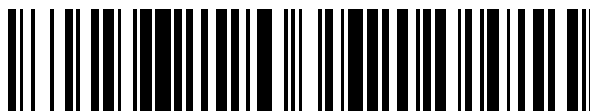


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 762 672**

51 Int. Cl.:

A61K 38/00 (2006.01)
A61K 38/46 (2006.01)
A61K 9/00 (2006.01)
C07K 14/705 (2006.01)
C12N 9/14 (2006.01)
C12N 9/16 (2006.01)
A61K 38/40 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **12.02.2015** **PCT/US2015/015662**
87 Fecha y número de publicación internacional: **27.08.2015** **WO15126729**
96 Fecha de presentación y número de la solicitud europea: **12.02.2015** **E 15708375 (9)**
97 Fecha y número de publicación de la concesión europea: **18.09.2019** **EP 3107562**

54 Título: **Proteínas de fusión de P97-IDS**

30 Prioridad:

19.02.2014 US 201461941896 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
25.05.2020

73 Titular/es:

BIOASIS TECHNOLOGIES INC. (100.0%)
130-10691 Shellbridge Way
Richmond, British Columbia V6X 2W8, CA

72 Inventor/es:

VITALIS, TIMOTHY Z. y
GABATHULER, REINHARD

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 762 672 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 de P97-IDS

Antecedentes

Campo técnico

- 5 La presente invención se refiere a proteínas de fusión entre p97 (melanotransferrina) e iduronato-2-sulfatasa (IDS, por sus siglas), y composiciones y métodos de uso relacionados con estas, por ejemplo, para facilitar el suministro de IDS a través de la barrera hematoencefálica (BBB, por sus siglas en inglés) y/o mejorar su penetración en el tejido en el CNS y/o tejidos periféricos y tratar y/o diagnosticar, de este modo, el síndrome de Hunter (Mucopolisacaridosis tipo II; MPS II, por sus siglas) y trastornos de almacenamiento lisosómico relacionados, incluidos aquellos que tienen un componente del sistema nervioso central (CNS, por sus siglas en inglés).

Descripción de la técnica relacionada

- 15 Las enfermedades del almacenamiento lisosómico (LSD, por sus siglas en inglés) surgen por la ausencia o actividad reducida de enzimas o proteínas específicas dentro de los lisosomas de una célula. Dentro de las células, el efecto de la actividad enzimática faltante se puede observar como una acumulación de un "material de almacenamiento" no degradado dentro del lisosoma intracelular. Este aumento provoca que los lisosomas se hinchen y fallen, lo que resulta en daño celular y tisular. Dado que las enfermedades de almacenamiento lisosómico típicamente tienen una etiología genética, muchos tejidos carecerán de la enzima en cuestión. Sin embargo, diferentes tejidos sufren la ausencia de la misma actividad enzimática de manera diferente. Cuán adversamente se verá afectado un tejido se determina, en cierta medida, por el grado en que dicho tejido genera el sustrato de la enzima faltante. A su vez, los tipos de tejidos más cargados por el almacenamiento, indican cómo se debe administrar el fármaco al paciente.

- 20 Se ha identificado y correlacionado una gran cantidad de enzimas de enfermedad de almacenamiento lisosómico con sus respectivas enfermedades. Después de que se ha identificado la enzima faltante o deficiente, el tratamiento se puede centrar en el problema de suministrar de manera eficaz la enzima de reemplazo a los tejidos afectados de un paciente. El síndrome de Hunter o Mucopolisacaridosis tipo II (MPS II) es un trastorno de almacenamiento lisosómico (LSD, por sus siglas en inglés) causado por una deficiencia de iduronato-2-sulfatasa (I2S o IDS). La I2S es una enzima lisosómica responsable del metabolismo de mucopolisacáridos. La deficiencia en la actividad de la enzima conduce finalmente a una variedad de patologías y a la muerte prematura. La terapia de reemplazo enzimático (ERT, por sus siglas en inglés) con I2S recombinante (Elaprase®) puede tratar síntomas periféricos, pero los pacientes sufren eventualmente demencia debido a que la enzima no puede cruzar la barrera hematoencefálica (BBB). Las formas desfosforiladas de proteínas de enfermedad de almacenamiento lisosómico (LSD), incluidas formas desfosforiladas de iduronato-2-sulfatasa (IDS o I2D) e iduronidasa (IDU), que tienen mayor capacidad de atravesar o penetrar la barrera hematoencefálica (BBB) con respecto a las formas fosforiladas de la proteína, y conjugados con p97 de estas se describen en la publicación internacional WO 2014/022515 A1.

- 35 La terapia de reemplazo enzimático (ERT) intravenosa puede ser beneficiosa para LSD tales como MPSII. Sin embargo, serían ventajosos medios para potenciar el suministro de la enzima terapéutica al lisosoma en dichas enfermedades en relación con un costo reducido y mayor eficacia terapéutica.

- 40 Como uno de los problemas, la barrera hematoencefálica (BBB) bloquea la transferencia libre de muchos agentes desde la sangre al cerebro. Por este motivo, no se espera que las LSD que presentan un aspecto neurológico significativo respondan a la ERT intravenosa. Para dichas enfermedades, serían muy deseables métodos para mejorar el suministro de la enzima a través de la BBB y hacia los lisosomas en las células afectadas.

Breve compendio

- 45 La presente invención se refiere a proteínas de fusión de p97 (melanotransferrina o MTf), que comprenden un polipéptido de iduronato-2-sulfatasa (IDS o I2S) fusionado al extremo N de un polipéptido de p97 y un enlazador peptídico opcional (L) entremedio, en donde el polipéptido de p97 consiste en la secuencia de aminoácidos DSSHAFTLDELRL. La presente invención se refiere, además, a proteínas de fusión de p97 (melanotransferrina), que comprenden un polipéptido de iduronato-2-sulfatasa (IDS) fusionado al extremo C de un polipéptido de p97 y un enlazador peptídico opcional (L) entremedio, en donde el polipéptido de p97 consiste en la secuencia de aminoácidos DSSHAFTLDELRL.

- 50 Ciertas proteínas de fusión comprenden el enlazador peptídico entremedio. En ciertas realizaciones, el enlazador peptídico se selecciona de uno o más de un enlazador rígido, un enlazador flexible y un enlazador enzimáticamente escindible. En ciertas realizaciones, el enlazador peptídico es un enlazador rígido, opcionalmente que comprende la secuencia (EAAAK)₁₋₃ (SEQ ID NOS:36-38), tal como EAAAKEAAAKEAAAK (SEQ ID NO:38). En algunas realizaciones, el enlazador peptídico es un enlazador flexible. En ciertas realizaciones, el enlazador peptídico es un enlazador enzimáticamente escindible.

En ciertas realizaciones, la proteína de fusión comprende una secuencia de péptido señal (SP, por sus siglas en inglés) en el extremo N, opcionalmente seleccionada de la Tabla 4. En algunas realizaciones, la proteína de fusión comprende la estructura: (a) SP-IDS-L-p97 o (b) SP-p97-L-IDS.

5 En realizaciones específicas, el SP comprende la secuencia MEWSWVFLFFLSVTTGVHS (SEQ ID NO:149) y la proteína de fusión de p97 comprende la estructura: (a) SP-p97-IDS o (b) SP-p97-L-IDS.

En ciertas realizaciones, el SP comprende la secuencia de SP de p97 humana MRGPSGALWLLALRTVLG (SEQ ID NO:39) y la proteína de fusión de p97 comprende la estructura: (a) SP-p97-IDS o (b) SP-p97-L-IDS.

10 En ciertas realizaciones, el SP comprende la secuencia de SP de IDS humana MPPPTGRGLLWLGLVSSVCVALG (SEQ ID NO:40) y la proteína de fusión de p97 comprende la estructura: (a) SP-IDS-p97 o (b) SP-IDS-L-p97.

15 En algunas realizaciones, la proteína de fusión comprende una etiqueta de purificación (TAG), opcionalmente seleccionada de la Tabla 5. En ciertas realizaciones, la proteína de fusión comprende la estructura: (a) SP-TAG-IDS-L-p97 o (b) SP-TAG-p97-L-IDS. En ciertas realizaciones, la etiqueta comprende una etiqueta de poli-histidina, opcionalmente una etiqueta de 10X poli-histidina. En algunas realizaciones, la etiqueta comprende una etiqueta FLAG DYKDDDDK (SEQ ID NO:122). En realizaciones específicas, la etiqueta comprende una etiqueta de poli-histidina, por ejemplo, una etiqueta de 10X poli-histidina y una etiqueta FLAG.

20 En ciertas realizaciones, la proteína de fusión comprende un sitio de proteasa (PS, por sus siglas en inglés), opcionalmente seleccionado de la Tabla 6. En realizaciones específicas, la proteína de fusión comprende la estructura: (a) SP-TAG-PS-IDS-L-p97 o (b) SP-TAG-PS-p97-enlazador-IDS. En realizaciones específicas, el sitio PS comprende el sitio de proteasa TEV ENLYFQG (SEQ ID NO:135).

En ciertas realizaciones, la proteína de fusión comprende la estructura (a) SP (MEWSWVFLFFLSVTTGVHS; SEQ ID NO:149)-HIS TAG-TEV PS-IDS-L rígido-p97 o (b) SP (MEWSWVFLFFLSVTTGVHS; SEQ ID NO: 149)-HIS TAG-TEV PS-p97-L rígido-IDS.

25 En realizaciones específicas, la proteína de fusión comprende la estructura (a) SP (MEWSWVFLFFLSVTTGVHS; SEQ ID NO: 149)-HIS TAG-TEV PS-IDS-(EAAAK)₃-p97 o (b) SP (MEWSWVFLFFLSVTTGVHS; SEQ ID NO: 149)-HIS TAG-TEV PS-p97-(EAAAK)₃-IDS.

En ciertas realizaciones, la proteína de fusión comprende la estructura (a) IDS SP-HIS TAG-TEV PS-IDS-L rígido-p97 o (b) p97 SP-HIS TAG-TEV PS-p97-L rígido-IDS.

30 En realizaciones específicas, la proteína de fusión comprende la estructura (a) IDS SP-10xHIS TAG-TEV PS-IDS-(EAAAK)₃-p97 (SEQ ID NO:29) o (b) p97 SP-10xHIS TAG-TEV PS-p97-(EAAAK)₃-IDS (SEQ ID NO:30).

35 En ciertas realizaciones, el polipéptido de IDS comprende, consiste o consiste esencialmente en (a) una secuencia de aminoácidos expuesta en SEQ ID NOs:31-35; (b) una secuencia de aminoácidos al menos 90 % idéntica a una secuencia expuesta en SEQ ID NOs:31-35; (c) o una secuencia de aminoácidos que difiere de SEQ ID NOs:31-35 por la adición, sustitución, inserción o eliminación de aproximadamente 1-50 aminoácidos. En algunas realizaciones, el polipéptido de IDS comprende, consiste o consiste esencialmente en la secuencia de aminoácidos expuesta en SEQ ID NO:32 o 33.

El polipéptido de p97 según la invención consiste en la secuencia de aminoácidos expuesta en SEQ ID NO:14 (MTfpep).

40 En ciertas realizaciones, la proteína de fusión comprende, consiste o consiste esencialmente en (a) una secuencia de aminoácidos expuesta en SEQ ID NOs: 140-142; (b) una secuencia de aminoácidos al menos 90 % idéntica a una secuencia expuesta en SEQ ID NO: 140-142; (c) o una secuencia de aminoácidos que difiere de SEQ ID NO: 140-142 por la adición, sustitución, inserción o eliminación de aproximadamente 1-50 aminoácidos. En realizaciones específicas, la proteína de fusión comprende, consiste o consiste esencialmente en una secuencia de aminoácidos expuesta en SEQ ID NO: 140-142.

45 También se incluyen polinucleótidos aislados que codifican la proteína de fusión de p97 descrita en la presente memoria. En algunas realizaciones, el polinucleótido aislado se optimiza por codón para su expresión en una célula hospedante. En ciertas realizaciones, la célula hospedante es una célula de mamífero, una célula de insecto, una célula de levadura o una célula bacteriana. El polinucleótido descrito en la presente memoria puede comprender una secuencia seleccionada de SEQ ID NO:143-147.

50 Algunas realizaciones incluyen células hospedantes recombinantes, que comprenden un polinucleótido aislado descrito en la presente memoria, donde el polinucleótido aislado está funcionalmente enlazado a uno o más elementos reguladores.

También se incluyen vectores que comprenden un polinucleótido aislado que codifica una proteína de fusión de p97 descrita en la presente memoria, que está funcionalmente enlazado a uno o más elementos reguladores.

También se incluyen células hospedantes recombinantes que comprenden un vector, polinucleótido aislado y/o proteína de fusión de p97 descrita en la presente memoria. En ciertas realizaciones, la célula hospedante es una célula de mamífero, una célula de insecto, una célula de levadura o una célula bacteriana. En realizaciones específicas, la célula de mamífero es una célula de ovario de hámster chino (CHO, por sus siglas en inglés), una célula HEK-293 o una célula de fibrosarcoma humano HT-1080.

Ciertas realizaciones incluyen composiciones farmacéuticas que comprenden un portador farmacéuticamente aceptable y una proteína de fusión de p97 según la invención, donde la composición farmacéutica es estéril y no pirógena.

También se incluyen métodos para el tratamiento de una enfermedad de almacenamiento lisosómico en un sujeto que lo necesita, que comprenden administrar al sujeto una proteína de fusión de p97 o composición farmacéutica según la invención. En ciertas realizaciones, la enfermedad de almacenamiento lisosómico es el síndrome de Hunter (MPS II). En ciertas realizaciones, la enfermedad de almacenamiento lisosómico tiene afectación del sistema nervioso central (CNS). En ciertas realizaciones, el sujeto está en riesgo de desarrollar afectación del CNS de la enfermedad de almacenamiento lisosómico. En ciertas realizaciones, el sujeto es un humano de sexo masculino. En ciertas realizaciones, la proteína de fusión de p97 o composición farmacéutica se administra mediante infusión intravenosa (IV) o inyección intraperitoneal (IP).

Breve descripción de los dibujos

Las Figuras 1A y 1B ilustran la estructura general de proteínas de fusión ilustrativas que tienen un péptido señal (SP), etiqueta de purificación o afinidad (TAG), sitio de proteasa (PS) para la eliminación del SP y la TAG, polipéptido de p97 (melanotransferrina), un enlazador (L) y un polipéptido de iduronato-2-sulfatasa (IDS).

La Figura 2 muestra la evaluación de la actividad enzimática de las proteínas de fusión I2S-MTf y MTf-I2S según se mide por su capacidad de hidrolizar el sustrato 4-Nitrocatecol Sulfato (PNCS) con respecto a IDS humana recombinante y un testigo negativo (fusión TZM-MTf). Se usó 1 ug de cada muestra en el ensayo de actividad enzimática y los datos presentados están normalizados con respecto a rhIDS.

La Figura 3 muestra la evaluación de la actividad enzimática de las proteínas de fusión MTfpep-I2S y I2S-MTfpep (con propéptido de I2S) según se mide por su capacidad de hidrolizar el sustrato PNCS con respecto a la fusión I2S-MTf y un testigo negativo (fusión TZM-MTf). Se usó 1 ug de cada muestra en el ensayo de actividad enzimática y los datos presentados están normalizados con respecto al blanco de sustrato.

La Figura 4 muestra una comparación de la actividad enzimática de las proteínas de fusión I2S-MTfpep (con propéptido de I2S) e I2S-MTfpep (sin propéptido de I2S) según se mide por su capacidad de hidrolizar el sustrato PNCS. Se usó 1 ug de cada muestra en el ensayo de actividad enzimática y los datos presentados están normalizados con respecto al blanco de sustrato.

La Figura 5 muestra la cuantificación de la distribución relativa de las proteínas de fusión MTfpep-I2S (con propéptido) e I2S-MTf entre capilares (C) y el parénquima (P) en el cerebro, con respecto a la señal total (T). La generación de imágenes por microscopía confocal cuantitativa muestra que ambas proteínas de fusión MTfpep-I2S y I2S-MTf se asociaron intensamente con tejidos parenquimatosos del CNS.

Descripción detallada

La puesta en práctica de la presente invención empleará, a menos que se indique específicamente lo contrario, métodos convencionales de biología molecular y técnicas de ADN recombinante que están dentro de la experiencia en la técnica, muchas de las cuales se describen a continuación con fines ilustrativos. Dichas técnicas se explican de manera completa en la literatura. Ver, *p. ej.*, Sambrook, et al., *Molecular Cloning: A Laboratory Manual* (3a edición, 2000); *DNA Cloning: A Practical Approach*, tomos I y II (D. Glover, ed.); *Oligonucleotide Synthesis* (N. Gait, ed., 1984); *Oligonucleotide Synthesis: Methods and Applications* (P. Herdewijn, ed., 2004); *Nucleic Acid Hybridization* (B. Hames & S. Higgins, eds., 1985); *Nucleic Acid Hybridization: Modern Applications* (Buzdin and Lukyanov, eds., 2009); *Transcription and Translation* (B. Hames & S. Higgins, eds., 1984); *Animal Cell Culture* (R. Freshney, ed., 1986); Freshney, R.I. (2005) *Culture of Animal Cells, a Manual of Basic Technique*, 5a ed. Hoboken NJ, John Wiley & Sons; B. Perbal, *A Practical Guide to Molecular Cloning* (3a edición 2010); Farrell, R., *RNA Methodologies: A Laboratory Guide for Isolation and Characterization* (3a edición, 2005).

Definiciones

A menos que se definan de cualquier otra manera, todos los términos técnicos y científicos usados en la presente memoria tienen el mismo significado que el entendido comúnmente por los expertos en la técnica a los cuales corresponde la invención. Aunque se pueden usar cualesquiera métodos y materiales similares o equivalentes a los descritos en la presente memoria en la puesta en práctica o evaluación de la presente invención, se describen ciertos métodos y materiales ilustrativos en la presente memoria. A los efectos de la presente invención, se definen los siguientes términos a continuación.

Los artículos "un/uno" y "una" se usan en la presente memoria para hacer referencia a uno o más de uno (*es decir*, a al menos uno) del objeto gramatical del artículo. A modo de ejemplo, "un elemento" significa un elemento o más de un elemento.

- 5 Se entiende por "aproximadamente" una cantidad, nivel, valor, número, frecuencia, porcentaje, dimensión, tamaño, monto, peso o longitud que varía tanto como 30, 25, 20, 15, 10, 9, 8, 7, 6, 5, 4, 3, 2 o 1 % con respecto a una cantidad, nivel, valor, número, frecuencia, porcentaje, dimensión, tamaño, monto, peso o longitud de referencia.

Según se usa en la presente memoria, se pretende que el término "aminoácido" signifique aminoácidos de origen natural y de origen no natural, así como análogos y miméticos de aminoácidos. Los aminoácidos de origen natural incluyen los 20 (L)-aminoácidos usados durante la biosíntesis proteica, así como otros tales como 4-hidroxiprolina, hidroxilisina, desmosina, isodesmosina, homocisteína, citrulina y ornitina, por ejemplo. Los aminoácidos de origen no natural incluyen, por ejemplo, (D)-aminoácidos, norleucina, norvalina, p-fluorofenilalanina, etionina y similares, que conoce un experto en la técnica. Los análogos de aminoácidos incluyen formas modificadas de aminoácidos de origen natural y de origen no natural. Dichas modificaciones pueden incluir, por ejemplo, sustitución o reemplazo de grupos y restos químicos en el aminoácido o la derivación del aminoácido. Los miméticos de aminoácidos incluyen, por ejemplo, estructuras orgánicas que exhiben propiedades funcionalmente similares, tales como características de carga y espaciado de carga del aminoácido de referencia. Por ejemplo, una estructura orgánica que imita Arginina (Arg o R) tendría un resto de carga positiva ubicado en un espacio molecular similar y que tendría el mismo grado de movilidad que el grupo e-amino en la cadena lateral del aminoácido Arg de origen natural. Los miméticos también incluyen estructuras restringidas para mantener un espaciado óptimo e interacciones de carga del aminoácido o de los grupos funcionales del aminoácido. Los expertos en la técnica saben o pueden determinar qué estructuras constituyen análogos de aminoácidos o miméticos de aminoácidos funcionalmente equivalentes.

A lo largo de la presente memoria descriptiva, a menos que el contexto lo requiera de cualquier otra manera, se entenderá que las palabras "comprenden", "comprende" y "que comprende/n" implican la inclusión de una etapa o elemento o grupo de etapas o elementos indicados, pero no la exclusión de ninguna otra etapa o elemento o grupo de etapas o elementos. Se entiende por "consiste en" que incluye, y se limita a, todo lo que sigue a la frase "consiste en". Por lo tanto, la frase "consiste en" indica que los elementos indicados son necesarios u obligatorios, y que ningún otro elemento puede estar presente. Se entiende por "consiste esencialmente en" que incluye cualesquiera elementos indicados después de la frase y se limita a otros elementos que no interfieren en ni contribuyen a la actividad o acción especificada en la descripción de los elementos indicados. Por lo tanto, la frase "consiste esencialmente en" indica que los elementos indicados son necesarios u obligatorios, pero que otros elementos son opcionales y pueden o no estar presentes, dependiendo de si afectan o no materialmente la actividad o acción de los elementos indicados.

Se pretende que el término "conjugado" haga referencia a la entidad formada como resultado de un acoplamiento o enlace covalente o no covalente de un agente u otra molécula, *p. ej.*, una molécula biológicamente activa, con un polipéptido de p97 o secuencia de p97. Un ejemplo de un polipéptido conjugado es una "proteína de fusión" o "polipéptido de fusión", es decir, un polipéptido que se crea a través de la unión de dos o más secuencias codificantes, que originalmente codificaban polipéptidos separados; la traducción de las secuencias codificantes unidas genera un polipéptido de fusión simple, típicamente con propiedades funcionales derivadas de cada uno de los polipéptidos separados.

40 Según se usan en la presente memoria, los términos "función" y "funcional" y similares se refieren a una función biológica, enzimática o terapéutica.

"Homología" se refiere al número porcentual de aminoácidos que son idénticos o constituyen sustituciones conservadoras. La homología se puede determinar mediante el uso de programas de comparación de secuencias tales como GAP (Deveraux et al., Nucleic Acids Research. 12, 387-395, 1984). De este modo, se podrían comparar secuencias de una longitud similar o sustancialmente diferente con las mencionadas en la presente memoria mediante la inserción de espacios en la alineación, tales como espacios que se determinan, por ejemplo, mediante el algoritmo de comparación usado por GAP.

Se entiende por "aislado" material que está sustancialmente o esencialmente libre de componentes que normalmente lo acompañan en su estado natural. Por ejemplo, un "péptido aislado" o un "polipéptido aislado" y similares, según se usa en la presente memoria, incluye el aislamiento *in vitro* de una molécula de péptido o polipéptido de su entorno natural celular, y de la asociación con otros componentes de la célula; *es decir*, no está asociado significativamente con sustancias *in vivo*.

El término "enlace", "enlazador", "resto enlazador" o "L" se usa en la presente memoria para hacer referencia a un enlazador que se puede usar para separar un polipéptido de p97 de un agente de interés, o para separar un primer agente de otro agente, por ejemplo, donde dos o más agentes se enlazan para formar un conjugado o proteína de fusión de p97. El enlazador puede ser fisiológicamente estable o puede incluir un enlazador liberable tal como un enlazador enzimáticamente degradable (*p. ej.*, enlazadores proteolíticamente escindibles). En ciertos aspectos, el enlazador puede ser un enlazador peptídico, por ejemplo, como parte de una proteína de fusión de p97. En ciertos

aspectos, el enlazador puede ser un enlazador no peptídico o un enlazador no proteínico. En ciertos aspectos, el enlazador puede ser una partícula, tal como una nanopartícula.

Los términos "modular" y "alterar" incluyen "aumentar", "potenciar" o "estimular", así como "disminuir" o "reducir" típicamente en una cantidad o grado estadísticamente significativo o fisiológicamente significativo con respecto a un testigo. Una cantidad "aumentada", "estimulada" o "potenciada" es típicamente una cantidad "estadísticamente significativa", y puede incluir un aumento que es 1,1, 1,2, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30 o más veces mayor (*p. ej.*, 500, 1000 veces) (incluidos todos los números enteros y puntos decimales entre y por encima de 1, *p. ej.*, 1,5, 1,6, 1,7, 1,8, etc.) que la cantidad producida sin la composición (*p. ej.*, la ausencia de una proteína de fusión de la invención) o una composición, muestra o sujeto de prueba testigo. Una cantidad "disminuida" o "reducida" es típicamente una cantidad "estadísticamente significativa", y puede incluir una disminución de 1 %, 2 %, 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 %, 10 %, 11 %, 12 %, 13 %, 14 %, 15 %, 16 %, 17 %, 18 %, 19 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 % o 100 % en la cantidad producida sin composición o mediante una composición testigo, incluidos todos los números enteros entremedio. Como un ejemplo no limitante, un testigo podría comparar la actividad, tal como la actividad enzimática, la cantidad o la velocidad de transporte/suministro a través de la barrera hematoencefálica, la velocidad y/o los niveles de distribución al tejido del sistema nervioso central y/o la $C_{m\acute{a}x}$ para el plasma, tejidos del sistema nervioso central o cualesquiera otros tejidos que no pertenecen al sistema nervioso central sistémicos o periféricos, de una proteína de fusión de p97 con respecto al agente/proteína solo. Otros ejemplos de comparaciones y cantidades "estadísticamente significativas" se describen en la presente memoria.

En ciertas realizaciones, la "pureza" de un agente dado (*p. ej.*, un conjugado de p97 tal como una proteína de fusión) se puede definir específicamente en una composición. Por ejemplo, ciertas composiciones pueden comprender un agente que es al menos 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o 100 % puro, incluidos todos los decimales entremedio, según se mide, por ejemplo, y de ningún modo como limitación, mediante cromatografía de líquidos de alta presión (HPLC, por sus siglas en inglés), una forma conocida de cromatografía en columna usada frecuentemente en bioquímica y química analítica para separar, identificar y cuantificar compuestos.

Los términos "polipéptido" y "proteína" se usan de manera intercambiable en la presente memoria para hacer referencia a un polímero de residuos aminoacídicos y a variantes y análogos sintéticos de los mismos. Por lo tanto, estos términos se aplican a polímeros aminoacídicos en lo que uno o más residuos aminoacídicos son aminoácidos de origen no natural sintéticos, tales como un análogo químico de un aminoácido de origen natural correspondiente, así como a polímeros aminoacídicos de origen natural. Los polipéptidos descritos en la presente memoria no se limitan a una longitud específica del producto; por lo tanto, se incluyen péptidos, oligopéptidos y proteínas dentro de la definición de polipéptido, y dichos términos se pueden usar de manera intercambiable en la presente memoria a menos que se indique específicamente de cualquier otra manera. Los polipéptidos descritos en la presente memoria también pueden comprender modificaciones postexpresión, tales como glicosilaciones, acetilaciones, fosforilaciones y similares, así como otras modificaciones conocidas en la técnica, tanto de origen natural como de origen no natural. Un polipéptido puede ser una proteína entera, o una subsecuencia, fragmento, variante o derivado de esta.

Una unión "fisiológicamente escindible" o "hidrolizable" o "degradable" es una unión que hace reacción con agua (es decir, se hidroliza) en condiciones fisiológicas. La tendencia de una unión a hidrolizarse en agua dependerá no solo del tipo general de enlace que conecta dos átomos centrales, sino también de los sustituyentes acoplados a estos átomos centrales. Los enlaces hidrolíticamente inestables o débiles adecuados incluyen, pero no se limitan a: éster de carboxilato, éster de fosfato, anhídrido, acetal, cetel, éter de aciloxialquilo, imina, ortoéster, tio éster, tiol éster, carbonato e hidrazona, péptidos y oligonucleótidos.

Un "enlazador liberable" incluye, pero no se limita a, un enlazador fisiológicamente escindible y un enlazador enzimáticamente degradable. Por lo tanto, un "enlazador liberable" es un enlazador que puede sufrir hidrólisis espontánea o escisión mediante algún otro mecanismo (*p. ej.*, catalizado por enzima, catalizado por ácido, catalizado por base, entre otros) en condiciones fisiológicas. Por ejemplo, un "enlazador liberable" puede implicar una reacción de eliminación que tiene una abstracción de base de un protón (*p. ej.*, un átomo de hidrógeno ionizable, H α), como la fuerza impulsora. A efectos de la presente memoria, un "enlazador liberable" es sinónimo de un "enlazador degradable". Un "enlace enzimáticamente degradable" incluye un enlace, *p. ej.*, una secuencia de aminoácidos que se somete a degradación mediante una o más enzimas, *p. ej.*, peptidasas o proteasas. En realizaciones específicas, un enlazador liberable tiene una semivida a pH 7,4, 25 °C, *p. ej.*, un pH fisiológico, temperatura corporal humana (*p. ej.*, in vivo), de aproximadamente 30 minutos, aproximadamente 1 hora, aproximadamente 2 hora, aproximadamente 3 horas, aproximadamente 4 horas, aproximadamente 5 horas, aproximadamente 6 horas, aproximadamente 12 horas, aproximadamente 18 horas, aproximadamente 24 horas, aproximadamente 36 horas, aproximadamente 48 horas, aproximadamente 72 horas o aproximadamente 96 horas o menos.

El término "secuencia de referencia" se refiere generalmente a una secuencia codificante de ácido nucleico o secuencia de aminoácidos, con la que se compara otra secuencia. Todas las secuencias polipeptídicas y polinucleotídicas descritas en la presente memoria se incluyen como secuencias de referencia, incluidas las descritas por nombre y las descritas en la Lista de secuencias.

Los términos "identidad de secuencia" o, por ejemplo, que comprenden una "secuencia 50 % idéntica a", según se usan en la presente memoria, hacen referencia a la medida en que dichas secuencias son idénticas en función de una ventana de comparación nucleótido por nucleótido o aminoácido por aminoácido. Por lo tanto, un "porcentaje de identidad de secuencia" se puede calcular al comparar dos secuencias alineadas de manera óptima en una ventana de comparación y determinar la cantidad de posiciones en las que la base de ácido nucleico idéntica (*p. ej.*, A, T, C, G, I) o el residuo aminoácido idéntico (*p. ej.*, Ala, Pro, Ser, Thr, Gly, Val, Leu, Ile, Phe, Tyr, Trp, Lys, Arg, His, Asp, Glu, Asn, Gln, Cys y Met) aparece en ambas secuencias para proporcionar la cantidad de posiciones coincidentes, dividir la cantidad de posiciones coincidentes entre la cantidad total de posiciones en la ventana de comparación (*es decir*, el tamaño de la ventana) y multiplicar el resultado por 100 para proporcionar el porcentaje de identidad de secuencia. Se incluyen nucleótidos y polipéptidos que tienen al menos aproximadamente 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 97 %, 98 %, 99 % o 100 % de identidad de secuencia con respecto a cualquiera de las secuencias de referencia descritas en la presente memoria (*ver, p. ej.*, la Lista de secuencias), típicamente donde la variante polipeptídica mantiene al menos una actividad biológica del polipéptido de referencia.

Los términos usados para describir relaciones secuenciales entre dos o más polinucleótidos o polipéptidos incluyen "secuencia de referencia", "ventana de comparación", "identidad de secuencia", "porcentaje de identidad de secuencia" e "identidad sustancial". Una "secuencia de referencia" tiene al menos 12, pero frecuentemente 15 a 18 y a menudo al menos 25 unidades monoméricas, que incluyen nucleótidos y residuos aminoácidos, de longitud. Dado que dos polinucleótidos pueden comprender cada uno (1) una secuencia (*es decir*, solo una porción de la secuencia polinucleotídica completa) que es similar entre los dos polinucleótidos, y (2) una secuencia que es divergente entre los dos polinucleótidos, las comparaciones de secuencia entre dos (o más) polinucleótidos típicamente se llevan a cabo al comparar secuencias de los dos polinucleótidos en una "ventana de comparación" para identificar y comparar regiones locales de similitud secuencial. Una "ventana de comparación" hace referencia a un segmento conceptual de al menos 6 posiciones contiguas, habitualmente, aproximadamente 50 a aproximadamente 100, más habitualmente, aproximadamente 100 a aproximadamente 150 en el que una secuencia se compara con una secuencia de referencia con la misma cantidad de posiciones contiguas después de que las dos secuencias se alinean de manera óptima. La ventana de comparación puede comprender adiciones o eliminaciones (*es decir*, espacios) de aproximadamente 20 % o menos en comparación con la secuencia de referencia (que no comprende adiciones ni eliminaciones) para la alineación óptima de las dos secuencias. La alineación óptima de secuencias para alinear una ventana de comparación se puede llevar a cabo mediante implementaciones informáticas de algoritmos (GAP, BESTFIT, FASTA y TFASTA en el Wisconsin Genetics Software Package Release 7.0, Genetics Computer Group, 575 Science Drive Madison, WI, EE. UU.) o por la inspección y la mejor alineación (*es decir*, que da como resultado el porcentaje de homología más alto en la ventana de comparación) generadas mediante cualesquiera de los diversos métodos seleccionados. También se puede consultar la familia de programas BLAST como, por ejemplo, la descrita por Altschul et al., Nucl. Acids Res. 25:3389, 1997. Se puede encontrar una descripción detallada del análisis de secuencias en la Unidad 19.3 de Ausubel et al., "Current Protocols in Molecular Biology", John Wiley & Sons Inc, 1994-1998, Capítulo 15.

Se entiende por "estadísticamente significativo" que es improbable que el resultado se produjera por casualidad. La significación estadística se puede determinar mediante cualquier método conocido en la técnica. Las mediciones de la significación usadas comúnmente incluyen el valor de la *p*, que es la frecuencia o probabilidad de que se produjera el evento observado, si la hipótesis nula fuera verdadera. Si el valor de la *p* obtenido es menor que el nivel de significación, entonces la hipótesis nula se rechaza. En casos simples, el nivel de significación se define por un valor de la *p* de 0,05 o menor.

El término "solubilidad" se refiere a la propiedad de una proteína de disolverse en un disolvente líquido y formar una disolución homogénea. La solubilidad típicamente se expresa como una concentración, ya sea en masa de soluto por unidad volumen de disolvente (g de soluto por kg de disolvente, g por dL (100 mL), mg/ml, etc.), molaridad, molalidad, fracción molar u otras descripciones de la concentración similares. La cantidad de soluto en equilibrio máximo que se puede disolver por la cantidad de disolvente es la solubilidad de dicho soluto en dicho disolvente en las condiciones especificadas, incluidas temperatura, presión, pH y la naturaleza del disolvente. En ciertas realizaciones, la solubilidad se mide a pH fisiológico, u otros pH, por ejemplo, a pH 5,0, pH 6,0, pH 7,0 o pH 7,4. En ciertas realizaciones, la solubilidad se mide en agua o un tampón fisiológico tal como PBS o NaCl (con o sin NaP). En realizaciones específicas, la solubilidad se mide a pH relativamente más bajo (*p. ej.*, pH 6,0) y sal relativamente más alta (*p. ej.*, 500mM NaCl y 10mM NaP). En ciertas realizaciones, la solubilidad se mide en un fluido biológico (disolvente) tal como sangre o suero. En ciertas realizaciones, la temperatura puede ser aproximadamente temperatura ambiente (*p. ej.*, aproximadamente 20, 21, 22, 23, 24, 25 °C) o aproximadamente temperatura corporal (~37 °C). En ciertas realizaciones, una proteína de fusión de polipéptido p97 tiene una solubilidad de al menos aproximadamente 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25 o 30 mg/ml a temperatura ambiente o a aproximadamente 37 °C.

Un "sujeto", según se usa en la presente memoria, incluye cualquier animal que exhibe un síntoma, o está en riesgo de exhibir un síntoma, que se puede tratar o diagnosticar con una proteína de fusión de p97 de la invención. Los sujetos adecuados (pacientes) incluyen animales de laboratorio (tales como ratón, rata, conejo o cobayo), animales de granja y animales domésticos o mascotas (tales como un gato o perro). Se incluyen primates no humanos y, preferiblemente, pacientes humanos.

"Sustancialmente" o "esencialmente" significa casi totalmente o completamente, por ejemplo, 95 %, 96 %, 97 %, 98 %, 99 % o más de alguna cantidad dada.

"Sustancialmente libre" se refiere a la ausencia casi completa o completa de una cantidad dada, por ejemplo, menos de aproximadamente 10 %, 5 %, 4 %, 3 %, 2 %, 1 %, 0,5 % o menos de alguna cantidad dada. Por ejemplo, ciertas composiciones pueden estar "sustancialmente libres" de proteínas celulares, membranas, ácidos nucleicos, endotoxinas u otros contaminantes.

"Tratamiento" o "tratar", según se usa en la presente memoria, incluye cualquier efecto deseable en los síntomas o la patología de una enfermedad o afección, y puede incluir incluso cambios o mejoras mínimos en uno o más marcadores medibles de la enfermedad o afección que se está tratando. "Tratamiento" o "tratar" no indica necesariamente la erradicación o cura completa de la enfermedad o afección, o síntomas asociados a estas. El sujeto que recibe este tratamiento es un cualquier sujeto que lo necesita. Los marcadores ilustrativos de mejora clínica serán evidentes para las personas con experiencia en la técnica.

El término "natural" se refiere a un gen o producto génico que tiene las características de dicho gen o producto génico cuando se asila de una fuente de origen natural. Un gen o producto génico natural (*p. ej.*, un polipéptido) es el que se observa más frecuentemente en una población y, por lo tanto, se designa arbitrariamente como la forma "normal" o "natural" del gen.

Proteínas de fusión

Realizaciones de la presente invención se refieren generalmente a proteínas de fusión que comprenden una secuencia polipeptídica de p97 (melanotransferrina; MTf) humana que consiste en la secuencia de aminoácidos DSSHAFTLDELRL y una secuencia polipeptídica de iduronato-2-sulfatasa (IDS o I2S), polinucleótidos que codifican las proteínas de fusión, células hospedantes y métodos para producir proteínas de fusión, y composiciones y métodos de uso relacionados con estas. Se describen proteínas de fusión ilustrativas (*p. ej.*, Tabla 1), secuencias polipeptídicas de p97 (*p. ej.*, Tabla 2) y secuencias polipeptídicas de IDS (*p. ej.*, Tabla 3) en la presente memoria. Los términos "p97" y "MTf" se usan de manera intercambiable en la presente memoria, así como los términos "IDS" e "I2S".

También se describen métodos y componentes ilustrativos para acoplar una secuencia polipeptídica de p97 a una secuencia de IDS. En ciertas realizaciones, la proteína de fusión de p97 comprende una o más secuencias de péptido señal (SP), etiquetas de purificación (TAG), sitios de escisión por proteasa (PS) y/o enlazadores peptídicos (L), incluida cualquier combinación de los anteriores, cuyos ejemplos se proporcionan en la presente memoria. Las variantes y fragmentos de cualesquiera de los anteriores también se describen en la presente memoria.

En ciertas realizaciones, la proteína de fusión de p97 comprende, consiste o consiste esencialmente en al menos una de las configuraciones ilustradas a continuación (extremo N > extremo C):

- IDS-p97
- p97-IDS
- IDS-L-p97
- p97-L-IDS
- SP-IDS-p97
- SP-p97-IDS
- SP-IDS-L-p97
- SP-P97-L-IDS
- SP-PS-IDS-p97
- SP-PS-P97-IDS
- SP-PS-IDS-L-p97
- SP-PS-p97-L-IDS
- SP-TAG-PS-IDS-p97
- SP-TAG-PS-p97-IDS
- SP-TAG-PS-IDS-L-p97

- SP-TAG-PS-p97-L-IDS
- TAG-IDS-p97
- TAG-p97-IDS
- TAG-IDS-L-p97
- 5 ▪ TAG-p97-L-IDS
- TAG-PS-IDS-p97
- TAG-PS-p97-IDS
- TAG-PS-IDS-L-p97
- TAG-PS-p97-L-IDS
- 10 ▪ IDS SP-HIS TAG-TEV PS-IDS-L rígido-p97
- IDS SP-HIS TAG-TEV PS-IDS-(EAAAK)₃-p97
- p97 SP-HIS TAG-TEV PS-p97-L rígido-IDS
- p97 SP-HIS TAG-TEV PS-p97-(EAAAK)₃-IDS

15 Las proteínas de fusión de estas y configuraciones relacionadas se pueden construir mediante el uso de cualesquiera de las secuencias de IDS, p97, L, SP, TAG o PS descritas en la presente memoria, incluidas variantes funcionales o activas y fragmentos de estas.

Los ejemplos específicos de proteínas de fusión de p97 se ilustran en la Tabla 1 a continuación.

Tabla 1: Proteínas de fusión de p97 ilustrativas		
Descripción	Secuencia	SEQ ID NO:
IDS SP-10xHIS TAG-TEV PS-IDS-L rígido-p97	MPPPRRTGRGLLWLGLVLSSVCVALGHHHHHHHHHENLYFQSETQANST TDALNVLLIIIVDDLRLPSLGCGDKLVRSPNIDQLASHSLLFQNAFAQQA VCAPSRVSFLTGRRPDTRLYDFNSYWRVHAGNFSTIPQYFKENGIVTM SVGKVFHHPGISSNHTDDSPYSWSFPPYHPSSEKYENTKTCRGPDELHA NLLCPVDVLDVPEGTLDPDKQSTEQAIQLLEKMKTSASPPFLAVGYHKPH IPFRYPKEFQKLYPLENITLAPDPEVPDGLPPVAYNPWMDIRQREDVQA LNISVPYGPPIPVDFQRKIRQSYFASVSYLDTQVGRLLSALDDLQLANST IIAFTSDHGWALGEHGEWAKYSNFDVATHVPLIFYVPGRTASLPEAGEK LFPYLDPPDSASQLMEPGRQSMDELVELVSLFPTLAGLAGIQVPPRCVP SFHVELCREGKNLLKHFRFRDLEEDPYLPGNPRELIAYSQYPRPSDIPQ WNSDKPSLKDIIKIMGYSIRTIDYRYTVWVGFPNDEFNANFSDIHAGELY FVDSDDLQDHNMYNDSQGGDLFQLLMPEAAAKEAAAKEAAAKGMEVRWC ATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIAAQEADAI TLDGGAIYEAGKEHGLKPVVGEVYDQEVGTSYYAVAVVRRSSHVTIDTL KGVKSCHTGINRTVGWNVVPGYLVESGRLSVMGCDVLKAVSDYFGGSCV PGAGETSYSESLCRLCRGDSSGEGVCDKSPLEYYDYSGAFRCIAEGAG DVAFFVKHSTVLENTDGKTLPSWGQALLSQDFELLCRDGSRADVTEWRQC HLARVPAHAVVVRADTDGGLIFRLNNEGQRLFSHEGSSFQMFSSSEAYGQ KDLLFKDSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPNRLPPYLRLWC VLSTPEIQKCGDMAVAFRRQRLKPEIQCVSAKSPQHCMERIQAEQVDAV TLSGEDIYTAGKTYGLVPAAGEHYAPEDSSNSYYVAVVRRDSSHAFTL DELRGKRSCHAGFGSPAGWDVPVGALIQRGFIRPKDCDVLTAVERSEFNA SCVPVNNPKNYPSLLCALCVGDEQGRNKC VGNSQERYYYGRGAFRCLVE NAGDVAFVRHTTVFDNTNGHNSEPWAAELRSEDYELLCPNGARAEVSQF AACNLAQIPPHAVMVRPDTNIFTVYGLLDKAQDLFGDDHNKNGFKMFDS SNYHGQDLLFKDATVRAVPVGEKTTYRGWLGLDYYAALEGMSQQCS	29
P97 SP-10xHIS TAG-TEV PS-p97-L		30

Tabla 1: Proteínas de fusión de p97 ilustrativas		
Descripción	Secuencia	SEQ ID NO:
rigido-IDS	<p>MRGPSGALWLLALRITVLGHHHHHHHHHHENLYFQGMEVWRCATSDPEQ HKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIAAQEADAITLDGGAI YEAGKEHGLKPVVGEVYDQEVGTSYYAVAVVRRSSSHVTIDTLKGVKSCH TGINRTVGWNVPGYLVESEGRLSVMGCDVLKAVSDYFGGSCVPGAGETS YSESLCRLCRGDSSGEGVCDKSPLEYYDYSGAFRCLAEGAGDVAFVKH STVLENTDGKTLPSWGQALLSQDFELLCRDGSRADVTEWRQCHLARVPA HAVVVRADTDGGLIFRLLNEGQRLFSHEGSSSQMFSSSEAYGQKDLLFKD STSELVPIATQTYEAWLGHEYLHAMKGLLCDPNRLPPYLRWCVLSTPEI QKCGDMAVAFRRQRLKPEIQCVSAKSPQHCHMERIQAEQVDAVTLSGEDI YTAGKTYGLVPAAGEHYAPEDSSNSYYVAVVRRDSSSHAFTLDELGRKR SCHAGFGSPAGWDVPGALIQRGFIRPKDCDVLTAVERSEFNASCVPVNN PKNYPSSSLCALCVGDEQGRNKCVCNSQERYYYGYRGAFRCLVENAGDVAF VRHTTVFDNTNGHNSEPWAAELRSEDYELLCPNGARAEVSQFAACNLAQ IPPHAVMVRPDTNIFTVYGLLDKAQDLFGDDHNKNGFKMFDSSSNYHGQD LLLFKDATVRAVPVGEKTTYRGWLGLDYVAALEGMSSQQCSEAAAKEAAK KEAAAKSETQANSTTDALNVLIIIVDDLRLPSLGCYGDKLVRSPNIDQLA SHSLLFQNAFAQQAVCAPSRVSFLTGRRPDTTRLYDFNSYWRVHAGNFS TIPQYFKENGYVTMSVGKVFHPGISSNHTDDSPYSWSFPPYHPSSEKYE NTKTCRGPDGELHANLLCPVDVLDVPEGTLPDKQSTEQAIIQLEKMKTS ASPFFLAVGYHKPHIPFRYPKEFKLYPLENITLAPDPEVDPGLPPVAY NPWMDIRQREDVQALNISVPYGPPIPVDFQRKIRQSYFASVSYLDTQVGR LLSALDDLQLANSTIIAFTSDHGWALGEHGEWAKYSNFDVATHVPLIFY VPGRTASLPEAGEKLFYLDPFDSASQLMEPGRQSMDELVELVSLFPTLA GLAGLQVPPRCVPVPSFHVELCREGKNLLKHFRFRDLEEDPYLPGNPREL IAYSQYPRPSDIPQWNSDKPSLKDIIKIMGYSIRTIDYRYTVWVGFPNDE FLANFSDIHAGELYFVDSDDLQDHNMYNDSQGGDLFQLLMP</p>	
<p>I2S-MTf</p> <p>(SP: Flag TAG y 10xHis TAG : TEV PS: IDS : L rigido: p97 soluble)</p>	<p>MEWSWVFLFFLSVTTGVHSDYKDDDDKEQKLISEEDLHHHHHHHHHHGG GGENLYFQGSETQANSTTDALNVLIIIVDDLRLPSLGCYGDKLVRSPNID QLASHSLLFQNAFAQQAVCAPSRVSFLTGRRPDTTRLYDFNSYWRVHAG NFSTIPQYFKENGYVTMSVGKVFHPGISSNHTDDSPYSWSFPPYHPSSE KYENTKTCRGPDGELHANLLCPVDVLDVPEGTLPDKQSTEQAIIQLEKMK TSASPFFLAVGYHKPHIPFRYPKEFKLYPLENITLAPDPEVDPGLPP VAYNPWMDIRQREDVQALNISVPYGPPIPVDFQRKIRQSYFASVSYLDTQ VGRLLSALDDLQLANSTIIAFTSDHGWALGEHGEWAKYSNFDVATHVPL IFYVPGRTASLPEAGEKLFYLDPFDSASQLMEPGRQSMDELVELVSLFP TLAGLAGLQVPPRCVPVPSFHVELCREGKNLLKHFRFRDLEEDPYLPGNP RELIAYSQYPRPSDIPQWNSDKPSLKDIIKIMGYSIRTIDYRYTVWVGFPN PDEFLANFSDIHAGELYFVDSDDLQDHNMYNDSQGGDLFQLLMPEAAAK EAAAKEAAKGMEVWRCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGT SADHCVQLIAAQEADAITLDGGAIYEAGKEHGLKPVVGEVYDQEVGTSY YAVAVVRRSSSHVTIDTLKGVKSCHTGINRTVGWNVPGYLVESEGRLSVM GCDVLKAVSDYFGGSCVPGAGETS YSESLCRLCRGDSSGEGVCDKSPLE RYYDYSGAFRCLAEGAGDVAFVKHSTVLENTDGKTLPSWGQALLSQDFE LLCRDGSRADVTEWRQCHLARVPAHAVVVRADTDGGLIFRLLNEGQRLF SHEGSSSQMFSSSEAYGQKDLLFKDSTSELVPIATQTYEAWLGHEYLHAM KGLLCDPNRLPPYLRWCVLSTPEIQKCGDMAVAFRRQRLKPEIQCVSAK SPQHCHMERIQAEQVDAVTLSGEDIYTAGKTYGLVPAAGEHYAPEDSSNS YYVAVVRRDSSSHAFTLDELGRKRSCHAGFGSPAGWDVPGALIQRGFI RPKDCDVLTAVERSEFNASCVPVNNPKNYPSSSLCALCVGDEQGRNKCVCN SQERYYYGYRGAFRCLVENAGDVAFVRHTTVFDNTNGHNSEPWAAELRSE DYELLCPNGARAEVSQFAACNLAQIPPHAVMVRPDTNIFTVYGLLDKAQ DLFQDDHNKNGFKMFDSSSNYHGQD LLLFKDATVRAVPVGEKTTYRGWLGL DYVAALEGMSSQQCS</p>	138

Tabla 1: Proteínas de fusión de p97 ilustrativas		
Descripción	Secuencia	SEQ ID NO:
MTf-I2S (SP : Flag TAG y 10xHIS TAG : TEV PS : P97 soluble : L rigido : IDS)	MEWSWVFLFFLSVTTGVHS DYKDDDDKEQKLI SEEDLHHHHHHHHHHGG GG ENLYFQG GMEVRCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTS ADHCVQLIAAQEADAITLDGGAIYEAGKEHGLKPVVGEVYDQEVGTSYY AVAVVRRSSHVITIDLKGVKSCHTGINRTVGWNVVPGYLVESGRLSVMG CDVLKAVSDYFGGSCVPGAGETSYSESLCRLCRGDSSGEGVCDKSPLE YYDYSGAFRCLAEGAGDVAFVKHSTVLENTDGKTLPSWGQALLSQDFEL LCRDGSRADVTEWRQCHLARVPAHAVVVRADTDGGLIFRLLNEGQRLFS HEGSSSQMFSSSEAYGQKDLLFKDSTSELVPIATQTYEAWLGHEYLHAMK GLLCDPNRLPPYLRLWCVLSTPEIQKCGDMAVAFRRQRLKPEIQCVSAKS PQHCMERIQAEQVDAVTLSGEDIYTAGKTYGLVPAAGEHYAPEDSSNSY YVVAVVRDSSSHAFTLDELRGKRSCHAGFGSPAGWDVPVGALIQRGFIR PKDCDVLTAVERSEFFNASCVPVNNPKNYPSSLCALCVGDEQGRNKC VGNS QERYYGYRGAFRCLVENAGDVAFVRHTTVFDNTNGHNSEPWAAELRSED YELLCPNGARAEVSQFAACNLAQIPPHAVMVREDTNIIFTVYGLLDKAQD LFGDDHNKNGFKMFSSNYHGQDLLFKDATVRAVPVGEKTTYRGWLGLD YVAALEGMSQQCS EAAAKEAAAKEAAAK SETQANSTTDALNVLLIIVD DLRPSLGCYGDKLVRSFNIDQLASHSLLFQNAFAQQAVCAPSRVSFLT RRPDTTRLYDFNSYWRVHAGNFSTIPQYFKENGYVTMSVGKVFHFGISS NHTDDSPYSWSFPPYHPSSEKYENTKTCRGPDGELHANLLCPVDVLDVP EGTLPDKQSTEQAIQLLEKMKTSASPFLLAVGYHKPHIPFRYPKEFQKL YPLENITLAPDPEVPDGLPPVAYNPWMDIRQREDVQALNISVPYGPPIPV DFQRKIRQSYFASVSYLDTQVGRLLSALDDLQLANSTIIAFTSDHGWA LGEHGEWAKYSNFDVATHVPLIFYVPGRTASLPEAGEKLPYLDPFDSAS QLMEPGRQSM DLVELVSLFPTLAGLAGLQVPPRCVPVPSFHVLCREGKN LLKHFRFRDLEEDPYLPGNPRELIAYSQYPRPSDIPQWNSDKPSLKD IKIMGYISIRTIDYRYTVWVGFPNDEF LANFSDIHAGELYFVDS DPLQDHNMYNDSQGGDLFQLLMP	139
MTfpep-I2S (SP : Flag TAG y 10xHIS TAG : TEV PS : MTfpep c/extremo C Y: L rigido : I2S)	MEWSWVFLFFLSVTTGVHS DYKDDDDKEQKLI SEEDLHHHHHHHHHHGG GG ENLYFQG DDSSHAFTLDELRY EAAAKEAAAKEAAAK SETQANSTTDAL NVLLIIVDDLRLPSLGCYGDKLVRSFNIDQLASHSLLFQNAFAQQAVCAP SRVSFLTGRRPDTTRLYDFNSYWRVHAGNFSTIPQYFKENGYVTMSVGK VFHFGISSNHTDDSPYSWSFPPYHPSSEKYENTKTCRGPDGELHANLLC PVDVLDVPEGTLPDKQSTEQAIQLLEKMKTSASPFLLAVGYHKPHIPFR YPKEFQKLYPLENITLAPDPEVPDGLPPVAYNPWMDIRQREDVQALNIS VPYGPPIPVDFQRKIRQSYFASVSYLDTQVGRLLSALDDLQLANSTIIAF TSDHGWA LGEHGEWAKYSNFDVATHVPLIFYVPGRTASLPEAGEKLPY LDPFDSASQLMEPGRQSM DLVELVSLFPTLAGLAGLQVPPRCVPVPSFHV ELCREGKNLLKHFRFRDLEEDPYLPGNPRELIAYSQYPRPSDIPQWNSD KPSLKD IKIMGYISIRTIDYRYTVWVGFPNDEF LANFSDIHAGELYFVDS DPLQDHNMYNDSQGGDLFQLLMP	140

Tabla 1: Proteínas de fusión de p97 ilustrativas		
Descripción	Secuencia	SEQ ID NO:
I2S-MTfpep (SP : Flag TAG y 10xHIS TAG : TEV PS : I2S : L rígido : MTfpep c/extremo C Y)	<u>MEWSWVFLFFLSVTTGVHS</u> DYKDDDDKEQKLI SEEDLHHHHHHHHHHGGG <u>ENLYFQG</u> SETQANSTTDALNVLLIIVDDLRLPSLGCYGDKLVRSPNIDQLASHSLLFQNAFAQQAVCAPSRVSFLTGRRPDTTRLYDFNSYWRVHAGNFSTIPQYFKENGYVTMSVGKVFHHPGISSNHTDDSPYSWSFPPYHPSSEKYENTKTCRGPDGELHANLLCPVDVLDVPEGTLDPDKQSTEQAIQLLEKMKTSASPPFLAVGYHKPHIPFRYPKEFQKLYPLENITLAPDPEVPDGLPPVAYNPWMDIRQREDVQALNISVPYGPPIPVDFQRKIRQSYFASVSYLDTQVGRLLSALDDLQLANSTIIAFTSDHGWALGEHGEWAKYSNFDVATHVPLIFYVPGRTASLPEAGEKLFYLDPFDSASQLMEPGRQSMDLVELVSLFPTLAGLAGLQVPPRCVPVPSFHVELCREGKNLLKHFRFRDLEEDPYLPGNPRELIAYSQYPRPSDIPQWNSDKPSLKDIIKIMGYSIRTIDYRYTVWVGFPDEFANFSDIHAGELYFVDSPLQDHNMYNDSQGGDLFQLLMP <u>EAAAKEAAK</u> DSSHAFTLDELRY	141
I2S-MTfpep (sin propéptido de I2S) SP : Flag TAG y 10xHIS TAG : TEV PS : I2S sin propéptido : L rígido : MTfpep c/extremo C Y)	<u>MEWSWVFLFFLSVTTGVHS</u> DYKDDDDKEQKLI SEEDLHHHHHHHHHHGGG <u>ENLYFQG</u> TDALNVLLIIVDDLRLPSLGCYGDKLVRSPNIDQLASHSLLFQNAFAQQAVCAPSRVSFLTGRRPDTTRLYDFNSYWRVHAGNFSTIPQYFKENGYVTMSVGKVFHHPGISSNHTDDSPYSWSFPPYHPSSEKYENTKTCRGPDGELHANLLCPVDVLDVPEGTLDPDKQSTEQAIQLLEKMKTSASPPFLAVGYHKPHIPFRYPKEFQKLYPLENITLAPDPEVPDGLPPVAYNPWMDIRQREDVQALNISVPYGPPIPVDFQRKIRQSYFASVSYLDTQVGRLLSALDDLQLANSTIIAFTSDHGWALGEHGEWAKYSNFDVATHVPLIFYVPGRTASLPEAGEKLFYLDPFDSASQLMEPGRQSMDLVELVSLFPTLAGLAGLQVPPRCVPVPSFHVELCREGKNLLKHFRFRDLEEDPYLPGNPRELIAYSQYPRPSDIPQWNSDKPSLKDIIKIMGYSIRTIDYRYTVWVGFPDEFANFSDIHAGELYFVDSPLQDHNMYNDSQGGDLFQLLMP <u>EAAAKEAAKEAAK</u> DSSHAFTLDELRY	142

Por lo tanto, en algunas realizaciones, la proteína de fusión comprende, consiste o consiste esencialmente en una secuencia de aminoácidos de la Tabla 1, o una variante y/o fragmento de esta.

