(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 993 140**

(51) Int. Cl.:

C07K 19/00 (2006.01)

C07K 16/28 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **01.04.2011 E 18159324 (5)**

(97) Fecha y número de publicación de la concesión europea: **24.07.2024 EP 3372617**

(54) Título: **Proteínas de fusión de unión, conjugados de proteína de fusión de unión-fármaco, conjugados de xten-fármaco y métodos de fabricación y uso de los mismos**

(30) Prioridad:

02.04.2010 US 341720 P
08.04.2010 US 341996 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
23.12.2024

(73) Titular/es:

AMUNIX PHARMACEUTICALS, INC. (100.00%)
2 Tower Pl 1100
South San Francisco, CA 94080-1826, US

(72) Inventor/es:

SCHELLENBERGER, VOLKER;
SILVERMAN, JOSHUA;
WANG, CHIA-WEI;
SPINK, BENJAMIN;
STEMMER, WILLEM P. y
ALTERS, SUSAN E.

(74) Agente/Representante:

LEHMANN NOVO, María Isabel

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 unión, conjugados de proteína de fusión de unión-fármaco, conjugados de xten-fármaco y métodos de fabricación y uso de los mismos

REFERENCIA CRUZADA A SOLICITUDES RELACIONADAS

Esta solicitud reivindica el beneficio de prioridad de la Solicitud Provisional de EE. UU. con n.º de serie 61/341.720, presentada el 2 de abril de 2010, y la 61/341.996, presentada el 8 de abril de 2010.

DECLARACIÓN RELATIVA A LA INVESTIGACIÓN PATROCINADA POR EL GOBIERNO FEDERAL

Esta invención se realizó con apoyo gubernamental bajo la subvención SBIR 2R44GM079873-02 otorgada por el Instituto Nacional de la Salud (NIH). El gobierno tiene ciertos derechos en la invención.

ANTECEDENTES DE LA INVENCIÓN

Los anticuerpos o inmunoglobulinas son moléculas que reconocen y se unen a antígenos o ligandos afines específicos. Debido a sus especificidades exclusivas, los anticuerpos, particularmente los anticuerpos monoclonales, se han utilizado ampliamente en el diagnóstico y el tratamiento de una variedad de enfermedades humanas.

Los anticuerpos completos comprenden dos cadenas pesadas unidas entre sí por puentes de disulfuro, y dos cadenas ligeras, estando cada cadena ligera unida a una de las cadenas pesadas por un puente de disulfuro. Cada cadena tiene un dominio variable (V_H o V_L) en el extremo amínico y uno o más dominios constantes en el extremo carboxílico; el dominio constante de la cadena ligera está alineado con, y unido por el disulfuro a, el primer dominio constante de la cadena pesada, y el dominio variable de la cadena ligera está alineado con el dominio variable de la cadena pesada. Cada uno de los dominios variables de la cadena pesada y ligera incluye regiones estructurales (FR) y regiones hipervariables y un puente de disulfuro intracatenario. (Véase, por ejemplo, Chothia et al., J. Mol. Biol. 186: 651-663 (1985); Novotny y Haber, Proc. Natl. Acad. Sci. USA 82: 45924596 (1985); Padlar et al., Mol. Immunol., 23 (9): 951-960 (1986); y S. Miller, J. Mol. Biol., 216: 965-973 (1990). Los anticuerpos pueden derivar de anticuerpos naturales o sintetizarse para que incluyan combinaciones de dominios variables de las cadenas pesadas y ligeras para formar un sitio de unión al antígeno. Los tipos de fragmentos de anticuerpos incluyen, por ejemplo, Fab, Fab', $F(ab')_2$, fragmentos Fv, scFv, Fd y Fd'. Sin embargo, la expresión de fragmentos de anticuerpos en hospedadores bacterianos, incluidos fragmentos de anticuerpos de dominio (dAb), fragmentos Fv, fragmentos Fv monocatenarios (scFv), fragmentos Fab, fragmentos Fab'2, y muchas proteínas que no son anticuerpos (tales como dominios FnIII) pueden dar como resultado la formación de cuerpos de inclusión en el citoplasma, lo que aumenta la complejidad y el coste de producción (Kou, G., et al., 2007, Protein Expr Purif. 52, 131; Cao, P., et al. 2006, Appl Microbiol Biotechnol., 73, 151; Chen, L.H et al., 2006, Protein Expr Purif.; 46, 495). Además, se ha averiguado que la estabilidad y/o los rendimientos de producción de los fragmentos scFv o Fab de los anticuerpos naturales producidos en células hospedadoras son insuficientes. Honneger et al., J. Mol. Biol., 309: 687-699 (2001), y la estabilidad de los fragmentos scFv no siempre está correlacionada con el rendimiento de expresión en el periplasma bacteriano (Wear et al., J. Mol. Biol., 305: 989-1010 (2001). Los muchos factores que afectan el rendimiento de la expresión periplásmica y/o a la estabilidad de los scFv aún no se comprenden completamente y pueden ser impredecibles. Además, los procedimientos para extraer proteínas periplásmicas no son tan robustos como la extracción desde el citoplasma, lo que contribuye a bajos rendimientos. Por lo tanto, debido a que sigue siendo necesario mejorar el proceso de producción de anticuerpos y fragmentos de anticuerpos, particularmente en forma soluble, es deseable encontrar proteínas alternativas que puedan unirse a antígenos y que puedan producirse con unos rendimientos mejorados en el cultivo celular, especialmente en un cultivo celular bacteriano. El documento WO2009/023270 divulga secuencias XTEN (que consisten en los aminoácidos glicina (G), aspartato (D), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P)) para ser fusionadas con diferentes proteínas, tales como anticuerpos que incluyen scFv.

SUMARIO DE LA INVENCIÓN

La presente invención se refiere generalmente a proteínas de fusión de unión novedosas y seleccionables, útiles como agentes para el tratamiento de cualquier enfermedad o afección que se mejore, atenúe o inhiba mediante la administración de proteínas que se unen a ciertas proteínas, hidratos de carbono o glicoproteínas diana asociados con la enfermedad, el trastorno o la afección. En particular, la presente invención proporciona proteínas de fusión de unión que comprenden polipéptidos recombinantes expandidos con una secuencia no repetitiva y una conformación no estructurada (XTEN) unidos a uno o más restos de direccionamiento polipeptídicos que presentan afinidad de unión por ciertas dianas. Las proteínas de fusión de unión de las realizaciones divulgadas en el presente documento presentan una o más o cualquier combinación de las propiedades y/o las realizaciones como se detalla en el presente documento.

En un primer aspecto, la invención proporciona una proteína de fusión de unión aislada que comprende un primer resto de direccionamiento con afinidad de unión específica a una primera diana y un primer polipéptido recombinante expandido (XTEN), en donde el XTEN comprende una secuencia de aminoácidos que tiene al menos un 90 % de

identidad de secuencia con una secuencia seleccionada del grupo que consiste en la SEQ ID NO: 41, la SEQ ID NO: 47, la SEQ ID NO: 54 y la SEQ ID NO: 59, en donde el XTEN comprende un motivo seleccionado del grupo que consiste en la SEQ ID NO: 6, la SEQ ID NO: 7, la SEQ ID NO: 8 y la SEQ ID NO: 9, en donde el primer resto de direccionamiento comprende un fragmento variable monocatenario (scFv) que se une a CD3.

En el presente documento se describen proteínas de fusión de unión aisladas que comprenden un polipéptido recombinante expandido (XTEN) unido a un resto de direccionamiento con afinidad de unión por una diana seleccionada de la tabla 1 o la tabla 2, en donde la proteína de fusión presenta una semivida terminal que es mayor de aproximadamente 48 h, o aproximadamente 72 h, o aproximadamente 96 h, o aproximadamente 120 h, o aproximadamente 10 días, o aproximadamente 21 días, o aproximadamente 30 días cuando se administra a un sujeto. En una realización de lo anterior, el XTEN se caracteriza por que la secuencia comprende al menos aproximadamente 144, o al menos aproximadamente 288, o al menos aproximadamente 576, o al menos aproximadamente 864, o al menos aproximadamente 1000, o al menos aproximadamente 1400, o al menos aproximadamente 2000, hasta aproximadamente 3000 restos de aminoácidos, constituyendo la suma de restos de glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) más de aproximadamente un 80 %, o aproximadamente un 85 %, o aproximadamente un 90 %, o aproximadamente un 95 %, o aproximadamente un 96 %, o aproximadamente un 97 %, o aproximadamente un 98 %, o aproximadamente un 99 %, o aproximadamente un 100 %, de la secuencia de aminoácidos total de XTEN, la secuencia del XTEN es sustancialmente no repetitiva en que (i) la secuencia del XTEN no contiene tres aminoácidos contiguos que sean idénticos a menos que los aminoácidos sean serina, (ii) al menos aproximadamente un 80 % de la secuencia del XTEN consiste en motivos de secuencia no solapados, comprendiendo cada uno de los motivos de secuencia de aproximadamente 9 a aproximadamente 14 restos de aminoácidos, en donde dos restos de aminoácidos contiguos no aparecen más de dos veces en cada uno de los motivos de secuencia; o (iii) la secuencia del XTEN tiene una puntuación de subsecuencia promedio menor de 3, la secuencia del XTEN carece de un epítopo de linfocitos T predicho cuando se analiza mediante el algoritmo TEPITOPE, en donde la predicción del algoritmo TEPITOPE para epítopos dentro de la secuencia del XTEN se basa en una puntuación de -6, -7, o -8, o -9, o -10, la secuencia del XTEN tiene una formación de espirales aleatorias superior al 90 %, o aproximadamente un 95 %, o aproximadamente un 96 %, o aproximadamente un 97 %, o aproximadamente un 98 %, o aproximadamente un 99 % de formación de espirales aleatorias según se determina mediante el algoritmo de GOR; y la secuencia del XTEN tiene menos de un 2 % de hélices alfa y un 2 % de láminas beta según se determina mediante el algoritmo Chou-Fasman. En otra realización de lo anterior, el XTEN se caracteriza además por que la suma de los restos de asparagina y glutamina es inferior al 10 % de la secuencia de aminoácidos total del XTEN, la suma de los restos de metionina y triptófano es inferior al 2 % de la secuencia de aminoácidos total del XTEN

En otra realización, el XTEN de la proteína de fusión de unión se caracteriza además por que ningún tipo de aminoácido constituye más de aproximadamente un 16 % o un 24 %, o aproximadamente un 30 %, de la secuencia del XTEN. En otra realización, el XTEN de la proteína de fusión de unión se caracteriza además por que al menos aproximadamente un 80 %, o al menos aproximadamente un 90 %, o al menos aproximadamente un 91 %, o al menos aproximadamente un 92 %, o al menos aproximadamente un 93 %, o al menos aproximadamente un 94 %, o al menos aproximadamente un 95 %, o al menos aproximadamente un 96 %, o al menos aproximadamente un 97 %, o al menos aproximadamente un 98 %, o al menos aproximadamente un 99 % de la secuencia del XTEN consiste en motivos de secuencia no solapados, en donde cada uno de los motivos de secuencia tiene aproximadamente 12 restos de aminoácidos y en donde la secuencia de dos restos de aminoácidos contiguos no aparece más de dos veces en cada uno de los motivos de secuencia, y los motivos de secuencia consisten en entre cuatro y seis tipos de aminoácidos seleccionados entre glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P). En una realización, más del 90 % de la secuencia del XTEN consiste en motivos de secuencia no solapados, en donde los motivos de secuencia son de una o más secuencias de la tabla 3. Descrita en el presente documento, la secuencia del XTEN presenta al menos aproximadamente un 80 %, o al menos aproximadamente un 90 %, o al menos aproximadamente un 91 %, o al menos aproximadamente un 92 %, o al menos aproximadamente un 93 %, o al menos aproximadamente un 94 %, o al menos aproximadamente un 95 %, o al menos aproximadamente un 96 %, o al menos aproximadamente un 97 %, o al menos aproximadamente un 98 %, o al menos aproximadamente un 99 % o un 100 % de identidad de secuencia con una secuencia seleccionada de una cualquiera de la tabla 4, la tabla 11, la tabla 12, la tabla 13, la tabla 14 o la tabla 15, cuando se alinean de manera óptima. Descrita en el presente documento, la invención proporciona una proteína de fusión de unión aislada que comprende una secuencia que tiene al menos aproximadamente un 80 %, o al menos aproximadamente un 90 %, o al menos aproximadamente un 91 %, o al menos aproximadamente un 92 %, o al menos aproximadamente un 93 %, o al menos aproximadamente un 94 %, o al menos aproximadamente un 95 %, o al menos aproximadamente un 96 %, o al menos aproximadamente un 97 %, o al menos aproximadamente un 98 %, o al menos aproximadamente un 99 % o un 100 % de identidad de secuencia con una secuencia seleccionada de una cualquiera de la tabla 25, la tabla 40 o la tabla 41.

Las presentes proteínas de fusión de unión presentan unas propiedades farmacocinéticas mejoradas. En una realización, la propiedad farmacocinética mejorada es una semivida terminal que es superior a aproximadamente 24 h, o superior a aproximadamente 48 h, o superior a aproximadamente 72 h, o superior a aproximadamente 96 h, o superior a aproximadamente 120 h, o superior a aproximadamente 144 h, o superior a aproximadamente 7 días, o superior a aproximadamente 10 días, o superior a aproximadamente 14 días, o superior a aproximadamente 21 días, cuando se administra a un sujeto, en donde las propiedades farmacocinéticas se determinan midiendo las concentraciones en sangre de la proteína de fusión con el tiempo después de la administración de una dosis a un

sujeto. En una realización, la propiedad farmacocinética mejorada engloba un aumento en la semivida terminal de al menos aproximadamente dos veces, o al menos aproximadamente tres veces, o al menos aproximadamente cuatro veces, o al menos aproximadamente cinco veces, o al menos aproximadamente seis veces, o al menos aproximadamente ocho veces, o al menos aproximadamente diez veces, o al menos aproximadamente 20 veces en comparación con el resto de direccionamiento no unido al XTEN y administrado a un sujeto a una dosis comparable.

Descrito en el presente documento, el resto de direccionamiento de la proteína de fusión de unión aislada se selecciona del grupo que consiste en anticuerpo, fragmento de anticuerpo, scFv, diacuerpo, anticuerpo de dominio, receptor de citocina y receptor de la superfamilia de las inmunoglobulinas. El primer resto de direccionamiento es un scFv. Descrito en el presente documento, el resto de direccionamiento es un scFv con afinidad de unión por Her2. En algunas realizaciones, la proteína de fusión de unión es multivalente, comprendiendo dos, o tres, o cuatro, o cinco, o seis, siete u ocho restos de direccionamiento. En una realización de lo anterior, el resto de direccionamiento multivalente es un scFv. En una realización, los restos de direccionamiento multivalentes pueden presentar una afinidad de unión específica a la misma diana, en donde las dianas se seleccionan de la tabla 1 o la tabla 2. En una realización, los restos de direccionamiento multivalentes pueden presentar una afinidad de unión específica a dos o más dianas, en donde las dianas se seleccionan de la tabla 1 o la tabla 2. En cualquiera de las realizaciones descritas en este párrafo, la constante de afinidad de unión (K_d) para el uno o más restos de direccionamiento de la presente proteína de fusión de unión y un ligando diana es menor de aproximadamente 10^{-4} M, alternativamente menor de aproximadamente 10^{-5} M, alternativamente menor de aproximadamente 10^{-6} M, alternativamente menor de aproximadamente 10^{-7} M, alternativamente menor de aproximadamente 10^{-8} M, alternativamente menor de aproximadamente 10^{-9} M, o menor de aproximadamente 10^{-10} M, o menor de aproximadamente 10^{-11} M o menor de aproximadamente 10^{-12} M.

En una realización, la proteína de fusión de unión puede comprender un segundo XTEN que tiene al menos aproximadamente 48 restos de aminoácidos unidos al extremo amínico de la proteína de fusión de unión, en donde la expresión de la proteína de fusión de unión en una célula hospedadora que comprende un vector de expresión que codifica la proteína de fusión de unión se está mejorada en comparación con la expresión en una célula hospedadora que comprende un vector de expresión que codifica una proteína de fusión de unión correspondiente que carece del segundo XTEN. En una realización de lo anterior, el segundo XTEN presenta al menos un 90 % de identidad de secuencia, o al menos aproximadamente un 91 %, o al menos aproximadamente un 92 %, o al menos aproximadamente un 93 %, o al menos aproximadamente un 94 %, o al menos aproximadamente un 95 %, o al menos aproximadamente un 96 %, o al menos aproximadamente un 97 %, o al menos aproximadamente un 98 %, o al menos aproximadamente un 99 % o un 100 % de identidad de secuencia con una secuencia del XTEN seleccionada entre AE48, AM 48, AE624, AE912 o AM923. En otra realización de lo anterior, la incorporación en una célula hospedadora de un polinucleótido que codifica una proteína de fusión de unión que comprende el XTEN aminoterminal puede dar como resultado un nivel de expresión que está mejorado en al menos un 50 %, o al menos aproximadamente un 75 %, o al menos aproximadamente un 100 %, o al menos aproximadamente un 150 %, o al menos aproximadamente un 200 % o más en comparación con los niveles de expresión en una célula hospedadora comparable con un polinucleótido que codifica una proteína de fusión de unión sin el XTEN aminoterminal.

La invención proporciona proteínas de fusión de unión en diferentes configuraciones. En el presente documento se describe una proteína de fusión de unión aislada de fórmula I:

$$(XTEN)_x\text{-}TM\text{-}(XTEN)_y \qquad I$$

en donde independientemente en cada aparición, XTEN es un polipéptido recombinante expandido que comprende más de aproximadamente 36 a aproximadamente 3000 aminoácidos con una secuencia sustancialmente no repetitiva en donde la suma de los restos de glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) constituyen más de aproximadamente un 80 % de la secuencia de aminoácidos total del XTEN, x es 0 o 1, y es 0 o 1, en donde $x + y \geq 1$, y TM es un resto de direccionamiento con afinidad de unión específica a una diana seleccionada de la tabla 1 o la tabla 2.

En el presente documento se describe una proteína de fusión de unión aislada de fórmula II:

$$(XTEN)_x\text{-}TM1\text{-}L\text{-}TM2\text{-}(XTEN)_y \qquad II$$

en donde independientemente en cada aparición, XTEN es un polipéptido recombinante expandido que comprende más de aproximadamente 36 a aproximadamente 3000 aminoácidos con una secuencia sustancialmente no repetitiva en donde la suma de los restos de glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) constituyen más de aproximadamente un 80 % de la secuencia de aminoácidos total del XTEN, x es 0 o 1, y es 0 o 1, en donde $x + y \geq 1$, TM1 es un resto de direccionamiento con afinidad de unión específica a una diana seleccionada de la tabla 1, TM2 es un resto de direccionamiento con afinidad de unión por una diana seleccionada de la tabla 1 o la tabla 2 que puede ser idéntico o puede ser diferente a TM1, y L es una secuencia conectora que tiene entre 1 y aproximadamente 300 restos de aminoácidos. En una realización, el conector puede ser una secuencia en la que al menos un 80 % de los restos están comprendidos por los aminoácidos glicina, serina y/o glutamato, tales como, pero sin limitación, una secuencia con aproximadamente un 80-100 % de identidad de secuencia con la secuencia GSGEGSEGEGGGEGSEGEGSGEGGEGEGSG (SEQ ID NO: 1) o una porción o un multímero de la misma.

En el presente documento se describe una proteína de fusión de unión aislada de fórmula III

TM1-XTEN-TM2 III

en donde independientemente en cada aparición, XTEN es un polipéptido recombinante expandido que comprende más de aproximadamente 400 a aproximadamente 3000 aminoácidos con una secuencia sustancialmente no repetitiva en donde la suma de los restos de glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) constituyen más de aproximadamente un 80 % de la secuencia de aminoácidos total del XTEN, x es 0 o 1, y es 0 o 1, en donde x + y ≥1, TM1 es un resto de direccionamiento con afinidad de unión específica a un ligando diana seleccionado de la tabla 1 o la tabla 2; y TM2 es un resto de direccionamiento con afinidad de unión por una diana seleccionada de la tabla 1 o la tabla 2 que puede ser idéntico o puede ser diferente a la diana unida por TM1,

En el presente documento se describe una proteína de fusión de unión multivalente con tres restos de unión de fórmula IV:

$(XTEN)_x$-TM1-L1-TM2-L2-TM3-$(XTEN)_y$ IV

en donde independientemente en cada aparición: XTEN es un polipéptido recombinante expandido como se ha descrito anteriormente, x es 0 o 1; y es 0 o 1, en donde x+y ≥1; TM1 es un resto de direccionamiento con afinidad de unión por un ligando diana seleccionado de la tabla 1 o la tabla 2; TM2 es un resto de direccionamiento con afinidad de unión por el ligando diana seleccionado de la tabla 1 o la tabla 2 que puede ser idéntico o puede ser diferente a TM1; TM3 es un resto de direccionamiento con afinidad de unión por el ligando diana seleccionado de la tabla 1 o la tabla 2 que puede ser idéntico o puede ser diferente a TM1 o TM2; L1 es una secuencia conectora que tiene entre 1 y aproximadamente 300 restos de aminoácidos como se ha descrito para la fórmula II, y en donde la secuencia conectora está unida covalentemente al extremo carboxílico de TM1 y al extremo amínico de TM2; y L2 es una secuencia conectora que puede ser idéntica o diferente a L1, que tiene entre 1 y aproximadamente 300 restos de aminoácidos como se ha descrito para la fórmula II, y en donde la secuencia conectora está unida covalentemente al extremo carboxílico de TM2 y al extremo amínico de TM3.

Descrita en el presente documento, la invención proporciona una proteína de fusión aislada con un único resto de direccionamiento, en donde el resto de direccionamiento presenta afinidad específica de unión a una diana seleccionada de la tabla 1 o la tabla 2. En una realización de lo anterior, la diana se selecciona entre IL17, IL17R, VRS, HER2, IL12, IL23, RANKL, NGF, CD80, CD86, CD3, CD40, EGFR, TNF-alfa, cMET, IL6R y elastasa. En el presente documento se describe una proteína de fusión aislada con múltiples restos de direccionamiento (por ejemplo, dos, o tres, o cuatro, o cinco, o seis, o siete o más restos de direccionamiento) en donde el resto de direccionamiento presenta afinidad de unión a una o más dianas seleccionadas de la tabla 1 o la tabla 2. Descritos en el presente documento, una más dianas se seleccionan entre IL17, IL17R, VRS, HER2, IL12, IL23, RANKL, NGF, CD80, CD86, CD3, CD40, EGFR, TNF-alfa, cMET, IL6R y elastasa. En una realización preferida de lo anterior, los dos o más restos de direccionamiento son scFv. En cualquiera de las realizaciones descritas anteriormente en este párrafo, la constante de afinidad de unión (K_d) para el uno o más restos de direccionamiento de la presente proteína de fusión de unión y un ligando diana es menor de aproximadamente 10^{-5} M, alternativamente menor de aproximadamente 10^{-6} M, alternativamente menor de aproximadamente 10^{-7} M, alternativamente menor de aproximadamente 10^{-8} M, alternativamente menor de aproximadamente 10^{-9} M, o menor de aproximadamente 10^{-10} M, o menor de aproximadamente 10^{-11} M, o menor de aproximadamente 10^{-12} M.

En algunas realizaciones, las proteínas de fusión de unión presentan un peso molecular aparente aumentado según se determina mediante cromatografía de exclusión por tamaño, en comparación con el peso molecular real, en donde el peso molecular aparente es de al menos aproximadamente 100 kD, 150 kD, 200 kD, 300 kD, 400 kD, 500 kD, 600 kD o 700 kD, mientras que el peso molecular real de la proteína de fusión es menor de aproximadamente 25 kD. Por consiguiente, las proteínas de fusión de unión pueden tener un peso molecular aparente que es aproximadamente 4 veces mayor, o aproximadamente 5 veces mayor, o aproximadamente 6 veces mayor, o aproximadamente 7 veces mayor, o aproximadamente 8 veces mayor que el peso molecular real de la proteína de fusión de unión. En una realización, la proteína de fusión de unión aislada de las realizaciones anteriores presenta un factor de peso molecular aparente en condiciones fisiológicas que es mayor de aproximadamente 4, o aproximadamente 5, o aproximadamente 6, o aproximadamente 7 o aproximadamente 8.

En otra realización, la invención proporciona la proteína de fusión de unión aislada de una cualquiera de las realizaciones anteriores, que comprende además una o más moléculas de un fármaco seleccionado de la tabla 9. En una realización, el fármaco está unido covalentemente mediante un agente de conjugación al XTEN, preferentemente a través de uno o más restos de aminoácidos de cisteína o lisina incorporados en el XTEN. En otra realización, el fármaco está unido covalentemente mediante un agente de conjugación al resto de direccionamiento, preferentemente a través de uno o más restos de aminoácidos de cisteína o lisina incorporados en el resto de direccionamiento. Los conjugados de fármaco y de proteína de fusión de unión pueden estar en diferentes configuraciones. En el presente documento se describe una composición de conjugado de proteína de fusión de unión-fármaco de fórmula V:

$$[(D\text{-}CL)_{z1}\text{-}XTEN]_x\text{-}TM\text{-}[XTEN\text{-}(CL\text{-}D)_{z2}]_y \qquad V$$

en donde independientemente en cada aparición: X es 0 o 1; y es 0 o 1; XTEN es un polipéptido recombinante expandido genomodificado con cisteína o lisina como se ha descrito anteriormente; TM es un resto de direccionamiento con afinidad de unión por un ligando diana seleccionado de la tabla 1 o la tabla 2 (que puede comprender más de un dominio de unión unido por conectores); CL es un agente de conjugación como se define en el presente documento; D es un resto de fármaco seleccionado de la tabla 9 o una sal, un ácido o un derivado farmacéuticamente aceptable del mismo; y cada uno de z1 y z2 es independientemente un número entero de 1 a 100. Algunas composiciones de conjugado de proteína de fusión de unión-fármaco de fórmula V ilustrativas pueden comprender XTEN que tienen de 1 a aproximadamente 100 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 50 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 40 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 20 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 10 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 5 aminoácidos genomodificados con cisteína o lisina que están disponibles para su conjugación con moléculas de fármacos.

En el presente documento se describe una composición de conjugado de proteína de fusión de unión-fármaco de fórmula VI:

$$[(D\text{-}CL)_{z1}\text{-}XTEN]_x\text{-}TM1\text{-}L\text{-}TM2\text{-}[XTEN\text{-}(CL\text{-}D)_{z2}]_y \qquad VI$$

en donde independientemente en cada aparición: X es 0 o 1, e y es 0 o 1; XTEN es un polipéptido recombinante expandido genomodificado con cisteína o lisina como se ha descrito; TM1 es un resto de direccionamiento con afinidad de unión por un ligando diana seleccionado de la tabla 1 o la tabla 2 (que puede comprender más de un dominio de unión unido por conectores); TM2 es un resto de direccionamiento con afinidad de unión por un ligando diana seleccionado de la tabla 1 o la tabla 2 (que puede comprender más de un dominio de unión unido por conectores) que puede ser idéntico o puede ser diferente a TM1; y L es una secuencia conectora que tiene entre 1 y aproximadamente 300 restos de aminoácidos, en donde la secuencia conectora está unida covalentemente al extremo carboxílico de TM1 y al extremo amínico de TM2; D es un resto de fármaco seleccionado de la tabla 9 o una sal, un ácido o un derivado farmacéuticamente aceptable del mismo; CL es un agente de conjugación como se define en el presente documento; y cada uno de z1 y z2 es independientemente un número entero de 0 a 100. Algunas composiciones de conjugado de proteína de fusión de unión-fármaco de fórmula VI ilustrativas pueden comprender XTEN que tienen de 1 a aproximadamente 100 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 50 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 40 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 20 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 10 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 5 aminoácidos genomodificados con cisteína o lisina que están disponibles para su conjugación con moléculas de fármacos.

En el presente documento se describen composiciones de las proteínas de fusión de unión aisladas de cualquiera de las realizaciones anteriores. En una realización, la invención proporciona una composición farmacéutica que comprende una proteína de fusión de unión de cualquiera de las realizaciones anteriores y al menos un portador farmacéuticamente aceptable. En el presente documento se describen kits que comprenden material de envasado y al menos un primer recipiente que comprende la composición farmacéutica de la realización anterior y una etiqueta que identifica la composición farmacéutica y las condiciones de almacenamiento y manipulación, y una hoja de instrucciones para la reconstitución y/o la administración de las composiciones farmacéuticas a un sujeto.

La invención proporciona además métodos de uso de las composiciones farmacéuticas que comprenden la proteína de fusión de cualquiera de las realizaciones anteriores en el tratamiento de una enfermedad, un trastorno o una afección en un sujeto que lo necesite, en donde la enfermedad, el trastorno o la afección se selecciona del grupo que consiste en carcinoma de mama, carcinoma de pulmón, carcinoma gástrico, carcinoma de esófago, carcinoma colorrectal, carcinoma de hígado, carcinoma de ovario, tecoma, arrenoblastoma, carcinoma cervicouterino, carcinoma endometrial, endometriosis, fibrosarcoma, coriocarcinoma, cáncer de cabeza y cuello, carcinoma nasofaríngeo, carcinoma laríngeo, hepatoblastoma, sarcoma de Kaposi, melanoma, carcinoma cutáneo, hemangioma, hemangioma cavernoso, hemangioblastoma, carcinoma de páncreas, retinoblastoma, astrocitoma, glioblastoma, schwannoma, oligodendroglioma, meduloblastoma, neuroblastoma, rabdomiosarcoma, sarcoma osteógeno, leiomiosarcomas, carcinoma de las vías urinarias, carcinoma de tiroides, tumor de Wilm, carcinoma de células renales y carcinoma de próstata.

En el presente documento se describe un método para tratar una enfermedad o una afección mediada por o asociada con una diana de la tabla 1 o la tabla 2, que comprende administrar la composición farmacéutica descrita anteriormente utilizando una cantidad terapéuticamente eficaz a un sujeto que la necesite. En una realización, la diana de la tabla 1 es un antígeno asociado al tumor, tal como, pero sin limitación, los antígenos de la tabla 2. En el presente documento se describe un método de tratamiento en donde la proteína de fusión de unión presenta unión al antígeno asociado al tumor HER2. En una realización del método, la administración de la composición farmacéutica utilizando una cantidad terapéuticamente eficaz a un sujeto tiene un efecto inhibidor del crecimiento en una célula tumoral. El método de tratamiento puede comprender administrar la composición farmacéutica por una vía apropiada, que incluye por vía subcutánea, intramuscular, intravítrea o intravenosa. En una realización, se administran múltiples dosis consecutivas

de la composición farmacéutica con una pauta posológica terapéuticamente eficaz, y pueden dar como resultado una mejora en al menos un parámetro medido relevante para la enfermedad, el trastorno o la afección metabólica. En una realización, la pauta posológica terapéuticamente eficaz se puede conseguir utilizando dos administraciones de composición farmacéutica por pauta posológica mensual durante el periodo de administración. En una realización, la pauta posológica terapéuticamente eficaz se logra utilizando una administración de una composición farmacéutica por pauta posológica mensual durante el periodo de administración. En otra realización, la proteína de fusión de unión tiene un efecto inhibidor del crecimiento en células SK-BR-3 en un ensayo de cultivo celular.

La invención proporciona ácidos nucleicos aislados que comprenden una secuencia polinucleotídica seleccionada de (a) un polinucleótido que codifica la proteína de fusión de unión de cualquiera de las realizaciones anteriores, o (b) el complemento del polinucleótido de (a). Descrito en el presente documento, el ácido nucleico aislado comprende una secuencia polinucleotídica que tiene al menos un 80 % de identidad de secuencia, o aproximadamente un 85 %, o al menos aproximadamente un 90 %, o aproximadamente un 91 %, o aproximadamente un 92 %, o aproximadamente un 93 %, o aproximadamente un 94 %, o aproximadamente un 95 %, o aproximadamente un 96 %, o aproximadamente un 97 %, o aproximadamente un 98 %, o aproximadamente un 99 % a aproximadamente un 100 % de identidad de secuencia con (a) una secuencia polinucleotídica que codifica un polipéptido seleccionado de una cualquiera de la tabla 25, la tabla 40 o la tabla 41; o (b) el complemento del polinucleótido de (a). La invención proporciona vectores de expresión que comprenden el ácido nucleico de cualquiera de las realizaciones descritas anteriormente en este párrafo. En una realización, el vector de expresión de lo anterior comprende además una secuencia reguladora recombinante unida operativamente a la secuencia polinucleotídica. En otra realización, la secuencia polinucleotídica de los vectores de expresión de lo anterior se fusiona en marco con un polinucleótido que codifica una secuencia de señalización de secreción, que puede ser una secuencia de señalización procariota. En una realización, la secuencia de señalización de secreción se selecciona entre las secuencias de señalización OmpA, DsbA y PhoA.

La invención proporciona una célula hospedadora, que puede comprender un vector de expresión divulgado en el párrafo anterior. En una realización, la célula hospedadora es una célula procariota, tal como, pero sin limitación, *E. coli*. En otra realización, la célula hospedadora es una célula eucariota, tal como, pero sin limitación, CHO.

La invención también proporciona células hospedadoras que comprenden un vector de expresión, en donde el vector de expresión codifica una proteína de fusión de unión que comprende un XTEN aminoterminal optimizado para su expresión. Descrito en el presente documento, el vector comprende una secuencia que codifica una secuencia polipeptídica que presenta al menos aproximadamente un 80 %, más preferentemente al menos aproximadamente un 90 %, más preferentemente al menos aproximadamente un 91 %, más preferentemente al menos aproximadamente un 92 %, más preferentemente al menos aproximadamente un 93 %, más preferentemente al menos aproximadamente un 94 %, más preferentemente al menos aproximadamente un 95 %, más preferentemente al menos aproximadamente un 96 %, más preferentemente al menos aproximadamente un 97 %, más preferentemente al menos aproximadamente un 98 %, más preferentemente al menos el 99 %, o presenta una identidad de secuencia del 100 % con la secuencia de aminoácidos de AE48. AM48, AE624, AE912 o AM923. En una realización, el nivel de expresión de la proteína de fusión de unión codificada en la célula hospedadora está mejorado en comparación con el nivel de expresión en una célula hospedadora correspondiente que comprende un vector de expresión que codifica una proteína de fusión de unión que carece del XTEN aminoterminal optimizado para su expresión. En una realización, el nivel de expresión está mejorado en al menos aproximadamente un 50 %, o aproximadamente un 75 %, o aproximadamente un 100 %, o aproximadamente un 150 %, o aproximadamente un 200 % o aproximadamente un 400 % en comparación con una proteína de fusión de unión correspondiente que no comprende la secuencia del XTEN aminoterminal.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

Las características y ventajas de la invención pueden explicarse adicionalmente mediante referencia a la siguiente descripción detallada y dibujos adjuntos, que exponen realizaciones ilustrativas.

La FIG. 1 muestra representaciones esquemáticas de una proteína de fusión de unión con scFv ilustrativa con un único resto de direccionamiento representado en una orientación de amino a carboxiterminal. La FIG. 1A muestra la configuración de una proteína de fusión de unión con scFv (100) con el resto de direccionamiento Fv monocatenario en una conformación "relajada", que comprende una secuencia del XTEN aminoterminal (101), una secuencia del dominio de unión VL (102), una secuencia conectora (103), una secuencia del dominio de unión VH (104) y una secuencia portadora del XTEN (105) en el extremo carboxílico. La FIG. 1B muestra el conector flexible que permite que los dos dominios de unión entren en asociación para formar el sitio de unión al antígeno del resto de direccionamiento scFv.

La FIG. 2 muestra representaciones esquemáticas de componentes ilustrativos de una proteína de fusión de unión con scFv con dos restos de direccionamiento representados en una orientación amino a carboxiterminal. La FIG. 2A muestra la configuración de una proteína de fusión de unión con scFv (107) con dos restos de direccionamiento Fv monocatenarios en una conformación "relajada", comprendiendo cada uno una secuencia del XTEN aminoterminal (101), una secuencia del dominio de unión VL (102), una secuencia

conectora (103), una secuencia del dominio de unión VH (104), una segunda secuencia conectora que une el primer y el segundo resto de direccionamiento (106) y una secuencia portadora del XTEN (105) en el extremo carboxílico. La FIG. 2B muestra los conectores flexibles que permiten que los dos dominios de unión entren en asociación para formar el sitio de unión al antígeno de los dos restos de direccionamiento scFv respectivos.

La FIG. 3 muestra representaciones esquemáticas de una proteína de fusión de unión de diacuerpo monómera ilustrativa que puede formar dos sitios de unión al antígeno. La FIG. 3A muestra la configuración de una proteína de fusión de unión de diacuerpo (108) con los dos restos de direccionamiento en una conformación "relajada", comprendiendo cada uno una secuencia del XTEN aminoterminal (101), una primera secuencia del dominio de unión VL (102), una secuencia conectora corta que une las VL y VH adyacentes (103), una primera secuencia del dominio de unión VH (104), una segunda secuencia conectora que une el primer y el segundo resto de direccionamiento (106), un segundo dominio de unión VH (109) previsto para emparejarse con el primer dominio de unión VL (104), un segundo VL (110) previsto para emparejarse con el primer VH (104) y una secuencia portadora del XTEN (105) en el extremo carboxílico. La FIG. 3B muestra el conector central flexible que permite que dos dominios de unión (VL y VH), desde los extremos opuestos de la molécula (en lugar de los dos dominios de unión adyacentes que están restringidos debido al conector corto), entren en asociación para formar los dos sitios de unión al antígeno, dando como resultado dos restos de direccionamiento de diacuerpo dentro de la proteína de fusión de unión monómera.

La FIG. 4 muestra representaciones esquemáticas de una proteína de fusión de unión a citocinas VEGF ilustrativa. La FIG. 4A muestra la configuración de la proteína de fusión de unión (114) con las regiones de unión de dímero en una conformación "relajada", cada una de las cuales comprende una secuencia del XTEN aminoterminal (101), una primera secuencia del dominio D2 (111), una secuencia conectora corta que une las unidades de dominio adyacentes (103), una primera secuencia del dominio D3 (112), una segunda secuencia conectora flexible que une el primer y el segundo dominios de unión (106), un segundo dominio de unión VH (109) previsto para emparejarse con el primer dominio de unión VL (104), un segundo dominio D2 (111), una secuencia conectora corta que une las unidades de dominio adyacentes (103), una segunda secuencia del dominio D3 (112) y una secuencia portadora del XTEN (105) en el extremo carboxílico. En el esquema, la diana dimérica VEGF (113) ha comenzado a asociarse con uno de los dominios de unión. La FIG. 4B muestra el conector central flexible que permite que dos dominios de unión de la molécula rodeen el VEGF dimérico (113) para formar el complejo de unión que secuestra la molécula del VEGF.

La FIG. 5 muestra representaciones esquemáticas de una proteína de fusión de unión al receptor ilustrativa que comprende dominios de unión similares a Ig. La proteína de fusión de unión (115) tiene una secuencia del XTEN aminoterminal (101) y en este caso, dos regiones de unión similares a Ig (116), una secuencia conectora flexible que une las regiones de unión similares a Ig adyacentes (106) que permitiría a las dos regiones de unión secuestrar el ligando diana, y una secuencia portadora del XTEN (105) en el extremo carboxílico.

La FIG. 6 muestra representaciones esquemáticas de genes ilustrativos que codifican proteínas de fusión de unión, todos representados en una orientación de 5' a 3', con todas las secuencias componentes conectadas en marco. La FIG. 6A muestra la configuración de un gen que codifica proteínas de fusión de unión con scFv (100) en dos configuraciones (de 5' a 3') con un único resto de direccionamiento, teniendo el primero nucleótidos que codifican una secuencia del dominio de unión VL (202), una secuencia conectora (203), una secuencia del dominio de unión VH (204) y una secuencia portadora del XTEN (205) en el extremo 3'. Debajo de eso está la configuración opuesta, con nucleótidos que codifican un XTEN aminoterminal (201), unidos a nucleótidos que codifican una secuencia del dominio de unión VL (202), una secuencia conectora (203) y una secuencia del dominio de unión VH La FIG. 6B muestra la configuración de un gen que codifica una proteína de fusión de unión de diacuerpo que puede formar dos sitios de unión al antígeno (208), y la construcción superior mostrada comprende, en el extremo 5', genes para una primera secuencia del dominio de unión VL (202), una secuencia conectora corta que une los genes VL y VH adyacentes (303), una primera secuencia del dominio de unión VH (204), una segunda secuencia conectora (que puede ser una XTEN) que une el primer y el segundo resto de direccionamiento (206), un segundo dominio de unión VH (209), un segundo dominio de unión VL (210) y una secuencia portadora del XTEN (205) en el extremo 3'. El gen inferior codifica, en el extremo 5', un XTEN aminoterminal (201), una primera secuencia del dominio de unión VL (202), una secuencia conectora corta que une los genes VL y VH adyacentes (203), una primera secuencia del dominio de unión VH (204), una segunda secuencia conectora (que puede ser un XTEN) que une el primer y el segundo resto de direccionamiento (206), un segundo dominio de unión VH (209), un segundo dominio de unión VL (210) y una secuencia portadora del XTEN (205) en el extremo 3'. La FIG. 6C muestra la configuración de un gen que codifica una proteína de unión a citocinas. El gen codifica, en el extremo 5', un XTEN aminoterminal (201), una primera secuencia del dominio de unión D2 (211), una secuencia conectora corta que une los genes del dominio adyacente (203), una primera secuencia del dominio de unión D3 (212), una segunda secuencia conectora (206) (que puede ser un XTEN) que une el primer y el segundo resto de direccionamiento, un segundo dominio de unión D2 (211), un segundo dominio de unión D3 (212) y una secuencia portadora del XTEN (205) en el extremo 3'. La FIG. 6D muestra tres configuraciones, de 5' a 3', de

genes que codifican una proteína de fusión de unión a anticuerpos de dominio (217). Los genes codifican, donde se muestre, un XTEN aminoterminal (201), la secuencia del dominio de unión VHH (218) y una secuencia portadora del XTEN (205) en el extremo 3'.

La FIG. 7 es un diagrama de flujo esquemático de etapas representativas en el ensamblaje, la producción y la evaluación de un XTEN.

La FIG. 8 es un diagrama de flujo esquemático de etapas representativas en el ensamblaje de una construcción de polinucleótido XTEN-resto de direccionamiento que codifica, en este caso, una proteína de fusión de unión anti-Her2. Los oligonucleótidos individuales **501** se hibridan en los motivos de secuencia **502**, tales como un motivo de 12 aminoácidos ("12-mero"), que se ligan a motivos de secuencia adicionales de una biblioteca para crear una fracción que engloba la longitud deseada del XTEN **504,** así como se ligan a una concentración más pequeña de un oligo que contiene los sitios de restricción BbsI y KpnI **503.** La fracción de productos de ligación resultante se purifica en gel y se corta la banda con la longitud deseada del XTEN, lo que da como resultado un gen XTEN aislado con una secuencia de tapón **505.** El gen XTEN se clona en un vector de relleno. En este caso, el vector codifica una secuencia de CBD opcional **506** y un gen GFP **508.** La digestión se realiza con BbsI/HindIII para eliminar **507** y **508** y colocar el codón de terminación. A continuación, el producto resultante se clona en un vector digerido con BsaI/HindIII que contiene un gen que codifica el svFv anti-Her2, lo que da como resultado que el gen **500** codifica una proteína de fusión de unión.

La FIG. 9 es un diagrama de flujo esquemático de etapas representativas en el ensamblaje de un gen que codifica una proteína de fusión de unión que comprende un resto de direccionamiento y XTEN, su expresión y recuperación como una proteína de fusión y su evaluación como un producto de proteína de fusión de unión candidato.

La FIG. 10 es una representación esquemática del diseño de vectores de expresión de proteína de fusión de unión anti-Her2 con diferentes estrategias de procesamiento. La FIG. 10A muestra un vector de expresión que codifica XTEN fusionado con el extremo 3' de la secuencia que codifica el resto de unión anti-Her2. Obsérvese que no se requieren secuencias de líder adicionales en este vector. La FIG. 10B representa un vector de expresión que codifica XTEN fusionado con el extremo 5' de la secuencia que codifica el resto de unión anti-Her2 con una secuencia líder CBD y un sitio de proteasa TEV. La FIG. 10C representa un vector de expresión como en la FIG. 10B donde el sitio de procesamiento de CBD y TEV se ha reemplazado por una secuencia líder aminoterminal optimizada (NTS). La FIG. 10D representa un vector de expresión que codifica una secuencia NTS, un XTEN, una secuencia que codifica el resto de unión anti-Her2, y a continuación, una segunda secuencia que codifica un XTEN.

La FIG. 11 muestra los resultados de los ensayos de expresión para las construcciones indicadas que comprenden las secuencias GFP y XTEN, realizados como se describe en el ejemplo 14. Los cultivos de expresión se ensayaron utilizando un lector de placas de fluorescencia (excitación a 395 nm, emisión a 510 nm) para determinar la cantidad de GFP indicadora presente, y los resultados se representan gráficamente como diagramas de cajas y bigotes.

La FIG. 12 muestra tres bibliotecas aleatorizadas usadas para el tercer y cuarto codón en las secuencias aminoterminales de clones de LCW546, LCW547 y LCW552. Las bibliotecas se diseñaron con el tercer y cuarto resto modificados de manera que todas las combinaciones de codones de XTEN permitidos estuvieran presentes en estas posiciones, como se muestra. Con el fin de incluir todos los codones XTEN permitidos para cada biblioteca, se diseñaron nueve pares de oligonucleótidos que codificaban 12 motivos de aminoácidos con diversidades de los codones en los terceros y cuartos restos, se hibridaron y se ligaron en el vector de relleno pCW0551 digerido con las enzimas de restricción NdeI/BsaI (Stuffer-XTEN_AM875-GFP), y se transformaron en células BL21Gold(DE3) de *E. coli* competentes para obtener colonias de las tres bibliotecas LCW0569, LCW0570 y LCW0571.

La FIG. 13 muestra un histograma de un nuevo ensayo de los 75 clones principales después de la etapa de optimización, como se describe en el ejemplo 15, para la señal de fluorescencia de la GFP con respecto a la construcción CBD_AM875 de referencia. Los resultados indicaron que varios clones eran ahora superiores a los clones de referencia observados en la FIG. 11.

La FIG. 14 es un esquema de un enfoque combinatorio emprendido para la unión de las preferencias de optimización de codones para dos regiones de los 48 aminoácidos aminoterminales. El enfoque creó 48meros novedosos en el extremo amínico de la proteína XTEN para la evaluación de la optimización de la expresión para las secuencias líder, para mejorar la expresión de las proteínas XTEN donde el XTEN es aminoterminal con respecto a los restos de direccionamiento.

La FIG. 15 muestra un gel de SDS-PAGE que confirma la expresión de los clones preferidos obtenidos de los experimentos de optimización de codones aminoterminales de XTEN, en comparación con los clones de

XTEN de referencia que comprenden secuencias líder CBD en el extremo amínico de las secuencias de la construcción.

La FIG. 16 muestra un gel de SDS-PAGE de muestras de un estudio de estabilidad de la proteína de fusión de XTEN_AE864 fusionada con el extremo amínico de GFP. La GFP-XTEN se incubó en plasma de macaco cangrejero y lisado renal de rata durante hasta 7 días a 37 °C, como se describe en el ejemplo 56. Además, también se evaluó la GFP-XTEN administrada a macacos cangrejeros. Las muestras se retiraron a los 0, 1 y 7 días y se analizaron mediante una SDS-PAGE, seguido de la detección mediante inmunoelectrotransferencia y la detección con anticuerpos contra la GFP.

La FIG. 17 muestra la caracterización de proteínas de fusión de unión con scFv multivalentes. La FIG. 17A es una representación esquemática de dos variaciones de proteínas de fusión de unión con scFv multivalentes, con una construcción monoespecífica a la izquierda que comprende dos restos de direccionamiento dirigidos a HER2 y una construcción biespecífica a la derecha que comprende un resto de direccionamiento a HER2 y un segundo resto de direccionamiento al EGFR. La FIG. 17B muestra una SDS-PAGE de los materiales después de la purificación, como se describe en el ejemplo 29. El carril 1 muestra los patrones de peso molecular, el carril 2 muestra el aHER2-XTEN-aHER2 purificado, y el carril 3 muestra el aHER2-XTEN-aEGFR biespecífico purificado. La FIG. 17C muestra los resultados del análisis de la SEC de aHER2-XTEN-aEGFR en comparación con los patrones de peso molecular, y demuestra que no se forman dímeros ni otros oligómeros de orden superior y que la proteína tiene un peso molecular aparente aproximado de aproximadamente 500 kDa, como se describe en el ejemplo 37.

La FIG. 18 muestra los resultados de un ensayo de actividad de unión de la proteína de fusión de unión biespecífica aHER2-XTEN-aEGFR a sus respectivas dianas utilizando un formato de ELISA, como se describe en el ejemplo 37.

La FIG. 19 muestra un esquema de dos construcciones de proteínas de fusión de unión con scFv con una etiqueta de GFP y los resultados de la citometría de flujo de los ensayos de unión celular. La FIG. 19A es un esquema de las construcciones aCD3-XTEN-GFP y aHER2-XTEN-GFP. La FIG. 19B muestra el resultado de la citometría de flujo en la que las dos construcciones se hicieron reaccionar individualmente con células Jurkat CD3+ y SK-BR3 HER2+, como se describe en el ejemplo 36, lo que demuestra la especificidad de unión de las respectivas construcciones hacia sus ligandos y la ausencia de unión a las dianas heterólogas.

La FIG. 20 muestra los resultados de los ensayos de caracterización de una proteína de fusión de unión con scFv biespecífica, como se describe en el ejemplo 36. La FIG. 20A muestra un gel de SDS-PAGE de la aHER2-aCD3-XTEN purificada. La FIG. 20B muestra el resultado de un análisis de cromatografía de exclusión por tamaño (SEC) de la aHER2-aCD3-XTEN en comparación con los patrones de peso molecular, y demuestra que no se forman dímeros ni otros oligómeros de orden superior y que la proteína tiene un peso molecular aparente aproximado de aproximadamente 500 kDa, aproximadamente cinco veces mayor que la masa indicada en el ensayo de SDS-PAGE de la FIG. 20A.

La FIG. 21 muestra los resultados de los ensayos de caracterización de una proteína de fusión de unión con scFv biespecífica, como se describe en el ejemplo 36. La FIG. 21A muestra los resultados de un ensayo de unión realizado mediante un ELISA que compara un scFv de anti-Her2 en el extremo amínico de una proteína de fusión de unión con un scFv biespecífico anti-CD3 en el extremo amínico y anti-Her2 en el extremo carboxílico de un XTEN, frente a pocillos recubiertos con HER2, mostrando que el biespecífico conserva la afinidad de unión por la diana HER2. Las FIG. 21B y C muestran los resultados de ensayos de unión celular de citometría de flujo utilizando las construcciones mono- y biespecíficas, respectivamente, frente a células SK-BR-3 que expresan HER2. Los resultados muestran una señal mayor para el scFv-XTEN anti-Hers aminoterminal (FIG. 21B) que la de la construcción biespecífica, lo que es coherente con la del estudio por ELISA, pero que la construcción biespecífica conserva, no obstante, una buena actividad de unión.

La FIG. 22 muestra los resultados de un ensayo de citometría de flujo para caracterizar la unión de aCD3-XTEN-GFP a células Jurkat positivas para CD3. El ensayo se realizó utilizando la detección de anticuerpos anti-GFP de aCD3-XTEN-GFP que reaccionó en presencia de un exceso molar de 10 veces de la aHER2-aCD3-XTEN, como se describe en el ejemplo 36. Los resultados muestran que el exceso de aHER2-aCD3-XTEN biespecífica desplaza competitivamente la proteína aCD3-XTEN-GFP monoespecífica y elimina el desplazamiento observado en la MFI.

La FIG. 23 muestra los resultados de un ensayo de destrucción de células tumorales en el que se incubaron concentraciones variables de la aHER2-XTEN-aCD3 biespecífica con células diana M21 o SK-BR-3 durante 24 h, como se describe en el ejemplo 36, seguido de una tinción con yoduro de propidio para medir la destrucción, se muestra en el gráfico de barras como la proporción de células tumorales muertas para las dos líneas celulares.

La FIG. 24 muestra representaciones esquemáticas de construcciones de proteínas de fusión de unión Vhh única y multivalente y su caracterización. La FIG. 24A muestra descripciones esquemáticas de construcciones Vhh anti-EGFR monómeras, dímeras, tetrámeras y hexámeras unidas por XTEN, con una GFP carboxiterminal. La FIG. 24B muestra la SDS-PAGE del lisado y las construcciones Vhh purificadas, como se describe en el ejemplo 38, que demuestra la pureza y el aumento escalonado en el peso molecular con unidades crecientes del resto de direccionamiento Vhh anti-EGFR.

La FIG. 25 muestra el resultado del análisis de cromatografía de exclusión por tamaño (SEC) de las construcciones de proteína de fusión de unión Vhh anti-EGFR monómera y multivalente de la FIG. 24 en comparación con los patrones de peso molecular, como se describe en el ejemplo 38. Los resultados demuestran unos aumentos proporcionales del peso molecular aparente aproximado de las construcciones con un número creciente de restos de direccionamiento Vhh, en comparación con los pesos moleculares determinados por SDS-PAGE en la FIG. 24 y enumerados en la tabla 26.

La FIG. 26 muestra los resultados de los ensayos de ELISA de caracterización de la unión de las construcciones de restos de direccionamiento monómeros y multivalentes de la proteína de fusión de unión Vhh anti-EGFR, como se describe en el ejemplo 38 (las construcciones se representan esquemáticamente en la FIG. 24A). La FIG. 26A muestra los resultados de la señal del ELISA generada utilizando unas concentraciones equimolares de las diferentes proteínas de fusión de unión Vhh anti-EGFR monómeras o multivalentes contra la diana de EGFR. La FIG. 26B muestra los resultados de curvas de unión para las mismas construcciones en diferentes diluciones, dando las formas multivalentes como resultado más señal que la construcción con un único resto de direccionamiento Vhh.

La FIG. 27 muestra los resultados de un ensayo de ELISA de caracterización de la unión de dos construcciones de proteína de fusión de unión con scFv anti-CD40, AC384 (cuadrados negros) y AC385 (cuadrados blancos), contra el CD40 humano, como se describe en el ejemplo 34.

La FIG. 28 muestra los resultados de los ensayos de caracterización para una proteína de fusión de unión anti-IL6R. La FIG. 29A muestra la uniformidad de la proteína purificada evaluada mediante SEC, que mostró un máximo monodisperso con una contaminación mínima. La FIG. 29B muestra los resultados de un ensayo de unión ELISA de la proteína de fusión de unión anti-IL6R contra la IL6R humana, como se describe en el ejemplo 33.

La FIG. 29 muestra los resultados de un ensayo de caracterización de unión ELISA de la proteína de unión anti-Her2 con dos restos de direccionamiento anti-Her2 contra el HER2 humano, como se describe en el ejemplo 35.

La FIG. 30 muestra el espectro de dicroísmo circular de UV cercano de Ex4-XTEN_AE864, realizado como se describe en el ejemplo 52.

La FIG. 31 muestra el perfil farmacocinético (concentraciones plasmáticas) en macacos cangrejeros después de dosis únicas de diferentes composiciones de GFP unidas a polipéptidos no estructurados de longitud variable, administradas por vía subcutánea o intravenosa, como se describe en el ejemplo 53. Las composiciones eran GFP-L288, GFP-L576, GFP-XTEN AF576, GFP-Y576 y XTEN AD836-GFP. Las muestras de sangre se analizaron en diferentes momentos después de la inyección, y la concentración de GFP en plasma se midió mediante un ELISA utilizando un anticuerpo policlonal contra la GFP para la captura y una preparación biotinilada del mismo anticuerpo policlonal para la detección. Los resultados se presentan como la concentración plasmática frente al tiempo (h) después de la administración y muestran, en particular, un aumento considerable en la semivida para XTEN_AD836-GFP, la composición con la longitud de secuencia más larga de XTEN. La construcción con la longitud de secuencia más corta, la GFP-L288, tuvo la semivida más corta.

La FIG. 32 muestra los resultados de un análisis cromatográfico de exclusión por tamaño de muestras de construcciones de glucagón-XTEN medidas frente a patrones de proteínas de peso molecular conocido, con el resultado del gráfico como absorbancia frente al volumen de retención, como se describe en el ejemplo 60. Las construcciones de glucagón-XTEN son 1) glucagón-Y288; 2) glucagón-Y144; 3) glucagón-Y72; y 4) glucagón-Y36. Los resultados indican un aumento en el peso molecular aparente al aumentar la longitud del componente XTEN.

La FIG. 33 muestra superposiciones de cromatogramas de SEC de tres proteínas de fusión conjugadas aHer2-XTEN(Cys)-AF680 con tres XTEN diferentes; AE864, AE576 y AE288 (de arriba a abajo), realizadas como se describe en el ejemplo 41. Los cromatogramas muestran que el Alexa Fluor 680 se conjugó con éxito con la proteína aHer2-XTEN(Cys), ya que el tiempo de retención de la absorbancia de Alexa Fluor 680 (A690) se superpone con el de la proteína aHer2-XTEN(Cys) (A280) en las tres muestras, y que las construcciones eluyeron de manera proporcional con respecto a la longitud del componente XTEN. Obsérvese que el tiempo de retención del colorante se desplaza significativamente desde el tiempo de elución

esperado del colorante libre a los ~52 minutos. Adicionalmente, la relativa ausencia de material que eluye por delante de los diferentes máximos de aHer2-XTEN(Cys) indica una ausencia de agregación y un producto monodisperso.

La FIG. 34 muestra el resultado de los ensayos de citometría de flujo para las tres proteínas de fusión descritas en la FIG. 33 en comparación con el conjugado Herceptin sin marcar y el control de IgG-AF680, midiendo la dispersión frontal y lateral frente a FL4 para Alexa680, como se describe en el ejemplo 41. La FIG. 34A es el resultado del ensayo aHer2-XTEN_AE864-AF680. La FIG. 34B es el resultado del ensayo aHer2-XTEN_AE576-AF680, y la FIG. 34C es el resultado del ensayo aHer2-XTEN_AE288-AF680.

La FIG. 35 muestra los resultados gráficos de los datos de la obtención de imágenes *in vivo* de ratones hembra nu/nu portadores de células tumorales SKOV3 a los que se les administró una única inyección de dosis alta o baja de aHer2-XTEN-AE-288-Cys-AF680, aHer2-XTEN-AE-576-Cys-AF680, aHer2-XTEN-AE-864-Cys-AF680 o Herceptin-AF680 de control, como se describe en el ejemplo 41.

La FIG. 36 muestra los resultados gráficos de los datos de la obtención de imágenes *ex vivo* de los mismos grupos de tratamiento que en la FIG. 36, que demuestran que todas las construcciones de proteína de fusión de unión aHer2-XTEN penetraron en los tejidos ensayados, mostrando la construcción AE864 XTEN más larga los niveles más altos en comparación con las otras dos, como se describe en el ejemplo 41.

La FIG. 37 es un esquema del diagrama de flujo lógico del algoritmo SegScore. En la figura se aplica la siguiente leyenda: i, j - contadores usados en los bucles de control que se ejecutan a través de toda la secuencia; HitCount- esta variable es un contador que realiza un seguimiento de cuántas veces una subsecuencia encuentra una subsecuencia idéntica en un bloque; SubSeqX - esta variable contiene la subsecuencia cuya redundancia se está comprobando; SubSeqY - esta variable contiene la subsecuencia contra la que se comprueba SubSeqX; BlockLen - esta variable contiene la longitud determinada por el usuario del bloque; SegLen - esta variable contiene la longitud de un segmento. El programa tiene una codificación rígida para generar puntuaciones para subsecuencias de longitudes 3, 4, 5, 6, 7, 8, 9 y 10; Block - Esta variable contiene una cadena de longitud BlockLen. La cadena está compuesta por letras de una secuencia del XTEN de entrada y está determinada por la posición del contador i; SubSeqList - esta es una lista que contiene todas las puntuaciones de subsecuencias generadas.

La FIG. 38 representa la aplicación del algoritmo SegScore a un hipotético XTEN de 11 aminoácidos con el fin de determinar la repetitividad. Una secuencia del XTEN que consiste en N aminoácidos se divide en subsecuencias N-S+1 de longitud S (S=3 en este caso). Se realiza una comparación por pares de todas las subsecuencias y se calcula el número promedio de subsecuencias idénticas para dar como resultado, en este caso, una puntuación de subsecuencia de 1,89.

La FIG. 39 ilustra el uso de secuencias XTEN donantes para producir secuencias XTEN truncadas. La FIG. 39A proporciona la secuencia de AG864, con la secuencia subrayada utilizada para generar una secuencia de AG576. La FIG. 39B proporciona la secuencia de AG864, con la secuencia subrayada utilizada para generar una secuencia de AG288. La FIG. 39C proporciona la secuencia de AG864, con la secuencia subrayada utilizada para generar una secuencia de AG144. La FIG. 39D proporciona la secuencia de AE864, con la secuencia subrayada utilizada para generar una secuencia de AE576. La FIG. 39E proporciona la secuencia de AE864, con la secuencia subrayada utilizada para generar una secuencia de AE288.

La FIG. 40 muestra diferentes ejemplos esquemáticos de conjugados de proteína y fármaco basados en XTEN, denominándose "D" el ligando conjugado químicamente de fármaco y reticulador. La FIG. 40A muestra tres conjugados de XTEN-fármaco, con 1, 2 y 4 moléculas de fármaco conjugadas con el XTEN. La FIG. 40B muestra cuatro configuraciones de BFP-D con el dominio de unión ("BD") unido al extremo amínico o carboxílico del XTEN de la proteína de fusión, y 1, 3 o 4 moléculas de fármaco conjugadas con el portador XTEN mediante agentes de conjugación. La FIG. 40C muestra cuatro configuraciones de BFP-D con el dominio de unión scFv en el extremo amínico o carboxílico del XTEN, y 1, 3 o 4 moléculas de fármaco conjugadas con el portador XTEN mediante agentes de conjugación. La FIG. 40D muestra cuatro configuraciones de BFP-D multivalente con dos dominios de unión ("BD") en el extremo amínico o carboxílico, o configuraciones con un XTEN aminoterminal y cualquiera de las moléculas de fármaco 1, 2, 3 o 4 conjugadas con el portador XTEN o XTEN aminoterminal mediante agentes de conjugación.

La FIG. 41 muestra diferentes ejemplos esquemáticos del proceso de conjugación para crear conjugados con múltiples ligandos de fármaco utilizando unas químicas de acoplamiento ortogonal. La FIG. 41A muestra un proceso de acoplamiento en dos etapas de un ligando de agente de conjugación y fármaco (D1) a una cisteína interna de un XTEN genomodificado con cisteína, y un segundo ligando de agente de conjugación y fármaco diferente (D2) al extremo amínico del XTEN. La FIG. 41B muestra un proceso de acoplamiento en dos etapas de un ligando a la cisteína interna de un XTEN genomodificado con cisteína y dos ligandos a grupos amino en un dominio de unión (BD) recombinante. La FIG. 41C muestra un proceso de acoplamiento en dos etapas de dos ligandos de fármaco (D1) a dos cisteínas interna de un XTEN genomodificado con cisteína y un

segundo ligando de fármaco (D2) al extremo amínico de XTEN. La FIG. 41D muestra un proceso de acoplamiento en dos etapas de dos ligandos de fármaco (D1) a las cisteínas internas del XTEN y 1 ligando de agente de conjugación y fármaco (D2) al extremo amínico y dos ligandos de agente de conjugación y fármaco D2 adicionales a restos de lisina internos del XTEN genomodificado.

La FIG. 42 muestra los resultados de los ensayos analíticos del XTEN conjugado con FITC reticulada, como se describe en el ejemplo 63. La FIG. 42A muestra la comigración en un gel fotografiado a través de una caja de luz UV para mostrar el gran PM aparente de las especies conjugadas que contienen FITC, también detectado por SEC a una DO214 (señal de la proteína) y una DO495 (señal de la FITC) en una columna de SEC, lo que indica el etiquetado exitoso del XTEN con una contaminación mínima de colorante libre. Los materiales por carril (de izquierda a derecha, después de los patrones de PM son: FITC-CL-CBD-XTEN marcado; FITC-CL-XTEN marcado; FITC-CL-XTEN purificado; FITC-CL-XTEN purificado; y FITC-CL-XTEN purificado. El gel se fotografió a través de una caja de luz UV para mostrar el PM aparente de la FITC de especies que contenían la FITC. La FIG. 42B muestra los resultados del análisis por SEC del XTEN conjugado con FITC, que muestra el solapamiento del resultado de materiales detectados a una DO214 y una DO495, y también el gran peso molecular aparente.

La FIG. 43 muestra los resultados de los análisis por SEC de las fracciones de elución máximas de los conjugados de la GFP reticulada con XTEN y de la GFP libre, como se describe en el ejemplo 64. La reticulación se confirmó mediante la comigración de la señal de la proteína a una DO214 y la señal de la GFP a una DO395 en la columna de SEC.

La FIG. 44 muestra los resultados de los ensayos farmacocinéticos de GFP-X-XTEN y FITC-X-XTEN ensayadas en macacos cangrejeros, como se describe en el ejemplo 65.

La FIG. 45 muestra una representación esquemática del proceso de química ortogonal para crear BFP-D que comprende un anti-HER2 scFv con paclitaxel conjugado con el XTEN, lo que da como resultado aHER2-XTEN-CL-paclitaxel. El paclitaxel se hace reaccionar en primer lugar con un conector activado, a continuación se conjuga con un XTEN genomodificado con cisteína, como se describe en el ejemplo 66.

Descripción detallada de la invención

Antes de describir las realizaciones de la invención, debe entenderse que dichas realizaciones se proporcionan a modo de ejemplo.

Salvo que se indique lo contrario, todos los términos técnicos y científicos utilizados en el presente documento tienen el mismo significado que entiende comúnmente un experto habitual en la materia a la que pertenece esta invención. Aunque en la práctica o los ensayos de la presente invención se pueden utilizar métodos y materiales similares o equivalentes a los descritos en el presente documento, a continuación se describen métodos y materiales ilustrativos. En caso de conflicto, prevalecerá la memoria descriptiva de la patente, incluidas las definiciones. Además, los materiales, métodos y ejemplos son meramente ilustrativos y no pretenden ser limitativos.

DEFINICIONES

Como se utiliza en el presente documento, los siguientes términos tienen los significados atribuidos a los mismos, salvo que se indique lo contrario.

Como se utiliza en la memoria descriptiva y en las reivindicaciones, las formas en singular "un", "uno/a" y "el/la" incluyen las referencias en plural a menos que el contexto indique claramente lo contrario. Por ejemplo, la expresión "una célula" incluye una pluralidad de células, incluyendo mezclas de las mismas.

Los términos "polipéptido", "péptido" y "proteína" se usan indistintamente en el presente documento para referirse a polímeros de aminoácidos de cualquier longitud. El polímero puede ser lineal o ramificado, puede comprender aminoácidos modificados y puede estar interrumpido por no aminoácidos. Los términos también engloban un polímero de aminoácidos que ha sido modificado, por ejemplo, mediante la formación de puentes de disulfuro, glicosilación, lipidación, acetilación, fosforilación o cualquier otra manipulación, tal como una conjugación con un componente de etiquetado.

Como se utiliza en el presente documento, el término "aminoácido" se refiere a aminoácidos naturales y/o no naturales o sintéticos, que incluyen, pero sin limitación, glicina y ambos isómeros ópticos D o L, y análogos de aminoácidos y peptidomiméticos. Para designar los aminoácidos se utilizan los códigos convencionales de una o tres letras.

La expresión "L-aminoácido natural" significa las formas que son isómeros ópticos L de glicina (G), prolina (P), alanina (A), valina (V), leucina (L), isoleucina (I), metionina (M), cisteína (C), fenilalanina (F), tirosina (Y), triptófano (W), histidina (H), lisina (K), arginina (R), glutamina (Q), asparagina (N), ácido glutámico (E), ácido aspártico (D), serina (S) y treonina (T).

La expresión "no natural", como se aplica en las secuencias y como se utiliza en el presente documento, significa secuencias polipeptídicas o polinucleotídicas que no tienen un compañero, no son complementarias o no tienen un alto grado de homología con una secuencia natural que se encuentra en un mamífero. Por ejemplo, un polipéptido o un fragmento de origen no natural puede compartir no más del 99 %, 98 %, 95 %, 90 %, 80 %, 70 %, 60 %, 50 % o incluso menos identidad de secuencia de aminoácidos en comparación con una secuencia natural cuando se alinean adecuadamente.

Los términos "hidrófilo" e "hidrófobo" se refieren al grado de afinidad que una sustancia tiene por el agua. Una sustancia hidrófila tiene una fuerte afinidad por el agua, tendiendo a disolverse en, mezclarse con o ser humedecida por el agua, mientras que una sustancia hidrófoba carece sustancialmente de afinidad por el agua, tendiendo a repeler y no absorber agua y tendiendo a no disolverse ni mezclarse con el agua ni a ser humedecida por el agua. Los aminoácidos pueden caracterizarse en función de su hidrofobia. Se han desarrollado varias escalas. Un ejemplo es una escala desarrollada por Levitt, M, et al., J Mol Biol (1976) 104: 59, que se recoge en Hopp, TP, et al., Proc Natl Acad Sci U S A (1981) 78: 3824. Algunos ejemplos de "aminoácidos hidrófilos" son arginina, lisina, treonina, alanina, asparagina, y glutamina. De particular interés son los aminoácidos hidrófilos aspartato, glutamato, y serina y glicina. Algunos ejemplos de "aminoácidos hidrófobos" son triptófano, tirosina, fenilalanina, metionina, leucina, isoleucina y valina.

Un "fragmento" es una forma truncada de una proteína biológicamente activa natural que conserva al menos una porción de la actividad terapéutica y/o biológica. Una "variante" es una proteína con homología de secuencia con la proteína biológicamente activa natural que conserva al menos una porción de la actividad terapéutica y/o biológica de la proteína biológicamente activa. Por ejemplo, una proteína variante puede compartir al menos un 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de la secuencia de aminoácidos con la proteína biológicamente activa de referencia. Como se utiliza en el presente documento, la expresión "resto de proteína biológicamente activo" incluye proteínas modificadas deliberadamente, como por ejemplo, por mutagénesis dirigida, inserciones, o accidentalmente a través de mutaciones.

Una "célula hospedadora" incluye una célula individual o un cultivo celular que puede ser o ha sido un receptor para los presentes vectores. Las células hospedadoras incluyen la descendencia de una única célula hospedadora. La descendencia puede no ser necesariamente completamente idéntica (en morfología o en genómica del complemento de ADN total) a la célula progenitora original debido a una mutación natural, accidental o deliberada. Una célula hospedadora incluye las células transfectadas *in vivo* con un vector de esta invención.

Cuando se utiliza "aislado" para describir los diferentes polipéptidos divulgados en el presente documento, significa un polipéptido que se ha identificado y separado y/o recuperado a partir de un componente de su entorno natural. Los componentes contaminantes de su entorno natural son materiales que normalmente interferirían con los usos diagnósticos o terapéuticos del polipéptido, y pueden incluir enzimas, hormonas y otros solutos proteicos o no proteicos. Como es evidente para los expertos en la técnica, un polinucleótido, un péptido, un polipéptido, una proteína, un anticuerpo o fragmentos de los mismos no naturales no requieren "aislamiento" para distinguirlos de su homólogo natural. Además, un polinucleótido, un péptido, un polipéptido, una proteína, un anticuerpo "concentrado", "separado" o "diluido", o fragmentos de los mismos, se distinguen de su compañero natural en que la concentración o el número de moléculas por volumen es generalmente mayor que la de su homólogo natural. En general, se considera que un polipéptido producido por medios recombinantes y expresado en una célula hospedadora está "aislado".

Un ácido nucleico que codifica un polinucleótido o un polipéptido "aislado" o un otro ácido nucleico que codifica otro polipéptido es una molécula de ácido nucleico que se identifica y se separa de al menos una molécula de ácido nucleico contaminante con la que normalmente está asociado en la fuente natural del ácido nucleico que codifica el polipéptido. Una molécula de ácido nucleico aislada que codifica un polipéptido es distinta de la forma o el entorno en el que se encuentra en la naturaleza. Por lo tanto, las moléculas aisladas de ácido nucleico que codifican un polipéptido se distinguen de la molécula específica de ácido nucleico que codifica un polipéptido tal como existe en las células naturales. Sin embargo, una molécula de ácido nucleico aislada que codifica un polipéptido incluye moléculas de ácido nucleico que codifican un polipéptido contenidas en células que normalmente expresan el polipéptido donde, por ejemplo, la molécula de ácido nucleico está en una ubicación cromosómica o extracromosómica diferente de la de las células naturales.

Una proteína "quimérica" contiene al menos un polipéptido de fusión que comprende regiones en una posición diferente en la secuencia que la que se produce en la naturaleza. Las regiones pueden existir normalmente en proteínas individuales y se unen en el polipéptido de fusión; o pueden existir normalmente en la misma proteína pero se colocan en una nueva disposición en el polipéptido de fusión. Una proteína quimérica puede crearse, por ejemplo, mediante síntesis química, o creando y traduciendo un polinucleótido en el que las regiones peptídicas están codificadas en la relación deseada.

"Conjugado" y "conjugación" se refiere a la unión covalente entre dos o más elementos químicos o componentes mediante una reacción química, en lugar de recombinantemente; por ejemplo, un fármaco y un XTEN.

Un "agente de conjugación" o "CL" significa un resto químico que comprende un enlace covalente, un conector de fármaco o una cadena de átomos que conjugan covalentemente un resto de fármaco con una proteína. Los componentes de agentes de conjugación pueden comprender uno o dos grupos reactivos para facilitar la conjugación de un fármaco y una proteína, y los grupos reactivos pueden haber sido bloqueados con grupos protectores para permitir la conjugación selectiva con un fármaco o reactivo proteico dado.

Un "grupo reactivo" es una estructura química o un grupo funcional que es parte de, o que puede acoplarse a, un reactivo para la conjugación. Algunos ejemplos de grupos reactivos son grupos amino, grupos carboxilo, grupos sulfhidrilo, grupos hidroxilo, grupos aldehído, grupos azida. Algunos grupos reactivos se pueden activar para facilitar el acoplamiento con un segundo grupo reactivo, tal como un componente de agente de conjugación. Algunos ejemplos de activación son la reacción de un grupo carboxilo con carbodiimida, la conversión de un grupo carboxilo en un éster activado, la conversión de un grupo carboxilo en una función azida o la conversión de un hidroxilo en un tiol.

La expresión "agente citotóxico", como se utiliza en el presente documento, se refiere a una sustancia que causa la destrucción de las células. La expresión pretende incluir isótopos radiactivos, fármacos y toxinas tales como toxinas de molécula pequeña o toxinas enzimáticamente activas de origen bacteriano, fúngico, vegetal o animal, incluidos análogos sintéticos y los derivados de los mismos.

La expresión "agente citostático" se refiere a una sustancia que tiene el efecto de limitar la función de las células, tal como limitar el crecimiento celular o la proliferación de células.

Un "fármaco" es un compuesto químico no proteico útil en el tratamiento de una enfermedad, un trastorno o una afección en un sujeto; por ejemplo, cáncer, enfermedades inflamatorias o afecciones tales como artritis reumatoide, trastornos metabólicos tales como diabetes, enfermedades infecciosas, enfermedades de los órganos tales como asma, o afecciones generalizadas tales como dolor e hipertensión, etc. La administración de un fármaco a un sujeto o a una célula puede, por ejemplo, dar como resultado un efecto farmacológico, un efecto terapéutico, un efecto inhibidor, o dar como resultado la muerte celular.

En el contexto de los polipéptidos, "fusión", "fusionado" y "unido" se usan indistintamente en el presente documento. Estos términos se refieren a la unión de dos componentes proteicos más, por cualquier medio que incluya una conjugación química o medios recombinantes. Por ejemplo, un promotor o un potenciador está unido operativamente a una secuencia codificante si afecta a la transcripción de la secuencia. Generalmente, "unido operativamente" significa que las secuencias de ADN que se están uniendo son contiguas y están en fase de lectura o en marco. Una "fusión en marco" se refiere a la unión de dos o más marcos abiertos de lectura (ORF) para formar un ORF continuo más largo, de manera que se mantenga el marco de lectura correcto de los ORF originales. Por lo tanto, la proteína de fusión recombinante resultante es una única proteína que contiene dos o más segmentos que corresponden a los polipéptidos codificados por los ORF originales (cuyos segmentos normalmente no están tan unidos en la naturaleza).

En el contexto de polipéptidos, una "secuencia lineal" o una "secuencia" es el orden de los aminoácidos de un polipéptido en una dirección de extremo amínico a carboxílico en la que los restos cercanos entre sí en la secuencia están contiguos en la estructura primaria del polipéptido. Una "secuencia parcial" es una secuencia lineal de parte de un polipéptido que se sabe que comprende restos adicionales en una o ambas direcciones.

"Heterólogo" significa derivado de una entidad genotípicamente distinta del resto de la entidad con la que se está comparando. Por ejemplo, una secuencia rica en glicina eliminada de su secuencia codificante natural y unida operativamente a una secuencia codificante distinta de la secuencia natural es una secuencia heteróloga rica en glicina. El término "heterólogo", como se aplica a un polinucleótido, un polipéptido, significa que el polinucleótido o el polipéptido deriva de una entidad genotípicamente distinta de la del resto de la entidad con la que se está comparando.

Las expresiones "polinucleótidos", "ácidos nucleicos", "nucleótidos" y "oligonucleótidos" se usan indistintamente. Se refieren a una forma polimérica de nucleótidos de cualquier longitud, ya sea desoxirribonucleótidos o ribonucleótidos, o análogos de los mismos. Los polinucleótidos pueden tener cualquier estructura tridimensional y pueden realizar cualquier función, conocida o desconocida. Los siguientes son algunos ejemplos no limitativos de polinucleótidos: regiones codificantes o no codificantes de un gen o un fragmento génico, locus definidos a partir del análisis de enlaces, exones, intrones, ARN mensajero (ARNm), ARN de transferencia, ARN ribosómico, ribozimas, ADNc, polinucleótidos recombinantes, polinucleótidos ramificados, plásmidos, vectores, ADN aislado de cualquier secuencia, ARN aislado de cualquier secuencia, sondas de ácido nucleico y cebadores. Un polinucleótido puede comprender nucleótidos modificados, tales como nucleótidos metilados y análogos de nucleótidos. Si están presentes, las modificaciones en la estructura del nucleótido pueden impartirse antes o después del ensamblaje del polímero. La secuencia de nucleótidos puede estar interrumpida por componentes no nucleotídicos. Un polinucleótido puede modificarse adicionalmente después de la polimerización, tal como mediante conjugación con un componente de etiquetado.

La expresión "complemento de un polinucleótido" designa una molécula polinucleotídica que tiene una secuencia de base complementaria y una orientación inversa en comparación con una secuencia de referencia, de modo que podría hibridar con una secuencia de referencia con una fidelidad completa.

"Recombinante", como se aplica a un polinucleótido, significa que el polinucleótido es el producto de varias combinaciones de etapas de clonación, restricción y/o ligación *in vitro*, y otros procedimientos que dan como resultado una construcción que potencialmente puede ser expresada en una célula hospedadora.

Los términos "gen" o "fragmento génico" se usan indistintamente en el presente documento. Se refieren a un polinucleótido que contiene al menos un marco abierto de lectura que es capaz de codificar una proteína particular después de ser transcrito y traducido. Un gen o un fragmento génico puede ser genómico o ADNc, siempre que el polinucleótido contenga al menos un marco abierto de lectura, que puede cubrir toda la región codificante o un segmento de la misma. Un "gen de fusión" es un gen compuesto por al menos dos polinucleótidos heterólogos que están unidos entre sí.

"Homología" u "homólogo" se refiere a la similitud de secuencia o intercambiabilidad entre dos o más secuencias polinucleotídicas o dos o más secuencias polipeptídicas. Cuando se utiliza un programa tal como BestFit para determinar la identidad de secuencia, la similitud o la homología entre dos secuencias de aminoácidos diferentes, se pueden usar los ajustes predeterminados o se puede seleccionar una matriz de puntuación apropiada, tal como blosum45 o blosum80, para optimizar las puntuaciones de identidad, similitud u homología. Preferentemente, los polinucleótidos homólogos son aquellos que hibridan en condiciones rigurosas como se define en el presente documento, y tienen al menos un 70 %, preferentemente al menos un 80 %, más preferentemente al menos un 90 %, más preferentemente un 95 %, más preferentemente un 97 %, más preferentemente un 98 % e incluso más preferentemente un 99 % de identidad de secuencia con esas secuencias.

Las expresiones "condiciones rigurosas" o "condiciones de hibridación rigurosas" incluyen la referencia a condiciones en las que un polinucleótido hibridará con su secuencia diana hasta un grado detectablemente mayor que con otras secuencias (por ejemplo, al menos 2 veces con respecto al fondo). Generalmente, la rigurosidad de la hibridación se expresa, en parte, mediante referencia a la temperatura y la concentración salina en las que se lleva a cabo la etapa de lavado. Normalmente, las condiciones rigurosas serán aquellas en las que la concentración salina es menor de aproximadamente 1,5 M de ion Na, normalmente de aproximadamente 0,01 a 1,0 M de concentración de ion Na (u otras sales) a un pH de 7,0 a 8,3, y la temperatura es de al menos aproximadamente 30 °C para polinucleótidos cortos (por ejemplo, de 10 a 50 nucleótidos) y al menos aproximadamente 60 °C. Para polinucleótidos largos (por ejemplo, de más de 50 nucleótidos), por ejemplo, las "condiciones rigurosas" pueden incluir una hibridación en formamida al 50 %, NaCl 1 M, 1 % de SDS a 37 °C y tres lavados durante 15 min cada uno en 0,1×SSC/1 % de SDS a entre 60 °C y 65 °C. Alternativamente, se pueden usar unas temperaturas de aproximadamente 65 °C, 60 °C, 55 °C o 42 °C. La concentración de SSC puede variar de aproximadamente 0,1 a 2×SSC, estando el SDS presente aproximadamente a un 0,1 %. Dichas temperaturas de lavado se seleccionan normalmente para que sean aproximadamente entre 5 °C y 20 °C más bajas que el punto de fusión térmico © para la secuencia específica a una fuerza iónica y pH definidos. La Tm es la temperatura (a una fuerza iónica y pH definidos) a la que el 50 % de la secuencia diana hibrida con una sonda perfectamente emparejada. Una ecuación para calcular Tm y las condiciones para la hibridación de ácidos nucleicos son bien conocidas y se pueden encontrar en Sambrook, J. et al (1989) Molecular Cloning: A Laboratory Manual, 2ª ed., vol. 1-3, Cold Spring Harbor Press, Plainview N.Y.; véanse específicamente el volumen 2 y el capítulo 9. Normalmente se usan reactivos de bloqueo para bloquear la hibridación no específica. Dichos reactivos de bloqueo incluyen, por ejemplo, ADN de esperma de salmón cizallado y desnaturalizado a aproximadamente 100-200 µg/ml. También se puede usar un disolvente orgánico, tal como formamida, a una concentración de aproximadamente el 35-50 % v/v, en algunas circunstancias particulares, tales como para hibridaciones de ARN:ADN. Las variaciones útiles en estas condiciones de lavado serán perfectamente evidentes para los expertos familiarizados con la materia.

Las expresiones "porcentaje de identidad" y "% de identidad", tal como se aplican a las secuencias polinucleotídicas, se refieren al porcentaje de coincidencias de restos entre al menos dos secuencias polinucleotídicas alineadas utilizando un algoritmo normalizado. Dicho algoritmo puede insertar, de una forma normalizada y reproducible, huecos en las secuencias que se están comparando con el fin de optimizar la alineación entre dos secuencias y, por lo tanto, conseguir una comparación más significativa de las dos secuencias. El porcentaje de identidad puede medirse a lo largo de una secuencia polinucleotídica definida completa, por ejemplo, como se define por un número de SEQ ID particular, o puede medirse a lo largo de una longitud más corta, por ejemplo, a lo largo de un fragmento tomado de una secuencia polinucleotídica más grande y definida, por ejemplo, un fragmento de al menos 45, al menos 60, al menos 90, al menos 120, al menos 150, al menos 210 o al menos 450 restos contiguos. Dichas longitudes son únicamente a modo de ejemplo, y se entiende que puede utilizarse cualquier longitud de fragmento soportada por las secuencias mostradas en el presente documento, en las tablas, las figuras o la lista de secuencias para describir una longitud sobre la cual se puede medir la identidad porcentual.

El "porcentaje (%) de identidad de una secuencia de aminoácidos" con respecto a las secuencias polipeptídicas identificadas en el presente documento, se define como el porcentaje de restos de aminoácidos de una secuencia de consulta que son idénticos a los restos de aminoácidos de una segunda secuencia polipeptídica de referencia o una porción de la misma, después de alinear las secuencias e introducir huecos, si fuera necesario, para conseguir el máximo porcentaje de identidad de secuencia, y sin considerar ninguna sustitución conservadora como parte de la identidad de secuencia. La alineación con el fin de determinar el porcentaje de identidad de secuencia de aminoácidos se puede conseguir de diversas maneras que están dentro de los conocimientos especializados en la técnica, por

ejemplo, utilizando un programa informático disponible para uso público tal como el programa informático BLAST, BLAST-2, ALIGN o Megalign (DNASTAR). Los expertos en la técnica pueden determinar los parámetros apropiados para medir la alineación, incluyendo los algoritmos necesarios para conseguir la alineación máxima a lo largo de toda la longitud de las secuencias que se están comparando. La identidad porcentual puede medirse a lo largo de toda la secuencia de polipéptidos definida, por ejemplo, como se define por un número de SEQ ID particular, o puede medirse a lo largo de una longitud más corta, por ejemplo, a lo largo de un fragmento tomado de una secuencia de polipéptidos más grande y definida, por ejemplo, un fragmento de al menos 15, al menos 20, al menos 30, al menos 40, al menos 50, al menos 70 o al menos 150 restos contiguos. Dichas longitudes son únicamente a modo de ejemplo, y se entiende que puede utilizarse cualquier longitud de fragmento soportada por las secuencias mostradas en el presente documento, en las tablas, las figuras o la lista de secuencias para describir una longitud sobre la cual se puede medir la identidad porcentual.

La expresión "no repetitividad", como se utiliza en el presente documento en el contexto de un polipéptido, se refiere a una ausencia o a un grado limitado de homología interna en una secuencia de péptidos o polipéptidos. La expresión "sustancialmente no repetitivo" puede significar, por ejemplo, que hay pocos o ningún caso de cuatro aminoácidos contiguos en la secuencia que sean tipos de aminoácidos idénticos, o que el polipéptido tiene una puntuación de subsecuencia (definida a continuación) de 10 o menos, o que no hay un patrón en el orden, del extremo amínico al carboxílico, de los motivos de la secuencia que constituyen la secuencia del polipéptido. El término "repetitividad", como se utiliza en el presente documento en el contexto de un polipéptido, se refiere al grado de homología interna en una secuencia de péptidos o polipéptidos. Por el contrario, una secuencia "repetitiva" puede contener múltiples copias idénticas de secuencias de aminoácidos cortas. Por ejemplo, una secuencia de polipéptidos de interés puede dividirse en secuencias de n-meros y se puede contar el número de secuencias idénticas. Las secuencias muy repetitivas contienen una gran fracción de secuencias idénticas, mientras que las secuencias no repetitivas contienen pocas secuencias idénticas. En el contexto de un polipéptido, una secuencia puede contener múltiples copias de secuencias más cortas de una longitud definida o variable, o motivos, en los que los propios motivos tienen secuencias no repetitivas, haciendo que el polipéptido completo sea sustancialmente no repetitivo. La longitud del polipéptido dentro del cual se mide la no repetitividad puede variar de 3 aminoácidos a aproximadamente 200 aminoácidos, aproximadamente de 6 a aproximadamente 50 aminoácidos, o de aproximadamente 9 a aproximadamente 14 aminoácidos. La "repetitividad" utilizada en el contexto de secuencias polinucleotídicas se refiere al grado de homología interna en la secuencia, tal como, por ejemplo, la frecuencia de secuencias nucleotídicas idénticas de una longitud dada. La repetitividad puede medirse, por ejemplo, analizando la frecuencia de secuencias idénticas.

Un "vector" es una molécula de ácido nucleico, preferentemente autorreplicante en un hospedador apropiado, que transfiere una molécula de ácido nucleico insertada dentro y/o entre las células hospedadoras. El término incluye vectores que funcionan principalmente para la inserción de ADN o ARN en una célula, la replicación de vectores que funcionan principalmente para la replicación de ADN o ARN, y vectores de expresión que funcionan para la transcripción y/o la traducción del ADN o el ARN. También se incluyen vectores que proporcionan más de una de las funciones anteriores. Un "vector de expresión" es un polinucleótido que, cuando se introduce en una célula hospedadora apropiada, se puede transcribir y traducir en un polipéptido o polipéptidos. Un "sistema de expresión" normalmente connota una célula hospedadora adecuada que comprende un vector de expresión que puede funcionar para producir un producto de expresión deseado.

"Resistencia a la degradación en suero", como se aplica a un polipéptido, se refiere a la capacidad de los polipéptidos de soportar la degradación en la sangre o en los componentes de la misma, lo que normalmente involucra a las proteasas del suero o el plasma. La resistencia a la degradación en suero se puede medir combinando la proteína con suero o plasma humano (o de ratón, rata, mono, según corresponda), normalmente durante un intervalo de días (por ejemplo, 0,25, 0,5, 1, 2, 4, 8, 16 días), normalmente aproximadamente a 37 °C. Las muestras para estos puntos temporales se pueden analizar en un ensayo de inmunoelectrotransferencia y la proteína se detecta con un anticuerpo. El anticuerpo puede estar en una etiqueta en la proteína. Si la proteína muestra una sola banda en la inmunoelectrotransferencia, donde el tamaño de la proteína es idéntico al de la proteína inyectada, entonces no se ha producido degradación. En este método ilustrativo, el punto temporal en el que se degrada el 50 % de la proteína, según las inmunoelectrotransferencias o técnicas equivalentes, es la semivida de degradación en suero o "semivida sérica" de la proteína.

El término "$t_{1/2}$", como se utiliza en el presente documento, significa la semivida terminal calculada como $\ln(2)/K_{el}$. La K_{el} es la constante de la tasa de eliminación terminal calculada mediante regresión lineal de la porción lineal terminal de la curva del logaritmo de la concentración frente al tiempo. La semivida se refiere normalmente al tiempo requerido para que la mitad de la cantidad de una sustancia administrada depositada en un organismo vivo sea metabolizada o eliminada mediante los procesos biológicos normales. Las expresiones "$t_{1/2}$", "semivida terminal", "semivida de eliminación" y "semivida en circulación" se usan indistintamente en el presente documento.

El "factor de peso molecular aparente" o "peso molecular aparente" son expresiones relacionadas que se refieren a una medida del aumento o disminución relativos en el peso molecular aparente que presenta una secuencia de aminoácidos particular. El factor de peso molecular aparente se determina utilizando una cromatografía de exclusión por tamaño (SEC) y métodos similares mediante la comparación con los patrones de proteínas globulares, y se mide en unidades de "kD aparente". El factor de peso molecular aparente es la relación entre el factor de peso molecular

aparente y el peso molecular real; este último se predice añadiendo, en función de la composición de aminoácidos, el peso molecular calculado de cada tipo de aminoácido en la composición.

El "radio hidrodinámico" o "radio de Stokes" es el radio eficaz (R_h en nm) de una molécula en una solución que se mide asumiendo que es un cuerpo que se mueve a través de la solución y opone resistencia por la viscosidad de la solución. En las realizaciones de la invención, las mediciones del radio hidrodinámico de las proteínas de fusión XTEN se correlacionan con el "factor de peso molecular aparente", que es una medida más intuitiva. El "radio hidrodinámico" de una proteína afecta a su tasa de difusión en solución acuosa, así como a su capacidad de migrar en geles de macromoléculas. El radio hidrodinámico de una proteína está determinado por su peso molecular, así como por su estructura, incluyendo su forma y su compacidad. Los métodos para determinar el radio hidrodinámico son bien conocidos en la técnica, tales como mediante el uso de una cromatografía de exclusión por tamaño (SEC), como se describe en las Patentes de EE. UU. n.º 6.406.632 y 7.294.513. La mayoría de las proteínas tienen una estructura globular, que es la estructura tridimensional más compacta que una proteína puede tener con el radio hidrodinámico más pequeño. Algunas proteínas adoptan una conformación aleatoria y abierta, no estructurada o "lineal" y, como resultado, tienen un radio hidrodinámico mucho mayor en comparación con las proteínas globulares normales con un peso molecular similar.

Las "condiciones fisiológicas" se refieren a un conjunto de condiciones en un hospedador vivo, así como a las condiciones *in vitro*, que incluyen temperatura, concentración salina, pH, que imitan esas condiciones de un sujeto vivo. Se han establecido varias condiciones fisiológicamente relevantes para usar en ensayos *in vitro*. Generalmente, un tampón fisiológico contiene una concentración fisiológica de sal y se ajusta a un pH neutro que varía de aproximadamente 6,5 a aproximadamente 7,8, y preferentemente de aproximadamente 7,0 a aproximadamente 7,5. En Sambrook et al. (1989) se recogen diversos tampones fisiológicos. La temperatura fisiológicamente relevante varía de aproximadamente 25 °C a aproximadamente 38 °C, y preferentemente de aproximadamente 35 °C a aproximadamente 37 °C.

Un "grupo reactivo" es una estructura química que puede acoplarse a un segundo grupo reactivo. Algunos ejemplos de grupos reactivos son grupos amino, grupos carboxilo, grupos sulfhidrilo, grupos hidroxilo, grupos aldehído, grupos azida. Algunos grupos reactivos se pueden activar para facilitar el acoplamiento con un segundo grupo reactivo. Algunos ejemplos no limitativos de activación son la reacción de un grupo carboxilo con carbodiimida, la conversión de un grupo carboxilo en un éster activado o la conversión de un grupo carboxilo en una función azida.

"Agente de liberación controlada", "agente de liberación lenta", "formulación prolongada" o "agente de liberación sostenida" se usan indistintamente para referirse a un agente capaz de prolongar la duración de la liberación de un polipéptido de la invención con respecto a la duración de liberación cuando el polipéptido se administra en ausencia del agente. Diferentes realizaciones de la presente invención pueden tener diferentes tasas de liberación, lo que da como resultado diferentes cantidades terapéuticas.

Las expresiones "antígeno", "antígeno diana" o "inmunógeno" se usan indistintamente en el presente documento para referirse a la estructura o determinante de unión a la que un fragmento de anticuerpo o un fragmento terapéutico basado en un anticuerpo se une o contra la que tiene especificidad.

Las expresiones "unión específica" o "unirse específicamente" se usan indistintamente en el presente documento para referirse al alto grado de afinidad de unión de un resto de direccionamiento o proteína de fusión de unión por su diana correspondiente. Normalmente, la unión específica medida con uno o más de los ensayos divulgados en el presente documento tendría una constante de disociación o K_d menor de aproximadamente 10^{-6} M.

La expresión "carga activa", como se utiliza en el presente documento, se refiere a una secuencia de proteína o péptidos que tiene actividad biológica o terapéutica; el compañero del farmacóforo de moléculas pequeñas. Algunos ejemplos de cargas activas incluyen, pero sin limitación, citocinas, enzimas, hormonas y factores sanguíneos y de crecimiento. Las cargas activas pueden comprender además restos fusionados genéticamente o conjugados químicamente tales como antineoplásicos, compuestos antivíricos, toxinas o agentes de contraste. Estos restos conjugados se pueden unir al resto del polipéptido a través de un conector que puede ser escindible o no escindible.

El término "antagonista", como se utiliza en el presente documento, incluye cualquier molécula que bloquea, inhibe o neutraliza parcial o completamente una actividad biológica de un polipéptido natural divulgado en el presente documento. Los métodos para identificar los antagonistas de un polipéptido pueden comprender poner en contacto un polipéptido natural con una molécula antagonista candidata y medir un cambio detectable en una o más actividades biológicas normalmente asociadas con el polipéptido natural. En el contexto de la presente invención, los antagonistas pueden incluir proteínas, ácidos nucleicos, hidratos de carbono, anticuerpos o cualquier otra molécula que disminuya el efecto de una proteína biológicamente activa.

El término "agonista" se usa en el sentido más amplio e incluye cualquier molécula que imite una actividad biológica de un polipéptido natural divulgado en el presente documento. Algunas moléculas agonistas adecuadas incluyen específicamente anticuerpos o fragmentos de anticuerpos agonistas, fragmentos o variantes de secuencia de aminoácidos de polipéptidos naturales, péptidos, moléculas orgánicas pequeñas, etc. Los métodos para identificar los

agonistas de un polipéptido natural pueden comprender poner en contacto un polipéptido natural con una molécula agonista candidata y medir un cambio detectable en una o más actividades biológicas normalmente asociadas con el polipéptido natural.

"Actividad", para los fines del presente documento, se refiere a una acción o un efecto de un componente de una proteína de fusión coherente con los de la proteína biológicamente activa natural correspondiente, en donde "actividad biológica" se refiere a una función o un efecto biológico *in vitro* o *in vivo* que incluye, pero no sin limitación, la unión al receptor, la actividad antagonista, la actividad agonista o la respuesta celular o fisiológica.

Como se utiliza en el presente documento, "tratamiento" o "tratar" o "paliar" o "mejorar" se usan indistintamente en el presente documento. Estos términos se refieren a un enfoque para obtener unos resultados beneficiosos o deseados que incluyen, pero no sin limitación, un beneficio terapéutico y/o un beneficio profiláctico. Por beneficio terapéutico se entiende la erradicación o la mejora del trastorno subyacente que se está tratando. También se logra un beneficio terapéutico con la erradicación o la mejora de uno o más de los síntomas fisiológicos asociados con el trastorno subyacente, de modo que se observa una mejora en el sujeto, a pesar de que el sujeto todavía puede estar afectado por el trastorno subyacente. Para un beneficio profiláctico, las composiciones pueden administrarse a un sujeto en riesgo de desarrollar una enfermedad particular, o a un sujeto que comunica uno o más de los síntomas fisiológicos de una enfermedad, incluso aunque no se haya hecho un diagnóstico de esta enfermedad.

Un "efecto terapéutico", como se utiliza en el presente documento, se refiere a un efecto fisiológico, pero sin limitación, la cura, mitigación, mejora o prevención de enfermedades en seres humanos u otros animales, o para mejorar de otro modo el bienestar físico o mental de seres humanos o animales, causado por un polipéptido de fusión de la invención, distinto de la capacidad de inducir la producción de un anticuerpo contra un epítopo antigénico poseído por la proteína biológicamente activa. La determinación de una cantidad terapéuticamente eficaz está dentro de la capacidad de los expertos en la materia, especialmente a la luz de la divulgación detallada proporcionada en el presente documento.

Las expresiones "cantidad terapéuticamente eficaz" y "dosis terapéuticamente eficaz", como se utilizan en el presente documento, se refieren a una cantidad de una proteína biológicamente activa, ya sea sola o como parte de una composición de proteína de fusión, que sea capaz de tener cualquier efecto detectable y beneficioso sobre cualquier síntoma, aspecto, parámetro medido o características de un estado patológico o afección cuando se administra en una dosis o en dosis repetidas a un sujeto. Dicho efecto no necesita ser absoluto para ser beneficioso.

La expresión "pauta posológica terapéuticamente eficaz", como se utiliza en el presente documento, se refiere a un cronograma para dosis administradas consecutivamente de una proteína biológicamente activa, ya sea sola o como parte de una composición de proteína de fusión, en donde las dosis se administran en unas cantidades terapéuticamente eficaces para dar como resultado un efecto beneficioso sostenido sobre cualquier síntoma, aspecto, parámetro medido o características de un estado patológico o afección.

I). TÉCNICAS GENERALES

La práctica de la presente descripción emplea, salvo que se indique lo contrario, técnicas convencionales de inmunología, bioquímica, química, biología molecular, microbiología, biología celular, genómica y ADN recombinante, que están dentro de los conocimientos de la técnica. *Véase* Sambrook, J. et al., "Molecular Cloning: A Laboratory Manual", 3ª edición, Cold Spring Harbor Laboratory Press, 2001; "Current protocols in molecular biology", F. M. Ausubel, et al. eds., 1987; la serie "Methods in Enzymology", Academic Press, San Diego, CA.; "PCR 2: a practical approach", M.J. MacPherson, B.D. Hames y G.R. Taylor eds., Oxford University Press, 1995; "Antibodies, a laboratory manual" Harlow, E. y Lane, D. eds., Cold Spring Harbor Laboratory,1988; "Goodman & Gilman's The Pharmacological Basis of Therapeutics", 11ª edición, McGraw-Hill, 2005; y Freshney, R.I., "Culture of Animal Cells: A Manual of Basic Technique", 4ª edición, John Wiley & Sons, Somerset, NJ, 2000.

II). COMPOSICIONES DE PROTEÍNA DE FUSIÓN DE UNIÓN

(a) La presente invención se refiere, en parte, a composiciones de proteína de fusión de unión ("BFP") que comprenden proteínas de fusión aisladas de restos de direccionamiento polipeptídicos unidos a uno o más polipéptidos recombinantes expandidos ("XTEN"). En particular, la invención proporciona composiciones aisladas de proteína de fusión de unión útiles en el tratamiento de enfermedades, trastornos o afecciones en las que el resto de direccionamiento puede estar dirigido a un antígeno, un ligando o un receptor implicado en, asociado con, o que modula una enfermedad, un trastorno o una afección, mientras que la porción portadora del XTEN puede estar diseñada para conferir una semivida deseada o una propiedad farmacéutica mejorada a la proteína de fusión de unión, como se describe más completamente a continuación. Las proteínas de fusión de unión de la presente invención pueden actuar como agonistas o como antagonistas. En una realización, la composición puede comprender además un segundo resto de direccionamiento o múltiples restos de direccionamiento que pueden tener afinidad de unión por la misma diana o una diana o diferente, lo que da como resultado proteínas de fusión de unión biespecíficas o multivalentes. La invención proporciona varias formas y configuraciones diferentes de restos de direccionamiento y XTEN. Las proteínas

de fusión de unión de las realizaciones divulgadas en el presente documento presentan una o más o cualquier combinación de las propiedades y/o las realizaciones como se detalla en el presente documento.

(b) Dianas

En general, los restos de direccionamiento de las presentes composiciones de proteína de fusión de unión presentan una especificidad de unión a una diana dada u otra característica biológica deseada cuando se utilizan *in vivo* o cuando se utilizan en un ensayo *in vitro*. Las presentes proteínas de fusión de unión que comprenden dos o más restos de direccionamiento pueden estar diseñadas para unirse a la misma diana, a diferentes epítopos en la misma diana o a diferentes dianas mediante la incorporación selectiva de restos de direccionamiento con afinidad de unión por los respectivos sitios de unión.

Las dianas a los que se pueden dirigir los presentes restos de direccionamiento de las composiciones de proteínas de fusión de unión incluyen citocinas, proteínas relacionadas con citocinas, receptores de citocinas, quimiocinas, receptores de quimiocinas, receptores o antígenos de la superficie celular, hormonas o proteínas o péptidos circulantes similares, oligonucleótidos o sustratos enzimáticos. Las dianas están generalmente asociadas con una enfermedad, un trastorno o una afección. Como se utiliza en el presente documento, "una diana asociada con una enfermedad, un trastorno o una afección" significa que la diana es expresada o sobreexpresada por células o tejidos patológicos, la diana causa o es un mediador o es un subproducto de la enfermedad, el trastorno o la afección, o la diana se encuentra generalmente en unas concentraciones más altas en un sujeto con la enfermedad, el trastorno o la afección, o la diana se encuentra en unas concentraciones más altas que las iniciales dentro o cerca de las áreas de la enfermedad, el trastorno o la afección en el sujeto. Un ejemplo no limitativo de lo anterior es la diana HER2, que está implicado en aproximadamente un 30 por ciento de los cánceres de mama debido a una amplificación del gen *HER2/neu* o a la sobreexpresión de su producto proteico. La sobreexpresión del receptor HER2 en cáncer de mama se asocia con una mayor recaída de la enfermedad y un peor pronóstico, y se usa un anticuerpo anti-Her2/neu humanizado en el tratamiento de los cánceres de mama que expresan el receptor HER2 (véase, por ejemplo, la Pat. de EE. UU. N.º 4.753.894).

Descritos en el presente documento, el uno o más restos de direccionamiento pueden tener afinidad de unión por las dianas seleccionadas entre, pero sin limitación, las dianas de la tabla 1.

Tabla 1: Dianas de los restos de direccionamiento

Diana
ABCF1; ACVR1; ACVR1B; ACVR2; ACVR2B; ACVRL1; ADORA2A; Agrecan; AGR2; AICDA; AIF1; AIG1; AKAP1; AKAP2; AMH; AMHR2; ANGPT1; ANGPT2; ANGPTL3; ANGPTL4; ANPEP; APC; APOC1; APRIL; AR; AZGP1 (cinc-a- glicoproteína); A4 integrina; B7; B7.1; B7.2; BAD; BAFF; BAG1; BAI1; BCL2; BCL6; BDNF; BLNK; BLR1 (MDR15); BlyS; BMP1; BMP2; BMP3B (GDF10); BMP4; BMP6; BMP8; BMPR1A; BMPR1B; BMPR2; BPAG1 (plectina); BRCA1; C19orf10 (IL27w); C3; C4A; C5; C5R1; CANT1; CASP1; CASP4; CAV1; CCBP2 (D6/JAB61); CCL1 (1-309); CCL11 (eotaxina); CCL13 (MCP-4); CCL15 (MIP-1d); CCL16 (HCC-4); CCL17 (TARC); CCL18 (PARC); CCL19 (MIP-3b); CCL2 (MCP-1); MCAF; CCL20 (MIP-3a); CCL21 (MIP-2); SLC; exodus-2; CCL22 (MDC/STC-1); CCL23 (MPIF-1); CCL24 (MPIF-2/eotaxina-2); CCL25 (TECK); CCL26 (eotaxina-3); CCL27 (CTACK/ILC); CCL28; CCL3 (MIP-1a); CCL4 (MIP-1b); CCL5 (RANTES); CCL7 (MCP-3); CCL8 (mcp-2); CCNA1; CCNA2; CCND1; CCNE1; CCNE2; CCR1 (CKR1/HM145); CCR2 (mcp-1RB/RA); CCR3 (CKR3/CMKBR3); CCR4; CCR5 (CMKBR5/ChemR13); CCR6 (CMKBR6/CKR-L3/STRL22/DRY6); CCR7 (CKR7/EBI1); CCR8 (CMKBR8/TER1/CKR-L1); CCR9 (GPR-9-6); CCRL1 (VSHK1); CCRL2 (L-CCR); CD164; CD19; CD1C; CD20; CD200; CD-22; CD24; CD28; CD3; CD37; CD38; CD3E; CD3G; CD3Z; CD4; CD11a (LFA-1 integrina alfaL); CD40; CD40L; CD44; CD45RB; CD52; CD69; CD72; CD74; CD79A; CD79B; CD8; CD80; CD81; CD83; CD86; CD340; CDH1 (E-cadherina); CDH10; CDH12; CDH13; CDH18; CDH19; CDH20; CDH5; CDH7; CDH8; CDH9; CDK2; CDK3; CDK4; CDK5; CDK6; CDK7; CDK9; CDKN1A (p21Wap1/Cip1); CDKN1B (p27Kip1); CDKN1C; CDKN2A (p16INK4a); CDKN2B; CDKN2C; CDKN3; CEBPB; CER1; CHGA; CHGB; quitinasa; CHST10; CKLFSF2; CKLFSF3; CKLFSF4; CKLFSF5; CKLFSF6; CKLFSF7; CKLFSF8; CLDN3; CLDN7 (claudina-7); CLN3; CLU (clusterina); cMET; CMKLR1; CMKOR1 (RDC1); CNR1; COL18A1; COL1A1; COL4A3; COL6A1; CR2; CRP; CSF1 (M-CSF); CSF2 (GM-CSF); CSF3 (GCSF); CTLA4; CTNNB1 (b-catenina); CTSB (catepsina B); CX3CL1 (SCYD1); CX3CR1 (V28); CXCL1 (GRO1); CXCL10(IP-10); CXCL11 (I-TAC/IP-9); CXCL12 (SDF1); CXCL13; CXCL14; CXCL16; CXCL2 (GRO2); CXCL3 (GRO3); CXCL5 (ENA-78/LIX); CXCL6 (GCP-2); CXCL9 (MIG); CXCR3 (GPR9/CKR-L2); CXCR4; CXCR6 (TYMSTR/STRL33/Bonzo); CYB5; CYC1; CYSLTR1; DAB2IP; DES; DKFZp451J0118; DNCL1; DPP4; E2F1; ECGF1; EDG1; EFNA1; EFNA3; EFNB2; EGF; EGFR; ELAC2; elastasa; ENG; ENO1; ENO2; ENO3; EPHB4; EPO; ERBB-2 (Her2); EREG; ERK8; ESR1; ESR2; F3 (TF); FADD; FasL; FASN; FCER1A; FCER2; FCGR3A; FGF; FGF1 (aFGF); FGF10; FGF11; FGF12; FGF12B; FGF13; FGF14; FGF16; FGF17; FGF18; FGF19; FGF2 (bFGF); FGF20;

Diana
FGF21; FGF22; FGF23; FGF3 (int-2); FGF4 (HST); FGF5; FGF6 (HST-2); FGF7 (KGF); FGF8; FGF9; FGFR3; FIGF (VEGFD); FIL1 (ÉPSILON); FIL1 (ZETA); FLJ12584; FLJ25530; FLRT1 (fibronectina); FLT1; FOS; FOSL1 (FRA-1); FY (DARC); GABRP (GABAa); GAGEB1; GAGEC1; GALNAC4S-6ST; GATA3; GDF5; GFI1; GGT1; GM-CSF; GNAS1; GNRH1; GPR2 (CCR10); GPR31; GPR44; GPR81 (FKSG80); GRCC10 (C10); GRP; GSN (Gelsolina); GSTP1; HAVCR2; HDAC4; HDAC5; HDAC7A; HDAC9; HER2; HGF; HIF1A; HIP1; histamina y receptores de histamina; HLA-A; HLA-DRA; HM74; HMOX1; HUMCYT2A; ICEBERG; ICOSL; ID2; IFN-a; IFNA1; IFNA2; IFNA4; IFNA5; IFNA6; IFNA7; IFNB1; IFNgamma; IFNW1; IGBP1; IGF1; IGF1R; IGF2; IGFBP2; IGFBP3; IGFBP6; IL-1; IL10; IL10RA; IL10RB; IL11; IL11RA; IL-12; IL12A; IL12B; IL12RB1; IL12RB2; IL13; IL13RA1; IL13RA2; IL14; IL15; IL15RA; IL16; IL17; IL17B; IL17C; IL17R; IL18; IL18BP; IL18R1; IL18RAP; IL19; IL1A; IL1B; IL1F10; IL1F5; IL1F6; IL1F7; IL1F8; IL1F9; IL1HY1; IL1R1; IL1R2; IL1RAP; IL1RAPL1; IL1RAPL2; IL1RL1; IL1RL2; IL1RN; IL2; IL20; IL20RA; IL21R; IL22; IL22R; IL22RA2; IL23; IL24; IL25; IL26; IL27; IL28A; IL28B; IL29; IL2RA; IL2RB; IL2RG; IL3; IL30; IL3RA; IL4; IL4R; IL5; IL5RA; IL6; IL6R; IL6ST (glicoproteína 130); IL7; IL7R; IL8; IL8RA; IL8RB; IL8RB; IL9; IL9R; ILK; INHA; INHBA; INSL3; INSL4; IRAK1; IRAK2; ITGA1; ITGA2; ITGA3; ITGA6 (integrina a6); ITGAV; ITGB3; ITGB4 (integrina b 4); JAG1; JAK1; JAK3; JUN; K6HF; KAI1; KDR; KITLG; KLF5 (GC Box BP); KLF6; KLK10; KLK12; KLK13; KLK14; KLK15; KLK3; KLK4; KLK5; KLK6; KLK9; KRT1; KRT19 (queratina 19); KRT2A; KRTHB6 (queratina específica del cabello de tipo II); LAMA5; LEP (leptina); LFA3; LIGHT; Lingo-p75; Lingo-Troy; LPS; LTA (TNF-b); LTB; LTB4R (GPR16); LTB4R2; LTBR; MACMARCKS; MAG u Omgp; MAP2K7 (c-Jun); MDK; MIB1; midcina; MIF; MIP-2; MKI67 (Ki-67); MMP2; MMP9; MS4A1; MSMB; MT3 (metalotionectina-III); MTSS1; MUC1 (mucina); MYC; MYD88; NCK2; neurocán; NFKB1; NFKB2; NGFB (NGF); NGFR; NgR-Lingo; NgR-Nogo66 (Nogo); NgR-p75; NgR-Troy; NME1 (NM23A); NOX5; NPPB; NR0B1; NR0B2; NR1D1; NR1D2; NR1H2; NR1H3; NR1H4; NRII2; NRII3; NR2C1; NR2C2; NR2E1; NR2E3; NR2F1; NR2F2; NR2F6; NR3C1; NR3C2; NR4A1; NR4A2; NR4A3; NR5A1; NR5A2; NR6A1; NRP1; NRP2; NT5E; NTN4; ODZ1; OPRD1; P2RX7; PAP; PART1; PATE; PAWR; PCA3; PCNA; PDGFA; PDGFB; PECAM1; PF4 (CXCL4); PGF; PGR; fosfacan; PIAS2; PIK3CG; PLAU (uPA); PLG; PLXDC1; PPBP (CXCL7); PPID; PR1; PRKCQ; PRKD1; PRL; PROC; PROK2; PSAP; PSCA; PTAFR; PTEN; PTGS2 (COX-2); PTN; RAC2 (p21Rac2); RANKL; RARB; RGS1; RGS13; RGS3; RNF110 (ZNF144); ROBO2; VRS; SI00A2; SCGB1D2 (lipofilina B); SCGB2A1 (mamaglobina 2); SCGB2A2 (mamaglobina 1); SCYE1 (citocina endotelial activadora de monocitos); SDF2; SERPINA1; SERPINA3; SERPINB5 (maspina); SERPINE1 (PAI-1); SERPINF1; SHBG; SLA2; SLC2A2; SLC33A1; SLC43A1; SLIT2; SPP1; SPRR1B (Spr1); ST6GAL1; STAB1; STAT6; STEAP; STEAP2; TB4R2; TBX21; TCP10; TDGF1; TEK; TGFA; TGFB1; TGFB111; TGFB2; TGFB3; TGFBI; TGFBR1; TGFBR2; TGFBR3; TH1L; THBS1 (tromboespondina-1); THBS2; THBS4; THPO; TIE (Tie-1); TIMP3; factor tisular; TLR10; TLR2; TLR3; TLR4; TLR5; TLR6; TLR7; TLR8; TLR9; TNF; TNF-a; TNFAIP2 (B94); TNFAIP3; TNFRSF11A; TNFRSF1A; TNFRSF1B; TNFRSF21; T'NFRSFS; TNFRSF6 (Fas); TNFRSF7; TNFRSF8; TNFRSF9; TNFSF10 (TRAIL); TNFSF11 (TRANCE); TNFSF12 (APO3L); TNFSF13 (April); TNFSF13B; TNFSF14 (HVEM-L); TNFSF15 (VEGI); TNFSF18; TNFSF4 (ligando OX40); TNFSF5 (ligando CD40); TNFSF6 (FasL); TNFSF7 (ligando CD27); TNFSF8 (ligando CD30); TNFSF9 (ligando 4-1BB); TOLLIP; receptores de tipo Toll (TLR1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 to TLR-13); TOP2A (topoisomerasa Iia); TP53; TPM1; TPM2; TRADD; TRAF1; TRAF2; TRAF3; TRAF4; TRAF5; TRAF6; TREM1; TREM2; TRPC6; TSLP; TWEAK; VAP1; VEGF; VEGFB; VEGFC; versicán; VHL C5; VLA-1; VLA-4; XCL1 (linfotactina); XCL2 (SCM-1b); XCR1 (GPR5/CCXCR1); YY1; ZFPM2.

Descritos en el presente documento, el uno o más restos de direccionamiento pueden tener afinidad de unión por uno o más antígenos asociados a tumores (TAA) que se sabe que son expresados en las células tumorales o cancerosas o están asociados de otro modo con tumores o cánceres. Los antígenos asociados a tumores se conocen en la técnica, y generalmente se consideran dianas celulares eficaces para el diagnóstico y la terapia del cáncer. En particular, los investigadores han tratado de identificar los TAA que son expresados específicamente en la superficie de uno o más tipos particulares de células cancerosas en comparación con una o más células normales no cancerosas, y ha dado lugar a la capacidad de dirigirse específicamente a las células cancerosas para su destrucción a través de terapias basadas en anticuerpos. Algunos ejemplos no limitativos de TAA se enumeran en la tabla 2. Descrita en el presente documento, una proteína de fusión de unión comprende un resto de direccionamiento con afinidad de unión por una diana de TAA seleccionado de la tabla 2. Descrita en el presente documento, la proteína de fusión de unión comprende dos o más restos de direccionamiento que tienen afinidad de unión por una o más dianas de TAA seleccionados entre las dianas de la tabla 2.

Tabla 2: Dianas de antígenos asociados a tumores

Dianas de TAA (sinónimos)	Número de registro y referencias
Her2 (ErbB2)	N.º de registro en el GenBank M11730; Pat. de EE. UU. N.º 5.869.445; documento WO2004048938; documento WO2004027049; documento WO2004009622; documento WO2003081210; documento WO2003089904; documento WO2003016475; documento US2003118592; documento WO2003008537; documento WO2003055439; documento WO2003025228; documento WO200222636; documento WO200212341; documento WO200213847; documento WO200214503; documento WO200153463; documento WO200141787; documento WO200044899; documento WO200020579; documento WO9630514; documento EP1439393; WO2004043361; documento WO2004022709; documento WO200100244
BMPR1B (receptor de la proteína morfogenética ósea de tipo IB)	N.º de registro en el GenBank NM_001203; documento WO2004063362; documento WO2003042661; documento US 2003134790; documento WO2002102235; documento WO2003055443; documento WO200299122; documento WO2003029421; documento WO2003024392; documento WO200298358' documento WO200254940; documento WO200259377; documento WO200230268
E16 (LAT1, SLC7A5)	N.º de registro en el GenBank NM_003486); documento WO2004048938; documento WO2004032842; documento WO2003042661; documento WO2003016475; documento WO200278524; documento WO200299074; documento WO200286443; documento WO2003003906; documento WO200264798; documento WO200014228; documento US2003224454; documento WO2003025138
STEAP1 (antígeno epitelial de seis transmembranarios de la próstata)	N.º de registro en el GenBank NM_012449; documento WO2004065577; documento WO2004027049; documento EP1394274; documento WO2004016225; documento WO2003042661; documento US2003157089; documento US2003185830; documento US2003064397; documento WO200289747618; documento WO2003022995
STEAP2 (antígeno epitelial de seis transmembranarios de la próstata 2)	N.º de registro en el GenBank AF455138; documento WO2003087306; documento US2003064397; documento WO200272596; documento WO200172962; documento WO2003104270; documento WO2003104270; documento US2004005598; documento WO2003042661; documento US2003060612; documento WO200226822; documento WO200216429
CA125/0772P (MUC16)	N.º de registro en el GenBank AF361486; documento WO2004045553; documento WO200292836; documento WO200283866; documento US2003124140
factor potenciador de megacariocitos (MPF, mesotelina)	N.º de registro en el GenBank NM_005823; documento WO2003101283; documento WO2002102235; documento WO2002101075; documento WO200271928; documento WO9410312
cotransportador NalPi de tipo IIb (NaPi3b)	N.º de registro en el GenBank NM_006424; documento WO2004022778; documento EP1394274; documento WO2002102235; documento EP875569; documento WO200157188; documento WO2004032842; documento WO200175177
Semaforina 5b (SEMA5B, SEMAG)	N.º de registro en el GenBank AB040878; documento WO2004000997; documento WO2003003984; documento WO200206339; documento WO200188133; documento WO2003054152; documento WO2003101400
Antígeno de células madre de cáncer de próstata (PSCA hlg)	N.º de registro en el GenBank AY358628; documento US2003129192; documento US2004044180; documento US2004044179; documento US2003096961; documento US2003232056; documento WO2003105758; documento US2003206918; documento EP1347046; documento WO2003025148
ETBR (receptor de endotelina de tipo B)	N.º de registro en el GenBank AY275463; documento WO2004045516; documento WO2004048938; documento WO2004040000; documento WO2003087768; documento WO2003016475; documento WO2003016475; documento WO200261087; documento WO2003016494; documento WO2003025138; documento WO200198351; documento EP522868; documento WO200177172; documento US2003109676; Pat. de EE. UU. N.º 6.518.404; Pat. de EE. UU. N.º 5.773.223; documento WO2004001004
TRPV4 (canal catiónico potencial receptor temporal, subfamilia V)	Sol. de Pat. de EE. UU. N.º 20090208514
CDC45L	N.º de registro en el GenBank AJ223728; Sol. de Pat. de EE. UU. N.º 20090208514

Dianas de TAA (sinónimos)	Número de registro y referencias
CRIPTO (CR, CR1, CRGF)	N.º de registro en el GenBank NP_003203 o NM_003212; documento US2003224411; documento WO2003083041; documento WO2003034984; documento WO200288170; documento WO2003024392; documento WO200216413; documento WO200222808; documento U.S. 5,854,399; Pat. de EE. UU. N.º 5.792.616
CD21 (CR2 (receptor del complemento 2) o C3DR (C3d/receptor del virus de Epstein Barr)	N.º de registro en el GenBank M26004; documento WO2004045520; documento US2004005538; documento WO2003062401; documento WO2004045520; documento WO9102536; documento WO2004020595
CD79b (CD79B, CD79β, IGB (beta asociada a inmunoglobulina), B29)	N.º de registro en el GenBank NM_000626 o 11038674; documento WO2004016225; documento WO2003087768; documento US2004101874; documento WO2003062401; documento WO200278524; documento US2002150573; Pat. de EE. UU. N.º 5.644.033; documento WO2003048202; documento WO 99/558658, Pat. de EE. UU. N.º 6.534.482; documento WO200055351
FcRH2 (IFGP4, IRTA4, SPAP1A (dominio SH2 que contiene la proteína de anclaje de la fosfatasa 1a), SPAP1B, SPAP1C)	N.º de registro en el GenBank NM_030764, AY358130; documento WO2004016225; documento WO2003077836; documento WO200138490; documento WO2003097803; documento WO2003089624
NCA (CEACAM6)	N.º de registro en el GenBank M18728; documento WO2004063709; documento EP1439393; documento WO2004044178; documento WO2004031238; documento WO2003042661; documento WO200278524; documento WO200286443; documento WO200260317
MDP (DPEP1)	N.º de registro en el GenBank BC017023; documento WO2003016475; documento WO200264798
IL20Rα (IL20Ra, ZCYTOR7)	N.º de registro en el GenBank AF184971; documento EP1394274; documento US2004005320; documento WO2003029262; documento WO2003002717; documento WO200222153; documento US2002042366; documento WO200146261; documento WO200146232; documento WO9837193
BECAN (proteína nuclear de Brevican)	N.º de registro en el GenBank AF229053; documento US2003186372; documento US2003186373; documento US2003119131; documento US2003119122; documento US2003119126; documento US2003119121; documento US2003119129; documento US2003119130; documento US2003119128; documento US2003119125; documento WO2003016475; documento WO200202634
EphB2R (DRT, ERK, Hek5, EPHT3, Tyro5)	N.º de registro en el GenBank NM_004442; documento WO2003042661; documento WO200053216; documento WO2004065576 (reivindicación 1); documento WO2004020583; documento WO2003004529; documento WO200053216
B7h (ASLG659)	N.º de registro en el GenBank AX092328; documento US20040101899; documento WO2003104399; documento WO2004000221; documento US2003165504; documento US2003124140; documento US2003065143; documento WO2002102235; documento US2003091580; documento WO200210187; documento WO200194641; documento WO200202624; documento US2002034749; documento WO200206317; documento WO200271928; documento WO200202587; documento WO200140269; documento WO200036107; documento WO2004053079; documento WO2003004989; documento WO200271928
PSCA (precursor del antígeno de células madre de próstata	N.º de registro en el GenBank AJ297436; documento WO2004022709; documento EP1394274; documento US2004018553; documento WO2003008537 (reivindicación 1); documento WO200281646; documento WO2003003906; documento WO200140309; documento US2001055751; documento WO200032752; documento WO9851805; documento WO9851824; documento WO9840403
BAFF-R (receptor del factor activador de linfocitos B, receptor BLyS 3, BR3)	N.º de registro en el GenBank AF116456; documento WO2004058309; documento WO2004011611; documento WO2003045422; documento WO2003014294; documento WO2003035846; documento WO200294852; documento WO200238766; documento WO200224909
CD22 (forma CD22-β del receptor de linfocitos B, BL-CAM, Lyb-8, Lyb8, SIGLEC-2, FLJ22814)	N.º de registro en el GenBank AK026467; documento WO2003072036

Dianas de TAA (sinónimos)	Número de registro y referencias
CD79a (alfa asociada a inmunoglobulina)	N.º de registro en el GenBank NP_001774.10; documento WO2003088808, documento US20030228319; documento WO2003062401; documento US2002150573; documento WO9958658; documento WO9207574; Pat. de EE. UU. N.º 5.644.033
CXCR5 (receptor 1 del linfoma de Burkitt)	N.º de registro en el GenBank NP_001707.1; documento WO2004040000; documento WO2004015426; documento US2003105292; Pat. de EE. UU. N.º 6.555.339; documento WO200261087; documento WO200157188; documento WO200172830; documento WO200022129; documento WO9928468; Pat. de EE. UU. N.º 5.440.021; documento WO9428931; documento WO9217497
HLA-DOB	N.º de registro en el GenBank NP_002111.1; documento WO9958658; Pat. de EE. UU. N.º 6.153.408; Pat. de EE. UU. N.º 5.976.551; Pat. de EE. UU. N.º 6.011.146
P2X5	N.º de registro en el GenBank NP_002552.2; documento WO2004047749; documento WO2003072035; documento WO200222660; documento WO2003093444; documento WO2003087768; documento WO2003029277
CD72 (antígeno de diferenciación de linfocitos B CD72, Lyb-2)	N.º de registro en el GenBank NP_001773.1; documento WO2004042346; documento WO2003026493; documento WO200075655
CD180 (LY64)	N.º de registro en el GenBank NP_005573.1; documento US2002193567; documento WO9707198; documento WO2003083047; documento WO9744452
FcRH1 (proteína similar al receptor Fc 1)	N.º de registro en el GenBank NP_443170.1) documento WO2003077836; documento WO200138490; documento WO2003089624; documento EP1347046; documento WO2003089624
IRTA2 (translocación del receptor de la superfamilia de las inmunoglobulinas asociada 2)	N.º de registro en el GenBank Humanas:AF343662, AF343663, AF343664, AF343665, AF369794, AF397453; documento WO2003024392; documento WO2003077836; documento WO200138490
TENB2 (TMEFF2, tomorregulina, TPEF, HPP1)	N.º de registro en el GenBank AF179274; AY358907, CAF85723, CQ782436; documento WO2004074320; documento WO2003042661; documento WO2003009814; documento EP1295944; documento WO200230268; documento WO200190304; documento US2004249130; documento US2004022727; documento WO2004063355; documento US2004197325; documento US2003232350; documento US2004005563; documento US2003124579
CS1 (CRACC, 19A, APEX-1, FOAP12)	N.º de registro en el GenBank NM 021181; documento US 20100168397
DLL4	N.º de registro en el GenBank NM 019074; documento US 20100303812
Lewis Y	ADB235860; documento US 7879983
CD40 (Bp50, CDW40, MGC9013, TNFRSF5, p50)	AL035662.65; documento US 6946129
OBA1 (5T4)	N.º de registro en el GenBank NP_001159864.1; documento US 20100021483
p97	Woodbury et al., 1980, Proc. Natl. Acad. Sci. USA 77: 2183-2186; Brown et al., 1981, J. Immunol. 127: 539- 546
antígeno carcinoembrionario (ACE)	N.º de registro en el GenBank NP_004354.2; documento US 6676924
TAG-72	documento US 7256004
ADN	
Neuropilina -1 (NRP1)	N.º de registro en el GenBank NP_001019799.1; documento US 20080213268
A33	N.º de registro en el GenBank NP_005805.1; documento US 7579187
Mucina-1 (MUC1)	N.º de registro en el GenBank NP_001018016.1; NP_001018017.1; documento US 7183388
Fibronectina ED-B	documento US 7785591

Dianas de TAA (sinónimos)	Número de registro y referencias
Antígeno de Thomsen-Friedenreich (TF)	documento US 7374755; documento US 20100297159

(c) Polipéptidos recombinantes expandidos

En el presente documento se describen composiciones de polipéptidos XTEN que son útiles como compañero de una proteína de fusión a los que se pueden unir uno o más restos de direccionamiento, lo que da como resultado una proteína de fusión de unión. Los XTEN son generalmente polipéptidos de longitud expandida con secuencias no naturales sustancialmente no repetitivas que están compuestas principalmente por aminoácidos hidrófilos pequeños, teniendo la secuencia un bajo grado o ninguna estructura secundaria ni terciaria en condiciones fisiológicas. Los XTEN tienen utilidad como compañeros de proteínas de fusión, ya que desempeñan diversas funciones, confiriendo ciertas propiedades farmacocinéticas, fisicoquímicas y farmacéuticas deseables, entre otras propiedades descritas a continuación, cuando se unen a un resto de direccionamiento para crear una proteína de fusión.

En algunas realizaciones, los XTEN son polipéptidos largos que tienen más de 100 hasta aproximadamente 3000 restos, y preferentemente de 400 a aproximadamente 3000 restos cuando se utilizan como portador, o acumulativamente cuando se utiliza más de una unidad XTEN en una única proteína de fusión con un resto de direccionamiento; por ejemplo, un conector y un portador o un XTEN aminoterminal y un portador. En otras realizaciones se pueden utilizar secuencias XTEN más cortas como conectores para unir componentes de las proteínas de fusión de unión o para mejorar la expresión en forma de un XTEN aminoterminal, como se describe más completamente a continuación.

Los criterios de selección del XTEN utilizado para crear las composiciones inventivas generalmente se refieren a atributos de propiedades físicas/químicas y estructura conformacional del XTEN que pueden usarse, a su vez, para conferir propiedades farmacéuticas y farmacocinéticas mejoradas a las composiciones. El XTEN de la presente invención puede presentar una o más de las siguientes propiedades ventajosas: flexibilidad conformacional, solubilidad acuosa mejorada, alto grado de resistencia a proteasas, baja inmunogenicidad, baja unión a receptores de mamífero y mayores radios hidrodinámicos (o de Stokes); propiedades que pueden hacerlos particularmente útiles como compañeros de proteínas de fusión y armazones para conjugados de fármacos. Algunos ejemplos no limitativos de las propiedades de las composiciones inventivas que pueden potenciarse mediante un XTEN incluyen aumentos en la solubilidad global y/o en la estabilidad metabólica, sensibilidad reducida a la proteólisis, inmunogenicidad reducida, tasa reducida de absorción cuando se administra por vía subcutánea o intramuscular, y unas propiedades farmacocinéticas mejoradas, tales como semivida terminal más larga y área bajo la curva (AUC) aumentada, absorción más lenta después de la inyección subcutánea o intramuscular (en comparación con agentes no unidos a XTEN administrados por una vía parenteral) de tal manera que la $C_{máx}$ es menor, lo que puede, a su vez, dar como resultado reducciones en los efectos adversos que, en conjunto, pueden dar como resultado un período de tiempo aumentado en el que una composición de proteína de fusión administrada a un sujeto conserva la actividad terapéutica.

En la técnica se conocen varios métodos y ensayos para determinar las propiedades físicas/químicas de las proteínas, tales como las composiciones que comprenden el XTEN inventivo; propiedades tales como la estructura secundaria o terciaria, la solubilidad, la agregación de proteínas, las propiedades de fusión, la contaminación y el contenido de agua. Dichos métodos incluyen centrifugación analítica, EPR, HPLC con intercambio iónico, HPLC de exclusión por tamaño, HPLC en fase inversa, dispersión de la luz, electroforesis capilar, dicroísmo circular, calorimetría diferencial de barrido, fluorescencia, HPLC con intercambio iónico, HPLC de exclusión por tamaño, IR, RMN, espectroscopia Raman, refractometría y espectroscopia UV/visible. Algunos métodos adicionales se divulgan en Arnau et al, Prot expr and Purif (2006) 48, 1-13. La aplicación de estos métodos a la invención estaría al alcance de una persona experta en la materia.

En una realización, los XTEN están diseñados para comportarse como secuencias peptídicas desnaturalizadas en condiciones fisiológicas, a pesar de la longitud expandida del polímero. Desnaturalizado describe el estado de un péptido en solución que se caracteriza por una gran libertad conformacional del esqueleto peptídico. La mayoría de los péptidos y las proteínas adoptan una conformación desnaturalizada en presencia de altas concentraciones de agentes desnaturalizantes o a una temperatura elevada. Los péptidos en la conformación desnaturalizada tienen, por ejemplo, unos espectros de dicroísmo circular (CD) característicos y se caracterizan por la ausencia de interacciones de largo alcance, según se determina mediante RMN. "Conformación desnaturalizada" y "conformación no estructurada" se usan como sinónimos en el presente documento. En una realización, la invención proporciona secuencias XTEN que, en condiciones fisiológicas, pueden parecerse a secuencias desnaturalizadas en gran parte desprovistas de estructura secundaria. En una realización, las secuencias XTEN pueden estar sustancialmente desprovistas de estructura secundaria en condiciones fisiológicas. "Muy desprovisto", como se utiliza en este contexto, significa que menos del 50 % de los restos de aminoácidos XTEN de la secuencia del XTEN contribuyen a una estructura secundaria medida o determinada por los medios descritos en el presente documento. "Sustancialmente desprovisto", como se utiliza en este contexto, significa que al menos aproximadamente un 60 %, o aproximadamente

un 70 %, o aproximadamente un 80 %, o aproximadamente un 90 %, o aproximadamente un 95 %, o al menos aproximadamente un 99 % de los restos de aminoácidos XTEN de la secuencia del XTEN no contribuyen a la estructura secundaria, medida o determinada por los medios descritos en el presente documento.

Se han establecido varios métodos en la técnica para discernir la presencia o ausencia de estructuras secundarias y terciarias en un polipéptido dado. En particular, la estructura secundaria se puede medir espectrofotométricamente, por ejemplo, mediante espectroscopia de dicroísmo circular en la región espectral de "UV lejano" (190-250 nm). Cada uno de los elementos estructurales secundarios, tales como la hélice alfa y la lámina beta, da lugar a una forma y una magnitud características de los espectros de CD. También se puede predecir una estructura secundaria para una secuencia polipeptídica a través de ciertos programas informáticos o algoritmos, tales como el conocido algoritmo de Chou-Fasman (Chou, P. Y., et al (1974) Biochemistry, 13: 222-45) y el algoritmo de Garnier-Osguthorpe-Robson ("GOR") (Garnier J, Gibrat JF, Robson B. (1996), GOR method for predicting protein secondary structure from amino acid sequence. Methods Enzymol 266:540-553), como se describe en la Publicación de Solicitud de Patente de EE. UU. n.º 20030228309A1. Para una secuencia dada, los algoritmos pueden predecir si existe alguna o ninguna estructura secundaria en absoluto, expresada como el total y/o el porcentaje de restos de la secuencia que forman, por ejemplo, hélices alfa o láminas beta, o el porcentaje de restos de la secuencia que se predice que darán como resultado la formación aleatoria de espirales (que carecen de estructura secundaria).

En una realización, las secuencias XTEN utilizadas en las presentes composiciones de proteína de fusión tienen un porcentaje de hélices alfa que varía del 0 % a menos de aproximadamente un 5 %, según se determina mediante el algoritmo de Chou-Fasman. En otra realización, las secuencias XTEN de las composiciones de proteína de fusión tienen un porcentaje de láminas beta que varía del 0 % a menos de aproximadamente un 5 % según se determina mediante el algoritmo de Chou-Fasman. En algunas realizaciones, las secuencias XTEN de las composiciones de proteínas de fusión tienen un porcentaje de hélices alfa que varía del 0 % a menos de aproximadamente un 5 %, y un porcentaje de láminas beta que varía del 0 % a menos de aproximadamente un 5 %, según se determina mediante el algoritmo de Chou-Fasman. En una realización, las secuencias XTEN de las composiciones de proteína de fusión tienen un porcentaje de hélices alfa menor de aproximadamente un 2 % y un porcentaje de láminas beta menor de aproximadamente un 2 %. Las secuencias XTEN de las composiciones de la proteína de fusión tienen un alto grado de porcentaje de espirales aleatorias, según se determina mediante el algoritmo de GOR. En algunas realizaciones, una secuencia del XTEN tiene al menos aproximadamente un 80 %, más preferentemente al menos aproximadamente un 90 %, más preferentemente al menos aproximadamente un 91 %, más preferentemente al menos aproximadamente un 92 %, más preferentemente al menos aproximadamente un 93 %, más preferentemente al menos aproximadamente un 94 %, más preferentemente al menos aproximadamente un 95 %, más preferentemente al menos aproximadamente un 96 %, más preferentemente al menos aproximadamente un 97 %, más preferentemente al menos aproximadamente un 98 % y lo más preferentemente al menos aproximadamente un 99 % de espirales aleatorias, según se determina mediante el algoritmo de GOR. En una realización, las secuencias XTEN de las composiciones de proteína de fusión tienen un porcentaje de hélices alfa que varía del 0 % a menos de aproximadamente un 5 % y un porcentaje de lámina beta que varía del 0 % a menos de aproximadamente un 5 % según se determina mediante el algoritmo de Chou-Fasman y al menos aproximadamente un 90 % de espiral aleatoria, según se determina mediante el algoritmo de GOR. En otra realización, las secuencias XTEN de las composiciones de proteína de fusión tienen un porcentaje de hélices alfa menor de aproximadamente un 2 % y un porcentaje de láminas beta menor de aproximadamente un 2 % y al menos aproximadamente un 90 % de espirales aleatorias, según se determina mediante el algoritmo de GOR.

1. <u>Secuencias no repetitivas</u>

Se contempla específicamente que las secuencias XTEN de las realizaciones de la proteína de fusión de unión son sustancialmente no repetitivas. En general, las secuencias repetitivas de aminoácidos tienen tendencia a agregarse o a formar estructuras de orden superior, como ejemplifican las secuencias repetitivas naturales tales como los colágenos y las cremalleras de leucina. Estos aminoácidos repetitivos también pueden tender a formar contactos que dan como resultado estructuras cristalinas o pseudocristalinas. Por el contrario, la baja tendencia de las secuencias no repetitivas a agregarse permite el diseño de unos XTEN de secuencia larga con una frecuencia relativamente baja de aminoácidos cargados que de otro modo sería probable que se agregaran si las secuencias fueran repetitivas. En una realización, las secuencias XTEN tienen más de aproximadamente 36 a aproximadamente 1000 restos de aminoácidos, o más de aproximadamente 100 a aproximadamente 3000 restos de aminoácidos en los que no hay tres aminoácidos contiguos en la secuencia que sean tipos de aminoácidos idénticos a menos que el aminoácido sea serina, en cuyo caso no más de tres aminoácidos contiguos son restos de serina. En la realización anterior, la secuencia del XTEN es "sustancialmente no repetitiva". En otra realización, como se describe más completamente a continuación, las secuencias XTEN de las composiciones comprenden motivos de secuencia no solapados de 9 a 14 restos de aminoácidos, en donde los motivos consisten en 4 a 6 tipos de aminoácidos seleccionados entre glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P), y en donde la secuencia de dos restos cualesquiera de aminoácidos contiguos en un motivo no se repite más de dos veces en el motivo de la secuencia. En la realización anterior, la secuencia del XTEN es "sustancialmente no repetitiva".

El grado de repetitividad de un polipéptido o un gen se puede medir mediante programas informáticos o algoritmos o por otros medios conocidos en la técnica. De acuerdo con la actual invención, en el presente documento se divulgan

algoritmos que se utilizarán para calcular el grado de repetitividad de un polipéptido particular, tal como un XTEN, y se proporcionan ejemplos de secuencias analizadas mediante los algoritmos (*véanse* los ejemplos, a continuación). En una realización, se puede calcular la repetitividad de un polipéptido de una longitud predeterminada (en lo sucesivo en el presente documento, "puntuación de la subsecuencia") de acuerdo con la fórmula dada por la ecuación I:

$$\text{Puntuación de la subsecuencia} = \frac{\sum_{i=1}^{m} Recuento_i}{m} \qquad I$$

en donde: m = (longitud de aminoácidos del polipéptido) - (longitud de aminoácidos de la subsecuencia) + 1; y

$Recuento_i$ = número acumulado de apariciones de cada subsecuencia única dentro de la secuencia$_i$

Se desarrolló un algoritmo denominado "SegScore" para aplicar la ecuación anterior para cuantificar la repetitividad de polipéptidos, tales como un XTEN, proporcionando la puntuación de subsecuencia en donde se analizan secuencias de una longitud de aminoácidos predeterminada para determinar la repetitividad determinando el número de veces (un "recuento") que una subsecuencia única de longitud "s" aparece en la longitud establecida, dividida por el número absoluto de subsecuencias dentro de la longitud predeterminada de la secuencia. La FIG. 37 representa un diagrama del flujo lógico del algoritmo SegScore, mientras que la FIG. 38 representa un esquema de cómo se obtiene una puntuación de subsecuencia para un XTEN ficticio con 11 aminoácidos y una longitud de subsecuencia de 3 restos de aminoácidos. Por ejemplo, una longitud de polipéptidos predeterminada de 200 restos de aminoácidos tiene 192 subsecuencias de 9 aminoácidos solapadas y 198 subsecuencias de 3-meros, pero la puntuación de subsecuencia de cualquier polipéptido determinado dependerá del número absoluto de subsecuencias únicas y de la frecuencia con que aparece cada subsecuencia única (lo que significa una secuencia de aminoácidos diferente) en la longitud predeterminada de la secuencia. En el contexto de la presente invención en donde el algoritmo se utiliza para determinar el grado de repetitividad en un polipéptido, la variable "longitud de aminoácidos del polipéptido" se establece en 200 aminoácidos, y la variable "longitud de aminoácidos de subsecuencia" se establece en 3 aminoácidos. Por lo tanto, la puntuación de subsecuencia será igual a la suma de apariciones de cada marco único de 3-meros a través de una secuencia consecutiva de 200 aminoácidos del polipéptido dividida por el número absoluto de subsecuencias únicas de 3-meros dentro de la secuencia de 200 aminoácidos. En el ejemplo 58 se presentan ejemplos de dichas puntuaciones de subsecuencias derivadas de los primeros 200 aminoácidos de polipéptidos repetitivos y no repetitivos.

En una realización, la presente invención proporciona proteínas de fusión de unión que comprenden un XTEN en las que el XTEN tiene una puntuación de subsecuencia menor de 10, o menor de 9, o menor de 8, o menor de 7, o menor de 6, o menor de 5 o menor. En otra realización, la invención proporciona proteínas de fusión de unión que comprenden dos XTEN más en las que al menos un XTEN tiene una puntuación de subsecuencia menor de 10, o menor de 9, o menor de 8, o menor de 7, o menor de 6, o menor de 5 o menor. En otra realización más, la invención proporciona proteínas de fusión de unión que comprenden al menos dos XTEN en las que cada XTEN individual de 36 o más aminoácidos tiene una puntuación de subsecuencia menor de 10, o menor de 9, o menor de 8, o menor de 7, o menor de 6, o menor de 5 o menor. En las realizaciones de este párrafo, el XTEN se caracteriza como "sustancialmente no repetitivo".

Se cree que la característica no repetitiva de XTEN de la presente invención contribuye a muchas de las propiedades fisicoquímicas y biológicas mejoradas de las proteínas de fusión de unión; solo o junto con la elección de los tipos particulares de aminoácidos que predominan en el XTEN de las composiciones divulgadas en el presente documento. Estas propiedades incluyen un mayor grado de expresión de la proteína de fusión en la célula hospedadora, una mayor estabilidad genética del gen que codifica el XTEN y un mayor grado de solubilidad y menor tendencia a la agregación de las proteínas de fusión de unión resultantes en comparación con las proteínas de fusión que comprenden polipéptidos que tienen secuencias repetitivas. Estas propiedades permiten una fabricación más eficiente, un menor coste de los productos y facilitan la formulación de preparados farmacéuticos que comprenden XTEN que contienen unas concentraciones de fármaco extremadamente altas, en algunos casos superiores a 100 mg/ml. Además, las secuencias de los polipéptidos XTEN de las realizaciones están diseñadas para tener un bajo grado de repetitividad interna, con el fin de reducir o eliminar sustancialmente la inmunogenicidad cuando se administran a un mamífero. Las secuencias de polipéptidos compuestas por motivos cortos y repetidos, en gran parte limitados a solo tres aminoácidos, tales como glicina, serina y glutamato, pueden dar como resultado unas concentraciones de anticuerpos relativamente altos cuando se administran a un mamífero a pesar de la ausencia de epítopos de linfocitos T predichos en estas secuencias. Esto puede ser causado por la naturaleza repetitiva de los polipéptidos, ya que se ha demostrado que los inmunógenos con epítopos repetidos, incluidos los agregados proteicos, los inmunógenos reticulados y los hidratos de carbono repetitivos, son muy inmunogénicos y pueden, por ejemplo, dar como resultado la reticulación de los receptores de los linfocitos B, causando la activación de los linfocitos B. (Johansson, J., et al. (2007) Vaccine,

25:1676-82; Yankai, Z., et al. (2006) Biochem Biophys Res Commun, 345: 1365-71; Hsu, C. T., et al. (2000) Cancer Res, 60: 3701-5); Bachmann MF, et al. Eur J Immunol. (1995) 25 (12): 3445-3451).

2. Motivos de secuencia ilustrativos

La presente invención engloba XTEN utilizados como compañeros de fusión que comprenden múltiples unidades de secuencias más cortas, o motivos, en los que las secuencias de aminoácidos de los motivos no son repetitivas. La propiedad no repetitiva se cumple a pesar del uso de un enfoque de "bloque de construcción" utilizando una biblioteca de motivos de secuencia que están multimerizados para crear las secuencias XTEN. Por lo tanto, mientras que una secuencia de XTEN puede consistir en múltiples unidades de tan pocos como cuatro tipos diferentes de motivos de secuencia, debido a que los propios motivos generalmente consisten en secuencias de aminoácidos no repetitivas, la secuencia global del XTEN está diseñada para hacer que la secuencia sea sustancialmente no repetitiva.

Descritos en el presente documento, los XTEN tienen una secuencia no repetitiva de más de aproximadamente 36 a aproximadamente 3000 restos de aminoácidos en donde al menos aproximadamente un 80 %, o al menos aproximadamente un 85 %, o al menos aproximadamente un 90 %, o al menos aproximadamente un 95 %, o al menos aproximadamente un 97 % o aproximadamente un 100 % de la secuencia del XTEN consiste en motivos de secuencia no solapados, en donde cada uno de los motivos tiene aproximadamente de 9 a 36 restos de aminoácidos. En otras realizaciones, al menos aproximadamente un 80 %, o al menos aproximadamente un 85 %, o al menos aproximadamente un 90 %, o al menos aproximadamente un 95 %, o al menos aproximadamente un 97 % o aproximadamente un 100 % de la secuencia del XTEN consiste en motivos de secuencia no solapados, en donde cada uno de los motivos tiene de 9 a 14 restos de aminoácidos. En otras realizaciones más, al menos aproximadamente un 80 %, o al menos aproximadamente un 85 %, o al menos aproximadamente un 90 %, o al menos aproximadamente un 95 %, o al menos aproximadamente un 97 % o aproximadamente un 100 % del componente de la secuencia del XTEN consiste en motivos de secuencia no solapados en donde cada uno de los motivos tiene 12 restos de aminoácidos. En estas realizaciones, se prefiere que los motivos de secuencia estén compuestos principal o exclusivamente por aminoácidos hidrófilos pequeños, de modo que la secuencia global tenga una característica flexible no estructurada. Algunos ejemplos de aminoácidos que se incluyen en los XTEN son, por ejemplo, arginina, lisina, treonina, alanina, asparagina, glutamina, aspartato, glutamato, serina y glicina. Como resultado de las variables de ensayo, tales como la optimización de codones, los polinucleótidos de ensamblaje que codifican motivos de secuencia, la expresión de proteínas, la distribución de carga y la solubilidad de la proteína expresada, y la estructura secundaria y terciaria, se descubrió que las composiciones XTEN con características mejoradas incluyen principalmente restos de glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P), en donde las secuencias están diseñadas para ser sustancialmente no repetitivas. Descritas en el presente documento, las secuencias XTEN tienen predominantemente de cuatro a seis tipos de aminoácidos seleccionados entre glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) o prolina (P) que están dispuestos en una secuencia sustancialmente no repetitiva que tiene una longitud mayor de entre aproximadamente 36 y aproximadamente 3000 restos de aminoácidos. Descritos en el presente documento, los XTEN tienen secuencias de más de aproximadamente 36 a aproximadamente 3000 restos de aminoácidos, en donde al menos aproximadamente un 80 % de la secuencia consiste en motivos de secuencia no solapados, en donde cada uno de los motivos tiene de 9 a 36 restos de aminoácidos, en donde cada uno de los motivos consiste en 4 a 6 tipos de aminoácidos seleccionados entre glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P), y en donde el contenido de un tipo de aminoácido cualquiera en el XTEN completo no supera el 30 %. En otras realizaciones, al menos aproximadamente un 90 % de la secuencia del XTEN consiste en motivos de secuencia no solapados, en donde cada uno de los motivos tiene de 9 a 36 restos de aminoácidos, en donde los motivos consisten en 4 a 6 tipos de aminoácidos seleccionados entre glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P), y en donde el contenido de un tipo de aminoácido cualquiera en el XTEN completo no supera el 30 %. En otras realizaciones, al menos aproximadamente un 90 % de la secuencia del XTEN consiste en motivos de secuencia no solapados, en donde cada uno de los motivos tiene 12 restos de aminoácidos que consisten en 4 a 6 tipos de aminoácidos seleccionados entre glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P), y en donde el contenido de un tipo de aminoácido cualquiera en el XTEN completo no supera el 30 %. En otras realizaciones más, al menos aproximadamente un 90 %, o aproximadamente un 91 %, o aproximadamente un 92 %, o aproximadamente un 93 %, o aproximadamente un 94 %, o aproximadamente un 95 %, o aproximadamente un 96 %, o aproximadamente un 97 %, o aproximadamente un 98 %, o aproximadamente un 99 %, hasta aproximadamente un 100 % de la secuencia del XTEN consiste en motivos de secuencia no solapados, en donde cada uno de los motivos tiene 12 restos de aminoácidos que consisten en glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P), y en donde el contenido de un tipo de aminoácido cualquiera en el XTEN completo no supera el 30 %.

Descritos en el presente documento, los XTEN comprenden secuencias no repetitivas de más de aproximadamente 36 a aproximadamente 3000 restos de aminoácidos en donde al menos aproximadamente un 80 %, o al menos aproximadamente un 90 %, o aproximadamente un 91 %, o aproximadamente un 92 %, o aproximadamente un 93 %, o aproximadamente un 94 %, o aproximadamente un 95 %, o aproximadamente un 96 %, o aproximadamente un 97 %, o aproximadamente un 98 %, o aproximadamente un 99 % de la secuencia consiste en motivos de secuencia no solapados de 9 a 14 restos de aminoácidos, en donde los motivos consisten en 4 a 6 tipos de aminoácidos seleccionados entre glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P), y en donde la secuencia de dos restos de aminoácidos contiguos en un motivo cualquiera no se repite más de dos veces en el motivo

de secuencia. En otras realizaciones, al menos aproximadamente un 90 %, o aproximadamente un 91 %, o aproximadamente un 92 %, o aproximadamente un 93 %, o aproximadamente un 94 %, o aproximadamente un 95 %, o aproximadamente un 96 %, o aproximadamente un 97 %, o aproximadamente un 98 %, o aproximadamente un 99 % de una secuencia de XTEN consiste en motivos de secuencia no solapados de 12 restos de aminoácidos, en donde los motivos consisten en cuatro a seis tipos de aminoácidos seleccionados entre glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P), y en donde la secuencia de dos restos cualquiera de aminoácidos contiguos en un motivo de secuencia cualquiera no se repite más de dos veces en el motivo de secuencia. En otras realizaciones, al menos aproximadamente un 90 %, o aproximadamente un 91 %, o aproximadamente un 92 %, o aproximadamente un 93 %, o aproximadamente un 94 %, o aproximadamente un 95 %, o aproximadamente un 96 %, o aproximadamente un 97 %, o aproximadamente un 98 %, o aproximadamente un 99 % de una secuencia de XTEN consiste en motivos de secuencia no solapados de 12 restos de aminoácidos, en donde los motivos consisten en glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P), y en donde la secuencia de dos restos cualquiera de aminoácidos contiguos en un motivo de secuencia cualquiera no se repite más de dos veces en el motivo de secuencia. En otras realizaciones más, los XTEN consisten en 12 motivos de secuencia de aminoácidos en donde los aminoácidos se seleccionan entre glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P), y en donde la secuencia de dos restos cualquiera de aminoácidos contiguos en un motivo de secuencia cualquiera no se repite más de dos veces en el motivo de secuencia, y en donde el contenido de un tipo de aminoácido cualquiera en el XTEN completo no supera el 30 %. En las realizaciones anteriores arriba descritas en este párrafo, las secuencias XTEN son "sustancialmente no repetitivas".

Descritas en el presente documento, la invención proporciona composiciones que comprenden una, o dos, o tres, o cuatro o más secuencias XTEN no repetitivas de aproximadamente 36 a aproximadamente 1000 restos de aminoácidos, o acumulativamente de aproximadamente 100 a aproximadamente 3000 restos de aminoácidos, en donde al menos aproximadamente un 80 %, o al menos aproximadamente un 90 %, o aproximadamente un 91 %, o aproximadamente un 92 %, o aproximadamente un 93 %, o aproximadamente un 94 %, o aproximadamente un 95 %, o aproximadamente un 96 %, o aproximadamente un 97 %, o aproximadamente un 98 %, o de aproximadamente un 99 % a aproximadamente un 100 % de la secuencia consiste en múltiples unidades de dos o más motivos de secuencia no solapados seleccionados entre las secuencias de aminoácidos de la tabla 3, en donde la secuencia global permanece sustancialmente no repetitiva. Descrito en el presente documento, el XTEN comprende motivos de secuencia no solapados en los que aproximadamente un 80 %, o al menos aproximadamente un 85 %, o al menos aproximadamente un 90 %, o aproximadamente un 91 %, o aproximadamente un 92 %, o aproximadamente un 93 %, o aproximadamente un 94 %, o aproximadamente un 95 %, o aproximadamente un 96 %, o aproximadamente un 97 %, o aproximadamente un 98 %, o aproximadamente un 99 % o aproximadamente un 100 % de la secuencia consiste en múltiples unidades de dos o más secuencias no solapadas seleccionadas entre una familia de motivo único seleccionada de la tabla 3, lo que da como resultado una secuencia de familia. Como se utiliza en el presente documento, "familia" significa que el XTEN tiene motivos seleccionados únicamente de una sola categoría de motivo de la tabla 3; es decir, un XTEN con AD, AE, AF, AG, AM, AQ, BC O BD, y que cualquier otro aminoácido del XTEN que no sea de un motivo de familia se selecciona para conseguir una propiedad necesaria, tal como para permitir la incorporación de un sitio de restricción por los nucleótidos codificantes, la incorporación de una secuencia de escisión o para conseguir un mejor enlace a un componente de la proteína de unión.

Tabla 3: Motivos de secuencia del XTEN de 12 aminoácidos y familias de motivos

Familia de motivos*	SECUENCIA DEL MOTIVO	SEQ ID NO:
AD	GESPGGSSGSES	2
AD	GSEGSSGPGESS	3
AD	GSSESGSSEGGP	4
AD	GSGGEPSESGSS	5
AE, AM	GSPAGSPTSTEE	6
AE, AM, AQ	GSEPATSGSETP	7
AE, AM, AQ	GTSESATPESGP	8
AE, AM, AQ	GTSTEPSEGSAP	9
AF, AM	GSTSESPSGTAP	10
AF, AM	GTSTPESGSASP	11
AF, AM	GTSPSGESSTAP	12
AF, AM	GSTSSTAESPGP	13
AG, AM	GTPGSGTASSSP	14

Familia de motivos*	SECUENCIA DEL MOTIVO	SEQ ID NO:
AG, AM	GSSTPSGATGSP	15
AG, AM	GSSPSASTGTGP	16
AG, AM	GASPGTSSTGSP	17
AQ	GEPAGSPTSTSE	18
AQ	GTGEPSSTPASE	19
AQ	GSGPSTESAPTE	20
AQ	GSETPSGPSETA	21
AQ	GPSETSTSEPGA	22
AQ	GSPSEPTEGTSA	23
BC	GSGASEPTSTEP	24
BC	GSEPATSGTEPS	25
BC	GTSEPSTSEPGA	26
BC	GTSTEPSEPGSA	27
BD	GSTAGSETSTEA	28
BD	GSETATSGSETA	29
BD	GTSESATSESGA	30
BD	GTSTEASEGSAS	31
*Designa secuencias de motivos individuales que, cuando se utilizan juntas en diferentes permutaciones, dan como resultado una "secuencia de familia"		

En algunas realizaciones de familias de XTEN, una secuencia de XTEN comprende múltiples unidades de motivos de secuencia no solapados de la familia de motivos AD, o una secuencia de XTEN comprende múltiples unidades de motivos de secuencia no solapados de la familia de motivos AE, o una secuencia de XTEN comprende múltiples unidades de motivos de secuencia no solapados de la familia de motivos AF, o una secuencia de XTEN comprende múltiples unidades de motivos de secuencia no solapados de la familia de motivos AG, o una secuencia de XTEN comprende múltiples unidades de motivos de secuencia no solapados de la familia de motivos AM, o una secuencia de XTEN comprende múltiples unidades de motivos de secuencia no solapados de la familia de motivos AQ, o una secuencia de XTEN comprende múltiples unidades de motivos de secuencia no solapados de la familia BC, o una secuencia de XTEN comprende múltiples unidades de motivos de secuencia no solapados de la familia BD, presentando el XTEN resultante el intervalo de homología descrito anteriormente. En otras realizaciones, el XTEN comprende múltiples unidades de secuencias de motivos de dos o más de las familias de motivos de la tabla 3, seleccionadas para conseguir las características fisicoquímicas deseadas, que incluyen propiedades tales como carga neta, ausencia de estructura secundaria o ausencia de repetitividad que puede conferirse mediante la composición de aminoácidos de los motivos, descrita más completamente a continuación. En las realizaciones de la divulgación descritas anteriormente en este párrafo, los motivos incorporados en el XTEN se pueden seleccionar y ensamblar utilizando los métodos descritos en el presente documento para conseguir un XTEN de aproximadamente 36 a aproximadamente 3000 restos de aminoácidos. En la tabla 4 se presentan ejemplos no limitativos de secuencias de la familia XTEN.

Tabla 4. Polipéptidos XTEN

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
AE42_1	TEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGS	32
AE42_2	PAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSG	33
AE42_3	SEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSP	34
AG42_1	GAPSPSASTGTGPGSGTASSSPGSSTPSGATGSPGPSGP	35
AG42_2	GPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGASP	36
AG42_3	SPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGA	37
AG42_4	SASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATG	38
AE48	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGS	39
AM48	MAEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGS	40
AE144	GSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEG SAPGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSTEEPSEGSAPGTSESA TPESGPGSEPATSGSETPGTSTEEPSEGSAP	41
AF 144	GTSTPESGSASPGTSPSGESSTAPGTSPSGESSTAPGTSSTAESPGPGTSTSESPSG TAPGSTSSTAESPGPGTSPSGESSTAPGTSTPESGSASPGSTSSTAESPGPGTSPSG ESSTAPGTSPSGESSTAPGTSPSGESSTAP	42
AG144_1	PGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSAST GTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASP GTSSTGSPGASPGTSSTGSPGTPGSGTASSS	43
AG144_2	SGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSP GSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGA TGSPGSSPSASTGTGPGSSPSASTGTGPGASP	44
AG144_3	GTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSAST GTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSP SASTGTGPGASPGTSSTGSPGASPGTSSTGSP	45
AG144_4		46

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	GTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSS TGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPS ASTGTGPGTPGSGTASSSPGSSTPSGATGSP	
AE288	GTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPE SGPGTSTEPSEGSAPGSPSTSTEEGTSTEPSEGSAPGTSESATPESGPGTS ESATPESGPGTSESATPESGPGSEPATSGSETPGSPATGSPTSTEE GTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEG SAP	47
AG288_1	ASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTP GSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSS PGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTS STGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGAS PGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSP GSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGS	48
AG288_2	PGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGT ASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSP SASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSP GSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGA TGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSTP SGATGS	49
AG288_3	GSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSAST GTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASP GTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTSSTGSP GTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSTPSGA TGSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPG SGTASSSP	50
AF504	GASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGA TGSPGSXPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPG SGTASSSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSP GASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSAST GTGPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASP GTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSP GSSPSASTGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTTSS	51
	TGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTP SGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSP	

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
AF540	GSTSSTAESPGPGTSTSTAESPGQSTSESPSGTAPGSTSSTAESPGPGSTSSTAES PGPGTSTPESGSASPGSTSESPSGTAPGTSPSGESSTAPGSTSESPSGTAPGSTSESPSGTAP PSGTAPGTSPSGESSTAPGSTSESPSGTAPGSTSPSGESSTAPGSTSESPSGTAPGSTSESPSGTAP SESPSGTAPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTSPSGTAP GTSTPESGSASPGSTSSTAESPGPGTSTPESGSASPGTSTSESPSGTAPGSTSESPSGTAP ASPGSTSESPSGTAPGTSTPESGSASPGTSTESPSGTAPGSTSESPSGTAPGSTSES PSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTESPSGSASPGTSTPESGSASPGST SESPSGTAPGSTSESPSGTAPGTSTPESGSASPGSTSESPSGTAPGSTSESPSGTAP GTSTPESGSASPGTSTESPSGSASPGTSTSSTAESPGPGTSTPESGSASPGTSTPESGS PGPGTSTPESGSASPGSTSSTAESPGTAP	52
AD576	GSSESGSSEGSGPGSKGEPSESGSSKGSSESGSESEGGPGSSESGSSEGGPGSSESGSSE GGPGSSESGSSEGGPGSSESGSSEGSPAGSPGTSTEPSEGSAPGSPAGSPTSTEEGTS GSSEGGPGSSESGSSEGGPGSSESGSSEGSPAGSPGTSTEPSEGSAPGSPAGSPTSTEEGTS ESPGCSSGSEPSGSGEGGPGSSESGSSEGGPGSGEPSESGSSGGGEPSES GSSGSEGSSGPGESSGSPAGSSGSESPSGSSGSGGEPPSESGSSGSGGGEE PSESGSSGSSESGSSEGGPGESPAGSGSSGSESPAGSSGSGESPAGSSKGSSESGSE SPGGSSGSESPAGSPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTS SSGSSEGSSEGGPAGSPGTSTEPSEGSAPGTSTEPSEGSSEGGPAGSPGTSTEEGTS SSGSSEGSSEGGPAGSPGTSTEPSEGSAPGTSTEPSEGSSEGGPAGSPGTSTEEGTS SSGSSEGSSEGGPAGSPGTSTEPSEGSAPGTSTEPSEGSSEGGPAGSPGTSTEEGTS GGEPSESGSSGSGGEPSESGSSGSESPAGSSKGSSESGSSEGPGESSKGSSESGSSEG GPGSEGSSGPGESS	53
AE576	GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEG SAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSPAGS PTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTS TEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGP GSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPE SGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEP SEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEEGSAPGTS TEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETP GTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPE SGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEP SEGSAP	54
AF576	GSTSSTAESPGPGTSTSTAESPGPGSTSTSESPSGTAPGSTSSTAESPGPGSTSSTAES PGPGTSTPESGSASPGSTSESPSGTAPGTSPSGESSTAPGSTSESPSGTAPGSTSESPSGTAP PSGTAPGTSPSGESSTAPGSTSESPSGTAPGSTSPSGESSTAPGSTSESPSGTAPGSTSESPSGTAP SESPSGTAPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTSPSGTAP GTSTPESGSASPGSTSSTAESPGPGSTSTPESGSASPGTSTSESPSGTAPGSTSESPSGTAP ASPGSTSESPSGTAPGTSTPESGSASPGTSTESPSGTAPGSTSESPSGTAPGSTSES PSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTESPSGSASPGTSTPESGSASPGST SESPSGTAPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTSPSGTAP GTSTPESGSASPGSTSSTAESPGPGSTSTPESGSASPGTSTPESGS SGSASP	55

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
AG576	PGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSTPSG ATGSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGSTPSG GSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSTPSG PGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSG ATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSS TPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGS PGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTS STGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTP GSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTG PGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSAST GTGPGASPGTSSTGS	56
AE624	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGSPAGSP TSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTST EPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPATS GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPA GSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAP	57
AD836	GSSESGSSEGGPGSSESGSSEGGPGESPGGSSGSESGSGGEPSESGSSGESPGGSS GSESGESPGGSSGSESGSSESGSSEGGPGSSESGSSEGGPGSSESGSSEGGPGESP GGSSGSESGESPGGSSGSESGESPGGSSGSESGSSEGGPGSSESGSSEGGPGESP GSSESGSSEGGPGSSESGSSEGGPGSSESGSSEGGPGESPGGSSGSESGGP SGSSGESPGGSSGSESGSGGEPSESGSGGPGESPGGSSGSESGSSES GSSEGGPGSGGPGESPSESGSSGSESGSSEGGPGESPGGSSGSESGSSES ESPGGSGSESGSGGEPSESGSSGSGSESGSSGSESGSSEGGPGSGGEPSES GSSGSGGEPSESGSSGSESGSGPGESSGEPSPGGSSGSESGSSEGGPGESSGSESGS SGPGESSGSGGEPSESGSSGSESGSSEGGPGESSGSESGSSEGGPGESSGSESGS SGGEPSESGSSGSESGSGPGESSGSESGSSEGGPGESSGSESGSSEGGPGESES GSSGSGGEPSESGSSGSESGSGPGESSGSESGSSEGGPGESSGSESGSSEGES SGPGESSGESPGGSSGSESGSSEGGPGSSESGSSEGGPGSSESGSSEGGPG SGGEPSESGSSGSESGSSEGGPGESPGGSSGSESGSSEGGPESPSGSSGSESGSSEGG GPGESPGGSSGSESGSGGEPSESGSSGESPGGSSGSESGSGGEPSESGSS	58
AE864		59

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEG SAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSEPAGS PTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTS TEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGP GSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGP SAPGSPAGSPTSTEEGTSESATPESGPGSPATSGSETPGTSTEPSEGSAPGTSTEP SEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEP TEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEP SAPGSPAGSPTSTEEGSPAGSPTSTEEGTSTESSATPESGPGTSTEP GTSESATPESGPGTSTEPSEGSAPGSPTSTEEGTSTESSATPESGPGTSTEP SEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTS ESATPESGPGTSTEPSEGSAPGSPTSTEEGSPAGSPTSTEEGTSTESSATPESGPGTSTEP GTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGSPTSTEESATPE SGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGTSTEP PTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTS TEPSEGSAP	
AF864	GSTSESPSGTAPGTSTEPSEGSSTAPGSTSESPSGTAPGTSTEPSGS ASPGTSTPESGSASPGSTSESPSGTAPGSTSESPSGTAPGSTSESPSGTAPGSTSES PSGTAPGTSPSGESSTAPGTSPSGESSTAPGSTSSTAESPSGPGTSPSGESSTAPGTS PSGESSTAPGSTSSTAESPSGPGTSTPESGSASPGTSTSPESGSASPGSTSSTAPGTS GSTSESPSGTAPGTSTPESGSASPGSTSSTAESPSGPGTSTPESGSASPGSTSSTAPGSG TAPGTSPSGESSTAPGSTSSTAESPGPGTSPSGESSTAPGTSTPESGSASPGSTSST AESPGPGSTSSTAESPGPGTSTSSTAESPGPGTSPSGESSTAPGSTSSTAPGSTSES SESPSGTAPGSTSESPSGTAPGTSTPESGPGTAPGTSTPESGSASPGSTVVVVSESPSGTAP GSTSESPSGTAPGSTSESPSGTAPGSTSESPSGTAPGSTSESPSGTAPGSTSESPSG TAPGTSTPESGSASPGTSPSGESSTAPGTSPSGESSTAPGSTSESPSGTAPGTSTPESGSASP ESSTAPGTSTPESGSASPGSTSESPSGTAPGSTSESPSGTAPGTSPSGESSTAPGST SESPSGTAPGTSTPESGSASPGTSTPESGSASPGSTSESPSGTAPGTSTPESGSASP GSTSSTAESPGPGSTSESPSGTAPGSTSESPSGTAPGTSPSGESSTAPGSTSESTAES PGPGTSPSGESSTAPGTSTPESGSASPGTSPSGESSTAPGTSPSGESSTAPGTSPSG ESSTAPGSTSSTAESPGPGTSSTAESPGPGTSPSGESSTAPGSSPSASTGTGPGSS	60
	TPSGATGSPGSSTPSGATGSP	
AG864	GASPGTSSTGSPGSPSASTGTGPGSPSASTGTGPGTPGSGTASSSPGSSTPSGA TGSPGSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPG SGTASSSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSP GASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASST GTGPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASP GTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSP GSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSS TGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTP SGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPG ASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTTSST GSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPS GATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGASTGTGPGSSPSASTGTGPGS SPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGT GPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGA TGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSS TPSGATGSPGASPGTSSTGSP	61

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
AM875	GTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEPGSTSSTAESPGPGTSTEPSEGS ASPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGSTPESGSASPGSEPAT SGSETPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAP TEPSEGSAPGTSTEPSEGSAPGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAP GTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPE SGPGTSTEPSEGSAPGSHEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGS GTASSSPGSSTPSGATGSPGTSTEPSEGSAPGTSTEPSEGSAPGSPEPATSGSETPGS PAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGSAASGAPSTGGTSESATPESG PGSPAGSPTSTEEGSPAGSPTSTEEGSTSSTAESPGPGSTSSTAESPGPGTSTEPSEGS STAPGTPGSGTASSSPGSSTPSGATGSPGSSTSAESPGPGTSTTEEGSSTAPGS ATPESGPGSEPATSGSETPGTSTSSTAESPGPGTSTTEPSEGSSTAPGS EPATSGSETPGSDPATSGSETPGTSTEPSEGSAPGSTSSTAESPGPGTSTEPSEGSAS PGSTSESPSGETAPGTSTEPSEGSAPGTSTEPSEGSAPGSTSTEPSEGSAPGSSTPSGA TGSPGSSPSASTGTGPGASPGTSSTSGSPGSTPGSTSESATPESGPGSPA GSPTSTEEGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTSGSPGTSESATPESGPG TSTEPSEGSAPGTSTEPSEGSAP	62
AE912	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGSPAGSP TSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSPGSAPGTST EPSEGSAPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGS TSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPTSTEEGTSTEPSEGS APGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPAT GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSPA GSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGSPAGSPTSTEECGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGS GPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGTSESATPESGPGTSESAT PESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSE SATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGS AP	63
AM923	MAEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGTSTEPS EGSAPGSEPATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTPESGSASPGSTS PSPSGTAPGSTSESPSGTAPGTSTPESGSASPGSTSTPESGSASPGSPATSGSETPG TSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGS APGTSTEPSEGSAPGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESAT PESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTST EPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGTASSSPG	64

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	SSTPSGATGSPGTSTEPSEGSAPGTSTEPSEGSAPGSPATSGSETPGSPAGSPTST EEGTPAGSPTSTEEGTSTEPSEGSAPGASASGAPSTGGTSESATPESGPGSPAGSP TSTEEGSPAGSPTSTEEGSTSTAESPGPGSTSSTAESPGPGTSTPESGSASPGSTSTEP SGPATSGSETPGTSTEPSGSAPGSTSTEEGTSTEPSEGSAPGTSTEPSEGSAP SEPATSGSETPGTSTEPSGSAPGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAP TPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGSSTPSGATGSPGSSP SASTGTGPGASPGTSSTGSPGSEPATSGSETPTGTSESATPPSGPGSPAGSPTSTEE GSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTSESATPPSGPGTSTPPSPG SAPGTSTEPSPGSAP	
AM1318	GTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTPESGS ASPGSTSESPSGSTAPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPGSEPAT SGSETPGTSSEATPLSGQPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTS TEPSEGSAPGTSTEPSEGSAPGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAP GTSESATPESGPGTSTEPSEGSAPGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAP SGPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGS GTASSPGSSTPSGATGSPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETTGS PAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGPEPTGPAPSGGSEPATSGSETP GTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSPTSTEEGSPAGSPTS TEEGTSESATPESGPGSEAGSPTSTEEGSPAGSPTSTEEGSTSSTAESPGPGSTSES PSGTAPGTSPSGESSTAPGSTSESPSGTAPGTSPSGESSTAPGTS TEPSEGSAPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGTSESATPESGP GTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSPSEGSAPGTSPSGESS TAPGTSPSGESSTAPGTSPSEGSAPGSPAGSPTSTEEGTSTEP SEGSAPGSSPSASTGTGPGSTPSGATGSPGSSTPSGATGSPGSS STPSGATGSPGASPGTSSTGSPGASASGAHSTGGTSPSGESSTAPGSTSSTAESPG PGTSPSGESSTAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSSPSAST GTGPGSSTPSGATGSPGTSSTSGSPGTSTEPSEGSASPGTSPSGESSTAPGTSPS GESSTAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGSTSESPSGTAPGTSPS TSESPSGTAPGTSTEPSEGSASPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSA PGSPAGSPSTSTEEGTSESATPESGPGSEPATSGSETPGSETPGSATGSPGASPGTSS TGSPGSSTPSGATGSPGSTSESPSGTAPGTSPSGESSTAPGSTSSTAESPGPGSSTP SGATGSPGASPGTSSTGSPGTPGSGTASSSPGSPAGSPTSTEEGSPAGSPTSTEEG TSTEPSEGSAP	65
BC 864	GTSTEPSEPGSAGTSTEPSEPGSAGSEPATSGTEPSAGSAGSEPTSTEPGSEPATSGT EPSGSEPATSGTEPSGSEPATSGTEPSGSGSASEPTSTEPSEPGSAGSEPAT SGTEPSGTSTEPSEPGSAGSEPATSGTEPSGSEPATSGTEPSGTSTEPSEPGSAGTS TEPSEPGSAGSEPATSGTEPSGSEPATSGTEPSGTSTEPSEPGSAGTSTEPSEP GTSEPSTSEPGAGSEPATSGTEPSGSEPATSGTEPSGTSTEPSEPGSAGTSTEPSEP GSAGSGASEPTSTEPGSGSEPATSGTEPSGSEPATSGTEPSGSEPA TSGTEPSGTSTEPSEPGSAGSEPATSGTEPSGSGSASEPTSTEPGTSTEPSEPGSAGS EPATSGTEPSGSGASEPTSTEPGTSTEPSEPGSAGSGASEPTSTEPGSEPATSGTEP SGSGASEPTSTEPGTSTEPSEPGSAGSEPATSGTEPSGTSTEPSEPGSAGSEPA TSGTEPSGTSTEPSEPGSAGTSTEPSEPGSAGTSTEPSEPGSAGT STESEPGSAGTSTEPSEPGSAGTSEPSTSEPGAGSASEPTSTEPGTSTEPSEPGS AGTSTEPSEPGSAGTSTEPSEPGSAGSEPATSGTEPSGSGASEPTSTEPGSEPATS GTEPASGSEPATSGTEPSGSEPATSGTEPSGTSEPSTSEPGAGSEP ATSGTEPSGSGASEPTSTEPGTSTEPSEPGSAGSEPATSGTEPSGSGASEPTSTEPG TSTEPSEPGSA	66
BD864		67

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	GSETATSGSETAGTSESATSESGAGSTAGSETSTEAGTSESATSESGAGSETATS GSETAGSETATSGSETAGTSTEASESGSASGTSTEASESGSASGTSESATSESGAGS ETATSGSETAGTSTEASESGSASGSTAGSETSTEAGTSESATSESGAGTSESATSES GAGSETATSGSETAGTSESATSESGAGTSTEASESGSASGSETATSGSETAGSETA TSGSETAGTSTEASESGSASGSTAGSETSTEAGTSESATSESGAGTSTEASESGSAS GSETATSGSETAGSTAGSETSTEAGSETATSGSETAGSETATSGSETAGTSESAT SESGAGTSESATSESGAGSETATSGSETAGTSESATSESGAGTSESATSESGAGS ETATSGSETAGSETATSGSETAGTSTEASESGSASGSTAGSETATSGSETATSGSE	
	ETAGTSESATSESGAGSTAGSETSTEAGSTAGSETSTEAGSTAGSETSTEAGTSY EASESGSASGSTAGSETSTEAGTAGSETSTEAGTSEASESGSASGSTAGSETSTE AGSETATSGSETAGTSTEASESGSASGTSESATSESGAGSETATSGSETAGTSESA TSESGAGTSESATSESGAGSETATSGSETAGTSESATSESGAGSETATSGSETAG TSTEASESGSASGTSTEASESGSASGSTAGSETSTEAGSTAGSETSTEAGSETASTSGS ETAGTSESATSESGAGTSESATSESGAGSETATSGSETATSGSETAGSET ATSGSETAGTSTEASESGSASGTSESATSESGAGSETATSGSETAGTSESET AGTSESATSESGAGTSESATSESGAGSETATSGSETA	
AE948	GTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEG SAPGSPAGSPTSTEEGTSTEPSESGSAPGSEPATSGSETPGTSESATPESGPGTSESA TPESGPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTS TEPSEGSAPGSEPATSGSETPGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETP GTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSESATP EPSGPGTSTEPSEGSAPGSEPATSGSETPGTSTEPSEGSAPGSEPATSGSETPGTS ESATPESGPGTSTEEGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGSPTS TEEGSPAGSPTSTEEGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGSPAGS PTSTEEGTSESATPESGPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTS ESATPESGPGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGSEPATSGSETP GSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSESAP TEEGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTS SEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTS TEPSEGSAPGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAP GTSTEPSEGSAPGTSESATPESGPGTSESATPESGP	68
AE1044	GSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTS TEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPAT SGSETPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSE PATSGSETPGTSTEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPE GTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSESATPE SGPGSTPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGTSTEPSEGSAPGTSESA TPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTS TPESGPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSESATPESGPGTS TEPSEGSAPGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSESATPESGP GTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTS TEEGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSEPAT SGSETPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTS EGSATPESGPGSEPATSGSETPGTSTEPSEGSAPGSPTSTEEGSPAGSPTSTEEGTS GTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGS ETPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSESATPESGPGTSTEP SEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTS TEPSEGSAPGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAP GTSTEPSEGSAPGTSESATPESGPGTSESATPESGP	69

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
AE1140	GSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGSEPATSGSETPGTSESATPE SGPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEP SEGSAPGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSE PATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEE GTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPE SGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEP SEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTS TEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAP GTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEG SAPGTSESATPESGPGTSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSESA TPESGPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSPAGSPTSTEEGSP AGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGSEPATSGSETP GTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSTEPSEG SAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSESATPESGPGTSTEPSEG SGSETPGSPAGSPTSTEEGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSE	70
	PATSGSETPGSEPATSGSETPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGS GTSTEPSEGSAPGTSTEPSEGSAPGSPATSGSETPGTSTEPSEGSAPGTSTEP SGPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSE	71
AE1236	GSPAGSPTSTEEGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGS ETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEP SEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGSPAGSPTSTEEGTS TEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAP GSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTS TEEGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGTSESATPESGPGSPAGS PTSTEEGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGTSESATPESGPGTS ESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEE GTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEG SAPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEP SGSETPGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEE AGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPTSTEEGSP GTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEG SAPGSSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTS PTSTEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGTS ESATPESGPGTSTEPSEGSAPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGTS GSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGSPAGSPTS TEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEEGSPAGSEP	72

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	GSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGSPPATSGS ETPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESA TPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGSE PATSGSETPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEE GTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEP TEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAP SEGSAPGSPEPATSGSETPGTSTEEGSPAGSPTSTEEGTSESATPESGPGTS TEPSEGSAPGTSTEEGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAP ETPGSEPATSGSETPGTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGSPPATSGS GTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSPEPATSGSETPGTSTEP TPESGPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPEPATSGSE TPESGPGSPAGSPTSTEEGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSA GSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSPEPATSGSETPGTSTEP SGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSP SEGSAPGSEPATSGSETPGSTPGTSTEPSEGSAPGTSESATPESGPGSP AGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSPTSTEEGSPAGSPTSTEE GTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPTS ETPGSEPATSGSETPGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTST	
AE1428	GSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGTSESATPE SGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEP SEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGPGSEPATSGSETPGTS ESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTS GTSTEPSEGSAPGTSTEEPSEGSAPGSPTSTEEGTSTEPSEGSAPGSEPATSGSETP SGPGTSESATPESGPGTSTEPSEGSAPGSPTSTEEGTSTEPSEGSAPGSPTSTEEGTS PTSTEEGTSTEPSEGSAPGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSP AGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSEPATSGSETP GTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESA ETPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPTSTEEGTSESA TPESGPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGTS TEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSPTSTEEGTS GSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPE SGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPTSTEEGTS PTSTEEGTSTEPSEGSAPGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSP AGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSEPATSGS ETPGTSESATPESGPGTSTEPSEGSAPGSPTSTEEGTSESATPESGPGTSTEPS EGSAPGTSESATPESGPGTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPS EGSAPGSPAGSPTSTEEGTSESATPESGPGTSTEEGTSESATPESGPGTSTEP SGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGTS	73
AE1524	GTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPTSTEPSEGSAPGSPTS TEEGSPAGSPTSTEEGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTS TEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSPAGSPTSTEE GTSTEPSEGSAPGSEPATSGSETPGTSTEPSEGSAPGSPTSTEEGTSTEPSEGSAP GSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSPAGSPTSTEEGTSTEP ETPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSEPAT SGSETPGTSTEPSEGSAPGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGSE PATSGSETPGTSTEPSEGSAPGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSE GTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPE ETPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSPTSTEEGTSTEP SEGSAPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTS TEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSPAGSPTSTEE GTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEE SGPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTS ESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGP GSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEG SAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSPA	74

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
AE1620	GSEPATSGSETPGTSTEPSEGSAPGSEPATSGSETPGTSTEPSEGSAPGTSESATPE SGPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESA TPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSETPGSPAGSPTSTEEGTTS ESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETP GTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGS ETPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGTSESA TPESGPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTS TEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEE GSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEG SAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGSPAGS PTSTEEGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTS TEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGSPAGSPTSTEE GTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSESATPE SGPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSTEP SEGSAPGSEPATSGSETPGSEPATSGSETPGTSESATPESGPGTSESATPESGPGTS TEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSESATPESGP GTSTEPSEGSAPGTSESATPESGPGTSTEEGTSTEPSEGSAPGSPAGSPTSTEEGTST	75
AE1716	GTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGSEPATSGS ETPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGTSESA TPESGPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEE ESATPESGPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGP GSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESP SGPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESA TPESGPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSPAGSPTSTEEGSP AGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGTSESATPESGP GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPTSTEEGTSTEPSEGSAP GSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSPTSTEEGTSTEPSEGSAP TEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGSPTSTEEGTST	76
AE1812	GTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSEPATSGS ETPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPTSTEEGTSTEP SEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSP AGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAP	77

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	GTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSATPESGPGSPAGSPTS TEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPAT SGSETPGTSTEPSEGSAPGSPAGSPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTS ESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAP GSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGTSESATPE SGPGSPAGSPTSTEEGSPAGSPTSTEEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTS TPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTS ESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEEGTSTEPSEGSAPGSPAGSPS GSEPATSGSETPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPE SGPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESA TPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTS ESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEEGTSTEPSEGSAPGSPAGSPS GSEPATSGSETPGTSTEPSEGSAPGSPTSTEEGTSTEPSEGSAPGSPAGSPTSTEEGTS EPTGSPAGSPTSTEEGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSEP	SEQ ID NO: 78
AE1908	GSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEG SAPGSPAGSPTSTEEGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSEPAT SGSETPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGTS ESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPG GSEPATSGSETPGTSTEEDSPSAPGSEPATSGSETPGSEPATSGSETPGTSTEPSEG SAPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGTSTEPSEG SGSETPGSEPATSGSETPGTSESATPESGPGTSESATPESGPGTSTEPSEGSAP TEPSEGSAPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAP GTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEG SAPGSPAGSPTSTEEGTSESATPESGPGTSESATPESGPGSEPATSGSETPGTSTEPSEG PATSGSETPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAP GSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAP ETPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSPAGS PTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSEPATSGSETPGTS TEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGP GSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGTSESATPE TEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGTSESATPE	79
AE2004 A	GTSTEPSEGSAPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGS ETPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEP SEGSAPGTSTEPSEGSAPGSPAGSPTSTEEDGTSESATPESGPGTSESATPESGPGSP AGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGSEPATSGSETP GTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTS TEEGSPAGSPTSTEEGTSESATPESGPGSPAGSPTSTEEPSEGSAPGSPAGSPAT SGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSE PATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGP GTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSESATPESGPGSPAGSPTS SGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGSPAGS PTSTEEGTSTEPSEGSAPGTSGSETPGTSESATPESGPGTSTEEGTSESATPESGPGTS ESATPESGPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGP GSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTS TEEGTSESATPESGPGTSTEPSEGSAPGSPTSTEEGTSTEPSEGSAPGTSESATPESGP	79
AG948		80

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	GSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTA SSSPGTPGSGTASSSPGTPGSGTASSSPGSSPSASTGTGPGTPGSGTASSSPGSSPS ASTGTGPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSTPSGATGSPG ASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSST GSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGASPGTSSTGSPGTPGS GTASSSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPG ASPGTSSTGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGAT GSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGTPGS	
	GTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPG SSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGAT GSPGASPGTSSTGSPGSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGSSTPS GATGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGS SPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASS SPGSSTPSGATGSPGSSPSASTGTGPGSSTPSGATGSPGSSPSASTGTGPGSSPSAS TGTGPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGTP GSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGS PGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGTP	
AG1044	GTPGSGTASSSPGTPGSGTASSSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSS TGSPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGASP GTSSTGSPGSSPSASTGTGPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGP GTPGSGTASSSPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGSSPSAST GTGPGTPGSGTASSSPGASPGTSSTGSPGTPGSGTASSSPGSSPSASTGTGPGSSP SASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSP GTPGSGTASSSPGSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGSSTPSGA TGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGASP GTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSP GSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTA SSSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGSSTP SGATGSPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPG ASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGTGPGTPGSGTAS SSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPGASPG TSSTGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGS STPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASS SPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSAS TGTGPGTPGSGTASSSPGSSPSASTGTGPGASPGTSSTGSPGSSTPSGATGSPGTP GSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGTPGSGTASS	81
AG1140		82

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	GASPGTSSTGSPGSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGASPGTSS TGSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGTPGSGTASSSPGTPG SGTASSSPGSSTPSGATGSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSP GSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTA SSSPGASPGTSSTGSPGSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGSSTP SGATGSPGSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPG TPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGT GPGTPGSGTASSSPGTPGSGTASSSPGASPGTSSTGSPGTPGSGTASSSPGTPGSG TASSSPGSSPSASTGTGPGASPGTSSTGSPGSSTPSGATGSPGASPGTSSTGSPGSS PSASTGTGPGTPGSGTASSSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGS PGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGT ASSSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGTPGSGTASSSPGSSP SASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSP GTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGTPGSGTA SSSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPGSSTP SGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSPSASTGTGPG ASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGSSPSASTGT GPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGT SSTGSPGSST	83
AG1236	GSSPSASTGTGPGTPGSGTASSSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTA SSSPGASPGTSSTGSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGTPGSG GTASSSPGTPGSGTASSSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGATGSPG ASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGTPGSGTASSSPGSSTPSGAT GSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGTPGSGTASSSPGASPG TSSTGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPG ASPGTSSTGSPGSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGASPGTSST GSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGS GTASSSPGSSTPSGATGSPGASPGTSSTGSPGSSTPSGATGSPGTPGSGTASSSPG SSPSASTGTGPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGAT GSPGASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGSSPSA	84
AG1332	STGTGPGTPGSGTASSSPGTPGSGTASSSPGASPGTSSTGSPGTPGSGTASSSPGA SPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGASPGTSSTGS PGSSPSASTGTGPGTPGSGTASSSPGTPGSGTASSSPGSSPSASTGTGPGTPGSGT ASSSPGASPGTSSTGSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGTP GSGTASSSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTG PGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSAST GTGPGASPGTSSTGSPGTPGSGTASSSPGSSPSASTGTGPGSSPSASTGTGPGASP GTSSTGSPGASP	

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	GSSTPSGATGSPGSSPSASTGTGPGTPGSGTASSSPGSSPSASTGTGPGASPGTSS TGSPGSSPSASTGTGPGTPSGGTASSSPGASPGTSSTGSPGTSPSGATGSPGSSPS ASTGTGPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPG SSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGAT GSPGSSTPSGATGSPGSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGTPGS GTASSSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGPGTPGSGTASSSPG SSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGPGTPGSGT GPGASPGTSSTGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGPGTPG SSTGSPGTPG	85
AG1428	GTPGSGTASSSPGPGSSTPSGATGSPGASPGTSSTGSPGSSTPSGATGSPGTPGSGTA SSSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGSSPS ASTGTGPGTPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPG ASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGASPGTSSTG GSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSSPSASTGTGPGSSPSA STGTGPGASPGTSSTGSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGA SPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSSPSASTGTGPGASPGTSSTGS PGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPGSSTPSGATGSPGSSPSAST GTGPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGASP GTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSP GSSTPSGATGSPGTPGSGTASSSPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGA TGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGASPGTSSTGSPGSSP GASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGSSPSAST GTGPGTPGSGTASSSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASP GTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSP GSSTPSGATGSPGTPGSGTASSSPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGA TGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGASPGTSSTGSPGSSP GTSSTGSPGASPGTSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGTPGSGT ASSSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGSS PSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSSPSASTGTG PGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSG ATGSPGSSTPSGATGSPGSSTPSGATGTGPGSSTPSGATGSPGSSTPSGATGSPGS STPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTA SSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSST PGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATG SPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGSSPSA STGTGPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGA SPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGAT GSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSST PSGATGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSS PGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSG ATGSPGSSTPSGATGSPGSSTPSGATGTGPGSSTPSGATGSPGSSTPSGATGSPGS STPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTA SSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSST PSGATGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSS PGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSG ATGSPGSSTPSGATGSPGSSTPSGATGTGPGSSTPSGATGSPGSSTPSGATGSPGS	86
AG1524	GSSTPSGATGSPGTPGSGTASSSPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGA TGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSTPSGATGSPGTPG SGTASSSPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGATGSPGSSPSASTGTGP GTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGTSSTGSPGTPGSGTA SSSPGSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGASPGTSSTGSPGASPG TSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGS SPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSSPSASTGT GPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSA SSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGTPGSGTASSSPG SGTASSSPGTPGSGTASSSPGASPGTSSTGSPGSSPSASTGTGPGSSTPSGATGSP GTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSG ATGSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGS STPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSG ATGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPG SSTPSGATGSPGSSTPSGAT

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	PSASTGTGPGTPGSGTASSSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTG PGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGT ASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGSSTP PSGATGSPGSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGSTPSGATGSPGSST GSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTA SSSPGASPGTSSTGSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGSSTP SGATGSPGTPG	
AG 1620	87	
	GSSTPSGATGSPGSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGTPGSGTA SSSPGASPGTSSTGSPGSSTPSGATGSPGASPGTSSTGSPGSGTASSSPGASPG TSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGS SPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTG SPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSG ATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPGSSPSASTGTGPGSS PSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSPSASTGTG PGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGPGSSTPSGATGSPGASPGTSSTGSPG SSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGSSPSASTGTGPGSSTPSGAT GSPGSSPSASTGTGPGSSTPSGATGSPGSSPSASTGTGPGTPGSGTASSSPGTPGS GTASSSPGSSTPSGATGSPGSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPG SSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPGASPGTSST GSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGTPGSGTASSSPGSSTPS GATGSPGSST	
AG1716	88	
	GASPGTSSTGSPGSSPSASTGTGPGSSTPSGATGSPGSSPSASTGTGPGTPGSGTA SSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSTPSGATGSPGTPGS GTASSSPGSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPG SSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGT GPGASPGTSSTGSPGSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGSSTP TGTGPGSSPSASTGTGPGTPGSGTASSSPGTPGSGTASSSPGSSTPSGATGSPGTP GSGTASSSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSS PGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSG ATGSPGASPGTSSTGSPGTPGSGTASSSPGSSPSASTGTGPGTPGSGTASSSPGAS PGTSSTGSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSP GTPGSGTASSSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPGTPGSGTA SSSPGASPGTSSTGSPGSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGTP TSSTGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGSSTPSGATGSPG PGSSGTASSSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGASPGTSSTG SPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSTPSGATGSPGTPGSG TASSSPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSS PSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGATGS PGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTS STGSPGTPG	
AG1812	89	

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	GSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTA SSSPGSSPSASTGTGPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGASPG TSSTGSPGSSPSASTGTGPGTPGSGTASSSPGTPGSGTASSSPGASPGTSSTGSPGS STPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTG SPGSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGATGSPGTPGSG TASSSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGSS TPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGTPGSGTASSSPGSSPSAST GTGPGTPGSGTASSSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGSST PSGATGSPGSSPSASTGTGPGASPSASTGTGPGTPGSGTASSSPGASPGTSSTGSP GASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGA TGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGTPG SGTASSSPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGSP	
	GASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSTPSGA TGSPGTPGSGTASSSPGSSPSASTGTGPGASPGTSSTGSPGSSTPSGATGSPGTPG SGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSP GSSPSASTGTGPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGA TGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSTPSGATGSPGSSTP SGATGSPGASP	90
AG1908	GSSPSASTGTGPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSPSAST GTGPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGTPG SGTASSSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSSPSASTGTGP GSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGASPGTSS TGSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGTPGSGTASSSPGTPG SGTASSSPGTPGSGTASSSPGSSPSASTGTGPGASPGTSSTGSPGATGSPGTSSTGSP GSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGA TGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGTPG SGTASSSPGTPGSGTASSSPGSSPSASTGTGPGASPGTSSTGSPGATGSPGTSSTGSP GSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGA TGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGTPG SGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPG SSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTAS SSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSA STGTGPGSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGASPGTSSTGSPGS STPSGATGSPGSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGT GPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGSSPSAS TGTGPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGSS PSASTGTGPGSSP	91
AG2004A		

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	GSSPSASTGTGPGSTPSGATGSPGSSTPSGATGSPGTPSGSTASSPGSSTPSGA TGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSPSASTGTGPGTPG SGTASSSPGASPGTSSTGSPGSSTPSGATGSPGTPGSGTASSSPGTPGSGTASSSPGSSIS GSSTPSGATGSPGSSTPSGATGSPGTSSTGSPGASPGTSSTGSPGATGSPGTSS TGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSSIS ASTGTGPGSSTPSGATGSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPG SSPSASTGTGPGASPGTSSTGSPGTPGSSTPSGASPGTSSTGSPGSSPSASTGST GPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGTPGSG TASSSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGSSIS TPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGSSIS PGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGSSPSAST GTGPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGSSPSAST GTSSTGSPGTPGSGTASSSPGTPSGTASSSPGSSTPSGATGSPGSSTPSGATGSP GTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGSSPSAST GTGPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGSSPSAST PSGATGSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSSPSASTGTGP GASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGPGSSPSASTGTGPGASPGTSS TGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGSSPS ASTGTGPGASP	92
AE72B	SPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPES GPGSEPATSGSETPG	93
AE72C	TSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTST EEGTSTEPSEGSAPG	94
AE108A	TEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEP SEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTS	95
AE108B	GSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPE SGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAP	96
AE144A	STEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETI PGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATP ESGPGSPAGSPTSTEEGSPAGSPTSTEEGS	97
AE144B	SEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGS	

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	APGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSAEGP TSTEEGSPAGSPTSTEEGTSTEPSEGSAPG	
AE180 A	TSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPA GSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPG SEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGS APGSEPATS	98
AE216 A	PESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSE SATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPG TSESATPESGPGTSESATPESGPGSEPATSGSETPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESAT	99
AE252 A	ESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSES ATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGT STEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSET PGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSG SETPGTSESATPESGPGTSTEPSE	100
AE288 A	TPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSE PATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPTSTEEGTSESATPESGP GSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEG SAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPAT SGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTS ESA	101
AE324 A	PESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTST EPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPG TSESATPESGPGTSTEPSEGSAPGSPTSTEEGTSESATPESGPGSEPATSGSETPG TPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESAT PESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPA GSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATS	102
AE360 A	PESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPA GSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPG TSESATPESGPGTSTEPSEGSAPGSPTSTEEGTSTEPSEGSAPGSPAGSPTST EEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSP TSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEP ATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPG SEPATSGSETPGTSESAT	103

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
AE396 A	PESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSPSATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG TSPSATPESGPGSEPATSGSETPGTSESATTPESGPGSEPATSGSETPGTSESATPES GPGTSTEEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSE SATPESGPGTTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSTEPSEGSAPGTSTEPS	104
AE432 A	EGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSE SATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPG SPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGS APGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGTSTEPSEGSAPGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSE SATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPG TSESATPESGPGTTSESATPESGPGSEPATSGSETPGSEPATSGSETPGTSESAT EEGTSTEPSEGSAPGTSTEPS	105
AE468 A	EGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPTSTEEGTSTEPSEGSAPGTSE SATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPG TSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTST EEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGTSE SATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPG	106
	TSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSE TPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATS GSETPGTSESAT	
AE504 A	EGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPA GSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGT SESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPG EEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGSEPATPGTSESATPESGPGPGTST EPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPTPGTSESATPESGPG SPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPES GPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPS EGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPS	107
AE540 A		108

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	TPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTS TEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAP GTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGS ETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESA TPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETP TEPSEGSAPGTSESATPESGPGSPAASGSETPGTSESATPESGPGSEPATSGSETP GTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGS EPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESA TPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSP AGSPTSTEEGTSTEPSEGSAPGTSTEP	
AE576 A	TPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTS ESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAP GTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPE SGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEP SEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTS ESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGP GSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPTSTEEGTSESATPE SGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGTSTEP SEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPG PATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETP GTSESA	109
AE612 A	GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPA GSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPTSTEEGTSES ATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPG GPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGTSESATPESGPGTSTEP PESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSE SATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPTSTEEGSPAGSPTSTEEG SPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSE TPGTSESATPESGPGSEPATSGSETPGTSESATPESGPAGSPAGSPTSTEEGSPAGSP TSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPA GSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPG SEPATSGSETPESGPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGS APGSEPATSGSETPGTSESAT	110
AE648 A	PESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTST EPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPG SEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGSTSTEPSEGSAPGTSTEPSEGSAPGSPTSTEEGTSTEPSEGS EGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGTSTEPS SATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEG SPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSE TPGTSESATPESGPGSEPATSGSETPGTSESATPESGPAGSPAGSPTSTEEGSPAGSP TSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPA GSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPG SEPATSGSETPESGPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGS APGSEPATSGSETPGTSESAT	111
AE684 A		112

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	EGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTST EPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPG TSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSE TPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPS EGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPTSTEEGTSTEPSEGSAPGTSE SATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPG TSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTST EEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPTSTEEGTSE SATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEG TSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSEPATSGSE TPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATS	
AE720 A	TSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGT STEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESG PGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSE GSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPA TSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGT STEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSA PGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATP ESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAG SPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGT SESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPTSTE EGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPT STEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPA TSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTE	113
AE756 A	TSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGT STEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESG PGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSE GSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPA TSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGT STEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSA PGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATP ESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAG SPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGT SESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTE EGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPT STEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPA TSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTE	114
AE792 A		115

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	EGSAPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSE SATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPG TSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSE TPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSEPAGSP TSTEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTST EPSEGSAPGTSTEPSEGSAHGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPG SPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPG GPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPAGSP TSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSE SATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPG TSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPG GPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPAGSP PESGPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTST EPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPS	
AE828 A	PESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSE SATPESGPGSEPATSGSETPGSEPATPESGPGSEPATSGSETPGTSESATPESGPG TSTEPSEGSAPGTSTEPSEGSAPGSPTSTEEGTSTEPSEGSAPGTSTEPSEGS APGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPS	116
AG72A	EGSAPGTSTEPSEGSAPGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESA SATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPG TSTEPSEGSAPGTSTEPSEGSAPGSPTSTEEGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESA GSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPA GSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPG SEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGS APGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSP TSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSE SATPESGPGSEPATSGSETPGSPAUSPTSTEEGTSTEPSEGSAPG TSTEPSEGSAPGSPATSGSETPGTSESAT	117
AG72B	GPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSPSASTGTGPGASPGT SSTGSPGTPGSGTASSs	118
AG72C	GSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSS TGSPGTPGSGTASSSP	119
AG108 A	SPSASTGTGPGASPGTSSTGSPGSPSASTGTGPGTPGSGTASSSPGSSTPSGATG SPGSSTPSGATGSPGA	120
	SASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGP GASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASP	

Nombre del XTEN	Secuencia de aminoacidos	SEQ ID NO:
AG108 B	PGTPGSGTASSSPGSSTPSGATGSPGTPASSSPGSSTPSGATGSPGSSTPSG ATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSS	121
AG144 A	PGSSPSASTGTGPGSSPSASTGTGPGTPGSGTPASSSPGSSTPSGATGSPGSSPSAST GTGPGASPGTSSTGSPGTPSGATGSPGSTPSGATGSPGTASSSPGASP GTSSTGSPGASPGTSSTGSPGTPGSGTASSS	122
AG144 B	PSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPSSASTGTGP GSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSAST GTGPGSSPSASTGTGPGASPGTSSTGSPGASP	123
AG180 A	TSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGS SPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASS SPEASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGASPGT SSTGSPGTPGS	124
AG216 A	TGTGSPGSSPSASTGTGPGTTPGSGTASSSPGSSTPSGATGSPGSPSASTGTGPGAS TGTSSTGSPGTPGSGTASSSPESSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSP GASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTA SSSPESSTPSGATGSPGSSPSASTTGTPGSSPSASTGTGPGSSTPSG	125
AG252 A	TSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGS SPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASS SPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGASPGT SSTGSPGTPGSGTASSSPGSSTPSGATGSPGSPSASTGTGPGSSPSASTGTGPGSS TPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGS	126
AG288 A	TSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGS SPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASS SPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGASPGT SSTGSPGTPGSGTASSSPGSSTPSGATGSPGSPSASTGTGPGSSPSASTGTGPGSS TPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGS PGTPGS	127
AG324 A		128

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	TSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPG ASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGAT GSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSASTGTGPGS STGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGS SPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGT GPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSTP	
AG360 A		129
	TSSTGSPGASPGPTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPG ASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASS	
	SSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSA STGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGS SPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATG SPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSG ATGSPGSSTPSGATGSPGASPG	
AG396 A		130
	GATGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGT PGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGTASPGTSSTGSPGASPGTSSTG SPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSG TASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSASTGTGPGSSPSASTGTGPGA SPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTG PGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTS STGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGAS PGT	
AG432 A		131
	GATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGS STPSGATGSPGSSTPSGATGSPGSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPG SPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGT SSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTP GSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTG PGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSAST GTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASP GTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPS	
AG468 A		132
	TSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPG ASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGS SSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASPGT TSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPG ASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTAS SSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSA STGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGS SPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATG SPGSSPSASTGTGPGASPG	
AG504 A		133

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	TSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPG ASPGTSSTGSPGASTGTGPGTPGSGTPSGATGSPGASPGTSSTGSPGTPGSGTAS SSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASPG TSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPG ASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTAS SSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSA STGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGS SPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATG SPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGSTPGSGTASSSPGSSTPP	
AG540 A	TSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPG ASPGTSSTGSPGASPGTSSTGSPGASTGSPGASTGSPGASPGTSSTGSPGAST GSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSA STGTGPGASPGTSSTGSPGASTGSPGSSTPSGATGSPGASPGTSSTGSPGA SPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGS PGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSG ATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGAS TPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGS PGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGT ASSSPGSSTPSGATGSPGSSTPSGATGSPGASPG	134
AG576 A	TSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGS SPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTG SPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSG ATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGAS PGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASTGTGPGAS GASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGA TGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSTPS ASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPG	135
	SSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGASPGTSSTG SPGATGSPGSSTPSGATGSPGASPGTSSTGSPGASTGSPGSSTPSGATGSPGSSTP GATGSPGASPG	
AG612 A	STGSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSST PSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGSP GASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSS TGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASP GTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASTGSPGSSTPSGATGSPGSTPSGA GASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGA TGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSST GTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSP GSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGA TGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTS	136
AG648 A		137

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	GTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPG SSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASS SSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGS GTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGATGSPG SSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGAT GSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGS GTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPG TPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTAS SSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPG TSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPG ASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSST GSPGSSPSASTGTGPGTPGSGTASSSPGSSTP	
AG684 A	TSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGS STPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTG SPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSAS TGTGPGTPTGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGS TPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGS PGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTS STGSPGTPGSGTTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGAS PGTSSTGSPGTPGSGTTASSSPGASPGTSSTPSGATGSPGTPGSGTASSSPGASTG GTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSGTASSSPGSSTPSGATGSP SASTGTGPGASPGTSSTGSPGASPGTSSTPSGATGSPGSSTPSGATGSPGSSTPSGA TGSPGA	138
AG720 A	TSSTGSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGS STPSGATGSPGSSTPSATGTGPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATG SPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGT SSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGA SPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGS PGASPGTSSTGSPGTPTGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASS PGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSP GSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGA TGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTPGSGATGSPGTSST GASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSAST GTGPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGASPG	139
AG756 A	TSSTGSPGSPSASTGTGPGSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGS SPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSS SPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTSST SSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSTPSGATGSPG TPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTG	140

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	PGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSAST GTGPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSST PSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSP GSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSS TGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASP GTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSP GTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSAST GTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSP SASTGTGPGASPGTSSTGSPGASPG	141
AG792 A	TSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGS SPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASS SPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGASPGT SSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSS TPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGS PGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSAST GTGPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASP GTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSS PSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSP GSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSPSAST GTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSP SASTGTGPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGP GASPG	142
AG828 A	TSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGS SPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASS SPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGASPGT SSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSS TPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGS PGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSAST GTGPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASP PSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSP GSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSS TGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASP GTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSP GTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSAST GTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSP SASTGTGPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGP GASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSTP	

Descrita en el presente documento, la composición de proteína de fusión de unión comprende una o más secuencias XTEN no repetitivas de aproximadamente 36 a aproximadamente 3000 restos de aminoácidos, en donde al menos aproximadamente un 80 %, o al menos aproximadamente un 90 %, o aproximadamente un 91 %, o aproximadamente un 92 %, o aproximadamente un 93 %, o aproximadamente un 94 %, o aproximadamente un 95 %, o aproximadamente un 96 %, o aproximadamente un 97 %, o aproximadamente un 98 %, o aproximadamente un 99 % a aproximadamente un 100 % de la secuencia consiste en 36 motivos de secuencia de aminoácidos no solapados seleccionados entre una o más de las secuencias de polipéptidos de las tablas 11-14, ya sea como una secuencia de familia, o donde los motivos se seleccionan entre dos o más familias de motivos.

