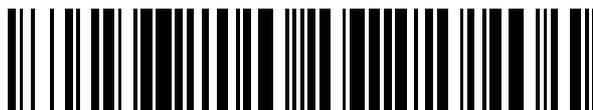


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 562 425**

51 Int. Cl.:

**C12N 15/62** (2006.01)

**A61K 47/48** (2006.01)

**C07K 7/23** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **16.05.2012 E 12722486 (3)**

97 Fecha y número de publicación de la concesión europea: **25.11.2015 EP 2710129**

54 Título: **Proteínas de fusión terapéuticas**

30 Prioridad:

**16.05.2011 GB 201108108**

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

**04.03.2016**

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 y  
HARPER, ELAINE**

74 Agente/Representante:

**DE ELZABURU MÁRQUEZ, Alberto**

**ES 2 562 425 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

## Proteínas de fusión terapéuticas

La presente invención se refiere a la construcción de una nueva clase de inhibidores de la secreción dirigidos ("Targeted Secretion Inhibitors", TSI), a un método para su activación y al producto activado.

5 La proteasas no citotóxicas son un grupo de proteasas bien reconocido que actúa sobre las células diana incapacitando la función celular. De modo importante, las proteasas no citotóxicas no matan a las células diana sobre las que actúan. Algunos de los ejemplos más conocidos de proteasas no citotóxicas incluyen las neurotoxinas clostridiales (por ejemplo, neurotoxina botulínica, que se comercializa con nombres tales como Dysport™, Neurobloc™, y Botox™) e IgA proteasas.

10 Las proteasas no citotóxicas actúan rompiendo proteolíticamente proteínas de transporte intracelular denominadas SNARE (por ejemplo, SNAP-25, VAMP, o sintaxia) (véase Gerald K (2002), "Cell and Molecular Biology" (4ª edición), John Wiley & Sons, Inc.). El acrónimo SNARE se deriva del inglés **S**oluble **N**SF **A**ttachment **R**eceptor (receptor de unión a NSF soluble), en donde NSF significa **N**-ethylmaleimide-**S**ensitive **F**actor (factor sensible a N-etilmaleimida). Las proteínas SNARE son esenciales para la formación de vesículas intracelulares y, por tanto, para la secreción de moléculas mediante el transporte de vesículas desde una célula. Por consiguiente, tras haber sido transportada a la

15 célula diana, la proteasa no citotóxica es capaz de inhibir la secreción celular de la célula diana.

Las proteasas no citotóxicas pueden emplearse en su forma nativa o sustancialmente nativa (es decir, como holotoxinas, tal como es el caso de Dysport™, Neurobloc™, y Botox™), en cuyo caso el transporte de las proteasas hasta los tipos celulares específicos se basa en (i) la administración localizada de la proteasa y/o (ii) la capacidad de

20 unión inherente de la proteasa nativa. Como alternativa, las proteasas no citotóxicas pueden emplearse en forma redirigida, en la que la proteasa nativa se modifica para incluir un ligando exógeno denominado resto de transporte dirigido (TM). El TM se selecciona para proporcionar especificidad de unión por una célula diana deseada y, como parte del proceso de redirección, puede eliminarse la porción de unión nativa de la proteasa no citotóxica.

El presente solicitante ha iniciado la exploración del concepto y el desarrollo de la tecnología de la redirección basada en neurotoxinas clostridiales, y las proteínas de fusión resultantes se denominan inhibidores de la secreción dirigidos ("Targeted Secretion Inhibitors", TSI).

25

La sustitución de TM puede realizarse mediante técnicas de conjugación química convencionales, que son muy conocidas por los expertos en la técnica. A este respecto, se hace referencia a Hermanson, G.T. (1996), Bioconjugate techniques, Academic Press; y a Wong, S.S. (1991), Chemistry of protein conjugation and cross-linking, CRC Press.

30

Sin embargo la conjugación química a menudo no es precisa. Por ejemplo, después de la conjugación, un TM puede haberse unido al resto del conjugado en más de un sitio de unión. La conjugación química también es difícil de controlar. Por ejemplo, un TM puede haberse unido al resto de la toxina modificada en un sitio de unión sobre el componente de proteasa y/o sobre el componente de translocación. Esto resulta problemático cuando solo se desea uno de dichos componentes (preferiblemente en un único sitio) por la eficacia terapéutica. Por tanto, la conjugación química produce una población mixta de moléculas de toxina modificadas, lo cual no resulta deseable.

35

Como alternativa a la conjugación química, la sustitución del TM puede realizarse mediante la preparación recombinante de una proteína de fusión de polipéptido monocatenario. La preparación de dicha molécula se describe en el documento WO98/07864. Sin embargo, los presentes inventores han descubierto que la metodología del documento WO98/07864 no resulta adecuada para todos los tipos de TM.

40

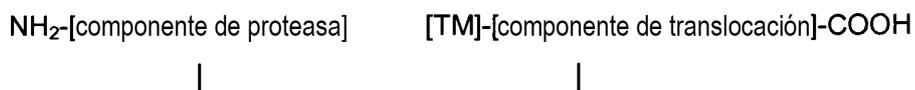
Un sistema alternativo al del documento WO98/07864 se describe en el documento WO2006/059093. Según el documento WO2006/059093, el TM se presenta centralmente (CP) dentro de la proteína de fusión monocatenaria, entre el componente de proteasa no citotóxica y el componente de dominio de translocación. Esto produce una proteína de fusión monocatenaria que tiene la siguiente estructura:

45 NH<sub>2</sub> - [componente de proteasa] - [TM] - [componente de translocación] - COOH

Las proteínas de fusión descritas anteriormente se activan mediante un tratamiento con una proteasa, que efectúa la ruptura en un sitio localizado en el C-terminal del componente de proteasa. Este proceso de activación produce una proteína dicatenaria que comprende el componente de proteasa unido covalentemente (a través de un enlace disulfuro) al componente de translocación de la proteína de fusión. En el caso del documento WO2006/059093, la molécula dicatenaria resultante tiene un TM que está unido mediante un péptido a través de su C-terminal al N-terminal del componente de dominio de translocación. Por consiguiente, la porción N-terminal del TM queda libre para actuar y unirse a un receptor deseado. Esta disposición es importante para las clases de TM que requieren un N-terminal libre o una porción N-terminal libre para unirse a su receptor.

50

Como ejemplo, después de la activación proteolítica, el documento WO2006/059093 proporciona polipéptidos que tienen la siguiente conformación dicatenaria:



5 En dicha conformación dicatenaria, los componentes de TM y de translocación se presentan en forma de una proteína de fusión monocatenaria, en la que el C-terminal de TM está unido mediante un péptido al N-terminal del componente de translocación.

10 Los presentes inventores han descubierto que los sistemas descritos en los documentos WO98/07864 y WO2006/059093 no resultan óptimos para la presentación de todos los tipos de TM y, como tales, puede dar como resultado la producción de proteínas de fusión que tengan una capacidad de unión no deseable/reducida por la célula diana prevista.

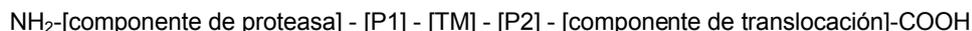
Por tanto, es necesario un sistema alternativo o mejorado para construir TSI.

La presente invención resuelve uno o más de los problemas mencionados anteriormente proporcionando una proteína de fusión de polipéptido monocatenario, según se indica en las reivindicaciones de la patente.

15 El sistema descrito en el documento WO2006/059093 proporciona TSI que tienen un TM con un N-terminal que está libre para interactuar con un sitio de unión sobre una célula diana. Sin embargo, los presentes inventores han descubierto que el sistema descrito en el documento WO2006/059093 no es adecuado para los TM que requieren un dominio N-terminal libre y un dominio C-terminal libre para interactuar con un sitio de unión sobre una célula diana.

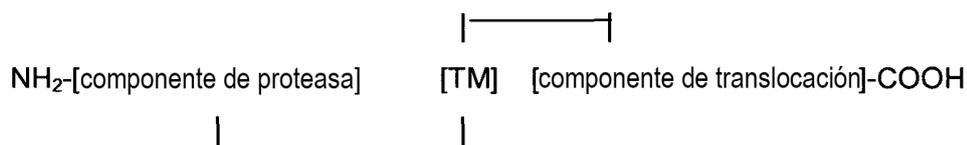
20 Así, en contraste con el documento WO2006/059093, la presente invención proporciona un sistema para proporcionar TSI en el que el componente de TM tiene un dominio N-terminal libre y un dominio C-terminal libre.

En una realización, la presente invención proporciona una proteína de fusión monocatenaria que tiene la siguiente orientación N-terminal a C-terminal, en la que P1 y P2 representan el primer y el segundo sitio de ruptura de proteasas:

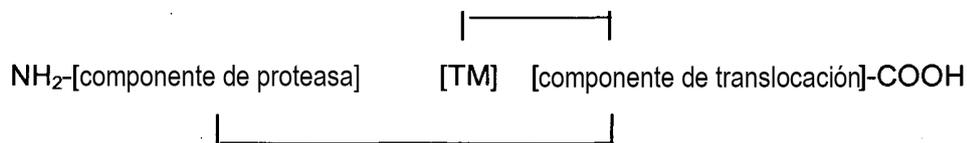


25 Después de la ruptura en el primer y segundo sitio de ruptura, dicha proteína de fusión monocatenaria adopta la siguiente estructura tricatenaria en la que los componentes de TM y de translocación están unidos covalentemente entre sí, y en la que:

A) el componente de proteasa está unido covalentemente al componente de TM:



30 o B) el componente de proteasa está unido covalentemente al componente de translocación:

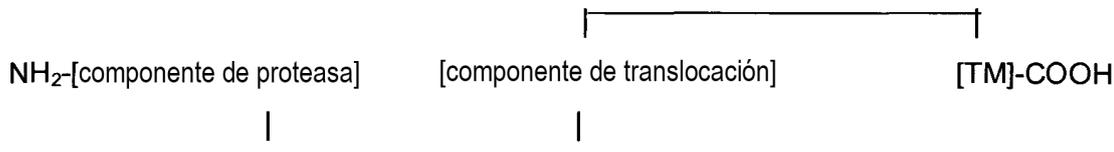


En otra realización, la presente invención proporciona una proteína de fusión monocatenaria que tiene la siguiente orientación N-terminal a C-terminal, en la que P1 y P2 representan el primer y el segundo sitio de ruptura de proteasas:

35  $\text{NH}_2\text{-[componente de proteasa] - [P1] - [componente de translocación] - [P2] - [TM]\text{-COOH}}$

Después de la ruptura en el primer y segundo sitio de ruptura, dicha proteína de fusión monocatenaria adopta la siguiente estructura tricatenaria, en la que los componentes de TM y de translocación están unidos covalentemente entre sí, y en la que:

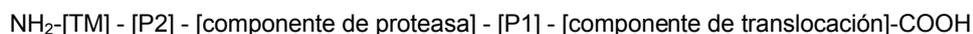
A) el componente de proteasa está unido covalentemente al componente de translocación:



o B) el componente de proteasa está unido covalentemente al componente de TM:

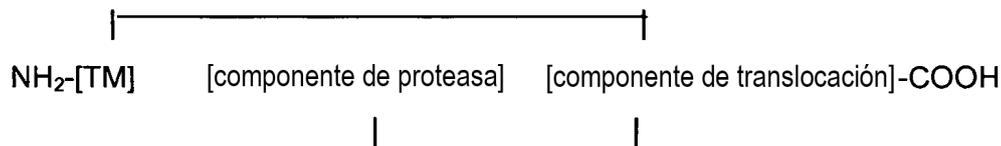


5 En otra realización, la presente invención proporciona una proteína de fusión monocatenaria que tiene la siguiente orientación N-terminal a C-terminal, en la que P1 y P2 representan el primer y el segundo sitio de ruptura de proteasas:

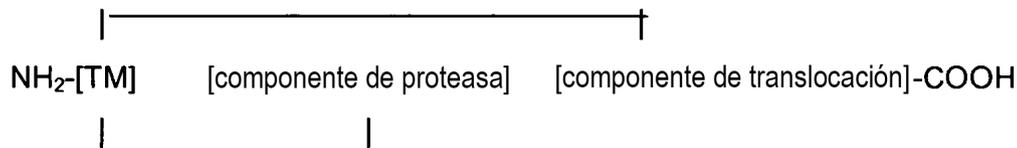


10 Después de la ruptura en el primer y segundo sitio de ruptura, dicha proteína de fusión monocatenaria adopta la siguiente estructura tricatenaria, en la que los componentes de TM y de translocación están unidos covalentemente entre sí, y en la que:

A) el componente de proteasa está unido covalentemente al componente de translocación:



o B) el componente de proteasa está unido covalentemente al componente de TM:



15 Durante el uso, un polipéptido TSI de la presente invención se une a una célula diana, y la unión es facilitada por el TM. El componente de dominio de translocación entonces realiza el transporte del componente de proteasa no citotóxica hacia el citosol de la célula diana. Una vez dentro, el componente de proteasa no citotóxica inhibe el proceso de fusión exocítica de la célula diana. Por tanto, mediante la inactivación del aparato de fusión exocítica de la célula diana, el polipéptido de la presente invención inhibe la secreción desde la célula. Por consiguiente, los polipéptidos TSI de la presente invención pueden emplearse para suprimir o tratar una diversidad de síntomas trastornos patofisiológicos que están relacionados con la secreción celular.

**La proteasa no citotóxica**

25 El componente biológicamente activo de los polipéptidos TSI de la presente invención es una proteasa no citotóxica. Así, tras haber sido transportada al citosol de una célula diana, el componente de proteasa no citotóxica realiza la ruptura SNARE dentro de la célula diana deseada. Puesto que las proteínas SNARE son un componente fundamental del proceso secretor dentro de las células diana de mamífero, su inactivación proteolítica inhibe/suprime la secreción desde dichas células diana.

30 Las proteasas no citotóxicas son una clase discreta de moléculas que no mata a las células; en lugar de ello, actúan inhibiendo procesos celulares distintos de la síntesis de proteínas. Las proteasas no citotóxicas son producidas por una diversidad de organismos superiores (por ejemplo, plantas y animales), siendo un ejemplo de organismo superior el escorpión brasileño. Además, las proteasas no citotóxicas son producidas por una diversidad de microorganismos, de forma notable bacterias, tales como *Clostridium sp.* y *Neisseria sp.*

35 La neurotoxinas clostridiales representan un grupo importante de moléculas de toxinas no citotóxicas y comprenden dos cadenas polipeptídicas unidas entre sí mediante un enlace disulfuro. Las dos cadenas se denominan la cadena

pesada (cadena H), que tiene una masa molecular de aproximadamente 100 kDa, y la cadena ligera (cadena L), que tiene una masa molecular de aproximadamente 50 kDa. La cadena L es la que posee una función proteasa y muestra una alta especificidad de sustrato por las proteínas asociadas a membranas de vesículas y/o plasmáticas (SNARE) implicadas en el proceso exocítico (por ejemplo, sinaptobrevina, syntaxina, SNAP y/o VAMP). Estos sustratos son componentes importantes de la maquinaria secretora de la célula.

*Neisseria sp.*, y de forma más notable la especie *N. gonorrhoeae*, producen moléculas de toxinas no citotóxicas similares desde el punto de vista funcional. Un ejemplo de una de estas proteasas no citotóxicas es la IgA proteasa (véase el documento WO99/58571). Los estreptococos producen IgA proteasas similares, como por ejemplo *Streptococcus pneumoniae*.

Así, en una realización, la proteasa no citotóxica de la presente invención puede ser una proteasa de neurotoxina clostridial o una IgA proteasa (véase, por ejemplo, el documento WO 99/032272). Otro ejemplo de proteasas no citotóxicas es la proteasa de veneno de escorpión, tal como del veneno del escorpión brasileño *Tityus serrulatus*, o la proteasa antareasa (véase, por ejemplo, el documento WO 2011/022357).

### El resto de transporte dirigido (TM)

El componente de TM de la presente invención es responsable de la unión del polipéptido de la presente invención al sitio de unión sobre una célula diana. Así, el componente de TM es un ligando a través del cual el polipéptido de la presente invención se une a una célula diana seleccionada.

En el contexto de la presente invención, la célula diana puede ser cualquier célula de mamífero (preferiblemente un ser humano). Así, el TM puede unirse a una célula no neuronal y/o a una célula neuronal.

El componente de TM de los polipéptidos de la presente invención tiene una porción N-terminal libre y una porción C-terminal libre. Así, en una realización, el TM es capaz de interactuar con el sitio de unión (por ejemplo, un receptor o aceptor) sobre una célula diana a través de la interacción entre una porción N-terminal del resto de transporte dirigido y un dominio del sitio de unión. En otra realización, el TM es capaz de producir una interacción dual, en la que una porción N-terminal del resto de transporte dirigido interacciona con un dominio del sitio de unión, y una porción C-terminal del resto de transporte dirigido interacciona con un dominio de un sitio de unión. En esta última realización, las porciones N- y C-terminales del TM pueden unirse al mismo dominio o a dominios diferentes de un sitio de unión y/o pueden unirse a dominios sobre sitios de unión diferentes.

Los TM adecuados para su uso en los polipéptidos de la presente invención incluyen citoquinas, factores del crecimiento, neuropéptidos, lectinas, y anticuerpos (este término incluye anticuerpos monoclonales, andamiajes de unión a proteínas, fragmentos de anticuerpos, tales como Fab, F(ab)<sub>2</sub>, Fv, ScFv, y anticuerpos monocatenarios, tales como de camélidos, etc.).

En una realización, el componente de TM comprende o consiste en un ligando peptídico (por ejemplo, una hormona peptídica) que se une a un receptor presente sobre una célula diana. En una realización, el ligando peptídico presenta una secuencia de aminoácidos de 5-200 restos aminoácidos consecutivos. Como ejemplo, dicho ligando peptídico consiste o comprende una secuencia de aminoácidos de 5-150 o 5-100 o 5-50 o 5-40 o 5-30 o 5-25 o 5-20 o 7-12 o aproximadamente 10 restos aminoácidos consecutivos.

El componente de TM comprende una porción N-terminal y una porción C-terminal. Cada una de dichas porciones generalmente comprende al menos 5, al menos 10, al menos 15, al menos 20, o al menos 25 restos aminoácidos consecutivos.

En una realización, el TM comprende o consiste en un ligando peptídico (o un análogo de este) que se une a un receptor seleccionado de MRGPRX<sub>1</sub> (por ejemplo, un receptor del péptido de la médula adrenal bovina (BAM)), un receptor de péptidos opioides, OPRM<sub>1</sub> u OPRD<sub>1</sub> (por ejemplo, un receptor de péptidos de beta-endorfina), BDKRB<sub>1</sub> o BDKRB<sub>2</sub> (por ejemplo, un receptor de péptidos de bradiquinina), OPRM<sub>1</sub> u OPRD<sub>1</sub> (por ejemplo, un receptor de péptidos de met- o leu-encefalina), OPRK<sub>1</sub> (por ejemplo, un receptor de péptidos de dinorfina), GALR<sub>1</sub>, GALR<sub>2</sub> o GALR<sub>3</sub> (por ejemplo, un receptor de péptidos de galanina), OPRL<sub>1</sub> (por ejemplo, un receptor de péptidos de nociceptina), y TACR<sub>1</sub>, TACR<sub>2</sub> o TACR<sub>3</sub> (por ejemplo, un receptor de péptidos de la sustancia P).

En una realización, el TM comprende o consiste en un ligando peptídico (o un análogo de este) seleccionado de un péptido de la médula adrenal bovina (BAM), un péptido opioide, un péptido de beta-endorfina, un péptido de bradiquinina, un péptido de met- o leu-encefalina, un péptido de dinorfina, un péptido de galanina, un péptido de nociceptina, y un péptido de la sustancia P.

En una realización, el TM comprende o consiste en un péptido de la hormona liberadora de gonadotropina (GnRH). La GnRH es una hormona peptídica de 10 aminoácidos. Los aminoácidos N-terminales de la GnRH desempeñan un papel en la activación del receptor, mientras que los aminoácidos C-terminales son necesarios para la unión de alta afinidad al receptor de GnRH (véase Flanagan, Millar e Illing (1997), Reviews of Reproduction, 2, 113-120. La

- función de GnRH *in vivo* es actuar sobre los receptores de GnRH localizados en la glándula pituitaria anterior y estimular la síntesis y liberación de gonadotropinas, tales como la hormona luteinizante (LH) y la hormona estimulante del folículo (FSH). La referencia a los péptidos de GnRH incluye todos sus fragmentos de unión funcionales, variantes y análogos. Como ejemplo, la expresión "péptido de GnRH" incluye un péptido de GnRH en el cual se ha insertado un aminoácido cisteína (flanqueado por dos restos aminoácidos aquirales, tales como glicina y/o alanina) como aminoácido de sustitución para la posición 6 del péptido de GnRH. La GnRH también se denomina hormona liberadora de la hormona luteinizante (LHRH). Otros ejemplos incluyen péptidos de GnRH I, péptidos de GnRH II y péptidos de GnRH III, así como el polipéptido precursor de GnRH de 92 aminoácidos de longitud completa y sus truncamientos.
- 5
- 10 En una realización, el TM comprende o consiste en un péptido del factor liberador de corticotrofina (CRF). El CRF es una hormona peptídica hipotalámica de 41 aminoácidos que interacciona con los receptores CRF<sub>1</sub> y CRF<sub>2</sub>. La principal función de CRF *in vivo* es estimular la liberación de ACTH desde los corticotropos dentro del lóbulo anterior de la pituitaria. La referencia a un péptido de CRF incluye el CRF de longitud completa, urocortina 1 y urocortina 2, así como sus fragmentos de unión funcionales, variantes y análogos.
- 15 En una realización, el TM comprende o consiste en un péptido liberador de gastrina (GRP). La GRP es una hormona peptídica de 27 aminoácidos. El GRP regula numerosas funciones de los sistemas gastrointestinal y nervioso central, que incluyen la liberación de hormonas gastrointestinales, la contracción de las células del músculo liso, y la proliferación de células epiteliales, y es un potente mitógeno para tejidos neoplásicos. La referencia a los péptidos de GRP incluye todos sus fragmentos de unión funcionales, variantes y análogos.
- 20 En una realización, el TM comprende o consiste en una neuromedina B. La neuromedina B es una hormona peptídica de 10 aminoácidos. La función de la neuromedina B es actuar sobre los receptores BB<sub>1</sub> *in vivo* y es un potente mitógeno y factor del crecimiento para el pulmón normal y neoplásico y para el tejido epitelial gastrointestinal. La referencia a los péptidos de neuromedina B incluye todos sus fragmentos de unión funcionales, variantes y análogos. La referencia a la neuromedina B incluye el péptido homólogo humano, la bombesina, e incluye la bombesina de longitud completa (un péptido de 14 aminoácidos aislado originariamente de la piel de una rana), así como sus truncamientos y análogos peptídicos.
- 25
- 30 En una realización, el TM comprende o consiste en gastrina o colecistoquinina (CCK). La gastrina es una hormona peptídica de 17 aminoácidos, y la CCK es una hormona peptídica de 8 aminoácidos. Tanto la gastrina como la colecistoquinina actúan sobre los receptores CCK1 y CCK2 *in vivo* principalmente dentro del sistema gastrointestinal y SNC para modular la secreción de enzimas pancreáticas y la contracción del músculo liso de la vesícula biliar y el estómago, la ansiedad, la analgesia, la excitación sexual y la actividad neuroléptica. La referencia a los péptidos de gastrina y colecistoquinina incluye todos sus fragmentos de unión funcionales, variantes y análogos.
- 35 En una realización, un TM comprende o consiste en un péptido de somatostatina (SST). Los ejemplos de TM de péptidos de SST adecuados incluyen SST de longitud completa y cortistatina (CST), así como sus truncamientos y análogos peptídicos, tales como BIM 23052, BIM 23056 o BIM23268; péptidos de octreótido, péptidos de lanreótido, BIM23027, CYN154806, BIM23027, péptidos de vapreótido, péptidos de seglitida, y SOM230. Estos TM se unen a los receptores sst, tales como los receptores sst<sub>1</sub>, sst<sub>2</sub>, sst<sub>3</sub>, sst<sub>4</sub> y sst<sub>5</sub>. SST y CST presentan una alta homología estructural, y se unen a todos los receptores sst conocidos. La referencia a los péptidos de SST o CST incluye todos sus fragmentos de unión funcionales, variantes y análogos.
- 40 En una realización, un TM comprende o consiste en un péptido de la hormona liberadora de la hormona del crecimiento (GHRH). La GHRH también denomina factor liberador de la hormona del crecimiento (GRF o GHRF) o somatocrina. Los péptidos de GHRH adecuados incluyen el péptido de GHRH de longitud completa (1-44), y sus truncamientos, tales como GHRH (1-27, 1-28, 1-29), GHRH (1-37), y GHRH (1-40,1-43)-OH, así como análogos peptídicos, tales como BIM 28011 o NC-9-96. La referencia a los péptidos de GHRH incluye todos sus fragmentos de unión funcionales, variantes y análogos.
- 45
- 50 En una realización, un TM comprende o consiste en un péptido del receptor activado por proteínasa (PAR), por ejemplo, PAR1. Los péptidos de PAR representan un subtipo exclusivo de receptores acoplados al receptor de proteína G 7-transmembrana porque están proteolíticamente modificados para exponer un nuevo N-terminal extracelular, que actúa como un ligando activador sujeto. Se han identificado agonistas de PAR1 (tales como TFLLR) que activan su receptor cognado. La referencia a los péptidos de PAR incluye todos sus fragmentos de unión funcionales, variantes y análogos.
- 55 En una realización, un TM comprende o consiste en una hormona paratiroidea (PTH). La PTH es un péptido que es liberado por la glándula paratiroides y se une al receptor PTH-1. Este receptor tiene una amplia distribución, pero es particularmente abundante en tejidos diana de PTH, predominantemente el riñón y en el hueso. La referencia a los péptidos de PTH incluye todos sus fragmentos de unión funcionales, variantes y análogos.
- En una realización, un TM comprende o consiste en un péptido que se une a una célula secretora de moco, o a una célula neuronal que controla o dirige la secreción de moco. Por ejemplo, el TM se une a (a) células que segregan mucinas, tales como células caliciformes epiteliales y células secretoras de moco de las glándulas submucósicas, (b)

células que segregan componentes acuosos del moco, tales como células Clara y células serosas y/o (c) células que controlan o dirigen la secreción de moco, tales como fibras C "sensoriales-eferentes" o fibras del sistema neural NANC. Se mencionan en concreto los siguientes TM peptídicos: VIP; agonistas del beta<sub>2</sub> adrenerreceptor; péptido liberador de gastrina; y péptido relacionado con el gen de calcitonina. La referencia a estos TM peptídicos incluye todos sus fragmentos de unión funcionales, variantes y análogos. Así, los TSI según esta realización tienen aplicación terapéutica para tratar la hipersecreción de moco, el asma y/o la enfermedad pulmonar obstructiva crónica.

En otra realización, el TM comprende o consiste en un péptido que se une a una célula endocrina. Se mencionan en concreto GHRH; hormona estimulante del tiroides (TSH); insulina, factor del crecimiento insulínico; hormona liberadora de TSH (protirelina); hormona liberadora de FSH/LH (gonadorelina); hormona liberadora de corticotrofina (CRH); y ACTH. La referencia a estos TM peptídicos incluye todos sus fragmentos de unión funcionales, variantes y análogos. Así, los TSI según esta realización tienen una aplicación terapéutica para tratar: la neoplasia endocrina, que incluye MEN; la tirotoxicosis y otras enfermedades dependientes de hipersecreciones de la tiroides; la acromegalia, la hiperprolactinemia, la enfermedad de Cushing y otras enfermedades dependientes de la hipersecreción de la pituitaria anterior; el hiperandrogenismo, la anovulación crónica y otras enfermedades asociadas con el síndrome del ovario poliquístico.

En otra realización, el TM comprende o consiste en un péptido que se une a una célula inflamatoria. Se mencionan en concreto los TM peptídicos (i) para células cebadas, tales como el dominio C4 del Fc IgE; (ii) para eosinófilos, tales como ligandos para el receptor del complemento C3a/C4a-R, antígenos reactivos hacia el receptor del complemento CR4; (iii) para macrófagos y monocitos, tales como el factor estimulante de macrófagos; (iv) para neutrófilos, tales como un antígeno asociado con el receptor del complemento iC3b, o IL8. La referencia a estos TM peptídicos incluye todos sus fragmentos de unión funcionales, variantes y análogos. Así, los TSI según esta realización tienen aplicación terapéutica para tratar alergias (rinitis alérgica estacional (fiebre del heno), conjuntivitis alérgica, rinitis vasomotora y alergia alimentaria), eosinofilia, asma, artritis reumatoide, lupus eritematoso sistémico, lupus eritematoso discoide, colitis ulcerosa, enfermedad de Crohn, hemorroides, prurito, glomerulonefritis, hepatitis, pancreatitis, gastritis, vasculitis, miocarditis, psoriasis, eccema, fibrosis crónica inducida por radiación, cicatrices pulmonares y otros trastornos fibróticos.

En otra realización, el TM comprende o consiste en un péptido que se une a una célula exocrina. Se menciona en concreto el péptido activador de la adenil ciclasa de la pituitaria (PACAP-38). La referencia a estos TM peptídicos incluye todos sus fragmentos de unión funcionales, variantes y análogos. Así, los TSI según esta realización tienen aplicación terapéutica para tratar la hipersecreción de moco de células secretoras de moco localizadas en el tracto alimentario, en particular localizadas en el colon.

En otra realización, el TM comprende o consiste en un péptido que se une a una célula inmunológica. Se mencionan en concreto ligandos, tales como la característica de superficie/fragmento del virus de Epstein Barr. La referencia a estos TM peptídicos incluye todos sus fragmentos de unión funcionales, variantes y análogos. Así, los TSI según esta realización tienen aplicación terapéutica para tratar la miastenia grave, artritis reumatoide, lupus eritematoso sistémico, lupus eritematoso discoide, trasplante de órganos, trasplante de tejidos, trasplante de fluidos, enfermedad de Graves, tirotoxicosis, diabetes autoinmunitaria, anemia hemolítica, púrpura trombocitopénica, neutropenia, hepatitis autoinmunitaria crónica, gastritis autoinmunitaria, anemia perniciosa, tiroiditis de Hashimoto, enfermedad de Addison, síndrome de Sjogren, cirrosis biliar primaria, polimiositis, escleroderma, esclerosis sistémica, pénfigo vulgar, pénfigoide ampuloso, miocarditis, carditis reumática, glomerulonefritis (tipo de Goodpasture), uveitis, orquitis, colitis ulcerosa, vasculitis, gastritis atrófica, anemia perniciosa, diabetes mellitus de tipo 1.

En otra realización, el TM comprende o consiste en un péptido que se une a una célula cardiovascular. Se mencionan en concreto la trombina y TRAP (péptido agonista del receptor de trombina), y los ligandos que se unen a células endoteliales cardiovasculares, tales como anticuerpos que reconocen el antígeno de superficie GP1 b. La referencia a estos TM peptídicos incluye todos sus fragmentos de unión funcionales, variantes y análogos. Así, los TSI según esta realización tienen aplicación terapéutica para tratar trastornos cardiovasculares y/o hipertensión

En otra realización, el TM comprende o consiste en un péptido que se une a una célula ósea. Se mencionan en la presente ligandos que se unen a los osteoblastos para el tratamiento de una enfermedad seleccionada de la osteopetrosis y la osteoporosis, que incluyen la calcitonina, y ligandos que se unen a osteoclastos, que incluyen factores de diferenciación de osteoclastos (por ejemplo, TRANCE, o RANKL u OPGL). La referencia a estos TM peptídicos incluye todos sus fragmentos de unión funcionales, variantes y análogos. Así, los TSI según esta realización tienen una aplicación terapéutica para tratar trastornos óseos.

Las secuencias de unión a integrina lineales y cíclicas representan otro grupo de TM peptídicos de la presente invención. Muchas integrinas reconocen la secuencia peptídica triple Arg-Gly-Asp (RGD) (Ruoslahti, 1996). El motivo RGD se encuentra en más de 100 proteínas, que incluyen fibronectina, tenascina, fibrinógeno y vitronectina. La interacción de RGD-integrina es aprovechada como un mecanismo conservado de entrada en la célula por muchos patógenos, que incluyen coxsackievirus (Roivaninen *et al.*, 1991) y adenovirus (Mathias *et al.*, 1994). Las secuencias peptídicas lineal y cíclica PLAEIDGIEL y CPLAEIDGIELC, respectivamente, han demostrado unirse e internalizar

ADN en células que expresan  $\alpha_9\beta_1$  integrina (Schneider *et al.*, 1999). La referencia a estos TM peptídicos incluye todos sus fragmentos de unión funcionales, variantes y análogos.

Otros TM de la presente invención incluyen los descubiertos mediante técnicas de presentación de fagos, en particular aquellos que se dirigen y son internalizados por células epiteliales de las vías respiratorias humanas. Estos incluyen los péptidos THALWHT lineal y cíclico (Jost *et al.*, 2001); LEBP-1 (QPFMQCLCLIIDASC), LEBP-2 (RNVPIFNDVYWIAF) y LEBP-3 (VFRVRPWYQSTSQS) (Wu *et al.*, 2003); CDSAFVTDWGRSMLC (Florea *et al.*, 2003); SERSMNF, YGLPHKF, PSGAARA, LPHKSMP, LQHSMP (Writer *et al.*, 2004); FLSKPP, HSQLST y STQAMFQ (Rahim *et al.*, 2003). La referencia a estos TM peptídicos incluye todos sus fragmentos de unión funcionales, variantes y análogos.

5 En una realización, el TM comprende o consiste en una primera y una segunda porción (por ejemplo, dominios). En una realización, la primera y la segunda porción del resto de transporte dirigido pueden derivarse del mismo ligando (por ejemplo, cualquiera de los ligandos de TM identificados anteriormente). La primera y la segunda porción pueden unirse al mismo sitio o a sitios diferentes sobre el mismo receptor. Como alternativa, la primera y la segunda porción pueden unirse a sitios sobre receptores diferentes.

15 La primera y la segunda porción del resto de transporte dirigido pueden derivarse de ligandos diferentes (por ejemplo, cualquiera de los ligandos de TM identificados anteriormente), que pueden unirse al mismo receptor o a receptores diferentes. Por consiguiente, el TM puede ser un híbrido de dos TM. La primera y la segunda porción pueden unirse al mismo sitio o a sitios diferentes sobre el mismo receptor. Como alternativa, la primera y la segunda porción pueden unirse a sitios sobre receptores diferentes.

20 El TM puede incluir también una tercera y/o posterior porción procedente de otros TM (por ejemplo, cualquiera de los ligandos de TM identificados anteriormente).

En una realización, la primera porción (por ejemplo, dominio) comprende o consiste en un ligando que se une a través de una porción N-terminal libre (por ejemplo un N-terminal libre) a su receptor diana. Un ejemplo de este ligando es un ligando que se une a un receptor de opioides (por ejemplo, un ligando que se une a un receptor ORL<sub>1</sub>, tal como un péptido opioide). Otros ejemplos de péptidos opioides incluyen nociceptina, dinorfina, beta-endorfina, y encefalina. Otros ligandos de TM peptídicos no opioides incluyen los péptidos de BAM, galanina, sustancia P, GnRH, CRF, GRP, neuromedina B, bombesina, gastrina, CCK, SST, CST, y GHRH (así como sus truncamientos, variantes y análogos).

30 En otra realización (o la misma), la segunda porción (por ejemplo, dominio) comprende o consiste en un ligando que se une a través de una porción C-terminal libre (por ejemplo, un C-terminal libre) a su receptor diana. Un ejemplo de este ligando es un ligando que se une a un receptor de bradiquinina (por ejemplo, un péptido de bradiquinina) o a un receptor de la sustancia P (por ejemplo, un péptido de la sustancia P). Otros ligandos de TM peptídicos incluyen los péptidos de BAM, galanina, sustancia P, GnRH, CRF, GRP, neuromedina B, bombesina, gastrina, CCK, SST, CST, y GHRH (así como sus truncamientos, variantes y análogos).

35 Como ejemplo, el TM híbrido incluye una primera porción que comprende o consiste en un péptido de nociceptina, y una segunda porción que comprende o consiste en un péptido de bradiquinina (o un péptido de la sustancia P). En otros ejemplos, la primera porción comprende o consiste en un péptido de nociceptina, y la segunda porción comprende o consiste en un péptido seleccionado de un péptido de BAM, un péptido opioide, un péptido de beta-endorfina, un péptido de bradiquinina, un péptido de encefalina, un péptido de dinorfina, un péptido de galanina, y un péptido de la sustancia P.

40 En otro ejemplo, el TM híbrido incluye una primera porción que comprende o consiste en un péptido de dinorfina, y una segunda porción que comprende o consiste en un péptido de bradiquinina (o un péptido de la sustancia P). En otros ejemplos, la primera porción comprende o consiste en un péptido de dinorfina, y la segunda porción comprende o consiste en un péptido seleccionado de un péptido de BAM, un péptido opioide, un péptido de beta-endorfina, un péptido de bradiquinina, un péptido de encefalina, un péptido de nociceptina, un péptido de galanina, y un péptido de la sustancia P.

45 En otro ejemplo, el TM híbrido incluye una primera porción que comprende o consiste en un péptido de galanina, y una segunda porción que comprende o consiste en un péptido de bradiquinina (o un péptido de la sustancia P). En otros ejemplos, la primera porción comprende o consiste en un péptido de galanina, y la segunda porción comprende o consiste en un péptido seleccionado de un péptido de BAM, un péptido opioide, un péptido de beta-endorfina, un péptido de bradiquinina, un péptido de encefalina, un péptido de nociceptina, un péptido de dinorfina, y un péptido de la sustancia P.

50 En otro ejemplo, el TM híbrido incluye una primera porción que comprende o consiste en un péptido de BAM, y una segunda porción que comprende o consiste en un péptido de bradiquinina (o un péptido de la sustancia P). En otros ejemplos, la primera porción comprende o consiste en un péptido de BAM, y la segunda porción comprende o consiste en un péptido seleccionado de un péptido opioide, un péptido de beta-endorfina, un péptido de bradiquinina, un péptido de encefalina, un péptido de nociceptina, un péptido de dinorfina, un péptido de galanina, y un péptido de la sustancia P.

En otro ejemplo, el TM híbrido incluye una primera porción que comprende o consiste en un péptido de beta-endorfina, y una segunda porción que comprende o consiste en un péptido de bradiquinina (o un péptido de la sustancia P). En otros ejemplos, la primera porción comprende o consiste en un péptido de beta-endorfina, y la segunda porción comprende o consiste en un péptido seleccionado de un péptido opioide, un péptido de BAM, un péptido de bradiquinina, un péptido de encefalina, un péptido de nociceptina, un péptido de dinorfina, un péptido de galanina, y un péptido de la sustancia P.

En otro ejemplo, el TM híbrido incluye una primera porción que comprende o consiste en un péptido de una encefalina (por ejemplo, leu- o met-encefalina), y una segunda porción que comprende o consiste en un péptido de bradiquinina (o un péptido de la sustancia P). En otros ejemplos, la primera porción comprende o consiste en un péptido de encefalina, y la segunda porción comprende o consiste en un péptido seleccionado de un péptido opioide, un péptido de beta-endorfina, un péptido de bradiquinina, un péptido de BAM, un péptido de nociceptina, un péptido de dinorfina, un péptido de galanina, y un péptido de la sustancia P.

En una realización, el TM comprende o consiste en una primera y una segunda porción (por ejemplo dominios) que son idénticos (o similares) y que, en combinación, proporcionan una interacción eficaz con el receptor sobre la célula diana. Además, también pueden incluirse porciones idénticas/similares (por ejemplo, una tercera y opcionalmente más, etc.). Así, en esta realización, los polipéptidos de la presente invención comprenden una estructura repetida (por ejemplo, TM-TM; TM-TM-TM etc.) del mismo TM (o de un TM similar).

Los ejemplos de dichas estructuras TM repetidas (por ejemplo, TM-TM; TM-TM-TM; etc.) son un TM seleccionado de un péptido opioide, un péptido de beta-endorfina, un péptido de bradiquinina, un péptido de BAM, un péptido de nociceptina, un péptido de dinorfina, un péptido de galanina, un péptido de encefalina, un péptido de la sustancia P, un péptido de GnRH, un péptido de CRF, un péptido de GRP, un péptido de neuromedina B, un péptido de bombesina, un péptido de gastrina, un péptido de CCK, un péptido de SST, un péptido de CST, y un péptido de GHRH (así como sus truncamientos, variantes y análogos).

