

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 371 916**

51 Int. Cl.:
C12N 15/56 (2006.01)
C12N 9/42 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **04820192 .5**
96 Fecha de presentación: **07.12.2004**
97 Número de publicación de la solicitud: **1702981**
97 Fecha de publicación de la solicitud: **20.09.2006**

54 Título: **CELULASA TOLERANTE A LOS TENSIOACTIVOS Y PROCEDIMIENTO PARA CONVERTIR LA MISMA.**

30 Prioridad:
08.12.2003 JP 2003409692

45 Fecha de publicación de la mención BOPI:
11.01.2012

45 Fecha de la publicación del folleto de la patente:
11.01.2012

73 Titular/es:
Meiji Seika Pharma Co., Ltd.
4-16, Kyobashi 2-chome Chuo-ku
Tokyo 104-8002, JP

72 Inventor/es:
WATANABE, Manabu;
YANAI, Koji y
TSUYUKI, Yumiko

74 Agente: **Carpintero López, Mario**

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).

ES 2 371 916 T3

DESCRIPCIÓN

Celulasa tolerante a los tensioactivos y procedimiento para convertir la misma

Campo técnico

5 La presente invención se refiere a un procedimiento para modificar el extremo N terminal de una proteína que tiene actividad de endoglucanasa (en especial, una celulasa perteneciente a la familia 45 y que tiene actividad de endoglucanasa) a ácido piroglutámico, para convertir la proteína en una celulasa con actividad de endoglucanasa cuya reducción en presencia de un tensioactivo es pequeña, y está relacionada con la celulasa.

Antecedentes de la técnica

10 Se dice que la biomasa de celulosa es el recurso más abundante entre los recursos naturales, y por lo tanto en varios campos resulta deseable una aplicación eficiente de los sistemas de celulasas que descomponen la biomasa de celulosa. En este proceso de desarrollo, se han purificado y caracterizado varias celulasas, y además se han clonado diversos genes de celulasa, que se han clasificado en familias mediante el análisis de la homología de secuencias (véase referencia no relacionada con patentes 1). En otro aspecto, las celulasas se utilizan basándose en sus propiedades, en diversos campos industriales, en particular en el campo de procesamiento de tejidos. Por
15 ejemplo, se lleva a cabo un tratamiento con celulasa para mejorar el tacto y/o el aspecto de tejidos que contienen celulosa, o para un "biolavado" que imparte un aspecto de "lavado a la piedra" a un tejido que contiene celulosa coloreado, proporcionando por consiguiente al tejido variaciones de color localizadas. Además, en el procedimiento para fabricar liocel, la celulasa se usa para retirar la pelusa generada durante el procedimiento de la superficie del tejido. Al respecto, el liocel es un tejido de celulosa regenerado derivado de la pulpa de madera, que recientemente
20 ha atraído la atención por sus propiedades (tales como la alta resistencia o la absorción de agua) y como un procedimiento de producción que provoca menos contaminación ambiental.

Hasta ahora, se ha considerado que la celulasa decompone la celulosa mediante el efecto colaborativo de varias enzimas, es decir, el efecto sinérgico. El grupo de celulasas que consiste en varias enzimas contiene enzimas que tienen propiedades inadecuadas para el campo del procesamiento de tejidos (tales como una enzima que disminuye la resistencia de la fibra). Por consiguiente, se ha llevado a cabo un intento para separar componentes enzimáticos adecuados para el procesamiento de tejidos del grupo de las celulasas, y para producir los componentes enzimáticos, utilizando técnicas de separación de proteínas y/o técnicas de ingeniería genética. En particular, se han sometido a un estudio serio las celulasas derivadas de microorganismos que pertenecen a los hongos filamentosos tales como el género *Trichoderma* o el género *Humicola*. Por ejemplo, como componentes de celulasa, se aislaron
25 CBH I, EG V (véase referencia de patente 1), NCE2, NCE4 y NCE5 en el género *Humicola*, y CBH I, CBH II, EG II y EG III en el género *Trichoderma*, y de este modo, se puede producir las preparaciones de celulasa que contienen como componentes principales uno o más componentes de celulasa específicos apropiados para cada propósito mediante la preparación de enzimas sobreexpresadas o enzimas monocomponentes usando técnicas de ingeniería genética. Además, se ha aclarado que las celulasas que pertenecen a la familia 45, tales como NCE4 (véase referencia de patente 2), NCE5 (véase referencia de patente 3), RCE1 (véase referencia de patente 4) o STCE1 (véase Solicitud Internacional N° PCT/JP2004/15733), son muy útiles en los campos anteriores.
35

En otro aspecto más, cuando se usan las celulasas como un detergente para la ropa, no solamente se desea una mejora cuantitativa de los componentes de celulasa usados sino también una cualitativa. Más particularmente, un detergente para la ropa contiene diversos tensioactivos, y una disolución obtenida mediante la solubilización del detergente para la ropa en agua es alcalina (pH 10 a pH 11). Por consiguiente, es necesario que las celulasas contenidas en un detergente para la ropa sean resistentes a diversos tensioactivos, en condiciones alcalinas. Como un informe en el que una se suprime la reducción de la actividad en presencia de un tensioactivo, Otzen, D. E. y col. informaron que cuando se introdujo una mutación en la secuencia interna de aminoácidos de Cel45 derivada de *Humicola insolens*, la actividad de la misma a pH 7 en presencia de benceno sulfonato de alquilo lineal (LAS) era
40 aproximadamente 3,3 veces superior a la del tipo natural (véase referencia no relacionada con patentes 2). Sin embargo, se encontró que la supresión de la reducción de la actividad en presencia del tensioactivo, proporcionada por la mutación, está limitada a Cel45 o a sus proteínas homólogas, y que no es aplicable a las endoglucanasas pertenecientes a la familia 45 que tienen baja homología con Cel45.

(referencia de patente 1) Publicación Internacional WO91/17243.

50 (referencia de patente 2) Publicación Internacional WO98/0366.

(referencia de patente 3) Publicación Internacional WO01/90375.

(referencia de patente 4) Publicación Internacional WO00/24879.

(referencia no relacionada con patentes 1) Henrissat B., Bairoch A. Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.* 316: 695-696 (1996).

(referencia no relacionada con patentes 2) Daniel E. Otzen, Lars Christiansen, Martin Schulein. A comparative study of the unfolding of the endoglucanase Cel45 from *Humicola insolens* in denaturant and surfactant. *Protein Sci.* 8: 1878-1887 (1999).

El documento WO 94/07998 se refiere a una variante de celulosa de una celulosa original que comprende un dominio de unión a celulosa, un dominio catalíticamente activo y una región que une los dos dominios mencionados. Para mejorar las propiedades de la variante de celulosa, uno o más restos de aminoácidos neutros en la superficie del dominio catalíticamente activo están sustituidos por uno o más restos de aminoácidos cargados negativamente, o uno o más restos de aminoácidos cargados positivamente en la superficie del dominio catalíticamente activo están sustituidos por uno o más restos de aminoácidos neutros o cargados negativamente. Además, uno o más restos de aminoácidos hidrófobos están sustituidos por uno o más restos de aminoácidos no hidrófobos, o en el que uno o más restos de aminoácidos están sustituidos por prolina. *J. Mol. Biol.* (1997) 272, 383-397 describe la estructura cristalina del dominio catalítico central de la endoglucanasa I de *Trichoderma reesei* y una comparación con enzimas relacionadas.

Un objeto de la presente invención es para proporcionar un procedimiento para convertir una proteína que tiene una actividad de endoglucanasa (en particular, una proteína perteneciente a la familia 45 y que tiene una actividad de endoglucanasa) en una proteína que tiene una actividad de endoglucanasa cuya reducción en presencia de un tensioactivo es pequeña; un vector que se utiliza en el procedimiento; una proteína que tiene una actividad de endoglucanasa cuya reducción en presencia de un tensioactivo es pequeña; y un polinucleótido que codifique la misma. Otro objeto de la presente invención es proporcionar, utilizando lo anterior, un microorganismo que produce de manera eficiente una proteína útil como una enzima para el lavado de ropa.

Medios para solucionar el problema

Los presentes inventores realizaron estudios intensivos, y como resultado, encontraron que las proteínas en las que se añadió ácido piroglutámico (a continuación denominado algunas veces pQ en el presente documento) o un péptido que contenía pQ al extremo N terminal de cada proteína perteneciente a la familia 45 y que tenía una endoglucanasa, es decir, celulasas con adición en el extremo N terminal, tenían una actividad de endoglucanasa que no presentaba una reducción significativa de la actividad en presencia de un tensioactivo, en comparación con las celulasas de tipo natural.

La característica de que se mantenga una actividad de endoglucanasa alta en presencia de un agente tensioactivo aniónico es particularmente útil en el caso de enzimas para el lavado de ropa, pero no se conoce una celulasa de tal tipo. Además, no se ha informado que las funciones inherentes a una enzima puedan mantenerse en presencia de un tensioactivo por medio de la adición a la misma de ácido piroglutámico o de un péptido que lo contenga.

En otro aspecto de la presente invención, con respecto a todas las celulasas en las que el aminoácido N terminal no está protegido y es deseable el mantenimiento de la actividad en presencia de un agente tensioactivo, resulta posible, mediante la adición a la celulasa de ácido piroglutámico o de un péptido que lo contiene, suprimir la reducción de la actividad en presencia de un tensioactivo. Las celulasas a las que se añade ácido piroglutámico o un péptido que lo contiene no están particularmente limitadas, pero resultan de preferencia las celulasas pertenecientes a la familia 45.

Por consiguiente, la presente invención incluye lo siguiente:

(1) un procedimiento para suprimir la reducción de la actividad de endoglucanasa en presencia de un tensioactivo, caracterizado por la modificación de una proteína que tiene la actividad de endoglucanasa en la que el extremo N terminal es un aminoácido diferente de ácido piroglutámico, a una proteína que tiene el extremo N terminal de ácido piroglutámico;

(2) el procedimiento de (1), en el que la modificación se lleva a cabo mediante la adición de ácido piroglutámico o un aminoácido que puede convertirse en ácido piroglutámico, o un péptido que tiene el extremo N terminal de ácido piroglutámico o un aminoácido que puede convertirse en ácido piroglutámico, al extremo N terminal de la proteína que tiene la actividad de endoglucanasa en la que el extremo N terminal es un aminoácido diferente de ácido piroglutámico;

(3) el procedimiento de (1), en el que la modificación se lleva a cabo mediante la sustitución del aminoácido N terminal o una región N terminal de la proteína que tiene la actividad de endoglucanasa en la que el extremo N terminal es un aminoácido diferente de ácido piroglutámico, por ácido piroglutámico o un aminoácido que puede convertirse en ácido piroglutámico, o un péptido que tiene el extremo N terminal de ácido piroglutámico o un aminoácido que puede convertirse en ácido piroglutámico;

(4) el procedimiento de uno cualquiera de (1) a (3), en el que la proteína que tiene la actividad de endoglucanasa en la que el extremo N terminal es un aminoácido diferente de ácido piroglutámico es una celulasa perteneciente a la familia 45;

(5) una proteína modificada que tiene actividad de endoglucanasa en la que el aminoácido N terminal se convierte en ácido piroglutámico por una modificación de aminoácidos;

(6) la proteína modificada de (5), que se puede obtener por el procedimiento de uno cualquiera de (1) a (4);

(7) una proteína seleccionada del grupo que consiste en:

- 5 (a) una proteína que comprende la secuencia de aminoácidos de la ID. SEC. Nº 2, 4, 38 o 40;
- (b) una proteína modificada que comprende una secuencia de aminoácidos en la que uno o varios aminoácidos se han eliminado, sustituido, insertado o añadido en la secuencia de aminoácidos de la ID. SEC. Nº 2, 4, 38 o 40, y que tiene una actividad de endoglucanasa cuya reducción de presencia de un tensioactivo es pequeña; y
- 10 (c) una proteína homóloga que comprende una secuencia de aminoácidos que tiene al menos el 85% de homología con una proteína que comprende la secuencia de aminoácidos de la ID. SEC. Nº 2, 4, 38 o 40, y que tiene una actividad de endoglucanasa cuya reducción en presencia de un tensioactivo es pequeña;

Efectos de la invención

De acuerdo con la presente invención, es posible producir de manera eficaz una celulasa nueva que es útil como una enzima para el lavado de ropa y que tiene una actividad de endoglucanasa cuya reducción en presencia de un tensioactivo es pequeña.

Mejor modo de llevar a cabo la invención

Proteína perteneciente a la familia 45 y que tiene actividad de endoglucanasa (denominada a continuación algunas veces en el presente documento "celulasa perteneciente a la familia 45").

Familia 45

20 El término "familia 45" como se utiliza en el presente documento, significa una proteína clasificada en la familia 45 según el análisis de grupos hidrófobos de enzimas activadoras de hidratos de carbono por B. Henrissat y A. Bairoch [Henrissat B., Bairoch A. Updating the sequence-based classification of glycosyl hydrolases. Biochem. J. 316: 695-696 (1996)].

Proteínas que tienen actividad de endoglucanasa

25 La expresión "proteína que tiene una actividad de endoglucanasa" como se utiliza en el presente documento, significa una enzima que presenta una actividad de endoglucanasa, es decir, endo-1,4-β-glucanasa (EC3.2.1.4), que hidroliza el enlace β-1,4-glucopiranosilo en β-1,4-glucano.

Tensioactivo

30 El término "tensioactivo", como se utiliza en el presente documento, es un componente detergente contenido en un detergente para la ropa, y que se clasifica ampliamente en tensioactivos aniónicos, catiónicos y no iónicos. Los tensioactivos aniónicos se utilizan comúnmente. Como tensioactivos aniónicos utilizados de preferencia en la presente invención, se pueden citar, por ejemplo, benceno sulfonato de alquilo lineal (denominado a continuación algunas veces en el presente documento "LAS").

Actividad de endoglucanasa

35 La expresión "actividad de endoglucanasa (denominada a continuación en el presente documento 'EGU') como se utiliza en el presente documento se define como una actividad enzimática obtenida por medio de la medición de la disminución de la viscosidad de una disolución de carboximetilcelulosa según el siguiente procedimiento. Como disolución de sustrato, se disolvió carboximetilcelulosa (Hercules) en un tampón Tris-HCl 0,1 mol/l (pH 10,0) (concentración final = 3,5%). A la disolución de sustrato (5 ml) previamente calentada a 40 °C durante 10 minutos, se le añadieron 0,15 ml de una disolución de la enzima, y a continuación se mezcló todo bien para llevar a cabo la reacción a 40 °C durante 30 minutos. La viscosidad de la mezcla de reacción se midió mediante un viscosímetro de tipo R (RE100; TOKI Sangyo Co., LTD) a 40 °C. Se define "1 unidad" de la actividad enzimática como una cantidad de enzima que reduce la viscosidad inicial al 1/2, en cada condición de la reacción. Como tensioactivo aniónico se utilizó benceno sulfonato de alquilo lineal (Wako Pure Chemical Industries, Co., Ltd.) y se añadió a la disolución de carboximetilcelulosa hasta una concentración final de 800 ppm.

Supresión de la reducción de la actividad de endoglucanasa en presencia de tensioactivo

50 "Que tiene actividad de endoglucanasa cuya reducción en presencia de un tensioactivo es pequeña", como se utiliza en el presente documento, significa que, cuando la proteína en la que está modificado el extremo N terminal (proteína del tipo modificación N terminal) de acuerdo con la presente invención se compara con la proteína original antes de realizar la modificación (denominada a continuación en el presente documento simplemente proteína

original) con respecto a la actividad de endoglucanasa en presencia del tensioactivo, la actividad de endoglucanasa de la proteína del tipo modificación N terminal es mayor que la de la proteína original.

Proteína original

La proteína original que se puede aplicar al procedimiento de la presente invención no está particularmente limitada, con la condición de que se trate de una proteína que tiene una actividad de endoglucanasa en la que el extremo N terminal es un aminoácido diferente de ácido piroglutámico. Como la proteína original, se pueden citar, por ejemplo, una celulasa (por ejemplo, endoglucanasa, celobiohidrolasa o β -glucosidasa) y resulta de preferencia una celulasa perteneciente a la familia 45. En este sentido, siempre que la proteína original tenga al menos una actividad de endoglucanasa, puede ser una proteína que tenga sólo la actividad de endoglucanasa, o una proteína que tenga otra u otras actividades enzimáticas además de la actividad de endoglucanasa. Además, la proteína original puede ser una proteína de origen natural o una proteína modificada genéticamente.

Fuente original de la celulasa perteneciente a la familia 45

La celulasa perteneciente a la familia 45 puede generarse por medio de técnicas de ingeniería genética de uso común, tales como técnicas de ADN recombinante o técnicas de síntesis de polipéptidos, o se pueden obtener a partir de una cepa natural aislada. Además, las celulasas incluyen variantes de celulasas de tipo natural pertenecientes a la familia 45. La celulasa perteneciente a la familia 45 se puede obtener a partir de un microorganismo tal como los hongos filamentosos o zygomycetes. Como hongos filamentosos, se pueden citar, por ejemplo, microorganismos pertenecientes al género *Humicola* (tal como *Humicola insolens*), al género *Trichoderma* (tal como *Trichoderma viride*), al género *Staphylotrichum* (tal como *Staphylotrichum coccosporum*) o del género *Myriococcum* (tal como *Myriococcum thermophilum*). Más concretamente, las celulasas derivadas de género *Humicola* incluyen, por ejemplo, CBH I, EG V, NCE2, NCE4 y NCE5; las celulasas derivadas del género *Trichoderma* incluyen, por ejemplo, CBH I, CBH II, EG II y EG III; las celulasas derivadas del género *Staphylotrichum* incluyen, por ejemplo, STCE1 y STCE3; y las celulasas derivadas del género *Myriococcum* incluyen, por ejemplo, MTE1. Como zygomycetes, se pueden citar, por ejemplo, microorganismos pertenecientes al género *Rhizopus* (tal como *Rhizopus oryzae*), al género *Mucor* (tal como *Mucor circinelloides*) o al género *Phycomyces* (tal como *Phycomyces nitens*). Más concretamente, se pueden citar, por ejemplo, RCE I, RCE II o RCE III derivadas de *Rhizopus oryzae*, MCE I o MCE II derivadas de *Mucor circinelloides*; o PCE I derivada de *Phycomyces nitens* (documento WO 00/24879).