En aquellas realizaciones en donde el componente XTEN de la proteína de fusión tiene menos del 100 % de sus aminoácidos que consisten en cuatro a seis aminoácidos seleccionados entre glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P), o menos del 100 % de la secuencia que consiste en las secuencias de una cualquiera de las tablas 4 u 11-14, los demás restos de aminoácidos se seleccionan entre cualquier otro de los 14 L-aminoácidos naturales, pero se seleccionan preferentemente entre aminoácidos hidrófilos, de modo que la secuencia del XTEN contenga al menos aproximadamente un 90 %, o al menos aproximadamente un 91 %, o al menos aproximadamente un 92 %, o al menos aproximadamente un 93 %, o al menos aproximadamente un 94 %, o al menos aproximadamente un 95 %, o al menos aproximadamente un 96 %, o al menos aproximadamente un 97 %, o al menos aproximadamente un 98 %, o al menos aproximadamente un 99 % de aminoácidos hidrófilos. Los aminoácidos XTEN que no son glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) están intercalados a lo largo de la secuencia del XTEN, están ubicados dentro de o entre los motivos de la secuencia, o se concentran en uno o más tramos cortos de la secuencia del XTEN. En dichos casos, cuando el componente XTEN de la proteína de fusión de unión comprende aminoácidos distintos de glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P), se prefiere que los aminoácidos no sean restos hidrófobos y no deben conferir sustancialmente una estructura secundaria del componente XTEN. Los restos hidrófobos que están menos favorecidos en la construcción del XTEN incluyen triptófano, fenilalanina, tirosina, leucina, isoleucina, valina y metionina. Adicionalmente, las secuencias XTEN se pueden diseñar para que contengan menos del 5 % o menos del 4 % o menos del 3 % o menos del 2 % o menos del 1 % o ninguno de los siguientes aminoácidos: cisteína (para evitar la formación y oxidación de disulfuros), metionina (para evitar la oxidación), asparagina y glutamina (para evitar la desamidación). Por lo tanto, en algunas realizaciones, el componente XTEN de la proteína de fusión que comprende otros aminoácidos además de glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) tendría una secuencia con menos del 5 % de los restos contribuyendo a hélices alfa y láminas beta según lo medido mediante el algoritmo de Chou-Fasman, y tendrían al menos un 90 %, o al menos aproximadamente un 95 % o más de formación de espirales aleatorias según lo medido mediante el algoritmo de GOR.

3. <u>Longitud de la secuencia</u>

En otro aspecto, la invención proporciona composiciones de proteína de fusión de unión que comprenden una proteína de unión y uno o más polipéptidos XTEN en donde la longitud de las secuencias XTEN se elige tomando como base la propiedad o la función que se quiere conseguir. Dependiendo de la propiedad o la función prevista, las composiciones de proteína de fusión de unión comprenden secuencias XTEN de una longitud corta o intermedia y/o XTEN más largas que pueden servir como portadores. Las presentes proteínas de fusión de unión de la divulgación engloban XTEN o fragmentos de XTEN con unas longitudes de aproximadamente 6, o aproximadamente 12, o aproximadamente 36, o aproximadamente 40, o aproximadamente 100, o aproximadamente 144, o aproximadamente 288, o aproximadamente 401, o aproximadamente 500, o aproximadamente 600, o aproximadamente 700, o aproximadamente 800, o aproximadamente 900, o aproximadamente 1000, o aproximadamente 1500, o aproximadamente 2000, o aproximadamente 2500, o hasta aproximadamente 3000 restos de aminoácidos de longitud. En otros casos, las secuencias XTEN de la divulgación pueden tener de aproximadamente 6 a aproximadamente 50, o de aproximadamente 100 a 150, de aproximadamente 150 a 250, de aproximadamente 250 a 400, de aproximadamente 400 a aproximadamente 500, de aproximadamente 500 a 900, de aproximadamente 900 a 1500, de aproximadamente 1500 a 2000 o de aproximadamente 2000 a aproximadamente 3000 restos de aminoácidos de longitud. En las realizaciones de las proteínas de fusión de unión de la divulgación, el uno o más XTEN o fragmentos de secuencias XTEN presentan individualmente al menos aproximadamente un 80 % de identidad de secuencia, o alternativamente un 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de identidad de secuencia en comparación con un motivo o un XTEN seleccionado de una cualquiera de las tablas 4 u 11-14, o un fragmento del mismo de longitud comparable. En algunas realizaciones, las proteínas de fusión de unión comprenden una primera y al menos una segunda secuencia de XTEN, en donde la longitud acumulada de los restos en las secuencias XTEN es mayor de aproximadamente 100 a aproximadamente 3000 restos de aminoácidos, y los XTEN pueden ser idénticos o pueden ser diferentes en secuencia o en longitud. Como se utiliza en el presente documento, la "longitud acumulada" pretende englobar la longitud total, en restos de aminoácidos, cuando se incorpora más de un XTEN en las proteínas de fusión de unión de las realizaciones. Descritas en el presente documento, cada una de las primeras y al menos la segunda secuencias presenta al menos aproximadamente un 80 % de identidad de secuencia, o alternativamente un 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de identidad de secuencia en comparación con una o más secuencias XTEN de la tabla 4, o fragmentos de las mismas.

Como se describe más completamente a continuación, se divulgan métodos en los que la proteína de fusión de unión se diseña seleccionando la longitud del XTEN para conferir una semivida diana u otra propiedad fisicoquímica a una proteína de fusión administrada a un sujeto. Cuando el XTEN se utiliza como portador, la invención aprovecha el descubrimiento de que al aumentar la longitud de los polipéptidos no repetitivos y no estructurados se mejora la naturaleza no estructurada de los XTEN y, en consecuencia, mejoran las propiedades biológicas y farmacocinéticas de las proteínas de fusión que comprenden el portador XTEN. Como se describe con más detalle en los ejemplos, los aumentos proporcionales en la longitud del XTEN, incluso si se crean mediante un orden repetido de motivos de secuencia de familia única (por ejemplo, los cuatro motivos AE de la tabla 3), dan como resultado una secuencia con un mayor porcentaje de formación de espirales aleatorias, según se determina mediante el algoritmo de GOR, o un contenido reducido de hélices alfa o de láminas beta, según se determina mediante el algoritmo Chou-Fasman, en comparación con longitudes más cortas del XTEN. Además, el aumento de la longitud del compañero de fusión polipeptídico no estructurado, como se describe en los ejemplos, da como resultado una proteína de fusión con un aumento desproporcionado en la semivida terminal en comparación con proteínas de fusión con compañeros polipeptídicos no estructurados con unas longitudes de secuencia más cortas. En general, las longitudes acumuladas de XTEN superiores a aproximadamente 400 restos incorporados en las composiciones de proteína de fusión de unión dan como resultado una semivida más larga en comparación con unas longitudes acumuladas más cortas; por ejemplo, menores de aproximadamente 280 restos.

En algunas realizaciones, donde el XTEN sirve principalmente como portador, la invención engloba composiciones de proteína de fusión de unión que comprenden una o más XTEN en donde la longitud de secuencia del XTEN acumulada de la proteína o proteínas de fusión es mayor de aproximadamente 100, o mayor de aproximadamente 200, o mayor de aproximadamente 400, o mayor de aproximadamente 500, o mayor de aproximadamente 600, o mayor de aproximadamente 800, o mayor de aproximadamente 900, o mayor de aproximadamente 1000 a aproximadamente 3000 restos de aminoácidos, en donde la proteína de fusión presenta unas propiedades farmacocinéticas mejoradas cuando se administra a un sujeto en comparación con una proteína de unión no unida a XTEN y administrada a una dosis comparable. Descritas en el presente documento, la una o más secuencias XTEN muestran al menos aproximadamente un 80 %, o al menos aproximadamente un 90 %, o al menos aproximadamente un 91 %, o al menos aproximadamente un 92 %, o al menos aproximadamente un 93 %, o al menos aproximadamente un 94 %, o al menos aproximadamente un 95 %, o al menos aproximadamente un 96 %, o al menos aproximadamente un 97 %, o al menos aproximadamente un 98 % o más de identidad con una secuencia seleccionada de la tabla 4, o fragmentos de la misma, y el resto de la secuencia o secuencias portadoras contienen al menos un 90 % de aminoácidos hidrófilos, y menos de aproximadamente un 2 % de la secuencia global consiste en aminoácidos hidrófobos o aromáticos o cisteína. Las propiedades farmacocinéticas mejoradas de las proteínas de fusión de unión en comparación con las proteínas de unión no unidas a XTEN se describen con más detalle a continuación.

4. <u>Segmentos XTEN</u>

En el presente documento se describen XTEN de longitudes cortas o intermedias, en donde la elección del XTEN confiere diferentes funciones o propiedades a las proteínas de fusión de unión. En particular, los diseños de configuración de la proteína de fusión de unión donde el XTEN sirve como un conector flexible o está diseñado para interferir con los receptores de aclaramiento, o cuando se utiliza una longitud corta o intermedia del XTEN para facilitar la penetración tisular o para variar la fuerza de las interacciones de la proteína de fusión de unión con su diana, o cuando es deseable distribuir la longitud acumulada de XTEN en al menos dos segmentos de longitud corta o intermedia, la invención proporciona proteínas de fusión de unión que comprenden una o más secuencias XTEN truncadas.

El XTEN de longitud corta o intermedia puede ser un XTEN o un fragmento de un XTEN de una longitud de aproximadamente 6 aminoácidos a aproximadamente 600 aminoácidos, o aproximadamente 12 a aproximadamente 288 aminoácidos, o aproximadamente 36 a aproximadamente 144 aminoácidos, o aproximadamente 42 a aproximadamente 96 aminoácidos de longitud. En la tabla 4 se presentan algunos ejemplos no limitativos de XTEN de longitud corta o intermedia contemplados para su inclusión en las realizaciones de proteínas de fusión de unión de la divulgación, pero también pueden incluir fragmentos de los motivos de la tabla 3 o fragmentos de las secuencias de la tabla 4 utilizados individualmente o unidos entre sí utilizando los métodos divulgados en el presente documento para conseguir un XTEN de una longitud dada, incluidas las longitudes englobadas por los intervalos divulgados anteriormente. En algunos ejemplos no limitativos, como se representa esquemáticamente en las FIG. 39A-C, la secuencia AG864 de 864 restos de aminoácidos se puede truncar para producir una AG144 con 144 restos, una AG288 con 288 restos, una AG576 con 576 restos, u otras longitudes intermedias, mientras que la secuencia AE864 (FIG. 39D-E) puede truncarse para producir una AE288 o una AE576 u otras longitudes intermedias. Se contempla específicamente que dicho enfoque se pueda utilizar con cualquiera de las realizaciones de XTEN descritas en el presente documento o con cualquiera de las secuencias enumeradas en la tabla 4 para dar como resultado un XTEN de una longitud deseada.

En el presente documento se describen XTEN de longitudes más largas en donde la secuencia es sustancialmente no repetitiva. La incorporación de XTEN de longitud más larga como portadores en proteínas de fusión de unión confiere propiedades mejoradas a las proteínas de fusión en comparación con los compañeros de fusión de longitud más corta XTEN, que incluyen tasas más lentas de absorción sistémica, mayor biodisponibilidad y semivida

aumentada después de la administración subcutánea o intramuscular a un sujeto, y semivida terminal o área bajo la curva más largas. En una realización, el XTEN de longitud más larga tiene más de aproximadamente 400, o más de aproximadamente 600, o más de aproximadamente 800, o más de aproximadamente 900, o más de aproximadamente 1000, o más de aproximadamente 1100, o más de aproximadamente 1200, o más de aproximadamente 1300, o más de aproximadamente 1400, o más de aproximadamente 1500, o más de aproximadamente 1600, o más de aproximadamente 1700, o más de aproximadamente 1800, o más de aproximadamente 1900, o más de aproximadamente 2000, hasta aproximadamente 3000 restos de aminoácidos o más de longitud, en donde el XTEN ensamblado es sustancialmente no repetitivo.

En algunas realizaciones, las proteínas de fusión de unión comprenden al menos dos segmentos XTEN en los que los segmentos XTEN pueden ser idénticos o pueden ser diferentes, en donde la longitud acumulada de los componentes XTEN es mayor que aproximadamente 100 a aproximadamente 3000 restos de aminoácidos y comprende al menos un segmento de secuencia de al menos aproximadamente 36 a aproximadamente 923, o al menos aproximadamente 42 a aproximadamente 875, o al menos aproximadamente 96 a aproximadamente 576, o al menos aproximadamente 100 a aproximadamente 288, o al menos aproximadamente 132 a aproximadamente 144 restos de aminoácidos en donde el segmento o segmentos de secuencia consisten en al menos cuatro, o al menos cinco, o al menos seis tipos diferentes de aminoácidos, y la suma de restos de glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) en el segmento o segmentos de secuencia constituye al menos aproximadamente un 80 %, o al menos aproximadamente un 85 %, o al menos aproximadamente un 90 %, o al menos aproximadamente un 91 %, o al menos aproximadamente un 92 %, o al menos aproximadamente un 93 %, o al menos aproximadamente un 94 %, o al menos aproximadamente un 95 %, o al menos aproximadamente un 96 %, o al menos aproximadamente un 97 %, o al menos aproximadamente un 98 %, o al menos aproximadamente un 99 % de la secuencia de aminoácidos total del segmento de secuencia, y al menos aproximadamente un 90 %, o al menos aproximadamente un 91 %, o al menos aproximadamente un 92 %, o al menos aproximadamente un 93 %, o al menos aproximadamente un 94 %, o al menos aproximadamente un 95 %, o al menos aproximadamente un 96 %, o al menos aproximadamente un 97 %, o al menos aproximadamente un 98 % del resto de la secuencia o secuencias XTEN consisten en aminoácidos hidrófilos, y menos de aproximadamente un 2 % del resto de la secuencia o secuencias XTEN consisten aminoácidos hidrófobos o aromáticos, o cisteína. En otra realización, la invención proporciona una proteína de fusión de unión aislada en donde la longitud acumulada del componente XTEN es mayor de aproximadamente 100 a aproximadamente 3000 restos de aminoácidos y comprende al menos un segmento de secuencia de al menos aproximadamente 36 a aproximadamente 923, o al menos aproximadamente 42 a aproximadamente 875, o al menos aproximadamente 96 a aproximadamente 576, o al menos aproximadamente 100 a aproximadamente 288, o al menos aproximadamente 132 a aproximadamente 144 restos de aminoácidos, en donde la suma en el segmento o segmentos de secuencia de los restos de glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) en el segmento o segmentos de secuencia constituye al menos aproximadamente un 90 %, o al menos aproximadamente un 91 %, o al menos aproximadamente un 92 %, o al menos aproximadamente un 93 %, o al menos aproximadamente un 94 %, o al menos aproximadamente un 95 %, o al menos aproximadamente un 96 %, o al menos aproximadamente un 97 %, o al menos aproximadamente un 98 %, o al menos aproximadamente un 99 % de la secuencia de aminoácidos total del segmento de secuencia, y en donde la puntuación de subsecuencia de un segmento o de los segmentos acumulados es menor de 10, o menor de 9, o menor de 8, o menor de 7, o menor de 6, o menor de 5, y al menos aproximadamente un 90 %, o al menos aproximadamente un 91 %, o al menos aproximadamente un 92 %, o al menos aproximadamente un 93 %, o al menos aproximadamente un 94 %, o al menos aproximadamente un 95 %, o al menos aproximadamente un 96 %, o al menos aproximadamente un 97 %, o al menos aproximadamente un 98 % del resto de la secuencia o secuencias XTEN consisten en aminoácidos hidrófilos, y menos de aproximadamente un 2 % del resto de la secuencia o secuencias XTEN consiste en aminoácidos hidrófobos, aromáticos o cisteína.

5. <u>Secuencias potenciadoras de la expresión del XTEN aminoterminal</u>

En una realización, la invención proporciona una secuencia de XTEN corta diseñada para ser incorporada como la porción aminoterminal de la proteína de fusión de unión, en donde la expresión de la proteína de fusión se mejora en una célula hospedadora transformada con un vector de expresión adecuado que comprende una secuencia líder aminoterminal optimizada (que codifica el XTEN aminoterminal) incorporada en el polinucleótido que codifica la proteína de fusión de unión. Se ha descubierto, como se describe en los ejemplos 14-17, que una célula hospedadora transformada con dicho vector de expresión que comprende una secuencia líder aminoterminal optimizada (NTS) en el gen de la proteína de fusión de unión da como resultado una expresión muy mejorada de la proteína de fusión de unión en comparación con la expresión de una proteína de fusión de unión correspondiente a partir de un polinucleótido que no comprende la NTS, y puede obviar la necesidad de incorporar una secuencia líder no XTEN utilizada para mejorar la expresión. En una realización de lo anterior, la invención proporciona proteínas de fusión de unión que comprenden una NTS en donde la expresión de la proteína de fusión de unión del gen codificante en una célula hospedadora se mejora aproximadamente un 50 %, o aproximadamente un 75 %, o aproximadamente un 100 %, o aproximadamente un 150 %, o aproximadamente un 200 %, o aproximadamente un 400 % en comparación con la expresión de una proteína de fusión de unión que no comprende la secuencia del XTEN aminoterminal (donde el gen codificante carece de la NTS).

En una realización de lo anterior, el polipéptido XTEN aminoterminal comprende una secuencia que presenta una identidad de secuencia de al menos aproximadamente un 80 %, más preferentemente al menos aproximadamente un

90 %, más preferentemente al menos aproximadamente un 91 %, más preferentemente al menos aproximadamente un 92 %, más preferentemente al menos aproximadamente un 93 %, más preferentemente al menos aproximadamente un 94 %, más preferentemente al menos aproximadamente un 95 %, más preferentemente al menos aproximadamente un 96 %, más preferentemente al menos aproximadamente un 97 %, más preferentemente al menos aproximadamente un 98 %, más preferentemente al menos el 99 % o el 100 % con la secuencia de aminoácidos de AE48 o AM48, siendo las secuencias respectivas como sigue:

AE48: MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGS (SEQ ID NO: 143)

AM48: MAEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGS (SEQ ID NO: 144)

En otra realización, el XTEN aminoterminal corto puede unirse a un XTEN de longitud más larga para formar la región aminoterminal de la proteína de fusión de unión, en donde la secuencia polinucleotídica que codifica el XTEN aminoterminal corto confiere la propiedad de expresión mejorada en la célula hospedadora, y en donde la larga longitud del XTEN expresado contribuye a las propiedades mejoradas del portador XTEN en la proteína de fusión, como se ha descrito anteriormente. En lo anterior, el XTEN corto puede unirse a cualquiera de los XTEN divulgados en el presente documento (por ejemplo, un XTEN de la tabla 4) y el XTEN resultante puede, a su vez, unirse al aminoterminal de cualquiera de los restos de direccionamiento divulgados en el presente documento (por ejemplo, un resto de direccionamiento dirigido a una diana de la tabla 1) como un componente de la proteína de fusión de unión. Alternativamente, los polinucleótidos que codifican el XTEN corto (o su complemento) pueden unirse a polinucleótidos que codifican cualquiera de los XTEN (o su complemento) divulgados en el presente documento, y el gen resultante que codifica el XTEN aminoterminal puede, a su vez, unirse al extremo 5' de polinucleótidos que codifican cualquiera de los restos de direccionamiento (o al extremo 3' de su complemento) divulgados en el presente documento. En las realizaciones preferidas de lo anterior, el polipéptido XTEN aminoterminal de longitud larga puede presentar al menos aproximadamente un 80 %, o al menos aproximadamente un 90 %, o al menos aproximadamente un 91 %, o al menos aproximadamente un 92 %, o al menos aproximadamente un 93 %, o al menos aproximadamente un 94 %, o al menos aproximadamente un 95 %, o al menos aproximadamente un 96 %, o al menos aproximadamente un 97 %, o al menos aproximadamente un 98 %, o al menos un 99 %, o presenta un 100 % de identidad de secuencia con una secuencia de aminoácidos seleccionada del grupo que consiste en las secuencias AE624, AE912 y AM923.

En cualquiera de las realizaciones de XTEN aminoterminal anteriores descritas anteriormente, el XTEN aminoterminal puede tener de aproximadamente uno a aproximadamente seis restos de aminoácidos adicionales, preferentemente seleccionados entre glicina, serina, treonina, glutamato, prolina y alanina, para acomodar los sitios de restricción de la endonucleasa que se emplean para unir los nucleótidos que codifican el XTEN aminoterminal al gen que codifica el resto de direccionamiento de la proteína de fusión. Algunos ejemplos no limitativos de aminoácidos compatibles con los sitios de restricción y los aminoácidos preferidos se enumeran en la tabla 6, a continuación. Los métodos para la generación de las secuencias aminoterminales y la incorporación en las proteínas de fusión de la invención se describen más detalladamente en los ejemplos.

6. <u>Carga neta</u>

En otras realizaciones, los polipéptidos XTEN tienen una característica no estructurada impartida mediante la incorporación de restos de aminoácidos con una carga neta y que contienen una proporción baja o ningún aminoácido hidrófobo en la secuencia del XTEN. La carga neta global y la densidad de carga neta se controlan modificando el contenido de aminoácidos cargados en las secuencias XTEN, ya sean positivos o negativos, con la carga neta representada normalmente como el porcentaje de aminoácidos del polipéptido que contribuyen a un estado cargado más allá de los restos que son cancelados por un resto con una carga opuesta. En algunas realizaciones, la densidad de carga neta del XTEN de las composiciones puede estar por encima de +0,1 o por debajo de -0,1 cargas/resto. Por "densidad de carga neta" de una proteína o un péptido se entiende en el presente documento la carga neta dividida por el número total de aminoácidos de la proteína o el propéptido. En otras realizaciones, la carga neta de un XTEN puede ser de aproximadamente un 0 %, aproximadamente un 1 %, aproximadamente un 2 %, aproximadamente un 3 %, aproximadamente un 4 %, y aproximadamente un 5 %, aproximadamente un 6 %, aproximadamente un 7 %, aproximadamente un 8 %, aproximadamente un 9 %, aproximadamente un 10 % aproximadamente un 11 %, aproximadamente un 12 %, aproximadamente un 13 %, aproximadamente un 14 %, aproximadamente un 15 %, aproximadamente un 16 %, aproximadamente un 17 %, aproximadamente un 18 %, aproximadamente un 19 % o aproximadamente un 20 % o más. En algunas realizaciones, la secuencia del XTEN comprende restos cargados separados por otros restos tales como serina o glicina, lo que conduce a una mejor expresión o comportamiento de purificación. Tomando como base la carga neta, algunos XTEN tienen un punto isoeléctrico (PI) de 1,0, 1,5, 2,0, 2,5, 3,0, 3,5, 4,0, 4,5, 5,0, 5,5, 6,0, o incluso 6,5. En una realización, el XTEN tendrá un punto isoeléctrico entre 1,5 y 4,5, y portará una carga neta negativa en condiciones fisiológicas.

Dado que la mayoría de los tejidos y superficies de un ser humano o un animal tienen una carga neta negativa, en algunas realizaciones, las secuencias XTEN están diseñadas para tener una carga neta negativa para minimizar las interacciones no específicas entre las composiciones que contienen XTEN y diferentes superficies, tales como vasos sanguíneos, tejidos sanos o diferentes receptores. Sin pretender quedar ligado a teoría alguna en particular, el XTEN puede adoptar conformaciones abiertas debido a la repulsión electrostática entre aminoácidos individuales del

polipéptido XTEN que portan individualmente una carga neta negativa y que se distribuyen a través de la secuencia del polipéptido XTEN. Dicha distribución de la carga neta negativa en las longitudes de secuencia de XTEN expandidas puede conducir a una conformación no estructurada que, a su vez, puede dar como resultado un aumento eficaz en el radio hidrodinámico. En las realizaciones preferidas, la carga negativa del presente XTEN se confiere mediante la incorporación de restos de ácido glutámico. Por ejemplo, cuando se desea un XTEN con una carga negativa, el XTEN puede seleccionarse únicamente de una secuencia de la familia de AE, que tiene aproximadamente una carga neta del 17 % debido al ácido glutámico incorporado, o puede incluir proporciones variables de motivos que contienen ácido glutámico de la tabla 3 para proporcionar el grado deseado de carga neta. Algunos ejemplos no limitativos de XTEN de AE incluyen, pero sin limitación AE36, AE42, AE48, AE144, AE288, AE576, AE624, AE864 y AE912, y las secuencias de polipéptidos de las tablas 4 o 12, o fragmentos de las mismas. En una realización, una secuencia del XTEN de las tablas 4 u 11-15 puede modificarse para que incluya restos adicionales de ácido glutámico para conseguir la carga neta negativa deseada. Por consiguiente, en una realización, la invención proporciona XTEN en el que las secuencias XTEN contienen aproximadamente un 1 %, un 2 %, un 4 %, un 8 %, un 10 %, un 15 %, un 17 %, un 20 %, un 25 % o incluso aproximadamente un 30 % de ácido glutámico. Generalmente, los restos de glutámico están espaciados uniformemente a lo largo de la secuencia del XTEN. En algunos casos, el XTEN puede contener aproximadamente 10-80, o aproximadamente 15-60, o aproximadamente 20-50 restos de glutámico por 20 kDa de XTEN que pueden dar como resultado un XTEN con restos cargados que tendrían pKa muy similar, lo que puede aumentar la homogeneidad de la carga del producto y refinar su punto isoeléctrico, mejorar las propiedades fisicoquímicas de la proteína de fusión de unión resultante para, y por lo tanto, simplificar los procedimientos de purificación. En una realización, la invención contempla la incorporación de restos de ácido aspártico en el XTEN, además de ácido glutámico, con el fin de conseguir una carga neta negativa.

En otros casos, donde no se desea una carga neta, el XTEN puede seleccionarse entre, por ejemplo, componentes XTEN de AG, tales como los motivos AG de la tabla 3, o los motivos AM de la tabla 3 que aproximadamente no tienen carga neta. Algunos ejemplos no limitativos de XTEN de AG incluyen, pero sin limitación, secuencias de polipéptidos AG42, AG144, AG288, AG576 y AG864 de las tablas 4 o 14, o fragmentos de las mismas. En otra realización, el XTEN puede comprender proporciones variables de motivos AE y AG con el fin de tener una carga neta que se considere óptima para un uso dado o para mantener una propiedad fisicoquímica dada.

Sin pretender quedar ligado a teoría alguna en particular, se espera que los XTEN de las composiciones de proteína de fusión de unión con la mayor carga neta tengan menos interacciones no específicas con las diferentes superficies cargadas negativamente, tales como vasos sanguíneos, tejidos o diferentes receptores, lo que contribuiría además a un aclaramiento activo reducido. Por el contrario, se cree que el XTEN de las composiciones de proteína de fusión de unión con una carga neta baja o nula tendría un mayor grado de interacción con superficies que pueden potenciar la actividad de la proteína de unión asociada, dada la contribución conocida de las células fagocíticas en el proceso inflamatorio del pulmón.

El XTEN de las composiciones de la presente invención generalmente no tiene, o tiene un contenido bajo de, aminoácidos cargados positivamente. En algunas realizaciones, el XTEN puede tener menos de aproximadamente un 10 % de restos de aminoácidos con una carga positiva, o menos de aproximadamente un 7 %, o menos de aproximadamente un 5 %, o menos de aproximadamente un 2 %, o menos de aproximadamente un 1 % de restos de aminoácidos con una carga positiva. Sin embargo, la invención contempla construcciones donde se incorpora un número limitado de aminoácidos con carga positiva, tales como lisina, en el XTEN para permitir la conjugación entre la amina épsilon de la lisina y un grupo reactivo de un péptido, un puente conector o un grupo reactivo de un fármaco o una molécula pequeña para conjugarse con la cadena principal del XTEN. En una realización de lo anterior, el XTEN tiene entre aproximadamente 1 a aproximadamente 100 restos de lisina, o de aproximadamente 1 a aproximadamente 70 restos de lisina, o de aproximadamente 1 a aproximadamente 50 restos de lisina, o de aproximadamente 1 a aproximadamente 30 restos de lisina, o de aproximadamente 1 a aproximadamente 20 restos de lisina, o de aproximadamente 1 a aproximadamente 10 restos de lisina, o de aproximadamente 1 a aproximadamente 5 restos de lisina, o alternativamente solo un único resto de lisina. Utilizando el XTEN que contiene lisina anterior, se construyen proteínas de fusión que comprenden XTEN, una proteína de unión, más un antineoplásico útil en el tratamiento de enfermedades o trastornos, en donde el número máximo de moléculas del agente incorporado en el componente XTEN está determinado por el número de lisinas o de otros aminoácidos con cadenas laterales reactivas (por ejemplo, cisteína) incorporadas en el XTEN.

A medida que los aminoácidos hidrófobos imparten estructura a un polipéptido, la invención prevé que el contenido de aminoácidos hidrófobos en el XTEN será normalmente inferior al 5 %, o inferior al 2 %, o inferior al 1 % de contenido de aminoácidos hidrófobos. En una realización, el contenido de aminoácidos de metionina y triptófano en el componente XTEN de una proteína de fusión de unión es normalmente inferior al 5 %, o inferior al 2 %, y lo más preferentemente inferior al 1 %. En otra realización, el XTEN tendrá una secuencia que tiene menos de un 10 % de restos de aminoácidos con una carga positiva, o menos de aproximadamente un 7 %, o menos de aproximadamente un 5 %, o menos de aproximadamente un 2 % de restos de aminoácidos con una carga positiva, la suma de restos de metionina y triptófano será inferior al 2 %, y la suma de restos de asparagina y glutamina será inferior al 10 % de la secuencia total del XTEN.

7. <u>Baja inmunogenicidad</u>

En otro aspecto, la invención proporciona composiciones en las que las secuencias XTEN tienen un bajo grado de inmunogenicidad o son sustancialmente no inmunogénicas. Varios factores pueden contribuir a la baja inmunogenicidad del XTEN, por ejemplo, la secuencia no repetitiva, la conformación no estructurada, el alto grado de solubilidad, el bajo grado o la ausencia de autoagregación, el bajo grado o la ausencia de sitios proteolíticos dentro de la secuencia, y el bajo grado o la ausencia de epítopos en la secuencia del XTEN.

Los epítopos conformacionales están formados por regiones de la superficie de la proteína que están compuestas por múltiples secuencias de aminoácidos discontinuas del antígeno proteico. El plegamiento preciso de la proteína lleva estas secuencias a unas configuraciones espaciales estables bien definidas, o epítopos, que pueden ser reconocidos como "extraños" por el sistema inmunitario humoral del hospedador, lo que da como resultado la producción de anticuerpos contra la proteína o desencadena una respuesta inmunitaria mediada por células. En este último caso, la respuesta inmunitaria a una proteína en un individuo está fuertemente influenciada por el reconocimiento del epítopo de los linfocitos T, que es una función de la especificidad de unión al péptido del alotipo HLA-DR de ese individuo. El acoplamiento de un complejo peptídico del MHC de clase II por parte de un receptor de linfocitos T afín en la superficie del linfocito T, junto con la reticulación de otros correceptores dados, tales como la molécula CD4, puede inducir un estado activado dentro del linfocito T. La activación conduce a la liberación de citocinas que activan adicionalmente otros linfocitos, tales como los linfocitos B, para producir anticuerpos o activan linfocitos T citolíticos como una respuesta inmunitaria celular completa.

La capacidad de un péptido para unirse a una molécula del MHC de clase II dada para su presentación en la superficie de una APC (célula presentadora del antígeno) depende de una serie de factores; más notablemente, de su secuencia primaria. En una realización, se puede conseguir un menor grado de inmunogenicidad diseñando secuencias XTEN que resistan el procesamiento de antígeno en las células presentadora del antígeno, y/o eligiendo unas secuencias que no se unan bien a los receptores del MHC. La invención proporciona proteínas de fusión de unión con polipéptidos XTEN sustancialmente no repetitivos diseñados para reducir la unión con receptores del MHC II, así como para evitar la formación de epítopos para la unión al receptor de linfocitos T o al anticuerpo, lo que da como resultado un bajo grado de inmunogenicidad. La evitación de la inmunogenicidad es, en parte, un resultado directo de la flexibilidad conformacional de las secuencias XTEN; es decir, la ausencia de una estructura secundaria debido a la selección y el orden de los restos de aminoácidos. Por ejemplo, son de particular interés las secuencias que tienen una baja tendencia a adoptar unas conformaciones plegadas de forma compacta en disolución acuosa o en condiciones fisiológicas que podrían dar como resultado epítopos conformacionales. La administración de proteínas de fusión que comprenden XTEN, utilizando prácticas terapéuticas y de administración convencionales, generalmente no daría como resultado la formación de anticuerpos neutralizantes contra la secuencia del XTEN, y también puede reducir la inmunogenicidad del compañero de fusión del resto de direccionamiento en las composiciones de proteína de fusión de unión.

En una realización, las secuencias XTEN utilizadas en las presentes proteínas de fusión pueden estar sustancialmente exentas de epítopos reconocidos por los linfocitos T humanos. La eliminación de dichos epítopos con el fin de generar menos proteínas inmunogénicas se ha divulgado previamente; véanse, por ejemplo, el documento WO 98/52976, el documento WO 02/079232 y el documento WO 00/3317. Se han descrito ensayos para epítopos de linfocitos T humanos (Stickler, M., et al. (2003) J Immunol Methods, 281: 95-108). De particular interés son las secuencias peptídicas que se pueden oligomerizar sin generar epítopos de linfocitos T o secuencias no humanas. Esto se puede conseguir probando repeticiones directas de estas secuencias para detectar la presencia de epítopos de linfocitos T y la aparición de secuencias de 6 a 15-meros y, en particular, de 9-meros que no son humanas, y alterando a continuación el diseño de la secuencia del XTEN para eliminar o alterar la secuencia del epítopo. En una realización, las secuencias XTEN son sustancialmente no inmunogénicas por la restricción del número de epítopos del XTEN previstos para unirse a los receptores del MHC. Con una reducción en el número de epítopos capaces de unirse a receptores del MHC, hay una reducción simultánea en el potencial de activación de los linfocitos T, así como de la función auxiliar de los linfocitos T, una activación o una regulación positiva de linfocitos B reducida y una producción de anticuerpos reducida. El bajo grado de epítopos de linfocitos T predichos puede determinarse mediante algoritmos de predicción de epítopos tales como, por ejemplo, TEPITOPE (Sturniolo, T., et al. (1999) Nat Biotechnol, 17: 555-61), como se muestra en el ejemplo 59. La puntuación del TEPITOPE de un marco peptídico dado dentro de una proteína es el logaritmo de la K_d (constante de disociación, afinidad, tasa de disociación) de la unión de ese marco peptídico a múltiples de los alelos del MHC humanos más comunes, tal como se divulga en Sturniolo, T. et al (1999) Nature Biotechnology 17: 555). La puntuación varía a lo largo de al menos 20 logaritmos, de aproximadamente 10 a aproximadamente -10 (correspondiente a unas restricciones de unión de $10e^{10}$ K_d a $10e^{-10}$ K_d), y se puede reducir evitando los aminoácidos hidrófobos que puedan servir como restos de anclaje durante la presentación del péptido en el MHC, tales como M, I, L, V, F. En algunas realizaciones, un componente XTEN incorporado en una proteína de fusión de unión no tiene un epítopo de linfocitos T predicho en una puntuación umbral del TEPITOPE de aproximadamente -5, o -6, o -7, o -8, o -9, o con una puntuación del TEPITOPE de -10. Como se utiliza en el presente documento, una puntuación de "-9" sería un umbral del TEPITOPE más riguroso que una puntuación de -5.

En otra realización, las secuencias XTEN inventivas, incluidas las incorporadas en las presentes proteínas de fusión de unión, pueden volverse sustancialmente no inmunogénicas mediante la restricción de los sitios proteolíticos conocidos de la secuencia del XTEN, reduciendo el procesamiento del XTEN en pequeños péptidos que pueden unirse

a receptores del MHC II. En otra realización, la secuencia del XTEN puede volverse sustancialmente no inmunogénica mediante el uso de una secuencia que está sustancialmente desprovista de estructura secundaria, confiriendo resistencia a muchas proteasas debido a la alta entropía de la estructura. Por consiguiente, la puntuación reducida del TEPITOPE y la eliminación de los sitios proteolíticos conocidos del XTEN pueden hacer que las composiciones XTEN, incluyendo el XTEN de las composiciones de proteína de fusión de unión, sean sustancialmente incapaces de ser unidas por receptores de mamífero, incluidos los del sistema inmunitario. En una realización, un XTEN de una proteína de fusión de unión puede tener una unión K_d >100 nM a un receptor de mamífero, o superior a K_d 500 nM, o superior a K_d 1 µm frente a una superficie celular de mamífero o un receptor polipeptídico circulante.

Adicionalmente, la secuencia no repetitiva y la correspondiente ausencia de epítopos del XTEN pueden limitar la capacidad de los linfocitos B para unirse a o ser activados por el XTEN. Una secuencia repetitiva es reconocida y puede formar contactos multivalentes incluso con unos pocos linfocitos B y, como consecuencia de la reticulación de múltiples receptores independientes de linfocitos T, puede estimular la proliferación de linfocitos B y la producción de anticuerpos. Por el contrario, aunque un XTEN puede crear contactos con muchos linfocitos B diferentes a lo largo de su secuencia expandida, cada linfocito B individual puede hacer solo uno o un pequeño número de contactos con un XTEN individual debido a la falta de repetitividad de la secuencia. Como resultado, los XTEN pueden tener normalmente una tendencia mucho menor a estimular la proliferación de linfocitos B y, por lo tanto, una respuesta inmunitaria. En una realización, la proteína de fusión de unión puede tener una inmunogenicidad reducida en comparación con el resto de direccionamiento correspondiente que no está fusionado. En una realización, la administración de hasta tres dosis parenterales de una proteína de fusión de unión a un mamífero puede dar como resultado una IgG anti-proteína de fusión de unión detectable a una dilución en suero de 1:100, pero no a una dilución de 1:1000. En otra realización, la administración de hasta tres dosis parenterales de una proteína de fusión de unión a un mamífero puede dar como resultado una IgG anti-resto de direccionamiento detectable a una dilución en suero de 1:100, pero no a una dilución de 1:1000. En otra realización, la administración de hasta tres dosis parenterales de una proteína de fusión de unión a un mamífero puede dar como resultado una IgG anti-XTEN detectable a una dilución en suero de 1:100, pero no a una dilución de 1:1000. En las realizaciones anteriores, el mamífero puede ser un ratón, una rata, un conejo o un macaco cangrejero.

Una característica adicional de los XTEN con secuencias no repetitivas con respecto a secuencias con un alto grado de repetitividad puede ser que los XTEN no repetitivos formen contactos más débiles con los anticuerpos. Los anticuerpos son moléculas multivalentes. Por ejemplo, las IgG tienen dos sitios de unión idénticos y las IgM contienen 10 sitios de unión idénticos. Por lo tanto, los anticuerpos contra las secuencias repetitivas pueden formar contactos multivalentes con dichas secuencias repetitivas con alta avidez, lo que puede afectar a la potencia y/o a la eliminación de dichas secuencias repetitivas. Por el contrario, los anticuerpos contra los XTEN no repetitivos pueden producir interacciones monovalentes, lo que da como resultado una menor probabilidad de aclaramiento inmunitario, de modo que las composiciones de proteína de fusión de unión pueden permanecer en circulación durante un mayor período de tiempo.

8. <u>Aumento del radio hidrodinámico</u>

En otro aspecto, la presente invención proporciona XTEN en los que los polipéptidos del XTEN pueden tener un alto radio hidrodinámico que confiere un correspondiente factor de peso molecular aparente aumentado a la proteína de fusión que incorpora el XTEN. Como se detalla en el ejemplo 40, la unión del XTEN a secuencias de restos de direccionamiento puede dar como resultado composiciones de proteína de fusión de unión que pueden tener unos radios hidrodinámicos aumentados, un factor de peso molecular aparente aumentado y un factor de peso molecular aparente aumentado en comparación con un resto de direccionamiento no unido a un XTEN. Por ejemplo, en aplicaciones terapéuticas en las que se desea una semivida prolongada, composiciones en las que se incorpora un XTEN con un alto radio hidrodinámico a una proteína de fusión que comprende uno o más restos de direccionamiento pueden agrandar eficazmente el radio hidrodinámico de la composición más allá del tamaño de poro glomerular de aproximadamente 3-5 nm (correspondiente a un peso molecular aparente de aproximadamente 70 kDa) (Caliceti. 2003. Pharmacokinetic and biodistribution properties of poly(ethylene glycol)-protein conjugates. Adv Drug Deliv Rev 55: 1261-1277), lo que da como resultado un aclaramiento renal reducido de las proteínas circulantes. El radio hidrodinámico de una proteína está determinado por su peso molecular, así como por su estructura, incluyendo su forma y su compacidad. Sin pretender quedar ligado a teoría alguna en particular, el XTEN puede adoptar conformaciones abiertas debido a la repulsión electrostática entre cargas individuales del péptido o la flexibilidad inherente impartida por los aminoácidos particulares de la secuencia que carecen del potencial para conferir una estructura secundaria. La conformación abierta, expandida y no estructurada del polipéptido XTEN puede tener un radio hidrodinámico proporcional mayor en comparación con polipéptidos con una longitud de secuencia y/o un peso molecular comparables que tienen estructura secundaria y/o terciaria, tales como las proteínas globulares clásicas. Los métodos para determinar el radio hidrodinámico son bien conocidos en la técnica, tales como mediante el uso de una cromatografía de exclusión por tamaño (SEC), como se describe en las Patentes de EE. UU. n.º 6.406.632 y 7.294.513. Como demuestran los resultados del ejemplo 40, la adición de longitudes crecientes de XTEN a un polipéptido con carga activa da como resultado aumentos proporcionales en los parámetros del radio hidrodinámico, el factor de peso molecular aparente y el factor de peso molecular aparente, lo que permite adaptar las proteínas de fusión de unión a valores de corte deseados de los factores de peso molecular aparente característicos o de los radios hidrodinámicos. Por consiguiente, en ciertas realizaciones, la proteína de fusión de unión se puede configurar con un

XTEN de modo que la proteína de fusión pueda tener un radio hidrodinámico de al menos aproximadamente 5 nm, o al menos aproximadamente 8 nm, o al menos aproximadamente 10 nm, o 12 nm, o al menos aproximadamente 15 nm. En las realizaciones anteriores, el gran radio hidrodinámico conferido por el XTEN a una proteína de fusión de unión puede conducir a un aclaramiento renal reducido de la proteína de fusión resultante, lo que conduce a un aumento correspondiente de la semivida terminal, un aumento en el tiempo de permanencia medio y/o una disminución en la tasa de aclaramiento renal.

En otra realización, un XTEN de una longitud y secuencia elegidas puede incorporarse selectivamente en una proteína de fusión de unión para crear una proteína de fusión que tendrá, en condiciones fisiológicas, un peso molecular aparente de al menos aproximadamente 150 kDa, o al menos aproximadamente 300 kDa, o al menos aproximadamente 400 kDa, o al menos aproximadamente 500 kDa, o al menos aproximadamente 600 kDa, o al menos aproximadamente 700 kDa, o al menos aproximadamente 800 kDa, o al menos aproximadamente 900 kDa, o al menos aproximadamente 1000 kDa, o al menos aproximadamente 1200 kDa, o al menos aproximadamente 1500 kDa, o al menos aproximadamente 1800 kDa, o al menos aproximadamente 2000 kDa, o al menos aproximadamente 2300 kDa o más. En otra realización, un XTEN de una longitud y secuencia elegidas puede unirse selectivamente a un resto de direccionamiento para dar como resultado una proteína de fusión de unión que tiene, en condiciones fisiológicas, un factor de peso molecular aparente de al menos tres, alternativamente de al menos cuatro, alternativamente de al menos cinco, alternativamente de al menos seis, alternativamente de al menos ocho, alternativamente de al menos 10, alternativamente de al menos 15, o un factor de peso molecular aparente de al menos 20 o mayor. En otra realización, la proteína de fusión de unión tiene, en condiciones fisiológicas, un factor de peso molecular aparente que es de aproximadamente 4 a aproximadamente 20, o es de aproximadamente 6 a aproximadamente 15, o es de aproximadamente 8 a aproximadamente 12, o es de aproximadamente 9 a aproximadamente 10 con respecto al peso molecular real de la proteína de fusión.

(d) Restos de direccionamiento

En otro aspecto de la invención, se divulgan restos de direccionamiento que pueden unirse a uno o más XTEN, lo que da como resultado composiciones de proteína de fusión de unión monómeras. "Restos de direccionamiento", como se utiliza en el presente documento, se refiere a polipéptidos que tienen afinidad de unión específica por un ligando diana tal como citocinas, quimiocinas, receptores de citocinas, receptores de quimiocinas, hormonas, receptores de superficie celular o antígenos o glicoproteínas, oligonucleótidos, sustratos enzimáticos, determinantes antigénicos u otros sitios de unión que pueden estar presentes en la circulación, o en la superficie o en el citoplasma de una célula diana. Algunas dianas ilustrativas no limitativas a las que se dirigen los restos de direccionamiento de las presentes composiciones se han divulgado anteriormente; por ejemplo, dianas seleccionadas entre la tabla 1 y la tabla 2. La invención proporciona múltiples categorías de restos de direccionamiento que se pueden unir a uno o más XTEN en diferentes configuraciones, lo que da como resultado las composiciones de proteína de fusión de unión inventivas. Como se describe más completamente a continuación, los restos de direccionamiento pueden derivar de o basarse en secuencias de anticuerpos, fragmentos de anticuerpos, receptores, dominios de unión similares a inmunoglobulina, o pueden ser completamente sintéticos. Las proteínas de fusión de unión pueden comprender uno o más sitios de unión al antígeno funcionales, haciendo estos últimos que la proteína de fusión de unión sea "multivalente". Un "sitio de unión al antígeno" de una proteína de fusión de unión es uno que es capaz de unirse a un antígeno diana con al menos una porción de la afinidad de unión del anticuerpo o receptor progenitor del que deriva el sitio de unión al antígeno. El propio sitio de unión al antígeno puede estar compuesto por más de un dominio de unión, unidos entre sí en las proteínas de fusión de unión. "Dominio de unión" significa una secuencia polipeptídica capaz de unirse a un antígeno o ligando, pero que puede requerir dominios de unión adicionales para unirse y/o secuestrar realmente el antígeno o ligando. Una CDR de un anticuerpo es un ejemplo de un dominio de unión. "Anticuerpo" se utiliza en toda la memoria descriptiva como un ejemplo prototípico de un resto de direccionamiento, pero no pretende ser limitativo.

Los métodos para medir la afinidad de unión y/u otra actividad biológica de las composiciones de proteína de fusión de unión de la invención pueden ser los divulgados en el presente documento o métodos generalmente conocidos en la técnica. Además, las propiedades fisicoquímicas de la proteína de fusión de unión se pueden medir para determinar el grado de solubilidad, la estructura y la retención de estabilidad. Se realizan ensayos que permiten la determinación de las características de unión de los restos de direccionamiento hacia un ligando, incluida la constante de disociación de unión (K_d, K_{on} y K_{off}), la semivida de disociación del complejo ligando-receptor, así como la actividad de la proteína de fusión de unión para inhibir la actividad biológica del ligando secuestrado en comparación con el ligando libre (valores de la CI_{50}). El término "K_d", como se utiliza en el presente documento, pretende referirse a la constante de disociación de una interacción anticuerpo-antígeno en particular, como es conocido en la técnica, y se aplicaría como parámetro de la afinidad de unión de un resto de direccionamiento a su ligando afín para las presentes composiciones. El término "K_{on}", como se utiliza en el presente documento, pretende referirse a la constante de tasa de asociación de un anticuerpo al antígeno para formar el complejo anticuerpo/antígeno, como es conocido en la técnica. El término "K_{off}", como se utiliza en el presente documento, pretende referirse a la constante de tasa de disociación para la disociación de un anticuerpo del complejo anticuerpo/antígeno, como es conocido en la técnica. El término "CI_{50}" se refiere a la concentración necesaria para inhibir la mitad de la respuesta biológica máxima del agonista del ligando, y se determina generalmente mediante ensayos de unión competitiva.

Se pueden utilizar técnicas tales como citometría de flujo o resonancia de plasmón superficial para detectar los eventos de unión. Los ensayos pueden comprender antígenos solubles o moléculas receptoras, o pueden determinar la unión a receptores expresados por células. Dichos ensayos pueden incluir ensayos basados en células, que incluyen ensayos de proliferación, muerte celular, apoptosis y migración celular. La afinidad de unión de las presentes composiciones por los ligandos diana se puede ensayar utilizando ensayos de unión o de unión competitiva, tales como ensayos Biacore con receptores unidos a chip o proteínas de unión o ensayos ELISA, como se describe en la Patente de EE. UU. 5.534.617, ensayos descritos en los Ejemplos del presente documento, ensayos de radiorreceptores u otros ensayos conocidos en la técnica. La constante de afinidad de unión se puede determinar a continuación utilizando métodos convencionales, tales como un análisis de Scatchard, como han descrito van Zoelen, et al., Trends Pharmacol Sciences (1998) 19)12): 487, u otros métodos conocidos en la técnica. Además, las bibliotecas de variantes de secuencia de restos de direccionamiento se pueden comparar con los correspondientes anticuerpos naturales o progenitores utilizando un ensayo ELISA de unión competitiva para determinar si tienen la misma especificidad y afinidad de unión que el anticuerpo progenitor o alguna fracción de los mismos, de modo que sean adecuados para su inclusión en las proteínas de fusión de unión. Los resultados de dichos ensayos se pueden utilizar en un proceso iterativo de modificación de secuencia de los restos de direccionamiento, seguido de ensayos de unión y caracterización fisicoquímica para guiar el proceso mediante el cual se seleccionan construcciones específicas con las propiedades deseadas.

En una realización, la invención proporciona proteínas de fusión de unión aisladas que inhiben competitivamente la unión de un anticuerpo a un ligando diana, según se determina mediante cualquier método conocido en la técnica para determinar la unión competitiva, tal como los inmunoensayos descritos en el presente documento. El anticuerpo puede incluir el anticuerpo progenitor del que deriva el resto de direccionamiento o un control positivo conocido que se sabe que se une al epítopo o ligando diana. En las realizaciones preferidas, la proteína de fusión de unión inhibe competitivamente la unión del control positivo al ligando en al menos un 90 %, al menos un 80 %, al menos un 70 %, al menos un 60 % o al menos un 50 % en un ensayo de unión competitiva frente al control positivo.

La invención proporciona proteínas de fusión de unión aisladas en las que la afinidad de unión de uno o más restos de direccionamiento por ligandos diana puede ser al menos aproximadamente un 1 %, o al menos aproximadamente un 10 %, o al menos aproximadamente un 20 %, o al menos aproximadamente un 30 %, o al menos aproximadamente un 40 %, o al menos aproximadamente un 50 %, o al menos aproximadamente un 60 %, o al menos aproximadamente un 70 %, o al menos aproximadamente un 80 %, o al menos aproximadamente un 90 %, o al menos aproximadamente un 95 %, o al menos aproximadamente un 99 %, o al menos aproximadamente un 99,9 % o más de la afinidad de un anticuerpo progenitor no unido al XTEN por el receptor o ligando diana. En una realización, la K_d entre el uno o más restos de direccionamiento de la presente proteína de fusión de unión y un ligando diana es menor de aproximadamente 10^{-4} M, alternativamente menor de aproximadamente 10^{-5} M, alternativamente menor de aproximadamente 10^{-6} M, alternativamente menor de aproximadamente 10^{-7} M, alternativamente menor de aproximadamente 10^{-8} M, alternativamente menor de aproximadamente 10^{-9} M, o menor de aproximadamente 10^{-10} M, o menor de aproximadamente 10^{-11} M, o menor de aproximadamente 10^{-12} M. En la realización anterior, la afinidad de unión de la proteína de fusión de unión hacia la diana se caracterizaría como "específica". La invención contempla proteínas de fusión de unión que comprenden dos o más restos de direccionamiento en los que las afinidades de unión por los restos de direccionamiento respectivos pueden estar independientemente entre los intervalos de valores de los anteriores. En cualquiera de las realizaciones anteriores del párrafo, el uno o más restos de direccionamiento de las presentes proteínas de fusión de unión se unen específicamente a una diana de la tabla 1 o la tabla 2.

Las proteínas de fusión de unión de la presente invención también pueden describirse o especificarse en términos de la reactividad cruzada del resto de direccionamiento. En una realización, la invención proporciona proteínas de fusión de unión que no se unen a ningún otro análogo, ortólogo u homólogo de una diana divulgada en el presente documento. En una realización, las proteínas de fusión de unión pueden unirse a polipéptidos con al menos un 95 %, al menos un 90 %, al menos un 85 %, al menos un 80 %, al menos un 75 %, al menos 70 %, al menos 65 %, al menos 60 %, al menos 55 % y al menos 50 % de identidad de secuencia (como se calcula utilizando métodos conocidos en la técnica y descritos en el presente documento) a una diana polipeptídica descrito en el presente documento.

Las proteínas de fusión de unión de la presente invención pueden actuar como agonistas o como antagonistas. Por ejemplo, la presente invención incluye proteínas de fusión de unión que comprenden restos de direccionamiento que alteran las interacciones receptor/ligando, ya sea parcial o completamente. La invención presenta tanto BFP específicas del receptor como BFP específicas del ligando. La invención también presenta proteínas de fusión de unión específicas del receptor que no impiden la unión del ligando, sino que impiden la activación del receptor. La activación del receptor, tal como la señalización celular, puede determinarse mediante técnicas descritas en el presente documento o conocidas de otro modo en la técnica. Por ejemplo, la activación del receptor puede determinarse detectando la fosforilación (por ejemplo, tirosina o serina/treonina) del receptor o de su sustrato mediante inmunoprecipitación seguida de técnicas de análisis convencionales de inmunoelectrotransferencia. En realizaciones específicas, se proporcionan proteínas de fusión de unión que pueden unirse e inhibir la actividad del ligando o del receptor en al menos un 90 %, al menos un 80 %, al menos un 70 %, al menos un 60 % o al menos un 50 % de la actividad en comparación con la actividad en ausencia de la proteína de fusión de unión.

La invención proporciona proteínas de fusión de unión específicas del receptor que impiden tanto la unión del ligando como la activación del receptor, así como proteínas de fusión de unión que reconocen el complejo receptor-ligando, pero, preferentemente, no reconocen específicamente el receptor no unido o el ligando no unido. En una realización, la invención proporciona proteínas de fusión de unión neutralizantes que se unen al ligando, formando así un complejo neutralizante que impide la unión del ligando al receptor, o, en otros casos, puede unirse al ligando pero no impide que el ligando se una al receptor, sin embargo, dan como resultado una activación reducida del receptor en comparación con el ligando no complejado. Además, se incluyen en la invención proteínas de fusión de unión que activan el receptor. Estas BFP pueden actuar como agonistas del receptor, es decir, potenciar o activar todas o un subconjunto de las actividades biológicas de la activación del receptor mediada por ligando.

En otra realización, la invención proporciona proteínas de fusión de unión aisladas en las que la proteína de fusión está diseñada para unirse con alta afinidad a un receptor diana, lo que da como resultado una actividad antagónica para el ligando natural. En dichos casos, las BFP pueden tener afinidad, pero ninguna eficacia, por sus receptores afines, de manera que su unión alterará la interacción e inhibirá la función de un agonista o un agonista inverso en los receptores. Normalmente, dicha actividad antagonista será de un tipo competitivo, y se puede determinar una K_i. Un ejemplo no limitativo de una BFP antagonista es una proteína de fusión que comprende un resto de direccionamiento configurado para unirse a un receptor de la IL-1 (IL-1R) de modo que la composición unida interfiere sustancialmente con la unión de la IL-1 α y/o la IL-1β al receptor de la IL-1. En ciertos casos, la interferencia por parte de una proteína de fusión de unión antagonista (tal como, pero sin limitación, una proteína de fusión de unión anti-IL-1R) con la unión del ligando natural a su receptor afín puede ser al menos aproximadamente un 1 %, o aproximadamente un 10 %, o aproximadamente un 20 %, o aproximadamente un 30 %, o aproximadamente un 40 %, o aproximadamente un 50 %, o aproximadamente un 60 %, o aproximadamente un 70 %, o aproximadamente un 80 %, o aproximadamente un 90 %, o aproximadamente un 95 %, o aproximadamente un 99 % o más. En otras realizaciones, la invención proporciona proteínas de fusión de unión aisladas (tales como, pero sin limitación, proteína de fusión de unión anti-IL-1R) en donde la unión de la proteína de fusión aislada a un receptor celular provoca menos del 20 %, o menos del 10 %, o menos del 5 % de activación de las vías de señalización de la célula con el antagonista de proteína de fusión de unión unido en comparación con las provocadas por el ligando natural.

En una realización, la invención proporciona proteínas de fusión de unión aisladas que comprenden restos de direccionamiento dirigidos a una o más dianas de citocinas, proteínas relacionadas con citocinas, receptores de citocinas, quimiocinas, receptores de quimiocinas, receptores de quimiocinas, receptores de la superficie celular, hormonas o proteínas o péptidos circulantes similares, oligonucleótidos o sustratos enzimáticos. En una realización, el uno o más restos de direccionamiento pueden tener una afinidad de unión específica hacia dianas seleccionadas entre, pero sin limitación, las dianas de la tabla 1. En otra realización, el uno o más restos de direccionamiento pueden tener afinidad de unión específica hacia dianas seleccionadas entre, pero sin limitación, las dianas de antígenos asociados a tumores de la tabla 2.

La presente invención proporciona una variedad de configuraciones de proteína de fusión de unión en las que las variaciones se basan en la inclusión del tipo y la posición relativa o el número de dominios de unión, así como la inclusión de conectores de longitud predeterminada y una o más secuencias XTEN. Por diseño, las composiciones de proteína de fusión de unión resultantes pueden ser monómeras o multivalentes, pueden unirse a un solo ligando o antígeno, o ser multiméricas en cuanto al número de unidades de unión englobadas en la proteína de fusión.

En una realización, la invención proporciona proteínas de fusión de unión que comprenden restos de direccionamiento capaces de unirse a un único ligando. En otra realización, las proteínas de fusión de unión de la invención son multivalentes, y los restos de direccionamiento se unen específicamente a al menos dos antígenos o ligandos diana diferentes ("bifuncionales" o "biespecíficos"), o epítopos diferentes en el mismo ligando. Las proteínas de fusión de unión multivalentes pueden diseñarse para que sean bifuncionales, ya que pueden incorporar dominios de unión heterólogos de diferentes anticuerpos "progenitores" y unirse a dos ligandos o antígenos diferentes con el fin de conseguir mejor una respuesta farmacológica deseada; por ejemplo, dimerización de receptores en una superficie celular diana que conduce a la señalización celular o, alternativamente, muerte celular, o modulación de una función biológica de una o más dianas. Se espera que la proteína de fusión de unión multimérica que conduce a la muerte celular, ya sea desencadenando una apoptosis o una necrosis, tenga utilidad, particularmente, en el tratamiento de enfermedades oncológicas. Algunos ejemplos no limitativos de pares de dianas contemplados como adecuados para proteínas de fusión de unión bifuncionales multivalentes incluyen: IGF1 y IGF2; IGF1/2 y Erb2B; VEGFR y EGFR; CD20 y CD3, CD138 y CD20, CD38 y CD20, CD38 y CD138, CD40 y CD20, CD138 y CD40, CD38 y CD40.

En una realización, las proteínas de fusión de unión de la divulgación se unen específicamente a al menos dos citocinas, linfocinas, monocinas y/u hormonas polipeptídicas. Algunos ejemplos no limitativos de pares de dianas a los que pueden unirse proteínas de fusión de unión bifuncionales se seleccionan entre, pero sin limitación, IL-1α e IL-1β; IL-12 e IL-18, TNFα e IL-23, TNFα e IL-13; TNF e IL-18; TNF e IL-12; TNF e IL-1beta; TNF y MIF; TNF e IL-17; y TNF e IL-15; TNF y VEGF; VEGFR y EGFR; IL-13 e IL-9; IL-13 e IL-4; IL-13 e IL-5; IL-13 e IL-25; IL-13 y TARC; IL-13 y MDC; IL-13 y MIF; IL-13 y TGF-β; IL-13 y agonista del LHR; IL-13 y CL25; IL-13 y SPRR2a; IL-13 y SPRR2b; IL-13 y ADAM8; y TNFα y PGE4, IL-13 y PED2, TNF y PEG2.

Algunos ejemplos de otros pares de dianas adecuados para proteínas de fusión de unión bifuncionales multivalentes incluyen, pero sin limitación, CD19 y CD20; CD-8 e IL-6; PDL-1 y CTLA-4; CTLA-4 y BTNO2; CSPGs y RGM A; IL-12 e IL-18; IL-12 y TWEAK; IL-13 y ADAM8; IL-13 y CL25; IL-13 e IL-1beta; IL-13 e IL-25; IL-13 e IL-4; IL-13 e IL-5; IL-13 e IL-9; IL-13 y agonista del LHR; IL-13 y MDC; IL-13 y MIF; IL-13 y PED2; IL-13 y SPRR2a; IL-13 y SPRR2b; IL-13 y TARC; IL-13 y TGF-β; IL-1α e IL-1β; MAG y RGM A; NgR y RGM A; NogoA y RGM A; OMGp y RGM A; RGM A y RGM B; Te38 y TNFα; TNFα e IL-12; TNFα e IL-12p40; TNFα e IL-13; TNFα e IL-15; TNFα e IL-17; TNFα e IL-18; TNFα e IL-1beta; TNFα e IL-23; TNFα y MIF; TNFα y PEG2; TNFα y PGE4; TNFα y VEGF; TNFα y ligando RANK; TNFα y Blys; TNFα y GP130; TNFα y CD-22; y TNFα y CTLA-4.

En otra realización, las proteínas de fusión de unión de la divulgación se unen específicamente a pares de dianas seleccionadas entre, pero sin limitación, CD138 y CD20; CD138 y CD40; CD19 y CD20; CD20 y CD3; CD38 & CD138 CD38 y CD20; CD38 y CD40; CD40 y CD20; CD-8 e IL-6; CSPGs y RGM A; CTLA4 y BTNO2; IGF1 y IGF2; IGF1/2 y Erb2B; IL-12 y TWEAK; IL-13 e IL-1β; MAG y RGM A; NgR y RGM A; NogoA y RGM A; OMGp y RGM A; PDL-1 y CTLA4; RGM A y RGM B; Te38 y TNFα; TNFα y Blys; TNFα y CD-22; TNFα y CTLA-4; TNFα y GP130; TNFα e IL-12p40; y TNFα y ligando RANK.

Los restos de direccionamiento de las proteínas de fusión de unión pueden derivar de uno o más fragmentos de diferentes anticuerpos monoclonales conocidos en la técnica. Algunos ejemplos no limitativos de dichos anticuerpos monoclonales incluyen, pero sin limitación, anticuerpo anti-TNF (Pat. de EE. UU. N.º 6.258.562), anticuerpo anti-IL-12 y o anti-IL-12p40 (Pat. de EE. UU. N.º 6.914.128); anticuerpo anti-IL-18 (US 2005/0147610 A1), anti-RANKL (Patente de EE. UU. n.º 7.411.050), anti-C5, anti-CBL, anti-CD147, anti-gp120, anti-VLA4, anti-CD11a, anti-CD18, anti-VEGF, anti-CD40L, anti-Id, anti-ICAM-1, anti-CXCL13, anti-CD2, anti-EGFR, anti-TGF-beta 2, anti-E-selectina, anti-Factor VII, anti-Her2/neu, anti-Fgp, anti-CD11/18, anti-CD14, anti-ICAM-3, anti-CD80, anti-CD4, anti-CD3, anti-CD23, anti-integrina beta2, anti-alfa4beta7, anti-CD52, anti-HLA DR, anti-CD22, anti-CD20, anti-MIF, anti-CD64 (FcR), anti-TCR alfa beta, anti-CD2, anti-Hep B, anti-CA 125, anti-EpCAM, anti-gp120, anti-CMV, anti-gpIIbIIIa, anti-IgE, anti-CD25, anti-CD33, anti-HLA, anti-integrina VNR, anti-IL-1alfa, anti-IL-1beta, anti-receptor de la IL-1, anti-receptor de la IL-2, anti-IL-4, anti-receptor de la IL4, anti-IL5, anti-receptor de la IL-5, anti-IL-6, anti-IL-8, anti-IL-9, anti-IL-13, anti-receptor de la IL-13, anti-IL-17 y anti-IL-23 (véase Presta LG. 2005 Selection, design, and engineering of therapeutic antibodies J Allergy Clin Immunol. 116: 731-6 y Clark, M., "Antibodies for Therapeutic Applications", Department of Pathology, Cambridge University, Reino Unido, 15 de octubre de 2000, publicado en línea en la página de presentación de M. Clark en el sitio web del Department of Pathology, Cambridge University).

En algunas realizaciones, los restos de direccionamiento derivan de uno o más fragmentos de anticuerpos monoclonales terapéuticos aprobados para usar en seres humanos o anticuerpos que han demostrado eficacia en ensayos clínicos o modelos preclínicos establecidos de enfermedades, trastornos o afecciones. Dichos anticuerpos terapéuticos incluyen, pero sin limitación, rituximab, IDEC/Genentech/Roche (véase, por ejemplo, la patente de EE. UU. N.º 5.736.137), un anticuerpo anti-CD20 quimérico utilizado en el tratamiento de muchos linfomas, leucemias y algunos trastornos autoinmunitarios; ofatumumab, un anticuerpo anti-CD20 aprobado para usar en la leucemia linfocítica crónica, y en desarrollo para el linfoma folicular no hodgkiniano, el linfoma difuso de linfocitos B grandes, la artritis reumatoide y la esclerosis múltiple remitente recurrente, que está siendo desarrollado por GlaxoSmithKline; lucatumumab (HCD122), un anticuerpo anti-CD40 desarrollado por Novartis para el linfoma no hodgkiniano o de Hodgkin (véase, por ejemplo, la Pat. de EE. UU. N.º 6.899.879), AME-133, un anticuerpo desarrollado por Applied Molecular Evolution que se une a células que expresan el CD20 para tratar el linfoma no hodgkiniano, veltuzumab (hA20), un anticuerpo desarrollado por Immunomedics, Inc. que se une a células que expresan el CD20 para tratar el púrpura trombocitopénico inmunitario, HumaLYM desarrollado por Intracel para el tratamiento del linfoma de linfocitos B de bajo grado, y ocrelizumab, desarrollado por Genentech, que es un anticuerpo monoclonal anti-CD20 para el tratamiento de la artritis reumatoide (véase, por ejemplo, la Solicitud de Patente de EE. UU. 20090155257), trastuzumab (véase, por ejemplo, la Pat. de EE. UU. N.º 5.677.171), un anticuerpo anti-Her2/neu humanizado aprobado para tratar el cáncer de mama desarrollado por Genentech; pertuzumab, un anticuerpo inhibidor de la dimerización anti-Her2 desarrollado por Genentech en el tratamiento de cánceres de próstata, mama y ovario; (véase, por ejemplo, la patente de EE. UU. N.º 4.753.894); cetuximab, un anticuerpo anti-EGRF utilizado para tratar el cáncer colorrectal metastásico con KRAS natural que expresa el receptor del factor de crecimiento epidérmico (EGFR), y el cáncer de cabeza y cuello, desarrollado por ImClone y BMS (véase la patente de EE. UU. N.º 4.943.533; documento PCT WO 96/40210); panitumumab, un anticuerpo monoclonal completamente humano específico para el receptor del factor de crecimiento epidérmico (también conocido como receptor del EGF, EGFR, ErbB-1 y Her1, actualmente comercializado por Amgen para el tratamiento del cáncer colorrectal metastásico (véase la patente de EE. UU. N.º 6.235.883); zalutumumab, un anticuerpo monoclonal de IgG1 completamente humano desarrollado por Genmab dirigido al receptor del factor de crecimiento epidérmico (EGFR) para el tratamiento del carcinoma epidermoide de cabeza y cuello (véase, por ejemplo, la Pat. de EE. UU. N.º 7.247.301); nimotuzumab, un anticuerpo quimérico contra el EGFR desarrollado por Biocon, YM Biosciences, Cuba y Oncoscience, Europa) en el tratamiento de carcinomas epidermoides de cabeza y cuello, cáncer de nasofaringe y glioma (véase, por ejemplo, la patente de EE. UU. N.º 5.891.996; la Pat. de EE. UU. N.º 6.506.883); alemtuzumab, un anticuerpo monoclonal humanizado contra el CD52 comercializado por Bayer Schering Pharma para el tratamiento de la leucemia linfocítica crónica (LLC), el linfoma cutáneo de linfocitos T (LCLT) y el linfoma de linfocitos T; muromonab-CD3, un anticuerpo anti-CD3 desarrollado por Ortho Biotech/Johnson & Johnson utilizado como inmunosupresor biológico administrado para reducir el rechazo agudo en pacientes con trasplantes de órganos; ibritumomab tiuxetán, un anticuerpo monoclonal anti-CD20

desarrollado por IDEC/Schering AG como tratamiento de algunas formas de linfoma no hodgkiniano de linfocitos B; gemtuzumab ozogamicina, un anticuerpo anti-CD33 (proteína p67) unido al quelante citotóxico tiuxetán, al que se une un isótopo radiactivo, desarrollado por Celltech/Wyeth utilizado para tratar la leucemia mielógena aguda; alefacept, una fusión Fc anti-LFA-3 desarrollada por Biogen que se utiliza para controlar la inflamación en la psoriasis en placas de moderada a grave; abciximab, fabricado a partir de los fragmentos Fab de un anticuerpo contra el receptor IIb/IIIa de la membrana plaquetaria desarrollado por Centocor/Lilly como inhibidor de la agregación plaquetaria utilizado principalmente durante y después de procedimientos coronarios; basiliximab, un anticuerpo monoclonal quimérico ratón-humano contra la cadena α (CD25) del receptor IL-2 de los linfocitos T, desarrollado por Novartis, utilizado para prevenir el rechazo en el trasplante de órganos; palivizumab, desarrollado por Medimmune; infliximab (REMICADE), un anticuerpo anti-TNFalfa desarrollado por Centocor/Johnson y Johnson, adalimumab (HUMIRA), un anticuerpo anti-TNFalfa desarrollado por Abbott, HUMICADE, un anticuerpo anti-TNFalfa desarrollado por Celltech, etanercept (ENBREL), una fusión Fc anti-TNFalfa desarrollada por Immunex/Amgen, ABX-CBL, un anticuerpo anti-CD147 desarrollado por Abgenix, ABX-IL8, un anticuerpo anti-IL8 desarrollado por Abgenix, ABX-MA1, un anticuerpo anti-MUC18 desarrollado por Abgenix, pemtumomab (R1549, 90Y-muHMFG1), un anti-MUC1 en desarrollo por Antisoma, Therex (R1550), un anticuerpo anti-MUC1 desarrollado por Antisoma, AngioMab (AS1405), desarrollado por Antisoma, HuBC-1, desarrollado por Antisoma, tioplatino (AS1407) desarrollado por Antisoma, ANTEGREN (natalizumab), un anticuerpo anti-alfa-4-beta-1 (VLA4) y alfa-4-beta-7 desarrollado por Biogen, AcM VLA-1, un anticuerpo anti-integrina VLA-1 desarrollado por Biogen, AcM LTBR, un anticuerpo anti-receptor de linfotoxina beta (LTBR) desarrollado por Biogen, CAT-152, un anticuerpo anti-TGF-Ø2 desarrollado por Cambridge Antibody Technology, J695, un anticuerpo anti-IL-12 desarrollado por Cambridge Antibody Technology y Abbott, CAT-192, un anticuerpo anti-TGFβ1 desarrollado por Cambridge Antibody Technology y Genzyme, CAT-213, un anticuerpo anti-Eotaxin! desarrollado por Cambridge Antibody Technology, LYMPHOSTAT-B, un anticuerpo anti-Blys desarrollado por Cambridge Antibody Technology y Human Genome Sciences Inc, TRAIL-R1mAb, un anticuerpo anti-TRAIL-R1 desarrollado por Cambridge Antibody Technology y Human Genome Sciences, Inc., bevacizumab (AVASTIN, rhuMAb-VEGF), un anticuerpo anti-VEGF desarrollado por Genentech, HERCEPTIN, un anticuerpo contra la familia de receptores HER desarrollado por Genentech, anti-factor tisular (ATF), un anticuerpo anti-factor tisular desarrollado por Genentech, XOLAIR (omalizumab), un anticuerpo anti-IgE desarrollado por Genentech, anticuerpo MLN-02 (antes LDP-02), desarrollado por Genentech y Millennium Pharmaceuticals, HUMAXCD4®, un anticuerpo anti-CD4 desarrollado por Genmab, tocilizuma, y anticuerpo anti-IL6R desarrollado por Chugai, HUMAX-IL15, un anticuerpo anti-IL15 desarrollado por Genmab y Amgen, HUMAX-Inflam, desarrollado por Genmab y Medarex, HUMAX-Cancer, un anticuerpo anti-Heparanasa I desarrollado por Genmab y Medarex y Oxford GlycoSciences, HUMAX- Lymphoma, desarrollado por Genmab y Amgen, HUMAX-TAC, desarrollado por Genmab, IDEC-131, y anticuerpo anti-CD40L desarrollado por IDEC Pharmaceuticals, IDEC-151 (clenoliximab), un anticuerpo anti-CD4 desarrollado por IDEC Pharmaceuticals, IDEC-114, un anticuerpo anti-CD80 desarrollado por IDEC Pharmaceuticals, IDEC-152, un anti-CD23 desarrollado por IDEC Pharmaceuticals, anticuerpos anti-factor de migración de macrófagos (MIF) desarrollados por IDEC Pharmaceuticals, BEC2, un anticuerpo anti-idiotípico desarrollado por Imclone, IMC-1C11, un anticuerpo anti-KDR desarrollado por Imclone, DC101, un anticuerpo anti-flk-1 desarrollado por Imclone, anticuerpos anti-cadherina VE desarrollados por Imclone, CEA-CIDE (labetuzumab), un anticuerpo anti-antígeno carcinoembrionario (CEA) desarrollado por Immunomedics, Yervoy (ipilimumab), un anticuerpo anti-CTLA4 desarrollado por Bristol-Myers Sequibb para el tratamiento del melanoma, LYMPHOCIDE (epratuzumab), un anticuerpo anti-CD22 desarrollado por Immunomedics, AFP-Cide, desarrollado por Immunomedics, MyelomaCide, desarrollado por Immunomedics, LkoCide, desarrollado por Immunomedics, ProstaCide, desarrollado por Immunomedics, MDX-010, un anticuerpo anti-CTLA4 desarrollado por Medarex, MDX-060, un anticuerpo anti-CD30 desarrollado por Medarex, MDX-070 desarrollado por Medarex, MDX-018 desarrollado por Medarex, OSIDEM (IDM-1), y anticuerpo anti-Her2 desarrollado por Medarex e Immuno-Designed Molecules,HUMAX®-CD4, un anticuerpo anti-CD4 desarrollado por Medarex y Genmab, HuMax-IL15, un anticuerpo anti-IL15 desarrollado por Medarex y Genmab, CNTO 148, un anticuerpo anti-TNFa desarrollado por Medarex y Centocor/J&J, CNTO 1275, un anticuerpo anticitocinas desarrollado por Centocor/J&J, MOR101 y MOR102, anticuerpos anti-molécula de adhesión intercelular-1 (ICAM-1) (CD54) desarrollados por MorphoSys, MOR201, un anticuerpo anti-receptor 3 del factor de crecimiento de fibroblastos (FGFR-3) desarrollado por MorphoSys, tremelimumab, un anticuerpo anti-CTLA-4 desarrollado por Pfizer, visilizumab, un anticuerpo anti-CD3 desarrollado por Protein Design Labs, HUZAF, un anticuerpo anti-interferón gamma desarrollado por Protein Design Labs, anti-integrina 5β1, desarrollado por Protein Design Labs, anti-IL-12, desarrollado por Protein Design Labs, ING-1, un anticuerpo anti-Ep-CAM desarrollado por Xoma,XOLAIR® (omalizumab) un anticuerpo humanizado anti-IgE desarrollado por Genentech y Novartis, y MLN01, un anticuerpo anti-integrina beta2 desarrollado por Xoma. Las secuencias de los anteriores anticuerpos se pueden obtener a partir de bases de datos, patentes o referencias bibliográficas disponibles públicamente.

1. <u>Restos de direccionamiento ilustrativos</u>

La siguiente sección proporciona una lista y una descripción no limitativas de restos de direccionamiento ilustrativos y su uso en proteínas de fusión de unión.

Anti-Her2:

En el presente documento se describe una proteína de fusión de unión anti-Her2 aislada. "Anti-Her2" significa un resto de direccionamiento que se une específicamente al dominio extracelular del receptor HER2/neu (también conocido

como proteína erbB-2), que incluye anticuerpos, fragmentos de anticuerpos, dímeros de fragmentos, trampas y otros polipéptidos con afinidad de unión por el dominio IV de la proteína HER2/erbB-2. El gen que codifica el HER2 se encuentra en la banda q21 del cromosoma 17, genera un ARN mensajero (ARNm) de 4,8 kb, y la proteína codificada por el gen HER2 es de 185 000 Dalton. En sujetos normales, los ligandos que se unen al receptor HER2 promueven la dimerización con otros receptores, lo que da como resultado la transducción de señales y la activación de la ruta PI3K/Akt y la ruta MAPK.

En aproximadamente un 25 % de los cánceres de mama, el gen *HER2* está amplificado de 2 a más de 20 veces en cada núcleo de las células tumorales con respecto al número de copias del cromosoma 17. La amplificación del gen *HER2* impulsa la expresión de proteínas, y el aumento resultante en el número de receptores en la superficie de las células tumorales promueve la activación del receptor, lo que conduce a la señalización, la división celular excesiva y la formación de tumores (Hicks, DG et al., HER2+ breast cancer: review of biologic relevance and optimal use of diagnostic tools. Am J Clin Pathol. (2008) 129 (2): 263-73).

El anti-Her2 utilizado como compañero de fusión con XTEN crea una composición de proteína de fusión de unión que puede tener utilidad terapéutica cuando se administra a un sujeto uniéndose al dominio extracelular del segmento extracelular del receptor HER2/neu. Dicha unión puede interferir en la dimerización del receptor y la activación resultante de la función intrínseca de tirosina cinasa del EGFR (Yarden et al, Biochemistry, (1988), 27, 3114-3118; Schlessinger, Biochemistry, (1988), 27, 3119-3123), con el resultado de que las células con receptores unidos experimentan una detención durante la fase G1 del ciclo celular, por lo que hay una proliferación reducida de células tumorales, así como supresión de la angiogénesis.

Un objeto de la divulgación es proporcionar novedosas proteínas de fusión de unión anti-Her2 que comprenden uno o más restos de unión que se unen específicamente a la proteína erbB-2 y que no se unen sustancialmente a células humanas normales, que pueden utilizarse para el tratamiento o la prevención de células tumorales que expresan la erbB-2 o para la detección inmunológica de células tumorales que expresan la erbB-2. Los restos de la CDR y la FR de un anticuerpo HER2 humanizado se han comunicado en Carter et al., Proc. Natl. ACAD. SCI. USA, 89: 4285 (1992). En una realización, las composiciones de anticuerpos anti-Her2 comprenden un único resto de direccionamiento anti-Her2 unido a al menos un primer XTEN. En otra realización, las composiciones anti-Her2 comprenden un primer y un segundo resto de direccionamiento anti-Her2, que puede ser el mismo o que puede unirse a diferentes epítopos de la proteína erbB-2. En una realización, el componente anti-Her2 de una proteína de fusión de unión comprende una o más regiones determinantes de la complementariedad (CDR) de trastuzumab capaces de unirse al dominio IV del segmento extracelular del receptor HER2/neu.

En el presente documento se describe un método para inhibir el crecimiento de células tumorales mediante la administración a un paciente de una cantidad terapéuticamente eficaz de una composición de proteína de fusión de unión anti-Her2 capaz de inhibir la función del receptor HER2. En el presente documento se describe la administración de una cantidad terapéuticamente eficaz de composición anti-Her2 capaz de inhibir la función del receptor del factor de crecimiento, y una cantidad terapéuticamente eficaz de un factor citotóxico. En el presente documento se describen métodos para el tratamiento y/o la prevención de tumores que sobreexpresan el receptor de la erbB-2 que comprenden la administración de una cantidad antitumoral eficaz de al menos una de las proteínas de fusión anti-Her2 divulgadas capaces de unirse a células cancerosas asociadas por la sobreexpresión de la proteína erbB-2. En el presente documento se describe un método para el tratamiento y/o la prevención de tumores que sobreexpresan el receptor de la erbB-2 que comprende la administración de cantidades terapéuticamente eficaces de proteína de fusión anti-Her2 que comprende un primer y un segundo resto de unión anti-Her2, que puede ser el mismo o que puede unirse a diferentes epítopos de la proteína erbB-2, capaz de inhibir la función del receptor HER2. Preferentemente, dichas combinaciones de restos de unión presentarán una actividad citotóxica mejor de la que se esperaría para la suma de la actividad citotóxica de los anticuerpos individuales a la misma concentración global de anticuerpos. Adicionalmente, uno o más de los anticuerpos administrados pueden conjugarse con un resto citotóxico, por ejemplo, un fármaco antitumoral, una toxina o un radionúclido.

Anti-VRS:

En el presente documento se describe una proteína de fusión de unión anti-VRS aislada. "Anti-VRS" significa un resto de direccionamiento que se une específicamente a antígenos de superficie del virus respiratorio sincicial (VRS). Un anti-VRS puede ser un anticuerpo o un fragmento del mismo que neutraliza el VRS, impidiendo su capacidad para establecer una infección en un mamífero, o que contribuye a la eliminación del VRS de un hospedador infectado. El anti-VRS puede usarse como compañero de fusión con un XTEN para crear una composición de proteína de fusión que tenga utilidad profiláctica o terapéutica cuando se administra a un sujeto, tal como un lactante en riesgo de infección por el VRS. En una realización, el componente anti-VRS de una proteína de fusión de unión comprende una o más regiones determinantes de la complementariedad (CDR) de palivizumab. Se han descrito anticuerpos contra el VRS en la patente de EE. UU. n.º 5.824.307.

Anti-cMet:

En el presente documento se describe una proteína de fusión de unión anti-cMet aislada. "Anti-cMet" significa un resto de direccionamiento que se une específicamente al receptor Met o del factor de crecimiento de hepatocitos (HGF). El MET es un proto-oncogén, teniendo el receptor codificado del factor de crecimiento de hepatocitos (HGFR) o cMet una actividad de tirosina-cinasa esencial para el desarrollo embrionario y la cicatrización de heridas. Tras la unión y estimulación del HGF, el MET induce varias respuestas biológicas que, en conjunto, dan lugar a un crecimiento invasivo. La activación anormal del MET en el cáncer se correlaciona con un mal pronóstico, donde el MET anómalamente activo desencadena el crecimiento tumoral, la angiogénesis y la formación de nuevos vasos sanguíneos que suministran nutrientes al tumor, y la propagación del cáncer a otros órganos (metástasis). El MET está desregulado en muchos tipos de neoplasias malignas humanas, incluidos cánceres de riñón, hígado, estómago, mama y cerebro. Un anti-cMET puede ser un resto de direccionamiento que se une específicamente a un receptor del HGF, sirviendo como antagonista del HGF. El anti-cMET puede usarse como compañero de fusión con un XTEN para crear una composición de proteína de fusión que tenga utilidad profiláctica o terapéutica cuando se administra a un sujeto para el tratamiento de tumores que expresan el MET. En una realización, el componente anti-cMET de una proteína de fusión de unión comprende una o más regiones determinantes de la complementariedad (CDR) del anticuerpo MetMab o PRO143966. Los anticuerpos contra el cMET y sus secuencias se han descrito en la Patente de EE. UU. n.º 5.686.292. El documento US 6.468.529, el documento US 7.476.724 y la Publicación de Solicitud de Patente de EE. UU. n.º 20070092520.

Anti-IL6R:

En el presente documento se describe una proteína de fusión de unión anti-IL6R aislada. "Anti-IL6R" significa un resto de direccionamiento que se une específicamente a un receptor de la IL-6. El anti-IL6R puede servir como antagonista de la IL-6. El anti-IL6R puede usarse como compañero de fusión con un XTEN para crear una composición de proteína de fusión que tenga utilidad profiláctica o terapéutica cuando se administra a un sujeto para afecciones inflamatorias, tales como artritis o enfermedad de Crohn. Se ha demostrado que el tocilizuma tiene utilidad clínica en la artritis reumatoide de moderada a grave, y ha sido aprobado por la FDA. En una realización, el componente anti-IL6R de una proteína de fusión de unión comprende una o más regiones determinantes de la complementariedad (CDR) de tocilizuma. Se han descrito anticuerpos contra el IL-6R en las Patentes de EE. UU. n.º 5.670.373, 5.795.965, 5.817.790 y 7.479.543.

Anti-IL17:

En el presente documento se describe una proteína de fusión de unión anti-IL17 aislada. "Anti-IL17" significa un resto de direccionamiento que se une específicamente a la citocina IL-17. La IL-17 es una citocina homodimérica unida por disulfuro de aproximadamente 32 kDa que es sintetizada y secretada únicamente por los linfocitos T CD4+ de memoria activados (revisado en Fossiez et al., Int. Rev. Immunol., 16: 541-551 (1998)). La interleucina (IL-17) es una citocina proinflamatoria de los linfocitos T que se expresa, por ejemplo, en el líquido sinovial de pacientes con artritis reumatoide. La IL-17 es un potente inductor de diferentes citocinas tales como el TNF y la IL-1, y se ha demostrado que la IL-17 tiene efectos aditivos o incluso sinérgicos con el TNF y la IL-1. La anti-IL17 puede usarse como compañero de fusión con un XTEN para crear una composición de proteína de fusión de unión que tenga utilidad profiláctica o terapéutica cuando se administra a un sujeto para afecciones inflamatorias, tales como artritis o enfermedad de Crohn, o en la esclerosis múltiple. El LY2439821 es un anticuerpo que ha mostrado utilidad, cuando se añade a FARME orales, para mejorar los signos y síntomas de la artritis reumatoide. En una realización, el componente anti-IL6R de una proteína de fusión de unión comprende una o más regiones determinantes de la complementariedad (CDR) de LY2439821. Se han descrito anticuerpos anti-IL17 en las Solicitudes de Patente de EE. UU. n.º 20050147609 y 20080269467, y en la publicación de solicitud PCT WO 2007/070750.

IL17R:

En el presente documento se describe una proteína de fusión de unión al IL17R aislada. "IL17R" significa un resto de direccionamiento que se une específicamente al receptor de citocinas para la IL-17. Algunos estudios han demostrado que los linfocitos T en contacto con una forma soluble del polipéptido del receptor de la IL-17 inhibieron la proliferación de los linfocitos T y la producción de IL-2 inducida por PHA, concanavalina A y anticuerpo monoclonal anti-TCR (Yao et al., J. Immunol., 155: 5483-5486 [1995]). Como la interleucina (IL-17) es una citocina proinflamatoria de los linfocitos T que es un potente inductor de diferentes citocinas, tales como el TNF y la IL-1, el IL17R puede usarse como un compañero de fusión con XTEN para crear una composición de proteína de fusión de unión para unirse y neutralizar la IL-17. El IL17R puede tener utilidad terapéutica cuando se administra a un sujeto para afecciones inflamatorias, tales como artritis reumatoide o enfermedad de Crohn. Se han clonado receptores y homólogos del IL7R, como se describe en la Patente de EE. UU. n.º 5.869.286.

Anti-IL12:

En el presente documento se describe una proteína de fusión de unión anti-IL12 aislada. "Anti-IL12" significa un resto de direccionamiento que se une específicamente a la citocina IL-12 y, en algunos casos, a la IL-23. La IL-12 biológicamente activa existe como un heterodímero que comprende 2 subunidades unidas covalentemente de 35 (p35) y 40 (p40) kD, esta última se conoce como IL-23. La IL-12 es una citocina que es una parte importante de la respuesta

inflamatoria, y estimula la producción de interferón-gamma (IFN-γ) y factor de necrosis tumoral-alfa (TNF-α) a partir de linfocitos T y linfocitos citolíticos naturales (NK), y reduce la supresión mediada por la IL-4 del IFN-γ. Los linfocitos T que producen IL-12 tienen un correceptor, CD30, que está asociado con la actividad de la IL-12. La IL-12 también se ha relacionado con la autoinmunidad y con la psoriasis, siendo significativa la interacción entre los linfocitos T y los queratinocitos de células madre que producen IL-12. El ustekinumab es un anticuerpo anti-IL12/23 que ha demostrado utilidad en el tratamiento de la psoriasis en placas de moderada a grave, y ha sido aprobado por la FDA. El anti-IL-12 se puede utilizar como compañero de fusión con XTEN para crear una composición de proteína de fusión que tenga utilidad terapéutica cuando se administra a un sujeto que padece afecciones inflamatorias, tales como, pero sin limitación, psoriasis, artritis reumatoide o enfermedad de Crohn. En una realización, el componente anti-IL12 de una proteína de fusión de unión comprende una o más regiones determinantes de la complementariedad (CDR) del anticuerpo ustekinumab. Los anticuerpos contra la IL-12 y su uso se han descrito en la Patente de EE. UU. n.º 7.279.157.

Anti-IL23:

En el presente documento se describe una proteína de fusión de unión anti-IL23 aislada. "Anti-IL23" significa un resto de direccionamiento que se une específicamente a la citocina IL-23. La IL-23 es el nombre dado a un factor que está compuesto por la subunidad p40 de la IL-12, y es una citocina proinflamatoria que es parte importante de la respuesta inflamatoria frente a la infección. La IL-23 promueve la regulación por aumento de la metaloproteasa de la matriz MMP9, aumenta la angiogénesis y reduce la infiltración de linfocitos T CD8+. Se ha demostrado que la IL-23 desempeña un papel en la psoriasis, la esclerosis múltiple y el intestino inflamatorio. El ustekinumab es un anticuerpo anti-IL23 que ha demostrado utilidad en la psoriasis. El anti-IL-23 se puede utilizar como compañero de fusión con XTEN para crear una composición de proteína de fusión que tenga utilidad terapéutica cuando se administra a un sujeto que padece afecciones inflamatorias, tales como, pero sin limitación, psoriasis, artritis reumatoide o enfermedad de Crohn. En una realización, el componente anti-IL23 de una proteína de fusión de unión comprende una o más regiones determinantes de la complementariedad (CDR) del anticuerpo ustekinumab. Se han descrito anticuerpos contra la IL-23 en las Patentes de EE. UU. n.º 7.491.391 y 7.247.711.

Anti-RANKL:

En el presente documento se describe una proteína de fusión de unión anti-RANKL aislada. "Anti-RANKL" significa un resto de direccionamiento que se une específicamente a la proteína RANKL (activador del receptor del ligando del factor nuclear kappa B o ligando RANK). La RANKL es una proteína que actúa como la señal principal para promover la eliminación ósea, impulsada por los osteoclastos, que descomponen el hueso. En muchas afecciones de pérdida ósea, la RANKL supera la defensa natural del cuerpo contra la destrucción ósea. El anti-RANKL se puede utilizar como compañero de fusión con el XTEN para crear una composición de proteína de fusión que tenga utilidad terapéutica cuando se administra a un sujeto al inhibir la maduración de los osteoclastos mediante la unión a la RANKL, protegiendo el hueso de la degradación y, por lo tanto, de la osteoporosis. Por lo tanto, la proteína de fusión de unión imita los efectos endógenos de la osteoprotegerina, otra proteína producida por los osteoblastos que actúa como receptor alternativo de la RANKL, modulando la actividad osteoclástica inducida por la RANKL. Los anticuerpos contra la RANKL, tales como denosumab, han demostrado eficacia en ensayos de fase III demostrados en la osteoporosis posmenopáusica. En una realización, el componente anti-RANKL de una proteína de fusión de unión comprende una o más regiones determinantes de la complementariedad (CDR) del anticuerpo denosumab. Se han descrito anticuerpos anti-RANKL en la Patente de EE. UU. n.º 7.411.050.

CTLA4:

En el presente documento se describe una proteína de fusión de unión al CTLA4 aislada. "CTLA4" significa un resto de direccionamiento que se une específicamente a CD80 y CD86 en las células presentadoras del antígeno, y puede unirse específicamente a B7. El CTLA4 se puede utilizar como compañero de fusión con XTEN para crear una composición de proteína de fusión que tenga utilidad terapéutica cuando se administra a un sujeto que padece afecciones inflamatorias, tales como, pero sin limitación, artritis reumatoide, psoriasis y en el trasplante de órganos. Belatacept es una proteína de fusión compuesta por el fragmento Fc de una inmunoglobulina lgG 1 humana unida al dominio extracelular del CTLA-4 que ha demostrado eficacia para proporcionar una supervivencia prolongada del injerto. En una realización, el componente de unión al CTLA4 de la proteína de fusión de unión comprende una o más regiones de unión de belatacept. La clonación y el uso de composiciones de CTLA4 se han descrito en las Patentes de EE. UU. n.º 5.434.131, 5.773.253, 5.851.795, 5.885.579, 7.094.874 y 7.439.230.

ANTI-CD3:

La invención proporciona una proteína de fusión de unión anti-CD3 aislada. "Anti-CD3" significa un resto de direccionamiento que se une específicamente al receptor CD3 de los linfocitos T. El correceptor de linfocitos T es un complejo proteico compuesto por cuatro cadenas distintas; una cadena CD3γ, una cadena CD3δ y dos cadenas CD3ε. Estas cadenas se asocian con una molécula conocida como receptor de linfocitos T (TCR) y la cadena de ζ para generar una señal de activación en los linfocitos T. Los anticuerpos monoclonales anti-CD3 suprimen las respuestas inmunitarias mediante el agotamiento temporal de linfocitos T y la modulación antigénica del complejo CD3/receptor

de linfocitos T. Por ejemplo, el tratamiento anti-CD3 de ratones diabéticos adultos no obesos (NOD), un modelo espontáneo de diabetes sacarina autoinmunitaria dependiente de insulina mediada por linfocitos T, puede inhibir el proceso autoinmunitario que da lugar a la diabetes. El uso de anticuerpos anti-CD3 para tratar enfermedades y trastornos se ha descrito, por ejemplo, en la Patente de EE. UU. n.º 4.515.893. En una realización, el componente de unión al CD3 de la proteína de fusión de unión comprende una o más regiones determinantes de la complementariedad (CDR) del anticuerpo muromonab-CD3.

ANTI-CD40:

En el presente documento se describe una proteína de fusión de unión anti-CD40 aislada. "Anti-CD40" significa un resto de direccionamiento que se une específicamente al receptor de superficie celular CD-40. El CD-40 es un receptor de superficie celular que desempeña un papel en las respuestas inmunitarias, así como en el crecimiento celular y la señalización de supervivencia cuando es activado por el ligando CD40 (CD40L). EL CD40 está habitualmente sobreexpresado y activado en neoplasias malignas de linfocitos B, tales como mieloma múltiple y linfoma. El anti-CD40 se puede utilizar como compañero de fusión con XTEN para crear una composición de proteína de fusión que puede tener utilidad terapéutica cuando se administra a un sujeto que padece diferentes cánceres, particularmente neoplasias malignas de linfocitos B. En una realización, el componente anti-CD40 de una proteína de fusión de unión comprende una o más regiones determinantes de la complementariedad (CDR) del anticuerpo lucatumumab. Se han descrito anticuerpos anti-CD40 en la Patente de EE. UU. n.º 7.445.780 y en las Solicitudes de Patente de EE. UU. n.º 20070110754 y 20080254026.

ANTI-TNFalfa:

En el presente documento se describe una proteína de fusión de unión anti-TNFalfa aislada. "Anti-TNFalfa" significa un resto de direccionamiento que se une específicamente a la citocina TNFalfa. El TNFalfa, o caquexina, es una citocina proinflamatoria implicada en la inflamación sistémica y es miembro de un grupo de citocinas que estimulan la reacción en fase aguda. El papel principal del TNF es en la regulación de las células inmunitarias. El TNF es producido principalmente por los macrófagos, pero también es producido por células linfoides, mastocitos, células endoteliales, miocitos cardíacos, tejido adiposo, fibroblastos y tejido neuronal. Se liberan grandes cantidades de TNF en respuesta a lipopolisacárido e interleucina-1 (IL-1). El TNF ha sido implicado en trastornos autoinmunitarios tales como artritis reumatoide, espondilitis anquilosante, enfermedad de Crohn, psoriasis y asma resistente, y desempeña un papel en el choque séptico y otras formas graves de respuesta inflamatoria aguda y SIRS. El anti-IL-TNFalfa se puede utilizar como compañero de fusión con XTEN para crear una composición de proteína de fusión que puede tener utilidad terapéutica en una amplia variedad de trastornos inflamatorios, incluyendo artritis reumatoide, espondilitis anquilosante, enfermedad de Crohn, psoriasis y asma resistente. Los anticuerpos anti-TNFalfa, tales como infliximab y etanercept, han demostrado eficacia en psoriasis, enfermedad de Crohn, espondilitis anquilosante, artritis psoriásica, artritis reumatoide y colitis ulcerosa. En una realización, el componente anti-TNFalfa de una proteína de fusión de unión comprende una o más regiones determinantes de la complementariedad (CDR) o regiones de unión de infliximab o de etanercept. Se han descrito anticuerpos anti-TNF en la Patente de EE. UU. n.º 6.790.444, y se han descrito anticuerpos quiméricos que comprenden un receptor del TNF en la Patente de EE. UU. n.º 5.605.690.

La invención proporciona composiciones de proteína de fusión de unión en las que las regiones de unión de los restos de direccionamiento ilustrativos mencionados anteriormente son variantes de secuencia. En la tabla 5 se muestran ejemplos de sustituciones conservadoras de aminoácidos en secuencias de polipéptidos. Sin embargo, en realizaciones de la proteína de fusión de unión en las que la identidad de secuencia del resto de direccionamiento es inferior al 100 % en comparación con una secuencia específica a la que se hace referencia o divulgada en el presente documento, la invención contempla la sustitución de cualquiera de los otros 19 L-aminoácidos naturales por un resto de aminoácido dado del resto de direccionamiento dado, que puede estar en cualquier posición dentro de la secuencia del resto de direccionamiento o región de unión del resto de direccionamiento, incluidos los restos de aminoácidos adyacentes. Si una sustitución cualquiera da como resultado un cambio indeseable en la actividad de unión, entonces se puede emplear uno de los aminoácidos alternativos y evaluarse la proteína de construcción mediante los métodos descritos en el presente documento (por ejemplo, los ensayos de los ejemplos), o utilizando cualquiera de las técnicas y directrices para mutaciones conservadoras y no conservadoras expuestas, por ejemplo, en la Patente de EE. UU. N.º 5.364.934, o utilizando métodos generalmente conocidos en la técnica. Además, las variantes pueden incluir, por ejemplo, polipéptidos en donde se añaden o eliminan uno o más restos de aminoácidos en el extremo amínico o carboxílico de la secuencia de aminoácidos referenciada o divulgada de un resto de direccionamiento que conserva parte, si no toda, la actividad de unión del resto de direccionamiento referenciado o divulgado; por ejemplo, la capacidad de unirse a una diana de la tabla 1 o la tabla 2.

Tabla 5: Sustituciones conservadoras de aminoácidos ilustrativas

Resto original	Sustituciones ilustrativas
Ala (A)	val; leu; ile
Arg (R)	Lys; Gln; Asn

Resto original	Sustituciones ilustrativas
Asn (N)	Gln; His;Llys; Arg
Asp (D)	Glu
Cys (C)	Ser
Gln (Q)	Asn
Glu (E)	Asp
Gly (G)	Pro
His (H)	Asn; Gln; Lys; Arg
Ile (I)	Leu; Val; Met; Ala; Phe; Norleucina
Leu (L)	Norleucina; Ile: Val; Met; Ala: Phe
Lys (K)	Arg' Gln; Asn
Met (M)	Leu; Phe; Ile
Phe (F)	Leu; Val; i=LIe; Ala
Pro (P)	Gly
Ser (S)	Thr
Thr (T)	Ser
Trp (W)	Tyr
Tyr (Y)	Trp; Phe: Thr; Ser
Val (V)	Ile; Leu; Met; Phe; Ala; Norleucina

2. <u>Formas ilustrativas de restos de direccionamiento</u>

La siguiente sección proporciona una lista y una descripción no limitativas de formas ilustrativas de restos de direccionamiento.

"Anticuerpo" o "anticuerpos", como se utiliza en el presente documento, se refiere a un resto de direccionamiento que consiste en uno o más polipéptidos sustancialmente codificados por genes de inmunoglobulina o fragmentos de genes de inmunoglobulina, y se utiliza en el sentido más amplio para cubrir anticuerpos monoclonales intactos, anticuerpos multiespecíficos (por ejemplo, anticuerpos biespecíficos) formados a partir de al menos dos anticuerpos intactos o fragmentos de los mismos, y fragmentos de anticuerpos siempre que presenten la actividad biológica deseada; por ejemplo, afinidad de unión por un ligando o antígeno diana.

Las inmunoglobulinas se pueden asignar a diferentes clases dependiendo de la secuencia de aminoácidos del dominio constante de sus cadenas pesadas. Hay cinco clases principales de inmunoglobulinas: IgA, IgD, IgE, IgG e IgM, y varias de éstas pueden dividirse además en subclases (isotipos), por ejemplo, IgG1, IgG2, IgG3, IgG4, IgA1 e IgA2. Los dominios constantes de la cadena pesada que corresponden a las diferentes clases de inmunoglobulinas se denominan α, δ, ε, γ y μ, respectivamente. Las estructuras de las subunidades y las configuraciones tridimensionales de las diferentes clases de inmunoglobulinas son bien conocidas.

El término "monoclonal" indica el carácter del anticuerpo o fragmento de anticuerpo del resto de direccionamiento como obtenido a partir de una población sustancialmente homogénea de anticuerpos o fragmentos, y no debe interpretarse como que requiere la producción del anticuerpo mediante un método en particular. Por ejemplo, aunque los anticuerpos monoclonales creados de acuerdo con los métodos de la presente invención pueden prepararse mediante el método de hibridoma, descrito por primera vez por Kohler et al., Nature, 256: 495 (1975), también pueden ser sintéticos elaborados mediante métodos de ADN recombinante (véase, por ejemplo, la Patente de EE. UU. N.º 4.816.567) y expresados en hospedadores de mamífero o no mamífero; por ejemplo, *E. coli.* La sustitución de células inmortalizadas por células bacterianas simplifica considerablemente los procedimientos para preparar grandes cantidades de moléculas de la proteína de fusión de unión inventiva. Además, un sistema de producción recombinante permite la capacidad de producir anticuerpos a medida y fragmentos de los mismos, o incluso bibliotecas para un cribado en función de atributos específicos. Por ejemplo, es posible producir moléculas quiméricas con nuevas combinaciones de funciones de unión y efectoras, anticuerpos humanizados y moléculas de unión al antígeno novedosas, incluidas proteínas de fusión de unión bifuncionales. Además, el uso de la amplificación mediante la reacción en cadena de la polimerasa (PCR) (Saiki, R. K., et al., Science 239, 487-491 (1988)) para introducir variaciones en la secuencia y aislar secuencias que producen anticuerpos a partir de células tiene un gran potencial

para acelerar la escala de tiempo bajo la cual se pueden aislar las especificidades. Los genes V_H y V_L amplificados se pueden clonar directamente en vectores para la expresión en bacterias o células de mamífero (Orlandi, R., et al., 1989, Proc. Natl. Acad. Sci., USA 86, 3833-3837; Ward, E. S., et al., 1989 supra; Larrick, J. W., et al., 1989, Biochem. Biophys. Res. Commun. 160, 1250-1255; Sastry, L. et al., 1989, Proc. Natl. Acad. Sci., USA, 86, 5728-5732). Los fragmentos de anticuerpos solubles secretados a partir de bacterias pueden seleccionarse a continuación en ensayos de unión descritos en el presente documento o en otros conocidos en la técnica, para seleccionar aquellas construcciones con actividades de unión suficientes para cumplir la aplicación.

Los anticuerpos monoclonales del presente documento incluyen específicamente anticuerpos "quiméricos" en los que una porción de la cadena pesada y/o de la ligera es idéntica a, o tiene un alto grado de homología con, las secuencias progenitoras correspondientes en anticuerpos derivados de una primera especie en particular, mientras que el resto de la cadena o cadenas es idéntico a, o tiene un alto grado de homología con, las secuencias de anticuerpos derivados de una segunda especie, en donde el anticuerpo resultante presenta la actividad biológica deseada; por ejemplo, afinidad de unión por el antígeno o ligando diana (Pat. de EE. UU. N.º 4.816.567; Morrison et al., Proc. Natl. ACAD. SCI. USA, 81: 6851-4855 (1984)).

El término "humanizado" significa formas de anticuerpos, incluidos fragmentos, que son quiméricas por incluir una secuencia mínima derivada de una inmunoglobulina no humana, pero que por lo demás comprenden una secuencia de inmunoglobulinas humanas. En la técnica se han descrito métodos para humanizar anticuerpos no humanos. Preferentemente, un anticuerpo humanizado tiene uno o más restos de aminoácidos introducidos en el mismo procedentes de una fuente que no es humana (por ejemplo, murina, de rata, o de primates no humanos) y que normalmente se toman a partir de un dominio variable de una cadena V_L o V_H que tiene la especificidad y afinidad deseadas por el ligando diana. La humanización se puede realizar esencialmente siguiendo el método de Winter y colaboradores (Jones et al., Nature, 321: 522-525 (1986); Riechmann et al., Nature, 332: 323-327 (1988); Verhoeyen et al., Science, 239: 1534-1536 (1988)), sustituyendo las secuencias de la región hipervariable por las secuencias correspondientes de un anticuerpo humano. Por consiguiente, dichos anticuerpos "humanizados" son anticuerpos quiméricos (véase, por ejemplo, la Pat. de EE. UU. N.º 4.816.567) en donde todo o una porción del dominio variable humano ha sido sustituido por la secuencia correspondiente de una especie no humana. En la práctica, los anticuerpos humanizados son normalmente anticuerpos humanos en los que algunos restos de CDR hipervariables y posiblemente algunos restos de FR son sustituidos por restos de sitios análogos de anticuerpos de roedores (u otras especies no humanas, por ejemplo, primates no humanos). En una realización, los anticuerpos humanizados comprenden restos que no se encuentran en el anticuerpo receptor o en el anticuerpo donante para, por ejemplo, aumentar la afinidad de unión o alguna otra propiedad. En general, los anticuerpos humanizados comprenden sustancialmente todos de al menos uno, y normalmente dos, dominios variables, en los que todos o sustancialmente todos los bucles hipervariables corresponden o tienen secuencias derivadas de las de una inmunoglobulina no humana, y todas o sustancialmente todas las regiones FR son las de una secuencia de inmunoglobulina humana. El anticuerpo humanizado puede comprender opcionalmente al menos una porción de una región constante de inmunoglobulina (Fc), preferentemente la de una inmunoglobulina humana.

Los restos de direccionamiento de las presentes composiciones pueden derivar de anticuerpos humanizados. La elección de los dominios variables humanos, tanto ligeros como pesados, que van a utilizarse en las composiciones es muy importante para reducir la antigenicidad del anticuerpo. Por ejemplo, la secuencia del dominio variable de un anticuerpo roedor puede cribarse frente a una biblioteca de secuencias del dominio variable humano conocidas con el fin de seleccionar una secuencia que sea menos probable que provoque una respuesta inmunitaria en el receptor. De manera correspondiente, la secuencia humana más cercana a la del roedor puede usarse como el marco humano (FR) para el anticuerpo humanizado (Sims et al., J. Immunol, 151: 2296 (1993); Chothia et al., J. Mol. Biol., 196: 901 (1987)). Se puede utilizar el mismo marco para varios anticuerpos humanizados diferentes (Carter et al., Proc. Natl. ACAD. SCI. USA, 89: 4285 (1992); Presta et al., J. Immnol., 151: 2623 (1993)).

Una propiedad adicional es que los restos de direccionamiento pueden humanizarse conservando una alta afinidad por el antígeno y otras propiedades biológicas favorables. Para conseguir este objetivo, de acuerdo con un método preferido, los restos de direccionamiento humanizados se preparan mediante un proceso iterativo de análisis de las secuencias progenitoras y diferentes productos humanizados conceptuales utilizando modelos tridimensionales de las secuencias progenitoras y humanizadas seguidos de ensayos. Los modelos tridimensionales de las inmunoglobulinas están habitualmente disponibles y son familiares para los expertos en la técnica. Hay disponibles programas informáticos que ilustran y muestran probables estructuras conformacionales tridimensionales de secuencias de inmunoglobulina candidatas seleccionadas. La inspección de estas presentaciones permite analizar el probable papel de los restos en el funcionamiento de la secuencia de inmunoglobulina candidata, es decir, el análisis de restos que influyen en la capacidad de la inmunoglobulina candidata para unirse a su antígeno. De esta manera, los restos de FR se pueden seleccionar y combinar del receptor y el donante utilizando técnicas de ADN recombinante convencionales de modo que se pueda conseguir la característica deseada, tal como una afinidad aumentada por el antígeno o antígenos diana. En una realización, se crean construcciones de proteína de fusión de unión en las que la secuencia que comprende dominios variables de la cadena pesada unidos se une a un dominio constante de la cadena pesada, y la secuencia que comprende dominios variables de la cadena ligera unidos se une a un dominio constante de la cadena ligera. Preferentemente, los dominios constantes son el dominio constante de la cadena pesada humana y el dominio constante de la cadena ligera humana, respectivamente. En una realización adicional de lo anterior, la proteína

de fusión puede diseñarse para que incluya porciones o la totalidad de una región bisagra con el fin de permitir la dimerización de la proteína de fusión de unión, y que puede comprender además un XTEN unido al extremo carboxílico de la región constante. En una realización alternativa, la proteína de fusión de unión puede diseñarse para que incorpore un Fc parcial sin bisagra y con un dominio CH2 que está truncado pero conserva la unión a FcRn con el fin de conferir una semivida terminal más larga a la construcción. En otra realización más, la proteína de fusión de unión puede diseñarse para que incorpore un Fc parcial sin bisagra, pero con un dominio CH2 y CH3, que puede dimerizar a través del dominio CH3. En las realizaciones anteriormente en el presente documento, un XTEN puede unirse al extremo amínico o carboxílico de la proteína de fusión, para mejorar una o más propiedades de la proteína de fusión de unión resultante.

Los "fragmentos de anticuerpos" comprenden una porción de un anticuerpo intacto o un homólogo sintético o quimérico, preferentemente la región de unión al antígeno o variable del anticuerpo intacto. Algunos ejemplos de fragmentos de anticuerpos incluyen moléculas tales como fragmentos Fab, fragmentos Fab', fragmentos $F(ab')_2$, fragmentos Fd, fragmentos Fabc, fragmentos Fd, fragmentos Fabc, anticuerpos de dominio (V_{HH}), moléculas de anticuerpos monocatenarios (scFv), diacuerpos, cadenas ligeras de anticuerpos individuales, cadenas pesadas de anticuerpos individuales, fusiones quiméricas entre cadenas de anticuerpos y otras moléculas, y similares.

Un "fragmento Fab" se refiere a una región de un anticuerpo que se une a antígenos. Un fragmento Fab está compuesto por un dominio constante y un dominio variable de cada una de las cadenas pesada y ligera. Estos dominios dan forma al parátopo (el sitio de unión al antígeno) en el extremo aminoterminal del monómero. Los dos dominios variables se unen al epítopo en sus antígenos específicos. Un fragmento Fab puede estar unido por un puente de disulfuro en el extremo carboxílico. Los fragmentos Fab se pueden generar *in vitro*. Por ejemplo, puede usarse la enzima papaína para escindir un monómero de inmunoglobulina en dos fragmentos Fab y un fragmento Fc. La enzima pepsina escinde por debajo de la región bisagra, por lo que se forma un fragmento $F(ab')_2$ y un fragmento Fc. Como se describe más completamente a continuación, las regiones variables de las cadenas pesada y ligera se pueden fusionar entre sí para formar un fragmento variable de cadena única (scFv), que conserva la especificidad original de la inmunoglobulina progenitora.

Las "cadenas ligeras" de los anticuerpos (inmunoglobulinas) de cualquier especie de vertebrado se pueden clasificar en uno de dos tipos claramente distintos, llamados kappa y lambda, tomando como base las secuencias de aminoácidos de sus dominios constantes.

El término "variable" se refiere al hecho de que porciones de los dominios variables difieren ampliamente en su secuencia entre los anticuerpos y confieren la especificidad de unión de cada anticuerpo particular para su antígeno particular. La variabilidad se concentra en tres segmentos denominados regiones determinantes de la complementariedad (CDR) o regiones hipervariables, en los dominios variables tanto de la cadena ligera como de la cadena pesada; es decir, CDR1, CDR2 y CDR3. En particular, las regiones CDR de los anticuerpos se pueden incorporar en restos de direccionamiento de las presentes composiciones, pero se pueden seleccionar individualmente a partir de uno o más anticuerpos para crear el dominio de unión. Las porciones más conservadas de los dominios variables se denominan regiones estructurales (FR), que también pueden ser incorporadas en restos de direccionamiento. Cada dominio variable de las cadenas pesadas y ligeras naturales comprende cuatro regiones FR, adoptando normalmente una configuración de lámina β, conectadas por tres CDR que forman bucles. Las CDR de cada cadena se mantienen juntas en estrecha proximidad por las regiones FR y, con las CDR de la otra cadena, contribuyen a la formación del sitio de unión al antígeno de los anticuerpos (véase Kabat et al., *NIH Publ. N.º* 91-3242, vol. I, páginas 647-669 (1991)). Los dominios constantes no están implicados directamente en la unión de un anticuerpo a un antígeno, sino que presentan o participan en diferentes funciones efectoras, tales como la toxicidad celular dependiente de anticuerpo.

Proteínas de fusión de unión con fragmentos variables monocatenarios

En un aspecto, la presente invención proporciona composiciones de proteína de fusión de unión con fragmentos variables monocatenarios. La expresión "fragmento variable monocatenario" o "scFv" significa un fragmento de anticuerpo que comprende un dominio V_H y un dominio V_L de un anticuerpo, en donde estos dominios están presentes en una única cadena polipeptídica, y generalmente están unidos por un conector polipeptídico entre los dominios que permite que el scFv forme la estructura deseada para la unión al antígeno. Los métodos para crear scFv son conocidos en la técnica (véase, por ejemplo, la Patente de EE. UU. 6.806.079; Bird et al. (1988) Science 242: 423-426; Huston et al (1988) PNAS 85: 5879-5883; Pluckthun en The Pharmacology of Monoclonal Antibodies, vol. 113, Rosenburg y Moore eds., Springer-Verlag, Nueva York, págs. 269-315 (1994)).

Un dominio de unión de las composiciones de proteína de fusión de unión con scFv de la invención puede tener la configuración del extremo amínico al extremo carboxílico VH-conector-VL o VL-conector-VH. Las proteínas de fusión de unión incluirían al menos una primera secuencia del XTEN y opcionalmente una segunda secuencia del XTEN unida al extremo amínico o carboxílico de la proteína de fusión (como se muestra en la FIG.1), lo que daría como resultado al menos las siguientes permutaciones de estructura (de amino a carboxiterminal); XTEN-VH-conector-VL; VH-conector-VL-XTEN; XTEN-VH-conector-VL-XTEN; XTEN-VL-conector-VH; VL-conector-VH-XTEN; XTEN-VL-conector-VH-XTEN. En una realización de lo anterior, la composición comprende un XTEN unido al extremo amínico

de la proteína de fusión, en donde la expresión de la proteína de fusión en una célula hospedadora transformada con un vector de expresión adecuado que comprende un polinucleótido que codifica la proteína de fusión que comprende la secuencia aminoterminal está mejorada en comparación con la expresión de una proteína de fusión correspondiente de un polinucleótido que no comprende la secuencia codificante del XTEN aminoterminal. En dichos casos, el XTEN aminoterminal podría tener únicamente la secuencia corta que mejora la expresión, o podría incluir además un XTEN largo de al menos más de 400 a aproximadamente 3000 restos de aminoácidos entre la parte aminoterminal y un dominio de unión, para conferir unas propiedades farmacocinéticas o farmacéuticas mejoradas a la proteína de fusión, como se ha descrito anteriormente. En una realización de lo anterior, la secuencia del XTEN aminoterminal comprende una secuencia que presenta al menos aproximadamente un 80 % de identidad de secuencia, o un 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de identidad de secuencia con una secuencia seleccionada entre AE48, AM48, AE624, AE913 y AM923. Descritos en el presente documento, los conectores, el XTEN aminoterminal, así como el XTEN portador largo pueden comprender una secuencia que puede ser un fragmento de o que presenta al menos aproximadamente un 80 % de identidad de secuencia, o un 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de identidad de secuencia con una secuencia seleccionada de una cualquiera de las tablas 4 u 11-15.

Los conectores utilizados para unirse a los componentes de las proteínas de fusión de unión deben ser de naturaleza flexible. En una realización, el conector que une los dominios de unión V_L y V_H que forman el sitio de unión al antígeno del resto de direccionamiento scFv puede tener de aproximadamente 15 a aproximadamente 30 restos de aminoácidos de longitud. En otra realización, el conector puede tener de aproximadamente 30 a aproximadamente 200 restos de aminoácidos, o aproximadamente 40 a aproximadamente 144 restos de aminoácidos, o aproximadamente 50 a aproximadamente 96 restos de aminoácidos. En cualquiera de las realizaciones descritas anteriormente en este párrafo, el conector puede ser una secuencia derivada de un fragmento de cualquiera de las secuencias XTEN de las tablas 4 u 11-15. En otra realización, el conector puede ser una secuencia en la que al menos el 80 % de los restos están comprendidos por aminoácidos glicina, serina y/o glutamato, tales como, pero sin limitación, una secuencia con aproximadamente un 80-100 % de identidad de secuencia con la secuencia GSGEGSEGEGGGEGSEGEGSGEGGEGEGSG (SEQ ID NO: 1) o una porción o un multímero de la misma.

En una realización, la invención proporciona proteínas de fusión de unión que comprenden dos o más restos de direccionamiento scFv (como se muestra en la FIG. 2). En una realización, los dos o más restos de direccionamiento scFv pueden ser idénticos. En otra realización, los dos o más restos de direccionamiento scFv pueden ser diferentes y pueden unirse a diferentes dianas o a diferentes epítopos en la misma diana. En las realizaciones anteriores, los dos o más restos de direccionamiento scFv pueden estar unidos mediante una secuencia conectora. La divulgación contempla que los conectores, así como el portador XTEN largo y el XTEN aminoterminal, pueden comprender una secuencia que puede ser un fragmento de o que presenta al menos aproximadamente un 80 % de identidad de secuencia, o un 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de identidad de secuencia con una secuencia seleccionada de una cualquiera de las tablas 4 u 11-15.

La invención contempla diferentes configuraciones de las proteínas de fusión de unión con scFv-XTEN multivalentes, con los dos o más restos de direccionamiento, conectores y uno o más XTEN en diferentes configuraciones de amino- a carboxiterminal; por ejemplo, TM1-L-TM2-XTEN, XTEN-TM1-L-TM2, XTEN-TM1-L-TM2-XTEN, etc.

La metodología general para el ensamblaje de los componentes, la expresión y la recuperación, seguida de la caracterización de la proteína de fusión de unión, se ilustra en la FIG. 9.

Proteínas de fusión de unión de diacuerpo

En el presente documento se describen composiciones de proteínas de fusión de unión de diacuerpo. El término "diacuerpo" o "diacuerpos", como se utiliza en el presente documento, se refiere a proteínas de fusión que comprenden fragmentos de anticuerpos con dos sitios de unión al antígeno, cuyos fragmentos comprenden un dominio variable de cadena pesada (V_H) conectado a un dominio variable de cadena ligera (V_L) mediante un conector corto en la misma cadena polipeptídica. Descrita en el presente documento, la composición de proteína de fusión de unión de diacuerpo es una única cadena polipeptídica que tiene dos restos de unión, cada uno con un V_L y un V_H interconectados por un conector flexible de aproximadamente 30 a aproximadamente 300 restos de aminoácidos que unen el extremo carboxílico de un par de dominios de unión con el extremo amínico del segundo par de dominios de unión, y uno o más XTEN en el extremo amínico y/o carboxílico (como se muestra en la FIG. 3). En este caso, la composición se crea mediante la incorporación selectiva de un conector que es demasiado corto para permitir el emparejamiento entre los dos dominios adyacentes, pero las regiones del resto de direccionamiento están conectadas por un conector flexible más largo para permitir la asociación entre los dominios de unión de los extremos opuestos de la proteína de fusión. Por ejemplo, se sabe que una molécula scFv con un conector de 3-12 restos de longitud no puede plegarse en un dominio Fv funcional, pero, generalmente, en su lugar se asocia con una segunda molécula scFv para formar un diacuerpo multivalente con dos sitios de unión, mientras que un conector de 1-2 restos tiende a favorecer la formación de un triacuerpo con tres sitios de unión (John L. Atwell et al., Protein Engineering (1999) 12 (7): 597-604). Por lo tanto, en el presente documento se describen proteínas de fusión de unión de diacuerpo en las que los pares de dominios

de unión adyacentes están unidos por 1-2 o 3-12 restos de aminoácidos, y los pares de dominios de unión están, a su vez, unidos por un conector XTEN flexible. En una realización de lo anterior, el conector flexible puede ser una secuencia de XTEN de aproximadamente 100 a aproximadamente 300 restos de aminoácidos. En otra realización, la secuencia conectora flexible puede comprender aminoácidos tales como glicina, serina y/o glutamato que componen aproximadamente un 80-100 % de la secuencia. En otra realización, el conector puede ser una secuencia en la que al menos el 80 % de los restos están comprendidos por aminoácidos glicina, serina y/o glutamato, tales como, pero sin limitación, una secuencia con aproximadamente un 80-100 % de identidad de secuencia con la secuencia GSGEGSEGEGGGEGSEGEGSGEGGEGEGSG (SEQ ID NO: 1) o una porción o un multímero de la misma.

En el presente documento se describe una proteína de fusión de unión de diacuerpo que puede ser monoespecífica, uniéndose a un tipo de antígeno. En el presente documento se describe una proteína de fusión de unión de diacuerpo biespecífica en la que los restos de direccionamiento están dirigidos a diferentes antígenos o diferentes epítopos del mismo antígeno. En las realizaciones anteriores, la especificidad de la fusión de unión de diacuerpo se determina mediante la incorporación de componentes V_H y V_L idénticos o diferentes respectivos.

Una ilustración de un ejemplo de proteínas de fusión de unión de diacuerpo se muestra en la FIG. 2, sin embargo, la divulgación contempla diferentes configuraciones de las proteínas de fusión de unión de diacuerpo; por ejemplo, un orden diferente de amino- a carboxiterminal de los dominios V_H y V_L o la deleción de una de las secuencias XTEN del extremo amínico o carboxílico.

En una realización de lo anterior, la composición comprende un XTEN unido al extremo amínico de la proteína de fusión de unión de diacuerpo, en donde la expresión de la proteína de fusión en una célula hospedadora transformada con un vector de expresión adecuado que comprende un polinucleótido que codifica la proteína de fusión que comprende la secuencia aminoterminal está mejorada en comparación con la expresión de una proteína de fusión correspondiente de un polinucleótido que no comprende la secuencia codificante del XTEN aminoterminal.

Además, la invención proporciona multímeros que son trivalentes (que tienen tres sitios de unión al antígeno), tetravalentes (que tienen cuatro sitios de unión al antígeno), y así sucesivamente. En otra realización de las proteínas de fusión de diacuerpo de la divulgación, los conectores, así como el portador XTEN largo y el XTEN aminoterminal, pueden comprender una secuencia que puede ser un fragmento de o que presenta al menos aproximadamente un 80 % de identidad de secuencia, o un 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de identidad de secuencia con una secuencia seleccionada de una cualquiera de las tablas 4 u 11-15.

La metodología general para el ensamblaje de los componentes, la expresión y la recuperación, seguida de la caracterización de la proteína de fusión de unión, se ilustra en la FIG. 9.

Proteínas de fusión de unión a anticuerpos de dominio

En el presente documento se describen proteínas de fusión monocatenarias que comprenden restos de direccionamiento a anticuerpos de dominio unidos a XTEN, en donde la proteína de fusión es capaz de unirse a un antígeno o ligando. Como se utiliza en el presente documento, "anticuerpo de dominio" o "V_{HH}" es una inmunoglobulina que tiene una región variable (V_H) con un sitio de unión al antígeno que solo permitirá el reconocimiento y la unión completa de un antígeno, unido a una región constante (C_H), pero desprovisto del primer dominio de la región constante (C_H1) y desprovisto de un dominio V_L. La V_{HH} no corresponde a fragmentos obtenidos, por ejemplo, mediante la degradación de un modelo de inmunoglobulina natural de cuatro cadenas. Las V_{HH} están desprovistas de cadenas ligeras, de modo que los dominios variables de sus cadenas pesadas tienen unas propiedades que difieren de las de la cadena pesada variable (V_H) de las inmunoglobulinas de cuatro cadenas, incluidos los sitios de interacción no normales con el dominio V_L o el C_H1. Las V_{HH} pueden adoptar una organización tridimensional que se distingue de la organización tridimensional convencional de los anticuerpos de cuatro cadenas, de acuerdo con la descripción que proporciona Chothier C. y Lesk A. M, (1987-J. Mol. Biol. 197, 901-917). Las V_{HH} pueden comprender inmunoglobulinas de tipo G, especialmente de clase 2 (IgG2) o de clase 3 (IgG3). Las secuencias de inmunoglobulinas V_{HH} progenitoras pueden derivar de ciertos animales, especialmente de miembros de la familia de los camélidos, o de tiburones, que a continuación pueden generarse en células hospedadoras mediante ingeniería genética o mediante síntesis química. Las células hospedadoras apropiadas incluyen bacterias (por ejemplo, *E. coli*) y células eucariotas, tales como levaduras o células animales, incluidas células de mamífero, o células vegetales.

Descrita en el presente documento, la proteína de fusión de unión puede comprender un resto de direccionamiento V_{HH}. Descrita en el presente documento, la proteína de fusión de unión puede comprender dos o más restos de direccionamiento V_{HH}; por ejemplo, dos, o tres, o cuatro, o cinco, o seis o más restos de direccionamiento V_{HH}. En dicho caso, la secuencia conectora entre los fragmentos de V_{HH} puede ser, por ejemplo, una secuencia correspondiente a un fragmento del dominio de bisagra de una inmunoglobulina (por ejemplo, el dominio de bisagra larga) o puede ser una secuencia de XTEN corta de aproximadamente 30 a aproximadamente 300 restos de aminoácidos, o puede ser una secuencia en la que al menos el 80 % de los restos comprenden los aminoácidos glicina, serina y/o glutamato, tal como, pero sin limitación, una secuencia con aproximadamente un 80-100 % de identidad de secuencia con la secuencia GSGEGSEGEGGGEGSEGEGSGEGGEGEGSG (SEQ ID NO: 1) o una

porción o un multímero de la misma. En una realización de lo anterior, las V_{HH} son heteroespecíficas con afinidad de unión por diferentes epítopos del mismo antígeno. En otra realización, las V_{HH} son heteroespecíficas con afinidad de unión por antígenos heterólogos. La divulgación contempla que la proteína de fusión de unión de V_{HH} comprendería una secuencia de XTEN adicional de al menos 400 a aproximadamente 3000 restos en los que XTEN conferiría unas propiedades mejoradas a la composición, como se ha descrito anteriormente. Por lo tanto, la divulgación contempla configuraciones de proteínas de fusión de unión que incluyen, pero sin limitación, V_{HH}-XTEN; XTEN-V_{HH}; V_{HH}-Conector-V_{HH}-XTEN; XTEN-V_{HH}-Conector-V_{HH}; y XTEN-V_{HH}-Conector-V_{HH}-XTEN, o multímeros de las mismas. En lo anterior, para aquellas configuraciones con dos V_{HH}, la V_{HH}- puede ser idéntica o puede ser diferente, en este último caso, uniéndose a dos ligandos/antígenos diferentes o epítopos diferentes en el mismo ligando/antígeno. En una realización de las anteriores configuraciones de proteína de fusión de unión de dominio, la composición comprende un XTEN unido al extremo amínico de la proteína de fusión, en donde la expresión de la proteína de fusión en una célula hospedadora transformada con un vector de expresión adecuado que comprende un polinucleótido que codifica la proteína de fusión que comprende la secuencia aminoterminal está mejorada en comparación con la expresión de una proteína de fusión correspondiente de un polinucleótido que no comprende la secuencia codificante del XTEN aminoterminal.

En otra realización de las proteínas de fusión de dominio de la divulgación, los conectores, así como el portador XTEN largo y el XTEN aminoterminal, pueden comprender una secuencia que puede ser un fragmento de o que presenta al menos aproximadamente un 80 % de identidad de secuencia, o un 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de identidad de secuencia con una secuencia seleccionada de una cualquiera de las tablas 4 u 11-15. Descrito en el presente documento, el conector puede ser una secuencia en la que al menos el 80 % de los restos están comprendidos por aminoácidos glicina, serina y/o glutamato, tales como, pero sin limitación, una secuencia con aproximadamente un 80-100 % de identidad de secuencia con la secuencia GSGEGSEGEGGGEGSEGEGSGEGGEGEGSG (SEQ ID NO: 1) o una porción o un multímero de la misma.

La metodología general para el ensamblaje de los componentes, la expresión y la recuperación, seguida de la caracterización de la proteína de fusión de unión, se ilustra en la FIG. 9.

Proteínas de fusión de unión a citocinas

En el presente documento se describen proteínas de fusión monómeras capaces de unirse a una citocina para formar un complejo no funcional, que comprenden dominios de unión similares a receptores unidos por un XTEN, en donde la proteína de fusión resultante tiene afinidad de unión por el ligando de citocina diana. Como se utiliza en el presente documento, "citocinas" significa una categoría de pequeñas moléculas de señalización proteicas solubles que se utilizan ampliamente en la comunicación celular, incluyendo, pero sin limitación, interleucinas, interferones, quimiocinas, factores de crecimiento, factores estimulantes de colonias, factores de necrosis tumoral, hormonas peptídicas y otros agentes inmunomoduladores que median en una variedad de efectos biológicos, que incluyen la regulación del crecimiento y la diferenciación de muchos tipos de células. La acción de las citocinas puede ser autocrina, paracrina y endocrina. Por ejemplo, la interleucina-17 humana es una citocina que estimula la expresión de la interleucina-6, la molécula de adhesión intracelular 1, la interleucina-8, el factor estimulante de colonias de granulocitos y macrófagos, y la expresión de la prostaglandina E2, y desempeña un papel en la maduración preferente de precursores hematopoyéticos CD34+ en neutrófilos (Yao et al., J. Immunol. 155: 5483 (1995); Fossiez et al., J. Exp. MED. 183: 2593 (1996)).

Las citocinas se unen generalmente a receptores de la superficie celular que, a su vez, desencadenan cascadas de señalización intracelular que alteran las funciones celulares. Éstas pueden incluir la regulación por aumento y/o la regulación por disminución de varios genes y sus productos de transcripción, lo que da como resultado la producción de otras citocinas, un aumento en el número de receptores de superficie para otras moléculas o la supresión de su propio efecto mediante una inhibición por retroalimentación. El término "receptor" designa una proteína asociada a células que se une a una molécula bioactiva y media en el efecto del ligando sobre la célula. "Receptor", como se utiliza en el presente documento, incluye, pero no sin limitación, receptores de citocinas y receptores de quimiocinas. Aunque muchos receptores de células naturales tienen dominios de unión que son similares a una inmunoglobulina, el término "receptor", como se utiliza en el presente documento, excluye específicamente las inmunoglobulinas que constituyen anticuerpos o receptores del MHC.

Los receptores que se unen a citocinas están compuestos normalmente por una o más proteínas de membrana integrales que tienen un componente extracelular que se une a la citocina con alta afinidad y transduce este evento de unión a la célula a través de las porciones transmembranaria y citoplásmica de las subunidades del receptor. La subunidad de unión al ligando de un receptor se denomina cadena alfa, mientras que otras subunidades de transducción de señales se denominan cadenas beta o cadenas gamma. Por lo tanto, la capacidad de secuestrar citocinas, hormonas o ligandos relacionados antes de unirse a sus receptores celulares naturales representa un mecanismo para alterar la patogenia de enfermedades, trastornos o afecciones.

Los receptores de citocinas se han agrupado en varias familias sobre la base de similitudes en sus dominios extracelulares de unión al ligando; el receptor de la superfamilia de las inmunoglobulinas, la familia de receptores de

citocinas de clase I, la familia de receptores de citocinas de clase II, la familia de receptores del TNF y la familia de receptores de quimiocinas. En el presente documento se describen restos de unión a citocinas que comprenden dominios de unión, porciones de dominios de unión o secuencias con identidad de secuencia con miembros de las familias de receptores anteriores (excepto para las inmunoglobulinas de anticuerpos, específicamente excluidas anteriormente) unidos a una o más secuencias XTEN, en donde el resto de unión a citocinas tiene al menos una porción de la afinidad de unión en comparación con el receptor natural y, cuando el ligando es secuestrado, es capaz de interferir en la unión del ligando diana al receptor natural. En el presente documento se describe una proteína de fusión de unión al receptor que comprende dominios de unión, porciones de dominios de unión o secuencias con identidad de secuencia con receptores de la superfamilia de las inmunoglobulinas unidos a una o más secuencias XTEN, en donde los dominios de unión tienen al menos una porción de la afinidad de unión en comparación con el receptor natural y, cuando el ligando es secuestrado, es capaz de interferir en la unión del ligando diana al receptor natural.

Por "resto de unión a citocinas" lo que se entiende es al menos esa porción mínima del dominio extracelular de unión al ligando de un receptor de citocinas necesaria para unirse a la citocina. Por ejemplo, mientras que algunos receptores de citocinas tienen los componentes denominados α y β, un experto en la técnica sabría qué componente del receptor es el componente de transducción de señales y cuál es el componente de unión al ligando. Por lo tanto, para llevar a la práctica la presente divulgación y crear una proteína de unión de alta afinidad por una citocina, un experto en la técnica crearía un ácido nucleico aislado que comprende una secuencia nucleotídica que codifica un primer componente de polipéptido de fusión que comprende la secuencia de aminoácidos de la porción de unión a citocinas del dominio extracelular del componente determinante de la especificidad del receptor; una secuencia nucleotídica que codifica un segundo componente de polipéptido de fusión que comprende la secuencia de aminoácidos de una porción de unión a citocinas del dominio extracelular que puede ser el mismo o puede ser diferente del primero, y una secuencia nucleotídica que codifica un tercer componente de polipéptido de fusión que comprende la secuencia de aminoácidos de un XTEN que une los componentes de unión para crear un gen que codifica una proteína de fusión de unión a citocinas de alta afinidad por la citocina diana. La metodología general para el ensamblaje de los componentes, la expresión y la recuperación, seguida de la caracterización de la proteína de fusión de unión, se ilustra en la FIG. 9.

Como se utiliza en el presente documento, la propia expresión "dominio de unión" puede incluir una estructura de múltiples dominios en donde dos o más dominios individuales deben estar estrechamente unidos para formar una única región de unión. Algunos ejemplos de lo anterior son los receptores naturales de VEGF, flt-1 y flk-1, teniendo ambos siete dominios similares a inmunoglobulina en cada uno de los dos brazos que forman las regiones extracelulares de unión al ligando de los receptores (Matthews, et al., PNAS 88: 9026 (1991)). Sin embargo, se ha demostrado que las proteínas recombinantes que comprenden dos o tres de los siete dominios pueden, cuando se emparejan con un segundo brazo correspondiente que comprende dos o tres dominios, unirse al VEGF con alta afinidad (véanse las Patentes de EE. UU. n.º 5.952.199, 7.374.757). Descritas en el presente documento, las proteínas de fusión de unión a citocinas VEGF comprenden pares de un dominio Ig 2 (D2) y un dominio Ig 3 (D3) derivados de uno o más receptores del VEGF, como se muestra en la FIG. 4. En una realización de lo anterior, el receptor del VEGF es Flt-1. En otra realización de lo anterior, el receptor del VEGF es Flk-1. En otra realización más de lo anterior, el receptor del VEGF es Flt-4. Las unidades de dominio pueden comprender dominios del receptor del VEGF conectados directamente entre sí, u opcionalmente, conectados a través de espaciadores o conectores, tales como fragmentos de XTEN.

"Dominio similar a inmunoglobulina" o "dominio similar a Ig" o "dominio de unión al ligando" se refiere a dominios independientes y distintos que se encuentran en la región extracelular de unión al ligando de los receptores de citocinas, y se pretende específicamente que el término englobe no solo el dominio natural completo, sino también variantes de inserción, deleción y sustitución de los mismos que conservan al menos una parte de la afinidad de unión del dominio natural. Será perfectamente evidente para los expertos familiarizados con la materia que se pueden obtener numerosas variantes de los dominios o combinaciones de los dominios de las proteínas de unión a citocinas que conservarán sustancialmente las mismas características funcionales que el dominio natural.

En el presente documento se divulgan proteínas de fusión de unión a citocinas que comprenden dominios de unión derivados de receptores multiméricos. Los receptores multiméricos incluyen homodímeros (por ejemplo, las isoformas αα y ββ del receptor del PDGF, el receptor de la eritropoyetina, MPL y el receptor de G-CSF), heterodímeros cuyas subunidades tienen, cada una, dominios de unión al ligando y efectores (por ejemplo, la isoforma αβ del receptor del PDGF), y multímeros que tienen subunidades componentes con funciones dispares (por ejemplo, receptores de IL-2, IL-3, IL-4, IL-5, IL-6, IL-7 y GM-CSF).

La divulgación contempla varias construcciones diferentes que emplean dominios de unión a citocinas conocidos por tener afinidad por ligandos diana de interés. Las realizaciones preferidas de la invención incluyen composiciones de proteína de fusión capaces de unirse a un ligando que comprenden al menos un primer y un segundo dominio de unión polipeptídico unido operativamente a un componente de unión, en donde el componente de unión comprende una secuencia de XTEN flexible. En una realización de lo anterior, la composición comprende además al menos un segundo XTEN como portador, en donde el portador puede mejorar las propiedades farmacocinéticas o farmacéuticas de la proteína de fusión. En otra realización de lo anterior, la composición comprende un XTEN unido al extremo

amínico de la proteína de fusión, en donde la expresión de la proteína de fusión en una célula hospedadora transformada con un vector de expresión adecuado que comprende un polinucleótido que codifica la proteína de fusión que comprende la secuencia aminoterminal está mejorada en comparación con la expresión de una proteína de fusión correspondiente de un polinucleótido que no comprende la secuencia codificante del XTEN aminoterminal.

En una característica particular de las proteínas de fusión de unión a citocinas, la característica no estructurada del conector XTEN permite que los dominios de unión respectivos adopten una configuración flexible para orientarse libre y correctamente con el ligando para una unión óptima. Cuando se utiliza "configuración flexible" en referencia a las composiciones de proteína de unión a citocinas de la presente divulgación, significa que las secuencias polipeptídicas de los receptores o dominios de la proteína de fusión no se mantienen en una configuración particular relativa entre sí, sino que son libres de moverse, en condiciones fisiológicas, en la medida en que los polipéptidos XTEN anclados a los polipéptidos receptores lo permitan. Tomando como base estas características, se cree que es más probable que un dominio de unión de la proteína de fusión se encuentre y se una al ligando diana, en condiciones fisiológicas, en comparación con una construcción en la que los dominios se mantienen en una orientación fija, tales como las trampas de receptores en las que los dominios de unión se fusionan y se mantienen en su lugar mediante la dimerización entre dominios Fc o cadenas pesadas de IgG (véase, por ejemplo, la patente de EE. UU. nº 7.417.134). Sin pretender quedar ligado a teoría alguna en particular, al encontrarse con el ligando diana en condiciones fisiológicas o de ensayo, los dominios de unión de la proteína de fusión inventiva pueden orientarse mutuamente sustancialmente como en un receptor de citocinas natural, adoptando una configuración restringida para secuestrar el ligando. Cuando se utiliza "secuestrar" o "secuestro" en referencia a la actividad de una proteína de fusión de unión a citocinas de la presente invención significa que la proteína de fusión se une al ligando diana para formar un complejo sustancialmente no funcional, lo que interfiere con su capacidad de unirse a un receptor natural. Como se utiliza en el presente documento, "complejo sustancialmente no funcional" significa que la actividad residual del ligando unido y/o su capacidad para unirse a su receptor afín sería menor de aproximadamente un 60 %, o menor de aproximadamente un 50 %, o menor de aproximadamente un 40 %, o menor de aproximadamente un 30 %, o menor de aproximadamente un 20 %, o menor de aproximadamente un 10 %, o menor de aproximadamente un 5 %, o menor de aproximadamente un 1 % en comparación con el ligando no unido con un receptor natural. Las proteínas de fusión de unión a citocinas son, por lo tanto, antagonistas.

En vista de lo anterior, las composiciones de proteína de fusión de unión pueden tener una conformación "abierta" (FIG. 4A) de modo que, cuando una molécula diana está muy cerca de la proteína de fusión de unión, la interacción con los dominios permite un cambio en la conformación en donde ambas unidades de dominio entran en asociación con el ligando diana, creando una conformación "cerrada" (FIG. 4B), secuestrando eficazmente la molécula diana.

Las proteínas de fusión de cualquiera de las realizaciones de proteína de unión a citocinas anteriores pueden comprender además una segunda y opcionalmente una tercera proteína XTEN, como se ilustra en la FIG. 4, para impartir las características mejoradas como portador o como XTEN aminoterminal, como se ha descrito anteriormente.

Descritas en el presente documento, las proteínas de unión a citocinas son proteínas de fusión multivalentes que comprenden dos dominios de región de unión que son homólogos, el conector XTEN, y pueden comprender además las secuencias XTEN adicionales en el extremo amínico y/o carboxílico.

En otra realización de las proteínas de fusión de unión a citocinas de la divulgación, los conectores, así como el portador XTEN largo y el XTEN aminoterminal, pueden comprender una secuencia que puede ser un fragmento de o que presenta al menos aproximadamente un 80 % de identidad de secuencia, o un 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de identidad de secuencia con una secuencia seleccionada de una cualquiera de las tablas 4 u 11-15.

Proteínas de fusión de unión al receptor

En el presente documento se describen proteínas de fusión de unión que comprenden al menos una primera región de unión que comprende un primer y opcionalmente un segundo dominio de unión derivado de dominios similares a Ig de receptores celulares unidos a al menos un primer XTEN. Un ejemplo no limitativo de lo anterior es una proteína de unión con dominios de unión similares a Ig derivados del receptor del VEGF, con el XTEN unido al extremo amínico o carboxílico de las regiones de unión, como se ilustra en las FIG. 4 y 5. En una realización particular de lo anterior, la proteína de fusión de unión puede comprender dos regiones de unión similares a Ig unidas por un conector XTEN corto de aproximadamente 20 a aproximadamente 200 restos de aminoácidos (por ejemplo, un fragmento de una secuencia de XTEN de la tabla 4), y un portador XTEN más largo, en donde las regiones de unión al VEGF pueden unir y secuestrar sustancialmente el VEGF dimérico. En otra realización, la proteína de fusión de unión puede comprender dos conjuntos de dos regiones de unión similares a Ig, que pueden ser idénticas o pueden ser dominios diferentes, unidas por un conector XTEN corto de aproximadamente 20 a aproximadamente 200 restos de aminoácidos, y uno o más portadores XTEN más largos, en donde las regiones de unión al VEGF pueden unir y secuestrar sustancialmente el VEGF dimérico. La proteína de fusión de unión al receptor también puede comprender un XTEN aminoterminal, también ilustrado en la FIG.6, en donde la expresión de la proteína de fusión en una célula hospedadora transformada con un vector de expresión adecuado que comprende un polinucleótido que codifica la proteína de fusión que comprende la secuencia aminoterminal está mejorada en comparación con la expresión de una

proteína de fusión correspondiente de un polinucleótido que no comprende la secuencia codificante del XTEN aminoterminal. En las realizaciones anteriores descritas anteriormente en este párrafo, los conectores, así como el portador XTEN largo y el XTEN aminoterminal, pueden comprender una secuencia que puede ser un fragmento de o que presenta al menos aproximadamente un 80 % de identidad de secuencia, o un 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o 100% de identidad de secuencia con una secuencia seleccionada de una cualquiera de las tablas 4 u 11-15. En otra realización, el conector puede ser una secuencia en la que al menos el 80 % de los restos están comprendidos por aminoácidos glicina, serina y/o glutamato, tales como, pero sin limitación, una secuencia con aproximadamente un 80-100 % de identidad de secuencia con la secuencia GSGEGSEGEGGGEGSEGEGSGEGGEGEGSG (SEQ ID NO: 1) o una porción o un multímero de la misma.

En una característica particular de las realizaciones descritas anteriormente, las proteínas de fusión de unión a citocinas se pueden producir como monómeros funcionalmente activos, una característica que se cree que requiere menos etapas de fabricación y da como resultado un producto más homogéneo en comparación con construcciones que requieren un proceso de dimerización (por ejemplo, conjugados de Fc) con el fin de recuperar una molécula funcional.

Anticuerpos biespecíficos

Los anticuerpos biespecíficos son anticuerpos que tienen especificidades de unión para al menos dos epítopos diferentes. Un ejemplo no limitativo de una proteína de fusión de unión biespecífica sería una con dos restos de direccionamiento que pueden unirse a dos epítopos diferentes de la proteína ErbB2. Por ejemplo, un resto de direccionamiento puede unirse a un epítopo en el Dominio 1 de la ErbB2, tal como el epítopo 7C2/7F3, el otro puede unirse a un epítopo de la ErbB2 diferente, por ejemplo el epítopo 4D5. Otras proteínas de fusión de unión biespecíficas de este tipo pueden combinar un sitio de unión a la ErbB2 con un sitio o sitios de unión para el EGFR, la ErbB3 y/o la ErbB4. Alternativamente, un resto de direccionamiento anti-ErbB2 puede combinarse con un resto de direccionamiento que se une a una molécula desencadenante en un leucocito, tal como una molécula del receptor de linfocitos T (por ejemplo, CD2 o CD3), o receptores Fc para IgG (FcyR), tales como FcyRI (CD64), FcyRII (CD32) y FcyRIII (CD16) para enfocar los mecanismos de defensa celular a la célula que expresa la ErbB2. Las proteínas de fusión de unión biespecíficas se pueden preparar dirigiendo restos que son scFv multivalentes, anticuerpos de dominio multivalente o diacuerpos.

(e) Configuraciones de proteínas de fusión de unión

La divulgación contempla diferentes configuraciones de proteínas de fusión de unión, que incluyen, pero sin limitación, aquellas que comprenden restos de direccionamiento caracterizados como scFv, diacuerpos, anticuerpos de dominio, anticuerpos bifuncionales o proteínas de fusión de unión a citocinas, unidos a uno o más XTEN, y que opcionalmente tienen uno o más conectores. Para proteínas de fusión de unión con un único resto de direccionamiento, la divulgación proporciona una proteína de fusión de unión monómera de fórmula I:

$$(XTEN)_x\text{-}TM\text{-}(XTEN)_y \qquad I$$

en donde independientemente en cada aparición: XTEN es un polipéptido recombinante expandido como se ha descrito anteriormente, x es 0 o 1; y es 0 o 1, en donde x+y ≥1; y TM es un resto de direccionamiento con afinidad de unión por una diana, preferentemente una diana seleccionada de la tabla 1 o la tabla 2. En una realización, el TM comprende dos o más dominios de unión que pueden unirse mediante una secuencia conectora de 1 a aproximadamente 300 restos de aminoácidos que tiene una conformación flexible y no estructurada. En una realización, la secuencia conectora es un XTEN que tiene al menos aproximadamente 12 a aproximadamente 300 aminoácidos que presenta una característica flexible y no estructurada, tal como un fragmento de un XTEN, de tal manera que los dominios de unión pueden así orientarse mutuamente entre sí y el ligando y adoptar una configuración restringida para unirse al ligando diana. En otra realización, el conector puede ser una secuencia en la que al menos el 80 % de los restos están comprendidos por aminoácidos glicina, serina y/o glutamato, tales como, pero sin limitación, una secuencia con aproximadamente un 80-100 % de identidad de secuencia con la secuencia GSGEGSEGEGGGEGSEGEGSGEGGEGEGSG (SEQ ID NO: 1) o una porción o un multímero de la misma. En otra realización, en donde el TM comprende dos o más dominios de unión, los dominios de unión pueden ser idénticos o pueden ser diferentes, y pueden comprender, por ejemplo, secuencias derivadas de dichas secuencias V_L o V_H, dominios de unión al receptor, o dominios de unión similares a Ig, según sea necesario, para unirse con suficiente afinidad al ligando diana.

En el presente documento se describe una proteína de fusión de unión multivalente con dos restos de unión de fórmula II:

$$(XTEN)_x\text{-}TM1\text{-}L\text{-}TM2\text{-}(XTEN)_y \qquad II$$

en donde independientemente en cada aparición: XTEN es un polipéptido recombinante expandido como se ha descrito anteriormente, x es 0 o 1; y es 0 o 1, en donde x+y ≥1; TM1 es un resto de direccionamiento con afinidad de

unión por una diana, preferentemente una diana seleccionada de la tabla 1 o la tabla 2; TM2 es un resto de direccionamiento con afinidad de unión por el ligando diana que puede ser idéntico o puede ser diferente a TM1; y L es una secuencia conectora que tiene entre 1 y aproximadamente 300 restos de aminoácidos en donde la secuencia conectora está unida covalentemente al extremo carboxílico de TM1 y al extremo amínico de TM2. En una realización, cada uno de los respectivos TM puede comprender dos o más dominios de unión que pueden estar unidos por una secuencia conectora adicional de 1 a aproximadamente 300 restos de aminoácidos que tiene una conformación flexible y no estructurada. En una realización, la secuencia conectora es un XTEN o un fragmento de un XTEN de una cualquiera de las tablas 4 u 11-15 que tiene al menos aproximadamente 12 a aproximadamente 300 aminoácidos que presentan una característica flexible y no estructurada, de modo que los dominios de unión pueden así orientarse mutuamente entre sí y el ligando y adoptar una configuración restringida para unirse al ligando diana. En otra realización, el conector puede ser una secuencia en la que al menos el 80 % de los restos están comprendidos por aminoácidos glicina, serina y/o glutamato, tales como, pero sin limitación, una secuencia con aproximadamente un 80-100 % de identidad de secuencia con la secuencia GSGEGSEGEGGGEGSEGEGSGEGGEGEGSG (SEQ ID NO: 1) o una porción o un multímero de la misma. En otra realización, en donde el TM comprende dos o más dominios de unión, los dominios de unión pueden ser idénticos o pueden ser diferentes, y pueden comprender, por ejemplo, secuencias derivadas de dichas secuencias V_L o V_H, dominios de unión al receptor, o dominios de unión similares a Ig, según sea necesario, para unirse con suficiente afinidad al ligando diana.

En el presente documento se describe una proteína de fusión de unión multivalente con dos restos de unión de fórmula III:

TM1-XTEN-TM2 III

en donde independientemente en cada aparición: XTEN es un polipéptido recombinante expandido como se ha descrito anteriormente; TM1 es un resto de direccionamiento con afinidad de unión por una diana, preferentemente una diana seleccionada seleccionado de la tabla 1 o la tabla 2; y TM2 es un resto de direccionamiento con afinidad de unión por una diana, preferentemente una diana seleccionada seleccionado de la tabla 1 o la tabla 2 que puede ser idéntico o puede ser diferente a TM1. En una realización, cada uno de los respectivos TM puede comprender dos o más dominios de unión que pueden estar unidos por una secuencia conectora adicional de 1 a aproximadamente 300 restos de aminoácidos que tiene una conformación flexible y no estructurada. En una realización, la secuencia conectora es un XTEN o un fragmento de un XTEN de una cualquiera de las tablas 4 u 11-15 que tiene al menos aproximadamente 12 a aproximadamente 300 aminoácidos que presentan una característica flexible y no estructurada, de modo que los dominios de unión pueden así orientarse mutuamente entre sí y el ligando y adoptar una configuración restringida para unirse al ligando diana. En otra realización, el conector puede ser una secuencia en la que al menos el 80 % de los restos están comprendidos por aminoácidos glicina, serina y/o glutamato, tales como, pero sin limitación, una secuencia con aproximadamente un 80-100 % de identidad de secuencia con la secuencia GSGEGSEGEGGGEGSEGEGSGEGGEGEGSG (SEQ ID NO: 1) o una porción o un multímero de la misma. En otra realización, en donde el TM comprende dos o más dominios de unión, los dominios de unión pueden ser idénticos o pueden ser diferentes, y pueden comprender, por ejemplo, secuencias derivadas de dichas secuencias V_L o V_H, dominios de unión al receptor, o dominios de unión similares a Ig, según sea necesario, para unirse con suficiente afinidad al ligando diana.

En el presente documento se describe una proteína de fusión de unión multivalente con tres restos de unión de fórmula IV:

$(XTEN)_x$-TM1-L1-TM2-L2-TM3-$(XTEN)_y$ IV

en donde independientemente en cada aparición: XTEN es un polipéptido recombinante expandido como se ha descrito anteriormente, x es 0 o 1; y es 0 o 1, en donde $x+y \geq 1$; TM1 es un resto de direccionamiento con afinidad de unión por una diana, preferentemente una diana seleccionada de la tabla 1 o la tabla 2; TM2 es un resto de direccionamiento con afinidad de unión por una diana, preferentemente una diana seleccionada de la tabla 1 o la tabla 2 que puede ser idéntico o puede ser diferente a TM1; TM3 es un resto de direccionamiento con afinidad de unión por una diana, preferentemente una diana seleccionada de la tabla 1 o la tabla 2 que puede ser idéntico o puede ser diferente a TM1 o TM2; L1 es una secuencia conectora que tiene entre 1 y aproximadamente 300 restos de aminoácidos como se ha descrito para la fórmula II, y en donde la secuencia conectora está unida covalentemente al extremo carboxílico de TM1 y al extremo amínico de TM2; y L2 es una secuencia conectora que puede ser idéntica o diferente a L1, que tiene entre 1 y aproximadamente 300 restos de aminoácidos como se ha descrito para la fórmula II, y en donde la secuencia conectora está unida covalentemente al extremo carboxílico de TM2 y al extremo amínico de TM3. En una realización, cada uno de los respectivos TM puede comprender dos o más dominios de unión que pueden estar unidos por una secuencia conectora adicional de 1 a aproximadamente 300 restos de aminoácidos que tiene una conformación flexible y no estructurada. En una realización, la secuencia conectora es un XTEN que tiene al menos aproximadamente 12 a aproximadamente 300 aminoácidos que presenta una característica flexible y no estructurada, de tal manera que los dominios de unión pueden así orientarse mutuamente entre sí y el ligando y adoptar una configuración restringida para unirse al ligando diana. En otra realización, en donde el TM comprende dos o más dominios de unión, los dominios de unión pueden ser idénticos o pueden ser diferentes, y pueden comprender,

por ejemplo, secuencias derivadas de dichas secuencias V_L o V_H, dominios de unión al receptor, o dominios de unión similares a Ig, según sea necesario, para unirse con suficiente afinidad al ligando diana.

La invención contempla configuraciones adicionales y alternativas de las realizaciones anteriores, que incluyen restos de direccionamiento adicionales, regiones de unión, conectores y XTEN configurados en permutaciones alternativas de orden, de amino- a carboxiterminal, para los diferentes componentes. Las proteínas de fusión de unión de las realizaciones divulgadas en el presente documento presentan una o más o cualquier combinación de las propiedades y/o las realizaciones como se detalla en el presente documento.

(f) Configuraciones de proteínas de fusión de unión con secuencias espaciadoras y de escisión

La divulgación contempla configuraciones de proteínas de fusión de unión que incluyen, pero sin limitación, aquellas que comprenden restos de direccionamiento caracterizados como scFv, diacuerpos, anticuerpos de dominio, anticuerpos bifuncionales o proteínas de fusión de unión a citocinas, en las que el XTEN puede estar unido a restos de direccionamiento mediante secuencias espaciadoras incorporadas en o adyacentes al XTEN que están diseñadas para incorporar o mejorar una funcionalidad o una propiedad a la composición, o como ayuda en el ensamblaje o la fabricación de las composiciones de proteína de fusión de unión. Dichas propiedades incluyen, pero sin limitación, la inclusión de una secuencia o secuencias de escisión para permitir la liberación de los componentes, la inclusión de aminoácidos compatibles con sitios de restricciones nucleotídicas para permitir la unión de nucleótidos que codifican un XTEN a nucleótidos que codifican un resto de direccionamiento o que facilitan la construcción de vectores de expresión, o para reducir el impedimento estérico en regiones de las proteínas de fusión.

Se puede introducir una secuencia espaciadora entre una secuencia de XTEN y un componente de resto de direccionamiento para disminuir el impedimento estérico, de tal manera que el componente de resto de direccionamiento pueda asumir su estructura terciaria deseada y/o interactuar apropiadamente con su diana. Algunos espaciadores y métodos de identificación de espaciadores deseables, pueden consultarse, por ejemplo, en George, et al (2003) Protein Engineering 15: 871-879. En una realización, el espaciador comprende una o más secuencias peptídicas que tienen entre 1-50 restos de aminoácidos de longitud, o aproximadamente 1-25 restos, o aproximadamente 1-10 restos de longitud. Las secuencias espaciadoras, exclusivas de los sitios de escisión, pueden comprender cualquiera de los 20 aminoácidos L naturales, y preferentemente tendrán unas propiedades similares al XTEN en que 1) comprenderán aminoácidos hidrófilos que están estéricamente sin obstáculos tales como, pero sin limitación, glicina (G), alanina (A), serina (S), treonina (T), glutamato (E), prolina (P) y aspartato (D); y 2) serán sustancialmente no repetitivas. En algunos casos, el espaciador puede ser poliglicinas o polialaninas, o es predominantemente una mezcla de combinaciones de restos de glicina, serina y alanina. En una realización, una secuencia espaciadora, exclusiva de aminoácidos del sitio de escisión, tiene aproximadamente 1 a 10 aminoácidos que consisten en aminoácidos seleccionados entre glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P), y están sustancialmente desprovistos de estructura secundaria; por ejemplo, menos de aproximadamente un 10 % o menos de aproximadamente un 5 % según se determina mediante los algoritmos de Chou-Fasman y/o de GOR. En una realización, la secuencia espaciadora es GPEGPS (SEQ ID NO: 145). En otra realización, la secuencia espaciadora es GPEGPS (SEQ ID NO: 145) ligada a una secuencia de escisión de la tabla 7. Además, las secuencias espaciadoras están diseñadas para evitar la introducción de epítopos de linfocitos T, determinación de los cuales se ha descrito anteriormente y en los ejemplos.

En una realización, la proteína de fusión de unión comprende una o más secuencias espaciadoras unidas en la unión o uniones entre la secuencia de carga activa y el uno o más XTEN incorporados en la proteína de fusión, en donde las secuencias espaciadoras comprenden aminoácidos que son compatibles con nucleótidos que codifican sitios de restricción. En otra realización, la proteína de fusión de unión comprende una o más secuencias espaciadoras unidas en la unión o uniones entre la secuencia de carga activa y el uno o más XTEN incorporados en la proteína de fusión, en donde las secuencias espaciadoras comprenden aminoácidos que son compatibles con nucleótidos que codifican sitios de restricción, y los aminoácidos y el uno o más aminoácidos de la secuencia espaciadora se eligen entre glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P). En otra realización, la proteína de fusión de unión comprende una o más secuencias espaciadoras unidas en la unión o uniones entre la secuencia de carga activa y el uno o más XTEN incorporados en la proteína de fusión, en donde las secuencias espaciadoras comprenden aminoácidos que son compatibles con nucleótidos que codifican sitios de restricción y la una o más secuencias espaciadoras se eligen entre las secuencias de la tabla 6. La secuencia exacta de cada secuencia espaciadora se elige para que sea compatible con sitios de clonación en vectores de expresión que se utilizan para una construcción de proteína de fusión de unión particular. Para realizaciones en las que hay un único XTEN unido al extremo amínico o carboxílico, solo se requeriría una única secuencia espaciadora en la unión de los dos componentes. Como sería evidente para un experto en la materia, las secuencias espaciadoras que comprenden aminoácidos compatibles con sitios de restricción podrían omitirse de la construcción cuando se genera sintéticamente un gen completo de proteína de fusión, en lugar de ligarse utilizando el resto de direccionamiento y genes codificantes de XTEN.

Tabla 6: Secuencias espaciadoras compatibles con sitios de restricción

Secuencia espaciadora	Enzima de restricción
GSPG (SEQ ID NO: 146)	BsaI
ETET (SEQ ID NO: 147)	BsaI
PGSSS (SEQ ID NO: 148)	BbsI
GAP	AscI
GPA	FseI
GPSGP (SEQ ID NO: 149)	SfiI
AAA	SacII
TG	AgeI
GT	KpnI

En algunas realizaciones, una secuencia espaciadora de una composición de proteína de fusión de unión comprende una o más secuencias de escisión, que son idénticas o diferentes, en donde la secuencia de escisión puede ser activada por una proteasa para liberar la secuencia o secuencias XTEN de la proteína de fusión. En una realización, la incorporación de la secuencia de escisión en la proteína de fusión de unión está diseñada para permitir la liberación de un resto de direccionamiento que se vuelve activo o más activo tras su liberación del componente XTEN. Las secuencias de escisión se ubican lo suficientemente cerca de las secuencias del resto de direccionamiento, generalmente a 18, o a 12, o a 6, o a 2 aminoácidos de la secuencia del resto de direccionamiento, de modo que cualquier resto restante unido al resto de direccionamiento después de la escisión no interfiere apreciablemente en la actividad (por ejemplo, tal como la unión a una diana) pero proporciona suficiente acceso a la proteasa para que pueda efectuar la escisión de la secuencia de escisión. En algunos casos, la proteína de fusión de unión que comprende las secuencias de escisión también tendrá uno o más aminoácidos de la secuencia espaciadora entre el resto de direccionamiento y la secuencia de escisión, o el XTEN y la secuencia de escisión, para facilitar el acceso de la proteasa a la secuencia de escisión; comprendiendo los aminoácidos espaciadores cualquier aminoácido natural, incluyendo glicina, serina y alanina como aminoácidos preferidos. En una realización, el sitio de escisión es una secuencia que puede ser escindida por una proteasa endógena del sujeto mamífero de modo que la proteína de fusión puede ser escindida después de la administración a un sujeto. En una realización de la construcción anterior, el resto de direccionamiento que se libera de la proteína de fusión mediante la escisión de la secuencia de escisión presenta al menos aproximadamente dos veces, o al menos aproximadamente tres veces, o al menos aproximadamente cuatro veces, o al menos aproximadamente cinco veces, o al menos aproximadamente seis veces, o al menos aproximadamente ocho veces, o al menos aproximadamente diez veces, o al menos aproximadamente 20 veces más actividad en comparación con la proteína de fusión de unión intacta.

Algunos ejemplos de sitios de escisión contemplados por la invención incluyen, pero sin limitación, una secuencia polipeptídica escindible por una proteasa endógena de mamífero seleccionada entre FXIa, FXIIa, calicreína, FVIIIa, FVIIIa, FVIIIa, FXa, FIIa (trombina), elastasa-2, granzima B, MMP-12, MMP-13, MMP-17 o MMP-20, o por proteasas que no son de mamífero tales como TEV, enterocinasa, proteasa PreScission™ (proteasa de rinovirus 3C) y sortasa A. En la técnica se conocen secuencias que se sabe que han sido escindidas por las proteasas anteriores y por otras. En la tabla 7 se presentan secuencias de escisión ilustrativas contempladas por la invención y los respectivos sitios de corte dentro de las secuencias, así como variantes de secuencia de las mismas.

En una realización, la invención proporciona proteínas de fusión de unión que comprenden una o más secuencias de escisión posicionadas operativamente para liberar el resto de direccionamiento de la proteína de fusión tras la escisión, en donde la una o más secuencias de escisión tiene al menos aproximadamente un 86 %, o al menos aproximadamente un 92 % o más de identidad de secuencia con una secuencia seleccionada de la tabla 7.

En algunas realizaciones, solo los dos o tres aminoácidos que flanquean ambos lados del sitio de corte (de cuatro a seis aminoácidos en total) son incorporados a la secuencia de escisión que, a su vez, es incorporada a las proteínas de fusión de las realizaciones. En otras realizaciones, la secuencia de escisión incorporada de la tabla 7 puede tener una o más deleciones o inserciones o una o dos o tres sustituciones de aminoácidos para cualquiera de uno o dos o tres aminoácidos de la secuencia conocida, en donde las deleciones, las inserciones o sustituciones dan como resultado una susceptibilidad reducida o mejorada, pero no una ausencia de susceptibilidad a la proteasa, lo que da como resultado una capacidad para adaptar la tasa de liberación del resto de direccionamiento del XTEN. Algunas sustituciones ilustrativas se muestran en la tabla 7.