En una realización, la primera y la segunda porción (y/o posteriores) del TM están separadas por una secuencia espaciadora, por ejemplo, una secuencia peptídica. En una realización, la primera y la segunda porción (y/o posteriores) pueden estar separadas por una secuencia de un máximo de 40 o un máximo de 35 o un máximo de 30 o un máximo de 25 o un máximo de 20 o un máximo de 15 o un máximo de 10 o un máximo de 5 restos aminoácidos. En una realización, la primera y la segunda porción (y/o posteriores) pueden estar separadas por una secuencia de 4, 3, 2, 1 o cero restos aminoácidos.

Las proteínas de fusión de la presente invención en general muestran una afinidad de unión reducida (en la región de hasta 100 veces) por las células diana, cuando se comparan con el correspondiente TM 'libre' (es decir, el TM aislado *per se*). Sin embargo, a pesar de esta observación, las proteínas de fusión de la presente invención demuestran una sorprendente buena eficacia. Esto puede atribuirse a dos características principales. En primer lugar, la proteasa no citotóxica es catalítica y, así, el efecto terapéutico de unas pocas moléculas se amplifica con rapidez dentro de una célula diana. En segundo lugar, los receptores presentes sobre las células diana solo necesitan actuar como puerta de acceso para la entrada del producto terapéutico y no necesitan ser estimuladas necesariamente hasta el nivel requerido para lograr una respuesta farmacológica mediada por receptor-ligando. Por consiguiente, las proteínas de fusión de la presente invención pueden administrarse a una dosificación que sea menor que la empleada para otros tipos de moléculas terapéuticas, que generalmente se administran en cantidades altas de microgramos a miligramos (incluso hasta cientos de miligramos). Por contraste, las proteínas de fusión de la presente invención pueden administrarse a dosificaciones mucho menores, generalmente al menos 10 veces menores, y más generalmente 100 veces menores.

### El dominio de translocación

El componente de translocación de la presente invención permite la translocación de la proteasa no citotóxica (o uno de sus fragmentos) hacia el interior de la célula diana de modo que la expresión funcional de la actividad proteasa se produce dentro del citosol de la célula diana. El componente de translocación preferiblemente es capaz de formar poros permeables a iones en las membranas lipídicas (por ejemplo, membranas endosómicas) bajo condiciones de bajo pH. El componente de translocación puede obtenerse a partir de una fuente de proteínas microbianas, por ejemplo, una fuente de proteínas bacterianas o víricas. Por tanto, en una realización, el componente de translocación comprende o consiste en un dominio de translocación de una enzima, tal como una toxina bacteriana. En otra realización, el dominio de translocación comprende o consiste en el dominio de translocación de una proteína vírica. En una realización, el componente de translocación de la presente invención puede comprender o consistir en una cadena H de neurotoxina clostridial, o uno de sus fragmentos, tal como el dominio H<sub>N</sub> (o uno de sus fragmentos de translocación) de una neurotoxina clostridial.

### El primer y segundo sitio de ruptura de proteasas

Los polipéptidos de la presente invención comprenden un primer sitio de ruptura de proteasas. El primer sitio de ruptura de proteasas permite la ruptura (por ejemplo, la ruptura controlada) de la proteína de fusión en una posición entre el componente de proteasa no citotóxica y el resto de la proteína de fusión. Este acontecimiento de ruptura

actúa para 'activar' la estructura monocatenaria (dominio de translocación-proteasa no citotóxica), y produce la formación de una estructura dicatenaria 'activada', en la que el componente de proteasa no citotóxica está unido covalentemente (por ejemplo, con enlace disulfuro) al resto de la proteína de fusión.

5 Los polipéptidos de la presente invención comprenden también un segundo sitio de ruptura de proteasas. El segundo sitio de ruptura de proteasas permite la ruptura (por ejemplo, la ruptura controlada) de la proteína de fusión en una posición entre el componente de resto de transporte dirigido y el componente de dominio de translocación. Este acontecimiento de ruptura actúa para separar la estructura monocatenaria (dominio de translocación-TM), y produce la formación de una estructura dicatenaria separada, en la que el componente de TM está unido covalentemente (por ejemplo, con enlace disulfuro) al componente de translocación de la proteína de fusión. Al hacer esto, el entorno estructural del componente de TM cambia, de modo que se presenta en una conformación en la que ambas porciones N-terminal y C-terminal (por ejemplo, dominios) ya no están unidas mediante un péptido al resto de la proteína de fusión y, por tanto, cada una es capaz de interactuar libremente (por ejemplo, unirse) a diferentes dominios de unión sobre uno o más receptores.

15 Así la ruptura proteolítica en el primer o el segundo sitio de ruptura de proteasas convierte la proteína de fusión de polipéptido monocatenario en un polipéptido dicatenario. En el caso de una reacción de ruptura en el primer sitio de ruptura de proteasas, el componente de proteasa no citotóxica sigue estando unido mediante un enlace covalente (por ejemplo, un enlace disulfuro) al componente de dominio de translocación y/o al componente de TM. Dicho enlace covalente puede ser indirecto, por ejemplo, a través de una o más moléculas espaciadoras o conectoras, que en sí mismas están unidas al componente de proteasa no citotóxica, al componente de TM y/o al componente de translocación. De modo similar, en el caso de una reacción de ruptura en el segundo sitio de ruptura de proteasas, el componente de dominio de translocación permanece unido al componente TM mediante un enlace covalente (por ejemplo, un enlace disulfuro). Dicho enlace covalente puede ser indirecto, por ejemplo, a través de una o más moléculas espaciadoras o conectoras, que en sí mismas están unidas al componente de translocación y/o de TM.

25 Cuando las reacciones de ruptura se producen en el primer y el segundo sitio de ruptura de proteasas, la proteína de fusión de polipéptido monocatenario se convierte en un polipéptido tricatenario.

El primer y el segundo sitio de ruptura de proteasas pueden introducirse y/o cualquier secuencia de ruptura inherente puede eliminarse al nivel del ADN por medios convencionales, tales como mediante mutagénesis específica dirigida a sitio. La selección para confirmar la presencia de secuencias de ruptura puede realizarse de forma manual o con la ayuda de un programa informático (por ejemplo, el programa MapDraw de DNASTAR, Inc.).

30 Aunque puede emplearse cualquier sitio de ruptura de proteasas para su uso como el primer sitio de ruptura de proteasas y/o para su uso como el segundo sitio de ruptura de proteasas en los polipéptidos de la presente invención, se prefieren los siguientes:

Enteroquinasas	(DDDDK↓)
Factor Xa	(IEGR↓/IDGR↓)
TEV (virus del grabado del tabaco)	(ENLYFQ↓G)
Trombina	(LVPR↓GS)
PreScission	(LEVLFQ↓GP)

35 Otros ejemplos no limitantes incluyen un sitio de ruptura de papaína de plantas y un sitio de ruptura de papaína de insectos, un sitio de ruptura de papaína de crustáceos, un sitio de ruptura de proteasas del rinovirus 3C humano, un sitio de ruptura de proteasas del enterovirus 3C humano, un sitio de ruptura de proteasas del virus del grabado del tabaco (TEV), un sitio de ruptura del virus del jaspeado del nervio del tabaco (TVMV), un sitio de ruptura de subtilisina, un sitio de ruptura de hidroxilamina, o un sitio de ruptura de caspasa 3.

40 La expresión "sitio de ruptura de proteasas" incluye una inteína, que es una secuencia de autorruptura. La reacción de autorruptura puede controlarse, por ejemplo, variando la concentración del agente reductor presente. La expresión "sitio de ruptura de proteasas" también incluye la secuencia de ruptura sobre la que actúa una proteasa no citotóxica (por ejemplo, una neurotoxina clostridial). Un ejemplo de dicho sitio de ruptura es una secuencia del sitio de ruptura de proteínas SNARE, proporcionándose ejemplos de las secuencias de reconocimiento del sitio de ruptura nativo para una gama de proteasas no citotóxicas hacia el final de la presente sección de la descripción.

45 El primer y el segundo sitio de ruptura pueden ser iguales o diferentes. El primer y el segundo sitio de ruptura pueden ser rotos por (solo) la misma proteasa o por (solo) diferentes proteasas.

Como un aspecto separado de la presente invención, los sitios de ruptura/proteasa de ruptura mencionados anteriormente pueden emplearse por separado como un sitio de ruptura/proteasa "destructivos" (tal como se analiza a continuación) si uno se incorpora en un polipéptido de la presente invención.

5 En una realización, en el polipéptido monocatenario, el componente de proteasa no citotóxica y el componente de dominio de translocación están unidos entre sí mediante un enlace disulfuro. Así, después de la ruptura del primer sitio de ruptura de proteasa, el polipéptido adopta una conformación dicatenaria, en la que los componentes de proteasa y de translocación permanecen unidos juntos mediante el enlace disulfuro. Esta reacción de ruptura en general se denomina la etapa de "activación", puesto que provoca que el componente de proteasa no citotóxica tenga una mayor actividad proteasa (por ejemplo, óptima).

10 En una realización, el componente de proteasa no citotóxica forma un enlace covalente con el componente de dominio de translocación de la proteína de fusión. Por ejemplo, en una realización, el resto aminoácido del componente de proteasa que forma el enlace covalente está localizado dentro de los últimos 20, preferiblemente dentro de los últimos 10 restos aminoácidos C-terminales del componente de proteasa. De modo similar, en una  
15 realización, el resto aminoácido dentro del componente de translocación que forma la segunda parte del enlace covalente puede estar localizado dentro de los primeros 20, preferiblemente dentro de los primeros 10 restos aminoácidos N-terminales del componente de translocación.

Las anteriores disposiciones del enlace covalente tienen la ventaja de que los componentes de proteasa y translocación están dispuestos de una manera similar a la de una proteasa no citotóxica nativa (por ejemplo, una neurotoxina clostridial nativa). Como comparación, haciendo referencia a la secuencia de aminoácidos primaria de una neurotoxina clostridial nativa, los respectivos restos aminoácidos cisteína están separados por entre 8 y 27  
20 restos aminoácidos (tomado de Popoff, M.R. y Marvaud, J.-C., 1999, Structural & genomic features of clostridial neurotoxins, capítulo 9, en The Comprehensive Sourcebook of Bacterial Protein Toxins, ed. Alouf & Freer:

Serotipo <sup>1</sup>	Secuencia	Longitud 'nativa' entre C-C
BoNT/A1	CVRGIITSKTKS----LDKGYNKALNDLC	23
BoNT/A2	CVRGIIPFKTKS----LDEGYNKALNDLC	23
BoNT/B	CKSVKAPG----- C	8
BoNT/C	CHKAIDGRS-----LYNKTLDC	15
BoNT/D	CLRLTK-----NSRDDSTC	12
BoNT/E	CKN-IVSVK-----GIRK---SIC	13
BoNT/F	CKS-VIPRK-----GTKAPP-RLC	15
BoNT/G	CKPVMYKNT-----GKSE----QC	13
TeNT	CKKIIPPTNIRENLYNRTASLTDLGGELC	27

<sup>1</sup>Información solo de cepas proteolíticas

25 En una realización, el componente de proteasa no citotóxica y el componente del primer sitio de ruptura de proteasas de una proteína de fusión monocatenaria de la presente invención están separados por un máximo de 30, 25, 20, 15 o 10 restos aminoácidos. En una realización, dichos dos componentes están separados dentro de la proteína de fusión monocatenaria por un máximo de 5, 4, 3, 2 o 1 restos aminoácidos. En otra realización, dichos dos componentes están separados dentro de la proteína de fusión monocatenaria por cero restos aminoácidos.

30 Así, en una realización, la proteasa no citotóxica y el primer sitio de ruptura de proteasas pueden separarse empleando una primera secuencia espaciadora, y dicha secuencia espaciadora está localizada N-terminal con respecto al primer sitio de ruptura de proteasas y C-terminal del componente de proteasa no citotóxica. En una realización, la primera secuencia espaciadora puede comprender todo o parte del primer sitio de ruptura de proteasas, o puede ser parte del componente de proteasa no citotóxica.

5 En una realización, el componente de dominio de translocación (o TM) y el componente del primer sitio de ruptura de proteasas de la proteína de fusión monocatenaria están separados por un máximo de 30, 25, 20, 15 o 10 restos aminoácidos. En una realización, dichos dos componentes están separados dentro de la proteína de fusión monocatenaria por un máximo de 5, 4, 3, 2 o 1 restos aminoácidos. En otra realización, dichos dos componentes están separados dentro de la proteína de fusión monocatenaria por cero restos aminoácidos.

10 Así, en una realización, el dominio de translocación (o TM) y el primer sitio de ruptura de proteasas pueden separarse por una segunda secuencia espaciadora, y dicha segunda secuencia espaciadora está localizada C-terminal con respecto al primer sitio de ruptura de proteasas y N-terminal del componente de dominio de translocación (o TM). La segunda secuencia espaciadora puede ser idéntica o diferente de la primera secuencia espaciadora que separa la proteasa no citotóxica y el primer sitio de ruptura de proteasas. En una realización, la segunda secuencia espaciadora puede comprender todo o parte del segundo sitio de ruptura de proteasas, o puede ser parte del componente de dominio de translocación.

15 En una realización, el componente de dominio de translocación y el componente del segundo sitio de ruptura de proteasas de la proteína de fusión monocatenaria están separados por un máximo de 30, 25, 20, 15 o 10 restos aminoácidos. En una realización, dichos dos componentes están separados dentro de la proteína de fusión monocatenaria por un máximo de 5, 4, 3, 2 o 1 restos aminoácidos. En otra realización, dichos dos componentes están separados dentro de la proteína de fusión monocatenaria por cero restos aminoácidos.

20 Así, en una realización, el dominio de translocación y el segundo sitio de ruptura de proteasas pueden estar separados por una tercera secuencia espaciadora, y dicha tercera secuencia espaciadora está localizada N-terminal o C-terminal con respecto al dominio de translocación. La tercera secuencia espaciadora puede ser idéntica (o diferente) de una o ambas de la primera y la segunda secuencia espaciadora. En una realización, la tercera secuencia espaciadora puede comprender todo o parte del segundo sitio de ruptura de proteasas, o puede ser parte del componente de dominio de translocación.

25 En una realización, el resto de transporte dirigido y el segundo sitio de ruptura de proteasas están separados por un máximo de 30, 25, 20, 15 o 10 aminoácidos. En una realización, dichos dos componentes están separados dentro de la proteína de fusión monocatenaria por un máximo de 5, 4, 3, 2 o 1 restos aminoácidos. En otra realización, dichos dos componentes están separados dentro de la proteína de fusión monocatenaria por cero restos aminoácidos.

30 Así, después de la ruptura en el segundo sitio de ruptura de proteasas, se produce un polipéptido con un resto de transporte dirigido que tiene un dominio N-terminal y un dominio C-terminal que están sustancialmente exentos del resto del conjugado. Esta disposición facilita la interacción de los componentes N-terminal y C-terminal del resto de transporte dirigido con un sitio de unión sobre una célula diana.

35 En una realización, el resto de transporte dirigido y el segundo sitio de ruptura de proteasas pueden estar separados por una cuarta secuencia espaciadora, y dicha cuarta secuencia espaciadora está localizada N-terminal o C-terminal con respecto al resto de transporte dirigido. La cuarta secuencia espaciadora puede ser idéntica (o diferente) de una, dos o todas de la primera, la segunda y la tercera secuencia espaciadora. En una realización, la cuarta secuencia espaciadora puede comprender todo o parte del segundo sitio de ruptura de proteasas, o puede ser parte del componente de dominio de translocación.

40 En una realización, la primera proteasa (mediante la cual puede romperse el primer sitio de ruptura de proteasas) es la misma que la segunda proteasa (mediante la cual puede romperse el segundo sitio de ruptura de proteasas).

Así, en una realización, el tratamiento de la proteína de fusión de polipéptido monocatenario con una única proteasa puede producir la ruptura del primero y del segundo sitio de ruptura de proteasas.

45 Puede emplearse una diversidad de moléculas espaciadoras diferentes en cualquiera de las proteínas de fusión de la presente invención. Los ejemplos de dichas moléculas espaciadoras incluyen GS5, GS10, GS15, GS20, GS25, y Hx27.

### **El enlace covalente**

50 Las proteínas de fusión de polipéptido de la presente invención comprenden dos enlaces covalentes: el primer de dichos enlaces se produce entre el componente de proteasa no citotóxica y el resto de la proteína de fusión; el segundo de dichos enlaces es entre el resto de transporte dirigido y el dominio de translocación. Después de la ruptura proteolítica en el respectivos primer y segundo sitio de ruptura de proteasas, dichos dos enlaces covalentes permanecen intactos. En una realización, los enlaces covalentes no son enlaces peptídicos (es decir, los enlaces covalentes son enlaces no peptídicos). Por ejemplo, en una realización, uno o ambos de dichos enlaces covalentes son enlaces disulfuro.

55 Después de la ruptura proteolítica en el segundo sitio de ruptura de proteasas, el enlace covalente permanece intacto. La ruptura en el segundo sitio de ruptura de proteasas tiene el efecto de exponer el N-terminal (o C-terminal)

del resto de transporte dirigido. Así, la ruptura en el segundo sitio de ruptura de proteasas produce un resto de transporte dirigido que tiene un N-terminal libre y un C-terminal libre.

5 Así, después de la ruptura en el segundo sitio de ruptura de proteasas, el componente de transporte dirigido ya no forma parte de la misma cadena polipeptídica que el componente de dominio de translocación, puesto que el enlace peptídico entre el resto de transporte dirigido y el dominio de translocación se ha roto. Sin embargo, el resto de transporte dirigido sigue unido al dominio de translocación debido a la presencia del enlace covalente.

El enlace covalente puede comprender cualquier enlace covalente capaz de formar o de ser formado entre dos restos aminoácidos en una cadena polipeptídica.

10 En una realización, el enlace covalente es un enlace disulfuro. Puede formarse un enlace disulfuro entre cualesquiera dos grupos tiol (es decir, -SH) presentes en el polipéptido. Como ejemplo, pueden formarse enlaces disulfuro entre dos restos cisteína (o sus variantes equivalentes desde el punto de vista funcional) localizados en una cadena polipeptídica.

15 Así, en una realización, un resto cisteína localizado en el componente de dominio de translocación forma un enlace covalente con otro resto cisteína localizado en el componente de resto de transporte dirigido. Dicho enlace disulfuro permanece intacto después de la ruptura en el segundo sitio de ruptura de proteasas.

20 Los restos aminoácidos localizados en el componente de dominio de translocación y en el componente de resto de transporte dirigido que están unidos por el enlace covalente pueden estar presentes de forma natural en dichos componentes. Así, en una realización, el enlace covalente se forma entre un resto aminoácido presente de forma natural en el componente de dominio de translocación y un resto aminoácido presente de forma natural en el componente de resto de transporte dirigido. Como alternativa, uno o ambos de dichos restos aminoácidos pueden introducirse en el componente de dominio de translocación y/o el componente de resto de transporte dirigido. Los restos aminoácidos pueden introducirse como sustituciones.

25 En una realización, el enlace covalente es un enlace disulfuro formado entre un resto cisteína presente de forma natural en el componente de dominio de translocación y un resto cisteína presente de forma natural en el componente de resto de transporte dirigido. En una realización alternativa, un resto cisteína se introduce específicamente en el dominio de translocación o en el resto de transporte dirigido, o ambos, para facilitar o permitir la formación de un enlace disulfuro entre estos dos componentes.

30 En una realización, uno o más restos cisteína se introducen en el dominio de TM y/o de translocación. Cuando se hace esto, el resto o restos cisteína introducidos pueden estar flanqueados por dos restos aminoácidos pequeños aquirales (tales como glicina y/o alanina). El uso de estos restos aminoácidos evita la formación de una estructura terciaria inmediata y facilita la formación del enlace disulfuro. Los restos aminoácidos pequeños aquirales pueden estar presentes de forma natural o pueden introducirse en el dominio de TM y/o de translocación.

35 En una realización, además del enlace covalente, entre el dominio de translocación y el resto de transporte dirigido se localiza un polipéptido corto (por ejemplo, de 1-20, o 1-10, o 5-10 restos aminoácidos) que proporciona una estructura polipeptídica secundaria. Dicha estructura polipeptídica secundaria ayuda a colocar el dominio de translocación y el resto de transporte dirigido, ayudando con ello (1) a la formación del enlace covalente entre el dominio de TM y de translocación y/o (2) a colocar el TM de modo que sus extremos C-terminal y N-terminal estén en dirección contraria al componente de translocación.

40 Así, en una realización, la estructura secundaria del polipéptido actúa para acercar el resto de transporte dirigido al dominio de translocación, provocando con ello que la formación del enlace covalente sea energéticamente más favorable.

En una realización, un polipéptido capaz de formar una estructura secundaria del polipéptido tal como se describió anteriormente es una secuencia de polipéptido que contiene al menos un resto aminoácido 'voluminoso', tal como un resto prolina.

45 Así, en una realización, entre el dominio de translocación y el resto de transporte dirigido se localiza un polipéptido que comprende al menos un resto aminoácido voluminoso. Dicho resto voluminoso ayuda a formar una curva en la cadena polipeptídica, de modo que parte del resto de transporte dirigido se acerca al dominio de translocación, que de otra forma no se acercaría.

50 El correspondiente enlace covalente entre el componente de proteasa no citotóxica y el resto de la proteína de fusión (por ejemplo, el componente de translocación y/o el componente de TM) puede formarse de la misma manera que se describió anteriormente para el enlace covalente entre el componente de dominio de translocación y el componente de TM. Tal como se describió anteriormente, pueden introducirse una o más estructuras secundarias y/o uno o más restos aminoácidos voluminosos.

55 En una realización, el enlace covalente entre el componente de proteasa no citotóxica y el resto de la proteína de fusión se produce entre el componente de proteasa no citotóxica y el componente de dominio de translocación. En

una realización, el enlace covalente entre el componente de proteasa no citotóxica y el componente de dominio de translocación emplea restos cisteína naturales localizados sobre los respectivos componentes, tal como, por ejemplo, uno o más de los restos cisteína ilustrados anteriormente en la sección de la descripción. Como alternativa, pueden introducirse uno o más restos cisteína apropiados en los respectivos componentes.

- 5 La proteína de fusión puede comprender uno o más marcadores de purificación, que están localizados en posición N-terminal con respecto al componente de proteasa y/o en posición C-terminal con respecto al componente de translocación.

Aunque puede emplearse cualquier marcador de purificación, se prefieren los siguientes:

- marcador His (por ejemplo, 6 × histidina), preferiblemente como marcador C-terminal y/o N-terminal,
- 10 marcador MBP (proteína de unión a maltosa), preferiblemente como marcador N-terminal,
- marcador GST (glutatión-S-transferasa), preferiblemente como marcador N-terminal,
- marcador His-MBP, preferiblemente como marcador N-terminal,
- marcador His-GST, preferiblemente como marcador N-terminal,
- marcador tiorredoxina, preferiblemente como marcador N-terminal,
- 15 marcador CBD (dominio de unión a quitina), preferiblemente como marcador N-terminal.

#### Aplicaciones terapéuticas

El componente de TM dirige la molécula terapéutica del inhibidor de la secreción dirigido (TSI) de la presente invención hacia la célula diana deseada.

- 20 Como ejemplo, el uso de los TM descritos en esta memoria descriptiva (tales como un péptido opioide, un péptido de beta-endorfina, un péptido de bradiquinina, un péptido de BAM, un péptido de nociceptina, un péptido de dinorfina, un péptido de galanina, un péptido de encefalina, un péptido de la sustancia P) dirige la molécula terapéutica del inhibidor de la secreción dirigido (TSI) de la presente invención hacia células sensibles al dolor (por ejemplo, aferentes sensoriales primarias). Por tanto, las proteínas de fusión resultantes proporcionan moléculas terapéuticas para suprimir el dolor; el solicitante se remite a los documentos WO2006/059093, WO2007/138339 y WO96/33273.

- 25 Los TM descritos en esta memoria descriptiva pueden utilizarse para dirigir las moléculas del inhibidor de la secreción dirigido (TSI) de la presente invención hacia células que estimulan la inflamación neurogénica. Por consiguiente, las moléculas del inhibidor de la secreción dirigido (TSI) de la presente invención proporcionan moléculas terapéuticas para suprimir la inflamación neurogénica; el solicitante se remite a los documentos WO2010/138395, WO2010/138392, WO2010/138387, WO2010138382 y WO2010/138379. Los TM preferidos para su uso en estas terapias y moléculas de TSI incluyen TM opioides, tales como nociceptina y dinorfina.
- 30

- Los TM descritos en esta memoria descriptiva pueden utilizarse para dirigir las moléculas del inhibidor de la secreción dirigido (TSI) de la presente invención hacia células que estimulan trastornos urogenitales-neurológicos, tales como una vejiga hiperactiva. Por consiguiente, las moléculas del inhibidor de la secreción dirigido (TSI) de la presente invención proporcionan moléculas terapéuticas para suprimir trastornos urogenitales-neurológicos, tales como una vejiga hiperactiva; el solicitante se remite a los documentos WO2010/138393, WO2010/138389, WO2010/138384, y WO2010/138366. Los TM preferidos para su uso en estas terapias y moléculas de TSI incluyen TM opioides, tales como nociceptina y dinorfina.
- 35

- Los TM tales como el péptido de la hormona liberadora de gonadotropina (GnRH), el péptido de CRF, el péptido de GRP, el péptido de neuromedina B, el péptido de bombesina, el péptido de gastrina, el péptido de CCK, el péptido de SST, el péptido de CST, y el péptido de GHRH pueden utilizarse para dirigir las moléculas de TSI de la presente invención a células que estimulan el cáncer o que son células cancerosas *per se*. Por consiguiente, las moléculas del inhibidor de la secreción dirigido (TSI) de la presente invención proporcionan moléculas terapéuticas para suprimir trastornos neuroendocrinos, tales como acromegalia y enfermedad de Cushing, y para suprimir el cáncer (por ejemplo, cáncer de pulmón, cáncer renal, cáncer de cerebro, cáncer de mama, cáncer pancreático, cáncer colorrectal, cáncer adrenal, cáncer esofágico, linfoma, leucemia, leucemia aguda, cáncer de vejiga, cáncer de huesos, cáncer de intestino, cáncer cervical, leucemia linfocítica crónica, linfoma de Hodgkin, cáncer de hígado, cáncer de piel, cáncer orofaríngeo, mieloma, cáncer de próstata, cáncer gástrico, cáncer testicular, cáncer uterino o sarcoma de Kaposi; el solicitante se remite a los documentos WO2009/150489, WO2009/150470 y WO2010/055358. Los TM preferidos para su uso en estas terapias y moléculas de TSI incluyen péptidos de GHRH,
- 40
- 45
- 50 péptidos de SST y péptidos de CST.

**Sitios de ruptura destructivos**

Los polipéptidos de la presente invención puede modificarse aún más para reducir o prevenir efectos secundarios no deseados asociados con la dispersión hacia áreas no deseadas. Según esta realización, el polipéptido comprende un sitio de ruptura destructivo. El sitio de ruptura destructivo se diferencia del sitio de 'activación' (es decir, de formación de una cadena doble) y del segundo sitio de ruptura de proteasas (es decir, de formación de un TM con dominios C-terminal y N-terminal libres). Dicho sitio de ruptura destructivo puede ser roto por una tercera proteasa y no por la primera ni la segunda proteasa. Además, cuando es roto en un sitio de ruptura destructivo por la tercera proteasa, el polipéptido de la invención tienen una menor potencia (por ejemplo, una menor capacidad de unión a la célula diana prevista, una menor actividad de translocación y/o una menor actividad proteasa no citotóxica). Como ejemplo, el solicitante se remite a los documentos WO 2010/094905 y WO 02/044199.

Así, según esta realización, la presente invención proporciona un polipéptido que puede inactivarse y/o destruirse de una manera controlable en una localización fuera de sitio.

En una realización, el sitio de ruptura destructivo es reconocido y roto por una tercera proteasa (es decir, una proteasa destructiva) seleccionada de una proteasa en circulación (por ejemplo, una proteasa extracelular, tal como una proteasa sérica o una proteasa de la cascada de coagulación sanguínea), una proteasa asociada a un tejido (por ejemplo, una metaloproteasa de matriz (MMP), tal como una MMP del músculo), y una proteasa intracelular (preferiblemente una proteasa que esté ausente en el célula diana). Así, en uso, si un polipéptido de la presente invención se aleja de su célula diana prevista y/o es captado por una célula que no es su diana, el polipéptido será inactivado por medio de la ruptura del sitio de ruptura destructivo (por la tercera proteasa).

Las metaloproteasas de matriz (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 rompe una diversidad de proteínas ancladas a la membrana de la superficie celular para "desprender" los dominios extracelulares. Otras MMP preferidas incluyen adamalisin, serralisin, y astacinas. Otro grupo de proteasas destructivas preferidas son las proteasas de sangre de mamífero, 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 serina proteasa asociada a MBP.

Según un segundo aspecto de la presente invención, se proporciona una secuencia de ácido nucleico que codifica la proteína de fusión de polipéptido descrita anteriormente.

En un aspecto preferido de la presente invención, la secuencia de ADN se prepara como parte de un vector de ADN, y dicho vector comprende un promotor y un terminador. La secuencia de ADN que codifica la proteína de fusión de polipéptido descrita anteriormente se localiza cadena abajo del promotor; el terminador se localiza cadena arriba de la secuencia de ácido nucleico.

En una realización preferida, el vector presenta un promotor seleccionado de:

Promotor	Agente de inducción	Condición de inducción típica
tac (híbrido)	IPTG	0,2 mM (0,05-2,0 mM)
AraBAD	L-arabinosa	0,2% (0,002-0,4%)
operador <i>lac</i> -T7	IPTG	0,2 mM (0,05-2,0 mM)

La construcción de ADN de la presente invención se diseña preferiblemente por medios electrónicos y después se sintetiza mediante técnicas de síntesis de ADN convencionales.

La información de la secuencia de ADN mencionada anteriormente se modifica opcionalmente para el sesgo de codones según el sistema de expresión de la célula hospedante (por ejemplo, *E. coli*) que se va a emplear finalmente.

El esqueleto de ADN preferiblemente se selecciona para detectar cualquier secuencia de ácido nucleico inherente que, tras transcribirse y traducirse, produciría una secuencia de aminoácidos que se corresponde con el sitio de ruptura de proteasas codificado por la segunda secuencia codificadora de péptido. Esta selección puede realizarse de forma manual o con la ayuda de un programa informático (por ejemplo, el programa MapDraw de DNASTAR, Inc.).

Según otra realización de la presente invención, se proporciona un método para preparar una proteína de fusión de polipéptido monocatenario, según se describió anteriormente, que comprende expresar una secuencia de ácido

nucleico que codifica la proteína de fusión descrita anteriormente, o un vector de ADN según se describió anteriormente, en una célula hospedante.

Según otra realización de la presente invención, se proporciona un método para preparar un agente no citotóxico, que comprende:

- 5 a. proporcionar una disolución que contienen una proteína de fusión de polipéptido monocatenario de la invención;
- b. añadir a dicha disolución una primera proteasa capaz de romper el primer sitio de ruptura de proteasas, y una segunda proteasa capaz de romper el segundo sitio de ruptura de proteasas;
- c. romper el primer sitio de ruptura de proteasas y el segundo sitio de ruptura de proteasas;
- formando con ello una proteína de fusión tricatenaria.

- 10 En una realización, la primera proteasa y la segunda proteasa se añaden secuencialmente. En una realización alternativa, la segunda proteasa se añade antes de la primera proteasa. En otra realización, la primera proteasa y la segunda proteasa se añaden simultáneamente.

Este aspecto proporciona un polipéptido tricatenario. Con más detalle, el polipéptido tricatenario resultante generalmente tiene una estructura en la que:

- 15 a. la primera cadena comprende la proteasa no citotóxica, o uno de sus fragmentos, y dicha proteasa o fragmento de proteasa es capaz de romper una proteína del aparato de fusión exocítica de una célula diana;
- b. la segunda cadena comprende el dominio de translocación que es capaz de translocar la proteasa o el fragmento de proteasa desde el interior de un endosoma, a través de la membrana endosómica y hacia el citosol de la célula diana;
- 20 c. la tercera cadena comprende el resto de transporte dirigido que es capaz de unirse a un sitio de unión sobre la célula diana, y dicho sitio diana es capaz de sufrir una endocitosis para ser incorporado en un endosoma dentro de la célula diana;
- d. la primera y la segunda cadena están unidos entre sí mediante un enlace disulfuro; y el segundo y el tercer dominio están unidos entre sí mediante un enlace covalente no peptídico.

## 25 **Transporte del polipéptidos**

Según otro aspecto de la presente invención, se proporciona una proteína de fusión de polipéptido monocatenario según se describió anteriormente, o un polipéptido no citotóxico según se describió anteriormente, para su uso para tratar, prevenir o mejorar un trastorno médico.

- 30 En uso, la presente invención emplea una composición farmacéutica que comprende un polipéptido, junto con al menos un componente seleccionado de un vehículo, excipiente, adyuvante, propelente y/o sal farmacéuticamente aceptable.

- 35 Los polipéptidos de la presente invención pueden formularse para la aplicación oral, parenteral, en infusión continua, mediante implante, inhalación o aplicación tópica. Las composiciones adecuadas para la inyección pueden estar en forma de disoluciones, suspensiones o emulsiones, o polvos secos que se disuelven o se suspenden en un vehículo adecuado antes del uso.

- 40 Los medios de transporte local pueden incluir un aerosol o cualquier otro pulverizado (por ejemplo, un nebulizador). A este respecto, una formulación en aerosol de un polipéptido permite el transporte a los pulmones y/u otras vías nasales y/o bronquiales o respiratorias. Una vía de administración preferida se selecciona de: sistémica (por ejemplo, intravenosa), laparoscópica y/o inyección localizada (por ejemplo, una inyección transesfenoidal directamente en una célula diana, tal como un tumor).

- 45 En el caso de formulaciones para inyección, se puede incluir una sustancia farmacéuticamente activa para ayudar a la retención o reducir la eliminación del polipéptido del sitio de la administración. Un ejemplo de dicha sustancia farmacéuticamente activa es un vasoconstrictor, tal como adrenalina. Esta formulación confiere la ventaja de aumentar el tiempo de residencia del polipéptido después de la administración y, así, aumenta y/o potencia su efecto.

- 50 Los intervalos de dosificación para la administración de los polipéptidos de la presente invención son los que producen el efecto terapéutico deseado. Se apreciará 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, el grado o la gravedad del trastorno del paciente, las contraindicaciones, si existen, y el criterio del médico encargado. Las variaciones en estos niveles de dosificación pueden ajustarse empleando rutinas empíricas convencionales para la optimización.

Las dosificaciones diarias adecuadas (por kg de peso del paciente) están en el intervalo de 0,0001-1 mg/kg, preferiblemente de 0,0001-0,5 mg/kg, más preferiblemente 0,002-0,5 mg/kg, y en particular preferiblemente 0,004-0,5 mg/kg. La dosificación unitaria puede variar de menos de 1 microgramo a 30 mg, pero generalmente estará en la región de 0,01 a 1 mg por dosis, que pueden administrarse a diario o preferiblemente con menos frecuencia, tal como una vez a la semana o cada seis meses. Un régimen de dosificación particularmente preferido se basa en 2,5 ng de polipéptido como la dosis 1X. A este respecto, las dosificaciones preferidas están en el intervalo de 1X-100X (concretamente, 2,5-250 ng).

Las formas de dosificación fluidas generalmente se preparan utilizando el polipéptido y un vehículo estéril apirógeno. El polipéptido, dependiendo del vehículo y la concentración utilizados, puede disolverse o suspenderse en el vehículo. Cuando se preparan disoluciones, el polipéptido puede disolverse en el vehículo, la disolución puede hacerse isotónica si es necesario mediante la adición de cloruro de sodio, y puede esterilizarse mediante filtración a través de un filtro estéril utilizando técnicas asépticas antes de ser introducido en viales o ampollas estériles adecuados y sellarse. Como alternativa, si la estabilidad de la disolución es adecuada, la disolución en sus recipientes sellados puede esterilizarse mediante un autoclave. De forma ventajosa, pueden disolverse en el vehículo ciertos aditivos, tales como agentes tamponantes, solubilizantes, estabilizantes, conservantes o bactericidas, agentes suspensores o emulgentes y/o agentes anestésicos locales.

Pueden prepararse polvos secos, que se disuelven o se suspenden en un vehículo adecuado antes del uso, introduciendo los ingredientes preesterilizados en un recipiente estéril utilizando una técnica aséptica en un área estéril. Como alternativa, los ingredientes pueden disolverse en recipientes adecuados utilizando una técnica aséptica en un área estéril. El producto después se liofiliza y los recipientes se sellan asépticamente.

Las suspensiones parenterales, adecuadas para la inyección intramuscular, subcutánea o intradérmica, se preparan sustancialmente de la misma manera, excepto que los componentes estériles se suspenden en el vehículo estéril, en lugar de ser disueltos, y la esterilización no puede realizarse mediante filtración. Los componentes pueden aislarse en un estado estéril o, como alternativa, pueden esterilizarse después del aislamiento, por ejemplo, mediante irradiación con rayos gamma.

De forma ventajosa, un agente suspensor, por ejemplo, polivinilpirrolidona, se incluye en las composiciones para facilitar la distribución uniforme de los componentes.

### Sección de definiciones

El resto de transporte dirigido (TM) significa cualquier estructura química que interacciona funcionalmente con un sitio de unión para provocar una asociación física entre el polipéptido de la invención y la superficie de una célula diana. El término TM incluye cualquier molécula (es decir, una molécula natural o uno de sus variantes química o físicamente modificados) que es capaz de unirse a un sitio de unión sobre la célula diana, y dicho sitio de unión es capaz de ser internalizado (por ejemplo, mediante la formación de endosomas), lo cual también se denomina endocitosis mediada por receptores. El TM puede poseer una función de translocación de membrana endosómica, en cuyo caso no es necesario que estén presentes un componente de TM y de dominio de translocación separados en un agente de la presente invención. A lo largo de la descripción anterior se han descrito TM específicos. La referencia a dichos TM es solo como ejemplos, y la presente invención incluye todos sus variantes y derivados, que conserven la capacidad de unión básica (es decir, dirigida) de los ejemplos de TM.

Tal como se mencionó previamente, los TM preferidos incluyen anticuerpos (por ejemplo, fragmentos de anticuerpos) y andamiajes de unión, en especial, anticuerpos/fragmentos y andamiajes disponibles en el mercado diseñados para unirse (por ejemplo, específicamente) a células diana.

Los andamiajes de proteínas representan una nueva generación de marcos de unión universales para complementar el repertorio en expansión de anticuerpos monoclonales terapéuticos y derivados, tales como scFv, moléculas de Fab, dAb (anticuerpos de dominio único), de camélidos, 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. Estos andamiajes incluyen, pero no se limitan a:

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

(ii) andamiajes basados en lipocalina - anticalinas (Skerra, 2008, "Alternative binding proteins: anticalins - harnessing the structural plasticity of the lipocalin ligand pocket to engineer novel binding activities", *FEBS J.*, 275:2677-2683);

(iii) andamiajes basados en fibronectina - adnectina (Dineen et al., 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 et al., 2005, "Multivalent avimer proteins evolved by exon shuffling of a family of human receptor domains", *Nat. Biotechnol.*, 23:1556-1561);

(v) andamiajes basados en anquirina - darpinas (Zahnd et al., 2006, "Selection and characterization of Her2 binding-designed ankyrin repeat proteins", J. Biol. Chem., 281:35167-35175); y

5 (vi) andamiajes de centirina - basados en un plegamiento de la proteína que tiene una significativa homología estructural con los dominios Ig con bucles que son análogos a CDR. Los dominios Ig son un módulo habitual en proteínas humanas y se han aplicado mucho como proteínas de andamiaje alternativas.

Los andamiajes de unión pueden utilizarse para dirigirse a tipos celulares concretos mediante la interacción con proteínas de la superficie celular específicas, receptores u otros epitopos de la superficie celular, tales como grupos azúcar. Estos andamiajes modificados pueden introducirse sobre los polipéptidos basados en una proteasa no citotóxica recombinante de la presente invención.

10 El TM de la presente invención se une (preferiblemente se une específicamente) a la célula diana en cuestión. La expresión "se une específicamente" significa preferiblemente que un TM concreto se une a la célula diana con una afinidad de unión ( $K_a$ ) de  $10^6 M^{-1}$  o mayor, preferiblemente  $10^7 M^{-1}$  o mayor, más preferiblemente  $10^8 M^{-1}$  o mayor, y lo más preferiblemente  $10^9 M^{-1}$  o mayor.