Ácido piroglutámico o péptido que contiene ácido piroglutámico

El término "ácido piroglutámico" como se utiliza en el presente documento significa ácido piroglutámico generado por ciclación de la glutamina o el ácido glutámico del extremo N terminal de una proteína madura. El ácido piroglutámico tiene la característica de que el grupo amino N terminal no está expuesto. La formación de piroglutamato se puede realizar in vivo o in vitro. In vivo, un polinucleótido que codifica una proteína modificada, que está genéticamente diseñada para que el extremo N terminal de una proteína madura sea glutamina o ácido glutámico, puede expresarse en una célula huésped para obtener una proteína de ciclación de piroglutamato. In vitro, una proteína que tiene el extremo N terminal de glutamina o ácido glutámico pueden ser tratada con una disolución ácida tal como una disolución de ácido fórmico para obtener una proteína que tiene el extremo N terminal de ácido piroglutámico.

El término "péptido" como se utiliza en el presente documento, significa un compuesto que consiste en uno o en varios aminoácidos en el que los aminoácidos están polimerizados por medio de enlaces peptídicos. Por lo tanto, la expresión "péptido que contiene ácido piroglutámico" como se utiliza en el presente documento, significa un péptido en el que el aminoácido N terminal es ácido piroglutámico. El péptido que contiene ácido piroglutámico consiste en dos o más, (varios) aminoácidos reticulados, por ejemplo, de 2 a 40 aminoácidos, de preferencia de 2 a 30 aminoácidos, de más preferencia de 2 a 20 aminoácidos, de más preferencia aún de 2 a 10, y de más preferencia aún de 2 a 5, de mayor preferencia de 2 a 4 aminoácidos. Los aminoácidos no están particularmente limitados, con la condición de que puedan ser utilizados por los expertos en la técnica para el propósito indicado.

En el procedimiento de la presente invención, un procedimiento para modificar una proteína original a una proteína que tiene el extremo N terminal de ácido piroglutámico no está particularmente limitado, con la condición de que la modificación de la proteína pueda llevarse a cabo. Como procedimiento, se pueden citar, por ejemplo, técnicas de ingeniería genética o técnicas químicas.

De acuerdo con las técnicas de ingeniería genética, una proteína original puede ser modificada, por ejemplo, mediante la realización de una adición y/o sustitución de un aminoácido o secuencia de aminoácidos adecuada por ingeniería genética. La modificación de proteínas puede llevarse a cabo por la adición por ingeniería genética de ácido piroglutámico o de un péptido que tiene el extremo N terminal de ácido piroglutámico, al extremo N terminal de una proteína original (es decir, una proteína que tiene una actividad de endoglucanasa en la que el extremo N terminal no es ácido piroglutámico). Más particularmente, la adición por medio de ingeniería genética puede comprender, por ejemplo, las etapas de:

- (1) añadir un polinucleótido que codifica un aminoácido que puede convertirse en ácido piroglutámico (tal como ácido glutámico o glutamina) o un polinucleótido que codifica un péptido que tiene el extremo N terminal de un

aminoácido que puede convertirse en ácido piroglutámico, al extremo 5' terminal de un polinucleótido que codifica una proteína original, y

(2) expresar el polinucleótido resultante en un huésped en el que puede llevarse a cabo la formación del piroglutamato del aminoácido N terminal.

5 La modificación de proteínas puede llevarse a cabo por la sustitución por ingeniería genética del aminoácido N terminal o de una región N terminal de una proteína original por ácido piroglutámico o un péptido que tiene el extremo N terminal de ácido piroglutámico. Más particularmente, la sustitución por medio de ingeniería genética puede comprender, por ejemplo, las etapas de:

10 (1) sustituir el extremo 5' terminal o una región que contiene el mismo de un polinucleótido que codifica una proteína original por un polinucleótido que codifica un aminoácido que puede convertirse en ácido piroglutámico (tal como ácido glutámico o glutamina) o un polinucleótido que codifica un péptido que tiene el extremo N terminal de un aminoácido que puede convertirse en ácido piroglutámico, y

(2) expresar el polinucleótido resultante en un huésped en el que puede llevarse a cabo la formación del piroglutamato del aminoácido N terminal.

15 Como formas de realización mediante el uso de técnicas químicas, se pueden citar, por ejemplo,

(a) una forma de realización en la que el ácido piroglutámico (o un péptido que tiene el extremo N terminal de ácido piroglutámico) se añade directamente al extremo N terminal de una proteína original;

20 (b) una forma de realización en la que un aminoácido que puede convertirse en ácido piroglutámico (o un péptido que tiene el extremo N terminal de un aminoácido que puede convertirse en ácido piroglutámico) se añade químicamente al extremo N terminal de una proteína original, y a continuación se lleva a cabo químicamente la formación del piroglutamato del aminoácido N terminal; o

(c) una forma de realización en la que el aminoácido N terminal de una proteína original que tiene el extremo N terminal de un aminoácido que puede convertirse en ácido piroglutámico se convierte químicamente en ácido piroglutámico.

25 **Procedimiento para añadir ácido piroglutámico o un péptido que contiene ácido piroglutámico al lado N terminal de la celulasa perteneciente a la familia 45**

El procedimiento de modificación se ilustrará además por medio de una forma de realización que utiliza celulasa perteneciente a la familia 45. Un procedimiento para añadir ácido piroglutámico o un péptido que contiene ácido piroglutámico al lado N terminal de la celulasa perteneciente a la familia 45 puede llevarse a cabo mediante técnicas de ingeniería genética. En una producción de celulasas de uso común, se puede unir una región codificadora de un polinucleótido que codifica la celulasa deseada de manera operativa entre un promotor y un terminador funcional en un huésped, tal como un hongo filamentoso, y a continuación puede introducirse la cassette de expresión resultante en el huésped. Además, se puede añadir un polinucleótido que codifica una secuencia señal de secreción funcional en la célula huésped en la cassette. Cuando se introduce la cassette en la célula huésped, la celulasa deseada es secretada en un medio, y se puede recoger fácilmente. En este caso, puede añadirse un aminoácido deseado al extremo N terminal de la celulasa, mediante la adición de un polinucleótido que codifica el aminoácido inmediatamente secuencia abajo de la secuencia señal de secreción. Además, la modificación del grupo amino del aminoácido N terminal se puede llevar a cabo utilizando una secuencia señal de secreción en un huésped. Por ejemplo, en cbh1 o cbh2 derivadas de *Trichoderma viride*, o NCE2 o NCE5 derivadas de *Humicola insolens*, el extremo N terminal es piroglutamato formado, la modificación puede llevarse a cabo mediante el uso de estas secuencias señal para la secreción y la expresión en *Trichoderma viride* o *Humicola insolens*. De acuerdo con la forma de realización de preferencia, la celulasa modificada perteneciente a la familia 45 preparada como se describió anteriormente presenta una característica ventajosa, es decir, tiene una actividad de endoglucanasa cuya reducción en presencia de un tensioactivo es pequeña.

45 Según otra forma de realización, todo puede ser sintetizado químicamente, dentro del ámbito de los conocimientos técnicos comunes de los expertos en la técnica. En este caso, la síntesis puede llevarse a cabo utilizando parte de una proteína de origen natural.

Proteína de la presente invención

La proteína de la presente invención se prepara mediante la modificación del extremo N terminal de una proteína original que tiene una actividad de endoglucanasa a ácido piroglutámico [por ejemplo, preparada por la obtención de celulasa perteneciente a la familia 45 y la adición de ácido piroglutámico (pQ) o un péptido que contiene ácido piroglutámico a un lado del aminoácido N terminal de la proteína madura de la misma] y tiene una actividad de endoglucanasa cuya reducción en presencia de un tensioactivo es pequeña. Además, la presente invención incluye una proteína que se puede preparar por el procedimiento descrito anteriormente y tiene una actividad de endoglucanasa cuya reducción en presencia de un tensioactivo es pequeña.

Más concretamente, la proteína de la presente invención incluye una proteína seleccionada del grupo que consiste en las siguientes proteínas:

(a) una proteína que comprende la secuencia de aminoácidos de la ID. SEC. Nº 2, 4, 38 o 40;

5 (b) una proteína modificada que comprende una secuencia de aminoácidos en la que uno o varios aminoácidos se han eliminado, sustituido, insertado o añadido en la secuencia de aminoácidos de la ID. SEC. Nº 2, 4, 38 o 40, y que tener una actividad de endoglucanasa cuya reducción en presencia de un tensioactivo es pequeña; y

(c) una proteína homóloga que comprende una secuencia de aminoácidos que tiene al menos el 85% de homología con una proteína que comprende la secuencia de aminoácidos de la ID. SEC. Nº 2, 4, 38 o 40, y que tiene una actividad de endoglucanasa cuya reducción en presencia de un tensioactivo es pequeña.

10 La secuencia de aminoácidos de la ID. SEC. Nº 2 es la secuencia de aminoácidos de NCE4 modificada en el extremo N terminal (véase Ejemplo A2) en la que se añadió un péptido que consiste en cinco aminoácidos (N terminal: ácido piroglutámico) al extremo N terminal de la endoglucanasa NCE4 derivada de *Humicola insolens* MN200-1.

15 La secuencia de aminoácidos de la ID. SEC. Nº 4 es la secuencia de aminoácidos de STCE1 modificada en el extremo N terminal (véase Ejemplo B2) en la que se sustituyó el aminoácido N terminal (Ala) de la endoglucanasa STCE1 derivada de *Staphylotrichum coccosporum* IFO 31817 por un péptido que consiste en cuatro aminoácidos (N terminal: ácido piroglutámico).

20 La secuencia de aminoácidos de la ID. SEC. Nº 38 es la secuencia de aminoácidos de la STCE1 modificada en el extremo N terminal (véase Ejemplo B3) en la que se añadió ácido piroglutámico al extremo N terminal de la endoglucanasa STCE1 derivada de *Staphylotrichum coccosporum* IFO 31817.

La secuencia de aminoácidos de la ID. SEC. Nº 40 es la secuencia de aminoácidos de la STCE1 modificada en el extremo N terminal (véase Ejemplo B4) en la que se añadió un péptido que consiste en cuatro aminoácidos (N terminal: ácido piroglutámico) al extremo N terminal de la endoglucanasa STCE1 derivada de *Staphylotrichum coccosporum* IFO 31817.

25 El término "proteína modificada", como se utiliza en el presente documento, significa una proteína que comprende una secuencia de aminoácidos en la que uno o una pluralidad de aminoácidos (por ejemplo, uno o varios aminoácidos) se han eliminado, sustituido, insertar o añadido en la secuencia de aminoácidos de ID. SEC. Nº 2, 4, 38 o 40, y que tiene una actividad de endoglucanasa cuya reducción en presencia de un tensioactivo es pequeña. El número de aminoácidos a modificar, tal como "eliminar, sustituir, insertar o añadir", es de preferencia de 1 a 30, de más preferencia de 1 a 10, de mayor preferencia de 1 a 6.

30 Además, la proteína modificada incluye una proteína que comprende una secuencia de aminoácidos en la que uno o varios aminoácidos son sustituidos de manera conservadora en la secuencia de aminoácidos de la ID. SEC. Nº 2, 4, 38 o 40, y que tiene una actividad de endoglucanasa cuya reducción en presencia de un tensioactivo es pequeña. La expresión "sustitución conservadora", como se utiliza en el presente documento, significa que uno o varios restos de aminoácidos contenidos en una proteína son sustituidos con diferentes aminoácidos que tienen propiedades químicas similares de modo que las actividades de la proteína no se modifican sustancialmente. Como sustitución conservadora, se pueden citar, por ejemplo, la sustitución de un resto de aminoácido hidrófobo por otro resto de aminoácido hidrófobo, o la sustitución de un resto de aminoácido polar por otro resto de aminoácido polar que tiene la misma carga. Los aminoácidos que tienen propiedades químicas similares y que pueden ser sustituidos de manera conservadora entre sí son conocidos por los expertos en la técnica. Más en particular, como aminoácidos no polares (hidrófobos), se pueden citar, por ejemplo, alanina, valina, isoleucina, leucina, prolina, triptófano, fenilalanina o metionina. Como aminoácidos polares (neutros), se pueden citar, por ejemplo, glicina, serina, treonina, tirosina, glutamina, asparagina o cisteína. Como aminoácidos básicos que tienen una carga positiva, se pueden citar, por ejemplo, arginina, histidina o lisina. Como aminoácidos ácidos que tienen una carga negativa, se pueden citar, por ejemplo, ácido aspártico o ácido glutámico.

35 La expresión "proteína homóloga", como se utiliza en el presente documento, significa una proteína que comprende una secuencia de aminoácidos que tiene al menos el 85% (de preferencia el 90% o más, de mayor preferencia el 95% o más) de homología (identidad de secuencia) con una proteína que comprende la secuencia de aminoácidos de ID. SEC. Nº 2, 4, 38 o 40, y que tiene una actividad de endoglucanasa cuya reducción en presencia de un tensioactivo es pequeña. La homología como se utiliza en el presente documento se muestra como el valor (identidad) calculado por FASTA3 [Science, 227, 1435-1441 (1985); Proc. Natl. Acad. Sci. USA, 85, 2444-2498 (1988); <http://www.ddbj.nig.ac.jp/E-mail/homology-j.html>], un programa de búsqueda de homología conocido, de conformidad con los parámetros por defecto.

50 Como se describió anteriormente, en la "proteína que consiste en la secuencia de aminoácidos de la ID. SEC. Nº 2" de la presente invención, la porción excepto el péptido N terminal que consiste en cinco aminoácidos deriva de la endoglucanasa NCE4. Además, en la "proteína que consiste en la secuencia de aminoácidos de la ID. SEC. Nº 4, 38 o 40" de la presente invención, la porción excepto el aminoácido o péptido N terminal deriva de la endoglucanasa

STCE1. Las endoglucanasas NCE4 y STCE1 pertenecen a la familia 45. Como endoglucanasa conocida perteneciente a la familia 45, se pueden citar, por ejemplo, NCE5 derivada del género *Humicola* (documento WO01/90375).

En las figuras 1 y 2, se muestra una alineación de las secuencias de aminoácidos de la endoglucanasa STCE1 [péptido señal (ID. SEC. Nº 43) y proteína madura (ID. SEC. Nº 44)], la endoglucanasa NCE4 [péptido señal (ID. SEC. Nº 45) y proteína madura (ID. SEC. Nº 46)] y la endoglucanasa NCE5 [péptido señal (ID. SEC. Nº 47) y proteína madura (ID. SEC. Nº 48)]. La figura 1 muestra la alineación de la mitad N terminal, y la Figura 2 muestra la mitad C terminal. El símbolo "*" en las figuras 1 y 2 indica un aminoácido común al de STCE1.

Como se muestra en las figuras 1 y 2, las endoglucanasas pertenecientes a la familia 45 contienen el dominio catalítico (1º a 207º) como un dominio común, y a veces contienen la región conectora (208º a 258º) y/o el dominio de unión a la celulosa (CDB) (259º a 295º). En este sentido, los números entre paréntesis después de los dominios anteriores representan el número de aminoácidos en la secuencia de aminoácidos (ID. SEC. Nº 44) de la endoglucanasa STCE1.

En cada región, hay muchos aminoácidos conservadores entre dos o más endoglucanasas en el dominio catalítico y el dominio de unión a la celulosa, pero no se observa ninguna región conservadora notable en la región del conector. Las regiones que contienen muchos aminoácidos conservadores (por ejemplo, el dominio catalítico o el dominio de unión a la celulosa, en particular, el dominio catalítico), o aminoácidos comunes contenidos en las regiones son consideradas regiones o aminoácidos importantes para la actividad enzimática de las endoglucanasas (tales como STCE1 o NCE4). Por lo tanto, cuando se lleva a cabo una modificación de aminoácidos (por ejemplo, supresión, sustitución, inserción y/o adición, en particular, sustitución conservadora) en una región o un aminoácido distinto de regiones o aminoácidos tan importantes, puede obtenerse con una alta posibilidad y sin excesiva experimentación una proteína modificada u homóloga que mantiene la actividad de la enzima.

Además, incluso cuando tiene lugar en las regiones que contienen muchos aminoácidos conservadores, una modificación de un aminoácido o aminoácidos no comunes entre dos o más endoglucanasas a aminoácido(s) diferente(s) [de preferencia aminoácido(s) que son similares y que pueden ser sustituidos de manera conservadora] probablemente podrá mantener la actividad de la enzima. Por lo tanto, por una de tales modificaciones, puede obtenerse con una alta probabilidad y sin excesiva experimentación una proteína modificada u homóloga que mantiene la actividad de la enzima.

En este sentido, aunque uno o más aminoácidos comunes en la región que contiene muchos aminoácidos conservadores se modifican a aminoácido(s) diferente(s), la actividad de la enzima a veces se mantiene. En particular, en una modificación a aminoácido(s) que son similares y que pueden ser sustituidos de manera conservadora, la posibilidad es mayor. La proteína modificada u homóloga de la presente invención incluye una proteína en la que uno o más aminoácidos contenidos en cualquier región, tal como el dominio catalítico, la región conectora o el dominio de unión a la celulosa, se han modificado, con la condición de que presente una actividad de endoglucanasa.

Polinucleótido que codifica la proteína de la presente invención

Se puede proporcionar un polinucleótido que codifica una proteína que comprende la secuencia de aminoácidos de la ID. SEC. Nº 2, 4, 38 o 40, o una de sus proteínas modificadas u homólogas (a continuación, en el presente documento denominada como la proteína de la presente invención). Cuando se da la secuencia de aminoácidos de una proteína, se puede seleccionar fácilmente una secuencia de nucleótidos que codifica la secuencia de aminoácidos, y por lo tanto pueden seleccionarse varias secuencias de nucleótidos que codifican la proteína de la presente invención. El término "polinucleótido" como se utiliza en el presente documento, incluye ADN y ARN, y resulta de preferencia el ADN.

El polinucleótido incluye un polinucleótido seleccionado del grupo que consiste en los siguientes polinucleótidos:

(a) un polinucleótido que comprende la secuencia de nucleótidos de la ID. SEC. Nº 1, 3, 37 o 39;

(b) un polinucleótido que comprende una secuencia de nucleótidos en el que uno o varios nucleótidos se han eliminado, sustituido, insertado o añadido en la secuencia de nucleótidos de la ID. SEC. Nº 1, 3, 37 o 39, y que codifica una proteína con una actividad de endoglucanasa cuya reducción en presencia de un tensioactivo es pequeña; y

(c) un polinucleótido que hibrida en condiciones rigurosas con un polinucleótido que consiste en la secuencia de nucleótidos de la ID. SEC. Nº 1, 3, 37 o 39, y que codifica una proteína con una actividad de endoglucanasa cuya reducción en presencia de un tensioactivo es pequeña.

