

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 814 558**

51 Int. Cl.:

**C07K 14/435** (2006.01)

**C12N 15/12** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **05.02.2014** **E 18184239 (4)**

97 Fecha y número de publicación de la concesión europea: **24.06.2020** **EP 3406629**

54 Título: **Proteínas de dominio de fibronectina tipo III con solubilidad mejorada**

30 Prioridad:

**06.02.2013 US 201361761452 P**

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

**29.03.2021**

73 Titular/es:

**BRISTOL-MYERS SQUIBB COMPANY (100.0%)**  
**Route 206 and Province Line Road**  
**Princeton, NJ 08543, US**

72 Inventor/es:

**YAMNIUK, AARON, P. y**  
**KRYSTEK, STANLEY, R.**

74 Agente/Representante:

**VALLEJO LÓPEZ, Juan Pedro**

**ES 2 814 558 T3**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

## DESCRIPCIÓN

Proteínas de dominio de fibronectina tipo III con solubilidad mejorada

5 La presente solicitud reivindica prioridad a la Solicitud Provisional de EE.UU. N.º 61/761.452.

## Antecedentes

10 La fibronectina es una proteína grande que desempeña papeles esenciales en la formación de la matriz extracelular y en las interacciones célula-célula; consiste en muchas repeticiones de tres tipos (tipos I, II y III) de pequeños dominios. Los dominios de fibronectina tipo III (Fn3) se hallan frecuentemente como partes de moléculas de adhesión celular, receptores de hormonas y citocinas de superficie celular, chaperonas y dominios de unión a hidratos de carbono. Un dominio de Fn3 de tipo silvestre es pequeño, monomérico, soluble y estable. Carece de enlaces disulfuro y, por lo tanto, es estable en condiciones reductoras. Para revisiones véase Bork y Doolittle, 1992, Proc Natl Acad Sci EE.UU. 89(19):8990-4; Bork et al., 1994, J Mol Biol. 242(4):309-20; Campbell y Spitzfaden, Structure 2(5):333-7 (1994); Harpez y Chothia, 1994, J Mol Biol. 238(4):528-39.

20 Los armazones basados en fibronectina son una familia de proteínas que tienen un plegamiento de tipo inmunoglobulina. Estas proteínas, que generalmente hacen uso de un armazón que deriva de una fibronectina tipo III (Fn3) o un dominio de tipo Fn3, funciona de una forma característica de los anticuerpos naturales o diseñados genéticamente (por ejemplo, anticuerpos policlonales, monoclonales o de cadena única) y contienen bucles que se localizan de forma análoga a las regiones determinantes de complementariedad (CDR) de un dominio variable del anticuerpo. Además, los armazones basados en fibronectina poseen ventajas estructurales. Específicamente, las estructuras de estos miméticos de anticuerpos se han optimizado con frecuencia para el plegamiento, estabilidad y solubilidad óptimas, incluso en condiciones que normalmente llevan a la pérdida de la estructura y la función en anticuerpos. Un ejemplo de proteínas de armazón basadas en fibronectina son las adnectinas (por Adnexus, una subsidiaria de propiedad total de Bristol Myers Squibb), que son una clase de elementos biológicos dirigidos que derivan del décimo dominio de tipo III (<sup>10</sup>Fn3) de la fibronectina humana. Los polipéptidos de adnectina de unión a albúmina de suero de tipo silvestre humano se describen en los documentos WO 2011140086 y GENESEQ:AZ63010. Se ha demostrado que las regiones de bucle de tipo CDR de los armazones basados en fibronectina se pueden modificar para dar lugar a una proteína capaz de unirse a una diana de interés. Por ejemplo, La Patente de los Estados Unidos N.º 7.115.396 describe proteínas del dominio Fn3 en las que las alteraciones de los bucles BC, DE y FG dan como resultado aglutinantes de TNF $\alpha$  de alta afinidad. La Patente de los Estados Unidos N.º 7.858.739 describe proteínas del dominio Fn3 en las que las alteraciones de los bucles BC, DE y FG dan como resultado aglutinantes de VEGFR2 de alta afinidad. La primera adnectina ensayada en un ensayo clínico, CT-322, se dirige al receptor 2 de factor de crecimiento endotelial vascular (VEGFR-2). Los estudios preclínicos (Mamluk, R. *et al.*, 2010, mAbs 2, 199-208) y de Fase I en CT-322 (Bloom, L. y Calabro, V., 2009, Drug Discov. Today 14, 949-955; Molckovsky, A. y Siu, L.L., 2008, J. Hematol. Oncol. 1, 20; Tolcher, A.W. *et al.*, 2011, Clin. Cancer Res. 17, 363-371) demostraron que se toleraba bien y que producía efectos farmacológicos esperados de la inhibición de la vía de VEGFR-2. Las proteínas de fusión recombinantes que comprenden un miembro proteico farmacológicamente activo y un miembro de dominio proteico farmacológicamente inactivo de origen humano, tal como un dominio <sup>10</sup>Fn3 se describen en el documento WO 2008/153745. Los armazones proteicos basados en una secuencia consenso de la repetición del dominio <sup>10</sup>Fn3 de la fibronectina humana se describen en el documento WO 2010/093627. El documento WO 2009/142773 se refiere a polipéptidos multivalentes que comprenden al menos dos dominios de armazón de fibronectina conectados a través de un enlazador de polipéptidos. Un estudio que explora los efectos de la elongación del bucle sobre la estabilidad del dominio <sup>10</sup>Fn3 se describe en Batori et al. (Protein Engineering 2002, 15(12): 1015-1020). William, M. *et al.*, 1994, J. Am. Chem. Soc. 116, 10797-10798 describe la síntesis química total de un dominio de <sup>10</sup>Fn3.

50 La agregación de proteínas puede ser un reto para los agentes terapéuticos basados en proteínas, tales como los armazones basados en fibronectina, en particular cuando se desean formulaciones de alta concentración de proteínas para la administración del fármaco. Además, la agregación puede suponer retos para la producción y la fabricación, llevar a actividades indeseables tales como agonismo no intencionado del receptor, y posiblemente afectar a la seguridad del fármaco por la inducción de una respuesta inmunológica. Véase Shire, S.J. *et al.*, 2004, J Pharm Sci. 93: 1390-402; Vazquez-Rey, M. y Lang, D.A., 2011, Biotechnol Bioeng 108: 1494-508; y Barker, M.P. *et al.*, 2010, Self Nonself 1: 314-322. Por consiguiente, es deseable generar y/o seleccionar agentes terapéuticos proteicos, tales como los armazones basados en fibronectina, que tienen una solubilidad potenciada o una propensión a la agregación reducida.

## Sumario

60 En un primer aspecto, la invención se refiere a polipéptidos que comprenden un resto de armazón basado en fibronectina que comprende un 10º dominio de fibronectina tipo III (<sup>10</sup>Fn3) modificado, en donde el dominio <sup>10</sup>Fn3 modificado comprende una secuencia de aminoácidos en la que el aminoácido correspondiente al resto 58 de la SEQ ID NO: 1 es un aminoácido hidrófilo que no es treonina, y en donde la solubilidad del dominio <sup>10</sup>Fn3 modificado se mejora en relación con la solubilidad de un dominio <sup>10</sup>Fn3 que comprende la misma secuencia de aminoácidos que el dominio <sup>10</sup>Fn3 modificado, excepto que el resto de aminoácido en la posición 58 de la SEQ ID NO: 1 es treonina, y en

donde el dominio  $^{10}\text{Fn3}$  modificado se une específicamente a una diana que no está unida por un dominio Fn3 humano seleccionado entre los establecidos en las SEQ ID NO:1-16, 65 y 66.

En un segundo aspecto, la invención se refiere a bibliotecas que comprenden una pluralidad de los polipéptidos del primer aspecto de la invención. También se desvelan en el presente documento métodos para identificar un polipéptido que se une a una diana que comprenden la exploración de una biblioteca descrita en el presente documento, y polipéptidos aislados identificados mediante tal método.

En un tercer aspecto, la invención se refiere a ácidos nucleicos que codifican los polipéptidos del primer aspecto de la invención.

### Breve descripción de las figuras

La **Figura 1** muestra una alineación de los dominios  $^7\text{Fn3}$ ,  $^{10}\text{Fn3}$  y  $^{14}\text{Fn3}$ . Las cadenas beta se muestran en negrita y subrayadas. Los restos de Thr que se corresponden con la T58 del  $^{10}\text{Fn3}$  humano de tipo silvestre están recuadrados.

La **Figura 2** muestra un ejemplo de curva de precipitación salina para 1 mg/ml de adnectina IGFR1. Los datos se ajustan a una función de curva sigmoidal a partir de la que se determina el punto medio de la transición de la precipitación salina ( $\text{AS}_m$ ).

La **Figura 3** muestra los datos de precipitación salina en sulfato de amonio para las adnectinas. A) Adnectinas EGFR1-7 y IGFR1. B) Adnectinas de EI-tándem. C) adnectinas EI-tándem PEGiladas (EI-tándem-PEG), en donde el recuadro muestra los datos con un eje x expandido. Los símbolos se corresponden con la identidad del dominio de adnectina anti-EGFR en cada molécula como sigue: EGFR1 (triángulo blanco  $\Delta$ ), EGFR2 (triángulo negro hacia arriba  $\blacktriangle$ ), EGFR3 (cuadrado blanco  $\square$ ), EGFR4 (cuadrado negro  $\blacksquare$ ), EGFR5 (círculo negro  $\bullet$ ), EGFR6 (triángulo negro hacia abajo  $\blacktriangledown$ ), EGFR7 (círculo blanco  $\circ$ ). IGFR1 se indica mediante (aspa x) en el panel A). D) Correlación entre los valores de  $\text{AS}_m$  para monoadnectinas y EI-tándems anti-EGFR (triángulos negros  $\blacktriangle$  con ajuste de línea sólida) o con adnectinas de EI-tándem PEGiladas (círculos negros  $\bullet$  con ajuste de línea punteada).

**Figura 4.** Efecto de la PEGilación sobre precipitación salina de un anticuerpo de dominio. A) Curva de precipitación salina para molécula de anticuerpo de dominio dAb1 (cuadrados negros  $\blacksquare$ ) o dAb1 PEGilado sobre un resto de cisteína libre con PEG lineal de 30 kDa (círculos blancos  $\circ$ ) o con PEG de ramificación única de 40 kDa (triángulos blancos hacia arriba  $\Delta$ ).

La **Figura 5** muestra datos de precipitación salina en sulfato de amonio para moléculas de fusión a Fc usando el método automático. A) La reproducibilidad del ensayo se demuestra con 3-4 experimentos replicados para Adn-Fc-A (círculo blanco  $\circ$ ), Adn-Fc-B (cuadrado blanco  $\square$ ) y Adn-Fc-K (triángulo blanco hacia arriba  $\Delta$ ). B) Datos de precipitación salina en sulfato de amonio para 2,1, 1,5, 0,9, 0,45, 0,3, 0,15 y 0,075 mg/ml de Adn-Fc-K tal como se detecta mediante A280 o mediante C) fluorescencia intrínseca. D) Dependencia logarítmica de valores de  $\text{AS}_m$  en la concentración de proteína ensayada, tal como se demuestra para Adn-Fc-K (círculo negro  $\bullet$ ), dAb3-1-IgG1\* (triángulo blanco hacia arriba  $\Delta$ ) y dAb5-4-IgG1\* (cuadrado blanco  $\square$ ).

**Figura 6.** Datos de solubilidad y de estabilidad para Adn-Fc. A) Cambio en las especies de alto peso molecular (% de APM) durante el tiempo para 1 mg/ml de moléculas de Adn-Fc tras la exposición a un estrés térmico de 30-37 °C, tal como se mide mediante SEC. B) % de APM para moléculas de Adn-Fc a elevadas concentraciones de proteína tal como se determina mediante SEC. C) Ejemplos de datos de precipitación salina para proteínas Adn-Fc. Los símbolos en los paneles A-C son tal como sigue: Adn-Fc-a (estrella blanca  $\star$ ), Adn-Fc-A (estrella negra  $\blackstar$ ), Adn-Fc-B (círculo blanco  $\circ$ ), Adn-Fc-b (círculo negro  $\bullet$ ), Adn-Fc-C (cuadrado blanco  $\square$ ), Adn-Fc-D (cuadrado negro  $\blacksquare$ ), Adn-Fc-E (diamante negro  $\blacklozenge$ ), Adn-Fc-F (aspa x), Adn-Fc-G (cruz +), Adn-Fc-H (triángulo blanco hacia abajo  $\blacktriangledown$ ), Adn-Fc-I (triángulo blanco hacia arriba  $\Delta$ ), Adn-Fc-J (triángulo negro hacia arriba  $\blacktriangle$ ), Adn-Fc-K (triángulo negro hacia abajo  $\blacktriangledown$ ), Adn-Fc-L (diamante blanco  $\lozenge$ ). D) Valores de  $\text{AS}_m$  para moléculas Adn-Fc, ordenados de izquierda a derecha por valores de  $\text{AS}_m$  en aumento a 0,45 mg/ml de [proteína]. Las barras de error representan la desviación estándar de 3-4 experimentos diferentes.

**Figura 7.** Solubilidad y estado oligomérico de moléculas dAb-Fc. C) Ejemplo de datos de precipitación salina para proteínas dAb-Fc. B) Dependencia de concentración de Rh para dAb-Fc seleccionados tal como se mide mediante DLS. Los símbolos en los paneles A y B son tal como sigue: dAb2-2-IgG1\*\* (triángulo blanco hacia abajo  $\blacktriangledown$ ), dAb2-3-IgG1\*\* (círculo negro  $\bullet$ ), dAb3-1-IgG1\* (triángulo blanco hacia arriba  $\Delta$ ), dAb5-4-IgG1\*-CHO (triángulo negro hacia abajo  $\blacktriangledown$ ), dAb6-2-IgG1\* (triángulo negro hacia arriba  $\blacktriangle$ ), dAb8-1-IgG1\* (diamante negro  $\blacklozenge$ ), dAb8-3-IgG1\* (círculo blanco  $\circ$ ), dAb8-2-IgG4 (cuadrado negro  $\blacksquare$ ), dAb9-1-IgG1\* (cuadrado blanco  $\square$ ). Las barras de error en el panel B representan la desviación estándar de 3-4 mediciones. C) Superposición de los valores de  $\text{AS}_m$  (triángulo negro hacia arriba  $\blacktriangle$ ) y de los valores de radio hidrodinámico ( $\circ$ ) para todas las moléculas dAb-Fc ensayadas.

**Figura 8.** % de APM para dAb-5-5-IgG1\* a elevadas concentraciones de proteína tal como se determina mediante SEC analítico.

Figura 9. Ejemplo de datos de precipitación salina para mAb con diferentes dominios Fc o diferentes dominios Fab.

Figura 10. Ejemplo de datos de precipitación salina de mutantes de adnectina anti-EGFR. Los símbolos representan los datos para EGFR8 de tipo silvestre,  $\text{AS}_m = 2,11 \text{ M}$  (círculo blanco  $\circ$ ), y los mutantes V54→A,  $\text{AS}_m = 2,25 \text{ M}$  (cuadrado blanco  $\square$ ), T58→E,  $\text{AS}_m = 2,17 \text{ M}$  (triángulo blanco hacia arriba  $\Delta$ ) y D77→A,  $\text{AS}_m = 1,96 \text{ M}$  (triángulo blanco hacia abajo  $\blacktriangledown$ ).

**Figura 11.** Solubilidad medida y predicha de mutantes de adnectina EGFR8. A) Solubilidad de proteínas EGFR8

de tipo silvestre y mutante en AS 2,0 M (barras en gris claro) o AS 2,2 M (barras en gris oscuro). Los datos de EGFR8 de tipo silvestre representan el promedio con barras de error de desviación estándar de tres mediciones independientes. B) Valores de propensión a la agregación espacial (PAE) para todos los mutantes de adnectina EGFR8, a R = 5 A (barras en gris claro) y R = 10 A (barras en gris oscuro).

5 **Figura 12.** Comparación del anticuerpo, el fragmento VH y la Adnectina que muestra el tamaño relativo de las moléculas. Diagramas de dibujo de (A) IgG1 (PDB ID 1IGT), (B) VHH (1F2X) con las CDR marcadas 1, 2, 3, (C) una Adnectina, con los bucles diversificados marcados BC, DE y FG y (D) superposición de una adnectina y VHH que muestra que las posiciones de los bucles BC, DE y FG son similares a las de las CDR 1, 2 y 3.

10 **Figura 13.** Unión de Adnectina 1 a EGFR. (A) El EGFR se representa como una superficie gris y se muestra unido a Adnectina 1 sobre el dominio I y la porción de Fv de cetuximab sobre el dominio III (PDB 1YY9; Li *et al.*, 2005). La adnectina y el cetuximab se muestran como dibujos con cadenas  $\beta$  (rojo) y estructura secundaria no repetitiva (cyan). (B) Restos de Adnectina 1 implicados en los contactos con EGFR. La estructura principal de la adnectina se muestra como un dibujo con el siguiente esquema de colores: cadenas  $\beta$  (rojo), estructura secundaria no repetitiva (naranja) y bucles diversificados (magenta). Los restos implicados en los contactos de los bucles diversificados se muestran con átomos de carbono magenta. Los restos implicados en los contactos del resto de la adnectina se muestran con átomos de carbono negros y regiones negras sobre el dibujo de la estructura secundaria. Nótese el tramo secuencial en la cadena D que se pone en contacto con EGFR. (C) Interacción entre adnectina y el dominio I de EGFR. La adnectina está en magenta y el dibujo de EGFR está en cyan. La superficie de interacción en EGFR muestra la definición más estricta de restos de contacto en naranja y la definición relajada de restos internos en amarillo. (D) Se muestra la superposición de las superficies de contacto de adnectina 1 y EGFR sobre el dominio I de EGFR. La adnectina 1 (Cyan) y el EGF (naranja) se representan como dibujos. Las únicas superficies de contacto y la superficie de solapamiento se muestran como Adnectina 1 (amarillo), EGF (rojo) y para ambos (magenta). (E) Interacciones de tipo cadena  $\beta$  entre los restos 15-18 del EGFR y los restos 76-79 de la adnectina 1 con los enlaces de hidrógeno de N... O=C formados entre Q16 N y D77 O, K79 N y Q16 O, y L17 N y K79 O, y un enlace de hidrógeno de cadena lateral entre T16 OG1 y D77 OD1.

25 **Figura 14.** Unión de Adnectina 2 a IL-23. (A) La IL-23 se representa como una superficie gris y se muestra unida a la adnectina 2 en la superficie de contacto entre las subunidades p40 y p19 y la porción Fv de 7G10 se muestra unida a la subunidad p19 (PDB 3D85; Beyer *et al.*, 2008). La adnectina y 7G10 se muestran como dibujos con el mismo código de color que la Figura 2A. (B) Restos de adnectina 2 implicados en los contactos con IL-23. Las adnectinas 1 (Figura 2B) y 2 se orientan de forma idéntica para permitir la comparación de las diferentes formas mediante inspección. El código de color es el mismo que en la Figura 2B. En la adnectina 2 las siguientes regiones hacen contacto con IL-23: la región N-terminal, la cadena C, el bucle CD, la cadena E y la cadena F. (C) Interacción entre la adnectina la IL-23 con el código de color de la superficie y la adnectina siendo igual que en la Figura 2C. El dominio p40 (cadena A) se muestra en un cyan más claro y el dominio p19 (cadena B) se muestra en un cyan más oscuro. (D) Una vista de la interacción Adnectina 2/IL-23 que implica solo los bucles diversificados y la región N-terminal. El mismo código de color que en la parte (C). (E) Una vista que muestra solo los restos 76-85 del bucle FG de adnectina 2 unida a IL-23.

30 **Figura 15.** Comparación de adnectina 1 y adnectina 2 con  $^{10}\text{Fn3}$ . (A) Secuencias de aminoácidos de las dos adnectinas frente al dominio parental  $^{10}\text{Fn3}$ . Las partes de los bucles BC, DE y FG que se diversificaron se subrayan en la figura, abarcando los restos 23-29, 52-55 y 77-86, respectivamente. (B, C) Dos vistas ortogonales superpuestas en los restos 1416 a 1509 de PDB 1FNF. Código de colores: 1FNF (azul), Adnectina 1 (rojo) y Adnectina 2 (cyan). Nótese la excelente superposición de las cadenas  $\beta$  del núcleo y los bucles AB y EF. El bucle DE, que es bastante corto, muestra una pequeña variación en estas estructuras. El bucle BC muestra una variación modesta. Por el contrario, el bucle FG muestra una variación importante en la posición incluso entre los bucles de igual longitud de Adnectina 2 y  $^{10}\text{Fn3}$ . En la adnectina 1, las cadenas  $\beta$  F y G se extienden más allá en la región diversificada; sin embargo, para resaltar la región diversificada, esos residuos se dibujaron como una estructura secundaria no repetitiva. Aunque los extremos N-terminales de  $^{10}\text{Fn3}$  y de adnectina 1 son similares, el de la adnectina 2 difiere de manera considerable.

45 **Figura 16.** Vistas en estereo de la densidad de electrones inicial con el modelo final de los bucles diversificados, que no se incluyeron en el modelo inicial, de (A) de adnectina 1 y (B) de adnectina 2.

50 **Figura 17.** Vistas en estereo de las partes C y D de la Figura 13. Ambas figuras rotadas 90° en z a partir de su representación en la Figura 13 para acomodar una distancia interocular de aproximadamente 60 mm para la visión en estereo. C. Interacción entre adnectina 1 y el dominio I de EGFR. Los dibujos de adnectina están en magenta y el EGFR está en cyan. La superficie de interacción en EGFR muestra los restos de contacto en naranja y los restos internos en amarillo. D. Se muestra la superposición de las superficies de contacto de adnectina 1 y EGFR sobre el dominio I de EGFR. La adnectina 1 (Cyan) y el EGF (azul) se representan como dibujos. Se muestran las superficies de contacto de Adnectina 1 (amarillo), EGF (rojo) y la superficie de solapamiento (magenta).

55 **Figura 18.** Visiones en estereo de las partes C, D y E de la Figura 14. Las Figuras C y D se rotan 90° en z a partir de su representación en la Figura 14 para acomodar una distancia interocular de aproximadamente 60 mm para la visión en estereo. C. Interacción entre Adnectina y IL-23. La adnectina se muestra como un dibujo (magenta). Los dibujos del dominio p40 (cadena A) se muestran en un cyan más claro y el dominio p19 (cadena B) se muestra en un cyan más oscuro. La superficie de interacción en IL-23 muestra los restos de contacto en naranja y los restos internos en amarillo. D. Una vista de la interacción Adnectina 2/IL-23 que implica solo los bucles diversificados y la región N-terminal. El mismo código de color que en la parte (C). E. Una vista de la interacción Adnectina 2/IL-23 que implica solo el bucle FG. El mismo código de color que en la parte (C).

65 **Figura 19.** Vistas en estereo de la superposición de variantes basadas en  $^{10}\text{Fn3}$ , en relación con la Figura 15,

partes B y C, pero que muestran variantes adicionales basadas en <sup>10</sup>F<sub>n3</sub>. B. Visión en estéreo de la superposición de los dominios basados en <sup>10</sup>F<sub>n3</sub> de 2OCF (naranja), 3CSB (rosa), 3CSG (frambuesa), 3K2M (cadena C) (verde), 3QHT (cadena B) (pizarra), 3QWQ (rojo) y 3QWR (cyan) en el dominio 10 1FNF (azul). Esta vista se rota 90° en z en comparación con la Figura 15B para acomodar una distancia interocular de aproximadamente 60 mm. C. Vista en estéreo de la superposición de los dominios basados en <sup>10</sup>F<sub>n3</sub> de 2OCF (naranja), 3CSB (rosa), 3CSG (frambuesa), 3K2M (cadena C) (verde), 3QHT (cadena B) (pizarra), 3QWQ (rojo) y 3QWR (cyan) en el dominio 10 1FNF (azul). Esta vista está en la misma orientación que la Figura 15C, pero en comparación con A (anteriormente), está rotada 90° en z y después 90° en y. Estas figuras resaltan que el núcleo de los dominios basados en <sup>10</sup>F<sub>n3</sub> se mantiene, pero los bucles diversificados pueden variar en conformación para adaptar la unión a la diana.

**Figura 20.** Estado oligomérico y estabilidad térmica de las adnectinas EGFR8. A) Cromatogramas de SEC para EGFR8 (carga de 75 ug), EGFR8-T58E (carga de 73 ug) y EGFR8-T58D (carga de 52 ug). B) Datos del termograma de DSC para muestras de 1 mg/ml de adnectinas EGFR8 en PBS a pH 7,1. C) Ajuste de datos de DSC para EGFR8. D) Ajuste de datos de DSC para EGFR8-T58E. E) Ajuste de datos de DSC para EGFR8-T58D.

**Figura 21.** Curvas de precipitación salina de sulfato de amonio para proteínas EGFR8 ensayadas a 0,33 mg/ml. Los símbolos representan los datos para EGFR8 de tipo silvestre, AS<sub>m</sub> = 2,08 M (círculo blanco o), EGFR8-T58E, AS<sub>m</sub> = 2,15 M (cuadrado blanco □), y EGFR8-T58D, AS<sub>m</sub> = 2,16 M (triángulo blanco hacia arriba Δ).

**Figura 22.** Propensión a la agregación de las adnectinas EGFR8. A) Relación entre la concentración de proteína medida frente a la concentración de proteína esperada (basándose en estimaciones visuales del volumen) para adnectinas concentradas a pequeña escala usando ultrafiltración. B) Porcentaje de especies de alto peso molecular (% de APM) en la fracción soluble de las muestras de adnectina concentrada tal como se mide mediante SEC.

**Figura 23.** Datos de estabilidad acelerada para adnectinas EGFR8 en PBS a pH 7,1. A) Concentración de proteína adnectina en el tiempo cero (t<sub>0</sub>) y después de 1 semana (1sem) o 2 semanas (2sem) a 40 °C, medida por A<sub>280</sub>. B) Cambios en la concentración de proteína soluble medida por la integración de los picos de SEC. C) Porcentaje de especies de alto peso molecular (% de APM) en la fracción soluble tal como se mide mediante SEC tras 1 semana (1sem) o 2 semanas (2sem) a 40 °C.

**Figura 24.** Alineamiento de secuencia de aminoácidos de EGFR4 y EGFR8 con las secuencias de bucle de unión a diana "BC", "DE" y "FG" indicadas.

**Figura 25.** Estado oligomérico y estabilidad térmica de adnectinas EGFR4. A) Cromatogramas de SEC para EGFR4 (carga de 50 ug) o EGFR4-T58E (carga de 50 ug). B) Datos del termograma de DSC para muestras de 1 mg/ml de adnectinas EGFR4 en PBS a pH 7,1. C) Ajuste de datos de DSC para EGFR4. D) Ajuste de datos de DSC para EGFR4-T58E.

**Figura 26.** Curvas de precipitación salina de sulfato de amonio para adnectinas EGFR4 ensayadas a 0,3 mg/ml. Los símbolos representan los datos para EGFR4 de tipo silvestre, AS<sub>m</sub> = 0,865 ± 0,014 M (círculo blanco o), EGFR4-T58E, AS<sub>m</sub> = 1,024 ± 0,003 (cuadrado blanco □).

**Figura 27.** Propensión a la agregación de las adnectinas EGFR4. A) Relación entre la concentración de proteína medida frente a la concentración de proteína esperada (basándose en estimaciones visuales del volumen) para adnectinas concentradas a pequeña escala usando ultrafiltración. B) Porcentaje de especies de alto peso molecular (% de APM) en la fracción soluble de las muestras de adnectina concentrada incubada toda la noche a temperatura ambiente tal como se mide mediante SEC.

## Descripción detallada

### Definiciones

El término "polipéptido" se refiere a cualquier secuencia de dos o más aminoácidos, independientemente de la longitud, de la modificación postraducciona l o de la función. Los polipéptidos pueden incluir aminoácidos naturales y aminoácidos no naturales tales como los descritos en la Patente de Estados Unidos N.º 6.559.126. Los polipéptidos también se pueden modificar mediante cualquiera de varias formas químicas habituales (por ejemplo, se puede modificar un aminoácido con un grupo protector; el aminoácido de carboxilo terminal se puede convertir en un grupo amida terminal; el resto de amino terminal se puede modificar con grupos para, por ejemplo, mejorar la lipofilia; o el polipéptido se puede glucosilar químicamente o modificar de otro modo para aumentar la estabilidad o la semivida *in vivo*). Las modificaciones de polipéptidos pueden incluir la unión de otra estructura tal como un compuesto cíclico u otra molécula al polipéptido y también puede incluir polipéptidos que contienen uno o más aminoácidos en una configuración alterada (es decir, R o S; o, L o D).

Tal como se usa en el presente documento, un "armazón basado en fibronectina" o proteína o resto de "FBS" (siglas del inglés *fibronectin based scaffold*) se refiere a proteínas o restos que se basan en un dominio de fibronectina, por ejemplo, la repetición de fibronectina tipo III ("<sup>10</sup>F<sub>n3</sub>"). Una proteína o un dominio (de una proteína) que se basa en una repetición de F<sub>n3</sub> se cita en el presente documento como "proteína F<sub>n3</sub>", "proteína de dominio de F<sub>n3</sub>" o "dominio de F<sub>n3</sub>". F<sub>n3</sub> es un pequeño dominio (de aproximadamente 10 kDa) que tiene la estructura de un plegamiento de inmunoglobulina (Ig) (es decir, una estructura de sándwich β de tipo Ig, que consiste en siete cadenas β y seis bucles). La fibronectina tiene 18 repeticiones de F<sub>n3</sub>, y aunque la homología de secuencia entre las repeticiones es baja, todas comparten una elevada similitud en la estructura terciaria. Los dominios F<sub>n3</sub> también están presentes en muchas proteínas que no son la fibronectina, tales como moléculas de adhesión, moléculas de superficie celular, por ejemplo, receptores de citocina y dominios de unión a hidratos de carbono. La expresión proteína o resto de "armazón basado en fibronectina" o "proteína F<sub>n3</sub>" o "dominio F<sub>n3</sub>" o "proteína de dominio F<sub>n3</sub>" pretende incluir proteínas o dominios

basados en dominios Fn3 a partir de estas otras proteínas (es decir, no fibronectina). Los dominios Fn3 a modo de ejemplo incluyen el 7<sup>o</sup>, 10<sup>o</sup> y 14<sup>o</sup> dominio de fibronectina tipo III, a los que se hace referencia en el presente documento como 7Fn3, 10Fn3 y 14Fn3, respectivamente. Tal como se usa en el presente documento, un "dominio Fn3" o "resto de Fn3" o "proteína de dominio Fn3" se refiere a la Fn3 de tipo silvestre (por ejemplo, las SEQ ID NO: 1-8, 10, 12, 14 o 16, 65 o 66) y variantes biológicamente activas de la misma, por ejemplo, las variantes biológicamente activas que se pueden unir específicamente a una diana, tal como EGFR, IL23 y IGF1R. Por ejemplo, las moléculas de 10Fn3 que se unen a dianas específicas se pueden seleccionar a partir de bibliotecas de 10Fn3 usando la técnica PROfusion, que se describe, por ejemplo, en el documento WO02/32925. Un dominio 10Fn3 de tipo silvestre puede comprender una de las secuencias de aminoácidos establecidas en las SEQ ID NO: 1-8, 10, 12, 14, 16. La secuencia de aminoácidos de 10Fn3 humano de tipo silvestre se establece en la SEQ ID NO: 1. Un dominio 7Fn3 de tipo silvestre puede comprender la secuencia de aminoácidos establecida en la SEQ ID NO: 65. Un dominio 14Fn3 de tipo silvestre puede comprender la secuencia de aminoácidos establecida en la SEQ ID NO: 66. Las variantes biológicamente activas de un dominio Fn3 incluyen dominios Fn3 que comprenden al menos, como máximo o aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, o 40 variaciones de aminoácidos, es decir, sustituciones, adiciones o deleciones, en relación con un dominio Fn3 que comprende una secuencia de aminoácidos seleccionada de las SEQ ID NO:1-16 y 65-71. Una variante biológicamente activa de un dominio Fn3 también puede comprender, o comprender como máximo, 1-3, 1-5, 1-10, 1-15, 1-20, 1-25, 1-30, o 1-40 cambios de aminoácidos en relación con un dominio Fn3 que comprende una secuencia de aminoácidos seleccionada de las SEQ ID NO: 1-16 y 65-71. En determinadas realizaciones, una variante biológicamente activa de un dominio Fn3 no comprende más de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, o 40 variaciones de aminoácidos, es decir, sustituciones, adiciones o deleciones, en relación con un dominio Fn3 que comprende una secuencia de aminoácidos seleccionada de las SEQ ID NO:1-16 y 65-71. Un(os) cambio(s) de aminoácidos puede(n) estar en una región de bucle y/o en una región no de bucle, por ejemplo, una cadena  $\beta$ . Las secuencias a modo de ejemplo de aminoácidos degenerados de 10Fn3 se proporcionan en el presente documento como las SEQ ID NO: 17-29.

La frase "que comprende una secuencia de aminoácidos basada en" una secuencia específica o primera secuencia pretende incluir secuencias de aminoácidos que provienen de la secuencia específica o de la primera secuencia, por ejemplo, mediante sustituciones, adiciones o deleciones de aminoácidos. Por ejemplo, una proteína que comprende una secuencia de aminoácidos basada en una secuencia de aminoácidos seleccionada de las SEQ ID NO:1-29 y 65-66 se refiere a una proteína que comprende una secuencia de aminoácidos que proviene de cualquiera de las SEQ ID NO: 1-29 y 65-66, incluyendo, por ejemplo, una proteína que comprende una secuencia de aminoácidos que difiere de una o más de las SEQ ID NO: 1-29 y 65-66 en una o más secuencias de bucle o no de bucle, tal como para obtener la unión a una diana deseada.

Una "región" de un dominio Fn3 (o resto) tal como se usa en el presente documento se refiere o bien a un bucle (AB, BC, CD, DE, EF y FG), una cadena  $\beta$  (A, B, C, D, E, F y G), el extremo N-terminal (por ejemplo, los restos de aminoácidos 1-7 de la SEQ ID NO: 1 o 6), o el extremo C-terminal (por ejemplo, los restos de aminoácidos 93-101 de la SEQ ID NO: 1 o los restos de los aminoácidos 93-95 de la SEQ ID NO: 6) de un dominio Fn3.

Un dominio Fn3 puede comprender, en orden desde el extremo N-terminal al extremo C-terminal, una cadena beta o de tipo beta, A; un bucle, AB; una cadena beta o de tipo beta, B; un bucle, BC; una cadena beta o de tipo beta, C; un bucle, CD; una cadena beta o de tipo beta, D; un bucle, DE; una cadena beta o de tipo beta, E; un bucle, EF; una cadena beta o de tipo beta, F; un bucle, FG; y una cadena beta o de tipo beta, G. Las siete cadenas  $\beta$  antiparalelas se disponen como dos láminas beta que forman un núcleo estable, mientras que crean dos "caras" compuestas de los bucles que conectan las cadenas beta o de tipo beta. Los bucles AB, CD y EF se localizan en una cara ("el polo sur") y los bucles BC, DE y FG se localizan en la cara opuesta ("el polo norte"). Cualquiera o todos los bucles AB, BC, CD, DE, EF y FG puede participar en o contribuir a la unión del ligando.

El término "región no de bucle" de un dominio Fn3 se refiere a una cadena  $\beta$ , el extremo N-terminal o el extremo C-terminal de un dominio Fn3. Una región no de bucle de un dominio Fn3 también se puede citar como una "región de armazón".

Un "bucle del polo norte" se refiere a cualquiera de los bucles BC, DE y FG de un dominio Fn3.

Un "bucle del polo sur" se refiere a cualquiera de los bucles AB, CD y EF de un dominio Fn3.

"Porcentaje (%) de identidad de secuencia de aminoácidos" se define en el presente documento como el porcentaje de restos de aminoácidos en una secuencia candidata que son idénticos a los restos de aminoácidos en una secuencia seleccionada, después de alinear las secuencias e introducir huecos, en caso necesario, para lograr el porcentaje máximo de identidad de secuencia y no considerar cualquier sustitución conservativa como parte de la identidad de secuencia. El alineamiento con el fin de determinar el porcentaje de identidad de secuencia de aminoácidos puede lograrse de varias maneras que se encuentran dentro de las capacidades de la técnica, por ejemplo, usando programas informáticos disponibles de manera pública, tales como los programas BLAST, BLAST-2, ALIGN, ALIGN-2 o Megalign (DNASTAR). Los expertos en la materia pueden determinar los parámetros adecuados para medir la alineación, incluyendo cualquier algoritmo necesario para conseguir la alineación máxima sobre toda la longitud de las

secuencias que se comparan.

A efectos del presente documento, el % de identidad de secuencia de aminoácidos de una secuencia de aminoácidos dada A para, con o contra una secuencia de aminoácidos dada B (que puede, como alternativa, citarse como una secuencia de aminoácidos A que tiene o comprende un % determinado de identidad de secuencia para, con o frente a una secuencia de aminoácidos B) se calcula del modo siguiente: 100 veces la parte X/Y donde X es el número de restos de aminoácidos registrados como coincidencias idénticas mediante un programa de alineamiento de secuencias, tal como BLAST, BLAST-2, ALIGN, ALIGN-2 o Megalign (DNASTAR), en esa alineación de A y B en el programa y donde Y es el número total de restos de aminoácidos en B. Se apreciará que cuando la longitud de la secuencia de aminoácidos A no es igual a la longitud de la secuencia de aminoácidos B, el % de identidad de secuencia de aminoácidos de A a B no será igual al % de identidad de secuencia de aminoácidos de B a A.

Tal como se usa en el presente documento, se considera que un resto de aminoácido en un polipéptido "contribuye a la unión" a una diana si (1) se descubre que cualquiera de los átomos que no sean de hidrógeno de la cadena lateral o de la cadena principal del resto está a cinco ángstroms de cualquier átomo de la diana de unión basándose en una estructura tridimensional del complejo determinada de forma experimental, y/o (2) la mutación del resto a su equivalente en Fn3 de tipo silvestre (por ejemplo, SEQ ID NO: 1), a alanina, o a un resto que tiene una cadena lateral de tamaño similar o más pequeña que el resto en cuestión, lleva a un aumento medido de la constante de disociación de equilibrio para la diana (por ejemplo, un aumento en la  $k_{on}$ ).

Tal como se usa en el presente documento, "Fc" abarca dominios que provienen de la región constante de una inmunoglobulina, preferentemente una inmunoglobulina humana, que incluye un fragmento, un análogo, una variante, un mutante o derivado de la región constante. Las inmunoglobulinas incluyen IgG1, IgG2, IgG3, IgG4 y otras clases tales como IgA, IgD, IgE e IgM. La región constante de una inmunoglobulina se define como un polipéptido de origen natural o producido de forma sintética homólogo a la región C-terminal de inmunoglobulina, y puede incluir un dominio CH1, una bisagra, un dominio CH2, un dominio CH3, o un dominio CH4, por separado o en combinación. El término "resto Fc" o "dominio Fc" tal como se usa en el presente documento se refiere a cualquier combinación de los dominios CH1, bisagra, CH2, CH3 y CH4. Por lo tanto, un "dominio Fc" o resto puede o no comprender una bisagra.

"Resto" se refiere a una parte de una proteína. Por ejemplo, una proteína de fusión puede comprender varios restos. En una realización, una proteína de fusión comprende un resto de armazón basado en fibronectina (Fn) y un resto de Fc.

El término "solubilidad mejorada" puede significar mayor proporción (por ejemplo, concentración, molalidad, fracción molar, o proporción molar) de un soluto de proteína en un disolvente designado, en condiciones de solución saturada, en comparación con otra proteína. Esto también se puede expresar como un mayor límite de solubilidad de una proteína. "Solubilidad mejorada" también puede referirse a una reducida agregación de proteínas, o a una menor proporción (por ejemplo, porcentaje, parte o relación) de especie de proteína agregada a una concentración de proteína dada en un disolvente designado, en comparación con otra proteína. Esto se puede expresar de forma alternativa como una reducida propensión de una proteína a la agregación. El término "agregación" se puede referir a la interacción o a la asociación entre dos o más moléculas de proteína, por ejemplo, proteínas en su conformación natural o moléculas de proteína mal plegadas. Las moléculas de interacción o de asociación se pueden citar como, por ejemplo, un "agregado", un "oligómero" (tal como, un dímero, un trímero, un tetramero u oligómeros de orden superior), una "especie de alto peso molecular", una "especie de orden superior", u otros términos relacionados que comúnmente entiende un experto en la materia. Se puede analizar la solubilidad de una proteína mediante cualquiera de las técnicas adecuadas. En algunas realizaciones, se puede analizar la solubilidad de una proteína mediante el ensayo de solubilidad de sulfato de amonio. En algunas realizaciones, se puede analizar la solubilidad de una proteína mediante ultrafiltración. En algunas realizaciones, se puede analizar la solubilidad de una proteína en condiciones de estrés acelerado.

En algunas realizaciones, la solubilidad de una proteína se puede considerar mejorada si la solubilidad de una proteína en un disolvente mejora en al menos el 1 %, 2 %, 5 %, 10 %, 15 %, 20 %, 30 % o el 50 % o más en comparación con la solubilidad de otra proteína en el disolvente. En algunas realizaciones, la solubilidad de una proteína se puede considerar mejorada si el límite de solubilidad de una proteína en un disolvente mejora en al menos 0,01 mg/ml, 0,1 mg/ml, 0,5 mg/ml, 1 mg/ml, 2 mg/ml, 5 mg/ml, 10 mg/ml, 20 mg/ml, 50 mg/ml, 100 mg/ml, 200 mg/ml o 500 mg/ml mayor que el límite de solubilidad de otra proteína en el disolvente. En algunas realizaciones, la solubilidad de una proteína se puede considerar mejorada si el porcentaje de especie de proteína agregada a una concentración dada es menor en al menos el 0,5 %, 1 %, 2 %, 5 %, 10 %, 20 %, el 50 % o más en comparación con el porcentaje de especie de proteína agregada de otra proteína en la concentración.

### Descripción general de la solubilidad de la proteína y ensayos

La agregación de proteínas es un complejo proceso que se puede originar mediante varios mecanismos diferentes (Morris, A.M. *et al.*, 2009, *Biochim Biophys Acta*. 1794, 375-97; Joubert, M.K. *et al.*, 2011, *J Biol Chem* 286, 25118-33; Weiss, W.F. *et al.*, 2009, *J Pharm Sci*. 98, 1246-77; Roberts, C.J. 2007, *Biotechnol Bioeng*. 98, 927-38; Hawe, A. *et al.*, 2009, *Eur J Pharm Sci*. 38, 79-87; Mahler, H.C. *et al.*, 2009, *J Pharm Sci*. 98, 2909-34; Philo, J.S. Y Arakawa, T.,

2009, *Curr Pharm Biotechnol.* 10, 348-51). Se pueden formar agregados, por ejemplo, debido a la autoasociación de la conformación natural, o estados estructuralmente alterados como un glóbulo fundido, estructuras desnaturalizadas, degradadas o químicamente modificadas. La agregación o solubilidad de la proteína también puede estar influenciada por las condiciones de la solución, la presencia de contaminantes, así como por varios factores externos tales como la temperatura, el estrés mecánico, o el estrés de congelación/descongelación. Además, los tipos de agregado pueden ser covalentes o no covalentes, y pueden diferir en la morfología, variando desde dímeros solubles u oligómeros de orden superior, a precipitados amorfos insolubles, o estructuras amiloides más regulares.