15 La referencia a TM en la presente memoria descriptiva incluye fragmentos y sus variantes, que conservan la capacidad de unirse a la célula diana en cuestión. Como ejemplo, un variante puede tener al menos 80%, preferiblemente al menos 90%, más preferiblemente al menos 95%, y lo más preferiblemente al menos 97 o al menos 99% de homología de secuencia de aminoácidos con el TM de referencia. Así, un variante puede incluir uno o más análogos de un aminoácido (por ejemplo, un aminoácido no natural) o un enlace sustituido. Además, como ejemplo, el término fragmento, cuando se emplea en relación con un TM, significa un péptido que tiene al menos diez, preferiblemente al menos veinte, más preferiblemente al menos treinta, y lo más preferiblemente al menos cuarenta restos aminoácidos en el TM de referencia. El término fragmento también se refiere a los variantes mencionados anteriormente. Así, como ejemplo, un fragmento de la presente invención puede comprender una secuencia peptídica que tenga al menos 10, 20, 30 o 40 aminoácidos, en el que la secuencia peptídica tiene al menos 80% de homología de secuencia frente a una correspondiente secuencia peptídica de aminoácidos (contiguos) del péptido de referencia.

25 Es habitual confirmar que un TM se une a la célula diana seleccionada. Por ejemplo, puede emplearse un experimento de desplazamiento radiactivo simple, en el que el tejido o las células representativas de una célula diana en cuestión se exponen a un TM marcado (por ejemplo, tritiado) en presencia de un exceso de TM no marcado. En este experimento, las proporciones relativas de unión no específica y específica pueden evaluarse, lo cual permite confirmar que el TM se une a la célula diana. Opcionalmente, el ensayo puede incluir uno o más antagonistas de la unión, y el ensayo puede comprender también observar la pérdida de unión del TM. Los ejemplos de este tipo de experimentos pueden encontrarse en Hulme, E.C. (1990), Receptor-binding studies, a brief outline, pp. 303-311, en Receptor biochemistry, A Practical Approach, ed. E.C. Hulme, Oxford University Press.

30 En el contexto de la presente invención, la referencia a un TM peptídico incluye sus análogos de péptidos, con la condición de que el análogo se una al mismo receptor que el correspondiente TM de 'referencia'. Dichos análogos pueden incluir restos sintéticos, tales como:

$\beta$ -Nal =  $\beta$ -naftilalanina

$\beta$ -Pal =  $\beta$ -piridilalanina

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

40 hArg(Et)<sub>2</sub> = N,N'-guanidino-(dimetil)-homoarginina

hArg(CH<sub>2</sub>CF<sub>3</sub>)<sub>2</sub> = N,N-guanidino-bis-(2,2,2,-trifluoroetil)-homoarginina

hArg(CH<sub>3</sub>, hexil) = N,N'-guanidino-(metil, hexil)-homoarginina

Lys(Me) = N<sup>ε</sup>-metil-lisina

Lys(iPr) = N<sup>ε</sup>-isopropil-lisina

45 AmPhe = aminometilfenilalanina

AChxAla = aminociclohexilalanina

Abu = ácido  $\alpha$ -aminobutírico

Tpo = 4-tiaprolina

MeLeu = N-metil-leucina

50 Orn = ornitina

Nle - norleucina

Nva = norvalina

Trp(Br) = 5-bromo-triptófano

Trp(F) = 5-fluoro-triptófano

5 Trp(NO<sub>2</sub>) = 5-nitro-triptófano

Gaba = ácido γ-aminobutírico

Bmp = J-mercaptopropionilo

Ac = acetilo

Pen - pencilamina

10 Los polipéptidos de la presente invención pueden carecer de un dominio H<sub>C</sub> o H<sub>CC</sub> funcional de una neurotoxina clostridial. Por consiguiente, dichos polipéptidos no son capaces de unirse a las membranas sinaptosómicas de rata (a través del componente H<sub>C</sub> clostridial) en los ensayos de unión descritos en Shone et al. (1985), Eur. J. Biochem., 151, 75-82. En una realización, los polipéptidos carecen de los últimos 50 aminoácidos C-terminales de una holotoxina de neurotoxina clostridial. En otra realización, los polipéptidos carecen de los últimos 100, 150, 200, 250, 15 o 300 restos aminoácidos C-terminales de una holotoxina de neurotoxina clostridial. Como alternativa, la actividad de unión de H<sub>C</sub> puede negarse/reducirse mediante mutagénesis; como ejemplo, referido a BoNT/A por comodidad, la modificación de una o dos mutaciones de restos aminoácidos (W1266 a L, e Y1267 a F) en el bolsillo de unión a gangliósidos provoca que la región H<sub>C</sub> pierda su función de unión al receptor. Pueden realizarse mutaciones análogos en los componentes que no son el péptido clostridial de serotipo A, por ejemplo, una construcción basada en toxina botulínica B con mutaciones (W1262 a L, e Y1263 a F) o toxina botulínica E (W1224 a L, e Y1225 a F). 20 Otras mutaciones en el sitio activo logran la misma ablación de la actividad de unión al receptor de H<sub>C</sub>, por ejemplo Y1267S en la toxina botulínica de tipo A y el correspondiente resto altamente conservado en otras neurotoxinas clostridiales. Los detalles de esta y otras mutaciones se describen en Rummel et al. (2004) (Molecular Microbiol., 51:631-634).

25 El péptido H<sub>C</sub> de una neurotoxina clostridial nativa comprende aproximadamente 400-440 restos aminoácidos, y consiste en dos dominios funcionalmente diferenciados de aproximadamente 25 kDa cada uno, concretamente la región N-terminal (denominada habitualmente el péptido o dominio de H<sub>CN</sub>) y la región C-terminal (denominada habitualmente el péptido o dominio de H<sub>CC</sub>). Además, está bien documentado que la región C-terminal (H<sub>CC</sub>), que constituyen 160-200 restos aminoácidos C-terminales, es responsable de la unión de una neurotoxina clostridial a sus receptores celulares naturales, concretamente a los terminales nerviosos en la zona de unión neuromuscular. 30 Así, la referencia a lo largo de esta memoria descriptiva a una cadena pesada clostridial que carece de un péptido (o dominio) de H<sub>C</sub> de cadena pesada funcional, de modo que dicha cadena pesada es incapaz de unirse a los receptores de la superficie celular a los cuales se une una neurotoxina clostridial nativa, significa que la cadena pesada clostridial simplemente carece de un péptido de H<sub>CC</sub> funcional. En otras palabras, la región del péptido de H<sub>CC</sub> está parcial o totalmente delecionada, o modificada de otra forma (por ejemplo, mediante un tratamiento químico o proteolítico convencional) para inactivar su capacidad de unión nativa por los terminales nerviosos en la zona de unión neuromuscular. 35

Así, en una realización, un péptido de H<sub>N</sub> clostridial de la presente invención carece de parte de una porción de péptido C-terminal (H<sub>CC</sub>) de una neurotoxina clostridial y, así, carece de la función de unión de H<sub>C</sub> de la neurotoxina clostridial nativa. Como ejemplo, en una realización, el péptido de H<sub>N</sub> clostridial C-terminalmente extendido carece de los 40 restos aminoácidos C-terminales, o los 60 restos aminoácidos C-terminales, o los 80 restos aminoácidos C-terminales, o los 100 restos aminoácidos C-terminales, o los 120 restos aminoácidos C-terminales, o los 140 restos aminoácidos C-terminales, o los 150 restos aminoácidos C-terminales, o los 160 restos aminoácidos C-terminales de una cadena pesada de neurotoxina clostridial. En otra realización, el péptido de H<sub>N</sub> clostridial de la presente invención carece de la porción de péptido C-terminal (H<sub>CC</sub>) entera de una neurotoxina clostridial y, así, carece de la función de unión de H<sub>C</sub> de la neurotoxina clostridial nativa. Como ejemplo, en una realización, el péptido de H<sub>N</sub> clostridial carece de los 165 restos aminoácidos C-terminales, o los 170 restos aminoácidos C-terminales, o los 175 restos aminoácidos C-terminales, o los 180 restos aminoácidos C-terminales, o los 185 restos aminoácidos C-terminales, o los 190 restos aminoácidos C-terminales, o los 195 restos aminoácidos C-terminales de una cadena pesada de neurotoxina clostridial. Como otro ejemplo, el péptido de H<sub>N</sub> clostridial de la presente invención carece de una secuencia de referencia de H<sub>CC</sub> clostridial seleccionada del grupo que consiste en: 40 45 50

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

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

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

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

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

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

neurotoxina botulínica de tipo B - restos aminoácidos (Y1106-E1297)

5 neurotoxina del tétanos - restos aminoácidos (Y1128-D1315).

Las secuencias de referencia identificadas anteriormente deben considerarse una guía, puesto que pueden producirse ligeras variaciones según los subserotipos.

10 La proteasa de la presente invención incluye todas las proteasas no citotóxicas que son capaces de romper una o más proteínas del aparato de fusión exocítica en células eucariotas. La proteasa de la presente invención preferiblemente es una proteasa bacteriana (o uno de sus fragmentos). Más preferiblemente, la proteasa bacteriana se selecciona de los géneros *Clostridium* o *Neisseria/Streptococcus* (por ejemplo, una cadena L clostridial o una IgA proteasa neisserial, preferiblemente de *N. gonorrhoeae* o *S. pneumoniae*).

15 La presente invención también incluye variantes de proteasas no citotóxicas (es decir, variantes de moléculas de proteasa naturales), con la condición de los variantes de proteasas aún demuestren la actividad proteasa necesaria. Como ejemplo, un variante puede tener al menos 70%, preferiblemente al menos 80%, más preferiblemente al menos 90%, y lo más preferiblemente al menos 95 o al menos 98% de homología de secuencia de aminoácidos con una secuencia de proteasa de referencia. Así, el término variante incluye proteasas no citotóxicas que tiene una mayor (o menor) actividad endopeptidasa; se menciona particularmente la mayor  $K_{cat}/K_m$  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 emplea en relación con una proteasa, generalmente significa un péptido que tiene al menos 20 150 , preferiblemente al menos 200, más preferiblemente al menos 250, y lo más preferiblemente al menos 300 restos aminoácidos de la proteasa de referencia. Con respecto al componente del 'fragmento' de TM (tal como se analizó anteriormente), los 'fragmentos' de proteasas de la presente invención incluyen fragmentos de variantes de proteasas basados en una secuencia de referencia.

25 La proteasa de la presente invención preferiblemente muestra una actividad serina o metaloproteasa (por ejemplo, actividad endopeptidasa). La proteasa es preferiblemente específica para una proteína SNARE (por ejemplo, SNAP-25, sinaptobrevina/VAMP, o sintaxina).

30 Se mencionan en particular los dominios de proteasa de neurotoxinas, por ejemplo, los dominios de proteasa de neurotoxinas bacterianas. Así, la presente invención incluye el uso de dominios de neurotoxinas, que aparecen en la naturaleza, así como las versiones preparadas de modo recombinante de dichas neurotoxinas que aparecen en la naturaleza. Los ejemplos de neurotoxinas son las producidas por clostridios, y la expresión neurotoxina clostridial incluye neurotoxinas producidas por *C. tetani* (TeNT), y por *C. botulinum* (BoNT), serotipos A-G, así como las neurotoxinas similares a BoNT muy relacionadas producidas por *C. baratii* y *C. butyricum*. Las abreviaturas mencionadas anteriormente se emplean a lo largo de la presente memoria descriptiva. Por ejemplo, la nomenclatura BoNT/A indica que la fuente de la neurotoxina es BoNT (serotipo A).

40 Los BoNT comparten una estructura común, siendo proteínas dicatenarias de aproximadamente 150 kDa, que consisten en una cadena pesada (cadena H) de aproximadamente 100 kDa unida covalentemente mediante un único enlace disulfuro a una cadena ligera (cadena L) de aproximadamente 50 kDa. La cadena H consiste en dos dominios, cada uno de aproximadamente 50 kDa. El dominio C-terminal ( $H_C$ ) es necesario para la unión neuronal de alta afinidad, mientras que se ha propuesto que el dominio N-terminal ( $H_N$ ) está implicado en la translocación de membrana. La cadena L es una metaloproteasa dependiente de cinc que es responsable de la ruptura del sustrato de proteína SNARE.

45 La expresión fragmento de cadena L significa un componente de la cadena L de una neurotoxina, en el que dicho fragmento muestra una actividad metaloproteasa y es capaz de romper proteolíticamente una vesícula y/o una proteína asociada a la membrana plasmática implicada en la exocitosis celular.

Los ejemplos de secuencias de proteasa (de referencia) adecuadas incluyen:

neurotoxina botulínica de tipo A - restos aminoácidos (1-448)

neurotoxina botulínica de tipo B - restos aminoácidos (1-440)

neurotoxina botulínica de tipo C - restos aminoácidos (1-441)

50 neurotoxina botulínica de tipo D - restos aminoácidos (1-445)

neurotoxina botulínica de tipo E - restos aminoácidos (1-422)

neurotoxina botulínica de tipo F - restos aminoácidos (1-439)

neurotoxina botulínica de tipo G - restos aminoácidos (1-441)

neurotoxina del tétanos - restos aminoácidos (1-457)

IgA proteasa - restos aminoácidos (1-959)\*

- 5 \* Pohlner, J. et al. (1987), Nature, 325, pp. 458-462.

La secuencia de referencia identificada anteriormente debe considerarse una guía, puesto que pueden producirse ligeras variaciones según los subserotipos. Como ejemplo, el documento US 2007/0166332 cita secuencias clostridiales ligeramente diferentes:

neurotoxina botulínica de tipo A - restos aminoácidos (M1-K448)

- 10 neurotoxina botulínica de tipo B - restos aminoácidos (M1-K441)

neurotoxina botulínica de tipo C - restos aminoácidos (M1-K449)

neurotoxina botulínica de tipo D - restos aminoácidos (M1-R445)

neurotoxina botulínica de tipo E - restos aminoácidos (M1-R422)

neurotoxina botulínica de tipo F - restos aminoácidos (M1-K439)

- 15 neurotoxina botulínica de tipo G - restos aminoácidos (M1-K446)

neurotoxina del tétanos - restos aminoácidos (M1-A457).

- 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 cadena ligera puedan dirigirse específicamente a los componentes centrales del aparato de liberación de neurotransmisores y, así, participar en la ejecución del mecanismo celular global mediante el cual una toxina clostridial rompe proteolíticamente un sustrato. Las cadenas ligeras de toxinas clostridiales tienen una longitud de aproximadamente 420-460 aminoácidos y comprenden un dominio enzimático. Las investigaciones han demostrado que no es necesaria la longitud completa de una cadena ligera de 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 son necesarios para la actividad enzimática. Como otro ejemplo no limitante, los primeros ocho aminoácidos de la cadena ligera de TeNT no son necesarios para la actividad enzimática. De forma similar, el carboxilo-terminal 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 son necesarios para la actividad enzimática. Como otro ejemplo no limitante, los últimos 31 aminoácidos de la cadena ligera de TeNT (restos 427-457) no son necesarios para la actividad enzimática. Así, los aspectos de esta realización pueden incluir cadenas ligeras de toxinas clostridiales que comprenden un dominio enzimático que tiene 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 que tiene una longitud, por ejemplo, de un máximo de 350 aminoácidos, un máximo de 375 aminoácidos, un máximo de 400 aminoácidos, un máximo de 425 aminoácidos, y un máximo de 450 aminoácidos.

En una realización, la proteasa no citotóxica rompe una proteína SNARE no neuronal, tal como una proteína SNAP-23. En una realización, la proteasa no citotóxica es una cadena L de la toxina botulínica modificada capaz de romper a SNAP-23. Un ejemplo de dicha cadena L modificada se describe en Chen y Barbieri, PNAS, vol. 106, n.º 23, pp. 9180-9184, 2009.

- 40 En una realización, la proteasa no citotóxica es una proteasa BoNT/A, BoNT/C o BoNT/E, y el motivo SNARE preferido es un motivo SNAP (por ejemplo, SNAP 25). En otra realización, la proteasa no citotóxica es una proteasa BoNT/B, BoNT/D, BoNT/F o BoNT/G o de la neurotoxina del tétanos (TeNT), y el motivo SNARE preferido es un motivo VAMP. En otra realización, la proteasa no citotóxica es una proteasa BoNT/C<sub>1</sub>, y el motivo SNARE preferido es un motivo de sintaxina.

- 45 Los polipéptidos de la presente invención, en especial el componente de proteasa de estos, pueden estar PEGilados; esto puede ayudar a aumentar la estabilidad, por ejemplo, la duración de la acción del componente de proteasa. La PEGilación se prefiere en particular cuando la proteasa comprende una proteasa BoNT/A, B o C<sub>1</sub>. La PEGilación preferiblemente incluye la adición de PEG al N-terminal del componente de proteasa. Como ejemplo, el N-terminal de una proteasa puede extenderse con uno o más restos aminoácidos (por ejemplo, cisteína), que pueden ser iguales o diferentes. Uno o más de dichos restos aminoácidos pueden tener su propia molécula de PEG unida (por ejemplo, unida covalentemente) a él. Un ejemplo de esta tecnología se describe en el documento WO2007/104567.

Un dominio de translocación es una molécula que permite la translocación de una proteasa hacia el interior de una célula diana, de modo que se produce una expresión funcional de la actividad proteasa dentro del citosol de la célula diana. Si cualquier molécula (por ejemplo, una proteína o un péptido) posee la función de translocación necesaria de la presente invención es algo que se puede confirmar mediante cualquiera de una serie de ensayos convencionales.

- 5 Por ejemplo, Shone C. (1987), describe un ensayo *in vitro* que emplea liposomas, que son expuestos a una molécula de ensayo. La presencia de la función de translocación necesaria se confirma mediante la liberación desde los liposomas de  $K^+$  y/o NAD marcado, que puede controlarse con facilidad [véase Shone C. (1987), Eur. J. Biochem., vol. 167(1): pp. 175-180]. Otro ejemplo lo proporciona Blaustein R. (1987), que describe un ensayo *in vitro* simple que emplea membranas de bicapas de fosfolípidos planas. Las membranas se exponen a una molécula de ensayo y se confirma la función de translocación necesaria mediante un aumento en la conductancia a través de dichas membranas [véase Blaustein (1987) FEBS Letts., vol. 226, n.º 1: pp. 115-120]. Otra metodología para permitir la evaluación de la fusión de membranas y, por tanto, la identificación de dominios de translocación adecuados para su 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.
- 10
- 15 La presente invención también incluye variantes de dominios de translocación, con la condición de que los variantes de los dominios sigan demostrando la actividad de translocación necesaria. Como ejemplo, un variante puede tener al menos 70%, preferiblemente al menos 80%, más preferiblemente al menos 90%, y lo más preferiblemente al menos 95% o al menos 98% de homología de secuencia de aminoácidos con un dominio de translocación de referencia. El término fragmento, cuando se emplea en relación con un dominio de translocación, significa un péptido que tiene al menos 20, preferiblemente al menos 40, más preferiblemente al menos 80, y lo más preferiblemente al menos 100 restos aminoácidos del dominio de translocación de referencia. En el caso de un dominio de translocación clostridial, el fragmento preferiblemente presenta al menos 100, preferiblemente al menos 150, más preferiblemente al menos 200, y lo más preferiblemente al menos 250 restos aminoácidos del dominio de translocación de referencia (por ejemplo, el dominio  $H_N$ ). Con respecto al componente del 'fragmento' de TM (tal como se analizó anteriormente), los 'fragmentos' de translocación de la presente invención incluyen fragmentos de variantes de dominios de translocación basados en las secuencias de referencia.
- 20
- 25

Se ha documentado que ciertos dominios de moléculas de toxinas bacterianas son capaces de formar poros. También se sabe que ciertos dominios de translocación de proteínas de fusión de membranas viralmente expresados son capaces de formar estos poros. Estos dominios pueden emplearse en la presente invención.

- 30 El dominio de translocación puede tener un origen clostridial, tal como el dominio  $H_N$  (o uno de sus componentes funcionales).  $H_N$  significa una porción o fragmento de la cadena H de una neurotoxina clostridial que es aproximadamente equivalente a la mitad amino-terminal de la cadena H, o el dominio correspondiente a ese fragmento en la cadena H intacta. A este respecto, puede resultar deseable eliminar la función de unión a la célula de  $H_C$ , esto se puede hacer mediante la delección de la secuencia de aminoácidos de  $H_C$  o  $H_{CC}$  (al nivel de síntesis de ADN, o al nivel de postsíntesis mediante un tratamiento con una nucleasa o proteasa). Como alternativa, la función  $H_C$  puede inactivarse mediante un tratamiento químico o biológico.
- 35

Los ejemplos de dominios de translocación (de referencia) adecuados incluyen:

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

La secuencia de referencia identificada anteriormente debe considerarse una guía, puesto que pueden producirse ligeras variaciones según los subserotipos. Como ejemplo, el documento US 2007/0166332 cita secuencias clostridiales ligeramente diferentes:

- neurotoxina botulínica de tipo A - restos aminoácidos (A449-K871)
- 50 neurotoxina botulínica de tipo B - restos aminoácidos (A442-S858)
- neurotoxina botulínica de tipo C - restos aminoácidos (T450-N866)

neurotoxina botulínica de tipo D - restos aminoácidos (D446-N862)

neurotoxina botulínica de tipo E - restos aminoácidos (K423-K845)

neurotoxina botulínica de tipo F - restos aminoácidos (A440-K864)

neurotoxina botulínica de tipo G - restos aminoácidos (S447-S863)

5 neurotoxina del tétanos - restos aminoácidos (S458-V879)

En el contexto de la presente invención, en aspectos de la presente invención pueden ser útiles una diversidad de regiones H<sub>N</sub> de toxinas clostridiales que comprenden un dominio de translocación, con la condición que estos fragmentos activos puedan facilitar la liberación de una proteasa no citotóxica (por ejemplo, una cadena L clostridial) desde vesículas intracelulares hacia el citoplasma de la célula diana y, así, participar en la ejecución del mecanismo celular global mediante el cual una toxina clostridial rompe proteolíticamente un sustrato. Las regiones H<sub>N</sub> de las cadenas pesadas de las toxinas clostridiales tienen una longitud de aproximadamente 410-430 aminoácidos y comprenden un dominio de translocación. Las investigaciones han demostrado que no es necesaria la longitud completa de la región H<sub>N</sub> de la cadena pesada de una toxina clostridial para la actividad de translocación del dominio de translocación. Así, los aspectos de esta realización pueden incluir regiones H<sub>N</sub> de toxinas clostridiales que comprenden un dominio de translocación que tiene 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<sub>N</sub> de toxinas clostridiales que comprenden un dominio de translocación que tiene una longitud, por ejemplo, de un máximo de 350 aminoácidos, un máximo de 375 aminoácidos, un máximo de 400 aminoácidos, y un máximo de 425 aminoácidos.

Para más detalles acerca de la base genética de la producción de toxinas en *Clostridium botulinum* y *C. tetani*, se remite a Henderson et al. (1997), en *The Clostridia: Molecular Biology and Pathogenesis*, Academic Press. El término H<sub>N</sub> incluye porciones H<sub>N</sub> de neurotoxinas naturales y porciones H<sub>N</sub> modificadas que tienen secuencias de aminoácidos que no aparecen en la naturaleza y/o restos aminoácidos sintéticos, con la condición de que las porciones de H<sub>N</sub> modificadas aún muestren la función de translocación mencionada anteriormente. Como alternativa, el dominio de translocación puede no tener un origen clostridial. Los ejemplos de orígenes de dominios de translocación que no son clostridiales (de referencia) incluyen, pero no se limitan al dominio de translocación de la toxina de la difteria [O'Keefe et al., *Proc. Natl. Acad. Sci. USA* (1992), 89, 6202-6206; Silverman et al., *J. Biol. Chem.* (1993), 269, 22524-22532; y London, E. (1992), *Biochem. Biophys. Acta.*, 1112, pp.25-51], los dominios de translocación de la exotoxina de tipo A de *Pseudomonas* [Prior et al., *Biochemistry* (1992), 31, 3555-3559], los dominios de translocación de la toxina del ántrax [Blanke et al., *Proc. Natl. Acad. Sci. USA* (1996), 93, 8437-8442], una diversidad de péptidos fusogénicos o hidrófobos con función de translocación [Plank et al., *J. Biol. Chem.* (1994), 269, 12918-12924; y Wagner et al. (1992), *PNAS*, 89, pp. 7934-7938], y péptidos anfífilicos [Murata et al. (1992), *Biochem.*, 31, pp.1986-1992]. El dominio de translocación puede reflejar el dominio de translocación presente en una proteína natural, o puede incluir variaciones en aminoácidos, con la condición de que las variaciones no destruyan la capacidad de translocación del dominio de translocación.

Los ejemplos concretos de dominios de translocación víricos (de referencia) adecuados para su uso en la presente invención incluyen ciertos dominios de translocación de proteínas de fusión de membranas expresadas víricamente. Por ejemplo, Wagner et al. (1992), y Murata et al. (1992) describen la función de translocación (es decir, la fusión de membranas y la formación de vesículas) de una serie de péptidos fusogénicos y anfífilicos derivados de la región N-terminal de la hemaglutinina del virus de la gripe. Otras proteínas de fusión de membranas víricamente expresadas conocidas por tener la actividad de translocación deseada son un dominio de translocación de un péptido fusogénico del virus del bosque de Semliki (SFV), un dominio de translocación de la glicoproteína 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 glicoproteína de la envuelta del virus espumoso. Las proteínas Aspíke víricamente codificadas tienen una aplicación concreta en el contexto de la presente invención, por ejemplo, la proteína E1 de SFV y la proteína G de VSV.

El uso de dominios de translocación (de referencia) incluye el uso de sus variantes de secuencia. Un variante puede comprender una o más sustituciones de aminoácidos conservativas y/o deleciones o inserciones de ácidos nucleicos, con la condición de que el variante posea la función de translocación necesaria. Un variante también puede comprender una o más sustituciones de aminoácidos y/o deleciones o inserciones de aminoácidos, con la condición de que el variante posea la función de translocación necesaria.

Los polipéptidos de la presente invención pueden comprender también un dominio facilitador de la translocación. Dicho dominio facilita el transporte de la proteasa no citotóxica hacia el citosol de la célula diana y se describe, por ejemplo, en los documentos WO 08/008803 y WO 08/008805.

Como ejemplo, los dominios facilitadores de la translocación adecuados incluyen un dominio de péptido fusogénico de virus envuelto, por ejemplo, los dominios de péptidos fusogénicos adecuados incluyen el dominio de péptido fusogénico del virus de la gripe (por ejemplo, el dominio de péptido fusogénico del virus de la gripe A de 23 aminoácidos), el dominio de péptido fusogénico de alfavirus (por ejemplo, el dominio de péptido fusogénico del virus del bosque de Semliki de 26 aminoácidos), el dominio de péptido fusogénico de vesiculovirus (por ejemplo, el

- dominio de péptido fusogénico del virus de la estomatitis vesicular de 21 aminoácidos), el dominio de péptido fusogénico de respirovirus (por ejemplo, el dominio de péptido fusogénico del virus de Sendai de 25 aminoácidos), el dominio de péptido fusogénico de morbilivirus (por ejemplo, el dominio de péptido fusogénico del virus del moquillo canino de 25 aminoácidos), el dominio de péptido fusogénico de avulavirus (por ejemplo, el dominio de péptido fusogénico del virus de la enfermedad de Newcastle de 25 aminoácidos), el dominio de péptido fusogénico de henipavirus (por ejemplo, el dominio de péptido fusogénico del virus Hendra de 25 aminoácidos), el dominio de péptido fusogénico de metapneumovirus (por ejemplo, el dominio de péptido fusogénico del metapneumovirus humano de 25 aminoácidos) o el dominio de péptido fusogénico de espumavirus, tal como el dominio de péptido fusogénico del virus espumoso de simio; o sus fragmentos o variantes.
- 5
- 10 Como otro ejemplo, un dominio facilitador de la translocación puede comprender un dominio  $H_{CN}$  de toxina clostridial o uno de sus fragmentos o variantes. Con más detalle, un dominio facilitador de la translocación de  $H_{CN}$  de una toxina clostridial puede tener una longitud de al menos 200 aminoácidos, al menos 225 aminoácidos, al menos 250 aminoácidos, al menos 275 aminoácidos. A este respecto, un dominio facilitador de la translocación de  $H_{CN}$  de una toxina clostridial preferiblemente tiene una longitud de un máximo de 200 aminoácidos, un máximo de 225 aminoácidos, un máximo de 250 aminoácidos, o un máximo de 275 aminoácidos. Los ejemplos específicos (de referencia) incluyen:
- 15
- neurotoxina botulínica de tipo A - restos aminoácidos (872-1110)
- neurotoxina botulínica de tipo B - restos aminoácidos (859-1097)
- neurotoxina botulínica de tipo C - restos aminoácidos (867-1111)
- 20
- neurotoxina botulínica de tipo D - restos aminoácidos (863-1098)
- neurotoxina botulínica de tipo E - restos aminoácidos (846-1085)
- neurotoxina botulínica de tipo F - restos aminoácidos (865-1105)
- neurotoxina botulínica de tipo G - restos aminoácidos (864-1105)
- neurotoxina del tétanos - restos aminoácidos (880-1127)
- 25
- Las anteriores posiciones en la secuencia pueden variar un poco según el serotipo/subtipo, y otros ejemplos de dominios  $H_{CN}$  de toxina clostridial (de referencia) adecuados incluyen:
- neurotoxina botulínica de tipo A - restos aminoácidos (874-1110)
- neurotoxina botulínica de tipo B - restos aminoácidos (861-1097)
- neurotoxina botulínica de tipo C - restos aminoácidos (869-1111)
- 30
- neurotoxina botulínica de tipo D - restos aminoácidos (865-1098)
- neurotoxina botulínica de tipo E - restos aminoácidos (848-1085)
- neurotoxina botulínica de tipo F - restos aminoácidos (867-1105)
- neurotoxina botulínica de tipo G - restos aminoácidos (866-1105)
- neurotoxina del tétanos - restos aminoácidos (882-1127)
- 35
- Cualquiera de los dominios facilitadores descritos anteriormente puede combinarse con cualquiera de los péptidos de dominio de translocación previamente descritos que son adecuados para su uso en la presente invención. Así, como ejemplo, un dominio facilitador que no sea clostridial puede combinarse con un péptido de dominio de translocación que no sea clostridial o con un péptido de dominio de translocación clostridial. Como alternativa, un dominio facilitador de la translocación de  $H_{CN}$  de una toxina clostridial puede combinarse con un péptido de dominio de translocación no clostridial. Como alternativa, un dominio facilitador de  $H_{CN}$  de una toxina clostridial puede combinarse con un péptido de dominio de translocación clostridial, cuyos ejemplos incluyen:
- 40
- neurotoxina botulínica de tipo A - restos aminoácidos (449-1110)
- neurotoxina botulínica de tipo B - restos aminoácidos (442-1097)
- neurotoxina botulínica de tipo C - restos aminoácidos (450-1111)
- 45
- neurotoxina botulínica de tipo D - restos aminoácidos (446-1098)
- neurotoxina botulínica de tipo E - restos aminoácidos (423-1085)

neurotoxina botulínica de tipo F - restos aminoácidos (440-1105)

neurotoxina botulínica de tipo G - restos aminoácidos (447-1105)

neurotoxina del tétanos - restos aminoácidos (458-1127)

**Homología de secuencia**

5 Puede utilizarse cualquiera de una diversidad de métodos de alineamiento de secuencias para determinar el porcentaje de identidad que incluyen, sin limitación, métodos globales, métodos locales y métodos híbridos, tales como, por ejemplo, métodos de estrategias de segmentos. Los protocolos para determinar el porcentaje de identidad son procedimientos rutinarios dentro del alcance de los expertos en la técnica. Los métodos globales alinean secuencias desde el principio al final de la molécula y determinan el mejor alineamiento añadiendo puntuaciones de parejas de restos individuales e imponiendo penalizaciones de huecos. Los métodos no limitantes incluyen, por ejemplo, CLUSTAL W, véase, por ejemplo, Julie D. Thompson et al., CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment Through Sequence Weighting, Position- Specific Gap Penalties and Weight Matrix Choice, 22(22), Nucleic Acids Research, 4673-4680 (1994); y el 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, 264(4), J. Mol. Biol., 823-838 (1996). Los métodos locales alinean secuencias identificando uno o más motivos conservados compartidos por todas las secuencias de entrada. Los métodos no limitantes incluyen, por ejemplo, Match-box, véase, por ejemplo, Eric Depiereux and Ernest Feytmans, Match-Box: A Fundamentally New Algorithm for the Simultaneous Alignment of Several Protein Sequences, 8(5), CABIOS, 501-509 (1992); muestreo de Gibbs, véase, por ejemplo, C. E. Lawrence et al., Detecting Subtle Sequence Signals: A Gibbs Sampling Strategy for Multiple Alignment, 262(5131), Science, 208-214 (1993); Align-M, véase, por ejemplo, Ivo Van Walle et al., Align-M - A New Algorithm for Multiple Alignment of Highly Divergent Sequences, 20(9), Bioinformatics, 1428-1435 (2004). Así, el porcentaje de identidad de secuencia se determina por métodos convencionales. Véase, por ejemplo, Altschul et al., Bull. Math. Bio. 48: 603-616, 1986, y Henikoff y Henikoff, Proc. Natl. Acad. Sci. USA, 89:10915-10919, 1992. Brevemente, dos secuencias de aminoácidos se alinean para optimizar las puntuaciones de alineamiento utilizando una penalización de apertura de hueco de 10, una penalización de extensión de hueco de 1, y la matriz de puntuación "Blosum 62" de Henikoff y Henikoff.

Los polipéptidos sustancialmente homólogos se caracterizan por tener una o más sustituciones, deleciones o adiciones de aminoácidos. Estos cambios preferiblemente son de naturaleza pequeña, es decir, sustituciones de aminoácidos conservativas (véase a continuación) y otras sustituciones que no afecten significativamente al plegamiento o la actividad del polipéptido; deleciones pequeñas, generalmente de uno a aproximadamente 30 aminoácidos; y pequeñas extensiones amino- o carboxilo-terminales, tales como un resto metionina amino-terminal, un péptido conector pequeño de hasta aproximadamente 20-25 restos, o un marcador de afinidad.

**Sustituciones de aminoácidos conservativas**

Básicos:	arginina; lisina; histidina
Ácidos:	ácido glutámico; ácido aspártico
Polares:	glutamina; asparagina
Hidrófobos:	leucina; isoleucina; valina
Aromáticos:	fenilalanina; triptófano; tirosina
Pequeños:	glicina; alanina; serina; treonina; metionina

35 Además de los 20 aminoácidos convencionales, los restos aminoácidos de los polipéptidos de la presente invención pueden ser sustituidos por aminoácidos no convencionales (tales como 4-hidroxiprolina, 6-N-metil-lisina, ácido 2-aminoisobutírico, isovalina y  $\alpha$ -metilserina). Los restos aminoácidos del polipéptido clostridial pueden ser sustituidos por un número limitado de aminoácidos no conservativos, aminoácidos que no son codificados por el código genético y aminoácidos no naturales. Los polipéptidos de la presente invención también pueden comprender restos aminoácidos que no aparecen en la naturaleza.

Los aminoácidos que no aparecen en la naturaleza 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 métodos para incorporar restos aminoácidos que no aparecen en la naturaleza en proteínas. Por ejemplo, puede emplearse un sistema *in vitro* en el que las mutaciones sin sentido son suprimidas utilizando ARNt supresores químicamente aminoacilados. Los métodos para sintetizar aminoácidos y aminoacilar ARNt son conocidos en la técnica. La transcripción y la traducción de plásmidos que contienen mutaciones sin sentido se realiza en un sistema sin células que comprende un extracto de *E. coli* S30 y enzimas disponibles en el mercado y otros reactivos. Las proteínas se purifican mediante cromatografía. Véase, por ejemplo, Robertson et al., J. Am. Chem. Soc., 113:2722, 1991; Ellman et al., Methods Enzymol., 202:301, 1991; Chung et al., Science, 259:806-809, 1993; y Chung et al., Proc. Natl. Acad. Sci. USA, 90:10145-10149, 1993). En un segundo método, la traducción se realiza en oocitos de *Xenopus* mediante microinyección de ARNm mutado y ARNt supresores químicamente aminoacilados (Turcatti et al., J. Biol. Chem., 271:19991-19998, 1996). Dentro de un tercer método, se cultivan células de *E. coli* en ausencia de un aminoácido natural que se va a reemplazar (por ejemplo, fenilalanina) y en presencia de los aminoácidos que no aparecen en la naturaleza deseados (por ejemplo, 2-azafenilalanina, 3-azafenilalanina, 4-azafenilalanina, o 4-fluorofenilalanina). El aminoácido que no aparece en la naturaleza se incorpora en el polipéptido en lugar de su homólogo natural. Véase, Koide et al., Biochem., 33:7470-7476, 1994. Los restos aminoácidos que aparecen en la naturaleza pueden convertirse en especies que no aparecen en la naturaleza mediante una modificación química *in vitro*. La modificación química puede combinarse con una mutagénesis específica dirigida a sitio para expandir aún más la gama de sustituciones (Wynn y Richards, Protein Sci., 2:395-403, 1993).

Los restos aminoácidos de los polipéptidos de la presente invención pueden ser sustituidos por un número limitado de aminoácidos no conservativos, aminoácidos que no son codificados por el código genético, aminoácidos que no aparecen en la naturaleza y aminoácidos no naturales.

Los aminoácidos esenciales en los polipéptidos de la presente invención pueden identificarse según procedimientos conocidos en la técnica, tales como la mutagénesis específica dirigida a sitio o la mutagénesis de barrido de alanina (Cunningham y Wells, Science, 244: 1081-1085, 1989). Los sitios de interacción biológica también pueden determinarse mediante el análisis físico de la estructura, según se determina mediante técnicas como la resonancia magnética nuclear, la cristalografía, la difracción de electrones o el marcaje de fotoafinidad, junto con la mutación de aminoácidos de sitio de contacto putativos. Véase, por ejemplo, de Vos et al., Science, 255:306-312, 1992; Smith et al., J. Mol. Biol., 224:899-904, 1992; Wlodaver et al., FEBS Lett., 309:59-64, 1992. Las identidades de los aminoácidos esenciales también pueden inferirse del análisis de las homologías con componentes relacionados (por ejemplo, los componentes de translocación o de proteasa) de los polipéptidos de la presente invención.

Pueden realizarse y ensayarse múltiples sustituciones de aminoácidos utilizando métodos conocidos de mutagénesis y selección, tales como los descritos por Reidhaar-Olson y Sauer (Science, 241:53-57, 1988) o Bowie y Sauer (Proc. Natl. Acad. Sci. USA, 86:2152-2156, 1989). Brevemente, estos autores describen métodos para aleatorizar simultáneamente dos o más posiciones en un polipéptido, seleccionar un polipéptido funcional, y después secuenciar los polipéptidos mutagenizados para determinar el espectro de sustituciones permitidas en cada posición. Otros métodos que pueden utilizarse incluyen la presentación de fagos (por ejemplo, Lowman et al., Biochem., 30:10832-10837, 1991; Ladner et al., patente de EEUU n.º 5.223.409; Huse, publicación WIPO WO 92/06204) y la mutagénesis específica dirigida a región (Derbyshire et al., Gene, 46:145, 1986; Ner et al., DNA, 7:127, 1988).

Pueden realizarse y ensayarse múltiples sustituciones de aminoácidos utilizando métodos conocidos de mutagénesis y selección, tales como los descritos por Reidhaar-Olson y Sauer (Science, 241:53-57, 1988) o Bowie y Sauer (Proc. Natl. Acad. Sci. USA, 86:2152-2156, 1989). Brevemente, estos autores describen métodos para aleatorizar simultáneamente dos o más posiciones en un polipéptido, seleccionar un polipéptido funcional, y después secuenciar los polipéptidos mutagenizados para determinar el espectro de sustituciones permitidas en cada posición. Otros métodos que pueden utilizarse incluyen la presentación de fagos (por ejemplo, Lowman et al., Biochem., 30:10832-10837, 1991; Ladner et al., patente de EEUU n.º 5.223.409; Huse, publicación WIPO WO 92/06204) y la mutagénesis específica dirigida a región (Derbyshire et al., Gene, 46:145, 1986; Ner et al., DNA, 7:127, 1988).