En la secuencia de nucleótidos que se describe en el punto anterior (b), el número de nucleótidos a eliminar, sustituir, insertar o añadir, es por ejemplo, de 1 a 90, de preferencia de 1 a 30, de más preferencia de 1 a 18, de mayor preferencia de 1 a 9.

La expresión "condiciones rigurosas", como se utiliza en el presente documento, significa las siguientes condiciones. De acuerdo con un protocolo adjunto de un sistema de detección y marcación directa de ADN/ARN ECL (Amersham), después de que un polinucleótido a probar se prehibrida a 42 °C durante una hora, se añade un polinucleótido que consiste en la secuencia de nucleótidos de la ID. SEC. N° 1, 3, 37 o 39, y la hibridación se lleva a cabo a 42 °C durante 14 a 16 horas. Después de la hibridación, se repite dos veces un tratamiento de lavado con 0,5 X SSC (1 X SSC; citrato de sodio 15 mmol/l, cloruro de sodio 150 mmol/l) con SDS al 0,4% y urea 6 mol/l a 42 °C durante 20 minutos, y se lleva a cabo dos veces un tratamiento de lavado con 5 X SSC a temperatura ambiente durante 5 minutos.

El polinucleótido incluye un polinucleótido de origen natural. Además, se puede sintetizar en su totalidad. Por otra parte, la síntesis puede llevarse a cabo utilizando parte del polinucleótido de origen natural. Por lo general, el polinucleótido de la presente invención pueden obtenerse por selección de una biblioteca genómica derivada de un microorganismo deseado según un procedimiento habitual de uso común en la ingeniería genética, por ejemplo, utilizando una sonda de ADN adecuada en base a la información de una secuencia de aminoácidos parcial.

Producción de la proteína de la presente invención

La proteína de la presente invención se pueden producir en células huésped mediante la transformación de la célula huésped con una molécula de polinucleótido (en particular, en forma de vector de expresión) que comprende un fragmento de polinucleótido que codifica la proteína de manera que la molécula de polinucleótido pueda ser replicada y el gen se pueda expresar en la célula huésped.

Se proporciona un vector de expresión que comprende un fragmento de polinucleótido que codifica la proteína de la presente invención, para que el fragmento de polinucleótido pueda ser replicado y la proteína se pueda expresar en un microorganismo huésped.

El vector de expresión se puede construir en base a un vector auto-replicante (tal como un plásmido), que existe como un elemento extracromosómico y que puede replicarse de manera independiente de la replicación de los cromosomas. Como alternativa, el vector de expresión puede ser un vector que está integrado en el cromosoma del microorganismo huésped y que se replica junto con los cromosomas, cuando el huésped se transforma con el vector. La construcción del vector de la presente invención puede llevarse a cabo por procedimientos habituales o procedimientos comúnmente utilizados en la ingeniería genética.

Para expresar una proteína que tiene una actividad deseada mediante la transformación de un microorganismo huésped con el vector de expresión de la presente invención, es preferible que el vector de expresión contenga, por ejemplo, un polinucleótido capaz de controlar la expresión, o un marcador genético para seleccionar los transformantes, además del fragmento de polinucleótido.

Como polinucleótido capaz de regular la expresión genética, se pueden utilizar varias señales para la regulación de la transcripción o de la traducción, tales como un promotor, un terminador o un polinucleótido que codifica un péptido señal para la secreción. La unión de estos polinucleótidos y la inserción de los mismos a un vector pueden llevarse a cabo por un procedimiento habitual.

El promotor no está particularmente limitado, con la condición de que muestre actividad transcripcional en un microorganismo huésped. El promotor se puede obtener como un polinucleótido que regula la expresión de un gen que codifica una proteína igual o diferente de la que deriva del microorganismo huésped. Por ejemplo, se puede usar un promotor tal como un operón de lactosa o un operón de triptófano en *Escherichia coli*; un promotor de un gen de la alcohol deshidrogenasa, un gen de la fosfatasa ácida, un gen de utilización de la galactosa o un gen de la gliceraldehído 3-fosfato deshidrogenasa en una levadura; y un promotor de un gen de la α amilasa, un gen de la glucoamilasa, un gen de la celobiohidrolasa o un gen de la gliceraldehído 3-fosfato deshidrogenasa en un hongo filamentoso.

El péptido señal no está particularmente limitado, con la condición de que contribuya a la secreción de proteínas en un microorganismo huésped. El péptido señal se puede obtener como un polinucleótido derivado de un gen que codifica una proteína igual o diferente de la que deriva del microorganismo huésped.

El marcador de selección puede ser adecuadamente seleccionado de acuerdo con el procedimiento de selección de transformantes. Como marcador de selección, en la presente invención se puede utilizar por ejemplo, un gen de resistencia a fármacos o un gen que complementa una mutación auxotrofa. Cuando el huésped es una bacteria, se puede utilizar por ejemplo, un gen de resistencia a la ampicilina, un gen de resistencia a la kanamicina o un gen de resistencia a la tetraciclina. Cuando el huésped es una levadura, se puede utilizar por ejemplo, un gen de biosíntesis de triptófano (*trp1*, *trpC*), un gen de biosíntesis de uracilo (*ura3*), un gen de biosíntesis de leucina (*leu2*) o un gen de biosíntesis de histidina (*his3*). Cuando el huésped es un hongo filamentoso, se puede utilizar por ejemplo, un gen de resistencia a la destomicina, un gen de utilización de nitratos (*niaD*), un gen de biosíntesis de arginina (*argB*), un gen de biosíntesis de uracilo (*pyr4*), un gen resistencia a la higromicina, un gen de resistencia a bialafos, un gen de resistencia a la bleomicina o un gen de resistencia a la aureobasidina.

Un sistema huésped-vector que puede ser utilizado en la presente invención no está particularmente limitado. Por ejemplo, se puede usar un sistema que utiliza *E. coli*, *Actinomyces*, levaduras u hongos filamentosos, o un sistema para la expresión de una proteína de fusión utilizando tales microorganismos. Cuando se utiliza una levadura como huésped, se pueden citar, por ejemplo, un microorganismo que pertenece al género *Saccharomyces*, *Hansenula* o *Pichia*, de preferencia *Saccharoztyces cerevisiae*. Cuando se utiliza un hongo filamentoso como huésped, se puede utilizar cualquier hongo filamentoso, de preferencia un microorganismo perteneciente al género *Humicola*, *Trichoderma*, *Aspergillus*, *Acremonium* o *Fusarium*, de más preferencia un microorganismo perteneciente al género *Humicola* o *Trichoderma*. Más concretamente, se pueden utilizar *Humicola insolens*, *Humicola thermoidea*, *Trichoderma viride*, *Trichoderma reesei*, *Trichoderma longibrachiatum*, *Staphylotrichum coccosporum*, *Aspergillus niger*, *Aspergillizs oryzae*, *Fusarium oxysporum* o *Acremonium cellulolyticus*, de preferencia *Humicola insolens* o *Trichoderma viride*.

La transformación de un microorganismo con el vector de expresión puede llevarse a cabo según un procedimiento habitual, tal como un procedimiento que utiliza iones de calcio, un procedimiento que utiliza iones de litio, un procedimiento de electroporación, un procedimiento con PEG, un procedimiento de *Agrobacterium* o un procedimiento con pistola de partículas, y se puede seleccionar un procedimiento adecuado según el huésped a transformar.

Se cultiva el transformante, y el cultivo resultante se utiliza para obtener la proteína de la presente invención de interés. El transformante se puede cultivar según un procedimiento habitual por medio de selección adecuada, por ejemplo, un medio o condiciones del cultivo.

El medio se puede complementar con, por ejemplo, fuentes de carbono o fuentes de nitrógeno que puede ser anabolizadas o utilizadas, respectivamente, por el transformante de la presente invención, sales inorgánicas, varias vitaminas, varios aminoácidos tales como ácido glutámico o asparagina, oligonutrientes, tales como nucleótidos, o agentes de selección tales como antibióticos. Además, se pueden añadir de manera adecuada sustancias orgánicas y/o inorgánicas que ayudan al crecimiento del transformante de la presente invención, o promueven la producción de la proteína de la presente invención. Además, de ser necesario, se puede utilizar un medio natural o un medio sintético que contenga adecuadamente otros nutrientes.

Se puede utilizar cualquier tipo de fuente de carbono y fuente de nitrógeno en el medio, con la condición de que puedan ser utilizadas por el transformante de la presente invención. Como fuente de carbono anabolizable, se pueden utilizar varios hidratos de carbono, por ejemplo, sacarosa, almidón, glicerina, glucosa, fructosa, maltosa, lactosa, celulosa, celobiosa, dextrina, aceites animales o aceites vegetales, o hidrolizados de los mismos. En general, la concentración de los mismos en el medio es de preferencia del 0,1% al 5%. Como fuente de nitrógeno utilizable, se pueden usar por ejemplo, componentes, animales o vegetales o extractos de los mismos, tales como peptona, extracto de carne, licor de maceración de maíz o polvo de soja desgrasada, sales de amonio de ácidos orgánicos tales como sales de amonio del ácido succínico o sales de amonio del ácido tartárico, urea u otros varios compuestos que contienen nitrógeno, tales como ácidos inorgánicos u orgánicos. Como sales inorgánicas, se pueden utilizar adecuadamente por ejemplo, las que pueden producir sodio, potasio, calcio, magnesio, cobalto, cloro, ácido fosfórico, ácido sulfúrico u otros iones.

Se puede utilizar adecuadamente cualquier medio que contenga otros componentes, por ejemplo, células, exudados o extractos de microorganismos tales como levaduras o polvos finos de plantas, con la condición de que no interfieran con el crecimiento de los transformantes de la presente invención ni con la producción y acumulación de la proteína de la presente invención. Cuando se cultiva una cepa mutante que tiene un requerimiento nutricional, por lo general se añade al medio una sustancia para satisfacer sus requerimientos nutricionales. En este sentido, es posible que tal nutriente no deba añadirse necesariamente cuando se utiliza un medio que contiene sustancias naturales.

Las condiciones de cultivo, tales como el componente del medio, la fluidez del medio, la temperatura de incubación, la velocidad de agitación o la tasa de aireación, se pueden seleccionar y controlar adecuadamente según el transformante a utilizar y las condiciones externas a fin de obtener los resultados preferibles. Si se formara espuma en un medio líquido, puede utilizarse adecuadamente un agente antiespumante, tal como aceites de silicona, aceites vegetales, aceites minerales o tensioactivos.

La proteína de la presente invención acumulada en el cultivo resultante se encuentra en el transformante y en el filtrado del cultivo. Por consiguiente, es posible recuperar la proteína de la presente invención, tanto del filtrado del cultivo como de las células transformantes mediante la separación del cultivo en cada fracción por medio de centrifugación.

La proteína de la presente invención se puede recuperar del filtrado del cultivo según un procedimiento habitual. Los procedimientos de recuperación de la proteína de la presente invención del cultivo se pueden realizar de manera individual, en combinación en un determinado orden, o en varias veces. Por ejemplo, se pueden utilizar la extracción por filtración, la centrifugación, la diálisis, la concentración, el secado, la congelación, la adsorción y la desorción, también medios para la separación basada en la diferencia de solubilidad en diversos disolventes (tales como la precipitación, el desalado, la cristalización, la recristalización, la disolución inversa o la cromatografía).

Además, la proteína de la presente invención se puede obtener a partir del cultivo dentro del transformante. Por ejemplo, la extracción del cultivo (por ejemplo, por tratamiento de molienda o alteración irreversible por presión), la recuperación (por ejemplo, filtración o centrifugación) o la purificación (por ejemplo, desalado o precipitación del disolvente) puede llevarse a cabo utilizando un procedimiento habitual.

5 La sustancia bruta resultante puede ser purificada según un procedimiento habitual, por ejemplo, por medio de una cromatografía en columna utilizando un vehículo tal como el dextrano, la celulosa, agarosa, polímeros sintéticos o gel de sílice. Se puede obtener una proteína pura de la presente invención de interés a partir del cultivo del transformante utilizando los procedimientos anteriormente mencionados, ya sea por separado o en una combinación adecuada.

10 Como se describió anteriormente, se puede proporcionar el proceso para la producción de la nueva proteína de la presente invención. El cultivo del transformante (incluyendo las condiciones de cultivo) se puede realizar de una manera sustancialmente similar al cultivo del huésped original utilizado para preparar el transformante. Como procedimiento para la recuperación de la proteína de interés después del cultivo del transformante, se pueden llevar a cabo los procedimientos de uso común.

15 **Cepas depositadas**

Se depositó la cepa de *Escherichia coli* JM109 transformada con el vector de expresión pEGD01 de la presente invención (FERM BP-5973) en el país en el Instituto Nacional Depositario de Organismos de Patentes Internacionales de Ciencia y Tecnología Industrial Avanzada [(Nombre anterior) Instituto Nacional de Biocencia y Agencia de Tecnología Humana de Ciencia y Tecnología Industrial (Dirección: AIST Tsukuba Central 6, 1-1, Higashi 1-chome Tukuba-shi, Ibaraki-ken 305-8566, Japón)] el 12 de julio de 1996, y fue trasladada a un depósito internacional el 13 de junio de 1997. El número de depósito internacional (un número entre paréntesis [] después del número de depósito internacional se incluye un número de depósito nacional) es FERM BP-5973 [FERM P-15729].

Se depositó la cepa de *Escherichia coli* JM109 transformada con el vector de expresión pMKD01 de la presente invención (FERM BP-5974) en el país en el Instituto Nacional Depositario de Organismos de Patentes Internacionales de Ciencia y Tecnología Industrial Avanzada [(Nombre anterior) Instituto Nacional de Biocencia y Agencia de Tecnología Humana de Ciencia y Tecnología Industrial (Dirección: AIST Tsukuba Central 6, 1-1, Higashi 1-chome Tukuba-shi, Ibaraki-ken 305-8566, Japón)] el 12 de julio de 1996, y fue trasladada a un depósito internacional el 13 de junio de 1997. El número de depósito internacional (un número entre paréntesis [] después del número de depósito internacional se incluye un número de depósito nacional) es FERM BP-5974 [FERM P-15730].

Se depositó la cepa de *Escherichia coli* JM109 transformada con el vector de expresión pCB1-M2XR (FERM BP-6045) en el país en el Instituto Nacional Depositario de Organismos de Patentes Internacionales de Ciencia y Tecnología Industrial Avanzada [(Nombre anterior) Instituto Nacional de Biocencia y Agencia de Tecnología Humana de Ciencia y Tecnología Industrial (Dirección: AIST Tsukuba Central 6, 1-1, Higashi 1-chome Tukuba-shi, Ibaraki-ken 305-8566, Japón)] el 9 de septiembre de 1996, y fue trasladada a un depósito internacional el 11 de agosto de 1997. El número de depósito internacional (un número entre paréntesis [] después del número de depósito internacional se incluye un número de depósito nacional) es FERM BP-6045 [FERM P-15840]. Al respecto, el plásmido pCB1-M2 de la presente invención no solo se puede obtener por el procedimiento descrito en los EJEMPLOS, sino también por auto-unión del plásmido pCB1-M2XR previamente digerido con la enzima de restricción XbaI.

Se depositó la cepa de *Humicola insolens* MN200-1 (FERM BP-5977), que puede ser utilizada como huésped para el vector de expresión de la presente invención, en el país en el Instituto Nacional Depositario de Organismos de Patentes Internacionales de Ciencia y Tecnología Industrial Avanzada [(Nombre anterior) Instituto Nacional de Biocencia y Agencia de Tecnología Humana de Ciencia y Tecnología Industrial (Dirección: AIST Tsukuba Central 6, 1-1, Higashi 1-chome Tukuba-shi, Ibaraki-ken 305-8566, Japón)] el 15 de julio de 1996, y fue trasladada a un depósito internacional el 13 de junio de 1997. El número de depósito internacional (un número entre paréntesis [] después del número de depósito internacional se incluye un número de depósito nacional) es FERM BP-5977 [FERM P-15736].

Se depositó la cepa de *Trichoderma viride* MC300-1 (FERM BP-6047), que puede ser utilizada como huésped para el vector de expresión de la presente invención, en el país en el Instituto Nacional Depositario de Organismos de Patentes Internacionales de Ciencia y Tecnología Industrial Avanzada [(Nombre anterior) Instituto Nacional de Biocencia y Agencia de Tecnología Humana de Ciencia y Tecnología Industrial (Dirección: AIST Tsukuba Central 6, 1-1, Higashi 1-chome Tukuba-shi, Ibaraki-ken 305-8566, Japón)] el 9 de septiembre de 1996, y fue trasladada a un depósito internacional el 11 de agosto de 1997. El número de depósito internacional (un número entre paréntesis [] después del número de depósito internacional se incluye un número de depósito nacional) es FERM BP-6047 [FERM P-15842].

Se depositó la cepa de *Escherichia coli* DH5 α transformada con el vector de expresión pUC118-STCEex (FERM BP-10127) en el país en el Instituto Nacional Depositario de Organismos de Patentes Internacionales de Ciencia y Tecnología Industrial Avanzada (Dirección: AIST Tsukuba Central 6, 1-1, Higashi 1-chome Tukuba-shi, Ibaraki-ken 305-8566, Japón) el 1 de diciembre de 2003, y fue trasladado a un depósito internacional el 15 de septiembre de

2004. El número de depósito internacional (un número entre paréntesis [] después del número de depósito internacional se incluye un número de depósito nacional) es FERM BP-10127 [FERM P-19602]. Al respecto, el plásmido pUC118-STCEex es un plásmido obtenido por la inserción del gen de STCE en el sitio BamHI del plásmido pUC118. El gen de la endoglucanasa STCE1 presente en el fragmento BamHI contiene un intrón y tiene la secuencia de la ID. SEC. N° 5.

Ejemplos

La presente invención ahora se ilustrará con más detalle por medio de los siguientes ejemplos, pero no significa que esté limitada a los mismos.