- 5 Secuencias de p97. Una secuencia polipeptídica de p97 usada en una composición y/o proteína de fusión descrita en la presente memoria comprende, consiste esencialmente, o consiste en una secuencia de referencia de p97 humana proporcionada en la Tabla 2 a continuación. También se incluyen variantes y fragmentos de esta. La secuencia polipeptídica de p97 usada en una composición y/o proteína de fusión de la invención consiste en la secuencia de p97 humana proporcionada como SEQ ID NO: 14 en la Tabla 2.

Tabla 2: Secuencias de p97 ilustrativas		
Descripción	Secuencia	SEQ ID NO:
p97 humana FL	<u>M</u> RGP SGALWLLALRTVLGGMEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIAAQEADAITLDGGAIYEAGKEHGLKPVVGEVYDQEVGTSYYAVAVVRRSSHVTIDTLKGVKSCHTGINRTVGWNVVPGYLVESGRLSVMGCDVLKAVSDYFGGSCVPGAGETSYSESLCRLCRGDSSGEGVCDKSPLEERYDYSGAFRCLAEGAGDVAFAVKHSTVLENTDGKTLPSWGQALLSQDFELLCRDGSRADVTEWRQCHLARVPAHAVVVRADTDGGLIFRLLNEGQRLFSHEGSSFQMFSSSEAYGQKDLLFKDSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPNRLPPYLRCVLSLTPETIQKCGDMAVAFRRQRLKPEIQCVSAKSPQHCMERIQAEQVDAVTLSGEDIYTAGKTYGLVPAAGEHYAPEDSSNSYVAVVVRDSSHAFTLDELRLGKRSCHAGFGSPAGWDVPVGALIQRGFIRPKDCDVLTAVERSEFFNASCVPVNNPKNYPSSLCALCVGDEQGRNKC VGNSQERYYG YRGAFRCLVENAGDVAFVRHTTVFDNTNGHNSPWA AELRSEDYELLCPNGARAEVSQFAACNLAQIPPHAVMVRPDTNIFTVYGLLDKAQDLFGDDHNKNGFKMFDSSNYHGQDLLFKDATVRAVPVGEKTTYRGWLGLDYVA ALEGMS SQCSGAAAPAPGAPLLPLLLPALAARLLPAL	1

Tabla 2: Secuencias de p97 ilustrativas		
Descripción	Secuencia	SEQ ID NO:
P97 humana soluble	GMEVRCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIA AQEADAITLDGGAIYEAGKEHGLKPVVGEVYDQEVGTSYYAVAVVRRSS HVTIDTLKGVKSCHTGINRTVGWNPVGYLVESGRLSVMGCDVLKAVSD YFGGSCVPGAGETSYSESLCRLCRGDSSGEGVCDKSPLEYYDYSGAFR CLAEAGADVAFVKHSTVLENTDGKTLPSWGQALLSQDFELLCRDGSRAD VTEWRQCHLARVPAHAVVVRADTDGGLIFRLLNEGQRLFSHEGSSSQMF SSEAYGQKDLLFKDSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPNRL PPYLRWCVLSTPEIQKCGDMAVAFRRQRLKPEIQCVSAKSPQHCMERIQ AEQVDAVTLSGEDIYTAGKTYGLVPAAGEHYAPEDSSNSYYVAVVRRD SSHAFTLDELRGKRSCHAGFGSPAGWDVPVGALIQRGFIRPKDCDVLTA VSEFFNASCVPVNNPKNYPSSLCALCVGDEQGRNKC VGNSQERYYYGYRG AFRCLVENAGDVAFVRHTTVFDNTNGHNSEPWAAELRSEDYELLCPNGA RAEVSQFAACNLAQIPPHAVMVRPDTNIFTVYGLLDKAQDLFGDDHNKN GFKMFDSSNYHGQDLLFKDATVRAVPVGEKTTYRGWLGLDYVAALLEGMS SQQCS	2
Fragmento de p97	WCATSDPEQHK	3
Fragmento de p97	RSSHVTIDTLK	4
Fragmento de p97	SSHVTIDTLKGVK	5
Fragmento de p97	LCRGDSSGEGVCDK	6
Fragmento de p97	GDSSGEGVCDKSPLER	7
Fragmento de p97	YYDYSGAFR	8
Fragmento de p97	ADVTEWR	9
Fragmento de p97	VPAHAVVVR	10
Fragmento de p97	ADTDGGLIFR	11
Fragmento de p97	CGDMAVAFR	12
Fragmento de p97	LKPEIQCVSAK	13
Fragmento de p97	DSSHAFTLDEL	14
Fragmento de p97	14	148
Fragmento de p97	SEDYELLCPNGAR	15
Fragmento de p97	AQDLFGDDHNKNKGFK	16
Fragmento de p97	FSSEAYGQKDLLFKDSTSELVPIATQTYEAWLGHEYLHAM	17
Fragmento de p97	ERIQAEQVDAVTLSGEDIYTAGKTYGLVPAAGEHYAPEDSSNSYYVAV VRRDSSHAFTLDELRGKRSCHAGFGSPAGWDVPVGALIQRGFIRPKDCD VLTAVSEFFNASCVPVNNPKNYPSSLCALCVGDEQGRNKC VGNSQERY GYRGAFRCLVENAGDVAFVRHTTVFDNTNGHNSEPWAAELRSEDYELL CPNGARAEVSQFAACNLAQIPPHAVM	18

Tabla 2: Secuencias de p97 ilustrativas		
Descripción	Secuencia	SEQ ID NO:
Fragmento de p97	VRPDTNIFTVYGLLDKAQDLFGDDHNKNGFKM	19
Fragmento de p97	GMEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIA AQEADAITLDGGAIYEAGKEHGLKPVVGEVYDQEVGTSYYAVAVVRRSS HVTIDTLKGVKSCHTGINRTVGWNPVGYLVESGRLSVMGCDVLKAVSD YFGGSCVPGAGETSYSESLCRLCRGDSGEGVCDKSPLERYDYSGAFR CLAEGAGDVAFVKHSTVLENTDGKTLPSWGQALLSQDFELLCRDGSRAD VTEWRQCHLARVPAHAVVVRADTDGGLIFRLLNEGQRLFSHEGSSSQMF SSEAYGQKDLLFKDSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPNRL PPYLRWCVLSTPEIQKCGDMAVAFRRLKPEIQCVSAKSPQHCMERIQ AEQVDAVTLSGEDIYTAGKTYGLVPAAGEHYAPEDSSNSYYVAVVRRD SSHAFTLDELRGKRSCHAGFGSPAGWDVPVGALIQRGFIRPKDCDVLTA VSEFFNASCVPVNNPKNPSSLCALCVGDEQGRNKC VGNSQERYYG YRG AFRCLVENAGDVAFVRHTTVFDNTN	20
Fragmento de p97	GHNSEPWAAELRSEDYELLCPN	21
Fragmento de p97	GARAEVSQFAACNLAQIPPHAVMVRPDTNIFTVYGLLDKAQDLFGDDHN KN	22
Fragmento de p97	GFKMFDSSNYHGQDLLFKDATVRAVPVGEKTTYRGWLGLDYVAALEGMS SQQC	23
Fragmento de p97	GMEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIA AQEADAITLDGGAIYEAGKEHGLKPVVGEVYDQEVGTSYYAVAVVRRSS HVTIDTLKGVKSCHTGINRTVGWNPVGYLVESGRLSVMGCDVLKAVSD YFGGSCVPGAGETSYSESLCRLCRGDSGEGVCDKSPLERYDYSGAFR CLAEGAGDVAFVKHSTVLENTDGKTLPSWGQALLSQDFELLCRDGSRAD VTEWRQCHLARVPAHAVVVRADTDGGLIFRLLNEGQRLFSHEGSSSQMF SSEAYGQKDLLFKDSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPNRL PPYLRWCVLSTPEIQKCGDMAVAFRRLKPEIQCVSAKSPQHCMERIQ AEQVDAVTLSGEDIYTAGKTYGLVPAAGEHYAPEDSSNSYYVAVVRRD SSHAFTLDELRGKRSCHAGFGSPAGWDVPVGALIQRGFIRPKDCDVLTA VSEFFNASCVPVNNPKNPSSLCALCVGDEQGRNKC VGNSQERYYG YRG AFRCLVENAGDVAFVRHTTVFDNTNGHNSEPWAAELRSEDYELLCPN	24
Fragmento de p97	GMEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIA AQEADAITLDGGAIYEAGKEHGLKPVVGEVYDQEVGTSYYAVAVVRRSS HVTIDTLKGVKSCHTGINRTVGWNPVGYLVESGRLSVMGCDVLKAVSD YFGGSCVPGAGETSYSESLCRLCRGDSGEGVCDKSPLERYDYSGAFR CLAEGAGDVAFVKHSTVLENTDGKTLPSWGQALLSQDFELLCRDGSRAD VTEWRQCHLARVPAHAVVVRADTDGGLIFRLLNEGQRLFSHEGSSSQMF SSEAYGQKDLLFKDSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPNRL PPYLRWCVLSTPEIQKCGDMAVAFRRLKPEIQCVSAKSPQHCMERIQ AEQVDAVTLSGEDIYTAGKTYGLVPAAGEHYAPEDSSNSYYVAVVRRD SSHAFTLDELRGKRSCHAGFGSPAGWDVPVGALIQRGFIRPKDCDVLTA VSEFFNASCVPVNNPKNPSSLCALCVGDEQGRNKC VGNSQERYYG YRG AFRCLVENAGDVAFVRHTTVFDNTNGHNSEPWAAELRSEDYELLCPNGA RAEVSQFAACNLAQIPPHAVMVRPDTNIFTVYGLLDKAQDLFGDDHNKN	25
Fragmento de p97	GHNSEPWAAELRSEDYELLCPNGARAEVSQFAACNLAQIPPHAVMVRPD TNIFTVYGLLDKAQDLFGDDHNKN	26
Fragmento de p97	GHNSEPWAAELRSEDYELLCPNGARAEVSQFAACNLAQIPPHAVMVRPD TNIFTVYGLLDKAQDLFGDDHNKNGFKMFDSSNYHGQDLLFKDATVRAV PVGEKTTYRGWLGLDYVAALEGMS SQQC	27

Tabla 2: Secuencias de p97 ilustrativas		
Descripción	Secuencia	SEQ ID NO:
Fragmento de p97	GARAEVSQFAACNLAQIPPHAVMVRPDTNIFTVYGLLDKAQDLFGDDHN KNGFKMFDSSNYHGQDLLFKDATVRAVPVGEKTTYRGWLGLDYVAALLEG MSSQQC	28

Una secuencia polipeptídica de p97 descrita en la presente memoria puede comprender una secuencia que tiene al menos 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad o homología, a lo largo de su longitud, con respecto a una secuencia de p97 humana en la Tabla 2, o un fragmento de esta.

- 5 La secuencia polipeptídica de p97 de la invención consiste en la SEQ ID NO:14 (MTfpep).

Una secuencia polipeptídica de p97 descrita en la presente memoria puede comprender un fragmento de una secuencia de p97 humana en la Tabla 2. Un fragmento polipeptídico de p97 descrito en la presente memoria puede tener aproximadamente, al menos aproximadamente, o hasta aproximadamente 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 700, 710, 720, 730 o más aminoácidos de longitud, incluidos todos los números enteros e intervalos entremedio, y que puede comprender toda o una porción de la secuencia de una secuencia de referencia de p97.

Un fragmento polipeptídico de p97 descrito en la presente memoria puede tener aproximadamente 5-700, 5-600, 5-500, 5-400, 5-300, 5-200, 5-100, 5-50, 5-40, 5-30, 5-25, 5-20, 5-15, 5-10, 10-700, 10-600, 10-500, 10-400, 10-300, 10-200, 10-100, 10-50, 10-40, 10-30, 10-25, 10-20, 10-15, 20-700, 20-600, 20-500, 20-400, 20-300, 20-200, 20-100, 20-50, 20-40, 20-30, 20-25, 30-700, 30-600, 30-500, 30-400, 30-300, 30-200, 30-100, 30-50, 30-40, 40-700, 40-600, 40-500, 40-400, 40-300, 40-200, 40-100, 40-50, 50-700, 50-600, 50-500, 50-400, 50-300, 50-200, 50-100, 60-700, 60-600, 60-500, 60-400, 60-300, 60-200, 60-100, 60-70, 70-700, 70-600, 70-500, 70-400, 70-300, 70-200, 70-100, 70-80, 80-700, 80-600, 80-500, 80-400, 80-300, 80-200, 80-100, 80-90, 90-700, 90-600, 90-500, 90-400, 90-300, 90-200, 90-100, 100-700, 100-600, 100-500, 100-400, 100-300, 100-250, 100-200, 100-150, 200-700, 200-600, 200-500, 200-400, 200-300 o 200-250 aminoácidos de longitud, y comprende toda o una porción de una secuencia de referencia de p97.

Las secuencias polipeptídicas de p97 descritas en la presente memoria incluyen secuencias de aminoácidos de p97, subsecuencias y/o variantes de p97 que son eficaces para transportar un agente de interés a través de la barrera hematoencefálica y hacia el sistema nervioso central (CNS). La variante o fragmento descrito en la presente memoria puede comprender el lóbulo N de p97 humana (residuos 20-361 de la SEQ ID NO:1). En aspectos específicos, la variante o fragmento descrito en la presente memoria puede comprender un sitio de unión a Fe³⁺ intacto y funcional.

Una secuencia polipeptídica de p97 descrita en la presente memoria puede ser una forma soluble de un polipéptido de p97 (ver Yang et al., Prot Exp Purif. 34:28-48, 2004), o un fragmento o variante de esta. En algunos aspectos, el polipéptido de p97 soluble descrito en la presente memoria puede tener una eliminación de todo o de una porción del dominio hidrófobo (residuos 710-738 de la SEQ ID NO:1), sola o en combinación con una eliminación de todo o de una porción del péptido señal (residuos 1-19 de la SEQ ID NO:1). En aspectos específicos, el polipéptido de p97 soluble descrito en la presente memoria puede comprender o consistir en la SEQ ID NO:2 (~residuos 20-710 o 20-711 de la SEQ ID NO:1), incluidas variantes y fragmentos de esta.

Para aquellos que emplean liposomas, la secuencia polipeptídica de p97 descrita en la presente memoria puede ser una forma soluble lipídica de un polipéptido de p97. Por ejemplo, algunos de estos y polipéptidos relacionados pueden incluir un polipéptido de p97 que comprende todo o una porción del dominio hidrófobo, opcionalmente con o sin el péptido señal.

El fragmento o variante de p97 descrito en la presente memoria puede ser capaz de unirse específicamente a un receptor de p97, un receptor de LRP1 y/o un receptor de LRP1B.

Las variantes y fragmentos de polipéptidos de p97 de referencia descritos en la presente memoria y otros polipéptidos de referencia se describen en mayor detalle más adelante.

Secuencias de iduronato-2-sulfatasa. En ciertas realizaciones, una secuencia polipeptídica de IDS (o I2S) usada en una proteína de fusión de la invención comprende, consiste esencialmente, o consiste en una o más secuencias de IDS humana ilustradas en la Tabla 3 a continuación.

Tabla 3: Secuencias de IDS ilustrativas		
Nombre	Secuencia	SEQ ID NO:
IDS humana de longitud completa (secuencia subrayada)	<u>MPPPRRTGRGLLWLGLVLSSVCVALGSETQANSTTDALNVLLIIIVDDLRLP</u> <u>SLGCGYGDKLVRSPNIDQLASHSLLFQNAFAQQAVCAPSRVSFLTGRRPD</u> <u>TTRYLYDFNSYWRVHAGNFSTIPQYFKENG YVTMSVGKVFHFGISSNHTD</u> <u>DSPYSWSFPPYHPSSEKYENTKTCRGPDGELHANLLCPVDVLDVPEGTL</u> <u>PDQKQSTEQAIQLLEKMKTSASPFLLAVGYHKPHIPFRYPKEFQKLYPLE</u> <u>NITLAPDPEVPDGLPPVAYNPWMDIRQREDVQALNISVPYGPPIPVDFQR</u> <u>KIRQSYFASVSYLDTQVGRLLSALDDLQLANSTIIAFTSDHGWALGEHG</u> <u>EWAKYSNFDVATHVPLIFYVPGRTASLPEAGEKLFYLDPFDSASQLME</u> <u>PGRQSM DLVLSLFP TLAGLAGLQVPPRCVPVPSFHVELCREGKNLLKH</u> <u>FRFRDLEEDPYLPGNPRELIAYSQYPRPSDIPQWNSDKPSLKD IKIMGY</u> <u>SIRTIDYRYTVWVGFPNDEF LANSFSDIHAGELYFVDS DPLQDHNMYNDS</u> <u>QGGDLFQLLMP</u>	31
IDS humana con propéptido secuencia (subrayada) pero sin secuencia señal	<u>SETQANSTTDALNVLLIIIVDDLRLPSLGCYGDKLVRSPNIDQLASHSLLF</u> <u>QNAFAQQAVCAPSRVSFLTGRRPDTTRYLYDFNSYWRVHAGNFSTIPQYF</u> <u>KENG YVTMSVGKVFHFGISSNHTD DSPYSWSFPPYHPSSEKYENTKTCR</u> <u>GPDGELHANLLCPVDVLDVPEGTL PDKQSTEQAIQLLEKMKTSASPFLL</u> <u>AVGYHKPHIPFRYPKEFQKLYPLENITLAPDPEVPDGLPPVAYNPWMDI</u> <u>RQREDVQALNISVPYGPPIPVDFQRKIRQSYFASVSYLDTQVGRLLSALD</u> <u>DLQLANSTIIAFTSDHGWALGEHGEWAKYSNFDVATHVPLIFYVPGRTA</u> <u>SLPEAGEKLFYLDPFDSASQLMEPGRQSM DLVLSLFP TLAGLAGLQ</u> <u>VPPRCVPVPSFHVELCREGKNLLKHFRFRDLEEDPYLPGNPRELIAYSQY</u> <u>PRPSDIPQWNSDKPSLKD IKIMGYSIRTIDYRYTVWVGFPNDEF LANSF</u> <u>SDIHAGELYFVDS DPLQDHNMYNDSQGGDLFQLLMP</u>	32
IDS humana sin propéptido secuencia señal	<u>TDALNVLLIIIVDDLRLPSLGCYGDKLVRSPNIDQLASHSLLFQNAFAQQA</u> <u>VCAPSRVSFLTGRRPDTTRYLYDFNSYWRVHAGNFSTIPQYFKENG YVTM</u> <u>SVGKVFHFGISSNHTD DSPYSWSFPPYHPSSEKYENTKTCRGPDGELHA</u> <u>NLLCPVDVLDVPEGTL PDKQSTEQAIQLLEKMKTSASPFLLAVGYHKPH</u> <u>IPFRYPKEFQKLYPLENITLAPDPEVPDGLPPVAYNPWMDIRQREDVQA</u> <u>LNISVPYGPPIPVDFQRKIRQSYFASVSYLDTQVGRLLSALDDLQLANST</u> <u>IIAFTSDHGWALGEHGEWAKYSNFDVATHVPLIFYVPGRTASLPEAGEK</u> <u>LFPYLDPFDSASQLMEPGRQSM DLVLSLFP TLAGLAGLQVPPRCVPV</u> <u>SFHVELCREGKNLLKHFRFRDLEEDPYLPGNPRELIAYSQYPRPSDIPQ</u> <u>WNSDKPSLKD IKIMGYSIRTIDYRYTVWVGFPNDEF LANSFSDIHAGELY</u> <u>FVDS DPLQDHNMYNDSQGGDLFQLLMP</u>	33
Cadena de 42 kDa de IDS humana	<u>TDALNVLLIIIVDDLRLPSLGCYGDKLVRSPNIDQLASHSLLFQNAFAQQA</u> <u>VCAPSRVSFLTGRRPDTTRYLYDFNSYWRVHAGNFSTIPQYFKENG YVTM</u> <u>SVGKVFHFGISSNHTD DSPYSWSFPPYHPSSEKYENTKTCRGPDGELHA</u> <u>NLLCPVDVLDVPEGTL PDKQSTEQAIQLLEKMKTSASPFLLAVGYHKPH</u> <u>IPFRYPKEFQKLYPLENITLAPDPEVPDGLPPVAYNPWMDIRQREDVQA</u> <u>LNISVPYGPPIPVDFQRKIRQSYFASVSYLDTQVGRLLSALDDLQLANST</u> <u>IIAFTSDHGWALGEHGEWAKYSNFDVATHVPLIFYVPGRTASLPEAGEK</u> <u>LFPYLDPFDSASQLMEPGRQSM DLVLSLFP TLAGLAGLQVPPRCVPV</u> <u>SFHVELCREGKNLLKHFRFRDLEEDPYLP</u>	34
Cadena de 14 kDa de IDS humana	<u>NPRELIAYSQYPRPSDIPQWNSDKPSLKD IKIMGYSIRTIDYRYTVWVG</u> <u>FNPDEF LANSFSDIHAGELYFVDS DPLQDHNMYNDSQGGDLFQLLMP</u>	35

- 5 También se incluyen variantes y fragmentos biológicamente activos de las secuencias de IDS en la Tabla 3 y la Lista de secuencias. En ciertos aspectos, un polipéptido de IDS biológicamente activo o variantes/fragmento de este hidrolizan los grupos 2-sulfato de las unidades L-iduronato 2-sulfato de sulfato de dermatán, sulfato de heparán y/o heparina, por ejemplo, a aproximadamente 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85

%, 90 %, 95 %, 100 %, 200 %, 300 %, 400 %, 500 % o más de la actividad de la IDS humana natural (p. ej., SEQ ID NO:31).

Enlazadores. Según se indicó anteriormente, ciertas proteínas de fusión pueden emplear uno o más grupos enlazadores, incluidos enlazadores peptídicos. Dichos enlazadores pueden ser enlazadores rígidos, enlazadores flexibles, enlazadores estables o enlazadores liberables, tales como enlazadores enzimáticamente escindibles. Ver, p. ej., Chen et al., Adv. Drug. Deliv. Ref., 65:1357-69, 2012.

Por ejemplo, para conjugados polipéptido-polipéptido, los enlazadores peptídicos pueden separar los componentes una distancia suficiente para garantizar que cada polipéptido se pliegue en sus estructuras secundarias y terciarias. Dicha secuencia de enlazador peptídico se puede incorporar en la proteína de fusión mediante el uso de técnicas estándares descritas en la presente memoria y conocidas en la técnica. Las secuencias de enlazador peptídico adecuadas se pueden elegir en función de los siguientes factores: (1) su capacidad para adoptar una conformación extendida rígida o flexible; (2) su incapacidad para adoptar una estructura secundaria que podrían interactuar con epítopos funcionales en el primer y segundo polipéptidos; y (3) la carencia de residuos hidrófobos o cargados que podrían hacer reacción con los epítopos funcionales del polipéptido. Las secuencias de aminoácidos que pueden emplearse de manera útil como enlazadores incluyen las descritas en Maratea et al., Gene 40:39-46, 1985; Murphy et al., Proc. Natl. Acad. Sci. USA 83:8258-8262, 1986; la patente estadounidense núm. 4 935 233 y la patente estadounidense núm. 4 751 180.

En ciertas realizaciones ilustrativas, un enlazador peptídico tiene entre aproximadamente 1 y 5 aminoácidos, entre 5 y 10 aminoácidos, entre 5 y 25 aminoácidos, entre 5 y 50 aminoácidos, entre 10 y 25 aminoácidos, entre 10 y 50 aminoácidos, entre 10 y 100 aminoácidos, o cualquier intervalo intermedio de aminoácidos. En otras realizaciones ilustrativas, un enlazador peptídico comprende aproximadamente 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50 o más aminoácidos de longitud. Los enlazadores particulares puede tener una longitud de aminoácidos general de aproximadamente 1-200 aminoácidos, 1-150 aminoácidos, 1-100 aminoácidos, 1-90 aminoácidos, 1-80 aminoácidos, 1-70 aminoácidos, 1-60 aminoácidos, 1-50 aminoácidos, 1-40 aminoácidos, 1-30 aminoácidos, 1-20 aminoácidos, 1-10 aminoácidos, 1-5 aminoácidos, 1-4 aminoácidos, 1-3 aminoácidos o aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 60, 70, 80, 90, 100 o más aminoácidos.

Un enlazador peptídico puede emplear uno cualquiera o más aminoácidos de origen natural, aminoácido(s) de origen no natural, análogos de aminoácidos y/o miméticos de aminoácido según se describe en otra parte en la presente memoria y se conoce en la técnica. Ciertas secuencias de aminoácidos que pueden emplearse de manera útil como enlazadores incluyen las descritas en Maratea et al., Gene 40:39-46, 1985; Murphy et al., PNAS USA. 83:8258-8262, 1986; la patente estadounidense núm. 4 935 233 y la patente estadounidense núm. 4 751 180. Las secuencias de enlazador peptídico particulares contienen residuos Gly, Ser y/o Asn. Se pueden emplear otros aminoácidos casi neutros, tales como Thr y Ala en la secuencia de enlazador peptídico, si se desea.

En realizaciones específicas, el enlazador es un enlazador rígido. Los ejemplos de enlazadores rígidos incluyen, sin limitación, (EAAAK)_x (SEQ ID NO:36) y A(EAAAK)_x ALEA(EAAAK)_x A (SEQ ID NO:41), y (Ala-Pro)_x donde x es 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 o más. Los ejemplos específicos de enlazadores rígidos incluyen EAAAK (SEQ ID NO:36), (EAAAK)₂ (SEQ ID NO:37), (EAAAK)₃ (SEQ ID NO:38), A(EAAAK)₄ ALEA(EAAAK)₄ A (SEQ ID NO:42), PAPAP (SEQ ID NO:43) y AEAAAKEAAAKA (SEQ ID NO:44).

En realizaciones específicas, el enlazador comprende, consiste o consiste esencialmente en (EAAAK)₃ o EAAAKEAAAKEAAAK (SEQ ID NO:38).

En algunas realizaciones, el enlazador es un enlazador flexible. En realizaciones específicas, el enlazador flexible es GGGGS (SEQ ID NO:45), (GGGGS)₂ (SEQ ID NO:46), (GGGGS)₃ (SEQ ID NO:47) o Gly₂₋₁₀ (SEQ ID NOS:48-54). A continuación, se proporcionan ejemplos adicionales de enlazadores flexibles.

Ciertos enlazadores ilustrativos incluyen enlazadores que contienen Gly, Ser y/o Asn, como los siguientes: enlazadores [G]_x, [S]_x, [N]_x, [GS]_x, [GGS]_x, [GSS]_x, [GSGS]_x (SEQ ID NO:55), [GGSG]_x (SEQ ID NO:56), [GGGS]_x (SEQ ID NO: 57), [GGGGS]_x (SEQ ID NO: 45), [GN]_x, [GGN]_x, [GNN]_x, [GNGN]_x (SEQ ID NO: 58), [GGNG]_x (SEQ ID NO: 59), [GGGN]_x (SEQ ID NO: 60), [GGGGN]_x (SEQ ID NO: 61), donde x es 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 o más. Otras combinaciones de estos y aminoácidos relacionados serán evidentes para las personas con experiencia en la técnica. En realizaciones específicas, el enlazador comprende o consiste en una secuencia [GGGGS]₃ (SEQ ID NO: 47) o GGGGSGGGSGGGGS (SEQ ID NO: 47).

En realizaciones específicas, la secuencia de enlazador comprende una secuencia de enlazador Gly₃, que incluye tres residuos glicina. En realizaciones específicas, los enlazadores flexibles se pueden diseñar racionalmente mediante el uso de un programa informático capaz de modelar sitios de unión a ADN y los propios péptidos (Desjarlais & Berg, PNAS. 90:2256-2260, 1993; y PNAS. 91:11099-11103, 1994) o mediante métodos de visualización en fagos.

Los enlazadores peptídicos pueden ser fisiológicamente estables o pueden incluir un enlazador liberable tal como un enlazador fisiológicamente degradable o enzimáticamente degradable (p. ej., enlazadores proteolíticamente o

enzimáticamente escindibles). En ciertas realizaciones, uno o más enlazadores liberables pueden dar como resultado una semivida más corta y aclaramiento más rápido de la proteína de fusión. Estas y realizaciones relacionadas se pueden usar, por ejemplo, para potenciar la solubilidad y tiempo de circulación en sangre de las proteínas de fusión de p97 en el torrente circulatorio, mientras también se suministra un agente en el torrente circulatorio (o a través de la BBB) que, después de la degradación del enlazador, está sustancialmente libre de la secuencia de p97. Estos aspectos son especialmente útiles en aquellos casos donde polipéptidos u otros agentes, cuando están fusionados de manera permanente a una secuencia de p97, demuestran actividad reducida. Mediante el uso de los enlazadores, según se proporcionan en la presente memoria, dichos polipéptidos pueden mantener su actividad terapéutica cuando están en forma conjugada o fusionada. De estas y otras maneras se pueden adaptar de manera más eficaz las propiedades de las proteínas de fusión de p97 para equilibrar la bioactividad y la semivida en circulación de los polipéptidos con el tiempo.

Los ejemplos específicos de enlazadores enzimáticamente escindibles incluyen, sin limitación, un enlazador escindible por Factor Xla/FVIIa (VSQTSKLTR▼AETVFPDV) (SEQ ID NO:62), un enlazador escindible por metaloproteasa-1 de matriz (PLG▼LWA) (SEQ ID NO:63), un enlazador escindible por proteasa de VIH (RVL▼AEA) (SEQ ID NO:64), un enlazador escindible por proteasa NS3 de virus de la hepatitis C (EDVVCC▼SMSY) (SEQ ID NO:65), un enlazador escindible por Factor Xa (GGIEGR/GS) (SEQ ID NO:66), un enlazador escindible por Furina (TRHRQPR▼GWE o AGNRVRR▼SVG o RRRRRRR▼R▼R) (SEQ ID NOS:67-69) y un enlazador escindible por Catepsina B (GFLG) (SEQ ID NO:70).

Los enlaces enzimáticamente degradables adecuados para su uso en realizaciones específicas incluyen, pero no se limitan a: una secuencia de aminoácidos escindida por una serina proteasa tal como trombina, quimotripsina, tripsina, elastasa, calicreína o subtilisina. Los ejemplos ilustrativos de secuencias de aminoácidos escindibles por trombina incluyen, pero no se limitan a: -Gly-Arg-Gly-Asp-(SEQ ID NO: 71), -Gly-Gly-Arg-, -Gly-Arg-Gly-Asp-Asn-Pro-(SEQ ID NO:72), -Gly-Arg-Gly-Asp-Ser-(SEQ ID NO: 73), -Gly-Arg-Gly-Asp-Ser-Pro-Lys-(SEQ ID NO: 74), -Gly-Pro-Arg-, -Val-Pro-Arg- y -Phe-Val-Arg-. Los ejemplos ilustrativos de secuencias de aminoácidos escindibles por elastasa incluyen, pero no se limitan a: -Ala-Ala-Ala-, -Ala-Ala-Pro-Val-(SEQ ID NO:75), -Ala-Ala-Pro-Leu-(SEQ ID NO: 76), -Ala-Ala-Pro-Phe-(SEQ ID NO: 77), -Ala-Ala-Pro-Ala-(SEQ ID NO: 78) y -Ala-Tyr-Leu-Val-(SEQ ID NO: 79).

Los enlaces enzimáticamente degradables adecuados para su uso en realizaciones específicas también incluyen secuencias de aminoácidos que se pueden escindir mediante una metaloproteinasa de matriz tal como collagenasa, estromelina y gelatinasa. Los ejemplos ilustrativos de secuencias de aminoácidos escindibles por metaloproteinasa de matriz incluyen, pero no se limitan a: -Gly-Pro-Y-Gly-Pro-Z-(SEQ ID NO: 80), -Gly-Pro-, Leu-Gly-Pro-Z-(SEQ ID NO: 81), -Gly-Pro-Ile-Gly-Pro-Z-(SEQ ID NO:82) y -Ala-Pro-Gly-Leu-Z-(SEQ ID NO: 83), donde Y Z son aminoácidos. Los ejemplos ilustrativos de secuencias de aminoácidos escindibles por collagenasa incluyen, pero no se limitan a: -Pro-Leu-Gly-Pro-D-Arg-Z-(SEQ ID NO: 84), -Pro-Leu-Gly-Leu-Leu-Gly-Z-(SEQ ID NO: 85), -Pro-Gln-Gly-Ile-Ala-Gly-Trp-(SEQ ID NO: 86), -Pro-Leu-Gly-Cys(Me)-His-(SEQ ID NO: 87), -Pro-Leu-Gly-Leu-Tyr-Ala-(SEQ ID NO:88), -Pro-Leu-Ala-Leu-Trp-Ala-Arg-(SEQ ID NO: 89) y -Pro-Leu-Ala-Tyr-Trp-Ala-Arg-(SEQ ID NO: 90), donde Z es un aminoácido. Un ejemplo ilustrativo de una secuencia de aminoácidos escindible por estromelina es -Pro-Tyr-Ala-Tyr-Tyr-Met-Arg-(SEQ ID NO: 91); y un ejemplo de una secuencia de aminoácidos escindible por gelatinasa es -Pro-Leu-Gly-Met-Tyr-Ser-Arg-(SEQ ID NO: 92).

Los enlaces enzimáticamente degradables adecuados para su uso en realizaciones específicas también incluyen secuencias de aminoácidos que se pueden escindir mediante una enzima convertidora de angiotensina tales como, por ejemplo, -Asp-Lys-Pro-, -Gly-Asp-Lys-Pro-(SEQ ID NO: 93) y -Gly-Ser-Asp-Lys-Pro-(SEQ ID NO: 94).

Los enlaces enzimáticamente degradables adecuados para su uso en realizaciones específicas también incluyen secuencias de aminoácidos que se pueden degradar mediante catepsina B, tales como, por ejemplo, -Val-Cit-, -Ala-Leu-Ala-Leu- (SEQ ID NO:95), -Gly-Phe-Leu-Gly- (SEQ ID NO:96) y -Phe-Lys-.

En ciertas realizaciones, sin embargo, uno cualquiera o más de los enlazadores no peptídicos o peptídicos son opcionales. Por ejemplo, las secuencias de enlazador pueden no ser necesarias en una proteína de fusión donde el primer y segundo polipéptidos tienen regiones de aminoácidos en el extremo N y/o el extremo C no esenciales que se pueden usar para separar los dominios funcionales y evitar la interferencia estérica.

Secuencias de péptido señal. En ciertas realizaciones, una proteína de fusión de p97 comprende una o más secuencias de péptido señal (SP). En realizaciones específicas, la secuencia de péptido señal es una secuencia señal de extremo N, es decir, la porción más cercana al extremo N de la proteína de fusión.

Los ejemplos específicos de secuencias señal se proporcionan en la Tabla 4 a continuación. Ver también Kober et al., Biotechnology and Bioengineering. 110:1164-73, 2013.

Tabla 4: Secuencias de péptido señal (SP) ilustrativas		
Proteína	Secuencia señal	SEQ ID NO:
P97 humana	MRGPSGALWLLALRTVLG	39

Tabla 4: Secuencias de péptido señal (SP) ilustrativas		
Proteína	Secuencia señal	SEQ ID NO:
IDS humana	MPPPRTRGRLLWLGLVLSSVCVALG	40
Cadena pesada de Ig	MEWSWVFLFFLSVTTGVHS	149
Precursor de cadena ligera kappa de Ig	MDMRAPAGIFGFLLVLPGYRS	97
Preproteína de seroalbúmina	MKWVTFISLLFLFSSAYS	98
Cadena pesada de Ig	MDWTWRVFCLLAVTPGAHP	99
Cadena ligera de Ig	MAWSPLFLTLLTHCAGSWA	100
Preproteína de azurocidina	MTRLTVLALLAGLASSRA	101
Precursor de cistatina-S	MARPLCTLLLLMATLAGALA	102
Precursor de tripsinógeno 2	MRSLVFVLLIGAAFA	103
Bloqueador de canal de potasio	MSRLFVFILIALFLSAIDVMS	104
Conotoxina alfa	MGMRMMFIMFMLVVLATTVVS	105
Alfa-galactosidasa (mutante m3)	MRAFLFLTACISLPGVFG	106
Celulasa	MKFQSTLLLLAAAGSALA	107
Proteinasa aspártica de nepenthes en-1	MASSLYSFLALLSIVYIFVAPTHS	108
Quitinasa ácida	MKTHYSSAILPILTLFVFLSINPSHG	109
K28 prepro-toxina	MESVSSLFNIFSTIMVNYKSLVLALLSVSNLKYARG	110
Precursor de zigocina toxina destructora	MKAAQILTASIVSLLPIYTSA	111
Toxina del cólera	MIKLKFGVFFTVLLSSAYA	112

Por lo tanto, en algunas realizaciones, el péptido señal comprende, consiste o consiste esencialmente al menos una secuencia de la Tabla 4. En algunas realizaciones, el péptido señal comprende la SEQ ID NO:149.

- 5 En realizaciones específicas, la secuencia de péptido señal corresponde a la proteína más cercana al extremo N (p97 o IDS) de la proteína de fusión. Es decir, en algunas realizaciones, la secuencia de péptido señal del extremo N es la secuencia de péptido señal de P97 humana (SEQ ID NO:39) y la proteína de fusión de p97 comprende la estructura general: p97 SP-p97-IDS. En otras realizaciones, la secuencia de péptido señal del extremo N es la secuencia de péptido señal de IDS humana (SEQ ID NO:40) y la proteína de fusión de p97 comprende la estructura general: IDS SP-IDS-p97. Opcionalmente, la proteína de fusión puede comprender, además, una o más etiquetas de purificación y/o sitios de proteasa, por ejemplo, entre la secuencia señal del extremo N y las porciones p97/IDS de la proteína de fusión, según se describen en otra parte en la presente memoria. Aquí, el sitio de proteasa se coloca típicamente en el extremo C de la secuencia señal o etiqueta de purificación para que el tratamiento con la proteasa correspondiente elimine la secuencia señal, la etiqueta de purificación y la mayor parte o todo el sitio de proteasa del extremo N de la proteína de fusión.
- 10
- 15 Etiquetas de purificación. En algunas realizaciones, la proteína de fusión comprende una o más etiquetas de purificación o afinidad (TAG o TAGs). Los ejemplos no limitantes de etiquetas de purificación incluyen etiquetas de poli-histidina (p. ej., etiquetas 6xHis), avidina, etiquetas FLAG, etiquetas de glutatión S-transferasa (GST), etiquetas de proteína de unión a maltosa, proteína de unión a quitina (CBP, por sus siglas en inglés) y otras. También se incluyen etiquetas epitópicas, que se unen a anticuerpos con alta afinidad, cuyos ejemplos incluyen etiquetas V5, etiquetas Myc y etiquetas HA. En ejemplos específicos, la etiqueta de purificación es una etiqueta de polihistidina (H₅₋₁₀), por ejemplo, H₅, H₆, H₇, H₈, H₉ o H₁₀ (SEQ ID NOS:113-118).
- 20

Los ejemplos no limitantes de etiquetas de purificación se proporcionan en la Tabla 5 a continuación.

Tabla 5: Etiquetas de purificación ilustrativas (TAG)		
Nombre	Secuencia	SEQ ID NO:
5X-HIS	HHHHH	113
6X-HIS	HHHHHH	114
7X-HIS	HHHHHHH	115
8X-HIS	HHHHHHHH	116
9X-HIS	HHHHHHHHH	117
10X-HIS	HHHHHHHHHH	118
Etiqueta Avi	GLNDIFEAQKIEWHE	119
Etiqueta de calmodulina	KRRWKKNFIAVSAANRFKKISSSGAL	120
Etiqueta de poliglutamato	EEEEEE	121
Etiqueta FLAG	DYKDDDDK	122
Etiqueta HA	YPYDVPDYA	123
Etiqueta MYC	EQKLISEEDL	124
Etiqueta S	KETAAAKFERQHMS	125
Etiqueta SPB	MDEKTTGWRGGHVVEGLAGELEQLRARLEHHPQGQREP	126
Softag 1	SLAELLNAGLGGS	127
Softag 3	TQDPSRVG	128
Etiqueta V5	GKPIPNPLLGLDST	129
Etiqueta Xpress	DLYDDDDK	130

Por lo tanto, en ciertas realizaciones, la etiqueta de purificación comprende, consiste o consiste esencialmente en al menos una secuencia de la Tabla 5. En realizaciones específicas, la etiqueta comprende una etiqueta FLAG y una HIS, por ejemplo, etiqueta de 10X-HIS.

- 5 Sitios de proteasa (PS). En algunas realizaciones, la proteína de fusión comprende uno o más sitios de proteasa. Opcionalmente, el uno o más sitios de proteasa se posicionan en el extremo C de la secuencia de etiqueta de purificación y/o péptido señal (si una cualquiera o ambas están presentes) para que el tratamiento con la proteasa correspondiente elimine la secuencia señal, la etiqueta de purificación y/o la mayor parte o todo el sitio de proteasa del extremo N de la proteína de fusión.
- 10 En realizaciones específicas, por ejemplo, donde la proteína de fusión comprende un enlazador enzimáticamente escindible, el sitio de proteasa típicamente difiere del enlazador enzimáticamente escindible, para que el tratamiento con la proteasa elimine cualesquiera secuencias terminales (p. ej., secuencias de péptido señal, etiqueta de purificación) sin escindir el enlazador peptídico entre las secuencias de p97 e IDS.

Los ejemplos no limitantes de sitios de proteasa se proporcionan en la Tabla 6 a continuación.

Tabla 6: Sitios de proteasa (PS) ilustrativos.		
Proteasa	Secuencia	SEQ ID NO:
Trombina	LVPR▼GS	131
Enteropeptidasa	DDDDK▼	132
Factor Xa	I(E/D)GR▼	133
Enterocinasa	DDDDK▼	134
Proteasa de TEV	ENLYFQ▼G	135
Proteasa 3C	LEVLFG▼GP	136

Tabla 6: Sitios de proteasa (PS) ilustrativos.		
Proteasa	Secuencia	SEQ ID NO:
de HRV		
Proteasa de SUMO (Ulp1)	GSLQDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEA FAKRQ GKEMDSLTFLYDGIQADQTPEDLDMEDNDIIEAHREQIGG	137
▼ Denota el sitio de escisión		

Por lo tanto, en ciertas realizaciones, el sitio de proteasa comprende, consiste o consiste esencialmente en al menos una secuencia de la Tabla 6. En realizaciones específicas, el sitio de proteasa comprende el sitio de proteasa de TEV (SEQ ID NO:135).

- 5 Secuencias variantes. Ciertas realizaciones incluyen variantes de las secuencias polipeptídicas y polinucleotídicas de referencia descritas en la presente memoria, ya estén descritas por nombre o por referencia a un identificador de secuencia, incluidas secuencias de p97, secuencias de IDS, secuencias de enlazador, secuencias de péptido señal, etiquetas de purificación y sitios de proteasa (ver, p. ej., las Tablas 1-6 y la Lista de secuencias). Las secuencias naturales o más prevalentes de estos polipéptidos se conocen en la técnica y se pueden usar como comparación para las variantes y fragmentos descritos en la presente memoria.

Una secuencia "variante", según se usa el término en la presente memoria, se refiere a una secuencia polipeptídica o polinucleotídica que difiere de una secuencia de referencia descrita en la presente memoria en una o más sustituciones, eliminaciones (p. ej., truncamientos), adiciones e/o inserciones. Ciertas variantes, por lo tanto, incluyen fragmentos de una secuencia de referencia descrita en la presente memoria. Los polipéptidos variantes son biológicamente activos, es decir, continúan teniendo la actividad enzimática o de unión de un polipéptido de referencia. Dichas variantes pueden originarse a partir de, por ejemplo, polimorfismo genético y/o a partir de manipulación humana.

