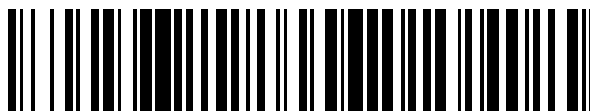


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 740 499**

51 Int. Cl.:

C12N 9/52

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **10.01.2013** **PCT/US2013/020940**

87 Fecha y número de publicación internacional: **18.07.2013** **WO13106510**

96 Fecha de presentación y número de la solicitud europea: **10.01.2013** **E 13736040 (0)**

97 Fecha y número de publicación de la concesión europea: **05.06.2019** **EP 2802652**

54 Título: **Enzima de Clostridium histolyticum**

30 Prioridad:

12.01.2012 US 201261585909 P

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

05.02.2020

73 Titular/es:

**ENDO GLOBAL VENTURES (100.0%)
22 Victoria Street, Canon's Court
Hamilton HM12, BM**

72 Inventor/es:

HERBER, WAYNE, K.

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 740 499 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

Enzima de *Clostridium histolyticum*

5 Antecedentes de la invención

El colágeno es el principal componente estructural de los organismos mamíferos y constituye una gran parte del contenido total de proteínas de la piel y otras partes del cuerpo del animal. En los seres humanos, es particularmente importante en los procesos de curación de heridas y envejecimiento natural. Diversos traumatismos de la piel, incluyendo quemaduras, cirugía e infección, se caracterizan por la acumulación de tejido fibroso rico en colágeno y por tener un contenido aumentado de proteoglucanos. Además del reemplazo del tejido normal que ha sido dañado o destruido, a veces se forman depósitos excesivos y desfigurantes de tejido nuevo durante el proceso de curación. El exceso de deposición de colágeno se ha atribuido a una alteración en el equilibrio entre la síntesis de colágeno y la degradación del colágeno.

Las enfermedades y afecciones asociadas a una deposición excesiva de colágeno y la acumulación errática de tejido fibroso rico en colágeno pueden denominarse "enfermedades mediadas por colágeno". La colagenasa, una enzima que tiene la capacidad específica de digerir colágeno, se ha utilizado para tratar una diversidad de enfermedades mediadas por colágeno, incluyendo, por ejemplo, contractura de Dupuytren, enfermedad de Peyronie, lipoma y capsulitis adhesiva. Las Patentes de los EE. UU. N.º 6.086.872 y 5.589.171 desvelan el uso de preparaciones de colagenasa en el tratamiento de la enfermedad de Dupuytren. La Patente de los EE. UU. N.º 6.022.539 desvela el uso de preparaciones de colagenasa en el tratamiento de la enfermedad de Peyronie. Las Patentes de los EE. UU. N.º 6.958.150 y 7.842.673 desvelan el uso de colagenasa para el tratamiento del lipoma. La Publicación de Solicitud de Patente de los EE. UU. 2006/020448A1 desvela el uso de colagenasa en el tratamiento de la capsulitis adhesiva.

El documento WO 2007/089851 desvela un medicamento que comprende una combinación de colagenasa I y colagenasa II purificadas de *Clostridium histolyticum*. El medicamento incluye colagenasa I y colagenasa II en una relación de aproximadamente 1 a 1, con una pureza de más de al menos el 95 %. También se desvelan procesos mejorados de fermentación y purificación para preparar el medicamento.

El documento WO 2011/073925 desvela un método para la producción de colagenasas recombinantes de *Clostridium histolyticum*. El método implica diseñar una secuencia nucleotídica que se optimiza para la expresión en *E. coli*.

La colagenasa para su uso en terapia puede obtenerse de una diversidad de fuentes incluyendo fuentes de mamíferos, fúngicas y bacterianas. Una fuente común de colagenasa en bruto proviene de un proceso de fermentación bacteriana, específicamente la fermentación de *Clostridium histolyticum* (*C. histolyticum*). La colagenasa en bruto obtenida de *C. histolyticum* puede purificarse usando cualquiera de una serie de técnicas cromatográficas.

Un inconveniente de la fermentación de bacterias es que se producirán diversas toxinas, que si están presentes en la composición terapéutica, serían perjudiciales para la salud del paciente. Por ejemplo, la fermentación de *C. histolyticum* da como resultado la síntesis de las toxinas hemolíticas alfa y épsilon, que pueden producir la lisis de eritrocitos (hemólisis), conduciendo potencialmente a una crisis hemolítica y una anemia hemolítica. Una crisis hemolítica se produce cuando hay una destrucción rápida de un gran número de eritrocitos junto con la incapacidad del cuerpo para reponer los eritrocitos lo suficientemente rápido para restablecer los niveles normales de eritrocitos.

Una crisis hemolítica provoca anemia hemolítica aguda (y con frecuencia grave) y puede provocar fatiga, dificultad para respirar, mareos, dolor de cabeza, frialdad en las manos y los pies, piel pálida, dolor en el pecho, ictericia, dolor en la parte superior del abdomen, úlceras y dolor en las piernas, reacciones graves a una transfusión de sangre, arritmias, agrandamiento del corazón e insuficiencia cardíaca. Con el fin de garantizar que la preparación terapéutica de colagenasa no contenga toxinas hemolíticas que puedan expresarse durante la fermentación de *C. histolyticum*, se presenta un método para liberar un fármaco antes de la administración a un paciente.

Como se ha mencionado anteriormente, la colagenasa para su uso en terapia puede obtenerse de una diversidad de fuentes tales como fuentes bacterianas (por ejemplo, de la fermentación de *C. histolyticum*). Sería útil desarrollar fuentes adicionales de colagenasa tales como formas recombinantes de enzimas colagenasa.

55 Sumario de la invención

En algunos aspectos, la presente invención se basa en el descubrimiento de secuencias polinucleotídicas mutadas que codifican colagenasa I y colagenasa II funcionales. La invención abarca, por tanto, ácido nucleico recombinante y polipéptidos que comprenden las secuencias polinucleotídicas o polipeptídicas novedosas. También se desvela un método para detectar la secreción de una toxina hemolítica mediante una cepa de producción bacteriana, en el que la cepa de producción produce una colagenasa, antes de la administración terapéutica de dicha colagenasa a un paciente y métodos de detección de la presencia de una toxina hemolítica en una composición de colagenasa.

En un primer aspecto, la invención se refiere a una molécula de ácido nucleico recombinante que comprende (a) un polinucleótido que tiene la secuencia de ácido nucleico de la SEQ ID NO: 1 (secuencia nucleotídica de colagenasa I) o el complemento de la SEQ ID NO: 1, y b) una secuencia reguladora heteróloga unida operativamente al

polinucleótido. También se desvela una molécula de ácido nucleico recombinante que consiste en un polinucleótido de la SEQ ID NO: 1. En una realización, la invención se refiere a una molécula de ácido nucleico recombinante que consiste en un polinucleótido de la SEQ ID NO: 1 y una secuencia reguladora heteróloga unida operativamente al polinucleótido.

En el presente documento también se desvela una molécula de ácido nucleico recombinante que comprende un polinucleótido que tiene la secuencia de la SEQ ID NO: 2 (secuencia nucleotídica de colagenasa II) o el complemento de la SEQ ID NO: 2. El ácido nucleico recombinante puede comprender adicionalmente un promotor heterólogo unido operativamente al polinucleótido. En el presente documento también se desvela una molécula de ácido nucleico recombinante que consiste en un polinucleótido de la SEQ ID NO: 2. En el presente documento también se desvela una molécula de ácido nucleico recombinante que consiste en un polinucleótido de la SEQ ID NO: 2 y una secuencia reguladora heteróloga unida operativamente al polinucleótido.

En un aspecto adicional, la invención proporciona un polipéptido recombinante codificado por la molécula de ácido nucleico recombinante del primer aspecto de la invención. En el presente documento también se desvelan polipéptidos recombinantes codificados por un ácido nucleico recombinante que comprende un polinucleótido que tiene la secuencia de la SEQ ID NO: 1 o la SEQ ID NO: 2.

En el presente documento también se desvela un casete de expresión que comprende un ácido nucleico recombinante, en el que el ácido nucleico comprende un polinucleótido que tiene la secuencia de la SEQ ID NO: 1 o la SEQ ID NO: 2.

En un aspecto adicional, la invención proporciona un vector que comprende la molécula de ácido nucleico de la invención. En el presente documento también se desvela un vector que comprende un ácido nucleico recombinante, en el que el ácido nucleico comprende un polinucleótido que tiene la secuencia de la SEQ ID NO: 1 o la SEQ ID NO: 2. En algunas realizaciones, el vector es un plásmido.

En un aspecto adicional, la invención proporciona una célula hospedadora recombinante que comprende el vector de la invención. También se desvela una célula hospedadora recombinante que comprende un vector o plásmido que comprende un polinucleótido que tiene la secuencia de la SEQ ID NO: 1 o la SEQ ID NO: 2. La invención también abarca un método de producción de colagenasa I que comprende cultivar la célula hospedadora de la invención en condiciones adecuadas para la expresión del ácido nucleico y recuperar la colagenasa I. En el presente documento también se desvela un método de producción de colagenasa I o colagenasa II que comprende cultivar la célula hospedadora en condiciones adecuadas para la expresión del ácido nucleico y recuperar la colagenasa I o colagenasa II. En el presente documento también se desvela una enzima colagenasa producida mediante el cultivo de la célula hospedadora recombinante.

En algunas realizaciones, la invención se refiere a una colagenasa I producida de forma recombinante que comprende la secuencia aminoacídica de la SEQ ID NO: 3 o la SEQ ID NO: 5. En el presente documento también se desvela una colagenasa II producida de forma recombinante que comprende la secuencia aminoacídica de la SEQ ID NO: 4 o una colagenasa II producida de forma recombinante que comprende la secuencia aminoacídica de la SEQ ID NO: 6.

En un aspecto adicional, la invención proporciona una composición farmacéutica que comprende un vehículo farmacéuticamente aceptable y el polipéptido recombinante de la invención. En el presente documento también se desvelan composiciones farmacéuticas que comprenden colagenasa I como se describe en el presente documento, colagenasa II como se describe en el presente documento o una combinación de las mismas. En determinados aspectos, la presente invención se refiere a una composición farmacéutica que comprende un vehículo farmacéuticamente aceptable y un polipéptido que comprende la secuencia amino de la SEQ ID NO: 3 o que comprende una combinación de un polipéptido que comprende la secuencia amino de la SEQ ID NO: 3 y un polipéptido que comprende la secuencia amino de la SEQ ID NO: 4. En el presente documento también se desvela una composición farmacéutica que comprende un vehículo farmacéuticamente aceptable y un polipéptido que comprende la secuencia aminoacídica de la SEQ ID NO: 4. En determinados aspectos adicionales, la presente invención se refiere a una composición farmacéutica que comprende un vehículo farmacéuticamente aceptable y un polipéptido que comprende la secuencia amino de la SEQ ID NO: 5 o que comprende una combinación de un polipéptido que comprende la secuencia amino de la SEQ ID NO: 5 y un polipéptido que comprende la secuencia amino de la SEQ ID NO: 6. En el presente documento también se desvela una composición farmacéutica que comprende un vehículo farmacéuticamente aceptable y un polipéptido que comprende la secuencia aminoacídica de la SEQ ID NO: 6. En el presente documento también se desvelan métodos de tratamiento de una enfermedad mediada por colágeno que comprende administrar una cantidad eficaz de colagenasa I, colagenasa II o una combinación de las mismas.

Como se ha mencionado anteriormente, en el presente documento se desvelan métodos de detección de la secreción de una toxina hemolítica mediante una cepa de producción bacteriana y métodos de detección de la presencia de una toxina hemolítica en una composición de colagenasa.

Una cepa bacteriana que produce colagenasa puede someterse a ensayo para determinar la producción de toxinas hemolíticas usando un ensayo de hemólisis. El ensayo de hemólisis puede realizarse usando un sustrato de agar

sangre.

En otro ejemplo, un producto de colagenasa se somete a ensayo para determinar la presencia de toxinas hemolíticas usando un ensayo de hemólisis. El ensayo de hemólisis puede realizarse usando un sustrato de agar sangre. El ensayo de hemólisis puede realizarse usando detección fotométrica de hemoglobina liberada. La ausencia de toxinas hemolíticas, según se determina mediante un ensayo de hemólisis o detección fotométrica, respaldaría la liberación del medicamento para la administración terapéutica.

Pueden analizarse diversas cepas de bacterias productoras de colagenasa para determinar la actividad hemolítica de acuerdo con un método que se desvela en el presente documento, en respaldo de la liberación de un medicamento de colagenasa para la administración terapéutica. Por ejemplo, miembros de los géneros *Actinobacillus*, *Actinomadura*, *Bacillus*, *Bacteroides*, *Bifidobacterium*, *Brucella*, *Capnocytophaga*, *Clostridium*, *Enterococcus*, *Escherichia*, *Eubacterium*, *Flavobacterium*, *Fusobacterium*, *Peptococcus*, *Peptostreptococcus*, *Porphyromonas*, *Prevotella*, *Proteus*, *Pseudomonas*, *Serratia*, *Staphylococcus*, *Streptomyces*, *Streptococcus*, *Treponema* y *Vibrio* pueden someterse a ensayo para determinar la actividad hemolítica de acuerdo con un método de la invención, en respaldo de la liberación de un medicamento de colagenasa para la administración terapéutica.

Un producto de colagenasa producido y purificado a partir de una cepa de bacterias productoras de colagenasa puede someterse a ensayo para determinar la actividad hemolítica de acuerdo con un método que se desvela en el presente documento, en respaldo de la liberación de un medicamento de colagenasa para la administración terapéutica. La cepa de producción puede seleccionarse entre, pero sin limitación, los géneros enumerados anteriormente o puede ser una cepa de *Escherichia coli* (*E. coli*), incluyendo formas de *E. coli* que se han transformado con formas recombinantes de colagenasa I y colagenasa II. La cepa de producción puede ser una cepa de *Clostridium perfringens* (*C. perfringens*) o una cepa de *C. histolyticum*.

La composición de colagenasa puede someterse a ensayo para determinar la actividad hemolítica de acuerdo con un método que se desvela en el presente documento, en el que la composición de colagenasa comprende una combinación de colagenasa I y colagenasa II purificadas de *C. histolyticum*. En el presente documento también se desvela un método de producción de un medicamento que consiste en colagenasa I y II de *C. histolyticum*, en el que dicho método comprende someter a ensayo una cepa de producción bacteriana para determinar la ausencia de una toxina hemolítica funcional y secretada de acuerdo con un método que se desvela en el presente documento.

En el presente documento también se desvela un método de purificación de una composición de colagenasa en bruto, en el que dicho método comprende purificar la composición mediante filtración y cromatografía en columna, seguido de confirmar la ausencia de una toxina hemolítica de acuerdo con un método que se describe en el presente documento.

En un aspecto adicional, la invención proporciona la composición farmacéutica de la invención para su uso en un método de tratamiento de un paciente que padece una afección mediada por colágeno. En el presente documento también se desvela un método de tratamiento de una afección mediada por colágeno en un paciente que lo necesite, en el que dicho método comprende administrar a dicho paciente una cantidad eficaz de un medicamento que comprende colagenasa, en el que la ausencia de una toxina hemolítica en dicho medicamento o en una cepa de producción bacteriana que produce dicha colagenasa se confirma de acuerdo con un método que se desvela en el presente documento antes de la administración de dicho medicamento a un paciente, y/o la formulación de la colagenasa en una composición farmacéutica.

También se describen kits para someter a ensayo la presencia o ausencia de toxinas hemolíticas en una muestra.

Breve descripción de los dibujos

Los anteriores y otros objetos, características y ventajas de la invención serán evidentes a partir de la siguiente descripción más particular de las realizaciones preferidas de la invención, como se ilustra en los dibujos adjuntos en los que caracteres con la misma referencia se refieren a las mismas partes a lo largo de las diversas vistas. Los dibujos no son necesariamente a escala, en su lugar, se pone énfasis en ilustrar los principios de la invención.

La FIG. 1 muestra una alineación de proteína de la toxina alfa de *Clostridium septicum* (*C. septicum*) con la supuesta toxina alfa de CLH_2834 y 2835 de *C. histolyticum*. La secuencia aminoacídica de la toxina alfa de *C. septicum* (SEQ ID NO: 7) es la secuencia superior en cada fila. El CLH_2834 y 2835 (SEQ ID NO: 8) de *C. histolyticum* es la secuencia inferior en cada fila. La secuencia subrayada, la secuencia sombreada es el extremo N-terminal de la toxina alfa de *C. septicum* madura. Los asteriscos sobre los aminoácidos muestran restos esenciales no conservados críticos para la funcionalidad (identifican la falta de apareamiento en la secuencia). El sombreado muestra restos esenciales conservados (confirma la identidad). La numeración de la secuencia se basa en la secuencia de *C. septicum*.

La FIG. 2 muestra placas de agar sangre de *C. septicum*. Las flechas indican la actividad beta hemolítica.

La FIG. 3 muestra la alineación de aminoácidos de la termolisina de *Bacillus proteolyticus* con la supuesta toxina delta de *C. histolyticum* CLH_2576. La secuencia superior en cada fila muestra la secuencia proteínica de la

termolisina de *Bacillus proteolyticus* (*B. proteolyticus*) (SEQ ID NO: 9). La secuencia inferior en cada fila es la secuencia de CLH_2576 de *C. histolyticum* (SEQ ID NO: 10). El sombreado de color verde muestra la región de proproteína. La numeración se basa en la secuencia de termolisina.

La FIG. 4 muestra la alineación de aminoácidos de prosequencia de termolisina de *B. proteolyticus* con la supuesta toxina delta de CLH_2576 de *C. histolyticum*. La secuencia superior en cada fila es la prosequencia de la proteína termolisina de *B. proteolyticus* (SEQ ID NO: 11). La secuencia inferior en cada fila es la prosequencia de CLH_2576 de *C. histolyticum* (SEQ ID NO: 12). Los asteriscos sobre los aminoácidos muestran los restos esenciales no conservados críticos para la funcionalidad (identifican la falta de apareamiento en la secuencia). El sombreado de color verde muestra los restos esenciales conservados (confirma la identidad). La numeración se basa en la secuencia de termolisina.

La FIG. 5 muestra la alineación de la proteína de secuencia madura de termolisina de *B. proteolyticus* con la supuesta toxina delta de CLH_2576 de *C. histolyticum*. La secuencia superior en cada fila es la secuencia madura de termolisina de *B. proteolyticus* (SEQ ID NO: 13). La secuencia inferior en cada fila es CLH_2576 de *C. histolyticum* (SEQ ID NO: 14). Los asteriscos sobre los aminoácidos muestran restos esenciales no conservados críticos para la funcionalidad (identifican la falta de apareamiento en la secuencia). El sombreado de color verde muestra restos esenciales conservados (confirma la identidad). La numeración se basa en la secuencia de la termolisina.

La FIG. 6 muestra la alineación de proteína de la perfringolisina de *C. perfringens* con la supuesta toxina épsilon de CLH_1920 de *C. histolyticum*. La secuencia superior en cada fila es la secuencia aminoacídica de la perfringolisina de *C. perfringens* (SEQ ID NO: 15). La secuencia inferior en cada fila es la secuencia aminoacídica de CLH_1920 de *C. histolyticum* (SEQ ID NO: 16). La estrella azul muestra el sitio de escisión de la peptidasa señal de la perfringolisina K43. Los asteriscos sobre los aminoácidos muestran restos esenciales no conservados críticos para la funcionalidad (identifican la falta de apareamiento en la secuencia). El sombreado de color verde muestra restos esenciales conservados (confirma la identidad). La numeración se basa en la secuencia de la perfringolisina.

La FIG. 7 muestra el fenotipo beta hemolítico de tetanolisina.

La FIG. 8 muestra la alineación de proteína de clostripaína de *C. histolyticum* con la supuesta toxina gamma de CLH_1861 de *C. histolyticum*. La secuencia superior en cada fila es la secuencia aminoacídica de clostripaína de *C. histolyticum* (SEQ ID NO: 17). La secuencia inferior en cada fila es la secuencia aminoacídica de CLH_1920 de *C. histolyticum* (SEQ ID NO: 18). Los asteriscos sobre los aminoácidos muestran restos esenciales no conservados críticos para la funcionalidad (identifican la falta de apareamiento en la secuencia). El sombreado de color verde muestra los restos esenciales conservados (confirma la identidad). La numeración basada en la secuencia X63673 de clostripaína.

La FIG. 9 muestra una comparación de la alineación de la secuencia aminoacídica traducida de colG y la secuencia aminoacídica de la SEQ ID NO: 3 (la secuencia aminoacídica traducida de CLH_1768 y 1769; la secuencia superior). Como se muestra en la Fig. 9, la proteína madura codificada por la secuencia aminoacídica de la SEQ ID NO: 3 difiere de la secuencia aminoacídica traducida de la secuencia aminoacídica de colG en tres aminoácidos. El extremo N-terminal de la proteína madura comienza en Ile 119 de la secuencia de la SEQ ID NO: 3. La secuencia aminoacídica de la proteína madura que comienza en Ile 119 de la SEQ ID NO: 3 es la SEQ ID NO: 5.

La FIG. 10 muestra una comparación de alineación de la secuencia aminoacídica traducida de colH y la SEQ ID NO: 4 (la secuencia aminoacídica traducida de CLH_2116; la secuencia inferior). Como se muestra en la Fig. 10, la proteína madura codificada por la secuencia aminoacídica de la SEQ ID NO: 4 difiere de la secuencia aminoacídica de colG traducida en ocho aminoácidos. El extremo N-terminal de la proteína madura comienza en Ala 31 en colG y en la SEQ ID NO: 4. La secuencia aminoacídica de la proteína madura que comienza en Ala 31 de la SEQ ID NO: 4 es la SEQ ID NO: 6.

La FIG. 11 muestra la secuencia nucleotídica de la SEQ ID NO: 1 (CLH_1768 y 1769; colagenasa I).

La FIG. 12 muestra la secuencia nucleotídica de la SEQ ID NO: 2 (CLH_2116; colagenasa II).

Las FIG. 13A y 13B muestran la secuencia aminoacídica y nucleotídica de la SEQ ID NO: 8 y ala SEQ ID NO: 21, respectivamente (CLH_2835 y CLH_2834; toxina alfa).

Las FIG. 14A y 14B muestran la secuencia aminoacídica y nucleotídica de la SEQ ID NO: 10 y la SEQ ID NO: 22, respectivamente (CLH_2576; toxina delta).

Las FIG. 15A y 15B muestran la secuencia aminoacídica y nucleotídica de la SEQ ID NO: 16 y la SEQ ID NO: 23, respectivamente (CLH_1920; toxina épsilon).

Las FIG. 16A y 16B muestran la secuencia aminoacídica y nucleotídica de la SEQ ID NO: 18 y la SEQ ID NO: 24, respectivamente (CLH_1861; toxina gamma).

La FIG. 17A muestra la secuencia aminoacídica de la SEQ ID NO: 3 (colG).

La FIG. 17B muestra la secuencia aminoacídica de la SEQ ID NO: 5.

La FIG. 18A muestra la secuencia aminoacídica de la SEQ ID NO: 4 (colH).

La FIG. 18B muestra la secuencia aminoacídica de la SEQ ID NO: 6.

Descripción detallada de la invención

A continuación se presenta una descripción de realizaciones preferidas de la invención.

Las palabras "un" o "una" pretenden abarcar uno o más, a menos que se especifique lo contrario. Por ejemplo, "una toxina hemolítica" se refiere a una o más toxinas hemolíticas.