### Resumen de los ejemplos

**Ejemplo 1** Creación de una proteína de LHD que incorpora un polipéptido de GnRH al C-terminal del dominio H<sub>N</sub>

**Ejemplo 2** Creación de una proteína de LHA que incorpora un polipéptido de GnRH al C-terminal del dominio H<sub>N</sub>

**Ejemplo 3** Creación de una proteína de LHD que incorpora un polipéptido de GnRH al C-terminal del dominio H<sub>N</sub>, en donde se incorporan dos sitios de reconocimiento de proteasas diferentes

**Ejemplo 4** Método de preparación de una proteína de LHD que incorpora un polipéptido de GnRH al C-terminal del dominio H<sub>N</sub>

**Ejemplo 5** Demostración de la presencia de un ligando unido covalentemente mediante análisis de la transferencia Western

**Ejemplo 6** Demostración de la presencia de un TM unido covalentemente mediante espectrometría de masas

**Ejemplo 7** Evaluación de la capacidad de unión de una proteína de LHD que incorpora un polipéptido de GnRH

**Ejemplo 8** Evaluación de la funcionalidad *in vitro* de una proteína de LHD que incorpora un polipéptido de GnRH

5 **Ejemplo 9** Creación de una proteína de LHD que incorpora un polipéptido de dinorfina y de bradiquinina al C-terminal del dominio H<sub>N</sub>

**Ejemplo 10** Creación de una proteína de LHA que incorpora un polipéptido de beta-endorfina y de bradiquinina al C-terminal del dominio H<sub>N</sub>

**Ejemplo 11** Creación de una proteína de LHD que incorpora dos polipéptidos de GHRH al C-terminal del dominio H<sub>N</sub>

10 **Ejemplo 12** Creación de una proteína de LHD que incorpora un polipéptido de GnRH al C-terminal del dominio H<sub>N</sub>, separado por 5 aminoácidos del segundo sitio de activación de proteasas

**Ejemplo 13** Creación de una proteína de LHA que incorpora un polipéptido liberador de gastrina al C-terminal del dominio H<sub>N</sub>

**Ejemplo 14** Método para tratar pacientes que padecen cáncer de próstata

15 **Ejemplo 15** Método para tratar pacientes que padecen una inflamación neurogénica

**Ejemplo 16** Método para tratar pacientes que padecen endometriosis

**Resumen de las figuras**

La figura 1 ilustra un análisis de SDS-PAGE de muestras activadas (eluidas en fracciones de imidazol 80 mM + 250 mM) del ejemplo 4 en condiciones reductoras y no reductoras.

20 **Resumen de SEQ ID NO**

Las siguientes SEQ ID NO pueden excluir cualquier resto aminoácido metionina inicial (o la correspondiente secuencia/codón de ácido nucleico N-terminal).

SEQ ID 1

Secuencia de ADN de LHD-GnRH

25 SEQ ID 2

Secuencia de proteína de LHD-GnRH

SEQ ID 3

Secuencia de ADN de LHA-GnRH

SEQ ID 4

30 Secuencia de proteína de LHA-GnRH

SEQ ID 5

Secuencia de ADN de LHD-GnRH con dos sitios de proteasa diferentes

SEQ ID 6

Secuencia de proteína de LHD-GnRH con dos sitios de proteasa diferentes

35 SEQ ID 7

Secuencia de ADN de LHD que incorpora un polipéptido de dinorfina y de bradiquinina al C-terminal del dominio H<sub>N</sub>

SEQ ID 8

Secuencia de proteína de LHD que incorpora un polipéptido de dinorfina y de bradiquinina al C-terminal del dominio H<sub>N</sub>

40

SEQ ID 9

Secuencia de ADN de LHA que incorpora un polipéptido de beta-endorfina y de bradiquinina al C-terminal del dominio H<sub>N</sub>

SEQ ID 10

5 Secuencia de proteína de LHA que incorpora un polipéptido de beta-endorfina y de bradiquinina al C-terminal del dominio H<sub>N</sub>

SEQ ID 11

Secuencia de ADN de LHD que incorpora dos polipéptidos de GHRH al C-terminal del dominio H<sub>N</sub>

SEQ ID 12

10 Secuencia de proteína de LHD que incorpora dos polipéptidos de GHRH al C-terminal del dominio H<sub>N</sub>

SEQ ID 13

Secuencia de ADN de LHD que incorpora GnRH al C-terminal del dominio H<sub>N</sub>

SEQ ID 14

Secuencia de proteína de LHD que incorpora GnRH al C-terminal del dominio H<sub>N</sub>

15 SEQ ID 15

Secuencia de ADN de LHA que incorpora un péptido liberador de gastrina al C-terminal del dominio H<sub>N</sub>

SEQ ID 16

Secuencia de proteína de LHA que incorpora un péptido liberador de gastrina al C-terminal del dominio H<sub>N</sub>

20 A continuación se ofrece la descripción de realizaciones específicas de la invención, ilustradas mediante los ejemplos.

**Ejemplo 1: Creación de una proteína de LHD que incorpora un polipéptido de GnRH al C-terminal del dominio H<sub>N</sub>**

25 Se analiza la secuencia primaria de una proteína quimérica construida mediante una fusión genética del fragmento LH<sub>N</sub> de BoNT/D y el péptido de 10 aminoácidos GnRH para detectar la presencia de tramos de aminoácidos que tengan semejanza con el sitio de reconocimiento prototípico para el factor Xa (IEGR). Puesto que no se encuentran tramos de este tipo, se elige emplear FXa como la proteasa para activar la proteína de fusión en la zona de unión LC-H<sub>N</sub> y también para romper el enlace peptídico entre el H<sub>N</sub> y el TM (GnRH).

El ADN optimizado para la expresión en *E. coli* se obtiene en el mercado en Entelechon (Alemania) para codificar una proteína de fusión que tiene la siguiente estructura, desde el N- al C-terminal:

- 30
- marcador de purificación N-terminal de 10 His,
  - un espaciador de 10 aminoácidos de asparagina,
  - la LC de BoNT/D,
  - un conector interdominio con una secuencia primaria similar a la que se encuentra en BoNT/A, modificado para que incorpore el tetrapéptido IEGR que es un sustrato para FXa,
- 35
- el H<sub>N</sub> de BoNT/D modificado para incorporar una Cys C-terminal,
  - un espaciador de Gly-Gly-Gly-Gly-Ser que incorpora una secuencia del péptido IEGR en el C-terminal,
  - un péptido GnRH de 10 aminoácidos modificado para incorporar un resto Cys en la posición 6 en lugar de la Gly natural (QHWSYCLRPG).

40 Se evaluó la utilización de codones de *E. coli* remitiéndose a programas informáticos, tales como Graphical Codon Usage Analyser (Geneart), y se evaluó el contenido en GC% y la proporción de utilización de codones remitiéndose a tablas de utilización de codones publicadas (por ejemplo, GenBank Release 143, 13 de septiembre de 2004) para asegurarse de que la construcción no incluya una mala utilización de codones. El ADN se incorporó en un vector de clonación convencional, por ejemplo pCR4, antes de la transformación en el hospedante de *E. coli*. La integridad del

ORF del ADN se comprobó mediante secuenciación. El ORF final se ilustra como SEQ ID 1, y la secuencia de aminoácidos del producto de la expresión se ilustra en SEQ ID 2.

**Ejemplo 2: Creación de una proteína de LHA que incorpora un polipéptido de GnRH al C-terminal del dominio H<sub>N</sub>**

- 5 Se analiza la secuencia primaria de una proteína quimérica construida mediante una fusión genética del fragmento LH<sub>N</sub> de BoNT/A y el péptido de 10 aminoácidos GnRH para detectar la presencia de tramos de aminoácidos que tengan semejanza con el sitio de reconocimiento prototípico para el factor Xa (IEGR). Puesto que no se encuentran tramos de este tipo, se elige emplear FXa como la proteasa para activar la proteína de fusión en la zona de unión LC-H<sub>N</sub> y también para romper el enlace peptídico entre el H<sub>N</sub> y el TM (GnRH).
- 10 El ADN optimizado para la expresión en *E. coli* se obtiene en el mercado en Entelechon (Alemania) para codificar una proteína de fusión que tiene la siguiente estructura, desde el N- al C-terminal:
- marcador de purificación N-terminal de 10 His,
  - un espaciador de 10 aminoácidos de asparagina,
  - la LC de BoNT/A,
  - 15 • un conector interdominio con una secuencia primaria similar a la que se encuentra en BoNT/A, modificado para que incorpore el tetrapéptido IEGR que es un sustrato para FXa,
  - el H<sub>N</sub> de BoNT/A modificado para incorporar una Cys C-terminal,
  - un espaciador de Gly-Gly-Gly-Gly-Ser que incorpora una secuencia del péptido IEGR en el C-terminal,
  - un péptido GnRH de 10 aminoácidos modificado para incorporar un resto Cys en la posición 6 en lugar de la Gly natural (QHWSYCLRPG).

20 Se evaluó la utilización de codones de *E. coli* remitiéndose a programas informáticos, tales como Graphical Codon Usage Analyser (Geneart), y se evaluó el contenido en GC% y la proporción de utilización de codones remitiéndose a tablas de utilización de codones publicadas (por ejemplo, GenBank Release 143, 13 de septiembre de 2004) para asegurarse de que la construcción no incluya una mala utilización de codones. El ADN se incorporó en un vector de clonación convencional, por ejemplo pCR4, antes de la transformación en el hospedante de *E. coli*. La integridad del ORF del ADN se comprobó mediante secuenciación. El ORF final se ilustra como SEQ ID 3, y la secuencia de aminoácidos del producto de la expresión se ilustra en SEQ ID 4.

**Ejemplo 3: Creación de una proteína de LHD que incorpora un polipéptido de GnRH al C-terminal del dominio H<sub>N</sub>, en donde se incorporan dos sitios de reconocimiento de proteasas diferentes**

- 30 Se analiza la secuencia primaria de una proteína quimérica construida mediante una fusión genética del fragmento LH<sub>N</sub> de BoNT/D y el péptido de 10 aminoácidos GnRH para detectar la presencia de tramos de aminoácidos que tengan semejanza con el sitio de reconocimiento prototípico para el factor Xa (IEGR) y la enteroquinasa (DDDDK). Puesto que no se encuentran tramos de este tipo, se elige emplear FXa como la proteasa para activar la proteína de fusión en la zona de unión LC-H<sub>N</sub> y la enteroquinasa para romper el enlace peptídico entre el H<sub>N</sub> y el TM (GnRH).
- 35 El ADN optimizado para la expresión en *E. coli* se obtiene en el mercado en Entelechon (Alemania) para codificar una proteína de fusión que tiene la siguiente estructura, desde el N- al C-terminal:
- marcador de purificación N-terminal de 10 His,
  - un espaciador de 10 aminoácidos de asparagina,
  - la LC de BoNT/D,
  - 40 • un conector interdominio con una secuencia primaria similar a la que se encuentra en BoNT/A, modificado para que incorpore el tetrapéptido IEGR que es un sustrato para FXa,
  - el H<sub>N</sub> de BoNT/D modificado para incorporar una Cys C-terminal,
  - un espaciador de Gly-Gly-Gly-Gly-Ser que incorpora una secuencia del péptido DDDDK en el C-terminal,
  - un péptido GnRH de 10 aminoácidos modificado para incorporar un resto Cys en la posición 6 en lugar de la Gly natural (QHWSYCLRPG).
- 45

Se evaluó la utilización de codones de *E. coli* remitiéndose a programas informáticos, tales como Graphical Codon Usage Analyser (Geneart), y se evaluó el contenido en GC% y la proporción de utilización de codones remitiéndose a tablas de utilización de codones publicadas (por ejemplo, GenBank Release 143, 13 de septiembre de 2004) para asegurarse de que la construcción no incluya una mala utilización de codones. El ADN se incorporó en un vector de clonación convencional, por ejemplo pCR4, antes de la transformación en el hospedante de *E. coli*. La integridad del ORF del ADN se comprobó mediante secuenciación. El ORF final se ilustra como SEQ ID 5, y la secuencia de aminoácidos del producto de la expresión se ilustra en SEQ ID 6.

#### **Ejemplo 4: Método de preparación de una proteína de LHD que incorpora un polipéptido de GnRH al C-terminal del dominio H<sub>N</sub>**

El ORF creado en el ejemplo 1 se clonó en un vector de expresión de *E. coli* (un vector pET (Novagen) que se había modificado para asegurar una deficiencia en la movilización) y se transformó en una cepa hospedante de *E. coli*, de modo más habitual BL21.

La expresión de la proteína de fusión de LHD-GnRH se logra utilizando el siguiente protocolo. Se inoculan 100 ml de TB modificado que contenía glucosa al 0,2% y ampicilina 100 µg/ml en un matraz de 250 ml con una única colonia procedente de la cepa de expresión de LHD-GnRH. Se hace crecer el cultivo a 37 °C, 225 rpm durante 16 horas. Se inoculan 2x1 l del TB modificado que contenía glucosa al 0,2% y ampicilina 100 µg/ml en un matraz de 2x2 l con 10 ml de cultivo que creció durante la noche. Se hacen crecer los cultivos a 37 °C hasta que se alcanza una OD 600 nm de aproximadamente 0,5, en cuyo momento se reduce la temperatura a 16 °C. Después de 1 hora se inducen los cultivos con IPTG 1 mM y se cultivan a 16 °C durante 16 horas más. La centrifugación del cultivo produce 35,2 g de pasta de células.

La purificación de la fusión de LHD-GnRH se logra mediante una cromatografía de afinidad. En detalle, se descongela un tubo Falcon que contenía 25 ml de HEPES 50 mM, pH 7,2, NaCl 200 mM y aproximadamente 10 g de pasta de células de *E. coli* BL21. Se sonica la pasta de células en hielo en 10 ciclos de 30 segundos de activación, 30 segundos de desconexión, con una potencia de 22 micrómetros, asegurándose de que la muestra se mantiene enfriada. Se centrifugan las células lisadas a 18.000 rpm, 4 °C durante 30 minutos. Se carga el sobrenadante sobre una columna quelante HisTrap HP (es suficiente una columna de 5 ml) equilibrada con HEPES 50 mM, pH 7,2, NaCl 200 mM. Después de la adición de imidazol 40 mM para lavar la proteína unida no específica, la proteína de fusión se eluyó con un gradiente discontinuo de imidazol 80 mM, imidazol 250 mM e imidazol 500 mM. Se dializa la proteína de fusión eluida frente a 5 l de HEPES 50 mM, pH 7,2, NaCl 200 mM a 4 °C durante la noche, y se mide la OD de la proteína de fusión dializada. Se añaden 10 U de proteína de fusión de factor Xa/mg y se incuba a 25 °C sin agitación durante la noche. Se carga sobre una columna quelante HisTrap HP (es suficiente una columna de 5 ml) equilibrada con HEPES 50 mM, pH 7,2, NaCl 200 mM. Se lava la columna hasta la línea de base con HEPES 50 mM, pH 7,2, NaCl 200 mM. Utilizando un gradiente discontinuo de imidazol 10 y 40 mM se lava la proteína unida no específica y se eluye la proteína de fusión con imidazol 100 mM. Se dializa la proteína de fusión eluida frente a 5 l de Tris 25 mM, NaCl 200 mM, pH 8,0 a 4 °C durante la noche y se concentra la fusión hasta aproximadamente 2 mg/ml, se forman partes alícuotas de la muestra y se congela a -20 °C. Se ensaya la proteína purificada empleando análisis de OD, BCA y de pureza.

Las muestras de la proteína activada se analizan mediante SDS-PAGE en condiciones reductoras y no reductoras. Las muestras eluidas en las fracciones de imidazol 80 mM y 250 mM se analizan (véase la figura 1).

#### **Ejemplo 5: Demostración de la presencia de un TM unido covalentemente mediante análisis de la transferencia Western**

La presencia del TM dentro de la proteína de fusión puede evaluarse mediante una diversidad de métodos. Un método consiste en emplear antisueros específicos contra el TM y visualizar el resultado mediante SDS-PAGE y análisis de la transferencia Western. Los anticuerpos contra TM pueden obtenerse en el mercado (por ejemplo, los anticuerpos anti-GnRH están disponibles en Abcam (AB76560) o Novus Biologicals (H00002796-B01), o pueden ser generados específicamente contra una secuencia peptídica concreta por un suministrador de servicios comercial.

Utilizando estas técnicas, se confirma la presencia de GnRH dentro de la proteína de fusión activada de longitud completa cuando se ensaya bajo condiciones no reductoras, pero no su presencia en el dominio H<sub>N</sub> cuando se ensaya bajo condiciones reductoras.

#### **Ejemplo 6: Demostración de la presencia de un TM unido covalentemente mediante espectrometría de masas**

La presencia del TM dentro de la proteína de fusión puede evaluarse mediante una diversidad de métodos. Un método es la utilización de una espectrometría de masas para determinar la masa de la proteína de fusión antes y después de la reducción.

Utilizando la proteína preparada según el ejemplo 4, diversas muestras de la proteína reducida y no reducida se extraen del SDS-PAGE (véase la figura 1) y se analizan mediante una espectrometría de masas (Intertek, Manchester).

La masa predicha de la proteína de fusión no activada y no reducida es de 105271 Da. La masa observada para las muestras fue de 105284 Da, una diferencia de solo 13 Da, que se encuentra dentro del error del equipo. Por tanto, se confirma la presencia de GnRH intacto en la proteína de fusión no activada y no reducida.

Muestra no reducida:			
Masa teórica:	Estructura correspondiente:	Masa observada:	Diferencia de masas:
105271 Da.	Longitud completa	105284 Da.	13 Da.

- 5 La masa predicha de la proteína de fusión activada y no reducida es de 105271 Da. La masa observada para las muestras fue de 105321 Da, una diferencia de solo 50 Da, que se encuentra dentro del error del equipo. Por tanto, se confirma la presencia de GnRH intacto en la proteína de fusión activada y no reducida.

Muestra no reducida:			
Masa teórica:	Estructura correspondiente:	Masa observada:	Diferencia de masas:
105271 Da.	Longitud completa	105321 Da.	50 Da.

- 10 Cuando se evalúan muestras reducidas de LC y del dominio H<sub>N</sub>, el dominio H<sub>N</sub> (que debería comprender H<sub>N</sub> + espaciador + sitio de activación) tiene una masa predicha de 49419 Da y una masa observada de 49421. Esto indica que el dominio H<sub>N</sub> reducido no conserva el péptido de GnRH. Este resultado es igual al predicho, puesto que la proteólisis y la reducción del enlace disulfuro libera la secuencia de GnRH del C-terminal del dominio H<sub>N</sub>.

Muestra reducida: masa de la subunidad Hn			
Masa teórica:	Estructura correspondiente:	Masa observada:	Diferencia de masas:
49419 Da.	Cadena pesada + espaciador + sitio de activación	49421 Da.	2 Da.

- 15 Estos datos demuestran que el ligando de GnRH se une a la proteína de fusión antes de la activación y la reducción, se une a la proteína de fusión después de la activación en ausencia del agente reductor, pero está ausente del dominio H<sub>N</sub> después de la activación y la reducción. Esto confirma que la proteína de fusión se ha activado correctamente en ambos sitios proteolíticos y que el ligando de GnRH está unido al dominio H<sub>N</sub> a través del enlace disulfuro modificado.

20 **Ejemplo 7: Evaluación de la capacidad de unión de una proteína de LHD que incorpora un polipéptido de GnRH**

La proteína preparada según el ejemplo 4 se evalúa para la funcionalidad de la interacción ligando-receptor utilizando uno de una serie de ensayos adecuados. Por ejemplo, los ensayos de unión de ligando-receptor de la hormona liberadora de gonadotropina GnRHR suministrados por Cisbio Bioassays son un ensayo de competición que cuantifica la actividad de unión en una muestra (<http://www.htf.com/products/gpcr/binding/ligands/inserts/C1TT1GNRH.pdf>). Como alternativa, en la bibliografía científica se indica una gama de ensayos de unión de acceso público (por ejemplo, Christopher E. Heise, Susan K. Sullivan y Paul D. Crowe, J. Biomol. Screen, 2007, 12: 235; DOI: 10.1177/1087057106297362). El uso de estos ensayos indica que el TM de GnRH es capaz de interactuar con el receptor diana.

Los datos indican que el TM de GnRH es capaz de interactuar con el receptor diana.

30 **Ejemplo 8: Evaluación de la funcionalidad *in vitro* de una proteína de LHD que incorpora un polipéptido de GnRH**

35 La proteína preparada según el ejemplo 4 se evalúa para determinar su capacidad para romper proteínas SNARE dentro de la célula diana. Brevemente, una línea celular alfa T3-1 (una línea celular gonadotrofa inmortalizada) que expresa altos niveles del receptor de la hormona liberadora de gonadotropina (GnRH) se incubó con un compuesto de la invención. Veinticuatro horas después, el material celular se recolecta y se analizan las proteínas SNARE

mediante una transferencia Western. Los datos indican que la proteína de fusión que comprende el TM de GnRH es capaz de interactuar con el receptor diana, lo cual conduce a la internalización y la ruptura de las proteínas SNARE intracelulares.

5 **Ejemplo 9: Creación de una proteína de LHD que incorpora un polipéptido de dinorfina y de bradiquinina al C-terminal del dominio H<sub>N</sub>**

10 Se analiza la secuencia primaria de una proteína quimérica construida mediante una fusión genética del fragmento LH<sub>N</sub> de BoNT/D y los péptidos dinorfina y bradiquinina para detectar la presencia de tramos de aminoácidos que tengan semejanza con el sitio de reconocimiento prototípico para el factor Xa (IEGR). Puesto que no se encuentran tramos de este tipo, se elige emplear FXa como la proteasa para activar la proteína de fusión en la zona de unión LC-H<sub>N</sub> y también para romper el enlace peptídico entre el H<sub>N</sub> y el péptido de dinorfina. Se construye un espaciador de 11 aminoácidos entre los péptidos de dinorfina y bradiquinina que incorpora una única Cys para facilitar la unión con disulfuro al H<sub>N</sub>.

El ADN optimizado para la expresión en *E. coli* se obtiene en el mercado en Entelechon (Alemania) para codificar una proteína de fusión que tiene la siguiente estructura, desde el N- al C-terminal:

- 15
- marcador de purificación N-terminal de 10 His,
  - un espaciador de 10 aminoácidos de asparagina,
  - la LC de BoNT/D,
  - un conector interdominio con una secuencia primaria similar a la que se encuentra en BoNT/A, modificado para que incorpore el tetrapéptido IEGR que es un sustrato para FXa,
- 20
- el H<sub>N</sub> de BoNT/D modificado para incorporar una Cys C-terminal,
  - un espaciador de Gly-Gly-Gly-Gly-Ser que incorpora una secuencia del péptido IEGR en el C-terminal,
  - un péptido de dinorfina de 17 aminoácidos,
  - un Gly-Gly-Gly-Gly-Ser-Cys-Gly-Gly-Gly-Ser de 11 aminoácidos,
  - un péptido de bradiquinina de 9 aminoácidos.

25 Se evaluó la utilización de codones de *E. coli* remitiéndose a programas informáticos, tales como Graphical Codon Usage Analyser (Geneart), y se evaluó el contenido en GC% y la proporción de utilización de codones remitiéndose a tablas de utilización de codones publicadas (por ejemplo, GenBank Release 143, 13 de septiembre de 2004) para asegurarse de que la construcción no incluya una mala utilización de codones. El ADN se incorporó en un vector de clonación convencional, por ejemplo pCR4, antes de la transformación en el hospedante de *E. coli*. La integridad del ORF del ADN se comprobó mediante secuenciación. El ORF final se ilustra como SEQ ID 7, y la secuencia de aminoácidos del producto de la expresión se ilustra en SEQ ID 8.

30

**Ejemplo 10: Creación de una proteína de LHA que incorpora un polipéptido de beta-endorfina y de bradiquinina al C-terminal del dominio H<sub>N</sub>**

35 Se analiza la secuencia primaria de una proteína quimérica construida mediante una fusión genética del fragmento LH<sub>N</sub> de BoNT/A y los péptidos beta-endorfina y bradiquinina para detectar la presencia de tramos de aminoácidos que tengan semejanza con el sitio de reconocimiento prototípico para el factor Xa (IEGR). Puesto que no se encuentran tramos de este tipo, se elige emplear FXa como la proteasa para activar la proteína de fusión en la zona de unión LC-H<sub>N</sub> y también para romper el enlace peptídico entre el H<sub>N</sub> y el péptido de beta-endorfina. Se construye un espaciador de 11 aminoácidos entre los péptidos de beta-endorfina y bradiquinina que incorpora una única Cys para facilitar la unión con disulfuro al H<sub>N</sub>.

40

El ADN optimizado para la expresión en *E. coli* se obtiene en el mercado en Entelechon (Alemania) para codificar una proteína de fusión que tiene la siguiente estructura, desde el N- al C-terminal:

- 45
- marcador de purificación N-terminal de 10 His,
  - un espaciador de 10 aminoácidos de asparagina,
  - la LC de BoNT/A,
  - un conector interdominio con una secuencia primaria similar a la que se encuentra en BoNT/A, modificado para que incorpore el tetrapéptido IEGR que es un sustrato para FXa,

- el H<sub>N</sub> de BoNT/A modificado para incorporar una Cys C-terminal,
  - un espaciador de Gly-Gly-Gly-Gly-Ser que incorpora una secuencia del péptido IEGR en el C-terminal,
  - un péptido de beta-endorfina de 31 aminoácidos,
  - un Gly-Gly-Gly-Gly-Ser-Cys-Gly-Gly-Gly-Gly-Ser de 11 aminoácidos,
- 5 • un péptido de bradiquinina de 9 aminoácidos.

Se evaluó la utilización de codones de *E. coli* remitiéndose a programas informáticos, tales como Graphical Codon Usage Analyser (Geneart), y se evaluó el contenido en GC% y la proporción de utilización de codones remitiéndose a tablas de utilización de codones publicadas (por ejemplo, GenBank Release 143, 13 de septiembre de 2004) para asegurarse de que la construcción no incluya una mala utilización de codones. El ADN se incorporó en un vector de clonación convencional, por ejemplo pCR4, antes de la transformación en el hospedante de *E. coli*. La integridad del ORF del ADN se comprobó mediante secuenciación. El ORF final se ilustra como SEQ ID 9, y la secuencia de aminoácidos del producto de la expresión se ilustra en SEQ ID 10.

**Ejemplo 11: Creación de una proteína de LHD que incorpora dos polipéptidos de GHRH al C-terminal del dominio H<sub>N</sub>**

Se analiza la secuencia primaria de una proteína quimérica construida mediante una fusión genética del fragmento LH<sub>N</sub> de BoNT/D y dos péptidos GHRH para detectar la presencia de tramos de aminoácidos que tengan semejanza con el sitio de reconocimiento prototípico para el factor Xa (IEGR). Puesto que no se encuentran tramos de este tipo, se elige emplear FXa como la proteasa para activar la proteína de fusión en la zona de unión LC-H<sub>N</sub> y también para romper el enlace peptídico entre el H<sub>N</sub> y el péptido de GHRH. Se construye un espaciador de 11 aminoácidos entre los dos péptidos de GHRH que incorpora una única Cys para facilitar la unión con disulfuro al H<sub>N</sub>.

El ADN optimizado para la expresión en *E. coli* se obtiene en el mercado en Entelechon (Alemania) para codificar una proteína de fusión que tiene la siguiente estructura, desde el N- al C-terminal:

- marcador de purificación N-terminal de 10 His,
  - un espaciador de 10 aminoácidos de asparagina,
- 25 • la LC de BoNT/D,
- un conector interdominio con una secuencia primaria similar a la que se encuentra en BoNT/A, modificado para que incorpore el tetrapéptido IEGR que es un sustrato para FXa,
  - el H<sub>N</sub> de BoNT/D modificado para incorporar una Cys C-terminal,
  - un espaciador de Gly-Gly-Gly-Gly-Ser que incorpora una secuencia del péptido IEGR en el C-terminal,
- 30 • un péptido de GHRH de 40 aminoácidos,
- un Gly-Gly-Gly-Gly-Ser-Cys-Gly-Gly-Gly-Gly-Ser de 11 aminoácidos,
  - un péptido de GHRH de 40 aminoácidos.

Se evaluó la utilización de codones de *E. coli* remitiéndose a programas informáticos, tales como Graphical Codon Usage Analyser (Geneart), y se evaluó el contenido en GC% y la proporción de utilización de codones remitiéndose a tablas de utilización de codones publicadas (por ejemplo, GenBank Release 143, 13 de septiembre de 2004) para asegurarse de que la construcción no incluya una mala utilización de codones. El ADN se incorporó en un vector de clonación convencional, por ejemplo pCR4, antes de la transformación en el hospedante de *E. coli*. La integridad del ORF del ADN se comprobó mediante secuenciación. El ORF final se ilustra como SEQ ID 11, y la secuencia de aminoácidos del producto de la expresión se ilustra en SEQ ID 12.

**Ejemplo 12: Creación de una proteína de LHD que incorpora un polipéptido de GnRH al C-terminal del dominio H<sub>N</sub>, separado por 5 aminoácidos del segundo sitio de activación de proteasas**

Se analiza la secuencia primaria de una proteína quimérica construida mediante una fusión genética del fragmento LH<sub>N</sub> de BoNT/D y el péptido de 10 aminoácidos GnRH para detectar la presencia de tramos de aminoácidos que tengan semejanza con el sitio de reconocimiento prototípico para el factor Xa (IEGR). Puesto que no se encuentran tramos de este tipo, se elige emplear FXa como la proteasa para activar la proteína de fusión en la zona de unión LC-H<sub>N</sub> y también para romper el enlace peptídico entre el H<sub>N</sub> y el espaciador con el N-terminal del TM (GnRH).

El ADN optimizado para la expresión en *E. coli* se obtiene en el mercado en Entelechon (Alemania) para codificar una proteína de fusión que tiene la siguiente estructura, desde el N- al C-terminal:

- marcador de purificación N-terminal de 10 His,
- un espaciador de 10 aminoácidos de asparagina,
- 5 • la LC de BoNT/D,
- un conector interdominio con una secuencia primaria similar a la que se encuentra en BoNT/A, modificado para que incorpore el tetrapéptido IEGR que es un sustrato para FXa,
- el H<sub>N</sub> de BoNT/D modificado para incorporar una Cys C-terminal,
- 10 • un espaciador de Gly-Gly-Gly-Gly-Ser-Ile-Glu-Gly-Arg-Gly-Gly-Gly-Gly-Ser que incorpora un péptido de IEGR,
- un péptido GnRH de 10 aminoácidos modificado para incorporar un resto Cys en la posición 6 en lugar de la Gly natural (QHWSYCLRPG).

Se evaluó la utilización de codones de *E. coli* remitiéndose a programas informáticos, tales como Graphical Codon Usage Analyser (Geneart), y se evaluó el contenido en GC% y la proporción de utilización de codones remitiéndose a tablas de utilización de codones publicadas (por ejemplo, GenBank Release 143, 13 de septiembre de 2004) para asegurarse de que la construcción no incluya una mala utilización de codones. El ADN se incorporó en un vector de clonación convencional, por ejemplo pCR4, antes de la transformación en el hospedante de *E. coli*. La integridad del ORF del ADN se comprobó mediante secuenciación. El ORF final se ilustra como SEQ ID 13, y la secuencia de aminoácidos del producto de la expresión se ilustra en SEQ ID 14.

#### 20 **Ejemplo 13: Creación de una proteína de LHA que incorpora un polipéptido liberador de gastrina al C-terminal del dominio H<sub>N</sub>**

Se analiza la secuencia primaria de una proteína química construida mediante una fusión genética del fragmento LH<sub>N</sub> de BoNT/A y el péptido liberador de gastrina de 27 aminoácidos (GRP) para detectar la presencia de tramos de aminoácidos que tengan semejanza con el sitio de reconocimiento prototípico para el factor Xa (IEGR). Puesto que no se encuentran tramos de este tipo, se elige emplear FXa como la proteasa para activar la proteína de fusión en la zona de unión LC-H<sub>N</sub> y también para romper el enlace peptídico entre el H<sub>N</sub> y el TM (GRP).

El ADN optimizado para la expresión en *E. coli* se obtiene en el mercado en Entelechon (Alemania) para codificar una proteína de fusión que tiene la siguiente estructura, desde el N- al C-terminal:

- marcador de purificación N-terminal de 10 His,
- 30 • un espaciador de 10 aminoácidos de asparagina,
- la LC de BoNT/A,
- un conector interdominio con una secuencia primaria similar a la que se encuentra en BoNT/A, modificado para que incorpore el tetrapéptido IEGR que es un sustrato para FXa,
- el H<sub>N</sub> de BoNT/A modificado para incorporar una Cys C-terminal,
- 35 • un espaciador de Gly-Gly-Gly-Gly-Ser que incorpora una secuencia del péptido IEGR en el C-terminal,
- un péptido liberador de gastrina de 28 aminoácidos modificado para que incorpore un resto Cys en la posición 17 en lugar de la Arg natural, y un resto Gly adicional en el C-terminal para sustituir la necesidad de una amidación C-terminal (VPLPAGGGTVLTKMYPCGNHWAVGHLMG).

Se evaluó la utilización de codones de *E. coli* remitiéndose a programas informáticos, tales como Graphical Codon Usage Analyser (Geneart), y se evaluó el contenido en GC% y la proporción de utilización de codones remitiéndose a tablas de utilización de codones publicadas (por ejemplo, GenBank Release 143, 13 de septiembre de 2004) para asegurarse de que la construcción no incluya una mala utilización de codones. El ADN se incorporó en un vector de clonación convencional, por ejemplo pCR4, antes de la transformación en el hospedante de *E. coli*. La integridad del ORF del ADN se comprobó mediante secuenciación. El ORF final se ilustra como SEQ ID 15, y la secuencia de aminoácidos del producto de la expresión se ilustra en SEQ ID 16.

**Ejemplo 14: Método para tratar pacientes que padecen cáncer de próstata**

Un hombre de 56 años que padece cáncer de próstata avanza hasta una situación en la que la terapia de privación de andrógenos ya no es suficiente para controlar la enfermedad. El hombre se trata mediante la administración local de una composición que comprende un TSI de la presente invención (en este ejemplo concreto, un TSI basado en un TM de un péptido de GnRH) en la vecindad de la próstata. Se controla el trastorno del paciente y, aproximadamente 2 meses después del tratamiento, el médico advierte una disminución en el tamaño tumoral, lo cual indica un tratamiento con éxito con la composición que comprende una molécula de la invención.

**Ejemplo 15: Método para tratar pacientes que padecen una inflamación neurogénica**

Una mujer de 62 años diagnosticada con artritis reumatoide se queja de hinchazón y rigidez en las articulaciones. Un médico determina que la hinchazón y la rigidez en las articulaciones son debidas a una inflamación neurogénica. La mujer se trata mediante la administración local de una composición que comprende un TSI de la presente invención (en este ejemplo, el TSI comprende un TM de opioides; se ensayan ejemplos paralelos con TSI que comprenden TM de nociceptina o dinorfina) en la vecindad del área afectada. Se controla el trastorno del paciente y, después de aproximadamente 1-3 días tras el tratamiento, la mujer indica que se ha reducido la hinchazón y la rigidez en las articulaciones. En los seguimientos de uno y tres meses, la mujer indica que continúa teniendo una menor hinchazón y rigidez en las articulaciones en el área tratada. Esta reducción en los síntomas de la inflamación neurogénica crónica indica un tratamiento con éxito con la composición que comprende una molécula de la invención.

**Ejemplo 16: Método para tratar pacientes que padecen endometriosis**

Una mujer de 39 años se presenta con dolor pélvico debido a una endometriosis que no puede tratarse adecuadamente con fármacos antiinflamatorios no esteroideos (NSAID) y anticonceptivos de estrógeno-progestina combinados. El médico le administra una composición que comprende un TSI de la presente invención (en este ejemplo, el TSI comprende un TM de opioides; se ensayan ejemplos paralelos con TSI que comprenden TM de nociceptina o dinorfina). Se controla el trastorno del paciente y, después de aproximadamente 1-3 días tras el tratamiento, la mujer indica que se ha reducido el dolor. En los seguimientos de uno y tres meses, la mujer indica que continúa teniendo menor dolor y una mayor libertad de movimientos. Esta reducción en los síntomas asociados con la endometriosis indica un tratamiento con éxito con la composición que comprende una molécula de la invención.

**Ejemplo 17: Método para tratar pacientes que padecen vejiga hiperactiva**

Un hombre de 58 años se queja de una mayor urgencia urinaria. Un médico diagnostica al paciente con vejiga hiperactiva con un componente neurológico que implica una actividad neuronal anómala. Al hombre se le trata inyectando uretroscópicamente una composición que comprende un TSI de la presente invención (en este ejemplo, el TSI comprende un TM de opioides; se ensayan ejemplos paralelos con TSI que comprenden TM de nociceptina o dinorfina). Dependiendo de la localización de la actividad neuronal anómala, la toxina puede administrarse, por ejemplo, en el detrusor, el cuello vesical que incluye los esfínteres uretrales externos e internos, el trigono, la cúpula vesical u otras áreas de la pared vesical y/u otras áreas que rodean a la vejiga, tales como la uretra, uréter, diafragma urogenital, músculos pélvicos inferiores, próstata, glándula bulbouretral, bulbo uretral, cruce de pene. Se controla el trastorno del paciente y, después de aproximadamente 1-3 días tras el tratamiento, el hombre indica que tiene una menor urgencia para orinar. En los seguimientos de uno y tres meses, el hombre indica que sigue teniendo menor urgencia para orinar. Esta reducción en los síntomas de una vejiga hiperactiva indica un tratamiento con éxito con la composición que comprende una molécula de la invención.

**Ejemplo 18: Método para tratar pacientes que padecen una inflamación neurogénica**

Una mujer de 62 años diagnosticada con artritis reumatoide se queja de hinchazón y rigidez en las articulaciones. Un médico determina que la hinchazón y la rigidez en las articulaciones son debidas a una inflamación neurogénica. La mujer se trata mediante la administración local de una composición que comprende un TSI de la presente invención en la vecindad del área afectada (en este ejemplo, el TSI comprende un TM de opioides; se ensayan ejemplos paralelos con TSI que comprenden TM de nociceptina o dinorfina). Se controla el trastorno del paciente y, después de aproximadamente 1-3 días tras el tratamiento, la mujer indica que se ha reducido la hinchazón y la rigidez en las articulaciones. En los seguimientos de uno y tres meses, la mujer indica que continúa teniendo una menor hinchazón y rigidez en las articulaciones en el área tratada. Esta reducción en los síntomas de la inflamación neurogénica crónica indica un tratamiento con éxito con la composición que comprende una molécula de la invención. Un tipo similar de administración local de una proteína, según se describe en la presente memoria descriptiva, puede utilizarse para tratar a un paciente que padezca una inflamación neurogénica crónica asociada con cualquier monoartritis, oligoartritis, o poliartritis, tal como, por ejemplo, osteoartritis, artritis idiopática juvenil, artritis séptica, una espondiloartropatía (que incluye la espondilitis anquilosante, artritis reactiva (síndrome de Reiter), artritis psoriática, artritis enteropática asociada con la enfermedad del intestino inflamatoria, enfermedad de Whipple o enfermedad de Behcet), una sinovitis, gota, pseudogota, o enfermedad de Still, así como una bursitis, una fiebre reumática, o una tenosinovitis. Además, también puede utilizarse la administración sistémica para administrar una composición que comprende una molécula de la invención para tratar la inflamación neurogénica crónica.