En muchos casos, una variante biológicamente activa contendrá una o más sustituciones conservadoras. Una "sustitución conservadora" es una en que un aminoácido se sustituye por otro aminoácido que tiene propiedades similares, de manera que un experto en la técnica de química de péptidos esperaría que la estructura secundaria y la naturaleza hidropática del polipéptido permanecieran sustancialmente incambiadas. Según se describió anteriormente, se pueden hacer modificaciones en la estructura de los polinucleótidos y polipéptidos de la presente invención y todavía obtener una molécula funcional que codifica un polipéptido variante o derivado con características deseables. Cuando se desea alterar la secuencia de aminoácidos de un polipéptido para crear una variante o porción equivalente, o incluso mejorada, de un polipéptido de la invención, un experto en la técnica típicamente cambiará uno o más de los codones de la secuencia de ADN codificante según la Tabla A a continuación.

Tabla A								
Aminoácidos			Codones					
Alanina	Ala	A	GCA	GCC	GCG	GCU		
Cisteína	Cys	C	UGC	UGU				
Ácido aspártico	Asp	D	GAC	GAU				
Ácido glutámico	Glu	E	GAA	GAG				
Fenilalanina	Phe	F	UUC	UUU				
Glicina	Gly	G	GGA	GGC	GGG	GGU		
Histidina	His	H	CAC	CAU				

Tabla A								
Aminoácidos			Codones					
Isoleucina	Ile	I	AUA	AUC	AUU			
Lisina	Lys	K	AAA	AAG				
Leucina	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU
Metionina	Met	M	AUG					
Asparagina	Asn	N	AAC	AAU				
Prolina	Pro	P	CCA	CCC	CCG	CCU		
Glutamina	Gln	Q	CAA	CAG				
Arginina	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU
Serina	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU
Treonina	Thr	T	ACA	ACC	ACG	ACU		
Valina	Val	V	GUA	GUC	GUG	GUU		
Triptófano	Trp	W	UGG					
Tirosina	Tyr	Y	UAC	UAU				

Por ejemplo, ciertos aminoácidos se pueden sustituir por otros aminoácidos en una estructura proteica sin pérdida apreciable de la capacidad de unión interactiva con estructuras tales como, por ejemplo, regiones de unión a antígeno de anticuerpos o sitios de unión en moléculas de sustrato. Dado que la capacidad interactiva y la naturaleza de una proteína son lo que definen la actividad biológica funcional de dicha proteína, ciertas sustituciones de secuencias de aminoácidos se pueden hacer en una secuencia proteica, y, evidentemente, su secuencia codificante de ADN y obtener, no obstante, una proteína con propiedades similares. Por lo tanto, se contempla que se pueden hacer diversos cambios en las secuencias peptídicas de las composiciones descritas o secuencias de ADN correspondientes que codifican dichos péptidos sin pérdida apreciable de su utilidad.

Al hacer dichos cambios, se puede tomar en cuenta el índice hidropático de los aminoácidos. La importancia del índice hidropático de los aminoácidos para conferir la función biológica interactiva a una proteína se entiende generalmente en la técnica (Kyte & Doolittle, 1982). Se acepta que el carácter hidropático relativo del aminoácido contribuye a la estructura secundaria de la proteína resultante que, a su vez, define la interacción de la proteína con otras moléculas, por ejemplo, enzimas, sustratos, receptores, ADN, anticuerpos, antígenos y similares. Se ha asignado a cada aminoácido un índice hidropático en función de sus características de hidrofobicidad y carga (Kyte & Doolittle, 1982). Estos valores son: isoleucina (+4,5); valina (+4,2); leucina (+3,8); fenilalanina (+2,8); cisteína (+2,5); metionina (+1,9); alanina (+1,8); glicina (-0,4); treonina (-0,7); serina (-0,8); triptófano (-0,9); tirosina (-1,3); prolina (-1,6); histidina (-3,2); glutamato (-3,5); glutamina (-3,5); aspartato (-3,5); asparagina (-3,5); lisina (-3,9); y arginina (-4,5). Se sabe en la técnica que ciertos aminoácidos se pueden sustituir por otros aminoácidos que tiene un índice o puntuación hidropático similar y todavía dar como resultado una proteína con actividad biológica similar, es decir, todavía obtener una proteína biológica funcionalmente equivalente. Al hacer dichos cambios, se prefiere la sustitución de aminoácidos cuyos índices hidropáticos están dentro de ± 2 , aquellos dentro de ± 1 se prefieren particularmente y aquellos dentro de $\pm 0,5$ se prefieren incluso más particularmente.

También se entiende en la técnica que la sustitución de aminoácidos similares se puede hacer de manera eficaz en función de la hidrofiliidad. La patente estadounidense 4 554 101 indica que la mayor hidrofiliidad promedio local de una proteína, según se rige por la hidrofiliidad de sus aminoácidos adyacentes, se correlaciona con una propiedad biológica de la proteína. Según se detalla en la patente estadounidense 4 554 101, se han asignado los siguientes

valores de hidrofiliidad a residuos aminoacídicos: arginina (+3,0); lisina (+3,0); aspartato (+3,0 \pm 1); glutamato (+3,0 \pm 1); serina (+0,3); asparagina (+0,2); glutamina (+0,2); glicina (0); treonina (-0,4); prolina (-0,5 \pm 1); alanina (-0,5); histidina (-0,5); cisteína (-1,0); metionina (-1,3); valina (-1,5); leucina (-1,8); isoleucina (-1,8); tirosina (-2,3); fenilalanina (-2,5); triptófano (-3,4). Se entiende que un aminoácido se puede sustituir por otro que tiene un valor de hidrofiliidad similar y todavía obtener una proteína biológicamente equivalente y, en particular, inmunológicamente equivalente. En dichos cambios, se prefiere la sustitución de aminoácidos cuyos valores de hidrofiliidad están dentro de ± 2 , aquellos dentro de ± 1 se prefieren particularmente y aquellos dentro de $\pm 0,5$ se prefieren incluso más particularmente.

Según se señaló anteriormente, las sustituciones de aminoácidos se basan generalmente, por lo tanto, en la similitud relativa de los sustituyentes de cadena lateral de aminoácido, por ejemplo, su hidrofobicidad, hidrofiliidad, carga, tamaño y similares. Las sustituciones ilustrativas que toman varias de las características anteriores en consideración son conocidas para los expertos en la técnica e incluyen: arginina y lisina; glutamato y aspartato; serina y treonina; glutamina y asparagina; y valina, leucina e isoleucina.

Las sustituciones de aminoácidos se pueden hacer, adicionalmente, en función de la similitud en la polaridad, carga, solubilidad, hidrofobicidad, hidrofiliidad y/o la naturaleza anfipática de los residuos. Por ejemplo, los aminoácidos con carga negativa incluyen ácido aspártico y ácido glutámico; los aminoácidos con carga positiva incluyen lisina y arginina; y los aminoácidos con grupos de extremo polar no cargados que tienen valores de hidrofiliidad similares incluyen leucina, isoleucina y valina; glicina y alanina; asparagina y glutamina; y serina, treonina, fenilalanina y tirosina. Otros grupos de aminoácidos que pueden representar cambios conservadores incluyen: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; y (5) phe, tyr, trp, his.

Una variante puede también, o alternativamente, contener cambios no conservadores. En una realización preferida, los polipéptidos variantes difieren de una secuencia natural o de referencia por la sustitución, eliminación o adición de menos de aproximadamente 10, 9, 8, 7, 6, 5, 4, 3, 2 aminoácidos, o incluso 1 aminoácido. Las variantes también (o alternativamente) se pueden modificar, por ejemplo, por la eliminación o adición de aminoácidos que influyen mínimamente en la inmunogenicidad, estructura secundaria, actividad enzimática y/o naturaleza hidropática del polipéptido.

En ciertas realizaciones, una secuencia polipeptídica tiene aproximadamente, al menos aproximadamente o hasta aproximadamente 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000 o más aminoácidos contiguos de longitud, incluidos todos los números enteros entremedio, y que puede comprender toda o una porción de una secuencia de referencia (*ver, p. ej.,* Lista de secuencias).

En otras realizaciones específicas, una secuencia polipeptídica consiste en aproximadamente o no más de aproximadamente 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000 o más aminoácidos contiguos, incluidos todos los números enteros entremedio, y que puede comprender toda o una porción de una secuencia de referencia (*ver, p. ej.,* Lista de secuencias).

En todavía otras realizaciones específicas, una secuencia polipeptídica tiene aproximadamente 10-1000, 10-900, 10-800, 10-700, 10-600, 10-500, 10-400, 10-300, 10-200, 10-100, 10-50, 10-40, 10-30, 10-20, 20-1000, 20-900, 20-800, 20-700, 20-600, 20-500, 20-400, 20-300, 20-200, 20-100, 20-50, 20-40, 20-30, 50-1000, 50-900, 50-800, 50-700, 50-600, 50-500, 50-400, 50-300, 50-200, 50-100, 100-1000, 100-900, 100-800, 100-700, 100-600, 100-500, 100-400, 100-300, 100-200, 200-1000, 200-900, 200-800, 200-700, 200-600, 200-500, 200-400 o 200-300 aminoácidos contiguos, incluidos todos los intervalos entremedio, y comprende toda o una porción de una secuencia de referencia. En ciertas realizaciones, la región del extremo C o el extremo N de cualquier polipéptido de referencia se puede truncar en aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750 o 800 o más aminoácidos, o en aproximadamente 10-50, 20-50, 50-100, 100-150, 150-200, 200-250, 250-300, 300-350, 350-400, 400-450, 450-500, 500-550, 550-600, 600-650, 650-700, 700-750, 750-800 o más aminoácidos, incluidos todos los números enteros e intervalos entremedio (*p. ej.,* 101, 102, 103, 104, 105), siempre que el polipéptido truncado conserve las propiedades de unión y/o la actividad del polipéptido de referencia. Típicamente, el fragmento biológicamente activo tiene no menos de aproximadamente 1 %, aproximadamente 5 %, aproximadamente 10 %, aproximadamente 25 %, o aproximadamente 50 % de una actividad del polipéptido de referencia biológicamente activo del que se deriva.

En general, las variantes exhibirán al menos aproximadamente 30 %, 40 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % de similitud o identidad de secuencia u homología de secuencia con respecto a una secuencia polipeptídica de referencia. Asimismo, se contemplan secuencias que difieren de las secuencias naturales u originales por la adición (*p. ej.*, adición en el extremo C, adición en el extremo N, ambas), eliminación, truncamiento, inserción o sustitución (*p. ej.*, sustitución conservadora) de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 o 100 aminoácidos (incluidos todos los números enteros e intervalos entremedio), pero que conservan las propiedades o actividades de una secuencia polipeptídica original o de referencia.

En algunas realizaciones, los polipéptidos variantes difieren con respecto a la secuencia de referencia en al menos uno, pero menos de 50, 40, 30, 20, 15, 10, 8, 6, 5, 4, 3 o 2 residuos aminoácidos. En otras realizaciones, los polipéptidos variantes difieren con respecto a la secuencia de referencia en al menos 1 %, pero menos de 20 %, 15 %, 10 % o 5 % de los residuos. (Si esta comparación requiere alineación, las secuencias se deben alinear para la máxima similitud. Las secuencias "en bucle" a partir de eliminaciones o inserciones, o emparejamientos incorrectos se consideran diferencias).

Los cálculos de similitud de secuencia o identidad de secuencia entre secuencias (los términos se usan de manera intercambiable en la presente memoria) se llevan a cabo de la siguiente manera. Para determinar el porcentaje de identidad de dos secuencias de aminoácidos, o de dos secuencias de ácido nucleico, las secuencias se alinean para una comparación óptima (*p. ej.*, se pueden introducir espacios en una o ambas de una primera y una segunda secuencia de aminoácidos o ácido nucleico para la alineación óptima y se pueden ignorar las secuencias no homólogas a los fines de la comparación). En ciertas realizaciones, la longitud de una secuencia de referencia alineada para comparación es de al menos 30 %, preferiblemente, al menos 40 %, más preferiblemente, al menos 50 %, 60 %, e incluso más preferiblemente, al menos 70 %, 80 %, 90 %, 100 % de la longitud de la secuencia de referencia. A continuación, se comparan los residuos aminoácidos o nucleótidos en las posiciones de aminoácidos o posiciones de nucleótidos correspondientes. Cuando una posición en la primera secuencia está ocupada por el mismo residuo aminoácido o nucleótido que la posición correspondiente en la segunda secuencia, las moléculas son idénticas en dicha posición.

El porcentaje de identidad entre las dos secuencias es una función de la cantidad de posiciones idénticas compartidas por las secuencias, tomando en cuenta la cantidad de espacios y la longitud de cada espacio, que se deben introducir para la alineación óptima de las dos secuencias.

La comparación de secuencias y la determinación el porcentaje de identidad entre dos secuencias se pueden lograr mediante el uso de un algoritmo matemático. En una realización preferida, el porcentaje de identidad entre dos secuencias de aminoácidos se determina mediante el uso del algoritmo de Needleman and Wunsch, (*J. %mol. Biol.* 48: 444-453, 1970) que se ha incorporado en el programa GAP en el paquete informático GCG, mediante el uso de una matriz Blossum 62 o una matriz PAM250 y un valor de espacio de 16, 14, 12, 10, 8, 6 o 4 y un valor de longitud de 1, 2, 3, 4, 5 o 6. En todavía otra realización preferida, el porcentaje de identidad entre dos secuencias de nucleótidos se determina mediante el uso del programa GAP en el paquete informático GCG, mediante el uso de una matriz NWSgapdna.CMP y un valor de espacio de 40, 50, 60, 70 u 80 y un valor de longitud de 1, 2, 3, 4, 5 o 6. Un conjunto de parámetros particularmente preferido (y el que se debe usar a menos que se especifique de cualquier otra manera) son una matriz de puntuación Blossum 62 con una penalización por espacio de 12, una penalización extendida de espacio de 4 y una penalización de espacio de desplazamiento de marco de 5.

El porcentaje de identidad entre dos secuencias de aminoácidos o nucleótidos se puede determinar mediante el uso del algoritmo de E. Meyers y W. Miller (*Cabios.* 4:11-17, 1989) que se ha incorporado en el programa ALIGN (versión 2.0), mediante el uso de una tabla de valor de residuo PAM120, una penalización por longitud de espacio de 12 y una penalización por espacio de 4.

Las secuencias de ácido nucleico y proteicas descritas en la presente memoria se pueden usar como una "secuencia de consulta" para llevar a cabo una búsqueda en bases de datos públicas para, por ejemplo, identificar otros miembros de la familia o secuencias relacionadas. Dichas búsquedas se pueden llevar a cabo mediante el uso de los programas NBLAST y XBLAST (versión 2.0) de Altschul, et al., (1990, *J. Mol. Biol.* 215: 403-10). Las búsquedas de nucleótidos por BLAST se pueden llevar a cabo con el programa NBLAST, puntuación = 100, longitud de palabra = 12 para obtener secuencias nucleotídicas homólogas a las moléculas de ácido nucleico de la invención. Las búsquedas de proteína por BLAST se pueden llevar a cabo con el programa XBLAST, puntuación = 50, longitud de palabra = 3 para obtener secuencias de aminoácidos homólogas a las moléculas proteicas de la invención. Para obtener alineaciones con espacios para comparación, se puede usar Gapped BLAST según se describe en Altschul et al., (*Nucleic Acids Res.* 25: 3389-3402, 1997). Cuando se usan los programas BLAST y Gapped BLAST, se pueden usar los parámetros predeterminados de los respectivos programas (*p. ej.*, XBLAST y NBLAST).

En una realización, según se indicó anteriormente, los polinucleótidos y/o polipéptidos se pueden evaluar mediante el uso de la herramienta de alineación BLAST. Una alineación local consiste simplemente en un par de segmentos de secuencia, uno de cada una de las secuencias que se van a comparar. Una modificación de los algoritmos de Smith-Waterman o Sellers encontrará todos los pares de segmentos cuyas puntuaciones no se pueden mejorar por

medio de extensión o recorte, llamados pares de segmentos con alta puntuación (HSP, por sus siglas en inglés). Los resultados de las alineaciones por BLAST incluyen mediciones estadísticas para indicar la probabilidad de que la puntuación de BLAST se pueda esperar solo por casualidad.

5 La puntuación bruta, S , se calcula a partir de la cantidad de espacios y sustituciones asociados a cada secuencia alineada en donde las puntuaciones de similitud más alta indican una alineación más significativa. Las puntuaciones de sustitución se proporcionan mediante una tabla de búsqueda (ver PAM, BLOSUM).

10 Las puntuaciones de espacio se calculan típicamente como la suma G , la penalización de apertura de espacio y L , la penalización de extensión de espacio. Para un espacio de longitud n , el costo del espacio sería $G + Ln$. La elección de los costos de espacio, G y L es empírica, pero es habitual elegir un valor alto para G (10-15), *p. ej.*, 11, y un valor bajo para L (1-2) *p. ej.*, 1.

15 La puntuación de trozo, S' , se deriva de la puntuación de alineación bruta S en la que se han tomado en cuenta las propiedades estadísticas del sistema de puntuación usado. Las puntuaciones de trozo se normalizan con respecto al sistema de puntuación, por lo tanto, se pueden usar para comparar puntuaciones de alineación de diferentes búsquedas. Los términos "puntuación de trozo" y "puntuación de similitud" se usan de manera intercambiable. La puntuación de trozo proporciona una indicación de cuán buena es la alineación; cuanto mayor es la puntuación, mejor es la alineación.

20 El valor E , o valor esperado, describe la probabilidad de que una secuencia con una puntuación similar aparezca en la base de datos por casualidad. Es una predicción de la cantidad de diferentes alineaciones con puntuaciones equivalentes a o mejores que S que se espera que aparezcan en una búsqueda de base de datos por casualidad. Cuanto menor es el valor E , más significativa es la alineación. Por ejemplo, una alineación que tiene un valor E de e^{-117} significa que es muy improbable que una secuencia con una puntuación similar aparezca simplemente por casualidad. Además, la puntuación esperada por alinear un par aleatorio de aminoácidos debe ser negativa, de lo contrario las alineaciones largas tenderían a tener una puntuación alta independientemente de si los segmentos alineados estuviesen relacionados. Además, el algoritmo BLAST usa una matriz de sustitución, nucleótido o aminoácido adecuada y para alineaciones con espacios usa penalizaciones de creación y extensión de espacios. Por ejemplo, la alineación y la comparación de secuencias polipeptídicas por BLAST típicamente se llevan a cabo mediante el uso de la matriz BLOSUM62, una penalización de existencia de espacio de 11 y una penalización de extensión de espacio de 1.

30 En una realización, las puntuaciones de similitud de secuencia se informan a partir de análisis por BLAST llevados a cabo mediante el uso de la matriz BLOSUM62, una penalización de existencia de espacio de 11 y una penalización de extensión de espacio de 1.

35 En una realización específica, las puntuaciones de identidad/similitud de secuencia proporcionadas en la presente memoria se refieren al valor obtenido mediante el uso de GAP Versión 10 (GCG, Accelrys, San Diego, Calif.) con los siguientes parámetros: % de identidad y % de similitud para una secuencia nucleotídica mediante el uso de un valor de GAP de 50 y un valor de Longitud de 3, y la matriz de puntuación *nwsgapdna.cmp*; % de identidad y % de similitud para una secuencia de aminoácidos mediante el uso de un valor de GAP de 8 y un valor de Longitud de 2, y la matriz de puntuación BLOSUM62 (Henikoff y Henikoff, PNAS USA. 89:10915-10919, 1992). GAP usa el algoritmo de Needleman and Wunsch (J Mol Biol. 48:443-453, 1970) para encontrar la alineación de dos secuencias completas que maximiza la cantidad de coincidencias y minimiza la cantidad de espacios.

40 En una realización específica, el polipéptido variante comprende una secuencia de aminoácidos que se puede alinear de manera óptima con una secuencia polipeptídica de referencia (ver, *p. ej.*, la Lista de secuencias) para generar puntuaciones de trozo o puntuaciones de similitud de secuencia por BLAST de al menos aproximadamente 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000 o más, incluidos todos los números enteros e intervalos entremedio, en donde la alineación BLAST usó la matriz BLOSUM62, una penalización por existencia de espacio de 11 y una penalización de extensión de espacio de 1.

50 Según se indicó anteriormente, un polipéptido de referencia se puede alterar de varias maneras que incluyen sustituciones, eliminaciones, truncamientos, adiciones e inserciones de aminoácidos. Los métodos para dichas manipulaciones se conocen generalmente en la técnica. Por ejemplo, las variantes de secuencia de aminoácidos de un polipéptido de referencia se pueden preparar mediante mutaciones en el ADN. Los métodos para mutagénesis y alteraciones de la secuencia nucleotídica se conocen en la técnica. Ver, por ejemplo, Kunkel (PNAS USA. 82: 488-492, 1985); Kunkel et al., (Methods in Enzymol. 154: 367-382, 1987), la patente estadounidense núm. 4 873 192, Watson, J. D. et al., ("Molecular Biology of the Gene," cuarta edición, Benjamin/Cummings, Menlo Park, Calif., 1987) y las referencias mencionadas en estas. Se puede encontrar orientación para sustituciones de aminoácidos adecuadas que no afectan la actividad biológica de la proteína de interés en el modelo de Dayhoff et al., (1978) Atlas of Protein Sequence and Structure (Natl. Biomed. Res. Found., Washington, D.C.).

Los métodos para barrer productos génicos en bibliotecas combinatorias elaboradas mediante dichas modificaciones y para barrer bibliotecas de ADNc para encontrar productos génicos que tienen una propiedad seleccionada se conocen en la técnica. Dichos métodos son adaptables para el barrido rápido de las bibliotecas génicas generadas mediante mutagénesis combinatoria de polipéptidos de referencia. Como ejemplo, la mutagénesis de ensamble recursivo (REM, por sus siglas en inglés), una técnica que potencia la frecuencia de mutantes funcionales en las bibliotecas, se puede usar en combinación con los ensayos de barrido para identificar variantes de polipéptidos (Arkin and Yourvan, PNAS USA 89: 7811-7815, 1992; Delgrave et al., Protein Engineering. 6: 327-331, 1993).

Polinucleótidos, células Hospedantes y métodos de producción. Ciertas realizaciones se refieren a polinucleótidos que codifican las proteínas de fusión descritas en la presente memoria, y vectores que comprenden dichos polinucleótidos, por ejemplo, donde los polinucleótidos están funcionalmente enlazados a uno o más elementos reguladores. También se incluyen células hospedantes recombinantes que comprenden dichos polinucleótidos, vectores, proteínas de fusión y métodos de producción recombinante de los anteriores.

Las proteínas de fusión se pueden preparar mediante el uso de técnicas estándar. Preferiblemente, sin embargo, una proteína de fusión se expresa como una proteína recombinante en un sistema de expresión, según se describe en la presente memoria y se conoce en la técnica. Las proteínas de fusión pueden contener una o múltiples copias de una secuencia de p97 y una o múltiples copias de una secuencia de IDS, presentes en cualquier disposición deseada.

Los polinucleótidos y polinucleótidos de fusión pueden contener una o múltiples copias de un ácido nucleico que codifica una secuencia polipeptídica de p97 y/o una o múltiples copias de un ácido nucleico que codifica una secuencia de IDS.

Para las proteínas de fusión, los componentes de secuencias de ADN que codifican la secuencia polipeptídica de p97, la secuencia de IDS de interés y, opcionalmente, un enlazador peptídico se puede ensamblar por separado y a continuación ligarse en un vector de expresión adecuado. El extremo hacia 3' de la secuencia de ADN que codifica un componente polipeptídico se puede ligar, con o sin un enlazador peptídico, al extremo hacia 5' de una secuencia de ADN que codifica el(los) otro(s) componente(s) polipeptídico(s) de manera que los marcos de lectura de las secuencias estén dentro del mismo marco. Las secuencias de ADN ligadas se enlazan funcionalmente a elementos reguladores de la transcripción y/o traducción adecuados. Los elementos reguladores responsables de la expresión del ADN habitualmente solo se ubican hacia 5' con respecto a la secuencia de ADN que codifica los primeros polipéptidos. De manera similar, los codones de terminación necesarios para detener las señales de terminación de traducción y transcripción están solo presentes hacia 3' con respecto a la secuencia de ADN que codifica el polipéptido más cercano al extremo C. Esto permite la traducción en un único polipéptido de fusión que conserva la actividad biológica de ambos polipéptidos componentes.

Se pueden aplicar técnicas similares, principalmente la disposición de los elementos reguladores tales como promotores, codones de terminación y señales de terminación de la transcripción, para la producción recombinante de proteínas que no tienen fusión.

Se pueden elegir o construir vectores adecuados que contengan secuencias reguladoras adecuadas, incluidas secuencias promotoras, secuencias de terminación, secuencias de poliadenilación, secuencias potenciadoras, genes marcadores y otras secuencias, según sea adecuado. Los vectores pueden ser plásmidos, fagos, p. ej., víricos o fagémidos, según sea adecuado. Para obtener detalles adicionales ver, por ejemplo, Molecular Cloning: a Laboratory Manual: 2a edición, Sambrook et al., 1989, Cold Spring Harbor Laboratory Press. Muchas técnicas conocidas y protocolos para la manipulación de ácido nucleico, por ejemplo, en la preparación de construcciones de ácido nucleico, mutagénesis, secuenciación, introducción de ADN en células y expresión génica, y análisis de proteínas, se describen en detalle en Current Protocols in Molecular Biology, segunda edición, Ausubel et al. eds., John Wiley & Sons, 1992, o actualizaciones posteriores de este.

Como lo entenderán los expertos en la técnica, puede ser ventajoso en algunos casos producir secuencias de nucleótidos que codifican polipéptidos que poseen codones de origen no natural. Por ejemplo, se pueden seleccionar los codones preferidos por un hospedante procariota o eucariota particular para aumentar la tasa de expresión proteica o para producir un transcrito de ARN recombinante que tiene propiedades deseables, tal como una semivida que es más extensa que la de un transcrito generado a partir de la secuencia de origen natural. Dichos polinucleótidos se denominan comúnmente "optimizados por codón". Cualesquiera de los polinucleótidos descritos en la presente memoria se pueden usar en una forma optimizada por codón. En ciertas realizaciones, un polinucleótido se puede optimizar por codón para su uso en bacterias específicas tales como *E. coli* o levadura tal como *S. cerevisiae* (ver, p. ej., Burgess-Brown et al., Protein Expr Purif. 59:94-102, 2008).

Las secuencias polinucleotídicas ilustrativas se proporcionan en la Tabla 7 a continuación.

Tabla 7. Secuencias polinucleotídicas ilustrativas		
Nombre	Secuencia polinucleotídica	SEQ ID NO:
I2S-MTf	<p>ATGGAATGGAGCTGGGTCTTTCTCTTCTTCTGTGTCAGTAACGACTGGTGTCCAC TCCGACTACAAGGACGACGACGACAAAGAGCAGAAGCTGATCTCCGAAGAGGAC CTGCACCACCATCATCACCATCACCACCATCACGGAGGCGGTGGAGAGAACCTG TACTTTTCAGGGCTCGGAAACTCAGGCCAACTCCACCACAGATGCACTCAACGTG CTGCTGATCATCGTAGATGACCTCCGACCTTCTCTGGGCTGTTACGGCGACAAG CTAGTACGGAGCCCCAACATCGACCAGCTCGCATCGCACTCTCTCCTATTCCAG AACGCATTGCGCCAGCAGGCTGTCTGTGCTCCCTCCCGAGTGTCTTCTCTCACG GGTCGGAGACCCGATACCACGAGGTTATATGACTTCAACTCATACTGGGCGGTG CATGCCGGTAACCTTTTCTACTATACCCCAGTATTTTAAAGAAAATGGCTATGTT ACAATGTCCGTTGGCAAGGTATTTTCATCCTGGTATTAGCAGCAACCACACAGAT GACTCTCCGTATAGCTGGTCATTCCCACCATAACCACCCCTCCAGCGAAAAGTAC GAAAACACAAAGACTTGCCGGGGGCCAGATGGCGAACTGCACGCAAACTCTGCTG TGCCCTGTAGATGTCTTGGACGTGCCCGAAGGTACTCTGCCCGACAAACAGTCC ACAGAACAGGCAATCCAACCTCCTTGAAAAGATGAAAACGAGCGCGTCCCCCTTC TTCCTCGCCGTGGGCTACCACAAGCCCCACATCCCGTTTAGATACCCCAAGGAA TTTCAGAACTGTACCCCTTGGAACATCACTCTCGCGCCCGACCCCGAAGTG CCAGACGGACTCCCTCCTGTTGCCTACAACCCTTGGATGGACATCAGACAACGT GAAGATGTGCAGGCCCTGAACATCTCAGTGCCTTACGGCCCCATTCCAGTTGAC TTCCAGAGGAAGATTTCGGCAGTCTTACTTCGCCTCCGTTAGTTACCTGGACACC CAAGTGGGTAGACTCCTGAGCGCCTTGGACGATCTCCAGCTCGCAAACAGCACC ATCATTGCCTTCACCAGCGACCATGGTTGGGCGCTGGGTGAACATGGAGAATGG GCTAAATATTCAAATTTTCGACGTTGCGACCCACGTCCCATTGATCTTCTACGTG CCTGGACGAACAGCCTCCTTGCCCTGAAGCCGGGGAAAAGTTGTTTCCATATCTG GACCCTTTCGATTCTGCGAGCCAACTCATGGAACCTGGGCGACAGAGCATGGAC CTGGTGGAACTGGTCAGTTTATTTCCAACCCCTGGCAGGCCTTGCAGGCCTCCAA GTTCCACCTCGGTGTCCCGTTCCCTCATTTCCACGTGCAACTCTGTCTCGGAAGGT AAAAACCTCCTCAAGCATTTTTCGTTTTCGGGACCTCGAAGAAGACCCATACCTG CCAGGGAATCCAAGGGAAGTATTGCTTACAGCCAGTACCCTAGACCTAGCGAC ATCCACAGTGGAACAGCGACAAGCCCTCCCTCAAGGACATTAATCATGGGT TATAGTATCCGGACATATGACTACAGGTATACCGTGTGGGTGGGTTCACCCCA GACGAATTTCTCGCCAATTTCTCCGACATCCACGCGGGCGAACTGTATTTTCGTT GATTCCGATCCACTGCAAGATCATAATATGTACAACGATAGTCAAGGGGGTGAC CTCTTCCAGTTGCTAATGCCAGAAGCCGCCGCGAAAAGAAGCCGCCGCAAAAGAA GCCGCTGCCAAAGGCATGGAAGTGCCTTGGTGCGCCACCTCTGACCCCGAGCAG CACAAGTGCAGCAACATGTCCGAGGCCCTCAGAGAGGCCGCGCATCCAGCCTTCT CTGCTGTGTGTGCGGGGCACCTCTGCCGACCATTGCGTGCAGCTGATCGCCGCC CAGGAAGCCGACGCTATCACACTGGATGGCGGCGCTATCTACGAGGCTGGCAAA GAGCACGGCTGAAGCCCGTCTGTGGGCGAGGTGTACGATCAGGAAGTGGGCACC TCCTACTACGCCGTGGCTGTCTGTGCGGAGATCCTCCACGTGACCATCGACACC CTGAAGGGCGTGAAGTCTGCCACACCGGCATCAACAGAACCCTGGGCTGGAAC GTGCCCGTGGGCTACCTGGTGAATCCGGCAGACTGTCCGTGATGGGCTGCGAC GTGCTGAAGGCCGTGTCGATTACTTCCGCGGCTCTTGTGTGCCTGGCGCTGGC GAGACATCCTACTCCGAGTCCCTGTGCAGACTGTGCAGGGGCGACTCTTCTGGC GAGGGCGTGTGCGACAAGTCCCCCTTGGAACGGTACTACGACTACTCCGGCGCC</p>	143

Tabla 7. Secuencias polinucleotídicas ilustrativas		
Nombre	Secuencia polinucleotídica	SEQ ID NO:
	<p> TTCAGATGCCTGGCTGAAGGTGCTGGCGACGTGGCCTTCGTGAAGCACTCCACC GTGCTGGAAAACACCGACGGCAAGACCCGCTTCTTGGGGCCAGGCACTGCTG TCCCAGGACTTCGAGCTGCTGTGCCGGGATGGCTCCAGAGCCGATGTGACAGAG TGGCGGCAGTGCCACCTGGCCAGAGTGCTGCTCATGCTGTGGTTCGTGCGCGCC GATACAGATGGCGGCCTGATCTTCCGGCTGCTGAACGAGGGCCAGCGGCTGTTC TCTCACGAGGGCTCCAGCTTCCAGATGTTCTCCAGCGAGGCCTACGGCCAGAAG GACCTGCTGTTCAAGGACTCCACCTCCGAGCTGGTGCCTATCGCCACCCAGACC TATGAGGCTTGGCTGGGCCACGAGTACCTGCACGCTATGAAGGGACTGCTGTGC GACCCCAACCGGCTGCCTCCTTATCTGAGGTGGTGCCTGCTGTCCACCCCGAG ATCCAGAAATGCGGCGATATGGCCGTGGCCTTTCGGCGGCAGAGACTGAAGCCT GAGATCCAGTGCGTGTCCGCCAAGAGCCCTCAGCACTGCATGGAACGGATCCAG GCCGAACAGGTGGACGCCGTGACACTGTCCGGCGAGGATATCTACACCGCCGGA AAGACCTACGGCCTGGTGCCAGCTGCTGGCGAGCATTACGCCCCTGAGGACTCC TCCAACAGCTACTACGTGGTGGCAGTCTGTGCGCCGGGACTCCTCTACGCCTTT ACCTTGATGAGCTGCGGGGCAAGAGAAGCTGTACGCGCGGCTTTGGAAGCCCT GCCGGATGGGATGTGCCTGTGGGCGCTCTGATCCAGCGGGGCTTCATCAGACCC AAGGACTGTGATGTGCTGACCGCCGTGTCTGAGTTCTTCAACGCCCTCCTGTGTG CCCGTGAACAACCCCAAGAACTACCCCTCCAGCCTGTGCGCCCTGTGTGTGGGA GATGAGCAGGGCCGGAACAAATGCGTGGGCAACTCCCAGGAAAGATATTACGGC TACAGAGGCGCCTTCCGGTGTCTGGTGGAAAACGCCGGGGATGTGGCTTTTGTG CGGCACACCACCGTGTTCGACAACACCAATGGCCACAACCTCCGAGCCTTGGGCC GCTGAGCTGAGATCCGAGGATTACGAACGTGTGTGTCCCAACGGCGCCAGGGCT GAGGTGTCCAGTTTGCCGCCCTGTAACCTGGCCAGATCCCTCCCCACGCTGTG ATGGTGCACCCGACACCAACATCTTACCGTGTACGGCCTGCTGGACAAGGCC CAGGATCTGTTCGGCGACGACCACAACAAGAACGGGTCAAGATGTTCGACTCC AGCAACTACCACGGACAGGATCTGCTGTTTAAAGATGCCACCGTGCGGGCCGTG CCAGTGGGCGAAAAGACCACCTACAGAGGATGGCTGGGACTGGACTACGTGGCC GCCCTGGAAGGCATGTCTCCAGCAGTGTTCCTGA </p>	

Tabla 7. Secuencias polinucleotídicas ilustrativas		
Nombre	Secuencia polinucleotídica	SEQ ID NO:
MTf-I2S	ATGGAATGGAGCTGGGTCTTTCTCTTCTTCTGTCAGTAACGACTGGTGTCCAC TCCGACTACAAGGACGACGACGACAAAGAGCAGAAGCTGATCTCCGAAGAGGAC CTGCACCACCATCATCACCATCACCACCATCACGGAGGCGGTGGAGAGAACCTG TACTTTCAGGGCGGCATGGAAGTGCCTTGGTGCGCCACCTCTGACCCCCGAGCAG CACAAGTGCGGCAACATGTCCGAGGCCCTTCAGAGAGGCCGGCATCCAGCCTTCT CTGCTGTGTGTGCGGGGCACCTCTGCCGACCATTGCGTGCAGCTGATCGCCGCC CAGGAAGCCGACGCTATCACACTGGATGGCGGCGCTATCTACGAGGCTGGCAAA GAGCACGGCCTGAAGCCCGTCTGTTGGGCGAGGTGTACGATCAGGAAGTGGGCACC TCCTACTACGCCGTGGCTGTCTGCGGAGATCCTCCACGCTGACCATCGACACC CTGAAGGGCGTGAAGTCTGCCACACCGGCATCAACAGAACCGTGGGCTGGAAC GTGCCCCGTGGGCTACCTGGTGAATCCGGCAGACTGTCCGTGATGGGCTGCGAC GTGCTGAAGGCCGTGTCCGATTACTTCGGCGGCTCTTGTGTGCCGTGGCGCTGGC GAGACATCCTACTCCGAGTCCCTGTGCAGACTGTGCAGGGGCGACTCTTCTGGC GAGGGCGTGTGCGACAAGTCCCTCTGGAACGGTACTACGACTACTCCGGCGCC TTCAGATGCCTGGCTGAAGGTGCTGGCGACGTGGCCTTCGTGAAGCACTCCACC GTGCTGGAACACCGACGGCAAGACCTGCCTTCTTGGGGCCAGGCACTGCTG TCCCAGGACTTCGAGCTGCTGTGCGGGATGGCTCCAGAGCCGATGTGACAGAG TGGCGGCAGTGCCACCTGGCCAGAGTGCCTGCTCATGCTGTGGTCTGCGCGCC GATACAGATGGCGGCTGATCTTCCGGCTGCTGAACGAGGGCCAGCGGCTGTTC TCTCACGAGGGCTCCAGCTTCCAGATGTTCTCCAGCGAGGCCTACGGCCAGAAG GACCTGCTGTTCAAGGACTCCACCTCCGAGCTGGTGCCTATCGCCACCCAGACC TATGAGGCTTGGCTGGGCCACGAGTACCTGCACGCTATGAAGGGACTGCTGTGC GACCCCAACCGGCTGCCTCCTTATCTGAGGTGGTGGCTGCTGTCCACCCCCGAG ATCCAGAAATGCGGCGATATGGCCGTGGCCTTTCGGCGGCAGAGACTGAAGCCT GAGATCCAGTGCCTGTCCGCCAAGAGCCCTCAGCACTGCATGGAACGGATCCAG GCCGAACAGGTGGACGCCGTGACACTGTCCGGCGAGGATATCTACACCGCCGGA AAGACCTACGGCTGGTGCCAGCTGCTGGCGAGCATTACGCCCCTGAGGACTCC TCCAACAGCTACTACGTGGTGGCAGTCGTGCGCCGGGACTCCTCTCACGCCTTT ACCCTGGATGAGCTGCGGGGCAAGAGAAGCTGTCACGCCGGCTTTGGAAGCCCT GCCGGATGGGATGTGCTGTGGGCGCTCTGATCCAGCGGGGCTTCATCAGACCC AAGGACTGTGATGTGCTGACCGCCGTGTCTGAGTTCTTCAACGCCCTCCTGTGTG CCCGTGAACAACCCCAAGAACTACCCCTCCAGCCTGTGCGCCCTGTGTGTGGGA GATGAGCAGGGCCGAACAAATGCGTGGGCAACTCCCAGGAAAGATATTACGGC	144

Tabla 7. Secuencias polinucleotídicas ilustrativas		
Nombre	Secuencia polinucleotídica	SEQ ID NO:
	<p>TACAGAGGCGCCTTCCGGTGTCTGGTGAAAAACGCCGGGGATGTGGCTTTTGTG CGGCACACCACCGTGTTCGACAACACCAATGGCCACAACCCGAGCCTTGGGCC GCTGAGCTGAGATCCGAGGATTACGAACCTGCTGTGTCCCAACGGCGCCAGGGCT GAGGTGTCCAGTTTGCCGCCTGTAACCTGGCCAGATCCCTCCCCACGCTGTG ATGGTGCACCCGACACCAACATCTTCACCGTGTACGGCCTGCTGGACAAGGCC CAGGATCTGTTTCGGCGACGACCACAACAAGAACGGGTTCAAGATGTTTCGACTCC AGCAACTACCACGGACAGGATCTGCTGTTTAAAGATGCCACCGTGC GGCGCCGTG CCAGTGGGCGAAAAGACCACCTACAGAGGATGGCTGGGACTGGACTACGTGGCC GCCCTGGAAGGCATGTCTCCAGCAGTGTTCCGAAGCCGCCGCGAAAAGAAGCC GCCGCAAAAAGAAGCCGCTGCCAAATCGGAAACTCAGGCCAACTCCACCACAGAT GCACTCAACGTGCTGCTGATCATCGTAGATGACCTCCGACCTTCTCTGGGCTGT TACGGCGACAAGCTAGTACGGAGCCCAACATCGACCAGCTCGCATCGCACTCT CTCTATTCCAGAACGCATTTCGCCAGCAGGCTGTCTGTGCTCCCTCCCGAGTG TCCTTCTCACGGGTGCGAGACCCGATACCACGAGGTTATATGACTTCAACTCA TACTGGCGCGTGCATGCCGGTAACCTTTTCTACTATACCCAGTATTTTAAAGAA AATGGCTATGTTACAATGTCCGTTGGCAAGGTATTTTCATCTGGTATTAGCAGC AACCACACAGATGACTCTCCGTATAGCTGGTCAATCCACCATAACCACCCCTCC AGCGAAAAGTACGAAAACACAAAGACTTGCCGGGGCCAGATGGCGAACTGCAC GCAAATCTGCTGTGCCTGTAGATGTCTTGGACGTGCCCCAAGGTACTCTGCCC GACAAACAGTCCACAGAACAGGCAATCCAACCTCCTTGAAAAGATGAAAACGAGC GCGTCCCCCTTCTTCTCGCCGTGGGCTACCACAAGCCCCACATCCCGTTTAGA TACCCCAAGGAATTTTACGAAACTGTACCCCTGGAAAACATCACTCTCGCGCCC GACCCCGAAGTGCCAGACGGACTCCCTCCTGTGCTTACAACCCCTGGATGGAC ATCAGACAACGTGAAGATGTGCAGGCCCTGAACATCTCAGTGCCTTACGGCCCC ATTCCAGTTGACTTCCAGAGGAAGATTCCGGCAGTCTTCTCGCCTCCGTTAGT TACCTGGACACCCAAGTGGGTAGACTCCTGAGCGCCTTGGACGATCTCCAGCTC GCAAACAGCACCATCATTGCCTTACCAGCGACCATGGTTGGGCGCTGGGTGAA CATGGAGAATGGGCTAAATATTCAAATTTTCGACGTTGCGACCCACGTCCCATTG ATCTTCTACGTGCCTGGACGAACAGCCTCCTTGCTGAAGCCGGGGAAAAGTTG TTTCCATATCTGGACCTTTTCGATTCTGCGAGCCAACTCATGGAACCTGGGCGA CAGAGCATGGACCTGGTGGAACTGGTCACTTTATTTCCAACCTGGCAGGCCTT GCAGGCCTCCAAGTTCACCTCGGTGTCCCGTTCCCTCATTTCCACGTGCAACTC TGTCGCGAAGGTAAAAACCTCCTCAAGCATTTTCGTTTTTCGGGACCTCGAAGAA GACCCATACCTGCCAGGGAATCCAAGGGAAGTATTGCCTACAGCCAGTACCCT AGACCTAGCGACATCCACAGTGAACAGCGACAAGCCCTCCCTCAAGGACATT AAAATCATGGGTATATAGTATCCGGACTATTGACTACAGGTATACCGTGTGGGTG GGTTTCAACCCAGACGAATTTCTCGCAATTTCTCCGACATCCACGCGGGCGAA CTGTATTTTCGTTGATTCCGATCCACTGCAAGATCATAATATGTACAACGATAGT CAAGGGGTGACCTCTTCCAGTTGCTAATGCCATGA</p>	
MTfpep-I2S	<p>ATGGAATGGAGCTGGGTCTTTCTCTTCTTCTGTGTCAGTAACGACTGGTGTCCAC TCCGACTACAAGGACGACGACGACAAGAGCAGAAGCTGATCTCCGAAGAGGAC CTGCACCACCATCATCACCATCACCACCATCAGGAGGCGGTGGAGAGAACCTG TACTTTTCAGGGCGACTCCTCTCACGCCCTTACCCTGGACGAGCTGCGGTACGAA GCCGCCGCGAAAAGAAGCCGCGCAAAAAGAAGCCGCTGCCAAATCGGAAACTCAG GCCAACTCCACCACAGATGCACTCAACGTGCTGCTGATCATCGTAGATGACCTC CGACCTTCTCTGGGCTGTTACGGCGACAAGCTAGTACGGAGCCCAACATCGAC CAGCTCGCATCGCACTCTCTCTTATTCAGAACGCATTCGCCCAGCAGGCTGTCT TGTGCTCCCTCCCGAGTGTCTTCTTCCAGGGTTCGGAGACCCGATACACGAGG TTATATGACTTCAACTCATACTGGCGCGTGCATGCCGGTAACCTTTTCTACTATA CCCCAGTATTTTAAAGAAAATGGCTATGTTACAATGTCCGTTGGCAAGGTATTT CATCCTGGTATTAGCAGCAACCACACAGATGACTCTCCGTATAGCTGGTCATT CCACCATAACCACCCCTCCAGCGAAAAGTACGAAAACACAAAGACTTGCCGGGGC CCAGATGGCGAACTGCACGCAAACTGCTGTGCGCTGTAGATGTCTTGGACGTG CCCGAAGGTACTCTGCCGACAAACAGTCCACAGAACAGGCAATCCAACCTCCTT GAAAAGATGAAAACGAGCGCGTCCCCCTTCTTCTCGCCGTGGGCTACCACAAG CCCCACATCCCGTTTTAGATACCCCAAGGAATTTTACGAAACTGTACCCCTTGAA AACATCACTCTCGCGCCCGACCCGAAGTGCCAGACGGACTCCCTCCTGTGTC TACAACCTTGGATGGACATCAGACAACGTGAAGATGTGCAGGCCCTGAACATC TCAGTGCCCTTACGGGCCCATTTCCAGTTGACTTCCAGAGGAAGATTTCGGCAGTCC TACTTCGCCTCCGTTAGTTACCTGGACACCCAAAGTGGGTAGACTCCTGAGCGCC TTGGACGATCTCCAGCTCGCAACAGCACCATCATTGCCCTTACCAGCGACCAT</p>	145

Tabla 7. Secuencias polinucleotídicas ilustrativas		
Nombre	Secuencia polinucleotídica	SEQ ID NO:
	GGTTGGGCGCTGGGTGAACATGGAGAATGGGCTAAATATTCAAATTTTCGACGTT GCGACCCACGTCCCATTGATCTTCTACGTGCCTGGACGAACAGCCTCCTTGCCT GAAGCCGGGGAAAAGTTGTTTCCATATCTGGACCCCTTTCGATTCTGCGAGCCAA CTCATGGAACCTGGGCGACAGAGCATGGACCTGGTGGAACCTGGTCACTTTATTT CCAACCTGGCAGGCCTTGCAGGCCTCCAAGTTCCACCTCGGTGTCCCGTTCCC TCATTCCACGTGCAACTCTGTGCGGAAGGTAAAAACCTCCTCAAGCATTTTCGT TTTCGGGACCTCGAAGAAGACCCATACCTGCCAGGGAATCCAAGGGAAGTATT GCCTACAGCCAGTACCCTAGACCTAGCGACATCCACAGTGAACAGCGACAAG CCCTCCCTCAAGGACATTAATAATCATGGGTTATAGTATCCGGACTATTGACTAC AGGTATACCGTGTGGGTGGGTTCACCCAGACGAATTTCTCGCCAATTTCTCC GACATCCACGCGGGCGAAGTGTATTTTCGTTGATTCCGATCCACTGCAAGATCAT AATATGTACAACGATAGTCAAGGGGTGACCTCTTCAGTTGCTAATGCCATGA	
I2S-MTfpep	ATGGAATGGAGCTGGGTCTTTCTCTTCTTCTGTGTCAGTAACGACTGGTGTCCAC TCCGACTACAAGGACGACGACGACAAAGAGCAGAAGCTGATCTCCGAAGAGGAC CTGCACCACCATCATCACCATCACCACCATCACGGAGGCGGTGGAGAGAACCTG TACTTTTCAGGGCTCGGAACTCAGGCCAAGTCCACCACAGATGCACTCAACGTG CTGCTGATCATCGTAGATGACCTCCGACCTTCTCTGGGCTGTTACGGCGACAAG CTAGTACGGAGCCCAAACATCGACCAGCTCGCATCGCACTCTCTCCTATTCCAG AACGCATTCGCCCAGCAGGCTGTCTGTGCTCCCTCCCGAGTGTCTTCTCTCACG GGTCCGAGACCCGATACCACGAGGTTATATGACTTCAACTCATACTGGCGCGTG CATGCCGGTAACCTTTCTACTATACCCAGTATTTTAAAGAAAATGGCTATGTT ACAATGTCCGTTGGCAAGGTATTTTCATCCTGGTATTAGCAGCAACCACACAGAT GACTCTCCGTATAGCTGGTCATTCCCACCATAACCACCCCTCCAGCGAAAAGTAC GAAAACACAAAGACTTGCCGGGGCCAGATGGCGAAGTGCACGCAAAATCTGCTG TGCCCTGTAGATGTCTTGACGCTGCCCCGAAGGTACTCTGCCGACAAAACAGTCC ACAGAACAGGCAATCCAACCTCTTGAAAAGATGAAAACGAGCGCGTCCCCCTTC TTCTTCGCGGTGGGCTACCACAAGCCCCACATCCCGTTTAGATACCCCAAGGAA TTTCAGAACTGTACCCCTGGAAAACATCACTCTCGCGCCCGACCCCGAAGTG CCAGACGGACTCCCTCCTGTTGCCTACAACCCTTGGATGGACATCAGACAACGT GAAGATGTGCAGGCCCTGAACATCTCAGTGCCTTACGGCCCCATTCCAGTTGAC TTCCAGAGGAAGATTCCGGCAGTCTTACTTCGCTCCGTTAGTTACCTGGACACC CAAGTGGGTAGACTCCTGAGCGCCTTGGACGATCTCCAGCTCGCAACACGACC ATCATTGCCTTACCAGCGACCATGGTTGGGCGCTGGGTGAACATGGAGAATGG GCTAAATATTCAAATTTTCGACGTTGCGACCCACGTCCCATTGATCTTCTACGTG CCTGGACGAACAGCCTCCTTGCCCTGAAGCCGGGGAAAAGTTGTTTCCATATCTG GACCTTTTCGATTCTGCGAGCCAACTCATGGAACCTGGGCGACAGAGCATGGAC CTGGTGGAACTGGTCACTTTATTTCCAACCCTGGCAGGCCTTGCAGGCCTCCAA GTTCCACCTCGGTGTCCCGTTCCCTCATTCCACGTGCAACTCTGTGCGGAAGGT AAAAACCTCCTCAAGCATTTTCGTTTTCGGGACCTCGAAGAAGACCCATACCTG CCAGGGAAATCCAAGGGAAGTATTGCTTACAGCCAGTACCTTAGACCTAGCGAC ATCCACAGTGAACACGCGACAAGCCCTCCCTCAAGGACATTAATAATCATGGGT TATAGTATCCGGACTATTGACTACAGGTATACCGTGTGGGTGGGTTTCAACCCA GACGAATTTCTCGCCAATTTCTCCGACATCCACGCGGGCGAAGTGTATTTTCGTT GATTCCGATCCACTGCAAGATCATAATATGTACAACGATAGTCAAGGGGGTGAC CTCTTCCAGTTGCTAATGCCAGAGGCGCTGTAAAGAGGCTGCCGCCAAAGAA GCCGCCGCTAAGGACTCCTCTCAGCCTTACCCCTGGACGAGCTGCGGTACTAA	146

Tabla 7. Secuencias polinucleotídicas ilustrativas		
Nombre	Secuencia polinucleotídica	SEQ ID NO:
I2S-MTfpep (sin propéptido de I2S)	ATGGAATGGAGCTGGGTCTTTCTCTTCTTCTGTCAGTAACGACTGGTGTCCAC TCCGACTACAAGGACGACGACGACAAAGAGCAGAAGCTGATCTCCGAAGAGGAC CTGCACCACCATCATCACCATCACCACCATCAGGAGGCGGTGGAGAGAACCTG TACTTTTCAGGGCACAGATGCACTCAACGTGCTGCTGATCATCGTAGATGACCTC CGACCTTCTCTGGGCTGTTACGGCGACAAGCTAGTACGGAGCCCAACATCGAC CAGCTCGCATCGCACTCTCTCCTATTCCAGAACGCATTGCGCCAGCAGGCTGTC TGTGCTCCCTCCCGAGTGTCTTCTCCTCACGGGTCGGAGACCCGATACCACGAGG TTATATGACTTCAACTCATACTGGCGCGTGCATGCCGGTAACTTTTCTACTATA CCCCAGTATTTTAAAGAAAATGGCTATGTTACAATGTCCGTGGCAAGGTATTT CATCCTGGTATTAGCAGCAACCACACAGATGACTCTCCGTATAGCTGGTCATTTC CCACCATACCACCCCTCCAGCGAAAAGTACGAAAACACAAGACTTGCCGGGGC CCAGATGGCGAACTGCACGCAAACTGTGCTGTGCCCTGTAGATGTCTTGGACGTG CCCGAAGGTACTCTGCCCCGACAAACAGTCCACAGAACAGGCAATCCAACCTCCTT GAAAAGATGAAAACGAGCGCGTCCCCCTTCTTCTCGCCGTGGGCTACCACAAG CCCCACATCCCGTTTATAGTACCCCAAGGAATTCAGAACTGTACCCCTGGAA AACATCACTCTCGCGCCCGACCCCGAAGTGCCAGACGGACTCCCTCCTGTTGCC TACAACCTTGGATGGACATCAGACAACGTGAAGATGTGCAGGCCCTGAACATC TCAGTGCCCTACGGCCCCATTCCAGTTGACTTCCAGAGGAAGATTCCGGCAGTCC TACTTCGCCTCCGTTAGTTACCTGGACACCCAAGTGGGTAGACTCCTGAGCGCC TTGGACGATCTCCAGCTCGCAAACAGCACCATCATTGCCTTCACCAGCGACCAT GGTTGGGCGCTGGGTGAACATGGAGAATGGGCTAAATATTCAAATTTTCGACGTT GCGACCCACGTCCCATTTGATCTTCTACGTGCCTGGACGAACAGCCTCCTTGCTT GAAGCCGGGGAAAAGTTGTTTCCATATCTGGACCCCTTTCGATTCTGCGAGCCAA CTCATGGAACCTGGGCGACAGAGCATGGACCTGGTGGAACTGGTCACTTTATTT CCAACCTGGCAGGCCCTTGCAGGCCTCCAAGTTCACCTCGGTGTCCCGTTCCC TCATTCCAGTTCGAACCTCTGTGCGGAAGGTAAAAACCTCCTCAAGCATTTTCGT TTTCGGGACCTCGAAGAAGACCCATACCTGCCAGGGAATCCAAGGGAAGTATT GCCTACAGCCAGTACCCTAGACCTAGCGACATCCACAGTGGAAACAGCGACAAG CCCTCCCTCAAGGACATTAATAATCATGGGTTATAGTATCCGGACTATTGACTAC AGGTATACCGTGTGGGTGGGTTTCAACCCAGACGAATTTCTCGCCAATTTCTCC GACATCCACGCGGGCGAACTGTATTTGTTGATTCCGATCCACTGCAAGATCAT AATATGTACAACGATAGTCAAGGGGGTGACCTCTTCCAGTTGCTAATGCCAGAG GCCGCTGCTAAAGAGGCTGCCGCCAAAGAAGCCGCCGCTAAGGACTCCTCTCAC GCCTTACCTTGGACGAGCTGCGGTACTAA	147

Por lo tanto, un polinucleótido que codifica una proteína de fusión o fusión de anticuerpo descrita en la presente memoria, o una porción de estas, puede comprender una o más secuencias polinucleotídicas de la Tabla 7 (p. ej., SEQ ID NOS:143-147), o un fragmento/variante de estas.

