

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 531 402**

51 Int. Cl.:

**C12N 15/00** (2006.01)

**C07K 19/00** (2006.01)

**C12N 1/21** (2006.01)

**C12N 5/071** (2010.01)

**C12N 9/52** (2006.01)

**C12N 15/09** (2006.01)

**C07K 14/78** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **06.11.2009 E 09827488 (9)**

97 Fecha y número de publicación de la concesión europea: **24.12.2014 EP 2363461**

54 Título: **Colagenasa de fusión a la que se une una marca de afinidad, y método para producir la misma**

30 Prioridad:

**19.11.2008 JP 2008295922**

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

**13.03.2015**

73 Titular/es:

**MEIJI SEIKA PHARMA CO., LTD. (50.0%)**  
**4-16, Kyobashi 2-chome Chuo-ku**  
**Tokyo 104-8002, JP y**  
**TOHOKU UNIVERSITY (50.0%)**

72 Inventor/es:

**FUKUSHIMA, TAKAYOSHI;**  
**YOKOYAMA, KENGO;**  
**MURASHIMA, KOICHIRO;**  
**GOTO, MASAFUMI y**  
**YAMAGATA, YOUHEI**

74 Agente/Representante:

**VALLEJO LÓPEZ, Juan Pedro**

**ES 2 531 402 T3**

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

## DESCRIPCIÓN

Colagenasa de fusión a la que se une una marca de afinidad, y método para producir la misma

5 **Campo técnico**

La presente invención se refiere a una colagenasa usada para aislar células (masa celular) de un órgano, tales como células de los islotes pancreáticos de un páncreas.

10 **Antecedentes de la técnica**

Como terapia para la curación para la diabetes, se ha conocido el trasplante de islotes pancreáticos, en el que los islotes pancreáticos (células que producen insulina) aislados de un páncreas tratado con una proteasa se trasplantan en la vena porta de un paciente diabético a través de goteo intravenoso. Este método de trasplante no requiere laparotomía para el paciente durante el trasplante, y por lo tanto recientemente ha atraído la atención como una terapia de curación, que es segura y sencilla, para la diabetes.

Los islotes pancreáticos se aíslan de un páncreas mediante el tratamiento del páncreas con una colagenasa y una metaloproteasa neutra. Como la colagenasa usada para este fin, es particularmente eficaz una colagenasa derivada de *Clostridium histolyticum* (Bibliografías de No Patente 1, 2).

Se ha averiguado que la colagenasa derivada de *Clostridium histolyticum* incluye dos tipos de enzimas que tienen diferente especificidad hacia el sustrato; colagenasa G y colagenasa H (Bibliografías de No Patente 1, 2). Los genes que codifican la colagenasa G y la colagenasa H ya se han aislado, y se ha encontrado que ambas colagenasas son enzimas de dominios múltiples que incluyen un dominio catalítico en el lado del extremo amino y un dominio de unión a colágeno (en lo sucesivo en el presente documento, denominado "CBD") en el lado del extremo carboxilo (Bibliografías de No Patente 1, 2).

Sin embargo, se ha encontrado el problema de que, cuando se produce una colagenasa con *Clostridium histolyticum*, parte de la colagenasa expresada se degrada (Bibliografía de No Patente 2), y el tratamiento de un páncreas con una colagenasa mezclada con la colagenasa degradada disminuye la calidad de los islotes pancreáticos aislados. Por lo tanto, para aislar los islotes pancreáticos, es deseable tratar un páncreas con una colagenasa obtenida retirando tal colagenasa degradada tanto como sea posible a partir del mismo.

Sin embargo, una colagenasa no degradada y una colagenasa degradada son similares en las propiedades fisicoquímicas, y por lo tanto ha sido difícil separar estas colagenasas con métodos tales como cromatografía de intercambio iónico o cromatografía hidrofóbica.

En tales antecedentes, se ha deseado un método para preparar una colagenasa por recogida selectiva de una colagenasa no degradada a partir de colagenasas derivadas de *Clostridium histolyticum*.

Se debería indicar que aún no se ha informado la degradación de una colagenasa derivada de *Clostridium histolyticum* con la colagenasa siendo expresada como una proteína recombinante en un huésped tal como *E. coli*.

45 [Lista de citas]

[Bibliografías de no patente]

50 [NPL 1] Yoshihara, K. *et. al.* Journal of Bacteriology. (1994), 176, 6489-6496

[NPL 2] Matsushita, O. *et. al.* Journal of Bacteriology. (1999), 181, 923-933

**Sumario de la invención**

55 BIOMOLECULAR NMR ASSIGNMENTS, vol. 2, nº 2, p. 127-129 desvela una proteína de fusión que comprende una colagenasa de *C. histolyticum* y una marca de GST. El dominio de unión a colágeno C-terminal de la colagenasa está condensado con el extremo C de la marca de GST. APL MICROBIOL BIOTECHNOL (2008) 19: 627-635 describe una colagenasa de fusión en la que una marca de afinidad que contiene seis Histidinas se une a los extremos carboxilo de la colagenasa expresada en *Clostridium perfringens*.

60 **Problema técnico**

La presente invención tiene como objetivo la recogida de forma selectiva de una colagenasa no degradada por retirada de una colagenasa degradada de colagenasas derivadas de *Clostridium histolyticum*.

65

**Solución al problema**

Con respecto a las colagenasas derivadas de *Clostridium histolyticum*, los presentes inventores prepararon colagenasas de fusión en las que marcas de afinidad unidas a los extremos carboxilo de las colagenasas se expresaban como proteínas recombinantes. Cuando una colagenasa derivada de *Clostridium histolyticum* se expresa como una proteína recombinante en un huésped tal como *E. coli*, no se ha predicho bien si la proteína recombinante se degrada por la acción de una proteasa en un huésped o no, ni cómo se degrada la proteína recombinante suponiendo que se degrade. Sin embargo, los presentes inventores purificaron la proteína expresada en un huésped con una columna de afinidad, y encontraron de forma accidental que se retiraba una colagenasa degradada y que se recogía de forma selectiva una sola colagenasa que tenía una actividad de colagenasa. Esto se basa en las siguientes suposiciones: una colagenasa derivada de *Clostridium histolyticum* expresada en un huésped se degradó en el CBD en las cercanías de la marca de afinidad fusionada, y por lo tanto la degradación del CBD permitió que la marca de afinidad se separara de la colagenasa; y en la etapa de purificación por cromatografía por afinidad, la colagenasa degradada no se adsorbía a la columna de afinidad sino que solo se adsorbía una colagenasa no degradada a la columna de afinidad.

En otras palabras, los presentes inventores encontraron que si una marca de afinidad se une a una colagenasa en una configuración específica, una sola colagenasa que tiene una actividad de colagenasa se puede recoger de forma selectiva en el proceso de purificación por afinidad sin que se adsorba la colagenasa degradada en un huésped. Este descubrimiento ha llevado a los inventores a la finalización de la presente invención.

La presente invención se refiere a un método para producir una colagenasa de fusión, en el que:

un cultivo obtenido mediante el cultivo de *E. coli* transformada por uno cualquiera de un ADN que codifica la colagenasa de fusión y un vector de expresión que comprende el ADN se purifica mediante cromatografía por afinidad que corresponde a una marca de afinidad para recoger de ese modo de forma selectiva la colagenasa de fusión que tiene un dominio de unión a colágeno, en el que la colagenasa de fusión tiene las siguientes características (i) a (ii):

- (i) la marca de afinidad se une directa o indirectamente a un extremo carboxilo de la colagenasa;  
(ii) la colagenasa se selecciona entre los siguientes (a) a (h):

- (a) una colagenasa que comprende una secuencia de aminoácidos de las posiciones -110 a 1008 de una cualquiera de las SEC ID N<sup>o</sup>: 1 y 2;  
(b) una colagenasa que comprende una secuencia de aminoácidos de las posiciones -110 a 1008 de una cualquiera de las SEC ID N<sup>o</sup>: 1 y 2 en la que se suprimen, sustituyen, insertan, o añaden de 1 a 30 aminoácidos;  
(c) una colagenasa que comprende una secuencia de aminoácidos que tiene una homología de un 80 % o superior con la secuencia de aminoácidos de las posiciones -110 a 1008 de una cualquiera de las SEC ID N<sup>o</sup>: 1 y 2;  
(d) una colagenasa en la que una secuencia de aminoácidos de las posiciones -110 a -1 de una cualquiera de las SEC ID N<sup>o</sup>: 1 y 2 se retira de una cualquiera de las colagenasas que se han descrito en (a) a (c);  
(e) una colagenasa que comprende una secuencia de aminoácidos de las posiciones -40 a 981 de una cualquiera de las SEC ID N<sup>o</sup>: 5 y 6;  
(f) una colagenasa que comprende una secuencia de aminoácidos de las posiciones -40 a 981 de una cualquiera de las SEC ID N<sup>o</sup>: 5 y 6 en la que se suprimen, sustituyen, insertan, o añaden de 1 a 30 aminoácidos;  
(g) una colagenasa que comprende una secuencia de aminoácidos que tiene una homología de un 80 % o superior con la secuencia de aminoácidos de las posiciones -40 a 981 de una cualquiera de las SEC ID N<sup>o</sup>: 5 y 6;  
(h) una colagenasa en la que una secuencia de aminoácidos de las posiciones -40 a -1 de una cualquiera de las SEC ID N<sup>o</sup>: 5 y 6 se retira de una cualquiera de las colagenasas que se han descrito en (e) a (g).

La colagenasa de fusión de acuerdo con la invención, en la que la marca de afinidad se une directa o indirectamente a un extremo carboxilo de la colagenasa.

La colagenasa de fusión de acuerdo con la invención, en la que la marca de afinidad es dos o más restos de histidina consecutivos.

La colagenasa de fusión de acuerdo con la invención, en la que la colagenasa se deriva de *Clostridium histolyticum*.

La colagenasa de fusión de acuerdo con la invención puede comprender una secuencia de aminoácidos de las posiciones -110 a 1021 de la SEC ID N<sup>o</sup>: 3. En particular, a partir de la colagenasa de fusión que se ha mencionado anteriormente se retira una secuencia de aminoácidos de las posiciones -110 a -1 de la SEC ID N<sup>o</sup>: 3.

La colagenasa de fusión de acuerdo con la invención, comprende una secuencia de aminoácidos de las posiciones -40 a 994 de la SEC ID N°: 7. En particular, a partir de la colagenasa de fusión que se ha mencionado anteriormente se retira una secuencia de aminoácidos de las posiciones -40 a -1 de la SEC ID N°: 7.

5 En el presente documento se describe un ADN que codifica la colagenasa de fusión de acuerdo con la invención.

En el presente documento se describe un ADN que comprende una secuencia de bases de las posiciones 1 a 3396 de la SEC ID N°: 4.

10 En el presente documento se describe un ADN que comprende una secuencia de bases de las posiciones 1 a 3105 de la SEC ID N°: 8.

En el presente documento se describe un vector de expresión que comprende el ADN que se ha mencionado anteriormente.

15 Además, en el presente documento se describe una célula huésped transformada con uno cualquiera del ADN y el vector de expresión que se han mencionado anteriormente. La célula huésped es *E. coli*.

20 En el presente documento se describe una colagenasa de fusión producida con el método de acuerdo con la invención.

### Efecto ventajoso de la invención

25 La presente invención proporciona una colagenasa de fusión en la que se une una marca de afinidad a una colagenasa derivada de *Clostridium histolyticum*, en la que la colagenasa y la marca de afinidad se unen entre sí de tal modo que un fragmento que tiene una actividad de colagenasa y la marca de afinidad se separan entre sí cuando la colagenasa de fusión expresada en un huésped se degrada por acción del huésped.

30 La presente invención también proporciona un ADN necesario para producir la colagenasa de fusión como una proteína recombinante de forma eficaz, y una célula huésped que expresa la colagenasa de fusión como una proteína recombinante. Además, la presente invención proporciona un método capaz de recoger de forma selectiva una sola colagenasa de fusión que tiene un CBD por retirada de una colagenasa en la que partió todo de un CBD se degrada a partir de una solución de cultivo obtenida por cultivo de una célula huésped que expresa la colagenasa de fusión. La presente invención conduce a una producción eficaz de una colagenasa de fusión que tiene un CBD.

35 **Breve descripción de las figuras**

[Fig. 1] La Figura 1 es una figura que muestra una estructura de un pColG plásmido. En la figura, los significados de los símbolos son tal como sigue a continuación (en la figura se subrayan los tipos oblicuos). colG: un gen de colagenasa G, PlacZ: un promotor de lacZ, y Amp<sup>r</sup>: un gen de resistencia a la ampicilina.

[Fig. 2] La Figura 2 es una figura que muestra una estructura de un pColG-His plásmido. En la figura, los significados de los símbolos son tal como sigue a continuación (en la figura se subrayan los tipos oblicuos). colG: un gen de colagenasa G, PlacZ: un promotor de lacZ, Amp<sup>r</sup>: un gen de resistencia a la ampicilina, MCS: un sitio de clonación múltiple, y 6XHis: una marca de histidina.

45 [Fig. 3] La Figura 3 es una imagen de electroforesis que muestra el resultado de la actividad de tinción realizada en un extracto de *E. coli* que tiene colagenasa G de fusión expresada.

[Fig. 4] La Figura 4 es una imagen de electroforesis que muestra el resultado de la actividad de tinción realizada en una solución de colagenasa G de fusión sometida a cromatografía por afinidad.

50 [Fig. 5] La Figura 5 es una figura que muestra una estructura de un plásmido pColH. En la figura, los significados de los símbolos son tal como sigue a continuación (en la figura se subrayan los tipos oblicuos). colH: un gen de colagenasa H, PlacZ: un promotor de lacZ, y Amp<sup>r</sup>: un gen de resistencia a la ampicilina.

[Fig. 6] La Figura 6 es una figura que muestra una estructura de un plásmido pColH-His. En la figura, los significados de los símbolos son tal como sigue a continuación (en la figura se subrayan los tipos oblicuos). colH: un gen de colagenasa H, PlacZ: un promotor de lacZ, Amp<sup>r</sup>: un gen de resistencia a la ampicilina, MCS: un sitio de clonación múltiple, y 6XHis: una marca de histidina.

[Fig. 7] La Figura 7 es una imagen de electroforesis que muestra el resultado de la actividad de tinción realizada en un extracto de *E. coli* que tiene colagenasa H de fusión expresada.

[Fig. 8] La Figura 8 es una imagen de electroforesis que muestra el resultado de la actividad de tinción realizada en una solución de colagenasa H de fusión sometida a cromatografía por afinidad.

60 **Descripción de realizaciones**

### (Colagenasa de fusión)

65 En la presente invención, una colagenasa de fusión se refiere a una proteína en la que una marca de afinidad se une directa o indirectamente a una colagenasa, es decir una colagenasa de fusión en la que la colagenasa y la marca de

afinidad se unen entre sí de tal modo que un fragmento que tiene una actividad de colagenasa y la marca de afinidad en la colagenasa de fusión se separan entre sí cuando la colagenasa de fusión expresada en un huésped se degrada por una acción de huésped. La forma de unión de tal colagenasa a la marca de afinidad es una unión preferentemente a un CBD de la colagenasa, y lo más preferentemente una unión al extremo carboxilo del CBD (es decir el extremo carboxilo de la colagenasa). En la presente invención, el "CBD" se refiere a regiones de de los segmentos 3a y 3b en la colagenasa G (posición 776 del extremo carboxilo de cada una de las SEC ID N°: 1 y 2), y una región de un segmento 3 en la colagenasa H (de la posición 864 al extremo carboxilo de la SEC ID N°: 5) (Bibliografía de No Patente 2).

En la presente invención, como una colagenasa se puede usar cualquier colagenasa siempre y cuando se puedan aislar islotes pancreáticos de un páncreas al tratarlos en combinación con una metaloproteasa neutra. De forma particular, se desea el uso de una colagenasa derivada de *Clostridium histolyticum*. Se puede usar cualquier colagenasa derivada de *Clostridium histolyticum*, pero se desea en particular el uso de la colagenasa G de una cualquiera de las SEC ID N°: 1 y 2, o la colagenasa H de una cualquiera de las SEC ID N°: 5 y 6. Además, tal colagenasa puede incluir una secuencia de aminoácidos que tiene retirada toda o parte de una secuencia de señal.

Aquí, la secuencia de aminoácidos de la colagenasa G puede ser cualquiera de las siguientes siempre y cuando se mantengan la actividad de colagenasa y la unión al colágeno: (i) una colagenasa que comprende toda o parte de una secuencia de aminoácidos de las posiciones -110 a 1008 de una cualquiera de las SEC ID N°: 1 y 2; (ii) una colagenasa que comprende una secuencia de aminoácidos de las posiciones -110 a 1008 de una cualquiera de las SEC ID N°: 1 y 2 en la que se suprimen, sustituyen, insertan, añaden uno o más aminoácidos; (iii) una colagenasa que comprende una secuencia de aminoácidos que tiene una homología de un 70 % o superior con la secuencia de aminoácidos de las posiciones -110 a 1008 de una cualquiera de las SEC ID N°: 1 y 2; y (iv) una colagenasa en la que toda o parte de una secuencia de señal que incluye una secuencia de aminoácidos de las posiciones -110 a -1 de una cualquiera de las SEC ID N° : 1 y 2 se retira a partir de una de las colagenasas que se describen en (i) a (iii).

De forma análoga, la secuencia de aminoácidos de la colagenasa H puede ser cualquiera de las siguientes siempre y cuando se mantengan la actividad de colagenasa y la unión al colágeno: (i) una colagenasa que comprende toda o parte de una secuencia de aminoácidos de las posiciones -40 a 981 de una cualquiera de las SEC ID N°: 5 y 6; (ii) una colagenasa que comprende una secuencia de aminoácidos de las posiciones -40 a 981 de una cualquiera de las SEC ID N°: 5 y 6 en la que se suprimen, sustituyen, insertan, añaden uno o más aminoácidos; (iii) una colagenasa que comprende una secuencia de aminoácidos que tiene una homología de un 70 % o superior con la secuencia de aminoácidos de las posiciones -40 a 981 de una cualquiera de las SEC ID N°: 5 y 6; y (iv) una colagenasa en la que toda o parte de una secuencia de señal que incluye una secuencia de aminoácidos de las posiciones -40 a -1 de una cualquiera de las SEC ID N°: 5 y 6 se retira a partir de una de las colagenasas que se describen en (i) a (iii).

Aquí, la "secuencia de aminoácidos en la que se suprimen, sustituyen, insertan, añaden uno o más aminoácidos" se refiere a que la secuencia de aminoácidos se modifica con métodos bien conocidos tales como mutagénesis dirigida al sitio, o sustitución o similares de aminoácidos con el mismo número de aminoácidos sustituidos de forma natural. El número de modificación de los aminoácidos es preferentemente de 1 a 50, más preferentemente 1 a 30, además preferentemente de 1 a 10, aún además preferentemente de 1 a 5, y lo más preferentemente de 1 a 2. Un ejemplo de la secuencia de aminoácidos modificada puede ser preferentemente una secuencia de aminoácidos que tiene una o más sustituciones conservadoras (preferentemente, de 1 a varias, o 1, 2, 3 o 4) de aminoácidos de la misma. Aquí, la "sustitución conservadora" se refiere a que al menos un resto de aminoácido está sustituido con otro resto de aminoácido químicamente similar. Los ejemplos incluyen un caso en el que un determinado resto hidrófobo está sustituido con otro resto hidrófobo, un caso en el que un determinado resto polar está sustituido con otro resto polar que tiene la misma carga, otros casos. En la técnica se conocen aminoácidos funcionalmente similares que se pueden someter a tal sustitución para cada aminoácido. Los ejemplos específicos de aminoácidos no polares (hidrófobos) incluyen alanina, valina, isoleucina, leucina, prolina, triptófano, fenilalanina, metionina, etc. los ejemplos específicos de aminoácidos polares (neutros) incluyen glicina, serina, treonina, tirosina, glutamina, asparagina, cisteína, etc. los ejemplos específicos de aminoácidos cargados de forma positiva (básicos) incluyen arginina, histidina, lisina, etc. Además, los ejemplos específicos de aminoácidos cargados de forma negativa (ácidos) incluyen ácido aspártico, ácido glutámico, etc.