La práctica de la presente invención empleará, a menos que se indique lo contrario, técnicas convencionales de cultivo celular, biología molecular, microbiología, biología celular e inmunología, que están bien dentro de las habilidades de la técnica. Dichas técnicas se explican por completo en la bibliografía. Véase, por ejemplo, Sambrook et al., 1989, "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor Laboratory Press; Ausubel et al. (1995), "Short Protocols in Molecular Biology", John Wiley and Sons; *Methods in Enzymology* (varios volúmenes); *Methods in Cell Biology* (varios volúmenes) y *Methods in Molecular Biology* (varios volúmenes).

A. Ácidos nucleicos y proteínas recombinantes

Una fuente importante de colagenasa es la fermentación de *C. histolyticum*. Una formulación inyectable que comprende colagenasa I y colagenasa II de *C. histolyticum* se comercializa con el nombre comercial XIAFLEX® y está aprobada por la Administración de Alimentos y Fármacos de los EE. UU. para el tratamiento de la contractura de Dupuytren. Se han descrito en la bibliografía secuencias aminoacídicas para colagenasa I y colagenasa II codificadas por los genes *colG* y *colH*, respectivamente. Por ejemplo, se describe *colG* en GenBank Acc. N.º D87215 y Matsushita et al. (1999), *Journal of Bacteriology* 181 (3): 923-933 y se ha descrito *colH* en GenBank Acc. N.º D29981 y Yoshihara et al. (1994), *Journal of Bacteriology* 176 (21): 6489-6496. La presente invención se basa parcialmente en el análisis de secuenciación de los genes que codifican colagenasa I y colagenasa II en una cepa de *C. histolyticum* (Clon 004 que se describe a continuación en los Ejemplos) que produce y secreta colagenasa I y colagenasa II funcionales. Se descubrió que las secuencias nucleotídicas de los genes que codifican la colagenasa I y la colagenasa II eran diferentes de las secuencias descritas en la bibliografía para *C. histolyticum* (por ejemplo, GenBank Acc. N.º D87125 y D29981) (SEQ ID NO: 19 y 20) (FIG. 9 y 10).

La colagenasa I y la colagenasa II son metaloproteasas y requieren cinc estrechamente unido y calcio débilmente unido para su actividad (Eddie L. Angleton y H. E. Van Wart, *Biochemistry* 1988, 27, 7406 - 7412). La colagenasa I y la colagenasa II tienen una amplia especificidad hacia todos los tipos de colágeno (Steinbrink, D; Bond, M y Van Wart, H; (1985), JBC, 260 págs. 2771-2776). La colagenasa I y la colagenasa II digieren el colágeno hidrolizando la región de hélice triple del colágeno en condiciones fisiológicas (Steinbrink, D; Bond, M y Van Wart, H; (1985), JBC, 260 págs. 2771-2776). Aunque cada colagenasa muestra una especificidad diferente (por ejemplo, cada una tiene una secuencia amino preferida diferente para la escisión), conjuntamente, tienen actividad sinérgica hacia el colágeno (Mandl, I., (1964), *Biochemistry*, 3: p. 1737-1741; Vos-Scheperkeuter, GH, (1997), *Cell Transplantation*, 6: págs. 403-412).

La invención abarca una molécula de ácido nucleico recombinante que comprende (a) un polinucleótido de la SEQ ID NO: 1 o el complemento de la SEQ ID NO: 1 y (b) una secuencia reguladora heteróloga unida operativamente al polinucleótido. También se desvela una molécula de ácido nucleico recombinante que comprende o que consiste en un polinucleótido de la SEQ ID NO: 2 o el complemento de la SEQ ID NO: 2. El ácido nucleico recombinante puede comprender adicionalmente un promotor heterólogo unido operativamente al polinucleótido.

La invención también abarca polipéptidos recombinantes codificados por la molécula de ácido nucleico recombinante del primer aspecto de la invención. En algunas realizaciones, el polipéptido recombinante comprende o consiste en la secuencia aminoacídica de la SEQ ID NO: 3. También se desvela un polipéptido recombinante que comprende o que consiste en la secuencia aminoacídica de la SEQ ID NO: 4. En realizaciones adicionales, el polipéptido recombinante comprende o consiste en la secuencia aminoacídica de la SEQ ID NO: 5 (la proteína colagenasa I madura, que comienza en Ile 119 de la SEQ ID NO: 3 en la FIG. 9). También se desvela un polipéptido recombinante que comprende o que consiste en la SEQ ID NO: 6 (la proteína colagenasa II madura, que comienza en Ala 31 de la SEQ ID NO: 4 en la FIG. 10).

En el presente documento se desvela un ácido nucleico recombinante que codifica un polipéptido que comprende o que consiste en la secuencia aminoacídica de la SEQ ID NO: 3 o la SEQ ID NO: 4 o la SEQ ID NO: 5 o la SEQ ID NO: 6. También se desvela un ácido nucleico recombinante que comprende una secuencia nucleotídica que codifica un polipéptido de la secuencia aminoacídica SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 o SEQ ID NO: 6.

Un ácido nucleico recombinante es una molécula de ácido nucleico que contiene, además de una secuencia polinucleotídica que se describe en el presente documento (por ejemplo, la secuencia polinucleotídica de la SEQ ID NO: 1 o la SEQ ID NO: 2), una secuencia nucleotídica codificante o no codificante heteróloga adicional. El término "heterólogo" significa que el polinucleótido se origina a partir de una especie diferente o de la misma especie, sin embargo, de otra ubicación en el genoma que dicha secuencia nucleotídica añadida. Los polipéptidos o proteínas recombinantes se refieren a polipéptidos o proteínas producidos usando técnicas recombinantes, por ejemplo, las proteínas o polipéptidos producidos a partir de células transformadas mediante una construcción de ácido nucleico exógeno que codifica el polipéptido o proteína deseado.

En el presente documento se desvelan ácidos nucleicos que comprenden la secuencia polinucleotídica de la SEQ ID NO: 1 o la SEQ ID NO: 2, en el que dicho ácido nucleico está unido operativamente a una secuencia reguladora. También se desvelan ácidos nucleicos que comprenden un polinucleótido que codifica un polipéptido que comprende la secuencia aminoacídica de la SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 o SEQ ID NO: 6, en el que dicho ácido nucleico está unido operativamente a una secuencia reguladora. Las secuencias reguladoras incluyen las secuencias

reguladoras que dirigen la expresión constitutiva de una secuencia nucleotídica en muchos tipos de células hospedadoras y/o aquellas que dirigen la expresión de la secuencia nucleotídica solo en determinadas células hospedadoras (por ejemplo, secuencias reguladoras específicas de tejido). Son ejemplos no limitantes de secuencias reguladoras promotores y potenciadores. Las secuencias reguladoras también incluyen otros elementos de control de la expresión, por ejemplo, los que se describen en Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Un ácido nucleico está "unido operativamente" a una secuencia reguladora cuando la molécula de ácido nucleico está unida a la secuencia reguladora de una manera que permite la expresión de la secuencia de ácido nucleico.

Una molécula de ácido nucleico que se describe en el presente documento puede fusionarse adicionalmente a una secuencia marcadora, por ejemplo, una secuencia que codifica un polipéptido para ayudar en el aislamiento o la purificación del polipéptido. Dichas secuencias incluyen, pero sin limitación, las que codifican una proteína de fusión de glutatión-S-transferasa (GST), las que codifican un marcador polipeptídico hemaglutina A (HA) de la gripe y las que codifican el péptido hexa-histidina, tal como le marcador proporcionado en un vector pQE (Qiagen, Inc.). En determinadas realizaciones, la invención se refiere a un polipéptido que comprende una secuencia aminoacídica de la SEQ ID NO: 3 o la SEQ ID NO: 5, en el que dicho polipéptido se fusiona con una secuencia aminoacídica marcadora, a condición de que el polipéptido esté codificado por la molécula de ácido nucleico recombinante de acuerdo con el primer aspecto de la invención. En el presente documento se desvela un polipéptido que comprende una secuencia aminoacídica de la SEQ ID NO: 4 o la SEQ ID NO: 6, en el que dicho polipéptido se fusiona con una secuencia aminoacídica marcadora.

En el presente documento se desvela un ácido nucleico que es una variante de la secuencia nucleotídica de la SEQ ID NO: 1 o la SEQ ID NO: 2. Un ácido nucleico variante es un ácido nucleico que incluye una sustitución, adición o supresión de nucleótidos con respecto a la secuencia nucleotídica de la SEQ ID NO: 1 o la SEQ ID NO: 2. En algunos aspectos, la variante es un ácido nucleico que codifica secuencias aminoacídicas idénticas o sustancialmente idénticas a las secuencias nucleotídicas de la SEQ ID NO: 1 o la SEQ ID NO: 2. Como comprenderá el experto en la materia, debido a la degeneración del código genético, varias secuencias de ácido nucleico diferentes pueden codificar una proteína dada. Por ejemplo, los codones GCA, GCC, GCG y GCU codifican cada uno el aminoácido alanina. Por tanto, por ejemplo, en cada posición en la que un codón especifica un aminoácido específico, el codón puede cambiarse a cualquiera de los codones correspondientes que codifican el mismo aminoácido sin alterar la secuencia aminoacídica del polipéptido codificado. Un experto en la materia comprenderá que cada codón en una secuencia nucleotídica (excepto AUG, que es el único codón para la metionina, y TGG, que por lo general es el único codón para el triptófano) puede modificarse para producir una molécula funcionalmente idéntica.

En determinadas realizaciones, la invención se refiere a un polipéptido que comprende o que consiste en la secuencia aminoacídica de la SEQ ID NO: 3 o la SEQ ID NO: 5, en el que uno o más aminoácidos se han suprimido o añadido, en el que el polipéptido posee la actividad de degradar o lisar colágeno, a condición de que el polipéptido esté codificado por la molécula de ácido nucleico recombinante de acuerdo con el primer aspecto de la invención. También se desvela un polipéptido que comprende o que consiste en la secuencia aminoacídica de la SEQ ID NO: 4 o la SEQ ID NO: 6, en el que uno o más aminoácidos se han suprimido o añadido, en el que el polipéptido posee la actividad de degradar o lisar colágeno. Los polipéptidos que se desvelan en el presente documento pueden comprender o consistir en la secuencia aminoacídica de la SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 o SEQ ID NO: 6, en la que uno o más restos de aminoácidos se han reemplazado por un resto de aminoácido diferente, en la que el polipéptido posee la actividad de degradar o lisar colágeno y en la que el polipéptido comprende o consiste en una secuencia aminoacídica que es diferente de las secuencias aminoacídicas de GenBank Acc. N.º D87125 (SEQ ID NO: 19) y D29981 (SEQ ID NO: 20). En determinados aspectos, cuando se reemplaza un aminoácido, el reemplazo es un cambio conservador de aminoácidos. Un cambio conservador de aminoácidos es, por ejemplo, la sustitución de un aminoácido no polar por otro aminoácido no polar o la sustitución de un aminoácido polar por otro aminoácido polar o la sustitución de un aminoácido cargado positivamente por otro aminoácido cargado positivamente y similares. Por ejemplo, los aminoácidos no polares (hidrófobos) incluyen alanina, leucina, isoleucina, valina, prolina, fenilalanina, triptófano y metionina; los aminoácidos polares neutros incluyen glicina, serina, treonina, cisteína, tirosina, asparagina y glutamina; los aminoácidos cargados positivamente (básicos) incluyen arginina, lisina e histidina; y los aminoácidos cargados negativamente (ácidos) incluyen ácido aspártico y ácido glutámico.

Un ácido nucleico aislado y un polipéptido aislado no están en la forma o el entorno en los que existen en la naturaleza. Por ejemplo, un ácido nucleico aislado es uno que está separado de los nucleótidos que normalmente flanquea la molécula de ácido nucleico en la naturaleza. Los ácidos nucleicos recombinantes y los ácidos nucleicos recombinantes dentro de un vector también son un ejemplo de un ácido nucleico aislado. Además, las moléculas de ácido nucleico aisladas incluyen moléculas de ácido nucleico recombinantes en células hospedadoras heterólogas, así como moléculas de ácido nucleico parcialmente o sustancialmente purificadas en solución.

Como se describe con mayor detalle a continuación, la invención también abarca células hospedadoras recombinantes, tales como células bacterianas, células fúngicas, células vegetales, células de insecto, células aviares, células de anfibios y células de mamíferos, que comprenden el vector de la invención.

Un casete de expresión es una secuencia nucleotídica que es capaz de afectar a la expresión de un gen estructural

(es decir, una secuencia de codificación de proteínas, tal como una colagenasa de la invención) en un hospedador compatible con dichas secuencias. Los casetes de expresión pueden incluir un promotor unido operativamente con la secuencia codificante del polipéptido; y, opcionalmente, con otras secuencias, por ejemplo, señales de terminación de la transcripción. También pueden usarse factores adicionales necesarios o útiles para efectuar la expresión, por ejemplo, potenciadores.

La invención también se refiere a vectores que comprenden un ácido nucleico de la invención. También se desvela un vector que comprende que el ácido nucleico es la SEQ ID NO: 1 o la SEQ ID NO: 2 o un complemento de las mismas. También se desvela un vector que comprende un ácido nucleico que codifica un polipéptido que tiene la secuencia aminoacídica de la SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 o SEQ ID NO: 6. Un "vector" es una molécula de ácido nucleico capaz de transportar otro ácido nucleico al que se ha unido. Un ejemplo no limitante de un vector es un plásmido que es un ADN bicatenario circular en el que puede ligarse un segmento de ADN adicional. Otro ejemplo de un vector es un vector vírico, en el que un segmento de ADN adicional se liga en el genoma vírico. Determinados vectores son capaces de replicarse de manera autónoma en una célula hospedadora en la que se introducen (por ejemplo, vectores bacterianos que tienen un origen de replicación bacteriano y vectores episómicos de mamífero). Otros vectores (por ejemplo, vectores no episómicos de mamífero) se integran en el genoma de una célula hospedadora tras su introducción en la célula hospedadora y, de este modo, se replican junto con el genoma del hospedador. Los vectores de expresión son capaces de dirigir la expresión de genes a los que están unidos operativamente. Dichos vectores de expresión incluyen, por ejemplo, plásmidos. La invención abarca otros vectores de expresión, tales como vectores víricos (por ejemplo, retrovirus defectuosos para la replicación, adenovirus y virus adeno-asociados) que son capaces de dirigir la expresión génica. Como apreciarán los expertos en la materia, el diseño del vector de expresión depende de varios factores, tales como la elección de la célula hospedadora que se ha de transformar, el nivel de expresión de la proteína deseada y similares. Los vectores de expresión incluyen una o más secuencias reguladoras que se seleccionan basándose en la célula hospedadora que se ha de usar para la expresión. Como se ha mencionado anteriormente, la secuencia reguladora se une operativamente al ácido nucleico que se ha de expresar, por ejemplo, un ácido nucleico de la invención. En algunas realizaciones, la secuencia reguladora es una secuencia reguladora nativa de la célula hospedadora transformada. Un vector de expresión puede comprender uno o más marcadores seleccionables, incluyendo, pero sin limitación, el gen que codifica la dihidrofolato reductasa y los genes que confieren resistencia a neomicina, tetraciclina, ampicilina, cloranfenicol, kanamicina y resistencia a estreptomycin.

Las células hospedadoras procariotas y eucariotas pueden transfectarse mediante los vectores que se describen en el presente documento. Las células hospedadoras que pueden transfectarse con los vectores de la presente invención incluyen, pero sin limitación, células bacterianas tales como *E. coli* (por ejemplo, cepas de *E. coli* K12), *Streptomyces*, *Pseudomonas*, *Serratia marcescens* y *Salmonella typhimurium*, células de insecto (baculovirus), incluyendo *Drosophila*, células fúngicas, tales como células de levadura, células vegetales y células de mamíferos, tales como timocitos, células de ovario de hámster chino (CHO), células COS y células de *Lactococcus lactis*. En algunas realizaciones, la célula hospedadora es una célula bacteriana. En otra realización más, la célula hospedadora es una cepa de *E. coli*. En una realización adicional más, la célula hospedadora es una célula de *Lactococcus lactis*. Se han descrito métodos para la producción de polipéptidos recombinantes en bacterias *Lactococcus lactis*, por ejemplo, en la Patente de los EE. UU. N.º 7.358.067. En una realización, la célula hospedadora es *Lactococcus lactis* y el ácido nucleico comprende la secuencia nucleotídica de la SEQ ID NO: 1 unida operativamente al promotor regulable por pH P170 y derivados del mismo. También se desvela una célula hospedadora de *Lactococcus lactis* en la que el ácido nucleico comprende la secuencia nucleotídica de la SEQ ID NO: 2 unida operativamente al promotor regulable por pH P170 y derivados del mismo. El promotor P170 y derivados del mismo se han descrito en detalle en los documentos WO 94/16086 y WO 98/10079.

La ligadura de la molécula de ácido nucleico en una construcción génica, tal como un vector de expresión y la transformación o transfección en células hospedadoras, ya sean eucariotas (levaduras, de ave, de insecto, vegetales o de mamíferos) o procariotas (células bacterianas), son procedimientos convencionales. Un vector que se describe en el presente documento puede introducirse en células procariotas o eucariotas usando técnicas convencionales de transformación o transfección, incluyendo, pero sin limitación, coprecipitación de fosfato de calcio o cloruro de calcio, transfección mediada por DEAE-dextrano, lipofección o electroporación. Los polipéptidos de la presente invención pueden aislarse o purificarse (por ejemplo, hasta homogeneidad) del cultivo celular recombinante mediante una diversidad de procesos.

La invención abarca métodos de producción de una colagenasa I funcional o una combinación de colagenasa I y II que comprende cultivar una célula hospedadora transformada o transfectada con un vector que comprende un ácido nucleico de la invención. También se desvelan métodos de producción de una colagenasa II funcional que comprende cultivar una célula hospedadora transformada o transfectada con un vector que comprende un ácido nucleico que se desvela en el presente documento. El método comprende adicionalmente aislar el polipéptido del medio o la célula hospedadora. Una colagenasa funcional es un polipéptido que tiene una actividad biológica de una colagenasa de origen natural, por ejemplo, una colagenasa que posee la capacidad de degradar colágeno.

El polipéptido puede aislarse mediante métodos incluyendo, pero sin limitación, cromatografía de intercambio aniónico o catiónico, precipitación en etanol, cromatografía de afinidad y cromatografía de líquidos de alta resolución (HPLC) o

una combinación de cualquiera de los mismos. El método particular utilizado dependerá de las propiedades del polipéptido y la selección de la célula hospedadora; los métodos apropiados serán fácilmente evidentes para los expertos en la materia.

En el presente documento se desvela un método de producción de collagenasa I o collagenasa II, comprendiendo dicho método las etapas de (i) construir una bacteria recombinante que comprende la secuencia nucleotídica de la SEQ ID NO: 1 o la SEQ ID NO: 2 o el polinucleótido que codifica la secuencia polipeptídica de la SEQ ID NO: 3, 4, 5 o 6 unida operativamente a una secuencia reguladora apropiada; (ii) cultivar dicha bacteria recombinante en condiciones adecuadas para expresar el gen y (iii) recoger de la bacteria recombinante, la collagenasa I o la collagenasa II. La collagenasa I y la collagenasa II pueden purificarse mediante una diversidad de métodos conocidos por los expertos en la materia, incluyendo cromatografía de afinidad de ligando de colorante, cromatografía de afinidad con heparina, precipitación en sulfato de amonio, cromatografía de hidroxilapatita, cromatografía de exclusión molecular, cromatografía de intercambio iónico y cromatografía de quelación de metales. La collagenasa I y la collagenasa II pueden purificarse mediante filtración y cromatografía en columna y las collagenasas purificadas I y II se combinan en una relación de aproximadamente 1 a 1 usando métodos descritos en la Patente de los EE. UU. N.º 7.811.250.

Los ejemplos de enfermedades mediadas por colágeno que pueden tratarse mediante las composiciones (que comprenden collagenasa I, collagenasa II o una combinación de las mismas codificadas por los ácidos nucleicos que se describen en el presente documento y/o que comprenden las secuencias aminoacídicas de la SEQ ID NO: 3 y/o la SEQ ID NO: 4) y los métodos que se desvelan en el presente documento incluyen, pero sin limitación, enfermedad de Dupuytren, enfermedad de Peyronie, hombro congelado (capsulitis adhesiva), queloides, cicatrices hipertróficas, cicatrices deprimidas, como las que son resultado del acné inflamatorio; adherencias postquirúrgicas, acné común, lipomas y afecciones desfigurantes tales como arrugas, formación de celulitis y fibrosis neoplásica. Las Patentes de los EE. UU. N.º 6.086.872 y 5.589.171 desvelan el uso de preparaciones de collagenasa en el tratamiento de la enfermedad de Dupuytren. La Patente de los EE. UU. N.º 6.022.539 desvela el uso de preparaciones de collagenasa en el tratamiento de la enfermedad de Peyronie.

Además de su uso en el tratamiento de enfermedades mediadas por colágeno, una composición que comprende un polipéptido recombinante que se describe en el presente documento también es útil para la disociación de tejido en células individuales y grupos de células ya que es útil en una amplia diversidad de aplicaciones de laboratorio, de diagnóstico y terapéuticas. Estas aplicaciones implican el aislamiento de muchos tipos de células para diversos usos, incluidas las células endoteliales microvasculares para la siembra de injertos vasculares sintéticos de diámetro pequeño, hepatocitos para terapia génica, detección de toxicología de fármacos y dispositivos de asistencia hepática extracorpórea, condrocitos para la regeneración del cartílago e islotes de Langerhans para el tratamiento de la diabetes mellitus dependiente de insulina. El tratamiento enzimático trabaja para fragmentar proteínas de la matriz extracelular y proteínas que mantienen el contacto de célula a célula. Puesto que el colágeno es el principal componente proteico de la ultraestructura tisular, la enzima collagenasa se ha usado con frecuencia para lograr la desintegración tisular deseada. En general, la composición que se desvela en el presente documento es útil para cualquier aplicación en la que se desee la retirada de células o la modificación de una matriz extracelular.

La invención abarca una composición farmacéutica que comprende un vehículo farmacéuticamente aceptable y el polipéptido recombinante de la invención. También se desvelan composiciones farmacéuticas que comprenden un vehículo farmacéuticamente aceptable y collagenasa I y/o collagenasa II producidas de acuerdo con un método que se describe en el presente documento. Las composiciones farmacéuticas pueden comprender collagenasa I que comprende o que consiste en la secuencia aminoacídica de la SEQ ID NO: 3 o la SEQ ID NO: 5. Las composiciones farmacéuticas que se desvelan en el presente documento pueden comprender collagenasa II que comprende o que consiste en la secuencia aminoacídica de la SEQ ID NO: 4 o la SEQ ID NO: 6. En el presente documento se desvela una composición farmacéutica que comprende un vehículo farmacéuticamente aceptable y una collagenasa I y una collagenasa II como se describe en el presente documento. En el presente documento se desvela una composición farmacéutica que comprende un vehículo farmacéuticamente aceptable y la collagenasa I y la collagenasa II a una relación de masa 1:1. La composición farmacéutica de la presente invención comprende una cantidad eficaz de una collagenasa formulada junto con uno o más vehículos o excipientes farmacéuticamente aceptables.

B. Métodos de detección de la presencia de una toxina hemolítica

En el presente documento también se desvelan métodos de detección de la presencia de una toxina hemolítica en una fermentación bacteriana, en los que la fermentación bacteriana produce una collagenasa. En el presente documento también se desvela un método para liberar un medicamento de collagenasa antes de la administración terapéutica de dicha sustancia farmacológica de collagenasa a un paciente que comprende detectar la presencia de una toxina hemolítica en la cepa de producción del medicamento. La expresión "cepa de producción de medicamentos", "cepa de producción", "cepa de producción de collagenasa" y "cepa de producción bacteriana" se usan indistintamente y se refieren a una cepa bacteriana de la que se obtiene una collagenasa. En el presente documento también se desvela un método para liberar un medicamento de collagenasa antes de la administración terapéutica de dicho medicamento de collagenasa a un paciente, que comprende detectar la presencia de una toxina hemolítica en el medicamento.

