCTTATGGCG  
 AGCGACACGACCTACCAGGAATTCACCTGTCTGGCAACGAGTTCTCTTTGATGTT  
 GATGTTTCGCAGCTGCCGTAAGTACTTACCATGAACCCCTGACGTATCTTCTTG  
 GGCTCCCAGCTGACTGGCCAATTTAAGGTGCGGCTTGAACGGAGCTCTCTACTTCGT  
 GTCCATGGACGCGGATGGTGGCGTGAGCAAGTATCCCACCAACACCGCTGGCGCCAA  
 GTACGGCACGGGTACTGTGACAGCCAGTGTCCCGCGATCTGAAGTTCATCAATGG  
 CCAGGCCAACGTTGAGGGCTGGGAGCCGTCAATCCAAACAACGAAAACACGGGCATTGG  
 AGGACACGGAAGCTGCTGCTCTGAGATGGATATCTGGGAGGCCAACTCCATCTCCGA  
 GGCTCTTACCCCCAACCTTGCACGACTGTGCGCCAGGAGATCTGCGAGGGTGATGG  
 GTGCGGCGGAACCTTACTCCGATAACAGATATGGCGGCACTTGCATCCCGATGGCTG  
 CGACTGGAACCCATAACCGCTGGGCAACACCAGCTTCTACGGCCCTGGCTCAAGCTT  
 TACCCTCGATAACCAAGAAATGACCGTTGTACCCAGTTTCGAGACGTCCGGTGC  
 CATCAACCGATACTATGTCCAGAATGGCGTCACTTTCCAGCAGCCCAACGCCGAGCT  
 TGGTAGTTACTCTGGCAACGAGCTCAACGATGATTACTGCACAGCTGAGGAGGCAGA  
 ATTCGGCGGATCCTCTTTCTCAGACAAGGGCGGCTGACTCAGTTCAAGAAGGCTAC  
 CTCTGGCGGATGGTTCTGGTTCATGAGTCTGTGGGATGATGTGAGTTTGATGGACAA  
 ACATGCGCGTTGACAAAGAGTCAAGCAGCTGACTGAGATGTTACAGTACTACGCCAA  
 CATGCTGTGGCTGGACTCCACCTACCCGACAAACGAGACCTCCTCCACACCCGGTGC  
 CGTGCGCGGAAGCTGCTCCACCAGCTCCGGTGTCCCTGCTCAGGTCGAATCTCAGTC  
 TCCCAACGCCAAGGTACCTTCTCCAACATCAAGTTCGGACCAATGGCAGCACCCGG  
 CAACCTTAGCGGCGGCAACCTCCCGGCGAAACCCGCTGGCACCACCACCACCCG  
 CCGCCAGCCACTACCACTGGAAGCTCTCCCGACCTACTAGTAAACGCGGTGGCGG  
 TGATATTCAAATGACACAATCTCCTTCTTCTGTGAGCCTCAGTGGGCGACCCGTG  
 GAGATTACTTGCCGCGCTCTCAGGACGTTAACACTGCCGTGCGATGGTACCAGCA  
 GAAGCCAGGCAAGGCGCCCAAGCTTCTGATTTACAGCGCTTCGTTCTGTACTCTGG  
 CGTGCCATCCCGCTTCTCTGGCAGCCGAAGCGGACCGGATTTACCCCTGACCATTT  
 GTCCCTGCAGCCGAGGATTTGCCACGTATTACTGCCAGCAGCACTACCCACTCC  
 ACCACCTTTGGCCAAGGAACGAGAGTCGAAATCACTCGCACGGTCGCTGCCCTTC  
 AGTCTTCATCTTCCCCCAGCGACGAACAGCTGAAGTCTGGTACGGCCAGCGTCGT  
 TTGCTTGCTTAATAACTTCTATCCGCGAGAGGCGAAGGTCCAATGGAAGGTTGATAA  
 CGTTCTGCAGTCCGGCAATTCGAGGAGAGCGTGACCGAGCAGGATTCAAAGGATAG  
 CACCTACTCACTCAGCAGCACCCCTGACGTTGTCCAAGGCCGATTACGAGAAGCATAA  
 GTTGATGTCATGCGAGGTCAACACCAGGGACTGTCAAGCCAGTTACCAAGTCGTT  
 CAATCGAGGCGAGTGCTAAGGCGCGCCGCGCGCCAGCTCCGTGCGAAAGCCTGACGC  
 ACCGGTAGATTCTTGGTGAGCCGTATCATGACGGCGGCGGGAGCTACATGGCCCCG  
 GGTGATTTATTTTTTTTGTATCTACTTCTGACCCTTTTCAAATATACGGTCAACTCA  
 TCTTTCACTGGAGATGCGGCTGCTTGGTATTGCGATGTTGTCAGCTTGGCAAATG  
 TGGCTTCGAAAACACAAAACGATTCTTAGTAGCCATGCATTTAAGATAACGGAA  
 TAGAAGAAAGAGGAAATTAATAAAAAAAAAAAAAAAAAACAACATCCCGTTCATAACCCGT  
 AGAATCGCCGCTCTTCGTGTATCCAGTACCAGTTTATTTTGAATAGCTCGCCCGCT  
 GGAGAGCATCCTGAATGCAAGTAACAACC