Tabla 7: Secuencias de escisión de proteasas

Proteasa que actúa sobre la secuencia	Secuencia de escisión ilustrativa	SEQ ID NO:	Sitio de corte mínimo*	SEQ ID NO:
FXIa	KLTR↓AET	150	KD/FL/T/R↓VA/VE/GT/GV	
FXIa	DFTR↓VVG	151	KD/FL/T/R↓VA/VE/GT/GV	
FXIIa	TMTR↓IVGG	152	NA	
Calicreína	SPFR↓STGG	153	-/-/FL/RY↓SR/RT/-/-	
FVIIa	LQVR↓IVGG	154	NA	
FIXa	PLGR↓IVGG	155	-/-/G/R↓-/-/-/-	
FXa	IEGR↓TVGG	156	IA/E/GFP/R↓STI/VFS/-/G	
FIIa (trombina	LTPR↓SLLV	157	-/-/PLA/R↓SAG/-/-/-	
Elastasa-2	LGPV↓SGVP	158	-/-/-/VIAT↓-/-/-/-	
Granzima-B	VAGD↓SLEE	159	V/-/-/D↓-/-/-/-	
MMP-12	GPAG↓LGGA	160	G/PA/-/G↓L/-/G/-	161
MMP-13	GPAG↓LRGA	162	G/P/-/G↓L/-/GA/-	163
MMP-17	APLG↓LRLR	164	-/PS/-/-↓LQ/-/LT/-	
MMP-20	PALP↓LVAQ	165	NA	
TEV	ENLYFQ↓G	166	ENLYFQ↓G/S	167
Enterocinasa	DDDK↓IVGG	168	DDDK↓IVGG	169
Proteasa 3C (PreScission™)	LEVLFQ↓GP	170	LEVLFQ↓GP	171
Sortasa A	LPKT↓GSES	172	L/P/KEAD/T↓G/-/EKS/S	173

↓indica el sitio de escisión NA: no aplicable

* la lista de múltiples aminoácidos antes, entre o después de una barra indica aminoácidos alternativos que pueden sustituirse en la posición; "-" indica que cualquier aminoácido puede sustituirse por el aminoácido correspondiente indicado en la columna central

(g) Métodos de uso de proteínas de fusión de unión

En el presente documento se describe un método para conseguir un efecto beneficioso en una enfermedad, un trastorno o una afección mediada por una proteína de fusión de unión. En una realización, la invención proporciona el uso de una proteína de fusión de unión derivada de un anticuerpo progenitor que se une a una diana seleccionada del grupo que consiste en las dianas de la tabla 1 en el tratamiento de una enfermedad, un trastorno o una afección en un sujeto que lo necesite mediante la administración de una cantidad terapéuticamente eficaz de la proteína de fusión de unión, en donde dicha administración conduce a la erradicación o la mejora de uno o más de los síntomas fisiológicos o clínicos asociados con el trastorno subyacente, de modo que se observa una mejora en el sujeto, a pesar de que el sujeto todavía puede estar afectado por el trastorno subyacente. En el presente documento se describe un método para tratar una enfermedad, un trastorno o una afección en un mamífero que comprende administrar al mamífero una cantidad terapéuticamente eficaz de una proteína de fusión de unión que comprende uno o más restos de direccionamiento dirigidos a una o más dianas seleccionadas de la tabla 1, unidos a una o más moléculas de secuencias XTEN y, opcionalmente, a uno o más conectores, para formar la proteína de fusión de unión, en donde la unión no altera sustancialmente las propiedades funcionales esenciales de la proteína de fusión de unión de afinidad de unión y semivida terminal sostenida o tasa de eliminación sérica reducida en comparación con la del resto de direccionamiento progenitor del que deriva la proteína de fusión de unión, y en donde la administración de la proteína de fusión de unión consigue un efecto terapéutico beneficioso. La cantidad eficaz puede producir un efecto beneficioso para ayudar a tratar (por ejemplo, curar o reducir la gravedad) o prevenir (por ejemplo, reducir la probabilidad de aparición o la gravedad) una enfermedad, un trastorno o una afección, tal como, pero no sin limitación un cáncer, una enfermedad o afección cardiovascular, una enfermedad infecciosa, una afección inflamatoria, una afección respiratoria, un rechazo de trasplante de órganos o una enfermedad metabólica mediada por o asociada con una o más dianas, preferentemente seleccionada de la tabla 1.

En una realización, el método comprende administrar una cantidad terapéuticamente eficaz de una composición farmacéutica que comprende una composición de proteína de fusión de unión que comprende uno o más restos de direccionamiento unidos a una o más secuencias XTEN y al menos un portador farmacéuticamente aceptable a un sujeto que lo necesite, lo que da como resultado una mejora en al menos un parámetro, condición fisiológica o resultado clínico mediado por el componente o componentes del resto de direccionamiento. El método contempla la administración de la composición farmacéutica por cualquier vía apropiada para la enfermedad, el trastorno o la afección que se está tratando, incluyendo por vía subcutánea, intramuscular, intravítrea o intravenosa.

Los métodos de la divulgación incluyen la administración de dosis consecutivas de una cantidad terapéuticamente eficaz de la composición farmacéutica durante un período de tiempo suficiente para conseguir y/o mantener el parámetro o efecto clínico deseado, y dichas dosis consecutivas de una cantidad terapéuticamente eficaz establecen la pauta posológica terapéuticamente eficaz para la composición farmacéutica; es decir, el programa para las dosis administradas consecutivamente, en donde las dosis se administran en cantidades terapéuticamente eficaces para dar como resultado un efecto beneficioso sostenido sobre cualquier signo o síntoma clínico, aspecto, parámetro medido o característica de un cuadro clínico o afección metabólica, que incluye, pero no sin limitación, los descritos en el presente documento.

Una cantidad terapéuticamente eficaz de la composición farmacéutica puede variar según factores tales como el cuadro clínico, la edad, el sexo y el peso del individuo, y la capacidad del anticuerpo o de la porción de anticuerpo para provocar una respuesta deseada en el individuo. Una cantidad terapéuticamente eficaz es también una en la que cualquier efecto tóxico o perjudicial de la proteína de fusión de unión se ve superado por los efectos terapéuticamente beneficiosos. Una cantidad profilácticamente eficaz se refiere a una cantidad de composición farmacéutica requerida durante el período de tiempo necesario para conseguir el resultado profiláctico deseado.

Para los métodos divulgados, se prefieren composiciones o composiciones farmacéuticas de proteína de unión de acción más larga, con el fin de mejorar la comodidad del paciente, aumentar el intervalo entre dosis y reducir la cantidad de fármaco necesaria para conseguir un efecto sostenido. En una realización, un método de tratamiento comprende la administración de una dosis terapéuticamente eficaz de una proteína de fusión de unión a un sujeto que lo necesite, que da como resultado un aumento en el tiempo invertido dentro de una ventana terapéutica establecida para la proteína de fusión de la composición farmacéutica en comparación con el componente o componentes del resto de direccionamiento correspondientes no unidos al XTEN y administrados a una dosis comparable a un sujeto. En una realización, el aumento en el tiempo invertido dentro de la ventana terapéutica es al menos aproximadamente tres veces, o al menos aproximadamente cuatro veces, o al menos aproximadamente cinco veces, o al menos aproximadamente seis veces, o al menos aproximadamente ocho veces, o al menos aproximadamente 10 veces, o al menos aproximadamente 20 veces, o al menos aproximadamente 40 veces en comparación con el componente de resto de direccionamiento correspondiente no unido al XTEN y administrado a una dosis comparable a un sujeto. Los métodos prevén además que la administración de múltiples dosis consecutivas de una composición farmacéutica administrada utilizando una pauta posológica terapéuticamente eficaz a un sujeto que lo necesite puede dar como resultado un aumento en el tiempo entre máximos de $C_{máx}$ consecutivos y/o cubetas de $C_{mín}$ para los niveles en sangre de la proteína de fusión de unión en comparación a los restos de direccionamiento correspondientes no unidos al XTEN. En la realización anterior, el aumento en el tiempo transcurrido entre máximos de $C_{máx}$ consecutivos y/o cubetas de $C_{mín}$ puede ser al menos aproximadamente tres veces, o al menos aproximadamente cuatro veces, o al menos aproximadamente cinco veces, o al menos aproximadamente seis veces, o al menos aproximadamente ocho veces, o al menos aproximadamente 10 veces, o al menos aproximadamente 20 veces, o al menos aproximadamente 40 veces en comparación con el componente o componentes del resto de direccionamiento correspondientes no unidos al XTEN y administrados utilizando una pauta posológica comparable establecida para ese resto de direccionamiento. En las realizaciones descritas anteriormente en este párrafo, la administración de la proteína de fusión o la composición farmacéutica puede dar como resultado una mejora en al menos un parámetro que se sabe que es útil para evaluar las enfermedades, afecciones o trastornos del sujeto) utilizando una dosis unitaria más baja en moles de proteína de fusión en comparación con el componente o componentes de resto de direccionamiento correspondientes no unidos al XTEN y administrados a una dosis unitaria o pauta posológica comparable a un sujeto.

En una realización, la administración de una proteína de fusión de unión o una composición farmacéutica puede dar como resultado una mejora en uno de los parámetros clínicos, bioquímicos o fisiológicos que es mayor que la lograda mediante la administración del componente de resto de direccionamiento no unido a XTEN, determinados utilizando el mismo ensayo o basados en un parámetro clínico medido. En otra realización, la administración de la proteína de fusión de unión o la composición farmacéutica puede dar como resultado una mejora de dos o más parámetros clínicos o relacionados con el metabolismo, cada uno mediado por uno de los diferentes restos de direccionamiento que, en conjunto, dan como resultado un efecto mejorado en comparación con el componente de resto de direccionamiento no unido al XTEN, determinados utilizando los mismos ensayos o basados en parámetros clínicos medidos. En otra realización, la administración de la proteína de fusión de unión o la composición farmacéutica puede dar como resultado una actividad en uno o más de los parámetros clínicos o bioquímicos o fisiológicos que es de mayor duración que la actividad de uno de los componentes del resto de direccionamiento único no unido al XTEN, determinado utilizando el mismo ensayo o basado en un parámetro clínico medido.

En una realización, la proteína de fusión de unión se utiliza para tratar trastornos mediados por el VEGF. En particular, en el presente documento se describe un método para tratar una enfermedad mediada por el VEGF en un paciente humano con una proteína de fusión de unión que comprende uno o más de los restos de direccionamiento que se unen al VEGF humano, en donde la unión reduce la capacidad del VEGF de unirse a su receptor afín. Dichas proteínas de fusión de unión pueden tener aplicaciones profilácticas y terapéuticas en un amplio espectro de trastornos mediados por el VEGF, incluidas las patologías apoyadas por la proliferación de vasos sanguíneos, es decir, la angiogénesis, de una manera similar a la aplicación de anticuerpos anti-VEGF en el tratamiento de dichas indicaciones patológicas que se conoce en la técnica, indicaciones de tratamiento que incluyen tumores sólidos ((Kim et al. Nature 362: 841-844 (1993); Warren et al J. Clin. Invertir. 95: 1789-1797 (1995); Borgström et al Cancer Res. 56: 4032-4039 (1996); y Melnyk et al Cancer Res. 56: 921-924 (1996)) y síndromes neovasculares intraoculares tales como retinopatías proliferativas y degeneración macular senil (AMD) (Adamis et al. Arco. Ophthalmol. 114: 66-71 (1996)).

Las proteínas de fusión que comprenden el XTEN de la invención pueden aproximarse a la farmacocinética *in vivo* (por ejemplo, la semivida terminal) del anticuerpo completo. Dadas estas características, se cree que las proteínas de fusión de unión de la invención que comprenden restos de direccionamiento anti-VEGF muestran las mismas actividades *in vivo* o sustancialmente similares que el anticuerpo monoclonal anti-VEGF completo en un intervalo de diferentes parámetros, incluidas las características farmacocinéticas y los criterios de valoración terapéuticos en un modelo tumoral animal, respaldando la utilidad de las proteínas de fusión de unión en el mismo amplio espectro de indicaciones de enfermedad neovascular que responde al tratamiento con anticuerpos completos anti-VEGF.

Cualquier proteína de fusión de unión divulgada en el presente documento que comprende un resto de direccionamiento derivado de un anticuerpo o un fragmento anti-VEGF puede utilizarse ventajosamente en un método para tratar una enfermedad o un trastorno mediado por el VEGF, tales como trastornos neovasculares. En el presente documento se describe un método para tratar un trastorno neovascular en un paciente humano que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión o una composición farmacéutica en donde al menos un resto de direccionamiento de la proteína de fusión de unión comprende un sitio de unión al antígeno que se une al VEGF humano.

En el presente documento se describe un método para tratar un trastorno tumoral sólido en un paciente humano que comprende administrar al paciente una cantidad eficaz de una proteína de fusión de unión o una composición farmacéutica en donde al menos un resto de direccionamiento de la proteína de fusión de unión comprende un sitio de unión al antígeno que se une al VEGF humano. En otra realización más, el trastorno tumoral sólido en el método anterior se selecciona del grupo que consiste en carcinomas de mama, carcinomas de pulmón, carcinomas gástricos, carcinomas de esófago, carcinomas colorrectales, carcinomas de hígado, carcinomas de ovario, tecomas, arrenoblastomas, carcinomas cervicouterinos, carcinoma endometrial, hiperplasia endometrial, endometriosis, fibrosarcomas, coriocarcinoma, cáncer de cabeza y cuello, carcinoma nasofaríngeo, carcinomas laríngeos, hepatoblastoma, sarcoma de Kaposi, melanoma, carcinomas de piel, hemangioma, hemangioma cavernoso, hemangioblastoma, carcinomas de páncreas, retinoblastoma, astrocitoma, glioblastoma, schwannoma, oligodendroglioma, meduloblastoma, neuroblastomas, rabdomiosarcoma, sarcoma osteógeno, leiomiosarcomas, carcinomas de las vías urinarias, carcinomas de tiroides, tumor de Wilm, carcinoma de células renales, carcinoma de próstata, proliferación vascular anormal asociada a facomatosis, edema (tal como el asociado con tumores cerebrales) y síndrome de Meigs.

En el presente documento se describe un método para tratar un trastorno neovascular intraocular en un paciente humano que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión o de una composición farmacéutica en donde al menos un resto de direccionamiento comprende un sitio de unión al antígeno que se une al VEGF humano. En una realización adicional, el trastorno neovascular intraocular se selecciona del grupo que consiste en retinopatías diabéticas y otras retinopatías proliferativas, incluidas retinopatía del prematuro, fibroplasia retrocristaliniana, glaucoma neovascular y degeneración macular senil.

En el presente documento se describe un método para inhibir la angiogénesis en un paciente humano que comprende administrar al paciente una cantidad eficaz de una proteína de fusión de unión en donde al menos un resto de direccionamiento de la composición comprende un sitio de unión al antígeno que se une al VEGF humano.

En otra realización, la proteína de fusión de unión se usa para tratar trastornos mediados por células que expresan el HER2. En el presente documento se describe un método para tratar una enfermedad humana mediada por células que expresan el HER2 con una composición de proteína de fusión de unión que deriva de un anticuerpo progenitor que se une al HER2. Dichas composiciones tienen aplicaciones profilácticas y terapéuticas en un amplio espectro de trastornos mediados por células que expresan el HER2, que incluyen patologías apoyadas por la proliferación de células que expresan el HER2, tales como cánceres caracterizados por la sobreexpresión del HER2, de una manera similar a la aplicación de anticuerpos anti-Her2 completos en el tratamiento de dichas indicaciones de enfermedad que es conocida en la técnica, indicaciones de tratamiento que incluyen cánceres de mama, de ovario y de pulmón que sobreexpresan el HER2. La elección de un resto de direccionamiento para una proteína de fusión de unión que va a usarse en el método se puede determinar mediante ensayos de unión *in vitro* o ensayos de destrucción celular *in vitro* como se describe en los ejemplos o es conocido en la técnica. Por ejemplo, una proteína de fusión de unión candidata contra el HER2 puede usarse en ensayos de citotoxicidad con cultivos celulares de estirpes de cáncer de mama

humano tales como MCF-7, CAMA-1, SKBR-3 y BT-20, tal como se describe en la Patente de EE. UU. n.º 4753894. En el presente documento se describe un método para tratar un trastorno mediado por células que expresan el HER2 en un paciente humano que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión o de una composición farmacéutica en donde al menos un resto de direccionamiento de la proteína de fusión de unión comprende un sitio de unión al antígeno que se une al HER2. El trastorno puede ser un trastorno proliferativo de células que expresan el HER2, que incluye un tumor benigno o maligno caracterizado por la sobreexpresión del receptor de la ErbB2, por ejemplo, un cáncer, tal como cáncer de mama, cáncer epidermoide, cáncer de pulmón microcítico, cáncer de pulmón no microcítico, cáncer gastrointestinal, cáncer de páncreas, glioblastoma, cáncer cervicouterino, cáncer de ovario, cáncer de hígado, cáncer de vejiga, hepatoma, cáncer de colon, cáncer colorrectal, carcinoma endometrial, carcinoma de glándula salival, cáncer de riñón, cáncer de hígado, cáncer de próstata, cáncer de vulva, cáncer de tiroides, carcinoma hepático y diferentes tipos de cáncer de cabeza y cuello. Además, la divulgación contempla el uso de la composición anterior en lugar del anticuerpo anti-Her2 completo en el tratamiento de cánceres que sobreexpresan el HER2, como se describe en la Pat. de EE. UU. N.º 5.725.856.

En el presente documento se describe un método para tratar trastornos mediados por células que expresan el CD20. En el presente documento se describe un método para tratar una enfermedad humana mediada por células que expresan el CD20 con una composición de proteína de fusión de unión que deriva de un anticuerpo progenitor que se une al CD20 humano. Dichas composiciones tienen aplicaciones profilácticas y terapéuticas en un amplio espectro de trastornos mediados por células que expresan el CD20, que incluyen patologías apoyadas por la proliferación de células que expresan el CD20, tales como cánceres de células que expresan el CD20, de una manera similar a la aplicación de anticuerpos anti-CD20 completos en el tratamiento de dichas indicaciones de enfermedad conocidas en la técnica, indicaciones de tratamiento que incluyen linfomas de linfocitos B, como se describe en la patente de EE. UU. N.º 6.682.734.

En el presente documento se describe un método para tratar trastornos mediados por células que expresan el CD18. En el presente documento se describe un método para tratar una enfermedad humana mediada por células que expresan el CD18 con una composición de proteína de fusión de unión que deriva de un anticuerpo progenitor que se une al CD18 humano. Dichas composiciones tienen aplicaciones profilácticas y terapéuticas en un amplio espectro de trastornos mediados por células que expresan el CD18, que incluyen patologías apoyadas por la adhesión de leucocitos, de una manera similar a la aplicación de anticuerpos anti-CD18 completos en el tratamiento de dichas indicaciones de enfermedad conocidas en la técnica, indicaciones de tratamiento que incluyen infarto agudo de miocardio e ictus. En el presente documento se describe un método para tratar un trastorno en un paciente humano mediado por una célula que expresa el CD18, que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión en donde al menos un resto de direccionamiento de la composición comprende un sitio de unión al antígeno que se une al CD 18 humano. En otra realización, el trastorno mediado por células que expresan el CD18 es un trastorno inflamatorio, tal como un trastorno isquémico de reperfusión, que incluye infarto agudo de miocardio e ictus. Además, la divulgación contempla el uso de la proteína de fusión de unión anterior en lugar del anticuerpo anti-CD 18 completo en el tratamiento del ictus, como se describe en la Publicación PCT WO 97/26912.

En el presente documento se describe un método para tratar un trastorno mediado por la LFA-1 en un ser humano, que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión en donde al menos un resto de direccionamiento de la composición comprende un resto de direccionamiento que se une al CD 18 humano. Además, la divulgación contempla el uso de la proteína de fusión de unión anterior en lugar de anticuerpo anti-CD18 completo en el tratamiento de un trastorno mediado por la LFA-1, tal como psoriasis y rechazo del injerto, en un paciente humano, como se describe en la Patente de EE. UU. N.º 5.622.700.

En el presente documento se describe un método para tratar trastornos mediado por células que expresan el CD11a. En el presente documento se describe un método para tratar una enfermedad humana mediada por una célula que expresa el CD11a con una composición de proteína de fusión de unión que deriva de un anticuerpo progenitor que se une al CD11a humano. Dichas composiciones tienen aplicaciones profilácticas y terapéuticas en un amplio espectro de trastornos mediados por células que expresan el CD11a, que incluyen patologías apoyadas por la adhesión de leucocitos, de una manera similar a la aplicación de anticuerpos anti-CD11a completos en el tratamiento de dichas indicaciones de enfermedad conocidas en la técnica, indicaciones de tratamiento que incluyen psoriasis, asma, rechazo del injerto y esclerosis múltiple. En el presente documento se describe un método para tratar un trastorno mediado por la LFA-1 en un ser humano, que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión en donde al menos un resto de direccionamiento de la composición comprende un sitio de unión al antígeno que se une al CD11a humano. Además, la divulgación contempla el uso de la proteína de fusión de unión anterior en lugar de anticuerpo anti-CD11a completo en el tratamiento de un trastorno mediado por la LFA-1, tal como psoriasis y rechazo del injerto, en un paciente humano, como se describe en la Patente de EE. UU. n.º 5.622.700. En otro aspecto, la divulgación contempla el uso de las proteínas de fusión de unión anteriores en lugar del anticuerpo anti-CD11a completo en el tratamiento de trastornos mediados por la LFA-1 en un paciente humano, como se describe en la Patente de EE. UU. n.º 6.037.454.

En el presente documento se describe un método para tratar trastornos mediados por la IgE. En el presente documento se describe un método para tratar un trastorno mediado por la IgE en un paciente humano con una composición de

proteína de fusión de unión que deriva de un anticuerpo progenitor que se une a la IgE humana. Dichas composiciones tienen aplicaciones profilácticas y terapéuticas en un amplio espectro de trastornos mediados por la IgE, incluidas patologías caracterizadas por una sobreproducción y/o una hipersensibilidad a la inmunoglobulina IgE, de una manera similar a la aplicación de anticuerpos anti-IgE en el tratamiento de dichas indicaciones de enfermedad conocidas en la técnica, indicaciones de tratamiento que incluyen enfermedades alérgicas, tales como asma alérgica y rinitis alérgica. En el presente documento se describe un método para tratar un trastorno mediado por la IgE en un paciente humano que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión descrita, en donde al menos un resto de direccionamiento comprende un sitio de unión al antígeno que se une a la IgE humana. En otra realización, el trastorno mediado por la IgE es una enfermedad alérgica. En otra realización más, el trastorno mediado por la IgE es asma alérgica. En otra realización más, el trastorno mediado por la IgE es rinitis alérgica.

En el presente documento se describe un método para tratar un trastorno mediado por la IgE en un paciente humano que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión en donde al menos un resto de direccionamiento de la composición comprende un sitio de unión al antígeno que compite con el Fc épsilonRI humano por la unión a la IgE humana. En el presente documento se describe un método para tratar un trastorno mediado por la IgE en un paciente humano que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión en donde al menos un fragmento de anticuerpo de la proteína de fusión de unión comprende un sitio de unión al antígeno que se une a la IgE unida a membrana en la superficie de los linfocitos B humanos, pero no se une a la IgE soluble unida al receptor de Fc épsilonRI en la superficie de los basófilos humanos. Además, la divulgación contempla el uso de cualquiera de las proteínas de fusión de unión anteriores en lugar del anticuerpo anti-IgE humana completo en el tratamiento de un trastorno mediado por la IgE, tal como enfermedades alérgicas que incluyen asma alérgica y rinitis alérgica, en un paciente humano, como se describe en la Solicitud PCT n.º WO 99/01556. En otro aspecto, la divulgación contempla el uso de cualquiera de las proteínas de fusión de unión anteriores en lugar del anticuerpo anti-IgE humana completo en el tratamiento del asma alérgica en un paciente humano como se describe en el documento WO 97/04807.

En el presente documento se describe un método para tratar una enfermedad alérgica en un paciente humano que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión en donde al menos un resto de direccionamiento de la proteína de fusión de unión comprende un sitio de unión al antígeno que compite con el Fc épsilonRI humano para unirse a la IgE humana. En el presente documento se describe un método para tratar una enfermedad alérgica en un paciente humano que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión en donde al menos un resto de direccionamiento de la proteína de fusión de unión comprende un sitio de unión al antígeno que se une a la IgE unida a membrana en la superficie de los linfocitos B humanos, pero no se une a la IgE soluble unida al receptor de Fc épsilonRI en la superficie de los basófilos humanos.

En el presente documento se describe un método para tratar el asma alérgica en un paciente humano que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión en donde al menos un resto de direccionamiento de la proteína de fusión de unión comprende un sitio de unión al antígeno que compite con el Fc épsilonRI humano por la unión a la IgE humana. En el presente documento se describe un método para tratar el asma alérgica en un paciente humano que comprende administrar al paciente una cantidad terapéuticamente eficaz de cualquier proteína de fusión de unión descrita en esta sección (ii), en donde al menos un resto de direccionamiento de la proteína de fusión de unión comprende un sitio de unión al antígeno que se une a la IgE unida a membrana en la superficie de los linfocitos B humanos, pero no se une a la IgE soluble unida al receptor de Fc épsilonRI en la superficie de los basófilos humanos.

Trastornos mediados por el TNF-α

En el presente documento se describe un método para tratar una enfermedad mediada por el TNF-α con una proteína de fusión de unión que deriva de un anticuerpo progenitor que se une al TNF-α humano. Dichas proteínas de fusión de unión pueden tener aplicaciones profilácticas y terapéuticas en un amplio espectro de trastornos mediados por el TNF-α, incluidos trastornos inflamatorios y trastornos inmunitarios, de una manera similar a la aplicación de anticuerpos anti-TNF-α humano completos en el tratamiento de indicaciones de enfermedades tales como enfermedad de Crohn, enfermedad inflamatoria intestinal y artritis reumatoide.

En el presente documento se describe un método para tratar un trastorno inflamatorio en un paciente humano que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión en donde al menos un resto de direccionamiento de la proteína de fusión de unión comprende un sitio de unión al antígeno que se une al TNF-α humano. En otra realización, el trastorno inflamatorio es la enfermedad de Crohn. En otra realización más, el trastorno inflamatorio es enfermedad inflamatoria intestinal. En otra realización más, el trastorno inflamatorio es artritis reumatoide. El uso de anticuerpos que se unen al TNF-α humano en el tratamiento de afecciones inflamatorias se ha descrito, por ejemplo, en las Patentes de EE. UU. n.º 5.672.347, 5.656.272 y 5.698.195.

Trastornos mediados por el factor tisular

En el presente documento se describe un método para tratar una enfermedad mediada por el factor tisular con una proteína de fusión de unión que deriva de un anticuerpo progenitor que se une al factor tisular humano. Dichas proteínas de fusión de unión pueden tener aplicaciones profilácticas y terapéuticas en un amplio espectro de trastornos mediados por el factor tisular, incluyendo patologías apoyadas por la coagulación sanguínea y en el tratamiento de indicaciones patológicas tales como trombosis venosa profunda, trombosis arterial, aterosclerosis, estenosis vascular, enfermedades isquémicas del miocardio, incluyendo infarto agudo de miocardio, reoclusión después de una angioplastia o aterectomía o tratamiento trombolítico para el infarto agudo de miocardio, angina de pecho, enfermedades isquémicas cerebrales que incluyen ictus, tromboflebitis venosa y embolia pulmonar En el presente documento se describe un método para tratar una enfermedad o un trastorno mediado por el factor tisular (tal como lo anterior) en un paciente humano que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión en donde al menos un resto de direccionamiento de la proteína de fusión de unión comprende un sitio de unión al antígeno que se une al factor tisular humano.

En el presente documento se describe un método para inhibir la coagulación sanguínea en un paciente humano que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión en donde al menos un resto de direccionamiento de la proteína de fusión de unión comprende un sitio de unión al antígeno que se une al factor tisular humano, impidiendo la unión del factor VII de la coagulación.

Trastornos mediados por células que expresan el EGFR

En el presente documento se describe un método para tratar una enfermedad humana mediada por células que expresan el EGFR con una de la proteína de fusión de unión que deriva de un anticuerpo progenitor que se une al EGFR humano (también conocido como ErbB-1 o Her1). Dichas proteínas de fusión de unión pueden tener aplicaciones profilácticas y terapéuticas en un amplio espectro de trastornos mediados por células que expresan el EGFR, que incluyen patologías apoyadas por la proliferación de células que expresan el EGFR, tales como cánceres caracterizados por la sobreexpresión del EGFR, que incluyen cánceres de mama, de ovario, de cabeza y cuello, de cerebro, de vejiga, de páncreas y de pulmón.

En el presente documento se describe un método para tratar un trastorno de proliferación celular en un paciente humano caracterizado por la sobreexpresión del EGFR que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión en donde al menos un resto de direccionamiento de la proteína de fusión de unión comprende un sitio de unión al antígeno que se une al EGFR humano. El trastorno puede ser un tumor benigno o maligno caracterizado por la sobreexpresión del EGFR, por ejemplo, un cáncer, tal como cáncer de mama, cáncer epidermoide, cáncer de pulmón microcítico, cáncer de pulmón no microcítico, cáncer gastrointestinal, cáncer de páncreas, glioblastoma, cáncer cervicouterino, cáncer de ovario, cáncer de hígado, cáncer de vejiga, hepatoma, cáncer de colon, cáncer colorrectal, carcinoma endometrial, carcinoma de glándula salival, cáncer de riñón, cáncer de hígado, cáncer de próstata, cáncer de vulva, cáncer de tiroides, carcinoma hepático y diferentes tipos de cáncer de cabeza y cuello.

Trastornos mediados por células que expresan el CD3

En el presente documento se describe un método para tratar una enfermedad o un trastorno humano mediado por células que expresan el CD3 con una proteína de fusión de unión que deriva de un anticuerpo progenitor que se une al CD3 humano. Dichas proteínas de fusión de unión pueden tener aplicaciones profilácticas y terapéuticas en un amplio espectro de trastornos mediados por células que expresan el CD3, que incluyen afecciones asociadas con la proliferación o la activación de células que expresan el CD3, tales como trastornos inmunitarios mediados por linfocitos T y rechazo del injerto en receptores de un trasplante. El uso de anticuerpos anti-CD3 para tratar enfermedades y trastornos se ha descrito, por ejemplo, en la Patente de EE. UU. n.º 4.515.893. En otro aspecto, la invención contempla el uso de la proteína de fusión de unión anterior en lugar del anticuerpo anti-CD3 humano completo en el tratamiento del rechazo agudo de aloinjerto en receptores de trasplante renal, como se ha descrito para ORTHOCLONE OKT3 muromonab-CD3 en Physician's Desk Reference, 52ª edición (1998), págs. 1971-1974.

Trastornos mediados por células que expresan el TAC

En el presente documento se describe un método para tratar una enfermedad humana mediada por células que expresan la cadena α (TAC) del receptor de la interleucina-2 con una proteína de fusión de unión que deriva de un anticuerpo progenitor que se une al TAC humano. Dichas proteínas de fusión de unión pueden tener aplicaciones profilácticas y terapéuticas en un amplio espectro de trastornos mediados por células que expresan el TAC, incluyendo afecciones creadas por la proliferación o la activación de células que expresan el TAC y trastornos inmunitarios mediados por linfocitos T o linfocitos B, incluido el rechazo del injerto en receptores de un trasplante.

En el presente documento se describe un método para tratar un trastorno en un paciente humano mediado por células que expresan el TAC, que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión en donde al menos un resto de direccionamiento de la proteína de fusión de unión comprende un sitio de unión al antígeno que se une al TAC humano. En otra realización, el trastorno mediado por células que expresan el TAC se caracteriza por la activación o la proliferación de linfocitos T o linfocitos B, incluidos trastornos

inmunitarios tales como rechazo del injerto en receptores de un trasplante, enfermedad del injerto contra el hospedador (GHVD), rechazo del injerto en receptores de un trasplante, tal como rechazo agudo del injerto en receptores de un trasplante renal, y enfermedades autoinmunitarias tales como diabetes de tipo I, esclerosis múltiple, artritis reumatoide, lupus eritematoso sistémico y miastenia grave. El uso de anticuerpos para tratar trastornos mediados por la cadena α del receptor de la interleucina-2 con anticuerpos se ha descrito en la Patente de EE. UU. N.º 5.693.761.

(h) Composiciones de proteína de fusión de unión-fármaco y composiciones de XTEN-fármaco

La presente invención se refiere en parte a composiciones de proteínas de fusión de unión unidas covalentemente a un fármaco, lo que da como resultado un conjugado de fármaco y proteína de fusión de unión ("BFP-D"). En el presente documento se describen composiciones de XTEN unidas covalentemente a un fármaco, lo que da como resultado un conjugado de XTEN-fármaco ("XTEN-D"). En particular, la divulgación proporciona composiciones aisladas de BFP-D y XTEN-D útiles en el tratamiento de enfermedades, trastornos o afecciones. En una realización, el BFP-D comprende uno o más restos de direccionamiento de la proteína de fusión de unión dirigidos a un antígeno, ligando o receptor implicado en, asociado con, o que modula una enfermedad, un trastorno o una afección, mientras que uno o más XTEN de la proteína de fusión de unión se puede diseñar para que sirva como un portador con el que se conjuga el fármaco, confiriendo una semivida deseada o una propiedad farmacéutica mejorada a la proteína de fusión de unión, como se describe más completamente a continuación, y el fármaco unido covalentemente puede administrarse selectivamente a una célula, un tejido o un órgano para que ejerza un efecto farmacológico, citotóxico o citostático. Por lo tanto, el BFP-D comprende generalmente uno o más de los siguientes componentes: 1) XTEN; 2) resto de direccionamiento; 3) agente de conjugación; y 4) fármaco. El XTEN-D comprende generalmente uno o más de los siguientes componentes: 1) XTEN; 2) agente de conjugación; y 3) fármaco.

Anteriormente se han descrito realizaciones ilustrativas de restos de direccionamiento, XTEN y proteínas de fusión de restos de direccionamiento y XTEN. La divulgación proporciona XTEN que sirven además como una plataforma con la que pueden conjugarse los fármacos, de modo que sirven como un "portador", confiriendo ciertas propiedades farmacocinéticas, químicas y farmacéuticas deseables a las composiciones, entre otras propiedades descritas a continuación.

En algunas realizaciones, el componente XTEN se genomodifica para incorporar un número definido de restos de aminoácidos que contienen grupos reactivos que pueden usarse para la conjugación con fármacos y/o con agentes de conjugación. En una realización, el aminoácido reactivo es cisteína ("XTEN genomodificado con cisteína"). En otra realización, el aminoácido reactivo es lisina, que contiene un grupo ε-amino hidrófilo cargado positivamente ("XTEN genomodificado con lisina"). Como se utiliza en el presente documento, un "XTEN genomodificado con cisteína" significa una proteína XTEN, como se ha definido anteriormente, que comprende además aproximadamente 1 a aproximadamente 100 aminoácidos cisteína, o de 1 a aproximadamente 50 aminoácidos cisteína, o de 1 a aproximadamente 40 aminoácidos cisteína, o de 1 a aproximadamente 20 aminoácidos cisteína, o de 1 a aproximadamente 10 aminoácidos de cisteína, o de 1 a aproximadamente 5 aminoácidos de cisteína que están disponibles para la conjugación con moléculas de fármaco. Como se utiliza en el presente documento, un "XTEN genomodificado con lisina" significa una proteína XTEN, como se ha definido anteriormente, que comprende además aproximadamente 1 a aproximadamente 100 aminoácidos de lisina, o de 1 a aproximadamente 50 aminoácidos de lisina, o de 1 a aproximadamente 40 aminoácidos genomodificados con lisina, o de 1 a aproximadamente 20 aminoácidos genomodificados con lisina, o de 1 a aproximadamente 10 aminoácidos genomodificados con lisina, o de 1 a aproximadamente 5 aminoácidos genomodificados con lisina que están disponibles para la conjugación con moléculas de fármaco.

Generalmente, los grupos de tiol de la cisteína del XTEN son más reactivos, es decir, más nucleófilos, hacia reactivos de conjugación electrófila que los grupos amino o hidroxilo. Los restos de cisteína se han introducido en las proteínas mediante técnicas de ingeniería genética para formar enlaces covalentes con ligandos o para formar nuevos puentes de disulfuro intramoleculares (Better et al (1994) J. Biol. Chem. 13: 9644-9650; Bernhard et al (1994) Bioconjugate Chem. 5: 126-132; Greenwood et al (1994) Therapeutic Immunology 1: 247-255; Tu et al (1999) Proc. Natl. Acad. Sci USA 96: 4862-4867; Kanno et al (2000) J. of Biotechnology, 76: 207-214; Chmura et al (2001) Proc. Nat. Acad. Sci. USA 98 (15): 8480-8484; Pat. de EE. UU. N.º 6.248.564).

En el presente documento se describe una composición aislada que comprende un XTEN genomodificado con cisteína conjugado mediante un agente de conjugación con una o más moléculas de fármaco, en donde el fármaco se selecciona de la tabla 9. En el presente documento se describe una composición aislada que comprende un XTEN genomodificado con cisteína dirigido conjugado mediante un agente de conjugación con una o más moléculas de fármaco, en donde el fármaco se selecciona de la tabla 9 y el XTEN genomodificado con cisteína dirigido comprende uno o más restos de direccionamiento que presentan afinidad de unión por una o más dianas seleccionadas de la tabla 1 o la tabla 2. En el presente documento se describe una composición aislada que comprende un XTEN genomodificado con lisina conjugado mediante un agente de conjugación con una o más moléculas de fármaco, en donde el fármaco se selecciona de la tabla 9. En el presente documento se describe una composición aislada que comprende un XTEN genomodificado con lisina dirigido conjugado mediante un agente de conjugación con una o más moléculas de fármaco, en donde el fármaco se selecciona de la tabla 9 y el XTEN genomodificado con cisteína dirigido comprende uno o más restos de direccionamiento que presentan afinidad de unión por una o más dianas

seleccionadas de la tabla 1 o la tabla 2. En una realización de lo anterior, se conjugaría un único compuesto de fármaco con el XTEN. En otra realización, se puede conjugar más de un compuesto de fármaco con el XTEN mediante la aplicación selectiva de los métodos de conjugación y los reactivos, utilizando los métodos descritos en el presente documento o los conocidos en la técnica.

En algunos casos, las composiciones de la invención incluyen un XTEN genomodificado con cisteína, donde los nucleótidos que codifican uno o más aminoácidos de un XTEN se reemplazan por un aminoácido cisteína para crear el gen XTEN genomodificado con cisteína. En otros casos, los oligonucleótidos que codifican uno o más motivos de aproximadamente 9 a aproximadamente 14 aminoácidos que comprenden codones que codifican una o más cisteínas están unidos en marco con otros oligos que codifican motivos XTEN o un XTEN completo para crear el gen XTEN genomodificado con cisteína. En una realización de lo anterior, donde la una o más cisteínas se insertan en una secuencia del XTEN durante la creación del gen XTEN, los nucleótidos que codifican la cisteína se pueden unir a codones que codifican los aminoácidos usados en el XTEN para crear un motivo de cisteína-XTEN con la cisteína o cisteínas en una posición definida utilizando los métodos descritos en el presente documento (véanse el ejemplo 61 y las FIG. 40-41), o mediante técnicas de biología molecular convencionales, y los motivos se ensamblan posteriormente en el gen que codifica el XTEN genomodificado con cisteína completo. En dichos casos, cuando, por ejemplo, se añaden nucleótidos que codifican una única cisteína al ADN que codifica un motivo seleccionado de la tabla 3, el motivo resultante tendría 13 aminoácidos, mientras que la incorporación de dos cisteínas daría como resultado un motivo que tiene 14 aminoácidos, etc. En otros casos, se puede crear *de novo* un motivo de cisteína y tener una longitud y un número predefinidos de aminoácidos de cisteína uniendo los nucleótidos que codifican cisteína a nucleótidos que codifican uno o más restos de aminoácidos usados en el XTEN (por ejemplo, G, S, T, E, P, A) en una posición definida, y ensamblar posteriormente los motivos codificantes mediante la hibridación con otras secuencias de motivos codificantes del XTEN en el gen que codifica el XTEN completo, como se describe en el presente documento y se ilustra en las FIG. 7-8. En los casos en los que se utiliza un XTEN genomodificado con lisina para crear las composiciones de la invención, los enfoques descritos anteriormente se realizarían con codones que codifican lisina en lugar de cisteína. Por lo tanto, mediante lo anterior, se puede crear un nuevo motivo XTEN que podría comprender aproximadamente 9-14 restos de aminoácidos y tener uno o más aminoácidos reactivos; es decir, cisteína o lisina. Algunos ejemplos no limitativos de motivos adecuados para usar en un XTEN genomodificado que contiene una sola cisteína o lisina son:

GGSPAGSCTSP (SEQ ID NO: 174)

GASASCAPSTG (SEQ ID NO: 175)

GPEPTCPAPSG (SEQ ID NO: 176)

GGSPAGSKTSP (SEQ ID NO: 177)

GASASKAPSTG (SEQ ID NO: 178)

En dichos casos, cuando un gen que codifica un XTEN con uno o más motivos de cisteína y/o de lisina se construye a partir de módulos XTEN existentes, el gen se puede diseñar y construir uniendo polinucleótidos de "bloque de construcción" existentes que codifican XTEN tanto cortos como largos; por ejemplo, AE48, AE144, AE288, AE576, AM48, AE864, AM875, AE912, AG864, o los nucleótidos que codifican los 36-meros de los ejemplos 1-4, etc., que se puede fusionar en marco con los nucleótidos que codifican los motivos que contienen cisteína y/o lisina para construir un XTEN genomodificado en el que las cisteínas y/o las lisinas reactivas están colocadas en una o más ubicaciones seleccionadas de la secuencia en la cantidad deseada. En la tabla 8 se proporcionan ejemplos no limitativos de dicho XTEN genomodificado.

Tabla 8: XTEN genomodificado con cisteína y lisina

Nombre del XTEN*	Secuencia de aminoácidos	SEQ ID NO:
AE144-Island_Cys1 - AE576	GSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGS APGSEPATSGSETPGSEPATSGSETPGTSESATPESGPGSAPGTSTEPSEGSAP ESGPGSEPATSGSETPGTSTEPSEGSAPGGGSPAGSCTSPGGSPAGSPTSTEEGTSES ATPESGPGTSTEPSEGSAPGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTS ESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPG TSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAP GTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGS APGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATP ESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPS EGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTE PSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTS ESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGTSEPATSG SPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG	179
AE912-Island_Cys2 - AE144	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGSPAGSPTS TEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGTSTEPSE GSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESA TPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTE PSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTS TEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEG TSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAP GTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTST	180
	EEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSG SETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSP TSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPA TSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSP AGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEG SPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGP GSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGS APGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGGASASGAPSTGGGSEPAT SGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGSEP ATSGSETPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGS EPATSGSETPGTSTEPSEGSAPG	181
AE576-Island_Cys1 - AE288		181

Nombre del XTEN*	Secuencia de aminoácidos	SEQ ID NO:
	GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG STEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPS EGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPG TSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSP AGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAP GSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPES GPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSPAGSPT STEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGGSPA GSETSPGGTSESATPESGPGTSTEEGTSESATPESGPGSEPATSGSETPGTS ESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPG TSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGP GTSESATPESGPGTSESATPESGPGSEPATSGSETPGSPAGSPTSI EEGTSTEPSEGSAPGSTSTEPSEGSAPGSTEPSEGSAPGSETPGTSESATPESGPGTSTEPSSI GSAPG	182
AE48-Island_Cys3 AE864	- MAEPAGSPTSTEEGTPGSGTASSSPGGSTPSGATGSPGASPGTSSTGSPGGPEPTGP APSGGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTE PSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSP AGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPTSTEEG TSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGP GSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPES GPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSG GTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPG GSAPGTSTEPSEGSAPGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPES GPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSG TSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGP GTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPEIS GPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPES GPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPT STEEGTSESATPESGPGTSTEPSEGSAPGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPG BGSAPGTSTEPSEGSAPGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG	183
AE288-Island_Cys1 AE288-Island_Cys1 AE288	- - GTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPES GPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATP ESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEA PSEGSAPGTSTEPSEGSAPGSPATSGSETPGSPAGSPTSTEEGTSTE SPAGSCTSPGGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETP GTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSE TPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPE SGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSPATSGSETPGSPAGSP TSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEP SEGSAPGGGSPAGSCTSPGGTTSESATPESGPGSEPATSGSETPGTSESATPESGPGSE PATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPG SEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAP GTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSE TPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPE SGPGTSTEPSEGSAPG	184
AE48-Island_Cys2- AE576-Island_Cys2 AE144	- MAEPAGSPTSTEEGTPGSGTASSSPGGSTPSGATGSPGASPGTSSTGSPGGPGASASCA PSTGGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTE	

Nombre del XTEN*	Secuencia de aminoácidos	SEQ ID NO:
	PSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSP AGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEG TSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGP GSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPES GPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSE GSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEP SEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSE SATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSP AGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG GASASCAPSTGGGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTST EEGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGTSTEPSE GSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPG	
AM48-Island_Cys1 AM875	- MAEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGGGSPAGS CTSPGGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTPE SGSASPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPGSEPA TSGSETPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTST EPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGT SESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGP GTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGTASS SPGSSTPSGATGSPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTS TEEGSPAGSPTSTEEGTSTEPSEGSAPGASASCAPSTGGTSESATPESGPGSPAGSPT STEEGSPAGSPTSTEEGSTSSTAESPGPGSTSESPSGTAPGTSPSGESSTAPGTPGSGT ASSSPGSSTPSGATGSPGSSPSASTGTGPGSEPATSGSETPGTSESATPESGPGSEPA TSGSETPGSTSSTAESPGPGSTSSTAESPGPGTSPSGESSTAPGSEPATSGSETPGSEP ATSGSETPGTSTEPSEGSAPGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAPGTS TEPSEGSAPGTSTEPSEGSAPGSTSTEPSEGSAPGSTPSGATGSPGSSPSASTGTGPG ASPGTSSTGSPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSSTPSGATGS PGSSPSASTGTGPGASPGTSSTGSPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEG SAPG	185
AM48-Island_Cys2 AM1296	-	186

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	MAEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGGASASCA PSTGGGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTP ESGSASPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPGSEP ATSGSETPGTSESATPESGPGASGPTSTEEGTSTEPSEGSAPGTSESATPESGPGT STEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAP GTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPES GPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGTA SSSPGSSTPSGATGSPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGSP TSTEEGSPAGSPTSTEEGTSTEPSEGSAPGPEPTGPAPSGGSEPATSGSETPGTSESA TPESGPGSPAGSPTSTEEGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSES ATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSTSESPSGTAPGTSTPESGSASPGTS SGESSTAPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPGS ESATPESGPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSESATPESGPG TSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSPSGESSTAPGTSPSGESSTAP GTSPSGESSTAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGSPSASTGT GPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSS TGSPGASASGAPSTGGTSPSGESSTAPGSTSSTAESPGPGTSPSGESSTAPGTSESAT PESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPSASTGTGPGSSTPSGATGSPGASPG TSSTGSPGTSTEPSEGSAASPGTSPSGESSTAPGTSPSGESSTAPGTSESATPESGPGSEP ATSGSETPGTSTEPSEGSAPGSTSESPSGTAPGTSTPESGSASPGTSPSGESSAPGSF AGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPG SEPATSGSETPGSSTPSGATGSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGA PGTSPSGESSTAPGSTSSTAESPGPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTAS SSPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSTSTEPS	SEQ ID NO:
AM48-Island_Cys3 AM875	MAEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGGASASCA APSGGGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTP ESGSASPGSTSESPSGTAPGSTSESPSGTAPGSTSTPESGSASPGSEP ATSGSETPGTSESATPESGPGASGPTSTEEGTSTEPSEGSAPGTSESATPESGPGT STEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAP GTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPES GPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGTA SSSPGSSTPSGATGSPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGSP	187
	TSTEEGSPAGSPTSTEEGTSTEPSEGSAPGPEPTGPAPSGGSEPATSGSETPGTSESA TPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSP AGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPES GPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGTA SSSPGSSTPSGATGSPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGSP	188
AM48-Island_Cys3 AM875-Island_Cys3 AM48	- - -	

Nombre del XTEN*	Secuencia de aminoácidos	SEQ ID NO:
	MAEPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPTSTEEGTSESATPESGP

Nombre del XTEN	Secuencia de aminoácidos	SEQ ID NO:
	GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSPAGSPT	
	STEEGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPA TSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSP AGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAP GSPAGSPTSTEEGTSTEPSEGSAPGTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPG GPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPT STEEGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGGGSPA GSKTSPGGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTS ESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPA TSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGSPT GTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPST EEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSE GSAPG	
AE48-Island_Lys2-AE864	MAEPAGSPTSTEGTPGSGTASSSPGGSSPSGATGGSPGASPGTSSTGSPGGASASKA PSTGGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTE PSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSP AGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEG TSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGP GSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESG GPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSE GSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSP SHSSAPESPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSE SATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSP AGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGP GTSTEDSGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPES GPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATP ESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPS HSSAPGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSEPATSGSE TPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGTSTEPSE GSAPG	192
AE288-Island_Lys1-AE2188-Island_Lys1-AE288	GTSESATPESGPGTSESATPESGPGSEPATSGSETPGTSESATPESGTPGSEPATSGSE TPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATP ESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESA TPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTE PSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGASAPGGG SPAGSKTSPGGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETP GTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSET TPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPE SGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEP TSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGTSTEP SEGSAPGGGSPAGSKTSPGGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPA PATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGT SEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAP GTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGTSESATPE TPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPE SGPGTSTEPSEGSAPG	193

Nombre del XTEN*	Secuencia de aminoácidos	SEQ ID NO:
AE48-island_Lys1-AE576-island_Lys1-AE144	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGGGSPAGS KTSPGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTE PSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSP AGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEG TSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGP GSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPES GPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSE GSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEP SEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSE SATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSP AGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG CGSPAGSKTSPGGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTST EEGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGTSTEPSE GSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPG	194
AM48-island_Lys1-AM875	MAEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGGGSPAGS	195
	KTSPGGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTP ESGSASPGTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPGSEP ATSGSETPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGT STEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAP GTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPES GPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGTA SSSPGSSTPSGATGSPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGSP TSTEEGSPAGSPTSTEEGTSTEPSEGSAPGASASGAPSTGGTSESATPESGPGSPAGS PTSTEEGSPAGSPTSTEEGSTSSTAESPGPGSTSESPSGTAPGTSPSGESSTAPGTPGS GTASSSPGSSTPSGATGSPGSSPSASTGTGPGSEPATSGSETPGTSESATPESGPGSE PATSGSETPGSTSSTAESPGPGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAPG EPATSGSETPGTSTEPSEGSAPGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSSTPSGATGSPGSSPSASTGTG PGASPGTSSTGSPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSSTPSGAT GSPGSSPSASTGTGPGASPGTSSTGSPGTSESATPESGPGTSTEPSEGSAPGTSTEPS EGSAPG	196
AM48-island_Lys2-AM1296		

Nombre del XTEN*	Secuencia de aminoácidos	SEQ ID NO:
	MAEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGGASASKA PSTGGGTSTEPSEGSAPGSPATSGSETPGSPAGSPTSTEEGTSTSTAESPGPGTSTP ESGSASPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGSTSTPESGSASPGSEP ATSGSETPGTSESATPESGPGSPAAGSPTSTEEGTSTSTPESGSAPGTSESATPESGPGT STEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTSTPESGSAPGTSTEPSEGSAP GTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESS GPGTSTEPSEGSAPGSPTSTEEGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSS SSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGGASASKA TSTEEGSPAGSPTSTEEGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSS TGSPGASASGASPSTGGTSPASGESSTATGSTESTLAESPGPGTSPAGATGSPGGASPG PESGPGTSTEPSEGSAPGTSTEPSEGSAPGSSPKSASTGTGPGSSTPSGATGSPGGASPG TSSTGSPGTSTEPSEGSAPGSPTSTEEGSPSGESSTAPGTSPSGESSTAPGTSPSGPGSEP ATSGSETPGTSTEPSEGSAPGSTSESPSGTAPGSTSESPSGTAPGSTSTPESGSASPGSEP AGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGS SEPATSGSETPGSSTPSGATGSPGASPGTSSTGSPGSSTPSGATGSPGSTPSGSTA PGTSSPSGESSTAPGSTSSTAESPGPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTAS SSPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPG	
AM48-Island_Lys2-AM875	MAEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGGASASKA PSTGGGTSTEPSEGSAPGSPATSGSETPGSPAGSPTSTEEGTSTSTAESPGPGTSTP ESGSASPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGSTSTPESGSASPGSEP ATSGSETPGTSESATPESGPGSPAAGSPTSTEEGTSTSTPESGSAPGTSESATPESGPGT STEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAP GTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPES GPGTSTEPSEGSAPGSPTSTEEGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSS SSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGGASASKA TSTEEGSPAGSPTSTEEGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSS PTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGSPTSTEEGTSESATPESGPGTSPAGS GTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPATSGSETPGTSESATPESGPGSP PATSGSETPGTSTEPSEGSAPGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAPG EPATSGSETPGTSTEPSEGSAPGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSSTPSGATGSPESSPSASTGTG PGASPGTSSTGSPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSSTPSGAT GSPKSSPSASTGTLPGASPGTSSTGSPGTSESATPESGPGTSTPEPSEGSAPGTSTEPS EGSAPG	197
AM48-Island_Lys2-AM875-Island_Lys2-AM648	MAEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGGASASKA PSTGGGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSTSTAESPGPGTSTP ESGSASPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGSTSTPESGSASPGSEP ATSGSETPGTSESATPESGPGSPAAGSPTSTEEGTSTSTPESGSASPGSTSESPSGTAPG	198

Nombre del XTEN*	Secuencia de aminoácidos	SEQ ID NO:
	STEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAP GTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSESATPES GPGTSTEPSEGSAPGSPEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGTA SSSPGSSTPSGATGSHGTSTEPSEGSAHGTSTEPSEGSAPGSEPATSGSETPGSPAGSP TSTEEGSPAGSPTSTEEGTSTEPSEGSAPGASASGAPSTKGTSESATPESGPGSPAGS PTSTEEGSPAGSPTSTEEGSTSSTAESPGPGTSPSGESSTAPGSEPATSGSETPGTS GTASSSPGSSTPSGATGSPGSSPSASTGTGPGSEPATSGSETPGTSESATPESGPGSE PATSGSETPGTSTEPSEGSAPGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSSTPSGATGSPGSPSASTGTG PGTASPGTSSTPSGPGSEPATSGSETPGTSPSAATPESGPGSPAGSPTSTEEGSSTPSGAT GSPGSSPSASTGTGPGASPGTSSTPSGPGTSSESATPESGPGTSTEPSEGSAPGTSTEPS EGSAPGAGASSAPSTGGAAEPAGSPTSTEEGASPGTSSTPSGPGSSTPSGATGSPGS STPSGATGSPG	
AE288-Island_Cys1 AE288-Island_Lys1- AE288	- GTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPES GPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATP ESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESA TPESGPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTE PSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGGG SPAGSCTSPEGTSESATPESGPKSEPATSGSETPGTSESATPESGPGSEPATSGSETP GTSESATPESGPGTSTEPSEGSAHGSPAGSPTSTEEGTSESATPESGPGSEPATSGSE TPGTSESATPESGPGTSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPE SGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSPAGSP TSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEP SEGSAPGGGSPAGSKTSPGGGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSE PATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSE SEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAP GTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSE TPGSPAGSPTSTEEGTSTEPSEGSAHGTSTEPSEGSAPGSEPATSGSETPGTSESATPE SGPGTSTEPSEGSAPG	199
AE48-Island_Cys1 AE144-Island_Cys1 AE144-Island_Cys1 AE144-Island_Cys1 AE144	- - - - - MAEPAGSPTSTEEGTPKGSTASSSPGSSTPSGATGSPGASPGTSSTGSPGGGSPAGS CTSPGGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTE PSEGSAPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTS ESATPESGPGSEPATSGSETPGTSTEPSEGSAPGGGSPAGSCTSPGSEPATSGSETPG TSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETP GSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSE TPGTSTEPSEGSAPGGGSPAASKTSPGSEPATSGSETPGTSESATPESGPGSEPATS GSETPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGSEPA TSGSETPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGGG SPAGSCTSPGGSPAATSGSETPGTSESATPESGPGKSEPATSGSETPKGSPAGSPTSTEE GTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGS APGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPG	200
AE48-Island_Lys1- AE144-Island_Lys1- AE144-Island_Lys1- AE144-Island_Lys1- AE144		201

Nombre del XTEN*	Secuencia de aminoácidos	SEQ ID NO:
	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGGGSPAGS KTSPGGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTE PSEGSAPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTS ESATPESGPGSEPATSGSETPGTSTEPSEGSAPGGGSPAGSKTSPGSEPATSGSETPG TSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETP GSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSE TPGTSTEPSEGSAPGGGSPAGSKTSPGGSEPATSGSETPGTSESATPESGPGSEPATS GSETPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGSEPA TSGSETPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGGG SPAGSKTSPGGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEE GTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGS APGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPG	

En otra realización, donde un gen XTEN completo existente va a ser modificado con nucleótidos que codifican uno o más restos de cisteína o de lisina reactivos, se puede crear un oligonucleótido que codifica una cisteína o una lisina y que presenta una homología parcial y puede hibridar con una o más secuencias cortas del XTEN, dando como resultado un evento de recombinación y sustitución de un codón de cisteína o de lisina por un codón existente del gen XTEN (véase, por ejemplo, el ejemplo 61 para una descripción de los métodos generales). En una realización, la recombinación da como resultado una sustitución por la secuencia de aminoácidos GGSPAGSCTSP. Sin embargo, los oligonucleótidos pueden diseñarse para colocar la cisteína (o la lisina) en una ubicación diferente del motivo o para incluir una segunda cisteína (o lisina) en el motivo. Los oligonucleótidos que codifican la cisteína o la lisina se pueden diseñar para que hibriden con un segmento de secuencia dado en diferentes puntos a lo largo de la secuencia del XTEN conocida. Por lo tanto, la invención contempla que se pueden crear múltiples construcciones génicas de XTEN con cisteínas o lisinas insertadas en diferentes ubicaciones dentro de la secuencia del XTEN mediante la selección de sitios de restricción dentro de la secuencia del XTEN y el diseño de los oligonucleótidos apropiados para la ubicación dada y que codifican una cisteína o una lisina, incluyendo el uso de oligonucleótidos diseñados que dan como resultado múltiples inserciones en la misma secuencia del XTEN. Mediante el diseño y la selección de uno o más de dichos oligonucleótidos en consideración de la secuencia conocida del XTEN, y el uso apropiado de los métodos de la invención, se puede estimar el número potencial de restos de cisteína o de lisina reactivos sustituidos insertados en el XTEN completo y a continuación confirmarse mediante la secuenciación del gen XTEN.

Los métodos de diseño, selección y preparación de la invención permiten la creación de XTEN genomodificados que son reactivos con funcionalidad electrófila. Estos métodos permiten además la creación de composiciones conjugadas de XTEN-fármaco con moléculas de fármaco en sitios designados, diseñados y selectivos, como se ilustra esquemáticamente en las FIG. 40-41. Los fármacos pueden acoplarse específicamente en el sitio y de manera eficiente al XTEN genomodificado con cisteína y al XTEN genomodificado con cisteína dirigido de la invención con un reactivo con tiol. Por ejemplo, los restos de cisteína reactiva de un XTEN genomodificado con cisteína permiten conjugar específicamente un resto de fármaco con cada cisteína de la secuencia del XTEN mediante la conjugación

con un grupo reactivo tiol tal como maleimida o haloacetilo. La figura 45 ilustra un ejemplo específico de la conjugación de pacitaxel con una proteína de fusión de unión anti-Her2 mediante este enfoque. Generalmente, la reactividad nucleófila de la funcionalidad tiol de un resto de cisteína hacia un grupo maleimida es aproximadamente 1000 veces mayor en comparación con cualquier otra funcionalidad de aminoácido de una proteína, tal como el grupo amino de restos de lisina o el grupo amino aminoterminal. La funcionalidad específica de tiol de reactivos de yodoacetilo y maleimida puede reaccionar con grupos amino, pero se requiere un pH más alto (>9,0) y unos tiempos de reacción más largos (Garman, 1997, Non-Radiactive Labelling: A Practical Approach, Academic Press, Londres). Normalmente, las reacciones de conjugación con cisteína se realizan adecuadamente a un pH por debajo de aproximadamente 7, utilizando unas temperaturas de reacción en el intervalo de aproximadamente 5 a aproximadamente 40 °C, y preferentemente en el intervalo de 10 hasta 30 °C (véase el documento U.S. 6.048.720).

1. Fármacos

Los fármacos que van a incorporarse en las composiciones de BFP-D de la invención y en las composiciones de XTEN-fármaco de la divulgación tienen una o más actividades farmacológicas. Los fármacos pueden incluir un agente citotóxico o citostático (por ejemplo, epaclitaxel, paclitaxel, docetaxel, doxetaxel, irinotecán, pemetrexed, cloranbucilo o gemcitabina), un agente antiinflamatorio, un opiode (por ejemplo, morfina, oxicodona, hidromorfona), un analgésico, un antiinfeccioso o un fluoróforo tal como un colorante fluorescente tal como fluoresceína o rodamina, un quelante para un metal radioterápico o de obtención de imágenes, una etiqueta o etiqueta de detección de peptidilo o no peptidilo. En particular, se contemplan fármacos que tienen una alta incidencia de efectos secundarios o toxicidad, o aquellos para los que se desea localización en el sitio de la enfermedad o patología, para su incorporación en los conjugados de BFP-D o de XTEN-fármaco descritos en el presente documento.

Algunos fármacos ilustrativos para su incorporación en las composiciones de la invención se exponen en la Farmacopea de los Estados Unidos, la Farmacopea Homeopática oficial de los Estados Unidos o el vademécum nacional de especialidades farmacéuticas, en la Physician's Desk Reference (PDR) y en el Orange Book mantenido por la Administración de Medicamentos y Alimentos de los Estados Unidos (FDA). Los fármacos preferidos son aquellos que tienen el grupo funcional reactivo necesario o aquellos que se pueden derivatizar fácilmente para proporcionar el grupo funcional reactivo para la conjugación, y conservarán al menos una parte de la actividad farmacológica del fármaco no conjugado cuando se conjuga con el XTEN. En una realización, el fármaco para la conjugación con el presente XTEN o las proteínas de fusión divulgadas en el presente documento es un agente seleccionado de la tabla 9, o una sal, un ácido o un derivado farmacéuticamente aceptable del mismo.

Tabla 9: Fármacos para la conjugación con un XTEN genomodificado

Fármacos
Erlotinib; bortezomib; fulvestrant; dutent (SU11248), letrozol; mesilato de imatinib; PTK787/ZK 222584; oxaliplatino; 5-FU (5-fluorouracilo), leucovorina, rapamicina; lapatinib; lonafarnib; sorafenib; gefitinib; tiotepa; ciclofosfamida; busulfano; improsulfano; piposulfano; benzodopa; carbocuona; meturedopa; uredopa; altretamina; trietilenmelamina; trietilenfosforamida; trietilentiofosforamida; trimetilomelamina; bulatacina; bulatacinona; camptotecina; topotecán; briostatina; calistatina; CC-1065; adozelesina; caliqueamicina; auristatina; carzelesina; bizelesina; criptofitinas (particularmente criptofitina 1 y criptofitina 8); dolastatina; duocarmicina; eleuterobina; pancratistatina; sarcodictiina; espongistatina; clorambucilo; clomapazina; colofosfamida; estramustina; ifosfamida; mecloretamina; clorhidrato de óxido de mecloretamina; melfalán; novembichina; fenesterina; prednimustina; trofosfamida; mostaza de uracilo; carmustina; clorozotocina; fotemustina; lomustina; nimustina; ranimnustina; caliqueamicina; dinemicina; dinemicina A; clodronato; esperamicina; cromóforo de neocarzinostatina; aclacinomicinas, actinomicina; antramicina; azaserina; bleomicina; cactinomicina; carabicina; carminomicina; carzinofilina; cromomicina; dactinomicina; daunorubicina; detorubicina; 6-diazo-5-oxo-L-norleucina; doxorubicina (incluyendo morfolindoxorubicina, cianomorfolindoxorubicina, 2-pirrolin-doxorubicina y desoxidoxorubicina); epirubicina; esorubicina; idarubicina; marcelomicina; mitomicina C; ácido micofenólico; nogalamicina; olivomicina; peplomicina; potfiromicina; puromicina; quelamicina; rodorubicina; estreptonigrina; estreptozocina; tubercidina; ubenimex; zinostatina; zorubicina; metotrexato; 5-fluorouracilo (5-FU); fdenopterina; metotrexato; pteropterina; trimetrexato; fludarabina; 6-mercaptopurina; tiamiprina; tioguanina; ancitabina; azacitidina; 6-azauridina; carmofur; citarabina; didesoxuridina; doxifluridina; enocitabina; floxuridina; calusterona; propionato de dromostanolona; epitiostanol; mepitiostano; testolactona; aminoglutetimida; mitotano; trilostano; ácido frolínico; aceglatona; aldofosfamida glicósido; ácido aminolevulínico; eniluracilo; amsacrina; bestrabucilo; bisantreno; edatraxato; defofamina; demecolcina; diazicuona; elformitina; acetato de eliptinio; una epotilona; etoglúcido; nitrato de galio; hidroxiurea; lentinano; lonidainina; maitansina; ansamitocinas; mitoguazona; mitoxantrona; mopidanmol; nitraerina; pentostatina; fenamet; pirarubicina; losoxantrona; ácido podofilínico; 2-etilhidrazida; procarbazina; razoxano; rizoxina; ribavirina; zidovudina; aciclovir; gangciclovir; vidarabina; idoxuridina; trifluridina; foscamet; amantadina; rimantadina; saquinavir; indinavir; ritonavir; interferones alfa y otros interferones; AZT; sizofurano; espirogermanio; ácido tenuazónico; triazicuona; 2;2',2''-triclorotrietilamina; toxina T-2; verracurina A; roridina A; anguidina); uretano; vindesina; dacarbazina; manomustina; mitobronitol; mitolactol; pipobromano; gacitosina; arabinósido ("Ara-C"); ciclofosfamida; tiotepa; taxoides; epaclitaxel; paclitaxel; docetaxel; doxetaxel; irinotecán; pemetrexed cloranbucilo; gemcitabina; 6-tioguanina; mercaptopurina; metotrexato; cisplati; carboplatino; vinblastina; platino; etopósido (VP-16); ifosfamida; mitoxantrona; vincristina; vinorelbina; novantrona; tenipósido; edatrexato; daunomicina; aminopterina; xeloda; ibandronato; CPT-11; inhibidor de la topoisomerasa RFS 2000; difluorometilomitina (DMFO); ácido retinoico; capecitabina; lidocaína; bupivacaína; memantina; donepezilo; rivastigmina; galantamina; morfina; oxicodona; hidromorfona; oximorfona; metopón; apomorfina; normorfina; etorfina; buprenorfina; meperidina;
lopermida; anileridina; etoheptazina; piminidina; betaprodina; difenoxilato; fentanilo; sufentanilo; alfentanilo; remifentanilo; levorfanol; dextrometorfano; fenazocina; pentazocina; ciclazocina; metadona; isometadona; propoxifeno; naloxona; naltrexona; treprostinilo; N-metilnaloxona; 6-amino-14-hidroxi-17-alilnordesomorfina; naltrendol; N-metilnaltrexona; nalbufina; butorfanol; ciclazocina; pentazocina,; nalmefeno; naltrindol; nor-binaltorfimina; oxilorfano; 6-amino-6-desoxo-naloxona; pentazocina; levalorfanmetilnaltrexona; buprenorfina; ciclorfano; levalorfano; ciclosporina; ciclosporina A; micofenilato de mofetilo; sirólimus; tacrólimus; prednisona; azatioprina; metotrexato; ciclofosfamida; prednisona; ácido aminocaproico; cloroquina; hidroxicloroquina; dexametasona; clorambucilo; danazol; bromocriptina

2. Conjugación: agentes y métodos de conjugación

La conjugación entre el polipéptido (ya sea un XTEN o un compañero de fusión, tal como un resto de direccionamiento) y el compuesto del fármaco, opcionalmente a través de un agente de conjugación, se puede hacer de acuerdo con métodos conocidos en la técnica, por ejemplo, como describe Bodanszky en Peptide Synthesis, John Wiley, Nueva York, 1976 y en el documento WO 96/12505; Harris y Zalipsky, eds., Poly(ethylene glycol) Chemistry and Biological Applications, AZC, Washington; R. F. Taylor, (1991), "Protein immobilisation. Fundamental and applications", Marcel Dekker, N.Y.; S. S. Wong, (1992), "Chemistry of Protein Conjugation and Crosslinking", CRC Press, Boca Raton; G. T. Hermanson et al., (1993), "Immobilized Affinity Ligand Techniques", Academic Press, N.Y.; así como en las Pat. de EE. UU. N.º 5.977.163, 6.262.107, 6.441.025, 7.026.440, 7.329.721, 7.528.202, 7.579.444, 7.659.361 y 7.851.437; las Publicaciones de Solicitudes de Patente de EE. UU. n.º 2002001628, 20020077290, 20040157782 y 20050238649; y las Publicaciones PCT n.º WO 99/49901, WO 97/33552, WO 01/26693 y WO 01/70275, o mediante los métodos divulgados en el presente documento. Los métodos ilustrativos de las patentes o referencias anteriores, o los descritos en el presente documento, pueden aplicarse generalmente a las diferentes proteínas de fusión de unión divulgadas en el presente documento, lo que da como resultado las composiciones de proteína de fusión de unión-fármaco, o únicamente el XTEN, lo que da como resultado las composiciones de XTEN-D y/o de BFP-D descritas en el presente documento.

Normalmente, la unión de un fármaco a una proteína u otra superficie se consigue utilizando un derivado del fármaco activado, es decir, un fármaco que tiene al menos un extremo activado adecuado para la reacción con un centro nucleófilo (por ejemplo, lisina, cisteína y restos similares de proteínas). Las moléculas de fármaco que tienen grupos terminales activados adecuados para la reacción con los grupos amino de las proteínas incluyen aquellos con grupos funcionales tales como aldehídos (Harris, J. M., Herati, R. S., Polym Prepr. (Am. Chem. Soc., Div. Polym. Chem), 32 (1), 154-155 (1991), anhídridos mixtos, ésteres de N-hidroxisuccinimida, carboniladazolidos y clorocianuratos (Herman, S., et al., Macromol. Chem. Phys. 195, 203-209 (1994)). Aunque se ha demostrado que muchas proteínas conservan la actividad durante la modificación, en algunos casos, la unión del fármaco a través de los grupos amino de las proteínas puede ser indeseable, tal como cuando la derivatización de restos de lisina específicos inactiva el farmacóforo de la proteína (Suzuki, T., et al., Biochimica et Biophysica Acta 788, 248-255 (1984)). Además, dado que muchas proteínas no XTEN poseen varios grupos amino disponibles/accesibles, los conjugados de fármaco resultantes formados son normalmente mezclas de especies mono-, di-, triconjugadas, y así sucesivamente, que pueden ser difícil de caracterizar y separar y también requerir mucho tiempo. Un método para evitar estos problemas es emplear un reactivo selectivo del sitio que se dirija a grupos funcionales distintos de las aminas. Una diana particularmente atractiva es el grupo tiol, que en las proteínas está presente en el aminoácido cisteína. Las cisteínas son normalmente menos abundantes en las proteínas que las lisinas, reduciendo así la probabilidad de desactivación de las proteínas tras la conjugación con estos aminoácidos que contienen tiol. Además, la conjugación con sitios de cisteína a menudo se puede llevar a cabo de una manera bien definida, lo que conduce a la formación de conjugados de polímeros de una única especie.

El reactivo de tiol puede ser un reactivo multifuncional de conjugación, un conector de fármacos, un reactivo de captura, es decir, un resto de afinidad, un reactivo de etiqueta (por ejemplo, un reactivo conector de biotina), una etiqueta de detección (por ejemplo, un reactivo fluoróforo), un reactivo de inmovilización en fase sólida (por ejemplo, SEPHAROSE™, poliestireno o vidrio) o un intermedio de agente de conjugación de fármacos. Un ejemplo de un reactivo de tiol es N-etil maleimida (NEM). En una realización ilustrativa, la reacción de un tiol-XTEN con un reactivo conector de biotina proporciona un tiol-XTEN biotinilado mediante el cual se puede detectar y medir la presencia y la reactividad del resto de cisteína genomodificado. La reacción de un tiol-XTEN con un reactivo de conjugación multifuncional proporciona al tiol-XTEN un agente de conjugación funcionalizado que puede hacerse reaccionar adicionalmente con un reactivo de resto de fármaco u otra etiqueta. En una realización, la reacción de un tiol-XTEN o un tiol-XTEN dirigido con un intermedio conector de fármaco proporciona un conjugado de fármaco y tiol-XTEN o un conjugado de fármaco y tiol-XTEN dirigido, respectivamente.

Se pueden utilizar varias químicas de unión para la conjugación, incluyendo compuestos de conjugación homo- o heterobifuncionales disponibles comercialmente, de acuerdo con métodos conocidos y disponibles en la técnica, tales como los descritos, por ejemplo, en Hermanson, Greg T., Bioconjugate Techniques, Academic Press, Inc., 1995, y en Wong, Shan S., Chemistry. Los agentes de conjugación adecuados para usar en la preparación de las composiciones de la divulgación están disponibles comercialmente en empresas como Sigma-Aldrich, o Thermo Fisher Scientific Inc. (Pierce Protein Research Products). De utilidad particular son los componentes de conjugación que están disponibles en forma activada y se pueden utilizar directamente para la conjugación. Algunos ejemplos de agentes de conjugación útiles son imidoésteres, halógenos activos, maleimida, disulfuro de piridilo y NHS-ésteres. Los agentes de conjugación homobifuncionales tienen dos grupos reactivos idénticos y se usan a menudo en un procedimiento de conjugación químico de una sola etapa. Algunos ejemplos son BS3 (un análogo de DSS soluble en agua no escindible), BSOCOES (reversible con una base), DMA (dimetil adipimidato-2HCl), DMP (dimetil pimelimidato-2HCl), DMS (dimetil suberimidato-2HCl), DSG (análogo de 5 carbonos de DSS), DSP (reactivo de Lomant), DSS (no escindible), DST (escindible con agentes oxidantes), DTBP (3,3'-ditiobispropionimidato de dimetilo-2HCl), DTSSP, EGS, sulfo-EGS, THPP, TSAT, PMPI (N-[p-maleimidofenil]isocianato), DFDNB (1,5-difluoro-2,4-dinitrobenceno) es especialmente útil para la conjugación entre pequeñas distancias espaciales (Komblatt, J. A. y Lake, D. F. (1980). Cross-linking of cytochrome oxidase subunits with difluorodinitrobenzene. Can J. Biochem. 58, 219-224).

Los agentes de conjugación homobifuncionales reactivos con sulfhidrilo son agentes de conjugación de proteínas homobifuncionales que reaccionan con sulfhidrilos y a menudo se basan en maleimidas y ácido maleámico, que reaccionan con grupos -SH, formando enlaces de tioéter estables. La reacción se puede realizar a un pH de aproximadamente 6 a aproximadamente 10 y, por ejemplo, a un pH de aproximadamente 6,0, 6,5, 7,0, 7,5, 8,0, 8,5, 9,0, 9,5 o 10, con un pH óptimo de aproximadamente 6,5 a aproximadamente 8. En las realizaciones basadas en agentes de conjugación de maleimida, se puede aumentar aún más la estabilidad de los conjugados, el anillo de maleimida se puede abrir intencionalmente mediante una apertura forzada por hidrólisis durante la reacción utilizando un pH de 7 a aproximadamente 9 y una temperatura de reacción de aproximadamente 10 °C a 45 °C, o aproximadamente 18 °C a aproximadamente 30 °C, para proporcionar conjugados donde el componente de maleimida se convierte en su forma de anillo abierto de ácido succinámico, más estable.

Para fármacos con agentes de conjugación de maleimida conjugados, el método incluye hacer reaccionar la proteína con un agente activo que posee un nucleófilo en condiciones eficaces para acoplar el conector de fármaco a la proteína. Un ejemplo de dicho agente de conjugación y reacción incluyen BM[PEO]3; un espaciador de poliéter de 8 átomos que reduce el potencial de precipitación del conjugado en aplicaciones de conjugación de sulfhidrilo a sulfhidrilo. BM[PEO]4 es similar pero con un espaciador de 11 átomos. BMB es un agente de conjugación no escindible con un espaciador de cuatro carbonos. BMDB crea un enlace que se puede escindir con peryodato. BMH es un agente

de conjugación homobifuncional reactivo de sulfhidrilo ampliamente utilizado. BMOE tiene un agente de conjugación especialmente corto. DPDPB y DTME son agentes de conjugación escindibles. HVBS no tiene el potencial de hidrólisis de las maleimidas. TMEA es otra opción. Los agentes de conjugación heterobifuncionales tienen dos grupos reactivos diferentes. Algunos ejemplos son NHS-ésteres y aminas/hidrazinas a través de una activación por EDC, AEDP, ASBA (fotorreactivo, yodable), EDC (carbodiimida soluble en agua). Algunos agentes de conjugación bifuncionales reactivos con amina-sulfhidrilo son AMAS, APDP, BMPS, EMCA, EMCS, GMBS, KMUA, LC-SMCC, LC-SPDP, MBS, SBAP, SIA (extra corto), SIAB, SMCC, SMPB, SMPH, SMPT, SPDP, Sulfo-EMCS, Sulfo-GMBS, Sulfo-KMUS, Sulfo-LC-SMPT, Sulfo-LC-SPDP, Sulfo-MBS, Sulfo-SIAB, Sulfo-SMCC, Sulfo-SMPB. Algunos agentes de conjugación bifuncionales reactivos con sulfhidrilo-carbonilo, tales como KMUH (ácido N-[k-maleimidoundecanoico]hidrazida), BMPH (hidrazida del ácido N-[β-maleimidopropiónico]), EMCH (hidrazida del [ácido N-e-maleimidocaproico]), MPBH (clorhidrato de hidrazida del ácido 4-(4-N-maleimidofenil)butírico) y PDPH (hidrazida de 3-(2-piridilditio)propionilo). Algunos agentes de conjugación heterobifuncionales reactivos del grupo amino son ANB-NOS, MSA, NHS-ASA, SADP, SAED, SAND, SANPAH, SASD, SFAD, Sulfo-HSAB, Sulfo-NHS-LC-ASA, Sulfo-SADP, Sulfo-SANPAH, TFCS. Algunos agentes de conjugación reactivos con arginina son, por ejemplo, APG, que reacciona específicamente con argininas a pH 7-8.

Para fármacos con agentes de conjugación de maleimida conjugados, el método incluye hacer reaccionar la proteína con un agente activo que posee un nucleófilo en condiciones eficaces para acoplar el conector de fármaco a la proteína del receptor de adición de Michael mediante una reacción de adición de tipo Michael para formar un conjugado de proteína-fármaco unido a polímero-succinimida. Un "receptor de adición de Michael", como entenderá un experto en la técnica, es un resto capaz de reaccionar con un reactivo nucleófilo para experimentar una reacción de adición nucleófila característica de una reacción de adición de Michael. Después de que se produzca la adición nucleófila, el resto del receptor de adición de Michael se denomina "aducto de adición de Michael". Normalmente, una adición de Michael es la adición nucleófila de un carbanión u otro nucleófilo a un compuesto carbonilo alfa, beta insaturado, tal como un tioácido, que se crea tratando los restos de cisteína del XTEN genomodificado para obtener un tiol, a continuación un tioácido. Alternativamente, los grupos amino épsilon del XTEN genomodificado con lisina se pueden tiolar utilizando reactivos de tiolación, por ejemplo, SPDP o iminotiolano, para crear el receptor de adición de Michael. Dichos métodos son conocidos en la técnica (véase, por ejemplo, la Pat. de EE. UU. N.º 5.708.146).

Los grupos funcionales del fármaco que se va a conjugar pueden incluir grupos funcionales de ácido carboxílico y grupos funcionales de cloroformiato, que son sitios reactivos útiles porque pueden reaccionar con los grupos amino épsilon de un XTEN genomodificado con lisina o un agente de conjugación para formar un enlace de amida. También es útil como sitio reactivo un grupo funcional de carbonato, tal como, entre otros, carbonato de p-nitrofenilo, que puede reaccionar con un grupo amino para formar un enlace de carbamato. Cuando el fármaco va a conjugarse a través de un grupo hidroxilo con el XTEN, los grupos hidroxilo terminales de la molécula del fármaco deben modificarse y/o proporcionarse en una forma activada, es decir, con grupos funcionales reactivos (algunos ejemplos de los cuales incluyen grupos amino primarios, aminoxi, aldehído, hidrazida (HZ), tiol, tiolato, succinato (SUC), succinato de succinimidilo (SS), succinato de succinimidilo "éster activo", succinamida de succinimidilo (SSA), propionato de succinimidilo (SPA), butanoato de succinimidilo (SBA), carboximetilato de succinimidilo (SCM), carbonato de benzotriazol (BTC), N-hidroxisuccinimida (NHS), aldehído, nitrofenilcarbonato (NPC) y tresilato (TRES)). Otros grupos funcionales reactivos adecuados de moléculas de fármaco incluyen acetal, aldehídos que tienen una longitud de carbonos de 1 a 25 carbonos (por ejemplo, acetaldehído, propionaldehído y butiraldehído), hidrato de aldehído, alquenilo, acrilato, metacrilato, acrilamida, sulfona activa, amina, hidrazida, tiol, ácidos alcanoicos (por ejemplo, ácido carboxílico, carboximetilo, ácido propanoico y ácido butanoico), haluro de ácido, isocianato, isotiocianato, maleimida, vinilsulfona, ditiopiridina, vinilpiridina, yodoacetamida, epóxido, glioxal, diona, mesilato, tosilato y tresilato.

El fármaco también puede conjugarse utilizando un radical heterociclo de un sistema anular. Los grupos heterociclilo incluyen un sistema anular en el que uno o más átomos del anillo son un heteroátomo, por ejemplo nitrógeno, oxígeno y azufre. El radical heterociclo comprende 1 a 20 átomos de carbono y 1 a 3 heteroátomos seleccionados entre N, O, P y S. Un heterociclo puede ser un monociclo que tiene 3 a 7 miembros de anillo (2 a 6 átomos de carbono y 1 a 3 heteroátomos seleccionados entre N, O, P y S) o un biciclo que tiene 7 a 10 miembros de anillo (4 a 9 átomos de carbono y 1 a 3 heteroátomos seleccionados entre N, O, P y S), por ejemplo: un sistema biciclo [4,5], [5,5], [5,6] o [6,6]. Algunos heterociclos se describen en Paquette, Leo A.; "Principles of Modern Heterocyclic Chemistry" (W. A. Benjamin, New York, 1968); "The Chemistry of Heterocyclic Compounds, A series of Monographs" (John Wiley & Sons, Nueva York, 1950 hasta la actualidad), en particular los volúmenes 13, 14, 16, 19 y 28; y en J. Am. Chem. Soc. (1960) 82: 5566. Algunos ejemplos de heterociclos que se pueden encontrar en fármacos adecuados para la conjugación incluyen, a modo de ejemplo y sin limitación, piridilo, dihidropiridilo, tetrahidropiridilo (piperidilo), tiazolilo, tetrahidrotiofenilo, tetrahidrotiofenilo oxidado con azufre, pirimidinilo, furanilo, tienilo, pirrolilo, pirazolilo, imidazolilo, tetrazolilo, benzofuranilo, tianaftalenilo, indolilo, indolenilo, quinolinilo, isoquinolinilo, bencimidazolilo, piperidinilo, 4-piperidonilo, pirrolidinilo, 2-pirrolidonilo, pirrolinilo, tetrahidrofuranilo, bis-tetrahidrofuranilo, tetrahidropiranilo, bis-tetrahidropiranilo, tetrahidroquinolinilo, tetrahidroisoquinolinilo, decahidroquinolinilo, octahidroisoquinolinilo, azocinilo, triazinilo, 6H-1,2,5-tiadiazinilo, 2H,6H-1,5,2-ditiazinilo, tienilo, tiantrenilo, piranilo, isobenzofuranilo, cromenilo, xantenilo, fenoxatinilo, 2H-pirrolilo, isotiazolilo, isoxazolilo, pirazinilo, piridazinilo, indolizinilo, isoindolilo, 3H-indolilo, 1H-indazolilo, purinilo, 4H-quinolizinilo, ftalazinilo, naftiridinilo, quinoxalinilo, quinazolinilo, cinolinilo, pteridinilo, 4Ah-carbazolilo, carbazolilo, β-carbolinilo, fenantridinilo, acridinilo, pirimidinilo, fenantrolinilo, fenazinilo, fenotiazinilo, furazanilo, fenoxazinilo, isocromanilo, cromanilo, imidazolidinilo, imidazolinilo, pirazolidinilo, pirazolinilo, piperazinilo,

indolinilo, isoindolinilo, quinuclidinilo, morfolinilo, oxazolidinilo, benzotriazolilo, bencisoxazolilo, oxindolilo, benzoxazolinilo e isatinoilo.

En algunos casos, las moléculas del fármaco están unidas al XTEN genomodificado con lisina o cisteína mediante agentes de conjugación que tienen dos sitios reactivos para unirse al fármaco y al XTEN. Los grupos de conjugación preferidos son aquellos que son relativamente estables a la hidrólisis en la circulación, son biodegradables y no son tóxicos cuando se escinden del conjugado. Además, el uso de agentes de conjugación puede proporcionar el potencial de conjugados con una flexibilidad aún mayor entre el fármaco y el XTEN, o proporcionar suficiente espacio entre el fármaco y el XTEN de modo que el XTEN no interfiera en la unión entre el farmacóforo y su sitio de unión. En una realización, un agente de conjugación tiene un sitio reactivo que tiene un grupo electrófilo que es reactivo con un grupo nucleófilo presente en un XTEN. Los nucleófilos preferidos incluyen tiol, tiolato y amino. El heteroátomo del grupo nucleófilo de un XTEN genomodificado con lisina o cisteína es reactivo con un grupo electrófilo de un agente de conjugación, y forma un enlace covalente con la unidad de conjugación. Algunos grupos electrófilos útiles para agentes de conjugación incluyen, pero sin limitación, grupos de maleimida y haloacetamida, y proporcionan un sitio conveniente para la unión al XTEN.

En otra realización, un agente de conjugación tiene un sitio reactivo que tiene un grupo nucleófilo que es reactivo con un grupo electrófilo presente en un fármaco. Algunos grupos electrófilos útiles de un fármaco incluyen, pero sin limitación, grupos hidroxilo, tiol, aldehído y cetona carbonilo. El heteroátomo de un grupo nucleófilo de un agente de conjugación puede reaccionar con un grupo electrófilo de un fármaco y formar un enlace covalente. Algunos grupos nucleófilos útiles en un agente de conjugación incluyen, pero sin limitación, hidrazida, oxima, amino, hidrazina, tiosemicarbazona, carboxilato de hidrazina y arilhidrazida. El grupo electrófilo de un fármaco proporciona un sitio conveniente para la unión a un agente de conjugación.

Para la conjugación de fármacos con el grupo amino épsilon de lisina del XTEN genomodificado con lisina, se puede emplear el uso de un fármaco reactivo-N-hidroxilsuccinimida, o ésteres tales como fármaco-propionato de succinimidilo o fármaco-butanoato de succinimida u otros conjugados de fármaco-succinimida. Alternativamente, los restos de lisina pueden usarse para introducir grupos sulfhidrilo libres mediante reacción con iminotiolano. Alternativamente, las lisinas de la sustancia de direccionamiento pueden estar unidas a un reactivo heterobifuncional que tiene un grupo de hidrazida o aldehído libre disponible para la conjugación con un agente activo. Los ésteres reactivos pueden acoplarse a pH fisiológico, pero los derivados menos reactivos normalmente requieren un pH más alto. También se pueden emplear bajas temperaturas si se está utilizando una proteína lábil. En condiciones de baja temperatura, se puede utilizar un tiempo de reacción más largo para la reacción de conjugación.

La conjugación de grupos amino con restos de lisina se ve facilitada por la diferencia entre los valores de pKa del grupo α-amino del aminoácido aminoterminal (aproximadamente 7,6 a 8,0) y el grupo ε-amino de lisina (aproximadamente 10). La conjugación del grupo amino terminal a menudo emplea fármacos reactivos-aldehídos (tales como fármaco-propionaldehído o fármaco-butilaldehído), que son más selectivos para las aminas y, por lo tanto, tienen menos probabilidades de reaccionar con, por ejemplo, el grupo imidazol de histidina. Además, los restos de lisinilo amino se hacen reaccionar con anhídridos succínicos u otros anhídridos de ácidos carboxílicos, o con carbonato de N,N'-disuccinimidilo (DSC), N,N'-carbonil diimidazol (CDI) o cloroformiato de p-nitrofenilo, para producir el carbonato de succinimidilo, carbamato de imidazol o carbonato de p-nitrofenilo activados, respectivamente. La derivatización con estos agentes tiene el efecto de invertir la carga de los restos de lisinilo. La conjugación de un fármaco-aldehído con el grupo amino terminal normalmente tiene lugar en un tampón adecuado realizado a un pH que permite aprovechar las diferencias en el pKa entre los grupos ε-amino de los restos de lisina y el grupo α-amino del resto aminoterminal de la proteína; normalmente, el pH para el acoplamiento se encuentra en el intervalo de aproximadamente pH 7 hasta aproximadamente 8. Se han descrito métodos útiles para la conjugación del grupo amino épsilon de lisina en la Pat. de EE. UU. N.º 4.904.584; y en la Pat. de EE. UU. N.º 6.048.720.

La persona con la experiencia habitual en la técnica será consciente de que el método de activación y/o la química de conjugación que va a usarse depende de los grupos reactivos del polipéptido XTEN, así como de los grupos funcionales del resto de fármaco (por ejemplo, siendo amino, hidroxilo, carboxilo, aldehído, sulfhidrilo, etc.), el grupo funcional del reactivo fármaco-agente de reticulación, o el grupo funcional del reactivo XTEN-agente de reticulación. La conjugación del fármaco puede dirigirse hacia la conjugación con todos los grupos de unión disponibles en el polipéptido XTEN genomodificado, tales como los grupos de unión genomodificados específicos en los restos de cisteína o los restos de lisina incorporados. Con el fin de controlar los reactivos de manera que la conjugación se dirija al sitio reactivo apropiado, la invención contempla el uso de grupos protectores. Un "grupo protector" es un resto que impide o bloquea la reacción de un grupo funcional químicamente reactivo particular en una molécula en ciertas condiciones de reacción. El grupo protector variará en función del tipo de grupo químicamente reactivo que se esté protegiendo, así como de las condiciones de reacción que se vayan a emplear, así como de la presencia de grupos reactivos adicionales en la molécula. Algunos ejemplos no limitativos de grupos funcionales que pueden protegerse incluyen grupos ácido carboxílico, grupos hidroxilo, grupos amino, grupos hidroxilo, grupos tiol y grupos carbonilo. Algunos grupos protectores representativos para ácidos carboxílicos e hidroxilos incluyen ésteres (tales como un éster de p-metoxibencilo), amidas e hidrazidas; para grupos amino, carbamatos (tales como terc-butoxicarbonilo) y amidas; para grupos hidroxilo, éteres y ésteres; para grupos tiol, tioéteres y tioésteres; para grupos carbonilo, acetales y cetales; y similares. Dichos grupos protectores son bien conocidos por los expertos en la técnica y se describen, por

ejemplo, en T. W. Greene y G. M. Wuts, Protecting Groups in Organic Synthesis, tercera edición, Wiley, Nueva York, 1999, y las referencias citadas en los mismos.

La conjugación se puede conseguir en una etapa o de manera escalonada (por ejemplo, como se describe en el documento WO 99/55377), tal como mediante la adición de un agente de conjugación intermedio de reacción, utilizando los agentes de conjugación divulgados en el presente documento o los conocidos en la técnica por ser útiles para la conjugación con restos de cisteína o de lisina de polipéptidos que se unirán a grupos funcionales reactivos en moléculas de fármaco.

En algunos casos, el método para conjugar un agente de conjugación con un XTEN genomodificado con cisteína puede prever que el XTEN sea pretratado con un agente reductor, tal como ditiotreitol (DTT) para reducir cualquier resto de disulfuro de cisteína para formar grupos tiol muy nucleófilos (-CH_2 SH) en la cisteína. El agente reductor se elimina posteriormente mediante cualquier método convencional, tal como desalación. Por lo tanto, el XTEN parcialmente reducido reacciona con compuestos de fármaco-conector o con reactivos de conjugación con grupos funcionales electrófilos tales como maleimida o α-halo carbonilo, de acuerdo con, por ejemplo, el método de conjugación de Klussman, et al (2004), Bioconjugate Chemistry 15 (4): 765-773. La conjugación de un agente de conjugación o un fármaco con un resto de cisteína tiene lugar normalmente en un tampón adecuado a un pH de 6-9 a unas temperaturas que varían de 4 °C a 25 °C durante períodos de hasta aproximadamente 16 horas. Alternativamente, los restos de cisteína se pueden derivatizar con un agente derivatizante orgánico. Los agentes y métodos de derivatización adecuados son bien conocidos en la técnica. Por ejemplo, lo más habitualmente, los restos de cisteinilo se hacen reaccionar con α-haloacetatos (y las aminas correspondientes), tales como ácido cloroacético o cloroacetamida, para dar carboximetilo o derivados de carboximetilo. Los restos de cisteinilo también se derivatizan mediante reacción con bromotrifluoroacetona, ácido α-bromo-β-(4-imidozoil)propiónico, fosfato de cloroacetilo, N-alquilmaleimidas, disulfuro de 3-nitro-2-piridilo, metil disulfuro de 2-piridilo, p-cloromercuribenzoato, 2-cloromercuri-4-nitrofenol o cloro-7-nitrobenzo-2-oxa-1,3-diazol.

En una realización, el XTEN se puede disolver en borato de sodio 500 mM y cloruro de sodio 500 mM a pH 8,0, y a continuación se trata con un exceso de ditiotreitol (DTT) 100 mM. Después de la incubación a 37 °C durante aproximadamente 30 minutos, el tampón se intercambia por elución sobre resina Sephadex G25 y se eluye con PBS con DTPA 1 mM. El valor de tiol/XTEN se comprueba determinando la concentración del XTEN reducido a partir de la absorbancia a 280 nm de la solución, y la concentración de tiol mediante reacción con DTNB (Aldrich, Milwaukee, Wis.) y la determinación de la absorbancia a 412 nm. El XTEN reducido disuelto en PBS se enfría en hielo. El fármaco-agente de conjugación, por ejemplo, MC-val-cit-PAB-MMAE en DMSO, disuelto en acetonitrilo y agua a una concentración conocida, se añade al XTEN reducido enfriado en PBS. Después de aproximadamente una hora, se añade un exceso de maleimida para inactivar la reacción y proteger cualquier grupo tiol de anticuerpo que no haya no reaccionado. La mezcla de reacción se concentra mediante ultrafiltración centrífuga y el XTEN-MC-vc-PAB-MMAE se purifica y desala mediante elución a través de resina G25 en PBS, se filtra a través de filtros de 0,2 µm en condiciones estériles y se mantiene en unas condiciones de almacenamiento adecuadas.

Dicho enfoque se puede utilizar para conjugar otros agentes reactivos de tiol con la cisteína del XTEN, en los que el grupo reactivo es, por ejemplo, una maleimida, una yodoacetamida, un disulfuro de piridilo u otro compañero de conjugación reactivo de tiol unido a un compañero farmacológico (Haugland, 2003, Molecular Probes Handbook of Fluorescent Probes and Research Chemicals, Molecular Probes, Inc.; Brinkley, 1992, Bioconjugate Chem. 3: 2; Garman, 1997, Non-Radioactive Labelling: A Practical Approach, Academic Press, Londres; Means (1990) Bioconjugate Chem. 1: 2; Hermanson, G. in Bioconjugate Techniques (1996) Academic Press, San Diego, págs. 40-55, 643-671). Las maleimidas, en particular, son útiles en la reticulación debido a su susceptibilidad a las adiciones a través del doble enlace, ya sea mediante adiciones de Michael o a través de reacciones de Diels-Alder. Las bismaleimidas son una clase de compuestos de maleimida con dos grupos maleimida conectados a través de una unidad molecular y se pueden utilizar como reactivos de conjugación.

En algunos casos, la conjugación se realiza en condiciones que aspiran a hacer reaccionar tantos grupos de unión al XTEN disponibles como sea posible con moléculas de fármaco o de conector de fármaco. Esto se consigue mediante un exceso molar adecuado del fármaco en relación con el polipéptido. Las relaciones molares clásicas entre moléculas de fármaco activado o de conector de fármaco y polipéptido son de hasta aproximadamente 1000-1, tal como de hasta aproximadamente 200-1 o de hasta aproximadamente 100-1. En algunos casos, la relación puede ser algo menor, sin embargo, tal como de hasta aproximadamente 50-1, 10-1 o 5-1. También pueden usarse relaciones equimolares.

En algunos casos, las composiciones conjugadas que contienen fármaco de la divulgación conservan al menos una parte de la actividad farmacológica en comparación con el correspondiente fármaco no conjugado. En una realización, el conjugado de fármaco conserva al menos aproximadamente un 1 %, o al menos aproximadamente un 5 %, o al menos aproximadamente un 10 %, o al menos aproximadamente un 20 %, o al menos aproximadamente un 30 %, o al menos aproximadamente un 40 %, o al menos aproximadamente un 50 %, o al menos aproximadamente un 60 %, o al menos aproximadamente un 70 %, o al menos aproximadamente un 80 %, o al menos aproximadamente un 90 %, o al menos aproximadamente un 95 % de la actividad farmacológica del fármaco no conjugado.

En otros casos, el fármaco puede derivatizarse a través de un grupo funcional reactivo que es importante para la actividad biológica del fármaco, inhibiendo o reduciendo así la actividad farmacológica del fármaco para convertir de este modo el fármaco en un conjugado derivado de peptidilo farmacológicamente inactivo o relativamente inactivo. En una realización, el agente de conjugación de profármaco contiene un resto peptídico específicamente adaptado para hacer de un conjugado de fármaco de la presente invención un sustrato selectivo susceptible a la escisión enzimática por parte de una o más proteasas, por ejemplo, preferentemente proteasas lisosómicas, tales como catepsina B, C o D. La reacción de escisión enzimática eliminará el agente de conjugación de profármaco del resto de fármaco y afectará a la liberación del fármaco en su forma farmacológicamente activa. Algunos enlaces hidrolíticamente degradables representativos en un conjugado de agente de conjugación-fármaco incluyen éster de carboxilato, éster de carbonato, éster de fosfato, anhídrido, acetal, cetal, aciloxialquil éter, imina, ortoéster y oligonucleótidos. Los ésteres tales como ésteres de carboxilato y de carbonato son enlaces particularmente preferidos. El enlace y la química de enlace particular empleados dependerán del agente activo particular, de la presencia de grupos funcionales diana y adicionales dentro del agente activo y similares; consideraciones que están dentro del conocimiento de un experto en la materia. Mediante dicho enfoque, las composiciones de conjugado de XTEN-fármaco y de proteína de fusión de unión-fármaco inventivas administradas a un sujeto pueden presentar una toxicidad o una frecuencia de efectos secundarios reducidas en comparación con el correspondiente fármaco libre administrado a un sujeto. Por lo tanto, la reducida toxicidad de las composiciones conjugadas de XTEN-fármaco y de proteína de fusión de unión-fármaco divulgadas en el presente documento puede permitir la administración de mayores cantidades de fármaco, sobre una base molar, en comparación con el fármaco no conjugado.

La invención contempla que el XTEN genomodificado, que incorpora restos de cisteína o de lisina, se pueda conjugar con cualquier resto de fármaco con un grupo funcional reactivo que se pueda unir covalentemente al XTEN a través de un grupo tiol de cisteína reactivo o épsilon amino, respectivamente, ya sea directamente o utilizando un agente de conjugación, como se ha descrito anteriormente. En otra realización, la invención proporciona composiciones de BFP-D en las que el fármaco puede estar conjugado con el resto de direccionamiento del componente de la proteína de unión mediante conjugación con restos de cisteína o de lisina existentes o incorporados, así como con grupos amino aminoterminales o grupos carboxilo carboxiterminales que puedan estar presentes.

(i) Liberación del fármaco

La invención proporciona composiciones de BFP-D en las que el fármaco puede ser liberado de la composición mediante mecanismos específicos o no específicos. En algunos casos, los fármacos pueden ser liberados mediante la degradación proteolítica de aquellas moléculas captadas por las células. En una realización, la porción XTEN del BFP-D o del XTEN-fármaco es rápidamente degradada por las proteasas intracelulares, liberando el fármaco del portador XTEN. En otros casos, el fármaco puede ser liberado por la degradación de un agente de conjugación seleccionado para su inclusión en las composiciones de BFP-D y de XTEN-fármaco tomando como base su susceptibilidad a la degradación. Por ejemplo, en la técnica se sabe que el uso de conectores escindibles con ácidos suaves puede promover la liberación de fármacos tomando como base la observación de que el pH dentro de los tumores era a menudo inferior al pH fisiológico normal. En un ejemplo no limitativo, la liberación del componente de fármaco conjugado podría mejorarse incorporando una hidrazona como una unidad escindible y uniendo un fármaco como doxorubicina al componente proteico (ya sea XTEN o el resto de direccionamiento) a través de un grupo tioéter, tal como se describe en Willner et al., Pat de EE. UU. N.º 5.708.146; y en Trail et al. Cure of xenografted human carcinomas by BR96-doxorubicin immunoconjugates. Science 261: 212-215 (1993). En otros casos, pueden incorporarse ciertos enlaces éster en el conector entre una proteína y el fármaco que son lábiles; algunos mediante enzimas. (Gillimard y Saragovi, Cancer Res. 61: 694-699 (2001)). Otros ejemplos de enlaces enzimáticamente susceptibles incluyen enlaces que contienen uretano o carbonato. Además, algunos enlaces hidrolíticamente inestables incluyen éster de carboxilato, éster de fosfato, anhídridos, acetales, cetales, aciloxialquil éter, iminas, ortoésteres, péptidos y oligonucleótidos. Tomando como base estas características, se cree que el ligando de fármaco se liberará rápidamente después de la internalización celular de los conjugados de BFP-D o de XTEN-fármaco.

(j) Configuraciones de las composiciones de BFP-D y XTEN-fármaco

La invención proporciona conjugados de proteína de fusión de unión-fármaco en diferentes configuraciones.

En el presente documento se describe una composición de conjugado de proteína de fusión de unión-fármaco de fórmula V:

$$[(D\text{-}CL)_{z1}\text{-}XTEN]_x\text{-}TM\text{-}[XTEN\text{-}(CL\text{-}D)_{z2}]_y \qquad V$$

en donde independientemente en cada aparición: X es 0 o 1; y es 0 o 1; XTEN es un polipéptido recombinante expandido genotecnológico con cisteína o lisina como se ha descrito anteriormente; TM es un resto de direccionamiento con afinidad de unión por un ligando diana seleccionado de la tabla 1 o la tabla 2 (que puede comprender más de un dominio de unión unido por conectores); CL es un agente de conjugación como se define en el presente documento; D es un resto de fármaco seleccionado de la tabla 9 o una sal, un ácido o un derivado farmacéuticamente aceptable del mismo; y cada uno de z1 y z2 es independientemente un número entero de 1 a 100. El número de restos de fármaco que pueden conjugarse a través de un agente de conjugación reactivo con una

molécula de XTEN genomodificada está limitado por el número de restos reactivos que se incorporan en el XTEN. Algunas composiciones de conjugado de proteína de fusión de unión-fármaco de fórmula V ilustrativas pueden comprender XTEN que tienen de 1 a aproximadamente 100 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 50 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 40 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 20 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 10 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 5 aminoácidos genomodificados con cisteína o lisina que están disponibles para su conjugación con moléculas de fármacos. En algunos casos, el conjugado de proteína de fusión de unión-fármaco conserva al menos aproximadamente un 1 %, o al menos aproximadamente un 5 %, o al menos aproximadamente un 10 %, o al menos aproximadamente un 20 %, o al menos aproximadamente un 30 %, o al menos aproximadamente un 40 %, o al menos aproximadamente un 50 %, o al menos aproximadamente un 60 %, o al menos aproximadamente un 70 %, o al menos aproximadamente un 80 %, o al menos aproximadamente un 90 %, o al menos aproximadamente un 95 % de la actividad farmacológica del fármaco no conjugado.

En el presente documento se describe una composición de conjugado de proteína de fusión de unión-fármaco de fórmula VI:

$$[(D\text{-}CL)_{z1}\text{-}XTEN]_x\text{-}TM1\text{-}L\text{-}TM2\text{-}[XTEN\text{-}(CL\text{-}D)_{z2}]_y \qquad VI$$

en donde independientemente en cada aparición: X es 0 o 1, e y es 0 o 1; XTEN es un polipéptido recombinante expandido genotecnológico de cisteína o lisina como se ha descrito; TM1 es un resto de direccionamiento con afinidad de unión por un ligando diana seleccionado de la tabla 1 o la tabla 2 (que puede comprender más de un dominio de unión unido por conectores); TM2 es un resto de direccionamiento con afinidad de unión por un ligando diana seleccionado de la tabla 1 o la tabla 2 (que puede comprender más de un dominio de unión unido por conectores) que puede ser idéntico o puede ser diferente a TM1; y L es una secuencia conectora que tiene entre 1 y aproximadamente 300 restos de aminoácidos, en donde la secuencia conectora está unida covalentemente al extremo carboxílico de TM1 y al extremo amínico de TM2; D es un resto de fármaco seleccionado de la tabla 9 o una sal, un ácido o un derivado farmacéuticamente aceptable del mismo; CL es un agente de conjugación como se define en el presente documento; y cada uno de z1 y z2 es independientemente un número entero de 0 a 100. Algunas composiciones de conjugado de proteína de fusión de unión-fármaco de fórmula VI ilustrativas pueden comprender XTEN que tienen de 1 a aproximadamente 100 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 50 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 40 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 20 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 10 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 5 aminoácidos genomodificados con cisteína o lisina que están disponibles para su conjugación con moléculas de fármacos. Además, la invención contempla composiciones adicionales que comprenden múltiples restos de direccionamiento y XTEN en diferentes permutaciones de configuraciones. En algunos casos, el conjugado de proteína de fusión de unión-fármaco conserva al menos aproximadamente un 1 %, o al menos aproximadamente un 5 %, o al menos aproximadamente un 10 %, o al menos aproximadamente un 20 %, o al menos aproximadamente un 30 %, o al menos aproximadamente un 40 %, o al menos aproximadamente un 50 %, o al menos aproximadamente un 60 %, o al menos aproximadamente un 70 %, o al menos aproximadamente un 80 %, o al menos aproximadamente un 90 %, o al menos aproximadamente un 95 % de la actividad farmacológica del fármaco no conjugado.

En el presente documento se describe una composición farmacológicamente activa de conjugado de XTEN-fármaco de fórmula VII:

$$(CL\text{-}D)_z\text{-}XTEN \qquad VII$$

en donde independientemente en cada aparición: XTEN es un polipéptido recombinante expandido genomodificado con cisteína o lisina como se ha descrito anteriormente; D es un resto de fármaco seleccionado de la tabla 9 o una sal, un ácido o un derivado farmacéuticamente aceptable del mismo; CL es un agente de conjugación como se define en el presente documento; y z es un número entero de 1 a 100. Algunas composiciones de conjugados de XTEN-fármaco de fórmula VII ilustrativas pueden comprender XTEN que tienen de 1 a aproximadamente 100 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 50 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 40 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 20 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 10 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 5 aminoácidos genomodificados con cisteína o lisina que están disponibles para su conjugación con moléculas de fármacos. En algunos casos, el conjugado de XTEN-fármaco conserva al menos aproximadamente un 1 %, o al menos aproximadamente un 5 %, o al menos aproximadamente un 10 %, o al menos aproximadamente un 20 %, o al menos aproximadamente un 30 %, o al menos aproximadamente un 40 %, o al menos aproximadamente un 50 %, o al menos aproximadamente un 60 %, o al menos aproximadamente un 70 %, o al menos aproximadamente un 80 %, o al menos aproximadamente un 90 %, o al menos aproximadamente un 95 % de la actividad farmacológica del fármaco no conjugado.

En el presente documento se describe una composición farmacológicamente activa de conjugado de XTEN-fármaco de fórmula VIII:

$XTEN\text{-}(CL\text{-}D)_z$ VIII

en donde independientemente en cada aparición: XTEN es un polipéptido recombinante expandido genomodificado con cisteína o lisina como se ha descrito anteriormente; D es un resto de fármaco seleccionado de la tabla 9 o una sal, un ácido o un derivado farmacéuticamente aceptable del mismo; CL es un agente de conjugación como se define en el presente documento; y z es un número entero de 1 a 100. Algunas composiciones de conjugados de XTEN-fármaco de fórmula VII ilustrativas pueden comprender XTEN que tienen de 1 a aproximadamente 100 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 50 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 40 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 20 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 10 aminoácidos genomodificados con cisteína o lisina, o de 1 a aproximadamente 5 aminoácidos genomodificados con cisteína o lisina que están disponibles para su conjugación con moléculas de fármacos. En algunos casos, el conjugado de XTEN-fármaco conserva al menos aproximadamente un 1 %, o al menos aproximadamente un 5 %, o al menos aproximadamente un 10 %, o al menos aproximadamente un 20 %, o al menos aproximadamente un 30 %, o al menos aproximadamente un 40 %, o al menos aproximadamente un 50 %, o al menos aproximadamente un 60 %, o al menos aproximadamente un 70 %, o al menos aproximadamente un 80 %, o al menos aproximadamente un 90 %, o al menos aproximadamente un 95 % de la actividad farmacológica del fármaco no conjugado.

La divulgación contempla que los BFP-D engloben composiciones en las que el componente o componentes del resto de direccionamiento de las presentes composiciones se pueda dirigir a cualquiera de las dianas específicas descritas en el presente documento, incluyendo una diana seleccionada de la tabla 1 o la tabla 2; el fármaco conjugado con la composición puede ser cualquiera de los fármacos de la tabla 9 o una sal, un ácido o un derivado farmacéuticamente aceptable del mismo, y el o los componentes de XTEN pueden ser XTEN genomodificados con cisteína o lisina derivados de, o que presentan una identidad de secuencia sustancial con, cualquiera de los XTEN de la tabla 4 o un fragmento o una variante de los mismos, y tiene componentes de conjugación que unen la molécula o moléculas de fármaco al componente proteico. En una realización de lo anterior, la composición de BFP-D puede tener uno o más XTEN genomodificados en los que el XTEN tiene aproximadamente un 80 % de identidad de secuencia con un XTEN seleccionado de la tabla 4, o alternativamente un 81 %, un 82 %, un 83 %, un 84 %, un 85 %, un 86 %, un 87 %, un 88 %, un 89 %, un 90 %, un 91 %, un 92 %, un 93 %, un 94 %, un 95 %, un 96 %, un 97 %, un 98 % o aproximadamente un 99 % de identidad de secuencia con un XTEN seleccionado de la tabla 4, y comprende uno o más restos de cisteína. En otra realización de lo anterior, la composición de BFP-D puede tener uno o más XTEN genomodificados en los que el XTEN tiene aproximadamente un 80 % de identidad de secuencia con un XTEN seleccionado de la tabla 4, o alternativamente un 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o aproximadamente un 99 % de identidad de secuencia con un XTEN seleccionado de la tabla 4, y comprende uno o más restos de lisina. En cualquiera de las realizaciones descritas anteriormente en este párrafo, el XTEN genomodificado puede presentar aproximadamente un 80 % de identidad de secuencia con un XTEN seleccionado de la tabla 8, o alternativamente un 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o presentar un 100 % de identidad de secuencia con un XTEN seleccionado de la tabla 8.

En una realización ilustrativa de un BFP-D, la composición de proteína de fusión de unión-fármaco comprende uno o más restos de direccionamiento anti-HER2, un XTEN genomodificado y una o más moléculas de paclitaxel unidas al XTEN mediante un agente de conjugación. En otra realización, la composición de proteína de fusión de unión-fármaco comprende uno o más restos de direccionamiento anti-HER2, un XTEN genomodificado y una o más moléculas de docetaxel unidas al XTEN mediante un agente de conjugación. En otra realización, la composición de proteína de fusión de unión-fármaco comprende uno o más restos de direccionamiento anti-HER2, un XTEN genomodificado y una o más moléculas de irinotecán unido al XTEN mediante un agente de conjugación. En cualquiera de las realizaciones anteriores del párrafo, la invención engloba composiciones que se pueden configurar de acuerdo con la fórmula V o la fórmula VI.

En una realización ilustrativa de un XTEN-D, la composición de XTEN-fármaco comprende un XTEN genomodificado y una o más moléculas de paclitaxel unidas al XTEN mediante un agente de conjugación. En otra realización, la composición de XTEN-fármaco comprende un XTEN genomodificado y una o más moléculas de docetaxel unidas al XTEN mediante un agente de conjugación. En otra realización, la composición de XTEN-fármaco comprende un XTEN genomodificado y una o más moléculas de irinotecán unidas al XTEN mediante un agente de conjugación. En cualquiera de las realizaciones anteriores del párrafo, la invención engloba composiciones que pueden se configurar de acuerdo con la fórmula VI o la fórmula VII.

Generalmente, las composiciones de conjugado de proteína de fusión de unión-fármaco de la invención conservan la capacidad de unión al antígeno de su resto de direccionamiento. En algunos casos, en donde los fármacos están restringidos a la porción portadora XTEN de la proteína de fusión, se reduce el impedimento estérico entre el resto de fármaco, el resto de direccionamiento y el antígeno o ligando diana. Por lo tanto, en las realizaciones preferidas, los conjugados de proteína de fusión de fusión-fármaco genomodificados son capaces de unirse, preferentemente específicamente, a los antígenos diana y administrar el fármaco en la ubicación diana. Dichos antígenos incluyen, por ejemplo, antígenos asociados a tumores (TAA), proteínas receptoras de la superficie celular y otras moléculas de la

superficie celular, proteínas transmembranarias, proteínas de señalización, factores reguladores de la supervivencia celular, factores reguladores de la proliferación celular, moléculas asociadas con el desarrollo o la diferenciación tisular, linfocinas, citocinas, moléculas implicadas en la regulación del ciclo celular, moléculas implicadas en la vasculogénesis o angiogénesis. Un antígeno al que una composición de proteína de fusión de unión-fármaco genomodificada es capaz de unirse puede ser miembro de un subconjunto de una de las categorías mencionadas anteriormente, en donde el otro subconjunto o subconjuntos de dicha categoría comprenden otras moléculas/antígenos que tienen una característica distinta (con respecto al antígeno de interés).

La divulgación también contempla que el XTEN-fármaco englobe composiciones en las que el fármaco conjugado con el XTEN pueda ser cualquiera de los fármacos de la tabla 9 o una sal, un ácido o un derivado farmacéuticamente aceptable del mismo, y el componente XTEN pueda ser un XTEN genomodificado con cisteína o lisina derivado de, o que presenta una identidad de secuencia sustancial con, cualquiera de los XTEN de la tabla 4 o un fragmento o una variante de los mismos, y tiene componentes de conjugación que unen la molécula o moléculas del fármaco al componente XTEN. En una realización de lo anterior, la composición de XTEN-fármaco puede tener XTEN genomodificados en los que el XTEN tiene aproximadamente un 80 % de identidad de secuencia con un XTEN seleccionado de la tabla 4, o alternativamente un 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o aproximadamente un 99 % de identidad de secuencia con un XTEN seleccionado de la tabla 4, y comprende uno o más restos de cisteína. En otra realización de lo anterior, la composición de XTEN-fármaco puede tener un XTEN genomodificado en los que el XTEN tiene aproximadamente un 80 % de identidad de secuencia con un XTEN seleccionado de la tabla 4, o alternativamente un 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o aproximadamente un 99 % de identidad de secuencia con un XTEN seleccionado de la tabla 4, y comprende uno o más restos de lisina. En algunos casos, el XTEN-fármaco engloba composiciones en las que el fármaco conjugado con el XTEN puede ser cualquiera de los fármacos de la tabla 9 o una sal, un ácido o un derivado farmacéuticamente aceptable del mismo, y el componente XTEN puede ser un XTEN genomodificado con cisteína o lisina derivado de, o que presenta una identidad de secuencia sustancial con, cualquiera de los XTEN genomodificados de la tabla 8 o un fragmento o una variante de los mismos, y tiene componentes de conjugación que unen la molécula o moléculas a los restos de cisteína o lisina del componente XTEN. En una realización de lo anterior, el XTEN genomodificado puede presentar aproximadamente un 80 % de identidad de secuencia con un XTEN seleccionado de la tabla 8, o alternativamente un 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o presenta un 100 % de identidad de secuencia con un XTEN seleccionado de la tabla 8.

(k) Métodos de uso de las composiciones de BFP-fármaco y de XTEN-fármaco

En el presente documento se describe un método para conseguir un efecto beneficioso en una enfermedad, un trastorno o una afección mediado por una composición de BFP-D o de XTEN-fármaco. En el presente documento se describe el uso de un BFP-D, en el que el resto de direccionamiento de la proteína de fusión de unión deriva de un anticuerpo progenitor que se une a una diana seleccionada del grupo que consiste en las dianas de la tabla 1 o la tabla 2, y el fármaco se selecciona de la tabla 9, en el tratamiento de una enfermedad, un trastorno o una afección a un sujeto que lo necesite mediante la administración de una cantidad terapéuticamente eficaz del BFP-D, en donde dicha administración conduce a la erradicación o la mejora de uno o más de los síntomas fisiológicos o clínicos asociados con el trastorno subyacente, de modo que se observa una mejora en el sujeto, a pesar de que el sujeto todavía puede estar afectado por el trastorno subyacente. En el presente documento se describe un método para tratar una enfermedad, un trastorno o una afección en un mamífero que comprende administrar al mamífero una cantidad terapéuticamente eficaz de un BFP-D que comprende uno o más restos de direccionamiento dirigidos a una o más dianas seleccionadas de la tabla 1 o la tabla 2, unidos a una o más moléculas de secuencias XTEN y, opcionalmente, a uno o más conectores, para formar el componente de la proteína de fusión de unión, en donde la unión no altera sustancialmente las propiedades funcionales esenciales de afinidad de unión y semivida terminal sostenida o tasa de eliminación sérica reducida en comparación con la del resto de direccionamiento progenitor del que deriva el componente de la proteína de fusión de unión, y en donde el fármaco conjugado con el BFP-D se selecciona de la tabla 9, en donde la administración del BFP-D a un sujeto que lo necesita consigue un efecto terapéutico beneficioso. La cantidad eficaz puede producir un efecto beneficioso para ayudar a tratar (por ejemplo, curar o reducir la gravedad) o prevenir (por ejemplo, reducir la probabilidad de aparición o la gravedad) una enfermedad, un trastorno o una afección, tal como, pero no sin limitación un cáncer, una enfermedad o afección cardiovascular, una enfermedad infecciosa, una afección inflamatoria, una afección respiratoria, un rechazo de trasplante de órganos o una enfermedad metabólica mediada por o asociada con una o más dianas, seleccionada de la tabla 1 o la tabla 2.

La incorporación de un fármaco en las proteínas de fusión inventivas proporciona composiciones mejoradas que pueden dar como resultado la cura, la mitigación, el tratamiento o la prevención de enfermedades, trastornos o afecciones en el ser humano o en otros animales. Los conjugados de fármacos representados por la fórmula V o la fórmula VI o la fórmula VII o la fórmula VIII de la presente invención son eficaces para los fines habituales para los que son eficaces los fármacos correspondientes. En una realización, las composiciones de BFP-D pueden tener una eficacia superior en comparación con el fármaco no conjugado debido a la capacidad, inherente al resto de direccionamiento, de transportar el fármaco a las células deseadas, donde es de particular beneficio. En los ejemplos, a continuación, se proporcionan realizaciones ilustrativas de los datos anteriores y representativos. En otra realización, las composiciones de BFP-D y XTEN-fármaco pueden tener una eficacia superior en comparación con el fármaco no

conjugado debido a la semivida terminal mejorada conferida por el portador XTEN. En otra realización, la invención proporciona composiciones de BFP-D que pueden tener una eficacia superior, una respuesta farmacológica mejorada y/o una toxicidad reducida en comparación con el fármaco no conjugado debido a la compartimentación diferencial de la composición en comparación con el fármaco no conjugado; por ejemplo, ausencia de penetración a través de la barrera hematoencefálica, las barreras nerviosas o las barreras citoplásmicas de células no diana. En los ejemplos, a continuación, se proporcionan realizaciones ilustrativas de los datos anteriores y representativos. En una ventaja particular de las composiciones inventivas, dichas propiedades mejoradas permiten formulaciones farmacéuticas o métodos de tratamiento de dosis más bajas utilizando una pauta posológica o dosis reducida, tanto debido a la administración dirigida a tejidos y células como debido a las propiedades farmacocinéticas mejoradas, lo que da como resultado un índice terapéutico superior; es decir, una eficacia mejorada con una toxicidad reducida. La divulgación prevé métodos de uso de las composiciones de conjugado en métodos terapéuticos y de diagnóstico, por ejemplo, para terapias dirigidas a tumores que tienen una tasa alterada de captación o difusión tisular en comparación con el fármaco activo solo.

En una realización, el método comprende administrar una cantidad terapéuticamente eficaz de una composición farmacéutica que comprende un BFP-D que comprende uno o más restos de direccionamiento unidos a una o más secuencias XTEN y al menos un portador farmacéuticamente aceptable a un sujeto que lo necesite, lo que da como resultado una mejora en al menos un parámetro, condición fisiológica o resultado clínico mediado por el componente o componentes del resto de direccionamiento. En otra realización, el método comprende administrar una cantidad terapéuticamente eficaz de una composición farmacéutica que comprende un fármaco, tal como, pero sin limitación, un fármaco seleccionado de la tabla 9, unido a una secuencia de XTEN y al menos un portador farmacéuticamente aceptable, a un sujeto que lo necesite, que da como resultado una mejora en al menos un parámetro, condición fisiológica o resultado clínico mediado por el componente o componentes del resto de direccionamiento. Los métodos contemplan la administración de la composición farmacéutica por cualquier vía apropiada para la enfermedad, el trastorno o la afección que se está tratando, incluyendo por vía subcutánea, intramuscular, intravítrea o intravenosa.

Los métodos de la divulgación pueden incluir la administración de dosis consecutivas de una cantidad terapéuticamente eficaz de la composición farmacéutica durante un período de tiempo suficiente para conseguir y/o mantener el parámetro o efecto clínico deseado, y dichas dosis consecutivas de una cantidad terapéuticamente eficaz establecen la pauta posológica terapéuticamente eficaz para la composición farmacéutica; es decir, el programa para las dosis administradas consecutivamente, en donde las dosis se administran en cantidades terapéuticamente eficaces para dar como resultado un efecto beneficioso sostenido sobre cualquier signo o síntoma clínico, aspecto, parámetro medido o característica de un cuadro clínico o afección metabólica, que incluye, pero no sin limitación, los descritos en el presente documento.

Una cantidad terapéuticamente eficaz de la composición farmacéutica puede variar según factores tales como el cuadro clínico, la edad, el sexo y el peso del individuo, y la capacidad del anticuerpo o de la porción de anticuerpo para provocar una respuesta deseada en el individuo. Una cantidad terapéuticamente eficaz es también una en la que cualquier efecto tóxico o perjudicial de la proteína de fusión de unión se ve superado por los efectos terapéuticamente beneficiosos. Una cantidad profilácticamente eficaz se refiere a una cantidad de composición farmacéutica requerida durante el período de tiempo necesario para conseguir el resultado profiláctico deseado.

Para los métodos divulgados, se prefieren composiciones de BFP-D o XTEN-fármaco o composiciones farmacéuticas que comprenden las composiciones de BFP-D o XTEN-fármaco de acción más prolongada, con el fin de mejorar la comodidad del paciente, aumentar el intervalo entre dosis y reducir la cantidad de fármaco necesaria para conseguir un efecto sostenido. En una realización, un método de tratamiento comprende la administración de una dosis terapéuticamente eficaz de un BFP-D o un XTEN-fármaco a un sujeto que lo necesite, que da como resultado un aumento en el tiempo invertido dentro de una ventana terapéutica establecida para el resto de direccionamiento o los componentes de fármaco de la composición farmacéutica en comparación con el componente de fármaco correspondiente no unido a la proteína de fusión y administrado a una dosis comparable a un sujeto. En algunos casos, el aumento en el tiempo invertido dentro de la ventana terapéutica es al menos aproximadamente tres veces, o al menos aproximadamente cuatro veces, o al menos aproximadamente cinco veces, o al menos aproximadamente seis veces, o al menos aproximadamente ocho veces, o al menos aproximadamente 10 veces, o al menos aproximadamente 20 veces, o al menos aproximadamente 40 veces en comparación con el resto de direccionamiento correspondiente o los componentes de fármaco no unidos a la proteína de fusión y administrados a una dosis comparable a un sujeto. Los métodos prevén además que la administración de múltiples dosis consecutivas de una composición farmacéutica administrada utilizando una pauta posológica terapéuticamente eficaz a un sujeto que lo necesite puede dar como resultado un aumento en el tiempo entre máximos de $C_{máx}$ consecutivos y/o cubetas de $C_{mín}$ para los niveles en sangre de la composición en comparación con el resto de direccionamiento correspondiente o los componentes de fármaco no unidos a la proteína de fusión. En la realización anterior, el aumento en el tiempo transcurrido entre máximos de $C_{máx}$ consecutivos y/o cubetas de $C_{mín}$ puede ser al menos aproximadamente tres veces, o al menos aproximadamente cuatro veces, o al menos aproximadamente cinco veces, o al menos aproximadamente seis veces, o al menos aproximadamente ocho veces, o al menos aproximadamente 10 veces, o al menos aproximadamente 20 veces, o al menos aproximadamente 40 veces en comparación con el componente o componentes de fármaco correspondientes no unidos a la proteína de fusión y administrados utilizando una pauta posológica comparable establecida para ese fármaco. En las realizaciones descritas anteriormente en este párrafo, la

administración de la proteína de fusión o la composición farmacéutica puede dar como resultado una mejora en al menos un parámetro que se sabe que es útil para evaluar las enfermedades, afecciones o trastornos del sujeto) utilizando una dosis unitaria más baja en moles de proteína de fusión en comparación con el correspondiente componente o componentes del resto de direccionamiento o el componente o componentes de fármaco no unidos a la proteína de fusión y administrados a una dosis unitaria o pauta posológica comparable a un sujeto.

En una realización, la administración de una composición farmacéutica de BFP-D o XTEN puede dar como resultado una mejora en uno de los parámetros clínicos, bioquímicos o fisiológicos que es mayor que la lograda mediante la administración del resto de direccionamiento o los componentes de fármaco no unidos a XTEN, determinados utilizando el mismo ensayo o basado en un parámetro clínico medido. En otra realización, la administración de la composición farmacéutica de BFP-D o XTEN puede dar como resultado una mejora de dos o más parámetros clínicos o relacionados con el metabolismo, cada uno mediado por uno de los diferentes restos de direccionamiento que, en conjunto, dan como resultado un efecto mejorado en comparación con el componente de resto de direccionamiento no unido al XTEN, determinados utilizando los mismos ensayos o basados en parámetros clínicos medidos. En otra realización, la administración de la proteína de fusión de unión o la composición farmacéutica puede dar como resultado una actividad en uno o más de los parámetros clínicos o bioquímicos o fisiológicos que es de mayor duración que la actividad de uno de los componentes del resto de direccionamiento único o de fármaco no unido al XTEN, determinado utilizando el mismo ensayo o basado en un parámetro clínico medido.

Las presentes composiciones de conjugados de BFP-D o de XTEN-fármaco pueden ser útiles en el tratamiento de diferentes enfermedades, trastornos y afecciones. En una realización, la enfermedad es cáncer, incluyendo carcinomas/tumores, melanomas, sarcomas, leucemias y linfomas, y gliomas. Algunas afecciones o trastornos hiperproliferativos ilustrativos incluyen tumores benignos o malignos, trastornos neuronales, gliales, astrocitales, hipotálamicos, glandulares, de macrófagos, epiteliales, estromales, blastocélicos, inflamatorios, angiogénicos e inmunitarios, incluyendo autoinmunitarios. Una o más composiciones de la invención pueden usarse junto con otro tratamiento o fármaco para tratar a un paciente. Como se utiliza en el presente documento, un "paciente" se refiere a cualquier mamífero, incluyendo un ser humano, y puede estar afectado por cualquier enfermedad, trastorno o afección que pueda tratarse eficazmente con un anticuerpo terapéutico o un fármaco, incluyendo enfermedades, trastornos o afecciones asociadas con las dianas de las tablas 1 o 2, o enfermedades, trastornos o afecciones tratadas convencionalmente con los fármacos de la tabla 9. El tratamiento con un conjugado puede ser un tratamiento primario para una enfermedad o un trastorno, o puede ser una terapia adyuvante para una enfermedad o un trastorno. Además, el tratamiento puede ser de una enfermedad existente o puede ser profiláctico. Por lo general, se pueden determinar las dosis y los cronogramas de dosis apropiados para las composiciones inventivas utilizando modelos experimentales y/o ensayos clínicos. Puede preferirse el uso de la dosis mínima que es suficiente para proporcionar una terapia eficaz en algunas circunstancias, tales como cuando el fármaco está asociado con reacciones adversas perjudiciales. Por lo general, se puede monitorizar la eficacia terapéutica o profiláctica en los pacientes utilizando ensayos adecuados para la afección que se está tratando o previniendo, que serán familiares para los expertos habituales en la técnica. Dichos métodos permitirán establecer la ventana terapéutica, así como la dosis máxima tolerada para las composiciones de la invención.

Las composiciones de XTEN-fármaco y/o las composiciones de proteína de fusión de unión-fármaco y las composiciones farmacéuticas que comprenden el BFP-D o el XTEN-D se pueden administrar por las vías y métodos y con los cronogramas de administración apropiados para la enfermedad, el trastorno o la afección dados. En ciertas realizaciones, las composiciones de XTEN-fármaco y/o las composiciones de proteína de fusión de unión-fármaco de la invención se administran como una infusión i.v. En otras realizaciones, las composiciones se administran por vía subcutánea o intramuscular. En otras realizaciones más, las composiciones se administran por vía oral. En el caso de la infusión intravenosa, la administración puede durar cualquier período de tiempo apropiado, que es fácilmente determinable y evaluable por un experto habitual en la materia. Por ejemplo, las infusiones pueden durar de aproximadamente una a aproximadamente 24 horas, aunque los tiempos de infusión más cortos o más largos entran dentro del alcance de la invención. En ciertas realizaciones, las infusiones se administran diariamente, semanalmente, cada dos semanas, cada 21 días o mensualmente. Los períodos de tiempo apropiados son conocidos por un experto en la técnica, y pueden determinarse tomando como base varios factores, incluyendo el tipo de terapia o de fármaco que se usa junto con el conjugado de XTEN-fármaco. Los facultativos con conocimientos habituales en la técnica de la medicina sabrán que la dosis que se administra a un paciente variará de acuerdo con la edad, el peso y el estado físico del paciente, la vía de administración, la enfermedad específica que se está tratando, la fase de la enfermedad y similares. Para cualquier sujeto particular, las pautas posológicas específicas (tanto la dosis como la frecuencia de administración) deben ser ajustadas para ese paciente por un profesional experto. Algunos ejemplos de diferentes intervalos de dosis y cronogramas de administración se proporcionan en la Pat. de EE. UU. N.º 5.670.537.

Las composiciones conjugadas de XTEN-fármaco dirigidas útiles en el tratamiento del cáncer incluyen, pero sin limitación, composiciones que comprenden uno o más restos de direccionamiento contra los receptores de la superficie celular y los antígenos asociados a tumores (TAA). Los antígenos asociados a tumores son conocidos en la técnica y se pueden preparar para usar en la generación de restos de direccionamiento utilizando métodos e información que son bien conocidos en la técnica. Algunos ejemplos no limitativos de polipéptidos asociados a tumores se indican en la tabla 2. En las realizaciones preferidas, las dianas a las que se dirigirían los restos de direccionamiento podrían incluir receptores expresados específicamente en la superficie de uno o más tipos particulares de células cancerosas

en comparación con una o más células normales no cancerosas. A menudo, dichos polipéptidos asociados a tumores se expresan más abundantemente en la superficie de las células cancerosas en comparación con la superficie de las células no cancerosas; por ejemplo, HER2 en ciertos cánceres de mama. La identificación de dichos polipéptidos de antígenos de la superficie celular asociados a tumores ha dado lugar a la capacidad de dirigirse específicamente a las células cancerosas para su destrucción a través de terapias basadas en anticuerpos.

Trastornos mediados por células que expresan el VEGF

Una composición de BFP-D que comprende un resto de direccionamiento derivado de un anticuerpo o un fragmento anti-VEGF puede utilizarse ventajosamente en un método para tratar una enfermedad o un trastorno mediado por el VEGF, tales como trastornos neovasculares. En el presente documento se describe un método para tratar un trastorno tumoral sólido en un paciente humano que comprende administrar al paciente una cantidad eficaz de un BFP-D o de una composición farmacéutica en donde al menos un resto de direccionamiento de la proteína de fusión de unión comprende un sitio de unión al antígeno que se une al VEGF humano, en donde la unión puede inhibir la vascularización del tumor, y el BFP-D comprende además un fármaco citotóxico conjugado, tal como un fármaco seleccionado de la tabla 9, para tratar un tumor sólido. En otra realización más, el trastorno tumoral sólido en el método anterior se selecciona del grupo que consiste en carcinomas de mama, carcinomas de pulmón, carcinomas gástricos, carcinomas de esófago, carcinomas colorrectales, carcinomas de hígado, carcinomas de ovario, tecomas, arrenoblastomas, carcinomas cervicouterinos, carcinoma endometrial, hiperplasia endometrial, endometriosis, fibrosarcomas, coriocarcinoma, cáncer de cabeza y cuello, carcinoma nasofaríngeo, carcinomas laríngeos, hepatoblastoma, sarcoma de Kaposi, melanoma, carcinomas de piel, hemangioma, hemangioma cavernoso, hemangioblastoma, carcinomas de páncreas, retinoblastoma, astrocitoma, glioblastoma, schwannoma, oligodendroglioma, meduloblastoma, neuroblastomas, rabdomiosarcoma, sarcoma osteógeno, leiomiosarcomas, carcinomas de las vías urinarias, carcinomas de tiroides, tumor de Wilm, carcinoma de células renales, carcinoma de próstata, proliferación vascular anormal asociada a facomatosis, edema (tal como el asociado con tumores cerebrales) y síndrome de Meigs.

En el presente documento se describe un método para tratar un trastorno neovascular intraocular en un paciente humano que comprende administrar al paciente una cantidad terapéuticamente eficaz de un BFP-D o una composición farmacéutica que comprende el BFP-D, en donde al menos un resto de direccionamiento comprende un sitio de unión al antígeno que se une al VEGF humano y el BFP-D comprende además un fármaco citotóxico conjugado, tal como un fármaco seleccionado de la tabla 9. En una realización adicional, el trastorno neovascular intraocular se selecciona del grupo que consiste en retinopatías diabéticas y otras retinopatías proliferativas, incluidas retinopatía del prematuro, fibroplasia retrocristaliniana, glaucoma neovascular y degeneración macular senil.

Trastornos mediados por células que expresan el HER2

En otra realización, la proteína de fusión de unión se usa para tratar trastornos mediados por células que expresan el HER2. En el presente documento se describe un método para tratar una enfermedad humana mediada por células que expresan el HER2 con una composición de proteína de fusión de unión que deriva de un anticuerpo progenitor que se une al HER2. Dichas composiciones tienen aplicaciones profilácticas y terapéuticas en un amplio espectro de trastornos mediados por células que expresan el HER2, que incluyen patologías apoyadas por la proliferación de células que expresan el HER2, tales como cánceres caracterizados por la sobreexpresión del HER2, de una manera similar a la aplicación de anticuerpos anti-Her2 completos en el tratamiento de dichas indicaciones de enfermedad que es conocida en la técnica, indicaciones de tratamiento que incluyen cánceres de mama, de ovario y de pulmón que sobreexpresan el HER2.

En el presente documento se describe un método para tratar un trastorno mediado por células que expresan el HER2 en un paciente humano que comprende administrar al paciente una cantidad terapéuticamente eficaz de un BFP-D o de una composición farmacéutica que comprende un BFP-D, en donde al menos un resto de direccionamiento de la proteína de fusión de unión comprende un sitio de unión al antígeno eso se une al HER2, y el BFP-D comprende además un fármaco citotóxico conjugado, tal como un fármaco seleccionado de la tabla 9. El trastorno puede ser un trastorno proliferativo de células que expresan el HER2, que incluye un tumor benigno o maligno caracterizado por la sobreexpresión del receptor de la ErbB2, por ejemplo, un cáncer, tal como cáncer de mama, cáncer epidermoide, cáncer de pulmón microcítico, cáncer de pulmón no microcítico, cáncer gastrointestinal, cáncer de páncreas, glioblastoma, cáncer cervicouterino, cáncer de ovario, cáncer de hígado, cáncer de vejiga, hepatoma, cáncer de colon, cáncer colorrectal, carcinoma endometrial, carcinoma de glándula salival, cáncer de riñón, cáncer de hígado, cáncer de próstata, cáncer de vulva, cáncer de tiroides, carcinoma hepático y diferentes tipos de cáncer de cabeza y cuello. Además, la divulgación contempla el uso del conjugado anterior en lugar del anticuerpo anti-Her2 completo en el tratamiento de cánceres que sobreexpresan el HER2, como se describe en la Pat. de EE. UU. N.º 5.725.856.

En un método ilustrativo no limitativo, en el presente documento se describe un método para inhibir el crecimiento de células tumorales administrando a un paciente una cantidad terapéuticamente eficaz de composición de BFP-D anti-HER2 capaz de inhibir la función del receptor HER2 mediante el componente del resto de direccionamiento y/o inhibiendo o destruyendo la célula HEr2 mediante el componente de fármaco, tal como, pero sin limitación epaclitaxel,

paclitaxel, docetaxel, doxetaxel, irinotecán, pemetrexed, cloranbucilo o gemcitabina, o un fármaco citotóxico adecuado seleccionado de la tabla 9.

Una realización adicional de la divulgación se refiere a la administración de una cantidad terapéuticamente eficaz de composición anti-HER2 capaz de inhibir la función del receptor del factor de crecimiento. Otro objeto más de la divulgación es proporcionar métodos para el tratamiento y/o la prevención de tumores que sobreexpresan el receptor de la erbB-2 que comprenden la administración de una cantidad antitumoral eficaz de al menos uno de los BFP-D anti-HER2 divulgados capaces de unirse a células cancerosas asociadas por la sobreexpresión de la proteína erbB-2. En el presente documento se describe un método para el tratamiento y/o la prevención de tumores que sobreexpresan el receptor de la erbB-2 que comprende la administración de cantidades terapéuticamente eficaces de BFP-D anti-Her2 que comprende un primer y un segundo resto de unión anti-Her2, que puede ser el mismo o que puede unirse a diferentes epítopos de la proteína erbB-2, capaz de inhibir la función del receptor HER2, y una o más moléculas de fármaco seleccionadas de la tabla 9. Preferentemente, dichas combinaciones de restos de unión y fármaco presentarán una actividad citotóxica mejor de la que se esperaría para la suma de la actividad citotóxica de los anticuerpos individuales y los fármacos administrados por separado a la misma o menor concentración global de los componentes individuales.

Trastornos mediados por células que expresan el EGFR

En el presente documento se describe un método para tratar una enfermedad humana mediada por células que expresan el EGFR con un BFP-D que deriva de un anticuerpo progenitor que se une al EGFR humano (también conocido como ErbB-1 o Her1) y comprende además un fármaco seleccionado de la tabla 9. Dicho BFP-D puede tener aplicaciones profilácticas y terapéuticas en un amplio espectro de trastornos mediados por células que expresan el EGFR, que incluyen patologías apoyadas por la proliferación de células que expresan el EGFR, tales como cánceres caracterizados por la sobreexpresión del EGFR, que incluyen cánceres de mama, de ovario, de cabeza y cuello, de cerebro, de vejiga, de páncreas y de pulmón.

En el presente documento se describe un método para tratar un trastorno de proliferación celular en un paciente humano caracterizado por la sobreexpresión del EGFR que comprende administrar al paciente una cantidad terapéuticamente eficaz de un BFP-D, en donde al menos un resto de direccionamiento de la proteína de fusión de unión comprende un sitio de unión al antígeno que se une para el EGFR humano y comprende además un fármaco citotóxico conocido por ser eficaz contra células portadoras del EGFR. El trastorno puede ser un tumor benigno o maligno caracterizado por la sobreexpresión del EGFR, por ejemplo, un cáncer, tal como cáncer de mama, cáncer epidermoide, cáncer de pulmón microcítico, cáncer de pulmón no microcítico, cáncer gastrointestinal, cáncer de páncreas, glioblastoma, cáncer cervicouterino, cáncer de ovario, cáncer de hígado, cáncer de vejiga, hepatoma, cáncer de colon, cáncer colorrectal, carcinoma endometrial, carcinoma de glándula salival, cáncer de riñón, cáncer de hígado, cáncer de próstata, cáncer de vulva, cáncer de tiroides, carcinoma hepático y diferentes tipos de cáncer de cabeza y cuello.

Trastornos mediados por células que expresan el CD20

En el presente documento se describe un método para tratar trastornos mediados por células que expresan el CD20. En el presente documento se describe un método para tratar una enfermedad humana mediada por células que expresan el CD20 con una composición de BFP-D que deriva de un anticuerpo progenitor que se une al CD20 humano y comprende además un fármaco citotóxico conocido por tener actividad contra las células que expresan el CD20, tal como un fármaco seleccionado de la tabla 9. Dichas composiciones tienen aplicaciones profilácticas y terapéuticas en un amplio espectro de trastornos mediados por células que expresan el CD20, que incluyen patologías apoyadas por la proliferación de células que expresan el CD20, tales como cánceres de células que expresan el CD20, de una manera similar a la aplicación de anticuerpos anti-CD20 completos en el tratamiento de dichas indicaciones de enfermedad conocidas en la técnica, indicaciones de tratamiento que incluyen linfomas de linfocitos B, como se describe en la patente de EE. UU. N.º 6.682.734.

Trastornos mediados por células que expresan el CD18

En el presente documento se describe un método para tratar trastornos mediados por células que expresan el CD18. En el presente documento se describe un método para tratar una enfermedad humana mediada por células que expresan el CD18 con una composición de BFP-D que deriva de un anticuerpo progenitor que se une al CD18 humano y comprende además un fármaco citotóxico conocido por tener actividad contra las células que expresan el CD 18, tal como un fármaco seleccionado de la tabla 9. Dichas composiciones tienen aplicaciones profilácticas y terapéuticas en un amplio espectro de trastornos mediados por células que expresan el CD18, que incluyen patologías apoyadas por la adhesión de leucocitos, de una manera similar a la aplicación de anticuerpos anti-CD18 completos en el tratamiento de dichas indicaciones de enfermedad conocidas en la técnica, indicaciones de tratamiento que incluyen infarto agudo de miocardio e ictus. En el presente documento se describe un método para tratar un trastorno en un paciente humano mediado por una célula que expresa el CD18, que comprende administrar al paciente una cantidad terapéuticamente eficaz de un BFP-D en donde al menos un resto de direccionamiento del conjugado comprende un sitio de unión al antígeno que se une al CD18 humano, y el BFP-D comprende además un fármaco conjugado conocido por tener

efectos beneficiosos en el tratamiento del infarto de miocardio y el ictus, tal como un fármaco seleccionado de la tabla 9. En otra realización, el trastorno mediado por células que expresan el CD18 es un trastorno inflamatorio, tal como un trastorno isquémico de reperfusión, que incluye infarto agudo de miocardio e ictus. Además, la divulgación contempla el uso de la proteína de fusión de unión anterior en lugar del anticuerpo anti-CD 18 completo en el tratamiento del ictus, como se describe en la Publicación PCT WO 97/26912.

En el presente documento se describe un método para tratar un trastorno mediado por la LFA-1 en un ser humano, que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión en donde al menos un resto de direccionamiento que se une al CD 18 humano, y el BFP-D comprende además un fármaco inmunosupresor conjugado, tal como un fármaco seleccionado de la tabla 9. Además, la divulgación contempla el uso de la proteína de fusión de unión anterior en lugar de anticuerpo anti-CD 18 completo en el tratamiento de un trastorno mediado por una LFA-1, tal como psoriasis y rechazo del injerto, en un paciente humano, como se describe en la Patente de EE. UU. N.º 5.622.700.

Trastornos mediados por células que expresan el CD11a

En el presente documento se describe un método para tratar trastornos mediado por células que expresan el CD11a. En el presente documento se describe un método para tratar una enfermedad humana mediada por una célula que expresa el CD11a con una composición de proteína de fusión de unión que deriva de un anticuerpo progenitor que se une al CD11a humano y comprende además un fármaco inmunosupresor o citotóxico conocido por tener actividad contra las células que expresan el CD11a. Dichas composiciones tienen aplicaciones profilácticas y terapéuticas en un amplio espectro de trastornos mediados por células que expresan el CD11a, que incluyen patologías apoyadas por la adhesión de leucocitos, de una manera similar a la aplicación de anticuerpos anti-CD11a completos en el tratamiento de dichas indicaciones de enfermedad conocidas en la técnica, indicaciones de tratamiento que incluyen psoriasis, asma, rechazo del injerto y esclerosis múltiple. En el presente documento se describe un método para tratar un trastorno mediado por la LFA-1 en un ser humano, que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión en donde al menos un resto de direccionamiento del conjugado comprende un sitio de unión al antígeno que se une al CD11a humano. Además, la divulgación contempla el uso de la proteína de fusión de unión anterior en lugar de anticuerpo anti-CD11a completo en el tratamiento de un trastorno mediado por una LFA-1, tal como psoriasis y rechazo del injerto, en un paciente humano, como se describe en la Patente de EE. UU. n.º 5.622.700. En otro aspecto, la divulgación contempla el uso de las proteínas de fusión de unión anteriores en lugar del anticuerpo anti-CD11a completo en el tratamiento de trastornos mediados por la LFA-1 en un paciente humano, como se describe en la Patente de EE. UU. n.º 6.037.454.

Trastornos mediados por células que expresan el CD3

En el presente documento se describe un método para tratar una enfermedad o un trastorno humano mediado por células que expresan el CD3 con una proteína de fusión de unión que deriva de un anticuerpo progenitor que se une al CD3 humano y comprende además un fármaco citotóxico o inmunosupresor conocido por tener actividad contra las células que expresan el CD3, tal como un fármaco seleccionado de la tabla 9. Dichas proteínas de fusión de unión pueden tener aplicaciones profilácticas y terapéuticas en un amplio espectro de trastornos mediados por células que expresan el CD3, que incluyen afecciones asociadas con la proliferación o la activación de células que expresan el CD3, tales como trastornos inmunitarios mediados por linfocitos T y rechazo del injerto en receptores de un trasplante. El uso de anticuerpos anti-CD3 para tratar enfermedades y trastornos se ha descrito, por ejemplo, en la Patente de EE. UU. n.º 4.515.893. En otro aspecto, la invención contempla el uso de la proteína de fusión de unión anterior en lugar del anticuerpo anti-CD3 humano completo en el tratamiento del rechazo agudo de aloinjerto en receptores de trasplante renal, como se ha descrito para ORTHOCLONE OKT3 muromonab-CD3 en Physician's Desk Reference, 52ª edición (1998), págs. 1971-1974.

Trastornos mediados por células que expresan el TAC

En el presente documento se describe un método para tratar una enfermedad humana mediada por células que expresan la cadena α (TAC) del receptor de la interleucina-2 con una proteína de fusión de unión que deriva de un anticuerpo progenitor que se une al TAC humano y comprende además un fármaco citotóxico conocido por tener actividad contra las células que expresan el TAC. Dichas proteínas de fusión de unión pueden tener aplicaciones profilácticas y terapéuticas en un amplio espectro de trastornos mediados por células que expresan el TAC, incluyendo afecciones creadas por la proliferación o la activación de células que expresan el TAC y trastornos inmunitarios mediados por linfocitos T o linfocitos B, incluido el rechazo del injerto en receptores de un trasplante.

En el presente documento se describe un método para tratar un trastorno en un paciente humano mediado por una célula que expresa el TAC, que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión en donde al menos un resto de direccionamiento de la proteína de fusión de unión comprende un sitio de unión al antígeno que se une al TAC humano, y el BFP-D comprende además un fármaco inmunosupresor o citotóxico conjugado, tal como un fármaco seleccionado de la tabla 9. En otra realización, el trastorno mediado por células que expresan el TAC se caracteriza por la activación o la proliferación de linfocitos T o linfocitos B, incluidos trastornos inmunitarios tales como rechazo del injerto en receptores de un trasplante, enfermedad

del injerto contra el hospedador (GHVD), rechazo del injerto en receptores de un trasplante, tal como rechazo agudo del injerto en receptores de un trasplante renal, y enfermedades autoinmunitarias tales como diabetes de tipo I, esclerosis múltiple, artritis reumatoide, lupus eritematoso sistémico y miastenia grave. El uso de anticuerpos para tratar trastornos mediados por la cadena α del receptor de la interleucina-2 con anticuerpos se ha descrito en la Patente de EE. UU. N.º 5.693.761.

Trastornos mediados por el TNF-α

En el presente documento se describe un método para tratar una enfermedad mediada por el TNF-α con una proteína de fusión de unión que deriva de un anticuerpo progenitor que se une al TNF-α humano y comprende además un fármaco citotóxico o antiinflamatorio, tal como un fármaco seleccionado de la tabla 9. Dichas proteínas de fusión de unión pueden tener aplicaciones profilácticas y terapéuticas en un amplio espectro de trastornos mediados por el TNF-α, incluidos trastornos inflamatorios y trastornos inmunitarios, de una manera similar a la aplicación de anticuerpos anti-TNF-α humano completos en el tratamiento de indicaciones de enfermedades tales como enfermedad de Crohn, enfermedad inflamatoria intestinal y artritis reumatoide.

En el presente documento se describe un método para tratar un trastorno inflamatorio en un paciente humano que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión en donde al menos un resto de direccionamiento de la proteína de fusión de unión comprende un sitio de unión al antígeno que se une al TNF-α humano. En otra realización, el trastorno inflamatorio es la enfermedad de Crohn. En otra realización más, el trastorno inflamatorio es enfermedad inflamatoria intestinal. En otra realización más, el trastorno inflamatorio es artritis reumatoide. El uso de anticuerpos que se unen al TNF-α humano en el tratamiento de afecciones inflamatorias se ha descrito, por ejemplo, en las Patentes de EE. UU. n.º 5.672.347, 5.656.272 y 5.698.195.

Trastornos mediados por el factor tisular

En el presente documento se describe un método para tratar una enfermedad mediada por el factor tisular con una proteína de fusión de unión derivada de un anticuerpo progenitor que se une al factor tisular humano, y el BFP-D comprende además un fármaco anticoagulante o antitrombótico conjugado, tal como un fármaco seleccionado de la tabla 9. Dichas proteínas de fusión de unión pueden tener aplicaciones profilácticas y terapéuticas en un amplio espectro de trastornos mediados por el factor tisular, incluyendo patologías apoyadas por la coagulación sanguínea y en el tratamiento de indicaciones patológicas tales como trombosis venosa profunda, trombosis arterial, aterosclerosis, estenosis vascular, enfermedades isquémicas del miocardio, incluyendo infarto agudo de miocardio, reoclusión después de angioplastia o aterectomía o tratamiento trombolítico del infarto agudo de miocardio, angina de pecho, enfermedades isquémicas cerebrales que incluyen ictus, tromboflebitis venosa y embolia pulmonar. En el presente documento se describe un método para tratar una enfermedad o un trastorno mediado por el factor tisular (tal como lo anterior) en un paciente humano que comprende administrar al paciente una cantidad terapéuticamente eficaz de una proteína de fusión de unión en donde al menos un resto de direccionamiento de la proteína de fusión de unión comprende un sitio de unión al antígeno que se une al factor tisular humano.

III). SECUENCIAS DE ADN DE LA INVENCIÓN

La presente invención proporciona ácidos polinucleicos aislados que codifican un XTEN y polipéptidos quiméricos de proteína de fusión de unión y secuencias complementarias de las moléculas de ácido polinucleico que codifican un XTEN y polipéptidos quiméricos de proteína de fusión de unión, incluidas variantes homólogas. En el presente documento se describen métodos para producir ácidos polinucleicos que codifican un XTEN y polipéptidos quiméricos de proteína de fusión de unión y secuencias complementarias de moléculas de ácido polinucleico que codifican polipéptidos quiméricos de proteína de fusión de unión, incluidas variantes homólogas. En general, y como se ilustra en las FIG. 7-9, los métodos para producir una secuencia polinucleotídica que codifica un XTEN o una proteína de fusión de unión y expresar el producto génico resultante incluyen ensamblar los nucleótidos que codifican los restos de direccionamiento y el XTEN (y cualquier secuencia conectora, si la hubiera), unir los componentes en marco, incorporar el gen codificante en un vector de expresión apropiado, transformar una célula hospedadora apropiada con el vector de expresión y hacer que la proteína de fusión sea expresada en la célula hospedadora transformada, produciendo así la proteína de fusión de unión biológicamente activa. Se pueden usar técnicas recombinantes convencionales de biología molecular para crear los polinucleótidos y los vectores de expresión de la presente invención.

De acuerdo con la invención, se pueden usar secuencias de ácido nucleico que codifican un XTEN y proteínas de fusión de unión para generar moléculas de ADN recombinantes que dirigen la expresión del XTEN y proteínas de fusión de unión en células hospedadoras apropiadas. Se prevé que varias estrategias de clonación sean adecuadas para realizar la presente invención, muchas de las cuales pueden usarse para generar una construcción que comprende un gen que codifica una composición de proteína de fusión de unión de la presente invención, o su complemento. En una realización, la estrategia de clonación se usaría para crear un gen que codifica un polipéptido XTEN. En otra realización, la estrategia de clonación se usaría para crear un gen que codifica una proteína de fusión de unión monómera que comprende al menos un primer resto de direccionamiento y al menos un primer polipéptido XTEN, o su complemento. En otra realización, la estrategia de clonación se usaría para crear un gen que codifica una

proteína de fusión de unión monómera que comprende un primer y un segundo resto de direccionamiento y al menos un primer XTEN, o su complemento. En otra realización, la estrategia de clonación se usaría para crear un gen que codifica una proteína de fusión de unión monómera que comprende al menos un primer y un segundo resto de direccionamiento, un conector y al menos un primer XTEN, o su complemento. En las realizaciones anteriores, el gen se usaría en un vector de expresión adecuado para transformar una célula hospedadora para la expresión de la proteína de fusión.

Al diseñar las secuencias XTEN deseadas, se descubrió que la naturaleza no repetitiva del XTEN de las composiciones inventivas se puede conseguir a pesar del uso de un enfoque molecular de "bloque de construcción" en la creación de las secuencias codificantes del XTEN. Esto se consiguió mediante el uso de una biblioteca de polinucleótidos que codifican motivos de secuencia que a continuación se multimerizan para crear los genes que codifican las secuencias XTEN (véanse las FIG. 7 y 8). Por lo tanto, mientras que el XTEN expresado puede consistir en múltiples unidades de tan solo cuatro motivos de secuencia diferentes, debido a que los propios motivos consisten en secuencias de aminoácidos no repetitivas, la secuencia general de XTEN se vuelve no repetitiva. Por consiguiente, en una realización, los polinucleótidos que codifican el XTEN comprenden múltiples polinucleótidos que codifican secuencias no repetitivas, o motivos, unidos operativamente en marco y en los que las secuencias de aminoácidos del XTEN expresadas resultantes no son repetitivas.

En un enfoque, se prepara en primer lugar una construcción que contiene la secuencia de ADN correspondiente a la proteína de fusión de unión. El ADN que codifica el resto de direccionamiento de las composiciones se puede obtener a partir de una biblioteca de ADNc preparada utilizando métodos convencionales a partir de tejido o células aisladas que se cree que poseen el ARNm del resto de direccionamiento y para expresarlo a un nivel detectable. Si fuera necesario, la secuencia codificante se puede obtener utilizando procedimientos convencionales de extensión con cebador como se describe en Sambrook, *et al., supra,* para detectar precursores e intermedios de procesamiento del ARNm que pueden no haber sido retrotranscritos en ADNc. Por consiguiente, el ADN se puede obtener convenientemente a partir de una biblioteca de ADNc preparada a partir de dichas fuentes. El resto de direccionamiento que codifica el gen o genes también se puede obtener a partir de una biblioteca genómica o crear mediante procedimientos sintéticos convencionales conocidos en la técnica (por ejemplo, síntesis automatizada de ácido nucleico) utilizando secuencias de ADN obtenidas a partir de bases de datos, patentes o referencias bibliográficas disponibles públicamente. Dichos procedimientos son bien conocidos en la técnica y están bien descritos en la bibliografía científica y de patentes. Por ejemplo, se pueden obtener secuencias a partir de números de registro de Chemical Abstracts Services (CAS) (publicados por la American Chemical Society) y/o números de registro del GenBank (por ejemplo, locus ID, NM_XXXXX, NP_XXXXX, and XP_XXXXX), identificadores de proteínas modelo disponibles a través de la página web del National Center for Biotechnology Information (NCBI), disponible en la WWW en ncbi.nlm.nih.gov que corresponden a entradas del Registro CAS o de la base de datos GenBank que contienen una secuencia de aminoácidos del resto de direccionamiento (por ejemplo, un anticuerpo) o de un fragmento o variante del resto de direccionamiento. En una realización, el gen que codifica la proteína de fusión de unión codifica una proteína de una cualquiera de las tablas 25, 38 o 39, o un fragmento o variante de la misma.

Un gen o un polinucleótido que codifica la porción del resto de direccionamiento de la presente proteína de fusión de unión, en el caso de una proteína de fusión expresada que comprenderá un único resto de direccionamiento, se puede clonar a continuación en una construcción, que puede ser un plásmido u otro vector bajo el control de las secuencias de transcripción y de traducción apropiadas para la expresión de proteínas de alto nivel en un sistema biológico. En una etapa posterior, un segundo gen o polinucleótido que codifica el XTEN se fusiona genéticamente con los nucleótidos que codifican el extremo amínico y/o carboxílico del gen del resto de direccionamiento, clonándolo en la construcción adyacente y en marco con el gen o genes que codifican el resto de direccionamiento. Esta segunda etapa se puede producir a través de una etapa de ligación o de multimerización. En las realizaciones anteriores descritas anteriormente en este párrafo, debe entenderse que las construcciones génicas que se crean pueden ser alternativamente el complemento de los genes respectivos que codifican las respectivas proteínas de fusión. Además, para las proteínas de fusión de unión que comprenden dos o más restos de direccionamiento y conectores, el gen o los polinucleótidos que codifican estos componentes se clonarían en la construcción adyacente y en marco con respecto a los otros componentes descritos anteriormente, dependiendo de la configuración final deseada de la proteína de fusión. En un aspecto particular de lo anterior, se descubrió que el uso de secuencias codificantes alternativas para restos de direccionamiento multivalentes (por ejemplo, dos o más) reduce el riesgo de recombinación homóloga durante la expresión. Por consiguiente, para proteínas de fusión de unión que tienen dominios de unión repetidos o restos de direccionamiento multivalente con las mismas secuencias o muy similares, la invención proporciona polinucleótidos codificantes para los dominios respectivos que tienen secuencias de ADN diferentes. En un ejemplo no limitativo de lo anterior, la invención proporciona una proteína de fusión de unión con restos de direccionamiento de tipo Ig diméricos, en donde los codones de ciertos aminoácidos del gen codificante de cada resto de direccionamiento son variados, en donde la incidencia de recombinación durante la expresión en un hospedador transformado se reduce en comparación con un hospedador comparable transformado con genes de direccionamiento que son idénticos.

El gen que codifica el XTEN se puede hacer en una o más etapas, ya sea completamente sintéticamente o mediante una síntesis combinada con procesos enzimáticos, tales como clonación mediada por enzimas de restricción, PCR y extensión por solapamiento. Los polipéptidos XTEN se pueden construir de manera que el gen codificante del XTEN

tenga una baja repetitividad. Los genes que codifican el XTEN con secuencias no repetitivas se pueden ensamblar a partir de oligonucleótidos utilizando técnicas convencionales de síntesis génica. El diseño génico se puede realizar utilizando algoritmos que optimizan el uso de codones y la composición de aminoácidos. En un método de la invención, se crea una biblioteca de construcciones polinucleotídicas que codifican XTEN relativamente cortas, y a continuación se ensamblan, como se ilustra en las FIG. 7 y 8. Esto puede ser una biblioteca de codones pura, de modo que cada miembro de la biblioteca tenga la misma secuencia de aminoácidos, pero son posibles muchas secuencias codificantes diferentes. Dichas bibliotecas pueden ensamblarse a partir de oligonucleótidos parcialmente aleatorizados y usarse para generar grandes bibliotecas de segmentos XTEN que comprenden los motivos de secuencia. El esquema de aleatorización se puede optimizar para controlar las elecciones de aminoácidos para cada posición, así como el uso de codones.

Bibliotecas de polinucleótidos

En otro aspecto, la divulgación proporciona bibliotecas de polinucleótidos que codifican secuencias XTEN que pueden usarse para ensamblar genes que codifican XTEN de una longitud y secuencia deseadas, que son útiles para la creación de genes que codifican proteínas de fusión de unión.

En ciertas realizaciones, las construcciones de biblioteca que codifican XTEN comprenden polinucleótidos que codifican segmentos de polipéptidos de una longitud fija. Como una etapa inicial, se puede ensamblar una biblioteca de oligonucleótidos que codifican motivos de 9-14 restos de aminoácidos. En una realización preferida, se ensamblan bibliotecas de oligonucleótidos que codifican motivos de 12 aminoácidos.

Los segmentos de secuencia que codifican XTEN se pueden dimerizar o multimerizar en secuencias codificantes más largas. La dimerización o la multimerización se puede realizar mediante ligación, extensión por solapamiento, ensamblaje por PCR o técnicas de clonación similares conocidas en la técnica. Este proceso puede repetirse varias veces hasta que las secuencias codificantes de XTEN resultantes hayan alcanzado la organización de la secuencia y la longitud deseada, proporcionando los genes codificantes del XTEN. Como se apreciará, una biblioteca de polinucleótidos que codifica 12 aminoácidos se puede dimerizar en una biblioteca de polinucleótidos que codifica 36 aminoácidos. A su vez, la biblioteca de polinucleótidos que codifica 36 aminoácidos se puede dimerizar en serie en una biblioteca que contiene longitudes sucesivamente más largas de polinucleótidos que codifican secuencias XTEN. En algunas realizaciones, se pueden ensamblar bibliotecas de polinucleótidos que codifican aminoácidos que se limitan a familias de secuencias XTEN específicas; por ejemplo, las secuencias AD, AE, AF, AG, AM, AQ, BC o BD de la tabla 3. En otras realizaciones, las bibliotecas pueden comprender secuencias que codifican dos o más de las secuencias de la familia de motivos de la tabla 3. Los nombres y las secuencias de las secuencias polinucleotídicas representativas y no limitativas de bibliotecas que codifican 36meros se presentan en las tablas 11-14, y los métodos usados para crearlas se describen más detalladamente en los ejemplos. Las bibliotecas se pueden utilizar, a su vez, para la dimerización o ligación en serie para conseguir bibliotecas de secuencias polinucleotídicas que codifican secuencias XTEN, por ejemplo, de 72, 144, 288, 576, 864, 912, 923, 1296 aminoácidos, o hasta una longitud total de aproximadamente 3000 aminoácidos, así como longitudes intermedias. En una realización, las secuencias de la biblioteca de polinucleótidos también pueden incluir bases adicionales usadas como "islas de secuenciación", descritas más completamente a continuación.

La FIG. 8 es un diagrama de flujo esquemático de etapas representativas y no limitativas en el ensamblaje de una construcción de polinucleótido XTEN y una construcción de polinucleótido de proteína de fusión de unión utilizado en las realizaciones de XTEN, proteína de fusión de unión y BFP-D de la invención. Los oligonucleótidos individuales 501 pueden hibridarse en los motivos de secuencia 502, tales como un motivo de 12 aminoácidos ("12-mero"), que posteriormente se ligan con un oligo que contiene los sitios de restricción BbsI y KpnI 503. Los motivos de secuencia adicionales de una biblioteca se hibridan en el 12-mero hasta que se logra la longitud deseada del gen XTEN 504. El gen XTEN se clona en un vector de relleno. El vector puede codificar opcionalmente una secuencia Flag 506 seguida de una secuencia de relleno que está flanqueada por los sitios BsaI, BbsI y KpnI 507 y, en este caso, un único gen de resto de direccionamiento (que codifica anti-Her2 en este ejemplo) 508, lo que da como resultado que el gen codifica una proteína de fusión de unión que comprende un único resto de direccionamiento **500**. En la tabla 10 se proporciona una lista no exhaustiva de los nombres y secuencias XTEN para polinucleótidos que codifican XTEN y secuencias precursoras.

Tabla 10: Secuencias de ADN de XTEN **y secuencias precursoras**

Nombre del XTEN	Secuencia de ADN	SEQ ID NO:
AE144	GGTAGCGAACCGGCAACTTCCGGCTCTGAAACCCCAGGTACTTCTGAAAGCGCTA CTCCTGAGTCTGGCCCAGGTAGCGAACCTGCTACCTCTGGCTCTGAAACCCCAGG TAGCCCGGCAGGCTCTCCGACTTCCACCGAGGAAGGTACCTCTACTGAACCTTCT GAGGGTAGCGCTCCAGGTAGCGAACCGGCAACCTCTGGCTCTGAAACCCCAGGTA GCGAACCTGCTACCTCCGGCTCTGAAACTCCAGGTAGCGAACCGGCTACTTCCGG TTCTGAAACTCCAGGTACCTCTACCGAACCTTCCGAAGGCAGCGCACCACGTACT TCTGAAAGCGCAACCCCTGAATCCGGTCCAGGTAGCGAACCGGCTACTTCTGGCT CTGAGACTCCAGGTACTTCTACCGAACCGTCCGAAGGTAGCGCACCA	202
AF144	GGTACTTCTACTCCGGAAAGCGGTTCCGCATCTCCAGGTACTTCTCCTAGCGGTGA ATCTTCTACTGCTCCAGGTACCTCTCCTAGCGGCGAATCTTCTACTGCTCCAGGTT CTACCAGCTCTACCGCTGAATCTCCTGGCCCAGGTTCTACCAGCGAATCCCCGTCT GGCACCGCACCAGGTTCTACTAGCTCTACCGCAGAATCTCCGGTCCAGGTACTT CCCCTAGCGGTGAATCTTCTACTGCTCCAGGTACCTCTACTCCGGAAAGCGGCTCC GCATCTCCAGGTTCTACTAGCTCTACTGCTGAATCTCCTGGTCCAGGTACCTCCCC TAGCGGCGAATCTTCTACTGCTCCAGGTACCTCTCCTAGCGGCGAATCTTCTACCG CTCCAGGTACCTCCCCTAGCGGTGAATCTTCTACCGCACCA	203
AE288	GGTACCTCTGAAAGCGCAACTCCTGAGTCTGGCCCAGGTAGCGAACCTGCTACCT CCGGCTCTGAGACTCCAGGTACCTCTGAAAGCGCAACCCCGGAATCTGGTCCAGG TAGCGAACCTGCAACCTCTGGCTCTGAAACCCCAGGTACCTCTGAAAGCGCTACT CCTGAATCTGGCCCAGGTACTTCTACTGAACCGTCCGAGGGCAGCGCACCAGGTA GCCCTGCTGGCTCTCCAACCTCCACCGAAGAAGGTACCTCTGAAAGCGCAACCCC TGAATCCGGCCCAGGTAGCGAACCGGCAACCTCCGGTTCTGAAACCCCAGGTACT TCTGAAAGCGCTACTCCTGAGTCCGGCCCAGGTAGCCCGGCTGGCTCTCCGACTT CCACCGAGGAAGGTAGCCCGGCTGGCTCTCCAACTTCTACTGAAGAAGGTACTTC TACCGAACCTTCCGAGGGCAGCGCACCAGGTACTTCTGAAAGCGCTACCCCTGAG TCCGGCCCAGGTACTTCTGAAAGCGCTACTCCTGAATCCGGTCCAGGTACTTCTGA AAGCGCTACCCCGGAATCTGGCCCAGGTAGCGAACCGGCTACTTCTGGTTCTGAA ACCCCAGGTAGCGAACCGGCTACCTCCGGTTCTGAAACTCCAGGTAGCCCAGCAG GCTCTCCGACTTCCACTGAGGAAGGTACTTCTACTGAACCTTCCGAAGGCAGCGC ACCAGGTACCTCTACTGAACCTTCTGAGGGCAGCGCTCCAGGTAGCGAACCTGCA ACCTCTGGCTCTGAAACCCCAGGTACCTCTGAAAGCGCTACTCCTGAATCTGGCC CAGGTACTTCTACTGAACCGTCCGAGGGCAGCGCACCA	204
AE576	GGTAGCCCGGCTGGCTCTCCTACCTCTACTGAGGAAGGTACTTCTGAAAGCGCTA CTCCTGAGTCTGGTCCAGGTACCTCTACTGAACCGTCCGAAGGTAGCGCTCCAGG TAGCCCAGCAGGCTCTCCGACTTCCACTGAGGAAGGTACTTCTACTGAACCTTCC GAAGGCAGCGCACCAGGTACCTCTACTGAACCTTCTGAGGGCAGCGCTCCAGGTA CTTCTGAAAGCGCTACCCCGGAATCTGGCCCAGCTAGCGAACCGGCTACTTCTGG TTCTGAAACCCCAGGTAGCGAACCGGCTACCTCCGGTTCTGAAACTCCAGGTAGC CCGGCAGGCTCTCCGACCTCTACTGAGGAAGGTACTTCTGAAAGCGCAACCCCGG AGTCCGGCCCAGGTACCTCTACCGAACCGTCTGAGGGCAGCGCACCAGGTACTTC TACCGAACCGTCCGAGGGTAGCGCACCAGGTAGCCCAGCAGGTTCTCCTACCTCC ACCGAGGAAGGTACTTCTACCGAACCGTCCGAGGGTAGCGCACCAGGTACCTCTA CTGAACCTTCTGAGGGCAGCGCTCCAGGTACTTCTGAAAGCGCTACCCCGGAGTC CGGTCCAGGTACTTCTACTGAACCGTCCGAAGGTAGCGCACCAGGTACTTCTGAA	205

Nombre del XTEN	Secuencia de ADN	SEQ ID NO:
	AGCGCAACCCCTGAATCCGGTCCAGGTAGCGAACCGGCTACTTCTGGCTCTGAGA CTCCAGGTACTTCTACCGAACCGTCCGAAGGTAGCGCACCAGGTACTTCTACTGA ACCGTCTGAAGGTAGCGCACCAGGTACTTCTGAAAGCGCAACCCCGGAATCCGGC CCAGGTACCTCTGAAAGCGCAACCCCGGAGTCCGGCCCAGGTAGCCCTGCTGGCT CTCCAACCTCCACCGAAGAAGGTACCTCTGAAAGCGCAACCCCTGAATCCGGCCC AGGTAGCGAACCGGCAACCTCCGGTTCTGAAACCCCAGGTACCTCTGAAAGCGCT ACTCCGGAGTCTGGCCCAGGTACCTCTACTGAACCGTCTGAGGGTAGCGCTCCAG GTACTTCTACTGAACCGTCCGAAGGTAGCGCACCAGGTACTTCTACCGAACCGTC CGAAGGCAGCGCTCCAGGTACCTCTACTGAACCTTCCGAGGGCAGCGCTCCAGGT ACCTCTACCGAACCTTCTGAAGGTAGCGCACCAGGTACTTCTACCGAACCGTCCG AGGGTAGCGCACCAGGTAGCCCAGCAGGTTCTCCTACCTCCACCGAGGAAGGTAC TTCTACCGAACCGTCCGAGGGTAGCGCACCAGGTACCTCTGAAAGCGCAACTCCT GAGTCTGGCCCAGGTAGCGAACCTGCTACCTCCGGCTCTGAGACTCCAGGTACCT CTGAAAGCGCAACCCCGGAATCTGGTCCAGGTAGCGAACCTGCAACCTCTGGCTC TGAAACCCCAGGTACCTCTGAAAGCGCTACTCCTGAATCTGGCCCAGGTACTTCT ACTGAACCGTCCGAGGGCAGCGCACCAGGTACTTCTGAAAGCGCTACTCCTGAGT CCGGCCCAGGTAGCCCGGCTGGCTCTCCGACTTCCACCGAGGAAGGTAGCCCGGC TGGCTCTCCAACTTCTACTGAAGAAGGTAGCCCGGCAGGCTCTCCGACCTCTACT GAGGAAGGTACTTCTGAAAGCGCAACCCCGGAGTCCGGCCCAGGTACCTCTACCG AACCGTCTGAGGGCAGCGCACCA	
AF576	GGTTCTACTAGCTCTACCGCTGAATCTCCTGGCCCAGGTTCCACTAGCTCTACCGC AGAATCTCCGGGCCCAGGTTCTACTAGCGAATCCCCTTCTGGTACCGCTCCAGGTT CTACTAGCTCTACCGCTGAATCTCCGGGTCCAGGTTCTACCAGCTCTACTGCAGAA TCTCCTGGCCCAGGTACTTCTACTCCGGAAAGCGGTTCCGCTTCTCCAGGTTCTAC CAGCGAATCTCCTTCTGGCACCGCTCCAGGTACCTCTCCTAGCGGCGAATCTTCTA CCGCTCCAGGTTCTACTAGCGAATCTCCTTCTGGCACTGCACCAGGTTCTACCAGC GAATCTCCTTCTGGCACCGCTCCAGGTACCTCTCCTAGCGGCGAATCTTCTACCGC TCCAGGTTCTACTAGCGAATCTCCTTCTGGCACTGCACCAGGTTCTACCAGCGAAT CTCCTTCTGGCACCGCTCCAGGTACCTCTCCTAGCGGCGAATCTTCTACCGCTCCA GGTTCTACTAGCGAATCTCCTTCTGGCACTGCACCAGGTTCTACTAGCGAATCTCC TTCTGGCACTGCACCAGGTTCTACCAGCGAATCTCCGTCTGGCACTGCACCAGGT ACCTCTACCCCTGAAAGCGGTTCCGCTTCTCCAGGTTCTACTAGCGAATCTCCTTC TGGTACCGCTCCAGGTACTTCTACCCCTGAAAGCGGCTCCGCTTCTCCAGGTTCCA CTAGCTCTACCGCTGAATCTCCGGGTCCAGGTTCTACTAGCTCTACTGCAGAATCT CCTGGCCCAGGTACCTCTACTCCGGAAAGCGGCTCTGCATCTCCAGGTACTTCTAC CCCTGAAAGCGGTTCTGCATCTCCAGGTTCTACTAGCGAATCCCCGTCTGGTACCG CACCAGGTACTTCTACCCCGGAAAGCGGCTCTGCTTCTCCAGGTACTTCTACCCCG GAAAGCGGCTCCGCATCTCCAGGTTCTACTAGCGAATCTCCTTCTGGTACCGCTCC AGGTTCTACCAGCGAATCCCCGTCTGGTACTGCTCCAGGTTCTACCAGCGAATCTC CTTCTGGTACTGCACCAGGTTCTACTAGCTCTACTGCAGAATCTCCTGGCCCAGGT ACCTCTACTCCGGAAAGCGGCTCTGCATCTCCAGGTACTTCTACCCCTGAAAGCG GTTCTGCATCTCCAGGTTCTACTAGCGAATCTCCTTCTGGCACTGCACCAGGTTCT ACCAGCGAATCTCCGTCTGGCACTGCACCAGGTACCTCTACCCCTGAAAGCGGTT CCGCTTCTCCAGGTTCTACTAGCGAATCTCCTTCTGGCACTGCACCAGGTTCTACC AGCGAATCTCCGTCTGGCACTGCACCAGGTACCTCTACCCCTGAAAGCGGTTCCG CTTCTCCAGGTACTTCTCCGAGCGGTGAATCTTCTACCGCACCAGGTTCTACTAGC TCTACCGCTGAATCTCCGGGCCCAGGTACTTCTCCGAGCGGTGAATCTTCTACTGC TCCAGGTTCCACTAGCTCTACTGCTGAATCTCCTGGCCCAGGTACTTCTACTCCGG AAAGCGGTTCCGCTTCTCCAGGTTCTACTAGCGAATCTCCGTCTGGCACCGCACCA GGTTCTACTAGCTCTACTGCAGAATCTCCTGGCCCAGGTACCTCTACTCCGGAAAG CGGCTCTGCATCTCCAGGTACTTCTACCCCTGAAAGCGGTTCTGCATCTCCA	206

Nombre del XTEN	Secuencia de ADN	SEQ ID NO:
AM875	GGTACTTCTACTGAACCGTCTGAAGGCAGCGCACCAGGTAGCGAACCGGCTACTT CCGGTTCTGAAACCCCAGGTAGCCCAGCAGGTTCTCCAACTTCTACTGAAGAAGG TTCTACCAGCTCTACCGCAGAATCTCCTGGTCCAGGTACCTCTACTCCGGAAAGCG GCTCTGCATCTCCAGGTTCTACTAGCGAATCTCCTTCTGGCACTGCACCAGGTTCT ACTAGCGAATCCCCGTCTGGTACTGCTCCAGGTACTTCTACTCCTGAAAGCGGTTC CGCTTCTCCAGGTACCTCTACTCCGGAAAGCGGTTCTGCATCTCCAGGTAGCGAA CCGGCAACCTCCGGCTCTGAAACCCCAGGTACCTCTGAAAGCGCTACTCCTGAAT	207
	CCGGCCCAGGTAGCCCGGCAGGTTCTCCGACTTCCACTGAGGAAGGTACCTCTAC TGAACCTTCTGAGGGCAGCGCTCCAGGTACTTCTGAAAGCGCTACCCCGGAGTCC GGTCCAGGTACTTCTACTGAACCGTCCGAAGGTAGCGCACCAGGTACTTCTACCG AACCGTCCGAGGGTAGCGCACCAGGTAGCCCAGCAGGTTCTCCTACCTCCACCGA GGAAGGTACTTCTACCGAACCGTCCGAGGGTAGCGCACCAGGTACTTCTACCGAA CCTTCCGAGGGCAGCGCACCAGGTACTTCTGAAAGCGCTACCCCTGAGTCCGGCC CAGGTACTTCTGAAAGCGCTACTCCTGAATCCGGTCCAGGTACCTCTACTGAACCT TCCGAAGGCAGCGCTCCAGGTACCTCTACCGAACCGTCCGAGGGCAGCGCACCAG GTACTTCTGAAAGCGCAACCCCTGAATCCGGTCCAGGTACTTCTACTGAACCTTCC GAAGGTAGCGCTCCAGGTAGCGAACCTGCTACTTCTGGTTCTGAAACCCCAGGTA GCCCGGCTGGCTCTCCGACCTCCACCGAGGAAGGTAGCTCTACCCCGTCTGGTGC TACTGGTTCTCCAGGTACTCCGGGCAGCGGTACTGCTTCTTCCTCTCCAGGTAGCT CTACCCCTTCTGGTGCTACTGGCTCTCCAGGTACCTCTACCGAACCGTCCGAGGGT AGCGCACCAGGTACCTCTACTGAACCGTCTGAGGGTAGCGCTCCAGGTAGCGAAC CGGCAACCTCCGGTTCTGAAACTCCAGGTAGCCCTGCTGGCTCTCCGACTTCTACT GAGGAAGGTAGCCCGGCTGGTTCTCCGACTTCTACTGAGGAAGGTACTTCTACCG AACCTTCCGAAGGTAGCGCTCCAGGTGCAAGCGCAAGCGGCGCGCCAAGCACGG GAGGTACTTCTGAAAGCGCTACTCCTGAGTCCGGCCCAGGTAGCCCGGCTGGCTC TCCGACTTCCACCGAGGAAGGTAGCCCGGCTGGCTCTCCAACTTCTACTGAAGAA GGTTCTACCAGCTCTACCGCTGAATCTCCTGGCCCAGGTTCTACTAGCGAATCTCC GTCTGGCACCGCACCAGGTACTTCCCCTAGCGGTGAATCTTCTACTGCACCAGGT ACCCCTGGCAGCGGTACCGCTTCTTCCTCTCCAGGTAGCTCTACCCCGTCTGGTGC TACTGGCTCTCCAGGTTCTAGCCCGTCTGCATCTACCGGTACCGGCCCAGGTAGCG AACCGGCAACCTCCGGCTCTGAAACTCCAGGTACTTCTGAAAGCGCTACTCCGGA ATCCGGCCCAGGTAGCGAACCGGCTACTTCCGGCTCTGAAACCCCAGGTTCCACC AGCTCTACTGCAGAATCTCCGGGCCCAGGTTCTACTAGCTCTACTGCAGAATCTCC GGGTCCAGGTACTTCTCCTAGCGGCGAATCTTCTACCGCTCCAGGTAGCGAACCG GCAACCTCTGGCTCTGAAACTCCAGGTAGCGAACCTGCAACCTCCGGCTCTGAAA CCCCAGGTACTTCTACTGAACCTTCTGAGGGCAGCGCACCAGGTTCTACCAGCTCT ACCGCAGAATCTCCTGGTCCAGGTACCTCTACTCCGGAAAGCGGCTCTGCATCTC CAGGTTCTACTAGCGAATCTCCTTCTGGCACTGCACCAGGTACTTCTACCGAACCG TCCGAAGGCAGCGCTCCAGGTACCTCTACTGAACCTTCCGAGGGCAGCGCTCCAG GTACCTCTACCGAACCTTCTGAAGGTAGCGCACCAGGTAGCTCTACTCCGTCTGGT GCAACCGGCTCCCCAGGTTCTAGCCCGTCTGCTTCCACTGGTACTGGCCCAGGTGC TTCCCCGGGCACCAGCTCTACTGGTTCTCCAGGTAGCGAACCTGCTACCTCCGGTT CTGAAACCCCAGGTACCTCTGAAAGCGCAACTCCGGAGTCTGGTCCAGGTAGCCC TGCAGGTTCTCCTACCTCCACTGAGGAAGGTAGCTCTACTCCGTCTGGTGCAACCG GCTCCCCAGGTTCTAGCCCGTCTGCTTCCACTGGTACTGGCCCAGGTGCTTCCCCG GGCACCAGCTCTACTGGTTCTCCAGGTACCTCTGAAAGCGCTACTCCGGAGTCTG GCCCAGGTACCTCTACTGAACCGTCTGAGGGTAGCGCTCCAGGTACTTCTACTGA ACCGTCCGAAGGTAGCGCACCA	

Nombre del XTEN	Secuencia de ADN	SEQ ID NO:
AE864	GGTAGCCCGGCTGGCTCTCCTACCTCTACTGAGGAAGGTACTTCTGAAAGCGCTA CTCCTGAGTCTGGTCCAGGTACCTCTACTGAACCGTCCGAAGGTAGCGCTCCAGG TAGCCCAGCAGGCTCTCCGACTTCCACTGAGGAAGGTACTTCTACTGAACCTTCC GAAGGCAGCGCACCAGGTACCTCTACTGAACCTTCTGAGGGCAGCGCTCCAGGTA CTTCTGAAAGCGCTACCCCGGAATCTGGCCCAGGTAGCGAACCGGCTACTTCTGG TTCTGAAACCCCAGGTAGCGAACCGGCTACCTCCGGTTCTGAAACTCCAGGTAGC CCGGCAGGCTCTCCGACCTCTACTGAGGAAGGTACTTCTGAAAGCGCAACCCCGG AGTCCGGCCCAGGTACCTCTACCGAACCGTCTGAGGGCAGCGCACCAGGTACTTC TACCGAACCGTCCGAGGGTAGCGCACCAGGTAGCCCAGCAGGTTCTCCTACCTCC ACCGAGGAAGGTACTTCTACCGAACCGTCCGAGGGTAGCGCACCAGGTACCTCTA CTGAACCTTCTGAGGGCAGCGCTCCAGGTACTTCTGAAAGCGCTACCCCGGAGTC CGGTCCAGGTACTTCTACTGAACCGTCCGAAGGTAGCGCACCAGGTACTTCTGAA AGCGCAACCCCTGAATCCGGTCCAGGTAGCGAACCGGCTACTTCTGGCTCTGAGA CTCCAGGTACTTCTACCGAACCGTCCGAAGGTAGCGCACCAGGTACTTCTACTGA ACCGTCTGAAGGTAGCGCACCAGGTACTTCTGAAAGCGCAACCCCCGGAATCCGGC CCAGGTACCTCTGAAAGCGCAACCCCGGAGTCCGGCCCAGGTAGCCCTGCTGGCT CTCCAACCTCCACCGAAGAAGGTACCTCTGAAAGCGCAACCCCTGAATCCGGCCC	208
	AGGTAGCGAACCGGCAACCTCCGGTTCTGAAACCCCAGGTACCTCTGAAAGCGCT ACTCCCGAGTCTGGCCCAGGTACCTCTACTGAACCGTCTGAGGGTAGCGCTCCAG GTACTTCTACTGAACCGTCCGAAGGTAGCGCACCAGGTACTTCTACCGAACCGTC CGAAGGCAGCGCTCCAGGTACCTCTACTGAACCTTCCGAGGGCAGCGCTCCAGGT ACCTCTACCGAACCTTCTGAAGGTAGCGCACCAGGTACTTCTACCGAACCGTCCG AGGGTAGCGCACCAGGTAGCCCAGCAGGTTCTCCTACCTCCACCGAGGAAGGTAC TTCTACCGAACCGTCCGAGGGTAGCGCACCAGGTACCTCTGAAAGCGCAACTCCT GAGTCTGGCCCAGGTAGCGAACCTGCTACCTCCGGCTCTGAGACTCCAGGTACCT CTGAAAGCGCAACCCCCGGAATCTGGTCCAGGTAGCGAACCTGCAACCTCTGGCTC TGAAACCCCAGGTACCTCTGAAAGCGCTACTCCTGAATCTGGCCCAGGTACTTCT ACTGAACCGTCCGAGGGCAGCGCACCAGGTACTTCTGAAAGCGCTACTCCTGAGT CCGGCCCAGGTAGCCCGGCTGGCTCTCCGACTTCCACCGAGGAAGGTAGCCCGGC TGGCTCTCCAACTTCTACTGAAGAAGGTAGCCCGGCAGGCTCTCCGACCTCTACT GAGGAAGGTACTTCTGAAAGCGCAACCCCGGAGTCCGGCCCAGGTACCTCTACCG AACCGTCTGAGGGCAGCGCACCAGGTACCTCTGAAAGCGCAACTCCTGAGTCTGG CCCAGGTAGCGAACCTGCTACCTCCGGCTCTGAGACTCCAGGTACCTCTGAAAGC GCAACCCCGGAATCTGGTCCAGGTAGCGAACCTGCAACCTCTGGCTCTGAAACCC CAGGTACCTCTGAAAGCGCTACTCCTGAATCTGGCCCAGGTACTTCTACTGAACC GTCCGAGGGCAGCGCACCAGGTAGCCCTGCTGGCTCTCCAACCTCCACCGAAGAA GGTACCTCTGAAAGCGCAACCCCTGAATCCGGCCCAGGTAGCGAACCGGCAACCT CCGGTTCTGAAACCCCAGGTACTTCTGAAAGCGCTACTCCTGAGTCCGGCCCAGG TAGCCCGGCTGGCTCTCCGACTTCCACCGAGGAAGGTAGCCCGGCTGGCTCTCCA ACTTCTACTGAAGAAGGTACTTCTACCGAACCTTCCGAGGCCAGCGCACCAGGTA CTTCTGAAAGCGCTACCCCTGAGTCCGGCCCAGGTACTTCTGAAAGCGCTACTCCT GAATCCGGTCCAGGTACTTCTGAAAGCGCTACCCCGGAATCTGGCCCAGGTAGCG AACCGGCTACTTCTGGTTCTGAAACCCCAGGTAGCGAACCGGCTACCTCCGGTTC TGAAACTCCAGGTAGCCCAGCAGGCTCTCCGACTTCCACTGAGGAAGGTACTTCT ACTGAACCTTCCGAAGGCAGCGCACCAGGTACCTCTACTGAACCTTCTGAGGGCA GCGCTCCAGGTAGCGAACCTGCAACCTCTGGCTCTGAAACCCCAGGTACCTCTGA AAGCGCTACTCCTGAATCTGGCCCAGGTACTTCTACTGAACCGTCCGAGGGCAGC GCACCA	

Nombre del XTEN	Secuencia de ADN	SEQ ID NO:
AF864	GGTTCTACCAGCGAATCTCCTTCTGGCACCGCTCCAGGTACCTCTCCTAGCGGCGA ATCTTCTACCGCTCCAGGTTCTACTAGCGAATCTCCTTCTGGCACTGCACCAGGTT CTACTAGCGAATCCCCGTCTGGTACTGCTCCAGGTACTTCTACTCCTGAAAGCGGT TCCGCTTCTCCAGGTACCTCTACTCCGGAAAGCGGTTCTGCATCTCCAGGTTCTAC CAGCGAATCTCCTTCTGGCACCGCTCCAGGTTCTACTAGCGAATCCCCGTCTGGTA CCGCACCAGGTACTTCTCCTAGCGGCGAATCTTCTACCGCACCAGGTTCTACTAGC GAATCTCCGTCTGGCACTGCTCCAGGTACTTCTCCTAGCGGTGAATCTTCTACCGC TCCAGGTACTTCCCCTAGCGGCGAATCTTCTACCGCTCCAGGTTCTACTAGCTCTA CTGCAGAATCTCCGGGCCCAGGTACCTCTCCTAGCGGTGAATCTTCTACCGCTCCA GGTACTTCTCCGAGCGGTGAATCTTCTACCGCTCCAGGTTCTACTAGCTCTACTGC AGAATCTCCTGGCCCAGGTACCTCTACTCCGGAAAGCGGCTCTGCATCTCCAGGT ACTTCTACCCCTGAAAGCGGTTCTGCATCTCCAGGTTCTACTAGCGAATCTCCTTC TGGCACTGCACCAGGTTCTACCAGCGAATCTCCGTCTGGCACTGCACCAGGTACC TCTACCCCTGAAAGCGGTTCCGCTTCTCCAGGTTCTACCAGCTCTACCGCAGAATC TCCTGGTCCAGGTACCTCTACTCCGGAAAGCGGCTCTGCATCTCCAGGTTCTACTA GCGAATCTCCTTCTGGCACTGCACCAGGTACTTCTCCGAGCGGTGAATCTTCTACC GCACCAGGTTCTACTAGCTCTACCGCTGAATCTCCGGGCCCAGGTACTTCTCCGAG CGGTGAATCTTCTACTGCTCCAGGTACCTCTACTCCTGAAAGCGGTTCTGCATCTC CAGGTTCCACTAGCTCTACCGCAGAATCTCCGGGCCCAGGTTCTACTAGCTCTACT GCTGAATCTCCTGGCCCAGGTTCTACTAGCTCTACTGCTGAATCTCCGGGTCCAGG TTCTACCAGCTCTACTGCTGAATCTCCTGGTCCAGGTACCTCCCCGAGCGGTGAAT CTTCTACTGCACCAGGTTCTACTAGCGAATCTCCTTCTGGCACTGCACCAGGTTCT ACCAGCGAATCTCCGTCTGGCACTGCACCAGGTACCTCTACCCCTGAAAGCGGTC CXXXXXXXXXXXXTGCAAGCGCAAGCGGCGCGCCAAGCACGGGAXXXXXXXXT AGCGAATCTCCTTCTGGTACCGCTCCAGGTTCTACCAGCGAATCCCCGTCTGGTAC TGCTCCAGGTTCTACCAGCGAATCTCCTTCTGGTACTGCACCAGGTTCTACTAGCG AATCTCCTTCTGGTACCGCTCCAGGTTCTACCAGCGAATCCCCGTCTGGTACTGCT	209
	CCAGGTTCTACCAGCGAATCTCCTTCTGGTACTGCACCAGGTACTTCTACTCCGGA AAGCGGTTCCGCATCTCCAGGTACTTCTCCTAGCGGTGAATCTTCTACTGCTCCAG GTACCTCTCCTAGCGGCGAATCTTCTACTGCTCCAGGTTCTACCAGCTCTACTGCT GAATCTCCGGGTCCAGGTACTTCCCCGAGCGGTGAATCTTCTACTGCACCAGGTA CTTCTACTCCGGAAAGCGGTTCCGCTTCTCCAGGTTCTACCAGCGAATCTCCTTCT GGCACCGCTCCAGGTTCTACTAGCGAATCCCCGTCTGGTACCGCACCAGGTACTT CTCCTAGCGGCGAATCTTCTACCGCACCAGGTTCTACTAGCGAATCCCCGTCTGGT ACCGCACCAGGTACTTCTACCCCGGAAAGCGGCTCTGCTTCTCCAGGTACTTCTAC CCCGGAAAGCGGCTCCGCATCTCCAGGTTCTACTAGCGAATCTCCTTCTGGTACCG CTCCAGGTACTTCTACCCCTGAAAGCGGCTCCGCTTCTCCAGGTTCCACTAGCTCT ACCGCTGAATCTCCGGGTCCAGGTTCTACCAGCGAATCTCCTTCTGGCACCGCTCC AGGTTCTACTAGCGAATCCCCGTCTGGTACCGCACCAGGTACTTCTCCTAGCGGC GAATCTTCTACCGCACCAGGTTCTACCAGCTCTACTGCTGAATCTCCGGGTCCAGG TACTTCCCCGAGCGGTGAATCTTCTACTGCACCAGGTACTTCTACTCCGGAAAGCG GTTCCGCTTCTCCAGGTACCTCCCCTAGCGGCGAATCTTCTACTGCTCCAGGTACC TCTCCTAGCGGCGAATCTTCTACCGCTCCAGGTACCTCCCCTAGCGGTGAATCTTC TACCGCACCAGGTTCTACTAGCTCTACTGCTGAATCTCCGGGTCCAGGTTCTACCA GCTCTACTGCTGAATCTCCTGGTCCAGGTACCTCCCCGAGCGGTGAATCTTCTACT GCACCAGGTTCTAGCCCTTCTGCTTCCACCGGTACCGGCCCAGGTAGCTCTACTCC GTCTGGTGCAACTGGCTCTCCAGGTAGCTCTACTCCGTCTGGTGCAACCGGCTCCC CA	
	XXXX se insertó en dos áreas donde no hay disponible información de la secuencia.	