Además, la "secuencia de aminoácidos que tiene una homología de un 70 % o superior" puede ser una secuencia de aminoácidos que tiene una homología de, preferentemente un 80 % o superior, más preferentemente un 85 % o superior, además preferentemente un 90 % o superior, aún además preferentemente un 95 % o superior, de forma particularmente preferente un 98 % o superior, y lo más preferentemente un 99 % o superior. El término "homología" con respecto a la secuencia de bases o a la secuencia de aminoácidos se usa como un significado del grado de correspondencia, entre secuencias a comparar, en restos de bases o aminoácidos que constituyen cada secuencia. Cada uno de los valores numéricos de la "homología" que se describe en la presente descripción puede ser cualquier valor numérico calculado usando un programa de búsqueda de homología conocido por los expertos en la materia, y se puede calcular fácilmente usando los parámetros por defecto (ajuste inicial) en FASTA, BLAST, etc., por ejemplo.

En la presente invención, como la marca de afinidad a fusionar directa o indirectamente a la colagenasa, se puede usar cualquier marca de afinidad siempre y cuando se pueda unir de forma selectiva a un determinado tipo de vehículo. Por ejemplo, se puede usar una marca de histidina que incluye dos o más restos de histidina consecutivos que se unen de forma selectiva a una columna de quelato de níquel que se describe en la Patente Japonesa N° 2686090, un dominio de unión a celulosa que se une de forma selectiva a una celulosa insoluble, un dominio de unión a maltosa que se une de forma selectiva a una resina de unión a maltosa, y similares. De forma particular, se desea el uso de la marca de histidina que tiene un peso molecular menor.

En la presente invención, la marca de afinidad se une directa o indirectamente a la colagenasa. Aquí, cuando la colagenasa y la marca de afinidad se unen indirectamente entre sí, la secuencia de aminoácidos que interviene entre ellas puede ser cualquier secuencia siempre y cuando no inhiba materialmente la actividad de la colagenasa de fusión y la unión al colágeno.

Los ejemplos de la colagenasa de fusión que satisface los requisitos anteriores incluyen una colagenasa de fusión que comprende toda una secuencia de aminoácidos de la SEC ID N°: 3 en la que la marca de histidina se une al extremo carboxilo de la colagenasa G, y una colagenasa de fusión que comprende toda una secuencia de aminoácidos de la SEC ID N°: 7 en la que la marca de histidina se une al extremo carboxilo de la colagenasa H. Además, los ejemplos incluyen una colagenasa de fusión que comprende parte de tal secuencia de aminoácidos siempre y cuando se mantenga la actividad de colagenasa, la unión al colágeno, y la unión a la marca de afinidad.

(Secuencia de aminoácidos que codifica un ADN de colagenasa de fusión)

En la presente invención, un ADN que codifica la colagenasa de fusión puede ser un ADN que comprende cualquier secuencia de bases siempre y cuando el ADN codifique la secuencia de aminoácidos de la colagenasa de fusión que se ha descrito anteriormente.

El ADN que codifica la colagenasa de fusión de la presente invención se puede obtener por síntesis química artificial. Además, el ADN que codifica la colagenasa de fusión se puede obtener por construcción de un ADN que codifica la colagenasa y un ADN que codifica la marca de afinidad de forma separada y uniéndolos en conjunto. En este caso, el gen de la colagenasa se puede amplificar por PCR usando un cebador sintetizado basándose en la secuencia del gen y con un molde de un ADN, tal como ADN genómico, ADNc, y plásmido, que incluye el gen. Por ejemplo, la colagenasa G o la colagenasa H derivadas de *Clostridium histolyticum* se pueden amplificar por PCR usando un cebador diseñado basándose en la secuencia del extremo en la posición 5' y el extremo en la posición 3' de la secuencia del gen de la colagenasa G que se describe en la Bibliografía de No Patente 1 o el gen de la colagenasa H que se describe en la Bibliografía de No Patente 2 y con un molde de un ADN genómico derivado de *Clostridium histolyticum*. Además, el ADN que codifica la marca de afinidad se puede amplificar por PCR con un molde de ADN genómico, ADNc, plásmido, o similar que incluye el gen que codifica la marca de afinidad. Por ejemplo, un ADN que codifica la marca de histidina se puede amplificar por PCR con un molde de pET-24a(+), un vector disponible en el mercado.

Los ejemplos del ADN que satisface los requisitos mencionados anteriormente incluyen un ADN que comprende toda o parte de la secuencia de bases de la SEC ID N°: 4 que codifica la colagenasa de fusión en la que la marca de histidina se une al extremo carboxilo de la colagenasa G; y un ADN que comprende toda o parte de la secuencia de bases de la SEC ID N°: 8 para unir la marca de histidina al extremo carboxilo de la colagenasa H.

(Vector de expresión, y célula huésped transformada por el vector de expresión)

La presente invención proporciona un vector de expresión que comprende el ADN que codifica la secuencia de aminoácidos de la colagenasa de fusión que se ha descrito anteriormente, que es capaz de replicar el ADN en una célula huésped, y que comprende una proteína codificada por la secuencia de ADN en un estado de expresión. El presente vector de expresión se puede construir basándose en un vector de autorreplicación, es decir, por ejemplo, un plásmido que existe en forma de un elemento extracromosómico independiente y que se replica independientemente de la replicación del cromosoma. Además, el presente vector de expresión puede ser cualquier vector que se introduce en una célula huésped, a continuación se incorpora en el genoma de la célula huésped, y se replica junto con el cromosoma incorporado con el mismo. Como un procedimiento y un método para construir el vector de acuerdo con la presente invención, se puede usar cualquier procedimiento y cualquier método usado habitualmente en el campo de la ingeniería genética.

El vector de expresión de acuerdo con la presente invención, para expresar la colagenasa de fusión cuando el vector de expresión se introduce realmente en la célula huésped, comprende de forma deseable, además del ADN que codifica la secuencia de aminoácidos de la colagenasa de fusión que se ha descrito anteriormente, una secuencia de ADN para regular la expresión del ADN, un marcador genético para seleccionar la célula huésped transformada, y similares. La secuencia de ADN para regular la expresión incluye una secuencia de ADN que codifica un promotor, terminador, un péptido de señal, y similares. El promotor no se limita de forma en particular siempre y cuando presente la actividad de transcripción en una célula huésped, y se pueda obtener como la secuencia de ADN para regular la expresión del gen que codifica una proteína homogénea o heterogénea con respecto a la célula huésped.

Además, el péptido de señal no se limita de forma en particular siempre y cuando contribuya a la secreción de la proteína en la célula huésped, y se pueda obtener a partir de la secuencia de ADN derivada del gen que codifica una proteína homogénea o heterogénea con respecto a la célula huésped. Además, en la presente invención el marcador genético se puede seleccionar de forma apropiada dependiendo del método para seleccionar el transformante, y se puede usar, por ejemplo, un gen que codifica la resistencia al fármaco y un gen que complementa la auxotrofia.

Además, de acuerdo con la presente invención, se proporciona una célula huésped transformada por este vector de expresión. Este sistema de huésped-vector no se limita de forma particular, y se puede usar un sistema que usa, por ejemplo, *E. coli*, actinomiceto, levadura, hongo filamentoso, célula animal, etc., o similares. De forma particular, como huésped se usa de forma deseable *E. coli*.

Además, la célula huésped se puede transformar con tales vectores de expresión de acuerdo con métodos usados habitualmente en la técnica.

(Cultivo de célula huésped que expresa colagenasa de fusión, y recogida selectiva de colagenasa de fusión que tiene CBD mediante cromatografía por afinidad)

En la presente invención, la colagenasa de fusión se puede obtener a partir de un cultivo que se obtiene mediante el cultivo de la célula huésped que expresa la colagenasa de fusión se puede cultivar en un medio adecuado. El cultivo y su condición para la célula huésped que expresa la colagenasa de fusión pueden ser básicamente los mismos y éstos se usan para la célula huésped.

La colagenasa de fusión secretada en la célula huésped que expresa la colagenasa de fusión o en la solución de cultivo se puede recoger adoptando métodos usados habitualmente en la técnica.

La colagenasa de fusión recogida de acuerdo con el método anterior se somete a cromatografía por afinidad para retirar una colagenasa degradada, recogiendo de ese modo de forma selectiva la colagenasa de fusión que tiene el CBD. Es necesario que la cromatografía por afinidad usada en este caso sea un método que corresponda a la marca de afinidad unida a la colagenasa. Por ejemplo, cuando la marca de histidina se une a la colagenasa, se usa una columna de quelato de níquel o similar a la que la marca de histidina se une de forma selectiva para purificación.

La extensión hasta la que se retira una colagenasa degradada mediante cromatografía por afinidad se puede evaluar sometiendo la solución de colagenasa de fusión a electroforesis usando gel con gelatina añadida al mismo y a continuación la actividad de tinción.

(Otros aspectos)

De acuerdo con el conocimiento de los presentes inventores acerca de que la expresión una colagenasa recombinante por *E. coli* o similar hace que parte o todo su CBD se degrade, la colagenasa degradada se puede eliminar y una colagenasa recombinante no degradada se puede purificar por afinidad de forma específica usando un anticuerpo que se une al CBD de la colagenasa recombinante. En este caso, es ventajoso en que no es necesario que la marca de afinidad se fusione con la colagenasa recombinante. De forma específica, la presente invención también proporciona un método para producir una colagenasa, caracterizado por que un cultivo obtenido mediante el cultivo de una célula huésped que tiene una colagenasa recombinante expresada se purifica con un anticuerpo que se une a un CBD de la colagenasa recombinante para recoger de ese modo de forma selectiva una colagenasa que tiene el CBD. Además, dado que el CBD y el colágeno tienen una afinidad entre sí, la colagenasa recombinante se puede purificar por afinidad de forma específica usando colágeno en lugar del anticuerpo que se ha descrito anteriormente. De forma específica, la presente invención también proporciona un método para producir una colagenasa, en el que un cultivo obtenido mediante el cultivo de una célula huésped que tiene una colagenasa recombinante expresada se purifica con colágeno para recoger de ese modo de forma selectiva una colagenasa que tiene el CBD.

### Ejemplos

La presente invención se describirá de forma más específica por medio de de Ejemplos, pero la presente invención no se va a limitar a los Ejemplos que siguen a continuación pero aún están dentro del punto esencial de la presente invención.

[Ejemplo 1] Expresión de colagenasa de fusión (colagenasa G de fusión) en la que se une marca de histidina a extremo carboxilo de la colagenasa G derivada de *Clostridium histolyticum*

(1-1) Preparación de fragmento de gen de colagenasa G

Se amplificó un gen de colagenasa G derivado de *Clostridium histolyticum* desde el extremo en la posición 5' al extremo la posición 3' del mismo mediante PCR con un molde de ADN genómico derivado de *Clostridium*

*histolyticum*. En este caso, los cebadores se diseñaron de modo que el extremo la posición 3' del gen amplificado de la colagenasa G sirvieran como una secuencia de reconocimiento de XbaI y adicionalmente se añadió una secuencia de reconocimiento de BamHI posteriormente al codón de parada del gen. Como resultado, se introdujeron mutaciones en la secuencia de aminoácidos en dos posiciones (aminoácidos en las posiciones 1007 y 1008 de la SEC ID N°: 1) en el plato del extremo carboxilo de la colagenasa G natural que se describe en la Bibliografía de No Patente 1. Por lo tanto, la secuencia de aminoácidos se modificó a la secuencia de aminoácidos de la SEC ID N°: 2.

Los cebadores usados en esta PCR fueron los que siguen a continuación .

10 colG-F:  
ATGAAAAAAAAATATTTTAAAGATTC (SEC ID N°: 9)  
colG-R:  
CCGGATCCTATCTAGATACCCTTAACT (SEC ID N°: 10)

15 El fragmento amplificado del gen de la colagenasa G se digirió con BamHI.

(1-2) Preparación de fragmento de ADN que incluye la región que codifica marca de histidina

20 Para preparar un ADN que codifica la marca de histidina que incluye seis restos de histidina consecutivos, el ADN que codifica la marca de histidina se amplificó mediante PCR con un molde de pET-24a(+), un vector disponible en el mercado. En el fragmento de ADN, además del ADN que codifica la marca de histidina, se incluyeron un sitio de clonación múltiple originado a partir del vector y un fragmento de gen que corresponde a un terminador T7. Los cebadores se diseñaron para que incluyeran la secuencia de reconocimiento de XbaI en el extremo en la posición 5' del fragmento de ADN amplificado y la secuencia de reconocimiento de BamHI en el extremo en la posición 3' del mismo.

Los cebadores usados en esta PCR fueron los que siguen a continuación.

30 His-F:  
GCTCTAGAAAGCTTGCGGCCGCACTCGA (SEC ID N°: 11)  
His-R:  
CGGGATCCGGATATAGTTCCTCCT (SEC ID N°: 12)

35 El fragmento de ADN amplificado que incluye la región que codifica la marca de histidina se digirió doblemente con XbaI y BamHI.

(1-3) Preparación de fragmento de promotor lacZ

40 Un promotor lacZ se preparó como un promotor para expresar la colagenasa de fusión. El fragmento de ADN se amplificó mediante PCR con un molde de pUC19. En este caso, los cebadores se diseñaron para que incluyeran una secuencia de reconocimiento de HindIII en el extremo en la posición 5' del fragmento de ADN amplificado.

Los cebadores usados en esta PCR fueron los que siguen a continuación.

45 lac-F:  
CCGGCAAGCTTGCCCAATACGCAAACCG (SEC ID N°: 13)  
lac-R:  
AGCTGTTTCCTGTGTGAA (SEC ID N°: 14)

50 El fragmento del promotor lacZ amplificado se digirió con HindIII.

(1-4) Construcción de vector de expresión para la colagenasa G de fusión a la que se une una marca de histidina

55 Se insertaron tres tipos de fragmentos de ADN preparados con el método que se ha descrito anteriormente en pBR322, un vector disponible en el mercado, con el fin de unirlos en el orden del promotor lacZ, el gen de la colagenasa G, y el fragmento de ADN que incluye la marca de histidina del extremo en la posición 5'. En primer lugar, el promotor lacZ y el gen de la colagenasa G se insertaron en pBR322. De forma específica, la región del promotor lacZ y el fragmento de gen de la colagenasa G preparado tal como se ha descrito anteriormente se fosforilaron, y a continuación se insertaron en pBR322 que se había digerido doblemente con HindIII y BamHI para construir pColG (Fig. 1). Se debe indicar que la región del promotor del gen lacZ (PlacZ) y el gen de colG se unieron entre sí en sus extremos truncados. Posteriormente, el ADN que codifica la marca de histidina preparado con el método que se ha descrito anteriormente se insertó en pColG que se había digerido doblemente con XbaI y BamHI para construir pColG-His (Fig. 2). El ADN que se insertó por último en pBR322 fue un ADN que tenía la secuencia de bases de las posiciones 1 a 3396 del Listado de Secuencias: 4.

65

(1-5) Preparación de *E. coli* que expresa la colagenasa G de fusión

pColG-His se transformó en *E. coli*, cepa  $\chi$ 1776 de *Escherichia coli*, de acuerdo con el método convencional, y se cultivó a 37 °C durante todo un día en agar LB al que se habían añadido 20 µg/ml de ácido diaminopimélico, 100 µg/ml de timidina, y 50 µg/ml de ampicilina. Por lo tanto, se preparó la *E. coli* que expresa la colagenasa G de fusión.

[Ejemplo 2] Cultivo de *E. coli* que expresa la colagenasa G de fusión, y recogida selectiva de la colagenasa G que tiene CBD

(2-1) Cultivo de *E. coli* que expresa la colagenasa G de fusión

La *E. coli* que expresa la colagenasa G de fusión obtenida en el Ejemplo 1 se inoculó en un matraz de Erlenmeyer 250 ml añadiendo 100 ml de un medio, y se cultivó con agitación a 200 rpm a 28 °C durante 16 horas. El medio usado en este cultivo era un medio de TB (triptona al 1,2 %, extracto de levadura al 2,4 %, hidrogenofosfato dipotásico al 0,94 %, dihidrogenofosfato potásico al 0,22 %, glicerol al 0,8 %) al que se añadieron 100 µg/ml de ácido diaminopimélico, 20 µg/ml de timidina, 50 µg/ml de ampicilina, y 0,1 mM de IPTG.

(2-2) Preparación de extracto de *E. coli* que expresa la colagenasa G de fusión

La solución de cultivo resultante obtenida en (2-1) se sometió a centrifugación para recoger células, seguido de lisis de las células recogidas en 10 ml de Reactivo de cultivo POP (fabricado por Merck & Co., Inc.) para extraer una proteína en las células. El sobrenadante obtenido por centrifugación del lisado se filtró con una membrana de 0,2 µm para retirar un gen recombinante, proporcionando de ese modo un extracto de la *E. coli* que expresa la colagenasa G de fusión.

(2-3) Recogida selectiva de la colagenasa G de fusión que tiene CBD mediante cromatografía por afinidad

Para retirar una colagenasa degradada del extracto de *E. coli* que expresa la colagenasa G de fusión obtenida en (2-2), el extracto se fraccionó con una columna de quelato de níquel que se sometió a cromatografía por afinidad para la marca de histidina. A 10 ml del extracto de la *E. coli* que expresa la colagenasa G de fusión preparada con el método que se ha descrito anteriormente, se añadieron 60 ml de un tampón (tampón de fosfato sódico 20 mM (pH 7,5) al que se añadieron NaCl 0,5 M e imidazol 20 mM) para unión a la columna de quelato de níquel. La mezcla se pasó a través de una columna de quelato de níquel de 100 ml equilibrada con un tampón para la unión a la columna de quelato de níquel. A continuación, la columna se lavó con una cantidad adecuada de un tampón para la unión a la columna de quelato de níquel para retirar una colagenasa degradada que no había sido capaz de ser adsorbida a la columna de quelato de níquel. A partir de ese momento, se pasaron a través de la columna 100 ml de un tampón para unión a la columna de quelato de níquel a la que se había añadido imidazol 500 mM. Por lo tanto, se recogió la colagenasa G de fusión que tenía el CBD.

(2-4) Confirmación de retirada de colagenasa degradada

Para confirmar que la colagenasa degradada se retiró del extracto de la colagenasa G de fusión, la actividad de tinción se realizó en el extracto de la *E. coli* que expresa la colagenasa G de fusión y la solución de la colagenasa G de fusión se sometió a la cromatografía por afinidad. La actividad de tinción se realizó en 0,25 µl del extracto de la *E. coli* que expresa la colagenasa G de fusión obtenida en (2-2) y 2,5 µl de la solución de la colagenasa G de fusión sometida a la cromatografía por afinidad obtenida en (2-3) con Zymogram-PAGE mini (fabricado por TEFCO). Como resultado, se observaron cinco bandas que indicaban actividad de proteasa para el extracto de la *E. coli* que expresa la colagenasa G de fusión (Fig. 3). Por otro lado, para la solución de la colagenasa G de fusión sometida a la cromatografía por afinidad, se observó una banda que indicaba el peso molecular máximo como una banda principal entre las cinco bandas mencionadas anteriormente (Fig. 4). Basándose en la comparación y el análisis de estas bandas, se encontró que parte o todo el CBD se degradaba cuando la colagenasa G de fusión se expresaba por *E. coli*. Los resultados anteriores mostraron que la colagenasa de fusión que tenía el CBD se recogió de forma selectiva satisfactoriamente purificando el extracto de la *E. coli* que expresa la colagenasa G de fusión mediante la cromatografía por afinidad y retirando de ese modo la colagenasa en la que se degradó parte o todo el CBD.