FIG. 3C

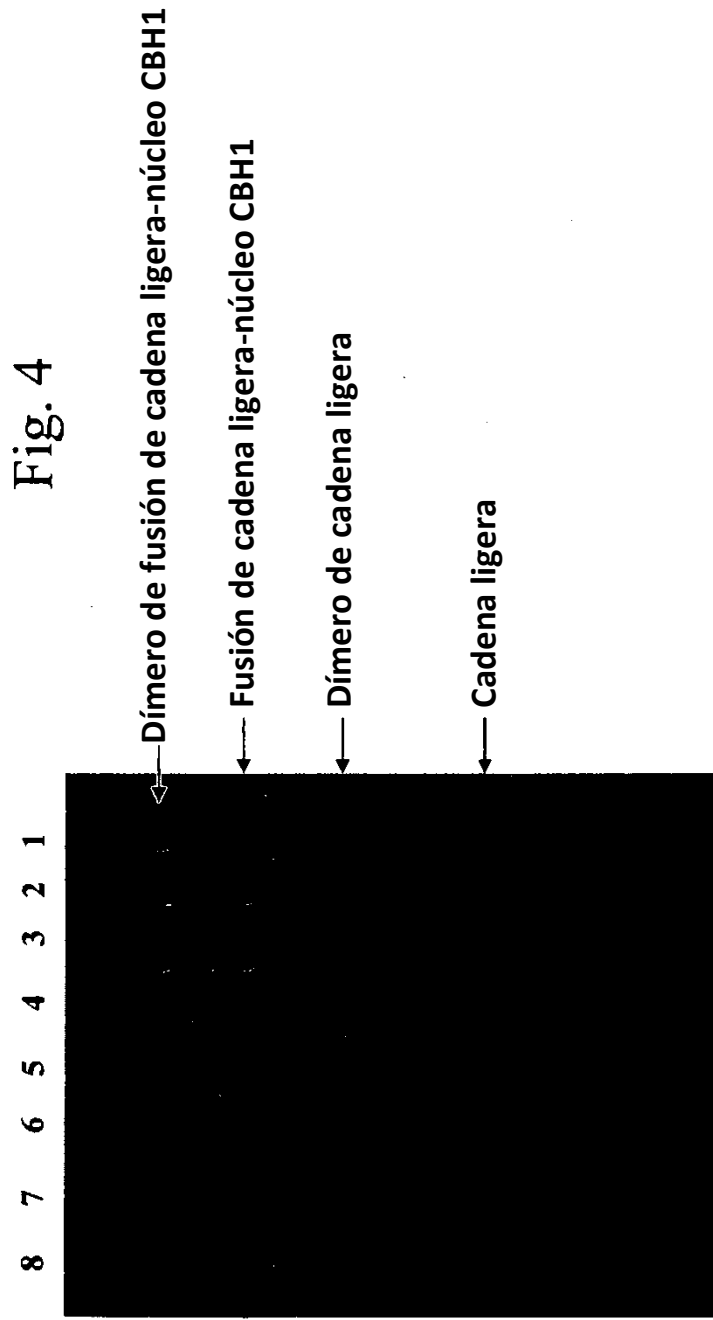
GTAGAGGCTGACACGGCAGGTGTTGCTAGGGAGCGTCGTGTTCTACAAGGCCAGACG  
 TCTTCGCGGTTGATATATATGTATGTTGACTGCAGGCTGCTCAGCGACGACAGTCA  
 AGTTCGCCCTCGCTGCTTGTGCAATAATCGCAGTGGGGAAGCCACACCGTGACTION  
 ATCTTTCAGTAAAGCTCTGTTGGTGTATATCAGCAATACACGTAATTTAAACTCGTT  
 AGCATGGGGCTGATAGCTTAATTACCGTTTACCAGTGCCGCGGTTCTGCAGCTTTCC  
 TTGGCCCGTAAAATTTCGGCGAAGCCAGCCAATCACCAGCTAGGCACCAGCTAAACCC  
 TATAATTAGTCTCTTATCAACACCATCCGCTCCCCGGGATCAATGAGGAGAATGAG  
 GGGGATGCGGGGCTAAAGAAGCCTACATAACCCTCATGCCAACTCCCAGTTTACT  
 CGTCGAGCCAACATCCTGACTATAAGCTAACACAGAATGCCTCAATCCTGGGAAGAA  
 CTGGCCGCTGATAAGCGCGCCCGCCTCGCAAAAACCATCCCTGATGAATGGAAAGTC  
 CAGACGCTGCCTGCGGAAGACAGCGTTATTGATTTCCCAAAGAAATCGGGGATCCTT  
 TCAGAGGCCGAACTGAAGATCACAGAGGCCCTCCGCTGCAGATCTTGTGTCCAAGCTG  
 GCGGCCGGAGAGTTGACCTCGGTGGAAGTTACGCTAGCATTCTGTAAACGGGCAGCA  
 ATCGCCCAGCAGTTAGTAGGGTCCCCTCTACCTCTCAGGGAGATGTAACAACGCCAC  
 CTTATGGGACTATCAAGCTGACGCTGGCTTCTGTGCAGACAAACTGCGCCCACGAGT  
 TCTTCCCTGACGCCGCTCTCGCGCAGGCAAGGGAACCTCGATGAATACTACGCAAAGC  
 ACAAGAGACCCGTTGGTCCACTCCATGGCCTCCCATCTCTCTCAAAGACCAGCTTC  
 GAGTCAAGGTACACCGTTGCCCTAAGTCGTTAGATGTCCCTTTTTGTGAGCTAACA  
 TATGCCACCAGGGCTACGAAACATCAATGGGCTACATCTCATGGCTAAACAAGTACG  
 ACGAAGGGGACTCGGTTCTGACAACCATGCTCCGCAAAGCCGGTGCCGTCTTCTACG  
 TCAAGACCTCTGTCCCGCAGACCCTGATGGTCTGCGAGACAGTCAACAACATCATCG  
 GGCGCACCGTCAACCCACGCAACAAGAACTGGTCGTGCGGCGGCAGTTCTGGTGGTG  