- 5 En algunas realizaciones, ácidos nucleicos o vectores que codifican un polipéptido de p97 objeto, un polipéptido de IDS y/o una fusión p97-IDS se introducen directamente en una célula hospedante, y la célula se incuba en condiciones suficientes para inducir la expresión del(de los) polipéptido(s) codificado(s). Por lo tanto, según ciertas realizaciones relacionados, se proporciona una célula hospedante recombinante que comprende un polinucleótido o un polinucleótido de fusión que codifica una o más proteínas de fusión descritas en la presente memoria, y que
- 10 opcionalmente comprende polinucleótidos exógenos adicionales.

- La expresión de una proteína de fusión en la célula hospedante se puede lograr al cultivar las células hospedantes recombinantes (que contienen el(los) polinucleótido(s)) en condiciones adecuadas. Después de la producción por expresión, el(los) polipéptido(s) y/o proteínas de fusión se pueden aislar y/o purificar mediante el uso de cualquier técnica adecuada y después usarse según se desee. El término "célula hospedante" se usa para hacer referencia a
- 15 una célula en la que se ha introducido, o que es capaz de tener introducida en ella, una secuencia de ácido nucleico que codifica uno o más de los polipéptidos descritos en la presente memoria, y que expresa, además, o es capaz de expresar un gen de interés seleccionado, tal como un gen que codifica cualquier polipéptido descrito en la presente memoria. El término incluye la progenie de la célula original, sea o no la progenie idéntica en morfología o en composición genética al genitor original, siempre que el gen seleccionado esté presente. Las células hospedantes se
- 20 pueden elegir por ciertas características, por ejemplo, la expresión de una aminoacil ARNt sintetasa(s) que puede incorporar aminoácidos no naturales en el polipéptido.

Los sistemas para la clonación y la expresión de una proteína en una variedad de células hospedantes diferentes se conocen. Las células hospedantes adecuadas incluyen células de mamífero, bacterias, levadura y sistemas de

baculovirus. Las líneas celulares de mamíferos disponibles en la técnica para la expresión de un polipéptido heterólogo incluyen células de ovario de hámster chino (CHO), células HeLa, células de riñón de bebé de hámster, células HEK-293, la línea celular de fibrosarcoma humano HT-1080 (ver, p. ej., Moran, Nat. Biotechnol. 28:1139-40, 2010), células de melanoma de ratón NSO y muchas otras. Los ejemplos adicionales de líneas celulares hospedantes de mamíferos útiles incluyen la línea de riñón de mono CV1 transformada por SV40 (COS-7, ATCC CRL 1651); la línea de riñón embrionario humano (células 293 o 293 subclonadas para crecimiento en cultivo de suspensión, Graham et al., J. Gen Virol. 36:59 (1977)); células de riñón de bebé de hámster (BHK, ATCC CCL 10); células de Sertoli de ratón (TM4, Mather, Biol. Reprod. 23:243-251 (1980)); células de riñón de mono (CV1 ATCC CCL 70); células de riñón de mono verde africano (VERO-76, ATCC CRL-1587); células de carcinoma de cuello de útero humano (HELA, ATCC CCL 2); células de riñón canino (MDCK, ATCC CCL 34); células de hígado de rata buffalo (BRL 3A, ATCC CRL 1442); células de pulmón humano (W138, ATCC CCL 75); células de hígado humano (Hep G2, HB 8065); tumor mamario de ratón (MMT 060562, ATCC CCL51); células TR1 (Mather et al., Annals N.Y. Acad. Sci. 383:44-68 (1982)); células MRC 5; células FS4; y una línea de hepatoma humano (Hep G2). Otras líneas celulares hospedantes de mamíferos útiles incluyen células de ovario de hámster chino (CHO), incluidas células DHFR-CHO (Urlaub et al., PNAS USA 77:4216 (1980)); y líneas celulares de mieloma tales como NSO y Sp2/O. Para acceder a una reseña de ciertas líneas celulares hospedantes de mamíferos adecuadas para la producción de polipéptidos, ver, p. ej., Yazaki and Wu, Methods in Molecular Biology, tomo 248 (B. K.C Lo, ed., Humana Press, Totowa, N.J., 2003), págs. 255-268. Ciertos sistemas de expresión en células de mamíferos preferidos incluyen los sistemas de expresión basados en células CHO y HEK293. Los sistemas de expresión en mamíferos pueden usar líneas celulares acopladas, por ejemplo, en matraces T, frascos rotativos o fábricas celulares, o cultivos en suspensión, por ejemplo, en centrifugadores de 1L y 5L, biorreactores de tanque con agitación de 5L, 14L, 40L, 100L y 200L o biorreactores WAVE de 20/50L y 100/200L, entre otros conocidos en la técnica.

Un hospedante bacteriano común preferido es *E. coli*. La expresión de proteínas en células procariotas tales como *E. coli* está bien establecida en la técnica. Para acceder a una reseña, ver, por ejemplo, Pluckthun, A. Bio/Technology. 9:545-551 (1991). La expresión en células eucariotas en cultivo también está disponible para los expertos en la técnica como una opción para la producción recombinante de polipéptidos (ver, Ref, Curr. Opinion Biotech. 4:573-576, 1993; y Trill et al., Curr. Opinion Biotech. 6:553-560, 1995). En realizaciones específicas, la expresión proteica se puede controlar mediante una ARN polimerasa T7 (p. ej., la serie de vectores pET). Estas y realizaciones relacionadas pueden usar la cepa hospedante para expresión BL21(DE3), un lisógeno λ DE3 de BL21 que admite la expresión mediada por T7 y carece de lon y ompT proteasas para una estabilidad mejorada de la proteína diana. También se incluyen cepas hospedantes de expresión que llevan plásmidos que codifican ARNt rara vez usados en *E. coli*, tales como las cepas Rosetta™ (DE3) y Rosetta 2 (DE3). La lisis celular y la manipulación de muestras también se pueden mejorar mediante el uso de reactivos tales como nucleasa Benzonase® y reactivo de extracción proteica BugBuster®. Para el cultivo celular, los medios autoinductores pueden mejorar la eficacia de muchos sistemas de expresión, incluidos sistemas de expresión de alto rendimiento. Los medios de este tipo (p. ej., el sistema de autoinducción Overnight Express™) gradualmente desencadenan la expresión proteica a través de desplazamiento metabólico sin la adición de agentes inductores artificiales tales como IPTG. Realizaciones específicas emplean etiquetas de hexahistidina (tales como fusiones His-Tag®) y posterior purificación mediante cromatografía de afinidad con inmovilización en metal (IMAC, por sus siglas en inglés) o técnicas relacionadas. En ciertos aspectos, sin embargo, las proteínas de grado clínico se pueden aislar a partir de cuerpos de inclusión de *E. coli*, sin o sin el uso de etiquetas de afinidad (ver, p. ej. Shimp et al., Protein Expr Purif. 50:58-67, 2006). Como ejemplo adicional, ciertas realizaciones pueden emplear un sistema de producción de alto rendimiento en *E. coli* inducido por choque frío, dado que la sobreexpresión de proteínas en *Escherichia coli* a baja temperatura mejora su solubilidad y estabilidad (ver, p. ej., Qing et al., Nature Biotechnology. 22:877-882, 2004).

Además, una cepa de célula hospedante se puede elegir por su capacidad para modular la expresión de las secuencias insertadas o para procesar la proteína expresada de la manera deseada. Dichas modificaciones del polipéptido incluyen, pero no se limitan a, modificaciones postraduccionales tales como acetilación, carboxilación, glicosilación, fosforilación, lipidación y acilación. El procesamiento postraduccion, que escinde una forma "prepro" de la proteína, también se puede usar para facilitar la correcta inserción, plegado y/o función. Se pueden elegir diferentes células hospedantes tales como levadura, CHO, HeLa, MDCK, HEK293 y W138, además de células bacterianas, que tienen o incluso carecen de maquinaria celular específica y mecanismos característicos para dichas actividades postraduccionales, para garantizar la correcta modificación y procesamiento de la proteína de fusión de interés.

Para una producción a largo plazo, de alto rendimiento de proteínas recombinantes, se prefiere generalmente una expresión estable. Por ejemplo, las líneas celulares que expresan de manera estable un polinucleótido de interés se pueden transformar mediante el uso de vectores de expresión que pueden contener orígenes de replicación víricos y/o elementos de expresión endógenos y un gen marcador seleccionable en el mismo o en un vector separado. Después de la introducción del vector, se puede dejar que las células crezcan durante aproximadamente 1-2 días en medios enriquecidos antes de pasarlas a medios selectivos. El objetivo del marcador seleccionable es conferir resistencia a la selección, y su presencia permite el crecimiento y la recuperación de células que expresan con éxito las secuencias introducidas. Los clones resistentes de células transformadas de manera estable se pueden proliferar mediante el uso de técnicas de cultivo de tejido adecuadas para el tipo celular. También se puede emplear la

producción transitoria, tal como por transfección o infección transitoria. Los sistemas de expresión en mamíferos ilustrativos que son adecuados para la producción transitoria incluyen sistemas basados en HEK293 y CHO.

Las células hospedantes transformadas con una secuencia polinucleotídica de interés se pueden cultivar en condiciones adecuadas para la expresión y recuperación de la proteína a partir del cultivo celular. Ciertas realizaciones específicas usan sistemas de expresión celular sin suero. Los ejemplos incluyen células HEK293 y células CHO que pueden crecer en medio sin suero (*ver, p. ej.,* Rosser et al., Protein Expr. Purif. 40:237-43, 2005; y la patente estadounidense número 6 210 922).

La(s) proteína(s) producida mediante una célula recombinante se puede(n) purificar y caracterizar según una variedad de técnicas conocidas en la técnica. Los sistemas ilustrativos para llevar a cabo la purificación proteica y analizar la pureza de la proteína incluyen cromatografía de líquidos de proteína rápida (FPLC, por sus siglas en inglés) (*p. ej.,* los sistemas de FPLC AKTA y Bio-Rad), cromatografía de líquidos a alta presión (HPLC, por sus siglas en inglés) (*p. ej.,* HPLC de Beckman and Waters). Las químicas ilustrativas para la purificación incluyen cromatografía de intercambio iónico (*p. ej.,* Q, S), cromatografía de exclusión por tamaño, gradientes de sal, purificación por afinidad (*p. ej.,* Ni, Co, FLAG, maltosa, glutatión, proteína A/G), filtración en gel, fase inversa, cromatografía de intercambio iónico en cerámica HyperD® y columnas de interacción hidrófoba (HIC, por sus siglas en inglés), entre otras conocidas en la técnica. También se incluyen métodos analíticos tales como SDS-PAGE (*p. ej.,* coomassie, tinción con plata), inmunotransferencia, Bradford y ELISA, que se pueden usar durante cualquier etapa del proceso de producción o purificación, típicamente para medir la pureza de la composición proteica.

Se incluyen, además, métodos para concentrar proteínas producidas de manera recombinante, *p. ej.,* proteínas de fusión. Los ejemplos incluyen liofilización, que se emplea típicamente cuando la disolución contiene pocos componentes solubles distintos de la proteína de interés. La liofilización se lleva a cabo frecuentemente después de un ciclo de HPLC y puede retirar la mayoría o todos los componentes volátiles de la mezcla. También se incluyen técnicas de ultrafiltración que emplean, típicamente, una o más membranas permeables selectivas para concentrar una disolución proteica. La membrana permite que el agua y moléculas pequeñas pasen a través de ella y retiene la proteína; la disolución se puede impulsar contra la membrana mediante una bomba mecánica, presión de gas o centrifugación, entre otras técnicas.

En ciertas realizaciones, las proteínas de fusión tienen una pureza de al menos aproximadamente 90 %, según se mide de acuerdo con técnicas rutinarias en la técnica. En ciertas realizaciones, tales como composiciones de diagnóstico o ciertas composiciones terapéuticas, las proteínas de fusión tienen una pureza de al menos aproximadamente 95 %. En realizaciones específicas, tales como composiciones terapéuticas o farmacéuticas, las proteínas de fusión tienen una pureza de al menos aproximadamente 97 % o 98 % o 99 %. En otras realizaciones, tales como cuando se usan como reactivos de referencia o investigación, las proteínas de fusión pueden tener una pureza menor, y pueden tener una pureza de al menos aproximadamente 50 %, 60 %, 70 % u 80 %. La pureza se puede medir en general o en relación con componentes seleccionados, tales como otras proteínas, *p. ej.,* pureza en función de una proteína.

En ciertas realizaciones, según se indicó anteriormente, las composiciones descritas en la presente están aproximadamente libres sustancialmente de endotoxinas, incluso, por ejemplo, aproximadamente 95 % libres de endotoxinas, preferiblemente, aproximadamente 99 % libres de endotoxinas y, más preferiblemente, aproximadamente 99,99 % libres de endotoxinas. La presencia de endotoxinas se puede detectar según técnicas rutinarias en la técnica, según se describen en la presente memoria. En realizaciones específicas, las proteínas de fusión se producen a partir de una célula eucariota tal como una célula de mamífero o humano en medios sustancialmente sin suero.

Métodos de uso y composiciones farmacéuticas

Ciertas realizaciones de la presente invención se refieren a métodos de uso de las proteínas de fusión de p97 descritas en la presente memoria. Los ejemplos de dichos métodos incluyen métodos de tratamiento y métodos de diagnóstico incluso, por ejemplo, el uso de proteínas de fusión de p97 para generación de imágenes con fines médicos de ciertos órganos/tejidos, tales como los del sistema nervioso. Algunas realizaciones incluyen métodos para diagnosticar y/o tratar trastornos o afecciones del sistema nervioso central (CNS), o trastornos o afecciones que tienen un componente del CNS. Aspectos particulares incluyen métodos para tratar un trastorno de almacenamiento lisosómico (LSD), incluidos los que tienen un componente del CNS.

Por consiguiente, ciertas realizaciones incluyen métodos para tratar a un sujeto que lo necesita, que comprenden administrar una proteína de fusión de p97 descrita en la presente memoria. También se incluyen métodos para suministrar una enzima IDS al sistema nervioso (*p. ej.,* tejidos del sistema nervioso central) de un sujeto, que comprenden administrar una composición que comprende una proteína de fusión de p97 descrita en la presente memoria. En algunas de estas y realizaciones relacionadas, los métodos aumentan la tasa de suministro del agente a los tejidos del sistema nervioso central con respecto, por ejemplo, al suministro mediante una composición que comprende una enzima IDS no fusionada.

- En algunos casos, el sujeto padece o está en riesgo de padecer una enfermedad de almacenamiento lisosómico. Por lo tanto, ciertos métodos se refieren al tratamiento de enfermedades de almacenamiento lisosómico en un sujeto que lo necesita, opcionalmente aquellas enfermedades de almacenamiento lisosómico asociadas al sistema nervioso central, o que presentan afectación del CNS. Las enfermedades de almacenamiento lisosómico ilustrativas incluyen mucopolisacaridosis tipo II (síndrome de Hunter). El síndrome de Hunter es un trastorno multisistémico ligado al cromosoma X caracterizado por una acumulación de glicosaminoglicanos (GAG, por sus siglas). La amplia mayoría de los individuos afectados son de sexo masculino; en raras ocasiones las mujeres portadoras manifiestan signos. La edad de aparición, la gravedad de la enfermedad y la velocidad de progresión pueden variar significativamente.
- En aquellos con enfermedad grave, la afectación del CNS (se manifiesta principalmente por el deterioro cognitivo progresivo), la enfermedad de las vías respiratorias progresiva y la enfermedad cardíaca habitualmente resultan en la muerte en la primera o segunda década de vida. Por lo tanto, ciertas realizaciones incluyen el tratamiento del síndrome de Hunter con afectación del CNS.
- En aquellos con enfermedad atenuada, el CNS no se ve afectado (o mínimamente), aunque el efecto de la acumulación de GAG en otros sistemas orgánicos puede ser tan grave como en los que padecen deterioro cognitivo progresivo. La supervivencia en los primeros años de la vida adulta con inteligencia normal es común en la forma atenuada de la enfermedad. Sin embargo, los sujetos con enfermedad atenuada todavía pueden beneficiarse de la administración de una proteína de fusión de p97-IDS que tiene mejor penetración en los tejidos del CNS, por ejemplo, para reducir el riesgo de progresión de un síndrome de Hunter atenuado a uno con afectación del CNS.
- Los signos adicionales en ambas formas del síndrome de Hunter incluyen: estatura baja; macrocefalia con o sin hidrocefalia comunicante; macroglosia; voz gruesa; pérdida auditiva conductiva y neurosensorial; hepatomegalia y/o esplenomegalia; disostosis múltiples y contracturas articulares que incluyen anquilosis de la articulación temporomandibular; estenosis espinal; y síndrome de túnel carpiano. Los sujetos que se someten a tratamiento con las proteínas de fusión descritas en la presente memoria, por lo tanto, pueden tener uno o más de estos signos del síndrome de Hunter.
- Mediante detección de GAG en la orina y series óseas radiográficas se puede establecer la presencia de una afección MPS, pero no son específicas para MPS II. El criterio de referencia para el diagnóstico de MPS II en una probando de sexo masculino es la actividad enzimática deficiente de iduronato sulfatasa (IDS) en leucocitos, fibroblastos o plasma en presencia de actividad normal de al menos una otra sulfatasa. Las pruebas genéticas moleculares de IDS, el único gen cuya mutación se sabe que está asociada al síndrome de Hunter, se pueden usar para confirmar el diagnóstico en un probando de sexo masculino con un fenotipo inusual o un fenotipo que no coincide con los resultados de la prueba de GAG.
- Los tratamientos comunes para el síndrome de Hunter incluyen terapia del desarrollo, ocupacional y física; derivación para la hidrocefalia; amigdalectomía; ventilación con presión positiva (CPAP o traqueostomía); liberación del túnel carpiano; remplazo de la válvula cardíaca; reparación de hernia inguinal. Por consiguiente, en ciertos aspectos, un sujeto para tratamiento mediante las proteínas de fusión descritas en la presente memoria puede estar por recibir, estar recibiendo o haber recibido uno o más de estos tratamientos.
- La monitorización de la enfermedad puede depender de la afectación del sistema orgánico y la gravedad de la enfermedad, y habitualmente incluye una evaluación cardíaca anual y ecocardiogramas; evaluaciones pulmonares que incluyen evaluar la función pulmonar; audiogramas; exploraciones oculares; evaluaciones del desarrollo; y exploraciones neurológicas. Los estudios adicionales pueden incluir estudios del sueño para la apnea obstructiva; velocidad de conducción nerviosa (NCV, por sus siglas en inglés) para evaluar el síndrome del túnel carpiano; evaluaciones para la hidrocefalia; evaluaciones ortopédicas para monitorizar la enfermedad de cadera. Por lo tanto, en algunos aspectos, un sujeto para tratamiento mediante las proteínas de fusión descritas en la presente memoria puede estar por recibir, estar recibiendo o haber recibido uno o más de estos protocolos de monitorización de la enfermedad.
- Para el uso *in vivo*, por ejemplo, para el tratamiento de una enfermedad humana, generación de imágenes para fines médicos o evaluación, las proteínas de fusión de p97 descritas en la presente memoria se incorporan generalmente en una composición farmacéutica antes de la administración. Una composición farmacéutica comprende una o más de las proteínas de fusión de p97 descritas en la presente memoria en combinación con un portador o excipiente fisiológicamente aceptable.
- Para preparar una composición farmacéutica, se mezcla una cantidad eficaz o deseada de una o más proteínas de fusión con cualquier portador(es) farmacéutico(s) o excipiente conocido para los expertos en la técnica por ser adecuado para el modo de administración específico. Un portador farmacéutico puede ser líquido, semilíquido o sólido. Las disoluciones o suspensiones usadas para aplicación parenteral, intradérmica, subcutánea o tópica pueden incluir, por ejemplo, un diluyente estéril (tal como agua), disolución salina (*p. ej.*, disolución salina tamponada con fosfato; PBS, por sus siglas en inglés), aceite fijo, polietilenglicol, glicerina, propilenglicol u otro disolvente sintético; agentes antimicrobianos (tales como alcohol bencílico y metilparabenos); antioxidantes (tales como ácido ascórbico y bisulfito sódico) y agentes quelantes (tales como ácido etilendiaminotetraacético (EDTA));

tampones (tales como acetatos, citratos y fosfatos). Si se administran por vía intravenosa (p. ej., mediante infusión IV), los portadores adecuados incluyen disolución salina fisiológica o disolución salina tamponada con fosfato (PBS), y disoluciones que contienen agentes espesantes y solubilizantes, tales como glucosa, polietilenglicol, polipropilenglicol y mezclas de estos.

- 5 La administración de proteínas de fusión descritas en la presente memoria, en forma pura o en una composición farmacéutica adecuada, se puede llevar a cabo a través de cualquier modo de administración aceptado de agentes que sirven para utilidades similares. Las composiciones farmacéuticas se pueden preparar al combinar una composición que contiene una proteína de fusión con un portador, diluyente o excipiente fisiológicamente aceptable adecuado, y se pueden formular en preparaciones en formas sólidas, semisólidas, líquidas o gaseosas, tales como comprimidos, cápsulas, polvos, gránulos, ungüentos, disoluciones, supositorios, inyecciones, inhalantes, geles, microesferas y aerosoles. Además, otros ingredientes farmacéuticamente activos (incluidas otras moléculas pequeñas según se describen en otra parte en la presente memoria) y/o excipientes adecuados tales como sales, tampones y estabilizadores pueden estar presentes dentro de la composición, pero no necesariamente.

- 10 La administración se puede lograr mediante una variedad de vías diferentes, incluidas oral, parenteral, nasal, intravenosa, intradérmica, subcutánea o tópica. Los modos de administración preferidos dependen de la naturaleza de la afección que se va a tratar o prevenir. Realizaciones específicas incluyen administración por infusión IV. Algunas realizaciones incluyen administración por inyección intraperitoneal (IP). También se incluyen combinaciones de estas.

- 15 Los portadores pueden incluir, por ejemplo, portadores, excipientes o estabilizadores farmacéuticamente aceptables que no son tóxicos para la célula o mamífero que se expone a ellos en las dosificaciones y concentraciones empleadas. A menudo, el portador fisiológicamente aceptable es una disolución tamponada por pH acuosa. Los ejemplos de portadores fisiológicamente aceptables incluyen tampones tales como fosfato, citrato y otros ácidos orgánicos; antioxidantes que incluyen ácido ascórbico; polipéptido de bajo peso molecular (menos de aproximadamente 10 residuos); proteínas, tales como seroalbúmina, gelatina o inmunoglobulinas; polímeros hidrófilos tales como polivinilpirrolidona; aminoácidos tales como glicina, glutamina, asparagina, arginina o lisina; monosacáridos, disacáridos y otros carbohidratos incluidos glucosa, manosa o dextrinas; agentes quelantes tales como EDTA; alcoholes de azúcar tales como manitol o sorbitol; contraiones formadores de sales tales como sodio; y/o tensioactivos no iónicos tales como polisorbato 20 (TWEEN™) polietilenglicol (PEG) y poloxámeros (PLURONICS™) y similares.

- 20 En ciertos aspectos, una proteína de fusión se une a o encapsula dentro de una partícula, p. ej., una nanopartícula, perla, formulación lipídica, partícula lipídica o liposoma, p. ej., inmunoliposoma. Las proteínas de fusión se pueden atrapar en microcápsulas preparadas, por ejemplo, mediante técnicas de coacervación o mediante polimerización interfacial (por ejemplo, microcápsulas de hidroximetilcelulosa o de gelatina y microcápsulas de poli-(metilmetacrilato), respectivamente), en sistemas de suministro de fármacos coloidales (por ejemplo, liposomas, microesferas de albúmina, microemulsiones, nanopartículas y nanocápsulas), o en macroemulsiones. Dichas técnicas se describen en Remington's Pharmaceutical Sciences, 16a edición, Oslo, A., Ed., (1980). La(s) partícula(s) o liposomas pueden comprender, además, otros agentes terapéuticos o de diagnóstico.

- 25 La dosificación exacta y la duración del tratamiento dependen de la enfermedad que se está tratando y se pueden determinar empíricamente mediante el uso de protocolos de prueba conocidos o mediante la evaluación de las composiciones en sistemas modelo conocidos en la técnica y la extrapolación a partir esta. También se pueden llevar a cabo ensayos clínicos controlados. Las dosificaciones pueden variar también según la gravedad de la afección que se va a aliviar. Una composición farmacéutica generalmente se formula y administra para ejercer un efecto terapéuticamente útil, mientras se minimizan los efectos secundarios indeseables. La composición se puede administrar una vez o se puede dividir en varias dosis más pequeñas para administrarlas en intervalos de tiempo.
- 30 Para cualquiera sujeto en particular, los regímenes de dosificación específicos se pueden ajustar con el tiempo según la necesidad del individuo.

- 35 Las vías de administración típicas de estas y composiciones farmacéuticas relacionadas incluyen, por lo tanto, sin limitación, oral, tópica, transdérmica, inhalación, parenteral, sublingual, bucal, rectal, vaginal e intranasal. El término parenteral, según se usa en la presente memoria, incluye inyecciones subcutáneas, técnicas de inyección o infusión intravenosa, intramuscular, intraesternal. Las composiciones farmacéuticas según ciertas realizaciones de la presente invención se formulan para permitir que los ingredientes activos contenidos en ellas estén biodisponibles después de la administración de la composición a un paciente. Las composiciones que se administrarán a un sujeto o paciente pueden tomar la forma de una o más dosificaciones unitarias donde, por ejemplo, un comprimido puede ser una dosificación unitaria simple, y un contenedor de un conjugado descrito en la presente memoria en forma de aerosol puede alojar una pluralidad dosificaciones unitarias. Los métodos reales para preparar dichas formas de dosificación son conocidos, o serán evidentes, para los expertos en esta técnica; por ejemplo, ver Remington: The Science and Practice of Pharmacy, 20a edición (Philadelphia College of Pharmacy and Science, 2000). La composición que se va a administrar típicamente contendrá una cantidad terapéuticamente eficaz de una proteína de fusión descrita en la presente memoria, para el tratamiento de una enfermedad o afección de interés.

Una composición farmacéutica puede estar en forma de un sólido o líquido. En una realización, el(los) portador(es) están particulados, de manera que las composiciones están, por ejemplo, en forma de comprimido o polvo. El(los) portador(es) pueden ser líquidos, con las composiciones que son, por ejemplo, un aceite oral, un líquido inyectable o un aerosol, que es útil, por ejemplo, en la administración por inhalación. Cuando se prevé la administración oral, la composición farmacéutica preferiblemente está en forma sólida o líquida, donde las formas semisólidas, semilíquidas, suspensión y gel se incluyen dentro de las formas consideradas en la presente memoria como sólidas o líquidas.

Como una composición sólida para administración oral, la composición farmacéutica se puede formular en polvo, gránulo, comprimido, pastilla, cápsula, goma de mascar, oblea o similares. Dicha composición sólida típicamente contendrá uno o más diluyentes inertes o portadores comestibles. Además, uno o más de los siguientes pueden estar presentes: aglutinantes tales como carboximetilcelulosa, etilcelulosa, celulosa microcristalina, goma tragacanto o gelatina; excipientes tales como almidón, lactosa o dextrinas, agentes desintegrantes tales como ácido algínico, alginato sódico, Primogel, almidón de maíz y similares; lubricantes tales como estearato de magnesio o Sterotex; fluidificantes tales como dióxido de silicio coloidal; agentes edulcorantes tales como sacarosa o sacarina; un agente saborizante tal como menta, salicilato de metilo o saborizante de naranja; y un agente colorante. Cuando la composición farmacéutica está en forma de cápsula, por ejemplo, una cápsula de gelatina, puede contener, además de los materiales del tipo indicado anteriormente, un portador líquido tal como polietilenglicol o aceite.

La composición farmacéutica puede estar en forma de un líquido, por ejemplo, un elixir, jarabe, disolución, emulsión o suspensión. El líquido puede ser para administración oral o para suministro mediante inyección, como dos ejemplos. Cuando está prevista para administración oral, la composición preferida contiene, además de los presentes compuestos, uno o más de un agente edulcorante, conservantes, tinte/colorante y potenciador de sabor. En una composición prevista para administración mediante inyección, se puede incluir uno o más de un tensioactivo, conservante, agente humectante, agente dispersante, agente de suspensión, tampón, estabilizador y agente isotónico.

Las composiciones farmacéuticas líquidas, ya sean disoluciones, suspensiones u otras formas similares, pueden incluir uno o más de los siguientes adyuvantes: diluyentes estériles tales como agua para inyección, disolución salina, preferiblemente disolución salina fisiológica, disolución de Ringer, cloruro de sodio isotónico, aceites fijos tales como mono o diglicéridos sintéticos que pueden servir como el disolvente o medio de suspensión, polietilenglicoles, glicerina, propilenglicol u otros disolventes; agentes antibacterianos tales como alcohol bencílico o metilparabeno; antioxidantes tales como ácido ascórbico o bisulfito sódico; agentes quelantes tales como ácido etilendiaminotetraacético; tampones tales como acetatos, citratos o fosfatos y agentes para el ajuste de la tonicidad tales como cloruro de sodio o dextrosa. La preparación parenteral se puede alojar en ampollas, jeringas descartables o viales de múltiples dosis hechos de vidrio o plástico. La disolución salina fisiológica es un adyuvante preferido. Una composición farmacéutica inyectable es preferiblemente estéril.

Una composición farmacéutica líquida prevista para administración parenteral u oral debería contener una cantidad de una proteína de fusión tal que se obtuviera una dosificación adecuada. Típicamente, esta cantidad es al menos 0,01 % del agente de interés en la composición. Cuando está prevista para administración oral, esta cantidad se puede variar entre 0,1 y aproximadamente 70 % del peso de la composición. Ciertas composiciones farmacéuticas orales contienen entre aproximadamente 4 % y aproximadamente 75 % del agente de interés. En ciertas realizaciones, las composiciones farmacéuticas y preparaciones según la presente invención se preparan de manera que una dosificación unitaria parenteral contenga entre 0,01 y 10 % en peso del agente de interés antes de la dilución.

La composición farmacéutica puede estar prevista para administración tópica, en tal caso el portador puede comprender, de manera adecuada, una base de disolución, emulsión, ungüento o gel. La base, por ejemplo, puede comprender uno o más de los siguientes: vaselina, lanolina, polietilenglicoles, cera de abeja, aceite mineral, diluyentes tales como agua y alcohol, y emulsionantes y estabilizadores. Puede haber agentes espesantes presentes en una composición farmacéutica para administración tópica. Si está prevista para administración transdérmica, la composición puede incluir un parche transdérmico o dispositivo de iontoforesis.

La composición farmacéutica puede estar prevista para administración rectal en forma de, por ejemplo, un supositorio, que se fundirá en el recto y liberará el fármaco. La composición para administración rectal puede contener una base oleaginosa tal como un excipiente no irritante adecuado. Dichas bases incluyen, sin limitación, lanolina, manteca de cacao y polietilenglicol.

La composición farmacéutica puede incluir diversos materiales que modifican la forma física de una dosificación unitaria sólida o líquida. Por ejemplo, la composición puede incluir materiales que forman una cubierta de recubrimiento alrededor de los ingredientes activos. Los materiales que forman la cubierta de recubrimiento son típicamente inertes y se pueden seleccionar de, por ejemplo, azúcar, laca y otros agentes de recubrimiento entérico. Alternativamente, los ingredientes activos se pueden encapsular en una cápsula de gelatina. La composición farmacéutica en forma sólida o líquida puede incluir un agente que se une al conjugado o agente y ayuda, de este modo, en el suministro del compuesto. Los agentes adecuados que pueden actuar de este modo incluyen anticuerpos monoclonales o policlonales, una o más proteínas o un liposoma.

La composición farmacéutica puede consistir esencialmente en dosificaciones unitarias que se pueden administrar como un aerosol. El término aerosol se usa para denotar una variedad de sistemas que varían de los de naturaleza coloidal a sistemas que consisten en envases presurizados. El suministro puede hacerse mediante un gas licuado o comprimido o mediante un sistema de bomba adecuado que dispensa los ingredientes activos. Los aerosoles se pueden suministrar en un sistema de fase simple, bifásicos o trifásicos para suministrar el(los) ingrediente(s) activo(s). El suministro del aerosol incluye el contenedor, activadores, válvulas, subcontenedores y similares necesarios, que juntos pueden formar un kit. Un experto en la técnica sin experimentación indebida puede determinar los aerosoles preferidos.

Las composiciones descritas en la presente memoria se pueden preparar con portadores que protegen las proteínas de fusión contra la eliminación rápida del cuerpo, tales como formulaciones o recubrimientos de liberación prolongada. Dichos portadores incluyen formulaciones de liberación controlada, tales como, pero sin limitarse a, implantes y sistemas de suministro microencapsulados, y polímeros biodegradables, biocompatibles, tales como etilenvinil acetato, polianhídridos, ácido poliglicólico, poliésteres, ácido poliláctico y otros conocidos para los expertos en la técnica.

Las composiciones farmacéuticas se pueden preparar mediante metodología conocida en la técnica farmacéutica. Por ejemplo, una composición farmacéutica prevista para administración mediante inyección puede comprender uno o más de sales, tampones y/o estabilizadores, con agua estéril, destilada para formar una disolución. Se puede agregar un tensioactivo para facilitar la formación de una disolución o suspensión homogénea. Los tensioactivos son compuestos que interactúan de manera no covalente con el conjugado para facilitar la disolución o suspensión homogénea del conjugado en el sistema de suministro acuoso.

Las composiciones se pueden administrar en una cantidad terapéuticamente eficaz, que variará dependiendo de una variedad de factores que incluyen la actividad del compuesto específico empleado; la estabilidad metabólica y la extensión de la acción del compuesto; la edad, peso corporal, salud general, sexo y dieta del paciente; el modo y tiempo de administración; la velocidad de excreción; la combinación de fármacos; la gravedad del trastorno o afección específico; y si el sujeto se somete a tratamiento. Generalmente, una dosis diaria terapéuticamente eficaz es (para un mamífero de 70 kg) de aproximadamente 0,001 mg/kg (es decir, ~ 0,07 mg) a aproximadamente 100 mg/kg (es decir, ~ 7,0 g); preferiblemente, una dosis terapéuticamente eficaz es (para un mamífero de 70 kg) de aproximadamente 0,01 mg/kg (es decir, ~ 0,7 mg) a aproximadamente 50 mg/kg (es decir, ~ 3,5 g); más preferiblemente, una dosis terapéuticamente eficaz es (para un mamífero de 70 kg) de aproximadamente 1 mg/kg (es decir, ~ 70 mg) a aproximadamente 25 mg/kg (es decir, ~ 1,75 g).

Las composiciones descritas en la presente memoria se pueden administrar también simultáneamente con, antes o después de la administración de uno o más agentes terapéuticos disintos, según se describió en la presente memoria. Por ejemplo, en una realización, el conjugado se administra con un agente antiinflamatorio. Los fármacos o agentes antiinflamatorios incluyen, pero no se limitan a, esteroides y glucocorticoides (incluidos betametasona, budesonida, dexametasona, acetato de hidrocortisona, hidrocortisona, hidrocortisona, metilprednisolona, prednisolona, prednisona, triamcinolona), fármacos antiinflamatorios no esteroideos (AINE) incluidas aspirina, ibuprofeno, naproxeno, metotrexato, sulfasalazina, leflunomida, medicamentos anti-TNF, ciclofosfamida y micofenolato.

Dicha politerapia puede incluir la administración de una única formulación de dosificación farmacéutica, que contiene un compuesto de la invención (es decir, proteína de fusión) y uno o más agentes activos adicionales, así como la administración de composiciones que comprenden conjugados de la invención y cada agente activo en su propia formulación de dosificación farmacéutica separada. Por ejemplo, una proteína de fusión según se describe en la presente memoria y el otro agente activo se pueden administrar al paciente juntos en una única composición de dosificación oral tal como un comprimido o cápsula, o cada agente se administra en formulaciones de dosificación orales separadas. De manera similar, una proteína de fusión según se describe en la presente memoria y el otro agente activo se pueden administrar al paciente juntos en una única composición de dosificación parenteral tal como en una disolución salina u otra disolución fisiológicamente aceptable, cada agente se administra en formulaciones de dosificación parenterales separadas. Cuando se usan formulaciones de dosificación separadas, las composiciones que comprenden las proteínas de fusión y uno o más agentes activos adicionales se pueden administrar esencialmente al mismo tiempo, es decir, concomitantemente, o en momentos separados en el tiempo, es decir, secuencialmente y en cualquier orden; se entiende que la politerapia incluye todos estos regímenes.

Las diversas realizaciones descritas en la presente memoria se pueden combinar para proporcionar realizaciones adicionales. Se pueden modificar aspectos de las realizaciones, si es necesario para emplear conceptos de las diversas patentes, solicitud y publicaciones para proporcionar todavía realizaciones adicionales.

Estos y otros cambios se pueden hacer a las realizaciones en virtud de la descripción detallada anteriormente.

Ejemplos

Ejemplo 1

Actividad in vitro de proteínas de fusión

Se prepararon y sometieron a prueba proteínas de fusión de p97 humana (melanotransferrina; MTf) y duronato-2-sulfatasa humana (IDS) para determinar la actividad enzimática in vitro. La Tabla E1 proporciona las secuencias de aminoácidos y la Tabla E2 proporciona las secuencias codificantes polinucleotídicas correspondientes de las proteínas de fusión que se prepararon y sometieron a prueba (SEQ ID NO: 138, 139, 143 y 144 son ejemplos de referencia).

5

Tabla E1. Secuencias polipeptídicas de proteínas de fusión		
Nombre	Secuencia	SEQ ID NO:
I2S-MTf (SP : TAG : PS: I2S : Enlazador : MTf soluble)	<p><u>MEWSWVFLFFLSVTTGVHS</u> DYKDDDDKEQKLI SEEDLHHHHHHHHHHGGGG<u>ENL</u> <u>YFQG</u> SETQANSTTDALNVLLIIVDDLRLPSLGCGDKLVRSPNIDQLASHSLLFQ NAFQAQQAVCAPSRVSFLTGRRPDTRLYDFNSYWRVHAGNFSTIPQYFKENG YV TMSVGKVFHPGISSNHTDDSPYSWSFPPYHPSSEKYENTKTCRGPDGELHANLL CPVDVLDVPEGTLDPDKQSTEQAIQLLEKMKTSASPFLLAVGYHKPHIPFRYPKE FQKLYPLENITLAPDPEVPDGLPPVAYNPWMDIRQREDVQALNISVPYGP I PVD FQRKIRQSYFASVSYLDTQVGRLLSALDDLQLANSTIIAFTSDHG WALGEHGEW AKYSNFDVATHVPLIFYVPGRTASLPEAGEKLFYLPDPFDSASQLMPEGRQSM D LVELVSLFPTLAGLAGLQVPPRCVPVPSFHVELCREGKNLLKHFRFRDLEEDPYL PGNPRELIAYSQYPRPSDIPQWNSDKPSLKD IKIMGYSIRTIDYRYTVWVG FNP DEFLANFSDIHAGELYFVDS DPLQDHNMYNDSQGGDLFQLLMP<u>EA</u><u>AAKE</u><u>AAKE</u> <u>AAAK</u> GMEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIAA QEADAITLDGGAIYEAGKEHGLKPVVGEVYDQEVGTSYYAVAVVRRSSHVTIDT LKGVKSCHTG INRTVGWNVPGYLVESGRLSVMGCDVLKAVSDYFGGSCVP GAG ETSYSESLCRLCRGDSSGEGVCDKSP LERYDYSGAFRCLAEGAGDVA FVKHST VLENTDGKTLPSWGQALLSQDFELLCRDGSRADVTEWRQCHLARVPAHAVV VRA DTDGGLIFRLLNEGQRLFSHEGSSFMFSSEAYGQKDLLFKDSTSELVPIATQT YEAWLGHEYLHAMKGLLCDPNRLPPYLRWCVLSTPEIQKCGDMAVAFRRQRLKP EIQCVSAKSPQHMERIQAEQVDAVTLSGEDIYTAGKTYGLVPAAGEHYAPEDS SNSYYVVAVVRRDSSHAFTLDELRGKRSCHAGFGSPAGWDVPVGALIQRGFIRP KDCDVLTA VSEFFNASCVPVNNPKNYPSSLCALCVGDEQGRNKC VGNSQERYYG YRGAFRCLVENAGDVA FVRHTTVFDNTNGHNSEPWAAELRSE DYELLC PN GARA EVSQFAACNLAQIPPHAVMVRPDTNIFTVYGLLDKAQDLFGDDHNKNGFKMFDS SNYHGQDLLFKDATVRAVPVGEKTTYRGWLG LDYVAAL EGMSSQ QCS</p>	138
MTf-I2S (SP : TAG : PS: MTf soluble : Enlazador : I2S)	<p><u>MEWSWVFLFFLSVTTGVHS</u> DYKDDDDKEQKLI SEEDLHHHHHHHHHHGGGG<u>ENL</u> <u>YFQG</u> GMEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIAA QEADAITLDGGAIYEAGKEHGLKPVVGEVYDQEVGTSYYAVAVVRRSSHVTIDT LKGVKSCHTG INRTVGWNVPGYLVESGRLSVMGCDVLKAVSDYFGGSCVP GAG ETSYSESLCRLCRGDSSGEGVCDKSP LERYDYSGAFRCLAEGAGDVA FVKHST VLENTDGKTLPSWGQALLSQDFELLCRDGSRADVTEWRQCHLARVPAHAVV VRA DTDGGLIFRLLNEGQRLFSHEGSSFMFSSEAYGQKDLLFKDSTSELVPIATQT YEAWLGHEYLHAMKGLLCDPNRLPPYLRWCVLSTPEIQKCGDMAVAFRRQRLKP EIQCVSAKSPQHMERIQAEQVDAVTLSGEDIYTAGKTYGLVPAAGEHYAPEDS SNSYYVVAVVRRDSSHAFTLDELRGKRSCHAGFGSPAGWDVPVGALIQRGFIRP KDCDVLTA VSEFFNASCVPVNNPKNYPSSLCALCVGDEQGRNKC VGNSQERYYG YRGAFRCLVENAGDVA FVRHTTVFDNTNGHNSEPWAAELRSE DYELLC PN GARA EVSQFAACNLAQIPPHAVMVRPDTNIFTVYGLLDKAQDLFGDDHNKNGFKMFDS SNYHGQDLLFKDATVRAVPVGEKTTYRGWLG LDYVAAL EGMSSQ QCS<u>EA</u><u>AAKEA</u> <u>AAKEAAAK</u> SETQANSTTDALNVLLIIVDDLRLPSLGCGDKLVRSPNIDQLASHS LLFQNAFAQQAVCAPSRVSFLTGRRPDTRLYDFNSYWRVHAGNFSTIPQYFKE NGYVTMSVGKVFHPGISSNHTDDSPYSWSFPPYHPSSEKYENTKTCRGPDGELH ANLLCPVDVLDVPEGTLDPDKQSTEQAIQLLEKMKTSASPFLLAVGYHKPHIPFR YPKEFQKLYPLENITLAPDPEVPDGLPPVAYNPWMDIRQREDVQALNISVPYGP I PVD FQRKIRQSYFASVSYLDTQVGRLLSALDDLQLANSTIIAFTSDHG WALGE HGEWAKYSNFDVATHVPLIFYVPGRTASLPEAGEKLFYLPDPFDSASQLMPEGR QSM DLVELVSLFPTLAGLAGLQVPPRCVPVPSFHVELCREGKNLLKHFRFRDLEE DPYLPGNPRELIAYSQYPRPSDIPQWNSDKPSLKD IKIMGYSIRTIDYRYTVWV GFNPDEFLANFSDIHAGELYFVDS DPLQDHNMYNDSQGGDLFQLLMP</p>	139