Ejemplo A1: Construcción de vectores de expresión para expresar celulasa modificada en su extremo N terminal y celulasa de tipo natural en *Humicola insolens*

(1) Construcción del vector de expresión pMKD01 para expresar celulasa modificada en su extremo N terminal

El plásmido pUC118 (Takara Bio) fue digerido utilizando la enzima BamHI. Se generaron extremos romos en el fragmento resultante con un DNA blunting kit (Takara Bio) y se auto-unió con un DNA ligation kit (Takara Bio) para obtener pUC118-BN. El pUC118-BN fue digerido con SphI, se generaron extremos romos y se auto-unió para obtener pUC118BSN. A continuación, se obtuvo un fragmento PstI-XbaI de 3,4 kb que comprendía el gen de celulasa NCE2 de *Humicola insolens* MN200-1 (FERM BP-5977) según el procedimiento descrito en la Publicación de Patente Japonesa no examinada (Kokai) N° 8-126492. El fragmento se unió con pUC118BSN previamente digerido con la misma enzima de restricción para obtener pUC118BSN-PX. Se introdujo la mutación dirigida al sitio en pUC118BSN-PX con un sistema de mutagénesis in vitro Sculptor (Amersham Bioscience) para obtener el plásmido pM21. Como oligonucleótidos para la mutagénesis, se usaron los oligonucleótidos MNC-02 y MNC-03.

MNC-02: 5'-GAGCGCCAGAACTGTGGATCCACTTGGTGAGCAATG-3' (SEC. N° 6)

MNC-03: 5'-TCCGCCGTTCTGAGCGGATCCAGGCGTTTGGCGCG-3' (SEC. N° 7)

El plásmido pM21 fue digerido utilizando la enzima BamHI y se unió con un fragmento de PCR digerido con BamHI de un gen de celobiohidrolasa (NCE3) derivado de *Humicola insolens* MN200-1 para obtener el plásmido pKM04. El fragmento de PCR de NCE3 se amplificó utilizando ADN genómico (documento WO98/03640) de *Humicola insolens* MN200-1 (FERM BP-5977) como molde y los siguientes cebadores MKA-05 y MKA-06.

MKA-05: 5'-GCCGCCAGCAGGCGGGATCCCTCACCACCGAGAGG-3' (SEC. N° 8)

MKA-06: 5'-TGATCGTCGAGTCAGGGATCCAGAATTTACAGGCAC-3' (SEC. N° 9)

A continuación, se utilizaron un promotor y un terminador (Mullaney, EJ y col., Mol. Gen. Genet., 199, 37-45, 1985) del gen *trpC* derivado de *Aspergillus nidulans* para preparar un gen capaz de expresar un gen de resistencia a la destomicina en *Humicola insolens*. El gen resultante se unió con el plásmido pKM04 previamente digerido con XbaI para construir el vector de expresión pMKD01 (FERM BP-5974) para expresar la celulasa modificada en el extremo N terminal. En pMKD01, se introdujeron los sitios de reconocimiento de BamHI a 10 pb secuencia abajo del extremo N terminal de la proteína madura del gen de NCE2 y a 3 pb secuencia abajo del codón de parada del mismo, y por lo tanto se pueden añadir cinco restos de aminoácidos al extremo N terminal de una celulasa deseada perteneciente a la familia 45.

Tabla 1

<u>BamHI</u>	<u>BamHI</u>
5'— GAGCGCCAGAACTGTGGATCCCTC—	—TGCCTGTAAg c g g a t c c a g g—3' (SEC. N° 10)
GluArgGlnAsnCysGlySerLeu—	—CysLeu (SEC. N° 11)
↑ extremo N terminal de proteína madura	

(2) Construcción del vector de expresión pJD01 para expresar la celulasa de tipo natural

Se introdujo la mutación dirigida al sitio utilizando el oligonucleótido pMN-Bam en el plásmido pM21 obtenido en el punto anterior (1) para obtener el plásmido pM21-m-A1.

pMN-Bam:

5'-GGTCAAACAAGTCTGTGCGGATCCCTGGGACAAGATGGCCAAGTTCTTCCTTAC-3' (SEC. N°: 12)

El plásmido pM21-m-A1 fue digerido con las enzimas de restricción HindIII y BamHI, y se recuperó el fragmento resultante de aproximadamente 1 kb. A continuación, el pMKD01 obtenido en el Ejemplo A1(1) fue digerido con HindIII y BamHI, y se recuperó un fragmento de aproximadamente 7 kb. Estos fragmentos se unieron para obtener el

plásmido pJD01. En este vector, se introdujeron los sitios de reconocimiento de BamHI a 15 pb secuencia arriba del sitio de iniciación de la traducción del gen de NCE2 y a 3 pb secuencia abajo del codón de parada del mismo, y por consiguiente se expresó una celulasa deseada perteneciente a la familia 45 desde el punto de iniciación de la transcripción.

Tabla 2

<u>Bam</u> HI		<u>Bam</u> HI
5'-tgc <u>ggatcct</u> tgggacaagATGGCC-	—	-CCGTTCTGAgc <u>ggatc</u> cagg-3' (SEC. Nº 13)
	Me tAla—	—Pr oPhe (SEC. Nº 14)

Ejemplo A2: Construcción de vectores de expresión para expresar NCE4 de tipo natural y NCE4 modificada en su extremo N terminal, transformación de *Humicola insolens* con los mismos y evaluación de las actividades

(1) Construcción de los vectores de expresión

Como un vector de expresión para NCE4 de tipo natural, se unió la región codificadora del gen de NCE4 con pJD01 obtenido en el Ejemplo A1(2) para la construcción del pN2EX-N4. En primer lugar, se amplificó la región codificadora del gen de NCE4 por medio de PCR utilizando ADN genómico de *Humicola insolens* MN200-1 como molde y NCE4-Ne y NCE4-Ce como cebadores. El producto de la PCR resultante, de aproximadamente 0,9 kpb, se digirió con BamHI y se unió con pJD01 previamente digerido con la misma enzima de restricción para la construcción del pN2EX-N4.

NCE4-Ne: 5'-GGGGGGATCCTGGGACAAGATGCGTTCTCCCTCTC-3' (SEC. Nº 15)

NCE4-Ce: 5'-GGGGGGATCCCTGCGTTTACAGGCACTGATGG-3' (SEC. Nº 16)

Como un vector de expresión para NCE4 modificada en su extremo N terminal, se unió la región codificadora de la proteína madura, en la que se había eliminado la secuencia señal para la secreción en el gen de NCE4, con el pMKD01 (FERM BP-5974) obtenido en el Ejemplo A1(1) para la construcción del pEGD01 (FERM BP-5973). En primer lugar, se amplificó la región codificadora de la parte de la proteína madura del gen de NCE4 por medio de PCR utilizando ADN genómico de *Humicola insolens* MN200-1 como molde y NCE4-Ns y NCE4-Cs como cebadores. El producto de PCR resultante, de aproximadamente 0,8 kpb, se digirió con BamHI y se unió con pMKD01 previamente digerido con la misma enzima de restricción para la construcción del pEGD01.

NCE4-Ns: 5'-CCGGTGTGGCCGGATCCGCTGATGGCAAG-3' (SEC. Nº 17)

NCE4-Cs: 5'-TAAGGCCCTCAAGATCCCTGCGTCTACAG-3' (SEC. Nº 18)

(2) Transformantes que contienen NCE4 de tipo natural o NCE4 modificada en su extremo N terminal

Se cultivó *Humicola insolens* MN200-1 en un medio S (glucosa al 3,0%, extracto de levadura al 2,0%, polipeptona al 0,1%, cloruro de calcio al 0,03% y sulfato de magnesio al 0,03%; pH 6,8) a 37 °C durante 24 horas. Los micelios resultantes se recogieron por centrifugación, se lavaron con sacarosa 0,5 mol/l y se suspendieron en 0,5 mol/l de disolución de sacarosa que contenía β-glucuronidasa 3 mg/ml (Sigma), quitinasa 1 mg/ml y zymolyase 20T 1 mg/ml (Seikagaku Corp.). La suspensión se agitó suavemente a 30 °C durante 60 a 90 minutos para generar protoplastos. Los protoplastos resultantes se recogieron por filtración y centrifugación, se lavaron con un tampón SUTC (sacarosa 0,5 mol/l, cloruro de calcio 10 mmol/l y Tris-HCl 10 mmol/l, pH 7,5) y se resuspendieron en SUTC para preparar una suspensión de protoplastos.

A 100 µl de la suspensión de protoplastos, se le añadieron 10 µl de cada disolución de ADN (pEGD01 o pN2EX-N4). Tras dejar en reposo cada uno en hielo durante 5 minutos, se añadieron 400 µl de una disolución de PEG (polietilenglicol 4000 al 60%, cloruro de calcio 10 mmol/l y Tris-HCl 10 mmol/l, pH 7,5) y todo se dejó en reposo sobre hielo durante 20 minutos. Después de lavar la suspensión de protoplastos resultante con el tampón SUTC, se cubrió la suspensión de protoplastos con un medio YMG (glucosa al 1%, extracto de levadura al 0,4%, extracto de malta al 0,2% y agar al 1%, pH 6,8) que contenía higromicina B 200 µg/ml junto con agar blando YMG. La placa se incubó a 37 °C durante 5 días para obtener los transformantes como colonias.

(3) Comparación de las actividades de endoglucanasa de NCE4 de tipo natural y NCE4 modificada en su extremo N terminal en los sobrenadantes de los cultivos

Los transformantes resultantes se cultivaron según el procedimiento descrito en el documento WO98/03667, para obtener los sobrenadantes de los cultivos. Los sobrenadantes de los cultivos se sometieron a SDS-PAGE para seleccionar los sobrenadantes de cultivo en los que se detectaba con claridad la banda de NCE4 de aproximadamente 43 kDa. Se midieron las actividades de endoglucanasa en el sobrenadante del cultivo a pH 10 en

ausencia y con presencia de LAS, y se compararon los grados de reducción en la actividad en presencia de LAS. Los resultados se muestran en la Tabla 3. En la tabla 3, se representan las actividades de endoglucanasa por medio de porcentajes considerando como 100 la actividad de endoglucanasa a pH 10.

Tabla 3

	Vector de expresión	EGU (%)	
		Sin LAS	Con LAS
NCE de tipo natural	pN2EX-N4	100	27,1
NCE modificada en su extremo N terminal	pEGD01	100	55,6

Como resultado, el sobrenadante del cultivo obtenido a partir del transformante de NCE4 modificada en su extremo N terminal mostró una pequeña reducción de la actividad en presencia de LAS. Tras una comparación de las actividades de endoglucanasa en presencia de LAS, se encontró que la actividad de endoglucanasa de NCE4 modificada en su extremo N terminal fue 2,1 veces mayor que la de NCE4 de tipo natural. Además, la actividad de endoglucanasa de *Humicola insolens* MN200-1 representó aproximadamente el 4% de los transformantes, por consiguiente se encontró que casi todas las actividades de endoglucanasa derivaron de las enzimas recombinantes expresadas por los vectores de expresión incorporados.

(4) Análisis de las secuencias de aminoácidos del extremo N terminal de NCE4 de tipo natural y NCE4 modificada en su extremo N terminal

Los sobrenadantes de los cultivos de los transformantes anteriores se sometieron a SDS-PAGE y las proteínas sometidas a electroforesis fueron transferidas eléctricamente a una membrana de PVDF (Immobilon-PSQ, Millipore). Las bandas de NCE4 de aproximadamente 43 kDa se recortaron de la mancha y se procesaron con el secuenciador Protein Sequencer Model 492 (Applied Biosystems) para analizar las secuencias de aminoácidos.

Como resultado, se confirmó que la secuencia de NCE4 de tipo natural obtenida del transformante pN2EX-N4 era coherente con la secuencia de ADN de NCE4 derivada de *Humicola insolens* MN200-1 descrita en el documento WO98/03667. Sin embargo, la secuencia de aminoácidos de NCE4 modificada en su extremo N terminal obtenida a partir del transformante de pEGD01 no pudo determinarse en el procedimiento anterior. Se utilizó una piroglutamato aminopeptidasa Pfu (Takara Bio) para eliminar el extremo N terminal modificado a partir de NCE4 modificada en su extremo N terminal, y a continuación se determinó el aminoácido (SEC. Nº 2).

Los resultados se resumieron en la Tabla 4, para comparar los aminoácidos del extremo N terminal y las secuencias de aminoácidos de extremo N terminal en NCE4 de tipo natural y en NCE4 con el extremo N terminal añadido, y se confirmó a partir de los resultados que cada NCE4 tenía la secuencia de aminoácidos del extremo N terminal esperada del correspondiente vector de expresión construido.

Tabla 4

	Aminoácido en el extremo N terminal	Secuencia de aminoácidos del extremo N terminal
NCE4 de tipo natural	Alanina (no modificado)	ADGKSTR (SEC. Nº 19)
NCE4 modificada en su extremo N terminal	Ácido piroglutámico (modificado)	pQNCGSADGKSTR (SEC. Nº 20)

Ejemplo B1: Construcción de vectores de expresión para *Trichoderma viride* para expresar celulasa modificada en su extremo N terminal y celulasa de tipo natural

(1) Construcción de los vectores de expresión

A partir de un fragmento de PstI de 7 kb que contenía el gen *cbh1* derivado de *Trichoderma viride* MC300-1 (obtenido según el documento WO98/11239) se clonaron un fragmento de HindIII de aproximadamente 3,1 kb que contenía una región promotora de *cbh1* y un fragmento de Sall de 2,7 kb que contenía una región terminadora de *cbh1* en el plásmido pUC119 para construir el plásmido pCB1-H4 y el plásmido pCB1-S1, respectivamente. Se infectó *E. coli* JM109 que contenía cada plásmido resultante con el fago ayudante M13KO7 para obtener cada ADN de hebra única. A los ADN de hebra única resultantes de los plásmidos pCB1-H4 y pCB1-S1 como moldes, se hibridaron oligonucleótidos fosforilados CBN-Stu y CBC-Xho, y se introdujeron mutaciones con un sistema de mutagénesis in vitro Sculptor (Amersham Bioscience), para la construcción de los plásmidos pCB1H4-19 y pCB1S1-17, respectivamente. Se unió un fragmento de aproximadamente 6 kb, que se había obtenido mediante la digestión del plásmido pCB1H4-19 con XbaI y XhoI, con un fragmento de aproximadamente 1,2 kb, que se había sido obtenido por digestión del plásmido pCB1S1-17 con XbaI y posteriormente sometiéndolo a digestión parcial con XhoI, para la construcción del pCB1-M.

CBN-Stu: 5'-GATACATGATGCGCAGGCCTTAGTCTGACTAGAATGC-3' (SEC. Nº 21)

CBC-Xho: 5'-GATCCTCAAGCTTTTGCTCGAGTACCTTACAAGCAC-3' (SEC. Nº 22)

A continuación, se clonó un fragmento de aproximadamente 2,7 kb, obtenido por digestión del pCB1-M con Sall, en el plásmido pUC119 para construir el plásmido pCB1-Salm. Al ADN de hebra única obtenido del plásmido pCB1-Salm se le hibridaron los oligonucleótidos fosforilados CB1-SmSph, CB1-Bam, y CB1-Pst, y se introdujeron las mutaciones con un sistema de mutagénesis in vitro Sculptor.

CB1-SmSph: 5'-GGAGGGTGCATGCCGACTGAGCCCCGGGCAGTAGCC-3' (SEC. Nº 23)

CB1-Bam: 5'-GCCGGGAGAGGATCCAGTGGAGG-3' (SEC. Nº 24)

CB1-Pst: 5'-GCTCGAGTACCTTACTGCAGGCACTGAGAG-3' (SEC. Nº 25)

A continuación, el plásmido pUC118 fue digerido utilizando las enzimas XbaI y EcoRI, se generaron extremos romos con el DNA blunting kit y se auto-unió para construir el plásmido pUC118-SBN. Al pUC118-SBN digerido con HindIII y Sall, se le insertó el fragmento del promotor de *cbh1* de aproximadamente 1,4 kb previamente digerido con las mismas enzimas de restricción. El plásmido resultante fue digerido con Sall, y se unió con la región codificadora de *cbh1* con las mutaciones introducidas y la región terminadora de aproximadamente 2,8 kb, para la construcción del pCB1-M2. Las mutaciones introducidas (excluyendo la mutación de BamHI) se muestran en la Tabla 5.

Tabla 5

Sa l I Stu I

5'-c t a g t c g a c t a a g g c c t g c g c a t c A T G T A T C A A A A G T T G G C C C T C A T C T C G G C C T T C T T G G C T A C T (SEC. Nº 26)

M e t T y r G l n L y s L e u A l a L e u I l l e S e r A l a P h e L e u A l a T h r (SEC. Nº 27)

Sma I Sph I Pst I Xho I Hind III (SEC. Nº 28)

5'-G C C C C G G C T C A G T C G G C A T G C A C C - - - C A G T C C T G C A G T A A g g t a c t c g a g c a a a a g c t t g a g - 3' (SEC. Nº 29)

A l a A r g A l a G l n S e r A l a C y s T h r - - - G l n C y s L e u G l n

↑ extremo N terminal de la proteína madura

Ejemplo B2: Construcción de vectores de expresión para expresar STCE1 modificada en su extremo N terminal y STCE1 de tipo natural, transformación de *Trichoderma viride* con los mismos y evaluación de las actividades

(1) Construcción de los vectores de expresión

Como un vector de expresión para STCE1 de tipo natural, se unió la región codificadora del gen de STCE1 con el plásmido pCB1-M2 obtenido en el Ejemplo B1 para la construcción del pCB-Ste.

En primer lugar, se amplificó la región codificadora del gen de STCE1 por medio de PCR utilizando el plásmido pUC118-STCEex (FERM BP-10127) como molde y la combinación de cebadores STCE1-TNERV y STCE1-TCET22I. El producto resultante de la PCR de aproximadamente 1 kpb se digirió con EcoRV y EcoT22I, y se unió con pCB1-M2 previamente digerido con StuI y PstI para la construcción del PCB-Ste.

STCE1-TNERV: GGGGATATCGCGCATCATGCGTTCTCCCCGTCCTC (SEC. Nº 30)

STCE1-TCET22I: GGGATGCATTTAAAGGCACTGCGAGTACCAGTC (SEC. Nº 31)

Como un vector de expresión para STCE1 modificada en el extremo N terminal, la región codificadora de la proteína madura, en la que se había eliminado la secuencia señal para la secreción en el gen de STCE1, se unió con pCB1-M2 obtenido en el Ejemplo B1 para la construcción del PCB-Sts.

En primer lugar, se amplificó la región codificadora de la proteína madura del gen de STCE1 por medio de PCR utilizando el plásmido pUC118-STCEex (FERM BP-10127) como molde y STCE1-TmNSph y STCE1-TCET22I como cebadores. El producto de la PCR resultante de aproximadamente 1 kpb fue digerido con SphI y EcoT22I, y se unió con pCB1-M2 previamente digerido con SphI y PstI para la construcción del PCB-SS.