SEQ ID

SEQ ID 1

atgcatcaccatcaccatcaccatcaccatcatgggagctCGAACAAtAACACAATAACAATAACAAtAA
Cggatccatgacgtggccagttaaggatttcaactactcagatcctgtaaatgacaacgatattctgtacc
ttcgcattccacaaaataaactgatcaccacaccagtcbaagcattcatgattactcaaacatttgggtc
attccagaacgcttttctagtgcacaaaatccgagtttatctaaacctccgctccgacgtccaaatatca
gagctattacgatccctcatatctcagtagcggacgacaaaaagatactttccttaaaggatcattaaac
tgtttaagcgtatataatgagcgcgatatcgggaaaaagttgattaattatcttggttggtgggtcccgctt
atgggagatagctctacccccgaagacacttttgattttaccgctcacaacgacaaacatcgcggtagagaa
gtttgagaacggatcgtggaagtcacaaacatcattacacctagcgtcttaatttttgggtccgctgcaa
acatcttagattatacagccagcctgactttgaggggcaacagtcgaatccgagtttgaaggttttgggt
accctgagcattctgaaagttgccccggaatttctgctcacttttcagatgtcaccagcaaccagagctc
agcagatttaggaaagtcaatttttgcagtgaccgggttatgactgagtcacgaaactgacgcaactctc
tgcatacaactgtatgggatcaaacatccccagtgacaaacgtaattcgtccccaggtgtctgaaggaTTTTc
tcacaggaagggccgaacgtccagttcgaaagagttgtataactttcgaggccctggacgtagagatcattcc
ccagattgagcgcagtcagctgaggaagcattgggccaataaaggataattgcaaacgcctgaaata
acattaacaaaacgattccatcttctggtgactcgaatattgataaataaagaaaatttttagcgagaaa
tataatttgataaagataatacaggttaacttttgggttaacattgacaaattcaactccctttacagtgta
tttgacgaatgtaatgagcgaagttgtgataagttcccaatacaacgtaagaatcgtacccattactctc
ctcgtcactacctgcccgttttgcgaaacatccttgacgataaataattacactattcgtgacggctttaa
ttgaccaacaagggcttcaatattgaaaattcaggccagaacattgaacgcaaccggccttgacagaact
gtcagtgatccctggtttgacctgtttaccaaagtctgctgcgacggcatcattacctccaaaactaaat
ctctgataagaaggtagaaaacaaagcgtgaacctgagtgattaaagtgaaaacaatcggctgccttat
gtgacagataaagatagcatttagtcaggagattttcgaaaaataaattatcactgacgaaaccaatgttca
gaattattcagataaattttcactggacgaaagcattcttagatggccaagtccgattaaccggaaattg
ttgatccgttactgccgaacgtgaatattggaaccgttaaaccctccctggcgaagagatcgtattttatgat
gacattacgaaatattgtagactaccttaattcttattactatttggaaagccagaactgtccaataacgt
ggaaaacattactctgaccacaagcgtggaagaggcttaggctactcaataagattataaccttctcc
cgtcgtggcggaaaaagtaaaaaggtgtgagcgtggtctgttccctcaactggcgaaatgaagttgtc
gaagactttaccagaaatattatgaaaaaggtacacctggataaaaatctccgacgtctcgggtattatccc
atataattggccctcgtttaaatacggtaaatagtgctgctgcccgggaattttaaccaggcctttgctaccg
cgggcgtcgcgttccctcctggagggcttccctgaatttactatcccggcgtcgggtgtttttacattttac
tcttccatccaggagcgtgagaaaaattcaaaaacctcgaaaaactgcctggagcagcgggtgaaacgctg
gaaagattcttcaatggatgggtgtcaaacctgggttatctcgcatcacgaccaattcaaccataatcaat
accagatgtagtagtctgtcgtaccagctgacgaccatcaaaactgagctggaatataaaaaag
tactctggttagcagataaggagaacatcaaaagccaggtggagaaaccttaagaaatagctggtggtgaaaat
ctctgaaagctatgaataacatcaaaaattcattcgtgaatgttcgggtgacgtacctgtcaagaatagc
tgccaaaagttattgatgaaactgaataaatttgatctgctgacaaaaccgaaacttatcaacctcatcgac
tcccacaacattatccttggggcgaagtgatcgtctgaagccaaagtaaacgagagctttgaaaaac
gatgcggtttaaataattttttcatataccaataaactcctgctgaaagatatcatcaatgaatatttcaaTC
TAGAaTGTggcgggtggcggtagcATCGAAGGTCGTcagcactggtcctattgctcgcctcgggttgataa

SEQ ID 2

MHHHHHHHHHSSNNNNNNNNNNGSMTWPKDFNYS DPVNDNDI LYLRI PQNKLI TTPVKAFMITQNIWV
IPERFSSDTNPSLSKPPRPTSKYQSYDPSYLSSTDEQKDTFLKGI IKLFKRINERDIGKKLINYL VVGSPF
MGDSS TPEDTFDFTRHTTNI AVEKFENGSWKV TNI ITPSVLI FGPLPNI LDYASLT LQGQQSNP SFEFGF
TLSILKVAPEFLTFSDVTSNQSSAVLGKSI FCM DPVIALMHELTHSLHQLYGINI PSDKRIRPQVSEGF
SQDGNVQFEELYTFGGLDVEI I PQIERSQLREKALGHYKDI AKRLNNINKTI PSSWI SNIDKYKKI FSEK
YNFDKDN TGNFVNI DKFN SLYSDL TNVMSEVVYSSQYNVKNRTHYFSRHYLPVFANI LDDNI YTI RDGFN
LTNKG FNIE NSGQNI ERN PALQKLSSES VVDLFTKVCVDGI ITS KTKSLI EGRNKALNLQCI KVKNNR LPY
VADKDSI SQEI FENKI ITDET NVQNYSDKFSLDES ILDGQVP INPEI VDP LLPNVNMEPLNL PGEI VFYD
DITKYVDY LNSYYLESQKLSNNVENITL TTSVEEALGYSNKI YTF LPSLAEKVNKGVQAGLFLN WANEV
EDFTTNIMK KDTLDKI SDVSVI I PYIGPALNIGNSALRGNFNQAFATAGVAFLEGFPEFTI PALGVFTFY
SSIQEREKII KTI ENCLEQRV KRKDSYQWMSNWL SRI TTQFNH INYQMYDSL SYQADAI KAKI DLEYKK
YSGSDKENI KSQVENLKN SLDVKI SEAMN NINKFIRECSVTYLFKNMLPKVIDELNKFDLRTKTELINLID
SHNI I LVGEVDRLKAKVNESFENTMPFNI FSYTNSLLKDI INEYFNLECGGGGSI EGRQHWSYCLRPG

5



gagctattacgatccctcatatctcagtagcggacgaacaaaaagatactttccttaaaggatcattaaac
tgtttaagcgtatataatgagcgcgatatcgggaaaaagtgattaattatcttgttgggttccccgttc
atggcgatagctctacccccgaagacacttttgattttaccogtcatagcaaaacatcgcggtagagaa
gtttgagaacggatcgtggaagtcacaaaacattacacctagcgtcttaatttttgggtccgtgcca
acatcttagattatacagccagcctgactttgaggggaacagtcgaatccgagtttcgaaggtttgg
accctgagcattctgaaagttgccccggaatttctgctcacttttccagatgtcaccagcaaccagagctc
agcagattaggaagtcattttttgcatggaccgggttatgactgatgacgaactgacgcactctc
tgcatacaactgtatgggatcaacatcccagtgacaaacgtattcgtccccaggtgtctgaaggattttc
tcacaggatgggcccgaacgtccagttcgaagagttgtatactttcggaggcctggaagtagagatcattc
ccagattgagcgcagtcagctgctgagaaggtcattgggccattataaggatattgcaaacgcctgaata
acattaacaaaaagattccatcttccgtggaatcgaatattgataaataaagaaaatttttagcgagaaa
tataattttgataaagataatacaggttaactttgtggttaacattgacaaaattcaactccctttacagtga
tttgacgaatgtaatgagcgaagttgtgtagttcccaatacaacgtaagaatcgtacccattactctc
ctcgtcactacctgccggttttcgcaacatccttgacgataataattacactattcgtgacggcttaac
ttgaccaacaagggcttcaatattgaaaattcaggccagaacattgaacgcaaccggccttgacagaaact
gtcagtggaatccgtggttgacctgtttaccaaaagtctgctgacggcatcattacctccaaaactaaat
ctctgatagaaggtagaaaacaaagcgtgaacctgacgtgtataaagtgaaaaacaatcggctgccttat
gtagcagataaagatagcattagtcagagattttcgaaaaataaattatcactgacgaaaccaatgttca
gaattattcagataaattttcactggacgaaagcatcttagatggccaagtccgattaacccggaattg
ttgatccggttactgccgaacgtgaatattggaaccttaaacctccctggcgaagagatcgtatttatgat
gacattacgaaatattgtgactaccttaattcttattactatttgaaagccagaactgtccaataacgt
ggaaaaacttactctgaccacaagcgtggaagagcttttaggctactcaataagattataccttctcc
cgtcgtggcggaaaaagtaaaaggtgtgcaggctggctgttctcaactgggcaatgaagttgtc
gaagactttaccacgaatattatgaaaaaggataccctggataaaatctccgacgtctcgggtattatccc
atataattggccctcgttataatcggtaatagtgctgctggggggaattttaaccaggcctttgctaccg
cggcgctcgcgttctcctggagggttctcctgaatttactatcccgccgctcgggtgttttacatttac
tcttccatccaggagcgtgagaaaattatcaaaaccatcgaaaactgcctggagcagcgggtgaaacgctg
gaaagattcttatacaatggatgggtgtcaaacgtggtattctcgcacacgacccaattcaaccatataatt
accagatgatgatagctctgctgacccaagctgacgccattaaagccaaaattgatctggaatataaaaag
tactctggtagcagataaggagaacatcaaaagccaggtggagaaccttaagaaatagctcggatgtgaaaat
ctctgaagctatgaataacattcaaaaattcattcgtgaatgttcgggtgacgtacctgtcaagaatagc
tgccaaaagttattgatgaactgaataaatttgatctgctgacccaaaaccgaaacttatcaacctcatcgac
tcccacaacattatccttgtggcggaagtgatcgtctgaaggccaaaagtaaaagagagctttgaaaatac
gatgcccgtttaaattttttcatataccaataaactccttgctgaaagataatcatcaatgaatatttcaaTC
TAGAAaTGTggcggtggcggttagcGACGATGACGATAAAacagcactggctcctattgcctgcccctggtgga
taa

SEQ ID 6

MHHHHHHHHHGSNNNNNNNNNGSMTWPVKDFNYSDPVNDNDILYLRI PQNKLIITTPVKAFMITQNIWV
IPERFSSDTNPSLSKPPRPTSKYQSYDPSYLSLSTDEQKDTFLKGI IKLFKRINERDIGKKLINYLTVGSPF
MGDSSTPEDTDFTRHTTNI AVEKFENGSWKVTNI ITPSVLIFGFLPNI LDYASLTLQGGQSNPSFEGFG
TLSILKVAPEFLLTFSDVTSNQSSAVLGKSI FCM DPVIALMHELTHSLHQLYGINI PSDKRIRPQVSEGGF
SQDGPVQFEELYTFGGLDVEI I PQIERSQLREKALGHYKDI AKRLNNINKTI PSSWISNIDKYKKI FSEK
YNFDKNTGNFVNI DKFNLSYSDLTNVMSEVVYSSQYNVKNRTHYFSRHYLPVFANILDDNIYT IRDGFN
LTNKGFNIE NSGQNIERNPALQKLSSESVDLFTKVCVDGII TSKTKSLIEGRNKALNLQCI KVKNNRLPY
VADKDSI SQEIFENKI ITDETNVQNYSDKFSLDES ILDGQVPINPEI VDP LLPNVNMEPI NLPGEEI VFYD
DITKYVDYLSNYYLESQLSNNVENITLTT SVEEALGYSNKI YTF LPSLAEKVNKG VQAGLFLNWANEVV
EDFTTNIMKKDTLDKISDVSVI IPYIGPALNIGNSALRGNFNAFATAGVAFLLEGFP EFTI PALGVFTFY
SSIQEREKII KTIENCLEQRVKRWKDSYQWVMSNWSLRI TTQFNHINYQMYDSL SYQADAI KAKIDLEYKK
YSGSDKENIKSQVENLKNSLDVKI SEAMNNINKFI RECSVTYLFKNMLPKVIDELNKF DLRTKTELNLIID
SHNII L VGEVDR LKAKVNESFENTMPFNI FSYTNN SLLKDI INEYFNLECGGGSDDDDKQHWSYCLRPG

SEQ ID 7

atgcatcaccatcaccatcaccatcaccatcatgggagctCGAACAAtACAACAATAACAATAACAAtAA
Cggatccatgacgtggccagtttaaggatttcaactactcagatcctgtaaatgacaacgatattctgtacc
ttcgattcccaaaaataaactgatcaccacaccagtc aaagcattcatgattactcaaaacatttgggtc
attccagaacgcttttctagtgcacaaaatccgagtttatctaaacctccgctccgacgtccaaatatca
gagctattacgatccctcatatctcagtagcggacgaacaaaaagatactttccttaaaggatcattaaac
tgtttaagcgtatataatgagcgcgatatcgggaaaaagttgataattatcttgttgggttccccgttc
atggcgatagctctacccccgaagacacttttgattttaccogtcatagcaaaacatcgcggtagagaa
gtttgagaacggatcgtggaagtcacaaaacattacacctagcgtcttaatttttgggtccgtgcca

5

acatcttagattatacagccagcctgactttgaggggcaacagtcgaatccgagtttgaagggtttgggt
accctgagcattctgaaagttgccccggaatcttctgctcacttttccagatgtcaccagcaaccagagctc
agcagtattaggaaagtcaattttttgcattggaccgggttatctgactgatgcacgaaactgacgcactctc
tgcatacaactgtatgggatcaacatccccagtgacaaacgtatctgccccagggtgtctgaaggattttcc
tcacaggatggggccgaacgtccagttcgaagagttgtatactttcggaggcctggacgtagagatcattcc
ccagattgagcgcagtcagctgctgagaaaggcattgggccattataaggatattgcaaaacgcctgaata
acattaacaaaacgattccatctctgctggatctcgaatattgataaatataagaaaattttagcgagaaa
tataattttgataaaagataatacaggttaacttttggttaacattgacaaaattcaactccctttacagtga
tttgacgaatgtaatgagcgaagttgtgtatagttcccaatacaacgtaagaatcgtaaccattactctct
ctcgtcactacctgcccgttttcgcaacatccttgacgataatatttactactattcgtgacggctttaac
ttgaccaacaagggttcaatattgaaaattcaggccagaacattgaaacgcaacccggccttgcaaaaact
gtcagtgaaatccgtggttgacctgtttacaaaagctcgtcgtcgaaggcatcattacctccaaaactaaat
ctctgataagaagtagaaaacaaagcgtgaacctgacgtgataaaagtgaacaaatcggtgccttat
gtagcagataaaagatagcattagtcaggagattttcgaaaataaaattatcactgacgaaaccaaattgta
gaattattcagataaaatctcactggacgaaagcatcttagatggccaagtccgattaaccgggaaattg
ttgatccgttactcggaaacgtgaataaggaaacggttaaacctcctggcgaagagatcgtattttatgat
gacattacgaaatagtgaggactaccttaattcttattactatttgaaagccagaaactgtccaataacgt
ggaaaacattactctgaccacaagcgtggaagaggctttaggctactcaaaataagatttatacctcctcc
cgtcgtggcggaaaaagtaaaaggtgtgacggctggctcgttccctcaactggcgaaatgaagttgtc
gaagactttaccacgaatattatgaaaaaggataccctggataaaaatctccgacgtctcggttatatccc
atataattggcctcgttcaaatatcggtaaatagtgctgctgctggggaattttaaccaggcctttgctaccg
cggcgtcgttccctcctcctgagggttctcctgaatttactatcccgccgctcggtgtttttacattttac
tcttccatccaggagcgtgagaaaattatcaaaacatcgaaaaactgcctggagcagcggtgaaacgctg
gaaagattcttatcaatggatgggtgtcaaacctgggtatctcgcacacgaccaattcaaccataatatt
accagatgatgatagctctcgtaccaaagctgacgccaataaaagcgaataatgatctggaatataaaaag
tactctggtagcagataaggagaacatcaaaagccaggtggagaaccttaagaa tagtctggatgtgaaaa
ctctgaagctatgaataaacatcaaaaattcattcgtgaaatgctcgtgacgtacctgttcaagaatagc
tgccaaaagttattgatgaactgaataaatttgatctgctgataccaaaaccgaaacttatcaacctcatcgac
tcccacaacattatccttgtggcgaaagtgatcgtctgaaggccaaagtaaacgagagctttgaaaatc
gatgcccgttataatatttttcatataccaataaactccttgcgtaagagatcatcaatgaatatttcaaTC
TAGAaTGTggcgggtggcggtagcATCGAAGGTCGTTATGGAGGTTTTTTGAGAAGGATACGACCAAAATTA
AAGTGGGATAATCAAgcgggtgggggtagtTGCggcgggtggcggttcgcgctccgcccgggtttctctccgtt
ccgttgataa

SEQ ID 8

MHHHHHHHHHGSNNNNNNNNNNGSMTWPKDFNYSDPVNDNDI LYLRI PQNKLITTPVKAFMITQNIWV
IPERFSSDTPNPSLSKPPRPTS KYQSYDPSYLSTDEQKDTFLKGI I KLFKRINERDIGKKLINYL VVGS PF
MGDSS TPEDT FDFTRHTTNI AVEKFENG SWKVTNII T P SVLI FGPLPNI LDYASLT LQGQSNP S FEGFG
T L S I L K V A P E F L L T F S D V T S N Q S S A V L G K S I F C M D P V I A L M H E L T H S L H Q L Y G I N I P S D K R I R P Q V S E G F F
S Q D G P N V Q F E E L Y F F G G L D V E I P Q I E R S Q L R E K A L G H Y K D I A K R L N N I N K T I P S S W I S N I D K Y K K I F S E K
Y N F D K D M T G N F V V N I D K F N S L Y S D L T N V M S E V V Y S Q Y N V K N R T H Y F S R H Y L P V F A N I L D D N I Y T I R D G F N
L T N K G F N I E N S G Q N I E R N P A L Q K L S S E S V V D L F T K V C V D G I I T S K T K S L I E G R N K A L N L Q C I K V K N N R L P Y
V A D K D S I S Q E I F E N K I I T D E T N V Q N Y S D K F S L D E S I L D G Q V P I N P E I V D P L L P N V N M E P L N L P G E E I V F Y D
D I T K Y V D Y L N S Y Y Y L E S Q K L S N N V E N I T L T T S V E E A L G Y S N K I Y T F L P S L A E K V N K G V Q A G L F L N W A N E V V
E D F T T N I M K K D T L D K I S D V S V I I P Y I G P A L N I G N S A L R G N F N Q A F A T A G V A F L L E G F P E F T I P A L G V F T F Y
S S I Q E R E K I I K T I E N C L E Q R V K R W K D S Y Q W M V S N W L S R I T T Q F N H I N Y Q M Y D S L S Y Q A D A I K A K I D L E Y K K
Y S G S D K E N I K S Q V E N L K N S L D V K I S E A M N 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 I L V G E V D 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 L K D I I N E Y F N L E C G G G G S I E G R Y G G F L R R I R P K L
K W D N Q G G G G S C G G G G S R P P G F S P F R

SEQ ID 9

atgcatcaccatcaccatcaccatcaccatcatgggagctCGAACAAtAACAAACAATAACAATAACAAtAA
Cggatccatggagttcgttaacaaacagttcaactataaagaccaggttaacgggtgttgacattgcttaca
tcaaaaatcccgaacgctggccagatgcagccggtaaaagcattcaaaaatccacaacaaaatctgggttatac
ccggaacgtgatacctttactaaccgggaagaaggtgacctgaaccgccaccggaagcgaacaggtgcc
ggtatcttactatgactccacctacctgtctaccgataacgaaaaggacaactacctgaaaggtgttacta
aactgttcgagcgtatcttactccaccgacctggccggtatgctgctgactagcatcgttcgcccgtatcccg
ttctggggcgggttctaccaatcgataccgaaactgaaagtaatcgacactaactgcatcaacggttatcagcc
ggacgggttccatcgttccgaagaactgaaacctgggtatcatcgcccgtctgctgataatcatccagttcg
agtgtaaagactttggtcagaagttctgaacctcaaccgtaacggctacgggtccactcagtaacatccgtt
ttctctccggacttccactcgggtttgaaagaatccctggaagtagacacgaaaccactgctggcggctgg

5

taaattcgcaactgatcctgcggttacctggctcagcaactgattcatgcaggccaccgctgtacggta
tcgccatcaatccgaaccgtgtcttcaaagttaacaccaacgctattacgagatgtccggctctggaagt
agcttcgaagaactgcgtaacttttggcggctcaccgacgctaaattcatcgcactctctgcaagaaaacgagtt
ccgtctgtactactataacaagttcaaaagatctcgcatccaccctgaacaaagcgaatccatcgtgggta
ccactgcttctctccagtagacagaagcgtttttaaagaaaaaacctgctcagcgaagacacctccggc
aaattctctgtagacaagttgaaattcgataaaactttacaaaaatgctgactgaaatttacaccgaagaca
cttcggttaagttctttaaagttctgaaccgcaaaacctatctgaaacttcgacaaggcagattcaaaatca
acatcgtgccgaaagtttaactacactatctacgatggtttcaacctgctgaacaccaacctggctgcta
tttaacggccagaacacggaaatcaacaacatgaacttcacaaaactgaaaaacttactggtctgttcga
gttttacaagctgctgtgctgcagcggcatcattacctccaaaactaaatctctgatagaaggtagaaca
aagcgtcgaacgacctctgtatcaaggttaacaactgggatttattcttcagcccgagtgaagacaacttc
accaacgacctgaacaaaggtgaagaaatcacctcagataactaacatcgaagcagccgaagaaaacatctc
gctggacctgatccagcagtagctactacctgacctttaaatttcgacacagagccggaacatttctatcgaaa
acctgagctctgatcatcggccagctggaactgatgccgaacatcgaacggttcccaaacggtgaaaaag
tacgagctggacaaatataccatggtccactacctgcgcgcgaggaattgaaacagggcaaatcccgat
cgcactgactaactccgttaacgaagctctgctcaaccgctccggtgtatacacttctctctagcagact
acgtgaaaaaggtcaacaaagcagactgaagctgcaactgttcttgggtgggtgaaacagcttgtttatgat
tttacgcagcagcagctccgaagtagtactactaccgacaaaattgaggatcactatcatcatcccgtacat
cggctccggctctgaacattggcaacatgctgtacaaagacgacttcttggcgactgatctctccgggtg
cgggtgatcctgctggagttcatcccggaaatcgccatcccggtagtgggcaacttctgctggtttcttac
attgcaaaaaggtctgactgtacaaacctcgaacacgctgagcaaacgtaacgaaaaatgggatga
agtttacaataatctcgtgaccaactgctggctaaaggttaatactcagatcgacctcatccgcaaaaaa
tgaaagaagcactggaaaaccaggcgggaagctaccaaggcaatcattaactaccagtacaaccagtagacc
gaggagaaaaaaacacatcaacttcaacatcgacgatctgtcctctaaactgaacgaatccatcaacaa
agctatgatcaacatcaacaagttcctgaaccagtgctctgtaagctatctgatgaactccatgatcccg
acgggtgtaaagctctggaggacttcgtagcgtctctgaaagacgcctctgaaatacatttacgacaac
cgtggcactctgatcggtcagggtgatcgtctgaaagcacaagtgaaacaatacctatcgaccgacatccc
tttccagctcagtaaatatgctgataaccaacgccttttgtccactTGTggcggtggcggtagcATCGAAG
GTCGTTACGGTGGTTTCATGACCTCTGAAAAATCTCAGACCCCGCTGGTTACCCTGTTCAAAAACGCTATC
ATCAAAAACGCTTACAAAAAGGTGAAGcggtgggggtagtTGCggcggtggcggttcgcgtccgccggg
tttctctccgttccggtgataa

SEQ ID 10

MHHHHHHHHHGGSSNNNNNNNNNNGSMFVFNKQFNKYKDPVNGVDIAYIKIPNAGQMOPVKAFKIHNNKIW
VIPERDFTNPEEGDLNPPPEAKQVPVSYDYDSTYLSNDNEKDNYLKGVTKLFEIYSTDLGRMLLTSIV
RGIPIFWGGSTIDTELKVIDTNCINVIQPDGYSRSEELNLVIGPSADIIQFECKSFGEVNLNLRNGYG
STQYIRFSPDFTFGFEESLEVDNPLLGAGKFATDPAVTLAHELHAGHRLYGIAINPNRVFKVNTNAY
YEMSGLEVSFEELRTFGGHDAKFIDSLQENEFRLYYNKFKDIASTLNKAKSIVGTASLQYMKNVFEKE
KYLLEDTSKGFSVDLKLKFDKLYKMLTEIYTEDNFVKFFKVLNRKTYLNFDKAVFKINIVPKVNYTIYD
GFNLRNINLAANFNGQNTIENNMNFTKLNFTGLFEFYKLLCVDGIITSKTKSLIEGRNKALNDLCIKV
NNWDLFFSPSEDNFTNDLNKGEETSDTNI EAAEENISLDLIQQYYLTFNFDNEPENISINLSSDIIG
QLELMPNIERFPNGKKEYELDKYTMFHYLRAQEFEHGKSRIALTNVNEALLNPSRVYTFSSDYVKKVN
KATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIIIPYIGPALNIGNMLYKDFVGLIFSGAVIL
LEFIPEIAIPVLGTFALVSYIANKVLTVQTI DNALSKRNEKWDEVYKYIVTNWLAKVNTQIDLIRKMKM
EALENQAEATKAIINYQYNQYTEEEKNNINFNIDDLSSKLNESINKAMININKFLNQC SVSYLMNSMIP
YGVKRLLEDFDASLKDALLKYIYDNRGTLLIGQVDRLLKDKVNTLSTDIPFQLSKYVDNQRLSTCGGGGS
IEGRYGGFMTSEKSTPLVTLFKNAIKNAYKKGEGGGGSCGGGSRPPGFS PFR

SEQ ID 11

atgcatcaccatcaccatcaccatcaccatcatgggagctCGAACAATAACAACAATAACAATAACAATAA
Cggatccatgacgtggccagtttaaggatttcaactactcagatcctgtaaatgacaacgatattctgtacc
ttcgcatcccaaaaataaactgatcaccacaccagtcaaagcattcatgattactcaaaacatttgggtc
attccagaacgcttttctagtgacacaaaatccgagtttatctaaacctccgctccgacgtccaaatatca
gagctattaccatccctcatctcagtagcggacgaaacaaaagatacttctctaaaggtatcattaaac
tgtttaagcgtatataatgagcgcgatcgggaaaaagttgattatcttctgttgggttcccgctc
atggggcagatagctctacccccgaagacacttttgattttaccgctacacgacaaaacatcgcggttagagaa
gtttgagaacggatcgtgaaagtcacaaacatctacacctagcgtcttaatttttggctccgctgcaa
acatcttagattatacagccagcctgactttgcaggggcaacagtcgaatccgagtttcaaggttttgg
acctgagcattctgaaagttgccccggaatttctgctcacttttccagatgtcaccgcaaccgagctc
agcagttattaggaagtcatttttctgcatggaccgggttattgactgatgcacgaactgacgcactctc
tgcataactctgatgggatcaacatcccagtgacaaacgtattctgctcccagggtgctgaaggattttc

5

tcacaggatgggcccgaacgtccaggttcgaagagttgtatactttcggaggcctggacgtagagatcattcc
ccagattgagcgcagtcagctgagcgtgagaagcattgggcccattataaggatattgcaaacgcctgaata
acattaacaaaacgattccatcttcgctggatctcgaatattgataaatataagaaaatttttagcgagaaa
tataattttgataaagataatacaggtaactttgtggtaacattgacaaaattcaactccctttacagtgga
tttgacgaatgtaatgagcgaagttgtgtatagttcccaatacaacgtaagaatcgtaccattactctc
ctcgtcactacctgcccgttttcgcgaacatccttgacgataatattacactattcgtgacggctttaac
ttgaccaacaagggcttcaatattgaaaattcaggccagaacattgaacgcaacccggccttgacagaaact
gtcagtgaaatccgtgggtgacctgtttacaaaagtctcgtcgaacggcatcattacctccaaaaactaat
ctctgatagaaggtagaaacaaagcgtgaacctgcagtgattaaagtgaaaacaatcggtgctccttat
gtagcagataaagatagcattagtcaggagattttcgaataaaaattatcactgacgaaacaaatgttca
gaattattcagataaaatttcactggacgaaagcatcttagatggccaagtccgattacccggaattg
ttgatccgttactgcccgaacgtgaatattgaaaccgttaaaccctccctggcgaagagatcgtattttatgat
gacattacgaaatattgtgactaccttaattcttattactatttgaaagccgaaactgtccaataacgt
ggaaaacattactctgaccacaagcgtggaagaggcttaggctactcaaaataagattatacctcctcc
cgtcgtggcggaaaagtaaaaggtgtgcaggctggctgttctcaactgggccaatgaagttgtc
gaagactttaccacgaatattatgaaaaaggataccctggataaaaatctccgacgtctcggttattatccc
atatttgccctgcttaaatatcggttaatagtcgctgcccggggaattttaaccaggcctttgctaccg
cgggctgcccgttccctcctggagggttctcctgaatttactatcccggcgtcgggtgtttttacattttac
tcttccatccaggagcgtgagaaaattatcaaaaccatcgaaaactgcctggagcagcgggtgaaacgctg
gaaagattcttataatggatgggtgtcaaaactggttatctcgcacacgacccaattcaaccatattaatt
accagatgtagatagtcgtaccgaagcgtgacgccaattaaagccaaaattgatctggaatataaaaag
tactctggtagcgataaaggagaacatcaaaaagccaggtggagaaccttaagaaatagtcggatgtgaaat
ctctgaagctatgaaataacattaacaaattcattcgtgaatgttcgggtgacgtacctgttcaagaatagc
tgccaaaagttattgatgaaactgaaataattgatctgctgacccaaaaccgacttatcaacctcatcgac
tccccacaacattatccttggggcgaagtggtatcgtcgaagccaaaagtaaacgagagctttgaaaatc
gatgcccgtttaatatttttcatataccaataactccttgctgaaagatatcatcaatgaatatttcaaTC
TAGAATGTggcgggtggcggtagcATCGAAGGTCGTACCGTGGATGCGATCTTCACTCAGTCTTACCGTAAA
GTTCTGGCGCAGCTGAGCGCTCGTAAACTGCTGCAGGATATCCTGAACCGTCAGCAGGGTGAACGTAACCA
GGAACAGGGCGTggcgggtgggggtagtTGCggcgggtggcggttcgCACGTGGATGCGATCTTCACTCAGT
CTTACCGTAAAGTTCTGGCCAGCTGAGCGCTCGTAAACTGCTGCAGGATATCCTGAACCGTCAGCAGGGT
GAACGTAACCGAAGGCGCTtgataa

SEQ ID 12

MHHHHHHHHHGSNNNNNNNNNNGSMTWPKDFNYSDFVNDNDI LYLRI PQNKLI TTPVKAFMITQNIWV
IPERFSSDTNPSLSKPPRPTSKYQSYDPSYLSSTDEQKDTFLKGI I KLFKRINERDI GKKLINYL VVGSPF
MGDSSTPEDTFDFTRHTTN IAVEKFENGSWKVTNI I T P SVLI FGPLPNI LDY TASLT LQGQSNPSFEGFG
T LSI LKVAPEFL LTFSDVT SNQSSAVLGKSI FCM DPVIALMHELTHSLHQLYGINI PSDKRI RPQVSEGF
SQDGNVQFEELYTFGGDLVEI I PQI ERSQ LREKALGHYKDI AKRLNNINKTI PSSWI SNI DKYKKI FSEK
YNFDKDNTGNFVNI DKFN SLYSDL TNVMSEVVYSQYNVKNRTHYFSRHYLPVFANI LDDNI YTI RDGFN
LTNKG FNI ENSQNI ERN PALQKLSSESVDLFTKVCVDGI I TSKTKSLIEGRNKALNLQCI KVKNNRLPY
VADKDSI SQEI FENKI ITDET NVQNYSDKFSLDESI LDGQVP INPEI VDP LLLPNVNM EPLNL PGEI VFYD
DITKYVDY LNSYYLESQKLSNNVENIT L TTSVEEALGYSNKI YTF L PSLAEKVNKG VQAGL FLN WAN EVV
EDFTTNIMKKD TLDKI SDVSVI I PYIG PALNIGNSALRGNFNQAFATAGVAF LLEGFPEFT I PALGVTFY
SSI QEREKI I KTI ENCLEQRV KRKDSYQWMSN WLSRI T TQFNHINYQMY DLSYQADAI KAKI DLEYKK
YSGSDKENIKSQVENLKN SLDVKI SEAMN NINKFIRECSVTYLFKNM LPKVI DELNKFDLRTKTELINLID
SHNI I LVGEVDRLKAKVNESFENTMPFNI FSYTNN SLLKDI INEYFNLECGGGGSI EGRHVD AI FTQSYRK
VLAQLSARKLLQDI LNRQQGERNQEQGAGGGGSCGGGSHVD AI FTQSYRKVLAQLSARKLLQDI LNRQQG
ERNQEQA

SEQ ID 13

atgcatcaccatcaccatcaccatcaccatcatgggagctCGAACAATACAACAATAACAATAACAATAA
Cggatccatgacgtggccagtttaaggatttcaactactcagatcctgtaaatgacaacgatattctgtacc
ttcgcattccacaaaataaactgatcaccacaccagtc aaagcattcatgattactcaaaacatttgggtc
attccagaacgcttttctagtacacaaaatccgagtttatctaaacctccgctcgcgacgtccaaatca
gagctattacgatccctcatatctcagtcaggacgacaaaaagatactttccttaaaggatcatataaac
tgtttaagcgtattaatgagcggatcgcgggaaaaagttgattaattatcttgttgggttccccgttc
atggcgatagctctacccccgaagacacttttgattttaccgctcatacgacaaacatcgcggtagagaa
gtttgagaacggatcgtggaagtcacaaacatcattacacctagcgtccttaatttttggctcgtgccaa
acatcttagattatacagccagcctgactttgcaggggcaacagtcgaatccgagtttcgaaggttttgg
accctgagcattctgaaagttgccccggaatttctgctcacttttccagatgtcaccagcaaccagagctc
agcagattaggaagtc aatttttgcattggaccgggttatgcaactgatgacgaactgacgcactcc

5

tgcatcaactgtatgggatcaacatccccagtgacaaacgtattcgtccccaggtgtctgaaggatttttc
tcacaggatgggcccgaacgtccagttcgaagagttgtatactttcggaggcctggacgtagagatcattcc
ccagattgagcgcagtcagctgctgagaaggcattgggcccattataaggataattgcaaaacgcctgaata
acattaacaaaacgattccatcttcgtggatctcgaatattgataaaatataagaaaatttttagcggagaaa
tataattttgataaaagataatcacaggaactttgtggtaacattgacaaattcaactccctttacagtgga
tttgacgaatgtaatgagcgaagttgtgtatagttcccaatacaacgtaagaatcgtaccattactctct
ctcgtcactacctgcccgttttcggaacatccttgacgataatatttactattcgtgacggctttaac
ttgaccaacaagggcttcaatattgaaaattcaggccagaacattgaaacgcaaccggccttgacagaact
gtcagtgatccgtggttgacctgtttaccaaagtctgctgacggcatcattacctcaaaaactaaat
ctctgataagaagtagaaaacaaagcgtgaaacctgacgtgtattaaagtgaaaacaatcggctgccttat
gtagcagataaaagatagcatttagtcaggagattttcgaaaataaaattatcactgacgaaaccaatgttca
gaattattcagataaaattttcactggacgaaagcatcttagatggccaagtccgattaaccggaaattg
ttgatccgcttactgcccgaacgtgaatatggaaccttaaacctcctggcgaagagatcgtattttatgat
gacattacgaaatattgtgactaccttaattcttattactatttggaagccagaaactgtccaataacgt
ggaaaacattactctgaccacaagcgtggaagaggctttaggctactcaataagatttataccttccctcc
cgtcgtggcggaaaaagtaaaataaggtgtgacggctggtctgttccactggcggaatgaagttgtc
gaagactttaccacgaatattatgaaaaaggataccctggataaaatctccgacgtctcggttattatccc
atataattggccctgctgtaaaatcggtaatagtgctgctgccccggaattttaaccaggcctttgctaccg
cggcgctgcgcttccctcctggagggcttccctgaatttactatcccgcgctcgggtgtttttacattttac
tcttccatccaggagcgtgagaaaattatcaaaacctogaaaactgcctggagcagcgggtgaaacgctg
gaaagattcttatcaatggatggtgtcaaacctggttatctcgcacacgaccaattcaaccataattaatt
accagatgtagatagctgtcgtaccgaactgacgccattaaagccaaaattgatctggaatataaaaag
tactctggttagcgataaggagaacatcaaaagccaggtggagaaccttaagaaatagctggatgtgaaaat
ctctgaaactatgaataacatcaaaaattcattcgtgaatgttcgggtgacgtacctgtcaagaatagc
tgccaaaagttattgatgaactgaataaatttgatctgcgtacaaaaccgaaacttatcaacctcatcgac
tcccacaacattatccttgtggcggaagtgatcgtctgaagccaaaagtaaacgagagctttgaaaatac
gatgcccgtttaatatttttcatataccaataactccttgctgaaagatatcatcaatgaatatttcaaTC
TAGAaTGTGGTGGCGGTGGCTCTGGTGGCGGGTCTATTGAAGGTGCGGGTGGCGGTGGCTCTGGTGGC
GGCGTCTCagcaactggtcctattgctcgcacctggttgataa

SEQ ID 14

MHHHHHHHHHGSNNNNNNNNNGSMTWPVKDFNYS DPVNDNDI LYLRI PQNKLI TTPVKAFMITQNIWV
I PERFFSSDTNPSLSKPPRPTSKYQSYDPSYLSSTDEQKDTFLKGI IKLFKRINERDIGKKLINYL VVGS PF
MGDSSSTPEDTDFDRHTTNI AVEKFEENGSWKVNTN ITPSVLI FGLPNI LDYASLT LQGQSNPS FEGFG
TLSILKVAPEFLTFSDVTSNQSSAVLGKSI FCM DPVIALMHELTS LHQLYGINI PSDKRIRPQVSEGF
SQDGPVNVQFEELYTFGLDVEI I PQIERSQLREKALGHYKDI AKRLNNINKTI PSSWISNIDKYKFI FSEK
YNFDKDN TGNFVNIDKFNLSYSDLTNVMSEVVYSQYNVKNRTHYFSRHYLPVFANI LDDNI YTI RDGFN
LTNKGFN IENSGQNI ERN PALQKLS SESVVDLFTKVCVDGI ITSKT KSLIEGRNKALNLQCI KVKNNR LPY
VADKDSI SQEI FENKI ITDET NVQNSDKFSLDES ILDGQVPINPEIVDPLLPNVNMEPLNLPGEEIVFYD
DITKYVDY LNSYLYLESQKLSNNVENITLTT SVEEALGYSNKI YTF LPLSLAEKVNKG VQAGLFLN WANEV
EDFTT NIMKKTLDKI SDVSVI I PYIGPALNIGNSALRGNFNQAFATAGVAFLLEGFP EFTI PALGVTFY
SSIQEREKI IKTI ENCLEQRV KRKDS YQWVMSN WLSRI TTFQFNHINYQMYDSL SYQADAI KAKI DLEYKK
YSGSDKENIKSQVENLKN SLDVKISEAMNNINKFI RECSVTYLFKNMLPKVIDELNKFDLRTKTELINLID
SHNI I LVGEVDRLKAKVNES FENTMPFNI FSYTNNS LKDI INEYFNLECGGGSGGGGSIEGRGGGSGG
GGSQHWSYCLRPG

SEQ ID 15

atgcatcaccatcaccatcaccatcaccatcatgggagctCGAACAAtACAACAATAACAATAACAAtAA
Cggatccatggagttcgttaacaaacagttcaactataaagaccagttaacgggtgttgacattgcttaca
tcaaaaatcccgaacgctggccagatgcagccggttaaaggcattcaaaaatccacaacaaaatctgggttatac
ccggaacgtgatacctttactaacccggaagaaggtgacctgaacccgccaccggaagcgaacaggtgcc
ggtatcttactatgactccacctacctgctcaccgataacgaaaaggacaactacctgaaaggtgttacta
aactgttcgagcgtatttactccaccgacctgggcccgtatgctgctgactagcatcgttcggcggtatcccg
ttctggggcggttctaccatcgataccgaactgaaagtaatcgacactaactgcatcaacggttatcagcc
ggacgggttccatcgttccgaagaactgaaacctggtgatcaccgcccgtctgctgatacatccagttcg
agtgtaaagactttggtcacgaagttctgaaacctcaaccgtaacggctacgggtccactcagtaacatccgt
ttctctccggacttcaacctcgggtttgagaatccctgggaagtagacacgaaaccactgctggcgctgg
taaatcgcgaactgatcctgcccgttacctggctcaggaactgatcattgcaggccaccgctgtacggta
tcgccatcaatccgaacggtgtcttcaaagttaacaccaacgctattacgagatgtccggctcgggaagtt
agcttcgaagaactgcgtactttggcggctcagcagcgtataatcattcgaactctctgcaagaaaacgagtt
ccgtctgtaactactataacaagttcaaaagatctcgcattccaccctgaacaaagcgaatccatcgtgggta

5

ccactgcttctctccagtagatgaagaaacggttttaagaaaaaacctgctcagcgaagacacctccggc  
 aaattctctgtagacaagttgaaattcgataaaactttacaaaatgctgactgaaatttacaccgaagacaa  
 cttcgttaagttctttaaagttctgaaaccgcaaaaactatctgaaacttcgacaaggcagatttcaaaatca  
 acatcgtgcccgaaggttaactacactatctacgatgggttcaacctgcgtaaacaccaacctggctgcta  
 tttaacggccagaacacggaatcaacaacatgaacttcacaaaactgaaaaacttcaactggtctgttcga  
 gttttacaagctgctgctgacggcatcattacctccaaaactaaatctctgatagaaggtagaaaca  
 aagcgtgaacgacctctgtatcaaggttaacaactgggatttattcttcagcccagtgagacaacttc  
 accaacgacctgaaacaagggtgaagaaatcacctcagataactaacatcgaagcagccgaagaaaaatctc  
 gctggacctgatccagcagtaactacctgacctttaaatttcgacaacgagccggaaaaacatttctatcgaaa  
 acctgagctctgatcatcggccagctggaactgatgcccgaacatcgaacgtttcccaaacggtaaaaag  
 tacgagctggacaaaataaccatgttccactacctcgcgcgcaggaatttgaacacggcaaatcccgat  
 cgcactgactaactccgttaacgaagctctgctcaaccgctccggtgatacaaccttctctctagcgact  
 acgtgaaaaaggtaacaaaaggcactgaagctgcaatgttcttgggttgggtgaacagcttggttatgat  
 tttaccgacgagcgtccgaagtatctactaccgacaaaattcgggatacactatcatcatcccgta  
 cgtccggctctgaaacattggcaacatgctgtacaaaagcagcttcttggcgcactgatcttctccggtg  
 cggatcctgctggagttcatcccgaaaatcgccatcccggtactgggcaaccttctctgctggtttcttac  
 attgcaaaaacagggtctgactgtacaaaaccatcgacaacgcgctgagcaaacgtaaacgaaaaatgggatga  
 agtttacaataatctgtagcaactggctggctaaggttaataactcagatcgacctcatccgcaaaaaaa  
 tgaagaagcactggaaaaccaggcgaagctaccaaggcaatcattaactaccagtaacaaccagtaacc  
 gaggaagaaaaaaacaactcaacatcgacgatctgctctctaaactgaacgaatccatcaaaa  
 agctatgatcaacatcaacaagttcctgaaccagtgctctgtaagctatctgtagaactccatgatccgct  
 acggtgttaaacgctctggaggacttcgatgctctgaaagcgcctgctgaaatacatttacgacaac  
 cgtggcactctgatcggctcaggtgatcgtctgaaggacaaagtgaacaataccttatcgaccgacatccc  
 tttcagctcagtaaatatgtcgataaccaacgccttttctccactTGTggcgggtggcggtagcATCGAAG  
 GTCGTGTTCCATTACCAGCAGGAGGAGGAACAGTAT'GACTAAAATGTATCCATgcGGAAATCACTGGGCA  
 GTGGGACATCTAATGGGAtgataa