Tabla E1. Secuencias polipeptídicas de proteínas de fusión		
Nombre	Secuencia	SEQ ID NO:
MTfpep-I2S (SP : TAG : PS : MTfpep : Enlazador : I2S)	MEWSWVFLFFLSVTTGVHS DYKDDDDKEQKLISEEDLHHHHHHHHHHGGGG ENL YFQG DSSHAFTLDELRY EAAAKEAAAKEAAAK SETQANSTTDALNVLLIIVDDL RPSLGCYGDCLVRSFNIDQLASHSLLFQNAFAQQAVCAPSRVSFLTGRRPDTTR LYDFNSYWRVHAGNFSTIPQYFKENGYVTMSVGKVFHPGISSNHTDDSPYSWSF PPYHPSSEKYENTKTCRGPDGELHANLLCPVDVLDVPEGTLDPKQSTEQAIQLL EKMKTSASPFLLAVGYHKPHIPFRYPKEFQKLYPLENITLAPDPEVPDGLPPVA YNPWMDIRQREDVQALNISVPYGPPIPVDFQRKIRQSYFASVSYLDTQVGRLLSA LDDLQLANSTIIAFTSDHGVALGEHGEWAKYSNFDVATHVPLIFYVPGRTASLP EAGEKLFYLDPFDSASQLMEPGRQSMDELVELVSLFPTLAGLAGLQVPPRCVPV SFHVELCREGKNLLKHFRFRDLEEDPYLPGNPRELIAYSQYPRPSDIPQWNSDK PSLKDIIKIMGYSIRTIDYRYTVWVGFPNDEFNANFSDIHAGELYFVDSPLQDH NMYNDSQGGDLFQLLMP	140
I2S-MTfpep (SP : TAG : PS : I2S : Enlazador : MTfpep)	MEWSWVFLFFLSVTTGVHS DYKDDDDKEQKLISEEDLHHHHHHHHHHGGGG ENL YFQG SETQANSTTDALNVLLIIVDDLRLPSLGCYGDCLVRSFNIDQLASHSLLFQ NAFAQQAVCAPSRVSFLTGRRPDTTRLYDFNSYWRVHAGNFSTIPQYFKENGYV TMSVGKVFHPGISSNHTDDSPYSWSFPPYHPSSEKYENTKTCRGPDGELHANLL CPVDVLDVPEGTLDPKQSTEQAIQLEKMKTSASPFLLAVGYHKPHIPFRYPKE FQKLYPLENITLAPDPEVPDGLPPVAYNPWMDIRQREDVQALNISVPYGPPIVD FQRKIRQSYFASVSYLDTQVGRLLSALDDLQLANSTIIAFTSDHGVALGEHGEW AKYSNFDVATHVPLIFYVPGRTASLPEAGEKLFYLDPFDSASQLMEPGRQSM LVELVSLFPTLAGLAGLQVPPRCVPVPSFHVELCREGKNLLKHFRFRDLEEDPYL PGNPRELIAYSQYPRPSDIPQWNSDKPSLKDIIKIMGYSIRTIDYRYTVWVGFPN DEFNANFSDIHAGELYFVDSPLQDHNMYNDSQGGDLFQLLMP EAAAKEAAAKE AAAK DSSHAFTLDELRY	141
I2S-MTfpep (sin propéptido de I2S) (SP : TAG : PS : sin propéptido de I2S : Enlazador : MTfpep)	MEWSWVFLFFLSVTTGVHS DYKDDDDKEQKLISEEDLHHHHHHHHHHGGGG ENL YFQG TDALNVLLIIVDDLRLPSLGCYGDCLVRSFNIDQLASHSLLFQNAFAQQAV CAPSRVSFLTGRRPDTTRLYDFNSYWRVHAGNFSTIPQYFKENGYVTMSVGKVF HPGISSNHTDDSPYSWSFPPYHPSSEKYENTKTCRGPDGELHANLLCPVDVLDV PEGTLDPKQSTEQAIQLEKMKTSASPFLLAVGYHKPHIPFRYPKEFQKLYPLE NITLAPDPEVPDGLPPVAYNPWMDIRQREDVQALNISVPYGPPIPVDFQRKIRQ SYFASVSYLDTQVGRLLSALDDLQLANSTIIAFTSDHGVALGEHGEWAKYSNFDV ATHVPLIFYVPGRTASLPEAGEKLFYLDPFDSASQLMEPGRQSMDELVELVSLF PTLAGLAGLQVPPRCVPVPSFHVELCREGKNLLKHFRFRDLEEDPYLPGNPRELI AYSQYPRPSDIPQWNSDKPSLKDIIKIMGYSIRTIDYRYTVWVGFPNDEFNANFS DIHAGELYFVDSPLQDHNMYNDSQGGDLFQLLMP EAAAKEAAAKEAAAK DSSH AFTLDELRY	142

Tabla E2. Secuencias codificantes polinucleotídicas de construcciones de fusión		
Nombre	Secuencia polinucleotídica	SEQ ID NO:

Tabla E2. Secuencias codificantes polinucleotídicas de construcciones de fusión		
Nombre	Secuencia polinucleotídica	SEQ ID NO:
I2S-MTf	<p>ATGGAATGGAGCTGGGTCTTTCTCTTCTTCTGTCAGTAACGACTGGTGTCCAC TCCGACTACAAGGACGACGACGACAAAGAGCAGAAGCTGATCTCCGAAGAGGAC CTGCACCACCATCATCACCATCACCACCATCACGGAGGCGGTGGAGAGAACCTG TACTTTCAGGGCTCGGAAACTCAGGCCAACTCCACCACAGATGCACTCAACGTG CTGCTGATCATCGTAGATGACCTCCGACCTTCTCTGGGCTGTTACGGCGACAAG CTAGTACGGAGCCCAAACATCGACCAGCTCGCATCGCACTCTCTCCTATTCCAG AACGCATTGCCCCAGCAGGCTGTCTGTGCTCCCTCCCGAGTGTCTTCTCCTCACG GGTCCGAGACCCGATACCACGAGGTTATATGACTTCAACTCATACTGGCGCGTG CATGCCGGTAACCTTTTCTACTATACCCAGTATTTTAAAGAAAATGGCTATGTT ACAATGTCCGTTGGCAAGGTATTTTCATCCTGGTATTAGCAGCAACCACACAGAT GACTCTCCGTATAGCTGGTCAATTCCCACCATACCACCCCTCCAGCGAAAAGTAC GAAAACACAAAGACTTGCCGGGGCCAGATGGCGAACTGCACGCAAACTCTGCTG TGCCCTGTAGATGTCTTGACGTGCCCCAAGGTACTCTGCCCCGACAAACAGTCC ACAGAACAGGCAATCCAACCTCCTTGAAAAGATGAAAACGAGCGCGTCCCCCTTC TTCTCGCCGTGGGCTACCACAAGCCCCACATCCCGTTTAGATACCCCAAGGAA TTTCAGAACTGTACCCCTTGAAAACATCACTCTCGCGCCCGACCCCGAAGTG CCAGACGGACTCCCTCCTGTTGCCTACAACCCCTTGGATGGACATCAGACAACGT GAAGATGTGCAGGCCCTGAACATCTCAGTGCCTTACGGCCCCATTCCAGTTGAC TTCCAGAGGAAGATTCGGCAGTCCTACTTCGCCTCCGTTAGTTACCTGGACACC CAAGTGGGTAGACTCCTGAGCGCCTTGGACGATCTCCAGCTCGCAAAACAGCACC ATCATTGCCTTCACCAGCGACCATGGTTGGGCGCTGGGTGAACATGGAGAATGG GCTAAATATTCAAATTTTCGACGTTGCGACCCACGTCCCATTTGATCTTCTACGTG CCTGGACGAACAGCCTCCTTGCCCTGAAGCCGGGGAAAAGTTGTTTCCATATCTG GACCTTTTCGATTCTGCGAGCCAACCTCATGGAACCTGGGCGACAGAGCATGGAC CTGGTGGAACTGGTCAGTTTATTTCCAACCCTGGCAGGCCTTGCAGGCCTCCAA GTTCCACCTCGGTGTCCCCTCCCTCATTCACGTCGAACTCTGTGCGGAAGGT AAAAACCTCCTCAAGCATTTTTCGTTTTTCGGGACCTCGAAGAAGACCCATACCTG CCAGGGAATCCAAGGGAACGATTGCCTACAGCCAGTACCCTAGACCTAGCGAC ATCCACAGTGGAACAGCGACAAGCCCTCCCTCAAGGACATTAAAAATCATGGGT TATAGTATCCGGACTATTGACTACAGGTATACCGTGTGGGTGGGTTCAACCCA</p>	143

Tabla E2. Secuencias codificantes polinucleotídicas de construcciones de fusión		
Nombre	Secuencia polinucleotídica	SEQ ID NO:
	<p>GACGAATTTCTCGCCAATTTCTCCGACATCCACGCGGGCGAACTGTATTTTCGTT GATTCCGATCCACTGCAAGATCATAATATGTACAACGATAGTCAAGGGGGTGAC CTCTTCCAGTTGCTAATGCCAGAAGCCGCCGCGAAAGAAGCCGCCGCAAAAGAA GCCGCTGCCAAAGGCATGGAAGTGCGTTGGTGCGCCACCTCTGACCCCCGAGCAG CACAAAGTGCGGCAACATGTCCGAGGCCCTCAGAGAGGCCGGCATCCAGCCTTCT CTGCTGTGTGTGCGGGGCACCTCTGCCGACCATTGCGTGCAGCTGATCGCCGCC CAGGAAGCCGACGCTATCACACTGGATGGCGGCGCTATCTACGAGGCTGGCAAA GAGCACGGCCTGAAGCCCGTCTGCGGCGAGGTGTACGATCAGGAAGTGGGCACC TCCTACTACGCCGTGGCTGTCTGTGCGGAGATCCTCCACGTGACCATCGACACC CTGAAGGGCGTGAAGTCTGCCACACCGGCATCAACAGAACCCTGGGCTGGAAC GTGCCCGTGGGCTACCTGGTGAATCCGGCAGACTGTCCGTGATGGGCTGCCGAC GTGCTGAAGGCCGTGTCCGATTACTTCGGCGGCTCTTGTGTGCCGTGGCGCTGGC GAGACATCCTACTCCGAGTCCCTGTGCAGACTGTGCAGGGGCGACTCTTCTGGC GAGGGCGTGTGCACAAGTCCCTCTGGAACGGTACTACGACTACTCCGGCGCC TTCAGATGCCTGGCTGAAGGTGCTGGCGACGTGGCCTTCGTGAAGCACTCCACC GTGCTGGAAAACACCGACGGCAAGACCCTGCCTTCTTGGGGCCAGGCACTGTG TCCCAGGACTTCGAGCTGCTGTGCCGGGATGGCTCCAGAGCCGATGTGACAGAG TGGCGGCAGTGCCACCTGGCCAGAGTGCCTGCTCATGCTGTGGTCTGCGCGCC GATACAGATGGCGGCCCTGATCTTCCGGCTGCTGAACGAGGGCCAGCGGCTGTTT TCTCACGAGGGCTCCAGCTTCCAGATGTTCTCCAGCGAGGCCTACGGCCAGAAG GACCTGCTGTTCAAGGACTCCACCTCCGAGCTGGTGCCTATCGCCACCCAGACC TATGAGGCTTGGCTGGGCCACGAGTACCTGCACGCTATGAAGGGACTGCTGTGC GACCCCAACCGGCTGCCTCCTTATCTGAGGTGGTGCCTGCTGTCCACCCCCGAG ATCCAGAAATGCGCGCATATGCGCGTGGCCTTTCGGCGGCAGAGACTGAAGCCT GAGATCCAGTGCGTGTCCGCCAAGAGCCCTCAGCACTGCATGGAACGGATCCAG GCCGAACAGGTGGACGCCGTGACACTGTCCGGCGAGGATATCTACACCGCCGGA AAGACCTACGGCCTGGTGCCAGCTGCTGGCGAGCATTACGCCCTGAGGACTCC TCCAACAGCTACTACGTGGTGGCAGTCTGTGCGCCGGGACTCCTCTCACGCCTTT ACCTGGATGAGCTGCGGGGCAAGAGAAGCTGTACGCCGGCTTTGGAAGCCCT GCCGGATGGGATGTGCTGTGGGCGCTCTGATCCAGCGGGGCTTCATCAGACCC AAGGACTGTGATGTGCTGACCGCCGTGTCTGAGTTCTTCAACGCCCTCCTGTGTG CCCGTGAACAACCCCAAGAACTACCCCTCCAGCCTGTGCGCCCTGTGTGTGGGA GATGAGCAGGGCCGGAACAAATGCGTGGGCAACTCCCAGGAAAGATATTACGGC TACAGAGGCGCCTTCCGGTGTCTGGTGGAAAACGCCGGGATGTGGCTTTTGTG CGGCACACCACCGTGTTCGACAACACCAATGGCCACAACCTCCGAGCCTTGGGCC GCTGAGCTGAGATCCGAGGATTACGAACGTCTGTGTCCCAACGGCGCCAGGGCT GAGGTGTCCAGTTTGGCCCTGTAACTGGCCAGATCCCTCCCCACGCTGTG ATGGTGCAGCCCGACACCAACATCTTACCGTGTACGGCCTGCTGGACAAGGCC CAGGATCTGTTCGGCGACGACCACAACAAGAACGGGTTCAAGATGTTTCGACTCC AGCAACTACCACGGACAGGATCTGCTGTTTAAAGATGCCACCGTGCGGGCCGTG CCAGTGGGCGAAAAGACCACCTACAGAGGATGGCTGGGACTGGACTACGTGGCC GCCCTGGAAGGCATGTCTCCAGCAGTGTTCCTGA</p>	

Tabla E2. Secuencias codificantes polinucleotídicas de construcciones de fusión		
Nombre	Secuencia polinucleotídica	SEQ ID NO:
MTf-I2S	ATGGAATGGAGCTGGGTCTTTCTCTTCTTCCTGTCAGTAACGACTGGTGTCCAC TCCGACTACAAGGACGACGACGACAAAGAGCAGAAGCTGATCTCCGAAGAGGAC CTGCACCACCATCATCACCATCACCACCATCACGGAGGCGGTGGAGAGAACCTG TACTTTCAGGGCGGCATGGAAGTGCGTTGGTGCGCCACCTCTGACCCCGAGCAG CACAAGTGCGGCAACATGTCCGAGGCCCTTCAGAGAGGCCGGCATCCAGCCTTCT CTGCTGTGTGTGCGGGGCACCTCTGCCGACCATTGCGTGCAGCTGATCGCCGCC CAGGAAGCCGACGCTATCACACTGGATGGCGGCGCTATCTACGAGGCTGGCAA GAGCACGGCCTGAAGCCCGTCGTGGGCGAGGTGTACGATCAGGAAGTGGGCACC TCCTACTACGCCGTGGCTGTCTGTGCGGAGATCCTCCACGTGACCATCGACACC CTGAAGGGCGTGAAGTCTGCCACACCGGCATCAACAGAACCGTGGGCTGGAAC GTGCCCGTGGGCTACCTGGTGAATCCGGCAGACTGTCCGTGATGGGCTGCGAC GTGCTGAAGGCCGTGTCCGATTACTTCGGCGGCTCTTGTGTGCCTGGCGCTGGC GAGACATCCTACTCCGAGTCCCTGTGCAGACTGTGCAGGGGCGACTCTTCTGGC GAGGGCGTGTGCGACAAGTCCCTCTGGAACGGTACTACGACTACTCCGGCGCC TTCAGATGCCTGGCTGAAGGTGCTGGCGACGTGGCCTTCGTGAAGCACTCCACC GTGCTGGAACACCGACGGCAAGACCCTGCCTTCTTGGGGCCAGGCACTGCTG TCCCAGGACTTCGAGCTGCTGTGCCGGGATGGCTCCAGAGCCGATGTGACAGAG TGGCGGCAGTGCCACCTGGCCAGAGTGCCTGCTCATGCTGTGGTCTGCGCGCC GATACAGATGGCGGCCTGATCTTCCGGCTGCTGAACGAGGGCCAGCGGCTGTTC	144

Tabla E2. Secuencias codificantes polinucleotídicas de construcciones de fusión		
Nombre	Secuencia polinucleotídica	SEQ ID NO:
	<p>TCTCACGAGGGCTCCAGCTTCCAGATGTTCTCCAGCGAGGCCTACGGCCAGAAG GACCTGCTGTTCAAGGACTCCACCTCCGAGCTGGTGCCTATCGCCACCCAGACC TATGAGGCTTGGCTGGGCCACGAGTACCTGCACGCTATGAAGGGACTGCTGTGC GACCCCAACCGGCTGCCTCCTTATCTGAGGTGGTGGCTGCTGTCCACCCCCGAG ATCCAGAAATGCGGCGATATGGCCGTGGCCTTTTCGGCGGCAGAGACTGAAGCCT GAGATCCAGTGCGTGTCCGCCAAGAGCCCTCAGCACTGCATGGAACGGATCCAG GCCGAACAGGTGGACGCCGTGACACTGTCCGGCGAGGATATCTACACGCCCGGA AAGACCTACGGCCTGGTGCCAGCTGCTGGCGAGCATTACGCCCTGAGGACTCC TCCAACAGCTACTACGTGGTGGCAGTGTGCGCCGGGACTCCTCTCACGCCTTT ACCTTGGATGAGCTGCGGGGCAAGAGAAGCTGTACGCCGGCTTTTGAAGCCCT GCCGGATGGGATGTGCTGTGGGCGCTCTGATCCAGCGGGGCTTCATCAGACCC AAGGACTGTGATGTGCTGACCGCGTGTCTGAGTTCTTCAACGCCCTCCTGTGTG CCCGTGAACAACCCCAAGAACTACCCCTCCAGCCTGTGCGCCCTGTGTGTGGGA GATGAGCAGGGCCGGAACAAATGCGTGGGCAACTCCAGGAAAGATATTACGGC TACAGAGGCGCCTTCCGGTGTCTGGTGGAAAACGCCGGGGATGTGGCTTTTGTG CGGCACACCACCGTGTTCGACAAACACCAATGGCCACAACCTCCGAGCCTTGGGCC GCTGAGCTGAGATCCGAGGATTACGAACGTGTGTGTCCCAACGGCGCCAGGGCT GAGGTGTCCAGTTTGGCGCCTGTAACCTGGCCCAGATCCCTCCCCACGCTGTG ATGGTGCGACCCGACACCAACATCTTACCGTGTACGGCCTGCTGGACAAGGCC CAGGATCTGTTTCGGCGACGACCACAACAAGAACGGGTTCAGATGTTTCGACTCC AGCAACTACCACGGACAGGATCTGCTGTTTAAAGATGCCACCGTGGGGCCGTG CCAGTGGGCGAAAAGACCACCTACAGAGGATGGCTGGGACTGGACTACGTGGCC GCCCTGGAAGGCATGTCTCCAGCAGTGTTCGAAGCCGCCGGAAGAAGCC GCCGCAAAAGAAGCCGCTGCCAAATCGGAAACTCAGGCCAACTCCACCACAGAT GCACTCAACGTGCTGCTGATCATCGTAGATGACCTCCGACCTTCTCTGGGCTGT TACGGCGACAAGCTAGTACGGAGCCCAACATCGACCAGCTCGCATCGCACTCT CTCCTATTCCAGAACGCATTTCGCCAGCAGGCTGTCTGTGCTCCCTCCCGAGTG TCCTTCCTCACGGGTGCGAGACCCGATACCACGAGGTTATATGACTTCAACTCA TACTGGCGCGTGCATGCCGTAACCTTTTCTACTATACCCAGTATTTTAAAGAA AATGGCTATGTTACAATGTCCGTGGCAAGGTATTTTCATCCTGGTATTAGCAGC AACCACACAGATGACTCTCCGTATAGCTGGTCATTCCCAACATACCACCCCTCC AGCGAAAAGTACGAAAACACAAGACTTGCCGGGGCCAGATGGCGAAGTGCAC GCAAATCTGCTGTGCCCTGTAGATGTCTTGGACGTGCCGGAAGGTACTCTGCC GACAAACAGTCCACAGAACAGGCAATCCAACCTCCTTGAAAAGATGAAAACGAGC GCGTCCCCCTTCTTCTCGCCGTGGGCTACCACAAGCCCCACATCCCGTTTAGA TACCCCAAGGAATTTAGAACTGTACCCCTGGAACATCACTCTCGCGCCC GACCCGAAGTGCCAGACGACTCCCTCCTGTTGCCCTACAACCCTTGGATGGAC ATCAGACAACGTGAAGATGTGCAGGCCCTGAACATCTCAGTGCCTTACGGCCCC ATTCCAGTTGACTTCCAGAGGAAGATTCCGCGAGTCTTACTTCGCCTCCGTTAGT TACCTGGACACCCAAAGTGGGTAGACTCCTGAGCGCCTTGGACGATCTCCAGCTC GCAAACAGCACCATCATTGCCCTTACCAGCGACCATGGTTGGGCGCTGGGTGAA CATGGAGAATGGGCTAAATATTCAAATTTTCGACGTTGCGACCCACGTCCCATTG ATCTTCTACGTGCCTGGACGAACAGCCTCCTTGCCCTGAAGCCGGGGAAAAGTTG TTTCCATATCTGGACCCCTTTCGATTCTGCGAGCCAACATCATGGAACCTGGGCGA CAGAGCATGGACCTGGTGGAACTGGTCAGTTTATTTCCAACCCTGGCAGGCCCT GCAGGCCTCCAAGTTCCACCTCGGTGTCCCGTTCCCTCATTTCCACGTCGAACCTC TGTCGCGAAGGTAAAAACCTCCTCAAGCATTTTCGTTTTTCGGGACCTCGAAGAA GACCCATACCTGCCAGGGAATCCAAGGGAAGTGAATGCCTACAGCCAGTACCCCT AGACCTAGCGACATCCACAGTGGAAACAGCGACAAGCCCTCCCTCAAGGACATT AAAATCATGGGTATAGTATCCGGACTATTGACTACAGGTATACCGTGTGGGTG GGTTTCAACCCAGACGAATTTCTCGCAATTTCTCCGACATCCACGCGGGCGAA CTGTATTTTCGTTGATTCCGATCCACTGCAAGATCATAATATGTACAACGATAGT CAAGGGGGTGACCTCTTCCAGTTGCTAATGCCATGA</p>	
MTfpep-I2S		145

Tabla E2. Secuencias codificantes polinucleotídicas de construcciones de fusión		
Nombre	Secuencia polinucleotídica	SEQ ID NO:
	<p>ATGGAATGGAGCTGGGTCTTTCTCTTCTTCTCCTGTCAGTAACGACTGGTGTCCAC TCCGACTACAAGGACGACGACGACAAAGAGCAGAAGCTGATCTCCGAAGAGGAC CTGCACCACCATCATCACCATCACCACCATCACGGAGGCGGTGGAGAGAACCTG TACTTTCAGGGCGACTCCTCTCACGCCTTCACCCTGGACGAGCTGCGGTACGAA GCCGCCGCGAAAGAAGCCGCCGCAAAAGAAGCCGCTGCCAAATCGGAAACTCAG GCCAACTCCACCACAGATGCACTCAACGTGCTGCTGATCATCGTAGATGACCTC CGACCTTCTCTGGGCTGTTACGGCGACAAGCTAGTACGGAGCCCCAACATCGAC CAGCTCGCATCGCACTCTCTCTTATTCAGAACGCATTGCCCCAGCAGGCTGTC TGTGCTCCCTCCCGAGTGTCTTCTCACGGGTCGGAGACCCGATACACGAGG TTATATGACTTCAACTCATACTGGCGCGTGTCATGCCGGTAACTTTTCTACTATA CCCCAGTATTTTAAAGAAAATGGCTATGTTACAATGTCCGTGGCAAGGTATTT CATCCTGGTATTAGCAGCAACCACACAGATGACTCTCCGTATAGCTGGTCATTC CCACCATACCACCCCTCCAGCGAAAAGTACGAAAACACAAAGACTTGCCGGGGC CCAGATGGCGAACTGCACGCAAATCTGCTGTGCCCTGTAGATGTCTTGGACGTG CCCGAAGGTACTCTGCCCCGACAAACAGTCCACAGAACAGGCAATCCAACCTCTT GAAAAGATGAAAACGAGCGCGTCCCCCTTCTTCTCGCCGTGGGCTACCACAAG CCCCACATCCCGTTTAGATACCCCCAAGGAATTTTCAGAACTGTACCCCTGGAA AACATCACTCTCGCGCCCGACCCGAAGTGCCAGACGGACTCCCTCTGTGTC TACAACCCCTGGATGGACATCAGACAACGTGAAGATGTGCAGGCCCTGAACATC TCAGTGCCTTACGGCCCCATTCCAGTTGACTTCCAGAGGAAGATTTCGGCAGTCC TACTTCGCCTCCGTAGTTACCTGGACACCCAAGTGGGTAGACTCCTGAGCGCC TTGGACGATCTCCAGCTCGCAAACAGCACCATCATTGCCTTCACCAGCGACCAT GGTTGGGCGCTGGGTGAACATGGAGAATGGGCTAAATATTCAAATTTTCGACGTT GCGACCCACGTCCCATTTGATCTTCTACGTGCCCTGGACGAACAGCCTCCTGTCCT GAAGCCGGGGAAAAGTTGTTTCCATATCTGGACCCCTTTCGATTCTGCGAGCCAA CTCATGGAACCTGGGCGACAGAGCATGGACCTGGTGGAACTGGTCAGTTTATTT CCAACCCCTGGCAGGCCTTGCAGGCCTCCAAGTTCCACCTCGGTGTCCCGTTCCC TCATTCCACGTGCAACTCTGTGCGGAAGGTAAAAACCTCCTCAAGCATTTTCGT TTTCGGGACCTCGAAGAAGACCCATACCTGCCAGGGAATCCAAGGGAAGTATT GCCTACAGCCAGTACCCTAGACCTAGCGACATCCACAGTGGAACAGCGACAAG CCCTCCCTCAAGGACATTAATAATCATGGGTTATAGTATCCGGACTATTGACTAC AGGTATACCGTGTGGGTGGGTTTCAACCCAGACGAATTTCTCGCCAATTTCTCC GACATCCACGCGGGCGAACTGTATTTTCGTTGATTCCGATCCACTGCAAGATCAT AATATGTACAACGATAGTCAAGGGGTGACCTCTTCCAGTTGCTAATGCCATGA</p>	

Tabla E2. Secuencias codificantes polinucleotídicas de construcciones de fusión		
Nombre	Secuencia polinucleotídica	SEQ ID NO:
I2S-MTfpep	<p>ATGGAATGGAGCTGGGTCTTTCTCTTCTCCTGTCTAGTAACGACTGGTGTCCAC TCCGACTACAAGGACGACGACGACAAAGAGCAGAAGCTGATCTCCGAAGAGGAC CTGCACCACCATCATCACCATCACCACCATCACGGAGCGGTGGAGAGAACCTG TACTTTCAGGGCTCGGAAACTCAGGCCAACTCCACCACAGATGCACTCAACGTG CTGCTGATCATCGTAGATGACCTCCGACCTTCTCTGGGCTGTTACGGCGACAAG CTAGTACGGAGCCCAAACATCGACCAGCTCGCATCGCACTCTCTCCTATTCCAG AACGCATTGCCCCAGCAGGCTGTCTGTGCTCCCTCCCGAGTGTCTTCTCCTCACG GGTCCGAGACCCGATACCACGAGGTTATATGACTTCAACTCATACTGGCGCGTG CATGCCGGTAACCTTTTCTACTATACCCAGTATTTTAAAGAAAATGGCTATGTT ACAATGTCCGTTGGCAAGGTATTTTCATCCTGGTATTAGCAGCAACCACACAGAT GACTCTCCGTATAGCTGGTCAATTTCCACCATAACCACCCTCCAGCGAAAAGTAC GAAAACACAAAGACTTGCCGGGGCCAGATGGCGAACTGCACGCAAATCTGCTG TGCCCTGTAGATGTCTTGGACGTGCCCCAAGGTACTCTGCCCCGACAAACAGTCC ACAGAACAGGCAATCCAACCTCTGAAAAGATGAAAACGAGCGCGTCCCCCTTC TTCTCGCCGTGGGCTACCACAAGCCCCACATCCCGTTTAGATACCCCAAGGAA TTTCAGAACTGTACCCCTTGAAAACATCACTCTCGCGCCCGACCCCGAAGTG CCAGACGGACTCCCTCCTGTTGCCCTACAACCTTGGATGGACATCAGACAACGT GAAGATGTGCAGGCCCTGAACATCTCAGTGCCTTACGGCCCCATTCCAGTTGAC TTCCAGAGGAAGATTCGGCAGTCTTACCTCGCCTCCGTAGTTACCTGGACACC CAAGTGGGTAGACTCCTGAGCGCCTTGGACGATCTCCAGCTCGCAAACAGCACC ATCATTGCCTTCACCAGCGACCATGGTTGGGCGCTGGGTGAACATGGAGAATGG GCTAAATATTCAAATTTTCGACGTTGCGACCCACGTCCCATTTGATCTTCTACGTG CCTGGACGAACAGCCTCCTTGCCCTGAAGCCGGGGAAAAGTTGTTTCCATATCTG GACCTTTTCGATTCTGCGAGCCAACTCATGGAACCTGGGCGACAGAGCATGGAC CTGGTGGAACCTGGTCAGTTTATTTCCAACCTTGGCAGGCCTTGCAGGCCTCCAA GTTCCACCTCGGTGTCCCCTTCCCTCATTCACGTCGAACTCTGTGCGGAAGGT AAAAACCTCCTCAAGCATTTTTCGTTTTTCGGGACCTCGAAGAAGACCCATACCTG CCAGGGAATCCAAGGGAACGATTGCTTACAGCCAGTACCCTAGACCTAGCGAC ATCCACAGTGGAACAGCGACAAGCCCTCCCTCAAGGACATTAAAAATCATGGGT TATAGTATCCGGACTATTGACTACAGGTATACCGTGTGGGTGGGTTTCAACCCA GACGAATTTCTCGCCAATTTCTCCGACATCCACGCGGGCGAACTGTATTTCGTT GATTCCGATCCACTGCAAGATCATAATATGTACAACGATAGTCAAGGGGGTGAC CTCTTCCAGTTGCTAATGCCAGAGGCCGCTGCTAAAGAGGCTGCCGCCAAAGAA GCCGCCGCTAAGGACTCCTCTCACGCCTTACCCTGGACGAGCTGCGGTACTAA</p>	146
I2S-MTfpep	<p>ATGGAATGGAGCTGGGTCTTTCTCTTCTCCTGTCTAGTAACGACTGGTGTCCAC</p>	147

Tabla E2. Secuencias codificantes polinucleotídicas de construcciones de fusión		
Nombre	Secuencia polinucleotídica	SEQ ID NO:
(sin propéptido de I2S)	<p>TCCGACTACAAGGACGACGACGACAAAGAGCAGAAGCTGATCTCCGAAGAGGAC CTGCACCACCATCATCACCATCACCACCATCACGGAGGCGGTGGAGAGAACCTG TACTTTCAGGGCACAGATGCACTCAACGTGCTGCTGATCATCGTAGATGACCTC CGACCTTCTCTGGGCTGTTACGGCGACAAGCTAGTACGGAGCCCCAAACATCGAC CAGCTCGCATCGCACTCTCTCCTATTCCAGAACGCATTGCCCCAGCAGGCTGTC TGTGCTCCCTCCCGAGTGTCTCTCCTCACGGGTCGGAGACCCGATACCCACGAGG TTATATGACTTCAACTCATACTGGCGCGTGCATGCCGGTAACTTTTCTACTATA CCCCAGTATTTTAAAGAAAATGGCTATGTTACAATGTCCGTTGGCAAGGTATTT CATCCTGGTATTAGCAGCAACCACACAGATGACTCTCCGTATAGCTGGTCATTTC CCACCATAACCACCCCTCCAGCGAAAAGTACGAAAACACAAAGACTTGCCGGGGG CCAGATGGCGAACTGCACGCAAAATCTGCTGTGCCCTGTAGATGTCTTGGACGTG CCCGAAGGTACTCTGCCCGACAACAGTCCACAGAACAGGCAATCCAATCCTT GAAAAGATGAAAACGAGCGCGTCCCCCTTCTTCTCGCCGTGGGCTACCACAAG CCCCACATCCCGTTTAGATACCCCAAGGAATTTTCAAGAACTGTACCCCTGGAA AACATCACTCTCGCGCCCGACCCCGAAGTGCCAGACGGACTCCCTCCTGTTGCC TACAACCTTGGATGGACATCAGACAACGTGAAGATGTGCAGGCCCTGAACATC TCAGTGCCCTTACGGCCCCATTCCAGTTGACTTCCAGAGGAAGATTCCGGCAGTCC TACTTCGCCTCCGTTAGTTACCTGGACACCCAAAGTGGGTAGACTCCTGAGCGCC TTGGACGATCTCCAGCTCGCAAACAGCACCATCATTGCCTTCACCAGCGACCAT GGTTGGGCGCTGGGTGAACATGGAGAATGGGCTAAATATTCAAATTTTCGACGTT GCGACCCACGTCCCATTGATCTTCTACGTGCCTGGACGAACAGCCTCCTTGCCCT GAAGCCGGGGAAAAGTTGTTTCCATATCTGGACCCCTTCGATTCTCGGAGCCAA CTCATGGAACCTGGGCGACAGAGCATGGACCTGGTGGAACTGGTCAGTTTATTT CCAACCTTGGCAGGCTTGCAGGCTTCCAAGTTCCACCTCGGTGTCCCGTTCCC TCATTCCACGTGCACTCTGTGCGGAAGGTAAAAACCTCCTCAAGCATTTTTCGT TTTCGGGACCTCGAAGAAGACCCATACCTGCCAGGGAATCCAAGGGAAGTATT GCCTACAGCCAGTACCCTAGACCTAGCGACATCCCACAGTGGAACAGCGACAAG CCCTCCCTCAAGGACATTAAATCATGGGTTATAGTATCCGGACTATTGACTAC AGGTATACCGTGTGGGTGGGTTTCAACCCAGACGAATTTCTCGCCAATTTCTCC GACATCCACGCGGGCGAACTGTATTTTCGTGATTCCGATCCACTGCAAGATCAT AATATGTACAACGATAGTCAAGGGGGTGACCTCTTCCAGTTGCTAATGCCAGAG GCCGCTGCTAAAGAGGCTGCCGCCAAAGAAGCCCGCGCTAAGGACTCCTCTCAC GCCTTCACCCCTGGACGAGCTGCGGTACTAA</p>	

Las proteínas recombinantes se prepararon y sometieron a prueba para determinar la actividad enzimática contra el sustrato 4-Nitrocatecol Sulfato (PNCS) con respecto a IDS humana recombinante y un testigo negativo (la fusión trastuzumab-MTf). Los resultados se muestran en la Figuras 2-4. Se usó un µg de cada muestra en el ensayo de actividad enzimática y los datos presentados están normalizados con respecto al blanco de sustrato.

La Figura 2 muestra la evaluación de la actividad enzimática de las proteínas de fusión I2S-MTf y MTf-I2S según se mide por su capacidad de hidrolizar el sustrato 4-Nitrocatecol Sulfato (PNCS) con respecto a IDS humana recombinante y un testigo negativo (fusión T2M-MTf). Estos datos muestran que las proteínas de fusión I2S-MTf y MTf-I2S no solo tuvieron actividad enzimática significativa, sino que también tuvieron mayor actividad enzimática con respecto a la IDS humana natural (sin fusión).

La Figura 3 muestra la evaluación de la actividad enzimática de las proteínas de fusión MTfpep-I2S y I2S-MTfpep (con propéptido de I2S) según se mide por su capacidad de hidrolizar el sustrato PNCS con respecto a la fusión I2S-MTf y un testigo negativo (fusión T2M-MTf). Estos datos muestran que las proteínas de fusión MTfpep-I2S y I2S-MTfpep no solo tuvieron actividad enzimática significativa, sino que también tuvieron mayor actividad enzimática con respecto a la proteína de fusión I2S-MTf significativamente activa (de la Figura 2) y, por lo tanto, mayor actividad enzimática con respecto a IDS humana natural (sin fusión).

La Figura 4 muestra una comparación de la actividad enzimática de las proteínas de fusión I2S-MTfpep (con propéptido de I2S) e I2S-MTfpep (sin propéptido de I2S) según se mide por su capacidad de hidrolizar el sustrato

PNCS. Estos datos muestran que las proteínas de fusión MTfpep-I2S y I2S-MTfpep no solo tuvieron actividad enzimática significativa, sino que también tuvieron mayor actividad enzimática relativamente a la IDS humana natural (sin fusión).

Ejemplo 2

5 Distribución in vivo de las fusiones I2S-MTf y MTfpep-I2S en el cerebro

La biodistribución en el cerebro de las proteínas de fusión I2S-MTf (ejemplo de referencia) y MTfpep-I2S en ratones se evaluó mediante generación de imágenes por microscopía confocal cuantitativa. Se administraron equivalentes de dosis terapéuticas de I2S-MTf y MTfpep-I2S en volumen de 100 μ L a ratones a través de inyección en la vena de la cola. Antes de la eutanasia, los ratones recibieron por inyección (i.v.) Lectina de tomate-FITC (40 μ g) durante 10 min para teñir la vasculatura cerebral. La sangre se aclaró mediante perfusión intracardiaca de 10 ml de disolución salina heparinizada a una velocidad de 1 ml por minuto. Los cerebros se extirparon y congelaron en OCT y se almacenaron a -80 °C. Los cerebros se montaron en Tissue Tek y se cortaron en secciones con un crióstato a -20 °C. Las secciones se montaron en portaobjetos para microscopio Superfrost Plus, se fijaron en Acetona/MeOH fríos (1:1) durante 10 minutos a temperatura ambiente y se lavaron con PBS. Se montaron cubiertas de vidrio sobre las secciones mediante el uso del reactivo para fluorescencia Prolong Gold con DAPI (sondas moleculares, P36931). Se llevaron a cabo microscopía confocal tridimensional (3D) y análisis cuantitativo.

La Figura 5 muestra la cuantificación de la distribución relativa de las proteínas de fusión MTfpep-I2S (con propéptido) e I2S-MTf entre capilares (C) y el parénquima (P) en el cerebro, con respecto a la señal total (T). La tinción significativa de tejidos parenquimatosos con respecto a los capilares confirma que las proteínas de fusión MTfpep-I2S y I2S-MTf fueron ambas capaces de cruzar la barrera hematoencefálica (BBB) y acumularse en tejidos del sistema nervioso central.

En resumen, los datos de los Ejemplos 1 y 2 muestran que las proteínas de fusión MTfpep-I2S (con propéptido) y I2S-MTf no solo son capaces de cruzar la BBB y acumularse en tejidos del CNS, sino que también tienen actividad enzimática significativamente mayor con respecto a la IDS humana recombinante natural (sin fusión).

25 Ejemplo 3

Actividad in vivo de las fusiones I2S-MTf y MTfpep-I2S en modelo de MPS II en ratón

Se evaluó la eficacia terapéutica de las proteínas de fusión I2S-MTf (ejemplo de referencia) y MTfpep-I2S en un modelo de síndrome de Hunter o Mucopolisacaridosis tipo II (MPS II) en ratón con respecto a Idursulfasa (Elaprase®), que se indica para el tratamiento del síndrome de Hunter. Estos estudios se diseñaron para evaluar el efecto de la administración intravenosa (IV) e intraperitoneal (IP) de las proteínas de fusión en la patología cerebral en un modelo de Mucopolisacaridosis II (MPSII) en ratón con gen inactivado.

Síndrome de Hunter. Según se indicó anteriormente, el síndrome de Hunter es una enfermedad recesiva ligada al cromosoma X causada por niveles insuficientes de la enzima iduronato 2-sulfatasa (IDS) lisosómica. Esta enzima escinde los restos de 2-O-sulfato terminales de los glicosaminoglicanos (GAG) sulfato de dermatán y sulfato de heparán. Debido a la actividad faltante o defectuosa de la enzima IDS en pacientes con síndrome de Hunter, los GAG se acumulan progresivamente en los lisosomas de una variedad de tipos celulares. Esto conduce a congestión celular, organomegalia, destrucción de tejido y disfunción de sistema orgánico.

Modelo en ratón. Los ratones IDS-KO tienen escasa o ninguna actividad de IDS en los tejidos y exhiben muchos de los efectos celulares y clínicos observados en el síndrome de Hunter, incluida mayor vacuolización de tejidos, niveles de GAG y excreción urinaria de GAG. Debido a la naturaleza recesiva, ligada al cromosoma X del síndrome de Hunter, todos los estudios de farmacología se llevaron a cabo en ratones macho. La reproducción animal se lleva a cabo según lo describen Garcia et al, 2007 (3). En resumen, se aparean hembras portadoras con ratones macho naturales de la cepa C57Bl/6, se producen ratones hembra heterogéneos y macho hemicigóticos con gen inactivado, así como hembras y machos naturales (WT, por sus siglas en inglés). Se obtienen ratones macho IDS-KO alternativamente al aparear hembras portadoras con ratones macho IDS-KO. El genotipo de todos los ratones usados en estos experimentos se confirma mediante reacción en cadena de la polimerasa de ADN obtenido a partir de corte en la cola. Todos los ratones IDS-KO son machos IKO (-/0) hemicigóticos de entre 12-13 semanas en el comienzo de la iniciación del tratamiento (no se usan ratones menores de 12 semanas en este estudio). Se usa un grupo de machos WT de la camada (+/0) no tratados como testigos.

Idursulfasa (Elaprase®). La idursulfasa es un fármaco usado para tratar el síndrome de Hunter (también denominado MPS-II) (ver Garcia et al., Mol Genet Metab. 91:183-90, 2007). Es una forma purificada de la enzima lisosómica iduronato-2-sulfatasa y se produce mediante tecnología de ADN recombinante en una línea celular humana. Diseño del estudio. El diseño del estudio se señala en la Tabla E3 a continuación.

Tabla E3						
Grupo	Animal	Ratones/Grupo	Nivel de dosis (mg/kg)	Volumen de dosis (mL/kg)	Régimen de tratamiento	Sacrificio
Vehículo (testigo)	WT	5	0	5-6	IV, una vez por semana durante 6 semanas	24 h después de la última inyección
Vehículo (testigo)	IDS-KO	5	0	5-6	IV, una vez por semana durante 6 semanas	24 h después de la última inyección
IDS (Elaprasa) (dosis alta)	IDS-KO	5	6 mg/kg	5-6	IV, bisemanalmente durante 6 semanas	24 h después de la última inyección
hMTf	IDS-KO	3-5	Equivalente molar a dosis de hMTf-IDS	5-6	IV, una vez por semana durante 6 semanas	24 h después de la última inyección
IDS-hMTf	IDS-KO	5	Equivalente de actividad a IDS (dosis alta)	5-6	IV, una vez por semana durante 6 semanas	24 h después de la última inyección
hMTfpep-IDS	IDS-KO	5	Equivalente de actividad a IDS (dosis alta)	5-6	IV, una vez por semana durante 6 semanas	24 h después de la última inyección

Todos los artículos de prueba y testigos de vehículo se administran mediante dos bolos lentos (uno por inyección IV y uno por IP) que se llevan a cabo una vez por semana durante un total de 6 semanas.

- 5 Los pesos corporales se determinan en la aleatorización durante el primer día de tratamiento y semanalmente a continuación. Se llevan a cabo observaciones clínicas a diario. Los animales se sacrifican aproximadamente 24 horas después del último tratamiento.

- 10 Se recogen los órganos seleccionados (cerebro, hígado, riñón y corazón) y se registran sus pesos. Los cerebros se conservan para análisis de histopatología e inmunotinción. Los otros tejidos se dividen en una mitad u órgano de par y se conservan para histopatología e inmunotinción de manera similar al cerebro. La otra mitad u órgano del par se congela en nitrógeno líquido y se almacena a -80 °C hasta que se someten a ensayo para determinar GAG.

Criterios de valoración del estudio: Los criterios de valoración primarios son los siguientes:

- Evaluación histológica: Tinción con hematoxilina y eosina de secciones del cerebro. Este método se usa para evaluar si el tratamiento tuvo un efecto de reducción de la cantidad/tamaño de vacuolas de almacenamiento celular observadas en ratones IDS-KO; y

- Evaluación inmunohistoquímica de la proteína de membrana asociada lisosómica-1 (LAMP-1) en secciones de cerebro: este método se usa para determinar si el tratamiento tuvo un efecto de reducción de la inmunorreactividad de LAMP-1 elevada que se observa en ratones IDS-KO.

5 Si es factible, también se emplean métodos cualitativos o semicualitativos para el análisis de los criterios de valoración 1-2 (tales como puntuación, mediciones de área, barridos de sección, etc.). El histopatólogo que lleva a cabo este análisis desconoce la asignación de los portaobjetos a los grupos del estudio. El área superficial del lisosoma se cuantifica al barrer áreas teñidas para LAMP1 (IHC) y se compara entre grupos experimentales.

Los criterios de valoración secundarios son los siguientes:

- Niveles de GAG en tejidos seleccionados (hígado, riñón y corazón);
- 10 • Tinción con H&E de tejidos seleccionados y detección de vacuolas de almacenamiento celular; y
- Evaluación inmunohistoquímica de los niveles de LAMP-1 en órganos/tejidos seleccionados.

Histopatología (tinción con H&E). Los tejidos se recogen y fijan en formalina tamponada neutra al 10 %, después se procesan y embeben en parafina. Se preparan y tiñen secciones de parafina de 5 µm con hematoxilina y eosina (H&E) mediante procedimientos estándares.

15 Inmunohistoquímica (LAMP-1). Los portaobjetos desparafinizados se incuban durante toda la noche con IgG anti-LAMP-1 de rata (Santa CruzBiotechnology) como el anticuerpo primario o IgG2a de rata como anticuerpo testigo (AbDSerotec, Raleigh, NC). Después de la incubación durante toda la noche a 2-8 °C, se agrega IgG de conejo anti-rata (H&L) adsorbida por ratón biotinilada (Vector Laboratories). Después de 30 minutos de incubación a 37 °C, las muestras se lavan y después se tratan con complejo avidina-biotina-peroxidasa (Vector Laboratories) durante 30 minutos. La proteína etiquetada se localiza mediante incubación con 3,39-diaminobenzidina. El área de células 20 positivas para LAMP-1 se analiza con el programa informático Image-Pro Plus (Media Cybernetics, Inc., Bethesda, MD).

25 Mediciones de GAG. Se preparan extractos de tejido al homogeneizar tejido en un tampón de lisis (10 mM Tris, 5 mM EDTA, Igepal CA-630 al 0,1 %, 2 mM Pefabloc SC) con un triturador de vidrio (Kontes Glass Company, Vineland, NJ) o un homogeneizador de tejido motorizado (PowerGen Model 125, Omni International, Warrenton, VA). A continuación, los homogenatos se someten a 5 ciclos de congelamiento-descongelamiento usando un baño de etanol/hielo seco y un baño de agua a 37 °C. Los residuos de tejido se granulan dos veces mediante centrifugación a temperatura ambiente a 2000 g durante 12 minutos y los sobrenadantes se recogen y se someten a ensayo para determinar la concentración de proteína total (mg/mL) usando el ensayo de ácido bicínico (BCA, por 30 sus siglas en inglés) (Pierce, Rockford, IL).

La concentración de GAG en orina y extractos de tejido se cuantifica mediante un ensayo colorimétrico usando tinte azul de 1,9-dimetilmetileno (DMB) y una curva estándar (1,56-25 µg/mL) preparada a partir de sulfato de dermatán (MP Biomedicals, Aurora, OH). Las muestras de orina se procesan en diluciones de 1/10, 1/20 y 1/40. Para evitar la interferencia de la proteína en el ensayo, las muestras de extracto de tejido se diluyen hasta concentraciones de 35 proteína <200 µg / mL. Las concentraciones de GAG en la orina se ajustan según las concentraciones de creatinina medidas con un kit disponible comercialmente (Sigma, St. Louis, MO, parte núm. 555A) para compensar las diferencias en la función renal y se expresan como µg de GAG/mg de creatinina. Los niveles de GAG en extractos de tejido se ajustan según la concentración de proteína (µg de GAG/mg de proteína) o gramos de tejido.