Como se usa en el presente documento, la frase "liberar un medicamento de colagenasa" significa confirmar la ausencia de una toxina hemolítica en el medicamento de colagenasa. Se entiende que las expresiones "fármaco", "medicamento" o "composición de colagenasa" pueden usarse indistintamente. Como también se usan en el presente documento, las expresiones "hemolisina" y "toxina hemolítica" se usan indistintamente y se refieren a una toxina que es responsable de la lisis de un eritrocito.

Se ha descubierto que la cepa de producción de colagenasa y el medicamento pueden someterse a ensayo para detectar la presencia o ausencia de actividad hemolítica, garantizando que la sustancia farmacológica de colagenasa proporciona una actividad enzimática altamente reproducible y óptima y un efecto terapéutico superior, disminuyendo al mismo tiempo el potencial de efectos secundarios. Se proporcionan métodos para someter a ensayo la cepa de producción o el medicamento para determinar la secreción o presencia de una toxina hemolítica funcional que puede estar copresente con la colagenasa en el medicamento. En el presente documento también se desvela un método de ensayo de una muestra de ensayo para detectar la presencia de una toxina hemolítica, en el que la muestra de ensayo comprende una cepa de producción bacteriana o una composición de colagenasa, que comprende incubar la muestra de ensayo con eritrocitos, seguido de la detección de lisis de eritrocitos.

Se describen métodos específicos para detectar la lisis de eritrocitos en la bibliografía, incluyendo, por ejemplo, 1) Ryan KJ y Ray CG. *Principles of laboratory diagnosis*. En *Sherris medical microbiology: an introduction to infectious diseases*. Ryan KJ, Ray CG y Sherris JC (eds.) McGraw-Hill Professional, 2004; 229-260; y 2) Eschbach E et al. *Improved erythrocyte lysis assay in microtitre plates for sensitive detection and efficient measurement of hemolytic compounds from ichthyotoxic algae*. *Journal of Applied Toxicology* 21, 513-519 (2001).

El método puede comprender incubar muestras de una cepa de producción de colagenasa, una colagenasa parcialmente purificada aislada de una cepa de producción de colagenasa o un medicamento de colagenasa en un sustrato de agar sangre y observar el agar sangre para determinar zonas de aclaramiento después del período de incubación, en el que una zona de aclaramiento indica la lisis de eritrocitos. Si se sometió a ensayo la cepa de producto bacteriano, la lisis de eritrocitos indica la secreción de una toxina hemolítica funcional de la cepa de producción bacteriana. Si se sometió a ensayo una colagenasa parcialmente purificada o un medicamento de colagenasa, la lisis de eritrocitos indica la presencia de una toxina hemolítica funcional en la colagenasa parcialmente purificada o en el medicamento de colagenasa. La cepa de producción puede ser una cepa de *C. histolyticum*. La ausencia de una zona de aclaramiento indica la ausencia de una toxina hemolítica. La ausencia observada de zonas de aclaramiento indica o confirma la ausencia de toxinas hemolíticas en la cepa de producción de colagenasa, en la colagenasa parcialmente purificada o en el medicamento de colagenasa, y permite que el medicamento se libere para la administración terapéutica.

El método puede comprender incubar eritrocitos con extractos tomados de una cepa de producción de colagenasa, o con una colagenasa parcialmente purificada aislada de una cepa de producción de colagenasa, o con un medicamento de colagenasa, seguido de un análisis fotométrico de la mezcla de incubación para la lisis de eritrocitos como se indica mediante la aparición de hemoglobina en la mezcla de incubación. Una toxina hemolítica lisará los eritrocitos, liberando hemoglobina en la muestra incubada. La detección fotométrica de la hemoglobina puede proporcionar un ensayo sensible para la presencia de toxinas hemolíticas. En un aspecto, los eritrocitos se incuban con extractos tomados de una cepa de producción de colagenasa o con una colagenasa parcialmente purificada aislada de una cepa de producción de colagenasa, o con un medicamento de colagenasa, y después analizando fotométricamente los extractos para detectar la presencia de hemoglobina a una longitud de onda de 540 nm. En otro aspecto, el análisis fotométrico se realiza a una longitud de onda de 414 nm. En otro aspecto más, la incubación y el análisis fotométrico pueden realizarse usando placas de microtitulación. La ausencia de hemoglobina y, por tanto, la ausencia de toxinas hemolíticas, permitiría la liberación del medicamento para la administración terapéutica a un paciente.

Las toxinas hemolíticas que se encuentran en *C. histolyticum* pertenecen a dos familias diferentes de hemolisinas: hemolisinas similares a aerolisina y hemolisinas lábiles a oxígeno. Las hemolisinas similares a aerolisina son sintetizadas por la bacteria como preproteínas inactivas que se secretan en el ambiente extracelular como protoxinas inactivas. Las protoxinas inactivas se unirán a los receptores en una membrana celular diana, por ejemplo, receptores en un eritrocito donde las protoxinas son escindidas en sus estructuras activas por proteasas. Una vez activadas, las toxinas se oligomerizan sobre la superficie celular en un complejo de preporos, seguido de la inserción de un barril beta en la membrana celular diana. El barril beta forma un poro en la membrana, permitiendo el influjo rápido de iones de calcio en la célula, con consecuencias tóxicas para la célula. La toxina alfa de *C. histolyticum* es muy probablemente una hemolisina similar a aerolisina, ya que se ha descubierto que comparte una homología significativa con la toxina alfa de *Clostridium septicum*, que es un miembro de la familia de toxinas similares a aerolisina y que posee actividad hemolítica (véase, por ejemplo, Ejemplo 1 a continuación).

La toxina épsilon de *C. histolyticum* y la tetanolisina de *Clostridium tetani* (*C. tetani*), se han descrito como hemolisinas lábiles al oxígeno [Hatheway CL. *Clin Microbiol Rev* 3 (1): 66-98 (1990)]. Se ha descubierto que la toxina épsilon de *C. histolyticum* comparte homología con la tetanolisina, que es un miembro de toxinas activadas por tiol, de barril beta, formadoras de poros con afinidad por el colesterol. Dichas proteínas son parte de una familia de citolisinas dependientes del colesterol (CDC). Estas proteínas son secretadas por la bacteria en el ambiente extracelular como proteínas monoméricas hidrosolubles donde se unen a membranas celulares diana, mediadas por la unión a colesterol.

La toxina se oligomeriza después sobre la superficie de la membrana para formar arcos y estructuras similares a anillo que son responsables de la citólisis. Se sabe que la toxina épsilon de *C. histolyticum* es una hemolisina lábil al oxígeno y es serológicamente similar a las hemolisinas lábiles al oxígeno producidas por otras cepas de *Clostridium*, tales como *C. tetani*, *C. novyi* y *C. septicum*.

En el presente documento también se desvela un método de detección de la presencia de toxina alfa de *C. histolyticum* en una cepa de producción bacteriana usando un ensayo que se describe en el presente documento. En el presente documento también se desvela un método de detección de la presencia de toxina alfa de *C. histolyticum* en un medicamento. En el presente documento también se desvela un método de detección de la presencia de toxina épsilon de *C. histolyticum* en una cepa de producción bacteriana. En el presente documento también se desvela un método de detección de la presencia de toxina épsilon de *C. histolyticum* en un medicamento. En el presente documento también se desvela un método de detección de la presencia de toxina alfa de *C. histolyticum* y toxina épsilon en una cepa de producción bacteriana. En el presente documento también se desvela un método de detección de la presencia de toxina alfa de *C. histolyticum* y toxina épsilon en un medicamento.

En el presente documento también se desvela un método de producción de un medicamento que consiste en colagenasa I y colagenasa II, en el que las colageninasas I y II se obtienen de *C. histolyticum* y en el que el método comprende las etapas de fermentar una cepa de *C. histolyticum* en la que la ausencia de una toxina hemolítica funcional secretada se ha confirmado mediante incubación de la cepa de producción con eritrocitos en condiciones adecuadas para la lisis de los eritrocitos por una toxina hemolítica, en el que la lisis de eritrocitos indica la secreción de una toxina hemolítica y en el que la ausencia de lisis de los eritrocitos indica la ausencia de una toxina hemolítica. En el presente documento también se desvela un método de producción de un medicamento que consiste en colagenasa I y colagenasa II, en el que las colageninasas I y II se obtienen de *C. histolyticum* y en el que el método comprende las etapas de confirmar la ausencia de una toxina hemolítica secretada funcional en el medicamento mediante incubación del medicamento con eritrocitos en condiciones adecuadas para la lisis de eritrocitos por una toxina hemolítica, en el que la lisis de eritrocitos indica la secreción de una toxina hemolítica y en el que la ausencia de lisis de los eritrocitos indica la ausencia de una toxina hemolítica.

En el presente documento también se desvelan métodos de purificación de una composición de colagenasa en bruto que comprende purificar la composición mediante filtración y cromatografía en columna, seguido de confirmar la ausencia de una toxina hemolítica mediante incubación de una muestra de la composición purificada con eritrocitos en condiciones adecuadas para la lisis de eritrocitos por una toxina hemolítica, en el que la lisis de eritrocitos indica la secreción de una toxina hemolítica y en el que la ausencia de lisis de los eritrocitos indica la ausencia de una toxina hemolítica.

Como se ha mencionado anteriormente, varias enfermedades y afecciones se asocian al depósito excesivo de colágeno y la acumulación errática de tejido fibroso rico en colágeno y pueden tratarse con medicamentos de colágeno. Dichas enfermedades y afecciones se denominan colectivamente en el presente documento "enfermedades mediadas por colágeno". En el presente documento se desvela un método de tratamiento de una enfermedad mediada por colágeno en un paciente que lo necesite, en el que la composición que comprende colagenasa se administra a dicho paciente y en el que, antes de dicha administración, dicha composición o cepa de producción bacteriana se somete a ensayo para determinar la presencia o ausencia de toxinas hemolíticas usando un método que se describe en el presente documento. Los ejemplos de afecciones mediadas por colágeno que pueden tratarse mediante las composiciones y métodos que se describen en el presente documento incluyen, pero sin limitación: enfermedad de Dupuytren; enfermedad de Peyronie; hombro congelado (capsulitis adhesiva), queloides; cicatrices hipertróficas; cicatrices deprimidas tales como las que son resultado del acné inflamatorio; adherencias postquirúrgicas; acné común; lipomas y afecciones desfigurantes tales como arrugas, formación de celulitis y fibrosis neoplásica. En determinados aspectos, la composición sometida a ensayo se administra a un paciente para tratar las enfermedades de Peyronie o Dupuytren o capsulitis adhesiva.

Con respecto a la cepa de producción que puede someterse a ensayo de acuerdo con un método que se describe en el presente documento, se sabe, por ejemplo, que la colagenasa es expresada por bacterias que son miembros de los géneros *Actinobacillus*, *Actinomadura*, *Bacillus*, *Bacteroides*, *Bifidobacterium*, *Brucella*, *Capnocytophaga*, *Clostridium*, *Enterococcus*, *Escherichia*, *Eubacterium*, *Flavobacterium*, *Fusobacterium*, *Peptococcus*, *Peptostreptococcus*, *Porphyromonas*, *Prevotella*, *Proteus*, *Pseudomonas*, *Serratia*, *Staphylococcus*, *Streptomyces*, *Streptococcus*, *Treponema* y *Vibrio*. La cepa de producción puede seleccionarse entre los géneros enumerados anteriormente. La cepa de producción puede ser una cepa de *E. coli*, incluyendo formas de *E. coli* que se han transformado con formas recombinantes de colagenasa I y colagenasa II. Se prefiere más que la cepa de producción sea una cepa de *C. perfringens*. Se prefiere mucho más que la cepa de producción sea una cepa de *C. histolyticum* (*C. his*).

La cepa de producción puede producir una composición de colagenasa que comprenda una mezcla de colagenasa I y colagenasa II. La cepa de producción que puede usarse para producir una mezcla de colagenasa I y colagenasa II es *C. histolyticum*. El medicamento de colagenasa puede comprender una mezcla de colagenasa I y colagenasa II de *C. histolyticum* altamente purificadas en una relación de masa de aproximadamente 1 a 1.

También se presentan kits para someter a ensayo la presencia de hemolisinas en una muestra, en los que una

hemolisina es una sustancia que provoca la lisis de eritrocitos. Los kits permiten la identificación de sustancias de ensayo que son hemolíticas o que contienen hemolisinas. Las sustancias de ensayo incluyen, pero sin limitación, sustancias químicas, biológicas y emisoras de radiación. El kit puede comprender materiales para someter a ensayo la presencia de hemolisinas en una muestra de ensayo incluyendo, por ejemplo, un kit que comprende eritrocitos y materiales de ensayo relacionados. El kit puede comprender una placa de Petri compuesta de agar sangre, un control positivo y un control negativo compuesto por una cepa bacteriana en la que los genes hemolíticos están mutados o inactivados y en el que no se producen proteínas hemolíticas funcionales. El kit puede comprender eritrocitos, placas de microtitulación, un control positivo y un control negativo compuesto por el medicamento.

Como se comprenderá, los kits y métodos inventivos pueden usarse para detectar la presencia o ausencia de hemolisinas en composiciones de colagenasa, en las que la colagenasa se obtiene de una bacteria.

La colagenasa en bruto obtenida de *C. histolyticum* puede purificarse mediante una diversidad de métodos conocidos por los expertos en la materia, incluyendo cromatografía de afinidad de ligando de colorante, cromatografía de afinidad con heparina, precipitación en sulfato de amonio, cromatografía de hidroxilapatita, cromatografía de exclusión molecular, cromatografía de intercambio iónico y cromatografía de quelación de metales. La colagenasa en bruto y parcialmente purificada está disponible en el mercado de muchas fuentes incluyendo Advance Biofactures Corp., Lynbrook, Nueva York. Los métodos de purificación de colagenasa en bruto obtenida de *C. histolyticum* también se describen en la Patente de los EE. UU. N.º 7.811.560 El procedimiento de purificación puede comprender las etapas de: a) filtrar la recolección en bruto a través de un filtro de cápsula de intercambio aniónico MUSTANG Q; b) añadir sulfato de amonio; preferentemente a una concentración final de 1 M; c) filtrar la recolección en bruto; preferentemente a través de un filtro de 0,45 µm; d) someter el filtrado a través de una columna HIC; preferentemente una fenil Sepharose 6FF (sub bajo); e) añadir leupeptina al filtrado; preferentemente a una concentración final de 0,2 mM al producto eluido post HIC; f) retirar el sulfato de amonio y mantener la leupeptina para una correcta unión de colagenasa I y colagenasa II con intercambio de tampón por TFF; preferentemente con intercambio de tampón por TFF; g) filtrar la mezcla de la etapa; (f) preferentemente a través de un filtro de 0,45 µm; h) separar colagenasa I y colagenasa II usando Q-Sepharose HP; i) preparar la concentración de TFF y la formulación para colagenasa I y colagenasa II por separado; en el que TFF es una filtración de flujo tangencial usando 10 y/o 30 K MWCO (corte de peso molecular) PES o RC-polieter-sulfona o membranas de filtro de celulosa regeneradas (TFF proporciona un medio para retener y concentrar proteínas seleccionadas e intercambiar proteínas de una solución tampón en otra); y j) filtrar a través de un sistema de filtración de 0,2 µm.

C. Toxinas alfa, beta, delta, épsilon y gamma de C. histolyticum

Las secuencias aminoacídicas de las toxinas alfa, delta y épsilon de *C. histolyticum* Clon 004 se muestran en las Figuras y son la SEQ ID NO: 8, SEQ ID NO: 12 y SEQ ID NO: 16, respectivamente. Las secuencias nucleotídicas de las toxinas alfa, delta y épsilon de *C. histolyticum* Clon 004 también se muestran en las Figuras y son la SEQ ID NO: 21, SEQ ID NO: 22 y SEQ ID NO: 23, respectivamente. Cada una de las secuencias aminoacídicas de la SEQ ID NO: 8, SEQ ID NO: 12, SEQ ID NO: 16 tienen características de secuencia que hacen que estas proteínas sean no funcionales y/o no se secreten.

Para la toxina gamma (clostripaína), solo existen tres diferencias de aminoácidos en comparación con la proteína modelo (véase la sección de Ejemplos) y ninguno de los restos de aminoácidos que se encuentra que difieren en la toxina gamma de *C. histolyticum* Clon 004 ha sido identificado como esencial para la actividad. Por tanto, se prevé que la toxina gamma de *C. histolyticum* Clon 004 (que tiene la secuencia aminoacídica de la SEQ ID NO: 18) se secreta y es funcional. La secuencia nucleotídica de la toxina gamma de *C. histolyticum* Clon 004 es la SEQ ID NO: 24.

Como se ha mencionado anteriormente, las toxinas beta que tienen secuencias aminoacídicas de la SEQ ID NO: 3 y la SEQ ID NO: 4 son completamente funcionales.

Como se comprenderá, pueden introducirse una o más mutaciones (por ejemplo, supresión o adición de uno o más restos de aminoácidos o restos de ácidos nucleicos) en las secuencias nucleotídicas y/o aminoacídicas de las toxinas alfa, beta, épsilon o gamma de *C. histolyticum* (SEQ ID NO: 8, SEQ ID NO: 12, SEQ ID NO: 16 y SEQ ID NO: 18, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23 y SEQ ID NO: 24). Puede introducirse una o más mutaciones con el fin de mejorar o perjudicar la actividad, función, producción y/o secreción de la toxina. Puede introducirse una mutación que haga que las toxinas alfa, beta y/o épsilon sean funcionales y/o se secreten. La secuencia de la toxina gamma (SEQ ID NO: 18) puede mutarse para hacer que la proteína no sea funcional y/o no se secrete.

En el presente documento también se desvelan métodos de producción de anticuerpos contra *C. histolyticum* o una toxina de *C. histolyticum* que comprende administrar a un sujeto una cantidad eficaz de una composición que comprende una proteína o péptido, en los que dicha proteína o péptido comprende una secuencia aminoacídica seleccionada entre el grupo que consiste en SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 y SEQ ID NO: 6, o un fragmento o variante de las mismas o una combinación de cualquiera de las mismas. Además, en el presente documento se desvelan métodos de estimulación de una respuesta inmunitaria a una toxina de *C. histolyticum* que comprende administrar a un sujeto

una cantidad eficaz de una composición que comprende una proteína o péptido que comprende una secuencia aminoacídica seleccionada entre el grupo que consiste en SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 y SEQ ID NO: 6, o un fragmento o variante de las mismas o una combinación de cualquiera de las mismas. En el presente documento se desvela una vacuna que comprende una cantidad eficaz de una proteína o péptido que comprende una secuencia aminoacídica seleccionada entre el grupo que consiste en SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 y SEQ ID NO: 6, o un fragmento o variante de las mismas o una combinación de cualquiera de las mismas. La proteína o péptido que comprende una secuencia aminoacídica seleccionada entre el grupo que consiste en SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 y SEQ ID NO: 6, o un fragmento o variante de las mismas, puede ser producido por una cepa de *C. histolyticum* o puede ser una proteína o péptido recombinante.

D. Composiciones farmacéuticas que comprenden collagenasa y métodos de tratamiento

En el presente documento se desvelan composiciones farmacéuticas que comprenden las secuencias de proteínas y proteínas recombinantes y, también, composiciones farmacéuticas que comprenden un medicamento de collagenasa sometido a ensayo de acuerdo con los métodos que se describen en el presente documento. Como se usa en el presente documento, la expresión "vehículo o excipiente farmacéuticamente aceptable" significa una carga, diluyente, material de encapsulación o adyuvante de formulación de cualquier tipo no tóxicos, inertes, sólidos, semisólidos o líquidos. "Tratamiento" o "tratamiento" incluye la administración de las composiciones, compuestos o agentes de aspectos de la presente invención para prevenir o retrasar la aparición de los síntomas, complicaciones o indicios bioquímicos de una enfermedad, aliviar o mejorar los síntomas o detener o inhibir el desarrollo adicional de la enfermedad, afección o trastorno. Una "cantidad terapéuticamente eficaz" o una "cantidad eficaz" es una cantidad que, sola o en combinación con uno o más de otros agentes activos, puede controlar, reducir, inhibir, mejorar, prevenir o afectar de otro modo a uno o más síntomas de una enfermedad o afección que se ha de tratar. En el contexto de producción de una respuesta inmunitaria o en la preparación de una vacuna, una "cantidad eficaz" abarca una cantidad eficaz para producir una respuesta inmunitaria, incluyendo la generación de anticuerpos contra un antígeno.

Algunos ejemplos de materiales que pueden servir como vehículos farmacéuticamente aceptables son azúcares tales como lactosa, glucosa y sacarosa; almidones, tales como almidón de maíz y almidón de patata; celulosa y derivados de la misma, tales como carboximetilcelulosa de sodio, etilcelulosa y acetato de celulosa; tragacanto en polvo; malta; gelatina; talco; glicoles, tales como propilenglicol; ésteres, tales como oleato de etilo y laurato de etilo; agar; agentes tamponantes, tales como hidróxido de magnesio e hidróxido de aluminio; ácido algínico; agua apirógena; solución salina isotónica; solución de Ringer; alcohol etílico y soluciones de tampón fosfato, así como otros lubricantes compatibles no tóxicos, tales como lauril sulfato de sodio y estearato de magnesio, así como agentes colorantes, agentes de liberación, agentes de recubrimiento, agentes perfumantes, conservantes y antioxidantes también pueden estar presentes en la composición, de acuerdo con el criterio del formulador.

Las composiciones de collagenasa también pueden prepararse mezclando un número específico de unidades de actividad o masas específicas de las enzimas preferentemente purificadas. La actividad de la collagenasa puede medirse por la capacidad de la enzima para hidrolizar el péptido sintético o el sustrato de colágeno. Los expertos en la materia reconocerán que los ensayos enzimáticos distintos de los desvelados en el presente documento también pueden usarse para definir y preparar composiciones enzimáticas funcionalmente equivalentes. La actividad collagenasa puede describirse, por ejemplo, en unidades SRC. Una unidad SRC solubilizará colágeno de cola de rata en material de reacción de ninhidrina equivalente a 1 nanomol de leucina por minuto, a 25 °C y pH 7,4. En determinadas realizaciones de la presente invención, la actividad collagenasa se describe en unidades ABC. Este ensayo de potencia de collagenasa se basa en la digestión de colágeno no desnaturalizado (de tendón bovino) a pH 7,2 y 37 °C durante 20-24 horas. El número de enlaces peptídicos escindidos se mide mediante reacción con ninhidrina. Los grupos amino liberados por un control de digestión con tripsina se restan. Una unidad ABC neta de collagenasa solubilizará material reactivo con ninhidrina equivalente a 1,09 nanomoles de leucina por minuto. 1 unidad SRC equivale a aproximadamente 6,3 unidades ABC.

En determinados aspectos, la sustancia farmacológica para la collagenasa inyectable consiste en dos collagenasas microbianas, denominadas collagenasa AUX I y collagenasa ABC I y collagenasa AUX II y collagenasa ABC II. Se entiende que la expresiones "Collagenasa I", "ABC I", "AUX I", "collagenasa AUX I" y "collagenasa ABC I" significan lo mismo y pueden usarse indistintamente. De forma similar, las expresiones "Collagenasa II", "ABC II", "AUX II", "collagenasa AUX II" y "collagenasa ABC II" se refieren a la misma enzima y también pueden usarse indistintamente. Estas collagenasas son secretadas por células bacterianas. Se aíslan y purifican del sobrenadante de cultivo de *C. histolyticum* mediante métodos cromatográficos. Ambas collagenasas son proteasas especiales y comparten el mismo número de EC (E.C 3.4.24.3).