Nombre del XTEN	Secuencia de ADN	SEQ ID NO:
AG864	GGTGCTTCCCCGGGCACCAGCTCTACTGGTTCTCCAGGTTCTAGCCCGTCTGCTTC TACTGGTACTGGTCCAGGTTCTAGCCCTTCTGCTTCCACTGGTACTGGTCCAGGTA CCCCGGGTAGCGGTACCGCTTCTTCTTCTCCAGGTAGCTCTACTCCGTCTGGTGCT ACCGGCTCTCCAGGTTCTAACCCTTCTGCATCCACCGGTACCGGCCCAGGTGCTTC TCCGGGCACCAGCTCTACTGGTTCTCCAGGTACCCCGGGCAGCGGTACCGCATCT TCTTCTCCAGGTAGCTCTACTCCTTCTGGTGCAACTGGTTCTCCAGGTACTCCTGG CAGCGGTACCGCTTCTTCTTCTCCAGGTGCTTCTCCTGGTACTAGCTCTACTGGTTC TCCAGGTGCTTCTCCGGGCACTAGCTCTACTGGTTCTCCAGGTACCCCGGGTAGCG GTACTGCTTCTTCCTCTCCAGGTAGCTCTACCCCTTCTGGTGCAACCGGCTCTCCA GGTGCTTCTCCGGGCACCAGCTCTACCGGTTCTCCAGGTACCCCGGGTAGCGGTA CCGCTTCTTCTTCTCCAGGTAGCTCTACTCCGTCTGGTGCTACCGGCTCTCCAGGTT CTAACCCTTCTGCATCCACCGGTACCGGCCCAGGTTCTAGCCCTTCTGCTTCCACC GGTACTGGCCCAGGTAGCTCTACCCCTTCTGGTGCTACCGGCTCCCCAGGTAGCTC TACTCCTTCTGGTGCAACTGGCTCTCCAGGTGCATCTCCGGGCACTAGCTCTACTG GTTCTCCAGGTGCATCCCCTGGCACTAGCTCTACTGGTTCTCCAGGTGCTTCTCCT GGTACCAGCTCTACTGGTTCTCCAGGTACTCCTGGCAGCGGTACCGCTTCTTCTTC TCCAGGTGCTTCTCCTGGTACTAGCTCTACTGGTTCTCCAGGTGCTTCTCCGGGCA CTAGCTCTACTGGTTCTCCAGGTGCTTCCCCGGGCACTAGCTCTACCGGTTCTCCA GGTTCTAGCCCTTCTGCATCTACTGGTACTGGCCCAGGTACTCCGGGCAGCGGTAC TGCTTCTTCCTCTCCAGGTGCATCTCCGGGCACTAGCTCTACTGGTTCTCCAGGTG CATCCCCTGGCACTAGCTCTACTGGTTCTCCAGGTGCTTCTCCTGGTACCAGCTCT ACTGGTTCTCCAGGTAGCTCTACTCCGTCTGGTGCAACCGGTTCCCCAGGTAGCTC TACTCCTTCTGGTGCTACTGGCTCCCCAGGTGCATCCCCTGGCACCAGCTCTACCG GTTCTCCAGGTACCCCGGGCAGCGGTACCGCATCTTCCTCTCCAGGTAGCTCTACC CCGTCTGGTGCTACCGGTTCCCCAGGTAGCTCTACCCCGTCTGGTGCAACCGGCTC CCCAGGTAGCTCTACTCCGTCTGGTGCAACCGGCTCCCCAGGTTCTAGCCCGTCTG CTTCCACTGGTACTGGCCCAGGTGCTTCCCCGGGCACCAGCTCTACTGGTTCTCCA GGTGCATCCCCGGGTACCAGCTCTACCGGTTCTCCAGGTACTCCTGGCAGCGGTA CTGCATCTTCCTCTCCAGGTGCTTCTCCGGGCACCAGCTCTACTGGTTCTCCAGGT GCATCTCCGGGCACTAGCTCTACTGGTTCTCCAGGTGCATCCCCTGGCACTAGCTC TACTGGTTCTCCAGGTGCTTCTCCTGGTACCAGCTCTACTGGTTCTCCAGGTACCC CTGGTAGCGGTACTGCTTCTTCCTCTCCAGGTAGCTCTACTCCGTCTGGTGCTACC GGTTCTCCAGGTACCCCGGGTAGCGGTACCGCATCTTCTTCTCCAGGTAGCTCTAC CCCGTCTGGTGCTACTGGTTCTCCAGGTACTCCGGGCAGCGGTACTGCTTCTTCCT CTCCAGGTAGCTCTACCCCTTCTGGTGCTACTGGCTCTCCAGGTAGCTCTACCCCG	210
	TCTGGTGCTACTGGCTCCCCAGGTTCTAGCCCTTCTGCATCCACCGGTACCGGTCC AGGTTCTAGCCCGTCTGCATCTACTGGTACTGGTCCAGGTGCATCCCCGGGCACTA GCTCTACCGGTTCTCCAGGTACTCCTGGTAGCGGTACTGCTTCTTCTTCTCCAGGT AGCTCTACTCCTTCTGGTGCTACTGGTTCTCCAGGTTCTAGCCCTTCTGCATCCACC GGTACCGGCCCAGGTTCTAGCCCGTCTGCTTCTACCGGTACTGGTCCAGGTGCTTC TCCGGGTACTAGCTCTACTGGTTCTCCAGGTGCATCTCCTGGTACTAGCTCTACTG GTTCTCCAGGTAGCTCTACTCCGTCTGGTGCAACCGGCTCTCCAGGTTCTAGCCCT TCTGCATCTACCGGTACTGGTCCAGGTGCATCCCCTGGTACCAGCTCTACCGGTTC TCCAGGTTCTAGCCCTTCTGCTTCTACCGGTACCGGTCCAGGTACCCCTGGCAGCG GTACCGCATCTTCCTCTCCAGGTAGCTCTACTCCGTCTGGTGCAACCGGTTCCCCA GGTAGCTCTACTCCTTCTGGTGCTACTGGCTCCCCAGGTGCATCCCCTGGCACCAG CTCTACCGGTTCTCCA	

Nombre del XTEN	Secuencia de ADN	SEQ ID NO:
AM923	ATGGCTGAACCTGCTGGCTCTCCAACCTCCACTGAGGAAGGTGCATCCCCGGGCA CCAGCTCTACCGGTTCTCCAGGTAGCTCTACCCCGTCTGGTGCTACCGGCTCTCCA GGTAGCTCTACCCCGTCTGGTGCTACTGGCTCTCCAGGTACTTCTACTGAACCGTC TGAAGGCAGCGCACCAGGTAGCGAACCGGCTACTTCCGGTTCTGAAACCCCAGGT AGCCCAGCAGGTTCTCCAACTTCTACTGAAGAAGGTTCTACCAGCTCTACCGCAG AATCTCCTGGTCCAGGTACCTCTACTCCGGAAAGCGGCTCTGCATCTCCAGGTTCT ACTAGCGAATCTCCTTCTGGCACTGCACCAGGTTCTACTAGCGAATCCCCGTCTGG TACTGCTCCAGGTACTTCTACTCCTGAAAGCGGTTCCGCTTCTCCAGGTACCTCTA CTCCGGAAAGCGGTTCTGCATCTCCAGGTAGCGAACCGGCAACCTCCGGCTCTGA AACCCCAGGTACCTCTGAAAGCGCTACTCCTGAATCCGGCCCAGGTAGCCCGGCA GGTTCTCCGACTTCCACTGAGGAAGGTACCTCTACTGAACCTTCTGAGGGCAGCG CTCCAGGTACTTCTGAAAGCGCTACCCCGGAGTCCGGTCCAGGTACTTCTACTGA ACCGTCCGAAGGTAGCGCACCAGGTACTTCTACCGAACCGTCCGAGGGTAGCGCA CCAGGTAGCCCAGCAGGTTCTCCTACCTCCACCGAGGAAGGTACTTCTACCGAAC CGTCCGAGGGTAGCGCACCAGGTACTTCTACCGAACCTTCCGAGGGCAGCGCACC AGGTACTTCTGAAAGCGCTACCCCTGAGTCCGGCCCAGGTACTTCTGAAAGCGCT ACTCCTGAATCCGGTCCAGGTACCTCTACTGAACCTTCCGAAGGCAGCGCTCCAG GTACCTCTACCGAACCGTCCGAGGGCAGCGCACCAGGTACTTCTGAAAGCGCAAC CCCTGAATCCGGTCCAGGTACTTCTACTGAACCTTCCGAAGGTAGCGCTCCAGGT AGCGAACCTGCTACTTCTGGTTCTGAAACCCCAGGTAGCCCGGCTGGCTCTCCGA CCTCCACCGAGGAAGGTAGCTCTACCCCGTCTGGTGCTACTGGTTCTCCAGGTACT CCGGGCAGCGGTACTGCTTCTTCCTCTCCAGGTAGCTCTACCCCTTCTGGTGCTAC TGGCTCTCCAGGTACCTCTACCGAACCGTCCGAGGGTAGCGCACCAGGTACCTCT ACTGAACCGTCTGAGGGTAGCGCTCCAGGTAGCGAACCGGCAACCTCCGGTTCTG AAACTCCAGGTAGCCCTGCTGGCTCTCCGACTTCTACTGAGGAAGGTAGCCCGGC TGGTTCTCCGACTTCTACTGAGGAAGGTACTTCTACCGAACCTTCCGAAGGTAGC GCTCCAGGTGCAAGCGCAAGCGGCGCGCCAAGCACGGGAGGTACTTCTGAAAGC GCTACTCCTGAGTCCGGCCCAGGTAGCCCGGCTGGCTCTCCGACTTCCACCGAGG AAGGTAGCCCGGCTGGCTCTCCAACTTCTACTGAAGAAGGTTCTACCAGCTCTAC CGCTGAATCTCCTGGCCCAGGTTCTACTAGCGAATCTCCGTCTGGCACCGCACCA GGTACTTCCCCTAGCGGTGAATCTTCTACTGCACCAGGTACCCCTGGCAGCGGTA CCGCTTCTTCCTCTCCAGGTAGCTCTACCCCGTCTGGTGCTACTGGCTCTCCAGGT TCTAGCCCGTCTGCATCTACCGGTACCGGCCCAGGTAGCGAACCGGCAACCTCCG GCTCTGAAACTCCAGGTACTTCTGAAAGCGCTACTCCGGAATCCGGCCCAGGTAG CGAACCGGCTACTTCCGGCTCTGAAACCCCAGGTTCCACCAGCTCTACTGCAGAA TCTCCGGGCCCAGGTTCTACTAGCTCTACTGCAGAATCTCCGGGTCCAGGTACTTC TCCTAGCGGCGAATCTTCTACCGCTCCAGGTAGCGAACCGGCAACCTCTGGCTCT GAAACTCCAGGTAGCGAACCTGCAACCTCCGGCTCTGAAACCCCAGGTACTTCTA CTGAACCTTCTGAGGGCAGCGCACCAGGTTCTACCAGCTCTACCGCAGAATCTCC TGGTCCAGGTACCTCTACTCCGGAAAGCGGCTCTGCATCTCCAGGTTCTACTAGCG AATCTCCTTCTGGCACTGCACCAGGTACTTCTACCGAACCGTCCGAAGGCAGCGC TCCAGGTACCTCTACTGAACCTTCCGAGGGCAGCGCTCCAGGTACCTCTACCGAA CCTTCTGAAGGTAGCGCACCAGGTAGCTCTACTCCGTCTGGTGCAACCGGCTCCC CAGGTTCTAGCCCGTCTGCTTCCACTGGTACTGGCCCAGGTGCTTCCCCGGGCACC AGCTCTACTGGTTCTCCAGGTAGCGAACCTGCTACCTCCGGTTCTGAAACCCCAG GTACCTCTGAAAGCGCAACTCCGGAGTCTGGTCCAGGTAGCCCTGCAGGTTCTCC	211
	TACCTCCACTGAGGAAGGTAGCTCTACTCCGTCTGGTGCAACCGGCTCCCCAGGT TCTAGCCCGTCTGCTTCCACTGGTACTGGCCCAGGTGCTTCCCCGGGCACCAGCTC TACTGGTTCTCCAGGTAGCGAACCTGCTACCTCCGGAGTCTGGCCCAGGTACCT CTACTGAACCGTCTGAGGGTAGCGCTCCAGGTACTTCTACTGAACCGTCCGAAGG TAGCGCACCA	

Nombre del XTEN	Secuencia de ADN	SEQ ID NO:
AE912	ATGGCTGAACCTGCTGGCTCTCCAACCTCCACTGAGGAAGGTACCCCGGGTAGCG GTACTGCTTCTTCCTCTCCAGGTAGCTCTACCCCTTCTGGTGCAACCGGCTCTCCA GGTGCTTCTCCGGGCACCAGCTCTACCGGTTCTCCAGGTAGCCCGGCTGGCTCTCC TACCTCTACTGAGGAAGGTACTTCTGAAAGCGCTACTCCTGAGTCTGGTCCAGGT ACCTCTACTGAACCGTCCGAAGGTAGCGCTCCAGGTAGCCCAGCAGGCTCTCCGA CTTCCACTGAGGAAGGTACTTCTACTGAACCTTCCGAAGGCAGCGCACCAGGTAC CTCTACTGAACCTTCTGAGGGCAGCGCTCCAGGTACTTCTGAAAGCGCTACCCCG GAATCTGGCCCAGGTAGCGAACCGGCTACTTCTGGTTCTGAAACCCCAGGTAGCG AACCGGCTACCTCCGGTTCTGAAACTCCAGGTAGCCCGGCAGGCTCTCCGACCTC TACTGAGGAAGGTACTTCTGAAAGCGCAACCCCGGAGTCCGGCCCAGGTACCTCT ACCGAACCGTCTGAGGGCAGCGCACCAGGTACTTCTACCGAACCGTCCGAGGGTA GCGCACCAGGTAGCCCAGCAGGTTCTCCTACCTCCACCGAGGAAGGTACTTCTAC CGAACCGTCCGAGGGTAGCGCACCAGGTACCTCTACTGAACCTTCTGAGGGCAGC GCTCCAGGTACTTCTGAAAGCGCTACCCCGGAGTCCGGTCCAGGTACTTCTACTG AACCGTCCGAAGGTAGCGCACCAGGTACTTCTGAAAGCGCAACCCCTGAATCCGG TCCAGGTAGCGAACCGGCTACTTCTGGCTCTGAGACTCCAGGTACTTCTACCGAA CCGTCCGAAGGTAGCGCACCAGGTACTTCTACTGAACCGTCTGAAGGTAGCGCAC CAGGTACTTCTGAAAGCGCAACCCCCGGAATCCGGCCCAGGTACCTCTGAAAGCGC AACCCCGGAGTCCGGCCCAGGTAGCCCTGCTGGCTCTCCAACCTCCACCGAAGAA GGTACCTCTGAAAGCGCAACCCCTGAATCCGGCCCAGGTAGCGAACCGGCAACCT CCGGTTCTGAAACCCCAGGTACCTCTGAAACCGCTACTCCGGAGTCTGGCCCAGG TACCTCTACTGAACCGTCTGAGGGTAGCGCTCCAGGTACTTCTACTGAACCGTCCG AAGGTAGCGCACCAGGTACTTCTACCGAACCGTCCGAAGGCAGCGCTCCAGGTAC CTCTACTGAACCTTCCGAGGGCAGCGCTCCAGGTACCTCTACCGAACCTTCTGAA GGTAGCGCACCAGGTACTTCTACCGAACCGTCCGAGGGTAGCGCACCAGGTAGCC CAGCAGGTTCTCCTACCTCCACCGAGGAAGGTACTTCTACCGAACCGTCCGAGGG TAGCGCACCAGGTACCTCTGAAAGCGCAACTCCTGAGTCTGGCCCAGGTAGCGAA CCTGCTACCTCCGGCTCTGAGACTCCAGGTACCTCTGAAAGCGCAACCCCCGGAAT CTGGTCCAGGTAGCGAACCTGCAACCTCTGGCTCTGAAACCCCAGGTACCTCTGA AAGCGCTACTCCTGAATCTGGCCCAGGTACTTCTACTGAACCGTCCGAGGGCAGC GCACCAGGTACTTCTGAAAGCGCTACTCCTGAGTCCGGCCCAGGTAGCCCGGCTG GCTCTCCGACTTCCACCGAGGAAGGTAGCCCGGCTGGCTCTCCAACTTCTACTGA AGAAGGTAGCCCGGCAGGCTCTCCGACCTCTACTGAGGAAGGTACTTCTGAAAGC GCAACCCCGGAGTCCGGCCCAGGTACCTCTACCGAACCGTCTGAGGGCAGCGCAC CAGGTACCTCTGAAAGCGCAACTCCTGAGTCTGGCCCAGGTAGCGAACCTGCTAC CTCCGGCTCTGAGACTCCAGGTACCTCTGAAAGCGCAACCCCGGAATCTGGTCCA GGTAGCGAACCTGCAACCTCTGGCTCTGAAACCCCAGGTACCTCTGAAAGCGCTA CTCCTGAATCTGGCCCAGGTACTTCTACTGAACCGTCCGAGGGCAGCGCACCAGG TAGCCCTGCTGGCTCTCCAACCTCCACCGAAGAAGGTACCTCTGAAAGCGCAACC CCTGAATCCGGCCCAGGTAGCGAACCGGCAACCTCCGGTTCTGAAACCCCAGGTA CTTCTGAAAGCGCTACTCCTGAGTCCGGCCCAGGTAGCCCGGCTGGCTCTCCGAC TTCCACCGAGGAAGGTAGCCCGGCTGGCTCTCCAACTTCTACTGAAGAAGGTACT TCTACCGAACCTTCCGAGGGCAGCGCACCAGGTACTTCTGAAAGCGCTACCCCTG AGTCCGGCCCAGGTACTTCTGAAAGCGCTACTCCTGAATCCGGTCCAGGTACTTCT GAAAGCGCTACCCCGGAATCTGGCCCAGGTAGCGAACCGGCTACTTCTGGTTCTG AAACCCCAGGTAGCGAACCGGCTACCTCCGGTTCTGAAACTCCAGGTAGCCCAGC AGGCTCTCCGACTTCCACTGAGGAAGGTACTTCTACTGAACCTTCCGAAGGCAGC GCACCAGGTACCTCTACTGAACCTTCTGAGGGCAGCGCTCCAGGTAGCGAACCTG CAACCTCTGGCTCTGAAACCCCAGGTACCTCTGAAAGCGCTACTCCTGAATCTGG CCCAGGTACTTCTACTGAACCGTCCGAGGGCAGCGCACCA	212
AM1296	GGTACTTCTACTGAACCGTCTGAAGGCAGCGCACCAGGTAGCGAACCGGCTACTT CCGGTTCTGAAACCCCAGGTAGCCCAGCAGGTTCTCCAACTTCTACTGAAGAAGG TTCTACCAGCTCTACCGCAGAATCTCCTGGTCCAGGTACCTCTACTCCGGAAAGCG	213

Nombre del XTEN	Secuencia de ADN	SEQ ID NO:
	GCTCTGCATCTCCAGGTTCTACTAGCGAATCTCCTTCTGGCACTGCACCAGGTTCT ACTAGCGAATCCCCGTCTGGTACTGCTCCAGGTACTTCTACTCCTGAAAGCGGTTC CGCTTCTCCAGGTACCTCTACTCCGGAAAGCGGTTCTGCATCTCCAGGTAGCGAA CCGGCAACCTCCGGCTCTGAAACCCCAGGTACCTCTGAAAGCGCTACTCCTGAAT CCGGCCCAGGTAGCCCGGCAGGTTCTCCGACTTCCACTGAGGAAGGTACCTCTAC TGAACCTTCTGAGGGCAGCGCTCCAGGTACTTCTGAAAGCGCTACCCCGGAGTCC GGTCCAGGTACTTCTACTGAACCGTCCGAAGGTAGCGCACCAGGTACTTCTACCG AACCGTCCGAGGGTAGCGCACCAGGTAGCCCAGCAGGTTCTCCTACCTCCACCGA GGAAGGTACTTCTACCGAACCGTCCGAGGGTAGCGCACCAGGTACTTCTACCGAA CCTTCCGAGGGCAGCGCACCAGGTACTTCTGAAAGCGCTACCCCTGAGTCCGGCC CAGGTACTTCTGAAAGCGCTACTCCTGAATCCGGTCCAGGTACCTCTACTGAACCT TCCGAAGGCAGCGCTCCAGGTACCTCTACCGAACCGTCCGAGGGCAGCGCACCAG GTACTTCTGAAAGCGCAACCCCTGAATCCGGTCCAGGTACTTCTACTGAACCTTCC GAAGGTAGCGCTCCAGGTAGCGAACCTGCTACTTCTGGTTCTGAAACCCCAGGTA GCCCGGCTGGCTCTCCGACCTCCACCGAGGAAGGTAGCTCTACCCCGTCTGGTGC TACTGGTTCTCCAGGTACTCCGGGCAGCGGTACTGCTTCTTCCTCTCCAGGTAGCT CTACCCCTTCTGGTGCTACTGGCTCTCCAGGTACCTCTACCGAACCGTCCGAGGGT AGCGCACCAGGTACCTCTACTGAACCGTCTGAGGGTAGCGCTCCAGGTAGCGAAC CGGCAACCTCCGGTTCTGAAACTCCAGGTAGCCCTGCTGGCTCTCCGACTTCTACT GAGGAAGGTAGCCCGGCTGGTTCTCCGACTTCTACTGAGGAAGGTACTTCTACCG AACCTTCCGAAGGTAGCGCTCCAGGTCCAGAACCAACGGGGCCGGCCCCAAGCG GAGGTAGCGAACCGGCAACCTCCGGCTCTGAAACCCCAGGTACCTCTGAAAGCGC TACTCCTGAATCCGGCCCAGGTAGCCCGGCAGGTTCTCCGACTTCCACTGAGGAA GGTACTTCTGAAAGCGCTACTCCTGAGTCCGGCCCAGGTAGCCCGGCTGGCTCTC CGACTTCCACCGAGGAAGGTAGCCCGGCTGGCTCTCCAACTTCTACTGAAGAAGG TACTTCTGAAAGCGCTACTCCTGAGTCCGGCCCAGGTAGCCCGGCTGGCTCTCCG ACTTCCACCGAGGAAGGTAGCCCGGCTGGCTCTCCAACTTCTACTGAAGAAGGTT CTACCAGCTCTACCGCTGAATCTCCTGGCCCAGGTTCTACTAGCGAATCTCCGTCT GGCACCGCACCAGGTACTTCCCCTAGCGGTGAATCTTCTACTGCACCAGGTTCTAC CAGCGAATCTCCTTCTGGCACCGCTCCAGGTTCTACTAGCGAATCCCCGTCTGGTA CCGCACCAGGTACTTCTCCTAGCGGCGAATCTTCTACCGCACCAGGTACTTCTACC GAACCTTCCGAGGGCAGCGCACCAGGTACTTCTGAAAGCGCTACCCCTGAGTCCG GCCCAGGTACTTCTGAAAGCGCTACTCCTGAATCCGGTCCAGGTAGCGAACCGGC AACCTCTGGCTCTGAAACCCCAGGTACCTCTGAAAGCGCTACTCCGGAATCTGGT CCAGGTACTTCTGAAAGCGCTACTCCGGAATCCGGTCCAGGTACCTCTACTGAAC CTTCTGAGGGCAGCGCTCCAGGTACTTCTGAAAGCGCTACCCCGGAGTCCGGTCC AGGTACTTCTACTGAACCGTCCGAAGGTAGCGCACCAGGTACCTCCCCTAGCGGC GAATCTTCTACTGCTCCAGGTACCTCTCCTAGCGGCGAATCTTCTACCGCTCCAGG TACCTCCCCTAGCGGTGAATCTTCTACCGCACCAGGTACTTCTACCGAACCGTCCG AGGGTAGCGCACCAGGTAGCCCAGCAGGTTCTCCTACCTCCACCGAGGAAGGTAC TTCTACCGAACCGTCCGAGGGTAGCGCACCAGGTTCTAGCCCTTCTGCTTCCACCG GTACCGGCCCAGGTAGCTCTACTCCGTCTGGTGCAACTGGCTCTCCAGGTAGCTCT ACTCCGTCTGGTGCAACCGGCTCCCCAGGTAGCTCTACCCCGTCTGGTGCTACCGG CTCTCCAGGTAGCTCTACCCCGTCTGGTGCAACCGGCTCCCCAGGTGCATCCCCGG GTACTAGCTCTACCGGTTCTCCAGGTGCAAGCGCAAGCGGCGCCGCCAAGCACGGG AGGTACTTCTCCGAGCGGTGAATCTTCTACCGCACCAGGTTCTACTAGCTCTACCG CTGAATCTCCGGGCCCAGGTACTTCTCCGAGCGGTGAATCTTCTACTGCTCCAGGT ACCTCTGAAAGCGCTACTCCGGAGTCTGGCCCAGGTACCTCTACTGAACCGTCTG AGGGTAGCGCTCCAGGTACTTCTACTGAACCGTCCGAAGGTAGCGCACCAGGTTC TAGCCCTTCTGCATCTACTGGTACTGGCCCAGGTAGCTCTACTCCTTCTGGTGCTA CCGGCTCTCCAGGTGCTTCTCCGGGTACTAGCTCTACCGGTTCTCCAGGTACTTCT ACTCCGGAAAGCGGTTCCGCATCTCCAGGTACTTCTCCTAGCGGTGAATCTTCTAC TGCTCCAGGTACCTCTCCTAGCGGCGAATCTTCTACTGCTCCAGGTACTTCTGAAA GCGCAACCCCTGAATCCGGTCCAGGTAGCGAACCGGCTACTTCTGGCTCTGAGAC TCCAGGTACTTCTACCGAACCGTCCGAAGGTAGCGCACCAGGTTCTACCAGCGAA TCCCCTTCTGGTACTGCTCCAGGTTCTACCAGCGAATCCCCTTCTGGCACCGCACC AGGTACTTCTACCCCTGAAAGCGGCTCCGCTTCTCCAGGTAGCCCGGCAGGCTCT CCGACCTCTACTGAGGAAGGTACTTCTGAAAGCGCAACCCCGGAGTCCGGCCCAG	

Nombre del XTEN	Secuencia de ADN	SEQ ID NO:
	GTACCTCTACCGAACCGTCTGAGGGCAGCGCACCAGGTAGCCCTGCTGGCTCTCC AACCTCCACCGAAGAAGGTACCTCTGAAAGCGCAACCCCTGAATCCGGCCCAGGT AGCGAACCGGCAACCTCCGGTTCTGAAACCCCAGGTAGCTCTACCCCGTCTGGTG CTACCGGTTCCCCAGGTGCTTCTCCTGGTACTAGCTCTACCGGTTCTCCAGGTAGC TCTACCCCGTCTGGTGCTACTGGCTCTCCAGGTTCTACTAGCGAATCCCCGTCTGG TACTGCTCCAGGTACTTCCCCTAGCGGTGAATCTTCTACTGCTCCAGGTTCTACCA GCTCTACCGCAGAATCTCCGGGTCCAGGTAGCTCTACCCCTTCTGGTGCAACCGG CTCTCCAGGTGCATCCCCGGGTACCAGCTCTACCGGTTCTCCAGGTACTCCGGGTA GCGGTACCGCTTCTTCCTCTCCAGGTAGCCCTGCTGGCTCTCCGACTTCTACTGAG GAAGGTAGCCCGGCTGGTTCTCCGACTTCTACTGAGGAAGGTACTTCTACCGAAC CTTCCGAAGGTAGCGCTCCA	
BC864	GGTACTTCCACCGAACCATCCGAACCAGGTAGCGCAGGTACTTCCACCGAACCAT CCGAACCTGGCAGCGCAGGTAGCGAACCGGCAACCTCTGGTACTGAACCATCAG GTAGCGGCGCATCCGAGCCTACCTCTACTGAACCAGGTAGCGAACCGGCTACCTC CGGTACTGAGCCATCAGGTAGCGAACCGGCAACTTCCGGTACTGAACCATCAGGT AGCGAACCGGCAACTTCCGGCACTGAACCATCAGGTAGCGGTGCATCTGAGCCGA CCTCTACTGAACCAGGTACTTCTACTGAACCATCTGAGCCGGGCAGCGCAGGTAG CGAACCAGCTACTTCTGGCACTGAACCATCAGGTACTTCTACTGAACCATCCGAA CCAGGTAGCGCAGGTAGCGAACCTGCTACCTCTGGTACTGAGCCATCAGGTAGCG AACCGGCTACCTCTGGTACTGAACCATCAGGTACTTCTACCGAACCATCCGAGCC TGGTAGCGCAGGTACTTCTACCGAACCATCCGAGCCAGGCAGCGCAGGTAGCGA ACCGGCAACCTCTGGCACTGAGCCATCAGGTAGCGAACCAGCAACTTCTGGTACT GAACCATCAGGTACTAGCGAGCCATCTACTTCCGAACCAGGTGCAGGTAGCGGCG CATCCGAACCTACTTCCACTGAACCAGGTACTAGCGAGCCATCCACCTCTGAACC AGGTGCAGGTAGCGAACCGGCAACTTCCGGCACTGAACCATCAGGTAGCGAACC GGCTACCTCTGGTACTGAACCATCAGGTACTTCTACCGAACCATCCGAGCCTGGT AGCGCAGGTACTTCTACCGAACCATCCCAGCCAGGCAGCGCAGGTAGCGGTGCAT CCGAGCCGACCTCTACTGAACCAGGTAGCGAACCAGCAACTTCTGGCACTGAGCC ATCAGGTAGCGAACCAGCTACCTCTGGTACTGAACCATCAGGTAGCGAACCGGCT ACTTCCGGCACTGAACCATCAGGTAGCGAACCAGCAACCTCCGGTACTGAACCAT CAGGTACTTCCACTGAACCATCCGAACCGGGTAGCGCAGGTAGCGAACCGGCAA CTTCCGGCACTGAACCATCAGGTAGCGGTGCATCTGAGCCGACCTCTACTGAACC AGGTACTTCTACTGAACCATCTGAGCCGGGCAGCGCAGGTAGCGAACCTGCAACC TCCGGCACTGAGCCATCAGGTAGCGGCGCATCTGAACCAACCTCTACTGAACCAG GTACTTCCACCGAACCATCTGAGCCAGGCAGCGCAGGTAGCGGCGCATCTGAACC AACCTCTACTGAACCAGGTAGCGAACCAGCAACTTCTGGTACTGAACCATCAGGT AGCGGCGCATCTGAGCCTACTTCCACTGAACCAGGTAGCGAACCGGCAACTTCCG GCACTGAACCATCAGGTAGCGGTGCATCTGAGCCGACCTCTACTGAACCAGGTAC TTCTACTGAACCATCTGAGCCGGGCAGCGCAGGTAGCGAACCGGCAACTTCCGGC ACTGAACCATCAGGTAGCGGTGCATCTGAGCCGACCTCTACTGAACCAGGTACTT CTACTGAACCATCTGAGCCGGGCAGCGCAGGTAGCGAACCAGCTACTTCTGGCAC TGAACCATCAGGTACTTCTACTGAACCATCCGAACCAGGTAGCGCAGGTAGCGAA CCTGCTACCTCTGGTACTGAGCCATCAGGTACTTCTACTGAACCATCCGAGCCGG GTAGCGCAGGTACTTCCACTGAACCATCTGAACCTGGTAGCGCAGGTACTTCCAC TGAACCATCCGAACCAGGTAGCGCAGGTACTTCTACTGAACCATCCGAGCCGGGT AGCGCAGGTACTTCCACTGAACCATCTGAACCTGGTAGCGCAGGTACTTCCACTG AACCATCCGAACCAGGTAGCGCAGGTACTAGCGAACCATCCACCTCCGAACCAG GCGCAGGTAGCGGTGCATCTGAACCGACTTCTACTGAACCAGGTACTTCCACTGA ACCATCTGAGCCAGGTAGCGCAGGTACTTCCACCGAACCATCCGAACCAGGTAGC GCAGGTACTTCCACCGAACCATCCGAACCTGGCAGCGCAGGTAGCGAACCGGCA ACCTCTGGTACTGAACCATCAGGTAGCGGTGCATCCGAGCCGACCTCTACTGAAC CAGGTAGCGAACCAGCAACTTCTGGCACTGAGCCATCAGGTAGCGAACCAGCTAC CTCTGGTACTGAACCATCAGGTAGCGAACCGGCAACCTCTGGCACTGAGCCATCA GGTAGCGAACCAGCAACTTCTGGTACTGAACCATCAGGTACTAGCGAGCCATCTA CTTCCGAACCAGGTGCAGGTAGCGAACCTGCAACCTCCGGCACTGAGCCATCAGG TAGCGGCGCATCTGAACCAACCTCTACTGAACCAGGTACTTCCACCGAACCATCT GAGCCAGGCAGCGCAGGTAGCGAACCTGCAACCTCCGGCACTGAGCCATCAGGT AGCGGCGCATCTGAACCAACCTCTACTGAACCAGGTACTTCCACCGAACCATCTG	214

Nombre del XTEN	Secuencia de ADN	SEQ ID NO:
	AGCCAGGCAGCGCA	
BD864	GGTAGCGAAACTGCTACTTCCGGCTCTGAGACTGCAGGTACTAGTGAATCCGCAA CTAGCGAATCTGGCGCAGGTAGCACTGCAGGCTCTGAGACTTCCACTGAAGCAGG TACTAGCGAGTCCGCAACCAGCGAATCCGGCGCAGGTAGCGAAACTGCTACCTCT GGCTCCGAGACTGCAGGTAGCGAAACTGCAACCTCTGGCTCTGAAACTGCAGGTA CTTCCACTGAAGCAAGTGAAGGCTCCGCATCAGGTACTTCCACCGAAGCAAGCGA AGGCTCCGCATCAGGTACTAGTGAGTCCGCAACTAGCGAATCCGGTGCAGGTAGC GAAACCGCTACCTCTGGTTCCGAAACTGCAGGTACTTCTACCGAGGCTAGCGAAG GTTCTGCATCAGGTAGCACTGCTGGTTCCGAGACTTCTACTGAAGCAGGTACTAG CGAATCTGCTACTAGCGAATCCGGCGCAGGTACTAGCGAATCCGCTACCAGCGAA TCCGGCGCAGGTAGCGAAACTGCAACCTCTGGTTCCGAGACTGCAGGTACTAGCG AGTCCGCTACTAGCGAATCTGGCGCAGGTACTTCCACTGAAGCTAGTGAAGGTTC TGCATCAGGTAGCGAAACTGCTACTTCTGGTTCCGAAACTGCAGGTAGCGAAACC GCTACCTCTGGTTCCGAAACTGCAGGTACTTCTACCGAGGCTAGCGAAGGTTCTG CATCAGGTAGCACTGCTGGTTCCGAGACTTCTACTGAAGCAGGTACTAGCGAGTC CGCTACTAGCGAATCTGGCGCAGGTACTTCCACTGAAGCTAGTGAAGGTTCTGCA TCAGGTAGCGAAACTGCTACTTCTGGTTCCGAAACTGCAGGTAGCACTGCTGGCT CCGAGACTTCTACCGAAGCAGGTAGCACTGCAGGTTCCGAAACTTCCACTGAAGC AGGTAGCGAAACTGCTACCTCTGGCTCTGAGACTGCAGGTACTAGCGAATCTGCT ACTAGCGAATCCGGCGCAGGTACTAGCGAATCCGCTACCAGCGAATCCGGCGCA GGTAGCGAAACTGCAACCTCTGGTTCCGAGACTGCAGGTACTAGCGAATCTGCTA CTAGCGAATCCGGCGCAGGTACTAGCGAATCCGCTACCAGCGAATCCGGCGCAG GTAGCGAAACTGCAACCTCTGGTTCCGAGACTGCAGGTAGCGAAACCGCTACCTC TGGTTCCGAAACTGCAGGTACTTCTACCGAGGCTAGCGAAGGTTCTGCATCAGGT AGCACTGCTGGTTCCGAGACTTCTACTGAAGCAGGTAGCGAAACTGCTACTTCCG GCTCTGAGACTGCAGGTACTAGTGAATCCGCAACTAGCGAATCTGGCGCAGGTAG CACTGCAGGCTCTGAGACTTCCACTGAAGCAGGTAGCACTGCTGGTTCCGAAACC TCTACCGAAGCAGGTAGCACTGCAGGTTCTGAAACCTCCACTGAAGCAGGTACTT CCACTGAGGCTAGTGAAGGCTCTGCATCAGGTAGCACTGCTGGTTCCGAAACCTC TACCGAAGCAGGTAGCACTGCAGGTTCTGAAACCTCCACTGAAGCAGGTACTTCC ACTGAGGCTAGTGAAGGCTCTGCATCAGGTAGCACTGCAGGTTCTGAGACTTCCA CCGAAGCAGGTAGCGAAACTGCTACTTCTGGTTCCGAAACTGCAGGTACTTCCAC TGAAGCTAGTGAAGGTTCCGCATCAGGTACTAGTGAGTCCGCAACCAGCGAATCC GGCGCAGGTAGCGAAACCGCAACCTCCGGTTCTGAAACTGCAGGTACTAGCGAAT CCGCAACCAGCGAATCTGGCGCAGGTACTAGTGAGTCCGCAACCAGCGAATCCG GCGCAGGTAGCGAAACCGCAACCTCCGGTTCTGAAACTGCAGGTACTAGCGAATC CGCAACCAGCGAATCTGGCGCAGGTAGCGAAACTGCTACTTCCGGCTCTGAGACT GCAGGTACTTCCACCGAAGCAAGCGAAGGTTCCGCATCAGGTACTTCCACCGAGG CTAGTGAAGGCTCTGCATCAGGTAGCACTGCTGGCTCCGAGACTTCTACCGAAGC AGGTAGCACTGCAGGTTCCGAAACTTCCACTGAAGCAGGTAGCGAAACTGCTACC TCTGGCTCTGAGACTGCAGGTACTAGCGAATCTGCTACTAGCGAATCCGGCGCAG GTACTAGCGAATCCGCTACCAGCGAATCCGGCGCAGGTAGCGAAACTGCAACCTC TGGTTCCGAGACTGCAGGTAGCGAAACTGCTACTTCCGGCTCCGAGACTGCAGGT AGCGAAACTGCTACTTCTGGCTCCGAAACTGCAGGTACTTCTACTGAGGCTAGTG AAGGTTCCGCATCAGGTACTAGCGAGTCCGCAACCAGCGAATCCGGCGCAGGTA GCGAAACTGCTACCTCTGGCTCCGAGACTGCAGGTAGCGAAACTGCAACCTCTGG CTCTGAAACTGCAGGTACTAGCGAATCTGCTACTAGCGAATCCGGCGCAGGTACT AGCGAATCCGCTACCAGCGAATCCGGCGCAGGTAGCGAAACTGCAACCTCTGGTT CCGAGACTGCA	215

Se puede clonar la biblioteca de genes que codifican XTEN en uno o más vectores de expresión conocidos en la técnica. Para facilitar la identificación de miembros de la biblioteca que expresan bien, se puede construir la biblioteca como fusión con una proteína indicadora. Algunos ejemplos no limitativos de genes indicadores adecuados son proteína fluorescente verde, luciferasa, fosfatasa alcalina y beta-galactosidasa. Mediante un cribado se pueden identificar secuencias XTEN cortas que pueden expresarse en alta concentración en el organismo hospedador de elección. Posteriormente, se puede generar una biblioteca de dímeros XTEN aleatorios y repetir el cribado para un

alto nivel de expresión. Posteriormente, se pueden cribar las construcciones resultantes para detectar una serie de propiedades, tales como nivel de expresión, estabilidad de la proteasa o unión al antisuero.

Un aspecto de la invención es proporcionar secuencias polinucleotídicas que codifican los componentes de la proteína de fusión, en donde la creación de la secuencia ha experimentado una optimización de codones. De particular interés es la optimización de codones con el objetivo de mejorar la expresión de las composiciones polipeptídicas y mejorar la estabilidad genética del gen codificante en los hospedadores de producción. Por ejemplo, la optimización de codones es de particular importancia para las secuencias XTEN que son ricas en glicina o que tienen secuencias de aminoácidos repetitivas. La optimización de codones se puede realizar utilizando programas informáticos (Gustafsson, C., et al (2004) Trends Biotechnol, 22: 346-53), algunos de los cuales minimizan la pausa ribosómica (Coda Genomics Inc.). En una realización, la optimización de codones se puede realizar construyendo bibliotecas de codones donde todos los miembros de la biblioteca codifican la misma secuencia de aminoácidos pero donde el uso de codones es variado. Dichas bibliotecas pueden cribarse para detectar miembros genéticamente estables y de alta expresión que sean particularmente adecuados para la producción a gran escala de productos que contienen XTEN. Al diseñar secuencias XTEN se pueden considerar varias propiedades. Se puede minimizar la repetitividad en las secuencias de ADN codificantes. Además, se puede evitar o minimizar el uso de codones que rara vez son usados por el hospedador de producción (por ejemplo, los codones de arginina AGG y AGA y un codón de leucina en *E. coli*). En el caso de *E. coli,* dos codones de glicina, GGA y GGG, rara vez se usan en proteínas muy expresadas. Por lo tanto, la optimización de codones del gen que codifica las secuencias XTEN puede ser muy deseable. Las secuencias de ADN que tienen un alto nivel de glicina tienden a tener un alto contenido de GC que puede conducir a inestabilidad o bajos niveles de expresión. Por lo tanto, cuando sea posible, se prefiere elegir codones de manera que el contenido de GC de la secuencia codificante del XTEN sea adecuado para el organismo de producción que se utilizará para fabricar el XTEN.

Opcionalmente, el gen codificante del XTEN completo puede comprender una o más islas de secuenciación. En este contexto, las islas de secuenciación son secuencias de tramo corto que son distintas de las secuencias de construcción de la biblioteca XTEN y que incluyen un sitio de restricción que no está presente o que se espera que esté presente en el gen codificante del XTEN completo. En una realización, una isla de secuenciación es la secuencia 5'-AGGTGCAAGCGCAAGCGGCGCGCCAAGCACGGGAGGT-3' (SEQ ID NO: 216). En otra realización, una isla de secuenciación es la secuencia 5'-AGGTCCAGAACCAACGGGGCCGGCCCCAAGCGGAGGT-3' (SEQ ID NO: 217).

Como alternativa, se pueden construir bibliotecas de codones donde todos los miembros de la biblioteca codifican la misma secuencia de aminoácidos pero donde el uso de codones es variado. Dichas bibliotecas pueden cribarse para detectar miembros genéticamente estables y de alta expresión que sean particularmente adecuados para la producción a gran escala de productos que contienen XTEN.

Opcionalmente, se pueden secuenciar clones de la biblioteca para eliminar aislados que contienen secuencias indeseables. La biblioteca inicial de secuencias XTEN cortas puede permitir alguna variación en la secuencia de aminoácidos. Por ejemplo, se pueden aleatorizar algunos codones de forma que un número de aminoácidos hidrófilos puedan aparecer en una posición particular.

Durante el proceso de multimerización iterativa se pueden cribar los miembros de la biblioteca resultante en busca de otras características como la solubilidad o la resistencia a las proteasas, además de una criba para la expresión de alto nivel.

En una realización de una construcción que codifica una proteína de fusión de unión, una vez que se selecciona el gen que codifica el XTEN con la longitud y propiedades deseadas, se fusiona genéticamente con los nucleótidos que codifican el extremo amínico y/o carboxílico del gen o genes del resto de direccionamiento clonándolo en la construcción adyacente y en marco con el gen que codifica el resto de direccionamiento o adyacente a una secuencia conectora. La invención proporciona diferentes permutaciones de lo anterior, dependiendo de la proteína de fusión de unión que se va a codificar. Por ejemplo, un gen que codifica una proteína de fusión de unión que comprende dos restos de direccionamiento, tales como los representados por la fórmula II, como se ha mostrado anteriormente, tendría polinucleótidos que codifican dos restos de direccionamiento, un conector, al menos un primer XTEN, y opcionalmente un segundo XTEN. La etapa de clonación de los genes del resto de direccionamiento en la construcción XTEN puede ocurrir a través de una etapa de ligación o de multimerización. Como se muestra en la FIG. 6, las construcciones que codifican proteínas de fusión de unión se pueden diseñar en diferentes configuraciones de los componentes; por ejemplo, XTEN **205,** VL **202,** VH **204** y secuencias conectoras **203** o **206.** En una realización, como se ilustra en la FIG. 6A, la construcción comprende secuencias polinucleotídicas complementarias de, o aquellas que codifican, un polipéptido monómero de componentes en el siguiente orden (de 5' a 3') VL 202, conector **203,** VH **204** y XTEN **205,** o el orden inverso. Como será evidente para los expertos en la técnica, en vista de la divulgación y de la FIG. 6, son posibles otras permutaciones o combinaciones de lo anterior.

La divulgación también engloba polinucleótidos que comprenden variantes polinucleotídicas que codifican XTEN que tienen un alto porcentaje de identidad de secuencia con (a) una secuencia polinucleotídica de la tabla 10, o (b) secuencias que son complementarias de los polinucleótidos de (a). Un polinucleótido con un alto porcentaje de identidad de secuencia es uno que tiene al menos aproximadamente un 80 % de identidad de la secuencia de ácido nucleico, alternativamente al menos aproximadamente un 81 %, alternativamente al menos aproximadamente un

82 %, alternativamente al menos aproximadamente un 83 %, alternativamente al menos aproximadamente un 84 %, alternativamente al menos aproximadamente un 85 %, alternativamente al menos aproximadamente un 86 %, alternativamente al menos aproximadamente un 87 %, alternativamente al menos aproximadamente un 88 %, alternativamente al menos aproximadamente un 89 %, alternativamente al menos aproximadamente un 90 %, alternativamente al menos aproximadamente un 91 %, alternativamente al menos aproximadamente un 92 %, alternativamente al menos aproximadamente un 93 %, alternativamente al menos aproximadamente un 94 %, alternativamente al menos aproximadamente un 95 %, alternativamente al menos aproximadamente un 96 %, alternativamente al menos aproximadamente un 97 %, alternativamente al menos aproximadamente un 98 % y alternativamente al menos aproximadamente un 99 % de identidad de la secuencia de ácido nucleico con (a) o (b) de las anteriores, o que puede hibridar con el polinucleótido diana o su complemento en condiciones rigurosas.

La homología, la similitud de secuencias o la identidad de secuencias de nucleótidos o de aminoácidos también se pueden determinar convencionalmente utilizando programas informáticos conocidos, tales como los programas de comparación por pares BestFit o Gap (GCG Wisconsin Package, Genetics Computer Group, 575 Science Drive, Madison, Wis. 53711). BestFit utiliza el algoritmo de homología local de Smith y Waterman (Advances in Applied Mathematics. 1981. 2: 482-489), para encontrar el mejor segmento de identidad o similitud entre dos secuencias. Gap realiza alineaciones globales: toda una secuencia con toda otra secuencia similar utilizando el método de Needleman y Wunsch, (Journal of Molecular Biology. 1970. 48: 443-453). Cuando se utiliza un programa de alineación de secuencias, tal como BestFit, para determinar el grado de homología de las secuencias, la similitud o la identidad, se pueden usar los ajustes predeterminados o se puede seleccionar una matriz de puntuación apropiada para optimizar las puntuaciones de identidad, de similitud o de homología.

Las secuencias de ácido nucleico que son "complementarias" son aquellas que son capaces de emparejarse por bases según las reglas de complementariedad convencionales de Watson-Crick. Como se utiliza en el presente documento, la expresión "secuencias complementarias" significa secuencias de ácidos nucleicos que son sustancialmente complementarias, como puede evaluarse mediante la misma comparación de nucleótidos expuesta anteriormente, o como se define como ser capaz de hibridar con los polinucleótidos que codifican las secuencias de la proteína de fusión de unión en condiciones rigurosas, tales como las descritas en el presente documento.

Los polinucleótidos resultantes que codifican las composiciones quiméricas de proteína de fusión de unión pueden a continuación clonarse individualmente en un vector de expresión. La secuencia de ácido nucleico se puede insertar en el vector mediante una variedad de procedimientos. En general, el ADN se inserta en un sitio o sitios de endonucleasa de restricción apropiados utilizando técnicas conocidas en la técnica. Los componentes del vector generalmente incluyen, pero sin limitación, una o más de una secuencia de señalización, un origen de replicación, uno o más genes marcadores, un elemento potenciador, un promotor y una secuencia de terminación de la transcripción. La construcción de vectores adecuados que contienen uno o más de estos componentes emplea técnicas de ligación convencionales que son conocidas por el experto. Dichas técnicas son bien conocidas en la técnica y están bien descritas en la bibliografía científica y de patentes.

Hay diferentes vectores a disposición del público. El vector puede estar, por ejemplo, en forma de un plásmido, un cósmido, una partícula vírica o un fago. Tanto los vectores de expresión como los de clonación contienen una secuencia de ácido nucleico que permite que el vector se replique en una o más células hospedadoras seleccionadas. Dichas secuencias de vector son bien conocidas para una variedad de bacterias, levaduras y virus. Los vectores de expresión útiles que se pueden utilizar incluyen, por ejemplo, segmentos de secuencias de ADN cromosómico, no cromosómico y sintético. Algunos vectores adecuados incluyen, pero sin limitación, derivados del SV40 y pcDNA y plásmidos bacterianos conocidos tales como col EI, pCRI, pBR322, PMAL-C2, pET, pGEX como se describe en Smith, et al., Gene 57: 31-40 (1988), pMB9 y derivados de los mismos, plásmidos tales como RP4, ADN de fagos tales como los numerosos derivados del fago I, tales como NM98 9, así como otro ADN de fagos tales como M13 y ADN de fagos filamentoso monocatenario; plásmidos de levadura tales como el plásmido de 2 micrómetros o derivados del plásmido 2m, así como vectores lanzadera de levadura centómeros y de integración; vectores útiles en células eucariotas tales como vectores útiles en células de insecto o de mamífero; vectores derivados de combinaciones de plásmidos y ADN de fago, tales como plásmidos que se han modificado para emplear ADN de fago o las secuencias de control de la expresión; y similares. Los requisitos son que los vectores sean replicables y viables en la célula hospedadora de elección. Según se desee, pueden usarse vectores de bajo o de alto número de copias.

Algunos promotores adecuados para usar en vectores de expresión con hospedadores procariotas incluyen los sistemas de promotores de β-lactamasa y lactosa [Chang et al., Nature, 275: 615 (1978); Goeddel et al., Nature, 281: 544 (1979)], fosfatasa alcalina, un sistema promotor del triptófano (trp) [Goeddel, Nucleic Acids Res., 8: 4057 (1980); documento EP 36.776], y promotores híbridos tales como el promotor tac [deBoer et al., Proc. Natl. ACAD. SCI. USA, 80: 21-25 (1983)]. Los promotores para usar en sistemas bacterianos también pueden contener una secuencia de Shine-Dalgarno (S.D.) unida operativamente al ADN que codifica la proteína de fusión de unión.

Por ejemplo, en un sistema de expresión de baculovirus, pueden utilizarse ambos vectores de transferencia sin fusión, tales como, pero sin limitación, pVL941 (sitio de clonación BarnHI, disponible en Summers, et al., Virology 84: 390-402 (1978)), pVL1393 (sitios de clonación BamHI, SmaI, XbaI, EcoRI, IVotI, XmaIII, BglII y PstI; Invitrogen), pVL1392 (sitio de clonación BglII, PstI, NotI, XmaIII, EcoRI, XbaII, SmaI y BamHI; Summers, et al., Virology 84: 390- 402 (1978)

e Invitrogen) y pBlueBacIII (sitio de clonación BamHI, BgIII, PstI, NcoI e HindI II, con cribado recombinante azul/blanco, Invitrogen) y vectores de transferencia de fusión tales como, pero sin limitación, pAc7 00 (sitios de clonación de BamHI y KpnI, en los que el sitio de reconocimiento de BamHI comienza con el codón de inicio; Summers, et al., Virology 84: 390-402 (1978)), pAc701 y pAc70-2 (igual que pAc700, con diferentes marcos de lectura), pAc360 [sitio de clonación BamHI a 36 pares de bases secuencia abajo de un codón de inicio de poliedrina; Invitrogen (1995)) y pBlueBacHisA, B, C (tres marcos de lectura diferentes con el sitio de clonación BamH I, Bgl II, PstI, Nco I e Hind III, un péptido aminoterminal para la purificación ProBond y cribado recombinante azul/blanco de placas; Invitrogen (220).

Los esquemas de plásmidos ilustrativos que contienen uno o más de los componentes descritos anteriormente se ilustran en la FIG. 10.

Los vectores de expresión de mamíferos pueden comprender un origen de replicación, un promotor y un potenciador adecuados, y también cualquier sitio de unión al ribosoma, sitio de poliadenilación, sitios donantes y aceptores de empalme, secuencias de terminación de la transcripción y secuencias no transcritas flanqueantes en 5' necesarios. Se pueden usar secuencias de ADN derivadas del empalme del SV40 y sitios de poliadenilación para proporcionar los elementos genéticos no transcritos requeridos. Los vectores de expresión de mamíferos contemplados para usar en la invención incluyen vectores con promotores inducibles, tales como los promotores de dihidrofolato reductasa, cualquier vector de expresión con un casete de expresión de DHFR o un vector de coamplificación de DHFR/metotrexato tal como pED (sitios de clonación de PstI, SaiI, SbaI, SmaI y EcoRI, expresando el vector tanto el gen clonado como la DHFR; Randal J. Kaufman, 1991, Randal J. Kaufman, Current Protocols in Molecular Biology, 16, 12 (1991)). Alternativamente, un vector de coamplificación de glutamina sintetasa/metionina sulfoximina, tal como pEE14 (sitios de clonación de HindIII, XbaII, SmaI, SbaI, EcoRI y SelI en los que el vector expresa la glutamina sintetasa y el gen clonado; Celltech). Se puede usar un vector que dirige la expresión episómica bajo el control del virus Epstein Barr (VEB) o el antígeno nuclear (EBNA), tal como pREP4 (sitios de clonación BamHI r SfH, XhoI, NotI, NheI, HindI II, NheI, PvuII y KpnI, promotor constitutivo del VRS-LTR, marcador de selección con higromicina; Invitrogen), pCEP4 (sitios de clonación BamHI, SfH, XhoI, NotI, NheI, HindIII, NheI, PvuII y KpnI, promotor génico temprano inmediato constitutivo del hCMV, marcador de selección con higromicina; Invitrogen), pMEP4 (sitios de clonación KpnI, PvuI, NheI, HindIII, NotI, XhoI, SfiI, BamHI, promotor inducible del gen de la metalotioneína H a, marcador de selección con higromicina, Invitrogen), pREP8 (sitios de clonación BamHI, XhoI, NotI, HindIII, NheI y KpnI, promotor del RSV-LTR, marcador de selección con histidinol; Invitrogen), pREP9 (sitios de clonación KpnI, NheI, Hind 111, NotI, Xho I, Sfi I, BamH I, promotor del RSV-LTR, marcador de selección G418; Invitrogen) y pEBVHis (promotor del RSV-LTR, marcador de selección con higromicina, péptido aminoterminal purificable mediante resina ProBond y escindido por enterocinasa; Invitrogen).

Algunos vectores de expresión de mamíferos seleccionables para usar en la invención incluyen, pero sin limitación, pRC/CMV (sitios de clonación Hind 111, BstXI, NotI, SbaI y ApaI, selección G418, Invitrogen), pRc/RSV (sitios de clonación Hind II, SpeI, BstXI, NotI, XbaI, selección G418, Invitrogen), y similares. Algunos vectores de expresión de mamíferos del virus variolovacunal (*véase*, por ejemplo, Randall J. Kaufman, Current Protocols in Molecular Biology 16,12 (Frederick M. Ausubel, et al., eds Wiley 1991) que pueden usarse en la presente invención incluyen, pero sin limitación, pSC11 (sitio de clonación SmaI, selección con TK- y beta-gal), pMJ601 (sitios de clonación Sal I, Sma I, A fII, NarI, BspMII, BamHI, ApaI, NheI, SacII, KpnI e HindIII; selección con TK- y -gal), pTKgptFIS (sitios de clonación EcoRI, PstI, SaIII, AccI, HindII, SbaI, BamHI y Hpa, selección con TK o XPRT), y similares.

Los sistemas de expresión de levaduras que también pueden utilizarse en la presente invención incluyen, pero sin limitación, el vector pYES2 sin fusión (sitios de clonación XJbaI, SphI, ShoI, NotI, GstXI, EcoRI, BstXI, BamHI, Sad, KpnI e HindIII, Invitrogen), la fusión pYESHisA, B, C (sitios de clonación XbaII, SphI, ShoI, NotI, BstXI, EcoRI, BamHI, Sad, KpnI e Hindi II, péptido aminoterminal purificado con resina ProBond y escindido con enterocinasa; Invitrogen), vectores pRS y similares.

Además, el vector de expresión que contiene la molécula polinucleotídica que codifica la proteína de unión quimérica puede incluir marcadores de selección con fármacos. Dichos marcadores ayudan en la clonación y en la selección o la identificación de los vectores que contienen las moléculas de ADN quiméricas. Por ejemplo, los genes que confieren resistencia a neomicina, puromicina, higromicina, inhibidor de dihidrofolato reductasa (DHFR), guanina fosforribosil transferasa (GPT), zeocina e histidinol son marcadores de selección útiles. Alternativamente, pueden emplearse enzimas, tales como la timidina cinasa (tk) o la cloranfenicol acetiltransferasa (CAT) del virus del herpes simple. También se pueden emplear marcadores inmunológicos. Puede emplearse cualquier marcador de selección conocido siempre que sea susceptible de ser expresado simultáneamente con el ácido nucleico que codifica un producto génico. Otros ejemplos de marcadores de selección son bien conocidos por un experto en la materia e incluyen indicadores tales como la proteína fluorescente verde mejorada (EGFP), la beta-galactosidasa (β-gal) o la cloranfenicol acetiltransferasa (CAT).

En una realización, el polinucleótido que codifica una composición de proteína de fusión de unión se puede fusionar en el extremo carboxílico con una secuencia de señalización aminoterminal apropiada para el sistema hospedador de expresión. Las secuencias de señalización normalmente se eliminan proteolíticamente de la proteína durante el proceso de translocación y secreción, generando un extremo amínico definido. Se ha descrito una amplia variedad de secuencias de señalización para la mayoría de los sistemas de expresión, incluidos los sistemas bacterianos, de

levadura, de insecto y de mamífero. A continuación en el presente documento se presenta una lista no limitativa de ejemplos preferidos para cada sistema de expresión. Las secuencias de señalización preferidas son OmpA, PhoA y DsbA para la expresión en *E. coli*. Los péptidos de señalización preferidos para la expresión en levaduras son ppL-alfa, DEX4, péptido de señalización de invertasa, péptido de señalización de fosfatasa ácida, CPY o INU1. Para la expresión en células de insecto, las secuencias de señalización preferidas son el precursor de la hormona adipocinética sexta, CP1, CP2, CP3, CP4, TPA, PAP o gp67. Para la expresión en mamífero, las secuencias de señalización preferidas son IL2L, SV40, IgG kappa e IgG lambda.

En otra realización, se puede fusionar una secuencia líder, que comprende potencialmente un dominio de proteína independiente bien expresado, con el extremo amínico de la secuencia de la proteína de fusión de unión, separada por un sitio de escisión de proteasa. Aunque puede usarse cualquier secuencia del péptido líder que no inhiba la escisión en el sitio proteolítico diseñado, las secuencias, en las realizaciones preferidas, comprenderán secuencias estables y bien expresadas de manera que la expresión y el plegamiento de la composición global no se vean afectados significativamente negativamente, y preferentemente la expresión, la solubilidad y/o la eficacia de plegamiento mejoran significativamente. En la bibliografía se ha descrito una amplia variedad de secuencias líder adecuadas. Una lista no limitativa de secuencias adecuadas incluye la proteína de unión a maltosa, el dominio de unión a celulosa, la glutatión S-transferasa, la etiqueta 6xHis (SEQ ID NO: 218), la etiqueta FLAG, la etiqueta de hemaglutinina y la proteína fluorescente verde. La secuencia líder también se puede mejorar aún más mediante la optimización de codones, especialmente en la segunda posición del codón después del codón de inicio ATG, mediante métodos bien descritos en la bibliografía y anteriormente en el presente documento.

Se conocen diferentes métodos enzimáticos *in vitro* para escindir proteínas en sitios específicos. Dichos métodos incluyen el uso de enterocinasa (DDDK) (SEQ ID NO: 219), Factor Xa (IDGR) (SEQ ID NO: 880), trombina (LVPRGS) (SEQ ID NO: 220), PreScission™ (LEVLFQGP) (SEQ ID NO: 221), proteasa TEV (EQLYFQG) (SEQ ID NO: 222), proteasa 3C (ETLFQGP) (SEQ ID NO: 223), Sortasa A (LPETG) (SEQ ID NO: 224), Granzima B (D/X, N/X, M/N o S/X), inteínas, SUMO, DAPasa (TAGZyme™), aminopeptidasa de *Aeromonas*, aminopeptidasa M y carboxipeptidasas A y B. Se divulgan métodos adicionales en Arnau, et al., Protein Expression and Purification 48: 1-13 (2006).

En otras realizaciones, puede incluirse una secuencia polinucleotídica optimizada que codifica al menos aproximadamente 20 a aproximadamente 60 aminoácidos con características de XTEN en el extremo amínico de la secuencia del XTEN para promover el inicio de la traducción para permitir la expresión de fusiones de XTEN en el extremo amínico de proteínas sin la presencia de un dominio auxiliar. En una ventaja de lo anterior, la secuencia no requiere una escisión posterior, reduciendo así el número de etapas para fabricar composiciones que contienen XTEN. Como se describe con más detalle en los ejemplos, la secuencia aminoterminal optimizada tiene atributos de una proteína no estructurada, pero puede incluir bases nucleotídicas que codifican aminoácidos seleccionados por su capacidad para promover el inicio de la traducción y una expresión mejorada. En una realización de lo anterior, el polinucleótido optimizado codifica una secuencia del XTEN con al menos aproximadamente un 90 % de identidad de secuencia con AE624. En otra realización de lo anterior, el polinucleótido optimizado codifica una secuencia del XTEN con al menos aproximadamente un 90 % de identidad de secuencia con AE912. En otra realización más de lo anterior, el polinucleótido optimizado codifica una secuencia del XTEN con al menos aproximadamente un 90 % de identidad de secuencia con AM923.

En otra realización, el sitio de proteasa de la construcción de la secuencia líder se elige de manera que sea reconocido por una proteasa *in vivo*. En esta realización, la proteína se purifica a partir del sistema de expresión mientras se conserva el líder evitando el contacto con una proteasa apropiada. A continuación, la construcción completa se inyecta en un paciente. Tras la inyección, la construcción entra en contacto con la proteasa específica para el sitio de escisión y es escindida por la proteasa. En caso de que la proteína no escindida sea sustancialmente menos activa que la forma escindida, este método tiene el efecto beneficioso de permitir unas dosis iniciales más altas al tiempo que evita la toxicidad, ya que la forma activa se genera lentamente *in vivo*. Algunos ejemplos no limitativos de proteasas *in vivo* que son útiles para esta aplicación incluyen FXIa, FXIIa, calicreína, FVIIa, FIXa, FXa, FIIa (trombina), elastasa-2, granzima B, MMP-12, MMP-13, MMP-17 o MMP-20, o mediante proteasas que no son de mamífero tales como TEV, enterocinasa, la proteasa PreScission™ (proteasa de rinovirus 3C) y sortasa A.

De esta manera, se genera una molécula de ADN quimérica que codifica una proteína de fusión de unión monómera dentro de la construcción. Opcionalmente, esta molécula de ADN quimérica puede transferirse o clonarse en otra construcción que sea un vector de expresión más apropiado. En este punto, una célula hospedadora capaz de expresar la molécula de ADN quimérica puede ser transformada con la molécula de ADN quimérica. Los vectores que contienen los segmentos de ADN de interés pueden transferirse a la célula hospedadora mediante métodos bien conocidos, dependiendo del tipo de hospedador celular. Por ejemplo, para células procariotas se utiliza habitualmente la transfección mediante cloruro de calcio, mientras que para otros hospedadores celulares puede utilizarse el tratamiento con fosfato de calcio, la lipofección o la electroporación. Otros métodos utilizados para transformar células de mamífero incluyen el uso de polibreno, fusión de protoplastos, liposomas, electroporación y microinyección. *Véase,* en general, Sambrook, *et al.,* más arriba.

La transformación puede ocurrir con o sin la utilización de un portador, tal como un vector de expresión. A continuación, la célula hospedadora transformada se cultiva en condiciones adecuadas para la expresión de la molécula de ADN quimérica que codifica la proteína de fusión de unión.

La presente invención también proporciona una célula hospedadora para la expresión de las composiciones de proteína de fusión monómeras divulgadas en el presente documento. Algunos ejemplos de células hospedadoras eucariotas adecuadas incluyen, pero sin limitación, células de mamífero, tales como células VERO, células HELA tales como ATCC n.º CCL2, estirpes de células CHO, células COS, células WI38, células BHK, células HepG2, células 3T3, células A549, células PC12, células K562, células 293, células Sf9 y células Cvl. Algunos ejemplos de células eucariotas que no son de mamífero adecuadas incluyen microbios eucariotas tales como hongos filamentosos o levaduras, que son hospedadores adecuados de clonación o de expresión para codificar vectores. *Saccharomyces cerevisiae* es un microorganismo hospedador eucariota inferior de uso habitual. Otros incluyen *Schizosaccharomyces pombe* (Beach y Nurse, Nature, 290: 140 [1981]; documento EP 139.383 publicado el 2 de mayo de 1985); hospedadores de *Kluyveromyces* (Pat. de EE. UU. N.º 4.943.529; Fleer et al., Bio/Technology, 9: 968-975 (1991)), tal como, por ejemplo, *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt et al., J. Bacteriol., 737 [1983]), *K. fragilis* (ATCC 12.424), *K. bulgaricus* (ATCC 16.045), *K. wickeramii* (ATCC 24.178), *K. waltii* (ATCC 56.500), *K. drosophilarum* (ATCC 36.906; Van den Berg et al., Bio/Technology, 8: 135 (1990), *K. thermotolerans* y *K. marxianus*; *yarrowia* (documento EP 402.226); *Pichia pastoris* (documento EP 183.070; Sreekrishna et al., J. Basic Microbiol., 28: 265-278 [1988]); *Candida*; *Trichoderma reesia* (documento EP 244.234); *Neurospora crassa* (Case et al., Proc. Natl. Acad. Sci. USA, 76: 5259-5263 [1979]); *Schwanniomyces* tales como *Schwanniomyces occidentalis* (documento EP 394.538 publicado el 31 de octubre de 1990); y hongos filamentosos tales como, por ejemplo, *Neurospora, Penicillium, Tolypocladium* (documento WO 91/00357 publicado el 10 de enero de 1991), y hospedadores de *Aspergillus* tales como *A. nidulans* (Ballance et al., Biochem. Biophys. Res. Commun., 112: 284-289 [1983]; Tilburn et al., Gene, 26: 205-221 [1983]; Yelton et al., Proc. Natl. Acad. Sci. USA, 81: 1470-1474 [1984]) y *A. niger* (Kelly y Hynes, EMBO J., 4: 475-479 [1985]). Las levaduras metilótropas son adecuadas en el presente documento e incluyen, pero sin limitación, levaduras capaces de crecer en metanol seleccionadas de los géneros que consisten en *Hansenula, Candida, Kloeckera, Pichia, Saccharomyces, Torulopsis,* y *Rhodotorula*. Se puede encontrar una lista de especies específicas que son ilustrativas de esta clase de levaduras en C. Anthony, The Biochemistry of Methylotrophs, 269 (1982).

Otras células adecuadas que pueden utilizarse en la presente invención incluyen, pero sin limitación, cepas de células hospedadoras procariotas tales como *Escherichia coli, (por ejemplo, la cepa DH5-α), Bacillus subtilis, Salmonella typhimurium,* o cepas de los géneros *Pseudomonas, Streptomyces y Staphylococcus.* Algunos ejemplos no limitativos de procariotas adecuados incluyen los de los géneros: *Actinoplanes; Archaeoglobus; Bdellovibrio; Borrelia; Chloroflexus; Enterococcus; Escherichia; Lactobacillus; Listeria; Oceanobacillus; Paracoccus; Pseudomonas; Staphylococcus; Streptococcus; Streptomyces; Thermoplasma; y Vibrio.* Algunos ejemplos no limitativos de cepas específicas incluyen: *Archaeoglobus fulgidus; Bdellovibrio bacteriovorus; Borrelia burgdorferi; Chloroflexus aurantiacus; Enterococcus faecalis; Enterococcus faecium; Lactobacillus johnsonii; Lactobacillus plantarum; Lactococcus lactis; Listeria innocua; Listeria monocytogenes; Oceanobacillus iheyensis; Paracoccus zeaxanthinifaciens; Pseudomonas mevalonii; Staphylococcus aureus; Staphylococcus epidermidis; Staphylococcus haemolyticus; Streptococcus agalactiae; Streptomyces griseolosporeus; Streptococcus mutans; Streptococcus pneumoniae; Streptococcus pyogenes; Thermoplasma acidophilum; Thermoplasma volcanium; Vibrio cholerae; Vibrio parahaemolyticus;* y *Vibrio vulnificus.*

Las células hospedadoras que contienen los polinucleótidos de interés se pueden cultivar en medios nutritivos convencionales (por ejemplo, mezcla de nutrientes de Ham) modificados según sea apropiado para activar promotores, seleccionar transformantes o amplificar genes. Las condiciones de cultivo, tales como la temperatura, el pH y similares, son las utilizadas previamente con la célula hospedadora seleccionada para la expresión, y serán evidentes para el experto habitual. Las células se recogen normalmente por centrifugación, se homogeneizan por medios físicos o químicos, y el extracto en bruto resultante se conserva para su posterior purificación. Para las composiciones secretadas por las células hospedadoras, el sobrenadante de la centrifugación se separa y se conserva para una posterior purificación. Las células microbianas empleadas en la expresión de proteínas pueden homogeneizarse mediante cualquier método conveniente, incluido un ciclo de congelación-descongelación, sonicación, homogeneización mecánica o el uso de agentes de lisado celular, todos los cuales son bien conocidos por los expertos en la técnica. Las realizaciones que implican una lisis celular pueden implicar el uso de un tampón que contenga inhibidores de proteasa que limitan la degradación después de la expresión de la molécula de ADN quimérica. Algunos inhibidores de proteasa adecuados incluyen, pero sin limitación, leupeptina, pepstatina o aprotinina. El sobrenadante puede precipitarse a continuación en concentraciones crecientes sucesivas de sulfato de amonio saturado.

La expresión génica puede medirse en una muestra directamente, por ejemplo, mediante el método de Southern convencional, el método Northern para cuantificar la transcripción del ARNm [Thomas, Proc. Natl. Acad. Sci. USA, 77: 5201-5205 (1980)], inmunotransferencia por puntos (análisis de ADN) o hibridación *in situ*, utilizando una sonda apropiadamente marcada, tomando como base las secuencias proporcionadas en el presente documento. Alternativamente, se pueden emplear anticuerpos que pueden reconocer dúplex específicos, incluidos dúplex de ADN, dúplex de ARN y dúplex híbridos de ADN-ARN o dúplex de ADN-proteína. Los anticuerpos a su vez pueden estar

marcados y el ensayo puede llevarse a cabo donde el dúplex está unido a una superficie, de modo que tras la formación del dúplex en la superficie, se puede detectar la presencia del anticuerpo unido al dúplex.

La expresión génica, alternativamente, se puede medir mediante métodos inmunológicos o fluorescentes, tales como una tinción inmunohistoquímica de células o secciones tisulares y ensayos de cultivo celular o de líquidos corporales o la detección de marcadores de selección, para cuantificar directamente la expresión del producto génico. Los anticuerpos útiles para la tinción inmunohistoquímica y/o el ensayo de líquidos de muestra pueden ser monoclonales o policlonales, y pueden prepararse en cualquier mamífero. Convenientemente, se pueden preparar anticuerpos contra un polipéptido de secuencia natural o contra un péptido sintético tomando como base las secuencias de ADN proporcionadas en el presente documento o contra una secuencia exógena fusionada a restos de direccionamiento y que codifica un epítopo específico. Algunos ejemplos de marcadores de selección son bien conocidos por un experto en la materia e incluyen indicadores tales como la proteína fluorescente verde mejorada (EGFP), la beta-galactosidasa (β-gal) o la cloranfenicol acetiltransferasa (CAT).

El producto o productos de la proteína de fusión de unión expresados pueden purificarse a través de métodos conocidos en la técnica o mediante métodos divulgados en el presente documento. Se pueden utilizar procedimientos tales como filtración en gel, purificación por afinidad, fraccionamiento salino, cromatografía de intercambio iónico, cromatografía de exclusión por tamaño, cromatografía de adsorción en hidroxiapatita, cromatografía de interacción hidrófoba y electroforesis en gel; cada uno adaptado para recuperar y purificar la proteína de fusión producida por las respectivas células hospedadoras. Alguna proteína de fusión de unión expresada puede requerir un replegamiento durante el aislamiento y la purificación. Los métodos de purificación se describen en Robert K. Scopes, Protein Purification: Principles and Practice, Charles R. Castor (ed.), Springer-Verlag 1994, y en Sambrook, *et al., supra.* Las separaciones de purificación en múltiples etapas también se describen en Baron, et al., Crit. Rev. Biotechnol. 10: 179-90 (1990) y en Below, et al., J. Chromatogr. A. 679: 67-83 (1994).

IV). COMPOSICIONES FARMACÉUTICAS

La presente invención proporciona composiciones farmacéuticas que comprenden proteínas de fusión de unión o conjugados de BFP-D. En una realización, la composición farmacéutica comprende la proteína de fusión de unión y al menos un portador farmacéuticamente aceptable. En otra realización, la composición farmacéutica comprende el BFP-D y al menos un portador farmacéuticamente aceptable. Descrita en el presente documento, la composición farmacéutica comprende el conjugado XTEN-fármaco y al menos un portador farmacéuticamente aceptable. Las composiciones farmacéuticas de la presente invención se pueden formular de acuerdo con métodos conocidos para preparar composiciones farmacéuticamente útiles, de modo que el polipéptido se combine en una mezcla con un vehículo portador farmacéuticamente aceptable, tal como soluciones acuosas o tampones, suspensiones y emulsiones farmacéuticamente aceptables. Algunos ejemplos de disolventes no acuosos incluyen propiletilenglicol, polietilenglicol y aceites vegetales. Las formulaciones terapéuticas se preparan para su conservación mezclando el principio activo que tiene el grado deseado de pureza con portadores, excipientes o estabilizantes fisiológicamente aceptables opcionales, como se describe en Remington's Pharmaceutical Sciences 16ª edición, Osol, A. Ed. (1980), en forma de formulaciones liofilizadas o soluciones acuosas. Además, las composiciones farmacéuticas también pueden contener otros compuestos farmacéuticamente activos o una pluralidad de composiciones de la invención.

Las composiciones farmacéuticas se pueden administrar para terapia por cualquier vía adecuada que incluye oral, rectal, nasal, tópica (incluyendo transdérmica, en aerosol, bucal y sublingual), vaginal, parenteral (incluyendo subcutánea, subcutánea con bomba de infusión, intramuscular, intravenosa e intradérmica), intravítrea y pulmonar. También se apreciará que la vía preferida variará dependiendo de la afección y la edad del receptor, y la enfermedad que se está tratando.

En las realizaciones preferidas, la composición farmacéutica se administra por vía parenteral. En esta realización, la composición puede suministrarse como un polvo liofilizado que se va a reconstituir antes de la administración. La composición también puede suministrarse en forma líquida, que puede administrarse directamente a un paciente. En una realización, la composición se suministra como un líquido en una jeringa precargada, de modo que un paciente puede autoadministrarse fácilmente la composición.

Las composiciones de la invención pueden formularse utilizando una variedad de excipientes. Algunos excipientes adecuados incluyen celulosa microcristalina (por ejemplo, Avicel PH102, Avicel PH101), polimetacrilato, poli(acrilato de etilo, metacrilato de metilo, cloruro de metacrilato de trimetilamonioetilo) (tal como Eudragit RS-30D), hidroxipropilmetilcelulosa (Methocel K100M, CR Methocel K100M Premium, Methocel E5, Opadry®), estearato de magnesio, talco, citrato de trietilo, dispersión acuosa de etilcelulosa (Surelease®) y sulfato de protamina. El agente de liberación lenta también puede comprender un portador, que puede comprender, por ejemplo, disolventes, medios de dispersión, recubrimientos, agentes antibacterianos y antifúngicos, agentes isotónicos y retardantes de la absorción. También se pueden utilizar sales farmacéuticamente aceptables en estos agentes de liberación lenta, por ejemplo, sales minerales tales como clorhidratos, bromhidratos, fosfatos o sulfatos, así como las sales de ácidos orgánicos tales como acetatos, proprionatos, malonatos o benzoatos. La composición también puede contener líquidos, tales como agua, solución salina, glicerol y etanol, así como sustancias tales como agentes humectantes, agentes emulsionantes o agentes tamponantes del pH. Los liposomas también pueden usarse como portador.

Para formulaciones líquidas, una propiedad deseada es que la formulación se suministre en una forma que pueda pasar a través de una aguja de calibre 25, 28, 30, 31, 32 para una administración intravenosa, intramuscular, intraarticular o subcutánea. También se pueden utilizar bombas de jeringa para administrar las composiciones farmacéuticas de la invención. Dichos dispositivos se describen en las Pat. de EE. UU. N.º 4.976.696; 4.933.185; 5.017.378; 6.309.370; 6.254.573; 4.435.173; 4.398.908; 6.572.585; 5.298.022; 5.176.502; 5.492.534; 5.318.540; y 4.988.337. Un experto en la técnica, considerando tanto la divulgación de esta invención como la divulgación de estas otras patentes, podría producir una bomba de jeringa para la liberación prolongada de las composiciones de la presente invención.

La administración a través de formulaciones transdérmicas se puede realizar utilizando métodos también conocidos en la técnica, que incluyen los descritos generalmente en, por ejemplo, las Pat. de EE. UU. N.º 5.186.938 y 6.183.770, 4.861.800, 6.743.211, 6.945.952, 4.284.444, y en el documento WO 89/09051. Un parche transdérmico es una realización particularmente útil con polipéptidos que tienen problemas de absorción. Se pueden hacer parches para controlar la liberación de principios activos permeables a la piel durante un periodo de 12 horas, 24 horas, 3 días y 7 días. En un ejemplo, se coloca un exceso diario de 2 veces de un polipéptido de la presente invención en un líquido no volátil. Las composiciones de la invención se proporcionan en forma de un líquido viscoso no volátil. La penetración a través de la piel de formulaciones específicas puede medirse mediante métodos convencionales en la técnica (por ejemplo, Franz et al., J. Invest. Derm. 64: 194-195 (1975)). Algunos ejemplos de parches adecuados son parches cutáneos de transferencia pasiva, parches cutáneos iontoforéticos o parches con microagujas tales como Nicoderm.

En otras realizaciones, la composición puede administrarse por vía intranasal, bucal o sublingual al cerebro para permitir la transferencia de los agentes activos a través de las vías olfativas al SNC y reducir la administración sistémica. Los dispositivos utilizados habitualmente para esta vía de administración se incluyen en la Pat. de EE. UU. N.º 6.715.485. Las composiciones administradas por esta vía pueden permitir una dosis aumentada en el SNC o una carga corporal total reducida, reduciendo los riesgos de toxicidad sistémica asociados con ciertos fármacos. La preparación de una composición farmacéutica para su administración en un dispositivo implantable subdérmicamente se puede realizar utilizando métodos conocidos en la técnica, tales como los descritos, por ejemplo, en las Pat. de EE. UU. N.º 3.992.518; 5.660.848; y 5.756.115.

V). KITS FARMACÉUTICOS

En el presente documento se describe un kit para facilitar el uso de las realizaciones de composición divulgadas en el presente documento. Descrito en el presente documento, el kit comprende, al menos en un primer recipiente: (a) una cantidad de una composición de proteína de fusión de unión suficiente para administrar en el tratamiento de un sujeto con una enfermedad, una afección o un trastorno; y (b) una cantidad de un portador farmacéuticamente aceptable; junto con una formulación lista para inyección o para reconstitución con agua estéril, tampón o glucosa; junto con una etiqueta que identifique el fármaco de la proteína de fusión de unión y las condiciones de conservación y manipulación, y/o una hoja de las indicaciones aprobadas para el fármaco y las instrucciones para la reconstitución y/o la administración del fármaco de la proteína de fusión de unión para usar en la prevención y/o el tratamiento de una indicación aprobada, la dosis apropiada e información de seguridad, e información que identifica el lote y la caducidad del fármaco.

Descrito en el presente documento, el kit comprende, al menos en un primer recipiente: (a) una cantidad de una composición de conjugado de proteína de fusión de unión-fármaco suficiente para administrar en el tratamiento de un sujeto con una enfermedad, una afección o un trastorno; y (b) una cantidad de un portador farmacéuticamente aceptable; junto en una formulación lista para inyección o para reconstitución con agua estéril, tampón o glucosa; junto con una etiqueta que identifique el conjugado de proteína de fusión de unión-fármaco y las condiciones de conservación y manipulación, y/o una hoja de las indicaciones aprobadas para el fármaco y las instrucciones para la reconstitución y/o la administración de las composiciones para usar en la prevención y/o el tratamiento de una indicación aprobada, la dosis apropiada e información de seguridad, e información que identifica el lote y la caducidad del fármaco.

Descrito en el presente documento, el kit comprende, al menos en un primer recipiente: (a) una cantidad de una composición de conjugado de XTEN-fármaco suficiente para administrar en el tratamiento de un sujeto con una enfermedad, una afección o un trastorno; y (b) una cantidad de un portador farmacéuticamente aceptable; junto con una formulación lista para inyección o para reconstitución con agua estéril, tampón o glucosa; junto con una etiqueta que identifique el conjugado de XTEN-fármaco y las condiciones de conservación y manipulación, y/o una hoja de las indicaciones aprobadas para el fármaco y las instrucciones para la reconstitución y/o la administración de las composiciones para usar en la prevención y/o el tratamiento de una indicación aprobada, la dosis apropiada e información de seguridad, e información que identifica el lote y la caducidad del fármaco.

En cualquiera de las realizaciones de los kits anteriores, el kit puede comprender un segundo recipiente que puede portar un diluyente adecuado para la presente composición, lo que proporcionará al usuario la concentración apropiada de la composición farmacéutica que va a administrarse al sujeto.

EJEMPLOS

Ejemplo 1: Construcción de segmentos del motivo XTEN_AD36

El siguiente ejemplo describe la construcción de una colección de genes con codones optimizados que codifican secuencias de motivos de 36 aminoácidos. Como primera etapa, se construyó un vector de relleno pCW0359 basado en un vector pET y que incluye un promotor T7. El pCW0359 codifica un dominio de unión a celulosa (CBD) y un sitio de reconocimiento de proteasa TEV seguido de una secuencia de relleno que está flanqueada por los sitios BsaI, BbsI y KpnI. Los sitios BsaI y BbsI se insertaron de manera que generaran salientes compatibles después de la digestión. La secuencia de relleno está seguida por una versión truncada del gen GFP y una etiqueta de His. La secuencia del relleno contiene codones de terminación y, por lo tanto, las células de *E. coli* que portan el plásmido de relleno pCW0359 forman colonias no fluorescentes. El vector de relleno pCW0359 se digirió con BsaI y KpnI para eliminar el segmento de relleno, y el fragmento de vector resultante se aisló mediante purificación en gel de agarosa. Las secuencias se denominaron XTEN_AD36, reflejando la familia de motivos AD. Sus segmentos tienen la secuencia de aminoácidos $[X]_3$ donde X es un péptido 12mero con las secuencias: GESPGGSSGSES (SEQ ID NO: 2), GSEGSSGPGESS (SEQ ID NO: 3), GSSESGSSEGGP (SEQ ID NO: 4) o GSGGEPSESGSS (SEQ ID NO: 5). El inserto se obtuvo hibridando los siguientes pares de pares de oligonucleótidos sintéticos fosforilados:

AD1for: AGGTGAATCTCCDGGTGGYTCYAGCGGTTCYGARTC (SEQ ID NO: 225)

AD1rev: ACCTGAYTCRGAACCGCTRGARCCACCHGGAGATTC (SEQ ID NO: 226)

AD2for: AGGTAGCGAAGGTTCTTCYGGTCCDGGYGARTCYTC (SEQ ID NO: 227)

AD2rev: ACCTGARGAYTCRCCHGGACCRGAAGAACCTTCGCT (SEQ ID NO: 228)

AD3for: AGGTTCYTCYGAAAGCGGTTCTTCYGARGGYGGTCC (SEQ ID NO: 229)

AD3rev: ACCTGGACCRCCYTCRGAAGAACCGCTTTCRGARGA (SEQ ID NO: 230)

AD4for: AGGTTCYGGTGGYGAACCDTCYGARTCTGGTAGCTC (SEQ ID NO: 231)

También hibridamos el oligonucleótido fosforilado 3KpnIstopperFor: AGGTTCGTCTTCACTCGAGGGTAC (SEQ ID NO: 232) y el oligonucleótido no fosforilado pr_3KpnIstopperRev: CCTCGAGTGAAGACGA (SEQ ID NO: 233). Los pares de oligonucleótidos hibridados se ligaron, lo que dio como resultado una mezcla de productos con longitud variable que representa el número variable de repeticiones de 12meros ligadas a un segmento BbsI/KpnI. Los productos correspondientes a la longitud de 36 aminoácidos se aislaron de la mezcla mediante una electroforesis en gel de agarosa preparativa y se ligaron en el vector de relleno pCW0359 digerido con BsaI/KpnI. La mayoría de los clones en la biblioteca resultante denominada LCW0401 mostraron fluorescencia verde después de la inducción, lo que muestra que la secuencia de XTEN_AD36 se había ligado en marco con el gen GFP y que la mayoría de las secuencias XTEN_AD36 tenían buenos niveles de expresión.

Cribamos 96 aislados de la biblioteca LCW0401 en busca de un alto nivel de fluorescencia estampándolos en una placa de agar que contenía IPTG. Los mismos aislados se evaluaron mediante una PCR y se identificaron 48 aislados que contenían segmentos con 36 aminoácidos, así como una fluorescencia fuerte. Estos aislados se secuenciaron y se identificaron 39 clones que contenían segmentos XTEN_AD36 correctos. Los nombres de archivo de las construcciones de nucleótidos y aminoácidos y las secuencias para estos segmentos se recogen en la tabla 11.

Tabla 11: Secuencias de ADN y de aminoácidos para motivos de 36-meros

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
LCW0401_001_GFP-N_A01.ab1	GSGGSPNFSGSGGESP GGRSGNSESGESPGGS GSES	234	GGTTCTGGTGGTGGTGGAAACCCTCCTGGAGGTCT GGTAGCTCAAGTCCAAATCCTCCGGGTGGC TCTTAGCGGTTCCGGATCCAGGTGGAAATCTC CTGGTGGTTCAAGCGGTCCGAGTCA	272
LCW0401_002_GFP-N_B01.ab1	GSEGSSGPGESSGSESP GGSSGSESGSSESGSS ESGP	235	GGTAGCGGAGGTGAAGGTCCTCTGGGTCCTGGC GAGTCTTCAGGTGGAATCTCCTCGGTGGTT CCAGCGGTTCTTCAATCAGGTTCCTCCGGA AAGCGGTTCTTCCGGAGGACCGGTCCA	273
LCW0401_003_GFP-N_C01.ab1	GSSESGASSTGGTGSST SGSSTGGPGESPGGSS GSGS	236	GGTTCCTCTGGAAACCCGGTACTCCGGAAG CTCGTCCAGTCCACTCTGGAAGCCGGTTC TTCTGGAGCCTCCTCCAGGTGGAAACTCC GGCGTCCCTCCAGCGGTCCGAGTCA	274
LCW0401_004_GFP-N_D01.ab1	GSGGGPSNSGSGESP SGSSEGGPVSGGPSE SGSS	237	GGTTCTGGTGGTCCTGGAAACCCTAACTCT GGTAGCTCAAGTTCTTCTGGAAAGCCGGTT CTCCCGAGGCGGTCCAGGTTCTGGTGGT GTCGAACCTTCCGGACCTTCTGGTTCTTCA	275
LCW0401_007_GFP-N_F01.ab1	GSSESGSSEGGPGSEG SSGPGESSGSESGPGP GPSS	238	GGTTCTTCCGGAAAAGGTGGTCCTCTGGAGG GTGGTCCAGGTAGCTCTTCAAGCGGTTCTTCCGG GTCCAGGTGGAGTCCTTCAAGCTGGCGGAAG GTTCTCCTCCGGTCCTGGTGGAAGTCCTCA	276
LCW0401_008_GFP-N_E01.ab1	GSSESGSSEGGPGESP GGSSGSESGSESGSGP GPSS	239	GGTTCCTCCGGAAAAGGTGGTTCTCCGGAGG GTGGTCCAGGTAGTCCTCCAGGTGGTTT CCAGCGGTTCTGGATCAAGCTGGCGGAAG GTTCTTCCGGGTCCAGGTGGAAGTCCTCA	277
LCW0401_012_GFP-N_H03.ab1	GSGGSPNFSGSGGSGG EPSESGSSESGSGSGP GPSS	240	GGTTCTGGTGGTCCTGGAAACCCTTCTCCTGGAGGTCCTGG GTAGCTCAAGTCCTCCGGTGGTCCTGGAACCCAT CCGGAATCTGGTAGCTTCAGCTCAGGTGGAAGC GTTCTTCCGGGTCCAGGTGGAGTCCTCA	278
LCW0401_015_GFP-N_A02.ab1		241		279

Nombre del archivo		Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
		GSSESGSSEGGPGSSEG SSGPGESSGSESPAGSS GSES		GGTTCTTCCGAAAGCGGTTCTTCTCGAAG GCGGTTCAGGTAGCGAAGGTTCTTCTG GTCCAGCGCAAACTTCAGGTGAATCTC CTGGTGGCTCCAGCGGTTCTGAGTCA	
LCW0401_016_N_B02.ab1	GFP-	GSSESGSSEGGPGSSESGSS SGSSEGGPGSSESGSS EGGP	242	GGTTCTTCCGAAAGCGGTTCTTCTCGAGG GCGGTTCAGGTTCCTCTGAAAGCGGTT CTTCCGAGGCGGTCCAGGTTCTTCTGA AAGCGGTTCTTCTGAGGGCGGTCCA	280
LCW0401_026_N_E02.ab1	GFP-	GSGGEPSESGSSGSEG SSGPGESSGSESPAGSS EGGP	243	GGTTCCGGTGGCGGCGAAACGATCCGAATCT GGTAGCTCCGGTTCTGAAGGTTCTTCT GGTCCAGGCGAAACTTCAGGTTCCTCTG AAAGCGGTTCTTCTGAGGGCGGTCCA	281
LCW0401_022_N_F02.ab1	GFP-	GSGGEPSESGSSGSSE SGSSEGGPGSGGEPSE SGSS	244	GGTTCTGGTGGTGGCGAACCGTCTGAATCT GGTAGCTCACGTTCTTCCGAAAGCGGT TCTTCTGAAGGTGGTCCAGGTTCCGGTG GCGAACCGTCTGAATCTGGTAGCGTCA	282
LCW0401_024_N_G02.ab1	GFP-	GSGGEPSESGSSGSSE SGSSEGGPGSSPGSSS GSES	245	GGTTCTGGTGGCGGAACCGTCTGAATCT GGTAGCTCAGGTTCTTCCGAAAGCGGT TCTTCTGAAGGTGGTCCAGGTTCTGAATCC CAGGTGGTTCTAGCGGTTCTGAATCA	283
LCW0401_026_N_H02.ab1	GFP-	GSGGEPSESGSSGESP GSSSGSESGSEGSSGP GESS	246	GGTTCTGGTGGCGGAACCGTCTGAGTCT GGTAGCTCAGGTGAATCTCCTGGTGGC TCCAGCGGTTCTGAATCAGGTAGCGAA GGTTCTTCTGGTCCTGGTGAATCTTCA	284
LCW0401_027_N_A03.ab1	GFP-	GSGGEPSESGSSGCESP GSSSGSESGSEGEPSE SGSS	247	GGTTCCGGTGGCGGAAACGTTCCGAATCT GGTAGCTCAGGTGAATCTCCGGGTGGT TCTAGCGGTTCTGAATCAGGTTCTGGTG GTGAACCTTCCGAGTCTGGTAGCTCA	285
LCW0401_028_N_B03.ab1	GFP-	GSSESGSSEGGPGSSE SGSSEGGPGSSESGSS EGGP	248	GGTTCTTCTGAAAGCGGTTCTTCTGAGG GCGGTTCAGGTTCTTCCGAAAGCGGTT CTTCCGAGGCGGTCCAGGTTCTTCCGA AAGCGGTTCTTCTGAGGGCGGTCCA	286

Nombre del archivo		Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
LCW0401_030_N_C03.ab1	GFP-	GESRGGSSGSESGSEG SSGPGESSGSGSESGP GESS	249	GGTGAATKTCCGGGTGGCTCCAGCGGT TCTGAGTCAGSTACKGAAGGTTCTTCC GGTCCGGGTGAGTCTCCAGSTAGCGAA GGTTCTTCGGGTCCTGGGTSAGTCTTCA	287
LCW0401_031_N_D03.ab1	GFP-	GSGGEPSESGSSGSGG EPSESGSSGSSESGSS EGGP	250	GGTTCTCGGTGGCGCAACCTTCCGAATCTG GTAGCTCAGGTTCCGGTGGTGAACCTTC TGAATCTGGTAGCTCAGGTTCTTCTGAA AGCCGGTTCTTCCGAGTCTGGTAGCTCA	288
LCW0401_033_E03.ab1	GFP-N	GSGGEPSESGSSGSGG EPSESGSSGSGESPSE SGSS	251	GGTTCTCGGTGGTGAACCTTCTGAATCTG GTAGCTCAGGTTCCGGTGGTGAACCTAT CCGAATCTGGTAGCTCAGGTTCCGGTG GTGAACCATCCGAGTCTGGTAGCTCA	289
LCW0401_037_N_F03.ab1	GFP-	GSGGEPSESGSSGSSE SGSSEEGGPGSEGSGGP GESS	252	GGTTCCGGTGGTGGTGAACCTTCTGAATCTG GTAGCTCAGGTTCCTCCGAAAGCGGTT CTTCTGAGGGCGGTCCAGGTAGCGAAG GTTCTTCTGGTCCTGGTAGCTCTTCA	290
LCW0401_038_N_G03.ab1	GFP-	GSKGEPSESGSSGSEG SSGPGESSGSGGEPSE SGSS	253	GGTTCCGGTGGTGTGAACCGTCCGGAGTCT GGTAGCTCAGGTACCGAAGGTTCTTCT GGTCCGGGTGAGTCTTCAGGTTCTGGTG GTGAACCTTCCGGAATCTGGTAGCTCA	291
LCW0401_039_N_H03.ab1	GFP-	GSKGEPSESGSSGESP GGSSGSESGSGGEPSE	254	GGTTCTGGTGGCGAACCGTCCGAAGTCT GGTAGCTCAGGTGAATCTCCTGGTGGTT	292
		SGSS	255	CCAGCGGTTCCGAGTCAGGTTCTGGTG GTGAACCTTCCGAATCTGGTAGCTCA	293
LCW0401_040_N_A04.ab1	GFP-	GSSESGSSEGGPGSGG EPSESGSSGSESGSS EGGP	255		

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
			GGTTCTCCGAAAGCGGTTCTTCCGAGG GCGGTCCAGGTTCTGGTGGTGAACCAT CTGAATCTGGTAGCTCAGGTTCTCTGA AAGCGGTTCTTCTGAAGGTGGTCCA	
LCW0401_042_N_C04.ab1 GFP-	GSECGSKQPGESSGIESP GGSSGSESGSEESSGP GESS	256	GGTAGCGAAGGTTCTTCTGGTCCTGGT GAGTCTTCAGGTGAATCTCCAGGTGGC TCTAGCGGTTCCGAGTCAGGTAGTGAA GGTTCTTCTGGTCCTGGCCTAGTCCTA	294
LCW0401_046_C04.ab1 GFP-N	GSSESGSSEQGPGSSE SGSSHGKPGSSESGSS EGGP	257	GGTTCTCTGAAAGCGGTTCTTCCGAAG GCGGTCCAGGTTCTTGGGAAAGTGGTT CTTCTGAGGGCGGTCCAGGTTCTCCGA AAGCGGTTCTTCTGAGGGTGGTCCA	295
LCW0401_047_N_E04.ab1 GFP-	GSGGEPSESGSSGESP GGSSGSESGESPGGSS GSES	258	GGTTCTGGTGGCGAAACCTTCCGAGTCTG GTAGCTCAGGTGAAATCTCTGGGGTGGTT CTAGCGGTTCCGAGTCAGGTGAATCTC CGGGCGGTTCCAGCGGTTCTGAGTCA	296
LCW0401_051_N_F04.ab1 GFP-	GSGGEPSESGSSGSEG SSGPGESSGESPGGSS GSES	259	GGTTCTGGTGGTGAAACCATCTCTGAGTCT GGTAGCTCAGGTAGCGGTGAAGGTTCTTCT GGTTCCAGCGGTGAGTCTTCTAGCGGTGAATCT CTGGTGGCCTCCAGCGGTTCTGAGTCA	297
LCW0401_053_N_H04.ab1 GFP-	GESPGSSSLSESGESP GGSSGSESGESPGGSS GSES	260	GGTGAATCTCCTGGTGGTTCTAGCGGTGT CCTAGTCAGGTGAATCTCCAAGTGGCT CTAGCGGTTCCGAGTCAGGTGAATCTC CTGGTGGTTCTAGCGGTTCTGAATCA	298
LCW0401_054_N_A05.ab1 GFP-	GSEGSSGPAGESSGSEG SSGPGESSGSEGESPE SGSS	261	GGTAGCGAAGGTTCTTCTGGTCCAGGT GAATCTTCAGGTGAGGGAGGTTCTTCTG GTCCTGGTGAATCTCCAGGTTCCGGTGG CGAACCATCTGAAATCTGGTAGCTCA	299
LCW0401_059_N_D05.ab1 GFP-	GSGGEPSESGSSGSEG SSGPGESSGESPGGSS GSES	262	GGTTCTGGTGGTGAACCATCCGAATCT GGTAGCTCAGGTGAAGGTTCTTCT GGTCCTGGCGAATCTCCAGGTGAATCTC CAGGTGGCTCTAGCGGTTCCGAATCA	300

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
LCW0401_060_N_E05.ab1	GFP-GSGGEPSESGSSGSSEGSSSEGGPGSGGEPSESGSS	263	GGTTCCGGTGGTGGTGGTAACGGTTCCGAATCTGGTAGCTCAGGTTCCTCTGGTAAAGAGGAGTCTTCCGAGGGTGGTTCCAGGTTCCGGTGTCTGAAGTCTGGTAGCTCA	301
LCW0401_061_N_F05.ab1	GFP-GSSESGSSEGGPGSGGEPSESGSSGSPGESGSSGSEGSSSGP	264	GGTTCTTCTGGTAAAGAGGAGTTGTTCTCTGGAGTCGGTCCAGGTTCTGGTGGTCACAATCATCTGGAACTCTGGTAGCTCAGGTAGGGAAGGTTCTTCTGGTCTCGGTGGTGAATCTTCA	302
LCW0401_063_N_H05.ab1	GFP-GSGGEPSESGSSGSSEGSSGPGESSGSEGSSGP	265	GGTACTGGTGGTGAACCGTCCGAATCTGGTAGCTCAGGTAGCGAAGGTTCTTCTGGTCTCGGTGGTGAATCTTCA	303
LCW0401_066_N_B06.ab1	GFP-GSGGEPSESGSSGSSEGSSSEGGPGSGGEPSESGSS	266	GGTTCTGGTGGCGGAACCATCCGAAGTCGGTAGCTCAGGTTCTTCTGGTAAAGAGGAGTTGTTCTCCGAAGGTGGTTCCAGGTTCTGGTGTGGTGAATCTTCA	304
LCW0401_067_N_C08.ab1	GFP-GSGGEPSESGSSGSSEGSSGPGESSGSEGSSGP	267	GGTTCCGGTGGCGGAACCTTCGAATCTGGAGCTCAGGTAGCGAAGGTTCTTCTGGTCTCGGTGGTGAATCTTCTTACGCGGTTCTGCGAATCAGGTGAATTCCAGGTGGTTCTAGCGGTTCCGAATCA	305
LCW0401_069_N_D06.ab1	GFP-GSGGEPSESGSSGSSEGSSSEGGPGSGGEPSESGSS	268	GGTTCCGGTGGTGGTAACATCTCGAGTCTGGTAGCTCAGGTTCCGGTGGCGGAACCGTCCGAATCTGGTAGCTCAGGTAGGGAAGGTTCTTCTGGTCTCGGTGGTGAATCTTCA	306
LCW0401_070_N_E06.ab1	GFP-GSFGSSGPGESSGSSEGSSEGGPGSGGEPSESGSS	269	GGTAGCGGAAGGTTCTTCTGGTCCGGGCGAATCCTCAGGTTCTCCCGAAGGTGGTTCTAGGGGTGGTGAATCTTCTTCCGGTCCTGGTGGTGAATCTTCA	307

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
LCW0401_078_N_F06.ab1	GFP-GSSESGSSEGGPGESP GGSSGSESGESPGGSS GSES	270	GGTTCCTCTGAAAGCGGGTTCTTCTGAAG GCGGTTCAGGTGAATCTCCGGGTGGCT CCAGCGGTTCTGAATCAGGTGAATCTC CTGGTGGCTCCAGCGGTTCCGAGTCA	308
LCW0401_079_N_G06.ab1	GFP-GSEGSSGPGESSGSEG SSGPGESSGSGGEPSE SGSS	271	GGTAGCGGAAGGTTCTTCTGGTCCAGGC GAGTCTTCAGGTAGCGAAGGTTCTTCC GGTCCTGGCGAGTCTTCAGGTTCCGGTG GCGAACCGTCCGAATCTGGTAGCTCA	309

Ejemplo 2: Construcción de segmentos XTEN_AE36

Se construyó una biblioteca de codones que codificaba secuencias XTEN de 36 aminoácidos de longitud. La secuencia del XTEN se denominó XTEN_AE36. Sus segmentos tienen la secuencia de aminoácidos $[X]_3$ donde X es un péptido 12mero con la secuencia: GSPAGSPTSTEE (SEQ ID NO: 6), GSEPATSGSETP (SEQ ID NO: 7), GTSESATPESGP (SEQ ID NO: 8) o GTSTEPSEGSAP (SEQ ID NO: 9). El inserto se obtuvo hibridando los siguientes pares de pares de oligonucleótidos sintéticos fosforilados:

AE1for: AGGTAGCCCDGCWGGYTCTCCDACYTCYACYGARGA (SEQ ID NO: 310)

AE1rev: ACCTTCYTCRGTRGARGTHGGAGARCCWGCHGGGCT (SEQ ID NO: 311)

AE2for: AGGTAGCGAACCKGCWACYTCYGGYTCTGARACYCC (SEQ ID NO: 312)

AE2rev: ACCTGGRGTYTCAGARCCRGARGTWGCMGGTTCGCT (SEQ ID NO: 313)

AE3for: AGGTACYTCTGAAAGCGCWACYCCKGARTCYGGYCC (SEQ ID NO: 314)

AE3rev: ACCTGGRCCRGAYTCMGGRGTWGCGCTTTCAGARGT (SEQ ID NO: 315)

AE4for: AGGTACYTCTACYGAACCKTCYGARGGYAGCGCWCC (SEQ ID NO: 316)

AE4rev: ACCTGGWGCGCTRCCYTCRGAMGGTTCRGTAGARGT (SEQ ID NO: 317)

También hibridamos el oligonucleótido fosforilado 3KpnIstopperFor: AGGTTCGTCTTCACTCGAGGGTAC (SEQ ID NO: 232) y el oligonucleótido no fosforilado pr_3KpnIstopperRev: CCTCGAGTGAAGACGA (SEQ ID NO: 233). Los pares de oligonucleótidos hibridados se ligaron, lo que dio como resultado una mezcla de productos con longitud variable que representa el número variable de repeticiones de 12meros ligadas a un segmento BbsI/KpnI. Los productos correspondientes a la longitud de 36 aminoácidos se aislaron de la mezcla mediante una electroforesis en gel de agarosa preparativa y se ligaron en el vector de relleno pCW0359 digerido con BsaI/KpnI. La mayoría de los clones en la biblioteca resultante denominada LCW0402 mostraron fluorescencia verde después de la inducción, lo que muestra que la secuencia de XTEN_AE36 se había ligado en marco con el gen GFP, y la mayoría de las secuencias XTEN_AE36 muestra una buena expresión.

Cribamos 96 aislados de la biblioteca LCW0402 en busca de un alto nivel de fluorescencia estampándolos en una placa de agar que contenía IPTG. Los mismos aislados se evaluaron mediante una PCR y se identificaron 48 aislados que contenían segmentos con 36 aminoácidos, así como una fluorescencia fuerte. Estos aislados se secuenciaron y se identificaron 37 clones que contenían segmentos XTEN_AE36 correctos. Los nombres de archivo de las construcciones de nucleótidos y aminoácidos y las secuencias para estos segmentos se recogen en la tabla 12.

Tabla 12: **Secuencias de ADN y de aminoácidos para motivos de 36-meros**

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
LCW0402_002_N_A07.ab1 GFP-	GSPAGSPTSTEEGT SESATPESQPGTSTE PSEGSAP	318	GGTAGCCCCGGCAGCGCTCTCCGACCTCTACT GAGGAAGGTACTTCTGAAAGCGCAACCCCG GAGTCCGGCCCAGGTACCTCTACCGAACCG TCTGAAGGTAGCGCAACCA	355
LCW0402_003_N_B07.ab1 GFP-	GTSTEPSEGSAPGT STEPSEGSAPGTSTE PSEGSAP	319	GGTACTTCTACCGAACCGTCTGAAGGCAGC GCTCCAGGTACCTCTACTGAACCTTCCGAG GGCAGCGCTCCAGGTACCTCTACCGAACCT TCTGAAGGTAGCGCAACCA	356
LCW0402_004_N_C07.ab1 GFP-	GTSTEPSEGSAPGT SESATPESQPGTSES ATPESGP	320	GGTACCTCTACCGAACCGTCTGAAGGTAGC GCACCAGGTACCTCTGAAAGCGGAACTCT GAGTCCGGTCCAGGTACTTCTGAAAGCGCA ACCCCGGAGTCTGGTCCA	357
LCW0402_005_N_D07.ab1 GFP-	GTSTEPSEGSAPGT SESATPESQPGTSES ATPESGP	321	GGTACTTCTACTGAAACCGTCTGAAGGTAGC GCACCAGGTACTTCTGAAAGCGCAACCCCG GAATCCGGCCCAGGTACCTCTGTGAAAGCGCA ACCCCGGAGTCCGGTCCA	358
LCW0402_006_N_E07.ab1 GFP-	GSEPATSGSETPGT SESATPESQPGSPA GSPTSTEE	322	GGTAGCGAACCGGCCAACCTCCGGCTCTGAA ACCCCAGGTACCTCTGAAAGCGCTACTCCT GAATCCGGCCCAGGTAGCCCCGCCAGGTCT CCGACTTCCACTGAGGAA	359
LCW0402_008_N_F07.ab1 GFP-	GTSESATPESGPGS EPATSGSETPGTSTE PSEGSAP	323	GGTACTTCTGAAAGCGCAACCCCTGAATCC GGTCCAGGTAGCGAAACCGGCTACTCTGGC TCTGAAACTCCAGGTACTTCTACCGAACCG TCCGAAGGTAGCGCACCA	360
LCW0402_009_N_G07.ab1 GFP-	GSPAGSPTSTEEGSP AGSPTSTEEGSEPA TSGSETP	324	GGTAGCCCCGGCTCTCCAACCTCCACT GAGGAAGGTAGCCCCGGCTCTGGTCTCTCCAACC TCCACTGAAGAAGGTAGCGAACCCGGCTACC TCCGGCTCTGAAACTCCA	361
LCW0402_011_N_A08.ab1 GFP-		325		362

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
	GSPAGSPTSTEEGT SESATPESGPGTSTE PSEGSAP		GGTAGGCCCGGCTGGCTCTCCTACCTCTACT GAGGAAGGTACTTCTGAAAGCGCTACTCCT GAGTCTGGTCCAGGTACCTCTACTGAACCG TCCGAAGGTAGCGCTCCA	
LCW0402_012_N_B08.ab1 GFP-	GSPAGSPTSTEEGSP AGSPTSTEEGTSTEP SEGSAP	326	GGTAGCCCTGGCTGGACTCTCCGACTCTACTG AGGAAGGTAGCCCGGCTGGTTCTCCGACTT CTACTGAGGAAGGTACTTCTACCGAACCTT CCGAAGGTAGCGCTCCA	363
LCW0402_013_N_C08.ab1 GFP-	GTSESATPESGPGT STEPSEGSAPGTSTE PSEGSAP	327	GGTACTTCTGAAAGCGCTACTCCTGAGTCC GGTCCAGGTACCTCTACCGAACCTTCCGAA CCGGCTCCTCAGGTACTTCTACTGAACCT TCTGAAGGTAGCGCTCCA	364
LCW0402_014_N_D08.ab1 GFP-	GTSTEPSEGSAPGSP AGSPTSTEEGTSTEP SEGSAP	328	GGTACCTCTACCGAACCTTCCGAAGGTAGC GCTCCAGGTAGCCCTGGCAGGTTCTCTACT TCCACTGAAGAAGGTACTTCTACCGAACCT	365
			TCTGAGGGTAGCGCACCA	
LCW0402_015_N_E08.ab1 GFP-	GSEPATSGSETPGSP AGSPTSTEEGTSES ATPESGP	329	GGTAGCGAACCGGGTACTTCCGGCTCTGAG ACTCCAGGTAGCCCTGGCTCTCCGAACC TCTACCGAAGAAGGTACCTCTGAAAGCGCT ACCCCTGAGCCTGGCCCA	366
LCW0402_016_N_F08.ab1 GFP-	GTSTEPSEGSAPGT SESATPESGPGTSES ATPESGP	330	GGTACTTCTACCGAACCTTCCGAGGGCAGC GCACCAGGTACTCTGAAAGCGCTACCCCT GAGTCCGGCCCCAGGTACTTCTGAAAGCGCT ACTCCTGAACCTGGTCCA	367
LCW0402_020_N_G08.ab1 GFP-	GTSTEPSEGSAPGS EPATSGSETPGSPA GSPTSTEE	331	GGTACTTCTACTGAACCGTCTGAAGGCAGC GCACCAGGTAGCGAACCCGGTACTTCCGGT TCTGAAACCCCAGGTAGCCCAGGTAGCCCT CCAACTTCTACTGAAGAA	368
LCW0402_023_N_A09.ab1 GFP-		332		369

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
	GSPAGSPTSTEEGT SESATPESGPGSEPA TSGSETP		GGTAGCCCTGCTGGCTCTCCAACCTCCACC GAAGGAGGTACCTCTGAAAGCGCAACCCCT GAATCCGGCCCAGGTAGCGAACCTGGCAAC CTCGGGTCTTCTGAAACCCA	
LCW0402_024_GFP-N_B09.ab1	GTSESATPESGPGSP AGSPTSTEEGSPAG SPTSTEE	333	GGTACTTCTGAAAGCGCTACTCCTGAGTCC GGTCCAGGTAGCCCGGAGTGGCTCTCCGACT TCCACCGAGGAAGGTAGCCCTGGCTCGGCTCT CCAACTTCTACTGAAGAA	370
LCW0402_025_GFP-N_C09.ab1	GTSTEPSEGSAPGT SESATPESGPGTSTE PSEGSAP	334	GGTACCTCTACTGAACCTTCTGAGGGCAGC GCTCCTAGGTACTTCTGAAAGCGCTACCCCG GAGTCCGGTCAGGTACTTCTACTGAACCG TCCGAAGCTAGCGCACCA	371
LCW0402_026_GFP-N_D09.ab1	GSPAGSPTSTEEGT STEPSEGSAPGSEPA TSGSETP	335	GGTAGCCCGGGCAGCTCTCCGACTTCCACC GAGGAAGGTACCTCTACTGAAGCCTCTGAG GCTAGCGGTCCAGGTAGCGAACCGGGCAACC TCTGGCTCTTCGAAACCCCA	372
LCW0402_027_GFP-N_E09.ab1	GSPAGSPTSTEEGT STEPSEGSAPGTSTE PSEGSAP	336	GGTAGCCCCAGCAGGCTCTCCGACTTCCACT GAGGAAGGTACTTCTACTGAAGCTTCTGAA GGCAGCGCACCAGGTACCTCTACTGAACCT TCTGGAGGCAGTGGCTCCA	373
LCW0402_032_GFP-N_H09.ab1	GSEPATSGSETPGT SESATPESGPGSPA GSPTSTEE	337	GGTAGCGAACCTGCTACCTCCGGTTCTGAA ACCCCAGGTACTTCTGAAAGCGGCAACTCCG GAGTCTGGTCCAGGTAGCCCTGGCAGGTCT CCTACTTCCACTGAAGAA	374
LCW0402_034_GFP-N_A10.ab1	GTSESATPESGPGT STEPSEGSAPGTSTE PSEGSAP	338	GGTACTCTGGAAAGCGGCACTCCGGAGTCT GGCCCAGGTACCTCTACTGAAACCGTCTGAG GGTAGCGGCTCCAGGTACTTCTACTGAACCG TCCGAAGCTACCGCACCA	375
LCW0402_036_GFP-N_C10.ab1	GSPAGSPTSTEEGT STEPSEGSAPGTSTE PSEGSAP	339		376

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
			GGTAGCCGGCTGGTTCTCTGGACTTCTCACC GAGGAAGGTACCTCTACTGAACCTTCTGAG GGTAGCGCTCCAGGTACCTCTACTGAACCT TCCGAAGGCAGCTCCA	
LCW0402_039_N_E10.ab1 GFP-	GTSTEPSEGSAPGT STEPSEGSAPGTSTE PSEGSAP	340	GGTACTTCTACTGAACCTTCTGAAGGCAGC GCTCCAGGTACTTCTACTGAACCTTCTGAA GGCAGCGCTCCAGGTACTTCTACTGAACCT TCCGAAGGTAGCGCCA	377
LCW0402_040_N_F10.ab1 GFP-	GSEPATSGSETPGT SESATPESGPGTSTE PSEGSAP	341	GGTAGCGAACCTGCAACCCTTGGCTCTGAA ACCCAGGTACCTCTGAAAGGCTACTCCT GAATCTGGCCAGGTACTTCTACTGAACCG TCCGACGGCAGGCACCA	378
LCW0402_041_N_G10.ab1 GFP-	GTSTEPSEGSAPGSP AGSPTSTEEGTSTEP SEGSAP	342	GGTACTTCTACCGAACCGTCTGAAGGTAGC GCAUCAGGTACCTCAGCAGGTTCTCCTACC TCCACCGAGGAAGGTACTTCTACCGAACCG TCCGAGGGTAGCGGCACCA	379
LCW0402_050_N_A11.ab1 GFP-	GSEPATSGSETPGT SESATPESGPGSEPA TSGSETP	343	GGTAGCGAACCTGGCACAACCTCCGGCTCTGAA ACTCCAGGTACTTCTGAAAGCGCTACTCCT GAATCCGGCCCAGGTAGCGAACCGGGCACT TCCGACTCTGAAACCCCA	380
LCW0402_051_N_B11.ab1 GFP-	GSEPATSGSETPGT SESATPESGPGSEPA TSGSETP	344	GGTAGCGAACCTGGCAACTTCCGGCTCTGAA ACCCCAGGTACTTCTGAAAGCGCTACTCCT GAGTCTGGCCCAGGTAGCGAACCTGCTACC TCTGGCTCTGAAACCCCA	381
LCW0402_059_N_E11.ab1 GFP-	GSEPATSGSETPGS EPATSGSETPGTSTE PSEGSAP	345	GGTAGCGAACCTGGCAACTTCTGGCTCTGAA ACTCCAGGTAGCGAACCTGCAACCTCCGGC TCTGAACCCCAGGTACTTCTACTGAACCT TCTGACGGCAGCGACCA	382
LCW0402_060_N_F11.ab1 GFP-	GTSESATPESGPGS EPATSGSETPGSEP ATSGSETP	346		383

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
			GGTACTTCTGAAAGCGCTACCCCGGAATCT GGCCCAGGTAGGGAAACGGGTACTTCTGGT TCTGAAACCCCAGGTAGCGAACCGGCTACC TCCGGTTCTGAAACTCCA	
LCW0402_061_GFP-N_G11.ab1	GTSTEPSEGSAPGT STEPSEGSAPGTSES ATPESGP	347	GGTACTTCTACTGAACCTTCGAAGGCAGC GCTCCAGGTACCTACGGAACGGTCCGAG GGCAGCGCACCAGCTACTCTGAAAGCGCA ACCCTGGAATCCGGTCCA	384
LCW0402_065_GFP-N_A12.ab1	GSEPATSGSETPGT SESATPESGPGTSES ATPESGP	348	GGTAGCGAACCGGCAACGCTCTGGCTCTGAA ACCCCAGGTACCTCTGAAAGCGCTACTCCG GAATCTGGTCCAGGTACTCTGAAAGCGCT ACTCCGGAAATCCGGTCCA	385
LCW0402_066_GFP-N_B12.ab1	GSEPATSGSETPGS EPATSGSETPGTSTE PSEGSAP	349	GGTAGCGAACCTGCTACCTCCGGCTCTGAA ACTCCAGGTACCTCTGAAAGCGCTACTTCCGGT TCTGAAACTCCAGGTACTTCTACCGAAACT TCCGGAAGCCAGCGCAACTCA	386
LCW0402_067_GFP-N_C12.ab1	GSEPATSGSETPGT STEPSEGSAPGSEPA TSGSETP	350	GGTAGCGAAACTGCTACTTCTGGTTCTCAA ACTCTAGGTACTTCTACCGAACCGTCTGAG GGTAGCGCTCCAGGTAGCGAACCTGCTACT TCTGGTTCTGAAACTCCA	387
LCW0402_068_GFP-N_D12.ab1	GTSTEPSEGSAPGT STEPSEGSAPGSEPA TSGSETP	351	GGTACCTCTACCGAACCGTCTGAAGGTGCAGC GCACCAGGTACCTCTACTGAACCGTCTGAG GGTAGCGCTCCAGGTAGCGAACCGGCTACC TCCGGTTCTGAAACTCCA	388
LCW0402_073_GFP-N_F12.ab1	GTSTEPSEGSAPGS EPATSGSETPGSEPA GSPTSTEE	352	GGTACTTCTACTGAACCTTCGAAGGTAGC GCTCCAGGTACCTACTGCTACTTCTGGT TCTGAAACCCCAGGTAGCCCGGCTGGCTCT CCGACTCCCACCTAAGAAA	389
LCW0402_074_GFP-N_G12.ab1	GSEPATSGSETPGSP AGSPTSTEEGTSES ATPESGP	353		390

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
			GGTAGCGGAACCGGCTACTTCCGGCTCTGAG ACTCCAGGTAGCCCAGCTGGTTCTTCCAACC TCTACTGAGGAAGGTACTCTGAAAGCGCT ACCCCTGAATCTGGTCCA	
LCW0402_075_GFP-N_H12.ab1	GTSESATPESGPGS EPATSGSETPGTSES ATPESGP	354	GGTACCTCTGAAAGCGCAACTCCTGAGTCT GGCCCAGGTAGCGAACCTGCTACCTCCGGC TCTGAGACTCCAGGTACCTCTGAAAGCGCA ACCCCGGAATCTGGTCCA	391

Ejemplo 3: Construcción de segmentos XTEN_AF36

Se construyó una biblioteca de codones que codificaba secuencias de 36 aminoácidos de longitud. Las secuencias se denominaron XTEN_AF36. Sus segmentos tienen la secuencia de aminoácidos $[X]_3$ donde X es un péptido 12mero con la secuencia: GSTSESPSGTAP (SEQ ID NO: 10), GTSTPESGSASP (SEQ ID NO: 11), GTSPSGESSTAP (SEQ ID NO: 12) o GSTSSTAESPGP (SEQ ID NO: 13). El inserto se obtuvo hibridando los siguientes pares de pares de oligonucleótidos sintéticos fosforilados:

AF1for: AGGTTCTACYAGCGAATCYCCKTCTGGYACYGCWCC (SEQ ID NO: 392)

AF1rev: ACCTGGWGCRGTRCCAGAMGGRGATTCGCTRGTAGA (SEQ ID NO: 393)

AF2for: AGGTACYTCTACYCCKGAAAGCGGYTCYGCWTCTCC (SEQ ID NO: 394)

AF2rev: ACCTGGAGAWGCRGARCCGCTTTCMGGRGTAGARGT (SEQ ID NO: 395)

AF3for: AGGTACYTCYCCKAGCGGYGAATCTTCTACYGCWCC (SEQ ID NO: 396)

AF3rev: ACCTGGWGCRGTAGAAGATTCRCCGCTMGGRGARGT (SEQ ID NO: 397)

AF4for: AGGTTCYACYAGCTCTACYGCWGAATCTCCKGGYCC (SEQ ID NO: 398)

AF4rev: ACCTGGRCCMGGAGATTCWGCRGTAGAGCTRGTRGA (SEQ ID NO: 399)

También hibridamos el oligonucleótido fosforilado 3KpnIstopperFor: AGGTTCGTCTTCACTCGAGGGTAC (SEQ ID NO: 232) y el oligonucleótido no fosforilado pr_3KpnIstopperRev: CCTCGAGTGAAGACGA (SEQ ID NO: 233). Los pares de oligonucleótidos hibridados se ligaron, lo que dio como resultado una mezcla de productos con longitud variable que representa el número variable de repeticiones de 12meros ligadas a un segmento BbsI/KpnI Los productos correspondientes a la longitud de 36 aminoácidos se aislaron de la mezcla mediante una electroforesis en gel de agarosa preparativa y se ligaron en el vector de relleno pCW0359 digerido con BsaI/KpnI. La mayoría de los clones en la biblioteca resultante denominada LCW0403 mostraron fluorescencia verde después de la inducción, lo que muestra que la secuencia de XTEN_AF36 se había ligado en marco con el gen GFP, y la mayoría de las secuencias XTEN_AF36 muestra una buena expresión.

Cribamos 96 aislados de la biblioteca LCW0403 en busca de un alto nivel de fluorescencia estampándolos en una placa de agar que contenía IPTG. Los mismos aislados se evaluaron mediante una PCR y se identificaron 48 aislados que contenían segmentos con 36 aminoácidos, así como una fluorescencia fuerte. Estos aislados se secuenciaron y se identificaron 44 clones que contenían segmentos XTEN_AF36 correctos. Los nombres de archivo de las construcciones de nucleótidos y aminoácidos y las secuencias para estos segmentos se recogen en la tabla 13.

Tabla 13: Secuencias de ADN y de aminoácidos para motivos de 36-meros

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
LGW0403_004_N_A01.ab1	GFP- GESTPESASPKSTIS PGKESSTAPGTPSPV ESSTAP	455	GGTAGTCTTGACTCCGAAAGGGGTTATGGCA TCTCCAGCTACTTCTCCTGAGGGTTGAATCTT CCTACTGCCGAGCTACTCCTCCAGCGGCG AATCTTGTACTGCTGCA	444
LGW0403_009_N_D01.ab1	GFP- GPSPGESSTAPGSTSPV PSGESSTAPGTPSPV ESSTAP	459	GGTACTTCTCCAGCGGTTATGGTGTTGCA TCTCCAGCTACTTCTCCTGAGGGTTGAATCTT CCTACTGCCGAGCTACTCCTCCAGCGGCG AATCTTGTACTGCTGCA	445
LGW0403_008_N_G01.ab1	GFP- GSTSSTAPGTPSPVE PSGESSTAPGTPSPV GSASP	458	GGTTCCACCTCTGCTACTCCTGAGGGTACTCCT TCTCCTGTTGAATCTGCTCCTTCTGGTTCT GCTTCTACTGCTGAGGGTACTCCTTCTCCTGTT GAATCTGCTCCTTCTGGTTCTGCTTCTACTGC TGAGGGTACTCCTTCTCCTGTTGAATCTGCT CCTTCTGGTTCTGCTTCTACTGCTGAGGGTA CTCCTTCTCCTGTTGAATCTGCTCCTTCTGG TTCTGCTTCTACTGCTGAGGGTACTCCTTCT CCTGTTGAATCTGCTCCTTCTGGTTCTGCTT CTACTGCTGAGGGTACTCCTTCTCCTGTTGA ATCTGCTCCTTCTGGTTCTGCTTCTACTGCT GAGGGTACTCCTTCTCCTGTTGAATCTGCTC CTTCTGGTTCTGCTTCTACTGCTGAGGGTAC TCCTTCTCCTGTTGAATCTGCTCCTTCTGGT TCTGCTTCTACTGCTGAGGGTACTCCTTCTC CTGTTGAATCTGCTCCTTCTGGTTCTGCTTC TACTGCTGAGGGTACTCCTTCTCCTGTTGAA TCTGCTCCTTCTGGTTCTGCTTCTACTGCTG AGGGTACTCCTTCTCCTGTTGAATCTGCTCC TTCTGGTTCTGCTTCTACTGCTGAGGGTACT CCTTCTCCTGTTGAATCTGCTCCTTCTGGTT CTGCTTCTACTGCTGAGGGTACTCCTTCTCC TGTTGAATCTGCTCCTTCTGGTTCTGCTTCT ACTGCTGAGGGTACTCCTTCTCCTGTTGAAT CTGCTCCTTCTGGTTCTGCTTCTACTGCTGA GGGTACTCCTTCTCCTGTTGAATCTGCTCCT TCTGGTTCTGCTTCTACTGCTGAGGGTACTC CTTCTCCTGTTGAATCTGCTCCTTCTGGTTC TGCTTCTACTGCTGAGGGTACTCCTTCTCCT GTTGAATCTGCTCCTTCTGGTTCTGCTTCTA CTGCTGAGGGTACTCCTTCTCCTGTTGAATC TGCTCCTTCTGGTTCTGCTTCTACTGCTGAG GGTACTCCTTCTCCTGTTGAATCTGCTCCTT CTGGTTCTGCTTCTACTGCTGAGGGTACTCC TTCTCCTGTTGAATCTGCTCCTTCTGGTTCT GCTTCTACTGCTGAGGGTACTCCTTCTCCTG TTGAATCTGCTCCTTCTGGTTCTGCTTCTAC TGCTGAGGGTACTCCTTCTCCTGTTGAATCT GCTCCTTCTGGTTCTGCTTCTACTGCTGAGG GTACTCCTTCTCCTGTTGAATCTGCTCCTTC TGGTTCTGCTTCTACTGCTGAGGGTACTCCT TCTCCTGTTGAATCTGCTCCTTCTGGTTCTG CTTCTACTGCTGAGGGTACTCCTTCTCCTGT TGAATCTGCTCCTTCTGGTTCTGCTTCTACT GCTGAGGGTACTCCTTCTCCTGTTGAATCTG CTCCTTCTGGTTCTGCTTCTACTGCTGAGGG TACTCCTTCTCCTGTTGAATCTGCTCCTTCT GGTTCTGCTTCTACTGCTGAGGGTACTCCTT CTCCTGTTGAATCTGCTCCTTCTGGTTCTGC TTCTACTGCTGAGGGTACTCCTTCTCCTGTT GAATCTGCTCCTTCTGGTTCTGCTTCTACTG CTGAGGGTACTCCTTCTCCTGTTGAATCTGC TCCTTCTGGTTCTGCTTCTACTGCTGAGGGT ACTCCTTCTCCTGTTGAATCTGCTCCTTCTG GTTCTGCTTCTACTGCTGAGGGTACTCCTTC TCCTGTTGAATCTGCTCCTTCTGGTTCTGCT TCTACTGCTGAGGGTACTCCTTCTCCTGTTG A

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
LOW0403_012_GFP-N_H07.ab1	GSTSESPSGTAPGTS PSGESSTAPGSTSESP SGTAP	407	GGTTCTAACAGGGAATCTCCTTCTGGCACC GCTTCAGGTACCTCTCTTAGCGGCGAATCT TCTACCGCTCCAGGTTCTACTAGCGGAATCTC CTTCTGGCACTGGCACCA	451
LOW0403_013_GFP-N_A02.ab1	GSTSSTAAESPGPGST SSTAAESPGPGTSPSG ESSTAP	408	GGTTCTACTAGCTCTACTTCTGAAGAATCTGCT GGCTCAGGTTCTACTAGCTCTACTGCAGAA TCTCCGGGTCCAGGTACTAGCTCTCCTAGCGGC GAATCTGCTACCGGTGCTCA	452
LOW0403_014_GFP-N_B02.ab1	GSTSSTAAESPGPGTS TPESGSASPGSTSESP SGTAP	409	GGTTCCACTAGCTCTACTTCTGCAGAATCTGCT GGCCCAGGTACCTCTACCCTGAAAGCGGC TCTGCATCTGCAGGTTCTACCAGCGAATCC CCGTCTGGTGCACCGCACCA	453
LOW0403_015_GFP-N_C02.ab1	GSTSSTAAESPGPGST SSTAAESPGPGTSPSG ESSTAP	410	GGTTCTACTAGCTCTCCACTCCTGAATCTCCGG GTCCAGGTTCTACTAGCTCTACTGCTGAAT CTCCTGGTCCAGGTACCTCCCCGAGCGGTG AATCTACTACTGGCACCA	454
LOW0403_017_GFP-N_D02.ab1	GSTSSTAAESPGPGST SESPSGTAPGSTSST AESPGP	411	GGTTCTAGCAGCTCTAGCGGTGAATCTACTAGTG GCCCAGGTTCTACCAGCGGAATCCCCGGTCTG GCACCGCACCAGGTTCTACTAGCTCTACCG CTGGAATCTGGTGGTCCA	455
LOW0403_018_GFP-N_E02.ab1	GSTSSTAAESPGPGST SSTAESPGPGSTSST AESPGP	412	GGTTCTACCAGCTCTACCGCAGAATCTCCT GGCCCAGGTTCCACTAGCTCTACCGGTGAA TCTCCTGGTCCAGGTTCTACTAGCTCTACCG CTGGAATCTCCTGGTGCCA	456
LOW0403_019_GFP-N_F02.ab1	GSTSESPSGTAPGST SSTAESPGPGSTSST AESPGP	413	GGTTCTAAGCGAATCCCCTTCTGGTACTG CTCCAGGTTCCACTAGCTCTACCGCTGAAT CTCCTGGTCCAGGTTCCACTAGCTCTACTGC AGAATCTCCTGGTCCA	457

5

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
LCW0403_023_N_H02.ab1 GFP-	GSTSESPSGTAPGST SESPSGTAPGSTSESP SGTAP	414	GGTTCTACTAGCGGAATCTCCTCCTCTGGTACCG CTCCAGGTTCTAGCAGTGGAATCTCCGGTCTG GTACTGCTCCAGGTTCTACCAGCGGAATCTC CTTCTGGTACTGCACCA	458
LCW0403_024_N_A03.ab1 GFP-	GSTSSTAESPGPGST SSTAESPGPGSTSST AESPGP	415	GGTTCCACCAGCTCTACTGCTGGAATCTCTG GCCCAGGTTCTACCAGCTCTACTGCTGAAT CTCCGGCCCAGGTTCCACCAGCTCTACC CTGAATCTGGTGGTTCA	459
LCW0403_025_N_B03.ab1 GFP-	GSTSSTAESPGPGST SSTAESPGPGTSPSG ESSTAP	416	GGTTCCACTAGCTCTACCGGCAGAAATCTCT GGTCCAGGTTCTACTAGCTCTACTGCTGAA TCTCCGGGTCCAGGTACCTCCCCTAGCGGC GAATCTTCTACCGCTCCA	460
LCW0403_028_N_D03.ab1 GFP-	GSSPSASTGTGPGSS TPSGATGSPGSSTPS GATGSP	417	GGTTCTACGCCTTCTGGTCTCCTACCGGTACCG TCCCAGGTAGCTCTACCTCCGTCTGGTCAA CTGGTCCTCCAGGTAGCTCTACTCCGTCTGG TGCAACCGGCTCCCCA	461
LCW0403_029_N_E03.ab1 GFP-	GTSPSGESSTAPGTS TPESGSASPGSTSST AESPGP	418	GGTACTTCCCCTACGGTGGAATCTCTACTG CTCCAGGTACCTCTACTCCGGAAACGCCCT CCGCAACTCCAGGTTCTACTAGCTCTACTGC TGAAACTGCTGGTGGTCCA	462
LCW0403_030_N_F03.ab1 GFP-	GSTSSTAESPGPGST SSTAESPGPGTSTPES GSASP	419	GGTTCTACTAGCTCTACCGGTGAAATCTCCG GCTCCAGGTTCTACCAGCTCTACTGCAGAA TCTCCTGGCCCAGGTACTTCTACTCCGGAA AGCGGTTCCGGCTTCTCCA	463
LCW0403_031_N_G03.ab1 GFP-	GTSPSGESSTAPGST SSTAESPGPGTSTPES GSASP	420	GGTACTTCTCCTACGGTGGTGAATCTTCTACCG CTCCAGGTTCTACCAGCTCTACTGCTGAATC TCCTGGCCCAGGTACTTCTACCCCGGAAAG CGGCTCCGCTTCTCCA	464
LCW0403_033_N_H03.ab1 GFP-		421		465

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
	GSTSESPSGTAPGST SSTAESPGPGSTSST AESPGP		GGTTCTACTAGCGGAAATCCCCCTTCTGGTACTG CACCAGGTTCTACCAGCTCTACTGCTGAAT CTCCGGGCCCAGGTTCCAGCAGCTCTACCG CAGAATCTCCTGGTCCA	
LCW0403_035_GFP_N_A04.ab1	GSTSSTAESPGPGST SESPSGTAPGSTSSI AESPGP	422	GGTTCCACCAGCTCTACCGCTGAATCTCCG GGCCCAGGTTCTACCAGCGAATCCCCTTCT GGCACTGCACCAGGTTCTACTAGCTCTACC GCAGAATCTCCTGGTCCA	466
LCW0403_036_GFP_N_B04.ab1	GSTSSTAESPGPGTS PSGESSTAPGTSTPES GSASP	423	GGTTCTACCAGCTCTACTGCTGAAACTCCG GGTCCAGCTACTCCCCGAGCGGTGAATCT TCTACTGCACCAGGTACTCTACTCCGGAA AGCGGTTCGCTTCTGCA	467
LCW0403_039_GFP_N_C04.ab1	GSTSESPSGTAPGST SESPSGTAPGTSPSG ESSTAP	424	GGTTCTACCAGCGAATCTCCTCTTCTGGCACC GCTCCAGGTTCTACTAGCGAATCTCCGACT GGTACCGGCAGCAGGTACTTCTCCTAGCGGC GAATCTTCTACCGCACCA	468
LCW0403_041_GFP_N_D04.ab1	GSTSESPSGTAPGST SESPSGTAPGTSTPES GSASP	425	GGTTCTACCAGCGGAATCCCCTCTGCTACT GCTCCAGGTTCTACCAGCGAATCTCCTTCT GGCACCGCACCAGGTACTTCTACCCCTGAA AGCGGCTCCGCTTCTCTCA	469
LCW0403_044_GFP_N_E04.ab1	GTSTPESGSASPGST SSTAESPGPGSTSST AESPGP	426	GGTACCTCTACTCCTCTGAAAGCGGTTCTGCA TCTCCAGGTTCCACTAGCTCTACTGCAGAA TCTCCGGGCCCAGGTTCTACTAGCTCTACTG CTGAATCTCCTGGTCCA	470
LCW0403_046_GFP_N_F04.ab1	GSTSESPSGTAPGST SESPSGTAPGTSPSG ESSTAP	427	GGTTCTACTAGGGGAATCCCCTTCTGGCACT GCACCAGGTTCTACTAGCGGAATCCCCTTCT GGTACCGCACCAGGTACTTCTCCGAGCGGC GAATCTTCTACTGGTCCA	471
LCW0403_047_GFP_N_G04.ab1	GSTSSTAESPGPGST SSTAESPGPGSTSESP SGTAP	428		472

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
			GGTTCTACTAGCTCTAAGGGTGAATCTTCTG GCCCAGGTTCCACTAGGTCTACCGCAGAAT CTCCGGCCCAGGTTCTACTAGCGAATCCC CTTCTGGTACCGCTCCA	
LCW0403_049_N_H04.ab1 GFP-	429 GSTSSTAESPGPGTSTPES SSTAESPGPGTSTPES GSASP		473 GGTTCTACCAGCTCTACTCCAGAATCTCCT GGCCCAGGTTCTACTAGCTCTACCGCAGAA TCTCCTGGTCCAGGTACCTCTACTCCTGAAA GCGGTTCCCCATCTCCA	473
LCW0403_051_N_A05.ab1 GFP-	430 GSTSSTAESPGPGST SSTAESPGPGSTSESP		474 GGTTCTACTAGCTCTACTGCTGAATCTCCGG GCTCCAGGTTCTACTAGGTCTATCTGCTGAAT	474
	SGTAP		CTCCGGGTCCAGGTTCTACTAGCGAATCTC CTTCTGGTACCGCTCCA	
LCW0403_053_N_B05.ab1 GFP-	431 GTSPSGESSTAPGSTSST SESPSGTAPGSTSST AESPGP		475 GGTACCTGCTCCTGAAGCGGTCCAATCTTCTACT GCACCAGGTTCTACTAGCGAATCCCCTACT GGTACTGGTCCAGGTTCCACCAGCTCTACT GCAGAATCTCCGGGTCCA	475
LCW0403_054_N_C05.ab1 GFP-	432 GSTSESPSGTAPGTS PSGESSTAPGSTSST AESPGP		476 GGTTCTACTAGCGAATCCCCTACTGGTACT GCTCCAGGTACTTCCCCTAGCGGTGAATCT TCTACTGCTCCAGGTTCTACCAGCTCTACCG CAGAATCTCCTGGGTCCA	476
LCW0403_057_N_D05.ab1 GFP-	433 GSTSSTAESPGPGST SESPSGTAPGTSPSG ESSTAP		477 GGTTCTACCAGCTCTACCGGTGAAATCTGTG GCTCCAGGTTCTACTAGCGAATCTCCGTCTG GCACCGGCAGCAGGTACTTCCCCTAGCGGTG AATCTTCTACTGGCAGCA	477
LCW0403_058_N_E05.ab1 GFP-	434 GSTSESPSGTAPGST SESPSGTAPGTSTPES GSASP		478 GGTTCTACTAGCGAATCTCCTTCTGGCACTG CACCAGGTTCTACCAGCGAATCTCCGTCTG GCACTGCAGCAGGTACCTCTACCCCTGAAA GCGGTTCCGGCTTCTCCA	478

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
LCW0403_060_GFP-N_F05.ab1	GTSTPESGSASPGSTSESPSGTAPGSTSSTAESPGP	435	GGTACCTCTACTCCGGAAACGGGTTCCCCATCTTCCAGGTTCTACCAGCGAATCCCGTCGTCTACTGGCACCGGCACCAGGTTCTACTAGCTCTACTGCTGAATCTCCGGGCCCA	479
LCW0403_063_GFP-N_G05.ab1	GSTSSTAESPGPGTSPSGESSTAPGTSPSGESSTAPGTSPSGESSTAP	436	GGTTCTACTAGCTCTACTGCAGAATCTCCGGGCCAGGTACCTCTTCTTAGCGGTGAATCTCTACCGCTCCAGGTACTTCTCCGAGGGGTGAATCTCTACCGCTCCA	480
LCW0403_064_GFP-N_H05.ab1	GTSPSGESSTAPGTSPSGESSTAPGTSPSGESSTAP	437	GGTACCTCCCTAGCGGCGAAATCTTCTACTGCTCCAGGTACCTCTTCTAGCGGCGAAACTCTACCGCTCCAGGTACTTCTCCGAGGGGTGAATCTCTACCGCTCCA	481
LCW0403_065_GFP-N_A06.ab1	GSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAP	438	GGTTCCACTAGCTCTACTGCTGAAATCTCCGGCCAGGTACTTCTACTGCTGGAAACGCCGTTCCCCTTCTCCAGGTTCTACTAGCGGAATCTCTCCGTCTGGCACCGCACCA	482
LCW0403_066_GFP-N_B06.ab1	GSTSSTAESPGPGTSPSGESSTAPGTSPSGESSTAP	439	GGTTCTACTAGCGCAATCTCCGTCTGGCACTGCTCCAGGTACTTCTCCTAGCGGCGAAACTCTACCGCTCCAGGTACTTCTCCGAGGGGTGAATCTCTACCGCTCCA	483
LCW0403_067_GFP-N_C06.ab1	GSTSESPSGTAPGTSTPESGSASPGSTSSTAESPGP	440	GGTTCTACTAGCGGAATCTCCTCTGTGGCACCGCTCCAGGTACTTCTCCTAGCGGCGAAAGCGGCTTCCGCTTCTCCAGGTTCCACTAGCTCTACCGCTGAATCTCCGGGCCCA	484
LCW0403_068_GFP-N_D06.ab1	GSTSSTAESPGPGTSTSSTAESPGPGSTSESPSGTAP	441	GGTTCCACTAGCTCTACTGGTGAATCTCCTGGCCCAGGTACTCTACCAGCTCTGAATCTCCTGGCCCAGGTTCTACCAGCGGGAATCTCCGTGTGGCACCGCACCA	485
LCW0403_069_GFP-N_E06.ab1		442		486

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
	GSTSESPSGTAPGTS TPESGSASPGTSTPES GSASP		GGTTCTACTAGCGAATCCCCGTCTGGTACC GCACCAGGTACTTCTACCCCGGAAAGCGGC TCTGCTTCTCCAGGTACTTCTACCCCGGAAA GCGGCTCCGCATCTCCA	
LCW0403_070_ GFP-N_F06.ab1	GSTSESPSGTAPGTS TPESGSASPGTSTPES GSASP	443	GGTTCTACTAGCGAAATCCCCGTCTGGTACT GCTCCAGGTACTTCTACTCCTGAAAGCGGT TCCGCTTCTCCAGGTACCTCTACTCCGGAA AGCGGTTCTGCATCTCCA	487

Ejemplo 4: Construcción de segmentos XTEN_AG36

Se construyó una biblioteca de codones que codificaba secuencias de 36 aminoácidos de longitud. Las secuencias se denominaron XTEN_AG36. Sus segmentos tienen la secuencia de aminoácidos $[X]_3$ donde X es un péptido 12mero con la secuencia: GTPGSGTASSSP (SEQ ID NO: 14), GSSTPSGATGSP (SEQ ID NO: 15), GSSPSASTGTGP (SEQ ID NO: 16) o GASPGTSSTGSP (SEQ ID NO: 17). El inserto se obtuvo hibridando los siguientes pares de pares de oligonucleótidos sintéticos fosforilados:

AG1for: AGGTACYCCKGGYAGCGGTACYGCWTCTTCYTCTCC (SEQ ID NO: 488)

AG1rev: ACCTGGAGARGAAGAWGCRGTACCGCTRCCMGGRGT (SEQ ID NO: 489)

AG2for: AGGTAGCTCTACYCCKTCTGGTGCWACYGGYTCYCC (SEQ ID NO: 490)

AG2rev: ACCTGGRGARCCRGTWGCACCAGAMGGRGTAGAGCT (SEQ ID NO: 491)

AG3for: AGGTTCTAGCCCKTCTGCWTCYACYGGTACYGGYCC (SEQ ID NO: 492)

AG3rev: ACCTGGRCCRGTACCRGTRGAWGCAGAMGGGCTAGA (SEQ ID NO: 493)

AG4for: AGGTGCWTCYCCKGGYACYAGCTCTACYGGTTCTCC (SEQ ID NO: 494)

AG4rev: ACCTGGAGAACCRGTAGAGCTRGTRCCMGGRGAWGC (SEQ ID NO: 495)

También hibridamos el oligonucleótido fosforilado 3KpnIstopperFor: AGGTTCGTCTTCACTCGAGGGTAC (SEQ ID NO: 232) y el oligonucleótido no fosforilado pr_3KpnIstopperRev: CCTCGAGTGAAGACGA (SEQ ID NO: 233). Los pares de oligonucleótidos hibridados se ligaron, lo que dio como resultado una mezcla de productos con longitud variable que representa el número variable de repeticiones de 12meros ligadas a un segmento BbsI/KpnI. Los productos correspondientes a la longitud de 36 aminoácidos se aislaron de la mezcla mediante una electroforesis en gel de agarosa preparativa y se ligaron en el vector de relleno pCW0359 digerido con BsaI/KpnI. La mayoría de los clones en la biblioteca resultante denominada LCW0404 mostraron fluorescencia verde después de la inducción, lo que muestra que la secuencia de XTEN_AG36 se había ligado en marco con el gen GFP, y la mayoría de las secuencias XTEN_AG36 muestra una buena expresión.

Cribamos 96 aislados de la biblioteca LCW0404 en busca de un alto nivel de fluorescencia estampándolos en una placa de agar que contenía IPTG. Los mismos aislados se evaluaron mediante una PCR y se identificaron 48 aislados que contenían segmentos con 36 aminoácidos, así como una fluorescencia fuerte. Estos aislados se secuenciaron y se identificaron 44 clones que contenían segmentos XTEN_AG36 correctos. Los nombres de archivo de las construcciones de nucleótidos y aminoácidos y las secuencias para estos segmentos se recogen en la tabla 14.

Tabla 14: Secuencias de ADN y de aminoácidos para motivos de 36-meros

Nombre del archivo		Secuencia aminoácidos	de	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
LCW0404_007_ N_A07.ab1	GFP-			496		540
		GASNKGPSSTGSPGLSGP SGTASSSPGTSSTPNGGA			GGTGGCGCATCCCTCGGGCGCACTGGCTCTACTGGCGGT GTTCTTGGAGGTGCTCCTGGCTAGCGGCTGGCTGC	
		TGSP			TGCTTCTTCTTCTGCAGGTAGCTCTACTGCTC CTTCTGGTGCTACTGGTTCTGCA	
LCW0404_003_ N_B07.ab1	GFP-			497		541
		GSSTPSGASPGTSSTGSP SASTGTGPGSSTPSGA			GGTAGCTCTACTGGTGGTCTAACTGG GCTTCTTCTAGGTTCTAGCGTTCCTTCTACTTCT CACCGGTAGCGGTCCTAGCGTTACTGGTACC	
		TGSP			CTTCTGGTGCTACTGGTTCTGCA	
LCW0404_006_ N_C07.ab1	GFP-			498		542
		GASNKGPSSTGSPGLSGP SASTGTGPGSSTPSGA			GGTGGCGCATCCCTCGGGCGCACTGGCTCTACTGGCGGT GTTCTTGGAGGTGCTCCTGGCTAGCGGCTGGCTGC	
		TGSP			AGTAGCACTGGCGGTCCTAGCGTTACTGGTACC GGTCCTGGTTCTCCTACTGGTTCTGCA	
LCW0404_007_ N_D07.ab1	GFP-			499		543
		GPNSGTASSSNGSST PSGATGSPGASPGTSS			GGTCCTAACTCTGGTGCTGCTTCTTCTAACGGTTCTTCT CCTTCTGGTGCTTCTCCTGGTACTAGCTCTACTGGTACT	
		GTOP			GGTGCTACTGGTCCTACTGGTTCTGCA	
LCW0404_008_ N_E07.ab1	GFP-			500		544
		GTPGSGTASSSPGASP GTSSTGSPNSSPGAST			GGTACTCCTGGTTCTGGTGCTGCTTCTTCTCCTGGTGCT CCTTCTGGTAGCGGTCCTGGTGCTTCTACTTCTCCTAGC	
		GTOP			TCTGGTCCTAACTCTTCTCCTGGTACTGCTTCT	
LCW0404_011_ N_F07.ab1	GFP-			501		545
		GASKPSSTPSGHGAST PSGATGSPGASPGTSS			GGTGCTAACTCTCCTTCTGGTCATGGTGCTACTTCTCCT CCTTCTGGTGCTTCTCCTGGTACTAGCTCTACTGGTACT	
		TGSP			GGTGCTACTGGTCCTACTGGTTCTGCA	
LCW0404_012_ N_G07.ab1	GFP-			502		546
		GTPGSGTASSSPGSST PSGATGSPGSSTPSGA			GGTACTCCTGGTTCTGGTGCTGCTTCTTCTCCTGGTTCT CCTTCTGCTAGCGGTCCTAGCGTTACTTCTCCTAGC	
		TGSP			TCTGGTCCTAACTCTTCTCCTGGTACTGCTTCT	

Nombre del archivo		Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
LCW0404_001_N_A07.ab1	GFP-	GASPGTSSTGSPGTPG SGTASSSPGSSTPSGA	498	GGTGCAATCCCCGGGCACTAGCTCTACCG GTTCTCCAGGTACTCCTGGTAGCGGTAC	540
		TGSP		TGCTTCTTCTCTCCAGGTAGCTCTACTC CTTCTGGTGCTACTGGTTCTCCA	
LCW0404_003_N_B07.ab1	GFP-	GSSTPSGATGSPGSSP SASTGTGPGSSTPSGA TGSP	497	GGTAGCTCTACCCCTTCTGGTGCTACCG GCTCTCCAGGTTCTAGCCCTTCTGCTTC TACCGGTACCGTTCCAGGTAGCTCTACC CCTTCTGGTGCTACTGGTTCTCCA	541
LCW0404_006_N_C07.ab1	GFP-	GASPGTSSTGSPGSSP SASTGTGPGSSTPSGA TGSP	496	GGTGCAATCTCCGGGTACTAGCTCTACCG GTTCTCCAGGTTCTAGCCATTCTGCTTCC ACTGGTACCGGTCCCAGGTAGCTCTACC CGTCTGGTGCTACTGGTTCTCCA	542
LCW0404_007_N_D07.ab1	GFP-	GTPGSGTASSSPGSST PSGATGSPGASPGTSS TGSP	499	GGTACTACCGGGCAGCCGGTACTGCTTCTT CCTCTCCAGGTAGCTCTACCCTTCTGG TGCAACTGGTTCCCCAGGTGCATCCCCT GGTACTAGCTCTAGCCGGTTCTCCA	543
LCW0404_009_N_E07.ab1	GFP-	GTPGSGTASSSPGASP GTSSTGSPGSDPSAST GTGP	500	GGTAGCCCTGGCAGCCGGTACTGCTTCTT CTTCTCCAGGTGCTTCCCTGGTACCAG CTCTAGCGGTTCTCCAGGTTCTAGACCT GGTGCATCCAGCGGTACTGGTTCTA	544
LCW0404_011_N_F07.ab1	GFP-	GASPGTSSTGSPGSST PSGATGSPGASPGTSS TGSP	501	GGTGCAATCTCCTGGTACCAGCTCTACCG GTTCTCCAGGTAGCTCTACTCCTTCTGG TGCTACTGGCTCTCCAGGTGCTTCCCCG GGTACCAGCTCTACCGGTTCTCCA	545
LCW0404_012_N_G07.ab1	GFP-	GTPGSGTASSSPGSST PSGATGSPGSSTPSGA TGSP	502	GGTACCCCCGGGCAGCGGTACTGCTTCTT CCTCTCCAGGTAGCTCTACCCCTTCTGG TGCTACCGGTTCCCCAGGTAGCTCTACC CCCTCTGGTGCTACTCCTTCTCCCA	546

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
LCW0404_014_N_H07.ab1	GFP- GASPGTSSTGSPGASP GTSSTGSPGASPGTSS TGSP	503	GGTGCATCTCCGGGCACTAGCTCTACTG GTTCTCCAGGTGCATCCCCTGGCACTAG CTCTACTGGTTCTCCAGGTGCTTCTGCTG GTACCAGCTCTACTGGTTCTCCA	547
LCW0404_015_N_A08.ab1	GFP- GSSTPSGATGSPGSSP SASTGTGPGASPGTSS TGSP	504	GGTAGCTCTACTGGTCTGGTGGAAACC GCTCCTCAGGTTCTAGCCTGTCTGCTTC CATGGTACTGGCCCAGGTGCTTCCCCG GGCACCAGCTCTACTGGTTCTCCA	548
LCW0404_016_N_B08.ab1	GFP- GSSTPSGATGSPGSST PSGATGSPGPTPGSGT ASSSP	505	GGTAGCTCTACTCTCTGGTGCTACCG GTTCCCCAGGTAGCTGTACTCCTTCTGG TGCTACTGGTTCCCCAGGTACTCCCGGGC ACCGGTACTGCTCTTCCTCTTCA	549
LCW0404_017_N_C08.ab1	GFP- GSSTPSGATGSPGSST PSGATGSPGASPGTSS TGSP	506	GGTAGCTCTACTGGTTCTGGTGGAAACC GTTCCCTAGGTAGCTCTACTCCTTCTGG TGCTACTGGCTCCCCAGGTGATCCCCT GGCACCAGCTCTACCGGTTCTCCA	550
LCW0404_018_N_D08.ab1	GFP- GTPGSGTASSSPGSSP SASTGTGPGSSTPSGA TGSP	507	GGTACTCTGGTAGCGGTACCGCATCTT CCTCTCCAGGTTCTAGCCCTTCTGCATCT ACCGGTACCGGTCCAGGTAGCTCTACTC CTTCTGGTGGTACTGGACTGTCTCA	551
LCW0404_023_N_F08.ab1	GFP- GASPGTSSTGSPGSSP SASTGTGPGTPGSGT ASSSP	508	GGTGCTTCCCCAGGCACTAGCTCTACCG GTTCTCCAGGTTCTAGCCCTTCTGCATCT ACTGGTACTGGCCCAGGTACTCCGGGCA GGGGTACTGCTCTTCTCTTCTCA	552
LCW0404_025_N_G08.ab1	GFP- GSSTPSGATGSPGSST PSGATGSPGASPGTSS TGSP	509	GGTAGCTCTACTCTCTGGTGCTACCG GCTCTCCAGGTAGCTCTACTCCTTCTGG TGCAACCGGCTCCCCAGGTGCTCTCCG GGTACCAGCTCTACTGGTTCTCCA	553
LCW0404_029_N_A09.ab1	GFP-	510		554

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
	GTPGSGTASSPGSST PSGATGSPGSSPSAST GTGP		GGTACCGCTGGCAGCGGTACCGGTCTCTT CCTCTCCAGGTAGCTCTACCTCCGTCTGG TGCTACTGGCTCTCCAGGTTCTAGCCCG TCTGCATCTACCAGTACCGGCCCA	
LCW0404_030_GFP-N_B09.ab1	511			555
	GSSTPSGATGSPGTTG SGTASSPGTPGSGTA SSSP		GGTAGCTCTACTCCTCTGGTGCAACCG GCTCCCTAGGTACCCCGGGCAGCGGTA CCGCATCTCCTCTCCAGGTACTCCGGC TAGCGGTACTGCTCTTCTTCTCCA	
LCW0404_031_GFP-N_C09.ab1	512			556
	GTPGSGTASSPGSST PSGATGSPGSSPSAST TGSP		GGTACCCTCGGGTAGCGGTACTCTCTTT CCTCTCCAGGTAGCTCTACCCTTCTGG TGCAACCGGCTCTCCAGGTCCTTCTCG GGCATCTAGCTCTACTGGTTCTCA	
LCW0404_034_GFP-N_D09.ab1	513			557
	GSSTPSGATGSPGSST PSGATGSPGASPGTSS TGSP		GGTAGCTCTACCCCTCTGGTGCTACCG GCTCTCCAGGTAGCTCTACCCCGTCTGG TGCAACCGGCTCTCCAGGTGGCATCTCCG GGTACTAGCCTCTACCCGGTCTCCA	
LCW0404_035_GFP-N_E09.ab1	514			558
	GASPGTSSTGSPGTPG SGTASSPGSSTPSGA TGSP		GGTGCTTCTCCGGGCAGCAGCTCTACCG GTTCTCCAGGTACCCCGGGCAGCGGTAC CGCATCTTCTCCAGGTAGCTCTACT CCTTCTGGTGCAACCGGTCTCCA	
LCW0404_036_GFP-N_F09.ab1	515			559
	GSSPASTCTGPGSST PSGATGSPGTPGSGT ASSSP		GGTTCTAGCCCCGTCTGCTTCCACCGGTA CTGGCCCAGGTAGCTCTACCCGTCTGG TGCAACTGGTTCTCCCAGGTACCCCTGGT AGCGGTACCGCTCTCTTCTTCTCCA	
LCW0404_037_GFP-N_G09.ab1	516			560
	GASPGTSSTGSPGSSP SASTGTGPGSSTPSGA TGSP		GGTGCTTCTCCGGGCAGCAGCTCTACTG GTTCTCCAGGTTCTAGCCCTTCTGCATC CACCGGTACCGGTTCCAGGTAGCTCTACC CCTTCTGGTGCAACCGGTCTCCA	
LCW0404_040_GFP-N_H09.ab1	517			561
	GASPGTSSTGSPGSST PSGATGSPGSSTPSGA TGSP		GGTGCATCCCCTCTACCAGCTCTACCG GTTCTCCAGGTAGCTCTACCCCGTCTGG TGCTACCGGCTCTCCAGGTAGCCTCTACC CCGTCTGGTGCTACTGGTCTCCA	

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
LCW0404_041_N_A10.ab1	GFP- GTPGSGTASSPGGSST PSGATGSPGSTPSGGT ASSSF	518	GGTACCCTGGTAGCGGTACTGCTTCTT CCTCTCCAGGTAGCTCTACTCCGGTCTGG TGCTACCGGTTCTCCAGGTACCCCGGGT AGCGGTACCGCAICTCTTCTCCA	562
LCW0404_043_N_C10.ab1	GFP- GSSPSASTGTGPGSST PSGATGSPGSTPSGA TGSP	519	GGTTCTAGCCCTTCTGCTTCTACCGGTA CTGGCCAGGTAGCTCTACCCCTTCTGG TGCTACCGGCTCTCCAGGTAGCTCTACT CCTTCTGGTGGTACTGGTACTGGTTCCA	563
LCW0404_045_N_D10.ab1	GFP- GASPGTSSTGSPGSSP SASTGTGPGSSPSAST GTGP	520	GGTGCTTCTCCTGGCACCAGGTCTACTG GTTCTCCAGGTTCTAGCCCTTCTGCTTCT ACCGGTACTGGTCCAGGTTCTAGCCCTT CTGCATCCACTGGTACTGGTTCCA	564
LCW0404_047_N_F10.ab1	GFP- GTPGSGTASSSPGGASP GTSSTPSGASPGTSS TGSP	521	GGTACTCCTGGCAGCGGTACCGCTTCTT CTTCTCCAGGTGCTCTCCTGGTACTAG CTCTACCGGTTCTCCAGGTGCTCTCCCG GGCACTAGCTCTACTGGTTCTCCA	565
LCW0404_048_N_G10.ab1	GFP- GSSTPSGATGSPGASP GTSSTGSPGSSTPSGA TGSP	522	GGTAGCTCTACCCCGTCTGGTGCTACCG GTTCTCCAGGTGCTCTCCTGGTACTAG CTCTACCGGTCCAGGTAGCTCTACC CCGTCTGGTGCTACTGGTCTCTCCA	566
LCW0404_049_N_H10.ab1	GFP- GSSTPSGATGSPGTPG SGTASSSPGSSTPSGA TGSP	523	GGTAGCTCTACCCCGGCTCTGGTGCTACTG GTTCTCCAGGTACTCCGGGCAGCGGTAC TGACTCTTCCTCTCCAGGTAGCTCTACC CTTCTGGTGGCTACTGGCTCTCCA	567
LCW0404_050_N_A11.ab1	GFP- GASPGTSSTGSPGSSP SASTGTGPGSSTPSGA TGSP	524	GGTGCTTCTCCTGGTACCAGCTCTACTG GTTCTCCAGGTTCTAGCCCTTCTGCTTCT ACCGGTACCGGTCTAGGTAGCTCTACTC CTTCTGGTGGTACCGGTTCTCCA	568

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
LCW0404_051_N_B11.ab1	GFP-	525		569
	GSSTPSGATGSPGSSST PSGATGSPGSSTPSGA		GGTAGCTCTACCCCGTCTGGTGGTACTG GCTCTCCAGGTAGCTCTACTCCTTCTGG	
	TGSP		TGCTACTGGTTCCCCAGGTAGCTCTACC CCGTCTGGTGCAACTGGCTCTCCA	
LCW0404_052_N_C11.ab1	GFP-	526		570
	GASPGTSSTGSPGTPG SGTASSSPGASPGTSS TGSP		GGTGCATCCCCGGGTACCAGCTCTACCG GTTCTCCAGGTACTCCTGGCAGCGGTAC TGCATCTTCTCTCCAGGTGCTTCTCCG GGCACCAGCTCTACTGGTTCTCCA	
LCW0404_053_N_D11.ab1	GFP-	527		571
	GSSTPSGATGSPGSSP SASTGTGPGASPGTSS TGSP		GGTAGCTCTACCCCTCTGGTGGCAACC GTTCTCCAGGTCTAGCCCGGTCTGCATC CACTGGTACCGGTCCAGGTGCTTCCGCT GGCACCAGCTCTACTGGTTCTCCA	
LCW0404_057_N_E11.ab1	GFP-	528		572
	GASEGTSSTGSPGSSI PSGATGSPGSSPSAST GTGP		GGTGCATCCGCTGGTACTAGCTCTACCC GTTCTCCAGGTAGCTCTACTCCGTCTGG TGCAGCCGGCTCTCCAGGTCTAGCCCT TCTGGATCTACCGGTACTGGTCCA	
LCW0404_060_N_F11.ab1	GFP-	529		573
	GTPGSGTASSSPGSSI PSGATGSPGASPGTSS TGSP		GGTACTCCTGGCAGCGGTACCGCATCTT CTTCTCAGGTAGCTCTACTCCGTCTGG TGCAACTGGTTCCCCAGGTGCTTCTCCG GGTACCAGCTCTACCGGTTCTCCA	
LCW0404_062_N_G11.ab1	GFP-	530		574
	GSSTPSGATGSPGTPG SGTASSSPGSSTPSGA TGSP		GGTAGCTCTACCCCGCTCTGGTGGCAACCG GCTCCCCAGGTACTCCTGGTAGCGGTAC CGCTTCTTCTCTCCAGGTAGCTCTACTC CGTCTGGTGGCTACCGGCTCCCCA	
LCW0404_066_N_H11.ab1	GFP-	531		575
	GSSTSASTGTGPGSSP SASTGTGPGASPGTSS TGSP		GGTTCTAGCGCCTTCTGCAATCCACCGGTA CCGGCCCAGGTTCTAGCCCTGTCGTTC TACCGGTACTGGTTCCAGGTGCTCTCCG GGTACTAGCTCTACTGGTTCTCCA	

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
LCW0404_067_N_A12.ab1 GFP-	GTKGSGTASSPGSSTPSGATGSPGSNPSASTGTGP	532	GGTACCCGGGTACCGGTACCGGTACCGTTCTTCTTCCAGGTAGCTCTACTCCGTCTGGCTACCGGCTCTCCAGGTTCTAACCTTTCTGCATACCGGTACCGGTTCCAGGGTACCGGTCCA	576
LCW0404_068_N_B12.ab1 GFP-	GSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGP	533	GGTTCTAGCTCCTTCTCCACATCTACTGGTACTGGCCAGGTAGCTCTACTCCTTCTCCTGGTGCTACCGGCTCTCCAGGTGCTTCTCCGGGTACTGGTTCTCCA	577
LCW0404_069_N_C12.ab1 GFP-	GSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSP	534	GGTAGCTCTACCCCTTCTGGTGCAACCGGCTGTCCAGGTGCATCCCGGGTACCAGACTCCCGGGTACCAGAGGTACCGCTTCTCCTCCTCTCA	578
LCW0404_070_N_D12.ab1 GFP-	GSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGA	535	GGTAGCTCTACTCCGTCTGGTGCAACCGGTTCCCAGGGTAGCTCTACCCCTTCTGGTGCAACCGGTACCCTTCTCCTCCTCTCA	579
LCW0404_073_N_E12.ab1 GFP-	GASPGTSSTGSPGTPGSGTASSPGSSTPSGATGSPGSSTPSGA	536	GGTGCTTCTCCTGGCACTAGCTCTACCGGTTCTCCAGGTACCCCTGGTAGCGGTACCGGTTCCTCCTCTCCAGGTACTGGTCCA	580
LCW0404_075_N_F12.ab1 GFP-	GSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGP	537	GGTAGCTCTACCCCGTCTGGTGCTACTGGCGCTCCCAGGTCTAGCCCTTCTGCATCCACCGGTACCCGGTCCAGGTTCTAGCCCGTCTGCATACTGGTACCGGTCCA	581
LCW0404_080_N_G12.ab1 GFP-	GASPGTSSTGSPGSSPSASTGTGPGTGP	538	GGTGCTTCCCCGGGCATCCAGCTCTACTGGGTTCTCCAGGTTCTAGCCCGTCTGCTTCTACTGGTCCAGGTTCTAGCCCTTCTGCACTGGTACCGGTTCCA	582
LCW0404_081_N_H12.ab1 GFP-		539		583

Nombre del archivo	Secuencia de aminoácidos	SEQ ID NO:	Secuencia de nucleótidos	SEQ ID NO:
	GASPGTSSTGSPGSSP SASTGTGPGTPGSGT ASSSP		GGTGCTTCCCCGGGTACCAGCTCTACCG GTTCTCCAGGTTCTAGCCCTTCTGCTTCT ACCGGTACCGGTCCAGGTACCCCTGGCA GCGGTACCGCATCTTCCTCTCCA	

Ejemplo 5: Construcción de XTEN_AE864

XTEN_AE864 se construyó a partir de la dimerización en serie de XTEN_AE36 a AE72, 144, 288, 576 y 864. Se construyó una colección de segmentos XTEN_AE72 a partir de 37 segmentos diferentes de XTEN_AE36. Se mezclaron cultivos de *E. coli* que portaban los 37 segmentos diferentes de 36 aminoácidos y se aisló el plásmido. Esta fracción de plásmidos se digirió con BsaI/NcoI para generar el fragmento pequeño como inserto. Se digirió la misma fracción de plásmidos con BbsI/NcoI para generar el fragmento grande como vector. Los fragmentos del inserto y el vector se ligaron dando como resultado una duplicación de la longitud, y la mezcla de ligación se transformó en células BL21Gold(DE3) para obtener colonias de XTEN_AE72.

Esta biblioteca de segmentos XTEN_AE72 se denominó LCW0406. Todos los clones de LCW0406 se combinaron y dimerizaron de nuevo utilizando el mismo proceso como se ha descrito anteriormente, lo que produjo la biblioteca LCW0410 de XTEN_AE144. Todos los clones de LCW0410 se combinaron y dimerizaron de nuevo utilizando el mismo proceso como se ha descrito anteriormente, lo que produjo la biblioteca LCW0414 de XTEN_AE288. Dos aislados, LCW0414.001 y LCW0414.002, se recogieron aleatoriamente de la biblioteca y se secuenciaron para verificar las identidades. Todos los clones de LCW0414 se combinaron y dimerizaron de nuevo utilizando el mismo proceso como se ha descrito anteriormente, lo que produjo la biblioteca LCW0418 de XTEN_AE576. Se cribaron 96 aislados de la biblioteca LCW0418 en busca de un alto nivel de fluorescencia de la GFP. Se secuenciaron 8 aislados con tamaños correctos de insertos mediante una PCR y fluorescencia fuerte, y se eligieron 2 aislados (LCW0418.018 y LCW0418.052) para un uso futuro tomando como base los datos de secuenciación y de expresión.

El clon específico pCW0432 de XTEN_AE864 se construyó combinando LCW0418.018 de XTEN_AE576 y LCW0414.002 de XTEN_AE288 utilizando el mismo proceso de dimerización como se ha descrito anteriormente.

Ejemplo 6: Construcción de XTEN_AM144

Se construyó una colección de segmentos XTEN_AM144 a partir de 37 segmentos diferentes de XTEN_AE36, 44 segmentos de XTEN _AF36 y 44 segmentos de XTEN_AG36.

Se mezclaron cultivos de *E. coli* que portaban los 125 segmentos diferentes de 36 aminoácidos y se aisló el plásmido. Esta fracción de plásmidos se digirió con BsaI/NcoI para generar el fragmento pequeño como inserto. Se digirió la misma fracción de plásmidos con BbsI/NcoI para generar el fragmento grande como vector. Los fragmentos del inserto y el vector se ligaron dando como resultado una duplicación de la longitud, y la mezcla de ligación se transformó en células BL21Gold(DE3) para obtener colonias de XTEN_AM72.

Esta biblioteca de segmentos XTEN_AM72 se denominó LCW0461. Todos los clones de LCW0461 se combinaron y dimerizaron de nuevo utilizando el mismo proceso como se ha descrito anteriormente, lo que produjo la biblioteca LCW0462. Se cribaron 1512 aislados de la biblioteca LCW0462 para determinar la expresión de proteínas. Las colonias individuales se transfirieron a placas de 96 pocillos y se cultivaron durante la noche como cultivos de inicio. Estos cultivos de inicio se diluyeron en medio de autoinducción reciente y se cultivaron durante 20-30 h. La expresión se midió utilizando un lector de placas de fluorescencia con excitación a 395 nm y emisión a 510 nm. 192 aislados mostraron expresión de alto nivel, y se sometieron a una secuenciación del ADN. La mayoría de los clones en la biblioteca LCW0462 mostraron una buena expresión y unas propiedades fisicoquímicas similares, lo que sugiere que la mayoría de las combinaciones de segmentos XTEN_AM36 producen secuencias XTEN útiles. Se eligieron 30 aislados de LCW0462 como una colección preferida de segmentos XTEN_AM144 para la construcción de proteínas multifuncionales que contienen múltiples segmentos XTEN. Los nombres de archivo de las construcciones de nucleótidos y aminoácidos y las secuencias para estos segmentos se recogen en la tabla 15.

