

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 537 338**

51 Int. Cl.:

A61K 38/48

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **16.12.2009** **E 09799683 (9)**

97 Fecha y número de publicación de la concesión europea: **29.04.2015** **EP 2398495**

54 Título: **Proteasas no citotóxicas modificadas**

30 Prioridad:

23.02.2009 GB 0903006

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

05.06.2015

73 Titular/es:

**IPSEN BIOINNOVATION LIMITED (100.0%)
Units 4-10 The Quadrant Barton Lane
Abingdon, Oxfordshire OX14 3YS, GB**

72 Inventor/es:

**CHADDock, JOHN, ANDREW y
FOSTER, KEITH, ALAN**

74 Agente/Representante:

PONTI SALES, Adelaida

ES 2 537 338 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

Proteasas no citotóxicas modificadas

5 **[0001]** La presente invención se refiere a proteasas no citotóxicas con eficacia mejorada y a la construcción de las mismas.

[0002] Las proteasas no citotóxicas son un grupo de proteasas claramente reconocido que actúan sobre las células diana incapacitando la función celular. Como aspecto importante, las proteasas no citotóxicas no destruyen
10 las células diana sobre las que actúan. Algunos de los ejemplos mejor conocidos de proteasas no citotóxicas incluyen las neurotoxinas clostridiales (por ejemplo, la neurotoxina botulínica, también conocida como BOTOX™) y las proteasas de IgA.

[0003] Las proteasas no citotóxicas actúan escindiendo proteolíticamente las proteínas de transporte
15 intracelular conocidas como proteínas SNARE (por ejemplo, SNAP-25, VAMP o syntaxina), véase Gerald K. (2002) "Cell and Molecular Biology" (4.ª edición) John Wiley & Sons, Inc. El acrónimo SNARE deriva del término: receptor de unión a NSF soluble (en inglés, Soluble NSF Attachment Receptor), en que NSF significa factor sensible a N-etilmaleimida. Las proteínas SNARE son esenciales para la formación de vesículas intracelulares y, por lo tanto, para la secreción de moléculas por una célula mediante el transporte de vesículas. Por consiguiente, una vez
20 suministrada a una célula diana, la proteasa no citotóxica es capaz de inhibir la secreción celular en la célula diana.

[0004] Las proteasas no citotóxicas pueden emplearse en su forma nativa o sustancialmente nativa (es decir, como holotoxinas, tales como BOTOX™), en cuyo caso, el direccionamiento de las proteasas a tipos de células específicos depende de (i) la administración localizada de la proteasa y/o (ii) la capacidad de unión intrínseca de la
25 proteasa nativa. Alternativamente, las proteasas no citotóxicas pueden emplearse en una forma redirigida en la que la proteasa nativa se modifica para incluir un ligando exógeno conocido como fracción de direccionamiento (TM, del inglés, Targeting Moiety). La TM se selecciona de modo que proporcione especificidad de unión para una célula diana deseada y, como parte del proceso de redireccionamiento, la porción de unión nativa de la proteasa no citotóxica puede eliminarse. La tecnología del redireccionamiento se describe, por ejemplo, en los documentos: EP-
30 B-0689459, EP-B-0939818, US 6.461.617, US 7.192.596, EP-B-0826051, US 5.989.545, US 6.395.513, US 6.962.703, EP-B-0996468, US 7.052.702, EP-B-1107794 y US 6.632.440.

[0005] A la vista de la gran distribución de las proteínas SNARE, las proteasas no citotóxicas se han
35 empleado con éxito en una amplia gama de tratamientos.

[0006] A modo de ejemplo, se menciona William J. Lipham, Cosmetic and Clinical Applications of Botulinum Toxin (Slack, Inc., 2004), que describe el uso de toxinas clostridiales tales como las neurotoxinas botulínicas (BoNT), BoNT/A, BoNT/B, BoNT/C1, BoNT/D, BoNT/E, BoNT/F y BoNT/G, y la neurotoxina tetánica (TeNT), para inhibir la
40 transmisión neuronal en una amplia diversidad de aplicaciones terapéuticas y cosméticas; como ejemplo, BOTOX™ está aprobada actualmente como producto terapéutico para las indicaciones siguientes: acalasia, espasticidad en adultos, fisura anal, dolor de espalda, blefaroespasma, bruxismo, distonía cervical, temblor hereditario, líneas glaberales o líneas faciales hiperkinéticas, dolor de cabeza, espasmo hemifacial, hiperactividad de la vejiga, hiperhidrosis, parálisis cerebral juvenil, esclerosis múltiple, trastornos mioclónicos, líneas labiales nasales, disfonía espasmódica, estrabismo y trastornos del nervio VII. Además, se han descrito tratamientos con toxinas clostridiales
45 para el tratamiento de trastornos neuromusculares (véase el documento US 6.872.397); para el tratamiento de trastornos uterinos (véase el documento US2004/0175399); para el tratamiento de úlceras y la enfermedad del reflujo gastroesofágico (véase el documento US2004/0086531); para el tratamiento de la distonía (véase el documento US 6.319.505); para el tratamiento de trastornos oculares (véase el documento US2004/0234532); para el tratamiento del blefaroespasma (véase el documento US2004/0151740); para el tratamiento del estrabismo
50 (véase el documento US2004/0126396); para el tratamiento del dolor (véanse los documentos US 6.869.610, US 6.641.820, US 6.464.986, US 6.113.915); para el tratamiento de la fibromialgia (véanse los documentos US 6.623.742, US2004/0062776); para el tratamiento del dolor en la región lumbar (véase el documento US2004/0037852); para el tratamiento de lesiones musculares (véase el documento 6.423.319); para el tratamiento del dolor de cabeza sinusal (véase el documento US 6.838.434); para el tratamiento del dolor de cabeza tensional
55 (véase el documento US 6.776.992); para el tratamiento del dolor de cabeza (véase el documento US 6.458.365); para la reducción de la migraña (véase el documento US 5.714.469); para el tratamiento de enfermedades cardiovasculares (véase el documento US 6.767.544); para el tratamiento de trastornos neurológicos tales como la enfermedad de Parkinson (véanse los documentos US 6.620.415, US 6.306.403); para el tratamiento de trastornos neuropsiquiátricos (véanse los documentos

[0007] US2004/0180061, US2003/0211121); para el tratamiento de trastornos endocrinos (véase el documento US 6.827.931); para el tratamiento de trastornos tiroideos (véase el documento US 6.740.321); para el tratamiento de glándulas sudoríparas bajo influencia colinérgica (véase el documento US 6.683.049); para el tratamiento de la diabetes (véanse los documentos US 6.337.075, US 6.416.765); para el tratamiento de un trastorno pancreático (véanse los documentos US 6.261.572, US 6.143.306); para el tratamiento de cánceres tales como tumores óseos (véanse los documentos US 6.565.870, US 6.368.605, US 6.139.845, US2005/0031648); para el tratamiento de trastornos auditivos (véanse los documentos US 6.358.926, US 6.265.379); para el tratamiento de trastornos autónomos tales como trastornos de los músculos gastrointestinales y disfunción de otros músculos lisos (véase el documento US 5.437.291); para el tratamiento de lesiones cutáneas asociadas con trastornos proliferativos de células cutáneas (véase el documento US 5,670,484); para la gestión de trastornos inflamatorios neurogénicos (véase el documento US 6.063.768); para la reducción de la caída del cabello y la estimulación de su crecimiento (véase el documento US 6.299.893); para el tratamiento de la boca caída (véase el documento US 6.358.917); para la reducción del apetito (véase el documento US2004/40253274); para tratamientos y procedimientos dentales (véase el documento US2004/0115139); para el tratamiento de trastornos y afecciones neuromusculares (véase el documento US2002/0010138); para el tratamiento de diversos trastornos y afecciones y el dolor asociado (véase el documento US2004/0013692); para el tratamiento del dolor (véase el documento WO96/33274); para el tratamiento de afecciones que resultan de la hipersecreción de moco, tales como asma y EPOC (véase el documento WO00/10598); para el tratamiento de afecciones no neuronales tales como inflamación, afecciones endocrinas, afecciones exocrinas, afecciones inmunológicas, afecciones cardiovasculares, afecciones óseas (véase el documento WO01/21213).

[0008] Se anticipa que el uso de proteasas no citotóxicas, tales como las neurotoxinas clostridiales (por ejemplo, BoNT y TeNT), en tratamientos terapéuticos y cosméticos en humanos y en otros mamíferos va a expandirse a una gama cada vez más amplia de enfermedades y dolencias que pueden beneficiarse de las propiedades de estas toxinas.

[0009] La administración de una proteasa no citotóxica (incluidos productos clínicos de neurotoxinas clostridiales nativas) puede ser problemática debido a la necesidad de dosis mayores requeridas para un efecto beneficioso. Unas dosis mayores pueden aumentar la probabilidad de que la proteasa se desplace, por ejemplo, a través de los líquidos intersticiales y los sistemas circulatorios (tales como el sistema cardiovascular y el sistema linfático) del cuerpo, lo que resulta en una dispersión no deseada de la proteasa a zonas que no son diana del tratamiento. Tal dispersión puede dar lugar a efectos secundarios no deseados, tales como la inhibición de la secreción celular en células que no son diana del tratamiento (por ejemplo, la inhibición de la liberación de neurotransmisores en neuronas que no son diana del tratamiento o la parálisis de un músculo que no es diana del tratamiento). A modo de ejemplo específico, en un paciente al que se administra una cantidad terapéuticamente eficaz de una BoNT en los músculos del cuello a causa de una tortícolis puede producirse disfagia debido a la dispersión de la proteasa en la bucofaringe. De manera similar, un paciente al que se administra una proteasa no citotóxica para tratar un trastorno neuromuscular puede sufrir una inactivación del tejido muscular no deseada debido a la dispersión de la proteasa en el músculo.

[0010] Al igual que con cualquier otro fármaco, existe un intervalo de dosificación terapéutica que identifica los límites inferior y superior de un tratamiento eficaz y seguro. Con frecuencia, el límite superior está determinado por el aumento de la importancia de los efectos fuera de la diana que conducen a efectos secundarios no deseados (potencialmente perjudiciales) del tratamiento con el fármaco. En el caso de las proteasas no citotóxicas (especialmente BoNT), esto podría conducir a la parálisis de la secreción celular en células fuera de la diana, lo que, a su vez, podría resultar letal.

[0011] El creciente uso clínico, terapéutico y cosmético de las proteasas no citotóxicas en tratamientos que requieren dosis mayores plantea a la industria farmacéutica una necesidad cada vez mayor de desarrollar medios para minimizar los efectos fuera de la diana, a la vez que se mantiene la potencia de la proteasa, de manera que el intervalo de dosis terapéuticas pueda aumentarse y así puedan proporcionarse a los pacientes dosis mayores que, a su vez, conduzcan a un aumento de la eficacia del tratamiento.

[0012] Por consiguiente, en la técnica existe la necesidad de nuevos tratamientos y/o nuevas opciones terapéuticas capaces de abordar específicamente los efectos no deseados del direccionamiento fuera de la diana. La presente invención aborda esta necesidad y soluciona uno o más de los problemas mencionados anteriormente.

[0013] En el documento WO02/044199, Lin, Wei-Jen, y col. tratan de resolver este problema proporcionando

neurotoxinas clostridiales modificadas para contener un sitio de escisión para proteasas sanguíneas (es decir, un sitio escindible por una proteasa presente en la sangre) en el dominio de unión de la neurotoxina, de manera que el contacto con una proteasa sanguínea inactiva selectivamente la neurotoxina. Dicho dominio de unión (denominado también dominio H_C) se ilustra en la figura 1B de Lin, Wei-Jen, y col. como la región que comienza en el resto aminoacídico 873. Sin embargo, la solución mencionada anteriormente que proporcionan Lin y col., presenta una serie de problemas y no resuelve adecuadamente el problema de los efectos del direccionamiento fuera de la diana. A este respecto, los presentes inventores han identificado que las neurotoxinas clostridiales en las que se ha eliminado (o inactivado de otro modo) el dominio de unión (H_C) todavía son tóxicas y pueden tener un efecto de inhibición sobre sus neuronas diana; esto lo confirma la figura 1 (véase el ejemplo 39) de la presente solicitud, que ilustra la escisión de proteínas SNARE por una molécula de neurotoxina clostridial (LH_N) que carece del dominio de unión (H_C). Otra deficiencia asociada con el documento WO02/044199 (Lin, Wei-Jen, y col.) es que la tecnología descrita se limita a moléculas de neurotoxinas clostridiales que tienen un dominio H_C (es decir, moléculas de holotoxinas clostridiales). Sin embargo, según se ha expuesto anteriormente, las proteasas no citotóxicas pueden emplearse en una forma redirigida en la que la proteasa nativa se modifica para incluir un ligando exógeno conocido como fracción de direccionamiento (TM), que proporciona especificidad de unión para una célula diana deseada. Por lo tanto, en el contexto de las proteasas no citotóxicas redirigidas, la descripción de Lin y col. no aborda el problema de los efectos del direccionamiento fuera de la diana.

[0014] La presente invención aborda las deficiencias de Lin y col. y proporciona proteasas no citotóxicas que reducen o evitan los efectos secundarios no deseados asociados con su dispersión en áreas no seleccionadas como diana. Estas y otras ventajas relacionadas son útiles para diversas aplicaciones clínicas, terapéuticas y cosméticas, tales como el tratamiento de trastornos neuromusculares, trastornos neuropáticos, trastornos oculares, dolor, lesiones musculares, dolor de cabeza, enfermedades cardiovasculares, trastornos neuropsiquiátricos, trastornos endocrinos, trastornos exocrinos, trastornos relacionados con la secreción de moco, tales como asma y EPOC, cánceres, trastornos auditivos y líneas faciales hiperkinéticas, así como otros trastornos en los que la administración de proteasas no citotóxicas a un mamífero puede producir un efecto beneficioso (por ejemplo, todos los tratamientos descritos en las páginas 2-3 de esta memoria descriptiva).

[0015] La presente invención proporciona polipéptidos según se especifican en las reivindicaciones. Además se proporcionan ácidos nucleicos codificantes de los polipéptidos de la invención.

[0016] Por lo tanto, la presente invención proporciona un polipéptido que puede inactivarse y/o destruirse controladamente en una localización fuera de la diana.

[0017] En más detalle, un primer aspecto de la presente descripción proporciona un polipéptido que comprende:

- a. una proteasa no citotóxica que es capaz de escindir una proteína SNARE;
- b. un dominio de translocación que es capaz de translocar la proteasa no citotóxica desde dentro de un endosoma de una célula de mamífero, a través de la membrana endosómica del mismo hasta el citosol de la célula de mamífero;
- c. un primer sitio de escisión destructivo que es escindible por una segunda proteasa y no por la proteasa no citotóxica y en que después de su escisión por la segunda proteasa, el polipéptido tiene una potencia reducida medible por una capacidad reducida de escindir dicha proteína SNARE y/o una capacidad reducida de translocar dicha proteasa no citotóxica a través de una membrana endosómica;
- d. una fracción de direccionamiento (TM) que se une a un sitio de unión presente en una célula neuronal de un mamífero, en que el sitio de unión es capaz de experimentar endocitosis para incorporarse en un endosoma dentro de la célula neuronal del mamífero; y
- e. con la condición de que dicho primer sitio de escisión destructivo no esté localizado dentro de dicha TM.

[0018] En una realización descrita en este documento, el sitio de escisión destructivo es reconocido y escindido por una segunda proteasa (es decir, una proteasa destructiva) seleccionada entre una proteasa circulante (por ejemplo, una proteasa extracelular, tal como una proteasa sérica o una proteasa de la cascada de coagulación sanguínea), una proteasa asociada a tejidos (por ejemplo, una metaloproteasa matricial (MMP), tal como una MMP muscular) y una proteasa intracelular (preferentemente una proteasa que está ausente en la célula diana).

[0019] Por lo tanto, en la práctica, si un polipéptido de la presente invención se dispersa fuera de su célula diana deseada y/o es absorbido por una célula no diana, el polipéptido es inactivado por escisión del sitio de escisión destructivo (por la segunda proteasa).

5

[0020] En una realización, el sitio de escisión destructivo es reconocido y escindido por una segunda proteasa que está presente dentro de un tipo de célula fuera de la diana. En esta realización, la célula fuera de la diana y la célula diana son preferentemente de distinto tipo. Alternativamente (o adicionalmente), el sitio de escisión destructivo es reconocido y escindido por una segunda proteasa que está presente en una localización fuera de la diana (por ejemplo, distal con respecto a la célula diana). Por consiguiente, cuanto se produce una escisión destructiva extracelularmente, la célula diana y la célula fuera de la diana pueden ser del mismo o de distinto tipo celular. A este respecto, la célula diana y la célula fuera de la diana pueden tener respectivamente un receptor al que se une el mismo polipéptido de la invención.

[0021] A modo de ejemplo, cuando se tratan trastornos musculares, un polipéptido de la presente invención se dirige a la célula diana deseada (por ejemplo, una neurona motora) e incluye un sitio destructivo de escisión proteásica que es escindido por una segunda proteasa presente dentro de y/o en proximidad inmediata al tejido muscular. Por consiguiente, el polipéptido demuestra efectos adversos mínimos en el tejido muscular y puede usarse en dosis mayores que las tolerables actualmente por los pacientes, lo que conduce a una mayor eficacia clínica.

20

[0022] El sitio de escisión destructivo de la presente invención proporciona la inactivación/destrucción del polipéptido cuando este está en una localización fuera de la diana. En este aspecto, la escisión en el sitio de escisión destructivo minimiza la potencia del polipéptido al reducir la capacidad intrínseca de dicho polipéptido (en comparación con un polipéptido idéntico que no presenta el mismo sitio de escisión destructivo o que presenta el mismo sitio destructivo pero en una forma no escindida) para translocar el componente no citotóxico (a través de la membrana endosómica de una célula de un mamífero en dirección al citosol) y/o efectuar la escisión de las proteínas SNARE.

[0023] En una realización, el polipéptido de la invención puede incluir un segundo (o siguiente) sitio de inactivación/destrucción. Dicho segundo (o siguiente) sitio puede localizarse en cualquier lugar dentro del polipéptido (incluido el componente TM). Dicho segundo (o siguiente) sitio puede ser escindido por la misma proteasa o por una diferente. Dicho segundo (o siguiente) sitio puede tener una secuencia aminoacídica de reconocimiento diferente del primer sitio de inactivación/destrucción y puede ser escindido por la misma proteasa o por una diferente.

[0024] La reducida capacidad de escisión de proteínas SNARE y/o la reducida capacidad de translocación mencionadas anteriormente pueden medirse fácilmente por comparación directa de un polipéptido de la invención con un polipéptido idéntico (aunque sin el mismo sitio de escisión destructivo o con el mismo sitio de escisión destructivo, pero en forma no escindida). En más detalle, el polipéptido de la invención y el equivalente correspondiente no escindido pueden ensayarse en paralelo en uno cualquiera de una diversidad de ensayos convencionales con células enteras o en ensayos acelulares. A modo de ejemplo, se mencionan los ejemplos 1-4. Durante dichos ensayos, el polipéptido de la invención se inactiva (por medio de su escisión en el sitio de escisión destructivo), mientras que el polipéptido equivalente retiene sustancialmente toda la potencia. Por lo tanto, en el contexto de la presente invención, cuando está escindido en el sitio de escisión destructivo, un polipéptido de la invención tiene menos del 50 % o menos del 25 %, menos del 10 % o menos del 5 %, menos del 1 % o menos del 0,5 %, menos del 0,1 % o menos del 0,01 % o menos del 0,001 % o menos del 0,0001 % de la capacidad de escisión de proteínas SNARE y/o una capacidad de translocación reducida en comparación con el polipéptido equivalente no escindido.

[0025] En el contexto de los ensayos con células enteras, la reducida capacidad de escisión de proteínas SNARE y/o la reducida capacidad de translocación pueden determinarse midiendo la escisión relativa de proteínas SNARE en una célula de un mamífero. Esto refleja la capacidad total del polipéptido para translocarse al citosol de una célula de un mamífero y posteriormente escindir una proteína SNARE. Existen diversos modos de medir la escisión de proteínas SNARE, tales como, por ejemplo, SDS-PAGE y inmunotransferencia, con posterior análisis densitométrico de los productos escindidos.

55

[0026] En el contexto de los ensayos acelulares, la potencia puede medirse con respecto a la escisión relativa de proteínas SNARE o con respecto a la función de translocación relativa (por ejemplo, la liberación de K⁺ o NAD de los liposomas o mediciones de la conductancia de la membrana).

[0027] Los sitios fuera de la diana preferidos (y por lo tanto, células de mamíferos preferidas para los ensayos anteriores) incluyen: células epiteliales, especialmente células epiteliales pulmonares, células neuronales que no son células neuronales motoras y células musculares.

5 **[0028]** En referencia al ejemplo 39, se proporciona una neurotoxina clostridial modificada (LH_N/C). Esta neurotoxina mimetiza la neurotoxina modificada de Lin y col. (es decir, según se expone en la parte de los antecedentes de esta memoria descriptiva) ya que carece de un dominio de unión H_C funcional. Dicha neurotoxina modificada se incubó en presencia de células de mamífero (por ejemplo, neuronas de médula espinal embrionaria no humana (eSCN)) para analizar su capacidad de demostrar una actividad residual de neurotoxina clostridial en
10 forma de la escisión de proteínas SNARE. En paralelo, se incubó de la misma manera una neurotoxina de control que constaba solamente del dominio endopeptidásico de la neurotoxina botulínica de tipo C (LC/C); por lo tanto, la neurotoxina de control carecía de un dominio de translocación H_N funcional. Cada uno de los dos polipéptidos se analizó después en cuanto a la escisión de proteínas SNARE en las células de prueba. De manera sorprendente, la neurotoxina clostridial modificada LH_N/C demostró la escisión de proteínas SNARE (véase la figura 1) y así se
15 confirmó que la inactivación del dominio de unión H_C de la neurotoxina botulínica no es adecuada para reducir la actividad fuera de la diana. En contraste, la neurotoxina de control (que carece de un componente de translocación funcional) demostró que carecía de la capacidad de escindir proteínas SNARE.

[0029] Según se menciona anteriormente, el polipéptido de la presente invención puede incluir uno o más (por
20 ejemplo, dos, tres, cuatro, cinco o más) sitios destructivos de escisión proteásica. Cuando se incluye más de un sitio de escisión destructivo, cada uno de los sitios puede ser igual o diferente. A este respecto, el uso de más de un sitio de escisión destructivo proporciona una mejora de la inactivación fuera de la diana. De manera similar, el uso de dos o más sitios de escisión destructivos diferentes proporciona una flexibilidad de diseño adicional. Por ejemplo, cuando se minimizan los efectos fuera de la diana en los tejidos musculares, el polipéptido de la presente invención puede
25 incluir dos sitios destructivos diferentes que son reconocidos y escindidos por dos proteasas asociadas a tejidos musculares diferentes.

[0030] El primer sitio(s) de escisión destructivo(s) puede(n) incorporarse en el componente proteásico no citotóxico o en el componente de translocación. El segundo (o siguientes) sitio(s) puede(n) incorporarse en cualquier
30 lugar del polipéptido. A este respecto, el(los) sitio(s) de escisión destructivo(s) se elige(n) para asegurar mínimos efectos adversos sobre la potencia del polipéptido (por ejemplo, por tener un mínimo efecto sobre el dominio de translocación y/o el dominio proteásico no citotóxico), mientras se asegura que el polipéptido es lábil fuera del sitio o célula diana.

35 **[0031]** Los sitios de escisión destructivos preferidos (más las correspondientes segundas proteasas) se listan en la tabla inmediatamente a continuación. Estos sitios de escisión listados son puramente ilustrativos y no se pretende que sean limitantes de la presente invención.

Segunda proteasa	Secuencia de reconocimiento del sitio de escisión destructivo	Variación tolerada de la secuencia de reconocimiento P4-P3-P2-P1-▼-P1'-P2'-P3'						
		P4	P3	P2	P1	P1'	P2'	P3'
Trombina	LVPR▼GS	A, F, G, I, L, T, V o M	A, F, G, I, L, T, V, W o A	P	R	No D o E	No D o E	---
Trombina	GR▼G			G	R	G		
Factor Xa	IEGR▼	A, F, G, I, L, T, V o M	D o E	G	R	---	---	---
ADAM17	PLAQA▼VRSSS							
Proteasa similar a tripsina de las vías respiratorias humanas (HAT)	SKGR▼SLIGRV							
ACE (peptidildipeptidasa A)		---	---	---	---	no P	No D o E	N/A

Elastasa (leucocitos)	MEA▼VTY	M, R	E	A, H	V, T	V, T, H	Y	---
Furina	RXR/KR▼	R	X	R o K	R			
Granzima	IEPD▼	I	E	P	D	---	---	---
Caspasa 1		F, W, Y, L	---	H, A, T	D	No P, E, D, Q, K o R	---	---
Caspasa 2	DVAD▼	D	V	A	D	No P, E, D, Q, K o R	---	---
Caspasa 3	DMQD▼	D	M	Q	D	No P, E, D, Q, K o R	---	---
Caspasa 4	LEVD▼	L	E	V	D	No P, E, D, Q, K o R	---	---
Caspasa 5		L o W	E	H	D	---	---	---
Caspasa 6		V	E	H o I	D	No P, E, D, Q, K o R	---	---
Caspasa 7	DEVD▼	D	E	V	D	No P, E, D, Q, K o R	---	---
Caspasa 8		I o L	E	T	D	No P, E, D, Q, K o R	---	---
Caspasa 9	LEHD▼	L	E	H	D	---	---	---
Caspasa 10	IEHD▼	I	E	H	D	---	---	---

[0032] La presente invención puede emplear sitios de escisión destructivos que pueden ser escindidos por una proteasa sanguínea de mamíferos, tales como trombina, factor de coagulación VIIa, factor de coagulación IXa, factor de coagulación Xa, factor de coagulación XIa, factor de coagulación XIIa, calicreína, proteína C y proteasa de serina asociada a MBP.

[0033] Lin y col. describen el uso de los sitios de escisión para la trombina o el factor Xa para inactivar el dominio de unión H_C de una holotoxina clostridial. Sin embargo, según se expone anteriormente, la inactivación de H_C es inadecuada para alcanzar la inactivación deseada fuera de la diana. Además, debido a la escasez de sitios de escisión desvelados, el procedimiento descrito por Lin y col. tiene una utilidad limitada, por ejemplo, en entornos fuera de la diana en los que la trombina y el factor Xa están ausentes (o solo presentes en bajas concentraciones).

[0034] Las metaloproteasas matriciales (MMP) son un grupo preferido de proteasas destructivas en el contexto de la presente invención. Dentro de este grupo, se prefiere ADAM17 (EC 3.4.24.86), también conocida como TACE, que escinde una diversidad de proteínas de la superficie celular ancladas a la membrana para la "eliminar" los dominios extracelulares. Adicionalmente, las MMP preferidas incluyen adamalisin, serralisin y astacin.

[0035] En una realización de la presente invención, dicho(s) sitio(s) de escisión destructivo(s) comprende(n) una secuencia de reconocimiento con al menos tres o cuatro, preferentemente cinco o seis, con mayor preferencia seis o siete y con especial preferencia al menos ocho restos aminoacídicos contiguos. A este respecto, cuanto más larga (en cuanto a restos aminoacídicos contiguos) sea la secuencia de reconocimiento, menor será la probabilidad de que se produzca una escisión inespecífica del sitio destructivo por una segunda proteasa accidentalmente.

[0036] Opcionalmente, el polipéptido de la presente invención incluye una fracción de direccionamiento (TM) que se une a un sitio de unión en una célula neuronal (por ejemplo, nerviosa), lo que proporciona selectividad al polipéptido para esta especie de célula diana sobre otras células. En una realización, la célula neuronal es una célula de la unión neuromuscular o una terminación presináptica del sistema nervioso periférico colinérgico.

[0037] El primero (y siguientes) sitio(s) de escisión destructivo(s) de la presente invención se introduce(n)

preferentemente en el componente proteásico y/o en el componente de translocación. De estos dos componentes, se prefiere el componente proteásico. Por consiguiente, el polipéptido puede inactivarse rápidamente por destrucción directa del componente proteásico no citotóxico y/o el componente de translocación. Estas posiciones de inserción son preferibles a una posición de inserción en la TM ya que, incluso en el caso de una inactivación total de la TM, se ha demostrado que el polipéptido resultante puede no mostrar una reducción de potencia adecuada en las células fuera de la diana [Chaddock, J.A. y col., Protein Expression Purification 2002, 25: 219-228 y Sutton, J. M. y col., Protein Expression & Purification 2005, 40(1): 31-41].

[0038] Por lo tanto, el polipéptido de la presente invención no comprende un sitio(s) de escisión destructivo(s) solamente en el componente de la fracción de direccionamiento del polipéptido. Sin el deseo de limitarse a ninguna teoría, se cree que el uso de un sitio de destrucción dentro del componente TM solamente no aborda la absorción inespecífica por las células fuera de la diana. El ejemplo 39 (véase también la figura 1) demuestra que un fragmento de la neurotoxina botulínica de tipo C que carece del dominio de unión H_C todavía es capaz de entrar en las eSCN y escindir su proteína SNARE sustrato (syntaxina). Otra posibilidad es que la escisión dentro del componente TM puede conducir a una TM con mayor afinidad de unión por las células fuera de la diana, por ejemplo, a través de la exposición de una región de mayor afinidad de unión dentro de la TM. En resumen, el uso de un sitio(s) de escisión destructivo(s) dentro del componente TM solamente se considera insatisfactorio. En primer lugar, no se aborda adecuadamente el direccionamiento fuera de la diana y, en segundo lugar, una vez suministrados a una célula fuera de la diana, los polipéptidos todavía son capaces de mostrar actividad de translocación (natural) y/o actividad de escisión de proteínas SNARE.

[0039] Se prefiere que la TM no tenga ningún sitio de escisión destructivo. A este respecto, se ha demostrado que el componente TM puede ser especialmente sensible a cambios conformacionales adversos (al insertar un sitio de escisión destructivo), que afectan negativamente a la unión deseada del polipéptido. Se ha demostrado que esto es un problema especial cuando la TM es la fracción de direccionamiento nativa de una neurotoxina clostridial (es decir, H_C).

[0040] Las TM adecuadas para uso en los polipéptidos de la presente invención incluyen citocinas, factores de crecimiento, neuropéptidos, lectinas, andamiajes de unión a proteínas y anticuerpos (este término incluye anticuerpos monoclonales y fragmentos de anticuerpos tales como Fab, F(ab)₂, Fv, ScFv, etc.).

[0041] La TM es un ligando (preferentemente un péptido) que se une a una célula neuronal, preferentemente a una célula neuronal de la unión neuromuscular. A este respecto, en una realización, la TM comprende el dominio de unión (H_{CC} o H_C) de una neurotoxina clostridial (por ejemplo, BoNT, TeNT, o de otro *Clostridium spp.*) o un fragmento del mismo que tiene la capacidad de unión nativa de la neurotoxina. El dominio H_C clostridial ha evolucionado para unirse de manera muy eficiente a receptores presentes en las células neuronales. De acuerdo con esta última realización, la presente invención proporciona el uso y los procedimientos correspondientes para la modificación de BOTOX™ para mejorar su utilidad clínica. A modo de ejemplo, las secuencias de referencia de H_{CC} clostridiales adecuados para la TM incluyen:

neurotoxina botulínica de tipo A - restos aminoacídicos (Y1111-L1296)

neurotoxina botulínica de tipo B - restos aminoacídicos (Y1098-E1291)

neurotoxina botulínica de tipo C - restos aminoacídicos (Y1112-E1291)

neurotoxina botulínica de tipo D - restos aminoacídicos (Y1099-E1276)

neurotoxina botulínica de tipo E - restos aminoacídicos (Y1086-K1252)

neurotoxina botulínica de tipo F - restos aminoacídicos (Y1106-E1274)

neurotoxina botulínica de tipo G - restos aminoacídicos (Y1106-E1297)

neurotoxina tetánica - restos aminoacídicos (Y1128-D1315).

[0042] Las secuencias de referencia identificadas anteriormente deben considerarse como guía, ya que pueden producirse ligeras variaciones según los subserotipos.

[0043] De manera similar, a modo de ejemplo, los dominios Hc clostridiales adecuados para la TM de las secuencias de referencia incluyen BoNT/A - N872-L1296; BoNT/B - E859-E1291; BoNT/C1-N867-E1291; BoNT/D - S863-E1276; BoNT/E - R846-K1252; BoNT/F - K865-E1274; BoNT/G - N864-E1297; y TeNT - I880-D1315.

5 **[0044]** En otra realización, la TM se selecciona para proporcionar una unión deseada a la unión neuromuscular. Las TM adecuadas se listan en el documento WO2006/099590 e incluyen: la hormona similar al glucagón, una neurohormona, una citocina neuroreguladora, una neurotrofina, un factor de crecimiento, una molécula señalizadora para la guía de axones, una proteína de unión a azúcares, un ligando que se une selectivamente a neurexina, un ligando para neurexina 2 α , un ligando para neurexina 2 β , un ligando para neurexina 3 α , un ligando para neurexina 3 β , una WNT, Ng-CAM(LI), NCAM, N-cadherina, un péptido PACAP tal como un péptido VIP, agrina-MUSK, un polipéptido de la membrana basal y una variante de cualquiera de los polipéptidos anteriores, una citocina neuroreguladora tal como el factor neurotrófico ciliar (CNTF), glicoforina A (GPA), el factor inhibidor de la leucemia (LIF), una interleucina (IL), onostatina M, cardiotropina 1 (CT-1), una citocina similar a la cardiotropina (CLC), una neuroleucina, VEGF, un factor de crecimiento similar a la insulina (IGF), un factor de crecimiento epidérmico (EGF) y una variante de cualquiera de las citocinas neuroreguladoras anteriores. Estas y otras TM se seleccionan para uso en la presente invención porque mimetizan la capacidad de unión de las neurotoxinas clostridiales.

10 **[0045]** Según se menciona anteriormente, el(los) sitio(s) de escisión destructivo(s) se introduce(n) con un efecto negativo mínimo sobre las propiedades biológicas del polipéptido (notablemente, la actividad endopeptidásica y/o la actividad de translocación a través de la membrana). A este respecto, se prefiere que cualquier disminución potencial de la potencia del polipéptido (en comparación con el mismo polipéptido sin dicho(s) sitio(s) de escisión destructivo(s)) sea inferior al 25 %, preferentemente inferior al 15 %, con mayor preferencia inferior al 5 % con respecto a la proteína original sin modificar. La potencia aquí puede medirse mediante un ensayo comparativo tal como se ilustra en los ejemplos 1-4.

15 **[0046]** Cuando se selecciona un sitio(s) de escisión destructivo(s) en el contexto de la presente invención, se prefiere que el(los) sitio(s) de escisión destructivo(s) no sea(n) sustrato(s) para ninguna proteasa que pueda usarse separadamente para la modificación postraducciona del polipéptido de la presente invención como parte de su proceso de obtención. A este respecto, las proteasas no citotóxicas de la presente invención emplean típicamente una acción de activación proteásica (a través de un sitio separado de escisión proteásica de "activación", que es estructuralmente distinto del sitio de escisión destructivo de la presente invención). La finalidad del sitio de escisión de activación es escindir un enlace peptídico entre el componente proteásico no citotóxico y el componente de translocación o el componente TM del polipéptido de la presente invención, para proporcionar un polipéptido bcatenario "activado", en el que dichos dos componentes están unidos entre sí por medio de un enlace disulfuro.

20 **[0047]** En una holotoxina clostridial natural, el sitio de escisión proteásica del bucle bcatenario se encuentra en la posición K-448-A449 para BoNT/A, K441-A442 para BoNT/B, K449-T450 para BoNT/C1, R445-D446 para BoNT/D, R422-K423 para BoNT/E, K439-A440 para BoNT/F, K446-S447 para BoNT/G y A457-S458 para TeNT. Por lo tanto, para contribuir a asegurar que el sitio de escisión destructivo de los polipéptidos de la presente invención no afecta negativamente al sitio de escisión de "activación" y la posterior formación de un enlace disulfuro, el primero se introduce preferentemente en el polipéptido de la presente invención en una posición a una distancia de al menos 20, al menos 30, al menos 40, al menos 50 y con mayor preferencia de al menos 60, al menos 70, al menos 80 restos aminoacídicos (contiguos) del sitio de escisión de "activación". A este respecto, el sitio de activación de un polipéptido de la invención puede alinearse fácilmente (mediante alineamiento simple de las secuencias primarias) con las posiciones del sitio de activación (listadas anteriormente) de la holotoxina clostridial.

25 **[0048]** El(los) sitio(s) de escisión destructivo(s) es(son) preferentemente exógeno(s) (es decir, manipulado(s)/artificiale(s)) con respecto a los componentes nativos del polipéptido. En otras palabras, preferentemente dichos sitios de escisión no son intrínsecos a los correspondientes componentes nativos del polipéptido. A modo de ejemplo, un componente proteásico o de translocación basado en las cadenas L o H (respectivamente) de BoNT/A puede manipularse de acuerdo con la presente invención para incluir un sitio(s) de escisión. Sin embargo, dicho(s) sitio(s) de escisión no estarán presentes en las correspondientes cadenas L y H nativas del BoNT.

30 **[0049]** En una realización preferida de la presente invención, el(los) sitio(s) de escisión destructivo(s) y el sitio de escisión de "activación" no son escindidos por la misma proteasa. En una realización, los dos sitios de escisión difieren entre sí en que al menos uno, con mayor preferencia al menos dos, con especial preferencia al menos tres, con la máxima preferencia al menos cuatro de los aminoácidos tolerados dentro de las secuencias de

reconocimiento respectivas es o son diferentes.

[0050] A modo de ejemplo, en el caso de un polipéptido quimérico que contiene un sitio de “activación” para el factor Xa entre los componentes clostridiales de la cadena L y H_N, se prefiere emplear un sitio(s) de escisión destructivo(s) distinto(s) del sitio para el factor Xa, que puede(n) insertarse en cualquier otro lugar en el componente de la cadena L y/o el componente H_N. En este escenario, el polipéptido puede modificarse para acomodar un sitio de “activación” alternativo entre los componentes de la cadena L y H_N (por ejemplo, un sitio de escisión para una enterocinasa), en cuyo caso puede incorporarse un sitio(s) de escisión independiente(s) para el factor Xa en cualquier otro lugar del polipéptido como sitio de escisión destructivo. Alternativamente, puede mantenerse el sitio de “activación” para el factor Xa existente entre los componentes de la cadena L y H_N y puede incorporarse un sitio de escisión alternativo, tal como un sitio de escisión para trombina, como sitio(s) de escisión destructivo(s).

[0051] Al identificar sitios adecuados dentro de la secuencia primaria de cualquiera de los componentes de la presente invención para la inclusión de un sitio(s) de escisión, es preferible seleccionar una secuencia primaria que se asemeje en gran medida al sitio(s) de escisión propuesto(s) que ha(n) de insertarse. Al hacerlo así, se introducen cambios estructurales mínimos en el polipéptido. A modo de ejemplo, los sitios de escisión comprenden típicamente al menos tres restos aminoacídicos contiguos. Por lo tanto, en una realización preferida, se selecciona un sitio de escisión que ya tenga (en la posición o posiciones correctas) al menos uno, preferentemente al menos dos de los restos aminoacídicos que se requieren para introducir el nuevo sitio de escisión. A modo de ejemplo, cuando ha de introducirse el sitio de escisión para caspasa 3 (DMQD), puede identificarse una posición de inserción preferida que ya incluya una secuencia primaria seleccionada, por ejemplo, entre Dxxx, xMxx, xxQx, xxxD, DMxx, DxQx, DxxD, xMQx, xMxD, xxQD, DMQx, xMQD, DxQD y DMxD.

[0052] Al analizar la estructura terciaria de la neurotoxina clostridial, los presentes inventores han identificado una serie de regiones expuestas (en particular regiones de bucle expuestas) adecuadas para la inserción de la(s) secuencia(s) de los sitios destructivos. Este análisis se ha basado principalmente en Chaddock y Marks (2006) en Cell & Molecular Life Sciences, 63: 540-551; con una referencia adicional a http://pathema.tigr.org/pathema/BoNT_structures.shtml; Lacy y Stevens, 1999, J. Mol. Biol., 291: 1091-1104; y la tabla siguiente.

Serotipo de BoNT	ID del PDB	Descripción del PDB
A	1E1H	Estructura cristalina de la cadena ligera de la neurotoxina botulínica de tipo A recombinante, endopeptidasa de Zn autoinhibidora
A	1XTF	Mutante E224Q Y366F de la neurotoxina BoNT/A
A	1XTG	Estructura cristalina de la neurotoxina BoNT/A complejada con la proteína 25 asociada a sinaptosomas
A	3BTA	Estructura cristalina de la neurotoxina botulínica de serotipo A
B	1EPW	Estructura cristalina de la neurotoxina de <i>Clostridium</i> de tipo B
B	1F31	Estructura cristalina de la neurotoxina B de <i>Clostridium botulinum</i> complejada con un trisacárido
B	1F82	Dominio catalítico de la neurotoxina botulínica de tipo B
B	1F83	Dominio catalítico de la neurotoxina botulínica de tipo B unido a sinaptobrevina II
B	1G9A	Estructura cristalina de la neurotoxina B de <i>Clostridium botulinum</i> complejada con un inhibidor (experimento 3)
B	1G9B	Estructura cristalina de la neurotoxina B de <i>Clostridium botulinum</i> complejada con un inhibidor (experimento 1)
B	1G9C	Estructura cristalina de la neurotoxina B de <i>Clostridium botulinum</i> complejada con un inhibidor (experimento 4)
B	1G9D	Estructura cristalina de la neurotoxina B de <i>Clostridium botulinum</i> complejada con un inhibidor (experimento 2)
B	1I1E	Estructura cristalina de la neurotoxina B de <i>Clostridium botulinum</i> complejada con doxorubicina
B	1S0B	Estructura cristalina de la neurotoxina botulínica de tipo B a pH 4,0
B	1S0C	Estructura cristalina de la neurotoxina botulínica de tipo B a pH 5,0
B	1S0D	Estructura cristalina de la neurotoxina botulínica de tipo B a pH 5,5
B	1S0E	Estructura cristalina de la neurotoxina botulínica de tipo B a pH 6,0
B	1S0F	Estructura cristalina de la neurotoxina botulínica de tipo B a pH 7,0
B	1S0G	Estructura cristalina de la neurotoxina botulínica de tipo B, forma apo

B	1Z0H	La hélice N-terminal se reorienta en el fragmento C recombinante de la neurotoxina de <i>Clostridium botulinum</i> de tipo B
B	2ETF	Estructura cristalina de la longitud completa de la cadena ligera de la neurotoxina botulínica (tipo B)
D	2FPQ	Estructura cristalina de la cadena ligera de la neurotoxina botulínica de tipo D
E	1T3A	Estructura cristalina del dominio catalítico de la neurotoxina de <i>Clostridium botulinum</i> de tipo E
E	1T3C	Mutante E212Q del dominio catalítico de la neurotoxina de <i>Clostridium botulinum</i> de tipo E
E	1ZKW	Estructura cristalina del mutante Arg347Ala del dominio catalítico de la neurotoxina botulínica E
E	1ZKX	Estructura cristalina de Glu158Ala/Thr159Ala/Asn160Ala, un mutante triple del dominio catalítico de la neurotoxina E de <i>Clostridium botulinum</i>
E	1ZL5	Estructura cristalina del mutante Glu335Gln del dominio catalítico de la neurotoxina E de <i>Clostridium botulinum</i>
E	1ZL6	Estructura cristalina del mutante Tyr350Ala del dominio catalítico de la neurotoxina E de <i>Clostridium botulinum</i>
E	1ZN3	Estructura cristalina del mutante Glu335Ala de la neurotoxina E de <i>Clostridium botulinum</i>
F	2A8A	Estructura cristalina de la cadena ligera de la neurotoxina de <i>Clostridium botulinum</i> de serotipo F
F	2A97	Estructura cristalina del dominio catalítico de la neurotoxina de <i>Clostridium botulinum</i> de serotipo F
G	1ZB7	Estructura cristalina de la cadena ligera de la neurotoxina botulínica de tipo G

[0053] La identificación PDB anterior se refiere al código de cuatro caracteres usado por el Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (<http://www.rcsb.org/pdb/home/home.do>) para identificar una entrada específica en la base de datos estructural.

5

[0054] Otras técnicas adicionales empleadas incluyen el uso de información del mapeo con péptidos/anticuerpos, por ejemplo, el mapeo con anticuerpos de sitios en la superficie de Hc/A (Dolimbek B. Z., 2007, Mol. Immunol. 44(5): 1029-41), H_N/A (Atassi M. Z., 2004, Protein J. 23(1): 39-52), Hc/A (Oshima M., 1998, Immunol. Lett. 60(1): 7-12; Bavari S. 1998, Vaccine 16(19): 1850-6), Hc/E (Kubota T, 1997, Appl. Environ. Microbiol. 63(4): 1214-8) - una lista de epítomos dentro de los serotipos de BoNT está a disposición pública y se mantiene en http://pathema.tigr.org/pathema/BoNT_epitopes.shtml; y el uso de software de predicción estructural para predecir la accesibilidad de disolventes a una región peptídica específica - algunos ejemplos de software a disposición pública incluyen Swiss Model (<http://swissmodel.expasy.org>), ESyPred3D (<http://www.fundp.ac.be/sciences/biologie/urbm/bioinfo/esypred>) y Geno3D (http://geno3d-pbil.ibcp.fr/cqi-bin/geno3d/automat.pl?page=/GENO3D/geno3d_home.html).

15

[0055] En una realización de la presente invención, el(los) sitio(s) de escisión destructivo(s) se introduce(n) en una o más de las siguientes posiciones, que se basan (por conveniencia) en la secuencia aminoacídica primaria de BoNT/A. Aunque las posiciones de inserción se identifican por referencia a BoNT/A, las secuencias aminoacídicas primarias de los correspondientes dominios proteásico y/o de translocación de BoNT/B-G, etc., pueden alinearse fácilmente con dichas posiciones de BoNT/A; a modo de ejemplo, se menciona el alineamiento de serotipos ilustrado en la figura 2.

20

[0056] Para el componente proteásico, se prefieren una o más de las posiciones siguientes: 27-31, 56-63, 73-75, 78-81, 99-105, 120-124, 137-144, 161-165, 169-173, 187-194, 202-214, 237-241, 243-250, 300-304, 323-335, 375-382, 391-400 y 413-423. La numeración anterior comienza preferentemente a partir del extremo N del componente proteásico de la presente invención. De estas posiciones, las más preferidas son 99-105 y/o 202-214. A este respecto, con referencia a la figura 2, las posiciones 99-105 corresponden a la secuencia "YSTDLGR" en el serotipo A, que equivale a la región "KSKPLGE" en el serotipo B, "NSREIGE" en el serotipo C₁, "NERDIGK" en el serotipo D, "NNNLSGG" en el serotipo E, "NSNPAGQ" en el serotipo F y "NSKPSGQ" en el serotipo G. De manera similar, con referencia a la figura 2, las posiciones 202-214 corresponden a la secuencia "VDTNPLLGAQKFA" en el serotipo A, que equivale a la región "NKGASIFNRRGYF" en el serotipo B, "DVGEGRFKSEFC" en el serotipo C₁, "NQSSAVLGKSIFC" en el serotipo D, "DNC---MN--EFI" en el serotipo E, "DN-TD--LFI" en el serotipo F y "ENKDTSFRRAYF" en el serotipo G. y "P" (202) usando la numeración en la parte superior de la figura 2 como y "P", respectivamente.

30
35

- [0057]** En una realización preferida, el(los) sitio(s) de escisión destructivo(s) se localiza(n) en una posición a más de ocho restos aminoacídicos, preferentemente a más de diez restos aminoacídicos, con mayor preferencia a más de 25 restos aminoacídicos, con especial preferencia a más de 50 restos aminoacídicos del extremo N del componente proteásico. De manera similar, en una realización preferida, el(los) sitio(s) de escisión destructivo(s) se localiza(n) en una posición a más de 20 restos aminoacídicos, preferentemente a más de 30 restos aminoacídicos, con mayor preferencia a más de 40 restos aminoacídicos, con especial preferencia a más de 50 restos aminoacídicos del extremo C del componente proteásico.
- 10 **[0058]** Para el componente de translocación se prefieren una o más de las posiciones siguientes: 474-479, 483-495, 507-543, 557-567, 576-580, 618-631, 643-650, 669-677, 751-767, 823-834 y 845-859. La numeración anterior considera una posición inicial de 449 para el extremo N del componente del dominio de translocación de la presente invención y una posición inicial de 871 para el extremo C del componente H_N. De estas posiciones, las más preferidas son 557-567 y/o 751-767. A este respecto, con referencia a la figura 2, las posiciones 557-567 corresponden a la secuencia "QEFEHGKSRIA" en el serotipo A, que equivale a la región "QTFPLDIRDIS" en el serotipo B, "QKLSDNVEDFT" en el serotipo C1, "QKLSNNVENIT" en el serotipo D, "QKVPEGENNVN" en el serotipo E, "QKAPEGESAIS" en el serotipo F y "QTFPSNIENLQ" en el serotipo G. De manera similar, con referencia a la figura 2, las posiciones 751-767 corresponden a la secuencia "YNQYTEEEKNINNID" en el serotipo A, que equivale a la región "YNIYSEKEKSNIN--IDFN" en el serotipo B, "YKKYSGSDKENIKS--QVE" en el serotipo C1, 15 "YKKYSGSDKENIKS--QVE" en el serotipo D, "YNSYTLEEKNELTNKYDIK" en el serotipo E, "YNNYTLDEKNRLRAEYNIY" en el serotipo F y "YNRYSEEDKMNN--IDFN" en el serotipo G.
- [0059]** En una realización preferida, el(los) sitio(s) de escisión destructivo(s) se localiza(n) en una posición a más de diez restos aminoacídicos, preferentemente a más de 25 restos aminoacídicos, con mayor preferencia a más de 40 restos aminoacídicos, con especial preferencia a más de 50 restos aminoacídicos del extremo N del componente de translocación. De manera similar, en una realización preferida, el(los) sitio(s) de escisión destructivo(s) se localiza(n) en una posición a más de diez restos aminoacídicos, preferentemente a más de 25 restos aminoacídicos, con mayor preferencia a más de 40 restos aminoacídicos, con especial preferencia a más de 50 restos aminoacídicos del extremo C del componente de translocación.
- 30 **[0060]** De acuerdo con un segundo aspecto de la presente invención, se proporciona un polipéptido no citotóxico para uso en el tratamiento de una serie de afecciones y enfermedades diversas. Tales afecciones y enfermedades cuentan con tratamientos establecidos (véase la parte de antecedentes de la presente memoria descriptiva) basados en polipéptidos no citotóxicos estrechamente relacionados (aunque sin modificar de acuerdo con la presente invención). Por consiguiente, la presente invención proporciona mejoras con respecto a dichos tratamientos mediante el uso de un polipéptido no citotóxico modificado que tiene un sitio de escisión destructivo, lo que reduce los efectos fuera de la diana.
- 35 **[0061]** En particular, la presente invención proporciona polipéptidos para uso en el tratamiento del estrabismo, blefaroespasma, bizqueo, distonía espasmódica y bucomandibular, tortícolis y otras aplicaciones de tratamientos de belleza (cosmética) que se benefician de la incapacitación de células/músculos (a través de la regulación por disminución o inactivación de las proteínas SNARE).
- [0062]** Se proporcionan otros tratamientos adicionales relacionados para tratar un trastorno o afección neuromuscular de la motilidad ocular, por ejemplo, estrabismo concomitante y vertical, parálisis del recto lateral, nistagmo, miopatía distiroidea, etc.; distonía, por ejemplo, distonías focales tales como tortícolis espasmódica, calambre de los escritores, blefaroespasma, distonía bucomandibular y los síntomas de las misma, por ejemplo, bruxismo, enfermedad de Wilson, distonía tardía, distonía laríngea, etc.; otras distonías, por ejemplo, temblores, tics, mioclonus segmentario; espasmos, tales como espasticidad debida a esclerosis múltiple crónica, espasticidad que resulta en un control anormal de la vejiga, por ejemplo, en pacientes con lesiones de la médula espinal, anismo, espasmo de la espalda, calambres musculares, etc.; dolores de cabeza tensionales, síndrome del músculo elevador de la pelvis; espina bífida, discinesia tardía; enfermedad de Parkinson y distonía (focal) de miembros y tartamudeo, etc.
- 50 **[0063]** En la práctica, un polipéptido de la invención se une a una estructura superficial (el sitio de unión), que se presenta sobre una célula diana y preferentemente es característico de la misma. Después de la unión, el polipéptido (al menos el componente proteásico del mismo) experimenta endocitosis en una vesícula y el componente de translocación dirige entonces el transporte del componente proteásico a través de la membrana endosómica al citosol de la célula diana. Una vez dentro de la célula diana, la proteasa no citotóxica inhibe los

procesos celulares de fusión exocítica, con lo que se inhibe la liberación/secreción en la célula diana.