LISTA DE SECUENCIAS

<110> BiOasis Technologies, Inc.
 Vitalis, Timothy
 Gabathuler, Reinhard

5

<120> PROTEÍNAS DE FUSIÓN DE P97-IDS

<130> BIOA-009/01WO

10

<150> US 61/941,896

<151> 19-02-2014

<160> 149

15

<170> PatentIn versión 3.5

<210> 1

<211> 738

<212> PRT

20

<213> Homo sapiens

<400> 1

Met Arg Gly Pro Ser Gly Ala Leu Trp Leu Leu Leu Ala Leu Arg Thr
 1 5 10 15

Val Leu Gly Gly Met Glu Val Arg Trp Cys Ala Thr Ser Asp Pro Glu
 20 25 30

Gln His Lys Cys Gly Asn Met Ser Glu Ala Phe Arg Glu Ala Gly Ile
 35 40 45

Gln Pro Ser Leu Leu Cys Val Arg Gly Thr Ser Ala Asp His Cys Val
 50 55 60

Gln Leu Ile Ala Ala Gln Glu Ala Asp Ala Ile Thr Leu Asp Gly Gly
 65 70 75 80

Ala Ile Tyr Glu Ala Gly Lys Glu His Gly Leu Lys Pro Val Val Gly
 85 90 95

Glu Val Tyr Asp Gln Glu Val Gly Thr Ser Tyr Tyr Ala Val Ala Val
 100 105 110

Val Arg Arg Ser Ser His Val Thr Ile Asp Thr Leu Lys Gly Val Lys
 115 120 125

Ser Cys His Thr Gly Ile Asn Arg Thr Val Gly Trp Asn Val Pro Val
 130 135 140

Gly Tyr Leu Val Glu Ser Gly Arg Leu Ser Val Met Gly Cys Asp Val
 145 150 155 160

ES 2 762 672 T3

Leu Lys Ala Val Ser Asp Tyr Phe Gly Gly Ser Cys Val Pro Gly Ala
 165 170 175
 Gly Glu Thr Ser Tyr Ser Glu Ser Leu Cys Arg Leu Cys Arg Gly Asp
 180 185 190
 Ser Ser Gly Glu Gly Val Cys Asp Lys Ser Pro Leu Glu Arg Tyr Tyr
 195 200 205
 Asp Tyr Ser Gly Ala Phe Arg Cys Leu Ala Glu Gly Ala Gly Asp Val
 210 215 220
 Ala Phe Val Lys His Ser Thr Val Leu Glu Asn Thr Asp Gly Lys Thr
 225 230 235 240
 Leu Pro Ser Trp Gly Gln Ala Leu Leu Ser Gln Asp Phe Glu Leu Leu
 245 250 255
 Cys Arg Asp Gly Ser Arg Ala Asp Val Thr Glu Trp Arg Gln Cys His
 260 265 270
 Leu Ala Arg Val Pro Ala His Ala Val Val Val Arg Ala Asp Thr Asp
 275 280 285
 Gly Gly Leu Ile Phe Arg Leu Leu Asn Glu Gly Gln Arg Leu Phe Ser
 290 295 300
 His Glu Gly Ser Ser Phe Gln Met Phe Ser Ser Glu Ala Tyr Gly Gln
 305 310 315 320
 Lys Asp Leu Leu Phe Lys Asp Ser Thr Ser Glu Leu Val Pro Ile Ala
 325 330 335
 Thr Gln Thr Tyr Glu Ala Trp Leu Gly His Glu Tyr Leu His Ala Met
 340 345 350
 Lys Gly Leu Leu Cys Asp Pro Asn Arg Leu Pro Pro Tyr Leu Arg Trp
 355 360 365
 Cys Val Leu Ser Thr Pro Glu Ile Gln Lys Cys Gly Asp Met Ala Val
 370 375 380
 Ala Phe Arg Arg Gln Arg Leu Lys Pro Glu Ile Gln Cys Val Ser Ala
 385 390 395 400
 Lys Ser Pro Gln His Cys Met Glu Arg Ile Gln Ala Glu Gln Val Asp
 405 410 415

ES 2 762 672 T3

Ala Val Thr Leu Ser Gly Glu Asp Ile Tyr Thr Ala Gly Lys Thr Tyr
420 425 430

Gly Leu Val Pro Ala Ala Gly Glu His Tyr Ala Pro Glu Asp Ser Ser
435 440 445

Asn Ser Tyr Tyr Val Val Ala Val Val Arg Arg Asp Ser Ser His Ala
450 455 460

Phe Thr Leu Asp Glu Leu Arg Gly Lys Arg Ser Cys His Ala Gly Phe
465 470 475 480

Gly Ser Pro Ala Gly Trp Asp Val Pro Val Gly Ala Leu Ile Gln Arg
485 490 495

Gly Phe Ile Arg Pro Lys Asp Cys Asp Val Leu Thr Ala Val Ser Glu
500 505 510

Phe Phe Asn Ala Ser Cys Val Pro Val Asn Asn Pro Lys Asn Tyr Pro
515 520 525

Ser Ser Leu Cys Ala Leu Cys Val Gly Asp Glu Gln Gly Arg Asn Lys
530 535 540

Cys Val Gly Asn Ser Gln Glu Arg Tyr Tyr Gly Tyr Arg Gly Ala Phe
545 550 555 560

Arg Cys Leu Val Glu Asn Ala Gly Asp Val Ala Phe Val Arg His Thr
565 570 575

Thr Val Phe Asp Asn Thr Asn Gly His Asn Ser Glu Pro Trp Ala Ala
580 585 590

Glu Leu Arg Ser Glu Asp Tyr Glu Leu Leu Cys Pro Asn Gly Ala Arg
595 600 605

Ala Glu Val Ser Gln Phe Ala Ala Cys Asn Leu Ala Gln Ile Pro Pro
610 615 620

His Ala Val Met Val Arg Pro Asp Thr Asn Ile Phe Thr Val Tyr Gly
625 630 635 640

Leu Leu Asp Lys Ala Gln Asp Leu Phe Gly Asp Asp His Asn Lys Asn
645 650 655

Gly Phe Lys Met Phe Asp Ser Ser Asn Tyr His Gly Gln Asp Leu Leu

ES 2 762 672 T3

660

665

670

Phe Lys Asp Ala Thr Val Arg Ala Val Pro Val Gly Glu Lys Thr Thr
675 680 685

Tyr Arg Gly Trp Leu Gly Leu Asp Tyr Val Ala Ala Leu Glu Gly Met
690 695 700

Ser Ser Gln Gln Cys Ser Gly Ala Ala Ala Pro Ala Pro Gly Ala Pro
705 710 715 720

Leu Leu Pro Leu Leu Leu Pro Ala Leu Ala Ala Arg Leu Leu Pro Pro
725 730 735

Ala Leu

<210> 2

<211> 692

<212> PRT

<213> Homo sapiens

<400> 2

Gly Met Glu Val Arg Trp Cys Ala Thr Ser Asp Pro Glu Gln His Lys
1 5 10 15

Cys Gly Asn Met Ser Glu Ala Phe Arg Glu Ala Gly Ile Gln Pro Ser
20 25 30

Leu Leu Cys Val Arg Gly Thr Ser Ala Asp His Cys Val Gln Leu Ile
35 40 45

Ala Ala Gln Glu Ala Asp Ala Ile Thr Leu Asp Gly Gly Ala Ile Tyr
50 55 60

Glu Ala Gly Lys Glu His Gly Leu Lys Pro Val Val Gly Glu Val Tyr
65 70 75 80

Asp Gln Glu Val Gly Thr Ser Tyr Tyr Ala Val Ala Val Val Arg Arg
85 90 95

Ser Ser His Val Thr Ile Asp Thr Leu Lys Gly Val Lys Ser Cys His
100 105 110

Thr Gly Ile Asn Arg Thr Val Gly Trp Asn Val Pro Val Gly Tyr Leu
115 120 125

Val Glu Ser Gly Arg Leu Ser Val Met Gly Cys Asp Val Leu Lys Ala

5

10

ES 2 762 672 T3

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

ES 2 762 672 T3

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

ES 2 762 672 T3

Met Phe Asp Ser Ser Asn Tyr His Gly Gln Asp Leu Leu Phe Lys Asp
645 650 655

Ala Thr Val Arg Ala Val Pro Val Gly Glu Lys Thr Thr Tyr Arg Gly
660 665 670

Trp Leu Gly Leu Asp Tyr Val Ala Ala Leu Glu Gly Met Ser Ser Gln
675 680 685

Gln Cys Ser Gly
690

<210> 3

<211> 11

5 <212> PRT

<213> Homo sapiens

<400> 3

Trp Cys Ala Thr Ser Asp Pro Glu Gln His Lys

1 5 10

10

<210> 4

<211> 11

<212> PRT

<213> Homo sapiens

15

<400> 4

Arg Ser Ser His Val Thr Ile Asp Thr Leu Lys

1 5 10

20

<210> 5

<211> 13

<212> PRT

<213> Homo sapiens

<400> 5

Ser Ser His Val Thr Ile Asp Thr Leu Lys Gly Val Lys

1 5 10

25

<210> 6

<211> 14

<212> PRT

<213> Homo sapiens

30

<400> 6

Leu Cys Arg Gly Asp Ser Ser Gly Glu Gly Val Cys Asp Lys

1 5 10

35

<210> 7

<211> 16

<212> PRT

<213> Homo sapiens

40

<400> 7

Gly Asp Ser Ser Gly Glu Gly Val Cys Asp Lys Ser Pro Leu Glu Arg

1 5 10 15

45

<210> 8

<211> 9

<212> PRT

<213> Homo sapiens

<400> 8
 Tyr Tyr Asp Tyr Ser Gly Ala Phe Arg
 1 5

5 <210> 9
 <211> 7
 <212> PRT
 <213> Homo sapiens

10 <400> 9
 Ala Asp Val Thr Glu Trp Arg
 1 5

15 <210> 10
 <211> 9
 <212> PRT
 <213> Homo sapiens

20 <400> 10
 Val Pro Ala His Ala Val Val Val Arg
 1 5

25 <210> 11
 <211> 10
 <212> PRT
 <213> Homo sapiens

30 <400> 11
 Ala Asp Thr Asp Gly Gly Leu Ile Phe Arg
 1 5 10

35 <210> 12
 <211> 9
 <212> PRT
 <213> Homo sapiens

40 <400> 12
 Cys Gly Asp Met Ala Val Ala Phe Arg
 1 5

45 <210> 13
 <211> 11
 <212> PRT
 <213> Homo sapiens

50 <400> 13
 Leu Lys Pro Glu Ile Gln Cys Val Ser Ala Lys
 1 5 10

55 <210> 14
 <211> 12
 <212> PRT
 <213> Homo sapiens

60 <400> 14
 Asp Ser Ser His Ala Phe Thr Leu Asp Glu Leu Arg
 1 5 10

65 <210> 15
 <211> 13
 <212> PRT
 <213> Homo sapiens

70 <400> 15
 Ser Glu Asp Tyr Glu Leu Leu Cys Pro Asn Gly Ala Arg
 1 5 10

ES 2 762 672 T3

<210> 16
 <211> 15
 <212> PRT
 <213> Homo sapiens
 5
 <400> 16
 Ala Gln Asp Leu Phe Gly Asp Asp His Asn Lys Asn Gly Phe Lys
 1 5 10 15
 <210> 17
 10 <211> 40
 <212> PRT
 <213> Homo sapiens
 <400> 17
 Phe Ser Ser Glu Ala Tyr Gly Gln Lys Asp Leu Leu Phe Lys Asp Ser
 1 5 10 15
 Thr Ser Glu Leu Val Pro Ile Ala Thr Gln Thr Tyr Glu Ala Trp Leu
 20 25 30
 Gly His Glu Tyr Leu His Ala Met
 15 35 40
 <210> 18
 <211> 221
 <212> PRT
 20 <213> Homo sapiens
 <400> 18

ES 2 762 672 T3

Glu Arg Ile Gln Ala Glu Gln Val Asp Ala Val Thr Leu Ser Gly Glu
 1 5 10 15
 Asp Ile Tyr Thr Ala Gly Lys Thr Tyr Gly Leu Val Pro Ala Ala Gly
 20 25 30
 Glu His Tyr Ala Pro Glu Asp Ser Ser Asn Ser Tyr Tyr Val Val Ala
 35 40 45
 Val Val Arg Arg Asp Ser Ser His Ala Phe Thr Leu Asp Glu Leu Arg
 50 55 60
 Gly Lys Arg Ser Cys His Ala Gly Phe Gly Ser Pro Ala Gly Trp Asp
 65 70 75 80
 Val Pro Val Gly Ala Leu Ile Gln Arg Gly Phe Ile Arg Pro Lys Asp
 85 90 95
 Cys Asp Val Leu Thr Ala Val Ser Glu Phe Phe Asn Ala Ser Cys Val
 100 105 110
 Pro Val Asn Asn Pro Lys Asn Tyr Pro Ser Ser Leu Cys Ala Leu Cys
 115 120 125
 Val Gly Asp Glu Gln Gly Arg Asn Lys Cys Val Gly Asn Ser Gln Glu
 130 135 140
 Arg Tyr Tyr Gly Tyr Arg Gly Ala Phe Arg Cys Leu Val Glu Asn Ala
 145 150 155 160
 Gly Asp Val Ala Phe Val Arg His Thr Thr Val Phe Asp Asn Thr Asn
 165 170 175
 Gly His Asn Ser Glu Pro Trp Ala Ala Glu Leu Arg Ser Glu Asp Tyr
 180 185 190
 Glu Leu Leu Cys Pro Asn Gly Ala Arg Ala Glu Val Ser Gln Phe Ala
 195 200 205
 Ala Cys Asn Leu Ala Gln Ile Pro Pro His Ala Val Met
 210 215 220

<210> 19
 <211> 32
 <212> PRT
 <213> Homo sapiens

<400> 19
 Val Arg Pro Asp Thr Asn Ile Phe Thr Val Tyr Gly Leu Leu Asp Lys
 1 5 10 15

Ala Gln Asp Leu Phe Gly Asp Asp His Asn Lys Asn Gly Phe Lys Met
 20 25 30

<210> 20
 <211> 564
 <212> PRT

ES 2 762 672 T3

<213> Homo sapiens

<400> 20

Gly Met Glu Val Arg Trp Cys Ala Thr Ser Asp Pro Glu Gln His Lys
1 5 10 15

Cys Gly Asn Met Ser Glu Ala Phe Arg Glu Ala Gly Ile Gln Pro Ser
20 25 30

Leu Leu Cys Val Arg Gly Thr Ser Ala Asp His Cys Val Gln Leu Ile
35 40 45

Ala Ala Gln Glu Ala Asp Ala Ile Thr Leu Asp Gly Gly Ala Ile Tyr
50 55 60

Glu Ala Gly Lys Glu His Gly Leu Lys Pro Val Val Gly Glu Val Tyr
65 70 75 80

Asp Gln Glu Val Gly Thr Ser Tyr Tyr Ala Val Ala Val Val Arg Arg
85 90 95

Ser Ser His Val Thr Ile Asp Thr Leu Lys Gly Val Lys Ser Cys His
100 105 110

Thr Gly Ile Asn Arg Thr Val Gly Trp Asn Val Pro Val Gly Tyr Leu
115 120 125

Val Glu Ser Gly Arg Leu Ser Val Met Gly Cys Asp Val Leu Lys Ala
130 135 140

Val Ser Asp Tyr Phe Gly Gly Ser Cys Val Pro Gly Ala Gly Glu Thr
145 150 155 160

5

ES 2 762 672 T3

Ser Tyr Ser Glu Ser Leu Cys Arg Leu Cys Arg Gly Asp Ser Ser Gly
 165 170 175
 Glu Gly Val Cys Asp Lys Ser Pro Leu Glu Arg Tyr Tyr Asp Tyr Ser
 180 185 190
 Gly Ala Phe Arg Cys Leu Ala Glu Gly Ala Gly Asp Val Ala Phe Val
 195 200 205
 Lys His Ser Thr Val Leu Glu Asn Thr Asp Gly Lys Thr Leu Pro Ser
 210 215 220
 Trp Gly Gln Ala Leu Leu Ser Gln Asp Phe Glu Leu Leu Cys Arg Asp
 225 230 235 240
 Gly Ser Arg Ala Asp Val Thr Glu Trp Arg Gln Cys His Leu Ala Arg
 245 250 255
 Val Pro Ala His Ala Val Val Val Arg Ala Asp Thr Asp Gly Gly Leu
 260 265 270
 Ile Phe Arg Leu Leu Asn Glu Gly Gln Arg Leu Phe Ser His Glu Gly
 275 280 285
 Ser Ser Phe Gln Met Phe Ser Ser Glu Ala Tyr Gly Gln Lys Asp Leu
 290 295 300
 Leu Phe Lys Asp Ser Thr Ser Glu Leu Val Pro Ile Ala Thr Gln Thr
 305 310 315 320
 Tyr Glu Ala Trp Leu Gly His Glu Tyr Leu His Ala Met Lys Gly Leu
 325 330 335
 Leu Cys Asp Pro Asn Arg Leu Pro Pro Tyr Leu Arg Trp Cys Val Leu
 340 345 350
 Ser Thr Pro Glu Ile Gln Lys Cys Gly Asp Met Ala Val Ala Phe Arg
 355 360 365
 Arg Gln Arg Leu Lys Pro Glu Ile Gln Cys Val Ser Ala Lys Ser Pro
 370 375 380
 Gln His Cys Met Glu Arg Ile Gln Ala Glu Gln Val Asp Ala Val Thr
 385 390 395 400
 Leu Ser Gly Glu Asp Ile Tyr Thr Ala Gly Lys Thr Tyr Gly Leu Val
 405 410 415

ES 2 762 672 T3

Pro Ala Ala Gly Glu His Tyr Ala Pro Glu Asp Ser Ser Asn Ser Tyr
420 425 430

Tyr Val Val Ala Val Val Arg Arg Asp Ser Ser His Ala Phe Thr Leu
435 440 445

Asp Glu Leu Arg Gly Lys Arg Ser Cys His Ala Gly Phe Gly Ser Pro
450 455 460

Ala Gly Trp Asp Val Pro Val Gly Ala Leu Ile Gln Arg Gly Phe Ile
465 470 475 480

Arg Pro Lys Asp Cys Asp Val Leu Thr Ala Val Ser Glu Phe Phe Asn
485 490 495

Ala Ser Cys Val Pro Val Asn Asn Pro Lys Asn Tyr Pro Ser Ser Leu
500 505 510

Cys Ala Leu Cys Val Gly Asp Glu Gln Gly Arg Asn Lys Cys Val Gly
515 520 525

Asn Ser Gln Glu Arg Tyr Tyr Gly Tyr Arg Gly Ala Phe Arg Cys Leu
530 535 540

Val Glu Asn Ala Gly Asp Val Ala Phe Val Arg His Thr Thr Val Phe
545 550 555 560

Asp Asn Thr Asn

<210> 21

<211> 22

5 <212> PRT

<213> Homo sapiens

<400> 21

Gly His Asn Ser Glu Pro Trp Ala Ala Glu Leu Arg Ser Glu Asp Tyr
1 5 10 15

Glu Leu Leu Cys Pro Asn
20

10

<210> 22

<211> 51

<212> PRT

<213> Homo sapiens

15

<400> 22

Gly Ala Arg Ala Glu Val Ser Gln Phe Ala Ala Cys Asn Leu Ala Gln
1 5 10 15

Ile Pro Pro His Ala Val Met Val Arg Pro Asp Thr Asn Ile Phe Thr
20 25 30

Val Tyr Gly Leu Leu Asp Lys Ala Gln Asp Leu Phe Gly Asp Asp His
35 40 45

Asn Lys Asn
50

ES 2 762 672 T3

<210> 23
 <211> 53
 <212> PRT
 5 <213> Homo sapiens

 <400> 23
 Gly Phe Lys Met Phe Asp Ser Ser Asn Tyr His Gly Gln Asp Leu Leu
 1 5 10 15

 Phe Lys Asp Ala Thr Val Arg Ala Val Pro Val Gly Glu Lys Thr Thr
 20 25 30

 Tyr Arg Gly Trp Leu Gly Leu Asp Tyr Val Ala Ala Leu Glu Gly Met
 35 40 45

 Ser Ser Gln Gln Cys
 50
 10 <210> 24
 <211> 586
 <212> PRT
 <213> Homo sapiens

 15 <400> 24
 Gly Met Glu Val Arg Trp Cys Ala Thr Ser Asp Pro Glu Gln His Lys
 1 5 10 15

 Cys Gly Asn Met Ser Glu Ala Phe Arg Glu Ala Gly Ile Gln Pro Ser
 20 25 30

 Leu Leu Cys Val Arg Gly Thr Ser Ala Asp His Cys Val Gln Leu Ile
 35 40 45

 Ala Ala Gln Glu Ala Asp Ala Ile Thr Leu Asp Gly Gly Ala Ile Tyr
 50 55 60

 Glu Ala Gly Lys Glu His Gly Leu Lys Pro Val Val Gly Glu Val Tyr

ES 2 762 672 T3

65				70				75				80			
Asp	Gln	Glu	Val	Gly 85	Thr	Ser	Tyr	Tyr	Ala 90	Val	Ala	Val	Val	Arg 95	Arg
Ser	Ser	His	Val 100	Thr	Ile	Asp	Thr	Leu 105	Lys	Gly	Val	Lys	Ser 110	Cys	His
Thr	Gly	Ile 115	Asn	Arg	Thr	Val	Gly 120	Trp	Asn	Val	Pro	Val 125	Gly	Tyr	Leu
Val	Glu 130	Ser	Gly	Arg	Leu	Ser 135	Val	Met	Gly	Cys	Asp 140	Val	Leu	Lys	Ala
Val 145	Ser	Asp	Tyr	Phe	Gly 150	Gly	Ser	Cys	Val	Pro 155	Gly	Ala	Gly	Glu	Thr 160
Ser	Tyr	Ser	Glu	Ser 165	Leu	Cys	Arg	Leu	Cys 170	Arg	Gly	Asp	Ser	Ser 175	Gly
Glu	Gly	Val	Cys 180	Asp	Lys	Ser	Pro	Leu 185	Glu	Arg	Tyr	Tyr	Asp 190	Tyr	Ser
Gly	Ala	Phe 195	Arg	Cys	Leu	Ala	Glu 200	Gly	Ala	Gly	Asp	Val 205	Ala	Phe	Val
Lys	His 210	Ser	Thr	Val	Leu	Glu 215	Asn	Thr	Asp	Gly	Lys 220	Thr	Leu	Pro	Ser
Trp 225	Gly	Gln	Ala	Leu	Leu 230	Ser	Gln	Asp	Phe	Glu 235	Leu	Leu	Cys	Arg	Asp 240
Gly	Ser	Arg	Ala	Asp 245	Val	Thr	Glu	Trp	Arg 250	Gln	Cys	His	Leu	Ala 255	Arg
Val	Pro	Ala	His 260	Ala	Val	Val	Val	Arg 265	Ala	Asp	Thr	Asp	Gly 270	Gly	Leu
Ile	Phe	Arg 275	Leu	Leu	Asn	Glu	Gly 280	Gln	Arg	Leu	Phe	Ser 285	His	Glu	Gly
Ser	Ser	Phe 290	Gln	Met	Phe	Ser 295	Ser	Glu	Ala	Tyr	Gly 300	Gln	Lys	Asp	Leu
Leu 305	Phe	Lys	Asp	Ser	Thr 310	Ser	Glu	Leu	Val	Pro 315	Ile	Ala	Thr	Gln	Thr 320

ES 2 762 672 T3

Tyr Glu Ala Trp Leu Gly His Glu Tyr Leu His Ala Met Lys Gly Leu
 325 330 335
 Leu Cys Asp Pro Asn Arg Leu Pro Pro Tyr Leu Arg Trp Cys Val Leu
 340 345 350
 Ser Thr Pro Glu Ile Gln Lys Cys Gly Asp Met Ala Val Ala Phe Arg
 355 360 365
 Arg Gln Arg Leu Lys Pro Glu Ile Gln Cys Val Ser Ala Lys Ser Pro
 370 375 380
 Gln His Cys Met Glu Arg Ile Gln Ala Glu Gln Val Asp Ala Val Thr
 385 390 395 400
 Leu Ser Gly Glu Asp Ile Tyr Thr Ala Gly Lys Thr Tyr Gly Leu Val
 405 410 415
 Pro Ala Ala Gly Glu His Tyr Ala Pro Glu Asp Ser Ser Asn Ser Tyr
 420 425 430
 Tyr Val Val Ala Val Val Arg Arg Asp Ser Ser His Ala Phe Thr Leu
 435 440 445
 Asp Glu Leu Arg Gly Lys Arg Ser Cys His Ala Gly Phe Gly Ser Pro
 450 455 460
 Ala Gly Trp Asp Val Pro Val Gly Ala Leu Ile Gln Arg Gly Phe Ile
 465 470 475 480
 Arg Pro Lys Asp Cys Asp Val Leu Thr Ala Val Ser Glu Phe Phe Asn
 485 490 495
 Ala Ser Cys Val Pro Val Asn Asn Pro Lys Asn Tyr Pro Ser Ser Leu
 500 505 510
 Cys Ala Leu Cys Val Gly Asp Glu Gln Gly Arg Asn Lys Cys Val Gly
 515 520 525
 Asn Ser Gln Glu Arg Tyr Tyr Gly Tyr Arg Gly Ala Phe Arg Cys Leu
 530 535 540
 Val Glu Asn Ala Gly Asp Val Ala Phe Val Arg His Thr Thr Val Phe
 545 550 555 560
 Asp Asn Thr Asn Gly His Asn Ser Glu Pro Trp Ala Ala Glu Leu Arg
 565 570 575
 Ser Glu Asp Tyr Glu Leu Leu Cys Pro Asn
 580 585

5 <210> 25
 <211> 637
 <212> PRT
 <213> Homo sapiens

10 <400> 25

ES 2 762 672 T3

Gly	Met	Glu	Val	Arg	Trp	Cys	Ala	Thr	Ser	Asp	Pro	Glu	Gln	His	Lys	1	5	10	15
Cys	Gly	Asn	Met	Ser	Glu	Ala	Phe	Arg	Glu	Ala	Gly	Ile	Gln	Pro	Ser	20	25	30	
Leu	Leu	Cys	Val	Arg	Gly	Thr	Ser	Ala	Asp	His	Cys	Val	Gln	Leu	Ile	35	40	45	
Ala	Ala	Gln	Glu	Ala	Asp	Ala	Ile	Thr	Leu	Asp	Gly	Gly	Ala	Ile	Tyr	50	55	60	
Glu	Ala	Gly	Lys	Glu	His	Gly	Leu	Lys	Pro	Val	Val	Gly	Glu	Val	Tyr	65	70	75	80
Asp	Gln	Glu	Val	Gly	Thr	Ser	Tyr	Tyr	Ala	Val	Ala	Val	Val	Arg	Arg	85	90	95	
Ser	Ser	His	Val	Thr	Ile	Asp	Thr	Leu	Lys	Gly	Val	Lys	Ser	Cys	His	100	105	110	
Thr	Gly	Ile	Asn	Arg	Thr	Val	Gly	Trp	Asn	Val	Pro	Val	Gly	Tyr	Leu	115	120	125	
Val	Glu	Ser	Gly	Arg	Leu	Ser	Val	Met	Gly	Cys	Asp	Val	Leu	Lys	Ala	130	135	140	
Val	Ser	Asp	Tyr	Phe	Gly	Gly	Ser	Cys	Val	Pro	Gly	Ala	Gly	Glu	Thr	145	150	155	160
Ser	Tyr	Ser	Glu	Ser	Leu	Cys	Arg	Leu	Cys	Arg	Gly	Asp	Ser	Ser	Gly	165	170	175	
Glu	Gly	Val	Cys	Asp	Lys	Ser	Pro	Leu	Glu	Arg	Tyr	Tyr	Asp	Tyr	Ser	180	185	190	
Gly	Ala	Phe	Arg	Cys	Leu	Ala	Glu	Gly	Ala	Gly	Asp	Val	Ala	Phe	Val	195	200	205	

ES 2 762 672 T3

Lys His Ser Thr Val Leu Glu Asn Thr Asp Gly Lys Thr Leu Pro Ser
 210 215 220
 Trp Gly Gln Ala Leu Leu Ser Gln Asp Phe Glu Leu Leu Cys Arg Asp
 225 230 235 240
 Gly Ser Arg Ala Asp Val Thr Glu Trp Arg Gln Cys His Leu Ala Arg
 245 250 255
 Val Pro Ala His Ala Val Val Val Arg Ala Asp Thr Asp Gly Gly Leu
 260 265 270
 Ile Phe Arg Leu Leu Asn Glu Gly Gln Arg Leu Phe Ser His Glu Gly
 275 280 285
 Ser Ser Phe Gln Met Phe Ser Ser Glu Ala Tyr Gly Gln Lys Asp Leu
 290 295 300
 Leu Phe Lys Asp Ser Thr Ser Glu Leu Val Pro Ile Ala Thr Gln Thr
 305 310 315 320
 Tyr Glu Ala Trp Leu Gly His Glu Tyr Leu His Ala Met Lys Gly Leu
 325 330 335
 Leu Cys Asp Pro Asn Arg Leu Pro Pro Tyr Leu Arg Trp Cys Val Leu
 340 345 350
 Ser Thr Pro Glu Ile Gln Lys Cys Gly Asp Met Ala Val Ala Phe Arg
 355 360 365
 Arg Gln Arg Leu Lys Pro Glu Ile Gln Cys Val Ser Ala Lys Ser Pro
 370 375 380
 Gln His Cys Met Glu Arg Ile Gln Ala Glu Gln Val Asp Ala Val Thr
 385 390 395 400
 Leu Ser Gly Glu Asp Ile Tyr Thr Ala Gly Lys Thr Tyr Gly Leu Val
 405 410 415
 Pro Ala Ala Gly Glu His Tyr Ala Pro Glu Asp Ser Ser Asn Ser Tyr
 420 425 430
 Tyr Val Val Ala Val Val Arg Arg Asp Ser Ser His Ala Phe Thr Leu
 435 440 445
 Asp Glu Leu Arg Gly Lys Arg Ser Cys His Ala Gly Phe Gly Ser Pro
 450 455 460

ES 2 762 672 T3

Ala Gly Trp Asp Val Pro Val Gly Ala Leu Ile Gln Arg Gly Phe Ile
465 470 475 480

Arg Pro Lys Asp Cys Asp Val Leu Thr Ala Val Ser Glu Phe Phe Asn
485 490 495

Ala Ser Cys Val Pro Val Asn Asn Pro Lys Asn Tyr Pro Ser Ser Leu
500 505 510

Cys Ala Leu Cys Val Gly Asp Glu Gln Gly Arg Asn Lys Cys Val Gly
515 520 525

Asn Ser Gln Glu Arg Tyr Tyr Gly Tyr Arg Gly Ala Phe Arg Cys Leu
530 535 540

Val Glu Asn Ala Gly Asp Val Ala Phe Val Arg His Thr Thr Val Phe
545 550 555 560

Asp Asn Thr Asn Gly His Asn Ser Glu Pro Trp Ala Ala Glu Leu Arg
565 570 575

Ser Glu Asp Tyr Glu Leu Leu Cys Pro Asn Gly Ala Arg Ala Glu Val
580 585 590

Ser Gln Phe Ala Ala Cys Asn Leu Ala Gln Ile Pro Pro His Ala Val
595 600 605

Met Val Arg Pro Asp Thr Asn Ile Phe Thr Val Tyr Gly Leu Leu Asp
610 615 620

Lys Ala Gln Asp Leu Phe Gly Asp Asp His Asn Lys Asn
625 630 635

<210> 26

<211> 73

5 <212> PRT

<213> Homo sapiens

<400> 26

Gly His Asn Ser Glu Pro Trp Ala Ala Glu Leu Arg Ser Glu Asp Tyr
1 5 10 15

Glu Leu Leu Cys Pro Asn Gly Ala Arg Ala Glu Val Ser Gln Phe Ala
20 25 30

Ala Cys Asn Leu Ala Gln Ile Pro Pro His Ala Val Met Val Arg Pro
35 40 45

10

Asp Thr Asn Ile Phe Thr Val Tyr Gly Leu Leu Asp Lys Ala Gln Asp
50 55 60

Leu Phe Gly Asp Asp His Asn Lys Asn
65 70

<210> 27

<211> 126

15 <212> PRT

<213> Homo sapiens

ES 2 762 672 T3

<400> 27

Gly His Asn Ser Glu Pro Trp Ala Ala Glu Leu Arg Ser Glu Asp Tyr
1 5 10 15

Glu Leu Leu Cys Pro Asn Gly Ala Arg Ala Glu Val Ser Gln Phe Ala
20 25 30

Ala Cys Asn Leu Ala Gln Ile Pro Pro His Ala Val Met Val Arg Pro
35 40 45

Asp Thr Asn Ile Phe Thr Val Tyr Gly Leu Leu Asp Lys Ala Gln Asp
50 55 60

Leu Phe Gly Asp Asp His Asn Lys Asn Gly Phe Lys Met Phe Asp Ser
65 70 75 80

Ser Asn Tyr His Gly Gln Asp Leu Leu Phe Lys Asp Ala Thr Val Arg
85 90 95

Ala Val Pro Val Gly Glu Lys Thr Thr Tyr Arg Gly Trp Leu Gly Leu
100 105 110

Asp Tyr Val Ala Ala Leu Glu Gly Met Ser Ser Gln Gln Cys
115 120 125

5 <210> 28
<211> 104
<212> PRT
<213> Homo sapiens

10 <400> 28
Gly Ala Arg Ala Glu Val Ser Gln Phe Ala Ala Cys Asn Leu Ala Gln
1 5 10 15

Ile Pro Pro His Ala Val Met Val Arg Pro Asp Thr Asn Ile Phe Thr
20 25 30

Val Tyr Gly Leu Leu Asp Lys Ala Gln Asp Leu Phe Gly Asp Asp His
35 40 45

Asn Lys Asn Gly Phe Lys Met Phe Asp Ser Ser Asn Tyr His Gly Gln
50 55 60

Asp Leu Leu Phe Lys Asp Ala Thr Val Arg Ala Val Pro Val Gly Glu
65 70 75 80

Lys Thr Thr Tyr Arg Gly Trp Leu Gly Leu Asp Tyr Val Ala Ala Leu
85 90 95

Glu Gly Met Ser Ser Gln Gln Cys
100

15 <210> 29
<211> 1272
<212> PRT
<213> Secuencia Artificial

ES 2 762 672 T3

<220>

<223> Proteína de fusión p97

<400> 29

Met Pro Pro Pro Arg Thr Gly Arg Gly Leu Leu Trp Leu Gly Leu Val
1 5 10 15

Leu Ser Ser Val Cys Val Ala Leu Gly His His His His His His His
20 25 30

His His His Glu Asn Leu Tyr Phe Gln Ser Glu Thr Gln Ala Asn Ser
35 40 45

Thr Thr Asp Ala Leu Asn Val Leu Leu Ile Ile Val Asp Asp Leu Arg
50 55 60

Pro Ser Leu Gly Cys Tyr Gly Asp Lys Leu Val Arg Ser Pro Asn Ile
65 70 75 80

Asp Gln Leu Ala Ser His Ser Leu Leu Phe Gln Asn Ala Phe Ala Gln
85 90 95

Gln Ala Val Cys Ala Pro Ser Arg Val Ser Phe Leu Thr Gly Arg Arg
100 105 110

Pro Asp Thr Thr Arg Leu Tyr Asp Phe Asn Ser Tyr Trp Arg Val His
115 120 125

5

ES 2 762 672 T3

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

ES 2 762 672 T3

Ser Leu Pro Glu Ala Gly Glu Lys Leu Phe Pro Tyr Leu Asp Pro Phe
 385 390 395 400

Asp Ser Ala Ser Gln Leu Met Glu Pro Gly Arg Gln Ser Met Asp Leu
 405 410 415

Val Glu Leu Val Ser Leu Phe Pro Thr Leu Ala Gly Leu Ala Gly Leu
 420 425 430

Gln Val Pro Pro Arg Cys Pro Val Pro Ser Phe His Val Glu Leu Cys
 435 440 445

Arg Glu Gly Lys Asn Leu Leu Lys His Phe Arg Phe Arg Asp Leu Glu
 450 455 460

Glu Asp Pro Tyr Leu Pro Gly Asn Pro Arg Glu Leu Ile Ala Tyr Ser
 465 470 475 480

Gln Tyr Pro Arg Pro Ser Asp Ile Pro Gln Trp Asn Ser Asp Lys Pro
 485 490 495

Ser Leu Lys Asp Ile Lys Ile Met Gly Tyr Ser Ile Arg Thr Ile Asp
 500 505 510

Tyr Arg Tyr Thr Val Trp Val Gly Phe Asn Pro Asp Glu Phe Leu Ala
 515 520 525

Asn Phe Ser Asp Ile His Ala Gly Glu Leu Tyr Phe Val Asp Ser Asp
 530 535 540

Pro Leu Gln Asp His Asn Met Tyr Asn Asp Ser Gln Gly Gly Asp Leu
 545 550 555 560

Phe Gln Leu Leu Met Pro Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys
 565 570 575

Glu Ala Ala Ala Lys Gly Met Glu Val Arg Trp Cys Ala Thr Ser Asp
 580 585 590

Pro Glu Gln His Lys Cys Gly Asn Met Ser Glu Ala Phe Arg Glu Ala
 595 600 605

Gly Ile Gln Pro Ser Leu Leu Cys Val Arg Gly Thr Ser Ala Asp His
 610 615 620

Cys Val Gln Leu Ile Ala Ala Gln Glu Ala Asp Ala Ile Thr Leu Asp
 625 630 635 640

ES 2 762 672 T3

Gly Gly Ala Ile Tyr Glu Ala Gly Lys Glu His Gly Leu Lys Pro Val
 645 650 655
 Val Gly Glu Val Tyr Asp Gln Glu Val Gly Thr Ser Tyr Tyr Ala Val
 660 665 670
 Ala Val Val Arg Arg Ser Ser His Val Thr Ile Asp Thr Leu Lys Gly
 675 680 685
 Val Lys Ser Cys His Thr Gly Ile Asn Arg Thr Val Gly Trp Asn Val
 690 695 700
 Pro Val Gly Tyr Leu Val Glu Ser Gly Arg Leu Ser Val Met Gly Cys
 705 710 715 720
 Asp Val Leu Lys Ala Val Ser Asp Tyr Phe Gly Gly Ser Cys Val Pro
 725 730 735
 Gly Ala Gly Glu Thr Ser Tyr Ser Glu Ser Leu Cys Arg Leu Cys Arg
 740 745 750
 Gly Asp Ser Ser Gly Glu Gly Val Cys Asp Lys Ser Pro Leu Glu Arg
 755 760 765
 Tyr Tyr Asp Tyr Ser Gly Ala Phe Arg Cys Leu Ala Glu Gly Ala Gly
 770 775 780
 Asp Val Ala Phe Val Lys His Ser Thr Val Leu Glu Asn Thr Asp Gly
 785 790 795 800
 Lys Thr Leu Pro Ser Trp Gly Gln Ala Leu Leu Ser Gln Asp Phe Glu
 805 810 815
 Leu Leu Cys Arg Asp Gly Ser Arg Ala Asp Val Thr Glu Trp Arg Gln
 820 825 830
 Cys His Leu Ala Arg Val Pro Ala His Ala Val Val Arg Ala Asp
 835 840 845
 Thr Asp Gly Gly Leu Ile Phe Arg Leu Leu Asn Glu Gly Gln Arg Leu
 850 855 860
 Phe Ser His Glu Gly Ser Ser Phe Gln Met Phe Ser Ser Glu Ala Tyr
 865 870 875 880
 Gly Gln Lys Asp Leu Leu Phe Lys Asp Ser Thr Ser Glu Leu Val Pro

ES 2 762 672 T3

[illegible]

ES 2 762 672 T3

Ala Gly Asp Val Ala Phe Val Arg His Thr Thr Val Phe Asp Asn
1130 1135 1140

Thr Asn Gly His Asn Ser Glu Pro Trp Ala Ala Glu Leu Arg Ser
1145 1150 1155

Glu Asp Tyr Glu Leu Leu Cys Pro Asn Gly Ala Arg Ala Glu Val
1160 1165 1170

Ser Gln Phe Ala Ala Cys Asn Leu Ala Gln Ile Pro Pro His Ala
1175 1180 1185

Val Met Val Arg Pro Asp Thr Asn Ile Phe Thr Val Tyr Gly Leu
1190 1195 1200

Leu Asp Lys Ala Gln Asp Leu Phe Gly Asp Asp His Asn Lys Asn
1205 1210 1215

Gly Phe Lys Met Phe Asp Ser Ser Asn Tyr His Gly Gln Asp Leu
1220 1225 1230

Leu Phe Lys Asp Ala Thr Val Arg Ala Val Pro Val Gly Glu Lys
1235 1240 1245

Thr Thr Tyr Arg Gly Trp Leu Gly Leu Asp Tyr Val Ala Ala Leu
1250 1255 1260

Glu Gly Met Ser Ser Gln Gln Cys Ser
1265 1270

<210> 30

<211> 1266

5 <212> PRT

<213> Secuencia Artificial

<220>

<223> Proteína de fusión p97

10

<400> 30

Met Arg Gly Pro Ser Gly Ala Leu Trp Leu Leu Leu Ala Leu Arg Thr
1 5 10 15

Val Leu Gly His His His His His His His His His His Glu Asn Leu
20 25 30

Tyr Phe Gln Gly Met Glu Val Arg Trp Cys Ala Thr Ser Asp Pro Glu
35 40 45

Gln His Lys Cys Gly Asn Met Ser Glu Ala Phe Arg Glu Ala Gly Ile

ES 2 762 672 T3

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

ES 2 762 672 T3

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

ES 2 762 672 T3

Cys Val Gly Asn Ser Gln Glu Arg Tyr Tyr Gly Tyr Arg Gly Ala Phe
 565 570 575
 Arg Cys Leu Val Glu Asn Ala Gly Asp Val Ala Phe Val Arg His Thr
 580 585 590
 Thr Val Phe Asp Asn Thr Asn Gly His Asn Ser Glu Pro Trp Ala Ala
 595 600 605
 Glu Leu Arg Ser Glu Asp Tyr Glu Leu Leu Cys Pro Asn Gly Ala Arg
 610 615 620
 Ala Glu Val Ser Gln Phe Ala Ala Cys Asn Leu Ala Gln Ile Pro Pro
 625 630 635 640
 His Ala Val Met Val Arg Pro Asp Thr Asn Ile Phe Thr Val Tyr Gly
 645 650 655
 Leu Leu Asp Lys Ala Gln Asp Leu Phe Gly Asp Asp His Asn Lys Asn
 660 665 670
 Gly Phe Lys Met Phe Asp Ser Ser Asn Tyr His Gly Gln Asp Leu Leu
 675 680 685
 Phe Lys Asp Ala Thr Val Arg Ala Val Pro Val Gly Glu Lys Thr Thr
 690 695 700
 Tyr Arg Gly Trp Leu Gly Leu Asp Tyr Val Ala Ala Leu Glu Gly Met
 705 710 715 720
 Ser Ser Gln Gln Cys Ser Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys
 725 730 735
 Glu Ala Ala Ala Lys Ser Glu Thr Gln Ala Asn Ser Thr Thr Asp Ala
 740 745 750
 Leu Asn Val Leu Leu Ile Ile Val Asp Asp Leu Arg Pro Ser Leu Gly
 755 760 765
 Cys Tyr Gly Asp Lys Leu Val Arg Ser Pro Asn Ile Asp Gln Leu Ala
 770 775 780
 Ser His Ser Leu Leu Phe Gln Asn Ala Phe Ala Gln Gln Ala Val Cys
 785 790 795 800
 Ala Pro Ser Arg Val Ser Phe Leu Thr Gly Arg Arg Pro Asp Thr Thr
 805 810 815

ES 2 762 672 T3

Arg Leu Tyr Asp Phe Asn Ser Tyr Trp Arg Val His Ala Gly Asn Phe
 820 825 830
 Ser Thr Ile Pro Gln Tyr Phe Lys Glu Asn Gly Tyr Val Thr Met Ser
 835 840 845
 Val Gly Lys Val Phe His Pro Gly Ile Ser Ser Asn His Thr Asp Asp
 850 855 860
 Ser Pro Tyr Ser Trp Ser Phe Pro Pro Tyr His Pro Ser Ser Glu Lys
 865 870 875 880
 Tyr Glu Asn Thr Lys Thr Cys Arg Gly Pro Asp Gly Glu Leu His Ala
 885 890 895
 Asn Leu Leu Cys Pro Val Asp Val Leu Asp Val Pro Glu Gly Thr Leu
 900 905 910
 Pro Asp Lys Gln Ser Thr Glu Gln Ala Ile Gln Leu Leu Glu Lys Met
 915 920 925
 Lys Thr Ser Ala Ser Pro Phe Phe Leu Ala Val Gly Tyr His Lys Pro
 930 935 940
 His Ile Pro Phe Arg Tyr Pro Lys Glu Phe Gln Lys Leu Tyr Pro Leu
 945 950 955 960
 Glu Asn Ile Thr Leu Ala Pro Asp Pro Glu Val Pro Asp Gly Leu Pro
 965 970 975
 Pro Val Ala Tyr Asn Pro Trp Met Asp Ile Arg Gln Arg Glu Asp Val
 980 985 990
 Gln Ala Leu Asn Ile Ser Val Pro Tyr Gly Pro Ile Pro Val Asp Phe
 995 1000 1005
 Gln Arg Lys Ile Arg Gln Ser Tyr Phe Ala Ser Val Ser Tyr Leu
 1010 1015 1020
 Asp Thr Gln Val Gly Arg Leu Leu Ser Ala Leu Asp Asp Leu Gln
 1025 1030 1035
 Leu Ala Asn Ser Thr Ile Ile Ala Phe Thr Ser Asp His Gly Trp
 1040 1045 1050
 Ala Leu Gly Glu His Gly Glu Trp Ala Lys Tyr Ser Asn Phe Asp

ES 2 762 672 T3

1055		1060		1065
Val Ala Thr His Val Pro Leu Ile Phe Tyr Val Pro Gly Arg Thr				
1070		1075		1080
Ala Ser Leu Pro Glu Ala Gly Glu Lys Leu Phe Pro Tyr Leu Asp				
1085		1090		1095
Pro Phe Asp Ser Ala Ser Gln Leu Met Glu Pro Gly Arg Gln Ser				
1100		1105		1110
Met Asp Leu Val Glu Leu Val Ser Leu Phe Pro Thr Leu Ala Gly				
1115		1120		1125
Leu Ala Gly Leu Gln Val Pro Pro Arg Cys Pro Val Pro Ser Phe				
1130		1135		1140
His Val Glu Leu Cys Arg Glu Gly Lys Asn Leu Leu Lys His Phe				
1145		1150		1155
Arg Phe Arg Asp Leu Glu Glu Asp Pro Tyr Leu Pro Gly Asn Pro				
1160		1165		1170
Arg Glu Leu Ile Ala Tyr Ser Gln Tyr Pro Arg Pro Ser Asp Ile				
1175		1180		1185
Pro Gln Trp Asn Ser Asp Lys Pro Ser Leu Lys Asp Ile Lys Ile				
1190		1195		1200
Met Gly Tyr Ser Ile Arg Thr Ile Asp Tyr Arg Tyr Thr Val Trp				
1205		1210		1215
Val Gly Phe Asn Pro Asp Glu Phe Leu Ala Asn Phe Ser Asp Ile				
1220		1225		1230
His Ala Gly Glu Leu Tyr Phe Val Asp Ser Asp Pro Leu Gln Asp				
1235		1240		1245
His Asn Met Tyr Asn Asp Ser Gln Gly Gly Asp Leu Phe Gln Leu				
1250		1255		1260
Leu Met Pro				
1265				

<210> 31
 <211> 550
 <212> PRT
 <213> Homo sapiens
 <400> 31

ES 2 762 672 T3

Met 1	Pro	Pro	Pro	Arg 5	Thr	Gly	Arg	Gly	Leu 10	Leu	Trp	Leu	Gly	Leu 15	Val
Leu	Ser	Ser	Val 20	Cys	Val	Ala	Leu	Gly 25	Ser	Glu	Thr	Gln	Ala 30	Asn	Ser
Thr	Thr	Asp 35	Ala	Leu	Asn	Val	Leu 40	Leu	Ile	Ile	Val	Asp 45	Asp	Leu	Arg
Pro	Ser 50	Leu	Gly	Cys	Tyr	Gly 55	Asp	Lys	Leu	Val	Arg 60	Ser	Pro	Asn	Ile
Asp 65	Gln	Leu	Ala	Ser	His 70	Ser	Leu	Leu	Phe	Gln 75	Asn	Ala	Phe	Ala	Gln 80
Gln	Ala	Val	Cys	Ala 85	Pro	Ser	Arg	Val	Ser	Phe	Leu	Thr	Gly	Arg 95	Arg
Pro	Asp	Thr	Thr 100	Arg	Leu	Tyr	Asp	Phe 105	Asn	Ser	Tyr	Trp	Arg 110	Val	His
Ala	Gly	Asn 115	Phe	Ser	Thr	Ile	Pro 120	Gln	Tyr	Phe	Lys	Glu 125	Asn	Gly	Tyr
Val	Thr 130	Met	Ser	Val	Gly	Lys 135	Val	Phe	His	Pro	Gly 140	Ile	Ser	Ser	Asn
His 145	Thr	Asp	Asp	Ser	Pro 150	Tyr	Ser	Trp	Ser	Phe 155	Pro	Pro	Tyr	His	Pro 160
Ser	Ser	Glu	Lys	Tyr 165	Glu	Asn	Thr	Lys	Thr 170	Cys	Arg	Gly	Pro	Asp 175	Gly
Glu	Leu	His	Ala 180	Asn	Leu	Leu	Cys	Pro 185	Val	Asp	Val	Leu	Asp 190	Val	Pro
Glu	Gly	Thr 195	Leu	Pro	Asp	Lys	Gln 200	Ser	Thr	Glu	Gln	Ala 205	Ile	Gln	Leu
Leu	Glu 210	Lys	Met	Lys	Thr	Ser 215	Ala	Ser	Pro	Phe	Phe 220	Leu	Ala	Val	Gly
Tyr 225	His	Lys	Pro	His	Ile 230	Pro	Phe	Arg	Tyr	Pro 235	Lys	Glu	Phe	Gln	Lys 240

ES 2 762 672 T3

Leu Tyr Pro Leu Glu Asn Ile Thr Leu Ala Pro Asp Pro Glu Val Pro
 245 250 255

Asp Gly Leu Pro Pro Val Ala Tyr Asn Pro Trp Met Asp Ile Arg Gln
 260 265 270

Arg Glu Asp Val Gln Ala Leu Asn Ile Ser Val Pro Tyr Gly Pro Ile
 275 280 285

Pro Val Asp Phe Gln Arg Lys Ile Arg Gln Ser Tyr Phe Ala Ser Val
 290 295 300

Ser Tyr Leu Asp Thr Gln Val Gly Arg Leu Leu Ser Ala Leu Asp Asp
 305 310 315 320

Leu Gln Leu Ala Asn Ser Thr Ile Ile Ala Phe Thr Ser Asp His Gly
 325 330 335

Trp Ala Leu Gly Glu His Gly Glu Trp Ala Lys Tyr Ser Asn Phe Asp
 340 345 350

Val Ala Thr His Val Pro Leu Ile Phe Tyr Val Pro Gly Arg Thr Ala
 355 360 365

Ser Leu Pro Glu Ala Gly Glu Lys Leu Phe Pro Tyr Leu Asp Pro Phe
 370 375 380

Asp Ser Ala Ser Gln Leu Met Glu Pro Gly Arg Gln Ser Met Asp Leu
 385 390 395 400

Val Glu Leu Val Ser Leu Phe Pro Thr Leu Ala Gly Leu Ala Gly Leu
 405 410 415

Gln Val Pro Pro Arg Cys Pro Val Pro Ser Phe His Val Glu Leu Cys
 420 425 430

Arg Glu Gly Lys Asn Leu Leu Lys His Phe Arg Phe Arg Asp Leu Glu
 435 440 445

Glu Asp Pro Tyr Leu Pro Gly Asn Pro Arg Glu Leu Ile Ala Tyr Ser
 450 455 460

Gln Tyr Pro Arg Pro Ser Asp Ile Pro Gln Trp Asn Ser Asp Lys Pro
 465 470 475 480

Ser Leu Lys Asp Ile Lys Ile Met Gly Tyr Ser Ile Arg Thr Ile Asp
 485 490 495

ES 2 762 672 T3

Tyr Arg Tyr Thr Val Trp Val Gly Phe Asn Pro Asp Glu Phe Leu Ala
500 505 510

Asn Phe Ser Asp Ile His Ala Gly Glu Leu Tyr Phe Val Asp Ser Asp
515 520 525

Pro Leu Gln Asp His Asn Met Tyr Asn Asp Ser Gln Gly Gly Asp Leu
530 535 540

Phe Gln Leu Leu Met Pro
545 550

<210> 32

<211> 525

<212> PRT

<213> Homo sapiens

<400> 32

Ser Glu Thr Gln Ala Asn Ser Thr Thr Asp Ala Leu Asn Val Leu Leu
1 5 10 15

Ile Ile Val Asp Asp Leu Arg Pro Ser Leu Gly Cys Tyr Gly Asp Lys
20 25 30

Leu Val Arg Ser Pro Asn Ile Asp Gln Leu Ala Ser His Ser Leu Leu
35 40 45

Phe Gln Asn Ala Phe Ala Gln Gln Ala Val Cys Ala Pro Ser Arg Val
50 55 60

Ser Phe Leu Thr Gly Arg Arg Pro Asp Thr Thr Arg Leu Tyr Asp Phe
65 70 75 80

Asn Ser Tyr Trp Arg Val His Ala Gly Asn Phe Ser Thr Ile Pro Gln
85 90 95

Tyr Phe Lys Glu Asn Gly Tyr Val Thr Met Ser Val Gly Lys Val Phe
100 105 110

His Pro Gly Ile Ser Ser Asn His Thr Asp Asp Ser Pro Tyr Ser Trp
115 120 125

Ser Phe Pro Pro Tyr His Pro Ser Ser Glu Lys Tyr Glu Asn Thr Lys
130 135 140

Thr Cys Arg Gly Pro Asp Gly Glu Leu His Ala Asn Leu Leu Cys Pro
145 150 155 160

5

10

ES 2 762 672 T3

Val	Asp	Val	Leu	Asp	Val	Pro	Glu	Gly	Thr	Leu	Pro	Asp	Lys	Gln	Ser	165	170	175
Thr	Glu	Gln	Ala	Ile	Gln	Leu	Leu	Glu	Lys	Met	Lys	Thr	Ser	Ala	Ser	180	185	190
Pro	Phe	Phe	Leu	Ala	Val	Gly	Tyr	His	Lys	Pro	His	Ile	Pro	Phe	Arg	195	200	205
Tyr	Pro	Lys	Glu	Phe	Gln	Lys	Leu	Tyr	Pro	Leu	Glu	Asn	Ile	Thr	Leu	210	215	220
Ala	Pro	Asp	Pro	Glu	Val	Pro	Asp	Gly	Leu	Pro	Pro	Val	Ala	Tyr	Asn	225	230	235
Pro	Trp	Met	Asp	Ile	Arg	Gln	Arg	Glu	Asp	Val	Gln	Ala	Leu	Asn	Ile	245	250	255
Ser	Val	Pro	Tyr	Gly	Pro	Ile	Pro	Val	Asp	Phe	Gln	Arg	Lys	Ile	Arg	260	265	270
Gln	Ser	Tyr	Phe	Ala	Ser	Val	Ser	Tyr	Leu	Asp	Thr	Gln	Val	Gly	Arg	275	280	285
Leu	Leu	Ser	Ala	Leu	Asp	Asp	Leu	Gln	Leu	Ala	Asn	Ser	Thr	Ile	Ile	290	295	300
Ala	Phe	Thr	Ser	Asp	His	Gly	Trp	Ala	Leu	Gly	Glu	His	Gly	Glu	Trp	305	310	315
Ala	Lys	Tyr	Ser	Asn	Phe	Asp	Val	Ala	Thr	His	Val	Pro	Leu	Ile	Phe	325	330	335
Tyr	Val	Pro	Gly	Arg	Thr	Ala	Ser	Leu	Pro	Glu	Ala	Gly	Glu	Lys	Leu	340	345	350
Phe	Pro	Tyr	Leu	Asp	Pro	Phe	Asp	Ser	Ala	Ser	Gln	Leu	Met	Glu	Pro	355	360	365
Gly	Arg	Gln	Ser	Met	Asp	Leu	Val	Glu	Leu	Val	Ser	Leu	Phe	Pro	Thr	370	375	380
Leu	Ala	Gly	Leu	Ala	Gly	Leu	Gln	Val	Pro	Pro	Arg	Cys	Pro	Val	Pro	385	390	395
Ser	Phe	His	Val	Glu	Leu	Cys	Arg	Glu	Gly	Lys	Asn	Leu	Leu	Lys	His	405	410	415