Tabla 15: Secuencias de ADN y de aminoácidos para los segmentos AM144

Clon	Secuencia recortada	SEQ ID NO:	Secuencia de proteínas	SEQ ID NO:
LDW462_f1		564		617
LDW462_f5		565		618
LDW462_f9		566		619
LDW462_f10		567		620

Clon	Secuencia recortada	SEQ ID NO:	Secuencia de proteínas	SEQ ID NO:
	GTTGTCCAGGTAGCTCTACTCGTTCTGGTGCAAAC GGTTCCCCAGGTAGCTCTACTCCTTCTGGTGCTACT GGCTCCCCACGTGCATCCCCTGGCACCAGCTCTAC CGGTTCCCA		GSPGSSTPSGATG SPGASPGTSSTGS P	
LCW462_115	GGTGGCTTCTCCGGGCACCAGCTCTACTGGTTCTCC AGGTTCTACGTCCTTCTGCACTCCACCGGTACCGGTC CAGGTACGTCTACTCCTTCTGGTGCAACCGGTCTCT CCAGGTACTTCTGCAAAGCGTACCCCGGAATCTGG CCCAGGTAGCGCAACCGGCTACTTCTGGTTCTGAAA CCCCAGGTAGCGCAACCGGTACTTCTGGTTCTGAA ACTCCACGTACTTCTGAAAGCGCTACTCCGCAGTC CGGTCCAGGTACCTCTACCGCAACCGTCCGAAGGCA GCGCTCCAGGTACTTCTACTGAACCTTCTGAGGGT AGCGCTCCAGGTACTTCTACCGAACCGTCCGAGGG TAGCGCACCAGGTACCCTACCGAACCGTCTGAGG GTAGCGCCCCAGGTAGCGAACCGCGCAACTCCGGT TCTGAAACTCCA	588	GASPGTSSTGSPG SSPSASTGTGPGS STPSGATGSPGTS ESATPESGPGSEP ATSGSETPGTSES TSGSETPGTSESA TPESGPGTSTEPS EGSAPGTSTEPSE GSAPGTSTEPSEG SAPGTSTEPSEGS APGSEPATSGSET P	621
LCW462_116	GGTACCTCTACCGAACCTTCCGAAGGTAGCGCTCC AGGTAGCCCCGCAGGTTCCTACTTCCACTGAGG AAGGTACTTCCTACCGAACCTTCTGAGGGTAGCGCA CCAGGTACTCTGAAAGCGCAACTCCTGAGTCTGG CCCAGGTAGCGAACCGCAACCCCGGAATCT CTCCAGGTACCTCTGAAAGCGCAACCCCGGAATCT GGTCCAGGTAGCCCCGGCTGGCTCTCCTACCTCTAC TGAAGGAAGGTACTTCTGAAAGCGCTACTCCTGAGT CTGGTCCAGGTACTTCTACTGAACCGTCCGAAGGT AGCGCTCCAGGTAGCGAACCTGCTACTTCTGGTTC TGAAACTCCAGGTAGCGAACCGTCCTACTTCTGGT GTAGCGCTCCAGGTAGCGAACCTGCTACTTCTGGT TCTGAAACTCCA	589	GTSTEPSEGSAPG SPAGSPTSTEEGT STEPSEGSAPGTS ESATPESGPGSEP ATSGSETPGTSES ATPESGPGSPAGS PTSTEEGTSESAT PESGPGTSTEPSE GSAPGSEPATSGS ETPGTSTEPSEGS APGSEPATSGSET P	622
LCW462_120	GGTACTTCTACCGAACGGTCTCGAAGGCAGCGCTCC AGGTACCTCTACTGAACCTTCCGAGGGCAGCGCTC CAGGTACCTCTACCGAACCTTCTGAAGGTAGCGCA CCAGGTACTTCTACTCGAACCGTCTGAAGGCAGCGC TCCAGGTACTTCTACTGAACCTTCTGAGGGCAGCGC CTCCAGGTACTCTACCGAACCTTCTGAAGGCAGCGC GCACCAGGTACTTCTACCGAACCTTCCGAGGGCAG CGCACCAGGTACTCTGTGAAAGCGCTACCCTGAAGT CCGGCCCAGGTACTTCTGAAAGCGCTACTCCTGAA TCCGGTCCAGGTACTTCTACTGAACCTTCCGAAGG TAGCGCTCCAGGTAGCGAACCTGCTACTTCTGGTT CTGAAACTCCAGGTAGCCCCGGCTGGCTCTCCTGATC TCCACCGAAGAA	590	GTSTEPSEGSAPG TSTEPSEGSAPGT STEPSEGSAPGTS TEPSEGSAPGTST EPSEGSAPGTSTE PSEGSAPGTSTEP SEGSAPGTSTEPS EGSAPGTSTEPSE GSAPGTSTEPSEG SAPGTSTEPSEGS APGSEPATSGSET P	623
LCW462_123		591		624

Clon	Secuencia recortada	SEQ ID NO:	Secuencia de proteínas	SEQ ID NO:
	GGTACTTCTACCGAAGGTGCGAGGGCAGCGCTCC AGGTACTTCTACTGAACCTTCTGAAGGCAGCGCTC CAGGTACTTCTACTGAACCTTCGGAAGGTAGCGCA CCAGGTTCTACCAGCGAATCCCTTCTGGTACTGC TCCAGGTTCTACCAGCGAATCTCCCTTCTGGCACCG CACCAGGTACTTCTACCCTGAAAGCGGCTCCGCT TCTCCAGGTAGCGAACCTGCAACCTCTGGCTCTGA AACCCCAGGTACCTCTGAAAGGGCTACTCCTGAAT CTGGTCCAGGTACTTCTACTGAACCGTCCGAACGGC AGCGCACCAGGTACTTCTACTGAACCGTCTGAACG TAGTCGCACCAGGTATCTTCTGAAAGCCGCAACCCCGG AATCCGGTCCAAGGTACCTCTGAAAGCGCAACCCCCG GAGTCCGGCCCA		GTSTEPSEGSAPG TSTEPSEGSAPGT STEPSEGSAPGST SESPSGTAPGSTS ESPSGTAPGTSTP ESGSASPGSEPAT SGSETPGTSESAT PESGPGTSTEPSE GSAPGTSTEPSEG SAPGTSESATPES GPGTSTESATPESG P	
LCW462_r24	GGTACCTCTACCCCTTCTGGTGCTACCGGCTCTCC AGGTTCTACCCCGGTCTGCTTCTACCGGTACCGGTC	592	GSSTPSGATGSPG SSPSASTGTGPGS	625
	CAGGTAGCTCTACCCCTTCTGGTGCTACTGGGTTCTC CAGGTAGCCCTGCTGGCTCTCCCGACTTCTACTGAG GAAGGTAGCCCCGCCTGGTGTTCTCCGACTTCTACTGA GGAAGGTACTTCTACCGAACCTTCCGAAGGTAGCG CTCCAGGTGCTTCCCCGGGCACTAGCTCTACCGGT TCTCCAGGTTCTAGCACCTTCTGCATCTACTGGTACT GGCCCAGGTACTCCGGCGCAGCGGTACTGCTTCTTC CTCTCCAGGTTCTACTAGCTCTACTGCTGAATCTCC TGGCCCAGGTACTTCTCCTAGCGGTGAATCTTCTA CCCGCTCCAGGTACCTCTACTCCGGAAAGCGGTTCT GCATCTCCA		STPSGATGSPGSP AGSPTSTEPGSPA GSPTSTEEGTSTE PSPGSAPGASPGT SSTGSPGSSPSAS TGTGSPGTPGSGT ASSSPGSTSSTAE SPGPGTSPSGESS TAPGTSTPESGSA SP	
LCW462_r27	GGTACCTCTACTGAACCTTCTGAGGGCAGCGCTCC AGGTACTTCTGAAAGCGCTACCCCGGAGTTCCGCTC CAGGTACTTCTACTGAACCGTCCGAAGGTAGCGCA CCAGGTACTTCTACTGAACCGTCTGAAGGTAGCGCC ACCAGGTACTTCTGAAAGCGCAACCCCGGAATCCG GCCCAGGTACCTCTGAAAGGCGCAACCCCGGAGTCC GGCCCAGGTACTCCTGGCAGCGGTACCGCTTCTTC TTCTCCAGGTGCTTCTCCTGGTACTAGCTCTACTGG TTCTCCAGGTGCTTCTCCGGGCACTAGCTCTACTG GTTCTCCAGGTAGCCCTGCTGGCTCTCCGACTTCTA CTGACGAAGGTAGCCCCGGTGGTTCTCCGACTTCT ACTGAGGAAGGTACTTCTACCGAACCTTCCGAAGG TAGCGCTCCA	593	GTSTEPSEGSAPG TSESATPESGPGT STEPSEGSAPGTS TEPSEGSAPGTSE SATPESGPGTSES ATPESGPGTPGSG TASSSPGASPGTS STGSPGASPGTSS TGSPGSPAGSPTS TEEGSPAGSPTST EEGTSTEPSEGSA P	626
LCW462_r28		594		627

Clon	Secuencia recortada	SEQ ID NO:	Secuencia de proteínas	SEQ ID NO:
	GGTAGCCCAGCAGGCTCTCCGACTTCCACTGAGGA AGGTACTTCTACTGAACCTTCCGAAGGCACGCGCAC CAGGTACCTCTACTGAACCTTCTGAGGGCAGCGCT CCAGGTACCTCTACCGAACCGTCTGAGGTAGCGC ACCAGGTACCTCTGAAAGCGCAACTCCTGAGTCCG GTCCAGGTACTTCTGAAAGCGCAACCCCGGAGTCT GGCCCAGGTACCCCGGGTAGCGGTACTGGTTCTTC CTCTCCAGGTAGCTTCTGGTGCAACCG GCTCTCCAGGTGCTTCTCCGGTGCACCAGTCTTACC GGTTCTCCAGGTACCTCTACTGAACTTCTGAGGG CAGCCCTCCAGGTACTTCTGAAAGCGCTACCCCGG AGTCCGGTCCAGGTACTTCTACTGAACCGTCCGAA GGTAGCGCAGCA		GSPAGSPTSTEEG TSTEPSEGSAPGT STEPSEGSAPGTS TEPSEGSAPGTSE SATPESGPGTSTS ATPESGPGTPGSG TASSSPGSSTPSG ATGSPGASPGTSS TGSPGTSTEPSEG SAPGTSESATPES GPGTSTEPSEGSA P	
LCW462_r38	GGTAGCGAACCGGCAACCTCCGGCTCTGAAACTCC AGGTACTTCTGAAAGCGCTACTCCCGGAATCCGGCC CAGGTAGCGAACCGGCTACTTCCGGCTCTGAAACC CCAGGTAGCTCTACCCCGTCTGGTGCAACCGGCTC CCCAGGTACTCCTGGTAGCGGTACCCGCTTCTTCTC TCCAGGTAGCTCTACTCCGTCTGGTGCTACCCGGCT CCCCAGGTGCATCTCCTGGTACCAGCTCTACCCGGT TCTCCAGGTAGCTCTACTCCTCTGGTGCTACTGGGC TCTCCAGCGTGCTCCCCCGGTAGCAGCTCTACCGGG TTCTCCAGGTAGCGAACCTGCTACTTCTGGTTCTCTG AAACTCCAGGTACTTCTACCCGAACCGTCCGAGGGT AGCCGTCCAGGTAGCGAACCTGCTACTTCTGGTTC TGAAACTCCA	585	GSEPATSGSETPG TSESATPESGPGS EPATSGSETPGSS TPSGATGSPGTPG SGTASSSPGSSTP SGATGSPGASPGT SSTGSPGSSTPSG ATGSPGASPGTSS TGSPGSEPATSGS ETPGTSTEPSEGS APGSEPATSGSET P	628
LCW462_r39	GGTACCTCTACTGAACCTTCCGAAGGCAGCGCTCC AGGTACCTCTACCGAACCGTCCGAGGCAGCGCA CCAGGTACTTCTGAAAGCGCAACCCCTGAATCCGG TCCAGGTAGCCCTGCTGGCTCTCGAACTTCTACTG AGGAAGCTACCTCCGGCTCTGGTTCTCTACTT GAGGAAGGTACTTCTACCGAACCTCTGAAGGTAG CGCTCCAGGTAGCCCGGCTGGTTCTCCGAACTTCCA CCGAGGAAGGTACCTCTACTGAACCTTCTGAGGGT	596	GTSTEPSEGSAPG TSTEPSEGSAPGT SESATPESGPGSP AGSPTSTEEGSPA GSPTSTEEGTSTE PSEGSAPGSPAGS PTSTEEGTSTEPS EGSAPGTSTEPSE	629
	AGCCGTCCAGGTACCTCTACTGAACCTTCTGAGGGT CAGCCCTCCAGGTGCTTCTCCCGGCACCAGCTCTA CTGGTTCTCCAGGTTCTAGCCCGTCTGCTTCTACTG GTACTGGTCCAGGTTCTAGCCCTTCTGCTTCCACTG GTACTGGTCCA	597	GSAPGASPGTSST GSPGSSPSASTGTG GPGSSPSASTGTG P	
LCW462_r41				630

Clon	Secuencia recortada	SEQ ID NO:	Secuencia de proteínas	SEQ ID NO:
	GGTAAGCTGTACCCGTCTGGTGGTGCTACCGGTTCCCC AGGTGCTTCTTCTGGTACTAGCTCTACCGGTTTCTCC AGGTAGCTCTACCCCGTCTGGTGCTACTGGCTCTC CAGGTAGCCCTCCTGGCTCTCCAACTCCCACCGAA GAAGGTACCTCTGAAAGCGCAACTCCTGAATCCG GCCTAGGTAGCGAACCGGCAACCTCCGGTTCTGAA ACCCCAGGTGCATGCTCCTGGTACTAGCTCTACTGG TTCTCCAGGTAGCTCTACTTCGTCTCTGCTGCAACCG GCTCTCCAGGTTCTAGCCCTTCTGCATCTACCGGTA CTGGTCCAGGTTCTACCAGCGAATCCCCTTCTGGT ACTGCTCCAGGTTCTACCAGCGAATCCCCTTCTGG CACCGCACCAGGTACTTCTACCCCTGAAAGCGCCT CCGCTTCTCCA		GSSTPSLATGSPG ASPGTSSTSGSPGS STPSGATGSPGSP AGSETSTEEGTSE SATPESGPGSEPA TSGSETPGTASPGT SSTGSPGSSSTPSG ATEGSPGSSPSAST GTAPGSTSESPSGT APGTSTPESGSAS P	
LCW462_f42	GGTTCTACCACGGAATTCCTTCTGGCACCGCTCC AGGTTCTACTAGCGGAATCCCGTCTGGTACCGCAC CAGGTACTTCTCCTAGCGGCGAATCTTCTACCGCA CCAGGTACCTCTGAAAGCGCTACTCCGGAGTCTGG CCCAGGTACTCTCTACTGAACCGCTCTGAGGGTACGG CTCCAGGTACTCTACTGAACCGGTCCCAAGGTACC GCACCTAGCTACCTCTACTGAACCTCTGAGGGCAG CGCTCCAGGTACTTCTGAAAGCGCTACTCCCGGAGT CCGGTCCAGGTACTTCTACTGAACCGTCCGAAGGT AGCGCACCAGGTAGCTCTACCCCTCCTGGTGCTAC CGGTTCCCCAGGTGCTTCTTCCTGGTACTAGCTCTAC CTGGCCTTCCA	588	GSTSESPSGTAPG STSSPSGTAPGST SPSGESSTAPGTS ESATPESGPGTST EPSPSGSAPGTSTP PSEGSAPGTSTEP SEGSAPGTSTEPSA PESGPGTSTEPSE GSAPGSSTPSGAT GSPGASPGTSSTG SPGSSTPSGATGS P	631
LCW462_f43	GGTTCTACTAGCTCTACTGCAGAATCTCCGGGCCC AGGTACCTTCCTAGCGGTGAATCTTCTACCGCTC CAGGTACTTCTCCAGCGGTGAATCTTCTACCGCT CCAGGTTCTACTAGCTCTACCGCTGAATCTCCGGG TCCAGGTTCTACCAGCTCTACTGCAGAATCTCCTG GCCCAGGTACTTCTACTCCGGAAAGCGGTTCCGGT TCTCCAGGTACTTCTCCTAGCGGTGAATCTTCTACC GCTCCAGGTTCTACCAGCTCTACTGCTGAATCTCCT GGCCCAGGTACTTCTACCCCGGAAAGCGGCTCCGC TTCTCCAGGTTCTACCAGCTCTACCGCTGAATCTCC TGGCCCAGGTTCTACTAGCGAATCTCCGTCTGGCA CCGCACCAGGTACTTCCCCTAGCGGTGAATCTCT ACTGCTACCA	589	GSTSSTAESPGPG TSPSGESSTAPGT SPSGESSTAPGST SSTAESPGPGSTS STAESPGPGTSTP ESGSASPGTSPSG ESSTAPGSTSSTA ESPGPGTSTPESG SASPGSTSSTAES PGPGSTSPSPSGT APGTSPSGESSTA P	632
LCW462_f45		600	GTSTPESGSASPG STSSPSGTAPGTS TSSTAESPGPGTS TEPSEGSAPGTST EPSEGSAPGTSTE ATPESGPGTSTEP TPESGPGTSTEPS EGSAPGTSTEPSE GSAPGTSESATPE SGPGTSTEPSEGS APGTSTEPSEGSA P	633

Clon	Secuencia recortada	SEQ ID NO:	Secuencia de proteínas	SEQ ID NO:
	GGTACCTCTACTCCGGAAAAGCCGTTCCCCATCTCC AGGTTCTACCAGCGAATCCCCGTCTGGCACCCCAC CAGGTTCTACTAGCTCTACTGCTGAATCTCCGGGC CCAGGTACCTCTACTGAACCTTTTGAAGCACGGC TCCAGGTACCTCTACTGAACCGTCTGAGGGCAGCG CACCAGGTACTTCTGAAAGCGGAACCCCTGAATCC GGTCCAGGTACCTCTGAAAGCGCTACTCCGCAGTC TGGCCCAGGTACCTCTACTGAACCGTCTGAGGGTA GCGCTCCAGGTACTTCTACTGAACCGTCCGAAGGT AGCGCACCAGGTACTTCTGAAAGGCACTACTCCGGA GTCCGGTCCAGGGTACCTCTACCCGAACCCTCCGAAG GCAGCGCTCCAGGGTACCTTCTACTGAACCTTCTGAG GGTAGCCCCTCCC			
LCW462_r47	GGTACCTCTACCGAACCGTCCGAGGGTAGCGCACC	601	GTSTEPSEGSAPG	634
	AGGTACCTCTACTGAACCGTCTGAGGGTAGCGCTC CAGGTAGCCAACCGGCAACCTCTCGGCTCTGAAACT CCAGGTACTTCTACTGAACCGTCTCAAGGTAGCCC ACCAGGTACTTCTGAAAGCGCAACCCCGGAATCCG GCCCAGGTACCTCTGAAAGCGCAACCCCGGAGTCC GGCCCAGGTGCATCTCCGGGTACTAGCTCTACCGG TTCTCCAGGTTCTAGCCCTTCTGCTTCCACTGGTAC CGGCCCAGGTAGCTCTACCCCGTCTGGTGCTACTG GTTCCCCAGGTACGCTCTACTCCTTCTGGTGCTACT GGGTCCCCA	602	TSTEPSEGSAPGS EPATSGSETPGTS TESEGSAPGTSE SATPESGPGTSIS ATPESGPGASPGT SSTGSPGSSPSAS TGTGPGSSTPSGA TGSPGSSTPSGAT GSPGSSTPSGATG SPGASPGTSSTGS P	635
LCW462_r54	GGTACCGAACCGGCAACCTCTGGCTCTGAAACTCC AGGTAGCGAACCTGCAACCTCTCGGCTCTGAAACCC CAGGTACTTCTACTGAACCTTCTGAGGGCAGCGCA CCAGGTAGCCAACCTGGCAACCTCTGGCTCTGAAAC CCCAGGTACCTCTGAAAGCGCTACTCCTGAATCTG GCCCAGGTACTTCTACTGAACCGTCCGAGGGCAGC GCACCAGGTAGCTCTACTCCGTCTGGTGCTACCGG CTCTCCAGGTAGCTCTACCCCTTCTGGTGCAACCG GCTCCCAGGTGCTTCTCCGGGTACCAGCTCTACT GGTTCTCCAGGTAGCTCTACCCCGTCTGGTGCTAC CGGTTCCCCAGGTGCTTCTCCTGGTACTAGCTCTAC CGGTTCTCCAGGTAGCTCTACCCCGTCTGGTGCTA CTGGCTCTCCA	603	GSEPATSGSETPG SEPATSGSETPGT STEPSEGSAPGSE PATSGSETPGTSE SATPESGPGTSTE PSEGSAPGSSTPS GATGSPGSSTPSG ATGSPGASPGTSS TGSPGSSTPSGAT GSPGASPGTSSTG SPGSSTPSGATGS P	636

Clon	Secuencia recortada	SEQ ID NO:	Secuencia de proteínas	SEQ ID NO:
	GGTACTTCTACCGAACCGTCCGAGGGCAGCGCTCC AGGTACTTCTACTGAACCTTCTGAAGGCAGCGCTCC CAGGTACTTCTACTGAACCTTCCGAAGGTAGCGCA CCAGGTACTTCTGAAAGCGCTACTCCGAGTCCCG TCCAGGTACTCTTACCGAACCGTCCGAAGGCAGCG CTCCAGGTACTTCTACTGAACCTTCTGAGGGTAGC GCTCCAGGTTCTACTAGCGAATCTCCGTCTGGCAC TGCTCCAGGTACTTCTCCTAGCCGGTGAATCTTCTAC CGCTCCAGGTACTTCTCCTAGCGCGAATCTTCTA CCGCTCCAGGTAGCCCGGCTGGCTCTCCTACCTCT ACTGAGGAAGGTACTTCTGAAAGCGCTACTCCTGA GTCTGGTCCAGGTACCTCTACTGAACCGTCCGAAG GTAGCGCTCCA	604	GTSTEPSEGSAPG SEPATSGSETPG PAGSPTSTEEGSP AGSPTSTEEGTSE SATPESGPGTSTE PSPGSAPGTSTEP SEGSAPGTSTEPS EGSAPGTSESATP ESGPGSSTPSGAT GSPGSSPSASTGT GPGASPGTSSTGS P	637
LCW462_r57	GGTACTTCTACTGAACCTTCCGAAGGTAGCGCTCC AGGTAGCGAACCTGCTACTTCTGGTTCTGAAACCC CAGGTAGCCCCGGCTGGCTCTCCGACCTCCACCGAG GAAGGTAGCCCCGGCAGCTCTCCGACCTCTACTGA GGAAGGTACTTCTGAAAGCGCAACCCCGGAGTCC GGCCCAGGTACCTCTACCGAACCGTCTGAGGGGCAG CGCACCAGGTACCTCTACTGAACCTTCCGAAGGCA GCGCTCCAGGTACCTCTACCGAACCGTCCGAGGGC AGCGGACCAGGTACTTCTGAAAGCGCAACCCCTCTGA ATCTGGTCCAGGTAGCTCTACTCCGTCTGTGTGCAA CCCGGCTCCCCAGGTTCTAGCCCGTCTGCTTCCACT GGTACTGGCCCAGGTGCTCCCCGGGCACCAGCTC TACTGGTTCTCCA	605	GSEPATSGSETPG SPAGSPTSTEEGT SESATPESGPGTS TEPSEGSAPGTST EPSPGSAPGTSTE ATPESGPGTSTPE SGSASPGSTSESP	638
	AACTCCAGGTTCTACCAGCGAAACCCCGTCTGGCA CCGCACCAGGTTCTACTAGCTCTACTGAAACCT CCGGCCTCAGGTACTTCTGAAAGCGCTACTCCGGGA GTCCGGTCCAGGTACCTCTACCGAACCGTCCGAAG GTAGCGCTCCAGGTACTTCTACTGAACCTTCTGAG GGTAGCGCTCCA	606	SCTAPGSTSSTAE SPGPGTSESATPE SGPGTSTEPSEGS APGTSTEPSEGSA P	639
LCW462_r64				

Clon	Secuencia recortada	SEQ ID NO:	Secuencia de proteínas	SEQ ID NO:
	GGTACTTCTACCGAACCGTCTGAGGCAGCGCTCC AGGTACTTCTACTGAACCTTCTGAAGGCAGCGCTC CAGGTACTTCTACTGAACCTTCTCGAAGGTAGCGCA CCAGGTACCTCTACCGAACCGTCTGAAGGTAGCGC ACCAGGTACCTCTGAAAGCGCAACTCTCTGAGTCCG GTCCAGGTACTTCTGAAAGCGCAAGCCTCGCAGTCT GGCCCAGGTACTCCTGGCAGTACCGCAGTCTC CTCTCCAGGTACTCTACTCCGTCTGGTGCAACTG GTTCCCCAGGTGCTTCTCCGGGTACCAGCTCTACC TCCTGGTCCAGGTACCTCTACTGCTGAATCTT CTACTGCTCCAGGTACCTCTACTGCTGAATCTT TCTGGTCTCCAA		GTSTPPSGSAPG TSTPPSEGSAPGT STEPSEGSAPGTS TEPSEGSAPGTSE SATPESGPGTSPS ATPESGPGTPGSG TASSPGSSTPPSG ATGSPASPGTSS TGSPGSTSSTAPS PGPGTSPSGESST APGTSTPESGSAS P	
LCW462_r67	GGTAGCGCCGGGGAGGCTCTCTCGACCTCTACTGAGGA AGTGTACTTCTGAAAGCGCAAGCCTCCGAGTCCTCTCC CAGGTACCTCTACCGAACCGTCTGAAGGGCAGCGCA CCAGGTACTTCTGAAAGCGCAAGCCTCTGAATCTGG TCCAGGTACTTCTACCGAACCGTCCCGAAGGTAGC GCACCAGGTACTCTACCGAACCGTCTGGTCTCCGAC CGAGGAAGGTACCTCTACTGAACCTTCTGAGGGTA GCGCTCCAGGTACCTCTACTGAACCTTCTGAAGGC AGCGCTCCAGGTACTTCTACCGAACCGTCCCGAGGG CAGCGCTCCAGGTACTTCTACTGAAGCTTCTGAAG GCAGCGCTCCAGGTACTTCTACTGAACCTTCCGAA GGTAGCGCACACCA	607	GSPAGSPTSTEEG TSESATPESGPGT STEPSEGSAPGTS ESATPESGPGSEP ATSGSETPGTSTE PSEGSAPGSPAGS PSTEPGTSTEPS EGSAPGTSTEPSE GSAPGTSTEPSEG SAPGTSTEPSEGS APGTSTTPESGSA P	640
LCW462_r69	GGTACTTCTACCGAGCCGGTGAAATCTCTACCGCACC ACGTTCTACTAGCTCTAGCGCTGAATCTCCGGGCC CAGGTACTTCTCGAGCCGGTGAATCTTCTACTGCT CCAGGTACCTCTGAAAGCGCTACTCCGGAGGTTGG CCCAGGTACCTCTACTGAACCGTCTGAGGGCAGCG CTCCAGGTACCTCTACTGAACCGTCCGAAAGCTAGC GCACCAGGTTCTAGCTCTCCTTCTGGTGCTACCG TGGCCCAGGTAGCTCTACTCCCTTCTGGTGCTACCG GCTCTCCAGGTGCTTCTCCGGGTACTAGCTCTACC GGTTCTCCAGGTACTCTACTCCGGAAAGCGGTTC CGCATCTCCAGGTACTTCTACTGCTGAATCTT CTACTGCTCCAGGTACCTCTAGCGGCGAATCT TCTACTGCTCCA	608	GTSPAGSETSTAPG SPSSPAESPGPGT SPSGESSTAPGTS ESATPESGPGTST EPSEGSAPGSPSA STGTPGSSTPSG ATGSPASPGTSS TGSPGTSTPESGS ASPGTSPSAPEST APGTSPSGESSTA P	641
LCW462_r70		609	GTSESATPESGPG TSTEPSEGSAPGT STEPSEGSAPGSP AGSPTSTEEGSPA GSPTSTEEGTSTE PSEGSAPGSPSPA STGTGPGSSTPSG ATGSPGSSTPSGA TGSPGSEPATSGS ETPGTSESATPES GPGSEPATSGSET P	642

Clon	Secuencia recortada	SEQ ID NO:	Secuencia de proteínas	SEQ ID NO:
	GGTACCTGTGAAAGCGCTACTCCGGAGTCTGGCCC AGGTACCTCTACTGCAACCGCTCTGAGGGTAGCGCTC CAGGTACTTCTACTGAACGGTCCGAAGGTAGCGCA CCAGGTAGCGCTCTGCTGCTCCGAACTTCTACTGA GGAAGGTAGCTCGGCTGGTTCTCGACTTCTACTG AGGAAGCGTACTTCTACCGAACTTCTCGAAGGTAGC GCTCCAGGTTCTAGCCCTTCTGCTTCCACCGGTACT GGCCCAGGTAGCTCTACCCTTCTGGTGCTACCGG CTCCCCAGGTAGCTCTACTCCTCTTCTGTTCCAACTG GCCTCCAGGTAGCGAACCGGCAACTTCCGGCTCT GAAACCCAAGTACTTCTGAAAGCGCTACTCCTCA CTCTGGCCTAGGTAGCGAACCTGCTACCTCTGGCT CTGAAACCCCA	610	GTSTEPSEGSAPG TSTEPSEGSAPGT STEPSEGSAPGSS TPSGATGSPGSSTP GTSSTGSPGSSTP SGATGSTGTSESA TPESGPGSEPATS GSETPGTSTEPSE GSAPGSTSESPSG TAPGSTSESPSGT APGTSTPESGSAS P	643
LCW462_r72				
	GGTACTTCTACTGAACCGTCTGGAAGCGCTCC AGGTACCTCTACTGAACTTCCGAGGGCAGCGCTC CAGGTACCTCTACCGAACTTCTGAAGGTAGCGCA CCAGGTAGCTTCTCCTCTGGTACTACCGGTTC TCCAGGTAGCTCTACCCGTCTGGTGCTACTGGCT TCTCCAGCTTCTCCTCTGGTACTACCGGTTC TCCAGGTAGCGAACCGTCTACTGCTGAATCTCCTGGC CCCAGGTAGCTCTACTCCTTCTGGTGCTACCGGCT CTCCAGGTAGCTCTACTCCTTCTGGTGCTACCGGCT CCCAGGTAGCTCTACTCCTTCTGGTGCTACCGGCT TCTCCAGGTAGCTTCTCCGGGTACTACTGGTT TCTCCAGGTAGCGAACCGGCAACCTCCGGCTCTGA AACCCTAGGTACCTCTTCGAAAGCGCTACTCCTCAAT CCGGCCCAGGTAGCCCGGCAGGTTCTCCGACTTCCC ACTGAGGAAGCGGTTCTACTAGCGAATCTCCTTCTGG CACTGCACCAAGGTTCTACCAAGCGAATCTCCCTCTG GCACTCCACCCAGGTACCTCTACCCCTGAAAGCGGT TCCCCTTCTCCC	611	GTSTPESGSASPG STSSTAESPGPGTS TSSTAESPGPGSS PSASTGTGPGSSST PSGATGSPGSSPAT SGSETPGTSESAT PESETPGSPAGSPT STPESGSTSESPSG TAPGSTSESPSGT APGTSTPESGSAS P	644
LCW462_r73				
		612	GSPAGSPTSTEEG TSESATPESGPGT STEPSEGSAPGST SESPSGTAPGSTS GSPSGTAPGTSPS GESSTAPGTSTEP SEGSAPGSPAGSP TSTEEGTSTEPSE GSAPGSPTAPTEG ETPGTSESATPES GPGSTEPSEGSA P	645
LCW462_r78				

Clon	Secuencia recortada	SEQ ID NO:	Secuencia de proteínas	SEQ ID NO:
	GGTAGCCCGGCTGGCTCTCCTACCTCTACTGAGGA AGGTACTTCTGAAGCCCTACTCCTGAGTCTGGTC CAGGTACCTCTCTACTGAACCGTCCGAAGGTACCGCT CCAGGTTCTACCAGCCAATCTCCTTCTGGCACCGC TCCAGGTTCTACTAGCGAATCCCCTCTGGTACCG CACCAGGTACTTCTCTAGCGGCGAATCTTCTACC GCACCAGGTACCTCTACCGAACCTTCCGAAGGTAG CGCTCCAGGTAGCCCGGCAGGTTCTCCTACTTCCA CTGAGGAAGGTACTTCTACCGAACCTTCTGAGGGT ACCCCACCAGGTAGCGAACTCTGGCACCCTCTGGCTC TGAAACCCCAGCTACCTCTGAAAGCGCTACTCCTG AATCTGGCCCAGGTACTTCTACTGAACCGTCCGAG GGCAGAGCACCA			
LCW462_r79	GGTACCTCTACCGAACCTTCCGAAGGTAGCGCTCC AGGTAGCCCGGCAGGTTCTCCTACTTCCACTGAGG AAGGTACTCTACCGAACCTTCTGAGGGTAGCGCA CCAGGTACCTCCCCTAGCGGCGAATCTTCTACTGC TCCAGGTACCTCTCCTAGCCGCGAATCTTCTACCG CTCCAGGTACCTCCCCTAGCCGTGAATCTTCTACC GCACCAGCTTCTACCAGCCGAATCCCCTTCTGGCA TGCTCCAGGTTCTACCAGCCGAATCCCCCTTCTGGCA CCGCACCAGGTACTTCTACCCCTGAAAGCGGCTCC GCTTCTCCAGGTAGCCGAACCTGCCAATCCTCTGGCTC TGAAACCCCAGGTACCTCTGAAAGCGCTACTCCTG AATCTGGCCCAGGTACTTCTACTGAACCGTCCGAG GGCAGGCACCA	613	GTSTEPSEGSAPG SPAGSPTSTEEGT STEPSEGSAPGTS PSGESSTAPGTSP SGESSTAPGTSPS GESSTAPGSTSES PSGTAPGSTSESP SGTAPGTSTPESG SASPGSSPATSGS ETPGTSESATPES GPGTSTEPSEGSA P	646
LCW462_r87	CGTAGCCAACCGGCAACCTCTGGCTCTCTGAAACCCC AGGTACCTCTCTGAAAGCGCTACTCCTGGAATCTGGTC CAGGTACTTCTCTGAAAGCGCTACTCCTGGAATCCTCGCT CCAGGTACTTCTCCGAGCGGTGAATCTTCTACCGC ACCAGGTTCTACTAGCTCTACCGCTGAATCTCCTGG GCCCAGGTACTTCTCCTGAGCCGGTGAATCTTCTACT	614	GSEPATSGSETPG TSESATPESGPGT SESATPESGPGTS PSGESSTAPGSTS STAESPGPGTSPS GESSTAPGSTSES	647
LCW462_r98	GCTCCAGGTTCTACTAGCGAATCCCCGTCTGGTAC TGCTCCAGGTACTTCCCTAGCGGTGAATCTTCTA CTGCTCCAGGTTCTACCAGCTCTACCGCAGAATCT CCGGGTCCAGGTAGCTCTACTCCGTCTGGTGCAAC CGGTTCCCCAGGTAGCTCTACCCCTTCTGGTGCA CCGGCTCCCCAGGTAGCTCTACCCCTTCTGGTGCA AACTGGCTCTCC	615	PSGTAPGTSPSGE SSTAPGSTSSTAE SPGPGSSTPSGAT GSPGSSTPSGATG SPGSSTPSGATGW LS	648

Clon	Secuencia recortada	SEQ ID NO:	Secuencia de proteínas	SEQ ID NO:
	GGTAGCCCTGCTGGCTCTCCGACTTCTACTGAGGA		GSPAGSPTSTEEG	
	AGGTAGCCCCGGCTGGTTCTCCGACTTCTACTGAGG		SPAGSPTSTEEGT	
	AAGGTACTTCTACCGAACCTTCCGAAGGTAGCCGCT		STEPSEGSAPGTS	
	CCAGGTACCTCTACTGAACCTTCCGAAGGCAGCCGC		TEPSEGSAPGTST	
	TCCAGGTACCTCTACCGAACCGTCCGAGGGCAGCCG		EPSEGSAPGTSES	
	CACCAGGTACTTCTGAAAGCGCAACCCCTGAATCC		ATPESGPGASPGT	
	GGTCCAGGTGCATCTCCTGGTACCAGCCTCTACCGG		SSTGSPGSSTPSG	
	TTCTCCAGGTAGCTCTACTCCTTCTGGTGCTACTGG		ATGSPGASPGTSS	
	CTCTCCAGGTGCTTCCCCGGGTACCAGCTCTACCG		TGSPGSSTPSGAT	
	GTTCTCCAGGTAGCTCTACCCCGTCTGGTGCTACT		GSPGTPGSGTASS	
	GGTTCTCCAGGTACTCCGGGCAGCGGTACTGCTTC		SPGSSTPSGATGS	
	TTCCTCTCCAGGTAGCTCTACCCCTTCTGGTGCTAC		P	
	TGGCTCTCCA			
LCW462_r89	GGTAGCTCTACCCCGTCTGGTGCTACTGGTTCTCC	616	GSSTPSGATGSPG	649
	AGGTACTCCGGGCAGCGGTACTGCTTCTTCCTCTC		TPGSGTASSPGS	
	CAGGTAGCTCTACCCCTTCTGGTGCTACTGGCTCTC		STPSGATGSPGSP	
	CAGGTAGCCCGGCTGGCTCTCCTACCTCTACTGAG		AGSPTSTEEGTSE	
	GAAGGTACTTCTGAAAGCGCTACTCCTGAGTCTGG		SATPESGPGTSTE	
	TCCAGGTACCTCTACTGAACCGTCCGAAGGTAGCG		PSEGSAPGTSESA	
	CTCCAGGTACCTCTGAAAGCGCAACTCCTGAGTCT		TPESGPGSEPATS	
	GGCCCAGGTAGCGAACCTGCTACCTCCGGCTCTGA		GSETPGTSESATP	
	GACTCCAGGTACCTCTGAAAGCGCAACCCCGGAAT		ESGPGTSTEPSEG	
	CTGGTCCAGGTACTTCTACTGAACCGTCTGAAGGT		SAPGTSESATPES	
	AGCGCACCAGGTACTTCTGAAAGCGCAACCCCGG		GPGTSESATPESG	
	AATCCGGCCCAGGTACCTCTGAAAGCGCAACCCCG		P	
	GAGTCCGGCCCA			

Ejemplo 7: Construcción de XTEN_AM288

La biblioteca completa LCW0462 se dimerizó como se describe en el ejemplo 6, lo que dio como resultado una biblioteca de clones XTEN_AM288 denominados LCW0463. Se cribaron 1512 aislados de la biblioteca LCW0463 utilizando el protocolo descrito en el ejemplo 6. Se secuenciaron 176 clones de alta expresión y se eligieron 40 segmentos XTEN_AM288 preferidos para la construcción de proteínas multifuncionales que contienen múltiples segmentos XTEN con 288 restos de aminoácidos.

Ejemplo 8: Construcción de XTEN_AM432

Generamos una biblioteca de segmentos XTEN_AM432 recombinando segmentos de la biblioteca LCW0462 de segmentos XTEN_AM144 y segmentos de la biblioteca LCW0463 de segmentos XTEN_AM288. Esta nueva biblioteca del segmento XTEN_AM432 se denominó LCW0464. Los plásmidos se aislaron a partir de cultivos de *E. coli* que portaban LCW0462 y LCW0463, respectivamente. Se cribaron 1512 aislados de la biblioteca LCW0464 utilizando el protocolo descrito en el ejemplo 6. Se secuenciaron 176 clones de alta expresión y se eligieron 39 segmentos XTEN_AM432 preferidos para la construcción de XTEN más largos y para la construcción de proteínas multifuncionales que contienen múltiples segmentos XTEN con 432 restos de aminoácidos.

En paralelo, construimos la biblioteca LMS0100 de segmentos XTEN_AM432 utilizando los segmentos preferidos de XTEN_AM144 y XTEN_AM288. El cribado de esta biblioteca produjo 4 aislados que se seleccionaron para una construcción adicional

Ejemplo 9: Construcción de XTEN_AM875

El vector de relleno pCW0359 se digirió con BsaI y KpnI para eliminar el segmento de relleno, y el fragmento de vector resultante se aisló mediante purificación en gel de agarosa.

Hibridamos el oligonucleótido fosforilado BsaI-AscI-KpnIforP: AGGTGCAAGCGCAAGCGGCGCGCCAAGCACGGGAGGTTCGTCTTCACTCGAGGGTAC (SEQ ID NO: 650) y el oligonucleótido no fosforilado BsaI-AscI-KpnIrev: CCTCGAGTGAAGACGAACCTCCCGTGCTTGGCGCGCCGCTTGCGCTTGC (SEQ ID NO: 651) para introducir la isla de secuenciación A (SI-A) que codifica los aminoácidos GASASGAPSTG (SEQ ID NO: 652) y tiene la secuencia de nucleótidos de reconocimiento GGCGCGCC de la enzima de restricción AscI en su interior. Los pares de oligonucleótidos hibridados se ligaron con el vector de relleno pCW0359 digerido con BsaI and KpnI preparado anteriormente para producir pCW0466 que contenía SI-A. A continuación, generamos una biblioteca de segmentos XTEN_AM443 recombinando 43 segmentos XTEN_AM432 preferidos del ejemplo 8 y segmentos SI-A de pCW0466 en el extremo carboxílico utilizando el mismo proceso de dimerización descrito en el ejemplo 5. Esta nueva biblioteca de segmentos XTEN_AM443 se denominó LCW0479.

Generamos una biblioteca de segmentos XTEN_AM875 recombinando segmentos de la biblioteca LCW0479 de segmentos XTEN_AM443 y 43 segmentos XTEN_AM432 preferidos del Ejemplo 8 utilizando el mismo proceso de dimerización descrito en el ejemplo 5. Esta nueva biblioteca del segmento XTEN_AM875 se denominó LCW0481.

Ejemplo 10: Construcción de XTEN_AM1318

Hibridamos el oligonucleótido fosforilado BsaI-FseI-KpnIforP: AGGTCCAGAACCAACGGGGCCGGCCCCAAGCGGAGGTTCGTCTTCACTCGAGGGTAC (SEQ ID NO: 653) y el oligonucleótido no fosforilado BsaI-AscI-KpnIrev: CCTCGAGTGAAGACGAACCTCCGCTTGGGGCCGGCCCCGTTGGTTCTGG (SEQ ID NO: 654) para introducir la isla de secuenciación B (SI-B) que codifica los aminoácidos GPEPTGPAPSG (SEQ ID NO: 655) y tiene la secuencia de nucleótidos de reconocimiento GGCCGGCC de la enzima de restricción FseI en su interior. Los pares de oligonucleótidos hibridados se ligaron con el vector de relleno pCW0359 digerido con BsaI and KpnI como se utilizó en el ejemplo 9 para producir pCW0467 que contenía SI-B. A continuación, generamos una biblioteca de segmentos XTEN_AM443 recombinando 43 segmentos XTEN_AM432 preferidos del ejemplo 8 y segmentos SI-B de pCW0467 en el extremo carboxílico utilizando el mismo proceso de dimerización descrito en el ejemplo 5. Esta nueva biblioteca de segmentos XTEN_AM443 se denominó LCW0480.

Generamos una biblioteca de segmentos XTEN_AM1318 recombinando segmentos de la biblioteca LCW0480 de segmentos XTEN_AM443 y segmentos de la biblioteca LCW0481 de segmentos XTEN_AM875 utilizando el mismo proceso de dimerización que en el ejemplo 5. Esta nueva biblioteca del segmento XTEN_AM1318 se denominó LCW0487.

Ejemplo 11: Construcción de XTEN_AD864

Usando las diversas rondas consecutivas de dimerización, ensamblamos una colección de secuencias XTEN_AD864 a partir de segmentos XTEN_AD36 indicados en el ejemplo 1. Estas secuencias se ensamblaron como se describe en el ejemplo 5. Se evaluaron varios aislados de XTEN_AD864 y se averiguó que mostraban buena expresión y excelente solubilidad en condiciones fisiológicas. Se secuenció una construcción intermedia de XTEN _AD576. Este clon se evaluó en un experimento FC en macacos cangrejeros y se midió una semivida de aproximadamente 20 h.

Ejemplo 12: Construcción de XTEN_AF864

Usando las diversas rondas consecutivas de dimerización, ensamblamos una colección de secuencias XTEN_AF864 a partir de segmentos XTEN_AF36 indicados en el ejemplo 3. Estas secuencias se ensamblaron como se describe en el ejemplo 5. Se evaluaron varios aislados de XTEN_AF864 y se averiguó que mostraban buena expresión y excelente solubilidad en condiciones fisiológicas. Se secuenció una construcción intermedia de XTEN_AF540. Este clon se evaluó en un experimento FC en macacos cangrejeros y se midió una semivida de aproximadamente 20 h. Un clon completo de XTEN_AF864 tuvo una excelente solubilidad y mostró semivida superior a 60 h en macacos cangrejeros. Se ensambló un segundo conjunto de secuencias XTEN_AF que incluía una isla de secuenciación como se describe en el ejemplo 9.

Ejemplo 13: Construcción de XTEN_AG864

Usando las diversas rondas consecutivas de dimerización, ensamblamos una colección de secuencias XTEN_AG864 a partir de segmentos XTEN_AD36 indicados en el ejemplo 1. Estas secuencias se ensamblaron como se describe en el ejemplo 5. Se evaluaron varios aislados de XTEN_AG864 y se averiguó que mostraban buena expresión y excelente solubilidad en condiciones fisiológicas. Un clon completo de XTEN_AG864 tuvo una excelente solubilidad y mostró semivida superior a 60 h en macacos cangrejeros.

Ejemplo 14: Construcción de extensiones aminoterminales de XTEN- construcción y cribado de bibliotecas de adición de 12meros

En este ejemplo se detalla una etapa en la optimización del extremo amínico de la proteína XTEN para promover el inicio de la traducción para permitir la expresión de fusiones XTEN en el extremo amínico de proteínas de fusión sin la presencia de un dominio auxiliar. Históricamente, la expresión de proteínas con XTEN en el extremo amínico ha sido mala, produciendo unos valores que serían esencialmente indetectables en el ensayo de fluorescencia con GFP (<25 % de la expresión con el dominio auxiliar de CBD aminoterminal). Para crear diversidad a nivel de codones, se seleccionaron y prepararon siete secuencias de aminoácidos con una diversidad de codones. Se diseñaron siete pares de oligonucleótidos que codifican 12 aminoácidos con diversidades de codones, se hibridaron y se ligaron en el vector de relleno pCW0551 digerido por la enzima de restricción NdeI/BsaI (Stuffer-XTEN_AM875-GFP), y se transformaron en células competentes BL21Gold(DE3) de *E. coli* para obtener colonias de siete bibliotecas. Los clones resultantes tienen 12meros de XTEN aminoterminales fusionados en marco con XTEN_AM875-GFP para permitir el uso de la fluorescencia de la GFP para cribar la expresión. Se recogieron colonias individuales de las siete bibliotecas creadas y se cultivaron durante la noche hasta saturación en 500 µl de medio súper caldo en una placa de 96 pocillos profundos. El número de colonias recogidas variaba entre aproximadamente la mitad y un tercio de la diversidad teórica de la biblioteca (véase la tabla 16).

Tabla 16: Diversidad teórica y números de muestreo para las bibliotecas de adición de 12meros. Los restos de aminoácidos con codones aleatorizados están subrayados.

Biblioteca	Familia de motivos	Secuencia de aminoácidos	SEQ ID NO:	Diversidad teórica	Número cribado
LCW546	AE12	MASPAGSPTSTEE	656	572	2 placas (168)
LCW547	AE12	MATSESATPESGP	657	1536	5 placas (420)
LCW548	AF12	MATSPSGESSTAP	658	192	2 placas (168)
LCW549	AF12	MESTSSTAESPGP	659	384	2 placas (168)
LCW552	AG12	MASSTPSGATGSP	660	384	2 placas (168)
LCW553	AG12	MEASPGTSSTGSP	661	384	2 placas (168)
LCW554	(Similar al CBD)	MASTPESGSSG	662	32	1 placa (84)

Los cultivos saturados durante la noche se utilizaron para inocular cultivos recientes de 500 µl en medio de autoinducción en el que se cultivaron durante la noche a 26 °C. A continuación, estos cultivos de expresión se ensayaron utilizando un lector de placas de fluorescencia (excitación a 395 nm, emisión a 510 nm) para determinar la cantidad de GFP indicadora presente (véanse en la FIG. 11 los resultados de los ensayos de expresión). Los resultados, representados como gráficos de cajas y bigotes, indican que si bien los niveles de expresión medianos fueron aproximadamente la mitad de los niveles de expresión en comparación con el dominio auxiliar CBD aminoterminal "de referencia", los mejores clones de las bibliotecas estaban mucho más cerca de los de referencia, lo que indica que se justificaba una mayor optimización alrededor de esas secuencias. Esto está en contraste con versiones anteriores de XTEN que tenían <25 % de los niveles de expresión del CBD aminoterminal de referencia. Los resultados también muestran que las bibliotecas que comienzan con aminoácidos MA tenían mejores niveles de expresión que las que comienzan con ME. Esto era más evidente cuando se observaron los mejores clones, que estaban más cerca de los de referencia, ya que en su mayoría comienzan con MA. De los 176 clones dentro del 33 % del CBD-AM875 de referencia, el 87 % comienzan con MA, donde únicamente el 75 % de las secuencias en las bibliotecas que comienzan con MA, una clara sobrerrepresentación de los clones que comienzan con MA al nivel más alto de expresión. Se secuenciaron 96 de los mejores clones para confirmar la identidad, y se seleccionaron doce secuencias (véase la tabla 17), 4 de LCW546, 4 de LCW547 y 4 de LCW552 para una optimización adicional.

Tabla 17: Secuencias avanzadas de ADN de 12meros

Clon	Secuencia de ADN	SEQ ID NO:
LCW546_02	ATGGCTAGTCCGGCTGGCTCTCCGACCTCCACTGAGGAAGGTACTTCTACT	663
LCW546_06	ATGGCTAGTCCTGCTGGCTCTCCAACCTCCACTGAGGAAGGTACTTCTACT	664
LCW546_07	ATGGCTAGTCCAGCAGGCTCTCCTACCTCCACCGAGGAAGGTACTTCTACT	665
LCW546_09	ATGGCTAGTCCTGCTGGCTCTCCGACCTCTACTGAGGAAGGTACTTCTACT	666

Clon	Secuencia de ADN	SEQ ID NO:
LCW547_03	ATGGCTACATCCGAAAGCGCAACCCCTGAGTCCGGTCCAGGTACTTCTACT	667
LCW547_06	ATGGCTACATCCGAAAGCGCAACCCCTGAATCTGGTCCAGGTACTTCTACT	668
LCW547_10	ATGGCTACGTCTGAAAGCGCTACTCCGGAATCTGGTCCAGGTACTTCTACT	669
LCW547_17	ATGGCTACGTCCGAAAGCGCTACCCCTGAATCCGGTCCAGGTACTTCTACT	670
LCW552_03	ATGGCTAGTTCTACCCCGTCTGGTGCAACCGGTTCCCCAGGTACTTCTACT	671
LCW552_05	ATGGCTAGCTCCACTCCGTCTGGTGCTACCGGTTCCCCAGGTACTTCTACT	672
LCW552_10	ATGGCTAGCTCTACTCCGTCTGGTGCTACTGGTTCCCCAGGTACTTCTACT	673
LCW552_11	ATGGCTAGTTCTACCCCTTCTGGTGCTACTGGTTCTCCAGGTACTTCTACT	674

Ejemplo 15: Construcción de extensiones aminoterminales de XTEN- Construcción y cribado de bibliotecas con los codones 3 y 4 optimizados

En este ejemplo se detalla una etapa en la optimización del extremo amínico de la proteína XTEN para promover el inicio de la traducción para permitir la expresión de fusiones XTEN en el extremo amínico de proteínas sin la presencia de un dominio auxiliar. Con las preferencias establecidas para los dos primeros codones (véase el ejemplo, más arriba), el tercer y el cuarto codones se aleatorizaron para determinar las preferencias. Se diseñaron tres bibliotecas, basadas en los mejores clones de LCW546, LCW547 y LCW552, con el tercer y cuarto restos modificados de manera que todas las combinaciones de codones XTEN permitidos estuvieran presentes en estas posiciones (véase la FIG. 12). Con el fin de incluir todos los codones XTEN permitidos para cada biblioteca, se diseñaron nueve pares de oligonucleótidos que codificaban 12 aminoácidos con diversidades de los codones en los terceros y cuartos restos, se hibridaron y se ligaron en el vector de relleno pCW0551 digerido con las enzimas de restricción NdeI/BsaI (Stuffer-XTEN_AM875-GFP), y se transformaron en células BL21Gold(DE3) de *E. coli* competentes para obtener colonias de las tres bibliotecas LCW0569-571. Con 24 codones XTEN, la diversidad teórica de cada biblioteca es de 576 clones únicos. Se recogieron un total de 504 colonias individuales de las tres bibliotecas creadas y se cultivaron durante la noche hasta saturación en 500 µl de medio súper caldo en una placa de 96 pocillos profundos. Esto proporcionó cobertura suficiente para comprender el rendimiento relativo de la biblioteca y las preferencias de secuencia. Los cultivos saturados durante la noche se utilizaron para inocular cultivos nuevos de 500 µl en medio de autoinducción en el que se cultivaron durante la noche a 26 °C. A continuación, estos cultivos de expresión se ensayaron utilizando un lector de placas de fluorescencia (excitación a 395 nm, emisión a 510 nm) para determinar la cantidad de GFP indicadora presente. Los 75 clones principales del cribado se secuenciaron y se volvieron a ensayar para determinar la expresión de la GFP indicadora frente a las muestras de referencia (véase la FIG. 13). 52 clones produjeron datos de secuenciación utilizables y se utilizaron para análisis posteriores. Los resultados se desglosaron por biblioteca, e indican que LCW546 era la biblioteca superior. Los resultados se presentan en la tabla 18. Sorprendentemente, se descubrió que las lecturas de fluorescencia iniciales para los mejores clones eran de ~900 UA, mientras que el valor de referencia del CBD aminoterminal era de solo ~600 UA. Esto indica que esta biblioteca había instituido una mejora de aproximadamente un 33 % con respecto a los mejores clones de la biblioteca anterior, que eran aproximadamente iguales en expresión con respecto a la referencia del CBD aminoterminal (ejemplo 14).

Tabla 18: **Comparación de la biblioteca de optimización de los codones tercero y cuarto**

	LCW569	LCW570	LCW571
N	21	15	16
Fluorescencia media (UA)	628	491	537
DE	173	71	232
CV	28%	15%	43%

Se observaron nuevas tendencias en los datos que mostraban las preferencias por codones particulares en la tercera y cuarta posición. Dentro de la biblioteca LCW569, el codón de glutamato GAA en la tercera posición y el codón de treonina ACT se asociaron con una mayor expresión, como puede verse en la tabla 19.

Tabla 19: Codones tercero y cuarto preferidos en LCW569

	3 = GAA	Resto	4 = ACT	Resto
N	8	13	4	17
Fluorescencia media (UA)	749	554	744	601
DE	234	47	197	162
CV	31%	9%	26%	27 %

Adicionalmente, la repetición del ensayo de los 75 clones principales indicó que varios eran ahora superiores a los clones de referencia.

Ejemplo 16: Construcción de extensiones aminoterminales de XTEN- construcción y cribado de bibliotecas combinatorias de 12meros y 36meros

En este ejemplo se detalla una etapa en la optimización del extremo amínico de la proteína XTEN para promover el inicio de la traducción para permitir la expresión de fusiones XTEN en el extremo amínico de proteínas sin la presencia de un dominio auxiliar. Una vez establecidas las preferencias por los dos primeros codones (véase el ejemplo, más arriba), se analizó el extremo amínico en un contexto más amplio combinando las 12 secuencias de 12meros seleccionadas (véase el ejemplo, más arriba) en el propio extremo amínico, seguidas de 125 segmentos de 36meros construidos previamente (véase el ejemplo, más arriba) de forma combinatoria. Esto creó 48meros novedosos en el extremo amínico de la proteína XTEN y permitió evaluar el impacto de las interacciones de largo alcance en el extremo amínico sobre la expresión de las secuencias más largas (FIG. 14). De manera similar a los procedimientos de dimerización utilizados para ensamblar los 36meros (véase el ejemplo, a continuación), los plásmidos que contenían los 125 segmentos de 36meros seleccionados se digirieron con las enzimas de restricción BbsI/NcoI y el fragmento apropiado se purificó en gel. El plásmido del clon AC94 (CBD-XTEN_AM875-GFP) también se digirió con BsaI/NcoI, y los fragmentos apropiados se purificaron en gel. Estos fragmentos se ligaron entre sí y se transformaron en células de *E. coli* BL21Gold(DE3) competentes para obtener colonias de la biblioteca LCW0579, que también sirvió como vector para la clonación adicional de 12 12meros seleccionados en el propio extremo amínico. Los plásmidos de LCW0579 se digirieron con NdeI/EcoRI/BsaI y los fragmentos apropiados se purificaron en gel. Se diseñaron 12 pares de oligonucleótidos que codificaban 12 secuencias seleccionadas de 12meros, se hibridaron y se ligaron con el vector LCW0579 digerido con NdeI/EcoRI/BsaI, y se transformaron en células de *E. coli* BL21Gold(DE3) competentes para obtener colonias de la biblioteca LCW0580. Con una diversidad teórica de 1500 clones únicos, se recogieron un total de 1512 colonias individuales de la biblioteca creada y se cultivaron durante la noche hasta saturación en 500 µl de medio súper caldo en una placa de 96 pocillos profundos. Esto proporcionó cobertura suficiente para comprender el rendimiento relativo de la biblioteca y las preferencias de secuencia. Los cultivos saturados durante la noche se utilizaron para inocular cultivos nuevos de 500 µl en medio de autoinducción que se cultivaron durante la noche a 26 °C. A continuación, estos cultivos de expresión se ensayaron utilizando un lector de placas de fluorescencia (excitación a 395 nm, emisión a 510 nm) para determinar la cantidad de GFP indicadora presente. Los 90 clones principales se secuenciaron y se volvieron a ensayar para evaluar la expresión de la GFP indicadora. 83 clones produjeron datos de secuenciación utilizables y se utilizaron para análisis posteriores. Los datos de secuenciación se utilizaron para determinar el 12mero líder que estaba presente en cada clon, y se evaluó el impacto de cada 12mero en la expresión. Los clones LCW546_06 y LCW546_09 destacaron por ser el extremo amínico superior (véase la tabla 20).

Tabla 20: Rendimiento relativo de los clones que comienzan con LCW546 06 y LCW459 09

	LCW546_06	Todos los demás	LCW546_09	Todos los demás
N	11	72	9	74
Fluorescencia media (UA)	1100	752	988	775
DE	275	154	179	202
CV	25%	20%	18%	26%

La secuenciación y la repetición del ensayo también revelaron varios casos de replicados independientes de la misma secuencia en los datos que produjeron unos resultados similares, aumentando así la confianza en el ensayo. Adicionalmente, 10 clones con 6 secuencias únicas eran superiores al clon de referencia. Se presentan en la tabla 21. Se apreció que éstas fueron las únicas apariciones de estas secuencias, y en ningún caso se produjo que una de estas secuencias apareciera y no lograra superar al clon de referencia. Estas seis secuencias progresaron para una mayor optimización.

Tabla 21: Clones combinatorios de 12meros y 36meros superiores al clon de referencia

Nombre del clon	Primeros 60 codones	SEQ ID NO:	Nombre del 12mero	Nombre del 36mero
LCW580_51	ATGGCTAGTCCTGCTGGCTCTCCAACC TCCACTGAGGAAGGTGCATCCCCGGGC ACCAGCTCTACCGGTTCTCCAGGTAGC TCTACCCCGTCTGGTGCTACCGGCTCT CCAGGTAGCTCTACCCCGTCTGGTGCT	675	LCW546_06	LCW0404_040
	ACTGGCTCTCCAGGTACTTCTACTGAA CCGTCTGAAGGCAGCGCA			
LCW580_81	ATGGCTAGTCCTGCTGGCTCTCCAACC TCCACTGAGGAAGGTGCATCCCCGGGC ACCAGCTCTACCGGTTCTCCAGGTAGC TCTACCCCGTCTGGTGCTACCGGCTCT CCAGGTAGCTCTACCCCGTCTGGTGCT ACTGGCTCTCCAGGTACTTCTACTGAA CCGTCTGAAGGCAGCGCA	676	LCW546_06	LCW0404_040
LCW580_38	ATGGCTAGTCCTGCTGGCTCTCCAACC TCCACTGAGGAAGGTACTTCTACCGAA CCGTCCGAGGGTAGCGCACCAGGTAG CCCAGCAGGTTCTCCTACCTCCACCGA GGAAGGTACTTCTACCGAACCGTCCGA GGGTAGCGCACCAGGTACTTCTACTGA ACCGTCTGAAGGCAGCGCA	677	LCW546_06	LCW0402_041
LCW580_63	ATGGCTAGTCCTGCTGGCTCTCCGACC TCTACTGAGGAAGGTACTTCTACTGAA CCGTCTGAAGGCAGCGCACCAGGTAG CGAACCGGCTACTTCCGGTTCTGAAAC CCCAGGTAGCCCAGCAGGTTCTCCAAC TTCTACTGAAGAAGGTACTTCTACTGA ACCGTCTGAAGGCAGCGCA	678	LCW546_09	LCW0402_020
LCW580_06	ATGGCTAGTCCTGCTGGCTCTCCAACC TCCACTGAGGAAGGTACCCCGGGTAGC GGTACTGCTTCTTCCTCTCCAGGTAGCT CTACCCCTTCTGGTGCAACCGGCTCTC CAGGTGCTTCTCCGGGCACCAGCTCTA CCGGTTCTCCAGGTACTTCTACTGAAC CGTCTGAAGGCAGCGCA	679	LCW546_06	LCW0404_031

Nombre del clon	Primeros 60 codones	SEQ ID NO:	Nombre del 12mero	Nombre del 36mero
LCW580_35	ATGGCTAGTCCTGCTGGCTCTCCGACC TCTACTGAGGAAGGTACTTCTACTGAA CCGTCTGAAGGCAGCGCACCAGGTAG CGAACCGGCTACTTCCGGTTCTGAAAC CCCAGGTAGCCCAGCAGGTTCTCCAAC TTCTACTGAAGAAGGTACTTCTACTGA ACCGTCTGAAGGCAGCGCA	680	LCW546_09	LCW0402_020
LCW580_67	ATGGCTAGTCCTGCTGGCTCTCCGACC TCTACTGAGGAAGGTACCTCCCCTAGC GGCGAATCTTCTACTGCTCCAGGTACC TCTCCTAGCGGCGAATCTTCTACCGCT CCAGGTACCTCCCCTAGCGGTGAATCT TCTACCGCACCAGGTACTTCTACTGAA CCGTCTGAAGGCAGCGCA	681	LCW546_09	LCW0403_064
LCW580_13	ATGGCTAGTCCTGCTGGCTCTCCGACC TCTACTGAGGAAGGTACCTCTACTCCG GAAAGCGGTTCCGCATCTCCAGGTTCT ACCAGCGAATCCCCGTCTGGCACCGCA CCAGGTTCTACTAGCTCTACTGCTGAA TCTCCGGGCCCAGGTACTTCTACTGAA CCGTCTGAAGGCAGCGCA	682	LCW546_09	LCW0403_060
LCW580_88	ATGGCTAGTCCTGCTGGCTCTCCGACC TCTACTGAGGAAGGTACCTCCCCTAGC GGCGAATCTTCTACTGCTCCAGGTACC TCTCCTAGCGGCGAATCTTCTACCGCT	683	LCW546_09	LCW0403_064
	CCAGGTACCTCCCCTAGCGGTGAATCT TCTACCGCACCAGGTACTTCTACTGAA CCGTCTGAAGGCAGCGCA			
LCW580_11	ATGGCTAGTCCTGCTGGCTCTCCGACC TCTACTGAGGAAGGTACCTCTACTCCG GAAAGCGGTTCCGCATCTCCAGGTTCT ACCAGCGAATCCCCGTCTGGCACCGCA CCAGGTTCTACTAGCTCTACTGCTGAA TCTCCGGGCCCAGGTACTTCTACTGAA CCGTCTGAAGGCAGCGCA	684	LCW546_09	LCW0403_060

Ejemplo 17: Construcción de extensiones aminoterminales de XTEN- construcción y cribado de bibliotecas combinatorias de 12meros y 36meros para XTEN-AM875 y XTEN-AE864

En este ejemplo se detalla una etapa en la optimización del extremo amínico de la proteína XTEN para promover el inicio de la traducción para permitir la expresión de fusiones XTEN en el extremo amínico de proteínas sin la presencia de un dominio auxiliar. Con las preferencias para los cuatro primeros codones (véanse los ejemplos anteriores, y para el mejor emparejamiento de 12meros y 36meros aminoterminales (véase el ejemplo, más arriba) establecidas, se emprendió un enfoque combinatorio para analizar la unión de estas preferencias. Esto creó 48meros novedosos en el extremo amínico de la proteína XTEN y permitió ensayar la confluencia las de conclusiones previas. Adicionalmente, se evaluó la capacidad de estas secuencias líder de ser una solución universal para todas las proteínas XTEN colocando los nuevos 48meros delante tanto de XTEN-AE864 como de XTEN-AM875. En lugar de usar los 125 clones del segmento de 36meros, los plásmidos de 6 clones seleccionados del segmento de 36meros con la mejor expresión de GFP de la biblioteca combinatoria se digirieron con NdeI/EcoRI/BsaI, y los fragmentos apropiados se purificaron en gel. Los plásmidos de los clones AC94 (CBD-XTEN_AM875-GFP) y AC104 (CBD-XTEN_AE864-GFP) se digirieron con NdeI/EcoRI/BsaI, y los fragmentos apropiados se purificaron en gel. Estos fragmentos se ligaron entre sí y se transformaron en células de *E. coli* BL21Gold(DE3) competentes para obtener colonias de las bibliotecas LCW0585 (-XTEN_AM875-GFP) y LCW0586 (-XTEN_AE864-GFP), que también podrían servir como vectores para la clonación adicional de 8 12meros seleccionados en el propio extremo amínico. Los plásmidos de LCW0585 y LCW0586 se digirieron con NdeI/EcoRI/BsaI y los fragmentos apropiados se purificaron en gel. Se diseñaron 8 pares de oligonucleótidos que codifican 8 secuencias de 12meros seleccionadas con la mejor expresión de GFP en el cribado anterior (generación 2), se hibridaron y se ligaron con los vectores LCW0585 y LCW0586 digeridos con NdeI/EcoRI/BsaI, y se transformaron en células de *E. coli* BL21Gold(DE3) competentes para obtener colonias de las bibliotecas finales LCW0587 (XTEN_AM923-GFP) y LCW0588 (XTEN _AE912-GFP). Con una diversidad teórica de 48 clones únicos, se recogieron un total de 252 colonias individuales de las bibliotecas creadas y se cultivaron durante la noche hasta saturación en 500 µl de medio súper caldo en una placa de 96 pocillos profundos. Esto proporcionó cobertura suficiente para comprender el rendimiento relativo de la biblioteca y las preferencias de secuencia. Los cultivos saturados durante la noche se utilizaron para inocular cultivos nuevos de 500 µl en medio de autoinducción en el que se cultivaron durante la noche a 26 °C. A continuación, estos cultivos de expresión se ensayaron utilizando un lector de placas de fluorescencia (excitación a 395 nm, emisión a 510 nm) para determinar la cantidad de GFP indicadora presente. Los 36 clones principales se secuenciaron y se volvieron a ensayar para evaluar la expresión de la GFP indicadora. 36 clones produjeron datos de secuenciación utilizables, y estos 36 se utilizaron para análisis posteriores. Los datos de secuenciación determinaron el 12mero, el tercer codón, el cuarto codón y el 36mero presentes en el clon, y revelaron que muchos de los clones eran replicados independientes de la misma secuencia. Adicionalmente, los resultados de la repetición del ensayo para estos clones tienen un valor cercano, lo que indica que el proceso de cribado era robusto. Se observaron preferencias por ciertas combinaciones en el extremo amínico, que daban sistemáticamente unos valores de fluorescencia más altos, aproximadamente un 50 % más altos que los controles de referencia (véanse las tablas 22 y 23). Estos datos respaldan la conclusión de que la inclusión de las secuencias que codifican el XTEN aminoterminal optimizado en los genes de la proteína de fusión confería una notable mejora en la expresión de las proteínas de fusión.

Tabla 22: Combinaciones aminoterminales preferidas para XTEN-AM875

Nombre del clon	Número de réplicas	12mero	36mero	Media	DE	CV
CBD-AM875	NA	NA	NA	1715	418	16 %
LCW587_08	7	LCW546_06_3=GAA	LCW404_40	2333	572	18 %
LCW587_17	5	LCW546_09_3=GAA	LCW403_64	2172	293	10 %

Tabla 23: Combinaciones aminoterminales preferidas para XTEN-AE864

Nombre del clon	Número de réplicas	12mero	36mero	Media	DE	CV
AC82	NA	NA	NA	1979	679	24 %
LCW588_14	8	LCW546_06_opt3	LCW404_31	2801	240	6 %
LCW588_27	2	LCW546_06_opt34	LCW404_40	2839	556	15 %

Cabe destacar que la combinación preferida para el extremo amínico del XTEN-AM875 y la combinación preferida para el del XTEN-AE864 no es la misma (tablas 22 y 23), lo que indica que hay interacciones más complejas a más de 150 bases del sitio de inicio que influyen en los niveles de expresión. Las secuencias para las secuencias nucleotídicas preferidas se enumeran en la tabla 24, y los clones preferidos se analizaron mediante una SDS-PAGE para confirmar independientemente la expresión (véase la FIG. 15). Las secuencias completas de XTEN_AM923 y XTEN_AE912 se seleccionaron para análisis adicionales.

Tabla 24: Secuencias de ADN preferidas para los primeros 48 restos de aminoácidos de XTEN-AM875 aminoterminal y XTEN-AE864

Nombre del clon	XTEN modificado	Secuencia de nucleótidos	SEQ ID NO:
LCW587_08	AM875		685

Nombre del clon	XTEN modificado	Secuencia de nucleótidos	SEQ ID NO:
		ATGGCTGAACCTGCTGGCTCTCCAACCTCCACTGAGGAAGGTGCA TCCCCGGGCACCAGCTCTACGGTTCTCCAGGTAGCTCTACCCCGT CTGGTGCTACCGGCTCTCCAGGTAGCTCTACCCCGTCTGGTGCTAC TGGCTCTCCAGGTACTTCTACTGAACCG TCTGAAGGCAGCGCA	
LCW587_17	AM875	ATGGCTGAACCTGCTGGCTCTCCGACCTCTACTGAGGAAGGTACC TCCCCTAGCGGCGAATCTTCTACTGCTCCAGGTACCTCTCCTAGCG GCGAATCTTCTACCGCTCCAGGTACCTCCCTAGCGGTGAATCTTC TACCGCACCAGGTACTTCTACTGAACCGTCTGAAGGCAGCGCA	686
LCW588_14	AE864	ATGGCTGAACCTGCTGGCTCTCCAACCTCCACTGAGGAAGGTACC CCGGGTAGCGGTACTGCTTCTTCCTCTCCAGGTAGCTCTACCCCTT CTGGTGCAACCGGCTCTCCAGGTGCTCTCCGGGCACCAGCTCTAC CGGTTCTCCAGGTAGCCCGGCTGGCTCTCCTACCTCTACTGAG	687
LCW588_27	AE864	ATGGCTGAAACTGCTGGCTCTCCAACCTCCACTGAGGAAGGTGCA TCCCCGGGCACCAGCTCTACGGTTCTCCAGGTAGCTCTACCCCGT CTGGTGCTACCGGCTCTCCAGGTAGCTCTACCCCGTCTGGTGCTAC TGGCTCTCCAGGTAGCCCGGCTGGCTCTCCTACCTCTACTGAG	688

Ejemplo 18: Construcción de genes y vectores CTLA4-XTEN

Los genes del dímero CTLA4 que codifican CTLA4(1-120)-XTEN_AE42-CTLA4(3-120) y CTLA4(1-125)-XTEN_AE42-CTLA4(3-125) fueron diseñados y sintetizados por GeneArt, que introdujo sitios de restricción NdeI y BbsI que son compatibles con los sitios NdeI y BsaI que flanquean el CBD (dominio de unión a la celulosa) en el vector de destino CBD-XTEN. El plásmido CBD-XTEN es un derivado del pET30 de Novagen con el formato CBD-XTEN_AE864, donde CBD es el relleno para la clonación. Las construcciones se generaron reemplazando el CBD del vector CBD-XTEN por los fragmentos que codifican dímeros CTLA4. El plásmido CBD-XTEN presenta un promotor T7 secuencia arriba de la secuencia del CBD, y una secuencia XTEN_AE864 fusionada en marco secuencia abajo de la secuencia del CBD. El CBD de relleno se eliminó mediante una digestión de restricción utilizando las endonucleasas NdeI y BsaI. Los fragmentos de ADN del dímero CTLA4 digeridos con las endonucleasas de restricción NdeI y BbsI se ligaron en el vector CBD-XTEN digerido con NdeI and BsaI utilizando ligasa T4 de ADN y se electroporaron en BL21-Gold(DE3) (Stratagene). Los transformantes se cribaron mediante ADN miniprep y las construcciones deseadas se confirmaron mediante la secuenciación del ADN. Los plásmidos finales producen los dímeros CTLA4 con el conector XTEN_AE42 fusionado con los genes XTEN_AE864 bajo el control de un promotor T7. Las secuencias de ADN y de aminoácidos resultantes se enumeran a continuación. SEQ ID NO: 182 y 184.

El conector XTEN_AE42 incluido en los plásmidos de dímeros CTLA4-XTEN_AE864 se eliminó mediante una digestión de restricción utilizando las endonucleasas AscI y FseI flanqueantes. En el otro extremo, la secuencia XTEN_AE158 con los mismos sitios de restricción AscI y FseI flanqueantes construidos en otro plásmido se digirió con las endonucleasas AscI y FseI y se ligó en los plásmidos de los dímeros-XTEN_AE864 de CTLA4 digeridos con AscI y FseI anteriores utilizando ligasa T4 de ADN, y se electroporó en BL21-Gold(DE3) (Stratagene). Los transformantes se cribaron mediante ADN miniprep y las construcciones deseadas se confirmaron mediante la secuenciación del ADN. Los plásmidos finales producen los dímeros CTLA4 con el conector XTEN_AE158 fusionado con los genes XTEN_AE864 bajo el control de un promotor T7. Las secuencias de ADN y de aminoácidos resultantes se presentan en la tabla 25, a continuación.

Ejemplo 19: Construcción de genes y vectores aIL6R-XTEN

Ligasa de ADN y se electroporaron en BL21-Gold(DE3) (Stratagene). Los plásmidos producen los genes sscFv-XTEN_AE864 del aIL6R bajo el control del promotor T7. El plásmido, con un sitio BsaI adicional introducido mediante una PCR, se digirió adicionalmente con NdeI y BsaI, se ligó al fragmento XTEN_AE48 digerido con NdeI y BsaI utilizando ligasa T4 de ADN y se electroporó en BL21-Gold (DE3). El plásmido final produce el gen XTEN_AE48-aIL6R scFv-XTEN_AE864 bajo el control de un promotor T7. Todos los transformantes se cribaron mediante ADN miniprep y las construcciones deseadas se confirmaron mediante la secuenciación del ADN. Las secuencias de ADN resultantes y el producto final codificado se proporcionan a continuación. Los genes que codifican el scFv del aIL6R en el extremo carboxílico se amplificaron mediante una PCR que introdujo los sitios de restricción BsaI/HindIII y BsaI/BbsI e HindIII que son compatibles con los sitios BbsI e HindIII que flanquean la GFP (proteína fluorescente verde) en el vector XTEN-GFP de destino. El plásmido XTEN-GFP es un derivado del pET30 de Novagen en el formato XTEN_AE912-GFP, donde GFP es el relleno para la clonación. Las construcciones se generaron reemplazando la GFP del vector XTEN-GFP por los fragmentos de la PCR que codifican el scFv del aIL6R. El plásmido XTEN-GFP presenta un promotor T7 secuencia arriba de la secuencia XTEN_AE912 y una secuencia GFP de relleno fusionada en marco secuencia abajo de la secuencia XTEN_AE912. La GFP de relleno se eliminó mediante una digestión de restricción utilizando las endonucleasas BbsI e HindIII. Las endonucleasas de restricción BsaI e HindIII digirieron los fragmentos de la PCR del scFv del aIL6R, que se ligaron en el vector XTEN-GFP digerido con BbsI e HindIII utilizando ligasa T4 de ADN y se electroporaron en BL21-Gold(DE3) (Stratagene). Los plásmidos producen los genes del scFv del XTEN_AE912-aIL6R bajo el control del promotor T7. El plásmido, con un sitio BbsI adicional introducido mediante una PCR, se digirió adicionalmente con BbsI e HindIII, se ligó al fragmento XTEN_AE144 digerido con BsaI e HindIII utilizando ligasa T4 de ADN y se electroporó en BL21-Gold (DE3). El plásmido final produce el gen XTEN_AE912-aIL6R scFv-XTEN_AE144 bajo el control de un promotor T7. Todos los transformantes se cribaron mediante ADN miniprep y las construcciones deseadas se confirmaron mediante la secuenciación del ADN. Las secuencias de ADN resultantes y las secuencias de aminoácidos codificadas se proporcionan en la tabla 25, a continuación.

Ejemplo 20: Construcción de genes y vectores anti-CD40-XTEN y XTEN-anti-CD40 [00516] Construcción anti-CD40-XTEN

Se diseñaron dos genes que codificaban anti-CD40 mediante una traducción inversa combinada con optimización de codones. Estos genes que codifican anti-CD40 fueron sintetizados por GeneArt (Regensburg, Alemania), que introdujo los sitios de restricción NdeI y BbsI compatibles con los sitios NdeI y BsaI que flanquean el relleno en el vector pCBD-XTEN_AE864 de destino. El plásmido pCBD-XTEN_AE864 es un derivado del pET30 de Novagen. Las construcciones se generaron sustituyendo la secuencia del CBD en el pCBD-XTEN_AE864 por los fragmentos codificantes anti-CD40. El pCBD-XTEN_AE864 presenta un promotor T7 secuencia arriba del CBD, seguido de una secuencia del XTEN fusionada en marco. Los fragmentos de ADN del anti-CD40 digeridos por restricción se ligaron en el vector pCBD-XTEN_AE864 escindido utilizando ligasa T4 de ADN y se electroporaron en BL21(DE3) Gold (Stratagene, La Jolla, CA). Los transformantes se cribaron mediante ADN miniprep y la construcción deseada se confirmó mediante la secuenciación del ADN. Los vectores finales producen el gen anti-CD40-XTEN_AE864 bajo el control de un promotor

T7. Las construcciones resultantes son: AC384, pBC0009; AC385, pBC0010. Las secuencias de ADN resultantes y las secuencias de aminoácidos codificadas se proporcionan en la tabla 25, a continuación.

Construcción de XTEN-anti-CD40

Se amplificaron dos genes que codifican el anti-CD40 mediante la reacción en cadena de la polimerasa (PCR) utilizando los cebadores anti-CD40forBsaI (anti-CD40_1forBsaI: ATAAAGGGTCTCCAGGTGAAATTGTTCTGACCCAATCTCC (SEQ ID NO: 689); anti-CD40_2forBsaI: ATAAAGGGTCTCCAGGTGAAATTGTTCTGACTCAATCTCCA (SEQ ID NO: 690) y anti-CD40revHindIII (anti-CD40_1revHindIII: AACTCGAAGCTTttaGCTAGACACAGTAACCAGAGT (SEQ ID NO: 691); anti-CD40_2revHindIII: AACTCGAAGCTTttaAGAGGATACGGTCACCAGAGT (SEQ ID NO: 692)), que introdujeron los sitios de restricción BsaI e HindIII que son compatibles con los sitios BbsI e HindIII que flanquean el relleno en el vector XTEN_AE912 de destino. El plásmido XTEN_AE912-GFP es un derivado del pET30 de Novagen. Las construcciones se generaron sustituyendo la secuencia del GFP en el XTEN_AE912-GFP por los fragmentos codificantes anti-CD40. El XTEN_AE912-GFP presenta un promotor T7 secuencia arriba del XTEN, seguido de una secuencia del GFP fusionada en marco. Los fragmentos de GFP se eliminaron mediante una digestión de restricción utilizando las endonucleasas BbsI e HindIII. Los fragmentos de ADN del anti-CD40 digeridos por restricción se ligaron en el vector XTEN_AE912-GFP escindido utilizando ligasa T4 de ADN y se electroporaron en BL21(DE3) Gold (Stratagene, La Jolla, CA). Los transformantes se cribaron mediante ADN miniprep y la construcción deseada se confirmó mediante la secuenciación del ADN. Los vectores finales producen el gen XTEN_AE912-anti-CD40 bajo el control de un promotor T7. Las construcciones resultantes son: AC386, pBC0011; AC387, pBC0012. Las secuencias de ADN resultantes y las secuencias de aminoácidos codificadas se proporcionan a continuación.

Ejemplo 21: Construcción de genes y vectores anti-Her2-XTEN y XTEN-anti-Her2

Construcción de anti-Her2-XTEN

El gen que codifica el scFv anti-Her2 tiene el formato VL-XTEN_Y30-VH, donde VL es la cadena ligera del fragmento de anticuerpo anti-Her2, VH es la cadena pesada del fragmento de anticuerpo y XTEN_Y30 es la secuencia GSGEGSEGEGGGEGSEGEGSGEGGEGEGSG (SEQ ID NO: 1) flanqueada por los sitios de restricción AgeI y KpnI. El gen se sintetizó y se clonó en un vector comercial para usarlo como plantilla en la PCR. El gen que codifica el scFv anti-Her2 se amplificó mediante la reacción en cadena de la polimerasa (PCR) utilizando los cebadores anti-Her2forNdeI (agatatacatATGGAAGACATTCAGATGACCCAGAGC (SEQ ID NO: 693)) y anti-Her2revBsaI (CCGGGCTACCTGGAGACCCGGAAACAGTTACCAGAGTACC (SEQ ID NO: 694)), que introdujeron los sitios de restricción NdeI y BsaI que son compatibles con los sitios NdeI y BsaI que flanquean el relleno en el vector pCBD-XTEN_AE864 de destino. El plásmido pCBD-XTEN_AE864 es un derivado del pET30 de Novagen. Las construcciones se generaron sustituyendo la secuencia del CBD en el pCBD-XTEN_AE864 por los fragmentos codificantes anti-Her2. El pCBD-XTEN_AE864 presenta un promotor T7 secuencia arriba del CBD, seguido de una secuencia del XTEN fusionada en marco. Los fragmentos de ADN del anti-Her2 digeridos por restricción se ligaron en el vector pCBD-XTEN_AE864 escindido utilizando ligasa T4 de ADN y se electroporaron en BL21(DE3) Gold (Stratagene, La Jolla, CA). Los transformantes se cribaron mediante ADN miniprep y la construcción deseada se confirmó mediante la secuenciación del ADN. Los vectores finales producen el gen anti-Her2-XTEN_AE864 bajo el control de un promotor T7. Las construcciones resultantes son: pBC0007, SeqID 140. Las secuencias de ADN resultantes y las secuencias de aminoácidos codificadas se proporcionan a continuación.

Construcción de XTEN-anti-Her2

El gen que codifica el svFV anti-Her2 (descrito anteriormente) se amplificó mediante la reacción en cadena de la polimerasa (PCR) utilizando los cebadores anti-Her2forBbsI (GCACCAGGTTCGTCTTCACTCGACATTCAGATGACCCAGAGC (SEQ ID NO: 695)) y anti-Her2revHindIII (AACTCGAAGCTTTCAGGAAACAGTTACCAGAGTACCTTG (SEQ ID NO: 696)), que introdujeron los sitios de restricción BbsI e HindIII que son compatibles con los sitios BbsI e HindIII que flanquean el relleno en el vector XTEN_AE912 de destino. El plásmido XTEN_AE912-GFP es un derivado del pET30 de Novagen. Las construcciones se generaron sustituyendo la secuencia del GFP en el XTEN_AE912-GFP por el fragmento codificante anti-Her2. El XTEN_AE912-GFP presenta un promotor T7 secuencia arriba del XTEN, seguido de una secuencia del GFP fusionada en marco. El fragmento de GFP se eliminó mediante una digestión de restricción utilizando endonucleasas BbsI e HindIII. El fragmento de ADN anti-Her2 digerido por restricción se ligó en el vector XTEN_AE912-GFP escindido utilizando ligasa T4 de ADN y se electroporó en BL21(DE3) Gold (Stratagene, La Jolla, CA). Los transformantes se cribaron mediante ADN miniprep y la construcción deseada se confirmó mediante la secuenciación del ADN. Los vectores finales producen el gen XTEN_AE912-anti-Her2 bajo el control de un promotor T7. Las construcciones resultantes son: pBC0008. Las secuencias de ADN resultantes y el producto final codificado se proporcionan en la tabla 25, a continuación.

Ejemplo 22: Construcción de genes y vectores anti-EGFR-XTEN

El gen que codifica el anti-EGFR se amplificó mediante la reacción en cadena de la polimerasa (PCR) a partir de una biblioteca, que introdujo sitios de restricción NdeI y BbsI que son compatibles con los sitios NdeI y BsaI que flanquean la primera etiqueta FLAG en el vector de destino FLAG-Y50-FLAG-HIS6 ("His6" divulgada como la SEQ ID NO: 218). Las construcciones se generaron reemplazando la secuencia FLAG del FLAG-Y50-FLAG-His6 ("His6" divulgada como la SEQ ID NO: 218) por los fragmentos que codifican el anti-EGFR. El FLAG-Y50-FLAG-His6 ("His6" divulgada como la SEQ ID NO: 218) presenta un promotor T7 secuencia arriba de la FLAG, seguido por la secuencia del Y50-FLAG-His6 ("His6" divulgada como la SEQ ID NO: 218) fusionada en marco. Los fragmentos de ADN del anti-EGFR digeridos por restricción se ligaron en el vector FLAG-Y50-FLAG-His6 escindido ("His6" divulgada como la SEQ ID NO: 218) utilizando ligasa T4 de ADN y se electroporaron en XL1 Blue. Los transformantes se cribaron mediante ADN miniprep y las construcciones deseadas se confirmaron mediante la secuenciación del ADN. Los vectores finales producen el gen anti-EGFR-Y50-FLAG-His6 ("His6" divulgada como la SEQ ID NO: 218) bajo el control de un promotor T7. La construcción resultante es: pMS0120. Las secuencias de ADN resultantes y las secuencias de aminoácidos codificadas se proporcionan en la tabla 25, a continuación.

Ejemplo 23: Construcción de genes y vectores anti-CD3-XTEN

El gen que codifica el anti-CD3 se amplificó mediante la reacción en cadena de la polimerasa (PCR) a partir de una biblioteca, que introdujo sitios de restricción NdeI y BbsI que son compatibles con los sitios NdeI y BsaI que flanquean el relleno en el vector de destino stuffer-Y288-GFP-His8 ("His8" divulgada como la SEQ ID NO: 697). Las construcciones se generaron reemplazando la secuencia de relleno en el stuffer-Y288-GFP-His8 ("His8" divulgada como la SEQ ID NO: 697) por los fragmentos que codifican el anti-CD3. Los fragmentos de ADN del anti-CD3 digeridos por restricción se ligaron en el vector stuffer-Y288-GFP-His8 ("His8" divulgada como la SEQ ID NO: 697) utilizando ligasa T4 de ADN y se electroporaron en BL21(DE3) Gold (Stratagene, La Jolla, CA). Los transformantes se cribaron mediante ADN miniprep y las construcciones deseadas se confirmaron mediante la secuenciación del ADN. Los vectores finales producen el gen anti-CD3-Y288-GFP-His bajo el control de un promotor T7. La construcción resultante es: pMS0185. Las secuencias de ADN resultantes y las secuencias de aminoácidos codificadas se proporcionan a continuación.

Ejemplo 24: Construcción de genes y vectores que comprenden múltiples scFv

Construcción de genes y vectores anti-Her2-Y288-anti-CD3-HA-His6 y anti-Her2-Y288-anti-EGFR-HA-His6

Los genes que codifican el anti-CD3 y el anti-EGFR se amplificaron mediante la reacción en cadena de la polimerasa (PCR), que introdujo sitios de restricción BbsI en ambos extremos de los fragmentos de ADN. Se realizó una reacción en cadena de la polimerasa (PCR) para introducir una etiqueta HA-His6 ("His6" divulgada como la SEQ ID NO: 218) con un sitio de restricción HindIII en el extremo 3'. Los fragmentos anti-CD3-HA-His6 ("His6" divulgada como la SEQ ID NO: 218) (o el anti-GFR-HA-His6 ("His6" divulgada como la SEQ ID NO: 218)) eran compatibles con los sitios BbsI e HindIII que flanquean la GFP-His8 en el vector de destino anti-Her2-Y288-GFP-His8 ("His8" divulgada como la SEQ ID NO: 697). Las construcciones se generaron reemplazando la secuencia GFP-His8 ("His8" divulgada como la SEQ ID NO: 697) en la secuencia anti-Her2-Y288-GFP-His8 ("His8" divulgada como la SEQ ID NO: 697) por la anti-CD3-HA-His6 ("His6" divulgada como la SEQ ID NO: 218) o fragmentos codificantes de anti-EGFR-HA-His6 ("His6" divulgada como la SEQ ID NO: 218). Los fragmentos anti-CD3-HA-His6 ("His6" divulgada como la SEQ ID NO: 218) o anti-EGFR-HA-His6 ("His6" divulgada como la SEQ ID NO: 218) digeridos por restricción se ligaron en el vector anti-Her2-Y288-GFP-His8 ("His8" divulgada como la SEQ ID NO: 697) utilizando ligasa T4 de ADN y se electroporaron en BL21(DE3) Gold (Stratagene, La Jolla, CA). Los transformantes se cribaron mediante ADN miniprep y las construcciones deseadas se confirmaron mediante la secuenciación del ADN. Los vectores finales producen el gen anti-Her2-Y288-anti-CD3-HA-His6 ("His6" divulgada como la SEQ ID NO: 218) y anti-Her2-Y288-anti-EGFR-HA-His6 ("His6" divulgada como la SEQ ID NO: 218) bajo el control de un promotor T7. Las construcciones resultantes son: pMS0183, AC48 y pMS0184, AC49. Las secuencias de ADN resultantes y el producto final codificado se proporcionan en la tabla 25, a continuación.

Construcción de genes y vectores anti-Her2-Y288-anti-CD3-HA-His8

El gen que codifica el anti-CD3 se amplificó mediante la reacción en cadena de la polimerasa (PCR), que introdujo sitios de restricción BbsI y SpeI que son compatibles con los sitios BbsI y SpeI que flanquean la GFP en el vector de destino anti-Her2-Y288-GFP_HA-His8 ("His8" divulgada como la SEQ ID NO: 697). Las construcciones se generaron reemplazando la secuencia de la GFP en la anti-Her2-Y288-GFP-HA-His8 ("His8" divulgada como la SEQ ID NO: 697) por los fragmentos que codifican el anti-CD3. Los fragmentos de anti-CD3 o ADN digeridos por restricción se ligaron en el vector anti-Her2-Y288-GFP-HA-His8 ("His8" divulgada como la SEQ ID NO: 697) utilizando ligasa T4 de ADN y se electroporaron en XL1 Blue. Los transformantes se cribaron mediante ADN miniprep y las construcciones deseadas se confirmaron mediante la secuenciación del ADN. Los vectores finales producen el gen anti-Her2-Y288-anti-CD3-HA-His8 ("His8" divulgada como la SEQ ID NO: 697) bajo el control de un promotor T7. La construcción resultante es: pMS0212, AC69. Las secuencias de ADN resultantes y las secuencias de aminoácidos codificadas se proporcionan en la tabla 25, a continuación.

Ejemplo 25: Construcción de proteínas de unión VHH del aEGFR multivalentes

Se construyó una biblioteca LCW0501 de EGFR_VHH-XTEN_AM144 utilizando una PCR de cuatro clones de la biblioteca con codones previamente optimizados LMS109.005, 020, 038 y 045 con las secuencias de aminoácidos y de ADN denominadas NdeI_BsaI-EGFR_VHH1-XTEN_AM144-GFP6~229-H8 (LMS109.005); NdeI_BsaI-EGFR_VHH1-XTEN_AM144-GFP6~229-H8 (LMS109.020); NdeI_BsaI_EGFR_VHH1-XTEN_AM144-GFP6~229-H8) (LMS109.038); y NdeI_BsaI-EGFR_VHH1-XTEN_AM144-GFP6-229-H8 (LMS 109.045). Las secuencias de aminoácidos y de ácidos nucleicos se proporcionan a continuación. LCW0501 tiene la biblioteca génica de EGFR_VHH-XTEN_AM144 con los sitios de restricción flanqueantes NdeI y BsaI y BbsI, fusionados con la GFP-8xHis-tag ("8xHis" divulgada como la SEQ ID NO: 697) en un vector derivado del pET30 de Novagen.

El plásmido de LCW0501 se digirió con BsaI/HindIII para generar el fragmento pequeño como inserto, y se digirió con BbsI/HindIII para generar el fragmento grande como vector. Los fragmentos del inserto y el vector se ligaron, y la mezcla de ligación se electroporó en células BL21-Gold (DE3) para obtener transformantes de LCW0502. LCW0502 es la biblioteca génica del dímero EGFR_VHH-XTEN_AM144 con los mismos sitios de restricción flanqueantes BsaI y BbsI fusionados con la GFP-8xHis-tag ("8xHis" divulgada como la SEQ ID NO: 697) en el mismo vector.

El plásmido de LCW0502 se digirió con BsaI/HindIII para generar el fragmento pequeño como inserto, y se digirió con BbsI/HindIII para generar el fragmento grande como vector. Los fragmentos del inserto y el vector se ligaron, y la mezcla de ligación se electroporó en células BL21-Gold (DE3) para obtener transformantes de LCW0503. LCW0503 es la biblioteca génica del tetrámero EGFR_VHH-XTEN_AM144 con los mismos sitios de restricción flanqueante BsaI y BbsI fusionados con la GFP-8xHis-tag ("8xHis" divulgada como la SEQ ID NO: 697) en el mismo vector.

El plásmido de LCW0502 se digirió con BsaI/HindIII para generar el pequeño fragmento a medida que el inserto y el plásmido de LCW0503 se digirieron con BbsI/HindIII para generar el fragmento grande como vector. Los fragmentos del inserto y el vector se ligaron, y la mezcla de ligación se electroporó en células BL21-Gold (DE3) para obtener transformantes de LCW0504. LCW0504 es la biblioteca génica del hexámero EGFR_VHH-XTEN_AM144 con los mismos sitios de restricción flanqueantes BsaI y BbsI fusionados con la GFP-8xHis-tag ("8xHis" divulgada como la SEQ ID NO: 697) en el mismo vector.

El plásmido de LCW0503 se digirió con BsaI/HindIII para generar el fragmento pequeño como inserto, y se digirió con BbsI/HindIII para generar el fragmento grande como vector. Los fragmentos del inserto y el vector se ligaron, y la mezcla de ligación se electroporó en células BL21-Gold (DE3) para obtener transformantes de LCW0505. LCW0505 es la biblioteca génica del octámero EGFR_VHH-XTEN_AM144 con los mismos sitios de restricción flanqueante BsaI y BbsI fusionados con la GFP-8xHis-tag ("8xHis" divulgada como la SEQ ID NO: 697) en el mismo vector.

Las bibliotecas LCW501, LCW502, LCW503, LCW504 y LCW505 se cribaron para determinar el mejor candidato de expresión para su evaluación. El cribado se realizó de la siguiente manera para todas las bibliotecas. Las colonias para la transformación se recogieron en 500 µl de cultivos de LB en 96 placas de pocillos profundos y se cultivaron hasta saturación durante la noche. Estos cultivos se almacenaron a 4 °C después de que se usaran 40 µl de estos cultivos para inocular 500 µl de medios de autoinducción, y estos cultivos se cultivaron a 26 °C durante >24 horas. Tras el crecimiento, se midió la fluorescencia de la GFP de 100 µl de estos cultivos de medios de autoinducción utilizando un lector de placas de fluorescencia. La fluorescencia de la GFP es proporcional a la expresión de la proteína y, por lo tanto, es una lectura de la expresión total. Se identificaron los clones de expresión más alta, y se inició un nuevo 1 ml durante la noche en SB a partir del saturado original durante la noche de ese clon. Se realizaron mini-preps con estos nuevos cultivos a plásmidos derivados. Las secuencias de ADN y de aminoácidos se proporcionan en la tabla 25, a continuación.

Tabla 25: Secuencias de ADN y de aminoácidos de construcciones de proteína de fusión de unión

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
CTLA4-AE36-CTLA4-AE864, AC389	ATGGCAATGCATGTTGCACAGCCTGCAGTTGTTCT GGCAAGCAGCCGTGGTATTGCCAGCTTTGTTTGTG AATATGCAAGTCCGGGTAAAGCAACCGAAGTTCG TGTTACCGTTCTGAGACAGGCAGATAGCCAGGTTA CCGAAGTTTGTGCAGCAACCTATATGATGGGTAAT GAACTGACCTTTCTGGATGATAGCATTTGTACCGG CACCAGCAGCGGTAATCAGGTTAATCTGACCATTC AGGGTCTGCGTGCAATGGATACCGGTCTGTATATT TGTAAAGTGGAACTGATGTATCCGCCTCCGTATTA TCTGGGTATTGGTAATGGCACCCAGATTTATGTTA TTGATCCGGAAGGCGCGCCAGGTACAAGCGAAAG CGCAACACCGGAAAGCGGTCCGGGTAGCGAACCG GCAACCAGCGGTAGCGAAACACCGGGTACATCAA CCGAACCGAGCGAAGGTAGCGCACCGGGGCCGGC CATGCATGTGGCCCAGCCAGCCGTGGTGCTGGCA AGTTCACGCGGTATTGCATCATTTGTGTGCGAATA TGCATCACCTGGTAAAGCCACAGAAGTGCGCGTA ACAGTACTGCGTCAGGCCGATTCACAGGTGACAG AAGTTTGCGCTGCCACATACATGATGGGCAACGA GCTGACATTCCTGGACGATTCAATTTGTACTGGTA CAAGCTCAGGCAATCAGGTGAACCTGACAATCCA AGGCCTGAGAGCTATGGACACAGGCCTGTACATCT	698	MAMHVAQPAVV LASSRGIASFVCE YASPGKATEVRV TVLRQADSQVTE VCAATYMMGNE LTFLDDSICTGTS SGNQVNLTIQGL RAMDTGLYICKV ELMYPPPYYLGIG NGTQIYVIDPEGA PGTSESATPESGP GSEPATSGSETPG TSTEPSEGSAPGP AMHVAQPAVVL ASSRGIASFVCEY ASPGKATEVRVT VLRQADSQVTEV CAATYMMGNEL TFLDDSICTGTSS GNQVNLTIQGLR AMDTGLYICKVE LMYPPPYYLGIG	723

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	GCAAAGTTGAGCTGATGTACCCTCCGCCTTATTAC TTAGGCATTGGCAACGGTACACAGATCTATGTGAT CGATCCTGAGGGAGGTAGCCCGGCTGGCTCTCCTA CCTCTACTGAGGAAGGTACTTCTGAAAGCGCTACT CCTGAGTCTGGTCCAGGTACCTCTACTGAACCGTC CGAAGGTAGCGCTCCAGGTAGCCCAGCAGGCTCT CCGACTTCCACTGAGGAAGGTACTTCTACTGAACC TTCCGAAGGCAGCGCACCAGGTACCTCTACTGAAC CTTCTGAGGGCAGCGCTCCAGGTACTTCTGAAAGC GCTACCCCGGAATCTGGCCCAGGTAGCGAACCGG CTACTTCTGGTTCTGAAACCCCAGGTAGCGAACCG GCTACCTCCGGTTCTGAAACTCCAGGTAGCCCGGC AGGCTCTCCGACCTCTACTGAGGAAGGTACTTCTG AAAGCGCAACCCCGGAGTCCGGCCCAGGTACCTC TACCGAACCGTCTGAGGGCAGCGCACCAGGTACT TCTACCGAACCGTCCGAGGGTAGCGCACCAGGTA GCCCAGCAGGTTCTCCTACCTCCACCGAGGAAGGT ACTTCTACCGAACCGTCCGAGGGTAGCGCACCAG GTACCTCTACTGAACCTTCTGAGGGCAGCGCTCCA GGTACTTCTGAAAGCGCTACCCCGGAGTCCGGTCC AGGTACTTCTACTGAACCGTCCGAAGGTAGCGCAC CAGGTACTTCTGAAAGCGCAACCCCTGAATCCGGT CCAGGTAGCGAACCGGCTACTTCTGGCTCTGAGAC TCCAGGTACTTCTACCGAACCGTCCGAAGGTAGCG CACCAGGTACTTCTACTGAACCGTCTGAAGGTAGC GCACCAGGTACTTCTGAAAGCGCAACCCCGGAAT CCGGCCCAGGTACCTCTGAAAGCGCAACCCCGGA GTCCGGCCCAGGTAGCCCTGCTGGCTCTCCAACCT CCACCGAAGAAGGTACCTCTGAAAGCGCAACCCC TGAATCCGGCCCAGGTAGCGAACCGGCAACCTCC GGTTCTGAAACCCCAGGTACCTCTGAAAGCGCTAC TCCGGAGTCTGGCCCAGGTACCTCTACTGAACCGT CTGAGGGTAGCGCTCCAGGTACTTCTACTGAACCG TCCGAAGGTAGCGCACCAGGTACTTCTACCGAACC GTCCGAAGGCAGCGCTCCAGGTACCTCTACTGAAC CTTCCGAGGGCAGCGCTCCAGGTACCTCTACCGAA CCTTCTGAAGGTAGCGCACCAGGTACTTCTACCGA ACCGTCCGAGGGTAGCGCACCAGGTAGCCCAGCA GGTTCTCCTACCTCCACCGAGGAAGGTACTTCTAC CGAACCGTCCGAGGGTAGCGCACCAGGTACCTCT GAAAGCGCAACTCCTGAGTCTGGCCCAGGTAGCG AACCTGCTACCTCCGGCTCTGAGACTCCAGGTACC TCTGAAAGCGCAACCCCGGAATCTGGTCCAGGTA GCGAACCTGCAACCTCTGGCTCTGAAACCCCAGGT ACCTCTGAAAGCGCTACTCCTGAATCTGGCCCAGG TACTTCTACTGAACCGTCCGAGGGCAGCGCACCAG GTACTTCTGAAAGCGCTACTCCTGAGTCCGGCCCA GGTAGCCCGGCTGGCTCTCCGACTTCCACCGAGGA AGGTAGCCCGGCTGGCTCTCCAACTTCTACTGAAG AAGGTAGCCCGGCAGGCTCTCCGACCTCTACTGAG GAAGGTACTTCTGAAAGCGCAACCCCGGAGTCCG GCCCAGGTACCTCTACCGAACCGTCTGAGGGCAG CGCACCAGGTACCTCTGAAAGCGCAACTCCTGAGT CTGGCCCAGGTAGCGAACCTGCTACCTCCGGCTCT GAGACTCCAGGTACCTCTGAAAGCGCAACCCCGG AATCTGGTCCAGGTAGCGAACCTGCAACCTCTGGC TCTGAAACCCCAGGTACCTCTGAAAGCGCTACTCC TGAATCTGGCCCAGGTACTTCTACTGAACCGTCCG		NGTQIYVIDPEGG SPAGSPTSTEEGT SESATPESGPGTS TEPSEGSAPGSPA GSPTSTEEGTSTE PSEGSAPGTSTEP SEGSAPGTSESAT PESGPGSEPATSG SETPGSEPATSGS ETPGSPAGSPTST EEGTSESATPESG PGTSTEPSEGSAP GTSTEPSEGSAPG SPAGSPTSTEEGT STEPSEGSAPGTS TEPSEGSAPGTSE SATPESGPGTSTE PSEGSAPGTSESA TPESGPGSEPATS GSETPGTSTEPSE GSAPGTSTEPSEG SAPGTSESATPES GPGTSESATPESG PGSPAGSPTSTEE GTSESATPESGPG SEPATSGSETPGT SESATPESGPGTS TEPSEGSAPGTST EPSEGSAPGTSTE PSEGSAPGTSTEP SEGSAPGTSTEPS EGSAPGTSTEPSE GSAPGSPAGSPTS TEEGTSTEPSEGS APGTSESATPESG PGSEPATSGSETP GTSESATPESGPG SEPATSGSETPGT SESATPESGPGTS TEPSEGSAPGTSE SATPESGPGSPAG SPTSTEEGSPAGS PTSTEEGSPAGSP TSTEEGTSESATP ESGPGTSTEPSEG SAPGTSESATPES GPGSEPATSGSET PGTSESATPESGP GSEPATSGSETPG TSESATPESGPGT STEPSEGSAPGSP AGSPTSTEEGTSE SATPESGPGSEPA TSGSETPGTSESA TPESGPGSPAGSP TSTEEGSPAGSPT STEEGTSTEPSEG SAPGTSESATPES	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	AGGGCAGCGCACCAGGTAGCCCTGCTGGCTCTCC AACCTCCACCGAAGAAGGTACCTCTGAAAGCGCA ACCCCTGAATCCGGCCCAGGTAGCGAACCGGCAA CCTCCGGTTCTGAAACCCCAGGTACTTCTGAAAGC GCTACTCCTGAGTCCGGCCCAGGTAGCCCGGCTGG CTCTCCGACTTCCACCGAGGAAGGTAGCCCGGCTG GCTCTCCAACTTCTACTGAAGAAGGTACTTCTACC GAACCTTCCGAGGGCAGCGCACCAGGTACTTCTG AAAGCGCTACCCCTGAGTCCGGCCCAGGTACTTCT GAAAGCGCTACTCCTGAATCCGGTCCAGGTACTTC TGAAAGCGCTACCCCGGAATCTGGCCCAGGTAGC GAACCGGCTACTTCTGGTTCTGAAACCCCAGGTAG CGAACCGGCTACCTCCGGTTCTGAAACTCCAGGTA GCCCAGCAGGCTCTCCGACTTCCACTGAGGAAGGT ACTTCTACTGAACCTTCCGAAGGCAGCGCACCAGG TACCTCTACTGAACCTTCTGAGGGCAGCGCTCCAG GTAGCGAACCTGCAACCTCTGGCTCTGAAACCCCA GGTACCTCTGAAAGCGCTACTCCTGAATCTGGCCC AGGTACTTCTACTGAACCGTCCGAGGGCAGCGCA CCAGGTTAA		GPGTSESATPESG PGTSESATPESGP GSEPATSGSETPG SEPATSGSETPGS PAGSPTSTEEGTS TEPSEGSAPGTST EPSEGSAPGSEPA TSGSETPGTSESA TPESGPGTSTEPS EGSAPG	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
CTLA4-AE36-CTLA4-AE864, AC390	ATGGCAATGCATGTTGCACAGCCTGCAGTTGTTCT GGCAAGCAGCCGTGGTATTGCCAGCTTTGTTTGTG AATATGCAAGTCCGGGTAAAGCAACCGAAGTTCG TGTTACCGTTCTGAGACAGGCAGATAGCCAGGTTA CCGAAGTTTGTGCAGCAACCTATATGATGGGTAAT GAACTGACCTTTCTGGATGATAGCATTTGTACCGG CACCAGCAGCGGTAATCAGGTTAATCTGACCATTC AGGGTCTGCGTGCAATGGATACCGGTCTGTATATT TGTAAAGTGGAACTGATGTATCCGCCTCCGTATTA TCTGGGTATTGGTAATGGCACCCAGATTTATGTTA TTGATCCGGAACCGTGTCCGGATAGCGGCGCGCC AGGTACAAGCGAAAGCGCAACACCGGAAAGCGGT CCGGGTAGCGAACCGGCAACCAGCGGTAGCGAAA CACCGGGTACATCAACCGAACCGAGCGAAGGTAG CGCACCGGGGCCGGCCATGCATGTGGCCCAGCCA GCCGTGGTGCTGGCAAGTTCACGCGGTATTGCATC ATTTGTGTGCGAATATGCATCACCTGGTAAAGCCA CAGAAGTGCGCGTAACAGTACTGCGTCAGGCCGA TTCACAGGTGACAGAAGTTTGCGCTGCCACATACA TGATGGGCAACGAGCTGACATTCCTGGACGATTCA ATTTGTACTGGTACAAGCTCAGGCAATCAGGTGAA CCTGACAATCCAAGGCCTGAGAGCTATGGACACA GGCCTGTACATCTGCAAAGTTGAGCTGATGTACCC TCCGCCTTATTACTTAGGCATTGGCAACGGTACAC AGATCTATGTGATCGATCCTGAACCTTGCCCTGAT TCAGGAGGTAGCCCGGCTGGCTCTCCTACCTCTAC TGAGGAAGGTACTTCTGAAAGCGCTACTCCTGAGT CTGGTCCAGGTACCTCTACTGAACCGTCCGAAGGT AGCGCTCCAGGTAGCCCAGCAGGCTCTCCGACTTC CACTGAGGAAGGTACTTCTACTGAACCTTCCGAAG GCAGCGCACCAGGTACCTCTACTGAACCTTCTGAG GGCAGCGCTCCAGGTACTTCTGAAAGCGCTACCCC GGAATCTGGCCCAGGTAGCGAACCGGCTACTTCTG GTTCTGAAACCCCAGGTAGCGAACCGGCTACCTCC GGTTCTGAAACTCCAGGTAGCCCGGCAGGCTCTCC GACCTCTACTGAGGAAGGTACTTCTGAAAGCGCA ACCCCGGAGTCCGGCCCAGGTACCTCTACCGAACC GTCTGAGGGCAGCGCACCAGGTACTTCTACCGAA	699	MAMHVAQPAVV LASSRGIASFVCE YASPGKATEVRV TVLRQADSQVTE VCAATYMMGNE LTFLDDSICTGTS SGNQVNLTIQGL RAMDTGLYICKV ELMYPPPYYLGIG NGTQIYVIDPEPC PDSGAPGTSESAT PESGPGSEPATSG SETPGTSTEPSEG SAPGPAMHVAQP AVVLASSRGIASF VCEYASPGKATE VRVTVLRQADSQ VTEVCAATYMM GNELTFLDDSICT GTSSGNQVNLTIQ GLRAMDTGLYIC KVELMYPPPYYL GIGNGTQIYVIDP EPCPDSGGSPAGS PTSTEEGTSESAT PESGPGTSTEPSE GSAPGSPAGSPTS TEEGTSTEPSEGS APGTSTEPSEGSA PGTSESATPESGP GSEPATSGSETPG SEPATSGSETPGS PAGSPTSTEEGTS ESATPESGPGTST EPSEGSAPGTSTE PSEGSAPGSPAGS PTSTEEGTSTEPS EGSAPGTSTEPSE	724

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	CCGTCCGAGGGTAGCGCACCAGGTAGCCCAGCAG GTTCTCCTACCTCCACCGAGGAAGGTACTTCTACC GAACCGTCCGAGGGTAGCGCACCAGGTACCTCTA CTGAACCTTCTGAGGGCAGCGCTCCAGGTACTTCT GAAAGCGCTACCCCCGGAGTCCGGTCCAGGTACTTC TACTGAACCGTCCGAAGGTAGCGCACCAGGTACTT CTGAAAGCGCAACCCCTGAATCCGGTCCAGGTAG CGAACCGGCTACTTCTGGCTCTGAGACTCCAGGTA CTTCTACCGAACCGTCCGAAGGTAGCGCACCAGGT ACTTCTACTGAACCGTCTGAAGGTAGCGCACCAGG TACTTCTGAAAGCGCAACCCCGGAATCCGGCCCA GGTACCTCTGAAAGCGCAACCCCGGAGTCCGGCC CAGGTAGCCCTGCTGGCTCTCCAACCTCCACCGAA GAAGGTACCTCTGAAAGCGCAACCCCTGAATCCG GCCCAGGTAGCGAACCGGCAACCTCCGGTTCTGA AACCCCAGGTACCTCTGAAAGCGCTACTCCGGAGT CTGGCCCAGGTACCTCTACTGAACCGTCTGAGGGT AGCGCTCCAGGTACTTCTACTGAACCGTCCGAAGG TAGCGCACCAGGTACTTCTACCGAACCGTCCGAAG GCAGCGCTCCAGGTACCTCTACTGAACCTTCCGAG GGCAGCGCTCCAGGTACCTCTACCGAACCTTCTGA AGGTAGCGCACCAGGTACTTCTACCGAACCGTCCG AGGGTAGCGCACCAGGTAGCCCAGCAGGTTCTCC TACCTCCACCGAGGAAGGTACTTCTACCGAACCGT CCGAGGGTAGCGCACCAGGTACCTCTGAAAGCGC AACTCCTGAGTCTGGCCCAGGTAGCGAACCTGCTA CCTCCGGCTCTGAGACTCCAGGTACCTCTGAAAGC GCAACCCCGGAATCTGGTCCAGGTAGCGAACCTG CAACCTCTGGCTCTGAAACCCCAGGTACCTCTGAA AGCGCTACTCCTGAATCTGGCCCAGGTACTTCTAC TGAACCGTCCGAGGGCAGCGCACCAGGTACTTCT GAAAGCGCTACTCCTGAGTCCGGCCCAGGTAGCC CGGCTGGCTCTCCGACTTCCACCGAGGAAGGTAGC CCGGCTGGCTCTCCAACTTCTACTGAAGAAGGTAG CCCGGCAGGCTCTCCGACCTCTACTGAGGAAGGTA CTTCTGAAAGCGCAACCCCGGAGTCCGGCCCAGG TACCTCTACCGAACCGTCTGAGGGCAGCGCACCA GGTACCTCTGAAAGCGCAACTCCTGAGTCTGGCCC AGGTAGCGAACCTGCTACCTCCGGCTCTGAGACTC CAGGTACCTCTGAAAGCGCAACCCCGGAATCTGG TCCAGGTAGCGAACCTGCAACCTCTGGCTCTGAAA CCCCAGGTACCTCTGAAAGCGCTACTCCTGAATCT GGCCCAGGTACTTCTACTGAACCGTCCGAGGGCA GCGCACCAGGTAGCCCTGCTGGCTCTCCAACCTCC ACCGAAGAAGGTACCTCTGAAAGCGCAACCCCTG AATCCGGCCCAGGTAGCGAACCGGCAACCTCCGG TTCTGAAACCCCAGGTACTTCTGAAAGCGCTACTC CTGAGTCCGGCCCAGGTAGCCCGGCTGGCTCTCCG ACTTCCACCGAGGAAGGTAGCCCGGCTGGCTCTCC AACTTCTACTGAAGAAGGTACTTCTACCGAACCTT CCGAGGGCAGCGCACCAGGTACTTCTGAAAGCGC TACCCCTGAGTCCGGCCCAGGTACTTCTGAAAGCG CTACTCCTGAATCCGGTCCAGGTACTTCTGAAAGC GCTACCCCGGAATCTGGCCCAGGTAGCGAACCGG CTACTTCTGGTTCTGAAACCCCAGGTAGCGAACCG GCTACCTCCGGTTCTGAAACTCCAGGTAGCCCAGC AGGCTCTCCGACTTCCACTGAGGAAGGTACTTCTA CTGAACCTTCCGAAGGCAGCGCACCAGGTACCTCT		GSAPGTSESATPE SGPGTSTEPSEGS APGTSESATPESG PGSEPATSGSETP GTSTEPSEGSAPG TSTEPSEGSAPGT SESATPESGPGTS ESATPESGPGSPA GSPTSTEEGTSES ATPESGPGSEPAT SGSETPGTSESAT PESGPGTSTEPSE GSAPGTSTEPSEG SAPGTSTEPSEGS APGTSTEPSEGSA PGTSTEPSEGSAP GTSTEPSEGSAPG SPAGSPTSTEEGT STEPSEGSAPGTS ESATPESGPGSEP ATSGSETPGTSES ATPESGPGSEPAT SGSETPGTSESAT PESGPGTSTEPSE GSAPGTSESATPE SGPGSPAGSPTST EEGSPAGSPTSTE EGSPAGSPTSTEE GTSESATPESGPG TSTEPSEGSAPGT SESATPESGPGSE PATSGSETPGTSE SATPESGPGSEPA TSGSETPGTSESA TPESGPGTSTEPS EGSAPGSPAGSPT STEEGTSESATPE SGPGSEPATSGSE TPGTSESATPESG PGSPAGSPTSTEE GSPAGSPTSTEEG TSTEPSEGSAPGT SESATPESGPGTS ESATPESGPGTSE SATPESGPGSEPA TSGSETPGSEPAT SGSETPGSPAGSP TSTEEGTSTEPSE GSAPGTSTEPSEG SAPGSEPATSGSE TPGTSESATPESG PGTSTEPSEGSAP G	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	ACTGAACCTTCTGAGGGCAGCGCTCCAGGTAGCG AACCTGCAACCTCTGGCTCTGAAACCCCAGGTACC TCTGAAAGCGCTACTCCTGAATCTGGCCCAGGTAC TTCTACTGAACCGTCCGAGGGCAGCGCACCAGGTT AA			

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
CTLA4-AE158-CTLA4-AE864, AC391	ATGGCAATGCATGTTGCACAGCCTGCAGTTGTTCT GGCAAGCAGCCGTGGTATTGCCAGCTTTGTTTGTG AATATGCAAGTCCGGGTAAAGCAACCGAAGTTCG TGTTACCGTTCTGAGACAGGCAGATAGCCAGGTTA CCGAAGTTTGTGCAGCAACCTATATGATGGGTAAT GAACTGACCTTTCTGGATGATAGCATTTGTACCGG CACCAGCAGCGGTAATCAGGTTAATCTGACCATTC AGGGTCTGCGTGCAATGGATACCGGTCTGTATATT TGTAAAGTGGAACTGATGTATCCGCCTCCGTATTA TCTGGGTATTGGTAATGGCACCCAGATTTATGTTA TTGATCCGGAAGGCGCGCCAAGCACGGGAGGTAC TTCTGAAAGCGCTACTCCGGAGTCCGGTCCAGGTA CCTCTACCGAACCGTCCGAAGGCAGCGCTCCAGGT ACTTCTACTGAACCTTCTGAGGGTAGCGCTCCAGG TACTTCTGAAAGCGCTACTCCGGAGTCCGGTCCAG GTACCTCTACCGAACCGTCCGAAGGCAGCGCTCCA GGTACTTCTACTGAACCTTCTGAGGGTAGCGCTCC AGGTACCTCTGAAAGCGCTACTCCGGAGTCTGGCC CAGGTACCTCTACTGAACCGTCTGAGGGTAGCGCT CCAGGTACTTCTACTGAACCGTCCGAAGGTAGCGC ACCAGGTACTTCTACCGAACCGTCCGAGGGTAGC GCACCAGGTAGCCCAGCAGGTTCTCCTACCTCCAC CGAGGAAGGTACTTCTACCGAACCGTCCGAGGGT AGCGCACCAGGTCCAGAACCAACGGGGCCGGCCA TGCATGTGGCCCAGCCAGCCGTGGTGCTGGCAAGT TCACGCGGTATTGCATCATTTGTGTGCGAATATGC ATCACCTGGTAAAGCCACAGAAGTGCGCGTAACA GTACTGCGTCAGGCCGATTCACAGGTGACAGAAG TTTGCGCTGCCACATACATGATGGGCAACGAGCTG ACATTCCTGGACGATTCAATTTGTACTGGTACAAG CTCAGGCAATCAGGTGAACCTGACAATCCAAGGC CTGAGAGCTATGGACACAGGCCTGTACATCTGCAA AGTTGAGCTGATGTACCCTCCGCCTTATTACTTAG GCATTGGCAACGGTACACAGATCTATGTGATCGAT CCTGAGGGAGGTAGCCCGGCTGGCTCTCCTACCTC TACTGAGGAAGGTACTTCTGAAAGCGCTACTCCTG AGTCTGGTCCAGGTACCTCTACTGAACCGTCCGAA GGTAGCGCTCCAGGTAGCCCAGCAGGCTCTCCGA CTTCCACTGAGGAAGGTACTTCTACTGAACCTTCC GAAGGCAGCGCACCAGGTACCTCTACTGAACCTTC TGAGGGCAGCGCTCCAGGTACTTCTGAAAGCGCTA CCCCGGAATCTGGCCCAGGTAGCGAACCGGCTAC TTCTGGTTCTGAAACCCCAGGTAGCGAACCGGCTA CCTCCGGTTCTGAAACTCCAGGTAGCCCGGCAGGC TCTCCGACCTCTACTGAGGAAGGTACTTCTGAAAG CGCAACCCCCGGAGTCCGGCCCAGGTACCTCTACCG AACCGTCTGAGGGCAGCGCACCAGGTACTTCTACC GAACCGTCCGAGGGTAGCGCACCAGGTAGCCCAG CAGGTTCTCCTACCTCCACCGAGGAAGGTACTTCT ACCGAACCGTCCGAGGGTAGCGCACCAGGTACCT CTACTGAACCTTCTGAGGGCAGCGCTCCAGGTACT TCTGAAAGCGCTACCCCGGAGTCCGGTCCAGGTAC TTCTACTGAACCGTCCGAAGGTAGCGCACCAGGTA	700	MAMHVAQPAVV LASSRGIASFVCE YASPGKATEVRV TVLRQADSQVTE VCAATYMMGNE LTFLDDSICTGTS SGNQVNLTIQGL RAMDTGLYICKV ELMYPPPYYLGIG NGTQIYVIDPEGA PSTGGTSESATPE SGPGTSTEPSEGS APGTSTEPSEGSA PGTSESATPESGP GTSTEPSEGSAPG TSTEPSEGSAPGT SESATPESGPGTS TEPSEGSAPGTST EPSEGSAPGTSTE PSEGSAPGSPAGS PTSTEEGTSTEPS EGSAPGPEPTGPA MHVAQPAVVLA SSRGIASFVCEYA SPGKATEVRVTV LRQADSQVTEVC AATYMMGNELT FLDDSICTGTSSG NQVNLTIQGLRA MDTGLYICKVEL MYPPPYYLGIGN GTQIYVIDPEGGS PAGSPTSTEEGTS ESATPESGPGTST EPSEGSAPGSPAG SPTSTEEGTSTEPS EGSAPGTSTEPSE GSAPGTSESATPE SGPGSEPATSGSE TPGSEPATSGSET PGSPAGSPTSTEE GTSESATPESGPG TSTEPSEGSAPGT STEPSEGSAPGSP AGSPTSTEEGTST EPSEGSAPGTSTE PSEGSAPGTSESA TPESGPGTSTEPS EGSAPGTSESATP ESGPGSEPATSGS ETPGTSTEPSEGS APGTSTEPSEGSA PGTSESATPESGP	725

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	CTTCTGAAAGCGCAACCCCTGAATCCGGTCCAGGT AGCGAACCGGCTACTTCTGGCTCTGAGACTCCAGG TACTTCTACCGAACCGTCCGAAGGTAGCGCACCAG GTACTTCTACTGAACCGTCTGAAGGTAGCGCACCA GGTACTTCTGAAAGCGCAACCCCGGAATCCGGCC CAGGTACCTCTGAAAGCGCAACCCCGGAGTCCGG CCCAGGTAGCCCTGCTGGCTCTCCAACCTCCACCG AAGAAGGTACCTCTGAAAGCGCAACCCCTGAATC CGGCCCAGGTAGCGAACCGGCAACCTCCGGTTCT GAAACCCCAGGTACCTCTGAAAGCGCTACTCCGG AGTCTGGCCCAGGTACCTCTACTGAACCGTCTGAG GGTAGCGCTCCAGGTACTTCTACTGAACCGTCCGA AGGTAGCGCACCAGGTACTTCTACCGAACCGTCCG AAGGCAGCGCTCCAGGTACCTCTACTGAACCTTCC GAGGGCAGCGCTCCAGGTACCTCTACCGAACCTTC TGAAGGTAGCGCACCAGGTACTTCTACCGAACCGT CCGAGGGTAGCGCACCAGGTAGCCCAGCAGGTTC TCCTACCTCCACCGAGGAAGGTACTTCTACCGAAC CGTCCGAGGGTAGCGCACCAGGTACCTCTGAAAG CGCAACTCCTGAGTCTGGCCCAGGTAGCGAACCTG CTACCTCCGGCTCTGAGACTCCAGGTACCTCTGAA AGCGCAACCCCCGGAATCTGGTCCAGGTAGCGAAC CTGCAACCTCTGGCTCTGAAACCCCAGGTACCTCT GAAAGCGCTACTCCTGAATCTGGCCCAGGTACTTC TACTGAACCGTCCGAGGGCAGCGCACCAGGTACT TCTGAAAGCGCTACTCCTGAGTCCGGCCCAGGTAG CCCGGCTGGCTCTCCGACTTCCACCGAGGAAGGTA GCCCGGCTGGCTCTCCAACTTCTACTGAAGAAGGT AGCCCGGCAGGCTCTCCGACCTCTACTGAGGAAG GTACTTCTGAAAGCGCAACCCCGGAGTCCGGCCC AGGTACCTCTACCGAACCGTCTGAGGGCAGCGCA CCAGGTACCTCTGAAAGCGCAACTCCTGAGTCTGG CCCAGGTAGCGAACCTGCTACCTCCGGCTCTGAGA CTCCAGGTACCTCTGAAAGCGCAACCCCGGAATCT GGTCCAGGTAGCGAACCTGCAACCTCTGGCTCTGA AACCCCAGGTACCTCTGAAAGCGCTACTCCTGAAT CTGGCCCAGGTACTTCTACTGAACCGTCCGAGGGC AGCGCACCAGGTAGCCCTGCTGGCTCTCCAACCTC CACCGAAGAAGGTACCTCTGAAAGCGCAACCCCT GAATCCGGCCCAGGTAGCGAACCGGCAACCTCCG GTTCTGAAACCCCAGGTACTTCTGAAAGCGCTACT CCTGAGTCCGGCCCAGGTAGCCCGGCTGGCTCTCC GACTTCCACCGAGGAAGGTAGCCCCGGCTGGCTCTC CAACTTCTACTGAAGAAGGTACTTCTACCGAACCT TCCGAGGGCAGCGCACCAGGTACTTCTGAAAGCG CTACCCCTGAGTCCGGCCCAGGTACTTCTGAAAGC GCTACTCCTGAATCCGGTCCAGGTACTTCTGAAAG CGCTACCCCCGGAATCTGGCCCAGGTAGCGAACCG GCTACTTCTGGTTCTGAAACCCCAGGTAGCGAACC GGCTACCTCCGGTTCTGAAACTCCAGGTAGCCCAG CAGGCTCTCCGACTTCCACTGAGGAAGGTACTTCT ACTGAACCTTCCGAAGGCAGCGCACCAGGTACCT CTACTGAACCTTCTGAGGGCAGCGCTCCAGGTAGC GAACCTGCAACCTCTGGCTCTGAAACCCCAGGTAC CTCTGAAAGCGCTACTCCTGAATCTGGCCCAGGTA CTTCTACTGAACCGTCCGAGGGCAGCGCACCAGGT TAA		GTSESATPESGPG SPAGSPTSTEEGT SESATPESGPGSE PATSGSETPGTSE SATPESGPGTSTE PSEGSAPGTSTEP SEGSAPGTSTEPS EGSAPGTSTEPSE GSAPGTSTEPSEG SAPGTSTEPSEGS APGSPAGSPTSTE EGTSTEPSEGSAP GTSESATPESGPG SEPATSGSETPGT SESATPESGPGSE PATSGSETPGTSE SATPESGPGTSTE PSEGSAPGTSESA TPESGPGSPAGSP TSTEEGSPAGSPT STEEGSPAGSPTS TEEGTSESATPES GPGTSTEPSEGSA PGTSESATPESGP GSEPATSGSETPG TSESATPESGPGS EPATSGSETPGTS ESATPESGPGTST EPSEGSAPGSPAG SPTSTEEGTSESA TPESGPGSEPATS GSETPGTSESATP ESGPGSPAGSPTS TEEGSPAGSPTST EEGTSTEPSEGSA PGTSESATPESGP GTSESATPESGPG TSESATPESGPGS EPATSGSETPGSE PATSGSETPGSPA GSPTSTEEGTSTE PSEGSAPGTSTEP SEGSAPGSEPATS GSETPGTSESATP ESGPGTSTEPSEG SAPG	
CTLA4-AE158-CTLA4-AE864, AC392	ATGGCAATGCATGTTGCACAGCCTGCAGTTGTTCT	701	MAMHVAQPAVV	726

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	GGCAAGCAGCCGTGGTATTGCCAGCTTTGTTTGTG AATATGCAAGTCCGGGTAAAGCAACCGAAGTTCG TGTTACCGTTCTGAGACAGGCAGATAGCCAGGTTA CCGAAGTTTGTGCAGCAACCTATATGATGGGTAAT GAACTGACCTTTCTGGATGATAGCATTTGTACCGG CACCAGCAGCGGTAATCAGGTTAATCTGACCATTC AGGGTCTGCGTGCAATGGATACCGGTCTGTATATT TGTAAAGTGGAACTGATGTATCCGCCTCCGTATTA TCTGGGTATTGGTAATGGCACCCAGATTTATGTTA TTGATCCGGAACCGTGTCCGGATAGCGGCGCGCC AAGCACGGGAGGTACTTCTGAAAGCGCTACTCCG GAGTCCGGTCCAGGTACCTCTACCGAACCGTCCGA AGGCAGCGCTCCAGGTACTTCTACTGAACCTTCTG AGGGTAGCGCTCCAGGTACTTCTGAAAGCGCTACT CCGGAGTCCGGTCCAGGTACCTCTACCGAACCGTC CGAAGGCAGCGCTCCAGGTACTTCTACTGAACCTT CTGAGGGTAGCGCTCCAGGTACCTCTGAAAGCGCT ACTCCGGAGTCTGGCCCAGGTACCTCTACTGAACC GTCTGAGGGTAGCGCTCCAGGTACTTCTACTGAAC CGTCCGAAGGTAGCGCACCAGGTACTTCTACCGA ACCGTCCGAGGGTAGCGCACCAGGTAGCCCAGCA GGTTCTCCTACCTCCACCGAGGAAGGTACTTCTAC CGAACCGTCCGAGGGTAGCGCACCAGGTCCAGAA CCAACGGGGCCGGCCATGCATGTGGCCCAGCCAG CCGTGGTGCTGGCAAGTTCACGCGGTATTGCATCA TTTGTGTGCGAATATGCATCACCTGGTAAAGCCAC AGAAGTGCGCGTAACAGTACTGCGTCAGGCCGAT TCACAGGTGACAGAAGTTTGCGCTGCCACATACAT GATGGGCAACGAGCTGACATTCCTGGACGATTCA ATTTGTACTGGTACAAGCTCAGGCAATCAGGTGAA CCTGACAATCCAAGGCCTGAGAGCTATGGACACA GGCCTGTACATCTGCAAAGTTGAGCTGATGTACCC TCCGCCTTATTACTTAGGCATTGGCAACGGTACAC AGATCTATGTGATCGATCCTGAACCTTGCCCTGAT TCAGGAGGTAGCCCGGCTGGCTCTCCTACCTCTAC TGAGGAAGGTACTTCTGAAAGCGCTACTCCTGAGT CTGGTCCAGGTACCTCTACTGAACCGTCCGAAGGT AGCGCTCCAGGTAGCCCAGCAGGCTCTCCGACTTC CACTGAGGAAGGTACTTCTACTGAACCTTCCGAAG GCAGCGCACCAGGTACCTCTACTGAACCTTCTGAG GGCAGCGCTCCAGGTACTTCTGAAAGCGCTACCCC CGAATCTGGCCCAGGTAGCGAACCGGCTACTTCTG GTTCTGAAACCCCAGGTAGCGAACCGGCTACCTCC GGTTCTGAAACTCCAGGTAGCCCGGCAGGCTCTCC GACCTCTACTGAGGAAGGTACTTCTGAAAGCGCA ACCCCGGAGTCCGGCCCAGGTACCTCTACCGAACC GTCTGAGGGCAGCGCACCAGGTACTTCTACCGAA CCGTCCGAGGGTAGCGCACCAGGTAGCCCAGCAG GTTCTCCTACCTCCACCGAGGAAGGTACTTCTACC GAACCGTCCGAGGGTAGCGCACCAGGTACCTCTA CTGAACCTTCTGAGGGCAGCGCTCCAGGTACTTCT GAAAGCGCTACCCCGGAGTCCGGTCCAGGTACTTC TACTGAACCGTCCGAAGGTAGCGCACCAGGTACTT CTGAAAGCGCAACCCCTGAATCCGGTCCAGGTAG CGAACCGGCTACTTCTGGCTCTGAGACTCCAGGTA CTTCTACCGAACCGTCCGAAGGTAGCGCACCAGGT ACTTCTACTGAACCGTCTGAAGGTAGCGCACCAGG TACTTCTGAAAGCGCAACCCCGGAATCCGGCCCA		LASSRGIASFVCE YASPGKATEVRV TVLRQADSQVTE VCAATYMMGNE LTFLDDSICTGTS SGNQVNLTIQGL RAMDTGLYICKV ELMYPPPYYLGIG NGTQIYVIDPEPC PDSGAPSTGGTSE SATPESGPGTSTE PSEGSAPGTSTEP SEGSAPGTSESAT PESGPGTSTEPSE GSAPGTSTEPSEG SAPGTSESATPES GPGTSTEPSEGSA PGTSTEPSEGSAP GTSTEPSEGSAPG SPAGSPTSTEEGT STEPSEGSAPGPE PTGPAMHVAQPA VVLASSRGIASFV CEYASPGKATEV RVTVLRQADSQV TEVCAATYMMG NELTFLDDSICTG TSSGNQVNLTIQG LRAMDTGLYICK VELMYPPPYYLGI GNGTQIYVIDPEP CPDSGGSPAGSPT STEEGTSESATPE SGPGTSTEPSEGS APGSPAGSPTSTE EGTSTEPSEGSAP GTSTEPSEGSAPG TSESATPESGPGS EPATSGSETPGSE PATSGSETPGSPA GSPTSTEEGTSES ATPESGPGTSTEP SEGSAPGTSTEPS EGSAPGSPAGSPT STEEGTSTEPSEG SAPGTSTEPSEGS APGTSESATPESG PGTSTEPSEGSAP GTSESATPESGPG SEPATSGSETPGT STEPSEGSAPGTS TEPSEGSAPGTSE SATPESGPGTSES ATPESGPGSPAGS PTSTEEGTSESAT PESGPGSEPATSG SETPGTSESATPE SGPGTSTEPSEGS	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	GGTACCTCTGAAAGCGCAACCCCGGAGTCCGGCC CAGGTAGCCCTGCTGGCTCTCCAACCTCCACCGAA GAAGGTACCTCTGAAAGCGCAACCCCTGAATCCG GCCCAGGTAGCGAACCGGCAACCTCCGGTTCTGA AACCCCAGGTACCTCTGAAAGCGCTACTCCGGAGT CTGGCCCAGGTACCTCTACTGAACCGTCTGAGGGT AGCGCTCCAGGTACTTCTACTGAACCGTCCGAAGG TAGCGCACCAGGTACTTCTACCGAACCGTCCGAAG GCAGCGCTCCAGGTACCTCTACTGAACCTTCCGAG GGCAGCGCTCCAGGTACCTCTACCGAACCTTCTGA AGGTAGCGCACCAGGTACTTCTACCGAACCGTCCG AGGGTAGCGCACCAGGTAGCCCAGCAGGTTCTCC TACCTCCACCGAGGAAGGTACTTCTACCGAACCGT CCGAGGGTAGCGCACCAGGTACCTCTGAAAGCGC AACTCCTGAGTCTGGCCCAGGTAGCGAACCTGCTA CCTCCGGCTCTGAGACTCCAGGTACCTCTGAAAGC GCAACCCCGGAATCTGGTCCAGGTAGCGAACCTG CAACCTCTGGCTCTGAAACCCCAGGTACCTCTGAA AGCGCTACTCCTGAATCTGGCCCAGGTACTTCTAC TGAACCGTCCGAGGGCAGCGCACCAGGTACTTCT GAAAGCGCTACTCCTGAGTCCGGCCCAGGTAGCC CGGCTGGCTCTCCGACTTCCACCGAGGAAGGTAGC CCGGCTGGCTCTCCAACTTCTACTGAAGAAGGTAG CCCGGCAGGCTCTCCGACCTCTACTGAGGAAGGTA CTTCTGAAAGCGCAACCCCGGAGTCCGGCCCAGG TACCTCTACCGAACCGTCTGAGGGCAGCGCACCA GGTACCTCTGAAAGCGCAACTCCTGAGTCTGGCCC AGGTAGCGAACCTGCTACCTCCGGCTCTGAGACTC CAGGTACCTCTGAAAGCGCAACCCCGGAATCTGG TCCAGGTAGCGAACCTGCAACCTCTGGCTCTGAAA CCCCAGGTACCTCTGAAAGCGCTACTCCTGAATCT GGCCCAGGTACTTCTACTGAACCGTCCGAGGGCA GCGCACCAGGTAGCCCTGCTGGCTCTCCAACCTCC ACCGAAGAAGGTACCTCTGAAAGCGCAACCCCTG AATCCGGCCCAGGTAGCGAACCGGCAACCTCCGG TTCTGAAACCCCAGGTACTTCTGAAAGCGCTACTC CTGAGTCCGGCCCAGGTAGCCCGGCTGGCTCTCCG ACTTCCACCGAGGAAGGTAGCCCGGCTGGCTCTCC AACTTCTACTGAAGAAGGTACTTCTACCGAACCTT CCGAGGGCAGCGCACCAGGTACTTCTGAAAGCGC TACCCCTGAGTCCGGCCCAGGTACTTCTGAAAGCG CTACTCCTGAATCCGGTCCAGGTACTTCTGAAAGC GCTACCCCGGAATCTGGCCCAGGTAGCGAACCGG CTACTTCTGGTTCTGAAACCCCAGGTAGCGAACCG GCTACCTCCGGTTCTGAAACTCCAGGTAGCCCAGC AGGCTCTCCGACTTCCACTGAGGAAGGTACTTCTA CTGAACCTTCCGAAGGCAGCGCACCAGGTACCTCT ACTGAACCTTCTGAGGGCAGCGCTCCAGGTAGCG AACCTGCAACCTCTGGCTCTGAAACCCCAGGTACC TCTGAAAGCGCTACTCCTGAATCTGGCCCAGGTAC TTCTACTGAACCGTCCGAGGGCAGCGCACCAGGTT AA		APGTSTEPSEGSA PGTSTEPSEGSAP GTSTEPSEGSAPG TSTEPSEGSAPGT STEPSEGSAPGSP AGSPTSTEEGTST EPSEGSAPGTSES ATPESGPGSEPAT SGSETPGTSESAT PESGPGSEPATSG SETPGTSESATPE SGPGTSTEPSEGS APGTSESATPESG PGSPAGSPTSTEE GSPAGSPTSTEEG SPAGSPTSTEEGT SESATPESGPGTS TEPSEGSAPGTSE SATPESGPGSEPA TSGSETPGTSESA TPESGPGSEPATS GSETPGTSESATP ESGPGTSTEPSEG SAPGSPAGSPTST EEGTSESATPESG PGSEPATSGSETP GTSESATPESGPG SPAGSPTSTEEGS PAGSPTSTEEGTS TEPSEGSAPGTSE SATPESGPGTSES ATPESGPGTSESA TPESGPGSEPATS GSETPGSEPATSG SETPGSPAGSPTS TEEGTSTEPSEGS APGTSTEPSEGSA PGSEPATSGSETP GTSESATPESGPG TSTEPSEGSAPG	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
AE912-alL6R scFv, AC341	ATGGCTGAACCTGCTGGCTCTCCAACCTCCACTGA GGAAGGTACCCCGGGTAGCGGTACTGCTTCTTCCT CTCCAGGTAGCTCTACCCCTTCTGGTGCAACCGGC TCTCCAGGTGCTTCTCCGGGCACCAGCTCTACCGG TTCTCCAGGTAGCCCGGCTGGCTCTCCTACCTCTA CTGAGGAAGGTACTTCTGAAAGCGCTACTCCTGAG	702	MAEPAGSPTSTEE GTPGSGTASSSPG SSTPSGATGSPGA SPGTSSTGSPGSP AGSPTSTEEGTSE SATPESGPGTSTE	727

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	TCTGGTCCAGGTACCTCTACTGAACCGTCCGAAGG		PSEGSAPGSPAGS	
	TAGCGCTCCAGGTAGCCCAGCAGGCTCTCCGACTT		PTSTEEGTSTEPS	
	CCACTGAGGAAGGTACTTCTACTGAACCTTCCGAA		EGSAPGTSTEPSE	
	GGCAGCGCACCAGGTACCTCTACTGAACCTTCTGA		GSAPGTSESATPE	
	GGGCAGCGCTCCAGGTACTTCTGAAAGCGCTACCC		SGPGSEPATSGSE	
	CGGAATCTGGCCCAGGTAGCGAACCGGCTACTTCT		TPGSEPATSGSET	
	GGTTCTGAAACCCCAGGTAGCGAACCGGCTACCTC		PGSPAGSPTSTEE	
	CGGTTCTGAAACTCCAGGTAGCCCGGCAGGCTCTC		GTSESATPESGPG	
	CGACCTCTACTGAGGAAGGTACTTCTGAAAGCGCA		TSTEPSEGSAPGT	
	ACCCCGGAGTCCGGCCCAGGTACCTCTACCGAACC		STEPSEGSAPGSP	
	GTCTGAGGGCAGCGCACCAGGTACTTCTACCGAA		AGSPTSTEEGTST	
	CCGTCCGAGGGTAGCGCACCAGGTAGCCCAGCAG		EPSEGSAPGTSTE	
	GTTCTCCTACCTCCACCGAGGAAGGTACTTCTACC		PSEGSAPGTSESA	
	GAACCGTCCGAGGGTAGCGCACCAGGTACCTCTA		TPESGPGTSTEPS	
	CTGAACCTTCTGAGGGCAGCGCTCCAGGTACTTCT		EGSAPGTSESATP	
	GAAAGCGCTACCCCGGAGTCCGGTCCAGGTACTTC		ESGPGSEPATSGS	
	TACTGAACCGTCCGAAGGTAGCGCACCAGGTACTT		ETPGTSTEPSEGS	
	CTGAAAGCGCAACCCCTGAATCCGGTCCAGGTAG		APGTSTEPSEGSA	
	CGAACCGGCTACTTCTGGCTCTGAGACTCCAGGTA		PGTSESATPESGP	
	CTTCTACCGAACCGTCCGAAGGTAGCGCACCAGGT		GTSESATPESGPG	
	ACTTCTACTGAACCGTCTGAAGGTAGCGCACCAGG		SPAGSPTSTEEGT	
	TACTTCTGAAAGCGCAACCCCGGAATCCGGCCCA		SESATPESGPGSE	
	GGTACCTCTGAAAGCGCAACCCCGGAGTCCGGCC		PATSGSETPGTSE	
	CAGGTAGCCCTGCTGGCTCTCCAACCTCCACCGAA		SATPESGPGTSTE	
	GAAGGTACCTCTGAAAGCGCAACCCCTGAATCCG		PSEGSAPGTSTEP	
	GCCCAGGTAGCGAACCGGCAACCTCCGGTTCTGA		SEGSAPGTSTEPS	
	AACCCCAGGTACCTCTGAAAGCGCTACTCCGGAGT		EGSAPGTSTEPSE	
	CTGGCCCAGGTACCTCTACTGAACCGTCTGAGGGT		GSAPGTSTEPSEG	
	AGCGCTCCAGGTACTTCTACTGAACCGTCCGAAGG		SAPGTSTEPSEGS	
	TAGCGCACCAGGTACTTCTACCGAACCGTCCGAAG		APGSPAGSPTSTE	
	GCAGCGCTCCAGGTACCTCTACTGAACCTTCCGAG		EGTSTEPSEGSAP	
	GGCAGCGCTCCAGGTACCTCTACCGAACCTTCTGA		GTSESATPESGPG	
	AGGTAGCGCACCAGGTACTTCTACCGAACCGTCCG		SEPATSGSETPGT	
	AGGGTAGCGCACCAGGTAGCCCAGCAGGTTCTCC		SESATPESGPGSE	
	TACCTCCACCGAGGAAGGTACTTCTACCGAACCGT		PATSGSETPGTSE	
	CCGAGGGTAGCGCACCAGGTACCTCTGAAAGCGC		SATPESGPGTSTE	
	AACTCCTGAGTCTGGCCCAGGTAGCGAACCTGCTA		PSEGSAPGTSESA	
	CCTCCGGCTCTGAGACTCCAGGTACCTCTGAAAGC		TPESGPGSPAGSP	
	GCAACCCCGGAATCTGGTCCAGGTAGCGAACCTG		TSTEEGSPAGSPT	
	CAACCTCTGGCTCTGAAACCCCAGGTACCTCTGAA		STEEGSPAGSPTS	
	AGCGCTACTCCTGAATCTGGCCCAGGTACTTCTAC		TEEGTSESATPES	
	TGAACCGTCCGAGGGCAGCGCACCAGGTACTTCT		GPGTSTEPSEGSA	
	GAAAGCGCTACTCCTGAGTCCGGCCCAGGTAGCC		PGTSESATPESGP	
	CGGCTGGCTCTCCGACTTCCACCGAGGAAGGTAGC		GSEPATSGSETPG	
	CCGGCTGGCTCTCCAACTTCTACTGAAGAAGGTAG		TSESATPESGPGS	
	CCCGGCAGGCTCTCCGACCTCTACTGAGGAAGGTA		EPATSGSETPGTS	
	CTTCTGAAAGCGCAACCCCGGAGTCCGGCCCAGG		ESATPESGPGTST	
	TACCTCTACCGAACCGTCTGAGGGCAGCGCACCA		EPSEGSAPGSPAG	
	GGTACCTCTGAAAGCGCAACTCCTGAGTCTGGCCC		SPTSTEEGTSESA	
	AGGTAGCGAACCTGCTACCTCCGGCTCTGAGACTC		TPESGPGSEPATS	
	CAGGTACCTCTGAAAGCGCAACCCCGGAATCTGG		GSETPGTSESATP	
	TCCAGGTAGCGAACCTGCAACCTCTGGCTCTGAAA		ESGPGSPAGSPTS	
	CCCCAGGTACCTCTGAAAGCGCTACTCCTGAATCT		TEEGSPAGSPTST	
	GGCCCAGGTACTTCTACTGAACCGTCCGAGGGCA		EEGTSTEPSEGSA	
	GCGCACCAGGTAGCCCTGCTGGCTCTCCAACCTCC		PGTSESATPESGP	
	ACCGAAGAAGGTACCTCTGAAAGCGCAACCCCTG		GTSESATPESGPG	
	AATCCGGCCCAGGTAGCGAACCGGCAACCTCCGG		TSESATPESGPGS	
	TTCTGAAACCCCAGGTACTTCTGAAAGCGCTACTC		EPATSGSETPGSE	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	CTGAGTCCGGCCCAGGTAGCCCGGCTGGCTCTCCG ACTTCCACCGAGGAAGGTAGCCCGGCTGGCTCTCC AACTTCTACTGAAGAAGGTACTTCTACCGAACCTT CCGAGGGCAGCGCACCAGGTACTTCTGAAAGCGC TACCCCTGAGTCCGGCCCAGGTACTTCTGAAAGCG CTACTCCTGAATCCGGTCCAGGTACTTCTGAAAGC GCTACCCCGGAATCTGGCCCAGGTAGCGAACCGG CTACTTCTGGTTCTGAAACCCCAGGTAGCGAACCG GCTACCTCCGGTTCTGAAACTCCAGGTAGCCCAGC AGGCTCTCCGACTTCCACTGAGGAAGGTACTTCTA CTGAACCTTCCGAAGGCAGCGCACCAGGTACCTCT ACTGAACCTTCTGAGGGCAGCGCTCCAGGTAGCG AACCTGCAACCTCTGGCTCTGAAACCCCAGGTACC TCTGAAAGCGCTACTCCTGAATCTGGCCCAGGTAC TTCTACTGAACCGTCCGAGGGCAGCGCACCAGGT GCTGATATTCAAATGACTCAATCTCCTTCTTCTCTG TCCGCATCTGTAGGCGACCGTGTAACCATCACTTG CCGTGCCTCCCAGGACATCTCCAGCTACCTGAACT GGTACCAGCAGAAGCCGGGCAAGGCTCCGAAACT GCTGATTTATTACACTAGCCGTCTGCATTCTGGTGT TCCGAGCCGCTTCTCCGGTTCTGGCAGCGGTACCG ATTTCACTTTTACTATCTCCAGCCTGCAACCGGAG GACATCGCGACGTACTATTGCCAGCAAGGTAATAC CCTGCCGTACACCTTCGGCCAAGGCACGAAAGTTG AAATCAAAACCGGTTCTGGCGAAGGCTCTGAAGG TGAAGGTGGTGGTGAAGGCTCTGAAGGTGAAGGA TCTGGTGAAGGTGGCGAAGGTGAGGGTTCTGGAT CCCAAGTTCAGCTGCAGGAATCTGGTCCGGGTCTG GTTCGTCCGTCTCAGACCCTGTCCCTGACCTGCAC GGTGTCCGGCTACTCTATTACCTCTGACCATGCGT GGTCCTGGGTCCGTCAGCCACCGGGTCGCGGTCTG GAGTGGATCGGCTACATCAGCTACAGCGGCATCA CCACTTACAACCCGTCCCTGAAAAGCCGTGTCACC ATGCTGCGTGACACCTCCAAAAATCAATTCTCCCT GCGCCTGAGCTCTGTGACGGCGGCCGACACTGCG GTGTACTACTGCGCTCGCAGCCTGGCGCGTACCAC TGCTATGGATTACTGGGGTCAGGGCAGCCTGGTAA CCGTCAGCAGCTAA		PATSGSETPGSPA GSPTSTEEGTSTE PSEGSAPGTSTEP SEGSAPGSEPATS GSETPGTSESATP ESGPGTSTEPSEG SAPGADIQMTQS PSSLSASVGDRVT ITCRASQDISSYL NWYQQKPGKAP KLLIYYTSRLHSG VPSRFSGSGSGTD FTFTISSLQPEDIA TYYCQQGNTLPY TFGQGTKVEIKT GSGEGSEGEGGG EGSEGEGSGEGG EGEGSGSQVQLQ ESGPGLVRPSQTL SLTCTVSGYSITS DHAWSWVRQPP GRGLEWIGYISYS GITTYNPSLKSRV TMLRDTSKNQFS LRLSSVTAADTA VYYCARSLARTT AMDYWGQGSLV TVSS	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
alL6R scFv-AE864, AC342	ATGGCTGATATTCAAATGACTCAATCTCCTTCTTCT CTGTCCGCATCTGTAGGCGACCGTGTAACCATCAC TTGCCGTGCCTCCCAGGACATCTCCAGCTACCTGA ACTGGTACCAGCAGAAGCCGGGCAAGGCTCCGAA ACTGCTGATTTATTACACTAGCCGTCTGCATTCTG GTGTTCCGAGCCGCTTCTCCGGTTCTGGCAGCGGT ACCGATTTCACTTTTACTATCTCCAGCCTGCAACC GGAGGACATCGCGACGTACTATTGCCAGCAAGGT AATACCCTGCCGTACACCTTCGGCCAAGGCACGA AAGTTGAAATCAAAACCGGTTCTGGCGAAGGCTCT GAAGGTGAAGGTGGTGGTGAAGGCTCTGAAGGTG AAGGATCTGGTGAAGGTGGCGAAGGTGAGGGTTC TGGATCCCAAGTTCAGCTGCAGGAATCTGGTCCGG GTCTGGTTCGTCCGTCTCAGACCCTGTCCCTGACC TGCACGGTGTCCGGCTACTCTATTACCTCTGACCA TGCGTGGTCCTGGGTCCGTCAGCCACCGGGTCGCG GTCTGGAGTGGATCGGCTACATCAGCTACAGCGGC ATCACCACTTACAACCCGTCCCTGAAAAGCCGTGT CACCATGCTGCGTGACACCTCCAAAAATCAATTCT CCCTGCGCCTGAGCTCTGTGACGGCGGCCGACACT	703	MADIQMTQSPSS LSASVGDRVTITC RASQDISSYLNW YQQKPGKAPKLL IYYTSRLHSGVPS RFSGSGSGTDFTF TISSLQPEDIATY YCQQGNTLPYTF GQGTKVEIKTGS GEGSEGEGGGEG SEGEGSGEGGEG EGSGSQVQLQES GPGLVRPSQTLSL TCTVSGYSITSDH AWSWVRQPPGR GLEWIGYISYSGI TTYNPSLKSRVT MLRDTSKNQFSL RLSSVTAADTAV YYCARSLARTTA	728

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	GCGGTGTACTACTGCGCTCGCAGCCTGGCGCGTAC		MDYWGQGSLVT	
	CACTGCTATGGATTACTGGGGTCAGGGCAGCCTGG		VSSGSPGSPAGSP	
	TAACCGTCAGCAGCGGGTCTCCAGGTAGCCCGGCT		TSTEEGTSESATP	
	GGCTCTCCTACCTCTACTGAGGAAGGTACTTCTGA		ESGPGTSTEPSEG	
	AAGCGCTACTCCTGAGTCTGGTCCAGGTACCTCTA		SAPGSPAGSPTST	
	CTGAACCGTCCGAAGGTAGCGCTCCAGGTAGCCC		EEGTSTEPSEGSA	
	AGCAGGCTCTCCGACTTCCACTGAGGAAGGTACTT		PGTSTEPSEGSAP	
	CTACTGAACCTTCCGAAGGCAGCGCACCAGGTAC		GTSESATPESGPG	
	CTCTACTGAACCTTCTGAGGGCAGCGCTCCAGGTA		SEPATSGSETPGS	
	CTTCTGAAAGCGCTACCCCGGAATCTGGCCCAGGT		EPATSGSETPGSP	
	AGCGAACCGGCTACTTCTGGTTCTGAAACCCCAGG		AGSPTSTEEGTSE	
	TAGCGAACCGGCTACCTCCGGTTCTGAAACTCCAG		SATPESGPGTSTE	
	GTAGCCCGGCAGGCTCTCCGACCTCTACTGAGGAA		PSEGSAPGTSTEP	
	GGTACTTCTGAAAGCGCAACCCCGGAGTCCGGCC		SEGSAPGSPAGSP	
	CAGGTACCTCTACCGAACCGTCTGAGGGCAGCGC		TSTEEGTSTEPSE	
	ACCAGGTACTTCTACCGAACCGTCCGAGGGTAGC		GSAPGTSTEPSEG	
	GCACCAGGTAGCCCAGCAGGTTCTCCTACCTCCAC		SAPGTSESATPES	
	CGAGGAAGGTACTTCTACCGAACCGTCCGAGGGT		GPGTSTEPSEGSA	
	AGCGCACCAGGTACCTCTACTGAACCTTCTGAGGG		PGTSESATPESGP	
	CAGCGCTCCAGGTACTTCTGAAAGCGCTACCCCGG		GSEPATSGSETPG	
	AGTCCGGTCCAGGTACTTCTACTGAACCGTCCGAA		TSTEPSEGSAPGT	
	GGTAGCGCACCAGGTACTTCTGAAAGCGCAACCC		STEPSEGSAPGTS	
	CTGAATCCGGTCCAGGTAGCGAACCGGCTACTTCT		ESATPESGPGTSE	
	GGCTCTGAGACTCCAGGTACTTCTACCGAACCGTC		SATPESGPGSPAG	
	CGAAGGTAGCGCACCAGGTACTTCTACTGAACCGT		SPTSTEEGTSESA	
	CTGAAGGTAGCGCACCAGGTACTTCTGAAAGCGC		TPESGPGSEPATS	
	AACCCCGGAATCCGGCCCAGGTACCTCTGAAAGC		GSETPGTSESATP	
	GCAACCCCGGAGTCCGGCCCAGGTAGCCCTGCTG		ESGPGTSTEPSEG	
	GCTCTCCAACCTCCACCGAAGAAGGTACCTCTGAA		SAPGTSTEPSEGS	
	AGCGCAACCCCTGAATCCGGCCCAGGTAGCGAAC		APGTSTEPSEGSA	
	CGGCAACCTCCGGTTCTGAAACCCCAGGTACCTCT		PGTSTEPSEGSAP	
	GAAAGCGCTACTCCGGAGTCTGGCCCAGGTACCTC		GTSTEPSEGSAPG	
	TACTGAACCGTCTGAGGGTAGCGCTCCAGGTACTT		TSTEPSEGSAPGS	
	CTACTGAACCGTCCGAAGGTAGCGCACCAGGTAC		PAGSPTSTEEGTS	
	TTCTACCGAACCGTCCGAAGGCAGCGCTCCAGGTA		TEPSEGSAPGTSE	
	CCTCTACTGAACCTTCCGAGGGCAGCGCTCCAGGT		SATPESGPGSEPA	
	ACCTCTACCGAACCTTCTGAAGGTAGCGCACCAGG		TSGSETPGTSESA	
	TACTTCTACCGAACCGTCCGAGGGTAGCGCACCAG		TPESGPGSEPATS	
	GTAGCCCAGCAGGTTCTCCTACCTCCACCGAGGAA		GSETPGTSESATP	
	GGTACTTCTACCGAACCGTCCGAGGGTAGCGCACC		ESGPGTSTEPSEG	
	AGGTACCTCTGAAAGCGCAACTCCTGAGTCTGGCC		SAPGTSESATPES	
	CAGGTAGCGAACCTGCTACCTCCGGCTCTGAGACT		GPGSPAGSPTSTE	
	CCAGGTACCTCTGAAAGCGCAACCCCGGAATCTG		EGSPAGSPTSTEE	
	GTCCAGGTAGCGAACCTGCAACCTCTGGCTCTGAA		GSPAGSPTSTEEG	
	ACCCCAGGTACCTCTGAAAGCGCTACTCCTGAATC		TSESATPESGPGT	
	TGGCCCAGGTACTTCTACTGAACCGTCCGAGGGCA		STEPSEGSAPGTS	
	GCGCACCAGGTACTTCTGAAAGCGCTACTCCTGAG		ESATPESGPGSEP	
	TCCGGCCCAGGTAGCCCGGCTGGCTCTCCGACTTC		ATSGSETPGTSES	
	CACCGAGGAAGGTAGCCCGGCTGGCTCTCCAACTT		ATPESGPGSEPAT	
	CTACTGAAGAAGGTAGCCCGGCAGGCTCTCCGAC		SGSETPGTSESAT	
	CTCTACTGAGGAAGGTACTTCTGAAAGCGCAACCC		PESGPGTSTEPSE	
	CGGAGTCCGGCCCAGGTACCTCTACCGAACCGTCT		GSAPGSPAGSPTS	
	GAGGGCAGCGCACCAGGTACCTCTGAAAGCGCAA		TEEGTSESATPES	
	CTCCTGAGTCTGGCCCAGGTAGCGAACCTGCTACC		GPGSEPATSGSET	
	TCCGGCTCTGAGACTCCAGGTACCTCTGAAAGCGC		PGTSESATPESGP	
	AACCCCGGAATCTGGTCCAGGTAGCGAACCTGCA		GSPAGSPTSTEEG	
	ACCTCTGGCTCTGAAACCCCAGGTACCTCTGAAAG		SPAGSPTSTEEGT	
	CGCTACTCCTGAATCTGGCCCAGGTACTTCTACTG		STEPSEGSAPGTS	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	AACCGTCCGAGGGCAGCGCACCAGGTAGCCCTGC TGGCTCTCCAACCTCCACCGAAGAAGGTACCTCTG AAAGCGCAACCCCTGAATCCGGCCCAGGTAGCGA ACCGGCAACCTCCGGTTCTGAAACCCCAGGTACTT CTGAAAGCGCTACTCCTGAGTCCGGCCCAGGTAGC CCGGCTGGCTCTCCGACTTCCACCGAGGAAGGTAG CCCGGCTGGCTCTCCAACTTCTACTGAAGAAGGTA CTTCTACCGAACCTTCCGAGGGCAGCGCACCAGGT ACTTCTGAAAGCGCTACCCCTGAGTCCGGCCCAGG TACTTCTGAAAGCGCTACTCCTGAATCCGGTCCAG GTACTTCTGAAAGCGCTACCCCGGAATCTGGCCCA GGTAGCGAACCGGCTACTTCTGGTTCTGAAACCCC AGGTAGCGAACCGGCTACCTCCGGTTCTGAAACTC CAGGTAGCCCAGCAGGCTCTCCGACTTCCACTGAG GAAGGTACTTCTACTGAACCTTCCGAAGGCAGCGC ACCAGGTACCTCTACTGAACCTTCTGAGGGCAGCG CTCCAGGTAGCGAACCTGCAACCTCTGGCTCTGAA ACCCCAGGTACCTCTGAAAGCGCTACTCCTGAATC TGGCCCAGGTACTTCTACTGAACCGTCCGAGGGCA GCGCACCAGGTTAA		ESATPESGPGTSE SATPESGPGTSES ATPESGPGSEPAT SGSETPGSEPATS GSETPGSPAGSPT STEEGTSTEPSEG SAPGTSTEPSEGS APGSEPATSGSET PGTSESATPESGP GTSTEPSEGSAPG	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
AE912-alL6R-AE144, AC361	ATGGCTGAACCTGCTGGCTCTCCAACCTCCACTGA GGAAGGTACCCCGGGTAGCGGTACTGCTTCTTCCT CTCCAGGTAGCTCTACCCCTTCTGGTGCAACCGGC TCTCCAGGTGCTTCTCCGGGCACCAGCTCTACCGG TTCTCCAGGTAGCCCGGCTGGCTCTCCTACCTCTA CTGAGGAAGGTACTTCTGAAAGCGCTACTCCTGAG TCTGGTCCAGGTACCTCTACTGAACCGTCCGAAGG TAGCGCTCCAGGTAGCCCAGCAGGCTCTCCGACTT CCACTGAGGAAGGTACTTCTACTGAACCTTCCGAA GGCAGCGCACCAGGTACCTCTACTGAACCTTCTGA GGGCAGCGCTCCAGGTACTTCTGAAAGCGCTACCC CGGAATCTGGCCCAGGTAGCGAACCGGCTACTTCT GGTTCTGAAACCCCAGGTAGCGAACCGGCTACCTC CGGTTCTGAAACTCCAGGTAGCCCGGCAGGCTCTC CGACCTCTACTGAGGAAGGTACTTCTGAAAGCGCA ACCCCGGAGTCCGGCCCAGGTACCTCTACCGAACC GTCTGAGGGCAGCGCACCAGGTACTTCTACCGAA CCGTCCGAGGGTAGCGCACCAGGTAGCCCAGCAG GTTCTCCTACCTCCACCGAGGAAGGTACTTCTACC GAACCGTCCGAGGGTAGCGCACCAGGTACCTCTA CTGAACCTTCTGAGGGCAGCGCTCCAGGTACTTCT GAAAGCGCTACCCCGGAGTCCGGTCCAGGTACTTC TACTGAACCGTCCGAAGGTAGCGCACCAGGTACTT CTGAAAGCGCAACCCCTGAATCCGGTCCAGGTAG CGAACCGGCTACTTCTGGCTCTGAGACTCCAGGTA CTTCTACCGAACCGTCCGAAGGTAGCGCACCAGGT ACTTCTACTGAACCGTCTGAAGGTAGCGCACCAGG TACTTCTGAAAGCGCAACCCCGGAATCCGGCCCA GGTACCTCTGAAAGCGCAACCCCGGAGTCCGGCC CAGGTAGCCCTGCTGGCTCTCCAACCTCCACCGAA GAAGGTACCTCTGAAAGCGCAACCCCTGAATCCG GCCCAGGTAGCGAACCGGCAACCTCCGGTTCTGA AACCCCAGGTACCTCTGAAAGCGCTACTCCGGAGT CTGGCCCAGGTACCTCTACTGAACCGTCTGAGGGT AGCGCTCCAGGTACTTCTACTGAACCGTCCGAAGG TAGCGCACCAGGTACTTCTACCGAACCGTCCGAAG GCAGCGCTCCAGGTACCTCTACTGAACCTTCCGAG GGCAGCGCTCCAGGTACCTCTACCGAACCTTCTGA	704	MAEPAGSPTSTEE GTPGSGTASSSPG SSTPSGATGSPGA SPGTSSTGSPGSP AGSPTSTEEGTSE SATPESGPGTSTE PSEGSAPGSPAGS PTSTEEGTSTEPS EGSAPGTSTEPSE GSAPGTSESATPE SGPGSEPATSGSE TPGSEPATSGSET PGSPAGSPTSTEE GTSESATPESGPG TSTEPSEGSAPGT STEPSEGSAPGSP AGSPTSTEEGTST EPSEGSAPGTSTE PSEGSAPGTSESA TPESGPGTSTEPS EGSAPGTSESATP ESGPGSEPATSGS ETPGTSTEPSEGS APGTSTEPSEGSA PGTSESATPESGP GTSESATPESGPG SPAGSPTSTEEGT SESATPESGPGSE PATSGSETPGTSE SATPESGPGTSTE PSEGSAPGTSTEP SEGSAPGTSTEPS EGSAPGTSTEPSE GSAPGTSTEPSEG SAPGTSTEPSEGS APGSPAGSPTSTE EGTSTEPSEGSAP GTSESATPESGPG	729

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	AGGTAGCGCACCAGGTACTTCTACCGAACCGTCCG AGGGTAGCGCACCAGGTAGCCCAGCAGGTTCTCC TACCTCCACCGAGGAAGGTACTTCTACCGAACCGT CCGAGGGTAGCGCACCAGGTACCTCTGAAAGCGC AACTCCTGAGTCTGGCCCAGGTAGCGAACCTGCTA CCTCCGGCTCTGAGACTCCAGGTACCTCTGAAAGC GCAACCCCGGAATCTGGTCCAGGTAGCGAACCTG CAACCTCTGGCTCTGAAACCCCAGGTACCTCTGAA AGCGCTACTCCTGAATCTGGCCCAGGTACTTCTAC TGAACCGTCCGAGGGCAGCGCACCAGGTACTTCT GAAAGCGCTACTCCTGAGTCCGGCCCAGGTAGCC CGGCTGGCTCTCCGACTTCCACCGAGGAAGGTAGC CCGGCTGGCTCTCCAACTTCTACTGAAGAAGGTAG CCCGGCAGGCTCTCCGACCTCTACTGAGGAAGGTA CTTCTGAAAGCGCAACCCCGGAGTCCGGCCCAGG TACCTCTACCGAACCGTCTGAGGGCAGCGCACCA GGTACCTCTGAAAGCGCAACTCCTGAGTCTGGCCC AGGTAGCGAACCTGCTACCTCCGGCTCTGAGACTC CAGGTACCTCTGAAAGCGCAACCCCGGAATCTGG TCCAGGTAGCGAACCTGCAACCTCTGGCTCTGAAA CCCCAGGTACCTCTGAAAGCGCTACTCCTGAATCT GGCCCAGGTACTTCTACTGAACCGTCCGAGGGCA GCGCACCAGGTAGCCCTGCTGGCTCTCCAACCTCC ACCGAAGAAGGTACCTCTGAAAGCGCAACCCCTG AATCCGGCCCAGGTAGCGAACCGGCAACCTCCGG TTCTGAAACCCCAGGTACTTCTGAAAGCGCTACTC CTGAGTCCGGCCCAGGTAGCCCGGCTGGCTCTCCG ACTTCCACCGAGGAAGGTAGCCCGGCTGGCTCTCC AACTTCTACTGAAGAAGGTACTTCTACCGAACCTT CCGAGGGCAGCGCACCAGGTACTTCTGAAAGCGC TACCCCTGAGTCCGGCCCAGGTACTTCTGAAAGCG CTACTCCTGAATCCGGTCCAGGTACTTCTGAAAGC GCTACCCCGGAATCTGGCCCAGGTAGCGAACCGG CTACTTCTGGTTCTGAAACCCCAGGTAGCGAACCG GCTACCTCCGGTTCTGAAACTCCAGGTAGCCCAGC AGGCTCTCCGACTTCCACTGAGGAAGGTACTTCTA CTGAACCTTCCGAAGGCAGCGCACCAGGTACCTCT ACTGAACCTTCTGAGGGCAGCGCTCCAGGTAGCG AACCTGCAACCTCTGGCTCTGAAACCCCAGGTACC TCTGAAAGCGCTACTCCTGAATCTGGCCCAGGTAC TTCTACTGAACCGTCCGAGGGCAGCGCACCAGGT GCTGATATTCAAATGACTCAATCTCCTTCTTCTCTG TCCGCATCTGTAGGCGACCGTGTAACCATCACTTG CCGTGCCTCCCAGGACATCTCCAGCTACCTGAACT CGTACCAGCAGAAGCCGGGCAAGGCTCCGAAACT GCTGATTTATTACACTAGCCGTCTGCATTCTGGTGT TCCGAGCCGCTTCTCCGGTTCTGGCAGCGGTACCG ATTTCACTTTTACTATCTCCAGCCTGCAACCGGAG GACATCGCGACGTACTATTGCCAGCAAGGTAATAC CCTGCCGTACACCTTCGGCCAAGGCACGAAAGTTG AAATCAAAACCGGTTCTGGCGAAGGCTCTGAAGC TGAAGGTGGTGGTGAAGGCTCTGAAGGTGAAGGA TCTGGTGAAGGTGGCGAAGGTGAGGGTTCTGGAT CCCAAGTTCAGCTGCAGGAATCTGGTCCGGGTCTG GTTCGTCCGTCTCAGACCCTGTCCCTGACCTGCAC GGTGTCCGGCTACTCTATTACCTCTGACCATGCGT GGTCCTGGGTCCGTCAGCCACCGGGTCGCGGTCTG GAGTGGATCGGCTACATCAGCTACAGCGGCATCA		SEPATSGSETPGT SESATPESGPGSE PATSGSETPGTSE SATPESGPGTSTE PSTGSAPGTSTESA TPESGPGSPAGSP TSTEEGSPAGSPT STEEGSPAGSPTS TEEGTSESATPES GPGTSTEPSEGSA PGTSESATPESGP GSEPATSGSETPG TSESATPESGPGS EPATSGSETPGTS ESATPESGPGTST EPSEGSAPGSPAG SPTSTEEGTSESA TPESGPGSEPATS GSETPGTSESATP ESGPGSPAGSPTS TEEGSPAGSPTST EEGTSTEPSEGSA PGTSESATPESGP GTSESATPESGPG TSESATPESGPGS EPATSGSETPGSE PATSGSETPGSPA GSPTSTEEGTSTE PSEGSAPGTSTEP SEGSAPGSEPATS GSETPGTSESATP ESGPGTSTEPSEG SAPGADIQMTQS PSSLSASVGDRVT ITCRASQDISSYL NWYQQKPGKAP KLLIYYTSRLHSG VPSRFSGSGSGTD FTFTISSLQPEDIA TYYCQQGNTLPY TFGQGTKVEIKT GSGEGSEGEGGG EGSEGEGSGEGG EGEGSGSQVQLQ ESGPGLVRPSQTL SLTCTVSGYSITS DHAWSWVRQPP GRGLEWIGYISYS GITTYNPSLKSRV TMLRDTSKNQFS LRLSSVTAADTA VYYCARSLARTT AMDYWGQGSLV TVSSGGTSESATP ESGPGTSTEPSEG SAPGTSTEPSEGS APGTSESATPESG PGTSTEPSEGSAP	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	CCACTTACAACCCGTCCCTGAAAAGCCGTGTCACC ATGCTGCGTGACACCTCCAAAAATCAATTCTCCCT GCGCCTGAGCTCTGTGACGGCGGCCGACACTGCG GTGTACTACTGCGCTCGCAGCCTGGCGCGTACCAC TGCTATGGATTACTGGGGTCAGGGCAGCCTGGTAA CCGTCAGCAGCGGAGGTACTTCTGAAAGCGCTACT CCGGAGTCCGGTCCAGGTACCTCTACCGAACCGTC CGAAGGCAGCGCTCCAGGTACTTCTACTGAACCTT CTGAGGGTAGCGCTCCAGGTACTTCTGAAAGCGCT ACTCCGGAGTCCGGTCCAGGTACCTCTACCGAACC GTCCGAAGGCAGCGCTCCAGGTACTTCTACTGAAC CTTCTGAGGGTAGCGCTCCAGGTACCTCTGAAAGC GCTACTCCGGAGTCTGGCCCAGGTACCTCTACTGA ACCGTCTGAGGGTAGCGCTCCAGGTACTTCTACTG AACCGTCCGAAGGTAGCGCACCAGGTACTTCTACC GAACCGTCCGAGGGTAGCGCACCAGGTAGCCCAG CAGGTTCTCCTACCTCCACCGAGGAAGGTACTTCT ACCGAACCGTCCGAGGGTAGCGCACCAGGTTAA		GTSTEPSEGSAPG TSESATPESGPGT STEPSEGSAPGTS TEPSEGSAPGTST EPSEGSAPGSPAG SPTSTEEGTSTEPS EGSAPG	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
AE48-alL6R-AE864, AC362	ATGGCTGAACCTGCTGGCTCTCCAACCTCCACTGA GGAAGGTACCCCGGGTAGCGGTACTGCTTCTTCCT CTCCAGGTAGCTCTACCCCTTCTGGTGCAACCGGC TCTCCAGGTGCTTCTCCGGGCACCAGCTCTACCGG TTCTCCAgggtctccaggtGATATTCAAATGACTCAATCT CCTTCTTCTCTGTCCGCATCTGTAGGCGACCGTGTA ACCATCACTTGCCGTGCCTCCCAGGACATCTCCAG CTACCTGAACTGGTACCAGCAGAAGCCGGGCAAG GCTCCGAAACTGCTGATTTATTACACTAGCCGTCT GCATTCTGGTGTTCCGAGCCGCTTCTCCGGTTCTG GCAGCGGTACCGATTTCACTTTTACTATCTCCAGC CTGCAACCGGAGGACATCGCGACGTACTATTGCC AGCAAGGTAATACCCTGCCGTACACCTTCGGCCAA GGCACGAAAGTTGAAATCAAAACCGGTTCTGGCG AAGGCTCTGAAGGTGAAGGTGGTGGTGAAGGCTC TGAAGGTGAAGGATCTGGTGAAGGTGGCGAAGGT GAGGGTTCTGGATCCCAAGTTCAGCTGCAGGAATC TGGTCCGGGTCTGGTTCGTCCGTCTCAGACCCTGT CCCTGACCTGCACGGTGTCCGGCTACTCTATTACC TCTGACCATGCGTGGTCCTGGGTCCGTCAGCCACC GGGTCGCGGTCTGGAGTGGATCGGCTACATCAGCT ACAGCGGCATCACCACTTACAACCCGTCCCTGAAA AGCCGTGTCACCATGCTGCGTGACACCTCCAAAAA TCAATTCTCCCTGCGCCTGAGCTCTGTGACGGCGG CCGACACTGCGGTGTACTACTGCGCTCGCAGCCTG GCGCGTACCACTGCTATGGATTACTGGGGTCAGGG CAGCCTGGTAACCGTCAGCAGCGGAGGTAGCCCG GCTGGCTCTCCTACCTCTACTGAGGAAGGTACTTC TGAAAGCGCTACTCCTGAGTCTGGTCCAGGTACCT CTACTGAACCGTCCGAAGGTAGCGCTCCAGGTAG CCCAGCAGGCTCTCCGACTTCCACTGAGGAAGGTA CTTCTACTGAACCTTCCGAAGGCAGCGCACCAGGT ACCTCTACTGAACCTTCTGAGGGCAGCGCTCCAGG TACTTCTGAAAGCGCTACCCCGGAATCTGGCCCAG GTAGCGAACCGGCTACTTCTGGTTCTGAAACCCCA GGTAGCGAACCGGCTACCTCCGGTTCTGAAACTCC AGGTAGCCCGGCAGGCTCTCCGACCTCTACTGAGG AAGGTACTTCTGAAAGCGCAACCCCGGAGTCCGG CCCAGGTACCTCTACCGAACCGTCTGAGGGCAGC GCACCAGGTACTTCTACCGAACCGTCCGAGGGTA	705	MAEPAGSPTSTEE GTPGSGTASSSPG SSTPSGATGSPGA SPGTSSTGSPGSP GDIQMTQSPSSLS ASVGDRVTITCR ASQDISSYLNWY QQKPGKAPKLLI YYTSRLHSGVPS RFSGSGSGTDFTF TISSLQPEDIATY YCQQGNTLPYTF GQGTKVEIKTGS GEGSEGEGGGEG SEGEGSGEGGEG EGSGSQVQLQES GPGLVRPSQTLSL TCTVSGYSITSDH AWSWVRQPPGR GLEWIGYISYSGI TTYNPSLKSRVT MLRDTSKNQFSL RLSSVTAADTAV YYCARSLARTTA MDYWGQGSLVT VSSGGSPAGSPTS TEEGTSESATPES GPGTSTEPSEGSA PGSPAGSPTSTEE GTSTEPSEGSAPG TSTEPSEGSAPGT SESATPESGPGSE PATSGSETPGSEP ATSGSETPGSPAG SPTSTEEGTSESA TPESGPGTSTEPS EGSAPGTSTEPSE GSAPGSPAGSPTS TEEGTSTEPSEGS APGTSTEPSEGSA	730

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	GCGCACCAGGTAGCCCAGCAGGTTCTCCTACCTCC ACCGAGGAAGGTACTTCTACCGAACCGTCCGAGG GTAGCGCACCAGGTACCTCTACTGAACCTTCTGAG GGCAGCGCTCCAGGTACTTCTGAAAGCGCTACCCC GGAGTCCGGTCCAGGTACTTCTACTGAACCGTCCG AAGGTAGCGCACCAGGTACTTCTGAAAGCGCAAC CCCTGAATCCGGTCCAGGTAGCGAACCGGCTACTT CTGGCTCTGAGACTCCAGGTACTTCTACCGAACCG TCCGAAGGTAGCGCACCAGGTACTTCTACTGAACC GTCTGAAGGTAGCGCACCAGGTACTTCTGAAAGC GCAACCCCGGAATCCGGCCCAGGTACCTCTGAAA GCGCAACCCCGGAGTCCGGCCCAGGTAGCCCTGC TGGCTCTCCAACCTCCACCGAAGAAGGTACCTCTG AAAGCGCAACCCCTGAATCCGGCCCAGGTAGCGA ACCGGCAACCTCCGGTTCTGAAACCCCAGGTACCT CTGAAAGCGCTACTCCGGAGTCTGGCCCAGGTACC TCTACTGAACCGTCTGAGGGTAGCGCTCCAGGTAC TTCTACTGAACCGTCCGAAGGTAGCGCACCAGGTA CTTCTACCGAACCGTCCGAAGGCAGCGCTCCAGGT ACCTCTACTGAACCTTCCGAGGGCAGCGCTCCAGG TACCTCTACCGAACCTTCTGAAGGTAGCGCACCAG GTACTTCTACCGAACCGTCCGAGGGTAGCGCACCA GGTAGCCCAGCAGGTTCTCCTACCTCCACCGAGGA AGGTACTTCTACCGAACCGTCCGAGGGTAGCGCA CCAGGTACCTCTGAAAGCGCAACTCCTGAGTCTGG CCCAGGTAGCGAACCTGCTACCTCCGGCTCTGAGA CTCCAGGTACCTCTGAAAGCGCAACCCCGGAATCT GGTCCAGGTAGCGAACCTGCAACCTCTGGCTCTGA AACCCCAGGTACCTCTGAAAGCGCTACTCCTGAAT CTGGCCCAGGTACTTCTACTGAACCGTCCGAGGGC AGCGCACCAGGTACTTCTGAAAGCGCTACTCCTGA GTCCGGCCCAGGTAGCCCGGCTGGCTCTCCGACTT CCACCGAGGAAGGTAGCCCGGCTGGCTCTCCAAC TTCTACTGAAGAAGGTAGCCCGGCAGGCTCTCCGA CCTCTACTGAGGAAGGTACTTCTGAAAGCGCAACC CCGGAGTCCGGCCCAGGTACCTCTACCGAACCGTC TGAGGGCAGCGCACCAGGTACCTCTGAAAGCGCA ACTCCTGAGTCTGGCCCAGGTAGCGAACCTGCTAC CTCCGGCTCTGAGACTCCAGGTACCTCTGAAAGCG CAACCCCGGAATCTGGTCCAGGTAGCGAACCTGC AACCTCTGGCTCTGAAACCCCAGGTACCTCTGAAA GCGCTACTCCTGAATCTGGCCCAGGTACTTCTACT GAACCGTCCGAGGGCAGCGCACCAGGTAGCCCTG CTGGCTCTCCAACCTCCACCGAAGAAGGTACCTCT GAAAGCGCAACCCCTGAATCCGGCCCAGGTAGCG AACCGGCAACCTCCGGTTCTGAAACCCCAGGTACT TCTGAAAGCGCTACTCCTGAGTCCGGCCCAGGTAG CCCGGCTGGCTCTCCGACTTCCACCGAGGAAGGTA GCCCGGCTGGCTCTCCAACTTCTACTGAAGAAGGT ACTTCTACCGAACCTTCCGAGGCAGCGCACCAG GTACTTCTGAAAGCGCTACCCCTGAGTCCGGCCCA GGTACTTCTGAAAGCGCTACTCCTGAATCCGGTCC AGGTACTTCTGAAAGCGCTACCCCGGAATCTGGCC CAGGTAGCGAACCGGCTACTTCTGGTTCTGAAACC CCAGGTAGCGAACCGGCTACCTCCGGTTCTGAAAC TCCAGGTAGCCCAGCAGGCTCTCCGACTTCCACTG AGGAAGGTACTTCTACTGAACCTTCCGAAGGCAGC GCACCAGGTACCTCTACTGAACCTTCTGAGGGCAG		PGTSESATPESGP GTSTEPSEGSAPG TSESATPESGPGS EPATSGSETPGTS TEPSEGSAPGTST EPSEGSAPGTSES ATPESGPGTSESA TPESGPGSPAGSP TSTEEGTSESATP ESGPGSEPATSGS ETPGTSESATPES GPGTSTEPSEGSA PGTSTEPSEGSAP GTSTEPSEGSAPG TSTEPSEGSAPGT STEPSEGSAPGTS TEPSEGSAPGSPA GSPTSTEEGTSTE PSEGSAPGTSESA TPESGPGSEPATS GSETPGTSESATP ESGPGSEPATSGS ETPGTSESATPES GPGTSTEPSEGSA PGTSESATPESGP GSPAGSPTSTEEG SPAGSPTSTEEGS PAGSPTSTEEGTS ESATPESGPGTST EPSEGSAPGTSES ATPESGPGSEPAT SGSETPGTSESAT PESGPGSEPATSG SETPGTSESATPE SGPGTSTEPSEGS APGSPAGSPTSTE EGTSESATPESGP GSEPATSGSETPG TSESATPESGPGS PAGSPTSTEEGSP AGSPTSTEEGTST EPSEGSAPGTSES ATPESGPGTSESA TPESGPGTSESAT PESGPGSEPATSG SETPGSEPATSGS ETPGSPAGSPTST EEGTSTEPSEGSA PGTSTEPSEGSAP GSEPATSGSETPG TSESATPESGPGT STEPSEGSAPG	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	CGCTCCAGGTAGCGAACCTGCAACCTCTGGCTCTG AAACCCCAGGTACCTCTGAAAGCGCTACTCCTGAA TCTGGCCCAGGTACTTCTACTGAACCGTCCGAGGG CAGCGCACCAGGTTAA			

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
anti-CD40-AE864, AC384	ATGGCTGAAATTGTTCTGACCCAATCTCCTGCAAC TCTGTCTCTGTCTCCAGGTGAACGCGCCACCCTGT CTTGTCGTGCGTCCCAGTCTATCTCTGATTATCTGC ATTGGTATCAGCAGAAACCTGGCCAGGCTCCGCG CCTGCTGATCTATTACGCCAGCCACAGCATCTCTG GTATCCCGGCTCGCTTCTCCGGCTCCGGCAGCGGC ACCGACTTCACTCTGACTATTAGCTCCCTGGAACC GGAGGATTTCGCAGTTTATTACTGTCAGCACGGTC ACTCCTACCCGTGGACCTTTGGTGGCGGCACCAAA GTTGAAATCAAAACCGGTTCTGGCGAAGGCTCTGA AGGTGAAGGTGGTGGTGAAGGCTCTGAAGGTGAA GGATCTGGTGAAGGTGGCGAAGGTGAGGGATCTG GTACCCAAGTCCAGCTGGTTCAGTCCGGCTCTGAA CTGAAGAAACCGGGCGCTTCTGTTAAAGTTAGCTG CAAAGCAAGCGGTTATGCCTTTACTACTACTGGTA TGCAGTGGGTCCGCCAGGCACCGGGTCAGGGCCT GGAGTGGATGGGCTGGATCAACACCCACTCTGGT GTCCCTAAATACGTTGAAGATTTCAAAGGCCGTTT CGTGTTCTCCCTGGACACTTCCGTCAGCACCGCGT ATCTGCAGATCAGCAGCCTGAAAGCTGAGGACAC CGCGGTTTATTACTGCGCGCGTAGCGGCAATGGTA ACTACGACCTGGCTTATTTCAAATACTGGGGTCAG GGCACTCTGGTTACTGTGTCTAGCGGAGGTAGCCC GGCTGGCTCTCCTACCTCTACTGAGGAAGGTACTT CTGAAAGCGCTACTCCTGAGTCTGGTCCAGGTACC TCTACTGAACCGTCCGAAGGTAGCGCTCCAGGTAG CCCAGCAGGCTCTCCGACTTCCACTGAGGAAGGTA CTTCTACTGAACCTTCCGAAGGCAGCGCACCAGGT ACCTCTACTGAACCTTCTGAGGGCAGCGCTCCAGG TACTTCTGAAAGCGCTACCCCGGAATCTGGCCCAG GTAGCGAACCGGCTACTTCTGGTTCTGAAACCCCA GGTAGCGAACCGGCTACCTCCGGTTCTGAAACTCC AGGTAGCCCGGCAGGCTCTCCGACCTCTACTGAGG AAGGTACTTCTGAAAGCGCAACCCCGGAGTCCGG CCCAGGTACCTCTACCGAACCGTCTGAGGGCAGC GCACCAGGTACTTCTACCGAACCGTCCGAGGGTA GCGCACCAGGTAGCCCAGCAGGTTCTCCTACCTCC ACCGAGGAAGGTACTTCTACCGAACCGTCCGAGG GTAGCGCACCAGGTACCTCTACTGAACCTTCTGAG GGCAGCGCTCCAGGTACTTCTGAAAGCGCTACCCC GGAGTCCGGTCCAGGTACTTCTACTGAACCGTCCG AAGGTAGCGCACCAGGTACTTCTGAAAGCGCAAC CCCTGAATCCGGTCCAGGTAGCGAACCGGCTACTT CTGGCTCTGAGACTCCAGGTACTTCTACCGAACCG TCCGAAGGTAGCGCACCAGGTACTTCTACTGAACC GTCTGAAGGTAGCGCACCAGGTACTTCTGAAAGC GCAACCCCCGGAATCCGGCCCAGGTACCTCTGAAA GCGCAACCCCCGGAGTCCGGCCCAGGTAGCCCTGC TGGCTCTCCAACCTCCACCGAAGAAGGTACCTCTG AAAGCGCAACCCCTGAATCCGGCCCAGGTAGCGA ACCGGCAACCTCCGGTTCTGAAACCCCAGGTACCT CTGAAAGCGCTACTCCGGAGTCTGGCCCAGGTACC TCTACTGAACCGTCTGAGGGTAGCGCTCCAGGTAC TTCTACTGAACCGTCCGAAGGTAGCGCACCAGGTA	706	MAEIVLTQSPATL SLSPGERATLSCR ASQSISDYLHWY QQKPGQAPRLLI YYASHSISGIPAR FSGSGSGTDFTLT ISSLEPEDFAVYY CQHGHSYPWTFG GGTKVEIKTGSG EGSEGEGGGEGS EGEGSGEGGEGE GSGTQVQLVQSG SELKKPGASVKV SCKASGYAFTTT GMQWVRQAPGQ GLEWMGWINTH SGVPKYVEDFKG RFVFSLDTSVSTA YLQISSLKAEDTA VYYCARSGNGN YDLAYFKYWGQ GTLVTVSSGGSP AGSPTSTEEGTSE SATPESGPGTSTE PSEGSAPGSPAGS PTSTEEGTSTEPS EGSAPGTSTEPSE GSAPGTSESATPE SGPGSEPATSGSE TPGSEPATSGSET PGSPAGSPTSTEE GTSESATPESGPG TSTEPSEGSAPGT STEPSEGSAPGSP AGSPTSTEEGTST EPSEGSAPGTSTE PSEGSAPGTSESA TPESGPGTSTEPS EGSAPGTSESATP ESGPGSEPATSGS ETPGTSTEPSEGS APGTSTEPSEGSA PGTSESATPESGP GTSESATPESGPG SPAGSPTSTEEGT SESATPESGPGSE PATSGSETPGTSE SATPESGPGTSTE PSEGSAPGTSTEP SEGSAPGTSTEPS EGSAPGTSTEPSE GSAPGTSTEPSEG SAPGTSTEPSEGS APGSPAGSPTSTE	731

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	CTTCTACCGAACCGTCCGAAGGCAGCGCTCCAGGT ACCTCTACTGAACCTTCCGAGGGCAGCGCTCCAGG TACCTCTACCGAACCTTCTGAAGGTAGCGCACCAG GTACTTCTACCGAACCGTCCGAGGGTAGCGCACCA GGTAGCCCAGCAGGTTCTCCTACCTCCACCGAGGA AGGTACTTCTACCGAACCGTCCGAGGGTAGCGCA CCAGGTACCTCTGAAAGCGCAACTCCTGAGTCTGG CCCAGGTAGCGAACCTGCTACCTCCGGCTCTGAGA CTCCAGGTACCTCTGAAAGCGCAACCCCGGAATCT GGTCCAGGTAGCGAACCTGCAACCTCTGGCTCTGA AACCCCAGGTACCTCTGAAAGCGCTACTCCTGAAT CTGGCCCAGGTACTTCTACTGAACCGTCCGAGGGC AGCGCACCAGGTACTTCTGAAAGCGCTACTCCTGA GTCCGGCCCAGGTAGCCCGGCTGGCTCTCCGACTT CCACCGAGGAAGGTAGCCCGGCTGGCTCTCCAAC TTCTACTGAAGAAGGTAGCCCGGCAGGCTCTCCGA CCTCTACTGAGGAAGGTACTTCTGAAAGCGCAACC CCGGAGTCCGGCCCAGGTACCTCTACCGAACCGTC TGAGGGCAGCGCACCAGGTACCTCTGAAAGCGCA ACTCCTGAGTCTGGCCCAGGTAGCGAACCTGCTAC CTCCGGCTCTGAGACTCCAGGTACCTCTGAAAGCG CAACCCCGGAATCTGGTCCAGGTAGCGAACCTGC AACCTCTGGCTCTGAAACCCCAGGTACCTCTGAAA GCGCTACTCCTGAATCTGGCCCAGGTACTTCTACT GAACCGTCCGAGGGCAGCGCACCAGGTAGCCCTG CTGGCTCTCCAACCTCCACCGAAGAAGGTACCTCT GAAAGCGCAACCCCTGAATCCGGCCCAGGTAGCG AACCGGCAACCTCCGGTTCTGAAACCCCAGGTACT TCTGAAAGCGCTACTCCTGAGTCCGGCCCAGGTAG CCCGGCTGGCTCTCCGACTTCCACCGAGGAAGGTA GCCCGGCTGGCTCTCCAACTTCTACTGAAGAAGGT ACTTCTACCGAACCTTCCGAGGGCAGCGCACCAG GTACTTCTGAAAGCGCTACCCCTGAGTCCGGCCCA GGTACTTCTGAAAGCGCTACTCCTGAATCCGGTCC AGGTACTTCTGAAAGCGCTACCCCGGAATCTGGCC CAGGTAGCGAACCGGCTACTTCTGGTTCTGAAACC CCAGGTAGCGAACCGGCTACCTCCGGTTCTGAAAC TCCAGGTAGCCCAGCAGGCTCTCCGACTTCCACTG AGGAAGGTACTTCTACTGAACCTTCCGAAGGCAGC GCACCAGGTACCTCTACTGAACCTTCTGAGGGCAG CGCTCCAGGTAGCGAACCTGCAACCTCTGGCTCTG AAACCCCAGGTACCTCTGAAAGCGCTACTCCTGAA TCTGGCCCAGGTACTTCTACTGAACCGTCCGAGGG CAGCGCACCA		EGTSTEPSEGSAP GTSESATPESGPG SEPATSGSETPGT SESATPESGPGSE PATSGSETPGTSE SATPESGPGTSTE PSEGSAPGTSESA TPESGPGSPAGSP TSTEEGSPAGSPT STEEGSPAGSPTS TEEGTSESATPES GPGTSTEPSEGSA PGTSESATPESGP GSEPATSGSETPG TSESATPESGPGS EPATSGSETPGTS ESATPESGPGTST EPSEGSAPGSPAG SPTSTEEGTSESA TPESGPGSEPATS GSETPGTSESATP ESGPGSPAGSPTS TEEGSPAGSPTST EEGTSTEPSEGSA PGTSESATPESGP GTSESATPESGPG TSESATPESGPGS EPATSGSETPGSE PATSGSETPGSPA GSPTSTEEGTSTE PSEGSAPGTSTEP SEGSAPGSEPATS GSETPGTSESATP ESGPGTSTEPSEG SAP	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
anti-CD40-AE864, AC385	ATGGCTGAAATTGTTCTGACTCAATCTCCAGCAAC TCTGTCTCTGTCTCCAGGTGAACGTGCAACCCTGT CTTGCCGTGCGTCCCAGTCCATCTCCGATTATCTG CATTGGTATCAGCAGAAACCGGGTCAGGCGCCTC GTCTGCTGATCTATTATGCGTCTCACTCCATTTCCG GTATCCCGGCACGTTTCTCTGGCAGCGGCAGCGGC ACCGATTTCACCCTGACGATCTCTTCTCTGGAACC GGAAGATTTCGCAGTCTATTATTGTCAGCATGGTC ACAGCTACCCGTGGACCTTCGGCGGTGGCACGAA AGTTGAAATCAAGACCGGTTCTGGCGAAGGCTCTG AAGGTGAAGGTGGTGGTGAAGGCTCTGAAGGTGA AGGATCTGGTGAAGGTGGCGAAGGTGAGGGATCT GGTACCCAGGTCCAGCTGGTTCAAAGCGGCTCTGA ACTGAAAAAGCCGGGTGCCTCTGTCAAAGTGTCTT	707	MAEIVLTQSPATL SLSPGERATLSCR ASQSISDYLHWY QQKPGQAPRLLI YYASHSISGIPAR FSGSGSGTDFTLT ISSLEPEDFAVYY CQHGHSYPWTFG GGTKVEIKTGSG EGSEGEGGGEGS EGEGSGEGGEGE GSGTQVQLVQSG SELKKPGASVKV SCKASGYAFTTT	732