[0064] El componente biológicamente activo de los polipéptidos de la presente invención es una proteasa no citotóxica. Las proteasas no citotóxicas son una clase separada de moléculas que no destruyen las células; en vez de ello, actúan inhibiendo procesos celulares distintos de la síntesis de proteínas. Las proteasas no citotóxicas son producidas como parte de una molécula de toxina de mayor tamaño por una diversidad de plantas y por una diversidad de microorganismos tales como *Clostridium* sp. y *Neisseria* sp.

[0065] Las neurotoxinas clostridiales representan un grupo principal de las moléculas de toxinas no citotóxicas y comprenden dos cadenas polipeptídicas unidas entre sí por un enlace disulfuro. Las dos cadenas se denominan cadena pesada (cadena H), que tiene una masa molecular de aproximadamente 100 kDa, y cadena ligera (cadena L), que tiene una masa molecular de aproximadamente 50 kDa. Es la cadena L la que tiene una función proteásica y muestra una gran especificidad de sustrato para las proteínas asociadas a vesículas y/o a la membrana plasmática (SNARE) implicadas en los procesos exocíticos (por ejemplo, sinaptobrevina, syntaxina o SNAP-25). Estos sustratos son componentes importantes de la maquinaria neurosecretora.

[0066] *Neisseria* sp., principalmente la especie *N. gonorrhoeae*, produce moléculas de toxinas no citotóxicas funcionalmente similares. Un ejemplo de una proteasa no citotóxica tal es la proteasa de IgA (véase el documento WO99/58571).

[0067] La elección de la TM determina la especificidad del polipéptido. A modo de ejemplo, el mismo receptor (o similar) puede estar presente en varias células diferentes, de manera que la TM se unirá a diferentes tipos de células. En este escenario, podría ser deseable solamente un único tipo de células como diana. De este modo, al emplear un segundo sitio de escisión proteásica (de "destrucción") en un polipéptido de la presente invención que es escindido por una proteasa específica para una o más de las células no deseadas (y/o para su entorno), es posible minimizar los efectos secundarios fuera de la diana en las células no deseadas.

[0068] En otra realización, los polipéptidos de la presente invención pueden comprender dos o más TM diferentes capaces de unirse a diferentes tipos de células diana. Alternativamente (o adicionalmente), pueden emplearse combinaciones de polipéptidos con diferentes TM para dirigirse coordinadamente a diferentes tipos de células diana.

Preparación de polipéptidos

[0069] Los polipéptidos de la presente invención comprenden cuatro componentes principales: una TM, una proteasa no citotóxica, un dominio de translocación y un sitio destructivo de escisión proteásica. Dichos polipéptidos comprenden holotoxinas no citotóxicas, tales como neurotoxinas clostridiales y, cuando hay una TM presente, quimeras redirigidas (denominadas frecuentemente proteasas redirigidas). La preparación de estas moléculas es convencional; a modo de ejemplo, se mencionan los documentos: WO94/21300, WO96/33273, WO98/07864, WO00/10598, WO01/21213, WO06/059093, WO00/62814, WO00/04926, WO93/15766, WO00/61192 y WO99/58571.

[0070] En más detalle, el componente TM de la presente invención puede fusionarse al componente proteásico o al componente de translocación de la presente invención. Dicha fusión tiene lugar preferentemente por medio de un enlace covalente, por ejemplo, un enlace covalente directo o mediante una molécula espaciadora/enlazante. El componente proteásico y el componente de translocación se unen entre sí preferentemente por medio de un enlace covalente, por ejemplo, un enlace covalente directo o mediante una molécula espaciadora/enlazante. Las moléculas espaciadoras/enlazantes adecuadas son bien conocidas en la técnica y comprenden típicamente una secuencia a base de aminoácidos con una longitud de entre cinco y 40, preferentemente de entre diez y 30 restos aminoácidos.

[0071] En la práctica, los polipéptidos tienen una conformación bicatenaria, en que el componente proteásico y el componente de translocación están unidos entre sí, preferentemente, mediante un enlace disulfuro.

[0072] Los polipéptidos de la presente invención pueden prepararse por técnicas convencionales de conjugación química bien conocidas por el experto en la técnica. A modo de ejemplo, se menciona Hermanson G. T. (1996), *Bioconjugate techniques*, Academic Press, and to Wong, S.S. (1991), *Chemistry of protein conjugation and cross-linking*, CRC Press.

[0073] Alternativamente, los polipéptidos pueden prepararse por preparación recombinante de una única proteína de fusión de polipéptidos (véase, por ejemplo, el documento WO98/07864). Esta técnica se basa en el mecanismo bacteriano *in vivo* por el que se prepara la neurotoxina clostridial nativa (es decir, la holotoxina) y resulta en una proteína de fusión que tiene la siguiente organización estructural "simplificada":

5 NH₂ - [componente proteásico] - [componente de translocación] - [TM] - COOH

[0074] De acuerdo con el documento WO98/07864, la TM se coloca hacia el extremo C-terminal de la proteína de fusión. La proteína de fusión se "activa" entonces por tratamiento con una proteasa que escinde un sitio entre el
10 componente proteásico y el componente de translocación. De este modo se produce una proteína bicatenaria que comprende el componente proteásico como una cadena polipeptídica independiente unida (mediante un puente disulfuro) a otra cadena polipeptídica independiente que contiene el componente de translocación y la TM.

[0075] El sistema del documento WO98/07864 es especialmente adecuado para la preparación de proteínas
15 de fusión con una TM que requiere un dominio C-terminal que esté "libre" para su interacción con un sitio de unión en una célula diana.

[0076] Para las proteínas de fusión con una TM que requiere un dominio N-terminal que esté "libre" para su interacción con un sitio de unión en una célula diana, puede emplearse un sistema modificado según se describe en
20 el documento WO06/059113.

[0077] En el sistema modificado, el componente TM de la proteína de fusión se localiza hacia la mitad de la secuencia lineal de la proteína de fusión, entre el sitio de escisión proteásica y el componente de translocación. Esto asegura que la TM está asociada al dominio de translocación (es decir, al igual que en la holotoxina clostridial
25 nativa), aunque en este caso, los dos componentes están en orden invertido con respecto a la holotoxina nativa. La posterior escisión en el sitio de escisión proteásica expone la porción N-terminal de la TM y proporciona la proteína de fusión de polipéptidos bicatenaria.

[0078] La(s) secuencia(s) de escisión proteásica mencionada(s) anteriormente puede(n) introducirse (y/o
30 eliminar cualquier secuencia de escisión intrínseca) como ADN por procedimientos convencionales, tales como la mutagénesis dirigida. El análisis para confirmar la presencia de las secuencias de escisión puede llevarse a cabo manualmente o con la ayuda de un software informático (por ejemplo, el programa MapDraw de DNASTAR, Inc.). Aunque puede emplearse cualquier sitio de escisión proteásica, (es decir, clostridial o no clostridial), se prefieren los siguientes (bien como el sitio de escisión "destrutivo" o como el sitio de escisión de "activación"):

35 Enterocinasa (DDDDK↓)
Factor Xa (IEGR↓ / IDGR↓)
40 TEV (virus del grabado del tabaco) (ENLYFQ↓G)
Trombina (LVPR↓GS)
Proteasa PreScission (LEVLFQ↓GP).

45 **[0079]** El término sitio de escisión proteásica incluye también una inteína que es una secuencia de autoescisión. La reacción de autoescisión es controlable, por ejemplo, modificando la concentración del agente reductor presente.

50 **[0080]** En una realización preferida, la proteína de fusión de la presente invención puede comprender una o más etiquetas de purificación localizadas en el extremo N y/o en el extremo C. Aunque puede emplearse cualquier etiqueta de purificación, se prefieren las siguientes:

His (por ejemplo, 6x histidina), preferentemente como etiqueta C-terminal y/o N-terminal
55 MBP (proteína de unión a maltosa), preferentemente como etiqueta N-terminal
GST (glutación-S-transferasa), preferentemente como etiqueta N-terminal

His-MBP, preferentemente como etiqueta N-terminal

GST-MBP, preferentemente como etiqueta N-terminal

5 Tiorredoxina, preferentemente como etiqueta N-terminal

CBD (dominio de unión a quitina), preferentemente como etiqueta N-terminal.

10 **[0081]** En la proteína de fusión pueden incluirse una o más moléculas peptídicas espaciadoras/enlazantes. Por ejemplo, puede emplearse un péptido espaciador entre una etiqueta de purificación y el resto de la molécula de la proteína de fusión.

[0082] Por lo tanto, un tercer aspecto de la presente invención proporciona una secuencia de ácido nucleico (por ejemplo, ADN) que codifica un polipéptido según se describe anteriormente.

15 **[0083]** Dicho ácido nucleico puede incluirse en forma de vector, tal como un plásmido, que puede incluir opcionalmente uno o más de entre un origen de replicación, un sitio de integración en un ácido nucleico, un promotor, un terminador y un sitio de unión a ribosomas.

20 **[0084]** La presente invención incluye también un procedimiento para la expresión de la secuencia de ácido nucleico descrita anteriormente (es decir, el tercer aspecto de la presente invención) en una célula hospedadora, en particular en *E. coli* o por medio de un sistema de expresión de baculovirus.

25 **[0085]** La presente descripción incluye además un procedimiento para la activación de un polipéptido de la presente invención, en que dicho procedimiento comprende la puesta en contacto del polipéptido con una proteasa que escinde el polipéptido en un sitio de reconocimiento (sitio de escisión) localizado entre el componente proteásico no citotóxico y el componente de translocación, lo que convierte al polipéptido en un polipéptido bicatenario en el que el componente proteásico no citotóxico y el componente de translocación están unidos entre sí por un enlace disulfuro. En una realización preferida, el sitio de reconocimiento no es nativo de una neurotoxina clostridial de
30 origen natural ni/o de una proteasa de IgA de origen natural.

Administración del polipéptido

35 **[0086]** En la práctica, la presente invención emplea una composición farmacéutica que comprende un polipéptido, junto con al menos un componente seleccionado de entre un vehículo farmacéuticamente aceptable, un excipiente, un adyuvante, un propulsor y/o una sal.

40 **[0087]** Los polipéptidos de la presente invención pueden formularse para administración por vía oral, parenteral, infusión continua, inhalación o tópica. Las composiciones adecuadas para inyección pueden estar en forma de disoluciones, suspensiones o emulsiones, o polvos secos que se disuelven o suspenden en un vehículo adecuado antes de su uso.

[0088] En el caso de un polipéptido que ha de administrarse localmente, el polipéptido puede formularse como crema (por ejemplo, para administración por vía tópica) o para inyección subdérmica.

45 **[0089]** El medio de administración local puede incluir un aerosol u otro pulverizador (por ejemplo, un nebulizador). A este respecto, la formulación de un polipéptido como aerosol permite su administración a los pulmones y/o otras vías nasales y/o bronquiales o respiratorias.

50 **[0090]** Los polipéptidos de la invención pueden administrarse a un paciente por inyección intratecal o epidural en la columna vertebral al nivel del segmento medular implicado en la innervación del órgano afectado.

[0091] Una vía de administración preferida es mediante inyección laparoscópica y/o localizada, especialmente intramuscular.

55 **[0092]** En el caso de las formulaciones para inyección, es opcional incluir una sustancia farmacéuticamente activa para contribuir a la retención o reducir la eliminación del polipéptido del sitio de administración. Un ejemplo de una sustancia activa tal es un vasoconstrictor tal como la adrenalina. Una formulación tal confiere la ventaja de aumentar el tiempo de permanencia del polipéptido después de su administración, lo que aumenta y/o potencia su

efecto.

- [0093]** Los intervalos de dosificación para la administración de los polipéptidos de la presente invención son aquellos que producen el efecto terapéutico deseado. Ha de apreciarse que el intervalo de dosificación requerido depende de la naturaleza concreta del polipéptido o la composición, la vía de administración, la naturaleza de la formulación, la edad del paciente, la naturaleza, extensión o gravedad de la afección del paciente, contraindicaciones, si es el caso, y el criterio del médico a cargo. Las variaciones en estos niveles de dosificación pueden ajustarse mediante procedimientos habituales empíricos de optimización.
- [0094]** Las dosis diarias adecuadas (por kg de peso del paciente) están en el intervalo 0,0001-1 ng/kg, preferentemente 0,0001-0,5 ng/kg, con mayor preferencia 0,002-0,5 ng/kg y con preferencia especial 0,004-0,5 ng/kg. La unidad de dosificación puede variar desde menos de 1 pg hasta 30 ng, pero típicamente estará en el intervalo de 0,01 a 1 ng por dosis, que puede administrarse diariamente o preferentemente con menor frecuencia, tal como semanal o semestralmente.
- [0095]** Un régimen de dosificación especialmente preferido se basa en 0,25 ng del polipéptido como dosis 1X. A este respecto, las dosis preferidas están en el intervalo 1X-100X (es decir 0,25-25 ng).
- [0096]** Las formas de dosificación líquidas se preparan típicamente utilizando el polipéptido y un vehículo estéril apirógeno. Dependiendo del vehículo y la concentración usados, el polipéptido puede disolverse o suspenderse en el vehículo. Al preparar las disoluciones, el polipéptido puede disolverse en el vehículo, la disolución puede hacerse isotónica si es necesario mediante la adición de cloruro de sodio y esterilizarse por filtración a través de un filtro estéril, usando técnicas asépticas, antes de introducirla en viales o ampollas estériles adecuados que después se cierran. Alternativamente, si la estabilidad de la disolución es adecuada, la disolución puede esterilizarse en autoclave en los envases cerrados. Ventajosamente, pueden disolverse en el vehículo aditivos tales como agentes tamponantes, solubilizantes, estabilizantes, conservantes o bactericidas, de suspensión o emulsionantes y/o agentes anestésicos locales.
- [0097]** Los polvos secos que se disuelven o suspenden en un vehículo adecuado antes de su uso pueden prepararse introduciendo los ingredientes previamente esterilizados en un envase estéril mediante técnicas asépticas en un área estéril. Alternativamente, los ingredientes pueden disolverse en los envases adecuados mediante técnicas asépticas en un área estéril. A continuación, el producto se liofiliza y los envases se cierran asépticamente.
- [0098]** Las suspensiones para vía parenteral, adecuadas para inyección intramuscular, subcutánea o intradérmica se preparan sustancialmente de la misma manera, excepto porque los componentes estériles se suspenden en el vehículo estéril, en lugar de disolverse y la esterilización no puede llevarse a cabo por filtración. Los componentes pueden aislarse en estado estéril o alternativamente, pueden esterilizarse después de su aislamiento, por ejemplo, mediante radiación gamma.
- [0099]** Ventajosamente, se incluye un agente de suspensión, por ejemplo, polivinilpirrolidona, en las composiciones para facilitar la distribución uniforme de los componentes.
- [0100]** La administración de acuerdo con la presente invención puede aprovechar diversas tecnologías de administración, incluida la encapsulación en micropartículas, los sistemas de administración mediante virus y el impacto de aerosoles a alta presión.

Sección de definiciones

- [0101]** La fracción de direccionamiento (TM) indica cualquier estructura química que interacciona funcionalmente con un sitio de unión para causar una asociación física entre el polipéptido de la invención y la superficie de una célula diana. El término TM abarca cualquier molécula (es decir, una molécula de origen natural o una variante de la misma modificada química o físicamente) que es capaz de unirse al sitio de unión en la célula diana, en que el sitio de unión es capaz de sufrir internalización (por ejemplo, formación de endosomas), también denominada endocitosis mediada por receptores. La TM puede tener una función de translocación a través de la membrana endosómica, en cuyo caso no es necesario que haya una TM y un dominio de translocación como componentes independientes en un agente de la presente invención. A lo largo de la descripción precedente, se han descrito TM específicas. La mención de dichas TM es solamente ejemplar y la presente invención abarca todas las variantes y derivados de las mismas que retienen la capacidad de unión (es decir, direccionamiento) básica de las

TM ejemplificadas.

[0102] Según se menciona previamente, las TM preferidas incluyen anticuerpos (por ejemplo, fragmentos de anticuerpos) y andamiajes de unión; especialmente anticuerpos/fragmentos y andamiajes disponibles comercialmente diseñados para su unión (por ejemplo, específicamente) a células nerviosas.

[0103] Los andamiajes proteínicos representan una nueva generación de armazones de unión universales para complementar la gama en expansión de anticuerpos monoclonales terapéuticos y derivados tales como scFv, moléculas Fab, dAb (anticuerpos de dominio único), diacuerpos y minicuerpos, cada uno de los cuales puede emplearse como TM de la presente invención. Los sistemas de andamiaje crean o modifican dominios de reconocimiento de proteínas conocidos mediante la creación de nuevos andamiajes o la modificación de dominios de unión a proteínas conocidos. Tales andamiajes incluyen, pero no se limitan a:

(i) andamiajes basados en la proteína A - aficuerpos (Nord K. y col. 1997 "Binding proteins selected from combinatorial libraries of an alpha-helical bacterial receptor domain". Nat. Biotechnol. 15: 772-777);

(ii) andamiajes basados en lipocaína - anticalinas (Skerra 2008 "Alternative binding proteins: anticalins - harnessing the structural plasticity of the lipocalin ligand pocket to engineer novel binding activities". FEBS J. 275: 2677-83);

(iii) andamiajes basados en fibronectina - adnectina (Dineen y col. 2008 "The Adnectin CT-322 is a novel VEGF receptor 2 inhibitor that decreases tumor burden in an orthotopic mouse model of pancreatic cancer". BMC Cancer 8: 352);

(iv) avímeros (Silverman y col. 2005 "Multivalent avimer proteins evolved by exon shuffling of a family of human receptor domains". Nat. Biotechnol. 23: 1556-61);

(v) andamiajes basados en ankirina - darpinas (Zahnd y col. 2006 "Selection and characterization of Her2 binding-designed ankyrin repeat proteins". J. Biol. Chem. 281: 35167-75); y

(vi) andamiajes de centirina - basados en un pliegue proteínico que tiene una homología estructural significativa con los dominios Ig, con bucles que son análogos a las CDR. Los dominios de Ig tienen un módulo común en las proteínas humanas y se han usado ampliamente como proteínas de andamiaje alternativas. Cada una de las publicaciones anteriores sobre "andamiajes" se incorpora en este documento (en su totalidad) por referencia a las mismas.

[0104] Los andamiajes de unión pueden usarse para el direccionamiento a tipos de células concretos a través de la interacción con proteínas específicas de la superficie celular, receptores u otros epítomos de la superficie celular tales como grupos de azúcares. Tales andamiajes modificados pueden manipularse para obtener polipéptidos recombinantes basados en proteasas no citotóxicas de la presente invención que se dirijan a tipos específicos de células nerviosas de interés.

[0105] La TM de la presente invención se une (preferentemente, se une específicamente) a la célula diana en cuestión. El término "se une específicamente" indica preferentemente que una TM dada se une a la célula diana con una afinidad de unión (K_a) de 10^6 M^{-1} o mayor, preferentemente de 10^7 M^{-1} o mayor, con mayor preferencia de 10^8 M^{-1} o mayor y con la máxima preferencia de 10^9 M^{-1} o mayor.

[0106] La referencia a una TM en la presente memoria descriptiva abarca fragmentos y variantes de la misma que retienen la capacidad de unirse a la célula diana en cuestión. A modo de ejemplo, una variante puede tener al menos el 80 %, preferentemente al menos el 90 %, con mayor preferencia al menos el 95 % y con la máxima preferencia al menos el 97 o al menos el 99 % de homología en la secuencia aminoacídica con la TM de referencia. Por lo tanto, una variante puede incluir uno o más análogos de un aminoácido (por ejemplo, un aminoácido no natural) o un enlace sustituido. Además, a modo de ejemplo, el término fragmento, cuando se usa en relación con una TM, indica un péptido que tiene al menos diez, preferentemente al menos 20, con mayor preferencia al menos 30 y con la máxima preferencia al menos 40 restos aminoacídicos de la TM de referencia. El término fragmento también se refiere a las variantes mencionadas anteriormente. Por lo tanto, a modo de ejemplo, un fragmento de la presente invención puede comprender una secuencia peptídica con al menos diez, 20, 30 o 40 aminoácidos, en que la secuencia peptídica tiene al menos el 80 % de homología de secuencia con respecto a la correspondiente secuencia peptídica de aminoácidos (contiguos) del péptido de referencia.

[0107] A modo de ejemplo, las TM de péptidos ErbB (por ejemplo, EGF) pueden modificarse para generar ligandos que son muteínas de ErbB con propiedades mejoradas tales como una mejor estabilidad. A modo de ejemplo, las TM muteínas de ErbB incluyen péptidos ErbB con modificaciones aminoacídicas tales como un resto de valina en la posición 46 o 47 (EGFVal46 o 47), que confieren estabilidad a las proteasas celulares. Las TM de ErbB pueden tener también aminoácidos eliminados o aminoácidos adicionales insertados. Esto incluye, pero no se limita a un EGF con una delección de los dos aminoácidos C-terminales y la sustitución neutra de un aminoácido en la posición 51 (en particular EGF51Gln51; véase el documento US20020098178A1) y un EGF con aminoácidos eliminados (por ejemplo, rEGF2-48, rEGF3-48 y rEGF4-48). Las TM que son fragmentos de ErbB pueden incluir fragmentos de TGF α que contienen regiones de hélice β pronosticada (por ejemplo, un péptido de la secuencia Ac-C-H-S-G-Y-V-G-A-R-C-O-OMe), fragmentos de EGF tales como [Ala20]EGF(14-31), y el péptido YHWYGYTPQNV1 o GE11. Todas las anteriores memorias de patente se incorporan en este documento por referencia a las mismas.

[0108] La confirmación de que una TM se une a la célula diana seleccionada en una cuestión rutinaria. Por ejemplo, puede emplearse un simple experimento de desplazamiento radiactivo en el que el tejido o las células representativas de la célula diana se exponen a la TM marcada (por ejemplo, tritiada) en presencia de un exceso de TM sin marcar. En un experimento tal, pueden determinarse las proporciones relativas de unión inespecífica y específica, lo que permite la confirmación de que la TM se une a la célula diana. Opcionalmente, el ensayo puede incluir uno o más antagonistas de la unión y el ensayo puede comprender además la observación de una pérdida de unión de la TM. Algunos ejemplos de este tipo de experimento pueden encontrarse en Hulme E. C. (1990), Receptor-binding studies, a brief outline, págs. 303-311, en Receptor biochemistry, A Practical Approach, ed. E. C. Hulme, Oxford University Press.

[0109] En el contexto de la presente invención, la referencia a una TM peptídica abarca los análogos peptídicos de la misma, siempre que el análogo se una al mismo receptor que la correspondiente TM de "referencia". Dichos análogos pueden incluir restos sintéticos tales como:

β -Nal = β -naftilalanina

β -Pal = β -piridilalanina

hArg(Bu) = N-guanidino(butil)homoarginina

hArg(Et)₂ = N,N'-guanidino(dimetil)homoarginina

hArg(CH₂CF₃)₂ = N,N'-guanidinobis-(2,2,2-trifluoroetil)homoarginina

hArg(CH₃, hexyl) = N,N'-guanidino(metil, hexil)homoarginina

Lys(Me) = N^e-metillisina

Lys(iPr) = N^e-isopropillisina

AmPhe = aminometilfenilalanina

AChxAla = aminociclohexilalanina

Abu = ácido α -aminobutírico

Tpo = 4-tiaproline

MeLeu = N-metilleucina

Orn = ornitina

Nle = norleucina

Nva = norvalina

Trp(Br) = 5-bromotriptófano

Trp(F) = 5-fluorotriptófano

Trp(NO₂) = 5-nitrotriptófano

5

Gaba = ácido γ-aminobutírico

Bmp = J-mercaptopropionilo

10 Ac = acetilo

Pen = pencilamina

[0110] Los polipéptidos de la presente invención pueden carecer de un dominio H_C (o H_{CC}) funcional de una neurotoxina clostridial, en cuyo caso, típicamente hay una TM no clostridial presente para unir el polipéptido a un sitio de unión en la célula nerviosa. El péptido H_C de una neurotoxina clostridial nativa comprende aproximadamente 400-440 restos aminoacídicos y consta de dos dominios funcionalmente distintos, cada uno de aproximadamente 25 kDa, es decir, la región N-terminal (denominada normalmente péptido o dominio H_{CN}) y la región C-terminal (denominada normalmente péptido o dominio H_{CC}). Se ha documentado detalladamente que la región C-terminal (H_{CC}), que consta de los 160-200 restos aminoacídicos C-terminales, es responsable de la unión de una neurotoxina clostridial a sus receptores celulares naturales, es decir, a las terminaciones nerviosas en la unión neuromuscular; este hecho ha sido también confirmado por las publicaciones anteriores. Por lo tanto, la referencia a lo largo de esta memoria descriptiva a una cadena pesada clostridial que carece de un péptido (o dominio) funcional H_C de la cadena pesada indica que la cadena pesada clostridial simplemente carece de un péptido H_{CC} funcional. En otras palabras, la región peptídica H_{CC} se ha eliminado parcial o totalmente o se ha modificado de otra manera (por ejemplo, mediante un tratamiento químico o proteolítico convencional) para inactivar su capacidad de unión nativa a las células nerviosas.

[0111] Alternativamente, un polipéptido de la presente invención puede contener un dominio H_C (o H_{CC}) de una neurotoxina clostridial como TM. Diversas regiones H_{CC} o H_C de neurotoxinas clostridiales que comprenden un dominio de unión pueden ser útiles en aspectos de la presente invención, con la condición de que estos fragmentos activos proporcionen la actividad de unión y la especificidad de unión de la neurotoxina natural. Las regiones H_C de las cadenas pesadas de las toxinas clostridiales tienen una longitud de aproximadamente 400-440 aminoácidos y comprenden un dominio de unión. La investigación ha demostrado que no es necesaria la longitud total de la región H_C de la cadena pesada de una toxina clostridial para la actividad de unión del dominio de unión. Por lo tanto, algunos aspectos de esta realización pueden incluir regiones H_C de toxinas clostridiales que comprenden un dominio de unión con una longitud, por ejemplo, de al menos 350 aminoácidos, al menos 375 aminoácidos, al menos 400 aminoácidos y al menos 425 aminoácidos. Otros aspectos de esta realización pueden incluir regiones H_C de toxinas clostridiales que comprenden un dominio de unión con una longitud, por ejemplo, de 350 aminoácidos como máximo, 375 aminoácidos como máximo, 400 aminoácidos como máximo o 425 aminoácidos como máximo.

[0112] La proteasa de la presente invención abarca todas las proteasas no citotóxicas que son capaces de escindir una o más proteínas del aparato de fusión exocítica en células eucariotas.

[0113] La proteasa de la presente invención es preferentemente una proteasa bacteriana (o un fragmento de la misma). Con mayor preferencia, la proteasa bacteriana se selecciona de los géneros *Clostridium* o *Neisseria/Streptococcus* (por ejemplo, una cadena L clostridial o una proteasa de IgA neiserial, preferentemente de *N. gonorrhoeae* o *S. pneumoniae*).

[0114] La presente invención también abarca proteasas no citotóxicas variantes (es decir, variantes de moléculas de proteasa de origen natural), siempre que las proteasas variantes aún demuestren la actividad proteásica requerida. A modo de ejemplo, una variante puede tener al menos el 70 %, preferentemente al menos el 80 %, con mayor preferencia al menos el 90 % y con la máxima preferencia al menos el 95 o al menos el 98 % de homología de secuencia aminoacídica con la secuencia de una proteasa de referencia. Por lo tanto, el término variante incluye proteasas no citotóxicas con una mayor (o menor) actividad endopeptidásica; se hace especial mención al aumento de la K_{cat}/K_m de los mutantes de BoNT/A Q161A, E54A y K165L, véase Ahmed S. A. (2008) Protein J. DOI 10.1007/s10930-007-9118-8. El término fragmento, cuando se usa en relación con una proteasa indica típicamente un péptido que tiene al menos 150, preferentemente al menos 200, con mayor preferencia al menos 250 y con la máxima preferencia al menos 300 restos aminoacídicos de la proteasa de referencia. Al igual

que con el componente TM que es un "fragmento" (expuesto anteriormente), los "fragmentos" de proteasa de la presente invención abarcan fragmentos de proteasas variantes basados en una secuencia de referencia.

[0115] La proteasa de la presente invención demuestra preferentemente una actividad serina o metaloproteásica (por ejemplo, actividad endopeptidásica). La proteasa es preferentemente específica para una proteína SNARE (por ejemplo, SNAP-25, sinaptobrevina/VAMP o syntaxina).

[0116] Se mencionan especialmente los dominios proteásicos de las neurotoxinas, por ejemplo, los dominios proteásico de neurotoxinas bacterianas. Por lo tanto, la presente invención abarca el uso de dominios de neurotoxinas de origen natural, así como versiones preparadas de manera recombinante de dichas neurotoxinas de origen natural.

[0117] Algunas neurotoxinas ejemplares son producidas por clostridios y el término neurotoxina clostridial abarca las neurotoxinas producidas por *C. tetani* (TeNT) y por *C. botulinum* (BoNT) de los serotipos A-G, así como las neurotoxinas estrechamente relacionadas similares a BoNT producidas por *C. baratii* y *C. butyricum*. Las abreviaturas mencionadas anteriormente se usan a lo largo de la presente memoria descriptiva. Por ejemplo, la nomenclatura BoNT/A denota la fuente de la neurotoxina como BoNT (serotipo A). A los otros serotipos de BoNT se aplica la nomenclatura correspondiente.

[0118] Las BoNT son las toxinas más potentes conocidas, con valores de dosis letal mediana (DL50) para ratones en el intervalo de 0,5 a 5 ng/kg, según el serotipo. Las BoNT son absorbidas por el tubo gastrointestinal y, después de entrar en la circulación general, se unen a la membrana presináptica de las terminaciones nerviosas colinérgicas e impiden la liberación de su neurotransmisor acetilcolina. BoNT/B, BoNT/D, BoNT/F y BoNT/G escinden la sinaptobrevina o proteína de membrana asociada a vesículas (VAMP); BoNT/C, BoNT/A y BoNT/E escinden la proteína asociada a sinaptosomas de 25 kDa (SNAP-25); y BoNT/C escinde la syntaxina.

[0119] Las BoNT tiene una estructura común y son proteínas bicatenarias de aproximadamente 150 kDa, que constan de una cadena pesada (cadena H) de 100 kDa unida covalentemente mediante un solo enlace disulfuro a una cadena ligera (cadena L) de aproximadamente 50 kDa. La cadena H consta de dos dominios, cada uno de ellos de aproximadamente 50 kDa. El dominio C-terminal (H_C) se requiere para la unión neuronal de alta afinidad, mientras que se sugiere que el dominio N-terminal (H_N) está implicado en la translocación a través de la membrana. La cadena L es una metaloproteasa dependiente de cinc responsable de la escisión de la proteína SNARE sustrato.

[0120] El término fragmento de la cadena L indica un componente de la cadena L de una neurotoxina, en que el fragmento demuestra actividad de metaloproteásica y es capaz de escindir proteolíticamente una proteína asociada a vesículas y/o a la membrana plasmática implicada en la exocitosis celular.

[0121] Algunos ejemplos de secuencias de proteasas (de referencia) adecuadas incluyen:

40	neurotoxina botulínica de tipo A	- restos aminoacídicos (1-448)
	neurotoxina botulínica de tipo B	- restos aminoacídicos (1-440)
	neurotoxina botulínica de tipo C	- restos aminoacídicos (1-441)
	neurotoxina botulínica de tipo D	- restos aminoacídicos (1-445)
	neurotoxina botulínica de tipo E	- restos aminoacídicos (1-422)
45	neurotoxina botulínica de tipo F	- restos aminoacídicos (1-439)
	neurotoxina botulínica de tipo G	- restos aminoacídicos (1-441)
	neurotoxina tetánica	- restos aminoacídicos (1-457)
	proteasa de IgA	- restos aminoacídicos (1-959)*

*Pohlner J. y col. (1987) Nature 325: 458-462.

[0122] Las secuencias de referencia identificadas anteriormente deben considerarse como guía, ya que pueden producirse ligeras variaciones según los subserotipos. A modo de ejemplo, el documento US 2007/0166332 cita secuencias clostridiales ligeramente diferentes:

55	neurotoxina botulínica de tipo A	- restos aminoacídicos (M1-K448)
	neurotoxina botulínica de tipo B	- restos aminoacídicos (M1-K441)
	neurotoxina botulínica de tipo C	- restos aminoacídicos (M1-K449)
	neurotoxina botulínica de tipo D	- restos aminoacídicos (M1-R445)

neurotoxina botulínica de tipo E	- restos aminoacídicos (M1-R422)
neurotoxina botulínica de tipo F	- restos aminoacídicos (M1-K439)
neurotoxina botulínica de tipo G	- restos aminoacídicos (M1-K446)
neurotoxina tetánica	- restos aminoacídicos (M1-A457)

5

[0123] Una diversidad de fragmentos de toxinas clostridiales que comprenden la cadena ligera pueden ser útiles en aspectos de la presente invención, con la condición de que estos fragmentos de la cadena ligera puedan dirigirse específicamente a los componentes centrales del aparato de liberación de neurotransmisores y así participar en ejecutar el mecanismo celular general por el que una toxina clostridial escinde proteolíticamente un sustrato. Las cadenas ligeras de las toxinas clostridiales tienen una longitud de aproximadamente 420-460 aminoácidos y comprenden un dominio enzimático. La investigación ha demostrado que no es necesaria la longitud total de la cadena ligera de una toxina clostridial para la actividad enzimática del dominio enzimático. Como ejemplo no limitante, los primeros ocho aminoácidos de la cadena ligera de BoNT/A no se requieren para la actividad enzimática. Como otro ejemplo no limitante, los primeros ocho aminoácidos de la cadena ligera de TeNT no se requieren para la actividad enzimática. Igualmente, el extremo carboxilo de la cadena ligera no es necesario para la actividad. Como ejemplo no limitante, los últimos 32 aminoácidos de la cadena ligera de BoNT/A (restos 417-448) no se requieren para la actividad enzimática. Como otro ejemplo no limitante, los últimos 31 aminoácidos de la cadena ligera de TeNT (restos 427-457) no se requieren para la actividad enzimática. Por lo tanto, algunos aspectos de esta realización pueden incluir cadenas ligeras de toxinas clostridiales que comprenden un dominio enzimático con una longitud, por ejemplo, de al menos 350 aminoácidos, al menos 375 aminoácidos, al menos 400 aminoácidos, al menos 425 aminoácidos y al menos 450 aminoácidos. Otros aspectos de esta realización pueden incluir cadenas ligeras de toxinas clostridiales que comprenden un dominio enzimático con una longitud, por ejemplo, de 350 aminoácidos como máximo, de 375 aminoácidos como máximo, de 400 aminoácidos como máximo, de 425 aminoácidos como máximo y de 450 aminoácidos como máximo.

25

[0124] Los polipéptidos de la presente invención, especialmente el componente proteásico de los mismos, pueden estar PEGilados; esto puede contribuir a aumentar la estabilidad, por ejemplo, la duración de la acción del componente proteásico. La PEGilación se prefiere especialmente cuando la proteasa comprende una proteasa de BoNT/A, B o C₁. La PEGilación incluye preferentemente la adición de PEG al extremo N del componente proteásico. A modo de ejemplo, el extremo N de una proteasa puede extenderse con uno o más restos aminoacídicos (por ejemplo, cisteína) que pueden ser iguales o diferentes. Uno o más de dichos restos aminoacídicos puede tener su propia molécula de PEG unida (por ejemplo, unida covalentemente) al mismo. Un ejemplo de esta tecnología se describe en el documento WO2007/104567, que se incorpora en su totalidad por referencia al mismo.

[0125] Un dominio de translocación es una molécula que permite la translocación de una proteasa al interior de una célula diana, de manera que tiene lugar una expresión funcional de la actividad proteásica dentro del citosol de la célula diana. Si una molécula cualquiera (por ejemplo, una proteína o un péptido) tiene la función de translocación de la presente invención requerida puede confirmarse mediante uno cualquiera de una serie de ensayos convencionales.

40

[0126] Por ejemplo, Shone C. (1987) describe un ensayo *in vitro* con el empleo de liposomas que se exponen a la molécula de prueba. La presencia de la función de translocación requerida se confirma por la liberación de K⁺ y/o NAD marcado de los liposomas, lo cual puede monitorizarse fácilmente [véase Shone C. (1987) Eur. J. Biochem.; vol. 167(1): págs. 175-180].

45

[0127] Otro ejemplo más lo proporciona Blaustein R. (1987), que describe un simple ensayo *in vitro* con el empleo de membranas planas de bicapas de fosfolípidos. Las membranas se exponen a la molécula de prueba y la función de translocación requerida se confirma por un aumento de la conductancia a través de dichas membranas [véase Blaustein (1987) FEBS Letts.; vol. 226, n.º 1: págs. 115-120].

50

[0128] Una metodología adicional para hacer posible la evaluación de la fusión membranaria y de este modo identificar dominios de translocación para uso en la presente invención se proporciona en Methods in Enzymology vol. 220 y 221, Membrane Fusion Techniques, partes A y B, Academic Press 1993.

[0129] La presente invención también abarca dominios de translocación variantes, siempre que los dominios variantes aún demuestren la actividad de translocación requerida. A modo de ejemplo, una variante puede tener al menos el 70 %, preferentemente al menos el 80 %, con mayor preferencia al menos el 90 % y con la máxima preferencia al menos el 95 % o al menos el 98 % de homología de secuencia aminoacídica con un dominio de translocación de referencia. El término fragmento, cuando se usa en relación con un dominio de translocación, indica

un péptido con al menos 20, preferentemente al menos 40, con mayor preferencia al menos 80 y con la máxima preferencia al menos 100 restos aminoacídicos del dominio de translocación de referencia. En el caso de un dominio de translocación clostridial, el fragmento tiene preferentemente al menos 100, preferentemente al menos 150, con mayor preferencia al menos 200 y con la máxima preferencia al menos 250 restos aminoacídicos del dominio de translocación de referencia (por ejemplo, el dominio H_N). Al igual que el componente TM que es un "fragmento" (expuesto anteriormente), los "fragmentos" de translocación de la presente invención abarcan fragmentos de dominios de translocación variantes basados en las secuencias de referencia.

[0130] Preferentemente, el dominio de translocación es capaz de formar poros permeables a iones en membranas lipídicas en condiciones de bajo pH. Ha resultado preferible el uso solo de aquellas porciones de la molécula de proteína capaces de la formación de poros dentro de la membrana endosómica.

[0131] El dominio de translocación puede obtenerse de una fuente de proteínas microbiana, en particular de una fuente de proteínas bacteriana o vírica. Por consiguiente, en una realización, el dominio de translocación es un dominio de translocación de una enzima, tal como una toxina bacteriana o una proteína vírica.

[0132] Se ha documentado detalladamente que ciertos dominios de moléculas de toxinas bacterianas son capaces de formar tales poros. También se sabe que ciertos dominios de translocación de proteínas de fusión membranaria expresadas por virus son capaces de formar tales poros. Tales dominios pueden emplearse en la presente invención.

[0133] El dominio de translocación puede ser de origen clostridial, tal como el dominio H_N (o un componente funcional del mismo). H_N indica una porción o fragmento de la cadena H de una neurotoxina clostridial aproximadamente equivalente a la mitad aminoterminal de la cadena H o el dominio correspondiente a este fragmento en la cadena H intacta. La cadena H carece de la función de unión natural del componente H_C de la cadena H. A este respecto, la función de H_C puede eliminarse por delección de la secuencia aminoacídica de H_C (en la síntesis del ADN o después de la síntesis del ADN, por un tratamiento con nucleasas o proteasas). Alternativamente, la función de H_C puede inactivarse por un tratamiento químico o biológico. Por lo tanto, la cadena H es incapaz de unirse al sitio de unión en una célula diana al que se une la neurotoxina clostridial nativa (es decir, la holotoxina).

[0134] Algunos ejemplos de dominios de translocación adecuados (de referencia) incluyen:

35	neurotoxina botulínica de tipo A	- restos aminoacídicos (449-871)
	neurotoxina botulínica de tipo B	- restos aminoacídicos (441-858)
	neurotoxina botulínica de tipo C	- restos aminoacídicos (441-866)
	neurotoxina botulínica de tipo D	- restos aminoacídicos (446-862)
	neurotoxina botulínica de tipo E	- restos aminoacídicos (423-845)
40	neurotoxina botulínica de tipo F	- restos aminoacídicos (440-864)
	neurotoxina botulínica de tipo G	- restos aminoacídicos (442-863)
	neurotoxina tetánica	- restos aminoacídicos (458-879)

[0135] Las secuencias de referencia identificadas anteriormente deben considerarse como guía, ya que pueden producirse ligeras variaciones según los subserotipos. A modo de ejemplo, el documento US 2007/0166332 cita secuencias clostridiales ligeramente diferentes:

50	neurotoxina botulínica de tipo A	- restos aminoacídicos (A449-K871)
	neurotoxina botulínica de tipo B	- restos aminoacídicos (A442-S858)
	neurotoxina botulínica de tipo C	- restos aminoacídicos (T450-N866)
	neurotoxina botulínica de tipo D	- restos aminoacídicos (D446-N862)
	neurotoxina botulínica de tipo E	- restos aminoacídicos (K423-K845)
	neurotoxina botulínica de tipo F	- restos aminoacídicos (A440-K864)
	neurotoxina botulínica de tipo G	- restos aminoacídicos (S447-S863)
55	neurotoxina tetánica	- restos aminoacídicos (S458-V879)

[0136] En el contexto de la presente invención, diversas regiones H_N clostridiales que comprenden un dominio de translocación pueden ser útiles en aspectos de la presente invención, con la condición de que estos fragmentos activos puedan facilitar la liberación de una proteasa no citotóxica (por ejemplo, un cadena L clostridial) desde las vesículas intracelulares al citoplasma de la célula diana y así participar en ejecutar el mecanismo celular general por

el que una toxina clostridial escinde proteolíticamente un sustrato. Las regiones H_N de las cadenas pesadas de las toxinas clostridiales tienen una longitud de aproximadamente 410-430 aminoácidos y comprenden un dominio de translocación. La investigación ha demostrado que no es necesaria la longitud total de la región H_N de la cadena pesada de una toxina clostridial para la actividad de translocación del dominio de translocación. Por lo tanto, algunos aspectos de esta realización pueden incluir regiones H_N de toxinas clostridiales que comprenden un dominio de translocación con una longitud, por ejemplo, de al menos 350 aminoácidos, al menos 375 aminoácidos, al menos 400 aminoácidos y al menos 425 aminoácidos. Otros aspectos de esta realización pueden incluir regiones H_N de toxinas clostridiales que comprenden un dominio de translocación con una longitud, por ejemplo, de 350 aminoácidos como máximo, 375 aminoácidos como máximo, 400 aminoácidos como máximo y 425 aminoácidos como máximo.

[0137] Para más detalles sobre la base genética de la producción de toxinas en *Clostridium botulinum* y *C. tetani*, se menciona Henderson y col. (1997) en *The Clostridia: Molecular Biology and Pathogenesis*, Academic press.

[0138] El término H_N abarca porciones H_N de neurotoxinas de origen natural y porciones H_N modificadas con secuencias aminoacídicas que no se presentan en la naturaleza y/o restos aminoacídicos sintéticos, siempre que las porciones H_N modificadas aún demuestren la función de translocación mencionada anteriormente.

[0139] Alternativamente, el dominio de translocación puede no ser de origen clostridial. Algunos ejemplos de dominios de translocación (de referencia) de origen no clostridial incluyen, pero no se limitan al dominio de translocación de la toxina diftérica [O'Keefe y col., *Proc. Natl. Acad. Sci. USA* (1992) 89: 6202-6206; Silverman y col., *J. Biol. Chem.* (1993) 269: 22524-22532; y London E. (1992) *Biochem. Biophys. Acta* 1112: 25-51], el dominio de translocación de la exotoxina de *Pseudomonas* de tipo A [Prior y col., *Biochemistry* (1992) 31: 3555-3559], los dominios de translocación de la toxina del ántrax [Blanke y col., *Proc. Natl. Acad. Sci. USA* (1996) 93: 8437-8442], diversos péptidos fusógenos o hidrófobos con función de translocación [Plank y col., *J. Biol. Chem.* (1994) 269: 12918-12924; y Wagner y col. (1992) *PNAS* 89: 7934-7938] y péptidos anfifílicos [Murata y col. (1992) *Biochem.* 31: 1986-1992]. El dominio de translocación puede reflejar el dominio de translocación presente en una proteína de origen natural o puede incluir variaciones de aminoácidos, siempre que las variaciones no destruyan la capacidad de translocación del dominio de translocación.

[0140] Algunos ejemplos concretos de dominios de translocación víricos (de referencia) para uso en la presente invención incluyen ciertos dominios de translocación de proteínas de fusión membranaria expresadas por virus. Por ejemplo, Wagner y col. (1992) y Murata y col. (1992) describen la función de translocación (es decir, fusión membranaria y vesiculación) de una serie de péptidos fusógenos y anfifílicos derivados de la región N terminal de la hemaglutinina del virus de la influenza. Otras proteínas de fusión membranaria expresadas por virus de las que se sabe que tienen la actividad de translocación deseada incluyen un dominio de translocación de un péptido fusógeno del virus del bosque de Semliki (SFV), un dominio de translocación de la glucoproteína G del virus de la estomatitis vesicular (VSV), un dominio de translocación de la proteína F del virus SER y un dominio de translocación de la glucoproteína de la cubierta del virus espumoso. Las proteínas Aspike codificadas por virus tienen una aplicación especial en el contexto de la presente invención, por ejemplo, la proteína E1 del SFV y la proteína G del VSV.

[0141] El uso de los dominios de translocación (de referencia) listados en la tabla (a continuación) incluye es uso de variantes de las secuencias de los mismos. Una variante puede comprender una o más sustituciones conservadoras en el ácido nucleico y/o deleciones o inserciones en el ácido nucleico, con la condición de que la variante tenga la función de translocación necesaria. Una variante puede comprender también una o más sustituciones aminoacídicas y/o deleciones o inserciones aminoacídicas, siempre que la variante tenga la función de translocación requerida.

Fuente del dominio de translocación	Restos aminoacídicos	Referencias
Toxina diftérica	194-380	Silverman y col., 1994, <i>J. Biol. Chem.</i> 269: 22524-22532; London E., 1992, <i>Biochem. Biophys. Acta</i> 1113: 25-51
Dominio II de la exotoxina de <i>Pseudomonas</i>	405-613	Prior y col., 1992, <i>Biochemistry</i> 31: 3555-3559 Kihara y Pastan, 1994, <i>Bioconj. Chem.</i> 5: 532-538
Hemaglutinina del virus de la influenza	GLFGAIAGFIENGWE GMIDGWYG y variantes de esta secuencia	Plank y col., 1994, <i>J. Biol. Chem.</i> 269: 12918-12924 Wagner y col., 1992, <i>PNAS</i> 89: 7934-7938 Murata y col., 1992, <i>Biochemistry</i> 31: 1986-1992

Proteína fusógena del virus del bosque de Semliki	Dominio de translocación	Kielian y col., 1996, J. Cell Biol. 134(4): 863-872
Glucoproteína G del virus de la estomatitis vesicular	118-139	Yao y col., 2003, Virology 310(2): 319-332
Proteína F del virus SER	Dominio de translocación	Seth a col., 2003, J. Virol. 77(11): 6520-6527
Glucoproteína de la cubierta del virus espumoso	Dominio de translocación	Picard-Maureau y col., 2003, J. Virol. 77(8): 4722-4730

[0142] Los polipéptidos de la presente invención pueden comprender además un dominio facilitador de la translocación. Dichos dominios facilitan la liberación de la proteasa no citotóxica al citosol de la célula diana y se describen, por ejemplo, en los documentos WO08/008803 y WO08/008805.

5

[0143] A modo de ejemplo, los dominios facilitadores de la translocación adecuados incluyen el dominio de un péptido fusógeno de un virus con cubierta; por ejemplo, los dominios de péptidos fusógenos adecuados incluyen el dominio del péptido fusógeno del virus de la influenza (por ejemplo, el dominio de 23 aminoácidos del péptido fusógeno del virus de la influenza A), el dominio del péptido fusógeno de alfavirus (por ejemplo, el dominio de 26 aminoácidos del péptido fusógeno del virus del bosque de Semliki), el dominio del péptido fusógeno de vesiculovirus (por ejemplo, el dominio de 21 aminoácidos del péptido fusógeno del virus de la estomatitis vesicular), el dominio del péptido fusógeno de respirovirus (por ejemplo, el dominio de 25 aminoácidos del péptido fusógeno del virus de Sendai), el dominio del péptido fusógeno de morbilivirus (por ejemplo, el dominio de 25 aminoácidos del péptido fusógeno del virus del moquillo canino), el dominio del péptido fusógeno de avulavirus (por ejemplo, el dominio de 25 aminoácidos del péptido fusógeno del virus de la enfermedad de Newcastle), el dominio del péptido fusógeno de henipavirus (por ejemplo, el dominio de 25 aminoácidos del péptido fusógeno del virus de Hendra), el dominio del péptido fusógeno de metaneumovirus (por ejemplo, el dominio de 25 aminoácidos del péptido fusógeno del metaneumovirus humano) o el dominio del péptido fusógeno de espumavirus tales como el dominio del péptido fusógeno del virus espumoso de los simios; o fragmentos o variantes de los mismos.

20

[0144] A modo de otros ejemplos, un dominio facilitador de la translocación puede comprender un dominio H_{CN} de una toxina clostridial o un fragmento o variante del mismo. Más detalladamente, un dominio facilitador de la translocación H_{CN} de una toxina clostridial puede tener una longitud de al menos 200 aminoácidos, al menos 225 aminoácidos o al menos 275 aminoácidos. A este respecto, un dominio facilitador de la translocación H_{CN} de una toxina clostridial tiene preferentemente una longitud de 200 aminoácidos como máximo, de 225 aminoácidos como máximo, de 250 aminoácidos como máximo o de 275 aminoácidos como máximo.