STCE1-TmNSph: GGGGCATGCGATGGCAAGTCGACCCGCTAC (SEC. Nº 32)

(2) Generación de la cepa 2 de *Trichoderma viride*

Se irradió una suspensión de esporas (aproximadamente 10^9 UFC/ml) de *Trichoderma viride* MC300-1 (FERM BP-6047) con UV de dos lámparas UV colocadas a una altura de 30 cm por encima de la misma, mientras que se agitaba suavemente. La suspensión se extendió en un medio de selección y se cultivó a 28 °C durante 7 días. Las

cepas cultivadas se seleccionaron obteniéndose la cepa 2 de *Trichoderma viride* que mostraba requerimiento de uracilo.

Al respecto, el medio de selección era un medio mínimo [dihidrogenofosfato de potasio al 0,2%, sulfato de amonio al 0,4%, urea al 0,03%, sulfato de magnesio heptahidratado al 0,03%, cloruro de calcio al 0,03%, glucosa al 0,5%, agar al 2,5% y oligoelementos al 0,01% (preparados por la disolución de 5 mg de sulfato de hierro heptahidratado, sulfato de manganeso heptahidratado 1,56 mg, sulfato de cinc heptahidratado 1,4 mg y cloruro de cobalto 2 mg en 1 litro de agua)] suplementado con uridina 10 µg/ml y ácido 5-fluoroorótico 1 mg/ml.

(3) Transformantes que contienen STCE1 de tipo natural o STCE1 modificada en su extremo N terminal

Se transformó *Trichoderma viride* con pCB-Ste o pCB-Sts obtenidos en el Ejemplo B2(1), mediante la realización de un procedimiento de co-transformación utilizando la cepa 2 de *Trichoderma viride* obtenida en el Ejemplo B2(2) y el plásmido marcador pPYR4 que contenía el gen *pyr4* derivado de *Neurospora crassa*, para obtener los transformantes como cepas cultivadas en el medio mínimo.

En primer lugar, se inoculó la cepa 2 de *Trichoderma viride* en un matraz cónico de 200 ml que contenía 50 ml de un medio de formación de micelios [extracto de levadura al 1%, extracto de malta al 1%, polipeptona al 2%, glucosa al 2,5%, hidrogenofosfato dipotásico al 0,1%, sulfato de magnesio heptahidratado al 0,05% y uridina al 0,0001% (pH 7,0)] y se cultivó a 28 °C durante 2 días. Las células resultantes se recogieron por centrifugación, se lavaron con sacarosa 0,5 mol/l y se suspendieron en disolución de sacarosa 0,5 mol/l que contenía β-glucuronidasa 3 mg/ml (Sigma), quitinasa 1 mg/ml y zymolyase 20T 1 mg/ml (Seikagaku Corp.). La suspensión se agitó suavemente a 30 °C durante 60 a 90 minutos para generar protoplastos. Los protoplastos resultantes se recogieron por filtración y centrifugación, se lavaron con tampón SUTC y se resuspendieron en SUTC para preparar una suspensión de protoplastos.

A 100 µl de la suspensión de protoplastos, se le añadieron 10 µl de cada disolución de ADN (pCB-Ste o pCB-Sts). Tras dejar cada una en reposo sobre hielo durante 5 minutos, se añadieron 400 µl de la disolución de PEG, y se dejó todo en reposo sobre hielo durante 20 minutos. Tras lavar la suspensión de protoplastos resultante con el tampón SUTC, se cubrió el medio mínimo que contenía 0,5 mol/l de sacarosa con la suspensión de protoplastos junto con agar blando. La placa se incubó a 28 °C durante 5 días y a continuación las colonias que se habían desarrollado fueron transferidas al medio mínimo, para obtener los transformantes como colonias desarrolladas en el mismo.

(4) Comparación de las actividades de endoglucanasa de STCE1 de tipo natural y STCE1 modificada en su extremo N terminal en los sobrenadantes de los cultivos

Los transformantes resultantes se cultivaron según el procedimiento descrito en el documento WO98/11239, para obtener los sobrenadantes de cultivo. Los sobrenadantes de los cultivos se sometieron a SDS-PAGE, para seleccionar los sobrenadantes de cultivo en los que se detectaba con claridad la banda de STCE1 de aproximadamente 45 kDa. Se midieron las actividades de endoglucanasa en los sobrenadantes de los cultivos a pH 10 en ausencia y con presencia de LAS, y se compararon los grados de reducción de la actividad en presencia de LAS. Los resultados se muestran en la Tabla 6. En la Tabla 6, se representan las actividades de endoglucanasa por medio de porcentajes considerando como 100 la actividad de endoglucanasa a pH 10.

Tabla 6

	Vector de expresión	EGU (%)	
		Sin LAS	con LAS
STCE1 de tipo natural	pCB-Ste	100	14,5
STCE1 modificada en su extremo N terminal	pCB-Sts	100	25,9

Como resultado, al comparar las actividades de endoglucanasa en presencia de LAS, se encontró que la actividad de endoglucanasa de STCE1 modificada en su extremo N terminal fue aproximadamente 1,8 veces superior a la de STCE1 de tipo natural. Además, no se detectó la actividad de endoglucanasa de la cepa 2 de *Trichoderma viride*, y por lo tanto se encontró que las actividades de endoglucanasa derivaban de las enzimas recombinantes expresadas por los vectores de expresión incorporados.

(5) Análisis de secuencias de aminoácidos del extremo N terminal de STCE1 de tipo natural y STCE1 modificada en su extremo N terminal

A partir de cada sobrenadante de cultivo de los transformantes anteriores, se preparó una disolución de los mismos que contenía Tris-HCl 0,05 mol/l (pH 8,0) / sulfato de sodio 1 mol/l que se adsorbió en un lecho de igual volumen (BV) de TOYOPEARL Butil-650S (Tosoh Corporation). Los vehículos se lavaron con tres BV de Tris-HCl 0,05 mol/l (pH 8,0) / sulfato de sodio 1 mol/l para eliminar los componentes no adsorbidos. La elución se realizó con tres BV de Tris-HCl 0,05 mol/l (pH 8,0) / sulfato de sodio 0,5 mol/l, y la disolución eluida se concentró por medio de Pellicon XL (valor de corte 5000, Millipore) y Ultrafree-15 (valor de corte 5000, Millipore). A partir del concentrado obtenido, se

5 preparó una disolución del mismo que contenía fosfato de sodio 0,1 mol/l (pH 5,8) / sulfato de amonio 1 mol/l. La disolución resultante se sometió a una cromatografía hidrófoba utilizando Resource PHE (6 ml, Amersham Bioscience), para recuperar una fracción no adsorbida. La fracción no adsorbida se concentró y se desaló por medio de PD-10 (Amersham Bioscience), y, a partir de la fracción, se preparó una disolución de la misma que contenía disolución de Tris-HCl 0,05 mol/l (pH 7,5). La disolución se sometió a Resource Q (6 ml, Amersham Bioscience), para recuperar una fracción no adsorbida. La fracción no adsorbida se desaló y, a partir de la fracción, se preparó una disolución de la misma que contenía disolución de acetato de sodio 0,05 mol/l (pH 5,0). La disolución se sometió a Resource S (6 ml, Amersham Bioscience), para recuperar una fracción no adsorbida. La fracción no adsorbida se concentró y se sometió a una cromatografía en gel utilizando Superdex 75pg (16 X 60 mm; Amersham Bioscience) para recuperar una fracción con un peso molecular fraccionado de aproximadamente 45.000. Cada fracción (derivada de cada sobrenadante de cultivo de los transformantes anteriores) contenía solo la banda única de 45 kDa.

15 Las fracciones purificadas se sometieron a SDS-PAGE, y las proteínas sometidas a electroforesis se transfirieron eléctricamente a una membrana de PVDF. Las bandas de STCE1 de aproximadamente 45 kDa se recortaron de la mancha y se procesaron con el secuenciador Protein Sequencer Model 492 (Applied Biosystems) para analizar las secuencias de aminoácidos.

Como resultado, se confirmó que la secuencia de STCE1 de tipo natural obtenida del transformante pCB-Ste era coherente con la secuencia de ADN (SEC. ID N° 5) de STCE1 derivada de *Staphylotrichum coccosporum* IFO (nombre actual: NBRC) 31817.

20 Sin embargo, la secuencia de aminoácidos de STCE1 obtenida del transformante de STCE1 modificada en su extremo N terminal no pudo determinarse en el procedimiento anterior. Se utilizó una piroglutamato aminopeptidasa Pfu (Takara Bio) para eliminar el extremo N terminal de la STCE1 modificada en su extremo N terminal, y a continuación se determinó el aminoácido (SEC. N° 4).

25 Los resultados se resumieron en la Tabla 7 para comparar los aminoácidos del extremo N terminal y las secuencias de aminoácidos del extremo N terminal de STCE1 de tipo natural y STCE1 modificada en su extremo N terminal. A partir de los resultados se confirmó que cada STCE1 tenía la secuencia de aminoácidos del extremo N terminal esperada del correspondiente vector de expresión construido.

Tabla 7

	Aminoácido del extremo N terminal	Secuencia de aminoácidos del extremo N terminal
STCE1 de tipo natural	Alanina (sin modificar)	ADGKSTR (SEC. N° 33)
STCE1 modificada en su extremo N terminal	Ácido piroglutámico (modificado)	pQSACDGKSTR (SEC. N° 34)

30 (6) Comparación de las actividades de endoglucanasa de STCE1 de tipo natural purificada y STCE1 modificada en su extremo N terminal purificada

35 Se midieron las actividades de endoglucanasa de STCE1 de tipo natural purificada y STCE1 sustituida en su extremo N terminal purificada obtenidas en el Ejemplo B2(5) a pH 10 en ausencia y con presencia de LAS, y se compararon los grados de reducción de la actividad en presencia de LAS. Los resultados se muestran en la Tabla 8. En la Tabla 8, se representan las actividades de endoglucanasa por medio de porcentajes considerando como 100 la actividad de endoglucanasa a pH 10.

Tabla 8

	EGU (%)	
	Sin LAS	Con LAS
STCE1 de tipo natural purificada	100	16,2
STCE1 modificada en su extremo N terminal purificada	100	30,0

40 Como resultado, al comparar las actividades de endoglucanasa en presencia de LAS, se encontró que la actividad de endoglucanasa de STCE1 modificada en su extremo N terminal purificada fue aproximadamente 1,9 veces superior a la de la STCE1 de tipo natural purificada.

Ejemplo B3: Cambio de la secuencia añadida al extremo N terminal de STCE1, construcción del vector, transformación de *Trichoderma viride* con el mismo y evaluación de las actividades

(1) Construcción del vector de expresión para la adición de ácido piroglutámico solo al extremo N terminal de proteína madura

Como un vector de expresión para STCE1 en la que solamente se añadió ácido piroglutámico, se unió la región codificadora de la proteína madura en la que se había eliminado la secuencia señal para la secreción en el gen de STCE1, con el plásmido pCB1-M2 obtenido en el Ejemplo B1 para la construcción del pCB-Stm11.

En primer lugar, se amplificó la región codificadora de la proteína madura del gen de STCE1 por medio de PCR utilizando el plásmido pUC118-STCEex (FERM BP-10127) como molde y STCE-TmNSma y STCE1-TCET22I como cebadores. El producto de la PCR resultante, de aproximadamente 1 kpb, se digirió con SmaI y EcoT22I y se unió con pCB1-M2, previamente digerido con SmaI y PstI, para la construcción del pCB-Stm11. *Trichoderma viride* transformado con el vector pCB-Stm11 construido de la manera descrita anteriormente produjo la STCE1 que tiene la secuencia de aminoácidos N terminal del pQADGKSTR (SEC. ID. Nº 41) [es decir, la secuencia en la que se añadió un ácido piroglutámico (pQ) al extremo N terminal de STCE1 de tipo natural].

STCE-TmNSma: GGCCCGGGCTCAGGCCGATGGCAAGTCGACCCG (SEC. Nº 35)

(2) Comparación de la actividad de endoglucanasa de STCE1 a la que se añadió ácido piroglutámico

Se transformó la cepa 2 de *Trichoderma viride* con pCB-Stm11 según el procedimiento descrito en el Ejemplo B2(3). Los transformantes resultantes se cultivaron de acuerdo con el documento WO98/11239 para obtener los sobrenadantes de cultivo. Los sobrenadantes de los cultivos se sometieron a SDS-PAGE para seleccionar un sobrenadante de cultivo en el que se detectaba con claridad la banda de STCE1 de aproximadamente 45 kDa. Se midió la actividad de endoglucanasa en el sobrenadante del cultivo a pH 10 en ausencia y con presencia de LAS, y se comparó el grado de reducción de la actividad en presencia de LAS con el de la STCE1 de tipo natural en el Ejemplo B2(4). Los resultados se muestran en la Tabla 9. En la Tabla 9, se representan las actividades de endoglucanasa por medio de porcentajes considerando como 100 la actividad de endoglucanasa a pH 10.

Tabla 9

		Vector expresión		EGU (%)	
				Sin LAS	Con LAS
STCE1 de tipo natural	pCB-Ste			100	14,5
STCE1 con pQ añadido	pCB-Stm11			100	23,0

Como resultado, al comparar las actividades de endoglucanasa en presencia de LAS, se encontró que la actividad de endoglucanasa de la STCE1 a la que se añadió ácido piroglutámico fue aproximadamente 1,6 veces superior a la de la STCE1 de tipo natural.

Ejemplo B4: Cambio de la secuencia añadida al extremo N terminal de STCE1, construcción del vector, transformación de *Trichoderma viride* con el mismo y evaluación de las actividades

(1) Construcción del vector de expresión para la adición de 4 aminoácidos que contienen pQ al extremo N terminal de la proteína madura

Como vector de expresión para la STCE1 a la que se añadieron cuatro aminoácidos que contenían ácido piroglutámico, se unió la región codificadora de proteína madura, en la que se había suprimido la secuencia señal para la secreción en el gen de STCE1, con el plásmido pCB1-M2 obtenido en el Ejemplo B1 para la construcción del pCB-Stm12.

En primer lugar, se amplificó la región codificadora de la proteína madura del gen de STCE1 por medio de PCR utilizando el plásmido pUC118-STCEex (FERM BP-10127) como molde y STCE-TmNSph-2 y STCE1-TCET22I como cebadores. El producto de la PCR resultante, de aproximadamente 1 kpb, se digirió con SphI y EcoT22I y se unió con pCB1-M2 previamente digerido con SphI y PstI para la construcción del pCB-Stm12. *Trichoderma viride* transformado con el vector pCB-Stm12 construido como se describió anteriormente produjo la STCE1 que tiene la secuencia de aminoácidos del extremo N terminal del pQSACADGKSTR (SEC. ID. Nº 42) [es decir, la secuencia en la que se añadió un péptido que consiste en 4 aminoácidos (el aminoácido N terminal es ácido piroglutámico) al extremo N terminal de la STCE1N de tipo natural].

STCE-TmNSph-2: GGGCATGCGCCGATGGCAAGTCGACCCGC (SEC. Nº 36)

(2) Comparación de la actividad de endoglucanasa de la STCE1 con 4 aminoácidos añadidos que contiene pQ

La cepa 2 de *Trichoderma viride* se transformó con pCB-Stm12 según el procedimiento descrito en el Ejemplo B2(3). Los transformantes resultantes se cultivaron de acuerdo con el documento WO98/11239 para obtener los sobrenadantes de cultivo. Los sobrenadantes de los cultivos se sometieron a SDS-PAGE, para seleccionar un sobrenadante de cultivo en el que se detectaba con claridad la banda de STCE1 de aproximadamente 45 kDa. Se

midió la actividad de endoglucanasa en el sobrenadante del cultivo a pH 10 en ausencia y con presencia de LAS y se comparó el grado de reducción de la actividad en presencia de LAS con el de la STCE1 de tipo natural en el Ejemplo B2(4). Los resultados se muestran en la Tabla 10. En la Tabla 10, se representan las actividades de endoglucanasa por medio de porcentajes considerando como 100 la actividad de endoglucanasa a pH 10.

5

Tabla 10

	Vector de expresión	EGU (%)	
		Sin LAS	Con LAS
STCE1 de tipo natural	pCB-Ste	100	14,5
STCE1 con 4-AAs añadidos que contiene pQ	pCB-Stm12	100	22,4

Como resultado, al comparar las actividades de endoglucanasa en presencia de LAS, se encontró que la actividad de endoglucanasa de la STCE1 con 4 aminoácidos añadidos que contiene pQ (STCE1 con 4 AAs añadidos que contiene pQ) fue aproximadamente 1,5 veces superior a la de la STCE1 de tipo natural.

10 Aplicabilidad industrial

La proteína de la presente invención presenta una pequeña reducción en la actividad de endoglucanasa en presencia de un tensioactivo, en comparación con las celulasas de tipo natural, y por consiguiente es útil como componente de un detergente para la ropa.

15 Aunque la presente invención ha sido descrita con referencia a formas de realización concretas, es posible realizar diversos cambios y modificaciones obvias para los expertos en la técnica sin apartarse del alcance de las reivindicaciones adjuntas.

Breve descripción de los dibujos

20 La figura 1 muestra los resultados de la comparación entre la secuencia de aminoácidos de la endoglucanasa STCE1 [el péptido señal (ID. SEC. N° 43) y la proteína madura (ID. SEC. N° 44)] y las secuencias de aminoácidos de endoglucanasas conocidas pertenecientes a la familia de 45, NCE4 [el péptido señal (ID. SEC. N° 45) y la proteína madura (ID. SEC. N° 46)] y NCE5 [el péptido señal (ID. SEC. N° 47) y la proteína madura (ID. SEC. N° 48)], con respecto a las secuencias de aminoácidos de la mitad N terminal.

La figura 2 muestra los resultados de la comparación que se describe en la figura 1, con respecto a las secuencias de aminoácidos de la mitad C terminal.