La collagenasa AUX I tiene una única cadena polipeptídica que consiste en aproximadamente 1000 aminoácidos con un peso molecular de 115 kDa. La collagenasa AUX II tiene también una única cadena polipeptídica que consiste en aproximadamente 1000 aminoácidos con un peso molecular de 110 kDa.

La sustancia farmacológica (concentrado de colagenasa) puede tener una relación de masa de aproximadamente 1 a 1 para la colagenasa AUX I y AUX II. El concentrado de colagenasa puede tener un coeficiente de extinción de 1,528.

Las composiciones farmacéuticas de la presente invención pueden administrarse por vía parenteral, por vía tópica o a través de un depósito implantado. El término "parenteral", como se usa en el presente documento, incluye la inyección subcutánea, intracutánea, intravenosa, intramuscular, intraarticular, intraarterial, intrasinovial, intraesternal, intratecal, intralesional e intracraneal o técnicas de infusión. En una realización preferida, la composición se inyecta en el tejido afectado. En el caso de las enfermedades de Peyronie o de Dupuytren o la capsulitis adhesiva, la composición se inyecta en el cordón de la mano o en la placa de Peyronies. La expresión "administración local" se define en el presente documento para que abarque dicha inyección directa en el tejido afectado. En determinados aspectos, la composición farmacéutica de la invención es una formulación inyectable. En determinados aspectos adicionales, la composición farmacéutica es una formulación tópica.

Asimismo, dependiendo del tratamiento, pueden obtenerse mejores resultados, en algunas circunstancias, inmovilizando el sitio de inyección después de la administración de la composición farmacéutica. Por ejemplo, el sitio de administración (por ejemplo, la mano), puede inmovilizarse durante 4 horas o más.

Pueden formularse preparaciones inyectables, por ejemplo, suspensiones acuosas u oleaginosas inyectables estériles, de acuerdo con la técnica conocida usando agentes dispersantes o humectantes y agentes de suspensión adecuados. La preparación inyectable estéril también puede ser una solución, suspensión o emulsión inyectable estéril en un diluyente o disolvente no tóxico parenteralmente aceptable, por ejemplo, como una solución en 1,3-butanodiol. Entre los vehículos y disolventes aceptables que pueden emplearse están el agua, la solución de Ringer, U.S.P. y solución de cloruro de sodio isotónica. Además, se emplean convencionalmente aceites no volátiles estériles como disolvente o medio de suspensión. Para este fin, puede emplearse cualquier aceite suave no volátil, incluyendo mono o diglicéridos sintéticos. Además, se usan ácidos grasos, tales como ácido oleico, en la preparación de inyectables.

Las formulaciones inyectables pueden esterilizarse, por ejemplo, mediante la filtración a través de un filtro de retención de bacterias o mediante la incorporación de agentes esterilizantes en forma de composiciones sólidas estériles que pueden disolverse o dispersarse en agua estéril u otro medio inyectable estéril antes de su uso. Las soluciones estériles también pueden liofilizarse para su uso posterior.

En algunas realizaciones, la composición que comprende colagenasa es una composición inyectable liofilizada formulada con sacarosa, Tris a un nivel de pH de aproximadamente 8,0. Generalmente, se incluye una fuente de calcio en la formulación, tal como cloruro de calcio.

Las formas de dosificación para la administración tópica o transdérmica de una composición farmacéutica de la presente invención incluyen pomadas, pastas, cremas, lociones, geles, polvos, soluciones, pulverizaciones, inhaladores o parches. El componente activo se premezcla en condiciones estériles con un vehículo farmacéuticamente aceptable y cualquier conservante o tampón necesario, según se requiera.

Las pomadas, pastas, cremas y geles pueden contener, además de un polipéptido de la presente invención, excipientes tales como grasas animales y vegetales, aceites, ceras, parafinas, almidón, tragacanto, derivados de celulosa, polietilenglicoles, siliconas, bentonitas, ácido silícico, talco y óxido de cinc, o mezclas de los mismos.

Los polvos y pulverizaciones pueden contener, además de los polipéptidos de la presente invención, excipientes tales como lactosa, talco, ácido silícico, hidróxido de aluminio, silicatos cálcicos y polvo de poliamida, o mezclas de estas sustancias. Las pulverizaciones pueden contener adicionalmente propulsores habituales tales como clorofluorohidrocarburos.

Los parches transdérmicos tienen la ventaja adicional de proporcionar una entrega controlada de un agente activo al cuerpo. Dichas formas de dosificación pueden prepararse disolviendo o dispensando el agente activo en el medio apropiado. Los potenciadores de la absorción también pueden usarse para aumentar el flujo del polipéptido a través de la piel. La velocidad puede controlarse proporcionando una membrana de control de la velocidad o dispersando el polipéptido de la invención en una matriz polimérica o gel.

La administración terapéutica del producto farmacéutico puede ser por vía parenteral, por vía tópica o a través de un depósito implantado. El término parenteral como se usa en el presente documento incluye la inyección subcutánea, intracutánea, intravenosa, intramuscular, intraarticular, intraarterial, intrasinovial, intraesternal, intratecal, intralesional e intracraneal o técnicas de infusión. La expresión "administración local" se define en el presente documento para que abarque dicha inyección directa. En una realización, la administración terapéutica de la composición farmacéutica es por inyección.

La administración terapéutica del producto farmacéutico en formas de dosificación para la administración tópica o transdérmica incluye pomadas, pastas, cremas, lociones, geles, polvos, soluciones, pulverizaciones, inhaladores o parches. El componente activo se premezcla en condiciones estériles con un vehículo farmacéuticamente aceptable y cualquier conservante o tampón necesario, según se requiera.

Las pomadas, pastas, cremas y geles pueden contener, además de un compuesto activo del medicamento, excipientes tales como grasas animales y vegetales, aceites, ceras, parafinas, almidón, tragacanto, derivados de celulosa, polietilenglicoles, siliconas, bentonitas, ácido silícico, talco y óxido de cinc, o mezclas de los mismos.

Los polvos y pulverizaciones pueden contener, además de los compuestos del medicamento, excipientes tales como lactosa, talco, ácido silícico, hidróxido de aluminio, silicatos cálcicos y polvo de poliamida, o mezclas de estas sustancias. Las pulverizaciones pueden contener adicionalmente propulsores habituales tales como clorofluorohidrocarburos.

Los parches transdérmicos tienen la ventaja añadida de proporcionar una entrega controlada de un compuesto al cuerpo. Dichas formas de dosificación pueden prepararse disolviendo o dispensando el compuesto en el medio adecuado. También pueden usarse potenciadores de la absorción para aumentar el flujo del compuesto a través de la piel. La velocidad puede controlarse proporcionando una membrana de control de la velocidad o dispersando el compuesto en una matriz polimérica o gel.

La invención se comprenderá mejor en relación con los siguientes ejemplos, que están destinados únicamente a ilustrar y no limitar el alcance de la invención. Serán evidentes para los expertos en la materia diversos cambios y modificaciones de las realizaciones desveladas y dichos cambios y modificaciones pueden hacerse sin apartarse del alcance de las reivindicaciones adjuntas.

Ejemplos

Ejemplo 1: Secuenciación del genoma de *C. histolyticum* y análisis de secuencias de toxinas

La escasez de estudios científicos relacionados con el *C. histolyticum* alfa, delta y épsilon (toxinas α , δ y ϵ) ha dado como resultado un conocimiento limitado acerca de la estructura proteínica de estas toxinas. Para abordar este déficit de conocimiento, se emprendió una iniciativa de secuenciación del genoma para comprender mejor el organismo de producción con un enfoque particular en la identificación de supuestos genes de toxinas. Como consecuencia de este esfuerzo, el genoma completo de la cepa de producción de Colagenasa de *Clostridium Histolyticum* (Clon 004) (Auxilium Product Operation, Malvern, PA) no se ha generado hasta recientemente, lo que representa, aparentemente, la primera vez que se ha publicado la secuencia del genoma de cualquier cepa de *C. histolyticum*.

Hubo tres etapas fundamentales implicadas en el proyecto de secuenciación del genoma. En primer lugar, se extrajo ADN genómico de un cultivo de Clon 004 y se envió a Creative Genomics para su secuenciación (Shirley, NY, EE. UU.). La secuencia del genoma de *C. histolyticum* Clon 004 se obtuvo usando métodos convencionales de la industria. En segundo lugar, los resultados obtenidos de la secuencia del genoma se analizaron usando métodos convencionales de bioinformática (análisis BLAST) con el fin de consultar la información de la secuencia contra las bases de datos de la secuencia del genoma. Esta segunda etapa dio como resultado la asignación de información de proteínas para cada gen de *C. histolyticum* que se identificó. El uso de dos bases de datos garantizó una evaluación exhaustiva, pero también sirvió como una segunda fuente para verificar la asignación de proteínas. La tercera etapa en el proyecto fue un análisis comparativo de la supuesta secuencia de toxina de *C. histolyticum* con la proteína asignada automáticamente por el análisis BLAST.

i. *C. Secuenciación del genoma de C. histolyticum e identificación de proteínas modelo*

Se enviaron muestras de ADN genómico aislado de una expansión de *C. histolyticum* (CLH) WCB derivada del Clon 004 a Creative Genomics (Shirley, NY, EUA) para la secuenciación del genoma. Creative Genomics empleó métodos convencionales utilizados para la determinación de secuencias de muestras de ADN genómico enviadas por clientes. La secuencia del genoma se generó a partir del sistema Roche/454 GS-FLX con química de titanio (secuenciación de fragmentos) acompañada con Illumina/Solexa Genome Analyzer. Se empleó ANI 3730xl para lograr el acabado del genoma mediante el andadura de cebadores. La secuencia completa del genoma de 2.842.906 pares de bases con un contenido de GC del 29,44 % se completó y estos valores eran típicos del tamaño del genoma y el contenido de GC obtenido para otros genomas clostridiales. A cada uno de los 2.887 marcos de lectura abiertos (ORF, por sus siglas en inglés) identificados se le asignó un número de CLH único. Cada uno de los supuestos 2.887 genes se investigó adicionalmente usando el análisis BLAST de las bases de datos GenBank y SwissProt, dando como resultado la asignación tentativa de los loci para toxinas beta, gamma, alfa y épsilon. Los resultados de la evaluación inicial se presentan en la Tabla 1. Por tanto, la asignación de proteínas modelo se completó como resultado de un análisis automatizado a través de una búsqueda exhaustiva de dos bases de datos. La asignación de proteínas modelo no se vio influenciada por la interpretación del operador.

Tabla 1: Asignación de proteínas modelo para supuestas toxinas CLH en función de la comparación con dos bases de datos de secuencias

Toxina	Nombre CLH	Nombre común	Proteína modelo
--------	------------	--------------	-----------------

alfa	CLH_2834 y 2835	Factor letal	aerolisina/hemolisina (toxina alfa de <i>C. septicum</i>)
beta	CLH_1768 y 1769	Colagenasa I	Colagenasa I de <i>colG</i>
beta	CLH_2116	Colagenasa II	Colagenasa II de <i>colH</i>
épsilon	CLH_1920	Hemolisina lábil al oxígeno	Perfringolisina de <i>C. perfringens</i> tetanolisina de <i>C. tetani</i>
gamma	CLH_1861	Clostripaína	Clostripaína de <i>C. histolyticum</i>

Una inspección de los resultados del análisis BLAST del genoma de *C. histolyticum* no reveló ningún OR codificante para una elastasa. Sin embargo, se han clasificado proteasas por MEROPS (<http://MEROPS.sanger.ac.uk/>) basándose en el criterio del grupo funcional más destacado en el sitio activo de esas proteasas. Usando este enfoque funcional basado en MEROPS, una elastasa cae en la familia de la peptidasa M4, de la que la termolisina (EC 3.4.24.27) es el miembro de la familia mejor estudiado y es el modelo clásico para dichas proteasas. Usando este conocimiento, una reinspección del resultado del análisis BLAST sugirió que *C. histolyticum* posee un solo ORF que comparte una homología significativa con la termolisina. Por tanto, el supuesto gen de la toxina delta dentro de *C. histolyticum* se ha asignado como un homólogo de termolisina de *B. proteolyticus*.

Los resultados de la evaluación inicial se presentan en la Tabla 2 a continuación basados en la comparación con dos bases de datos de secuencia.

Tabla 2: Asignación de proteínas modelo para supuestas toxinas de CLH

Toxina	Nombre CLH	Nombre común	Proteína modelo
alfa	CLH_2834 y 2835	Factor letal	toxina alfa de <i>C. septicum</i>
beta	CLH_1768 y 1769	Colagenasa I	Colagenasa I de <i>colG</i>
beta	CLH_2116	Colagenasa II	Colagenasa II de <i>colH</i>
delta	CLH_2576	Elastasa	termolisina de <i>B. thermoproteolyticus</i>
épsilon	CLH_1920	Hemolisina lábil al oxígeno	Perfringolisina de <i>C. perfringens</i>
gamma	CLH_1861	Clostripaína	Clostripaína de <i>C. histolyticum</i>

ii. Alineamientos y análisis de secuencias de proteínas

Para identificar el péptido señal, la muestra y las secuencias de control se analizaron en un programa denominado SignalP para identificar secuencias de péptidos señal potenciales (Nielsen (2004), J. Glasgow et al., eds., *Proc. Sixth Int. Conf. on Intelligent Systems for Molecular Biology*, 122-130. AAAI Press, 1998). Un péptido señal generalmente se encuentra dentro de los primeros setenta aminoácidos (o la región N-terminal) de la secuencia de proteína y actúa como una secuencia señal para la enzima que se ha de secretar. El péptido señal se escinde y la secuencia de proteína resultante es la proteína madura. Usando SignalP, el usuario puede identificar la ubicación del sitio de escisión del péptido señal con el fin de identificar el extremo N terminal de la proteína madura. Para algunas secuencias de muestra, en particular la toxina alfa y la toxina beta (AUX-I), sin embargo, solo se identificó la proteína madura, no la secuencia proteínica completa, incluyendo la secuencia del péptido señal. Un examen adicional reveló que el procedimiento de fragmentación de secuencia empleado separó la secuencia del péptido señal de la proteína madura. La proteína madura y las secuencias de señal se volvieron a ensamblar y se procesaron a través de la herramienta de alineación.

Una vez que se recogieron todas las secuencias de proteínas, se construyeron alineaciones de secuencia por pares usando MATLAB 7.0.10 (The MathWorks, Inc., 2010). Las alineaciones de secuencia por pares son comparaciones directas de dos secuencias para determinar las similitudes y diferencias entre dos secuencias. Tanto el control como las secuencias de muestra se cargaron en MATLAB y se realizó una alineación usando el algoritmo Needleman-Wunsch y la matriz de puntuación BLOSUM50. El algoritmo y la matriz de puntuación ayudan a ensamblar la alineación, ya que el algoritmo dicta el valor de cada apareamiento o apareamiento incorrecto de aminoácidos basándose en la matriz de puntuación e incorpora valores de hueco cuando es necesario. Los huecos pueden producirse por múltiples razones, incluyendo, pero sin limitación, dos secuencias que tienen longitudes variables y para garantizar que los aminoácidos apropiados se emparejen entre sí. La matriz de puntuación se basa en las tasas de sustitución observadas con frecuencia entre las secuencias y sirve para calificar la similitud o disimilitud entre dos secuencias (National Center for Biotechnology Information).

La revisión de Hatheway (1990) (*Clin Microbiol Rev* 3: 66-98) indicó que las cinco toxinas eran proteínas secretadas (exosustancias) y las cinco toxinas tenían una funcionalidad identificable. Esta información se usó para realizar el análisis de las supuestas toxinas CLH. Para analizar la función proteínica de las supuestas toxinas CLH, se seleccionaron varias proteínas modelo basándose en los hallazgos de la bibliografía y los resultados de BLAST. Los

1. Toxina alfa (α)

Existe poca información relacionada con la toxina alfa de *C. histolyticum* siguiendo el trabajo de Bowen (1952) (*Yale J Biol Med* 25: 124-138) en la bibliografía. Por tanto, era de interés la interrogación de la secuencia del genoma para los supuestos genes de toxinas. Un análisis preliminar del genoma sugirió que *C. histolyticum* poseía un único ORF que comparte una homología de aminoácidos significativa con la toxina alfa de *C. septicum* según se determinó por el análisis BLAST de dos bases de datos. Por tanto, el supuesto gen de la toxina alfa dentro de *C. histolyticum* se ha asignado como un homólogo de la toxina alfa de *C. septicum*. Estudiado ampliamente por el laboratorio Rodney K. Tweten, la toxina alfa de *C. septicum* se clasificó como un miembro de la familia de toxinas similares a aerolisina. En particular, la toxina alfa de *C. septicum* posee actividad hemolítica (Ballard et al. (1992), *Infect Immun* 60: 784-790; Melton-Witt et al. (2006), *Biochem* 45: 14347-14354) y es distinta de las hemolisinas lábiles al oxígeno como se describe para la toxina ϵ de *C. histolyticum* (Hatheway (1990), *Clin Microbiol Rev* 3: 66-98).

La toxina alfa de *C. septicum* se elabora como una preproteína inactiva que se procesa en el entorno extracelular como una protoxina inactiva. Después, la protoxina se une a los receptores en la membrana celular donde se dividen en sus estructuras activas por proteasas (por lo general furina). Un sitio de consenso de furina dentro de la toxina es esencial para la activación por proteasas eucariotas. La activación implica la escisión de 40-45 aminoácidos del extremo C-terminal. En ausencia de la escisión C-terminal, la toxina alfa de *C. septicum* no es funcional. La toxina alfa de *C. septicum* de longitud completa es hemolítica (Ballard et al. 1992). La toxina activa tiene aproximadamente 41,3 kDa (Gordon et al. 1997). Una vez activadas, las toxinas se oligomerizan en la superficie celular en un complejo de preporos seguido de la inserción de un barril beta en la membrana.

La toxina alfa de *C. septicum* modelo consiste en tres dominios distintos denominados: D1, D2 y D3. El dominio D1 está implicado en la unión del receptor y la oligomerización, mientras que el dominio D2 contribuye a la estructura de horquilla anfipática. El dominio D3 tiene una región de propéptido D3 que incluye un péptido carboxilo terminal corto escindido en el sitio de activación AT conocido (R398) y funciona como una chaperona intramolecular que evita la oligomerización prematura de la toxina alfa. Usando la mutagénesis de saturación, las sustituciones de aminoácidos individuales dentro de cada dominio han permitido la determinación de aquellos restos esenciales para la actividad biológica (Melton-Witt et al., 2006). De manera importante, el ensayo funcional usó un ensayo de viabilidad celular para determinar las dosis de DL₅₀. Por tanto, el efecto relativo de las sustituciones de aminoácidos individuales dentro de la región codificantes entera se evaluó usando un ensayo funcional.

Para comprender adicionalmente la estructura primaria de la toxina alfa CLH, la alineación de proteínas, realizada en MATLAB, de la proteína modelo (toxina alfa de *C. septicum*) se realizó con la toxina alfa CLH. Los resultados se presentan en la FIG. 1.

La supuesta toxina alfa CLH traducida tiene una secuencia señal identificable y tiene una probabilidad muy alta de ser una proteína secretada. Por tanto, se consigue el primer criterio de una exosustancia. Existe una homología positiva del 75 % entre la secuencia proteica de la toxina alfa de *C. septicum* y la secuencia proteica de la toxina alfa de CLH. Se identificaron múltiples regiones de alta homología entre la toxina alfa modelo y la supuesta toxina alfa CLH. Dichas regiones y restos de aminoácidos esenciales están resaltados en sombreado de color verde en la FIG. 1.

En particular, la alineación muestra múltiples diferencias en los restos de aminoácidos esenciales que, basándose en el trabajo de Melton-Witt et al. (2006) (*Biochem* 45: 14347-14354), individualmente hacen que la proteína CLH_2834 y 2835 no sea funcional. Comenzando con la región N-terminal de la proteína madura, falta una región de secuencia de 17 aminoácidos en la secuencia de la toxina alfa CLH que se encuentra a aproximadamente 20 aminoácidos corriente abajo del sitio de escisión del supuesto péptido señal. Dentro de este tramo de 17 aminoácidos, se identificó un resto W74 en la toxina alfa de *C. septicum* como un resto crítico en el bucle 1 (L1). La falta de 17 aminoácidos del dominio D1 en la versión de secuencia CLH sugiere una estructura alterada para este dominio con respecto a un tipo silvestre y una interrupción de la funcionalidad de unión al receptor.

Dentro de la región C-terminal de la proteína, varios cambios en los restos de aminoácidos también hacen que la proteína CLH no sea funcional. El aminoácido T302 en la toxina alfa de *C. septicum* se reemplazó por Prolina en la toxina alfa de CLH. El resto E303 en la toxina alfa de *C. septicum* se reemplazó por Treonina en la toxina alfa de CLH. Los estudios de Melton-Witt et al. (2006) (*Biochem* 45: 14347-14354) indicaron que cada una de estas modificaciones dará como resultado individualmente un 0 % de letalidad. Es de destacar la comparación del sitio de activación, o sitio de escisión de furina, entre las dos secuencias. La toxina alfa de *C. septicum* presenta un sitio de escisión de consenso de furina que comienza con K391 y termina en R398. Esta región se ajusta a la secuencia de escisión de consenso de furina Arg-X-Lys/Arg-Arg aunque la secuencia de escisión mínima es Arg-X-X-Arg. La supuesta toxina alfa de CLH tiene un resto de Glutamina en lugar de Arginina en la posición análoga R398. Por tanto, el sitio de activación de *C.*

septicum posee la secuencia aminoacídica, DKRRGKRSVDS, con R398 identificado como resto crítico. La secuencia homóloga de toxina alfa de CLH en el péptido D3 es NTSST-EQNVEV; comenzando con N367 de la SEQ ID. NO: 8. Por tanto, el supuesto sitio de escisión de furina de toxina alfa de *C. histolyticum* parece no ser funcional y esta proteína, incluso si se expresa, no puede procesarse por contacto con furina proteasa de células eucariotas para generar una toxina funcional. Los hallazgos del análisis comparativo de secuencia aminoacídica se resumen en la Tabla 3.

Tabla 3: Resumen de comparación de alineación de secuencias aminoacídicas para la supuesta toxina alfa de CLH

Proteína			Efecto sobre la función
α toxina de <i>C. septicum</i>		CLH2834 y 2835	
Resto de aminoácido esencial			
W 74		Falta	Unión al receptor interrumpida
T302		P	Falta de letalidad
E303		T	Falta de letalidad
K391-R398		T---Q	Incapaz de activación

El resumen del análisis de alineación de secuencias sugiere que la supuesta toxina alfa de CLH posee un número significativo de diferencias de restos de aminoácidos que podrían hacer que la proteína madura sea no funcional. El enlace fenotípico a la funcionalidad de la toxina alfa es la demostración de la actividad hemolítica. De manera importante, la cepa de producción de colagenasa de *Clostridium Histolyticum* no presenta actividad hemolítica cuando se siembra en placas en agar sangre. Los resultados de una evaluación hemolítica de agar sangre se ilustran en la FIG. 2.

Panel A de la FIG. 2 muestra los resultados obtenidos cuando una muestra de expansión celular de *C. histolyticum* Clon 004 se cultiva en agar sangre. No hay evidencia de ningún fenotipo beta hemolítico. Por el contrario, el panel B de la FIG. 2 muestra los resultados obtenidos cuando una muestra de *C. septicum* se cultiva en agar sangre. Existen pruebas claras de que la beta hemólisis se extiende mucho más allá del área de aplicación de la muestra, como se indica en el Panel C. Las imágenes presentadas no representan adecuadamente la diferencia cualitativa observada cuando se observan los artículos de ensayo. La aparición de beta hemólisis es fácilmente discernible y la falta total de hemólisis en la placa de *C. histolyticum* contrasta con la zona amplia de hemólisis observada cuando se inspecciona el cultivo de *C. septicum* (productor de toxina α).