25

[0145] Algunos ejemplos (de referencia) específicos incluyen:

30	neurotoxina botulínica de tipo A	- restos aminoacídicos (872-1110)
	neurotoxina botulínica de tipo B	- restos aminoacídicos (859-1097)
	neurotoxina botulínica de tipo C	- restos aminoacídicos (867-1111)
	neurotoxina botulínica de tipo D	- restos aminoacídicos (863-1098)
	neurotoxina botulínica de tipo E	- restos aminoacídicos (846-1085)
35	neurotoxina botulínica de tipo F	- restos aminoacídicos (865-1105)
	neurotoxina botulínica de tipo G	- restos aminoacídicos (864-1105)
	neurotoxina tetánica	- restos aminoacídicos (880-1127)

[0146] Las posiciones de las secuencias anteriores pueden variar ligeramente según el serotipo/subtipo y otros ejemplos de dominios H_{CN} de toxinas clostridiales (de referencia) adecuados) incluyen:

40

45	neurotoxina botulínica de tipo A	- restos aminoacídicos (874-1110)
	neurotoxina botulínica de tipo B	- restos aminoacídicos (861-1097)
	neurotoxina botulínica de tipo C	- restos aminoacídicos (869-1111)
	neurotoxina botulínica de tipo D	- restos aminoacídicos (865-1098)
	neurotoxina botulínica de tipo E	- restos aminoacídicos (848-1085)
	neurotoxina botulínica de tipo F	- restos aminoacídicos (867-1105)
	neurotoxina botulínica de tipo G	- restos aminoacídicos (866-1105)

neurotoxina tetánica

- restos aminoacídicos (882-1127)

[0147] Cualquiera de los dominios facilitadores de la translocación descritos anteriormente puede combinarse con cualquiera de los dominios peptídicos de translocación descritos anteriormente que son adecuados para uso en la presente invención. Por lo tanto, a modo de ejemplo, puede combinarse un dominio facilitador no clostridial con un dominio peptídico de translocación no clostridial o con un dominio peptídico de translocación clostridial. Alternativamente, dominio facilitador de la translocación H_{CN} de una toxina clostridial puede combinarse con un dominio peptídico de translocación no clostridial. Alternativamente, dominio facilitador H_{CN} de una toxina clostridial puede combinarse con un dominio peptídico de translocación clostridial, algunos ejemplos de lo cual incluyen:

10

neurotoxina botulínica de tipo A

- restos aminoacídicos (449-1110)

neurotoxina botulínica de tipo B

- restos aminoacídicos (442-1097)

neurotoxina botulínica de tipo C

- restos aminoacídicos (450-1111)

neurotoxina botulínica de tipo D

- restos aminoacídicos (446-1098)

15

neurotoxina botulínica de tipo E

- restos aminoacídicos (423-1085)

neurotoxina botulínica de tipo F

- restos aminoacídicos (440-1105)

neurotoxina botulínica de tipo G

- restos aminoacídicos (447-1105)

neurotoxina tetánica

- restos aminoacídicos (458-1127)

20 Homología de secuencia

[0148] Para determinar el porcentaje de identidad, puede usarse cualquiera de una diversidad de procedimientos de alineamiento de secuencias, incluidos, sin limitación, procedimientos globales, procedimientos locales y procedimientos híbridos, tales como, por ejemplo, procedimientos de aproximación por segmentos. Los protocolos para determinar el porcentaje de identidad son procedimientos habituales dentro del alcance de un experto en la técnica. Los procedimientos globales alinean las secuencias desde el principio hasta el final de la molécula y determinan el mejor alineamiento sumando las puntuaciones de pares de restos individuales e imponiendo penalizaciones por huecos. Algunos procedimientos no limitantes incluyen, por ejemplo CLUSTAL W, véase, por ejemplo, Julie D. Thompson y col., CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment Through Sequence Weighting, Position- Specific Gap Penalties and Weight Matrix Choice, Nucleic Acids Research 22(22): 4673-4680 (1994); y refinamiento iterativo, véase, por ejemplo, Osamu Gotoh, Significant Improvement in Accuracy of Multiple Protein. Sequence Alignments by Iterative Refinement as Assessed by Reference to Structural Alignments, J. Mol. Biol. 264(4): 823-838 (1996). Los procedimientos locales alinean las secuencias identificando uno o más motivos conservados en común en todas las secuencias de prueba. Algunos procedimientos no limitantes incluyen, por ejemplo, Match-box, véase, por ejemplo, Eric Depiereux y Ernest Feytmans, Match-Box: A Fundamentally New Algorithm for the Simultaneous Alignment of Several Protein Sequences, CABIOS 8(5): 501-509 (1992); muestreo de Gibbs, véase, por ejemplo, C. E. Lawrence y col., Detecting Subtle Sequence Signals: A Gibbs Sampling Strategy for Multiple Alignment, Science 262(5131): 208-214 (1993); Align-M, véase, por ejemplo, Ivo Van Walle y col., Align-M - A New Algorithm for Multiple Alignment of Highly Divergent Sequences, Bioinformatics 20(9):1428-1435 (2004).

[0149] Por lo tanto, el porcentaje de identidad de secuencia se determina por procedimientos convencionales. Véase, por ejemplo, Altschul y col., Bull. Math. Bio. 48: 603-16, 1986 y Henikoff y Henikoff, Proc. Natl. Acad. Sci. USA 89: 10915-19, 1992. Brevemente, dos secuencias aminoacídicas se alinean para optimizar las puntuaciones de alineamiento con una penalización de hueco abierto de 10, una penalización de extensión de hueco de 1 y la matriz de puntuación "blosum 62" de Henikoff y Henikoff (cita anterior) según se muestra a continuación (los aminoácidos se indican mediante los códigos de una letra estándar).

Puntuaciones de alineamiento para determinar la identidad de secuencia

50

[0150]

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
A	4																			
R	-1	5																		
N	-2	0	6																	
D	-2	-2	1	6																
C	0	-3	-3	-3	9															
Q	-1	1	0	0	-3	5														
E	-1	0	0	2	-4	2	5													
G	0	-2	0	-1	-3	-2	-2	6												
H	-2	0	1	-1	-3	0	0	-2	8											
I	-1	-3	-3	-3	-1	-3	-3	-4	-3	4										
L	-1	-2	-3	-4	-1	-2	-3	-4	-3	2	4									
K	-1	2	0	-1	-3	1	1	-2	-1	-3	-2	5								
M	-1	-1	-2	-3	-1	0	-2	-3	-2	1	2	-1	5							
F	-2	-3	-3	-3	-2	-3	-3	-3	-1	0	0	-3	0	6						
P	-1	-2	-2	-1	-3	-1	-1	-2	-2	-3	-3	-1	-2	-4	7					
S	1	-1	1	0	-1	0	0	0	-1	-2	-2	0	-1	-2	-1	4				
T	0	-1	0	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-2	-1	1	5			
W	-3	-3	-4	-4	-2	-2	-3	-2	-2	-3	-2	-3	-1	1	-4	-3	-2	11		
Y	-2	-2	-2	-3	-2	-1	-2	-3	2	-1	-1	-2	-1	3	-3	-2	-2	2	7	
V	0	-3	-3	-3	-1	-2	-2	-3	-3	3	1	-2	1	-1	-2	-2	0	-3	-1	4

[0151] El porcentaje de identidad se calcula entonces como:

$$5 \quad \frac{\text{Total de correspondencias idénticas}}{[\text{longitud de la secuencia más larga más el número de huecos introducidos en la secuencia más larga para alinear las dos secuencias}]} \times 100$$

[0152] Los polipéptidos sustancialmente homólogos se caracterizan como los que tienen una o más sustituciones, deleciones o adiciones de aminoácidos. Estos cambios son preferentemente de poca importancia, es decir, sustituciones de aminoácidos conservadoras (véase a continuación) y otras sustituciones que no afectan significativamente al plegamiento o la actividad del polipéptido; pequeñas deleciones, típicamente de uno a aproximadamente 30 aminoácidos; y pequeñas extensiones amino o carboxiterminales, tales como un resto de metionina aminoterminal, un pequeño péptido enlazante de hasta aproximadamente 20-25 restos o una etiqueta de afinidad.

Sustituciones de aminoácidos conservadoras

[0153]

5	Aminoácidos básicos:	arginina lisina histidina
	Aminoácidos ácidos:	ácido glutámico ácido aspártico
	Aminoácidos polares:	glutamina asparagina
10	Aminoácidos hidrófobos:	leucina isoleucina valina
	Aminoácidos aromáticos:	fenilalanina triptófano
	Aminoácidos pequeños:	tirosina glicina alanina serina treonina metionina
20		

[0154] Además de los 20 aminoácidos estándar, algunos aminoácidos no estándar (tales como 4-hidroxiprolina, 6-*N*-metilisina, ácido 2-aminoisobutírico, isovalina y α -metilserina) pueden sustituir a los restos aminoacídicos de los polipéptidos de la presente invención. Los restos aminoacídicos de los polipéptidos clostridiales pueden sustituirse por un número limitado de aminoácidos no conservadores, aminoácidos que no están codificados por el código genético y aminoácidos no naturales. Los polipéptidos de la presente invención pueden comprender también restos aminoacídicos de origen no natural.

[0155] Los aminoácidos de origen no natural incluyen, sin limitación, *trans*-3-metilprolina, 2,4-metanoprolina, *cis*-4-hidroxiprolina, *trans*-4-hidroxiprolina, *N*-metilglicina, alotreonina, metiltreonina, hidroxietilcisteína, hidroxietilhomocisteína, nitroglutamina, homoglutamina, ácido pipercolico, *terc*-leucina, norvalina, 2-azafenilalanina, 3-azafenilalanina, 4-azafenilalanina y 4-fluorofenilalanina. En la técnica se conocen varios procedimientos para la incorporación en proteínas de restos aminoacídicos de origen no natural. Por ejemplo, puede emplearse un sistema *in vitro*, en el que se suprimen mutaciones terminadoras mediante ARNt supresores aminoacilados químicamente. Los procedimientos para la síntesis de aminoácidos y la aminoacilación de ARNt son conocidos en la técnica. La transcripción y traducción de plásmidos que contienen mutaciones terminadoras se lleva a cabo en un sistema acelular que comprende un extracto de *E. coli* S30 y enzimas y otros reactivos disponibles comercialmente. Las proteínas se purifican por cromatografía. Véase, por ejemplo, Robertson y col., J. Am. Chem. Soc. 113: 2722, 1991; Ellman y col., Methods Enzymol. 202: 301, 1991; Chung y col., Science 259: 806-9, 1993; y Chung y col., Proc. Natl. Acad. Sci. USA 90: 10145-9, 1993. En un segundo procedimiento, la traducción se lleva a cabo en ovocitos de *Xenopus* por microinyección de ARNm mutado y ARNt supresores aminoacilados químicamente (Turcatti y col., J. Biol. Chem. 271: 19991-8, 1996). En un tercer procedimiento, se cultivan células de *E. coli* en ausencia del aminoácido natural que ha de sustituirse (por ejemplo, fenilalanina) y en presencia del(os) aminoácido(s) de origen no natural deseado(s) (por ejemplo, 2-azafenilalanina, 3-azafenilalanina, 4-azafenilalanina o 4-fluorofenilalanina). El aminoácido de origen no natural se incorpora en el polipéptido en lugar del equivalente natural. Véase, Koide y col., Biochem. 33: 7470-6, 1994. Los restos aminoacídicos de origen natural pueden convertirse a la especie de origen no natural por modificación química *in vitro*. La modificación química puede combinarse con mutagénesis dirigida para ampliar aún más la gama de sustituciones (Wynn y Richards, Protein Sci. 2: 395-403, 1993).

[0156] Los restos aminoacídicos de los polipéptidos de la presente invención pueden sustituirse por un número limitado de aminoácidos no conservadores, aminoácidos que no están codificados por el código genético, aminoácidos de origen no natural y aminoácidos no naturales.

[0157] Los aminoácidos esenciales en los polipéptidos de la presente invención pueden identificarse de acuerdo con procedimientos conocidos en la técnica, tales como la mutagénesis dirigida o la mutación por barrido con alanina (Cunningham y Wells, Science 244: 1081-5, 1989). Los sitios de interacción biológica también pueden determinarse por análisis físico de la estructura, según se determina por técnicas tales como resonancia magnética nuclear, cristalografía, difracción de electrones o marcado por fotoafinidad, en combinación con la mutación de aminoácidos del sitio de contacto potencial. Véase, por ejemplo, Vos y col, Science 255: 306-12, 1992; Smith y col.,

J. Mol. Biol. 224: 899-904, 1992; Wlodaver y col., FEBS Lett. 309: 59-64, 1992. Las identidades de los aminoácidos esenciales pueden inferirse también del análisis de homologías con componentes relacionados (por ejemplo, componentes de translocación o proteásico) de los polipéptidos de la presente invención.

5 **[0158]** Es posible llevar a cabo y analizar múltiples sustituciones de aminoácidos mediante procedimientos conocidos de mutagénesis y cribado, tales como los descritos por Reidhaar-Olson y Sauer (Science 241: 53-7, 1988) o Bowie y Sauer (Proc. Natl. Acad. Sci. USA 86: 2152-6, 1989). Brevemente, estos autores desvelan procedimientos para aleatorizar simultáneamente dos o más posiciones en un polipéptido, con la selección de un polipéptido funcional y la secuenciación posterior de los polipéptidos mutagenizados para determinar el espectro de
10 sustituciones permisibles en cada posición. Otros procedimientos que pueden usarse incluyen la exposición en fagos (por ejemplo, Lowman y col., Biochem. 30: 10832-7, 1991; Ladner y col., patente de los EE. UU. n.º 5.223.409; Huse, publicación WIPO WO92/06204) y mutagénesis dirigida de regiones (Derbyshire y col., Gene 46: 145, 1986; Ner y col., DNA 7: 127, 1988).

15 **[0159]** Es posible llevar a cabo y analizar múltiples sustituciones de aminoácidos mediante procedimientos conocidos de mutagénesis y cribado, tales como los descritos por Reidhaar-Olson y Sauer (Science 241: 53-7, 1988) o Bowie y Sauer (Proc. Natl. Acad. Sci. USA 86: 2152-6, 1989). Brevemente, estos autores desvelan procedimientos para aleatorizar simultáneamente dos o más posiciones en un polipéptido, con la selección de un polipéptido funcional y la secuenciación posterior de los polipéptidos mutagenizados para determinar el espectro de
20 sustituciones permisibles en cada posición. Otros procedimientos que pueden usarse incluyen la exposición en fagos (por ejemplo, Lowman y col., Biochem. 30: 10832-7, 1991; Ladner y col., patente de los EE. UU. n.º 5.223.409; Huse, publicación WIPO WO92/06204) y mutagénesis dirigida de regiones (Derbyshire y col., Gene 46: 145, 1986; Ner y col., DNA 7: 127, 1988).

25 **[0160]** A continuación se presenta una breve descripción de las figuras, que ilustran aspectos y/o realizaciones de la presente invención.

La figura 1 ilustra la sorprendente actividad de neurotoxina retenida por una neurotoxina clostridial modificada (LH_N). Dicha neurotoxina modificada carece de un dominio de unión H_C funcional y, por lo tanto, es equivalente a las
30 neurotoxinas clostridiales modificadas descritas por Lin y col., (documento WO02/044199). En contraste, no se detectó actividad de neurotoxina en una neurotoxina clostridial modificada (LC/C) que carece de un dominio de translocación H_N funcional.

La figura 2 ilustra un simple alineamiento de homología de secuencia aminoacídica para los diversos serotipos de BoNT. En este alineamiento pueden compararse los restos o regiones aminoacídicos de un serotipo (por ejemplo, del serotipo A) con los correspondientes restos/regiones en los demás serotipos por medio de un simple
35 alineamiento vertical.

La figura 3 muestra el análisis de SDS-PAGE de la purificación de una proteína quimérica L(#FXa)HC-EGF. El carril
40 1 ilustra el lisado celular clarificado; el carril 2 ilustra el flujo a través de la columna; el carril 3 muestra la fracción eluida después del lavado de la columna; los carriles 4, 6-12 son las fracciones eluidas tras la adición de imidazol 250 mM. El carril 5 son los marcadores de masa molecular (Benchmark).

La figura 4 muestra el análisis de SDS-PAGE de la proteólisis de una proteína quimérica L(#FXa)HC-EGF por FXa.
45 Los carriles 1 y 2 ilustran la proteína purificada antes de la activación con enterocinasa en la unión LC-H_N, en ausencia y en presencia de DTT, respectivamente; los carriles 3 y 4 ilustran la proteína después de la activación con enterocinasa en la unión LC-H_N, en ausencia y en presencia de DTT, respectivamente; el carril 5 son los marcadores de masa molecular (Benchmark); los carriles 6 y 7 ilustran la proteína activada con enterocinasa tratada con el factor Xa, en ausencia y en presencia de DTT, respectivamente; los carriles 8 y 9 ilustran el resultado del tratamiento con
50 FXa de la proteína no activada con enterocinasa, en ausencia y en presencia de DTT, respectivamente. La fragmentación de la proteína después del tratamiento con el factor Xa se observa claramente en los carriles 6 y 8.

La figura 5 muestra el análisis de SDS-PAGE de la purificación de una proteína quimérica LH(#FXa)C-EGF (según se prepara en el ejemplo 20). El carril 1 ilustra el lisado celular clarificado; el carril 2 son los marcadores de masa
55 molecular (Benchmark); el carril 3 ilustra el flujo a través de la columna; el carril 4 ilustra la fracción eluida después del lavado de la columna; los carriles 5-12 son las fracciones eluidas tras la adición de imidazol 250 mM.

La figura 6 muestra el análisis de SDS-PAGE de la purificación de una proteína quimérica LH(#FXa)C-EGF (según se prepara en el ejemplo 21). El carril 1 ilustra el lisado celular clarificado; el carril 2 ilustra el flujo a través de la

columna; el carril 3 son los marcadores de masa molecular (Benchmark); el carril 4 ilustra la fracción eluida después del lavado de la columna; los carriles 5-12 son las fracciones eluidas tras la adición de imidazol 250 mM.

La figura 7 muestra el análisis de SDS-PAGE de la proteólisis de una proteína quimérica LH(#FXa)C-EGF (según se prepara en el ejemplo 20) por FXa. Los carriles 1 y 3 ilustran la proteína purificada antes de la activación con enterocinasa en la unión LC-H_N, en ausencia y en presencia de DTT, respectivamente; los carriles 4 y 5 ilustran la proteína después de la activación con enterocinasa en la unión LC-H_N, en ausencia y en presencia de DTT, respectivamente; los carriles 6 y 7 ilustran la proteína activada con enterocinasa tratada con el factor Xa, en ausencia y en presencia de DTT, respectivamente; los carriles 8 y 9 ilustran el resultado del tratamiento con FXa de la proteína no activada con enterocinasa, en ausencia y en presencia de DTT, respectivamente. La fragmentación de la proteína después del tratamiento con el factor Xa se observa claramente en los carriles 7 y 9. El carril 2 son los marcadores de masa molecular (Benchmark).

La figura 8 muestra un análisis de inmunotransferencia de la proteólisis de una proteína quimérica LH(#FXa)C-EGF (según se prepara en el ejemplo 20) por FXa. Los carriles 1 y 3 ilustran la proteína purificada antes de la activación con enterocinasa en la unión LC-H_N, en ausencia y en presencia de DTT, respectivamente; los carriles 4 y 5 ilustran la proteína después de la activación con enterocinasa en la unión LC-H_N, en ausencia y en presencia de DTT, respectivamente; los carriles 6 y 7 ilustran la proteína activada con enterocinasa tratada con el factor Xa, en ausencia y en presencia de DTT, respectivamente; los carriles 8 y 9 ilustran el resultado del tratamiento con FXa de la proteína no activada con enterocinasa, en ausencia y en presencia de DTT, respectivamente. La fragmentación de la proteína después del tratamiento con el factor Xa se observa claramente en los carriles 6, 7, 8 y 9 por visualización de una banda inmunorreactiva con histidina correspondiente al tamaño anticipado. El carril 2 son los marcadores de masa molecular adecuados para detección por tinción (Benchmark). El carril 10 son los marcadores de masa molecular adecuados para visualización por inmunotransferencia (Magic Markers).

La figura 9 muestra el análisis de SDS-PAGE de la proteólisis de una proteína quimérica LH(#FXa)C-EGF (según se prepara en el ejemplo 21) por FXa. Los carriles 1 y 2 ilustran la proteína purificada antes de la activación con enterocinasa en la unión LC-H_N, en ausencia y en presencia de DTT, respectivamente; los carriles 4 y 5 ilustran la proteína después de la activación con enterocinasa en la unión LC-H_N, en ausencia y en presencia de DTT, respectivamente; los carriles 6 y 7 ilustran la proteína activada con enterocinasa tratada con el factor Xa, en ausencia y en presencia de DTT, respectivamente; los carriles 8 y 9 ilustran el resultado del tratamiento con FXa de la proteína no activada con enterocinasa, en ausencia y en presencia de DTT, respectivamente. La fragmentación de la proteína después del tratamiento con el factor Xa se observa claramente en los carriles 7 y 9. El carril 3 son los marcadores de masa molecular (Benchmark).

La figura 10 muestra el análisis de SDS-PAGE de la purificación de una proteína quimérica L(#Thr)HC-EGF (según se prepara en el ejemplo 19). El carril 1 son los marcadores de masa molecular (Benchmark); el carril 2 ilustra el lisado celular clarificado; el carril 3 ilustra el flujo a través de la columna; el carril 4 ilustra la fracción eluida después del lavado de la columna; los carriles 5-12 son las fracciones eluidas tras la adición de imidazol 250 mM.

La figura 11 muestra el análisis de SDS-PAGE de la proteólisis de una proteína quimérica L(#Thr)HC-EGF (según se prepara en el ejemplo 19) por trombina. El carril 1 son los marcadores de masa molecular (Benchmark). Los carriles 2 y 3 ilustran la proteína purificada antes de la activación con enterocinasa en la unión LC-H_N, en ausencia y en presencia de DTT, respectivamente; los carriles 4 y 5 ilustran la proteína después de la activación con enterocinasa en la unión LC-H_N, en ausencia y en presencia de DTT, respectivamente; los carriles 6 y 7 ilustran la proteína activada con enterocinasa tratada con el factor Xa, en ausencia y en presencia de DTT, respectivamente; los carriles 8 y 9 ilustran el resultado del tratamiento con FXa de la proteína no activada con enterocinasa, en ausencia y en presencia de DTT, respectivamente. La fragmentación de la proteína después del tratamiento con trombina se observa claramente en los carriles 6, 7, 8 y 9.

La figura 12 muestra un análisis de inmunotransferencia de la proteólisis de una proteína quimérica L(#Thr)HC-EGF (según se prepara en el ejemplo 19) por trombina. El carril 1 son los marcadores de masa molecular (Benchmark), que son poco visibles en la inmunotransferencia. Los carriles 2 y 3 ilustran la proteína purificada antes de la activación con enterocinasa en la unión LC-H_N, en ausencia y en presencia de DTT, respectivamente; los carriles 4 y 5 ilustran la proteína después de la activación con enterocinasa en la unión LC-H_N, en ausencia y en presencia de DTT, respectivamente; los carriles 6 y 7 ilustran la proteína activada con enterocinasa tratada con el factor Xa, en ausencia y en presencia de DTT, respectivamente; los carriles 8 y 9 ilustran el resultado del tratamiento con FXa de la proteína no activada con enterocinasa, en ausencia y en presencia de DTT, respectivamente. La fragmentación de la proteína después del tratamiento con trombina, que produce un fragmento de aproximadamente 85 kDa que

retiene el dominio EGF pero carece de aproximadamente 20 kDa del extremo N de LC, se observa claramente en los carriles 8 y 9.

La figura 13 muestra el análisis de SDS-PAGE de la purificación de una proteína quimérica L(#Thr)HA-EGF (según se prepara en el ejemplo 24). El carril 1 ilustra el lisado celular clarificado; el carril 2 ilustra el flujo a través de la columna; el carril 3 ilustra la fracción eluida después del lavado de la columna; el carril 4 son los marcadores de masa molecular (Benchmark); los carriles 5-11 son las fracciones eluidas tras la adición de imidazol 250 mM.

La figura 14 muestra el análisis de SDS-PAGE de la proteólisis de una proteína quimérica L(#Thr)HA-EGF (según se prepara en el ejemplo 24) por trombina. El carril 4 son los marcadores de masa molecular (Benchmark). Los carriles 1 y 2 ilustran la proteína purificada antes de la activación con enterocinasa en la unión LC-H_N, en ausencia y en presencia de DTT, respectivamente; los carriles 3 y 5 ilustran la proteína después de la activación con enterocinasa en la unión LC-H_N, en ausencia y en presencia de DTT, respectivamente; los carriles 6 y 7 ilustran la proteína activada con enterocinasa tratada con trombina, en ausencia y en presencia de DTT, respectivamente; los carriles 8 y 9 ilustran el resultado del tratamiento con trombina de la proteína no activada con enterocinasa, en ausencia y en presencia de DTT, respectivamente. La fragmentación de la proteína después del tratamiento con trombina se observa claramente en los carriles 6, 7, 8 y 9.

La figura 15 muestra un análisis de inmunotransferencia de la proteólisis de una proteína quimérica L(#Thr)HA-EGF (según se prepara en el ejemplo 24) por trombina. El carril 4 son los marcadores de masa molecular (Benchmark). Los carriles 1 y 2 ilustran la proteína purificada antes de la activación con enterocinasa en la unión LC-H_N, en ausencia y en presencia de DTT, respectivamente; los carriles 3 y 5 ilustran la proteína después de la activación con enterocinasa en la unión LC-H_N, en ausencia y en presencia de DTT, respectivamente; los carriles 6 y 7 ilustran la proteína activada con enterocinasa tratada con trombina, en ausencia y en presencia de DTT, respectivamente; los carriles 8 y 9 ilustran el resultado del tratamiento con trombina de la proteína no activada con enterocinasa, en ausencia y en presencia de DTT, respectivamente. La fragmentación de la proteína después del tratamiento con trombina se observa claramente en los carriles 7 y 9.

La figura 16 ilustra el resultado obtenido después de la exposición de SCN a L(FXa)HC-EGF tratada con FXa en comparación con L(FXa)HC-EGF sin tratar. La proteína tratada con el factor Xa es claramente menos eficaz en la escisión de la sintaxina que la proteína sin tratar con FXa. Por lo tanto, la invención hace posible una reducción de la eficacia de la proteína modificada.

La figura 17 ilustra el resultado obtenido después de la exposición de SCN a L(Thr)HA-EGF tratada con trombina 10 nM en comparación con L(Thr)HA-EGF sin tratar 10 nM. La proteína tratada con trombina es claramente menos eficaz en la escisión de SNAP-25 que la proteína sin tratar con trombina. Por lo tanto, la invención hace posible una reducción de la eficacia de la proteína modificada.

[0161] A continuación se presenta la descripción de realizaciones específicas de la invención, ilustradas por los ejemplos.

[0162] Ejemplo 1 - Evaluación de los polipéptidos de la invención cuando se exponen a células (musculares) de mamífero.

[0163] Ejemplo 2 - Evaluación de los polipéptidos de la invención cuando se exponen a células de mamífero, en que los polipéptidos han sido expuestos primeramente a proteasas circulantes.

[0164] Ejemplo 3 - Evaluación de la actividad catalítica de los polipéptidos de la invención.

[0165] Ejemplo 4 - Evaluación de la capacidad de translocación de los polipéptidos de la invención.

[0166] Ejemplo 5 - Creación de una proteína quimérica LHC-EGF que incorpora un sitio de reconocimiento para el factor Xa en LC.

[0167] Ejemplo 6 - Purificación de una proteína quimérica LHC-EGF que incorpora un sitio de reconocimiento para el factor Xa en LC.

[0168] Ejemplo 7 - Demostración del aumento de la sensibilidad a proteasas en una proteína quimérica LHC-EGF que incorpora un sitio de reconocimiento para el factor Xa en LC. **Ejemplo 8** - Creación de una proteína

quimérica LHC-EGF que incorpora un sitio de reconocimiento para trombina en LC.

[0169] Ejemplo 9 - Creación de una proteína quimérica LHA-EGF que incorpora un sitio de reconocimiento para trombina en LC.

[0170] Ejemplo 10 - Creación de una proteína quimérica LHC-EGF que incorpora un sitio de reconocimiento para furina en LC.

[0171] Ejemplo 11 - Creación de una proteína quimérica LHA-EGF que incorpora un sitio de reconocimiento para el factor Xa en el dominio H_N.

[0172] Ejemplo 12 - Creación de una proteína quimérica LHA-EGF que incorpora un sitio de reconocimiento para ADAM17 en el dominio LC.

[0173] Ejemplo 13 - Creación de una proteína BoNT/A recombinante que incorpora un sitio de reconocimiento para ADAM17 en LC.

[0174] Ejemplo 14 - Creación de una proteína BoNT/A recombinante que incorpora un sitio de reconocimiento para furina en H_N.

[0175] Ejemplo 15 - Tratamiento de un paciente que padece distonía (tortícolis espasmódica).

[0176] Ejemplo 16 - Tratamiento de un paciente que padece blefaroespasma.

[0177] Ejemplo 17 - Creación de una proteína quimérica LHC-EGF que incorpora un sitio de reconocimiento para el factor Xa en la posición 210 de LC.

[0178] Ejemplo 18 - Creación de una proteína quimérica LHC-EGF que incorpora un sitio de reconocimiento para trombina en la posición 195 de LC.

[0179] Ejemplo 19 - Creación de una proteína quimérica LHC-EGF que incorpora un sitio de reconocimiento para trombina en la posición 210 de LC.

[0180] Ejemplo 20 - Creación de una proteína quimérica LHC-EGF que incorpora un sitio de reconocimiento para el factor Xa en la posición 742 del dominio H_N.

[0181] Ejemplo 21 - Creación de una proteína quimérica LHC-EGF que incorpora un sitio de reconocimiento para el factor Xa en la posición 750 del dominio H_N.

[0182] Ejemplo 22 - Creación de una proteína quimérica LHC-EGF que incorpora un sitio de reconocimiento para trombina en la posición 750 del dominio H_N.

[0183] Ejemplo 23 - Creación de una proteína quimérica LHD-VIPr que incorpora un sitio de reconocimiento para el factor Xa en la posición 798 del dominio H_N.

[0184] Ejemplo 24 - Creación de una proteína quimérica LHA-EGF que incorpora un sitio de reconocimiento para trombina en el dominio LC.

[0185] Ejemplo 25 - Demostración de la escisión específica de una proteína quimérica LHC-EGF manipulada para incorporar un sitio de reconocimiento para el factor Xa en LC.

[0186] Ejemplo 26 - Demostración de la escisión específica de una proteína quimérica LHC-EGF manipulada para incorporar un sitio de reconocimiento para el factor Xa en H_N.

[0187] Ejemplo 27 - Demostración de la escisión específica de una proteína quimérica LHC-EGF manipulada para incorporar un sitio de reconocimiento para trombina en LC.

[0188] Ejemplo 28 - Demostración de la escisión específica de una proteína quimérica LHA-EGF manipulada para incorporar un sitio de reconocimiento para trombina en LC.

[0189] Ejemplo 29 - Demostración de la reducida actividad celular *in vitro* de una proteína manipulada para incorporar un sitio de escisión proteásica para el factor Xa en el dominio LC de L(FXa)HC-EGF.

5 **[0190] Ejemplo 30** - Demostración de la reducida actividad celular *in vitro* de una proteína manipulada para incorporar un sitio de escisión proteásica para trombina en el dominio LC de L(Thr)HA-EGF.

[0191] Ejemplo 31 - Creación de una proteína BoNT/A recombinante que incorpora un sitio de reconocimiento para trombina en LC.

10

[0192] Ejemplo 32 - Creación de una proteína BoNT/A recombinante que incorpora un sitio de reconocimiento para el factor Xa en LC.

15 **[0193] Ejemplo 33** - Creación de una proteína BoNT/A recombinante que incorpora un sitio de reconocimiento para el factor Xa en H_N.

[0194] Ejemplo 34 - Creación de una proteína BoNT/E recombinante que incorpora un sitio de reconocimiento para trombina en LC.

20 **[0195] Ejemplo 35** - Creación de una proteína BoNT/E recombinante que incorpora un sitio de reconocimiento para el factor Xa en H_N.

[0196] Ejemplo 36 - Creación de una proteína quimérica LHE-VIPr que incorpora un sitio de reconocimiento para trombina en LC.

25

[0197] Ejemplo 37 - Creación de una proteína quimérica LHE-VIPr que incorpora un sitio de reconocimiento para el factor Xa en H_N.

30 **[0198] Ejemplo 38** - Creación de una proteína quimérica LHE-VIPr que incorpora un sitio de reconocimiento para el factor Xa en LC.

[0199] Ejemplo 39 - Escisión de una proteína SNARE por una neurotoxina clostridial modificada (LH_N) que tiene las propiedades descritas por Lin y col. (documento WO02/044199).

35 **Resumen de secuencias SEQ ID NO**

[0200]

- | | | |
|----|-----------|---------------------------------------|
| 40 | SEQ ID 1 | secuencia de ADN de LHC-EGF |
| | SEQ ID 2 | secuencia proteínica de LHC-EGF |
| | SEQ ID 3 | secuencia de ADN de L(#FXa)HC-EGF |
| 45 | SEQ ID 4 | secuencia proteínica de L(#FXa)HC-EGF |
| | SEQ ID 5 | secuencia de ADN de L(#Thr)HC-EGF |
| | SEQ ID 6 | secuencia proteínica de L(#Thr)HC-EGF |
| 50 | SEQ ID 7 | secuencia de ADN de LHA-EGF |
| | SEQ ID 8 | secuencia proteínica de LHA-EGF |
| 55 | SEQ ID 9 | secuencia de ADN de L(#Thr)HA-EGF |
| | SEQ ID 10 | secuencia proteínica de L(#Thr)HA-EGF |
| | SEQ ID 11 | secuencia de ADN de L(#furina)HC-EGF |

	SEQ ID 12	secuencia proteínica de L(#furina)HC-EGF
	SEQ ID 13	secuencia de ADN de LH(#FXa)A-EGF
5	SEQ ID 14	secuencia proteínica de LH(#FXa)A-EGF
	SEQ ID 15	secuencia de ADN de L(#ADAM17)HA-EGF
10	SEQ ID 16	secuencia proteínica de L(#ADAM17)HA-EGF
	SEQ ID 17	secuencia de ADN de LHA-Hc/A
	SEQ ID 18	secuencia proteínica de LHA-Hc/A
15	SEQ ID 19	secuencia de ADN de L(#ADAM17)HA- Hc/A
	SEQ ID 20	secuencia proteínica de L(#ADAM17)HA- Hc/A
20	SEQ ID 21	secuencia de ADN de L(#furina)HA-Hc/A
	SEQ ID 22	secuencia proteínica de L(#furina)HA-Hc/A
	SEQ ID 23	secuencia de ADN de L(#FXa)HC-EGF (SXN1975)
25	SEQ ID 24	secuencia proteínica de L(#FXa)HC-EGF (SXN1975)
	SEQ ID 25	secuencia proteínica de L(#Thr)HC-EGF (SXN1931)
30	SEQ ID 26	secuencia proteínica de L(#Thr)HC-EGF (SXN1932)
	SEQ ID 27	secuencia proteínica de LH(#FXa)C-EGF (SXN1937)
	SEQ ID 28	secuencia proteínica de LH(#FXa)C-EGF (SXN1938)
35	SEQ ID 29	secuencia proteínica de LH(#Thr)C-EGF (SXN1939)
	SEQ ID 30	secuencia proteínica de LH(#FXa)D-VIPr (SXN1930)
40	SEQ ID 31	secuencia proteínica de L(#Thr)HA-EGF (SXN1974)
	SEQ ID 32	secuencia proteínica de L(#Thr)HA-Hc/A
	SEQ ID 33	secuencia proteínica de L(#FXa)HA-Hc/A
45	SEQ ID 34	secuencia proteínica de LH(FXa)A-Hc/A
	SEQ ID 35	secuencia proteínica de L(#Thr)HE-Hc/E
50	SEQ ID 36	secuencia proteínica de LH(#FXa)E-Hc/E
	SEQ ID 37	secuencia proteínica de L(#Thr)HE-VIPr
	SEQ ID 38	secuencia proteínica de LH(#FXa)E-VIPr
55	SEQ ID 39	secuencia proteínica de L(#FXa)HE-VIPr (mutación K228D)

Ejemplos

Ejemplo 1 - Evaluación de los polipéptidos de la invención cuando se exponen a células musculares de mamífero

- 5 **[0201]** Una proteína purificada creada de acuerdo con el ejemplo 13 se incuba en presencia de células musculares de mamífero (cultivo primario de músculo liso coronario o células HSkMC (150-05f) (disponibles en la colección ECACC)). Paralelamente, un segundo polipéptido (idéntico al primer polipéptido excepto porque no tiene el mismo sitio de escisión destructivo) se incuba en condiciones idénticas en presencia del mismo tipo de células de prueba.
- 10 **[0202]** A continuación, se evalúa la escisión por ADAM17 (intrínseca al cultivo primario de músculo liso coronario o células HSkMC) de cada uno de los dos polipéptidos mediante SDS-PAGE y posterior análisis de inmunotransferencia. A este respecto, la mayor escisión observada para el primer polipéptido en comparación con la observada para el segundo polipéptido confirma la inactivación controlable de la presente invención.

15 **Ejemplo 2 - Evaluación de los polipéptidos de la invención cuando se exponen a células de mamífero, en que los polipéptidos han sido expuesto primeramente a una proteasa circulante**

- 20 **[0203]** Se toma un primer polipéptido (SEQ ID 4; preparado de acuerdo con el ejemplo 5 de la presente invención) y se incuba en presencia de las células diana, en que el polipéptido ha sido expuesto primeramente a proteasas circulantes (por ejemplo, el factor Xa, trombina) *in vitro*. Paralelamente, un segundo polipéptido (SEQ ID 2; idéntico al primer polipéptido, excepto porque carece del sitio de escisión proteásica) se incuba de la misma manera que el primer polipéptido.
- 25 **[0204]** A continuación, cada uno de los polipéptidos se evalúa en cuanto a la escisión de syntaxina en neuronas de médula espinal embrionaria no humana (eSCN). A este respecto, la menor escisión observada para el primer polipéptido en comparación con la observada para el segundo polipéptido confirma la inactivación controlable de la presente invención.

Ejemplo 3 - Evaluación de la actividad catalítica de los polipéptidos de la invención

- 30 **[0205]** Un primer polipéptido (SEQ ID 10; preparado de acuerdo con el ejemplo 9 de la presente invención) se incuba *in vitro* en presencia de una proteasa (trombina) que escinde el polipéptido en un sitio de escisión destructivo introducido en el dominio proteásico del polipéptido. Paralelamente, un segundo polipéptido (SEQ ID 8; idéntico al primer polipéptido, excepto porque carece del sitio de escisión proteásica) se incuba de la misma manera en presencia de la misma proteasa.
- 35 **[0206]** A continuación, cada uno de los dos polipéptidos se expone a un sistema acelular *in vitro* (según se describe por Hallis y col., 1996, J. Clin. Microbiol. 34: 1934-1938) que contiene SNAP-25 inmovilizada y se mide la escisión de la proteína SNAP-25 con antiseros específicos producidos contra el producto de escisión. A este respecto, una menor escisión de la proteína SNARE por el primer polipéptido en comparación con la observada para el segundo polipéptido confirma la inactivación controlable de la presente invención.
- 40

Ejemplo 4 - Evaluación de la capacidad de translocación de los polipéptidos de la invención

- 45 **[0207]** Un primer polipéptido (de acuerdo con la presente invención) se incuba en presencia de una proteasa que escinde el polipéptido en un primer sitio de escisión destructivo introducido en el dominio de translocación (por ejemplo, H_N). Paralelamente, un segundo polipéptido (idéntico al primer polipéptido, excepto por que carece del sitio de escisión proteásica) se incuba de manera idéntica en presencia de la misma proteasa.
- 50 **[0208]** A continuación, cada uno de los dos polipéptidos se expone en un sistema *in vitro* a una membrana bicapa lipídica y se mide el transporte a través de la membrana. Por ejemplo, Shone C. (1987) describe un ensayo *in vitro* con el empleo de liposomas que se exponen a la molécula de prueba. La presencia de la función de translocación requerida se confirma por la liberación de K⁺ y/o NAD marcado de los liposomas, lo cual puede monitorizarse fácilmente [véase Shone C. (1987) Eur. J. Biochem.; vol. 167(1): págs. 175-180]. Otro ejemplo más lo proporciona Blaustein R. (1987), que describe un simple ensayo *in vitro* con el empleo de membranas planas de bicapas de fosfolípidos. Las membranas se exponen a la molécula de prueba y la función de translocación requerida se confirma por un aumento de la conductancia a través de dichas membranas [véase Blaustein (1987) FEBS Letts.; vol. 226, n.º 1: págs. 115-120].
- 55

[0209] Este procedimiento se aplica al estudio de la inactivación proteásica del dominio H_N de BoNT del serotipo D. La proteína del ejemplo 23 se expresa y purifica y se expone al factor Xa, lo que resulta en la escisión de la proteína dentro del dominio H_N. A continuación, la proteína escindida se evalúa en el sistema *in vitro* descrito anteriormente y se compara con la proteína que no ha sido tratada con el factor Xa. El experimento determina que el transporte a través de la membrana del polipéptido tratado con el factor Xa es significativamente menor que el del polipéptido sin tratar.

Ejemplo 5 - Creación de una proteína quimérica LHC-EGF que incorpora un sitio de reconocimiento para el factor Xa en LC

[0210] La secuencia primaria de una proteína quimérica construida por la fusión genética del fragmento LH_N de BoNT/C y la secuencia del factor de crecimiento epidérmico humano (SEQ ID 2) se revisa para determinar la presencia de porciones de la secuencia aminoacídica con similitud con el sitio de reconocimiento prototípico para el factor Xa (IEGR). Un simple análisis de caracteres de texto de la secuencia primaria identificó la secuencia ²¹⁰GEGR₂₁₃ dentro del dominio LC. La localización del péptido en la estructura terciaria de LC/C se predice a partir del examen de la localización de la secuencia peptídica homóloga en LC/A, con el uso de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA) como guía. Se usa un software disponible públicamente (tal como FirstGlance in Jmol (<http://molvis.sdsc.edu/fqii/fq.htm?mol=3bta>)) para identificar que la secuencia peptídica equivalente de LC/A se localiza en la superficie de LC. Por lo tanto, esta localización es una región de buena accesibilidad para proteasas.

[0211] Se llevó a cabo una mutagénesis dirigida de SEQ ID 1 (que codifica SEQ ID 2) con un cebador diseñado para cambiar el codón de G₂₁₀ (GGC) a uno que codifica Ile (ATC), mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene). El uso de codones de *E. coli* se determinó por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se determinaron por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resultara en una mala utilización de codones. El ADN mutagenizado se incorporó en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprobó por secuenciación. La ORF final que incorpora el sitio para el factor Xa se ilustra como SEQ ID 3 y la secuencia aminoacídica del producto de expresión se ilustra en SEQ ID 4.

Ejemplo 6 - Purificación de una proteína quimérica LHC-EGF que incorpora un sitio de reconocimiento para el factor Xa en LC

[0212] La ORF creada en el ejemplo 17 se clonó en un vector de expresión en *E. coli* (un vector pET (Novagen) modificado para asegurar una deficiencia de movilización) y se transformó en una cepa hospedadora de *E. coli*, principalmente BL21. El vector se modificó para incluir la expresión de una etiqueta de histidina en el extremo N de la ORF de LHC-EGF.

[0213] La expresión de la proteína de fusión LHC-EGF se lleva a cabo mediante el protocolo siguiente. Se inoculan 100 ml de TB modificado que contiene glucosa al 0,2 % y 100 µg/ml de ampicilina en una botella de 250 ml con una única colonia de la cepa de expresión de LHC-EGF. El cultivo se incuba a 37 °C, a 225 rpm durante 16 horas. Se inocula 1 l de TB modificado que contiene glucosa al 0,2 % y 100 µg/ml de ampicilina en una botella de 2 l con 10 ml del cultivo incubado durante la noche. Los cultivos se incuban a 37 °C hasta alcanzar una DO₆₀₀ nm de aproximadamente 0,5, en cuyo momento se reduce la temperatura a 16 °C. Después de una hora se inducen los cultivos con IPTG 1 mM y se incuban a 16 °C durante otras 16 horas.

[0214] La purificación de la fusión LHC-EGF se lleva a cabo por cromatografía de afinidad. En detalle, se descongela un tubo Falcon que contiene 25 ml de HEPES 50 mM, pH 7,2, NaCl 200 mM y aproximadamente 10 g de pasta de células *E. coli* B21. La pasta de células se trata con ultrasonidos en hielo durante diez ciclos de 30 segundos, con 30 segundos de pausa a una potencia de 22 µm, asegurando que la muestra se mantiene fría. Las células lisadas se centrifugan a 18.000 rpm y 4 °C durante 30 minutos. El sobrenadante se carga en una columna quelante cargada con NiSO₄ 0,1 M (20-30 ml por columna son suficientes) y equilibrada con HEPES 50 mM, pH 7,2, NaCl 200 mM. Mediante un gradiente de etapas de imidazol 10 y 40 mM, se elimina por lavado la proteína no unida específicamente y la proteína de fusión se eluye con imidazol 100 mM. La proteína de fusión eluida se dializa frente a 5 l de HEPES 50 mM, pH 7,2, NaCl 200 mM a 4 °C durante la noche y se mide la DO de la proteína de fusión dializada. Se añaden 6,4 ng de enterocinasa por mg de la proteína de fusión y se incuba a 25 °C en reposo durante la noche. Se carga una columna quelante cargada con NiSO₄ 0,1 M (20-30 ml por columna son suficientes) y equilibrada con HEPES 50 mM, pH 7,2, NaCl 200 mM. La columna se lava hasta el valor de referencia con HEPES

50 mM, pH 7,2, NaCl 200 mM. Mediante un gradiente de etapas de imidazol 10 y 40 mM, se elimina por lavado la proteína no unida específicamente y la proteína de fusión se eluye con imidazol 100 mM. La proteína de fusión eluida se dializa frente a 5 l de HEPES 50 mM, pH 7,2, NaCl 200 mM a 4 °C durante la noche y se concentra a aproximadamente 2 mg/ml y la muestra se distribuye en alícuotas y se congela a -20 °C. La proteína purificada se examina mediante DO y análisis BCA y de pureza.

Ejemplo 7 - Demostración del aumento de la sensibilidad a proteasas en una proteína quimérica LHC-EGF que incorpora un sitio de reconocimiento para el factor Xa en LC

10 [0215] La proteína quimérica purificada del ejemplo 6 se evalúa en cuanto a su estabilidad en presencia de proteasas mediante la metodología descrita en los ejemplos 2 y 3. En resumen, la proteína quimérica LHC-EGF se expone a un intervalo de concentraciones de la proteasa Factor Xa (obtenida, por ejemplo de New England Biolabs n.º P8010L) *in vitro* durante un periodo de 1-120 minutos. La proteólisis se termina por la adición de un inhibidor específico del factor Xa (por ejemplo, dansil-glu-gly-arg-clorometilcetona (CALBIOCHEM, n.º 251700)). Se usa una
15 proteína quimérica LHC-EGF de control que no incluye el sitio adicional para el factor Xa para comparar el efecto de la proteasa en la actividad de LC (mediante el ejemplo 3) y la funcionalidad de la quimera al exponerla a las células diana (mediante el ejemplo 2 y midiendo la escisión de syntaxina en neuronas de la médula espinal embrionaria no humana (eSCN)).

20 **Ejemplo 8 - Creación de una proteína quimérica LHC-EGF que incorpora un sitio de reconocimiento para trombina en LC**

[0216] La secuencia primaria de una proteína quimérica construida por la fusión genética del fragmento LH_N de BoNT/C y la secuencia del factor de crecimiento epidérmico humano (SEQ ID 2) se revisa para determinar la
25 presencia de porciones de la secuencia de aminoácidos con similitud con el sitio de reconocimiento prototípico para trombina (LVPRGS). Un simple análisis de caracteres de texto de la secuencia primaria identificó la secuencia ₁₉₄ISPRFM₁₉₉ dentro del dominio LC. La localización del péptido en la estructura terciaria de LC/C se predice a partir del examen de la localización de la secuencia peptídica homóloga en LC/A, con el uso de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA) como guía. Se usa un software disponible públicamente (tal como FirstGlance in Jmol (<http://molvis.sdsc.edu/fqii/fq.htm?mol=3bta>)) para identificar que la secuencia peptídica equivalente de LC/A se
30 localiza cerca de la superficie de LC. Por lo tanto, esta localización es una región de buena accesibilidad para proteasas.

[0217] Una mutagénesis dirigida de SEQ ID 1 (que codifica SEQ ID 2) con un cebador diseñado para cambiar
35 los codones de S₁₉₅ a Val (TCT a GTT) y de M₁₉₉ a Ser (AGT a TCC) cambia la región ₁₉₄ISPRFM₁₉₉ a IVPRFS para convertirla en un sustrato para la escisión por trombina. La mutagénesis se llevó a cabo mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene). El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas
40 publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La ORF final que incorpora el sitio para trombina se ilustra como SEQ ID 5 y la secuencia aminoácida del producto de expresión se ilustra en SEQ ID 6.

45 **Ejemplo 9 - Creación de una proteína quimérica LHA-EGF que incorpora un sitio de reconocimiento para trombina en LC**

[0218] La secuencia primaria de una proteína quimérica construida por la fusión genética del fragmento LH_N de BoNT/A y la secuencia del factor de crecimiento epidérmico humano (SEQ ID 8) se revisa para determinar la
50 presencia de porciones de la secuencia aminoácida con similitud con el sitio de reconocimiento prototípico para trombina (GRG). Un simple análisis de caracteres de texto de la secuencia primaria identificó la secuencia ₁₀₃GRM₁₀₅ dentro del dominio LC. La localización del péptido en la estructura terciaria de LC/A se predice a partir de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA) como guía. Se usa un software disponible públicamente (tal como FirstGlance in Jmol (<http://molvis.sdsc.edu/fqii/fq.htm?mol=3bta>)) para identificar que la secuencia peptídica de
55 LC/A se localiza en la superficie. Por lo tanto, esta localización es una región de buena accesibilidad para proteasas.

[0219] Se llevó a cabo una mutagénesis dirigida de SEQ ID 7 (que codifica la ORF de SEQ ID 8) con un cebador diseñado para cambiar el codón de Met₁₀₅ (ATG) a uno que codifica Gly (GGT), mediante las técnicas

moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene). El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La ORF final que incorpora el sitio para trombina se ilustra como SEQ ID 9 y la secuencia aminoacídica del producto de expresión se ilustra en SEQ ID 10.

10 Ejemplo 10 - Creación de una proteína quimérica LHC-EGF que incorpora un sitio de reconocimiento para furina en LC

[0220] La secuencia primaria de una proteína quimérica construida por la fusión genética del fragmento LH_N de BoNT/C y la secuencia del factor de crecimiento epidérmico humano (SEQ ID 2) se revisa para determinar la presencia de porciones de la secuencia aminoacídica con similitud con el sitio de reconocimiento prototípico para furina (RXR▼K/R). Un simple análisis de caracteres de texto de la secuencia primaria identificó la secuencia ²¹⁰GEGR₂₁₃ dentro del dominio LC. La localización del péptido en la estructura terciaria de LC/C se predice a partir del examen de la localización de la secuencia peptídica homóloga en LC/A, con el uso de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA) como guía. Se usa un software disponible públicamente (tal como FirstGlance in Jmol (<http://molvis.sdsc.edu/fjii/fq.htm?mol=3bta>)) para identificar que la secuencia peptídica equivalente de LC/A se localiza en la superficie de LC. Por lo tanto, esta localización es una región de buena accesibilidad para proteasas.

[0221] Se llevó a cabo una mutagénesis dirigida de SEQ ID 1 (que codifica SEQ ID 2) con un cebador diseñado para cambiar la región peptídica GEGR a RSRR mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene). El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La ORF final que incorpora el sitio para furina se ilustra como SEQ ID 11 y la secuencia aminoacídica del producto de expresión se ilustra en SEQ ID 12.

35 Ejemplo 11 - Creación de una proteína quimérica LHA-EGF que incorpora un sitio de reconocimiento para el factor Xa en el dominio H_N

[0222] La secuencia primaria de una proteína quimérica construida por la fusión genética del fragmento LH_N de BoNT/A y la secuencia del factor de crecimiento epidérmico humano (SEQ ID 8) se revisa para determinar la presencia de porciones de la secuencia aminoacídica con similitud con el sitio de reconocimiento prototípico para el factor Xa (IEGR). Un simple análisis de caracteres de texto de la secuencia primaria identificó la secuencia ⁵⁶²GKSR₅₆₅ dentro del dominio H_N. La localización del péptido en la estructura terciaria de H_N/A se predice a partir de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA) como guía. Se usa un software disponible públicamente (tal como FirstGlance in Jmol (<http://molvis.sdsc.edu/fjii/fq.htm?mol=3bta>)) para identificar que la secuencia peptídica de H_N se localiza en la superficie. Por lo tanto, esta localización es una región de buena accesibilidad para proteasas.

[0223] Se llevó a cabo una mutagénesis dirigida de SEQ ID 7 (que codifica SEQ ID 8) con un cebador diseñado para cambiar la región peptídica GKSR a IEGR mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene). El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La ORF final que incorpora el sitio para el factor Xa se ilustra como SEQ ID 13 y la secuencia aminoacídica del producto de expresión se ilustra en SEQ ID 14.