 AGGGTGCGATCGTTGGGATTCTGTGGTGGCGTCATCGGTGTAGGAACGGATATCGGTG  
 GCTCGATTGAGTGCAGGCGCGCTCAACTTCCTGTACGGTCTAAGGCCGAGTACG  
 GGCGGCTGCCGATGCAAAGATGGCGAACAGCATGGAGGGTCAGGAGAGGTCACACA  
 CGCTGTGCGGGCCGATTACGCACTCTGTTGAGGGTGAGTCCTTCGCCCTCTTCTCT  
 TTTCTGTCTATAACCAGGCCTCCACTGTCTCTCTCTTTCTTGCTTTTATACTATATA  
 CGAGACCCGCGAGTCACTGATGAAGTATGTTAGACCTCCGCTCTTCACCAAATCCGT  
 CCTCGGTGAGGACCATGGAATACGACTCCAAGGTCATCCCCATGCCCTGGCGCCA  
 GTCCGAGTCGGACATTATTGCCTCCAAGATCAAGAACGGCGGGCTCAATATCGGCTA  
 CTACAACCTCGACGGCAATGTCCTTCCACACCCTCCTATCCTGCGCGGCGTGGAAAC  
 CACCGTCGCCGCACTCGCCAAAGCCGGTCAACCGTGACCCCGTGACGCCATACAA  
 GCACGATTTGCGCCACGATCTCATCTCCATATCTACGCGGCTGACGGCAGCGCCGA  
 CGTAATGCGCGATATCAGTGCATCCGGCGAGCCGGCGATTCCAAATATCAAAGACCT  
 ACTGAAACCCGAACATCAAAGCTGTTAAACATGAACGAGCTCTGGGACACGCATCTCCA  
 GAAGTGGAATTACCAGATGGAGTACCTTGAGAAATGGCGGGAGGCTGAAGAAAAGGC  
 CGGGAAGGAACTGGACGCCATCATCGCGCCGATTACGCCTACCGCTGCGGTACGGCA  
 TGACCAGTTCGGTACTATGGGTATGCCTCTGTGATCAACCTGCTGGATTTACAGAG  
 CGTGGTTGTTCCGGTTACCTTTGCGGATAAGAACATCGATAAGAAGAATGAGAGTTT  
 CAAGGCCGTTAGTGAGCTTGATGCCCTCGTGCAGGAAGAGTATGATCCGGAGGCGTA  
 CCATGGGGCACCGGTTGCAGTGCAGGTTATCGGACGGAGACTCAGTGAAGAGAGGAC  
 GTTGGCGATTGCAGAGGAAGTGGGGAAGTTGCTGGGAAATGTGGTACTCCATAGCT  
 AATAAGTGTGAGATAGCAATTTGCACAAGAAATCAATACCAGCAACTGTAAATAAGC  
 GCTGAAGTGACCATGCCATGCTACGAAAGAGCAGAAAAAACCTGCCGTAGAACCGA  
 AGAGATATGACACGCTTCCATCTCTCAA