2. Toxina delta (δ)

Hatheway et al. (1990) (*Clin Microbiol Rev* 3: 66-98) ha definido la toxina δ de *C. histolyticum* como una elastasa, basándose principalmente en la comunicación de investigación inicial de Takahashi, et al. (1970) (*BBRC* 39: 1058-1064). Aparentemente no se han publicado más estudios sustanciales sobre esta toxina desde entonces. Takahashi et al. aislaron cuatro fracciones que demostraban actividad elastasa de *C. histolyticum* usando ultrafiltración diferencial. El enfoque principal estaba en una fracción que pasaba a través de membranas de membranas de corte nominales de 50 kDa pero quedaba retenida por membranas con un corte nominal de 10 kDa.

La termolisina es una metaloproteasa de cinc con un peso molecular de enzima madura de 34,6 kDa. De manera importante, la termolisina es una proteína modelo para una clase de proteínas que contienen una presecuencia empleada en la secreción (péptido señal) pero también una prosecuencia larga de aproximadamente 200 restos de aminoácidos que tiene dos tercios del tamaño de la proteína madura. Las enzimas similares a la termolisina se elaboran como preproteínas inactivas y la prosecuencia cumple una función como inhibidor de la enzima madura y también como chaperona para garantizar el plegamiento adecuado de la enzima (O'Donohue et al. (1996), *JBC* 271: 26477-26481). La prosecuencia se retira autocatalíticamente por la porción de enzima madura de la molécula en el entorno extracelular. Por tanto, la vía de maduración de las enzimas similares a termolisina incluye: una etapa de secreción, la presencia de una forma pro-madura en la matriz extracelular, la escisión de la prosecuencia y la presencia de una enzima activa madura.

La alineación de secuencias génicas para termolisina y CLH_2576, la supuesta toxina delta de *C. histolyticum*, se ilustra en la FIG. 3. Esta imagen muestra la secuencia aminoacídica prepromadura de longitud completa como una sola unidad que se transcribe teóricamente como un polipéptido único. Los 28 aminoácidos iniciales en el extremo N terminal de la termolisina se muestran yuxtapuestos a la prosecuencia sombreada en color verde que termina en Ser232. La secuencia aminoacídica maduros sin sombrear comienza con Ile233. Usando el programa SignalP, se prevé que la termolisina y los polipéptidos CLH_2576 se secretan. La supuesta toxina delta traducida de *C. histolyticum* tiene una secuencia señal identificable y una probabilidad muy alta de ser una proteína secretada. Existe una homología positiva del 65 % entre la secuencia proteínica de la termolisina y la secuencia proteínica de la toxina delta

de CLH.

Para comprender la naturaleza de las formas pro y maduras de ambas proteínas, las regiones individuales se analizaron como secuencias distintas con respecto a la funcionalidad. La alineación de la prosequencia se representa en la FIG. 4. Existe una homología positiva del 57 % entre las dos formas de prosequencia. Demidyuk et al. realizaron una revisión reciente del análisis estructural primario de las prosequencias de más de 100 proteasas de tipo termolisina. (2008) (*Proteína J* 27: 343-354). Estos investigadores observaron que existía una variabilidad considerable dentro de las prosequencias, denominadas de forma alternativa precursores o propéptidos. Las prosequencias fueron más tolerantes a las mutaciones en comparación con las enzimas maduras correspondientes. Sin embargo, se observaron regiones que presentan un alto grado de conservación y sustituciones en restos clave que pueden alterar drásticamente la función. Los restos sombreados en color verde en la FIG. 3 identifican los restos de aminoácidos que son críticos para que funcione la prosequencia. No se observan diferencias entre las secuencias de la termolisina y de CLH_2756. Dos restos correspondientes a Ile183 y Arg184 en la secuencia de termolisina están sombreados en color amarillo; sin embargo, las sustituciones en la secuencia CLH_2756 son aminoácidos similares que probablemente no den como resultado ninguna alteración de la función.

De manera importante, existe una región de no homología en el extremo C terminal de las prosequencias, como se ilustra mediante el sombreado de color amarillo de la secuencia CLH_2756 que comienza con Ser185. Esta región es el sitio de autocatálisis y sugiere que la secuencia CLH_2756 no es un sustrato aceptable para la escisión por el sitio activo de la enzima madura. La importancia de los restos de aminoácidos alrededor del sitio de escisión fue investigada por Wetmore et al. (1994) (*Mol Microbiol* 12: 747-759), usando la proteasa neutra similar a la termolisina de *Bacillus cereus* como enzima modelo. Estos investigadores determinaron que el procesamiento era particularmente sensible a la naturaleza del aminoácido tres restos corriente arriba del sitio de escisión. Se identificó una secuencia de consenso para la secuencia alrededor del sitio de procesamiento de proproteína y las alteraciones en los restos clave dieron como resultado la no exportación o el no procesamiento de la proteína a una enzima madura funcional. Las características clave de la secuencia de consenso fueron: la presencia de un resto no polar en la posición P₃ (Gly, Ala, Ile, Leu o Val), un resto polar o Pro en la posición P₁ (Pro, Ser, His, Glu) y un resto no polar en la posición P₁'. Adicionalmente, se ha demostrado que la maduración de la protermolisina se produce entre un resto de serina y uno de isoleucina (O'Donohue et al. (1994), *Biochem J.* 300: 599-603). Para explorar la alineación de secuencias alrededor del sitio de escisión, puede realizarse una evaluación de secuencia comparativa de los sitios de procesamiento de proproteína para la termolisina y para CLH_2756 mediante inspección. Es evidente que la secuencia aminoacídica de CLH_2756 en el área de procesamiento de proproteína no contiene la disposición de aminoácidos apropiada para permitir la autocatálisis. Cuando se realiza un ejercicio teórico para interrogar la secuencia amino de CLH_2756 para determinar si el sitio de procesamiento de proproteína está razonablemente cerca del sitio previsto basándose en la alineación de la secuencia, queda claro que ningún ajuste permite identificar la secuencia aminoacídica adecuada. El desplazamiento de los restos del sitio de procesamiento de proproteína 2 hacia el lado C terminal permite la disposición adecuada de aminoácidos que no violan las reglas de Wetmore et al. Sin embargo, la regla Ser-Ile de O'Donohue et al. (1994) (*Biochem J.* 300: 599-603) no está presente. Por tanto, se concluye que la forma de proproteína del polipéptido CLH_2756 no es un sustrato adecuado para la autocatálisis. El efecto neto es que la enzima madura activa no está presente en el caldo de células de *C. histolyticum* (Clon 004).

Para explorar las formas maduras de ambas proteínas, la alineación comparativa de secuencias se representa en la FIG. 5. Una inspección de la alineación de secuencias en la FIG. 5 sugiere que se han conservado muchos aminoácidos esenciales. En particular, la secuencia AHETHAVTD que comienza con Ala40 de la SEQ ID. NO: 13 se ha identificado como componente del sitio activo para la termolisina y la alta homología que muestra CLH_2756 sugiere que la toxina delta de CLH es un miembro de la clase termolisina de proteasas (Kooi et al. (1996), *J Med Microbiol* 45: 219-225; Kooi, et al., (1997), *Infect Immun* 65: 472-477). Se han identificado múltiples restos sombreados en color verde como esenciales para la unión o la catálisis. Una diferencia notable entre las secuencias de las dos moléculas es la región GGI que comienza con G135 en la termolisina. Este tramo de restos de aminoácidos está altamente conservado en proteasas de tipo termolisina sin función definida asignada (Frigerio et al. (1997), *Protein Eng* 10: 223-230). La región CLH_2756 correspondiente posee varias diferencias significativas en esta secuencia. Sin embargo, el alto grado general de homología y la conservación de restos de aminoácidos esenciales confirman la selección de CLH_2756 como toxina delta con una masa molecular prevista de aproximadamente 35 kDa. Esta evaluación se alinea con la información presentada por Takahashi et al (1970) (*BBRC* 39: 1058-1064).

En resumen, la supuesta toxina delta de CLH se ha identificado mediante análisis de secuencia genómica. Sin embargo, el interrogatorio de esta secuencia sugiere que no se producirá la escisión de la proproteína, lo que hace que esta molécula no funcione. Por tanto, se deduce que la toxina δ, si se expresa y secreta en el derivado Clon 004 de *C. histolyticum* ATCC 21000, no es funcional.

3. Toxina epsilon (ε)

MacLennan et al. (1962) (*Bact Rev* 26: 176-274) y Hatheway describieron la toxina ε de *C. histolyticum* como una hemolisina lábil al oxígeno serológicamente similar a las producidas por otras cepas de *Clostridium*, tales como *C. tetani*, *C. novyi* y *C. septicum*. Bowen (1952) (*Yale J Biol Med* 25: 124-138) demostró que la toxina ε se expresó durante la fase exponencial y se degradó durante la fase estacionaria como se observó para la actividad de la toxina α, y se

degradó de manera similar mediante proteinasas *in vitro*.

Una inspección de los resultados del análisis BLAST del genoma de *C. histolyticum* identificó un ORF que codificaba una hemolisina que estaba en la misma clase que la perfringolisina y la tetanolisina, que son miembros de proteínas formadoras de poros activadas por tiol con afinidad por el colesterol. Dichas proteínas son parte de una familia de citolisinas dependientes del colesterol (CDC) y todas presentan secuencias de proteínas distintivas y estructuras únicas. Se han identificado más de 25 proteínas CDC con secuencias proteicas completas disponibles. Las CDC son un grupo de toxinas formadoras de poros de β barril secretadas por diversas especies de bacterias Grampositivas, todas en el intervalo de peso molecular de 50-60 kDa. La CDC prototípica es la perfringolisina, que sirve como proteína modelo para todas las CDC (Heuck et al. 2007, *JBC* 282: 22629-22637). La organización típica de una CDC incluye una secuencia de señal escindible para facilitar las exportaciones al entorno extracelular como proteína monomérica hidrosoluble. Posteriormente, la forma monomérica plegada se une a una membrana eucariota diana, mediada por la unión al colesterol y después se oligomeriza en la superficie de la membrana para formar arcos y estructuras similares a anillo que son responsables de la citólisis. Las CDC también se conocen como citolisinas activadas por tiol y se describieron originalmente como hemolisinas (Billington et al., 2000).

La alineación de la secuencia génica para perfringolisina y CLH_1920, la supuesta toxina épsilon, se ilustra en la FIG. 6. Esta imagen muestra la secuencia proteínica de longitud completa (pre más madura) como una única unidad que se transcribe teóricamente como un único polipéptido. Los 29 aminoácidos iniciales en el extremo N terminal de la perfringolisina se ilustran con una estrella de color azul sobre Lys 29 en el sitio de escisión de la peptidasa señal. El análisis SignalP de la secuencia CLH_1920 no identificó un sitio de escisión de péptido señal reconocible y se predijo que era una proteína no secretada. Existe una homología positiva del 84 % entre la secuencia proteínica de la perfringolisina y la secuencia proteínica de la supuesta toxina épsilon CLH_1920.

Los restos de aminoácidos sombreados en color verde denotan aminoácidos esenciales que se conservan entre las dos proteínas. De manera importante, la secuencia de 11 aminoácidos ECTGLAWWWWR, comienza con el ácido glutámico 458 de la SEQ ID NO: 15, es una región esencial que se denomina secuencia undecapeptídica. Junto con el alto grado de homología dentro de la secuencia designada como la región de la proteína madura, esta secuencia undecapeptídica sirve para identificar la proteína CLH_1920 como una CDC. Por tanto, se esperaría que la proteína CLH_1920, si se elabora como una proteína secretada, tenga funcionalidad hemolítica. Una única región de no homología entre las dos proteínas se resalta en sombreado de color amarillo. De manera importante, se ha demostrado que el extremo C terminal de las CDC es crítico para la unión al colesterol (Shimada et al., 1999, *JBC* 274: 18536-18542). El proceso de hemólisis por CDC implica dos etapas críticas antes de la formación de poros: unión e inserción en la membrana. Shimada, et al. (1999) (*JBC* 274: 18536-18542) demostraron que cambios modestos en el extremo C terminal afectaron a la etapa de unión. Una alteración de los aminoácidos 3' terminales reduce gravemente la unión al colesterol medida mediante un método ELISA. La actividad hemolítica correspondiente en los eritrocitos se redujo o se eliminó coordinadamente dependiendo de la gravedad del cambio de aminoácidos C terminal. Una inspección del extremo C terminal de la secuencia de CLH_1920 muestra algunas diferencias significativas en comparación con la secuencia de la perfringolisina.

Como se resume en la Tabla 4, la actividad hemolítica de la supuesta toxina épsilon de *C. histolyticum* puede estar ausente debido a dos características de la secuencia teórica de aminoácidos. En primer lugar, se prevé que la molécula no se secretará; por tanto, la molécula no estaría disponible para la interacción con las células diana. En segundo lugar, el extremo C terminal de la proteína CLH_1920 no posee una región homóloga para la unión al colesterol, lo que sugiere que un elemento importante asociado a la hemólisis puede ser defectuoso.

Tabla 4: Resumen de comparación de alineación de secuencias aminoacídicas para la supuesta toxina épsilon de CLH

Proteína			Efecto sobre la función
<i>Perfringolisina</i>		CLH 1920	
Región		Característica	
Extremo N		Secuencia de escisión de peptidasa señal faltante	No secretada
Extremo C		No consenso	Falta de unión a colesterol/no activación

5 Los estudios de toxicidad no clínica no demostraron indicaciones clínicas ni morfológicas de los efectos de la hemolisina *in vivo*. Los datos generados por la administración en bolo local e IV respaldan la ausencia de toxinas hemolíticas tales como la toxina ϵ .

10 La ausencia de toxinas hemolíticas puede verificarse mediante el cultivo en placa del material de ensayo en agar sangre que se realiza habitualmente al final de cada fermentación de *C. histolyticum* Clon 004, lo que también confirma la ausencia de crecimiento extraño. La expresión de toxinas hemolíticas da como resultado la lisis de las células sanguíneas y, por tanto, da como resultado la formación de halos distintivos alrededor de las colonias que producen hemolisinas. La cepa de producción de colagenasa de *C. histolyticum* no produce halos o zonas de aclaramiento (véase la FIG. 2) que respalda la ausencia de toxina ϵ y cualquier otra entidad hemolítica en la cepa de producción. Para verificar la función hemolítica de una CDC, se aplicó tetanolisina disponible en el mercado al agar sangre para imitar el ensayo de cultivo en placa habitual. Los resultados se ilustran en la FIG. 7, que muestra el fenotipo beta hemolítico observado cuando se aplican 10 μ l de una solución 10 μ g/ml de tetanolisina en solución salina tamponada con fosfato a la superficie del agar sangre, después se incubó durante 24 horas a 37 °C. Por tanto, si una CDC funcional estuviera presente en el material de ensayo, se debería observar el fenotipo beta hemolítico.

20 4. Clostripaina o toxina gamma (γ)

La toxina gamma de *C. histolyticum* se ha descrito como clostripaina, una endopeptidasa de cisteína (EC 3.4.22.8). Dargatz, et al. (1993) (*Mol Gen Genet* 240: 140-145) clonaron y secuenciaron el gen de *C. histolyticum* para clostripaina y esta información se depositó en GenBank con el número de acceso X63673 (http://www.ncbi.nlm.nih.gov/nuccore/X63673.1). Para comprender la estructura primaria de la toxina gamma de CLH_1861, se realizó la alineación de la secuencia proteínica por MATLAB de la proteína modelo (clostripaina de *C. histolyticum*) con la secuencia teórica de la toxina gamma de CLH_1861. Los resultados se presentan en la FIG. 8.

30 Una inspección de la FIG. 8 muestra un grado muy alto de homología (99 %) entre la clostripaina modelo y la secuencia obtenida del análisis del genoma. De hecho, solo hay 3 diferencias de aminoácidos, ninguno de los cuales son restos identificados como esenciales para la actividad. Los aminoácidos críticos identificados en los estudios de la bibliografía como esenciales para la funcionalidad se muestran en sombreado de color verde. El análisis SignalP de ambas proteínas indicó que la puntuación alta de secreción y el sitio de escisión de señal se representaron con una estrella de color azul (Labrou et al. (2004). *Eur J Biochem* 271: 983-992). Por tanto, se podría predecir que la molécula CLH_1861 sería secretada y funcional. Se realizó un análisis residual de clostripaina como parte de la liberación de rutina.

40 El análisis de clostripaina respalda los méritos del enfoque de alineación de secuencias para las toxinas de *C. histolyticum* en general. Se podría predecir que la presencia de un gen de toxina funcional se traduciría necesariamente en una secuencia aminoacídica que compartía un alto grado de homología con una proteína modelo conocida. Adicionalmente, la conservación de restos de aminoácidos esenciales también sería una característica de un gen de toxina funcional.

45 La información obtenida del análisis de la secuencia del genoma proporcionó pruebas de que había presentes loci para supuestas toxinas alfa, delta y épsilon. Un análisis posterior de la estructura primaria teórica de cada toxina indicó que se predijeron formas no funcionales de cada toxina como consecuencia de defectos clave en la secuencia aminoacídica de cada toxina. En particular, las toxinas alfa y épsilon pueden asignarse como homólogos de dos clases de moléculas hemolíticas formadoras de poros. Como final de la fermentación, se siembran en placa muestras de cada lote en agar sangre como parte de un ensayo de pureza habitual. La falta de halos o zonas de aclaramiento 50 alrededor de las colonias confirma la ausencia de actividad hemolítica en el cultivo y la fermentación. En consecuencia, la ausencia de halos hemolíticos alrededor del extremo de muestras de fermentación demuestra la ausencia de toxinas tanto α como ϵ de forma continua.

55 La Tabla 5 muestra los resultados del análisis de secuencia y la funcionalidad prevista. Los resultados confirman por qué el Clon 004 ha demostrado funcionalmente la ausencia de toxicidad y la falta de actividad hemolítica.

Tabla 5: Resumen - Estado previsto de exosustancias de *C. histolyticum* Clon 4

Toxina	Nombre CLH	Resultado de secuencia	Funcionalidad prevista
alfa	CLH_2834 y 2835	Restos de aa críticos faltantes	No funcional; correlacionados a través de ausencia de actividad hemolítica en placas de agar sangre
delta	CLH_2576	Secuencia de escisión de proproteína de consenso faltante	No funcional
épsilon	CLH_1920	Sitio de escisión de peptidasa señal Secuencia de unión a colesterol defectuosa y no de consenso	No secretada, no funcional correlacionada a través de ausencia de actividad hemolítica en placas de agar sangre
gamma	CLH_1861	Clostripaína	Funcional

5. Análisis de secuencia de *C. histolyticum* de toxinas beta (colagenasa I y colagenasa II)

El análisis de secuencia de los supuestos loci de toxina beta de *C. histolyticum* se presenta en las FIG. 9 y 10. Como se muestra en la FIG. 9, la secuencia aminoacídica de la colagenasa I madura del clon 004 (CLH_1768 y 1769; SEQ ID NO: 3) difiere de la secuencia de colG traducida (SEQ ID NO: 19) en tres aminoácidos. La FIG. 10 muestra que la secuencia aminoacídica de la colagenasa II madura del clon 004 (CLH_2116; SEQ ID NO: 4) difiere de la secuencia de colH traducida (SEQ ID NO: 20) en ocho aminoácidos. Ambas colagenasas son completamente funcionales.

La patente y la bibliografía científica a la que se hace referencia en el presente documento establecen el conocimiento que está disponible para los expertos en la materia.

Referencias

- Nielsen et al. (2004) En J. Glasgow et al., eds., *Proc. Sixth Int. Conf. on Intelligent Systems for Molecular Biology*, 122-130. AAAI Press, 1998.
- Hatheway (1990) *Clin Microbiol Rev* 3: 66-98.
- Ballard et al. (1992) *Infect Immun* 60: 784-790.
- Melton-Witt et al. (2006) *Biochem* 45: 14347-14354.
- Gordon et al. (1997) *Infect immun* 65: 4130-4134.
- Takahashi et al. (1970) *BBRC* 39: 1058-1064.
- O'Donohue y Beaumont (1996) *JBC* 271: 26477-26481.
- Demidyuk et al. (2008) *Protein J* 27: 343-354.
- Wetmore et al. (1994) *Mol Microbiol* 12: 747-759.
- O'Donohue et al. (1994) *Biochem J.* 300: 599-603.
- Kooi y Sokol (1996) *J Med Microbiol* 45: 219-225.
- Kooi et al. (1997) *Infect Immun* 65: 472-477.
- Frigerio et al. (1997) *Protein Eng* 10: 223-230.
- MacLennan (1962) *Bact Rev* 26: 176-274.
- Bowen (1952) *Yale J Biol Med* 25: 124-138.
- Heuck et al. (2007) *JBC* 282: 22629-22637.
- Billington et al. (2000) *FEMS Microbiol Lett* 182: 197-205.
- Shimada et al. (1999) *JBC* 274: 18536-18542
- Dargatz et al. (1993) *Mol Gen Genet* 240: 140-145.
- Labrou y Rigden (2004) *Eur J Biochem* 271: 983-992.

Si bien la presente invención se ha mostrado y descrito particularmente con referencias a realizaciones preferidas de la misma, los expertos en la materia comprenderán que se pueden realizar diversos cambios en la forma y los detalles sin apartarse del alcance de la invención abarcada por las reivindicaciones adjuntas.

REIVINDICACIONES

- 5 1. Una molécula de ácido nucleico recombinante que comprende: a) un polinucleótido que tiene la secuencia de ácido nucleico de la SEQ ID NO: 1 o el complemento de la SEQ ID NO: 1, y b) una secuencia reguladora heteróloga unida operativamente al polinucleótido.
- 10 2. La molécula de ácido nucleico recombinante de la reivindicación 1, en la que la secuencia reguladora heteróloga es un promotor heterólogo unido operativamente a la secuencia de ácido nucleico de la SEQ ID NO: 1 o al complemento de la SEQ ID NO: 1.
- 15 3. La molécula de ácido nucleico recombinante de la reivindicación 1 o la reivindicación 2, en la que la molécula de ácido nucleico recombinante se fusiona con una secuencia marcadora.
4. La molécula de ácido nucleico recombinante de la reivindicación 3, en la que la secuencia marcadora codifica un polipéptido seleccionado entre el grupo que consiste en: una proteína de fusión de glutatión-S-transferasa (GST), un marcador polipeptídico de hemaglutina A (HA) de gripe y péptido de hexa-histidina.
- 20 5. Un vector que comprende la molécula de ácido nucleico de cualquier reivindicación anterior.
6. El vector de la reivindicación 5, en el que el vector es un plásmido.
7. Una célula hospedadora recombinante que comprende el vector de la reivindicación 5 o la reivindicación 6.
- 25 8. La célula hospedadora recombinante de la reivindicación 7, en la que la célula hospedadora se selecciona entre el grupo que consiste en una célula bacteriana, una célula fúngica, una célula de insecto, una célula vegetal y una célula de mamífero.
- 30 9. La célula hospedadora recombinante de la reivindicación 8, en la que la célula hospedadora se selecciona entre el grupo que consiste en *E. coli*, *Streptomyces*, *Pseudomonas*, *Serratia marcescens*, *Salmonella typhimurium*, una célula de levadura, células vegetales, timocitos, célula de ovario de hámster chino (CHO), células COS y *Lactococcus lactis*.
10. Un método de producción de colagenasa I que comprende cultivar la célula de una cualquiera de las reivindicaciones 7-9 en condiciones adecuadas para la expresión del ácido nucleico y recuperar la colagenasa I.
- 35 11. Un polipéptido recombinante codificado por la molécula de ácido nucleico recombinante de acuerdo con una cualquiera de las reivindicaciones 1-4.
- 40 12. Una composición farmacéutica que comprende un vehículo farmacéuticamente aceptable y el polipéptido recombinante de la reivindicación 11.
13. La composición farmacéutica de la reivindicación 12 para su uso en un método de tratamiento de un paciente que padece una afección mediada por colágeno.
- 45 14. La composición farmacéutica para el uso de la reivindicación 13, en la que la afección se selecciona entre el grupo que consiste en contractura de Dupuytren, enfermedad de Peyronie, capsulitis adhesiva, epicondilitis lateral, síndrome del túnel carpiano, celulitis y lipoma.