Se puede usar una serie de técnicas para analizar la solubilidad de las proteínas y para detectar los agregados de diferentes tamaños y morfologías, así como para caracterizar los mecanismos de la agregación (Wang, W. 2005, *Int J Pharm* 289, 1-30; Singh, S.K. *et al.*, 2010, *J Pharm Sci.* 99, 3302-21; Kamerzell, T.J. *et al.*, 2011, *Adv Drug Deliv Rev.* 63, 1118-59; Zolls, S. *et al.*, 2012, *J Pharm Sci.* 101, 914-35). Las técnicas basadas en los principios de dispersión de la luz, espectroscopía, electroforesis, cromatografía y muchos otros se pueden usar para caracterizar la agregación de proteínas, la estabilidad y la solubilidad o para comparar la propensión a la agregación o la solubilidad de diferentes moléculas de proteína entre sí, o una molécula en diferentes condiciones de solución. Las proteínas se pueden concentrar antes del análisis usando métodos tales como ultrafiltración o liofilización. O se pueden usar métodos de menor escala para concentrar las proteínas para análisis de solubilidad y agregación, que incluyen el uso de presión osmótica, congelación/descongelación, y evaporación del disolvente (Shire, S.J. *et al.*, 2004, *J Pharm Sci.* 93, 1390-402). La agregación también se puede estudiar a concentraciones de proteína más bajas en condiciones de estrés acelerado tales como elevada temperatura, tensión de cizalla o estrés de mezclado (Hawe, A. *et al.*, 2012, *J Pharm Sci.* 101, 895-913), utilizando ensayos de unión de colorante fluorescente de alto rendimiento (Kayser, V. *et al.*, 2012, *Biotechnol J.* 7, 127-32; He, F. *et al.*, 2010, *J Pharm Sci.* 99, 1707-20), o incluso predecir usando métodos de simulación computacional (Cafisch, A., 2006, *Curr Opin Chem Biol.* 10, 437-44; Fernandez-Escamilla, A.M. *et al.*, 2004, *Nat Biotechnol.* 22, 1302-6; Cellmer, T. *et al.*, 2007, *Trends Biotechnol.* 25, 254-61).

La propensión a la agregación de las proteínas también se puede analizar basándose en su solubilidad relativa en presencia de uno o más agentes cosmotrópicos, por ejemplo, sulfato de amonio (AS, de *ammonium sulfate*) u otros cosmotropos tales como fosfato ( $\text{PO}_4^{3-}$ ) o fluoruro ( $\text{F}^-$ ), o en presencia de polímeros de exclusión por volumen tales como polietilenglicol (PEG). El mecanismo de la precipitación de proteínas inducido por AS o "precipitación salina", puede implicar la unión fuerte de moléculas de agua mediante el anión sulfato cosmotrópico polar, que deshidrata las superficies de la proteína, aumenta el potencial químico de la proteína y provoca que las moléculas de proteína se agreguen en un precipitado amorfo (Baldwin, R.L., 1996, *Biophys J.* 71, 2056-63). Debido a que las superficies hidrofóbicas de la proteína están preferentemente deshidratadas sobre las superficies polares, la autoasociación de proteínas inducida por AS se puede dirigir mediante la interacción de las superficies hidrofobas expuestas, similar a las fuerzas que dirigen la agregación en ausencia de AS (Young, L. *et al.*, 1994, *Protein Sci.* 3, 717-29; Arunachalam, J. y Gautham, N., 2008, *Proteins.* 71, 2012-25; Fink, A.L., 1998, *Fold Des.* 3, R9-23). La teoría de la determinación de la solubilidad relativa de la proteína usando AS se revisa por Trevino *et al.* (2008, *J Pharm Sci.* 97, 4155-66), quienes también describen su aplicación en la determinación de la solubilidad relativa de una serie de variantes de ribonucleasa (Trevino, S.R. *et al.*, 2007, *J Mol Biol.* 366, 449-60). En este documento, se describen un método de escala de banco manual así como un método automatizado basado en la valoración en torno a este concepto, y se pueden usar para determinar la solubilidad relativa de la proteína.

### Dominios modificados de fibronectina tipo III (Fn3) con solubilidad mejorada

En el presente documento se desvelan polipéptidos que comprenden un dominio Fn3 modificado, en donde el dominio Fn3 modificado comprende una modificación de aminoácidos que mejora la solubilidad del polipéptido en relación con un polipéptido (o dominio Fn3) que no comprende la modificación de aminoácidos. La modificación de aminoácidos puede estar en un bucle o en una cadena  $\beta$ , por ejemplo, una cadena  $\beta$  E, del dominio Fn3. Los polipéptidos de la invención comprenden un dominio  $^{10}\text{Fn3}$  modificado, en donde el dominio  $^{10}\text{Fn3}$  modificado comprende una secuencia de aminoácidos en la que el aminoácido correspondiente al resto seleccionado 58 de la SEQ ID NO: 1 es un aminoácido hidrófilo que no es un resto de treonina, y en el que la solubilidad del dominio  $^{10}\text{Fn3}$  modificado se mejora en relación con la solubilidad de un dominio  $^{10}\text{Fn3}$  que comprende la misma secuencia de aminoácidos que el dominio  $^{10}\text{Fn3}$  modificado, excepto que el resto de aminoácido correspondiente al resto 58 de la SEQ ID NO: 1 es treonina.

La SEQ ID NO:1 es la secuencia del dominio  $^{10}\text{Fn3}$  humano de tipo silvestre establecida en: **VSDVPRDLEVVAATPTSLISWDAPAVTVRYRITYGETGGNSPVQEFVPGSKSTATISGLKPGVDYTITVYAVTGRG DSPASSKPI SINYPTEIDKPSQ** (SEQ ID NO: 1) (los bucles AB, CD y EF están subrayados; los bucles BC, DE y FG se resaltan en negrita; las cadenas  $\beta$  se localizan entre cada una de las regiones bucle; y las regiones de N-terminal y C-terminal se muestran en cursiva). En la SEQ ID NO:1, el bucle AB se refiere a los restos 14-17, el bucle BC se refiere a los restos 23-31, el bucle CD se refiere a los restos 37-47, el bucle DE se refiere a los restos 51-56, el bucle EF se refiere a los restos 63-67, y el bucle FG se refiere a los restos 75-87. Los bucles BC, DE y FG se alinean a lo largo de una cara de la molécula, es decir, el "polo norte", y los bucles AB, CD y EF se alinean a lo largo de la cara opuesta de la molécula, es decir, el "polo sur". En la SEQ ID NO: 1, la cadena  $\beta$  A se refiere a los restos 8-13, la cadena  $\beta$  B se refiere a los restos 18-22, la cadena  $\beta$  C se refiere a los restos 32-36, la cadena  $\beta$  D se refiere a los restos 48-50, la cadena  $\beta$  E se refiere a los restos 57-62, la cadena  $\beta$  F se refiere a los restos 68-74, y la cadena  $\beta$  G se refiere a los restos 88-92. Las cadenas  $\beta$  se unen entre sí a través del correspondiente bucle, por ejemplo, las

cadena A y B están conectadas a través del bucle AB en la formación de la cadena  $\beta$  A, el bucle AB, la cadena  $\beta$  B, etc. Las regiones N-terminal y/o C-terminal de la SEQ ID NO: 1 (en cursiva anteriormente), se pueden eliminar o alterar para generar una molécula que conserva la actividad biológica o para introducir actividad de unión a diana.

#### 5 Aglutinantes que tienen un bucle de Polo Norte y de Polo Sur modificados

En algunas realizaciones, los polipéptidos que comprenden una mutación de mejora de solubilidad, es decir, una sustitución en la posición T58 por un aminoácido hidrófilo que no sea una treonina, comprende un dominio <sup>10</sup>Fn3 que tiene (i) una modificación en la secuencia de aminoácidos de al menos un bucle del polo norte seleccionado de los bucles BC, DE y FG relativos al bucle correspondiente del dominio Fn3 humano de tipo silvestre, por ejemplo, un dominio <sup>10</sup>Fn3 (SEQ ID NO: 1 o 5), y (ii) una modificación en la secuencia de aminoácidos de al menos un bucle del polo sur seleccionado de los bucles AB, CD y EF relativos al bucle correspondiente del dominio Fn3 humano de tipo silvestre, por ejemplo, el dominio <sup>10</sup>Fn3 (SEQ ID NO: 1 o 5). Los bucles modificados del polo norte y del polo sur pueden contribuir a la unión a la misma diana. Se contemplan diversas combinaciones de los bucles modificados del polo norte y del polo sur. Por ejemplo, un Fn3, por ejemplo, <sup>10</sup>Fn3, puede comprender un bucle modificado del polo norte y un bucle modificado del polo sur, un bucle modificado del polo norte y dos bucles modificados del polo sur, dos bucles modificados del polo norte y un bucle modificado del polo sur, dos bucles modificados del polo norte y dos bucles modificados del polo sur, tres bucles modificados del polo norte y un bucle modificado del grupo del polo sur, etc., en donde cada uno de los bucles modificados contribuyen a la unión a la misma diana. Las combinaciones a modo de ejemplo de los bucles del polo norte y del polo sur que se pueden modificar incluyen, por ejemplo, el bucle CD (polo sur) y el bucle FG (polo norte), el bucle CD (polo sur) y el bucle DE (polo norte), el bucle EF (polo sur) y el bucle FG (polo norte), el bucle AB (polo sur) y el bucle FG (polo norte), o el bucle DE (polo norte) y el bucle EF (polo sur). Otra combinación de bucles ejemplar es el bucle CD (polo sur), el bucle DE (polo norte) y el bucle EF (polo sur). Otra combinación de bucles ejemplar más es el bucle DE (polo norte) y una de más de los bucles AB, CD y EF (polo sur). Los bucles modificados pueden tener modificaciones de la secuencia a través de todo el bucle o solo en una parte del bucle. Además, uno o más de los bucles modificados puede tener inserciones o deleciones de manera que la longitud del bucle es variada en relación con la longitud del bucle correspondiente de la secuencia de tipo silvestre. En determinadas realizaciones, las regiones adicionales en el dominio Fn3 (es decir, además de los bucles del polo norte y del polo sur), tal como las cadenas  $\beta$ , las regiones de N-terminal y/o de C-terminal, también se pueden modificar en la secuencia en relación con el dominio Fn3 de tipo silvestre, y tales modificaciones adicionales también pueden contribuir a la unión a la diana.

La presente divulgación incluye adicionalmente restos de armazón basados en fibronectina que comprenden un dominio <sup>10</sup>Fn3 que se define generalmente siguiendo la secuencia:

VSDVPRDLEVAA(X)<sub>u</sub>LLISW(X)<sub>v</sub>YRITY(X)<sub>w</sub>FTV(X)<sub>x</sub>ATISGL(X)<sub>y</sub>YTI TVYA(X)<sub>z</sub>ISINYRT (SEQ ID NO: 17), o mediante la secuencia que tiene la SEQ ID NO: 18-29. En las SEQ ID NO: 17-29, el bucle AB se representa mediante (X)<sub>u</sub>, el bucle BC se representa mediante (X)<sub>v</sub>, el bucle CD se representa mediante (X)<sub>w</sub>, el bucle DE se representa mediante (X)<sub>x</sub>, el bucle EF se representa mediante (X)<sub>y</sub> y el bucle FG se representa mediante X<sub>z</sub>. X representa, independientemente, cualquier aminoácido, y el subíndice tras la X representa un número entero del número de aminoácidos. Por ejemplo, u, v, w, x, y y z pueden ser cada uno un número entero seleccionado, independientemente, de 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7 y 6-7. Las secuencias de las cadenas beta (subrayadas) pueden tener en cualquier parte de 0 a 10, de 0 a 8, de 0 a 6, de 0 a 5, de 0 a 4, de 0 a 3, de 0 a 2, o de 0 a 1 sustituciones, deleciones o adiciones a través de las 7 regiones de armazón en relación con los correspondientes aminoácidos mostrados en las SEQ ID NO: 17-29. Las secuencias de las cadenas beta pueden tener en cualquier parte de 0 a 10, de 0 a 8, de 0 a 6, de 0 a 5, de 0 a 4, de 0 a 3, de 0 a 2, o de 0 a 1 sustituciones, por ejemplo, sustituciones conservativas, a través de las 7 regiones de armazón en relación con los correspondientes aminoácidos mostrados en las SEQ ID NO: 17-29. En determinadas realizaciones, los restos de aminoácidos del núcleo hidrófobo (restos en negrita en la SEQ ID NO: 17 anteriormente) son fijos, y cualquier sustitución, sustituciones conservativas, deleciones o adiciones tienen lugar en otros restos que no sean los restos de aminoácidos hidrófobos del núcleo. En algunas realizaciones, los restos hidrófobos del núcleo de los polipéptidos proporcionados en el presente documento no se han modificado en relación con el dominio <sup>10</sup>Fn3 de tipo silvestre (por ejemplo, la SEQ ID NO: 1).

En algunas realizaciones, la secuencia de aminoácidos del dominio Fn3 modificado puede ser al menos el 40 %, 50 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 97 %, el 98 % o el 99 % idéntica a la de un dominio Fn3 de tipo silvestre, por ejemplo, un dominio Fn3 humano de la SEQ ID NO:1-16, 65 o 66. En algunas realizaciones, la secuencia de aminoácidos del dominio Fn3 modificado puede ser al menos el 50% idéntica a la de un dominio Fn3 de tipo silvestre. En algunas realizaciones, la secuencia de aminoácidos del dominio Fn3 modificado puede ser al menos el 65 % idéntica a la de un dominio Fn3 de tipo silvestre. En algunas realizaciones, la secuencia de aminoácidos del dominio Fn3 modificado puede ser al menos el 80 % idéntica a la de un dominio Fn3 de tipo silvestre. En algunas realizaciones, la secuencia de aminoácidos del dominio Fn3 modificado puede ser al menos el 90 % idéntica a la de un dominio Fn3 de tipo silvestre. En determinadas realizaciones, uno o más de los bucles no se modificarán en relación con la secuencia del correspondiente bucle de la secuencia de tipo silvestre y/o una o más de las cadenas  $\beta$  no se modificará en relación con la secuencia de la correspondiente cadena  $\beta$  de la secuencia de tipo silvestre. En determinadas realizaciones, cada una de las cadenas de tipo beta de un dominio <sup>10</sup>Fn3 en un resto Fn3 puede

comprender, consistir esencialmente en, o consistir en una secuencia de aminoácidos que es al menos el 80 %, 85 %, 90 %, 95 % o el 100 % idéntica a la secuencia de una correspondiente cadena beta o de tipo beta de la SEQ ID NO: 1. En algunas realizaciones, las variaciones en las regiones de la cadena  $\beta$  pueden no alterar la estabilidad del polipéptido en condiciones fisiológicas.

5 En algunas realizaciones, la región no de bucle del dominio  $^{10}\text{Fn3}$  se puede modificar mediante una o más sustituciones conservativas. Se puede alterar hasta un 3 %, un 5 %, un 10 %, un 20 % o incluso un 30 % o más de los aminoácidos en el dominio  $^{10}\text{Fn3}$  mediante sustitución conservativa sin alterar sustancialmente la afinidad del  $^{10}\text{Fn3}$  por un ligando. En determinadas realizaciones, las regiones no de bucle, por ejemplo, las cadenas  $\beta$  pueden comprender en cualquier parte de 0-15, 0-10, 0-8, 0-6, 0-5, 0-4, 0-3, 1-15, 1-10, 1-8, 1-6, 1-5, 1-4, 1-3, 2-15, 2-10, 2-8, 2-6, 2-5, 2-4, 5-15, o 5-10 sustituciones conservativas de aminoácidos. En realizaciones a modo de ejemplo, la modificación del armazón puede reducir la afinidad de unión del Fn3, por ejemplo,  $^{10}\text{Fn3}$ , el aglutinante para un ligando en menos de 100 veces, 50 veces, 25 veces, 10 veces, 5 veces o 2 veces. Puede ser que tales cambios puedan alterar la inmunogenicidad del Fn3 *in vivo*, y cuando se reduce la inmunogenicidad, tales cambios pueden ser deseables. Tal como se usa en el presente documento, "sustituciones conservativas" son restos que son físicamente o funcionalmente similares a los correspondientes restos de referencia. Es decir, una sustitución conservativa y su resto de referencia tienen un tamaño, forma, carga eléctrica, propiedades químicas similares, incluida la capacidad de formar enlaces covalentes o de hidrógeno, o similares. Las sustituciones conservativas a modo de ejemplo incluyen aquellas que cumplen los criterios definidos para una mutación puntual aceptada en Dayhoff et al., Atlas of Protein Sequence and Structure 5:345-352 (1978 & Supp.). Los ejemplos de sustituciones conservativas incluyen sustituciones en los siguientes grupos: (a) valina, glicina; (b) glicina, alanina; (c) valina, isoleucina, leucina; (d) ácido aspártico, ácido glutámico; (e) asparagina, glutamina; (f) serina, treonina; (g) lisina, arginina, metionina; y (h) fenilalanina, tirosina.

25 En algunas realizaciones, la secuencia de aminoácidos del dominio Fn3 modificado puede diferir de un Fn3 de tipo silvestre, por ejemplo, el dominio  $^{10}\text{Fn3}$  en, como mucho, 50, 45, 40, 35, 30, 25, 20, 15, 10 o 5 aminoácidos. En algunas realizaciones, la secuencia de aminoácidos del dominio Fn3 modificado puede diferir de un dominio Fn3 de tipo silvestre en, como mucho, 30 aminoácidos. En algunas realizaciones, la secuencia de aminoácidos del dominio Fn3 modificado puede diferir de un dominio Fn3 de tipo silvestre en, como mucho, 20 aminoácidos. En algunas realizaciones, la secuencia de aminoácidos del dominio Fn3 modificado puede diferir de un dominio Fn3 de tipo silvestre en, como mucho, 10 aminoácidos.

35 En algunas realizaciones, el dominio Fn3 modificado puede comprender al menos una variación de aminoácido seleccionada de una sustitución, delección y adición en al menos un bucle en comparación con un dominio Fn3 de tipo silvestre, por ejemplo, un dominio  $^{10}\text{Fn3}$  humano de la SEQ ID NO:1-16, 65 o 66. En algunas realizaciones, el dominio Fn3 modificado puede comprender al menos una variación de aminoácido en cada uno de al menos dos bucles en comparación con un dominio Fn3 de tipo silvestre. En algunas realizaciones, el dominio Fn3 modificado puede comprender al menos una variación de aminoácido en cada uno de al menos tres bucles en comparación con un dominio Fn3 de tipo silvestre. En algunas realizaciones, el dominio Fn3 modificado puede comprender al menos una variación de aminoácido en al menos un bucle seleccionado de bucles del polo norte (bucles BC, DE y FG), en comparación con un dominio de Fn3 de tipo silvestre. En algunas realizaciones, el dominio Fn3 modificado puede comprender al menos una variación de aminoácido en al menos un bucle seleccionado de bucles del polo sur (bucles AB, CD y EF), en comparación con un dominio de Fn3 de tipo silvestre. En algunas realizaciones, el dominio Fn3 modificado puede comprender al menos una variación de aminoácido seleccionada de sustitución, delección y adición, en al menos una región no de bucle y al menos un bucle, en comparación con un dominio de Fn3 de tipo silvestre. En los polipéptidos de la invención, el dominio  $^{10}\text{Fn3}$  modificado comprende al menos una variación de aminoácido en una cadena  $\beta$  en comparación con un dominio  $^{10}\text{Fn3}$  de tipo silvestre. En algunas realizaciones, el dominio Fn3 modificado puede comprender al menos una variación de aminoácido en cada uno de al menos dos cadenas  $\beta$  en comparación con un dominio Fn3 de tipo silvestre. En algunas realizaciones, el dominio Fn3 modificado puede comprender al menos una variación de aminoácido en cada uno de al menos tres cadenas  $\beta$  en comparación con un dominio Fn3 de tipo silvestre.

50 En algunas realizaciones, el dominio Fn3 comprende un bucle, AB; un bucle, BC; un bucle, CD; un bucle, DE; un bucle, EF; y un bucle, FG; y tiene al menos un bucle seleccionado de los bucles AB, BC, CD, DE, EF y FG con una secuencia de aminoácidos alterada en relación con la secuencia del correspondiente bucle del dominio  $^{10}\text{Fn3}$  humano de la SEQ ID NO: 1-16. En algunas realizaciones, se alteran los bucles BC, DE y FG. En determinadas realizaciones, se alteran los bucles AB, CD y EF. En determinadas realizaciones, el bucle FG es el único bucle que se altera. En otras realizaciones, se alteran los bucles CD y FG, y de forma opcional, no se alteran otros bucles. En determinadas realizaciones, se alteran los bucles CD y EF, y de forma opcional, no se alteran otros bucles. En algunas realizaciones, se combinan una o más alteraciones específicas del armazón con una o más alteraciones del bucle. Con "alteración" se refiere a una o más alteraciones en la secuencia de aminoácidos en relación con una secuencia de molde (es decir, el correspondiente dominio de fibronectina humana de tipo silvestre) e incluye adiciones, delecciones y sustituciones de aminoácidos.

65 En algunas realizaciones, los polipéptidos pueden comprender un dominio Fn3 modificado en el que las regiones sin bucle comprenden una secuencia de aminoácidos que es al menos el 80, 85, 90, 95, 98 o el 100 % idéntica a las regiones sin bucle de un dominio Fn3 de tipo silvestre (por ejemplo, un dominio Fn3 humano de la SEQ ID NO:1-16,

65 o 66), en el que al menos un bucle seleccionado de AB, BC, CD, DE, EF y FG se puede alterar. Por ejemplo, en determinadas realizaciones, el bucle AB puede tener hasta 4 sustituciones de aminoácidos, hasta 10 inserciones de aminoácidos, hasta 3 deleciones de aminoácidos o una combinación de las mismas; el bucle BC puede tener hasta 10 sustituciones de aminoácidos, hasta 4 deleciones de aminoácidos, hasta 10 inserciones de aminoácidos o una combinación de las mismas; el bucle CD puede tener hasta 6 sustituciones de aminoácidos, hasta 10 inserciones de aminoácidos, hasta 4 deleciones de aminoácidos o una combinación de las mismas; el bucle DE puede tener hasta 6 sustituciones de aminoácidos, hasta 4 deleciones de aminoácidos, hasta 13 inserciones de aminoácidos o una combinación de las mismas; el bucle EF puede tener hasta 5 sustituciones de aminoácidos, hasta 10 inserciones de aminoácidos, hasta 3 deleciones de aminoácidos o una combinación de las mismas; y/o el bucle FG puede tener hasta 12 sustituciones de aminoácidos, hasta 11 deleciones de aminoácidos, hasta 25 inserciones de aminoácidos o una combinación de las mismas.

Se debería de entender que no todos los restos en una región de bucle o no de bucle necesitan modificarse para lograr un dominio de unión de Fn3 que tiene una gran afinidad por una diana deseada. Además, también se pueden hacer inserciones y deleciones en las regiones bucle mientras que aún se producen dominios de unión de Fn3 de alta afinidad. Por consiguiente, en algunas realizaciones, uno o más bucles seleccionados de AB, BC, CD, DE, EF y FG se pueden alargar o acortar en longitud en relación con el correspondiente bucle, un Fn3 de tipo silvestre. En cualquiera de los polipéptidos dados, uno o más bucles se pueden extender en longitud, uno o más bucles se pueden reducir en longitud o combinaciones de los mismos. En algunas realizaciones, la longitud de un bucle dado se puede extender en 2-25, 2-20, 2-15, 2-10, 2-5, 5-25, 5-20, 5-15, 5-10, 10-25, 10-20, o 10-15 aminoácidos. En algunas realizaciones, la longitud de un bucle dado se puede reducir en 1-15, 1-12, 1-10, 1-5, 1-3, 1-2, 2-10, o 2-5 aminoácidos. En particular, el bucle FG de <sup>10</sup>Fn3 es de 13 restos de longitud, mientras que el correspondiente bucle en las cadenas pesadas de los anticuerpos varía de 4-28 restos. Para optimizar la unión al antígeno en los polipéptidos que dependen del FG para la unión a la diana, por lo tanto, la longitud del bucle FG de Fn3 se puede alterar en longitud así como en la secuencia para obtener la mayor flexibilidad y afinidad posibles en la unión a la diana.

En algunas realizaciones, las secuencias de aminoácidos de las regiones N-terminal y/o C-terminal del dominio Fn3 modificado se pueden modificar mediante deleción, sustitución o inserción en relación con las secuencias de aminoácidos de las correspondientes regiones de un dominio Fn3 de tipo silvestre (por ejemplo, un dominio Fn3 humano de la SEQ ID NO: 1-16, 65 o 66). También se pueden añadir secuencias adicionales al extremo N-terminal o al extremo C-terminal del dominio Fn3 modificado. Por ejemplo, en algunas realizaciones, la extensión de N-terminal puede comprender una secuencia de aminoácidos seleccionada del grupo que consiste en: M, MG y G. En algunas realizaciones, la secuencia de aminoácidos de la cola de C-terminal del dominio Fn3 modificado se puede modificar o truncar en relación con la secuencia de aminoácidos de la cola de C-terminal de un dominio Fn3 de tipo silvestre (por ejemplo, un dominio Fn3 humano de la SEQ ID NO: 1-16, 65 o 66). En algunas realizaciones, la secuencia de aminoácidos del extremo N-terminal del dominio Fn3 modificado se puede modificar o truncar en relación con la secuencia de aminoácidos del extremo N-terminal de un dominio Fn3 de tipo silvestre.

En determinadas realizaciones, la secuencia de aminoácidos de los primeros 1, 2, 3, 4, 5, 6, 7, 8 o 9 restos del dominio Fn3 modificado se pueden modificar o deleccionar en los polipéptidos proporcionados en el presente documento en relación con la secuencia de los correspondientes aminoácidos en un dominio Fn3 de tipo silvestre. En algunas realizaciones, los aminoácidos correspondientes a los aminoácidos 1-8 de la SEQ ID NO: 1 se pueden reemplazar por una región N-terminal alternativa que tiene de 1-20, 1-15, 1-10, 1-8, 1-5, 1-4, 1-3, 1-2 o 1 aminoácidos de longitud. Las regiones a modo de ejemplo de N-terminal alternativas incluyen M, MG, G, MGVSDVPRDL (SEQ ID NO: 30) y GVSDVPRDL (SEQ ID NO: 31), o truncamientos N-terminales de cualquiera de las SEQ ID NO: 30 y 31. Otras regiones adecuadas de N-terminal alternativas incluyen, por ejemplo, X<sub>n</sub>SDVPRDL (SEQ ID NO: 32), X<sub>n</sub>DVPRDL (SEQ ID NO: 33), X<sub>n</sub>VPRDL (SEQ ID NO: 34), X<sub>n</sub>PRDL (SEQ ID NO: 35), X<sub>n</sub>RDL (SEQ ID NO: 36), X<sub>n</sub>DL (SEQ ID NO: 37), o X<sub>n</sub>L, en donde n = 0, 1 aminoácidos, en donde cuando n = 1, X es Met o Gly, y cuando n = 2, X es Met-Gly. Cuando se añade una secuencia de Met-Gly al extremo N-terminal de un dominio <sup>10</sup>Fn3, la M se puede escindir, dejando una G en el extremo N-terminal. En otras realizaciones, la región de N-terminal alternativa comprende la secuencia de aminoácidos MASTSG (SEQ ID NO: 38).

En determinadas realizaciones, la secuencia de aminoácidos correspondiente a los aminoácidos 93-101, 94-101, 95-101, 96-101, 97-101, 98-101, 99-101, 100-101 o 101 de la SEQ ID NO: 1 se puede deleccionar o modificar en los polipéptidos proporcionados en el presente documento en relación con la secuencia de los aminoácidos correspondientes en el dominio <sup>10</sup>Fn3 humano de tipo silvestre (SEQ ID NO: 1). En realizaciones a modo de ejemplo, los aminoácidos correspondientes a los aminoácidos 95-101 de la SEQ ID NO: 1 pueden reemplazarse con una región C-terminal alternativa que tiene de 1-20, 1-15, 1-10, 1-8, 1-5, 1-4, 1-3, 1-2 o 1 aminoácidos de longitud. Los ejemplos específicos de secuencias alternativas de la región C-terminal incluyen, por ejemplo, polipéptidos que comprenden, que consisten esencialmente en, o que consisten en, EIEK (SEQ ID NO: 39), EGSGC (SEQ ID NO: 40), EIEKPCQ (SEQ ID NO: 41), EIEKPSQ (SEQ ID NO: 42), EIEKP (SEQ ID NO: 43), EIEKPS (SEQ ID NO: 44), EIEKPC (SEQ ID NO: 45) o HHHHHH (SEQ ID NO: 46). En algunas realizaciones, la región C-terminal alternativa comprende EIDK (SEQ ID NO: 47), y en realizaciones particulares, la región C-terminal alternativa es EIDKPCQ (SEQ ID NO: 48) o EIDKPSQ (SEQ ID NO: 49).

En determinadas realizaciones, el dominio Fn3 modificado puede tener tanto una secuencia alternativa de la región

N-terminal como una secuencia alternativa de la región de C-terminal.

En algunas realizaciones, para alterar la unión a integrina se puede mutar o delecionar al menos un resto del motivo de unión a integrina "arginina-glicina-ácido aspártico" (RGD) (por ejemplo, los aminoácidos 78-80 de la SEQ ID NO:1).

5 En algunas realizaciones, el bucle FG de los polipéptidos proporcionados en el presente documento no contiene un sitio de unión a integrina RGD. En una realización, la secuencia de RGD se puede reemplazar por una secuencia de aminoácido polar-aminoácido neutro-aminoácido ácido (en la dirección de N-terminal a C-terminal). En otra realización, la secuencia de RGD se puede reemplazar con SGE. En aún otra realización, la secuencia de RGD se reemplaza con RGE (véase, por ejemplo, la SEQ ID NO: 16). El polipéptido de la invención se une de forma específica a una diana a la que no se une mediante un dominio Fn3 de tipo silvestre, en particular, el dominio Fn3 humano de tipo silvestre que tiene, por ejemplo, SEQ ID NO: 1-16, 65 o 66.

15 En algunas realizaciones, el polipéptido se puede unir a una diana deseada con una  $K_D$  de menos de 500 nM, 400 nM, 300 nM, 200 nM, 100 nM, 50 nM, 20 nM, 10 nM, 5 nM, 1 nM, 500 pM, 100 pM o menos. En algunas realizaciones, el polipéptido se une a una diana deseada con una  $K_D$  de entre 1 pM y 1  $\mu$ M, entre 100 pM y 500 nM, entre 1 nM y 500 nM, o entre 1 nM y 100 nM. En algunas realizaciones, el polipéptido se une a una diana deseada con una  $K_D$  de menos de 500 nM. En algunas realizaciones, el polipéptido se une a una diana deseada con una  $K_D$  de menos de 100 nM.

20 En algunas realizaciones, un polipéptido puede comprender una secuencia de aminoácidos que es al menos el 40 %, 50 %, 60 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 97 %, el 98 % o el 99 % idéntica a una secuencia de aminoácidos seleccionada del grupo de secuencias que consiste en las SEQ ID NO: 1-16, y el polipéptido se une de forma específica a una diana, por ejemplo, con una  $K_D$  de menos de 1000 nM, 500 nM, 100 nM, 50 nM, 10 nM, 5 nM, 1 nM, 500 pM, 100 pM o menos. El polipéptido puede comprender cambios (o alteraciones) de aminoácidos en uno o más bucles y una o más cadenas.

25

#### Aglutinantes que tienen modificaciones en la región del bucle y del armazón

También se proporcionan en el presente documento dominios  $^{10}$ Fn3 que tienen una mutación que mejora la solubilidad, es decir, una sustitución en la posición T58 por un aminoácido hidrófilo que no sea treonina, y que tenga combinaciones de modificaciones de bucle y andamiaje. En particular, la solicitud proporciona polipéptidos que comprenden una solubilidad que mejora la mutación y un dominio  $^{10}$ Fn3 que comprende (i) una modificación en la secuencia de aminoácidos de al menos uno de los bucles AB, BC, CD, DE, EF o FG y (ii) una modificación en la secuencia de aminoácidos de al menos una región de armazón (es decir, una modificación en al menos una cadena  $\beta$ , la región N-terminal, y/o la región C-terminal) incluyendo la sustitución en la posición T58 por un aminoácido hidrófilo que no es treonina, en donde el(los) bucle(s) modificado(s) y la(s) región(es) de armazón modificada(s) contribuyen a la unión a la misma diana. En realizaciones a modo de ejemplo, las modificaciones en la región del armazón se localizan adyacentes a las modificaciones en una región de bucle, por ejemplo, si se modifica en bucle AB, las mutaciones del armazón pueden tender a localizarse en la cadena  $\beta$  A y/o en la cadena  $\beta$  B, que son adyacentes al bucle AB en la secuencia lineal del dominio  $^{10}$ Fn3. En otras realizaciones, se puede hallar una agrupación de modificaciones juntas en regiones de bucle y de armazón que son adyacentes entre sí en la secuencia lineal del dominio Fn3. Por ejemplo, los aglutinantes de Fn3 que tienen tanto modificaciones en el bucle como en el armazón, pueden tener agrupaciones de modificaciones de aminoácidos en las siguientes combinaciones de regiones de bucle y de armazón que son adyacentes entre sí en la secuencia lineal del dominio Fn3: cadena  $\beta$ /bucle/cadena  $\beta$ , bucle/cadena  $\beta$ /bucle, bucle/cadena  $\beta$ /bucle/cadena  $\beta$ , región terminal/cadena  $\beta$ /bucle, o bucle/cadena  $\beta$ /región terminal, etc. Por ejemplo, los dominios Fn3 que tienen nuevas combinaciones de modificaciones de bucle y de armazón pueden ser agrupaciones de modificaciones de manera que sobre un tramo de 20 aminoácidos contiguos, al menos 15 de los aminoácidos están modificados en relación con el tipo silvestre. En otras realizaciones, al menos 17 de 20, 18 de 20, 17 de 25, 20 de 25 o 25 de 30 restos en un tramo contiguo están modificados en relación con la secuencia del dominio Fn3 de tipo silvestre sobre el tramo correspondiente de aminoácidos. En determinadas realizaciones, un dominio Fn3 dado puede tener dos o tres agrupaciones de modificaciones separadas mediante tramos de secuencia no modificada (es decir, de tipo silvestre). Para cualquier región dada (es decir, un bucle, una cadena  $\beta$  o una región terminal) que se modifica, se puede modificar todo o solo una parte de la región en relación con la secuencia de tipo silvestre. Cuando se modifica una región de cadena  $\beta$ , preferentemente los restos hidrófobos del núcleo permanecen sin modificar (es decir, de tipo silvestre) y se modifican uno o más restos no del núcleo en la cadena  $\beta$ .

55

#### Aglutinantes del "lado oeste"

60 En algunas realizaciones, la solicitud proporciona dominios  $^{10}$ Fn3 que comprenden una mutación que mejora la solubilidad, es decir, una sustitución en la posición T58 por un aminoácido hidrófilo que no sea treonina, y que tenga una cara de unión a lo largo del "lado oeste" de la molécula ("aglutinantes del lado oeste" o "aglutinantes WS"). Los aglutinantes WS tal como se describen en el presente documento, comprenden un dominio  $^{10}$ Fn3 que tiene un bucle CD modificado y un bucle FG modificado, en comparación con las correspondientes secuencias de los bucles CD y FG establecidas en la SEQ ID NO: 1 o 5. El bucle CD y el bucle FG contribuyen ambos a la unión a la misma diana. En determinadas realizaciones, los aglutinantes WS pueden comprender modificaciones adicionales en una o más regiones en el dominio Fn3. Por ejemplo, Los aglutinantes WS pueden comprender modificaciones del armazón en una o más de las regiones de cadena  $\beta$  adyacentes a los bucles CD y/o FG. En particular, los aglutinantes pueden

65

comprender modificaciones de secuencia en una o más de la cadena  $\beta$  C, la cadena  $\beta$  D, la cadena  $\beta$  F y/o la cadena  $\beta$  G. Las modificaciones del armazón a modo de ejemplo incluyen modificaciones en una o más posiciones de la región del armazón que se corresponden con las posiciones de los aminoácidos: 33, 35, 49, 69, 71, 73, 89 y/o 91 de la SEQ ID NO: 1 o 5. Los aglutinantes WS también pueden comprender modificaciones en el bucle BC, particularmente en la parte C-terminal del bucle BC. En una realización, los dos últimos restos del bucle BC (es decir, que se corresponden con los aminoácidos 30 y 31 en el dominio  $^{10}\text{Fn3}$  de tipo silvestre) se modifican en relación con la secuencia de tipo silvestre. Todo o una parte de las modificaciones adicionales del bucle y del armazón pueden contribuir a la unión a la diana en conjunto con los bucles modificados CD y FG. Preferentemente, los restos hidrófobos del núcleo no se modifican en relación con la secuencia de tipo silvestre.

En determinadas realizaciones, un aglutinante WS tiene un bucle CD que es de aproximadamente 3-11, 4-9 o 5 restos de longitud; un bucle FG que es de aproximadamente 1-10, por ejemplo, 6 o 5, restos de longitud; una cadena C que es de aproximadamente 6-14, 8-11 o 9 restos de longitud; y/o una cadena F que es de aproximadamente 9-11 o 10 restos de longitud. Las posiciones 31, 33, 35 y 37-39 de la cadena beta C se pueden alterar en relación con la secuencia de tipo silvestre. Las posiciones 32, 34 y 36 de la cadena beta C se pueden alterar en relación con la secuencia de tipo silvestre. Las posiciones 67, 69, 71 y 73 de la cadena beta F se pueden alterar en relación con la secuencia de tipo silvestre. Las posiciones 68, 70 y 72 de la cadena beta F pueden ser restos hidrófobos. Un aglutinante WS puede comprender sustituciones de aminoácidos en las posiciones 30, 31, 32, 33, 34, 35, 36, 37, 38 y/o 39, tales como las posiciones 31, 33, 35, 37, 38 y/o 39, por ejemplo, las posiciones 31 y/o 33, de la SEQ ID NO: 1 o 5. Un aglutinante WS puede comprender sustituciones de aminoácidos en las posiciones 44, 45, 46, 47, 48, 49, 50 y/o 51, tales como las posiciones 44, 45, 47 y/o 49, de la SEQ ID NO: 1 o 5. Un aglutinante WS puede comprender sustituciones de aminoácidos en las posiciones 40, 41, 42, 43, 44 y/o 45 de la SEQ ID NO: 1 o 5. Un aglutinante WS puede comprender sustituciones de aminoácidos en las posiciones 67, 68, 69, 70, 71, 72, 73, 74, 75 y/o 76, tales como las posiciones 67, 69, 71, 73 y/o 76 o las posiciones 71, 73, 75 y/o 76, de la SEQ ID NO: 1 o 5. Un aglutinante WS puede comprender sustituciones de aminoácidos en las posiciones 76, 77, 78, 79, 81, 82, 83, 84, 85 y/o 86, tales como las posiciones 84 y/o 85 de la SEQ ID NO: 1 o 5. Un aglutinante WS puede comprender sustituciones de aminoácidos en las posiciones 85, 86, 87, 88, 89, 90, 91, 92, 93 y/o 94 de la SEQ ID NO: 1 o 5. Un aglutinante WS puede comprender sustituciones de aminoácidos en las posiciones 31, 33, 47, 49, 73 y/o 75 de la SEQ ID NO: 1 o 5. Un aglutinante WS puede comprender un bucle C que comprende de 4-9 variados, por ejemplo, aminoácidos de tipo no silvestre; un bucle FG que comprende de 5-6 variados, por ejemplo, aminoácidos de tipo no silvestre; y en donde los aminoácidos 31, 33, 35, 37-39, 67, 69, 71, 73 y 76 son de tipo no silvestre. Los aminoácidos "de tipo no silvestre" son aminoácidos que no son los que se encuentran en la misma posición en la molécula de  $^{10}\text{Fn3}$  humana de tipo silvestre (que tiene, por ejemplo, SEQ ID NO: 1 o 5).

Los aglutinantes WS a modo de ejemplo incluyen aquellos que tienen un aminoácido de tipo silvestre o mutado en las posiciones 30, 31, 33, 35, 37, 38, 46, 47, 49, 50, 67, 69, 71, 73, 75, 76, 84, 85, 86, 87, 89 o 91. Por ejemplo, un diseño de aglutinante WS puede comprender una o más modificaciones en los aminoácidos 39-45 del bucle CD y una o más modificaciones de aminoácidos en los aminoácidos 77-83 en el bucle FG (diseño WS-LI1), y en donde una molécula  $^{10}\text{Fn3}$  que tiene ese diseño se une específicamente a una molécula diana (y opcionalmente no comprende una secuencia RGD). Un diseño de aglutinante WS puede comprender el diseño de WS-LI1 y, como mucho, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 17, 20 o 25 modificaciones de aminoácidos adicionales en los bucles o en las cadenas. Por ejemplo, un diseño de aglutinante WS puede comprender el diseño de WS-LI1 y, como mucho, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 17, 20 o 25 modificaciones de aminoácidos adicionales en las posiciones de los aminoácidos tales como las posiciones de aminoácidos 37, 38, 46, 47, 75, 76 y 85-88. Otras modificaciones de aminoácidos que se pueden incluir son las de las posiciones 30, 31, 33, 35, 49, 50, 67, 69, 71, 73, 89 y 91.

En ciertos aglutinantes WS desvelados en el presente documento, al menos o como mucho 10, 20, 30, 40, 50 o 50 aminoácidos de la secuencia del diseño no varían, por ejemplo, no varían por sustitución. Por ejemplo, uno o más de los siguientes aminoácidos se conservan como los aminoácidos de la molécula  $^{10}\text{Fn3}$  humana de tipo silvestre: los aminoácidos en las posiciones 1-29, 32, 34, 36, 48, 51-66, 68, 70, 72, 88, 90 y 92-101.

#### Aglutinantes "frontales"

En algunas realizaciones, los polipéptidos proporcionados en el presente documento comprenden un dominio  $^{10}\text{Fn3}$  que comprende una mutación que mejora la solubilidad, es decir, una sustitución en la posición T58 por un aminoácido hidrófilo que no sea treonina, y que tenga modificaciones en los bucles CD, DE y, en algunos casos, EF, en donde todas las modificaciones de bucle contribuyen a la unión a la diana. Estos polipéptidos se citan como "aglutinantes frontales". Los aglutinantes frontales pueden comprender adicionalmente modificaciones en una o más regiones del armazón, particularmente en regiones del armazón que flanquean o que son adyacentes a una región de bucle modificada. Por ejemplo, los aglutinantes frontales pueden comprender una modificación del armazón en una o más de cadena  $\beta$  C, cadena  $\beta$  D y/o cadena  $\beta$  E en relación con las secuencias de las correspondientes cadenas  $\beta$  del dominio Fn3 de tipo silvestre, por ejemplo, el dominio  $^{10}\text{Fn3}$  humano (SEQ ID NO: 1 o 5). Preferentemente, los restos hidrófobos del núcleo no se modifican en relación con la secuencia de tipo silvestre. Las modificaciones del armazón a modo de ejemplo que pueden estar presentes en los aglutinantes frontales, incluyen modificaciones en una o más posiciones que se corresponden con las posiciones de aminoácidos 36, 49 y/o 50 de la SEQ ID NO: 1 o 5. Tales modificaciones del armazón pueden contribuir a la unión a la diana junto con los bucles modificados. En determinadas realizaciones, los aglutinantes frontales pueden comprender agrupaciones de modificaciones que abarcan varias

regiones de bucle y de cadenas del dominio <sup>10</sup>F<sub>n</sub>3. En particular, los aglutinantes frontales pueden comprender modificaciones en al menos 15, 20, 24, 25 o 27 de los 31 restos entre los aminoácidos que se corresponden con los restos 36 a 66 del F<sub>n</sub>3 de tipo silvestre, por ejemplo, el dominio <sup>10</sup>F<sub>n</sub>3 humano (SEQ ID NO: 1 o 5). Las modificaciones del bucle y/o de la cadena pueden incluir sustituciones, deleciones y/o inserciones de aminoácidos o combinaciones de las mismas. En realizaciones a modo de ejemplo, el bucle CD se extiende en longitud o se reduce en longitud en relación con el bucle CD del F<sub>n</sub>3, por ejemplo, el dominio <sup>10</sup>F<sub>n</sub>3 humano de tipo silvestre (SEQ ID NO: 1 o 5).