Ejemplo 12 - Creación de una proteína quimérica LHA-EGF que incorpora un sitio de reconocimiento para ADAM17 en el dominio LC

[0224] La secuencia primaria de una proteína quimérica construida por la fusión genética del fragmento LH_N de BoNT/A y la secuencia del factor de crecimiento epidérmico humano (SEQ ID 8) se revisa para determinar la presencia de porciones de la secuencia aminoacídica con similitud con el sitio de reconocimiento prototípico para ADAM17 (PLAQAVRSSS). Un simple análisis de caracteres de texto de la secuencia primaria identifica una región de la estructura (206PLL₂₁₅GAGKFAT dentro del dominio LC) que puede someterse a ingeniería de proteínas. La localización del péptido en la estructura terciaria de LC se predice a partir de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA) como guía. Se usa un software disponible públicamente (tal como FirstGlance in Jmol (<http://molvis.sdsc.edu/fqii/fq.htm?mol=3bta>)) para identificar que la secuencia peptídica de LC se localiza en la superficie. Por lo tanto, esta localización es de buena accesibilidad para proteasas.

[0225] Se llevó a cabo una mutagénesis dirigida de SEQ ID 7 (que codifica SEQ ID 8) mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene). La mutagénesis de LC se realizó para modificar la región codificante de 206PLL₂₁₅GAGKFAT a PLAQAVRSSS.

[0226] El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La ORF final que incorpora el sitio para ADAM17 se ilustra como SEQ ID 15 y la secuencia aminoacídica del producto de expresión se ilustra en SEQ ID 16.

Ejemplo 13 - Creación de una proteína BoNT/A recombinante que incorpora un sitio de reconocimiento para ADAM17 en LC

[0227] La secuencia primaria de una BoNT/A recombinante con actividad endopeptidásica que contiene un sitio de activación proteásica específico para enterocinasa incorporado (SEQ ID 18) se revisa para determinar la presencia de porciones de la secuencia aminoacídica con similitud con el sitio de reconocimiento prototípico para ADAM17 (PLAQAVRSSS). Un simple análisis de caracteres de texto de la secuencia primaria identifica una región de la estructura de BoNT (206PLL₂₁₅GAGKFAT dentro del dominio LC) que puede someterse a ingeniería de proteínas. La localización del péptido en la estructura terciaria de LC se predice a partir de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA) como guía. Se usa un software disponible públicamente (tal como FirstGlance in Jmol (<http://molvis.sdsc.edu/fqii/fq.htm?mol=3bta>)) para identificar que la secuencia peptídica de LC se localiza en la superficie. Por lo tanto, esta localización es de buena accesibilidad para proteasas.

[0228] Se llevó a cabo una mutagénesis dirigida de SEQ ID 17 (que codifica SEQ ID 18) mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene). La mutagénesis de LC se realizó para modificar la región codificante de 206PLL₂₁₅GAGKFAT a PLAQAVRSSS.

[0229] El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La ORF final que incorpora el sitio para ADAM17 se ilustra como SEQ ID 19 y la secuencia aminoacídica del producto de expresión se ilustra en SEQ ID 20.

Ejemplo 14 - Creación de una proteína BoNT/A recombinante que incorpora un sitio de reconocimiento para furina en H_N

[0230] La secuencia primaria de una BoNT/A recombinante con actividad endopeptidásica que contiene un sitio de activación proteásica específico para enterocinasa incorporado (SEQ ID 18) se revisa para determinar la presencia de porciones de la secuencia aminoacídica con similitud con el sitio de reconocimiento prototípico para furina (RXR▼K/R). Un simple análisis de caracteres de texto de la secuencia primaria identifica la secuencia 563KSR₅₆₅ dentro del dominio H_N que puede someterse a ingeniería de proteínas. La localización del péptido en la estructura terciaria del dominio H_N se predice a partir de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA) como guía. Se usa un software disponible públicamente (tal como FirstGlance in Jmol (<http://molvis.sdsc.edu/fqii/fq.htm?mol=3bta>)) para identificar que la secuencia peptídica de H_N se localiza en la

superficie. Por lo tanto, esta localización es de buena accesibilidad para proteasas.

[0231] Una mutagénesis dirigida de SEQ ID 17 (que codifica SEQ ID 18) con un cebador diseñado para cambiar el codón de K₅₆₃ (AAA) a Arg (CGT) e insertar un resto de Arg (CGC) después del R₅₆₅ existente cambia la secuencia ₅₆₃KSR₅₆₅ a RSRR, que es un sustrato para la escisión por furina. La mutagénesis se llevó a cabo mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene).

[0232] El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La ORF final que incorpora el sitio para ADAM17 se ilustra como SEQ ID 21 y la secuencia aminoacídica del producto de expresión se ilustra en SEQ ID 22.

Ejemplo 15 - Tratamiento de un paciente que padece distonía (tortícolis espasmódica)

[0233] Un paciente varón que padecía tortícolis espasmódica, manifestada por contracciones espasmódicas o tónicas de la musculatura del cuello que producían desviaciones anormales estereotipadas de la cabeza, con la barbilla girada hacia un lado y con elevación del hombro hacia el lado al que se giraba la cabeza, había sido tratado previamente con una cantidad terapéuticamente eficaz de BoNT/A en los músculos del cuello para la tortícolis, pero presentó disfagia por dispersión de la proteasa en la bucofaringe. El paciente se trata a continuación por inyección con hasta aproximadamente 300 unidades o más de un polipéptido de la presente invención (por ejemplo, una neurotoxina botulínica de tipo A modificada para incluir un sitio sensible a la proteasa factor Xa) en los músculos distónicos del cuello. Después de tres a siete días, los síntomas se alivian sustancialmente y el paciente puede mantener la cabeza y el hombro en una posición normal durante al menos tres meses. Después del tratamiento con la neurotoxina modificada, el paciente no experimenta ninguna disfagia. Al utilizar la toxina botulínica de tipo A modificada, el médico puede inyectar más producto en el área que requiere el tratamiento sin temor a un aumento de los efectos secundarios. El aumento de la dosis conduce a un aumento de la duración de la acción y, por lo tanto, a una mejora del tratamiento.

Ejemplo 16 - Tratamiento de un paciente que padece blefaroespasmó

[0234] Una paciente de 58 años con blefaroespasmó se trata por inyección de entre aproximadamente 1 a aproximadamente 5 unidades de un polipéptido de la presente invención (por ejemplo, una neurotoxina botulínica de tipo A modificada para incluir un sitio sensible a la proteasa ADAM17, según se describe en el ejemplo 13) en el músculo ocular orbicular pretarsal lateral del párpado superior y el ocular orbicular pretarsal lateral del párpado inferior, en que la cantidad inyectada varía en función del tamaño del músculo que ha de inyectarse y de la extensión de la parálisis muscular deseada. En aproximadamente uno a siete días se produce un alivio del blefaroespasmó. Al utilizar la toxina botulínica de tipo A modificada, el médico puede inyectar más producto en el área que requiere el tratamiento sin temor a un aumento de los efectos secundarios. El aumento de la dosis conduce a un aumento de la duración de la acción y, por lo tanto, a una mejora del tratamiento.

Ejemplo 17 - Creación de una proteína química LHC-EGF que incorpora un sitio de reconocimiento para el factor Xa en la posición 210 de LC [SXN101975]

[0235] La secuencia primaria de una proteína química construida por la fusión genética del fragmento LHN de BoNT/C y la secuencia del factor de crecimiento epidérmico humano (SEQ ID 2) se revisa para determinar la presencia de porciones de la secuencia aminoacídica con similitud con el sitio de reconocimiento prototípico para el factor Xa (IEGR). Se identifica un sitio para la inserción de un sitio para el factor Xa en la secuencia primaria ₂₁₀GEGR dentro del dominio LC. La localización del péptido en la estructura terciaria de LC/C se predice a partir del examen de la localización de la secuencia peptídica homóloga en LC/A, con el uso de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA) como guía. Se usa un software disponible públicamente (tal como FirstGlance in Jmol (<http://molvis.sdsc.edu/fqii/fq.htm?mol=3bta>)) para identificar que la secuencia peptídica equivalente de LC/A se localiza cerca de la superficie de LC. Por lo tanto, esta localización es una región de buena accesibilidad para proteasas.

[0236] Se llevó a cabo una mutagénesis dirigida de SEQ ID 1 (que codifica SEQ ID 2) con un cebador

diseñado para cambiar el codón de G₂₁₀ a I para obtener un sustrato para la escisión por el factor Xa. La mutagénesis se llevó a cabo mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene). El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La ORF final que incorpora el sitio para el factor Xa se ilustra como SEQ ID 23 y la secuencia aminoacídica del producto de expresión se ilustra en SEQ ID 24.

Ejemplo 18 - Creación de una proteína quimérica LHC-EGF que incorpora un sitio de reconocimiento para trombina en la posición 195 de LC [SXN101931]

[0237] La secuencia primaria de una proteína quimérica construida por la fusión genética del fragmento LHN de BoNT/C y la secuencia del factor de crecimiento epidérmico humano (SEQ ID 2) se revisa para determinar la presencia de porciones de la secuencia aminoacídica con similitud con el sitio de reconocimiento prototípico para trombina. Se identifica un sitio para la inserción de un sitio para trombina en la secuencia primaria ¹⁹⁴ISPRFM₁₉₉ dentro del dominio LC. La localización del péptido en la estructura terciaria de LC/C se predice a partir del examen de la localización de la secuencia peptídica homóloga en LC/A, con el uso de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA) como guía. Se usa un software disponible públicamente (tal como FirstGlance in Jmol (<http://molvis.sdsc.edu/fqii/fq.htm?mol=3bta>)) para identificar que la secuencia peptídica equivalente de LC/A se localiza cerca de la superficie de LC. Por lo tanto, esta localización es una región de buena accesibilidad para proteasas.

[0238] Una mutagénesis dirigida de SEQ ID 1 (que codifica SEQ ID 2) con un cebador diseñado para cambiar los codones de S₁₉₅ a Val (TCT a GTT) y de M₁₉₉ a Ser (AGT a TCC) cambia la región ¹⁹⁴ISPRFM₁₉₉ a IVPRFS para convertirla en sustrato para la escisión por trombina. La mutagénesis se llevó a cabo mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene). El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La ORF final que incorpora el sitio para trombina se usa para codificar la secuencia aminoacídica del producto de expresión ilustrado en SEQ ID 25.

Ejemplo 19 - Creación de una proteína quimérica LHC-EGF que incorpora un sitio de reconocimiento para trombina en la posición 210 de LC [SXN101932]

[0239] La secuencia primaria de una proteína quimérica construida por la fusión genética del fragmento LHN de BoNT/C y la secuencia del factor de crecimiento epidérmico humano (SEQ ID 2) se revisa para determinar la presencia de porciones de la secuencia aminoacídica con similitud con el sitio de reconocimiento prototípico para trombina. Se identifica un sitio para la inserción de un sitio para trombina en la secuencia primaria ²¹⁰GEGRFS dentro del dominio LC. La localización del péptido en la estructura terciaria de LC/C se predice a partir del examen de la localización de la secuencia peptídica homóloga en LC/A, con el uso de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA) como guía. Se usa un software disponible públicamente (tal como FirstGlance in Jmol (<http://molvis.sdsc.edu/fqii/fq.htm?mol=3bta>)) para identificar que la secuencia peptídica equivalente de LC/A se localiza cerca de la superficie de LC. Por lo tanto, esta localización es una región de buena accesibilidad para proteasas.

[0240] Se llevó a cabo una mutagénesis dirigida de SEQ ID 1 (que codifica SEQ ID 2) con un cebador diseñado para cambiar los codones de ²¹¹EGR a TPR para crear la secuencia GTPRFS que es un sustrato para la escisión por trombina. La mutagénesis se llevó a cabo mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene). El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar,

por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La ORF final que incorpora el sitio para trombina se usa para codificar la secuencia aminoacídica del producto de expresión ilustrado en SEQ ID 26.

5 Ejemplo 20 - Creación de una proteína quimérica LHC-EGF que incorpora un sitio de reconocimiento para el factor Xa en la posición 742 del dominio H_N [SXN101937]

[0241] Se revisa la secuencia primaria de una proteína quimérica construida por la fusión genética del fragmento LH_N de BoNT/C y la secuencia del factor de crecimiento epidérmico humano (SEQ ID 2) y se identifica un sitio para la inserción de un sitio para el factor Xa en la secuencia primaria ⁷⁴²IDLE₇₅₅ dentro del dominio H_N. La localización del péptido en la estructura terciaria de H_N/C se predice a partir del examen de la localización de la secuencia peptídica homóloga en H_N/A, con el uso de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA) como guía. Se usa un software disponible públicamente (tal como FirstGlance in Jmol (<http://molvis.sdsc.edu/fqii/fq.htm?mol=3bta>)) para identificar que la secuencia peptídica equivalente de H_N/A se localiza cerca de la superficie de H_N. Por lo tanto, esta localización es una región de buena accesibilidad para proteasas.

[0242] Se llevó a cabo una mutagénesis dirigida de SEQ ID 1 (que codifica SEQ ID 2) con un cebador diseñado para cambiar los codones de ⁷⁴²LE a GR para obtener un sustrato para la escisión por el factor Xa. La mutagénesis se llevó a cabo mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene). El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La ORF final que incorpora el sitio para trombina se usa para codificar la secuencia aminoacídica del producto de expresión ilustrado en SEQ ID 27.

30 Ejemplo 21 - Creación de una proteína quimérica LHC-EGF que incorpora un sitio de reconocimiento para el factor Xa en la posición 750 del dominio H_N [SXN101938]

[0243] Se revisa la secuencia primaria de una proteína quimérica construida por la fusión genética del fragmento LH_N de BoNT/C y la secuencia del factor de crecimiento epidérmico humano (SEQ ID 2) y se identifica un sitio para la inserción de un sitio para el factor Xa en la secuencia primaria ⁷⁵⁰SGSD₇₅₃ dentro del dominio H_N. La localización del péptido en la estructura terciaria de H_N/C se predice a partir del examen de la localización de la secuencia peptídica homóloga en H_N/A, con el uso de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA) como guía. Se usa un software disponible públicamente (tal como FirstGlance in Jmol (<http://molvis.sdsc.edu/fqii/fq.htm?mol=3bta>)) para identificar que la secuencia peptídica equivalente de H_N/A se localiza cerca de la superficie de H_N. Por lo tanto, esta localización es una región de buena accesibilidad para proteasas.

[0244] Una mutagénesis dirigida de SEQ ID 1 (que codifica SEQ ID 2) con un cebador diseñado para cambiar los codones de ⁷⁵⁰SGSD a IDGR da lugar a un sustrato para la escisión por el factor Xa. La mutagénesis se llevó a cabo mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene). El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La ORF final que incorpora el sitio para el factor Xa se usa para codificar la secuencia aminoacídica del producto de expresión ilustrado en SEQ ID 28.

55 Ejemplo 22 - Creación de una proteína quimérica LHC-EGF que incorpora un sitio de reconocimiento para trombina en la posición 750 del dominio H_N [SXN101939]

[0245] Se revisa la secuencia primaria de una proteína quimérica construida por la fusión genética del fragmento LH_N de BoNT/C y la secuencia del factor de crecimiento epidérmico humano (SEQ ID 2) y se identifica un sitio para la inserción de un sitio para trombina en la secuencia primaria ⁷⁵⁰SGSD₇₅₃ dentro del dominio H_N. La

localización del péptido en la estructura terciaria de H_N/C se predice a partir del examen de la localización de la secuencia peptídica homóloga en H_N/A, con el uso de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA) como guía. Se usa un software disponible públicamente (tal como FirstGlance in Jmol (<http://molvis.sdsc.edu/fqii/fq.htm?mol=3bta>)) para identificar que la secuencia peptídica equivalente de H_N/A se localiza cerca de la superficie de H_N. Por lo tanto, esta localización es una región de buena accesibilidad para proteasas.

[0246] Se llevó a cabo una mutagénesis dirigida de SEQ ID 1 (que codifica SEQ ID 2) con un cebador diseñado para cambiar los codones de SGSD a GVPR para obtener un sustrato para la escisión por trombina. La mutagénesis se llevó a cabo mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene). El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La ORF final que incorpora el sitio para trombina se usa para codificar la secuencia aminoacídica del producto de expresión ilustrado en SEQ ID 29.

20 **Ejemplo 23 - Creación de una proteína quimérica LHD-VIPr que incorpora un sitio de reconocimiento para el factor Xa en la posición 798 del dominio H_N [SXN101930]**

[0247] Se revisa la secuencia primaria de una proteína quimérica construida por la fusión genética del fragmento LH_N de BoNT/D y un análogo del péptido intestinal vasoactivo humano (VIPr) y se identifica un sitio para la inserción de un sitio para el factor Xa en la secuencia primaria ⁷⁹⁸SGSD dentro del dominio H_N. La localización del péptido en la estructura terciaria de H_N/D se predice a partir del examen de la localización de la secuencia peptídica homóloga en H_N/A, con el uso de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA) como guía. Se usa un software disponible públicamente (tal como FirstGlance in Jmol (<http://molvis.sdsc.edu/fqii/fq.htm?mol=3bta>)) para identificar que la secuencia peptídica equivalente de H_N/A se localiza cerca de la superficie de H_N. Por lo tanto, esta localización es una región de buena accesibilidad para proteasas.

[0248] Se llevó a cabo una mutagénesis dirigida del gen con un cebador diseñado para cambiar los codones de ⁷⁹⁸SGSD a IDGR para obtener un sustrato para la escisión por el factor Xa. La mutagénesis se llevó a cabo mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene). El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La ORF final que incorpora el sitio para el factor Xa se usa para codificar la secuencia aminoacídica del producto de expresión ilustrado en SEQ ID 30.

45 **Ejemplo 24 - Creación de una proteína quimérica LHA-EGF que incorpora un sitio de reconocimiento para trombina en el dominio LC [SXN1974]**

[0249] La secuencia primaria de una proteína quimérica construida por la fusión genética del fragmento LH_N de BoNT/A y la secuencia del factor de crecimiento epidérmico humano (SEQ ID 8) se revisa para determinar la presencia de porciones de la secuencia aminoacídica con similitud con el sitio de reconocimiento prototípico para trombina (GRG). Un simple análisis de caracteres de texto de la secuencia primaria identifica la secuencia ¹⁰³GRM₁₀₅ dentro del dominio LC. La localización del péptido en la estructura terciaria de LC se predice a partir de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA) como guía. Se usa un software disponible públicamente (tal como FirstGlance in Jmol (<http://molvis.sdsc.edu/fqii/fq.htm?mol=3bta>)) para identificar que la secuencia peptídica de LC se localiza en la superficie. Por lo tanto, esta localización es una región de buena accesibilidad para proteasas.

[0250] Se llevó a cabo una mutagénesis dirigida de SEQ ID 7 (que codifica SEQ ID 8) con un cebador diseñado para cambiar la región peptídica GRM a GRG, mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene). El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el

contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La ORF final que incorpora el sitio para trombina se usa para codificar la secuencia aminoacídica del producto de expresión ilustrado en SEQ ID 31.

Ejemplo 25 - Demostración de la escisión específica de una proteína quimérica LHC-EGF manipulada para incorporar un sitio de reconocimiento para el factor Xa en LC [SXN1975]

10

[0251] Se construye una molécula nueva que incorpora un sitio de reconocimiento para el factor Xa en el dominio LC de LHC-EGF de acuerdo con el ejemplo 17. La proteína del ejemplo 17 se expresa y purifica mediante una metodología similar a la descrita en el ejemplo 6: La metodología se adaptó para su uso en un sistema de purificación AKTA Xpress. Esencialmente, los lisados clarificados de *E. coli* se cargaron en una columna HisTrap FF Crude de 5 ml en el sistema Xpress. El programa se ajustó para lavar las columnas con 10 volúmenes de columna del tampón de unión (Tris pH 8,0 50 mM, NaCl 200 mM) y 10 volúmenes de columna de imidazol 40 mM en el tampón de unión (recogidos conjuntamente con el flujo a través de la columna). La elución se llevó a cabo con 5 volúmenes de columna de imidazol 250 mM en el tampón de unión. La proteína se recogió en un bucle y se conservó hasta que el sistema estuvo listo para desalar (en Tris pH 8,0 50 mM, NaCl 150 mM). La proteína desalada se recogió en una placa de 96 pocillos de 2 ml. La figura 3 ilustra la purificación de LHC-EGF a partir de *E. coli*.

[0252] La proteína se trata con la proteasa factor Xa mediante la metodología descrita en el ejemplo 7 y las muestras se analizan por SDS-PAGE. La figura 4 ilustra la escisión de la proteína en presencia del factor Xa. Se observan productos de escisión en las muestras no reducidas y reducidas. La masa estimada de los productos de escisión está de acuerdo con el punto de escisión anticipado en la proteína manipulada.

Ejemplo 26 - Demostración de la escisión específica de una proteína quimérica LHC-EGF manipulada para incorporar un sitio de reconocimiento para el factor Xa en H_N [SXN1937 y SXN1938]

[0253] Se construye una molécula nueva que incorpora un sitio de reconocimiento para el factor Xa en el dominio H_N de LHC-EGF de acuerdo con el ejemplo 20 y se construye una segunda molécula nueva que incorpora un sitio de reconocimiento para el factor Xa en una localización diferente dentro del dominio H_N de LHC-EGF de acuerdo con el ejemplo 21. Las proteínas de los ejemplos 20 y 21 se expresan y purifican mediante una metodología similar a la descrita en el ejemplo 24. La figura 5 ilustra la purificación de la LHC-EGF del ejemplo 20 a partir de *E. coli* y la figura 6 ilustra la purificación de la LHC-EGF del ejemplo 21 a partir de *E. coli*.

[0254] La proteína del ejemplo 20 se trata con la proteasa factor Xa mediante la metodología descrita en el ejemplo 7 y las muestras se analizan por SDS-PAGE. La figura 7 ilustra la escisión de la proteína en presencia del factor Xa, evaluada por tinción de los geles de SDS-PAGE. La figura 8 ilustra el perfil de las muestras evaluadas por inmunotransferencia con anticuerpos dirigidos contra la etiqueta His para detectar la presencia de dicha etiqueta His. La masa estimada de los productos de escisión está de acuerdo con del punto de escisión anticipado en la proteína manipulada.

[0255] La proteína del ejemplo 21 se trata con la proteasa factor Xa mediante la metodología descrita en el ejemplo 7 y las muestras se analizan por SDS-PAGE. La figura 9 ilustra la escisión de la proteína en presencia del factor Xa. La masa estimada de los productos de escisión está de acuerdo con del punto de escisión anticipado en la proteína manipulada.

Ejemplo 27 - Demostración de la escisión específica de una proteína quimérica LHC-EGF manipulada para incorporar un sitio de reconocimiento para trombina en LC [SXN1932]

50

[0256] Se construye una molécula nueva que incorpora un sitio de reconocimiento para trombina en el dominio LC de LHC-EGF de acuerdo con el ejemplo 19. La proteína del ejemplo 19 se expresa y purifica mediante una metodología similar a la descrita en el ejemplo 25. La figura 10 ilustra la purificación de LHC-EGF a partir de *E. coli*.

[0257] La proteína se trata con la proteasa trombina mediante la metodología descrita en el ejemplo 7 y las muestras se analizan por SDS-PAGE. La figura 11 ilustra la escisión de la proteína en presencia de trombina, evaluada por SDS-PAGE. La figura 12 ilustra la escisión de la proteína en presencia de trombina, evaluada por

inmunotransferencia con anticuerpos dirigidos contra EGF. La masa estimada de los productos de escisión está de acuerdo con del punto de escisión anticipado en la proteína manipulada.

Ejemplo 28 - Demostración de la escisión específica de una proteína quimérica LHA-EGF manipulada para incorporar un sitio de reconocimiento para trombina en LC [SXN1974]

[0258] Se construye una molécula nueva que incorpora un sitio de reconocimiento para el factor Xa en el dominio LC de LHA-EGF de acuerdo con el ejemplo 24. La proteína del ejemplo 24 se expresa y purifica mediante una metodología similar a la descrita en el ejemplo 25. La figura 13 ilustra la purificación de LHA-EGF a partir de *E. coli*.

[0259] La proteína se trata con la proteasa trombina mediante la metodología descrita en el ejemplo 7 y las muestras se analizan por SDS-PAGE. La figura 14 ilustra la escisión de la proteína en presencia de trombina. La figura 15 ilustra el perfil de inmunotransferencia del mismo PAGE, con un anticuerpo dirigido contra EGF como anticuerpo primario. La masa estimada de los productos de escisión está de acuerdo con del punto de escisión anticipado en la proteína manipulada.

Ejemplo 29 - Demostración de la reducida actividad celular *in vitro* de una proteína manipulada para incorporar un sitio de escisión proteásica para el factor Xa en el dominio LC de LHC-EGF [SXN1975]

[0260] El producto proteínico del ejemplo 25 se expresa y purifica. La proteína purificada se expone a la proteasa FXa antes de su evaluación en un ensayo de neuronas de médula espinal (SCN) *in vitro*. La preparación de las SCN es una técnica bien establecida y se describe en la bibliografía [B. R. Ransom, E. Neale, M. Henkart, P. N. Bullock, P. G. Nelson, Mouse spinal cord in cell culture. I. Morphology and intrinsic neuronal electrophysiologic properties, J. Neurophysiol. 40 (1977): 1132-1150; S. C. Fitzgerald, A Dissection and Tissue Culture Manual of the Nervous System, Alan R. Liss Inc, Nueva York, 1989]. La proteína de prueba se prepara en diversas concentraciones por dilución en el medio de cultivo. Las SCN se exponen a las proteínas de prueba durante 24 horas antes de la retirada del medio y la preparación del material celular para su análisis por SDS-PAGE e inmunotransferencia. Después de la separación de las proteínas celulares en geles de poliacrilamida Novex con Tris-glicina al 4-20 %, las proteínas se transfieren a nitrocelulosa y a continuación se detecta la presencia de la proteína SNARE apropiada con anticuerpos obtenidos de fuentes comerciales. En este caso, los anticuerpos fueron específicos para la proteína SNARE syntaxina.

[0261] Con referencia a la figura 16, la proteína tratada con el factor Xa es claramente menos eficaz en la escisión de syntaxina que la proteína no tratada con FXa. Por lo tanto, la invención hace posible una reducción de la eficacia de la proteína modificada.

Ejemplo 30 - Demostración de la reducida actividad celular *in vitro* de una proteína manipulada para incorporar un sitio de escisión proteásica para trombina en el dominio LC de LHA-EGF [SXN1974]

[0262] El producto proteínico del ejemplo 24 se expresa y purifica. La proteína purificada se expone a la proteasa trombina antes de su evaluación en un ensayo de neuronas de médula espinal (SCN) *in vitro*. La preparación de las SCN es una técnica bien establecida y se describe en la bibliografía [B. R. Ransom, E. Neale, M. Henkart, P. N. Bullock, P. G. Nelson, Mouse spinal cord in cell culture. I. Morphology and intrinsic neuronal electrophysiologic properties, J. Neurophysiol. 40 (1977): 1132-1150; S. C. Fitzgerald, A Dissection and Tissue Culture Manual of the Nervous System, Alan R. Liss Inc, Nueva York, 1989]. La proteína de prueba se prepara en diversas concentraciones por dilución en el medio de cultivo. Las SCN se exponen a las proteínas de prueba durante 24 horas antes de la retirada del medio y la preparación del material celular para su análisis por SDS-PAGE e inmunotransferencia. Después de la separación de las proteínas celulares en geles de poliacrilamida Novex con Tris-glicina al 4-20 %, las proteínas se transfieren a nitrocelulosa y a continuación se detecta la presencia de la proteína SNARE apropiada con anticuerpos obtenidos de fuentes comerciales. En este caso, los anticuerpos fueron específicos para la proteína SNARE SNAP-25. La figura 17 demuestra la escisión de SNAP-25 por la L(Thr)HA-EGF tratada con trombina en comparación con L(Thr)HA-EGF sin tratar.

Ejemplo 31 - Creación de una proteína BoNT/A recombinante que incorpora un sitio de reconocimiento para trombina en LC

[0263] La secuencia primaria de una BoNT/A recombinante con actividad endopeptidásica que contiene un sitio de activación proteásica específico para enterocinasa incorporado (SEQ ID 18) se revisa para determinar la

presencia de porciones de la secuencia aminoacídica con similitud con el sitio de reconocimiento prototípico para trombina (GRG). Un simple análisis de caracteres de texto de la secuencia primaria identificó la secuencia ¹⁰³GRM₁₀₅ dentro del dominio LC que puede someterse a ingeniería de proteínas. La localización del péptido en la estructura terciaria de H_N se predice a partir de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA) como guía. Se usa un software disponible públicamente (tal como FirstGlance in Jmol (<http://molvis.sdsc.edu/fqii/fq.htm?mol=3bta>)) para identificar que la secuencia peptídica de H_N se localiza en la superficie. Por lo tanto, esta localización es de buena accesibilidad para proteasas.

[0264] Una mutagénesis dirigida de SEQ ID 17 (que codifica SEQ ID 18) con un cebador diseñado para cambiar el codón de M₁₀₅ a G cambia la secuencia ¹⁰³GRM₁₀₅ a GRG, que es un sustrato para la escisión por trombina. La mutagénesis se llevó a cabo mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene).

[0265] El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La secuencia aminoacídica final del producto de expresión se ilustra en SEQ ID 32.

Ejemplo 32 - Creación de una proteína BoNT/A recombinante que incorpora un sitio de reconocimiento para el factor Xa en LC

[0266] La secuencia primaria de una BoNT/A recombinante con actividad endopeptidásica que contiene un sitio de activación proteásica específico para enterocinasa incorporado (SEQ ID 18) se revisa para determinar la presencia de porciones de la secuencia aminoacídica con similitud con el sitio de reconocimiento prototípico para el factor Xa (IEGR). Un simple análisis de caracteres de texto de la secuencia primaria identificó la secuencia IDSL dentro del dominio LC que puede someterse a ingeniería de proteínas. La localización del péptido en la estructura terciaria del dominio LC se predice a partir de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA) como guía. Se usa un software disponible públicamente (tal como FirstGlance in Jmol (<http://molvis.sdsc.edu/fqii/fq.htm?mol=3bta>)) para identificar que la secuencia peptídica de LC se localiza en la superficie. Por lo tanto, esta localización es de buena accesibilidad para proteasas.

[0267] Una mutagénesis dirigida de SEQ ID 17 (que codifica SEQ ID 18) con un cebador diseñado para cambiar los codones de ²⁷⁶SL a GR cambia la secuencia IDSL a IDGR, que es un sustrato para la escisión por el factor Xa. La mutagénesis se llevó a cabo mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene).

[0268] El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La secuencia aminoacídica final del producto de expresión se ilustra en SEQ ID 33.

Ejemplo 33 - Creación de una proteína BoNT/A recombinante que incorpora un sitio de reconocimiento para el factor Xa en H_N

[0269] La secuencia primaria de una BoNT/A recombinante con actividad endopeptidásica que contiene un sitio de activación proteásica específico para enterocinasa incorporado (SEQ ID 18) se revisa para determinar la presencia de porciones de la secuencia aminoacídica con similitud con el sitio de reconocimiento prototípico para el factor Xa (IEGR). Un simple análisis de caracteres de texto de la secuencia primaria identificó la secuencia ⁵⁶²GKSR₅₆₅ dentro del dominio H_N que puede someterse a ingeniería de proteínas. La localización del péptido en la estructura terciaria del dominio H_N se predice a partir de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA) como guía. Se usa un software disponible públicamente (tal como FirstGlance in Jmol (<http://molvis.sdsc.edu/fqii/fq.htm?mol=3bta>)) para identificar que la secuencia peptídica de LC se localiza en la superficie. Por lo tanto, esta localización es de buena accesibilidad para proteasas.

[0270] Se llevó a cabo una mutagénesis dirigida de SEQ ID 17 (que codifica SEQ ID 18) con un cebador diseñado para cambiar la región peptídica GKSR a IEGR, que es un sustrato para la escisión por el factor Xa. La mutagénesis se llevó a cabo mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene).

[0271] El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La secuencia aminoacídica final del producto de expresión se ilustra en SEQ ID 34.

15 **Ejemplo 34 - Creación de una proteína BoNT/E recombinante que incorpora un sitio de reconocimiento para trombina en LC**

[0272] La secuencia primaria de una BoNT/E recombinante con actividad endopeptidásica (código de acceso de nucleótidos AM695755; número de Uniprot A8Y867) se revisa para determinar la presencia de porciones de la secuencia aminoacídica con similitud con el sitio de reconocimiento prototípico para trombina (LVPRGS). Un simple análisis de caracteres de texto de la secuencia primaria identificó la secuencia ¹⁸⁶FSPEYS¹⁹¹ dentro del dominio LC que puede someterse a ingeniería de proteínas. La localización del péptido en la estructura terciaria del dominio H_N se predice a partir de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA) como guía. Se usa un software disponible públicamente (tal como FirstGlance in Jmol (<http://molvis.sdsc.edu/fqii/fq.htm?mol=3bta>)) para identificar que la secuencia peptídica de LC se localiza en la superficie. Por lo tanto, esta localización es de buena accesibilidad para proteasas.

[0273] Se lleva a cabo una mutagénesis dirigida con un cebador diseñado para cambiar la región peptídica FSPEYS a IVPRFS, que es un sustrato para la escisión por trombina. La mutagénesis se llevó a cabo mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene).

[0274] El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La secuencia aminoacídica final del producto de expresión se ilustra en SEQ ID 35.

40 **Ejemplo 35 - Creación de una proteína BoNT/E recombinante que incorpora un sitio de reconocimiento para el factor Xa en H_N**

[0275] La secuencia primaria de una BoNT/E recombinante con actividad endopeptidásica (código de acceso de nucleótidos AM695755; número de Uniprot A8Y867) se revisa para detectar un sitio de inserción potencial para un péptido de reconocimiento por el factor Xa (IEGR). La comparación de la secuencia primaria de BoNT/E con la de BoNT/A y la correspondiente localización del péptido en la estructura terciaria del dominio H_N, que se predice a partir de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA), permite concluir que la región ⁷²⁷TLEE es adecuada para su conversión en IEGR por ingeniería de proteínas.

[0276] Se lleva a cabo una mutagénesis dirigida con un cebador diseñado para cambiar la región peptídica TLEE a IEGR, que es un sustrato para la escisión por el factor Xa. La mutagénesis se llevó a cabo mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene).

[0277] El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora

en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La secuencia aminoacídica final del producto de expresión se ilustra en SEQ ID 36.

5 Ejemplo 36 - Creación de una proteína quimérica LHE-VIPr que incorpora un sitio de reconocimiento para trombina en LC

[0278] La secuencia primaria de una proteína quimérica construida por la fusión genética del fragmento LH_N de BoNT/E y un análogo de la secuencia del péptido intestinal vasoactivo humano (VIPr) se revisa para determinar la presencia de porciones de la secuencia aminoacídica con similitud con el sitio de reconocimiento prototípico para trombina (GRG). Un simple análisis de caracteres de texto de la secuencia primaria identificó la secuencia ¹⁰³GGI¹⁰⁵ dentro del dominio LC de la quimera. La localización del péptido en la estructura terciaria de LC/E se predice a partir de la estructura cristalina de rayos X de LC/E (PDB: 1T3A) como guía. Se usa un software disponible públicamente (tal como Jmol (<http://www.rcsb.org/pdb/explore/jmol.do?structureId=1T3A&ionNumber=1>)) para identificar que la secuencia peptídica de LC se localiza cerca de la superficie. Por lo tanto, esta localización es una región de buena accesibilidad para proteasas.

[0279] Se lleva a cabo una mutagénesis dirigida con un cebador diseñado para cambiar la región peptídica GGI a GRG, mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene). El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La ORF final que incorpora el sitio para trombina se usa para codificar la secuencia aminoacídica del producto de expresión ilustrado en SEQ ID 37.

30 Ejemplo 37 - Creación de una proteína quimérica LHE-VIPr que incorpora un sitio de reconocimiento para el factor Xa en H_N

[0280] La secuencia primaria de una proteína quimérica construida por la fusión genética del fragmento LH_N de BoNT/E y un análogo de la secuencia del péptido intestinal vasoactivo humano (VIPr) se revisa para determinar la presencia de porciones de la secuencia aminoacídica con similitud con el sitio de reconocimiento prototípico para el factor Xa (IEGR).

[0281] Un simple análisis de caracteres de texto de la secuencia primaria identificó la secuencia ⁵⁸⁵GENN dentro del dominio H_N. La localización del péptido en la estructura terciaria de H_N/E se predice a partir de la estructura cristalina de rayos X de BoNT/A (PDB: 3BTA) como guía. Se usa un software disponible públicamente (tal como FirstGlance in Jmol (<http://molvis.sdsc.edu/fqii/fq.htm?mol=3bta>)) para identificar que la secuencia peptídica de LC se localiza en la superficie. Por lo tanto, esta localización es una región de buena accesibilidad para proteasas.

[0282] Se lleva a cabo una mutagénesis dirigida con un cebador diseñado para cambiar la región peptídica GENN a IEGR, mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene). El uso de codones de *E. coli* se evalúa por referencia a programas de software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La ORF final que incorpora el sitio para el factor Xa se usa para codificar la secuencia aminoacídica del producto de expresión ilustrado en SEQ ID 38.

55 Ejemplo 38 - Creación de una proteína quimérica LHE-VIPr que incorpora un sitio de reconocimiento para el factor Xa en LC

[0283] La secuencia primaria de una proteína quimérica construida por la fusión genética del fragmento LH_N de BoNT/E (que incorpora un dominio de reconocimiento del sustrato mutado (K228D)) y un análogo de la secuencia

del péptido intestinal vasoactivo humano (VIPr) se revisa para determinar la presencia de porciones de la secuencia aminoacídica expuestas en la superficie de la proteína que pueden manipularse para asemejarse al sitio de reconocimiento prototípico para el factor Xa (IEGR).

5 [0284] El análisis de la secuencia primaria identificó la secuencia ²⁶⁸VAQY dentro del dominio LC. La localización del péptido en la estructura terciaria de LC/E se predice a partir de la estructura cristalina de rayos X de BoNT/E (PDB: 1T3A) como guía. Se usa un software disponible públicamente (tal como Jmol (<http://www.rcsb.org/pdb/explore/imol.do?structureId=1T3A&bionumber=1>)) para identificar que la secuencia peptídica de LC se localiza en la superficie. Por lo tanto, esta localización es una región de buena accesibilidad para
10 proteasas.

[0285] Se lleva a cabo una mutagénesis dirigida con un cebador diseñado para cambiar la región peptídica VAQY a IEGR, mediante las técnicas moleculares estándar para mutagénesis (por ejemplo, la metodología de mutagénesis Quickchange de Stratagene). El uso de codones de *E. coli* se evalúa por referencia a programas de
15 software tales como Graphical Codon Usage Analyser (Geneart) y el contenido de GC en % y la relación del uso de codones se evalúan por referencia a tablas publicadas de uso de codones (por ejemplo, GenBank, edición 143, 13 de septiembre de 2004) para asegurar que la mutagénesis no resulta en una mala utilización de codones. El ADN mutagenizado se incorpora en un vector de clonación estándar, por ejemplo, pCR4, antes de su transformación en el hospedador *E. coli*. La integridad del ADN de la ORF se comprueba por secuenciación. La ORF final que incorpora
20 el sitio para el factor Xa se usa para codificar la secuencia aminoacídica del producto de expresión ilustrado en SEQ ID 39.

Ejemplo 39 - Escisión de una proteína SNARE por una neurotoxina clostridial modificada (LH_N) que tiene las propiedades descritas por Lin y col. (documento WO02/044199)

25 [0286] Se prepararon neuronas de médula espinal embrionaria por disección de ratas Sprague Dawley E15, que se disociaron antes de sembrarlas en placas de 96 pocillos recubiertas con Matrigel a una densidad de 125.000 células por pocillo en el medio (MEM tamponado con bicarbonato de sodio, suero de caballo inactivado al 5 %, D-glucosa al 0,6 %, suplemento de medio N1 al 2 %, 40 ng/ml de corticosterona, 20 ng/ml de triyodotironina).

30 [0287] Después de tres semanas, las células se incubaron con medio fresco que contenía la cadena ligera recombinante del serotipo C (LC/C) o una neurotoxina clostridial modificada que constaba de la cadena de translocación y la cadena ligera del serotipo C (LH_N/C) en concentraciones semilogarítmicas entre 180 nM y 0,18 nM durante 24 horas a 37 °C en una atmósfera humidificada con CO₂ al 5 %.

35 [0288] Las células se lisaron con tampón de carga de SDS-PAGE que contenía DTT. Las proteínas se separaron por SDS-PAGE (Tris-Bis al 12 %), se transfirieron a una membrana de nitrocelulosa y la syntaxina se detectó con un anticuerpo de conejo dirigido contra la syntaxina 2 (Synaptic Systems, n.º de catálogo 110022). El anticuerpo unido se detectó con un anticuerpo dirigido contra IgG de conejo conjugado con peroxidasa y el sustrato West dura para la obtención de señales fluorescentes. Las imágenes se escanearon y cuantificaron mediante un
40 software de Syngene (GeneTools). La figura 1 muestra la syntaxina escindida como porcentaje de la syntaxina total y confirma una actividad de neurotoxina para la neurotoxina clostridial modificada que carece de un dominio de unión H_C funcional (LH_N/C), pero ninguna actividad de neurotoxina detectable para la neurotoxina clostridial modificada que carece de un dominio de translocación H_N funcional (LC/C).

LISTADO DE SECUENCIAS

[0289]

SEQ ID NO: 1

ATGATTTCCGAATTTGGCTCGGAGTTCATGCCAATTACGATTAACAATTTTAACTATAGTGATCCGGTGG
ATAATAAAAAACATTTTATACCTGGATACCCACTTGAATACTCTTGCCAATGAGCCTGAAAAAGCCTTTCG
CATAACGGGTAAACATTTGGGTCATTCCGGACCGTTTTAGCCGGAACCTAACCCTAATCTGAATAAACCT
CCGCGTGTACAGTCTCCGAAAAGTGGGTATTACGATCCAAATTATCTGAGTACCGATTCAGACAAGGATA
CGTTTCTGAAAGAAATCATAAACTTTTCAAAAGAATCAACTCCCGTGAAATCGGTGAAGAGCTGATCTA
CCGTCTGTCGACGGACATTCCTTTTCCGGGAAACAATAACACTCCCATTAATACCTTCGACTTTGATGTC
GATTTCAACTCAGTCGATGTGAAAACCCGCCAGGGTAATAACTGGGTAAAACTGGATCCATTAACCCGT

5

CCGTTATTATCACAGGTCCTCGTGAAAATATTATAGATCCTGAGACCTCCACGTTCAAGCTGACGAATAA
 CACTTTTGCGGCACAGGAAGGGTTTGGTGCCCTTTCAATTATCTCTATCTCTCCGCGCTTCATGTTAACG
 TATTCTAACGCAACCAACGATGTTGGCGAGGGCCGCTTCAGCAAAAGTGAATTCTGTATGGATCCCATTC
 TGATCTTGATGCATGAGCTTAACCACGCTATGCATAATCTTTATGGTATTGCAATCCCAAACGATCAGAC
 GATCTCCAGCGTTACATCTAACATATTCTACAGCCAATATAATGTGAAGCTCGAATATGCAGAGATTTAC
 GCCTTCGGTGGGCCGACCATTGACCTCATTCCAAAGTCTGCCCCTAAGTACTTTGAGGAAAAAGCGTTGG
 ATTACTATCGTAGCATCGCGAAACGCCTGAATTCAATTACAACCTGCAAACCCATCTAGCTTCAACAAATA
 CATCGGAGAATATAAAACAAAAGCTGATACGCAATATCGCTTTGTGGTGAATCGTCCGGGGAAGTGACA
 GTTAATCGAAATAAATTTGTTGAACTCTATAATGAATTAACGCAGATCTTCACAGAATTTAATTATGCTA
 AAATCTATAATGTACAGAACCGGAAAATTTATCTCAGTAATGTATACACACCGGTGACTGCTAACATTCT
 GGACGATAACGTCTACGATATTCAAAATGGCTTTAATATCCCGAAGAGCAACTTGAATGTCCTCTTCATG
 GGGCAGAACTTGTACGTAACCCAGCGCTGCGAAAAGTTAACCCAGAAAATATGTTGTACCTCTTTACAA
 AATTCTGTGTAGACGCCGACGATGACGATAAACTGTACAACAAAACCCCTGCAATGCCGTGAACCTCTGGT
 TAAGAACACCGACCTGCCGTTTATTGGGGACATCAGTGATGTCAAAACGGATATTTTTCTTCGGAAGGAT
 ATTAATGAGGAAACCGAAGTGATACCTGACAATGTGTGCGGTAGATCAGGTAATCCTGAGTAAGAACACCA
 GCGAGCATGGGCAGCTGGATCTGTTGTATCCGAGCATTGACAGCGAGTCGGAATACTGCCCGGCGAAAA
 TCAAGTTTTTTATGACAATCGGACCCAGAATGTTGATTATCTGAATAGTTACTATTACTTGGAGAGCCAA
 AAATTATCAGATAATGTGGAAGACTTTACCTTTACCCGGTCTATCGAAGAGGCGCTGGATAACAGCGCGA
 AAGTTTACACTTATTTTCCACGCTCGCAAACAAAGTTAATGCTGGCGTACAGGGTGGATTATTTCTTAT
 GTGGGCGAATGATGTGGTAGAGGACTTTACAACCAACATCCTGCGCAAAGACACTTTAGACAAAATTTCT
 GACGTCTCGGCCATTATCCCGTATATAGGTCCGGCCTTAAACATAAGCAATTCGGTTCGCCGTGGCAACT
 TCACAGAAGCCTTCGCTGTGACTGGTGTGACCATTCTGTTGGAAGCATTTCCTGAGTTTACGATCCCGGC
 TCTGGGCGCATTGTGAATTTACTCTAAAGTTCAGGAACGAAATGAAATTATAAAACTATCGATAATTGC
 CTGGAACAGCGTATCAAGAGATGGAAGGATTCCTATGAGTGGATGATGGGGACCTGGCTGTCAAGAATTA
 TCACACAGTTTAAATAACATATCCTATCAAATGTATGATAGCTTAACTATCAAGCAGGAGCGATAAAGGC
 GAAAATTGACCTGGAATACAAGAAATATTCTGGTTCCGATAAAGAGAATATTAAAAGTCAGGTGGAAAAT
 CTGAAAATAGTTTAGATGTCAAAATTTCTGAGGCGATGAATAACATTAACAAATTCATCCGCGAGTGCA
 GTGTAACCTATTTGTTTAAAGAATATGTTACCCAAAGTTATCGACGAACTGAATGAATTTGATCGTAATAC
 CAAAGCCAAATTGATCAACCTCATCGACTCTCATAACATCATTCTGGTGGGAGAAGTCGACAAACTGAAA
 GCTAAGGTGAATAACAGCTTCCAGAATACAATTCCGTTTAAATATTTTCTCATACACCAATAACTCGCTGC
 TTAAAGATATTATCAACGAATATTTAATCTGGAGGGTGGCGGTGGCAGTGGCGGTGGCGGATCCGGCGG
 TGGCGGTAGCGCACTGGATAATTCAGATTCCGAATGTCCACTGTACACGATGGTTATTGTCTTCATGAT
 GCGGTGTGCATGTATATAGAAGCGTTAGATAAATACGCTTGCAACTGCGTGGTTGGCTATATCGGCGAAC
 GTTGTCAGTATCGTGATTTAAAGTGGTGGGAATTACGCTAATGA

SEQ ID NO: 2

ISEFGSEFMPITINNFNYS DPVDNKNILYLDTHLNTLANEPEKAFRITGNIWVIPDRFSRNSNP NLNKP
RVTS PKSGYYDPNYLSTDSDKDTFLKEIIKLFKRINSREIGEELIYRLSTDIPFPGNNNTPI NTFDFD
FNSVDVKTRQGNNWVK TGSINPSV IITGPRENIIDPETSTFKLTNNTFAAQEGFGALSIISISPRFMLTY
SNATNDVGEGRFSKSEFCMDPILILMH ELNHAMHNLYGIAIPNDQTISSVTSNIFYSQYNVKLEYAEIYA
FGGPTIDLIPKSARKYFEEKALDY YRSIAKRLNSITTANPSSFNKYIGEYKQKLIRKYRFVVESSGEVTV
NRNKFVELYNELTQIFTEFN YAKIYNVQNRKIYLSNVYTPVTANILDDNVYDIQNGFNIPKSNLNVLFMG
QNL SRNPALRKVPENMLYLFTKFCVDADDDDKLYNKT LQCRELLVKNTDLPFIGDISDVKT DIFLRKDI
NEETEVI PDNVSDQVILSKNTSEHGQLDLLYPSIDSESEILPGENQVFYDNRTQNV DYLSYYYYLESQK

LSDNVEDFTFTRSIEEALDNSAKVYTYFPTLANKVNAGVQGGLFLMWANDVVEDFTTNILRKDTL DKISD
VSAIIPYIGPALNISNSVRRGNFTEAFV TGVTTILLEAFPEFTIPALGAFVIYSKVQERNEI IKTIDNCL
EQRIKRWKDSYEWMMGTWLSRIITQFN NISYQMYDSLNYQAGAIKAKIDLEYKKYSGSDKENIKSQVENL
KNSLDVKISEAMNNINKFIRECSVTYLFKNM LPKVIDELNEFDRNTKAKLINLIDSHNIILVGEVDKLKA
KVNNSFQNTIPFNIFSYTNNSLLKDIINEYFNLEGGGGSGGGGSGGGGSALDNSDSECPLSHDGYCLHDG
VCMYIEALDKYACNCVVG YIGERCQYRDLKWWELR