[Ejemplo 3] Expresión de colagenasa de fusión en la que se une una marca de histidina a extremo de carboxilo de la colagenasa H derivada de *Clostridium histolyticum*

(3-1) Preparación de fragmento de gen de colagenasa H

Se amplificó un gen de la colagenasa H derivado de *Clostridium histolyticum* desde el extremo en la posición 5' al extremo en la posición 3' del mismo por PCR con un molde de ADN genómico derivado de *Clostridium histolyticum*. En este caso, los cebadores se diseñaron de modo que el extremo en la posición 3' del gen de la colagenasa H amplificada sirviera como una secuencia de reconocimiento de *Xba*I y adicionalmente se añadió una secuencia de reconocimiento de *Bam*HI después del codón de parada del gen. Como resultado, se introdujo una mutación en la secuencia de aminoácidos en una posición (un aminoácido en la posición 980 de la SEC ID N°: 5) en el lado del

extremo carboxilo de la secuencia de aminoácidos de la colagenasa H natural que se describe en la Bibliografía de No Patente 2. Por lo tanto, la secuencia de aminoácidos se modificó a la secuencia de aminoácidos de la SEC ID N°: 6.

5 Los cebadores usados en esta PCR fueron los que siguen a continuación.

colH-F:  
ATGAAAAGGAAATGTTTATC (SEC ID N°: 15)

10 colH-R:  
CCGGATCCTATCTAGATACTGAACCTT (SEC ID N°: 16)

El fragmento del gen de la colagenasa H amplificada se digirió con BamHI.

(3-2) Preparación de fragmento de ADN que incluye región que codifica marca de histidina

15 Se preparó un fragmento de ADN que incluye la región que codifica la marca de histidina con el mismo método que en el Ejemplo 1.

(3-3) Preparación del fragmento del promotor lacZ

20 Se preparó un fragmento del promotor lacZ con el mismo método que en el Ejemplo 1.

(3-4) Construcción de vector de expresión para colagenasa H de fusión a la que se une una marca de histidina

25 Se insertaron tres tipos de fragmentos de ADN preparados con el método que se ha descrito anteriormente en pBR322, un vector disponible en el mercado, con el fin de unirlos en el orden de promotor lacZ, el gen de la colagenasa H, y el fragmento de ADN que incluye la marca de histidina desde el extremo en la posición 5'. En primer lugar, para insertar el promotor lacZ y el gen de la colagenasa H en pBR322, la región del promotor lacZ y el fragmento del gen de la colagenasa H preparado tal como se ha descrito anteriormente se fosforilaron, y a continuación se insertaron en pBR322 que se había digerido doblemente con HindIII y BamHI para construir pColH (Fig. 5). Se debe indicar que la región del promotor del gen lacZ y el gen de la colagenasa H se unieron entre sí en sus extremos truncados. Posteriormente, el fragmento de ADN que incluía el ADN que codifica la marca de histidina preparada con el método que se ha descrito anteriormente se insertó en pColH que se había digerido doblemente con XbaI y BamHI para construir pColH-His (Fig. 6). El ADN que se había insertado por último en pBR322 era un ADN que tenía la secuencia de bases de las posiciones 1 a 3105 del Listado de Secuencias: 8.

(3-5) Preparación de *E. coli* que expresa la colagenasa H de fusión

La *E. coli* que expresa la colagenasa H de fusión se creó con el mismo método que en el Ejemplo 1.

40 [Ejemplo 4] Cultivo de *E. coli* que expresa la colagenasa H de fusión, y recogida selectiva de la colagenasa H que tiene CBD

(4-1) Cultivo de *E. coli* que expresa la colagenasa H de fusión

45 La *E. coli* que expresa la colagenasa H de fusión obtenida en el Ejemplo 3 se cultivó con el método que se ha descrito en el Ejemplo 2 para obtener una solución de cultivo.

(4-2) Preparación de extracto de *E. coli* que expresa la colagenasa H de fusión

50 Se obtuvo un extracto de la *E. coli* que expresa la colagenasa H de fusión a partir de la solución de cultivo resultante obtenida en (4-1) con el método que se ha descrito en el Ejemplo 2.

(4-3) Recogida selectiva de la colagenasa H de fusión que tiene CBD mediante cromatografía por afinidad

55 El extracto de la *E. coli* que expresa la colagenasa H de fusión obtenida en (4-2) se sometió a cromatografía por afinidad con el método que se ha descrito en el Ejemplo 2 para recoger de este modo la colagenasa H de fusión que tiene el CBD.

(4-4) Confirmación de retirada de colagenasa degradada

60 Para confirmar en qué medida se retiró la colagenasa degradada a partir de la colagenasa H de fusión, la actividad de tinción se realizó después de electroforesis con el método que se ha descrito en el Ejemplo 2. Como resultado, se observaron cuatro bandas que indicaban actividad de proteasa para el extracto de la *E. coli* que expresa la colagenasa H de fusión (Fig. 7). Por otro lado, para la solución de la colagenasa H de fusión sometida a la cromatografía por afinidad, se observó una banda que indicaba el peso molecular máximo como una banda principal

entre las cuatro bandas mencionadas anteriormente (Fig. 8). Basándose en la comparación y el análisis de estas bandas, se encontró que se degradaba parte o todo el CBD en la colagenasa H de fusión expresada por *E. coli*. Los resultados anteriores mostraron que la colagenasa H de fusión que tenía el CBD se recogió satisfactoriamente de forma selectiva mediante purificación del extracto de la *E. coli* que expresa la colagenasa H de fusión mediante la cromatografía por afinidad y de ese modo se retiró la colagenasa en la que se degrada parte o todo el CBD.

### Aplicabilidad Industrial

De acuerdo con la presente invención, se puede producir una colagenasa como una proteína recombinante con pureza elevada y sin incluir un producto degradado causado por una acción de un huésped. El uso de una colagenasa producida con el método de la presente invención permite, por ejemplo, que los islotes pancreáticos se separen de un páncreas sin disminuir la calidad de los islotes pancreáticos, contribuyendo de ese modo en gran medida al trasplante de islotes pancreáticos en pacientes diabéticos.

### 15 LISTADO DE SECUENCIAS

<110> MEIJI SEIKA KAISHA, LTD. TOHOKU UNIVERSITY

20 <120> Una colagenasa con una etiqueta de afinidad y su método de preparación

<130> M0833

<150> JP 2008-295922

25 <151> 19-11-2008

<160> 16

<170> PatentIn versión 3.1

30

<210> 1

<211> 1118

<212> PRT

<213> *Clostridium histolyticum*

35

<220>

<221> SEÑAL

<222> (-110)..(-1)

<223>

40

<220>

<221> mat\_peptide

<222> (1)..(1008)

<223>

45

<400> 1

Met Lys Lys Asn Ile Leu Lys Ile Leu Met Asp Ser Tyr Ser Lys  
-110                    -105                    -100

Glu Ser Lys Ile Gln Thr Val Arg Arg Val Thr Ser Val Ser Leu Leu  
-95                    -90                    -85                    -80

Ala Val Tyr Leu Thr Met Asn Thr Ser Ser Leu Val Leu Ala Lys Pro  
                  -75                    -70                    -65

Ile Glu Asn Thr Asn Asp Thr Ser Ile Lys Asn Val Glu Lys Leu Arg

-60                      -55                      -50

Asn Ala Pro Asn Glu Glu Asn Ser Lys Lys Val Glu Asp Ser Lys Asn  
 -45                      -40                      -35

Asp Lys Val Glu His Val Lys Asn Ile Glu Glu Ala Lys Val Glu Gln  
 -30                      -25                      -20

Val Ala Pro Glu Val Lys Ser Lys Ser Thr Leu Arg Ser Ala Ser Ile  
 -15                      -10                      -5                      -1 1

Ala Asn Thr Asn Ser Glu Lys Tyr Asp Phe Glu Tyr Leu Asn Gly Leu  
                     5                      10                      15

Ser Tyr Thr Glu Leu Thr Asn Leu Ile Lys Asn Ile Lys Trp Asn Gln  
                     20                      25                      30

Ile Asn Gly Leu Phe Asn Tyr Ser Thr Gly Ser Gln Lys Phe Phe Gly  
                     35                      40                      45

Asp Lys Asn Arg Val Gln Ala Ile Ile Asn Ala Leu Gln Glu Ser Gly  
 50                      55                      60                      65

Arg Thr Tyr Thr Ala Asn Asp Met Lys Gly Ile Glu Thr Phe Thr Glu  
                     70                      75                      80

Val Leu Arg Ala Gly Phe Tyr Leu Gly Tyr Tyr Asn Asp Gly Leu Ser  
                     85                      90                      95

Tyr Leu Asn Asp Arg Asn Phe Gln Asp Lys Cys Ile Pro Ala Met Ile  
                     100                      105                      110

Ala Ile Gln Lys Asn Pro Asn Phe Lys Leu Gly Thr Ala Val Gln Asp  
                     115                      120                      125

Glu Val Ile Thr Ser Leu Gly Lys Leu Ile Gly Asn Ala Ser Ala Asn  
 130                    135                    140                    145

Ala Glu Val Val Asn Asn Cys Val Pro Val Leu Lys Gln Phe Arg Glu  
                   150                    155                    160

Asn Leu Asn Gln Tyr Ala Pro Asp Tyr Val Lys Gly Thr Ala Val Asn  
                   165                    170                    175

Glu Leu Ile Lys Gly Ile Glu Phe Asp Phe Ser Gly Ala Ala Tyr Glu  
                   180                    185                    190

Lys Asp Val Lys Thr Met Pro Trp Tyr Gly Lys Ile Asp Pro Phe Ile  
                   195                    200                    205

Asn Glu Leu Lys Ala Leu Gly Leu Tyr Gly Asn Ile Thr Ser Ala Thr  
 210                    215                    220                    225

Glu Trp Ala Ser Asp Val Gly Ile Tyr Tyr Leu Ser Lys Phe Gly Leu  
                   230                    235                    240

Tyr Ser Thr Asn Arg Asn Asp Ile Val Gln Ser Leu Glu Lys Ala Val  
                   245                    250                    255

Asp Met Tyr Lys Tyr Gly Lys Ile Ala Phe Val Ala Met Glu Arg Ile  
                   260                    265                    270

Thr Trp Asp Tyr Asp Gly Ile Gly Ser Asn Gly Lys Lys Val Asp His  
                   275                    280                    285

Asp Lys Phe Leu Asp Asp Ala Glu Lys His Tyr Leu Pro Lys Thr Tyr  
 290                    295                    300                    305

Thr Phe Asp Asn Gly Thr Phe Ile Ile Arg Ala Gly Asp Lys Val Ser

310	315	320
Glu Glu Lys Ile Lys Arg Leu Tyr Trp Ala Ser Arg Glu Val Lys Ser		
325	330	335
Gln Phe His Arg Val Val Gly Asn Asp Lys Ala Leu Glu Val Gly Asn		
340	345	350
Ala Asp Asp Val Leu Thr Met Lys Ile Phe Asn Ser Pro Glu Glu Tyr		
355	360	365
Lys Phe Asn Thr Asn Ile Asn Gly Val Ser Thr Asp Asn Gly Gly Leu		
370	375	380
Tyr Ile Glu Pro Arg Gly Thr Phe Tyr Thr Tyr Glu Arg Thr Pro Gln		
390	395	400
Gln Ser Ile Phe Ser Leu Glu Glu Leu Phe Arg His Glu Tyr Thr His		
405	410	415
Tyr Leu Gln Ala Arg Tyr Leu Val Asp Gly Leu Trp Gly Gln Gly Pro		
420	425	430
Phe Tyr Glu Lys Asn Arg Leu Thr Trp Phe Asp Glu Gly Thr Ala Glu		
435	440	445
Phe Phe Ala Gly Ser Thr Arg Thr Ser Gly Val Leu Pro Arg Lys Ser		
450	455	460
Ile Leu Gly Tyr Leu Ala Lys Asp Lys Val Asp His Arg Tyr Ser Leu		
470	475	480
Lys Lys Thr Leu Asn Ser Gly Tyr Asp Asp Ser Asp Trp Met Phe Tyr		
485	490	495

Asn Tyr Gly Phe Ala Val Ala His Tyr Leu Tyr Glu Lys Asp Met Pro  
 500                      505                      510

Thr Phe Ile Lys Met Asn Lys Ala Ile Leu Asn Thr Asp Val Lys Ser  
 515                      520                      525

Tyr Asp Glu Ile Ile Lys Lys Leu Ser Asp Asp Ala Asn Lys Asn Thr  
 530                      535                      540                      545

Glu Tyr Gln Asn His Ile Gln Glu Leu Ala Asp Lys Tyr Gln Gly Ala  
 550                      555                      560

Gly Ile Pro Leu Val Ser Asp Asp Tyr Leu Lys Asp His Gly Tyr Lys  
 565                      570                      575

Lys Ala Ser Glu Val Tyr Ser Glu Ile Ser Lys Ala Ala Ser Leu Thr  
 580                      585                      590

Asn Thr Ser Val Thr Ala Glu Lys Ser Gln Tyr Phe Asn Thr Phe Thr  
 595                      600                      605

Leu Arg Gly Thr Tyr Thr Gly Glu Thr Ser Lys Gly Glu Phe Lys Asp  
 610                      615                      620                      625

Trp Asp Glu Met Ser Lys Lys Leu Asp Gly Thr Leu Glu Ser Leu Ala  
 630                      635                      640

Lys Asn Ser Trp Ser Gly Tyr Lys Thr Leu Thr Ala Tyr Phe Thr Asn  
 645                      650                      655

Tyr Arg Val Thr Ser Asp Asn Lys Val Gln Tyr Asp Val Val Phe His  
 660                      665                      670

Gly Val Leu Thr Asp Asn Ala Asp Ile Ser Asn Asn Lys Ala Pro Ile

675                      680                      685

Ala Lys Val Thr Gly Pro Ser Thr Gly Ala Val Gly Arg Asn Ile Glu  
690                      695                      700                      705

Phe Ser Gly Lys Asp Ser Lys Asp Glu Asp Gly Lys Ile Val Ser Tyr  
710                      715                      720

Asp Trp Asp Phe Gly Asp Gly Ala Thr Ser Arg Gly Lys Asn Ser Val  
725                      730                      735

His Ala Tyr Lys Lys Ala Gly Thr Tyr Asn Val Thr Leu Lys Val Thr  
740                      745                      750

Asp Asp Lys Gly Ala Thr Ala Thr Glu Ser Phe Thr Ile Glu Ile Lys  
755                      760                      765

Asn Glu Asp Thr Thr Thr Pro Ile Thr Lys Glu Met Glu Pro Asn Asp  
770                      775                      780                      785

Asp Ile Lys Glu Ala Asn Gly Pro Ile Val Glu Gly Val Thr Val Lys  
790                      795                      800

Gly Asp Leu Asn Gly Ser Asp Asp Ala Asp Thr Phe Tyr Phe Asp Val  
805                      810                      815

Lys Glu Asp Gly Asp Val Thr Ile Glu Leu Pro Tyr Ser Gly Ser Ser  
820                      825                      830

Asn Phe Thr Trp Leu Val Tyr Lys Glu Gly Asp Asp Gln Asn His Ile  
835                      840                      845

Ala Ser Gly Ile Asp Lys Asn Asn Ser Lys Val Gly Thr Phe Lys Ser  
850                      855                      860                      865

Thr Lys Gly Arg His Tyr Val Phe Ile Tyr Lys His Asp Ser Ala Ser  
 870 875 880

Asn Ile Ser Tyr Ser Leu Asn Ile Lys Gly Leu Gly Asn Glu Lys Leu  
 885 890 895

Lys Glu Lys Glu Asn Asn Asp Ser Ser Asp Lys Ala Thr Val Ile Pro  
 900 905 910

Asn Phe Asn Thr Thr Met Gln Gly Ser Leu Leu Gly Asp Asp Ser Arg  
 915 920 925

Asp Tyr Tyr Ser Phe Glu Val Lys Glu Glu Gly Glu Val Asn Ile Glu  
 930 935 940 945

Leu Asp Lys Lys Asp Glu Phe Gly Val Thr Trp Thr Leu His Pro Glu  
 950 955 960

Ser Asn Ile Asn Asp Arg Ile Thr Tyr Gly Gln Val Asp Gly Asn Lys  
 965 970 975

Val Ser Asn Lys Val Lys Leu Arg Pro Gly Lys Tyr Tyr Leu Leu Val  
 980 985 990

Tyr Lys Tyr Ser Gly Ser Gly Asn Tyr Glu Leu Arg Val Asn Lys  
 995 1000 1005

5 <210> 2  
 <211> 1118  
 <212> PRT  
 <213> *Clostridium histolyticum*

10 <220>  
 <221> SEÑAL  
 <222> (-110)..(-1)

15 <220>  
 <221> mat\_peptide  
 <222> (1)..(1008)

<400> 2

ES 2 531 402 T3

Met Lys Lys Asn Ile Leu Lys Ile Leu Met Asp Ser Tyr Ser Lys  
-110 -105 -100

Glu Ser Lys Ile Gln Thr Val Arg Arg Val Thr Ser Val Ser Leu Leu  
-95 -90 -85 -80

Ala Val Tyr Leu Thr Met Asn Thr Ser Ser Leu Val Leu Ala Lys Pro  
-75 -70 -65

Ile Glu Asn Thr Asn Asp Thr Ser Ile Lys Asn Val Glu Lys Leu Arg  
-60 -55 -50

Asn Ala Pro Asn Glu Glu Asn Ser Lys Lys Val Glu Asp Ser Lys Asn  
-45 -40 -35

Asp Lys Val Glu His Val Lys Asn Ile Glu Glu Ala Lys Val Glu Gln  
-30 -25 -20

Val Ala Pro Glu Val Lys Ser Lys Ser Thr Leu Arg Ser Ala Ser Ile  
-15 -10 -5 -1 1

Ala Asn Thr Asn Ser Glu Lys Tyr Asp Phe Glu Tyr Leu Asn Gly Leu  
5 10 15

Ser Tyr Thr Glu Leu Thr Asn Leu Ile Lys Asn Ile Lys Trp Asn Gln  
20 25 30

Ile Asn Gly Leu Phe Asn Tyr Ser Thr Gly Ser Gln Lys Phe Phe Gly  
35 40 45

Asp Lys Asn Arg Val Gln Ala Ile Ile Asn Ala Leu Gln Glu Ser Gly  
 50                    55                    60                    65

Arg Thr Tyr Thr Ala Asn Asp Met Lys Gly Ile Glu Thr Phe Thr Glu  
                   70                    75                    80

Val Leu Arg Ala Gly Phe Tyr Leu Gly Tyr Tyr Asn Asp Gly Leu Ser  
                   85                    90                    95

Tyr Leu Asn Asp Arg Asn Phe Gln Asp Lys Cys Ile Pro Ala Met Ile  
                   100                    105                    110

Ala Ile Gln Lys Asn Pro Asn Phe Lys Leu Gly Thr Ala Val Gln Asp  
                   115                    120                    125

Glu Val Ile Thr Ser Leu Gly Lys Leu Ile Gly Asn Ala Ser Ala Asn  
 130                    135                    140                    145

Ala Glu Val Val Asn Asn Cys Val Pro Val Leu Lys Gln Phe Arg Glu  
                   150                    155                    160