ES 2 762 672 T3

Phe Arg Phe Arg Asp Leu Glu Glu Asp Pro Tyr Leu Pro Gly Asn Pro
420 425 430

Arg Glu Leu Ile Ala Tyr Ser Gln Tyr Pro Arg Pro Ser Asp Ile Pro
435 440 445

Gln Trp Asn Ser Asp Lys Pro Ser Leu Lys Asp Ile Lys Ile Met Gly
450 455 460

Tyr Ser Ile Arg Thr Ile Asp Tyr Arg Tyr Thr Val Trp Val Gly Phe
465 470 475 480

Asn Pro Asp Glu Phe Leu Ala Asn Phe Ser Asp Ile His Ala Gly Glu
485 490 495

Leu Tyr Phe Val Asp Ser Asp Pro Leu Gln Asp His Asn Met Tyr Asn
500 505 510

Asp Ser Gln Gly Gly Asp Leu Phe Gln Leu Leu Met Pro
515 520 525

<210> 33

<211> 517

<212> PRT

<213> Homo sapiens

<400> 33

Thr Asp Ala Leu Asn Val Leu Leu Ile Ile Val Asp Asp Leu Arg Pro
1 5 10 15

Ser Leu Gly Cys Tyr Gly Asp Lys Leu Val Arg Ser Pro Asn Ile Asp
20 25 30

Gln Leu Ala Ser His Ser Leu Leu Phe Gln Asn Ala Phe Ala Gln Gln
35 40 45

Ala Val Cys Ala Pro Ser Arg Val Ser Phe Leu Thr Gly Arg Arg Pro
50 55 60

Asp Thr Thr Arg Leu Tyr Asp Phe Asn Ser Tyr Trp Arg Val His Ala
65 70 75 80

Gly Asn Phe Ser Thr Ile Pro Gln Tyr Phe Lys Glu Asn Gly Tyr Val
85 90 95

Thr Met Ser Val Gly Lys Val Phe His Pro Gly Ile Ser Ser Asn His
100 105 110

ES 2 762 672 T3

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

ES 2 762 672 T3

355 360 365
 Glu Leu Val Ser Leu Phe Pro Thr Leu Ala Gly Leu Ala Gly Leu Gln
 370 375 380
 Val Pro Pro Arg Cys Pro Val Pro Ser Phe His Val Glu Leu Cys Arg
 385 390 395 400
 Glu Gly Lys Asn Leu Leu Lys His Phe Arg Phe Arg Asp Leu Glu Glu
 405 410 415
 Asp Pro Tyr Leu Pro Gly Asn Pro Arg Glu Leu Ile Ala Tyr Ser Gln
 420 425 430
 Tyr Pro Arg Pro Ser Asp Ile Pro Gln Trp Asn Ser Asp Lys Pro Ser
 435 440 445
 Leu Lys Asp Ile Lys Ile Met Gly Tyr Ser Ile Arg Thr Ile Asp Tyr
 450 455 460
 Arg Tyr Thr Val Trp Val Gly Phe Asn Pro Asp Glu Phe Leu Ala Asn
 465 470 475 480
 Phe Ser Asp Ile His Ala Gly Glu Leu Tyr Phe Val Asp Ser Asp Pro
 485 490 495
 Leu Gln Asp His Asn Met Tyr Asn Asp Ser Gln Gly Gly Asp Leu Phe
 500 505 510
 Gln Leu Leu Met Pro
 515
 <210> 34
 <211> 422
 5 <212> PRT
 <213> Homo sapiens
 <400> 34
 Thr Asp Ala Leu Asn Val Leu Leu Ile Ile Val Asp Asp Leu Arg Pro
 1 5 10 15
 Ser Leu Gly Cys Tyr Gly Asp Lys Leu Val Arg Ser Pro Asn Ile Asp
 20 25 30
 Gln Leu Ala Ser His Ser Leu Leu Phe Gln Asn Ala Phe Ala Gln Gln
 35 40 45
 Ala Val Cys Ala Pro Ser Arg Val Ser Phe Leu Thr Gly Arg Arg Pro
 10

ES 2 762 672 T3

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

ES 2 762 672 T3

Ala Leu Gly Glu His Gly Glu Trp Ala Lys Tyr Ser Asn Phe Asp Val
305 310 315 320

Ala Thr His Val Pro Leu Ile Phe Tyr Val Pro Gly Arg Thr Ala Ser
325 330 335

Leu Pro Glu Ala Gly Glu Lys Leu Phe Pro Tyr Leu Asp Pro Phe Asp
340 345 350

Ser Ala Ser Gln Leu Met Glu Pro Gly Arg Gln Ser Met Asp Leu Val
355 360 365

Glu Leu Val Ser Leu Phe Pro Thr Leu Ala Gly Leu Ala Gly Leu Gln
370 375 380

Val Pro Pro Arg Cys Pro Val Pro Ser Phe His Val Glu Leu Cys Arg
385 390 395 400

Glu Gly Lys Asn Leu Leu Lys His Phe Arg Phe Arg Asp Leu Glu Glu
405 410 415

Asp Pro Tyr Leu Pro Gly
420

<210> 35

<211> 95

<212> PRT

<213> Homo sapiens

<400> 35

Asn Pro Arg Glu Leu Ile Ala Tyr Ser Gln Tyr Pro Arg Pro Ser Asp
1 5 10 15

Ile Pro Gln Trp Asn Ser Asp Lys Pro Ser Leu Lys Asp Ile Lys Ile
20 25 30

Met Gly Tyr Ser Ile Arg Thr Ile Asp Tyr Arg Tyr Thr Val Trp Val
35 40 45

Gly Phe Asn Pro Asp Glu Phe Leu Ala Asn Phe Ser Asp Ile His Ala
50 55 60

Gly Glu Leu Tyr Phe Val Asp Ser Asp Pro Leu Gln Asp His Asn Met
65 70 75 80

Tyr Asn Asp Ser Gln Gly Gly Asp Leu Phe Gln Leu Leu Met Pro
85 90 95

<210> 36

<211> 5

<212> PRT

<213> Secuencia Artificial

<220>

<223> Enlazador peptídico rígido

<400> 36

Glu Ala Ala Ala Lys
1 5

<210> 37
 <211> 10
 <212> PRT
 5 <213> Secuencia Artificial

 <220>
 <223> Enlazador peptídico rígido

 10 <400> 37
 Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys
 1 5 10

 <210> 38
 <211> 15
 15 <212> PRT
 <213> Secuencia Artificial

 <220>
 <223> Enlazador peptídico rígido

 20 <400> 38
 Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys
 1 5 10 15

 <210> 39
 <211> 19
 25 <212> PRT
 <213> Homo sapiens

 <400> 39
 Met Arg Gly Pro Ser Gly Ala Leu Trp Leu Leu Leu Ala Leu Arg Thr
 1 5 10 15

 30 Val Leu Gly

 <210> 40
 <211> 25
 <212> PRT
 35 <213> Homo sapiens

 <400> 40
 Met Pro Pro Pro Arg Thr Gly Arg Gly Leu Leu Trp Leu Gly Leu Val
 1 5 10 15

 Leu Ser Ser Val Cys Val Ala Leu Gly
 20 25

 40 <210> 41
 <211> 16
 <212> PRT
 <213> Secuencia Artificial

 45 <220>
 <223> Enlazador peptídico rígido

 <400> 41
 Ala Glu Ala Ala Ala Lys Ala Leu Glu Ala Glu Ala Ala Ala Lys Ala
 1 5 10 15

 50 <210> 42
 <211> 46
 <212> PRT
 <213> Secuencia Artificial

 55

<220>
 <223> Enlazador peptídico rígido

<400> 42
 Ala Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys
 1 5 10 15

Glu Ala Ala Ala Lys Ala Leu Glu Ala Glu Ala Ala Ala Lys Glu Ala
 20 25 30

5 Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Ala
 35 40 45

<210> 43
 <211> 5
 <212> PRT
 10 <213> Secuencia Artificial

<220>
 <223> Enlazador peptídico rígido

15 <400> 43
 Pro Ala Pro Ala Pro
 1 5

<210> 44
 <211> 12
 20 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Enlazador peptídico rígido

25 <400> 44
 Ala Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Ala
 1 5 10

<210> 45
 30 <211> 5
 <212> PRT
 <213> Secuencia Artificial

<220>
 35 <223> Enlazador peptídico flexible

<400> 45
 Gly Gly Gly Gly Ser
 1 5

40 <210> 46
 <211> 10
 <212> PRT
 <213> Secuencia Artificial

45 <220>
 <223> Enlazador peptídico flexible

<400> 46
 Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
 1 5 10

50 <210> 47
 <211> 15
 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Enlazador peptídico flexible

5 <400> 47
 Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
 1 5 10 15

<210> 48
 <211> 4
 10 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Enlazador peptídico flexible

15 <400> 48
 Gly Gly Gly Gly
 1

20 <210> 49
 <211> 5
 <212> PRT
 <213> Secuencia Artificial

25 <220>
 <223> Enlazador peptídico flexible

<400> 49
 Gly Gly Gly Gly Gly
 30 1 5

<210> 50
 <211> 6
 <212> PRT
 35 <213> Secuencia Artificial

<220>
 <223> Enlazador peptídico flexible

40 <400> 50
 Gly Gly Gly Gly Gly Gly
 1 5

<210> 51
 <211> 7
 45 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Enlazador peptídico flexible

50 <400> 51
 Gly Gly Gly Gly Gly Gly Gly
 1 5

55 <210> 52
 <211> 8
 <212> PRT
 <213> Secuencia Artificial

60 <220>
 <223> Enlazador peptídico flexible

<400> 52
 Gly Gly Gly Gly Gly Gly Gly Gly
 1 5

5 <210> 53
 <211> 9
 <212> PRT
 <213> Secuencia Artificial

10 <220>
 <223> Enlazador peptídico flexible

<400> 53
 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
 1 5

15 <210> 54
 <211> 10
 <212> PRT
 <213> Secuencia Artificial

20 <220>
 <223> Enlazador peptídico flexible

<400> 54
 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
 1 5 10

25 <210> 55
 <211> 4
 <212> PRT
 <213> Secuencia Artificial

30 <220>
 <223> Enlazador peptídico

<400> 55
 Gly Ser Gly Ser
 1

35 <210> 56
 <211> 4
 <212> PRT
 <213> Secuencia Artificial

40 <220>
 <223> Enlazador peptídico

45 <400> 56
 Gly Gly Ser Gly
 1

50 <210> 57
 <211> 4
 <212> PRT
 <213> Secuencia Artificial

55 <220>
 <223> Enlazador peptídico

<400> 57
 Gly Gly Gly Ser
 1

60

<210> 58
 <211> 4
 <212> PRT
 <213> Secuencia Artificial
 5
 <220>
 <223> Enlazador peptídico
 <400> 58
 Gly Asn Gly Asn
 10 1
 <210> 59
 <211> 4
 <212> PRT
 15 <213> Secuencia Artificial
 <220>
 <223> Enlazador peptídico
 20 <400> 59
 Gly Gly Asn Gly
 1
 <210> 60
 <211> 4
 25 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> Enlazador peptídico
 30 <400> 60
 Gly Gly Gly Asn
 1
 <210> 61
 <211> 5
 35 <212> PRT
 <213> Secuencia Artificial
 <220>
 40 <223> Enlazador peptídico
 <400> 61
 Gly Gly Gly Gly Asn
 1 5
 45 <210> 62
 <211> 17
 <212> PRT
 <213> Secuencia Artificial
 50 <220>
 <223> Enlazador escindible por el Factor Xla/FVIIa
 <400> 62
 Val Ser Gln Thr Ser Lys Leu Thr Arg Ala Glu Thr Val Phe Pro Asp
 1 5 10 15
 55 Val
 <210> 63
 <211> 6
 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Enlazador escindible por la metaloproteasa-1 de matriz

5 <400> 63
 Pro Leu Gly Leu Trp Ala
 1 5

 <210> 64
 <211> 6
 10 <212> PRT
 <213> Secuencia Artificial

 <220>
 <223> Enlazador escindible por la proteasa del HIV

15 <400> 64
 Arg Val Leu Ala Glu Ala
 1 5

 <210> 65
 20 <211> 10
 <212> PRT
 <213> Secuencia Artificial

 <220>
 25 <223> Enlazador escindible por la proteasa NS3 del virus de la Hepatitis C

 <400> 65
 Glu Asp Val Val Cys Cys Ser Met Ser Tyr
 1 5 10

30 <210> 66
 <211> 8
 <212> PRT
 <213> Secuencia Artificial

35 <220>
 <223> Enlazador escindible por el Factor Xa

 <400> 66
 Gly Gly Ile Glu Gly Arg Gly Ser
 1 5

40 <210> 67
 <211> 10
 <212> PRT
 <213> Secuencia Artificial

45 <220>
 <223> Enlazador escindible por Furina

 <400> 67
 Thr Arg His Arg Gln Pro Arg Gly Trp Glu
 50 1 5 10

 <210> 68
 <211> 10
 <212> PRT
 55 <213> Secuencia Artificial

 <220>
 <223> Enlazador escindible por Furina

60 <400> 68

Ala Gly Asn Arg Val Arg Arg Ser Val Gly
1 5 10

<210> 69
<211> 9
5 <212> PRT
<213> Secuencia Artificial

<220>
<223> Enlazador escindible por Furina
10

<400> 69
Arg Arg Arg Arg Arg Arg Arg Arg Arg
1 5

<210> 70
15 <211> 4
<212> PRT
<213> Secuencia Artificial

<220>
20 <223> Enlazador escindible por Catepsina B

<400> 70
Gly Phe Leu Gly
1

<210> 71
25 <211> 4
<212> PRT
<213> Secuencia Artificial

<220>
30 <223> Enlazador escindible por Trombina

<400> 71
Gly Arg Gly Asp
1

<210> 72
35 <211> 6
<212> PRT
<213> Secuencia Artificial

<220>
40 <223> Enlazador escindible por Trombina

<400> 72
Gly Arg Gly Asp Asn Pro
45 1 5

<210> 73
<211> 5
<212> PRT
50 <213> Secuencia Artificial

<220>
<223> Enlazador escindible por Trombina

<400> 73
Gly Arg Gly Asp Ser
55 1 5

<210> 74
60 <211> 7
<212> PRT

<213> Secuencia Artificial

<220>

<223> Enlazador escindible por Trombina

5

<400> 74
 Gly Arg Gly Asp Ser Pro Lys
 1 5

<210> 75

10 <211> 4
 <212> PRT
 <213> Secuencia Artificial

<220>

15 <223> Enlazador escindible por Elastasa

<400> 75
 Ala Ala Pro Val
 1

20 <210> 76
 <211> 4
 <212> PRT
 <213> Secuencia Artificial

25 <220>
 <223> Enlazador escindible por Elastasa

<400> 76
 Ala Ala Pro Leu
 1

30 <210> 77
 <211> 4
 <212> PRT
 <213> Secuencia Artificial

35 <220>
 <223> Enlazador escindible por Elastasa

<400> 77
 Ala Ala Pro Phe
 1

40 <210> 78
 <211> 4
 <212> PRT

45 <213> Secuencia Artificial

<220>
 <223> Enlazador escindible por Elastasa

50 <400> 78
 Ala Ala Pro Ala
 1

<210> 79
 <211> 4

55 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Enlazador escindible por Elastasa

60 <400> 79

Ala Tyr Leu Val
1
 <210> 80
 <211> 6
 5 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> Enlazador escindible por una metaloproteinasa de matriz
 10 <220>
 <221> VARIANT
 <222> (3)..(3)
 <223> Xaa = Cualquier aminoácido
 15 <220>
 <221> VARIANT
 <222> (6)..(6)
 <223> Xaa = Cualquier aminoácido
 20 <400> 80
Gly Pro Xaa Gly Pro Xaa
1 5
 <210> 81
 25 <211> 4
 <212> PRT
 <213> Secuencia Artificial
 <220>
 30 <223> Enlazador escindible por una metaloproteinasa de matriz
 <220>
 <221> VARIANT
 <222> (4)..(4)
 35 <223> Xaa = Cualquier aminoácido
 <400> 81
Leu Gly Pro Xaa
1
 40 <210> 82
 <211> 6
 <212> PRT
 <213> Secuencia Artificial
 45 <220>
 <223> Enlazador escindible por una metaloproteinasa de matriz
 <220>
 <221> VARIANT
 50 <222> (6)..(6)
 <223> Xaa = Cualquier aminoácido
 <400> 82
Gly Pro Ile Gly Pro Xaa
1 5
 55 <210> 83
 <211> 5
 <212> PRT
 <213> Secuencia Artificial
 60 <220>
 <223> Enlazador escindible por una metaloproteinasa de matriz

<220>
 <221> VARIANT
 <222> (5)..(5)
 5 <223> Xaa = Cualquier aminoácido

 <400> 83
 Ala Pro Gly Leu Xaa
 1 5

 10 <210> 84
 <211> 7
 <212> PRT
 <213> Secuencia Artificial

 15 <220>
 <223> Enlazador escindible por collagenasa

 <220>
 <221> VARIANT
 20 <222> (7)..(7)
 <223> Xaa = Cualquier aminoácido

 <400> 84
 Pro Leu Gly Pro Asp Arg Xaa
 1 5
 25 <210> 85
 <211> 7
 <212> PRT
 <213> Secuencia Artificial

 30 <220>
 <223> Enlazador escindible por collagenasa

 <220>
 35 <221> VARIANT
 <222> (7)..(7)
 <223> Xaa = Cualquier aminoácido

 <400> 85
 Pro Leu Gly Leu Leu Gly Xaa
 40 1 5

 <210> 86
 <211> 7
 <212> PRT
 45 <213> Secuencia Artificial

 <220>
 <223> Enlazador escindible por collagenasa

 50 <400> 86
 Pro Gln Gly Ile Ala Gly Trp
 1 5

 <210> 87
 <211> 5
 55 <212> PRT
 <213> Secuencia Artificial

 <220>
 <223> Enlazador escindible por collagenasa
 60 <400> 87

Pro Leu Gly Cys His
1 5

<210> 88
<211> 6
5 <212> PRT
<213> Secuencia Artificial

<220>
<223> Enlazador escindible por collagenasa

10 <400> 88
Pro Leu Gly Leu Tyr Ala
1 5

<210> 89
15 <211> 7
<212> PRT
<213> Secuencia Artificial

<220>
20 <223> Enlazador escindible por collagenasa

<400> 89
Pro Leu Ala Leu Trp Ala Arg
1 5

25 <210> 90
<211> 7
<212> PRT
<213> Secuencia Artificial

30 <220>
<223> Enlazador escindible por collagenasa

<400> 90
Pro Leu Ala Tyr Trp Ala Arg
1 5

35 <210> 91
<211> 7
<212> PRT
<213> Secuencia Artificial

40 <220>
<223> Enlazador escindible por estromelisina

<400> 91
Pro Tyr Ala Tyr Tyr Met Arg
45 1 5

<210> 92
<211> 7
<212> PRT
50 <213> Secuencia Artificial

<220>
<223> Enlazador escindible por gelatinasa

55 <400> 92
Pro Leu Gly Met Tyr Ser Arg
1 5

<210> 93
<211> 4
60 <212> PRT

<213> Secuencia Artificial

<220>

<223> Enlazador escindible por una enzima convertidora de angiotensina

5 <400> 93
Gly Asp Lys Pro
1

10 <210> 94
<211> 5
<212> PRT
<213> Secuencia Artificial

15 <220>
<223> Enlazador escindible por una enzima convertidora de angiotensina

<400> 94
Gly Ser Asp Lys Pro
1 5

20 <210> 95
<211> 4
<212> PRT
<213> Secuencia Artificial

25 <220>
<223> Enlazador escindible por catepsina B

<400> 95
Ala Leu Ala Leu
1

30 <210> 96
<211> 4
<212> PRT
<213> Secuencia Artificial

35 <220>
<223> Enlazador escindible por catepsina B

<400> 96
Gly Phe Leu Gly
1

40 <210> 97
<211> 22
<212> PRT

45 <213> Homo sapiens

<400> 97
Met Asp Met Arg Ala Pro Ala Gly Ile Phe Gly Phe Leu Leu Val Leu
1 5 10 15

Phe Pro Gly Tyr Arg Ser
20

50 <210> 98
<211> 18
<212> PRT
<213> Homo sapiens

55 <400> 98

ES 2 762 672 T3

Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala
1 5 10 15

Tyr Ser

<210> 99

<211> 19

5 <212> PRT

<213> Homo sapiens

<400> 99

Met Asp Trp Thr Trp Arg Val Phe Cys Leu Leu Ala Val Thr Pro Gly
1 5 10 15

Ala His Pro

10

<210> 100

<211> 19

<212> PRT

<213> Homo sapiens

15

<400> 100

Met Ala Trp Ser Pro Leu Phe Leu Thr Leu Ile Thr His Cys Ala Gly
1 5 10 15

Ser Trp Ala

<210> 101

20

<211> 19

<212> PRT

<213> Homo sapiens

<400> 101

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

Ser Arg Ala

25

<210> 102

<211> 20

<212> PRT

30

<213> Homo sapiens

<400> 102

Met Ala Arg Pro Leu Cys Thr Leu Leu Leu Leu Met Ala Thr Leu Ala
1 5 10 15

Gly Ala Leu Ala
20

35

<210> 103

<211> 15

<212> PRT

<213> Homo sapiens

40

<400> 103

Met Arg Ser Leu Val Phe Val Leu Leu Ile Gly Ala Ala Phe Ala
1 5 10 15

<210> 104

<211> 22

45

<212> PRT

<213> Homo sapiens

ES 2 762 672 T3

<400> 104
 Met Ser Arg Leu Phe Val Phe Ile Leu Ile Ala Leu Phe Leu Ser Ala
 1 5 10 15

 Ile Ile Asp Val Met Ser
 20

 5 <210> 105
 <211> 21
 <212> PRT
 <213> Homo sapiens

 10 <400> 105
 Met Gly Met Arg Met Met Phe Ile Met Phe Met Leu Val Val Leu Ala
 1 5 10 15

 Thr Thr Val Val Ser
 20

 <210> 106
 <211> 18
 15 <212> PRT
 <213> Homo sapiens

 <400> 106
 Met Arg Ala Phe Leu Phe Leu Thr Ala Cys Ile Ser Leu Pro Gly Val
 1 5 10 15

 Phe Gly
 20
 <210> 107
 <211> 18
 <212> PRT
 <213> Homo sapiens

 25 <400> 107
 Met Lys Phe Gln Ser Thr Leu Leu Leu Ala Ala Ala Ala Gly Ser Ala
 1 5 10 15

 Leu Ala
 30 <210> 108
 <211> 24
 <212> PRT
 <213> Homo sapiens

 <400> 108
 Met Ala Ser Ser Leu Tyr Ser Phe Leu Leu Ala Leu Ser Ile Val Tyr
 1 5 10 15

 Ile Phe Val Ala Pro Thr His Ser
 35 20

 <210> 109
 <211> 26
 <212> PRT
 40 <213> Homo sapiens

 <400> 109

ES 2 762 672 T3

Met Lys Thr His Tyr Ser Ser Ala Ile Leu Pro Ile Leu Thr Leu Phe
1 5 10 15

Val Phe Leu Ser Ile Asn Pro Ser His Gly
20 25

<210> 110

<211> 36

5 <212> PRT

<213> Homo sapiens

<400> 110

Met Glu Ser Val Ser Ser Leu Phe Asn Ile Phe Ser Thr Ile Met Val
1 5 10 15

Asn Tyr Lys Ser Leu Val Leu Ala Leu Leu Ser Val Ser Asn Leu Lys
20 25 30

Tyr Ala Arg Gly
35

10

<210> 111

<211> 21

<212> PRT

<213> Homo sapiens

15

<400> 111

Met Lys Ala Ala Gln Ile Leu Thr Ala Ser Ile Val Ser Leu Leu Pro
1 5 10 15

Ile Tyr Thr Ser Ala
20

20

<210> 112

<211> 19

<212> PRT

<213> Homo sapiens

<400> 112

Met Ile Lys Leu Lys Phe Gly Val Phe Phe Thr Val Leu Leu Ser Ser
1 5 10 15

25

Ala Tyr Ala

<210> 113

<211> 5

<212> PRT

30

<213> Secuencia Artificial

<220>

<223> Etiqueta de purificación

35

<400> 113

His His His His His

1 5

<210> 114

<211> 6

40

<212> PRT

<213> Secuencia Artificial

<220>

<223> Etiqueta de purificación

<400> 114
 His His His His His His
 1 5

5 <210> 115
 <211> 7
 <212> PRT
 <213> Secuencia Artificial

10 <220>
 <223> Etiqueta de purificación

<400> 115
 His His His His His His His
 1 5

15 <210> 116
 <211> 8
 <212> PRT
 <213> Secuencia Artificial

20 <220>
 <223> Etiqueta de purificación

<400> 116
 His His His His His His His His
 25 1 5

<210> 117
 <211> 9
 <212> PRT
 30 <213> Secuencia Artificial

<220>
 <223> Etiqueta de purificación

35 <400> 117
 His His His His His His His His His
 1 5

<210> 118
 <211> 10
 40 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Etiqueta de purificación

45 <400> 118
 His His His His His His His His His His
 1 5 10

<210> 119
 50 <211> 15
 <212> PRT
 <213> Secuencia Artificial

<220>
 55 <223> Etiqueta de purificación - AviTag

<400> 119
 Gly Leu Asn Asp Ile Phe Glu Ala Gln Lys Ile Glu Trp His Glu
 1 5 10 15

60 <210> 120

<211> 26
 <212> PRT
 <213> Secuencia Artificial

5 <220>
 <223> Etiqueta de purificación - Etiqueta de calmodulina

<400> 120
 Lys Arg Arg Trp Lys Lys Asn Phe Ile Ala Val Ser Ala Ala Asn Arg
 1 5 10 15
 10 Phe Lys Lys Ile Ser Ser Ser Gly Ala Leu
 20 25

<210> 121
 <211> 6
 15 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Etiqueta de purificación - Etiqueta de poliglutamato

20 <400> 121
 Glu Glu Glu Glu Glu Glu
 1 5

25 <210> 122
 <211> 8
 <212> PRT
 <213> Secuencia Artificial

30 <220>
 <223> Etiqueta de purificación - Etiqueta FLAG

<400> 122
 Asp Tyr Lys Asp Asp Asp Asp Lys
 1 5

35 <210> 123
 <211> 9
 <212> PRT
 <213> Secuencia Artificial

40 <220>
 <223> Etiqueta de purificación - Etiqueta HA

<400> 123
 Tyr Pro Tyr Asp Val Pro Asp Tyr Ala
 1 5

45 <210> 124
 <211> 10
 <212> PRT
 <213> Secuencia Artificial

50 <220>
 <223> Etiqueta de purificación - Etiqueta MYC

<400> 124
 Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu
 55 1 5 10

<210> 125
 <211> 15
 <212> PRT
 60 <213> Secuencia Artificial

<220>
 <223> Etiqueta de purificación - Etiqueta S

<400> 125
 Lys Glu Thr Ala Ala Lys Phe Glu Arg Gln His Met Asp Ser
 5 1 5 10 15

<210> 126
 <211> 38
 <212> PRT
 10 <213> Secuencia Artificial

<220>
 <223> Etiqueta de purificación - Etiqueta de SPB

15 <400> 126
 Met Asp Glu Lys Thr Thr Gly Trp Arg Gly Gly His Val Val Glu Gly
 1 5 10 15

Leu Ala Gly Glu Leu Glu Gln Leu Arg Ala Arg Leu Glu His His Pro
 20 25 30

Gln Gly Gln Arg Glu Pro
 35

<210> 127
 <211> 13
 <212> PRT
 20 <213> Secuencia Artificial

<220>
 <223> Etiqueta de purificación - Softag 1

25 <400> 127
 Ser Leu Ala Glu Leu Leu Asn Ala Gly Leu Gly Gly Ser
 1 5 10

<210> 128
 <211> 8
 <212> PRT
 30 <213> Secuencia Artificial

<220>
 <223> Etiqueta de purificación - Softag 3

35 <400> 128
 Thr Gln Asp Pro Ser Arg Val Gly
 1 5

40 <210> 129
 <211> 14
 <212> PRT
 <213> Secuencia Artificial

45 <220>
 <223> Etiqueta de purificación - Etiqueta V5

<400> 129
 Gly Lys Pro Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr
 1 5 10

50 <210> 130
 <211> 8
 <212> PRT
 <213> Secuencia Artificial

```

<220>
<223> Etiqueta de purificación - Etiqueta Xpress

5  <400> 130
    Asp Leu Tyr Asp Asp Asp Asp Lys
    1                               5

    <210> 131
    <211> 6
10  <212> PRT
    <213> Homo sapiens

    <400> 131
    Leu Val Pro Arg Gly Ser
    1                               5
15

    <210> 132
    <211> 5
    <212> PRT
    <213> Homo sapiens
20

    <400> 132
    Asp Asp Asp Asp Lys
    1                               5

    <210> 133
    <211> 4
25  <212> PRT
    <213> Homo sapiens

    <220>
30  <221> CARACTERÍSTICA MISCELÁNEA
    <222> (2)..(2)
    <223> Xaa = Glu o Asp

    <400> 133
35  Ile Xaa Gly Arg

    1

    <210> 134
    <211> 5
40  <212> PRT
    <213> Homo sapiens

    <400> 134
    Asp Asp Asp Asp Lys
45  1                               5

    <210> 135
    <211> 7
    <212> PRT
50  <213> Homo sapiens

    <400> 135
    Glu Asn Leu Tyr Phe Gln Gly
    1                               5

55  <210> 136
    <211> 8
    <212> PRT
    <213> Homo sapiens

60  <400> 136

```

Leu Glu Val Leu Phe Gln Gly Pro
 1 5
 <210> 137
 <211> 100
 5 <212> PRT
 <213> Homo sapiens
 <400> 137
 Gly Ser Leu Gln Asp Ser Glu Val Asn Gln Glu Ala Lys Pro Glu Val
 1 5 10 15
 Lys Pro Glu Val Lys Pro Glu Thr His Ile Asn Leu Lys Val Ser Asp
 20 25 30
 Gly Ser Ser Glu Ile Phe Phe Lys Ile Lys Lys Thr Thr Pro Leu Arg
 35 40 45
 Arg Leu Met Glu Ala Phe Ala Lys Arg Gln Gly Lys Glu Met Asp Ser
 50 55 60
 Leu Thr Phe Leu Tyr Asp Gly Ile Glu Ile Gln Ala Asp Gln Thr Pro
 65 70 75 80
 10 Glu Asp Leu Asp Met Glu Asp Asn Asp Ile Ile Glu Ala His Arg Glu
 85 90 95
 Gln Ile Gly Gly
 100
 <210> 138
 <211> 1289
 15 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> Proteína de fusión p97
 20 <400> 138

ES 2 762 672 T3

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

ES 2 762 672 T3

180	185	190
Pro Ser Ser Glu Lys Tyr Glu Asn Thr Lys Thr Cys Arg Gly Pro Asp 195 200 205		
Gly Glu Leu His Ala Asn Leu Leu Cys Pro Val Asp Val Leu Asp Val 210 215 220		
Pro Glu Gly Thr Leu Pro Asp Lys Gln Ser Thr Glu Gln Ala Ile Gln 225 230 235 240		
Leu Leu Glu Lys Met Lys Thr Ser Ala Ser Pro Phe Phe Leu Ala Val 245 250 255		
Gly Tyr His Lys Pro His Ile Pro Phe Arg Tyr Pro Lys Glu Phe Gln 260 265 270		
Lys Leu Tyr Pro Leu Glu Asn Ile Thr Leu Ala Pro Asp Pro Glu Val 275 280 285		
Pro Asp Gly Leu Pro Pro Val Ala Tyr Asn Pro Trp Met Asp Ile Arg 290 295 300		
Gln Arg Glu Asp Val Gln Ala Leu Asn Ile Ser Val Pro Tyr Gly Pro 305 310 315 320		
Ile Pro Val Asp Phe Gln Arg Lys Ile Arg Gln Ser Tyr Phe Ala Ser 325 330 335		
Val Ser Tyr Leu Asp Thr Gln Val Gly Arg Leu Leu Ser Ala Leu Asp 340 345 350		
Asp Leu Gln Leu Ala Asn Ser Thr Ile Ile Ala Phe Thr Ser Asp His 355 360 365		
Gly Trp Ala Leu Gly Glu His Gly Glu Trp Ala Lys Tyr Ser Asn Phe 370 375 380		
Asp Val Ala Thr His Val Pro Leu Ile Phe Tyr Val Pro Gly Arg Thr 385 390 395 400		
Ala Ser Leu Pro Glu Ala Gly Glu Lys Leu Phe Pro Tyr Leu Asp Pro 405 410 415		
Phe Asp Ser Ala Ser Gln Leu Met Glu Pro Gly Arg Gln Ser Met Asp 420 425 430		

ES 2 762 672 T3

Leu Val Glu Leu Val Ser Leu Phe Pro Thr Leu Ala Gly Leu Ala Gly
 435 440 445
 Leu Gln Val Pro Pro Arg Cys Pro Val Pro Ser Phe His Val Glu Leu
 450 455 460
 Cys Arg Glu Gly Lys Asn Leu Leu Lys His Phe Arg Phe Arg Asp Leu
 465 470 475 480
 Glu Glu Asp Pro Tyr Leu Pro Gly Asn Pro Arg Glu Leu Ile Ala Tyr
 485 490 495
 Ser Gln Tyr Pro Arg Pro Ser Asp Ile Pro Gln Trp Asn Ser Asp Lys
 500 505 510
 Pro Ser Leu Lys Asp Ile Lys Ile Met Gly Tyr Ser Ile Arg Thr Ile
 515 520 525
 Asp Tyr Arg Tyr Thr Val Trp Val Gly Phe Asn Pro Asp Glu Phe Leu
 530 535 540
 Ala Asn Phe Ser Asp Ile His Ala Gly Glu Leu Tyr Phe Val Asp Ser
 545 550 555 560
 Asp Pro Leu Gln Asp His Asn Met Tyr Asn Asp Ser Gln Gly Gly Asp
 565 570 575
 Leu Phe Gln Leu Leu Met Pro Glu Ala Ala Ala Lys Glu Ala Ala Ala
 580 585 590
 Lys Glu Ala Ala Ala Lys Gly Met Glu Val Arg Trp Cys Ala Thr Ser
 595 600 605
 Asp Pro Glu Gln His Lys Cys Gly Asn Met Ser Glu Ala Phe Arg Glu
 610 615 620
 Ala Gly Ile Gln Pro Ser Leu Leu Cys Val Arg Gly Thr Ser Ala Asp
 625 630 635 640
 His Cys Val Gln Leu Ile Ala Ala Gln Glu Ala Asp Ala Ile Thr Leu
 645 650 655
 Asp Gly Gly Ala Ile Tyr Glu Ala Gly Lys Glu His Gly Leu Lys Pro
 660 665 670
 Val Val Gly Glu Val Tyr Asp Gln Glu Val Gly Thr Ser Tyr Tyr Ala
 675 680 685

ES 2 762 672 T3

Val 690	Ala	Val	Val	Arg	Arg	Ser 695	Ser	His	Val	Thr	Ile 700	Asp	Thr	Leu	Lys
Gly 705	Val	Lys	Ser	Cys	His 710	Thr	Gly	Ile	Asn	Arg 715	Thr	Val	Gly	Trp	Asn 720
Val	Pro	Val	Gly	Tyr 725	Leu	Val	Glu	Ser	Gly 730	Arg	Leu	Ser	Val	Met	Gly
Cys	Asp	Val	Leu 740	Lys	Ala	Val	Ser	Asp 745	Tyr	Phe	Gly	Gly	Ser	Cys	Val
Pro	Gly	Ala 755	Gly	Glu	Thr	Ser	Tyr 760	Ser	Glu	Ser	Leu	Cys 765	Arg	Leu	Cys
Arg	Gly 770	Asp	Ser	Ser	Gly	Glu 775	Gly	Val	Cys	Asp	Lys 780	Ser	Pro	Leu	Glu
Arg 785	Tyr	Tyr	Asp	Tyr	Ser 790	Gly	Ala	Phe	Arg	Cys 795	Leu	Ala	Glu	Gly	Ala 800
Gly	Asp	Val	Ala	Phe 805	Val	Lys	His	Ser	Thr 810	Val	Leu	Glu	Asn	Thr 815	Asp
Gly	Lys	Thr	Leu 820	Pro	Ser	Trp	Gly	Gln 825	Ala	Leu	Leu	Ser	Gln 830	Asp	Phe
Glu	Leu 835	Leu	Cys	Arg	Asp	Gly	Ser 840	Arg	Ala	Asp	Val	Thr 845	Glu	Trp	Arg
Gln 850	Cys	His	Leu	Ala	Arg	Val 855	Pro	Ala	His	Ala	Val 860	Val	Val	Arg	Ala
Asp 865	Thr	Asp	Gly	Gly	Leu 870	Ile	Phe	Arg	Leu	Leu 875	Asn	Glu	Gly	Gln	Arg 880
Leu	Phe	Ser	His 885	Glu	Gly	Ser	Ser	Phe	Gln 890	Met	Phe	Ser	Ser	Glu 895	Ala
Tyr	Gly	Gln	Lys 900	Asp	Leu	Leu	Phe	Lys 905	Asp	Ser	Thr	Ser	Glu 910	Leu	Val
Pro	Ile 915	Ala	Thr	Gln	Thr	Tyr	Glu 920	Ala	Trp	Leu	Gly	His 925	Glu	Tyr	Leu
His 930	Ala	Met	Lys	Gly	Leu	Leu 935	Cys	Asp	Pro	Asn	Arg 940	Leu	Pro	Pro	Tyr

ES 2 762 672 T3

Leu	Arg	Trp	Cys	Val	Leu	Ser	Thr	Pro	Glu	Ile	Gln	Lys	Cys	Gly	Asp	
945					950					955					960	
Met	Ala	Val	Ala	Phe	Arg	Arg	Gln	Arg	Leu	Lys	Pro	Glu	Ile	Gln	Cys	
				965					970						975	
Val	Ser	Ala	Lys	Ser	Pro	Gln	His	Cys	Met	Glu	Arg	Ile	Gln	Ala	Glu	
			980					985						990		
Gln	Val	Asp	Ala	Val	Thr	Leu	Ser	Gly	Glu	Asp	Ile	Tyr	Thr	Ala	Gly	
		995					1000						1005			
Lys	Thr	Tyr	Gly	Leu	Val	Pro	Ala	Ala	Gly	Glu	His	Tyr	Ala	Pro		
	1010					1015						1020				
Glu	Asp	Ser	Ser	Asn	Ser	Tyr	Tyr	Val	Val	Ala	Val	Val	Arg	Arg		
	1025					1030						1035				
Asp	Ser	Ser	His	Ala	Phe	Thr	Leu	Asp	Glu	Leu	Arg	Gly	Lys	Arg		
	1040					1045					1050					
Ser	Cys	His	Ala	Gly	Phe	Gly	Ser	Pro	Ala	Gly	Trp	Asp	Val	Pro		
	1055					1060					1065					
Val	Gly	Ala	Leu	Ile	Gln	Arg	Gly	Phe	Ile	Arg	Pro	Lys	Asp	Cys		
	1070					1075					1080					
Asp	Val	Leu	Thr	Ala	Val	Ser	Glu	Phe	Phe	Asn	Ala	Ser	Cys	Val		
	1085					1090					1095					
Pro	Val	Asn	Asn	Pro	Lys	Asn	Tyr	Pro	Ser	Ser	Leu	Cys	Ala	Leu		
	1100					1105					1110					
Cys	Val	Gly	Asp	Glu	Gln	Gly	Arg	Asn	Lys	Cys	Val	Gly	Asn	Ser		
	1115					1120					1125					
Gln	Glu	Arg	Tyr	Tyr	Gly	Tyr	Arg	Gly	Ala	Phe	Arg	Cys	Leu	Val		
	1130					1135					1140					
Glu	Asn	Ala	Gly	Asp	Val	Ala	Phe	Val	Arg	His	Thr	Thr	Val	Phe		
	1145					1150					1155					
Asp	Asn	Thr	Asn	Gly	His	Asn	Ser	Glu	Pro	Trp	Ala	Ala	Glu	Leu		
	1160					1165					1170					
Arg	Ser	Glu	Asp	Tyr	Glu	Leu	Leu	Cys	Pro	Asn	Gly	Ala	Arg	Ala		

ES 2 762 672 T3

1175

1180

1185

Glu Val Ser Gln Phe Ala Ala Cys Asn Leu Ala Gln Ile Pro Pro
1190 1195 1200

His Ala Val Met Val Arg Pro Asp Thr Asn Ile Phe Thr Val Tyr
1205 1210 1215

Gly Leu Leu Asp Lys Ala Gln Asp Leu Phe Gly Asp Asp His Asn
1220 1225 1230

Lys Asn Gly Phe Lys Met Phe Asp Ser Ser Asn Tyr His Gly Gln
1235 1240 1245

Asp Leu Leu Phe Lys Asp Ala Thr Val Arg Ala Val Pro Val Gly
1250 1255 1260

Glu Lys Thr Thr Tyr Arg Gly Trp Leu Gly Leu Asp Tyr Val Ala
1265 1270 1275

Ala Leu Glu Gly Met Ser Ser Gln Gln Cys Ser
1280 1285

<210> 139

<211> 1289

<212> PRT

<213> Secuencia Artificial

<220>

<223> Proteína de fusión p97

<400> 139

Met Glu Trp Ser Trp Val Phe Leu Phe Phe Leu Ser Val Thr Thr Gly
1 5 10 15

Val His Ser Asp Tyr Lys Asp Asp Asp Asp Lys Glu Gln Lys Leu Ile
20 25 30

Ser Glu Glu Asp Leu His His His His His His His His His Gly
35 40 45

Gly Gly Gly Glu Asn Leu Tyr Phe Gln Gly Gly Met Glu Val Arg Trp
50 55 60

Cys Ala Thr Ser Asp Pro Glu Gln His Lys Cys Gly Asn Met Ser Glu
65 70 75 80

Ala Phe Arg Glu Ala Gly Ile Gln Pro Ser Leu Leu Cys Val Arg Gly
85 90 95

ES 2 762 672 T3

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

ES 2 762 672 T3

340	345	350
Ser Ser Glu Ala Tyr Gly Gln Lys Asp Leu Leu Phe Lys Asp Ser Thr 355 360 365		
Ser Glu Leu Val Pro Ile Ala Thr Gln Thr Tyr Glu Ala Trp Leu Gly 370 375 380		
His Glu Tyr Leu His Ala Met Lys Gly Leu Leu Cys Asp Pro Asn Arg 385 390 395 400		
Leu Pro Pro Tyr Leu Arg Trp Cys Val Leu Ser Thr Pro Glu Ile Gln 405 410 415		
Lys Cys Gly Asp Met Ala Val Ala Phe Arg Arg Gln Arg Leu Lys Pro 420 425 430		
Glu Ile Gln Cys Val Ser Ala Lys Ser Pro Gln His Cys Met Glu Arg 435 440 445		
Ile Gln Ala Glu Gln Val Asp Ala Val Thr Leu Ser Gly Glu Asp Ile 450 455 460		
Tyr Thr Ala Gly Lys Thr Tyr Gly Leu Val Pro Ala Ala Gly Glu His 465 470 475 480		
Tyr Ala Pro Glu Asp Ser Ser Asn Ser Tyr Tyr Val Val Ala Val Val 485 490 495		
Arg Arg Asp Ser Ser His Ala Phe Thr Leu Asp Glu Leu Arg Gly Lys 500 505 510		
Arg Ser Cys His Ala Gly Phe Gly Ser Pro Ala Gly Trp Asp Val Pro 515 520 525		
Val Gly Ala Leu Ile Gln Arg Gly Phe Ile Arg Pro Lys Asp Cys Asp 530 535 540		
Val Leu Thr Ala Val Ser Glu Phe Phe Asn Ala Ser Cys Val Pro Val 545 550 555 560		
Asn Asn Pro Lys Asn Tyr Pro Ser Ser Leu Cys Ala Leu Cys Val Gly 565 570 575		
Asp Glu Gln Gly Arg Asn Lys Cys Val Gly Asn Ser Gln Glu Arg Tyr 580 585 590		

ES 2 762 672 T3

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

ES 2 762 672 T3

Arg Val His Ala Gly Asn Phe Ser Thr Ile Pro Gln Tyr Phe Lys Glu
 850 855 860
 Asn Gly Tyr Val Thr Met Ser Val Gly Lys Val Phe His Pro Gly Ile
 865 870 875 880
 Ser Ser Asn His Thr Asp Asp Ser Pro Tyr Ser Trp Ser Phe Pro Pro
 885 890 895
 Tyr His Pro Ser Ser Glu Lys Tyr Glu Asn Thr Lys Thr Cys Arg Gly
 900 905 910
 Pro Asp Gly Glu Leu His Ala Asn Leu Leu Cys Pro Val Asp Val Leu
 915 920 925
 Asp Val Pro Glu Gly Thr Leu Pro Asp Lys Gln Ser Thr Glu Gln Ala
 930 935 940
 Ile Gln Leu Leu Glu Lys Met Lys Thr Ser Ala Ser Pro Phe Phe Leu
 945 950 955 960
 Ala Val Gly Tyr His Lys Pro His Ile Pro Phe Arg Tyr Pro Lys Glu
 965 970 975
 Phe Gln Lys Leu Tyr Pro Leu Glu Asn Ile Thr Leu Ala Pro Asp Pro
 980 985 990
 Glu Val Pro Asp Gly Leu Pro Pro Val Ala Tyr Asn Pro Trp Met Asp
 995 1000 1005
 Ile Arg Gln Arg Glu Asp Val Gln Ala Leu Asn Ile Ser Val Pro
 1010 1015 1020
 Tyr Gly Pro Ile Pro Val Asp Phe Gln Arg Lys Ile Arg Gln Ser
 1025 1030 1035
 Tyr Phe Ala Ser Val Ser Tyr Leu Asp Thr Gln Val Gly Arg Leu
 1040 1045 1050
 Leu Ser Ala Leu Asp Asp Leu Gln Leu Ala Asn Ser Thr Ile Ile
 1055 1060 1065
 Ala Phe Thr Ser Asp His Gly Trp Ala Leu Gly Glu His Gly Glu
 1070 1075 1080
 Trp Ala Lys Tyr Ser Asn Phe Asp Val Ala Thr His Val Pro Leu
 1085 1090 1095

ES 2 762 672 T3

Ile Phe Tyr Val Pro Gly Arg Thr Ala Ser Leu Pro Glu Ala Gly
1100 1105 1110

Glu Lys Leu Phe Pro Tyr Leu Asp Pro Phe Asp Ser Ala Ser Gln
1115 1120 1125

Leu Met Glu Pro Gly Arg Gln Ser Met Asp Leu Val Glu Leu Val
1130 1135 1140

Ser Leu Phe Pro Thr Leu Ala Gly Leu Ala Gly Leu Gln Val Pro
1145 1150 1155

Pro Arg Cys Pro Val Pro Ser Phe His Val Glu Leu Cys Arg Glu
1160 1165 1170

Gly Lys Asn Leu Leu Lys His Phe Arg Phe Arg Asp Leu Glu Glu
1175 1180 1185

Asp Pro Tyr Leu Pro Gly Asn Pro Arg Glu Leu Ile Ala Tyr Ser
1190 1195 1200

Gln Tyr Pro Arg Pro Ser Asp Ile Pro Gln Trp Asn Ser Asp Lys
1205 1210 1215

Pro Ser Leu Lys Asp Ile Lys Ile Met Gly Tyr Ser Ile Arg Thr
1220 1225 1230

Ile Asp Tyr Arg Tyr Thr Val Trp Val Gly Phe Asn Pro Asp Glu
1235 1240 1245

Phe Leu Ala Asn Phe Ser Asp Ile His Ala Gly Glu Leu Tyr Phe
1250 1255 1260

Val Asp Ser Asp Pro Leu Gln Asp His Asn Met Tyr Asn Asp Ser
1265 1270 1275

Gln Gly Gly Asp Leu Phe Gln Leu Leu Met Pro
1280 1285

<210> 140

<211> 611

5 <212> PRT

<213> Secuencia Artificial

<220>

<223> Proteína de fusión p97

10

<400> 140

ES 2 762 672 T3

Met 1	Glu	Trp	Ser 5	Trp	Val	Phe	Leu	Phe	Phe 10	Leu	Ser	Val	Thr	Thr 15	Gly
Val	His	Ser	Asp 20	Tyr	Lys	Asp	Asp	Asp 25	Asp	Lys	Glu	Gln	Lys 30	Leu	Ile
Ser	Glu	Glu 35	Asp	Leu	His	His	His 40	His	His	His	His	His 45	His	His	Gly
Gly	Gly 50	Gly	Glu	Asn	Leu	Tyr 55	Phe	Gln	Gly	Asp	Ser 60	Ser	His	Ala	Phe
Thr 65	Leu	Asp	Glu	Leu	Arg 70	Tyr	Glu	Ala	Ala	Ala 75	Lys	Glu	Ala	Ala	Ala 80
Lys	Glu	Ala	Ala	Ala 85	Lys	Ser	Glu	Thr	Gln 90	Ala	Asn	Ser	Thr	Thr 95	Asp
Ala	Leu	Asn	Val 100	Leu	Leu	Ile	Ile	Val 105	Asp	Asp	Leu	Arg	Pro 110	Ser	Leu
Gly	Cys	Tyr 115	Gly	Asp	Lys	Leu	Val 120	Arg	Ser	Pro	Asn	Ile 125	Asp	Gln	Leu
Ala	Ser 130	His	Ser	Leu	Leu	Phe 135	Gln	Asn	Ala	Phe	Ala 140	Gln	Gln	Ala	Val
Cys 145	Ala	Pro	Ser	Arg	Val 150	Ser	Phe	Leu	Thr	Gly 155	Arg	Arg	Pro	Asp	Thr 160
Thr	Arg	Leu	Tyr	Asp 165	Phe	Asn	Ser	Tyr	Trp 170	Arg	Val	His	Ala	Gly 175	Asn
Phe	Ser	Thr	Ile 180	Pro	Gln	Tyr	Phe	Lys 185	Glu	Asn	Gly	Tyr	Val 190	Thr	Met
Ser	Val	Gly 195	Lys	Val	Phe	His	Pro 200	Gly	Ile	Ser	Ser	Asn 205	His	Thr	Asp
Asp	Ser 210	Pro	Tyr	Ser	Trp	Ser 215	Phe	Pro	Pro	Tyr	His 220	Pro	Ser	Ser	Glu
Lys 225	Tyr	Glu	Asn	Thr	Lys 230	Thr	Cys	Arg	Gly	Pro 235	Asp	Gly	Glu	Leu	His 240
Ala	Asn	Leu	Leu	Cys 245	Pro	Val	Asp	Val	Leu 250	Asp	Val	Pro	Glu	Gly 255	Thr