5 SEQ ID NO: 3

ATGATTTCCGAATTTGGCTCGGAGTTCATGCCAATTACGATTAACAATTTTAACTATAGTGATCCGGTGG
 ATAATAAAAACATTTTATACCTGGATACCCACTTGAATACTCTTGCCAATGAGCCTGAAAAAGCCTTTTCG
 CATAACGGGTAAACATTTGGGTCAATCCGGACCGTTTTAGCCGGAACCTCTAACCCCTAATCTGAATAAACCT
 CCGCGTGTACGTCTCCGAAAAGTGGGTATTACGATCCAAATTATCTGAGTACCGATTGAGACAAGGATA
 CGTTTCTGAAAGAAATCATAAACTTTTCAAAAGAATCAACTCCCGTGAAATCGGTGAAGAGCTGATCTA
 CCGTCTGTGACGGACATTCTTTTCCGGGAAACAATAAACTCCCATTAAATACCTTCGACTTTGATGTC
 GATTTCAACTCAGTCGATGTGAAAAACCGCCAGGGTAATAACTGGGTAAAACTGGATCCATTAACCCGT
 CCGTTATTATCACAGGTCTCGTGAAAATATTATAGATCCTGAGACCTCCACGTTCAAGCTGACGAATAA
 CACTTTTGCGGCACAGGAAGGGTTTGGTGCCCTTTCAATTATCTCTATCTCTCCGCGCTTCATGTTAACG
 TATTCTAACGCAACCAACGATGTTATCGAGGGCCGCTTCAGCAAAAGTGAATTCTGTATGGATCCCATTTC
 TGATCTTGATGCATGAGCTTAACCACGCTATGCATAATCTTTATGGTATTGCAATCCCAAACGATCAGAC
 GATCTCCAGCGTTACATCTAACATATTCTACAGCCAATATAATGTGAAGCTCGAATATGCAGAGATTTAC
 GCCTTCGGTGGGCCGACCATTGACCTCATTCCAAAGTCTGCCCGTAAGTACTTTGAGGAAAAAGCGTTGG
 ATTACTATCGTAGCATCGCGAAACGCCTGAATTCAATTACAACCTGCAAACCCATCTAGCTTCAACAAATA
 CATCGGAGAATATAAACAAAAGCTGATACGCAAAATATCGCTTTGTGGTGAATCGTCCGGGGAAGTGACA
 GTTAATCGAAATAAATTTGTTGAACTCTATAATGAATTAACGCAGATCTTCACAGAATTTAATTATGCTA
 AAATCTATAATGTACAGAACCGGAAAATTTATCTCAGTAATGTATACACACCGGTGACTGCTAACATTCT
 GGACGATAACGTCTACGATATTCAAAATGGCTTTAATATCCCGAAGAGCAACTTGAATGTCTCTTTCATG
 GGGCAGAACTTGTACGTAACCCAGCGCTGCGAAAAGTTAACCCAGAAAATATGTTGTACCTCTTTACAA
 AATTCTGTGTAGACGCCGACGATGACGATAAACTGTACAACAAAACCCCTGCAATGCCGTGAACCTTCTGGT
 TAAGAACACCGACCTGCCGTTTATTGGGGACATCAGTGATGTCAAACCGGATATTTTCTTCGGAAGGAT
 ATTAATGAGGAAACCGAAGTGATACCTGACAATGTGTGGTAGATCAGGTAATCCTGAGTAAGAACACCA
 GCGAGCATGGGCAGCTGGATCTGTTGTATCCGAGCATTGACAGCGAGTCGGAAATACTGCCCGGCGAAAA
 TCAAGTTTTTTTATGACAATCGGACCCAGAATGTTGATTATCTGAATAGTTACTATTACTTGGAGAGCCAA
 AAATTATCAGATAATGTGGAAGACTTTACCTTTACCCGGTCTATCGAAGAGGCGCTGGATAACAGCGCGA
 AAGTTTACACTTATTTTCCACGCTCGCAACAAAGTTAATGCTGGCGTACAGGGTGGATTATTTCTTAT
 GTGGGCGAATGATGTGGTAGAGGACTTTACAACCAACATCCTGCGCAAAGACACTTTAGACAAAATTTCT
 GACGTCTCGGCCATTATCCCGTATATAGGTCCGGCCTTAAACATAAGCAATTCGGTTCGCCGTGGCAACT
 TCACAGAAGCCTTCGCTGTGACTGGTGTGACCATTTCTGTTGGAAGCATTTCTGAGTTTACGATCCCGGC
 TCTGGGCGCATTTGTAATTTACTCTAAAGTTCAGGAACGAAATGAAATTATAAAAACTATCGATAATTGC
 CTGGAACAGCGTATCAAGAGATGGAAGGATTCTATGAGTGGATGATGGGGACCTGGCTGTCAAGAATTA
 TCACACAGTTTAATAACATATCCTATCAAATGTATGATAGCTTAACTATCAAGCAGGAGCGATAAAGGC
 GAAAATTGACCTGGAATACAAGAAATATTCTGGTTCGGATAAAGAGAATATTAAAAGTCAGGTGGAAAAT
 CTGAAAAATAGTTTAGATGTCAAATTTCTGAGGCGATGAATAACATTAACAAATTCATCCGCGAGTGCA
 GTGTAACCTTATTTGTTTAAGAATATGTTACCCAAAGTTATCGACGAACTGAATGAATTTGATCGTAATAC
 CAAAGCCAAATTGATCAACCTCATCGACTCTCATAACATCATTCTGGTGGGAGAAGTCGACAAACTGAAA

GCTAAGGTGAATAACAGCTTCCAGAATACAATTCCGTTTAATATTTTCTCATACACCAATAACTCGCTGC
TTAAAGATATTATCAACGAATATTTTAATCTGGAGGGTGGCGGTGGCAGTGGCGGTGGCGGATCCGGCGG
TGGCGGTAGCGCACTGGATAATTCAGATTCCGAATGTCCACTGTCACACGATGGTTATTGTCTTCATGAT
GGCGTGTGCATGTATATAGAAGCGTTAGATAAATACGCTTGCAACTGCGTGGTTGGCTATATCGGCGAAC
GTTGTCACTATCGTGATTTAAAGTGGTGGGAATTACGCTAATGA

SEQ ID NO: 4

MISEFGSEFMPITINNFNYSDPVDNKNILYLDTHLNTLANEPEKAFRITGNIWVIPDRFSRNSNPNLNKP
PRVTSKSGYYDPNYLSTDSKDTFLKEI IKLFKRINSREIGEELIYRLSTDIPFPGNNNTPINTFDFDV
DFNSVDVKTRQGNWVKTSINPSVITGPRENIIDPETSTFKLTNNTFAAQEGFGALSIIISISPRFMLT
YSNATNDVIEGRFSKSEFCMDPILILMHELNHAMHNLGYIAIPNDQTISSVTSNIFYSQYNVKLEYAEIY
AFGGPTIDLIPKSARKYFEEKALDYYSIAKRLNSITTANPSSFNKYIGEYKQKLIRKYRFVVESSGEVT
VNRNKFVELYNELTQIFTEFNKYIYNVQNRKIYLSNVYTPVTANILDDNVYDIQNGFNI PKSNLNLVLFM
GQNL SRNPALRKVNPENMLYLFTKFCVDADDDDKLYNKTLQCRELLVKNTDLPFIGDISDVKTDFLRKD
INEETEVI PDNVSDQVILSKNTSEHGQDL LYP SIDSESEILPGENQVFYDNRTQNV DY LNSYYLESQ
KLSDNVEDFTFTRSIEEALDNSAKVYTYFPTLANKVNAGVQGGFLMWANDVVEDFTTNILRKDTLDKIS
DVSAIIPYIGPALNISNSVRRGNFTEAFVAVTGVTILLEAFPEFTIPALGAFVIYSKVQERNEIIKTIDNC
LEQRIKRWKDSYEWMMGTWLSRIITQFN NISYQMYDSLNYQAGAIKAKIDLEYKKYSGSDKENIKSQVEN
LKNSLDVKISEAMNNINKFIRECSVTYLFKNMLPKVIDELNEFDRNTKAKLINLIDSHNIILVGEVDKLK
AKVNNSFQNTIPFNIFSYTNNSLLKDIINEYFNLEGGGGSGGGGSGGGGSALDNSDSECPLSHDGYCLHD
5 GVCMYIEALDKYACNCVVG YIGERCQYRDLKWWELR

SEQ ID NO: 5

ATGATTTCCGAATTTGGCTCGGAGTTCATGCCAATTACGATTAACAATTTTAACTATAGTGATCCGGTGG
 ATAATAAAAACATTTTATACCTGGATACCCACTTGAATACTCTTGCCAATGAGCCTGAAAAAGCCTTTTCG
 CATAACGGGTAAACATTTGGGTCAATCCGGACCGTTTTAGCCGGAACCTCTAACCCCTAATCTGAATAAACCT
 CCGCGTGTACGTCTCCGAAAAGTGGGTATTACGATCCAAATTATCTGAGTACCGATTTCAGACAAGGATA
 CGTTTCTGAAAGAAATCATAAACTTTTCAAAAGAATCAACTCCCGTGAAATCGGTGAAGAGCTGATCTA
 CCGTCTGTGACGGACATTCCCTTTTCCGGGAAACAATAACACTCCCATTAAATACCTTCGACTTTGATGTC
 GATTTCAACTCAGTCGATGTGAAAACCCGCCAGGGTAATAACTGGGTAAAACTGGATCCATTAACCCGT
 CCGTTATTATCACAGGTCTCTGTGAAAATATTATAGATCCTGAGACCTCCACGTTCAAGCTGACGAATAA
 CACTTTTGCGGCACAGGAAGGGTTTGGTGCCCTTTCAATTATCTCTATCGTTCGCGCTTCTCCTTAACG
 TATTCTAACGCAACCAACGATGTTGGCGAGGGCCGCTTCAGCAAAAGTGAATTCTGTATGGATCCCATTC
 TGATCTTGATGCATGAGCTTAACCACGCTATGCATAATCTTTATGGTATTGCAATCCCAAACGATCAGAC
 GATCTCCAGCGTTACATCTAACATATTCTACAGCCAATATAATGTGAAGCTCGAATATGCAGAGATTTAC
 GCCTTCGGTGGGCCGACCATTGACCTCATTCCAAAGTCTGCCCGTAAGTACTTTGAGGAAAAAGCGTTGG
 ATTACTATCGTAGCATCGCGAAACGCCTGAATTCAATTACAACCTGCAAACCCATCTAGCTTCAACAAATA
 CATCGGAGAATATAAACAAAAGCTGATACGCAAATATCGCTTTGTGGTTCGAATCGTCCGGGGAAGTGACA
 GTTAATCGAAATAAATTTGTTGAACTCTATAATGAATTAACGCAGATCTTCACAGAATTTAATTATGCTA
 AAATCTATAATGTACAGAACCGGAAAATTTATCTCAGTAATGTATACACACCGGTGACTGCTAACATTCT
 GGACGATAACGTCTACGATATTCAAAATGGCTTTAATATCCCGAAGAGCAACTTGAATGTCCTCTTCATG
 GGGCAGAACTTGTCACGTAACCCAGCGCTGCGAAAAGTTAACCCAGAAAATATGTTGTACCTCTTTACAA
 AATTCTGTGTAGACGCCGACGATGACGATAAACTGTACAACAAAACCTGCAATGCCGTGAACCTTCTGGT
 TAAGAACACCGACCTGCCGTTCAATTGGGGACATCAGTGATGTCAAACGGATATTTTTCTTCGGAAGGAT

ATTAATGAGGAAACCGAAGTGATACCTGACAATGTGTCGGTAGATCAGGTAATCCTGAGTAAGAACACCA
GCGAGCATGGGCAGCTGGATCTGTTGTATCCGAGCATTGACAGCGAGTCGGAAATACTGCCCGGCGAAAA
TCAAGTTTTTTATGACAATCGGACCCAGAATGTTGATTATCTGAATAGTTACTATTACTTGGAGAGCCAA
AAATTATCAGATAATGTGGAAGACTTTACCTTTACCCGGTCTATCGAAGAGGCGCTGGATAACAGCGCGA
AAGTTTACACTTATTTTCCACGCTCGCAAAACAAAGTTAATGCTGGCGTACAGGGTGGATTATTTCTTAT
GTGGGCGAATGATGTGGTAGAGGACTTTACAACCAACATCCTGCGCAAAGACACTTTAGACAAAATTTCT
GACGTCTCGGCCATTATCCCGTATATAGGTCCGGCCTTAAACATAAGCAATTCGGTTCGCCGTGGCAACT
TCACAGAAGCCTTCGCTGTGACTGGTGTGACCATTCTGTTGGAAGCATTTCCTGAGTTTACGATCCCGGC
TCTGGGCGCATTGTGAATTTACTCTAAAGTTCAGGAACGAAATGAAATTATAAAAACTATCGATAATTGC
CTGGAACAGCGTATCAAGAGATGGAAGGATTCCCTATGAGTGGATGATGGGGACCTGGCTGTCAAGAATTA
TCACACAGTTTAATAACATATCCTATCAAATGTATGATAGCTTAACTATCAAGCAGGAGCGATAAAGGC
GAAAATTGACCTGGAATACAAGAAATATTCTGGTTCGGATAAAGAGAATATTAAGAGTCAGGTGGAAT
CTGAAAAATAGTTTAGATGTCAAATTTCTGAGGCGATGAATAACATTAACAAATTCATCCGCGAGTGCA
GTGTAACCTATTTGTTTAAGAATATGTTACCCAAAGTTATCGACGAAGTGAATGAATTTGATCGTAATAC
CAAAGCCAAATTGATCAACCTCATCGACTCTCATAACATCATTCTGGTGGGAGAAGTCGACAAACTGAAA
GCTAAGGTGAATAACAGCTTCCAGAATACAATTCGGTTAATATTTCTCATAACCAATAACTCGCTGC
TTAAAGATATTATCAACGAATATTTAATCTGGAGGGTGGCGGTGGCAGTGGCGGTGGCGGATCCGGCGG
TGGCGGTAGCGCACTGGATAATTCAGATTCCGAATGTCCACTGTCACACGATGGTTATTGTCTTCATGAT
GGCGTGTGCATGTATATAGAAGCGTTAGATAAATACGCTTGCAACTGCGTGGTTGGCTATATCGGCCGAAC
GTTGTCAAGTATCGTGATTTAAAGTGGTGGGAATTACGCTAATGA

SEQ ID NO: 6

MISEFGSEFMPITINNFNYS DPVDNKNILYLDTHLNTLANEPEKA FRITGNIWVIPDRFSRNSNP NLNKP
PRVTS PKSGYYDPNYLSTDSKD TFLKEI IKLFKRINSREIGEELIYRLSTDIPFPGNNTPI NTFDFDV
DFNSVDVKTRQGNWVK TGSINPSV IITGP RENIIDPETSTFKLTNNTFAAQEGFGALSII SISVPRFSLT
YSNATNDVGEGRFSKSEFCMDPILILMH ELNHAMHNLYGIAIPNDQTISSVTSNIFYSQYNVKLEYAEIY
AFGGPTIDLIPKSARKYFEEKALDYRSIAKRLNSITTANPSSFNKYIGEYKQKLIRKYRFVVESSGEVT
VNRNKFVELYNELTQIFTEFN YAKIYNVQNRKIYLSNVYTPVTANILDDNVYDIQNGFNIPKSNLNVLFM
GQNLSRNPALRKVPENMLYLFTKFCVDADDDDKLYNKT LQCRELLVKNTDLPFIGDISDVKT DIFLRKD
INEETEVI PDNVSDQVILSKNTSEHGQLDLLYPSIDSESEILPGENQVFYDNRTQNV DYLN SYYLESQ
KLSDNVEDFTFTRSIEEALDNSAKVYTYFPTLANKVNAGVQGLFLMWANDVVEDFTTNILRKDTLDKIS
DVS AIIPIYIGPALNISNSVRRGNFTEAFAVTGV TILLEAFPEFTIPALGAFVIYSKVQERNEI IKTIDNC
LEQRIKRWKDSYEWMMGTWLSRIITQFN NISYQMYDSLNYQAGAIKAKIDLEYKKYSGSDKENIKS QVEN
LKNSLDVKISEAMNNINKFIRECSVTYLFKNMLPKVIDELNEFDRNTKAKLINLIDSHNIILVGEVDKLK
AKVNNSFQNTIPFNIFSYTNN SLLKDIINEYFNLEGGGSGGGSGGGGSALDNSDSECPLSHDGYCLHD
5 GVCMYIEALDKYACNCVVG YIGERCQYRDLKWWELR

SEQ ID NO: 7

atgggatccatggagttcgttaaaaaacagttcaactataaagacccagttaacggtggtgacattgctt
acatcaaaatcccgaacgctggccagatgcagccggtaaaggcattcaaaatccacaacaaaatctgggt
tatcccggaacgtgatacctttactaaccggaagaagggtgacctgaaccgccaccggaagcgaaacag
gtgccggtatcttactatgactccacctacctgtctaccgataacgaaaaggacaactacctgaaagggtg
ttactaaactgttcgagcgtatttactccaccgacctgggccgtatgctgctgactagcatcgttcgcg
tatcccggttctggggcggttctaccatcgataccgaactgaaagtaatcgacactaactgcatcaacggt

attcagccggacgggttcctatcggtccgaagaactgaacctgggtgatcatcggcccgctctgctgatatca
tccagttcagagtgaagagctttgggtcacgaagttctgaacctcaccgtaacggctacggttccactca
gtacatccgtttctctccggacttcaccttcggttttgaagaatccctggaagtagacacgaacccactg
ctgggcgctggtaaattcgcaactgatcctgcggttacctgggtcacgaactgattcatgcaggccacc
gcctgtacgggtatcgccatcaatccgaaccgtgtcttcaaagttaacaccaacgcgtattacgagatgtc
cgggtctggaagttagcttcgaagaactgcgtacttttggcggtcacgacgctaaattcatcgactctctg
caagaaaacgagttccgtctgtactactataacaagttcaaagatatcgcatccaccctgaacaaagcga
aatccatcggtgggtaccactgcttctctccagtacatgaagaacggtttttaagaaaaatacctgctcag
cgaagacacctccggcaaattctctgtagacaagttgaaattcgataaactttacaaaatgctgactgaa
atttacaccgaagacaacttcggttaagttctttaagttctgaaccgcaaacctatctgaacttcgaca
aggcagatttcaaaatcaacatcggtccgaaagttaactacactatctacgatggtttcaacctgcgtaa
caccaacctgggtgctaattttaacggccagaacacggaaatcaacaacatgaacttcacaaaactgaaa
aacttcactgggtctgttcgagttttacaagctgctgtgcGTGACGGCATCATTACCTCCAAAACATAAT
CTGACGATGACGATAAAAACAAAGCGCTGAACCTGCAGtgatcaaggttaacaactgggatttattctt
cagcccgagtgaagacaacttcaccaacgacctgaacaaaggtgaagaaatcacctcagataactaacatc
gaagcagccgaagaaaacatctcgctggacctgatccagcagtaactacctgacctttaatttcgacaacg
agccggaaaacatttctatcgaaaacctgagctctgatcatcgggccagctggaactgatgccgaacat
cgaacgtttcccaaacggtaaaaagtacgagctggacaaatataccatgttccactacctgcgcgcgag
gaatttgaacacggcaaatcccgatatcgactgactaactccgttaacgaagctctgctcaacctgccc
gtgtatacaccttcttctctagcgactacgtgaaaaaggtcaacaaagcgactgaagctgcaatgttctt
gggttgggttgaacagcttggttatgattttaccgacgagacgtccgaagtatctactaccgacaaaatt
gcggatatcactatcatcatcccgatcatcggtccggctctgaacattggcaacatgctgtacaaagacg
acttcggttggcgactgatcttctccgggtgcggtgatcctgctggagttcatcccggaatcgccatccc
ggactgggcacctttgctctggtttcttacattgcaacaaggttctgactgtacaaacctcgacaac
gcgctgagcaaacgtaacgaaaaatgggatgaagtttacaatatatcgtagaccaactggctggctaagg
ttaatactcagatcgacctcatccgcaaaaaaatgaaagaagcactggaaaaccaggcggaagctaccaa
ggcaatcattaactaccagtacaaccagtacaccgaggaagaaaaaacaacatcaacttcaacatcgac
gatctgtcctctaaactgaacgaatccatcaacaaagctatgatcaacatcaacaagtccctgaaccagt
gctctgtaagctatctgatgaactccatgatcccgtagggtgttaaactgctggaggacttcgatgcgtc
tctgaaagacgccctgctgaaatacatttacgacaaccgtggcactctgatcggtcaggttgatcgtctg
aaggacaaagtgaacaataccttatcgaccgacatcccttttcagctcagtaaatatgtcgataaccaac
gccttttgtccactctagaaggtggcggtgggtccgggtggcggtgggtcagggcgggggcggttagcgact
agacaactctgactctgaatgcccgctgtctcacgacggttactgctgcacgacggtgtttgcatgtac
atcgaagctctggacaaatagcttgcaactgcgttggttggttacatcggtgaacggtgccagtagcgtg
acctgaaatggtgggaactgcgtgcgctagaagcaCACCATCATCACcaccatcaccatcaccattaatg
a

SEQ ID NO: 8

MGSMEFVNKQFNYKDPVNGVDIAYIKIPNAGQMOPVKAFKIHNKIWWIPERDTFTNPEEGDLNPPPEAKQ
VPVSYDSTYLSTDNEKDNYLKGVTKLFERIYSTDLGRMLLTSIVRGIPFWGGSTIDTELKVIDTNCINV
IQPDGSYRSEELNLVIIGPSADIIQFECKSFGHEVLNLTRNGYGSTQYIRFSPDFTFGFEESLEVDTNPL
LGAGKFATDPAVTLAHELIIHAGHRLYGIAINPNRVFKVNTNAYYEMSGLEVSFEELRTFGGHDAKFIDSL
QENEFRLYYYNKFKDIASTLNKAKSIVGTTASLQYMKNVFKEKYLLSEDTSGKFVSVDKLFKDKLYKMLTE
IYTEDNFVKFFKVLNRKTYLNFDKAVFKINIVPKVNYTIYDGFNLRNTNLAANFNGQNTTEINNMNFTKLK

NFTGLFEFYKLLCVDGIIITSKTSDDDDKNKALNLQCIKVNNDLFFSPSEDNFTNDLNKGEEITSDTNI
EAAEENISLDLIQQYYLTFNFDNEPENISIENLSSDIIGQLELMPNIEFPPNGKKYELDKYTMFHYLRAQ
EFEHGKSRIALTNSVNEALLNPSRVYTFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKI
ADITIIIPYIGPALNIGNMLYKDDFVGALIFSGAVILLEFIPEIAIPVLGTFALVSYIANKVLTQVTDN
ALSKRNEKWDEVYKYIVTNWLAKVNTQIDLIRKKMKEALENQAETKAIINYQYNQYTEEEKNNINFNID
DLSSKLNESINKAMININKFLNQCSVSYLMNSMIPYGVKRLEDFDASLKDALLKYIYDNRGTLIGQVDRL
KDKVNNTLSTDIPFQLSKYVDNQRLSTLEGGGGSGGGGSGGGGSALDNSDSECPLSHDGYCLHDGVCMY

5 IEALDKYACNCVVGYYIGERCQYRDLKWWELRALEAHHHHHHHHHH

SEQ ID NO: 9

atgggatccatggagttcgttaacaaacagttcaactataaagacccagttaacggtgttgacattgctt
 acatcaaaatcccgaacgctggccagatgcagccggtaaaggcattcaaaatccacaacaaaatctgggt
 tatcccggaaacgtgatacctttactaaccgggaagaaggtgacctgaacccgccaccggaagcgaaacag
 gtgccggtatcttactatgactccacctacctgtctaccgataacgaaaaggacaactacctgaaaggtg
 ttactaaactgttcgagcgatatttactccaccgacctgggcccgtggtctgctgactagcatcggttcgagg
 tatcccgttctggggcggttctaccatcgataccgaactgaaagtaatcgacactaactgcatcaacggtt
 attcagccgggacggttcctatcggtccgaagaactgaacctgggtgatcatcgccccgtctgctgatatca
 tccagttcgagtgaagagctttggtcacgaagttctgaacctcaccgtaacgggtacggttccactca
 gtacatccgtttctctccggacttcaccttcggttttgaagaatccctggaagtagacacgaacccactg
 ctgggcgctggtaaattcgcaactgatcctgcggttaccctgggtcacgaactgattcatgcaggccacc
 gcctgtacgggtatcgccatcaatccgaaccgtgtcttcaaagttaacaccaacgcgtattacgagatgtc
 cggctctggaagtagcttcgaagaactgcgtacttttggcggtcacgacgctaaattcatcgactctctg
 caagaaaacgagttccgtctgtactactataacaagttcaaagatatcgcatccaccctgaacaaagcga
 aatccatcggtgggtaccactgcttctctccagtacatgaagaacggtttttaagaaaaatacctgctcag
 cgaagacacctccggcaaatctctctgtagacaagttgaaattcgataaaactttacaaaatgctgactgaa
 atttacaccgaagacaacttcggttaagttctttaagttctgaaccgcaaaacctatctgaacttcgaca
 aggcagatttcaaaatcaacatcggtgccgaaagttaactacactatctacgatgggttcaacctgcgtaa
 caccaacctggctgctaattttaacggccagaacacggaaatcaacaacatgaacttcacaaaactgaaa
 aacttcactgggtctgttcgagttttacaagctgctgtgcGTGACGGCATCATTACCTCCAAAATAAAT
 CTGACGATGACGATAAAAACAAAGCGCTGAACCTGCAGTgtatcaaggttaacaactgggatttattctt
 cagcccagtgagacaacttcaccaacgacctgaacaaaggtgaagaaatcacctcagatactaactc
 gaagcagccgaagaaaacatctcgctggacctgatccagcagtagctacctgacctttaatttcgacaacg
 agccggaaaacatttctatcgaaaacctgagctctgatcatcggccagctggaactgatgccgaacat
 cgaacggttcccaaacggtaaaaagtacgagctggacaaatataccatgttccactacctgcgcgcgcgag
 gaatttgaacacggcaaatcccgatcgcaactgactaactccgttaacgaagctctgctcaacctgccc
 gtgtatacaccttcttctctagcgactacgtgaaaaaggtcaacaaagcgactgaagctgcaatgttctt
 gggttgggttgaacagcttggttatgattttaccgacgagacgtccgaagtatctactaccgacaaaatt
 gcggatatcactatcatcatcccgtagatcggtccggctctgaacattggcaacatgctgtacaaagacg
 acttcggtggcgactgatcttctccgggtgcggtgatcctgctggagttcatcccggaatcgccatccc
 ggtactgggcacctttgctctggtttcttacattgcaaacaaggttctgactgtacaaacctcgacaac
 gcgctgagcaaacgtaacgaaaaatgggatgaagtttacaaatatatcgtagaccaactggctggctaagg
 ttaatactcagatcgacctcatccgcaaaaaaatgaaagaagcactggaaaaccaggcggaagctaccaa
 ggcaatcattaactaccagtacaaccagtacaccgaggaagaaaaaaacaacatcaacttcaacatcgac
 gatctgtcctctaaactgaacgaatccatcaacaaagctatgatcaacatcaacaagttcctgaaccagt

gctctgtaagctatctgatgaactccatgatcccgtagcgtgttaaactctggaggacttcgatgcgtc
tctgaaagacgccctgctgaaatacattttacgacaaccgtggcactctgatcggtcaggttgatcgtctg
aaggacaaagtgaacaataccttatcgaccgacatcccttttcagctcagtaaataatgtcgataaccaac
gccttttgtccactctagaaggtggcggtgggtccggtggcggtggctcagggcgggggcggttagcgact
agacaactctgactctgaatgcccgctgtctcacgacggttactgctgcacgacggtgtttgcatgtac
atcgaagctctggacaaatacgttgcaactgcgttgttggttacatcggtgaacgttgccagtagcgtg
acctgaaatggtgggaactgcgtgcgctagaagcaCACCATCATCACcaccatcaccatcaccattaatg
a

SEQ ID NO: 10

MGSMEFVNKQFNYKDPVNGVDIAYIKIPNAGQMOPVKAFFKIHNKIWVIPERDTFTNPEEGDLNPPPEAKQ
VPVSYDSTYLSTDNEKDNLYLKGVTKLFERIYSTDLGRGLLTSIVRGIPFWGGSTIDTELKVIDTNCINV
IQPDGSYRSEELNLVIIIGPSADIIQFECKSFGEVLNLTRNGYGSTQYIRFSPDFTFGFEESLEVDTNPL
LGAGKFATDPAVTLAHELIHAGHRLYGIAINPNRVFKVNTNAYYEMSGLEVSFEELRTFGGHDAKFIDSL
QENEFRLYYNKFKDIASTLNKAKSIVGTASLQYMKNVFKEKYLSEDTSGKFSVDKLFKLYKMLTE
IYTEDNFVKFFKVLNRKTYLNFDAVKINIVPKVNYTIYDGFNLNNTNLAANFNGQNTNINNMNFTKLK
NFTGLFEFYKLLCVDGIITSKTKSDDDDKNKALNLQCIKVNNDLFFSPSEDNFTNDLNKGEEITSDTNI
EAAEENISLDLIQQYYLTFNFDNEPENISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQ
EFEHGKSRIALTNSVNEALLNPSRVYTFSSDYVKKVNKATEAMFLGWVEQLVYDFTDETSEVSTTDKI
ADITIIIPYIGPALNIGNMLYKDDFVGALIFSGAVILLEFIPEIAIPVLGTFALVSYIANKVLTQVTDN
ALSKRNEKWDEVYKIVTNWLAKVNTQIDLRKKMKEALENQAETKAIINYQYNQYTEEEKNNINFNID
DLSSKLNESINKAMININKFLNQCSVSYLMNSMIPYGVKRLEDFDASLKDALLKYIYDNRGTIGQVDR
KDKVNNTLSTDIPFQLSKYVDNQRLSTLEGGGGSGGGSGGGGSALDSDSECPLSHDGYCLHDGVC
5 IEALDKYACNCVVGYYIGERCQYRDLKWWELEAHHHHHHHHHH

SEQ ID NO: 11

ISEFGSEFMPITINNFNYSFVNDKNILYLDTHLNTLANEPEKAFRITGNIWVIPDRFSRNSNPNLNKP
RVTSPKSGYYDPNYLSTDSKDTFLKEIIFLKRINSREIGEELIYRLSTDIPFPGNNNTPIINTFDFD
FNSVDVKTRQGNNWVKTGSINPSVIIITGPRENIDPETSTFKLTNNTFAAQEGFGALSIIISISPRF
SNATNDVRSRRFSKSEFCMDPILILMHELNHAMHNLGIAIPNDQTISSVTSNIFYSQYNVKLEYAEIYA
FGGPTIDLIPKSARKYFEEKALDYRSIAKRLNSITTANPSSFNKYIGEYKQKLIRKYRFVVESSGEVTV
NRNKFVELYNELTQIFTEFNKYIYNVQNRKIYLSNVYTPVTANILDDNVYDIQNGFNIKSNLNLVFMG
QNLSRNPALRKVPENMLYLFTKFCVDADDDDKLYNKTLCRELLVKNTDLFFIGDISDVKTDFLRKDI
NEETEVIPOVNSVDQVILSKNTSEHGQDLLYPSIDSESEILPGENQVFDNRTQNVYDLYNSYYLESQK
LSDNVEDFTFTRSIEEALDNSAKVYTYFPTLANKVNAGVQGGFLMWANDVVEDFTTNILRKDTLDKISD
VSAIIPYIGPALNISNSVRRGNFTEAFVGTIVTILLEAFPEFTIPALGAFVIYSKVQERNEIKTIDNCL
EQRIKRWKDSYEWMMGTWLSRIITQFNNSYQMYDSLNYQAGAIKAKIDLEYKKYSGSDKENIKSQVENL
KNSLDVKISEAMNNINKFIRECSVTYLFKNMLPKVIDELNEFDRNTKAKLINLIDSHNIIILVGEVDK
KVNSNFQNTIPFNISYTNNSLLKDIINEYFNLEGGGGSGGGSGGGGSALDSDSECPLSHDGYCLHDG
VCMYIEALDKYACNCVVGYYIGERCQYRDLKWWELEAHHHHHHHHHH

10

ES 2 537 338 T3

SEQ ID NO: 12

atgggatccatggagttcgttaacaaacagttcaactataaagacccagttaacggtggtgacattgctt
acatcaaaatcccgaacgctggccagatgcagccggtaaaggcattcaaaatccacaacaaaatctgggt

5

tatcccggaacgtgatacctttactaaccggaagaaggtgacctgaacccgccaccggaagcgaaacag
 gtgccggtatcttactatgactccacctacctgtctaccgataacgaaaaggacaactacctgaaaggtg
 ttactaaactgttcgagcgtatcttactccaccgacctgggccgtatgctgctgactagcatcggttcgcgg
 tatcccggttctggggcggttctaccatcgataccgaactgaaagtaatcgacactaactgcatcaacggt
 attcagccggacggttcctatcggtccgaagaactgaacctggatgatcatcgggccgtctgctgatatca
 tccagttcgagtgtgaagagctttgggtcacgaagttctgaacctcaccgtaacggctacggttccactca
 gtacatccggtttctctccggacttcaccttcggttttgaagaatccctggaagtagacacgaacccactg
 ctggggcgctggtaaattcgcaactgatcctgcggttacctgggtcacgaactgattcatgcaggccacc
 gcctgtacgggtatcgccatcaatccgaaccgtgtcttcaaagttaacaccaacgcgtattacgagatgtc
 cgggtctggaagttagcttcgaagaactgcgtacttttggcggtcacgacgctaaattcatcgactctctg
 caagaaaacgagttccgtctgtactactataacaagttcaaagatatcgcatccacctgaacaaagcga
 aatccatcggtgggtaccactgcttctctccagtacatgaagaacggtttttaagaaaaataacctgctcag
 cgaagacacctccggcaaatctctgttagacaagttgaaattcgataaactttacaaaatgctgactgaa
 atttacaccgaagacaacttcggttaagttctttaagttctgaaccgcaaacctatctgaacttcgaca
 aggcagtattcaaaaatcaacatcggtgccgaaagttaactacactatctacgatgggtttcaacctgcgtaa
 caccaacctgggtgctaattttaacggccagaacacggaaatcaacaacatgaacttcacaaaactgaaa
 aacttcactgggtctgttcgagttttacaagctgctgtgcGTCGACGGCATCATTACCTCCAAAATAAAT
 CTGACGATGACGATAAAAAACAAAGCGCTGAACCTGCAGtgatcaaggttaacaactgggattttattctt
 cagcccagtggaagacaacttcaccaacgacctgaacaaaggtgaagaaatcacctcagataactaacatc
 gaagcagccgaagaaaacatctcgctggacctgatccagcagtaactacctgacctttaatttcgacaacg
 agccggaaaacattttctatcgaaaacctgagctctgatcatcggccagctggaactgatgccgaacat
 cgaacggtttcccaaacggtaaaaagtagagctggacaaatataccatgttccactacctgcgcgcgcgag
 gaatttgaacacATcGaaGGccgtatcgactgactaactccgttaacgaagctctgctcaacctgccc
 gtgtatacaccttcttctctagcgactacgtgaaaaaggtcaacaaagcgactgaagctgcaatgttctt
 ggggttgggttgaacagcttggttatgattttaccgacgagacgtccgaagtatctactaccgacaaaatt
 gcggatatcactatcatcatcccgtagatcggtccggctctgaacattggcaacatgctgtacaaagacg
 acttcggttggcgactgatcttctccggtgcggtgatcctgctggagttcatcccggaaatcgccatccc
 ggtactgggcacctttgctctgggtttcttacattgcaaaacaaggttctgactgtacaaacctcgacaac
 gcgctgagcaaacgtaacgaaaaatgggatgaagttacaaatatatcgtagaccaactggctggctaagg
 ttaatactcagatcgacctcatccgcaaaaaaatgaaagaagcactggaaaaccaggcggaagctacca
 ggcaatcattaactaccagtacaaccagtacaccgaggaagaaaaaacaacatcaacttcaacatcgac
 gatctgtcctctaaactgaacgaatccatcaacaaagctatgatcaacatcaacaagttcctgaaccagt
 gctctgtaagctatctgatgaactccatgatcccgtagcgtgttaaactgctggaggacttcgatgcgtc
 tctgaaagacgccctgctgaaatacatttacgacaaccgtggcactctgatcggtcaggttgatcgtctg
 aaggacaaagtgaacaataccttatcgaccgacatcccttttcagctcagtaaataatgtcgataaccaac
 gccttttgtccactctagaaggtggcggtgggtccggtggcggtggctcaggcgggggcggttagcgact
 agacaactctgactctgaatgcccgctgtctcacgacgggttactgcctgcacgacgggtgtttgcatgtac
 atcgaagctctggacaaatacgttgcgaactgcgttgttgggttacatcggtgaacggtgccagtagcgtg
 acctgaaatggtgggaactgcgtgcgctagaagcaCACCATCATCACcaccatcaccatcaccattaatg
 a

SEQ ID NO: 13

GSMEFVNKQFNYKDPVNGVDIAYIKIPNAGQMOPVKAFKIHNKIWVIPERDTFTNP EEGDLNPPPEAKQV
 PVSYYDSTYLSTDNEKDNYLKGVTKLFERIYSTDLGRMLLTSIVRGIPFWGGSTIDTELKVIDTNCINVI
 QPDGSYRSEELNLVIIGPSADIIQFECKSFGHEVLNLTRNGYGSTQYIRFSPDFTFGFEESLEVDTNPLL
 GAGKFATDPAVTLAHELIIHAGHRLYGIAINPNRVFKVNTNAYYEMSGLEVSFEELRTFGGHDAKFIDSLQ
 ENEFRLYYYNKFKDIASTLNKAKSIVGTTASLQYMKNVFKEKYLLSEDTSGKFSVDKLFKDKLYKMLTEI
 YTEDNEFVKFFKVLNRKTYLNFDKAVFKINIVPKVNYTIYDGFNLRNTNLAANFNGQNT EINNMF TKLKN
 FTGLFEFYKLLCVDGIITSKTSDDDDKNKALNLQCIKVNNDLFFSPSEDNFTNDLNKGEEITS DTNIE
 AAENISLDLIQQYYLT FNFDPENIS IENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQE
 FEHIEGRIALTNSVNEALLNPSRVYTFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIA
 DITIIIPYIGPALNIGNMLYKDDFVGALIFSGAVILLEFIPEIAIPVLGTFALVSYIANKVLT VQTIDNA
 LSKRNEKWDEVYKYIVTNWLAKVNTQIDLRKKMKEALENQAATKAIINYQYNQYTEEEKNNINFNIDD
 LSSKLNESINKAMININKFLNQCSVSYLMNSMIPYGVKRLED FDASLKDALLKYIYDNRGTLIGQVDRLK
 DKVNNTLSTDIPFQLSKYVDNQRLSTLEGGGGSGGGGSGGGGSALDNSDSECPLSHDGYCLHDGVCMYI
 5 EALDKYACNCVVG YIGERCQYRDLKWWELRALEAHHHHHHHHHH

SEQ ID NO: 14

atgggatccatggagttcggttaacaaacagttcaactataaagacccagttaacggtgttgacattgctt
 acatcaaaatcccgaacgctggccagatgcagccggtaaaggcattcaaaatccacaacaaaatctgggt
 tatcccgaacgtgatacctttactaaccgggaagaagggtgacctgaacccgccaccggaagcgaaacag
 gtgccggtatcttactatgactccacctacctgtctaccgataacgaaaaggacaactacctgaaagggtg
 ttactaaactgttcgagcgtatcttactccaccgacctgggccgtatgctgctgactagcatcgcttcgcgg
 tatcccgttctggggcggttctaccatcgataccgaactgaaagtaatcgacactaactgcatcaacggtt
 attcagccgggacggttcctatcgttccgaagaactgaacctgggtgatcatcggcccgctctgctgatatca
 tccagttcgagtgaagagctttgggtcacgaagttctgaacctcacccgtaacgggtacggttccactca
 gtacatccggtttctctccggacttcaccttcggttttgaagaatccctggaagtagacacgaacccactg
 GCGCAGgctGTTTCGTTCTTCTTgatcctgcggttaccctgggtcacgaactgattcatgcaggccacc
 gcctgtacggtatcgccatcaatccgaaccgtgtcttcaaagttaacaccaacgcgtattacgagatgtc
 cgggtctggaagttagcttcgaagaactgcgtacttttggcggtcacgacgctaaattcatcgactctctg
 caagaaaacgagttccgtctgtactactataacaagttaaaagatatcgcatccaccctgaacaaagcga
 aatccatcggtgggtaccactgcttctctccagtacatgaagaacggtttttaaagaaaaatacctgctcag
 cgaagacacctccggcaaattctctgtagacaagttgaaattcgataaactttacaaaatgctgactgaa
 atttacaccgaagacaacttcggttaagttctttaaagttctgaaccgcaaaaccttatctgaacttcgaca
 aggcagatttcaaaatcaacatcggtgccgaaagttaactacactatctacgatgggtttcaacctgcgtaa
 caccaacctggctgctaattttaacggccagaacacggaaatcaacaacatgaacttcacaaaactgaaa
 aacttcactgggtctgttcgagttttacaagctgctgtgcGTCGACGGCATCATTACCTCCAAAACATAAT
 CTGACGATGACGATAAAAACAAAGCGCTGAACCTGCAGtgatcaagggttaacaactgggattttattctt
 cagcccagagtgaagacaacttcaccaacgacctgaacaaagggtgaagaaatcacctcagataactaacatc
 gaagcagccgaagaaaacatctcgctggacctgatccagcagtaactacctgacctttaatttcgacaacg
 agccggaaaacatttctatcgaaaacctgagctctgatcatcggccagctggaactgatgccgaacat
 cgaacggtttcccaaacggtaaaaagtacgagctggacaaatataccatgttccactacctgcgcgcgag
 gaatttgaacacggcaaatcccgatcgccactgactaactccgttaacgaagctctgctcaaccgctccc
 gtgtatacaccttcttcttagcgactacgtgaaaaagggtcaacaaagcgactgaagctgcaatgttctt
 ggggtgggttgaacagcttggttatgattttaccgacgagacgtccgaagtatctactaccgacaaaatt
 gcggatatcactatcatcatcccgctacatcggtccgggtctgaacattggcaacatgctgtacaaagacg
 acttcggtggcgactgatcttctccggtgcggtgatcctgctggagttcatcccggaatcgccatccc
 ggtactgggcacctttgctctgggttcttacattgcaacaagggttctgactgtacaaaccatcgacaac

gcgctgagcaaacgtaacgaaaaatgggatgaagtttacaaatatatcgtgaccaactggctggctaagg
ttaatactcagatcgacctcatccgcaaaaaaatgaaagaagcactggaaaaccaggcggaagctaccaa
ggcaatcattaactaccagtacaaccagtagacccgaggaagaaaaaacaacatcaacttcaacatcgac
gatctgtcctctaaactgaacgaatccatcaacaaagctatgatcaacatcaacaagttcctgaaccagt
gctctgtaagctatctgatgaactccatgatcccgtagcgtgttaaacgtctggaggacttcgatgcgtc
tctgaaagacgccctgctgaaatacatttacgacaaccgtggcactctgatcggtcaggttgatcgtctg
aaggacaaagtgaacaataccttatcgaccgacatcccttttcagctcagtaaatatgtcgataaccaac
gccttttgtccactctagaaggtggcgggtgggtccggtggcgggtggctcaggcggggcggttagcgcact
agacaactctgactctgaatgcccgctgtctcacgacgggttactgcctgcacgacgggtgtttgcatgtac
atcgaagctctggacaaatacgttgcaactgcgttggttggttacatcggtgaacggtgccagtagcgtg
acctgaaatggtgggaactgcgtgcgctagaagcaCACCATCATCACcaccatcaccatcaccattaatg
a

SEQ ID NO: 15

5

GSMEFVNKQFNYKDPVNGVDIAYIKIPNAGQMOPVKAFKIHNKIWVIPERDTFTNPEEGDLNPPPEAKQV
PVSYYDSTYLSTDNEKDNLYLKGVTCLFERIYSTDLGRMLLTSIVRGIPFWGGSTIDTELKVIDTNCINVI
QPDGSYRSEELNLVIGPSADIIQFECKSFGHEVLNLTRNGYGSTQYIRFSPDFTFGFEESLEVDTNPLA
QAVRSSSDPAVTLAHELIIHAGHRLYGIAINPNRVFKVNTNAYYEMSGLEVSFEELRTFGGHDAKFIDSLQ
ENEFRLYYYNKFKDIASTLNKAKSIVGTTASLQYMKNVFKEKYLLSEDTSGKFSVDKLKFDKLYKMLTEI
YTEDNFVKFFKVLNRKTYLNFDAVKINIVPKVNYTIYDGFNLRTNLAANFNGQNTNINNMNFTKLKN
FTGLFEFYKLLCVDGIIITSKTSDDDDKNKALNLQCIKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIE
AAEENISLDLIQYYLTFFNFDNEPENISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQE
FEHGKSRIALTNSVNEALLNPSRVYTFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIA
DITIIIPYIGPALNIGNMLYKDDFVGALIFSGAVILLEFIPEIAIPVLGTFALVSYIANKVLTQVIDNA
LSKRNEKWDEVYKYIVTNWLAKVNTQIDLIRKKMKEALENQAEATKAIINYQYNQYTEEEKNNINFNIDD
LSSKLNESINKAMININKFLNQCSVSYLMNSMIPYGVKRLEDFDASLKDALLKYIYDNRGTLIGQVDRLK
DKVNNTLSTDIPFQLSKYVDNQRLSTLEGGGGSGGGGSGGGGSALDNSDSECPLSHDGYCLHDGVCMYI
EALDKYACNCVVGYIGERCQYRDLKWWELRALEAHHHHHHHHHH

SEQ ID NO: 16

atgggatccatggagttcgtaacaaacagttcaactataaagacccagttaacgggttgacattgcttacatcaaaatcccgaacgc
 tggccagatgcagccggtaaaggcattcaaaatccacaacaaaatctgggtatcccggaacgtgataccttactaaccggaag
 aaggtagctgaacccgccaccggaagcgaaacagggtccggtatcttactatgactccacctacctgtctaccgataacgaaaag
 gacaactacctgaaagggttactaaactgttcgagcgtatttactccaccgacctgggccgtatgctgctgactagcatcgttcgcggt
 atcccgttctggggcggttctaccatcgataccgaactgaaagtaatcgacactaactgcatcaacgttattcagccggacgggtcctat
 cgttccgaagaactgaacctgggtatcatcggtccgtctgctgatalcatccagttcgagtgaagagctttggtcacgaagttctgaac
 ctacccgtaaacggctacgggtccactcagtagatccgtttctctccggacttcacctcgggtttgaagaatccctggaagtagacacga
 acccactgctgggcgctggtaaattcgcaactgatcctgcggttaccctgggtcacgaactgattcatgcaggccaccgcctgtacgggt
 atcgccatcaatccgaaccgtgtcttcaaagtaacaccaacgcgtattacgagatgtccggtctggaagttagcttgaagaactgcg
 tacttttggcggtcacgacgctaaattcatcgactctctgcaagaaaacgagttccgtctgtactactataacaagttcaaagatatcg
 atccacccctgaacaaagcgaaatccatcggtgggtaccactgcttctccagtacatgaagaacgttttaagaaaaataacctgtca

gcgaagacacctccggcaaattctctgtagacaagtgaaattcgataaactttacaaaatgctgactgaaatttacaccgaagacaa
 cttcgttaagttcttaagttctgaaccgcaaacctatctgaacttcgacaaggcagattcaaatcaacatcgtgccgaagttaa
 ctacactatctacgatggttcaacctgcgtaaccaacctggctgctaatttaacggccagaacacggaaatcaacaacatgaac
 ttcacaaaactgaaaaacttcactggctgttcgagttttacaagctgctgtgcGTCGACGGCATCATTACCTCCAAAA
 CTAATCTGACGATGACGATAAAAAACAAAGCGCTGAACCTGCAGtgatcaaggtaacaactgggattt
 attctcagcccgagtgaaagacaacttcaccaacgacctgaacaaagggaagaaatcacctcagatactaacatcgaagcagcc
 gaagaaaacatctcgtggacctgatccagcagctactacctgacctttaatttcgacaacgagccggaaaacatttctatcgaaaacc
 tgagctctgatatcatcgccagctggaactgatgccgaacatcgaacgtttcccaaacggtaaaaagtagcagctggacaaatata
 ccatgttccactacctgcgcgcgaggaattgaacacggcaaatcccgatcgcactgactaactccgttaacgaagctctgtctcaa
 cccgtcccgtgtatacaccttctctctagcagctacgtgaaaaaggtaacaaagcgactgaagctgcaatgttctgggtgggtga
 acagctgtttatgattttaccgacgagacgtccgaagatctactaccgacaaaattgcgatctactatcatcatccgtacatcggt
 ccggctctgaacattggcaacatgctgtacaaagacgactcgttggcgactgatcttctccggtgcggtgatcctgctggagttcatc
 ccggaatcgccatcccggtactgggcaccttctgtctgttcttcatctgcaacaaggttctgactgtacaaaccatcgacaacgc
 gctgagcaaacgtaacgaaaaatgggatgaagttacaaatatctgtgaccaactggctggctaaggtaataactcagatcgacctc
 atccgcaaaaaaatgaagaagcactggaaaaccaggcggaagctaccaaggcaatcattaactaccagtacaaccagtacac
 cgaggaagaaaaaacaacatcaactcaacatcgacgatctgtccttaactgaacgaatccatcaacaagctatgatcaac
 atcaacaagttctgaaccagtgtctgtgaagctatctgatgaactccatgatcccgtaggtgttaaacgtctggaggactcgtatgcg
 tctctgaaagacgccctgtgaaatacatttacgacaaccgtggcactctgatcggtaggtgatcgtctgaaggacaaagtgaaca
 ataccttatcgaccgacatccctttcagctcagtaaatatgtcgataaccaacgcctttgtccactttcaccgaatacatcaaaaacat
 catcaacaccagtctagaaATCCTGAACCTGCGTTACGAATCTAACCACCTGATCGACCTGTCTCG
 TTACGCTTCTAAAATCAACATCGGTTCTAAAGTTAACTTCGACCCGATCGACAAAAACAG
 ATCCAGCTGTTCAACCTGGAATCTTCTAAAATCGAAGTTATCCTGAAAAACGCTATCGTTTA
 CAACTCTATGTACGAAAACCTTCTCTACCTCTTCTGGATTCTGATCCCGAAATACTTTAACT
 CTATCTCTCTGAACAACGAATACACCATCATCAACTGCATGGAAAACAACCTCTGGTTGGAA
 AGTTTCTCTGAACCTACGGTGAAATCATCTGGACCCTGCAAGACACCCAGGAAATCAAACAG
 CGTGTGTTTTCAAATACTCTCAGATGATCAACATCTCTGACTACATCAACCGTTGGATCTT
 CGTTACCATCACCAACAACCGTCTGAACAACCTCTAAAATCTACATCAACGGTCGTCTGATC
 GACCAGAAACCGATCTCTAACCTGGGTAACATCCACGCTTCTAACAACATCATGTTCAAAC
 TGGACGGTTGCCGTGACACCCACCGTTACATCTGGATCAAATACTTCAACCTGTTGACAA
 AGAACTGAACGAAAAAGAAATCAAAGACCTGTACGACAACAGTCTAACTCTGGTgcactagtg
 ATTTTGAAGGACTTTTGGGGCGACTATCTCCAGTACGACAAACCTTACTATATGCTGAATT
 TGTATGATCCCAACAAATATGTGGATGTGAATAACGTTGGTATTAGGGGTTACATGTATTT
 GAAGGGTCCAAGGGGGTCAGTCATGACAACCAATATCTACTTAAATTCCTCTCTTTACCGA
 GGGACAAAATTCATTATCAAAAAGTATGCTAGTGGAATAAAGATAATATAGTCAGAAACAA
 TGATCGCGTTTACATTAAACGTGGTAGTCAAAAATAAGGAGTATAGACTAGCTACGAATGCA
 TCGCAGGCGGGAGTGGAGAAGATACTGAGCGCACTAGAAATACCTGACGTAGGAACTTA
 AGCCAGGTTGTCGTTATGAAATCAAAGAACGATCAAGGAATTACTAATAAGTGTAAGATGA

ACTTACAAGATAACAATGGCAATGATATAGGCTTCATCGGGTTTCATCAATTTAACAACATA
GCGAAACTCGTAGCCTCTAACTGGTACAACCGTCAAATCGAACGAAGTTCCCGTACTCTA
GGTTGCTCGTGGGAGTTCATCCCAGTAGACGACGGGTGGGGCGAACGGCCGCTTgcgctag
caCACCATCATCACcaccatcaccatcaccattaatga

SEQ ID NO: 17

HMGSMEFVNKQFNYKDPVNGVDIAYIKIPNAGQMGPVKAFKIHNIWVIPERDTFTNPEEGDLN
PPPEAKQVPVSYDSTYLSTDNEKDNLYKGVTKLFERIYSTDLGRMLLSIVRGIPFWGGSTIDT
ELKVIDTNCINVIQPDGSYRSEELNLVIIGPSADIIQFECKSFGHEVLNLTRNGYGSTQYIRFSPDF
TFGFEESELEVDTNPLLGAAGKATDPAVTLAHELHAGHRLYGIAINPNRVFKVNTNAYYEMSGLE
VSFEELRTFGGHDAKFIDSLQENEFRLYYYNFKDIASTLNKAKSIVGTTASLQYMKNVFKEKYL
LSEDTSGKFSVDKLFKDKLYKMLTEIYTEDNFVKFFKVLNRKTYLNFDKAVFKINIVPKVNYTIYD
GFNLNRNTNLAANFNGQNTNINNMNFTKLKNFTGLFEFYKLLCVDGIITSKTKSDDDDKNKALNL
QCIKVNNDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQYYLTNFDNEPENISIE
NLSSDIIGQLELMPNIEFPNGKKYELDKYTMFHYLRAQEFHKGSRALTNSVNEALLNPSRVY
TFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIIIPYIGPALNIGNMLYK
DDFVGALIFSGAVILLEFIPEIAIPVLGTFALVSYIANKVLTQVTDNALSQRNEKWDEVYKYIVTN
WLAKVNTQIDLRKKMKEALENQAETKAIINYQYNQYTEEEKNINFNIDDLSSKLNESINKAMI
NINKFLNQCSVSYLMNSMIPYGVKRLEDFDASLKDALLKYIYDNRGTLIGQVDRDKDKVNNTLST
DIPFQLSKYVDNQRLSTFTEYIKNIINTSLEILNLRYESNHLIDLSRYASKINIGSKVNFDPIDKNQI
QLFNLESSKIEVILKNAIVYNSMYENFSTSFWIRIPKYFNSISLNNEYTIINCMENNSGWKVS LNY
GEIIWTLQDTQEIKQRVVFYKYSQMINISDYINRWIFVTITNNRLNNSKIYINGRLIDQKPISNLGNIH
ASNIMFKLDGCRDTHRYIWIKYFNLFDEKELNEKEIKDLYDNQSN SGALVILKDFWGDYLYQYDK
PYYMLNLYDPNKYVDVNNVGIRGYMYLKGPRGSVMTTNIYLNSSLYRGTKFIIKKYASGNKDNIV
RNNDRVYINVVKNKEYRLATNASQAGVEKILSALEIPDVGNLSQVVVMKSKNDQGITNKCKMN
LQDNNGNDIGFIGFHQFNNAIKLVASNWYNRQIERSSRTLGCSEWEIFVDDGWGERPLALAHH
HHHHHHHH