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	GCAAGGCAAGCGGCTACGCGTTTACGACCACCGG CATGCAGTGGGTCCGTCAGGCCCCGGGCCAGGGT CTGGAATGGATGGGCTGGATCAACACCCATTCTGG CGTACCGAAATACGTTGAAGATTTCAAAGGCCGTT TCGTGTTCTCCCTGGATACGTCCGTTTCCACCGCCT ACCTGCAGATCTCTTCCCTGAAAGCAGAAGATACT GCGGTGTACTATTGCGCACGTAGCGGCAACGGCA ACTACGACCTGGCCTACTTCAAATACTGGGGTCAG GGTACTCTGGTGACCGTATCCTCTGGAGGTAGCCC GGCTGGCTCTCCTACCTCTACTGAGGAAGGTACTT CTGAAAGCGCTACTCCTGAGTCTGGTCCAGGTACC TCTACTGAACCGTCCGAAGGTAGCGCTCCAGGTAG CCCAGCAGGCTCTCCGACTTCCACTGAGGAAGGTA CTTCTACTGAACCTTCCGAAGGCAGCGCACCAGGT ACCTCTACTGAACCTTCTGAGGGCAGCGCTCCAGG TACTTCTGAAAGCGCTACCCCGCAATCTGGCCCAG GTAGCGAACCGGCTACTTCTGGTTCTGAAACCCCA GGTAGCGAACCGGCTACCTCCGGTTCTGAAACTCC AGGTAGCCCGGCAGGCTCTCCGACCTCTACTGAGG AAGGTACTTCTGAAAGCGCAACCCCCGGAGTCCGG CCCAGGTACCTCTACCGAACCGTCTGAGGGCAGC GCACCAGGTACTTCTACCGAACCGTCCGAGGGTA GCGCACCAGGTAGCCCAGCAGGTTCTCCTACCTCC ACCGAGGAAGGTACTTCTACCGAACCGTCCGAGG GTAGCGCACCAGGTACCTCTACTGAACCTTCTGAG GGCAGCGCTCCAGGTACTTCTGAAAGCGCTACCCC GGAGTCCGGTCCAGGTACTTCTACTGAACCGTCCG AAGGTAGCGCACCAGGTACTTCTGAAAGCGCAAC CCCTGAATCCGGTCCAGGTAGCGAACCGGCTACTT CTGGCTCTGAGACTCCAGGTACTTCTACCGAACCG TCCGAAGGTAGCGCACCAGGTACTTCTACTGAACC GTCTGAAGGTAGCGCACCAGGTACTTCTGAAAGC GCAACCCCGGAATCCGGCCCAGGTACCTCTGAAA GCGCAACCCCGGAGTCCGGCCCAGGTAGCCCTGC TGGCTCTCCAACCTCCACCGAAGAAGGTACCTCTG AAAGCGCAACCCCTGAATCCGGCCCAGGTAGCGA ACCGGCAACCTCCGGTTCTGAAACCCCAGGTACCT CTGAAAGCGCTACTCCGGAGTCTGGCCCAGGTACC TCTACTGAACCGTCTGAGGGTAGCGCTCCAGGTAC TTCTACTGAACCGTCCGAAGGTAGCGCACCAGGTA CTTCTACCGAACCGTCCGAAGGCAGCGCTCCAGGT ACCTCTACTGAACCTTCCGAGGGCAGCGCTCCAGG TACCTCTACCGAACCTTCTGAAGGTAGCGCACCAG GTACTTCTACCGAACCGTCCGAGGGTAGCGCACCA GGTAGCCCAGCAGGTTCTCCTACCTCCACCGAGGA AGGTACTTCTACCGAACCGTCCGAGGGTAGCGCA CCAGGTACCTCTGAAAGCGCAACTCCTGAGTCTGG CCCAGGTAGCGAACCTGCTACCTCCGGCTCTGAGA CTCCAGGTACCTCTGAAAGCGCAACCCCGGAATCT GGTCCAGGTAGCGAACCTGCAACCTCTGGCTCTGA AACCCCAGGTACCTCTGAAAGCGCTACTCCTGAAT CTGGCCCAGGTACTTCTACTGAACCGTCCGAGGGC AGCGCACCAGGTACTTCTGAAAGCGCTACTCCTGA GTCCGGCCCAGGTAGCCCGGCTGGCTCTCCGACTT CCACCGAGGAAGGTAGCCCGGCTGGCTCTCCAAC TTCTACTGAAGAAGGTAGCCCGGCAGGCTCTCCGA CCTCTACTGAGGAAGGTACTTCTGAAAGCGCAACC CCGGAGTCCGGCCCAGGTACCTCTACCGAACCGTC		GMQWVRQAPGQ GLEWMGWINTH SGVPKYVEDFKG RFVFSLDTSVSTA YLQISSLKAEDTA VYYCARSGNGN YDLAYFKYWGQ GTLVTVSSGGSP AGSPTSTEEGTSE SATPESGPGTSTE PSEGSAPGSPAGS PTSTEEGTSTEPS EGSAPGTSTEPSE GSAPGTSESATPE SGPGSEPATSGSE TPGSEPATSGSET PGSPAGSPTSTEE GTSESATPESGPG TSTEPSEGSAPGT STEPSEGSAPGSP AGSPTSTEEGTST EPSEGSAPGTSTE PSEGSAPGTSESA TPESGPGTSTEPS EGSAPGTSESATP ESGPGSEPATSGS ETPGTSTEPSEGS APGTSTEPSEGSA PGTSESATPESGP GTSESATPESGPG SPAGSPTSTEEGT SESATPESGPGSE PATSGSETPGTSE SATPESGPGTSTE PSEGSAPGTSTEP SEGSAPGTSTEPS EGSAPGTSTEPSE GSAPGTSTEPSEG SAPGTSTEPSEGS APGSPAGSPTSTE EGTSTEPSEGSAP GTSESATPESGPG SEPATSGSETPGT SESATPESGPGSE PATSGSETPGTSE SATPESGPGTSTE PSEGSAPGTSESA TPESGPGSPAGSP TSTEEGSPAGSPT STEEGSPAGSPTS TEEGTSESATPES GPGTSTEPSEGSA PGTSESATPESGP GSEPATSGSETPG TSESATPESGPGS EPATSGSETPGTS ESATPESGPGTST EPSEGSAPGSPAG	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	TGAGGGCAGCGCACCAGGTACCTCTGAAAGCGCA ACTCCTGAGTCTGGCCCAGGTAGCGAACCTGCTAC CTCCGGCTCTGAGACTCCAGGTACCTCTGAAAGCG CAACCCCGGAATCTGGTCCAGGTAGCGAACCTGC AACCTCTGGCTCTGAAACCCCAGGTACCTCTGAAA GCGCTACTCCTGAATCTGGCCCAGGTACTTCTACT GAACCGTCCGAGGGCAGCGCACCAGGTAGCCCTG CTGGCTCTCCAACCTCCACCGAAGAAGGTACCTCT GAAAGCGCAACCCCTGAATCCGGCCCAGGTAGCG AACCGGCAACCTCCGGTTCTGAAACCCCAGGTACT TCTGAAAGCGCTACTCCTGAGTCCGGCCCAGGTAG CCCGGCTGGCTCTCCGACTTCCACCGAGGAAGGTA GCCCGGCTGGCTCTCCAACTTCTACTGAAGAAGGT ACTTCTACCGAACCTTCCGAGGGCAGCGCACCAG GTACTTCTGAAAGCGCTACCCCTGAGTCCGGCCCA GGTACTTCTGAAAGCGCTACTCCTGAATCCGGTCC AGGTACTTCTGAAAGCGCTACCCCGGAATCTGGCC CAGGTAGCGAACCGGCTACTTCTGGTTCTGAAACC CCAGGTAGCGAACCGGCTACCTCCGGTTCTGAAAC TCCAGGTAGCCCAGCAGGCTCTCCGACTTCCACTG AGGAAGGTACTTCTACTGAACCTTCCGAAGGCAGC GCACCAGGTACCTCTACTGAACCTTCTGAGGGCAG CGCTCCAGGTAGCGAACCTGCAACCTCTGGCTCTG AAACCCCAGGTACCTCTGAAAGCGCTACTCCTGAA TCTGGCCCAGGTACTTCTACTGAACCGTCCGAGGG CAGCGCACCA		SPTSTEEGTSESA TPESGPGSEPATS GSETPGTSESATP ESGPGSPAGSPTS TEEGSPAGSPTST EEGTSTEPSEGSA PGTSESATPESGP GTSESATPESGPG TSESATPESGPGS EPATSGSETPGSE PATSGSETPGSPA GSPTSTEEGTSTE PSEGSAPGTSTEP SEGSAPGSEPATS GSETPGTSESATP ESGPGTSTEPSEG SAP	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
AE912-anti-CD40, AC386	ATGGCTGAACCTGCTGGCTCTCCAACCTCCACTGA GGAAGGTACCCCGGGTAGCGGTACTGCTTCTTCCT CTCCAGGTAGCTCTACCCCTTCTGGTGCAACCGGC TCTCCAGGTGCTTCTCCGGGCACCAGCTCTACCGG TTCTCCAGGTAGCCCGGCTGGCTCTCCTACCTCTA CTGAGGAAGGTACTTCTGAAAGCGCTACTCCTGAG TCTGGTCCAGGTACCTCTACTGAACCGTCCGAAGG TAGCGCTCCAGGTAGCCCAGCAGGCTCTCCGACTT CCACTGAGGAAGGTACTTCTACTGAACCTTCCGAA GGCAGCGCACCAGGTACCTCTACTGAACCTTCTGA GGGCAGCGCTCCAGGTACTTCTGAAAGCGCTACCC CGGAATCTGGCCCAGGTAGCGAACCGGCTACTTCT GGTTCTGAAACCCCAGGTAGCGAACCGGCTACCTC CGGTTCTGAAACTCCAGGTAGCCCGGCAGGCTCTC CGACCTCTACTGAGGAAGGTACTTCTGAAAGCGCA ACCCCGGAGTCCGGCCCAGGTACCTCTACCGAACC GTCTGAGGGCAGCGCACCAGGTACTTCTACCGAA CCGTCCGAGGGTAGCGCACCAGGTAGCCCAGCAG GTTCTCCTACCTCCACCGAGGAAGGTACTTCTACC GAACCGTCCGAGGGTAGCGCACCAGGTACCTCTA CTGAACCTTCTGAGGGCAGCGCTCCAGGTACTTCT GAAAGCGCTACCCCGGAGTCCGGTCCAGGTACTTC TACTGAACCGTCCGAAGGTAGCGCACCAGGTACTT CTGAAAGCGCAACCCCTGAATCCGGTCCAGGTAG CGAACCGGCTACTTCTGGCTCTGAGACTCCAGGTA CTTCTACCGAACCGTCCGAAGGTAGCGCACCAGGT ACTTCTACTGAACCGTCTGAAGGTAGCGCACCAGG TACTTCTGAAAGCGCAACCCCGGAATCCGGCCCA GGTACCTCTGAAAGCGCAACCCCGGAGTCCGGCC CAGGTAGCCCTGCTGGCTCTCCAACCTCCACCGAA GAAGGTACCTCTGAAAGCGCAACCCCTGAATCCG GCCCAGGTAGCGAACCGGCAACCTCCGGTTCTGA	708	MAEPAGSPTSTEE GTPGSGTASSSPG SSTPSGATGSPGA SPGTSSTGSPGSP AGSPTSTEEGTSE SATPESGPGTSTE PSEGSAPGSPAGS PTSTEEGTSTEPS EGSAPGTSTEPSE GSAPGTSESATPE SGPGSEPATSGSE TPGSEPATSGSET PGSPAGSPTSTEE GTSESATPESGPG TSTEPSEGSAPGT STEPSEGSAPGSP AGSPTSTEEGTST EPSEGSAPGTSTE PSEGSAPGTSESA TPESGPGTSTEPS EGSAPGTSESATP ESGPGSEPATSGS ETPGTSTEPSEGS APGTSTEPSEGSA PGTSESATPESGP GTSESATPESGPG SPAGSPTSTEEGT SESATPESGPGSE PATSGSETPGTSE SATPESGPGTSTE PSEGSAPGTSTEP SEGSAPGTSTEPS	733

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	AACCCCAGGTACCTCTGAAAGCGCTACTCCGGAGT CTGGCCCAGGTACCTCTACTGAACCGTCTGAGGGT AGCGCTCCAGGTACTTCTACTGAACCGTCCGAAGG TAGCGCACCAGGTACTTCTACCGAACCGTCCGAAG GCAGCGCTCCAGGTACCTCTACTGAACCTTCCGAG GGCAGCGCTCCAGGTACCTCTACCGAACCTTCTGA AGGTAGCGCACCAGGTACTTCTACCGAACCGTCCG AGGGTAGCGCACCAGGTAGCCCAGCAGGTTCTCC TACCTCCACCGAGGAAGGTACTTCTACCGAACCGT CCGAGGGTAGCGCACCAGGTACCTCTGAAAGCGC AACTCCTGAGTCTGGCCCAGGTAGCGAACCTGCTA CCTCCGGCTCTGAGACTCCAGGTACCTCTGAAAGC GCAACCCCGGAATCTGGTCCAGGTAGCGAACCTG CAACCTCTGGCTCTGAAACCCCAGGTACCTCTGAA AGCGCTACTCCTGAATCTGGCCCAGGTACTTCTAC TGAACCGTCCGAGGGCAGCGCACCAGGTACTTCT GAAAGCGCTACTCCTGAGTCCGGCCCAGGTAGCC CGGCTGGCTCTCCGACTTCCACCGAGGAAGGTAGC CCGGCTGGCTCTCCAACTTCTACTGAAGAAGGTAG CCCGGCAGGCTCTCCGACCTCTACTGAGGAAGGTA CTTCTGAAAGCGCAACCCCGGAGTCCGGCCCAGG TACCTCTACCGAACCGTCTGAGGGCAGCGCACCA GGTACCTCTGAAAGCGCAACTCCTGAGTCTGGCCC AGGTAGCGAACCTGCTACCTCCGGCTCTGAGACTC CAGGTACCTCTGAAAGCGCAACCCCGGAATCTGG TCCAGGTAGCGAACCTGCAACCTCTGGCTCTGAAA CCCCAGGTACCTCTGAAAGCGCTACTCCTGAATCT GGCCCAGGTACTTCTACTGAACCGTCCGAGGGCA GCGCACCAGGTAGCCCTGCTGGCTCTCCAACCTCC ACCGAAGAAGGTACCTCTGAAAGCGCAACCCCTG AATCCGGCCCAGGTAGCGAACCGGCAACCTCCGG TTCTGAAACCCCAGGTACTTCTGAAAGCGCTACTC CTGAGTCCGGCCCAGGTAGCCCGGCTGGCTCTCCG ACTTCCACCGAGGAAGGTAGCCCGGCTGGCTCTCC AACTTCTACTGAAGAAGGTACTTCTACCGAACCTT CCGAGGGCAGCGCACCAGGTACTTCTGAAAGCGC TACCCCTGAGTCCGGCCCAGGTACTTCTGAAAGCG CTACTCCTGAATCCGGTCCAGGTACTTCTGAAAGC GCTACCCCGGAATCTGGCCCAGGTAGCGAACCGG CTACTTCTGGTTCTGAAACCCCAGGTAGCGAACCG GCTACCTCCGGTTCTGAAACTCCAGGTAGCCCAGC AGGCTCTCCGACTTCCACTGAGGAAGGTACTTCTA CTGAACCTTCCGAAGGCAGCGCACCAGGTACCTCT ACTGAACCTTCTGAGGGCAGCGCTCCAGGTAGCG AACCTGCAACCTCTGGCTCTGAAACCCCAGGTACC TCTGAAAGCGCTACTCCTGAATCTGGCCCAGGTAC TTCTACTGAACCGTCCGAGGGCAGCGCACCAGGT GAAATTGTTCTGACCCAATCTCCTGCAACTCTGTC TCTGTCTCCAGGTGAACGCGCCACCCTGTCTTGTC GTGCGTCCCAGTCTATCTCTGATTATCTGCATTGGT ATCAGCAGAAACCTGGCCAGGCTCCGCGCCTGCT GATCTATTACGCCAGCCACAGCATCTCTGGTATCC CGGCTCGCTTCTCCGGCTCCGGCAGCGGCACCGAC TTCACTCTGACTATTAGCTCCCTGGAACCGGAGGA TTTCGCAGTTTATTACTGTCAGCACGGTCACTCCTA CCCGTGGACCTTTGGTGGCGGCACCAAAGTTGAAA TCAAAACCGGTTCTGGCGAAGGCTCTGAAGGTGA AGGTGGTGGTGAAGGCTCTGAAGGTGAAGGATCT		EGSAPGTSTEPSE GSAPGTSTEPSEG SAPGTSTEPSEGS APGSPAGSPTSTE EGTSTEPSEGSAP GTSESATPESGPG SEPATSGSETPGT SESATPESGPGSE PATSGSETPGTSE SATPESGPGTSTE PSEGSAPGTSESA TPESGPGSPAGSP TSTEEGSPAGSPT STEEGSPAGSPTS TEEGTSESATPES GPGTSTEPSEGSA PGTSESATPESGP GSEPATSGSETPG TSESATPESGPGS EPATSGSETPGTS ESATPESGPGTST EPSEGSAPGSPAG SPTSTEEGTSESA TPESGPGSEPATS GSETPGTSESATP ESGPGSPAGSPTS TEEGSPAGSPTST EEGTSTEPSEGSA PGTSESATPESGP GTSESATPESGPG TSESATPESGPGS EPATSGSETPGSE PATSGSETPGSPA GSPTSTEEGTSTE PSEGSAPGTSTEP SEGSAPGSEPATS GSETPGTSESATP ESGPGTSTEPSEG SAPGEIVLTQSPA TLSLSPGERATLS CRASQSISDYLH WYQQKPGQAPR LLIYYASHSISGIP ARFSGSGSGTDFT LTISSLEPEDFAV YYCQHGHSYPW TFGGGTKVEIKT GSGEGSEGEGGG EGSEGEGSGEGG EGEGSGTQVQLV QSGSELKKPGAS VKVSCKASGYAF TTTGMQWVRQA PGQGLEWMGWI NTHSGVPKYVED FKGRFVFSLDTSV STAYLQISSLKAE DTAVYYCARSGN	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	GGTGAAGGTGGCGAAGGTGAGGGATCTGGTACCC AAGTCCAGCTGGTTCAGTCCGGCTCTGAACTGAAG AAACCGGGCGCTTCTGTTAAAGTTAGCTGCAAAGC AAGCGGTTATGCCTTTACTACTACTGGTATGCAGT GGGTCCGCCAGGCACCGGGTCAGGGCCTGGAGTG GATGGGCTGGATCAACACCCACTCTGGTGTCCCTA AATACGTTGAAGATTTCAAAGGCCGTTTCGTGTTC TCCCTGGACACTTCCGTCAGCACCGCGTATCTGCA GATCAGCAGCCTGAAAGCTGAGGACACCGCGGTT TATTACTGCGCGCGTAGCGGCAATGGTAACTACGA CCTGGCTTATTTCAAATACTGGGGTCAGGGCACTC TGGTTACTGTGTCTAGC		GNYDLAYFKYW GQGTLVTVS	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
AE912-anti-CD40, AC387,	ATGGCTGAACCTGCTGGCTCTCCAACCTCCACTGA GGAAGGTACCCCGGGTAGCGGTACTGCTTCTTCCT CTCCAGGTAGCTCTACCCCTTCTGGTGCAACCGGC TCTCCAGGTGCTTCTCCGGGCACCAGCTCTACCGG TTCTCCAGGTAGCCCGGCTGGCTCTCCTACCTCTA CTGAGGAAGGTACTTCTGAAAGCGCTACTCCTGAG TCTGGTCCAGGTACCTCTACTGAACCGTCCGAAGG TAGCGCTCCAGGTAGCCCAGCAGGCTCTCCGACTT CCACTGAGGAAGGTACTTCTACTGAACCTTCCGAA GGCAGCGCACCAGGTACCTCTACTGAACCTTCTGA GGGCAGCGCTCCAGGTACTTCTGAAAGCGCTACCC CGGAATCTGGCCCAGGTAGCGAACCGGCTACTTCT GGTTCTGAAACCCCAGGTAGCGAACCGGCTACCTC CGGTTCTGAAACTCCAGGTAGCCCGGCAGGCTCTC CGACCTCTACTGAGGAAGGTACTTCTGAAAGCGCA ACCCCGGAGTCCGGCCCAGGTACCTCTACCGAACC GTCTGAGGGCAGCGCACCAGGTACTTCTACCGAA CCGTCCGAGGGTAGCGCACCAGGTAGCCCAGCAG GTTCTCCTACCTCCACCGAGGAAGGTACTTCTACC GAACCGTCCGAGGGTAGCGCACCAGGTACCTCTA CTGAACCTTCTGAGGGCAGCGCTCCAGGTACTTCT GAAAGCGCTACCCCGGAGTCCGGTCCAGGTACTTC TACTGAACCGTCCGAAGGTAGCGCACCAGGTACTT CTGAAAGCGCAACCCCTGAATCCGGTCCAGGTAG CGAACCGGCTACTTCTGGCTCTGAGACTCCAGGTA CTTCTACCGAACCGTCCGAAGGTAGCGCACCAGGT ACTTCTACTGAACCGTCTGAAGGTAGCGCACCAGG TACTTCTGAAAGCGCAACCCCGGAATCCGGCCCA GGTACCTCTGAAAGCGCAACCCCGGAGTCCGGCC CAGGTAGCCCTGCTGGCTCTCCAACCTCCACCGAA GAAGGTACCTCTGAAAGCGCAACCCCTGAATCCG GCCCAGGTAGCGAACCGGCAACCTCCGGTTCTGA AACCCCAGGTACCTCTGAAAGCGCTACTCCGGAGT CTGGCCCAGGTACCTCTACTGAACCGTCTGAGGGT AGCGCTCCAGGTACTTCTACTGAACCGTCCGAAGG TAGCGCACCAGGTACTTCTACCGAACCGTCCGAAG GCAGCGCTCCAGGTACCTCTACTGAACCTTCCGAG GGCAGCGCTCCAGGTACCTCTACCGAACCTTCTGA AGGTAGCGCACCAGGTACTTCTACCGAACCGTCCG AGGGTAGCGCACCAGGTAGCCCAGCAGGTTCTCC TACCTCCACCGAGGAAGGTACTTCTACCGAACCGT CCGAGGGTAGCGCACCAGGTACCTCTGAAAGCGC AACTCCTGAGTCTGGCCCAGGTAGCGAACCTGCTA CCTCCGGCTCTGAGACTCCAGGTACCTCTGAAAGC GCAACCCCGGAATCTGGTCCAGGTAGCGAACCTG CAACCTCTGGCTCTGAAACCCCAGGTACCTCTGAA	709	MAEPAGSPTSTEE GTPGSGTASSSPG SSTPSGATGSPGA SPGTSSTGSPGSP AGSPTSTEEGTSE SATPESGPGTSTE PSEGSAPGSPAGS PTSTEEGTSTEPS EGSAPGTSTEPSE GSAPGTSESATPE SGPGSEPATSGSE TPGSEPATSGSET PGSPAGSPTSTEE GTSESATPESGPG TSTEPSEGSAPGT STEPSEGSAPGSP AGSPTSTEEGTST EPSEGSAPGTSTE PSEGSAPGTSESA TPESGPGTSTEPS EGSAPGTSESATP ESGPGSEPATSGS ETPGTSTEPSEGS APGTSTEPSEGSA PGTSESATPESGP GTSESATPESGPG SPAGSPTSTEEGT SESATPESGPGSE PATSGSETPGTSE SATPESGPGTSTE PSEGSAPGTSTEP SEGSAPGTSTEPS EGSAPGTSTEPSE GSAPGTSTEPSEG SAPGTSTEPSEGS APGSPAGSPTSTE EGTSTEPSEGSAP GTSESATPESGPG SEPATSGSETPGT SESATPESGPGSE PATSGSETPGTSE SATPESGPGTSTE PSEGSAPGTSESA TPESGPGSPAGSP TSTEEGSPAGSPT STEEGSPAGSPTS	734

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	AGCGCTACTCCTGAATCTGGCCCAGGTACTTCTAC TGAACCGTCCGAGGGCAGCGCACCAGGTACTTCT GAAAGCGCTACTCCTGAGTCCGGCCCAGGTAGCC CGGCTGGCTCTCCGACTTCCACCGAGGAAGGTAGC CCGGCTGGCTCTCCAACTTCTACTGAAGAAGGTAG CCCGGCAGGCTCTCCGACCTCTACTGAGGAAGGTA CTTCTGAAAGCGCAACCCCGGAGTCCGGCCCAGG TACCTCTACCGAACCGTCTGAGGGCAGCGCACCA GGTACCTCTGAAAGCGCAACTCCTGAGTCTGGCCC AGGTAGCGAACCTGCTACCTCCGGCTCTGAGACTC CAGGTACCTCTGAAAGCGCAACCCCGGAATCTGG TCCAGGTAGCGAACCTGCAACCTCTGGCTCTGAAA CCCCAGGTACCTCTGAAAGCGCTACTCCTGAATCT GGCCCAGGTACTTCTACTGAACCGTCCGAGGGCA GCGCACCAGGTAGCCCTGCTGGCTCTCCAACCTCC ACCGAAGAAGGTACCTCTGAAAGCGCAACCCCTG AATCCGGCCCAGGTAGCGAACCGGCAACCTCCGG TTCTGAAACCCCAGGTACTTCTGAAAGCGCTACTC CTGAGTCCGGCCCAGGTAGCCCGGCTGGCTCTCCG ACTTCCACCGAGGAAGGTAGCCCGGCTGGCTCTCC AACTTCTACTGAAGAAGGTACTTCTACCGAACCTT CCGAGGGCAGCGCACCAGGTACTTCTGAAAGCGC TACCCCTGAGTCCGGCCCAGGTACTTCTGAAAGCG CTACTCCTGAATCCGGTCCAGGTACTTCTGAAAGC GCTACCCCCGGAATCTGGCCCAGGTAGCGAACCGG CTACTTCTGGTTCTGAAACCCCAGGTAGCGAACCG GCTACCTCCCGGTTCTGAAACTCCAGGTAGCCCAGC AGGCTCTCCGACTTCCACTGAGGAAGGTACTTCTA CTGAACCTTCCGAAGGCAGCGCACCAGGTACCTCT ACTGAACCTTCTGAGGGCAGCGCTCCAGGTAGCG AACCTGCAACCTCTGGCTCTGAAACCCCAGGTACC TCTGAAAGCGCTACTCCTGAATCTGGCCCAGGTAC TTCTACTGAACCGTCCGAGGGCAGCGCACCAGGT GAAATTGTTCTGACTCAATCTCCAGCAACTCTGTC TCTGTCTCCAGGTGAACGTGCAACCCTGTCTTGCC GTGCGTCCCAGTCCATCTCCGATTATCTGCATTGG TATCAGCAGAAACCGGGTCAGGCGCCTCGTCTGCT GATCTATTATGCGTCTCACTCCATTTCCGGTATCCC GGCACGTTTCTCTGGCAGCGGCAGCGGCACCGATT TCACCCTGACGATCTCTTCTCTGGAACCGGAAGAT TTCGCAGTCTATTATTGTCAGCATGGTCACAGCTA CCCGTGGACCTTCGGCGGTGGCACGAAAGTTGAA ATCAAGACCGGTTCTGGCGAAGGCTCTGAAGGTG AAGGTGGTGGTGAAGGCTCTGAAGGTGAAGGATC TGGTGAAGGTGGCGAAGGTGAGGGATCTGGTACC CAGGTCCAGCTGGTTCAAAGCGGCTCTGAACTGAA AAAGCCGGGTGCCTCTGTCAAAGTGTCTTGCAAGG CAAGCGGCTACGCGTTTACGACCACCGGCATGCA GTGGGTCCGTCAGGCCCCGGGCCAGGGTCTGGAA TGGATGGGCTGGATCAACACCCATTCTGGCGTACC GAAATACGTTGAAGATTTCAAAGGCCGTTTCGTGT TCTCCCTGGATACGTCCGTTTCCACCGCCTACCTG CAGATCTCTTCCCTGAAAGCAGAAGATACTGCGGT GTACTATTGCGCACGTAGCGGCAACGGCAACTAC GACCTGGCCTACTTCAAATACTGGGGTCAGGGTAC TCTGGTGACCGTATCCTCT		TEEGTSESATPES GPGTSTEPSEGSA PGTSESATPESGP GSEPATSGSETPG TSESATPESGPGS EPATSGSETPGTS ESATPESGPGTST EPSEGSAPGSPAG SPTSTEEGTSESA TPESGPGSEPATS GSETPGTSESATP ESGPGSPAGSPTS TEEGSPAGSPTST EEGTSTEPSEGSA PGTSESATPESGP GTSESATPESGPG TSESATPESGPGS EPATSGSETPGSE PATSGSETPGSPA GSPTSTEEGTSTE PSEGSAPGTSTEP SEGSAPGSEPATS GSETPGTSESATP ESGPGTSTEPSEG SAPGEIVLTQSPA TLSLSPGERATLS CRASQSISDYLH WYQQKPGQAPR LLIYYASHSISGIP ARFSGSGSGTDFT LTISSLEPEDFAV YYCQHGHSYPW TFGGGTKVEIKT GSGEGSEGEGGG EGSEGEGSGEGG EGEGSGTQVQLV QSGSELKKPGAS VKVSCKASGYAF TTTGMQWVRQA PGQGLEWMGWI NTHSGVPKYVED FKGRFVFSLDTSV STAYLQISSLKAE DTAVYYCARSGN GNYDLAYFKYW GQGTLVTVS	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
anti-Her2-AE864	ATGGAAGACATTCAGATGACCCAGAGCCCGTCCT CCCTGAGCGCTTCTGTTGGCGACCGCGTGACCATC	710	MEDIQMTQSPSSL SASVGDRVTITCR	735

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	ACCTGCCGTGCTTCCCAGGATGTTAACACCGCTGT AGCTTGGTATCAACAGAAACCGGGCAAAGCACCG AAACTGCTGATCTACTCTGCTTCCTTTCTGTATAGC GGTGTTCCGTCTCGTTTCAGCGGCTCTCGTAGCGG TACGGATTTTACTCTGACGATCAGCTCTCTGCAGC CGGAGGACTTCGCTACCTACTACTGCCAGCAGCAC TACACCACCCCGCCTACCTTTGGTCAGGGCACCAA AGTGGAAATCAAGACCGGTTCTGGCGAAGGCTCT GAAGGTGAAGGTGGTGGTGAAGGCTCTGAAGGTG AAGGATCTGGTGAAGGTGGCGAAGGTGAGGGATC TGGTACCGAGGTCCAGCTGGTTGAGTCTGGCGGCG GTCTGGTCCAACCTGGTGGCTCCCTGCGCCTGTCT TGCGCAGCGTCCGGCTTTAATATCAAAGATACGTA CATTCACTGGGTCCGCCAGGCACCGGGCAAAGGC CTGGAATGGGTTGCTCGTATCTACCCGACTAACGG TTATACCCGTTATGCAGACAGCGTAAAGGGTCGCT TCACGATCTCCGCGGATACCTCCAAAAACACCGCA TACCTGCAAATGAACTCTCTGCGTGCGGAAGATAC TGCCGTGTACTACTGCTCTCGCTGGGGCGGTGACG GTTTCTATGCAATGGACTACTGGGGTCAAGGTACT CTGGTAACTGTTTCCGGGTCTCCAGGTAGCCCGGC TGGCTCTCCTACCTCTACTGAGGAAGGTACTTCTG AAAGCGCTACTCCTGAGTCTGGTCCAGGTACCTCT ACTGAACCGTCCGAAGGTAGCGCTCCAGGTAGCC CAGCAGGCTCTCCGACTTCCACTGAGGAAGGTACT TCTACTGAACCTTCCGAAGGCAGCGCACCAGGTAC CTCTACTGAACCTTCTGAGGGCAGCGCTCCAGGTA CTTCTGAAAGCGCTACCCCGGAATCTGGCCCAGGT AGCGAACCGGCTACTTCTGGTTCTGAAACCCCAGG TAGCGAACCGGCTACCTCCGGTTCTGAAACTCCAG GTAGCCCGGCAGGCTCTCCGACCTCTACTGAGGAA GGTACTTCTGAAAGCGCAACCCCGCAGTCCGGCC CAGGTACCTCTACCGAACCGTCTGAGGGCAGCGC ACCAGGTACTTCTACCGAACCGTCCGAGGGTAGC GCACCAGGTAGCCCAGCAGGTTCTCCTACCTCCAC CGAGGAAGGTACTTCTACCGAACCGTCCGAGGGT AGCGCACCAGGTACCTCTACTGAACCTTCTGAGGG CAGCGCTCCAGGTACTTCTGAAAGCGCTACCCCGG AGTCCGGTCCAGGTACTTCTACTGAACCGTCCGAA GGTAGCGCACCAGGTACTTCTGAAAGCGCAACCC CTGAATCCGGTCCAGGTAGCGAACCGGCTACTTCT GGCTCTGAGACTCCAGGTACTTCTACCGAACCGTC CGAAGGTAGCGCACCAGGTACTTCTACTGAACCGT CTGAAGGTAGCGCACCAGGTACTTCTGAAAGCGC AACCCCGGAATCCGGCCCAGGTACCTCTGAAAGC GCAACCCCGGAGTCCGGCCCAGGTAGCCCTGCTG GCTCTCCAACCTCCACCGAAGAAGGTACCTCTGAA AGCGCAACCCCTGAATCCGGCCCAGGTAGCGAAC CGGCAACCTCCGGTTCTGAAACCCCAGGTACCTCT GAAAGCGCTACTCCGGAGTCTGGCCCAGGTACCTC TACTGAACCGTCTGAGGGTAGCGCTCCAGGTACTT CTACTGAACCGTCCGAAGGTAGCGCACCAGGTAC TTCTACCGAACCGTCCGAAGGCAGCGCTCCAGGTA CCTCTACTGAACCTTCCGAGGGCAGCGCTCCAGGT ACCTCTACCGAACCTTCTGAAGGTAGCGCACCAGG TACTTCTACCGAACCGTCCGAGGGTAGCGCACCAG GTAGCCCAGCAGGTTCTCCTACCTCCACCGAGGAA GGTACTTCTACCGAACCGTCCGAGGGTAGCGCACC		ASQDVNTAVAW YQQKPGKAPKLL IYSASFLYSGVPS RFSGSRSGTDFTL TISSLQPEDFATY YCQQHYTTPPTF GQGTKVEIKTGS GEGSEGEGGGEG SEGEGSGEGGEG EGSGTEVQLVES GGGLVQPGGSLR LSCAASGFNIKDT YIHWVRQAPGKG LEWVARIYPTNG YTRYADSVKGRF TISADTSKNTAYL QMNSLRAEDTAV YYCSRWGGDGF YAMDYWGQGTL VTVSGSPGSPAGS PTSTEEGTSESAT PESGPGTSTEPSE GSAPGSPAGSPTS TEEGTSTEPSEGS APGTSTEPSEGSA PGTSESATPESGP GSEPATSGSETPG SEPATSGSETPGS PAGSPTSTEEGTS ESATPESGPGTST EPSEGSAPGTSTE PSEGSAPGSPAGS PTSTEEGTSTEPS EGSAPGTSTEPSE GSAPGTSESATPE SGPGTSTEPSEGS APGTSESATPESG PGSEPATSGSETP GTSTEPSEGSAPG TSTEPSEGSAPGT SESATPESGPGTS ESATPESGPGSPA GSPTSTEEGTSES ATPESGPGSEPAT SGSETPGTSESAT PESGPGTSTEPSE GSAPGTSTEPSEG SAPGTSTEPSEGS APGTSTEPSEGSA PGTSTEPSEGSAP GTSTEPSEGSAPG SPAGSPTSTEEGT STEPSEGSAPGTS ESATPESGPGSEP ATSGSETPGTSES ATPESGPGSEPAT SGSETPGTSESAT PESGPGTSTEPSE	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	AGGTACCTCTGAAAGCGCAACTCCTGAGTCTGGCC CAGGTAGCGAACCTGCTACCTCCGGCTCTGAGACT CCAGGTACCTCTGAAAGCGCAACCCCGGAATCTG GTCCAGGTAGCGAACCTGCAACCTCTGGCTCTGAA ACCCCAGGTACCTCTGAAAGCGCTACTCCTGAATC TGGCCCAGGTACTTCTACTGAACCGTCCGAGGGCA GCGCACCAGGTACTTCTGAAAGCGCTACTCCTGAG TCCGGCCCAGGTAGCCCGGCTGGCTCTCCGACTTC CACCGAGGAAGGTAGCCCGGCTGGCTCTCCAACTT CTACTGAAGAAGGTAGCCCGGCAGGCTCTCCGAC CTCTACTGAGGAAGGTACTTCTGAAAGCGCAACCC CGGAGTCCGGCCCAGGTACCTCTACCGAACCGTCT GAGGGCAGCGCACCAGGTACCTCTGAAAGCGCAA CTCCTGAGTCTGGCCCAGGTAGCGAACCTGCTACC TCCGGCTCTGAGACTCCAGGTACCTCTGAAAGCGC AACCCCCGGAATCTGGTCCAGGTAGCGAACCTGCA ACCTCTGGCTCTGAAACCCCAGGTACCTCTGAAAG CGCTACTCCTGAATCTGGCCCAGGTACTTCTACTG AACCGTCCGAGGGCAGCGCACCAGGTAGCCCTGC TGGCTCTCCAACCTCCACCGAAGAAGGTACCTCTG AAAGCGCAACCCCTGAATCCGGCCCAGGTAGCGA ACCGGCAACCTCCGGTTCTGAAACCCCAGGTACTT CTGAAAGCGCTACTCCTGAGTCCGGCCCAGGTAGC CCGGCTGGCTCTCCGACTTCCACCGAGGAAGGTAG CCCGGCTGGCTCTCCAACTTCTACTGAAGAAGGTA CTTCTACCGAACCTTCCGAGGGCAGCGCACCAGGT ACTTCTGAAAGCGCTACCCCTGAGTCCGGCCCAGG TACTTCTGAAAGCGCTACTCCTGAATCCGGTCCAG GTACTTCTGAAAGCGCTACCCCGGAATCTGGCCCA GGTAGCGAACCGGCTACTTCTGGTTCTGAAACCCC AGGTAGCGAACCGGCTACCTCCGGTTCTGAAACTC CAGGTAGCCCAGCAGGCTCTCCGACTTCCACTGAG GAAGGTACTTCTACTGAACCTTCCGAAGGCAGCGC ACCAGGTACCTCTACTGAACCTTCTGAGGGCAGCG CTCCAGGTAGCGAACCTGCAACCTCTGGCTCTGAA ACCCCAGGTACCTCTGAAAGCGCTACTCCTGAATC TGGCCCAGGTACTTCTACTGAACCGTCCGAGGGCA GCGCACCA		GSAPGTSESATPE SGPGSPAGSPTST EEGSPAGSPTSTE EGSPAGSPTSTEE GTSESATPESGPG TSTEPSEGSAPGT SESATPESGPGSE PATSGSETPGTSE SATPESGPGSEPA TSGSETPGTSESA TPESGPGTSTEPS EGSAPGSPAGSPT STEEGTSESATPE SGPGSEPATSGSE TPGTSESATPESG PGSPAGSPTSTEE GSPAGSPTSTEEG TSTEPSEGSAPGT SESATPESGPGTS ESATPESGPGTSE SATPESGPGSEPA TSGSETPGSEPAT SGSETPGSPAGSP TSTEEGTSTEPSE GSAPGTSTEPSEG SAPGSEPATSGSE TPGTSESATPESG PGTSTEPSEGSAP	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
anti-Her2-AE576	ATGGAAGACATTCAGATGACCCAGAGCCCGTCCT CCCTGAGCGCTTCTGTTGGCGACCGCGTGACCATC ACCTGCCGTGCTTCCCAGGATGTTAACACCGCTGT AGCTTGGTATCAACAGAAACCGGGCAAAGCACCG AAACTGCTGATCTACTCTGCTTCCTTTCTGTATAGC GGTGTTCCGTCTCGTTTCAGCGGCTCTCGTAGCGG TACGGATTTTACTCTGACGATCAGCTCTCTGCAGC CGGAGGACTTCGCTACCTACTACTGCCAGCAGCAC TACACCACCCCGCCTACCTTTGGTCAGGGCACCAA AGTGGAAATCAAGACCGGTTCTGGCGAAGGCTCT GAAGGTGAAGGTGGTGGTGAAGGCTCTGAAGGTG AAGGATCTGGTGAAGGTGGCGAAGGTGAGGGATC TGGTACCGAGGTCCAGCTGGTTGAGTCTGGCGGCG GTCTGGTCCAACCTGGTGGCTCCCTGCGCCTGTCT TGCGCAGCGTCCGGCTTTAATATCAAAGATACGTA CATTCACTGGGTCCGCCAGGCACCGGGCAAAGGC CTGGAATGGGTTGCTCGTATCTACCCGACTAACGG TTATACCCGTTATGCAGACAGCGTAAAGGGTCGCT TCACGATCTCCGCGGATACCTCCAAAAACACCGCA TACCTGCAAATGAACTCTCTGCGTGCGGAAGATAC	711	MEDIQMTQSPSSL SASVGDRVTITCR ASQDVNTAVAW YQQKPGKAPKLL IYSASFLYSGVPS RFSGSRSGTDFTL TISSLQPEDFATY YCQQHYTTPPTF GQGTKVEIKTGS GEGSEGEGGGEG SEGEGSGEGGEG EGSGTEVQLVES GGGLVQPGGSLR LSCAASGFNIKDT YIHWVRQAPGKG LEWVARIYPTNG YTRYADSVKGRF TISADTSKNTAYL QMNSLRAEDTAV YYCSRWGGDGF	736

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	TGCCGTGTACTACTGCTCTCGCTGGGGCGGTGACG GTTTCTATGCAATGGACTACTGGGGTCAAGGTACT CTGGTAACTGTTTCCGGGTCTCCAGGTAGCCCGGC TGGCTCTCCTACCTCTACTGAGGAAGGTACTTCTG AAAGCGCTACTCCTGAGTCTGGTCCAGGTACCTCT ACTGAACCGTCCGAAGGTAGCGCTCCAGGTAGCC CAGCAGGCTCTCCGACTTCCACTGAGGAAGGTACT TCTACTGAACCTTCCGAAGGCAGCGCACCAGGTAC CTCTACTGAACCTTCTGAGGGCAGCGCTCCAGGTA CTTCTGAAAGCGCTACCCCGGAATCTGGCCCAGGT AGCGAACCGGCTACTTCTGGTTCTGAAACCCCAGG TAGCGAACCGGCTACCTCCGGTTCTGAAACTCCAG GTAGCCCGGCAGGCTCTCCGACCTCTACTGAGGAA GGTACTTCTGAAAGCGCAACCCCGGAGTCCGGCC CAGGTACCTCTACCGAACCGTCTGAGGGCAGCGC ACCAGGTACTTCTACCGAACCGTCCGAGGGTAGC GCACCAGGTAGCCCAGCAGGTTCTCCTACCTCCAC CGAGGAAGGTACTTCTACCGAACCGTCCGAGGGT AGCGCACCAGGTACCTCTACTGAACCTTCTGAGGG CAGCGCTCCAGGTACTTCTGAAAGCGCTACCCCGG AGTCCGGTCCAGGTACTTCTACTGAACCGTCCGAA GGTAGCGCACCAGGTACTTCTGAAAGCGCAACCC CTGAATCCGGTCCAGGTAGCGAACCGGCTACTTCT GGCTCTGAGACTCCAGGTACTTCTACCGAACCGTC CGAAGGTAGCGCACCAGGTACTTCTACTGAACCGT CTGAAGGTAGCGCACCAGGTACTTCTGAAAGCGC AACCCCCGGAATCCGGCCCAGGTACCTCTGAAAGC GCAACCCCGGAGTCCGGCCCAGGTAGCCCTGCTG GCTCTCCAACCTCCACCGAAGAAGGTACCTCTGAA AGCGCAACCCCTGAATCCGGCCCAGGTAGCGAAC CGGCAACCTCCGGTTCTGAAACCCCAGGTACCTCT GAAAGCGCTACTCCGGAGTCTGGCCCAGGTACCTC TACTGAACCGTCTGAGGGTAGCGCTCCAGGTACTT CTACTGAACCGTCCGAAGGTAGCGCACCAGGTAC TTCTACCGAACCGTCCGAAGGCAGCGCTCCAGGTA CCTCTACTGAACCTTCCGAGGGCAGCGCTCCAGGT ACCTCTACCGAACCTTCTGAAGGTAGCGCACCAGG TACTTCTACCGAACCGTCCGAGGGTAGCGCACCAG GTAGCCCAGCAGGTTCTCCTACCTCCACCGAGGAA GGTACTTCTACCGAACCGTCCGAGGGTAGCGCACC AGGTACCTCTGAAAGCGCAACTCCTGAGTCTGGCC CAGGTAGCGAACCTGCTACCTCCGGCTCTGAGACT CCAGGTACCTCTGAAAGCGCAACCCCGGAATCTG GTCCAGGTAGCGAACCTGCAACCTCTGGCTCTGAA ACCCCAGGTACCTCTGAAAGCGCTACTCCTGAATC TGGCCCAGGTACTTCTACTGAACCGTCCGAGGGCA GCGCACCAGGTACTTCTGAAAGCGCTACTCCTGAG TCCGGCCCAGGTAGCCCGGCTGGCTCTCCGACTTC CACCGAGGAAGGTAGCCCGGCTGGCTCTCCAACTT CTACTGAAGAAGGTAGCCCGGCAGGCTCTCCGAC CTCTACTGAGGAAGGTACTTCTGAAAGCGCAACCC CGGAGTCCGGCCCAGGTACCTCTACCGAACCGTCT GAGGGCAGCGCACCA		YAMDYWGQGTL VTVSGSPGSPAGS PTSTEEGTSESAT PESGPGTSTEPSE GSAPGSPAGSPTS TEEGTSTEPSEGS APGTSTEPSEGSA PGTSESATPESGP GSEPATSGSETPG SEPATSGSETPGS PAGSPTSTEEGTS ESATPESGPGTST EPSEGSAPGTSTE PSEGSAPGSPAGS PTSTEEGTSTEPS EGSAPGTSTEPSE GSAPGTSESATPE SGPGTSTEPSEGS APGTSESATPESG PGSEPATSGSETP GTSTEPSEGSAPG TSTEPSEGSAPGT SESATPESGPGTS ESATPESGPGSPA GSPTSTEEGTSES ATPESGPGSEPAT SGSETPGTSESAT PESGPGTSTEPSE GSAPGTSTEPSEG SAPGTSTEPSEGS APGTSTEPSEGSA PGTSTEPSEGSAP GTSTEPSEGSAPG SPAGSPTSTEEGT STEPSEGSAPGTS ESATPESGPGSEP ATSGSETPGTSES ATPESGPGSEPAT SGSETPGTSESAT PESGPGTSTEPSE GSAPGTSESATPE SGPGSPAGSPTST EEGSPAGSPTSTE EGSPAGSPTSTEE GTSESATPESGPG TSTEPSEGSAP	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
anti-Her2-AE288	ATGGAAGACATTCAGATGACCCAGAGCCCGTCCT CCCTGAGCGCTTCTGTTGGCGACCGCGTGACCATC ACCTGCCGTGCTTCCCAGGATGTTAACACCGCTGT AGCTTGGTATCAACAGAAACCGGGCAAAGCACCG AAACTGCTGATCTACTCTGCTTCCTTTCTGTATAGC	712	MEDIQMTQSPSSL SASVGDRVTITCR ASQDVNTAVAW YQQKPGKAPKLL IYSASFLYSGVPS	737
	GGTGTTCCGTCTCGTTTCAGCGGCTCTCGTAGCGG TACGGATTTTACTCTGACGATCAGCTCTCTGCAGC CGGAGGACTTCGCTACCTACTACTGCCAGCAGCAC TACACCACCCCGCCTACCTTTGGTCAGGGCACCAA AGTGGAAATCAAGACCGGTTCTGGCGAAGGCTCT GAAGGTGAAGGTGGTGGTGAAGGCTCTGAAGGTG AAGGATCTGGTGAAGGTGGCGAAGGTGAGGGATC TGGTACCGAGGTCCAGCTGGTTGAGTCTGGCGGCG GTCTGGTCCAACCTGGTGGCTCCCTGCGCCTGTCT TGCGCAGCGTCCGGCTTTAATATCAAAGATACGTA CATTCACTGGGTCCGCCAGGCACCGGGCAAAGGC CTGGAATGGGTTGCTCGTATCTACCCGACTAACGG TTATACCCGTTATGCAGACAGCGTAAAGGGTCGCT TCACGATCTCCGCGGATACCTCCAAAAACACCGCA TACCTGCAAATGAACTCTCTGCGTGCGGAAGATAC TGCCGTGTACTACTGCTCTCGCTGGGGCGGTGACG GTTTCTATGCAATGGACTACTGGGGTCAAGGTACT CTGGTAACTGTTTCCGGGTCTCCAGGTACCTCTGA AAGCGCAACTCCTGAGTCTGGCCCAGGTAGCGAA CCTGCTACCTCCGGCTCTGAGACTCCAGGTACCTC TGAAAGCGCAACCCCGGAATCTGGTCCAGGTAGC GAACCTGCAACCTCTGGCTCTGAAACCCCAGGTAC CTCTGAAAGCGCTACTCCTGAATCTGGCCCAGGTA CTTCTACTGAACCGTCCGAGGGCAGCGCACCAGGT AGCCCTGCTGGCTCTCCAACCTCCACCGAAGAAGG TACCTCTGAAAGCGCAACCCCTGAATCCGGCCCAG GTAGCGAACCGGCAACCTCCGGTTCTGAAACCCC AGGTACTTCTGAAAGCGCTACTCCTGAGTCCGGCC CAGGTAGCCCGGCTGGCTCTCCGACTTCCACCGAG GAAGGTAGCCCGGCTGGCTCTCCAACTTCTACTGA AGAAGGTACTTCTACCGAACCTTCCGAGGGCAGC GCACCAGGTACTTCTGAAAGCGCTACCCCTGAGTC CGGCCCAGGTACTTCTGAAAGCGCTACTCCTGAAT CCGGTCCAGGTACTTCTGAAAGCGCTACCCCGGAA TCTGGCCCAGGTAGCGAACCGGCTACTTCTGGTTC TGAAACCCCAGGTAGCGAACCGGCTACCTCCGGTT CTGAAACTCCAGGTAGCCCAGCAGGCTCTCCGACT TCCACTGAGGAAGGTACTTCTACTGAACCTTCCGA AGGCAGCGCACCAGGTACCTCTACTGAACCTTCTG AGGGCAGCGCTCCAGGTAGCGAACCTGCAACCTC TGGCTCTGAAACCCCAGGTACCTCTGAAAGCGCTA CTCCTGAATCTGGCCCAGGTACTTCTACTGAACCG TCCGAGGGCAGCGCACCA		RFSGSRSGTDFTL TISSLQPEDFATY YCQQHYTTPPTF GQGTKVEIKTGS GEGSEGEGGGEG SEGEGSGEGGEG EGSGTEVQLVES GGGLVQPGGSLR LSCAASGFNIKDT YIHWVRQAPGKG LEWVARIYPTNG YTRYADSVKGRF TISADTSKNTAYL QMNSLRAEDTAV YYCSRWGGDGF YAMDYWGQGTL VTVSGSPGTSESA TPESGPGSEPATS GSETPGTSESATP ESGPGSEPATSGS ETPGTSESATPES GPGTSTEPSEGSA PGSPAGSPTSTEE GTSESATPESGPG SEPATSGSETPGT SESATPESGPGSP AGSPTSTEEGSPA GSPTSTEEGTSTE PSEGSAPGTSESA TPESGPGTSESAT PESGPGTSESATP ESGPGSEPATSGS ETPGSEPATSGSE TPGSPAGSPTSTE EGTSTEPSEGSAP GTSTEPSEGSAPG SEPATSGSETPGT SESATPESGPGTS TEPSEGSAP	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
AE912-anti-Her2	ATGGCTGAACCTGCTGGCTCTCCAACCTCCACTGA GGAAGGTACCCCGGGTAGCGGTACTGCTTCTTCCT CTCCAGGTAGCTCTACCCCTTCTGGTGCAACCGGC TCTCCAGGTGCTTCTCCGGGCACCAGCTCTACCGG TTCTCCAGGTAGCCCGGCTGGCTCTCCTACCTCTA CTGAGGAAGGTACTTCTGAAAGCGCTACTCCTGAG TCTGGTCCAGGTACCTCTACTGAACCGTCCGAAGG TAGCGCTCCAGGTAGCCCAGCAGGCTCTCCGACTT CCACTGAGGAAGGTACTTCTACTGAACCTTCCGAA GGCAGCGCACCAGGTACCTCTACTGAACCTTCTGA GGGCAGCGCTCCAGGTACTTCTGAAAGCGCTACCC CGGAATCTGGCCCAGGTAGCGAACCGGCTACTTCT GGTTCTGAAACCCCAGGTAGCGAACCGGCTACCTC CGGTTCTGAAACTCCAGGTAGCCCGGCAGGCTCTC CGACCTCTACTGAGGAAGGTACTTCTGAAAGCGCA	713	MAEPAGSPTSTEE GTPGSGTASSSPG SSTPSGATGSPGA SPGTSSTGSPGSP AGSPTSTEEGTSE SATPESGPGTSTE PSEGSAPGSPAGS PTSTEEGTSTEPS EGSAPGTSTEPSE GSAPGTSESATPE SGPGSEPATSGSE TPGSEPATSGSET PGSPAGSPTSTEE GTSESATPESGPG TSTEPSEGSAPGT	738

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	ACCCCGGAGTCCGGCCCAGGTACCTCTACCGAACC GTCTGAGGGCAGCGCACCAGGTACTTCTACCGAA CCGTCCGAGGGTAGCGCACCAGGTAGCCCAGCAG GTTCTCCTACCTCCACCGAGGAAGGTACTTCTACC GAACCGTCCGAGGGTAGCGCACCAGGTACCTCTA CTGAACCTTCTGAGGGCAGCGCTCCAGGTACTTCT GAAAGCGCTACCCCGGAGTCCGGTCCAGGTACTTC TACTGAACCGTCCGAAGGTAGCGCACCAGGTACTT CTGAAAGCGCAACCCCTGAATCCGGTCCAGGTAG CGAACCGGCTACTTCTGGCTCTGAGACTCCAGGTA CTTCTACCGAACCGTCCGAAGGTAGCGCACCAGGT ACTTCTACTGAACCGTCTGAAGGTAGCGCACCAGG TACTTCTGAAAGCGCAACCCCGGAATCCGGCCCA GGTACCTCTGAAAGCGCAACCCCGGAGTCCGGCC CAGGTAGCCCTGCTGGCTCTCCAACCTCCACCGAA GAAGGTACCTCTGAAAGCGCAACCCCTGAATCCG GCCCAGGTAGCGAACCGGCAACCTCCGGTTCTGA AACCCCAGGTACCTCTGAAAGCGCTACTCCGGAGT CTGGCCCAGGTACCTCTACTGAACCGTCTGAGGGT AGCGCTCCAGGTACTTCTACTGAACCGTCCGAAGG TAGCGCACCAGGTACTTCTACCGAACCGTCCGAAG GCAGCGCTCCAGGTACCTCTACTGAACCTTCCGAG GGCAGCGCTCCAGGTACCTCTACCGAACCTTCTGA AGGTAGCGCACCAGGTACTTCTACCGAACCGTCCG AGGGTAGCGCACCAGGTAGCCCAGCAGGTTCTCC TACCTCCACCGAGGAAGGTACTTCTACCGAACCGT CCGAGGGTAGCGCACCAGGTACCTCTGAAAGCGC AACTCCTGAGTCTGGCCCAGGTAGCGAACCTGCTA CCTCCGGCTCTGAGACTCCAGGTACCTCTGAAAGC GCAACCCCGGAATCTGGTCCAGGTAGCGAACCTG CAACCTCTGGCTCTGAAACCCCAGGTACCTCTGAA AGCGCTACTCCTGAATCTGGCCCAGGTACTTCTAC TGAACCGTCCGAGGGCAGCGCACCAGGTACTTCT GAAAGCGCTACTCCTGAGTCCGGCCCAGGTAGCC CGGCTGGCTCTCCGACTTCCACCGAGGAAGGTAGC CCGGCTGGCTCTCCAACTTCTACTGAAGAAGGTAG CCCGGCAGGCTCTCCGACCTCTACTGAGGAAGGTA CTTCTGAAAGCGCAACCCCGGAGTCCGGCCCAGG TACCTCTACCGAACCGTCTGAGGGCAGCGCACCA GGTACCTCTGAAAGCGCAACTCCTGAGTCTGGCCC AGGTAGCGAACCTGCTACCTCCGGCTCTGAGACTC CAGGTACCTCTGAAAGCGCAACCCCGGAATCTGG TCCAGGTAGCGAACCTGCAACCTCTGGCTCTGAAA CCCCAGGTACCTCTGAAAGCGCTACTCCTGAATCT GGCCCAGGTACTTCTACTGAACCGTCCGAGGGCA GCGCACCAGGTAGCCCTGCTGGCTCTCCAACCTCC ACCGAAGAAGGTACCTCTGAAAGCGCAACCCCTG AATCCGGCCCAGGTAGCGAACCGGCAACCTCCGG TTCTGAAACCCCAGGTACTTCTGAAAGCGCTACTC CTGAGTCCGGCCCAGGTAGCCCGGCTGGCTCTCCG ACTTCCACCGAGGAAGGTAGCCCGGCTGGCTCTCC AACTTCTACTGAAGAAGGTACTTCTACCGAACCTT CCGAGGGCAGCGCACCAGGTACTTCTGAAAGCGC TACCCCTGAGTCCGGCCCAGGTACTTCTGAAAGCG CTACTCCTGAATCCGGTCCAGGTACTTCTGAAAGC GCTACCCCGGAATCTGGCCCAGGTAGCGAACCGG CTACTTCTGGTTCTGAAACCCCAGGTAGCGAACCG GCTACCTCCGGTTCTGAAACTCCAGGTAGCCCAGC		STEPSEGSAPGSP AGSPTSTEEGTST EPSEGSAPGTSTE PSEGSAPGTSESA TPESGPGTSTEPS EGSAPGTSESATP ESGPGSEPATSGS ETPGTSTEPSEGS APGTSTEPSEGSA PGTSESATPESGP GTSESATPESGPG SPAGSPTSTEEGT SESATPESGPGSE PATSGSETPGTSE SATPESGPGTSTE PSEGSAPGTSTEP SEGSAPGTSTEPS EGSAPGTSTEPSE GSAPGTSTEPSEG SAPGTSTEPSEGS APGSPAGSPTSTE EGTSTEPSEGSAP GTSESATPESGPG SEPATSGSETPGT SESATPESGPGSE PATSGSETPGTSE SATPESGPGTSTE PSEGSAPGTSESA TPESGPGSPAGSP TSTEEGSPAGSPT STEEGSPAGSPTS TEEGTSESATPES GPGTSTEPSEGSA PGTSESATPESGP GSEPATSGSETPG TSESATPESGPGS EPATSGSETPGTS ESATPESGPGTST EPSEGSAPGSPAG SPTSTEEGTSESA TPESGPGSEPATS GSETPGTSESATP ESGPGSPAGSPTS TEEGSPAGSPTST EEGTSTEPSEGSA PGTSESATPESGP GTSESATPESGPG TSESATPESGPGS EPATSGSETPGSE PATSGSETPGSPA GSPTSTEEGTSTE PSEGSAPGTSTEP SEGSAPGSEPATS GSETPGTSESATP ESGPGTSTEPSEG SAPGSSSLDIQMT QSPSSLSASVGDR VTITCRASQDVN	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	AGGCTCTCCGACTTCCACTGAGGAAGGTACTTCTA CTGAACCTTCCGAAGGCAGCGCACCAGGTACCTCT ACTGAACCTTCTGAGGGCAGCGCTCCAGGTAGCG AACCTGCAACCTCTGGCTCTGAAACCCCAGGTACC TCTGAAAGCGCTACTCCTGAATCTGGCCCAGGTAC TTCTACTGAACCGTCCGAGGGCAGCGCACCAGGTT CGTCTTCACTCGACATTCAGATGACCCAGAGCCCG TCCTCCCTGAGCGCTTCTGTTGGCGACCGCGTGAC CATCACCTGCCGTGCTTCCCAGGATGTTAACACCG CTGTAGCTTGGTATCAACAGAAACCGGGCAAAGC ACCGAAACTGCTGATCTACTCTGCTTCCTTTCTGTA TAGCGGTGTTCCGTCTCGTTTCAGCGGCTCTCGTA GCGGTACGGATTTTACTCTGACGATCAGCTCTCTG CAGCCGGAGGACTTCGCTACCTACTACTGCCAGCA GCACTACACCACCCCGCCTACCTTTGGTCAGGGCA CCAAAGTGGAAATCAAGACCGGTTCTGGCGAAGG CTCTGAAGGTGAAGGTGGTGGTGAAGGCTCTGAA GGTGAAGGATCTGGTGAAGGTGGCGAAGGTGAGG GATCTGGTACCGAGGTCCAGCTGGTTGAGTCTGGC GGCGGTCTGGTCCAACCTGGTGGCTCCCTGCGCCT GTCTTGCGCAGCGTCCGGCTTTAATATCAAAGATA CGTACATTCACTGGGTCCGCCAGGCACCGGGCAA AGGCCTGGAATGGGTTGCTCGTATCTACCCGACTA ACGGTTATACCCGTTATGCAGACAGCGTAAAGGGT CGCTTCACGATCTCCGCGGATACCTCCAAAAACAC CGCATACCTGCAAATGAACTCTCTGCGTGCGGAAG ATACTGCCGTGTACTACTGCTCTCGCTGGGGCGGT GACGGTTTCTATGCAATGGACTACTGGGGTCAAGG TACTCTGGTAACTGTTTCC		TAVAWYQQKPG KAPKLLIYSASFL YSGVPSRFSGSRS GTDFTLTISSLQP EDFATYYCQQHY TTPPTFGQGTKVE IKTGSGEGSEGEG GGEGSEGEGSGE GGEGEGSGTEVQ LVESGGGLVQPG GSLRLSCAASGF NIKDTYIHWVRQ APGKGLEWVARI YPTNGYTRYADS VKGRFTISADTSK NTAYLQMNSLRA EDTAVYYCSRW GGDGFYAMDYW GQGTLVTVS	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
anti-EGFR-Y576-FLAG-HIS6 ("His6" divulgada como la SEQ ID NO: 218)	ATGGAGGATATTCTGCTGACGCAAAGCCCTGTTAT TCTGTCTGTTAGCCCGGGTGAGCGCGTTAGCTTCA GCTGCCGTGCATCTCAGAGCATTGGCACGAACATT CATTGGTATCAACAACGTACCAACGGTAGCCCGC GTCTGCTGATTAAATACGCATCCGAATCTATCTCT GGTATCCCGTCTCGCTTCAGCGGTTCTGGTAGCGG CACCGACTTTACCCTGAGCATTAACTCTGTAGAAA GCGAAGATATTGCGGATTACTACTGCCAGCAGAA CAACAACTGGCCGACTACTTTTGGTGCAGGTACTA AACTGGAACTGAAAACCGGTTCTGGCGAAGGCTC TGAAGGTGAAGGTGGTGGTGAAGGCTCTGAAGGT GAAGGATCTGGTGAAGGTGGCGAAGGTGAGGGAT CTGGTACCCAAGTGCAGCTGAAACAGAGCGGTCC GGGTCTGGTGCAACCATCCCAGTCTCTGTCTATTA CCTGTACCGTTAGCGGTTTCTCCCTGACCAACTAC GGTGTTCACTGGGTTCGCCAGTCCCCAGGCAAAGG CCTGGAATGGCTGGGCGTTATTTGGTCCGGCGGCA ATACGGATTATAACACCCCGTTCACCTCTCGTCTG TCTATCAACAAAGATAATTCTAAAAGCCAGGTATT CTTCAAGATGAACTCTCTGCAGAGCAATGACACCG CCATCTACTATTGCGCTCGTGCCCTGACTTACTAC GATTACGAGTTCGCATATTGGGGCCAGGGCACTCT GGTGACCGTTTCCGGAGGTGAGGGTTCTGGCGAA GGTTCCGAAGGTGAGGGCTCCGAAGGATCTGGCG AAGGTGAGGGTTCCGAAGGTTCTGGCGAAGGTGA AGGCGGTTCTGAGGGATCCGAAGGTGAAGGCTCC GAAGGATCTGGCGAAGGTGAAGGTGGTGAAGGTT CTGGCGAAGGTGAGGGATCTGGCGAAGGCTCTGA AGGTGAAGGTGGTGGTGAAGGCTCTGAAGGTGAA	714	MEDILLTQSPVIL SVSPGERVSFSCR ASQSIGTNIHWY QQRTNGSPRLLIK YASESISGIPSRFS GSGSGTDFTLSIN SVESEDIADYYC QQNNNWPTTFGA GTKLELKTGSGE GSEGEGGGEGSE GEGSGEGGEGEG SGTQVQLKQSGP GLVQPSQSLSITC TVSGFSLTNYGV HWVRQSPGKGLE WLGVIWSGGNTD YNTPFTSRLSINK DNSKSQVFFKMN SLQSNDTAIYYC ARALTYYDYEFA YWGQGTLVTVS GGEGSGEGSEGE GSEGSGEGEGSE GSGEGEGGSEGS EGEGSEGSGEGE GGEGSGEGEGSG EGSEGEGGGEGS EGEGSGEGGEGE GSEGGSEGEGGS	739

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	GGATCTGGTGAAGGTGGCGAAGGTGAGGGATCTG AAGGCGGCTCCGAAGGTGAAGGCGGATCTGAAGG CGGCGAAGGTGAAGGTTCCGAAGGTTCTGGTGAA GGTGAAGGATCTGAAGGTGGCTCCGAAGGTGAAG GATCTGAAGGCGGTTCCGAAGGTGAGGGCTCTGA AGGTTCTGGCGAAGGTGAAGGCTCTGAAGGATCT GGTGAAGGTGAAGGTTCCGAAGGTTCTGGTGAAG GTGAAGGTTCCGAAGGTTCTGGCGAAGGTGAAGG TTCTGAAGGTGGCTCTGAAGGTGAAGGCGGCTCTG AAGGATCCGAAGGTGAAGGTTCTGGTGAAGGCTC TGAAGGTGAAGGCGGCTCTGAGGGTTCCGAAGGT GAAGGCGGAGGCGAAGGTTCTGAAGGTGAGGGAT CTGGTGAAGGTTCTGAAGGTGAAGGCGGTTCTGA AGGTTCCGAAGGTGAAGGTGGCTCTGAGGGATCC GAAGGTGAAGGTGGCGAAGGATCTGGTGAAGGTG AAGGTTCTGAAGGTTCTGGCGAAGGTGAGGGTTCT GGCGAAGGTTCCGAAGGTGAGGGCTCCGAAGGAT CTGGCGAAGGTGAGGGTTCCGAAGGTTCTGGCGA AGGTGAAGGCGGTTCTGAGGGATCCGAAGGTGAG GGTTCTGGCGAAGGTTCCGAAGGTGAGGGCTCCG AAGGATCTGGCGAAGGTGAGGGTTCCGAAGGTTC TGGCGAAGGTGAAGGCGGTTCTGAGGGATCCGAA GGTGAAGGCGGTTCTGAAGGTTCCGAAGGTGAAG GTGGCTCTGAGGGATCCGAAGGTGAAGGTGGCGA AGGATCTGGTGAAGGTGAAGGTTCTGAAGGTTCTG GCGAAGGTGAGGGTTCTGGCGAAGGTTCCGAAGG TGAGGGCTCCGAAGGATCTGGCGAAGGTGAGGGT TCCGAAGGTTCTGGCGAAGGTGAAGGCGGTTCTG AGGGATCCGAAGGTGAAGGCTCCGAAGGATCTGG CGAAGGTGAAGGTGGTGAAGGTTCTGGCGAAGGT GAGGGATCTGGCGAAGGCTCTGAAGGTGAAGGTG GTGGTGAAGGCTCTGAAGGTGAAGGTTCCGAAGG TTCTGGTGAAGGTGAAGGTTCCGAAGGTTCTGGCG AAGGTGAAGGTTCTGAAGGTGGCTCTGAAGGTGA AGGCGGCTCTGAAGGATCCGAAGGTGAAGGATCT GAAGGTGGCTCCGAAGGTGAAGGATCTGAAGGCG GTTCCGAAGGTGAGGGCTCTGAAGGTTCTGGCGA AGGTGAAGGCTCTGAAGGATCTGGTGAAGGTGAA GGATCTGGCGAAGGCTCCGAAGGTGAAGGCGGTT CTGAAGGTGGCGAAGGTGAAGGATCTGAAGGTGG TTCCGAAGGTGAGGGATCTGAAGGTGGCTCTGAA GGTGAAGGTGGCGAAGGTTCTGGCGAAGGTGAAG GTGGAGGCGAAGGTTCTGAAGGTGAAGGTTCCGA AGGTTCTGGTGAAGGTGAGGGATCTGGCGAAGGT TCTGAAGGTGATTATAAAGACGATGACGATAAAG GTGGTTCTCATCACCATCACCATCACTAA		EGGEGEGSEGSG EGEGSEGGSEGE GSEGGSEGEGSE GSGEGEGSEGSG EGEGSEGSGEGE GSEGSGEGEGSE GGSEGEGGSEGS EGEGSGEGSEGE GGSEGSEGEGGG EGSEGEGSGEGSE GEGGSEGSEGEG GSEGSEGEGGEG SGEGEGSEGSGE GEGSGEGSEGEG SEGSGEGEGSEGS GEGEGGSEGSEG EGSGEGSEGEGSE GSGEGEGSEGSG EGEGGSEGSEGE GGSEGSEGEGGS EGSEGEGGEGSG EGEGSEGSGEGE GSGEGSEGEGSE GSGEGEGSEGSG EGEGGSEGSEGE GSEGSGEGEGGE GSGEGEGSGEGS EGEGGGEGSEGE GSEGSGEGEGSE GSGEGEGSEGGS EGEGGSEGSEGE GSEGGSEGEGSE GGSEGEGSEGSG EGEGSEGSGEGE GSGEGSEGEGGS EGGEGEGSEGGS EGEGSEGGSEGE GGEGSGEGEGGG EGSEGEGSEGSGE GEGSGEGSEGDY KDDDDKGGSHH HHHH	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
anti-CD3-Y288-GFP-HIS8 ("His8" divulgada como la SEQ ID NO: 697)	ATGAAAGACATCCAGATGACCCAGTCTCCTTCCTC TCTGTCCGCGTCCGTGGGCGACCGTGTTACTATCA CCTGCTCCGCCTCCTCTTCTGTCAGCTACATGAACT GGTATCAGCAGACTCCTGGCAAAGCTCCAAAACG TTGGATTTACGATACGTCCAAGCTGGCCTCCGGCG TACCAAGCCGTTTCTCTGGCTCTGGCAGCGGCACG GATTACACCTTCACTATTTCTAGCCTGCAGCCTGA AGATATTGCCACCTATTACTGCCAACAATGGTCCT CCAATCCTTTTACCTTTGGTCAGGGCACTAAGCTG CAGATTACTCGCACCGGTTCTGGCGAAGGCTCTGA AGGTGAAGGTGGTGGTGAAGGCTCTGAAGGTGAA GGATCTGGTGAAGGTGGCGAAGGTGAGGGATCTG	715	MKDIQMTQSPSS LSASVGDRVTITC SASSSVSYMNWY QQTPGKAPKRWI YDTSKLASGVPS RFSGSGSGTDYTF TISSLQPEDIATY YCQQWSSNPFTF GQGTKLQITRTGS GEGSEGEGGGEG SEGEGSGEGGEG EGSGTQVQLVQS	740

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	GTACCCAGGTCCAACTGGTTCAATCCGGCGGCGGT GTAGTTCAACCGGGTCGCTCTCTGCGTCTTTCCTGC AAGGCGTCCGGTTACACTTTCACGCGTTACACCAT GCACTGGGTCCGTCAGGCTCCTGGTAAAGGTCTGG AATGGATTGGCTATATCAACCCGTCTCGCGGCTAT ACCAACTATAACCAGAAATTCAAAGATCGTTTTAC GATTTCCACTGATAAATCCAAAAGCACCGCATTCC TCCAAATGGACAGCCTGCGTCCGGAAGACACGGC GGTTTATTATTCCGCCCGTTACTACGATGACCACT ACTGCCTGGATTATTGGGGCCAAGGCACTCCAGTA ACCGTGAGCAGCGGAGGTGAGGGTTCTGGCGAAG GTTCCGAAGGTGAGGGCTCCGAAGGATCTGGCGA AGGTGAGGGTTCCGAAGGTTCTGGCGAAGGTGAA GGCGGTTCTGAGGGATCCGAAGGTGAAGGCGGTT CTGAGGGATCTGAAGGTGAAGGTGGCTCTGAAGG ATCTGAAGGTGAGGGATCTGGTGAAGGTTCTGAA GGTGAAGGCGGCTCTGAGGGTTCTGAAGGTGAAG GATCTGGTGAAGGTTCCGAAGGTGAGGGTTCTGA AGGTGGTTCTGAAGGTGAAGGCGGTTCTGAGGGTT CTGAAGGTGAGGGTTCTGGCGAAGGTTCCGAAGG TGAAGGCGGCGAAGGTGGATCTGAAGGTGAGGGC TCCGAAGGATCTGGCGAAGGTGAAGGTTCTGGCG AAGGTTCCGAAGGTGAAGGTTCTGAAGGATCTGG CGAAGGTGAGGGTTCTGGCGAAGGTTCCGAAGGT GAGGGCTCCGAAGGATCTGGCGAAGGTGAGGGTT CCGAAGGTTCTGGCGAAGGTGAAGGCGGTTCTGA GGGATCCGAAGGTGAAGGCTCCGAAGGATCTGGC GAAGGTGAAGGTGGTGAAGGTTCTGGCGAAGGTG AGGGATCTGGCGAAGGCTCTGAAGGTGAAGGTGG TGGTGAAGGCTCTGAAGGTGAAGGATCTGGTGAA GGTGGCGAAGGTGAGGGATCTGAAGGCGGCTCCG AAGGTGAAGGCGGATCTGAAGGCGGCGAAGGTGA AGGTTCCGAAGGTTCTGGTGAAGGTGAAGGATCT GAAGGTGGCTCCGAAGGTGAAGGATCTGAAGGCG GTTCCGAAGGTGAGGGCTCTGAAGGTTCTGGCGA AGGTGAAGGCTCTGAAGGATCTGGTGAAGGTTCG TCTTCACTCGAGGGTACCGAACTTTTCACTGGAGT TGTCCCAATTCTTGTTGAATTAGATGGTGATGTTA ATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGA AGGTGATGCAACATACGGAAAACTTACCCTTAAAT TTATTTGCACTACTGGAAAACTACCTGTTCCATGG CCAACACTTGTCACTACTTTCTCTTATGGTGTTCAA TGCTTTTCCCGTTATCCGGATCACATGAAACGGCA TGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATG TACAGGAACGCACTATATCTTTCAAAGATGACGGG AACTACAAGACGCGTGCTGAAGTCAAGTTTGAAG GTGATACCCTTGTTAATCGTATCGAGTTAAAAGGT ATTGATTTTAAAGAAGATGGAAACATTCTCGGACA CAAACTCGAGTACAACTATAACTCACACAATGTAT ACATCACGGCAGACAAACAAAAGAATGGAATCAA AGCTAACTTCAAAATTCGCCACAACATTGAAGATG GATCCGTTCAACTAGCAGACCATTATCAACAAAAT ACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGA CAACCATTACCTGTCGACACAATCTGCCCTTTCGA AAGATCCCAACGAAAAGCGTGACCACATGGTCCT TCTTGAGTTTGTAACTGCTGCTGGGATTGGTGGCT CTCATCACCATCACCATCACCATCACTAA		GGGVVQPGRSLR LSCKASGYTFTR YTMHWVRQAPG KGLEWIGYINPSR GYTNYNQKFKDR FTISTDKSKSTAF LQMDSLRPEDTA VYYSARYYDDH YCLDYWGQGTP VTVSSGGEGSGE GSEGEGSEGSGE GEGSEGSGEGEG GSEGSEGEGGSE GSEGEGGSEGSE GEGSGEGSEGEG GSEGSEGEGSGE GSEGEGSEGGSE GEGGSEGSEGEG SGEGSEGEGGEG GSEGEGSEGSGE GEGSGEGSEGEG SEGSGEGEGSGE GSEGEGSEGSGE GEGSEGSGEGEG GSEGSEGEGSEGS GEGEGGEGSGEG EGSGEGSEGEGG GEGSEGEGSGEG GEGEGSEGGSEG EGGSEGGEGEGS EGSGEGEGSEGG SEGEGSEGGSEGE GSEGSGEGEGSE GSGEGSSSLEGTE LFTGVVPILVELD GDVNGHKFSVSG EGEGDATYGKLT LKFICTTGKLPVP WPTLVTTFSYGV QCFSRYPDHMKR HDFFKSAMPEGY VQERTISFKDDG NYKTRAEVKFEG DTLVNRIELKGID FKEDGNILGHKL EYNYNSHNVYIT ADKQKNGIKANF KIRHNIEDGSVQL ADHYQQNTPIGD GPVLLPDNHYLS TQSALSKDPNEK RDHMVLLEFVTA AGIGGSHHHHHH HH	

Nombre del clon	**Secuencia de ADN**	**SEQ ID NO:**	**Secuencia de aminoácidos**	**SEQ ID NO:**
anti-**Her2-Y288-anti-CD3-HA-His6, AC48** ("His6" divulgada como la SEQ ID NO: 218)	ATGGAGGACATTCAGATGACCCAGAGCCCGTCCT	716	MEDIQMTQSPSSL	741

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	CCCTGAGCGCTTCTGTTGGCGACCGCGTGACCATC ACCTGCCGTGCTTCCCAGGATGTTAACACCGCTGT AGCTTGGTATCAACAGAAACCGGGCAAAGCACCG AAACTGCTGATCTACTCTGCTTCCTTTCTGTATAGC GGTGTTCCGTCTCGTTTCAGCGGCTCTCGTAGCGG TACGGATTTTACTCTGACGATCAGCTCTCTGCAGC CGGAGGACTTCGCTACCTACTACTGCCAGCAGCAC TACACCACCCCGCCTACCTTTGGTCAGGGCACCAA AGTGGAAATCAAGACCGGTTCTGGCGAAGGCTCT GAAGGTGAAGGTGGTGGTGAAGGCTCTGAAGGTG AAGGATCTGGTGAAGGTGGCGAAGGTGAGGGATC TGGTACCGAGGTCCAGCTGGTTGAGTCTGGCGGCG GTCTGGTCCAACCTGGTGGCTCCCTGCGCCTGTCT TGCGCAGCGTCCGGCTTTAATATCAAAGATACGTA CATTCACTGGGTCCGCCAGGCACCGGGCAAAGGC CTGGAATGGGTTGCTCGTATCTACCCGACTAACGG TTATACCCGTTATGCAGACAGCGTAAAGGGTCGCT TCACGATCTCCGCGGATACCTCCAAAAACACCGCA TACCTGCAAATGAACTCTCTGCGTGCGGAAGATAC TGCCGTGTACTACTGCTCTCGCTGGGGCGGTGACG GTTTCTATGCAATGGACTACTGGGGTCAAGGTACT CTGGTAACTGTTTCCGGAGGTGAGGGTTCTGGCGA AGGTTCCGAAGGTGAGGGCTCCGAAGGATCTGGC GAAGGTGAGGGTTCCGAAGGTTCTGGCGAAGGTG AAGGCGGTTCTGAGGGATCCGAAGGTGAAGGCGG TTCTGAGGGATCTGAAGGTGAAGGTGGCTCTGAA GGATCTGAAGGTGAGGGATCTGGTGAAGGTTCTG AAGGTGAAGGCGGCTCTGAGGGTTCTGAAGGTGA AGGATCTGGTGAAGGTTCCGAAGGTGAGGGTTCT GAAGGTGGTTCTGAAGGTGAAGGCGGTTCTGAGG GTTCTGAAGGTGAGGGTTCTGGCGAAGGTTCCGAA GGTGAAGGCGGCGAAGGTGGATCTGAAGGTGAGG GCTCCGAAGGATCTGGCGAAGGTGAAGGTTCTGG CGAAGGTTCCGAAGGTGAAGGTTCTGAAGGATCT GGCGAAGGTGAGGGTTCTGGCGAAGGTTCCGAAG GTGAGGGCTCCGAAGGATCTGGCGAAGGTGAGGG TTCCGAAGGTTCTGGCGAAGGTGAAGGCGGTTCTG AGGGATCCGAAGGTGAAGGCTCCGAAGGATCTGG CGAAGGTGAAGGTGGTGAAGGTTCTGGCGAAGGT GAGGGATCTGGCGAAGGCTCTGAAGGTGAAGGTG GTGGTGAAGGCTCTGAAGGTGAAGGATCTGGTGA AGGTGGCGAAGGTGAGGGATCTGAAGGCGGCTCC GAAGGTGAAGGCGGATCTGAAGGCGGCGAAGGTG AAGGTTCCGAAGGTTCTGGTGAAGGTGAAGGATC TGAAGGTGGCTCCGAAGGTGAAGGATCTGAAGGC GGTTCCGAAGGTGAGGGCTCTGAAGGTTCTGGCG AAGGTGAAGGCTCTGAAGGATCTGGTGAAGGTGA CATCCAGATGACCCAGTCTCCTTCCTCTCTGTCCG CGTCCGTGGGCGACCGTGTTACTATCACCTGCTCC GCCTCCTCTTCTGTCAGCTACATGAACTGGTATCA GCAGACTCCTGGCAAAGCTCCAAAACGTTGGATTT ACGATACGTCCAAGCTGGCCTCCGGCGTACCAAG CCGTTTCTCTGGCTCTGGCAGCGGCACGGATTACA CCTTCACTATTTCTAGCCTGCAGCCTGAAGATATT GCCACCTATTACTGCCAACAATGGTCCTCCAATCC TTTTACCTTTGGTCAGGGCACTAAGCTGCAGATTA CTCGCACCGGTTCTGGCGAAGGCTCTGAAGGTGAA GGTGGTGGTGAAGGCTCTGAAGGTGAAGGATCTG		SASVGDRVTITCR ASQDVNTAVAW YQQKPGKAPKLL IYSASFLYSGVPS RFSGSRSGTDFTL TISSLQPEDFATY YCQQHYTTPPTF GQGTKVEIKTGS GEGSEGEGGGEG SEGEGSGEGGEG EGSGTEVQLVES GGGLVQPGGSLR LSCAASGFNIKDT YIHWVRQAPGKG LEWVARIYPTNG YTRYADSVKGRF TISADTSKNTAYL QMNSLRAEDTAV YYCSRWGGDGF YAMDYWGQGTL VTVSGGEGSGEG SEGEGSEGSGEGE GSEGSGEGEGGS EGSEGEGGSEGSE GEGGSEGSEGEG SGEGSEGEGGSE GSEGEGSGEGSE GEGSEGGSEGEG GSEGSEGEGSGE GSEGEGGEGGSE GEGSEGSGEGEG SGEGSEGEGSEGS GEGEGSGEGSEG EGSEGSGEGEGSE GSGEGEGGSEGS EGEGSEGSGEGE GGEGSGEGEGSG EGSEGEGGGEGS EGEGSGEGGEGE GSEGGSEGEGGS EGGEGEGSEGSG EGEGSEGGSEGE GSEGGSEGEGSE GSGEGEGSEGSG EGDIQMTQSPSSL SASVGDRVTITCS ASSSVSYMNWY QQTPGKAPKRWI YDTSKLASGVPS RFSGSGSGTDYTF TISSLQPEDIATY YCQQWSSNPFTF GQGTKLQITRTGS GEGSEGEGGGEG SEGEGSGEGGEG EGSGTQVQLVQS GGGVVQPGRSLR LSCKASGYTFTR	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	GTGAAGGTGGCGAAGGTGAGGGATCTGGTACCCA GGTCCAACTGGTTCAATCCGGCGGCGGTGTAGTTC AACCGGGTCGCTCTCTGCGTCTTTCCTGCAAGGCG TCCGGTTACACTTTCACGCGTTACACCATGCACTG GGTCCGTCAGGCTCCTGGTAAAGGTCTGGAATGGA TTGGCTATATCAACCCGTCTCGCGGCTATACCAAC TATAACCAGAAATTCAAAGATCGTTTTACGATTTC CACTGATAAATCCAAAAGCACCGCATTCCTCCAAA TGGACAGCCTGCGTCCGGAAGACACGGCGGTTTAT TATTCCGCCCGTTACTACGATGACCACTACTGCCT GGATTATTGGGGCCAAGGCACTCCAGTAACCGTG AGCAGCGGCGGTTATCCTTATGATGTTCCAGACTA TGCAGGTGGCTCTCATCACCATCACCATCACTGA		YTMHWVRQAPG KGLEWIGYINPSR GYTNYNQKFKDR FTISTDKSKSTAF LQMDSLRPEDTA VYYSARYYDDH YCLDYWGQGTP VTVSSGGYPYDV PDYAGGSHHHHH H	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
anti-Her2-Y288-anti-EGFR-HA-His6, AC49 ("His6" divulgada como la SEQ ID NO: 218)	ATGGAGGACATTCAGATGACCCAGAGCCCGTCCT CCCTGAGCGCTTCTGTTGGCGACCGCGTGACCATC ACCTGCCGTGCTTCCCAGGATGTTAACACCGCTGT AGCTTGGTATCAACAGAAACCGGGCAAAGCACCG AAACTGCTGATCTACTCTGCTTCCTTTCTGTATAGC GGTGTTCCGTCTCGTTTCAGCGGCTCTCGTAGCGG TACGGATTTTACTCTGACGATCAGCTCTCTGCAGC CGGAGGACTTCGCTACCTACTACTGCCAGCAGCAC TACACCACCCCGCCTACCTTTGGTCAGGGCACCAA AGTGGAAATCAAGACCGGTTCTGGCGAAGGCTCT GAAGGTGAAGGTGGTGGTGAAGGCTCTGAAGGTG AAGGATCTGGTGAAGGTGGCGAAGGTGAGGGATC TGGTACCGAGGTCCAGCTGGTTGAGTCTGGCGGCG GTCTGGTCCAACCTGGTGGCTCCCTGCGCCTGTCT TGCGCAGCGTCCGGCTTTAATATCAAAGATACGTA CATTCACTGGGTCCGCCAGGCACCGGGCAAAGGC CTGGAATGGGTTGCTCGTATCTACCCGACTAACGG TTATACCCGTTATGCAGACAGCGTAAAGGGTCGCT TCACGATCTCCGCGGATACCTCCAAAAACACCGCA TACCTGCAAATGAACTCTCTGCGTGCGGAAGATAC TGCCGTGTACTACTGCTCTCGCTGGGGCGGTGACG GTTTCTATGCAATGGACTACTGGGGTCAAGGTACT CTGGTAACTGTTTCCGGAGGTGAGGGTTCTGGCGA AGGTTCCGAAGGTGAGGGCTCCGAAGGATCTGGC GAAGGTGAGGGTTCCGAAGGTTCTGGCGAAGGTG AAGGCGGTTCTGAGGGATCCGAAGGTGAAGGCGG TTCTGAGGGATCTGAAGGTGAAGGTGGCTCTGAA GGATCTGAAGGTGAGGGATCTGGTGAAGGTTCTG AAGGTGAAGGCGGCTCTGAGGGTTCTGAAGGTGA AGGATCTGGTGAAGGTTCCGAAGGTGAGGGTTCT GAAGGTGGTTCTGAAGGTGAAGGCGGTTCTGAGG GTTCTGAAGGTGAGGGTTCTGGCGAAGGTTCCGAA GGTGAAGGCGGCGAAGGTGGATCTGAAGGTGAGG GCTCCGAAGGATCTGGCGAAGGTGAAGGTTCTGG CGAAGGTTCCGAAGGTGAAGGTTCTGAAGGATCT GGCGAAGGTGAGGGTTCTGGCGAAGGTTCCGAAG GTGAGGGCTCCGAAGGATCTGGCGAAGGTGAGGG TTCCGAAGGTTCTGGCGAAGGTGAAGGCGGTTCTG AGGGATCCGAAGGTGAAGGCTCCGAAGGATCTGG CGAAGGTGAAGGTGGTGAAGGTTCTGGCGAAGGT GAGGGATCTGGCGAAGGCTCTGAAGGTGAAGGTG GTGGTGAAGGCTCTGAAGGTGAAGGATCTGGTGA AGGTGGCGAAGGTGAGGGATCTGAAGGCGGCTCC GAAGGTGAAGGCGGATCTGAAGGCGGCGAAGGTG AAGGTTCCGAAGGTTCTGGTGAAGGTGAAGGATC	717	MEDIQMTQSPSSL SASVGDRVTITCR ASQDVNTAVAW YQQKPGKAPKLL IYSASFLYSGVPS RFSGSRSGTDFTL TISSLQPEDFATY YCQQHYTTPPTF GQGTKVEIKTGS GEGSEGEGGGEG SEGEGSGEGGEG EGSGTEVQLVES GGGLVQPGGSLR LSCAASGFNIKDT YIHWVRQAPGKG LEWVARIYPTNG YTRYADSVKGRF TISADTSKNTAYL QMNSLRAEDTAV YYCSRWGGDGF YAMDYWGQGTL VTVSGGEGSGEG SEGEGSEGSGEGE GSEGSGEGEGGS EGSEGEGGSEGSE GEGGSEGSEGEG SGEGSEGEGGSE GSEGEGSGEGSE GEGSEGGSEGEG GSEGSEGEGSGE GSEGEGGEGGSE GEGSEGSGEGEG SGEGSEGEGSEGS GEGEGSGEGSEG EGSEGSGEGEGSE GSGEGEGGSEGS EGEGSEGSGEGE GGEGSGEGEGSG EGSEGEGGGEGS EGEGSGEGGEGE GSEGGSEGEGGS EGGEGEGSEGSG EGEGSEGGSEGE GSEGGSEGEGSE GSGEGEGSEGSG	742

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	TGAAGGTGGCTCCGAAGGTGAAGGATCTGAAGGC GGTTCCGAAGGTGAGGGCTCTGAAGGTTCTGGCG AAGGTGAAGGCTCTGAAGGATCTGGTGAAGGTGA GGATATTCTGCTGACGCAAAGCCCTGTTATTCTGT CTGTTAGCCCGGGTGAGCGCGTTAGCTTCAGCTGC CGTGCATCTCAGAGCATTGGCACGAACATTCATTG GTATCAACAACGTACCAACGGTAGCCCGCGTCTGC TGATTAAATACGCATCCGAATCTATCTCTGGTATC CCGTCTCGCTTCAGCGGTTCTGGTAGCGGCACCGA CTTTACCCTGAGCATTAACTCTGTAGAAAGCGAAG ATATTGCGGATTACTACTGCCAGCAGAACAACAAC TGGCCGACTACTTTTGGTGCAGGTACTAAACTGGA ACTGAAAACCGGTTCTGGCGAAGGCTCTGAAGGT GAAGGTGGTGGTGAAGGCTCTGAAGGTGAAGGAT CTGGTGAAGGTGGCGAAGGTGAGGGATCTGGTAC CCAAGTGCAGCTGAAACAGAGCGGTCCGGGTCTG GTGCAACCATCCCAGTCTCTGTCTATTACCTGTAC CGTTAGCGGTTTCTCCCTGACCAACTACGGTGTTC ACTGGGTTCGCCAGTCCCCAGGCAAAGGCCTGGA ATGGCTGGGCGTTATTTGGTCCGGCGGCAATACGG ATTATAACACCCCGTTCACCTCTCGTCTGTCTATCA ACAAAGATAATTCTAAAAGCCAGGTATTCTTCAAG ATGAACTCTCTGCAGAGCAATGACACCGCCATCTA CTATTGCGCTCGTGCCCTGACTTACTACGATTACG AGTTCGCATATTGGGGCCAGGGCACTCTGGTGACC GTTTCCGGCGGTTATCCTTATGATGTTCCAGACTAT GCAGGTGGCTCTCATCACCATCACCATCACTGA		EGEDILLTQSPVIL SVSPGERVSFSCR ASQSIGTNIHWY QQRTNGSPRLLIK YASESISGIPSRFS GSGSGTDFTLSIN SVESEDIADYYC QQNNNWPTTFGA GTKLELKTGSGE GSEGEGGGEGSE GEGSGEGGEGEG SGTQVQLKQSGP GLVQPSQSLSITC TVSGFSLTNYGV HWVRQSPGKGLE WLGVIWSGGNTD YNTPFTSRLSINK DNSKSQVFFKMN SLQSNDTAIYYC ARALTYYDYEFA YWGQGTLVTVS GGYPYDVPDYAG GSHHHHHH	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
anti-Her2-Y288-anti-CD3-HA-His8, AC69 ("His8" divulgada como la SEQ ID NO: 697)	ATGGAGGACATTCAGATGACCCAGAGCCCGTCCT CCCTGAGCGCTTCTGTTGGCGACCGCGTGACCATC ACCTGCCGTGCTTCCCAGGATGTTAACACCGCTGT AGCTTGGTATCAACAGAAACCGGGCAAAGCACCG AAACTGCTGATCTACTCTGCTTCCTTTCTGTATAGC GGTGTTCCGTCTCGTTTCAGCGGCTCTCGTAGCGG TACGGATTTTACTCTGACGATCAGCTCTCTGCAGC CGGAGGACTTCGCTACCTACTACTGCCAGCAGCAC TACACCACCCCGCCTACCTTTGGTCAGGGCACCAA AGTGGAAATCAAGACCGGTTCTGGCGAAGGCTCT GAAGGTGAAGGTGGTGGTGAAGGCTCTGAAGGTG AAGGATCTGGTGAAGGTGGCGAAGGTGAGGGATC TGGTACCGAGGTCCAGCTGGTTGAGTCTGGCGGCG GTCTGGTCCAACCTGGTGGCTCCCTGCGCCTGTCT TGCGCAGCGTCCGGCTTTAATATCAAAGATACGTA CATTCACTGGGTCCGCCAGGCACCGGGCAAAGGC CTGGAATGGGTTGCTCGTATCTACCCGACTAACGG TTATACCCGTTATGCAGACAGCGTAAAGGGTCGCT TCACGATCTCCGCGGATACCTCCAAAAACACCGCA TACCTGCAAATGAACTCTCTGCGTGCGGAAGATAC TGCCGTGTACTACTGCTCTCGCTGGGGCGGTGACG GTTTCTATGCAATGGACTACTGGGGTCAAGGTACT CTGGTAACTGTTTCCGGGTCTCCAGGTGAGGGTTC TGGCGAAGGTTCCGAAGGTGAGGGCTCCGAAGGA TCTGGCGAAGGTGAGGGTTCCGAAGGTTCTGGCG AAGGTGAAGGCGGTTCTGAGGGATCCGAAGGTGA AGGCGGTTCTGAGGGATCTGAAGGTGAAGGTGGC TCTGAAGGATCTGAAGGTGAGGGATCTGGTGAAG GTTCTGAAGGTGAAGGCGGCTCTGAGGGTTCTGAA GGTGAAGGATCTGGTGAAGGTTCCGAAGGTGAGG GTTCTGAAGGTGGTTCTGAAGGTGAAGGCGGTTCT	718	MEDIQMTQSPSSL SASVGDRVTITCR ASQDVNTAVAW YQQKPGKAPKLL IYSASFLYSGVPS RFSGSRSGTDFTL TISSLQPEDFATY YCQQHYTTPPTF GQGTKVEIKTGS GEGSEGEGGGEG SEGEGSGEGGEG EGSGTEVQLVES GGGLVQPGGSLR LSCAASGFNIKDT YIHWVRQAPGKG LEWVARIYPTNG YTRYADSVKGRF TISADTSKNTAYL QMNSLRAEDTAV YYCSRWGGDGF YAMDYWGQGTL VTVSGSPGEGSG EGSEGEGSEGSGE GEGSEGSGEGEG GSEGSEGEGGSE GSEGEGGSEGSE GEGSGEGSEGEG GSEGSEGEGSGE GSEGEGSEGGSE GEGGSEGSEGEG SGEGSEGEGGEG	743

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	GAGGGTTCTGAAGGTGAGGGTTCTGGCGAAGGTT CCGAAGGTGAAGGCGGCGAAGGTGGATCTGAAGG TGAGGGCTCCGAAGGATCTGGCGAAGGTGAAGGT TCTGGCGAAGGTTCCGAAGGTGAAGGTTCTGAAG GATCTGGCGAAGGTGAGGGTTCTGGCGAAGGTTC CGAAGGTGAGGGCTCCGAAGGATCTGGCGAAGGT GAGGGTTCCGAAGGTTCTGGCGAAGGTGAAGGCG GTTCTGAGGGATCCGAAGGTGAAGGCTCCGAAGG ATCTGGCGAAGGTGAAGGTGGTGAAGGTTCTGGC GAAGGTGAGGGATCTGGCGAAGGCTCTGAAGGTG AAGGTGGTGGTGAAGGCTCTGAAGGTGAAGGATC TGGTGAAGGTGGCGAAGGTGAGGGATCTGAAGGC GGCTCCGAAGGTGAAGGCGGATCTGAAGGCGGCG AAGGTGAAGGTTCCGAAGGTTCTGGTGAAGGTGA AGGATCTGAAGGTGGCTCCGAAGGTGAAGGATCT GAAGGCGGTTCCGAAGGTGAGGGCTCTGAAGGTT CTGGCGAAGGTGAAGGCTCTGAAGGATCTGGTGA AGGTTCGTCTTCACTCGAGGGTACCAAAGACATCC AGATGACCCAGTCTCCTTCCTCTCTGTCCGCGTCC GTGGGCGACCGTGTTACTATCACCTGCTCCGCCTC CTCTTCTGTCAGCTACATGAACTGGTATCAGCAGA CTCCTGGCAAAGCTCCAAAACGTTGGATTTACGAT ACGTCCAAGCTGGCCTCCGGCGTACCAAGCCGTTT CTCTGGCTCTGGCAGCGGCACGGATTACACCTTCA CTATTTCTAGCCTGCAGCCTGAAGATATTGCCACC TATTACTGCCAACAATGGTCCTCCAATCCTTTTACC TTTGGTCAGGGCACTAAGCTGCAGATTACTCGCAC CGGTTCTGGCGAAGGCTCTGAAGGTGAAGGTGGT GGTGAAGGCTCTGAAGGTGAAGGATCTGGTGAAG GTGGCGAAGGTGAGGGATCTGGTACCCAGGTCCA ACTGGTTCAATCCGGCGGCGGTGTAGTTCAACCGG GTCGCTCTCTGCGTCTTTCCTGCAAGGCGTCCGGTT ACACTTTCACGCGTTACACCATGCACTGGGTCCGT CAGGCTCCTGGTAAAGGTCTGGAATGGATTGGCTA TATCAACCCGTCTCGCGGCTATACCAACTATAACC AGAAATTCAAAGATCGTTTTACGATTTCCACTGAT AAATCCAAAAGCACCGCATTCCTCCAAATGGACA GCCTGCGTCCGGAAGACACGGCGGTTTATTATTCC GCCCGTTACTACGATGACCACTACTGCCTGGATTA TTGGGGCCAAGGCACTCCAGTAACCGTGAGCAGC ACTAGTGGCGGTTATCCTTATGATGTTCCAGACTA TGCAGGTGGCTCTCATCACCATCACCATCACCACC ATTGA		GSEGEGSEGSGE GEGSGEGSEGEG SEGSGEGEGSGE GSEGEGSEGSGE GEGSEGSGEGEG GSEGSEGEGSEGS GEGEGGEGSGEG EGSGEGSEGEGG GEGSEGEGSGEG GEGEGSEGGSEG EGGSEGGEGEGS EGSGEGEGSEGG SEGEGSEGGSEGE GSEGSGEGEGSE GSGEGSSSLEGTK DIQMTQSPSSLSA SVGDRVTITCSAS SSVSYMNWYQQ TPGKAPKRWIYD TSKLASGVPSRFS GSGSGTDYTFTIS SLQPEDIATYYCQ QWSSNPFTFGQG TKLQITRTGSGEG SEGEGGGEGSEG EGSGEGGEGEGS GTQVQLVQSGGG VVQPGRSLRLSC KASGYTFTRYTM HWVRQAPGKGL EWIGYINPSRGYT NYNQKFKDRFTIS TDKSKSTAFLQM DSLRPEDTAVYY SARYYDDHYCLD YWGQGTPVTVSS TSGGYPYDVPDY AGGSHHHHHHH H	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
EGFR_V HH1-AM144-GFP6-22 9-H8, LMS109.005	ATGAAAGGGTCTCCAGGTGAAGTACAGCTTCAAG AATCTGGTGGTGGTCTTGTCCAGGCGGGCGATTCC CTGCGCCTGTCTTGTCTGGTCTCTGGTCGTTCATTT AACAGCTATACCATGGGCTGGTTCCGCCAAGCACC GGGCAAGGAACGTGAATTCGTAGCAGCTATTCTCT GGTCCGGTCCTACGACCTACTATGCTGACTCTGTA AAAGGTCGCTTCACCATCTCCCGTGATAACGCCAA AAACACCGTATATCTTCAGATGAACTCTCTGAAAC CGGAGGACACGGCCGTGTACTATTGTGCCGCTGCG CTGGGTGTACTGGTGCTAGCGCCTGGTAATGTCTA CAGCTATTGGGGTCAAGGTACCCAGGTCACGGTA AGCTCCGCGCATCATGGAGGTACCCCGGGCAGCG GTACCGCATCTTCCTCTCCAGGTAGCTCTACCCCG TCTGGTGCTACCGGTTCCCCAGGTAGCTCTACCCC GTCTGGTGCAACCGGCTCCCCAGGTAGCCCGGCTG	719	MKGSPGEVQLQE SGGGLVQAGDSL RLSCLVSGRSFNS YTMGWFRQAPG KEREFVAAILWS GPTTYYADSVKG RFTISRDNAKNTV YLQMNSLKPEDT AVYYCAAALGV LVLAPGNVYSYW GQGTQVTVSSAH HGGTPGSGTASSS PGSSTPSGATGSP GSSTPSGATGSPG SPAGSPTSTEEGT	744
	GCTCTCCTACCTCTACTGAGGAAGGTACTTCTGAA AGCGCTACTCCTGAGTCTGGTCCAGGTACCTCTAC TGAACCGTCCGAAGGTAGCGCTCCAGGTTCTAGCC CTTCTGCATCCACCGGTACCGGCCCAGGTTCTAGC CCGTCTGCTTCTACCGGTACTGGTCCAGGTGCTTCT CCGGGTACTAGCTCTACTGGTTCTCCAGGTACCTC TACCGAACCGTCCGAGGGTAGCGCACCAGGTACC TCTACTGAACCGTCTGAGGGTAGCGCTCCAGGTAG CGAACCGGCAACCTCCGGTTCTGAAACTCCAGGTT CGTCTTCACTCGAGGGTACCGAACTTTTCACTGGA GTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTT AATGGGCACAAATTTTCTGTCAGTGGAGAGGGTG AAGGTGATGCAACATACGGAAAACTTACCCTTAA ATTTATTTGCACTACTGGAAAACTACCTGTTCCAT GGCCAACACTTGTCACTACTTTCTCTTATGGTGTTC AATGCTTTTCCCGTTATCCGGATCACATGAAACGG CATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTA TGTACAGGAACGCACTATATCTTTCAAAGATGACG GGAACTACAAGACGCGTGCTGAAGTCAAGTTTGA AGGTGATACCCTTGTTAATCGTATCGAGTTAAAAG GTATTGATTTTAAAGAAGATGGAAACATTCTCGGA CACAAACTCGAGTACAACTATAACTCACACAATGT ATACATCACGGCAGACAAACAAAAGAATGGAATC AAAGCTAACTTCAAAATTCGCCACAACATTGAAG ATGGATCCGTTCAACTAGCAGACCATTATCAACAA AATACTCCAATTGGCGATGGCCCTGTCCTTTTACC AGACAACCATTACCTGTCGACACAATCTGCCCTTT CGAAAGATCCCAACGAAAAGCGTGACCACATGGT CCTTCTTGAGTTTGTAACTGCTGCTGGGATTGGTG GCTCTCATCACCATCACCATCACCATCACTAA		SESATPESGPGTS TEPSEGSAPGSSP SASTGTGPGSSPS ASTGTGPGASPG TSSTGSPGTSTEP SEGSAPGTSTEPS EGSAPGSEPATSG SETPGSSSLEGTE LFTGVVPILVELD GDVNGHKFSVSG EGEGDATYGKLT LKFICTTGKLPVP WPTLVTTFSYGV QCFSRYPDHMKR HDFFKSAMPEGY VQERTISFKDDG NYKTRAEVKFEG DTLVNRIELKGI FKEDGNILGHKL EYNYNSHNVYIT ADKQKNGIKANF KIRHNIEDGSVQL ADHYQQNTPIGD GPVLLPDNHYLS TQSALSKDPNEK RDHMVLLEFVTA AGIGGSHHHHHH HH	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
EGFR_V HH1-AM144-GFP6-22 9-H8, LMS109.020	ATGAAAGGGTCTCCAGGTGAAGTGCAGCTTCAAgA ATCTGGTGGTGGTCTGGTACAAGCCGGTGATTCTC TGCGCCTGTCTTGTCTGGTCTCCGGTCGCTCTTTTA ACAGCTATACCATGGGCTGGTTCCGCCAGGCACCA GGCAAAGAGCGTGAATTCGTAGCAGCTATCCTGTG GTCTGGTCCGACTACCTACTATGCTGACTCTGTAA AGGGTCGCTTCACGATTTCCCGTGATAACGCCAAA AACACGGTGTATCTACAAATGAATTCTCTGAAACC GGAGGACACTGCCGTTTACTATTGTGCCGCTGCGC TGGGTGTACTGGTGCTTGCCCCTGGTAATGTATAC AGCTATTGGGGTCAAGGTACGCAAGTTACCGTGA GCTCTGCGCATCATGGAGGTACTTCTACCGAACCG TCCGAGGGCAGCGCTCCAGGTACTTCTACTGAACC TTCTGAAGGCAGCGCTCCAGGTACTTCTACTGAAC CTTCCGAAGGTAGCGCACCAGGTTCTACCAGCGA ATCCCCTTCTGGTACTGCTCCAGGTTCTACCAGCG AATCCCCTTCTGGCACCGCACCAGGTACTTCTACC CCTGAAAGCGGCTCCGCTTCTCCAGGTAGCGAACC TGCAACCTCTGGCTCTGAAACCCCAGGTACCTCTG AAAGCGCTACTCCTGAATCTGGCCCAGGTACTTCT ACTGAACCGTCCGAGGGCAGCGCACCAGGTACTT CTACTGAACCGTCTGAAGGTAGCGCACCAGGTACT TCTGAAAGCGCAACCCCGGAATCCGGCCCAGGTA CCTCTGAAAGCGCAACCCCGGAGTCCGGCCCAGG TTCGTCTTCACTCGAGGGTACCGAACTTTTCACTG GAGTTGTCCCAATTCTTGTTGAATTAGATGGTGAT GTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGG	720	MKGSPGEVQLQE SGGGLVQAGDSL RLSCLVSGRSFNS YTMGWFRQAPG KEREFVAAILWS GPTTYYADSVKG RFTISRDNAKNTV YLQMNSLKPEDT AVYYCAAALGV LVLAPGNVYSYW GQGTQVTVSSAH HGGTSTEPSEGSA PGTSTEPSEGSAP GTSTEPSEGSAPG STSESPSGTAPGS TSESPSGTAPGTS TPESGSASPGSEP ATSGSETPGTSES ATPESGPGTSTEP SEGSAPGTSTEPS EGSAPGTSESATP ESGPGTSESATPE SGPGSSSLEGTEL FTGVVPILVELDG DVNGHKFSVSGE GEGDATYGKLTL KFICTTGKLPVPW	745
	TGAAGGTGATGCAACATACGGAAAACTTACCCTTA AATTTATTTGCACTACTGGAAAACTACCTGTTCCA TGGCCAACACTTGTCACTACTTTCTCTTATGGTGTT CAATGCTTTTCCCGTTATCCGGATCACATGAAACG GCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTT ATGTACAGGAACGCACTATATCTTTCAAAGATGAC GGGAACTACAAGACGCGTGCTGAAGTCAAGTTTG AAGGTGATACCCTTGTTAATCGTATCGAGTTAAAA GGTATTGATTTTAAAGAAGATGGAAACATTCTCGG ACACAAACTCGAGTACAACTATAACTCACACAAT GTATACATCACGGCAGACAAACAAAAGAATGGAA TCAAAGCTAACTTCAAAATTCGCCACAACATTGAA GATGGATCCGTTCAACTAGCAGACCATTATCAACA AAATACTCCAATTGGCGATGGCCCTGTCCTTTTAC CAGACAACCATTACCTGTCGACACAATCTGCCCTT TCGAAAGATCCCAACGAAAAGCGTGACCACATGG TCCTTCTTGAGTTTGTAACTGCTGCTGGGATTGGTG GCTCTCATCACCATCACCATCACCATCACTAA		PTLVTTFSYGVQ CFSRYPDHMKRH DFFKSAMPEGYV QERTISFKDDGN YKTRAEVKFEGD TLVNRIELKGIDF KEDGNILGHKLE YNYNSHNVYITA DKQKNGIKANFK IRHNIEDGSVQLA DHYQQNTPIGDG PVLLPDNHYLST QSALSKDPNEKR DHMVLLEFVTAA GIGGSHHHHHHH H	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
EGFR_V HH1-AM144-GFP6-22 9-H8, LMS109.038	ATGAAAGGGTCTCCAGGTGAGGTTCAACTTCaAgA ATCTGGTGGTGGTCTAGTACAAGCCGGCGACTCCC TGCGCCTGTCTTGTCTGGTCTCCGGTCGTTCTTTTA ACAGCTATACCATGGGCTGGTTCCGCCAAGCTCCG GGCAAAGAACGTGAATTCGTAGCAGCTATTCTCTG GTCTGGTCCTACCACCTACTATGCTGACTCTGTAA AGGGCCGTTTCACTATCTCCCGTGATAACGCCAAA AACACTGTCTATCTGCAGATGAATTCTCTGAAACC GGAGGACACCGCAGTATACTATTGCGCAGCTGCG CTGGGTGTACTGGTGCTCGCTCCAGGTAATGTATA CAGCTATTGGGGTCAAGGTACGCAAGTCACGGTA AGCTCTGCGCATCATGGAGGTACCCCGGGCAGCG GTACCGCATCTTCCTCTCCAGGTAGCTCTACCCCG TCTGGTGCTACCGGTTCCCCAGGTAGCTCTACCCC GTCTGGTGCAACCGGCTCCCCAGGTAGCCCGGCTG GCTCTCCTACCTCTACTGAGGAAGGTACTTCTGAA AGCGCTACTCCTGAGTCTGGTCCAGGTACCTCTAC TGAACCGTCCGAAGGTAGCGCTCCAGGTTCTAGCC CTTCTGCATCCACCGGTACCGGCCCAGGTTCTAGC CCGTCTGCTTCTACCGGTACTGGTCCAGGTGCTTCT CCGGGTACTAGCTCTACTGGTTCTCCAGGTACCTC TACCGAACCGTCCGAGGGTAGCGCACCAGGTACC TCTACTGAACCGTCTGAGGGTAGCGCTCCAGGTAG CGAACCGGCAACCTCCGGTTCTGAAACTCCAGGTT CGTCTTCACTCGAGGGTACCGAACTTTTCACTGGA GTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTT AATGGGCACAAATTTTCTGTCAGTGGAGAGGGTG AAGGTGATGCAACATACGGAAAACTTACCCTTAA ATTTATTTGCACTACTGGAAAACTACCTGTTCCAT GGCCAACACTTGTCACTACTTTCTCTTATGGTGTTC AATGCTTTTCCCGTTATCCGGATCACATGAAACGG CATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTA TGTACAGGAACGCACTATATCTTTCAAAGATGACG GGAACTACAAGACGCGTGCTGAAGTCAAGTTTGA AGGTGATACCCTTGTTAATCGTATCGAGTTAAAAG GTATTGATTTTAAAGAAGATGGAAACATTCTCGGA CACAAACTCGAGTACAACTATAACTCACACAATGT ATACATCACGGCAGACAAACAAAAGAATGGAATC AAAGCTAACTTCAAAATTCGCCACAACATTGAAG ATGGATCCGTTCAACTAGCAGACCATTATCAACAA	721	MKGSPGEVQLQE SGGGLVQAGDSL RLSCLVSGRSFNS YTMGWFRQAPG KEREFVAAILWS GPTTYYADSVKG RFTISRDNAKNTV YLQMNSLKPEDT AVYYCAAALGV LVLAPGNVYSYW GQGTQVTVSSAH HGGTPGSGTASSS PGSSTPSGATGSP GSSTPSGATGSPG SPAGSPTSTEEGT SESATPESGPGTS TEPSEGSAPGSSP SASTGTGPGSSPS ASTGTGPGASPG TSSTGSPGTSTEP SEGSAPGTSTEPS EGSAPGSEPATSG SETPGSSSLEGTE LFTGVVPILVELD GDVNGHKFSVSG EGEGDATYGKLT LKFICTTGKLPVP WPTLVTTFSYGV QCFSRYPDHMKR HDFFKSAMPEGY VQERTISFKDDG NYKTRAEVKFEG DTLVNRIELKGID FKEDGNILGHKL EYNYNSHNVYIT ADKQKNGIKANF KIRHNIEDGSVQL ADHYQQNTPIGD GPVLLPDNHYLS TQSALSKDPNEK	746
	AATACTCCAATTGGCGATGGCCCTGTCCTTTTACC AGACAACCATTACCTGTCGACACAATCTGCCCTTT CGAAAGATCCCAACGAAAAGCGTGACCACATGGT CCTTCTTGAGTTTGTAACTGCTGCTGGGATTGGTG GCTCTCATCACCATCACCATCACCATCACTAA		RDHMVLLEFVTA AGIGGSHHHHHH HH	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
EGFR_V HH1-AM144-GFP6-22 9-H8, LMS109.045	ATGAAAGGGTCTCCAGGTGAAGTGCAACttcaAgAA TCTGGTGGTGGTCTGGTACAAGCTGGTGACTCTCT GCGCCTGTCTTGTCTGGTCTCCGGTCGTTCCTTCAA TAGCTATACCATGGGCTGGTTCCGCCAAGCGCCTG GCAAAGAGCGTGAATTCGTAGCAGCAATCCTTTGG TCCGGTCCAACTACCTACTATGCTGACTCTGTAAA AGGTCGCTTCACCATCTCCCGTGATAACGCCAAAA ACACTGTTTATCTACAAATGAATTCTCTGAAACCG GAGGACACCGGCTGTTTACTACTGTGCTGCCGCGCT GGGTGTACTGGTGCTCGCACCAGGTAATGTGTACA GCTATTGGGGTCAAGGTACCCAGGTGACGGTCAG CTCTGCGCATCATGGAGGTAGCCCGGCAGGCTCTC CGACCTCTACTGAGGAAGGTACTTCTGAAAGCGCA ACCCCGGAGTCCGGCCCAGGTACCTCTACCGAACC GTCTGAGGGCAGCGCACCAGGTACTTCTGAAAGC GCAACCCCTGAATCCGGTCCAGGTAGCGAACCGG CTACTTCTGGCTCTGAGACTCCAGGTACTTCTACC GAACCGTCCGAAGGTAGCGCACCAGGTAGCCCGG CTGGTTCTCCGACTTCCACCGAGGAAGGTACCTCT ACTGAACCTTCTGAGGGTAGCGCTCCAGGTACCTC TACTGAACCTTCCGAAGGCAGCGCTCCAGGTACTT CTACCGAACCGTCCGAGGGCAGCGCTCCAGGTAC TTCTACTGAACCTTCTGAAGGCAGCGCTCCAGGTA CTTCTACTGAACCTTCCGAAGGTAGCGCACCAGGT TCGTCTTCACTCGAGGGTACCGAACTTTTCACTGG AGTTGTCCCAATTCTTGTTGAATTAGATGGTGATG TTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGT GAAGGTGATGCAACATACGGAAAACTTACCCTTA AATTTATTTGCACTACTGGAAAACTACCTGTTCCA TGGCCAACACTTGTCACTACTTTCTCTTATGGTGTT CAATGCTTTTCCCGTTATCCGGATCACATGAAACG GCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTT ATGTACAGGAACGCACTATATCTTTCAAAGATGAC GGGAACTACAAGACGCGTGCTGAAGTCAAGTTTG AAGGTGATACCCTTGTTAATCGTATCGAGTTAAAA GGTATTGATTTTAAAGAAGATGGAAACATTCTCGG ACACAAACTCGAGTACAACTATAACTCACACAAT GTATACATCACGGCAGACAAACAAAAGAATGGAA TCAAAGCTAACTTCAAAATTCGCCACAACATTGAA GATGGATCCGTTCAACTAGCAGACCATTATCAACA AAATACTCCAATTGGCGATGGCCCTGTCCTTTTAC CAGACAACCATTACCTGTCGACACAATCTGCCCTT TCGAAAGATCCCAACGAAAAGCGTGACCACATGG TCCTTCTTGAGTTTGTAACTGCTGCTGGGATTGGTG GCTCTCATCACCATCACCATCACCATCACTAA	722	MKGSPGEVQLQE SGGGLVQAGDSL RLSCLVSGRSFNS YTMGWFRQAPG KEREFVAAILWS GPTTYYADSVKG RFTISRDNAKNTV YLQMNSLKPEDT AVYYCAAALGV LVLAPGNVYSYW GQGTQVTVSSAH HGGSPAGSPTSTE EGTSESATPESGP GTSTEPSEGSAPG TSESATPESGPGS EPATSGSETPGTS TEPSEGSAPGSPA GSPTSTEEGTSTE PSEGSAPGTSTEP SEGSAPGTSTEPS EGSAPGTSTEPSE GSAPGTSTEPSEG SAPGSSSLEGTEL FTGVVPILVELDG DVNGHKFSVSGE GEGDATYGKLTL KFICTTGKLPVPW PTLVTTFSYGVQ CFSRYPDHMKRH DFFKSAMPEGYV QERTISFKDDGN YKTRAEVKFEGD TLVNRIELKGIDF KEDGNILGHKLE YNYNSHNVYITA DKQKNGIKANFK IRHNIEDGSVQLA DHYQQNTPIGDG PVLLPDNHYLST QSALSKDPNEKR DHMVLLEFVTAA GIGGSHHHHHHH H	747

Ejemplo 26: Purificación de alL6R-XTEN

Se cultivaron *E. coli* de una única colonia que contenga: AC341, AC342, AC361 o AC362 hasta saturación en medios 2xYT. Se utilizaron 10 ml de este cultivo saturado durante la noche para inocular 500 ml de medios 2xYT, y este cultivo se cultivó hasta una DO600 de entre 0,4 y 0,5 a 37 °C, se transfirió a 26 °C y se indujo con IPTG 1 mM. A continuación, el cultivo se cultivó durante la noche (15-17 horas) y se recogió mediante centrifugación a 10 000 rpm en un rotor Sorvall SLA-3000. Los sedimentos se almacenaron hasta su uso a -80 °C. La pasta celular se resuspendió en 25 ml de tampón de lisis (acetato de sodio 20 mM, cloruro de sodio 50 mM, pH 4,5), se lisó mediante sonicación y se clarificó por rotación a 10 000 rpm a 4 °C en un rotor Sorvall SS34. Las muestras se acidificaron adicionalmente mediante la adición de ácido acético. La muestra se clarificó adicionalmente mediante centrifugación, y el sobrenadante se cargó en una columna DE52 de 15 ml equilibrada con acetato de sodio 20 mM, cloruro de sodio 50 mM, pH 4,5. La columna se lavó con 2 volúmenes de columna de acetato de sodio 20 mM, cloruro de sodio 50 mM, pH 4,5, se lavó con cuatro

volúmenes de columna de acetato de sodio 20 mM, cloruro de sodio 100 mM, pH 4,5, y se eluyó con cuatro volúmenes de columna de acetato de sodio 20 mM, cloruro de sodio 150 mM, pH 4,5. Se añadió sulfato de sodio a las fracciones de elución hasta una concentración final de 1 M, y la muestra se cargó en una columna fenil HIC. La columna se lavó con cinco volúmenes de columna de acetato de sodio 20 mM, sulfato de sodio 1 M a pH 4,5, y se eluyó con cuatro volúmenes de columna de acetato de sodio 20 mM, sulfato de sodio 0,5 M a pH 4,5. Las muestras se intercambiaron con tampón de ensayo, se les asignaron los números de lote AP342, AP343, AP344 y AP345 y se almacenaron congeladas a -80 °C.

Ejemplo 27: Purificación de aHER2-XTEN-GFP

Se cultivó AC62 hasta saturación durante la noche en 2xYT + kanamicina. Se inocularon dos matraces de 500 ml de 2xYT con 3 ml de este saturado durante la noche y se cultivaron hasta una DO600 de ~0,8 a 37 °C. El cultivo se llevó a 25 °C y a continuación se indujo con IPTG 1 mM. El cultivo se indujo durante la noche. El sedimento celular se recogió por centrifugación a 5000 rpm, en un Sorvall SLA-3000 a 4 °C como el sedimento EP52. A continuación, el sedimento celular se resuspendió en Tris 30 mM a pH 8,0, NaCl 500 mM, imidazol 20 mM. Las células se lisaron mediante sonicación y a continuación se clarificaron mediante centrifugación a 15 000 rpm en un rotor Sorvall SS34. El sobrenadante se cargó en una columna de Ni-NTA de 5 ml, se lavó con 20 volúmenes de columna de Tris 30 mM a pH 8,0, NaCl 500 mM, imidazol 20 mM, y se eluyó con Tris 30 mM a pH 8,0, NaCl 500 mM, imidazol 600 mM. Las fracciones de elución se diluyeron a 1:2 con histidina 20 mM a pH 5,6 y se cargaron en una columna de intercambio aniónico equilibrada con histidina 20 mM a pH 5,6. La columna se lavó con histidina 20 mM a pH 5,6, 20 volúmenes de columna de histidina 20 mM a pH 5,6, NaCl 150 mM, 20 volúmenes de columna de histidina 20 mM a pH 5,6, NaCl 300 mM, y a continuación se eluyó con histidina 20 mM a pH 5,6, NaCl 600 mM. A esta proteína se le asignó el n.º de lote AP60 y se almacenó congelada.

Ejemplo 28: Purificación de aCD3-XTEN-GFP

Se cultivó AC50 hasta saturación durante la noche en 2xYT + kanamicina. Se inoculó un matraz de 500 ml de 2xYT con 3 ml de este saturado durante la noche y se cultivaron hasta una DO600 de ~-0,8 a 37 °C. El cultivo se llevó a 25 °C y a continuación se indujo con IPTG 0,2 mM. El cultivo se indujo durante la noche. El sedimento celular se recogió mediante centrifugación a 5000 rpm en un Sorvall SLA-3000 a 4 °C. A continuación, el sedimento celular se resuspendió en Tris 30 mM a pH 8,0, NaCl 500 mM, imidazol 20 mM. Las células se lisaron mediante sonicación y a continuación se clarificaron mediante centrifugación a 15 000 rpm en un rotor Sorvall SS34. El sobrenadante se cargó en una columna de Ni-NTA, se lavó con 10 volúmenes de columna de Tris 30 mM a pH 8,0, NaCl 500 mM, imidazol 20 mM, y se eluyó con Tris 30 mM a pH 8,0, NaCl 500 mM, imidazol 600 mM. A esta proteína se le asignó el n.º de lote AP43.

Ejemplo 29: Purificación de aHER2-XTEN-aEGFR

Se cultivó AC49 hasta saturación durante la noche en 2xYT + kanamicina. Se inoculó un matraz de 500 ml de 2xYT con 3 ml de este saturado durante la noche y se cultivaron hasta una DO600 de ~0,8 a 37 °C. El cultivo se llevó a 25 °C y a continuación se indujo con IPTG 0,2 mM. El cultivo se indujo durante la noche. El sedimento celular se recogió por centrifugación a 5000 rpm, en un Sorvall SLA-3000 a 4 °C como el sedimento EP36. A continuación, el sedimento celular se resuspendió en Tris 30 mM a pH 8,0, NaCl 500 mM, imidazol 20 mM. Las células se lisaron mediante sonicación y a continuación se clarificaron mediante centrifugación a 15 000 rpm en un rotor Sorvall SS34. El sobrenadante se cargó en una columna de Ni-NTA de 5 ml, se lavó con 10 volúmenes de columna de Tris 30 mM a pH 8,0, NaCl 500 mM, imidazol 20 mM, y se eluyó con Tris 30 mM a pH 8,0, NaCl 500 mM, imidazol 600 mM. Las fracciones de elución se diluyeron a 1:3 con histidina 20 mM a pH 5,6 y se cargaron en una columna con DEAE de 1 ml equilibrada con histidina 20 mM a pH 5,6, NaCl 200 mM. La columna se lavó con cinco volúmenes de columna de histidina 20 mM a pH 5,6, NaCl 400 mM y cinco volúmenes de columna de histidina 20 mM a pH 5,6, NaCl 600 mM, y a continuación se eluyó con histidina 20 mM a pH 5,6, NaCl 700 mM. A esta proteína se le asignó el n.º de lote AP41. La proteína aHER2-XTEN-aEGFR purificada final se sometió a una SDS-PAGE no reductora utilizando gel Bis-Tris NuPAGE al 4-12 % de Invitrogen según las especificaciones del fabricante. Los resultados, como se muestra en el carril 3 de la FIG. 17B (en comparación con aHER2-XTEN-aHER2-XTEN en el carril 2) demuestran que la proteína se recuperó con una pureza >95 % mediante el proceso detallado anteriormente. Los resultados del análisis mediante SEC del material purificado, realizado como se describe en el ejemplo 36, se muestran en la FIG. 17C, y demuestran que la proteína de fusión de unión se recuperó en forma monómera y tiene un peso molecular aparente de aproximadamente 500 kDa, muy superior a su peso molecular real.

Ejemplo 30: Purificación de aCD3-XTEN-aHER2

Se cultivó AC59 hasta saturación durante la noche en 2xYT + kanamicina. Se inoculó un cultivo de 10 l de 2xYT en bolsas de balanceo con este saturado durante la noche y se cultivaron hasta una DO600 de 1,3 a 37 °C. El cultivo se llevó a 25 °C y a continuación se indujo con IPTG 1 mM. El cultivo se indujo durante la noche. El sedimento celular se recogió por centrifugación a 5000 rpm, en un Sorvall SLA-3000 a 4 °C y se almacenó como el sedimento EP50. A continuación, el sedimento celular se resuspendió en Tris 20 mM a pH 8,0, NaCl 50 mM, más inhibidores de proteasa completos de Roche. Las células se lisaron mediante sonicación y a continuación se clarificaron mediante

centrifugación a 15 000 rpm en un rotor Sorvall SSR-15. El sobrenadante se cargó en una columna de Ni-NTA de 60 ml, se lavó con 11 volúmenes de columna de Tris 20 mM a pH 8,0, NaCl 50 mM; 10 volúmenes de columna de Tris 20 mM a pH 8,0, NaCl 500 mM; 10 volúmenes de columna de Tris 20 mM a pH 8,0, triton X-114 al 1 %; 10 volúmenes de columna de Tris 20 mM a pH 8,0, imidazol 5 mM; 10 volúmenes de columna de Tris 20 mM a pH 8,0, NaCl 50 mM, imidazol 10 mM, y se eluyó con Tris 20 mM a pH 8,0, NaCl 50 mM, imidazol 200 mM. Las fracciones de elución se diluyeron con histidina 20 mM a pH 6,2, y se extrajeron con triton-X114 para eliminar la endotoxina de la siguiente manera: llevar hasta un 5 % de detergente en hielo, calentar hasta 37 °C hasta vapor, separar la fase por centrifugación a la temperatura ambiente a 3000 rpm en una centrifugadora de mesa de Sorvall, transferir la fase acuosa superior, repetir el proceso 3 veces. La capa acuosa final se diluyó con histidina 20 mM a pH 6,2 y se cargó en una columna Q-sepharose FF de 7,5 ml equilibrada con histidina 20 mM a pH 6,2, NaCl 20 mM. La columna se lavó con 11 volúmenes de columna de histidina 20 mM a pH 6,2, NaCl 20 mM, 10 volúmenes de columna de histidina 20 mM a pH 6,2, NaCl 50 mM y 10 volúmenes de columna de histidina 20 mM a pH 6,2, NaCl 100 mM, a continuación se eluyó con histidina 20 mM a pH 5,6, NaCl 600 mM. A continuación, la proteína se purificó adicionalmente en una columna mono-Q en un sistema purificador AKTA. La columna y el sistema se higienizaron con 10 volúmenes de columna de NaOH 0,5 y la proteína se cargó en histidina 20 mM a pH 6,0 y se eluyó con un gradiente salino lineal. Las fracciones viables se agruparon y se extrajeron con triton como anteriormente, para reducir aún más la endotoxina. A esta proteína se le asignó el n.º de lote AP58.

Ejemplo 31: Purificación de aHER2-XTEN-aHER2

Se cultivó AC47 hasta saturación durante la noche en 2xYT + kanamicina. Se inoculó un matraz de 500 ml de 2xYT con 3 ml de este saturado durante la noche y se cultivaron hasta una DO600 de ~0,8 a 37 °C. El cultivo se llevó a 25 °C y a continuación se indujo con IPTG 1 mM. El cultivo se indujo durante la noche. El sedimento celular se recogió por centrifugación a 5000 rpm, en un Sorvall SLA-3000 a 4 °C y se almacenó como EP35. A continuación, el sedimento celular se resuspendió en Tris 30 mM a pH 8,0, NaCl 500 mM, imidazol 20 mM. Las células se lisaron mediante sonicación y a continuación se clarificaron mediante centrifugación a 15 000 rpm en un rotor Sorvall SS34. El sobrenadante se cargó en una columna de Ni-NTA de 5 ml, se lavó con 10 volúmenes de columna de Tris 30 mM a pH 8,0, NaCl 500 mM, imidazol 20 mM, y se eluyó con Tris 30 mM a pH 8,0, NaCl 500 mM, imidazol 600 mM. Las fracciones de elución se diluyeron a 1:3 con histidina 20 mM a pH 5,6 y se cargaron en una columna con DEAE de 1 ml equilibrada con histidina 20 mM a pH 5,6, NaCl 200 mM. La columna se lavó con cinco volúmenes de columna de histidina 20 mM a pH 5,6, NaCl 400 mM y cinco volúmenes de columna de histidina 20 mM a pH 5,6, NaCl 600 mM, y a continuación se eluyó con histidina 20 mM a pH 5,6, NaCl 700 mM. A esta proteína se le asignó el n.º de lote AP40.

Ejemplo 32: Purificación de las proteínas de unión Vhh del aEGFR multivalentes

Las proteínas de las construcciones LCW501.001, LCW502.009, LCW503.004, LCW504.004 y LCW505.002 se expresaron de la siguiente manera: Se cultivaron nueve placas de 96 pocillos con 0,5 ml en cada pocillo para cada construcción. Las placas se llenaron inoculando 40 ml de saturado durante la noche en SB en 1 l de medios de autoinducción y a continuación utilizando el Q-fill para cargar las placas. Las muestras se sellaron con una membrana transpirable y a continuación se colocaron a 26 C agitando durante la noche a 300 rpm. A continuación, las placas se agruparon en un cubo grande y se vertieron en un único frasco de centrífuga para la recolección. Los cultivos se rotaron a 10 000 rpm en la centrífuga RC-5B de Sorvall durante 20 minutos (rotor GS-3) para sedimentar las células. Se obtuvieron aproximadamente 15 g de masa de células. Las células se resuspendieron en tampón de lisis (fosfato 20 mM a pH 7,4, NaCl 500 mM, imidazol 5 mM, EDTA 0,5 mM y 1 mini comprimido de proteasa Roche completo por 10 ml). A continuación, las muestras se lisaron mediante sonicación (5x 20 segundos encendido, 40 segundos apagado al ajuste de potencia 8 con la sonda grande). Las muestras se mantuvieron en hielo, pero aun así se calentaron hasta -25 durante el proceso. A continuación, los lisados se rotaron a 15 000 rpm en el rotor SS-34 utilizando una centrífuga RC-5B de Sorvall para clarificarlos.

Se equilibraron columnas de Ni-NTA de 2 ml con 10 VC de tampón de lisis, y a continuación se cargaron los lisados de LCW503, LCW504 y LCW505. Las columnas se lavaron con 10 VC de tampón de lisis, 3 VC de PBS + imidazol 5 mM para reducir la sal, y a continuación se eluyeron con PBS + imidazol 300 mM. La elución se siguió a través de la GFP, de modo que se capturó toda la fluorescencia en la fracción. Las eluciones se llevaron a cabo durante la noche. Al día siguiente se equilibraron columnas de Ni-NTA de 3x3 ml (LCW503, LCW504, LCW505) y de 2x5 ml (LCW501 y LCW502) con 10 VC de tampón de lisis, y a continuación se cargaron los lisados de LCW503, LCW504 y LCW505. Las columnas se lavaron con 50 ml de tampón de lisis, 15 ml de PBS + imidazol 5 mM para reducir la sal, y a continuación se eluyeron con PBS + imidazol 300 mM. La elución se siguió a través de la GFP, de modo que se capturó toda la fluorescencia en la fracción. Las eluciones del primer día de 503, 504 y 505 se agruparon con las eluciones del segundo día.

A continuación, las muestras se purificaron mediante una cromatografía de intercambio aniónico. Se usó una columna mono Q de 1 ml para todos los preps. El tampón A fue siempre Tris 20 mM a pH 8,0 y el tampón B fue siempre Tris 20 mM a pH 8,0 y NaCl 1 M. Se utilizó el sistema purificador AKTA con un súper bucle de 10 ml. Todos los desarrollos implicaron una carga con inyección de 12 ml (independientemente del volumen de muestra), un lavado con 2 VC de tampón A, un gradiente hasta el 50 % de B sobre 20 VC y, finalmente, un lavado con 5 VC con un 100 % de B. La columna se volvió a equilibrar a continuación con tampón A entre los desarrollos. El súper bucle se desmontó y se

limpió, y la tubuladura se lavó para evitar la contaminación por arrastre entre desarrollos. Las fracciones se agruparon y se almacenaron durante la noche a 4C. Al día siguiente, las muestras se concentraron en un concentrador Amicon ultra de 30 000 MWCO a 4000 rpm en el Sorvall T21. La muestra concentrada se cargó en columnas de filtración en gel. Las columnas fueron TSK3000 para LCW501 y LCW502, y TSK4000 para LCW503, LCW504 y LCW505. Las fracciones se agruparon, se concentraron y se congelaron para almacenarlas a -80 °C.

Ejemplo 33: Caracterización de aIL6R-XTEN

Se cultivaron *E. coli* de una única colonia de AC342 hasta saturación en medios 2xYT. Se utilizaron 10 ml de este cultivo saturado durante la noche para inocular 500 ml de medios 2xYT, y este cultivo se cultivó hasta una DO600 de entre 0,4 y 0,5 a 37 °C, se transfirió a 26 °C y se indujo con IPTG 1 mM. A continuación, el cultivo se cultivó durante la noche (15-17 horas) y se recogió mediante centrifugación a 10 000 rpm en un rotor Sorvall SLA-3000. Los sedimentos se almacenaron hasta su uso a -80 °C. La pasta celular se resuspendió en 25 ml de tampón de lisis (acetato de sodio 20 mM, cloruro de sodio 50 mM, pH 4,5), se lisó mediante sonicación y se clarificó por rotación a 10 000 rpm a 4 °C en un rotor Sorvall SS34. Las muestras se acidificaron adicionalmente mediante la adición de ácido acético. La muestra se clarificó adicionalmente mediante centrifugación, y el sobrenadante se cargó en una columna DE52 de 15 ml equilibrada con acetato de sodio 20 mM, cloruro de sodio 50 mM, pH 4,5. La columna se lavó con 2 volúmenes de columna de acetato de sodio 20 mM, cloruro de sodio 50 mM, pH 4,5, se lavó con cuatro volúmenes de columna de acetato de sodio 20 mM, cloruro de sodio 100 mM, pH 4,5, y se eluyó con cuatro volúmenes de columna de acetato de sodio 20 mM, cloruro de sodio 150 mM, pH 4,5. La uniformidad se evaluó mediante una SEC, que mostró un máximo monodisperso con una contaminación mínima (FIG. 28A). La unión al IL6R se evaluó mediante un ELISA de la siguiente manera: las placas se recubrieron con 100 ng/pocillo de IL6R humano durante 1 hora a 37 °C, se bloquearon con BSA al 3 % en PBS durante una hora a 37 °C, se unieron mediante la adición de una serie de diluciones de AC342 a la placa y se incubaron durante una hora a la temperatura ambiente, se lavaron 3 veces con PBST, se detectaron con anticuerpo anti-XTEN biotinilado durante una hora a la temperatura ambiente, se lavaron tres veces con PBST, se amplificaron con estreptavidina-HRP durante una hora a la temperatura ambiente, se lavaron tres veces con PBST, se revelaron con sustrato de TMB y se leyeron a 405 nm utilizando un lector de placas. Los valores del ELISA se representaron frente a la concentración (escala logarítmica), lo que proporcionó la determinación de un valor de la CE50 de 5 nM (FIG. 28B).

Ejemplo 34: Caracterización de aCD40-XTEN

Se expresaron dos construcciones de aCD40-XTEN, AC384 y AC385, cultivando un cultivo hasta saturación en 2xYT y utilizando a continuación este cultivo para inocular un matraz de 500 ml de 2xYT. Este segundo matraz se cultivó hasta una D600 de entre 0,6 y 1,0 a 37 °C, se transfirió a 26 °C y se indujo con IPTG 1 mM. La inducción se dejó durante la noche y la pasta de células se recogió mediante centrifugación la mañana que fluía. Estos sedimentos se almacenaron como EP220 y EP221. Se tomaron pequeñas muestras del sedimento, se lisaron utilizando bugbuster y se clarificaron utilizando una microcentrífuga. Estos lisados clarificados se diluyeron en serie por un factor de tres en PBST, para usarlos en un ELISA de unión. El ELISA se realizó de la siguiente manera: Las placas se recubrieron con 100 ng/pocillo de CD40-fusión Fc humano durante 1 hora a 37 °C, se bloquearon con BSA al 3 % en PBS durante una hora a 37 °C, se unieron añadiendo una serie de diluciones de AC384 o AC385 a la placa y se incubaron durante una hora a la temperatura ambiente, se lavaron 3 veces con PBST, se detectaron con anticuerpo anti-XTEN biotinilado durante una hora a la temperatura ambiente, se lavaron tres veces con PBST, se amplificaron con estreptavidina-HRP durante una hora a la temperatura ambiente, se lavaron tres veces con PBST, se revelaron con sustrato de TMB y se leyeron a 405 nm utilizando un lector de placas. Los valores del ELISA se representaron gráficamente frente a la dilución de lisado representada gráficamente en una escala logarítmica, y se determinó un valor de la CE50, con unos resultados casi idénticos (FIG. 27). Combinando esto con una concentración estimada de 10 µm en el lisado, las construcciones aCD40-XTEN tienen una Kd de aproximadamente 50 nM para el CD40 humano.

Ejemplo 35: Caracterización de aHER2-XTEN-aHER2

La unión de la proteína aHER2-XTEN-aHER2 al HER2 humano se evaluó mediante un ELISA de la siguiente manera: Las placas se recubrieron con 100 ng/pocillo de HER2-fusión Fc humano durante la noche a 4 °C, se bloquearon con BSA al 1 % en PBS durante una hora a la temperatura ambiente, se unieron añadiendo una serie de diluciones de aHER2-XTEN-aHER2 a la placa e incubando durante una hora a la temperatura ambiente, se lavaron 3 veces con PBST, se detectaron con anticuerpo anti-HA conjugado con HRP durante una hora a la temperatura ambiente, lavado tres veces con PBST, se lavaron tres veces con PBST, se revelaron con sustrato ABTS y se leyeron a 405 nm utilizando un lector de placas. Los valores del ELISA se representaron gráficamente frente a la concentración representada gráficamente en una escala logarítmica, y se determinó una CE50 de 0,5 nM (FIG. 18).

Ejemplo 36: Caracterización de aCD3-XTEN-aHER2

Para confirmar la especificidad de unión de las proteínas de fusión de unión expresadas y purificadas a restos de direccionamiento scFv, se evaluaron dos preparaciones dirigidas a las dianas HER2 o CD3. Las proteínas de fusión de unión de aHER2 unidas al extremo amínico de XTEN, con la GFP unida al extremo carboxílico del XTEN ("aHER2-XTEN-GFP"), y el aCD3 unido al extremo amínico de XTEN, con la GFP unida al extremo carboxílico del XTEN ("aCD3-

XTEN-GFP"), fueron los productos experimentales cuya capacidad para unirse a sus respectivas dianas se evaluó en células Jurkat portadoras de CD3 y células SK-BR-3 portadoras de HER2. Las células Jurkat y las células SK-BR3 se incubaron con las proteínas de fusión scFv-XTEN-GFP indicadas. Las aCD3-XTEN-GFP y aHER2-XTEN-GFP unidas se detectaron con un anticuerpo policlonal anti-GFP específico para la GFP fusionada con la construcción, y un anticuerpo secundario conjugado con FITC apropiado, utilizando citometría de flujo. Los resultados de la citometría de flujo, mostrados gráficamente en la FIG. 19B, demuestran la unión específica de aCD3-XTEN-GFP a las células Jurkat y la unión específica de aHER2-XTEN-GFP a las células SK-BR-3, respectivamente, mientras que no se detectó unión no específica en la configuración opuesta, lo que indica la falta de reactividad cruzada por parte de las construcciones.

Se recuperó una proteína de fusión de unión biespecífica con restos de direccionamiento scFv tanto a HER2 como a CD3 ("aCD3-XTEN-aHER2", lo que coincide con la configuración de extremo amínico a extremo carboxílico de los componentes de la proteína de fusión) mediante el proceso de purificación detallado anteriormente, y la proteína se caracterizó. La proteína aCD3-XTEN-aHER2 se sometió a una SDS-PAGE no reductora utilizando gel Bis-Tris NuPAGE al 4-12 % de Invitrogen según las especificaciones del fabricante. Los resultados, mostrados en la FIG. 20A, demuestran que la proteína era >95 % pura, según se analizó mediante la SDS-PAGE. La FIG. 20B muestra el resultado de un análisis de cromatografía de exclusión por tamaño (SEC) de la aHER2-aCD3-XTEN en comparación con los patrones de peso molecular, y demuestra que no se forman dímeros ni otros oligómeros de orden superior y que la proteína tiene un peso molecular aparente aproximado de aproximadamente 500 kDa, aproximadamente cinco veces mayor que el derivado del ensayo de la SDS-PAGE.

Se realizó un ensayo de ELISA para detectar la unión directa de la aCD3-XTEN-aHER2 a pocillos recubiertos con HER2-Fc de la siguiente manera. El dominio extracelular de HER2 fusionado con el dominio Fc de la IgG humana se recubrió en los pocillos de una placa de microtitulación. A continuación, se aplicó una serie de diluciones de aCD3-XTEN-aHER2 a los pocillos recubiertos. Después de dos horas, la aCD3-XTEN-aHer2 no unida se eliminó por lavado y cualquier proteína unida se detectó con un anticuerpo anti-HA conjugado con HRP. Los resultados de los ensayos de unión se muestran en la FIG. 21A. La proteína de fusión de control con aHER2 aminoterminal se une a HER2-Fc con una Kd aparente de 80 pM, mientras que la fusión carboxiterminal de aCD3-XTEN-aHER2 se une con una Kd aparente de 3,4 nM. Por lo tanto, el componente anti-Her2 de la construcción aCD3-XTEN-aHER2 todavía conserva una buena afinidad de unión, aunque es menor que la construcción aminoterminal.

El análisis por citometría de flujo de aHER2-XTEN (FIG. 19) y aCD3-XTEN-aHER2 (FIG. 21C) muestra la unión a las células SK-BR-3 que expresan el HER2. La detección se realizó a través de un anticuerpo anti-HA y un anticuerpo secundario apropiado. La unión de aCD3-XTEN-GFP a las células Jurkat positivas para CD3 se midió mediante una citometría de flujo, utilizando la detección con un anticuerpo anti-GFP, en presencia de un exceso molar de 10 veces de aHER2-aCD3-XTEN. Los resultados demuestran claramente que el exceso de aCD3-XTEN-aHER2 biespecífica desplaza competitivamente a la proteína aCD3-XTEN-GFP monoespecífica y elimina el desplazamiento observado de la MFI (FIG. 22). Por lo tanto, hemos genomodificado una proteína de fusión de unión biespecífica aCD3-XTEN-aHER2 que interactúa con células tumorales que expresan el HER2 y linfocitos T positivos para CD3.

Para demostrar la capacidad de la proteína de fusión de unión aCD3-XTEN-aHER2 para dirigirse y destruir células cancerosas, se realizó un ensayo utilizando la construcción para dirigir la actividad citotóxica de los linfocitos T hacia células tumorales. Aquí, ensayamos la eficacia del candidato aCD3-XTEN-aHER2 utilizando células de melanoma M21 y de cáncer de mama SK-BR-3. Las células tumorales, que expresan el receptor diana, se marcaron con un fluoróforo permeable a la membrana, $DiOC_{18}$. A continuación, se mezclaron monomorfonucleares periféricos (PBMC) con las células tumorales marcadas (relación célula efectora:célula tumoral 10:1) y la proteína de fusión de unión aCD3-XTEN-aHER2. Después de 24 horas, las suspensiones celulares se recogieron mediante centrifugación y las células muertas se marcaron con yoduro de propidio. Las células se analizaron en un microscopio Zeiss Axiovert 100. Se capturaron imágenes de las células diana ($DiOC_{18}$) y de las muertas (yoduro de propidio) con una cámara SPOT CCD utilizando los juegos de filtros de excitación y de emisión apropiados. El análisis de las imágenes se realizó utilizando el programa ImageJ. El número de células diana por imagen se determinó utilizando el comando Analizar partículas utilizando la imagen del canal de fluorescencia verde ($DiOC_{18}$). La fracción de estas células que estaban muertas se determinó solapando el canal de fluorescencia rojo (yoduro de propidio) utilizando el plugin de colocalización y contando el número de células verdes que también eran rojas. La proporción de células tumorales muertas, en función de la concentración de aHER2-aCD3-XTEN, se muestra en la FIG. 23. Se ajustó un modelo lineal general con una función de enlace logit a los datos, con la concentración de la proteína de fusión aCD3-XTEN-aHER2 como variable independiente. El modelo encontró que la concentración de aCD3-XTEN-aHER2 es un factor predictivo significativo de muerte de las células tumorales a un valor p de <0,001 para las células diana tanto M21 como SK-BR-3. Los resultados respaldan la conclusión de que las proteínas de fusión de unión con restos de direccionamiento dirigidos a antígenos asociados a tumores pueden tener actividad contra células tumorales.

Ejemplo 37: Caracterización de aHER2-XTEN-aEGFR

Se realizó una cromatografía de exclusión por tamaño de la aHER2-XTEN-aEGFR recuperada para evaluar sus características monómeras. Los resultados del análisis de la SEC demuestran que la aHER2-XTEN-aEGFR (FIG. 17C) no dimerizó ni formó otros oligómeros de orden superior. Como se muestra en la FIG. 17C, el peso molecular aparente es aproximadamente cinco veces mayor que su masa calculada (aproximadamente 107 kDa) o la masa estimada a

partir de la SDS-PAGE de la FIG. 17B, debido al componente XTEN; aquí, el volumen de elución de aHER2-XTEN-aEGFR se compara con un conjunto de patrones de peso molecular globular (de izquierda a derecha: 670, 158, 44, 17 y 1.4 kDa). La actividad de unión de la aHER2-XTEN-aEGFR a sus respectivas dianas se ensayó en un formato de ELISA. Se recubrieron los dominios extracelulares de HER2 o EGFR fusionados con el dominio Fc de la IgG humana en los pocillos de una placa de microtitulación. A continuación, se aplicó una serie de diluciones de aHER2-XTEN-aEGFR a los pocillos recubiertos. Después de dos horas, la aHER2-XTEN-aEGFR no unida se eliminó por lavado y cualquier proteína unida se detectó con un anticuerpo Anti-HA conjugado con HRP. El resultado indica que la construcción BPXTEN bifuncional era capaz de unirse a cada diana a unas concentraciones esencialmente equivalentes. (FIG. 18).

Ejemplo 38: Caracterización de proteínas de fusión de unión al Vhh del aEGFR multivalentes

La caracterización de proteínas de fusión de unión al dominio se realizó con construcciones del Vhh aEGFR-XTEN que tenían un número creciente de restos de direccionamiento, como se representa esquemáticamente en la FIG. 24A. Se realizaron SDS-PAGE y SEC con cuatro construcciones; LCW501.001, LCW502.009, LCW503.004 y LCW504.004. Un claro escalonamiento entre las construcciones en el análisis de SDS-PAGE, mostrado en la FIG. 24B, demuestra la longitud creciente de la proteína a medida que aumenta el número de dominios de unión anti-EGFR. Esto se refleja en la SEC que, en la FIG. 25, muestra un aumento en el radio hidrodinámico de las proteínas a medida que aumenta el número de dominios anti-EGFR. Los valores derivados para el factor de peso molecular aparente, el factor de peso molecular aparente y los radios hidrodinámicos se presentan en la tabla 26. Obsérvese que estos radios hidrodinámicos son característicamente grandes para una proteína de este peso molecular debido a la naturaleza no estructurada de los dominios del conector XTEN, lo que da como resultado un aumento del peso molecular aparente para cada una de las proteínas de fusión de unión al dominio.

También se caracterizaron las LCW501.001, LCW502.009, LCW503.004 y LCW504.004 en dos experimentos de ELISA diferentes. En el primero, se recubrieron placas Costar 3690 con 100 ng de EGFR-Fc en 50 µl de PBS durante la noche a 4 °C, se bloquearon con BSA al 3 % en PBS durante 1 hora, se unieron con proteínas de unión Vhh del aEGFR multivalentes desde 1 µm hacia abajo utilizando diluciones de 5 veces en tampón de unión (BSA al 1 % en PBST) durante 2 horas a la TA con agitación a 80 rpm, se lavaron tres veces con PBST, se detectaron con una dilución 1:5000 de anti-GFP-HRP de cabra durante 1 hora a la temperatura ambiente con agitación a 80 rpm, se lavaron cinco veces con PBST, se revelaron con sustrato ABTS-H_2O_2 y se leyeron a 405 nm. Se determinaron las curvas de unión para cada una de las diferentes especies de multímeros. Todos los multímeros de orden superior mostraron una unión más apretada que el monómero, lo que indica un efecto de avidez y que múltiples sitios de unión estaban activos en las moléculas multivalentes (FIG. 26A). La segunda caracterización basada en ELISA fue como sigue: La placa Costar 3690 se recubrió con 200 ng de anti-GFP de cabra en 50 µl de PBS durante la noche a 4 °C, se bloqueó con BSA al 3 % en PBS durante 1 hora, se unió con proteínas de unión Vhh de aEGFR multivalentes a partir de 3 µm hacia abajo utilizando diluciones de 5 veces en tampón de unión (BSA al 1 % en PBST) durante 2 horas a la TA con agitación a 80 rpm, se lavaron tres veces con PBST, se detectaron con 100 ng/pocillo (~20 nM) de EGFR-Fc biotinilado durante 1 hora a la temperatura ambiente con agitación a 80 rpm, se lavaron tres veces con PBST, se amplificaron con una dilución a 1:5000 de estreptavidina-HRP en tampón de unión y ser revelaron con sustrato ABTS-H2O2, y se leyeron a 405 nm. Los multímeros de mayor orden mostraron una capacidad de unión al EGFR-Fc aumentada (FIG. 26B), lo que indica que el aumento del número de módulos de unión al EGFR en la proteína aumenta el número de sitios de unión disponibles dentro de la proteína. Estos datos sugieren que todos los sitios de unión de las proteínas de fusión de unión, incluidos los tetrámeros y hexámeros de orden superior, son activos y capaces de unirse al EGFR.

Ejemplo 39: Purificación de aHER2-XTEN

Se cultivó AC51 hasta saturación durante la noche en 2xYT + kanamicina. Se inoculó un matraz de 500 ml de 2xYT con 3 ml de este saturado durante la noche y se cultivaron hasta una DO600 de ~0,8 a 37 °C. El cultivo se llevó a 25 °C y a continuación se indujo con IPTG 0,2 mM. El cultivo se indujo durante la noche. El sedimento celular se recogió mediante centrifugación a 5000 rpm en un Sorvall SLA-3000 a 4 °C. A continuación, el sedimento celular se resuspendió en Tris 30 mM a pH 8,0, NaCl 500 mM, imidazol 20 mM. Las células se lisaron mediante sonicación y a continuación se clarificaron mediante centrifugación a 15 000 rpm en una centrífuga Sorvall. El sobrenadante se cargó en una columna de Ni-NTA, se lavó con 10 volúmenes de columna de Tris 30 mM a pH 8,0, NaCl 500 mM, imidazol 20 mM, y se eluyó con Tris 30 mM a pH 8,0, NaCl 500 mM, imidazol 600 mM. A esta proteína se le asignó el n.º de lote AP44.

Ejemplo 40: Cromatografía analítica de exclusión por tamaño de las proteínas de fusión XTEN

Se realizó un análisis cromatográfico de exclusión por tamaño de las proteínas de fusión de unión que contenían diferentes restos de direccionamiento y proteínas recombinantes no estructuradas de longitud variable para determinar el efecto sobre el XTEN del aumento del peso molecular aparente. Un ensayo ilustrativo utilizó una columna TSKGel-G4000 SWXL (7,8 mm x 30 cm) en la que se separaron 40 µg de proteína de fusión de glucagón purificada a una concentración de 1 mg/ml a un caudal de 0,6 ml/min en fosfato 20 mM a pH 6,8, NaCl 114 mM. Los perfiles cromatográficos se monitorizaron utilizando una DO214 nm y una DO 280nm. La calibración de columna para todos los ensayos se realizó utilizando un patrón de calibración de exclusión por tamaño de BioRad; los marcadores incluyen

tiroglobulina (670 kDa), gammaglobulina bovina (158 kDa), ovoalbúmina de pollo (44 kDa), mioglobina equina (17 kDa) y vitamina B12 (1,35 kDa). Los perfiles representativos de la SEC de las proteínas de fusión de unión se muestran como superposiciones en las FIG. 17C, 20 y 25. Tomando como base los análisis de la SEC para todas las construcciones evaluadas, los factores de peso molecular aparente, el factor de peso molecular aparente (expresado como la relación entre el factor de peso molecular aparente y el peso molecular calculado) y el radio hidrodinámico (R_H en nm) se muestran en la tabla 26. Los datos muestran que el peso molecular aparente de cada compuesto es proporcional a la longitud de la secuencia del XTEN unida. Esto es particularmente evidente en el caso de las construcciones Vhh del EGFR que incluían unidades crecientes de EGRF_VHH-AM144_XTEN, pasando de monómero a dímero a tetrámero a hexámero, y mostraron aumentos en el factor de peso molecular aparente a medida que la longitud acumulada del XTEN aumentó con cada adición de los conectores AM144. Los datos también muestran que el peso molecular aparente de cada construcción es significativamente mayor de lo esperado para una proteína globular (como se muestra mediante la comparación con las proteínas patrón desarrolladas en el mismo ensayo). Adicionalmente, la incorporación de compañeros de fusión XTEN con 244 aminoácidos en total o más en proteínas de fusión con restos de direccionamiento dio como resultado un radio hidrodinámico de 7 nm o más; mucho más allá del tamaño de poro glomerular de aproximadamente 3-5 nm. Por consiguiente, se concluye que las proteínas de fusión de unión que comprenden restos de direccionamiento y XTEN tendrían un aclaramiento renal reducido, lo que contribuiría a una semivida terminal aumentada y mejoraría el efecto terapéutico o biológico con respecto a las correspondientes proteínas con restos de direccionamiento no fusionadas.

Tabla 26: Análisis mediante SEC de diferentes polipéptidos

Nombre de la construcción	XTEN o compañero de fusión	Proteína terapéutica	PM real (kDa)	PM aparente (kDa)	Factor de peso molecular aparente	R_H (nm)
XTEN_AE912-aIL6R scFv	AE912	Anti-IL6R	111	883	7,9	8,7
aHER2-XTEN_Y288-aHER2	Y288	Anti-HER2	83	419	5,0	7,4
Monómero EGFR_Vhh-XTEN	AM144	Anti-EGFR	54	107	2	4,9
Dímero EGFR_Vhh-XTEN	AM144 X2	Anti-EGFR	81	329	4,1	7,0
Tetrámero EGFR_Vhh-XTEN	AM144 X4	Anti-EGFR	135	1022	7,5	9,1
Hexámero EGFR_Vhh-XTEN	AM144 X6	Anti-EGFR	189	1802	9,5	10,1

Ejemplo 41: Creación, purificación y caracterización de conjugados de proteína de fusión de unión-construcciones de aHer2-XTEN con fluoróforo AF680 conjugado

Para ensayar la viabilidad y la utilidad de los conjugados de proteína de fusión de unión-fármaco, las proteínas de fusión de unión se conjugaron con un fluoróforo que se caracterizó a continuación por afinidad de unión y se utilizó en modelos animales preclínicos para evaluar la capacidad de los conjugados de distribuirse sistémicamente después de la inyección y unirse al tejido tumoral diana.

aHer2-XTEN(AE864-Cvs)-AF680

Las *E. coli* que contenían el gen AC452 en un plásmido se cultivaron en cultivo líquido hasta saturación durante la noche y, a continuación, se utilizaron 90 ml de este cultivo para inocular 4,5 l de medios de inducción pho, divididos uniformemente entre 9 matraces de 4 l. Los cultivos se cultivaron a 26 °C en presencia de 10 µg/ml de tetraciclina y se indujeron a medida que se agotó el fosfato del medio. La proteína se transportó al periplasma de las células hospedadoras a través de una secuencia de señalización STII fusionada con el extremo amínico de la proteína codificada, que posteriormente fue eliminada mediante una modificación postraduccional en las células. El sedimento celular se recogió mediante centrifugación durante 20 min a 4000 rpm con un rotor SLA-3000. Parte del sedimento celular (45 de los 90 g totales) se resuspendió en 121 ml de PBS más imidizol 10 mM, y se añadieron 13,5 ml de BugBuster y DNasa. Las células se lisaron con el agitador vorticial periódicamente durante un período de 90 minutos. A continuación, el lisado se clarificó mediante centrifugación a 15 000 rpm durante 30 minutos. A continuación, el lisado clarificado se cargó en una columna de quelato Toyopearl de 85 ml cargada con NiSO4 100 mM, y se equilibró con PBS más imidizol 10 mM. La columna se lavó con 5 volúmenes de columna de PBS más imidizol 10 mM, y la proteína se eluyó con 3 volúmenes de columna de PBS más imidizol 250 mM y a continuación se arrastró con 1,2 volúmenes

de columna de PBS más imidizol 500 mM. El eluido se redujo con TCEP 0,3 mM y a continuación se marcó durante 3 horas a la temperatura ambiente con Alexa Fluor 680 en una relación de 0,1 mg de colorante por mg de proteína. A continuación, la muestra se transfirió a 4 °C para su almacenamiento durante la noche. A continuación, la muestra se diluyó 2,7 veces con agua para reducir la conductividad y se cargó en una columna macrocapQ, previamente higienizada con NaOH y equilibrada con Tris 20 mM a pH 7,5, NaCl 50 mM. La columna se lavó con 5 volúmenes de columna de Tris 20 mM a pH 7,5, NaCl 50 mM, y se eluyó con un gradiente lineal de 10 volúmenes de columna desde NaCl 150 mM hasta NaCl 300 mM, ambos con Tris 20 mM a pH 7,5 como tampón. Las fracciones de elución agrupadas se concentraron utilizando un ultra concentrador Amicon con una membrana de 10 000 MWCO y se almacenaron a -80 °C, se asignó el n.º de lote AP502.

aHer2-XTEN(AE576-Cys) -AF680

Las *E. coli* que contenían el gen AC451 en un plásmido se cultivaron en cultivo líquido hasta saturación durante la noche y, a continuación, se utilizaron 90 ml de este cultivo para inocular 10 l de medios de inducción pho, divididos uniformemente entre 20 matraces de 4 l. Los cultivos se cultivaron a 26 °C en presencia de 10 µg/ml de tetraciclina y se indujeron a medida que se agotó el fosfato del medio. La proteína se transportó al periplasma de las células hospedadoras a través de una secuencia de señalización STII fusionada con el extremo amínico de la proteína codificada, que posteriormente fue eliminada mediante una modificación postraduccional en las células. El sedimento celular se recogió mediante centrifugación durante 20 min a 4000 rpm con un rotor SLA-3000. El sedimento celular (39 g) se resuspendió en 96,3 ml de PBS más imidizol 10 mM, y se añadieron 10,7 ml de BugBuster y DNasa. Las células se lisaron con el agitador vorticial periódicamente durante un período de 55 minutos. A continuación, el lisado se clarificó mediante centrifugación a 15 000 rpm durante 25 minutos. A continuación, el lisado clarificado se cargó en una columna de quelato Toyopearl de 85 ml cargada con NiSO4 100 mM, y se equilibró con PBS más imidizol 10 mM. La columna se lavó con 5 volúmenes de columna de PBS más imidizol 10 mM, y la proteína se eluyó con 3 volúmenes de columna de PBS más imidizol 250 mM y a continuación se arrastró con 1,2 volúmenes de columna de PBS más imidizol 500 mM. El eluido se redujo con TCEP 0,3 mM y a continuación se marcó durante 3 horas a la temperatura ambiente con Alexa Fluor 680 en una relación de 0,1 mg de colorante por mg de proteína. A continuación, la muestra se transfirió a 4 °C para su almacenamiento durante la noche. A continuación, la muestra se diluyó 2,7 veces con agua para reducir la conductividad y se cargó en una columna macrocapQ, previamente higienizada con NaOH y equilibrada con Tris 20 mM a pH 7,5, NaCl 50 mM. La columna se lavó con 5 volúmenes de columna de Tris 20 mM a pH 7,5, NaCl 50 mM, y se eluyó con un gradiente lineal de 10 volúmenes de columna desde NaCl 150 mM hasta NaCl 300 mM, ambos con Tris 20 mM a pH 7,5 como tampón. Las fracciones de elución agrupadas se concentraron utilizando un ultra concentrador Amicon con una membrana de 10 000 MWCO y se almacenaron a -80 °C, se asignó el n.º de lote AP486.

aHer2-XTEN(AE288-Cys) -AF680

Las *E. coli* que contenían el gen AC450 en un plásmido se cultivaron en cultivo líquido hasta saturación durante la noche y, a continuación, se utilizaron 90 ml de este cultivo para inocular 10 l de medios de inducción pho, divididos uniformemente entre 20 matraces de 4 l. Los cultivos se cultivaron a 26 °C en presencia de 10 µg/ml de tetraciclina y se indujeron a medida que se agotó el fosfato del medio. La proteína se transportó al periplasma de las células hospedadoras a través de una secuencia de señalización STII fusionada con el extremo amínico de la proteína codificada, que posteriormente fue eliminada mediante una modificación postraduccional en las células. El sedimento celular se recogió mediante centrifugación durante 60 min a 4000 rpm con un rotor SLA-3000. El sedimento celular (36 g) se resuspendió en 97 ml de PBS más imidizol 10 mM, y se añadieron 10,7 ml de BugBuster y DNasa. Las células se lisaron con el agitador vorticial periódicamente durante un período de 55 minutos. A continuación, el lisado se clarificó mediante centrifugación a 15 000 rpm durante 25 minutos. A continuación, el lisado clarificado se cargó en una columna de quelato Toyopearl de 85 ml cargada con NiSO4 100 mM, y se equilibró con PBS más imidizol 10 mM. La columna se lavó con 5 volúmenes de columna de PBS más imidizol 10 mM, y la proteína se eluyó con 3 volúmenes de columna de PBS más imidizol 250 mM y a continuación se arrastró con 1,2 volúmenes de columna de PBS más imidizol 500 mM. El eluido se redujo con TCEP 0,3 mM y a continuación se marcó durante 3 horas a la temperatura ambiente con Alexa Fluor 680 en una relación de 0,1 mg de colorante por mg de proteína. A continuación, la muestra se transfirió a 4 °C para su almacenamiento durante la noche. A continuación, la muestra se diluyó 2,7 veces con agua para reducir la conductividad y se cargó en una columna macrocapQ, previamente higienizada con NaOH y equilibrada con Tris 20 mM a pH 7,5, NaCl 50 mM. La columna se lavó con 5 volúmenes de columna de Tris 20 mM a pH 7,5, NaCl 50 mM, y se eluyó con un gradiente lineal de 10 volúmenes de columna desde NaCl 150 mM hasta NaCl 300 mM, ambos con Tris 20 mM a pH 7,5 como tampón. Las fracciones de elución agrupadas se concentraron utilizando un ultra concentrador Amicon con una membrana de 10 000 MWCO y se almacenaron a -80 °C, se asignó el n.º de lote AP481.

Análisis mediante SEC de aHer2-XTEN(AE288-Cys)-AF680, aHer2-XTEN(AE576-Cys)-AF680, aIler2-XTEN(AE576-Cys)-AF680

La confirmación de que Alexa Fluor 680 se había conjugado con las proteínas XTEN se proporcionó mediante una cromatografía analítica de exclusión por tamaño. La absorbancia característica a 690 nm del colorante eluiría muy tarde en el cromatograma de la SEC si estuviera libre en solución, mientras que si estuviera conjugada con la proteína

XTEN eluiría antes en el volumen característico del XTEN. Las tres construcciones se ensayaron utilizando un BioSep S4000 (7,8 x 6000 mm, phenomenex) desarrollado en un sistema purificador Akta, y el cromatograma para cada una de las tres construcciones se controló a 280 (FIG. 33, líneas continuas) y a 690 nm (FIG. 33, líneas punteadas). Para cada una de las tres construcciones, los máximos de absorbancia a 690 nm coeluyeron con el máximo a 280 nm, lo que indica que el Alex Fluor 680 estaba conjugado con la proteína de fusión XTEN. No se observó colorante libre en el tiempo de elución previsto de aproximadamente 52 minutos. Los cromatogramas también reflejan eluciones proporcionales al aumento de la masa aparente contribuido por la longitud de los respectivos componentes XTEN.

Análisis mediante citometría de flujo de conjugados aHer2-XTEN

La unión específica de una proteína de fusión de unión a una estirpe celular tumoral se evaluó utilizando las construcciones de resto de direccionamiento anti-HER2 unido a diferentes longitudes de XTEN, con un fluoróforo AF680 unido covalentemente al XTEN, como se ha descrito anteriormente. Las construcciones aHer2-XTEN-Cys-AF680 que se unen al Her2 se evaluaron mediante citometría de flujo en la estirpe celular tumoral Her2+, SKOV3, de la siguiente manera: las células SKOV3 se cultivaron a partir de reservas congeladas y se subcultivaron 2-3 veces. Las células se resuspendieron en 31 tubos y se incubaron a 1×10^6 células/tubo con Fc Block (BD BioSciences) durante 10 minutos en hielo. Se preincubaron quince tubos con Herceptin 1000 nM sin marcar (para bloquear la unión como control) durante 20 minutos en hielo. A continuación, todos los tubos (15 incubados con Herceptin sin marcar, 15 no incubados con Herceptin sin marcar) se incubaron con los 100 nM de los aHer2-XTEN-AE-288-Cys-AF680, aHer2-XTEN-AE-576-Cys-AF680 adecuadamente marcados, y el conjugado aHer2-XTEN-AE-864-Cys-AF680 o el control de isotipo marcado durante 20 minutos en hielo, se lavaron dos veces con tampón de flujo, los sedimentos se resuspendieron en 0,2 ml de tampón de flujo y a continuación se desarrollaron en un citómetro de flujo Accuro C6. Se recogieron datos para la dispersión frontal y lateral y la dispersión lateral frente a FL4 para Alexa680. Los datos (FIG. 34) se presentan como histogramas de cada conjugado de aHer2-XTEN-AE-864-Cys-AF680 (FIG. 34A), aHer2-XTEN-AE-576-Cys-AF680 (FIG. 34B) y aHer2-XTEN-AE-288-Cys-AF680 (FIG. 34C) solapado con Herceptin bloqueado y no bloqueado. Los datos muestran que Herceptin bloqueó completamente la unión del conjugado aHer2-XTEN-AE-288-Cys-AF680, aHer2-XTEN-AE-576-Cys-AF680 y aHer2-XTEN-AE-864-Cys-AF680 a las células SKOV3 Her2+ *in vitro*, lo que demuestra la unión específica de las construcciones conjugadas para la diana HER2 en las células tumorales.

Imágenes *in vivo* y *ex vivo* de conjugados aHer2-XTEN

Se evaluó el direccionamiento y la biodistribución de las construcciones aHer2-XTEN-Cys-AF680, preparadas como se ha descrito anteriormente, al tumor Her2+ utilizando imágenes por fluorescencia *in vivo*, seguido de *ex vivo*. Los grupos de control incluyeron ratones a los que se inyectó Herceptin-Alexa 680 marcado con fluorescencia y ratones a los que se inyectó aHer2-XTEN-864-Alexa 680 pero bloqueado con Herceptin una hora antes de la inyección.

A los ratones nu/nu hembra que portaban células tumorales SKOV3 se les administró una única inyección de dosis alta o baja de aHer2-XTEN-AE-288-Cys-AF680, aHer2-XTEN-AE-576-Cys-AF680, aHer2-XTEN-AE-864-Cys-AF680 o Herceptin-AF680 de control. Se realizaron escaneos de cuerpo completo antes de la inyección y a continuación aproximadamente a las 8, 24, 48 y 72 horas después de la inyección. Después del punto temporal de las 72 horas, todos los grupos con dosis alta se sacrificaron y se tomaron imágenes *ex vivo* de los tumores, hígado, pulmón, corazón, bazo y riñones utilizando imágenes de fluorescencia *ex vivo*. Las imágenes de fluorescencia se obtuvieron utilizando un sistema de obtención de imágenes óptico IVIS 50 (Caliper Life Sciences, Hopkinton, MA). Los filtros de excitación Cy5.5 (615-665 nm) y de emisión (695-770 nm) se seleccionaron para que coincidieran con las longitudes de onda de los agentes de fluorescencia. Se utilizó un empaquetado pequeño e intermedio del chip CCD y el tiempo de exposición fue de entre 5-20 segundos, para obtener al menos varios miles de recuentos a partir de las señales que se observaban en cada ratón en la imagen y para evitar la saturación del chip CCD. Para normalizar las imágenes para la cuantificación, se adquirió una imagen de fluorescencia de fondo utilizando filtros de excitación y de emisión de fondo para la región espectral Cy5.5.

Los datos de la obtención de imágenes *in vivo* se muestran en la FIG. 35 y la tabla 27. Varios de los grupos mostraron señales fluorescentes específicas en el tumor, por encima del nivel del fondo autofluorescente. Había señales significativas evidentes en la mayoría de los grupos de agentes de ensayo de aHer2-XTEN con el nivel de dosis más alto (6,7 nmol/ratón) (grupos 1, 3 y 5), así como en el grupo de control positivo al que se administró herceptin marcado (grupo 7). Además, hubo tendencias menores que sugieren la detección de los agentes en los tumores para algunos de los grupos donde se administraron los niveles de dosis más bajos. Los datos también mostraron un valor máximo un 50 % mayor para aHer2-XTEN-864-Alexa 680 (grupo 1) en comparación con aHer2-XTEN-576-Alexa 680 (grupo 3) y aHer2-XTEN-288-Alexa 680 (grupo 5). Mientras que la herceptin marcada mostró el máximo de unión al tumor 8 h después de la administración, los agentes de ensayo mostraron generalmente un máximo de unión tardío, aproximadamente 24-48 h después de la administración. aHer2-XTEN-288-Alexa 680 mostró la cinética de direccionamiento más rápida, así como el aclaramiento más rápido.

Los datos de la obtención de imágenes *ex vivo* mostrados en la FIG. 36 y la tabla 28 resumen las señales de fluorescencia medias totales medidas por grupo y tipo de tejido. La obtención de imágenes de tumores en los grupos 2 y 9 mostró aproximadamente 6 veces mayor señal de fluorescencia total en el grupo 2 (0,67 nmol de aHer2-XTEN-864-Alexa 680) en comparación con el grupo 9 (0,67 nmol de aHer2-XTEN-864-Alexa 680 + exceso de 100 veces de

Herceptin sin marcar administrado 1 hora antes), lo que indica el direccionamiento específico al tumor Her2+. Los resultados del conjunto de tejidos ampliado (tumor, corazón, pulmones, bazo, hígado, riñones) indicaron que el grupo 7 (3,3 nmol de Herceptin-Alexa680) mostró las mayores señales en la mayoría de los tejidos, pero ese grupo 1 (6,7 nmol de aHer2-XTEN-864-Alexa 680), con la molécula de XTEN más larga, mostró unas señales totales casi comparables en la mayoría de los casos excepto en el hígado, en el que el grupo 7 con Herceptin tenía aproximadamente el doble de la señal fluorescente en comparación con el grupo 1 con aHer2-XTEN-864. El grupo 3 (6,7 nmol de aHer2-XTEN-576-Alexa 680) y el grupo 5 (6,7 nmol de aHer2-XTEN-864-Alexa 680) mostraron las señales más bajas en todos los tejidos en comparación con los grupos 1 y 7.

Tabla 27: Fluorescencia *in vivo* por grupo de señales máximas medias y señal a las 72 h

N.º de grupo	Material de ensayo	Nivel de dosis (nmol/ratón)	Eficiencia media de fluorescencia (media del grupo normalizada a los valores previos al tratamiento)		
			Máximo	Momento del máximo	Nivel a las 72 h
Grupo 1	aHer2-XTEN-864-Alexa 680	6,7	20,3	24h	11,2
Grupo 2	aHer2-XTEN-864-Alexa 680	0,67	5,3	48h	3,5
Grupo 3	aHer2-XTEN-576-Alexa 680	6,7	14,0	24h	6,9
Grupo 4	aHer2-XTEN-576-Alexa 680	0,67	3,6	48h	3,4
Grupo 5	aHer2-XTEN-288-Alexa 680	6,7	13,4	8h	4,2
Grupo 6	aHer2-XTEN-288-Alexa 680	0,67	3,5	48h	3,1
Grupo 7	Control de Herceptin-Alexa680	3,3	40,3	8h	18,6
Grupo 8	Control de Herceptin-Alexa680	0,33	5,3	48h	4,4
Grupo 9	aHer2-XTEN-864-Alexa 680 + Herceptin sin marcar	0,67 + 100x en exceso 1 h antes	6,0	72 h	6,0

Tabla 28: Sumario por grupo de las señales totales medias en imágenes de tejidos tomadas *ex vivo*.

N.º de grupo	Material de ensayo	Nivel de dosis (nmol/ratón)	Eficiencia total de fluorescencia (media del grupo) ($x10^6$)					
			Tumor	Corazón	Pulmones	Bazo	Hígado	Riñones
Grupo 1	aHer2-XTEN-864-Alexa 680	6,7	42	28	130	16	180	120
Grupo 2	aHer2-XTEN-864-Alexa 680	0,67	4,2	-	-	-	-	-
Grupo 3	aHer2-XTEN-576-Alexa 680	6,7	27	6,8	24	3,4	48	31
Grupo 5	aHer2-XTEN-288-Alexa 680	6,7	7,2	1,9	5,6	2,1	20	34
Grupo 7	Control de Herceptin-Alexa680	3,3	69	32	150	25	370	110
Grupo 9	aHer2-XTEN-864-Alexa 680 + Herceptin sin marcar	0,67 + 100x en exceso 1 h antes	0,7	-	-	-	-	-

Ejemplo 42: Análisis farmacocinético de CTLA4-XTEN

La farmacocinética *in vivo* de las construcciones CTLA4-XTEN se puede evaluar utilizando métodos convencionales para composiciones de proteínas. La farmacocinética se evalúa en múltiples especies, sin embargo, se prefieren ratones, ratas, macacos cangrejeros y perros debido a su uso habitual en la predicción de la farmacocinética humana. Las composiciones de construcciones de CTLA4-XTEN, o CTLA4 como comparador, se proporcionan normalmente en un tampón acuoso compatible con la administración *in vivo* (por ejemplo: solución salina tamponada con fosfato o solución salina tamponada con Tris). Las composiciones se administrarían a las dosis apropiadas y por múltiples vías: lo más preferentemente, por vía intravenosa o subcutánea. Se recogerían muestras de sangre en los puntos temporales apropiados que variarían entre 0,08 y 504 horas, y se procesarían hasta plasma. Las muestras de plasma se pueden analizar para determinar la concentración de los productos experimentales mediante un ensayo de ELISA. El análisis se realiza normalmente utilizando un formato de ELISA en sándwich. Los anticuerpos policlonales anti-XTEN de conejo o el anti-resto direccionamiento se recubren sobre pocillos de una placa de ELISA. Los pocillos se bloquean, se lavan y las muestras de plasma se incuban a continuación en los pocillos a diluciones variables para permitir la captura del compuesto por los anticuerpos recubiertos. A continuación, los pocillos se lavan ampliamente y la proteína unida se detecta utilizando una preparación biotinilada de un anticuerpo policlonal anti CTLA4 o anticuerpo anti-XTEN y estreptavidina HRP. A continuación, se calculan las concentraciones del producto experimental en cada punto temporal comparando la respuesta colorimétrica a cada dilución sérica con una curva patrón. A continuación, se calculan los parámetros farmacocinéticos utilizando el paquete del programa informático WinNonLin. Se espera que los resultados respalden el hallazgo de que la adición de un XTEN a CTLA4 puede aumentar en gran medida la semivida terminal en comparación con el resto de direccionamiento no unido a XTEN, y mejorar también otras propiedades farmacocinéticas.

Ejemplo 43: Análisis farmacocinético de IL6R-XTEN

La farmacocinética *in vivo* de las construcciones IL6R-XTEN se puede evaluar utilizando métodos convencionales para composiciones de proteínas. La farmacocinética se evalúa en múltiples especies, sin embargo, se prefieren ratones, ratas, macacos cangrejeros y perros debido a su uso habitual en la predicción de la farmacocinética humana. Las composiciones de construcciones de IL6R-XTEN, o IL6R como comparador, se proporcionan normalmente en un tampón acuoso compatible con la administración *in vivo* (por ejemplo: solución salina tamponada con fosfato o solución salina tamponada con Tris). Las composiciones se administrarían a las dosis apropiadas y por múltiples vías: lo más preferentemente, por vía intravenosa o subcutánea. Se recogerían muestras de sangre en los puntos temporales apropiados que variarían entre 0,08 y 504 horas, y se procesarían hasta plasma. Las muestras de plasma se analizan para determinar la concentración de los productos experimentales mediante un ensayo de ELISA. El análisis se realiza normalmente utilizando un formato de ELISA en sándwich. Los anticuerpos policlonales anti-XTEN de conejo o los anticuerpos contra el resto de direccionamiento se recubren sobre pocillos de una placa de ELISA. Los pocillos se bloquean, se lavan y las muestras de plasma se incuban a continuación en los pocillos a diluciones variables para permitir la captura del compuesto por los anticuerpos recubiertos. A continuación, los pocillos se lavan ampliamente y la proteína unida se detecta utilizando una preparación biotinilada de un anticuerpo policlonal anti-IL6R o anticuerpo anti-XTEN y estreptavidina HRP. A continuación, se calculan las concentraciones del producto experimental en cada punto temporal comparando la respuesta colorimétrica a cada dilución sérica con una curva patrón. A continuación, se calculan los parámetros farmacocinéticos utilizando el paquete del programa informático WinNonLin. Se espera que los resultados respalden el hallazgo de que la adición de un XTEN a IL6R puede aumentar en gran medida la semivida terminal en comparación con el resto de direccionamiento no unido a XTEN, y mejorar también otros parámetros farmacocinéticos.

Ejemplo 44: Análisis farmacocinético de CD40-XTEN

La farmacocinética *in vivo* de las construcciones CD40-XTEN se puede evaluar utilizando métodos convencionales para composiciones de proteínas. La farmacocinética se evalúa en múltiples especies, sin embargo, se prefieren ratones, ratas, macacos cangrejeros y perros debido a su uso habitual en la predicción de la farmacocinética humana. Las composiciones de construcciones de CD40-XTEN, o CD40 como comparador, se proporcionan normalmente en un tampón acuoso compatible con la administración *in vivo* (por ejemplo: solución salina tamponada con fosfato o solución salina tamponada con Tris). Las composiciones se administrarían a las dosis apropiadas y por múltiples vías: lo más preferentemente, por vía intravenosa o subcutánea. Se recogerían muestras de sangre en los puntos temporales apropiados que variarían entre 0,08 y 504 horas, y se procesarían hasta plasma. Las muestras de plasma se analizan para determinar la concentración de los productos experimentales mediante un ensayo de ELISA. El análisis se realiza normalmente utilizando un formato de ELISA en sándwich. Los anticuerpos policlonales anti-XTEN de conejo o los anticuerpos contra el resto de direccionamiento se recubren sobre pocillos de una placa de ELISA. Los pocillos se bloquean, se lavan y las muestras de plasma se incuban a continuación en los pocillos a diluciones variables para permitir la captura del compuesto por los anticuerpos recubiertos. A continuación, los pocillos se lavan ampliamente y la proteína unida se detecta utilizando una preparación biotinilada de un anticuerpo policlonal anti CD40 o anticuerpo anti-XTEN y estreptavidina HRP. A continuación, se calculan las concentraciones del producto experimental en cada punto temporal comparando la respuesta colorimétrica a cada dilución sérica con una curva patrón. A continuación, se calculan los parámetros farmacocinéticos utilizando el paquete del programa informático WinNonLin. Se espera que los resultados respalden el hallazgo de que la adición de un XTEN a CD40 puede aumentar

en gran medida la semivida terminal en comparación con el resto de direccionamiento no unido a XTEN, y mejorar también otros parámetros farmacocinéticos.

Ejemplo 45: Análisis preclínico de CTLA4-XTEN

El CTLA4 está implicado en el suministro de la segunda señal coestimuladora requerida para la activación óptima de los linfocitos T. Como tal, la actividad farmacológica *in vivo* de las construcciones de CTLA4-XTEN puede evaluarse utilizando modelos preclínicos de enfermedades inflamatorias autoinmunitarias humanas. Los modelos apropiados para los ensayos preclínicos de eficacia incluyen, pero sin limitación, artritis inducida por colágeno, un modelo de artritis reumatoide humana, lupus eritematoso sistémico, un modelo del lupus humano, y encefalomielitis alérgica experimental, un modelo de esclerosis múltiple humana. Los ensayos preclínicos de eficacia también se pueden realizar en modelos de trasplante, tales como aloinjerto de órgano sólido o trasplante de islotes. Estos modelos se pueden desarrollar en múltiples especies utilizando métodos equivalentes a los utilizados para abatacept. Las composiciones de CTLA4-XTEN se proporcionan en un tampón acuoso compatible con la administración *in vivo* (por ejemplo: solución salina tamponada con fosfato o solución salina tamponada con Tris). Las composiciones se administrarían a las dosis, frecuencia de administración, cronograma de administración y vía de administración apropiados, según se optimice para el modelo particular. Las lecturas de eficacia para la inflamación podrían incluir mediciones articulares, inhibición de la respuesta inmunitaria humoral primaria y secundaria, infiltración de células inmunitarias medida por histopatología, proteinuria, entre otros. Se espera que los resultados respalden el hallazgo de que las construcciones de CTLA4-XTEN pueden ser más eficaces para inhibir la respuesta inflamatoria en comparación con CTLA4 y/o un equivalente en potencia a una dosis comparable de CTLA4 con una administración menos frecuente o más conveniente.

Ejemplo 46: Análisis preclínico de anti-IL6R-XTEN

La IL6 desempeña un papel importante en la patogenia de la artritis reumatoide. E anti-IL6R inhibe la unión de la IL6 a su receptor y neutraliza las acciones de la IL6. Como tal, la actividad farmacológica *in vivo* de las construcciones anti-IL6R-XTEN puede evaluarse utilizando modelos preclínicos de enfermedades inflamatorias autoinmunitarias humanas, en particular, de artritis reumatoide. Los modelos apropiados para los ensayos preclínicos de eficacia incluyen, pero sin limitación, artritis inducida por colágeno en primates no humanos, un modelo de artritis reumatoide humana. Estos modelos se pueden desarrollar utilizando métodos equivalentes a los utilizados para tocilizumab. Las composiciones de anti-IL6R-XTEN se proporcionan en un tampón acuoso compatible con la administración *in vivo* (por ejemplo: solución salina tamponada con fosfato o solución salina tamponada con Tris). Las composiciones se administrarían a las dosis, frecuencia de administración, cronograma de administración y vía de administración apropiados, según se optimice para el modelo particular. Las lecturas de eficacia para la inflamación podrían incluir mediciones articulares, hinchazón, inhibición de la respuesta inmunitaria humoral primaria y secundaria, infiltración de células inmunitarias medida por histopatología, análisis bioquímico de la sangre, entre otros. Se espera que los resultados respalden el hallazgo de que las construcciones anti-IL6R4-XTEN pueden ser más eficaces para inhibir la respuesta inflamatoria en comparación con anti-IL6R y/o un equivalente en potencia a unas dosis comparables de anti-IL6R con una administración menos frecuente o más conveniente.

Ejemplo 47: Análisis preclínico de anti-CD40-XTEN

La desregulación de la vía de coestimulación CD40-CD40L desempeña un papel importante en la patogenia de la enfermedad inflamatoria y autoinmunitaria humana. Como tal, la actividad farmacológica *in vivo* de las construcciones de anti-CD40-XTEN puede evaluarse utilizando modelos preclínicos de enfermedades inflamatorias autoinmunitarias humanas. Los modelos apropiados para los ensayos preclínicos de eficacia incluyen, pero sin limitación, artritis inducida por colágeno, un modelo de artritis reumatoide humana, lupus eritematoso sistémico, un modelo del lupus humano, y encefalomielitis alérgica experimental, un modelo de esclerosis múltiple humana. Los ensayos preclínicos de eficacia también se pueden realizar en modelos de trasplante, tales como aloinjerto de órgano sólido o trasplante de islotes. Estos modelos se pueden desarrollar utilizando métodos equivalentes a los utilizados para otras terapias de direccionamiento a CD40 o CD40L. Las composiciones de anti-CD40-XTEN se pueden en un tampón acuoso compatible con la administración *in vivo* (por ejemplo: solución salina tamponada con fosfato o solución salina tamponada con Tris). Las composiciones se pueden administrar a las dosis, frecuencia de administración, cronograma de administración y vía de administración apropiados, según se optimice para el modelo particular. Las lecturas de eficacia para la inflamación podrían incluir mediciones articulares, inhibición de la respuesta inmunitaria humoral primaria y secundaria, infiltración de células inmunitarias medida por histopatología, proteinuria, entre otros. Se espera que los resultados respalden el hallazgo de que las construcciones anti-CD40-XTEN pueden ser más eficaces para inhibir la respuesta inflamatoria en comparación con anti-CD40 y/o un equivalente en potencia a unas dosis comparables de anti-CD40 con una administración menos frecuente o más conveniente.

Ejemplo 48: Aplicaciones clínicas de CTLA4-XTEN

El CTLA4 está implicado en el suministro de la segunda señal coestimuladora requerida para la activación óptima de los linfocitos T. El CTLA4-XTEN se puede utilizar para tratar enfermedades autoinmunitarias mediadas por los linfocitos T, tales como artritis reumatoide y psoriasis, en ensayos clínicos utilizando una metodología similar a Orencia. Se

espera que la fusión de XTEN a CTLA4 para crear una composición de proteína de fusión de unión mejore la semivida de la proteína recombinante, permitiendo así una dosis global más baja por paciente con mejoras posteriores en la conveniencia (permitiendo la administración subcutánea, reduciendo la frecuencia de administración, etc.) y el coste (se requiere menos fármaco por dosis).

Se podrían realizar ensayos clínicos en pacientes que padezcan artritis reumatoide. Los ensayos clínicos pueden diseñarse de manera que se pueda verificar la eficacia y las ventajas de las composiciones de CTLA4-XTEN en seres humanos. Dichos estudios en pacientes comprenderían tres fases. En primer lugar, se realizaría un estudio de seguridad y farmacocinética de fase I en pacientes adultos para determinar la dosis máxima tolerada y la farmacocinética y la farmacodinámica en seres humanos. Estos estudios iniciales podrían realizarse en pacientes con artritis reumatoide y definirían posibles toxicidades y acontecimientos adversos de los que se debe hacer un seguimiento en futuros estudios. El esquema del estudio sería utilizar dosis únicas en aumento de las composiciones de CTLA4-XTEN y medir los parámetros bioquímicos, FC y clínicos. Esto permitiría la determinación de la dosis máxima tolerada y establecer el umbral y las concentraciones máximas en la dosis y el fármaco circulante que constituyen la ventana terapéutica que se utilizará en los ensayos posteriores de fase II y fase III realizados para las indicaciones diana, para determinar la eficacia y la tolerabilidad de las composiciones de CTLA4-XTEN.