#### Aglutinantes "posteriores"

En algunas realizaciones, los polipéptidos proporcionados en el presente documento comprenden un dominio <sup>10</sup>F<sub>n</sub>3 que comprende una mutación que mejora la solubilidad, es decir, una sustitución en la posición T58 por un aminoácido hidrófilo que no sea treonina, y que tenga modificaciones en los bucles EF y FG, en donde las modificaciones de los bucles contribuyen a la unión a la misma diana. Estos polipéptidos se citan como "aglutinantes posteriores" en el presente documento. Los aglutinantes posteriores pueden comprender modificaciones adicionales en otras regiones de bucle y/o de armazón. Por ejemplo, un aglutinante posterior puede contener modificaciones en al menos una parte del bucle AB, preferentemente la parte N-terminal del bucle AB. En una realización ejemplar, los dos primeros aminoácidos del bucle AB (es decir, los que se corresponden con los restos de aminoácidos 14 y 15 del dominio <sup>10</sup>F<sub>n</sub>3 de tipo silvestre) se modifican en relación con la secuencia de tipo silvestre. En determinadas realizaciones, un aglutinante posterior también puede contener una o más modificaciones del armazón, particularmente, modificaciones en una o más regiones del armazón que son adyacentes a una región de bucle modificada. Por ejemplo, los aglutinantes posteriores pueden contener una o más modificaciones en una o más de la cadena β A, la cadena β G, la región N-terminal y/o la región C-terminal. Preferentemente, los restos hidrófobos del núcleo no se modifican en relación con la secuencia de tipo silvestre. Las modificaciones del armazón a modo de ejemplo incluyen modificaciones en una o más posiciones que se corresponden con las posiciones de aminoácidos 1-7, 9-13, 89, 91, 93 y/o 94 de la SEQ ID NO: 1 o 5. Una o más modificaciones adicionales del bucle y/o del armazón pueden contribuir a la unión a la diana junto con los bucles EF y FG modificados. Las modificaciones adecuadas de la región del bucle y/o del armazón incluyen sustituciones, deleciones y/o inserciones de aminoácidos o combinaciones de las mismas. En determinadas realizaciones, la secuencia de aminoácidos del bucle FG se extiende en longitud o se reduce en longitud en relación con el bucle FG del dominio <sup>10</sup>F<sub>n</sub>3 humano de tipo silvestre (SEQ ID NO: 1 o 5).

En determinadas realizaciones, un aglutinante posterior puede comprender una agrupación de restos de aminoácidos modificados sobre un tramo contiguo de varias regiones en el dominio <sup>10</sup>F<sub>n</sub>3. Por ejemplo, al menos 14 de los primeros 15 restos de aminoácidos del dominio <sup>10</sup>F<sub>n</sub>3 se pueden modificar en relación con los correspondientes restos en el F<sub>n</sub>3 de tipo silvestre, por ejemplo, dominio <sup>10</sup>F<sub>n</sub>3 humano (SEQ ID NO: 1 o 5), y/o al menos 15 de los 18 restos entre los aminoácidos correspondientes a los restos 80 a 97 (o 94) del F<sub>n</sub>3 de tipo silvestre, por ejemplo, el dominio <sup>10</sup>F<sub>n</sub>3 humano (SEQ ID NO: 1 o 5) se puede modificar en relación con los correspondientes restos en la secuencia de tipo silvestre.

#### Aglutinantes del "polo sur"

En determinadas realizaciones, la solicitud proporciona polipéptidos que comprenden un dominio <sup>10</sup>F<sub>n</sub>3, en donde el dominio <sup>10</sup>F<sub>n</sub>3 comprende una mutación que mejora la solubilidad, es decir, una sustitución en la posición T58 por un aminoácido hidrófilo que no sea treonina, y modificaciones en las secuencias de aminoácidos de la cadena β A, el bucle AB, la cadena β B, el bucle CD, la cadena β E, el bucle EF y la cadena β F, en relación con las secuencias de las correspondientes regiones de la secuencia de tipo silvestre. Estos polipéptidos se citan como "aglutinantes del polo sur" o "aglutinantes SP" (SP, siglas del inglés *South Pole*, Polo Sur) en el presente documento. Los bucles y las cadenas modificadas contribuyen a la unión a la misma diana. La secuencia de aminoácidos del bucle CD se puede extender en longitud o reducir en longitud en relación con el bucle CD del F<sub>n</sub>3 de tipo silvestre, por ejemplo, el dominio <sup>10</sup>F<sub>n</sub>3 humano (SEQ ID NO: 1 o 5). Los aglutinantes del polo sur pueden comprender modificaciones adicionales en la cadena β G y/o en la región C-terminal en relación con la secuencia de la correspondiente región de la secuencia de tipo silvestre. En realizaciones a modo de ejemplo, los aglutinantes del polo sur pueden comprender una o más modificaciones en los aminoácidos que se corresponden con las posiciones 11, 12, 19, 60, 61, 69, 91, 93 y 95-97 de la secuencia de tipo silvestre.

#### Aglutinantes del "Noroeste"

En algunas realizaciones, la solicitud proporciona polipéptidos que comprenden un dominio <sup>10</sup>F<sub>n</sub>3 que comprende una mutación que mejora la solubilidad, es decir, una sustitución en la posición T58 por un aminoácido hidrófilo que no sea treonina, y que comprende los bucles modificados en BC, DE y FG, en comparación con las correspondientes secuencias de los bucles BC, DE y FG establecidas en la SEQ ID NO: 1 o 5, así como modificaciones adicionales en una o más de los restos de la cadena β C, la cadena β D, la cadena β F y la cadena β G. Las modificaciones de la región del bucle y de la cadena β contribuyen juntas a la unión a la diana. Estas proteínas se citan como "aglutinantes del Noroeste" o "aglutinantes NW" (NW, siglas del inglés *Northwest*, Noroeste) en el presente documento. En realizaciones a modo de ejemplo, los aglutinantes NW comprenden una o más modificaciones del armazón en una cualquiera de, o una combinación de, posiciones de aminoácidos que se corresponden con las posiciones de la región del armazón R33, T49, Y73 y S89 de la SEQ ID NO: 1 o 5. Las modificaciones adecuadas en las regiones del bucle y

del armazón incluyen sustituciones, deleciones y/o inserciones de aminoácidos o combinaciones de los mismos. En determinadas realizaciones, uno o más de los bucles BC, DE y FG se extienden en longitud o se reducen en longitud, o combinaciones de los mismos, en relación con la secuencia de tipo silvestre. En una realización, cada uno de los bucles BC, DE y FG se extienden en longitud o se reducen en longitud, o combinaciones de los mismos, en relación con la secuencia de tipo silvestre (por ejemplo, la SEQ ID NO: 1 o 5). En determinadas realizaciones, solo se modifica una parte del bucle BC, en particular, la parte de C-terminal, en relación con la secuencia de tipo silvestre. Por ejemplo, el bucle BC se puede modificar solo en los restos de aminoácidos que se corresponden con los aminoácidos 27-31 del bucle BC de tipo silvestre, mientras que el resto del bucle BC (es decir, el que se corresponde con los restos 23-26 del bucle de tipo silvestre) se deja sin modificar.

#### Aglutinantes del "Noreste"

En algunas realizaciones, la solicitud proporciona polipéptidos que comprenden un dominio <sup>10</sup>F<sub>n</sub>3 que comprende una mutación que mejora la solubilidad, es decir, una sustitución en la posición T58 por un aminoácido hidrófilo que no es treonina, y que comprende un bucle BC, DE y FG modificado, así como una o más modificaciones adicionales en una cualquiera de, o combinación de, la región N-terminal, la cadena β A, la cadena β B y/o la cadena β E. Estas proteínas se citan como "aglutinantes del noreste" o "aglutinantes NE" (NE del inglés *Northeast*, Noreste) en el presente documento. En realizaciones a modo de ejemplo, los aglutinantes NE se modifican a una cualquiera de, o una combinación de, los aminoácidos que se corresponden con las posiciones de la región del armazón 1-7, E9, L19 y/o S21 de la secuencia de tipo silvestre (SEQ ID NO: 1 o 5). La combinación de las regiones modificadas del bucle y del armazón contribuye a la unión a la diana.

#### Aglutinantes "frontales del sur"

En algunas realizaciones, la solicitud proporciona polipéptidos que comprenden un dominio <sup>10</sup>F<sub>n</sub>3 que comprende una mutación que mejora la solubilidad, es decir, una sustitución en la posición T58 por un aminoácido hidrófilo que no sea treonina, y que comprenden modificaciones en uno o más de los bucles AB, CD, DE y EF, así como modificaciones adicionales en una o más de la cadena β B, la cadena β D y/o la cadena β E. Estas proteínas se citan como "aglutinantes frontales del sur" en el presente documento. La combinación de los restos modificados del bucle y de la cadena contribuye a la unión a la diana. En realizaciones a modo de ejemplo, se puede modificar un aglutinante frontal del sur en una o más posiciones de aminoácidos que se corresponden con las posiciones de la región del armazón L19, T49, S60, y/o G61 de la SEQ ID NO: 1 o 5 y/o en una o más posiciones de aminoácidos correspondientes a las posiciones de la región de bucle T14-S17, P51, T56, G40-E47, y/o K63-G65 de la SEQ ID NO: 1 o 5. En realizaciones a modo de ejemplo, un aglutinante frontal del sur se puede extender en longitud o reducir en longitud en el bucle AB, entre los aminoácidos que se corresponden con los restos 18 y 20 de la secuencia de tipo silvestre y/o en el bucle CD.

#### Aglutinantes "AG"

En algunas realizaciones, la solicitud proporciona polipéptidos que comprenden un dominio <sup>10</sup>F<sub>n</sub>3 que comprende una mutación que mejora la solubilidad, es decir, una sustitución en la posición T58 por un aminoácido hidrófilo que no es treonina, y que comprende una cadena β modificada A y una cadena β G, en comparación con la cadena correspondiente de la SEQ ID NO: 1 o 5. Estas proteínas se citan como "aglutinantes AG" o "cadena AG" en el presente documento. En determinadas realizaciones, los aglutinantes de la cadena AG comprenden agrupaciones de modificaciones en las partes de N-terminal y de C-terminal del dominio <sup>10</sup>F<sub>n</sub>3, mientras que la parte media del F<sub>n</sub>3 permanece sin modificar. Por ejemplo, un aglutinante de cadena AG puede comprender modificaciones en 16 de 19 de los primeros 19 aminoácidos en el dominio <sup>10</sup>F<sub>n</sub>3 (es decir, correspondiente a las posiciones de aminoácidos 1-19 de la SEQ ID NO: 1 o 5) y modificaciones en 13-17 de 18 de los últimos 18 aminoácidos en el dominio <sup>10</sup>F<sub>n</sub>3 (es decir, correspondiente a las posiciones de aminoácidos 84-101 de la SEQ ID NO: 1) o en 14-18 de 22 de los últimos 22 aminoácidos en el dominio <sup>10</sup>F<sub>n</sub>3 (es decir, correspondiente a las posiciones de aminoácidos 80-101 de la SEQ ID NO: 1). En realizaciones a modo de ejemplo, un aglutinante AG puede comprender modificaciones en una o más posiciones que se corresponden con las posiciones 1-7, 9, 11-17, 19, 84-89 y 91-97 de la SEQ ID NO: 1. Preferentemente, las regiones modificadas en un aglutinante AG contribuyen a la unión a la misma diana.

#### Aglutinantes del "Suroeste"

En algunas realizaciones, la solicitud proporciona polipéptidos que comprenden un dominio <sup>10</sup>F<sub>n</sub>3 que comprende una mutación que mejora la solubilidad, es decir, una sustitución en la posición T58 por un aminoácido hidrófilo que no sea treonina, y que comprende un bucle modificado CD y EF, así como modificaciones adicionales en una cualquiera de, o combinación de restos que se corresponden con las posiciones 69 o 91-97 de la SEQ ID NO: 1. Estas proteínas se citan como "aglutinantes del Suroeste" o "aglutinantes SW" (SW, siglas del inglés *Southwest*, Suroeste) en el presente documento. Las regiones modificadas del bucle y del armazón contribuyen a la unión a la diana.

#### Proteínas que tienen inmunogenicidad reducida

La solicitud describe adicionalmente polipéptidos que tienen inmunogenicidad reducida que comprenden un dominio <sup>10</sup>F<sub>n</sub>3 en donde una parte del bucle BC se deja como de tipo silvestre. Preferentemente, tales polipéptidos tienen una

menor inmunogenicidad en relación con un polipéptido equivalente con modificaciones en una mayor parte del bucle BC. Por ejemplo, la parte de N-terminal del bucle BC se deja como de tipo silvestre. Por ejemplo, los primeros 1, 2, 3, 4, 5 o 5 restos del bucle BC se pueden dejar como de tipo silvestre, mientras que se pueden modificar los restos restantes de C-terminal del bucle BC. En los diseños de Fn3 que tienen al menos una parte de la región de N-terminal del bucle BC como de tipo silvestre, también puede ser deseable dejar todo o una parte de la cadena  $\beta$  B y/o de la cadena  $\beta$  C sin modificar en relación con la secuencia de tipo silvestre, particularmente las partes de la cadena  $\beta$  B y/o de la cadena  $\beta$  C que son adyacentes al bucle BC (es decir, la parte de C-terminal de la cadena  $\beta$  B y/o la parte de N-terminal de la cadena  $\beta$  C). Por ejemplo, los dominios Fn3 que tienen la secuencia de tipo silvestre en una parte de N-terminal del bucle BC e inmunogenicidad reducida pueden no tener ninguna modificación en la región de N-terminal, la cadena  $\beta$  A, el bucle AB y la cadena  $\beta$  B. En diseños de Fn3 con una parte del bucle BC como de tipo silvestre, la parte modificada del bucle BC puede contribuir a la unión a la diana junto con las modificaciones en otras regiones del dominio <sup>10</sup>Fn3.

La solicitud también describe polipéptidos que tienen inmunogenicidad reducida que comprenden dominios Fn3, en donde anclaje de HLA en la región de cadena  $\beta$  B/bucle BC/cadena  $\beta$  C (el "anclaje de BC") se ha eliminado o destruido (por ejemplo, modificado en relación con la secuencia de tipo silvestre de una forma que reduce la afinidad de unión a uno o más receptores de HLA). Por ejemplo, el anclaje de BC se puede eliminar o destruir modificando el Fn3, por ejemplo, el dominio <sup>10</sup>Fn3 en una o más posiciones que se corresponden con las posiciones L19, S21, R33 y/o T35 de SEQ ID NO: 1 o 5. Cuando el anclaje de BC se ha eliminado o destruido, es posible modificar la secuencia del bucle BC sin aumentar de manera significativa el potencial inmunogénico de la región BC. Por consiguiente, muchos de estos diseños de Fn3 tienen modificaciones en el bucle BC además de las modificaciones en la cadena  $\beta$  B y/o en la cadena  $\beta$  C. El bucle BC puede contribuir a la unión a la diana, de manera opcional, en combinación con modificaciones en otras regiones del dominio Fn3. Las modificaciones en la cadena  $\beta$  B y/o en la cadena  $\beta$  C pueden o no contribuir a la unión a la diana.

#### Mutaciones que mejoran la solubilidad de Fn3

En algunas realizaciones, un dominio Fn3 modificado comprende una secuencia de aminoácidos que se basa en una secuencia seleccionada de las SEQ ID NO: 1-29 y 65-66, en donde el dominio Fn3 comprende una mutación que mejora la solubilidad. Una mutación que mejora la solubilidad puede estar en una región de bucle o no de bucle. Por ejemplo, una mutación que mejora la solubilidad puede estar en una localización que se corresponde con al menos uno de los restos 24, 27, 54, 58, 78, 80 y 83 de la SEQ ID NO: 1. En algunas realizaciones, una mutación que mejora la solubilidad se localiza en una región no de bucle, tal como una cadena  $\beta$ . En los polipéptidos de la invención, una mutación que mejora la solubilidad se localiza en una cadena  $\beta$  E, a saber, en la Treonina (T) 58 que está sustituida por un aminoácido hidrófilo que no es treonina.

En algunas realizaciones, una mutación que mejora la solubilidad, por ejemplo, una sustitución por un aminoácido hidrófilo que no es treonina en el resto 58, en un dominio Fn3 que se une de manera específica a una diana, no contribuye a la unión del dominio Fn3 a la diana. En algunas realizaciones, el aminoácido que se muta para mejorar la solubilidad, por ejemplo, T58, no contribuye a la unión del dominio Fn3 a su diana. En algunas realizaciones, el aminoácido que se muta para mejorar la solubilidad, por ejemplo, T58, no está en contacto con la diana.

En algunas realizaciones, una mutación que mejora la solubilidad, por ejemplo, una sustitución por un aminoácido hidrófilo que no es treonina en el resto 58, en un dominio Fn3 que se une de manera específica a una diana, contribuye a la unión del dominio Fn3 a la diana. En algunas realizaciones, el aminoácido que se muta para mejorar la solubilidad, por ejemplo, T58, contribuye a la unión del dominio Fn3 a su diana. En algunas realizaciones, el aminoácido que se muta para mejorar la solubilidad, por ejemplo, T58, está en contacto con la diana.

La solubilidad del dominio Fn3 modificado del polipéptido de la invención se mejora en relación con la solubilidad de un dominio Fn3 que comprende la misma secuencia de aminoácidos salvo en que el aminoácido que se corresponde con el resto 58 de la SEQ ID NO: 1 es treonina.

En algunas realizaciones, el dominio Fn3 modificado puede comprender una secuencia de aminoácidos que se basa en una secuencia seleccionada de las SEQ ID NO: 1-29 y 65-66, en donde el aminoácido correspondiente al resto 58 de la SEQ ID NO: 1 está mutado a un aminoácido hidrófilo que no es treonina y en donde la solubilidad del dominio Fn3 modificado se mejora con respecto a la solubilidad de un dominio Fn3 en donde el aminoácido correspondiente al resto 58 de la SEQ ID NO: 1 es treonina.

En algunas realizaciones, el dominio Fn3 modificado puede comprender una secuencia de aminoácidos que es un 75 %, 80 %, 85 %, 90 %, 95 % o un 98 % idéntica a una secuencia de aminoácidos seleccionada de las SEQ ID NO: 1-29 y 65-71.

En algunas realizaciones, el dominio Fn3 modificado puede comprender una secuencia de aminoácidos seleccionada de las SEQ ID NO: 17-29, en donde el aminoácido correspondiente al resto 58 de la SEQ ID NO: 1 está mutado a Glu (E) o Asp (D), y en donde la solubilidad del dominio Fn3 modificado se mejora con respecto a la solubilidad de un dominio Fn3 en donde el aminoácido correspondiente al resto 58 de la SEQ ID NO: 1 es treonina.

En algunas realizaciones, el dominio Fn3 modificado puede comprender una secuencia de aminoácidos seleccionada de las SEQ ID NO: 67-68, en donde el aminoácido correspondiente al resto 58 de la SEQ ID NO: 1 está mutado a un aminoácido hidrófilo que no es treonina y en donde la solubilidad del dominio Fn3 modificado se mejora con respecto a la solubilidad de un dominio Fn3 en donde el aminoácido correspondiente al resto 58 de la SEQ ID NO: 1 es treonina. En algunas realizaciones, el dominio Fn3 modificado puede comprender una secuencia de aminoácidos seleccionada de las SEQ ID NO: 69-71.

Además, en el presente documento se describen dominios Fn3 modificados que se basan en un dominio <sup>10</sup>F<sub>n3</sub> (por ejemplo, uno cualquiera de los dominios <sup>10</sup>F<sub>n3</sub> humanos de la SEQ ID NO: 1-16, o uno cualquiera de los dominios <sup>10</sup>F<sub>n3</sub> humanos de la SEQ ID NO:17-29), un dominio <sup>7</sup>F<sub>n3</sub> (por ejemplo, un dominio <sup>7</sup>F<sub>n3</sub> humano de la SEQ ID NO:65) o un dominio <sup>14</sup>F<sub>n3</sub> (por ejemplo, un dominio <sup>14</sup>F<sub>n3</sub> humano de la SEQ ID NO:66), en donde el aminoácido que se corresponde con el resto 58 de la SEQ ID NO: 1 se puede mutar. En los polipéptidos de la invención, el dominio Fn3 modificado es un dominio <sup>10</sup>F<sub>n3</sub> modificado en el que el aminoácido correspondiente al resto 58 de la SEQ ID NO: 1 está mutado a un aminoácido hidrófilo que no es treonina.

Además, en el presente documento se desvelan dominios Fn3 modificados en los que el aminoácido correspondiente al resto 58 de la SEQ ID NO: 1 está mutado a cualquier aminoácido excepto Thr (T). En los polipéptidos de la invención, el aminoácido correspondiente al resto 58 de la SEQ ID NO: 1 está mutado a un aminoácido hidrófilo que no es treonina, tal como un aminoácido seleccionado de Gln (Q), Glu (E) y Asp (D). En algunas realizaciones, el aminoácido correspondiente al resto 58 de la SEQ ID NO: 1 puede mutar a un aminoácido seleccionado de Glu (E) y Asp (D). En algunas realizaciones, el aminoácido correspondiente al resto 58 de la SEQ ID NO: 1 se puede mutar a Glu (E) en un dominio <sup>10</sup>F<sub>n3</sub>. En algunas realizaciones, el aminoácido correspondiente al resto 58 de la SEQ ID NO: 1 se puede mutarse a Asp (D) en un dominio <sup>10</sup>F<sub>n3</sub>.

En algunas realizaciones, el dominio Fn3 comprende una mutación que mejora la solubilidad, es decir, una sustitución por un aminoácido hidrófilo que no es treonina en el resto 58, y comprende un aminoácido en la posición 19, 21 y/o 50 que corresponde al resto de tipo silvestre en esa posición, por ejemplo, Leu para el resto 19, Ser para el resto 21 y Ser para el resto 60. En algunas realizaciones, el dominio Fn3 comprende una mutación que mejora la solubilidad, es decir, una sustitución por un aminoácido hidrófilo que no es treonina en el resto 58, y una secuencia de aminoácidos en la que 1, 2, 3 o 4 de aminoácidos en las posiciones 63-66 se corresponden con el resto de tipo silvestre en esa posición (por ejemplo, K para el resto 63, P para el resto 64, G para el resto 65, V para el resto 66). En determinadas realizaciones, el dominio Fn3 puede comprender una mutación que mejora la solubilidad, es decir, una sustitución por un aminoácido hidrófilo que no es treonina en el resto 58, y una secuencia de aminoácidos en la que al menos 1, 2, 3, 4 o 5 de los restos de los bucles AB, CD, DE, EF o de los 7 aminoácidos de N-terminal se corresponden con el resto de tipo silvestre en esa posición.

En determinadas realizaciones, una proteína del dominio Fn3 que comprende una mutación que mejora la solubilidad no comprende una modificación de aminoácidos en el bucle AB en relación con la secuencia de aminoácidos del dominio Fn3 de tipo silvestre. En determinadas realizaciones, una proteína del dominio Fn3 que comprende una mutación que mejora la solubilidad no comprende una modificación de aminoácidos en el bucle BC en relación con la secuencia de aminoácidos del dominio Fn3 de tipo silvestre. En determinadas realizaciones, una proteína del dominio Fn3 que comprende una mutación que mejora la solubilidad no comprende una modificación de aminoácidos en el bucle CD en relación con la secuencia de aminoácidos del dominio Fn3 de tipo silvestre. En determinadas realizaciones, una proteína del dominio Fn3 que comprende una mutación que mejora la solubilidad no comprende una modificación de aminoácidos en el bucle DE en relación con la secuencia de aminoácidos del dominio Fn3 de tipo silvestre. En determinadas realizaciones, una proteína del dominio Fn3 que comprende una mutación que mejora la solubilidad no comprende una modificación de aminoácidos en el bucle EF en relación con la secuencia de aminoácidos del dominio Fn3 de tipo silvestre. En determinadas realizaciones, una proteína del dominio Fn3 que comprende una mutación que mejora la solubilidad no comprende una modificación de aminoácidos en el bucle FG en relación con la secuencia de aminoácidos del dominio Fn3 de tipo silvestre. En determinadas realizaciones, una proteína del dominio Fn3 que comprende una mutación que mejora la solubilidad no comprende una modificación de aminoácidos en los dominios N-terminal o C-terminal en relación con la secuencia de aminoácidos del dominio Fn3 de tipo silvestre. En determinadas realizaciones, una proteína del dominio Fn3 que comprende una mutación que mejora la solubilidad no comprende una modificación de aminoácidos en la cadena β A en relación con la secuencia de aminoácidos del dominio Fn3 de tipo silvestre. En determinadas realizaciones, una proteína del dominio Fn3 que comprende una mutación que mejora la solubilidad no comprende una modificación de aminoácidos en la cadena β B en relación con la secuencia de aminoácidos del dominio Fn3 de tipo silvestre. En determinadas realizaciones, una proteína del dominio Fn3 que comprende una mutación que mejora la solubilidad no comprende una modificación de aminoácidos en la cadena β C en relación con la secuencia de aminoácidos del dominio Fn3 de tipo silvestre. En determinadas realizaciones, una proteína del dominio Fn3 que comprende una mutación que mejora la solubilidad no comprende una modificación de aminoácidos en la cadena β D en relación con la secuencia de aminoácidos del dominio Fn3 de tipo silvestre. En determinadas realizaciones, una proteína del dominio Fn3 que comprende una mutación que mejora la solubilidad no comprende una modificación de aminoácidos en la cadena β F en relación con la secuencia de aminoácidos del dominio Fn3 de tipo silvestre. En determinadas realizaciones, una proteína del dominio Fn3 que comprende una mutación que mejora la solubilidad no comprende una modificación de aminoácidos

en la cadena  $\beta$  G en relación con la secuencia de aminoácidos del dominio Fn3 de tipo silvestre. En determinadas realizaciones, una proteína del dominio Fn3 que comprende una mutación que mejora la solubilidad no comprende una modificación de aminoácido en ninguno de los dos bucles, de las dos cadenas, los dos bucles y una cadena, los dos bucles y las dos cadenas, o un bucle y dos cadenas en relación con la secuencia de aminoácidos del dominio Fn3 de tipo silvestre.

En determinadas realizaciones, el bucle BC de una proteína de dominio Fn3 que comprende una mutación que mejora la solubilidad tiene la secuencia de aminoácidos del correspondiente bucle del dominio Fn3 humano de tipo silvestre, por ejemplo, el dominio <sup>10</sup>F<sub>n3</sub> (SEQ ID NO: 1). En determinadas realizaciones, los 2 primeros restos del bucle BC de una proteína de dominio Fn3 que comprende una mutación que mejora la solubilidad son los mismos que los restos correspondientes en el bucle BC del dominio Fn3 humano de tipo silvestre, por ejemplo, el dominio <sup>10</sup>F<sub>n3</sub> (SEQ ID NO: 1). En determinadas realizaciones, los 3 primeros restos del bucle BC de una proteína de dominio Fn3 que comprende una mutación que mejora la solubilidad son los mismos que los restos correspondientes en el bucle BC del dominio Fn3 humano de tipo silvestre, por ejemplo, el dominio <sup>10</sup>F<sub>n3</sub> (SEQ ID NO: 1). En determinadas realizaciones, los 4 primeros restos del bucle BC de una proteína de dominio Fn3 que comprende una mutación que mejora la solubilidad son los mismos que los restos correspondientes en el bucle BC del dominio Fn3 humano de tipo silvestre, por ejemplo, el dominio <sup>10</sup>F<sub>n3</sub> (SEQ ID NO: 1). En determinadas realizaciones, los 5 primeros restos del bucle BC de una proteína de dominio Fn3 que comprende una mutación que mejora la solubilidad son los mismos que los restos correspondientes en el bucle BC del dominio Fn3 humano de tipo silvestre, por ejemplo, el dominio <sup>10</sup>F<sub>n3</sub> (SEQ ID NO: 1).

Los cambios de aminoácidos en las regiones de bucle y/o no de bucle de un dominio Fn3 que se hacen para mejorar la solubilidad de un dominio Fn3 preferentemente no afectan de manera significativa a la actividad biológica del dominio Fn3. Por ejemplo, una mutación que mejora la solubilidad preferentemente no reduce de manera significativa la afinidad de unión (por ejemplo, K<sub>d</sub>) del dominio Fn3 a su diana deseada. En algunas realizaciones, una mutación que mejora la solubilidad, por ejemplo, una mutación T58, puede reducir la afinidad de unión (K<sub>d</sub>) de un dominio Fn3 en menos de un 1 %, 3 %, 5 %, 10 %, 20 %, 30 %, 50 %, 70 % o un 90 %. En algunas realizaciones, una mutación que mejora la solubilidad puede no reducir significativamente la estabilidad (por ejemplo, T<sub>m</sub>) de un dominio o proteína de Fn3 que comprende un dominio Fn3. En algunas realizaciones, una mutación que mejora la solubilidad, por ejemplo, una mutación T58, puede reducir la estabilidad (por ejemplo, T<sub>m</sub>) de un dominio o proteína Fn3 que comprende un dominio Fn3 en menos de un 1 %, 3 %, 5 %, 10 %, 20 %, 30 %, 50 %, 70 % o 90 %, o en relación con la T<sub>m</sub>, menos de 0,1 °C, 0,3 °C, 0,5 °C, 0,7 °C, 1 °C, 2 °C, 3 °C, 4 °C, 5 °C, 6 °C, 7 °C, 8 °C, 9 °C, 10 °C, 11 °C, 12 °C o 13 °C. En algunas realizaciones, una mutación que mejora la solubilidad no reduce de manera significativa bien la estabilidad o la afinidad de unión de una proteína Fn3.

De acuerdo con la invención, en un dominio <sup>10</sup>F<sub>n3</sub>, el aminoácido correspondiente al resto 58 de la SEQ ID NO: 1, está mutado a un aminoácido hidrófilo que no es un resto de treonina tal como Gln (Q), Glu (E) y Asp (D), en donde la solubilidad del dominio Fn3 modificado se mejora con respecto a la solubilidad del dominio <sup>10</sup>F<sub>n3</sub> en el que el aminoácido correspondiente al resto 58 de la SEQ ID NO: 1 es treonina, y en el que la mutación puede reducir la estabilidad (por ejemplo, T<sub>m</sub>) del dominio <sup>10</sup>F<sub>n3</sub> en menos del 1 %, 3 %, 5 %, 10 %, 20 %, 30 % o 50 %, o en relación con la T<sub>m</sub>, menos de 1 °C, 2 °C, 3 °C, 4 °C, 5 °C, 6 °C, 7 °C, 8 °C, 9 °C, 10 °C, 11 °C, 12 °C o 13 °C. En algunas realizaciones, en un dominio Fn3, el aminoácido correspondiente al resto 58 de la SEQ ID NO: 1 puede mutarse a Glu (E), en donde la solubilidad del dominio Fn3 modificado se puede mejorar en relación con la solubilidad del dominio Fn3 en donde el aminoácido correspondiente al resto 58 de la SEQ ID NO: 1 es treonina, y en donde la mutación puede reducir la estabilidad (por ejemplo, T<sub>m</sub>) del dominio Fn3 en menos del 5 %, 10 % o 15 %, o con respecto a la T<sub>m</sub>, menos de 1 °C, 2 °C, 3 °C, 4 °C, 5 °C o 10 °C. En algunas realizaciones, en un dominio Fn3, el aminoácido correspondiente al resto 58 de la SEQ ID NO: 1 puede mutar a Asp (D), en donde la solubilidad del dominio Fn3 modificado se puede mejorar en relación con la solubilidad del dominio Fn3 en donde el aminoácido correspondiente al resto 58 de la SEQ ID NO: 1 es treonina, y en donde la mutación puede reducir la estabilidad (por ejemplo, T<sub>m</sub>) del dominio Fn3 en menos del 15 % o el 20 %, o con respecto a la T<sub>m</sub>, menos de 5 °C, 10 °C o 15 °C.

En determinadas realizaciones, las mutaciones descritas en el presente documento mejoran la solubilidad de una proteína que es una proteína multivalente que comprende dos o más dominios Fn3, por ejemplo, los dominios <sup>10</sup>F<sub>n3</sub>. Por ejemplo, una proteína multivalente puede comprender 2, 3 o más dominios Fn3, por ejemplo, los dominios <sup>10</sup>F<sub>n3</sub>, que se asocian de manera covalente. En realizaciones a modo de ejemplo, la proteína puede ser una proteína biespecífica o dimérica que comprende dos dominios <sup>10</sup>F<sub>n3</sub>. Una mutación que mejora la solubilidad, por ejemplo, una sustitución por un aminoácido hidrófilo que no es treonina en la posición T58, puede estar presente en uno o más de los dominios Fn3 de una proteína Fn3 multimérica. En determinadas realizaciones, cada dominio Fn3 de una proteínas Fn3 multimérica puede comprender una mutación que mejora la solubilidad, por ejemplo, una sustitución por un aminoácido hidrófilo que no es treonina en la posición T58.

En algunas realizaciones, el polipéptido descrito en el presente documento comprende adicionalmente al menos un resto farmacocinético (PK) seleccionado de: un resto de polioxialquileno, una proteína de unión a albúmina sérica humana, ácido siálico, albúmina sérica humana, transferrina, IgG, una proteína de unión a IgG y un Fc (o un fragmento del mismo). En algunas realizaciones, el resto de PK es el resto de polioxialquileno. En algunas realizaciones, el resto de polioxialquileno es polietilenglicol (PEG). En algunas realizaciones, el resto de PEG está unido de manera covalente

al polipéptido a través de un aminoácido de Cys o de Lys. En algunas realizaciones, el PEG está entre aproximadamente 0,5 kDa y aproximadamente 100 kDa. En algunas realizaciones, el resto de PK es un fragmento de Fc.

5 Tal como se describe en el presente documento, los dominios Fn3 modificados se pueden usar para unirse a cualquier diana de interés para el tratamiento de enfermedades o trastornos. Las enfermedades o trastornos que se pueden  
 10 tratar estarán dictados por la especificidad de unión de los dominios Fn3. El polipéptido de la invención se une específicamente a una diana que no está unida por un dominio Fn3 humano de la SEQ ID NO: 1-16, 65 o 66. Las dianas a modo de ejemplo incluyen, por ejemplo, TNF-alfa, VEGFR2, PCSK9, IL-23, EGFR, IGF1R, DLL4, IL-17 y  
 15 PXR. Simplemente como un ejemplo, los dominios Fn3 modificados que se unen a TNF-alfa se pueden usar para tratar trastornos autoinmunes tales como artritis reumatoide, enfermedad inflamatoria del intestino, psoriasis y asma. Los dominios Fn3 modificados que se unen a IL-17 se pueden usar para tratar asma; y los dominios Fn3 modificados que se unen a DLL4 o EGFR se pueden usar para tratar trastornos o enfermedades hiperproliferativas asociadas con  
 20 angiogénesis no deseada, tales cánceres o tumores.

La solicitud también describe métodos para administrar los polipéptidos descritos en el presente documento a un sujeto, por ejemplo, un ser humano. De manera conveniente, las proteínas son farmacéuticamente aceptables para un mamífero, en particular un ser humano. Una composición "farmacéuticamente aceptable" se refiere a una composición que se administra a un animal sin consecuencias médicas adversas significativas.

Además, la solicitud describe composiciones farmacéuticamente aceptables que comprenden el polipéptido descrito en el presente documento. La composición puede ser esencialmente sin pirógenos. La composición puede estar sustancialmente libre de contaminación microbiana, haciéndola adecuada para la administración *in vivo*. La composición se puede formular, por ejemplo, para la administración (IV), intraperitoneal (IP) o subcutánea (SubQ). La  
 25 composición puede comprender un vehículo fisiológicamente aceptable. El pH de la composición puede estar entre 2-9, 3-8, 4-7.5, 4-7, 4-6.5, o entre 4-5.5, o a aproximadamente 4,0, 4,5, 5,0, o 5,5. La concentración del polipéptido puede ser 1-1000, 1-500, 1-200, 1-100, 1-50, 1-20, 1-10, 1-5, o 0,5-2 mg/ml en la composición.

La invención se refiere a un ácido nucleico que codifica los polipéptidos de la invención. También se incluyen los vectores que contienen los polinucleótidos para tales polipéptidos. Los vectores adecuados incluyen, por ejemplo, vectores de expresión. Además, la solicitud describe una célula, que comprende un polinucleótido, un vector, o un vector de expresión, que codifica un polipéptido descrito en el presente documento. Las secuencias se optimizan preferentemente para maximizar la expresión en el tipo celular usado. La expresión puede ser en una célula bacteriana, tal como *E. coli*. Como alternativa, la expresión puede ser en una célula de mamífero. La célula puede expresar un polipéptido que comprende un dominio Fn3 modificado tal como se describe en el presente documento. Los polinucleótidos que codifican polipéptidos descritos en el presente documento pueden estar optimizados por codón para la expresión en el tipo celular seleccionado. También se describen métodos para producir un polipéptido tal como se describe en el presente documento, que comprenden cultivar una célula hospedadora que comprende un ácido nucleico, un vector o un vector de expresión que codifica el polipéptido descrito, y recuperar el polipéptido expresado del cultivo.

En determinadas realizaciones, la solicitud proporciona bibliotecas que comprenden una pluralidad de los polipéptidos de la invención. Las bibliotecas desveladas en el presente documento pueden comprender, por ejemplo, al menos  $10^4$ ,  $10^5$ ,  $10^6$ ,  $10^7$ ,  $10^8$ ,  $10^9$ ,  $10^{10}$ ,  $10^{11}$ ,  $10^{12}$ ,  $10^{13}$ , o  $10^{14}$  o más polipéptidos, comprendiendo cada opción una secuencia de aminoácidos diferente. También se desvelan métodos para identificar un polipéptido que se une a una diana de interés de una de las bibliotecas descritas en el presente documento. Por ejemplo, un método de exploración de la biblioteca puede comprender, por ejemplo, poner en contacto una biblioteca de los polipéptidos descritos en el presente documento con una diana de interés, y aislar los miembros de la biblioteca que se unen a la diana (por ejemplo, con una afinidad particular o en condiciones de lavado adecuadas). La etapa de aislamiento se puede llevar a cabo usando cualquier método adecuado, tal como una presentación en fago o una presentación en ARNm. De manera similar, la unión a la diana se puede llevar a cabo usando cualquier método adecuado tal como la inmovilización de la diana sobre un soporte sólido (por ejemplo, una columna, una microplaca, perlas, etc.) y mezclando la diana inmovilizada con la biblioteca en condiciones adecuadas para permitir la unión de la proteína. Los miembros de la biblioteca unidos después se pueden separar de los miembros de la biblioteca no unidos para producir una proteína Fn3 aislada que se une a la diana. En determinadas realizaciones, el método de aislamiento puede implicar repetidas rondas de las etapas de unión a la diana y de aislamiento.

También se desvelan polipéptidos aislados identificados mediante un método descrito en el presente documento. En algunas realizaciones, el polipéptido aislado se puede unir a la diana con una  $K_D$  de menos de 1500 nM (1,5  $\mu$ M), 1000 nM (1  $\mu$ M), 500 nM, 400 nM, 300 nM, 200 nM, 100 nM, 50 nM, 20 nM, 10 nM, 5 nM, 1 nM, 500 pM, 100 pM o menos. En algunas realizaciones, el polipéptido se puede unir a la diana con una  $K_D$  de entre 1 pM y 1  $\mu$ M, entre 100 pM y 500 nM, entre 1 nM y 500 nM, o entre 1 nM y 100 nM.