25 Texto libre en el listado de secuencias

Las características de "secuencia artificial" se describen en el identificador numérico <223> en el Listado de secuencias. Cada una de las secuencias de nucleótidos de las ID. SEC. N° 6-10, 12-13, 15-18, 21-26, 28, 30-32 y 35-36 son: cebador MNC-02, cebador MNC-03, cebador MKA-05, cebador MKA-06, pMKD01, cebador pMN-Bam, pJD01, cebador NCE4-Ne, cebador NCE4-Ce, cebador NCE4-N, cebador NCE4-C, cebador CBN-Stu, cebador 30 CBC-Xho, CB1-SmSph, cebador CB1-Bam, cebador CB1-Pst, pCB1-M2, pCB1-M2, cebador STCE1-TNERV, cebador STCE1-TCET22I, cebador STCE1-TmNSph, cebador STCE1-TmNSma y cebador STCE1-TmNSph-2. Cada una de las secuencias de aminoácidos de las ID. SEC. N° 11, 14, 27 y 29 son: pMKD01, pJD01, pCB1-M2 y pCB1-M2.

LISTADO DE SECUENCIAS

35 <110> MEIJI SEIKA KAISHA, LTD.

<120> Celulosa tolerante a los detergentes y procedimiento para modificar la misma

<130> MEJ-719

<150> JP 2003-409692

<151> 08-12-2003

40 <160> 48

<170> PatentIn version 3.1

<210> 1

<211> 870

<212> ADN

<213> Humicola insolens MN200-1

$\langle 220 \rangle$

<221> CDS

5 <221> (1) .. (870)

<223>

$\langle 220 \rangle$

<221> característica miscelánea

$\langle 221 \rangle$ (1) .. (3)

10 <223> Ácido piroglutámico

$\langle 220 \rangle$

<221> fuente

<222> (16) .. (870)

<223> Humicola insolens MN200-1

15 <400> 1

cag aac tgt gga tcc gct gat ggc aag tcc acc cgc tac tgg gac tgc 48
Gln Asn Cys Gly Ser Ala Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys
1 5 10 15

tgc aag cct tog tgc ggc tgg gcc aag aag gct ccc gtg aac cag cct 96
Cys Lys Pro Ser Cys Gly Trp Ala Lys Lys Ala Pro Val Asn Gln Pro

20	25	30	
gtc ttc tcc tgc aac gcc aac ttc cag cgt ctc act gac ttc gac gcc			144
Val Phe Ser Cys Asn Ala Asn Phe Gln Arg Leu Thr Asp Phe Asp Ala			
35	40	45	
aag tcc ggc tgc gag ccg ggc ggt gtc gcc tac tcg tgc gcc gac cag			192
Lys Ser Gly Cys Glu Pro Gly Gly Val Ala Tyr Ser Cys Ala Asp Gln			
50	55	60	
acc cca tgg gct gtg aac gac gac ttc gcg ttc ggt ttt gct gcc acc			240
Thr Pro Trp Ala Val Asn Asp Asp Phe Ala Phe Gly Phe Ala Ala Thr			
65	70	75	80
tct att gcc ggc agc aat gag gcg ggc tgg tgc tgc gcc tgc tac gag			288
Ser Ile Ala Gly Ser Asn Glu Ala Gly Trp Cys Cys Ala Cys Tyr Glu			
85	90	95	
ctc acc ttc aca tcc ggt cct gtt gct ggc aag aag atg gtc gtc cag			336
Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Lys Met Val Val Gln			
100	105	110	
tcc acc agc act ggc ggt gat ctt ggc agc aac cac ttc gat ctc aac			384
Ser Thr Ser Thr Gly Gly Asp Leu Gly Ser Asn His Phe Asp Leu Asn			
115	120	125	
atc ccc ggc ggc ggc gtc ggc atc ttc gac gga tgc act ccc cag ttc			432
Ile Pro Gly Gly Gly Val Gly Ile Phe Asp Gly Cys Thr Pro Gln Phe			
130	135	140	
ggc ggt ctg ccc ggc cag cgc tac ggc ggc atc tcg tcc cgc aac gag			480
Gly Gly Leu Pro Gly Gln Arg Tyr Gly Gly Ile Ser Ser Arg Asn Glu			
145	150	155	160
tgc gat cgg ttc ccc gac gcc ctc aag ccc ggc tgc tac tgg cgc ttc			528
Cys Asp Arg Phe Pro Asp Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe			
165	170	175	
gac tgg ttc aag aac gcc gac aac ccg agc ttc agc ttc cgt cag gtc			576
Asp Trp Phe Lys Asn Ala Asp Asn Pro Ser Phe Ser Phe Arg Gln Val			
180	185	190	
caa tgc cca gcc gag ctc gtc gct cgc acc gga tgc cgc cgc aac gac			624
Gln Cys Pro Ala Glu Leu Val Ala Arg Thr Gly Cys Arg Arg Asn Asp			

ES 2 371 916 T3

195	200	205	
gac ggc aac ttc oct gcc gtc cag atc ccc tcc agc agc acc agc tct			672
Asp Gly Asn Phe Pro Ala Val Gln Ile Pro Ser Ser Ser Thr Ser Ser			
210	215	220	
ccg gtc ggc cag cct acc agt acc agc acc acc tcc acc tcc acc acc			720
Pro Val Gly Gln Pro Thr Ser Thr Ser Thr Thr Ser Thr Ser Thr Thr			
225	230	235	240
tgc agc ccg ccc gtc cag cct acg act ccc agc ggc tgc act gct gag			768
Ser Ser Pro Pro Val Gln Pro Thr Thr Pro Ser Gly Cys Thr Ala Glu			
245	250	255	
agg tgg gct cag tgc ggc ggc aat ggc tgg agc ggc tgc acc acc tgc			816
Arg Trp Ala Gln Cys Gly Gly Asn Gly Trp Ser Gly Cys Thr Thr Cys			
260	265	270	
gtc gct ggc agc acc tgc acg aag att aat gac tgg tac cat cag tgc			864
Val Ala Gly Ser Thr Cys Thr Lys Ile Asn Asp Trp Tyr His Gln Cys			
275	280	285	
ctg taa			870
Leu			

<210> 2

<211> 289

<212> PRT

5 <213> Humicola insolens MN200-1

<220>

<221> característica miscelánea

<221> (1) .. (3)

<223> Ácido piroglutámico

10 <400> 2

Gln	Asn	Cys	Gly	Ser	Ala	Asp	Gly	Lys	Ser	Thr	Arg	Tyr	Trp	Asp	Cys
1				5				10						15	

Cys Lys Pro Ser Cys Gly Trp Ala Lys Lys Ala Pro Val Asn Gln Pro

ES 2 371 916 T3

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

	195		200		205														
	Asp	Gly	Asn	Phe	Pro	Ala	Val	Gln	Ile	Pro	Ser	Ser	Ser	Thr	Ser	Ser			
	210						215					220							
	Pro	Val	Gly	Gln	Pro	Thr	Ser	Thr	Ser	Thr	Thr	Ser	Thr	Ser	Thr	Thr			
	225					230					235								240
	Ser	Ser	Pro	Pro	Val	Gln	Pro	Thr	Thr	Pro	Ser	Gly	Cys	Thr	Ala	Glu			
					245					250					255				
	Arg	Trp	Ala	Gln	Cys	Gly	Gly	Asn	Gly	Trp	Ser	Gly	Cys	Thr	Thr	Cys			
				260					265					270					
	Val	Ala	Gly	Ser	Thr	Cys	Thr	Lys	Ile	Asn	Asp	Trp	Tyr	His	Gln	Cys			
			275					280					285						

Leu

<210> 3

<211> 897

<212> ADN

5 <213> Staphylotrichum coccosporum IF0 31817

<220>

<221> CDS

<221> (1) .. (897)

<223>

10 <220>

<221> característica miscelánea

<221> (1) .. (3)

<223> Ácido piroglutámico

<220>

15 <221> fuente

<222> (13) .. (897)

<223> Staphylotrichum coccosporum IF0 31817

ES 2 371 916 T3

<400> 3

cag tgc gca tgc gat ggc aag tcc acc cgc tac tgg gac tgc tgc aag	48
Gln Ser Ala Cys Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys Cys Lys	
1 5 10 15	
cct tgc tgc tgc tgg ccc ggc aag gcc tgc gtg aac cag ccc gtc ttc	96
Pro Ser Cys Ser Trp Pro Gly Lys Ala Ser Val Asn Gln Pro Val Phe	
20 25 30	
gcc tgc agc gcc aac ttc cag cgc atc agc gac ccc aac gtc aag tgc	144
Ala Cys Ser Ala Asn Phe Gln Arg Ile Ser Asp Pro Asn Val Lys Ser	
35 40 45	
ggc tgc gac ggc ggc tcc gcc tac gcc tgc gcc gac cag acc ccg tgg	192
Gly Cys Asp Gly Gly Ser Ala Tyr Ala Cys Ala Asp Gln Thr Pro Trp	
50 55 60	
gcc gtc aac gac aac ttc tgc tac ggc ttc gcc gcc acg tcc atc tgc	240
Ala Val Asn Asp Asn Phe Ser Tyr Gly Phe Ala Ala Thr Ser Ile Ser	
65 70 75 80	
ggc ggc aac gag gcc tgc tgg tgc tgt ggc tgc tac gag ctg acc ttc	288
Gly Gly Asn Glu Ala Ser Trp Cys Cys Gly Cys Tyr Glu Leu Thr Phe	
85 90 95	
acc tgc ggc ccc gtc gct ggc aag acc atg gtt gtc cag tcc acc tgc	336
Thr Ser Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser Thr Ser	
100 105 110	
acc ggc ggc gac ctc ggc acc aac cac ttc gac ctg gcc atg ccc ggt	384
Thr Gly Gly Asp Leu Gly Thr Asn His Phe Asp Leu Ala Met Pro Gly	
115 120 125	
ggt ggt gtc ggc atc ttc gac ggc tgc tgc ccc cag ttc ggc ggc ctc	432
Gly Gly Val Gly Ile Phe Asp Gly Cys Ser Pro Gln Phe Gly Gly Leu	
130 135 140	
gcc ggc gac cgc tac ggc ggc gtc tgc tgc cgc agc cag tgc gac tgc	480

Ala Gly Asp Arg Tyr Gly Gly Val Ser Ser Arg Ser Gln Cys Asp Ser	
145 150 155 160	
ttc ccc gcc gcc ctc aag ccc ggc tgc tac tgg cgc ttc gac tgg ttc	528
Phe Pro Ala Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe Asp Trp Phe	
165 170 175	
aag aac gcc gac aac ccg acc ttc acc ttc cgc cag gtc cag tgc ccg	576
Lys Asn Ala Asp Asn Pro Thr Phe Thr Phe Arg Gln Val Gln Cys Pro	
180 185 190	
tcg gag ctc gtc gcc cgc acc ggc tgc cgc cgc aac gac gac ggc aac	624
Ser Glu Leu Val Ala Arg Thr Gly Cys Arg Arg Asn Asp Asp Gly Asn	
195 200 205	
ttc ccc gtc ttc acc cct ccc tcg ggc ggt cag tcc tcc tcg tct tcc	672
Phe Pro Val Phe Thr Pro Pro Ser Gly Gly Gln Ser Ser Ser Ser Ser	
210 215 220	
tcc tcc agc agc gcc aag ccc acc tcc acc tcc acc tcg acc acc tcc	720
Ser Ser Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr Ser Thr Thr Ser	
225 230 235 240	
acc aag gct acc tcc acc acc tcg acc gcc tcc agc cag acc tcg tcg	768
Thr Lys Ala Thr Ser Thr Thr Ser Thr Ala Ser Ser Gln Thr Ser Ser	
245 250 255	
tcc acc ggc ggc ggc tgc gcc gcc cag cgc tgg gcg cag tgc ggc ggc	816
Ser Thr Gly Gly Gly Cys Ala Ala Gln Arg Trp Ala Gln Cys Gly Gly	
260 265 270	
atc ggg ttc tcg ggc tgc acc acg tgc gtc agc ggc acc acc tgc aac	864
Ile Gly Phe Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr Cys Asn	
275 280 285	
aag cag aac gac tgg tac tcg cag tgc ctt taa	897
Lys Gln Asn Asp Trp Tyr Ser Gln Cys Leu	
290 295	

<210> 4

<211> 298

<212> PRT

5 <213> Staphylotrichum coccosporum IF0 31817

<220>

<221> característica miscelánea

<221> (1) .. (3)

<223> Ácido piroglutámico

5 <400> 4

Gln Ser Ala Cys Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys Cys Lys
1 5 10 15

Pro Ser Cys Ser Trp Pro Gly Lys Ala Ser Val Asn Gln Pro Val Phe
20 25 30

Ala Cys Ser Ala Asn Phe Gln Arg Ile Ser Asp Pro Asn Val Lys Ser
35 40 45

Gly Cys Asp Gly Gly Ser Ala Tyr Ala Cys Ala Asp Gln Thr Pro Trp
50 55 60

Ala Val Asn Asp Asn Phe Ser Tyr Gly Phe Ala Ala Thr Ser Ile Ser
65 70 75 80

Gly Gly Asn Glu Ala Ser Trp Cys Cys Gly Cys Tyr Glu Leu Thr Phe
85 90 95

Thr Ser Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser Thr Ser
100 105 110

Thr Gly Gly Asp Leu Gly Thr Asn His Phe Asp Leu Ala Met Pro Gly
115 120 125

Gly Gly Val Gly Ile Phe Asp Gly Cys Ser Pro Gln Phe Gly Gly Leu
130 135 140

ES 2 371 916 T3

Ala Gly Asp Arg Tyr Gly Gly Val Ser Ser Arg Ser Gln Cys Asp Ser
145 150 155 160

Phe Pro Ala Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe Asp Trp Phe
165 170 175

Lys Asn Ala Asp Asn Pro Thr Phe Thr Phe Arg Gln Val Gln Cys Pro
180 185 190

Ser Glu Leu Val Ala Arg Thr Gly Cys Arg Arg Asn Asp Asp Gly Asn
195 200 205

Phe Pro Val Phe Thr Pro Pro Ser Gly Gly Gln Ser Ser Ser Ser Ser
210 215 220

Ser Ser Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr Ser Thr Thr Ser
225 230 235 240

Thr Lys Ala Thr Ser Thr Thr Ser Thr Ala Ser Ser Gln Thr Ser Ser
245 250 255

Ser Thr Gly Gly Gly Cys Ala Ala Gln Arg Trp Ala Gln Cys Gly Gly
260 265 270

Ile Gly Phe Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr Cys Asn
275 280 285

Lys Gln Asn Asp Trp Tyr Ser Gln Cys Leu
290 295

<210> 5

<211> 1037

<212> ADN

5 <213> Staphylotrichum coccosporum IF0 31817

<220>

<221> péptido señal

<221> (1) .. (63)

<223>

5 <220>

<221> exón

<221> (64) .. (333)

<223>

<220>

10 <221> Intrón

<221> (334) .. (419)

<223>

<220>

<221> exón

15 <221> (420) .. (1037)

<223>

<400> 5

atgcgttcct ccccgctcct ccgcacggcc ctggccgctg cccctccccct ggccgccctc	60
gct gcc gat ggc aag tcg acc cgc tac tgg gac tgt tgc aag ccg tcg	108
Ala Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys Cys Lys Pro Ser	
1 5 10 15	
tgc tcg tgg ccc ggc aag gcc tcg gtg aac cag ccc gtc ttc gcc tgc	156
Cys Ser Trp Pro Gly Lys Ala Ser Val Asn Gln Pro Val Phe Ala Cys	
20 25 30	
agc gcc aac ttc cag cgc atc agc gac ccc aac gtc aag tcg ggc tgc	204
Ser Ala Asn Phe Gln Arg Ile Ser Asp Pro Asn Val Lys Ser Gly Cys	
35 40 45	
gac ggc ggc tcc gcc tac gcc tgc gcc gac cag acc ccg tgg gcc gtc	252
Asp Gly Gly Ser Ala Tyr Ala Cys Ala Asp Gln Thr Pro Trp Ala Val	
50 55 60	
aac gac aac ttc tcg tac ggc ttc gcc gcc acg tcc atc tcg ggc ggc	300
Asn Asp Asn Phe Ser Tyr Gly Phe Ala Ala Thr Ser Ile Ser Gly Gly	
65 70 75	

aac gag gcc tgg tgg tgc tgt ggc tgc tac gag tgagtgttc ccccccccc	353
Asn Glu Ala Ser Trp Cys Cys Gly Cys Tyr Glu	
80 85 90	
ccccccccac ccccggttcg gtcccttgcc gtgccttctt cataactaac gctaccccc	413
tccagg ctg acc ttc acc tgc ggc ccc gtc gct ggc aag acc atg gtt	461
Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Thr Met Val	
95 100	
gtc cag tcc acc tog acc ggc ggc gac ctc ggc acc aac cac ttc gac	509
Val Gln Ser Thr Ser Thr Gly Gly Asp Leu Gly Thr Asn His Phe Asp	
105 110 115 120	
ctg gcc atg ccc ggt ggt ggt gtc ggc atc ttc gac ggc tgc tgc ccc	557
Leu Ala Met Pro Gly Gly Gly Val Gly Ile Phe Asp Gly Cys Ser Pro	
125 130 135	
cag ttc ggc ggc ctc gcc ggc gac cgc tac ggc ggc gtc tgc tgc cgc	605
Gln Phe Gly Gly Leu Ala Gly Asp Arg Tyr Gly Gly Val Ser Ser Arg	
140 145 150	
agc cag tgc gac tgc ttc ccc gcc gcc ctc aag ccc ggc tgc tac tgg	653
Ser Gln Cys Asp Ser Phe Pro Ala Ala Leu Lys Pro Gly Cys Tyr Trp	
155 160 165	
cgc ttc gac tgg ttc aag aac gcc gac aac ccg acc ttc acc ttc cgc	701
Arg Phe Asp Trp Phe Lys Asn Ala Asp Asn Pro Thr Phe Thr Phe Arg	
170 175 180	
cag gtc cag tgc ccg tgc gag ctc gtc gcc cgc acc ggc tgc cgc cgc	749
Gln Val Gln Cys Pro Ser Glu Leu Val Ala Arg Thr Gly Cys Arg Arg	
185 190 195 200	
aac gac gac ggc aac ttc ccc gtc ttc acc cct ccc tgc ggc ggt cag	797
Asn Asp Asp Gly Asn Phe Pro Val Phe Thr Pro Pro Ser Gly Gly Gln	
205 210 215	
tcc tcc tgc tot tcc tcc tcc agc agc gcc aag ccc acc tcc acc tcc	845
Ser Ser Ser Ser Ser Ser Ser Ser Ser Ala Lys Pro Thr Ser Thr Ser	
220 225 230	
acc tgc acc acc tcc acc aag gct acc tcc acc acc tgc acc gcc tcc	893