**SEQ ID 16**

MHHHHHHHHHGSNNNNNNNNNNGSMFVNKQFNKYKDPVNGVDIAYIKIPNAGQMOPVKAFKIHNKIWI  
 PERDTFTNPEEGDLNPPPEARQVPVSYDSTYLS'DNEKDNYLKGVTKLFERISTDLGRMLLTSIVRGI P  
 FWGGSTIDTELKVIDTNCINVIQPDGSYRSEELNLVIGPSADIIQFECKSFGHEVLNLRNGYGSTQYIR  
 FSPDFTFGFEESLEVDTNPLLGAGKFATDPAVTLAHELIIHAGHRLYGIAINPNRVFKVNTNAYYEMSGLEV  
 SFEELRTFGGHDAKFIDSLQENEFRLYYNKFKDIASTLNKAKSIVGTTASLQYMKNVFKEKYLLEDTS  
 KFSVDKLFKLYKMLTEIYEDNFVKFFKVLNRKTYLNFDKAVFKINIVPKVNYTIYDGFNLRNTNLAAN  
 FNGQNTIENNNMFTKLNFTGLFEFYKLLCVDGIITSKTKSLIEGRNKALNDLCIKVNNWDLFFSPSEDNF  
 TNDLNKGEIETSDTNI EAAEENISLDLIQQYYLTFNFDNEPENIS IENLSSDI IGQLELMPNIEF'PNGK  
 YELDKYTMFHYLRAQEFEHGKSRIALTNVNEALLNPSRVYTFSSDYVKKVNKATEAAMFLGWVEQLVYD  
 FTDETSEVSTTDKIADITIIIPYIGPALNIGNMLYKDDFVGALIFSGAVILLEFIPEIAIPVLGTFALVSY  
 IANKVLTVQTIDNALS KRNEKWDEVYKYIVTNWLAKVNTQIDLIRKKMKEALENQAEATKAIINYQYNQYT  
 EEEKNNINENIDDLSSKLNESINKAMININKFLNQCSVSYLMNSMIPYGVKRLEDFFDASLKDALLKYIDN  
 RGTLLIGQVDRDKKVNNTLSTDIPFQLSKYVDNQRLSTCGGGGSI EGRVLPAGGGTVLTKMYP CGNHWA  
 VGHLMG

**Lista de secuencias**

5	<110> Syntaxin Limited Chaddock, John Harper, Elaine	
	<120> Proteínas de fusión terapéuticas	
	<130> P35972W0-MRM	
10	<140> Aún no asignado	
	<141> 16-05-2012	
	<150> GB1108108.0	
15	<151> 16-05-2011	
	<160> 16	
	<170> PatentIn versión 3.5	
20	<210> 1	
	<211> 2769	
	<212> ADN	
	<213> Secuencia artificial	
25	<220>	
	<223> Secuencia sintética	
	<400> 1	
	atgcatcacc atcaccatca ccatcaccat catggggagct cgaacaataa caacaataac	60
	aataacaata acggatccat gacgtggcca gttaaggatt tcaactactc agatcctgta	120
	aatgacaacg atattctgta ccttcgcatt ccacaaaata aactgatcac cacaccagtc	180
	aaagcattca tgattactca aaacatttgg gtcattccag aacgcttttc tagtgacaca	240
	aatccgagtt tatctaaacc tccgcgtccg acgtccaaat atcagagcta ttacgatccc	300
	tcatatctca gtacggacga acaaaaagat actttcctta aaggatcat taaactgttt	360
	aagcgtatta atgagcgcga tatcgggaaa aagttgatta attatcttgt tgtgggttcc	420
	ccgttcatgg gcgatagctc taccgccgaa gacacttttg attttaccg tcatacgaca	480
	aacatcgcgg tagagaagtt tgagaacgga tcgtggaaag tcacaaacat cattacacct	540
	agcgtcttaa tttttggtcc gctgccaaac atcttagatt atacagccag cctgactttg	600
	caggggcaac agtcgaatcc gagtttcgaa ggttttggtta ccctgagcat tctgaaagtt	660
	gccccggaat ttctgctcac tttttcagat gtcaccagca accagagctc agcagtatta	720
	ggaaagtcaa ttttttgcac ggaccgggtt attgcaactga tgcacgaact gacgcactct	780
	ctgcatcaac tgotatggat caacatcccc agtgacaaac gtattcgtcc ccaggtgtct	840
	gaaggatttt tctcacagga tgggccgaac gtccagttcg aagagttgta tactttcgga	900
	ggcctggacg tagagatcat tccccagatt gagcgcagtc agctgcgtga gaaggcattg	960
	ggccattata aggatattgc aaaacgcctg aataacatta acaaacgat tccatcttcg	1020
	tggatctcga atattgataa atataagaaa atttttagcg agaaatataa ttttgataaa	1080
	gataatacag gtaactttgt ggtaacatt gacaaattca actcccttta cagtgatttg	1140
30	acgaatgtaa tgagcgaagt tgtgtatagt tccaataca acgttaagaa tcgtacctat	1200

ES 2 562 425 T3

tacttctctc gtcactacct gccggttttc gcgaacatcc ttgacgataa tatttacact 1260  
 attcgtgacg gctttaactt gaccaacaag ggcttcaata ttgaaaattc aggccagaac 1320  
 attgaacgca acccggcctt gcagaaactg tcgagtgaat ccgtggttga cctgtttacc 1380  
 aaagtctgcg tcgacggcat cattacctcc aaaactaaat ctctgataga aggtagaaac 1440  
 aaagcgtga acctgcagtg tattaagtg aaaaacaatc ggctgcctta tntagcagat 1500  
 aaagatagca ttagtcagga gattttcgaa aataaaatta tcaactgacga aaccaatggt 1560  
 cagaattatt cagataaatt ttcactggac gaaagcatct tagatggcca agttccgatt 1620  
 aacccggaaa ttgttgatcc gttactgccg aacgtgaata tggaaaccgtt aaacctccct 1680  
 ggcaagaga tcgtatttta tgatgacatt acgaaatag tggactacct taattcttat 1740  
 tactatttgg aaagccagaa actgtccaat aacgtggaaa acattactct gaccacaagc 1800  
 gtggaagagg ctttaggcta ctcaaataag atttatacct tcctcccgtc gctggcggaa 1860  
 aaagtaaata aagggtgca ggctggctctg ttcctcaact gggcgaatga agttgtcgaa 1920  
 gactttacca cgaatattat gaaaaaggat accctggata aaatctccga cgtctcgggt 1980  
 attatcccat atattggccc tgcgttaaat atcggtaata gtgcgctgcg ggggaathtt 2040  
 aaccaggcct ttgctaccgc gggcgctcgc ttcctcctgg agggctttcc tgaatttact 2100  
 atcccggcgc tcggtgtttt tacattttac tcttccatcc aggagcgtga gaaaattatc 2160  
 aaaaccatcg aaaactgcct ggagcagcgg gtgaaacgct ggaaagattc ttatcaatgg 2220  
 atgggtgcaa actggttatc tcgcatcacg acccaattca accatattaa ttaccagatg 2280  
 tatgatagtc tgtcgtacca agctgacgcc attaaagcca aaattgatct ggaatataaa 2340  
 aagtactctg gtacgataa ggagaacatc aaaagccagg tggagaacct taagaatagt 2400  
 ctggatgga aaatctctga agctatgaat aacattaaca aattcattcg tgaatgttcg 2460  
 gtgacgtacc tgttcaagaa tatgctgcca aaagttattg atgaactgaa taaatttgat 2520  
 ctgctacca aaaccgaact tatcaacctc atcgactccc acaacattat ccttgtgggc 2580  
 gaagtggatc gtctgaaggc caaagtaaac gagagctttg aaaatacgat gccgtttaat 2640  
 attttttcat ataccaataa ctcttctgctg aaagatatca tcaatgaata tttcaatcta 2700  
 gaatgtggcg gtggcggtag catcgaaggt cgtcagcact ggtcctattg cctgcccct 2760  
 ggttgataa 2769

<210> 2  
 <211> 921  
 <212> PRT  
 <213> Secuencia artificial

5

<220>  
 <223> Secuencia sintética

10

<400> 2

Met His Gly Ser Ser Asn Asn  
 1 5 10 15

ES 2 562 425 T3

Asn Asn Asn Asn Asn Asn Asn Asn Gly Ser Met Thr Trp Pro Val Lys  
 20 25 30  
 Asp Phe Asn Tyr Ser Asp Pro Val Asn Asp Asn Asp Ile Leu Tyr Leu  
 35 40 45  
 Arg Ile Pro Gln Asn Lys Leu Ile Thr Thr Pro Val Lys Ala Phe Met  
 50 55 60  
 Ile Thr Gln Asn Ile Trp Val Ile Pro Glu Arg Phe Ser Ser Asp Thr  
 65 70 75 80  
 Asn Pro Ser Leu Ser Lys Pro Pro Arg Pro Thr Ser Lys Tyr Gln Ser  
 85 90 95  
 Tyr Tyr Asp Pro Ser Tyr Leu Ser Thr Asp Glu Gln Lys Asp Thr Phe  
 100 105 110  
 Leu Lys Gly Ile Ile Lys Leu Phe Lys Arg Ile Asn Glu Arg Asp Ile  
 115 120 125  
 Gly Lys Lys Leu Ile Asn Tyr Leu Val Val Gly Ser Pro Phe Met Gly  
 130 135 140  
 Asp Ser Ser Thr Pro Glu Asp Thr Phe Asp Phe Thr Arg His Thr Thr  
 145 150 155 160  
 Asn Ile Ala Val Glu Lys Phe Glu Asn Gly Ser Trp Lys Val Thr Asn  
 165 170 175  
 Ile Ile Thr Pro Ser Val Leu Ile Phe Gly Pro Leu Pro Asn Ile Leu  
 180 185 190  
 Asp Tyr Thr Ala Ser Leu Thr Leu Gln Gly Gln Gln Ser Asn Pro Ser  
 195 200 205  
 Phe Glu Gly Phe Gly Thr Leu Ser Ile Leu Lys Val Ala Pro Glu Phe  
 210 215 220  
 Leu Leu Thr Phe Ser Asp Val Thr Ser Asn Gln Ser Ser Ala Val Leu  
 225 230 235 240  
 Gly Lys Ser Ile Phe Cys Met Asp Pro Val Ile Ala Leu Met His Glu  
 245 250 255  
 Leu Thr His Ser Leu His Gln Leu Tyr Gly Ile Asn Ile Pro Ser Asp  
 260 265 270  
 Lys Arg Ile Arg Pro Gln Val Ser Glu Gly Phe Phe Ser Gln Asp Gly  
 275 280 285  
 Pro Asn Val Gln Phe Glu Glu Leu Tyr Thr Phe Gly Gly Leu Asp Val

ES 2 562 425 T3

290                      295                      300  
 Glu Ile Ile Pro Gln Ile Glu Arg Ser Gln Leu Arg Glu Lys Ala Leu  
 305                      310                      315                      320  
 Gly His Tyr Lys Asp Ile Ala Lys Arg Leu Asn Asn Ile Asn Lys Thr  
                                  325                      330                      335  
 Ile Pro Ser Ser Trp Ile Ser Asn Ile Asp Lys Tyr Lys Lys Ile Phe  
                                  340                      345                      350  
 Ser Glu Lys Tyr Asn Phe Asp Lys Asp Asn Thr Gly Asn Phe Val Val  
                                  355                      360                      365  
 Asn Ile Asp Lys Phe Asn Ser Leu Tyr Ser Asp Leu Thr Asn Val Met  
                                  370                      375                      380  
 Ser Glu Val Val Tyr Ser Ser Gln Tyr Asn Val Lys Asn Arg Thr His  
                                  385                      390                      395                      400  
 Tyr Phe Ser Arg His Tyr Leu Pro Val Phe Ala Asn Ile Leu Asp Asp  
                                  405                      410                      415  
 Asn Ile Tyr Thr Ile Arg Asp Gly Phe Asn Leu Thr Asn Lys Gly Phe  
                                  420                      425                      430  
 Asn Ile Glu Asn Ser Gly Gln Asn Ile Glu Arg Asn Pro Ala Leu Gln  
                                  435                      440                      445  
 Lys Leu Ser Ser Glu Ser Val Val Asp Leu Phe Thr Lys Val Cys Val  
                                  450                      455                      460  
 Asp Gly Ile Ile Thr Ser Lys Thr Lys Ser Leu Ile Glu Gly Arg Asn  
                                  465                      470                      475                      480  
 Lys Ala Leu Asn Leu Gln Cys Ile Lys Val Lys Asn Asn Arg Leu Pro  
                                  485                      490                      495  
 Tyr Val Ala Asp Lys Asp Ser Ile Ser Gln Glu Ile Phe Glu Asn Lys  
                                  500                      505                      510  
 Ile Ile Thr Asp Glu Thr Asn Val Gln Asn Tyr Ser Asp Lys Phe Ser  
                                  515                      520                      525  
 Leu Asp Glu Ser Ile Leu Asp Gly Gln Val Pro Ile Asn Pro Glu Ile  
                                  530                      535                      540  
 Val Asp Pro Leu Leu Pro Asn Val Asn Met Glu Pro Leu Asn Leu Pro  
                                  545                      550                      555                      560  
 Gly Glu Glu Ile Val Phe Tyr Asp Asp Ile Thr Lys Tyr Val Asp Tyr  
                                  565                      570                      575

ES 2 562 425 T3

Leu Asn Ser Tyr Tyr Tyr Leu Glu Ser Gln Lys Leu Ser Asn Asn Val  
 580 585 590  
 Glu Asn Ile Thr Leu Thr Thr Ser Val Glu Glu Ala Leu Gly Tyr Ser  
 595 600 605  
 Asn Lys Ile Tyr Thr Phe Leu Pro Ser Leu Ala Glu Lys Val Asn Lys  
 610 615 620  
 Gly Val Gln Ala Gly Leu Phe Leu Asn Trp Ala Asn Glu Val Val Glu  
 625 630 635 640  
 Asp Phe Thr Thr Asn Ile Met Lys Lys Asp Thr Leu Asp Lys Ile Ser  
 645 650 655  
 Asp Val Ser Val Ile Ile Pro Tyr Ile Gly Pro Ala Leu Asn Ile Gly  
 660 665 670  
 Asn Ser Ala Leu Arg Gly Asn Phe Asn Gln Ala Phe Ala Thr Ala Gly  
 675 680 685  
 Val Ala Phe Leu Leu Glu Gly Phe Pro Glu Phe Thr Ile Pro Ala Leu  
 690 695 700  
 Gly Val Phe Thr Phe Tyr Ser Ser Ile Gln Glu Arg Glu Lys Ile Ile  
 705 710 715 720  
 Lys Thr Ile Glu Asn Cys Leu Glu Gln Arg Val Lys Arg Trp Lys Asp  
 725 730 735  
 Ser Tyr Gln Trp Met Val Ser Asn Trp Leu Ser Arg Ile Thr Thr Gln  
 740 745 750  
 Phe Asn His Ile Asn Tyr Gln Met Tyr Asp Ser Leu Ser Tyr Gln Ala  
 755 760 765  
 Asp Ala Ile Lys Ala Lys Ile Asp Leu Glu Tyr Lys Lys Tyr Ser Gly  
 770 775 780  
 Ser Asp Lys Glu Asn Ile Lys Ser Gln Val Glu Asn Leu Lys Asn Ser  
 785 790 795 800  
 Leu Asp Val Lys Ile Ser Glu Ala Met Asn Asn Ile Asn Lys Phe Ile  
 805 810 815  
 Arg Glu Cys Ser Val Thr Tyr Leu Phe Lys Asn Met Leu Pro Lys Val  
 820 825 830  
 Ile Asp Glu Leu Asn Lys Phe Asp Leu Arg Thr Lys Thr Glu Leu Ile  
 835 840 845

ES 2 562 425 T3

Asn Leu Ile Asp Ser His Asn Ile Ile Leu Val Gly Glu Val Asp Arg  
850 855 860

Leu Lys Ala Lys Val Asn Glu Ser Phe Glu Asn Thr Met Pro Phe Asn  
865 870 875 880

Ile Phe Ser Tyr Thr Asn Asn Ser Leu Leu Lys Asp Ile Ile Asn Glu  
885 890 895

Tyr Phe Asn Leu Glu Cys Gly Gly Gly Ser Ile Glu Gly Arg Gln  
900 905 910

His Trp Ser Tyr Cys Leu Arg Pro Gly  
915 920

<210> 3

<211> 2739

<212> ADN

<213> Secuencia artificial

<220>

<223> Secuencia sintética

<400> 3

atgcatcacc atcaccatca ccatcaccat catgggagct cgaacaataa caacaataac	60
aataacaata acggatccat ggagttcgtt aacaaacagt tcaactataa agaccagtt	120
aacggtgttg acattgctta catcaaaatc ccgaacgctg gccagatgca gccggtaaag	180
gcattcaaaa tccacaacaa aatctggggt atcccggaaac gtgatacctt tactaacccg	240
gaagaagggtg acctgaaccc gccaccggaa gcgaaacagg tgccggtatc ttactatgac	300
tccacctacc tgtctaccga taacgaaaag gacaactacc tgaaagggtg tactaaactg	360
ttcgagcgta tttactccac cgacctgggc cgtatgctgc tgactagcat cgttcgcggt	420
atcccgttct ggggcggttc taccatcgat accgaactga aagtaatcga cactaactgc	480
atcaacgtta ttcagccgga cggttcctat cgttccgaag aactgaacct ggtgatcatc	540
ggcccgtctg ctgatatcat ccagttcgag tgtaagagct ttggtcacga agttctgaac	600
ctcaccgta acggctacgg ttccactcag tacatccgtt tctctccgga cttcaccttc	660
ggttttgaag aatccctgga agtagacacg aaccactgc tgggcgctgg taaattcgca	720
actgatcctg cggttaccct ggctcacgaa ctgattcatg caggccaccg cctgtacggt	780
atcgccatca atccgaaccg tgtcttcaaa gttaacacca acgcgtatta cgagatgtcc	840
ggctctggaag ttagcttcga agaactgcgt acttttgcg gtcacgacgc taaattcatc	900
gactctctgc aagaaaacga gttccgtctg tactactata acaagttaa agatatcgca	960
tccaccctga acaaagcgaa atccatcgtg ggtaccactg cttctctcca gtacatgaag	1020
aacgttttta aagaaaaata cctgctcagc gaagacacct ccggcaaatt ctctgtagac	1080
aagttgaaat tcgataaact ttacaaaatg ctgactgaaa ttacaccga agacaacttc	1140
gttaagttct ttaaagttct gaaccgcaaa acctatctga acttcgacaa ggcagtattc	1200

ES 2 562 425 T3

aaaatcaaca tcgtgccgaa agttaactac actatctacg atggtttcaa cctgcgtaac 1260  
 accaactcgg ctgctaattt taacggccag aacacggaaa tcaacaacat gaacttcaca 1320  
 aaactgaaaa acttcactgg tctgttcgag ttttacaagc tgctgtgctg cgacggcctc 1380  
 attacctcca aaactaaatc tctgatagaa ggtagaaaca aagcgtgaa cgacctctgt 1440  
 atcaaggtta acaactggga tttattcttc agccccagtg aagacaactt caccaacgac 1500  
 ctgaacaaag gtgaagaaat cacctcagat actaacatcg aagcagccga agaaaacatc 1560  
 tcgctggacc tgatccagca gtactacctg acctttaatt tcgacaacga gccggaaaaac 1620  
 atttctatcg aaaacctgag ctctgatatc atcggccagc tggaactgat gccgaacatc 1680  
 gaacgtttcc caaacggtaa aaagtacgag ctggacaaat ataccatggt cccactacctg 1740  
 cgcgcgcagg aatttgaaca cggcaaatcc cgtatcgac tgactaactc cgttaacgaa 1800  
 gctctgctca acccgtcccg tgtatacacc ttcttctcta gcgactacgt gaaaaaggtc 1860  
 aacaaagcga ctgaagctgc aatgttcttg gggtgggttg aacagcttgt ttatgatttt 1920  
 accgacgaga cgtccgaagt atctactacc gacaaaattg cggatatcac tatcatcatc 1980  
 ccgtacatcg gtccggctct gaacattggc aacatgctgt acaaagacga cttcgttggc 2040  
 gcactgatct tctccggtgc ggtgatcctg ctggagttca tcccggaaat cgccatcccg 2100  
 gtactgggca cctttgctct ggtttcttac attgcaaaca aggttctgac tgtacaaacc 2160  
 atcgacaacg cgctgagcaa acgtaacgaa aaatgggatg aagtttacia atatatcgtg 2220  
 accaactggc tggctaaggt taatactcag atcgacctca tccgcaaaaa aatgaaagaa 2280  
 gcactggaaa accaggcgga agctaccaag gcaatcatta actaccagta caaccagtac 2340  
 accgaggaag aaaaaaacia catcaacttc aacatcgacg atctgtctc taaactgaac 2400  
 gaatccatca acaaagctat gatcaacatc aacaagttcc tgaaccagtg ctctgtaagc 2460  
 tatctgatga actccatgat cccgtacggt gttaaactgc tggaggactt cgatgcgtct 2520  
 ctgaaagacg ccttgctgaa atacatttac gacaaccgtg gcactctgat cggtcagggt 2580  
 gatcgtctga aggacaaagt gaacaatacc ttatcgaccg acatcccttt tcagctcagt 2640  
 aaatatgctg ataaccaacg ccttttgtcc acttgtggcg gtggcggtag catcgaaggt 2700  
 cgtcagcact ggtcctattg cctgcgccct ggttgataa 2739

<210> 4  
 <211> 911  
 5 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> Secuencia sintética

10 <400> 4

Met His Gly Ser Ser Asn Asn  
 1 5 10 15

Asn Asn Asn Asn Asn Asn Asn Asn Gly Ser Met Glu Phe Val Asn Lys

ES 2 562 425 T3

20					25					30					
Gln	Phe	Asn	Tyr	Lys	Asp	Pro	Val	Asn	Gly	Val	Asp	Ile	Ala	Tyr	Ile
		35					40					45			
Lys	Ile	Pro	Asn	Ala	Gly	Gln	Met	Gln	Pro	Val	Lys	Ala	Phe	Lys	Ile
	50					55					60				
His	Asn	Lys	Ile	Trp	Val	Ile	Pro	Glu	Arg	Asp	Thr	Phe	Thr	Asn	Pro
65					70					75					80
Glu	Glu	Gly	Asp	Leu	Asn	Pro	Pro	Pro	Glu	Ala	Lys	Gln	Val	Pro	Val
				85					90					95	
Ser	Tyr	Tyr	Asp	Ser	Thr	Tyr	Leu	Ser	Thr	Asp	Asn	Glu	Lys	Asp	Asn
			100					105					110		
Tyr	Leu	Lys	Gly	Val	Thr	Lys	Leu	Phe	Glu	Arg	Ile	Tyr	Ser	Thr	Asp
		115					120					125			
Leu	Gly	Arg	Met	Leu	Leu	Thr	Ser	Ile	Val	Arg	Gly	Ile	Pro	Phe	Trp
	130					135					140				
Gly	Gly	Ser	Thr	Ile	Asp	Thr	Glu	Leu	Lys	Val	Ile	Asp	Thr	Asn	Cys
145					150					155					160
Ile	Asn	Val	Ile	Gln	Pro	Asp	Gly	Ser	Tyr	Arg	Ser	Glu	Glu	Leu	Asn
				165					170					175	
Leu	Val	Ile	Ile	Gly	Pro	Ser	Ala	Asp	Ile	Ile	Gln	Phe	Glu	Cys	Lys
			180					185					190		
Ser	Phe	Gly	His	Glu	Val	Leu	Asn	Leu	Thr	Arg	Asn	Gly	Tyr	Gly	Ser
		195					200					205			
Thr	Gln	Tyr	Ile	Arg	Phe	Ser	Pro	Asp	Phe	Thr	Phe	Gly	Phe	Glu	Glu
	210					215					220				
Ser	Leu	Glu	Val	Asp	Thr	Asn	Pro	Leu	Leu	Gly	Ala	Gly	Lys	Phe	Ala
225					230					235					240
Thr	Asp	Pro	Ala	Val	Thr	Leu	Ala	His	Glu	Leu	Ile	His	Ala	Gly	His
				245					250					255	
Arg	Leu	Tyr	Gly	Ile	Ala	Ile	Asn	Pro	Asn	Arg	Val	Phe	Lys	Val	Asn
			260					265					270		
Thr	Asn	Ala	Tyr	Tyr	Glu	Met	Ser	Gly	Leu	Glu	Val	Ser	Phe	Glu	Glu
		275					280					285			
Leu	Arg	Thr	Phe	Gly	Gly	His	Asp	Ala	Lys	Phe	Ile	Asp	Ser	Leu	Gln
	290					295					300				

ES 2 562 425 T3

Glu Asn Glu Phe Arg Leu Tyr Tyr Tyr Asn Lys Phe Lys Asp Ile Ala  
 305 310 315 320  
 Ser Thr Leu Asn Lys Ala Lys Ser Ile Val Gly Thr Thr Ala Ser Leu  
 325 330 335  
 Gln Tyr Met Lys Asn Val Phe Lys Glu Lys Tyr Leu Leu Ser Glu Asp  
 340 345 350  
 Thr Ser Gly Lys Phe Ser Val Asp Lys Leu Lys Phe Asp Lys Leu Tyr  
 355 360 365  
 Lys Met Leu Thr Glu Ile Tyr Thr Glu Asp Asn Phe Val Lys Phe Phe  
 370 375 380  
 Lys Val Leu Asn Arg Lys Thr Tyr Leu Asn Phe Asp Lys Ala Val Phe  
 385 390 395 400  
 Lys Ile Asn Ile Val Pro Lys Val Asn Tyr Thr Ile Tyr Asp Gly Phe  
 405 410 415  
 Asn Leu Arg Asn Thr Asn Leu Ala Ala Asn Phe Asn Gly Gln Asn Thr  
 420 425 430  
 Glu Ile Asn Asn Met Asn Phe Thr Lys Leu Lys Asn Phe Thr Gly Leu  
 435 440 445  
 Phe Glu Phe Tyr Lys Leu Leu Cys Val Asp Gly Ile Ile Thr Ser Lys  
 450 455 460  
 Thr Lys Ser Leu Ile Glu Gly Arg Asn Lys Ala Leu Asn Asp Leu Cys  
 465 470 475 480  
 Ile Lys Val Asn Asn Trp Asp Leu Phe Phe Ser Pro Ser Glu Asp Asn  
 485 490 495  
 Phe Thr Asn Asp Leu Asn Lys Gly Glu Glu Ile Thr Ser Asp Thr Asn  
 500 505 510  
 Ile Glu Ala Ala Glu Glu Asn Ile Ser Leu Asp Leu Ile Gln Gln Tyr  
 515 520 525  
 Tyr Leu Thr Phe Asn Phe Asp Asn Glu Pro Glu Asn Ile Ser Ile Glu  
 530 535 540  
 Asn Leu Ser Ser Asp Ile Ile Gly Gln Leu Glu Leu Met Pro Asn Ile  
 545 550 555 560  
 Glu Arg Phe Pro Asn Gly Lys Lys Tyr Glu Leu Asp Lys Tyr Thr Met  
 565 570 575

ES 2 562 425 T3

Phe His Tyr Leu Arg Ala Gln Glu Phe Glu His Gly Lys Ser Arg Ile  
 580 585 590  
 Ala Leu Thr Asn Ser Val Asn Glu Ala Leu Leu Asn Pro Ser Arg Val  
 595 600 605  
 Tyr Thr Phe Phe Ser Ser Asp Tyr Val Lys Lys Val Asn Lys Ala Thr  
 610 615 620  
 Glu Ala Ala Met Phe Leu Gly Trp Val Glu Gln Leu Val Tyr Asp Phe  
 625 630 635 640  
 Thr Asp Glu Thr Ser Glu Val Ser Thr Thr Asp Lys Ile Ala Asp Ile  
 645 650 655  
 Thr Ile Ile Ile Pro Tyr Ile Gly Pro Ala Leu Asn Ile Gly Asn Met  
 660 665 670  
 Leu Tyr Lys Asp Asp Phe Val Gly Ala Leu Ile Phe Ser Gly Ala Val  
 675 680 685  
 Ile Leu Leu Glu Phe Ile Pro Glu Ile Ala Ile Pro Val Leu Gly Thr  
 690 695 700  
 Phe Ala Leu Val Ser Tyr Ile Ala Asn Lys Val Leu Thr Val Gln Thr  
 705 710 715 720  
 Ile Asp Asn Ala Leu Ser Lys Arg Asn Glu Lys Trp Asp Glu Val Tyr  
 725 730 735  
 Lys Tyr Ile Val Thr Asn Trp Leu Ala Lys Val Asn Thr Gln Ile Asp  
 740 745 750  
 Leu Ile Arg Lys Lys Met Lys Glu Ala Leu Glu Asn Gln Ala Glu Ala  
 755 760 765  
 Thr Lys Ala Ile Ile Asn Tyr Gln Tyr Asn Gln Tyr Thr Glu Glu Glu  
 770 775 780  
 Lys Asn Asn Ile Asn Phe Asn Ile Asp Asp Leu Ser Ser Lys Leu Asn  
 785 790 795 800  
 Glu Ser Ile Asn Lys Ala Met Ile Asn Ile Asn Lys Phe Leu Asn Gln  
 805 810 815  
 Cys Ser Val Ser Tyr Leu Met Asn Ser Met Ile Pro Tyr Gly Val Lys  
 820 825 830  
 Arg Leu Glu Asp Phe Asp Ala Ser Leu Lys Asp Ala Leu Leu Lys Tyr  
 835 840 845  
 Ile Tyr Asp Asn Arg Gly Thr Leu Ile Gly Gln Val Asp Arg Leu Lys  
 850 855 860  
 Asp Lys Val Asn Asn Thr Leu Ser Thr Asp Ile Pro Phe Gln Leu Ser  
 865 870 875 880  
 Lys Tyr Val Asp Asn Gln Arg Leu Leu Ser Thr Cys Gly Gly Gly Gly  
 885 890 895  
 Ser Ile Glu Gly Arg Gln His Trp Ser Tyr Cys Leu Arg Pro Gly  
 900 905 910

ES 2 562 425 T3

<210> 5  
 <211> 2772  
 <212> ADN  
 5 <213> Secuencia artificial

<220>  
 <223> Secuencia sintética

10 <400> 5

```

atgcatcacc atcaccatca ccatcaccat catgggagct cgaacaataa caacaataac      60
aataacaata acggatccat gacgtggcca gttaaggatt tcaactactc agatcctgta      120
aatgacaacg atattctgta ccttcgcatt ccacaaaata aactgatcac cacaccagtc      180
aaagcattca tgattactca aaacatttgg gtcattccag aacgcttttc tagtgacaca      240
aatccgagtt tatctaaacc tccgcgtccg acgtccaaat atcagagcta ttacgatccc      300
tcatatctca gtacggacga acaaaaagat actttcctta aaggatcat taaactgttt      360
aagcgtatta atgagcgcga tatcgggaaa aagttgatta attatcttgt tgtgggttcc      420
ccgttcatgg gcgatagctc taccgccgaa gacacttttg attttaccg tcatacgaca      480
aacatcgcgg tagagaagtt tgagaacgga tcgtggaaag tcacaaacat cattacacct      540
agcgtcttaa tttttggtcc gctgccaaac atcttagatt atacagccag cctgactttg      600
caggggcaac agtcgaatcc gagtttcgaa ggttttggtg ccctgagcat tctgaaagtt      660
gccccggaat ttctgctcac tttttcagat gtcaccagca accagagctc agcagtatta      720
ggaaagtcaa ttttttgcag ggacccgggtt attgcaactga tgcacgaact gacgactct      780
ctgcatcaac tgtatgggat caacatcccc agtgacaaac gtattcgtcc ccagggtgtct      840
gaaggatttt tctcacagga tgggccgaac gtccagttcg aagagttgta tactttcgga      900
ggcctggacg tagagatcat tccccagatt gagcgcagtc agctgctgta gaaggcattg      960
ggccattata aggatattgc aaaacgcctg aataacatta acaaaacgat tccatcttcg     1020
tggatctcga atattgataa atataagaaa atttttagcg agaaatataa ttttgataaa     1080
gataatacag gtaactttgt ggtaaacatt gacaaattca actcccttta cagtgatttg     1140
acgaatgtaa tgagcgaagt tgtgtatagt tccaataca acgtaagaa tcgtacccat     1200
tacttctctc gtcactacct gccggtttcc gcgaacatcc ttgacgataa tatttacact     1260
attcgtgacg gctttaactt gaccaacaag ggcttcaata ttgaaaattc aggccagaac     1320
    
```

ES 2 562 425 T3

attgaacgca acccggcctt gcagaaactg tcgagtgaaat ccggtggttga cctgtttacc 1380  
 aaagtctgcy tcgacggcat cattacctcc aaaactaaat ctctgataga aggtagaaac 1440  
 aaagcgtga acctgcagtg tattaagtg aaaaacaatc ggctgcctta ttagcagat 1500  
 aaagatagca ttagtcagga gattttcgaa aataaaatta tcaactgacga aaccaatggt 1560  
 cagaattatt cagataaatt ttcactggac gaaagcatct tagatggcca agttccgatt 1620  
 aaccggaaa ttgttgatcc gttactgccg aacgtgaata tggaccggtt aaacctcct 1680  
 ggcaagaga tcgtatttta tgatgacatt acgaaatag tggactacct taattcttat 1740  
 tactatttgg aaagccagaa actgtccaat aacgtggaaa acattactct gaccacaagc 1800  
 gtggaagagg ctttaggcta ctcaaataag atttatacct tcctcccgtc gctggcggaa 1860  
 aaagtaaata aaggtgtgca ggctggtctg ttctcaact gggcgaatga agttgtcgaa 1920  
 gactttacca cgaatattat gaaaaaggat accctggata aaatctccga cgtctcgggt 1980  
 attatcccat atattggccc tgcgttaaat atcggtaata gtgcgctgcy ggggaatttt 2040  
 aaccaggcct ttgctaccgc gggcgtcgcg ttctcctgg agggctttcc tgaatttact 2100  
 atcccggcgc tcggtgtttt tacattttac tcttccatcc aggagcgtga gaaaattatc 2160  
 aaaaccatcg aaaactgcct ggagcagcgg gtgaaacgct ggaaagattc ttatcaatgg 2220  
 atggtgtcaa actggttattc tcgcatcacg acccaattca accatattaa ttaccagatg 2280  
 tatgatagtc tgcgtacca agctgacgcc attaaagcca aaattgatct ggaatataaa 2340  
 aagtactctg gtagcgataa ggagaacatc aaaagccagg tggagaacct taagaatagt 2400  
 ctggatgtga aaatctctga agctatgaat aacattaaca aattcattcg tgaatgttcg 2460  
 gtgacgtacc tgttcaagaa tatgctgcca aaagttattg atgaactgaa taaatttgat 2520  
 ctgctacca aaaccgaaact tatcaacctc atcgactccc acaacattat ccttggtggc 2580  
 gaagtggatc gtctgaaggg caaagtaaac gagagcttg aaaatacgat gccgtttaat 2640  
 atttttcat ataccaataa ctcttgctg aaagatatca tcaatgaata tttcaatcta 2700  
 gaatgtggcg gtggcggtag cgacgatgac gataaacagc actggtccta ttgcctgcy 2760  
 cctggttgat aa 2772

<210> 6  
 <211> 922  
 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> Secuencia sintética

<400> 6

Met His His His His His His His His His Gly Ser Ser Asn Asn  
 1 5 10 15

Asn Asn Asn Asn Asn Asn Asn Gly Ser Met Thr Trp Pro Val Lys  
 20 25 30

ES 2 562 425 T3

Asp Phe Asn Tyr Ser Asp Pro Val Asn Asp Asn Asp Ile Leu Tyr Leu  
 35 40 45  
 Arg Ile Pro Gln Asn Lys Leu Ile Thr Thr Pro Val Lys Ala Phe Met  
 50 55 60  
 Ile Thr Gln Asn Ile Trp Val Ile Pro Glu Arg Phe Ser Ser Asp Thr  
 65 70 75 80  
 Asn Pro Ser Leu Ser Lys Pro Pro Arg Pro Thr Ser Lys Tyr Gln Ser  
 85 90 95  
 Tyr Tyr Asp Pro Ser Tyr Leu Ser Thr Asp Glu Gln Lys Asp Thr Phe  
 100 105 110  
 Leu Lys Gly Ile Ile Lys Leu Phe Lys Arg Ile Asn Glu Arg Asp Ile  
 115 120 125  
 Gly Lys Lys Leu Ile Asn Tyr Leu Val Val Gly Ser Pro Phe Met Gly  
 130 135 140  
 Asp Ser Ser Thr Pro Glu Asp Thr Phe Asp Phe Thr Arg His Thr Thr  
 145 150 155 160  
 Asn Ile Ala Val Glu Lys Phe Glu Asn Gly Ser Trp Lys Val Thr Asn  
 165 170 175  
 Ile Ile Thr Pro Ser Val Leu Ile Phe Gly Pro Leu Pro Asn Ile Leu  
 180 185 190  
 Asp Tyr Thr Ala Ser Leu Thr Leu Gln Gly Gln Gln Ser Asn Pro Ser  
 195 200 205  
 Phe Glu Gly Phe Gly Thr Leu Ser Ile Leu Lys Val Ala Pro Glu Phe  
 210 215 220  
 Leu Leu Thr Phe Ser Asp Val Thr Ser Asn Gln Ser Ser Ala Val Leu  
 225 230 235 240  
 Gly Lys Ser Ile Phe Cys Met Asp Pro Val Ile Ala Leu Met His Glu  
 245 250 255  
 Leu Thr His Ser Leu His Gln Leu Tyr Gly Ile Asn Ile Pro Ser Asp  
 260 265 270  
 Lys Arg Ile Arg Pro Gln Val Ser Glu Gly Phe Phe Ser Gln Asp Gly  
 275 280 285  
 Pro Asn Val Gln Phe Glu Glu Leu Tyr Thr Phe Gly Gly Leu Asp Val  
 290 295 300  
 Glu Ile Ile Pro Gln Ile Glu Arg Ser Gln Leu Arg Glu Lys Ala Leu