## SECUENCIAS

65 Dominios  $^{10}$ Fn3 de TS:

VSDVPRDLEVVAATPTSLLISWDAPAVTVRYRITYGETGGNSPVQEF  
TVPGSKSTATISGLKPGVDYTITVYAVTGRGDSPASSKPISINYRTEIDKPSQ (SEQ  
ID NO: 1)

Dominio <sup>10</sup>F<sub>n3</sub> de la SEQ ID NO: 1 con D97E:

5

VSDVPRDLEVVAATPTSLLISWDAPAVTVRYRITYGETGGNSPVQEF  
TVPGSKSTATISGLKPGVDYTITVYAVTGRGDSPASSKPISINYRTEIEKPSQ (SEQ  
ID NO: 2)

Versión 1 de la secuencia del núcleo del dominio <sup>10</sup>F<sub>n3</sub> de TS:

LEVVAATPTSLLISWDAPAVTVRYRITYGETGGNSPVQEFTVPGSKST  
ATISGLKPGVDYTITVYAVTGRGDSPASSKPISINY (SEQ ID NO: 3)

10

Versión 2 de la secuencia del núcleo del dominio <sup>10</sup>F<sub>n3</sub> de TS:

EVVAATPTSLLISWDAPAVTVRYRITYGETGGNSPVQEFTVPGSKST  
ATISGLKPGVDYTITVYAVTGRGDSPASSKPISINYRT (SEQ ID NO: 4)

15

Versión 3 de la secuencia del núcleo del dominio <sup>10</sup>F<sub>n3</sub> de TS:

VSDVPRDLEVVAATPTSLLISWDAPAVTVRYRITYGETGGNSPVQEF  
TVPGSKSTATISGLKPGVDYTITVYAVTGRGDSPASSKPISINYRT (SEQ ID NO: 5)

20

Versión 4 de la secuencia del núcleo del dominio <sup>10</sup>F<sub>n3</sub> de TS:

VSDVPRDLEVVAATPTSLLISWDAPAVTVRYRITYGETGGNSPVQEF  
TVPGSKSTATISGLKPGVDYTITVYAVTGRGDSPASSKPISINYRTE (SEQ ID NO:  
6)

25

Versión 5 de la secuencia del núcleo del dominio <sup>10</sup>F<sub>n3</sub> de TS:

VSDVPRDLEVVAATPTSLLISWDAPAVTVRYRITYGETGGNSPVQEF  
TVPGSKSTATISGLKPGVDYTITVYAVTGRGDSPASSKPISINYRTEI (SEQ ID NO:  
7)

Versión 6 de la secuencia del núcleo del dominio <sup>10</sup>F<sub>n3</sub> de TS:

VSDVPRDLEVVAATPTSLLISWDAPAVTVRYRITYGETGGNSPVQEF  
TVPGSKSTATISGLKPGVDYTITVYAVTGRGDSPASSKPISINYRTEID (SEQ ID  
NO: 8)

30

ES 2 814 558 T3

Versión 7 de la secuencia del núcleo del dominio <sup>10</sup>F<sub>n</sub>3 (versión 6 con D97E):

VSDVPRDLEVVAATPTSLLISWDAPAVTVRYRITYGETGGNSPVQEF  
TVPGSKSTATISGLKPGVDYTITVYAVTGRGDSPASSKPISINYRTEIE (SEQ ID  
NO: 9)

5

Versión 8 de la secuencia del núcleo del dominio <sup>10</sup>F<sub>n</sub>3 de TS:

VSDVPRDLEVVAATPTSLLISWDAPAVTVRYRITYGETGGNSPVQEF  
TVPGSKSTATISGLKPGVDYTITVYAVTGRGDSPASSKPISINYRTEIDK (SEQ ID  
NO: 10)

10

Versión 9 de la secuencia del núcleo del dominio <sup>10</sup>F<sub>n</sub>3 (versión 8 con D97E):

VSDVPRDLEVVAATPTSLLISWDAPAVTVRYRITYGETGGNSPVQEF  
TVPGSKSTATISGLKPGVDYTITVYAVTGRGDSPASSKPISINYRTEIEK (SEQ ID  
NO: 11)

15

Versión 10 de la secuencia del núcleo del dominio <sup>10</sup>F<sub>n</sub>3 de TS:

VSDVPRDLEVVAATPTSLLISWDAPAVTVRYRITYGETGGNSPVQEF  
TVPGSKSTATISGLKPGVDYTITVYAVTGRGDSPASSKPISINYRTEIDKP (SEQ ID  
NO: 12)

Versión 11 de la secuencia del núcleo del dominio <sup>10</sup>F<sub>n</sub>3 (versión 10 con D97E):

VSDVPRDLEVVAATPTSLLISWDAPAVTVRYRITYGETGGNSPVQEF  
TVPGSKSTATISGLKPGVDYTITVYAVTGRGDSPASSKPISINYRTEIEKP (SEQ ID  
NO: 13)

20

Versión 12 de la secuencia del núcleo del dominio <sup>10</sup>F<sub>n</sub>3 de TS:

VSDVPRDLEVVAATPTSLLISWDAPAVTVRYRITYGETGGNSPVQEF  
TVPGSKSTATISGLKPGVDYTITVYAVTGRGDSPASSKPISINYRTEIDKPS (SEQ ID  
NO: 14)

25

Versión 13 de la secuencia del núcleo del dominio <sup>10</sup>F<sub>n</sub>3 (versión 12 con D97E):

VSDVPRDLEVVAATPTSLLISWDAPAVTVRYRITYGETGGNSPVQEF  
TVPGSKSTATISGLKPGVDYTITVYAVTGRGDSPASSKPISINYRTEIEKPS (SEQ ID  
NO: 15)

Dominio <sup>10</sup>F<sub>n3</sub> de TS con sustitución D80E

VSDVPRDLEVVAATPTSLLISWDAPAVTVRYYRITYGETGGNSPVQEF  
TVPGSKSTATISGLKPGVDYTTITVYAVTGRGESPASSKPI SINYRTEIDKPSQ (SEQ  
ID NO: 16)

5 Secuencia degenerada del núcleo del dominio <sup>10</sup>F<sub>n3</sub> de TS:

VSDVPRDLEVVA A(X)<sub>u</sub>LLISW(X)<sub>v</sub>YRITY(X)<sub>w</sub>FTV(X)<sub>x</sub>ATISGL(X)<sub>y</sub>YTIT  
VYA(X)<sub>z</sub>ISINYRT (SEQ ID NO: 17)

VSDVPRDLEVVA A(X)<sub>u</sub>LLISW(X)<sub>v</sub>YRITY(X)<sub>w</sub>FTV(X)<sub>x</sub>ATISGL(X)<sub>y</sub>YTIT  
VYA(X)<sub>z</sub>ISINYRTE (SEQ ID NO: 18)

10

VSDVPRDLEVVA A(X)<sub>u</sub>LLISW(X)<sub>v</sub>YRITY(X)<sub>w</sub>FTV(X)<sub>x</sub>ATISGL(X)<sub>y</sub>YTIT  
VYA(X)<sub>z</sub>ISINYRTEI (SEQ ID NO: 19)

VSDVPRDLEVVA A(X)<sub>u</sub>LLISW(X)<sub>v</sub>YRITY(X)<sub>w</sub>FTV(X)<sub>x</sub>ATISGL(X)<sub>y</sub>YTIT  
VYA(X)<sub>z</sub>ISINYRTEID (SEQ ID NO: 20)

VSDVPRDLEVVA A(X)<sub>u</sub>LLISW(X)<sub>v</sub>YRITY(X)<sub>w</sub>FTV(X)<sub>x</sub>ATISGL(X)<sub>y</sub>YTIT  
VYA(X)<sub>z</sub>ISINYRTEIE (SEQ ID NO: 21)

15

VSDVPRDLEVVA A(X)<sub>u</sub>LLISW(X)<sub>v</sub>YRITY(X)<sub>w</sub>FTV(X)<sub>x</sub>ATISGL(X)<sub>y</sub>YTIT  
VYA(X)<sub>z</sub>ISINYRTEIDK (SEQ ID NO: 22)

VSDVPRDLEVVA A(X)<sub>u</sub>LLISW(X)<sub>v</sub>YRITY(X)<sub>w</sub>FTV(X)<sub>x</sub>ATISGL(X)<sub>y</sub>YTIT  
VYA(X)<sub>z</sub>ISINYRTEIEK (SEQ ID NO: 23)

20

VSDVPRDLEVVA A(X)<sub>u</sub>LLISW(X)<sub>v</sub>YRITY(X)<sub>w</sub>FTV(X)<sub>x</sub>ATISGL(X)<sub>y</sub>YTIT  
VYA(X)<sub>z</sub>ISINYRTEIDKP (SEQ ID NO: 24)

VSDVPRDLEVVA A(X)<sub>u</sub>LLISW(X)<sub>v</sub>YRITY(X)<sub>w</sub>FTV(X)<sub>x</sub>ATISGL(X)<sub>y</sub>YTIT  
VYA(X)<sub>z</sub>ISINYRTEIEKP (SEQ ID NO: 25)

VSDVPRDLEVVA A(X)<sub>u</sub>LLISW(X)<sub>v</sub>YRITY(X)<sub>w</sub>FTV(X)<sub>x</sub>ATISGL(X)<sub>y</sub>YTIT  
VYA(X)<sub>z</sub>ISINYRTEIDKPS (SEQ ID NO: 26)

25

VSDVPRDLEVVAAX<sub>u</sub>LLISW(X)<sub>v</sub>YRITY(X)<sub>w</sub>FTV(X)<sub>x</sub>ATISGL(X)<sub>y</sub>YTIT  
VYA(X)<sub>z</sub>ISINYRTEIEKPS (SEQ ID NO: 27)

VSDVPRDLEVVAAX<sub>u</sub>LLISW(X)<sub>v</sub>YRITY(X)<sub>w</sub>FTV(X)<sub>x</sub>ATISGL(X)<sub>y</sub>YTIT  
VYA(X)<sub>z</sub>ISINYRTEIDKPSQ (SEQ ID NO: 28)

5 VSDVPRDLEVVAAX<sub>u</sub>LLISW(X)<sub>v</sub>YRITY(X)<sub>w</sub>FTV(X)<sub>x</sub>ATISGL(X)<sub>y</sub>YTIT  
VYA(X)<sub>z</sub>ISINYRTEIEKPSQ (SEQ ID NO: 29)

- MGVSDVPRDL (SEQ ID NO: 30)
- GVSDVPRDL (SEQ ID NO: 31)
- 10 X<sub>n</sub>SDVPRDL (SEQ ID NO: 32)
- X<sub>n</sub>DVPRDL (SEQ ID NO: 33)
- X<sub>n</sub>VPRDL (SEQ ID NO: 34)
- X<sub>n</sub>PRDL (SEQ ID NO: 35)
- X<sub>n</sub>RDL (SEQ ID NO: 36)
- 15 X<sub>n</sub>DL (SEQ ID NO: 37)
- MASTSG (SEQ ID NO: 38)
- EIEK (SEQ ID NO: 39)
- EGSGC (SEQ ID NO: 40)
- EIEKPCQ (SEQ ID NO: 41)
- 20 EIEKPSQ (SEQ ID NO: 42)
- EIEKP (SEQ ID NO: 43)
- EIEKPS (SEQ ID NO: 44)
- EIEKPC (SEQ ID NO: 45)
- HHHHHH (SEQ ID NO: 46)
- EIDK (SEQ ID NO: 47)
- 25 EIDKPCQ (SEQ ID NO: 48)
- EIDKPSQ (SEQ ID NO: 49)

MPAPDRLRFTNETPSSLLISWTPPRVQITGYIIRYGPVGS DGRVKEFTVP  
PSVSSATITGLKPGTEYTISVIALKDNQESEPLRGRVTTGG (FibconB; SEQ ID NO:  
50)

- 30 TPSS (SEQ ID NO: 51)
- TPPRVQI (SEQ ID NO: 52)
- VGSDGR (SEQ ID NO: 53)
- PSVS (SEQ ID NO: 54)
- GLKPG (SEQ ID NO: 55)
- 35 KDNQSEPEP (SEQ ID NO: 56)

LDAPDQLQVTNVTDTSITVSWTPPSATITGYRITYTPSNGPGEPKELTVP  
PSSTSVTITGITPGVEYVVS VYALKDNQESPPLVGTCTT (SEQ ID NO: 57)

LPAPKNLVVSEVTEDSLRLSWTAPDAAFDSFLIQYQESEKVG EAINLTV  
PGSERSYDLTGLKPGTEYTVSIYGVKGGHRSNPLSAEFTT (SEQ ID NO: 58)

- 40 TEDS (SEQ ID NO: 59)
- TAPDAAF (SEQ ID NO: 60)
- SEKVG E (SEQ ID NO: 61)
- GSER (SEQ ID NO: 62)

GLKPG (SEQ ID NO: 63)  
KGGHRN (SEQ ID NO: 64)  
Dominio <sup>7</sup>F<sub>n3</sub> de tipo silvestre:

PLSPPTNLHLEANPDTGVLTVSWERSTTPDITGYRITTTPTNGQQGNSL  
EEVVHADQSSCTFDNLSPGLEYNVSVYTVKDDKESVPISDTIIP (SEQ ID NO: 65)

5

Dominio <sup>14</sup>F<sub>n3</sub> de tipo silvestre:

NVSPRRARVTDATETTITISWRKTETITGFQVDAVPANGQTPIQRTIK  
PDVRSYTTGLQPGTDYKIYLYTLNDNARSSPVVIDAST (SEQ ID NO: 66)

10

Dominio <sup>10</sup>F<sub>n3</sub> de EGFR8:

MGVSDVPRDLEVVAATPTSLLISWDSGRGSYQYYRITYGETGGNSPVQ  
EFTVPGPVHTATISGLKPGVDYTTITVYAVTDHKPHADGPHTYHESPISINYRTEID  
KPSQ (SEQ ID NO: 67)

15

Dominio <sup>10</sup>F<sub>n3</sub> de EGFR4:

MGVSDVPRDLEVVAATPTSLLISWYWEGLPYQYYRITYGETGGNSPV  
QEFTVPRDVNTATISGLKPGVDYTTITVYAVTDWYNPDTHEYIYHTIPISINYRTEID  
KPSQ (SEQ ID NO: 68)

20

Dominio <sup>10</sup>F<sub>n3</sub> de EGFR8-T58E:

MGVSDVPRDLEVVAATPTSLLISWDSGRGSYQYYRITYGETGGNSPVQ  
EFTVPGPVHTAEISGLKPGVDYTTITVYAVTDHKPHADGPHTYHESPISINYRTEID  
KPSQ (SEQ ID NO: 69)

Dominio <sup>10</sup>F<sub>n3</sub> de EGFR8-T58D:

MGVSDVPRDLEVVAATPTSLLISWDSGRGSYQYYRITYGETGGNSPVQ  
EFTVPGPVHTADISGLKPGVDYTTITVYAVTDHKPHADGPHTYHESPISINYRTEID  
KPSQ (SEQ ID NO: 70)

25

Dominio <sup>10</sup>F<sub>n3</sub> de EGFR4-T58E:

MGVSDVPRDLEVVAATPTSLLISWYWEGLPYQYYRITYGETGGNSPV  
QEFTVPRDVNTAEISGLKPGVDYTTITVYAVTDWYNPDTHEYIYHTIPISINYRTEID  
KPSQ (SEQ ID NO: 71)

30

### Ejemplos

La invención descrita en el presente documento se entenderá más fácilmente con referencia a los siguientes ejemplos que se incluyen meramente con fines ilustrativos de determinados aspectos y realizaciones de la presente divulgación, y no pretenden limitar de ningún modo la invención.

35

**Ejemplo 1: Solubilidad de proteínas basada en un ensayo basado en cosmótopo**

Este ejemplo describe la propensión a la agregación de las proteínas basado en su solubilidad relativa en sulfato de amonio (AS).

El efecto del AS sobre la solubilidad de las proteínas se sabe desde hace más de 80 años (Green, A.A., 1931, J. Biol. Chem. 93, 495-516), cuando se ha usado más comúnmente para la partición y la purificación de proteínas. Se cree que el mecanismo de la precipitación de proteínas inducido por AS o "precipitación salina" implica la unión fuerte de moléculas de agua mediante el anión sulfato cosmótopo polar, que deshidrata las superficies de la proteína, aumenta el potencial químico de la proteína y provoca que las moléculas de proteína se agreguen en un precipitado amorfo (Baldwin, R.L., 1996, Biophys J. 71:2056-63). Debido a que las superficies hidrofóbicas de la proteína están preferentemente deshidratadas sobre las superficies polares, la autoasociación de proteínas inducida por AS se dirige mediante la interacción de las superficies hidrofobas expuestas, similar a las fuerzas que dirigen la agregación en ausencia de AS (Young, L. *et al.*, 1994, Protein Sci. 3, 717-29; Arunachalam, J. y Gautham, N., 2008, Proteins. 71, 2012-25; Fink, A.L. 1998, Fold Des. 3, R9-23).

En este ejemplo se describe un método alrededor de este concepto y su aplicación en el análisis de la solubilidad o la propensión a la agregación de los compuestos terapéuticos de proteínas de varios formatos diferentes que incluyen adnectinas de dominio único, adnectinas multidominio/multiespecíficas, fusiones de adnectina-Fc (Adn-Fc), fusiones de anticuerpo de dominio (dAb)-Fc (dAb-Fc) y anticuerpos monoclonales (mAb). Los datos de solubilidad con AS se correlacionaron con la propensión a la agregación observada mediante otros métodos establecidos (véase, Yamniuk, A.P. *et al.*, 2013, J Pharm. Sci. 102, 2424-2439).

**MATERIALES Y MÉTODOS****Proteínas:**

Las adnectinas anti-EGFR de dominio único, las adnectinas de tándem biespecífico anti-EGFR/IGF1R (EI-tándems) y las adnectinas de tándem biespecífico anti-EGFR/IGFR PEGiladas se expresaron y se purificaron tal como se ha descrito anteriormente (Emanuel, S. L. *et al.*, 2011, MAbs 3, 38-48). Las fusiones Adnectina-Fc y dAb-Fc, así como los mAb se expresaron en células HEK293-6E o CHO-S y se purificaron mediante cromatografía convencional de afinidad con proteína A seguida de cromatografía preparativa de exclusión por tamaño (SEC, del inglés *size exclusion chromatography*). Las proteínas de fusión Adn-Fc Adn-Fc-a, Adn-Fc-b y Adn-Fc-e se sobreexpresaron en *E. coli*, y cada proteína se replegó resuspendiendo la parte insoluble del lisado en guanidina-HCl 6 M a pH 10,0 seguido por diálisis frente a tampón a pH neutro. La proteína replegada resultante se purificó tal como se describe anteriormente para las fusiones expresadas de mamífero. Se determinó que todas las proteínas eran puras a más del 95 % mediante SDS-PAGE.

**Ensayo de solubilidad en sulfato de amonio a escala de banco:**

Los estudios de solubilidad en AS a escala de banco se llevaron a cabo a temperatura ambiente, y salvo que se especifique lo contrario, el tampón fue NaPO<sub>4</sub> 10 mM, NaCl 130 mM a pH 7,1. Las muestras se prepararon mezclando solución madre de proteína con tampón y AS 3,5 M en tampón combinado y pH, para producir una serie de 8-12 muestras de [proteína] idéntica y aumentar [AS] en un intervalo deseado. Tal como se observa en otros estudios (Trevino, S.R. *et al.*, 2007, J Mol Biol. 366, 449-60; Schein, C.H., 1990, Biotechnology 8, 308-17), la precipitación de la proteína fue rápida (<5 min), en gran medida reversible tras la dilución, y la concentración de proteína soluble permaneció estable durante horas hasta días tras la eliminación de proteína precipitada mediante centrifugación o filtración. En algunos casos, se aprovechó la reversibilidad de la precipitación generando muestras de una concentración de proteína dada a las concentraciones de AS más bajas y más altas deseadas, y mezclando estas dos muestras en volúmenes apropiados para generar las concentraciones de AS intermedias de la serie de titulación. Las muestras se incubaron durante aproximadamente 10 minutos a temperatura ambiente, seguido bien por centrifugación o por filtración para eliminar la proteína precipitada, y se determinó la concentración de la proteína usando la absorbancia a 280 nm (A<sub>280</sub>) de 2-3 ul de muestra en un instrumento NanoDrop 2000 (Thermo Scientific).

**Ensayo automático de solubilidad en sulfato de amonio:**

Las muestras del ensayo automático de solubilidad en AS se prepararon usando un instrumento Tecan Genesis Freedom 200 a temperatura ambiente. En primer lugar, se aspiraron 125 ul de solución madre de proteína de un tubo o placa estándar y se dispensaron 15 ul en cada uno de los 8 pocillos en una placa de 384 pocillos (Whatman Uniplate n.º 7701-5101). A continuación, se aspiraron 40 ul de 8 soluciones de AS diferentes en tampón de una placa de 96 pocillos profundos y se dispensaron 35 ul en los 8 pocillos de proteína con mezcla suave inmediata. Tras la preparación de la última muestra en un conjunto dado, todas las muestras se incubaron durante 5 minutos adicionales a temperatura ambiente. Después, las muestras se mezclaron suavemente y se aspiraron 47 ul de la mezcla proteína/AS y se dispensaron 45 ul en una placa de filtro de 384 pocillos (PALL AcroPrep384 100 ul, 0,45 um, GHP n.º PN5071) colocada encima de una placa de UV de fondo plano y claro de 384 pocillos de Corning (Corning UV 384 pocillos n.º

3675). La proteína precipitada se filtró mediante centrifugación durante 5 minutos a 21 °C. La proteína soluble se detectó usando un lector de placas SpectroMax M5, midiendo o bien la absorbancia a 280 nm ( $A_{280}$ ) o la fluorescencia intrínseca usando una longitud de onda de excitación de 280 nm y una longitud de onda de emisión de 350 nm. Ambos métodos de detección generalmente dan resultados comparables, pero la detección mediante  $A_{280}$  se implementó como el método por defecto para el análisis.

#### **Análisis de datos de solubilidad en sulfato de amonio y ajuste de la curva:**

La solubilidad de una proteína en AS en la región de precipitación salina se puede describir mediante la relación lineal (Ec. 1) (Green, A.A., 1931, J. Biol. Chem. 93, 495-516):

$$\log S = \beta - K_s [AS] \quad (1)$$

en donde S es la solubilidad de la proteína (concentración de la proteína),  $\beta$  es la solubilidad teórica a cero molar de AS,  $K_s$  es la constante de precipitación salina, y [AS] es la concentración de AS. Teóricamente, la solubilidad de las diferentes proteínas se puede evaluar comparando los valores de  $\beta$  extrapolados de los datos en la región de precipitación salina. Sin embargo, a las bajas concentraciones de proteínas y volúmenes deseados para la selección sistemática de alto rendimiento, hubo datos limitados dentro de la región de precipitación salina y, por lo tanto, una considerable incertidumbre en la extrapolación, lo que hace que la comparación del valor  $\beta$  no sea fiable. Por lo tanto, los datos de precipitación salina se ajustaron a una función de curva sigmoideal (Ec. 2):

$$Y = \frac{A_1 - A_2}{1 + e^{(x - AS_m)/dx}} \quad (2)$$

en donde  $A_1$  es el valor de Y inicial (absorbancia o fluorescencia),  $A_2$  es el valor de Y final, dx es la pendiente y  $AS_m$  es el valor del punto medio. El análisis de  $AS_m$  permitió una rápida comparación cuantitativa de la solubilidad relativa de las diferentes proteínas usando bien  $A_{280}$  o la detección por fluorescencia, y eliminó la necesidad de convertir los datos a unidades de concentración de proteína. En general, los errores de ajuste de la curva calculados estaban en un intervalo similar a la desviación estándar típica para múltiples mediciones independientes de una proteína dada (aproximadamente 0,01-0,04 M). El análisis de los valores de  $AS_m$  es también similar a la determinación de los valores del punto medio de PEG descrito recientemente por Gibson, T.J. et al (2011, J Pharm Sci. 100, 1009-21)

#### **Estudios de estabilidad acelerada:**

Los estudios de estabilidad acelerada (degradación forzada) se llevaron a cabo para varias moléculas de adnectina-Fc incubando 1 mg/ml de muestras de proteína en fosfato de sodio ( $NaH_2PO_4$ ) 50 mM a pH 7,0 y/o succinato de sodio ( $Na_2C_4H_4O_4$ ) 50 mM a pH 6,0 durante 2-4 semanas a temperaturas que varían de 30-37 °C. En todos los casos, las temperaturas de incubación fueron al menos 15° por debajo de la temperatura de fusión ( $T_m$ , del inglés *melting temperature*, temperatura de fusión) de otras moléculas tal como se determina mediante calorimetría de barrido diferencial o fluorescencia de barrido térmico, con el fin de minimizar las vías de degradación del estado no natural. En diversos puntos temporales durante la incubación, se eliminaron las alícuotas y se sometieron a un análisis de cromatografía analítica por exclusión de tamaño (SEC) para determinar el porcentaje de especies de monómero, de alto peso molecular (APM) y de bajo peso molecular (BPM).

#### **Ultrafiltración:**

Los estudios de ultrafiltración con moléculas de adnectina anti-EGFR se realizaron en concentradores VivaSpin 3000 MWCO a temperatura ambiente. Se concentraron lentamente 0,5 - 1,0 mg/ml de proteína en PBS a pH 7,1 usando la fuerza centrífuga hasta un objetivo de 30 - 40 mg/ml con mezclado periódico para evitar la formación del gradiente, y después se incubaron las muestras a 4 °C durante 5 días. La proteína precipitada se eliminó mediante centrifugación, y la concentración de la proteína soluble se midió mediante  $A_{280}$  en un instrumento Nanodrop 2000, con el estado de agregación medido con DLS tal como se describe a continuación. Los estudios de ultrafiltración para las moléculas de adnectina-Fc o dAb-Fc se realizaron a 4 °C en concentradores VivaSpin 10.000 MWCO usando presión neumática positiva regulada a 15 psi. Para minimizar la formación de gradiente, los concentradores se colocaron en una placa oscilante ajustada a 60 rpm para la mezcla continua, y las muestras se retiraron periódicamente y se mezclaron suavemente con una pipeta. Normalmente, las muestras de partida fueron de 1 - 7 mg/ml de proteína y se concentraron hasta aproximadamente 50 mg/ml o más durante el trascurso de 6-8 horas. Se retiraron 20 - 40 ul en diversos puntos temporales y se almacenaron toda la noche a 4 °C, seguido por la centrifugación para eliminar cualquiera de los precipitados, y después se determinó la concentración de las proteínas mediante  $A_{280}$  y el estado de agregación de las muestras no diluidas se caracterizó mediante SEC analítico.

#### **Dispersión de luz dinámica:**

Los estudios de dispersión de luz dinámica (DLS, del inglés *dynamic light scattering*) se realizaron en un lector de placas Wyatt DynaPro en placas de 384 pocillos a 25 °C, usando muestras de proteína en  $NaPO_4$  10 mM, NaCl 130

mM a pH 7,1. Los parámetros experimentales típicos fueron 20 adquisiciones de 5 s cada una por medición y las mediciones se registraron al menos por triplicado y se promediaron para dar los valores documentados. Las funciones de autocorrelación de intensidad se ajustaron usando el algoritmo de "Regularización" en el programa informático Dynamics (Wyatt Technologies).

5

## RESULTADOS

### Solubilidad de adnectinas en sulfato de amonio:

10 Se realizaron estudios iniciales de solubilidad en AS a escala de banco usando un conjunto de moléculas de adnectina de dominio único (monoadnectinas) con bucles diseñados genéticamente para unirse de manera específica al receptor EGFR (EGFR1-7) o al receptor IGF1R (IGFR1) (Emanuel, S.L. *et al.*, 2011, MAbs 3, 38-48). Un ejemplo de curva de precipitación salina para 1 mg/ml de IGF1R se muestra en la **Figura 2**. A baja [AS], la [proteína] (concentración de proteína) medida fue la misma que la [proteína] que se ensayó (1 mg/ml), lo que indica que no ha tenido lugar la precipitación. Sin embargo, a mayor [AS] (concentración de AS), la solubilidad de IGFR1 disminuyó, dando como resultado la precipitación o "precipitación salina", y un correspondiente descenso en la [proteína] medida. Los datos se ajustaron a una función de curva sigmoïdal para obtener el valor del punto medio de la precipitación salina (AS<sub>m</sub>), tal como se describe en Materiales y Métodos.

20 Las curvas de salado para EGFR1-7 así como IGFR1 se muestran en **Figura 3A**, y los valores de AS<sub>m</sub> junto con otros datos biofísicos se resumen en la **Tabla 1**. A pesar de ser proteínas monoméricas de tamaño y estructura similar, cada una de las adnectinas tenía propiedades únicas de precipitación salina, con valores de AS<sub>m</sub> que varían desde 0,67 M para la adnectina menos soluble (EGFR4, SEQ ID NO:68) hasta AS<sub>m</sub> = 2,01 M para la adnectina más soluble (EGFR6). Para determinar si estos valores de AS<sub>m</sub> eran predictivos de la propensión a la agregación observada usando otros métodos, se realizaron experimentos de ultrafiltración a pequeña escala concentrando cada adnectina de EGFR hasta una concentración diana de 30 - 40 mg/ml. La concentración de proteína soluble se midió mediante A<sub>280</sub> y el estado de agregación se evaluó usando DLS. Estos datos mostraron que las adnectinas con menores valores de AS<sub>m</sub> también tenían límites de solubilidad más bajos (EGFR2, EGFR3, EGFR4) y/o mayores niveles de agregación (EGFR2, EGFR4, EGFR5) a elevadas concentraciones de proteína (**Tabla 1**). Por el contrario, los datos de solubilidad no muestran correlación con los parámetros biofísicos tales como la estabilidad térmica o con parámetros predichos a partir de la secuencia de aminoácidos tales como el pI, la carga o la hidropatía.

35 Se realizaron estudios similares de precipitación salina a escala de banco usando cada una de las adnectinas EGFR1-7 fusionadas en línea con IGF1R1 (EI-tándems), así como para los EI-tándems unidos de manera covalente con un polietilenglicol ramificado de 40 kDa en el extremo C-terminal (EI-tándem-PEG) que se hizo para extender la semivida sérica (Emanuel *et al.*, 2011, mAbs 3, 38-48). La pendiente de las curvas de precipitación salina para los EI-tándems fue similar para las monoadnectinas (**Figura 3B**). Sin embargo, en todos los EI-tándem PEGilados se dio la precipitación salina en un intervalo [AS] muy estrecho de aproximadamente 0,7 - 0,9 M y tuvieron transiciones de precipitación salina mucho más pronunciadas (**Figura 3C**). También se observó un comportamiento similar para otras proteínas PEGiladas (**Figura 4**), que probablemente se deba a efectos adicionales de exclusión de volumen del resto PEG adjunto (Stevenson, C.L. y Hageman, M.J., 1995, Pharm Res. 12, 1671-6). Curiosamente, a pesar de las diferencias en los valores de solubilidad absoluta en AS para los diferentes formatos de adnectina, la clasificación de la solubilidad relativa basada en la identidad de un dominio de adnectina anti-EGFR se conservó ampliamente entre los formatos (**Figura 3D**), lo que sugiere que la solubilidad en AS de las monoadnectinas es predictiva de la solubilidad para las moléculas biespecíficas multidominio o de formato PEG más complejas.

### Solubilidad en sulfato de amonio de fusiones adnectina-Fc:

50 Para determinar si la solubilidad en AS era predictiva de la propensión a la agregación de otros formatos terapéuticos de proteína tales como fusiones a Fc, se usó un método automático de solubilidad en AS basado en placa para generar curvas de precipitación salina usando o bien la absorbancia en UV (A<sub>280</sub>) o la detección por fluorescencia usando tan solo 50 microgramos de proteína para generar una curva de 8 puntos (**Figura 5B y C**). Las comparaciones de solubilidad relativa descritas a continuación para Adn-Fcs, dAb-Fcs y los mAb se realizaron utilizando concentraciones de proteína idénticas de 0,45 mg/ml.

55

Se aplicó el método automático de solubilidad en AS a una serie de moléculas de Adn-Fc con diversos bucles de unión seleccionados de manera específica para unirse a diferentes proteínas de dianas terapéuticas, cada una de las cuales estaba fusionada a un dominio de Fc de IgG1 (Adn-Fc). La mayoría de los Adn-Fc se produjeron en células de mamífero y se indican con letras mayúsculas (es decir, Adn-Fc-A), pero también se produjeron unas pocas moléculas en *E. coli*, y se indican con minúsculas (es decir, Adn-Fc-a). La propensión a la agregación de muchas de estas moléculas de Adn-Fc se ha caracterizado previamente a baja concentración (1 mg/ml) en condiciones aceleradas de estrés térmico, en donde tres moléculas (Adn-A, Adn-a y Adn-L) demostraron tener una propensión a la agregación significativamente mayor que las otras (**Figura 6A**). Aquellas moléculas que demostraron una baja propensión a la agregación en el estudio de estabilidad acelerada se ampliaron y se concentraron mediante ultrafiltración, y se midieron los niveles de alto peso molecular (APM) soluble mediante SEC a una elevada concentración de proteína. Estos experimentos identificaron dos moléculas adicionales, Adn-B y Adn-b, con una propensión a la agregación

65

significativamente mayor en comparación con las otras (**Figura 6B**).

Las propiedades de precipitación salina de las moléculas Adn-Fc se caracterizaron usando el método automático, y el comportamiento de precipitación salina se comparó con los datos de estabilidad acelerada y de ultrafiltración. Se descubrió que las moléculas de Adn-Fc precipitan durante un intervalo bastante amplio de [AS], con valores de AS<sub>m</sub> de 0,67 - 1,33 M (**Figura 6C y D**). Las moléculas producidas en células HEK y CHO tuvieron curvas de precipitación salina indistinguibles (no mostradas). Sin embargo, se descubrió que las proteínas producidas por *E. coli* eran o bien de solubilidad superior (Adn-Fc-b), menor (Adn-Fc-e) o similar (Adn-Fc-a) en comparación con las moléculas producidas por los mamíferos, Adn-Fc-B, Adn-Fc-E y Adn-Fc-A respectivamente (**Figura 6D**). Las tres moléculas (Adn-Fc-A, Adn-Fc-a, Adn-Fc-L) con la propensión a la agregación más alta en los estudios de estabilidad acelerada, también tuvieron los valores de AS<sub>m</sub> más bajos (AS<sub>m</sub> = 0,67 - 0,76 M). Adn-Fc-B y Adn-Fc-b también tuvieron una menor solubilidad en AS (AS<sub>m</sub> = 0,94 - 1,03 M) en comparación con las otras moléculas ensayadas (AS<sub>m</sub> = 1,15-1,33 M), que es consistente con mayores niveles de agregación observados para estas moléculas a concentraciones de proteína mayores. Por lo tanto, el ensayo de solubilidad en AS fue capaz de identificar fácilmente las mismas moléculas Adn-Fc de "elevada propensión a la agregación" que se identificaron en los estudios de estabilidad acelerada y de ultrafiltración.

### Solubilidad de fusiones dAb-Fc:

Para determinar si el método de solubilidad en AS es predictivo de la solubilidad relativa de un agente terapéutico de proteína que no sean las adnectinas, se ensayó una serie de moléculas de fusión de anticuerpo de dominio y Fc (dAb-Fc) usando el método automático. Los dAb eran de varias familias de secuencia diferente con bucles de unión seleccionados específicamente para unirse a diferentes dianas terapéuticas. La nomenclatura de estas moléculas incluye un identificador para la familia de la secuencia (ej. "3" en dAb3-1) incluso para la variante de secuencia específica dentro de la familia (ejemplo "1" en dAb3-1). Se produjeron moléculas de dAb como fusiones bien al dominio IgG4 Fc humano que contiene una bisagra que estabiliza la mutación S228p (Newman, R. *et al.*, 2001, Clin Immunol. 98, 164-74) (citada como IgG4), o fusionada a uno de muchos dominios de IgG1 Fc que se identifican con asteriscos, por ejemplo, IgG1\*, IgG1\*\*, IgG1\*\*\*, etc. Muchas de estas moléculas dAb-Fc (dAb3, dAb5, dAb6, familias dAb9) podrían purificarse como monómeros estables según lo juzgado por SEC-MALS (no se muestra) o DLS (**Figura 7B y C**), mientras que otras familias demostraron propiedades indeseables que incluyen una mayor agregación (familias dAb2 y dAb8) (**Figura 7B y C**), y/o interacción con varias columnas analíticas de SEC (familias dAb2, dAb8 y dAb9).

Los ejemplos de curvas de precipitación salina para las moléculas dAb-Fc representativas se muestran en la **Figura 7A**, los datos del radio hidrodinámico (Rh) dependiente de la concentración de proteínas para moléculas seleccionadas se muestran en la **Figura 7B** y los valores de AS<sub>m</sub> se comparan con el Rh medido a una concentración de proteína de 1 mg/ml en la **Figura 7C**. Como en los Adn-Fc descritos anteriormente, se descubrió que los dAb-Fc tienen un amplio intervalo de propiedades de precipitación salina, con valores de AS<sub>m</sub> de 0,64 - 1,40 M. En comparación con los datos para las moléculas que contienen el mismo dAb pero diferentes dominios Fc (familias dAb3, dAb5 y dAb8) se muestra que las propiedades de precipitación salina principalmente dependen de la identidad del dominio dAb y son principalmente independientes del dominio Fc (**Figura 7C**). Las moléculas monoméricas dAb-Fc de las familias dAb3, dAb5 y dAb6 tenían los valores de AS<sub>m</sub> más altos (1,11 - 1,40 M), incluyendo las siete moléculas de la familia dAb6-Fc (AS<sub>m</sub> = 1,33 - 1,40 M) que tenían una elevada identidad de secuencia (98,6 - 99,7 %) y propiedades biofísicas muy similares. Por otro lado, todas las moléculas dAb-Fc de las familias dAb2, dAb8 y dAb9 que demostraron una mayor agregación y/o "pegajosidad" hacia las columnas analíticas de SEC, se descubrió que tenían valores de AS<sub>m</sub> más bajos (0,64 - 1,05 M) (**Figura 7C**). Uno de los dAb-Fc con solubilidad en AS favorable (dAb5-4-IgG1\*) se produjo a gran escala para generar suficiente material para estudios de solubilidad de ultrafiltración y se confirmó la baja propensión a la agregación, con menos del 2 % de las especies de APM observadas a concentraciones tan altas como 100 mg/ml (**Figura 8**).

### Solubilidad de anticuerpos monoclonales:

Se ensayó la solubilidad de AS de un panel de anticuerpos monoclonales (mAb) que tenía o bien el mismo dominio Fc pero diferentes dominios Fab de unión a diana, o bien los mismos dominios Fab fusionados a diferentes dominios Fc. Los dominios Fc en este conjunto de muestras incluyeron la IgG1 de tipo silvestre, varios dominios modificados de IgG1 Fc (IgG1\*, IgG1\*\*, IgG1\*\*\*), IgG4-S228P (identificada como IgG4), IgG2a, o una IgG2a modificada (identificada como IgG2a\*). Los datos de sal para los mAb se muestran en la **Figura 9** y los valores de AS<sub>m</sub> y los identificadores Fab/Fc se enumeran en la **Tabla 2**. En comparación con las adnectinas, los Adn-Fc y los dAb-Fc, se descubrió que todos los mAb tenían precipitación salina en un intervalo de AS más estrecho, con valores de AS<sub>m</sub> de 1,31 - 1,62 M. A pesar de este intervalo estrecho, hubo diferencias claras y reproducibles en la solubilidad de las diferentes moléculas, que principalmente dependían de la identidad de los dominios Fab y eran menos dependientes del dominio Fc. Por ejemplo, los valores de AS<sub>m</sub> para mAb-A que tiene 3 dominios Fc diferentes fueron todos 1,54-1,62 M, y para mAb-B que tiene 5 dominios Fc diferentes fueron todos 1,38-1,51 M (**Tabla 2**). Esto implica que la agregación de mAb puede estar dirigida por la autoasociación de los dominios Fab y probablemente las CDR de unión a diana que son las regiones más variables de estas moléculas.

### Aplicación del ensayo de solubilidad en sulfato de amonio en la optimización del fármaco:

Tal como se describe en el presente documento, se usó una estrategia de barrido de Alanina para examinar la contribución de los restos del bucle de unión a la energía de interacción para una adnectina que se une al receptor EGFR (citado como EGFR8, que comprende un dominio  $^{10}\text{Fn3}$  de la SEQ ID NO: 67). Todos los restos en los bucles de unión a diana BC, DE y FG se sustituyeron individualmente con Ala, y se evaluaron los efectos sobre la unión a la diana. Además, un resto adyacente al bucle DE, T58, también se identificó basándose en la modelación molecular como un sitio en el que la energía de unión podría modificarse a través de la mutación. Por lo tanto, se crearon tres mutantes adicionales (T58→E, T58→D, T58→Q) y se estudió su unión a la diana. Aquellos mutantes de T58 y de Ala también se ensayaron para determinar si cualquiera de las mutaciones podría mejorar la solubilidad de EGFR8.

La adnectina EGFR8 de tipo silvestre (SEQ ID NO:67) ensayada a 0,33 mg/ml, se descubrió que tenía precipitación salina entre aproximadamente 1,9 - 2,3 M de AS, con  $\text{ASm} = 2,1 \text{ M}$  (**Figura 10**). Los estudios iniciales con mutantes seleccionados indicaron que los mutantes generalmente tenían precipitación salina en un intervalo similar al de la proteína de tipo silvestre. Por lo tanto, para seleccionar rápidamente los mutantes con un mínimo consumo de proteína, cada proteína se ensayó solo a una [AS] cerca del inicio (2,0 M) y cerca de la base (2,2 M) de la transición de la precipitación salina para EGFR8 de tipo silvestre, en donde los datos a 2,0 M podrían documentar principalmente moléculas con solubilidad reducida y los datos a 2,2 M podrían identificar moléculas con solubilidad más alta. Este experimento identificó mutantes con solubilidad más baja tales como D77→A, K79→A y E85→A y mutantes con solubilidad más alta que incluyen S24→A, G27→A, V54→A, T58→E, T58→D, H78→A, P80→A y Y83→A, en comparación con la adnectina EGFR8 de tipo silvestre (**Figura 11A**). Las tendencias en solubilidad para un subconjunto de estas moléculas se confirmaron generando curvas de titulación completas (**Figura 10**). Las mutaciones con solubilidad reducida generalmente fueron las de los restos hidrófilos tales como Asp, Glu o Lys a Ala, mientras que las que mejoraban la solubilidad generalmente fueron las mutaciones de los restos más hidrófobos a Ala, o la introducción de un resto hidrófilo Glu o Asp en el lugar de T58.

Dado que se determinó la estructura de EGFR8, las propensiones a la agregación específicas de los restos se calcularon para todos los restos mutados en la proteína usando el algoritmo de "propensión a la agregación espacial" (PAE) (Chennamsetty, N. *et al.*, 2009, Proc. Natl. Acad. Sci. EE.UU. 106, 11937-42). Los valores de PAE se calcularon a alta resolución usando un radio de 5 Å para identificar los restos específicos propensos a la agregación, así como a baja resolución con  $R = 10 \text{ Å}$  para identificar parches hidrófobos propensos a la agregación más grandes. Los datos de PAE para los restos mutados se muestran en la **Figura 11B**. Aunque se descubrió que la mayoría de la adnectina era altamente hidrófila (valores de PAE negativos), en el bucle de unión DE se predijo que tiene un carácter hidrófobo moderado (valores PAE positivos) y posiblemente propenso a la agregación. De los restos en este parche hidrófobo, las mutaciones que se predijo que reducían la hidrofobicidad y, por lo tanto, la agregación (por ejemplo, V54→A, T58→E y T58→D), la mayoría, de hecho, aumentó significativamente la solubilidad de EGFR8 en el ensayo de solubilidad en AS (**Figura 11A y Figura 10**). Los datos de solubilidad medidos también identificaron mutaciones adicionales tales como H78→A o Y83→A en las que la solubilidad se podría mejorarse mediante mutación, que no se predijeron solo mediante PAE. A la vista de los datos de unión a la diana basados en la estructura (Ejemplo 2), algunas de las mutaciones que mejoran la solubilidad (V54→A, H78→A, Y83→A) dieron como resultado reducciones indeseables en afinidad/potencia, probablemente debido a la retirada de interacciones hidrófobas importantes con la proteína diana de EGFR. Sin embargo, algunas otras mutaciones que mejoran la solubilidad (S24→A, G27→A, T58→E y T58→D) dieron como resultado una unión/potencia sin cambios o mejorada, lo que sugiere que estas mutaciones podrían ser modificaciones útiles para optimizar tanto la solubilidad como la potencia de la adnectina.

Tabla 1: Resumen de datos experimentales y de las propiedades teóricas de las adnectinas anti-EGFR.

Adnectina EGFR	ASm (M) <sup>1</sup>	Rh (nm)		Conc (mg/ml) <sup>3</sup>	Rh (nm) <sup>2</sup>	Tm (oC) <sup>4</sup>	Puntuación "GRAVY" de hidropatía <sup>5</sup>	pl teórico	Carga teórica a pH 7
		Conc (mg/ml)	Rh (nm)						
EGPR1	1,85	0,9	1,7 ± 0,1	33	2,3 ± 0,2	71	-0,58	5,9	-4,2
EGFR2	1,64	1,0	1,9 ± 0,1	22*	5,9 ± 0,7	59	-0,54	6,2	-2,3
EGFR3	1,34	1,1	1,8 ± 0,5	14*	2,5 ± 1,5	82	-0,57	6,0	-3,3
EGFR4	0,67	0,6	1,9 ± 0,1	4,9*	3,0 ± 0,1	77	-0,50	5,4	-6,3
EGFR5	1,45	1,4	2,0 ± 0,1	40	4,4 ± 0,0	78	-0,43	5,9	-3,4
EGFR6	2,01	0,9	1,9 ± 0,0	36	2,5 ± 0,0	74	-0,56	5,9	-4,2
EGFR7	1,53	1,2	1,7 ± 0,1	35	2,7 ± 0,1	n.d. <sup>6</sup>	-0,48	5,9	-4,2

<sup>1</sup> - a concentración de proteína de 0,28 mg/ml.

<sup>2</sup> - Rh determinado mediante dispersión de luz dinámica

<sup>3</sup> - los valores con un asterisco (\*) indican un límite de solubilidad aparente

<sup>4</sup> - Tm = temperatura de fusión

<sup>5</sup> - puntuación de hidropatía GRAVY calculada para los restos variables del bucle de unión a la diana. Todos los restos del armazón fueron idénticos en cada adnectina.