Asn Leu Asn Gln Tyr Ala Pro Asp Tyr Val Lys Gly Thr Ala Val Asn  
                   165                    170                    175

Glu Leu Ile Lys Gly Ile Glu Phe Asp Phe Ser Gly Ala Ala Tyr Glu  
                   180                    185                    190

Lys Asp Val Lys Thr Met Pro Trp Tyr Gly Lys Ile Asp Pro Phe Ile  
                   195                    200                    205

Asn Glu Leu Lys Ala Leu Gly Leu Tyr Gly Asn Ile Thr Ser Ala Thr  
 210                    215                    220                    225

Glu Trp Ala Ser Asp Val Gly Ile Tyr Tyr Leu Ser Lys Phe Gly Leu  
                   230                    235                    240

Tyr Ser Thr Asn Arg Asn Asp Ile Val Gln Ser Leu Glu Lys Ala Val  
245 250 255

Asp Met Tyr Lys Tyr Gly Lys Ile Ala Phe Val Ala Met Glu Arg Ile  
260 265 270

Thr Trp Asp Tyr Asp Gly Ile Gly Ser Asn Gly Lys Lys Val Asp His  
275 280 285

Asp Lys Phe Leu Asp Asp Ala Glu Lys His Tyr Leu Pro Lys Thr Tyr  
290 295 300 305

Thr Phe Asp Asn Gly Thr Phe Ile Ile Arg Ala Gly Asp Lys Val Ser  
310 315 320

Glu Glu Lys Ile Lys Arg Leu Tyr Trp Ala Ser Arg Glu Val Lys Ser  
325 330 335

Gln Phe His Arg Val Val Gly Asn Asp Lys Ala Leu Glu Val Gly Asn  
340 345 350

Ala Asp Asp Val Leu Thr Met Lys Ile Phe Asn Ser Pro Glu Glu Tyr  
355 360 365

Lys Phe Asn Thr Asn Ile Asn Gly Val Ser Thr Asp Asn Gly Gly Leu  
370 375 380 385

Tyr Ile Glu Pro Arg Gly Thr Phe Tyr Thr Tyr Glu Arg Thr Pro Gln  
390 395 400

Gln Ser Ile Phe Ser Leu Glu Glu Leu Phe Arg His Glu Tyr Thr His  
405 410 415

Tyr Leu Gln Ala Arg Tyr Leu Val Asp Gly Leu Trp Gly Gln Gly Pro  
 420                      425                      430

Phe Tyr Glu Lys Asn Arg Leu Thr Trp Phe Asp Glu Gly Thr Ala Glu  
 435                      440                      445

Phe Phe Ala Gly Ser Thr Arg Thr Ser Gly Val Leu Pro Arg Lys Ser  
 450                      455                      460                      465

Ile Leu Gly Tyr Leu Ala Lys Asp Lys Val Asp His Arg Tyr Ser Leu  
 470                      475                      480

Lys Lys Thr Leu Asn Ser Gly Tyr Asp Asp Ser Asp Trp Met Phe Tyr  
 485                      490                      495

Asn Tyr Gly Phe Ala Val Ala His Tyr Leu Tyr Glu Lys Asp Met Pro  
 500                      505                      510

Thr Phe Ile Lys Met Asn Lys Ala Ile Leu Asn Thr Asp Val Lys Ser  
 515                      520                      525

Tyr Asp Glu Ile Ile Lys Lys Leu Ser Asp Asp Ala Asn Lys Asn Thr  
 530                      535                      540                      545

Glu Tyr Gln Asn His Ile Gln Glu Leu Ala Asp Lys Tyr Gln Gly Ala  
 550                      555                      560

Gly Ile Pro Leu Val Ser Asp Asp Tyr Leu Lys Asp His Gly Tyr Lys  
 565                      570                      575

Lys Ala Ser Glu Val Tyr Ser Glu Ile Ser Lys Ala Ala Ser Leu Thr  
 580                      585                      590

Asn Thr Ser Val Thr Ala Glu Lys Ser Gln Tyr Phe Asn Thr Phe Thr  
 595                      600                      605

Leu Arg Gly Thr Tyr Thr Gly Glu Thr Ser Lys Gly Glu Phe Lys Asp  
610                    615                    620                    625

Trp Asp Glu Met Ser Lys Lys Leu Asp Gly Thr Leu Glu Ser Leu Ala  
                  630                    635                    640

Lys Asn Ser Trp Ser Gly Tyr Lys Thr Leu Thr Ala Tyr Phe Thr Asn  
                  645                    650                    655

Tyr Arg Val Thr Ser Asp Asn Lys Val Gln Tyr Asp Val Val Phe His  
                  660                    665                    670

Gly Val Leu Thr Asp Asn Ala Asp Ile Ser Asn Asn Lys Ala Pro Ile  
                  675                    680                    685

Ala Lys Val Thr Gly Pro Ser Thr Gly Ala Val Gly Arg Asn Ile Glu  
690                    695                    700                    705

Phe Ser Gly Lys Asp Ser Lys Asp Glu Asp Gly Lys Ile Val Ser Tyr  
                  710                    715                    720

Asp Trp Asp Phe Gly Asp Gly Ala Thr Ser Arg Gly Lys Asn Ser Val  
                  725                    730                    735

His Ala Tyr Lys Lys Ala Gly Thr Tyr Asn Val Thr Leu Lys Val Thr  
                  740                    745                    750

Asp Asp Lys Gly Ala Thr Ala Thr Glu Ser Phe Thr Ile Glu Ile Lys  
                  755                    760                    765

Asn Glu Asp Thr Thr Thr Pro Ile Thr Lys Glu Met Glu Pro Asn Asp  
770                    775                    780                    785

Asp Ile Lys Glu Ala Asn Gly Pro Ile Val Glu Gly Val Thr Val Lys  
790 795 800

Gly Asp Leu Asn Gly Ser Asp Asp Ala Asp Thr Phe Tyr Phe Asp Val  
805 810 815

Lys Glu Asp Gly Asp Val Thr Ile Glu Leu Pro Tyr Ser Gly Ser Ser  
820 825 830

Asn Phe Thr Trp Leu Val Tyr Lys Glu Gly Asp Asp Gln Asn His Ile  
835 840 845

Ala Ser Gly Ile Asp Lys Asn Asn Ser Lys Val Gly Thr Phe Lys Ser  
850 855 860 865

Thr Lys Gly Arg His Tyr Val Phe Ile Tyr Lys His Asp Ser Ala Ser  
870 875 880

Asn Ile Ser Tyr Ser Leu Asn Ile Lys Gly Leu Gly Asn Glu Lys Leu  
885 890 895

Lys Glu Lys Glu Asn Asn Asp Ser Ser Asp Lys Ala Thr Val Ile Pro  
900 905 910

Asn Phe Asn Thr Thr Met Gln Gly Ser Leu Leu Gly Asp Asp Ser Arg  
915 920 925

Asp Tyr Tyr Ser Phe Glu Val Lys Glu Glu Gly Glu Val Asn Ile Glu  
930 935 940 945

Leu Asp Lys Lys Asp Glu Phe Gly Val Thr Trp Thr Leu His Pro Glu  
950 955 960

Ser Asn Ile Asn Asp Arg Ile Thr Tyr Gly Gln Val Asp Gly Asn Lys  
965 970 975

Val Ser Asn Lys Val Lys Leu Arg Pro Gly Lys Tyr Tyr Leu Leu Val  
980 985 990

Tyr Lys Tyr Ser Gly Ser Gly Asn Tyr Glu Leu Arg Val Ser Arg  
995 1000 1005

5 <210> 3  
<211> 1131  
<212> PRT  
<213> *Clostridium histolyticum*

10 <220>  
<221> SEÑAL  
<222> (-110)..(-1)

15 <220>  
<221> mat\_peptide  
<222> (1)..(1021)  
<223>

<400> 3

Met Lys Lys Asn Ile Leu Lys Ile Leu Met Asp Ser Tyr Ser Lys  
-110 -105 -100

Glu Ser Lys Ile Gln Thr Val Arg Arg Val Thr Ser Val Ser Leu Leu  
-95 -90 -85 -80

Ala Val Tyr Leu Thr Met Asn Thr Ser Ser Leu Val Leu Ala Lys Pro  
-75 -70 -65

Ile Glu Asn Thr Asn Asp Thr Ser Ile Lys Asn Val Glu Lys Leu Arg  
-60 -55 -50

Asn Ala Pro Asn Glu Glu Asn Ser Lys Lys Val Glu Asp Ser Lys Asn  
-45 -40 -35

Asp Lys Val Glu His Val Lys Asn Ile Glu Glu Ala Lys Val Glu Gln  
 -30                      -25                      -20

Val Ala Pro Glu Val Lys Ser Lys Ser Thr Leu Arg Ser Ala Ser Ile  
 -15                      -10                      -5                      -1 1

Ala Asn Thr Asn Ser Glu Lys Tyr Asp Phe Glu Tyr Leu Asn Gly Leu  
                     5                      10                      15

Ser Tyr Thr Glu Leu Thr Asn Leu Ile Lys Asn Ile Lys Trp Asn Gln  
                     20                      25                      30

Ile Asn Gly Leu Phe Asn Tyr Ser Thr Gly Ser Gln Lys Phe Phe Gly  
                     35                      40                      45

Asp Lys Asn Arg Val Gln Ala Ile Ile Asn Ala Leu Gln Glu Ser Gly  
 50                      55                      60                      65

Arg Thr Tyr Thr Ala Asn Asp Met Lys Gly Ile Glu Thr Phe Thr Glu  
                     70                      75                      80

Val Leu Arg Ala Gly Phe Tyr Leu Gly Tyr Tyr Asn Asp Gly Leu Ser  
                     85                      90                      95

Tyr Leu Asn Asp Arg Asn Phe Gln Asp Lys Cys Ile Pro Ala Met Ile  
                     100                      105                      110

Ala Ile Gln Lys Asn Pro Asn Phe Lys Leu Gly Thr Ala Val Gln Asp  
                     115                      120                      125

Glu Val Ile Thr Ser Leu Gly Lys Leu Ile Gly Asn Ala Ser Ala Asn  
 130                      135                      140                      145

Ala Glu Val Val Asn Asn Cys Val Pro Val Leu Lys Gln Phe Arg Glu  
                     150                      155                      160

Asn Leu Asn Gln Tyr Ala Pro Asp Tyr Val Lys Gly Thr Ala Val Asn  
 165 170 175

Glu Leu Ile Lys Gly Ile Glu Phe Asp Phe Ser Gly Ala Ala Tyr Glu  
 180 185 190

Lys Asp Val Lys Thr Met Pro Trp Tyr Gly Lys Ile Asp Pro Phe Ile  
 195 200 205

Asn Glu Leu Lys Ala Leu Gly Leu Tyr Gly Asn Ile Thr Ser Ala Thr  
 210 215 220 225

Glu Trp Ala Ser Asp Val Gly Ile Tyr Tyr Leu Ser Lys Phe Gly Leu  
 230 235 240

Tyr Ser Thr Asn Arg Asn Asp Ile Val Gln Ser Leu Glu Lys Ala Val  
 245 250 255

Asp Met Tyr Lys Tyr Gly Lys Ile Ala Phe Val Ala Met Glu Arg Ile  
 260 265 270

Thr Trp Asp Tyr Asp Gly Ile Gly Ser Asn Gly Lys Lys Val Asp His  
 275 280 285

Asp Lys Phe Leu Asp Asp Ala Glu Lys His Tyr Leu Pro Lys Thr Tyr  
 290 295 300 305

Thr Phe Asp Asn Gly Thr Phe Ile Ile Arg Ala Gly Asp Lys Val Ser  
 310 315 320

Glu Glu Lys Ile Lys Arg Leu Tyr Trp Ala Ser Arg Glu Val Lys Ser  
 325 330 335

Gln Phe His Arg Val Val Gly Asn Asp Lys Ala Leu Glu Val Gly Asn  
 340 345 350

Ala Asp Asp Val Leu Thr Met Lys Ile Phe Asn Ser Pro Glu Glu Tyr  
 355 360 365

Lys Phe Asn Thr Asn Ile Asn Gly Val Ser Thr Asp Asn Gly Gly Leu  
 370 375 380 385

Tyr Ile Glu Pro Arg Gly Thr Phe Tyr Thr Tyr Glu Arg Thr Pro Gln  
 390 395 400

Gln Ser Ile Phe Ser Leu Glu Glu Leu Phe Arg His Glu Tyr Thr His  
 405 410 415

Tyr Leu Gln Ala Arg Tyr Leu Val Asp Gly Leu Trp Gly Gln Gly Pro  
 420 425 430

Phe Tyr Glu Lys Asn Arg Leu Thr Trp Phe Asp Glu Gly Thr Ala Glu  
 435 440 445

Phe Phe Ala Gly Ser Thr Arg Thr Ser Gly Val Leu Pro Arg Lys Ser  
 450 455 460 465

Ile Leu Gly Tyr Leu Ala Lys Asp Lys Val Asp His Arg Tyr Ser Leu  
 470 475 480

Lys Lys Thr Leu Asn Ser Gly Tyr Asp Asp Ser Asp Trp Met Phe Tyr  
 485 490 495

Asn Tyr Gly Phe Ala Val Ala His Tyr Leu Tyr Glu Lys Asp Met Pro  
 500 505 510

Thr Phe Ile Lys Met Asn Lys Ala Ile Leu Asn Thr Asp Val Lys Ser  
 515 520 525

Tyr Asp Glu Ile Ile Lys Lys Leu Ser Asp Asp Ala Asn Lys Asn Thr  
 530                    535                    540                    545

Glu Tyr Gln Asn His Ile Gln Glu Leu Ala Asp Lys Tyr Gln Gly Ala  
                   550                    555                    560

Gly Ile Pro Leu Val Ser Asp Asp Tyr Leu Lys Asp His Gly Tyr Lys  
                   565                    570                    575

Lys Ala Ser Glu Val Tyr Ser Glu Ile Ser Lys Ala Ala Ser Leu Thr  
                   580                    585                    590

Asn Thr Ser Val Thr Ala Glu Lys Ser Gln Tyr Phe Asn Thr Phe Thr  
                   595                    600                    605

Leu Arg Gly Thr Tyr Thr Gly Glu Thr Ser Lys Gly Glu Phe Lys Asp  
 610                    615                    620                    625

Trp Asp Glu Met Ser Lys Lys Leu Asp Gly Thr Leu Glu Ser Leu Ala  
                   630                    635                    640

Lys Asn Ser Trp Ser Gly Tyr Lys Thr Leu Thr Ala Tyr Phe Thr Asn  
                   645                    650                    655

Tyr Arg Val Thr Ser Asp Asn Lys Val Gln Tyr Asp Val Val Phe His  
                   660                    665                    670

Gly Val Leu Thr Asp Asn Ala Asp Ile Ser Asn Asn Lys Ala Pro Ile  
                   675                    680                    685

Ala Lys Val Thr Gly Pro Ser Thr Gly Ala Val Gly Arg Asn Ile Glu  
 690                    695                    700                    705

Phe Ser Gly Lys Asp Ser Lys Asp Glu Asp Gly Lys Ile Val Ser Tyr  
710 715 720

Asp Trp Asp Phe Gly Asp Gly Ala Thr Ser Arg Gly Lys Asn Ser Val  
725 730 735

His Ala Tyr Lys Lys Ala Gly Thr Tyr Asn Val Thr Leu Lys Val Thr  
740 745 750

Asp Asp Lys Gly Ala Thr Ala Thr Glu Ser Phe Thr Ile Glu Ile Lys  
755 760 765

Asn Glu Asp Thr Thr Thr Pro Ile Thr Lys Glu Met Glu Pro Asn Asp  
770 775 780 785

Asp Ile Lys Glu Ala Asn Gly Pro Ile Val Glu Gly Val Thr Val Lys  
790 795 800

Gly Asp Leu Asn Gly Ser Asp Asp Ala Asp Thr Phe Tyr Phe Asp Val  
805 810 815

Lys Glu Asp Gly Asp Val Thr Ile Glu Leu Pro Tyr Ser Gly Ser Ser  
820 825 830

Asn Phe Thr Trp Leu Val Tyr Lys Glu Gly Asp Asp Gln Asn His Ile  
835 840 845

Ala Ser Gly Ile Asp Lys Asn Asn Ser Lys Val Gly Thr Phe Lys Ser  
850 855 860 865

Thr Lys Gly Arg His Tyr Val Phe Ile Tyr Lys His Asp Ser Ala Ser  
870 875 880

Asn Ile Ser Tyr Ser Leu Asn Ile Lys Gly Leu Gly Asn Glu Lys Leu  
885 890 895

Lys Glu Lys Glu Asn Asn Asp Ser Ser Asp Lys Ala Thr Val Ile Pro  
 900 905 910

Asn Phe Asn Thr Thr Met Gln Gly Ser Leu Leu Gly Asp Asp Ser Arg  
 915 920 925

Asp Tyr Tyr Ser Phe Glu Val Lys Glu Glu Gly Glu Val Asn Ile Glu  
 930 935 940 945

Leu Asp Lys Lys Asp Glu Phe Gly Val Thr Trp Thr Leu His Pro Glu  
 950 955 960

Ser Asn Ile Asn Asp Arg Ile Thr Tyr Gly Gln Val Asp Gly Asn Lys  
 965 970 975

Val Ser Asn Lys Val Lys Leu Arg Pro Gly Lys Tyr Tyr Leu Leu Val  
 980 985 990

Tyr Lys Tyr Ser Gly Ser Gly Asn Tyr Glu Leu Arg Val Ser Arg Lys  
 995 1000 1005

Leu Ala Ala Ala Leu Glu His His His His His His  
 1010 1015 1020

5 <210> 4  
 <211> 3396  
 <212> ADN  
 <213> *Clostridium histolyticum*

10 <220>  
 <221> CDS  
 <222> (1)..(3396)

<400> 4

atg aaa aaa aat att tta aag att ctt atg gat agt tat tct aaa gaa 48  
 Met Lys Lys Asn Ile Leu Lys Ile Leu Met Asp Ser Tyr Ser Lys Glu