FIG. 3D

GGAAGAATCCCTTCAGGGTTGCGTTTCCAGTCTAGACACGTATAACGGCACAAGTGT  
 CTCTCACCAAATGGGTTATATCTCAAATGTGATCTAAGGATGGAAAGCCAGAATCT  
 AGGCCTATTAATATTCCGGAGTATACGTAGCCGGCTAACGTAAACAACCGGTACCTC  
 TAGAACTATAGCTAGCATGCGCAAATTTAAAGCGCTGATATCGATCGCGCGCAGATC  
 CATATATAGGGCCCCGGTTATAAATTACCTCAGGTCGACGTCCCATGGCCATTTCGAAT  
 TCGTAATCATGGTCATAGCTGTTTCCCTGTGTGAAATTGTTATCCGCTCACAATTCAC  
 CACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGC  
 TAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGCG  
 TGCCAGCTGCATTAATGAATCGGCCAACCGCGGGGAGAGGCGGTTTTCGCTATTGGG  
 CGCTCTTCCGCTTCCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCCGGCTGCGGCGA  
 GCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAAC  
 GCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCC  
 GCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGA  
 CGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAACCAGGCGTTTTCCC  
 CCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTG  
 TCCGCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTAT  
 CTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGGTT  
 CAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGA  
 CACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGGTAT  
 GTAGGCGGTGCTACAGAGTTCCTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGA  
 ACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGT  
 AGCTCTTGATCCGGCAAACAAACCCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAG  
 CAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACG  
 GGGTCTGACGCTCAGTGGAACGAAACTCACGTTAAGGGATTTTGGTCATGAGATTA  
 TCAAAAAGGATCTTACCTAGATCCTTTTAAATTAATAAATGAAGTTTTAAATCAATC  
 TAAAGTATATATAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA  
 CCTATCTCAGCGATCTGTCTATTTTCGTTTCAATAGTTGCCTGACTCCCCGTCGTG  
 TAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCG  
 CGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCCAGCCAGCCGGAAGG  
 GCCGAGCGCAGAAGTGGTCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGT  
 TGCCGGGAAGCTAGAGTAAGTAGTTCCGCAAGTTAATAGTTTGGCAACGTTGTTGCC  
 ATTGCTACAGGCATCGTGGTGTACGCTCGTCTTGGTATGGCTTCATTCAGCTCC  
 GGTTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTT  
 AGCTCCTTCGTCCTCCGATCGTTGTGAGAAGTAAGTTGGCCGAGTGTATCACTC  
 ATGGTTATGGCAGCACTGCATAATTCTTACTGTGATGCCATCCGTAAGATGCTTT  
 TCTGTGACTGGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCG  
 AGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTA  
 AAAGTGCTCATCATTTGAAAACGTTCTTCGGGGCGAAAACCTCAAGGATCTTACCG  
 CTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAAGTATCTTACGATCT  
 TTTACTTTACAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGAAAA  
 AAGGGAATAAGGCGCACCGAAATGTTGAATACTCATACTCTTCTTTTCAATAT  
 TATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATT  
 TAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGAC  
 GTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAAATAGGCGTATCACGAGG  
 CCCTTTCGTCCTCGCGGTTTTCGGTGATGA