<sup>6</sup> - no determinado

**Tabla 2:** Valores de ASm para mAb que tienen diferentes dominios Fab o Fc.

Anticuerpo	Fab	Fc	ASm <sup>1</sup>
mAb-A-IgG1	A	IgG1	1,59 ± 0,01
mAb-A-IgG1*	A	IgG1*	1,62
mAb-A-IgG4	A	IgG4	1,54, 1,56
mAb-B-IgG1	B	IgG1	1,50
mAb-B-IgG1**	B	IgG1**	1,38
mAb-B-IgG2	B	IgG2	1,51
mAb-B-IgG2*	B	IgG2*	1,40
mAb-B-IgG4	B	IgG4	1,40
mAb-C-IgG1****	C	IgG1****	1,32, 1,33
mAb-D-IgG1	D	IgG1	1,61
mAb-E-IgG1	E	IgG1	1,31
mAb-F-IgG1****	F	IgG1****	1,38
mAb-G-IgG4	G	IgG4	1,53
mAb-H-IgG4	H	IgG4	1,43 ± 0,05

<sup>1</sup> - los valores de ASm para experimentos realizados tres o más veces se representan como la media ± la desviación estándar.

**Ejemplo 2: Estructura de la adnectina /complejo proteico**

- 5 Las adnectinas que se unen de manera específica al receptor de factor de crecimiento epidérmico (EGFR) o a la interleucina 23 (IL-23), dos dianas validadas terapéuticamente, se generaron usando la técnica de presentación en ARNm descrita anteriormente (XU *et al.*, 2002, Chem. Biol. 9, 933-942). Estas adnectinas inhiben la unión de su diana al receptor afin de la diana o al ligando y también bloquean la señalización intracelular de las interacciones diana/ligando en ensayos basados en células. Se seleccionó una adnectina representativa que bloqueaba cada diana para la cocrystalización con su diana para identificar la naturaleza de los contactos.

15 El análisis de las estructuras de las adnectinas de unión a EGFR y a IL-23 genera múltiples perspectivas en las interacciones moleculares entre las variantes basadas en <sup>10</sup>Fn3 y sus dianas. Muchos de los restos del bucle diversificado entran en contacto con la diana. De forma más interesante, no todos los restos en los tres bucles diversificados (**Figura 12C**) estaban en o cerca de la superficie de contacto del dominio basado en <sup>10</sup>Fn3/diana. Por el contrario, varios restos de tipo silvestre fuera de los bucles diversificados interactúan con la proteína diana. Las conformaciones de la cadena principal de las dos adnectinas de unión a diana y de <sup>10</sup>Fn3 de tipo silvestre fueron similares, pero las conformaciones de los tres bucles diversificados y del extremo N-terminal fueron diferentes del tipo silvestre para facilitar la unión a la respectiva superficie de la proteína diana.

20

**PROCEDIMIENTOS EXPERIMENTALES:****Selección de adnectina**

- 25 Los aglutinantes de adnectina iniciales frente a EGFR y IL-23 se obtuvieron usando PROfusión (Xu *et al.*, 2002, Chem. Biol. 9, 933-942; Getmanova *et al.*, 2006, Chem. Biol. 13, 549-556) también conocido como presentación en ARNm. Se diseñaron librerías de partida de <sup>10</sup>Fn3 para aleatorizar las posiciones subrayadas indicadas en la **Figura 15A** usando trímeros de fosforamiditas (Glen Research). Todas las posiciones se aleatorizaron con una mezcla de trímeros que representa el 10 % de Tyr y cantidades iguales de todos los otros aminoácidos salvo Trp, Phe y Cys, que se omitieron de la mezcla. La exploración posterior identificó una adnectina anti-EGFR parental y una adnectina anti-IL-23 parental de interés. La adnectina anti-EGFR parental contenía una mutación R30Q, una posición que no estaba aleatorizada en el diseño de la biblioteca original. Después se generaron bibliotecas de optimización basadas en derivaciones parentales en las que cada bucle se realeatorizó con la mezcla anterior de codones mientras que se mantenían los otros dos bucles constantes. La PROfusión se llevó a cabo en estas bibliotecas de bucle único hasta que se recuperó la unión a las dianas, y entonces, los bucles aleatorizados se recombinaron seguido por rondas adicionales de PROfusión con concentraciones de diana más bajas para obtener Adnectina 1 (SEQ ID NO:67, EGFR8 en el Ejemplo 1) y Adnectina 2. La adnectina 1 contiene un bucle FG que es cinco aminoácidos más largo que el de <sup>10</sup>Fn3, que no se incluyó en el diseño de la biblioteca de optimización. Tanto la mutación R30Q inicial como la longitud

35

más larga del bucle FG probablemente se generaron debido a cualquier error durante las etapas de PCT de la PROfusión o en la síntesis de oligonucleótidos.

#### Ensayos de purificación y de actividad.

5 La expresión y la purificación de adnectinas anti-IL-23 y anti-EGFR fue análoga a la descrita anteriormente (Mamluk *et al.*, 2010, mAbs 2, 199-208). La inhibición de las actividades de EGFR y IL-23 mediante estas proteínas se midió mediante ensayos de competición de ELISA de unión de interleucina-23 y el receptor natural re IL-23 (IL-23R). Las placas Nunc Maxisorp (Thermo Fisher Scientific, Dinamarca) se recubrieron toda la noche con 50  $\mu$ l de IL-23R-Fc humana recombinante (R&D Systems, Minneapolis, MN), 4  $\mu$ g/ml en PBS a 4 °C. Las placas se lavaron con PBS que contiene Tween-20 al 0,05 % en p/v usando un lavador de placas automático (Biotek, VT). Se usó tampón OptEIA (BD Bioscience, CA) como agente de bloqueo y diluyente del ensayo. Las diluciones de adnectina que varían de 28 pM a 200 nM se preincubaron con IL-23 1 nM durante una hora antes de la transferencia a placas recubiertas con IL-23R-Fc bloqueado. Tras una incubación de 30 minutos se detectó IL-23 unido mediante un anti-IL-23 (GeneTex, CA) y anti-ratón-HRP (R&D Systems, MN) seguido de la adición de TMB (3,3',5,5'-tetrametilbenzidina) (BD Bioscience, CA). Se calculó la inhibición porcentual usando un patrón conocido de neutralización de adnectina IL-23 para definir la inhibición al 100 % y un patrón de no unión a adnectina como un control negativo. Se generaron  $CI_{50}$  a partir del promedio de cuatro ejecuciones con una aplicación de ajuste interno de curva. La inhibición de adnectina de la fosforilación de EGFR en tirosina 1068 se determinó usando un ensayo ELISA de células H292 *in vitro* tal como se describe en otra parte (Emanuel *et al.*, 2011, mAbs 3, 38-48).

#### Determinación por resonancia de plasmón superficial (RPS) de constantes de unión adnectina.

25 La  $K_D$  para la Adnectina 2 se determinó mediante resonancia de plasmón superficial (RSP) en un instrumento Biacore T100 (GE Healthcare, Piscataway, NJ), inyectando una serie de concentraciones de la adnectina en tres densidades de IL-23 humana inmovilizada en modo cinética de ciclo único sin regeneración de la superficie.

30 La  $K_D$  para la unión de adnectina 1 a EGFR-Fc recombinante (que contiene los aminoácidos 25 a 645 del ectodominio EGFR, R&D Systems) capturada sobre anticuerpo IgG antihumano de ratón (GE Healthcare) se evaluó mediante RPS. La IgG anti-humano se inmovilizó en células de flujo 1-4 de microplacas del sensor CM5 de acuerdo con las instrucciones del fabricante hasta un promedio de 7500-10000 RU. Todas las medidas cinéticas se llevaron a cabo en HBS-P (HEPES 10 mM, NaCl 150 mM, tensoactivo P20 al 0,05 %) a 37 ° C con  $MgCl_2$  3 M como solución de regeneración. La cinética de la asociación adnectina-EGFR se controló durante 250 segundos seguido por la disociación durante hasta 3000 segundos con concentraciones de adnectina de 0,78-100 nM. Se calcularon los parámetros cinéticos para ambos usando el programa informático Biacore T100.

#### Expresión y purificación de IL-23 humana.

40 Se creó una construcción bicistónica para expresar las subunidades p40 y p19 de IL-23 mediante la clonación de la subunidad p19 en el vector pFastBac Dual (Invitrogen) bajo el control del promotor PpH y la subunidad p40 usando hp19-pFastBac Dual bajo el control del promotor Pp10. Se expresó IL-23 humana en células Sf9 que secretan la proteína IL-23 al medio de crecimiento. El medio que contiene IL-23 se concentró y se intercambió el tampón en PBS o en solución salina tamponada con Tris usando la filtración por flujo tangencial. La IL-23 activa se purificó por afinidad a partir de este concentrado por medio de una nueva columna de afinidad de adnectina que consiste en proteína de adnectina anti-IL-23 purificada unida covalentemente mediante acoplamiento de amina primaria a resina de Sepharose 4 Fast Flow activada por CNBr (GE Healthcare) de acuerdo con las instrucciones del fabricante y empleando una incubación nocturna de enlaces a 4 ° C. Los medios concentrados, intercambiados con tampón se pasaron a través de una columna empaquetada con esta resina de afinidad a un caudal lineal de 20 cm/h. El lecho de la columna se lavó con cinco volúmenes de columna o solo con tampón. La IL-23 altamente purificada se eluyó con acetato 0,1 M, a pH 4,0, y el NaCl 1,0 M y el eluato se neutralizaron en cuanto a pH con 1/10 volúmenes de Tris HCl, a pH 8,0. La muestra se purificó adicionalmente usando una columna preparativa de cromatografía de exclusión por tamaño Superdex 200 (SEC) equilibrada y procesada con tampón HBS.

#### Purificación de EGFR.

55 El EGFR humano (restos 1-642) con un marcador His en C-terminal se expresó en células Sf9. El medio secretado que contiene EGFR humano se concentró y se intercambió el tampón en Tris-HCl 25 mM, a pH 8,0, y NaCl 250 mM, glicerol al 5 % (v/v) usando la filtración de flujo tangencial. El EGFR humano se purificó mediante cromatografía Ni-NTA seguido por la cromatografía de exclusión por tamaño en una colima Superdex 200 (GE Healthcare) y se combinaron las fracciones que se corresponden con el monómero de EGFR.

#### Preparación purificación de los complejos EGFR/Adnectina 1 y IL-23/adnectina 2.

65 Se mezcló EGFR humano con adnectina anti-EGFR a una proporción molar de 1:6 y se incubó en hielo durante 2 horas. Se purificó un complejo EGFR: Adnectina 1 a 1:1 usando SEC sobre una columna Superdex 200 equilibrada y procesada en HEPES 25 mM, a pH 7,5 y NaCl 200 mM. El complejo purificado se concentró usando un concentrador

de corte Vivaspin de 5 kDa hasta 20 mg/ml.

El complejo IL-23 humana/Adnectina 2 se purificó de manera análoga con las siguientes modificaciones usando IL-23 purificada tal como se describe anteriormente: La IL-23 y la adnectina 2 se mezclaron en una proporción molar de 1:3 y se incubaron toda la noche a 4 °C; el complejo final se concentró hasta 12 mg/ml.

#### Cristalización de complejos proteicos.

El complejo EGFR/adnectina 1 se cristalizó a 20 °C usando el método de difusión de vapor de gota colgante mezclando 1 µl de complejo proteico con 1 µl de solución de reservorio que contiene Tacsimato al 60 % (Hampton Research, Aliso Viejo, CA). El complejo IL-23/Adnectina 2 con un exceso molar de una vez de adnectina 2 se cristalizó del mismo modo a 20 °C, pero se usó citrato de trisodio 1 M, NaCl 0,2 M, Tris 0,1 M, a pH 7,0 para la adnectina 2. La calidad del cristal del complejo IL-23/adnectina 2 se mejoró usando la siembra de cristales.

#### Recolección y procesamiento de datos.

Los datos para el complejo EGFR/adnectina 1 se recogieron en la línea de luz 21ID-G en LS-CAT en Advanced Photon Source en Argonne National Laboratory. La longitud de onda usada fue de 0,979 Å y el detector fue un Rayonix MX-300. Los datos se indexaron, se integraron y se escalaron con HKL2000 (Otwinowski y Minor, 1997, *Methods Enzymol.* 276, 307-326). Los datos para IL-23/adnectina 2 se recogieron en la línea de luz 17ID en IMCA-CAT en Advanced Photon Source en Argonne National Laboratory. La longitud de onda usada fue de 1,0 Å y el detector fue un MAR 165 CCD. Los datos se indexaron, se integraron y se escalaron con D\*TREK (Pflugrath, 1999, *Acta Crystallogr. D Biol. Crystallogr.* 55, 1718-1725). El grupo espacial, por parámetros de la celda unidad y las estadísticas de la colección de datos para ambos conjuntos de datos se enumeran en la **Tabla 7**.

#### Reemplazo molecular.

Un modelo para las adnectinas provino de PDB 1FNF delecionando los bucles BC, EF y FG. El modelo de EGFR se basó en dominios de 1NQL. El modelo para la IL-23 fue una estructura determinada en una forma cristalina diferente de las publicadas en la bibliografía (3DUH, 3D87). PHASER (McCoy *et al.*, 2007, *J. Appl. Crystallogr.* 40, 658-674) se usó para el reemplazo molecular. Cuando PHASER falló al encontrar la adnectina en el complejo IL-23/adnectina 2, se usó con éxito una búsqueda hexadimensional usando la función de traducción de AMoRe (Sheriff *et al.*, 1999, *J. Appl. Crystallogr.* 32, 98-101; Navaza y Vernoslova, 1995, *Acta Crystallogr. Sect. A* 51, 445-449; y CCP4, 1994, *Acta Crystallogr. D Biol. Crystallogr.* 50, 760-763.).

#### Construcción del modelo, refinamiento y análisis.

COOT (Emsley *et al.*, 2010, *Acta Crystallogr. D Biol. Crystallogr.* 66, 486-501) se utilizó para la construcción de modelos. El refinado se llevó a cabo con autoBUSTER (BUSTER, versión 2.13.0. Cambridge, Reino Unido: Global Phasing Ltd.). Las estadísticas del refinado se enumeran en la **Tabla 7**. La densidad electrónica inicial para los bucles diversificados con el modelo final se muestran en la **Figura 16**. Los gráficos de la presentación se produjeron con PyMOL v1.4 (Schrödinger, LLC). El área de superficie interna se calculó con el programa MS (Connolly, 1983, *J. Appl. Crystallogr.* 16, 548-558) usando una esfera de sonda de 1,7 Å, los restos del contacto se enumeraron tal como se define por Sheriff, 1993 (*Immunomethods* 3, 191-196), Sheriff *et al.*, 1987 (*J. Mol. Biol.* 197, 273-296), y ambos usan radios atómicos extendidos tal como se define por Gelin y Karplus, 1979 (*Biochemistry* 18, 1256-1268).

#### Estimaciones de energías libre de restos y energías de interacción.

Los modelos atómicos de los complejos se optimizaron utilizando el flujo de trabajo del asistente de preparación de proteínas en MAESTRO 9.0.211 (Schrödinger, LLC. 2009). La estimación de energía libre de Gibbs se calculó tal como se describe anteriormente (Novotny *et al.*, 1989, *Biochemistry* 28, 4735-4749; Krystek *et al.*, 1993, *J. Mol. Biol.* 234, 661-679) y se implementaron en una secuencia de comandos usando MAESTRO (Maestro, versión 9.0, Schrodinger, LLC, Nueva York, NY, 2009).

#### Cálculos energéticos

Las **Tablas 8-10** contienen las energías calculadas para todos los restos mutantes y de tipo silvestre. Para la mutagénesis de barrido de alanina, las energías de interacción y las estimaciones de energía libre se compararon con los valores experimentales para la unión de diana de adnectina. El cambio en la energía de interacción del tipo silvestre al mutante se documenta en las **Tablas 3-5** junto con los datos experimentales. Para todos los restos de contacto en cada complejo se crearon modelos individuales de homología sustituyendo todos los aminoácidos en cada posición de contacto. Para cada uno de los modelos se analizaron los cálculos energéticos de la energía de interacción y las estimaciones de energía libre centrándose en los aminoácidos que se corresponden con las posiciones de contacto en las adnectinas. Los valores calculados para los modelos seleccionados se presentan en as **Tablas 8-10**.

#### Estimaciones de energías libre de restos y energías de interacción.

Los modelos atómicos de los complejos se optimizaron utilizando el flujo de trabajo del asistente de preparación de proteínas en MAESTRO 9.0.211 (Schrödinger, LLC. 2009). Durante este proceso, los estados de protonación de la cadena lateral, los tautómeros de histidina y los rotámeros Chi terminales para las cadenas laterales de histidina, asparagina y glutamina se optimizan. La última etapa en el flujo de trabajo es la minimización restringida del complejo (desviación cuadrática media de 0,3 Å) que permite la optimización sutil del complejo dentro del campo de fuerzas OPLS\_2005. Se crearon modelos de proteína para cada proteína mutante usando el protocolo de refinado de cadena lateral PRIME seguido por dos etapas de minimización. La primera minimización se aplicó solo a las cadenas laterales para el subconjunto de restos que estaban a 5Å de un sitio de mutación. La etapa de minimización final se aplicó al mismo subconjunto de restos pero incluyó la estructura principal de los restos en la minimización.

La estimación de la energía libre de Gibbs se calculó tal como se describe anteriormente y se implementó en una secuencia de comandos usando MAESTRO (Maestro, versión 9.0, Schrodinger, LLC, Nueva York, NY, 2009.). Las energías de interacción de los restos se determinaron usando el campo de fuerzas OPLS\_2005 tal como se implementa en el comando de interacciones del componente (Schrödinger, LLC) usando Macromodel (Macromodel, versión 9.7, Schrodinger, LLC, Nueva York, NY, 2009.). La secuencia de comandos calcula la energía de interacción de la mecánica molecular entre un conjunto de restos y emite los términos VDW individuales y la contribución electrostática. Para el componente electrostático, se utilizó un dieléctrico dependiente de la distancia con una constante de 4,0, similar a los cálculos de energía libre.

#### Números de registro.

Las coordenadas y las amplitudes de la estructura se han depositado en el banco de datos de proteínas RCSB con los códigos de identificación 3QWQ (EGFR/Adnectina 1) y 3QWR (IL-23/Adnectina 2).

## RESULTADOS

### Identificación de adnectinas antagonistas de EGFR y de IL-23.

Las adnectinas que se unen a y bloquean la actividad de bien EGFR (Adnectina 1) o IL-23 (Adnectina 2) se identificaron usando la técnica de selección bioquímica de presentación en ARNm en la que una proteína se une de manera covalente a sus secuencias codificantes de ácido nucleico (Xu *et al.*, 2002, Chem. Biol. 9, 933-942; Getmanova *et al.*, 2006, Chem. Biol. 13, 549-556; Roberts y Szostak, 1997, Proc. Natl. Acad. Sci., EE.UU. 94, 12297-12302) (**Figura 15A** y Procedimientos experimentales). La adnectina 1 se unió a la fusión de Fc del ectodominio EGFR con una  $K_D$  de 2 nM e inhibió la fosforilación de EGFR inducida por EGF en células H292 con una  $CI_{50}$  de aproximadamente 50 nM. La adnectina 2 se unió a IL-23 inmovilizada con una  $K_D$  de 2 nM y compitió con la interacción IL-23/IL-23R con una  $CI_{50}$  de 1 nM en un ensayo de competición de unión del receptor bioquímico. La adnectina 1 se cocrystalizó con EGFR y la adnectina 2 se cocrystalizó con IL-23 para determinar la base estructural de estas actividades.

### Descripción general de la estructura del complejo EGFR/Adnectina 1.

La EGFR es una molécula notablemente flexible con diferentes interacciones entre los cuatro dominios. Las estructuras depositadas en la PDB (1IVO, 1NQL, 1YY9) muestran que los dominios I y III son relativamente rígidas mientras que partes de los dominios II y IV pueden adoptar múltiples conformaciones que orientan el dominio III de manera diferente con respecto al dominio I (Ogiso *et al.*, 2002, Cell 110, 775-787; Ferguson *et al.*, 2003, Mol. Cell 11, 507-517; Li *et al.*, 2005, Cancer Cell 7, 301-311). La estructura de EGFR en el complejo EGFR/adnectina 1 se asemeja más a 1NQL, que es un complejo de EGFR con EGF a bajo pH y se considera que es una forma inactiva del receptor. Sin embargo, la superimposición del dominio I de los dos complejos lleva a los centros del dominio III de los dos complejos a estar separados por aproximadamente 25 Å, mostrando una vez más la notable flexibilidad de EGFR. La adnectina unida al dominio EGFR I (**Figura 13A**) solapando el sitio de unión de EGF en el dominio I de EGFR en sus formas activa (1IVO) o inactiva (1NQL) y, por lo tanto, unión de EGF con impedimento estérico (**Figura 13D, 17D**). Este sitio es un sitio radicalmente diferente de los anticuerpos que se unen a EGFR en el dominio III. Cetuximab (1YY9) y necitumumab (3B2V) se unen esencialmente al mismo sitio (**Figura 13A**), pero están orientados de manera diferente, y matuzamab (3C09) se une a un sitio distinto en el dominio III (Li *et al.*, 2005, Cancer Cell 7, 301-311; Li *et al.*, 2008, Structure 16, 216-227; Schmiedel *et al.*, 2008, Cancer Cell 13, 365-373).

La adnectina 1 descrita en el presente documento tenía una inserción de 5 restos en el bucle FG. Para mantener un esquema de numeración constante, a los restos insertados en el bucle FG se les dio letras de inserción (**Figura 15A**) análogas a las ideadas para las inmunoglobulinas por Kabat *et al.*, 1991 (Sequences of Proteins of Immunological Interest, 5ª ed. (Bethesda: National Institutes of Health)), y similares a las usadas por Gilbreth *et al.*, 2008 (J. Mol. Biol. 381, 407-418) para el bucle BC.

### Interacciones específicas de adnectina 1 con EGFR.

La interacción entre adnectina 1 y el dominio I de EGFR dio como resultado aproximadamente 520 Å<sup>2</sup> sobre la adnectina y aproximadamente 590 Å<sup>2</sup> sobre el dominio I de EGFR que se introdujeron mediante la interacción (**Figura**

**13C, 17C).** El tamaño de estas superficies de interacción fue hacia el extremo más pequeño típico de las superficies de contacto de anticuerpo/proteína antigénica (Sheriff, 1993, *Immunomethods* 3, 191-196). También fueron más pequeñas que las vistas para otras variantes basadas en <sup>10</sup>Fn3 (**Tabla 3**; Gilbreth *et al.*, 2008, *J. Mol. Biol.* 381, 407-418; Gilbreth *et al.*, 2011, *Proc. Natl. Acad. Sci., EE.UU.* 108, 7751-7756; Huang *et al.*, 2009, *J. Mol. Biol.* 392 1221-1231; Wojcik *et al.*, 2010, *Nature Struct. Mol. Biol.* 17, 519-527), pero las afinidades por adnectina 1 y los anticuerpos estudiados eran del mismo orden de magnitud. El tamaño relativamente pequeño de la interacción se debe a la superficie convexa en EGFR que consiste en bucles que conectan las dos láminas β en un borde de un dominio de sándwich β que interactúa con la superficie convexa de la adnectina (**Figura 13C**). No obstante, el estadístico Sc (Lawrence y Colman, 1993, *J. Mol. Biol.* 234, 946-950), que es una medida de la complementariedad de las superficies de unión, para este complejo fue 0,71, que cae hacia el extremo más bajo del intervalo de proteasa/inhibidores de proteasa (0,71-0,76) y de las superficies de contacto oligoméricas (0,70-0,74) y por encima del intervalo observado para los complejos anticuerpo/antígeno (0,66-0,68) (Lawrence y Colman, 1993, citado anteriormente) y ligeramente por encima de la media de los complejos dominio basado en <sup>10</sup>Fn3/proteína (**Tabla 3**). Esto sugiere que aunque la superficie de interacción fue relativamente pequeña para el complejo adnectina 1/EGFR, fue un ajuste más complementario que el visto con anticuerpos y sus antígenos.

Las principales interacciones de la adnectina tuvieron lugar a través del bucle FG (aproximadamente 175 Å<sup>2</sup> de superficie interior), la cadena D (-175 Å<sup>2</sup>), el bucle DE (-70 Å<sup>2</sup>), la cadena C (-50 Å<sup>2</sup>) y el bucle BC (-40 Å<sup>2</sup>) (**Figura 13C, 17C**). Los restos de adnectina con contactos de van der Waals (VDW) que dependen de los radios (Sheriff, 1993; Sheriff *et al.*, 1987a) para EGFR fueron: **Bucle BC:** Gln30; **Cadena C:** Tyr31; **Cadena D:** Glu47, Phe48, Thr49, Val50, Pro51; **Bucle DE:** Val54; **Bucle FG:** Asp77, His78, Lys79, Ala79C, His81 (**Figuras 13B, C, E, 17C, 18D, 18E**). Una característica de esta superficie de contacto fue que las extensas interacciones entre la secuencia de la estructura principal de tipo silvestre de la cadena D en la adnectina y EGFR contribuyeron con tanta superficie y tantos restos como la interacción del bucle diversificado FG. Los restos aromáticos (His, Phe, Tyr) constituyeron aproximadamente 1/4 de la superficie total de interacción con adnectina (-120 Å<sup>2</sup>), pero solo estaba implicada una tirosina (31), que era parte del armazón de adnectina y cuya cadena lateral estaba paralela a la superficie EGFR bastante plana. El extremo N-terminal y la mayoría de los bucles BC y DE estaban en el lado distal de la adnectina de EGFR y, por lo tanto, no estaban implicados en la interacción. Tuvo lugar una interacción de tipo lámina β entre la adnectina 1 y la región N-terminal de E, que presentaba enlaces de hidrógeno de 3 N...O=C y una cadena lateral para el enlace de hidrógeno de cadena lateral (**Figura 13E**).

#### Mutagénesis de barrido de alanina de EGFR/Adnectina 1.

Los mutantes de alanina de sitio único (Wells, 1991, *Methods Enzymol.* 202, 390-411) se crearon para 30 de los 101 restos de la adnectina. Todos los restos en los bucles BC, DE y FG que se diversificaron como parte del proceso de selección se incluyeron, así como cualquier otro resto que estaba en contacto (véase anteriormente) con EGFR (**Tabla 4, Tabla 8**). La mutación de los siguientes residuos condujo a una unión significativamente disminuida: Tyr 29, Gln 30, Tyr 31, Gly 52, Val 54, Asp 77, Lys 79, His 81 y Tyr 83. Sorprendentemente, Tyr 29, Gly 52 y Tyr 83 no interactúan directamente con EGFR. Esto se puede explicar basándose en las interacciones estructurales en la adnectina. Se puede esperar que la cadena lateral de Tyr29 empaquetada en el interior del bucle BC y frente a la cadena F y la mutación a Ala alteren el bucle BC. Dado que Gly52 era parte de un giro de Pro-Gly-Pro, el cambio de Gly a Ala podría haber alterado este elemento estructural. La cadena lateral de Tyr83 formó una interacción borde frente a cara con la cadena lateral de His81 y un enlace de hidrógeno con la cadena lateral de Glu85. La cadena lateral de His81 a su vez interactúa con la Thr15 de EGFR. La actividad disminuida resultante de la mutación Tyr83 → Ala puede deberse a que ayuda a mantener la posición de His81.

Antes de la mutagénesis de barrido de alanina, la energía de la interacción se calculó basándose en los restos. Estos cálculos identificaron correctamente Gln 30, Tyr 31 y Asp 77 como importantes para la interacción, pero también identificaron que Pro 51 es importante, aunque la mutación a Ala no afectó a la unión (**Tabla 4, Tabla 8**). Los restos restantes que interactuaron con EGFR y que se identificaron mediante barrido de alanina, Val 54, Lys 79 e His 81, todos mostraron cambios en la dirección correcta (**Tabla 4**), pero la magnitud del cambio no alcanzó el umbral considerado significativo (3 kcal/mol).

#### Mutagénesis de restos de contacto de EGFR/adnectina 1.

Seis restos, Tyr31, Glu47, Thr49, Pro51, Thr58 y Ala79C, se direccionaron para estudios adicionales de mutación basados en las predicciones de energía de interacción calculada a partir de modelos para explorar la posibilidad de que las mutaciones pudieran aumentar la afinidad de unión a adnectina 1. La mayoría de las mutaciones, que normalmente eran intentos de crear enlaces de hidrógeno o interacciones cargadas, no cambiaron sustancialmente la afinidad en comparación con el parental (**Tabla 5, Tabla 9**). Sin embargo, las tres mutaciones al resto 49 para aumentar las interacciones de van der Waals convirtiendo una Thr a una Val, Ile o Tyr, aumentaron la afinidad de la adnectina de 8 a 40 veces (**Tabla 5, Tabla 9**). Dado que la cadena lateral de la Thr49 no estaba en contacto con EGFR, la mutación a una cadena lateral mayor (Ile, Tyr) podría haber creado contactos hidrófobos adicionales con restos de EGFR localizados en la depresión superficial cerca de Thr 49 (**Figura 13C**). En el caso de Thr49→Val, la sustitución de un metilo por el hidroxilo que apuntaba hacia un ambiente hidrófobo, puede explicar la mayor afinidad.

Se mutó Tyr 31 a Ser, Leu y Phe para ver la predicción de que los restos pequeños (Ser) llevarían a una reducción de la afinidad, pero los restos más grandes tales como Leu y Phe, que posiblemente conservan las interacciones de VDW podrían tener un efecto favorable sobre la unión. Phe y, sorprendentemente, Ser tuvieron un efecto pequeño, pero la Leu, llevó a una unión mucho más débil a pesar de los cálculos energéticos y de modelado que sugerían que sería muy similar al parental.

La cadena lateral de Glu 47 no estaba implicada en las interacciones con EGFR, pero el modelado mostró que estaba lo suficientemente cerca como para que una mutación a Arg o Lys posiblemente formase un puente de sal con Glu 90 de EGFR. Aunque los cálculos energéticos sugirieron que este puente de sal sería favorable, experimentalmente ni afectó de forma material a la unión, ni hizo una mutación a Asp, que cambiase la posición relativa de la carga.

Pro 51 fue el primer resto del bucle DE y estaba en el contacto de VDW de la cadena lateral de Leu 69 de EGFR y los átomos de oxígeno del carbonilo de la cadena principal de Ser 99 y Tyr 101 de EGFR. El modelado mostró que Leu posiblemente mejora los contactos de VDW o que Thr formó un enlace de hidrógeno. La Leu en la posición 51 no fue significativamente diferente que el parental, pero Thr llevó a una unión débil, quizás porque se alteró la estructura Pro 51-Gly 52-Pro 53 al comienzo del bucle DE.

La cadena lateral de Thr 49 no tuvo contactos significativos con EGFR, pero el modelado sugirió la mutación a Ile, Val o Tyr, puede permitir que estos restos mantengan los enlaces de hidrógeno de la estructura principal y establecer interacciones adicionales de empaquetamiento con los restos Leu 14 y Leu 69 de EGFR que forman un parche de superficie hidrófoba. De las tres mutaciones, se predijo que Tyr sería energéticamente favorable en comparación con el parental, mientras que se predijo que la Val y la Ile eran similares al parental. No obstante, la mutación de Thr 49 a Val o a Ile presentó un aumento de 12 y de 37 veces respectivamente, en la afinidad de unión mientras que la mutación a tirosina produjo un aumento de 8 veces en la afinidad de unión.

La Thr 58 no estaba en contacto con EGFR en la estructura cristalina, pero el modelado sugirió que la sustitución con un resto ácido podría llevar a un puente de sal con Lys 105 aunque el átomo NZ no se colocase en el modelo cristalográfico debido a una densidad de electrones inadecuada. A pesar de la predicción de grandes cambios favorables en la energía de interacción, ninguna de las tres sustituciones (Glu, Gln, o Asp) llevó a ningún cambio significativo en la actividad. Sin embargo, Las sustituciones en esta posición dieron como resultado una solubilidad mejorada de la proteína, tal como se describe en el presente documento (véase el **Ejemplo 1**).

Ala 79C estaba de forma marginal, en contacto con EGFR, pero el modelado mostró que un resto más largo podría llevar a interacciones adicionales y los cálculos energéticos sugirieron un aumento considerable en la afinidad de unión para cinco mutaciones (Arg, Asn, Glu, Leu y Tyr). Sin embargo, las cinco mutaciones mostraron una ligera, pero insignificante, disminución de la actividad.

### Descripción general de la estructura del complejo IL-23/Adnectina 2.

La IL-23 es una proteína de dos subunidades que consiste en una subunidad p40 que comparte con la IL-12 y una subunidad p19 que es distinta de la subunidad p35 de la IL-12. La subunidad p40 consiste en tres dominios de láminas  $\beta$  de 7 cadenas de tipo Ig, mientras que la p19 consiste en un paquete de 4 hélices. La adnectina 2 se unió a la unión de las subunidades p40 y p19 y produjo interacciones considerables con ambas subunidades, incluidos los dominios 2 y 3 de la subunidad p40 (**Figuras 14A, C, D, E, 18C y 18E**). A pesar de las interacciones con p40, la adnectina 2 no inhibió la unión o la señalización de IL-12 (datos no mostrados). Además, aunque los bucles diversificados estaban hacia el centro de la superficie de contacto, las interacciones se extendieron junto con las cadenas  $\beta$  lejos de los bucles diversificados e incluyeron el bucle CD en el extremo opuesto de la molécula. Este sitio cóncavo es probablemente inaccesible para los Fv que son mucho más grandes, que consisten en dos dominios de subunidades separadas y seis bucles hipervariables. De hecho, el sitio de unión de adnectina era considerablemente distinto al del complejo de anticuerpos conocido para IL-23 (PDB 3D85), que se une solo a la subunidad p19 (Beyer *et al.*, 2008, J. Mol. Biol. 382, 942-955) (**Figura 14A**).

### Interacciones específicas de adnectina 2 con IL-23.

La interacción entre adnectina 2 y la IL-23 fue bastante grande, penetrando aproximadamente 1320 Å<sup>2</sup> sobre la superficie de adnectina y aproximadamente 1370 Å<sup>2</sup> sobre la superficie de IL-23 (**Figuras 14C, 18C**). Esta cantidad de área de superficie interna fue más grande que la mayoría de las interacciones anticuerpo/antígeno y, puede reflejar la naturaleza cóncava del sitio de unión sobre IL-23. La superficie interna fue también mucho más grande que para cualquier otro complejo de dominio basado en <sup>10</sup>Fn3 (**Tabla 3**). A pesar de la gran superficie de interacción, la afinidad de la adnectina 2 por IL-23 era del mismo orden de magnitud que los anticuerpos para sus antígenos de proteína. El estadístico Sc para este complejo fue de 0,73, lo que sugirió que era más complementario que los complejos anticuerpo/antígeno estudiados por Lawrence y Colman, 1993 (J. Mol. Biol. 234, 946-950) y el segundo más largo de los complejos de dominio basado en <sup>10</sup>Fn3/proteína (**Tabla 3**).

Las principales interacciones tuvieron lugar a través de los bucles FG (de aproximadamente 610 Å<sup>2</sup>) y BC (de aproximadamente 380 Å<sup>2</sup>), pero la mayoría de los segmentos de la estructura secundaria tenían al menos algo de

área de superficie interna por la interacción (**Figuras 14C, 18C**). Se ha descubierto que los siguientes restos de adnectina están en contacto con IL-23: Región de N-terminal: Pro5, Arg6, Asp7, **Bucle BC**: Glu23, His24, Asp25, Tyr26, Pro27, Tyr28, Arg30, **Cadena C**: Tyr31, Arg33; **Bucle CD**: Gly40, Asn42, Val45; **Cadena F**: Tyr73, Val75; **Bucle FG**: Thr76, Ser77, Ser78, Tyr79, Lys80, Tyr81, Asp82, Met83, Gln84, Tyr85, Pro87 (**Figuras 14B-E, 18C-E**). De esta lista destacaron cuatro puntos. En primer lugar, el número de restos de interacción fue grande y vinieron de muchas de las cadenas  $\beta$  y de los bucles. En segundo lugar, no hubo contactos entre el bucle DE diversificado e IL-23. En tercer lugar, un gran número (7) de restos de tirosina estaba implicado en la interacción. La frecuente aparición de tirosina se ha observado para los anticuerpos, los fragmentos VH y las variantes basadas en  $^{10}\text{Fn3}$  que interactúan con antígenos (Padlan, 1990, *Proteins* 7, 112-124; Mian *et al.*, 1991, *J. Mol. Biol.* 217, 133-151; Kosiakoff y Koide, 2008, *Curr. Opin. Struct. Biol.* 18, 499-506; Koide y Sidhu, 2009, *ACS Chem. Biol.* 4, 325-334), y es presumiblemente debido a la pérdida relativamente baja de entropía debido a relativamente pocos ángulos diedros que se inmovilizan en comparación con la gran superficie que los residuos de tirosina pueden contribuir, que asciende a un total de aproximadamente  $450 \text{ \AA}^2$  en este caso. Además, varias de estas tirosinas (20, 79, 81, 85) parecían encajar en grietas en la superficie de IL-23 (**Figuras 14D-E**). En cuarto lugar, un gran número (11) de restos no diversificados estaban implicados en interacciones directas con IL-23. Los restos en esta cuarta categoría incluyeron 2 de los 7 restos de Tyr y restos en el extremo N-terminal. Aunque la densidad de electrones solo fue interpretable para parte del extremo N-terminal, estaba claro que el extremo N-terminal no apunta en la dirección de los bucles BC, DE y FG, y lo hacía en el  $^{10}\text{Fn3}$  de tipo silvestre, sino que más bien invertía la dirección y apuntaba hacia el extremo opuesto de la molécula.

## 20 Mutagénesis de restos de contacto de IL-23/adnectina 2.

Cuatro aminoácidos, Tyr28, Tyr73, Tyr81 y Pro87 se mutaron a alanina para demostrar que los restos energéticamente importantes se podrían predecir. Este resultó ser el caso de Tyr28, Tyr81 y Pro87, pero no el de Tyr73, que tuvo un pequeño efecto cuando se mutó a alanina (**Tabla 6**). Tyr28 se localizaba en el centro del bucle BC y forma contactos significativos con los aminoácidos que estaban en el extremo terminal de la hélice A del dominio p19 de la IL-23, por ejemplo, interacciones borde frente a cara con Trp26 e His29. De manera similar, la Tyr81 que se localizaba en el centro del bucle FG tenía contactos significativos con la subunidad p40 de la IL-23, por ejemplo, Ser204. La Pro87, que se localizaba en el extremo C-terminal del bucle FG, se puede requerir para mantener la conformación del bucle FG y estaba en contacto con los restos Gly100 y Pro101 de la subunidad p40 de IL-23. Aunque se predijo que Tyr73 contribuía en aproximadamente 6 kcal (**Tabla 10**) a la interacción, esto fue menos de la mitad que se predijo para la Tyr28 (aproximadamente 16 kcal) y para Tyr81 (aproximadamente 13 kcal) (**Tabla 10**). En la estructura minimizada, la cadena lateral de Tyr73 formó un enlace de hidrógeno con el oxígeno del carbonilo de Lys99 de la subunidad p40 de IL-23, pero el mutante Tyr73→Ala demostró que no era un resto energético clave.

## 35 Interacciones de restos de contacto mutados en el complejo IL-23/ adnectina 2

Algunas mutaciones se hicieron en un intento para mejorar la afinidad mediante interacciones específicas en el complejo IL-23/adnectina 2. Cuatro posiciones en ATI-000929, Thr 35, Val 45, Tyr 73 y Val 75, que estaban todas sobre la misma cara del pliegue del sándwich  $\beta$  (cadenas  $\beta$  C, D, F y G) de la adnectina, se seleccionaron para la mutagénesis en un intento de mejorar la afinidad a través de interacciones con la subunidad p40 (**Tabla 6**). La Thr 35 y la Val 45 estaban cerca entre sí y se mutaron en un intento de crear una interacción con Lys 99 de la subunidad p40 de la IL-23. En cada posición, se hicieron de forma individual las cuatro mismas mutaciones (asparagina, glutamina, ácido aspártico y ácido glutámico), con la esperanza de formar interacciones electrostáticas con Lys 99. Los restos ácidos se favorecieron especialmente mediante modelado. Sin embargo, no se vio mejoras en la unión en ninguna posición. El mutante Thr 35→Asp mostró una reducción significativa en la unión. Esto se puede deber a interacciones desfavorables que podría tener la Asp 35 con Glu 49 de la adnectina. Se mutó la Val 75 en un intento para formar un enlace de hidrógeno con el oxígeno del carbonilo de Pro 101 de la subunidad p40 de IL-23. Ninguna mutación de este resto alteró la unión. Tyr 73, como se ha señalado anteriormente, está en contacto con los restos 99-101 de la subunidad p40 de la IL-23 así como forma un enlace de hidrógeno con el oxígeno carbonilo de Lys 99. El modelado de Tyr 73 sugirió que un mutante de arginina podría formar un enlace de hidrógeno adicional con Glu 100 y elevados contactos con los restos 99-101 de la subunidad p40 de la IL-23. Tres mutaciones en esta posición no tuvieron efecto, pero Tyr 73→Gln llevó a un modesto, aumento en la actividad de aproximadamente 4 veces.

## 55 Comparación de las estructuras de adnectina 1 y adnectina 2 con $^{10}\text{Fn3}$ .

Las comparaciones estructurales demostraron que la molécula de tipo silvestre ( $^{10}\text{Fn3}$ ; 1FNF restos 1416-1509) tiene una topología muy similar a la de adnectina 1 y adnectina 2 cuando está unido a sus moléculas diana (**Figuras 14 y 19**), incluyendo una excelente superposición de la lámina  $\beta$  central y dos de los tres bucles (AB y DE) distales de los que se diversificaron para la selección (**Figuras 14A y 19A**). Los bucles BC y DE de la adnectina 1 y de la adnectina 2 eran idénticos en longitud al tipo silvestre. En estas estructuras, el bucle DE corto mostró una variación mínima, mientras que el bucle BC presentó más variación cuando se comparó con la estructura  $^{10}\text{Fn3}$ . La cadena lateral de Lys52 en el bucle DE de adnectina 2 estaba lo suficientemente cerca al bucle BC como para que pueda estar implicado en la estabilización de la posición desplazada de ese bucle en comparación con el  $^{10}\text{Fn3}$  de tipo silvestre. Estas grandes variaciones fueron en el bucle de FG, donde, en la estructura cristalina 1FNP, el motivo RGD natural estaba implicado en un contacto cristalino y que ese contacto era probablemente responsable de su orientación en esa estructura. El bucle FG de adnectina 1 fue 5 restos más larga que la de cualquiera de  $^{10}\text{Fn3}$  o adnectina 2 y las cadenas

F y G se extendieron. Sin embargo, los residuos implicados en las cadenas  $\beta$  extendidas se mostraron como tubos en lugar de flechas en la **Figura 15** para destacar los restos diversificados. En la adnectina **2**, el bucle adoptó todavía una conformación diferente cuando se unió a IL-23. Finalmente, en el extremo N-terminal fue flexible en adnectinas. En  $^{10}\text{Fn3}$ , la posición se dicta mediante el enlace a  $^9\text{Fn3}$ . En la adnectina 1, el extremo N-terminal tenía una conformación relativamente similar dado que no interactuó con EGFR. Por otro lado, el extremo N-terminal de adnectina **2** estaba plegado hacia afuera en comparación con los otros dos extremos de N-terminal, quizás para evitar la colisión con IL-23. Por lo tanto, estas estructuras muestran que los bucles de adnectina pueden adoptar conformaciones diferentes de las del tipo silvestre dependiendo de la interacción proteína/proteína.