ES 2 762 672 T3

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

ES 2 762 672 T3

500

505

510

Tyr Leu Pro Gly Asn Pro Arg Glu Leu Ile Ala Tyr Ser Gln Tyr Pro
515 520 525

Arg Pro Ser Asp Ile Pro Gln Trp Asn Ser Asp Lys Pro Ser Leu Lys
530 535 540

Asp Ile Lys Ile Met Gly Tyr Ser Ile Arg Thr Ile Asp Tyr Arg Tyr
545 550 555 560

Thr Val Trp Val Gly Phe Asn Pro Asp Glu Phe Leu Ala Asn Phe Ser
565 570 575

Asp Ile His Ala Gly Glu Leu Tyr Phe Val Asp Ser Asp Pro Leu Gln
580 585 590

Asp His Asn Met Tyr Asn Asp Ser Gln Gly Gly Asp Leu Phe Gln Leu
595 600 605

Leu Met Pro
610

<210> 141

<211> 611

<212> PRT

<213> Secuencia Artificial

<220>

<223> Proteína de fusión p97

<400> 141

Met Glu Trp Ser Trp Val Phe Leu Phe Phe Leu Ser Val Thr Thr Gly
1 5 10 15

Val His Ser Asp Tyr Lys Asp Asp Asp Asp Lys Glu Gln Lys Leu Ile
20 25 30

Ser Glu Glu Asp Leu His His His His His His His His His His Gly
35 40 45

Gly Gly Gly Glu Asn Leu Tyr Phe Gln Gly Ser Glu Thr Gln Ala Asn
50 55 60

Ser Thr Thr Asp Ala Leu Asn Val Leu Leu Ile Ile Val Asp Asp Leu
65 70 75 80

Arg Pro Ser Leu Gly Cys Tyr Gly Asp Lys Leu Val Arg Ser Pro Asn
85 90 95

ES 2 762 672 T3

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

ES 2 762 672 T3

340	345	350
Asp Leu Gln Leu Ala Asn Ser Thr Ile Ile Ala Phe Thr Ser Asp His		
355	360	365
Gly Trp Ala Leu Gly Glu His Gly Glu Trp Ala Lys Tyr Ser Asn Phe		
370	375	380
Asp Val Ala Thr His Val Pro Leu Ile Phe Tyr Val Pro Gly Arg Thr		
385	390	400
Ala Ser Leu Pro Glu Ala Gly Glu Lys Leu Phe Pro Tyr Leu Asp Pro		
	405	410
Phe Asp Ser Ala Ser Gln Leu Met Glu Pro Gly Arg Gln Ser Met Asp		
	420	425
Leu Val Glu Leu Val Ser Leu Phe Pro Thr Leu Ala Gly Leu Ala Gly		
	435	440
Leu Gln Val Pro Pro Arg Cys Pro Val Pro Ser Phe His Val Glu Leu		
	450	455
Cys Arg Glu Gly Lys Asn Leu Leu Lys His Phe Arg Phe Arg Asp Leu		
465	470	475
Glu Glu Asp Pro Tyr Leu Pro Gly Asn Pro Arg Glu Leu Ile Ala Tyr		
	485	490
Ser Gln Tyr Pro Arg Pro Ser Asp Ile Pro Gln Trp Asn Ser Asp Lys		
	500	505
Pro Ser Leu Lys Asp Ile Lys Ile Met Gly Tyr Ser Ile Arg Thr Ile		
	515	520
Asp Tyr Arg Tyr Thr Val Trp Val Gly Phe Asn Pro Asp Glu Phe Leu		
	530	535
Ala Asn Phe Ser Asp Ile His Ala Gly Glu Leu Tyr Phe Val Asp Ser		
	545	550
Asp Pro Leu Gln Asp His Asn Met Tyr Asn Asp Ser Gln Gly Gly Asp		
	565	570
Leu Phe Gln Leu Leu Met Pro Glu Ala Ala Ala Lys Glu Ala Ala Ala		
	580	585
Lys Glu Ala Ala Ala Lys Asp Ser Ser His Ala Phe Thr Leu Asp Glu		
	595	600
Leu Arg Tyr		
610		

5 <210> 142
 <211> 603
 <212> PRT

<213> Secuencia Artificial

<220>

<223> Proteína de fusión p97

5

<400> 142

Met Glu Trp Ser Trp Val Phe Leu Phe Phe Leu Ser Val Thr Thr Gly
1 5 10 15

Val His Ser Asp Tyr Lys Asp Asp Asp Lys Glu Gln Lys Leu Ile
20 25 30

Ser Glu Glu Asp Leu His His His His His His His His His Gly
35 40 45

Gly Gly Gly Glu Asn Leu Tyr Phe Gln Gly Thr Asp Ala Leu Asn Val
50 55 60

Leu Leu Ile Ile Val Asp Asp Leu Arg Pro Ser Leu Gly Cys Tyr Gly
65 70 75 80

Asp Lys Leu Val Arg Ser Pro Asn Ile Asp Gln Leu Ala Ser His Ser
85 90 95

Leu Leu Phe Gln Asn Ala Phe Ala Gln Gln Ala Val Cys Ala Pro Ser
100 105 110

Arg Val Ser Phe Leu Thr Gly Arg Arg Pro Asp Thr Thr Arg Leu Tyr
115 120 125

Asp Phe Asn Ser Tyr Trp Arg Val His Ala Gly Asn Phe Ser Thr Ile
130 135 140

Pro Gln Tyr Phe Lys Glu Asn Gly Tyr Val Thr Met Ser Val Gly Lys
145 150 155 160

Val Phe His Pro Gly Ile Ser Ser Asn His Thr Asp Asp Ser Pro Tyr
165 170 175

Ser Trp Ser Phe Pro Pro Tyr His Pro Ser Ser Glu Lys Tyr Glu Asn

ES 2 762 672 T3

180	185	190
Thr Lys Thr Cys Arg Gly Pro Asp Gly Glu Leu His Ala Asn Leu Leu 195 200 205		
Cys Pro Val Asp Val Leu Asp Val Pro Glu Gly Thr Leu Pro Asp Lys 210 215 220		
Gln Ser Thr Glu Gln Ala Ile Gln Leu Leu Glu Lys Met Lys Thr Ser 225 230 235 240		
Ala Ser Pro Phe Phe Leu Ala Val Gly Tyr His Lys Pro His Ile Pro 245 250 255		
Phe Arg Tyr Pro Lys Glu Phe Gln Lys Leu Tyr Pro Leu Glu Asn Ile 260 265 270		
Thr Leu Ala Pro Asp Pro Glu Val Pro Asp Gly Leu Pro Pro Val Ala 275 280 285		
Tyr Asn Pro Trp Met Asp Ile Arg Gln Arg Glu Asp Val Gln Ala Leu 290 295 300		
Asn Ile Ser Val Pro Tyr Gly Pro Ile Pro Val Asp Phe Gln Arg Lys 305 310 315 320		
Ile Arg Gln Ser Tyr Phe Ala Ser Val Ser Tyr Leu Asp Thr Gln Val 325 330 335		
Gly Arg Leu Leu Ser Ala Leu Asp Asp Leu Gln Leu Ala Asn Ser Thr 340 345 350		
Ile Ile Ala Phe Thr Ser Asp His Gly Trp Ala Leu Gly Glu His Gly 355 360 365		
Glu Trp Ala Lys Tyr Ser Asn Phe Asp Val Ala Thr His Val Pro Leu 370 375 380		
Ile Phe Tyr Val Pro Gly Arg Thr Ala Ser Leu Pro Glu Ala Gly Glu 385 390 395 400		
Lys Leu Phe Pro Tyr Leu Asp Pro Phe Asp Ser Ala Ser Gln Leu Met 405 410 415		
Glu Pro Gly Arg Gln Ser Met Asp Leu Val Glu Leu Val Ser Leu Phe 420 425 430		

ES 2 762 672 T3

Pro Thr Leu Ala Gly Leu Ala Gly Leu Gln Val Pro Pro Arg Cys Pro
435 440 445

Val Pro Ser Phe His Val Glu Leu Cys Arg Glu Gly Lys Asn Leu Leu
450 455 460

Lys His Phe Arg Phe Arg Asp Leu Glu Glu Asp Pro Tyr Leu Pro Gly
465 470 475 480

Asn Pro Arg Glu Leu Ile Ala Tyr Ser Gln Tyr Pro Arg Pro Ser Asp
485 490 495

Ile Pro Gln Trp Asn Ser Asp Lys Pro Ser Leu Lys Asp Ile Lys Ile
500 505 510

Met Gly Tyr Ser Ile Arg Thr Ile Asp Tyr Arg Tyr Thr Val Trp Val
515 520 525

Gly Phe Asn Pro Asp Glu Phe Leu Ala Asn Phe Ser Asp Ile His Ala
530 535 540

Gly Glu Leu Tyr Phe Val Asp Ser Asp Pro Leu Gln Asp His Asn Met
545 550 555 560

Tyr Asn Asp Ser Gln Gly Gly Asp Leu Phe Gln Leu Leu Met Pro Glu
565 570 575

Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Asp Ser
580 585 590

Ser His Ala Phe Thr Leu Asp Glu Leu Arg Tyr
595 600

<210> 143
<211> 3870
5 <212> DNA
<213> Secuencia Artificial

<220>
10 <223> Polinucleótido que codifica la proteína de fusión p97

<400> 143	
atggaatgga gctgggtctt tctcttcttc ctgtcagtaa cgactggtgt ccactccgac	60
tacaaggacg acgacgacaa agagcagaag ctgatctccg aagaggacct gcaccacat	120
catcaccatc accaccatca cggaggcggg ggagagaacc tgtactttca gggctcggaa	180
actcaggcca actccaccac agatgcactc aacgtgctgc tgatcatcgt agatgacctc	240
cgaccttctc tgggctgtta cggcgacaag ctagtacgga gcccaaacat cgaccagctc	300

ES 2 762 672 T3

gcatcgact ctctcctatt ccagaacgca ttcgcccagc aggtgtctg tgetccctcc 360
 cgagtgtcct tcctcacggg tcggagaccc gataccacga ggttatatga cttcaactca 420
 tactggcgcg tgcattgccg taacttttct actatacccc agtattttta agaaaatggc 480
 tatgttacaa tgtccgttgg caaggtatct catcctggta ttagcagcaa ccacacagat 540
 gactctccgt atagctgggc attcccacca taccaccctt ccagcgaaaa gtacgaaaac 600
 acaaagactt gccggggccc agatggcgaa ctgcacgcaa atctgctgtg ccctgtagat 660
 gtcttggagc tgcccgaagg tactctgccc gacaaacagt ccacagaaca ggcaatccaa 720
 ctcttgaaa agatgaaaac gagcgctgcc ccttcttctc tcgccgtggg ctaccacaag 780
 ccccatatcc cgttttagata ccccaaggaa ttccagaaac tgtacccctt ggaaaacatc 840
 actctcgcg cccgacccga agtgccagac ggactccctc ctgttgccca caacccttgg 900
 atggacatca gacaacgtga agatgtgcag gccctgaaca tctcagtgcc ttacggcccc 960
 attccagttg acttccagag gaagattcgg cagtcctact tcgcctccgt tagttacctg 1020
 gacacccaag tgggttagact cctgagcgcc ttggacgata tccagctcgc aaacagcacc 1080
 atcattgcct tcaccagcga ccatggttgg gcgctgggtg aacatggaga atgggctaaa 1140
 tattcaaatt tcgacgttgc gaccacgctc ccattgatct tctacgtgcc tggacgaaca 1200
 gcctccttgc ctgaagccgg ggaaaagtgt ttcccatatc tggacccttt cgattctgcg 1260
 agccaactca tggaaacctg gcgacagagc atggacctgg tggaaactgt cagtttatct 1320
 ccaaccctgg caggccttgc aggcctccaa gttccacctc ggtgtccgtt tccctcatc 1380
 cacgtcgaac tctgtcgcga aggtaaaaac ctctcaagc attttctgtt tcgggacctc 1440
 gaagaagacc catacctgcc agggaatcca agggaactga ttgcctacag ccagtaccct 1500
 agacctagcg acatcccaca gtggaacagc gacaagccct ccctcaagga cattaaaatc 1560
 atgggttata gtatccggac tattgactac aggtataccg tgtgggtggg ttccaacca 1620
 gacgaatttc tcgccaattt ctccgacatc cagcgggcg aactgtatct cgttgattcc 1680
 gatccactgc aagatcataa tatgtacaac gatagtcaag ggggtgacct cttccagttg 1740
 ctaatgccag aagccgccgc gaaagaagcc gccgcaaaag aagccgctgc caaaggcatg 1800
 gaagtgcgtt ggtgcgccac ctctgacccc gagcagcaca agtgccgcaa catgtccgag 1860
 gccttcagag aggcggcat ccagccttct ctgctgtgtg tgcggggcac ctctgccgac 1920
 cattgcgtgc agctgatcgc cggccaggaa gccgacgcta tcacactgga tggcggcgct 1980
 atctacgagg ctggcaaaag gcacggcctg aagcccgctg tggcgaggt gtacgatcag 2040
 gaagtgggca cctcctacta cggcgtggct gtctgtcgga gatcctcca cgtgaccatc 2100
 gacaccctga agggcggtga gtctgccac accggcatca acagaaccgt gggtggaac 2160
 gtgcccgtgg gctacctggg ggaatccggc agactgtccg tgatgggctg cgacgtgctg 2220

ES 2 762 672 T3

```

aaggccgtgt ccgattactt cggcggctct tgtgtgcctg gcgctggcga gacatcctac 2280
tccgagtgcc tgtgcagact gtgcaggggc gactcttctg gcgagggcgt gtgcgacaag 2340
tcccctctgg aacggtacta cgactactcc ggcgccttca gatgcctggc tgaaggtgct 2400
ggcgacgtgg ccttcgtgaa gcactccacc gtgctggaaa acaccgacgg caagaccctg 2460
ccttcttggg gccaggcact gctgtcccag gacttcgagc tgctgtgccg ggatggctcc 2520
agagccgatg tgacagagtg gcggcagtg cacttgcca gagtgccctg tcatgctgtg 2580
gtcgtgcgcg ccgatacaga tggcggcctg atcttccggc tgctgaacga gggccagcgg 2640
ctgttctctc acgagggctc cagcttccag atgttctcca gcgaggccta cggccagaag 2700
gacctgctgt tcaaggactc cactccgag ctggtgccta tcgccacca gacctatgag 2760
gcttggtggt gccacgagta cctgcacgct atgaaggagc tgctgtgcga cccaaccgg 2820
ctgcctcctt atctgaggtg gtgcgtgctg tccacccccg agatccagaa atgcggcgat 2880
atggccgtgg cctttcggcg gcagagactg aagcctgaga tccagtgcgt gtccgccaag 2940
agccctcagc actgcatgga acggatccag gccgaacagg tggacgccgt gacactgtcc 3000
ggcgaggata tctacaccgc cggaaagacc tacggcctgg tgccagctgc tggcgagcat 3060
tacgcccctg aggactcctc caacagctac tacgtggtgg cagtcgtgcg ccgggactcc 3120
tctcacgcct ttaccctgga tgagctgcgg ggcaagagaa gctgtcacgc cggctttgga 3180
agccctgccg gatgggatgt gcctgtgggc gctctgatcc agcggggctt catcagaccc 3240
aaggactgtg atgtgctgac cgccgtgtct gagttcttca acgcctcctg tgtgcccggtg 3300
aacaacccca agaactaccc ctccagcctg tgcgccctgt gtgtgggaga tgagcagggc 3360
cggaacaaat gcgtgggcaa ctcccaggaa agatattacg gctacagagg cgccttccgg 3420
tgtctggtgg aaaacgccgg ggatgtggct tttgtgcggc acaccaccgt gttcgacaac 3480
accaatggcc acaactccga gccttgggcc gctgagctga gatccgagga ttacgaactg 3540
ctgtgtccca acggcgccag ggctgaggtg tcccagtttg ccgcctgtaa cctggcccag 3600
atccctcccc acgctgtgat ggtgcgaccc gacaccaaca tcttcaccgt gtacggcctg 3660
ctggacaagg cccaggatct gttcggcgac gaccacaaca agaacgggtt caagatgttc 3720
gactccagca actaccacgg acaggatctg ctgtttaaaag atgccaccgt gcgggccgtg 3780
ccagtgggcg aaaagaccac ctacagagga tggctgggac tggactacgt ggccgccctg 3840
gaaggcatgt cctcccagca gtgttcctga 3870

```

<210> 144

<211> 3870

<212> DNA

<213> Secuencia Artificial

<220>

<223> Polinucleótido que codifica la proteína de fusión p97

<400> 144

ES 2 762 672 T3

atggaatgga gctgggtctt tctcttcttc ctgtcagtaa cgactggtgt ccactccgac	60
tacaaggacg acgacgacaa agagcagaag ctgatctccg aagaggacct gcaccaccat	120
catcaccatc accaccatca cggaggcggt ggagagaacc tgtactttca gggcgcatg	180
gaagtgcgtt ggtgcgccac ctctgacccc gagcagcaca agtgcggcaa catgtccgag	240
gccttcagag aggccggcat ccagccttct ctgctgtgtg tgcggggcac ctctgccgac	300
cattgcgtgc agctgatcgc cgcccaggaa gccgacgcta tcacactgga tggcggcgt	360
atctacgagg ctggcaaaga gcacggcctg aagcccgtcg tggcgaggt gtacgatcag	420
gaagtgggca cctcctacta cgccgtggct gtcgtgcgga gatcctcca cgtgaccatc	480
gacaccctga agggcggtgaa gtcctgccac accggcatca acagaaccgt gggctggaac	540
gtgcccgtag gctacctggt ggaatccggc agactgtccg tgatgggctg cgacgtgctg	600
aaggccgtgt ccgattactt cggcggtctt tgtgtgcctg gcgctggcga gacatcctac	660
tccgagtccc tgtgcagact gtgcaggggc gactcttctg gcgagggcgt gtgcgacaag	720
tcccctctgg aacggtacta cgactactcc ggccgcttca gatgcctggc tgaagtgct	780
ggcgacgtgg ccttcgtgaa gcactccacc gtgctggaac acaccgacgg caagaccctg	840
ccttcttggg gccaggcact gctgtcccag gacttcgagc tgctgtgccg ggatggctcc	900
agagccgatg tgacagagtg gcggcagtcg cacctggcca gagtgcctgc tcatgctgtg	960
gtcgtgcgcg ccgatacaga tggcggcctg atcttccggc tgctgaacga gggccagcgg	1020
ctgttctctc acgagggctc cagcttccag atgttctcca gcgaggccta cggccagaag	1080
gacctgctgt tcaaggactc cacctccgag ctggtgccta tcgccacca gacctatgag	1140
gcttggtggg gccacgagta cctgcacgct atgaagggac tgctgtgcga cccaaccgg	1200
ctgcctcctt atctgaggtg gtgcgtgctg tccacccccg agatccagaa atgcggcgat	1260
atggccgtgg cctttcggcg gcagagactg aagcctgaga tccagtgcgt gtccgccaag	1320
agccctcagc actgcatgga acggatccag gccgaacagg tggacgccgt gacactgtcc	1380
ggcgaggata tctacaccgc cggaaagacc tacggcctgg tgccagctgc tggcgagcat	1440
tacgcccctg aggactcctc caacagctac tacgtggtgg cagtcgtgcg ccgggactcc	1500
tctcacgcct ttaccctgga tgagctgcgg ggcaagagaa gctgtcacgc cggctttgga	1560
agccctgccg gatgggatgt gcctgtgggc gctctgatcc agcggggctt catcagaccc	1620
aaggactgtg atgtgctgac cgccgtgtct gagttcttca acgcctcctg tgtgcccgtag	1680
aacaacccca agaactaccc ctccagcctg tgcgccctgt gtgtgggaga tgagcagggc	1740
cggaacaaat gcgtgggcaa ctcccaggaa agatattacg gctacagagg cgccttccgg	1800

ES 2 762 672 T3

tgtctggtgg	aaaacgccgg	ggatgtggct	tttgtgcggc	acaccaccgt	gttcgacaac	1860
accaatggcc	acaactccga	gccttggggc	gctgagctga	gatccgagga	ttacgaactg	1920
ctgtgtccca	acggcgccag	ggctgaggtg	tcccagtttg	ccgcctgtaa	cctggcccag	1980
atccctcccc	acgctgtgat	ggcgacccc	gacaccaaca	tcttcaccgt	gtacggcctg	2040
ctggacaagg	cccaggatct	gttcggcgac	gaccacaaca	agaacggggt	caagatgttc	2100
gactccagca	actaccacgg	acaggatctg	ctgtttaaag	atgccaccgt	gcggggccgtg	2160
ccagtgggcg	aaaagaccac	ctacagagga	tggctgggac	tggactacgt	ggccgccctg	2220
gaaggcatgt	cctcccagca	gtgttccgaa	gccgcccgcg	aagaagccgc	cgcaaaagaa	2280
gccgctgcc	aatcggaaac	tcaggccaac	tccaccacag	atgcactcaa	cgtgctgctg	2340
atcatcgtag	atgacctccg	accttctctg	ggctgttacg	gcgacaagct	agtacggagc	2400
ccaaacatcg	accagctcgc	atcgcaactc	ctcctattcc	agaacgcatt	cgcccagcag	2460
gctgtctgtg	ctccctcccg	agtgtccttc	ctcacgggtc	ggagaccgga	taccacgagg	2520
ttatatgact	tcaactcata	ctggcgcggtg	catgccggta	acttttctac	tataccccag	2580
tattttaaag	aaaatggcta	tgttacaatg	tccgttgga	aggtatttca	tcctggtatt	2640
agcagcaacc	acacagatga	ctctccgtat	agctggtcat	tcccaccata	ccaccctcc	2700
agcgaaggt	acgaaaacac	aaagacttgc	cggggcccag	atggcgaaact	gcacgcaaact	2760
ctgctgtgcc	ctgtagatgt	cttggaactg	cccgaaggta	ctctgcccga	caaacagtcc	2820
acagaacagg	caatccaact	ccttgaaaag	atgaaaacga	gcgcgtcccc	cttcttcctc	2880
gccgtgggct	accacaagcc	ccacatcccc	tttagatacc	ccaaggaatt	tcagaaactg	2940
tacccctctg	aaaacatcac	tctcgcgccc	gaccccgaa	tgccagacgg	actccctcct	3000
gttgccctaca	acccttggat	ggacatcaga	caacgtgaag	atgtgcaggc	cctgaacatc	3060
tcagtgcctt	acggccccat	tccagttgac	ttccagagga	agattcggca	gtcctacttc	3120
gcctccgtta	gttacctgga	cacccaagtg	ggtagactcc	tgagcgccct	ggacgatctc	3180
cagctcgcaa	acagcaccat	cattgccttc	accagcgacc	atggttgggc	gctgggtgaa	3240
catggagaat	gggctaaata	ttcaaatttc	gacgttgcca	cccacgtccc	attgatcttc	3300
tacgtgcctg	gacgaacagc	ctccttgcc	gaagccgggg	aaaagttggt	tccatatctg	3360
gaccctttcg	attctgcgag	ccaactcatg	gaacctgggc	gacagagcat	ggacctggtg	3420
gaactggtca	gtttatttcc	aaccctggca	ggccttgacg	gcctccaagt	tccacctcgg	3480
tgtcccgttc	cctcattcca	cgtcgaactc	tgtcggaag	gtaaaaacct	cctcaagcat	3540
tttcgttttc	gggacctcga	agaagacca	tacctgccag	ggaatccaag	ggaactgatt	3600
gcctacagcc	agtaccctag	acctagcgac	atcccacagt	ggaacagcga	caagccctcc	3660
ctcaaggaca	ttaaaatcat	gggttatagt	atccggacta	ttgactacag	gtataccgtg	3720
tgggtggggt	tcaaccacga	cgaatttctc	gccaaatttct	ccgacatcca	cgcgggcgaa	3780
ctgtatttctg	ttgattccga	tccactgcaa	gatcataata	tgtacaacga	tagtcaaggg	3840
ggtgacctct	tccagttgct	aatgccatga				3870

5 <210> 145
<211> 1836

<212> DNA
<213> Secuencia Artificial

<220>

5 <223> Polinucleótido que codifica la proteína de fusión p97

<400> 145

atggaatgga gctgggtctt tctcttcttc ctgtcagtaa cgactgggtg ccactccgac	60
tacaaggacg acgacgacaa agagcagaag ctgatctccg aagaggacct gcaccacccat	120
catcaccatc accaccatca cggaggcggt ggagagaacc tgtactttca gggcgactcc	180
tctcacgcct tcaccctgga cgagctgcgg tacgaagccg ccgcgaaaga agccgccgca	240
aaagaagccg ctgccaaatc ggaaactcag gccaaactcca ccacagatgc actcaacgtg	300
ctgctgatca tcgtagatga cctccgacct tctctgggct gttacggcga caagctagta	360
cggagcccaa acatcgacca gctcgcatcg cactctctcc tattccagaa cgcattcgcc	420
cagcaggctg tctgtgctcc tcctcgagtg tccttctca cgggtcggag acccgatacc	480
acgaggttat atgacttcaa ctcatctagg cgctgcatg ccggtaactt ttctactata	540
ccccagtatt ttaaagaaaa tggctatggt acaatgtccg ttggcaaggt atttcacct	600
ggtattagca gcaaccacac agatgactct ccgtatagct ggtcattccc accataccac	660
ccctccagcg aaaagtacga aaacacaaag acttgccggg gccagatgg cgaactgcac	720
gcaaactctg tgtgccctgt agatgtcttg gacgtgcccg aaggtactct gcccgacaaa	780
cagtccacag aacaggcaat ccaactcctt gaaaagatga aaacgagcgc gtcccccttc	840
ttcctcgccg tgggctacca caagccccac atcccgttta gataccccaa ggaatttcag	900
aaactgtacc ccctggaaaa catcactctc gcgcccagacc ccgaagtgcc agacggactc	960
cctcctgttg cctacaacc ttggatggac atcagacaac gtgaagatgt gcaggccctg	1020
aacatctcag tgccttacgg cccattcca gttgacttcc agaggaagat tcggcagtc	1080
tacttcgcct ccgttagtta cctggacacc caagtgggta gactcctgag cgccttgga	1140
gatctccagc tcgcaaacag caccatcatt gccttcacca gcgaccatgg ttgggcgctg	1200
ggtgaacatg gagaatgggc taaatattca aatttcgacg ttgcgaccca cgtcccattg	1260
atcttctacg tgcctggacg aacagcctcc ttgcctgaag ccggggaaaa gttgtttcca	1320
tatctggacc ctttcgatcc tgcgagccaa ctcatggaac ctgggacgaca gagcatggac	1380
ctggtggaac tggtcagttt atttccaacc ctggcaggcc ttgcaggcct ccaagttcca	1440
cctcgggtgc ccgttccttc attccacgtc gaactctgtc gcgaaggtaa aaacctcctc	1500
aagcattttc gttttcgga cctcgaagaa gaccataacc tgccaggga tccaaggga	1560
ctgattgcct acagccagta ccctagacct agcgacatcc cacagtggaa cagcgacaag	1620
ccctccctca aggacattaa aatcatgggt tatagtatcc ggactattga ctacaggtat	1680
accgtgtggg tgggtttcaa ccagacgaa tttctcgcca atttctccga catccacgcg	1740
ggcgaactgt atttcggtga ttccgatcca ctgcaagatc ataatatgta caacgatagt	1800
caagggggtg acctcttcca gttgctaata ccatga	1836

<210> 146

<211> 1836

10

<212> DNA
<213> Secuencia Artificial

<220>

5 <223> Polinucleótido que codifica la proteína de fusión p97

<400> 146

atggaatgga gctgggtctt tctcttcttc ctgtcagtaa cgactggtgt ccactccgac	60
tacaaggacg acgacgacaa agagcagaag ctgatctccg aagaggacct gcaccacat	120
catcaccatc accaccatca cggaggcggt ggagagaacc tgtactttca gggctcggaa	180
actcaggcca actccaccac agatgcactc aacgtgctgc tgatcatcgt agatgacctc	240
cgaccttctc tgggctgtta cggcgacaag ctagtacgga gcccaaacat cgaccagctc	300
gcatcgact ctctcctatt ccagaacgca ttcgccagc aggtgtctg tgctccctcc	360
cgagtgtcct tcctcacggg tcggagaccc gataccacga ggttatatga cttcaactca	420
tactggcgcg tgcattgccc taacttttct actatacccc agtattttta agaaaatggc	480
tatgttaciaa tgcctgttgg caaggtatct catcctggta ttagcagcaa ccacacagat	540
gactctccgt atagctggtc attccacca taccaccct ccagcgaaaa gtacgaaaac	600
acaaagactt gccggggccc agatggcgaa ctgcacgcaa atctgctgtg ccctgtagat	660
gtcttgacg tgcctgaagg tactctgccc gacaaacagt ccacagaaca ggcaatccaa	720
ctccttgaaa agatgaaaac gagcgctcc cccttcttcc tcgacctggg ctaccacaag	780
ccccacatcc cgtttagata cccaaggaa ttccagaaac tgtacccctt ggaaaacatc	840
actctcgcg ccgaccccg agtgccagac ggactccctc ctgttgccca caacccttg	900
atggacatca gacaacgtga agatgtgcag gccctgaaca tctcagtgcc ttacggcccc	960
attccagttg acttccagag gaagattcgg cagtccact tcgctccgt tagttacctg	1020
gacacccaag tgggtagact cctgagcgcc ttggacgatc tccagctcgc aaacagcacc	1080
atcattgcct tcaccagcga ccatggttgg gcgctgggtg aacatggaga atgggctaaa	1140
tattcaaatt tcgacgttgc gacccacgtc ccattgatct tctacgtgcc tggacgaaca	1200
gcctccttgc ctgaagccgg ggaaaagtgt tttccatatc tggacccttt cgattctgcg	1260
agccaactca tggaaacctg gcgacagagc atggacctgg tggaaactgg cagtttattt	1320
ccaaccctgg caggccttgc aggcctccaa gttccacctc ggtgtcccgt tccctcatc	1380
cacgtcgaac tctgtcgca aggtaaaaac ctctcaagc attttcgttt tcgggacctc	1440
gaagaagacc catacctgcc agggaatcca agggaactga ttgcctacag ccagtacct	1500
agacctagcg acatcccaca gtggaacagc gacaagccct ccctcaagga cattaaaatc	1560
atgggttata gtatccggac tattgactac aggtataccg tgtgggtggg tttcaaccca	1620
gacgaatttc tcgccaattt ctccgacatc cagcgggcg aactgtattt cgttgattcc	1680
gatccactgc aagatcataa tatgtacaac gatagtcaag ggggtgacct cttccagttg	1740
ctaattgccag aggcgctgc taaagaggct gccgccaag aagccgccc taaggactcc	1800
tctcagcct tcacctgga cgagctgcgg tactaa	1836

<210> 147

<211> 1812

10

<212> DNA
<213> Secuencia Artificial

<220>

5 <223> Polinucleótido que codifica la proteína de fusión p97

<400> 147

atggaatgga gctgggtctt tctcttcttc ctgtcagtaa cgactggtgt ccactccgac	60
tacaaggacg acgacgacaa agagcagaag ctgatctccg aagaggacct gcaccaccat	120
catcaccatc accacccatca cggaggcggt ggagagaacc tgtactttca gggcacagat	180
gcactcaacg tgtgtctgat catcgtagat gacctccgac cttctctggg ctgttacggc	240
gacaagctag tacggagccc aaacatcgac cagctcgcat cgcactctct cctattccag	300
aacgcattcg cccagcaggc tgtctgtgct ccctcccgag tgccttcct cacgggtcgg	360
agacccgata ccacgaggtt atatgacttc aactcatact ggcgcggtgca tgccggtaac	420
ttttctacta taccocagta ttttaaagaa aatggctatg ttacaatgtc cgttggcaag	480
gtatttcatc ctggtattag cagcaaccac acagatgact ctccgtatag ctggtcattc	540
ccaccatacc acccctccag cgaaaagtac gaaaacacaa agacttgccg gggcccagat	600
ggcgaaactgc acgcaaactc gctgtgccct gtagatgtct tggacgtgcc cgaaggtact	660
ctgcccagaca aacagtccac agaacaggca atccaactcc ttgaaaagat gaaaacgagc	720
gcgctccccct tcttcctcgc cgtgggctac cacaagcccc acatcccggt tagatacccc	780
aaggaatttc agaaactgta cccctcgaa aacatcactc tcgcgcccga ccccgaaagt	840
ccagacggac tccctcctgt tgcctacaac ccttggtatg acatcagaca acgtgaagat	900
gtgcaggccc tgaacatctc agtgccttac ggccccattc cagttgactt ccagaggaag	960
attcggcagt cctacttcgc ctccgttagt tacctggaca cccaagtggg tagactcctg	1020
agcgcccttg acgatctcca gctcgcaaac agcaccatca ttgccttcac cagcgacctat	1080
ggttggggcg tgggtgaaca tggagaatgg gctaaatatt caaatttcga cgttgcgacc	1140
cacgtcccat tgatcttcta cgtgcctgga cgaacagcct ccttgcctga agccggggaa	1200
aagttgtttc catatctgga ccctttcgat tctcgagacc aactcatgga acctgggcga	1260
cagagcatgg acctggtgga actggtcagt ttatttccaa ccctggcagg ccttgcaggc	1320
ctccaagttc cacctcgggtg tcccgttccc tcattccacg tcgaactctg tcgcgaaggt	1380
aaaaacctcc tcaagcattt tcgttttcgg gacctogaag aagaccata cctgccaggg	1440
aatccaaggg aactgattgc ctacagccag taccctagac ctagcgacat cccacagtgg	1500
aacagcgaca agccctccct caaggacatt aaaatcatgg gttatagtat ccggactatt	1560
gactacaggt ataccgtgtg ggtgggtttc aaccagacg aatttctcgc caatttctcc	1620
gacatccacg cggggaact gtatttcgtt gattccgatc cactgcaaga tcataatatg	1680
tacaacgata gtcaaggggg tgacctcttc cagttgctaa tgccagaggc cgctgctaaa	1740
gaggctgccg ccaagaagc cgccgctaag gactcctctc acgccttcac cctggacgag	1800
ctgcggtact aa	1812

<210> 148

<211> 13

10

ES 2 762 672 T3

<212> PRT
<213> Homo sapiens

<400> 148
Asp Ser Ser His Ala Phe Thr Leu Asp Glu Leu Arg Tyr
5 1 5 10

<210> 149
<211> 19
<212> PRT
10 <213> Homo sapiens

<400> 149
Met Glu Trp Ser Trp Val Phe Leu Phe Phe Leu Ser Val Thr Thr Gly
1 5 10 15

Val His Ser

15

REIVINDICACIONES

1. Una proteína de fusión de p97 (melanotransferrina) que comprende un polipéptido de iduronato-2-sulfatasa (IDS) fusionado al extremo N de un polipéptido de p97 y un enlazador peptídico opcional (L) entremedio, en donde el polipéptido de p97 consiste en la secuencia de aminoácidos DSSHAFTLDELRL.
- 5 2. Una proteína de fusión de p97 (melanotransferrina) que comprende un polipéptido de iduronato-2-sulfatasa (IDS) fusionado al extremo C de un polipéptido de p97 y un enlazador peptídico opcional (L) entremedio, en donde el polipéptido de p97 consiste en la secuencia de aminoácidos DSSHAFTLDELRL.
3. La proteína de fusión de la reivindicación 1 o la reivindicación 2, en donde el polipéptido de iduronato-2-sulfatasa (IDS) comprende la secuencia de aminoácidos de SEQ ID NO:32 o SEQ ID NO:33.
- 10 4. La proteína de fusión de la reivindicación 1 o la reivindicación 2, en donde el enlazador comprende la secuencia de aminoácidos (EAAAK)₁₋₃.
5. La proteína de fusión de la reivindicación 1 y la reivindicación 2, en donde el enlazador tiene la secuencia de aminoácidos (EAAAK)₃.
- 15 6. La proteína de fusión de la reivindicación 1 y la reivindicación 2, en donde la proteína de fusión comprende una secuencia de péptido señal (SP) en el extremo N.
7. La proteína de fusión de la reivindicación 1 y la reivindicación 2, en donde la proteína de fusión comprende una etiqueta de purificación (TAG).
8. La proteína de fusión de la reivindicación 1 y la reivindicación 2, en donde la proteína de fusión comprende un sitio de proteasa (PS).
- 20 9. Un polinucleótido aislado que codifica una proteína de fusión de una cualquiera de las reivindicaciones 1-8 o un vector que comprende el polinucleótido aislado que está funcionalmente enlazado a uno o más elementos reguladores.
10. El polinucleótido aislado o vector de la reivindicación 9, que se optimiza por codón para su expresión en una célula hospedante.
- 25 11. El polinucleótido aislado o vector de la reivindicación 10, donde la célula hospedante es una célula de mamífero, una célula de insecto, una célula de levadura o una célula bacteriana.
12. Una célula hospedante recombinante que comprende una proteína de fusión de una cualquiera de las reivindicaciones 1-8 o un polinucleótido aislado de una cualquiera de las reivindicaciones 9-11 que está funcionalmente enlazado a uno o más elementos reguladores, o un vector de una cualquiera de las reivindicaciones 9-11.
- 30 13. La célula hospedante recombinante de la reivindicación 12, donde la célula hospedante es una célula de mamífero, una célula de insecto, una célula de levadura o una célula bacteriana.
14. La célula hospedante recombinante de la reivindicación 13, donde la célula de mamífero es una célula de ovario de hámster chino (CHO), una célula HEK-293 o una célula de fibrosarcoma humano HT-1080.
- 35 15. Una composición farmacéutica que comprende un portador farmacéuticamente aceptable y una proteína de fusión de una cualquiera de las reivindicaciones 1-8, donde la composición farmacéutica es estéril y no pirógena.
16. Una composición farmacéutica de la reivindicación 15 para su uso en el tratamiento de una enfermedad de almacenamiento lisosómico en un sujeto que lo necesita.
- 40 17. La composición para su uso según la reivindicación 16, donde la enfermedad de almacenamiento lisosómico es el síndrome de Hunter (MPS II).
18. La composición para su uso según la reivindicación 16 o 17, donde la enfermedad de almacenamiento lisosómico presenta afectación del sistema nervioso central (CNS).
19. La composición para su uso según la reivindicación 16 o 17, donde el sujeto está en riesgo de desarrollar afectación del CNS de la enfermedad de almacenamiento lisosómico.
- 45 20. La composición para su uso según una cualquiera de las reivindicaciones 16-19, donde el sujeto es un humano de sexo masculino.
21. La composición para su uso según una cualquiera de las reivindicaciones 16-20, donde la proteína de fusión o composición farmacéutica se administra mediante infusión intravenosa (IV).

A



B



FIG. 1

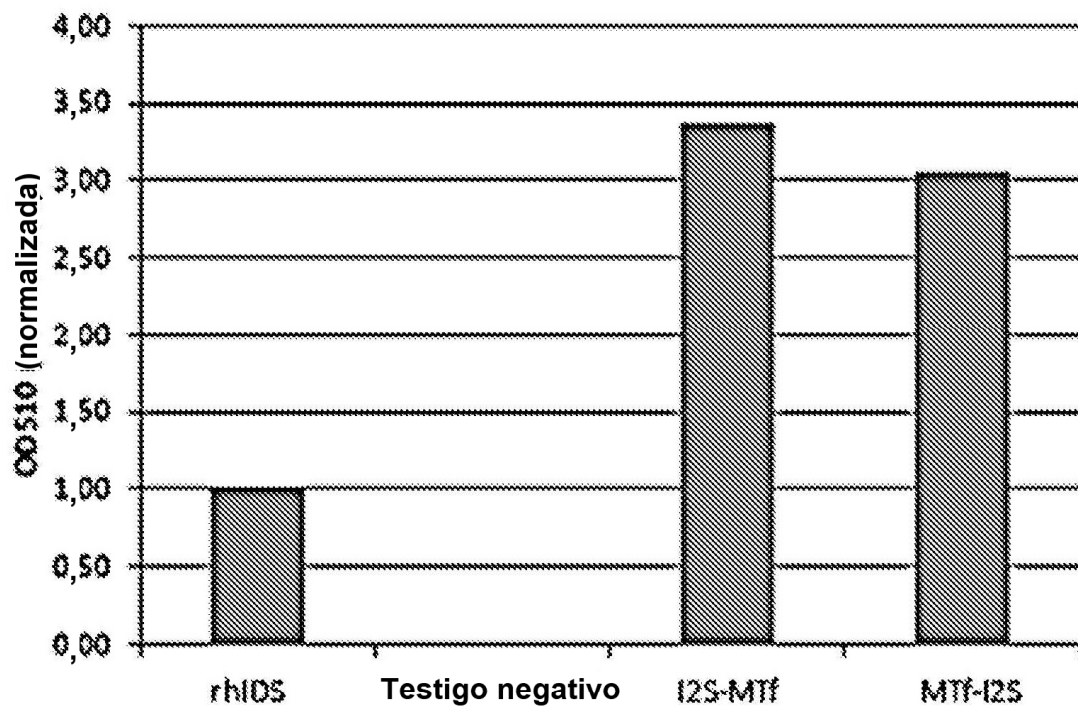


FIG. 2

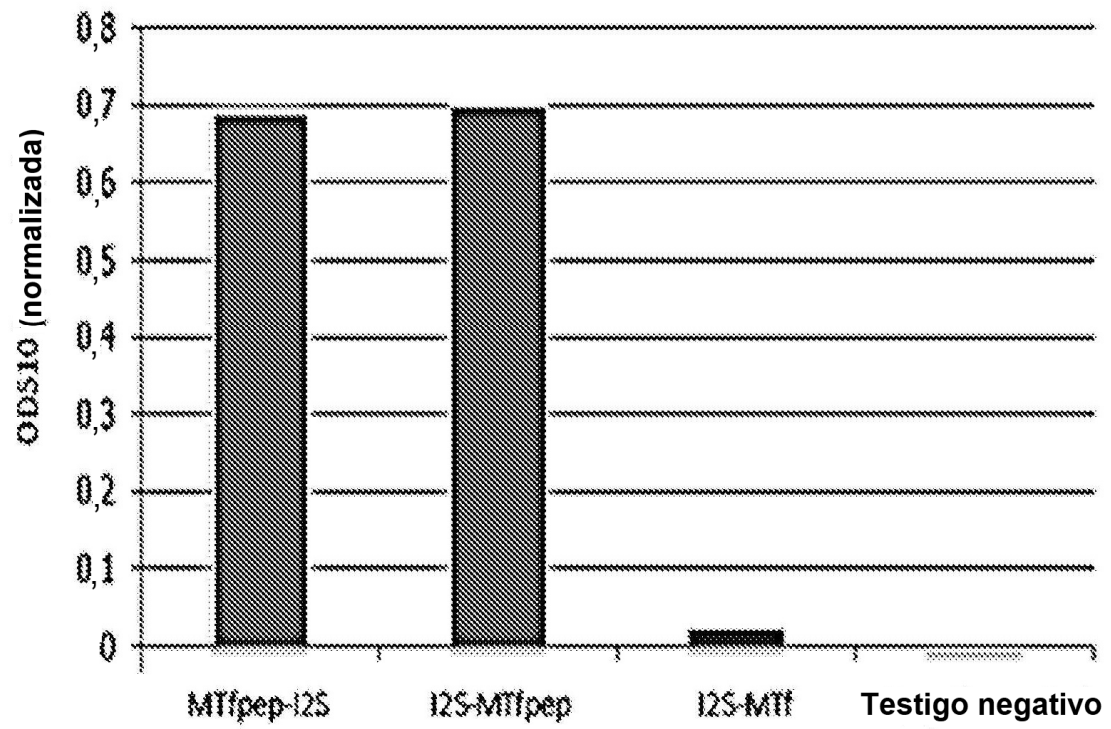


FIG. 3

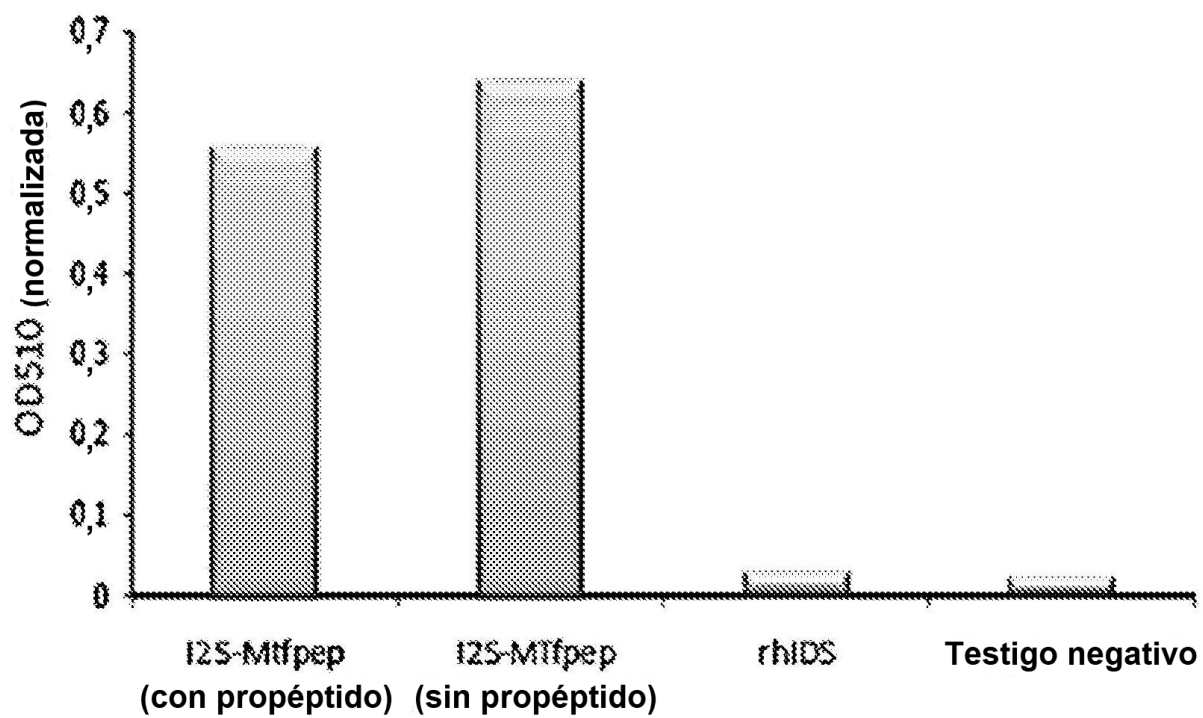
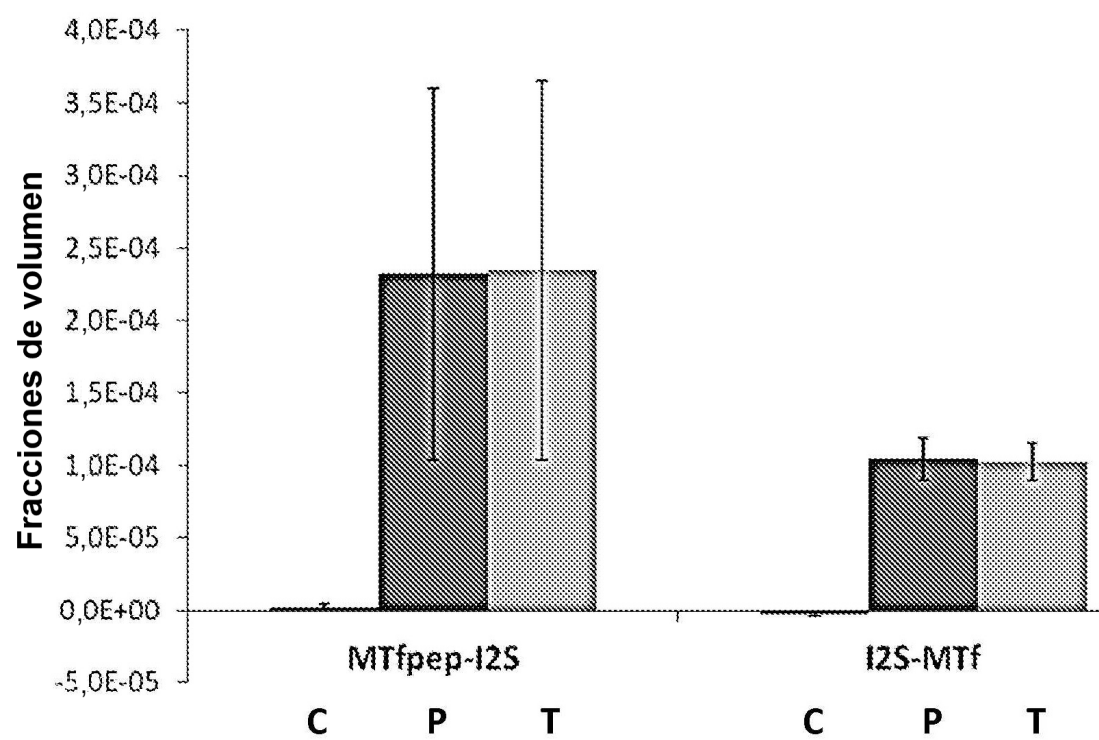


FIG. 4

**FIG. 5**