5

SEQ ID NO: 18

atgggatccatggagttcgttaaaaaacagttcaactataaagacccagttaacggtgttgacattgctt
 acatcaaaatcccgaacgctggccagatgcagccggtaaaggcattcaaaatccacaacaaaatctgggt
 tatcccgggaacgtgatacctttactaaccgggaagaagggtgacctgaacccgccaccggaagcgaaacag
 gtgccggtatcttactatgactccacctacctgtctaccgataacgaaaaggacaactacctgaaagggtg
 ttactaaactgttcgagcgtatttactccaccgacctgggccgtatgctgctgactagcatcgttcgcg
 tatcccgttctggggcggttctaccatcgataccgaactgaaagtaatcgacactaactgcatcaacggt
 attcagccggacggttcctatcgttccgaagaactgaacctgggtgatcatcggcccgctctgctgatatca
 tccagttcgagtgaagagctttggtcacgaagttctgaacctcaccgtaacggctacggttccactca
 gtacatccgtttctctccggacttcaccttcggttttgaagaatccctggaagtagacacgaacccactg
 GCGCAGgctGTTCGTTTCCTCTTCTgatcctgcggttaccctggctcacgaactgattcatgcaggccacc

gcctgtacggtatcgccatcaatccgaaccgtgtcttcaaagttaacaccaacgcgtattacgagatgtc
 cggctctggaagtttagcttcgaagaactgcgtacttttgccggtcacgacgctaaattcatcgactctctg
 caagaaaaagagttccgtctgtactactataacaagttcaaagatatcgcatccaccctgaacaaagcga
 aatccatcgtgggtaccactgcttctctccagtacatgaagaacgtttttaagaaaaatacctgctcag
 cgaagacacctccggcaaattctctgtagacaagttgaaattcgataaaactttacaaaatgctgactgaa
 atttacaccgaagacaacttcgttaagttctttaagttctgaaccgcaaaacctatctgaacttcgaca
 aggcagtattcaaaatcaacatcgtgccgaaagttaactacactatctacgatggtttcaacctgcgtaa
 caccaacctggctgctaattttaacggccagaaacacggaaatcaacaacatgaacttcacaaaactgaaa
 aacttcactggctgttctcaggttttacaagctgctgtgcGTGACGGCATCATACCTCCAAAATAAAT
 CTGACGATGACGATAAAAAACAAAGCGCTGAACCTGCAGtgatcaaggttaacaactgggattttctt
 cagcccagagtgaagacaacttcaccaacgacctgaacaaaggtgaagaaatcacctcagataactaacatc
 gaagcagccgaagaaaaacatctcgtggacctgatccagcagtaactacctgacctttaatttcgacaacg
 agccggaaaaacatttctatcgaaaacctgagctctgatcatcgccagctggaactgatgccgaacat
 cgaacgtttcccaaacggtaaaaagtagagctggacaaatataccatgttccactacctgcgcgcgag
 gaatttgaacacggcaaattcccgatcgactgactaactccgttaacgaagctctgctcaacctgccc
 gtgtatacaccttcttctcagcgactacgtgaaaaaggtcaacaaagcgactgaagctgcaatgttctt
 ggggtgggttgaacagcttgtttatgattttaccgacgagacgtccgaagtatctactaccgacaaaatt
 gcggatatcactatcatcatcccgatcatcggtccggctctgaacattggcaacatgctgtacaagacg
 acttcgttggcgactgatcttctccggtgcggtgatcctgctggagttcatcccgaaaatcgccatccc
 ggtactgggcaccttctgctctggtttcttacattgcaacaaggttctgactgtacaaacctcgacaac
 gcgctgagcaaacgtaacgaaaaatgggatgaagtttacaatatatcgtgaccaactggctggctaagg
 ttaatactcagatcgacctcatccgcaaaaaatgaaagaagcactggaaaaccaggcggaagctaccaa
 ggcaatcattaactaccagtacaaccagtacaccgaggaagaaaaaacaacatcaacttcaacatcgac
 gatctgtcctctaaactgaacgaatccatcaacaaagctatgatcaacatcaacaagttcctgaaccagt
 gctctgtaagctatctgatgaactccatgatcccgtagggtgttaaacgtctggaggacttcgatgcgtc
 tctgaaagacgcccgtctgaaatacatttacgacaaccgtggcactctgatcggtcaggttgatcgtctg
 aaggacaaagtgaacaataccttatcgaccgacatcccttttcagctcagtaaatatgtcgataaccaac
 gccttttgtccactttcaccgaatacatcaaaaacatcatcaacaccagtctagaaATCCTGAACCTGCG
 TTACGAATCTAACCACTGATCGACCTGTCTCGTTACGCTTCTAAAAACAACATCGGTTCTAAAGTTAAC
 TTCGACCCGATCGACAAAAACAGATCCAGCTGTTCAACCTGGAATCTTCTAAAAATCGAAGTTATCCTGA
 AAAACGCTATCGTTTACAACCTCTATGTACGAAAACTTCTCTACCTCTTTCTGGATTTCGTATCCCGAAAATA
 CTTTAACTCTATCTCTCTGAACAACGAATACACCATCATCAACTGCATGGAAAACAACCTCTGGTTGGAAA
 GTTTCTCTGAACCTACGGTGAAATCATCTGGACCTGCAAGACACCCAGGAAATCAAACAGCGTGTGTTT
 TCAAATACTCTCAGATGATCAACATCTCTGACTACATCAACCGTTGGATCTTCGTTACCATCACCAACAA
 CCGTCTGAACAACCTCTAAAATCTACATCAACGGTCTGATCGACCAGAAACCGATCTCTAACCTGGGT
 AACATCCACGCTTCTAACAACATCATGTTCAAACCTGGACGGTTGCCGTGACACCCACCGTTACATCTGGA
 TCAAATACTTCAACCTGTTTCGACAAAGAACTGAACGAAAAAGAAATCAAAGACCTGTACGACAACCAAGTC
 TAACTCTGGTgcactagtgATTTTGAAGGACTTTTGGGGCGACTATCTCCAGTACGACAAACCTTACTAT
 ATGCTGAATTTGTATGATCCCAACAAATATGTGGATGTGAATAACGTTGGTATTAGGGGTTACATGTATT
 TGAAGGGTCCAAGGGGGTCAGTCATGACAACCAATATCTACTTAAATTCCTCTCTTTACCGAGGGACAAA
 ATTCATTATCAAAAAGTATGCTAGTGGAATAAAGATAATATAGTCAGAAACAATGATCGCGTTTACATT
 AACGTGGTAGTCAAAAATAAGGAGTATAGACTAGCTACGAATGCATCGCAGGCGGGAGTGGAGAAGATAC
 TGAGCGCACTAGAAATACCTGACGTAGGAACTTAAGCCAGGTTGTCGTTATGAAATCAAAGAACGATCA
 AGGAATTACTAATAAGTGTAAGATGAACCTACAAGATAACAATGGCAATGATATAGGCTTCATCGGGTTT

CATCAATTTAACAACATAGCGAAACTCGTAGCCTCTAACTGGTACAACCGTCAAATCGAACGAAGTTCCTCC
GTACTCTAGGTTGCTCGTGGGAGTTCATCCCAGTAGACGACGGGTGGGGCGAACGGCCGCTTgcgctagc
aCACCATCATCACcaccatcaccatcaccattaatga

SEQ ID NO: 19

GSMEFVNKQFNYKDPVNGVDIAYIKIPNAGQMOPVKAFKIHNKIWIPIPERDTFTNPEEGDLNPPPEAKQV
PVSYYDSTYLSTDNEKDNYLKGVTCLFERIYSTDLGRMLLTSIVRGIPFWGGSTIDTELKVIDTNCINVI
QPDGSYRSEELNLVIIGPSADIIQFECKSFGHEVLNLTRNGYGSTQYIRFSPDFTFGFEESLEVDTNPLA
QAVRSSSDPAVTLAHELHAGHRLYGIAINPNRVFKVNTNAYYEMSGLEVSFEELRTFGGHDAKFIDSLQ
ENEFRLYYYNKFKDIASTLNKAKSIVGTTASLQYMKNVFKEKYLLSEDTSGKFSVDKLFKDKLYKMLTEI
YTEDNFVKFFKVLNRKTYLNFDAVKINIVPKVNYTIYDGFNLRTNLAANFNGQNTTEINNMNFTKLKN
FTGLFEFYKLLCVDGIITSKTKSDDDDKNKALNLQCIKVNNDLFFSPSEDNFTNDLNKGEEITSDTNIE
AAEENISLDLIQQYYLTFNFDNEPENISIENLSSDIIGQLELMPNIEFPPNGKKYELDKYTMFHYLRAQE
FEHGKSRIALTNSVNEALLNPSRVYTFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIA
DITIIIPYIGPALNIGNMLYKDDFVGALIFSGAVILLEFIPEIAIPVLGTFALVSYIANKVLTVQTIIDNA
LSKRNEKWDEVYKYIVTNWLAKVNTQIDLIRKKMKEALENQAATKAIINYQYNQYTEEEKNINFNIDD
LSSKLNESINKAMININKFLNQCSVSYLMNSMIPYGVKRLDFDASLKDALLKYIYDNRGTLIGQVDRLK
DKVNNTLSTDIPFQLSKYVDNQRLSTFTTEYIKNIINTSLEILNLRYESNHLIDLSRYASKINIGSKVNF
DPIDKNQIQLFNLESSKIEVILKNAIVYNSMYENFSTSFWIRIPKYFNSISLNNEYTIINCMENNSGWKV
SLNYGEIIWTLQDTQEIKQRVVFKYSQMINISDYINRWIFVTITNNRLNNSKIYINGRLIDQKPISNLGN
IHASNNIMFKLDGCRDTHRYIWIKYFNLFDEKELNEKEIKDLYDNQSNAGALVILKDFWGDYLYDKPYYM
LNLYDPNKYVDVNNVGIRGYMYLKGPRGSVMTTNIYLNSSLYRGTKFIKKYASGNKDNIVRNNDRVYIN
VVVKNEKEYRLATNASQAGVEKILSALEIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFH
5 QFNNIAKLVASNWYNRQIERSSRTLGCSEWFI PVDDGWERPLALAHHHHHHHHHH

SEQ ID NO: 20

atgggatccatggagttcgttaacaaacagttcaactataaagacccagttaacggtgttgacattgctt
 acatcaaaatcccgaacgctggccagatgcagccggtaaaggcattcaaaatccacaacaaaatctgggt
 tatcccgggaacgtgatacctttactaaccgggaagaaggtgacctgaacccgccaccggaagcgaaacag
 gtgccggtatcttactatgactccacctacctgtctaccgataacgaaaaggacaactacctgaaaggtg
 ttactaaactgttcgagcgtatcttactccaccgacctgggccgtatgctgctgactagcatcgcttcgagg
 tatcccgttctggggcggttctaccatcgataccgaactgaaagtaatcgacactaactgcatcaacggtt
 attcagccggacggttcctatcggtccgaagaactgaacctgggtgatcatcggcccggtctgctgatatca
 tccagttcgagtgaagagctttgggtcacgaagttctgaacctcaccgtaacgggtacggttccactca
 gtacatccgtttctctccggaacttcaccttcggttttgaagaatccctggaagtagacacgaacccactg
 ctgggcgctggtaaattcgcaactgatcctgcggttaccctgggtcacgaactgattcatgcaggccacc
 gcctgtacggtatcgccatcaatccgaaccggtgtcttcaaagttaacaccaacgcgtattacgagatgtc
 cggctctggaagttagcttcgaagaactgcgtacttttggcggtcacgacgctaaattcatcgactctctg
 caagaaaacgagttccgtctgtactactataacaagttcaaagatatcgcatccacctgaacaaagcga
 aatccatcgtgggtaccactgcttctctccagtacatgaagaacgtttttaaagaaaaatacctgctcag
 cgaagacacctccggcaaatctctctgtagacaagttgaaattcgataaaactttacaaaatgctgactgaa
 atttacaccgaagacaacttcgttaagttctttaaagttctgaaccgcaaaacctatctgaacttcgaca
 aggcagtattcaaaatcaacatcgtgccgaaagttaactacactatctacgatgggtttcaacctgcgtaa
 caccaacctggctgctaattttaacggccagaacacggaaatcaacaacatgaacttcacaaaactgaaa

aacttcactggtctgttcgagttttacaagctgctgtgcGTCGACGGCATCATTACCTCCAAAACCTAAAT
CTGACGATGACGATAAAAAACAAAGCGCTGAACCTGCAGTgtatcaagggttaacaactgggattttattctt
cagcccgagtgaagacaacttcaccaacgacctgaacaaagggtgaagaaatcacctcagataactaacatc
gaagcagccgaagaaaacatctcgtgtggacctgatccagcagtactacctgacctttaatttcgacaacg
agccggaaaaacatttctatcgaaaacctgagctctgatcatcgccagctggaactgatgccgaacat
cgaacgtttcccaaacggtaaaaagtagcagctggacaaatataccatgttccactacctgcgcgcgcag
gaatttgaacacggcCGTtcccgTCGcatcgactgactaactccgttaacgaagctctgctcaaccctg
cccggtgtatacaccttcttctctagcgactacgtgaaaaagggtcaacaaagcgactgaagctgcaatgtt
cttgggttgggttgaacagcttgtttatgattttaccgacgagacgtccgaagtatctactaccgacaaa
attgcgatatcactatcatcatcccgtagatcggtccggctctgaacattggcaacatgctgtacaaag
acgacttcgttggcgactgatcttctccggtgcggtgatccgtgctggagttcatcccggaatcgccat
cccggtactgggcacctttgctctggtttcttacattgcaacaagggttctgactgtacaaaccatcgac
aacgcgctgagcaaacgtaacgaaaaatgggatgaagtttacaatatatcgtagaccaactggctggcta
agggttaatactcagatcgacctcatccgcaaaaaatgaaagaagcactggaaaaccaggcggaagctac
caaggcaatcattaactaccagtacaaccagtacaccgaggaagaaaaaacacatcaacttcaacatc
gacgatctgtcctctaaactgaacgaatccatcaacaaagctatgatcaacatcaacaagttcctgaacc
agtgtctgtgaagctatctgatgaactccatgatcccgtagcgtgttaaacgtctggaggacttcgatgc
gtctctgaaagacgcctgctgaaatacatttacgacaaccgtggcactctgatcggtcaggttgatcgt
ctgaaggacaaagtgaacaataccttatcgaccgacatcccttttcagctcagtaaatatgtcgataacc
aacgccttttgtccactttcacccgaatacatcaaaaacatcatcaacaccagtctagaaATCCTGAACCT
GCGTTACGAATCTAACACCTGATCGACCTGTCTCGTTACGCTTCTAAAAATCAACATCGGTTCTAAAGTT
AACTTCGACCCGATCGACAAAAACAGATCCAGCTGTTCAACCTGGAATCTTCTAAAAATCGAAGTTATCC
TGAAAAACGCTATCGTTTACAACTCTATGTACGAAAACCTTCTCTACCTCTTTCTGGATTCTGATCCCGAA
ATACTTTAACTCTATCTCTCTGAACAACGAATACACCATCATCAACTGCATGGAAAACAACTCTGGTTGG
AAAGTTTCTCTGAACCTACGGTGAAATCATCTGGACCCTGCAAGACACCCAGGAAATCAAACAGCGTGTTG
TTTTCAAATACTCTCAGATGATCAACATCTCTGACTACATCAACCGTTGGATCTTCGTTACCATCACCAA
CAACCGTCTGAACAACCTCTAAATCTACATCAACGGTCGTCTGATCGACCAGAAACCGATCTCTAACCTG
GGTAACATCCACGCTTCTAACAACATCATGTTCAAACCTGGACGGTTGCCGTGACACCCACCGTTACATCT
GGATCAAATACTTCAACCTGTTTCGACAAAGAACTGAACGAAAAAGAAATCAAAGACCTGTACGACAACCA
GTCTAACTCTGGTgcactagtATTTTGAAGGACTTTTGGGGCGACTATCTCCAGTACGACAAACCTTAC
TATATGCTGAATTTGTATGATCCCAACAAATATGTGGATGTGAATAACGTTGGTATTAGGGGTTACATGT
ATTTGAAGGGTCCAAGGGGGTCAGTCATGACAACCAATATCTACTTAAATTCCTCTCTTTACCGAGGGAC
AAAATTCATTATCAAAAAGTATGCTAGTGGAATAAAGATAATATAGTCAGAAACAATGATCGCGTTTAC
ATTAACGTGGTAGTCAAAAATAAGGAGTATAGACTAGCTACGAATGCATCGCAGGCGGGAGTGGAGAAGA
TACTGAGCGCACTAGAAATACCTGACGTAGGAAACTTAAGCCAGGTTGTCGTTATGAAATCAAAGAACGA
TCAAGGAATTACTAATAAGTGTAAAGATGAACCTACAAGATAACAATGGCAATGATATAGGCTTCATCGGG
TTTCATCAATTTAACAACATAGCGAAACCTCGTAGCCTCTAACTGGTACAACCGTCAAATCGAACGAAGTT
CCCGTACTCTAGGTTGCTCGTGGGAGTTCATCCAGTAGACGACGGGTGGGGCGAACGGCCGCTTgcgct
agcaCACCATCATCACcaccatcaccatcaccattaatga

SEQ ID NO: 21

GSMEFVNKQFNYKDPVNGVDIAYIKIPNAGQMOPVKAFKIHNKI WVIPERDTFTNPEEGDLNPPPEAKQV
PVSYYDSTYLSTDNEKDNYLKGVTCLFERIYSTDLGRMLLTSIVRGIPFWGGSTIDTELKVIDTNCINVI
QPDGSYRSEELNLVIIGPSADIIQFECKSFGHEVLNLTRNGYGSTQYIRFSPDFTFGFEESLEVDTNP LL

GAGKFATDPAVTLAHELIIHAGHRLYGIAINPNRVFKVNTNAYYEMSGLEVSFEELRTFGGHDAKFIDSLQ
ENEFRLYYYNKFKDIASTLNKAKSIVGTTASLQYMKNVFKEKYLLSEDTSGKFSVDKLFKDKLYKMLTEI
YTEDNFVKFFKVLNRKTYLNFDAVFKINIVPKVNYTIYDGFNLRTNLAANFNGQNTTEINNMNFTKLKN
FTGLFEFYKLLCVDGIITSKTSDDDDKNKALNLQCIKVNNDLFFSPSEDNFTNDLNKGEEITS DTNIE
AAEENISLDLIQQYYLTFNFDNEPENISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQE
FEHGRSRRIALTNSVNEALLNPSRVYTFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKI
ADITIIIPYIGPALNIGNMLYKDDFVGALIFSGAVILLEFIPEIAIPVLGTFALVSYIANKVLT VQTIDN
ALSKRNEKWDEVYKYIVTNWLAKVNTQIDLIRKKMKEALENQAEATKAIINYQYNQYTEEEKNNINFNID
DLSSKLNESINKAMININKFLNQCSVSYLMNSMIPYGVKRLEDFDASLKDALLKYIYDNRGTLIGQVDRL
KDKVNNTLSTDIPFQLSKYVDNQRLSTFTEYIKNIINTSLEILNLRYESNHLIDLSRYASKINIGSKVN
FDPIDKNQIQLFNLESSKIEVILKNAIVNSMYENFSTSFWIRIPKYFNSISLNNEYTIINCMENNSGWK
VSLNYGEIIWTLQDTQEIKQRVVFKYSQMINISDYINRWIFVTITNNRLNNSKIYINGRLIDQKPISNLG
NIHASNNIMFKLDGCRDTHRYIWI KYFNLFDKELNEKEIKDLYDNQSN SGALVILKDFWGDYLYDKPYY
MLNLYDPNKYVDVNNVGIRGYMYLKGPRGSVMTTNIYLNSSLYRGTKFIKKYASGNKDNIVRNNDRVYI
NVVVKNKEYRLATNASQAGVEKILSALEIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGF
5 HQFNNI AKLVASNWYNRQIERSSRTLGC SWEFIPVDDGWGERPLALAH HHHHHHHHHH

SEQ ID NO: 22

atgccgatcaccatcaacaacttcaactacagcgatccggtggataa caaaaaacatcctgtacctggata
 cccatctgaataccctggcgaacgaaccggaaaaagcggttctgtatcacgggaacatttgggttattcc
 ggatcggttttagccgtaacagcaaccgaatctgaataaacggccgctgttaccagcccgaaaagcggt
 tattacgatccgaactatctgagcaccgatagcgataaagataccttcttgaaagaaatcatcaaactgt
 tcaaacgcatcaacagccgtgaaattggcgaagaactgatctatcgctgagcaccgatattccgtttcc
 gggcaacaacaacaccccgatcaacacctttgatttcgatgtggatttcaacagcggttgatgttaaaacc
 cgccagggtaacaattgggtgaaaaccggcagcattaacccgagcggtgattattaccgggtccgcgcgaaa
 acattattgatccggaaaccagcacctttaaaactgaccaacaacacctttgcggcgaggaaggttttgg
 cgcgctgagcattattagcattagcccgcgctttatgctgacctatagcaacgcgaccaacgatgttatt
 gaaggccgtttcagcaaaagcgaattttgcatggaccgatcctgatcctgatgcatgaactgaacctg
 cgatgcataacctgtatggcatcgcgattccgaacgatcagaccattagcagcgtagcagcaacatctt
 ttacagccagtacaacgtgaaactggaatatgcggaaatctatgcgtttggcggtccgaccattgatctg
 attccgaaaagcgcgcgcaataacttcgaagaaaaagcgctggattactatcgagcattgcgaaacgtc
 tgaacagcattaccaccggaatccgagcagcttcaacaaatatatcggcgaatataaacagaaactgat
 ccgcaaatatcgctttgtggtggaaagcagcggcgaagtaccgttaaccgcaataaattcgtggaactg
 tacaacgaactgaccagatcttcaccgaatttaactatgcgaaaatctataacgtgcagaaccgtaaaa
 tctacctgagcaacgtgtataccccgggtgaccgcgaatattctggatgataacgtgtacgatatccagaa
 cggtttaacatcccgaagcaacctgaacgttctgtttatgggccagaacctgagccgtaatccggcg
 ctgcgtaaaagtgaaccggaaaacatgctgtacctgttcaccaaattttgcGTCGAcGCGGACGATGACG
 ATAAACTGTACAACAAAACCTGCAGtgctcgatgaactgctggtgaaaaacaccgatctgccgtttattgg
 cgatatcagcgatgtgaaaaccgatatcttctgcgcaaagatatcaacgaagaaaccgaagtgatcccg
 gataacgtgagcggttgatcaggtgatcctgagcaaaaacaccagcgaacatggtcagctggatctgctgt
 atccgagcattgatagcgaaagcgaaattctgcccggcgaaaaccaggtgttttacgataaccgtaccca
 gaacgtggattacctgaacagctattactacctggaaagccagaaactgagcgataacgtggaagatttt
 acctttacccgcagcattgaagaagcgctggataacagcgcgaaagtttacacctattttccgacctgg
 cgaacaaagttaatgcgggtgttcagggcggtctgtttctgatgtgggcgaacgatgtggtggaagattt

caccaccaacatcctgcgtaaagataccctggataaaatcagcgatgtagcgcgattattccgtatatt
 ggtccggcgctgaacattagcaatagcgtgcgtcgtggcaattttaccgaagcgtttgcggttaccggtg
 tgaccattctgctggaagcgtttccggaatttaccattccggcgctgggtgcgtttgtgatctatagcaa
 agtgcaggaacgcaacgaaatcatcaaaaccatcgataactgcctggaacagcgattataaacgctggaaa
 gatagctatgaatggatgatgggcacctggctgagccgtattatcacccagttcaacaacatcagctacc
 agatgtacgatagcctgaactatcaggcgggtgcgattaaagcgaaaatcgatctggaatacaaaaaata
 cagcggcgagcgataaagaaaacatcaaaagccagggttgaaaacctgaaaaacagcctggatgtgaaaatt
 agcgaagcgatgaataacatcaaaaaattcatccgcgaatgcagcgtagacctacctgttcaaaaacatgc
 tgccgaaagtgcgatgaactgaacgaatttgatcgcaacaccaaagcgaaaactgatcaacctgatcga
 tagccacaacattattctgggtgggcgaagtggataaactgaaagcgaaagttaacaacagcttccagaac
 accatcccgtttaacatcttcagctataccaacaacagcctgctgaaagatatcatcaacgaataacttca
 atctagaaggtggcggtgggtccggtggcggtggctcaggcggggcggtagcgactagacaactctga
 ctctgaatgcccgctgtctcacgacggttactgcctgcacgacggtgtttgcatgtacatcgaagctctg
 gacaaatacgcttgcaactgcgttgttggttacatcggtgaacggttgccagtaccgtgacctgaaatggt
 gggaactgcgt

SEQ ID NO: 23

MPITINNFNYS DPVDNKNILYLDTHLNTLANEPEKA FRITGNIWVI PDRFSRNSNP NLNKP PRTVSPKSG
 YYDPNYLSTDS DKDTFLKEI IKLFKRINSREIGEELIYRLSTDIPFP GNNNTPI NTFDFDVFNSVDVKT
 RQGNWVKTGSINPSVII TGPRENIIDPETSTFKLTNNTFAAQEGFGALSII SISPRFMLTYSNATNDVI
 EGRFSKSEFCMDPILILMHELNHAMHNLYGIAI PNDQTISSVTSNIFYSQYNVKLEYAEIYAFGGPTIDL
 IPKSARKYFEKALDYRSIAKRLNSITTANPSSFNKYIGEYKQKLIRKYRFVVESSGEVTVNRNKFVEL
 YNELTQIFTEFN YAKIYNVQNRKIYLSNVYTPVTANILDDNVYDIQNGFNIPKSNLNVLFMGQNL SRNPA
 LRKVPENMLYLFTKFCVDADDDDKLYNKTLCRELLVKNTDLPFIGDISDVKTDIFLRKDINEETEVI P
 DNVSDQVILSKNTSEHGQLDLLYPSIDSESEILPGENQVFYDNRTQNV DYLNSYYLESQKLSDNVEDF
 TFTRSIEEALDNSAKVYTYFPTLANKVNAGVQGGLFLMWANDVVEFTTNILRKDTLDKISDVSAIIPYI
 GPALNISNSVRRGNFTEAFVGTGVTILLEAFPEFTIPALGAFVIYSKVQERNEIIKTIDNCLEQRIKRWK
 DSYEWMMGTWLSRIITQFNNISYQMYDSLNYQAGAIKAKIDLEYKKYSGSDKENIKSQVENLKNSLDVKI
 SEAMNNINKFIRECSVTYLFKNMLPKVIDELNEFD RNTKAKLINLIDSHNII LVGEVDKLKAKVNNSFQN
 TIPFNIFSYTNNSLLKDI INEYFNLEGGGGSGGGGSGGGGSALDNSDSECPLSHDGYCLHDGVC MYIEAL
 5 DKYACNCVVGYIGERCQYRDLKWWELR

SEQ ID NO: 24

MPITINNENYSDPVDNKNILYLDTHLNTLANEPEKAFRITGNIWVIPDRFSRNSNPNLNKPPRVTSPKSG
 YYDPNYLSTDSKDTFLKEIIKLFKRINSREIGEELIYRLSTDIPFPGNNNTPINTFDFDVFNSVDVKT
 RQGNWVKTGSINPSVIITGPRENIIDPETSTFKLTNNTFAAQEGFGALSIIISIVPRFSLTYSNATNDVG
 EGRFSKSEFCMDPILILMHENHAMHNLYGIAIPNDQTISSVTSNIFYSQYNVKLEYAEIYAFGGPTIDL
 IPKSARKYFEEKALDYRSIAKRLNSITTANPSSFNKYIGEYKQKLIRKYRFVVESSGEVTVNRNKFVEL
 YNELTQIFTEFNYAKIYNVQNRKIYLSNVYTPVTANILDDNVYDIQNGFNIPKSNLNVLFMGQNLSRNP
 LRKVPENMLYLFTKFCVDADDDDKLYNKTLQCRELLVKNTDLPFIGDISDVKTDFLRKDINEETEVY
 YPDNVSDQVILSKNTSEHGQLDLLYPSIDSESEILPGENQVFYDNRTQNVLYLSYYYLESQKLSDNVE
 DFTFTRSIEEALDNSAKVYTYFPTLANKVNAGVQGGFLMWANDVVEDFTTNILRKDTLDKISDVSAIIP
 YIGPALNISNSVRRGNFTEAFVGTILLEAFPEFTIPALGAFVIYSKVQERNEIIKTIDNCLEQRIKR

WKDSYEWMMGTWLSRIITQFNNISYQMYDSLNYQAGAIKAKIDLEYKKYSGSDKENIKSQVENLKNSLDV
 KISEAMNNINKFIRECSVTYLFKNMLPKVIDELNEFDRNTKAKLINLIDSHNIIILVGEVDKLLAKVNNSF
 QNTIPFNIFSNTNSLLKDIINEYFNLEGGGGSGGGSGGGGSALDNSDSECPLSHDGYCLHDGVCMYIE
 ALDKYACNCVVGYIGERCQYRDLKWWELR

5 SEQ ID NO: 25

MPITINNENYSDPVDNKNILYLDTHLNTLANEPEKAFRITGNIWVIPDRFSRNSNPNLNKPPRVTSPKSG
 YYDPNYLSTDSKDTFLKEIIKLFKRINSREIGEELIYRLSTDIPFPGNNNTPINTFDFDVFNSVDVKT
 RQGNWVKTGSINPSVIITGPRENIIDPETSTFKLTNNTFAAQEGFGALSIIISIPRFLTYSNATNDVG
 TPRFSKSEFCMDPILILMHENHAMHNLYGIAIPNDQTISSVTSNIFYSQYNVKLEYAEIYAFGGPTIDL
 IPKSARKYFEEKALDYRSIAKRLNSITTANPSSFNKYIGEYKQKLIRKYRFVVESSGEVTVNRNKFVEL
 YNELTQIFTEFNYAKIYNVQNRKIYLSNVYTPVTANILDDNVYDIQNGFNIPKSNLNVLFMGQNLSRNP
 LRKVPENMLYLFTKFCVDADDDDKLYNKTLQCRELLVKNTDLPFIGDISDVKTDFLRKDINEETEVY
 YPDNVSDQVILSKNTSEHGQLDLLYPSIDSESEILPGENQVFYDNRTQNVLYLSYYYLESQKLSDNVE
 DFTFTRSIEEALDNSAKVYTYFPTLANKVNAGVQGGFLMWANDVVEDFTTNILRKDTLDKISDVSAIIP
 YIGPALNISNSVRRGNFTEAFVGTILLEAFPEFTIPALGAFVIYSKVQERNEIIKTIDNCLEQRIKR
 WKDSYEWMMGTWLSRIITQFNNISYQMYDSLNYQAGAIKAKIDLEYKKYSGSDKENIKSQVENLKNSLDV
 KISEAMNNINKFIRECSVTYLFKNMLPKVIDELNEFDRNTKAKLINLIDSHNIIILVGEVDKLLAKVNNSF
 QNTIPFNIFSNTNSLLKDIINEYFNLEGGGGSGGGSGGGGSALDNSDSECPLSHDGYCLHDGVCMYIE
 ALDKYACNCVVGYIGERCQYRDLKWWELR

SEQ ID NO: 26

10

MPITINNFNYS DPVDNKNILYLDTHLNTLANEPEKAFRITGNIWVIPDRFSRNSNP NLNKP PRTVSPKSG
 YYDPNYLSTDSKDTFLKEIIKLFKRINSREIGEELIYRLSTDIPFPGNNNTPINTDFDVFNSVDVKT
 RQGNNWVK TGSINPSVII TGPRENIIDPETSTFKLTNNTFAAQEGFGALSIISISPRFMLTYSNATNDVG
 EGRFSKSEFCMDPILILMHENHAMHNLYGIAIPNDQTISSVTSNIFYSQYNVKLEYAEIYAFGGPTIDL
 IPKSARKYFEEKALDYRSIAKRLNSITTANPSSFNKYIGEYKQKLIRKYRFVVESSGEVTVNRNKFVEL
 YNELTQIFTEFN YAKIYNVQNRKIYLSNVYTPVTANILDDNVYDIQNGFNIPKSNLNVLFMGQNL SRNPA
 LRKVPENMLYLFTKFCVDADDDDKLYNKT LQCRELLVKNTDLPFIGDISDVKTDIFLRKDINEETEVIY
 YPDNVSDQVILSKNTSEHGQLDLLYPSIDSESEILPGENQVFYDNRTQNV DYLSYYYLESQKLSDNVE
 DFTFTRSIEEALDNSAKVYTYFPTLANKVNAGVQGGFLMWANDVVEDFTTNILRKDTLDKISDVSAIIP
 YIGPALNISNSVRRGNFTEAFVGTILLEAFPEFTIPALGAFVIYSKVQERNEI IKTIDNCLEQRIKR
 WKDSYEWMMGTWLSRIITQFN NISYQMYDSLNYQAGAIKAKIDGRYKKYSGSDKENIKSQVENLKNSLDV
 KISEAMNNINKFIRECSVTYLFKNMLPKVIDELNEFDRNTKAKLINLIDSHNII LVGEVDK LKAKVNNSF
 QNTIPFNIFSYTNNSLLKDIINEYFNLEGGGGSGGGGSGGGGSALDNSDSECPLSHDGYCLHDGVC MYIE
 ALDKYACNCVVG YIGERCQYRDLKWWELR

SEQ ID NO: 27

MPITINNFNYS DPVDNKNILYLDTHLNTLANEPEKAFRITGNIWVIPDRFSRNSNP NLNKP PRTVSPKSG
 YYDPNYLSTDSKDTFLKEIIKLFKRINSREIGEELIYRLSTDIPFPGNNNTPINTDFDVFNSVDVKT
 RQGNNWVK TGSINPSVII TGPRENIIDPETSTFKLTNNTFAAQEGFGALSIISISPRFMLTYSNATNDVG
 EGRFSKSEFCMDPILILMHENHAMHNLYGIAIPNDQTISSVTSNIFYSQYNVKLEYAEIYAFGGPTIDL
 5 IPKSARKYFEEKALDYRSIAKRLNSITTANPSSFNKYIGEYKQKLIRKYRFVVESSGEVTVNRNKFVEL
 YNELTQIFTEFN YAKIYNVQNRKIYLSNVYTPVTANILDDNVYDIQNGFNIPKSNLNVLFMGQNL SRNPA
 LRKVPENMLYLFTKFCVDADDDDKLYNKT LQCRELLVKNTDLPFIGDISDVKTDIFLRKDINEETEVIY
 YPDNVSDQVILSKNTSEHGQLDLLYPSIDSESEILPGENQVFYDNRTQNV DYLSYYYLESQKLSDNVE
 DFTFTRSIEEALDNSAKVYTYFPTLANKVNAGVQGGFLMWANDVVEDFTTNILRKDTLDKISDVSAIIP
 YIGPALNISNSVRRGNFTEAFVGTILLEAFPEFTIPALGAFVIYSKVQERNEI IKTIDNCLEQRIKR
 WKDSYEWMMGTWLSRIITQFN NISYQMYDSLNYQAGAIKAKIDLEYKKYIDGRKENIKSQVENLKNSLDV
 KISEAMNNINKFIRECSVTYLFKNMLPKVIDELNEFDRNTKAKLINLIDSHNII LVGEVDK LKAKVNNSF
 QNTIPFNIFSYTNNSLLKDIINEYFNLEGGGGSGGGGSGGGGSALDNSDSECPLSHDGYCLHDGVC MYIE
 ALDKYACNCVVG YIGERCQYRDLKWWELR

SEQ ID NO: 28

10

MPITINNFNYS D PVDNKNILYLDTHLNTLANEPEKA FRITGNIWVIPDRFSRNSNP NLNKP RVTSPKSG
 YYDPNYLSTDS DKDTFLKEI IKLFKRINSREIGEELIYRLSTDIPFPGNNNTPI NTFDFD VDFNSVDVKT
 RQGNNWVK TGSINPSVII TGP RENIIDPETSTFKLTNNTFAAQEGFGALSIISISPRFMLTYSNATNDVG
 EGRFSKSEFCMDPILILMH ELNHAMHNLYGIAIPNDQTISSVTSNIFYSQYNVKLEYAEIYAFGGPTIDL
 IPKSARKYFEEKALDY YRSIAKRLNSITTANPSSFNKYIGEYKQKLIRKYRFVVESSGEVTVNRNKFVEL
 YNELTQIFTEFN YAKIYNVQNRKIYLSNVYTPVTANILDDNVYDIQNGFNIPKSNLNLVLFMGQNL SRNPA
 LRKVN PENMLYLFTKFCVDADDDDKLYNKT LQCRELLVKN TDLPFIGDISDVKT DIFLRKDINEETEVIY
 YPDNVSV DQVILSKNTSEHGQLDLLYPSIDSESEILPGENQVFYDNRTQNV DYLSYYYLESQKLSDNVE
 DFTFTRSIEEALDNSAKVYTYFPTLANKVNAGVQGGFLMWANDVVEDFTTNILRKDTLDKISDVSAIIP
 YIGPALNISNSVRRGNFTEAFAVTGVTILLEAFPEFTIPALGAFVIYSKVQERNEI IKTI DNCLEQRIKR
 WKDSYEWMMGTWLSRIITQFN NISYQMYDSLNYQAGAIKAKIDLEYKKYGVPRKENIKSQVENLKNSLDV
 KISEAMNNINK FIRECSVTYLFKNMLPKVIDELNEFDRNTKAKLINLIDSHNII LVGEVDKLKAKVNNSF
 QNTIPFNIFSYTNNSLLKDIINEYFNLEGGGGSGGGGSGGGGSALDNSDSECPLSHDGYCLHDGVC MYIE
 ALDKYACNCVVG YIGERCQYRDLKWWELR

SEQ ID NO: 29

MTWPVKDFNYS D PVDNDNDILYLRIPQNK LITTPVKAFMITQNIWVIPERFSSDTNP SLSKPPRPTS KYQS
 YYDPSYLSTDEQKDTFLKGI IKLFKRINERDIGKKLINYL VVGSPFMGDSSTPEDTFDFTRHTTNI AVEK
 FENGSWKVTNIITPSVLI FGPLPNILDY TASLT LQGQSNPSFEGFGTLSILKVAPEFLLTFSDVTSNQS
 SAVLGKSIFCMDPVIALMH ELTSLHQLYGINIPSDKRIRPQVSEGFFSQDGP NVQFEELYTFGGLDVEI
 IPQIERSQLREKALGHYKDI AKRLNNINKTIPSSWISNIDKYKKIFSEKYNFDKDNTGNFVVNIDKFNSL
 YSDLTNVMSEVVYSSQYNVKNRTHYFSRHYLPVFANILDDNIYTIRDGFNL TNKGFN IENSGQNIERNPA
 LQKLSSESVDLFTKVCVDGGGGSADDDDKHSDAVFTDNYTRLRRQLAVRRYLNSILNALAGGGGSGGGG
 SGGGGSALALQCIKVKNR LPYVADKDSISQEIFENKIITDETNVQNYSDKFSLDESILDGQVPINPEIV
 DPLLPNVNMEPLNLPGE EIVFYDDITKYVDYLSYYYLESQKLSNNVENITLTTSVEEALGYSNKIYTF L
 PSLAEKVNKG VQAGLFLNWANEVVEDFTTNIMKKDTLDKISDVSVIIPYIGPALNIGNSALRGNFNQAFA
 TAGVAFLLLEGFP EFTIPALGVFTFYSSIQEREKIIKTIENCLEQRVKRWKDSYQWMVSNWLSRIITQFNH
 INYQMYDSL SYQADAIAKAKIDLEYKKYIDGRKENIKSQVENLKNSLDVKISEAMNNINK FIRECSVTYLF
 KNMLPKVIDELNKFDLRTKTELINLIDSHNII LVGEVDRLKAKVNESFENTMPFNIFSYTNNSLLKDIIN
 EYFNLEA

5

SEQ ID NO: 30

MEFVNKQFNYKDPVNGVDIAYIKIPNAGQMOPVKAFKIHNKIWIWIPERDTFTNPEEGDLNPPPEAKQVPV
 SYDSTYLSTDNEKDNYLKGVTCLFERIYSTDLGRGLLTSIVRGIPFWGGSTIDTELKVIDTNCINVIQP
 DGSYRSEELNLVIIIGPSADIIQFECKSFGHEVLNLTRNGYGSTQYIRFSPDFTFGFEESLEVDTNPLLGA
 GKFATDPAVTLAHELIIHAGHRLYGIAINPNRVFKVNTNAYYEMSGLEVSFEELRTFGGHDAKFIDSLQEN
 EFRLYYYNKFKDIASTLNKAKSIVGTTASLQYMKNVFKEKYLLSEDTSGKFSVDKLFKLYKMLTEIYT
 EDNFVKFFKVLNRKTYLNFDAVKINIVPKVNYTIYDGFNLRTNLAANFNGQNTTEINNMNFTKLKNFT
 GLFEFYKLLCVDGIITSKTKSDDDDKNKALNLQCIKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAA
 EENISLDLIQQYYLTFNFDNEPENISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFE
 HGKSRIALTNSVNEALLNPSRVYTFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADI
 TIIIPYIGPALNIGNMLYKDDFVGALIFSGAVILLEFIPEIAIPVLGTFALVSYIANKVLTQVQIDNALS
 KRNEKWDEVYKYIVTNWLAKVNTQIDLIRKKMKEALENQAATKAIINYQYNQYTEEEKNNINFNIDDL
 SKLNEKINAMININKFLNQCSVSYLMNSMIPYGVKRLEDFDASLKDALLKYIYDNRGTLIGQVDRKDK
 VNNTLSTDIPFQLSKYVDNQRLSTLEGGGGSGGGSGGGGSALDNSDSECPLSHDQYCLHDGVCMYIEA
 LDKYACNCVVGYIGERCQYRDLKWWELR

SEQ ID NO: 31

MEFVNKQFNYKDPVNGVDIAYIKIPNAGQMOPVKAFKIHNKIWIWIPERDTFTNPEEGDLNPPPEAKQVPV
 SYDSTYLSTDNEKDNYLKGVTCLFERIYSTDLGRGLLTSIVRGIPFWGGSTIDTELKVIDTNCINVIQP
 DGSYRSEELNLVIIIGPSADIIQFECKSFGHEVLNLTRNGYGSTQYIRFSPDFTFGFEESLEVDTNPLLGA
 GKFATDPAVTLAHELIIHAGHRLYGIAINPNRVFKVNTNAYYEMSGLEVSFEELRTFGGHDAKFIDSLQEN
 EFRLYYYNKFKDIASTLNKAKSIVGTTASLQYMKNVFKEKYLLSEDTSGKFSVDKLFKLYKMLTEIYT
 EDNFVKFFKVLNRKTYLNFDAVKINIVPKVNYTIYDGFNLRTNLAANFNGQNTTEINNMNFTKLKNFT
 GLFEFYKLLCVDGIITSKTKSDDDDKNKALNLQCIKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAA
 EENISLDLIQQYYLTFNFDNEPENISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFE
 HGKSRIALTNSVNEALLNPSRVYTFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADI
 TIIIPYIGPALNIGNMLYKDDFVGALIFSGAVILLEFIPEIAIPVLGTFALVSYIANKVLTQVQIDNALS
 KRNEKWDEVYKYIVTNWLAKVNTQIDLIRKKMKEALENQAATKAIINYQYNQYTEEEKNNINFNIDDL
 SKLNEKINAMININKFLNQCSVSYLMNSMIPYGVKRLEDFDASLKDALLKYIYDNRGTLIGQVDRKDK
 VNNTLSTDIPFQLSKYVDNQRLSTLEGGGGSGGGSGGGGSALDNSDSECPLSHDQYCLHDGVCMYIEA
 LDKYACNCVVGYIGERCQYRDLKWWELR

5

SEQ ID NO: 32

HMGSMEFVNKQFNYKDPVNGVDIAYIKIPNAGQMOPVKAFKIHNKIWVIPERDTFTNPEEGDLNPPPEAK
 QVPVSYDDSTYLSTDNEKDNYLKGVTCLFERIYSTDLGRGLLTSIVRGIPFWGGSTIDTELKVIDTNCIN
 VIQPDGSYRSEELNLVIIGPSADIIQFECKSFGEVLNLTRNGYGSTQYIRFSPDFTFGFEESLEVDTNP
 LLGAGKFATDPAVTLAHELIIHAGHRLYGIAINPNRVFKVNTNAYYEMSGLEVSFEELRTFGGHDAKFIDS
 LQENEFRLYYYNKFKDIASLTNKAISIVGTTASLQYMKNVFKEKYLLSEDTSGKFSVDKLFKDKLYKMLT
 EIYTEDNFVKFFKVLNRKTYLNFDAVKINIVPKVNYTIYDGFNLRTNLAANFNGQNTTEINNMNFTKL
 KNFTGLFEFYKLLCVDGIITSKTSDDDDKNKALNLQCIKVNNDLFFSPSEDNFTNDLNKGEEITSNTN
 IEAAEENISLDLIQQYYLTNFNDNEPENISIEENLSSDIIGQLELMPNIEFPPNGKKYELDKYTMFHYLRA
 QEFEGHKSRIALTNSVNEALLNPSRVYTFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDK
 IADITIIIPYIGPALNIGNMLYKDDFVGALIFSGAVILLEFIPEIAIPVLGTFALVSYIANKVLTVQTID
 NALSKRNEKWDEVYKYIVTNWLAKVNTQIDILIRKKMKEALENQAATKAIINYQYNQYTEEEKNNINFNI

DDLSSKLNESINKAMININKFLNQCSVSYLMNSMIPYGVKRLEDFDASLKDALLKYIYDNRGTLIGQVDR
 LKDKVNNTLSTDIPFQLSKYVDNQRLSTFTEYIKNIINTSLEILNLRYESNHLIDLSRYASKINIGSKV
 NFDPIDKNQIQLFNLESSKIEVILKNAIVNSMYENFSTSFWIRIPKYFNSISLNNEYTIINCMENNSGW
 KVSILNYGEIIWTLQDTQEIKQRVVFKYSQMINISDYINRWIFVTITNNRLNNSKIYINGRLIDQKPISNL
 GNIHASNNIMFKLDGCRDTHRYIWIKYFNLFDKELNEKEIKDLYDNQSNAGALVILKDFWGDYLYDKPY
 YMLNLYDPNKYVDVNNVGIRGYMYLKGPRGSVMTTNIYLNSSLYRGTKFIKKYASGNKDNIVRNNDRVY
 INVVKVKNKEYRLATNASQAGVEKILSALEIPDVGNLQVVMKSKNDQGITNKCKMNLQDNNNGNDIGFIG
 FHQFNNAIKLVASNWYNRQIERSSRTLGCSEWFI PVDDGWGERPLALAHHHHHHHHHH

5 SEQ ID NO: 33

HMGSMEFVNKQFNYKDPVNGVDIAYIKIPNAGQMOPVKAFKIHNKIWVIPERDTFTNPEEGDLNPPPEAK
 QVPVSYDDSTYLSTDNEKDNYLKGVTCLFERIYSTDLGRMLLTSIVRGIPFWGGSTIDTELKVIDTNCIN
 VIQPDGSYRSEELNLVIIGPSADIIQFECKSFGEVLNLTRNGYGSTQYIRFSPDFTFGFEESLEVDTNP
 LLGAGKFATDPAVTLAHELIIHAGHRLYGIAINPNRVFKVNTNAYYEMSGLEVSFEELRTFGGHDAKFIDG
 RQENEFRLYYYNKFKDIASLTNKAISIVGTTASLQYMKNVFKEKYLLSEDTSGKFSVDKLFKDKLYKMLT
 EIYTEDNFVKFFKVLNRKTYLNFDAVKINIVPKVNYTIYDGFNLRTNLAANFNGQNTTEINNMNFTKL
 KNFTGLFEFYKLLCVDGIITSKTSDDDDKNKALNLQCIKVNNDLFFSPSEDNFTNDLNKGEEITSNTN
 IEAAEENISLDLIQQYYLTNFNDNEPENISIEENLSSDIIGQLELMPNIEFPPNGKKYELDKYTMFHYLRA
 QEFEGHKSRIALTNSVNEALLNPSRVYTFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDK
 IADITIIIPYIGPALNIGNMLYKDDFVGALIFSGAVILLEFIPEIAIPVLGTFALVSYIANKVLTVQTID
 NALSKRNEKWDEVYKYIVTNWLAKVNTQIDILIRKKMKEALENQAATKAIINYQYNQYTEEEKNNINFNI
 DDLSSKLNESINKAMININKFLNQCSVSYLMNSMIPYGVKRLEDFDASLKDALLKYIYDNRGTLIGQVDR
 LKDKVNNTLSTDIPFQLSKYVDNQRLSTFTEYIKNIINTSLEILNLRYESNHLIDLSRYASKINIGSKV
 NFDPIDKNQIQLFNLESSKIEVILKNAIVNSMYENFSTSFWIRIPKYFNSISLNNEYTIINCMENNSGW
 KVSILNYGEIIWTLQDTQEIKQRVVFKYSQMINISDYINRWIFVTITNNRLNNSKIYINGRLIDQKPISNL
 GNIHASNNIMFKLDGCRDTHRYIWIKYFNLFDKELNEKEIKDLYDNQSNAGALVILKDFWGDYLYDKPY
 YMLNLYDPNKYVDVNNVGIRGYMYLKGPRGSVMTTNIYLNSSLYRGTKFIKKYASGNKDNIVRNNDRVY
 INVVKVKNKEYRLATNASQAGVEKILSALEIPDVGNLQVVMKSKNDQGITNKCKMNLQDNNNGNDIGFIG
 FHQFNNAIKLVASNWYNRQIERSSRTLGCSEWFI PVDDGWGERPLALAHHHHHHHHHH

SEQ ID NO: 34

HMGSMEFVNKQFNYKDPVNGVDIAYIKIPNAGQMOPVKAFKIHNKIWVIPERDTFTNPEEGDLNPPPEAK
QVPVSYDSTYLSTDNEKDNYLKGVTKLFERIYSTDLGRMLLTSIVRGIPFWGGSTIDTELKVIDTNCIN
VIQPDGSYRSEELNLVIIGPSADIIQFECKSFGHEVLNLTRNGYGSTQYIRFSPDFTFGFEESLEVDTNP
LLGAGKFATDPAVTLAHELIHAGHRLYGIAINPNRVFKVNTNAYYEMSGLEVSFEELRTFGGHDAKFIDS
LQENEFRLYYYNKFKDIASTLNKAKSIVGTTASLQYMKNVFKEKYLLEDTSKGKFSVDKLFKLYKMLT
EIYTEDNFVKFFKVLNRKTYLNFDAVKINIVPKVNYTIYDGFNLNNTNLAANFNGQNTNINMNFVKL
KNFTGLFEFYKLLCVDGIITSKTKSDDDDKNKALNLQCIKVNNDLFFSPSEDNFTNDLNKGEEITSNTN
IEAAEENISLDLIQQYYLTFNFDNEPENISIENLSSDIIGQLELMPNIEFPPNGKKYELDKYTMFHYLRA
QEFEGHIEGRIALTNSVNEALLNPSRVYTFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDK
IADITIIIPYIGPALNIGNMLYKDDFVGALIFSGAVILLEFIPEIAIPVLGTFALVSYIANKVLTQVTID
NALSKRNEKWDEVYKYIVTNWLAKVNTQIDLIRKKMKEALENQAATKAIINYQYNQYTEEEKNNINFNI
DDLSSKLNESINKAMININKFLNQCSVSYLMNSMIPYGVKRLEDFDASLKDALLKYIYDNRGTLIGQVDR
LKDKVNNTLSTDIPFQLSKYVDNQRLSTFTTEYIKNIINTSLEILNLRYESNHLIDLSRYASKINIGSKV