#### 10 Datos cristalográficos.

El grupo espacial, por parámetros de la celda unidad y las estadísticas de la colección de datos para ambos complejos se enumeran en la Tabla 7. La densidad de electrones inicial para los bucles diversificados, que se excluyen del modelo de reemplazo molecular, se muestra para el modelo final para ambos complejos en la Figura 16.

**Tabla 3. Comparación de área de superficie interior, contactos y complementariedad de superficie para complejos de dominio basados en  $^{10}\text{Fn3}$**

ID de PDB	Proteína	Par	Proteína diana			Dominio basado en $^{10}\text{Fn3}$			Tipo de contactos			Sc
			Área, A <sup>2</sup>	N.º de res.	N.º de átomo	Área, A <sup>2</sup>	N.º de res.	N.º de átomo	N.º de enlaces de H	N.º de enlaces de sal	N.º de interacciones de VDW	
20CF	RE $\alpha$		900	22	47	870	17	51	6	2	77	0,66
3CSB	PUM		650	16	59	600	14	51	7	0	121	0,70
3CSG	PUM		750	15	56	680	15	53	11	1	125	0,64
3K2M	SH2	AD	570	16	47	580	10	48	11	2	105	0,72
		BC	570	13	48	530	12	49	11	3	109	0,76
3QHT	ySUMO	AC	600	13	48	560	12	43	8	0	87	0,71
		BD	610	12	46	570	13	35	9	0	74	0,67
3QWQ	EGFR		585	14	39	520	13	41	8	0	75	0,71
3QWR	IL-23		1370	38	97	1320	29	99	16	4	197	0,73

Proteínas: RE $\alpha$  - receptor de estrógenos  $\alpha$ ; PUM = proteína de unión a maltosa; SH2 - homología 2 de *Src* para cinasa Abelson; ySUMO - modificador relacionado con ubiquitina pequeña de levadura

Pares: En casos de múltiples complejos por unidad asimétrica todos están tabulados. Las letras indican los nombres de las cadenas en los pares.

Área: calculada mediante el método de Connolly, 1983 y redondeada a los 10 A<sup>2</sup> más cercanos.

N.º de res, N.º de átomo: Número de restos y número de átomos en contacto tal como se calcula mediante el método de Sheriff *et al.*, 1987a; Sheriff, 1993.

N.º de enlaces de H, N.º de enlaces de sal, N.º de interacciones VDW: Número de enlaces de hidrógeno por pares de átomos e interacciones de van der Waals tal como se calcula mediante el método de Sheriff *et al.*, 1987a; Sheriff, 1993. El número de enlaces de sal se tabula por resto.

Sc: Complementariedad de superficie calculada mediante el método de Lawrence y Colman, 1993

**Tabla 4. Barrido de alanina en EGFR**

Mutación	Elemento estructural secundario	K <sub>D</sub> , nM	Actividad (parental/mutante)	Cl <sub>50</sub> de pEGFR (parental/mutante)	$\Delta$ Energía de interacción, kcal/mol
Parental		1,8	1	1	0
D23→A	Bucle BC	2,4	0,7	0,3	0
S24→A	Bucle BC	1,3	1,4	0,7	0
G25→A	Bucle BC	2,0	0,9	0,3	0
R26→A	Bucle BC	2,0	0,9	0,8	0
G27→A	Bucle BC	1,1	1,6	1,4	0
S28→A	Bucle BC	1,5	1,2	0,5	0
Y29→A	Bucle BC	14	0,1	0,02	0
Q30→A	Bucle BC	>50	<0,04	N.E.	<b>3</b>
Y31→A	Cadena C	N.D.		N.E.	<b>5</b>
F48→A	Cadena D	3,9	0,5	0,1	1

(continuación)

Mutación	Elemento estructural secundario	K <sub>D</sub> , nM	Actividad (parental/mutante)	CI <sub>50</sub> de pEGFR (parental/mutante)	Δ Energía de interacción, kcal/mol
T49→ A	Cadena D	1,5	1,2	0,1	-1
V50→ A	Cadena D	3,4	0,5	0,1	1
P51→ A	Cadena D	1,8	1	N.E.	<b>3</b>
G52→ A	Bucle DE	14	0,1	0,04	0
P53→ A	Bucle DE	2,8	0,6	0,3	1
V54→ A	Bucle DE	17	0,1	0,03	2
H55→ A	Bucle DE	3,0	0,6	0,6	0
D77→ A	Bucle FG	N.D.		N.D.	<b>6</b>
H78→ A	Bucle FG	5,5	0,3	0,1	2
K79→ A	Bucle FG	17	0,1	0,0	1
P79A→ A	Bucle FG	4,0	0,4	0,1	0
H79B→ A	Bucle FG	3,9	0,5	0,1	0
H79D→ A	Bucle FG	2,8	0,6	0,2	1
G79E→ A	Bucle FG	4,4	0,4	0,1	0
P80→ A	Bucle FG	4,2	0,4	0,1	0
H81→ A	Bucle FG	10	0,2	0,0	1
T82→ A	Bucle FG	6,1	0,3	0,1	0
Y83→ A	Bucle FG	N.D.		N.D.	0
H84→ A	Bucle FG	2,1	0,8	0,4	0
E85→ A	Bucle FG	4,9	0,4	0,05	0

N.D. no detectado; N.E. no ensayado.

La CI<sub>50</sub> de pEGFR es la concentración inhibitoria a la que la fosforilación de EGFR se inhibe al 50 %.

En la ΔEnergía de interacción, los números en negrita de la columna indican que se espera que sean pérdidas significativas de energía de interacción. Véase también la Tabla 8.

5

**Tabla 5. Mutantes de EGFR que tratan de mejorar la unión**

Mutación	Elemento estructural secundario	K <sub>D</sub> , nM	Actividad (parental/mutante)	CI <sub>50</sub> de pEGFR (parental/mutante)	Δ Energía de interacción, kcal/mol
Parental		1,8	1		1
Y31→ F	Bucle BC	2,0	0,9		N.E. 2
Y31→ S	Bucle BC	0,70	2,5		N.E. 5
Y31→ L	Bucle BC	N.D.			N.D. 2
E47→ R	Cadena D	0,83	2,1		N.E. <b>-4</b>
E47→ K	Cadena D	0,9	2,0		N.E. <b>-6</b>
E47→ D	Cadena D	0,78	2,3		N.E. 0
T49→ I	Cadena D	0,05	37		4,2 0
T49→ V	Cadena D	0,14	12		3,1 -1
T49→ Y	Cadena D	0,21	8,4		1,0 -3
P51→ T	Cadena D	N.D.			N.E. -1
P51→ L	Cadena D	1,5	1,2		N.E. -1
T58→ Q	Cadena E	2,2	0,8		0,3 0
T58→ E	Cadena E	1,3	1,3		0,7 <b>-9</b>
T58→ D	Cadena E	1,1	1,7		1,4 <b>-7</b>
A79C→ L	Bucle FG	4,0	0,4		N.E. <b>-3</b>

(continuación)

Mutación	Elemento estructural secundario	K <sub>D</sub> , nM	Actividad (parental/mutante)	Cl <sub>50</sub> de pEGFR (parental/mutante)	Δ Energía de interacción, kcal/mol
A79C→ N	Bucle FG	2,3	0,8	N.E.	<b>-4</b>
A79C→ Y	Bucle FG	3,2	0,5	N.E.	<b>-8</b>
A79C→ R	Bucle FG	4,3	0,4	N.E.	<b>-11</b>
A79C→ E	Bucle FG	2,7	0,7	N.E.	<b>-8</b>

N.D. no detectado; N.E. no ensayado.

La Cl<sub>50</sub> de pEGFR es la concentración inhibitora a la que la fosforilación de EGFR se inhibe al 50 %.

En la ΔEnergía de interacción, los números en negrita de la columna indican que se espera que sean pérdidas significativas de energía de interacción, y lo subrayado indica que se espera que sean ganancias significativas de energía de interacción.

5 Véase también la Tabla 9.

**Tabla 6. Cl<sub>50</sub> de IL-23**

Mutación	Elemento estructural secundario	Cl <sub>50</sub> , nM	Actividad (parental/mutante)	Δ Energía de interacción, kcal/mol
Parental		1,0	1	
Y28→ A	Bucle BC	13	0,08	<u>10</u>
Y73→ A	Cadena F	0,65	1,5	<u>5</u>
Y81→ A	Bucle FG	35	0,03	7
P87→ A	Bucle FG	8,3	0,1	2
T35→ N	Cadena C	1,7	0,6	0
T35→ Q	Cadena C	1,0	1	-2
T35→ E	Cadena C	0,6	1,7	<b>-6</b>
T35→ D	Cadena C	49	0,02	<b>-7</b>
V45→ N	Cadena D	0,54	1,8	1
V45→ Q	Cadena D	0,34	2,9	0
V45→ E	Cadena D	0,51	1,9	<b>-6</b>
V45→ D	Cadena D	0,83	1,2	<b>-6</b>
Y73→ N	Cadena F	1,2	0,8	2
Y73→ Q	Cadena F	0,26	3,8	3
Y73→ R	Cadena F	1,2	0,8	-2
V75→ Y	Cadena F	2,2	0,4	-1
V75→ Q	Cadena F	2,3	0,4	0
V75→ K	Cadena F	5,7	0,2	-1

10 En la ΔEnergía de interacción, los números en negrita de la columna indican que se espera que sean pérdidas significativas de energía de interacción, y lo subrayado indica que se espera que sean ganancias significativas de energía de interacción.

Véase también la Tabla 10.

Tabla 7. Colección de datos y estadísticas del refinado

	EGFR/Adnectina 1	IL-23/Adnectina 2
<b>Colección de datos</b>		
Grupo espacial	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	I2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Dimensiones de la celda		
<i>a, b, c</i> (Å)	68,0, 72,1, 262,0	77,7 A, 91,7, 225,8
$\alpha, \beta, \gamma$ (°)	90, 90, 90	90, 90, 90
Resolución (Å)	50-2,75 (2,85-2,75)	42,47-3,25 (3,37-3,25)
<i>R</i> <sub>sim</sub>	0,069 (0,552)	0,096 (0,320)
// $\sigma$	23,1 (3,1)	7,6 (3,2)
Completo (%)	99,6 (100,0)	98,0 (99,5)
Redundancia	4,8 (5,0)	4,0 (3,9)
<b>Refinado</b>		
Resolución (Å)	49,47-2,75	42,48-3,25
N.º de reflejos	34.224	12.815
<i>R</i> <sub>trabajo</sub> / <i>R</i> <sub>libre</sub>	0,202 / 0,246	0,234 / 0,264
N.º de átomos	5610	4040
factores <i>B</i>	65	101
Proteína	63	101
Carbohidratos	100	115
Agua	43	68
Desviaciones R.m.s.		
Longitudes del enlace (Å)	0,010	0,010
Ángulos del enlace (°)	1,4	1,4
Estadísticas del gráfico de Ramachandran <sup>a</sup>		
Más favorecido (%)	84,5	85,8
Adicional permitido (%)	13,7	11,3
No permitido (%)	0,8	1,1

Se usó un cristal para cada complejo. <sup>a</sup>  
Tal como se define por Laskowski *et al.*, 1993.

Tabla 8. Cálculos energéticos para la exploración de alanina del complejo EGFR/Adnectina 1 (resumido en la Tabla 4).

Mutación	Mutante					Tipo silvestre				
	E.I.	VDW	ELE	% I	$\Delta G$	E.I.	VDW	ELE	% I	$\Delta G$
D23→A	0	0	0	0	0	0	0	0	0	0
S24→A	0	0	0	0	0	0	0	0	0	0
G25→A	0	0	0	0	0	0	0	0	0	0
R26→A	0	0	0	0	0	0	0	0	0	0
G27→A	0	0	0	0	0	0	0	0	0	0
S28→A	0	0	0	0	0	0	0	0	0	0
Y29→A	0	0	0	0	0	0	0	0	0	0
Q30→A	-2,9	-2,6	-0,3	70	0,2	-6,1	-4,9	-1,2	75	-2,4
Y31→A	-3,7	-3,8	0,1	65	-0,8	-6,1	-6	-0,1	78	-1,3
F48→A	-1	-1,2	0,2	36	0,6	-1,9	-2,2	0,3	55	-0,5
T49→A	-4	-0,9	-3,1	50	-2,7	-5,2	-2,2	-2,9	70	-3,2
V50→A	-2	-1,9	-0,1	43	0	-3,1	-3	-0,1	99	-0,7
P51→A	-7,5	-6,1	-1,5	86	-1,5	-6,1	-5,8	-0,3	92	0,5
G52→A	-0,4	-0,3	-0,1	0	-0,4	-0,3	-0,2	-0,1	0	-1,8
P53→A	-0,7	-0,6	-0,1	6	-0,7	-1,2	-1,1	-0,1	14	0
V54→A	-1,3	-1,6	0,3	60	-1,3	-3,3	-3,5	0,2	95	0,3
H55→A	0	0	0	0	0	0	0	0	0	0

(continuación)

Mutación	Mutante					Tipo silvestre				
	E.I.	VDW	ELE	% I	$\Delta G$	E.I.	VDW	ELE	% I	$\Delta G$
D77→A	-4,7	-3,3	-1,3	87	-1,3	-10,3	-3,8	-6,5	96	-7,2
H78→A	-4	-4,2	0,2	63	0,9	-6	-6,3	0,3	53	0,1
K79→A	-1,2	-0,7	-0,5	93	0	-2,1	-2,4	0,3	28	-1,7
P79A→A	-2,5	-2,6	0,1	27	0,6	-2,4	-2,7	0,3	27	0,6
H79B→A	-1,8	-1,8		15	0,2	-1,9	-2,2	0,3	7	-0,7
H79D→A	0	0	0	0	0	-1,3	-0,1	-1,2	0	-2,4
G79E→A	0	0	0	0	0	0	0	0	0	0
P80→A	0	0	0	0	0	0	0	0	0	0
H81→A	0	0	0	0	0	-1,5	-1,6	0,1	19	-0,8
T82→A	0	0	0	0	0	-0,1	-0,1		0	-1,2
Y83→A	0	0	0	0	0	-0,1	-0,1	0	0	-1,8
H84→A	0	0	0	0	0	0	0	0	0	0
E85→A	0	0	0	0	0	0	0	0	0	0

\*Columnas: E.I. = energía de interacción; VDW = interacciones de Van der Waals; ELE = interacciones electrostáticas; % I = % de superficie interna

**Tabla 9. Cálculos energéticos para mutaciones a ciertos restos de contacto del complejo EGFR/Adnectina 1**  
(Resumido en la Tabla 5).

Mutación	Mutante					Tipo silvestre				
	E.I.	VDW	ELE	% I	$\Delta G$	E.I.	VDW	ELE	% I	$\Delta G$
Y31→F	-3,7	-3,8	0,1	65	-0,8	-6,1	-6	-0,1	78	-1,3
Y31→S	-1,5	-1,6	0,1	93	-0,9					
Y31→L	-3,9	-4	0,1	87	-0,4					
E47→R	-5,2	-1,7	-3,5	55	-4,3	-1	-0,5	-0,5	8	-2,2
E47→K	-6,7	-1,3	-5,4	27	-7					
E47→D	-0,8	-0,4	-0,4	17	-1,6					
T49→I	-5,3	-3,3	-2	67	-2,5	-5,2	-2,2	-2,9	70	-3,2
T49→V	-4,3	-2,3	-2	60	-2,2					
T49→Y	-5,3	-3,3	-2	67	-2,5					
P51→T	-5,3	-3,3	-2	67	-2,5	-6,1	-5,8	-0,3	92	0,5
P51→A	-4,3	-2,3	-2	60	-2,2					
P51→L	-6,6	-6,2	-0,4	86	-0,6					
T58→Q	-0,2	-0,1	-0,1	1	-1,9	-0,6	-0,2	-0,4	2	-1,6
T58→E	-9,7	-1,1	-8,6	38	-9,8					
T58→D	-7,9	-0,8	-7,1	32	-7,7					
A79C→L	-3,7	-3,6	-0,1	66	-0,1	-1,1	-1,1	0	63	1
A79C→N	-5	-3,2	-1,8	50	-1,5					
A79C→Y	-9,7	-8,8	-0,9	77	-1,3					
A79C→R	-12,3	-7,6	-4,7	81	-2,5					
A79C→E	-9,2	-6,7	-2,5	77	-1,4					

\*Columnas: E.I. = energía de interacción; VDW = interacciones de Van der Waals; ELE = interacciones electrostáticas; % I = % de superficie interna

**Tabla 10. Cálculos energéticos para restos de contacto del complejo Adnectina 2/IL-23**  
(Resumido en la Tabla 6).

Mutación	Mutante					Tipo silvestre				
	E.I.	VDW	ELE	% I	$\Delta G$	E.I.	VDW	ELE	% I	$\Delta G$
Y28→A										
Y73→A										
Y81→A										
P87→A										
T35→N	0	0	0	11	-1	-0,8	-0,6	-0,2	25	-0,5
T35→Q	-3,2	-1,8	-1,4	77	-2,7					
T35→E	-6,4	-1	-5,4	61	-6,7					
T35→D	-7,5	-0,5	-7	78	-8					
V45→N	-0,2	-0,3	0,1	5	-1,1	-0,8	-0,6	-0,2	25	-0,5
V45→Q	-0,4	-0,3	-0,1	7	-1,9					
V45→E	-7	-0,9	-6,1	42	-7,3					
V45→D	-6,9	-0,5	-6,4	42	-6,9					
Y73→N	-3,9	-1,3	-2,6	76	-3,4	-5,6	-2,5	-3,1	84	-3,8
Y73→Q	-2,8	-0,6	-2,2	60	-3,4					
Y73→R	-7,6	-4,7	-2,9	75	-4,1					
V75→Y	-1,3	-1,1	-0,2	20	-1,8	-0,7	-0,7	0	28	-0,6
V75→Q	-1,1	-0,7	-0,4	24	-2,1					
V75→K	-1,4	-1,5	0,1	32	-1,7					

\*Columnas: E.I. = energía de interacción; VDW = interacciones de Van der Waals; ELE = interacciones electrostáticas; % I = % de superficie interna

### Ejemplo 3: Solubilidad de mutantes de adnectina

5

#### PROCEDIMIENTOS EXPERIMENTALES:

EGFR8 (SEQ ID NO:67), EGFR8-T58E (SEQ ID NO:69), EGFR8-T58D (SEQ ID NO:70), EGFR4 (SEQ ID NO:68) y EGFR4-T58E (SEQ ID NO: 71) Las adnectinas se expresaron y purificaron tal como se describe anteriormente en los Ejemplos 1 y 2. Todas las proteínas se sometieron a diálisis en NaPO<sub>4</sub> 10 mM, NaCl 130 mM (PBS) a pH 7,1 y se confirmó que las muestras eran más del 97 % monoméricas usando la cromatografía analítica de exclusión por tamaño en una columna Zenix-C SEC-300 en tampón que contiene K<sub>2</sub>HPO<sub>4</sub> 200 mM, NaCl 150 mM, a pH 6,8, más azida Na al 0,02 %, con rendimiento de 0,35 ml/min.

Los análisis de estabilidad térmica de EGFR8, EGFR8-T58E, EGFR8-T58D, EGFR4 y EGFR4-T58E se llevaron a cabo mediante calorimetría de barrido diferencial (DSC) usando un instrumento de DSC MicroCal Capillary. Para estabilizar las condiciones de partida del instrumento de DSC y obtener un historial término constante, se registraron múltiples barridos de tampón PBS solo a pH 7,1 tanto en la muestra como en la célula de referencia antes del análisis de la muestra. Los barridos de la muestra contenían 1,0 mg/ml de adnectina en la célula de muestra y el tampón de PBS con pH 7,1 emparejado en la célula de referencia. Todos los barridos se llevaron a cabo desde 10-100 °C a una tasa de barrido de 60/h usando un periodo de termostato preciclo de 15 minutos. Los datos se analizaron usando el programa informático de análisis MicroCal Origin.

Se realizaron experimentos de solubilidad en sulfato de amonio para las adnectinas EGFR4 usando el método de ensayo automático descrito en el Ejemplo 1. Los experimentos de solubilidad en sulfato de amonio para las proteínas EGFR8 se realizaron usando el método a escala de banco descrito en el Ejemplo 1 que permite la titulación a mayores concentraciones de sulfato de amonio que el método automático.

Se realizaron experimentos de ultrafiltración a pequeña escala para EGFR8, EGFR8-T58E y EGFR8-T58D concentrando lentamente las adnectinas a 2-8 °C en concentradores VivaSpin 3000 MWCO con mezcla periódica para evitar la formación de gradientes. Los volúmenes se controlaron visualmente durante el proceso de concentración para permitir la estimación de la concentración de la muestra esperada basándose en la reducción del volumen, y se extrajeron alícuotas en diversos puntos para el análisis. Estas alícuotas se incubaron toda la noche a 2-8 °C y se retiró cualquier material insoluble mediante centrifugación antes de la determinación de la concentración de proteínas solubles mediante A<sub>280</sub> y del estado oligomérico mediante SEC usando una columna Zenix-C SEC-300 en tampón que contiene K<sub>2</sub>HPO<sub>4</sub> 200 mM, NaCl 150 mM, a pH 6,8, más azida Na al 0,02 %, con rendimiento de 0,35 ml/min.

Los primeros experimentos de ultrafiltración a pequeña escala usando EGFR4 y EGFR4-T58E realizados a 2-8 °C se vieron obstaculizados por tasas muy lentas de reducción de volumen. Por lo tanto, EGFR4 y EGFR4-T58E se concentraron cerca de la temperatura ambiente (21 °C), donde se observó una tasa más rápida de reducción de volumen. Los volúmenes se controlaron visualmente durante el proceso de concentración para permitir la estimación de la concentración de la muestra esperada basándose en la reducción del volumen, y se extrajeron alícuotas en diversos puntos para el análisis. Para hacer un seguimiento de las observaciones de la tasa de ultrafiltración dependiente de la temperatura, cada alícuota se dividió en dos alícuotas hermanas, con un conjunto incubado toda la noche a 2-8 °C y el otro conjunto incubado toda la noche a temperatura ambiente. A la mañana siguiente, se retiró el material insoluble mediante centrifugación y, después, se determinó la concentración de proteína soluble mediante  $A_{280}$  y el estado oligomérico caracterizado mediante SEC usando una columna Zenix-C SEC-300 en tampón que contiene K<sub>2</sub>HPO<sub>4</sub> 200 mM, NaCl 150 mM, a pH 6,8, más azida Na al 0,02 %, con rendimiento de 0,35 ml/min.

Se realizaron estudios de estabilidad acelerada para EGFR8, EGFR8-T58E y EGFR8-T58D incubando muestras de 46 mg/ml de cada adnectina a 40 °C durante dos semanas. Las alícuotas se retiraron en el momento cero (inmediatamente antes del comienzo de la incubación a 40 °C), y a 1 semana y a las 2 semanas. Las alícuotas se centrifugaron para eliminar material insoluble y después se determinó la concentración de proteína soluble mediante  $A_{280}$  y se caracterizó el estado oligomérico mediante SEC usando una columna Zenix-C SEC-300 en tampón que contiene K<sub>2</sub>HPO<sub>4</sub> 200 mM, NaCl 150 mM, a pH 6,8, más azida Na al 0,02 %, con rendimiento de 0,35 ml/min.

## 20 RESULTADOS:

Los primeros experimentos de solubilidad en sulfato de amonio con EGFR8 (**Figuras 10 y 11**) sugirieron que las mutaciones T58E y T58D mejoran la solubilidad de la adnectina EGFR8. Para caracterizar adicionalmente el efecto de estas mutaciones, el EGFR8 de tipo silvestre así como los mutantes EGFR8-T58E y EGFR8-T58D se expresaron y se purificaron a mayor escala para generar suficiente material para estudios biofísicos más detallados. Las proteínas purificadas eran de una pureza adecuada para estudios biofísicos, con más del 98 % de monómeros para cada muestra tal como se mide mediante SEC (**Figura 20A**). La estabilidad térmica global para EGFR8 de tipo silvestre y EGFR8-T58E también fue similar, con la curva de fusión para EGFR8 consistiendo en una transición térmica amplia con  $T_{inicial}$  cerca de 50 °C y  $T_m$  de 74,0 °C y EGFR8-T58E teniendo un perfil de fusión de forma gaussiana más simétrico con  $T_{inicial}$  cerca de 50 °C y  $T_m$  de 71,7 °C (**Figura 20B-D**). Sin embargo, la estabilidad térmica de EGFR8-T58D fue considerablemente menor, con  $T_{inicial}$  cerca de 43 °C y  $T_m$  de 64,4 °C (**Figura 20B y 20E**), lo que sugiere que la mutación T58 → D puede causar un cambio conformacional en la adnectina que reduce la estabilidad de la proteína.

El impacto de las mutaciones T58E y T58D sobre la agregación de EGFR8 se estudió a continuación usando el ensayo de solubilidad en sulfato de amonio. Las curvas de precipitación salina en sulfato de amonio generadas usando las preparaciones a mayor escala de material de adnectina purificada confirmaron que la solubilidad en sulfato de amonio aumentó para cada mutante en comparación con la proteína de tipo silvestre de EGFR8 (**Figura 21**).

Como medida ortogonal de propensión a la agregación para el método de AS, los presentes inventores realizaron experimentos de ultrafiltración a pequeña escala usando EGFR8, EGFR8-T58E y EGFR8-T58D tal como se describe anteriormente. Se descubrió que cada una de las tres adnectinas tiene alta solubilidad (más de 100 mg/ml) (**Figura 22A**). La buena concordancia entre la [proteína] observada y la [proteína] esperada tanto para EGFR8 como para EGFR8-T58E sugiere que el límite de solubilidad para cada una de estas moléculas es mayor que las concentraciones logradas en este experimento. Las concentraciones observadas para EGFR8-T58D también se registraron bien con las concentraciones esperadas salvo por la alícuota final (la más concentrada) para la que la concentración medida (112 mg/ml) era menos que la concentración esperada (aproximadamente  $150 \pm 25$  mg/ml). La recuperación más baja para EGFR8-T58D a una concentración elevada sugiere que esta proteína puede tener una mayor propensión a la agregación que EGFR8 y EGFR8-T58E en estas condiciones, mientras que la propensión a la agregación de EGFR8-T58E es indistinguible de la adnectina EGFR8 de tipo silvestre. Los datos de SEC presentaron pequeños aumentos similares en APM soluble para cada proteína como una función de la concentración de proteína, conservando cada proteína más del 97 % de monoméricos (menos del 3 % de APM) incluso a concentraciones mayores de 100 mg/ml (**Figura 22B**).

Dado que las diferencias de solubilidad entre el EGFR8 de tipo silvestre y EGFR8-T58E fueron difíciles de resolver en los estudios de ultrafiltración debido a la alta solubilidad de cada molécula, los presentes inventores trataron de diferenciar la propensión a la agregación de las adnectinas usando estudios de estabilidad acelerada. Aquí, las muestras de "momento cero" ( $t_0$ ) de 46 mg/ml de EGFR8, EGFR8-T58E y EGFR8-T58D se prepararon en PBS a pH 7,1, y confirmaron ser de más del 98 % de monómero mediante SEC (**Figura 23C**). Las muestras después se incubaron a 40 °C durante 2 semanas y en el momento de 1 semana (1s) y de 2 semanas (2s) se retiró el material insoluble mediante centrifugación y la parte soluble se analizó mediante  $A_{280}$  y SEC para determinar la concentración de proteína y el estado oligomérico. Estos datos demuestran que la concentración de EGFR8 se redujo en aproximadamente el 60 % a aproximadamente 19 mg/ml tras 1 semana a 40 °C, y después se redujo a 16 mg/ml en el punto temporal de 2s (**Figura 23A**). La concentración de EGFR8-T58D disminuyó incluso más que en la adnectina EGFR8 de tipo silvestre, a 11 mg/ml en 1sem y 9 mg/ml en 2sem. Por otro lado, la concentración de EGFR8-T58E se redujo a solo 40 mg/ml en 1sem y a 36 mg/ml en 2sem (**Figura 23A**). Estos cambios en la concentración de proteína soluble se confirmaron mediante los datos del área de pico de SEC integrada (**Figura 23B**). Los datos de SEC también

demonstraron que la fracción soluble de cada muestra permaneció monomérica a más del 98 %, con una pequeña reducción en el % de APM a lo largo del tiempo que probablemente indica una mayor propensión a la agregación para las especies de APM en comparación con la adnectina monomérica (**Figura 23C**). En conjunto, estos datos indican que EGFR8-T58E es significativamente más resistente a la agregación que EGFR8 o EGFR8-T58D en estas condiciones de estrés acelerado.

Los datos que comparan las propiedades de agregación de EGFR8 con EGFR8-T58E (**Figuras 21-23**) sugieren que la mutación T58E reduce la propensión a la agregación de esta adnectina en particular. Para determinar si esta mutación T58E tiene un impacto favorable similar sobre las propiedades de agregación de otras adnectinas, los presentes inventores expresaron y purificaron las cantidades de multimiligramos de adnectinas EGFR4 de tipo silvestre y EGFR4-T58E para estudios biofísicos. La adnectina EGFR4 se seleccionó para el estudio porque anteriormente los presentes inventores habían descubierto que esta adnectina era altamente propensa a la agregación en ambos estudios de precipitación salina AS (**Figura 3A**) y en experimentos de ultrafiltración (**tabla 1**), y por lo tanto, anticiparon que cualquier efecto de la mutación T58E sería más fácilmente detectable para EGFR4 que para las adnectinas menos propensas a la agregación. A pesar de la unión a la misma proteína diana de EGFR, la adnectina EGFR4 tenía diferentes secuencias de bucle de unión a diana que la EGFR8 (**Figura 24**).

Las proteínas EGFR4 y EGFR4-T58E purificadas demostraron ser, mediante SEC, de una pureza adecuada para estudios biofísicos, con más del 97 % de monómeros para cada muestra (**Figura 25A**). El termograma de DSC para la EGFR4 de tipo silvestre demostró que la adnectina tenía una elevada estabilidad térmica, con una  $T_{inicial}$  de  $\sim 60$  °C con dos transiciones claramente resolubles; una transición menor con  $T_m = 74,3$  °C y una transición dominante con  $T_m = 85,3$  °C (**Figura 25B-C**). Los datos de DSC para EGFR4-T58E mostraron una  $T_{inicial}$  similar cerca de  $\sim 60$  °C, y los datos del termograma para EGFR4 también se describieron mejor mediante dos transiciones superpuestas con una transición menor que tiene  $T_m1 = 77,4$  °C y una transición dominante con  $T_m2 = 83,8$  °C (**Figura 25B y D**). Por lo tanto, como EGFR8 (**Figura 20**), los datos de DSC sugieren que la estabilidad térmica de EGFR4 no se redujo significativamente por la mutación T58→E.

Las curvas de precipitación salina en sulfato de amonio para las preparaciones a gran escala de EGFR4 de tipo silvestre y EGFR4-T58E muestran que la mutación T58E aumenta la solubilidad de la proteína en AS, con  $AS_m = 0,865 \pm 0,014$  M para EGFR4 de tipo silvestre y  $AS_m = 1,024 \pm 0,003$  para EGFR4-T58E (promedio y desviación estándar de mediciones por cuadruplicado), cuando se ensaya a 0,3 mg/ml de [proteína] (**Figura 26**).

Los estudios de ultrafiltración a pequeña escala se realizaron con EGFR4 y EGFR4-T58E para examinar su propensión a la agregación. Los estudios de ultrafiltración inicial realizados a 2-8 °C no fueron exitosos tanto para la adnectina de tipo silvestre como para la adnectina mutante debido a tasas muy lentas de reducción de volumen, lo que sugiere posiblemente una alta propensión a la agregación de cada proteína a estas bajas temperaturas. Sin embargo, ambas adnectinas se concentraron a tasas más rápidas cuando el estudio de ultrafiltración se llevó a cabo cerca de la temperatura ambiente (21 °C). Por lo tanto, para investigar este fenómeno dependiente de la temperatura, las muestras se concentraron mediante ultrafiltración a temperatura ambiente y las alícuotas se almacenaron bien a 2-8 °C o a temperatura ambiente durante toda la noche. A la mañana siguiente se retiró el material insoluble mediante centrifugación (y a 2-8 °C o a temperatura ambiente, respectivamente) y la concentración de proteína soluble se midió mediante  $A_{280}$  y se caracterizó el estado oligomérico mediante SEC. Los datos para el EGFR4 de tipo silvestre demostraron que la mayoría de la proteína precipitó durante el almacenamiento a 2-8 °C, tal que la concentración más alta de EGFR4 soluble después del almacenamiento nocturno fue de 5,2 mg/ml (**Figura 27A**), de acuerdo con nuestro estudio anterior para EGFR4 realizado en condiciones similares (véase la **Tabla 1**). EGFR4-T58E presentó una tendencia similar en la concentración al tipo silvestre, excepto en el punto de datos de la concentración más alta donde permaneció ligeramente más EGFR4-T58E en solución (7,4 mg/ml) en comparación con el tipo silvestre. Las recuperaciones para ambas proteínas mejoraron significativamente en las alícuotas incubadas a temperatura ambiente, y se descubrió que EGFR4-T58E en particular tenía una concentración aproximadamente dos veces mayor (22 mg/ml) que la EGFR4 de tipo silvestre (12 mg/ml) en las alícuotas de concentración más altas (**Figura 27A**). Los datos de SEC mostraron tendencias similares para aumentar los APM en función de la concentración de proteína para EGFR4 y EGFR4-T58E (**Figura 27B**). Estos datos demuestran que la mutación T58→E redujo la propensión a la agregación de EGFR4.

Diversas modificaciones a los modelos y métodos de la invención, además de las que se describen en el presente documento, serán evidentes para los expertos en la materia a partir de la descripción anterior y las enseñanzas.

#### LISTADO DE SECUENCIAS

<110> BRISTOL-MYERS SQUIBB COMPANY

<120> PROTEÍNAS DE DOMINIO DE FIBRONECTINA TIPO III CON SOLUBILIDAD MEJORADA

<130> MXI-557US

<140> US 14/766.036

<141> 05/08/2015  
 <150> PCT/US2014/014887  
 <151> 05/02/2014  
 5 <150> US 61/761.452  
 <151> 06/02/2013  
 <160> 87  
 10 <170> PatentIn versión 3.5  
 <210> 1  
 <211> 101  
 15 <212> PRT  
 <213> *Homo sapiens*  
 <400> 1  
 Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Thr Pro Thr  
 1 5 10 15  
 Ser Leu Leu Ile Ser Trp Asp Ala Pro Ala Val Thr Val Arg Tyr Tyr  
 20 25 30  
 Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln Glu Phe  
 35 40 45  
 Thr Val Pro Gly Ser Lys Ser Thr Ala Thr Ile Ser Gly Leu Lys Pro  
 50 55 60  
 Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Gly Arg Gly Asp  
 65 70 75 80  
 Ser Pro Ala Ser Ser Lys Pro Ile Ser Ile Asn Tyr Arg Thr Glu Ile  
 85 90 95  
 Asp Lys Pro Ser Gln  
 100  
 20  
 <210> 2  
 <211> 101  
 <212> PRT  
 25 <213> Secuencia artificial  
 <220>  
 <223> Polipéptido sintético  
 30 <400> 2

ES 2 814 558 T3

Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Thr Pro Thr  
 1 5 10 15  
 Ser Leu Leu Ile Ser Trp Asp Ala Pro Ala Val Thr Val Arg Tyr Tyr  
 20 25 30  
 Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln Glu Phe  
 35 40 45  
 Thr Val Pro Gly Ser Lys Ser Thr Ala Thr Ile Ser Gly Leu Lys Pro  
 50 55 60  
 Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Gly Arg Gly Asp  
 65 70 75 80  
 Ser Pro Ala Ser Ser Lys Pro Ile Ser Ile Asn Tyr Arg Thr Glu Ile  
 85 90 95  
 Glu Lys Pro Ser Gln  
 100

5

<210> 3  
 <211> 85  
 <212> PRT  
 <213> *Homo sapiens*  
 <400> 3

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

10

<210> 4  
 <211> 86  
 <212> PRT  
 <213> *Homo sapiens*

15

ES 2 814 558 T3

<400> 4

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

5 <210> 5  
 <211> 94  
 <212> PRT  
 <213> *Homo sapiens*

10 <400> 5

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

15 <210> 6  
 <211> 95  
 <212> PRT  
 <213> *Homo sapiens*

<400> 6

ES 2 814 558 T3

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

5  
 <210> 7  
 <211> 96  
 <212> PRT  
 <213> *Homo sapiens*  
 <400> 7

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

15  
 <210> 8  
 <211> 97  
 <212> PRT  
 <213> *Homo sapiens*  
 <400> 8

ES 2 814 558 T3

Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Thr Pro Thr  
 1 5 10 15

Ser Leu Leu Ile Ser Trp Asp Ala Pro Ala Val Thr Val Arg Tyr Tyr  
 20 25 30

Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln Glu Phe  
 35 40 45

Thr Val Pro Gly Ser Lys Ser Thr Ala Thr Ile Ser Gly Leu Lys Pro  
 50 55 60

Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Gly Arg Gly Asp  
 65 70 75 80

Ser Pro Ala Ser Ser Lys Pro Ile Ser Ile Asn Tyr Arg Thr Glu Ile  
 85 90 95

Asp

<210> 9  
 <211> 97  
 <212> PRT  
 <213> *Homo sapiens*

5

<400> 9

Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Thr Pro Thr  
 1 5 10 15

Ser Leu Leu Ile Ser Trp Asp Ala Pro Ala Val Thr Val Arg Tyr Tyr  
 20 25 30

Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln Glu Phe  
 35 40 45

Thr Val Pro Gly Ser Lys Ser Thr Ala Thr Ile Ser Gly Leu Lys Pro  
 50 55 60

Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Gly Arg Gly Asp  
 65 70 75 80

Ser Pro Ala Ser Ser Lys Pro Ile Ser Ile Asn Tyr Arg Thr Glu Ile  
 85 90 95

10

Glu

<210> 10

ES 2 814 558 T3

<211> 98  
 <212> PRT  
 <213> *Homo sapiens*

5 <400> 10

Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Thr Pro Thr  
 1 5 10 15

Ser Leu Leu Ile Ser Trp Asp Ala Pro Ala Val Thr Val Arg Tyr Tyr  
 20 25 30

Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln Glu Phe  
 35 40 45

Thr Val Pro Gly Ser Lys Ser Thr Ala Thr Ile Ser Gly Leu Lys Pro  
 50 55 60

Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Gly Arg Gly Asp  
 65 70 75 80

Ser Pro Ala Ser Ser Lys Pro Ile Ser Ile Asn Tyr Arg Thr Glu Ile  
 85 90 95

Asp Lys

10 <210> 11  
 <211> 98  
 <212> PRT  
 <213> Secuencia artificial

15 <220>  
 <223> Polipéptido sintético

<400> 11

Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Thr Pro Thr  
 1 5 10 15

Ser Leu Leu Ile Ser Trp Asp Ala Pro Ala Val Thr Val Arg Tyr Tyr  
 20 25 30

Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln Glu Phe  
 35 40 45

Thr Val Pro Gly Ser Lys Ser Thr Ala Thr Ile Ser Gly Leu Lys Pro  
 50 55 60

20

ES 2 814 558 T3

Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Gly Arg Gly Asp  
65 70 75 80

Ser Pro Ala Ser Ser Lys Pro Ile Ser Ile Asn Tyr Arg Thr Glu Ile  
85 90 95

Glu Lys

5 <210> 12  
<211> 99  
<212> PRT  
<213> *Homo sapiens*  
<400> 12

Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Thr Pro Thr  
1 5 10 15

Ser Leu Leu Ile Ser Trp Asp Ala Pro Ala Val Thr Val Arg Tyr Tyr  
20 25 30

Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln Glu Phe  
35 40 45

Thr Val Pro Gly Ser Lys Ser Thr Ala Thr Ile Ser Gly Leu Lys Pro  
50 55 60

Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Gly Arg Gly Asp  
65 70 75 80

Ser Pro Ala Ser Ser Lys Pro Ile Ser Ile Asn Tyr Arg Thr Glu Ile  
85 90 95

Asp Lys Pro

10 <210> 13  
<211> 99  
<212> PRT  
15 <213> Secuencia artificial  
  
<220>  
<223> Polipéptido sintético  
  
20 <400> 13

Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Thr Pro Thr  
1 5 10 15

ES 2 814 558 T3

Ser Leu Leu Ile Ser Trp Asp Ala Pro Ala Val Thr Val Arg Tyr Tyr  
 20 25 30

Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln Glu Phe  
 35 40 45

Thr Val Pro Gly Ser Lys Ser Thr Ala Thr Ile Ser Gly Leu Lys Pro  
 50 55 60

Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Gly Arg Gly Asp  
 65 70 75 80

Ser Pro Ala Ser Ser Lys Pro Ile Ser Ile Asn Tyr Arg Thr Glu Ile  
 85 90 95

Glu Lys Pro

<210> 14  
 <211> 100  
 <212> PRT  
 <213> *Homo sapiens*

5

<400> 14

Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Thr Pro Thr  
 1 5 10 15

Ser Leu Leu Ile Ser Trp Asp Ala Pro Ala Val Thr Val Arg Tyr Tyr  
 20 25 30

Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln Glu Phe  
 35 40 45

Thr Val Pro Gly Ser Lys Ser Thr Ala Thr Ile Ser Gly Leu Lys Pro  
 50 55 60

Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Gly Arg Gly Asp  
 65 70 75 80

Ser Pro Ala Ser Ser Lys Pro Ile Ser Ile Asn Tyr Arg Thr Glu Ile  
 85 90 95

Asp Lys Pro Ser  
 100

10

<210> 15  
 <211> 100  
 <212> PRT  
 <213> Secuencia artificial

15

<220>

ES 2 814 558 T3

<223> Polipéptido sintético

<400> 15

```

Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Thr Pro Thr
1          5          10          15

Ser Leu Leu Ile Ser Trp Asp Ala Pro Ala Val Thr Val Arg Tyr Tyr
          20          25          30

Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln Glu Phe
          35          40          45

Thr Val Pro Gly Ser Lys Ser Thr Ala Thr Ile Ser Gly Leu Lys Pro
          50          55          60

Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Gly Arg Gly Asp
65          70          75          80

Ser Pro Ala Ser Ser Lys Pro Ile Ser Ile Asn Tyr Arg Thr Glu Ile
          85          90          95

Glu Lys Pro Ser
          100
    
```

5

<210> 16  
 <211> 101  
 <212> PRT  
 <213> *Homo sapiens*

10

<400> 16

```

Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Thr Pro Thr
1          5          10          15

Ser Leu Leu Ile Ser Trp Asp Ala Pro Ala Val Thr Val Arg Tyr Tyr
          20          25          30

Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln Glu Phe
          35          40          45

Thr Val Pro Gly Ser Lys Ser Thr Ala Thr Ile Ser Gly Leu Lys Pro
          50          55          60

Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Gly Arg Gly Glu
65          70          75          80

Ser Pro Ala Ser Ser Lys Pro Ile Ser Ile Asn Tyr Arg Thr Glu Ile
          85          90          95
    
```