1	5	10	15	
tct aaa att caa act gta cgt agg gtt acg agt gta tca ctt tta gcg	96			
Ser Lys Ile Gln Thr Val Arg Arg Val Thr Ser Val Ser Leu Leu Ala				
20	25	30		
gta tat ctt act atg aat act tca agt tta gtt tta gca aaa cca ata	144			
Val Tyr Leu Thr Met Asn Thr Ser Ser Leu Val Leu Ala Lys Pro Ile				
35	40	45		
gaa aat act aat gat act agt ata aaa aat gtg gag aaa tta aga aat	192			
Glu Asn Thr Asn Asp Thr Ser Ile Lys Asn Val Glu Lys Leu Arg Asn				
50	55	60		
gct cca aat gaa gag aat agt aaa aag gta gaa gat agt aaa aat gat	240			
Ala Pro Asn Glu Glu Asn Ser Lys Lys Val Glu Asp Ser Lys Asn Asp				
65	70	75	80	
aag gta gaa cat gtg aaa aat ata gaa gag gca aag gtt gag caa gtt	288			
Lys Val Glu His Val Lys Asn Ile Glu Glu Ala Lys Val Glu Gln Val				
85	90	95		
gca ccc gaa gta aaa tct aaa tca act tta aga agt gct tct ata gcg	336			
Ala Pro Glu Val Lys Ser Lys Ser Thr Leu Arg Ser Ala Ser Ile Ala				
100	105	110		
aat act aat tct gag aaa tat gat ttt gag tat tta aat ggt ttg agc	384			
Asn Thr Asn Ser Glu Lys Tyr Asp Phe Glu Tyr Leu Asn Gly Leu Ser				
115	120	125		
tat act gaa ctt aca aat tta att aaa aat ata aag tgg aat caa att	432			
Tyr Thr Glu Leu Thr Asn Leu Ile Lys Asn Ile Lys Trp Asn Gln Ile				
130	135	140		
aat ggt tta ttt aat tat agt aca ggt tct caa aag ttc ttt gga gat	480			
Asn Gly Leu Phe Asn Tyr Ser Thr Gly Ser Gln Lys Phe Phe Gly Asp				
145	150	155	160	
aaa aat cgt gta caa gct ata att aat gct tta caa gaa agt gga aga	528			
Lys Asn Arg Val Gln Ala Ile Ile Asn Ala Leu Gln Glu Ser Gly Arg				
165	170	175		
act tac act gca aat gat atg aag ggt ata gaa act ttc act gag gtt	576			
Thr Tyr Thr Ala Asn Asp Met Lys Gly Ile Glu Thr Phe Thr Glu Val				
180	185	190		

tta aga gct ggt ttt tat tta ggg tac tat aat gat ggt tta tct tat 624  
 Leu Arg Ala Gly Phe Tyr Leu Gly Tyr Tyr Asn Asp Gly Leu Ser Tyr  
 195 200 205

tta aat gat aga aac ttc caa gat aaa tgt ata cct gca atg att gca 672  
 Leu Asn Asp Arg Asn Phe Gln Asp Lys Cys Ile Pro Ala Met Ile Ala  
 210 215 220

att caa aaa aat cct aac ttt aag cta gga act gca gtt caa gat gaa 720  
 Ile Gln Lys Asn Pro Asn Phe Lys Leu Gly Thr Ala Val Gln Asp Glu  
 225 230 235 240

gtt ata act tct tta gga aaa cta ata gga aat gct tct gct aat gct 768  
 Val Ile Thr Ser Leu Gly Lys Leu Ile Gly Asn Ala Ser Ala Asn Ala  
 245 250 255

gaa gta gtt aat aat tgt gta cca gtt cta aaa caa ttt aga gaa aac 816  
 Glu Val Val Asn Asn Cys Val Pro Val Leu Lys Gln Phe Arg Glu Asn  
 260 265 270

tta aat caa tat gct cct gat tac gtt aaa gga aca gct gta aat gaa 864  
 Leu Asn Gln Tyr Ala Pro Asp Tyr Val Lys Gly Thr Ala Val Asn Glu  
 275 280 285

tta att aaa ggt att gaa ttc gat ttt tct ggt gct gca tat gaa aaa 912  
 Leu Ile Lys Gly Ile Glu Phe Asp Phe Ser Gly Ala Ala Tyr Glu Lys  
 290 295 300

gat gtt aag aca atg cct tgg tat gga aaa att gat cca ttt ata aat 960  
 Asp Val Lys Thr Met Pro Trp Tyr Gly Lys Ile Asp Pro Phe Ile Asn  
 305 310 315 320

gaa ctt aag gcc tta ggt cta tat gga aat ata aca agt gca act gag 1008  
 Glu Leu Lys Ala Leu Gly Leu Tyr Gly Asn Ile Thr Ser Ala Thr Glu  
 325 330 335

tgg gca tct gat gtt gga ata tac tat tta agt aaa ttc ggt ctt tac 1056  
 Trp Ala Ser Asp Val Gly Ile Tyr Tyr Leu Ser Lys Phe Gly Leu Tyr  
 340 345 350

tca act aac cga aat gac ata gta cag tca ctt gaa aag gct gta gat 1104  
 Ser Thr Asn Arg Asn Asp Ile Val Gln Ser Leu Glu Lys Ala Val Asp  
 355 360 365

atg tat aag tat ggt aaa ata gcc ttt gta gca atg gag aga ata act 1152  
 Met Tyr Lys Tyr Gly Lys Ile Ala Phe Val Ala Met Glu Arg Ile Thr

370                    375                    380  
 tgg gat tat gat ggg att ggt tct aat ggt aaa aag gtg gat cac gat    1200  
 Trp Asp Tyr Asp Gly Ile Gly Ser Asn Gly Lys Lys Val Asp His Asp  
 385                    390                    395                    400  
  
 aag ttc tta gat gat gct gaa aaa cat tat ctg cca aag aca tat act    1248  
 Lys Phe Leu Asp Asp Ala Glu Lys His Tyr Leu Pro Lys Thr Tyr Thr  
                   405                    410                    415  
  
 ttt gat aat gga acc ttt att ata aga gca ggg gat aag gta tcc gaa    1296  
 Phe Asp Asn Gly Thr Phe Ile Ile Arg Ala Gly Asp Lys Val Ser Glu  
                   420                    425                    430  
  
 gaa aaa ata aaa agg cta tat tgg gca tca aga gaa gtg aag tct caa    1344  
 Glu Lys Ile Lys Arg Leu Tyr Trp Ala Ser Arg Glu Val Lys Ser Gln  
                   435                    440                    445  
  
 ttc cat aga gta gtt ggc aat gat aaa gct tta gag gtg gga aat gcc    1392  
 Phe His Arg Val Val Gly Asn Asp Lys Ala Leu Glu Val Gly Asn Ala  
                   450                    455                    460  
  
 gat gat gtt tta act atg aaa ata ttt aat agc cca gaa gaa tat aaa    1440  
 Asp Asp Val Leu Thr Met Lys Ile Phe Asn Ser Pro Glu Glu Tyr Lys  
 465                    470                    475                    480  
  
 ttt aat acc aat ata aat ggt gta agc act gat aat ggt ggt cta tat    1488  
 Phe Asn Thr Asn Ile Asn Gly Val Ser Thr Asp Asn Gly Gly Leu Tyr  
                   485                    490                    495  
  
 ata gaa cca aga ggg act ttc tac act tat gag aga aca cct caa caa    1536  
 Ile Glu Pro Arg Gly Thr Phe Tyr Thr Tyr Glu Arg Thr Pro Gln Gln  
                   500                    505                    510  
  
 agt ata ttt agt ctt gaa gaa ttg ttt aga cat gaa tat act cac tat    1584  
 Ser Ile Phe Ser Leu Glu Glu Leu Phe Arg His Glu Tyr Thr His Tyr  
                   515                    520                    525  
  
 tta caa gcg aga tat ctt gta gat ggt tta tgg ggg caa ggt cca ttt    1632  
 Leu Gln Ala Arg Tyr Leu Val Asp Gly Leu Trp Gly Gln Gly Pro Phe  
                   530                    535                    540  
  
 tat gaa aaa aat aga tta act tgg ttt gat gaa ggt aca gct gaa ttc    1680  
 Tyr Glu Lys Asn Arg Leu Thr Trp Phe Asp Glu Gly Thr Ala Glu Phe  
 545                    550                    555                    560

ttt gca gga tct acc cgt aca tct ggt gtt tta cca aga aaa tca ata 1728  
 Phe Ala Gly Ser Thr Arg Thr Ser Gly Val Leu Pro Arg Lys Ser Ile  
 565 570 575

tta gga tat ttg gct aag gat aaa gta gat cat aga tac tca tta aag 1776  
 Leu Gly Tyr Leu Ala Lys Asp Lys Val Asp His Arg Tyr Ser Leu Lys  
 580 585 590

aag act ctt aat tca ggg tat gat gac agt gat tgg atg ttc tat aat 1824  
 Lys Thr Leu Asn Ser Gly Tyr Asp Asp Ser Asp Trp Met Phe Tyr Asn  
 595 600 605

tat gga ttt gca gtt gca cat tac cta tat gaa aaa gat atg cct aca 1872  
 Tyr Gly Phe Ala Val Ala His Tyr Leu Tyr Glu Lys Asp Met Pro Thr  
 610 615 620

ttt att aag atg aat aaa gct ata ttg aat aca gat gtg aaa tct tat 1920  
 Phe Ile Lys Met Asn Lys Ala Ile Leu Asn Thr Asp Val Lys Ser Tyr  
 625 630 635 640

gat gaa ata ata aaa aaa tta agt gat gat gca aat aaa aat aca gaa 1968  
 Asp Glu Ile Ile Lys Lys Leu Ser Asp Asp Ala Asn Lys Asn Thr Glu  
 645 650 655

tat caa aac cat att caa gag tta gca gat aaa tat caa gga gca ggc 2016  
 Tyr Gln Asn His Ile Gln Glu Leu Ala Asp Lys Tyr Gln Gly Ala Gly  
 660 665 670

ata cct cta gta tca gat gat tac tta aaa gat cat gga tat aag aaa 2064  
 Ile Pro Leu Val Ser Asp Asp Tyr Leu Lys Asp His Gly Tyr Lys Lys  
 675 680 685

gca tct gaa gta tat tct gaa att tca aaa gct gct tct ctt aca aac 2112  
 Ala Ser Glu Val Tyr Ser Glu Ile Ser Lys Ala Ala Ser Leu Thr Asn  
 690 695 700

act agt gta aca gca gaa aaa tct caa tat ttt aac aca ttc act tta 2160  
 Thr Ser Val Thr Ala Glu Lys Ser Gln Tyr Phe Asn Thr Phe Thr Leu  
 705 710 715 720

aga gga act tat aca ggt gaa act tct aaa ggt gaa ttt aaa gat tgg 2208  
 Arg Gly Thr Tyr Thr Gly Glu Thr Ser Lys Gly Glu Phe Lys Asp Trp  
 725 730 735

gat gaa atg agt aaa aaa tta gat gga act ttg gag tcc ctt gct aaa 2256  
 Asp Glu Met Ser Lys Lys Leu Asp Gly Thr Leu Glu Ser Leu Ala Lys

740	745	750	
aat tct tgg agt gga tac aaa act tta aca gca tac ttt acg aat tat 2304			
Asn Ser Trp Ser Gly Tyr Lys Thr Leu Thr Ala Tyr Phe Thr Asn Tyr			
755	760	765	
aga gtt aca agc gat aat aaa gtt caa tat gat gta gtt ttc cat ggg 2352			
Arg Val Thr Ser Asp Asn Lys Val Gln Tyr Asp Val Val Phe His Gly			
770	775	780	
gtt tta aca gat aat gcg gat att agt aac aat aag gct cca ata gca 2400			
Val Leu Thr Asp Asn Ala Asp Ile Ser Asn Asn Lys Ala Pro Ile Ala			
785	790	795	800
aag gta act gga cca agc act ggt gct gta gga aga aat att gaa ttt 2448			
Lys Val Thr Gly Pro Ser Thr Gly Ala Val Gly Arg Asn Ile Glu Phe			
805	810	815	
agt gga aaa gat agt aaa gat gaa gat ggt aaa ata gta tca tat gat 2496			
Ser Gly Lys Asp Ser Lys Asp Glu Asp Gly Lys Ile Val Ser Tyr Asp			
820	825	830	
tgg gat ttt ggc gat ggt gca act agt aga ggc aaa aat tca gta cat 2544			
Trp Asp Phe Gly Asp Gly Ala Thr Ser Arg Gly Lys Asn Ser Val His			
835	840	845	
gct tac aaa aaa gca gga aca tat aat gtt aca tta aaa gta act gac 2592			
Ala Tyr Lys Lys Ala Gly Thr Tyr Asn Val Thr Leu Lys Val Thr Asp			
850	855	860	
gat aag ggt gca aca gct aca gaa agc ttt act ata gaa ata aag aac 2640			
Asp Lys Gly Ala Thr Ala Thr Glu Ser Phe Thr Ile Glu Ile Lys Asn			
865	870	875	880
gaa gat aca aca aca cct ata act aaa gaa atg gaa cct aat gat gat 2688			
Glu Asp Thr Thr Thr Pro Ile Thr Lys Glu Met Glu Pro Asn Asp Asp			
885	890	895	
ata aaa gag gct aat ggt cca ata gtt gaa ggt gtt act gta aaa ggt 2736			
Ile Lys Glu Ala Asn Gly Pro Ile Val Glu Gly Val Thr Val Lys Gly			
900	905	910	
gat tta aat ggt tct gat gat gct gat acc ttc tat ttt gat gta aaa 2784			
Asp Leu Asn Gly Ser Asp Asp Ala Asp Thr Phe Tyr Phe Asp Val Lys			
915	920	925	

gaa gat ggt gat gtt aca att gaa ctt cct tat tca ggg tca tct aat 2832  
 Glu Asp Gly Asp Val Thr Ile Glu Leu Pro Tyr Ser Gly Ser Ser Asn  
 930 935 940

ttc aca tgg tta gtt tat aaa gag gga gac gat caa aac cat att gca 2880  
 Phe Thr Trp Leu Val Tyr Lys Glu Gly Asp Asp Gln Asn His Ile Ala  
 945 950 955 960

agt ggt ata gat aag aat aac tca aaa gtt gga aca ttt aaa tct aca 2928  
 Ser Gly Ile Asp Lys Asn Asn Ser Lys Val Gly Thr Phe Lys Ser Thr  
 965 970 975

aaa gga aga cat tat gtg ttt ata tat aaa cac gat tct gct tca aat 2976  
 Lys Gly Arg His Tyr Val Phe Ile Tyr Lys His Asp Ser Ala Ser Asn  
 980 985 990

ata tcc tat tct tta aac ata aaa gga tta ggt aac gag aaa ttg aag 3024  
 Ile Ser Tyr Ser Leu Asn Ile Lys Gly Leu Gly Asn Glu Lys Leu Lys  
 995 1000 1005

gaa aaa gaa aat aat gat tct tct gat aaa gct aca gtt ata cca 3069  
 Glu Lys Glu Asn Asn Asp Ser Ser Asp Lys Ala Thr Val Ile Pro  
 1010 1015 1020

aat ttc aat acc act atg caa ggt tca ctt tta ggt gat gat tca 3114  
 Asn Phe Asn Thr Thr Met Gln Gly Ser Leu Leu Gly Asp Asp Ser  
 1025 1030 1035

aga gat tat tat tct ttt gag gtt aag gaa gaa ggc gaa gtt aat 3159  
 Arg Asp Tyr Tyr Ser Phe Glu Val Lys Glu Glu Gly Glu Val Asn  
 1040 1045 1050

ata gaa cta gat aaa aag gat gaa ttt ggt gta aca tgg aca cta 3204  
 Ile Glu Leu Asp Lys Lys Asp Glu Phe Gly Val Thr Trp Thr Leu  
 1055 1060 1065

cat cca gag tca aat att aat gac aga ata act tac gga caa gtt 3249  
 His Pro Glu Ser Asn Ile Asn Asp Arg Ile Thr Tyr Gly Gln Val  
 1070 1075 1080

gat ggt aat aag gta tct aat aaa gtt aaa tta aga cca gga aaa 3294  
 Asp Gly Asn Lys Val Ser Asn Lys Val Lys Leu Arg Pro Gly Lys  
 1085 1090 1095

tat tat cta ctt gtt tat aaa tac tca gga tca gga aac tat gag 3339  
 Tyr Tyr Leu Leu Val Tyr Lys Tyr Ser Gly Ser Gly Asn Tyr Glu



Ser Asp Ala Lys Glu Phe Tyr Gly Asn Lys Thr Arg Met Ser Phe Ile  
 45 50 55

Met Asp Glu Ile Gly Arg Arg Ala Pro Gln Tyr Thr Glu Ile Asp His  
 60 65 70

Lys Gly Ile Pro Thr Leu Val Glu Val Val Arg Ala Gly Phe Tyr Leu  
 75 80 85

Gly Phe His Asn Lys Glu Leu Asn Glu Ile Asn Lys Arg Ser Phe Lys  
 90 95 100

Glu Arg Val Ile Pro Ser Ile Leu Ala Ile Gln Lys Asn Pro Asn Phe  
 105 110 115 120

Lys Leu Gly Thr Glu Val Gln Asp Lys Ile Val Ser Ala Thr Gly Leu  
 125 130 135

Leu Ala Gly Asn Glu Thr Ala Pro Pro Glu Val Val Asn Asn Phe Thr  
 140 145 150

Pro Ile Leu Gln Asp Cys Ile Lys Asn Ile Asp Arg Tyr Ala Leu Asp  
 155 160 165

Asp Leu Lys Ser Lys Ala Leu Phe Asn Val Leu Ala Ala Pro Thr Tyr  
 170 175 180

Asp Ile Thr Glu Tyr Leu Arg Ala Thr Lys Glu Lys Pro Glu Asn Thr  
 185 190 195 200

Pro Trp Tyr Gly Lys Ile Asp Gly Phe Ile Asn Glu Leu Lys Lys Leu  
 205 210 215

Ala Leu Tyr Gly Lys Ile Asn Asp Asn Asn Ser Trp Ile Ile Asp Asn  
 220 225 230

Gly Ile Tyr His Ile Ala Pro Leu Gly Lys Leu His Ser Asn Asn Lys  
235 240 245

Ile Gly Ile Glu Thr Leu Thr Glu Val Met Lys Val Tyr Pro Tyr Leu  
250 255 260

Ser Met Gln His Leu Gln Ser Ala Asp Gln Ile Lys Arg His Tyr Asp  
265 270 275 280

Ser Lys Asp Ala Glu Gly Asn Lys Ile Pro Leu Asp Lys Phe Lys Lys  
285 290 295

Glu Gly Lys Glu Lys Tyr Cys Pro Lys Thr Tyr Thr Phe Asp Asp Gly  
300 305 310

Lys Val Ile Ile Lys Ala Gly Ala Arg Val Glu Glu Glu Lys Val Lys  
315 320 325

Arg Leu Tyr Trp Ala Ser Lys Glu Val Asn Ser Gln Phe Phe Arg Val  
330 335 340

Tyr Gly Ile Asp Lys Pro Leu Glu Glu Gly Asn Pro Asp Asp Ile Leu  
345 350 355 360

Thr Met Val Ile Tyr Asn Ser Pro Glu Glu Tyr Lys Leu Asn Ser Val  
365 370 375

Leu Tyr Gly Tyr Asp Thr Asn Asn Gly Gly Met Tyr Ile Glu Pro Glu  
380 385 390

Gly Thr Phe Phe Thr Tyr Glu Arg Glu Ala Gln Glu Ser Thr Tyr Thr  
395 400 405

Leu Glu Glu Leu Phe Arg His Glu Tyr Thr His Tyr Leu Gln Gly Arg  
 410                      415                      420

Tyr Ala Val Pro Gly Gln Trp Gly Arg Thr Lys Leu Tyr Asp Asn Asp  
 425                      430                      435                      440

Arg Leu Thr Trp Tyr Glu Glu Gly Gly Ala Glu Leu Phe Ala Gly Ser  
 445                      450                      455

Thr Arg Thr Ser Gly Ile Leu Pro Arg Lys Ser Ile Val Ser Asn Ile  
 460                      465                      470

His Asn Thr Thr Arg Asn Asn Arg Tyr Lys Leu Ser Asp Thr Val His  
 475                      480                      485

Ser Lys Tyr Gly Ala Ser Phe Glu Phe Tyr Asn Tyr Ala Cys Met Phe  
 490                      495                      500

Met Asp Tyr Met Tyr Asn Lys Asp Met Gly Ile Leu Asn Lys Leu Asn  
 505                      510                      515                      520

Asp Leu Ala Lys Asn Asn Asp Val Asp Gly Tyr Asp Asn Tyr Ile Arg  
 525                      530                      535

Asp Leu Ser Ser Asn Tyr Ala Leu Asn Asp Lys Tyr Gln Asp His Met  
 540                      545                      550

Gln Glu Arg Ile Asp Asn Tyr Glu Asn Leu Thr Val Pro Phe Val Ala  
 555                      560                      565

Asp Asp Tyr Leu Val Arg His Ala Tyr Lys Asn Pro Asn Glu Ile Tyr  
 570                      575                      580