Thr	Ser	Thr	Thr	Ser	Thr	Lys	Ala	Thr	Ser	Thr	Thr	Ser	Thr	Ala	Ser	
		235					240					245				
agc	cag	acc	tgc	tgc	tcc	acc	ggc	ggc	ggc	tgc	gcc	gcc	cag	cgc	tgg	941
Ser	Gln	Thr	Ser	Ser	Ser	Thr	Gly	Gly	Gly	Cys	Ala	Ala	Gln	Arg	Trp	
	250					255				260						
gcg	cag	tgc	ggc	ggc	atc	ggg	ttc	tgc	ggc	tgc	acc	acg	tgc	gtc	agc	989
Ala	Gln	Cys	Gly	Gly	Ile	Gly	Phe	Ser	Gly	Cys	Thr	Thr	Cys	Val	Ser	
265					270				275					280		
ggc	acc	acc	tgc	aac	aag	cag	aac	gac	tgg	tac	tgc	cag	tgc	ctt	tga	1037
Gly	Thr	Thr	Cys	Asn	Lys	Gln	Asn	Asp	Trp	Tyr	Ser	Gln	Cys	Leu		
			285					290					295			

<210> 6

<211> 36

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Cebador MNC-02

<400> 6

gagcgccaga actgtggatc cacttggtga gcaatg 36

10 <210> 7

<211> 35

<212> ADN

<213> Secuencia artificial

<220>

15 <223> Descripción de la secuencia artificial: Cebador MNC-03

<400> 7

tccgccgttc tgagcggatc caggcgttg gcgcg 35

<210> 8

<211> 36

20 <212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Cebador MKA-05

<400> 8

25 gccgcccagc aggcgggatc cctcaccacc gagagg 36

<210> 9
 <211> 36
 <212> ADN
 <213> Secuencia artificial
 5 <220>
 <223> Descripción de la secuencia artificial: Cebador MKA-06
 <400> 9
 tgatcgtcga gtcaggatc cagaatttac aggcac 36
 <210> 10
 10 <211> 44
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: pMKD01
 15 <220>
 <221> característica miscelánea
 <221> (24) .. (25)
 <223> bases no consecutivas
 <400> 10
 20 gagcgccaga actgtggatc cctctgcctg taagcggatc cagg 44
 <210> 11
 <211> 10
 <212> PRT
 <213> Secuencia artificial
 25 <220>
 <223> Descripción de la secuencia artificial: pMKD01
 <220>
 <221> NON_CONS
 <221> (8) .. (9)
 30 <223>
 <220>
 <221> MOD_RES
 <221> (1) .. (1)
 <223> ÁCIDO PIRROLIDONA CARBOXÍLICO
 35

<400> 11

Glu	Arg	Gln	Asn	Cys	Gly	Ser	Leu	Cys	Leu
1				5					10

<210> 12

<211> 53

5 <212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Cebador pMN-Bam

<400> 12

10 ggtcaaacaa gtctgtgcgg atcctgggac aagatggcca agttcttct tac 53

<210> 13

<211> 44

<212> ADN

<213> Secuencia artificial

15 <220>

<223> Descripción de la secuencia artificial: pJD01

<220>

<221> característica miscelánea

<221> (24) .. (25)

20 <223> bases no consecutivas

<400> 13

tgcggatcct gggacaagat ggccccgttc tgagcggatc cagg 44

<210> 14

<211> 4

25 <212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: pJD01

<220>

30 <221> NON_CONS

<221> (2) .. (3)

<223>

35

<400> 14

Met Ala Pro Phe
1

<210> 15

<211> 37

5 <212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Cebador NCE4-Ne

<400> 15

10 ggggggatcc tgggacaaga tgcgttcctc ccctctc 37

<210> 16

<211> 32

<212> ADN

<213> Secuencia artificial

15 <220>

<223> Descripción de la secuencia artificial: Cebador NCE4-Ce

<400> 16

gggggggatcc ctgcgtttac aggcactgat gg 32

<210> 17

20 <211> 30

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Cebador NCE4-Ns

25 <400> 17

ccggtgttgg ccggatccgc tgatggcaag 30

<210> 18

<211> 30

<212> ADN

30 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Cebador NCE4-Cs

<400> 18

taaggccctc aaggatccct gcgtctacag 30

35 <210> 19

<211> 7

<212> PRT

<213> Humicola insolens MN200-1

<400> 19

	Ala	Asp	Gly	Lys	Ser	Thr	Arg
	1				5		

5

<210> 20

<211> 12

<212> PRT

<213> Humicola insolens MN200-1

10

<220>

<221> MOD_RES

<221> (1) .. (1)

<223> ÁCIDO PIRROLIDONA CARBOXÍLICO

<220>

15

<221> MUTÁGENO

<221> (1) .. (5)

<223>

<400> 20

	Gln	Asn	Cys	Gly	Ser	Ala	Asp	Gly	Lys	Ser	Thr	Arg
	1				5					10		

20

<210> 21

<211> 36

<212> ADN

<213> Secuencia artificial

<220>

25

<223> Descripción de la secuencia artificial: Cebador CBN-Stu

<400> 21

gatacatgat ggcgaggcct tagtcgacta gaatgc 36

<210> 22

<211> 36

30

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Cebador CBc-Xho

<400> 22

gacccaag ctttgctcg agtacctac aagcac 36

<210> 23

<211> 35

5 <212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Cebador CB1-SmSph

<400> 23

10 ggagggtgca tgccgactga gcccgggcag tagcc 35

<210> 24

<211> 23

<212> ADN

<213> Secuencia artificial

15 <220>

<223> Descripción de la secuencia artificial: Cebador CB1-Bam

<400> 24

gccgggagag gatccagtgg agg 23

<210> 25

20 <211> 30

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Cebador CB1-Pst

25 <400> 25

gctcgagtac ctactgcag gcactgagag 30

<210> 26

<211> 66

<212> ADN

30 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: pCB1-M2

35

<400> 26

ctagtcgact aaggcctgcg catcatgtat caaaagttgg cccatcatctc ggccttcttg 60

gctact 66

<210> 27

<211> 14

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: pCB1-M2

<400> 27

Met Tyr Gln Lys Leu Ala Leu Ile Ser Ala Phe Leu Ala Thr
1 5 10

10

<210> 28

<211> 61

<212> ADN

<213> Secuencia artificial

15 <220>

<223> Descripción de la secuencia artificial: pCB1-M2

<210>

<211> característica miscelánea

<212> (24) .. (25)

20 <213> bases no consecutivas

<400> 28

gcccgggctc agtcggcatg caccagtgct ctgcagtaag gtactcgagc aaaagcttga 60

g 61

<210> 29

<211> 12

25 <212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: pCB1-M2

<220>

30 <221> NON_CONS

<221> (8) .. (9)

<223>

<400> 29

Ala	Arg	Ala	Gln	Ser	Ala	Cys	Thr	Gln	Cys	Leu	Gln
1				5					10		

5 <210> 30

<211> 37

<212> ADN

<213> Secuencia artificial

<220>

10 <223> Descripción de la secuencia artificial: Cebador STCE1-TNERV

<400> 30

ggggatatcg cgcatcatgc gttcctcccc cgtcctc 37

<210> 31

<211> 33

15 <212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Cebador STCE1-TCET221

<400> 31

20 gggatgcatt taaaggcact gcgagtacca gtc 33

<210> 32

<211> 30

<212> ADN

<213> Secuencia artificial

25 <220>

<223> Descripción de la secuencia artificial: Cebador STCE1-TmNSph

<400> 32

ggggcatgcg atggcaagtc gacccgtac 30

<210> 33

30 <211> 7

<212> PRT

<213> Staphylotrichum coccosporum IFO 31817

35

<400> 33

Ala Asp Gly Lys Ser Thr Arg
1 5

<210> 34

<211> 10

5 <212> PRT

<213> Staphylotrichum coccosporum IFO 31817

<220>

<221> MOD_RES

<221> (1) .. (1)

10 <223> ÁCIDO PIRROLIDONA CARBOXÍLICO

<220>

<221> MUTÁGENO

<221> (1) .. (4)

<223>

15 <400> 34

Gln Ser Ala Cys Asp Gly Lys Ser Thr Arg
1 5 10

<210> 35

<211> 33

<212> ADN

20 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Cebador STCE1-TmNSma

<400> 35

ggcccgggct caggccgatg gcaagtcgac ccg 33

25 <210> 36

<211> 29

<212> ADN

<213> Secuencia artificial

<220>

30 <223> Descripción de la secuencia artificial: Cebador STCE1-TmNSph-2

<400> 36

gggcatgcgc cgatggcaag tcgacccgc 29

<210> 37

<211> 891

<212> ADN

<213> Staphylotrichum coccosporum IF0 31817

<220>

<221> CDS

5 <221> (1) .. (891)

<223>

<220>

<221> característica miscelánea

<221> (1) .. (3)

10 <223> Ácido piroglutámico

<220>

<221> fuente

<222> (4) .. (891)

<223> Staphylotrichum coccosporum IF0 31817

15 <400> 37

cag gcc gat ggc aag tcc acc cgc tac tgg gac tgc tgc aag cct tcg 48

Gln	Ala	Asp	Gly	Lys	Ser	Thr	Arg	Tyr	Trp	Asp	Cys	Cys	Lys	Pro	Ser	
1				5					10					15		
tgc	tgc	tgg	ccc	ggc	aag	gcc	tgc	gtg	aac	cag	ccc	gtc	ttc	gcc	tgc	96
Cys	Ser	Trp	Pro	Gly	Lys	Ala	Ser	Val	Asn	Gln	Pro	Val	Phe	Ala	Cys	
			20					25					30			
agc	gcc	aac	ttc	cag	cgc	atc	agc	gac	ccc	aac	gtc	aag	tgc	ggc	tgc	144
Ser	Ala	Asn	Phe	Gln	Arg	Ile	Ser	Asp	Pro	Asn	Val	Lys	Ser	Gly	Cys	
		35					40					45				
gac	ggc	ggc	tcc	gcc	tac	gcc	tgc	gcc	gac	cag	acc	ccg	tgg	gcc	gtc	192
Asp	Gly	Gly	Ser	Ala	Tyr	Ala	Cys	Ala	Asp	Gln	Thr	Pro	Trp	Ala	Val	
	50					55					60					
aac	gac	aac	ttc	tgc	tac	ggc	ttc	gcc	gcc	acg	tcc	atc	tgc	ggc	ggc	240
Asn	Asp	Asn	Phe	Ser	Tyr	Gly	Phe	Ala	Ala	Thr	Ser	Ile	Ser	Gly	Gly	
65					70					75					80	
aac	gag	gcc	tgc	tgg	tgc	tgt	ggc	tgc	tac	gag	ctg	acc	ttc	acc	tgc	288
Asn	Glu	Ala	Ser	Trp	Cys	Cys	Gly	Cys	Tyr	Glu	Leu	Thr	Phe	Thr	Ser	
				85					90					95		
ggc	ccc	gtc	gct	ggc	aag	acc	atg	gtt	gtc	cag	tcc	acc	tgc	acc	ggc	336
Gly	Pro	Val	Ala	Gly	Lys	Thr	Met	Val	Val	Gln	Ser	Thr	Ser	Thr	Gly	
			100					105						110		
ggc	gac	ctc	ggc	acc	aac	cac	ttc	gac	ctg	gcc	atg	ccc	ggt	ggt	ggt	384
Gly	Asp	Leu	Gly	Thr	Asn	His	Phe	Asp	Leu	Ala	Met	Pro	Gly	Gly	Gly	
		115					120					125				
gtc	ggc	atc	ttc	gac	ggc	tgc	tgc	ccc	cag	ttc	ggc	ggc	ctc	gcc	ggc	432
Val	Gly	Ile	Phe	Asp	Gly	Cys	Ser	Pro	Gln	Phe	Gly	Gly	Leu	Ala	Gly	
	130					135					140					
gac	cgc	tac	ggc	ggc	gtc	tgc	tgc	cgc	agc	cag	tgc	gac	tgc	ttc	ccc	480
Asp	Arg	Tyr	Gly	Gly	Val	Ser	Ser	Arg	Ser	Gln	Cys	Asp	Ser	Phe	Pro	
145					150					155					160	
gcc	gcc	ctc	aag	ccc	ggc	tgc	tac	tgg	cgc	ttc	gac	tgg	ttc	aag	aac	528
Ala	Ala	Leu	Lys	Pro	Gly	Cys	Tyr	Trp	Arg	Phe	Asp	Trp	Phe	Lys	Asn	
				165					170					175		
gcc	gac	aac	ccg	acc	ttc	acc	ttc	cgc	cag	gtc	cag	tgc	ccg	tgc	gag	576

Ala	Asp	Asn	Pro	Thr	Phe	Thr	Phe	Arg	Gln	Val	Gln	Cys	Pro	Ser	Glu	
			180					185					190			
ctc	gtc	gcc	cgc	acc	ggc	tgc	cgc	cgc	aac	gac	gac	ggc	aac	ttc	ccc	624
Leu	Val	Ala	Arg	Thr	Gly	Cys	Arg	Arg	Asn	Asp	Asp	Gly	Asn	Phe	Pro	
		195					200					205				
gtc	ttc	acc	cct	ccc	tgc	ggc	ggt	cag	tcc	tcc	tgc	tct	tcc	tcc	tcc	672
Val	Phe	Thr	Pro	Pro	Ser	Gly	Gly	Gln	Ser	Ser	Ser	Ser	Ser	Ser	Ser	
	210					215					220					
agc	agc	gcc	aag	ccc	acc	tcc	acc	tcc	acc	tgc	acc	acc	tcc	acc	aag	720
Ser	Ser	Ala	Lys	Pro	Thr	Ser	Thr	Ser	Thr	Ser	Thr	Thr	Ser	Thr	Lys	
225					230					235					240	
gct	acc	tcc	acc	acc	tgc	acc	gcc	tcc	agc	cag	acc	tgc	tgc	tcc	acc	768
Ala	Thr	Ser	Thr	Thr	Ser	Thr	Ala	Ser	Ser	Gln	Thr	Ser	Ser	Ser	Thr	
				245					250					255		
ggc	ggc	ggc	tgc	gcc	gcc	cag	cgc	tgg	gcg	cag	tgc	ggc	ggc	atc	ggg	816
Gly	Gly	Gly	Cys	Ala	Ala	Gln	Arg	Trp	Ala	Gln	Cys	Gly	Gly	Ile	Gly	
			260					265					270			
ttc	tgc	ggc	tgc	acc	acg	tgc	gtc	agc	ggc	acc	acc	tgc	aac	aag	cag	864
Phe	Ser	Gly	Cys	Thr	Thr	Cys	Val	Ser	Gly	Thr	Thr	Cys	Asn	Lys	Gln	
		275					280					285				
aac	gac	tgg	tac	tgc	cag	tgc	ctt	taa								891
Asn	Asp	Trp	Tyr	Ser	Gln	Cys	Leu									
	290					295										

<210> 38

<211> 296

<212> PRT

5 <213> Staphylotrichum coccosporum IF0 31817

<220>

<221> característica miscelánea

<221> (1) .. (3)

<223> Ácido piroglutámico

ES 2 371 916 T3

<400> 38

Gln Ala Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys Cys Lys Pro Ser
1 5 10 15

Cys Ser Trp Pro Gly Lys Ala Ser Val Asn Gln Pro Val Phe Ala Cys
20 25 30

Ser Ala Asn Phe Gln Arg Ile Ser Asp Pro Asn Val Lys Ser Gly Cys
35 40 45

Asp Gly Gly Ser Ala Tyr Ala Cys Ala Asp Gln Thr Pro Trp Ala Val
50 55 60

Asn Asp Asn Phe Ser Tyr Gly Phe Ala Ala Thr Ser Ile Ser Gly Gly
65 70 75 80

Asn Glu Ala Ser Trp Cys Cys Gly Cys Tyr Glu Leu Thr Phe Thr Ser
85 90 95

Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser Thr Ser Thr Gly
100 105 110

Gly Asp Leu Gly Thr Asn His Phe Asp Leu Ala Met Pro Gly Gly Gly
115 120 125

Val Gly Ile Phe Asp Gly Cys Ser Pro Gln Phe Gly Gly Leu Ala Gly
130 135 140

Asp Arg Tyr Gly Gly Val Ser Ser Arg Ser Gln Cys Asp Ser Phe Pro
145 150 155 160

Ala Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe Asp Trp Phe Lys Asn
165 170 175

Ala Asp Asn Pro Thr Phe Thr Phe Arg Gln Val Gln Cys Pro Ser Glu
180 185 190

Leu Val Ala Arg Thr Gly Cys Arg Arg Asn Asp Asp Gly Asn Phe Pro
195 200 205

Val Phe Thr Pro Pro Ser Gly Gly Gln Ser Ser Ser Ser Ser Ser Ser
210 215 220

Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr Ser Thr Thr Ser Thr Lys
225 230 235 240

Ala Thr Ser Thr Thr Ser Thr Ala Ser Ser Gln Thr Ser Ser Ser Thr
245 250 255

Gly Gly Gly Cys Ala Ala Gln Arg Trp Ala Gln Cys Gly Gly Ile Gly
260 265 270

Phe Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr Cys Asn Lys Gln
275 280 285

Asn Asp Trp Tyr Ser Gln Cys Leu
290 295

<210> 39

<211> 900

<212> ADN

5 <213> Staphylotrichum coccosporum IF0 31817

<220>

<221> CDS

<221> (1) .. (900)

<223>

10 <220>

<221> fuente

<222> (13) .. (90)

<223> Ácido piroglutámico

<220>

<221> característica miscelánea

<221> (1) .. (3)