15

# ES 2 814 558 T3

## Asp Lys Pro Ser Gln 100

- 5 <210> 17  
<211> 166  
<212> PRT  
<213> Secuencia artificial
- 10 <220>  
<223> Polipéptido sintético
- 15 <220>  
<221> misc\_feature  
<222> (1)..(166)  
<223> /nota = "véase la memoria descriptiva tal como se presenta para una descripción detallada de las sustituciones y realizaciones preferidas"
- 20 <220>  
<221> MOD\_RES  
<222> (14)..(33)  
<223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- 25 <220>  
<221> MOD\_RES  
<222> (39)..(58)  
<223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- 30 <220>  
<221> MOD\_RES  
<222> (64)..(83)  
<223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- 35 <220>  
<221> MOD\_RES  
<222> (87)..(106)  
<223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- 40 <220>  
<221> MOD\_RES  
<222> (113)..(132)  
<223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- 45 <220>  
<221> MOD\_RES  
<222> (140)..(159)  
<223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- 50 <220>  
<221> MOD\_RES  
<222> (140)..(159)  
<223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- <400> 17

ES 2 814 558 T3

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

- 5 <210> 18
- <211> 167
- <212> PRT
- <213> Secuencia artificial
- 10 <220>
- <223> Polipéptido sintético
- 15 <220>
- <221> MOD\_RES
- <222> (14)..(33)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- 20 <220>
- <221> MOD\_RES
- <222> (39)..(58)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15

ES 2 814 558 T3

, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

- 5 <220>  
<221> MOD\_RES  
<222> (64)..(83)  
<223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- 10 <220>  
<221> MOD\_RES  
<222> (87)..(106)  
<223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- 15 <220>  
<221> MOD\_RES  
<222> (113)..(132)  
<223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- 20 <220>  
<221> MOD\_RES  
<222> (140)..(159)  
<223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- 25 <220>  
<221> MOD\_RES  
<222> (140)..(159)  
<223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- <400> 18

Val	Ser	Asp	Val	Pro	Arg	Asp	Leu	Glu	Val	Val	Ala	Ala	Xaa	Xaa	Xaa
1				5					10					15	
	Xaa														
			20					25					30		
Xaa	Leu	Leu	Ile	Ser	Trp	Xaa									
		35					40					45			
Xaa	Tyr	Arg	Ile	Thr	Tyr	Xaa									
	50					55					60				
Xaa															
65					70					75					80
Xaa	Xaa	Xaa	Phe	Thr	Val	Xaa									
			85						90					95	
Xaa	Ala	Thr	Ile	Ser	Gly	Leu									
			100					105					110		

30



ES 2 814 558 T3

Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Xaa Xaa Xaa  
 1 5 10 15

Xaa  
 20 25 30

Xaa Leu Leu Ile Ser Trp Xaa  
 35 40 45

Xaa Tyr Arg Ile Thr Tyr Xaa  
 50 55 60

Xaa  
 65 70 75 80

Xaa Xaa Xaa Phe Thr Val Xaa  
 85 90 95

Xaa Ala Thr Ile Ser Gly Leu  
 100 105 110

Xaa  
 115 120 125

Xaa Xaa Xaa Xaa Tyr Thr Ile Thr Val Tyr Ala Xaa Xaa Xaa Xaa Xaa  
 130 135 140

Xaa Ile  
 145 150 155 160

Ser Ile Asn Tyr Arg Thr Glu Ile  
 165

- 5 <210> 20
- <211> 169
- <212> PRT
- <213> Secuencia artificial
- 10 <220>
- <223> Polipéptido sintético
- 15 <220>
- <221> MOD\_RES
- <222> (14)..(33)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- 20 <220>
- <221> MOD\_RES
- <222> (39)..(58)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

ES 2 814 558 T3

<220>  
 <221> MOD\_RES  
 <222> (64)..(83)  
 5 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15  
 , 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes  
  
 <220>  
 <221> MOD\_RES  
 10 <222> (87)..(106)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15  
 , 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes  
  
 <220>  
 <221> MOD\_RES  
 15 <222> (113)..(132)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15  
 , 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes  
  
 <220>  
 <221> MOD\_RES  
 20 <222> (140)..(159)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15  
 , 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes  
  
 25 <400> 20

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

ES 2 814 558 T3

Xaa Ala Thr Ile Ser Gly Leu  
 100 105 110

Xaa  
 115 120 125

Xaa Xaa Xaa Xaa Tyr Thr Ile Thr Val Tyr Ala Xaa Xaa Xaa Xaa Xaa  
 130 135 140

Xaa Ile  
 145 150 155 160

Ser Ile Asn Tyr Arg Thr Glu Ile Asp  
 165

- 5 <210> 21
- <211> 169
- <212> PRT
- <213> Secuencia artificial
  
- 10 <220>
- <223> Polipéptido sintético
  
- 15 <220>
- <221> MOD\_RES
- <222> (14)..(33)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
  
- 20 <220>
- <221> MOD\_RES
- <222> (39)..(58)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
  
- 25 <220>
- <221> MOD\_RES
- <222> (64)..(83)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
  
- 30 <220>
- <221> MOD\_RES
- <222> (87)..(106)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
  
- 35 <220>
- <221> MOD\_RES
- <222> (113)..(132)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
  
- 40 <220>
- <221> MOD\_RES
- <222> (140)..(159)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
  
- 45 <400> 21

ES 2 814 558 T3

Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Xaa Xaa Xaa  
 1 5 10 15

Xaa  
 20 25 30

Xaa Leu Leu Ile Ser Trp Xaa  
 35 40 45

Xaa Tyr Arg Ile Thr Tyr Xaa  
 50 55 60

Xaa  
 65 70 75 80

Xaa Xaa Xaa Phe Thr Val Xaa  
 85 90 95

Xaa Ala Thr Ile Ser Gly Leu  
 100 105 110

Xaa  
 115 120 125

Xaa Xaa Xaa Xaa Tyr Thr Ile Thr Val Tyr Ala Xaa Xaa Xaa Xaa Xaa  
 130 135 140

Xaa Ile  
 145 150 155 160

Ser Ile Asn Tyr Arg Thr Glu Ile Glu  
 165

- 5 <210> 22
- <211> 170
- <212> PRT
- <213> Secuencia artificial
- 10 <220>
- <223> Polipéptido sintético
- 15 <220>
- <221> MOD\_RES
- <222> (14).. (33)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- 20 <220>
- <221> MOD\_RES
- <222> (39)..(58)
- <223> Cualquier aminoácido

# ES 2 814 558 T3

- 5 <220>  
 <221> MOD\_RES  
 <222> (64)..(83)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
  
- 10 <220>  
 <221> MOD\_RES  
 <222> (87)..(106)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
  
- 15 <220>  
 <221> MOD\_RES  
 <222> (113)..(132)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
  
- 20 <220>  
 <221> MOD\_RES  
 <222> (140)..(159)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
  
- 25 <400> 22

Val	Ser	Asp	Val	Pro	Arg	Asp	Leu	Glu	Val	Val	Ala	Ala	Xaa	Xaa	Xaa
1				5					10					15	
	Xaa														
			20					25					30		
	Xaa	Leu	Leu	Ile	Ser	Trp	Xaa								
			35					40					45		
	Xaa	Tyr	Arg	Ile	Thr	Tyr	Xaa								
		50					55				60				
	Xaa														
						70				75					80

ES 2 814 558 T3

Xaa Xaa Xaa Phe Thr Val Xaa  
85 90 95

Xaa Ala Thr Ile Ser Gly Leu  
100 105 110

Xaa  
115 120 125

Xaa Xaa Xaa Xaa Tyr Thr Ile Thr Val Tyr Ala Xaa Xaa Xaa Xaa Xaa  
130 135 140

Xaa Ile  
145 150 155 160

Ser Ile Asn Tyr Arg Thr Glu Ile Asp Lys  
165 170

- 5 <210> 23
- <211> 170
- <212> PRT
- <213> Secuencia artificial
  
- 10 <220>
- <223> Polipéptido sintético
  
- 15 <220>
- <221> MOD\_RES
- <222> (14)..(33)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
  
- 20 <220>
- <221> MOD\_RES
- <222> (39)..(58)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
  
- 25 <220>
- <221> MOD\_RES
- <222> (64)..(83)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
  
- 30 <220>
- <221> MOD\_RES
- <222> (87)..(106)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
  
- 35 <220>
- <221> MOD\_RES
- <222> (113)..(132)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
  
- 40 <220>
- <221> MOD\_RES

ES 2 814 558 T3

<222> (140)..(159)

<223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

5 <400> 23

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

10 <210> 24  
 <211> 171  
 <212> PRT  
 <213> Secuencia artificial

15 <220>  
 <223> Polipéptido sintético

<220>  
 <221> MOD\_RES  
 <222> (14)..(33)  
 20 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

# ES 2 814 558 T3

<220>  
 <221> MOD\_RES  
 <222> (39)..(58)  
 5 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15  
 , 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

<220>  
 <221> MOD\_RES  
 10 <222> (64)..(83)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15  
 , 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

<220>  
 <221> MOD\_RES  
 15 <222> (87)..(106)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15  
 , 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

<220>  
 <221> MOD\_RES  
 20 <222> (113)..(132)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15  
 , 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

<220>  
 <221> MOD\_RES  
 25 <222> (140)..(159)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15  
 , 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

30 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15  
 , 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

<400> 24

Val	Ser	Asp	Val	Pro	Arg	Asp	Leu	Glu	Val	Val	Ala	Ala	Xaa	Xaa	Xaa
1				5					10					15	
	Xaa														
			20					25					30		
	Xaa	Leu	Leu	Ile	Ser	Trp	Xaa								
		35					40					45			
	Xaa	Tyr	Arg	Ile	Thr	Tyr	Xaa								
	50					55				60					

35

ES 2 814 558 T3

Xaa  
65 70 75 80

Xaa Xaa Xaa Phe Thr Val Xaa  
85 90 95

Xaa Ala Thr Ile Ser Gly Leu  
100 105 110

Xaa  
115 120 125

Xaa Xaa Xaa Xaa Tyr Thr Ile Thr Val Tyr Ala Xaa Xaa Xaa Xaa Xaa  
130 135 140

Xaa Ile  
145 150 155 160

Ser Ile Asn Tyr Arg Thr Glu Ile Asp Lys Pro  
165 170

- <210> 25
- <211> 171
- 5 <212> PRT
- <213> Secuencia artificial
- <220>
- 10 <223> Polipéptido sintético
- <220>
- <221> MOD\_RES
- <222> (14)..(33)
- 15 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- <220>
- 20 <221> MOD\_RES
- <222> (39)..(58)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- <220>
- 25 <221> MOD\_RES
- <222> (64)..(83)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- <220>
- 30 <221> MOD\_RES
- <222> (87)..(106)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- <220>
- 35 <221> MOD\_RES
- <222> (113)..(132)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

ES 2 814 558 T3

<220>

<221> MOD\_RES

<222> (140)..(159)

5 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

<400> 25

Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Xaa Xaa Xaa  
1 5 10 15

Xaa  
20 25 30

Xaa Leu Leu Ile Ser Trp Xaa  
35 40 45

Xaa Tyr Arg Ile Thr Tyr Xaa  
50 55 60

Xaa  
65 70 75 80

Xaa Xaa Xaa Phe Thr Val Xaa  
85 90 95

Xaa Ala Thr Ile Ser Gly Leu  
100 105 110

Xaa  
115 120 125

Xaa Xaa Xaa Xaa Tyr Thr Ile Thr Val Tyr Ala Xaa Xaa Xaa Xaa Xaa  
130 135 140

Xaa Ile  
145 150 155 160

Ser Ile Asn Tyr Arg Thr Glu Ile Glu Lys Pro  
165 170

10

<210> 26

<211> 172

<212> PRT

15 <213> Secuencia artificial

<220>

<223> Polipéptido sintético

20

<220>

<221> MOD\_RES

## ES 2 814 558 T3

<222> (14).. (33)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

5

<220>  
 <221> MOD\_RES  
 <222> (39)..(58)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

10

<220>  
 <221> MOD\_RES  
 <222> (64).. (83)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

15

<220>  
 <221> MOD\_RES  
 <222> (87)..(106)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

20

<220>  
 <221> MOD\_RES  
 <222> (113).. (132)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

25

<220>  
 <221> MOD\_RES  
 <222> (140)..(159)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

30

<400> 26

	<b>Val</b>	<b>Ser</b>	<b>Asp</b>	<b>Val</b>	<b>Pro</b>	<b>Arg</b>	<b>Asp</b>	<b>Leu</b>	<b>Glu</b>	<b>Val</b>	<b>Val</b>	<b>Ala</b>	<b>Ala</b>	<b>Xaa</b>	<b>Xaa</b>	<b>Xaa</b>
	1				5					10					15	
	<b>Xaa</b>															
				20					25					30		
	<b>Xaa</b>	<b>Leu</b>	<b>Leu</b>	<b>Ile</b>	<b>Ser</b>	<b>Trp</b>	<b>Xaa</b>									

ES 2 814 558 T3

35	40	45																			
Xaa	Tyr	Arg	Ile	Thr	Tyr	Xaa															
	50											60									
Xaa																					
65						70					75					80					
Xaa	Xaa	Xaa	Phe	Thr	Val	Xaa															
				85							90					95					
Xaa	Ala	Thr	Ile	Ser	Gly	Leu															
			100							105					110						
Xaa																					
			115							120					125						
Xaa	Xaa	Xaa	Xaa	Tyr	Thr	Ile	Thr	Val	Tyr	Ala	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa					
			130							135					140						
Xaa	Ile																				
145						150								155		160					
Ser	Ile	Asn	Tyr	Arg	Thr	Glu	Ile	Asp	Lys	Pro	Ser										
				165							170										

- <210> 27
- <211> 172
- 5 <212> PRT
- <213> Secuencia artificial
  
- <220>
- 10 <223> Polipéptido sintético
  
- <220>
- <221> MOD\_RES
- <222> (14)..(33)
- 15 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
  
- <220>
- <221> MOD\_RES
- <222> (39)..(58)
- 20 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
  
- <220>
- <221> MOD\_RES
- <222> (64)..(83)
- 25 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
  
- <220>
- 30 <221> MOD\_RES
- <222> (87)..(106)

ES 2 814 558 T3

<223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

<220>

5 <221> MOD\_RES

<222> (113)..(132)

<223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

10 <220>

<221> MOD\_RES

<222> (140)..(159)

<223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

15

<400> 27

Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Xaa Xaa Xaa  
1 5 10 15

Xaa  
20 25 30

Xaa Leu Leu Ile Ser Trp Xaa  
35 40 45

Xaa Tyr Arg Ile Thr Tyr Xaa  
50 55 60

Xaa  
65 70 75 80

Xaa Xaa Xaa Phe Thr Val Xaa  
85 90 95

Xaa Ala Thr Ile Ser Gly Leu  
100 105 110

Xaa  
115 120 125

Xaa Xaa Xaa Xaa Tyr Thr Ile Thr Val Tyr Ala Xaa Xaa Xaa Xaa Xaa  
130 135 140

Xaa Ile  
145 150 155 160

Ser Ile Asn Tyr Arg Thr Glu Ile Glu Lys Pro Ser  
165 170

20

<210> 28

ES 2 814 558 T3

- <211> 173  
 <212> PRT  
 <213> Secuencia artificial
- 5 <220>  
 <223> Polipéptido sintético
- <220>  
 <221> MOD\_RES  
 10 <222> (14)..(33)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- <220>  
 <221> MOD\_RES  
 15 <222> (39)..(58)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- <220>  
 <221> MOD\_RES  
 20 <222> (64)..(83)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- <220>  
 <221> MOD\_RES  
 25 <222> (87)..(106)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- <220>  
 <221> MOD\_RES  
 30 <222> (113)..(132)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- <220>  
 <221> MOD\_RES  
 35 <222> (140)..(159)  
 <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
- <400> 28  
 45

Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Xaa Xaa Xaa  
 1 5 10 15

ES 2 814 558 T3

Xaa  
 20 25 30

Xaa Leu Leu Ile Ser Trp Xaa  
 35 40 45

Xaa Tyr Arg Ile Thr Tyr Xaa  
 50 55 60

Xaa  
 65 70 75 80

Xaa Xaa Xaa Phe Thr Val Xaa  
 85 90 95

Xaa Ala Thr Ile Ser Gly Leu  
 100 105 110

Xaa  
 115 120 125

Xaa Xaa Xaa Xaa Tyr Thr Ile Thr Val Tyr Ala Xaa Xaa Xaa Xaa Xaa  
 130 135 140

Xaa Ile  
 145 150 155 160

Ser Ile Asn Tyr Arg Thr Glu Ile Asp Lys Pro Ser Gln  
 165 170

- 5 <210> 29
- <211> 173
- <212> PRT
- <213> Secuencia artificial
  
- 10 <220>
- <223> Polipéptido sintético
  
- 15 <220>
- <221> MOD\_RES
- <222> (14)..(33)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
  
- 20 <220>
- <221> MOD\_RES
- <222> (39)..(58)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes
  
- 25 <220>
- <221> MOD\_RES
- <222> (64)..(83)
- <223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15

# ES 2 814 558 T3

, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

<220>

<221> MOD\_RES

5 <222> (87)..(106)

<223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

<220>

10 <221> MOD\_RES

<222> (113)..(132)

<223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

<220>

15 <221> MOD\_RES

<222> (140)..(159)

<223> Cualquier aminoácido y esta región pueden abarcar 2-20, 2-15, 2-10, 2-8, 5-20, 5-15, 5-10, 5-8, 6-20, 6-15, 6-10, 6-8, 2-7, 5-7, o 6-7 restos, en donde algunas posiciones pueden estar ausentes

20

<400> 29

Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Xaa Xaa Xaa  
1 5 10 15

Xaa  
20 25 30

Xaa Leu Leu Ile Ser Trp Xaa  
35 40 45

Xaa Tyr Arg Ile Thr Tyr Xaa  
50 55 60

Xaa  
65 70 75 80

Xaa Xaa Xaa Phe Thr Val Xaa  
85 90 95

Xaa Ala Thr Ile Ser Gly Leu  
100 105 110

Xaa  
115 120 125

Xaa Xaa Xaa Xaa Tyr Thr Ile Thr Val Tyr Ala Xaa Xaa Xaa Xaa Xaa

ES 2 814 558 T3

130

135

140

Xaa Ile  
145 150 155 160

Ser Ile Asn Tyr Arg Thr Glu Ile Glu Lys Pro Ser Gln  
165 170

5 <210> 30  
<211> 10  
<212> PRT  
<213> Secuencia artificial

10 <220>  
<223> Péptido sintético  
<400> 30

Met Gly Val Ser Asp Val Pro Arg Asp Leu  
1 5 10

15 <210> 31  
<211> 9  
<212> PRT  
<213> Secuencia artificial

20 <220>  
<223> Péptido sintético  
<400> 31

Gly Val Ser Asp Val Pro Arg Asp Leu  
1 5

25 <210> 32  
<211> 9  
<212> PRT  
<213> Secuencia artificial

30 <220>  
<223> Péptido sintético

35 <220>  
<221> MOD\_RES  
<222> (1)..(2)  
<223> Esta región puede abarcar 0-2 restos seleccionados de Met, Gly o "Met Gly" en donde algunas posiciones pueden estar ausentes

40 <220>  
<221> VARIANTE  
<222> (1)..(2)  
<223> /sustitución=" "

45 <220>  
<221> misc\_feature  
<222> (1)..(8)  
50 <223> / nota="Las variantes de restos dadas en la secuencia no tienen preferencia con respecto a aquellas en las anotaciones para variantes de posiciones"

<220>  
<221> misc\_feature

<222> (1)..(8)

<223> /nota = "véase la memoria descriptiva tal como se presenta para una descripción detallada de las sustituciones y realizaciones preferidas"

5 <400> 32

**Met Gly Ser Asp Val Pro Arg Asp Leu**  
**1 5**

<210> 33

10 <211> 8

<212> PRT

<213> Secuencia artificial

<220>

15 <223> Péptido sintético

<220>

<221> MOD\_RES

<222> (1)..(2)

20 <223> Esta región puede abarcar 0-2 restos seleccionados de Met, Gly o "Met Gly" en donde algunas posiciones pueden estar ausentes

<220>

<221> VARIANTE

25 <222> (1)..(2)

<223> /sustitución=" "

<220>

<221> misc\_feature

30 <222> (1)..(8)

<223> /nota="Las variantes de restos dadas en la secuencia no tiene preferencia respecto a aquellas en las anotaciones para variantes de posición"

<220>

35 <221> misc\_feature

<222> (1)..(8)

<223> /nota = "véase la memoria descriptiva tal como se presenta para una descripción detallada de las sustituciones y realizaciones preferidas"

40 <400> 33

**Met Gly Asp Val Pro Arg Asp Leu**  
**1 5**

<210> 34

45 <211> 7

<212> PRT

<213> Secuencia artificial

<220>

50 <223> Péptido sintético

<220>

<221> MOD\_RES

<222> (1)..(2)

55 <223> Esta región puede abarcar 0-2 restos seleccionados de Met, Gly o "Met Gly" en donde algunas posiciones pueden estar ausentes

<220>

<221> VARIANTE

60 <222> (1)..(2)

<223> /sustitución=" "

<220>



pueden estar ausentes

- 5 <220>  
<221> VARIANTE  
<222> (1)..(2)  
<223> /sustitución=" "
- 10 <220>  
<221> misc\_feature  
<222> (1)..(8)  
<223> /nota="Las variantes de restos dadas en la secuencia no tiene preferencia respecto a aquellas en las anotaciones para variantes de posición"
- 15 <220>  
<221> misc\_feature  
<222> (1)..(8)  
<223> /nota = "véase la memoria descriptiva tal como se presenta para una descripción detallada de las sustituciones y realizaciones preferidas"
- 20 <400> 36

**Met Gly Arg Asp Leu**  
**1 5**

- 25 <210> 37  
<211> 4  
<212> PRT  
<213> Secuencia artificial
- 30 <220>  
<223> Péptido sintético
- 35 <220>  
<221> MOD\_RES  
<222> (1)..(2)  
<223> Esta región puede abarcar 0-2 restos seleccionados de Met, Gly o "Met Gly" en donde algunas posiciones pueden estar ausentes
- 40 <220>  
<221> VARIANTE  
<222> (1)..(2)  
<223> /sustitución=" "
- 45 <220>  
<221> misc\_feature  
<222> (1)..(8)  
<223> /nota="Las variantes de restos dadas en la secuencia no tiene preferencia respecto a aquellas en las anotaciones para variantes de posición"
- 50 <220>  
<221> misc\_feature  
<222> (1)..(8)  
<223> /nota = "véase la memoria descriptiva tal como se presenta para una descripción detallada de las sustituciones y realizaciones preferidas"
- 55 <400> 37

**Met Gly Asp Leu**  
**1**

- 60 <210> 38  
<211> 6  
<212> PRT  
<213> Secuencia artificial

ES 2 814 558 T3

<220>  
<223> Péptido sintético

5 <400> 38

Met Ala Ser Thr Ser Gly  
1 5

10 <210> 39  
<211> 4  
<212> PRT  
<213> Secuencia artificial

15 <220>  
<223> Péptido sintético

<400> 39

Glu Ile Glu Lys  
1

20 <210> 40  
<211> 5  
<212> PRT  
<213> Secuencia artificial

25 <220>  
<223> Péptido sintético

<400> 40

30 Glu Gly Ser Gly Cys  
1 5

35 <210> 41  
<211> 7  
<212> PRT  
<213> Secuencia artificial

40 <220>  
<223> Péptido sintético

<400> 41

Glu Ile Glu Lys Pro Cys Gln  
1 5

45 <210> 42  
<211> 7  
<212> PRT  
<213> Secuencia artificial

50 <220>  
<223> Péptido sintético

<400> 42

55 Glu Ile Glu Lys Pro Ser Gln  
1 5

<210> 43

ES 2 814 558 T3

<211> 5  
 <212> PRT  
 <213> Secuencia artificial  
 5 <220>  
 <223> Péptido sintético  
 <400> 43  
  
 Glu Ile Glu Lys Pro  
 1 5  
 10  
 <210> 44  
 <211> 6  
 <212> PRT  
 15 <213> Secuencia artificial  
 <220>  
 <223> Péptido sintético  
 20 <400> 44  
  
 Glu Ile Glu Lys Pro Ser  
 1 5  
 25  
 <210> 45  
 <211> 6  
 <212> PRT  
 <213> Secuencia artificial  
 30 <220>  
 <223> Péptido sintético  
 <400> 45  
  
 Glu Ile Glu Lys Pro Cys  
 1 5  
 35  
 <210> 46  
 <211> 6  
 <212> PRT  
 <213> Secuencia artificial  
 40 <220>  
 <223> Marcador sintético 6xHis  
 <400> 46  
 45  
  
 His His His His His His  
 1 5  
 50  
 <210> 47  
 <211> 4  
 <212> PRT  
 <213> Secuencia artificial  
 55 <220>  
 <223> Péptido sintético  
 <400> 47

ES 2 814 558 T3

Glu Ile Asp Lys  
1

5 <210> 48  
<211> 7  
<212> PRT  
<213> Secuencia artificial

10 <220>  
<223> Péptido sintético  
<400> 48

Glu Ile Asp Lys Pro Cys Gln  
1 5

15 <210> 49  
<211> 7  
<212> PRT  
<213> Secuencia artificial

20 <220>  
<223> Péptido sintético  
<400> 49

Glu Ile Asp Lys Pro Ser Gln  
1 5

25  
30 <210> 50  
<211> 91  
<212> PRT  
<213> Secuencia artificial

<220>  
<223> Polipéptido sintético  
35 <400> 50

ES 2 814 558 T3

Met Pro Ala Pro Thr Asp Leu Arg Phe Thr Asn Glu Thr Pro Ser Ser  
 1 5 10 15

Leu Leu Ile Ser Trp Thr Pro Pro Arg Val Gln Ile Thr Gly Tyr Ile  
 20 25 30

Ile Arg Tyr Gly Pro Val Gly Ser Asp Gly Arg Val Lys Glu Phe Thr  
 35 40 45

Val Pro Pro Ser Val Ser Ser Ala Thr Ile Thr Gly Leu Lys Pro Gly  
 50 55 60

Thr Glu Tyr Thr Ile Ser Val Ile Ala Leu Lys Asp Asn Gln Glu Ser  
 65 70 75 80

Glu Pro Leu Arg Gly Arg Val Thr Thr Gly Gly  
 85 90

5 <210> 51  
 <211> 4  
 <212> PRT  
 <213> Secuencia artificial

10 <220>  
 <223> Péptido sintético  
 <400> 51

Thr Pro Ser Ser  
 1

15 <210> 52  
 <211> 7  
 <212> PRT  
 <213> Secuencia artificial

20 <220>  
 <223> Péptido sintético  
 <400> 52

25 Thr Pro Pro Arg Val Gln Ile  
 1 5

30 <210> 53  
 <211> 6  
 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> Péptido sintético  
 <400> 53

Val Gly Ser Asp Gly Arg  
 1 5

ES 2 814 558 T3

5 <210> 54  
 <211> 4  
 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> Péptido sintético

10 <400> 54

Pro Ser Val Ser  
 1

15 <210> 55  
 <211> 5  
 <212> PRT  
 <213> Secuencia artificial

20 <220>  
 <223> Péptido sintético

<400> 55

Gly Leu Lys Pro Gly  
 1 5

25 <210> 56  
 <211> 8  
 <212> PRT  
 <213> Secuencia artificial

30 <220>  
 <223> Péptido sintético

<400> 56

35 Lys Asp Asn Gln Glu Ser Glu Pro  
 1 5

40 <210> 57  
 <211> 89  
 <212> PRT  
 <213> Secuencia artificial

45 <220>  
 <223> Polipéptido sintético

<400> 57

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

ES 2 814 558 T3

Ile Thr Val Ser Trp Thr Pro Pro Ser Ala Thr Ile Thr Gly Tyr Arg  
 20 25 30

Ile Thr Tyr Thr Pro Ser Asn Gly Pro Gly Glu Pro Lys Glu Leu Thr  
 35 40 45

Val Pro Pro Ser Ser Thr Ser Val Thr Ile Thr Gly Ile Thr Pro Gly  
 50 55 60

Val Glu Tyr Val Val Ser Val Tyr Ala Leu Lys Asp Asn Gln Glu Ser  
 65 70 75 80

Pro Pro Leu Val Gly Thr Cys Thr Thr  
 85

<210> 58  
 <211> 89  
 <212> PRT  
 <213> Secuencia artificial

5

<220>  
 <223> Polipéptido sintético

10

<400> 58

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

Leu Arg Leu Ser Trp Thr Ala Pro Asp Ala Ala Phe Asp Ser Phe Leu  
 20 25 30

Ile Gln Tyr Gln Glu Ser Glu Lys Val Gly Glu Ala Ile Asn Leu Thr  
 35 40 45

Val Pro Gly Ser Glu Arg Ser Tyr Asp Leu Thr Gly Leu Lys Pro Gly  
 50 55 60

Thr Glu Tyr Thr Val Ser Ile Tyr Gly Val Lys Gly Gly His Arg Ser  
 65 70 75 80

Asn Pro Leu Ser Ala Glu Phe Thr Thr  
 85

<210> 59  
 <211> 4  
 <212> PRT  
 <213> Secuencia artificial

15

<220>  
 <223> Péptido sintético

20

<400> 59

ES 2 814 558 T3

Thr Glu Asp Ser  
1

5 <210> 60  
<211> 7  
<212> PRT  
<213> Secuencia artificial

10 <220>  
<223> Péptido sintético  
<400> 60

Thr Ala Pro Asp Ala Ala Phe  
1 5

15 <210> 61  
<211> 6  
<212> PRT  
<213> Secuencia artificial

20 <220>  
<223> Péptido sintético  
<400> 61

Ser Glu Lys Val Gly Glu  
1 5

30 <210> 62  
<211> 4  
<212> PRT  
<213> Secuencia artificial

35 <220>  
<223> Péptido sintético  
<400> 62

Gly Ser Glu Arg  
1

40 <210> 63  
<211> 5  
<212> PRT  
<213> Secuencia artificial

45 <220>  
<223> Péptido sintético  
<400> 63

Gly Leu Lys Pro Gly  
1 5

50 <210> 64  
<211> 7  
<212> PRT  
<213> Secuencia artificial

55 <220>

ES 2 814 558 T3

<223> Péptido sintético

<400> 64

Lys Gly Gly His Arg Ser Asn  
1 5

5

<210> 65

<211> 93

<212> PRT

10 <213> *Homo sapiens*

<400> 65

Pro Leu Ser Pro Pro Thr Asn Leu His Leu Glu Ala Asn Pro Asp Thr  
1 5 10 15

Gly Val Leu Thr Val Ser Trp Glu Arg Ser Thr Thr Pro Asp Ile Thr  
20 25 30

Gly Tyr Arg Ile Thr Thr Thr Pro Thr Asn Gly Gln Gln Gly Asn Ser  
35 40 45

Leu Glu Glu Val Val His Ala Asp Gln Ser Ser Cys Thr Phe Asp Asn  
50 55 60

Leu Ser Pro Gly Leu Glu Tyr Asn Val Ser Val Tyr Thr Val Lys Asp  
65 70 75 80

Asp Lys Glu Ser Val Pro Ile Ser Asp Thr Ile Ile Pro  
85 90

15

<210> 66

<211> 89

<212> PRT

20 <213> *Homo sapiens*

<400> 66

Asn Val Ser Pro Pro Arg Arg Ala Arg Val Thr Asp Ala Thr Glu Thr  
1 5 10 15

Thr Ile Thr Ile Ser Trp Arg Thr Lys Thr Glu Thr Ile Thr Gly Phe  
20 25 30

Gln Val Asp Ala Val Pro Ala Asn Gly Gln Thr Pro Ile Gln Arg Thr  
35 40 45

Ile Lys Pro Asp Val Arg Ser Tyr Thr Ile Thr Gly Leu Gln Pro Gly

ES 2 814 558 T3

50

55

60

Thr Asp Tyr Lys Ile Tyr Leu Tyr Thr Leu Asn Asp Asn Ala Arg Ser  
65 70 75 80

Ser Pro Val Val Ile Asp Ala Ser Thr  
85

<210> 67  
<211> 108  
<212> PRT  
<213> *Homo sapiens*

5

<400> 67

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

Pro Thr Ser Leu Leu Ile Ser Trp Asp Ser Gly Arg Gly Ser Tyr Gln  
20 25 30

Tyr Tyr Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln  
35 40 45

Glu Phe Thr Val Pro Gly Pro Val His Thr Ala Thr Ile Ser Gly Leu  
50 55 60

Lys Pro Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Asp His  
65 70 75 80

Lys Pro His Ala Asp Gly Pro His Thr Tyr His Glu Ser Pro Ile Ser  
85 90 95

Ile Asn Tyr Arg Thr Glu Ile Asp Lys Pro Ser Gln  
100 105

10

<210> 68  
<211> 108  
<212> PRT  
<213> *Homo sapiens*

15

<400> 68

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

Pro Thr Ser Leu Leu Ile Ser Trp Tyr Trp Glu Gly Leu Pro Tyr Gln  
20 25 30

Tyr Tyr Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln  
35 40 45

ES 2 814 558 T3

Glu Phe Thr Val Pro Arg Asp Val Asn Thr Ala Thr Ile Ser Gly Leu  
 50 55 60

Lys Pro Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Asp Trp  
 65 70 75 80

Tyr Asn Pro Asp Thr His Glu Tyr Ile Tyr His Thr Ile Pro Ile Ser  
 85 90 95

Ile Asn Tyr Arg Thr Glu Ile Asp Lys Pro Ser Gln  
 100 105

5 <210> 69  
 <211> 108  
 <212> PRT  
 <213> *Homo sapiens*

10 <400> 69

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

Pro Thr Ser Leu Leu Ile Ser Trp Asp Ser Gly Arg Gly Ser Tyr Gln  
 20 25 30

Tyr Tyr Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln  
 35 40 45

Glu Phe Thr Val Pro Gly Pro Val His Thr Ala Glu Ile Ser Gly Leu  
 50 55 60

Lys Pro Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Asp His  
 65 70 75 80

Lys Pro His Ala Asp Gly Pro His Thr Tyr His Glu Ser Pro Ile Ser  
 85 90 95

Ile Asn Tyr Arg Thr Glu Ile Asp Lys Pro Ser Gln  
 100 105

15 <210> 70  
 <211> 108  
 <212> PRT  
 <213> *Homo sapiens*

<400> 70

20 Met Gly Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Thr  
 1 5 10 15

ES 2 814 558 T3

Pro Thr Ser Leu Leu Ile Ser Trp Asp Ser Gly Arg Gly Ser Tyr Gln  
 20 25 30

Tyr Tyr Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln  
 35 40 45

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

Lys Pro Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Asp His  
 65 70 75 80

Lys Pro His Ala Asp Gly Pro His Thr Tyr His Glu Ser Pro Ile Ser  
 85 90 95

Ile Asn Tyr Arg Thr Glu Ile Asp Lys Pro Ser Gln  
 100 105

<210> 71  
 <211> 108  
 <212> PRT  
 <213> *Homo sapiens*

5

<400> 71

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

Pro Thr Ser Leu Leu Ile Ser Trp Tyr Trp Glu Gly Leu Pro Tyr Gln  
 20 25 30

Tyr Tyr Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln  
 35 40 45

Glu Phe Thr Val Pro Arg Asp Val Asn Thr Ala Glu Ile Ser Gly Leu  
 50 55 60

Lys Pro Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Asp Trp  
 65 70 75 80

Tyr Asn Pro Asp Thr His Glu Tyr Ile Tyr His Thr Ile Pro Ile Ser  
 85 90 95

Ile Asn Tyr Arg Thr Glu Ile Asp Lys Pro Ser Gln  
 100 105

10

<210> 72  
 <211> 5  
 <212> PRT  
 <213> *Homo sapiens*

15



ES 2 814 558 T3

<400> 78

Val Thr Ser Ser Tyr Lys Tyr Asp Met Gln Tyr  
 1 5 10

5

<210> 79  
 <211> 115  
 <212> PRT  
 <213> Secuencia artificial

10

<220>  
 <223> Polipéptido sintético

15

<400> 79

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

Pro Thr Ser Leu Leu Ile Ser Trp Asp Ser Gly Arg Gly Ser Tyr Gln  
 20 25 30

Tyr Tyr Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln  
 35 40 45

Glu Phe Thr Val Pro Gly Pro Val His Thr Ala Thr Ile Ser Gly Leu  
 50 55 60

Lys Pro Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Asp Asp  
 65 70 75 80

His Lys Pro His Ala Asp Gly Pro His Thr Tyr His Glu Ser Pro Ile  
 85 90 95

Ser Ile Asn Tyr Arg Thr Glu Ile Asp Lys Pro Ser Gln His His His  
 100 105 110

His His His  
 115

20

<210> 80  
 <211> 109  
 <212> PRT  
 <213> Secuencia artificial

25

<220>  
 <223> Polipéptido sintético

<400> 80

ES 2 814 558 T3

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

Pro Thr Ser Leu Leu Ile Ser Trp Glu His Asp Tyr Pro Tyr Arg Arg  
 20 25 30

Tyr Tyr Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln  
 35 40 45

Glu Phe Thr Val Pro Lys Asp Val Asp Thr Ala Thr Ile Ser Gly Leu  
 50 55 60

Lys Pro Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Ser Ser  
 65 70 75 80

Tyr Lys Tyr Asp Met Gln Tyr Ser Pro Ile Ser Ile Asn Tyr Arg Thr  
 85 90 95

Glu Ile Asp Lys Pro Ser Gln His His His His His His  
 100 105

<210> 81

<211> 112

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> Polipéptido sintético

<400> 81

ES 2 814 558 T3

Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Thr Pro Thr  
 1 5 10 15

Ser Leu Leu Ile Ser Trp Tyr Trp Glu Gly Leu Pro Tyr Gln Tyr Tyr  
 20 25 30

Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln Glu Phe  
 35 40 45

Thr Val Pro Arg Asp Val Asn Thr Ala Thr Ile Ser Gly Leu Lys Pro  
 50 55 60

Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Asp Trp Tyr Asn  
 65 70 75 80

Pro Asp Thr His Glu Tyr Ile Tyr His Thr Ile Pro Ile Ser Ile Asn  
 85 90 95

Tyr Arg Thr Glu Ile Asp Lys Pro Ser Gln His His His His His His

100

105

110

- 5 <210> 82
- <211> 112
- <212> PRT
- <213> Secuencia artificial
- 10 <220>
- <223> Polipéptido sintético
- <400> 82

ES 2 814 558 T3

Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Thr Pro Thr  
 1 5 10 15  
 Ser Leu Leu Ile Ser Trp Asp Ser Gly Arg Gly Ser Tyr Gln Tyr Tyr  
 20 25 30  
 Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln Glu Phe  
 35 40 45  
 Thr Val Pro Gly Pro Val His Thr Ala Thr Ile Ser Gly Leu Lys Pro  
 50 55 60  
 Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Asp His Lys Pro  
 65 70 75 80  
 His Ala Asp Gly Pro His Thr Tyr His Glu Ser Pro Ile Ser Ile Asn  
 85 90 95  
 Tyr Arg Thr Glu Ile Asp Lys Pro Ser Gln His His His His His His  
 100 105 110

5 <210> 83  
 <211> 5  
 <212> PRT  
 <213> *Homo sapiens*

<400> 83

10 Glu Phe Thr Val Pro  
 1 5

15 <210> 84  
 <211> 6  
 <212> PRT  
 <213> *Homo sapiens*

<400> 84

20 Glu His Asp Tyr Pro Tyr  
 1 5

25 <210> 85  
 <211> 10  
 <212> PRT  
 <213> *Homo sapiens*

<400> 85

30 Thr Ser Ser Tyr Lys Tyr Asp Met Gln Tyr  
 1 5 10

<210> 86  
 <211> 114  
 <212> PRT  
 <213> Secuencia artificial

ES 2 814 558 T3

<220>  
<223> Sintético

5 <400> 86

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

Pro Thr Ser Leu Leu Ile Ser Trp Asp Ser Gly Arg Gly Ser Tyr Gln  
20 25 30

Tyr Tyr Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln  
35 40 45

Glu Phe Thr Val Pro Gly Pro Val His Thr Ala Thr Ile Ser Gly Leu  
50 55 60

Lys Pro Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Asp His  
65 70 75 80

Lys Pro His Ala Asp Gly Pro His Thr Tyr His Glu Ser Pro Ile Ser  
85 90 95

Ile Asn Tyr Arg Thr Glu Ile Asp Lys Pro Ser Gln His His His His  
100 105 110

His His

10 <210> 87  
<211> 112  
<212> PRT  
<213> Secuencia artificial

15 <220>  
<223> Sintético

<220>  
<221> misc\_feature  
<222> (23)..(28)  
20 <223> Xaa puede ser cualquier aminoácido de origen natural

<220>  
<221> misc\_feature  
<222> (52)..(53)  
25 <223> Xaa puede ser cualquier aminoácido de origen natural

<220>  
<221> misc\_feature  
<222> (55)..(55)  
30 <223> Xaa puede ser cualquier aminoácido de origen natural

<220>  
<221> misc\_feature  
<222> (78)..(85)  
35 <223> Xaa puede ser cualquier aminoácido de origen natural

<220>  
 <221> misc\_feature  
 <222> (87)..(87)  
 5 <223> Xaa puede ser cualquier aminoácido de origen natural

<220>  
 <221> misc\_feature  
 <222> (90).. (91)  
 10 <223> Xaa puede ser cualquier aminoácido de origen natural

<400> 87

Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Thr Pro Thr  
 1 5 10 15

Ser Leu Leu Ile Ser Trp Xaa Xaa Xaa Xaa Xaa Tyr Gln Tyr Tyr  
 20 25 30

Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln Glu Phe  
 35 40 45

Thr Val Pro Xaa Xaa Val Xaa Thr Ala Thr Ile Ser Gly Leu Lys Pro  
 50 55 60

Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Asp Xaa Xaa Xaa  
 65 70 75 80

Xaa Xaa Xaa Xaa Xaa His Xaa Tyr His Xaa Xaa Pro Ile Ser Ile Asn  
 85 90 95

Tyr Arg Thr Glu Ile Asp Lys Pro Ser Gln His His His His His His  
 100 105 110

15

## REIVINDICACIONES

1. Un polipéptido que comprende restos de armazones a base de fibronectina que comprenden un 10<sup>o</sup> dominio de fibronectina tipo III (<sup>10</sup>F<sub>n3</sub>) modificado, en donde el dominio <sup>10</sup>F<sub>n3</sub> modificado comprende una secuencia de aminoácidos en la que el aminoácido correspondiente a la posición 58 de la SEQ ID NO: 1 es un aminoácido hidrófilo que no es un resto de treonina, en donde la solubilidad del dominio <sup>10</sup>F<sub>n3</sub> modificado está mejorada en relación con la solubilidad de un dominio <sup>10</sup>F<sub>n3</sub> que comprende la misma secuencia de aminoácidos que el dominio <sup>10</sup>F<sub>n3</sub> modificado, excepto que el resto de aminoácido en la posición 58 de la SEQ ID NO: 1 es treonina; y en donde el dominio <sup>10</sup>F<sub>n3</sub> modificado se une específicamente a una diana a la que no está unido mediante un dominio de F<sub>n3</sub> humano seleccionado de los establecidos en las SEQ ID NO: 1-16, 65 y 66.
2. El polipéptido de la reivindicación 1, en donde la secuencia de aminoácidos del dominio <sup>10</sup>F<sub>n3</sub> modificado es al menos el 50 % o al menos el 70 % idéntica a la de un dominio F<sub>n3</sub> humano seleccionado de los establecidos en las SEQ ID NO: 1-16, 65 y 66.
3. El polipéptido de las reivindicaciones 1 o 2, en donde la secuencia de aminoácidos del dominio <sup>10</sup>F<sub>n3</sub> modificado difiere en, como mucho, 40 aminoácidos de un dominio F<sub>n3</sub> humano seleccionado de los establecidos en la SEQ ID NO: 1-16, 65 y 66.
4. El polipéptido de una cualquiera de las reivindicaciones 1-3, en donde el dominio <sup>10</sup>F<sub>n3</sub> modificado comprende al menos una variación de aminoácido seleccionada de sustitución, delección y adición en al menos un bucle en comparación con un dominio F<sub>n3</sub> humano seleccionado de los establecidos en las SEQ ID NO: 1-16, 65 y 66.
5. El polipéptido de la reivindicación 4, en donde el dominio <sup>10</sup>F<sub>n3</sub> modificado comprende al menos una variación de aminoácido en cada uno de al menos dos bucles en comparación con un dominio F<sub>n3</sub> humano seleccionado de los establecidos en la SEQ ID NO: 1-16, 65 y 66.
6. El polipéptido de la reivindicación 5, en donde el dominio <sup>10</sup>F<sub>n3</sub> modificado comprende al menos una variación de aminoácido en cada uno de al menos tres bucles en comparación con un dominio F<sub>n3</sub> humano seleccionado de los establecidos en la SEQ ID NO: 1-16, 65 y 66.
7. El polipéptido de una cualquiera de las reivindicaciones 4-6, en donde el dominio <sup>10</sup>F<sub>n3</sub> modificado comprende al menos una variación de aminoácido en al menos un bucle seleccionado de bucles del polo norte (bucles BC, DE y FG), en comparación con un dominio F<sub>n3</sub> humano seleccionado de los establecidos en las SEQ ID NO: 1-16, 65 y 66.
8. El polipéptido de una cualquiera de las reivindicaciones 4-7, en donde el dominio <sup>10</sup>F<sub>n3</sub> modificado comprende al menos una variación de aminoácido en al menos un bucle seleccionado de bucles del polo sur (bucles AB, CD y EF), en comparación con un dominio F<sub>n3</sub> humano seleccionado de los establecidos en las SEQ ID NO: 1-16, 65 y 66.
9. El polipéptido de una cualquiera de las reivindicaciones 1-8, en donde el dominio <sup>10</sup>F<sub>n3</sub> modificado comprende al menos una variación de aminoácido seleccionada de sustitución, delección y adición, en al menos una región no de bucle y al menos un bucle, en comparación con un dominio F<sub>n3</sub> humano seleccionado de los establecidos en las SEQ ID NO: 1-16, 65 y 66.
10. El polipéptido de una cualquiera de las reivindicaciones 1-9, en donde el dominio <sup>10</sup>F<sub>n3</sub> modificado se une de manera específica a una diana con una K<sub>d</sub> de menos de 500 nM o menos de 100 nM.
11. El polipéptido de una cualquiera de las reivindicaciones 1-10, en donde el aminoácido correspondiente a la posición 58 de la SEQ ID NO: 1 en el dominio <sup>10</sup>F<sub>n3</sub> modificado es un resto de ácido aspártico (D).
12. El polipéptido de una cualquiera de las reivindicaciones 1-11, que comprende adicionalmente al menos uno de los restos farmacocinéticos (PK) seleccionados de: un resto de polioxialquileo, una proteína de unión a albúmina sérica humana, ácido siálico, albúmina sérica humana, IgG, una proteína de unión a IgG, transferrina y un fragmento Fc.
13. El polipéptido de la reivindicación 12, en donde el resto de PK es un Fc o un fragmento de Fc.
14. Una biblioteca que comprende una pluralidad de polipéptidos de una cualquiera de las reivindicaciones 1-13.
15. Un ácido nucleico que codifica el polipéptido de cualquiera de las reivindicaciones 1-13.