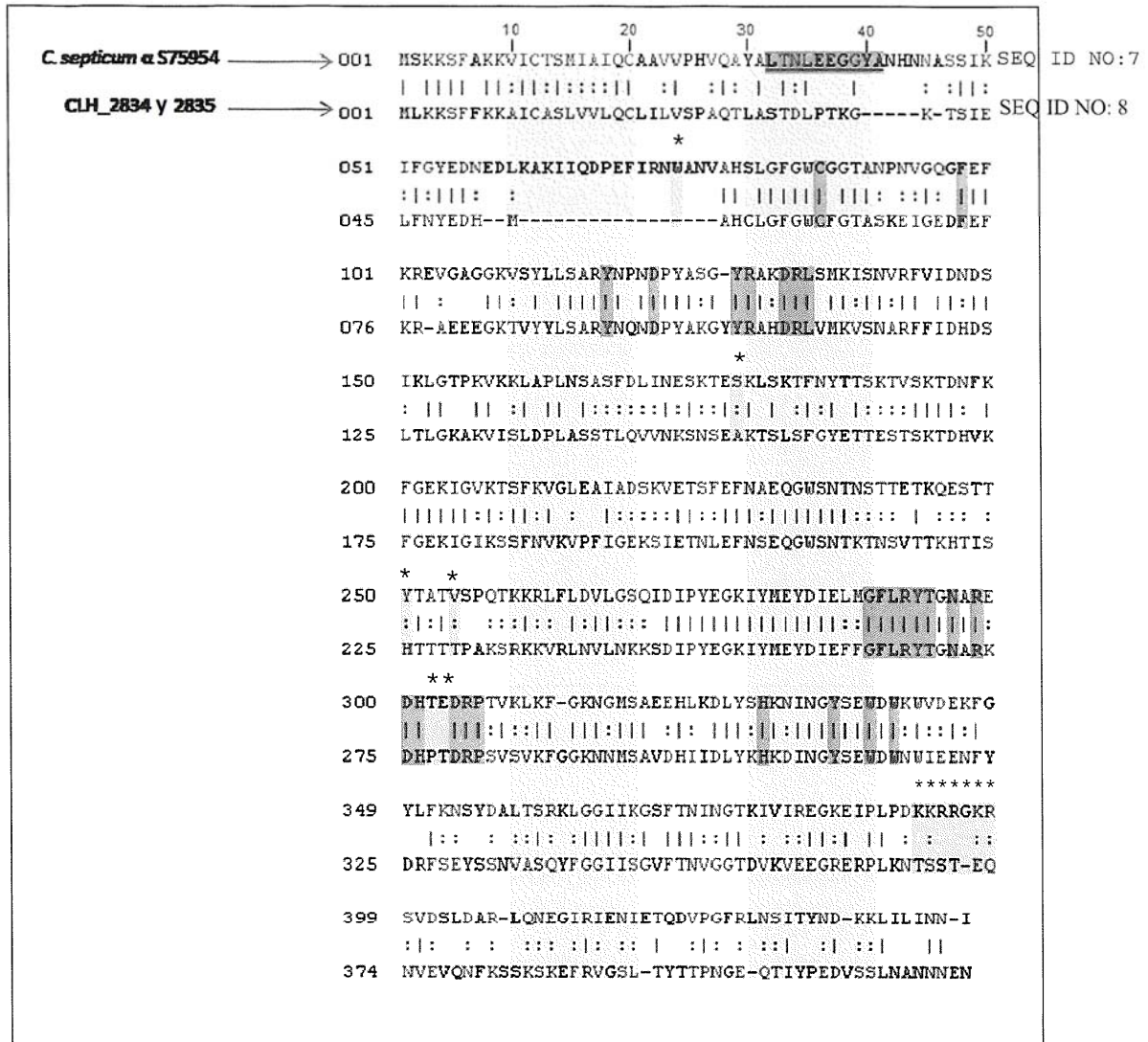
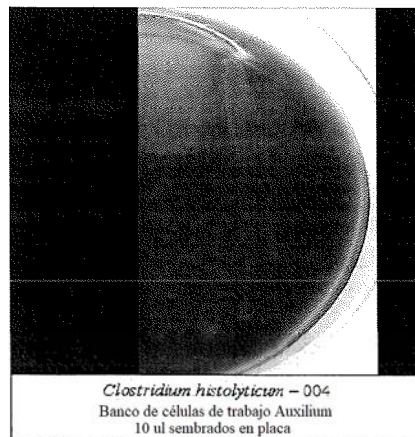
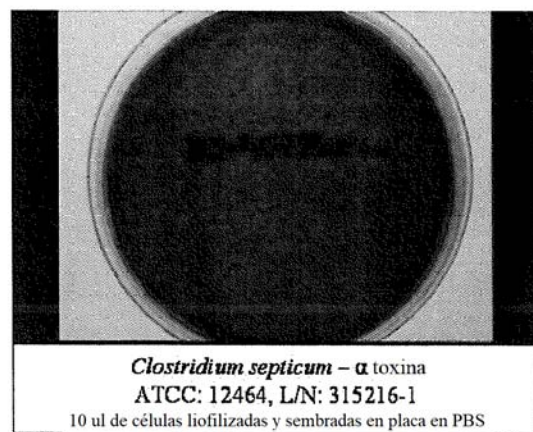


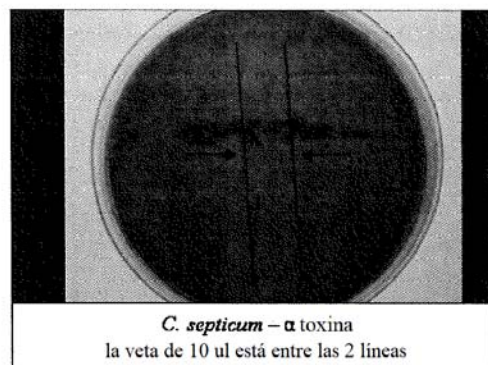
FIG. 1



Panel A



Panel B



Panel C

FIG. 2: Siembra en placa de agar sangre de *C. septicum*. Las flechas indican actividad beta hemolítica

26

```

10      20      30      40      50
Termolisina Pro  —————→ 001 STEHVTVNQQFQTPQFISGDL LKVN GTSPEELVYQYVEKNENKFKFHENA SEQ ID NO: 11
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
CLH_2576 Pro    —————→ 001 KPPKD KESISVLKTDLEKTKNIKSN NKEGDD-VTKVW-K--SALKEEGNL SEQ ID NO: 12

051 KDTLQLKEKKNDNLGFTFMR FQOTYKGIPVFGAVVTSHV-KDGLT LTALSG
    | : : : : : | : | : | : : | : | : : : |
047 GD-FKVDNKETDVGK GKHLRSQMFIDGIPVYGSQV I IHTNKD GQVYSVNG

100 TLIPNLDTKGSLKSGKKLSEKQARDIAEKDLVANVTREKVEPEYEQGKDTEF
    : :   | : : : : : | | | : | : : : : :
096 K-VDKQPKAQSFKNRVRIKDDKA I K I AEDSLGKEIKKN-KNY-HSESKLY
                                **                               ****

150 VVYVNGDEASLATVVVNLNFLTPEPGNWLYIIDAVD GKI LKNFNQLDAAKP
    : | | | : : : : : | : | : : : : | : | : : : : :
143 LYKVNGLDQPV-YLVKISSTEPEASFWMHMFVSAENGKIVDKYNAL-SCQA
    *****

200 GDVKS
    : :
191 THAQ-

```

FIG. 4

```

      10          20          30          40          50
Termolisina madura → 001 ITGTSIVGVGRGLVDQKNINITYSTVYVLQDNTR-GNG-IFTYDAKYRT SEQ ID NO: 13
    :|::: | :||: ::| : ||: ||| :|| :||||:
CLH_2576 madura → 001 VRGVNSSGEHK-IILNGMFE-N---G-RYFLADSTRPSNGYILTIDAMNQE SEQ ID NO: 14

049 -TLPGSLUADADNQFFASYDAPAVDAHYAGVTYDYKVNHNRLSYDGNN
   :||||::: :|::: :||||: :|||||| || :|:::
045 YGFPGSLFENLTGIFDSDRQKAGVDAHNNLTQVYDYKVNLRNDSFDGKG

                                     ** *
098 AAIRSSVHYSQGYNNAFWNGSQMVYGDGDGQTFLPLSGGIDVVAHELTHA
   |:| ||| ::| ||||| |::| || | : :|:||:|
095 ASIISSVHVGNNLNNAFWNGRQILFGDGDGVTFSNLAKCLEVTAHEFTHA

                                   *
148 VTDTYTAGLIYQNESGAINEAISDFCTLVEFYANKNPDWIEIGEDVYTPGI
   ||: ||| | :||:| ||:| : :| :|||:|:::
145 VTQSTAGLEYRFQSGALNEAFSDILG--IAVHSDPN-DWIEIGEDIYTPNV

198 SGDSLRSMSDPKYGDPDHYSKR-Y-TGTQDNGGVHINSGLINKAAAYLIS
   :||:| ||:| | :||: : | :||:| ||| ||||:
192 AGDALRSNSNPRLYRPQPDHMKBYLVUDYSMDKGGVHYNCGIPNKAAAYLN-

246 QGGTHYGVSVVGIGRDKLKIFYRALTQVLTPTSNTFSQLRAAAVQSATDL
   | : | |::| ||:| ||| ||| :| :| |||
241 --G-K---E-V---GKDSHAKIYYHALVNYLTPQSTFEDARNNAVVSATDL

296 YGSTSQE-VASVKQAFD-AVG---VK
   :| :| | :|: | :| |
282 HGENSKEHKLAIKS WADVGVGEEAVR

```

FIG. 5

Perfringolisina → 001 H-I--R-FKK-TK--LIASI-AMALC--L---FS-QPVISFSKDI-TD-- SEQ ID NO: 15
 | | : : : : : | | : : : | : : : : : | : : : : : | : : : : :
 CLH_1920 → 001 MKITKKGLRSLRLMLITHITGLTYNYHLGSSFNQNRVVLANPNTKTDNL SEQ ID NO: 16

034 -KNQS--IDSGISSLSYNRNEVLASNGDKIESFVPKEGKKTGNKFIVVER
 | | : : : : : | | : : : : : | | : : : : : | | : : : : : | | : : : : :
 051 IKNNSEIDEIDEKIYGLSDPYKILSYNGEKVENFVPAECSENSGKFTVIKR

081 QKRSLTTPVDISIIDSVDNRTYPGALQLADKAFVENRPTILNVKRPIN
 : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
 101 EKKNISDSTTDISIMDSINDRTYPGAIQLANRDLIENKPNLISCERKPIT

131 INIDLPLGLGENSIK-VDDPTYGKVSQAIDELVSKUNEKYSSTHTLPART
 : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
 151 ISVDLPGM-GEDGKKVVSPTYSVNSAINYLLDTUNSKYSSKYTTIPTRM

180 QYSESMVYSKSQLSSALNVNAKVLNSLGVDFAVANNEKKVHILAYKQI
 : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
 200 NYSDTMVYSKSQLSTNFGCNFKLSKSLNIDFDSIFKGEKKAHILSYKQI

230 FYTVSADLPKNPSDLFDDSVTFNDLKQKGVSEAPPLMVSNVAYGRTIYV
 | | : : : : : | | : : : : : | | : : : : : | | : : : : : | | : : : : :
 250 FYTVSVDGPNRPSDLFGYSVTSKSLALRGVMNDNPPAYVSNVAYGRTVYV

280 KLETTSSSKDVQAAFKALIKNTDIKNSQYKDIYENSSFTAVVLGGDAQE
 | | : : : : : | | : : : : : | | : : : : : | | : : : : : | | : : : : :
 300 KLETTSSSKDVQAAFKALVENQDISSNAEYKDIINQSSFTATVVLGGGAQK

330 HNKVVTKDFDEIRKVIKDNATFSTKNPAYPISYTSVFLKDNSVAAVHNKT
 | | : : : : : | | : : : : : | | : : : : : | | : : : : : | | : : : : :
 350 HNKVVTKDFDVIRNIKNNSVYSPQNGYPISYTSVFLKDNKIATVNNRT

380 DYIETTSTEYSGKINLDHSGAYVAQFEVANDVSYDKEGNEVLTHKTWD
 : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
 400 EYIETTATEYDSGKIMLDHSGVYVAQFEVTVDEVSYDKQNEIEHKSWS

430 GNYQDKTAHYSTVIPLEANARNIRIKARECTGLAWEUURDVIS EYDVPLT
 | | : : : : : | | : : : : : | | : : : : : | | : : : : : | | : : : : :
 450 GNNSDRTAHFNTELYLKGNAARNISIKARECTGLAWEUURTVVD AKNLPV

480 NNINVSINGTTLTPGSSITYN
 : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
 500 KERKLSINGTTLTPRYSNEEK

FIG. 6

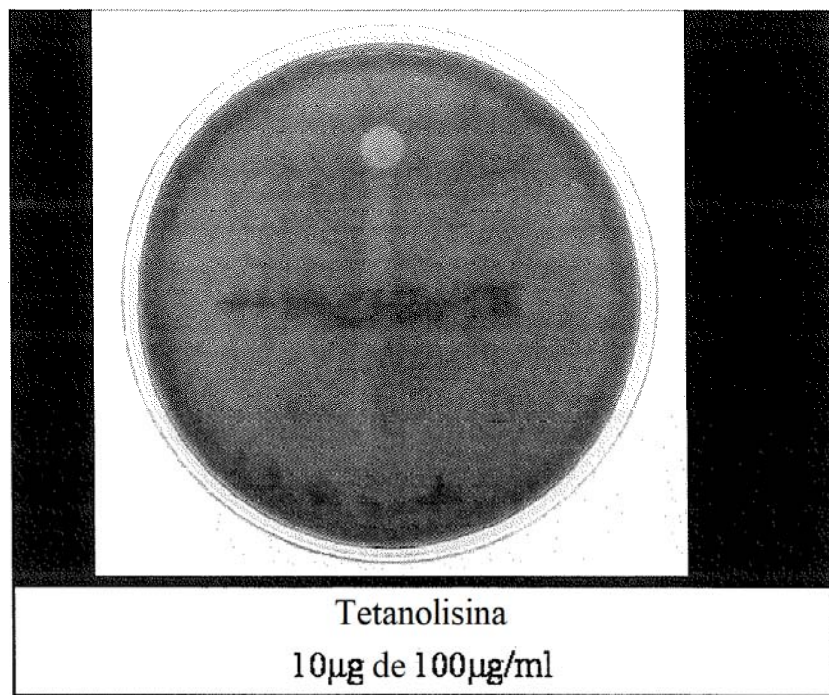


FIG. 7

```

SEQ ID NO: 17
Locus de clostripaina X63673 → 001 MLRRKYSTLLHTALITTSFLNSKPVYANPVTKSKDNNLKEVQQVTSKSHK
(SEQ ID NO: 18)
CLH_1861 → 001 MLRRKYSTLLHTALITTSFLNSKPVYANPVTKSKDNNLKEVQQVTSKSHK

051 NKNQKVTIMYYCDADNNLEGSLLNDIEEMRTGYKDSPNLNLIALVDRSPR
051 NKNQKVTIMYYCDADNNLEGSLLNDIEEMRTGYKDSPNLNLIALVDRSPR

101 YSSDEKVLGEDFSDTRLYKIEHNKANRLDGKNEFPEISTTSKYEANMGDP
101 YSSDEKVLGEDFSDTRLYKIELNKANRLDGKNEFPEISTTSKYEANMGDP

151 EVLKKFIDYCKSNYEADKYVLIMANHGGGAREKSNPRLNRAICWDDSNLD
151 EVLKKFIDYCKSNYEADKYVLIMANHGGGAREKSNPRLNRAICWDDSNLD

201 KNGEADCLYNGEISDHLTERQSVDLLAFDACLHGTAEVAYQYRPGNGGFS
201 KNGEADCLYNGEISDHLTERQSVDLLAFDACLHGTAEVAYQYRPGNGGFS

251 ADTLVASSPVVWPGFGFKYDKIFDRIKAGGGTNNEDDLTLGGKEQNFDPAT
251 ADTLVASSPVVWPGFGFKYDKIFDRIKAGGGTNNEDDLTLGGKEQNFDPAT

301 ITNEQLGALFVEEQRDSTHANGRYDQHLSFYDLKKAESVKRAIDNLAVNL
301 ITNEQLGALFVEEQRDSTHANGRYDQHLSFYDLKKAESVKRAIDNLAVNL

351 SNENKKSEIEKLRCGSIHTDLMHYFDEYSEGEWVEYYPYFDVYDLCEKINK
351 SNENKKSEIEKLRCGSIHTDLMHYFDEYSEGEWVEYYPYFDVYDLCEKINK

401 SENFSSKTKDLASAMNKLNEMIVYSFGDPSNNFKEGKNGLSIFLPNGDK
401 SENFSSKTKDLASAMNKLNEMIVYSFGDPSNNFKEGKNGLSIFLPNGDK

451 KYSTYYTSTKIPHWTMQSWYNSIDTVKYGLNPYGLSWCKDGDPEINKV
451 KYSTYYTSTKIPHWTMQSWYNSIDTVKYGLNPYGLSWCKDGDPEINKV

501 GNWFELLDsUFDKTNVDVTGGVNHVQW
501 GNWFELLDsUFDKTNVDVTGGVNHVQW

```

FIG. 8

Identities = 1113/1118 (100%), Positivos = 1118/1118 (100%)

SEQ ID NO: 3 0001 MKKNILKILMDSYSKESKIQTVRRTSVSLLAAYLTMTNTSSLVLAKPIENTNDTSIKNVEKLRN
 |||||||||||||||||||||||||||||:|||||||||||||||||||||||||||

SEQ ID NO: 19 0001 MKKNILKILMDSYSKESKIQTVRRTSVSLLAVYLTMTNTSSLVLAKPIENTNDTSIKNVEKLRN
 Extremo N terminal de proteína madura Ile 111

0065 APNEENSKKVEDSKNDKVEHVENIEEAKVEQVAPEVKSSTLRSASIANNTSEKYDFEYLNGLS
 |||||||||||||||||||||||||:|||||||||||||||||||||||||||

0065 APNEENSKKVEDSKNDKVEHVKNIEEAKVEQVAPEVKSSTLRSASIANNTSEKYDFEYLNGLS

0129 YTELTNLIKNIKWNQINGLFNYSTGSKQKFFGDKNRVQAIIN&LQESGRITYTANDMKGIETTFEV
 |||||||||||||||||||||||||||||:|||||||||||||||||||||||||||

0129 YTELTNLIKNIKWNQINGLFNYSTGSKQKFFGDKNRVQAIIN&LQESGRITYTANDMKGIETTFEV

0193 LRAGFYLGYYNDGLSYLNDRNFDQKCIPIAMIAIQKNPNFKLGTAVQDEVITSLGKLIGNASANA
 |||||||||||||||||||||||||||||:|||||||||||||||||||||||||||

0193 LRAGFYLGYYNDGLSYLNDRNFDQKCIPIAMIAIQKNPNFKLGTAVQDEVITSLGKLIGNASANA

0257 EVVMNCVPLKQFRENLNQYAPDVVKGTAVNELIKGIEFDGSGAAVEKDVKTMPWYKIDPFIN
 |||||||||||||||||||||||||||||:|||||||||||||||||||||||||||

0257 EVVMNCVPLKQFRENLNQYAPDVVKGTAVNELIKGIEFDGSGAAVEKDVKTMPWYKIDPFIN

0321 ELKALGLYGNITSATEWASDVGIYYLSKFGLYSTNRNDIVQSLEKAVDMYKYGKIAFVAMERIT
 |||||||||||||||||||||||||||||:|||||||||||||||||||||||||||

0321 ELKALGLYGNITSATEWASDVGIYYLSKFGLYSTNRNDIVQSLEKAVDMYKYGKIAFVAMERIT

0385 WDYDGIGSNGKKVDHDKFLDDAEKHYPKTYTFDNGTFIIRAGEKVSEEEKIKRLYWASREVKSQ
 |||||||||||||||||||||||||||||:|||||||||||||||||||||||||||

0385 WDYDGIGSNGKKVDHDKFLDDAEKHYPKTYTFDNGTFIIRAGDKVSEEEKIKRLYWASREVKSQ

0449 FHRVVGNDKALEVGNADDVLTMKIFNSPEEYKFNTNINGVSTDNNGGLYIEPRGTFYTYERTPQQ
 |||||||||||||||||||||||||||||:|||||||||||||||||||||||||||

0449 FHRVVGNDKALEVGNADDVLTMKIFNSPEEYKFNTNINGVSTDNNGGLYIEPRGTFYTYERTPQQ

0513 SIFSLEELFRHEYTHYLQARYLVDGLWGQGPFEYKNRLTWFDGTAEFFAGSTRTSGVLPKRSI
 |||||||||||||||||||||||||||||:|||||||||||||||||||||||||||

0513 SIFSLEELFRHEYTHYLQARYLVDGLWGQGPFEYKNRLTWFDGTAEFFAGSTRTSGVLPKRSI

FIG. 9

```

0577  LGYLAKDKVDHRYSLKKTLSNGYDDSDWMFYNYGFAVAHYLYEKDMPTFIKMNKAILNTDVKSY
      |||
0577  LGYLAKDKVDHRYSLKKTLSNGYDDSDWMFYNYGFAVAHYLYEKDMPTFIKMNKAILNTDVKSY

0641  DEIIKKLSDDANKNTEYQNHIOELADKYQGAGIPLVSDDYLDHGYKKASEVYSEISKAASLTN
      |||
0641  DEIIKKLSDDANKNTEYQNHIOELADKYQGAGIPLVSDDYLDHGYKKASEVYSEISKAASLTN

0705  TSVTAEKSYQYFNTFTLRGTYTGETSKGEFKDWDEMSKKLDGTLESLAKNSWSGYKTLTAYFTNY
      |||
0705  TSVTAEKSYQYFNTFTLRGTYTGETSKGEFKDWDEMSKKLDGTLESLAKNSWSGYKTLTAYFTNY

0769  RVTSDNKVQYDVVFHGVLTADNADISNNKAPIAKVTGPSTGAVGRNIEFSGKDSKDEDGKIVSYD
      |||
0769  RVTSDNKVQYDVVFHGVLTADNADISNNKAPIAKVTGPSTGAVGRNIEFSGKDSKDEDGKIVSYD

0833  WDFGDGATSRGKNSVHAYKKTGTYNVTLKVTDDKGATATESFTIEIKNEDTTTPITKEMEPNDD
      |||:|||||
0833  WDFGDGATSRGKNSVHAYKKAAGTYNVTLKVTDDKGATATESFTIEIKNEDTTTPITKEMEPNDD
      -

0897  IKEANGPIVEGVTVKGDLNGSDDADTFYFDVKEDGDVTIELPYSGSSNFTWLVIYKEGDDQNHIA
      |||
0897  IKEANGPIVEGVTVKGDLNGSDDADTFYFDVKEDGDVTIELPYSGSSNFTWLVIYKEGDDQNHIA

0961  SGIDKNNSKVGTFKATKGRHYVFIYKHDSASNISYSLNIKGLGNEKLKEKENNDSSDKATVIPN
      |||:|||||
0961  SGIDKNNSKVGTFKSTKGRHYVFIYKHDSASNISYSLNIKGLGNEKLKEKENNDSSDKATVIPN
      -

1025  FNTTMQGSLLGDDSRDYYSFEVKEEGEVNIELDKKDEFVGTWTLHPESNINDRITYGQVDGNKV
      |||
1025  FNTTMQGSLLGDDSRDYYSFEVKEEGEVNIELDKKDEFVGTWTLHPESNINDRITYGQVDGNKV

1089  SNKVKLRLPGKYLLVYKYSGSGNYELRVNK
      |||
1089  SNKVKLRLPGKYLLVYKYSGSGNYELRVNK

```

FIG. 9 (cont.)