ES 2 562 425 T3

Glu Asn Ile Thr Leu Thr Thr Ser Val Glu Glu Ala Leu Gly Tyr Ser  
 595 600 605  
 Asn Lys Ile Tyr Thr Phe Leu Pro Ser Leu Ala Glu Lys Val Asn Lys  
 610 615 620  
 Gly Val Gln Ala Gly Leu Phe Leu Asn Trp Ala Asn Glu Val Val Glu  
 625 630 635 640  
 Asp Phe Thr Thr Asn Ile Met Lys Lys Asp Thr Leu Asp Lys Ile Ser  
 645 650 655  
 Asp Val Ser Val Ile Ile Pro Tyr Ile Gly Pro Ala Leu Asn Ile Gly  
 660 665 670  
 Asn Ser Ala Leu Arg Gly Asn Phe Asn Gln Ala Phe Ala Thr Ala Gly  
 675 680 685  
 Val Ala Phe Leu Leu Glu Gly Phe Pro Glu Phe Thr Ile Pro Ala Leu  
 690 695 700  
 Gly Val Phe Thr Phe Tyr Ser Ser Ile Gln Glu Arg Glu Lys Ile Ile  
 705 710 715 720  
 Lys Thr Ile Glu Asn Cys Leu Glu Gln Arg Val Lys Arg Trp Lys Asp  
 725 730 735  
 Ser Tyr Gln Trp Met Val Ser Asn Trp Leu Ser Arg Ile Thr Thr Gln  
 740 745 750  
 Phe Asn His Ile Asn Tyr Gln Met Tyr Asp Ser Leu Ser Tyr Gln Ala  
 755 760 765  
 Asp Ala Ile Lys Ala Lys Ile Asp Leu Glu Tyr Lys Lys Tyr Ser Gly  
 770 775 780  
 Ser Asp Lys Glu Asn Ile Lys Ser Gln Val Glu Asn Leu Lys Asn Ser  
 785 790 795 800  
 Leu Asp Val Lys Ile Ser Glu Ala Met Asn Asn Ile Asn Lys Phe Ile  
 805 810 815  
 Arg Glu Cys Ser Val Thr Tyr Leu Phe Lys Asn Met Leu Pro Lys Val  
 820 825 830  
 Ile Asp Glu Leu Asn Lys Phe Asp Leu Arg Thr Lys Thr Glu Leu Ile  
 835 840 845  
 Asn Leu Ile Asp Ser His Asn Ile Ile Leu Val Gly Glu Val Asp Arg  
 850 855 860

ES 2 562 425 T3

Leu Lys Ala Lys Val Asn Glu Ser Phe Glu Asn Thr Met Pro Phe Asn  
865 870 875 880

Ile Phe Ser Tyr Thr Asn Asn Ser Leu Leu Lys Asp Ile Ile Asn Glu  
885 890 895

Tyr Phe Asn Leu Glu Cys Gly Gly Gly Gly Ser Asp Asp Asp Asp Lys  
900 905 910

Gln His Trp Ser Tyr Cys Leu Arg Pro Gly  
915 920

<210> 7

<211> 2850

5 <212> ADN

<213> Secuencia artificial

<220>

<223> Secuencia sintética

10

<400> 7

atgcatcacc	atcaccatca	ccatcaccat	catgggagct	cgaacaataa	caacaataac	60
aataacaata	acggatccat	gacgtggcca	gttaaggatt	tcaactactc	agatcctgta	120
aatgacaacg	atattctgta	ccttcgcatt	ccacaaaata	aactgatcac	cacaccagtc	180
aaagcattca	tgattactca	aaacatttgg	gtcattccag	aacgcttttc	tagtgacaca	240
aatccgagtt	tatctaaacc	tccgcgtccg	acgtccaaat	atcagagcta	ttacgatccc	300
tcatatctca	gtacggacga	acaaaaagat	actttcctta	aaggatcat	taaactgttt	360
aagcgtatta	atgagcgcga	tatcgggaaa	aagttgatta	attatcttgt	tgtgggttcc	420
ccgttcatgg	gcgatagctc	tacccccgaa	gacacttttg	atcttaccg	tcatacgaca	480
aacatcgcg	tagagaagtt	tgagaacgga	tcgtggaaag	tcacaaacat	cattacacct	540
agcgtcttaa	tttttggtcc	gctgccaaac	atcttagatt	atacagccag	cctgactttg	600
caggggcaac	agtcgaatcc	gagtttcgaa	ggttttggta	ccctgagcat	tctgaaagtt	660
gccccggaat	ttctgctcac	tttttcagat	gtcaccagca	accagagctc	agcagtatta	720
ggaaagtcaa	ttttttgcat	ggacccgggt	attgcaactga	tgacgaact	gacgcactct	780
ctgcatcaac	tgtatgggat	caacatcccc	agtgacaaac	gtattcgtcc	ccaggtgtct	840
gaaggatttt	tctcacagga	tgggccgaac	gtccagttcg	aagagttgta	tactttcgga	900
ggcctggacg	tagagatcat	tccccagatt	gagcgcagtc	agctgcgtga	gaaggcattg	960
ggcattata	aggatattgc	aaaacgcctg	aataacatta	acaaaacgat	tccatcttcg	1020
tggatctcga	atattgataa	atataagaaa	atTTTTtagcg	agaaatataa	ttttgataaa	1080
gataatacag	gtaactttgt	ggttaacatt	gacaaattca	actcccttta	cagtgatttg	1140
acgaatgtaa	tgagcgaagt	tgtgtatagt	tccaataca	acgtaagaa	tcgtacccat	1200
tacttctctc	gtcactacct	gccggttttc	gcgaacatcc	ttgacgataa	tatttacact	1260
attcgtgacg	gctttaactt	gaccaacaag	ggcttcaata	ttgaaaattc	aggccagaac	1320

ES 2 562 425 T3

attgaacgca acccggcctt gcagaaactg tcgagtgaaat ccgtggttga cctgtttacc 1380  
aaagtctgcy tcgacggcat cattacctcc aaaactaaat ctctgataga aggtagaaac 1440  
aaagcgtga acctgcagtg tattaaagtg aaaaacaatc ggctgcctta tgtagcagat 1500  
aaagatagca ttagtcagga gattttcgaa aataaaatta tcaactgacga aaccaatggt 1560  
cagaattatt cagataaatt ttcactggac gaaagcatct tagatggcca agttccgatt 1620  
aaccggaaa ttgttgatcc gttactgccg aacgtgaata tggaaaccgtt aaacctccct 1680  
ggcgaagaga tcgtatttta tgatgacatt acgaaatag tggactacct taattcttat 1740  
tactatttgg aaagccagaa actgtccaat aacgtggaaa acattactct gaccacaagc 1800  
gtggaagagg ctttaggcta ctcaaataag atttatacct tcctcccgtc gctggcggaa 1860  
aaagtaaata aagggtgaga ggctggctcg ttcctcaact gggcgaatga agttgtcgaa 1920  
gactttacca cgaatattat gaaaaaggat accctggata aaatctccga cgtctcgggt 1980  
attatcccat atattggccc tgcgttaaat atcggtaata gtgcgtgcy ggggaatttt 2040  
aaccaggcct ttgctaccgc gggcgtcgcg ttcctcctgg agggccttcc tgaatttact 2100  
atccccgcy tcggtgtttt tacattttac tcttccatcc aggagcgtga gaaaattatc 2160  
aaaaccatcg aaaactgcct ggagcagcgg gtgaaacgct ggaaagattc ttatcaatgg 2220  
atggtgtcaa actggttatc tcgcatcacg acccaattca accatattaa ttaccagatg 2280  
tatgatagtc tgtcgtacca agctgacgcc attaaagcca aaattgatct ggaatataaa 2340  
aagtactctg gtacgataa ggagaacatc aaaagccagg tggagaacct taagaatagt 2400  
ctggatgga aaatctctga agctatgaat aacattaaca aattcattcg tgaatgttcg 2460  
gtgacgtacc tgttcaagaa tatgctgcca aaagttattg atgaactgaa taaatttgat 2520  
ctgctacca aaaccgaact tatcaacctc atcgactccc acaacattat ccttgtgggc 2580  
gaagtggatc gtctgaaggc caaagtaaac gagagctttg aaaatagat gccgtttaat 2640  
atTTTTTcat ataccaataa ctcttgcctg aaagatatca tcaatgaata tttcaatcta 2700  
gaatgtggcg gtggcggtag catcgaaggt cgttatggag gttttttgag aaggatagca 2760  
ccaaaattaa agtgggataa tcaaggcggg gggggtagtt gcggcggtag cggttcgcyt 2820  
ccgccgggtt tctctccgtt ccgttgataa 2850

<210> 8  
<211> 948  
5 <212> PRT  
<213> Secuencia artificial

<220>  
10 <223> Secuencia sintética

<400> 8

Met His Gly Ser Ser Asn Asn  
1 5 10 15

Asn Asn Asn Asn Asn Asn Asn Asn Gly Ser Met Thr Trp Pro Val Lys

ES 2 562 425 T3

20					25					30					
Asp	Phe	Asn	Tyr	Ser	Asp	Pro	Val	Asn	Asp	Asn	Asp	Ile	Leu	Tyr	Leu
		35					40					45			
Arg	Ile	Pro	Gln	Asn	Lys	Leu	Ile	Thr	Thr	Pro	Val	Lys	Ala	Phe	Met
	50					55					60				
Ile	Thr	Gln	Asn	Ile	Trp	Val	Ile	Pro	Glu	Arg	Phe	Ser	Ser	Asp	Thr
65					70					75					80
Asn	Pro	Ser	Leu	Ser	Lys	Pro	Pro	Arg	Pro	Thr	Ser	Lys	Tyr	Gln	Ser
				85					90					95	
Tyr	Tyr	Asp	Pro	Ser	Tyr	Leu	Ser	Thr	Asp	Glu	Gln	Lys	Asp	Thr	Phe
			100					105					110		
Leu	Lys	Gly	Ile	Ile	Lys	Leu	Phe	Lys	Arg	Ile	Asn	Glu	Arg	Asp	Ile
		115					120					125			
Gly	Lys	Lys	Leu	Ile	Asn	Tyr	Leu	Val	Val	Gly	Ser	Pro	Phe	Met	Gly
	130					135					140				
Asp	Ser	Ser	Thr	Pro	Glu	Asp	Thr	Phe	Asp	Phe	Thr	Arg	His	Thr	Thr
145					150					155					160
Asn	Ile	Ala	Val	Glu	Lys	Phe	Glu	Asn	Gly	Ser	Trp	Lys	Val	Thr	Asn
				165					170					175	
Ile	Ile	Thr	Pro	Ser	Val	Leu	Ile	Phe	Gly	Pro	Leu	Pro	Asn	Ile	Leu
			180					185					190		
Asp	Tyr	Thr	Ala	Ser	Leu	Thr	Leu	Gln	Gly	Gln	Gln	Ser	Asn	Pro	Ser
		195				200						205			
Phe	Glu	Gly	Phe	Gly	Thr	Leu	Ser	Ile	Leu	Lys	Val	Ala	Pro	Glu	Phe
	210					215					220				
Leu	Leu	Thr	Phe	Ser	Asp	Val	Thr	Ser	Asn	Gln	Ser	Ser	Ala	Val	Leu
225					230					235					240
Gly	Lys	Ser	Ile	Phe	Cys	Met	Asp	Pro	Val	Ile	Ala	Leu	Met	His	Glu
				245					250					255	
Leu	Thr	His	Ser	Leu	His	Gln	Leu	Tyr	Gly	Ile	Asn	Ile	Pro	Ser	Asp
			260					265					270		
Lys	Arg	Ile	Arg	Pro	Gln	Val	Ser	Glu	Gly	Phe	Phe	Ser	Gln	Asp	Gly
		275					280					285			
Pro	Asn	Val	Gln	Phe	Glu	Glu	Leu	Tyr	Thr	Phe	Gly	Gly	Leu	Asp	Val
	290					295					300				

ES 2 562 425 T3

Glu Ile Ile Pro Gln Ile Glu Arg Ser Gln Leu Arg Glu Lys Ala Leu  
 305 310 315 320  
 Gly His Tyr Lys Asp Ile Ala Lys Arg Leu Asn Asn Ile Asn Lys Thr  
 325 330 335  
 Ile Pro Ser Ser Trp Ile Ser Asn Ile Asp Lys Tyr Lys Lys Ile Phe  
 340 345 350  
 Ser Glu Lys Tyr Asn Phe Asp Lys Asp Asn Thr Gly Asn Phe Val Val  
 355 360 365  
 Asn Ile Asp Lys Phe Asn Ser Leu Tyr Ser Asp Leu Thr Asn Val Met  
 370 375 380  
 Ser Glu Val Val Tyr Ser Ser Gln Tyr Asn Val Lys Asn Arg Thr His  
 385 390 395  
 Tyr Phe Ser Arg His Tyr Leu Pro Val Phe Ala Asn Ile Leu Asp Asp  
 405 410 415  
 Asn Ile Tyr Thr Ile Arg Asp Gly Phe Asn Leu Thr Asn Lys Gly Phe  
 420 425 430  
 Asn Ile Glu Asn Ser Gly Gln Asn Ile Glu Arg Asn Pro Ala Leu Gln  
 435 440 445  
 Lys Leu Ser Ser Glu Ser Val Val Asp Leu Phe Thr Lys Val Cys Val  
 450 455 460  
 Asp Gly Ile Ile Thr Ser Lys Thr Lys Ser Leu Ile Glu Gly Arg Asn  
 465 470 475 480  
 Lys Ala Leu Asn Leu Gln Cys Ile Lys Val Lys Asn Asn Arg Leu Pro  
 485 490 495  
 Tyr Val Ala Asp Lys Asp Ser Ile Ser Gln Glu Ile Phe Glu Asn Lys  
 500 505 510  
 Ile Ile Thr Asp Glu Thr Asn Val Gln Asn Tyr Ser Asp Lys Phe Ser  
 515 520 525  
 Leu Asp Glu Ser Ile Leu Asp Gly Gln Val Pro Ile Asn Pro Glu Ile  
 530 535 540  
 Val Asp Pro Leu Leu Pro Asn Val Asn Met Glu Pro Leu Asn Leu Pro  
 545 550 555 560  
 Gly Glu Glu Ile Val Phe Tyr Asp Asp Ile Thr Lys Tyr Val Asp Tyr  
 565 570 575

ES 2 562 425 T3

Leu Asn Ser Tyr Tyr Tyr Leu Glu Ser Gln Lys Leu Ser Asn Asn Val  
 580 585 590

Glu Asn Ile Thr Leu Thr Thr Ser Val Glu Glu Ala Leu Gly Tyr Ser  
 595 600 605

Asn Lys Ile Tyr Thr Phe Leu Pro Ser Leu Ala Glu Lys Val Asn Lys  
 610 615 620

Gly Val Gln Ala Gly Leu Phe Leu Asn Trp Ala Asn Glu Val Val Glu  
 625 630 635 640

Asp Phe Thr Thr Asn Ile Met Lys Lys Asp Thr Leu Asp Lys Ile Ser  
 645 650 655

Asp Val Ser Val Ile Ile Pro Tyr Ile Gly Pro Ala Leu Asn Ile Gly  
 660 665 670

Asn Ser Ala Leu Arg Gly Asn Phe Asn Gln Ala Phe Ala Thr Ala Gly  
 675 680 685

Val Ala Phe Leu Leu Glu Gly Phe Pro Glu Phe Thr Ile Pro Ala Leu  
 690 695 700

Gly Val Phe Thr Phe Tyr Ser Ser Ile Gln Glu Arg Glu Lys Ile Ile  
 705 710 715 720

Lys Thr Ile Glu Asn Cys Leu Glu Gln Arg Val Lys Arg Trp Lys Asp  
 725 730 735

Ser Tyr Gln Trp Met Val Ser Asn Trp Leu Ser Arg Ile Thr Thr Gln  
 740 745 750

Phe Asn His Ile Asn Tyr Gln Met Tyr Asp Ser Leu Ser Tyr Gln Ala  
 755 760 765

Asp Ala Ile Lys Ala Lys Ile Asp Leu Glu Tyr Lys Lys Tyr Ser Gly  
 770 775 780

Ser Asp Lys Glu Asn Ile Lys Ser Gln Val Glu Asn Leu Lys Asn Ser  
 785 790 795 800

Leu Asp Val Lys Ile Ser Glu Ala Met Asn Asn Ile Asn Lys Phe Ile  
 805 810 815

Arg Glu Cys Ser Val Thr Tyr Leu Phe Lys Asn Met Leu Pro Lys Val  
 820 825 830

Ile Asp Glu Leu Asn Lys Phe Asp Leu Arg Thr Lys Thr Glu Leu Ile  
 835 840 845

ES 2 562 425 T3

Asn Leu Ile Asp Ser His Asn Ile Ile Leu Val Gly Glu Val Asp Arg  
 850 855 860  
 Leu Lys Ala Lys Val Asn Glu Ser Phe Glu Asn Thr Met Pro Phe Asn  
 865 870 875 880  
 Ile Phe Ser Tyr Thr Asn Asn Ser Leu Leu Lys Asp Ile Ile Asn Glu  
 885 890 895  
 Tyr Phe Asn Leu Glu Cys Gly Gly Gly Gly Ser Ile Glu Gly Arg Tyr  
 900 905 910  
 Gly Gly Phe Leu Arg Arg Ile Arg Pro Lys Leu Lys Trp Asp Asn Gln  
 915 920 925  
 Gly Gly Gly Gly Ser Cys Gly Gly Gly Gly Ser Arg Pro Pro Gly Phe  
 930 935 940

Ser Pro Phe Arg  
 945

<210> 9  
 <211> 2862  
 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> Secuencia sintética

<400> 9

atgcatcacc atcaccatca ccatcaccat catgggagct cgaacaataa caacaataac	60
aataacaata acggatccat ggagttcggt aacaaacagt tcaactataa agaccagtt	120
aacggtgttg acattgctta catcaaaatc ccgaacgctg gccagatgca gccggtaaag	180
gcattcaaaa tccacaacaa aatctggggt atcccggaac gtgatacctt tactaacccg	240
gaagaagggtg acctgaacct gccaccgga gcgaaacagg tgccggtatc ttactatgac	300
tccacctacc tgtctaccga taacgaaaag gacaactacc tgaaagggtg tactaaactg	360
ttcgagcgta tttactccac cgacctgggc cgtatgctgc tgactagcat cgttcgcggt	420
atcccgttct ggggcgggtc taccatcgat accgaactga aagtaatcga cactaactgc	480
atcaacgtta ttcagccgga cggttcctat cgttcggaag aactgaacct ggtgatcatc	540
ggcccgtctg ctgatatcat ccagttcgag tgtaagagct ttggtcacga agttctgaac	600
ctcaccgta acggctacgg ttccactcag tacatccggt tctctccgga cttcaccttc	660
ggttttgaag aatccctgga agtagacacg aaccactgc tgggcgctgg taaattcgca	720
actgatcctg cggttaccct ggctcacgaa ctgattcatg caggccaccg cctgtacggt	780
atgccatca atccgaaccg tgtcttcaaa gttaacacca acgcgtatta cgagatgtcc	840
ggctctggaag ttagcttcga agaactgcgt acttttggcg gtcacgacgc taaattcatc	900
gactctctgc aagaaaacga gttccgtctg tactactata acaagttcaa agatatcgca	960

ES 2 562 425 T3

tccaccctga acaaagcgaa atccatcgtg ggtaccactg cttctctcca gtacatgaag 1020  
aacgttttta aagaaaaata cctgctcagc gaagacacct cggcaaatt ctctgtagac 1080  
aagttgaaat tcgataaact ttacaaaatg ctgactgaaa tttacaccga agacaacttc 1140  
gttaagttct ttaaagttct gaaccgcaaa acctatctga acttcgacaa ggcagtattc 1200  
aaaatcaaca tcgtgccgaa agttaactac actatctacg atggtttcaa cctgcgtaac 1260  
accaacctgg ctgctaattt taacggccag aacacggaaa tcaacaacat gaacttcaca 1320  
aaactgaaaa acttcactgg tctgttcgag ttttacaagc tgctgtgctg cgacggcatc 1380  
attacctcca aaactaaatc tctgatagaa ggtagaaaca aagcgtgaa cgacctctgt 1440  
atcaaggtta acaactggga tttattcttc agcccagtg aagacaactt caccaacgac 1500  
ctgaacaaag gtgaagaat cacctcagat actaacatcg aagcagccga agaaaaatc 1560  
tcgctggacc tgatccagca gtactacctg acctttaatt tcgacaacga gccgaaaac 1620  
atctctatcg aaaacctgag ctctgatatc atcggccagc tggaactgat gccgaacatc 1680  
gaacgtttcc caaacggtaa aaagtacgag ctggacaaat ataccatgtt cactacctg 1740  
cgcgcgcagg aatttgaaca cggcaaatcc cgtatcgac tgactaactc cgttaacgaa 1800  
gctctgctca acccgccccg tgtatacacc ttcttctcta gcgactacgt gaaaaaggtc 1860  
aacaagcga ctgaagctgc aatgttcttg ggttgggtg aacagcttg ttatgatatt 1920  
accgacgaga cgtccgaagt atctactacc gacaaaattg cggatatcac tatcatcatc 1980  
ccgtacatcg gtccggctct gaacattggc aacatgctgt acaaagacga cttcgttggc 2040  
gcactgatct tctccggtgc ggtgatcctg ctggagttca tcccggaaat cgccatcccg 2100  
gtactgggca cctttgctct ggtttcttac attgcaaaca aggttctgac tgtacaaacc 2160  
atcgacaacg cgctgagcaa acgtaacgaa aaatgggatg aagtttaca atatatcgtg 2220  
accaactggc tggctaaggt taatactcag atcgacctca tccgcaaaaa aatgaaagaa 2280  
gcactggaaa accagcgga agctaccaag gcaatcatta actaccagta caaccagtac 2340  
accgaggaag aaaaaaaca catcaacttc aacatcgac atctgtcctc taaactgaac 2400  
gaatccatca acaaagctat gatcaacatc aacaagttcc tgaaccagtg ctctgtaagc 2460  
tatctgatga actccatgat cccgtacggt gttaaacgtc tggaggactt cgatgctct 2520  
ctgaaagacg cctgctgaa atacatttac gacaaccgtg gcaactctgat cggtcagggt 2580  
gatcgtctga aggacaaagt gaacaatacc ttatcgaccg acatcccttt tcagctcagt 2640  
aaatatgctg ataaccaacg ccttttgtcc acttgtggcg gtggcggtag catcgaaggt 2700  
cgttacggtg gtttcatgac ctctgaaaaa tctcagaccg cgctggttac cctgttcaaa 2760  
aacgctatca tcaaaaacgc ttacaaaaaa ggtgaaggcg gtgggggtag ttgcggcgg 2820  
ggcggttcgc gtccgccggg tttctctccg ttccgttgat aa 2862

<210> 10  
<211> 952  
5 <212> PRT  
<213> Secuencia artificial

<220>  
10 <223> Secuencia sintética

<400> 10

ES 2 562 425 T3

Met His Gly Ser Ser Asn Asn  
1 5 10 15

Asn Asn Asn Asn Asn Asn Asn Asn Gly Ser Met Glu Phe Val Asn Lys  
20 25 30

Gln Phe Asn Tyr Lys Asp Pro Val Asn Gly Val Asp Ile Ala Tyr Ile  
35 40 45

Lys Ile Pro Asn Ala Gly Gln Met Gln Pro Val Lys Ala Phe Lys Ile  
50 55 60

His Asn Lys Ile Trp Val Ile Pro Glu Arg Asp Thr Phe Thr Asn Pro  
65 70 75 80

Glu Glu Gly Asp Leu Asn Pro Pro Pro Glu Ala Lys Gln Val Pro Val  
85 90 95

Ser Tyr Tyr Asp Ser Thr Tyr Leu Ser Thr Asp Asn Glu Lys Asp Asn  
100 105 110

Tyr Leu Lys Gly Val Thr Lys Leu Phe Glu Arg Ile Tyr Ser Thr Asp  
115 120 125

Leu Gly Arg Met Leu Leu Thr Ser Ile Val Arg Gly Ile Pro Phe Trp  
130 135 140

Gly Gly Ser Thr Ile Asp Thr Glu Leu Lys Val Ile Asp Thr Asn Cys  
145 150 155 160

Ile Asn Val Ile Gln Pro Asp Gly Ser Tyr Arg Ser Glu Glu Leu Asn  
165 170 175

Leu Val Ile Ile Gly Pro Ser Ala Asp Ile Ile Gln Phe Glu Cys Lys  
180 185 190

Ser Phe Gly His Glu Val Leu Asn Leu Thr Arg Asn Gly Tyr Gly Ser  
195 200 205

Thr Gln Tyr Ile Arg Phe Ser Pro Asp Phe Thr Phe Gly Phe Glu Glu  
210 215 220

Ser Leu Glu Val Asp Thr Asn Pro Leu Leu Gly Ala Gly Lys Phe Ala  
225 230 235 240

Thr Asp Pro Ala Val Thr Leu Ala His Glu Leu Ile His Ala Gly His  
245 250 255

ES 2 562 425 T3

Arg Leu Tyr Gly Ile Ala Ile Asn Pro Asn Arg Val Phe Lys Val Asn  
 260 265 270  
 Thr Asn Ala Tyr Tyr Glu Met Ser Gly Leu Glu Val Ser Phe Glu Glu  
 275 280 285  
 Leu Arg Thr Phe Gly Gly His Asp Ala Lys Phe Ile Asp Ser Leu Gln  
 290 295 300  
 Glu Asn Glu Phe Arg Leu Tyr Tyr Tyr Asn Lys Phe Lys Asp Ile Ala  
 305 310 315 320  
 Ser Thr Leu Asn Lys Ala Lys Ser Ile Val Gly Thr Thr Ala Ser Leu  
 325 330 335  
 Gln Tyr Met Lys Asn Val Phe Lys Glu Lys Tyr Leu Leu Ser Glu Asp  
 340 345 350  
 Thr Ser Gly Lys Phe Ser Val Asp Lys Leu Lys Phe Asp Lys Leu Tyr  
 355 360 365  
 Lys Met Leu Thr Glu Ile Tyr Thr Glu Asp Asn Phe Val Lys Phe Phe  
 370 375 380  
 Lys Val Leu Asn Arg Lys Thr Tyr Leu Asn Phe Asp Lys Ala Val Phe  
 385 390 395 400  
 Lys Ile Asn Ile Val Pro Lys Val Asn Tyr Thr Ile Tyr Asp Gly Phe  
 405 410 415  
 Asn Leu Arg Asn Thr Asn Leu Ala Ala Asn Phe Asn Gly Gln Asn Thr  
 420 425 430  
 Glu Ile Asn Asn Met Asn Phe Thr Lys Leu Lys Asn Phe Thr Gly Leu  
 435 440 445  
 Phe Glu Phe Tyr Lys Leu Leu Cys Val Asp Gly Ile Ile Thr Ser Lys  
 450 455 460  
 Thr Lys Ser Leu Ile Glu Gly Arg Asn Lys Ala Leu Asn Asp Leu Cys  
 465 470 475 480  
 Ile Lys Val Asn Asn Trp Asp Leu Phe Phe Ser Pro Ser Glu Asp Asn  
 485 490 495  
 Phe Thr Asn Asp Leu Asn Lys Gly Glu Glu Ile Thr Ser Asp Thr Asn  
 500 505 510  
 Ile Glu Ala Ala Glu Glu Asn Ile Ser Leu Asp Leu Ile Gln Gln Tyr  
 515 520 525

ES 2 562 425 T3

Tyr Leu Thr Phe Asn Phe Asp Asn Glu Pro Glu Asn Ile Ser Ile Glu  
 530 535 540

Asn Leu Ser Ser Asp Ile Ile Gly Gln Leu Glu Leu Met Pro Asn Ile  
 545 550 555 560

Glu Arg Phe Pro Asn Gly Lys Lys Tyr Glu Leu Asp Lys Tyr Thr Met  
 565 570 575

Phe His Tyr Leu Arg Ala Gln Glu Phe Glu His Gly Lys Ser Arg Ile  
 580 585 590

Ala Leu Thr Asn Ser Val Asn Glu Ala Leu Leu Asn Pro Ser Arg Val  
 595 600 605

Tyr Thr Phe Phe Ser Ser Asp Tyr Val Lys Lys Val Asn Lys Ala Thr  
 610 615 620

Glu Ala Ala Met Phe Leu Gly Trp Val Glu Gln Leu Val Tyr Asp Phe  
 625 630 635 640

Thr Asp Glu Thr Ser Glu Val Ser Thr Thr Asp Lys Ile Ala Asp Ile  
 645 650 655

Thr Ile Ile Ile Pro Tyr Ile Gly Pro Ala Leu Asn Ile Gly Asn Met  
 660 665 670

Leu Tyr Lys Asp Asp Phe Val Gly Ala Leu Ile Phe Ser Gly Ala Val  
 675 680 685

Ile Leu Leu Glu Phe Ile Pro Glu Ile Ala Ile Pro Val Leu Gly Thr  
 690 695 700

Phe Ala Leu Val Ser Tyr Ile Ala Asn Lys Val Leu Thr Val Gln Thr  
 705 710 715 720

Ile Asp Asn Ala Leu Ser Lys Arg Asn Glu Lys Trp Asp Glu Val Tyr  
 725 730 735

Lys Tyr Ile Val Thr Asn Trp Leu Ala Lys Val Asn Thr Gln Ile Asp  
 740 745 750

Leu Ile Arg Lys Lys Met Lys Glu Ala Leu Glu Asn Gln Ala Glu Ala  
 755 760 765

Thr Lys Ala Ile Ile Asn Tyr Gln Tyr Asn Gln Tyr Thr Glu Glu Glu  
 770 775 780

Lys Asn Asn Ile Asn Phe Asn Ile Asp Asp Leu Ser Ser Lys Leu Asn  
 785 790 795 800

Glu Ser Ile Asn Lys Ala Met Ile Asn Ile Asn Lys Phe Leu Asn Gln

ES 2 562 425 T3

	805		810		815			
Cys	Ser Val	Ser Tyr	Leu Met	Asn Ser	Met Ile	Pro Tyr	Gly Val	Lys
		820		825			830	
Arg	Leu Glu	Asp Phe	Asp Ala	Ser Leu	Lys Asp	Ala Leu	Leu Lys	Tyr
	835			840		845		
Ile	Tyr Asp	Asn Arg	Gly Thr	Leu Ile	Gly Gln	Val Asp	Arg Leu	Lys
	850		855			860		
Asp	Lys Val	Asn Asn	Thr Leu	Ser Thr	Asp Ile	Pro Phe	Gln Leu	Ser
865			870		875			880
Lys	Tyr Val	Asp Asn	Gln Arg	Leu Leu	Ser Thr	Cys Gly	Gly Gly	Gly
		885			890		895	
Ser	Ile Glu	Gly Arg	Tyr Gly	Gly Gly	Phe Met	Thr Ser	Glu Lys	Ser
		900			905		910	Gln
Thr	Pro Leu	Val Thr	Leu Phe	Lys Asn	Ala Ile	Ile Ile	Lys Asn	Ala Tyr
	915			920			925	
Lys	Lys Gly	Glu Gly	Gly Gly	Gly Gly	Ser Cys	Gly Gly	Gly Gly	Ser Arg
	930		935			940		
Pro	Pro Gly	Phe Ser	Pro Phe	Arg				
945			950					

<210> 11  
 <211> 3012  
 5 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> Secuencia sintética

10 <400> 11

atgcatcacc atcaccatca ccatcaccat catgggagct cgaacaataa caacaataac	60
aataacaata acggatccat gacgtggcca gttaaggatt tcaactactc agatcctgta	120
aatgacaacg atattctgta ccttcgcatt ccacaaaata aactgatcac cacaccagtc	180
aaagatttca tgattactca aaacatttgg gtcattccag aacgcttttc tagtgacaca	240
aatccgagtt tatctaaacc tccgcgtccg acgtccaaat atcagagcta ttacgatccc	300
tcatatctca gtacggacga acaaaaagat actttcctta aaggtatcat taaactgttt	360
aagcgtatta atgagcgcga tatcgggaaa aagttgatta attatcttgt tgtgggttcc	420
ccgttcattg gcgatagctc taccgccgaa gacacttttg attttaccg tcatacgaca	480
aacatcgcg tagagaagtt tgagaacgga tcgtggaaag tcacaaacat cattacacct	540
agcgtcttaa tttttggtcc gctgccaaac atcttagatt atacagccag cctgactttg	600
caggggcaac agtcgaatcc gagtttcgaa ggttttggtta ccctgagcat tctgaaagtt	660

ES 2 562 425 T3

gccccggaat ttctgctcac tttttcagat gtcaccagca accagagctc agcagtatta 720  
 ggaaagtcaa ttttttgcac ggacccgggtt attgcaactga tgcacgaact gacgcaactct 780  
 ctgcatcaac tgtatgggat caacatcccc agtgacaaac gtattcgtcc ccagggtgtct 840  
 gaaggatttt tctcacagga tgggccgaac gtccagttcg aagagttgta tactttcgga 900  
 ggcctggacg tagagatcat tccccagatt gagcgcagtc agctgcgtga gaaggcattg 960  
 ggccattata aggatattgc aaaacgcctg aataacatta acaaaacgat tccatcttcg 1020  
 tggatctcga atattgataa atataagaaa atttttagcg agaaatataa ttttgataaa 1080  
 gataatacag gtaactttgt ggtaaacatt gacaaattca actcccctta cagtgatttg 1140  
 acgaatgtaa tgagcgaagt tgtgtatagt tcccaatata acgtaagaa tcgtacccat 1200  
 tacttctctc gtcactacct gccggtttcc gcgaacatcc ttgacgataa tatttacact 1260  
 attcgtgacg gctttaactt gaccaacaag ggcttcaata ttgaaaattc aggccagaac 1320  
 attgaacgca acccggcctt gcagaaactg tcgagtgaat ccgtggttga cctgtttacc 1380  
 aaagtctgcy tgcacggcat cattacctcc aaaactaat ctctgataga aggtagaac 1440  
 aaagcgtga acctgcagtg tattaagtg aaaaacaatc ggctgcctta thtagcagat 1500  
 aaagatagca ttagtcagga gatcttcgaa aataaaatta tcaactgacga aaccaatggt 1560  
 cagaattatt cagataaatt ttcactggac gaaagcatct tagatggcca agttccgatt 1620  
 aaccggaaa ttgttgatcc gttactgccg aacgtgaata tggaaaccgtt aaacctccct 1680  
 ggcgaagaga tcgtatctta tgatgacatt acgaaatag tggactacct taattcttat 1740  
 tactatttgg aaagccagaa actgtccaat aacgtggaaa acattactct gaccacaagc 1800  
 gtggaagagg ctttaggcta ctcaaataag atttatacct tcctcccgtc gctggcggaa 1860  
 aaagtaaata aaggtgtgca ggctggctg ttctcaact gggcgaatga agttgtcgaa 1920  
 gactttacca cgaatattat gaaaaaggat accctggata aaatctccga cgtctcgggt 1980  
 attatcccat atattggccc tgcgttaaat atcggtaata gtgcgctgcy ggggaatttt 2040  
 aaccaggcct ttgctaccgc gggcgtcgcg ttctctctgg agggctttcc tgaatttact 2100  
 atcccgccgc tcggtgtttt tacattttac tcttccatcc aggagcgtga gaaaattatc 2160  
 aaaaccatcg aaaactgcct ggagcagcgg gtgaaacgct ggaaagattc ttatcaatgg 2220  
 atggtgtcaa actggttata tgcgatcacg acccaattca accatattaa ttaccagatg 2280  
 tatgatagtc tgcgtacca agctgacgcc attaaagcca aaattgatct ggaatataaa 2340  
 aagtactctg gtgagcgtga ggagaacatc aaaagccagg tggagaacct taagaatagt 2400  
 ctggatgtga aaatctctga agctatgaat aacattaaca aattcattcg tgaatgttcg 2460  
 gtgacgtacc tgttcaagaa tatgctgcca aaagttattg atgaaactgaa taaatttgat 2520  
 ctgctacca aaaccgaact tatcaacctc atcgaactccc acaacattat cctgtggggc 2580  
 gaagtggatc gcttgaaggc caaagtaaac gagagctttg aaaatacgaat gccgtttaat 2640  
 attttttcat ataccaataa ctcttgctg aaagatatca tcaatgaata tttcaatcta 2700  
 gaatgtggcg gtggcggtag catcgaaggt cgtcacgtgg atgcgatctt cactcagtct 2760  
 taccgtaaag ttctggcgca gctgagcgtc cgtaaactgc tgcaggatat cctgaaccgt 2820  
 cagcagggtg aacgtaacca ggaacagggc gctggcggtg ggggtagttg cggcgggtggc 2880  
 ggttcgcacg tggatgcat ctctactcag tcttaccgta aagttctggc gcagctgagc 2940  
 gctcgtaaac tgcctcagga taccctgaac cgtcagcagg gtgaacgtaa ccaggaacag 3000  
 ggcgcttgat aa 3012

ES 2 562 425 T3

<210> 12  
 <211> 1002  
 <212> PRT

5 <213> Secuencia artificial

<220>  
 <223> Secuencia sintética

10 <400> 12

Met His Gly Ser Ser Asn Asn  
 1 5 10 15  
 Asn Asn Asn Asn Asn Asn Asn Asn Gly Ser Met Thr Trp Pro Val Lys  
 20 25 30  
 Asp Phe Asn Tyr Ser Asp Pro Val Asn Asp Asn Asp Ile Leu Tyr Leu  
 35 40 45  
 Arg Ile Pro Gln Asn Lys Leu Ile Thr Thr Pro Val Lys Ala Phe Met  
 50 55 60  
 Ile Thr Gln Asn Ile Trp Val Ile Pro Glu Arg Phe Ser Ser Asp Thr  
 65 70 75 80  
 Asn Pro Ser Leu Ser Lys Pro Pro Arg Pro Thr Ser Lys Tyr Gln Ser  
 85 90 95  
 Tyr Tyr Asp Pro Ser Tyr Leu Ser Thr Asp Glu Gln Lys Asp Thr Phe  
 100 105 110  
 Leu Lys Gly Ile Ile Lys Leu Phe Lys Arg Ile Asn Glu Arg Asp Ile  
 115 120 125  
 Gly Lys Lys Leu Ile Asn Tyr Leu Val Val Gly Ser Pro Phe Met Gly  
 130 135 140  
 Asp Ser Ser Thr Pro Glu Asp Thr Phe Asp Phe Thr Arg His Thr Thr  
 145 150 155 160  
 Asn Ile Ala Val Glu Lys Phe Glu Asn Gly Ser Trp Lys Val Thr Asn  
 165 170 175  
 Ile Ile Thr Pro Ser Val Leu Ile Phe Gly Pro Leu Pro Asn Ile Leu  
 180 185 190

ES 2 562 425 T3

Asp Tyr Thr Ala Ser Leu Thr Leu Gln Gly Gln Gln Ser Asn Pro Ser  
 195 200 205  
 Phe Glu Gly Phe Gly Thr Leu Ser Ile Leu Lys Val Ala Pro Glu Phe  
 210 215 220  
 Leu Leu Thr Phe Ser Asp Val Thr Ser Asn Gln Ser Ser Ala Val Leu  
 225 230 235  
 Gly Lys Ser Ile Phe Cys Met Asp Pro Val Ile Ala Leu Met His Glu  
 245 250 255  
 Leu Thr His Ser Leu His Gln Leu Tyr Gly Ile Asn Ile Pro Ser Asp  
 260 265 270  
 Lys Arg Ile Arg Pro Gln Val Ser Glu Gly Phe Phe Ser Gln Asp Gly  
 275 280 285  
 Pro Asn Val Gln Phe Glu Glu Leu Tyr Thr Phe Gly Gly Leu Asp Val  
 290 295 300  
 Glu Ile Ile Pro Gln Ile Glu Arg Ser Gln Leu Arg Glu Lys Ala Leu  
 305 310 315 320  
 Gly His Tyr Lys Asp Ile Ala Lys Arg Leu Asn Asn Ile Asn Lys Thr  
 325 330 335  
 Ile Pro Ser Ser Trp Ile Ser Asn Ile Asp Lys Tyr Lys Lys Ile Phe  
 340 345 350  
 Ser Glu Lys Tyr Asn Phe Asp Lys Asp Asn Thr Gly Asn Phe Val Val  
 355 360 365  
 Asn Ile Asp Lys Phe Asn Ser Leu Tyr Ser Asp Leu Thr Asn Val Met  
 370 375 380  
 Ser Glu Val Val Tyr Ser Ser Gln Tyr Asn Val Lys Asn Arg Thr His  
 385 390 395 400  
 Tyr Phe Ser Arg His Tyr Leu Pro Val Phe Ala Asn Ile Leu Asp Asp  
 405 410 415  
 Asn Ile Tyr Thr Ile Arg Asp Gly Phe Asn Leu Thr Asn Lys Gly Phe  
 420 425 430  
 Asn Ile Glu Asn Ser Gly Gln Asn Ile Glu Arg Asn Pro Ala Leu Gln  
 435 440 445  
 Lys Leu Ser Ser Glu Ser Val Val Asp Leu Phe Thr Lys Val Cys Val  
 450 455 460