Ser Glu Ile Ser Glu Val Ala Lys Leu Lys Asp Ala Lys Ser Glu Val  
 585                      590                      595                      600

Lys Lys Ser Gln Tyr Phe Ser Thr Phe Thr Leu Arg Gly Ser Tyr Thr  
605 610 615

Gly Gly Ala Ser Lys Gly Lys Leu Glu Asp Gln Lys Ala Met Asn Lys  
620 625 630

Phe Ile Asp Asp Ser Leu Lys Lys Leu Asp Thr Tyr Ser Trp Ser Gly  
635 640 645

Tyr Lys Thr Leu Thr Ala Tyr Phe Thr Asn Tyr Lys Val Asp Ser Ser  
650 655 660

Asn Arg Val Thr Tyr Asp Val Val Phe His Gly Tyr Leu Pro Asn Glu  
665 670 675 680

Gly Asp Ser Lys Asn Ser Leu Pro Tyr Gly Lys Ile Asn Gly Thr Tyr  
685 690 695

Lys Gly Thr Glu Lys Glu Lys Ile Lys Phe Ser Ser Glu Gly Ser Phe  
700 705 710

Asp Pro Asp Gly Lys Ile Val Ser Tyr Glu Trp Asp Phe Gly Asp Gly  
715 720 725

Asn Lys Ser Asn Glu Glu Asn Pro Glu His Ser Tyr Asp Lys Val Gly  
730 735 740

Thr Tyr Thr Val Lys Leu Lys Val Thr Asp Asp Lys Gly Glu Ser Ser  
745 750 755 760

Val Ser Thr Thr Thr Ala Glu Ile Lys Asp Leu Ser Glu Asn Lys Leu  
765 770 775

Pro Val Ile Tyr Met His Val Pro Lys Ser Gly Ala Leu Asn Gln Lys  
 780 785 790

Val Val Phe Tyr Gly Lys Gly Thr Tyr Asp Pro Asp Gly Ser Ile Ala  
 795 800 805

Gly Tyr Gln Trp Asp Phe Gly Asp Gly Ser Asp Phe Ser Ser Glu Gln  
 810 815 820

Asn Pro Ser His Val Tyr Thr Lys Lys Gly Glu Tyr Thr Val Thr Leu  
 825 830 835 840

Arg Val Met Asp Ser Ser Gly Gln Met Ser Glu Lys Thr Met Lys Ile  
 845 850 855

Lys Ile Thr Asp Pro Val Tyr Pro Ile Gly Thr Glu Lys Glu Pro Asn  
 860 865 870

Asn Ser Lys Glu Thr Ala Ser Gly Pro Ile Val Pro Gly Ile Pro Val  
 875 880 885

Ser Gly Thr Ile Glu Asn Thr Ser Asp Gln Asp Tyr Phe Tyr Phe Asp  
 890 895 900

Val Ile Thr Pro Gly Glu Val Lys Ile Asp Ile Asn Lys Leu Gly Tyr  
 905 910 915 920

Gly Gly Ala Thr Trp Val Val Tyr Asp Glu Asn Asn Asn Ala Val Ser  
 925 930 935

Tyr Ala Thr Asp Asp Gly Gln Asn Leu Ser Gly Lys Phe Lys Ala Asp  
 940 945 950

Lys Pro Gly Arg Tyr Tyr Ile His Leu Tyr Met Phe Asn Gly Ser Tyr  
 955 960 965

Met Pro Tyr Arg Ile Asn Ile Glu Gly Ser Val Gly Arg  
 970 975 980

5 <210> 6  
 <211> 1021  
 <212> PRT  
 <213> *Clostridium histolyticum*

10 <220>  
 <221> SEÑAL  
 <222> (-40)..(-1)

15 <220>  
 <221> mat\_peptide  
 <222> (1)..(981)

<400> 6

Met Lys Arg Lys Cys Leu Ser Lys Arg Leu Met Leu Ala Ile Thr Met  
 -40 -35 -30 -25

Ala Thr Ile Phe Thr Val Asn Ser Thr Leu Pro Ile Tyr Ala Ala Val  
 -20 -15 -10

Asp Lys Asn Asn Ala Thr Ala Ala Val Gln Asn Glu Ser Lys Arg Tyr  
 -5 -1 1 5

Thr Val Ser Tyr Leu Lys Thr Leu Asn Tyr Tyr Asp Leu Val Asp Leu  
 10 15 20

Leu Val Lys Thr Glu Ile Glu Asn Leu Pro Asp Leu Phe Gln Tyr Ser  
 25 30 35 40

Ser Asp Ala Lys Glu Phe Tyr Gly Asn Lys Thr Arg Met Ser Phe Ile  
 45 50 55

Met Asp Glu Ile Gly Arg Arg Ala Pro Gln Tyr Thr Glu Ile Asp His

60                      65                      70

Lys Gly Ile Pro Thr Leu Val Glu Val Val Arg Ala Gly Phe Tyr Leu  
           75                      80                      85

Gly Phe His Asn Lys Glu Leu Asn Glu Ile Asn Lys Arg Ser Phe Lys  
           90                      95                      100

Glu Arg Val Ile Pro Ser Ile Leu Ala Ile Gln Lys Asn Pro Asn Phe  
       105                      110                      115                      120

Lys Leu Gly Thr Glu Val Gln Asp Lys Ile Val Ser Ala Thr Gly Leu  
           125                      130                      135

Leu Ala Gly Asn Glu Thr Ala Pro Pro Glu Val Val Asn Asn Phe Thr  
           140                      145                      150

Pro Ile Leu Gln Asp Cys Ile Lys Asn Ile Asp Arg Tyr Ala Leu Asp  
           155                      160                      165

Asp Leu Lys Ser Lys Ala Leu Phe Asn Val Leu Ala Ala Pro Thr Tyr  
           170                      175                      180

Asp Ile Thr Glu Tyr Leu Arg Ala Thr Lys Glu Lys Pro Glu Asn Thr  
       185                      190                      195                      200

Pro Trp Tyr Gly Lys Ile Asp Gly Phe Ile Asn Glu Leu Lys Lys Leu  
           205                      210                      215

Ala Leu Tyr Gly Lys Ile Asn Asp Asn Asn Ser Trp Ile Ile Asp Asn  
           220                      225                      230

Gly Ile Tyr His Ile Ala Pro Leu Gly Lys Leu His Ser Asn Asn Lys  
           235                      240                      245

Ile Gly Ile Glu Thr Leu Thr Glu Val Met Lys Val Tyr Pro Tyr Leu  
 250                      255                      260

Ser Met Gln His Leu Gln Ser Ala Asp Gln Ile Lys Arg His Tyr Asp  
 265                      270                      275                      280

Ser Lys Asp Ala Glu Gly Asn Lys Ile Pro Leu Asp Lys Phe Lys Lys  
                     285                      290                      295

Glu Gly Lys Glu Lys Tyr Cys Pro Lys Thr Tyr Thr Phe Asp Asp Gly  
                     300                      305                      310

Lys Val Ile Ile Lys Ala Gly Ala Arg Val Glu Glu Glu Lys Val Lys  
                     315                      320                      325

Arg Leu Tyr Trp Ala Ser Lys Glu Val Asn Ser Gln Phe Phe Arg Val  
                     330                      335                      340

Tyr Gly Ile Asp Lys Pro Leu Glu Glu Gly Asn Pro Asp Asp Ile Leu  
 345                      350                      355                      360

Thr Met Val Ile Tyr Asn Ser Pro Glu Glu Tyr Lys Leu Asn Ser Val  
                     365                      370                      375

Leu Tyr Gly Tyr Asp Thr Asn Asn Gly Gly Met Tyr Ile Glu Pro Glu  
                     380                      385                      390

Gly Thr Phe Phe Thr Tyr Glu Arg Glu Ala Gln Glu Ser Thr Tyr Thr  
                     395                      400                      405

Leu Glu Glu Leu Phe Arg His Glu Tyr Thr His Tyr Leu Gln Gly Arg  
                     410                      415                      420

Tyr Ala Val Pro Gly Gln Trp Gly Arg Thr Lys Leu Tyr Asp Asn Asp

425                    430                    435                    440  
 Arg Leu Thr Trp Tyr Glu Glu Gly Gly Ala Glu Leu Phe Ala Gly Ser  
                           445                    450                    455  
 Thr Arg Thr Ser Gly Ile Leu Pro Arg Lys Ser Ile Val Ser Asn Ile  
                           460                    465                    470  
 His Asn Thr Thr Arg Asn Asn Arg Tyr Lys Leu Ser Asp Thr Val His  
                           475                    480                    485  
 Ser Lys Tyr Gly Ala Ser Phe Glu Phe Tyr Asn Tyr Ala Cys Met Phe  
                           490                    495                    500  
 Met Asp Tyr Met Tyr Asn Lys Asp Met Gly Ile Leu Asn Lys Leu Asn  
 505                    510                    515                    520  
 Asp Leu Ala Lys Asn Asn Asp Val Asp Gly Tyr Asp Asn Tyr Ile Arg  
                           525                    530                    535  
 Asp Leu Ser Ser Asn Tyr Ala Leu Asn Asp Lys Tyr Gln Asp His Met  
                           540                    545                    550  
 Gln Glu Arg Ile Asp Asn Tyr Glu Asn Leu Thr Val Pro Phe Val Ala  
                           555                    560                    565  
 Asp Asp Tyr Leu Val Arg His Ala Tyr Lys Asn Pro Asn Glu Ile Tyr  
                           570                    575                    580  
 Ser Glu Ile Ser Glu Val Ala Lys Leu Lys Asp Ala Lys Ser Glu Val  
 585                    590                    595                    600  
 Lys Lys Ser Gln Tyr Phe Ser Thr Phe Thr Leu Arg Gly Ser Tyr Thr  
                           605                    610                    615

Gly Gly Ala Ser Lys Gly Lys Leu Glu Asp Gln Lys Ala Met Asn Lys  
620 625 630

Phe Ile Asp Asp Ser Leu Lys Lys Leu Asp Thr Tyr Ser Trp Ser Gly  
635 640 645

Tyr Lys Thr Leu Thr Ala Tyr Phe Thr Asn Tyr Lys Val Asp Ser Ser  
650 655 660

Asn Arg Val Thr Tyr Asp Val Val Phe His Gly Tyr Leu Pro Asn Glu  
665 670 675 680

Gly Asp Ser Lys Asn Ser Leu Pro Tyr Gly Lys Ile Asn Gly Thr Tyr  
685 690 695

Lys Gly Thr Glu Lys Glu Lys Ile Lys Phe Ser Ser Glu Gly Ser Phe  
700 705 710

Asp Pro Asp Gly Lys Ile Val Ser Tyr Glu Trp Asp Phe Gly Asp Gly  
715 720 725

Asn Lys Ser Asn Glu Glu Asn Pro Glu His Ser Tyr Asp Lys Val Gly  
730 735 740

Thr Tyr Thr Val Lys Leu Lys Val Thr Asp Asp Lys Gly Glu Ser Ser  
745 750 755 760

Val Ser Thr Thr Thr Ala Glu Ile Lys Asp Leu Ser Glu Asn Lys Leu  
765 770 775

Pro Val Ile Tyr Met His Val Pro Lys Ser Gly Ala Leu Asn Gln Lys  
780 785 790

Val Val Phe Tyr Gly Lys Gly Thr Tyr Asp Pro Asp Gly Ser Ile Ala

795                      800                      805

Gly Tyr Gln Trp Asp Phe Gly Asp Gly Ser Asp Phe Ser Ser Glu Gln  
 810                      815                      820

Asn Pro Ser His Val Tyr Thr Lys Lys Gly Glu Tyr Thr Val Thr Leu  
 825                      830                      835                      840

Arg Val Met Asp Ser Ser Gly Gln Met Ser Glu Lys Thr Met Lys Ile  
 845                      850                      855

Lys Ile Thr Asp Pro Val Tyr Pro Ile Gly Thr Glu Lys Glu Pro Asn  
 860                      865                      870

Asn Ser Lys Glu Thr Ala Ser Gly Pro Ile Val Pro Gly Ile Pro Val  
 875                      880                      885

Ser Gly Thr Ile Glu Asn Thr Ser Asp Gln Asp Tyr Phe Tyr Phe Asp  
 890                      895                      900

Val Ile Thr Pro Gly Glu Val Lys Ile Asp Ile Asn Lys Leu Gly Tyr  
 905                      910                      915                      920

Gly Gly Ala Thr Trp Val Val Tyr Asp Glu Asn Asn Asn Ala Val Ser  
 925                      930                      935

Tyr Ala Thr Asp Asp Gly Gln Asn Leu Ser Gly Lys Phe Lys Ala Asp  
 940                      945                      950

Lys Pro Gly Arg Tyr Tyr Ile His Leu Tyr Met Phe Asn Gly Ser Tyr  
 955                      960                      965

Met Pro Tyr Arg Ile Asn Ile Glu Gly Ser Val Ser Arg  
 970                      975                      980

<212> PRT  
 <213> *Clostridium histolyticum*

5 <220>  
 <221> SEÑAL  
 <222> (-40)..(-1)

10 <220>  
 <221> mat\_peptide  
 <222> (1)..(994)

<400> 7

Met Lys Arg Lys Cys Leu Ser Lys Arg Leu Met Leu Ala Ile Thr Met  
 -40 -35 -30 -25

Ala Thr Ile Phe Thr Val Asn Ser Thr Leu Pro Ile Tyr Ala Ala Val  
 -20 -15 -10

Asp Lys Asn Asn Ala Thr Ala Ala Val Gln Asn Glu Ser Lys Arg Tyr  
 -5 -1 1 5

Thr Val Ser Tyr Leu Lys Thr Leu Asn Tyr Tyr Asp Leu Val Asp Leu  
 10 15 20

Leu Val Lys Thr Glu Ile Glu Asn Leu Pro Asp Leu Phe Gln Tyr Ser  
 25 30 35 40

Ser Asp Ala Lys Glu Phe Tyr Gly Asn Lys Thr Arg Met Ser Phe Ile  
 45 50 55

Met Asp Glu Ile Gly Arg Arg Ala Pro Gln Tyr Thr Glu Ile Asp His  
 60 65 70

Lys Gly Ile Pro Thr Leu Val Glu Val Val Arg Ala Gly Phe Tyr Leu  
 75 80 85

Gly Phe His Asn Lys Glu Leu Asn Glu Ile Asn Lys Arg Ser Phe Lys  
90 95 100

Glu Arg Val Ile Pro Ser Ile Leu Ala Ile Gln Lys Asn Pro Asn Phe  
105 110 115 120

Lys Leu Gly Thr Glu Val Gln Asp Lys Ile Val Ser Ala Thr Gly Leu  
125 130 135

Leu Ala Gly Asn Glu Thr Ala Pro Pro Glu Val Val Asn Asn Phe Thr  
140 145 150

Pro Ile Leu Gln Asp Cys Ile Lys Asn Ile Asp Arg Tyr Ala Leu Asp  
155 160 165

Asp Leu Lys Ser Lys Ala Leu Phe Asn Val Leu Ala Ala Pro Thr Tyr  
170 175 180

Asp Ile Thr Glu Tyr Leu Arg Ala Thr Lys Glu Lys Pro Glu Asn Thr  
185 190 195 200

Pro Trp Tyr Gly Lys Ile Asp Gly Phe Ile Asn Glu Leu Lys Lys Leu  
205 210 215

Ala Leu Tyr Gly Lys Ile Asn Asp Asn Asn Ser Trp Ile Ile Asp Asn  
220 225 230

Gly Ile Tyr His Ile Ala Pro Leu Gly Lys Leu His Ser Asn Asn Lys  
235 240 245

Ile Gly Ile Glu Thr Leu Thr Glu Val Met Lys Val Tyr Pro Tyr Leu  
250 255 260

Ser Met Gln His Leu Gln Ser Ala Asp Gln Ile Lys Arg His Tyr Asp  
 265                    270                    275                    280

Ser Lys Asp Ala Glu Gly Asn Lys Ile Pro Leu Asp Lys Phe Lys Lys  
                   285                    290                    295

Glu Gly Lys Glu Lys Tyr Cys Pro Lys Thr Tyr Thr Phe Asp Asp Gly  
                   300                    305                    310

Lys Val Ile Ile Lys Ala Gly Ala Arg Val Glu Glu Glu Lys Val Lys  
                   315                    320                    325

Arg Leu Tyr Trp Ala Ser Lys Glu Val Asn Ser Gln Phe Phe Arg Val  
                   330                    335                    340

Tyr Gly Ile Asp Lys Pro Leu Glu Glu Gly Asn Pro Asp Asp Ile Leu  
 345                    350                    355                    360

Thr Met Val Ile Tyr Asn Ser Pro Glu Glu Tyr Lys Leu Asn Ser Val  
                   365                    370                    375

Leu Tyr Gly Tyr Asp Thr Asn Asn Gly Gly Met Tyr Ile Glu Pro Glu  
                   380                    385                    390

Gly Thr Phe Phe Thr Tyr Glu Arg Glu Ala Gln Glu Ser Thr Tyr Thr  
                   395                    400                    405

Leu Glu Glu Leu Phe Arg His Glu Tyr Thr His Tyr Leu Gln Gly Arg  
                   410                    415                    420

Tyr Ala Val Pro Gly Gln Trp Gly Arg Thr Lys Leu Tyr Asp Asn Asp  
 425                    430                    435                    440

Arg Leu Thr Trp Tyr Glu Glu Gly Gly Ala Glu Leu Phe Ala Gly Ser  
                   445                    450                    455

Thr Arg Thr Ser Gly Ile Leu Pro Arg Lys Ser Ile Val Ser Asn Ile  
460 465 470

His Asn Thr Thr Arg Asn Asn Arg Tyr Lys Leu Ser Asp Thr Val His  
475 480 485

Ser Lys Tyr Gly Ala Ser Phe Glu Phe Tyr Asn Tyr Ala Cys Met Phe  
490 495 500

Met Asp Tyr Met Tyr Asn Lys Asp Met Gly Ile Leu Asn Lys Leu Asn  
505 510 515 520

Asp Leu Ala Lys Asn Asn Asp Val Asp Gly Tyr Asp Asn Tyr Ile Arg  
525 530 535

Asp Leu Ser Ser Asn Tyr Ala Leu Asn Asp Lys Tyr Gln Asp His Met  
540 545 550

Gln Glu Arg Ile Asp Asn Tyr Glu Asn Leu Thr Val Pro Phe Val Ala  
555 560 565

Asp Asp Tyr Leu Val Arg His Ala Tyr Lys Asn Pro Asn Glu Ile Tyr  
570 575 580

Ser Glu Ile Ser Glu Val Ala Lys Leu Lys Asp Ala Lys Ser Glu Val  
585 590 595 600

Lys Lys Ser Gln Tyr Phe Ser Thr Phe Thr Leu Arg Gly Ser Tyr Thr  
605 610 615

Gly Gly Ala Ser Lys Gly Lys Leu Glu Asp Gln Lys Ala Met Asn Lys  
620 625 630

Phe Ile Asp Asp Ser Leu Lys Lys Leu Asp Thr Tyr Ser Trp Ser Gly  
 635                      640                      645

Tyr Lys Thr Leu Thr Ala Tyr Phe Thr Asn Tyr Lys Val Asp Ser Ser  
 650                      655                      660

Asn Arg Val Thr Tyr Asp Val Val Phe His Gly Tyr Leu Pro Asn Glu  
 665                      670                      675                      680

Gly Asp Ser Lys Asn Ser Leu Pro Tyr Gly Lys Ile Asn Gly Thr Tyr  
 685                      690                      695