Se llevaría a cabo un estudio clínico de fase II en pacientes con artritis que recibirían CTLA4-XTEN o una proteína antiinflamatoria adecuada para determinar una dosis apropiada para aliviar al menos un síntoma asociado con la artritis reumatoide, que incluye reducir la hinchazón articular, la sensibilidad articular, la inflamación, la rigidez matutina y el dolor, o al menos un marcador biológico indirecto asociado con la artritis reumatoide, que incluye reducir las tasas de sedimentación de eritrocitos y los niveles séricos de proteína C reactiva y/o de receptor de la IL2. Además, se recopilarían datos de seguridad relacionados con los acontecimientos adversos. Se realizaría un estudio de eficacia de fase III en donde los pacientes con artritis recibirían CTLA4-XTEN, un control positivo o un placebo diario, quincenal o semanal (u otro cronograma de administración que se considere apropiado dadas las propiedades farmacocinéticas y farmacodinámicas del compuesto) durante un período prolongado de tiempo. Los pacientes se evaluarían para detectar los síntomas iniciales de la actividad de la enfermedad antes de recibir cualquier tratamiento, que incluye hinchazón articular, sensibilidad articular, inflamación, rigidez matutina, la actividad de la enfermedad evaluada por el paciente y el médico, así como la discapacidad evaluada, por ejemplo, mediante una evaluación normalizada del Cuestionario de Salud (HAQ) y el dolor. Las evaluaciones iniciales adicionales podrían incluir tasas de sedimentación de eritrocitos (ESR), niveles séricos de proteína C reactiva (CRP) y receptor de IL-2 soluble (IL-2r). La respuesta clínica al tratamiento podría evaluarse utilizando los criterios establecidos por el American College of Rheumatology (ACR), tales como el criterio ACR20; es decir, si hubo una mejora del 20 por ciento en los recuentos de articulaciones sensibles e hinchadas y una mejora del 20 por ciento en tres de los cinco síntomas restantes medidos, tales como cambios en la enfermedad global según el paciente y el médico, dolor, discapacidad, y un reactivo de fase aguda (Felson, D. T., et al., 1993 Arthritis and Rheumatism 36: 729-740; Felson, D. T., et al., 1995 Arthritis and Rheumatism 38: 1-9). De manera similar, un sujeto cumpliría el criterio ACR50 o ACR70 si hubo una mejora del 50 o el 70 por ciento, respectivamente, en los recuentos de articulaciones sensibles e hinchadas y una mejora del 50 o 70 por ciento, respectivamente, en tres de los cinco síntomas restantes medidos, tales como cambios en la enfermedad global según el paciente y el médico, dolor, discapacidad física y un reactivo de fase aguda tal como la PCR o la ESR. Además, se podrían medir posibles biomarcadores de la actividad de la enfermedad, incluyendo factor reumatoide, CRP, ESR, IL-2R soluble, ICAM-1 soluble, E-selectina soluble y MMP-3. Los resultados de eficacia se determinarían utilizando métodos estadísticos convencionales. En este estudio también se realizaría un seguimiento de los marcadores de toxicidad y los acontecimientos adversos para verificar que el compuesto es seguro cuando se utiliza de la manera descrita.

Ejemplo 49: Aplicaciones clínicas de IL6R-XTEN

La IL6 desempeña un papel importante en la patogenia de la artritis reumatoide. E anti-IL6R inhibe la unión de la IL6 a su receptor y neutraliza las acciones de la IL6. El anti-IL6R-XTEN se puede utilizar para tratar enfermedades autoinmunitarias mediadas por los linfocitos T, tales como la artritis reumatoide, en ensayos clínicos utilizando una metodología similar a Actemra. Se espera que la fusión de XTEN a anti-IL6R para crear una proteína de fusión de unión mejore la semivida de la proteína recombinante, permitiendo así una dosis global más baja por paciente con mejoras posteriores en la conveniencia (permitiendo la administración subcutánea, reduciendo la frecuencia de administración, etc.) y el coste (se requiere menos fármaco por dosis). El anti-IL6R-XTEN también puede proporcionar una ventaja de seguridad sobre la terapia anti-IL6R existente.

Se podrían realizar ensayos clínicos en pacientes que padezcan artritis reumatoide. Los ensayos clínicos pueden diseñarse de manera que se pueda verificar la eficacia y las ventajas de las composiciones de anti-IL6R-XTEN en seres humanos. Dichos estudios en pacientes comprenderían tres fases. En primer lugar, se realizaría un estudio de seguridad y farmacocinética de fase I en pacientes adultos para determinar la dosis máxima tolerada y la farmacocinética y la farmacodinámica en seres humanos. Estos estudios iniciales podrían realizarse en pacientes con artritis reumatoide y definirían posibles toxicidades y acontecimientos adversos de los que se debe hacer un seguimiento en futuros estudios. El esquema del estudio sería utilizar dosis únicas en aumento de las composiciones de anti-IL6R-XTEN y medir los parámetros bioquímicos, FC y clínicos. Esto permitiría la determinación de la dosis máxima tolerada y establecer el umbral y las concentraciones máximas en la dosis y el fármaco circulante que

constituyen la ventana terapéutica que se utilizará en los ensayos posteriores de fase II y fase III realizados para las indicaciones diana, para determinar la eficacia y la tolerabilidad de las composiciones de anti-IL6R-XTEN.

Se llevaría a cabo un estudio clínico de fase II en pacientes con artritis que recibirían anti-IL6R-XTEN o una proteína antiinflamatoria adecuada para determinar una dosis apropiada para aliviar al menos un síntoma asociado con la artritis reumatoide, que incluye reducir la hinchazón articular, la sensibilidad articular, la inflamación, la rigidez matutina y el dolor, o al menos un marcador biológico indirecto asociado con la artritis reumatoide, que incluye reducir las tasas de sedimentación de eritrocitos y los niveles séricos de proteína C reactiva y/o de receptor de la IL2. Además, se recopilarían datos de seguridad relacionados con los acontecimientos adversos. Se realizaría un estudio de eficacia de fase III en donde los pacientes con artritis recibirían anti-IL6R-XTEN, un control positivo o un placebo diario, quincenal o semanal (u otro cronograma de administración que se considere apropiado dadas las propiedades farmacocinéticas y farmacodinámicas del compuesto) durante un período prolongado de tiempo. Los pacientes se evaluarían para detectar los síntomas iniciales de la actividad de la enfermedad antes de recibir cualquier tratamiento, que incluye hinchazón articular, sensibilidad articular, inflamación, rigidez matutina, la actividad de la enfermedad evaluada por el paciente y el médico, así como la discapacidad evaluada, por ejemplo, mediante una evaluación normalizada del Cuestionario de Salud (HAQ) y el dolor. Las evaluaciones iniciales adicionales podrían incluir tasas de sedimentación de eritrocitos (ESR), niveles séricos de proteína C reactiva (CRP) y receptor de IL-2 soluble (IL-2r). La respuesta clínica al tratamiento podría evaluarse utilizando los criterios establecidos por el American College of Rheumatology (ACR), tales como el criterio ACR20; es decir, si hubo una mejora del 20 por ciento en los recuentos de articulaciones sensibles e hinchadas y una mejora del 20 por ciento en tres de los cinco síntomas restantes medidos, tales como cambios en la enfermedad global según el paciente y el médico, dolor, discapacidad, y un reactivo de fase aguda (Felson, D. T., et al., 1993 Arthritis and Rheumatism 36: 729-740; Felson, D. T., et al., 1995 Arthritis and Rheumatism 38: 1-9). De manera similar, un sujeto cumpliría el criterio ACR50 o ACR70 si hubo una mejora del 50 o el 70 por ciento, respectivamente, en los recuentos de articulaciones sensibles e hinchadas y una mejora del 50 o 70 por ciento, respectivamente, en tres de los cinco síntomas restantes medidos, tales como cambios en la enfermedad global según el paciente y el médico, dolor, discapacidad física y un reactivo de fase aguda tal como la PCR o la ESR. Además, se podrían medir posibles biomarcadores de la actividad de la enfermedad, incluyendo factor reumatoide, CRP, ESR, IL-2R soluble, ICAM-1 soluble, E-selectina soluble y MMP-3. Los resultados de eficacia se determinarían utilizando métodos estadísticos convencionales. En este estudio también se realizaría un seguimiento de los marcadores de toxicidad y los acontecimientos adversos para verificar que el compuesto es seguro cuando se utiliza de la manera descrita.

Ejemplo 50: Aplicaciones clínicas de anti-CD40-XTEN

La desregulación de la vía de coestimulación CD40-CD40L desempeña un papel importante en la patogenia de la enfermedad inflamatoria y autoinmunitaria humana. El CD40 está sobreexpresado en células presentadoras del antígeno en una variedad de afecciones autoinmunitarias que incluyen artritis reumatoide, psoriasis, enfermedad inflamatoria intestinal y diabetes de tipo 1, y su ligando, CD154, está sobreexpresado en los linfocitos T en muchas de estas mismas enfermedades autoinmunitarias. Una proteína de fusión de unión de anti-CD40-XTEN podría usarse para evaluar la eficacia en enfermedades inflamatorias autoinmunitarias y su capacidad para inducir tolerancia al trasplante en ensayos clínicos utilizando una metodología similar a los anticuerpos anti-CD40, actualmente en ensayos clínicos. Se espera que la fusión de XTEN a anti-CD40 para crear una proteína de fusión de unión mejore la semivida de la proteína recombinante, permitiendo así una dosis global más baja por paciente con mejoras posteriores en la conveniencia (frecuencia de administración reducida, etc.) y el coste (se requiere menos fármaco por dosis).

Se podrían realizar ensayos clínicos en pacientes que padezcan cualquiera de una variedad de afecciones inflamatorias y autoinmunitarias, tales como, sin limitación, artritis reumatoide, lupus eritematoso, psoriasis, enfermedad inflamatoria intestinal, esclerosis múltiple, etc. o para trasplantes. Los ensayos clínicos pueden diseñarse de manera que se pueda verificar la eficacia y las ventajas de las composiciones de anti-CD40-XTEN en seres humanos. Dichos estudios en pacientes comprenderían tres fases. En primer lugar, se realizaría un estudio de seguridad y farmacocinética de fase I en pacientes adultos para determinar la dosis máxima tolerada y la farmacocinética y la farmacodinámica en seres humanos. Estos estudios definirían posibles toxicidades y acontecimientos adversos de los que se debe hacer un seguimiento en futuros estudios. El esquema del estudio sería utilizar dosis únicas en aumento de las composiciones de anti-CD40-XTEN y medir los parámetros bioquímicos, FC y clínicos. Esto permitiría la determinación de la dosis máxima tolerada y establecer el umbral y las concentraciones máximas en la dosis y el fármaco circulante que constituyen la ventana terapéutica que se utilizará en los ensayos posteriores de fase II y fase III realizados para las indicaciones diana, para determinar la eficacia y la tolerabilidad de las composiciones de anti-CD40-XTEN.

Se llevaría a cabo un estudio clínico de fase II en pacientes con artritis que recibirían anti-CD40-XTEN o una proteína antiinflamatoria adecuada para determinar una dosis apropiada para aliviar al menos un síntoma asociado con la artritis reumatoide, que incluye reducir la hinchazón articular, la sensibilidad articular, la inflamación, la rigidez matutina y el dolor, o al menos un marcador biológico indirecto asociado con la artritis reumatoide, que incluye reducir las tasas de sedimentación de eritrocitos y los niveles séricos de proteína C reactiva y/o de receptor de la IL2. Además, se recopilarían datos de seguridad relacionados con los acontecimientos adversos. Se realizaría un estudio de eficacia de fase III en donde los pacientes con artritis recibirían anti-CD40-XTEN, un control positivo o un placebo diario,

quincenal o semanal (u otro cronograma de administración que se considere apropiado dadas las propiedades farmacocinéticas y farmacodinámicas del compuesto) durante un período prolongado de tiempo. Los pacientes se evaluarían para detectar los síntomas iniciales de la actividad de la enfermedad antes de recibir cualquier tratamiento, que incluye hinchazón articular, sensibilidad articular, inflamación, rigidez matutina, la actividad de la enfermedad evaluada por el paciente y el médico, así como la discapacidad evaluada, por ejemplo, mediante una evaluación normalizada del Cuestionario de Salud (HAQ) y el dolor. Las evaluaciones iniciales adicionales podrían incluir tasas de sedimentación de eritrocitos (ESR), niveles séricos de proteína C reactiva (CRP) y receptor de IL-2 soluble (IL-2r). La respuesta clínica al tratamiento podría evaluarse utilizando los criterios establecidos por el American College of Rheumatology (ACR), tales como el criterio ACR20; es decir, si hubo una mejora del 20 por ciento en los recuentos de articulaciones sensibles e hinchadas y una mejora del 20 por ciento en tres de los cinco síntomas restantes medidos, tales como cambios en la enfermedad global según el paciente y el médico, dolor, discapacidad, y un reactivo de fase aguda (Felson, D. T., et al., 1993 Arthritis and Rheumatism 36: 729-740; Felson, D. T., et al., 1995 Arthritis and Rheumatism 38: 1-9). De manera similar, un sujeto cumpliría el criterio ACR50 o ACR70 si hubo una mejora del 50 o el 70 por ciento, respectivamente, en los recuentos de articulaciones sensibles e hinchadas y una mejora del 50 o 70 por ciento, respectivamente, en tres de los cinco síntomas restantes medidos, tales como cambios en la enfermedad global según el paciente y el médico, dolor, discapacidad física y un reactivo de fase aguda tal como la PCR o la ESR. Además, se podrían medir posibles biomarcadores de la actividad de la enfermedad, incluyendo factor reumatoide, CRP, ESR, IL-2R soluble, ICAM-1 soluble, E-selectina soluble y MMP-3. Los resultados de eficacia se determinarían utilizando métodos estadísticos convencionales. En este estudio también se realizaría un seguimiento de los marcadores de toxicidad y los acontecimientos adversos para verificar que el compuesto es seguro cuando se utiliza de la manera descrita.

Ejemplo 51: Aplicaciones clínicas de aHER2-XTEN-aCD3

El antígeno Her2 está sobreexpresado en un gran número de neoplasias malignas sólidas. La expresión es particularmente alta en muchas células de cáncer de mama. Herceptin ha sido aprobado para el tratamiento de cánceres de mama positivos para HER2. Podría evaluarse una proteína de fusión de unión de anti-Her2-XTEN-anti-CD3 para determinar la eficacia en el tratamiento de la misma población de pacientes. Los ensayos clínicos pueden diseñarse de manera que se pueda verificar la eficacia y las ventajas de las composiciones de aHER2-XTEN-aCD3 en seres humanos. Dichos estudios en pacientes comprenderían tres fases. En primer lugar, se realizaría un estudio de seguridad y farmacocinética de fase I en pacientes adultos para determinar la dosis máxima tolerada y la farmacocinética y la farmacodinámica en seres humanos. Estos estudios definirían posibles toxicidades y acontecimientos adversos de los que se debe hacer un seguimiento en futuros estudios. El esquema del estudio sería utilizar dosis únicas en aumento de las composiciones de aHER2-XTEN-aCD3 y medir los parámetros bioquímicos, FC y clínicos. Esto permitiría la determinación de la dosis máxima tolerada y establecer el umbral y las concentraciones máximas en la dosis y el fármaco circulante que constituyen la ventana terapéutica que se utilizará en los ensayos posteriores de fase II y fase III realizados para las indicaciones diana, para determinar la eficacia y la tolerabilidad de las composiciones de aHER2-XTEN-aCD3.

Ejemplo 52: Caracterización de la estructura secundaria de la proteína de fusión que comprende XTEN

Se evaluó una proteína de fusión que consiste en XTEN AE864 unida a una carga activa de exenatida para determinar el grado de estructura secundaria mediante espectroscopia de dicroísmo circular. La espectroscopia de DC se realizó con un espectropolarímetro Jasco J-715 (Jasco Corporation, Tokio, Japón) equipado con controlador de temperatura Jasco Peltier (TPC-348WI). La concentración de proteína se ajustó a 0,2 mg/ml en fosfato de sodio 20 mM a pH 7,0, NaCl 50 mM. Los experimentos se llevaron a cabo utilizando celdas de cuarzo HELLMA con un recorrido óptico de 0,1 cm. Los espectros DC se adquirieron a 5°, 25°, 45° y 65 °C y se procesaron utilizando el programa informático J-700 versión 1.08.01 (Build 1) Jasco para Windows. Las muestras se equilibraron a cada temperatura durante 5 min antes de realizar las mediciones de DC. Todos los espectros se registraron por duplicado de 300 nm a 185 nm utilizando un ancho de banda de 1 nm y una constante de tiempo de 2 s, a una velocidad de barrido de 100 nm/min. El espectro de DC mostrado en la FIG. 30 no muestra datos sugestivos de una estructura secundaria estable y es coherente con un polipéptido no estructurado.

Ejemplo 53: Farmacocinética de polipéptidos expandidos fusionados con GFP en macacos cangrejeros

La farmacocinética de GFP-L288, GFP-L576, GFP-XTEN AF576, GFP-XTEN Y576 y XTEN_AD836-GFP se ensayó en macacos cangrejeros para determinar el efecto de la composición y la longitud de los polipéptidos no estructurados en los parámetros FC. Las muestras de sangre se analizaron en diferentes momentos después de la inyección, y la concentración de GFP en plasma se midió mediante un ELISA utilizando un anticuerpo policlonal contra la GFP para la captura y una preparación biotinilada del mismo anticuerpo policlonal para la detección. Los resultados se resumen en la FIG. 31. Muestran un sorprendente aumento de la semivida al aumentar la longitud de la secuencia del XTEN. Por ejemplo, se determinó una semivida de 10 h para GFP-XTEN_L288 (con 288 restos de aminoácidos en el XTEN). Al duplicar la longitud del compañero de fusión del polipéptido no estructurado hasta 576 aminoácidos, la semivida aumentó hasta 20-22 h para múltiples construcciones de proteína de fusión; es decir, GFP-XTEN_L576, GFP-XTEN_AF576, GFP-XTEN_Y576. Un aumento adicional de la longitud del compañero de fusión del polipéptido no estructurado hasta 836 restos dio como resultado una semivida de 72-75 h para XTEN_AD836-GFP. Por lo tanto, al

aumentar la longitud del polímero en 288 restos, de 288 a 576 restos, la semivida *in vivo* aumentó en aproximadamente 10 h. Sin embargo, al aumentar la longitud del polipéptido en 260 restos, de 576 restos a 836 restos, la semivida aumentó en más de 50 h. Estos resultados muestran que hay un sorprendente umbral de longitud del polipéptido no estructurado que da como resultado un aumento superior al proporcional en la semivida *in vivo*. Por lo tanto, se espera que las proteínas de fusión que comprenden polipéptidos expandidos no estructurados tengan la propiedad de una farmacocinética mejorada en comparación con polipéptidos de longitudes más cortas.

Ejemplo 54: Aumento de la solubilidad y la estabilidad de una carga activa peptídica mediante la unión a XTEN

Con el fin de evaluar la capacidad del XTEN para mejorar las propiedades fisicoquímicas de solubilidad y estabilidad, se prepararon y evaluaron proteínas de fusión de glucagón más XTEN de longitud más corta. Los productos experimentales se prepararon en solución salina tamponada con Tris a pH neutro, y la caracterización de la solución de Gcg-XTEN se realizó mediante una HPLC en fase inversa y una cromatografía de exclusión por tamaño, para afirmar que la proteína era homogénea y no estaba agregada en solución. Los datos se presentan en la tabla 29. Con fines comparativos, se midió el límite de solubilidad del glucagón no modificado en el mismo tampón a 60 µM (0,2 mg/ml), y el resultado demuestra que para todas las longitudes de XTEN añadido, se consiguió un aumento sustancial en la solubilidad. Es importante destacar que, en la mayoría de los casos, las proteínas de fusión de glucagón-XTEN se prepararon para conseguir las concentraciones diana y no se evaluaron para determinar los límites máximos de solubilidad para la construcción dada. Sin embargo, en el caso del glucagón unido al AF-144 XTEN, se determinó el límite de solubilidad, con el resultado de que se logró un aumento de 60 veces en la solubilidad, en comparación con el glucagón no unido a XTEN. Además, se evaluó la estabilidad del glucagón-AF144 CFXTEN y se encontró que era estable en formulación líquida durante al menos 6 meses en condiciones refrigeradas y durante aproximadamente un mes a 37 °C (datos no mostrados).

Los datos respaldan la conclusión de que la unión de polipéptidos XTEN cortos a una proteína biológicamente activa tal como glucagón puede mejorar notablemente las propiedades de solubilidad de la proteína mediante la proteína de fusión resultante, así como conferir estabilidad a las concentraciones de proteína más altas.

Tabla 29: Solubilidad de las construcciones de glucagón-XTEN

Producto experimental	**Solubilidad**
Glucagón	60 µM
Glucagón-Y36	>370 µM
Glucagón-Y72	>293 µM
Glucagón-AF108	>145 µM
Glucagón-AF120	>160 µM
Glucagón-Y144	>497 µM
Glucagón-AE144	>467 µM
Glucagón-AF144	>3600 µM
Glucagón-Y288	>163 µM

Ejemplo 55: Proteínas de fusión de unión con secuencias de escisión

XTEN carboxiterminal liberable por el FXIa

Se puede crear una proteína de fusión que consista en una proteína XTEN fusionada con el extremo carboxílico de un resto de direccionamiento con una secuencia de liberación del sitio de escisión del XTEN colocada entre el resto de direccionamiento y los componentes del XTEN. En este caso, la secuencia de escisión del sitio de liberación puede incorporarse al XTEN que contiene una secuencia de aminoácidos que es reconocida y escindida por la proteasa FXIa (EC 3.4.21.27, Uniprot P03951). Específicamente, la secuencia de aminoácidos KLTRAET (SEQ ID NO: 748) es cortada después de la arginina de la secuencia por la proteasa FXIa. El FXI es la proteasa procoagulante ubicada inmediatamente antes del FVIII en la vía de coagulación intrínseca o activada por contacto. El FXIa activo se produce a partir del FXI mediante una escisión proteolítica del zimógeno por parte del FXIIa. La producción del FXIa está estrechamente controlada y aparece únicamente cuando es necesaria la coagulación para una hemostasia adecuada. Por lo tanto, mediante la incorporación de la secuencia de escisión KLTRAET (SEQ ID NO: 748), el dominio XTEN se elimina del resto de direccionamiento simultáneamente con la activación de la vía de coagulación intrínseca cerca del resto de direccionamiento-XTEN. Esto crea una situación donde la proteína de fusión de resto de direccionamiento-XTEN se procesa de una manera adicional durante la activación de la vía intrínseca.

XTEN carboxiterminal liberable por la elastasa-2

Se puede crear una proteína de fusión que consista en una proteína XTEN fusionada con el extremo carboxílico de un resto de direccionamiento con una secuencia de liberación del sitio de escisión del XTEN colocada entre el resto de direccionamiento y los componentes del XTEN. En este caso, el sitio de liberación contiene una secuencia de aminoácidos que es reconocida y escindida por la proteasa elastasa-2 (EC 3.4.21.37, Uniprot P08246). Específicamente, la secuencia LGPVSGVP (SEQ ID NO: 749) [Rawlings N.D., et al. (2008) *Nucleic Acids Res.,* **36**: D320], es cortada después de la posición 4 de la secuencia. La elastasa es expresada constitutivamente por los neutrófilos y está presente en todo momento en la circulación, pero particularmente durante la inflamación aguda. Por lo tanto, a medida que circula el resto de direccionamiento-XTEN de vida larga, una fracción de éste se escinde, particularmente localmente durante las respuestas inflamatorias, creando una reserva de resto de direccionamiento de vida más corta que se utilizará en el sitio de la inflamación.

XTEN carboxiterminal liberable por la MMP-12

Se puede crear una proteína de fusión que consista en una proteína XTEN fusionada con el extremo carboxílico de un resto de direccionamiento con una secuencia de liberación del sitio de escisión del XTEN colocada entre el resto de direccionamiento y los componentes del XTEN. En este caso, el sitio de liberación contiene una secuencia de aminoácidos que es reconocida y escindida por la proteasa MMP-12 (EC 3.4.24.65, Uniprot P39900). Específicamente, la secuencia GPAGLGGA (SEQ ID NO: 750) [Rawlings N.D., et al. (2008) Nucleic Acids Res., 36: D320], es cortada después de la posición 4 de la secuencia. La MMP-12 es expresada constitutivamente en la sangre completa. Por lo tanto, a medida que circula el AAT-XTEN de vida larga, una fracción de éste se escinde, creando una reserva de AAT de vida más corta que se podrá utilizar. En una característica deseable de la composición inventiva, esto crea un depósito de profármaco circulante que libera constantemente una cantidad profiláctica del resto de direccionamiento, liberándose unas cantidades más altas durante una respuesta inflamatoria.

XTEN carboxiterminal liberable por la MMP-13

Se puede crear una proteína de fusión que consista en una proteína XTEN fusionada con el extremo carboxílico de un resto de direccionamiento con una secuencia de liberación del sitio de escisión del XTEN colocada entre el resto de direccionamiento y los componentes del XTEN. En este caso, el sitio de liberación contiene una secuencia de aminoácidos que es reconocida y escindida por la proteasa MMP-13 (EC 3.4.24.-, Uniprot P45452). Específicamente, la secuencia GPAGLRGA (SEQ ID NO: 751) [Rawlings N.D., et al. (2008) *Nucleic Acids Res.,* **36**: D320], es cortada después de la posición 4. La MMP-13 es expresada constitutivamente en la sangre completa. Por lo tanto, a medida que circula el resto de direccionamiento-XTEN de vida larga, una fracción de éste se escinde, creando una reserva de AAT de vida más corta que se podrá utilizar. En una característica deseable de la composición inventiva, esto crea un depósito de profármaco circulante que libera constantemente una cantidad profiláctica del resto de direccionamiento, liberándose unas cantidades más altas durante una respuesta inflamatoria.

XTEN carboxiterminal liberable por la MMP-17

Se puede crear una proteína de fusión que consista en una proteína XTEN fusionada con el extremo carboxílico de un resto de direccionamiento con una secuencia de liberación del sitio de escisión del XTEN colocada entre el resto de direccionamiento y los componentes del XTEN. En este caso, el sitio de liberación contiene una secuencia de aminoácidos que es reconocida y escindida por la proteasa MMP-20 (EC.3.4.24.-, Uniprot Q9ULZ9). Específicamente, la secuencia APLGLRLR (SEQ ID NO: 752) [Rawlings N.D., et al. (2008) Nucleic Acids Res., 36: D320], es cortada después de la posición 4 de la secuencia. La MMP-17 es expresada constitutivamente en la sangre completa. Por lo tanto, a medida que circula el resto de direccionamiento-XTEN de vida larga, una fracción de éste se escinde, creando una reserva de resto de direccionamiento de vida más corta que se podrá utilizar. En una característica deseable de la composición inventiva, esto crea un depósito de profármaco circulante que libera constantemente una cantidad profiláctica del resto de direccionamiento, liberándose unas cantidades más altas durante una respuesta inflamatoria.

XTEN carboxiterminal liberable por la MMP-20

Se puede crear una proteína de fusión que consista en una proteína XTEN fusionada con el extremo carboxílico de un resto de direccionamiento con una secuencia de liberación del sitio de escisión del XTEN colocada entre el resto de direccionamiento y los componentes del XTEN. En este caso, el sitio de liberación contiene una secuencia de aminoácidos que es reconocida y escindida por la proteasa MMP-20 (EC.3.4.24.-, Uniprot O60882). Específicamente, la secuencia PALPLVAQ (SEQ ID NO: 753) [Rawlings N.D., et al. (2008) Nucleic Acids Res., 36: D320], es cortada después de la posición 4 (representada por la flecha). La MMP-20 es expresada constitutivamente en la sangre completa. Por lo tanto, a medida que circula el resto de direccionamiento-XTEN de vida larga, una fracción de éste se escinde, creando una reserva de resto de direccionamiento de vida más corta que se podrá utilizar. En una característica deseable de la composición inventiva, esto crea un depósito de profármaco circulante que libera constantemente una cantidad profiláctica del resto de direccionamiento, liberándose unas cantidades más altas durante una respuesta inflamatoria.

Ejemplo 56: Estabilidad en suero del XTEN

Una proteína de fusión que contenía XTEN_AE864 fusionada con el extremo amínico de la GFP se incubó en plasma de mono y lisado renal de rata durante hasta 7 días a 37 °C. Las muestras se retiraron en el momento 0, día 1 y día 7, y se analizaron mediante una SDS PAGE, seguido de una detección utilizando una inmunoelectrotransferencia y una detección con anticuerpos contra la GFP, como se muestra en la FIG. 16. La secuencia de XTEN_AE864 mostró signos insignificantes de degradación durante 7 días en plasma. Sin embargo, XTEN_AE864 se degradó rápidamente en el lisado renal de rata durante 3 días. La estabilidad *in vivo* de la proteína de fusión se ensayó en muestras de plasma en donde se inmunoprecipitó la GFP_AE864 y se analizó mediante una SDS PAGE como se ha descrito anteriormente. Las muestras que se retiraron hasta 7 días después de la inyección mostraron muy pocos signos de degradación. Los resultados demuestran la resistencia de la proteína de fusión de unión a la degradación debida a proteasas séricas; un factor en la mejora de las propiedades farmacocinéticas de las proteínas de fusión de unión.

Ejemplo 57: Análisis de secuencias para la estructura secundaria mediante algoritmos de predicción

La estructura secundaria de las secuencias de aminoácidos se puede evaluar a través de ciertos programas informáticos o algoritmos, tales como el conocido algoritmo de Chou-Fasman (Chou, P. Y., et al. (1974) Biochemistry, 13: 222-45) y el método de Garnier-Osguthorpe-Robson, o "GOR" (Garnier J, Gibrat JF, Robson B. (1996). Método de GOR para predecir la estructura secundaria de la proteína a partir de la secuencia de aminoácidos. Methods Enzymol 266: 540-553). Para una secuencia dada, los algoritmos pueden predecir si existe alguna o ninguna estructura secundaria en absoluto, expresada como los restos totales y/o el porcentaje de la secuencia que forma, por ejemplo, hélices alfa o láminas beta, o el porcentaje de restos de la secuencia que se predice que dará como resultado la formación de espirales aleatorias.

Se han evaluado varias secuencias representativas de "familias" de XTEN utilizando dos herramientas de algoritmo para los métodos de Chou-Fasman y de GOR, para evaluar el grado de estructura secundaria en estas secuencias. La herramienta de Chou-Fasman fue proporcionada por William R. Pearson y la Universidad de Virginia, en el sitio de Internet "Biosupport", URL ubicada en la WWW en .fasta.bioch.virginia.edu/fasta_www2/fasta_www.cgi?rm=misc1 tal como existía el 19 de junio de 2009. La herramienta de GOR fue proporcionada por Pole Informatique Lyonnais en el sitio de Internet Network Protein Sequence Analysis, URL ubicada en la WWW en .npsa-pbil.ibcp.fr/cgi-bin/secpred_gor4.pl tal como existía el 19 de junio de 2008.

Como primera etapa en los análisis, se analizó una única secuencia del XTEN mediante los dos algoritmos. La composición AE864 es un XTEN con 864 restos de aminoácidos creados a partir de múltiples copias de cuatro motivos de secuencia de 12 aminoácidos que consisten en los aminoácidos G, S, T, E, P y A. Los motivos de la secuencia se caracterizan por el hecho de que hay una repetitividad limitada dentro de los motivos y dentro de la secuencia global, porque la secuencia de dos aminoácidos consecutivos cualesquiera no se repite más de dos veces en un motivo de 12 aminoácidos cualquiera, y porque no hay tres aminoácidos contiguos completos en el XTEN que sean idénticos. Se analizaron sucesivamente porciones más largas de la secuencia de AF 864 del extremo amínico mediante los algoritmos de Chou-Fasman y de GOR (este último requiere una longitud mínima de 17 aminoácidos). Las secuencias se analizaron introduciendo las secuencias de formato FASTA en las herramientas de predicción y ejecutando el análisis. Los resultados de los análisis se presentan en la tabla 30.

Los resultados indican que, según los cálculos de Chou-Fasman, los XTEN cortos de las familias AE y AG, de hasta al menos 288 restos de aminoácidos, no tienen hélices alfa ni láminas beta, pero las cantidades del porcentaje predicho de espirales aleatorias por el algoritmo de GOR varían del 78-99 %. Al aumentar las longitudes de los XTEN de 504 restos a más de 1300, el XTEN analizado mediante el algoritmo de Chou-Fasman había predicho unos porcentajes de hélices alfa o láminas beta de 0 a aproximadamente un 2 %, mientras que los porcentajes calculados de espirales aleatorias aumentaron hasta a partir del 94-99 %. Los XTEN con hélices alfa o láminas beta eran aquellas secuencias con una o más instancias de tres restos de serina contiguos, lo que dio como resultado la formación predicha de láminas beta. Sin embargo, incluso estas secuencias todavía tenían aproximadamente un 99 % de formación de espirales aleatorias.

El análisis respalda la conclusión de que: 1) se predice que los XTEN creados a partir de múltiples motivos de secuencia de G, S, T, E, P y A que tienen una repetitividad limitada en cuanto a aminoácidos contiguos tienen unas cantidades muy bajas de hélices alfa y láminas beta; 2) que al aumentar la longitud del XTEN no aumenta apreciablemente la probabilidad de formación de hélices alfa o láminas beta; y 3) que al aumentar progresivamente la longitud de la secuencia del XTEN mediante la adición de 12-meros no repetitivos que consisten en los aminoácidos G, S, T, E, P y A da como resultado un aumento del porcentaje de formación de espirales aleatorias. Tomando como base las numerosas secuencias evaluadas mediante estos métodos, se concluye que se espera que los XTEN creados a partir de los motivos de secuencia de G, S, T, E, P y A que tienen una repetitividad limitada (definida como no más de dos aminoácidos contiguos idénticos en un motivo cualquiera) tengan una estructura secundaria muy limitada. Con la excepción de los motivos que contienen tres serinas contiguas, se cree que puede usarse cualquier orden o combinación de motivos de secuencia de la tabla 3 para crear un polipéptido XTEN que dará como resultado una secuencia de XTEN que estará sustancialmente desprovista de estructura secundaria, y que los efectos de tres serinas

contiguas se mejoran aumentando la longitud del XTEN. Se espera que dichas secuencias tengan las características descritas en las realizaciones CFXTEN de la invención divulgadas en el presente documento.

Tabla 30: Cálculos predictivos de CHOU-FASMAN y GOR de secuencias de polipéptidos

NOMBRE DE LA SECUENCIA	Secuencia	SEQ ID NO:	N.º de restos	Cálculo de Chou-Fasman	Cálculo de GOR
AE36: LCW0402_002	GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAP	754	36	Total de restos: H: 0 E: 0	94,44 %
				porcentaje: H: 0,0 E: 0,0	
AE36: LCW0402_003	GTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAP	755	36	Total de restos: H: 0 E: 0	94,44 %
				porcentaje: H: 0,0 E: 0,0	
AG36: LCW0404_001	GASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSP	756	36	Total de restos: H: 0 E: 0	77,78 %
				porcentaje: H: 0,0 E: 0,0	
AG36: LCW0404_003	GSSTPSGATGSPGSSPSASTGTGPGSSTPSGATGSP	757	36	Total de restos: H: 0 E: 0	83,33 %
				porcentaje: H: 0,0 E: 0,0	
AE42_1	TEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGS	758	42	Total de restos: H: 0 E: 0	90,48 %
				porcentaje: H: 0,0 E: 0,0	
AE42_1	TEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGS	759	42	Total de restos: H: 0 E: 0	90,48 %
				porcentaje: H: 0,0 E: 0,0	
AG42_1	GAPSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGPSGP	760	42	Total de restos: H: 0 E: 0	88,10 %
				porcentaje: H: 0,0 E: 0,0	
AG42_2	GPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGASP	761	42	Total de restos: H: 0 E: 0	88,10 %
				porcentaje: H: 0,0 E: 0,0	
AE144		762	144	Total de restos: H: 0 E: 0	98,61 %

NOMBRE DE LA SECUENCIA	Secuencia	SEQ ID NO:	N.º de restos	Cálculo de Chou-Fasman	Cálculo de GOR
	GSEPATSGSETPGTSESATP ESGPGSEPATSGSETPGSPA GSPTSTEEGTSTEPSEGSAP GSEPATSGSETPGSEPATSG SETPGSEPATSGSETPGTST EPSEGSAPGTSESATPESGP GSEPATSGSETPGTSTEPSE GSAP			porcentaje: H: 0,0 E: 0,0	
AG144_1	PGSSPSASTGTGPGSSPSAS TGTGPGTPGSGTASSSPGSS TPSGATGSPGSSPSASTGTG PGASPGTSSTGSPGTPGSGT ASSSPGSSTPSGATGSPGTP GSGTASSSPGASPGTSSTGS PGASPGTSSTGSPGTPGSGT ASSS	763	144	Total de restos: H: 0 E: 0	91,67 %
				porcentaje: H: 0,0 E: 0,0	
AE288	GTSESATPESGPGSEPATSG SETPGTSESATPESGPGSEP ATSGSETPGTSESATPESGP GTSTEPSEGSAPGSPAGSPT STEEGTSESATPESGPGSEP ATSGSETPGTSESATPESGP GSPAGSPTSTEEGSPAGSPT STEEGTSTEPSEGSAPGTSE SATPESGPGTSESATPESGP GTSESATPESGPGSEPATSG SETPGSEPATSGSETPGSPA GSPTSTEEGTSTEPSEGSAP GTSTEPSEGSAPGSEPATSG SETPGTSESATPESGPGTST EPSEGSAP	764	288	Total de restos: H: 0 E: 0	99,31 %
				porcentaje: H: 0,0 E: 0,0	
AG288_2	GSSPSASTGTGPGSSPSAST GTGPGTPGSGTASSSPGSST PSGATGSPGSSPSASTGTGP GASPGTSSTGSPGTPGSGT ASSSPGSSTPSGATGSPGTP GSGTASSSPGASPGTSSTGS PGASPGTSSTGSPGTPGSGT ASSSPGSSTPSGATGSPGAS PGTSSTGSPGTPGSGTASSS PGSSTPSGATGSPGSSPSAS TGTGPGSSPSASTGTGPGSS TPSGATGSPGSSTPSGATGS PGASPGTSSTGSPGASPGTS STGSPGASPGTSSTGSPGTP GSGTASSSP	765	288	Total de restos: H: 0 E: 0	92,71
				porcentaje: H: 0,0 E: 0,0	
AF504		766	504	Total de restos: H: 0 E: 0	94,44 %

NOMBRE DE LA SECUENCIA	Secuencia	SEQ ID NO:	N.º de restos	Cálculo de Chou-Fasman	Cálculo de GOR
	GASPGTSSTGSPGSSPSAST GTGPGSSPSASTGTGPGTP GSGTASSSPGSSTPSGATGS PGSNPSASTGTGPGASPGTS STGSPGTPGSGTASSSPGSS TPSGATGSPGTPGSGTASSS PGASPGTSSTGSPGASPGTS STGSPGTPGSGTASSSPGSS TPSGATGSPGASPGTSSTGS PGTPGSGTASSSPGSSTPSG ATGSPGSNPSASTGTGPGSS PSASTGTGPGSSTPSGATGS PGSSTPSGATGSPGASPGTS STGSPGASPGTSSTGSPGAS PGTSSTGSPGTPGSGTASSS PGASPGTSSTGSPGASPGTS STGSPGASPGTSSTGSPGSS PSASTGTGPGTPGSGTASSS PGASPGTSSTGSPGASPGTS STGSPGASPGTSSTGSPGSS TPSGATGSPGSSTPSGATGS PGASPGTSSTGSPGTPGSGT ASSSPGSSTPSGATGSPGSS TPSGATGSPGSSTPSGATGS PGSSPSASTGTGPGASPGTS STGSP			porcentaje: H: 0,0 E: 0,0	
AD 576		767	576	Total de restos: H: 7 E: 0	99,65 %
	GSSESGSSEGGPGSGGEPSE SGSSGSSESGSSEGGPGSSE SGSSEGGPGSSESGSSEGGP GSSESGSSEGGPGSSESGSS EGGPGESPGGSSGSESGSE GSSGPGESSGSSESGSSEGG PGSSESGSSEGGPGSSESGS SEGGPGSGGEPSESGSSGES PGGSSGSESGESPGGSSGSE SGSGGEPSESGSSGSSESGS SEGGPGSGGEPSESGSSGSG GEPSESGSSGSEGSSGPGES SGESPGGSSGSESGSGGEPS ESGSSGSGGEPSESGSSGSG GEPSESGSSGSSESGSSEGG PGESPGGSSGSESGESPGGS SGSESGESPGGSSGSESGES PGGSSGSESGESPGGSSGSE SGSSESGSSEGGPGSGGEPS ESGSSGSEGSSGPGESSGSS ESGSSEGGPGSGGEPSESGS SGSSESGSSEGGPGSGGEPS ESGSSGESPGGSSGSESGES PGGSSGSESGSSESGSSEGG PGSGGEPSESGSSGSSESGS SEGGPGSGGEPSESGSSGSG GEPSESGSSGESPGGSSGSE SGSEGSSGPGESSGSSESGS SEGGPGSEGSSGPGESS			porcentaje: H: 1,2 E: 0,0	
AE576		768	576	Total de restos: H: 2 E: 0	99,65 %

NOMBRE DE LA SECUENCIA	Secuencia	SEQ ID NO:	N.º de restos	Cálculo de Chou-Fasman	Cálculo de GOR
	GSPAGSPTSTEEGTSESATP ESGPGTSTEPSEGSAPGSPA			porcentaje: H: 0,4 E: 0,0	
	GSPTSTEEGTSTEPSEGSAP GTSTEPSEGSAPGTSESATP ESGPGSEPATSGSETPGSEP ATSGSETPGSPAGSPTSTEE GTSESATPESGPGTSTEPSE GSAPGTSTEPSEGSAPGSPA GSPTSTEEGTSTEPSEGSAP GTSTEPSEGSAPGTSESATP ESGPGTSTEPSEGSAPGTSE SATPESGPGSEPATSGSETP GTSTEPSEGSAPGTSTEPSE GSAPGTSESATPESGPGTSE SATPESGPGSPAGSPTSTEE GTSESATPESGPGSEPATSG SETPGTSESATPESGPGTST EPSEGSAPGTSTEPSEGSAP GTSTEPSEGSAPGTSTEPSE GSAPGTSTEPSEGSAPGTST EPSEGSAPGSPAGSPTSTEE GTSTEPSEGSAPGTSESATP ESGPGSEPATSGSETPGTSE SATPESGPGSEPATSGSETP GTSESATPESGPGTSTEPSE GSAPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEE GSPAGSPTSTEEGTSESATP ESGPGTSTEPSEGSAP				
AG576		769	576	Total de restos: H: 0 E: 3	99,31 %

NOMBRE DE LA SECUENCIA	Secuencia	SEQ ID NO:	N.º de restos	Cálculo de Chou-Fasman	Cálculo de GOR
	PGTPGSGTASSSPGSSTPSG ATGSPGSSPSASTGTGPGSS PSASTGTGPGSSTPSGATGS PGSSTPSGATGSPGASPGTS STGSPGASPGTSSTGSPGAS PGTSSTGSPGTPGSGTASSS PGASPGTSSTGSPGASPGTS STGSPGASPGTSSTGSPGSS PSASTGTGPGTPGSGTASSS PGASPGTSSTGSPGASPGTS STGSPGASPGTSSTGSPGSS TPSGATGSPGSSTPSGATGS PGASPGTSSTGSPGTPGSGT ASSSPGSSTPSGATGSPGSS TPSGATGSPGSSTPSGATGS PGSSPSASTGTGPGASPGTS STGSPGASPGTSSTGSPGTP GSGTASSSPGASPGTSSTGS PGASPGTSSTGSPGASPGTS STGSPGASPGTSSTGSPGTP GSGTASSSPGSSTPSGATGS PGTPGSGTASSSPGSSTPSG ATGSPGTPGSGTASSSPGSS TPSGATGSPGSSTPSGATGS PGSSPSASTGTGPGSSPSAS TGTGPGASPGTSSTGSPGTP GSGTASSSPGSSTPSGATGS PGSSPSASTGTGPGSSPSAS TGTGPGASPGTSSTGS			porcentaje: H: 0,4 E: 0,5	
AF540	GSTSSTAESPGPGSTSSTAE SPGPGSTSESPSGTAPGSTS	770	540	Total de restos: H: 2 E: 0	99,65

NOMBRE DE LA SECUENCIA	Secuencia	SEQ ID NO:	N.º de restos	Cálculo de Chou-Fasman	Cálculo de GOR
	STAESPGPGSTSSTAESPGP GTSTPESGSASPGSTSESPS GTAPGTSPSGESSTAPGSTS ESPSGTAPGSTSESPSGTAP GTSPSGESSTAPGSTSESPS GTAPGSTSESPSGTAPGTSP SGESSTAPGSTSESPSGTAP GSTSESPSGTAPGSTSESPS GTAPGTSTPESGSASPGSTS ESPSGTAPGTSTPESGSASP GSTSSTAESPGPGSTSSTAE SPGPGTSTPESGSASPGTST PESGSASPGSTSESPSGTAP GTSTPESGSASPGTSTPESG SASPGSTSESPSGTAPGSTS ESPSGTAPGSTSESPSGTAP GSTSSTAESPGPGTSTPESG SASPGTSTPESGSASPGSTS ESPSGTAPGSTSESPSGTAP GTSTPESGSASPGSTSESPS GTAPGSTSESPSGTAPGTST PESGSASPGTSPSGESSTAP GSTSSTAESPGPGTSPSGES STAPGSTSSTAESPGPGTST PESGSASPGSTSESPSGTAP			porcentaje: H: 0,4 E: 0,0	
AD836		771	836	Total de restos: H: 0 E: 0	98,44 %

NOMBRE DE LA SECUENCIA	Secuencia	SEQ ID NO:	N.º de restos	Cálculo de Chou-Fasman	Cálculo de GOR
	GSSESGSSEGGPGSSESGSS EGGPGESPGGSSGSESGSG GEPSESGSSGESPGGSSGSE SGESPGGSSGSESGSSESGS SEGGPGSSESGSSEGGPGSS ESGSSEGGPGESPGGSSGSE SGESPGGSSGSESGESPGGS SGSESGSSESGSSEGGPGSS ESGSSEGGPGSSESGSSEGG PGSSESGSSEGGPGSSESGS SEGGPGSSESGSSEGGPGSG GEPSESGSSGESPGGSSGSE SGESPGGSSGSESGSGGEPS ESGSSGSEGSSGPGESSGSS ESGSSEGGPGSGGEPSESGS SGSEGSSGPGESSGSSESGS SEGGPGSGGEPSESGSSGES PGGSSGSESGSGGEPSESGS SGSGGEPSESGSSGSSESGS SEGGPGSGGEPSESGSSGSG GEPSESGSSGSEGSSGPGES SGESPGGSSGSESGSEGSSG PGESSGSEGSSGPGESSGSG GEPSESGSSGSSESGSSEGG PGSSESGSSEGGPGESPGGS SGSESGSGGEPSESGSSGSE GSSGPGESSGESPGGSSGSE SGSEGSSGPGSSESGSSEGG PGSGGEPSESGSSGSEGSSG PGESSGSEGSSGPGESSGSE GSSGPGESSGSGGEPSESGS SGSGGEPSESGSSGESPGGS SGSESGESPGGSSGSESGSG			porcentaje: H: 0,0 E: 0,0	
	GEPSESGSSGSEGSSGPGES SGESPGGSSGSESGSSESGS SEGGPGSSESGSSEGGPGSS ESGSSEGGPGSGGEPSESGS SGSSESGSSEGGPGESPGGS SGSESGSGGEPSESGSSGSS ESGSSEGGPGESPGGSSGSE SGSGGEPSESGSSGESPGGS SGSESGSGGEPSESGSS				
AE864		772	864	Total de restos: H: 2 E: 3	99,77 %

NOMBRE DE LA SECUENCIA	Secuencia	SEQ ID NO:	N.º de restos	Cálculo de Chou-Fasman	Cálculo de GOR
	GSPAGSPTSTEEGTSESATP ESGPGTSTEPSEGSAPGSPA GSPTSTEEGTSTEPSEGSAP GTSTEPSEGSAPGTSESATP ESGPGSEPATSGSETPGSEP ATSGSETPGSPAGSPTSTEE GTSESATPESGPGTSTEPSE GSAPGTSTEPSEGSAPGSPA GSPTSTEEGTSTEPSEGSAP GTSTEPSEGSAPGTSESATP ESGPGTSTEPSEGSAPGTSE SATPESGPGSEPATSGSETP GTSTEPSEGSAPGTSTEPSE GSAPGTSESATPESGPGTSE SATPESGPGSPAGSPTSTEE GTSESATPESGPGSEPATSG SETPGTSESATPESGPGTST EPSEGSAPGTSTEPSEGSAP GTSTEPSEGSAPGTSTEPSE GSAPGTSTEPSEGSAPGTST EPSEGSAPGSPAGSPTSTEE GTSTEPSEGSAPGTSESATP ESGPGSEPATSGSETPGTSE SATPESGPGSEPATSGSETP GTSESATPESGPGTSTEPSE GSAPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEE GSPAGSPTSTEEGTSESATP ESGPGTSTEPSEGSAPGTSE SATPESGPGSEPATSGSETP GTSESATPESGPGSEPATSG SETPGTSESATPESGPGTST EPSEGSAPGSPAGSPTSTEE GTSESATPESGPGSEPATSG SETPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEE GTSTEPSEGSAPGTSESATP ESGPGTSESATPESGPGTSE SATPESGPGSEPATSGSETP GSEPATSGSETPGSPAGSPT STEEGTSTEPSEGSAPGTST EPSEGSAPGSEPATSGSETP GTSESATPESGPGTSTEPSE GSAP			porcentaje: H: 0,2 E: 0,4	
AF864	GSTSESPSGTAPGTSPSGES STAPGSTSESPSGTAPGSTS ESPSGTAPGTSTPESGSASP GTSTPESGSASPGSTSESPS GTAPGSTSESPSGTAPGTSP	773	875	Total de restos: H: 2 E: 0	95,20 %
				porcentaje: H: 0,2 E: 0,0	

NOMBRE DE LA SECUENCIA	Secuencia	SEQ ID NO:	N.º de restos	Cálculo de Chou-Fasman	Cálculo de GOR
	SGESSTAPGSTSESPSGTAP GTSPSGESSTAPGTSPSGES STAPGSTSSTAESPGPGTSP SGESSTAPGTSPSGESSTAP GSTSSTAESPGPGTSTPESG SASPGTSTPESGSASPGSTS ESPSGTAPGSTSESPSGTAP GTSTPESGSASPGSTSSTAE SPGPGTSTPESGSASPGSTS ESPSGTAPGTSPSGESSTAP GSTSSTAESPGPGTSPSGES STAPGTSTPESGSASPGSTS STAESPGPGSTSSTAESPGP GSTSSTAESPGPGSTSSTAE SPGPGTSPSGESSTAPGSTS ESPSGTAPGSTSESPSGTAP GTSTPESGPXXXGASASGA PSTXXXXSESPSGTAPGSTS ESPSGTAPGSTSESPSGTAP GSTSESPSGTAPGSTSESPS GTAPGSTSESPSGTAPGTST PESGSASPGTSPSGESSTAP GTSPSGESSTAPGSTSSTAE SPGPGTSPSGESSTAPGTST PESGSASPGSTSESPSGTAP GSTSESPSGTAPGTSPSGES STAPGSTSESPSGTAPGTST PESGSASPGTSTPESGSASP GSTSESPSGTAPGTSTPESG SASPGSTSSTAESPGPGSTS ESPSGTAPGSTSESPSGTAP GTSPSGESSTAPGSTSSTAE SPGPGTSPSGESSTAPGTST PESGSASPGTSPSGESSTAP GTSPSGESSTAPGTSPSGES STAPGSTSSTAESPGPGSTS STAESPGPGTSPSGESSTAP GSSPSASTGTGPGSSTPSGA TGSPGSSTPSGATGSP				
AG864		774	864	Total de restos: H: 0 E: 0	94,91 %

NOMBRE DE LA SECUENCIA	Secuencia	SEQ ID NO:	N.º de restos	Cálculo de Chou-Fasman	Cálculo de GOR
	GASPGTSSTGSPGSSPSAST GTGPGSSPSASTGTGPGTP GSGTASSSPGSSTPSGATGS PGSSPSASTGTGPGASPGTS STGSPGTPGSGTASSSPGSS TPSGATGSPGTPGSGTASSS PGASPGTSSTGSPGASPGTS STGSPGTPGSGTASSSPGSS TPSGATGSPGASPGTSSTGS PGTPGSGTASSSPGSSTPSG ATGSPGSSPSASTGTGPGSS PSASTGTGPGSSTPSGATGS PGSSTPSGATGSPGASPGTS STGSPGASPGTSSTGSPGAS PGTSSTGSPGTPGSGTASSS PGASPGTSSTGSPGASPGTS STGSPGASPGTSSTGSPGSS PSASTGTGPGTPGSGTASSS PGASPGTSSTGSPGASPGTS			porcentaje: H: 0,0 E: 0,0	
	STGSPGASPGTSSTGSPGSS TPSGATGSPGSSTPSGATGS PGASPGTSSTGSPGTPGSGT ASSSPGSSTPSGATGSPGSS TPSGATGSPGSSTPSGATGS PGSSPSASTGTGPGASPGTS STGSPGASPGTSSTGSPGTP GSGTASSSPGASPGTSSTGS PGASPGTSSTGSPGASPGTS STGSPGASPGTSSTGSPGTP GSGTASSSPGSSTPSGATGS PGTPGSGTASSSPGSSTPSG ATGSPGTPGSGTASSSPGSS TPSGATGSPGSSTPSGATGS PGSSPSASTGTGPGSSPSAS TGTGPGASPGTSSTGSPGTP GSGTASSSPGSSTPSGATGS PGSSPSASTGTGPGSSPSAS TGTGPGASPGTSSTGSPGA SPGTSSTGSPGSSTPSGATG SPGSSPSASTGTGPGASPGT SSTGSPGSSPSASTGTGPGT PGSGTASSSPGSSTPSGATG SPGSSTPSGATGSPGASPGT SSTGSP				
AM875		775	875	Total de restos: H: 7 E: 3	98,63 %

NOMBRE DE LA SECUENCIA	Secuencia	SEQ ID NO:	N.º de restos	Cálculo de Chou-Fasman	Cálculo de GOR
	GTSTEPSEGSAPGSEPATSG SETPGSPAGSPTSTEEGSTS STAESPGPGTSTPESGSASP GSTSESPSGTAPGSTSESPS GTAPGTSTPESGSASPGTST PESGSASPGSEPATSGSETP GTSESATPESGPGSPAGSPT STEEGTSTEPSEGSAPGTSE SATPESGPGTSTEPSEGSAP GTSTEPSEGSAPGSPAGSPT STEEGTSTEPSEGSAPGTST EPSEGSAPGTSESATPESGP GTSESATPESGPGTSTEPSE GSAPGTSTEPSEGSAPGTSE SATPESGPGTSTEPSEGSAP GSEPATSGSETPGSPAGSPT STEEGSSTPSGATGSPGTPG SGTASSSPGSSTPSGATGSP GTSTEPSEGSAPGTSTEPSE GSAPGSEPATSGSETPGSPA GSPTSTEEGSPAGSPTSTEE GTSTEPSEGSAPGASASGA PSTGGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEE GSTSSTAESPGPGSTSESPS GTAPGTSPSGESSTAPGTPG SGTASSSPGSSTPSGATGSP GSSPSASTGTGPGSEPATSG SETPGTSESATPESGPGSEP ATSGSETPGSTSSTAESPGP GSTSSTAESPGPGTSPSGES STAPGSEPATSGSETPGSEP ATSGSETPGTSTEPSEGSAP			porcentaje: H: 0,8 E: 0,3	
	GSTSSTAESPGPGTSTPESG SASPGSTSESPSGTAPGTST EPSEGSAPGTSTEPSEGSAP GTSTEPSEGSAPGSSTPSGA TGSPGSSPSASTGTGPGASP GTSSTGSPGSEPATSGSETP GTSESATPESGPGSPAGSPT STEEGSSTPSGATGSPGSSP SASTGTGPGASPGTSSTGSP GTSESATPESGPGTSTEPSE GSAPGTSTEPSEGSAP				
AM1318		776	1318	Total de restos: H: 7 E: 0	99,17 %

NOMBRE DE LA SECUENCIA	Secuencia	SEQ ID NO:	N.º de restos	Cálculo de Chou-Fasman	Cálculo de GOR
	GTSTEPSEGSAPGSEPATSG SETPGSPAGSPTSTEEGSTS STAESPGPGTSTPESGSASP GSTSESPSGTAPGSTSESPS GTAPGTSTPESGSASPGTST PESGSASPGSEPATSGSETP GTSESATPESGPGSPAGSPT STEEGTSTEPSEGSAPGTSE SATPESGPGTSTEPSEGSAP GTSTEPSEGSAPGSPAGSPT STEEGTSTEPSEGSAPGTST EPSEGSAPGTSESATPESGP GTSESATPESGPGTSTEPSE GSAPGTSTEPSEGSAPGTSE SATPESGPGTSTEPSEGSAP GSEPATSGSETPGSPAGSPT STEEGSSTPSGATGSPGTPG SGTASSSPGSSTPSGATGSP GTSTEPSEGSAPGTSTEPSE GSAPGSEPATSGSETPGSPA GSPTSTEEGSPAGSPTSTEE GTSTEPSEGSAPGPEPTGPA PSGGSEPATSGSETPGTSES ATPESGPGSPAGSPTSTEEG TSESATPESGPGSPAGSPTS TEEGSPAGSPTSTEEGTSES ATPESGPGSPAGSPTSTEEG SPAGSPTSTEEGSTSSTAES PGPGSTSESPSGTAPGTSPS GESSTAPGSTSESPSGTAPG STSESPSGTAPGTSPSGESS TAPGTSTEPSEGSAPGTSES ATPESGPGTSESATPESGPG SEPATSGSETPGTSESATPE SGPGTSESATPESGPGTSTE PSEGSAPGTSESATPESGPG TSTEPSEGSAPGTSPSGESS TAPGTSPSGESSTAPGTSPS GESSTAPGTSTEPSEGSAPG SPAGSPTSTEEGTSTEPSEG SAPGSSPSASTGTGPGSSTP SGATGSPGSSTPSGATGSPG SSTPSGATGSPGSSTPSGAT GSPGASPGTSSTGSPGASAS GAPSTGGTSPSGESSTAPGS TSSTAESPGPGTSPSGESST APGTSESATPESGPGTSTEP			porcentaje: H: 0,7 E: 0,0	

NOMBRE DE LA SECUENCIA	Secuencia	SEQ ID NO:	N.º de restos	Cálculo de Chou-Fasman	Cálculo de GOR
	SEGSAPGTSTEPSEGSAPGS SPSASTGTGPGSSTPSGATG SPGASPGTSSTGSPGTSTPE SGSASPGTSPSGESSTAPGT SPSGESSTAPGTSESATPES GPGSEPATSGSETPGTSTEP SEGSAPGSTSESPSGTAPGS TSESPSGTAPGTSTPESGSA SPGSPAGSPTSTEEGTSESA TPESGPGTSTEPSEGSAPGS PAGSPTSTEEGTSESATPES GPGSEPATSGSETPGSSTPS GATGSPGASPGTSSTGSPGS STPSGATGSPGSTSESPSGT APGTSPSGESSTAPGSTSST AESPGPGSSTPSGATGSPGA SPGTSSTGSPGTPGSGTASS SPGSPAGSPTSTEEGSPAGS PTSTEEGTSTEPSEGSAP				
AM923		777	924	Total de restos: H: 4 E: 3	98,70 %

NOMBRE DE LA SECUENCIA	Secuencia	SEQ ID NO:	N.º de restos	Cálculo de Chou-Fasman	Cálculo de GOR
	MAEPAGSPTSTEEGASPGT SSTGSPGSSTPSGATGSPGS STPSGATGSPGTSTEPSEGS APGSEPATSGSETPGSPAGS PTSTEEGSTSSTAESPGPGT STPESGSASPGSTSESPSGT APGSTSESPSGTAPGTSTPE SGSASPGTSTPESGSASPGS EPATSGSETPGTSESATPES GPGSPAGSPTSTEEGTSTEP SEGSAPGTSESATPESGPGT STEPSEGSAPGTSTEPSEGS APGSPAGSPTSTEEGTSTEP SEGSAPGTSTEPSEGSAPGT SESATPESGPGTSESATPES GPGTSTEPSEGSAPGTSTEP SEGSAPGTSESATPESGPGT STEPSEGSAPGSEPATSGSE TPGSPAGSPTSTEEGSSTPS GATGSPGTPGSGTASSSPGS STPSGATGSPGTSTEPSEGS APGTSTEPSEGSAPGSEPAT SGSETPGSPAGSPTSTEEGS PAGSPTSTEEGTSTEPSEGS APGASASGAPSTGGTSESA TPESGPGSPAGSPTSTEEGS PAGSPTSTEEGSTSSTAESP GPGSTSESPSGTAPGTSPSG ESSTAPGTPGSGTASSSPGS STPSGATGSPGSSPSASTGT GPGSEPATSGSETPGTSESA TPESGPGSEPATSGSETPGS TSSTAESPGPGSTSSTAESP GPGTSPSGESSTAPGSEPAT SGSETPGSEPATSGSETPGT STEPSEGSAPGSTSSTAESP GPGTSTPESGSASPGSTSES PSGTAPGTSTEPSEGSAPGT STEPSEGSAPGTSTEPSEGS			porcentaje: H: 0,4 E: 0,3	
	APGSSTPSGATGSPGSSPSA STGTGPGASPGTSSTGSPGS EPATSGSETPGTSESATPES GPGSPAGSPTSTEEGSSTPS GATGSPGSSPSASTGTGPG ASPGTSSTGSPGTSESATPE SGPGTSTEPSEGSAPGTSTE PSEGSAP				
AE912		778	913	Total de restos: H: 8 E: 3	99,45 %

NOMBRE DE LA SECUENCIA	Secuencia	SEQ ID NO:	N.º de restos	Cálculo de Chou-Fasman	Cálculo de GOR
	MAEPAGSPTSTEEGTPGSG TASSSPGSSTPSGATGSPGA SPGTSSTGSPGSPAGSPTST EEGTSESATPESGPGTSTEP SEGSAPGSPAGSPTSTEEGT STEPSEGSAPGTSTEPSEGS APGTSESATPESGPGSEPAT SGSETPGSEPATSGSETPGS PAGSPTSTEEGTSESATPES GPGTSTEPSEGSAPGTSTEP SEGSAPGSPAGSPTSTEEGT STEPSEGSAPGTSTEPSEGS APGTSESATPESGPGTSTEP SEGSAPGTSESATPESGPGS EPATSGSETPGTSTEPSEGS APGTSTEPSEGSAPGTSESA TPESGPGTSESATPESGPGS PAGSPTSTEEGTSESATPES GPGSEPATSGSETPGTSESA TPESGPGTSTEPSEGSAPGT STEPSEGSAPGTSTEPSEGS APGTSTEPSEGSAPGTSTEP SEGSAPGTSTEPSEGSAPGS PAGSPTSTEEGTSTEPSEGS APGTSESATPESGPGSEPAT SGSETPGTSESATPESGPGS EPATSGSETPGTSESATPES GPGTSTEPSEGSAPGTSESA TPESGPGSPAGSPTSTEEGS PAGSPTSTEEGSPAGSPTST EEGTSESATPESGPGTSTEP SEGSAPGTSESATPESGPGS EPATSGSETPGTSESATPES GPGSEPATSGSETPGTSESA TPESGPGTSTEPSEGSAPGS PAGSPTSTEEGTSESATPES GPGSEPATSGSETPGTSESA TPESGPGSPAGSPTSTEEGS PAGSPTSTEEGTSTEPSEGS APGTSESATPESGPGTSESA TPESGPGTSESATPESGPGS EPATSGSETPGSEPATSGSE TPGSPAGSPTSTEEGTSTEP SEGSAPGTSTEPSEGSAPGS EPATSGSETPGTSESATPES GPGTSTEPSEGSAP			porcentaje: H: 0,9 E: 0,3	
BC 864		779		Total de restos: H: 0 E: 0	99,77 %
	GTSTEPSEPGSAGTSTEPSE PGSAGSEPATSGTEPSGSG ASEPTSTEPGSEPATSGTEP			porcentaje: H: 0 E: 0	

NOMBRE DE LA SECUENCIA	Secuencia	SEQ ID NO:	N.º de restos	Cálculo de Chou-Fasman	Cálculo de GOR
	SGSEPATSGTEPSGSEPATS GTEPSGSGASEPTSTEPGTS TEPSEPGSAGSEPATSGTEP SGTSTEPSEPGSAGSEPATS GTEPSGSEPATSGTEPSGTS TEPSEPGSAGTSTEPSEPGS AGSEPATSGTEPSGSEPATS GTEPSGTSEPSTSEPGAGSG ASEPTSTEPGTSEPSTSEPG AGSEPATSGTEPSGSEPATS GTEPSGTSTEPSEPGSAGTS TEPSEPGSAGSGASEPTSTE PGSEPATSGTEPSGSEPATS GTEPSGSEPATSGTEPSGSE PATSGTEPSGTSTEPSEPGS AGSEPATSGTEPSGSGASEP TSTEPGTSTEPSEPGSAGSE PATSGTEPSGSGASEPTSTE PGTSTEPSEPGSAGSGASEP TSTEPGSEPATSGTEPSGSG ASEPTSTEPGSEPATSGTEP SGSGASEPTSTEPGTSTEPS EPGSAGSEPATSGTEPSGSG ASEPTSTEPGTSTEPSEPGS AGSEPATSGTEPSGTSTEPS EPGSAGSEPATSGTEPSGTS TEPSEPGSAGTSTEPSEPGS AGTSTEPSEPGSAGTSTEPS EPGSAGTSTEPSEPGSAGTS TEPSEPGSAGTSEPSTSEPG AGSGASEPTSTEPGTSTEPS EPGSAGTSTEPSEPGSAGTS TEPSEPGSAGSEPATSGTEP SGSGASEPTSTEPGSEPATS GTEPSGSEPATSGTEPSGSE PATSGTEPSGSEPATSGTEP SGTSEPSTSEPGAGSEPATS GTEPSGSGASEPTSTEPGTS TEPSEPGSAGSEPATSGTEP SGSGASEPTSTEPGTSTEPS EPGSA				
* H: hélice alfa, E: lámina beta					

Ejemplo 58: Análisis de repetitividad de las secuencias de polipéptidos

La repetitividad de las secuencias de aminoácidos de los polipéptidos se puede evaluar cuantificando el número de veces que aparece una subsecuencia más corta dentro del polipéptido global. Por ejemplo, un polipéptido de 200 restos de aminoácidos tiene 192 subsecuencias de 9 aminoácidos solapadas (o "marcos" de 9-meros), pero el número de subsecuencias de 9-meros únicas dependerá de la cantidad de repetitividad dentro de la secuencia. En el presente análisis, se evaluó la repetitividad de diferentes secuencias sumando la aparición de todas las subsecuencias únicas de 3-meros para cada marco de 3 aminoácidos a lo largo de los primeros 200 aminoácidos de la porción de polímero divididas por el número absoluto de subsecuencias únicas de 3-meros dentro de la secuencia de 200 aminoácidos. La puntuación de la subsecuencia resultante es un reflejo del grado de repetitividad dentro del polipéptido. Las secuencias de la tabla 31 se analizaron mediante el algoritmo SegScore (FIG. 37), que aplica la ecuación I a los primeros 200 aminoácidos de un polipéptido. Los resultados, mostrados en la tabla 31, indican que los polipéptidos no estructurados que consisten en 2 o 3 tipos de aminoácidos tienen altas puntuaciones de subsecuencias, mientras que los que consisten en 12 motivos de aminoácidos de los seis aminoácidos G, S, T, E, P y A con un bajo grado de repetitividad interna tienen unas puntuaciones de subsecuencias inferiores a 10 y, en algunos casos, inferiores a 5. Por ejemplo, la secuencia L288 tiene dos tipos de aminoácidos y tiene secuencias cortas y muy repetitivas, lo que da como resultado

una puntuación de la subsecuencia de 50,0. El polipéptido J288 tiene tres tipos de aminoácidos, pero también tiene secuencias cortas y repetitivas, lo que da como resultado una puntuación de la subsecuencia de 33,3. Y576 también tiene tres tipos de aminoácidos, pero no está hecho de repeticiones internas, reflejadas en la puntuación de la subsecuencia de 15,7 a lo largo de los primeros 200 aminoácidos. W576 consiste en cuatro tipos de aminoácidos, pero tiene un mayor grado de repetitividad interna, por ejemplo, "GGSG" (SEQ ID NO: 780), lo que da como resultado una puntuación de la subsecuencia de 23,4. El AD576 consiste en cuatro tipos de 12 motivos de aminoácidos, cada uno de los cuales consiste en cuatro tipos de aminoácidos. Debido al bajo grado de repetitividad interna de los motivos individuales, la puntuación global de la subsecuencia a lo largo de los primeros 200 aminoácidos es de 13,6. Por el contrario, los XTEN que consisten en cuatro motivos que contienen seis tipos de aminoácidos, cada uno con un bajo grado de repetitividad interna, tienen unas puntuaciones de la subsecuencia más bajas; es decir, AE864 (6,1), AF864 (7,5) y AM875 (4,5).

Conclusiones: Los resultados indican que la combinación de motivos de subsecuencias de 12 aminoácidos, consistente cada uno en cuatro a seis tipos de aminoácidos que son esencialmente no repetitivos, en un polipéptido XTEN más largo, da como resultado una secuencia global que no es repetitiva. Esto es a pesar del hecho de que cada motivo de subsecuencia puede usarse varias veces a lo largo de la secuencia. Por el contrario, los polímeros creados a partir de cifras más pequeñas de tipos de aminoácidos dieron como resultado puntuaciones de subsecuencias más altas, aunque la secuencia real se puede adaptar para reducir el grado de repetitividad para dar como resultado unas puntuaciones de subsecuencias más bajas.

Tabla 31: Cálculos de la puntuación de la subsecuencia de secuencias de polipéptidos

Nombre de la secuencia	Secuencia de aminoácidos	SEQ ID NO:	Puntuación
J288	GSGGEGGSGGGEGGGSGGGEGGGSGGGEGGGSGGEGGGSGGGEGGS GGEGGSGGEGGSGGEGGSGGGEGGSGGEGGSGGGEGGSGGEGGSGG EGGSGGEGGSGGGEGGGSGGEGGGSGGGEGGGSGGEGGSGGGEGGSGEG GSGGEGGSGGGEGGGSGGEGGGSGGGEGGGSGGGEGGGSGGGEGGS GGEGGGSGGEGGSGGGEGGGSGGEGGGSGGGEGGGSGGGEGGSGG EGGSGGEGGGSGGEGGGSGGEGGGSGGGEGGGSGGGEGGGSGGGEG	781	33,3
K288	GEGEGGGEGGGEGEGGGEGGGEGGGEGGGEGGGEGGGEGGGEGGGEGEG GGEGGEGEGGGEGGGEGGGEGGGEGGGEGGGEGGGEGGGEGGGEGGGEGG EGEGGGEGEGGGEGGGEGEGGGEGGGEGGGEGGGEGGGEGGGEGGGEGGG GEGGEGEGGGEGGGEGEGGGEGGGEGGGEGGGEGGGEGGGEGGGGE GEGGGEGGGEGGGGEGGGEGEGGGGEGGGEGGGEGGGEGGGEGGGGG EGGEGEGGGEGGGEGGGGEGGGEGGGEGGGEGGGEGGGEGGG	782	46,9
L288	SSESSESSSSESSESSESSSSESSESSESSESSESSESSESSESSSES SSESSESSSSESSESSESSSSESSESSESSESSESSESSESSESSSES SSESSESSSSESSESSESSSSESSESSESSESSESSESSESSESSSES SSESSESSSSESSESSESSSSESSESSESSESSESSESSESSESSSES SSESSESSSSESSESSESSSSESSESSESSESSESSESSESSESSSES	783	50,0
Y288		784	26,8

Nombre de la secuencia	Secuencia de aminoácidos	SEQ ID NO:	Puntuación
	GEGSGEGSGEGEGSGEGSGEGSGEGEGGGSGEGEGGGSGEGSGEGSGEG SEGSGSGEGSGEGGGSGEGSGEGSGEGEGSGEGEGGGSGEGSGEGSGE GSEGEGGGEGGSGEGEGSGEGSGEGSGEGSGEGEGSGEGSGEGSGEGS		
	EGSGEGEGSGEGSGEGEGGSGEGSGEGEGSGEGEGGGSGEGSGEGEGE GGGEGSGEGEGSGEGEGGGSGEGSGEGEGSGEGSGEGEGSGEGSGE GEGSGEGGGSGEGSGEGSGEGEGSGSGE		
O576	GGKPEGGKPEGGKPEGGGKPGGKPEGEGEGKPGGKPEGGKPGGGKPEGGKPEGG KPEGGSKPEGGEGKPGGKPEGEGKPEGGKPGGGGGKPGGKPEGEGKPE GGEGGKPEGGKPEGKPGGKPEGEGKPEGGKPGGGKPEGGKPEGGKPEG GGEGGKPEGGKPEGKPGGKPEGGKPEGGKPGGGKPEGGKPEGGKPE GGGGKPEGGGEGKPGGGKPGEGKPGGGKPEGGKPGGGKPEGGKPEGG GKPGGGGKPEGGKPEGKPGGKPEGGKPGGGKPEGGKPGGGKPEGGKPE GKPGGGKPGGGGKPGEGKPGGGKPEGGKPGGGKPEGGKPGGGKPEGG KPEGGKPGEGGKPGGKPEGGKPGGKPEGGKPGGGKPEGGKPGGGKPE GGGGKPGEGGKPGGKPEGGKPGGGKPEGGKPGGGKPEGGKPGGGKPG GKPGGGGKPEGGKPEGGKPGGGKPEGGKPGGGKPEGGKPGGGKPEG	785	18,5
U576	GEGKPGGKPGGGGKPGGKPEGGKPGGGKPGGKPGGGKPGGGKPEGGKP EGGKGGKPGGGGKPGGKPGGGEGGKPGGKPEGGKPGGKPEGGKPEG GGGGKPGGKPGGGKPGGGKPGGKPGGGKPGGKPEGGKPGGGKPEGGGG GKPGGKPEGGGGKPGGKPGGGKPGGKPGGGKPGGKPEGGKPGGGKP GGKPGGGKPGGGKPGGGKPGGGKPGGGKPGGGKPGGGKPGGGKPEGGGG GGKPGGGKPGGGKPGGGKPGGGKPGGGKPGGGKPGGGKPGGGKPEGGGG KPEGGKPGGGGKPGGKPGGGEGGKPGGGKPGGGKPGGGKPGGGKPE GGGGKPGGGKPGGKPGGGKPGGGKPGGGKPGGGKPGGGKPGGGKPG KPGGGGGKPGGGKPGGGKPGGGKPGGGKPGGGKPGGGKPGGGKPG KPEGGSKPGGKPGGSKPGGKPGGGKPGGGKPGGGKPGGGKPGGGKPG PG	786	16,1
W576	GGSGKPGKPGGSGSGKPGSGKPGGGSGKPGSGKPGGGSGKPGSGKPGGGSG KPGSGKPGGGKPGSGSGKPGGKPGGSGSGKPGKPGSGGGKPGSGKPG SGKPGGGSGGKPGKPGSGKPGGSGKPGGGSGKPGKPGSGGGKPGSGKPG GSGGKPGSGKPGGSGKPGSGKPGGGSGKPGSGKPGSGGGKPGSGKPG GSGGKPGSGKPGGSGKPGSGKPGGGSGKPGSGKPGSGGGKPGSGKPG PGSGSGKPGGGKPGSGSGKPGGGSGKPGSGKPGSGGGKPGSGKPGGGk GKPGSGGGSGKPGKPGSGSGKPGGGSGKPGSGKPGSGGGKPGSGKPG GSGGSGGKPGKPGSGGGKPGSGSGKPGGGKPGSGKPGSGGGKPGSGKPG GGKPGKPGGSGKPGSGKPGGGSGKPGSGKPGSGGGKPGSGKPGGGSG GKPGKPGGSGKPGSGKPGGSGKPGGGSGKPGSGKPGSGGGKPGSGKPG SGKPGGGKPGSGKPGGSGKPGKPGSGKPGSGGGKPGSGKPGSGKPG PGGGSGKPGKPGSGKPGGGKPGSGKPGSGGGKPGSGKPGSGKPG	787	23,4
Y576		788	15,7

Nombre de la secuencia	Secuencia de aminoácidos	SEQ ID NO:	Puntuación
	GEGSNEGSNEGEGSEGSKGDAEEGSKEGSKGDAEEGDSDSSEGEGDAGAEGS GEGEGSGEGSEGSEGEGGGEASEGEDSEGEGGSEGEGGSEGAEEGEGSF GSGEGDGSGSGGSEGSEGGSEGSEGSGEGEGSGEGEGSGSGSGAGEF SEGSGEGEGSEGSEGDAEGGSEGSDAEGEGGSDASEGDAGGEGSDAI EGSGEGSEGEGGSEGSEGEGGSEGSEGSGEGEGEGSGEGS EGEGSEGSGEGEGSEGSGEGSDAGEGSGEGEGSDASGEGDGSEGS GEGEGGSEGSEGSDAEGSDAEGSDAEGEGSEGSGEGEGSGEGEG EGSEGDAEGGDAEGSDAEGSDAEGSDAEGSDAEGSGEGEG EGSEGGSEGSEGSEGSGEGSEGSGEGSGEGEGSGEGEGSGEGEG GSE	788	13,6
AD576	GSSESGSSEGGPGSGGEPSESGSSGSSESAGSSEGGPGSSESGS SEGGPGSSESGSSEGGPGSSESGSSEGGPGESPGGSSGSESGSSGPGESSGS SESGSSEGGPGESSESGSSGGPGESPGGSSGSESGSSGPGESSGS SESGESPGGSSGSESGSGGPGESPSESGSSGSESGSSGPGESSGSSGGG EPSESGSSGSESGSGPGESSGSEPGGSSGSESGSSGPGESPSESGSSGGPGESSGS SGSGGEPSESGSSGSESGSSGSESGSSPGGSSGSESGSSGPGESSGSSGG SSGSESGESPGGSSGSESGESPGGSSGSESGSSEGGPGSSGGEEPSESGSSG	789	13,6
	SESGSSGPGESSGSSGSESGSSEGGQPGSGGEPSESGSSGSESEGGPGSSGGEPSE SGSSGESPGGSSGSESGESPGGSSGSESGSSEGGPGSSGGEPSESGSSGSS ESGSSEGSPGSGGEPSESGSSGSGGEPSESGSSGSESGGPGSSGGEPSESGSSGSS SSGSSESGSSEGGPGSGESGSSGPGESS	789	13,6
AE576	AGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEP SEGSAPGTSTEPSGSAPGTSESATPESGPGSEPATSGSETPGTSAGSETPG SPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPACGPT STEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTS ESATPESGPGSEPATSGSETPKGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPES GPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPATSGSETPGTSAES ATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAP GTSTEPSEGSAPGTSTEPSEGSAPGSPTSTEEGTSTEPSEGSAPGTSESATP PESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSEGPGT STEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTS TEEGTSESATPESGPGTSTEPSEGSAP	790	6,1
AF540	GSTSSTAESPGPGSTSSTAESPGPGSTSESPSGTAPGSTSSTAESPGPGSTSSTSAE SPGPGTSTPESGSASPGSTSESPSGTAPGTSPSGESSTAPGSTSESPSGTAPGTSPSGESSTAP ESPSGTAPGTSPSGESSTAPGSTSESPSGTAPGSTSESPSGTAPGTSPSGESSTAP GSTSESPSGTAPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGSTSESPSPGTA GTAPGTSTPESGSASPGSTSSTAESPGPGSTSSTAESPGPGTSTPESGSASPGTS TPESGSASPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPGSTSESPSGTA PGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPGTSTPES GSASPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGSTSESPSGTAPGST SESPSGTAPGTSTPESGSASPGTSPSGESSTAPGSTSSTAESPGPGTSPSGESSTA PGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAP	791	8,8

Nombre de la secuencia	Secuencia de aminoácidos	SEQ ID NO:	Puntuación
AF504	GASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGSSPSASTGTGPGSSPSASTGTGPG ASTGSPGSNPSASTGTGPGASPGTSSTGSPGTPGSHTASSPGTSPGQATGSPGSSPSASTGTGPG TPGSGTASSPGTASPGTSSTGSPGTASPGTSSTGSPGTPGSGTASSPGSSTPSGATG TGSPGASPGTSSTGSPGTPGSGTASSPGSSTPSGATGSPGSPSASTGTGPGSS PSASTGTGPGSSTPSGATGSPGTPSSTGSPGASPGTPSSTGSPGASPGTSSTG SPGASPGTSSTGSPGTPGSGTASSPGASPGTSSTGSPGASPGTSSTGSPGASPG TSSTGSPGSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSP GASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGA SPGTTSSTGSP	792	7,0
AE864	GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPS EGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGS PAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTS TEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGHGTSTEPSEGSAPGTSE SATPESKPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESG PGTSESATPESKPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESA TPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATP ESGPGSEPATSGSETPGTSESATPESGPGTSESATPESGPGTSESATPESGPGTS TEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGPSAGSPTSTEEGSPAGSPTST EEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPTSTEE GTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSP TSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGS EPATSGSETPGSEPATSGSETPGTSTEPSEGSAP	793	6,1
AF864	GSTSESPSGTAPGTSPSGESSTAPGSTSESPSGTAPGSTSESPSGTAPGSTSTPESG SASPGTSTPESGSASPGSTPESPSGTAPGSTSESPSGTAPGTSPSGESSTAPGSTS ESPSGTAPGTSPSGESSTAPGTSPSGESSTAPGSTSSTAESPSGPGTSPSGESSTAP GTSPSGESSTAPGSTSSTAESPGPGTSTPESGSASPGTSTPESGSASPGSTSESPS GTAPGSTSESPSGTAPGTSTPESGSASPGSTSSTAESPGPGTSTPESGSASPGST SESPSGTAPGTSPSGESSTAPGSTSSTAESPGPGTSPSGESSTAPGTSTPESGSAS PGSTSSTAESPGPGSTSSTAESPGPGSTSSTAESPGPGSTSSTAESPGPGTSPSGE SSTAPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGPXXXGASASGAPSTXXX	794	7,5
AG868	XSESPSGTAPGSTSESPSGTAPGSTSESPSGTAPGSTSESPSGTAPGSTSESPSGT APGSTSESPSGTAPGTSTPESGSASPGTSPSGESSTAPGTSPSGESSTAPGSTSST AESPGPGTSPSGESSTAPGTSTPESGSASPGSTSESPSGTAPGSTSESPSGTAPG TSPSGESSTAPGSTSESPSGTAPGTSTPESGSASPGSTSESPSGSASPGSTSESPSG TAPGTSTPESGSASPGSTSSTAESPGPGTSPSGESSTAPGSTSESPSGTAPGTSPS GESSTAPGSTSSTAESPGPGTSPSGESSTAPGTSTPESGSASPGTSPSGESSTAP GTSPSGESSTAPGTSPSGESSTAPGSTSSTAESPGPGSTSSTAESPGPGTSPSGES STAPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGSP	795	7,5

Nombre de la secuencia	Secuencia de aminoácidos	SEQ ID NO:	Puntuación
	GGSPGASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGSSPSASTGTGPGSG7ASSPGSS		
	TPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGAT		
	GSPGTPSGATGSPGTSSTGSPGTPGSGTASPGTSSTGSPGTPGSGTASSSPGSST		
	PSGATGSPGASPGTSSTGSPGTPGSGTASSPGASPGTSSTGSPGAPGSSTPSGAT		
	GPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGASPG		
	TSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSP		
	GASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGASPGTS		
	STGSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGT		
	PGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTG		
	TGPGASPGTSSTGSPGATGSPGTSSTGSPGTPGSGTASSPGASPGTSSTGSPGASP		
	GTSSTGSPGTPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGS		
	PGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGASPGTS		
	GASTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTS		
	STGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGT		
	PGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSP		
AM875	GTSTPSPSGSAPGSEPATSGSETPGSPAGSPTSTEEGSTSTSTADSPGTGTSTPTS	796	4,5
	GSASPGTSSTSPSGTAPGTSTSPSPAGTAPGTSTPESGSASPGTSTPESGSASPGSSE		
	PATSGSETPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSTESATPES		
	GPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTE		
	PSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTESATPES		
	GPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTE		
	PSEGSAPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAP		
	GTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGSTSTEE		
	GTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGSTSTEE		
	4TGTSEPGTPGSGTASSSPGSSTPSGATGSPGTSTEPSEGSAPGTSTEPSEGSAPGS		
	EPATSGSETPGSPAGSPTSTEEGSPASGSPTSTEEGSTSTEPSEGSAPGPEPTGPAP		
	SLASPEPATSESETPGTSESATPESGPASPAGSPTSTEEGSTSTEPSEGSAPGPEPTG		
	SPTSTEEGSPAGSPTSTEEGTSESATPESGPGSPAGSPTSTEEGSPASGSPTSTEE		
	GSTPSSTAESPGPGSTSESPSGTAPGTSPSPSESSTAPGTSPSPSESSTAPGSTSESPS		
	GTAPGTSPSPSESESTAPGTSTEPSEGSAPGTSPSPSESSTAPGSTSESATPESGPGSE		
	PATPSESETPGTSESATPESGPGTSTEPSEGSAPGSPSESASTGTGPGSTPSGATGSPG		
	SGTPSSTAESPGPGSTSESPSGTAPGTSPSPSESSTAPGTSPSPSESSTAPGSTSESPG		
	STGGTSPSPSESSTAPGSTSSTAESPGPGTSPSPSESSTAPGTSPESATPESGPGTSST		
	TPSSTGSAPGTSTEPSEGSAPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGS		

Nombre de la secuencia	Secuencia de aminoácidos	SEQ ID NO:	Puntuación
	PGTSTPESGSASPGTSPSGESSTAPGTSPSGESSTAPGTSESATPESGPGSEPATS GSETPGTSTEPSEGSAPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGSP AGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPES GPGSEPATSGSETPGSSTPSGATGSPGASPGTSSTPSGATGSPGSTSE SPSGTAPGTSPSGESSTAPGSTSSTAESPGPGSSTPSGATGSPGASPGTSSTGSP GTPGSGTASSSPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAP		

Ejemplo 59: Cálculo de las puntuaciones TEPITOPE

Las puntuaciones TEPITOPE de la secuencia peptídica de 9meros se pueden calcular añadiendo potenciales de bolsillo como se describe en Sturniolo [Sturniolo, T., et al. (1999) Nat Biotechnol, 17: 555]. En el presente ejemplo, se calcularon puntuaciones Tepitope individuales para alelos de HLA individuales. La tabla 32 muestra como ejemplo los potenciales de bolsillo para HLA*0101B, que se producen con alta frecuencia en la población de raza blanca. Para calcular la puntuación TEPITOPE de un péptido con secuencia P1-P2-P3-P4-P5-P6-P7-P8-P9, se añadieron los potenciales de bolsillo individuales correspondientes de la tabla 32. La puntuación HLA*0101B de un péptido de 9meros con la secuencia FDKLPRTSG (SEQ ID NO: 798) sería la suma de 0, -1,3, 0, 0,9, 0, -1,8, 0,09, 0, 0.

Para evaluar las puntuaciones TEPITOPE de péptidos largos se puede repetir el proceso para todas las subsecuencias de 9meros de las secuencias. Este proceso se puede repetir para las proteínas codificadas por otros alelos de HLA. Las tablas 33-36 dan potenciales de bolsillo para los productos proteicos de alelos de HLA que se producen con alta frecuencia en la población de raza blanca.

Las puntuaciones TEPITOPE calculadas por este método varían de aproximadamente -10 a +10. Sin embargo, péptidos de 9meros que carecen de un aminoácido hidrófobo (FKLMVWY (SEQ ID NO: 799)) en la posición P1 han calculado unas puntuaciones TEPITOPE en el intervalo de -1009 a -989. Este valor es biológicamente insignificante y refleja el hecho de que un aminoácido hidrófobo sirve como resto de anclaje para la unión al HLA, y se considera que los péptidos carecen de un resto hidrófobo en P1 no son proteínas de unión al HLA. Debido a que la mayoría de las secuencias XTEN carecen de restos hidrófobos, todas las combinaciones de subsecuencias de 9meros tendrán unas TEPITOPE en el intervalo de -1009 a -989. Este método confirma que los polipéptidos XTEN pueden tener pocos o ningún epítopo de linfocitos T predicho.

Tabla 32: Potencial de bolsillo para el alelo HLA*0101B.

Aminoácido	P1	P2	P3	P4	P5	P6	P7	P8	P9
A	-999	0	0	0	-	0	0	-	0
C	-999	0	0	0	-	0	0	-	0
D	-999	-1,3	-1,3	-2,4	-	-2,7	-2	-	-1,9
E	-999	0,1	-1,2	-0,4	-	-2,4	-0,6	-	-1,9
F	0	0,8	0,8	0,08	-	-2,1	0,3	-	-0,4
G	-999	0,5	0,2	-0,7	-	-0,3	-1,1	-	-0,8
H	-999	0,8	0,2	-0,7	-	-2,2	0,1	-	-1,1
I	-1	1,1	1,5	0,5	-	-1,9	0,6	-	0,7
K	-999	1,1	0	-2,1	-	-2	-0,2	-	-1,7
L	-1	1	1	0,9	-	-2	0,3	-	0,5
M	-1	1,1	1,4	0,8	-	-1,8	0,09	-	0,08
N	-999	0,8	0,5	0,04	-	-1,1	0,1	-	-1,2
P	-999	-0,5	0,3	-1,9	-	-0,2	0,07	-	-1,1
Q	-999	1,2	0	0,1	-	-1,8	0,2	-	-1,6
R	-999	2,2	0,7	-2,1	-	-1,8	0,09	-	-1
S	-999	-0,3	0,2	-0,7	-	-0,6	-0,2	-	-0,3
T	-999	0	0	-1	-	-1,2	0,09	-	-0,2
V	-1	2,1	0,5	-0,1	-	-1,1	0,7	-	0,3
W	0	-0,1	0	-1,8	-	-2,4	-0,1	-	-1,4
Y	0	0,9	0,8	-1,1	-	-2	0,5	-	-0,9

Tabla 33: Potencial de bolsillo para el alelo HLA*0301B.

Aminoácido	P1	P2	P3	P4	P5	P6	P7	P8	P9
A	-999	0	0	0	-	0	0	-	0

Aminoácido	P1	P2	P3	P4	P5	P6	P7	P8	P9
C	-999	0	0	0	-	0	0	-	0
D	-999	-1,3	-1,3	2,3	-	-2,4	-0,6	-	-0,6
E	-999	0,1	-1,2	-1	-	-1,4	-0,2	-	-0,3
F	-1	0,8	0,8	-1	-	-1,4	0,5	-	0,9
G	-999	0,5	0,2	0,5	-	-0,7	0,1	-	0,4
H	-999	0,8	0,2	0	-	-0,1	-0,8	-	-0,5
I	0	1,1	1,5	0,5	-	0,7	0,4	-	0,6
K	-999	1,1	0	-1	-	1,3	-0,9	-	-0,2
L	0	1	1	0	-	0,2	0,2	-	-0
M	0	1,1	1,4	0	-	-0,9	1,1	-	1,1
N	-999	0,8	0,5	0,2	-	-0,6	-0,1	-	-0,6
P	-999	-0,5	0,3	-1	-	0,5	0,7	-	-0,3
Q	-999	1,2	0	0	-	-0,3	-0,1	-	-0,2
R	-999	2,2	0,7	-1	-	1	-0,9	-	0,5
S	-999	-0,3	0,2	0,7	-	-0,1	0,07	-	1,1
T	-999	0	0	-1	-	0,8	-0,1	-	-0,5
V	0	2,1	0,5	0	-	1,2	0,2	-	0,3
W	-1	-0,1	0	-1	-	-1,4	-0,6	-	-1
Y	-1	0,9	0,8	-1	-	-1,4	-0,1	-	0,3

Tabla 34: Potencial de bolsillo para el alelo HLA*0401B.

Aminoácido	P1	P2	P3	P4	P5	P6	P7	P8	P9
A	-999	0	0	0	-	0	0	-	0
C	-999	0	0	0	-	0	0	-	0
D	-999	-1,3	-1,3	1,4	-	-1,1	-0,3	-	-1,7
E	-999	0,1	-1,2	1,5	-	-2,4	0,2	-	-1,7
F	0	0,8	0,8	-0,9	-	-1,1	-1	-	-1
G	-999	0,5	0,2	-1,6	-	-1,5	-1,3	-	-1
H	-999	0,8	0,2	1,1	-	-1,4	0	-	0,08
I	-1	1,1	1,5	0,8	-	-0,1	0,08	-	-0,3
K	-999	1,1	0	-1,7	-	-2,4	-0,3	-	-0,3
L	-1	1	1	0,8	-	-1,1	0,7	-	-1
M	-1	1,1	1,4	0,9	-	-1,1	0,8	-	-0,4
N	-999	0,8	0,5	0,9	-	1,3	0,6	-	-1,4
P	-999	-0,5	0,3	-1,6	-	0	-0,7	-	-1,3
Q	-999	1,2	0	0,8	-	-1,5	0	-	0,5
R	-999	2,2	0,7	-1,9	-	-2,4	-1,2	-	-1
S	-999	-0,3	0,2	0,8	-	1	-0,2	-	0,7
T	-999	0	0	0,7	-	1,9	-0,1	-	-1,2
V	-1	2,1	0,5	-0,9	-	0,9	0,08	-	-0,7

Aminoácido	P1	P2	P3	P4	P5	P6	P7	P8	P9
W	0	-0,1	0	-1,2	-	-1	-1,4	-	-1
Y	0	0,9	0,8	-1,6	-	-1,5	-1,2	-	-1

Tabla 35: Potencial de bolsillo para el alelo HLA*0701B.

Aminoácido	P1	P2	P3	P4	P5	P6	P7	P8	P9
A	-999	0	0	0	-	0	0	-	0
C	-999	0	0	0	-	0	0	-	0
D	-999	-1,3	-1,3	-1,6	-	-2,5	-1,3	-	-1,2
E	-999	0,1	-1,2	-1,4	-	-2,5	0,9	-	-0,3
F	0	0,8	0,8	0,2	-	-0,8	2,1	-	2,1
G	-999	0,5	0,2	-1,1	-	-0,6	0	-	-0,6
H	-999	0,8	0,2	0,1	-	-0,8	0,9	-	-0,2
I	-1	1,1	1,5	1,1	-	-0,5	2,4	-	3,4
K	-999	1,1	0	-1,3	-	-1,1	0,5	-	-1,1
L	-1	1	1	-0,8	-	-0,9	2,2	-	3,4
M	-1	1,1	1,4	-0,4	-	-0,8	1,8	-	2
N	-999	0,8	0,5	-1,1	-	-0,6	1,4	-	-0,5
P	-999	-0,5	0,3	-1,2	-	-0,5	-0,2	-	-0,6
Q	-999	1,2	0	-1,5	-	-1,1	1,1	-	-0,9
R	-999	2,2	0,7	-1,1	-	-1,1	0,7	-	-0,8
S	-999	-0,3	0,2	1,5	-	0,6	0,4	-	-0,3
T	-999	0	0	1,4	-	-0,1	0,9	-	0,4
V	-1	2,1	0,5	0,9	-	0,1	1,6	-	2
W	0	-0,1	0	-1,1	-	-0,9	1,4	-	0,8
Y	0	0,9	0,8	-0,9	-	-1	1,7	-	1,1

Tabla 36: Potencial de bolsillo para el alelo HLA*1501B.

Aminoácido	P1	P2	P3	P4	P5	P6	P7	P8	P9
A	-999	0	0	0	-	0	0	-	0
C	-999	0	0	0	-	0	0	-	0
D	-999	-1,3	-1,3	-0,4	-	-0,4	-0,7	-	-1,9
E	-999	0,1	-1,2	-0,6	-	-1	-0,7	-	-1,9
F	-1	0,8	0,8	2,4	-	-0,3	1,4	-	-0,4
G	-999	0,5	0,2	0	-	0,5	0	-	-0,8
H	-999	0,8	0,2	1,1	-	-0,5	0,6	-	-1,1
I	0	1,1	1,5	0,6	-	0,05	1,5	-	0,7
K	-999	1,1	0	-0,7	-	-0,3	-0,3	-	-1,7
L	0	1	1	0,5	-	0,2	1,9	-	0,5
M	0	1,1	1,4	1	-	0,1	1,7	-	0,08
N	-999	0,8	0,5	-0,2	-	0,7	0,7	-	-1,2

Aminoácido	P1	P2	P3	P4	P5	P6	P7	P8	P9
P	-999	-0,5	0,3	-0,3	-	-0,2	0,3	-	-1,1
Q	-999	1,2	0	-0,8	-	-0,8	-0,3	-	-1,6
R	-999	2,2	0,7	0,2	-	1	-0,5	-	-1
S	-999	-0,3	0,2	-0,3	-	0,6	0,3	-	-0,3
T	-999	0	0	-0,3	-	-0	0,2	-	-0,2
V	0	2,1	0,5	0,2	-	-0,3	0,3	-	0,3
W	-1	-0,1	0	0,4	-	-0,4	0,6	-	-1,4
Y	-1	0,9	0,8	2,5	-	0,4	0,7	-	-0,9

Ejemplo 60: Cromatografía analítica de exclusión por tamaño de proteínas de fusión XTEN con diversas cargas activas

Se realizaron análisis cromatográficos de exclusión por tamaño en proteínas de fusión que contenían diferentes proteínas terapéuticas y proteínas recombinantes no estructuradas de longitud creciente. Un ensayo ilustrativo utilizó una columna TSKGel-G4000 SWXL (7,8 mm x 30 cm) en la que se separaron 40 µg de proteína de fusión de glucagón purificada a una concentración de 1 mg/ml a un caudal de 0,6 ml/min en fosfato 20 mM a pH 6,8, NaCl 114 mM. Los perfiles cromatográficos se monitorizaron utilizando una DO214 nm y una DO 280nm. La calibración de columna para todos los ensayos se realizó utilizando un patrón de calibración de exclusión por tamaño de BioRad; los marcadores incluyen tiroglobulina (670 kDa), gammaglobulina bovina (158 kDa), ovoalbúmina de pollo (44 kDa), mioglobina equina (17 kDa) y vitamina B12 (1,35 kDa). Los perfiles cromatográficos representativos de glucagón-Y288, glucagón-Y144, glucagón-Y72, glucagón-Y36 se muestran como un solapamiento en la FIG. 32. Los datos muestran que el peso molecular aparente de cada compuesto es proporcional a la longitud de la secuencia del XTEN unida. Sin embargo, los datos también muestran que el peso molecular aparente de cada construcción es significativamente mayor de lo esperado para una proteína globular (como se muestra mediante la comparación con las proteínas patrón desarrolladas en el mismo ensayo). Tomando como base los análisis de la SEC para todas las construcciones evaluadas, los pesos moleculares aparentes, el factor de peso molecular aparente (expresado como la relación entre el factor de peso molecular aparente y el peso molecular calculado) y el radio hidrodinámico (R_H en nm) se muestran en la tabla 37. Los resultados indican que la incorporación de diferentes XTEN de 576 aminoácidos o más confiere un peso molecular aparente para la proteína de fusión de aproximadamente 339 kDa a 760, y que los XTEN de 864 aminoácidos o más confieren un peso molecular aparente mayor de aproximadamente 800 kDa. Los resultados de los aumentos proporcionales en el peso molecular aparente con respecto al peso molecular real eran coherentes para las proteínas de fusión creadas con XTEN de varias familias de motivos diferentes; es decir, AD, AE, AF, AG y AM, con aumentos de al menos cuatro veces y relaciones tan altas como de aproximadamente 17 veces. Además, la incorporación de compañeros de fusión XTEN con 576 aminoácidos o más en las proteínas de fusión con las diferentes cargas activas (y 288 restos en el caso de glucagón fusionado con Y288) dio como resultado un radio hidrodinámico de 7 nm o más; mucho más allá del tamaño de poro glomerular de aproximadamente 3-5 nm. Por consiguiente, se espera que las proteínas de fusión que comprenden crecimiento y XTEN tengan un aclaramiento renal reducido, contribuyendo a una semivida terminal aumentada y mejorando el efecto terapéutico o biológico con respecto a una proteína correspondiente de carga activa biológica no fusionada.

Tabla 37: Análisis mediante SEC de diferentes polipéptidos

Nombre de la construcción	**XTEN o compañero de fusión**	**Proteína terapéutica**	**PM real (kDa)**	**PM aparente (kDa)**	**Factor de peso molecular aparente**	**R_H (nm)**
AC14	Y288	Glucagón	28,7	370	12,9	7,0
AC28	Y144	Glucagón	16,1	117	7,3	5,0
AC34	Y72	Glucagón	9,9	58,6	5,9	3,8
AC33	Y36	Glucagón	6,8	29,4	4,3	2,6
AC89	AF120	Glucagón	14,1	76,4	5,4	4,3
AC88	AF108	Glucagón	13,1	61,2	4,7	3,9
AC73	AF144	Glucagón	16,3	95,2	5,8	4,7
AC53	AG576	GFP	74,9	339	4,5	7,0

Nombre de la construcción	XTEN o compañero de fusión	Proteína terapéutica	PM real (kDa)	PM aparente (kDa)	Factor de peso molecular aparente	R_H (nm)
AC39	AD576	GFP	76,4	546	7,1	7,7
AC41	AE576	GFP	80,4	760	9,5	8,3
AC52	AF576	GFP	78,3	526	6,7	7,6
AC398	AE288	FVII	76,3	650	8,5	8,2
AC404	AE864	FVII	129	1900	14,7	10,1
AC85	AE864	Exendina-4	83,6	938	11,2	8,9
AC114	AM875	Exendina-4	82,4	1344	16,3	9,4
AC143	AM875	hGH	100,6	846	8,4	8,7
AC227	AM875	IL-1ra	95,4	1103	11,6	9,2
AC228	AM1318	IL-1ra	134,8	2286	17,0	10,5

Ejemplo 61: Construcción de CBD-XTEN-Cys, un XTEN genomodificado con cisteína

Se introdujo, mediante oligos hibridados, una isla de cisteína (CysIsland) que codifica la secuencia de aminoácidos GGSPAGSCTSP (SEQ ID NO: 174) que contiene una cisteína en el vector CBD-relleno-GFP para obtener CBD-CysIsland-GFP, donde CysIsland está flanqueada por los sitios de restricción BsaI y BbsI. El vector CBD-relleno-GFP es un derivado del pET30 de Novagen con un sitio de reconocimiento de proteasa TEV entre el CBD y el relleno. Las construcciones se generaron previamente reemplazando el relleno en el vector CBD-relleno-GFP por genes que codifican XTEN_AE288 y XTEN_AE576. El plásmido de CBD-XTEN_AE288-GFP se digirió con BsaI/NcoI para generar el fragmento pequeño como inserto. El plásmido de CBD-CysIsland-GFP se digirió con BbsI/NcoI para generar el fragmento grande como vector. Los fragmentos del inserto y el vector se ligaron, y la mezcla de ligación se electroporó en células BL21-Gold (DE3) para obtener transformantes de CBD-CysIsland-XTEN_AE288-GFP. De manera similar, el plásmido de CBD-CysIsland-XTEN_AE288-GFP se digirió con BsaI/NcoI para generar el pequeño fragmento como inserto. El plásmido de CBD-XTEN_AE576-GFP se digirió con BbsI/NcoI para generar el fragmento grande como vector. Los fragmentos del inserto y el vector se ligaron, y la mezcla de ligación se electroporó en células BL21-Gold (DE3) para obtener transformantes de CBD-XTEN_AE576-CysIsland-XTEN_AE288-GFP. Finalmente, el plásmido de CBD-XTEN_AE576-CysIsland-XTEN_AE288-GFP se digirió con BbsI/HindIII para eliminar la GFP y ligarlo a oligos hibridados para el codón de terminación, y la mezcla de ligación se electroporó en células BL21-Gold (DE3) para obtener transformantes de CBD-XTEN_AE576-CysIsland-XTEN_AE288, que tiene las secuencias de ADN y aminoácidos codificadas de la siguiente tabla 38. Se pueden crear construcciones adicionales con cisteínas insertadas en diferentes ubicaciones dentro de la secuencia del XTEN mediante la selección de los sitios de restricción apropiados para la ubicación dada, incluyendo inserciones múltiples. El método también podría utilizarse para crear XTEN genomodificados con lisina mediante la sustitución de los codones que codifican lisina por los que codifican cisteína en los oligonucleótidos.

Tabla 38: Secuencia de ADN y de aminoácidos del XTEN genomodificado con Cys

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
CBD-TEV-AES76-CysIsland-AE288	ATGGCAAATACACCGGTATCAGGCAATTTGAAGGTT GAATTCTACAACAGCAATCCTTCAGATACTACTAAC TCAATCAATCCTCAGTTCAAGGTTACTAATACCGGA AGCAGTGCAATTGATTTGTCCAAACTCACATTGAGA TATTATTATACAGTAGACGGACAGAAAGATCAGACC TTCTGGGCTGACCATGCTGCAATAATCGGCAGTAAC GGCAGCTACAACGGAATTACTTCAAATGTAAAAGGA ACATTTGTAAAAATGAGTTCCTCAACAAATAACGCA GACACCTACCTTGAAATCAGCTTTACAGGCGGAACT CTTGAACCGGGTGCACATGTTCAGATACAAGGTAGA TTTGCAAAGAATGACTGGAGTAACTATACACAGTCA AATGACTACTCATTCAAGTCTGCTTCACAGTTTGTTG AATGGGATCAGGTAACAGCATACTTGAACGGTGTTC TTGTATGGGGTAAAGAACCCGGTGGCAGTGTAGTAG GTTCAGGTTCAGGATCCGAAAATCTGTATTTTCAGG GTGGGTCTCCAGGTAGCCCGGCTGGCTCTCCTACCTC TACTGAGGAAGGTACTTCTGAAAGCGCTACTCCTGA GTCTGGTCCAGGTACCTCTACTGAACCGTCCGAAGG TAGCGCTCCAGGTAGCCCAGCAGGCTCTCCGACTTC CACTGAGGAAGGTACTTCTACTGAACCTTCCGAAGG CAGCGCACCAGGTACCTCTACTGAACCTTCTGAGGG CAGCGCTCCAGGTACTTCTGAAAGCGCTACCCCGGA ATCTGGCCCAGGTAGCGAACCGGCTACTTCTGGTTC TGAAACCCCAGGTAGCGAACCGGCTACCTCCGGTTC TGAAACTCCAGGTAGCCCGGCAGGCTCTCCGACCTC TACTGAGGAAGGTACTTCTGAAAGCGCAACCCCGGA GTCCGGCCCAGGTACCTCTACCGAACCGTCTGAGGG CAGCGCACCAGGTACTTCTACCGAACCGTCCGAGGG	800	MANTPVSGNL KVEFYNSNPSD TTNSINPQFKV TNTGSSAIDLS KLTLRYYYTV DGQKDQTFWA DHAAHGSNGS YNGITSNVKGT FVKMSSSTNNA DTYLEISFTGG TLEPGAHVQIQ GRFAKNDWSN YTQSNDYSFKS ASQFVEWDQV TAYLNGVLVW GKEPGGSVVGS GSGSENLYFQG GSPGSPAGSPT STEEGTSESATP ESGPGTSTEPSE GSAPGSPAGSP TSTEEGTSTEPS EGSAPGTSTEP SEGSAPGTSES ATPESGPGSEP ATSGSETPGSE PATSGSETPGSP AGSPTSTEEGT	801

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	TAGCGCACCAGGTAGCCCAGCAGGTTCTCCTACCTC		SESATPESGPG	
	CACCGAGGAAGGTACTTCTACCGAACCGTCCGAGGG		TSTEPSEGSAP	
	TAGCGCACCAGGTACCTCTACTGAACCTTCTGAGGG		GTSTEPSEGSA	
	CAGCGCTCCAGGTACTTCTGAAAGCGCTACCCCGGA		PGSPAGSPTSTE	
	GTCCGGTCCAGGTACTTCTACTGAACCGTCCGAAGG		EGTSTEPSEGS	
	TAGCGCACCAGGTACTTCTGAAAGCGCAACCCCTGA		APGTSTEPSEG	
	ATCCGGTCCAGGTAGCGAACCGGCTACTTCTGGCTC		SAPGTSESATP	
	TGAGACTCCAGGTACTTCTACCGAACCGTCCGAAGG		ESGPGTSTEPSE	
	TAGCGCACCAGGTACTTCTACTGAACCGTCTGAAGG		GSAPGTSESAT	
	TAGCGCACCAGGTACTTCTGAAAGCGCAACCCCGGA		PESGPGSEPATS	
	ATCCGGCCCAGGTACCTCTGAAAGCGCAACCCCGGA		GSETPGTSTEPS	
	GTCCGGCCCAGGTAGCCCTGCTGGCTCTCCAACCTC		EGSAPGTSTEP	
	CACCGAAGAAGGTACCTCTGAAAGCGCAACCCCTGA		SEGSAPGTSES	
	ATCCGGCCCAGGTAGCGAACCGGCAACCTCCGGTTC		ATPESGPGTSE	
	TGAAACCCCAGGTACCTCTGAAAGCGCTACTCCGGA		SATPESGPGSP	
	GTCTGGCCCAGGTACCTCTACTGAACCGTCTGAGGG		AGSPTSTEEGT	
	TAGCGCTCCAGGTACTTCTACTGAACCGTCCGAAGG		SESATPESGPGS	
	TAGCGCACCAGGTACTTCTACCGAACCGTCCGAAGG		EPATSGSETPG	
	CAGCGCTCCAGGTACCTCTACTGAACCTTCCGAGGG		TSESATPESGP	
	CAGCGCTCCAGGTACCTCTACCGAACCTTCTGAAGG		GTSTEPSEGSA	
	TAGCGCACCAGGTACTTCTACCGAACCGTCCGAGGG		PGTSTEPSEGS	
	TAGCGCACCAGGTAGCCCAGCAGGTTCTCCTACCTC		APGTSTEPSEG	
	CACCGAGGAAGGTACTTCTACCGAACCGTCCGAGGG		SAPGTSTEPSE	
	TAGCGCACCAGGTACCTCTGAAAGCGCAACTCCTGA		GSAPGTSTEPS	
	GTCTGGCCCAGGTAGCGAACCTGCTACCTCCGGCTC		EGSAPGTSTEP	
	TGAGACTCCAGGTACCTCTGAAAGCGCAACCCCGGA		SEGSAPGSPAG	
	ATCTGGTCCAGGTAGCGAACCTGCAACCTCTGGCTC		SPTSTEEGTSTE	
	TGAAACCCCAGGTACCTCTGAAAGCGCTACTCCTGA		PSEGSAPGTSES	
	ATCTGGCCCAGGTACTTCTACTGAACCGTCCGAGGG		ATPESGPGSEP	
	CAGCGCACCAGGTACTTCTGAAAGCGCTACTCCTGA		ATSGSETPGTS	
	GTCCGGCCCAGGTAGCCCGGCTGGCTCTCCGACTTC		ESATPESGPGS	
	CACCGAGGAAGGTAGCCCGGCTGGCTCTCCAACTTC		EPATSGSETPG	
	TACTGAAGAAGGTAGCCCGGCAGGCTCTCCGACCTC		TSESATPESGP	
	TACTGAGGAAGGTACTTCTGAAAGCGCAACCCCGGA		GTSTEPSEGSA	
	GTCCGGCCCAGGTACCTCTACCGAACCGTCTGAGGG		PGTSESATPES	
	CAGCGCACCAGGTGGTAGCCCGGCTGGCTCTTGTAC		GPGSPAGSPTS	
	CTCTCCAGGTACCTCTGAAAGCGCAACTCCTGAGTC		TEEGSPAGSPT	
	TGGCCCAGGTAGCGAACCTGCTACCTCCGGCTCTGA		STEEGSPAGSP	
	GACTCCAGGTACCTCTGAAAGCGCAACCCCGGAATC		TSTEEGTSESA	
	TGGTCCAGGTAGCGAACCTGCAACCTCTGGCTCTGA		TPESGPGTSTEP	
	AACCCCAGGTACCTCTGAAAGCGCTACTCCTGAATC		SEGSAPGGSPA	
	TGGCCCAGGTACTTCTACTGAACCGTCCGAGGGCAG		GSCTSPGTSES	
	CGCACCAGGTAGCCCTGCTGGCTCTCCAACCTCCAC		ATPESGPGSEP	
	CGAAGAAGGTACCTCTGAAAGCGCAACCCCTGAATC		ATSGSETPGTS	
	CGGCCCAGGTAGCGAACCGGCAACCTCCGGTTCTGA		ESATPESGPGS	
	AACCCCAGGTACTTCTGAAAGCGCTACTCCTGAGTC		EPATSGSETPG	
	CGGCCCAGGTAGCCCCGCTGGCTCTCCGACTTCCAC		TSESATPESGP	
	CGAGGAAGGTAGCCCGGCTGGCTCTCCAACTTCTAC		GTSTEPSEGSA	
	TGAAGAAGGTACTTCTACCGAACCTTCCGAGGGCAG		PGSPAGSPTSTE	
	CGCACCAGGTACTTCTGAAAGCGCTACCCCTGAGTC		EGTSESATPES	
	CGGCCCAGGTACTTCTGAAAGCGCTACTCCTGAATC		GPGSEPATSGS	
	CGGTCCAGGTACTTCTGAAAGCGCTACCCCGGAATC		ETPGTSESATPE	
	TGGCCCAGGTAGCGAACCGGCTACTTCTGGTTCTGA		SGPGSPAGSPT	
	AACCCCAGGTAGCGAACCGGCTACCTCCGGTTCTGA		STEEGSPAGSP	
	AACTCCAGGTAGCCCAGCAGGCTCTCCGACTTCCAC		TSTEEGTSTEPS	
	TGAGGAAGGTACTTCTACTGAACCTTCCGAAGGCAG		EGSAPGTSESA	
	CGCACCAGGTACCTCTACTGAACCTTCTGAGGGCAG		TPESGPGTSES	
	CGCTCCAGGTAGCGAACCTGCAACCTCTGGCTCTGA		ATPESGPGTSE	

Nombre del clon	Secuencia de ADN	SEQ ID NO:	Secuencia de aminoácidos	SEQ ID NO:
	AACCCCAGGTACCTCTGAAAGCGCTACTCCTGAATC TGGCCCAGGTACTTCTACTGAACCGTCCGAGGGCAG CGCACCAGGTTAA		SATPESGPGSEP ATSGSETPGSE PATSGSETPGSP AGSPTSTEEGT STEPSEGSAPG TSTEPSEGSAP GSEPATSGSET PGTSESATPES GPGTSTEPSEG SAPG	

Ejemplo 62: Purificación de CBD-XTEN-Cys

Las *E. coli* que contenían AC292 en un plásmido se cultivaron hasta saturación durante la noche en 2xYT, y a continuación se utilizaron 200 ml de este cultivo para inocular un cultivo de 25 l de medios 2xYT en una bolsa de balanceo. Ambos cultivos estaban en presencia de 50 µg/ml de kanamicina. El segundo cultivo se cultivó hasta una DO600 de ~1,0 a 37 °C, se enfrió hasta 26 °C y se indujo con 12 ml de IPTG 1 M durante la noche. El sedimento celular se recolectó a 4000 rpm con un rotor SLA-3000 rotando durante 20 minutos. El sedimento celular (184 g) se resuspendió en 736 ml de Tris 20 mM a pH 6,8, NaCl 50 mM. Las células resuspendidas se lisaron con un microfluidificador a 137 895,15 kPa (20 000 psi), y a continuación se calentaron a 75 °C durante 15 minutos, seguido de un enfriamiento rápido en hielo durante 30 minutos. A continuación, el lisado se clarificó por centrifugación. A continuación, el lisado clarificado se cargó en una columna DE52, previamente higienizada con NaOH y equilibrada con Tris 20 mM a pH 6,8, NaCl 50 mM. La columna se lavó con 5 volúmenes de columna de Tris 20 mM a pH 6,8, NaCl 50 mM, 5 volúmenes de columna de Tris 20 mM a pH 6,8, NaCl 150 mM, y se eluyó con 5 volúmenes de columna de Tris 20 mM a pH 6,8, NaCl 250 mM. Las fracciones de elución agrupadas se cargaron a continuación en una columna macrocapQ, previamente higienizada con NaOH y equilibrada con Tris 20 mM a pH 6,8, NaCl 50 mM. La columna se lavó con 9 volúmenes de columna de Tris 20 mM a pH 6,8, NaCl 50 mM, 9 volúmenes de columna de Tris 20 mM a pH 6,8, NaCl 100 mM, y se eluyó con 9 volúmenes de columna de Tris 20 mM a pH 6,8, NaCl 250 mM. Las fracciones de elución agrupadas se ajustaron hasta un 15 % p/v de sulfato de sodio, y a continuación se cargaron en una columna de octilsepharose FF previamente higienizada con NaOH y equilibrada con Tris a pH 7,5. La columna se lavó con 4 volúmenes de columna de Tris 20 mM a pH 7,5, 15 % p/v de sulfato de sodio, y se eluyó con 4 volúmenes de columna de Tris 20 mM a pH 7,5, 5 % p/v de sulfato de sodio. La muestra se almacenó a 4 °C y se le dio el n.º de lote AP197. El XTEN genomodificado con cisteína purificado podría servir entonces como un reactivo adecuado para la conjugación con un fármaco, tal como un fármaco de la tabla 5, lo que da como resultado un conjugado de XTEN-fármaco.

Ejemplo 63: Conjugación y purificación de FITC-X-XTEN

La proteína purificada derivada de AC272, lote n.º AP197, se marcó con FITC maleimida. La muestra se redujo incubando a la temperatura ambiente con TCEP 5 mM durante 1 hora. A continuación, la muestra se desaló con PBS utilizando columnas DG-10. La muestra se marcó añadiendo un exceso molar de 25 veces de FITC-maleimida en DMSO e incubando a la temperatura ambiente durante 2 horas. Obsérvese que el volumen se ajustó de tal manera que la concentración de DMSO era <5 % del disolvente total. La reacción se inactivó añadiendo DTT 2 mM, y a continuación la muestra se digirió durante la noche con proteasa TEV. La muestra se diluyó dos veces con Tris 20 mM a pH 7,5 y se cargó en una columna macrocapQ, previamente higienizada con NaOH y equilibrada con Tris 20 mM a pH 7,5. La columna se lavó con 5 volúmenes de columna de Tris 20 mM a pH 7,5, NaCl 135 mM, 5 volúmenes de columna de Tris 20 mM a pH 7,5, NaCl 175 mM, y se eluyó con 5 volúmenes de columna de Tris 20 mM a pH 7,5, NaCl 250 mM. A continuación, las fracciones de elución agrupadas se digirieron con TEV durante 60 horas a 4C para completar la digestión. A continuación, las muestras digeridas se pasaron dos veces por una columna perloza de 1 ml previamente higienizada con NaOH y equilibrada con Tris 20 mM a pH 7,5, NaCl 135 mM. Para eliminar cualquier FITC libre, la muestra se dializó a continuación contra Tris 20 mM a pH 7,5, NaCl 135 mM, utilizando una membrana de 10 000 MWCO. La comigración de la señal de proteína a una DO214 y la señal de FITC a una DO495 en una columna de SEC indica una conjugación exitosa del XTEN con el marcador, con una contaminación mínima de colorante libre (FIG. 42B). La conjugación exitosa también está indicada por los grandes PM aparentes de la proteína con fluorescencia FITC en la SDS-PAGE (FIG. 42A).

Ejemplo 64: Purificación de GFP-X-XTEN

La GFP (AC219) se conjugó químicamente con el XTEN mediante un agente de conjugación bifuncional con un grupo amino reactivo para acoplarse a las lisinas de la GFP y un grupo cisteína reactivo para acoplarse a la cisteína libre genomodificada del XTEN en AC292. La GFP se marcó con un agente de conjugación bifuncional de sulfo-SMCC incubando a la temperatura ambiente durante 2 horas. La proteína se desaló con PBS utilizando columnas DG-10 para eliminar el sulfo-SMCC libre. La proteína purificada derivada de AC272, lote n.º AP197, se redujo y se desaló con PBS en columnas DG-10 y se mezcló con la GFP marcada para permitir la conjugación. La reacción de conjugación se inactivó con DTT 2 mM y se añadió TEV para eliminar el dominio CBD en una incubación durante la noche a 4 °C. Al día siguiente se añadió TEV adicional para completar la digestión con una incubación adicional de 60 horas a 4 °C. Después de la digestión con TEV, la muestra se diluyó hasta 100 ml con Tris 20 mM a pH 7,5 y se cargó en una columna macrocapQ, previamente higienizada con NaOH y equilibrada con Tris 20 mM a pH 7,5. La columna se lavó con 5 volúmenes de columna de Tris 20 mM a pH 7,5, 5 volúmenes de columna de Tris 20 mM a pH 7,5, NaCl 50 mM, 5 volúmenes de columna de Tris 20 mM a pH 7,5, NaCl 100 mM, 5 volúmenes de columna de Tris 20 mM a pH 7,5, NaCl 150 mM, 5 volúmenes de columna de Tris 20 mM a pH 7,5, NaCl 200 mM, 5 volúmenes de columna de Tris 20 mM a pH 7,5, NaCl 250 mM, 5 volúmenes de columna de Tris 20 mM a pH 7,5, NaCl 300 mM y 5 volúmenes de columna de Tris 20 mM a pH 7,5, NaCl 500 mM. Las fracciones de elución máxima se agruparon y se almacenaron a 4 °C. La conjugación se confirmó mediante la comigración de la señal de la proteína a una DO214 y la señal de la GFP a una DO395 en una columna de SEC, mostrándose el resultado de la SEC como superposiciones en la FIG. 43.

Ejemplo 65: Farmacocinética de los conjugados GFP-XTEN y FITC-XTEN

La farmacocinética de los conjugados GFP-XTEN y FITC-XTEN preparados como se describe en los ejemplos anteriores se ensayó en macacos cangrejeros. Se administraron GFP-XTEN y FITC-XTEN a macacos cangrejeros macho por vía i.v. a 2 mg/kg y con volúmenes de dosis de 0,77 y 0,68 ml, respectivamente. Se recogieron muestras de sangre (1,0 ml) en tubos heparinizados previamente enfriados en los puntos temporales de predosis, 2, 4, 8, 24, 48, 72, 96, 120, 168, 216, 264, 336, 388, 432, 504 horas, y se procesaron hasta plasma. La cuantificación se realizó mediante un ensayo de ELISA utilizando el anticuerpo anti-XTEN tanto para la captura como para la detección en el caso de GFP-XTEN, y para la captura anti-XTEN y para la detección anti-FITC en el caso de FITC-XTEN. Se realizó un análisis no compartimental en WinNonLin con todos los puntos temporales incluidos en el ajuste para determinar los parámetros FC. Los resultados farmacocinéticos se resumen en la tabla 39 y la FIG. 44. Los datos muestran que los XTEN pueden prolongar la semivida de moléculas con las que estén conjugados químicamente de una manera comparable a las fusiones genéticas con cargas activas de un tamaño similar.

Tabla 39: Narámetros FC:

Construcción	Cmáx (ng/ml)	AUC (h*ng/ml)	T ½ (h)
GFP-X-XTEN (AP197d)	52800	8220000	107
FITC-X-XTEN AP197e	18900	3930000	842

Ejemplo 66: Preparación del conjugado BFP-D anti-Her2-XTEN-paclitaxel

El proceso de conjugación que se puede emplear se ilustra genéricamente en la FIG. 45. Los fármacos pueden conjugarse con XTEN utilizando los métodos de la divulgación o los de, por ejemplo, la Publicación de Patente de EE. UU. n.º 2009/0074704. Mediante los métodos, se puede hacer reaccionar en primer lugar el paclitaxel con un conector activado con un grupo funcional adecuado, como se muestra en la FIG. 45. El conjugado resultante se puede purificar mediante métodos adecuados y convencionales conocidos en la técnica, que incluyen HPLC, cromatografía de exclusión por tamaño, filtración en gel, cromatografía de intercambio iónico, o combinaciones de los mismos. Posteriormente, el conjugado de paclitaxel activado se incubaría con aHer2-XTEN-Cys que se habría expresado y purificado como se describe en el ejemplo 24, utilizando al menos un exceso de 25 veces del conjugado de paclitaxel para garantizar que todos los sitios reactivos de cisteína del XTEN estén conjugados. La conjugación podría realizarse esencialmente como se describe en el ejemplo 64. El exceso de paclicaxel se eliminaría de la mezcla de reacción mediante diálisis, y el producto purificado final se concentraría y almacenaría para su uso posterior.

Ejemplo 67: Aplicaciones clínicas de las composiciones de anti-Her2-XTEN-pactlitaxel

El antígeno Her2 está sobreexpresado en un gran número de neoplasias malignas sólidas. La expresión es particularmente alta en muchas células de cáncer de mama. Herceptin ha sido aprobado para el tratamiento de cánceres de mama positivos para Her2. La divulgación contempla que el anti-Her2-XTEN-paclitaxel se pueda utilizar para el tratamiento de la misma población de pacientes. Los ensayos clínicos pueden diseñarse de manera que se pueda verificar la eficacia y las ventajas de las composiciones de anti-Her2-XTEN-paclitaxel en seres humanos. Dichos estudios en pacientes comprenderían tres fases. En primer lugar, se realizaría un estudio de seguridad y farmacocinética de fase I en pacientes adultos para determinar la dosis máxima tolerada y la farmacocinética y la farmacodinámica en seres humanos. Estos estudios definen posibles toxicidades y acontecimientos adversos de los que se debe hacer un seguimiento en futuros estudios. El esquema del estudio sería utilizar dosis únicas en aumento

de las composiciones de aHer2-XTEN-paclitaxel y medir los parámetros bioquímicos, FC y clínicos. Esto permitiría la determinación de la dosis máxima tolerada y establecer el umbral y las concentraciones máximas en la dosis y el fármaco circulante que constituyen la ventana terapéutica que se utilizará en los ensayos posteriores de fase II y fase III realizados para las indicaciones diana, para determinar la eficacia y la tolerabilidad de las composiciones de aHER2-XTEN-paclitaxel.

Tabla 40: Proteínas de fusión de unión con restos de direccionamiento dirigidos a dianas individuales

Nombre*	Secuencia de proteínas	SEQ ID NO:
CTLA4_4-AM875	MHVAQPAVVLASSRGIASFVCEYASPGKATEVRVTVLRQADSQVTEVCAATY MMGNELTFLDDSICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMYPPPYYLGI GNGTQIYVIDPEGGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSTSSTAE SPGPGTSTPESGSASPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTPE SGSASPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTS ESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAP GTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEG SAPGTSESATPESGPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSSTPS GATGSPGTPGSGTASSSPGSSTPSGATGSPGTSTEPSEGSAPGTSTEPSEGSAPGS EPATSGSETPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGASASGAPST GGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSTSSTAESPGPGSTSESPS GTAPGTSPSGESSTAPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSEP ATSGSETPGTSESATPESGPGSEPATSGSETPGSTSSTAESPGPGSTSSTAESPGPG TSPSGESSTAPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGSTSSTAESP GPGTSTPESGSASPGSTSESPSGTAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPS EGSAPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSEPATSGSETPGTSE SATPESGPGSPAGSPTSTEEGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPG TSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAP	802
AE912-CTLA4	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGSPAGSP TSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTST EPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATS GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPA GSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPES GPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSE SATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGS APGMHVAQPAVVLASSRGIASFVCEYASPGKATEVRVTVLRQADSQVTEVCAA TYMMGNELTFLDDSICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMYPPPYY LGIGNGTQIYVIDPEG	803

Nombre*	Secuencia de proteínas	SEQ ID NO:
CTLA4-AE36-CTLA4-AE864	MAMHVAQPAVVLASSRGIASFVCEYASPGKATEVRVTVLRQADSQVTEVCAA TYMMGNELTFLDDSICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMYPPPYY LGIGNGTQIYVIDPEGAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGPA MHVAQPAVVLASSRGIASFVCEYASPGKATEVRVTVLRQADSQVTEVCAATY MMGNELTFLDDSICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMYPPPYYLGI GNGTQIYVIDPEGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPT STEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPA TSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGS PAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSA PGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATP ESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGT STEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESG PGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSE GSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSES	804
	ATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGS EPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESG PGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSE GSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPA TSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGT SESATPESGPGTSTEPSEGSAPG	
CTLA4-AE158-CTLA4-AE864	MAMHVAQPAVVLASSRGIASFVCEYASPGKATEVRVTVLRQADSQVTEVCAA TYMMGNELTFLDDSICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMYPPPYY LGIGNGTQIYVIDPEGAPSTGGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAP GTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEG SAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGPEPTG PAMHVAQPAVVLASSRGIASFVCEYASPGKATEVRVTVLRQADSQVTEVCAAT YMMGNELTFLDDSICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMYPPPYYL GIGNGTQIYVIDPEGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGS PTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSE PATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAP GSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEG SAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESA TPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTS ESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAP GTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPE SGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEP SEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTS ESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGP GSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPE SGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEP SEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSE PATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETP GTSESATPESGPGTSTEPSEGSAPG	805

Nombre*	Secuencia de proteínas	SEQ ID NO:
CTLA4-AE158-CTLA4-AE864	MAMHVAQPAVVLASSRGIASFVCEYASPGKATEVRVTVLRQADSQVTEVCAA TYMMGNELTFLDDSICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMYPPPYY LGIGNGTQIYVIDPEPCPDSGAPSTGGTSESATPESGPGTSTEPSEGSAPGTSTEPS EGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTST EPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPG PEPTGPAMHVAQPAVVLASSRGIASFVCEYASPGKATEVRVTVLRQADSQVTE VCAATYMMGNELTFLDDSICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMYP PPYYLGIGNGTQIYVIDPEPCPDSGG SPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSP TSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTST EPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPG SEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPES GPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPS EGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTST EPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPG TSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPES GPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPS EGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSE SATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPG TSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPES GPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSP TSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTST EPSEGSAPG	806
AE912-aIL6R	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGSPAGSP TSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTST	807
	EPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATS GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPA GSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPES GPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSE SATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGS APGADIQMTQSPSSLSASVGDRVTITCRASQDISSYLNWYQQKPGKAPKLLIYY TSRLHSGVPSRFSGSGSGTDFTFTISSLQPEDIATYYCQQGNTLPYTFGQGTKVEI KTGSGEGSEGEGGGEGSEGEGSGEGGEGEGSGSQVQLQESGPGLVRPSQTLSLT CTVSGYSITSDHAWSWVRQPPGRGLEWIGYISYSGITTYNPSLKSRVTMLRDTS KNQFSLRLSSVTAADTAVYYCARSLARTTAMDYWGQGSLVTVSS	

Nombre*	Secuencia de proteínas	SEQ ID NO:
alL6R-AE864	MADIQMTQSPSSLSASVGDRVTITCRASQDISSYLNWYQQKPGKAPKLLIYYTS RLHSGVPSRFSGSGSGTDFTFTISSLQPEDIATYYCQQGNTLPYTFGQGTKVEIKT GSGEGSEGEGGGEGSEGEGSGEGGEGEGSGSQVQLQESGPGLVRPSQTLSLTCT VSGYSITSDHAWSWVRQPPGRGLEWIGYISYSGITTYNPSLKSRVTMLRDTSKN QFSLRLSSVTAADTAVYYCARSLARTTAMDYWGQGSLVTVSSGSPGSPAGSPT STEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTE PSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGT SESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSA PGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSG SETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAG SPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGT STEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSA PGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATP ESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAG SPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGT SESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESG PGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATP ESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSES ATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGT STEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSA PG	808
AE912-alL6R-AE144	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGSPAGSP TSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTST EPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATS GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPA GSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPES GPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSE SATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGS	809
	APGADIQMTQSPSSLSASVGDRVTITCRASQDISSYLNWYQQKPGKAPKLLIYY TSRLHSGVPSRFSGSGSGTDFTFTISSLQPEDIATYYCQQGNTLPYTFGQGTKVEI KTGSGEGSEGEGGGEGSEGEGSGEGGEGEGSGSQVQLQESGPGLVRPSQTLSLT CTVSGYSITSDHAWSWVRQPPGRGLEWIGYISYSGITTYNPSLKSRVTMLRDTS KNQFSLRLSSVTAADTAVYYCARSLARTTAMDYWGQGSLVTVSSGGTSESATP ESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTE PSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS	

Nombre*	Secuencia de proteínas	SEQ ID NO:
AE48-aIL6R-AE864	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGSPGDIQ MTQSPSSLSASVGDRVTITCRASQDISSYLNWYQQKPGKAPKLLIYYTSRLHSG VPSRFSGSGSGTDFTFTISSLQPEDIATYYCQQGNTLPYTFGQGTKVEIKTGSGEG SEGEGGGEGSEGEGSGEGGEGEGSGSQVQLQESGPGLVRPSQTLSLTCTVSGYS ITSDHAWSWVRQPPGRGLEWIGYISYSGITTYNPSLKSRVTMLRDTSKNQFSLR LSSVTAADTAVYYCARSLARTTAMDYWGQGSLVTVSSGGSPAGSPTSTEEGTS ESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAP GTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPE SGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEP SEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTS TEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEE GTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEG SAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGS PTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSE PATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEE GSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPE SGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEP SEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSP AGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGP GTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEG SAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG	810
aIL6R_VL-AF144-aIL6R_VH-AM875	DIQMTQSPSSLSASVGDRVTITCRASQDISSYLNWYQQKPGKAPKLLIYYTSRLH SGVPSRFSGSGSGTDFTFTISSLQPEDIATYYCQQGNTLPYTFGQGTKVEIKGGTS TPESGSASPGTSPSGESSTAPGTSPSGESSTAPGSTSSTAESPGPGSTSESPSGTAP GSTSSTAESPGPGTSPSGESSTAPGTSTPESGSASPGSTSSTAESPGPGTSPSGESST APGTSPSGESSTAPGTSPSGESSTAPGQVQLQESGPGLVRPSQTLSLTCTVSGYSI TSDHAWSWVRQPPGRGLEWIGYISYSGITTYNPSLKSRVTMLRDTSKNQFSLRL SSVTAADTAVYYCARSLARTTAMDYWGQGSLVTVSSGGTSTEPSEGSAPGSEP ATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAPG STSESPSGTAPGTSTPESGSASPGTSTPESGSASPGSEPATSGSETPGTSESATPES GPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPS EGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSE SATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPG SEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATG SPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSPAGSP TSTEEGTSTEPSEGSAPGASASGAPSTGGTSESATPESGPGSPAGSPTSTEEGSPA GSPTSTEEGSTSSTAESPGPGSTSESPSGTAPGTSPSGESSTAPGTPGSGTASSSPG SSTPSGATGSPGSSPSASTGTGPGSEPATSGSETPGTSESATPESGPGSEPATSGSE TPGSTSSTAESPGPGSTSSTAESPGPGTSPSGESSTAPGSEPATSGSETPGSEPATS GSETPGTSTEPSEGSAPGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAPGTST EPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSSTPSGATGSPGSSPSASTGTGPG ASPGTSSTGSPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSSTPSGATG SPGSSPSASTGTGPGASPGTSSTGSPGTSESATPESGPGTSTEPSEGSAPGTSTEPS EGSAPG	811
AM923-aIL6R_VH-AF144-aIL6R_VL	MAEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGTSTEPS EGSAPGSEPATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTPESGSASPGSTS ESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPGSEPATSGSETPG TSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGS APGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESAT	812

Nombre*	Secuencia de proteínas	SEQ ID NO:
	PESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTST EPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGTASSSPG SSTPSGATGSPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTST EEGSPAGSPTSTEEGTSTEPSEGSAPGASASGAPSTGGTSESATPESGPGSPAGSP TSTEEGSPAGSPTSTEEGSTSSTAESPGPGSTSESPSGTAPGTSPSGESSTAPGTPG SGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSEPATSGSETPGTSESATPESGPG SEPATSGSETPGSTSSTAESPGPGSTSSTAESPGPGTSPSGESSTAPGSEPATSGSE TPGSEPATSGSETPGTSTEPSEGSAPGSTSSTAESPGPGTSTPESGSASPGSTSESPS GTAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSSTPSGATGSPGSSPS ASTGTGPGASPGTSSTGSPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGS STPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTSESATPESGPGTSTEPSEGSA PGTSTEPSEGSAPGQVQLQESGPGLVRPSQTLSLTCTVSGYSITSDHAWSWVRQ PPGRGLEWIGYISYSGITTYNPSLKSRVTMLRDTSKNQFSLRLSSVTAADTAVYY CARSLARTTAMDYWGQGSLVTVSSGGTSTPESGSASPGTSPSGESSTAPGTSPS GESSTAPGSTSSTAESPGPGSTSESPSGTAPGSTSSTAESPGPGTSPSGESSTAPGT STPESGSASPGSTSSTAESPGPGTSPSGESSTAPGTSPSGESSTAPGTSPSGESSTAP GDIQMTQSPSSLSASVGDRVTITCRASQDISSYLNWYQQKPGKAPKLLIYYTSRL HSGVPSRFSGSGSGTDFTFTISSLQPEDIATYYCQQGNTLPYTFGQGTKVEIKG	
aIL6R_VL-Conector_AE42 - aIL6R_VH-AD576	DIQMTQSPSSLSASVGDRVTITCRASQDISSYLNWYQQKPGKAPKLLIYYTSRLH SGVPSRFSGSGSGTDFTFTISSLQPEDIATYYCQQGNTLPYTFGQGTKVEIKGGA PGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGPAGQVQLQESGPGLVRPS QTLSLTCTVSGYSITSDHAWSWVRQPPGRGLEWIGYISYSGITTYNPSLKSRVT MLRDTSKNQFSLRLSSVTAADTAVYYCARSLARTTAMDYWGQGSLVTVSSGG SSESGSSEGGPGSGGEPSESGSSGSSESGSSEGGPGSSESGSSEGGPGSSESGSSEG GPGSSESGSSEGGPGSSESGSSEGGPGESPGGSSGSESGSEGSSGPGESSGSSESGS SEGGPGSSESGSSEGGPGSSESGSSEGGPGSGGEPSESGSSGESPGGSSGSESGESP GGSSGSESGSGGEPSESGSSGSSESGSSEGGPGSGGEPSESGSSGSGGEPSESGSS GSEGSSGPGESSGESPGGSSGSESGSGGEPSESGSSGSGGEPSESGSSGSGGEPSES GSSGSSESGSSEGGPGESPGGSSGSESGESPGGSSGSESGESPGGSSGSESGESPG GSSGSESGESPGGSSGSESGSSESGSSEGGPGSGGEPSESGSSGSEGSSGPGESSGS SESGSSEGGPGSGGEPSESGSSGSSESGSSEGGPGSGGEPSESGSSGESPGGSSGSE SGESPGGSSGSESGSSESGSSEGGPGSGGEPSESGSSGSSESGSSEGGPGSGGEPSE SGSSGSGGEPSESGSSGESPGGSSGSESGSEGSSGPGESSGSSESGSSEGGPGSEGS SGPGESSG	813

Nombre*	Secuencia de proteínas	SEQ ID NO:
AM923-aIL6R_VH-Conector_AM1 50-aIL6R_VH	MAEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGTSTEPS EGSAPGSEPATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTPESGSASPGSTS ESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPGSEPATSGSETPG TSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGS APGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESAT PESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTST EPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGTASSSPG SSTPSGATGSPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTST EEGSPAGSPTSTEEGTSTEPSEGSAPGASASGAPSTGGTSESATPESGPGSPAGSP TSTEEGSPAGSPTSTEEGSTSSTAESPGPGSTSESPSGTAPGTSPSGESSTAPGTPG SGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSEPATSGSETPGTSESATPESGPG SEPATSGSETPGSTSSTAESPGPGSTSSTAESPGPGTSPSGESSTAPGSEPATSGSE TPGSEPATSGSETPGTSTEPSEGSAPGSTSSTAESPGPGTSTPESGSASPGSTSESPS GTAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSSTPSGATGSPGSSPS ASTGTGPGASPGTSSTGSPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGS STPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTSESATPESGPGTSTEPSEGSA PGTSTEPSEGSAPGQVQLQESGPGLVRPSQTLSLTCTVSGYSITSDHAWSWVRQ PPGRGLEWIGYISYSGITTYNPSLKSRVTMLRDTSKNQFSLRLSSVTAADTAVYY CARSLARTTAMDYWGQGSLVTVSSGGAPSTGGTSESATPESGPGTSTEPSEGSA PGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATP ESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTE PSEGSAPGPEPTGPAGQVQLQESGPGLVRPSQTLSLTCTVSGYSITSDHAWSWV RQPPGRGLEWIGYISYSGITTYNPSLKSRVTMLRDTSKNQFSLRLSSVTAADTAV YYCARSLARTTAMDYWGQGSLVTVSSG	814
aIL6R_VL-AE42-aIL6R _VH-AM875	DIQMTQSPSSLSASVGDRVTITCRASQDISSYLNWYQQKPGKAPKLLIYYTSRLH SGVPSRFSGSGSGTDFTFTISSLQPEDIATYYCQQGNTLPYTFGQGTKVEIKGGA PGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGPAGQVQLQESGPGLVRPS QTLSLTCTVSGYSITSDHAWSWVRQPPGRGLEWIGYISYSGITTYNPSLKSRVT MLRDTSKNQFSLRLSSVTAADTAVYYCARSLARTTAMDYWGQGSLVTVSSGG TSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTPESGSA SPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPGSEPATS GSETPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTST EPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPG TSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPES GPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGT ASSSPGSSTPSGATGSPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPA GSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGASASGAPSTGGTSESATPESGPG SPAGSPTSTEEGSPAGSPTSTEEGSTSSTAESPGPGSTSESPSGTAPGTSPSGESST APGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSEPATSGSETPGTSESAT PESGPGSEPATSGSETPGSTSSTAESPGPGSTSSTAESPGPGTSPSGESSTAPGSEP ATSGSETPGSEPATSGSETPGTSTEPSEGSAPGSTSSTAESPGPGTSTPESGSASPG STSESPSGTAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSSTPSGATG SPGSSPSASTGTGPGASPGTSSTGSPGSEPATSGSETPGTSESATPESGPGSPAGSP TSTEEGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTSESATPESGPGTST EPSEGSAPGTSTEPSEGSAPG	815

Nombre*	Secuencia de proteínas	SEQ ID NO:
AM923-aIL6R_VH-AE42-aCD40_VL	MAEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGTSTEPS EGSAPGSEPATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTPESGSASPGSTS ESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPGSEPATSGSETPG TSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGS APGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESAT PESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTST EPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGTASSSPG SSTPSGATGSPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTST EEGSPAGSPTSTEEGTSTEPSEGSAPGASASGAPSTGGTSESATPESGPGSPAGSP TSTEEGSPAGSPTSTEEGSTSSTAESPGPGSTSESPSGTAPGTSPSGESSTAPGTPG SGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSEPATSGSETPGTSESATPESGPG SEPATSGSETPGSTSSTAESPGPGSTSSTAESPGPGTSPSGESSTAPGSEPATSGSE TPGSEPATSGSETPGTSTEPSEGSAPGSTSSTAESPGPGTSTPESGSASPGSTSESPS GTAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSSTPSGATGSPGSSPS ASTGTGPGASPGTSSTGSPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGS STPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTSESATPESGPGTSTEPSEGSA PGTSTEPSEGSAPGQVQLQESGPGLVRPSQTLSLTCTVSGYSITSDHAWSWVRQ PPGRGLEWIGYISYSGITTYNPSLKSRVTMLRDTSKNQFSLRLSSVTAADTAVYY CARSLARTTAMDYWGQGSLVTVSSGGAPGTSESATPESGPGSEPATSGSETPGT STEPSEGSAPGPAGMAEIVLTQSPATLSLSPGERATLSCRASQSISDYLHWYQQK PGQAPRLLIYYASHSISGIPARFSGSGSGTDFTLTISSLEPEDFAVYYCQHGHSYP WTFGGGTKVEIKG	816
aIL6R_VL-Y32-aIL6R_VH-AM1296	DIQMTQSPSSLSASVGDRVTITCRASQDISSYLNWYQQKPGKAPKLLIYYTSRLH SGVPSRFSGSGSGTDFTFTISSLQPEDIATYYCQQGNTLPYTFGQGTKVEIKGTGS GEGSEGEGGGEGSEGEGSGEGGEGEGSGSGQVQLQESGPGLVRPSQTLSLTCTV SGYSITSDHAWSWVRQPPGRGLEWIGYISYSGITTYNPSLKSRVTMLRDTSKNQ FSLRLSSVTAADTAVYYCARSLARTTAMDYWGQGSLVTVSSGGTSTEPSEGSA PGSEPATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTPESGSASPGSTSESPS GTAPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPGSEPATSGSETPGTSES ATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGT STEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESG PGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSE GSAPGSEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGTASSSPGSSTP SGATGSPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGS PAGSPTSTEEGTSTEPSEGSAPGPEPTGPAPSGGSEPATSGSETPGTSESATPESGP GSPAGSPTSTEEGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPE SGPGSPAGSPTSTEEGSPAGSPTSTEEGSTSSTAESPGPGSTSESPSGTAPGTSPSG	817

Nombre*	Secuencia de proteínas	SEQ ID NO:
	ESSTAPGSTSESPSGTAPGSTSESPSGTAPGTSPSGESSTAPGTSTEPSEGSAPGTS ESATPESGPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSESATPESGP GTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSPSGESSTAPGTSPSGESS TAPGTSPSGESSTAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGSSPSA STGTGPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGA SPGTSSTGSPGASASGAPSTGGTSPSGESSTAPGSTSSTAESPGPGTSPSGESSTAP GTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSSPSASTGTGPGSSTPSGAT GSPGASPGTSSTGSPGTSTPESGSASPGTSPSGESSTAPGTSPSGESSTAPGTSESA TPESGPGSEPATSGSETPGTSTEPSEGSAPGSTSESPSGTAPGSTSESPSGTAPGTS TPESGSASPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEE GTSESATPESGPGSEPATSGSETPGSSTPSGATGSPGASPGTSSTGSPGSSTPSGAT GSPGSTSESPSGTAPGTSPSGESSTAPGSTSSTAESPGPGSSTPSGATGSPGASPGT SSTGSPGTPGSGTASSSPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPG	
aCD40-AE864	MAEIVLTQSPATLSLSPGERATLSCRASQSISDYLHWYQQKPGQAPRLLIYYASH SISGIPARFSGSGSGTDFTLTISSLEPEDFAVYYCQHGHSYPWTFGGGTKVEIKTG SGEGSEGEGGGEGSEGEGSGEGGEGEGSGTQVQLVQSGSELKKPGASVKVSCK ASGYAFTTTGMQWVRQAPGQGLEWMGWINTHSGVPKYVEDFKGRFVFSLDTS VSTAYLQISSLKAEDTAVYYCARSGNGNYDLAYFKYWGQGTLVTVSSGGSPA GSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPG TSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTST EEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPS EGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEP ATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPG SPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGS APGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPS EGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSE SATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPG SPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGS APGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSE SATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPG TSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPS EGSAP	818

Nombre*	Secuencia de proteínas	SEQ ID NO:
AE912-aCD40	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGSPAGSP TSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTST EPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATS GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPA GSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPES GPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSE SATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGS APGEIVLTQSPATLSLSPGERATLSCRASQSISDYLHWYQQKPGQAPRLLIYYAS HSISGIPARFSGSGSGTDFTLTISSLEPEDFAVYYCQHGHSYPWTFGGGTKVEIKT GSGEGSEGEGGGEGSEGEGSGEGGEGEGSGTQVQLVQSGSELKKPGASVKVSC KASGYAFTTTGMQWVRQAPGQGLEWMGWINTHSGVPKYVEDFKGRFVFSLD TSVSTAYLQISSLKAEDTAVYYCARSGNGNYDLAYFKYWGQGTLVTVS	819
AE912-aCD40_VH-AF144-aCD40_VL	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGSPAGSP TSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTST EPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG	820
	TSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATS GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPA GSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPES GPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSE SATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGS APGQVQLVQSGSELKKPGASVKVSCKASGYAFTTTGMQWVRQAPGQGLEWM GWINTHSGVPKYVEDFKGRFVFSLDTSVSTAYLQISSLKAEDTAVYYCARSGN GNYDLAYFKYWGQGTLVTVSSGGTSTPESGSASPGTSPSGESSTAPGTSPSGESS TAPGSTSSTAESPGPGSTSESPSGTAPGSTSSTAESPGPGTSPSGESSTAPGTSTPE SGSASPGSTSSTAESPGPGTSPSGESSTAPGTSPSGESSTAPGTSPSGESSTAPGMA EIVLTQSPATLSLSPGERATLSCRASQSISDYLHWYQQKPGQAPRLLIYYASHSIS GIPARFSGSGSGTDFTLTISSLEPEDFAVYYCQHGHSYPWTFGGGTKVEIKG	

Nombre*	Secuencia de proteínas	SEQ ID NO:
aCD40_VL-AE144-aCD40_VH-AE576	MAEIVLTQSPATLSLSPGERATLSCRASQSISDYLHWYQQKPGQAPRLLIYYASH SISGIPARFSGSGSGTDFTLTISSLEPEDFAVYYCQHGHSYPWTFGGGTKVEIKGG SEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGS APGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTSESAT PESGPGSEPATSGSETPGTSTEPSEGSAPGQVQLVQSGSELKKPGASVKVSCKAS GYAFTTTGMQWVRQAPGQGLEWMGWINTHSGVPKYVEDFKGRFVFSLDTSVS TAYLQISSLKAEDTAVYYCARSGNGNYDLAYFKYWGQGTLVTVSSGGSPAGSP TSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTST EPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATS GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPA GSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG	821
aCD40_VL-AE42-aCD40_VH-BC864	MAEIVLTQSPATLSLSPGERATLSCRASQSISDYLHWYQQKPGQAPRLLIYYASH SISGIPARFSGSGSGTDFTLTISSLEPEDFAVYYCQHGHSYPWTFGGGTKVEIKGG APGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGPAGQVQLVQSGSELKK PGASVKVSCKASGYAFTTTGMQWVRQAPGQGLEWMGWINTHSGVPKYVEDF KGRFVFSLDTSVSTAYLQISSLKAEDTAVYYCARSGNGNYDLAYFKYWGQGTL VTVSSGGTSTEPSEPGSAGTSTEPSEPGSAGSEPATSGTEPSGSGASEPTSTEPGSE PATSGTEPSGSEPATSGTEPSGSEPATSGTEPSGSGASEPTSTEPGTSTEPSEPGSA GSEPATSGTEPSGTSTEPSEPGSAGSEPATSGTEPSGSEPATSGTEPSGTSTEPSEP GSAGTSTEPSEPGSAGSEPATSGTEPSGSEPATSGTEPSGTSEPSTSEPGAGSGAS EPTSTEPGTSEPSTSEPGAGSEPATSGTEPSGSEPATSGTEPSGTSTEPSEPGSAGT STEPSEPGSAGSGASEPTSTEPGSEPATSGTEPSGSEPATSGTEPSGSEPATSGTEP SGSEPATSGTEPSGTSTEPSEPGSAGSEPATSGTEPSGSGASEPTSTEPGTSTEPSE PGSAGSEPATSGTEPSGSGASEPTSTEPGTSTEPSEPGSAGSGASEPTSTEPGSEPA TSGTEPSGSGASEPTSTEPGSEPATSGTEPSGSGASEPTSTEPGTSTEPSEPGSAGS EPATSGTEPSGSGASEPTSTEPGTSTEPSEPGSAGSEPATSGTEPSGTSTEPSEPGS AGSEPATSGTEPSGTSTEPSEPGSAGTSTEPSEPGSAGTSTEPSEPGSAGTSTEPSE PGSAGTSTEPSEPGSAGTSTEPSEPGSAGTSEPSTSEPGAGSGASEPTSTEPGTSTE PSEPGSAGTSTEPSEPGSAGTSTEPSEPGSAGSEPATSGTEPSGSGASEPTSTEPGS EPATSGTEPSGSEPATSGTEPSGSEPATSGTEPSGSEPATSGTEPSGTSEPSTSEPG AGSEPATSGTEPSGSGASEPTSTEPGTSTEPSEPGSAGSEPATSGTEPSGSGASEPT STEPGTSTEPSEPGSAG	822

Nombre*	Secuencia de proteínas	SEQ ID NO:
AM923-aCD40_VH-AE42-BD864	MAEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGTSTEPS EGSAPGSEPATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTPESGSASPGSTS ESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPGSEPATSGSETPG TSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGS APGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESAT PESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTST EPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGTASSSPG SSTPSGATGSPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTST EEGSPAGSPTSTEEGTSTEPSEGSAPGASASGAPSTGGTSESATPESGPGSPAGSP TSTEEGSPAGSPTSTEEGSTSSTAESPGPGSTSESPSGTAPGTSPSGESSTAPGTPG SGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSEPATSGSETPGTSESATPESGPG SEPATSGSETPGSTSSTAESPGPGSTSSTAESPGPGTSPSGESSTAPGSEPATSGSE TPGSEPATSGSETPGTSTEPSEGSAPGSTSSTAESPGPGTSTPESGSASPGSTSESPS GTAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSSTPSGATGSPGSSPS ASTGTGPGASPGTSSTGSPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGS STPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTSESATPESGPGTSTEPSEGSA PGTSTEPSEGSAPGQVQLVQSGSELKKPGASVKVSCKASGYAFTTTGMQWVRQ APGQGLEWMGWINTHSGVPKYVEDFKGRFVFSLDTSVSTAYLQISSLKAEDTA VYYCARSGNGNYDLAYFKYWGQGTLVTVSSGGAPGTSESATPESGPGSEPATS GSETPGTSTEPSEGSAPGPAGGSETATSGSETAGTSESATSESGAGSTAGSETSTE AGTSESATSESGAGSETATSGSETAGSETATSGSETAGTSTEASEGSASGTSTEA SEGSASGTSESATSESGAGSETATSGSETAGTSTEASEGSASGSTAGSETSTEAGT SESATSESGAGTSESATSESGAGSETATSGSETAGTSESATSESGAGTSTEASEGS ASGSETATSGSETAGSETATSGSETAGTSTEASEGSASGSTAGSETSTEAGTSES ATSESGAGTSTEASEGSASGSETATSGSETAGSTAGSETSTEAGSTAGSETSTEA GSETATSGSETAGTSESATSESGAGTSESATSESGAGSETATSGSETAGTSESATS ESGAGTSESATSESGAGSETATSGSETAGSETATSGSETAGTSTEASEGSASGST AGSETSTEAGSETATSGSETAGTSESATSESGAGSTAGSETSTEAGSTAGSETSTE AGSTAGSETSTEAGTSTEASEGSASGSTAGSETSTEAGSTAGSETSTEAGTSTEA SEGSASGSTAGSETSTEAGSETATSGSETAGTSTEASEGSASGTSESATSESGAGS ETATSGSETAGTSESATSESGAGTSESATSESGAGSETATSGSETAGTSESATSES GAGSETATSGSETAGTSTEASEGSASGTSTEASEGSASGSTAGSETSTEAGSTAG SETSTEAGSETATSGSETAGTSESATSESGAGTSESATSESGAGSETATSGSETAG SETATSGSETAGSETATSGSETAGTSTEASEGSASGTSESATSESGAGSETATSGS ETAGSETATSGSETAGTSESATSESGAGTSESATSESGAGSETATSGSETAG	823
aCD40 _VL-Y32-aCD40_VH-AE576	MAEIVLTQSPATLSLSPGERATLSCRASQSISDYLHWYQQKPGQAPRLLIYYASH SISGIPARFSGSGSGTDFTLTISSLEPEDFAVYYCQHGHSYPWTFGGGTKVEIKGT GSGEGSEGEGGGEGSEGEGSGEGGGEGEGSGSGQVQLVQSGSELKKPGASVKVS CKASGYAFTTTGMQWVRQAPGQGLEWMGWINTHSGVPKYVEDFKGRFVFSL DTSVSTAYLQISSLKAEDTAVYYCARSGNGNYDLAYFKYWGQGTLVTVSSGGS PAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSA PGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPT STEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTE PSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGS EPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESG PGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSE GSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTE PSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGT SESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESG PGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSE GSAPG	824

Nombre*	Secuencia de proteínas	SEQ ID NO:
AE912-aCD40_VH-AF144-aCD40_VL	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGSPAGSP TSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTST EPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATS GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPA GSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG	825
	TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPES GPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSE SATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGS APGQVQLVQSGSELKKPGASVKVSCKASGYAFTTTGMQWVRQAPGQGLEWM GWINTHSGVPKYVEDFKGRFVFSLDTSVSTAYLQISSLKAEDTAVYYCARSGN GNYDLAYFKYWGQGTLVTVSSGGTSTPESGSASPGTSPSGESSTAPGTSPSGESS TAPGSTSSTAESPGPGSTSESPSGTAPGSTSSTAESPGPGTSPSGESSTAPGTSTPE SGSASPGSTSSTAESPGPGTSPSGESSTAPGTSPSGESSTAPGTSPSGESSTAPGMA EIVLTQSPATLSLSPGERATLSCRASQSISDYLHWYQQKPGQAPRLLIYYASHSIS GIPARFSGSGSGTDFTLTISSLEPEDFAVYYCQHGHSYPWTFGGGTKVEIKG	
aCD40_VL-AE144-aCD40_VH-AE576	MAEIVLTQSPATLSLSPGERATLSCRASQSISDYLHWYQQKPGQAPRLLIYYASH SISGIPARFSGSGSGTDFTLTISSLEPEDFAVYYCQHGHSYPWTFGGGTKVEIKGG SEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGS APGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTSESAT PESGPGSEPATSGSETPGTSTEPSEGSAPGQVQLVQSGSELKKPGASVKVSCKAS GYAFTTTGMQWVRQAPGQGLEWMGWINTHSGVPKYVEDFKGRFVFSLDTSVS TAYLQISSLKAEDTAVYYCARSGNGNYDLAYFKYWGQGTLVTVSSGGSPAGSP TSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTST EPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATS GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPA GSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG	826

Nombre*	Secuencia de proteínas	SEQ ID NO:
aHER2-AE864	MEDIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSA SFLYSGVPSRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIK TGSGEGSEGEGGGEGSEGEGSGEGGEGEGSGTEVQLVESGGGLVQPGGSLRLS CAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYTRYADSVKGRFTISADT SKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGTLVTVSGSPGSP AGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAP GTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTS TEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEP SEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSE PATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGP GSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEG SAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEP SEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTS ESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGP GSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEG SAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESA TPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTS ESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGP GTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTS TEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEP SEGSAP	827
AE912-aHER2	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGSPAGSP TSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTST EPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS	828
	APGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATS GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPA GSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPES GPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSE SATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGS APGSSSLDIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLL IYSASFLYSGVPSRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTK VEIKTGSGEGSEGEGGGEGSEGEGSGEGGEGEGSGTEVQLVESGGGLVQPGGSL RLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYTRYADSVKGRFTIS ADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGTLVTVS	

Nombre*	Secuencia de proteínas	SEQ ID NO:
aHer2_VL-AE42-aHer2_VH-AM1296	MEDIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSA SFLYSGVPSRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIK GGAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGPAGEVQLVESGGGL VQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYTRYADSV KGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGTL VTVSGGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTST PESGSASPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPG SEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPES GPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS EGSAPGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSE SATPESGPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPG TPGSGTASSSPGSSTPSGATGSPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSE TPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGPEPTGPAPSGGSEPATSG SETPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSPAGSPTSTEEGSPAG SPTSTEEGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSSTAESPGPGS TSESPSGTAPGTSPSGESSTAPGSTSESPSGTAPGSTSESPSGTAPGTSPSGESSTAP GTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGTSESATPE SGPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSPSG ESSTAPGTSPSGESSTAPGTSPSGESSTAPGTSTEPSEGSAPGSPAGSPTSTEEGTS TEPSEGSAPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSP GSSTPSGATGSPGASPGTSSTGSPGASASGAPSTGGTSPSGESSTAPGSTSSTAES PGPGTSPSGESSTAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSSPSA STGTGPGSSTPSGATGSPGASPGTSSTGSPGTSTPESGSASPGTSPSGESSTAPGTS PSGESSTAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGSTSESPSGTAP GSTSESPSGTAPGTSTPESGSASPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEG SAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGSSTPSGATGSPGASPGT SSTGSPGSSTPSGATGSPGSTSESPSGTAPGTSPSGESSTAPGSTSSTAESPGPGSS TPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSPAGSPTSTEEGSPAGSPTSTEE GTSTEPSEGSAPG	829
AE912-aHER2_VL-AF144-aHER2_VH	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGSPAGSP TSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTST EPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATS GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPA GSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG	830

Nombre*	Secuencia de proteínas	SEQ ID NO:
	TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPES GPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSE SATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGS APGMEDIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLI YSASFLYSGVPSRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTK VEIKGGTSTPESGSASPGTSPSGESSTAPGTSPSGESSTAPGSTSSTAESPGPGSTS ESPSGTAPGSTSSTAESPGPGTSPSGESSTAPGTSTPESGSASPGSTSSTAESPGPG TSPSGESSTAPGTSPSGESSTAPGTSPSGESSTAPGEVQLVESGGGLVQPGGSLRL SCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYTRYADSVKGRFTISAD TSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGTLVTVSG	
AE48-aHER2_VH-AE144-aHER2_VL-AE576	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGEVQLVE SGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYTR YADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYW GQGTLVTVSGGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTST EEGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSTEPS EGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGMEDIQMTQSPSSLS ASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVPSRFSGSR SGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKGGSPAGSPTSTEEGT SESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSA PGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATP ESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTE PSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGT STEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTE EGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSE GSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAG SPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGS EPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTE EGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG	831
aHer2_VH_1 -AE288-aHer2_VL_1-AF576	EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYP TNGYTRYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFY AMDYWGQGTLVTVSGGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEP ATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPG SEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGS APGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATS GSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSE SATPESGPGTSTEPSEGSAPGMEDIQMTQSPSSLSASVGDRVTITCRASQDVNTA VAWYQQKPGKAPKLLIYSASFLYSGVPSRFSGSRSGTDFTLTISSLQPEDFATYY CQQHYTTPPTFGQGTKVEIKGGSTSSTAESPGPGSTSSTAESPGPGSTSESPSGTA PGSTSSTAESPGPGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAPGTSPSGESS TAPGSTSESPSGTAPGSTSESPSGTAPGTSPSGESSTAPGSTSESPSGTAPGSTSESP SGTAPGTSPSGESSTAPGSTSESPSGTAPGSTSESPSGTAPGSTSESPSGTAPGTST PESGSASPGSTSESPSGTAPGTSTPESGSASPGSTSSTAESPGPGSTSSTAESPGPG TSTPESGSASPGTSTPESGSASPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSAS PGSTSESPSGTAPGSTSESPSGTAPGSTSESPSGTAPGSTSSTAESPGPGTSTPESGS ASPGTSTPESGSASPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGSTSESP SGTAPGSTSESPSGTAPGTSTPESGSASPGTSPSGESSTAPGSTSSTAESPGPGTSP SGESSTAPGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAPGSTSSTAESPGPG TSTPESGSASPGTSTPESGSASPG	832

Nombre*	Secuencia de proteínas	SEQ ID NO:
aEGFR-Y576	MEDILLTQSPVILSVSPGERVSFSCRASQSIGTNIHWYQQRTNGSPRLLIKYASESI SGIPSRFSGSGSGTDFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLELKTGS GEGSEGEGGGEGSEGEGSGEGGEGEGSGTQVQLKQSGPGLVQPSQSLSITCTVS GFSLTNYGVHWVRQSPGKGLEWLGVIWSGGNTDYNTPFTSRLSINKDNSKSQV FFKMNSLQSNDTAIYYCARALTYYDYEFAYWGQGTLVTVSGGEGSGEGSEGE GSEGSGEGEGSEGSGEGEGGSEGSEGEGSEGSGEGEGGEGSGEGEGSGEGSEGE GGGEGSEGEGSGEGGEGEGSEGGSEGEGGSEGGEGEGSEGSGEGEGSEGGSEG EGSEGGSEGEGSEGSGEGEGSEGSGEGEGSEGSGEGEGSEGSGEGEGSEGGSEG	833
	EGGSEGSEGEGSGEGSEGEGGSEGSEGEGGGEGSEGEGSGEGSEGEGGSEGSEG EGGSEGSEGEGGEGSGEGEGSEGSGEGEGSGEGSEGEGSEGSGEGEGSEGSGEG EGGSEGSEGEGSGEGSEGEGSEGSGEGEGSEGSGEGEGGSEGSEGEGGSEGSEG EGGSEGSEGEGGEGSGEGEGSEGSGEGEGSGEGSEGEGSEGSGEGEGSEGSGEG EGGSEGSEGEGSEGSGEGEGGEGSGEGEGSGEGSEGEGGGEGSEGEGSEGSGEG EGSEGSGEGEGSEGGSEGEGGSEGSEGEGSEGGSEGEGSEGGSEGEGSEGSGEG EGSEGSGEGEGSGEGSEGEGGSEGGEGEGSEGGSEGEGSEGGSEGEGGEGSGEG EGGGEGSEGEGSEGSGEGEGSGEGSEG	
aEGFR VL-AE42-aEGFR VH-AD836	MEDILLTQSPVILSVSPGERVSFSCRASQSIGTNIHWYQQRTNGSPRLLIKYASESI SGIPSRFSGSGSGTDFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLELKGGA PGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGPAGQVQLKQSGPGLVQPS QSLSITCTVSGFSLTNYGVHWVRQSPGKGLEWLGVIWSGGNTDYNTPFTSRLSI NKDNSKSQVFFKMNSLQSNDTAIYYCARALTYYDYEFAYWGQGTLVTVSGGS SESGSSEGGPGSSESGSSEGGPGESPGGSSGSESGSGGEPSESGSSGESPGGSSGSE SGESPGGSSGSESGSSESGSSEGGPGSSESGSSEGGPGSSESGSSEGGPGESPGGSS GSESGESPGGSSGSESGESPGGSSGSESGSSESGSSEGGPGSSESGSSEGGPGSSES GSSEGGPGSSESGSSEGGPGSSESGSSEGGPGSSESGSSEGGPGSGGEPSESGSSG ESPGGSSGSESGESPGGSSGSESGSGGEPSESGSSGSEGSSGPGESSGSSESGSSEG GPGSGGEPSESGSSGSEGSSGPGESSGSSESGSSEGGPGSGGEPSESGSSGESPGG SSGSESGSGGEPSESGSSGSGGEPSESGSSGSSESGSSEGGPGSGGEPSESGSSGSG GEPSESGSSGSEGSSGPGESSGESPGGSSGSESGSEGSSGPGESSGSEGSSGPGESS GSGGEPSESGSSGSSESGSSEGGPGSSESGSSEGGPGESPGGSSGSESGSGGEPSES GSSGSEGSSGPGESSGESPGGSSGSESGSEGSSGPGSSESGSSEGGPGSGGEPSES GSSGSEGSSGPGESSGSEGSSGPGESSGSEGSSGPGESSGSGGEPSESGSSGSGGE PSESGSSGESPGGSSGSESGESPGGSSGSESGSGGEPSESGSSGSEGSSGPGESSGE SPGGSSGSESGSSESGSSEGGPGSSESGSSEGGPGSSESGSSEGGPGSGGEPSESGS SGSSESGSSEGGPGESPGGSSGSESGSGGEPSESGSSGSSESGSSEGGPGESPGGSS GSESGSGGEPSESGSSGESPGGSSGSESGSGGEPSESGSSG	834

Nombre*	Secuencia de proteínas	SEQ ID NO:
AM923-aEGFR VL-AM150-aEGFR_VH_ 1	MAEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGTSTEPS EGSAPGSEPATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTPESGSASPGSTS ESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPGSEPATSGSETPG TSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGS APGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESAT PESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTST EPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGTASSSPG SSTPSGATGSPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTST EEGSPAGSPTSTEEGTSTEPSEGSAPGASASGAPSTGGTSESATPESGPGSPAGSP TSTEEGSPAGSPTSTEEGSTSSTAESPGPGSTSESPSGTAPGTSPSGESSTAPGTPG SGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSEPATSGSETPGTSESATPESGPG SEPATSGSETPGSTSSTAESPGPGSTSSTAESPGPGTSPSGESSTAPGSEPATSGSE TPGSEPATSGSETPGTSTEPSEGSAPGSTSSTAESPGPGTSTPESGSASPGSTSESPS GTAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSSTPSGATGSPGSSPS ASTGTGPGASPGTSSTGSPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGS STPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTSESATPESGPGTSTEPSEGSA PGTSTEPSEGSAPGMEDILLTQSPVILSVSPGERVSFSCRASQSIGTNIHWYQQRT NGSPRLLIKYASESISGIPSRFSGSGSGTDFTLSINSVESEDIADYYCQQNNNWPT TFGAGTKLELKGGAPSTGGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGT SESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSA PGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGPEPTGPA GQVQLKQSGPGLVQPSQSLSITCTVSGFSLTNYGVHWVRQSPGKGLEWLGVIW SGGNTDYNTPFTSRLSINKDNSKSQVFFKMNSLQSNDTAIYYCARALTYYDYEF AYWGQGTLVTVSG	835
aEGFR VH-AF144-aEGFR VL-AF864	QVQLKQSGPGLVQPSQSLSITCTVSGFSLTNYGVHWVRQSPGKGLEWLGVIWS GGNTDYNTPFTSRLSINKDNSKSQVFFKMNSLQSNDTAIYYCARALTYYDYEFA YWGQGTLVTVSGGTSTPESGSASPGTSPSGESSTAPGTSPSGESSTAPGSTSSTAE SPGPGSTSESPSGTAPGSTSSTAESPGPGTSPSGESSTAPGTSTPESGSASPGSTSST AESPGPGTSPSGESSTAPGTSPSGESSTAPGTSPSGESSTAPGMEDILLTQSPVILS VSPGERVSFSCRASQSIGTNIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSGT	836
	DFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLELKGGSTSESPSGTAPGTSP SGESSTAPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPG STSESPSGTAPGSTSESPSGTAPGTSPSGESSTAPGSTSESPSGTAPGTSPSGESSTA PGTSPSGESSTAPGSTSSTAESPGPGTSPSGESSTAPGTSPSGESSTAPGSTSSTAES PGPGTSTPESGSASPGTSTPESGSASPGSTSESPSGTAPGSTSESPSGTAPGTSTPES GSASPGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAPGTSPSGESSTAPGSTS STAESPGPGTSPSGESSTAPGTSTPESGSASPGSTSSTAESPGPGSTSSTAESPGPG STSSTAESPGPGSTSSTAESPGPGTSPSGESSTAPGSTSESPSGTAPGSTSESPSGTA PGTSTPESGPXXXGASASGAPSTXXXXSESPSGTAPGSTSESPSGTAPGSTSESPS GTAPGSTSESPSGTAPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSPS GESSTAPGTSPSGESSTAPGSTSSTAESPGPGTSPSGESSTAPGTSTPESGSASPGS TSESPSGTAPGSTSESPSGTAPGTSPSGESSTAPGSTSESPSGTAPGTSTPESGSASP GTSTPESGSASPGSTSESPSGTAPGTSTPESGSASPGSTSSTAESPGPGSTSESPSGT APGSTSESPSGTAPGTSPSGESSTAPGSTSSTAESPGPGTSPSGESSTAPGTSTPES GSASPGTSPSGESSTAPGTSPSGESSTAPGTSPSGESSTAPGSTSSTAESPGPGSTS STAESPGPGTSPSGESSTAPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGSPG	

Nombre*	Secuencia de proteínas	SEQ ID NO:
AE912-aEGFR VH-Y32-aEGFR VL	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGSPAGSP TSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTST EPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATS GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPA GSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPES GPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSE SATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGS APGQVQLKQSGPGLVQPSQSLSITCTVSGFSLTNYGVHWVRQSPGKGLEWLGV IWSGGNTDYNTPFTSRLSINKDNSKSQVFFKMNSLQSNDTAIYYCARALTYYDY EFAYWGQGTLVTVSGTGSGEGSEGEGGGEGSEGEGSGEGGEGEGSGSGMEDIL LTQSPVILSVSPGERVSFSCRASQSIGTNIHWYQQRTNGSPRLLIKYASESISGIPS RFSGSGSGTDFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLELKG	837
aEGFR VL-Conector_Y32-aEGFR VH-BC864	MEDILLTQSPVILSVSPGERVSFSCRASQSIGTNIHWYQQRTNGSPRLLIKYASESI SGIPSRFSGSGSGTDFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLELKGTG SGEGSEGEGGGEGSEGEGSGEGGEGEGSGSGQVQLKQSGPGLVQPSQSLSITCT VSGFSLTNYGVHWVRQSPGKGLEWLGVIWSGGNTDYNTPFTSRLSINKDNSKS QVFFKMNSLQSNDTAIYYCARALTYYDYEFAYWGQGTLVTVSGGTSTEPSEPG SAGTSTEPSEPGSAGSEPATSGTEPSGSGASEPTSTEPGSEPATSGTEPSGSEPATS GTEPSGSEPATSGTEPSGSGASEPTSTEPGTSTEPSEPGSAGSEPATSGTEPSGTST EPSEPGSAGSEPATSGTEPSGSEPATSGTEPSGTSTEPSEPGSAGTSTEPSEPGSAG SEPATSGTEPSGSEPATSGTEPSGTSEPSTSEPGAGSGASEPTSTEPGTSEPSTSEP GAGSEPATSGTEPSGSEPATSGTEPSGTSTEPSEPGSAGTSTEPSEPGSAGSGASE PTSTEPGSEPATSGTEPSGSEPATSGTEPSGSEPATSGTEPSGSEPATSGTEPSGTS TEPSEPGSAGSEPATSGTEPSGSGASEPTSTEPGTSTEPSEPGSAGSEPATSGTEPS GSGASEPTSTEPGTSTEPSEPGSAGSGASEPTSTEPGSEPATSGTEPSGSGASEPTS TEPGSEPATSGTEPSGSGASEPTSTEPGTSTEPSEPGSAGSEPATSGTEPSGSGASE PTSTEPGTSTEPSEPGSAGSEPATSGTEPSGTSTEPSEPGSAGSEPATSGTEPSGTS TEPSEPGSAGTSTEPSEPGSAGTSTEPSEPGSAGTSTEPSEPGSAGTSTEPSEPGSA GTSTEPSEPGSAGTSEPSTSEPGAGSGASEPTSTEPGTSTEPSEPGSAGTSTEPSEP GSAGTSTEPSEPGSAGSEPATSGTEPSGSGASEPTSTEPGSEPATSGTEPSGSEPAT SGTEPSGSEPATSGTEPSGSEPATSGTEPSGTSEPSTSEPGAGSEPATSGTEPSGSG ASEPTSTEPGTSTEPSEPGSAGSEPATSGTEPSGSGASEPTSTEPGTSTEPSEPGSA G	838

Nombre*	Secuencia de proteínas	SEQ ID NO:
aEGFR VH-AE288-aEGFR VL-BD864	QVQLKQSGPGLVQPSQSLSITCTVSGFSLTNYGVHWVRQSPGKGLEWLGVIWS GGNTDYNTPFTSRLSINKDNSKSQVFFKMNSLQSNDTAIYYCARALTYYDYEFA YWGQGTLVTVSGGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATS GSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEP ATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPG TSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSE TPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESAT PESGPGTSTEPSEGSAPGMEDILLTQSPVILSVSPGERVSFSCRASQSIGTNIHWY QQRTNGSPRLLIKYASESISGIPSRFSGSGSGTDFTLSINSVESEDIADYYCQQNN NWPTTFGAGTKLELKGGSETATSGSETAGTSESATSESGAGSTAGSETSTEAGT SESATSESGAGSETATSGSETAGSETATSGSETAGTSTEASEGSASGTSTEASEGS ASGTSESATSESGAGSETATSGSETAGTSTEASEGSASGSTAGSETSTEAGTSESA TSESGAGTSESATSESGAGSETATSGSETAGTSESATSESGAGTSTEASEGSASGS ETATSGSETAGSETATSGSETAGTSTEASEGSASGSTAGSETSTEAGTSESATSES GAGTSTEASEGSASGSETATSGSETAGSTAGSETSTEAGSTAGSETSTEAGSETA TSGSETAGTSESATSESGAGTSESATSESGAGSETATSGSETAGTSESATSESGAG TSESATSESGAGSETATSGSETAGSETATSGSETAGTSTEASEGSASGSTAGSETS TEAGSETATSGSETAGTSESATSESGAGSTAGSETSTEAGSTAGSETSTEAGSTA GSETSTEAGTSTEASEGSASGSTAGSETSTEAGSTAGSETSTEAGTSTEASEGSAS GSTAGSETSTEAGSETATSGSETAGTSTEASEGSASGTSESATSESGAGSETATS GSETAGTSESATSESGAGTSESATSESGAGSETATSGSETAGTSESATSESGAGSE TATSGSETAGTSTEASEGSASGTSTEASEGSASGSTAGSETSTEAGSTAGSETSTE AGSETATSGSETAGTSESATSESGAGTSESATSESGAGSETATSGSETAGSETAT SGSETAGSETATSGSETAGTSTEASEGSASGTSESATSESGAGSETATSGSETAGS ETATSGSETAGTSESATSESGAGTSESATSESGAGSETATSGSETAG	839
aCD3-Y288	MKDIQMTQSPSSLSASVGDRVTITCSASSSVSYMNWYQQTPGKAPKRWIYDTS KLASGVPSRFSGSGSGTDYTFTISSLQPEDIATYYCQQWSSNPFTFGQGTKLQIT RTGSGEGSEGEGGGEGSEGEGSGEGGEGEGSGTQVQLVQSGGGVVQPGRSLRL SCKASGYTFTRYTMHWVRQAPGKGLEWIGYINPSRGYTNYNQKFKDRFTISTD KSKSTAFLQMDSLRPEDTAVYYSARYYDDHYCLDYWGQGTPVTVSSGGEGSG EGSEGEGSEGSGEGEGSEGSGEGEGGSEGSEGEGGSEGSEGEGSGEGSEGEGSG EGSEGEGGSEGSEGEGSGEGSEGEGSEGGSEGEGGSEGSEGEGSGEGSEGEGGE GGSEGEGSEGSGEGEGSGEGSEGEGSEGSGEGEGSGEGSEGEGSEGSGEGEGSE GSGEGEGGSEGSEGEGSEGSGEGEGGEGSGEGEGSGEGSEGEGGGEGSEGEGS GEGGEGEGSEGGSEGEGGSEGGEGEGSEGSGEGEGSEGGSEGEGSEGGSEGEGS EGSGEGEGSEGSGEGS	840

Nombre*	Secuencia de proteínas	SEQ ID NO:
aCD3_VL-AF144-aCD3_VH-AE576	MKDIQMTQSPSSLSASVGDRVTITCSASSSVSYMNWYQQTPGKAPKRWIYDTS KLASGVPSRFSGSGSGTDYTFTISSLQPEDIATYYCQQWSSNPFTFGQGTKLQIT RGGTSTPESGSASPGTSPSGESSTAPGTSPSGESSTAPGSTSSTAESPGPGSTSESP SGTAPGSTSSTAESPGPGTSPSGESSTAPGTSTPESGSASPGSTSSTAESPGPGTSP SGESSTAPGTSPSGESSTAPGTSPSGESSTAPGQVQLVQSGGGVVQPGRSLRLSC KASGYTFTRYTMHWVRQAPGKGLEWIGYINPSRGYTNYNQKFKDRFTISTDKS KSTAFLQMDSLRPEDTAVYYSARYYDDHYCLDYWGQGTPVTVSSGGSPAGSP TSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTST EPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATS GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPA GSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG	841
AM923-aCD3_VH-AM150-aCD3_VL	MAEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGTSTEPS EGSAPGSEPATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTPESGSASPGSTS ESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPGSEPATSGSETPG TSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGS APGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESAT	842
	PESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTST EPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGTASSSPG SSTPSGATGSPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTST EEGSPAGSPTSTEEGTSTEPSEGSAPGASASGAPSTGGTSESATPESGPGSPAGSP TSTEEGSPAGSPTSTEEGSTSSTAESPGPGSTSESPSGTAPGTSPSGESSTAPGTPG SGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSEPATSGSETPGTSESATPESGPG SEPATSGSETPGSTSSTAESPGPGSTSSTAESPGPGTSPSGESSTAPGSEPATSGSE TPGSEPATSGSETPGTSTEPSEGSAPGSTSSTAESPGPGTSTPESGSASPGSTSESPS GTAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSSTPSGATGSPGSSPS ASTGTGPGASPGTSSTGSPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGS STPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTSESATPESGPGTSTEPSEGSA PGTSTEPSEGSAPGQVQLVQSGGGVVQPGRSLRLSCKASGYTFTRYTMHWVRQ APGKGLEWIGYINPSRGYTNYNQKFKDRFTISTDKSKSTAFLQMDSLRPEDTAV YYSARYYDDHYCLDYWGQGTPVTVSSGGAPSTGGTSESATPESGPGTSTEPSE GSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSES ATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGT STEPSEGSAPGPEPTGPAGMKDIQMTQSPSSLSASVGDRVTITCSASSSVSYMNW YQQTPGKAPKRWIYDTSKLASGVPSRFSGSGSGTDYTFTISSLQPEDIATYYCQQ WSSNPFTFGQGTKLQITRG	

Nombre*	Secuencia de proteínas	SEQ ID NO:
aCD3_VL-Conector_AE42 - aCD3_VH-AM875	MKDIQMTQSPSSLSASVGDRVTITCSASSSVSYMNWYQQTPGKAPKRWIYDTS KLASGVPSRFSGSGSGTDYTFTISSLQPEDIATYYCQQWSSNPFTFGQGTKLQIT RGGAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGPAGQVQLVQSGGG VVQPGRSLRLSCKASGYTFTRYTMHWVRQAPGKGLEWIGYINPSRGYTNYNQ KFKDRFTISTDKSKSTAFLQMDSLRPEDTAVYYSARYYDDHYCLDYWGQGTPV TVSSGGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTST PESGSASPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPG SEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPES GPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS EGSAPGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSE SATPESGPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPG TPGSGTASSSPGSSTPSGATGSPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSE TPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGASASGAPSTGGTSESAT PESGPGSPAGSPTSTEEGSPAGSPTSTEEGSTSSTAESPGPGSTSESPSGTAPGTSP SGESSTAPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSEPATSGSETPG TSESATPESGPGSEPATSGSETPGSTSSTAESPGPGSTSSTAESPGPGTSPSGESST APGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGSTSSTAESPGPGTSTPES GSASPGSTSESPSGTAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSST PSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSEPATSGSETPGTSESATPESGPG SPAGSPTSTEEGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTSESATPES GPGTSTEPSEGSAPGTSTEPSEGSAPG	843
AE912-aCD3_VL-AE144-aCD3_VH	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGSPAGSP TSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTST EPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATS GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPA GSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPES GPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSE SATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEG TSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGS APGMKDIQMTQSPSSLSASVGDRVTITCSASSSVSYMNWYQQTPGKAPKRWIY DTSKLASGVPSRFSGSGSGTDYTFTISSLQPEDIATYYCQQWSSNPFTFGQGTKL	844
	QITRGGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTST EPSEGSAPGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPG TSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGQVQLVQSGGGVVQPGRSLR LSCKASGYTFTRYTMHWVRQAPGKGLEWIGYINPSRGYTNYNQKFKDRFTIST DKSKSTAFLQMDSLRPEDTAVYYSARYYDDHYCLDYWGQGTPVTVSSG	

Nombre*	Secuencia de proteínas	SEQ ID NO:
AE48-aCD3_VH-Conector_Y32-aCD3_VL-AE864	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGQVQLV QSGGGVVQPGRSLRLSCKASGYTFTRYTMHWVRQAPGKGLEWIGYINPSRGYT NYNQKFKDRFTISTDKSKSTAFLQMDSLRPEDTAVYYSARYYDDHYCLDYWG QGTPVTVSSGTGSGEGSEGEGGGEGSEGEGSGEGGEGEGSGSGMKDIQMTQSP SSLSASVGDRVTITCSASSSVSYMNWYQQTPGKAPKRWIYDTSKLASGVPSRFS GSGSGTDYTFTISSLQPEDIATYYCQQWSSNPFTFGQGTKLQITRGGSPAGSPTST EEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS EGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSE SATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPG TSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSE TPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSP TSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTST EPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPG SPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPES GPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSP TSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSE SATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPG TSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPES GPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESAT PESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTST EPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG	845
AM48-aCD3_VL-Conector_AE42 -aCD3_VH-AM875	MAEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGMKDIQ MTQSPSSLSASVGDRVTITCSASSSVSYMNWYQQTPGKAPKRWIYDTSKLASG VPSRFSGSGSGTDYTFTISSLQPEDIATYYCQQWSSNPFTFGQGTKLQITRGGAP GTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGPAGQVQLVQSGGGVVQPG RSLRLSCKASGYTFTRYTMHWVRQAPGKGLEWIGYINPSRGYTNYNQKFKDRF TISTDKSKSTAFLQMDSLRPEDTAVYYSARYYDDHYCLDYWGQGTPVTVSSGG TSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTPESGSA SPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPGSEPATS GSETPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTST EPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPG TSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPES GPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGT ASSSPGSSTPSGATGSPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPA GSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGASASGAPSTGGTSESATPESGPG SPAGSPTSTEEGSPAGSPTSTEEGSTSSTAESPGPGSTSESPSGTAPGTSPSGESST APGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSEPATSGSETPGTSESAT PESGPGSEPATSGSETPGSTSSTAESPGPGSTSSTAESPGPGTSPSGESSTAPGSEP ATSGSETPGSEPATSGSETPGTSTEPSEGSAPGSTSSTAESPGPGTSTPESGSASPG STSESPSGTAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSSTPSGATG SPGSSPSASTGTGPGASPGTSSTGSPGSEPATSGSETPGTSESATPESGPGSPAGSP TSTEEGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTSESATPESGPGTST EPSEGSAPGTSTEPSEGSAPG	846