ES 2 562 425 T3

Asp Gly Ile Ile Thr Ser Lys Thr Lys Ser Leu Ile Glu Gly Arg Asn  
 465 470 475 480  
 Lys Ala Leu Asn Leu Gln Cys Ile Lys Val Lys Asn Asn Arg Leu Pro  
 485 490 495  
 Tyr Val Ala Asp Lys Asp Ser Ile Ser Gln Glu Ile Phe Glu Asn Lys  
 500 505 510  
 Ile Ile Thr Asp Glu Thr Asn Val Gln Asn Tyr Ser Asp Lys Phe Ser  
 515 520 525  
 Leu Asp Glu Ser Ile Leu Asp Gly Gln Val Pro Ile Asn Pro Glu Ile  
 530 535 540  
 Val Asp Pro Leu Leu Pro Asn Val Asn Met Glu Pro Leu Asn Leu Pro  
 545 550 555 560  
 Gly Glu Glu Ile Val Phe Tyr Asp Asp Ile Thr Lys Tyr Val Asp Tyr  
 565 570 575  
 Leu Asn Ser Tyr Tyr Tyr Leu Glu Ser Gln Lys Leu Ser Asn Asn Val  
 580 585 590  
 Glu Asn Ile Thr Leu Thr Thr Ser Val Glu Glu Ala Leu Gly Tyr Ser  
 595 600 605  
 Asn Lys Ile Tyr Thr Phe Leu Pro Ser Leu Ala Glu Lys Val Asn Lys  
 610 615 620  
 Gly Val Gln Ala Gly Leu Phe Leu Asn Trp Ala Asn Glu Val Val Glu  
 625 630 635 640  
 Asp Phe Thr Thr Asn Ile Met Lys Lys Asp Thr Leu Asp Lys Ile Ser  
 645 650 655  
 Asp Val Ser Val Ile Ile Pro Tyr Ile Gly Pro Ala Leu Asn Ile Gly  
 660 665 670  
 Asn Ser Ala Leu Arg Gly Asn Phe Asn Gln Ala Phe Ala Thr Ala Gly  
 675 680 685  
 Val Ala Phe Leu Leu Glu Gly Phe Pro Glu Phe Thr Ile Pro Ala Leu  
 690 695 700  
 Gly Val Phe Thr Phe Tyr Ser Ser Ile Gln Glu Arg Glu Lys Ile Ile  
 705 710 715 720  
 Lys Thr Ile Glu Asn Cys Leu Glu Gln Arg Val Lys Arg Trp Lys Asp  
 725 730 735

ES 2 562 425 T3

Ser Tyr Gln Trp Met Val Ser Asn Trp Leu Ser Arg Ile Thr Thr Gln  
740 745 750

Phe Asn His Ile Asn Tyr Gln Met Tyr Asp Ser Leu Ser Tyr Gln Ala  
755 760 765

Asp Ala Ile Lys Ala Lys Ile Asp Leu Glu Tyr Lys Lys Tyr Ser Gly  
770 775 780

Ser Asp Lys Glu Asn Ile Lys Ser Gln Val Glu Asn Leu Lys Asn Ser  
785 790 795 800

Leu Asp Val Lys Ile Ser Glu Ala Met Asn Asn Ile Asn Lys Phe Ile  
805 810 815

Arg Glu Cys Ser Val Thr Tyr Leu Phe Lys Asn Met Leu Pro Lys Val  
820 825 830

Ile Asp Glu Leu Asn Lys Phe Asp Leu Arg Thr Lys Thr Glu Leu Ile  
835 840 845

Asn Leu Ile Asp Ser His Asn Ile Ile Leu Val Gly Glu Val Asp Arg  
850 855 860

Leu Lys Ala Lys Val Asn Glu Ser Phe Glu Asn Thr Met Pro Phe Asn  
865 870 875 880

Ile Phe Ser Tyr Thr Asn Asn Ser Leu Leu Lys Asp Ile Ile Asn Glu  
885 890 895

Tyr Phe Asn Leu Glu Cys Gly Gly Gly Gly Ser Ile Glu Gly Arg His  
900 905 910

Val Asp Ala Ile Phe Thr Gln Ser Tyr Arg Lys Val Leu Ala Gln Leu  
915 920 925

Ser Ala Arg Lys Leu Leu Gln Asp Ile Leu Asn Arg Gln Gln Gly Glu  
930 935 940

Arg Asn Gln Glu Gln Gly Ala Gly Gly Gly Gly Ser Cys Gly Gly Gly  
945 950 955 960

Gly Ser His Val Asp Ala Ile Phe Thr Gln Ser Tyr Arg Lys Val Leu  
965 970 975

Ala Gln Leu Ser Ala Arg Lys Leu Leu Gln Asp Ile Leu Asn Arg Gln  
980 985 990

Gln Gly Glu Arg Asn Gln Glu Gln Gly Ala  
995 1000

<210> 13  
<211> 2814  
5 <212> ADN  
<213> Secuencia artificial

<220>  
<223> Secuencia sintética

10 <400> 13

ES 2 562 425 T3

atgcatcacc atcaccatca ccatcaccat catgggagct cgaacaataa caacaataac	60
aataacaata acggatccat gacgtggcca gttaaggatt tcaactactc agatcctgta	120
aatgacaacg atattctgta ccttcgcatt ccacaaaata aactgatcac cacaccagtc	180
aaagcattca tgattactca aaacatttgg gtcattccag aacgcttttc tagtgacaca	240
aatccgagtt tatctaaacc tccgcgtccg acgtccaaat atcagagcta ttacgatccc	300
tcatatctca gtacggacga acaaaaagat actttcctta aaggatcat taaactgttt	360
aagcgtatta atgagcgcga tatcgggaaa aagttgatta attatcttgt tgtgggttcc	420
ccgttcattg gcgatagctc taccctcgaa gacacttttg attttaccg tcatacgaca	480
aacatcgcg tagagaagtt tgagaacgga tcgtggaaag tcacaaacat cattacacct	540
agcgtcttaa tttttggtcc gctgccaaac atcttagatt atacagccag cctgactttg	600
caggggcaac agtccaatcc gagtttcgaa ggttttgta ccctgagcat tctgaaagtt	660
gccccggaat tctgctcac ttttccagat gtcaccagca accagagctc agcagtatta	720
ggaaagtcaa tttttgcat ggaccgggtt attgactga tgcacgaact gacgactct	780
ctgcatcaac tgtatgggat caacatcccc agtgacaaac gtattcgtcc ccagggtgtct	840
gaaggatttt tctcacagga tgggccgaa gtccagttcg aagagttgta tactttcgga	900
ggcctggacg tagagatcat tccccagatt gagcgcagtc agctgcgtga gaaggcattg	960
ggccattata aggatattgc aaaacgcctg aataacatta acaaaacgat tccatcttcg	1020
tggatctcga atattgataa atataagaaa atttttagcg agaaatataa ttttgataaa	1080
gataatacag gtaactttgt ggtaacatt gacaaattca actcccttta cagtgatttg	1140
acgaatgtaa tgagcgaagt tgtgtatagt tccaataca acgtaagaa tcgtacctat	1200
tacttctctc gtcactacct gccggtttcc gcgaacatcc ttgacgataa tatttacct	1260
attcgtgacg gctttaactt gaccaacaag ggcttcaata ttgaaaattc aggccagaac	1320
attgaacgca acccggcctt gcagaaactg tcgagtgaat ccgtggttga cctgtttacc	1380
aaagtctgcg tcgacggcat cattacctcc aaaactaaat ctctgataga aggtagaaac	1440
aaagcgtga acctgcagtg tattaagtg aaaaacaatc ggctgcctta tgtagcagat	1500
aaagatagca ttagtcagga gatcttcgaa aataaaatta tcaactgacga aaccaatggt	1560
cagaattatt cagataaatt ttcactggac gaaagcatct tagatggcca agttccgatt	1620
aaccgggaaa ttgttgatcc gttactgccg aacgtgaata tggaaccggt aaacctcct	1680
ggcgaagaga tcgtatctta tgatgacatt acgaaatag tggactacct taattcttat	1740
tactatcttg aaagccagaa actgtccaat aacgtggaaa acattactct gaccacaagc	1800
gtggaagagg ctttaggcta ctcaaataag atttatacct tcctcccgtc gctggcggaa	1860

ES 2 562 425 T3

aaagtaaata aagggtgtgca ggctggctctg ttcctcaact gggcgaatga agttgtcgaa 1920  
gactttacca cgaatattat gaaaaaggat accctggata aaatctccga cgtctcggtt 1980  
attatcccat atattggccc tgcgttaaat atcggtaata gtgcgctgcg ggggaatttt 2040  
aaccaggcct ttgctaccgc gggcgtcgcg ttcctcctgg agggctttcc tgaatttact 2100  
atcccggcgc tcggtgtttt tacattttac tcttccatcc aggagcgtga gaaaattatc 2160  
aaaaccatcg aaaactgcct ggagcagcgg gtgaaacgct ggaaagattc ttatcaatgg 2220  
atgggtgcaa actggttatc tcgcatcacg acccaattca accatattaa ttaccagatg 2280  
tatgatagtc tgtcgtacca agctgacgcc attaaagcca aaattgatct ggaatataaa 2340  
aagtactctg gttagcgtgaa ggagaacatc aaaagccagg tggagaacct taagaatagt 2400  
ctggatgtga aaatctctga agctatgaat aacattaaca aattcattcg tgaatgttcg 2460  
gtgacgtacc tgttcaagaa tatgctgcc aagttattg atgaactgaa taaatttgat 2520  
ctgctacca aaaccgaact tatcaacctc atcgactccc acaacattat ccttgtgggc 2580  
gaagtggatc gtctgaaggc caaagtaaac gagagctttg aaaatacgat gccgtttaat 2640  
atTTTTTcat ataccaataa ctcttctgctg aaagatatca tcaatgaata tttcaatcta 2700  
gaatgtggtg gcggtggctc tgggtggcgc ggttctattg aaggtcgcg tggcgggtggc 2760  
tctggtggcg gcggttctca gcactggtcc tattgcctgc gccctggttg ataa 2814

<210> 14  
<211> 936  
5 <212> PRT  
<213> Secuencia artificial

<220>  
10 <223> Secuencia sintética

<400> 14

Met His Gly Ser Ser Asn Asn  
1 5 10 15  
Asn Asn Asn Asn Asn Asn Asn Asn Gly Ser Met Thr Trp Pro Val Lys  
20 25 30  
Asp Phe Asn Tyr Ser Asp Pro Val Asn Asp Asn Asp Ile Leu Tyr Leu  
35 40 45  
Arg Ile Pro Gln Asn Lys Leu Ile Thr Thr Pro Val Lys Ala Phe Met  
50 55 60  
Ile Thr Gln Asn Ile Trp Val Ile Pro Glu Arg Phe Ser Ser Asp Thr  
65 70 75 80  
Asn Pro Ser Leu Ser Lys Pro Pro Arg Pro Thr Ser Lys Tyr Gln Ser  
85 90 95  
Tyr Tyr Asp Pro Ser Tyr Leu Ser Thr Asp Glu Gln Lys Asp Thr Phe

ES 2 562 425 T3

100					105					110					
Leu	Lys	Gly	Ile	Ile	Lys	Leu	Phe	Lys	Arg	Ile	Asn	Glu	Arg	Asp	Ile
		115					120					125			
Gly	Lys	Lys	Leu	Ile	Asn	Tyr	Leu	Val	Val	Gly	Ser	Pro	Phe	Met	Gly
	130				135						140				
Asp	Ser	Ser	Thr	Pro	Glu	Asp	Thr	Phe	Asp	Phe	Thr	Arg	His	Thr	Thr
145					150					155					160
Asn	Ile	Ala	Val	Glu	Lys	Phe	Glu	Asn	Gly	Ser	Trp	Lys	Val	Thr	Asn
				165					170					175	
Ile	Ile	Thr	Pro	Ser	Val	Leu	Ile	Phe	Gly	Pro	Leu	Pro	Asn	Ile	Leu
			180					185					190		
Asp	Tyr	Thr	Ala	Ser	Leu	Thr	Leu	Gln	Gly	Gln	Gln	Ser	Asn	Pro	Ser
		195					200						205		
Phe	Glu	Gly	Phe	Gly	Thr	Leu	Ser	Ile	Leu	Lys	Val	Ala	Pro	Glu	Phe
	210					215					220				
Leu	Leu	Thr	Phe	Ser	Asp	Val	Thr	Ser	Asn	Gln	Ser	Ser	Ala	Val	Leu
225					230					235					240
Gly	Lys	Ser	Ile	Phe	Cys	Met	Asp	Pro	Val	Ile	Ala	Leu	Met	His	Glu
				245					250					255	
Leu	Thr	His	Ser	Leu	His	Gln	Leu	Tyr	Gly	Ile	Asn	Ile	Pro	Ser	Asp
			260					265					270		
Lys	Arg	Ile	Arg	Pro	Gln	Val	Ser	Glu	Gly	Phe	Phe	Ser	Gln	Asp	Gly
		275					280					285			
Pro	Asn	Val	Gln	Phe	Glu	Glu	Leu	Tyr	Thr	Phe	Gly	Gly	Leu	Asp	Val
	290					295					300				
Glu	Ile	Ile	Pro	Gln	Ile	Glu	Arg	Ser	Gln	Leu	Arg	Glu	Lys	Ala	Leu
305					310					315					320
Gly	His	Tyr	Lys	Asp	Ile	Ala	Lys	Arg	Leu	Asn	Asn	Ile	Asn	Lys	Thr
				325					330					335	
Ile	Pro	Ser	Ser	Trp	Ile	Ser	Asn	Ile	Asp	Lys	Tyr	Lys	Lys	Ile	Phe
			340					345					350		
Ser	Glu	Lys	Tyr	Asn	Phe	Asp	Lys	Asp	Asn	Thr	Gly	Asn	Phe	Val	Val
		355					360					365			
Asn	Ile	Asp	Lys	Phe	Asn	Ser	Leu	Tyr	Ser	Asp	Leu	Thr	Asn	Val	Met
	370					375					380				

ES 2 562 425 T3

Ser Glu Val Val Tyr Ser Ser Gln Tyr Asn Val Lys Asn Arg Thr His  
 385 390 395 400  
 Tyr Phe Ser Arg His Tyr Leu Pro Val Phe Ala Asn Ile Leu Asp Asp  
 405 410 415  
 Asn Ile Tyr Thr Ile Arg Asp Gly Phe Asn Leu Thr Asn Lys Gly Phe  
 420 425 430  
 Asn Ile Glu Asn Ser Gly Gln Asn Ile Glu Arg Asn Pro Ala Leu Gln  
 435 440 445  
 Lys Leu Ser Ser Glu Ser Val Val Asp Leu Phe Thr Lys Val Cys Val  
 450 455 460  
 Asp Gly Ile Ile Thr Ser Lys Thr Lys Ser Leu Ile Glu Gly Arg Asn  
 465 470 475 480  
 Lys Ala Leu Asn Leu Gln Cys Ile Lys Val Lys Asn Asn Arg Leu Pro  
 485 490 495  
 Tyr Val Ala Asp Lys Asp Ser Ile Ser Gln Glu Ile Phe Glu Asn Lys  
 500 505 510  
 Ile Ile Thr Asp Glu Thr Asn Val Gln Asn Tyr Ser Asp Lys Phe Ser  
 515 520 525  
 Leu Asp Glu Ser Ile Leu Asp Gly Gln Val Pro Ile Asn Pro Glu Ile  
 530 535 540  
 Val Asp Pro Leu Leu Pro Asn Val Asn Met Glu Pro Leu Asn Leu Pro  
 545 550 555 560  
 Gly Glu Glu Ile Val Phe Tyr Asp Asp Ile Thr Lys Tyr Val Asp Tyr  
 565 570 575  
 Leu Asn Ser Tyr Tyr Tyr Leu Glu Ser Gln Lys Leu Ser Asn Asn Val  
 580 585 590  
 Glu Asn Ile Thr Leu Thr Thr Ser Val Glu Glu Ala Leu Gly Tyr Ser  
 595 600 605  
 Asn Lys Ile Tyr Thr Phe Leu Pro Ser Leu Ala Glu Lys Val Asn Lys  
 610 615 620  
 Gly Val Gln Ala Gly Leu Phe Leu Asn Trp Ala Asn Glu Val Val Glu  
 625 630 635 640  
 Asp Phe Thr Thr Asn Ile Met Lys Lys Asp Thr Leu Asp Lys Ile Ser  
 645 650 655

ES 2 562 425 T3

Asp Val Ser Val Ile Ile Pro Tyr Ile Gly Pro Ala Leu Asn Ile Gly  
 660 665 670

Asn Ser Ala Leu Arg Gly Asn Phe Asn Gln Ala Phe Ala Thr Ala Gly  
 675 680 685

Val Ala Phe Leu Leu Glu Gly Phe Pro Glu Phe Thr Ile Pro Ala Leu  
 690 695 700

Gly Val Phe Thr Phe Tyr Ser Ser Ile Gln Glu Arg Glu Lys Ile Ile  
 705 710 715 720

Lys Thr Ile Glu Asn Cys Leu Glu Gln Arg Val Lys Arg Trp Lys Asp  
 725 730 735

Ser Tyr Gln Trp Met Val Ser Asn Trp Leu Ser Arg Ile Thr Thr Gln  
 740 745 750

Phe Asn His Ile Asn Tyr Gln Met Tyr Asp Ser Leu Ser Tyr Gln Ala  
 755 760 765

Asp Ala Ile Lys Ala Lys Ile Asp Leu Glu Tyr Lys Lys Tyr Ser Gly  
 770 775 780

Ser Asp Lys Glu Asn Ile Lys Ser Gln Val Glu Asn Leu Lys Asn Ser  
 785 790 795 800

Leu Asp Val Lys Ile Ser Glu Ala Met Asn Asn Ile Asn Lys Phe Ile  
 805 810 815

Arg Glu Cys Ser Val Thr Tyr Leu Phe Lys Asn Met Leu Pro Lys Val  
 820 825 830

Ile Asp Glu Leu Asn Lys Phe Asp Leu Arg Thr Lys Thr Glu Leu Ile  
 835 840 845

Asn Leu Ile Asp Ser His Asn Ile Ile Leu Val Gly Glu Val Asp Arg  
 850 855 860

Leu Lys Ala Lys Val Asn Glu Ser Phe Glu Asn Thr Met Pro Phe Asn  
 865 870 875 880

Ile Phe Ser Tyr Thr Asn Asn Ser Leu Leu Lys Asp Ile Ile Asn Glu  
 885 890 895

Tyr Phe Asn Leu Glu Cys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser  
 900 905 910

Ile Glu Gly Arg Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gln His  
 915 920 925

Trp Ser Tyr Cys Leu Arg Pro Gly  
 930 935

5 <210> 15  
 <211> 2793  
 <212> ADN  
 <213> Secuencia artificial

10 <220>  
 <223> Secuencia sintética

ES 2 562 425 T3

<400> 15

atgcatcacc atcaccatca ccatcaccat catgggagct cgaacaataa caacaataac	60
aataacaata acggatccat ggagttcgtt aacaaacagt tcaactataa agacccagtt	120
aacggtggtg acattgctta catcaaaatc ccgaacgctg gccagatgca gccggtaaag	180
gcattcaaaa tccacaacaa aatctggggtt atcccggaac gtgatacctt tactaacccg	240
gaagaagggtg acctgaacct gccaccggaa gcgaaacagg tgccggatc ttactatgac	300
tccacctacc tgtctaccga taacgaaaag gacaactacc tgaaagggtg tactaaactg	360
ttcgagcgta tttactccac cgacctgggc cgtatgctgc tgactagcat cgttcgcggt	420
atcccgttct ggggcggttc taccatcgat accgaactga aagtaatcga cactaactgc	480
atcaacgtta ttcagccgga cggttcctat cgttccgaag aactgaacct ggtgatcatt	540
ggccccgtctg ctgatattcat ccagttcgag tgtaagagct ttggtcacga agttctgaac	600
ctcaccgta acggctacgg ttccactcag tacatccgtt tctctccgga cttcaccttc	660
ggttttgaag aatccctgga agtagacacg aaccactgc tgggcgctgg taaattcgca	720
actgatcctg cggttaccct ggctcacgaa ctgattcatg caggccaccg cctgtacggt	780
atcgccatca atccgaaccg tgtcttcaaa gttaacacca acgcgtatta cgagatgtcc	840
ggtctggaag ttgacttcga agaactgcgt acttttgcg gtcacgacgc taaattcatc	900
gactctctgc aagaaaacga gttccgtctg tactactata acaagttaa agatatcgca	960
tccaccctga acaaagcgaa atccatcgtg ggtaccactg cttctctcca gtacatgaag	1020
aacgttttta aagaaaata cctgctcagc gaagacacct ccggcaaatt ctctgtagac	1080
aagttgaaat tcgataaact ttacaaaatg ctgactgaaa tttacaccga agacaacttc	1140
gttaagtctt ttaaagtctt gaaccgcaa acctatctga acttcgacaa gccagatttc	1200
aaaatcaaca tcgtgccgaa agttaactac actatctacg atggtttcaa cctgcgtaac	1260
accaacctgg ctgctaattt taacggccag aacacggaaa tcaacaacat gaacttcaca	1320
aaactgaaaa acttcaactg tctgttcgag ttttacaagc tgctgtgctg cgacggcatc	1380
attacctcca aaactaaatc tctgatagaa ggtagaaaca aagcctgaa cgacctctgt	1440
atcaaggtta acaactggga tttattcttc agcccagtg aagacaactt caccaacgac	1500
ctgaacaaag gtgaagaaat cacctcagat actaacatcg aagcagccga agaaaacatc	1560
tcgctggacc tgatccagca gtactacctg acctttaatt tcgacaacga gccggaaaac	1620
atctctatcg aaaacctgag ctctgatatc atcggccagc tggaactgat gccgaacatc	1680

ES 2 562 425 T3

gaacgtttcc caaacggtaa aaagtacgag ctggacaaat ataccatggt cactacctg 1740  
 cgcgcgagg aatttgaaca cggcaaatcc cgtatcgac tgactaactc cgттаacgaa 1800  
 gctctgctca acccgtcccg tgtatacacc ttcttctcta gcgactacgt gaaaaaggtc 1860  
 aacaaagcga ctgaagctgc aatgttcttg ggttgggttg aacagcttgt ttatgatttt 1920  
 accgacgaga cgtccgaagt atctactacc gacaaaattg cggatatac tatcatcatc 1980  
 ccgtacatcg gtccggctct gaacattggc aacatgctgt acaaagacga cttcgttggc 2040  
 gcactgatct tctccggtgc ggtgatcctg ctggagttca tcccggaaat cgccatcccc 2100  
 gtactgggca cctttgctct ggtttcttac attgcaaaca aggttctgac tgtacaaacc 2160  
 atcgacaacg cgctgagcaa acgtaacgaa aaatgggatg aagtttaca atatatcgtg 2220  
 accaactggc tggctaaggt taatactcag atcgacctca tccgcaaaaa aatgaaagaa 2280  
 gcactggaaa accaggcgga agctaccaag gcaatcatta actaccagta caaccagtac 2340  
 accgaggaag aaaaaaaca catcaacttc aacatcgacg atctgtcctc taaactgaac 2400  
 gaatccatca acaaagctat gatcaacatc aacaagttcc tgaaccagtg ctctgtaagc 2460  
 tatctgatga actccatgat cccgtacggt gttaaactgc tggaggactt cgatgctct 2520  
 ctgaaagacg ccttgctgaa atacatttac gacaaccgtg gcactctgat cggtcagggt 2580  
 gatcgtctga aggacaaagt gaacaatacc ttatcgaccg acatcccttt tcagctcagt 2640  
 aaatatgctg ataaccaacg ccttttgtcc acttgtggcg gtggcggtag catcgaaggt 2700  
 cgtgttccat taccagcagg aggaggaaca gtattgacta aaatgtatcc atgcggaaat 2760  
 cactgggcag tgggacatct aatgggatga taa 2793

<210> 16  
 <211> 929  
 5 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> Secuencia sintética

10 <400> 16

Met His Gly Ser Ser Asn Asn  
 1 5 10 15  
 Asn Asn Asn Asn Asn Asn Asn Gly Ser Met Glu Phe Val Asn Lys  
 20 25 30  
 Gln Phe Asn Tyr Lys Asp Pro Val Asn Gly Val Asp Ile Ala Tyr Ile  
 35 40 45  
 Lys Ile Pro Asn Ala Gly Gln Met Gln Pro Val Lys Ala Phe Lys Ile  
 50 55 60  
 His Asn Lys Ile Trp Val Ile Pro Glu Arg Asp Thr Phe Thr Asn Pro  
 65 70 75 80

ES 2 562 425 T3

Glu Glu Gly Asp Leu Asn Pro Pro Pro Glu Ala Lys Gln Val Pro Val  
 85 90 95  
 Ser Tyr Tyr Asp Ser Thr Tyr Leu Ser Thr Asp Asn Glu Lys Asp Asn  
 100 105 110  
 Tyr Leu Lys Gly Val Thr Lys Leu Phe Glu Arg Ile Tyr Ser Thr Asp  
 115 120 125  
 Leu Gly Arg Met Leu Leu Thr Ser Ile Val Arg Gly Ile Pro Phe Trp  
 130 135 140  
 Gly Gly Ser Thr Ile Asp Thr Glu Leu Lys Val Ile Asp Thr Asn Cys  
 145 150 155 160  
 Ile Asn Val Ile Gln Pro Asp Gly Ser Tyr Arg Ser Glu Glu Leu Asn  
 165 170 175  
 Leu Val Ile Ile Gly Pro Ser Ala Asp Ile Ile Gln Phe Glu Cys Lys  
 180 185 190  
 Ser Phe Gly His Glu Val Leu Asn Leu Thr Arg Asn Gly Tyr Gly Ser  
 195 200 205  
 Thr Gln Tyr Ile Arg Phe Ser Pro Asp Phe Thr Phe Gly Phe Glu Glu  
 210 215 220  
 Ser Leu Glu Val Asp Thr Asn Pro Leu Leu Gly Ala Gly Lys Phe Ala  
 225 230 235 240  
 Thr Asp Pro Ala Val Thr Leu Ala His Glu Leu Ile His Ala Gly His  
 245 250 255  
 Arg Leu Tyr Gly Ile Ala Ile Asn Pro Asn Arg Val Phe Lys Val Asn  
 260 265 270  
 Thr Asn Ala Tyr Tyr Glu Met Ser Gly Leu Glu Val Ser Phe Glu Glu  
 275 280 285  
 Leu Arg Thr Phe Gly Gly His Asp Ala Lys Phe Ile Asp Ser Leu Gln  
 290 295 300  
 Glu Asn Glu Phe Arg Leu Tyr Tyr Tyr Asn Lys Phe Lys Asp Ile Ala  
 305 310 315 320  
 Ser Thr Leu Asn Lys Ala Lys Ser Ile Val Gly Thr Thr Ala Ser Leu  
 325 330 335  
 Gln Tyr Met Lys Asn Val Phe Lys Glu Lys Tyr Leu Leu Ser Glu Asp  
 340 345 350  
 Thr ser Gly Lys Phe Ser Val Asp Lys Leu Lys Phe Asp Lys Leu Tyr

ES 2 562 425 T3

355	360	365																				
Lys	Met	Leu	Thr	Glu	Ile	Tyr	Thr	Glu	Asp	Asn	Phe	Val	Lys	Phe	Phe							
	370					375					380											
Lys	Val	Leu	Asn	Arg	Lys	Thr	Tyr	Leu	Asn	Phe	Asp	Lys	Ala	Val	Phe							
385					390					395					400							
Lys	Ile	Asn	Ile	Val	Pro	Lys	Val	Asn	Tyr	Thr	Ile	Tyr	Asp	Gly	Phe							
				405					410					415								
Asn	Leu	Arg	Asn	Thr	Asn	Leu	Ala	Ala	Asn	Phe	Asn	Gly	Gln	Asn	Thr							
			420				425						430									
Glu	Ile	Asn	Asn	Met	Asn	Phe	Thr	Lys	Leu	Lys	Asn	Phe	Thr	Gly	Leu							
		435					440						445									
Phe	Glu	Phe	Tyr	Lys	Leu	Leu	Cys	Val	Asp	Gly	Ile	Ile	Thr	Ser	Lys							
	450					455					460											
Thr	Lys	Ser	Leu	Ile	Glu	Gly	Arg	Asn	Lys	Ala	Leu	Asn	Asp	Leu	Cys							
465					470					475					480							
Ile	Lys	Val	Asn	Asn	Trp	Asp	Leu	Phe	Phe	Ser	Pro	Ser	Glu	Asp	Asn							
				485					490					495								
Phe	Thr	Asn	Asp	Leu	Asn	Lys	Gly	Glu	Glu	Ile	Thr	Ser	Asp	Thr	Asn							
			500					505					510									
Ile	Glu	Ala	Ala	Glu	Glu	Asn	Ile	Ser	Leu	Asp	Leu	Ile	Gln	Gln	Tyr							
		515					520					525										
Tyr	Leu	Thr	Phe	Asn	Phe	Asp	Asn	Glu	Pro	Glu	Asn	Ile	Ser	Ile	Glu							
	530					535					540											
Asn	Leu	Ser	Ser	Asp	Ile	Ile	Gly	Gln	Leu	Glu	Leu	Met	Pro	Asn	Ile							
545					550					555					560							
Glu	Arg	Phe	Pro	Asn	Gly	Lys	Lys	Tyr	Glu	Leu	Asp	Lys	Tyr	Thr	Met							
				565					570					575								
Phe	His	Tyr	Leu	Arg	Ala	Gln	Glu	Phe	Glu	His	Gly	Lys	Ser	Arg	Ile							
			580					585					590									
Ala	Leu	Thr	Asn	Ser	Val	Asn	Glu	Ala	Leu	Leu	Asn	Pro	Ser	Arg	Val							
		595					600					605										
Tyr	Thr	Phe	Phe	Ser	Ser	Asp	Tyr	Val	Lys	Lys	Val	Asn	Lys	Ala	Thr							
	610					615						620										
Glu	Ala	Ala	Met	Phe	Leu	Gly	Trp	Val	Glu	Gln	Leu	Val	Tyr	Asp	Phe							
625					630					635					640							

ES 2 562 425 T3

Thr Asp Glu Thr Ser Glu Val Ser Thr Thr Asp Lys Ile Ala Asp Ile  
 645 650 655  
 Thr Ile Ile Ile Pro Tyr Ile Gly Pro Ala Leu Asn Ile Gly Asn Met  
 660 665 670  
 Leu Tyr Lys Asp Asp Phe Val Gly Ala Leu Ile Phe Ser Gly Ala Val  
 675 680 685  
 Ile Leu Leu Glu Phe Ile Pro Glu Ile Ala Ile Pro Val Leu Gly Thr  
 690 695 700  
 Phe Ala Leu Val Ser Tyr Ile Ala Asn Lys Val Leu Thr Val Gln Thr  
 705 710 715  
 Ile Asp Asn Ala Leu Ser Lys Arg Asn Glu Lys Trp Asp Glu Val Tyr  
 725 730 735  
 Lys Tyr Ile Val Thr Asn Trp Leu Ala Lys Val Asn Thr Gln Ile Asp  
 740 745 750  
 Leu Ile Arg Lys Lys Met Lys Glu Ala Leu Glu Asn Gln Ala Glu Ala  
 755 760 765  
 Thr Lys Ala Ile Ile Asn Tyr Gln Tyr Asn Gln Tyr Thr Glu Glu Glu  
 770 775 780  
 Lys Asn Asn Ile Asn Phe Asn Ile Asp Asp Leu Ser Ser Lys Leu Asn  
 785 790 795 800  
 Glu Ser Ile Asn Lys Ala Met Ile Asn Ile Asn Lys Phe Leu Asn Gln  
 805 810 815  
 Cys Ser Val Ser Tyr Leu Met Asn Ser Met Ile Pro Tyr Gly Val Lys  
 820 825 830  
 Arg Leu Glu Asp Phe Asp Ala Ser Leu Lys Asp Ala Leu Leu Lys Tyr  
 835 840 845  
 Ile Tyr Asp Asn Arg Gly Thr Leu Ile Gly Gln Val Asp Arg Leu Lys  
 850 855 860  
 Asp Lys Val Asn Asn Thr Leu Ser Thr Asp Ile Pro Phe Gln Leu Ser  
 865 870 875 880  
 Lys Tyr Val Asp Asn Gln Arg Leu Leu Ser Thr Cys Gly Gly Gly Gly  
 885 890 895  
 Ser Ile Glu Gly Arg Val Pro Leu Pro Ala Gly Gly Gly Thr Val Leu  
 900 905  
 Thr Lys Met Tyr Pro Cys Gly Asn His Trp Ala Val Gly His Leu Met  
 915 920 925  
 Gly

## REIVINDICACIONES

1.- Una proteína de fusión de polipéptido monocatenario, que comprende:

(a) una proteasa no citotóxica capaz de romper una proteína del aparato de fusión exocítica de una célula diana;

5 (b) un resto de transporte dirigido que es capaz de unirse a un sitio de unión sobre la célula diana, y dicho sitio diana es capaz de sufrir una endocitosis para ser incorporado en un endosoma dentro de la célula diana;

(c) un dominio de translocación que es capaz de translocar la proteasa desde el interior de un endosoma, a través de la membrana endosómica y hacia el citosol de la célula diana;

10 (d) un primer sitio de ruptura de proteasas, en cuyo sitio la proteína de fusión puede ser rota por una primera proteasa, en la que el primer sitio de ruptura de proteasas está localizado entre la proteasa no citotóxica y el dominio de translocación;

(e) un segundo sitio de ruptura de proteasas, en cuyo sitio la proteína de fusión puede ser rota por una segunda proteasa, en la que el segundo sitio de ruptura de proteasas está localizado entre el dominio de translocación y el resto de transporte dirigido; y

15 (f) un enlace covalente entre el resto de transporte dirigido y el dominio de translocación, en la que, después de la ruptura proteolítica en el segundo sitio de ruptura de proteasas, el resto de transporte dirigido permanece unido al dominio de translocación por dicho enlace covalente;

20 en la que, después de la ruptura en el primer y segundo sitio de ruptura de proteasas, el resto de transporte dirigido es capaz de interaccionar con el sitio de unión sobre la célula diana a través de una interacción entre un dominio N-terminal del resto de transporte dirigido y un dominio del sitio de unión, y simultáneamente a través de una interacción entre un dominio C-terminal del resto de transporte dirigido y un dominio del sitio de unión.

2.- La proteína de fusión de la reivindicación 1, en la que el dominio de translocación está localizado entre la proteasa no citotóxica y el resto de transporte dirigido; o

25 en la que el resto de transporte dirigido está localizado entre la proteasa no citotóxica y el dominio de translocación, y el primer sitio de ruptura de proteasas está localizado también entre la proteasa no citotóxica y el resto de transporte dirigido.

3.- La proteína de fusión de cualquiera de las reivindicaciones anteriores, en la que la proteasa no citotóxica está localizada en el N-terminal de la proteína; y/o

en la que el enlace covalente es un enlace disulfuro.

30 4.- La proteína de fusión de cualquiera de las reivindicaciones anteriores, en la que, entre el dominio de translocación y el resto de transporte dirigido, está localizado un polipéptido corto que proporciona una estructura secundaria del polipéptido, y en la que dicha estructura secundaria del polipéptido actúa para acercar parte del resto de transporte dirigido al dominio de translocación, provocando con ello que la formación del enlace covalente sea energéticamente más favorable.

35 5.- La proteína de fusión de cualquiera de las reivindicaciones anteriores, en la que el resto de transporte dirigido comprende un primer y un segundo dominio, y en la que el primer y segundo dominio están separados por un máximo de 10 restos aminoácidos, preferiblemente un máximo de 5 aminoácidos, y más preferiblemente un máximo de cero restos aminoácidos; preferiblemente

en la que el primer y segundo dominio del resto de transporte dirigido puede derivarse de ligandos a diferentes receptores.

40 6.- La proteína de fusión de cualquiera de las reivindicaciones anteriores, en la que el resto de transporte dirigido comprende o consiste en un péptido seleccionado del grupo que consiste en: un péptido de la hormona liberadora de gonadotropina (GnRH), un péptido opioide, un péptido de beta-endorfina, un péptido de bradiquinina, un péptido de BAM, un péptido de nociceptina, un péptido de dinorfina, un péptido de galanina, un péptido de encefalina, un péptido de la sustancia P, un péptido del factor liberador de corticotropina (CRF), un péptido liberador de gastrina (GRP), un péptido de neuromedina B, un péptido de bombesina, un péptido de gastrina, un péptido de CCK, un péptido de somatostatina (SST), un péptido de cortistatina (CST), un péptido de la hormona liberadora de la hormona del crecimiento (GHRH), un péptido de PAR, un péptido de la hormona paratiroidea (PTH), un péptido vasointestinal (VIP), un péptido de un agonista del beta<sub>2</sub> adrenergico, un péptido liberador de gastrina, un péptido relacionado con el gen de la calcitonina, un péptido de la hormona estimulante del tiroides (TSH), un péptido de insulina, un péptido del factor del crecimiento insulínico, un péptido de gonadorrelina, un péptido de la hormona liberadora de corticotrofina (CRH), un péptido de la hormona adrenocorticotrópica (ACTH), un péptido activador de la adenil ciclasa de la pituitaria (PACAP).

50

- 7.- La proteína de fusión de cualquiera de las reivindicaciones anteriores, en la que la proteasa no citotóxica y el primer sitio de ruptura de proteasas están separados por un máximo de 10 restos aminoácidos, preferiblemente un máximo de 5 restos aminoácidos, y más preferiblemente por cero restos aminoácidos; y/o
- 5 en la que el dominio de translocación y el primer sitio de ruptura de proteasas están separados por un máximo de 10 restos aminoácidos, preferiblemente un máximo de 5 restos aminoácidos, y más preferiblemente por cero restos aminoácidos; y/o
- en la que el dominio de translocación y el segundo sitio de ruptura de proteasas están separados por un máximo de 10 restos aminoácidos, preferiblemente un máximo de 5 restos aminoácidos, y más preferiblemente por cero restos aminoácidos; y/o
- 10 en la que el resto de transporte dirigido y el segundo sitio de ruptura de proteasas están separados por un máximo de 10 restos aminoácidos, preferiblemente un máximo de 5 restos aminoácidos, y más preferiblemente por cero restos aminoácidos.
- 8.- La proteína de fusión de cualquiera de las reivindicaciones anteriores, en la que la primera proteasa y la segunda proteasa son iguales.
- 15 9.- La proteína de fusión de cualquiera de las reivindicaciones anteriores, en la que la proteasa no citotóxica es una cadena L de una neurotoxina clostridial, o uno de sus fragmentos; y/o
- en la que el dominio de translocación es un dominio H<sub>N</sub> de una neurotoxina clostridial, o uno de sus fragmentos.
- 10.- Una secuencia de ácido nucleico que codifica una proteína de fusión de polipéptido según cualquiera de las reivindicaciones anteriores.
- 20 11.-. Un vector de ADN, que comprende un promotor, una secuencia de ácido nucleico según la reivindicación 10, y un terminador; en el que dicha secuencia de ácido nucleico está localizada cadena abajo del promotor; y en la que dicho terminador está localizado cadena abajo de la secuencia de ácido nucleico.
- 12.- La hebra de ADN complementario de la secuencia de ADN de la reivindicación 10.
- 25 13.- Un método para preparar una proteína de fusión de polipéptido monocatenario según una cualquiera de las reivindicaciones 1-9, que comprende expresar una secuencia de ácido nucleico según la reivindicación 10, o un vector de ADN según la reivindicación 11, en una célula hospedante.
- 14.- Un método para preparar un agente no citotóxico, que comprende:
- a. proporcionar una disolución que contiene una proteína de fusión de polipéptido monocatenario según una cualquiera de las reivindicaciones 1-9;
- 30 b. añadir a dicha disolución una primera proteasa capaz de romper el primer sitio de ruptura de proteasas, y una segunda proteasa capaz de romper el segundo sitio de ruptura de proteasas; y
- c. romper el primer sitio de ruptura de proteasas y el segundo sitio de ruptura de proteasas;
- formando con ello una proteína de fusión tricatenaria.
- 15.- Un polipéptido no citotóxico que puede obtenerse mediante el método de la reivindicación 14, en el que el polipéptido es un polipéptido tricatenario, y en el que:
- 35 a. la primera cadena comprende la proteasa no citotóxica capaz de romper una proteína del aparato de fusión exocítica de una célula diana;
- b. la segunda cadena comprende el dominio de translocación que es capaz de translocar la proteasa no citotóxica desde el interior de un endosoma, a través de la membrana endosómica y hacia el citosol de la célula diana;
- 40 c. la tercera cadena comprende el resto de transporte dirigido que es capaz de unirse a un sitio de unión sobre la célula diana, y dicho sitio diana es capaz de sufrir una endocitosis para ser incorporado en un endosoma dentro de la célula diana;
- d. la primera y la segunda cadena están unidos entre sí mediante un enlace disulfuro; y el segundo y el tercer dominio están unidos entre sí mediante un enlace covalente.
- 45 16.- Una proteína de fusión de polipéptido monocatenario según una cualquiera de las reivindicaciones 1-9, o un polipéptido no citotóxico según la reivindicación 15, para su uso para tratar, prevenir o mejorar un trastorno médico; preferiblemente

en la que el trastorno médico se selecciona de dolor, inflamación neurogénica (crónica), trastornos urogenitales-neurológicos, tales como vejiga hiperactiva, cáncer de próstata, cáncer de pulmón, cáncer de mama o cáncer colorrectal.

Figura 1