FIG. 3E

CGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGC  
GGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCTG  
GGGCTGGCTTAAGTATGCGGCATCAGAGCAGATTGTAAGTGCAGAGTGCACCATAAAAT  
TGTAACGTTAATATTTTGTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATT  
TTTTAACCAATAGGCCGAAATCGGCAAATCCCTTATAAATCAAAGAATAGCCCGA  
GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAGAAGCTGGA  
CTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCACTACGTGAACC  
ATCACCCAAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCC  
TAAAGGGAGCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAA  
GGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTAC  
GCTGCGCGTAACCACCACACCCGCCGCTTAATGCGCCGCTACAGGGCGCGTACTA  
TGGTTGCTTTGACGTATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATAC  
CGCATCAGGCGCCATTGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTG  
CGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTA  
AGTTGGGTAACGCCAGGGTTTTCCAGTCACGACGTTGTAACGACGGCCAGTGCC



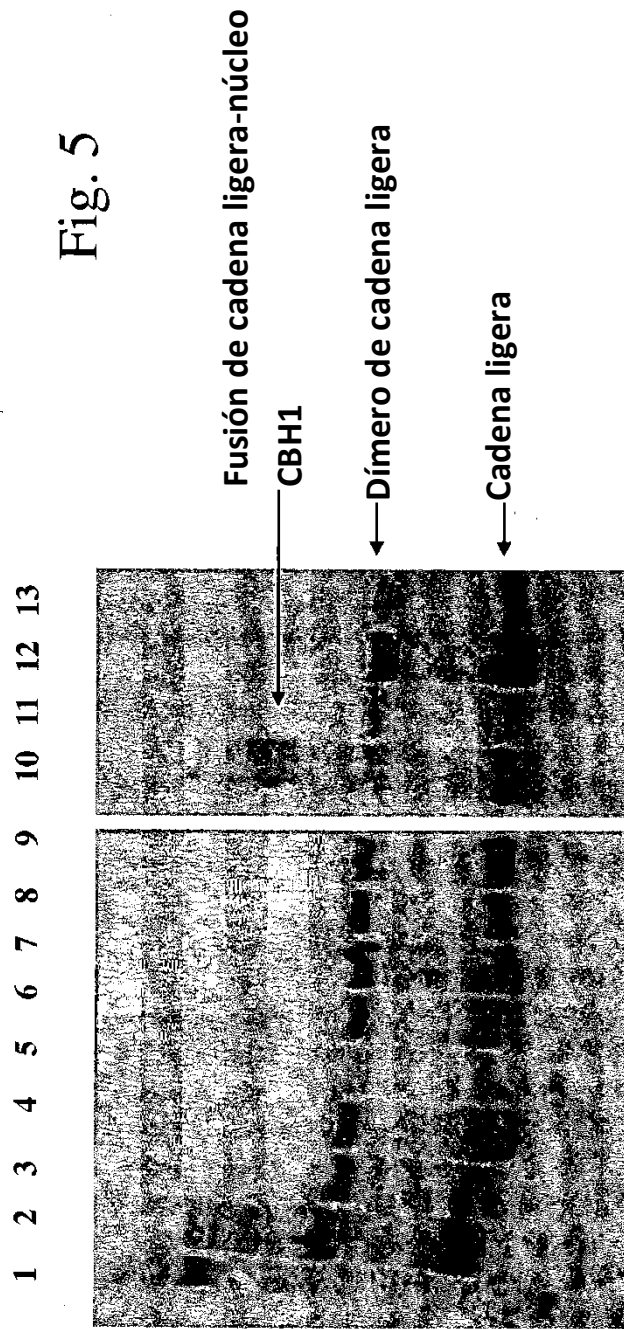


FIG. 6

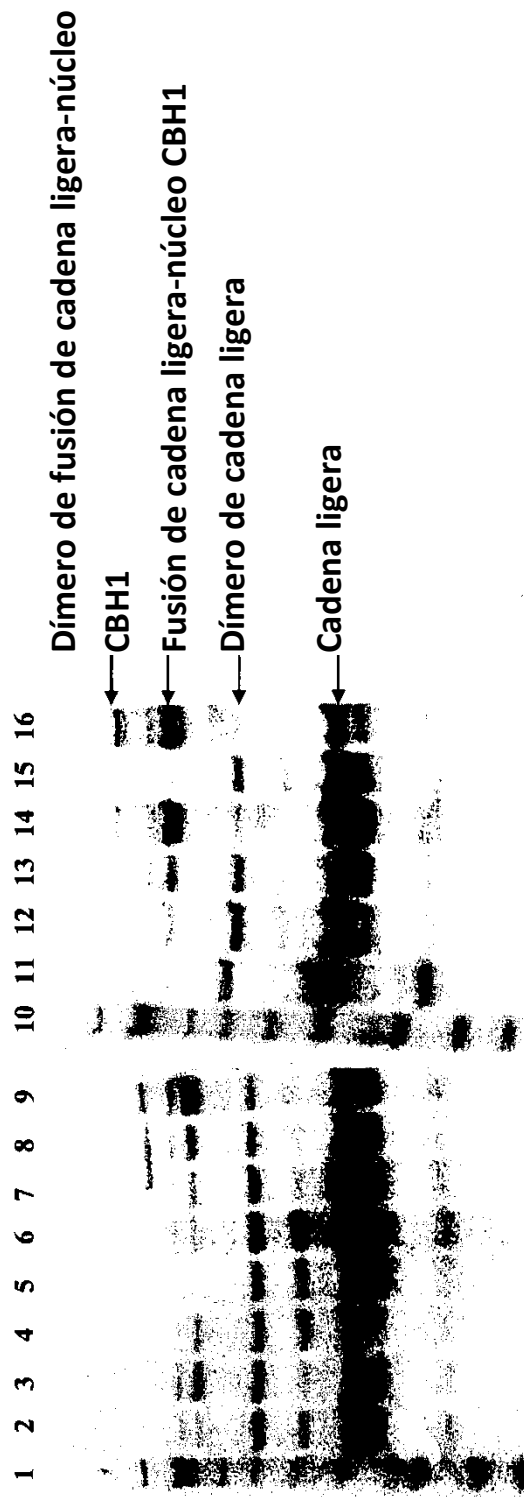
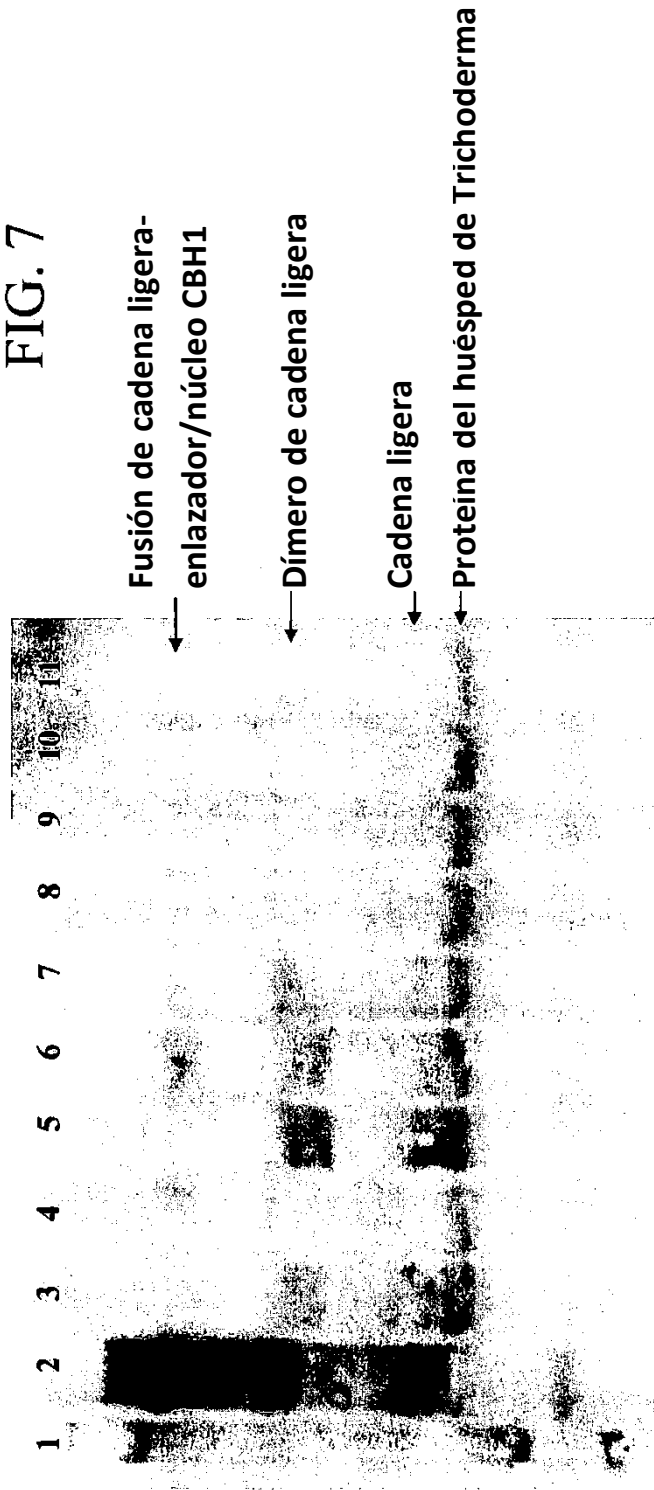