Lys Gly Thr Glu Lys Glu Lys Ile Lys Phe Ser Ser Glu Gly Ser Phe  
 700                      705                      710

Asp Pro Asp Gly Lys Ile Val Ser Tyr Glu Trp Asp Phe Gly Asp Gly  
 715                      720                      725

Asn Lys Ser Asn Glu Glu Asn Pro Glu His Ser Tyr Asp Lys Val Gly  
 730                      735                      740

Thr Tyr Thr Val Lys Leu Lys Val Thr Asp Asp Lys Gly Glu Ser Ser  
 745                      750                      755                      760

Val Ser Thr Thr Thr Ala Glu Ile Lys Asp Leu Ser Glu Asn Lys Leu  
 765                      770                      775

Pro Val Ile Tyr Met His Val Pro Lys Ser Gly Ala Leu Asn Gln Lys  
 780                      785                      790

Val Val Phe Tyr Gly Lys Gly Thr Tyr Asp Pro Asp Gly Ser Ile Ala  
 795                      800                      805

Gly Tyr Gln Trp Asp Phe Gly Asp Gly Ser Asp Phe Ser Ser Glu Gln  
 810                      815                      820

Asn Pro Ser His Val Tyr Thr Lys Lys Gly Glu Tyr Thr Val Thr Leu  
 825                    830                    835                    840

Arg Val Met Asp Ser Ser Gly Gln Met Ser Glu Lys Thr Met Lys Ile  
                   845                    850                    855

Lys Ile Thr Asp Pro Val Tyr Pro Ile Gly Thr Glu Lys Glu Pro Asn  
                   860                    865                    870

Asn Ser Lys Glu Thr Ala Ser Gly Pro Ile Val Pro Gly Ile Pro Val  
                   875                    880                    885

Ser Gly Thr Ile Glu Asn Thr Ser Asp Gln Asp Tyr Phe Tyr Phe Asp  
                   890                    895                    900

Val Ile Thr Pro Gly Glu Val Lys Ile Asp Ile Asn Lys Leu Gly Tyr  
 905                    910                    915                    920

Gly Gly Ala Thr Trp Val Val Tyr Asp Glu Asn Asn Asn Ala Val Ser  
                   925                    930                    935

Tyr Ala Thr Asp Asp Gly Gln Asn Leu Ser Gly Lys Phe Lys Ala Asp  
                   940                    945                    950

Lys Pro Gly Arg Tyr Tyr Ile His Leu Tyr Met Phe Asn Gly Ser Tyr  
                   955                    960                    965

Met Pro Tyr Arg Ile Asn Ile Glu Gly Ser Val Ser Arg Lys Leu Ala  
                   970                    975                    980

Ala Ala Leu Glu His His His His His His  
 985                    990

<210> 8  
 <211> 3105  
 <212> ADN  
 <213> *Clostridium histolyticum*

<220>  
 <221> CDS  
 <222> (1)..(3105)

5 <400> 8

atg aaa agg aaa tgt tta tct aaa agg ctt atg tta gct ata aca atg 48  
 Met Lys Arg Lys Cys Leu Ser Lys Arg Leu Met Leu Ala Ile Thr Met  
 1 5 10 15

gct aca ata ttt aca gtg aac agt aca tta cca att tat gca gct gta 96  
 Ala Thr Ile Phe Thr Val Asn Ser Thr Leu Pro Ile Tyr Ala Ala Val  
 20 25 30

gat aaa aat aat gca aca gca gct gta caa aat gaa agt aag agg tat 144  
 Asp Lys Asn Asn Ala Thr Ala Ala Val Gln Asn Glu Ser Lys Arg Tyr  
 35 40 45

aca gta tca tat tta aag act tta aat tat tat gac tta gta gat ttg 192  
 Thr Val Ser Tyr Leu Lys Thr Leu Asn Tyr Tyr Asp Leu Val Asp Leu  
 50 55 60

ctt gtt aag act gaa att gag aat tta cca gac ctt ttt cag tat agt 240  
 Leu Val Lys Thr Glu Ile Glu Asn Leu Pro Asp Leu Phe Gln Tyr Ser  
 65 70 75 80

tca gat gca aaa gag ttc tat gga aat aaa act cgt atg agc ttt atc 288  
 Ser Asp Ala Lys Glu Phe Tyr Gly Asn Lys Thr Arg Met Ser Phe Ile  
 85 90 95

atg gat gaa att ggt aga agg gca cct cag tat aca gag ata gat cat 336  
 Met Asp Glu Ile Gly Arg Arg Ala Pro Gln Tyr Thr Glu Ile Asp His  
 100 105 110

aaa ggt att cct act tta gta gaa gtt gta aga gct gga ttt tac tta 384  
 Lys Gly Ile Pro Thr Leu Val Glu Val Val Arg Ala Gly Phe Tyr Leu  
 115 120 125

gga ttc cat aac aag gaa ttg aat gaa ata aac aag agg tct ttt aaa 432  
 Gly Phe His Asn Lys Glu Leu Asn Glu Ile Asn Lys Arg Ser Phe Lys  
 130 135 140

gaa agg gta ata cct tet ata tta gca att caa aaa aat cct aat ttt 480  
 Glu Arg Val Ile Pro Ser Ile Leu Ala Ile Gln Lys Asn Pro Asn Phe  
 145 150 155 160

aaa cta ggt act gaa gtt caa gat aaa ata gta tct gca aca gga ctt 528  
 Lys Leu Gly Thr Glu Val Gln Asp Lys Ile Val Ser Ala Thr Gly Leu  
 165 170 175

tta gct ggt aat gaa aca gcg cct cca gaa gtt gta aat aat ttt aca 576  
 Leu Ala Gly Asn Glu Thr Ala Pro Pro Glu Val Val Asn Asn Phe Thr  
 180 185 190

cca ata ctt caa gac tgt ata aag aat ata gac aga tac gct ctt gat 624  
 Pro Ile Leu Gln Asp Cys Ile Lys Asn Ile Asp Arg Tyr Ala Leu Asp  
 195 200 205

gat tta aag tca aaa gca tta ttt aat gtt tta gct gca cct acc tat 672  
 Asp Leu Lys Ser Lys Ala Leu Phe Asn Val Leu Ala Ala Pro Thr Tyr  
 210 215 220

gat ata act gag tat tta aga gct act aaa gaa aaa cca gaa aac act 720  
 Asp Ile Thr Glu Tyr Leu Arg Ala Thr Lys Glu Lys Pro Glu Asn Thr  
 225 230 235 240

cct tgg tat ggt aaa ata gat ggg ttt ata aat gaa ctt aaa aag tta 768  
 Pro Trp Tyr Gly Lys Ile Asp Gly Phe Ile Asn Glu Leu Lys Lys Leu  
 245 250 255

gct ctt tat gga aaa ata aat gat aat aac tct tgg ata ata gat aac 816  
 Ala Leu Tyr Gly Lys Ile Asn Asp Asn Asn Ser Trp Ile Ile Asp Asn  
 260 265 270

ggt ata tat cat ata gca cct tta ggg aag tta cat agc aat aat aaa 864  
 Gly Ile Tyr His Ile Ala Pro Leu Gly Lys Leu His Ser Asn Asn Lys  
 275 280 285

ata gga ata gaa act tta aca gag gtt atg aaa gtt tat cct tat tta 912  
 Ile Gly Ile Glu Thr Leu Thr Glu Val Met Lys Val Tyr Pro Tyr Leu  
 290 295 300

agt atg caa cat tta caa tca gca gat caa att aag cgt cat tat gat 960  
 Ser Met Gln His Leu Gln Ser Ala Asp Gln Ile Lys Arg His Tyr Asp  
 305 310 315 320

tca aaa gat gct gaa gga aac aaa ata cct tta gat aag ttt aaa aag 1008  
 Ser Lys Asp Ala Glu Gly Asn Lys Ile Pro Leu Asp Lys Phe Lys Lys

325	330	335	
gaa gga aaa gaa aaa tac tgt cca aaa act tat aca ttt gat gat gga 1056			
Glu Gly Lys Glu Lys Tyr Cys Pro Lys Thr Tyr Thr Phe Asp Asp Gly			
340	345	350	
aaa gta ata ata aaa gct ggt gct aga gta gaa gaa gaa aaa gtt aaa 1104			
Lys Val Ile Ile Lys Ala Gly Ala Arg Val Glu Glu Glu Lys Val Lys			
355	360	365	
aga cta tac tgg gca tca aag gaa gtt aac tct caa ttc ttt aga gta 1152			
Arg Leu Tyr Trp Ala Ser Lys Glu Val Asn Ser Gln Phe Phe Arg Val			
370	375	380	
tac gga ata gac aaa cca tta gaa gaa ggt aat cca gat gat ata tta 1200			
Tyr Gly Ile Asp Lys Pro Leu Glu Glu Gly Asn Pro Asp Asp Ile Leu			
385	390	395	400
aca atg gtt atc tac aac agt ccc gaa gaa tat aaa ctc aat agt gtt 1248			
Thr Met Val Ile Tyr Asn Ser Pro Glu Glu Tyr Lys Leu Asn Ser Val			
405	410	415	
cta tac gga tat gat act aat aat ggt ggt atg tat ata gag cca gaa 1296			
Leu Tyr Gly Tyr Asp Thr Asn Asn Gly Gly Met Tyr Ile Glu Pro Glu			
420	425	430	
gga act ttc ttc acc tat gaa aga gaa gct caa gaa agc aca tac aca 1344			
Gly Thr Phe Phe Thr Tyr Glu Arg Glu Ala Gln Glu Ser Thr Tyr Thr			
435	440	445	
tta gaa gaa tta ttt aga cat gaa tat aca cat tat ttg caa gga aga 1392			
Leu Glu Glu Leu Phe Arg His Glu Tyr Thr His Tyr Leu Gln Gly Arg			
450	455	460	
tat gca gtt cca gga caa tgg gga aga aca aaa ctt tat gac aat gat 1440			
Tyr Ala Val Pro Gly Gln Trp Gly Arg Thr Lys Leu Tyr Asp Asn Asp			
465	470	475	480
aga tta act tgg tat gaa gaa ggt gga gca gaa tta ttt gca ggt tct 1488			
Arg Leu Thr Trp Tyr Glu Glu Gly Gly Ala Glu Leu Phe Ala Gly Ser			
485	490	495	
act aga act tct gga ata tta cca aga aag agt ata gta tca aat att 1536			
Thr Arg Thr Ser Gly Ile Leu Pro Arg Lys Ser Ile Val Ser Asn Ile			
500	505	510	

cat aat aca aca aga aat aat aga tat aag ctt tca gac act gta cat 1584  
 His Asn Thr Thr Arg Asn Asn Arg Tyr Lys Leu Ser Asp Thr Val His  
 515 520 525

tct aaa tat ggt gct agt ttt gaa ttc tat aat tat gca tgt atg ttt 1632  
 Ser Lys Tyr Gly Ala Ser Phe Glu Phe Tyr Asn Tyr Ala Cys Met Phe  
 530 535 540

atg gat tat atg tat aat aaa gat atg ggt ata tta aat aaa cta aat 1680  
 Met Asp Tyr Met Tyr Asn Lys Asp Met Gly Ile Leu Asn Lys Leu Asn  
 545 550 555 560

gat ctt gca aaa aat aat gat gtt gat gga tat gat aat tat att aga 1728  
 Asp Leu Ala Lys Asn Asn Asp Val Asp Gly Tyr Asp Asn Tyr Ile Arg  
 565 570 575

gat tta agt tct aat tat gct tta aat gat aaa tat caa gat cat atg 1776  
 Asp Leu Ser Ser Asn Tyr Ala Leu Asn Asp Lys Tyr Gln Asp His Met  
 580 585 590

cag gag cgc ata gat aat tat gaa aat tta aca gtg cct ttt gta gct 1824  
 Gln Glu Arg Ile Asp Asn Tyr Glu Asn Leu Thr Val Pro Phe Val Ala  
 595 600 605

gat gat tat tta gta agg cat gct tat aag aac cct aat gaa att tat 1872  
 Asp Asp Tyr Leu Val Arg His Ala Tyr Lys Asn Pro Asn Glu Ile Tyr  
 610 615 620

tct gaa ata tct gaa gta gca aaa tta aag gat gct aag agt gaa gtt 1920  
 Ser Glu Ile Ser Glu Val Ala Lys Leu Lys Asp Ala Lys Ser Glu Val  
 625 630 635 640

aag aaa tca caa tat ttt agt acc ttt act ttg aga ggt agt tac aca 1968  
 Lys Lys Ser Gln Tyr Phe Ser Thr Phe Thr Leu Arg Gly Ser Tyr Thr  
 645 650 655

ggt gga gca tct aag ggg aaa tta gaa gat caa aaa gca atg aat aag 2016  
 Gly Gly Ala Ser Lys Gly Lys Leu Glu Asp Gln Lys Ala Met Asn Lys  
 660 665 670

ttt ata gat gat tca ctt aag aaa tta gat acg tat tct tgg agt ggg 2064  
 Phe Ile Asp Asp Ser Leu Lys Lys Leu Asp Thr Tyr Ser Trp Ser Gly  
 675 680 685

tat aaa act tta act gct tat ttc act aat tat aaa gtt gac tct tca 2112  
 Tyr Lys Thr Leu Thr Ala Tyr Phe Thr Asn Tyr Lys Val Asp Ser Ser

690	695	700	
aat aga gtt act tat gat gta gta ttc cac gga tat tta cca aac gaa 2160			
Asn Arg Val Thr Tyr Asp Val Val Phe His Gly Tyr Leu Pro Asn Glu			
705	710	715	720
ggt gat tcc aaa aat tca tta cct tat ggc aag atc aat gga act tac 2208			
Gly Asp Ser Lys Asn Ser Leu Pro Tyr Gly Lys Ile Asn Gly Thr Tyr			
	725	730	735
aag gga aca gag aaa gaa aaa atc aaa ttc tct agt gaa ggc tct ttc 2256			
Lys Gly Thr Glu Lys Glu Lys Ile Lys Phe Ser Ser Glu Gly Ser Phe			
	740	745	750
gat cca gat ggt aaa ata gtt tct tat gaa tgg gat ttc gga gat ggt 2304			
Asp Pro Asp Gly Lys Ile Val Ser Tyr Glu Trp Asp Phe Gly Asp Gly			
	755	760	765
aat aag agt aat gag gaa aat cca gag cat tca tat gac aag gta gga 2352			
Asn Lys Ser Asn Glu Glu Asn Pro Glu His Ser Tyr Asp Lys Val Gly			
	770	775	780
act tat aca gtg aaa tta aaa gtt act gat gac aag gga gaa tct tca 2400			
Thr Tyr Thr Val Lys Leu Lys Val Thr Asp Asp Lys Gly Glu Ser Ser			
785	790	795	800
gta tct act act act gca gaa ata aag gat ctt tca gaa aat aaa ctt 2448			
Val Ser Thr Thr Thr Ala Glu Ile Lys Asp Leu Ser Glu Asn Lys Leu			
	805	810	815
cca gtt ata tat atg cat gta cct aaa tcc gga gcc tta aat caa aaa 2496			
Pro Val Ile Tyr Met His Val Pro Lys Ser Gly Ala Leu Asn Gln Lys			
	820	825	830
gtt gtt ttc tat gga aaa gga aca tat gac cca gat gga tct atc gca 2544			
Val Val Phe Tyr Gly Lys Gly Thr Tyr Asp Pro Asp Gly Ser Ile Ala			
	835	840	845
gga tat caa tgg gac ttt ggt gat gga agt gat ttt agc agt gaa caa 2592			
Gly Tyr Gln Trp Asp Phe Gly Asp Gly Ser Asp Phe Ser Ser Glu Gln			
	850	855	860
aac cca agc cat gta tat act aaa aaa ggt gaa tat act gta aca tta 2640			
Asn Pro Ser His Val Tyr Thr Lys Lys Gly Glu Tyr Thr Val Thr Leu			
865	870	875	880

aga gta atg gat agt agt gga caa atg agt gaa aaa act atg aag att 2688  
 Arg Val Met Asp Ser Ser Gly Gln Met Ser Glu Lys Thr Met Lys Ile  
 885 890 895

aag att aca gat ccg gta tat cca ata ggc act gaa aaa gaa cca aat 2736  
 Lys Ile Thr Asp Pro Val Tyr Pro Ile Gly Thr Glu Lys Glu Pro Asn  
 900 905 910

aac agt aaa gaa act gca agt ggt cca ata gta cca ggt ata cct gtt 2784  
 Asn Ser Lys Glu Thr Ala Ser Gly Pro Ile Val Pro Gly Ile Pro Val  
 915 920 925

agt gga acc ata gaa aat aca agt gat caa gat tat ttc tat ttt gat 2832  
 Ser Gly Thr Ile Glu Asn Thr Ser Asp Gln Asp Tyr Phe Tyr Phe Asp  
 930 935 940

gtt ata aca cca gga gaa gta aaa ata gat ata aat aaa tta ggg tac 2880  
 Val Ile Thr Pro Gly Glu Val Lys Ile Asp Ile Asn Lys Leu Gly Tyr  
 945 950 955 960

gga gga gct act tgg gta gta tat gat gaa aat aat aat gca gta tct 2928  
 Gly Gly Ala Thr Trp Val Val Tyr Asp Glu Asn Asn Asn Ala Val Ser  
 965 970 975

tat gcc act gat gat ggg caa aat tta agt gga aag ttt aag gca gat 2976  
 Tyr Ala Thr Asp Asp Gly Gln Asn Leu Ser Gly Lys Phe Lys Ala Asp  
 980 985 990

aaa cca ggt aga tat tac atc cat ctt tac atg ttt aat ggt agt tat 3024  
 Lys Pro Gly Arg Tyr Tyr Ile His Leu Tyr Met Phe Asn Gly Ser Tyr  
 995 1000 1005

atg cca tat aga att aat ata gaa ggt tca gta tct aga aag ctt 3069  
 Met Pro Tyr Arg Ile Asn Ile Glu Gly Ser Val Ser Arg Lys Leu  
 1010 1015 1020

gcg gcc gca ctc gag cac cac cac cac cac tga 3105  
 Ala Ala Ala Leu Glu His His His His His His  
 1025 1030

<210> 9  
 <211> 25  
 <212> ADN  
 <213> Artificial

5

<220>  
 <223> Un cebador para PCR (colG-F)

10

<400> 9

ES 2 531 402 T3

atgaaaaaaa atatttaaa gattc 25

5 <210> 10  
<211> 27  
<212> ADN  
<213> Artificial

10 <220>  
<223> Un cebador para PCR (colG-R)

<400> 10  
ccggatccta tctagatacc cttact 27

15 <210> 11  
<211> 28  
<212> ADN  
<213> Artificial

20 <220>  
<223> Un cebador para PCR (His-F)

<400> 11  
gctctagaaa gcttgcgcc gcactcga 28

25 <210> 12  
<211> 24  
<212> ADN  
<213> Artificial

30 <220>  
<223> Un cebador para PCR (His-R)

35 <400> 12  
cgggatccgg atatagtcc tcct 24

40 <210> 13  
<211> 28  
<212> ADN  
<213> Artificial

45 <220>  
<223> Un cebador para PCR (lac-F)

<400> 13  
ccggcaagct tgccaatac gcaaaccg 28

50 <210> 14  
<211> 18  
<212> ADN  
<213> Artificial

55 <220>  
<223> Un cebador para PCR (lac-R)

<400> 14  
agctgttcc tgtgtgaa 18

60 <210> 15  
<211> 20  
<212> ADN  
<213> Artificial

65 <220>  
<223> Un cebador para PCR (colH-F)

<400> 15

# ES 2 531 402 T3

atgaaaagga aatgtttatc 20

<210> 16

<211> 27

5 <212> ADN

<213> Artificial

<220>

10 <223> Un cebador para PCR (colH-R)

<400> 16

ccggatccta tctagatact gaacctt 27

**REIVINDICACIONES**

1. Un método para producir una colagenasa de fusión,  
en el que:

5 un cultivo obtenido mediante el cultivo de *E. coli* transformada por uno cualquiera de un ADN que codifica la colagenasa de fusión y un vector de expresión que comprende el ADN se purifica mediante cromatografía por afinidad que corresponde a una marca de afinidad para recoger de ese modo de forma selectiva la colagenasa de fusión que tiene un dominio de unión a colágeno, en donde la colagenasa de fusión tiene las siguientes características (i) a (ii):

- (i) la marca de afinidad se une directa o indirectamente a un extremo carboxilo de la colagenasa;
- (ii) la colagenasa se selecciona entre los siguientes (a) a (h):

- 15 (a) una colagenasa que comprende una secuencia de aminoácidos de las posiciones -110 a 1008 de una cualquiera de las SEC ID N<sup>o</sup>: 1 y 2;
- (b) una colagenasa que comprende una secuencia de aminoácidos de las posiciones -110 a 1008 de una cualquiera de las SEC ID N<sup>o</sup>: 1 y 2 en la que se suprimen, sustituyen, insertan o añaden de 1 a 30 aminoácidos;
- 20 (c) una colagenasa que comprende una secuencia de aminoácidos que tiene una homología de un 80 % o superior con la secuencia de aminoácidos de las posiciones -110 a 1008 de una cualquiera de las SEC ID N<sup>o</sup>: 1 y 2;
- (d) una colagenasa en la que una secuencia de aminoácidos de las posiciones -110 a -1 de una cualquiera de las SEC ID N<sup>o</sup>: 1 y 2 se retira de una cualquiera de las colagenasas que se han descrito en
- 25 (a) a (c);
- (e) una colagenasa que comprende una secuencia de aminoácidos de las posiciones -40 a 981 de una cualquiera de las SEC ID N<sup>o</sup>: 5 y 6;
- (f) una colagenasa que comprende una secuencia de aminoácidos de las posiciones -40 a 981 de una cualquiera de las SEC ID N<sup>o</sup>: 5 y 6 en la que se suprimen, sustituyen, insertan o añaden de 1 a 30 aminoácidos;
- 30 (g) una colagenasa que comprende una secuencia de aminoácidos que tiene una homología de un 80 % o superior con la secuencia de aminoácidos de las posiciones -40 a 981 de una cualquiera de las SEC ID N<sup>o</sup>: 5 y 6;
- (h) una colagenasa en la que una secuencia de aminoácidos de las posiciones -40 a -1 de una cualquiera de las SEC ID N<sup>o</sup>: 5 y 6 se retira de una cualquiera de las colagenasas que se han descrito en (e) a (g).