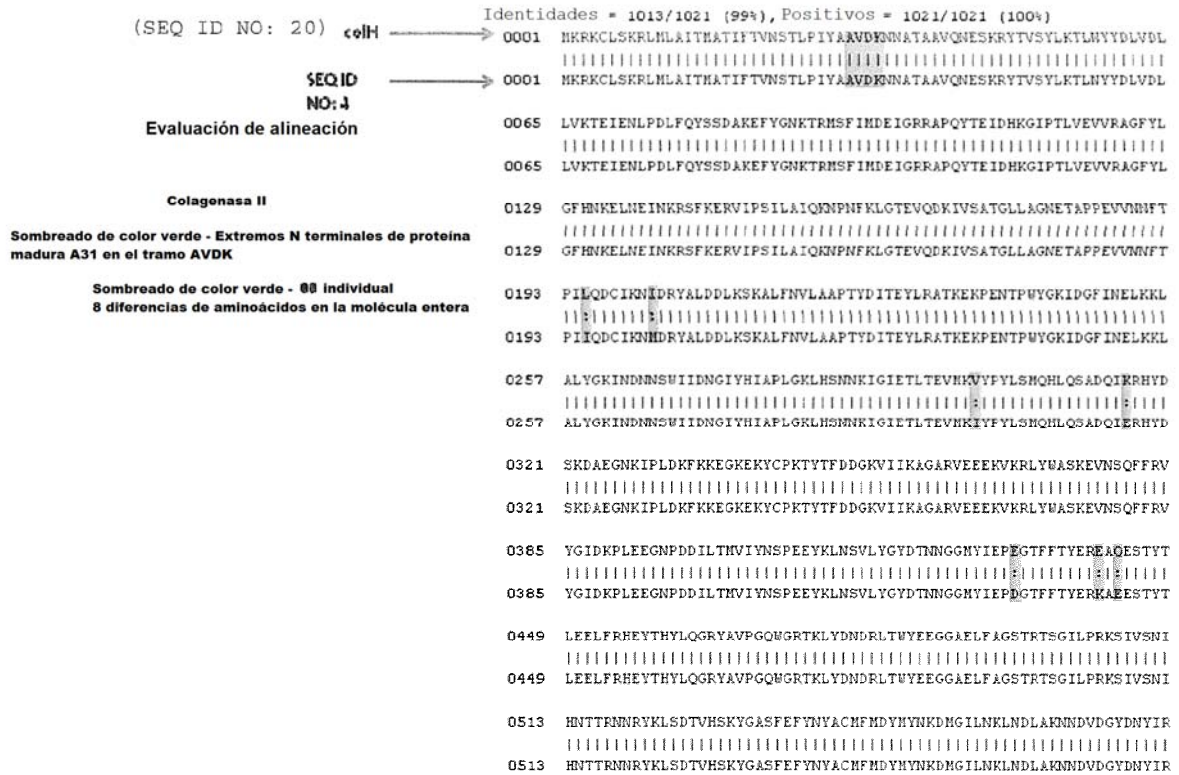


FIG. 10

```

celM → 0577 DLSSNYALNDKYQDHHQERIDNYENLTVPFVADDYLVRHAYKNPNEIYSEISEVAKLKDAKSEV
          |||||:|||||
SEQ ID NO: 4 → 0577 DLSSNHALNDKYQDHHQERIDNYENLTVPFVADDYLVRHAYKNPNEIYSEISEVAKLKDAKSEV

0641 KKSQYFSTFTLRGSYTGASKGKLEDQKAMNKFIDDSLKKLDTYSWSGYKTLTAYFTNYKVDSS
          |||||:|||||
0641 KKSQYFSTFTLRGSYTGASKGKLEDQKAMNKFIDDSLKKLDTYSWSGYKTLTAYFTNYKVDSS

0705 NRVTYDVVFHGYLPNEGDSKNSLPYGKINGTYKGTEKEKIKFSSEGSFDPDGKIVSYEWDFDGD
          |||||:|||||
0705 NRVTYDVVFHGYLPNEGDSKNSLPYGKINGTYKGTEKEKIKFSSEGSFDPDGKIVSYEWDFDGD

0769 NKSNEENPEHSYDKVGTYYTVKLKVTDDKGESSVSTTTAEIKDLSENKLPVIYMHVPKSGALNQK
          |||||:|||||
0769 NKSNEENPEHSYDKVGTYYTVKLKVTDDKGESSVSTTTAEIKDLSENKLPVIYMHVPKSGALNQK

0833 VVFYGKGTYPDGSIAGYQWDFDGSDFSSEQNPSHVYTKKGEYTVTLRVMDSSGQMSEKTMKI
          |||||:|||||
0833 VVFYGKGTYPDGSIAGYQWDFDGSDFSSEQNPSHVYTKKGEYTVTLRVMDSSGQMSEKTMKI

0897 KITDPVYPIGTEKEPNNSKETASGPIVPGIPVSGTIENTSDQDYFYFDVITPGEVKIDINKLGY
          |||||:|||||
0897 KITDPVYPIGTEKEPNNSKETASGPIVPGIPVSGTIENTSDQDYFYFDVITPGEVKIDINKLGY

0961 GGATWVVYDENNNVASYATDDGQNLSGKFKADKPGRYYIHLYMFNGSYMPYRINIEGSVGR
          |||||:|||||
0961 GGATWVVYDENNNVASYATDDGQNLSGKFKADKPGRYYIHLYMFNGSYMPYRINIEGSVGR

```

FIG. 10 (continuación)

SEQ ID NO: 1

ATGAAAAAAAAATATTTTAAAGATTCTTATGGATAGTTATTCTAAAGAATC
 TAAAATTCAAACGTACGTAGGGTTACGAGTGTATCACTTTTAGCGGCAT
 ATCTTACTATGAATACTTCAAGTTTAGTTTTAGCAAAACCAATAGAAAAT
 ACTAATGATACTAGTATAAAAAATGTGGAGAAATTAAGAAATGCTCCAAA
 TGAAGAGAATAGTAAAAAGGTAGAAGATAGTAAAAATGATAAGGTAGAAC
 ATGTGGAAAATATAGAAGAGGCCAAAAGTTGAGCAAGTTGCACCCGAAGTA
 AAATCTAAATCAACTTTAAGAAGTGCTTCTATAGCGAATACTAATTCTGA
 GAAATATGATTTTGAGTATTTAAATGGTTTGAGCTATACTGAACTTACAA
 ATTTAATTAAAAATATAAAGTGGAATCAAATTAATGGTTTATTTAATTAT
 AGTACAGGTTCTCAAAAGTTCTTTGGAGATAAAAAATCGTGTACAAGCTAT
 AATTAATGCTTTACAAGAAAGTGGAAGAACTTACACTGCAAATGATATGA
 AGGGTATAGAACTTTCACTGAGGTTTTAAGAGCTGGTTTTTATTTAGGG
 TACTATAATGATGGTTTATCTTATTTAAATGATAGAACTTCCAAGATAA
 ATGTATACCTGCAATGATTGCAATTCAAAAAAATCCTAACTTTAAGCTAG
 GAACTGCAGTTCAAGATGAAGTTATAACTTCTTTAGGAAAACATAAGGA
 AATGCTTCTGCTAATGCTGAAGTAGTTAATAATTGTGTACCAGTTCTAAA
 ACAATTTAGAGAAAACCTTAAATCAATATGCTCCTGATTACGTAAAGGAA
 CAGCTGTAAATGAATTAATTAAGGTATTGAATTCGATTTTTCTGGTGCT
 GCATATGAAAAAGATGTAAAGACAATGCCTTGGTATGGAAAAATTGATCC
 ATTTATAAATGAACTTAAGGCCTTAGGTCTATATGGAAATATAACAAGTG
 CAACTGAGTGGGCATCTGATGTTGGAATATACTATTTAAGTAAATTCGGT
 CTTTACTCAACTAACCGAAATGACATAGTACAGTCACTTGAAAAGGCTGT
 AGATATGTATAAGTATGGTAAAATAGCCTTTGTAGCAATGGAGAGAATAA
 CTTGGGATTATGATGGGATTGGTTCTAATGGTAAAAAGGTGGATCACGAT
 AAGTTCTTAGATGATGCTGAAAAACATTATCTGCCAAAGACATATACTTT
 TGATAATGGAACCTTTATTATAAGAGCAGGGGAGAAGGTATCCGAAGAAA
 AAATAAAAAAGGCTATATTGGGCATCAAGAGAAGTGAAGTCTCAATTCCAT
 AGAGTAGTTGGCAATGATAAAGCTTTAGAGGTGGGAAATGCCGATGATGT
 TTTAACTATGAAAATATTTAATAGCCCAGAAGAATATAAATTTAATACCA
 ATATAAATGGTGTAAGCACTGATAATGGTGGTCTATATATAGAACCAAGA
 GGGACTTTCTACACTTATGAGAGAACACCTCAACAAAGTATATTTAGTCT
 TGAAGAATTGTTTAGACATGAATATACTCACTATTTACAAGCGAGATATC
 TTGTAGATGGTTTATGGGGACAAGGTCCATTTTATGAAAAAATAGATTA
 ACTTGGTTTGATGAAGGTACAGCTGAATTCTTTGCAGGATCTACCCGTAC
 ATCTGGTGTTTTACCAAGAAAATCAATATTAGGATATTTGGCTAAGGATA
 AAGTAGATCATAGATACTCATTAAAGAAGACTCTTAATTCAGGGTATGAT
 GACAGTGATTGGATGTTCTATAATTATGGATTGTCAGTTGCACATTATCT
 ATATGAAAAAGATATGCCTACATTTATTAAGATGAATAAAGCTATATTGA
 ATACAGATGTGAAATCTTATGATGAAATAATAAAAAAATTAAGTGATGAT
 GCAAATAAAAAATACAGAATATCAAAACCATATTCAAGAGTTAGCAGATAA
 ATATCAAGGAGCAGGCATACCTCTAGTATCAGATGATTACTTAAAAGATC
 ATGGATATAAGAAAGCATCTGAAGTATATTCTGAAATTTCAAAAGCTGCT
 TCTCTTACAAACACTAGTGTAACAGCAGAAAAATCTCAATATTTTAACAC
 ATTCACTTTAAGAGGAACTTATACAGGTGAACTTCTAAAGGTGAATTTA
 AAGATTGGGATGAAATGAGTAAAAAATTAGATGGAACTTTGGAGTCCCTT
 GCTAAAAATTCTTGGAGTGGATACAAAACCTTTAACAGCATACTTTACGAA
 TTATAGAGTTACAAGCGATAATAAAGTTCAATATGATGTAGTTTTCCATG
 GGGTTTTAACAGATAATGCGGATATTAGTAACAATAAGGCTCCAATAGCA
 AAGGTAACCTGGACCAAGCACTGGTGCTGTAGGAAGAAATATTGAATTTAG
 TGGAAAAGATAGTAAAGATGAAGATGGTAAAATAGTATCATATGATTGGG

FIG. 11

ATTTTGGCGATGGTGCAACTAGTAGAGGCCAAAAATTCAGTACATGCTTAC
 AAAAAAACAGGAACATATAATGTTACATTAAAAGTAACTGACGATAAGGG
 TGCAACAGCTACAGAAAGCTTTACTATAGAAATAAAGAACGAAGATACAA
 CAACACCTATAACTAAAGAAATGGAACCTAATGATGATATAAAAGAGGCT
 AATGGTCCAATAGTTGAAGGTGTTACTGTAAAAGGTGATTTAAATGGTTC
 TGATGATGCTGATACCTTCTATTTTGATGTAAAAGAAGATGGTGATGTTA
 CAATTGAACTTCCTTATTCAGGGTCATCTAATTTACATGGTTAGTTTAT
 AAAGAGGGAGACGATCAAAACCATATTGCAAGTGGTATAGATAAGAATAA
 CTCAAAAGTTGGAACATTTAAAGCTACAAAAGGAAGACATTATGTGTTTA
 TATATAAACACGATTCTGCTTCAAATATATCCTATTCTTTAAACATAAAA
 GGATTAGGTAACGAGAAATTGAAGGAAAAAGAAAATAATGATTCTTCTGA
 TAAAGCTACAGTTATACCAAATTTCAATACCACTATGCAAGGTTCACTTT
 TAGGTGATGATTCAAGAGATTATTATTCTTTTGAGGTAAAGGAAGAAGGC
 GAAGTTAATATAGAACTAGATAAAAAGGATGAATTTGGTGTAACATGGAC
 ACTACATCCAGAGTCAAATATTAATGACAGAATAACTTACGGACAAGTTG
 ATGGTAATAAGGTATCTAATAAAGTTAAATTAAGACCAGGAAAATATTAT
 CTACTTGTTTATAAATACTCAGGATCAGGAACTATGAGTTAAGGGTAAA
 TAAATAA

FIG. 11 (cont.)

SEQ ID NO: 2

ATGAAAAGGAAATGTTTATCTAAAAGGCTTATGTTAGCTATAACAATGGC
 TACAATATTTACAGTGAACAGTACATTACCAATTTATGCAGCTGTAGATA
 AAAATAATGCAACAGCAGCTGTACAAAATGAAAGTAAGAGGTATACAGTA
 TCATATTTAAAGACTTTAAATTATTATGACTTAGTAGATTTGCTTGTTAA
 GACTGAAATTGAGAATTTACCAGACCTTTTTCAGTATAGTTCAGATGCAA
 AAGAGTTCTATGGAAATAAACTCGTATGAGCTTTATCATGGATGAAATT
 GGTAGAAGGGCACCACAGTATACAGAGATAGATCATAAAGGTATTCCTAC
 TTTAGTAGAAGTTGTAAGAGCTGGATTTTACTTAGGATTCCATAACAAGG
 AATTGAATGAAATAAATAAGAGGTCTTTTAAAGAAAGGGTAATACCTTCT
 ATATTGGCAATTCAAAAAAATCCTAATTTTAACTAGGTACTGAAGTTCA
 AGATAAAATAGTATCTGCAACAGGACTTTTAGCTGGTAATGAAACAGCGC
 CTCCAGAAGTTGTAAATAATTTTACACCAATAATTCAAGACTGTATCAAA
 AATATGGACAGATATGCTCTTGATGATTTAAAGTCAAAAGCATTATTTAA
 TGTTTTAGCTGCACCTACCTATGATATAACTGAGTATTTAAGAGCTACTA
 AAGAAAAACCAGAAAACACTCCTTGGTATGGTAAAATAGATGGGTTTATA
 AATGAACTTAAAAAGTTAGCTCTTTATGGAAAAATAAATGATAATAACTC
 TTGGATAATAGATAATGGTATATATCATATAGCACCTTTAGGGAAGTTAC
 ATAGCAATAATAAAATAGGAATAGAACTTTAACAGAGGTTATGAAGATA
 TATCCTTATTTAAGTATGCAACATTTACAATCAGCAGATCAAATTGAGCG
 TCATTATGATTCAAAGATGCTGAAGGAAATAAAATACCTTTAGATAAGT
 TAAAAAAGGAAGGAAAAGAGAAATACTGTCCAAAACTTATACATTTGAT
 GATGGAAAAGTAATAATAAAAGCTGGTGCTAGGGTAGAAGAAGAAAAAGT
 TAAAAAGACTATACTGGGCATCAAAGGAAGTTAACTCTCAATTCTTTAGGG
 TATATGGAATAGACAAACCATTAGAAGAAGGTAATCCAGATGATATATTA
 ACAATGGTTATCTACAACAGTCCTGAAGAATATAAACTTAATAGTGTTCT
 ATACGGATATGATACTAATAATGGTGGTATGTATATAGAGCCAGATGGAA
 CTTTCTTCACATATGAAAGAAAAGCTGAAGAAAGCACATACACATTAGAA
 GAATTATTTAGACATGAATATACACACTATTTACAAGGAAGATATGCAGT
 TCCTGGTCAATGGGGAAGAACAAAACCTTTATGACAATGATAGATTAAGTT
 GGTATGAAGAAGGTGGAGCAGAATTATTTGCAGGTTCTACTAGAAGTTCT
 GGAATATTACCAAGAAAGAGTATAGTATCAAATATTCATAATACAACAAG
 AAATAATAGATATAAGCTTTTCAGACACTGTACATTCTAAATATGGTGCTA
 GTTTTGAATTCTATAATTATGCATGTATGTTTATGGATTATATGTATAAT
 AAAGATATGGGTATATTAATAAACTAAATGATCTTGCAAAAAATAATGA
 TGTTGATGGATATGATAATTATATTAGAGATTTAAGTTCTAATCATGCTT
 TAAATGATAAATATCAAGATCATATGCAGGAGCGCATAGATAATTATGAA
 AATTTAACAGTGCCTTTTGTAGCTGATGATTATTTAGTAAGACATGCTTA
 TAAGAACCCTAATGAAATTTATTTCTGAAATATCTGAAGTAGCAAAATTAA
 AGGATGCTAAGAGTGAAGTTAAGAAATCACAATATTTTAGTACCTTTACT
 TTGAGAGGTAGTTACACAGGTGGAGCATCTAAGGGGAAATTAGAAGATCA
 AAAAGCAATGAATAAGTTTATAGATGATTCCTTAAGAAATTAGATACGT
 ATTCTTGGAGTGGGTATAAACTTTAACTGCTTATTTCACTAATTATAAA
 GTTGACTCTTCAAATAGAGTTACTTATGATGTAGTATTCCACGGATATTT
 ACCAAACGAAGGTGATTCCAAAAATTCATTACCTTATGGCAAGATCAATG
 GAACTTACAAGGGAACAGAGAAAGAAAAAATCAAATTCTCTAGTGAAGGC
 TCTTTTCGATCCAGATGGTAAAAATAGTTTCTTATGAATGGGATTTTCGGAGA
 TGGTAATAAGAGTAATGAGGAAAATCCAGAGCATTATCATATGACAAGGTAG
 GAACTTATACAGTGAAATTTAAAGTTACTGATGACAAGGGAGAATCTTCA
 GTATCTACTACTACTGCAGAAATAAAGGATCTTTCAGAAAAATAAACTTCC
 AGTTATATATATGCATGTACCTAAATCCGGAGCCTTAAATCAAAAAGTTG
 TTTTCTATGGAAAAGGAACATATGACCCAGATGGATCTATCGCAGGATAT
 CAATGGGACTTTGGTGATGGAAGTGATTTTAGCAGTGAACAAAACCCAAG

FIG. 12

CCATGTATATACTAAAAAAGGTGAATATACTGTAAACATTAAGAGTAATGG
ACAGTAGTGGACAAATGAGTGAAAAAACTATGAAGATTAAGATTACAGAT
CCGGTATATCCAATAGGCACTGAAAAAGAACCAAATAACAGTAAAGAAAC
TGCAAGTGGTCCAATAGTACCAGGTATACCTGTTAGTGGAACCATAGAAA
ATACAAGTGATCAAGATTATTTCTATTTTGATGTTATAACACCAGGAGAA
GTAAAAATAGATATAAATAAATTAGGGTACGGAGGAGCTACTTGGGTAGT
ATATGATGAAAATAATAATGCAGTATCTTATGCCACTGATGATGGGCAAA
ATTTAAGTGGAAAGTTTAAGGCAGATAAACCAGGTAGATATTACATCCAT
CTTTACATGTTTAATGGTAGTTATATGCCATATAGAATTAATATAGAAGG
TTCAGTAGGAAGATAA

FIG. 12 (cont.)