5

NFDPIDKNQIQLFNLESSKIEVILKNAIVYNSMYENFSTSFWIRIPKYFNSISLNNEYTIINCMENNSGW
KVS LN YGEIIWTLQDTQEIKQRVVFKYSQMINISDYINRWIFVTITNNRLNNSKIYINGRLIDQKPISNL
GNIHASNNIMFKLDGCRDTHRYIWIKYFNLFDKELNEKEIKDLYDNQSN SGALVILKDFWGDYLYDKPY
YMLNLYDPNKYVDVNNVGIRGYMYLKGPRGSVMTTNIYLNSSLYRGTKFIKKYASGNKDNIVRNDRVY
INVVVKNKEYRLATNASQAGVEKILSALEIPDVGNLQVVMKSKNDQGITNKCKMNLQDNNGNDIGFIG
FHQFNNAIKLVASNWNRYQIERSSRTLGCSEFIPVDDGWERPLALAHHHHHHHHHH

SEQ ID NO: 35

MPKINSFNYNDPVNDRTILYIKPGGCQEFYKSFNIMKNIWIIPERNVIGTTPQDFHPPTSLKNGDSSYYD
 PNYLQSDEEKDRFLKIVTKIFNRINNNLSGGILLEELSKANPYLGNDNTPDNQFHIGDASAVEIKFSNGS
 QHILLPNVIIMGAEPDLFETNSSNISLRNNYMPSNHGFGSIAIVTIVPRFSFRFNDNSINEFIQDPALTL
 MHELIHSLHGLYGAKGITTTCIITQQQNPLITNRKGINIEEFLTGGNDLNIITVAQYNDIYTNNLNDYR
 KIASKLSKVQVSNPQLNPYKDI FQEKYGLDKDASGIYSVNINKFDDILKKLYSFTEFDLATKFQVKCRET
 YIGQYKYFKLSNLLNDSIYNISEGYNNLNKVNFRGQANLNPRI IKPITGRGLVKKIIRFCKNIVSVKG
 IRKSICIEINNGELFFVASSENSYNDNINTPKEIDDTVTSNNNYENDLDQVILNFNSEAPGLSDEKLNL
 TIQNDAYIPKYDSNGTSDIEQHDVNELNVFFYLDAQVPEGENNVLNTSSIDTALLEQPKIYTFSSSEFI
 NNVNKPVQAALFVSWIQQVLVDFTTEANQKSTVDKIADISIVVPYIGLALNIGNEAQKGNFKDALELLGA
 GILLEFEPELLIPTILVFTIKSFLGSSDNKNKVIKAINNALKERDEKWKEVYSFIVSNWMTKINTQFNKR
 KEQMYQALQONQVNAIKTIIIESKYNSYTLLEEKNELTNKYDIKQIENELNQKVSIAMNNIDRFLTESSISYL
 MKLINEVKINKLREYDENVKTYLLNYIIQHGSILGESQQELNSMVTDTLNN SIPFKLSSYTDKILISYF
 NKFFKRIKSSSVLNMRYKNDKYVDTSYGDSNININGDVYKYPTNKNQFGIYNDKLSEVNISQNDYIIYDN
 KYKNFSISFWVRIPNYDNKIVNVNNEYTIINC MRDNNSGWKVS LNHN EI IWTLQDNAGINQKLAFNYGNA
 NGISDYINKWIFVTITNDR LGDSKLYINGNLIDQKSILNLGNIHVSDN ILFKIVNCSYTRYIGIRYFNIF
 DKELDETEIQTLYSNEPNTN ILKDFWGNLYLLYDKEYYLLNVLPNNFIDRRKDSTLSINNIRSTILLANR
 LYSGIKVKIQRVNNSSTNDNLVRKNDQVYIN FVASKTHLFPLYADTATTNKEKTIKISSSGNRFNQVVVM
 NSVGNNCTMNFKNNGNNIGLLGFKADTVVASTWYYTHMRDHTNSNGCFWNFI SEEHGWQEK

SEQ ID NO: 36

MPKINSFNYNDPVNDRTILYIKPGGCQEFYKSFNIMKNIWIIPERNVIGTTPQDFHPPTSLKNGDSSYYD
 PNYLQSDEEKDRFLKIVTKIFNRINNNLSGGILLEELSKANPYLGNDNTPDNQFHIGDASAVEIKFSNGS
 QHILLPNVIIMGAEPDLFETNSSNISLRNNYMPSNHGFGSIAIVTFSPEYSFRFNDNSINEFIQDPALTL
 MHELIHSLHGLYGAKGITTTCIITQQQNPLITNRKGINIEEFLTGGNDLNIITVAQYNDIYTNNLNDYR
 KIASKLSKVQVSNPQLNPYKDI FQEKYGLDKDASGIYSVNINKFDDILKKLYSFTEFDLATKFQVKCRET
 YIGQYKYFKLSNLLNDSIYNISEGYNNLNKVNFRGQANLNPRI IKPITGRGLVKKIIRFCKNIVSVKG
 IRKSICIEINNGELFFVASSENSYNDNINTPKEIDDTVTSNNNYENDLDQVILNFNSEAPGLSDEKLNL
 TIQNDAYIPKYDSNGTSDIEQHDVNELNVFFYLDAQVPEGENNVLNTSSIDTALLEQPKIYTFSSSEFI
 NNVNKPVQAALFVSWIQQVLVDFTTEANQKSTVDKIADISIVVPYIGLALNIGNEAQKGNFKDALELLGA
 GILLEFEPELLIPTILVFTIKSFLGSSDNKNKVIKAINNALKERDEKWKEVYSFIVSNWMTKINTQFNKR
 KEQMYQALQONQVNAIKTIIIESKYNSYIEGRKNELTNKYDIKQIENELNQKVSIAMNNIDRFLTESSISYL
 MKLINEVKINKLREYDENVKTYLLNYIIQHGSILGESQQELNSMVTDTLNN SIPFKLSSYTDKILISYF
 NKFFKRIKSSSVLNMRYKNDKYVDTSYGDSNININGDVYKYPTNKNQFGIYNDKLSEVNISQNDYIIYDN
 KYKNFSISFWVRIPNYDNKIVNVNNEYTIINC MRDNNSGWKVS LNHN EI IWTLQDNAGINQKLAFNYGNA
 5 NGISDYINKWIFVTITNDR LGDSKLYINGNLIDQKSILNLGNIHVSDN ILFKIVNCSYTRYIGIRYFNIF
 DKELDETEIQTLYSNEPNTN ILKDFWGNLYLLYDKEYYLLNVLPNNFIDRRKDSTLSINNIRSTILLANR
 LYSGIKVKIQRVNNSSTNDNLVRKNDQVYIN FVASKTHLFPLYADTATTNKEKTIKISSSGNRFNQVVVM
 NSVGNNCTMNFKNNGNNIGLLGFKADTVVASTWYYTHMRDHTNSNGCFWNFI SEEHGWQEK

SEQ ID NO: 37

10

MGSMPKINSFNNDPVDNDRITILYIKPGGCQEFYKSFNIMKNIWIIIPERNVIGTTPQDFHPPTSLKNGDSS
 YYDPNYLQSDEEKDRFLKIVTKIFNRINNNLSGRGLLEELSKANPYLGNDNTPDNQFHIGDASAVEIKFS
 NGSQHILLPNVIIMGAEPDLFETNSSNISLRNNYMPSNHGFGSIAIVTFSPEYSFRFNDNSINEFIQDPA
 LTMHELIHSLHGLYGAKGITTTCIITQQQNPLITNRKGINIEEFLTFGGNDLNIITVAQYNDIYTNNLN
 DYRKIASKLSKVQVSNPQLNPYKDIHQEYGLDKDASGIYSVNINKFDDILKKLYSFTEFDLATKFQVKC
 RETYIGQYKYFKLSNLLNDSIYNISEGYNNLNKVNFRGQANLNPRIIKPITGRGLVKKIIRFCVDGGG
 GSADDDDKHSDAVFTDNYTRLRRQLAVRRYLNSILNALAGGGGSGGGGSGGGGSALVLQCIENNGELFF
 VASENSYNDNINTPKEIDDTVTSNNNYENDLDQVILNFNSESAPGLSDEKLNLTIONDAYIPKYDSNGT
 SDIEQHDVNELNVFFYLDAQVPEGENNVNLTSSIDTALLEQPKIYTFSSSEFINNVNKPQAAALFVSWI
 QQVLVDFTEANQKSTVDKIADISIVVPYIGLALNIGNEAQKGNFKDALELLGAGILLEFEPELLIPTIL
 VFTIKSFLGSSDNKNKVIKAINNALKERDEKWKEVYSFIVSNWMTKINTQFNKRKEQMYQALQNVNAIK
 TIIESKYSYTLLEKNELTNKYDIKQIENELNQKVSAMNNIDRFLTESSISYLMKIINEVKINKLREYD
 ENVKTYLLNYIIQHGSILGESQQELNSMVTDTLNN SIPFKLSSYTDDKILISYFNKFFKG

SEQ ID NO: 38

MGSMPKINSFNNDPVDNDRITILYIKPGGCQEFYKSFNIMKNIWIIIPERNVIGTTPQDFHPPTSLKNGDSS
 YYDPNYLQSDEEKDRFLKIVTKIFNRINNNLSGGILLEELSKANPYLGNDNTPDNQFHIGDASAVEIKFS
 NGSQHILLPNVIIMGAEPDLFETNSSNISLRNNYMPSNHGFGSIAIVTFSPEYSFRFNDNSINEFIQDPA
 LTMHELIHSLHGLYGAKGITTTCIITQQQNPLITNRKGINIEEFLTFGGNDLNIITVAQYNDIYTNNLN
 DYRKIASKLSKVQVSNPQLNPYKDIHQEYGLDKDASGIYSVNINKFDDILKKLYSFTEFDLATKFQVKC
 RETYIGQYKYFKLSNLLNDSIYNISEGYNNLNKVNFRGQANLNPRIIKPITGRGLVKKIIRFCVDGGG
 GSADDDDKHSDAVFTDNYTRLRRQLAVRRYLNSILNALAGGGGSGGGGSGGGGSALVLQCIENNGELFF
 VASENSYNDNINTPKEIDDTVTSNNNYENDLDQVILNFNSESAPGLSDEKLNLTIONDAYIPKYDSNGT
 SDIEQHDVNELNVFFYLDAQVPEIEGRVNLTSSIDTALLEQPKIYTFSSSEFINNVNKPQAAALFVSWI
 QQVLVDFTEANQKSTVDKIADISIVVPYIGLALNIGNEAQKGNFKDALELLGAGILLEFEPELLIPTIL
 VFTIKSFLGSSDNKNKVIKAINNALKERDEKWKEVYSFIVSNWMTKINTQFNKRKEQMYQALQNVNAIK
 TIIESKYSYTLLEKNELTNKYDIKQIENELNQKVSAMNNIDRFLTESSISYLMKIINEVKINKLREYD
 5 ENVKTYLLNYIIQHGSILGESQQELNSMVTDTLNN SIPFKLSSYTDDKILISYFNKFFKG

SEQ ID NO: 39

MGSMPKINSFNNDPVDNDRITILYIKPGGCQEFYKSFNIMKNIWIIIPERNVIGTTPQDFHPPTSLKNGDSS
 YYDPNYLQSDEEKDRFLKIVTKIFNRINNNLSGGILLEELSKANPYLGNDNTPDNQFHIGDASAVEIKFS
 NGSQHILLPNVIIMGAEPDLFETNSSNISLRNNYMPSNHGFGSIAIVTFSPEYSFRFNDNSINEFIQDPA
 LTMHELIHSLHGLYGADGITTTCIITQQQNPLITNRKGINIEEFLTFGGNDLNIITIEGRNDIYTNNLN
 DYRKIASKLSKVQVSNPQLNPYKDIHQEYGLDKDASGIYSVNINKFDDILKKLYSFTEFDLATKFQVKC
 RETYIGQYKYFKLSNLLNDSIYNISEGYNNLNKVNFRGQANLNPRIIKPITGRGLVKKIIRFCVDGGG
 GSADDDDKHSDAVFTDNYTRLRRQLAVRRYLNSILNALAGGGGSGGGGSGGGGSALVLQCIENNGELFF
 VASENSYNDNINTPKEIDDTVTSNNNYENDLDQVILNFNSESAPGLSDEKLNLTIONDAYIPKYDSNGT

10

SDIEQHDVNELNVFFYLDAQVPEGENNVNLTSSIDTALLEQPKIYTFFSSEFINNVNKPVQAALFVSWI
QQVLVDFTTEANQKSTVDKIADISIVVPYIGLALNIGNEAQKGNFKDALELLGAGILLEFEPELLIPTIL
VFTIKSFLGSSDNKNKVIKAINNALKERDEKWKEVYSFIVSNWMTKINTQFNKRKEQMYQALQNQVNAIK
TIIESKYNSYTL EEKNELTNKYDIKQIENELNQKVS IAMNNIDRFLT ESSI SYLMKI INEVKINKLREYD
ENVKTYLLNYIIQHGSILGESQQELNSMVTDTLNNSIPFKLSSYTDDKILISYFNKFFKG

REIVINDICACIONES

1. Un polipéptido que comprende:

5 (a) una proteasa no citotóxica que es capaz de escindir una proteína SNARE;

(b) un dominio de translocación que es capaz de translocar la proteasa no citotóxica desde dentro de un endosoma de una célula de mamífero, a través de la membrana endosómica del mismo hasta el citosol de la célula de mamífero;

10

(c) un primer sitio de escisión destructivo que es escindible por una segunda proteasa y no por la proteasa no citotóxica y en que después de su escisión por la segunda proteasa, el polipéptido tiene una potencia reducida medible por una capacidad reducida de escindir dicha proteína SNARE y/o una capacidad reducida de translocar dicha proteasa no citotóxica, en que dicho primer sitio de escisión destructivo es escindible por una proteasa seleccionada entre: una proteasa circulante, una proteasa sérica, una proteasa de la cascada de coagulación sanguínea; una proteasa asociada a tejidos, una metaloproteasa matricial (MMP), una MMP muscular; o una proteasa intracelular ausente en las células nerviosas;

15

(d) una fracción de direccionamiento (TM) que se une a un sitio de unión presente en una célula nerviosa de la unión neuromuscular, en que el sitio de unión es capaz de experimentar endocitosis para incorporarse en un endosoma dentro de la célula nerviosa; y

20

(e) con la condición de que dicho primer sitio de escisión destructivo no esté localizado dentro de dicha TM.

25

2. Un polipéptido de acuerdo con la reivindicación 1, en que la proteasa es trombina, el factor Xa, ADAM17, la proteasa similar a tripsina de las vías respiratorias humanas (HAT), ACE (peptidildipeptidasa A), elastasa, furina, granzima, caspasa, preferentemente una de las caspasas 1-10, una metaloproteasa matricial (MMP), una TACE, una adamalisina, una serralisina, una astacina, el factor de coagulación VIIa, el factor de coagulación IXa, el factor de coagulación XIa, el factor de coagulación XIIa, calicreína, la proteína C o la proteasa de serina asociada a MBP.

30

3. Un polipéptido de acuerdo con cualquiera de las reivindicaciones precedentes, en que la proteasa no citotóxica comprende una cadena L de una neurotoxina clostridial o un fragmento de la misma que es capaz de escindir una proteína SNARE; o en que la proteasa no citotóxica comprende una proteasa de IgA o un fragmento de la misma que es capaz de escindir una proteína SNARE.

35

4. Un polipéptido de acuerdo con cualquiera de las reivindicaciones 1-3, en que el polipéptido comprende al menos un primer sitio de escisión destructivo presente en el componente proteásico no citotóxico y/o al menos un primer sitio de escisión destructivo presente en el dominio de translocación.

40

5. Un polipéptido de acuerdo con cualquiera de las reivindicaciones precedentes, en que el dominio de translocación comprende un dominio de translocación de una neurotoxina clostridial.

6. Un polipéptido de acuerdo con cualquiera de las reivindicaciones precedentes, en que la TM comprende un dominio H_C o H_{CC} de una neurotoxina clostridial o un fragmento del mismo que es capaz de unirse a una célula neuronal; o en que la TM comprende un péptido seleccionado de entre una hormona similar al glucagón, una neurohormona, una citocina neurorreguladora, una neurotrofina, un factor de crecimiento, una molécula señalizadora para la guía de axones, una proteína de unión a azúcares, un ligando que se une selectivamente a neurexina, un ligando para neurexina 2 α , un ligando para neurexina 2 β , un ligando para neurexina 3 α , un ligando para neurexina 3 β , una WNT, Ng-CAM(LI), NCAM, N-cadherina, un péptido PACAP tal como un péptido VIP, agrina-MUSK, un polipéptido de la membrana basal y una variante de cualquiera de los polipéptidos anteriores; una citocina neurorreguladora tal como el factor neurotrófico ciliar (CNTF), una glicoforina A (GPA), un factor inhibidor de la leucemia (LIF), una interleucina (IL), una onostatina M, una cardiotropina 1 (CT-1), una citocina similar a cardiotropina (CLC), una neuroleucina, VEGF, un factor de crecimiento similar a la insulina (IGF), un factor de crecimiento epidérmico (EGF) y una variante de cualquiera de las citocinas neurorreguladoras anteriores.

45

50

55

7. Un polipéptido de acuerdo con cualquiera de las reivindicaciones precedentes, en que después de la escisión en el sitio de escisión destructivo, el polipéptido tiene una capacidad reducida para escindir una proteína SNARE.

8. Un ácido nucleico que codifica un polipéptido de acuerdo con cualquiera de las reivindicaciones precedentes.
- 5 9. Un polipéptido de acuerdo con cualquiera de las reivindicaciones 1-7, para uso en la supresión de una afección o enfermedad seleccionada entre estrabismo, blefaroespasma, bizqueo, distonía (por ejemplo, distonía espasmódica, distonía bucomandibular, distonía focal, distonía tardía, distonía laríngea, distonía de miembros), tortícolis (por ejemplo, tortícolis espasmódica), aplicaciones para tratamientos de belleza (cosmética) que se benefician de la incapacitación de células/músculos (a través de la regulación por disminución o inactivación de las
- 10 proteínas SNARE), un trastorno o afección neuromuscular de la motilidad ocular (por ejemplo, estrabismo concomitante, estrabismo vertical, parálisis del recto lateral, nistagmo, miopatía distiroidea), calambre de los escribientes, blefaroespasma, bruxismo, enfermedad de Wilson, temblores, tics, mioclonus segmentario; espasmos, espasticidad debida a esclerosis múltiple crónica, espasticidad que resulta en un control anormal de la vejiga, anismo, espasmo de la espalda, calambres musculares, dolores de cabeza tensionales, síndrome del músculo
- 15 elevador de la pelvis, espina bífida, discinesia tardía, enfermedad de Parkinson y tartamudeo.

Figura 1

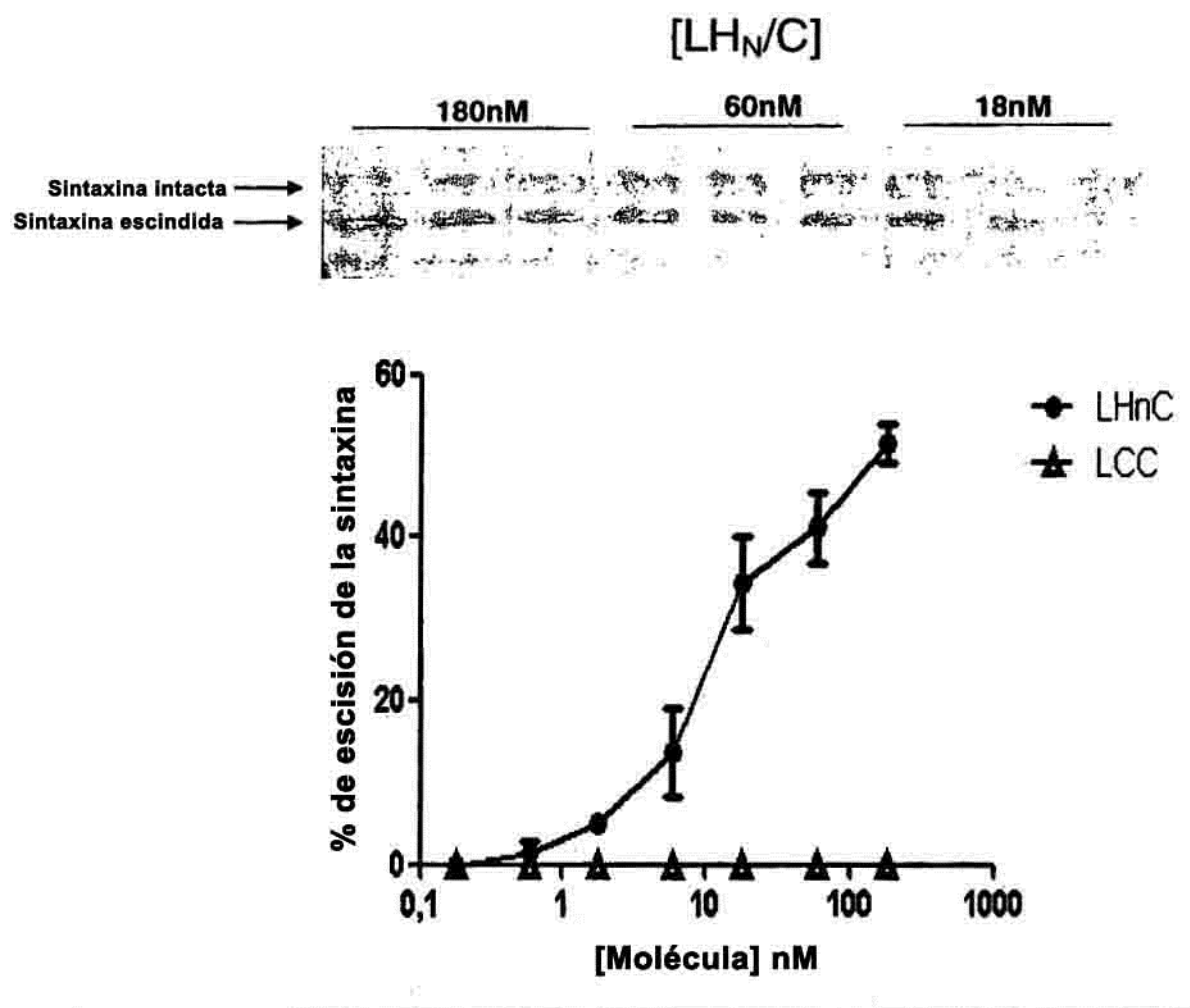


Figura 2

LHA : MEFVNKQFNKDFVNGVDIAYIKIPN--AGQMOPVRAFKIHKNIWIPIPERDTFTNPEEGDLNPPPEAKQVPV--SYDSTYLTSDNEKDNLYLKVTKLFEIYSTDLGRMLTST : 111
 LHB : MPVTINNPNYNDPIDNNNIIMWEPFARGTGRYYKAFKIITRINIWIIPERTFGYKPB--DFNKSSGIFNRDVCYDDPDYLTNDKKNIFLQTMWIKLFRNIRKSKPLGEKLEMI : 112
 LHC : MPITINNPNYSDVDNKNILYDLTHLNTLANEPEKAFRITGNIWIIPDRSRNSEN--LNKPRVTSPKS--GYDDPNYLTSDSDKDTFLKEIILKFRINSREIGELIYRL : 110
 LHD : MTWPKDFNYSDFVNDNDILYRIPOKLIITTPVKAFMITQNIWIIPERSDWTNPS--LSKPRPTSKYO--SYDDPSYLTSDDEKDTFLKGIILKFRINERDIGKKLINYL : 110
 LHE : MPK--INSEFNPNVNDRTILYIKP--GGCQEFYKSFNIMKNIWIIPERNVIGTTPQ--DFHPPTSLKNGDS--SYDDNYLQSDDEKDRFLKIVTKLFRINNNLSGGILLEEL : 107
 LHF : MPVNINNFNNDPNNFTTILYMKMPYEDSNKYKAFETMDNVMIIPERNIIGKKPS--DFYPPISLSDGSS--AYDDPNYLTDAEKDRFLKIVTKLFRINNNINSPAGQVLLLEI : 111
 LHG : MPVNIKNFNPNYNDPNNDDIIMWEPFNDPGCTYYKAFRIIDRIWIIVPERTYGQPD--QFNASTGVFSKDVVEYDPTYLKTDAEKDKFLKIVTKLFRINNSKPSQORLLDMI : 112
 M FNY DP61 I 6 KaF I 6W66PeR p YDp YL 31 K1 5Lk 6 K6F RI G L6 6

LHA : VRGIPFWGG--STIDTELKVIDTN--CINVIQPD--GSYR--SEELNLVIIGPSADIIOFECKSFGEV----LNLTRNGYGSTQYIRFSPDFTFGPEESLEVDNTNPLLGAGKFA : 214
 LHB : INGIPYLGDRRVPLEEFNTNTASVTVNKLIISNPGEVERKKGIFANLIIFGPGPVLNENETIDIGION----HFASREGFGGIMQMKFCPEYVSVFNNVQENKSGASIFNRRGYF : 221
 LHC : STDIPFPGNNWTPINTDFDVFDFNSVDVKTQGNWVKTSINPSVITIGPRENIIDPETSTFKLTNN--TFAAOEGFGALSIISISPRFMLTYSNATNDVGEGRFSKSEFC : 220
 LHD : VVGSPFMGDSSTPEDTDFTRHTTNTIAVEKFENGSKWVTNIITPSVLIFGLPDLNLDYASLTQGGQ--SNPSFEGFGTSLILKVAPEFLLTFSVTSNQSOSAVLGKSIFC : 220
 LHE : SKANPYLGNDNTPDNQPHIGDAS--AVEIKFSN--GSQD--ILLPNVIMGAEPDLFETNSSNISLRNN--YMPSNHRFGSIAIVTFSPEYSFRFN--DNC-----MN--EFI : 203
 LHF : KNGKPYLGNDHTAVNEFCANNRSTSEIKESN--GTTD--SMLNLVILGPGPNILECSTFPVRIFPNNIAVDPSKGFSGIQLMSFSTEYEVAFN--DN-----TD--LFI : 210
 LHG : VDAIPYLGNASTPPDKFAANVANVSINKKIIQPGAEDQIKGLMTNLIIFGPGPVLSDNFTDSMIMNG----HSPISEGFGARMIRFCPSCLNVFNNVQENKDTSTFSRRAYF : 221
 P5 G t f 6 661 Gp 6 5 5 5

LHA : TDPAVTLAHELTHAGHRLYGIAI--NPNRVFKVNTNAYYEMSGLEVSFEELRTFGGHDKAFIDSLQENEFRLYYNKKFKDIASLTN--K--AKSIVGTTASLQVMKNVFEKELYLS : 324
 LHB : SDPALILMHHELTHVHLGLYGIK--VDDLPIVPNEKKFFMQSDTAIQAEELYTFGGQDPSIITPSTDKSIYDKVLQNFGRIVDRINKV--LVCISDPNININIKYKNFKDKYKVF : 331
 LHC : MDPIILMHHELHAMNLYGIAIPNDQVTSNPNYFQYVNVKLEVAEYAFGGPTIDILIPKSARKYFEKALDYVRSIAKRLNSITANPSSEKNYIGEVKQKLIRKYRFV : 333
 LHD : MDPVIALMHHELTHSLHQLYGINIPSKRIRPQVSEGFSSQDGNPNVQFEELYTFGGDLDELIPQERSQLREKALGHYKDIKRLNINNKTIIPSSWISNIDKYKKITSEKYNFD : 333
 LHE : QDPALTLMHHELTHSLHGLYGAGK--ITTKYTTIQKNPLIINIRGNIIEEFLTFGGTDLNLTSAQSDNITYTLNLLADYKKIASKLS--KVQVSNPLNPN--YKDVFEAKYGLD : 310
 LHF : ADPAISLAHELTHVHLGLYGAGK--VTNKKVILEVDQALMAEKDIKEEFTFGGQDLNLTITNSTNQKIYVILLSNTYTAIASRLS--QVNRNNSALNTYTYKFNFFQWKYGLD : 319
 LHG : ADPALTLMHHELTHVHLGLYGIK--ISNLPITPNTKREFFMQHSDPQAEELYTFGGHDPSVISPSTDMNINYNKALQNFQDIANRLN--I--VSSAQGGSGIDISLYKQIYKNKYDFV : 330
 DP 6 L HEL H H LYG ee tFGG d I 1 5 Ia L yK

LHA : EDTSGKFSVDKLFKDKLYKMLTEIYTEDNFVKFFKVLNRRKTYLNFDKAVFKIN-IVPKVNTIYDGFNLRNTNLAAFNQONTENNMTKLKNFTGLFEFYKLCVDGII- : 435
 LHB : EDSEKYSIDVESFDKLYKSLMFGFTETNIAENYIKTRASYFSDSLPVVKIKNLLDNEIYTIIEGPNISDKMEKEYRQNKANKQAYEEIS-KEHLAVYKIOMCVDEEK- : 442
 LHC : KESGTVNRRKNEFVLYNELTQIFTEFNAYKIVNQNRKTYLSNRYTFVTAN-ILDDNYDIQGFNPKSNLNVFMQNLNRNPAKRVNP--ENMLYLFTRFCVDAID- : 442
 LHD : VDNNTGVAVNIDKFNLSYSLTNMSEVYSSQYNVNRTHVFSRHLVPFAN-ILDDNYTIRQGFNLTKNLFNSQNIERNPALOKLSS--ESVVDLFTKVCVDKSEE : 443
 LHE : KDSAGIYSVNINKFNDFIKKLYS-FTEFDLTKFQVQKCRQTYIGQYK-FKLSNLLNDSIYNISEGININ--NLKVNFRQGANLNPRIITPIT-GRGLVKKIIIRPCVD--- : 414
 LHF : QDSNGYTVNISKFNAIYKLF-SFTECDLAQKQVKNRSYLFHFKP-FRLDLLDDNYSISSEGFNIG--SLRVNNGQNIINLSRIVGPIP-DNGLVERFVGLCVD--- : 423
 LHG : EDPNGKYSVDKDKLYKALMFGFTETNLAGEYGIKTRYSYFSEYLPPIKTEKLLDNTIYTONEGFNIAKSKLTKTEFGNQNKAVNKEAYEEIS-LEHLVIYRIAMCKP--V- : 439
 d G 61 kf 65 L 3E 5 6 R Y 66 Y i G5N6 GQN N 6 Cvd

 LHA : -TSKTSLSIEGRNKALNLOC--IKVNMWDLFFSPSEDNFTNLDNKGEEITSDTNI EAAEENISLDLQOYVLTNFDNPEPENISLENSSDIQOLELMPNIERF--PNGKKY : 543
 LHB : -LYDDDDKDRWGS---SLQC--IDVNDLFFIADKNSFSDLSKNERIEYNTQSVNIENDPPIN--ELLDLTLISKIE-LPSENTESLTDNF-VDVPVYEQ--PAIKKI : 542
 LHC : -----GRSLYN---KTLQCRELLVXNTDLPF IGDISDVKTDFIRKRDINEETEVYIYDPNVSDV---QVILS-KNTSEHGQLDLIYPSIDSESE---ILPGENOVFYDNRTO : 539
 LHD : KLYDDDDKDRWG---SSLOC--IKVNMRLPVADKDSISQIEFENKIITDETWNQYSDKFSLD--ESILDGQVPINPEIVDPELLPVNMEPL---NLPGEIIVFYDDITK : 545
 LHE : -IEGRKGIR-----KLOC--IEINNGELFFVASENSYNDNINTPKIEDDTVTNNNYKNDLD---QVILNFSESA PG-LSDEKLNLTIOND-AYIIPKYDSNGTSDIEQH : 512
 LHF : -IEGRKGITK-----NLOC--IKVNMRLFFVASENSYNGINGINPKIEIDDTTNNNYKNDLD---EVLIDYNSDAIPN-LSSRLNLTQOND-SYVPKYDSNGTSEIKEY : 521
 LHG : -MYKNVTGK-----SEQC--IIVNNEDLFFIANKDSFSKDLAKAETIANTQNTIENNFSID---QLILDNDLSSGID-LPNEWTEFTNFDDIDIPVVIKQ--SALKKI : 535
 1QC 6 6 N L 5 T 61 2 iL 6

 LHA : ELDKYTMFHYLRAQEFEGHKSRIALTNSVNEALLNPSRVYTFSSDYVKKVKNKATEAMFLGWVEQLVYDFTDETSEVSTTDKIADIITIIIPYIGPALNIGNMLYKDDFVGAL : 656
 LHB : FTDENTIFQYLYSQTFPLDIRDISLTSFDDALLFNKVSFFSMDYIKRTANKVVEAGLFAGWVQIVNDFVIEANKSNTMDKIADISLIVPYIIGLALNVGNETAKGNFENAF : 655
 LHC : NVDYLSYYYLESQKLSDNVEDFTFTRSIEEALDNSAKVYTYFP-TLANKVNAVQVGLFLMANDVVEDFTTNILRKDTLDKISDVSAIIPYIIGPALNISNVRGNFTFAP : 651
 LHD : YVDYLSYYYLESQKLSNNVENITLTSVEEALGYNKIYTFLP-SLAEKVNKGVOAGLFLNANEVVEDFTTNIMKOTLDKISDVSAIIPYIIGPALNISNVRGNFTFAP : 657
 LHE : DVNELNVFFYLDQAKVPEGENNVNLTSSIDTALLQEPKIYTFSSDFINNVNKPQVQALFVSWIQOVLVDFTTEANQKSTVDKIADISIVVPYIIGLALNIGNEAOKGNFKDAL : 625
 LHF : TVDKLNVFFYLYAQAPAGESAISLTSVNTALLDASKVYTFSSDFINTVKNKPQVQALFVSWIQOVLVDFTTEATQKSTTDKIADISLIVPVVGLALNIGNEVQGNFKKEAI : 634
 LHG : FVDGDSLFYELHAQTFFPSNIENLQLTNSLNDALRNNNKVVYTFSTNLVEKANTVVGASLFVNVWKGVIDDFTSESTQKSTIDKVSVDVSIIPYIIGPALNIGNEVQGNFKKNAF : 648
 1 5 YL Q 1T S AL 46Y35f N v a 6F W 66 DFT T DK6 D63 66PY6G ALN6GN 4 1F A

 LHA : IFSGAVILLEFIPEIAPVLGTFALVSYIA---NKVLTVQITIDNALSKRNEKWDVVKYIVTNWLAKVNTQIDLIRKKMKEALENOAETKAIINVOYQNTVEEKNIN--F : 764
 LHB : EIAGASTILLEFIPEIAPVLGTFALVSYID---NKKKIITIDNALTKRNEKWSMDYGLIVAQWLSTVNTQFYTIKEGMKALNYCAQALPEIKYRYNIYSEKESKNIN--I : 763
 LHC : ATGVGTFILLEFPEFTIPALGAFVIYSKVQ---EREKIITIDNCLBQRIKRWKDSYEWMTGTLSSRIITQFNINISQMYDSLSNYQAGAIKADILEYKKYSGSKENIKS-- : 759
 LHD : ATAGVAFLLLEGPEFTIPALGAFVTFYSSIQ---EREKIITIDNCLBQRIKRWKDSYEWMTGTLSSRIITQFNINISQMYDSLSNYQAGAIKADILEYKKYSGSKENIKS-- : 765
 LHE : ELLGAGILLEFPEPELLIPTILVFTIKSFLGSSDNKKNVKAINNALKERDEKWEVYSFIVSNMTKINTQFNKRKEQMYQALQNVNAIKTIIIESKYSNTYLEEKNELTNKY : 738
 LHF : ELLGAGILLEFPEPELLIPTILVFTIKSFLGSSDNKKNVKAINNALKERDEKWEVYSFIVSNMTKINTQFNKRKEQMYQALQNVNAIKTIIIESKYSNTYLEEKNELTNKY : 747
 LHG : EIGGAATIMEFIPELIVPIVGFPTLESYVG---NKGHIIMTISNALKKRDQKWTDMYGLIVSQWLSTVNTQFYTIKERMVYNALNQSQAIEKIIEDQYNRSEEDKMNIN--I : 756
 G iL6E PE 6P 6 F S 6 4 6 I N L R 4W Y 66 W 6 6 TQf My L Q a I Y Y3 K 6

800 * 820 * 840 * 860 * 880 * 900
 LHA : NIDDLSSKLNESINKAMININKFLNQC SVSYLMNSMIPYGVKRLDFDASLKDALLKYIYDNRGTLIGQVD-RLKDKVNNTLSTDIQFQSKYVDMQORLLSTFTTEYIK----- : 871
 LHB : DFNDINSKLDNGINQALDNINNFINGC SVSYLMNKLPLAVEKLDLDFDNTLAKNLLNIDENKLIILIGSAE--YEKSKVNKYLKTINPDLSTYDNTLKTIMFNLKYN----- : 870
 LHC : QVENKNSLDVKISEAMNNINKFIRECSVTYLFKNMPLKVIDELNEDFTKAKLINLIDSHNIIILVGEVD--KLKAVNNSFQNTIPFNISYTNNSLLKDIINEYFN----- : 866
 LHD : QVENKNSLDVKISEAMNNINKFIRECSVTYLFKNMPLKVIDELNEDFTKAKLINLIDSHNIIILVGEVD--KLKAVNNSFQNTIPFNISYTNNSLLKDIINEYFN----- : 872
 LHE : DIKQIENELNOKVSIAMNNIDRFLTESSISYLMKIINEVKINKLREYDENKTYLLNYIIQHGSILGSSQ--ELNSMVTDTLNNISIPFKLSYTDKILISVFNKFFK----- : 845
 LHF : NIYSIKELNOKVSIAMNNIDRFLTESSISYLMKLINEAKINKLSEYDKRVQVYLLNYIIENSTLGTSSVPELNNLVSNLNNISPFELSEYTDKILIIHLIRFYK----- : 855
 LHG : DFNDIDFKLANOSINLAINNIDF INQC SISYLMNRMIPLA VKLKD FDDNLKRDLEIYDTNELYLLDEVN--ILKSKVNRHLKXDSIPFDLSLYTKTILIQVFNYYIS----- : 863
 6 L1 6 A6 NI1 F6 S63YL 6 6 L 5D k L6 I L 1 V 6PF 6 Yt 1 L 5

 LHA : MEFFVNKQFNVKDPVNGVDIAVIKIPN--AGQMOPVKAFKIHKNKIWI PERDTFTNPEEGDLNPPPEAKQVPU--SYIDSTYLSYTDNEKDNLYKGVTKLPERIYSTDLGRMLLTSI : 111
 LHB : MPVTINNFYNDPIDNNNIIIMPEPPARGTKYKAFKIDTRIWI I PERYTFGYPKPE--DFNKS SGI FNRDVCBEYDDYLDNTDNDKKNIFLQTMIKLFRNRIKSKPLGEKLEMI : 112
 LHC : MPITINNFYNDPIDNNNIIIMPEPPARGTKYKAFKIDTRIWI I PERYTFGYPKPE--DFNKS SGI FNRDVCBEYDDYLDNTDNDKKNIFLQTMIKLFRNRIKSKPLGEKLEMI : 110
 LHD : MTWPVRDFNYPVNDNDILYRI PQNKLITTPVAFMI TQNIWI PERESSDTNPS--LSKPRPTSKYQ--SYDPSYLSYTDQKDTFLKIIKLFKRIINERDICKKLINYL : 110
 LHE : MPK--INSFNNDPNDRTILYIKP---GGQEFYFSNIMKNIIWIIPERNVIGTPPQ--DPHPPTSLKNGDS--SYDPSYLSYTDQKDTFLKIIKLFKRIINERDICKKLINYL : 107
 LHF : MPVINNFYNDPIDNNNIIIMPEPPARGTKYKAFKIDTRIWI I PERYTFGYPKPE--DFNKS SGI FNRDVCBEYDDYLDNTDNDKKNIFLQTMIKLFRNRIKSKPLGEKLEMI : 111
 LHG : MPVINNFYNDPIDNNNIIIMPEPPARGTKYKAFKIDTRIWI I PERYTFGYPKPE--DFNKS SGI FNRDVCBEYDDYLDNTDNDKKNIFLQTMIKLFRNRIKSKPLGEKLEMI : 112
 MP6T6N1FNYNDP6111116Y646P514G3N25YKAF4ITIN6W66PER535G3K2P2GDFNKP3S6TS4DSCSYDPIYL331E2K145LKT6IKGFNRINS4P6G24L6EE6

 LHA : VRGIPFWCG--STIDTELKVIDTN--CINVIQPD--GSYR--SEELNLI VIIGPSADIIQFECKSFGEV---LNLTRNGYGVSTQYIREFSPDFTFGFEESLEVDVTNPLLGAKFA : 214
 LHB : INGIPIYLGDRRVPLEEFNTNINIASVTNKLISNPGEVERKKGIFANLIIIFGPGVLNENETIDIGIQN----HFASREGFGGIMQMKFCPEYVSFVNNVQENKGA SIFNRRGYF : 221
 LHC : STDIPFGNNNTPIINTPFDVDFNSVDVKTQGNWVKTSINPVIITGPRENIIDPTSTFKLTNN---TFAAQEGFGALSIISIPRFLTYSNATNDVGEGRFSKSBFC : 220
 LHD : VVGSPFMGDSSTPDTDFTRHTTNIAVEKEFENGSKWVTNIIITPSVLIIFGPLNIDYTASLTILQGOQ---SNPSFEGFGTILKVAPEFLITFDVTSNQSSAVILGKSIFC : 220
 LHE : SKANPVLGNDTPDNQPHIGDAS--AVEIKFSN--GSQD---ILLPNVILIMGAEPDLFETNSNLSLNN---YMSNHRFGSIAITVTSPEYFRFN---DNC---MN---EFI : 203
 LHF : KNGKPYLGNDHTAVNEFCANNRSTSEIKESN--GTTD---SMLNVLILGPGPNLECSFTFPVRIFPNNAIYDPSEKFGSGIQLMSFSTEYEVAFN---DN---TD---LFI : 210
 LHG : VDAIPYLGNDHTAVNEFCANNRSTSEIKESN--GTTD---SMLNVLILGPGPNLECSFTFPVRIFPNNAIYDPSEKFGSGIQLMSFSTEYEVAFN---DN---TD---LFI : 210
 61GIP56G11STP612F1611A3T3616K2S1PG35D4TKS66PN66I6GPGP1662NE3S36462NNNIAY1P32EG5G36G664FSP565T5NIVT11K33S6614SE5I

 LHA : TDPAVTLAHELHAGHRLYGI A1--NPNRVFKVNTNAYYEMSGLEVSFEELRTFGCHDAKFIDSLQENEFRLYYNKKFKDIASTLN--K--AKSIVGTTASLOVMKNVFEKYLIS : 324
 LHB : SDPALILMHELHAGHRLYGIK--VDDLPIVNEKFFMQSTDAIQAEELYTFGQDPSTIITPSTDKSYKVLQNFPGIVDRNLKV--LVCISDPNININIKNFKDKYFV : 331
 LHC : MPDILILMHELHAGHRLYGI A1PNDQTISSVTNIFYSYVNVKLEVAETIYAFGGTIDILIPKARKYFEKALDYVRSIAKRLNSITITANPSSENVKYGKOKLIRKYRV : 333
 LHD : MDPVIALMHELHAGHRLYGI A1PNDQTISSVTNIFYSYVNVKLEVAETIYAFGGTIDILIPKARKYFEKALDYVRSIAKRLNSITITANPSSENVKYGKOKLIRKYRV : 333
 LHE : QDPALILMHELHAGHRLYGI A1PNDQTISSVTNIFYSYVNVKLEVAETIYAFGGTIDILIPKARKYFEKALDYVRSIAKRLNSITITANPSSENVKYGKOKLIRKYRV : 310
 LHF : ADPALILMHELHAGHRLYGI A1PNDQTISSVTNIFYSYVNVKLEVAETIYAFGGTIDILIPKARKYFEKALDYVRSIAKRLNSITITANPSSENVKYGKOKLIRKYRV : 319
 LHG : ADPALILMHELHAGHRLYGI A1PNDQTISSVTNIFYSYVNVKLEVAETIYAFGGTIDILIPKARKYFEKALDYVRSIAKRLNSITITANPSSENVKYGKOKLIRKYRV : 330
 ADPA63LMHELHAGHRLYGIKIP63146662N32G556QS341625EE6YTFGGQD61613SST1K26YEK6L1154DTAS4LMN6NKNVS6S13N6161YK165KEK1FD

LHA : EDTSGKFSVDKLFKDKLYKMLTEIYTEDNFVVFVKVLRNRYTLNFDKAVFKIN-IVPKVNTIYDGNLNRNTLAAFNFGONTEINNMFVKLNKFTGLPEFYKLLCVDGII- : 435
LHB : EDESGKSIDVESFDKLYKSLMFGFETNIAENYKIKTRASFSDSIPVVKIKNLLDNEIYTIIEGPNISDKMEKEBYRGONKAINQAYEIS-KEHLAVYKIQMCVDEEK- : 442
LHC : VESGVEYVNRNKFVLYNELTQIFTEFNAYAKIYNQNRKTHYLSNVTLPVFAN-ILDDNIYTIQNGFNIPKSNLNLVFMGQNLNRPALKNP--ENMLYLTFTKFCVDAID- : 442
LHD : KONTGFWVNIIDKFNLSYDLTNMSEVYSSQYNVKNRTHYLSNVTLPVFAN-ILDDNIYTIQNGFNIPKSNLNLVFMGQNLNRPALKNP--ENMLYLTFTKFCVDAID- : 443
LHE : KDAAGIYVNIINKNFNDIFKLYS-FTFEDLRTKQVCRQTYGQYK-FKLSNLNDSIYNISEGYNIN--NLKVNFRGQANLNPRIITPITGRITGLVVKIIRFCVD---- : 414
LHF : QDSNGYTVNISKFNATYKLPFS-FTFEDLRTKQVCRQTYGQYK-FKLSNLNDSIYNISEGYNIN--NLKVNFRGQANLNPRIITPITGRITGLVVKIIRFCVD---- : 423
LHG : EDPNGKYVDKDKFDKLYKMLMFGFETNLAGYGIKTRYSYFSEYLPPIKTEKLLDNTIYQNEGPNIAKMLKTEFGNQKAVNKEAYEIS-LEHLVYIRIAMCKP--V- : 439
2D33GK536161KFLK65K653ET16AKK5K6KNRK3Y6S25KPP646N16611N6Y3J2EG5N631K1646E5NQ6E6NP265KE63NKEG66E55646CVDI6E

LHA : -TSKTKSLIEGRNALALQC--IKVNNWDLFFSPSEDNFTNDLNGKBEITSPTNIEAAEENISLDTLQYVLTFFNFDNEPENISLENLSSDIIGQLMLPNIERF--PNGKRY : 543
LHB : -LYDDDDKDRWGS--SLQC--IDVDNEDLFFADKNSFSDLSKNERIEYNTQSVIENDFPIN--ELILDYDLSKIE-LPSENTESLTDFN-VDVPVYEKO--PAIKKI : 542
LHC : ------GRSLYN--KTLQCRELLAVNRNLDLFFDIQSDVKTIDIFLRKIDNEETEVYIYPDNVSD--QVILS-KNTSEHGQDLDLPSIDSESE--ILPGENQVYFDDITK : 539
LHD : KLYDDDDKDRWG--SSLC--IKVNNRNPVYADKDSISOBIFENKIITDETVNQVSDKFSLD--ESILDGQVPIPELVDPLLPVNMWEP--NLPGEEIVFYDDITK : 545
LHE : -IEGRKGIR-----KLQC--IEINNGELFFVASENSYNDNINTPKEIDTITSNNYENDLD--QVILNFNSESAPG-LSDEKLNLTIOND-AIYIPKYSNNGTSIDIEQH : 512
LHF : -IEGRKGIR-----KLQC--IKVNNRDLFFVASENSYNDNINTPKEIDTITSNNYENDLD--QVILNFNSESAPG-LSDEKLNLTIOND-AIYIPKYSNNGTSIDIEQH : 521
LHG : -MYKNTGK-----SEQC--ILVNNEDLFFTANKDSFKDLAKAETIAYNTONTNENFSD--OLILDNLDLSSGID-LPNEENTEPTNFDDIDIPVYIKO--SALKKI : 535
K6YD1KG44R5GNKAK3LQCREB6K61NEDLF56A1K1S53116FK3E4ITD1T263NYE11FS611LQ261L15163SEPEQ6SSENLN3132N1D6Y6PKYE4QGF5164KY

LHA : ELDKYTMFHYLRAQEFHEHGSRIALTNSVNEALLNPSRVYTFSSDYVVKVKNKATEAAMFLGWVQVLYDFDTESETSVSTTDKIADITIIIPYIGPALNIGNMLYKDDFVGAL : 656
LHB : FTDDNTIFQYLYSQTFPLDIRDISLTSSFDLALLFSNKVYSPFSDYIKTANKVVEAGLFGWVKQIVNDVIEANKSNMTDKIADISLIVPIYIGLALNVGNETAKGNFENAF : 655
LHC : NVDYLNYSYLESQKLSNVEDFTFRSIEEALDNSAKVYTYPP-TLANKVNAGVQGLFLMANDVVEDEFTTNILRKDTLTKISDVSAIIPYIGPALNIGNMLYKDDFVGAL : 651
LHD : YVDYLNYSYLESQKLSNVENITLTTSVEEALGYSNKIYTFELP-SLAEKVNKGQVAGLFLMANDVVEDEFTTNILRKDTLTKISDVSAIIPYIGPALNIGNMLYKDDFVGAL : 657
LHE : DVNELANVFFYLDQAQKPEGENNVNLTSSIDTALLDQKPIYTFPSSEFINNVNKPQVQALFVSIQVQVLYDFDTESETSVSTTDKIADISLIVPIYIGLALNVGNETAKGNFENAF : 625
LHF : TVDKLANVFFYLDQAQKPEGENNVNLTSSIDTALLDQKPIYTFPSSEFINNVNKPQVQALFVSIQVQVLYDFDTESETSVSTTDKIADISLIVPIYIGLALNVGNETAKGNFENAF : 634
LHG : FVDGDSLFEYLAHQK6PE16E163LT3S61EALL1SN46Y35FS3156NKNVKNRPV2AA6F6S6W622661DFT3EA32KST6DK6AD63666PY6GPPALN6GNE6Q4G1FK1AF : 648
561KLN655YLEAQK6PE16E163LT3S61EALL1SN46Y35FS3156NKNVKNRPV2AA6F6S6W622661DFT3EA32KST6DK6AD63666PY6GPPALN6GNE6Q4G1FK1AF

LHA : IFSGAVILLEFIPBIAIPVLGTFALVSYIA---NKVLTVQITDINALSKRNEKWEVYKIVITNWLAKVNTQIDLIRKMKKEALENQAEATKALINYQVNYTEEEKNNIN--F : 764
LHB : ETAGASILLEFIPBIAIPVLGTFALVSYIA---NKVLTVQITDINALSKRNEKWEVYKIVITNWLAKVNTQIDLIRKMKKEALENQAEATKALINYQVNYTEEEKNNIN--F : 763
LHC : AVTGVTILLEAFPEFTIPALGAFVYKSVQ---ERNEIKITINCLEQRIKRWKDSYEMMGWTLRSITQFNINYSQMYDSINYQAGAKAKIDLEYKYSKSDKENIKS-- : 759
LHD : ETAGAVILLEAFPEFTIPALGAFVYKSVQ---ERNEIKITINCLEQRIKRWKDSYEMMGWTLRSITQFNINYSQMYDSINYQAGAKAKIDLEYKYSKSDKENIKS-- : 765
LHE : ELGAGILLEAFPEFTIPALGAFVYKSVQ---ERNEIKITINCLEQRIKRWKDSYEMMGWTLRSITQFNINYSQMYDSINYQAGAKAKIDLEYKYSKSDKENIKS-- : 738
LHF : ELGAGILLEAFPEFTIPALGAFVYKSVQ---ERNEIKITINCLEQRIKRWKDSYEMMGWTLRSITQFNINYSQMYDSINYQAGAKAKIDLEYKYSKSDKENIKS-- : 747
LHG : EIGGAAIILMEFIPBIAIPVLGTFALVSYIA---NKVLTVQITDINALSKRNEKWEVYKIVITNWLAKVNTQIDLIRKMKKEALENQAEATKALINYQVNYTEEEKNNIN--F : 756
E6AGAAIILMEFIPBIAIPVLGTFALVSYIA---NKVLTVQITDINALSKRNEKWEVYKIVITNWLAKVNTQIDLIRKMKKEALENQAEATKALINYQVNYTEEEKNNIN--F