2. El método de producción de acuerdo con la reivindicación 1, en el que la colagenasa de fusión es una colagenasa de fusión en la que la marca de afinidad es dos o más restos de histidina consecutivos.

40 3. El método de producción de acuerdo con la reivindicación 1, en el que la colagenasa de fusión comprende una secuencia de aminoácidos de las posiciones -110 a 1021 de la SEC ID N<sup>o</sup>: 3.

45 4. El método de producción de acuerdo con la reivindicación 3, en el que la colagenasa de fusión es una colagenasa de fusión a partir de la que se retira una secuencia de aminoácidos de las posiciones -110 a -1 de la SEC ID N<sup>o</sup>: 3.

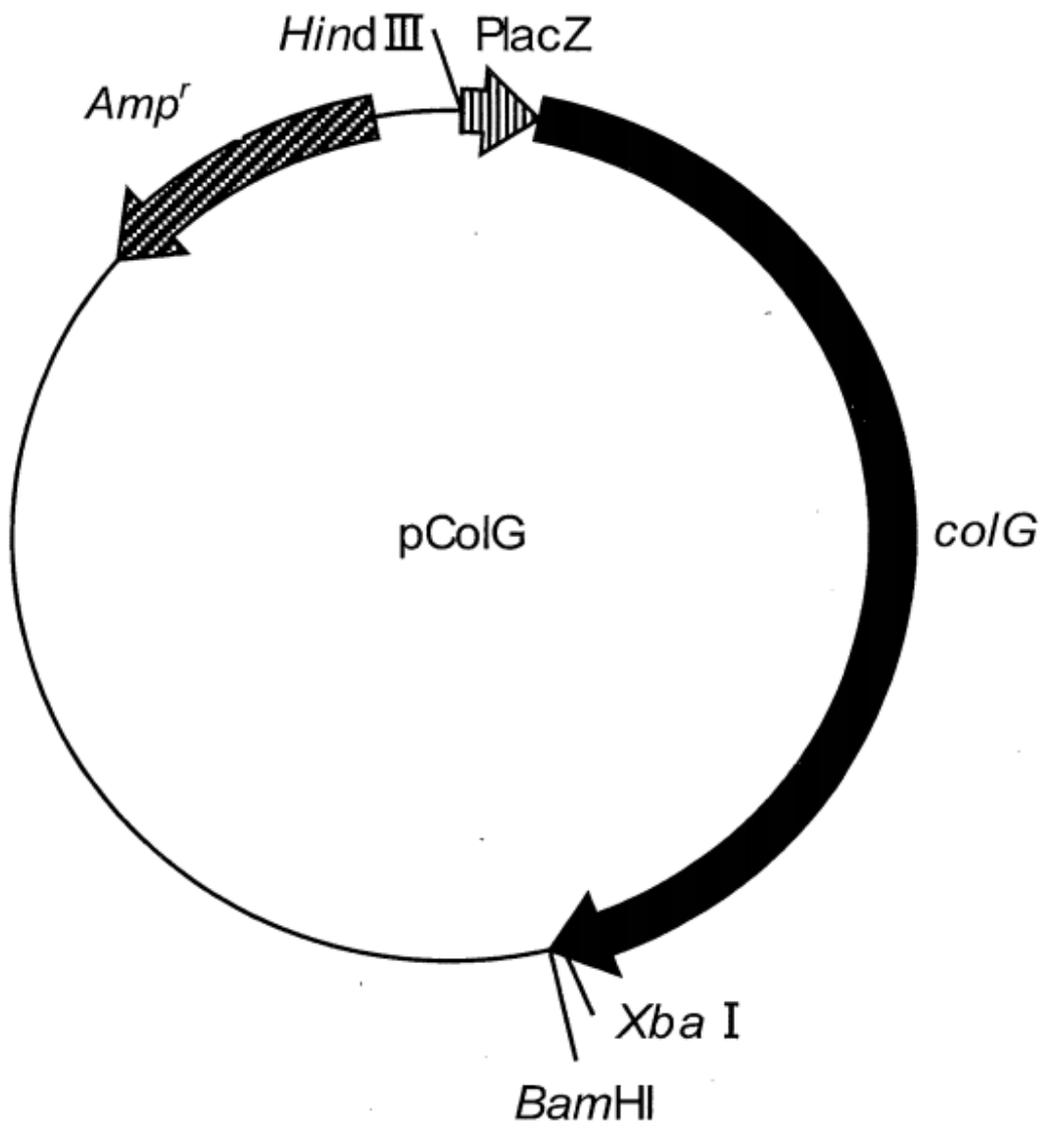
5. El método de producción de acuerdo con la reivindicación 1, en el que la colagenasa de fusión comprende una secuencia de aminoácidos de las posiciones -40 a 994 de la SEC ID N<sup>o</sup>: 7.

50 6. El método de producción de acuerdo con la reivindicación 5, en el que la colagenasa de fusión es una colagenasa de fusión a partir de la que se retira una secuencia de aminoácidos de las posiciones -40 a -1 de la SEC ID N<sup>o</sup>: 7.

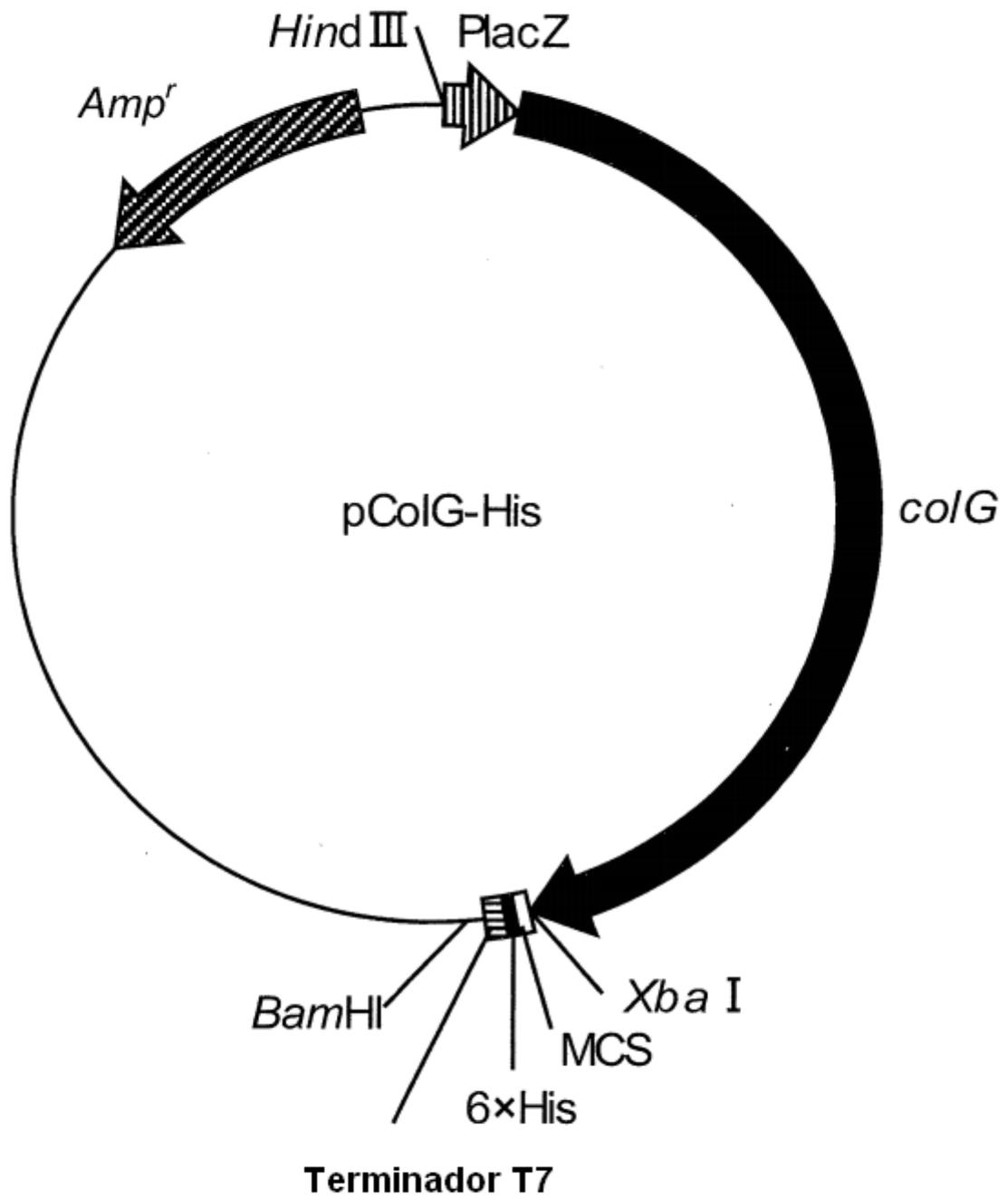
7. El método de producción de acuerdo con la reivindicación 1, en el que el ADN que codifica la colagenasa de fusión es un ADN que comprende una secuencia de bases de las posiciones 1 a 3396 de la SEC ID N<sup>o</sup>: 4.

55 8. El método de producción de acuerdo con la reivindicación 1, en el que el ADN que codifica la colagenasa de fusión es un ADN que comprende una secuencia de bases de las posiciones 1 a 3105 de la SEC ID N<sup>o</sup>: 8.

[Fig. 1]



[Fig. 2]



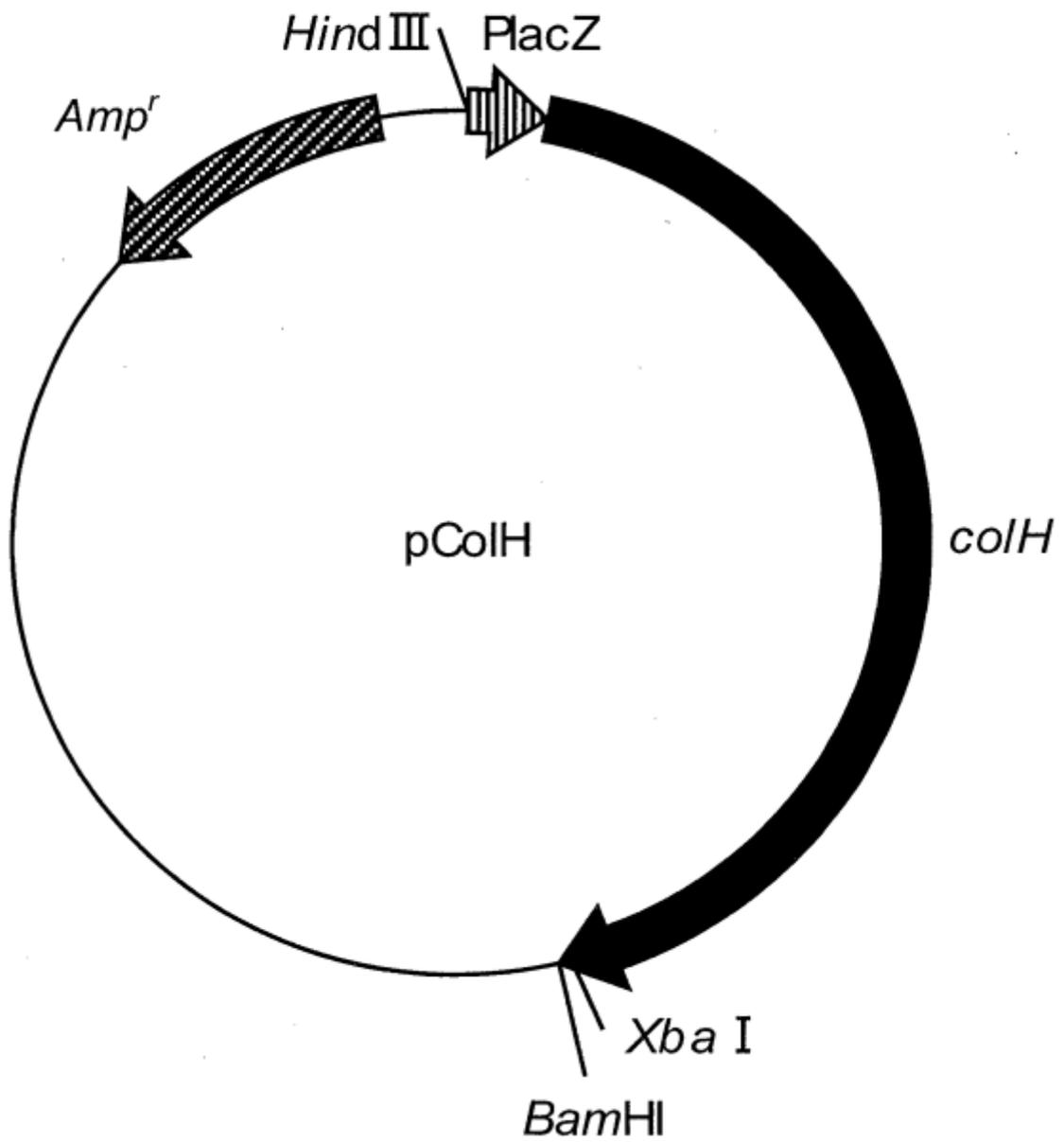
[Fig. 3]



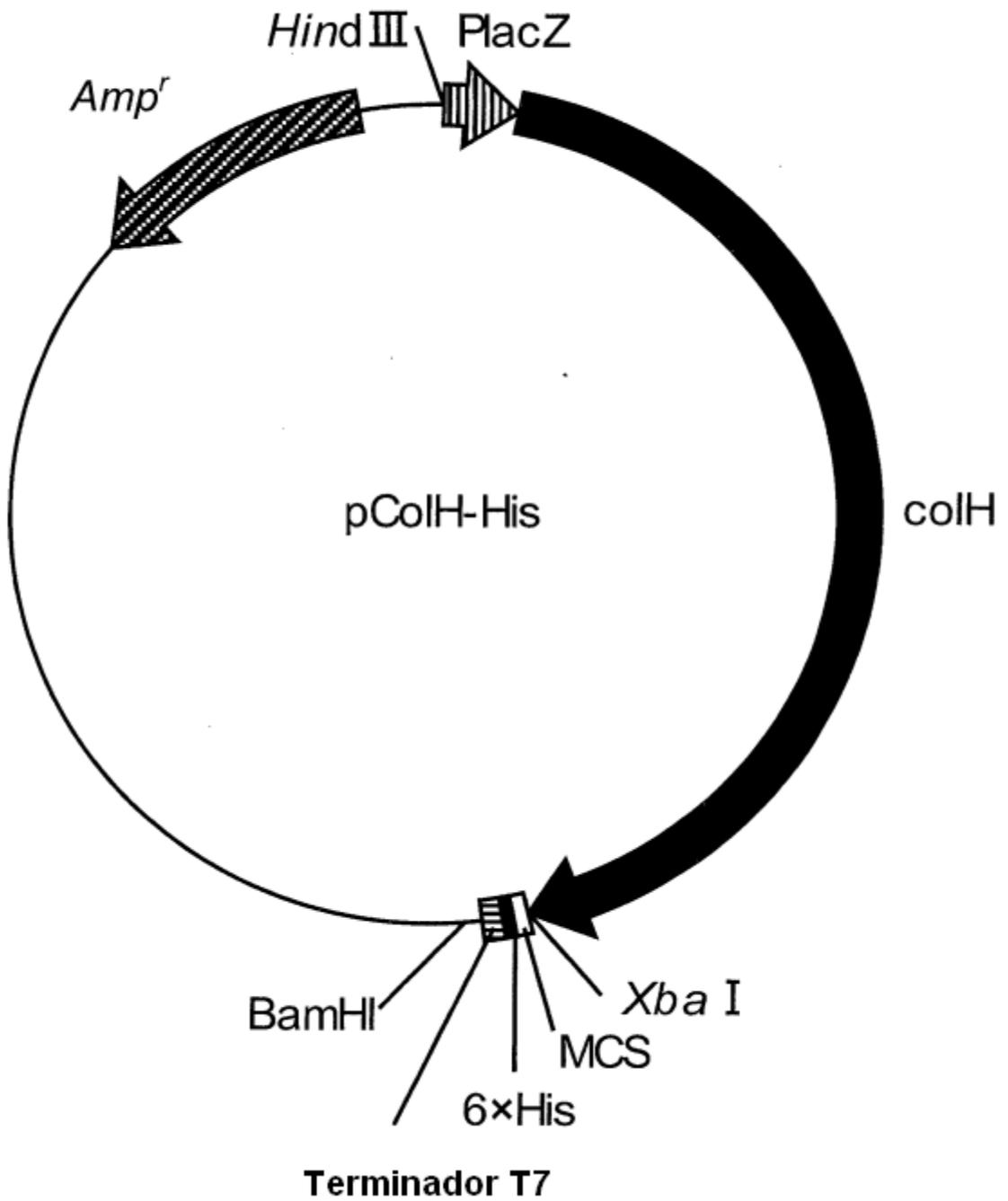
[Fig. 4]



[Fig. 5]



[Fig. 6]



[Fig. 7]



[Fig. 8]