Nombre*	Secuencia de proteínas	SEQ ID NO:
aEGRF_VHH -Conector_Y32-aEGRF_VHH 1-BC864	EVQLQESGGGLVQAGDSLRLSCLVSGRSFNSYTMGWFRQAPGKEREFVAAILW SGPTTYYADSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVYYCAAALGVLVL APGNVYSYWGQGTQVTVSSAHHGTGSGEGSEGEGGGEGSEGEGSGEGGEGEG SGSGEVQLQESGGGLVQAGDSLRLSCLVSGRSFNSYTMGWFRQAPGKEREFVA AILWSGPTTYYADSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVYYCAAALG VLVLAPGNVYSYWGQGTQVTVSSAHHGGTSTEPSEPGSAGTSTEPSEPGSAGSE PATSGTEPSGSGASEPTSTEPGSEPATSGTEPSGSEPATSGTEPSGSEPATSGTEPS GSGASEPTSTEPGTSTEPSEPGSAGSEPATSGTEPSGTSTEPSEPGSAGSEPATSGT EPSGSEPATSGTEPSGTSTEPSEPGSAGTSTEPSEPGSAGSEPATSGTEPSGSEPAT SGTEPSGTSEPSTSEPGAGSGASEPTSTEPGTSEPSTSEPGAGSEPATSGTEPSGSE	847
	PATSGTEPSGTSTEPSEPGSAGTSTEPSEPGSAGSGASEPTSTEPGSEPATSGTEPS GSEPATSGTEPSGSEPATSGTEPSGSEPATSGTEPSGTSTEPSEPGSAGSEPATSGT EPSGSGASEPTSTEPGTSTEPSEPGSAGSEPATSGTEPSGSGASEPTSTEPGTSTEP SEPGSAGSGASEPTSTEPGSEPATSGTEPSGSGASEPTSTEPGSEPATSGTEPSGSG ASEPTSTEPGTSTEPSEPGSAGSEPATSGTEPSGSGASEPTSTEPGTSTEPSEPGSA GSEPATSGTEPSGTSTEPSEPGSAGSEPATSGTEPSGTSTEPSEPGSAGTSTEPSEP GSAGTSTEPSEPGSAGTSTEPSEPGSAGTSTEPSEPGSAGTSTEPSEPGSAGTSEPS TSEPGAGSGASEPTSTEPGTSTEPSEPGSAGTSTEPSEPGSAGTSTEPSEPGSAGSE PATSGTEPSGSGASEPTSTEPGSEPATSGTEPSGSEPATSGTEPSGSEPATSGTEPS GSEPATSGTEPSGTSEPSTSEPGAGSEPATSGTEPSGSGASEPTSTEPGTSTEPSEP GSAGSEPATSGTEPSGSGASEPTSTEPGTSTEPSEPGSAG	
aEGRF VHH -AE144-aEGRF_VHH -BD864	EVQLQESGGGLVQAGDSLRLSCLVSGRSFNSYTMGWFRQAPGKEREFVAAILW SGPTTYYADSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVYYCAAALGVLVL APGNVYSYWGQGTQVTVSSAHHGGSEPATSGSETPGTSESATPESGPGSEPATS GSETPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGSEP ATSGSETPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPG EVQLQESGGGLVQAGDSLRLSCLVSGRSFNSYTMGWFRQAPGKEREFVAAILW SGPTTYYADSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVYYCAAALGVLVL APGNVYSYWGQGTQVTVSSAHHGGSETATSGSETAGTSESATSESGAGSTAGS ETSTEAGTSESATSESGAGSETATSGSETAGSETATSGSETAGTSTEASEGSASGT STEASEGSASGTSESATSESGAGSETATSGSETAGTSTEASEGSASGSTAGSETST EAGTSESATSESGAGTSESATSESGAGSETATSGSETAGTSESATSESGAGTSTE ASEGSASGSETATSGSETAGSETATSGSETAGTSTEASEGSASGSTAGSETSTEA GTSESATSESGAGTSTEASEGSASGSETATSGSETAGSTAGSETSTEAGSTAGSE TSTEAGSETATSGSETAGTSESATSESGAGTSESATSESGAGSETATSGSETAGTS ESATSESGAGTSESATSESGAGSETATSGSETAGSETATSGSETAGTSTEASEGS ASGSTAGSETSTEAGSETATSGSETAGTSESATSESGAGSTAGSETSTEAGSTAG SETSTEAGSTAGSETSTEAGTSTEASEGSASGSTAGSETSTEAGSTAGSETSTEAG TSTEASEGSASGSTAGSETSTEAGSETATSGSETAGTSTEASEGSASGTSESATSE SGAGSETATSGSETAGTSESATSESGAGTSESATSESGAGSETATSGSETAGTSES ATSESGAGSETATSGSETAGTSTEASEGSASGTSTEASEGSASGSTAGSETSTEA GSTAGSETSTEAGSETATSGSETAGTSESATSESGAGTSESATSESGAGSETATS GSETAGSETATSGSETAGSETATSGSETAGTSTEASEGSASGTSESATSESGAGS ETATSGSETAGSETATSGSETAGTSESATSESGAGTSESATSESGAGSETATSGSE TAG	848

Nombre*	Secuencia de proteínas	SEQ ID NO:
aEGRF VHH -AF144-aEGRF_VHH -AE576	EVQLQESGGGLVQAGDSLRLSCLVSGRSFNSYTMGWFRQAPGKEREFVAAILW SGPTTYYADSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVYYCAAALGVLVL APGNVYSYWGQGTQVTVSSAHHGGTSTPESGSASPGTSPSGESSTAPGTSPSGE SSTAPGSTSSTAESPGPGSTSESPSGTAPGSTSSTAESPGPGTSPSGESSTAPGTSTP ESGSASPGSTSSTAESPGPGTSPSGESSTAPGTSPSGESSTAPGTSPSGESSTAPGE VQLQESGGGLVQAGDSLRLSCLVSGRSFNSYTMGWFRQAPGKEREFVAAILWS GPTTYYADSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVYYCAAALGVLVLA PGNVYSYWGQGTQVTVSSAHHGGSPAGSPTSTEEGTSESATPESGPGTSTEPSE GSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPA TSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGT STEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESG PGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSE GSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPA TSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGT STEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSA PGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATP ESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAG SPTSTEEGTSESATPESGPGTSTEPSEGSAPG	849
aEGRF_VHH Conector_AE42 aEGRF VHH	EVQLQESGGGLVQAGDSLRLSCLVSGRSFNSYTMGWFRQAPGKEREFVAAILW SGPTTYYADSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVYYCAAALGVLVL APGNVYSYWGQGTQVTVSSAHHGGAPGTSESATPESGPGSEPATSGSETPGTST EPSEGSAPGPAGEVQLQESGGGLVQAGDSLRLSCLVSGRSFNSYTMGWFRQAP GKEREFVAAILWSGPTTYYADSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVY	850
-AM875	YCAAALGVLVLAPGNVYSYWGQGTQVTVSSAHHGGTSTEPSEGSAPGSEPATS GSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAPGSTS ESPSGTAPGTSTPESGSASPGTSTPESGSASPGSEPATSGSETPGTSESATPESGPG SPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGS APGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESAT PESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGSEP ATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPG TSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSPAGSPTST EEGTSTEPSEGSAPGASASGAPSTGGTSESATPESGPGSPAGSPTSTEEGSPAGSP TSTEEGSTSSTAESPGPGTSTESPSGTAPGTSPSGESSTAPGTPGSGTASSSPGSST PSGATGSPGSSPSASTGTGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPG STSSTAESPGPGSTSSTAESPGPGTSPSGESSTAPGSEPATSGSETPGSEPATSGSE TPGTSTEPSEGSAPGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAPGTSTEPSE GSAPGTSTEPSEGSAPGTSTEPSEGSAPGSSTPSGATGSPGSSPSASTGTGPGASP GTSSTGSPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSSTPSGATGSPG SSPSASTGTGPGASPGTSSTGSPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGS APG	

Nombre*	Secuencia de proteínas	SEQ ID NO:
aEGRF_VHH Conector_AM150-aEGRF_VHH -AM1296	EVQLQESGGGLVQAGDSLRLSCLVSGRSFNSYTMGWFRQAPGKEREFVAAILW SGPTTYYADSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVYYCAAALGVLVL APGNVYSYWGQGTQVTVSSAHHGGAPSTGGTSESATPESGPGTSTEPSEGSAPG TSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPES GPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPS EGSAPGPEPTGPAGEVQLQESGGGLVQAGDSLRLSCLVSGRSFNSYTMGWFRQ APGKEREFVAAILWSGPTTYYADSVKGRFTISRDNAKNTVYLQMNSLKPEDTA VYYCAAALGVLVLAPGNVYSYWGQGTQVTVSSAHHGGTSTEPSEGSAPGSEP ATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAPG STSESPSGTAPGTSTPESGSASPGTSTPESGSASPGSEPATSGSETPGTSESATPES GPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPS EGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSE SATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPG SEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATG SPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSPAGSP TSTEEGTSTEPSEGSAPGPEPTGPAPSGGSEPATSGSETPGTSESATPESGPGSPAG SPTSTEEGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGS PAGSPTSTEEGSPAGSPTSTEEGSTSSTAESPGPGSTSESPSGTAPGTSPSGESSTA PGSTSESPSGTAPGSTSESPSGTAPGTSPSGESSTAPGTSTEPSEGSAPGTSESATP ESGPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSESATPESGPGTSTE PSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSPSGESSTAPGTSPSGESSTAPGT SPSGESSTAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGSSPSASTGTG PGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSS TGSPGASASGAPSTGGTSPSGESSTAPGSTSSTAESPGPGTSPSGESSTAPGTSESA TPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSSPSASTGTGPGSSTPSGATGSPGAS PGTSSTGSPGTSTPESGSASPGTSPSGESSTAPGTSPSGESSTAPGTSESATPESGP GSEPATSGSETPGTSTEPSEGSAPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGS ASPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESA TPESGPGSEPATSGSETPGSSTPSGATGSPGASPGTSSTGSPGSSTPSGATGSPGST SESPSGTAPGTSPSGESSTAPGSTSSTAESPGPGSSTPSGATGSPGASPGTSSTGSP GTPGSGTASSSPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPG	851

*** "a" antes del nombre de la proteína diana = anti**

Tabla 41: Proteínas de fusión de unión con restos de direccionamiento dirigidos a dianas diferentes

Nombre*	Secuencia de proteínas	SEQ ID NO:
aHer2-Y288-aEGFR	MEDIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVP SRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKTGSGEGSEGEGGGEG SEGEGSGEGGEGEGSGTEVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGL	852
	EWVARIYPTNGYTRYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFY AMDYWGQGTLVTVSGGEGSGEGSEGEGSEGSGEGEGSEGSGEGEGSEGSGEGEGGSEGSEGEGGSEGSEG EGGSEGSEGEGSGEGSEGEGGSEGSEGEGSGEGSEGEGSEGGSEGEGGSEGSEGEGSGEGSEG EGGEGGSEGEGSEGSGEGEGSGEGSEGEGSEGSGEGEGSGEGSEGEGSEGSGEGEGSEGSGEG EGGSEGSEGEGSEGSGEGEGGEGSGEGEGSGEGSEGEGGGEGSEGEGSGEGGEGEGSEGGSE GEGGSEGGEGEGSEGSGEGEGSEGGSEGEGSEGGSEGEGSEGSGEGEGSEGSGEGEDILLTQS PVILSVSPGERVSFSCRASQSIGTNIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSGTDFT LSINSVESEDIADYYCQQNNNWPTTFGAGTKLELKTGSGEGSEGEGGGEGSEGEGSGEGGE GEGSGTQVQLKQSGPGLVQPSQSLSITCTVSGFSLTNYGVHWVRQSPGKGLEWLGVIWSGG NTDYNTPFTSRLSINKDNSKSQVFFKMNSLQSNDTAIYYCARALTYYDYEFAYWGQGTLVT VS	
aHer2-Y288-aCD3	MEDIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVP SRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKTGSGEGSEGEGGGEG SEGEGSGEGGEGEGSGTEVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGL EWVARIYPTNGYTRYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFY AMDYWGQGTLVTVSGSPGEGSGEGSEGEGSEGSGEGEGSEGSGEGEGGSEGSEGEGGSEGS EGEGGSEGSEGEGSGEGSEGEGGSEGSEGEGSGEGSEGEGSEGGSEGEGGSEGSEGEGSGEGS EGEGGEGGSEGEGSEGSGEGEGSGEGSEGEGSEGSGEGEGSGEGSEGEGSEGSGEGEGSEGSG EGEGGSEGSEGEGSEGSGEGEGGEGSGEGEGSGEGSEGEGGGEGSEGEGSGEGGEGEGSEGG SEGEGGSEGGEGEGSEGSGEGEGSEGGSEGEGSEGGSEGEGSEGSGEGEGSEGSGEGSSSLEG TKDIQMTQSPSSLSASVGDRVTITCSASSSVSYMNWYQQTPGKAPKRWIYDTSKLASGVPS RFSGSGSGTDYTFTISSLQPEDIATYYCQQWSSNPFTFGQGTKLQITRTGSGEGSEGEGGGE GSEGEGSGEGGEGEGSGTQVQLVQSGGGVVQPGRSLRLSCKASGYTFTRYTMHWVRQAPG KGLEWIGYINPSRGYTNYNQKFKDRFTISTDKSKSTAFLQMDSLRPEDTAVYYSARYYDDH YCLDYWGQGTPVTVSSTSG	853

Nombre*	Secuencia de proteínas	SEQ ID NO:
aCD3_VL_ 1-AE48-aCD3_VH-AE144-aHer2_VL-AE48-aHer2_VH-AE864	MKDIQMTQSPSSLSASVGDRVTITCSASSSVSYMNWYQQTPGKAPKRWIYDTSKL ASGVPSRFSGSGSGTDYTFTISSLQPEDIATYYCQQWSSNPFTFGQGTKLQITRGM AEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGQVQLVQSG GGVVQPGRSLRLSCKASGYTFTRYTMHWVRQAPGKGLEWIGYINPSRGYTNYNQ KFKDRFTISTDKSKSTAFLQMDSLRPEDTAVYYSARYYDDHYCLDYWGQGTPVT VSSGGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEP SEGSAPGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTSE SATPESGPGSEPATSGSETPGTSTEPSEGSAPGMEDIQMTQSPSSLSASVGDRVTIT CRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVPSRFSGSRSGTDFTLTISSL QPEDFATYYCQQHYTTPPTFGQGTKVEIKGMAEPAGSPTSTEEGTPGSGTASSSPG SSTPSGATGSPGASPGTSSTGSPGEVQLVESGGGLVQPGGSLRLSCAASGFNIKDT YIHWVRQAPGKGLEWVARIYPTNGYTRYADSVKGRFTISADTSKNTAYLQMNSL RAEDTAVYYCSRWGGDGFYAMDYWGQGTLVTVSGGSPAGSPTSTEEGTSESATP ESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESA TPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTE PSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTS ESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPG TSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGP GSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSE GSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESA TPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPA GSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTS ESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEG TSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEE GTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSE TPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSG SETPGTSESATPESGPGTSTEPSEGSAPG	854
aCD3_VH-AE48-aCD3_VL-AE144-aHer2_VH-AE48-aHer2_VL-AE864	QVQLVQSGGGVVQPGRSLRLSCKASGYTFTRYTMHWVRQAPGKGLEWIGYINPS RGYTNYNQKFKDRFTISTDKSKSTAFLQMDSLRPEDTAVYYSARYYDDHYCLDY WGQGTPVTVSSGMAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTS STGSPGMKDIQMTQSPSSLSASVGDRVTITCSASSSVSYMNWYQQTPGKAPKRWI	855
Nombre*	**Secuencia de proteínas NO:**	**SEQ ID NO:**

Nombre*	Secuencia de proteínas	SEQ ID NO:
	YDTSKLASGVPSRFSGSGSGTDYTFTISSLQPEDIATYYCQQWSSNPFTFGQGTKL QITRGGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEP SEGSAPGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTSE SATPESGPGSEPATSGSETPGTSTEPSEGSAPGEVQLVESGGGLVQPGGSLRLSCAA SGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYTRYADSVKGRFTISADTSKNTA YLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGTLVTVSGMAEPAGSPTST EEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGMEDIQMTQSPSSLSASVG DRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVPSRFSGSRSGTDF TLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKGGSPAGSPTSTEEGTSESATP ESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESA TPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTE PSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTS ESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPG TSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGP GSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSE GSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESA TPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPA GSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTS ESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEG TSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEE GTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSE TPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSG SETPGTSESATPESGPGTSTEPSEGSAPG	

Nombre*	Secuencia de proteínas	SEQ ID NO:
aHer2_VL-AM48-aHer2_VH-AF144-aCD3_VL-AM48-aCD3_VH-AM1296	MEDIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASF LYSGVPSRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKGM AEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGEVQLVESG GGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYTRYAD SVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGTL VTVSGGTSTPESGSASPGTSPSGESSTAPGTSPSGESSTAPGSTSSTAESPGPGSTSES PSGTAPGSTSSTAESPGPGTSPSGESSTAPGTSTPESGSASPGSTSSTAESPGPGTSPS GESSTAPGTSPSGESSTAPGTSPSGESSTAPGMKDIQMTQSPSSLSASVGDRVTITC SASSSVSYMNWYQQTPGKAPKRWIYDTSKLASGVPSRFSGSGSGTDYTFTISSLQP EDIATYYCQQWSSNPFTFGQGTKLQITRGMAEPAGSPTSTEEGASPGTSSTGSPGS STPSGATGSPGSSTPSGATGSPGQVQLVQSGGGVVQPGRSLRLSCKASGYTFTRYT MHWVRQAPGKGLEWIGYINPSRGYTNYNQKFKDRFTISTDKSKSTAFLQMDSLR PEDTAVYYSARYYDDHYCLDYWGQGTPVTVSSGGTSTEPSEGSAPGSEPATSGSE TPGSPAGSPTSTEEGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAPGSTSESPSG TAPGTSTPESGSASPGTSTPESGSASPGSEPATSGSETPGTSESATPESGPGSPAGSP TSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAG SPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTST EPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGSEPATSGSETPGSP AGSPTSTEEGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTSTEPSEGSAPG TSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAP GPEPTGPAPSGGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGTSESATPESG PGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSPAGSPTSTEEGSPAGSPTST EEGSTSSTAESPGPGSTSESPSGTAPGTSPSGESSTAPGSTSESPSGTAPGSTSESPSG TAPGTSPSGESSTAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSEPATS GSETPGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGTSTEP SEGSAPGTSPSGESSTAPGTSPSGESSTAPGTSPSGESSTAPGTSTEPSEGSAPGSPA GSPTSTEEGTSTEPSEGSAPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGSPGS STPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGASASGAPSTGGTSPSGESSTAPG STSSTAESPGPGTSPSGESSTAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAP GSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGTSTPESGSASPGTSPSGESST APGTSPSGESSTAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGSTSESPSG TAPGSTSESPSGTAPGTSTPESGSASPGSPAGSPTSTEEGTSESATPESGPGTSTEPSE GSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGSSTPSGATGSPGASPGT SSTGSPGSSTPSGATGSPGSTSESPSGTAPGTSPSGESSTAPGSTSSTAESPGPGSSTP	856
	SGATGSPGASPGTSSTGSPGTPGSGTASSSPGSPAGSPTSTEEGSPAGSPTSTEEGTS TEPSEGSAPG	

Nombre*	Secuencia de proteínas	SEQ ID NO:
aHer2_VH AM48-aHer2_VL AF144-aCD3_VH AM48-aCD3_VL-AM875 - -	EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPT NGYTRYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAM DYWGQGTLVTVSGMAEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPS GATGSPGMEDIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPK LLIYSASFLYSGVPSRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGT KVEIKGGTSTPESGSASPGTSPSGESSTAPGTSPSGESSTAPGSTSSTAESPGPGSTSE SPSGTAPGSTSSTAESPGPGTSPSGESSTAPGTSTPESGSASPGSTSSTAESPGPGTSP SGESSTAPGTSPSGESSTAPGTSPSGESSTAPGQVQLVQSGGGVVQPGRSLRLSCK ASGYTFTRYTMHWVRQAPGKGLEWIGYINPSRGYTNYNQKFKDRFTISTDKSKST AFLQMDSLRPEDTAVYYSARYYDDHYCLDYWGQGTPVTVSSGMAEPAGSPTST EEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGMKDIQMTQSPSSLSASVG DRVTITCSASSSVSYMNWYQQTPGKAPKRWIYDTSKLASGVPSRFSGSGSGTDYT FTISSLQPEDIATYYCQQWSSNPFTFGQGTKLQITRGGTSTEPSEGSAPGSEPATSGS ETPGSPAGSPTSTEEGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAPGSTSESPS GTAPGTSTPESGSASPGTSTPESGSASPGSEPATSGSETPGTSESATPESGPGSPAGS PTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPA GSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTS TEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGSEPATSGSETPG SPAGSPTSTEEGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTSTEPSEGSA PGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGS APGASASGAPSTGGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSTSSTAES PGPGSTSESPSGTAPGTSPSGESSTAPGTPGSGTASSSPGSSTPSGATGSPGSSPSAST GTGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSTSSTAESPGPGSTSST AESPGPGTSPSGESSTAPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGSTSS TAESPGPGTSTPESGSASPGSTSESPSGTAPGTSTEPSEGSAPGTSTEPSEGSAPGTST EPSEGSAPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSEPATSGSETPGT SESATPESGPGSPAGSPTSTEEGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSP GTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPG	857
aHer2_VL--Conector_AE 42-aHer2_VH-AE288-aEGFR V L--Conector_A M150-aEGFR V H-AD576 -	MEDIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASF LYSGVPSRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKGGA PGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGPAGEVQLVESGGGLVQPGG SLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYTRYADSVKGRFTIS ADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGTLVTVSGGTS ESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPG TSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGP GSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPES GPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSE GSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGMEDIL LTQSPVILSVSPGERVSFSCRASQSIGTNIHWYQQRTNGSPRLLIKYASESISGIPSRF SGSGSGTDFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLELKGGAPSTGGTSE SATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGT STEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAP GSPAGSPTSTEEGTSTEPSEGSAPGPEPTGPAGQVQLKQSGPGLVQPSQSLSITCTV SGFSLTNYGVHWVRQSPGKGLEWLGVIWSGGNTDYNTPFTSRLSINKDNSKSQV FFKMNSLQSNDTAIYYCARALTYYDYEFAYWGQGTLVTVSGGSSESGSSEGGPG SGGEPSESGSSGSSESGSSEGGPGSSESGSSEGGPGSSESGSSEGGPGSSESGSSEGG PGSSESGSSEGGPGESPGGSSGSESGSEGSSGPGESSGSSESGSSEGGPGSSESGSSE GGPGSSESGSSEGGPGSGGEPSESGSSGESPGGSSGSESGESPGGSSGSESGSGGEPS ESGSSGSSESGSSEGGPGSGGEPSESGSSGSGGEPSESGSSGSEGSSGPGESSGESPG GSSGSESGSGGEPSESGSSGSGGEPSESGSSGSGGEPSESGSSGSSESGSSEGGPGES PGGSSGSESGESPGGSSGSESGESPGGSSGSESGESPGGSSGSESGESPGGSSGSESG SSESGSSEGGPGSGGEPSESGSSGSEGSSGPGESSGSSESGSSEGGPGSGGEPSESGS SGSSESGSSEGGPGSGGEPSESGSSGESPGGSSGSESGESPGGSSGSESGSSESGSSE GGPGSGGEPSESGSSGSSESGSSEGGPGSGGEPSESGSSGSGGEPSESGSSGESPGGS SGSESGSEGSSGPGESSGSSESGSSEGGPGSEGSSGPGESSG	858

Nombre*	Secuencia de proteínas	SEQ ID NO:
aHer2_VH--Conector_AM150-aHer2_VL-AE288-aEGFR V H--Conector_AE42-aEGFR_V L-AE576	EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPT	859
	NGYTRYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAM DYWGQGTLVTVSGGAPSTGGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPG TSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAP GTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGPEPTGPAGM EDIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFL YSGVPSRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKGGTS ESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPG TSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGP GSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPES GPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSE GSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGQVQL KQSGPGLVQPSQSLSITCTVSGFSLTNYGVHWVRQSPGKGLEWLGVIWSGGNTD YNTPFTSRLSINKDNSKSQVFFKMNSLQSNDTAIYYCARALTYYDYEFAYWGQG TLVTVSGGAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGPAGMEDILLT QSPVILSVSPGERVSFSCRASQSIGTNIHWYQQRTNGSPRLLIKYASESISGIPSRFSG SGSGTDFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLELKGGSPAGSPTSTEE GTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGS APGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATP ESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS EGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTE PSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTS ESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEE GTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSE TPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTS TEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG	

Nombre*	Secuencia de proteínas	SEQ ID NO:
AE624-aHer2_VL--Conector_AE 42-aHer2_VH-AE288-aEGFR V L--Conector_A M150-aEGFR_V H	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGMEDIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVPSRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKGGAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGPAGEVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYTRYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGTLVTVSGGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGMEDILLTQSPVILSVSPGERVSFSCRASQSIGTNIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSGTDFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLELKGGAPSTGGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGPEPTGPAGQVQLKQSGPGLVQPSQSLSITCTVSGFSLTNYGVHWVRQSPGKGLEWLGVIWSGGNTDYNTPFTSRLSINKDNSKSQVFFKMNSLQSNDTAIYYCARALTYYDYEFAYWGQGTLVTVSG	860
AE912-aHer2_VH--Conector_A M150-aHer2_VL-AE288-aEGFR V H--Conector_AE 42-aEGFR_V L	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTE	861

Nombre*	Secuencia de proteínas	SEQ ID NO:
	PSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTS TEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEG TSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAP GTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSG SETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSP TSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPA TSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSP AGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEG SPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGP GSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGS APGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGEVQLVESGGGLVQPGGSL RLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYTRYADSVKGRFTISAD TSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGTLVTVSGGAPST GGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEG SAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSE GSAPGSPAGSPTSTEEGTSTEPSEGSAPGPEPTGPAGMEDIQMTQSPSSLSASVGDR VTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVPSRFSGSRSGTDFTLT ISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKGGTSESATPESGPGSEPATSGSET PGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPT STEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATS GSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPA TSGSETPGTSESATPESGPGTSTEPSEGSAPGQVQLKQSGPGLVQPSQSLSITCTVS GFSLTNYGVHWVRQSPGKGLEWLGVIWSGGNTDYNTPFTSRLSINKDNSKSQVF FKMNSLQSNDTAIYYCARALTYYDYEFAYWGQGTLVTVSGGAPGTSESATPESG PGSEPATSGSETPGTSTEPSEGSAPGPAGMEDILLTQSPVILSVSPGERVSFSCRASQ SIGTNIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSGTDFTLSINSVESEDIADY YCQQNNNWPTTFGAGTKLELKG	

Nombre*	Secuencia de proteínas	SEQ ID NO:
AM923-aEGFR V L--Conector_A M150-aEGFR_V H-AF144-aHer2_VH--Conector_A M150-aHer2_VL	MAEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGTSTEPSEG SAPGSEPATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTPESGSASPGSTSESPS GTAPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPGSEPATSGSETPGTSESA TPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTE PSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTS ESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPG SEPATSGSETPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGS PGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSPAGSPTST EEGTSTEPSEGSAPGASASGAPSTGGTSESATPESGPGSPAGSPTSTEEGSPAGSPTS TEEGSTSSTAESPGPGTSTESPSGTAPGTSPSGESSTAPGTPGSGTASSSPGSSTPSGA TGSPGSSPSASTGTGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSTSST AESPGPGSTSSTAESPGPGTSPSGESSTAPGSEPATSGSETPGSEPATSGSETPGTSTE PSEGSAPGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAPGTSTEPSEGSAPGTST EPSEGSAPGTSTEPSEGSAPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGS EPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSSTPSGATGSPGSSPSASTGTGP GASPGTSSTGSPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGMEDILLTQS PVILSVSPGERVSFSCRASQSIGTNIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSG SGTDFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLELKGGAPSTGGTSESATP ESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPS EGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAG SPTSTEEGTSTEPSEGSAPGPEPTGPAGQVQLKQSGPGLVQPSQSLSITCTVSGFSLT NYGVHWVRQSPGKGLEWLGVIWSGGNTDYNTPFTSRLSINKDNSKSQVFFKMNS LQSNDTAIYYCARALTYYDYEFAYWGQGTLVTVSGGTSTPESGSASPGTSPSGES STAPGTSPSGESSTAPGSTSSTAESPGPGSTSESPSGTAPGSTSSTAESPGPGTSPSGE SSTAPGTSTPESGSASPGSTSSTAESPGPGTSPSGESSTAPGTSPSGESSTAPGTSPSG ESSTAPGEVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEW VARIYPTNGYTRYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGG DGFYAMDYWGQGTLVTVSGGAPSTGGTSESATPESGPGTSTEPSEGSAPGTSTEP SEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTST	862
	EPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGP EPTGPAGMEDIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPK LLIYSASFLYSGVPSRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGT KVEIKG	

* "a" antes del nombre del resto proteico diana = anti

REIVINDICACIONES

1. Una proteína de fusión de unión aislada que comprende un primer resto de direccionamiento con afinidad de unión específica a una primera diana y un primer polipéptido recombinante expandido (XTEN), en donde el XTEN comprende una secuencia de aminoácidos que tiene al menos un 90 % de identidad de secuencia con una secuencia seleccionada del grupo que consiste en la SEQ ID NO: 41, la SEQ ID NO: 47, la SEQ ID NO: 54 y la SEQ ID NO: 59, en donde el XTEN comprende un motivo seleccionado del grupo que consiste en la SEQ ID NO: 6, la SEQ ID NO: 7, la SEQ ID NO: 8 y la SEQ ID NO: 9, en donde el primer resto de direccionamiento comprende un fragmento variable monocatenario (scFv) que se une a CD3.

2. La proteína de fusión de unión aislada de la reivindicación 1, en donde la proteína de fusión está configurada de acuerdo con la fórmula I:

$$(XTEN)_x\text{-}TM\text{-}(XTEN)_y \qquad I$$

en donde x es 0 o 1; y es 0 o 1, en donde x+y ≥1, y en donde TM es el resto de direccionamiento entre dos XTEN.

3. La proteína de fusión de unión aislada de la reivindicación 1, que comprende además un segundo resto de direccionamiento con unión específica a una segunda diana diferente de dicho primera diana, en donde el segundo resto de direccionamiento comprende un segundo scFv, en donde la segunda diana se selecciona de la Tabla 1.

4. La proteína de fusión de unión aislada de la reivindicación 3, en donde la primera diana es CD3 y la segunda diana es HER2.

5. La proteína de fusión de unión aislada de una cualquiera de las reivindicaciones 1, 3 o 4 que comprende además una secuencia de escisión.

6. La proteína de fusión de unión aislada de una cualquiera de las reivindicaciones 1, 3 o 4 que comprende además una o más moléculas de un fármaco seleccionado de la Tabla 9.

7. La proteína de fusión de unión aislada de una cualquiera de las reivindicaciones 1, 3, 4 o 6 para su uso en un método para tratar una enfermedad, un trastorno o una afección en un sujeto, en donde la enfermedad, el trastorno o la afección se selecciona del grupo que consiste en carcinoma de mama, carcinoma de pulmón, carcinoma gástrico, carcinoma de esófago, carcinoma colorrectal, carcinoma de hígado, carcinoma de ovario, tecoma, arrenoblastoma, carcinoma cervicouterino, carcinoma endometrial, endometriosis, fibrosarcoma, coriocarcinoma, cáncer de cabeza y cuello, carcinoma nasofaríngeo, carcinoma laríngeo, hepatoblastoma, sarcoma de Kaposi, melanoma, carcinoma cutáneo, hemangioma, hemangioma cavernoso, hemangioblastoma, carcinoma de páncreas, retinoblastoma, astrocitoma, glioblastoma, schwannoma, oligodendroglioma, meduloblastoma, neuroblastoma, rabdomiosarcoma, sarcoma osteógeno, leiomiosarcomas, carcinoma de las vías urinarias, carcinoma de tiroides, tumor de Wilm, carcinoma de células renales y carcinoma de próstata.

8. La proteína de fusión de unión aislada para su uso de acuerdo con la reivindicación 7, en donde el método comprende administrar una dosis terapéuticamente eficaz de la proteína de fusión de unión a un sujeto que lo necesite.

9. Una composición farmacéutica que comprende la proteína de unión de una cualquiera de las reivindicaciones 1, 3, 4 o 6 y al menos un portador farmacéuticamente aceptable.

10. Un ácido nucleico aislado que comprende una secuencia polinucleotídica seleccionada de (a) un polinucleótido que codifica la proteína de fusión de una cualquiera de las reivindicaciones 1, 3, 4 o 5, o (b) el complemento del polinucleótido de (a).

11. Un vector de expresión que comprende el ácido nucleico aislado de la reivindicación 10.

12. Una célula hospedadora que comprende el vector de expresión de la reivindicación 11.

13. La proteína de fusión de unión aislada de la reivindicación 3, en donde el segundo resto de direccionamiento se une a la diana con una constante de disociación menor de 10^{-7} M.

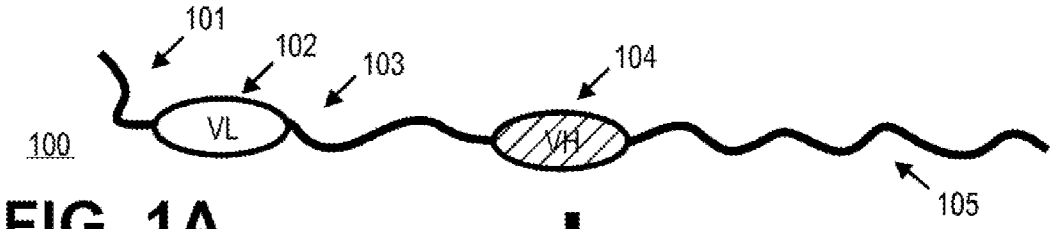


FIG. 1A

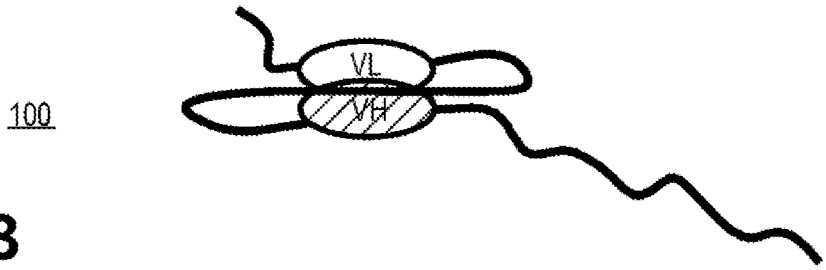


FIG. 1B

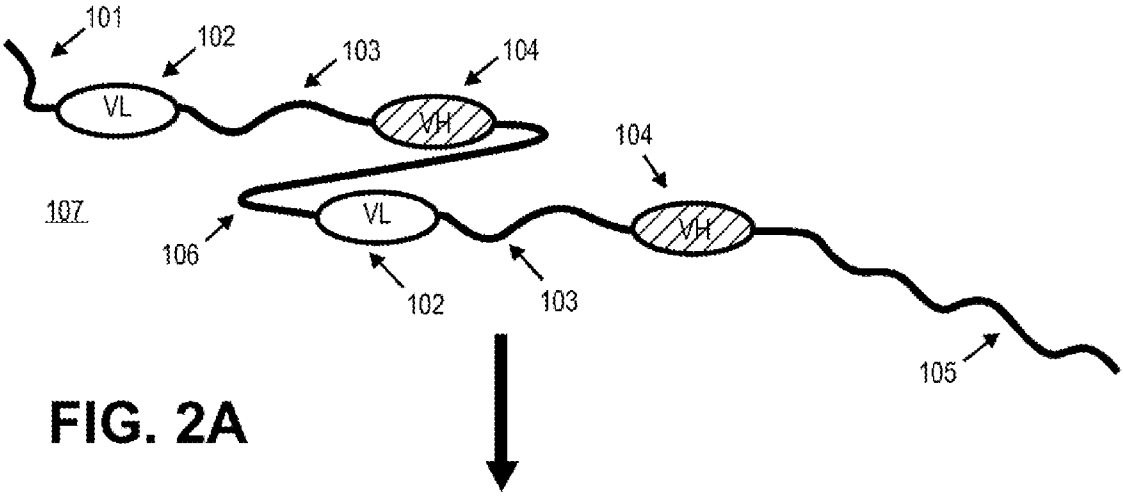


FIG. 2A

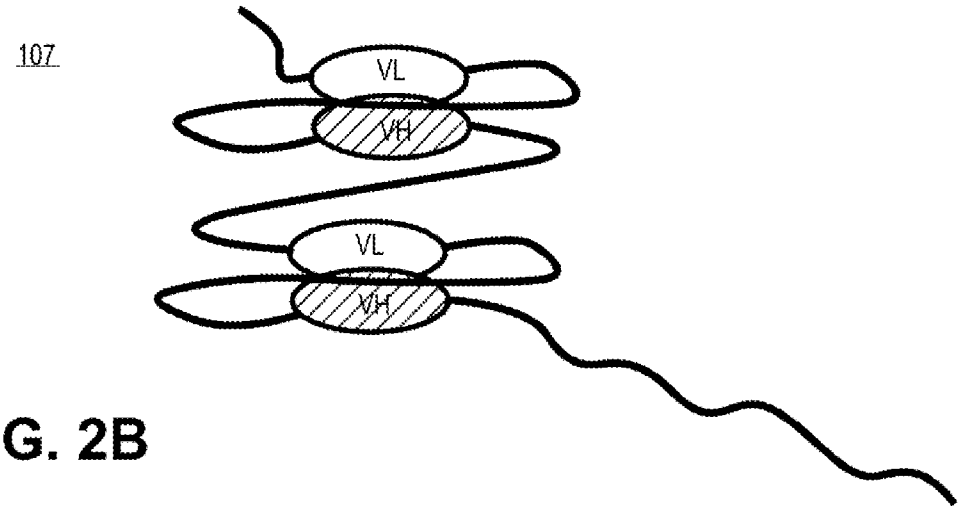


FIG. 2B

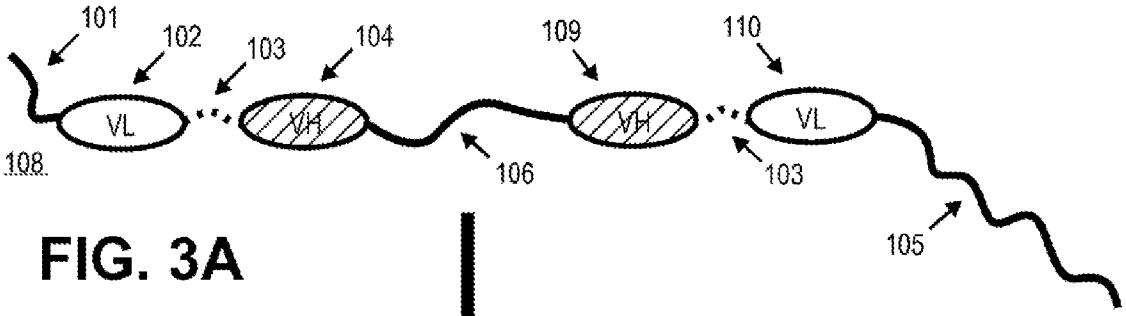

101
102
103
104
109
110
VL
VH
VH
VL
108
106
103
105

FIG. 3A

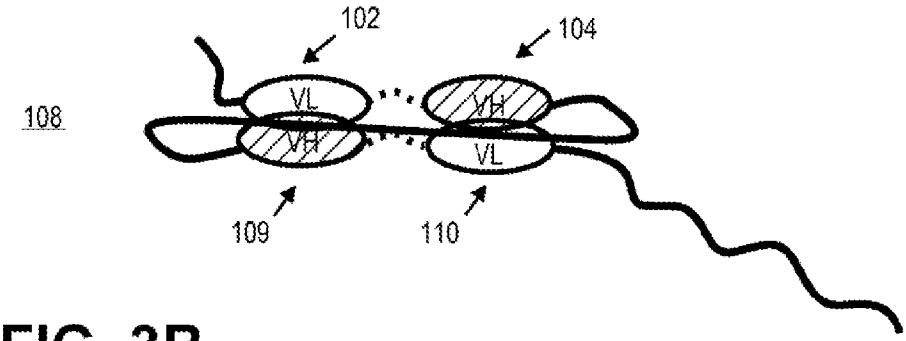

102
104
108
VL
VH
VH
VL
109
110

FIG. 3B

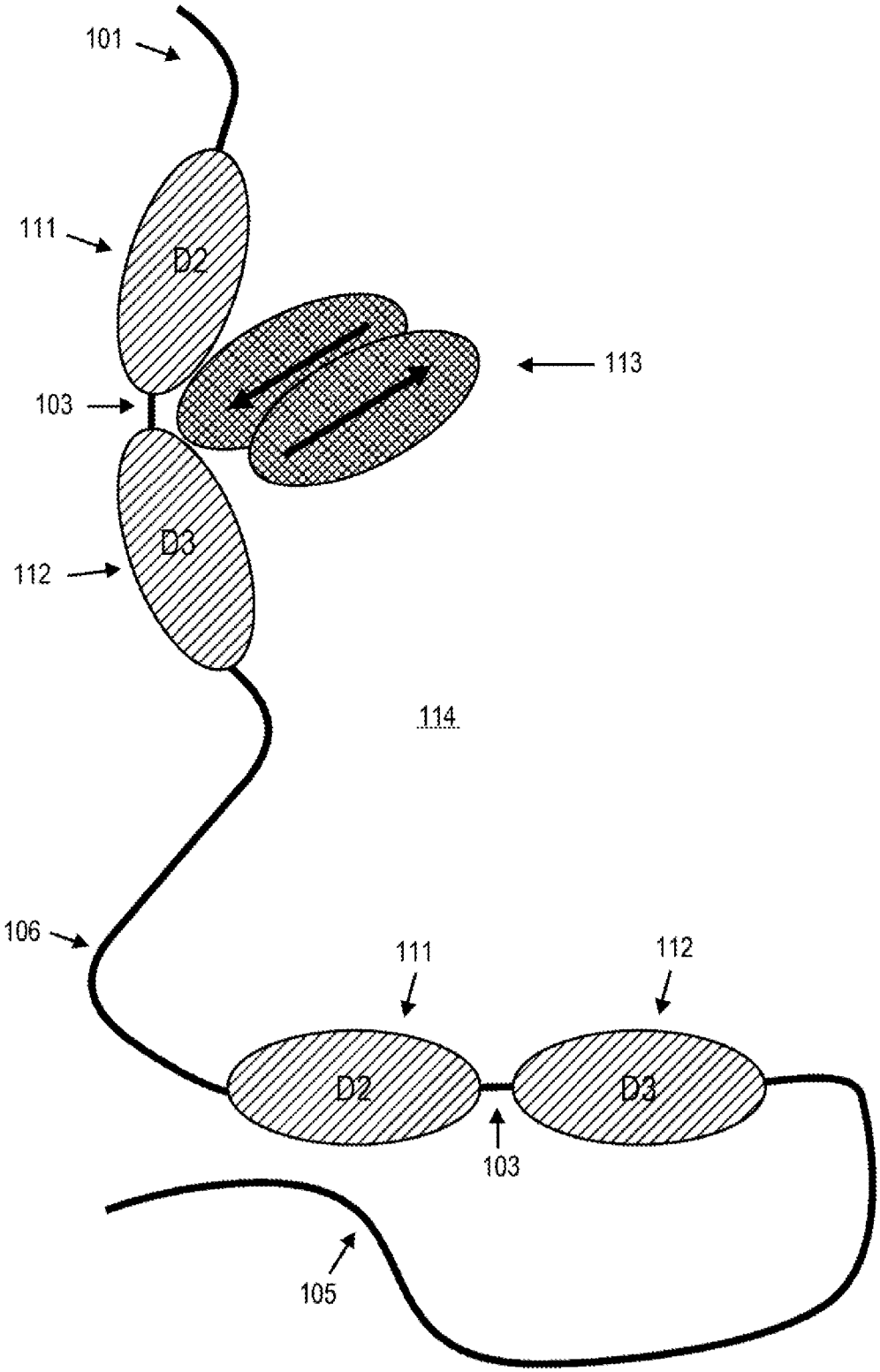


FIG. 4A

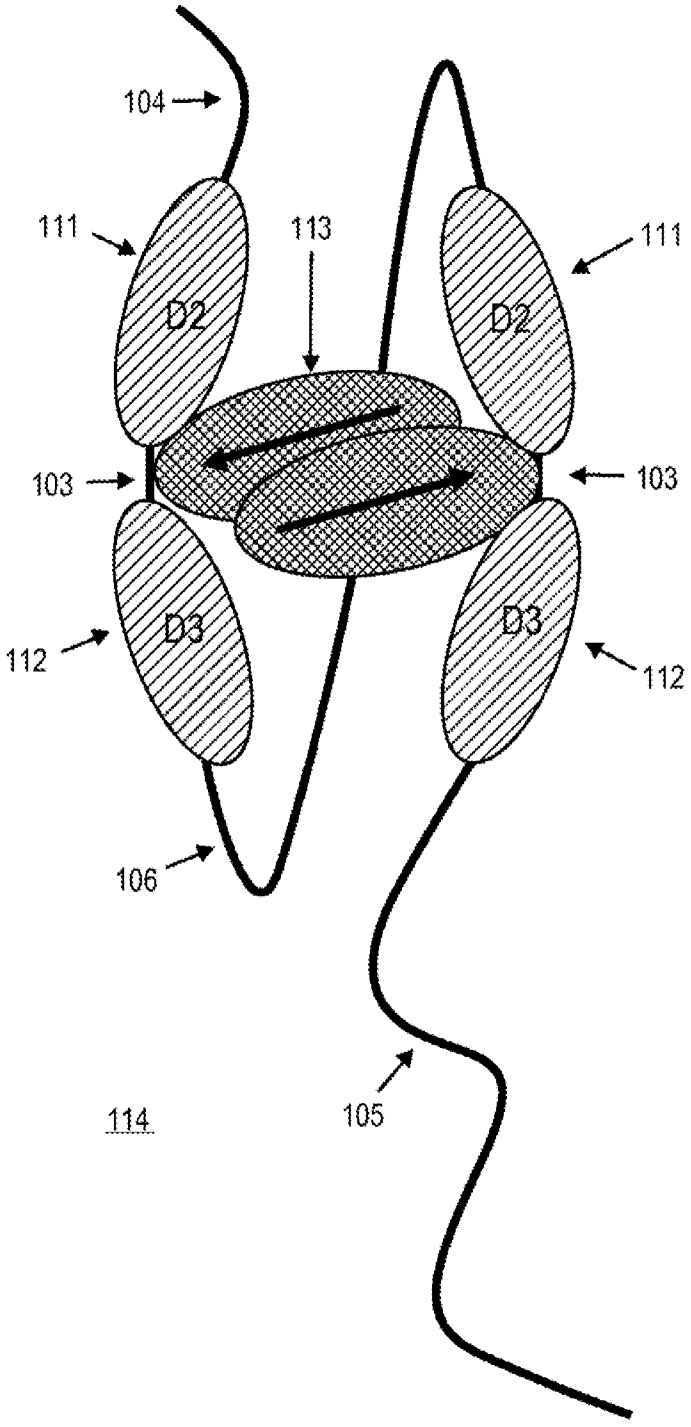


FIG. 4B

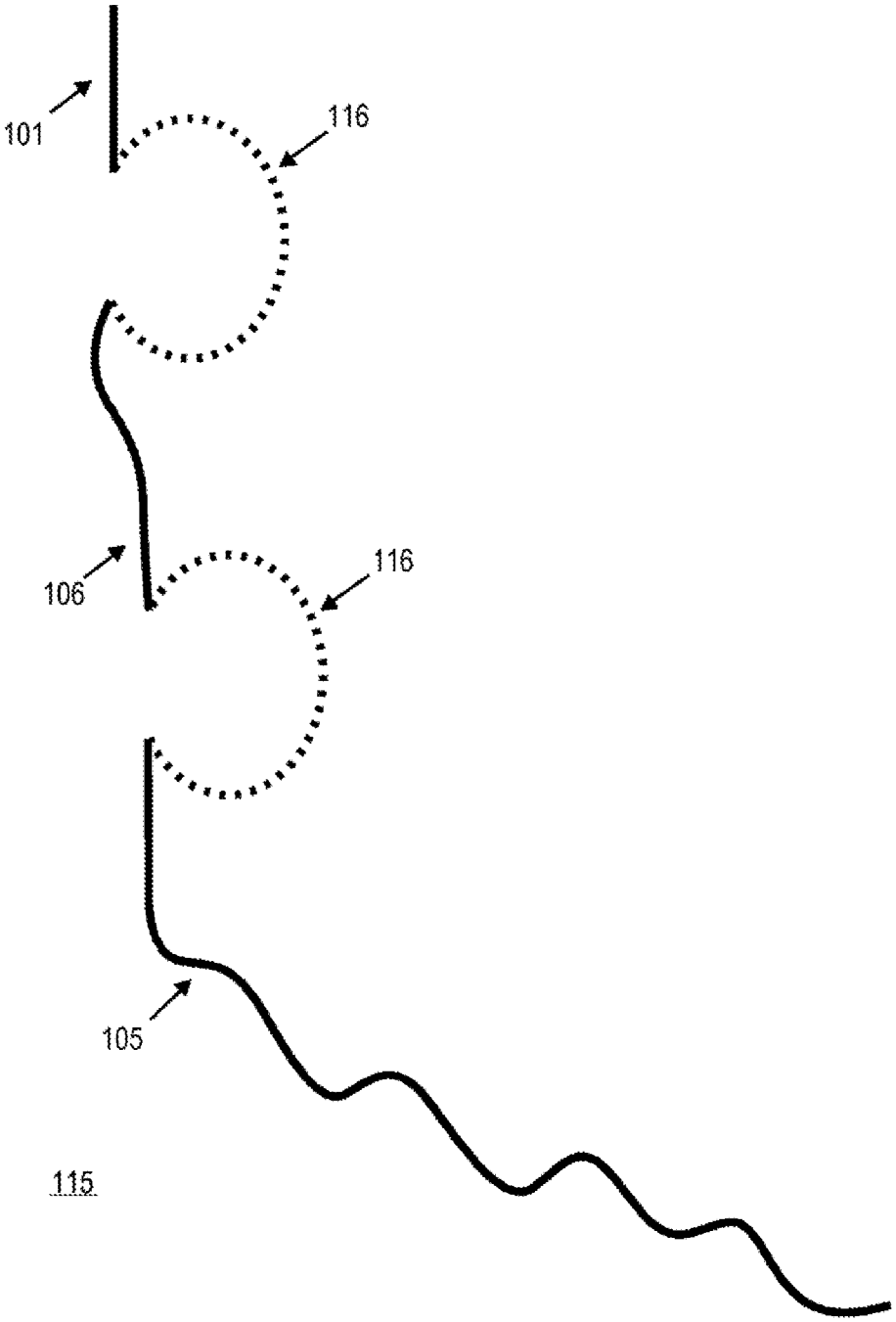


FIG. 5

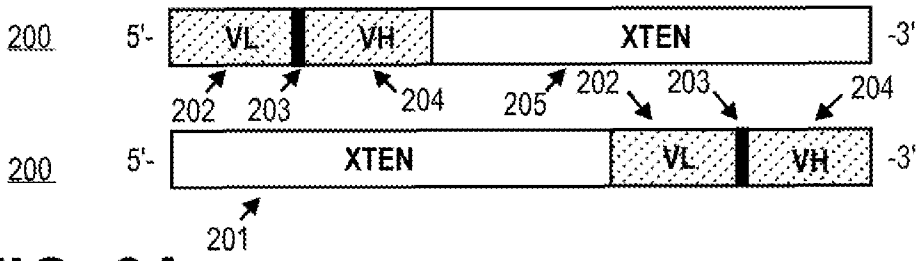


FIG. 6A

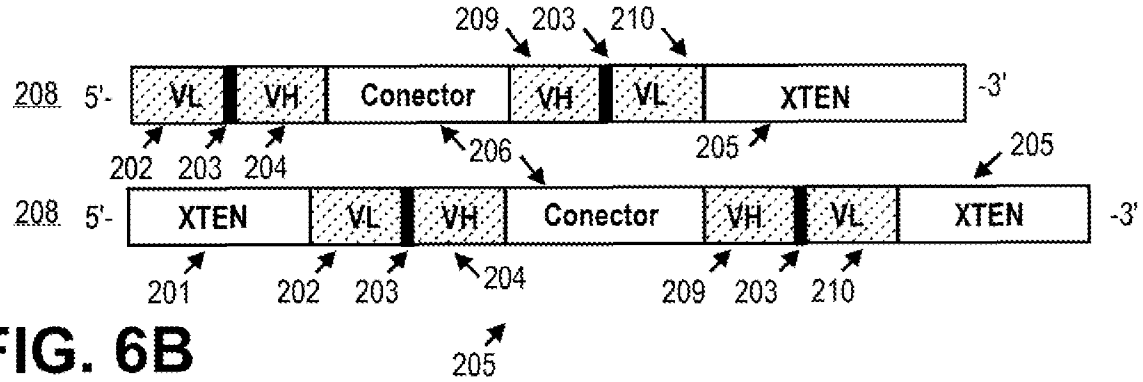


FIG. 6B

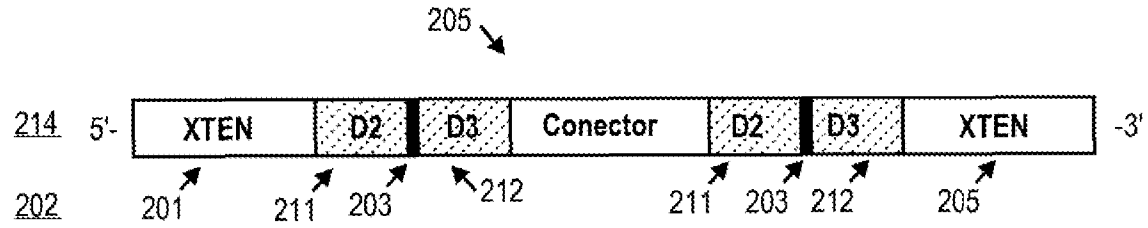


FIG. 6C

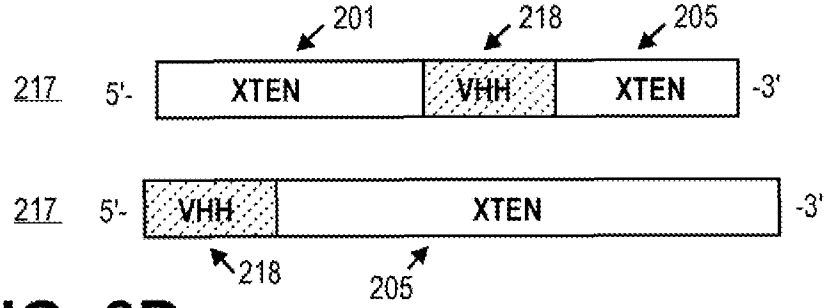


FIG. 6D

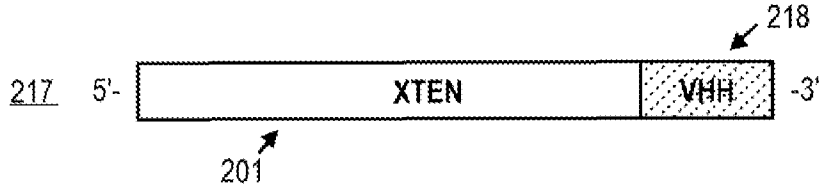


FIG. 6F

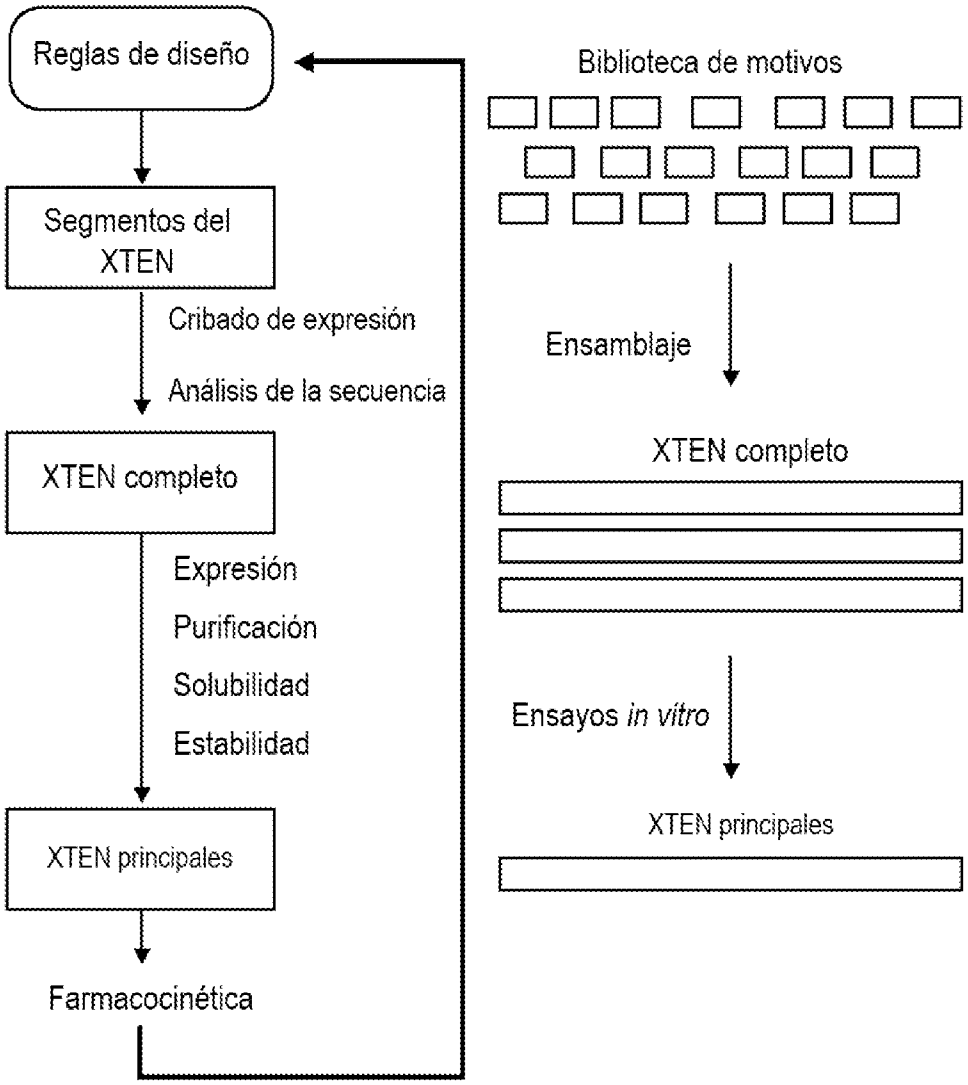

Reglas de diseño
Segmentos del XTEN
Cribado de expresión
Análisis de la secuencia
XTEN completo
Expresión
Purificación
Solubilidad
Estabilidad
XTEN principales
Farmacocinética
Biblioteca de motivos
Ensamblaje
XTEN completo
Ensayos *in vitro*
XTEN principales

FIG. 7

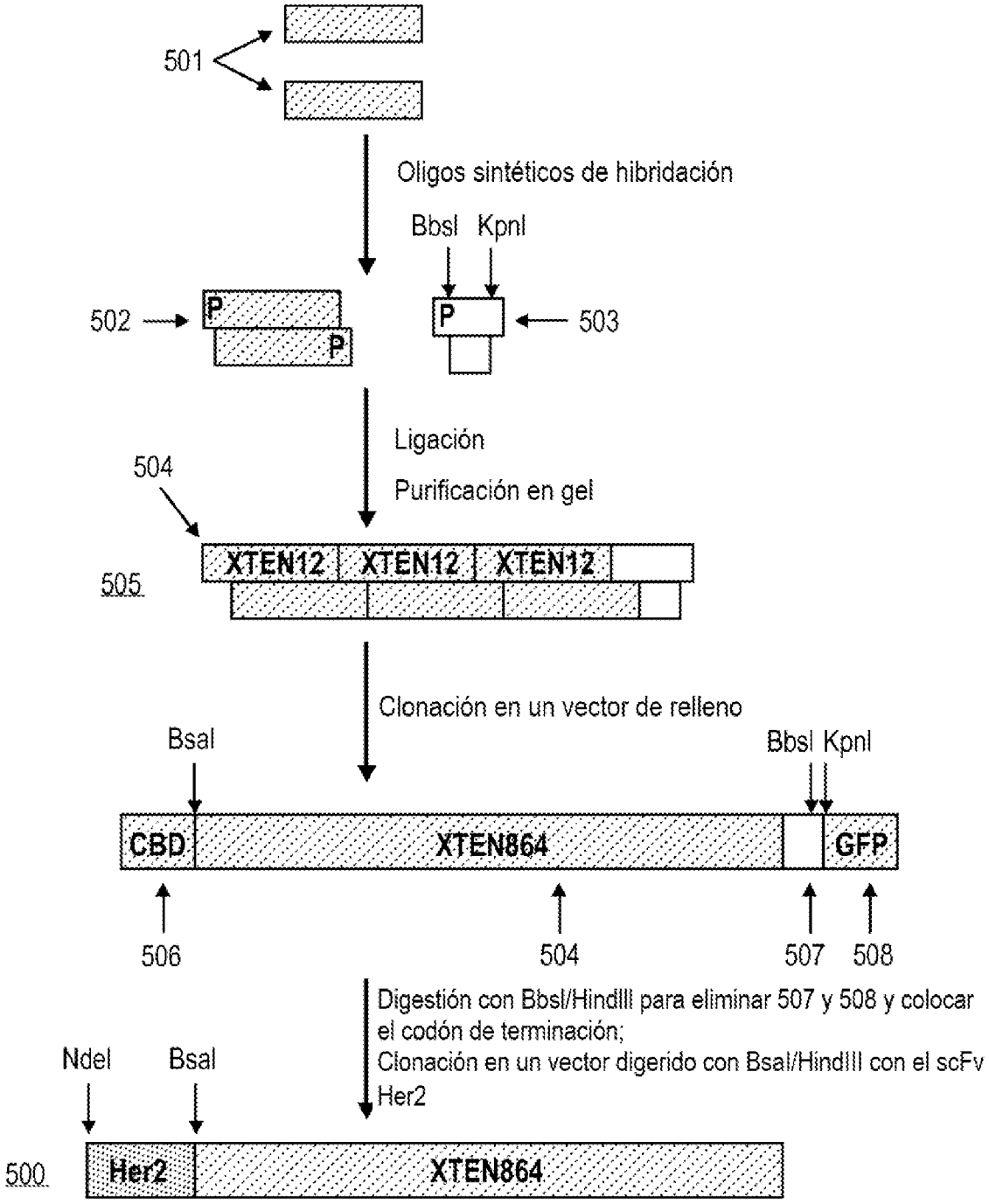

501
Oligos sintéticos de hibridación
BbsI KpnI
502
P
P
P
503
Ligación
Purificación en gel
504
505
XTEN12
XTEN12
XTEN12
Clonación en un vector de relleno
BsaI
BbsI KpnI
CBD
XTEN864
GFP
506
504
507
508
Digestión con BbsI/HindIII para eliminar 507 y 508 y colocar el codón de terminación;
Clonación en un vector digerido con BsaI/HindIII con el scFv Her2
NdeI
BsaI
500
Her2
XTEN864

FIG. 8

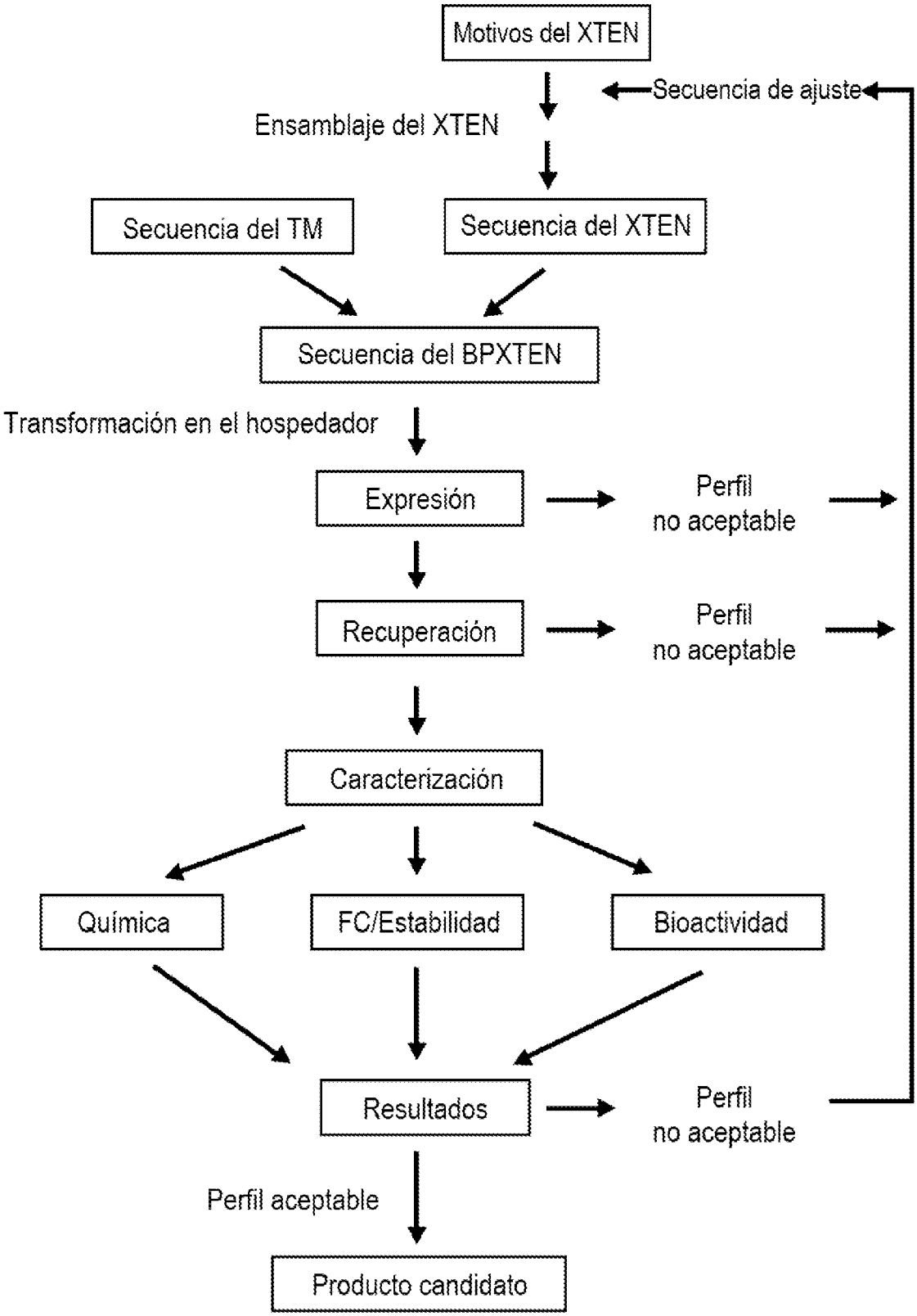

Motivos del XTEN
Secuencia de ajuste
Ensamblaje del XTEN
Secuencia del TM
Secuencia del XTEN
Secuencia del BPXTEN
Transformación en el hospedador
Expresión
Perfil no aceptable
Recuperación
Perfil no aceptable
Caracterización
Química
FC/Estabilidad
Bioactividad
Resultados
Perfil no aceptable
Perfil aceptable
Producto candidato

FIG. 9

FIG. 10A

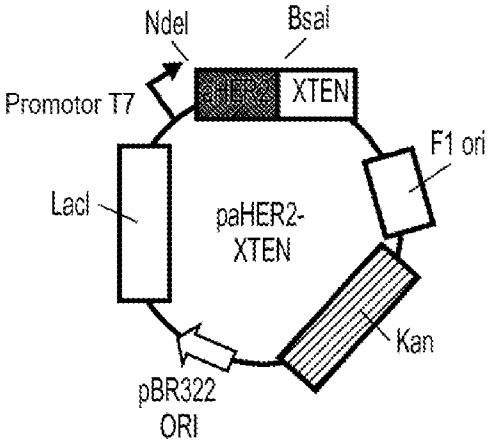


FIG. 10B

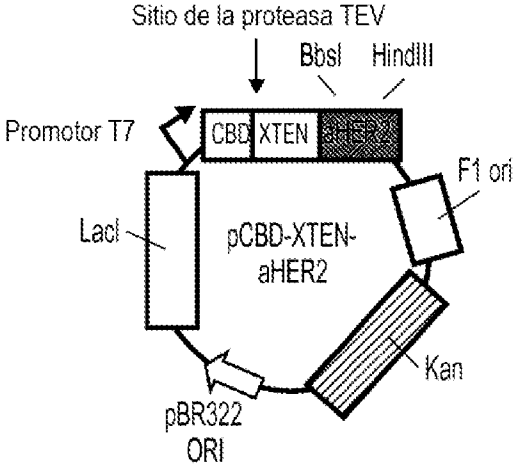


FIG. 10C

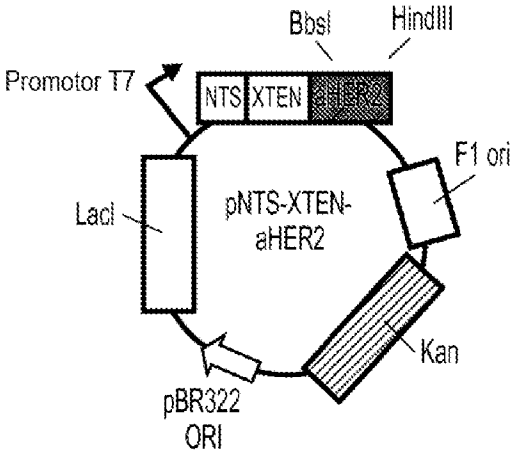
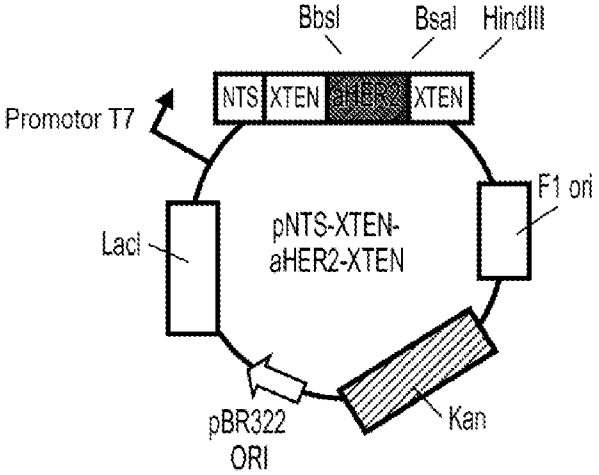


FIG. 10D



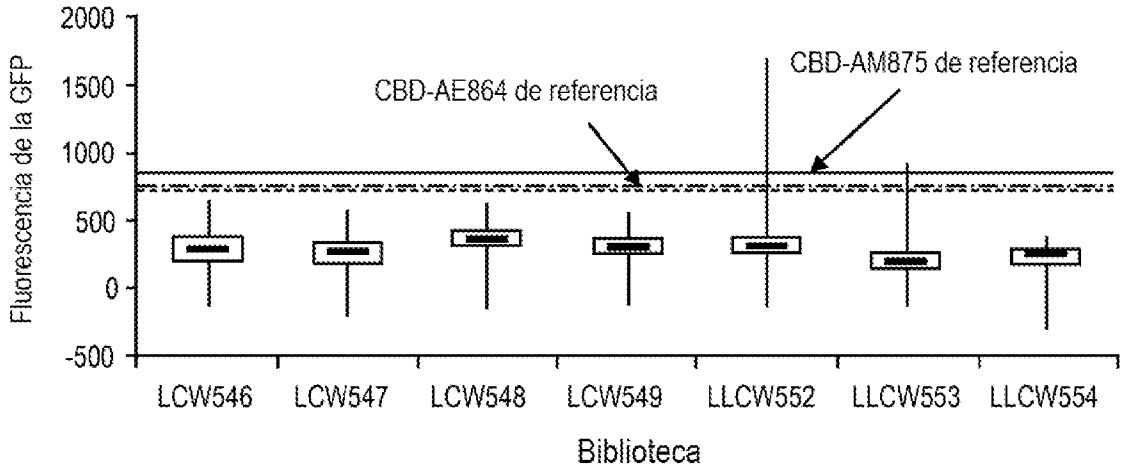
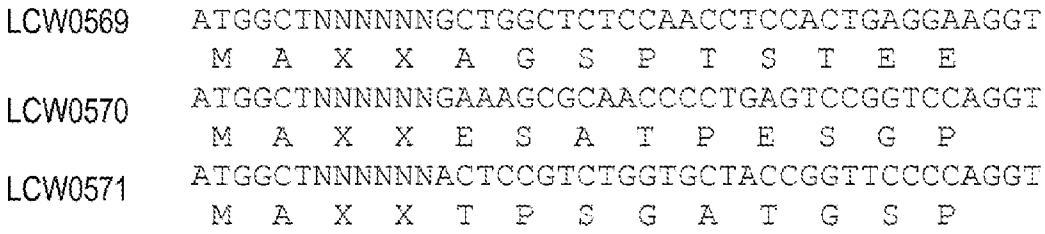

2000
1500
1000
500
0
-500
Fluorescencia de la GFP
CBD-AE864 de referencia
CBD-AM875 de referencia
LCW546
LCW547
LCW548
LCW549
LLCW552
LLCW553
LLCW554
Biblioteca

FIG. 11



X = APST,	GS	o	GE
TCAG/C/TCAG,	AG/G/TC	o	G/AG/AG
Diversidad:	16	4	4

- Las bibliotecas del lote 2 se basan en los 3 mejores clones del cribado del lote 1.
- Se incluyen los 24 codones para los 6 aminoácidos G, E, S, P, A, T.
- Cada nueva biblioteca se compone de 3 x 3 = 9 pares de oligos hibridados.

FIG. 12

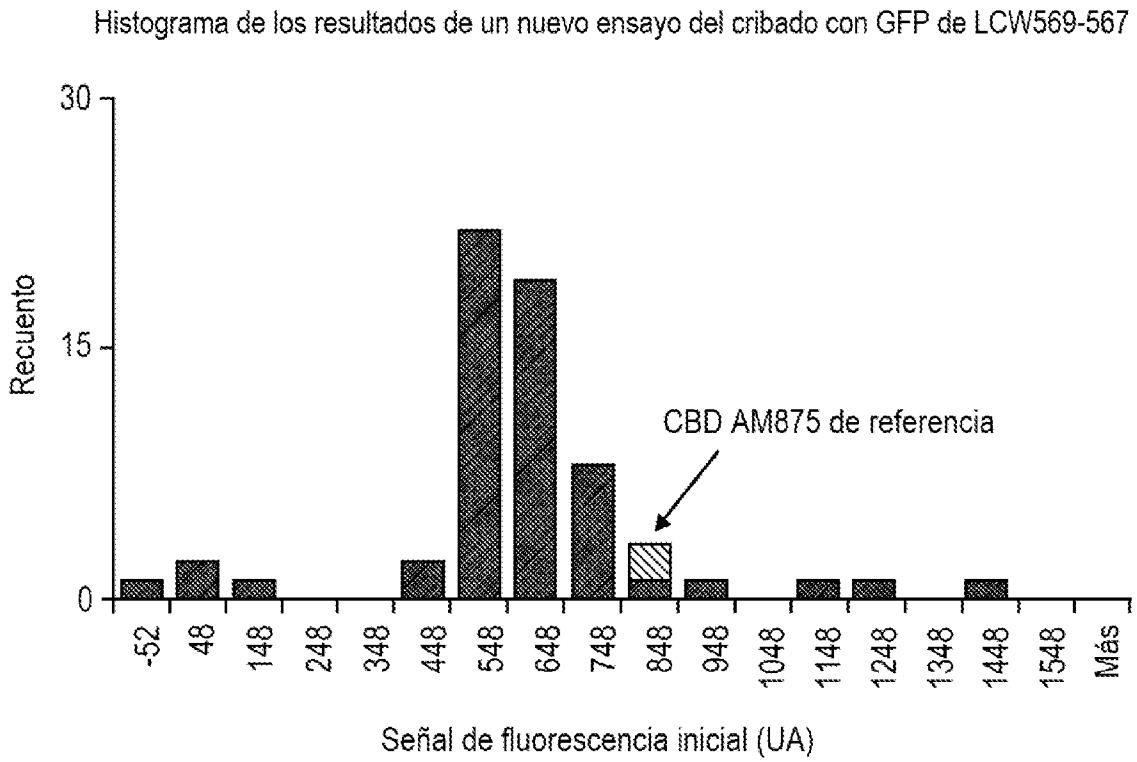

Histograma de los resultados de un nuevo ensayo del cribado con GFP de LCW569-567
Recuento
30
15
0
CBD AM875 de referencia
-52
48
148
248
348
448
548
648
748
848
948
1048
1148
1248
1348
1448
1548
Más
Señal de fluorescencia inicial (UA)

FIG. 13

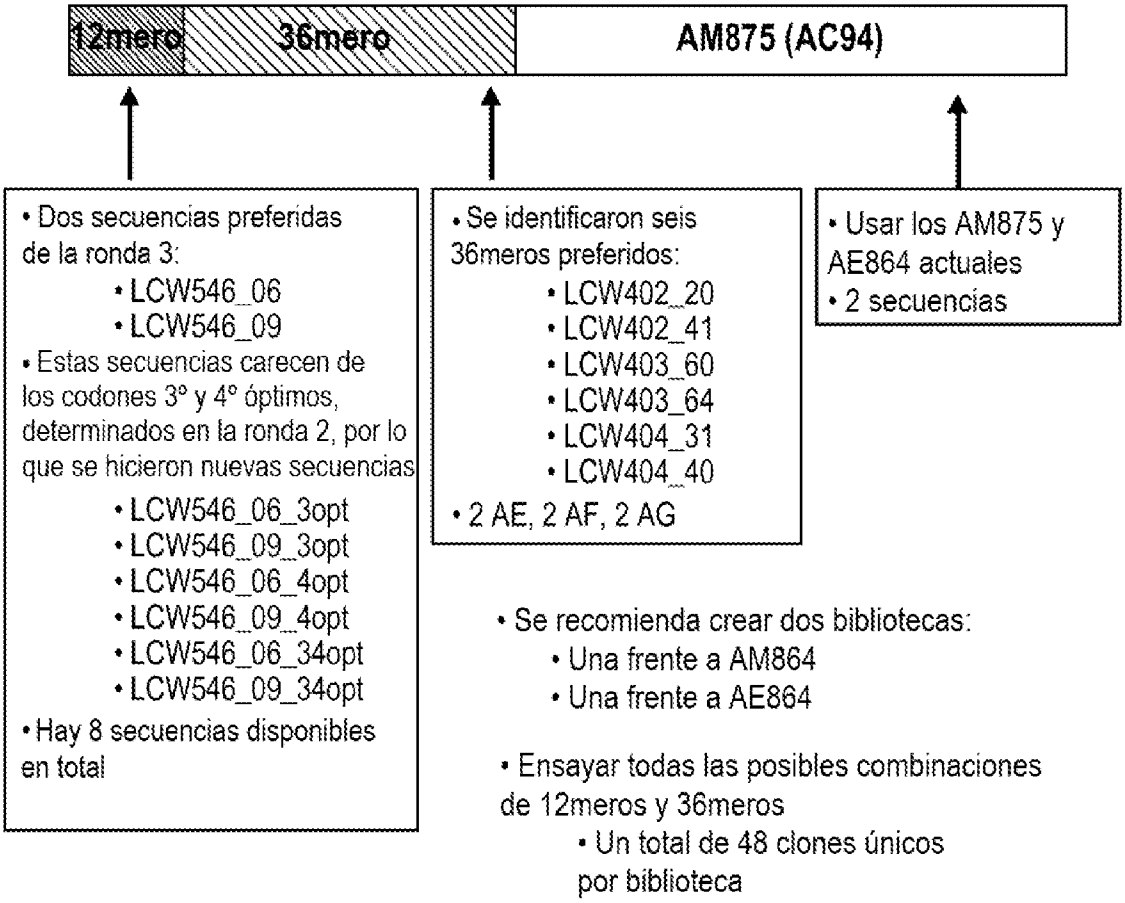

12mero
36mero
AM875 (AC94)
• Dos secuencias preferidas de la ronda 3:
• LCW546_06
• LCW546_09
• Estas secuencias carecen de los codones 3° y 4° óptimos, determinados en la ronda 2, por lo que se hicieron nuevas secuencias
• LCW546_06_3opt
• LCW546_09_3opt
• LCW546_06_4opt
• LCW546_09_4opt
• LCW546_06_34opt
• LCW546_09_34opt
• Hay 8 secuencias disponibles en total
• Se identificaron seis 36meros preferidos:
• LCW402_20
• LCW402_41
• LCW403_60
• LCW403_64
• LCW404_31
• LCW404_40
• 2 AE, 2 AF, 2 AG
• Usar los AM875 y AE864 actuales
• 2 secuencias
• Se recomienda crear dos bibliotecas:
• Una frente a AM864
• Una frente a AE864
• Ensayar todas las posibles combinaciones de 12meros y 36meros
• Un total de 48 clones únicos por biblioteca

FIG. 14

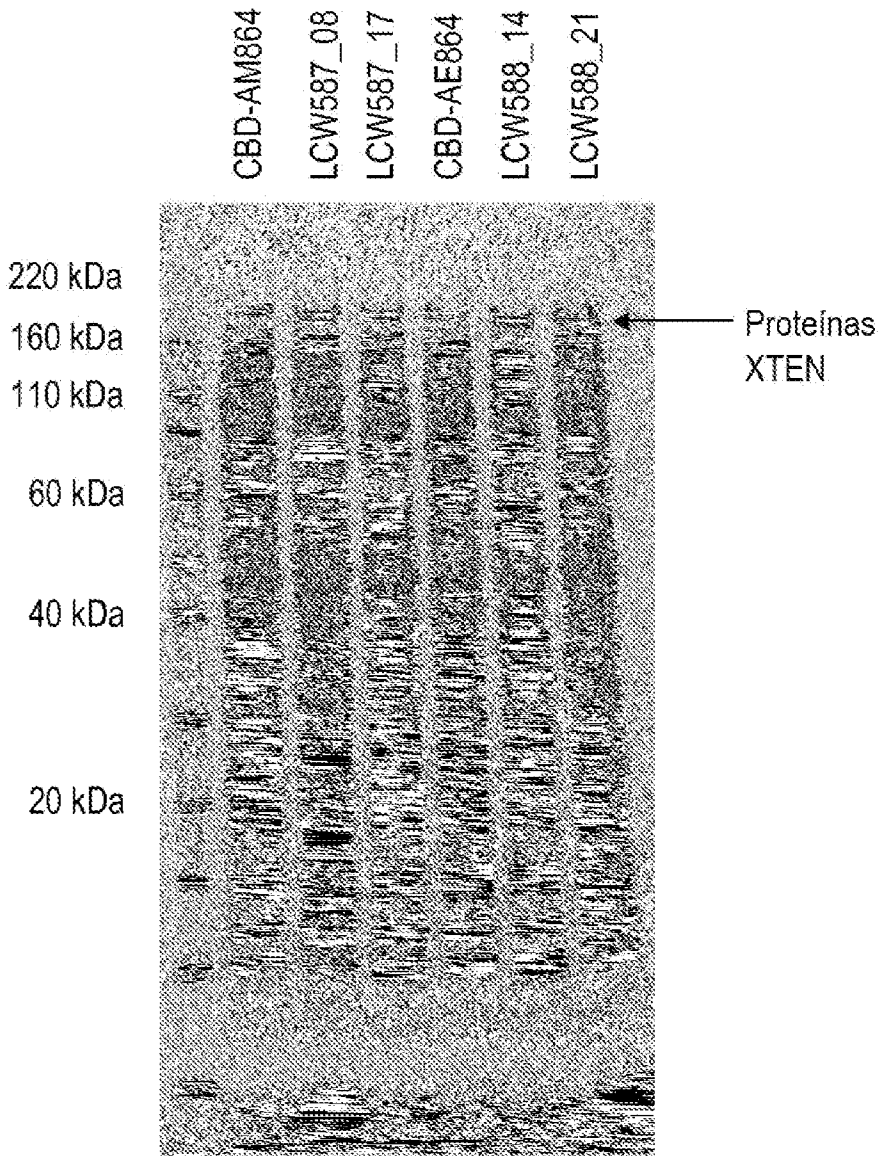

CBD-AM864
LCW587_08
LCW587_17
CBD-AE864
LCW588_14
LCW588_21
220 kDa
160 kDa
110 kDa
60 kDa
40 kDa
20 kDa
Proteínas XTEN

FIG. 15

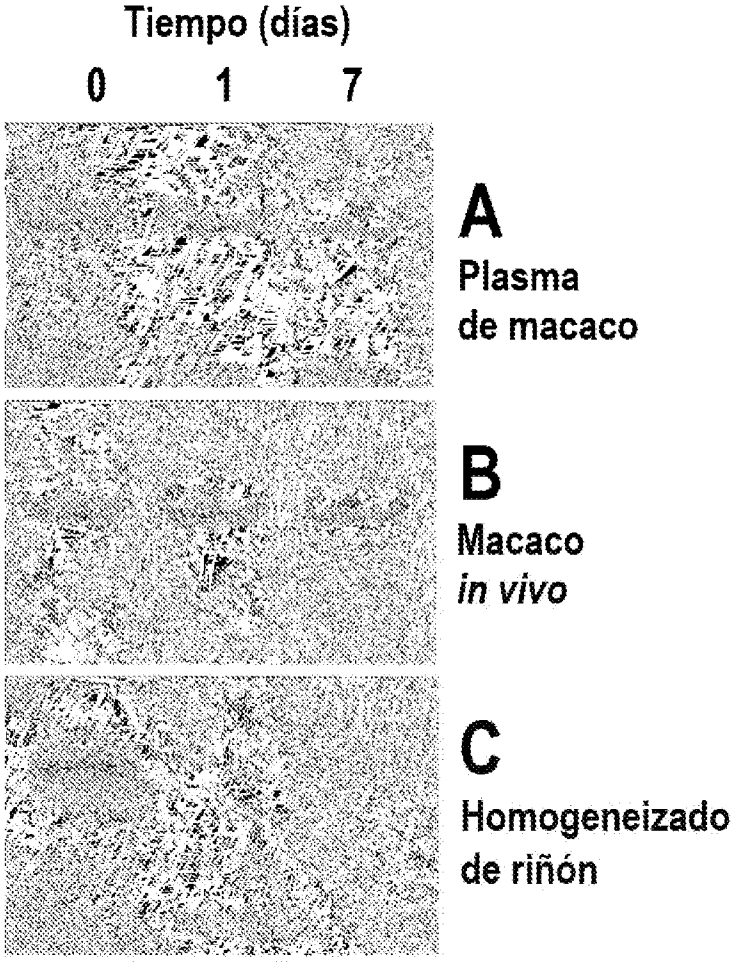

Tiempo (días)
0
1
7
A
Plasma
de macaco
B
Macaco
in vivo
C
Homogeneizado
de riñón

FIG. 16

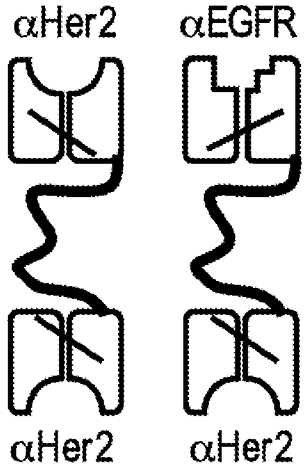


FIG. 17A

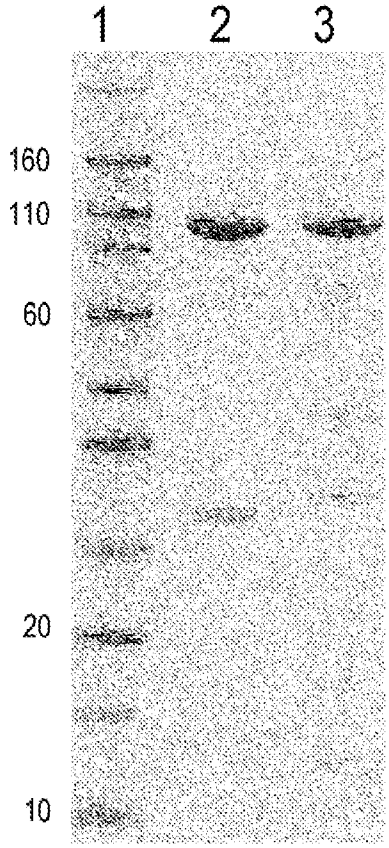


FIG. 17B

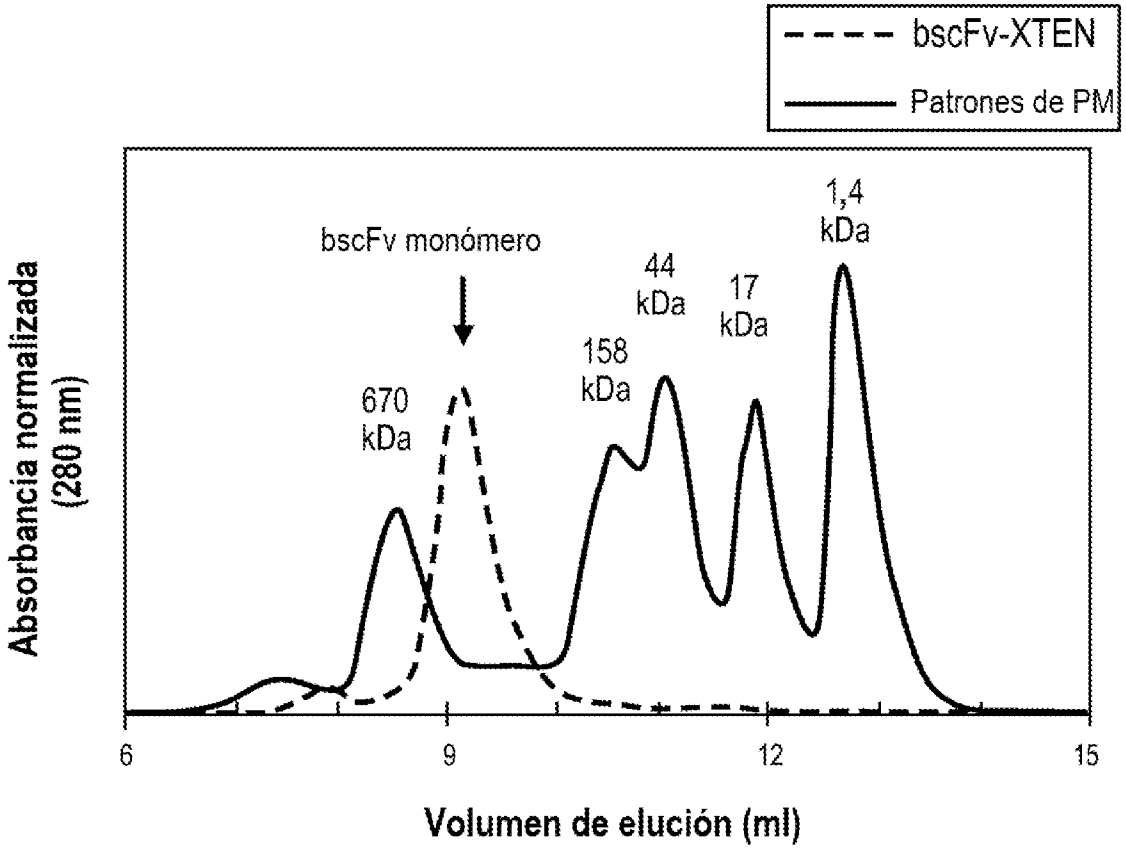

bscFv-XTEN
Patrones de PM
bscFv monómero
670 kDa
158 kDa
44 kDa
17 kDa
1,4 kDa
Absorbancia normalizada (280 nm)
6
9
12
15
Volumen de elución (ml)

FIG. 17C

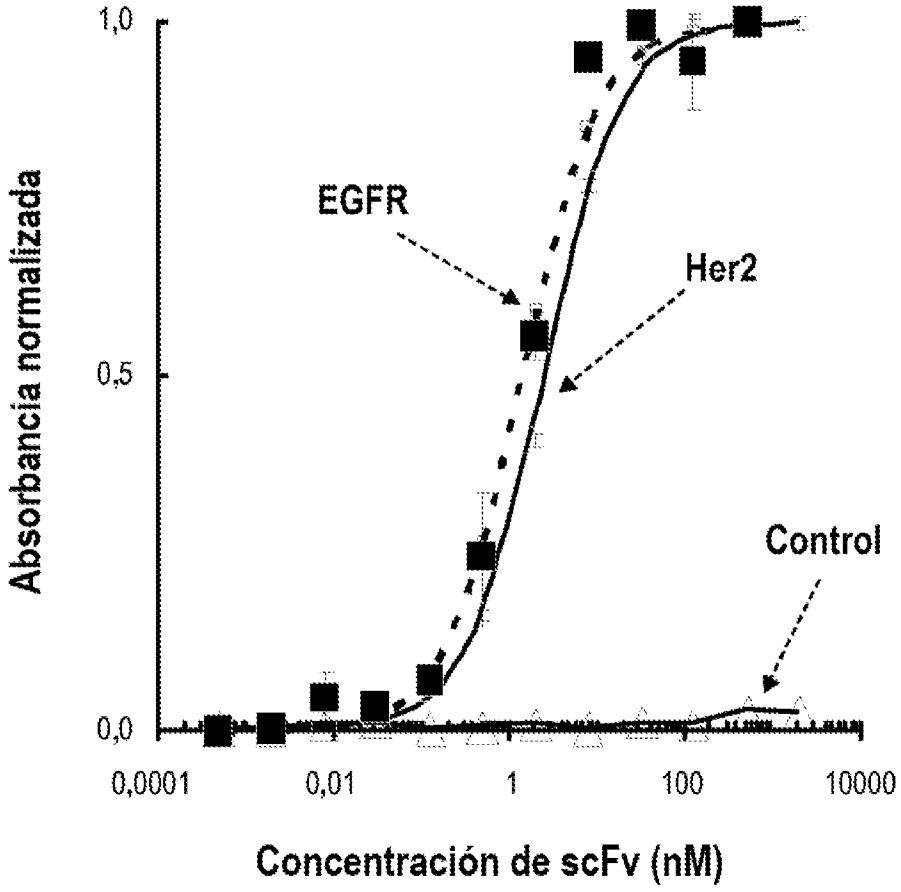

1,0
0,5
0,0
Absorbancia normalizada
EGFR
Her2
Control
0,0001
0,01
1
100
10000
Concentración de scFv (nM)

FIG. 18

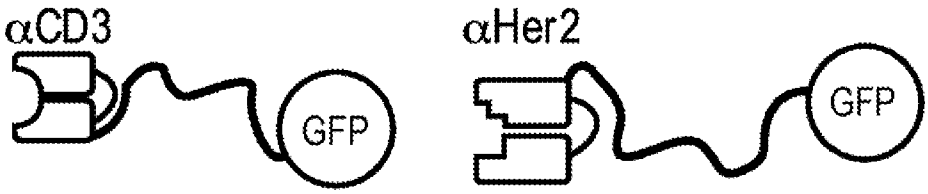


FIG. 19A

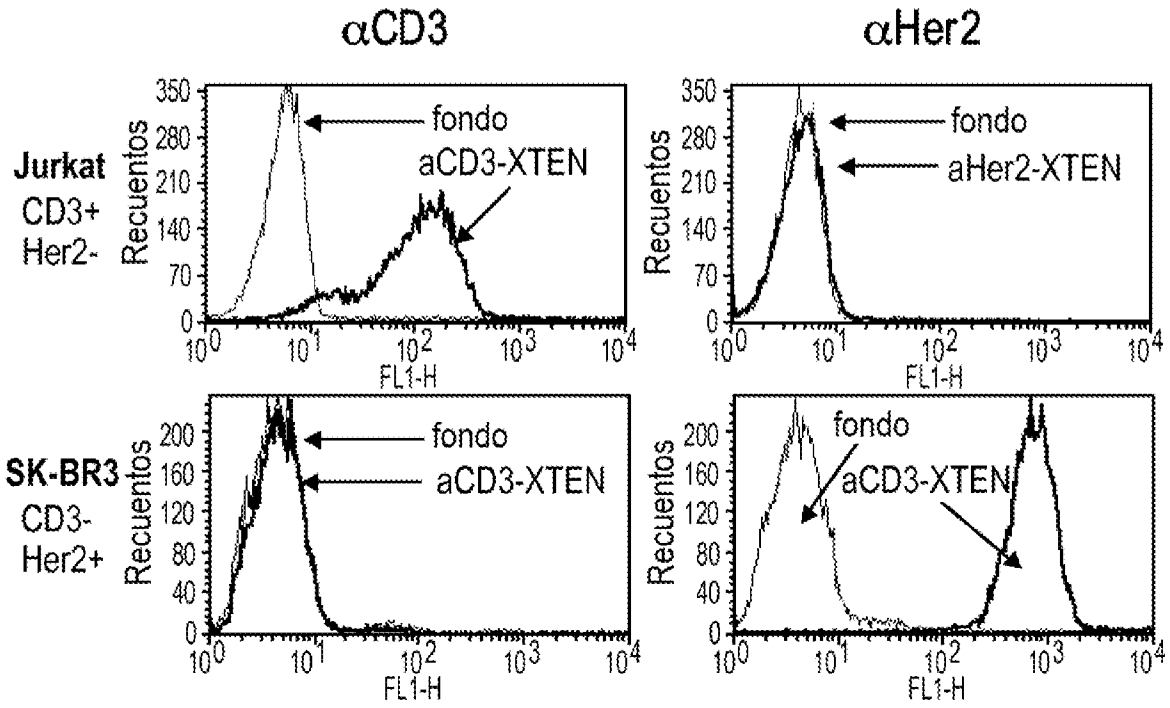


FIG. 19B

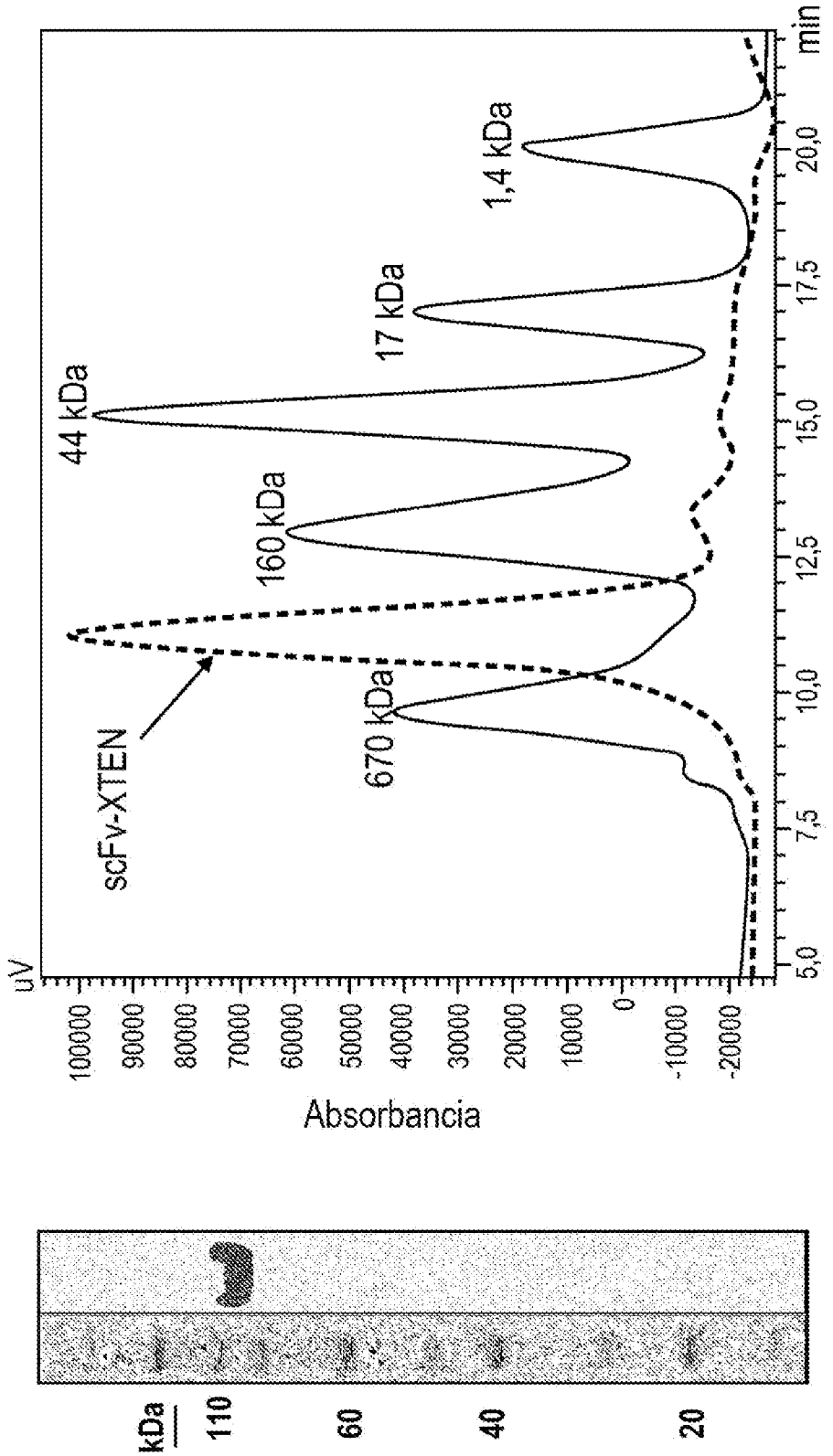

uV
100000
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
-10000
-20000
Absorbancia
5,0
7,5
10,0
12,5
15,0
17,5
20,0
min
Volumen de elución
44 kDa
160 kDa
670 kDa
17 kDa
1,4 kDa
scFv-XTEN
kDa
110
60
40
20

FIG. 20A

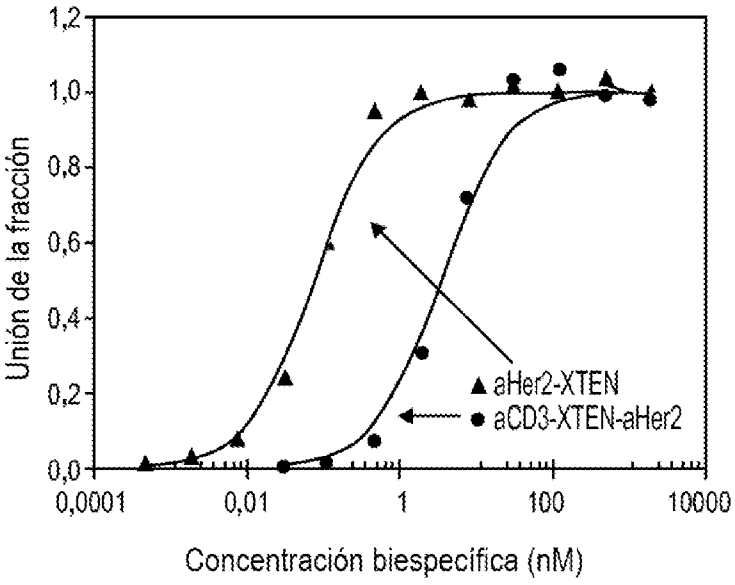


FIG. 21A

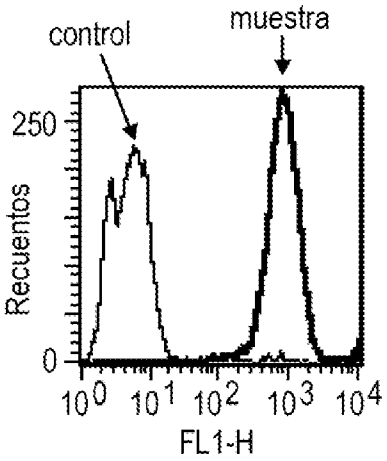


FIG. 21B

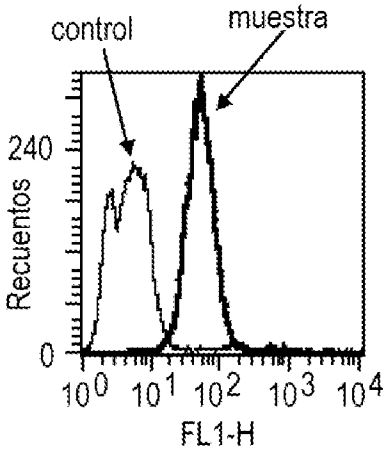


FIG. 21C

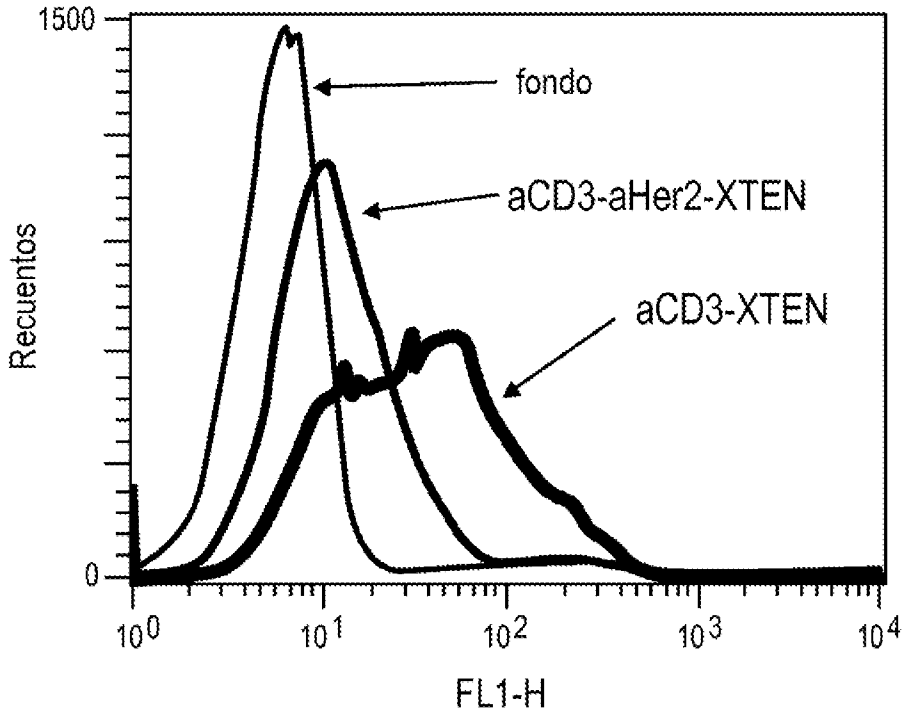

1500
0
Recuentos
fondo
aCD3-aHer2-XTEN
aCD3-XTEN
10^0
10^1
10^2
10^3
10^4
FL1-H

FIG. 22

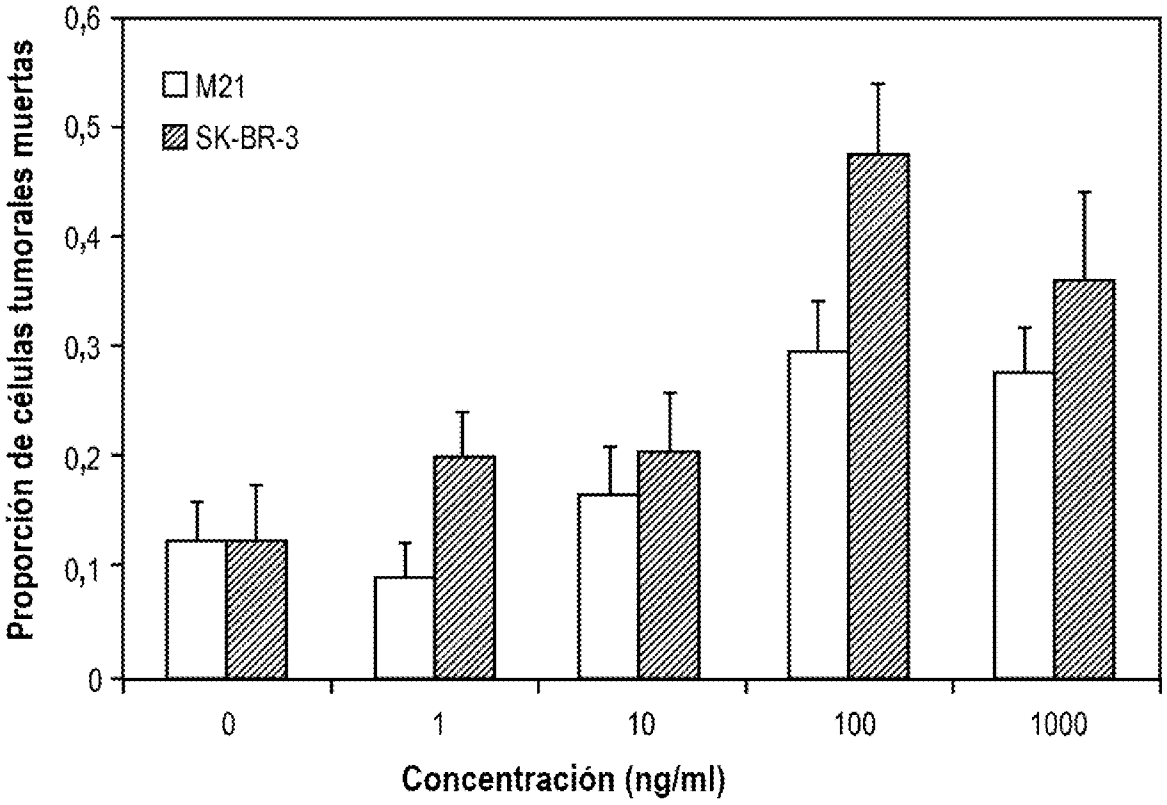

0,6
0,5
0,4
0,3
0,2
0,1
0
M21
SK-BR-3
Proporción de células tumorales muertas
0
1
10
100
1000
Concentración (ng/ml)

FIG. 23

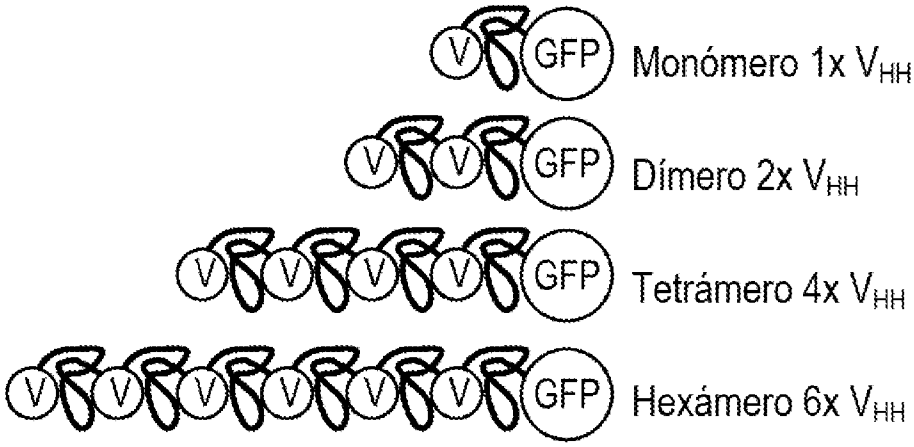


FIG. 24A

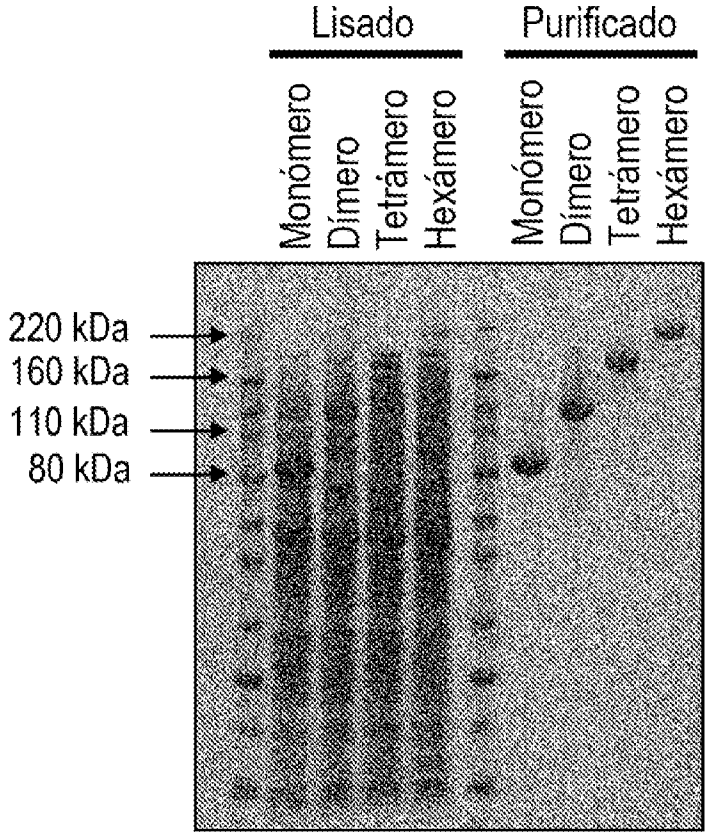


FIG. 24B

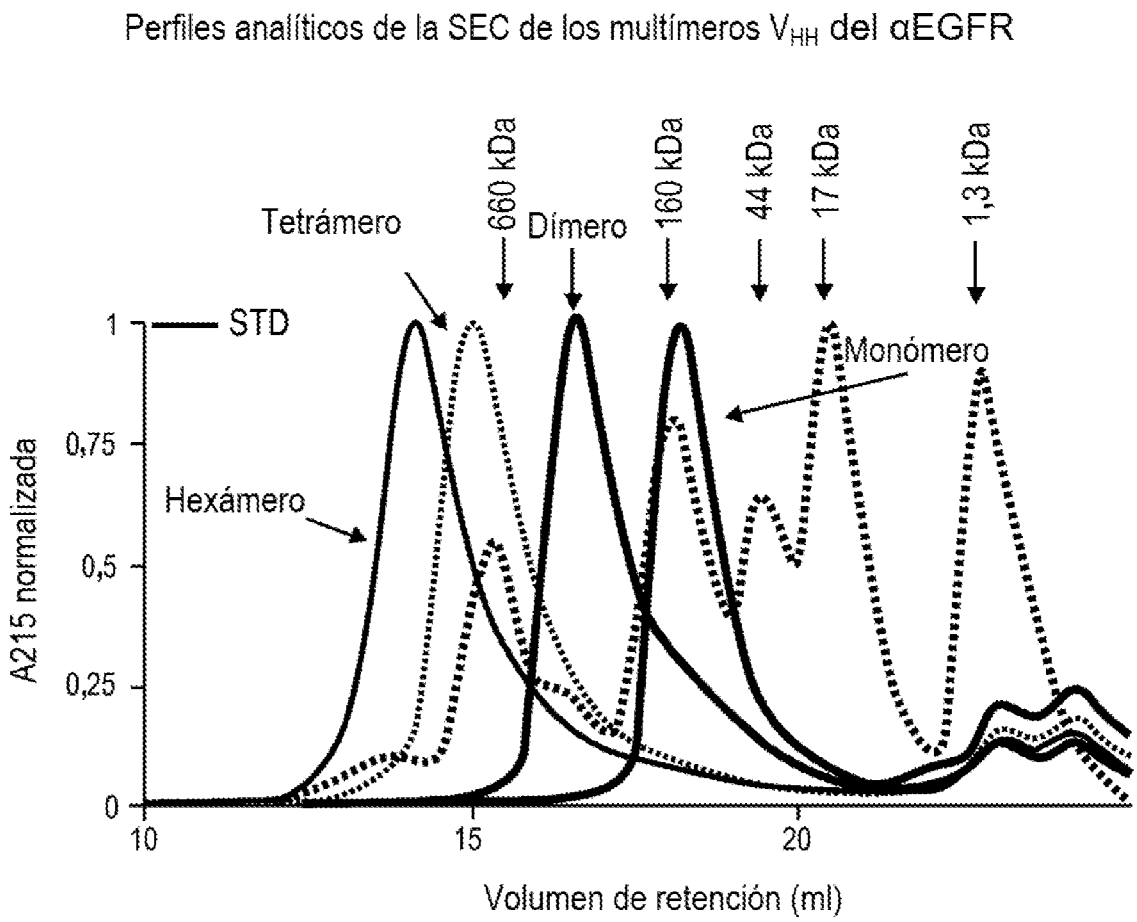

Perfiles analíticos de la SEC de los multímeros V_{HH} del αEGFR
660 kDa
160 kDa
44 kDa
17 kDa
1,3 kDa
Tetrámero
Dímero
Monómero
Hexámero
STD
A215 normalizada
1
0,75
0,5
0,25
0
10
15
20
Volumen de retención (ml)

FIG. 25

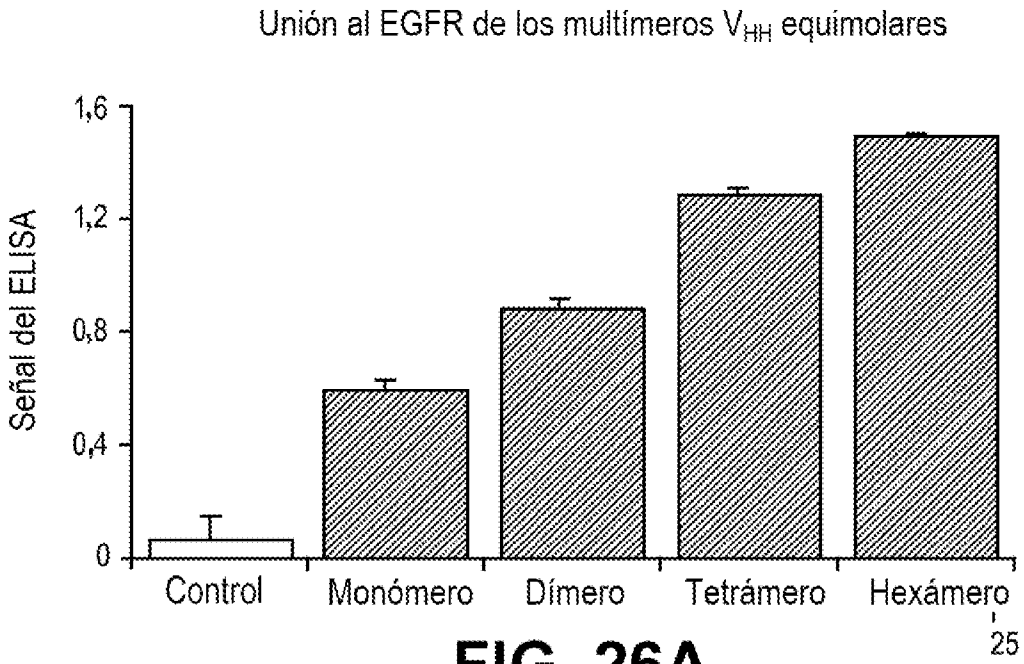


FIG. 26A

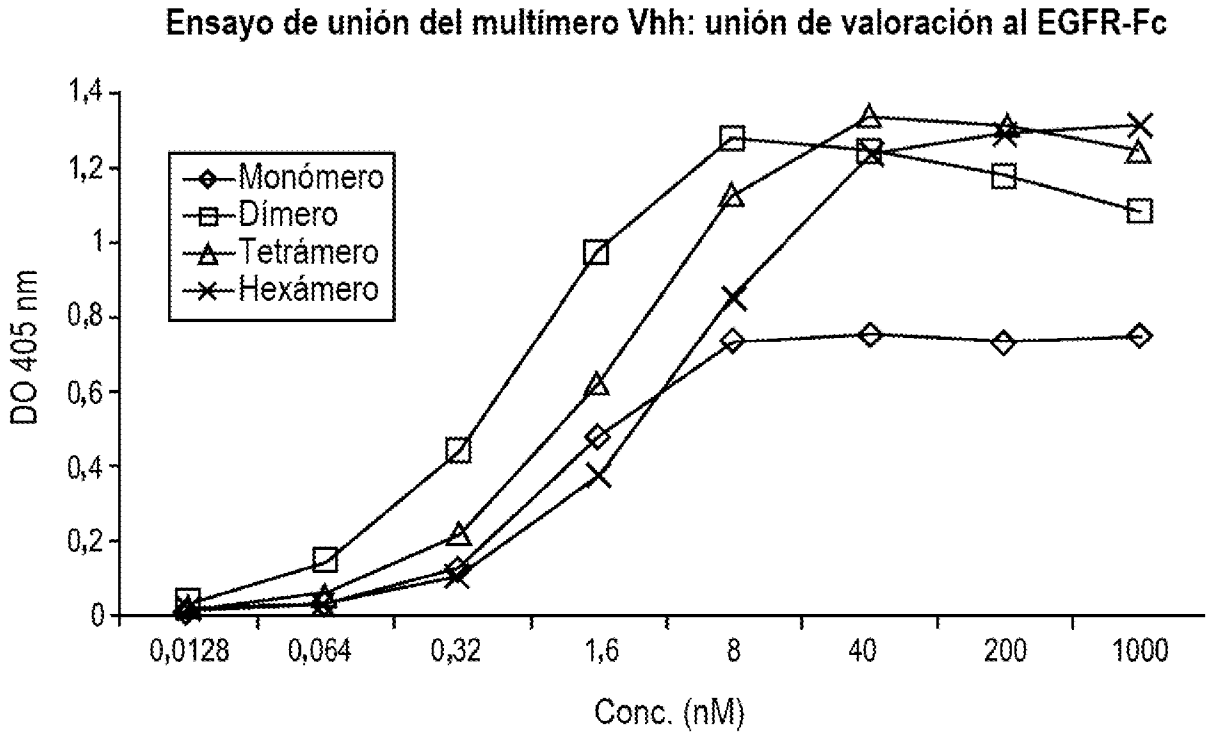


FIG. 26B

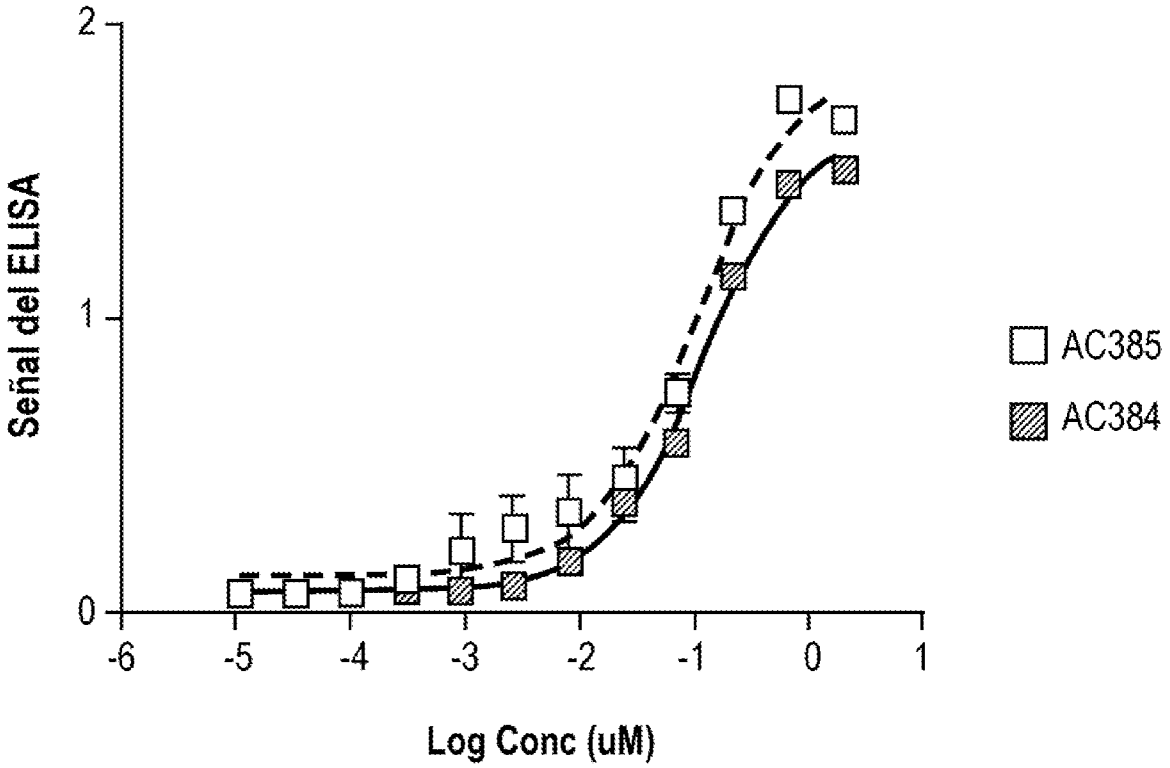


FIG. 27

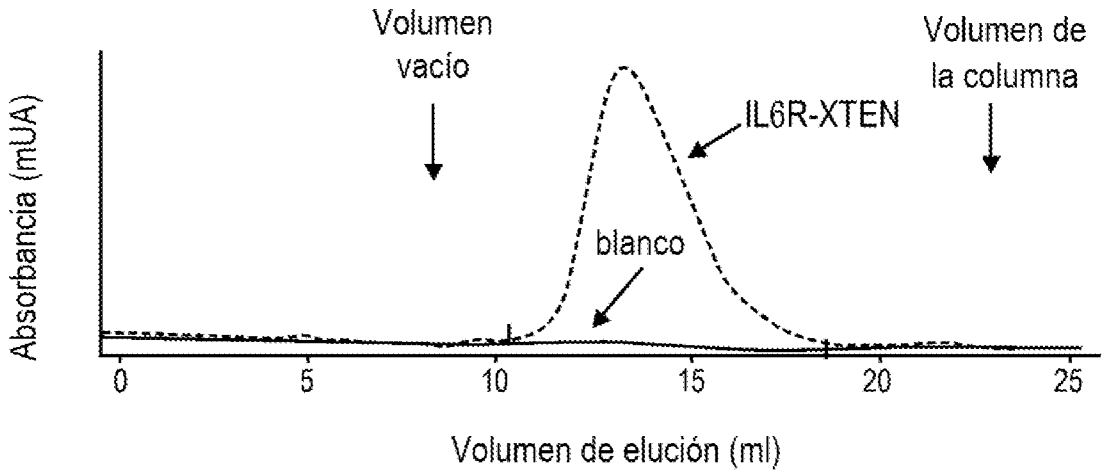


FIG. 28A

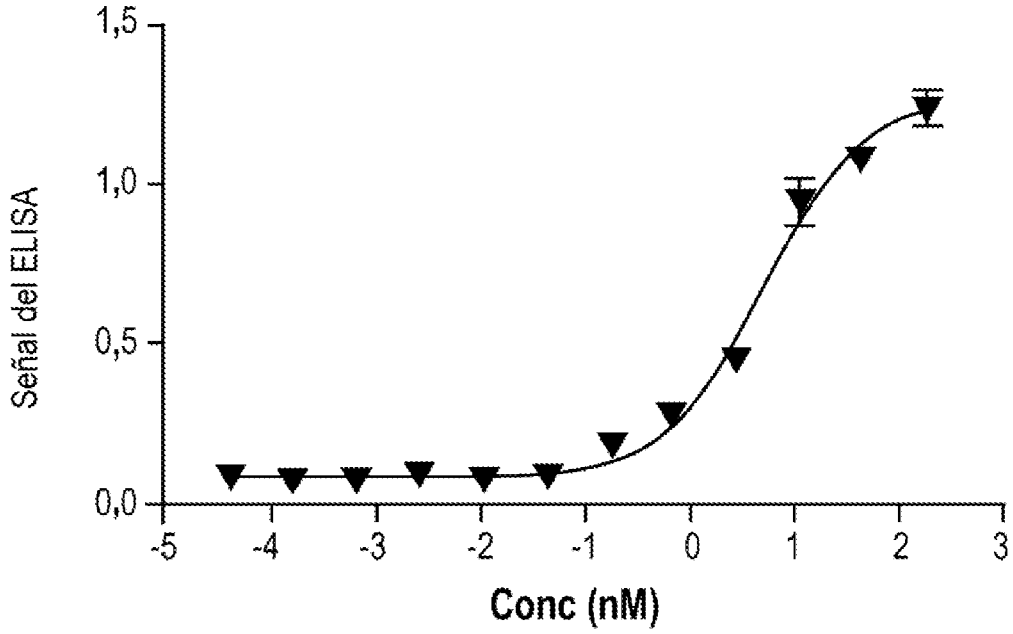


FIG. 28B

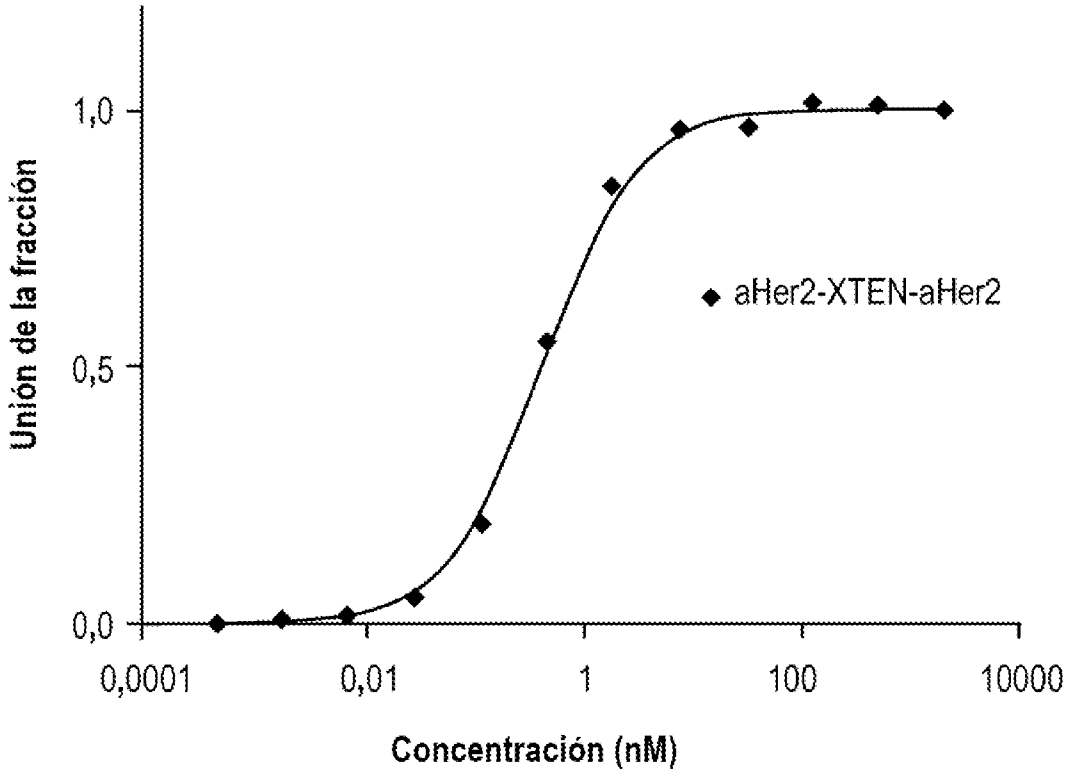

1,0
0,5
0,0
Unión de la fracción
0,0001
0,01
1
100
10000
Concentración (nM)
aHer2-XTEN-aHer2

FIG. 29

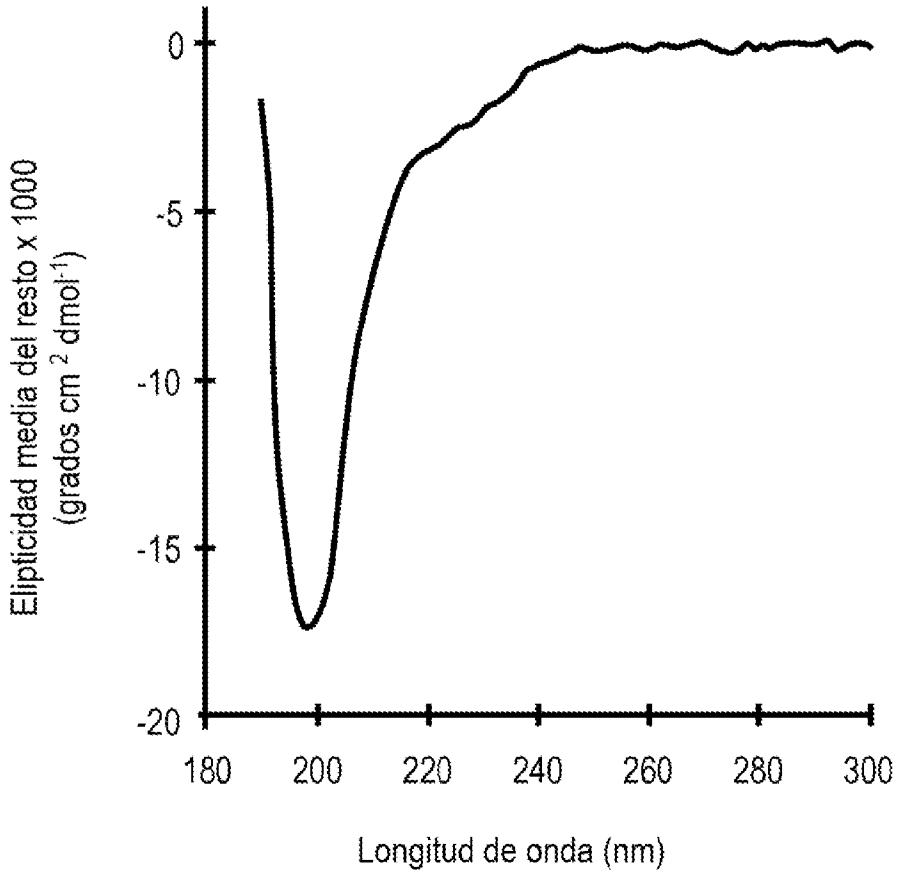

0
-5
-10
-15
-20
180
200
220
240
260
280
300
Elipticidad media del resto x 1000
(grados cm 2 dmol-1)
Longitud de onda (nm)

FIG. 30

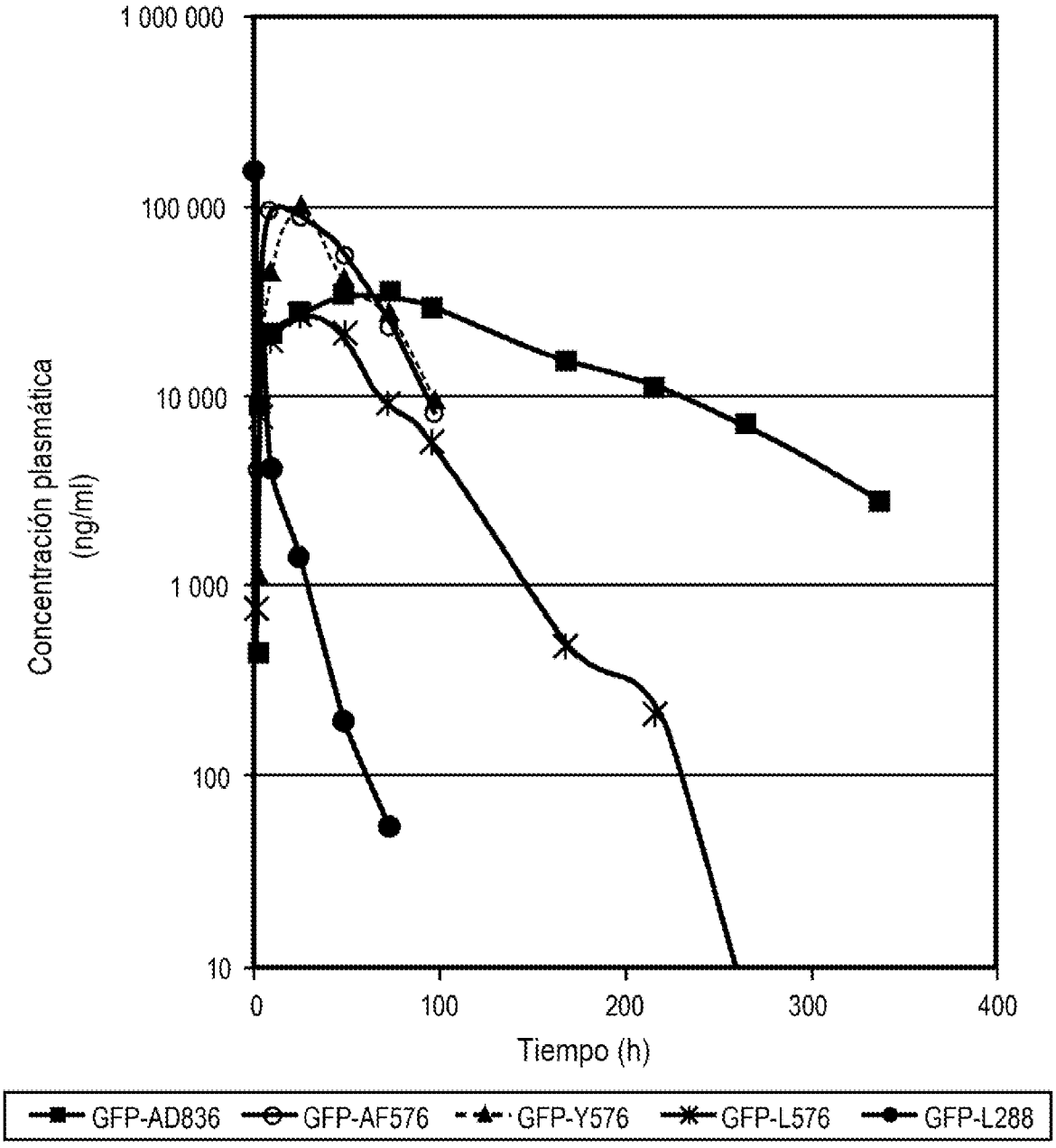
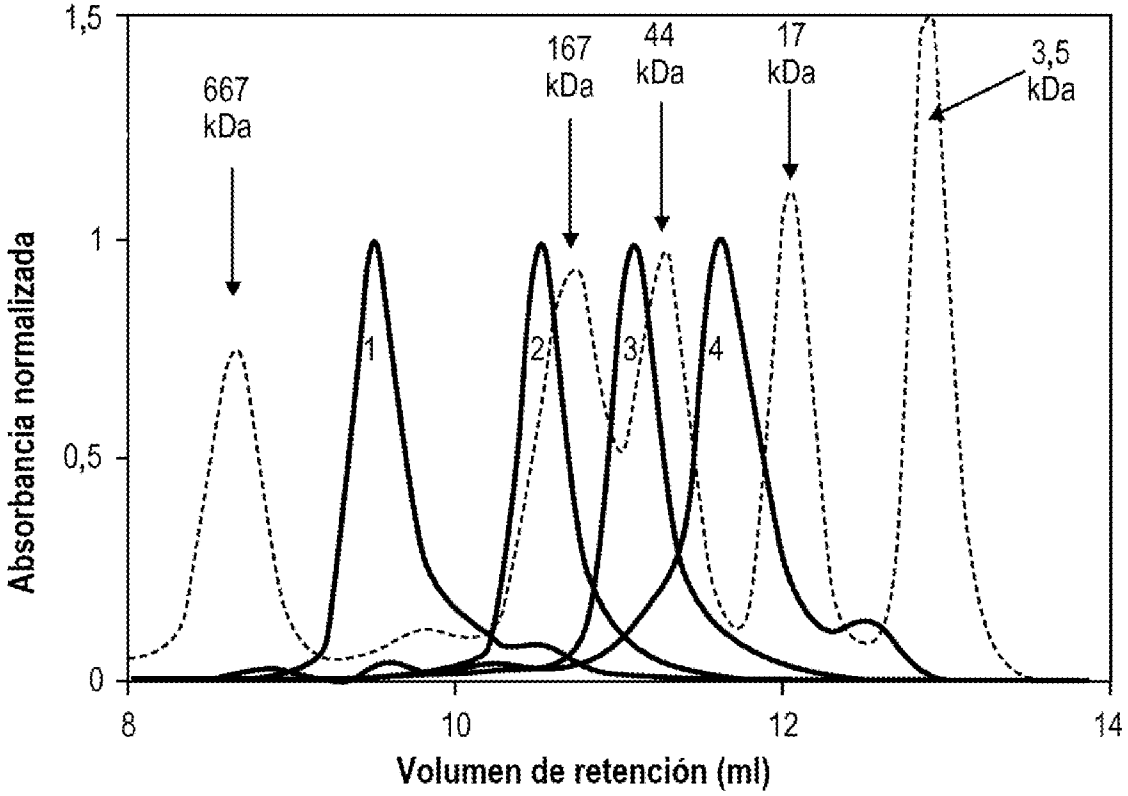

1 000 000
100 000
10 000
1 000
100
10
Concentración plasmática (ng/ml)
0
100
200
300
400
Tiempo (h)
GFP-AD836
GFP-AF576
GFP-Y576
GFP-L576
GFP-L288

FIG. 31



1. Glucagón-Y288
2. Glucagón-Y144
3. Glucagón-Y72
4. Glucagón-Y36

------ = Patrones

FIG. 32

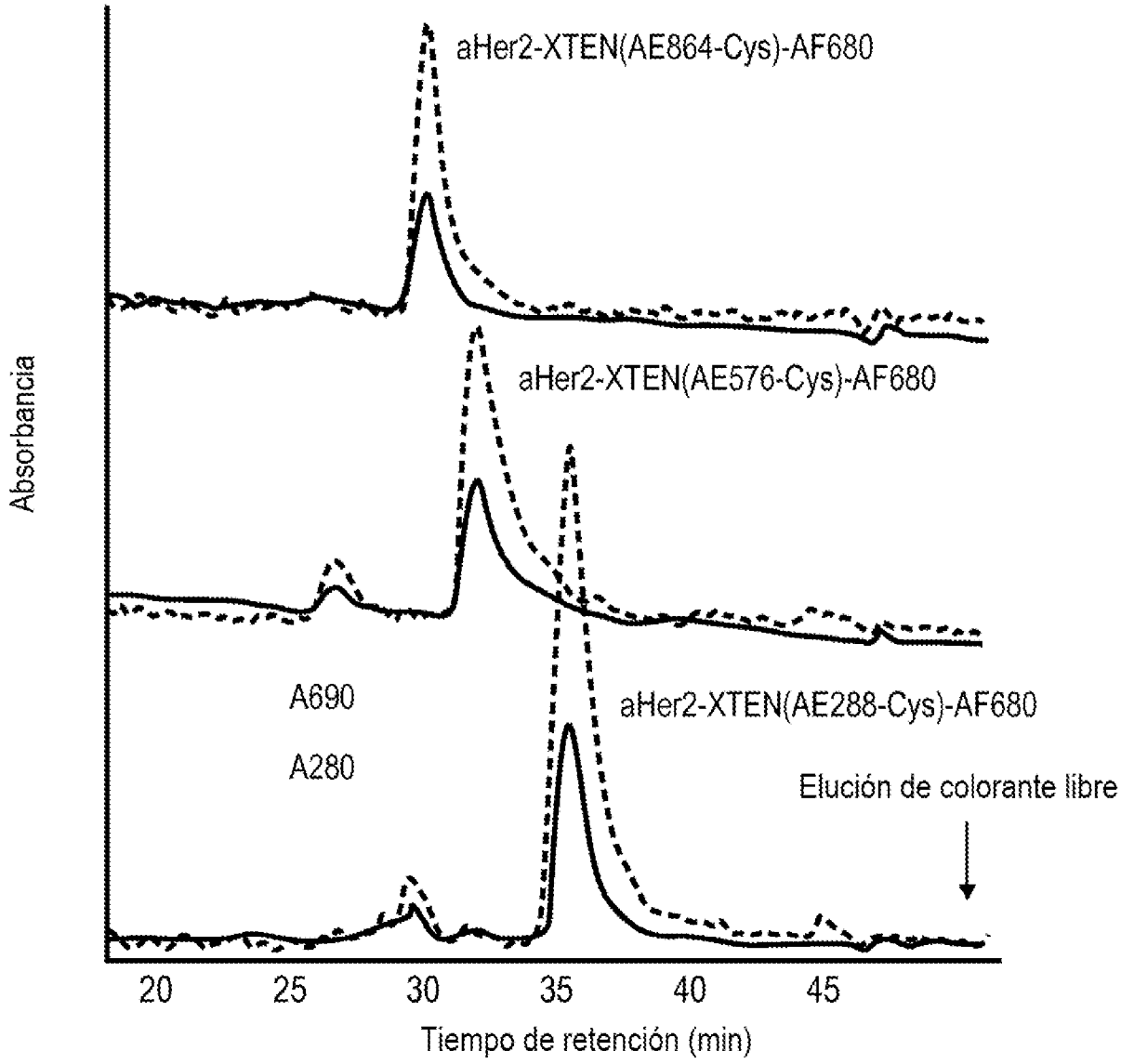

aHer2-XTEN(AE864-Cys)-AF680
aHer2-XTEN(AE576-Cys)-AF680
aHer2-XTEN(AE288-Cys)-AF680
A690
A280
Elución de colorante libre
Absorbancia
20
25
30
35
40
45
Tiempo de retención (min)

FIG. 33

FIG. 34A

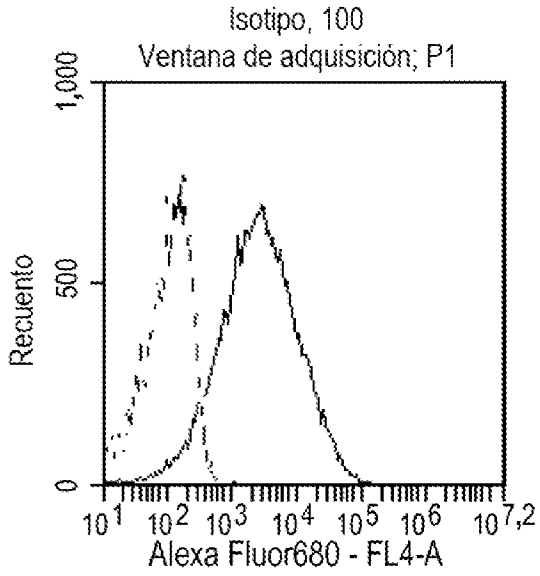


FIG. 34B

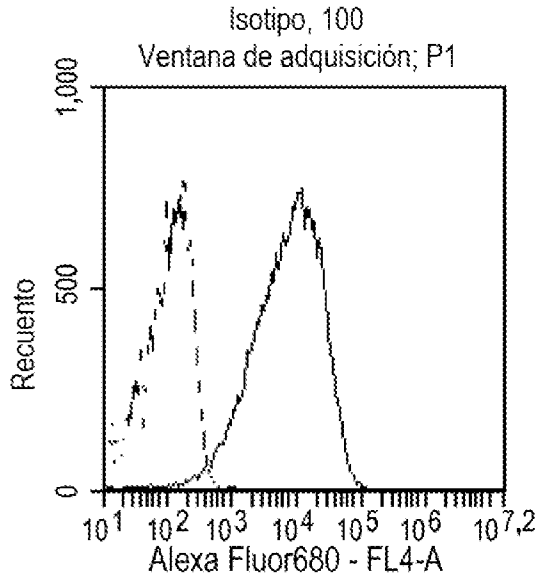
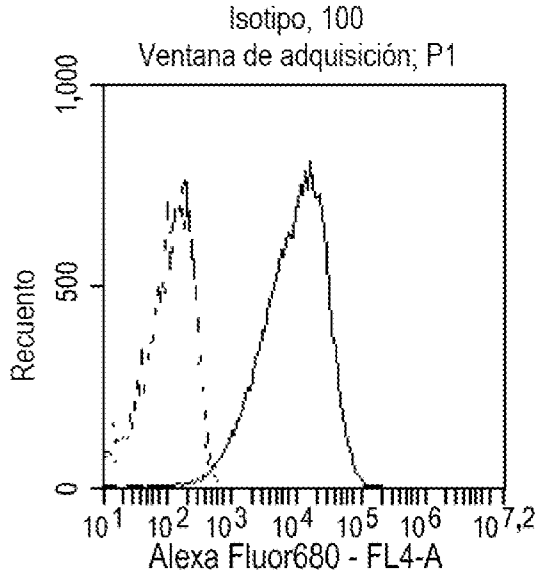


FIG. 34C



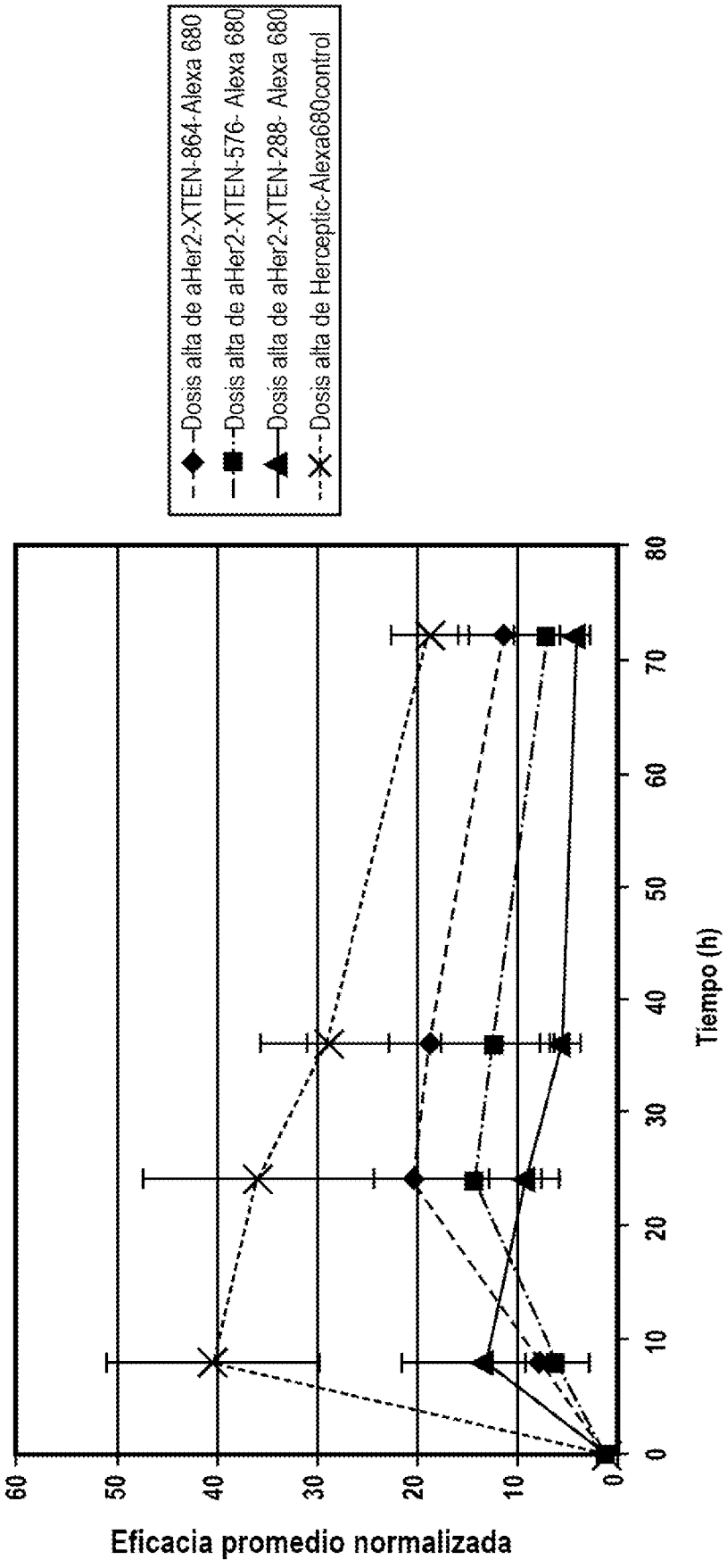

Dosis alta de aHer2-XTEN-864-Alexa 680
Dosis alta de aHer2-XTEN-576- Alexa 680
Dosis alta de aHer2-XTEN-288- Alexa 680
Dosis alta de Herceptic-Alexa680control
Eficacia promedio normalizada
Tiempo (h)
0
10
20
30
40
50
60
0
10
20
30
40
50
60
70
80

FIG. 35

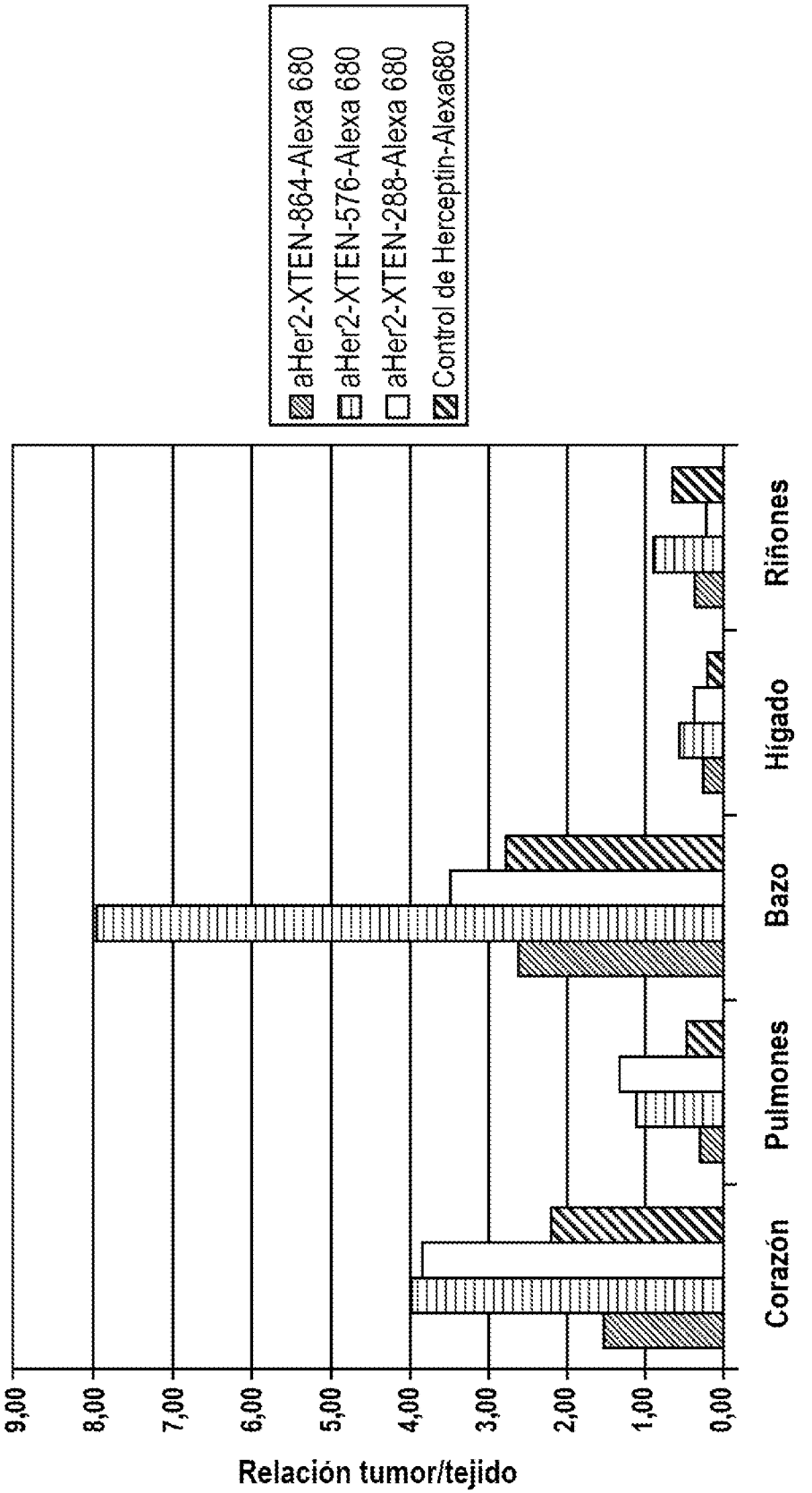

aHer2-XTEN-864-Alexa 680
aHer2-XTEN-576-Alexa 680
aHer2-XTEN-288-Alexa 680
Control de Herceptin-Alexa680
9,00
8,00
7,00
6,00
5,00
4,00
3,00
2,00
1,00
0,00
Relación tumor/tejido
Corazón
Pulmones
Bazo
Hígado
Riñones

FIG. 36

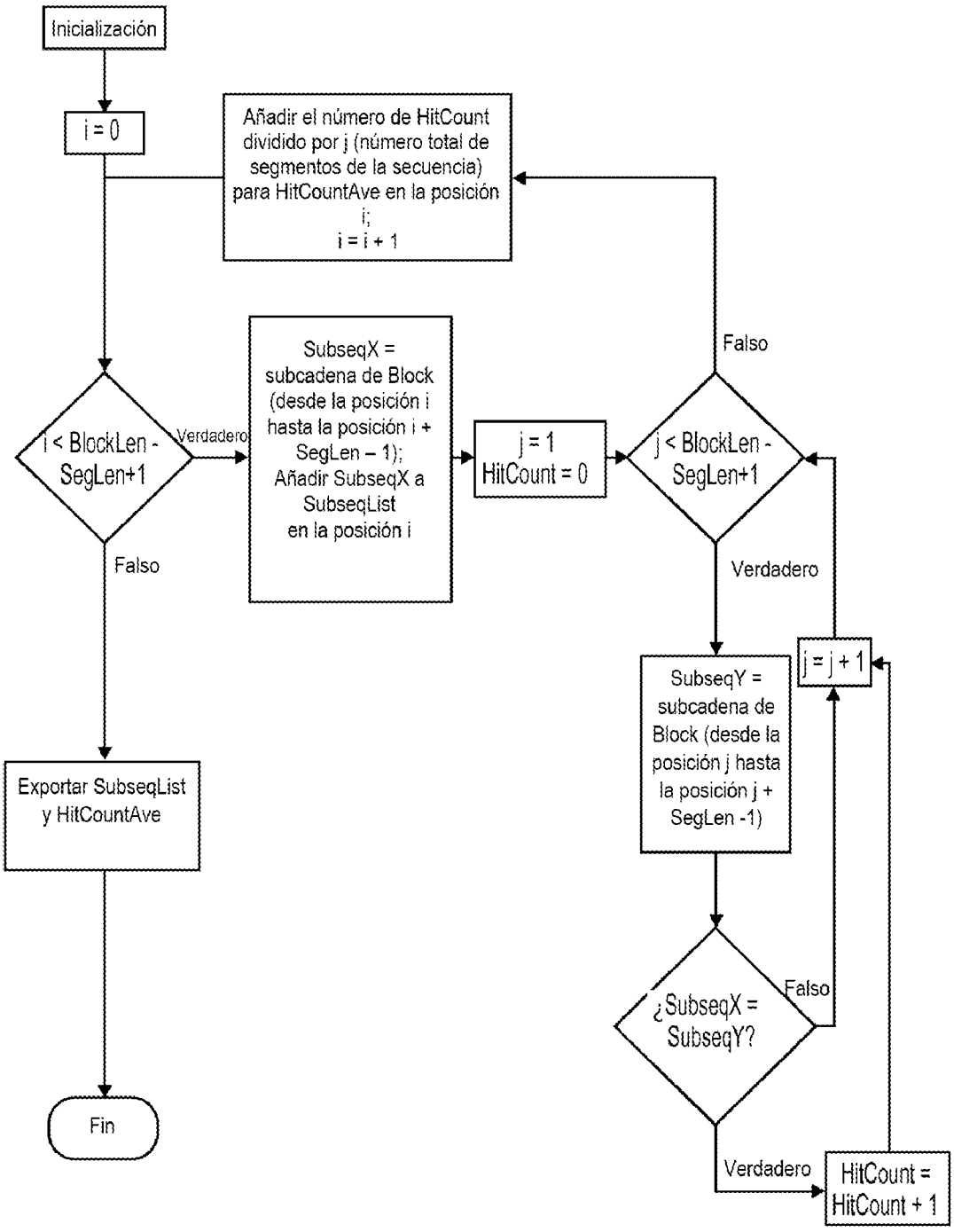

Inicialización
i = 0
Añadir el número de HitCount dividido por j (número total de segmentos de la secuencia) para HitCountAve en la posición i;
i = i + 1
i < BlockLen - SegLen+1
Verdadero
Falso
SubseqX = subcadena de Block (desde la posición i hasta la posición i + SegLen – 1); Añadir SubseqX a SubseqList en la posición i
j = 1
HitCount = 0
j < BlockLen - SegLen+1
Falso
Verdadero
SubseqY = subcadena de Block (desde la posición j hasta la posición j + SegLen -1)
j = j + 1
¿SubseqX = SubseqY?
Falso
Verdadero
HitCount = HitCount + 1
Exportar SubseqList y HitCountAve
Fin

FIG. 37

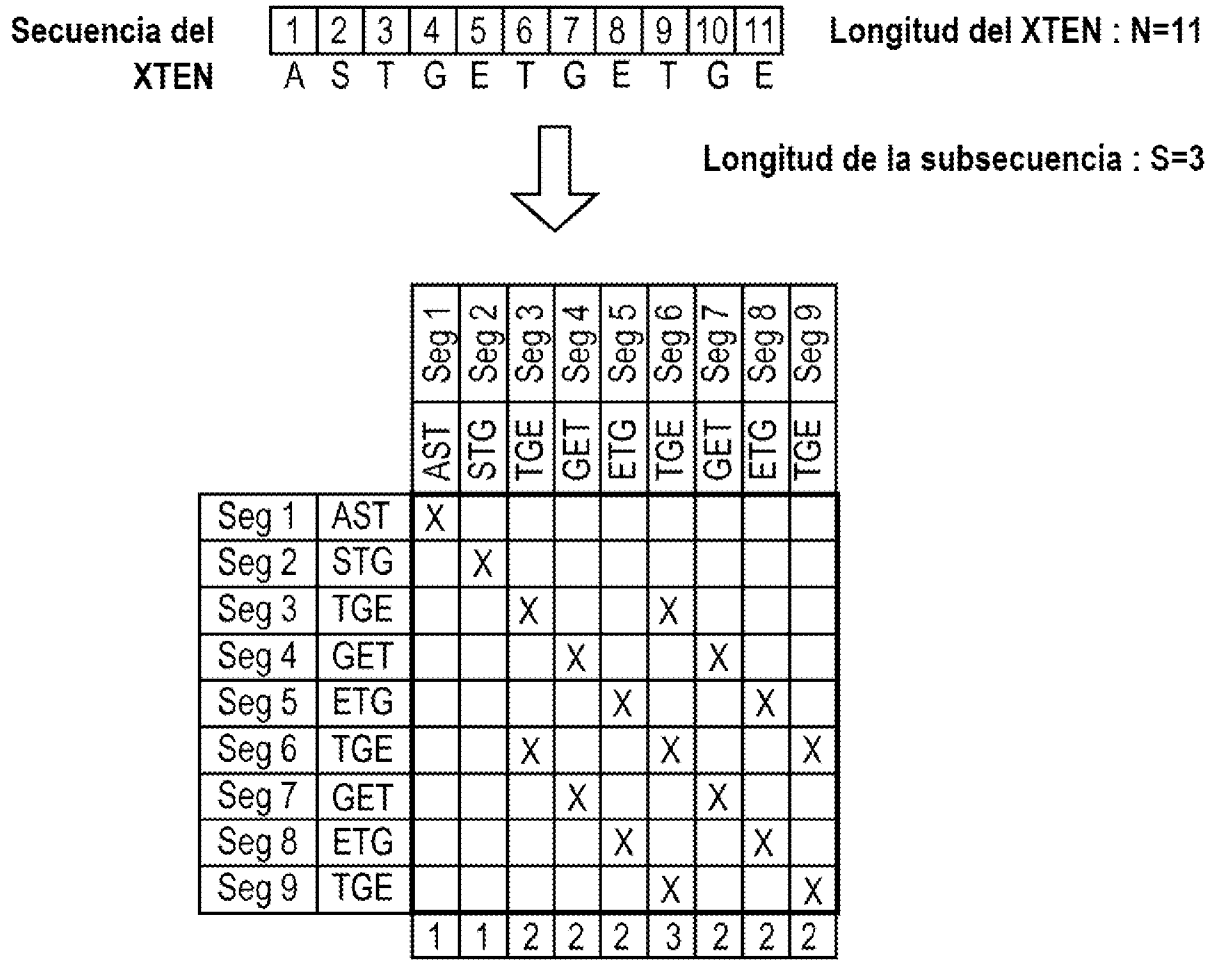

Secuencia del XTEN
1 2 3 4 5 6 7 8 9 10 11
A S T G E T G E T G E
Longitud del XTEN : N=11
Longitud de la subsecuencia : S=3
Seg 1 Seg 2 Seg 3 Seg 4 Seg 5 Seg 6 Seg 7 Seg 8 Seg 9
AST STG TGE GET ETG TGE GET ETG TGE
Seg 1 AST X
Seg 2 STG X
Seg 3 TGE X X
Seg 4 GET X X
Seg 5 ETG X X
Seg 6 TGE X X X
Seg 7 GET X X
Seg 8 ETG X X
Seg 9 TGE X X
1 1 2 2 2 3 2 2 2
Puntuación de la subsecuencia = 1,89

FIG. 38

AG864

GASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTG
PGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTG
SPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTG
TGPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSS
TGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGT
ASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGASPGT
SSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASPG
TSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASP
GTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSS
TPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGS
STPSGATGSPGSSPSASTGTGPGSSPSACTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPG
SSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSP
GASPGTSSTGSP

AG576

PGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATG
SPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSST
GSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSS
TGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSG
ATGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGT
SSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGS
GTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSP
SASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGAS
PGTSSTGS

FIG. 39A

AG864

GASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTG
PGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTG
SPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTG
TGPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSS
TGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGT
ASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGASPGT
SSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASPG
TSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASP
GTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSS
TPSGATGSPGSSTPSGATGSPGSSPSATASSSPGSSTPSGATGSSTGTGPGSSPSASTGTGPGASPGTSST
GSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSS
TGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSG
ATGSPGSSTPSGATGSPGASPGTSSTGSP

AG288

ASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSST
PSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPG
SSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGT
GPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTS
STGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGS

FIG. 39B

AG864

GASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTG
PGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTG
SPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTG
TGPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSS
TGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGT
ASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGASPGT
SSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASPG
TSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASP
GTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSS
TPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGS
STPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPG
SSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSP
GASPGTSSTGSP

AG144

PGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSS
TGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGTPGS
GTASSS

FIG. 39C

AE864

GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSA
PGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGS
APGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEG
SAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATP
ESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPS
EGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEP
SEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTE
PSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTST
EPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTS
TEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGS
PAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPG
SEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGP
GTSTEPSEGSAP

AE576

GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEG
SAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEP
SEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTS
TEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGP
GTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEG
SAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGS
PTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTS
ESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEE
GTSESATPESGPGTSTEPSEGSAP

FIG. 39D

AE864

GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSA
PGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGS
APGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEG
SAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATP
ESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPS
EGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEP
SEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTE
PSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTST
EPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTS
TEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGS
PAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPG
SEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGP
GTSTEPSEGSAP

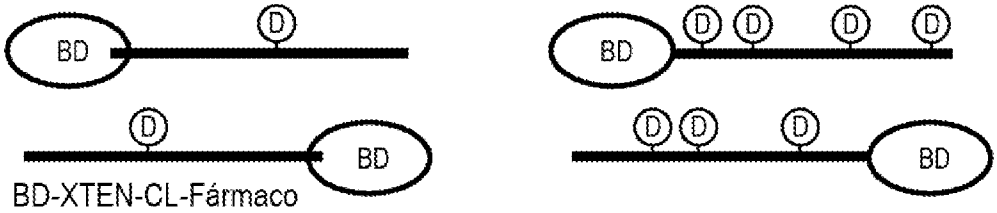
AE288

GTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEG
SAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGS
PTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSE
PATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGP
GTSTEPSEGSAP

FIG. 39E

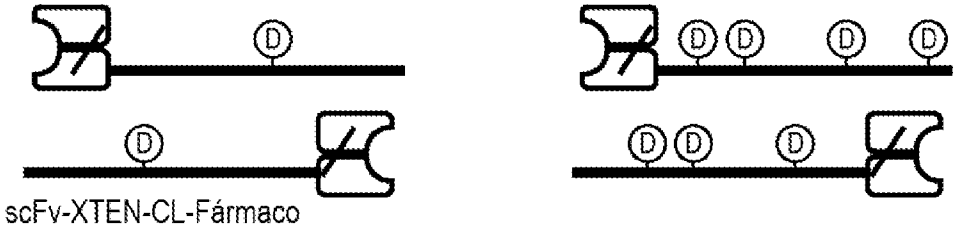
XTEN-CL-Fármaco

FIG. 40A



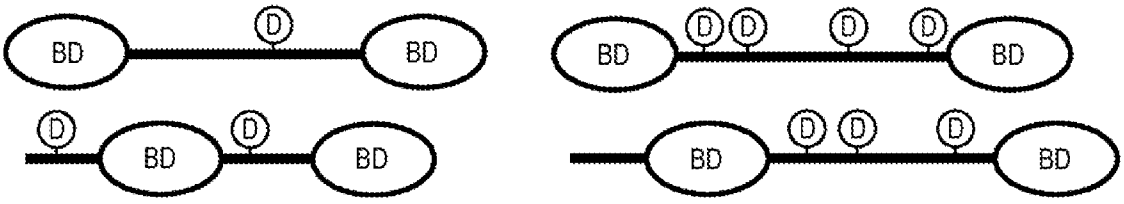
BD-XTEN-CL-Fármaco

FIG. 40B



scFv-XTEN-CL-Fármaco

FIG. 40C



Conjugados multivalentes

FIG. 40D

FIG. 41A

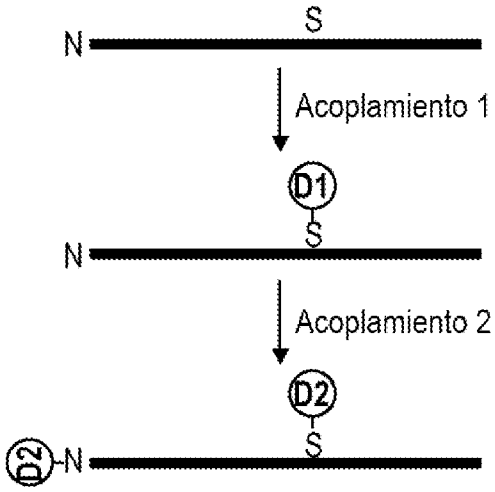


FIG. 41B

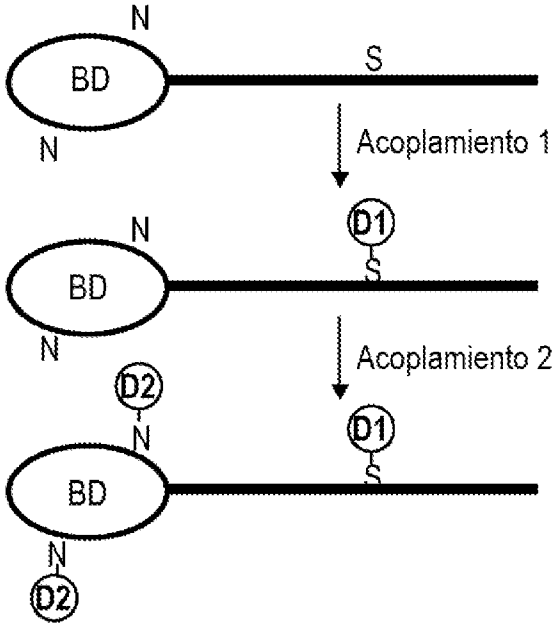


FIG. 41C

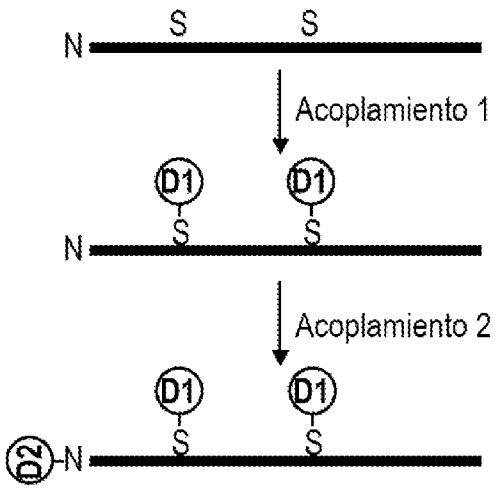
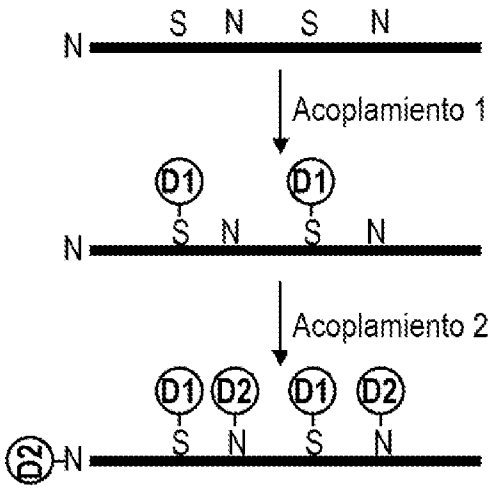


FIG. 41D



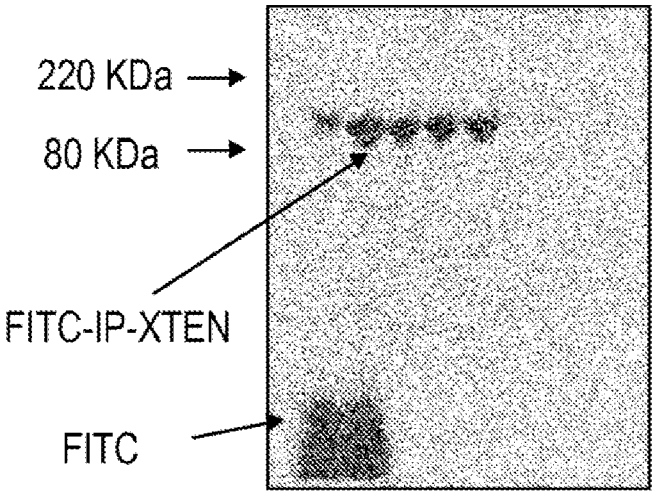


FIG. 42A

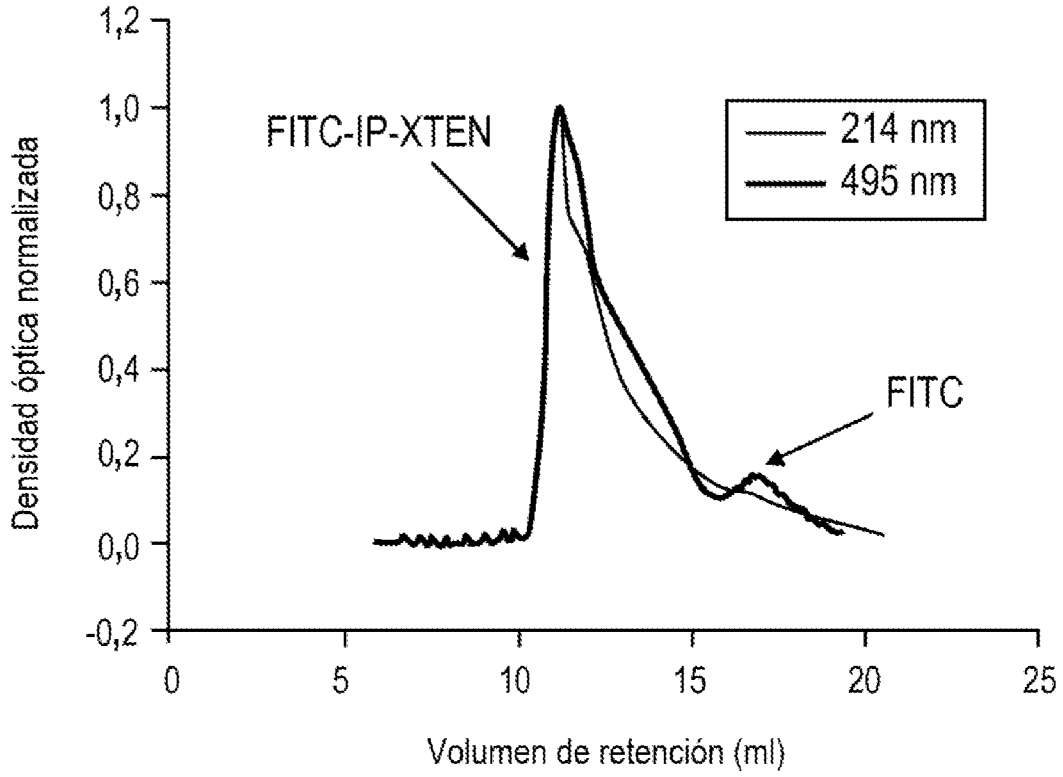


FIG. 42B

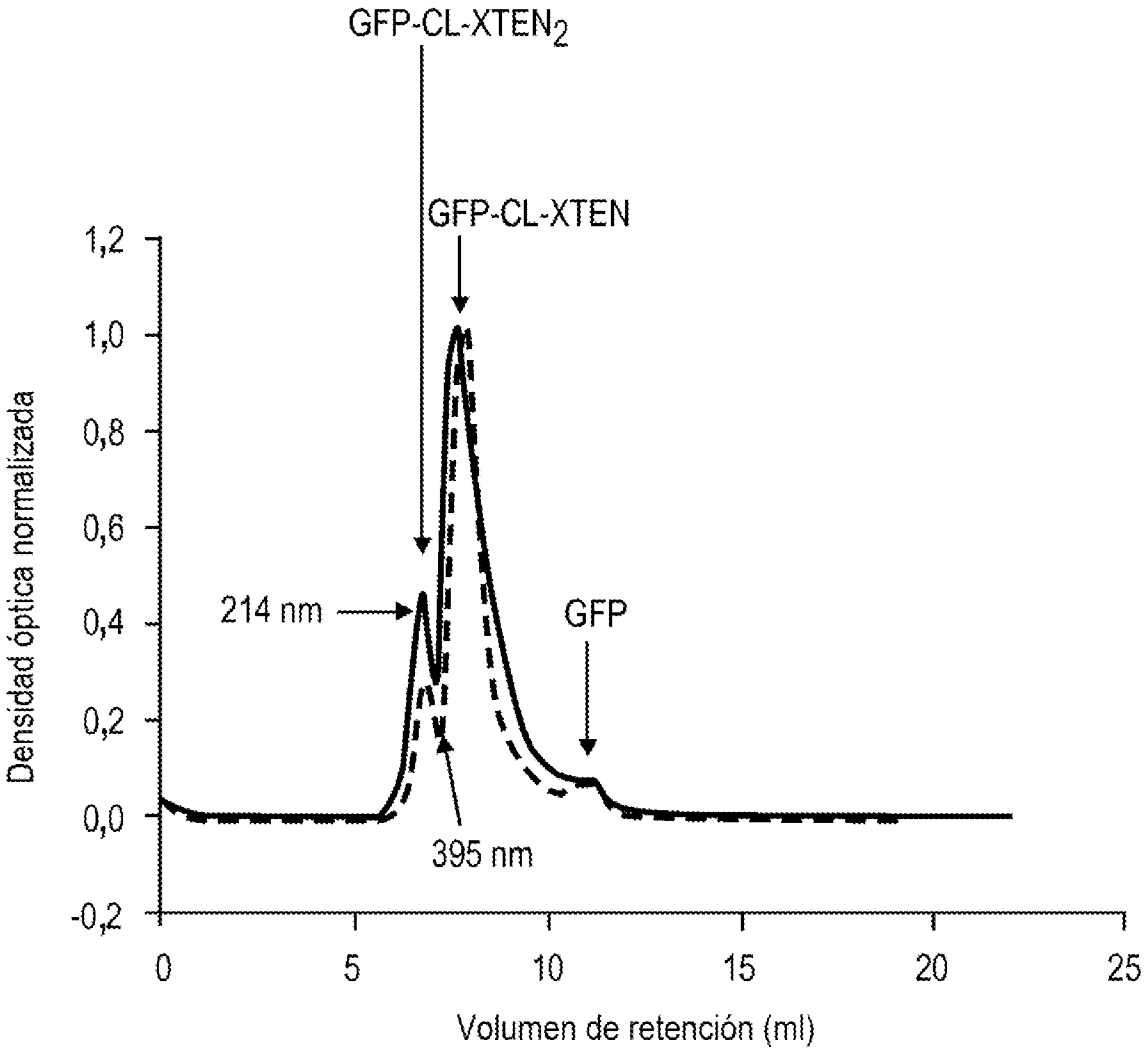

GFP-CL-XTEN2
GFP-CL-XTEN
214 nm
GFP
395 nm
Densidad óptica normalizada
1,2
1,0
0,8
0,6
0,4
0,2
0,0
-0,2
0
5
10
15
20
25
Volumen de retención (ml)

FIG. 43

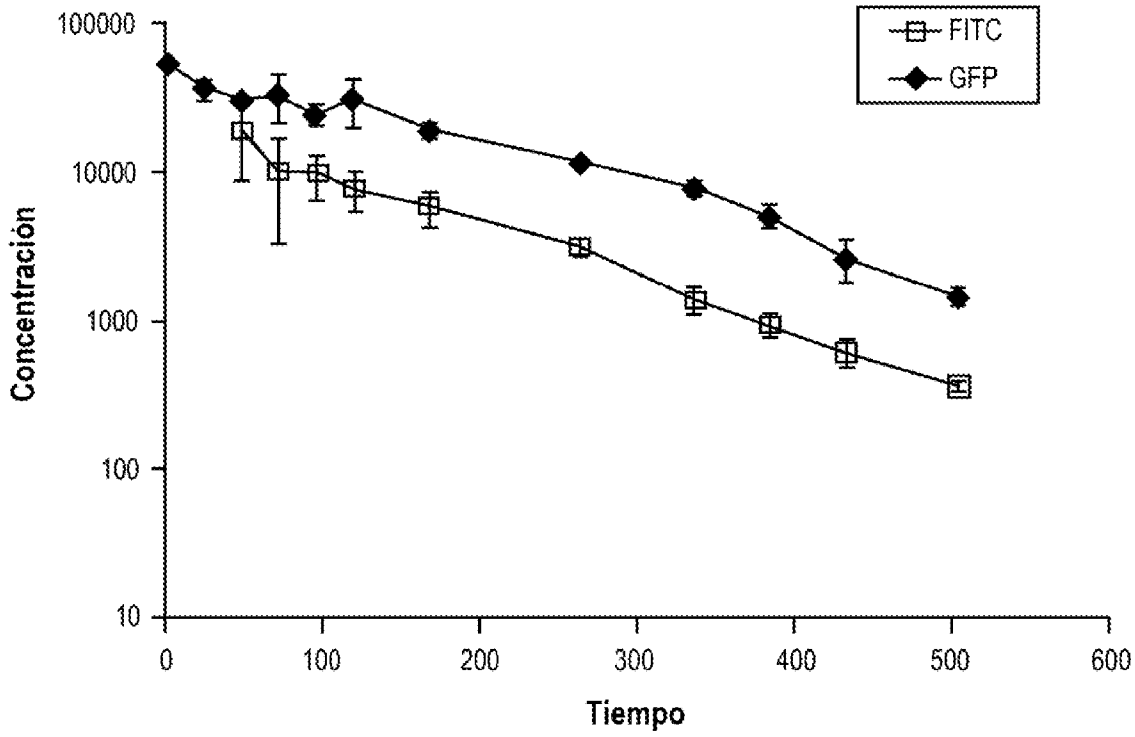

100000
10000
1000
100
10
Concentración
0
100
200
300
400
500
600
Tiempo
FITC
GFP

FIG. 44

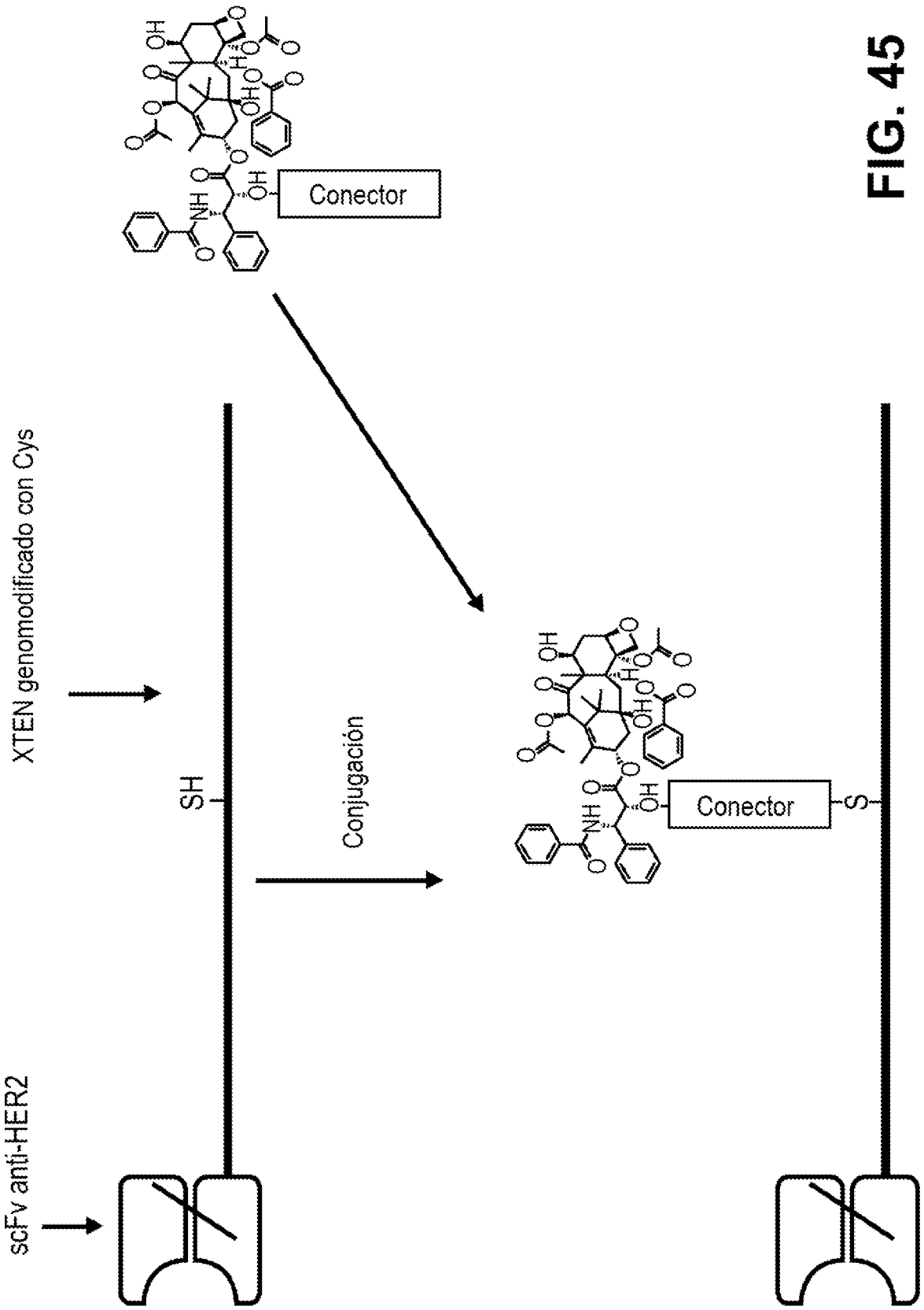

Conector
XTEN genomodificado con Cys
SH
Conjugación
scFv anti-HER2
Conector
S

FIG. 45