5 <223> Ácido piroglutámico

<400> 39

cag tgc gca tgc gcc gat ggc aag tcc acc cgc tac tgg gac tgc tgc	48
Gln Ser Ala Cys Ala Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys Cys	
1 5 10 15	
aag cct tgc tgc tgc tgg ccc ggc aag gcc tgc gtg aac cag ccc gtc	96
Lys Pro Ser Cys Ser Trp Pro Gly Lys Ala Ser Val Asn Gln Pro Val	
20 25 30	
ttc gcc tgc agc gcc aac ttc cag cgc atc agc gac ccc aac gtc aag	144
Phe Ala Cys Ser Ala Asn Phe Gln Arg Ile Ser Asp Pro Asn Val Lys	
35 40 45	
tgc ggc tgc gac ggc ggc tcc gcc tac gcc tgc gcc gac cag acc ccg	192
Ser Gly Cys Asp Gly Gly Ser Ala Tyr Ala Cys Ala Asp Gln Thr Pro	
50 55 60	
tgg gcc gtc aac gac aac ttc tgc tac ggc ttc gcc gcc acg tcc atc	240
Trp Ala Val Asn Asp Asn Phe Ser Tyr Gly Phe Ala Ala Thr Ser Ile	
65 70 75 80	
tgc ggc ggc aac gag gcc tgc tgg tgc tgt ggc tgc tac gag ctg acc	288
Ser Gly Gly Asn Glu Ala Ser Trp Cys Cys Gly Cys Tyr Glu Leu Thr	
85 90 95	
ttc acc tgc ggc ccc gtc gct ggc aag acc atg gtt gtc cag tcc acc	336
Phe Thr Ser Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser Thr	
100 105 110	
tgc acc ggc ggc gac ctc ggc acc aac cac ttc gac ctg gcc atg ccc	384
Ser Thr Gly Gly Asp Leu Gly Thr Asn His Phe Asp Leu Ala Met Pro	
115 120 125	

ggt ggt ggt gtc ggc atc ttc gac ggc tgc tcg ccc cag ttc ggc ggc Gly Gly Gly Val Gly Ile Phe Asp Gly Cys Ser Pro Gln Phe Gly Gly	432
130 135 140	
ctc gcc ggc gac cgc tac ggc ggc gtc tcg tcg cgc agc cag tgc gac Leu Ala Gly Asp Arg Tyr Gly Gly Val Ser Ser Arg Ser Gln Cys Asp	480
145 150 155 160	
tcg ttc ccc gcc gcc ctc aag ccc ggc tgc tac tgg cgc ttc gac tgg Ser Phe Pro Ala Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe Asp Trp	528
165 170 175	
ttc aag aac gcc gac aac ccg acc ttc acc ttc cgc cag gtc cag tgc Phe Lys Asn Ala Asp Asn Pro Thr Phe Thr Phe Arg Gln Val Gln Cys	576
180 185 190	
ccg tcg gag ctc gtc gcc cgc acc ggc tgc cgc cgc aac gac gac ggc Pro Ser Glu Leu Val Ala Arg Thr Gly Cys Arg Arg Asn Asp Asp Gly	624
195 200 205	
aac ttc ccc gtc ttc acc cct ccc tcg ggc ggt cag tcc tcc tcg tct Asn Phe Pro Val Phe Thr Pro Pro Ser Gly Gly Gln Ser Ser Ser Ser	672
210 215 220	
tcc tcc tcc agc agc gcc aag ccc acc tcc acc tcc acc tcg acc acc Ser Ser Ser Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr Ser Thr Thr	720
225 230 235 240	
tcc acc aag gct acc tcc acc acc tcg acc gcc tcc agc cag acc tcg Ser Thr Lys Ala Thr Ser Thr Thr Ser Thr Ala Ser Ser Gln Thr Ser	768
245 250 255	
tcg tcc acc ggc ggc ggc tgc gcc gcc cag cgc tgg gcg cag tgc ggc Ser Ser Thr Gly Gly Gly Cys Ala Ala Gln Arg Trp Ala Gln Cys Gly	816
260 265 270	
ggc atc ggg ttc tcg ggc tgc acc acg tgc gtc agc ggc acc acc tgc Gly Ile Gly Phe Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr Cys	864
275 280 285	
aac aag cag aac gac tgg tac tcg cag tgc ctt taa Asn Lys Gln Asn Asp Trp Tyr Ser Gln Cys Leu	900
290 295	

<210> 40

<211> 299

<212> PRT

<213> Staphylotrichum coccosporum IF0 31817

<220>

<221> característica miscelánea

5 <221> (1) .. (3)

<223> Ácido piroglutámico

<400> 40

Gln Ser Ala Cys Ala Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys Cys
1 5 10 15

Lys Pro Ser Cys Ser Trp Pro Gly Lys Ala Ser Val Asn Gln Pro Val
20 25 30

Phe Ala Cys Ser Ala Asn Phe Gln Arg Ile Ser Asp Pro Asn Val Lys
35 40 45

Ser Gly Cys Asp Gly Gly Ser Ala Tyr Ala Cys Ala Asp Gln Thr Pro
50 55 60

Trp Ala Val Asn Asp Asn Phe Ser Tyr Gly Phe Ala Ala Thr Ser Ile
65 70 75 80

Ser Gly Gly Asn Glu Ala Ser Trp Cys Cys Gly Cys Tyr Glu Leu Thr
85 90 95

Phe Thr Ser Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser Thr
100 105 110

Ser Thr Gly Gly Asp Leu Gly Thr Asn His Phe Asp Leu Ala Met Pro
115 120 125

ES 2 371 916 T3

Gly Gly Gly Val Gly Ile Phe Asp Gly Cys Ser Pro Gln Phe Gly Gly
130 135 140

Leu Ala Gly Asp Arg Tyr Gly Gly Val Ser Ser Arg Ser Gln Cys Asp
145 150 155 160

Ser Phe Pro Ala Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe Asp Trp
165 170 175

Phe Lys Asn Ala Asp Asn Pro Thr Phe Thr Phe Arg Gln Val Gln Cys
180 185 190

Pro Ser Glu Leu Val Ala Arg Thr Gly Cys Arg Arg Asn Asp Asp Gly
195 200 205

Asn Phe Pro Val Phe Thr Pro Pro Ser Gly Gly Gln Ser Ser Ser Ser
210 215 220

Ser Ser Ser Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr Ser Thr Thr
225 230 235 240

Ser Thr Lys Ala Thr Ser Thr Thr Ser Thr Ala Ser Ser Gln Thr Ser
245 250 255

Ser Ser Thr Gly Gly Gly Cys Ala Ala Gln Arg Trp Ala Gln Cys Gly
260 265 270

Gly Ile Gly Phe Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr Cys
275 280 285

Asn Lys Gln Asn Asp Trp Tyr Ser Gln Cys Leu
290 295

<211> 8

<212> PRT

<213> Staphylotrichum coccosporum IFO 31817

<220>

5 <221> MOD_RES

<221> (1) .. (1)

<223> ÁCIDO PIRROLIDONA CARBOXÍLICO

<220>

<221> MUTÁGENO

10 <221> (1) .. (1)

<223>

<400> 41

Gln Ala Asp Gly Lys Ser Thr Arg
1 5

<210> 42

15 <211> 11

<212> PRT

<213> Staphylotrichum coccosporum IFO 31817

<220>

<221> MOD_RES

20 <221> (1) .. (1)

<223> ÁCIDO PIRROLIDONA CARBOXÍLICO

<220>

<221> MUTÁGENO

<221> (1) .. (4)

25 <223>

<400> 42

Gln Ser Ala Cys Ala Asp Gly Lys Ser Thr Arg

1

5

10

<210> 43

<211> 21

30 <212> PRT

<213> Staphylotrichum coccosporum IFO 31817

ES 2 371 916 T3

<400> 43

Met Arg Ser Ser Pro Val Leu Arg Thr Ala Leu Ala Ala Ala Leu Pro
1 5 10 15

Leu Ala Ala Leu Ala
20

<210> 44

<211> 295

5 <212> PRT

<213> Staphylotrichum coccosporum IFO 31817

<400> 44

Ala Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys Cys Lys Pro Ser Cys
1 5 10 15

Ser Trp Pro Gly Lys Ala Ser Val Asn Gln Pro Val Phe Ala Cys Ser
20 25 30

Ala Asn Phe Gln Arg Ile Ser Asp Pro Asn Val Lys Ser Gly Cys Asp
35 40 45

Gly Gly Ser Ala Tyr Ala Cys Ala Asp Gln Thr Pro Trp Ala Val Asn
50 55 60

Asp Asn Phe Ser Tyr Gly Phe Ala Ala Thr Ser Ile Ser Gly Gly Asn
65 70 75 80

ES 2 371 916 T3

Glu Ala Ser Trp Cys Cys Gly Cys Tyr Glu Leu Thr Phe Thr Ser Gly
85 90 95

Pro Val Ala Gly Lys Thr Met Val Val Gln Ser Thr Ser Thr Gly Gly
100 105 110

Asp Leu Gly Thr Asn His Phe Asp Leu Ala Met Pro Gly Gly Gly Val
115 120 125

Gly Ile Phe Asp Gly Cys Ser Pro Gln Phe Gly Gly Leu Ala Gly Asp
130 135 140

Arg Tyr Gly Gly Val Ser Ser Arg Ser Gln Cys Asp Ser Phe Pro Ala
145 150 155 160

Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe Asp Trp Phe Lys Asn Ala
165 170 175

Asp Asn Pro Thr Phe Thr Phe Arg Gln Val Gln Cys Pro Ser Glu Leu
180 185 190

Val Ala Arg Thr Gly Cys Arg Arg Asn Asp Asp Gly Asn Phe Pro Val
195 200 205

Phe Thr Pro Pro Ser Gly Gly Gln Ser Ser Ser Ser Ser Ser Ser Ser
210 215 220

Ser Ala Lys Pro Thr Ser Thr Ser Thr Ser Thr Thr Ser Thr Lys Ala
225 230 235 240

Thr Ser Thr Thr Ser Thr Ala Ser Ser Gln Thr Ser Ser Ser Thr Gly
245 250 255

ES 2 371 916 T3

Gly Gly Cys Ala Ala Gln Arg Trp Ala Gln Cys Gly Gly Ile Gly Phe
260 265 270

Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr Cys Asn Lys Gln Asn
275 280 285

Asp Trp Tyr Ser Gln Cys Leu
290 295

<210> 45

<211> 20

<212> PRT

5 <213> Humicola insolens

<400> 45

Met Arg Ser Ser Pro Leu Leu Arg Ser Ala Val Val Ala Ala Leu Pro
1 5 10 15

Val Leu Ala Leu
20

<210> 46

<211> 286

10 <212> PRT

<213> Humicola insolens

<400> 46

Ala Ala Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys Cys Lys Pro Ser
1 5 10 15

Cys Gly Trp Ala Lys Lys Ala Pro Val Asn Gln Pro Val Phe Ser Cys
20 25 30

Asn Ala Asn Phe Gln Arg Leu Thr Asp Phe Asp Ala Lys Ser Gly Cys

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

ES 2 371 916 T3

210

215

220

Pro Thr Ser Thr Ser Thr Thr Ser Thr Ser Thr Thr Ser Ser Pro Pro
225 230 235 240

Val Gln Pro Thr Thr Pro Ser Gly Cys Thr Ala Glu Arg Trp Ala Cys
245 250 255

Gln Cys Gly Gly Asn Gly Trp Ser Gly Cys Thr Thr Cys Val Ala Gly
260 265 270

Ser Thr Cys Thr Lys Ile Asn Asp Trp Tyr His Gln Cys Leu
275 280 285

<210> 47

<211> 17

<212> PRT

5 <213> Humicola insolens

<400> 47

Met Gln Leu Pro Leu Thr Thr Leu Leu Thr Leu Leu Pro Ala Leu Ala
1 5 10 15

Ala

<210> 48

<211> 206

10 <212> PRT

<213> Humicola insolens

<400> 48

Ala Gln Ser Gly Ser Gly Arg Thr Thr Arg Tyr Trp Asp Cys Cys Lys
1 5 10 15

ES 2 371 916 T3

Pro Ser Cys Ala Trp Pro Gly Lys Gly Pro Ala Pro Val Arg Thr Cys
20 25 30

Asp Arg Trp Asp Asn Pro Leu Phe Asp Gly Gly Asn Thr Arg Ser Gly
35 40 45

Cys Asp Ala Gly Gly Gly Ala Tyr Met Cys Ser Asp Gln Ser Pro Trp
50 55 60

Ala Val Ser Asp Asp Leu Ala Tyr Gly Trp Ala Ala Val Asn Ile Ala
65 70 75 80

Gly Ser Asn Glu Arg Gln Trp Cys Cys Ala Cys Tyr Glu Leu Thr Phe
85 90 95

Thr Ser Gly Pro Val Ala Gly Lys Arg Met Ile Val Gln Ala Ser Asn
100 105 110

Thr Gly Gly Asp Leu Gly Asn Asn His Phe Asp Ile Ala Met Pro Gly
115 120 125

Gly Gly Val Gly Ile Phe Asn Ala Cys Thr Asp Gln Tyr Gly Ala Pro
130 135 140

Pro Asn Gly Trp Gly Gln Arg Tyr Gly Gly Ile Ser Gln Arg His Glu
145 150 155 160

Cys Asp Ala Phe Pro Glu Lys Leu Lys Pro Gly Cys Tyr Trp Arg Phe
165 170 175

Asp Trp Phe Leu Asn Ala Asp Asn Pro Ser Val Asn Trp Arg Gln Val
180 185 190

Ser Cys Pro Ala Glu Ile Val Ala Lys Ser Gly Cys Ser Arg
195 200 205

REIVINDICACIONES

1. Un procedimiento para suprimir una reducción en la actividad de endoglucanasa en presencia de un tensioactivo, **caracterizado por** la modificación de una proteína que tiene la actividad de endoglucanasa, en la que el aminoácido N terminal es diferente de ácido piroglutámico, a una proteína que tiene el extremo N terminal de ácido piroglutámico.

2. El procedimiento de la reivindicación 1, en el que la modificación se lleva a cabo añadiendo ácido piroglutámico o un aminoácido que puede convertirse en ácido piroglutámico, o un péptido que tiene el extremo N terminal de ácido piroglutámico o un aminoácido que puede convertirse en ácido piroglutámico, al extremo N terminal de la proteína que tiene la actividad de endoglucanasa en la que el extremo N terminal es un aminoácido diferente de ácido piroglutámico, en el que el aminoácido que se puede convertir en ácido piroglutámico se selecciona de ácido glutámico y glutamina.

3. El procedimiento de la reivindicación 1, en el que la modificación se lleva a cabo sustituyendo el aminoácido N terminal o una región N terminal de la proteína que tiene la actividad de endoglucanasa en la que el aminoácido N terminal es diferente de ácido piroglutámico por ácido piroglutámico o un aminoácido que puede convertirse en ácido piroglutámico, o un péptido que tiene el extremo N terminal de ácido piroglutámico o un aminoácido que puede convertirse en ácido piroglutámico, en el que el aminoácido que puede convertirse en ácido piroglutámico se selecciona de ácido glutámico y glutamina.

4. El procedimiento según una cualquiera de las reivindicaciones 1 a 3, en el que la proteína que tiene la actividad de endoglucanasa en la que el extremo N terminal es un aminoácido diferente de ácido piroglutámico es una celulasa perteneciente a la familia 45.

5. Una proteína seleccionada del grupo que consiste en:

(a) una proteína que comprende la secuencia de aminoácidos de la ID. SEC. Nº 2, 4, 38 o 40, en la que el extremo N terminal es ácido piroglutámico;

(b) una proteína modificada que comprende una secuencia de aminoácidos en la que el extremo N terminal es ácido piroglutámico y en la que se han eliminado, sustituido, insertado o añadido de 1 a 30 aminoácidos en la secuencia de aminoácidos de la ID. SEC. Nº 2, 4, 38 o 40, y que tiene una actividad de endoglucanasa; y

(c) una proteína que comprende una secuencia de aminoácidos que tiene al menos el 85% de identidad de secuencia con una proteína que comprende la secuencia de aminoácidos de la ID. SEC. Nº 2, 4, 38 o 40, en la que el extremo N terminal es ácido piroglutámico y que tiene una actividad de endoglucanasa.

Figura 1

NCE4	MRSSPLLRSVVAALPVLAL—————AADGKSTRYWDCCCKPSCGWAK	21

STCE1	MRSSPVLRTALAAALPLAALA—————ADGKSTRYWDCCCKPSCSWPG	20
	* *****	
NCE5	MQPLTTLLTLLPALAA—————AQSGSGRTTRYWDCCCKPSCAWPG	23
	<u>Péptido señal</u> → <u>Dominio catalítico</u> →	
NCE4	KAPVNQPVFSCNANFQRLTDF-DAKSGCEPGGVAYSCADQTPWAVNDFA	70
	** ***** * ***** * ***** ** ** ***** *	
STCE1	KASVNQPVFACSANFORISDP-NVKS GCD-GGSAYACADQTPWAVNDNFS	68
	* ** * * * ***** ** ** * ** ***** *	
NCE5	KGPA—PVRTCDRWDNPLFDGGNTRSGCDAGGGAYMCSDQSPWAVSDDLA—	71
NCE4	FGFAATS IAGSNEAGWCCACYELTFTSGPVAGKKMVVQSTSTGGDLGSNH	120
	***** * *** *** ***** ***** *	
STCE1	YGFAATS ISGGNEASWCCGCYELTFTSGPVAGKTMVVQSTSTGGDLGTNH	118
	** ** * * ** *** ***** * ** ***** *	
NCE5	YGWAAVNIAGSNERQWCCACYELTFTSGPVAGKRM I VQASNTGGDLGNNH	121
NCE4	FDLNI PGGGVG I FDGCTPQFGGLP—GQRYGG I SSRNECDRFPDALKPG	167
	*** ***** ***** * ***** ** ** *****	
STCE1	FDLAMPGGGVG I FDGCS PQFGGLA—GDRYGGVSSRSQCDSFPAALKPG	165
	** ***** * * * * ***** * * ** ** *****	
NCE5	FDIAMPGGGVG I FNACTDQYGAPPNGWGQRYGG I SQRHECD AFPEKLKPG	171

Figura 2

NCE4	CYWRFDWFKNADNPSFSFRQVQCPAELVARTGCRRNDDGNFPAVQIPSSS	217
	***** * ***** ***** **	
STCE1	CYWRFDWFKNADNPTFTFRQVQCPSELVARTGCRRNDDGNFPVFTPPSGG	215
	***** ***** *** ** * ** ** *	
NCE5	CYWRFDWFLNADNPSVNRQVSCPAEIVAKSGCSR-----	206
	Conector →	
NCE4	TSSPVGQPTSTSTTSTTSSPPVQPTTPS-----GCTAERWA	255
	** * ***** * * ** * ** *	
STCE1	QSSSSSSSSSAKPTSTSTTSTKATSTTSTASSQTSSSTGGGCAAQRWA	265
NCE5	-----	
	DUC →	
NCE4	CQCGGNGWSGCTTCVAGSTCTKINDWYHQCL	286
	***** * ***** * ** * *****	
STCE1	CQCGGIGFSGCTTCVSGTTCNKQNDWYSQCL	295
NCE5	-----	