CLH_2835 y CLH_2834 (SEQ ID NO: 8)

MLKKSFFKKAICASLVVLQCLILVSPAQTLASTDLPTKGKTSIELFNYEDHMAHCLGFG
WCFGTASKEIGEDFEFKRAEEEGKTVYYLSARYNQNDPYAKGYRAHDRLVMKVSNA
RFFIDHDSLTLGKAKVISLDPLASSTLQVVNKSNSEAKTSLSGYETTESTSKTDHVKFGE
KIGIKSSFNVKVPFIGEKSIETNLEFNSEQGWSNTKTNSVTTKHTISHTTTTPAKSRKKVR
LNVLNKKSDIPYEGKIYMEYDIEFFGFLRYTGNARKDHPTDRPSVSVKFGGKNNMSAVD
HIIDLYKHKDINGYSEWDWNWIEENFYDRFSEYSSNVASQYFGGIISGVFTNVGGTDVK
VEEGRERPLKNTSSTEQNVEVQNFKSSKSKEFRVGS�TYTTPNGEQTIYPEDVSSLNANN
NEN

FIG. 13A

CLH_2834&2835 (SEQ ID NO: 21)

ATGTTAAAAAATCTTTTTTAAAAAGGCAATTTGCGCATCTTTGGTGGT
GCTACAATGTTTGATATTAGTGTCACCAGCTCAAACATTGGCATCAACAG
ATTTGCCGACAAAAGGAAAAAAGTTCAATTGAACTATTTAACTATGAAGAT
CATTAAATGGCTCATTGTTTGGGATTTGGATGGTGCTTCGGTACAGCATC
AAAAGAAATAGGGGAAGATTTTGAATTTAAAAGAGCAGAAGAAGAAGGAA
AAACAGTATATTATTTATCAGCTAGATACAATCAAATGATCCTTACGCT
AAAGGCTATTATCGCGCGCATGATAGGCTTGTTATGAAGGTTAGTAATGC
TAGGTTTTTTATCGATCATGATTCATTAAGTTTAGGAAAAGCTAAAGTTA
TAAGTCTAGATCCACTGGCATCATCAACTCTTCAAGTAGTAAATAAAAGT
AATTCTGAAGCTAAAACATCATTATCTTTTGGATATGAACTACTGAAAG
TACTTCCAAAACGGATCACGTTAAATTCGGAGAAAAAATTGGAATTAAGT
CATCATTTAATGTTAAAGTTCCATTTATAGGAGAAAAATCAATTGAAACA
AATCTTGAATTCAATTCAGAGCAGGGTTGGTCCAATACGAAAACCTAACTC
TGTAACCTACTAAACATACAATTTCTCAACAACAACAACACCTGCAAAGA
GCAGGAAAAAGGTACGATTAAATGTTCTTAATAAAAAGTCCGACATACCA
TATGAGGGTAAGATATATATGGAATATGATATAGAGTTTTTTGGTTTTTT
AAGATATACTGGAAATGCGCGTAAAGATCATCCTACAGATAGACCTAGTG
TATCAGTAAAATTTGGGGGAAAAAATAATATGAGTGCGGTAGATCATATT
ATAGATTTGTACAAGCATAAAGATATTAATGGCTATTCAGAATGGGATTG
GAATTGGATTGAAGAAAATTTTTATGATAGATTTAGTGAATATTCATCTA
ATGTTGCTAGTCAATATTTTGGGGGCATTATTTCTGGTGTATTTACTAAT
GTGGGTGGAACAGATGTAAAAGTTGAAGAAGGTAGAGAAAGGCCACTTAA
AAATACAAGTTCTACAGAACAAAATGTCTGAAGTACAGAATTTTAAAAGCT
CTAAATCTAAAGAGTTTAGAGTGGGTAGTTTAACATATACTACTCCTAAT
GGAGAACAGACCATATATCCTGAAGACGTATCATCTCTTAACGCTAACAA
CAATGAGAATTAA

FIG. 13B

CLH_2576 (SEQ ID NO: 10)

MKKKFLSFIIISAIISLNISSMTVGAKQVKEIKPPKDKESISVLKTDLEKTKNIKSNNKEGDD
VTKVVKSALKEEGLGDFKVDNKETDVKGKKHLRSQMFDGIPVYGSQVIIHTNKGQ
VYSVNGKVVDKQPKAQSFKNRVRIKDDKAIKIAEDSLGKEIKKNKNYHSESKLYLYKVN
GDLQPVYLVKISSSTEPEASFWMFVSAENGKIVDKYNALSCQATHAQVRGVNSSGEHKI
LNGMFENG RYFLADSTRPSNGYILTYDANNQEYGFPGSLFSNLTGIFDSDRQKAGVDAH
HNLTQVYDYKYNVLNRDSFDGKGASIISSVHVGNLNLNNAFWNGRQILFGDGDGVTF
SNLAKCLEVTAHEFTHAVTQSTAGLEYRFQSGALNEAFSDILGIAVHSDPNDEIGEDIYTP
NVAGDALRSMSPRLYRQPDHMKDYLYWDYSMDKGGVHYNSGIPNKAAYLMGKEV
GKDSMAKIYYHALVNYLTPQSTFEDARNAVVSALDHLHGENSKEHKLAIKSWADVGVG
EEAVR

FIG. 14A

CLH_2576 (SEQ ID NO: 22)

ATGAAAAAAAAAATTTTAAAGTTTTATTATTATTTCTGCCATATCACTTAA
 CATTTCTTCTATGACTGTGGGGGCAAAGCAAGTGAAAGAAATCAAACCTC
 CAAAAGATAAAGAATCTATTTCTGTATTAACAAACAGATTTAGAAAAAAC
 AAGAATATAAAATCTAATAATAAGGAGGGGGATGATGTAACAAAAGTAGT
 TAAGAGTGCTTTAAAAGAAGAAGGCAATTTAGGAGATTTTAAGGTTGATA
 ATAAAGAACTGATGTAAAAGGTAAAAAGCACTTGCGTTCACAAATGTTT
 ATAGATGGTATTCCTGTATATGGTAGTCAAGTTATAATTCATACTAATAA
 AGATGGACAAGTATATAGCGTAAATGGAAAAGTAGATAAACAGCCTAAAG
 CTCAATCTTTTAAAGAACCGTGTAAGGATTAAGGACGATAAAGCTATTA
 ATAGCAGAAGACAGTTTAGGTAAGGAAATAAAGAAAAACAAAAATTATCA
 TTCTGAAAGTAAGTTGTACCTATACAAGGTTAATGGAGATTTACAACCTG
 TGTATTTGGTAAAGATATCATCTACAGAACCAGAAGCTTCATTTTGGCAT
 ATGTTTGTAAGTGCTGAAAATGGAAAGATAGTTGATAAGTATAATGCTTT
 ATCATGCCAAGCTACACATGCTCAAGTAAGAGGAGTTAATAGCAGTGGAG
 AGCATAAAATCTTAAATGGTATGTTTGAAAATGGAAGATATTTTTTAGCA
 GATTCAACAAGACCTTCAAATGGATATATATTAACATATGATGCTAATAA
 CCAAGAGTATGGTTTCCCAGGTAGCTTATTTAGTAATTTAACAGGCATTT
 TTGATAGTGATAGACAAAAGGCAGGAGTAGATGCTCACCATAATCTAACT
 CAAGTATATGATTATTATAAAAAATGTTTTAAATAGAGATAGTTTTGATGG
 AAAAGGTGCTAGTATAATATCTTCTGTGCATGTAGGAAATAATTTAAATA
 ATGCTTTCTGGAATGGTAGACAAATACTTTTTGGTGATGGAGACGGAGTT
 ACATTTAGTAACCTAGCAAAATGTTTAGAAGTTACTGCCCATGAATTTAC
 ACATGCAGTTACTCAAAGTACTGCAGGTCTAGAATATAGATTTCAATCTG
 GTGCTCTAAATGAAGCTTTTTCTGATATTTTAGGTATAGCTGTTCACAGT
 GATCCAAATGATTGGGAAATTGGAGAAGATATATACACTCCTAATGTAGC
 AGGAGATGCTTTAAGAAGTATGTCAAATCCTAGATTATATAGACAACCAG
 ACCATATGAAGGACTATTTATATTGGGATTATTCAATGGATAAAGGTGGA
 GTTCATTATAATTCAGGTATTCCAAATAAAGCAGCTTATTTGATGGGAAA
 AGAAGTTGGAAAAGATTCAATGGCTAAAATTTATTATCATGCTTTAGTGA
 ATTATTTAACTCCTCAAAGTACATTTGAAGATGCTAGAAATGCAGTAGTA
 TCATCTGCAATAGATTTACATGGTGAGAATAGTAAAGAACATAAACTTGC
 TATAAAATCTTGGGCAGATGTAGGCGTTGGAGAAGAGGCAGTAAGATAA

FIG. 14B

CLH_1920 (SEQ ID NO: 16)

MKITTGGRLSRLMLITMITGLTYNYHLGSSFNNGNRVVLANPNTKTDNLIKNNNSDEIDE
KIYGLSYDPYKILSYNGEKVENFVPAECSSENSGKFTVIKREKKNISDSTDDISIMDSINDRT
YPGAIQLANRDLIENKPNLISCERKPITISVDLPGMGEDGKKVVNSPTYSSVNSAINYLLD
TWNSKYSSKYTIPTRMNYSDTMVYSKSQLSTMFGCNFKTLSKSLNIDFDSIFKGEKKAMI
LSYKQIFYTVSVDGPNRPSDLFGYSVTSKSLALKGVNNDNPPAYVSNVAYGRTVYVKLE
TTSKSSKVKA AFKALVENQDISSNAEYKDIINQSSFTATVLGGGAQKHNKVVTKDFDVI
RNIIKNNSVYSPQNP GYPISYTSTFLKDNKIATVNNRTEYIETTATEYDSGKIMLDHSGVY
VAQFEVTWDEVSYDKQGNEIIEHKSWGNNSDRTAHFNTELYLKGNARNISIKAKECTG
LAWEWWRTVVD AKNLPLVKERKLSIWGTTLYPRYSMEEK

FIG. 15A

CLH_1920 (SEQ ID NO: 23)

ATGAAGATTACAAAGAAAGGCTTAAGATCATTATCACGCTTAATGTTAAT
TACTATGATAACAGGATTAACATACAATTATCACCTAGGTAGTAGCTTTA
ATGGGAATCGAGTAGTACTTGCAAATCCAAATACAAAAACAGATAATTTA
ATTAAGAATAATAGTGATGAAATAGACGAAAAGATTTATGGATTGTCTTA
TGATCCATATAAAATATTATCTTATAATGGAGAAAAGGTTGAAAACCTTTG
TTCCAGCTGAATGTTCCGAGAATTCGGGAAAATTTACTGTAATAAAACGT
GAAAAGAAAAATATTTCAAGATTCAACTACAGATATTTCAATAATGGATTC
AATAAATGATAGAAGCTTATCCTGGTGCTATACAAGTAGCAAATAGGGATC
TTATAGAAAATAAGCCTAATTTAATTCATGCGAGAGAAAACCTATTACT
ATAAGTGTTGATTTACCTGGTATGGGTGAGGATGGGAAAAAGGTTGTAA
TTCTCCAACATACTCTTCAGTTAATTCAGCAATAAATTATTTGCTAGATA
CATGGAATTCAAAATATTCATCTAAATATACTATACCTACAAGGATGAAT
TATTCTGATACTATGGTGTATAGTAAATCACAGTTATCTACAATGTTTGG
ATGTAACCTTTAAACCTTAAAGTAAATCCTTAAATATAGATTTTGATTCTA
TATTTAAAGGCGAAAAAAAGGCTATGATTCTATCATATAAACAAATTTTC
TACACAGTGAGTGATAGATGGACCTAATCGCCCATCAGATTTATTTGGTTA
CAGTGTAACCTTCTAAGAGCTTAGCTTTAAAGGAGTAAATAATGATAATC
CTCCAGCATACGTTTCCAATGTTGCATATGGTAGAAGCTGTTTATGTAAAA
CTAGAGACAACATCTAAGAGTTCAAAGGTTAAAGCAGCATTAAAGGCATT
AGTAGAGAATCAAGATATAAGTAGTAATGCAGAATATAAAGACATAATAA
ATCAAAGTTCATTTACAGCTACTGTTCTAGGTGGAGGAGCACAAAACAC
AATAAAGTAGTTACTAAAGATTTTCGATGTAATAAGAAATATTATTAATAA
TAATTCAGTATATAGCCCAAAAAATCCTGGATATCCTATTTTCATATACAA
GTACATTTTTTAAAGACAATAAAATAGCAACTGTAAACAATAGAACAGAA
TATATAGAAACAAGTGAACAGAAATACGATAGCGGCAAAATAATGCTTGA
CCATAGTGGAGTTTATGTTGCTCAATTTGAAGTAACCTGGGATGAAGTTA
GTTATGACAAACAAGGAAATGAAATAATTGAGCATAAATCTTGGTCTGGA
AACAATAGTGATAGAACAGCTCACTTTAATACAGAACTATATTTAAAGG
AAATGCAAGAAACATTTCTATAAAAGCAAAAGAATGTACAGGCCTTGCTT
GGGAATGGTGGAGAACTGTTGTAGATGCTAAAAATTTACCACTTGTAATA
GAAAGAAAGTTATCAATATGGGGTACAACATTATATCCTAGATATTCTAT
GGAAGAGAAATAA

FIG. 15B

CLH_1861 (SEQ ID NO: 18)

MLRRKVSTLLMTALITTSFLNSKPVYANPVTKSKDNNLKEVQQVISKSNKNKNQKVTIM
YYCDADNNLEGSLNDIEEMKTGYKDSPNLNLIALVDRSPRYSSDEKVLGEDFS DTRLY
KIELNKANRLDGKNEFPEISTTSKYEANMGDPEVLKKFIDYCKSNYEADKYVLIMANHG
GGAREKSNPRLNRAICWDDSNLDKNGEADCLYMGEISDHLTEKQSVDLLAFDACLMT
AEVAYQYRPGNGGFSADTLVASSPVVWGPFGKYDKIFDRIKAGGGTNNEDDLTLGGKE
QNFDPATITNEQLGALFVEEQRDSTHANGRYDQHLSFYDLKKAESVKRAIDNLAVNLSN
ENKKSEIEKLRGSGIHTDLMHYFDEYSEGEWVEYPYFDVYDLCEKINKSENFSSKTKDL
ASNAMNKLNEMIVYSFGDPSNNFKEGKNGLSTFLPNGDKKYSTYYTSTKIPHWTMQSW
YNSIDTVKYGLNPYGKLSWCKDGQDPEINKVGNWFELLD SWFDKTNDVTGGVNHYQ
W

FIG. 16A

CLH_1861 (SEQ ID NO: 24)

ATGttaagaagaaaagtatcaacactattaatgacagctt
 tgataactacttcatttttaattccaaacccgtatatgcaaatccagtaactaaatcca
 aggataataacttaaaagaagtacaacaagttataagcaagagtaataaaaacaaaatcaaaaagtaactattgtactattgcgacgcag
 ataataacttgga
 ggaagctattaaatgatatcgaggaaatgaaaacaggatataaggatagtcctaattta
 aatttaattgctctgtagacagatccccaagatagcagtgacgaaaaagttaggt
 gaagattttagtatacacgctttataagattgaactcaataaggcaaatagattagac
 ggtaaaaaatgaatttcagaaataagtagtagtaaatatgaagctaactgggggat
 cctgaagttcttaaaaaatttattgattattgtaaatctaattatgaggctgataaatat
 gtgcttataatggcfaatcatgggggtgcaagggaataatcaatccaagattaaat
 agagcaatttgcctggatgatagtaaccttgataaaaatgggaagcagactgcctttat
 atgggtgaaatttcagatcatttaacagaaaaacaatcagttgatttactgcctttgat
 gcatgccttatgggaactgcagaagtgcgtatcagtatagaccaggtaattggaggattt
 tctgccgatacttttagttgcttcaagcccagtagtttgggtcctggattcaaatatgat
 aagatttgcataaggataaaaagctggagggaactaataatgaggatgatttaacttta
 ggtggtaagaacaaaactttgatcctgcaaccattaccaatgagcaattaggtgcatta
 tttgtagaagagcaaagagactcaacacatgccaatggcgtatgatcaacactaagc
 ttttatgatttaagaagctgaatcagtaaaaagagccatagataatttagctgttaat
 ctaagtaatgaaaacaaaaatctgaaattgaaaaattaagaggaagtggaaatcataca
 gatttaatgcattacttcgatgaatattctgaaggagaatgggtgaatatccttatttt
 gacgtgtatgattatgtgaaaaataaataaaagtgaatttttagtagtaaaactaaa
 gatttagcttcaaatgctatgaataaattaaatgaaatgatagtttattcttttgagac
 cctagtaataattttaagaaggaaaaaatggattgagtacattctacctaattggagat
 aaaaaatattcaacttattatacatcaaccaagatacctcattggactatgcaaagtgg
 tataattcaatagatacagttaaatatggattgaatccttacggaaaattaagttggtgt
 aaagatggacaagatcctgaaataaataaagtggaaattggttgaacttctagattct
 tggtttgataaaactaatgatgtaactggaggagttatcattaccaatggTAA

FIG. 16B

SEQ ID NO: 3 (colG)

MKKNILKILMDSYSKESKIQTVRRVTSVSLAAYLTMNTSSLVLAKPIENTNDTSIKNVE
 KLRNAPNEENSKKVEDSKNDKVEHVENIEEAKVEQVAPEVKSKSTLRASASANTNSEKY
 DFEYLNGLSYTELTLNIKNIKWNQINGLFNYSTGSQKFFGDKNRVQAIINALQESGRITY
 ANDMKGIETFEVLRAGFYLGYYNDGLSYLNDRNFDKQKIPAMIAIQKNPNFKLGTAVQ
 DEVITSLGKLIGNASANAEEVNNCVPVLKQFRENLNQYAPDYVKGTAVNELIKGIEFDFS
 GAAYEKDVKTMPWYGKIDPFINELKALGLYGNITSATEWASDVGIYYLSKFGLYSTNRN
 DIVQSLEKAVDMYKYGKIAFVAMERITWDYDGIGSNGKKVDHDKFLDDAEKHLYPKT
 YTFDNGTFIIRAGEKVSEEKIKRLYWASREVKSQFHRVVGNDKALEVGNADDVLTMKIF
 NSPEEYKFNTNINGVSTDNGGLYIEPRGTFYTYERTPQQSIFSLEELFRHEYTHYLQARYL
 VDGLWGQGPFEYKRNRLTWFEDEGTAEFFAGSTRTSGVLPRKSILGYLAKDKVDHRYSLK
 KTLNSGYDDSDWMFYNYGFAVAHYLYEKDMPTFIKMNKAILNTDVKSYSYDEIHKLSDD
 ANKNTEYQNHIEQELADKYQGAGIPLVSDDYLDHGYKKASEVYSEISKAASLTNTSVTA
 EKSQYFNTFTLRGTYTGETSKGEFKDWDEMSKKLDGTLES LAKNSWSGYKTLTAYFTN
 YRVTS DNKVQYDVVFHGVLT DNADISNNKAPIAKVTGPSTGAVGRNIEFS GKDSKDED
 GKIVSYDWD FGDGATSRGKNSVHAYKKTGTYNVTLKVTDDKGATATESFTIEIKNEDT
 TTPITKEMEPNDDIKEANGPIVEGVTVKGD LNGSDDADTFYFDVKEDGDVTIELPYSGSS
 NFTWL VYKEGDDQNHIASGIDKNNSKVGTFKATKGRHYVFIYKHDSASNISYSLNIKGL
 GNEKLKEKENNDSSDKATVIPNFNTTMQGSLLGDDSRDYYSFEVKEEGEVNIELDKKDE
 FGVTWTLHPESNINDRITYGQVDGNKVS NKVKLRPGKY YLLVYKYS GSGNYELRVNK

FIG. 17A

SEQ ID NO: 5

IANTNSEKYDFEYLNGLSYTELTNLIKNIKWNQINGLFNYSTGSQKFFGDKNRVQAIINA
 LQESGRITYTANDMKGIETFEVLRAGFYLGYYNDGLSYLNDRNFQDKCIPAMIAIQKNP
 NFKLGTAVQDEVITSLGKLIGNASANAEEVNNCPVLKQFRENLNQYAPDYVKGTAVN
 ELIKGIEFDFSGAAYEKDVKTMPWYGKIDPFINELKALGLYGNITSATEWASDVGIYYLS
 KFLGYSTNRNDIVQSLEKAVDMYKYGKIAFVAMERITWDYDGIGSNGKKVDHDKFLDD
 AEKHLYLPKTYTFDNGTFIIRAGEKVSEEKIKRLYWASREVKSQFHRVVGNDKALEVGNA
 DDVLTMKIFNSPEEYKFNTNINGVSTDNGGLYIEPRGTFYTYERTPQQSIFSLEELFRHEY
 THYLQARYLVDGLWGQGPFYENRLTWFDGTAEFFAGSTRTSGVLPRKSILGYLAKD
 KVDHRYSLKKTLSNGYDDSDWMFYNYGFAVAHYLYEKDMPTFIKMNAKILNTDVKSY
 DEIKKLSDDANKNTEYQNHQELADKYQGAGIPLVSDDYLDHGYKKASEVYSEISKA
 ASLTNTSVTAESQYFNTFTLRGTYTGETSKGEFKDWDEMSKKLDGTLESKAKNSWSG
 YKTLTAYFTNYRVTSNKKVQYDVVFHGVLTNDADISNNKAPIAKVTGPSTGAVGRNIEF
 SGKDSKDEDGKIVSYDWDGFGDGSRGKNSVHAYKKTGTYNVTLKVTDDKGATATES
 FTIEIKNEDTTTPITKEMEPNDDIKEANGPIVEGVTVKGDNLGSDDADTFYFDVKEDGDV
 TIELPYSGSSNFTWLVIYKEGDDQNHASGIDKNNKSVGTFKATKGRHYVFIYKHDSASNI
 SYSLNIKGLGNEKLKEKENNDSSDKATVIPNFNTTMQGSLLGDDSRDYYSFEVKEEGEV
 NIELDKKDEFVGTWTLHPESNINDRITYGQVDGNKVSNNKVKLRPGKYLLVYKYSKSG
 NYELRVNK

FIG. 17B

SEQ ID NO: 4 (colH)

MKRKCLSKRLMLAITMATIFTVNSTLPIYAAVDKNNATAAVQNESKRYTVSYLKTLY
 YDLVDLLVKTEIENLPDLFQYSSDAKEFYGNKTRMSFIMDEIGRRAPQYTEIDHKGIPTL
 VEVVRAGFYLGFNKELNEINKRSFKERVIPSILAIQKNPNFKLGTEVQDKIVSATGLLAG
 NETAPPEVVNNFTPIIQDCIKNMDRYALDDLKSKALFNVLAAPTYDITEYL RATKEKPEN
 TPWYGKIDGFINELKKLALYGKINDNNSWIIDNGIYHIAPLGKLHSNNKIGIETLTEVMKI
 YPYLSMQHLQSADQIERHYDSKDAEGNKIPLDKFKKEGKEKYCPKTYTFDDGKVIIKAG
 ARVEEEKVKRLYWASKEVNSQFFRVYGIDKPLEEGNPDDILTMVIYNSPEEYKLNSVLY
 GYDTNNGGMYIEPDGTFFTYERKAEESTYTLLEELFRHEYTHYLQGRYAVPGQWGRTKL
 YDNDRLTWYEEGGAELFAGSTRTSGILPRKSIVSNIHNTTRNNRYKLSDTVHSGYASFE
 FYNACMFMDYMYNKMGMILNKLNDLAKNNDVDGYDNYIRDLSSNHALNDKYQDH
 MQERIDNYENLTVPFVADDYLVRHAYKNPNEIYSEISEVAKLKDAKSEVKKSQYFSTFT
 LRGSYTGGAASKGKLEDQKAMNKFIDDSLKKLDTYSWSGYKTLTAYFTNYKVDSSNRVT
 YDVVFHGYLPNEGDSKNSLPYGKINGTYKGTEKEKIKFSSEGSFDPDGKIVSYEWDFGD
 GNKSNEENPEHSYDKVGTYTVKLKVTDDKGESSVSTTTAEIKDLSENKLPVIYMHVPKS
 GALNQKVVFYGGTYDPDGSAGYQWDFGDGSDFSSEQNPSHVYTKKGEYTVTLRVM
 DSSGQMSEKTMKIKITDPVYPIGTEKEPNNSKETASGPVPGIPVSGTIENTSQDYFYFD
 VITPGEVKIDINKLGYGGATWVVYDENNNNAVSYATDDGQNLSGKFKADKPGRYYIHLY
 MFNGSYMPYRINIEGSVGR

FIG. 18A

SEQ ID NO: 6

AVDKNNATAAVQNESKRYTVSYLKTNLNYYDLVDLLVKTEIENLPDLFQYSSDAKEFYG
 NKTRMSFIMDEIGRRAPQYTEIDHKGIPTLVEVVVRAGFYLGPHNKELNEINKRSFKERVIP
 SILAIQKNPNFKLGTEVQDKIVSATGLLAGNETAPPEVVNNFTPIHQDCIKNMDRYALDDL
 KSKALFNVLAAPTYDITEYLRATKEKPENTPWYKGIDGFINEKLLALYGKINDNNSWII
 DNGIYHIAPLGKLHSNNKIGIETLTEVMKIYPYLSMQHLQSADQIERHYDSKDAEGNKIP
 LDKFKKEGKEKYCPKTYTFDDGKVIIKAGARVEEEKVKRLYWASKEVNSQFFRVYGID
 KPLEEGNPDDILTMVIYNSPEEYKLNSVLYGYDTNNGGMYIEPDGTFFTYERKAEESTYT
 LEELFRHEYTHYLQGRYAVPGQWGRTKLYDNDRLTWYEEGGAELFAGSTRTSGILPRK
 SIVSNIHNTTRNNRYKLSDTVHSKYGASFIFYNYACMFMDYMYNKMGMILNKLNDLA
 KNNDVDGYDNYIRDLSSNHALNDKYQDHMQERIDNYENLTVPFVADDYLVRHAYKNP
 NEIYSEISEVAKLKDAKSEVKKSQYFSTFTLRGSYTGASKGKLEDQKAMNKFIDDSLK
 KLDTYSWSGYKTLTAYFTNYKVDSSNRVTYDVVFHGYLPNEGDSKNSLPYGKINGTYK
 GTEKEKIKFSSEGSFDPDGKIVSYEWDGFGDNKSNEENPEHSYDKVGTYTVKLKVTDDK
 GESSVSTTTAEIKDLSNKLPIYMHVPKSGALNQKVVFYGKGTYDPDGSAGYQWDFG
 DGSDFSSEQNPSHVYTKKGEYTVTLRVMDSGQMSEKTMKIKITDPVYPIGTEKEPNNS
 KETASGPIVPGIPVSGTIENTSQDYFYFDVITPGEVKIDINKLGYGGATWVVYDENNNA
 VSYATDDGQNLSGKFKADKPGRYYIHLVYMFNGSYMPYRINIEGSVGR

FIG. 18B