FIG. 7



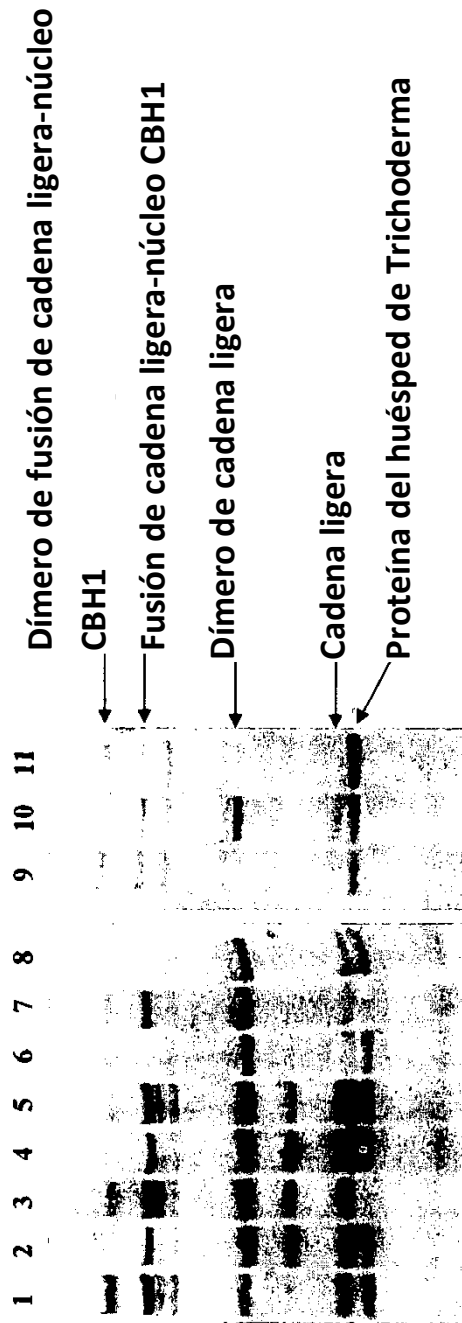


FIG. 8



FIG. 9