LHA :	NIDDLSSKLNESINKAMINININKFLAQC	SVSYLMN	SMPYGVKRL	EDF	ASLKO	ALLKYI	YDNR	GTL	IG	QVD	-RLKDKV	NN	TL	STD	IPFQ	L	SKV	VD	NQ	RL	LL	ST	F	T	E	Y	I	K	-	:	871																																																																												
LHB :	DFND	INS	K	IAN	EG	I	Q	A	I	D	N	I	N	N	F	I	N	G	S	V	S	Y	L	M	K	K	M	I	P	L	A	V	E	K	L	D	F	D	N	T	L	K	N	L	N	Y	I	D	E	N	K	D	Y	L	I	G	S	A	E	-	Y	E	K	S	V	N	K	Y	L	A	T	I	M	P	F	D	L	S	I	Y	T	N	D	T	I	L	I	E	M	F	N	K	Y	N	S	-	:	870									
LHC :	QVEN	L	K	N	S	L	D	V	K	I	S	E	A	M	N	I	N	K	F	I	R	E	C	S	V	T	Y	L	F	K	N	M	L	P	K	V	I	D	E	L	N	E	F	D	R	N	T	K	A	K	L	I	N	L	I	D	S	H	N	I	L	V	G	E	V	-	K	L	K	A	K	V	N	N	S	F	O	N	T	I	P	F	N	I	F	S	Y	T	N	N	S	L	K	D	I	I	N	E	Y	F	N	-	:	866			
LHD :	QVEN	L	K	N	S	L	D	V	K	I	S	E	A	M	N	I	N	K	F	I	R	E	C	S	V	T	Y	L	F	K	N	M	L	P	K	V	I	D	E	L	N	K	F	D	L	R	T	K	T	E	L	I	N	L	I	D	S	H	N	I	L	V	G	E	V	-	R	L	K	A	K	V	N	E	S	F	E	N	T	M	P	F	N	I	F	S	Y	T	N	N	S	L	K	D	I	I	N	E	Y	F	N	-	:	872			
LHE :	DI	K	O	I	E	N	E	L	N	Q	V	S	I	A	M	N	I	D	R	F	L	T	E	S	S	I	S	Y	L	M	K	I	N	E	V	K	I	N	K	L	R	E	Y	D	E	N	V	K	T	Y	L	L	N	I	I	O	H	G	S	I	L	G	E	S	Q	-	E	L	N	S	V	T	D	T	L	N	N	S	I	P	F	K	L	S	S	Y	T	D	D	K	I	L	I	S	F	N	K	F	K	-	:	845					
LHF :	NI	Y	S	I	K	E	E	L	N	K	V	S	L	A	M	Q	N	I	D	R	F	L	T	E	S	S	I	S	Y	L	M	K	L	I	N	E	A	K	I	N	K	L	S	E	Y	D	K	R	V	N	Q	Y	L	L	N	I	L	E	N	S	T	L	G	T	S	S	V	P	E	L	N	N	L	V	S	N	T	L	N	N	S	I	P	F	E	L	S	E	Y	T	N	D	K	I	L	I	H	I	L	I	R	F	Y	K	-	:	855
LHG :	DF	N	D	I	F	K	L	N	Q	S	I	N	L	A	I	N	N	I	D	D	F	I	N	Q	C	S	I	S	Y	L	M	N	R	M	I	P	L	A	V	K	K	L	D	F	D	N	L	K	R	D	L	L	E	Y	I	D	T	N	E	L	L	D	E	V	N	-	I	L	K	S	V	N	R	H	L	K	D	S	I	P	P	O	L	S	L	Y	T	K	T	I	L	I	Q	V	F	N	N	Y	I	S	-	:	863				
	16116	R	N	K	L	12	K	6	S	2	A	61	N	I	14	F	6	N	2	C	S	63	Y	L	M	K	4	6	6	P	6	A	61	4	L	4	E	5	D	4	N	6	K	4	1	L	6	N	Y	I	D	2	N	N	6	I	L	6	G	2	V	1	P	4	L	K	S	V	N	13	L	K	1	3	6	P	1	6	S	S	Y	T	11	3	6	L	6	2	6	F	N	E	5	5	K		:										

Figura 3

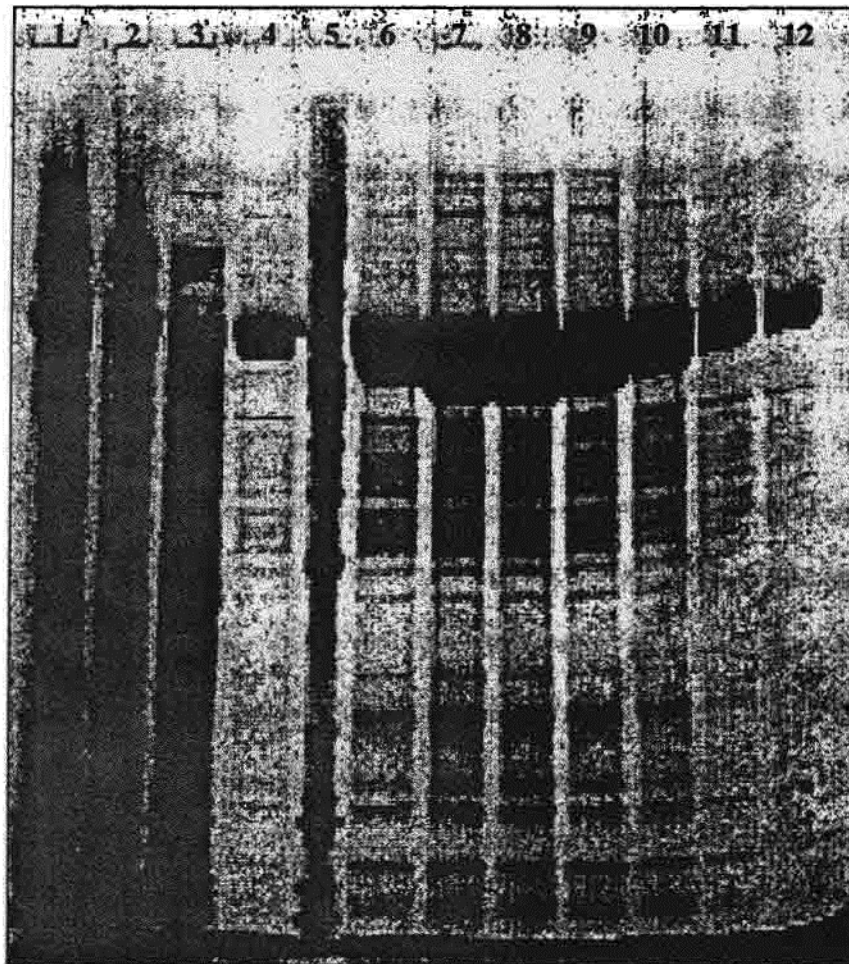


Figura 4

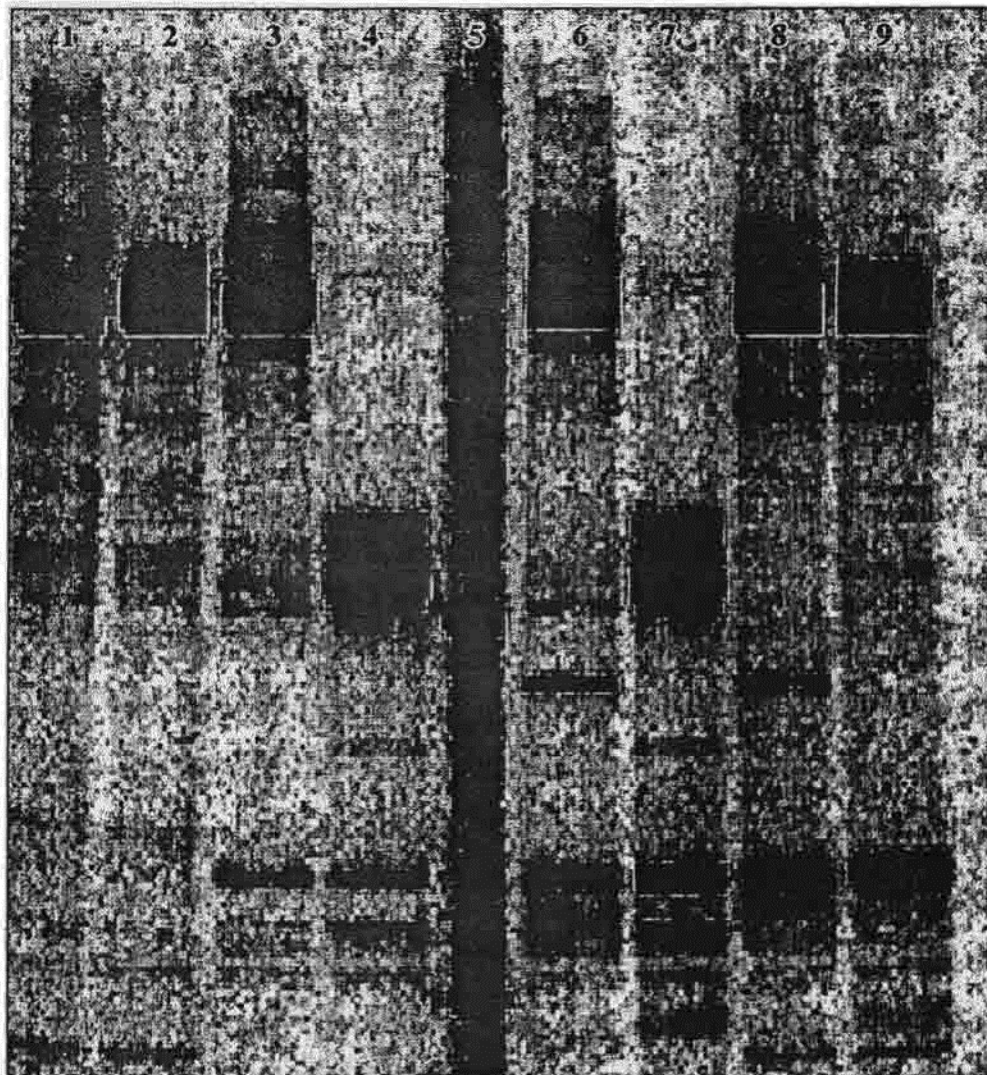


Figura 5

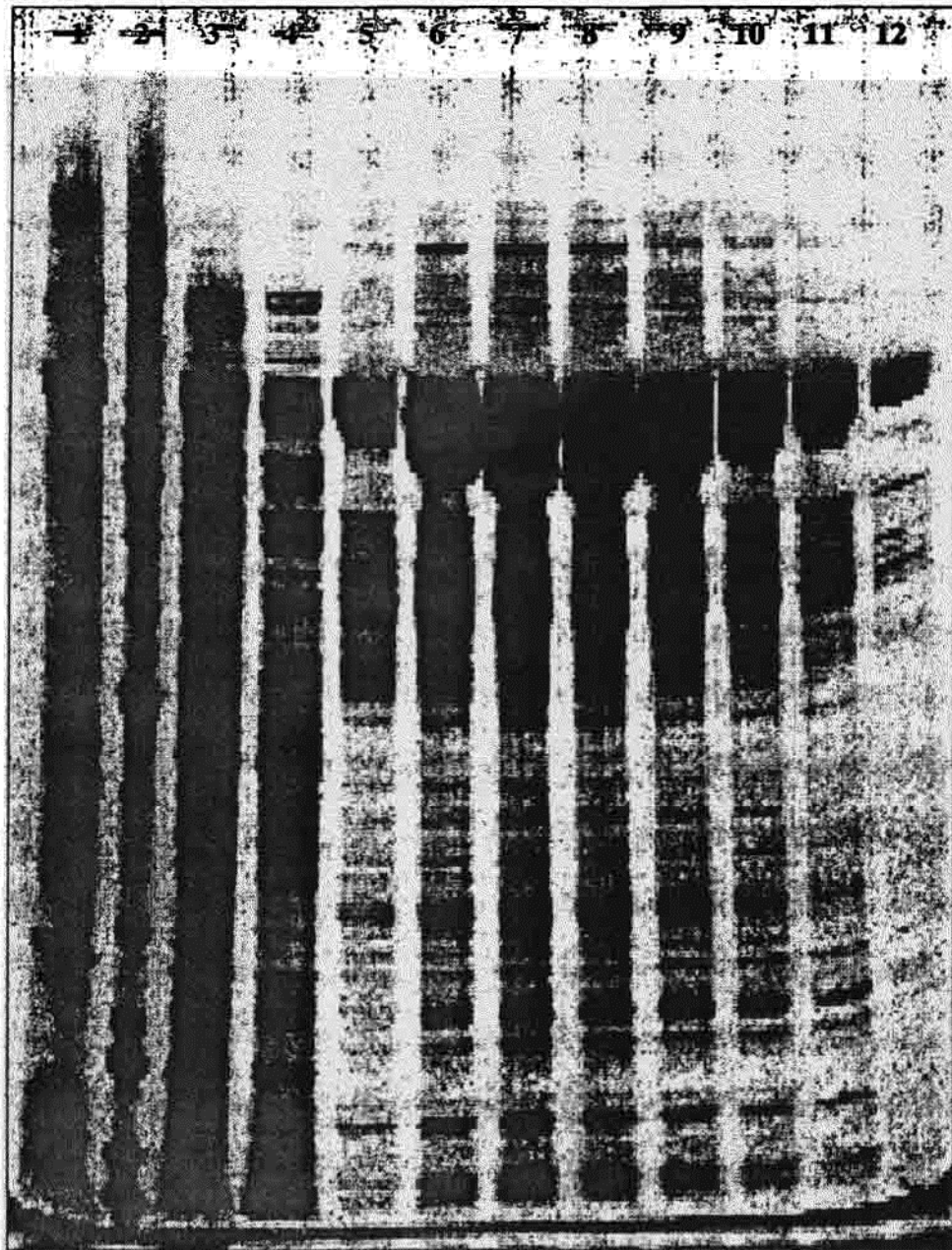


Figura 6

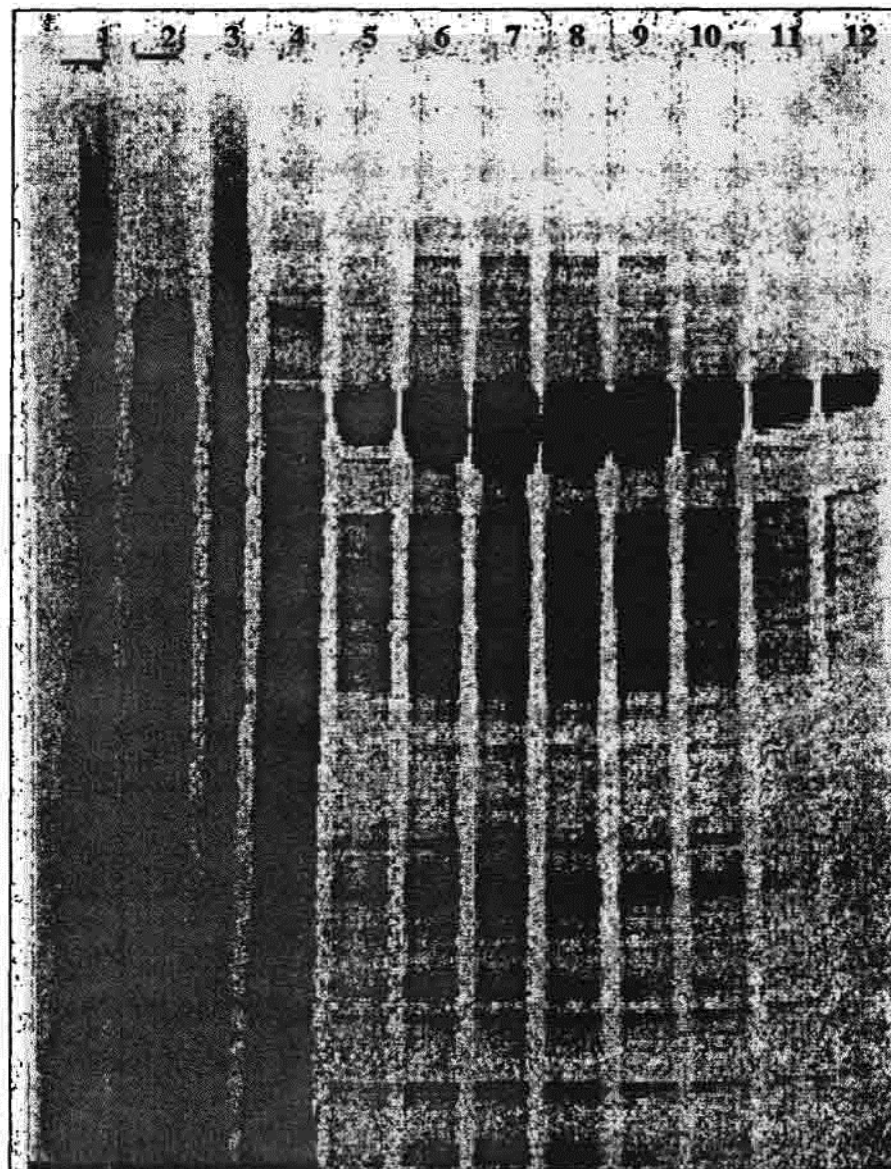


Figura 7

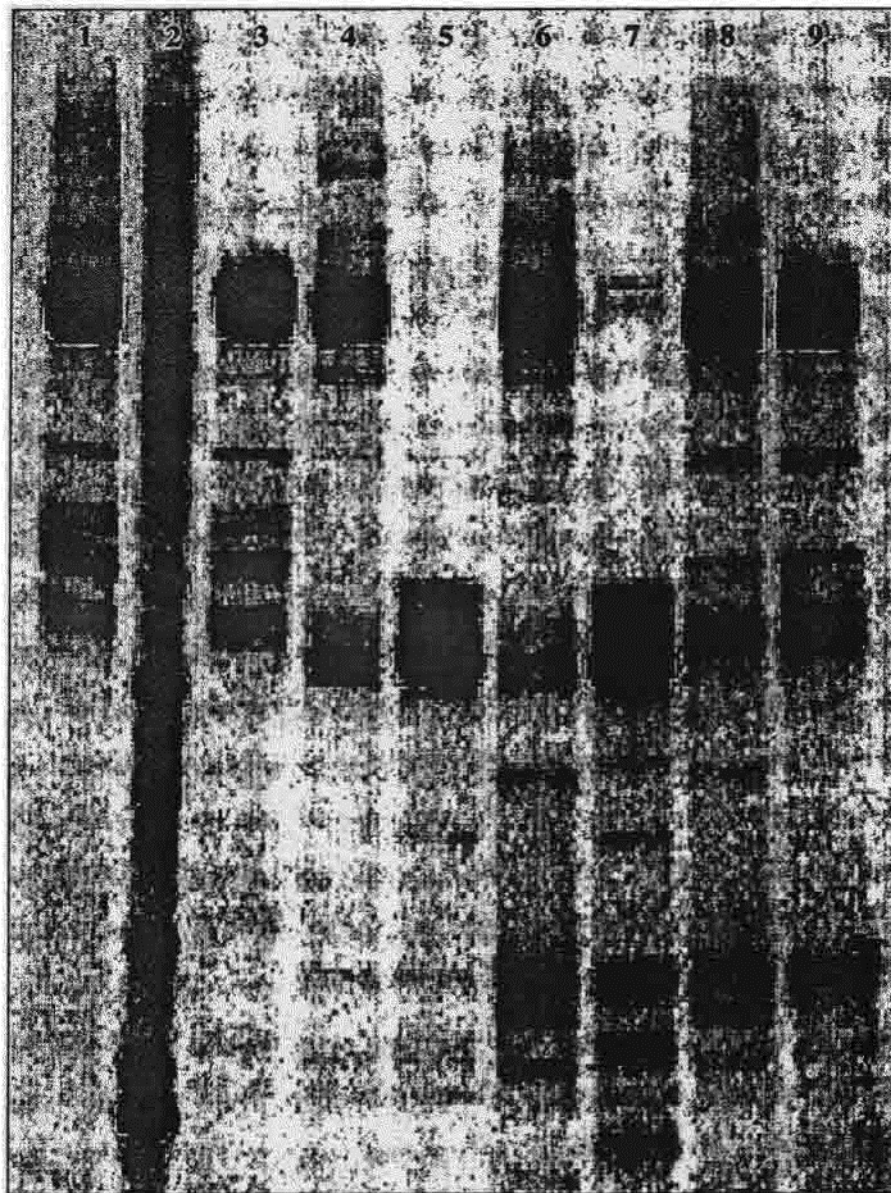


Figura 8

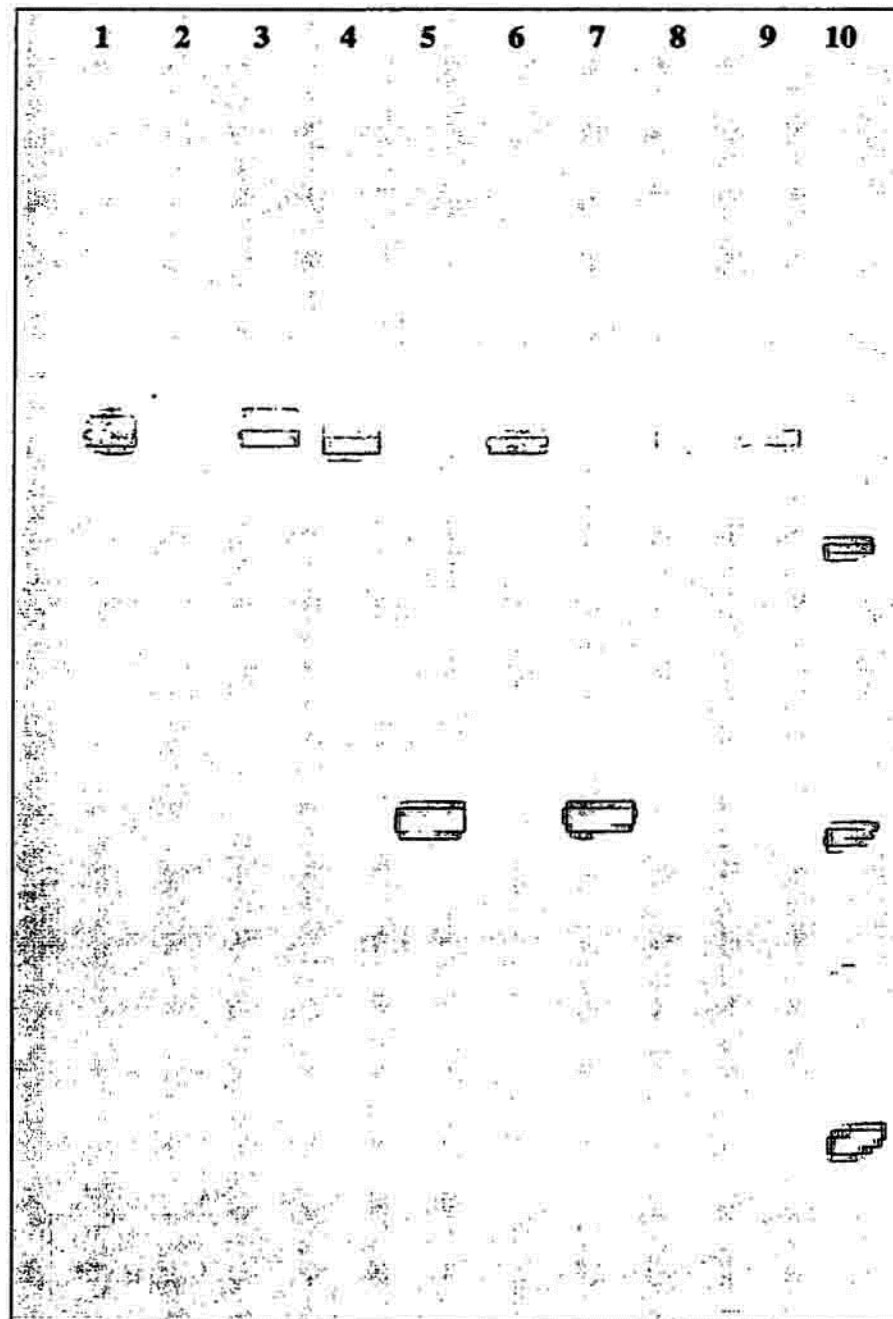


Figura 9

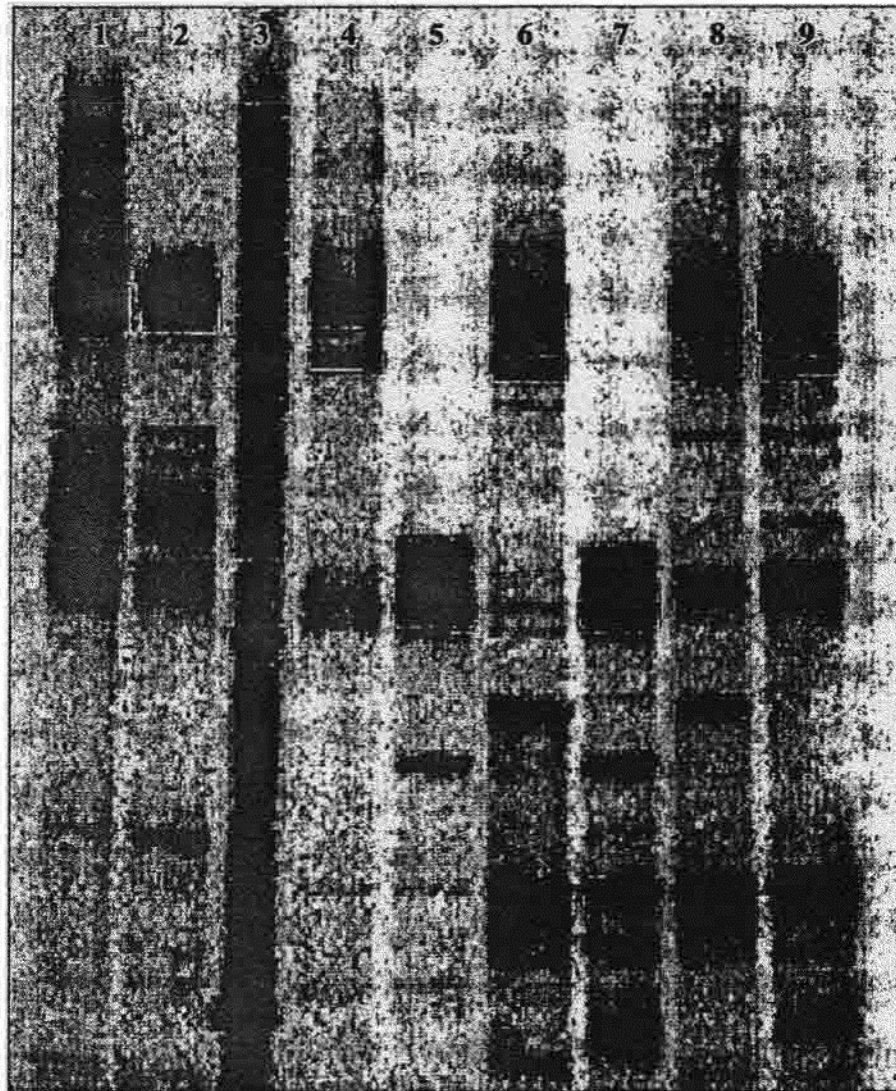


Figura 10

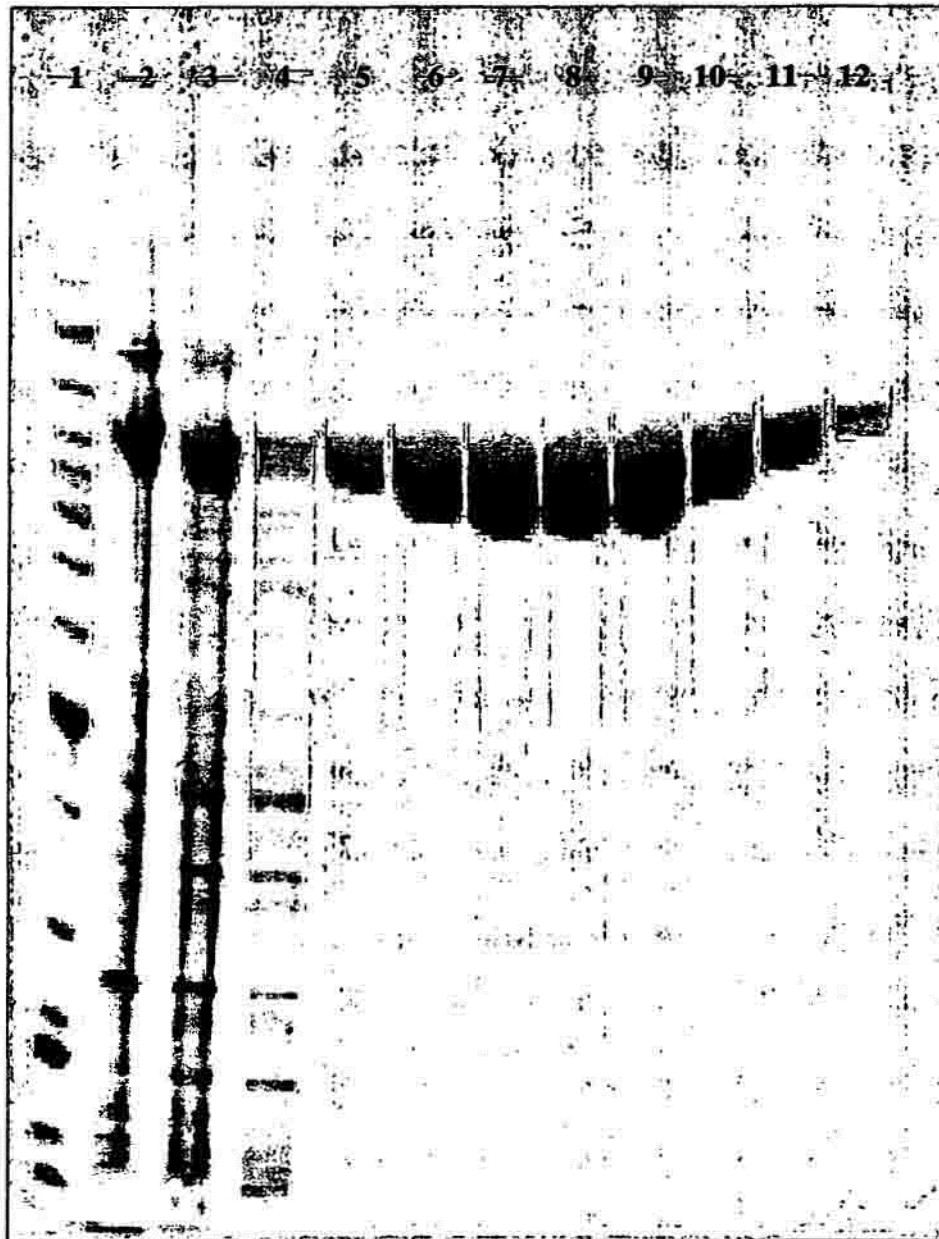


Figura 11

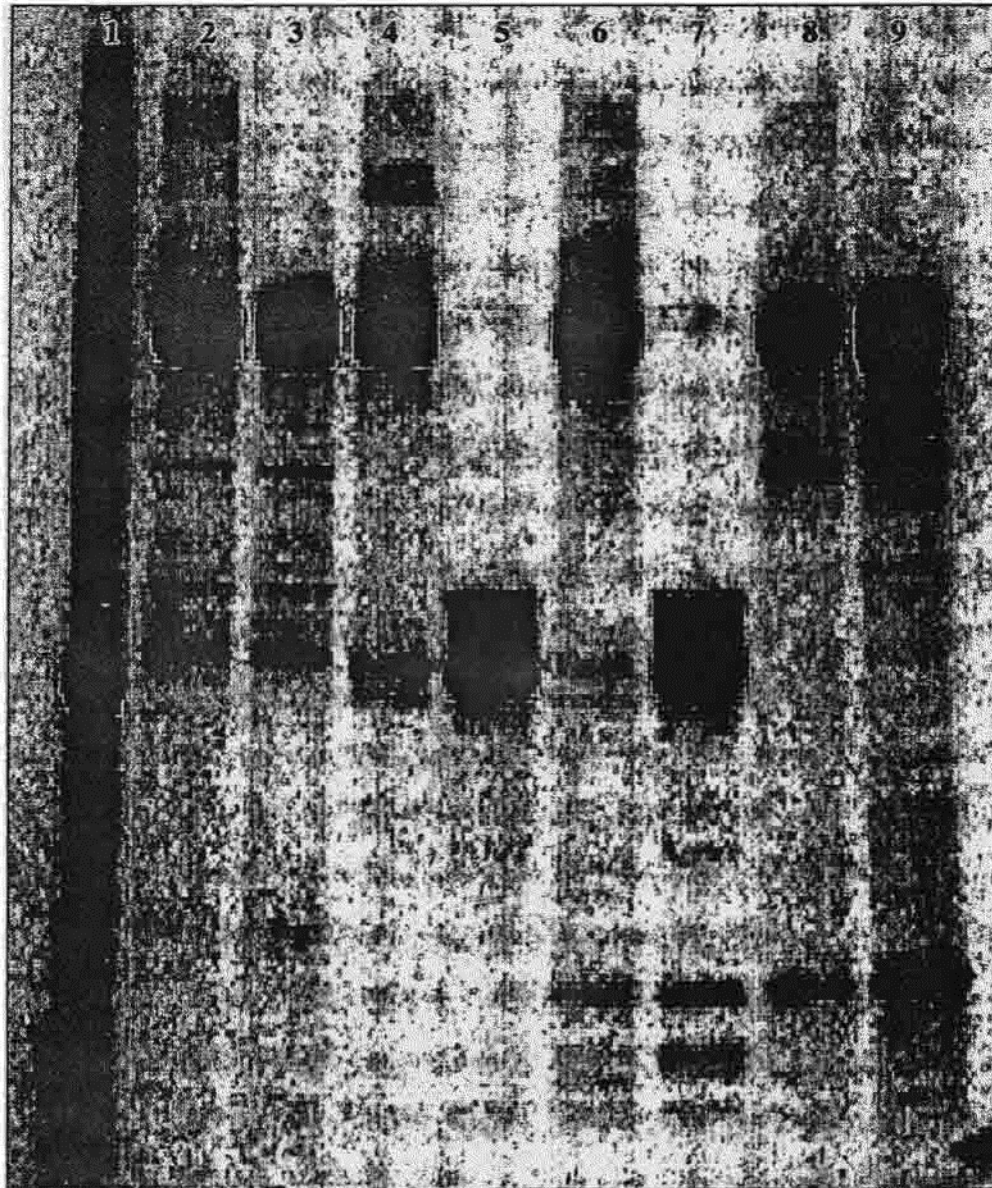


Figura 12

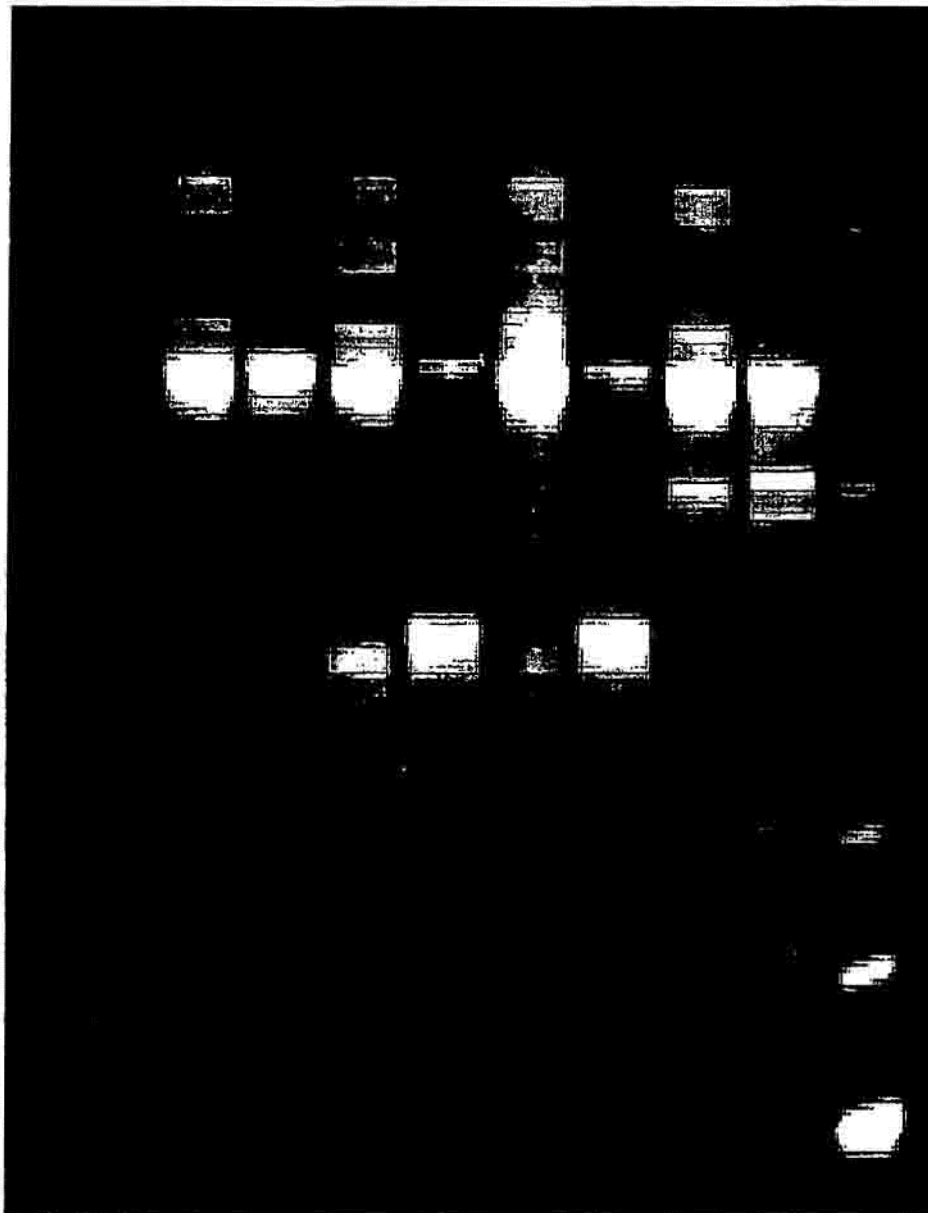


Figura 13

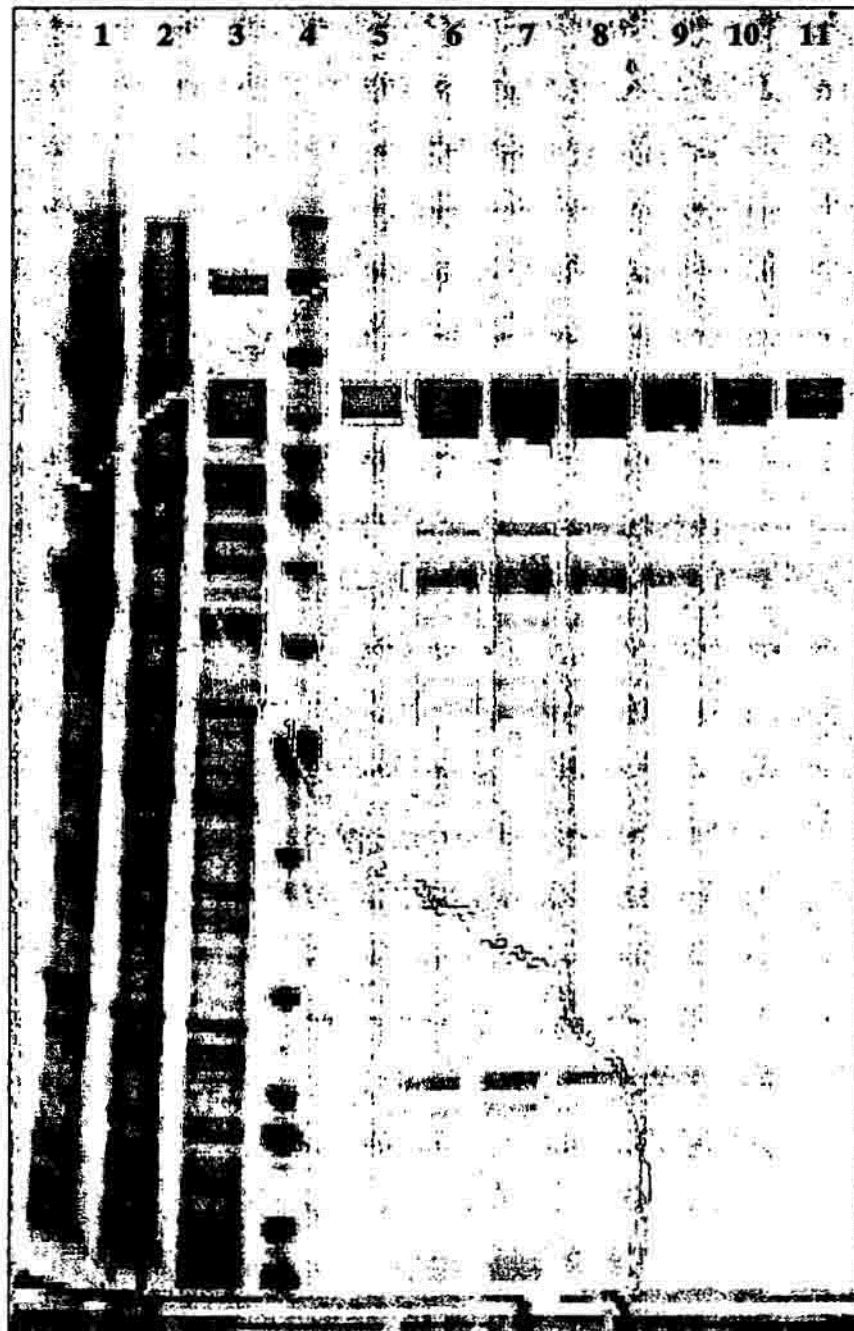


Figura 14

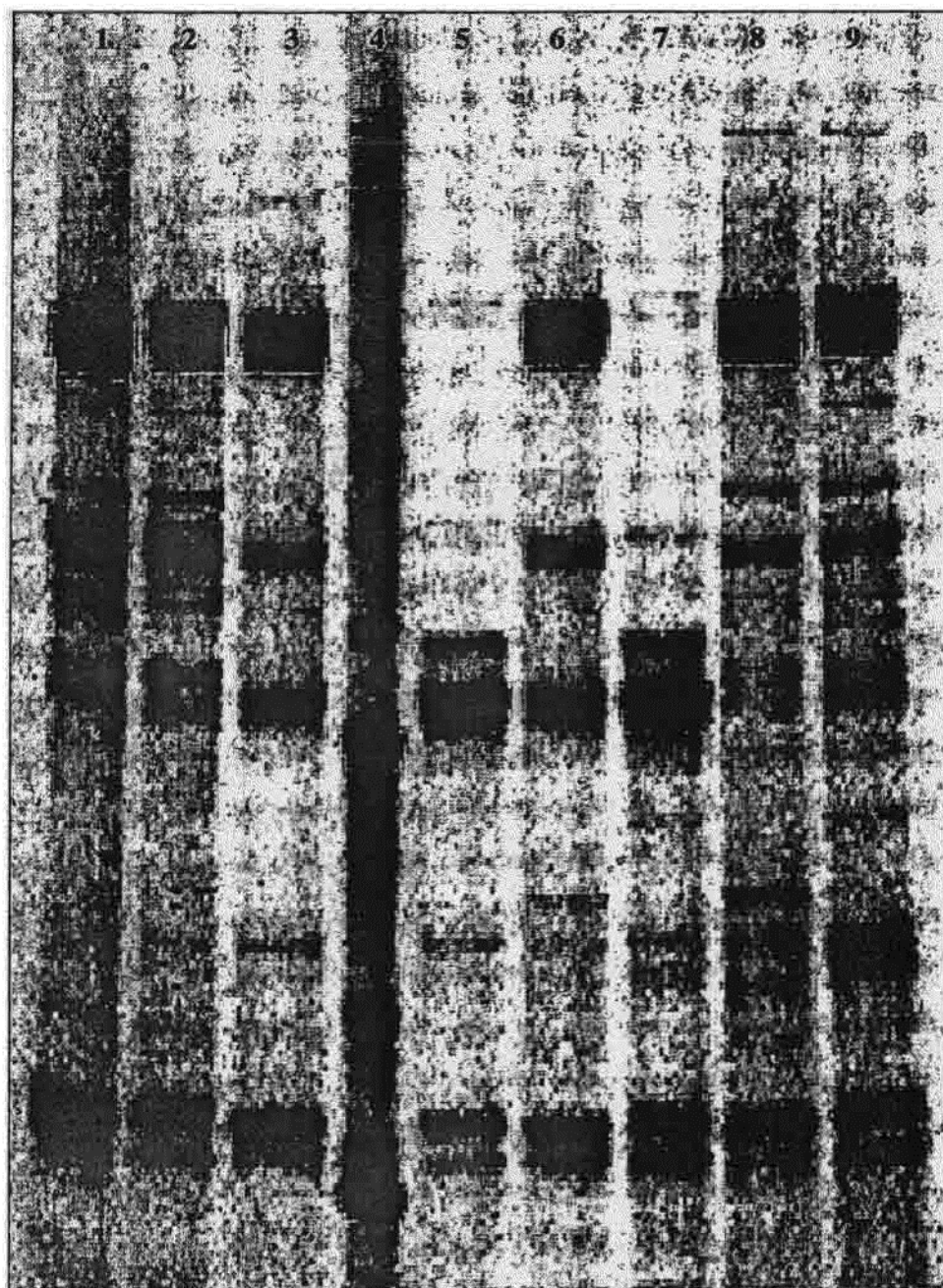


Figura 15

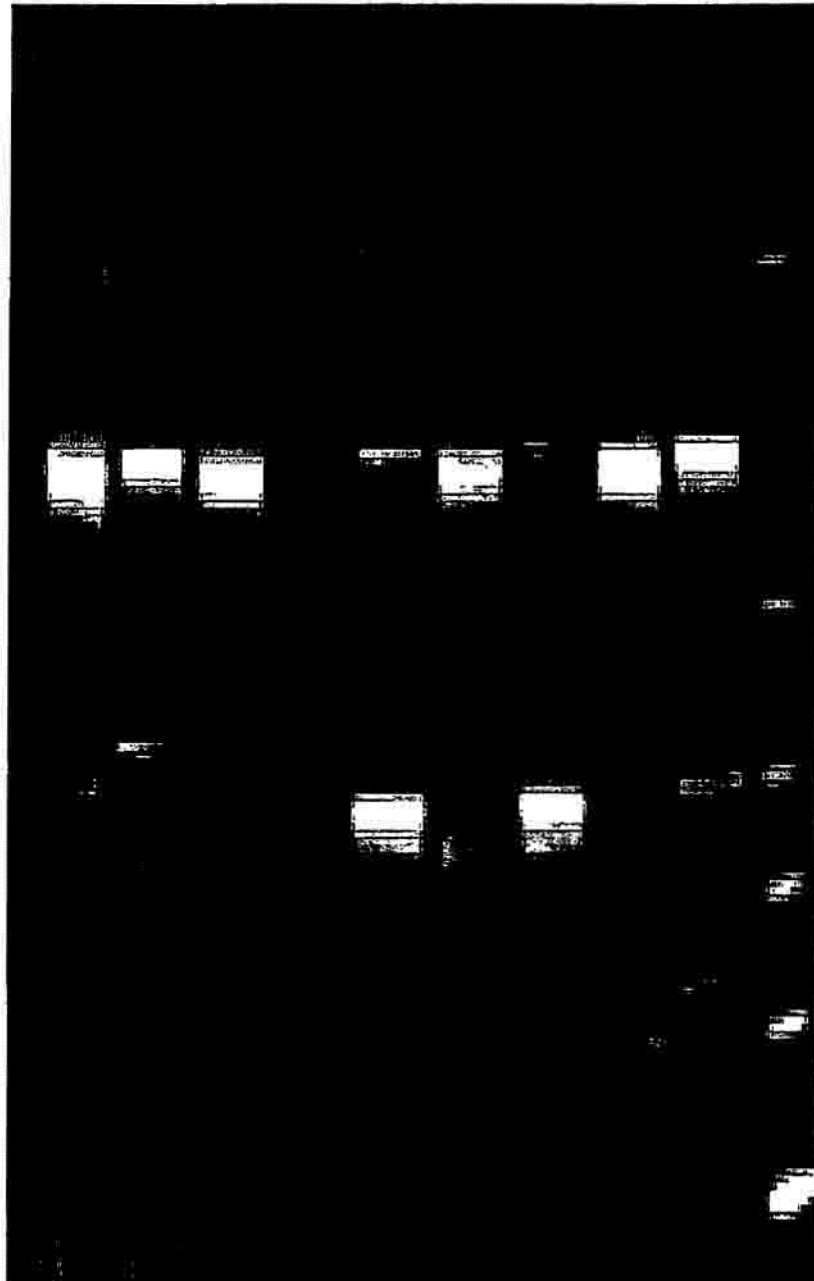


Figura 16

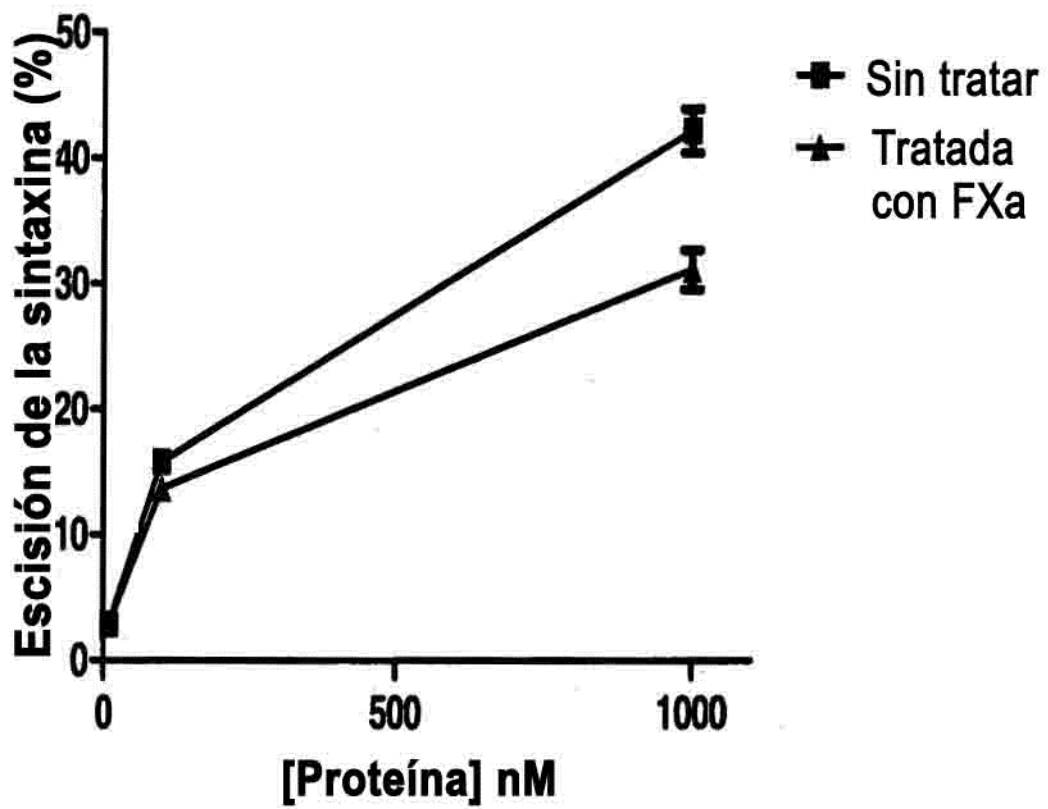


Figura 17