<sup>14</sup> Fn3	01	NVSPRRR <b>ARVTD</b> A <b>TET</b> - <b>TITIS</b> WR-TKTETITG <b>FQVDA</b> VPANGQ--TP	44
<sup>10</sup> Fn3	01	VSDVPRD <b>LEVVA</b> ATPT- <b>SL</b> LISWD-APAVTVRY <b>YRITY</b> GETGGN-SPV	45
<sup>7</sup> Fn3	01	PLSPPTNL <b>HLEAN</b> PD <b>TG</b> VL <b>TVSW</b> ERSTTPDITG <b>YRITT</b> TPTNGQQGNS	44
		A    AB    B    BC    C    CD	
<sup>14</sup> Fn3	45	I <b>QRTIK</b> PDVRS <b>Y</b> TITGLQPGTDY <b>KIY</b> LYTLNDNARS----SP <b>VVIDA</b> ST	89
<sup>10</sup> Fn3	46	Q <b>EFTV</b> PGSKST <b>ATIS</b> GLKPGVDY <b>TITV</b> YAVTGRGDSPASSK <b>PISIN</b> YRT	94
<sup>7</sup> Fn3	49	LE <b>EVV</b> HADQSS <b>CTFD</b> NLSPGLE <b>YNVSV</b> YTVKDDKES----VP <b>ISDTI</b> IP	93
		D   DE   E   EF   F   FG   G	

Figura 1

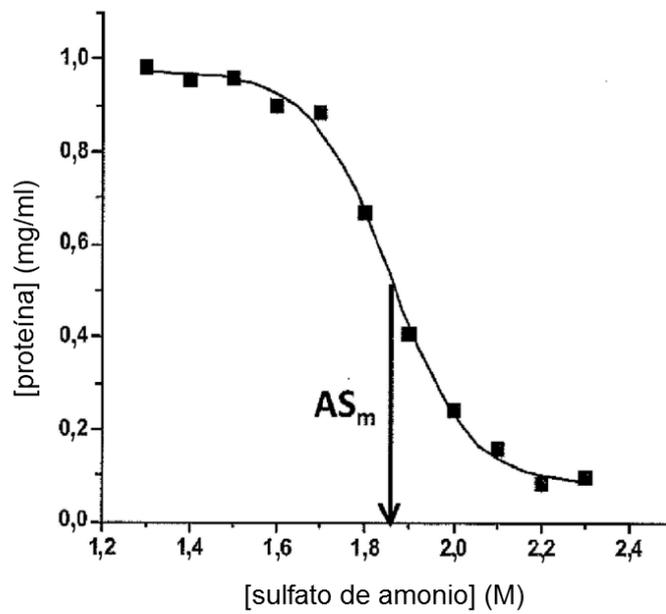


Figura 2

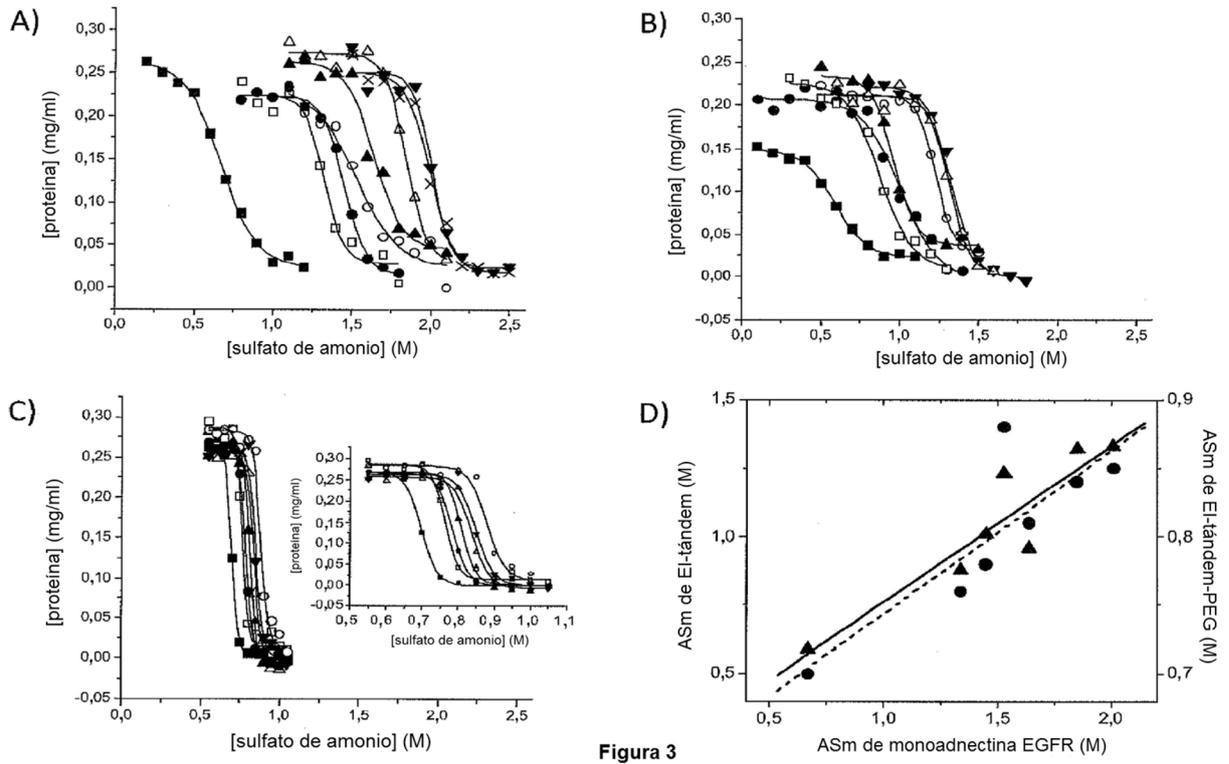


Figura 3

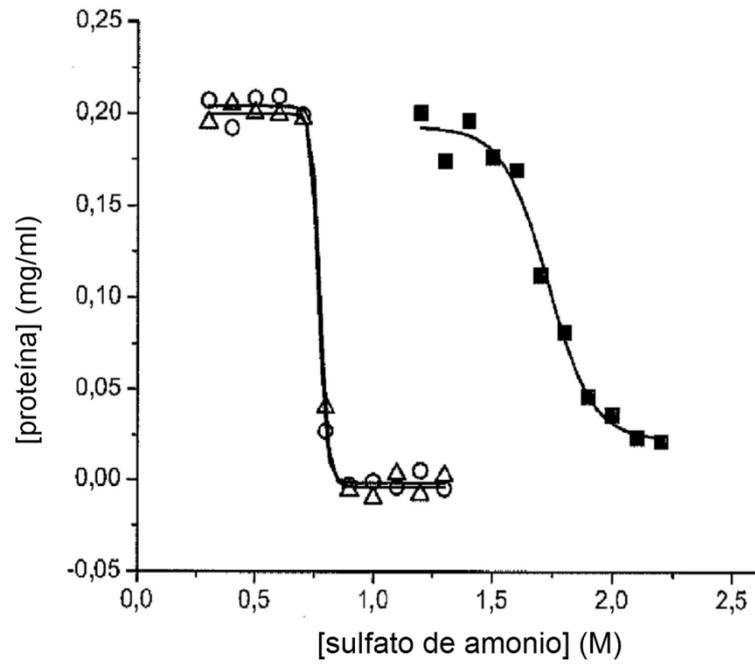


Figura 4

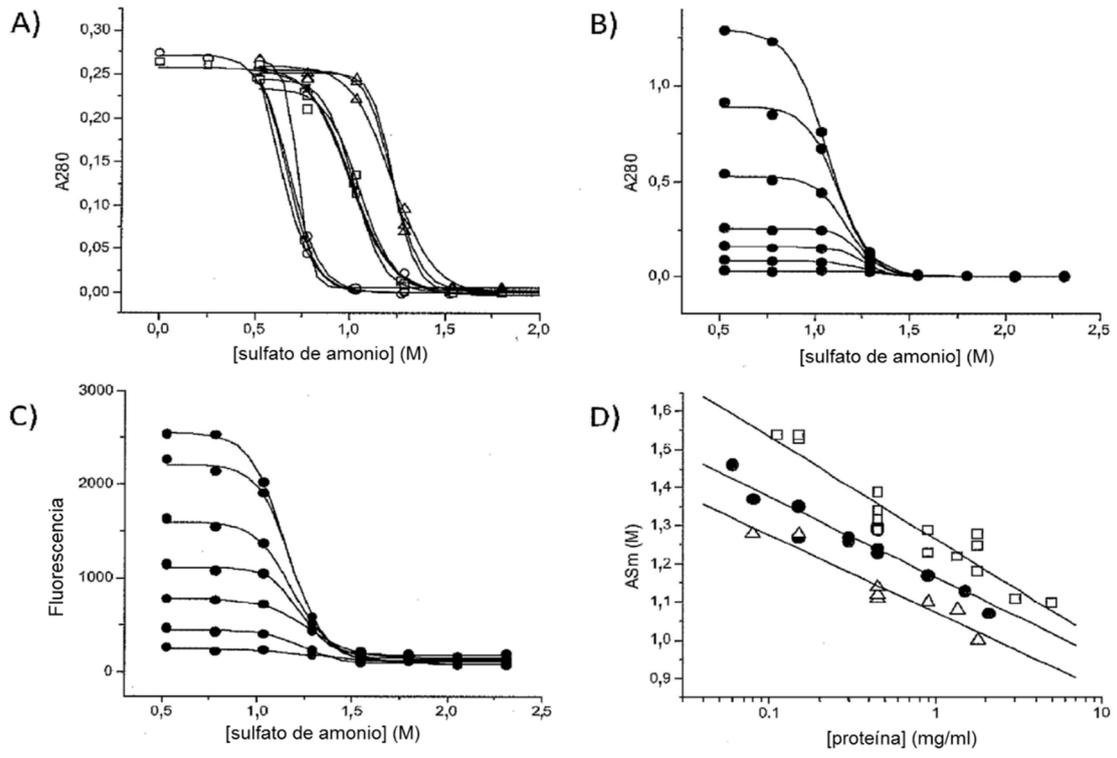


Figura 5

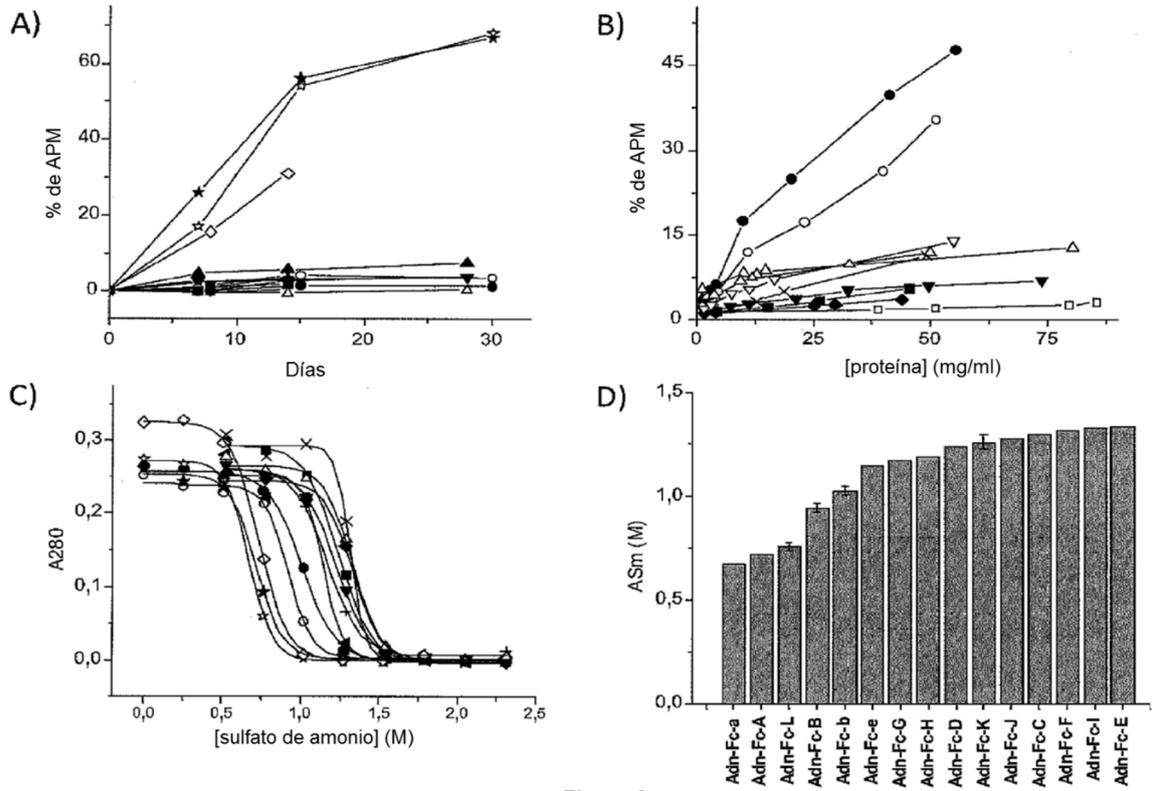


Figura 6

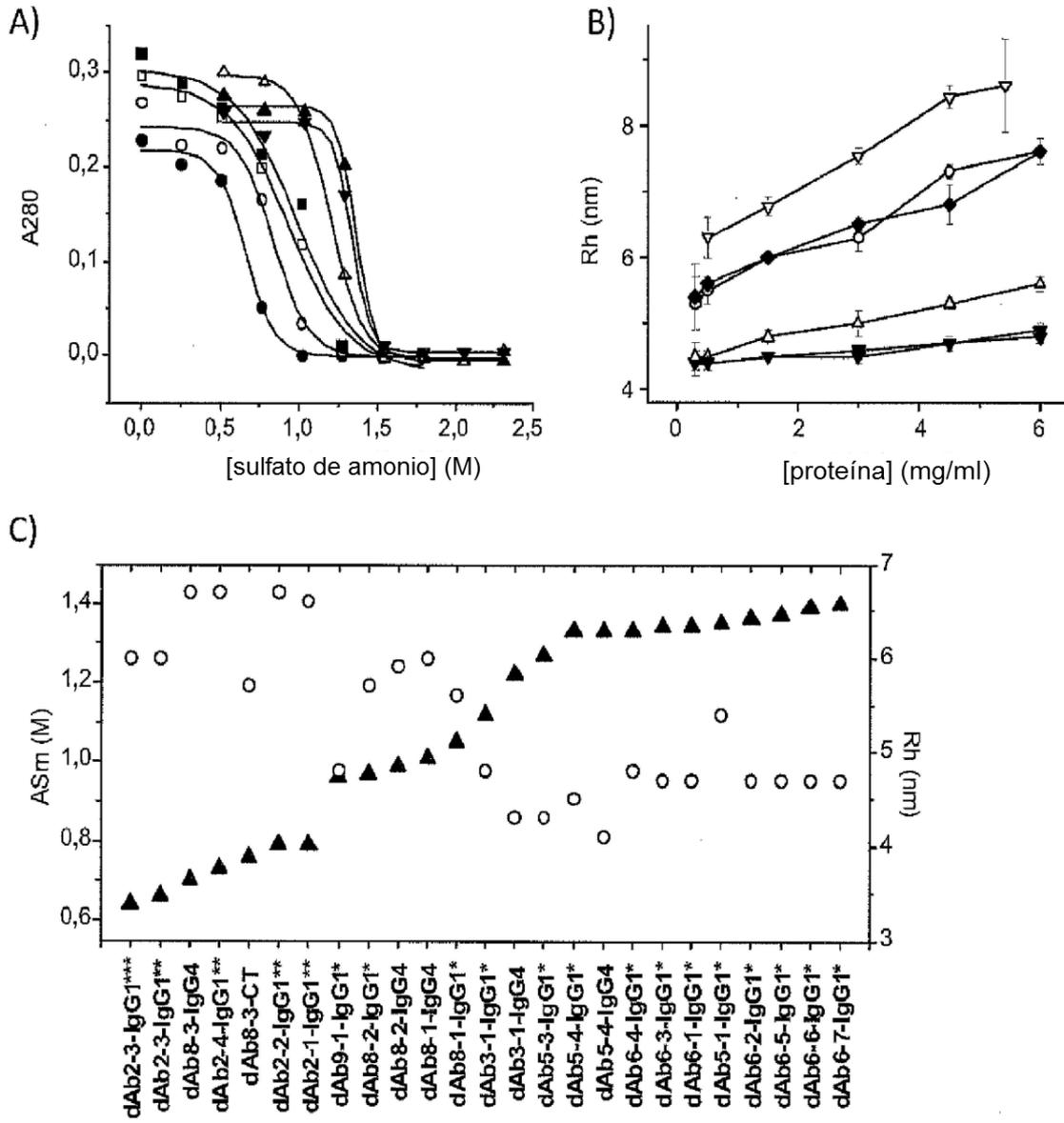


Figura 7

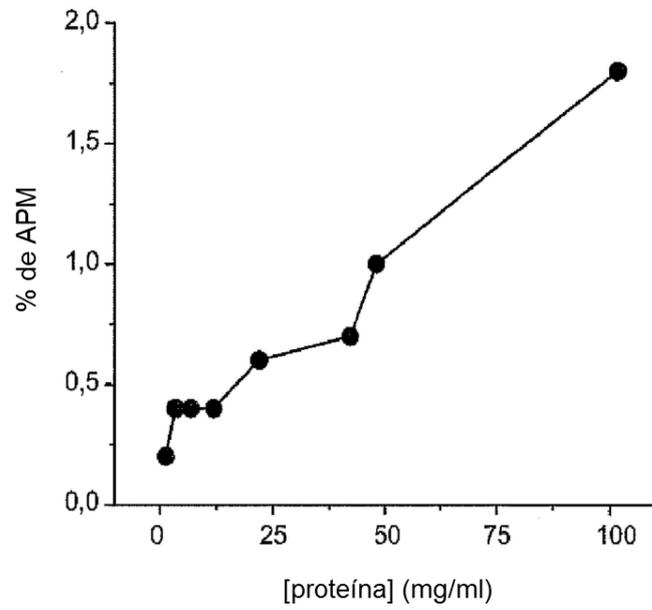


Figura 8

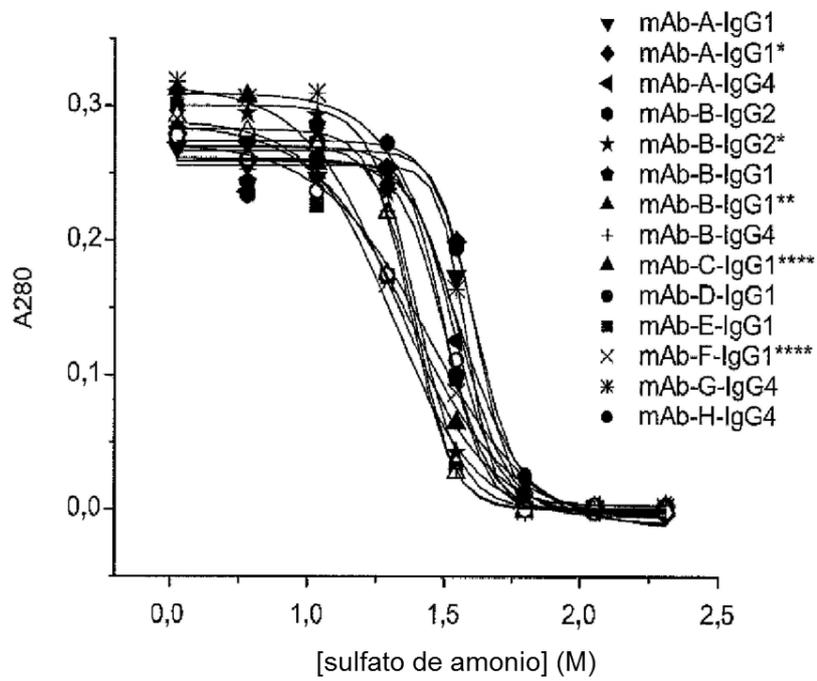


Figura 9

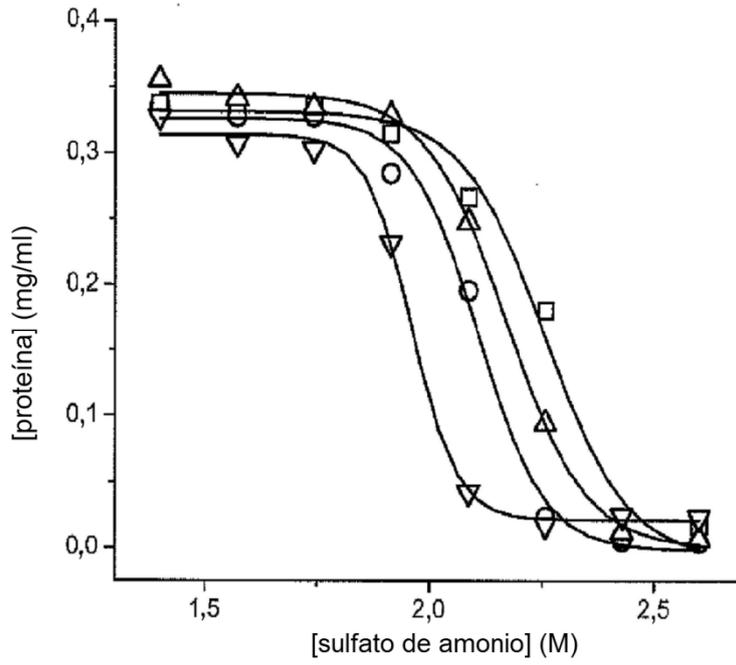
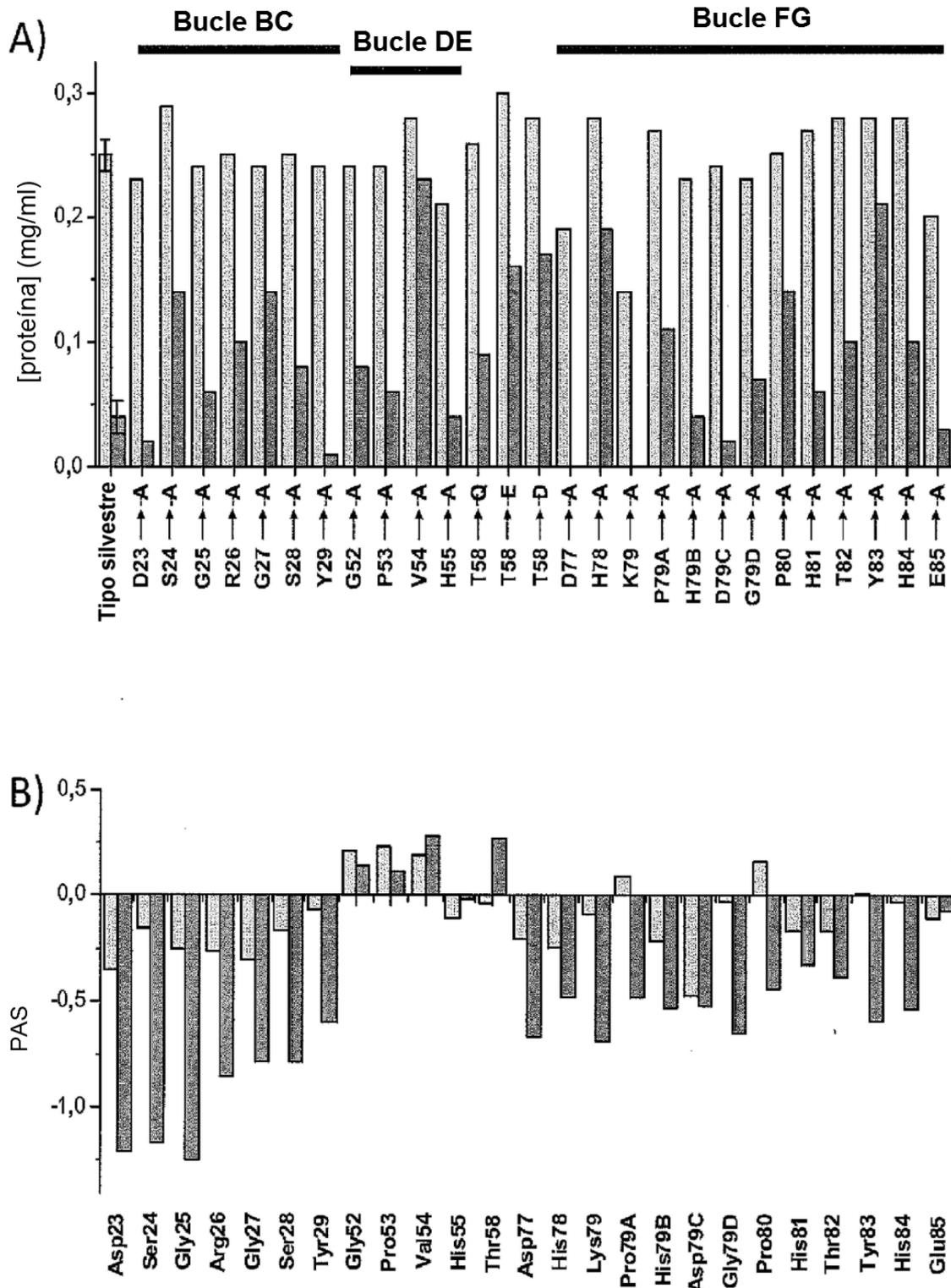


Figura 10



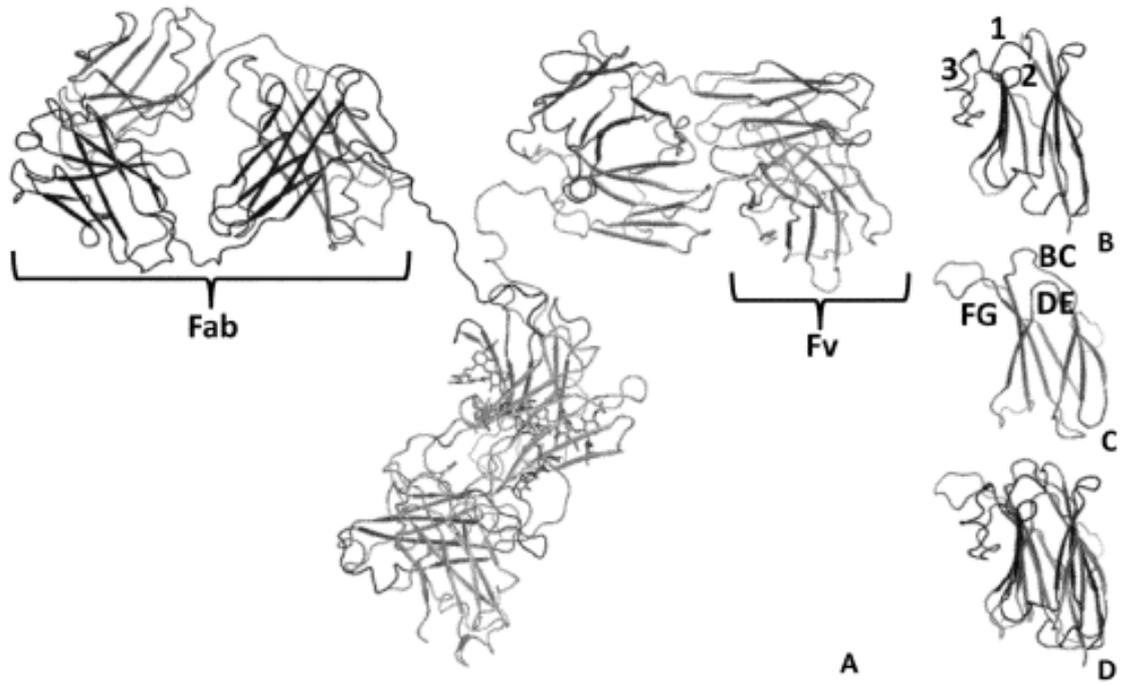


Figura 12

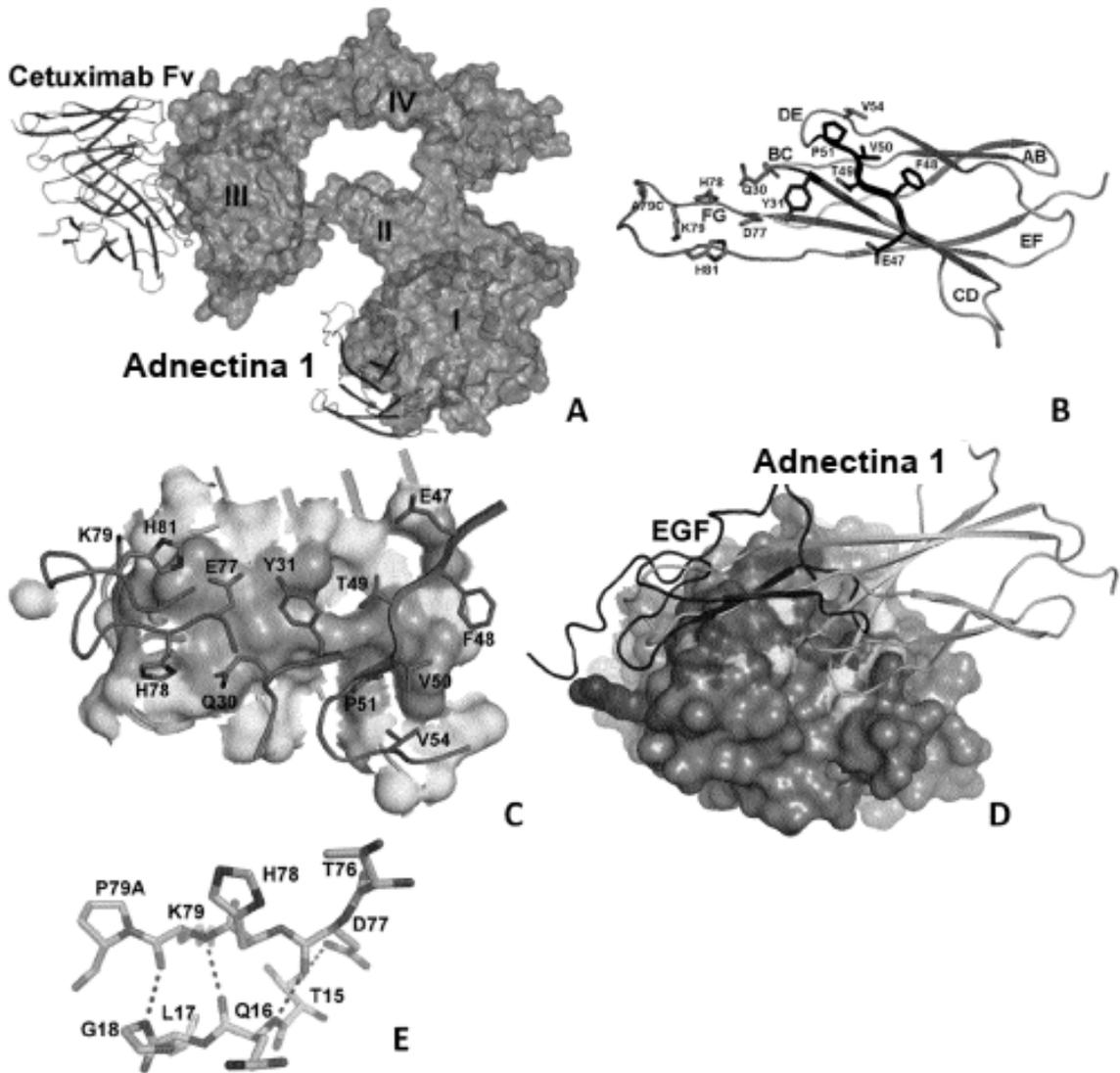


Figura 13



	-2	10	20	30	40	50
		-		-		-
<sup>10</sup> Fn3		VSDVPRDLEVVAATPTSLLISWDAPAVTVRYYRITYGETGGNSPVQEFV				
Adnectina 1		MGVSDVPRDLEVVAATPTSLLISW <u>DSGRGSY</u> QYYRITYGETGGNSPVQEFV				
Adnectina 2		MGVSDVPRDLEVVAATPTSLLISW <u>ENDYPYR</u> RYRITYGETGGNSPVQEFV				
Cadenas β		-A- -	-B-	- - -C- - -		-D- -
		60	70	80	90	100
		-		-		-
				ABCDE		
<sup>10</sup> Fn3		PGSKSTATISGLKPGVDYITITVYAVTGRG		DSPASSKPISINYRTEIDKPS		
Adnectina 1		<u>PGPVH</u> TATISGLKPGVDYITITVYAVT <u>DHKPHADGPHTYHES</u> SPISINYRTEIDKPS				
Adnectina 2		<u>PKDVD</u> TATISGLKPGVDYITITVYAVT <u>SSY</u>		<u>KYDMQYS</u> SPISINYRTEIDKPS		
Cadenas β			E-	- -F- - -		-G-
		107				
		-				
<sup>10</sup> Fn3						
Adnectina 1		QHHEHHH				
Adnectina 2		QHHEHHH				

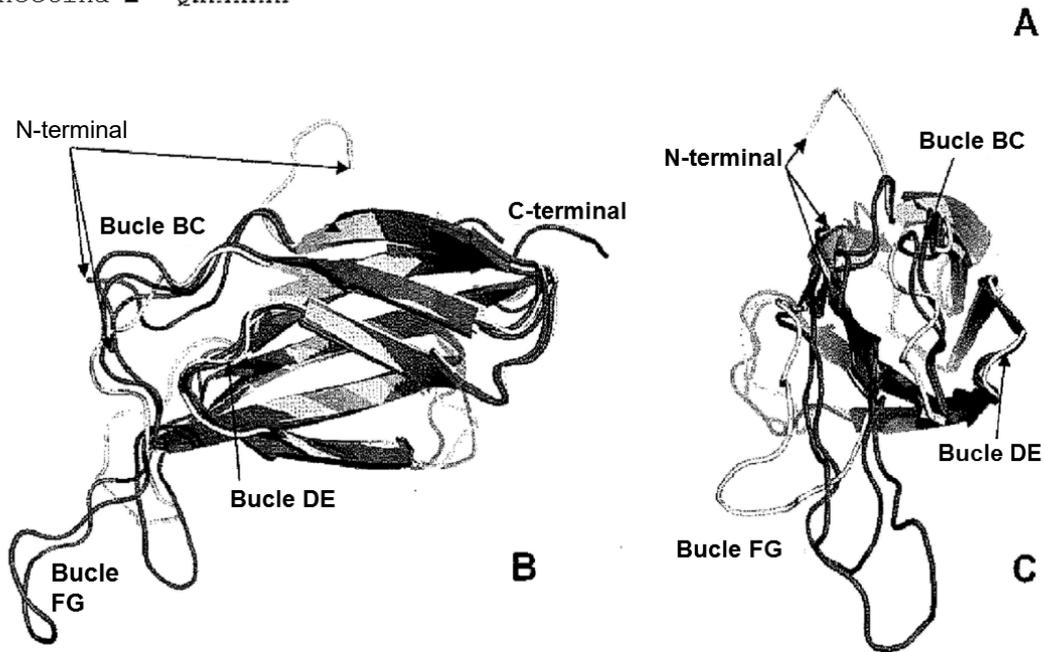
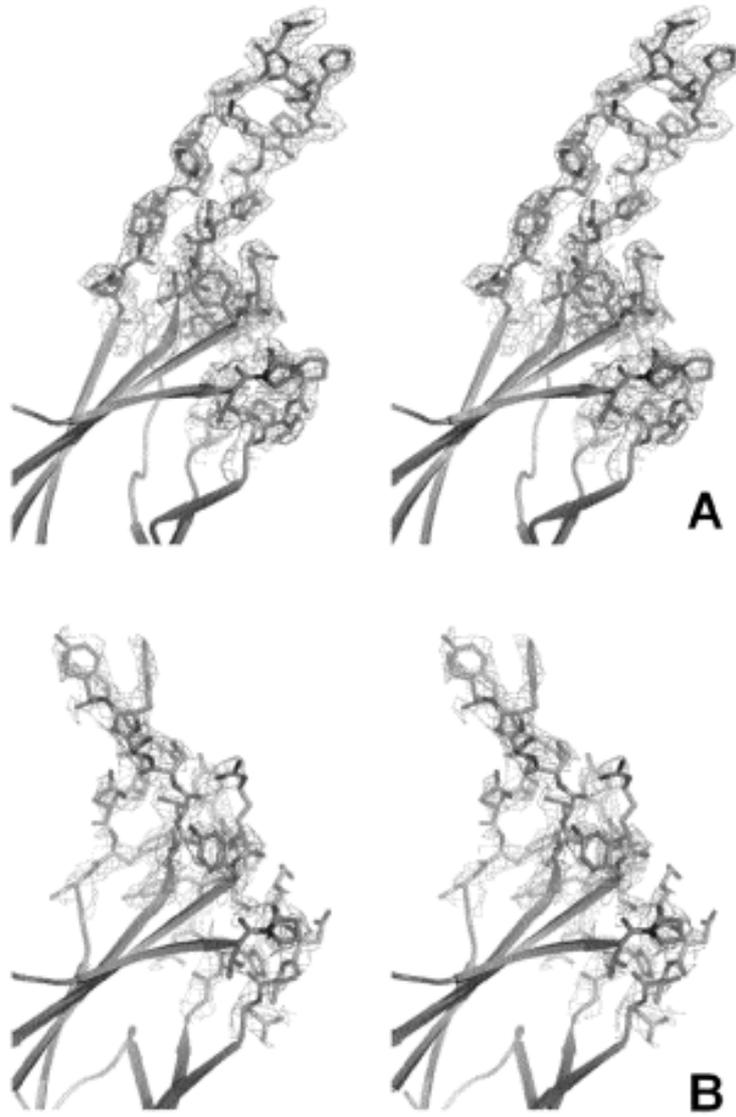
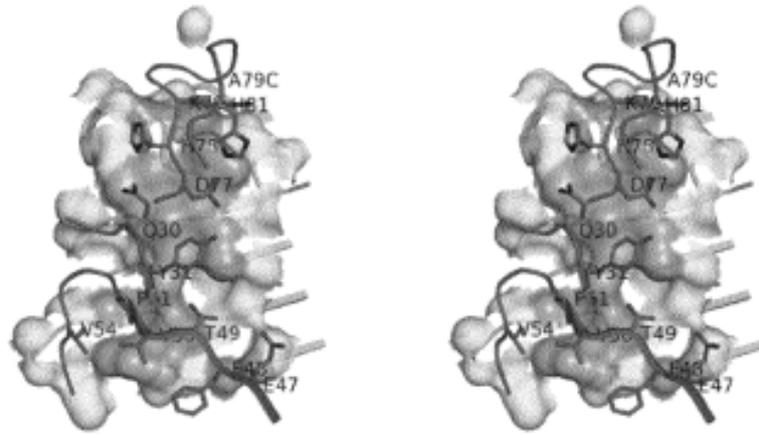


Figura 15



**Figura 16**

C.



D.

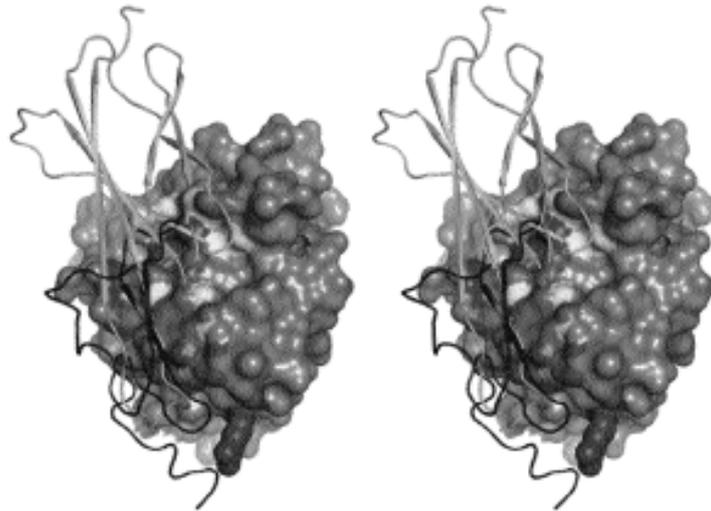
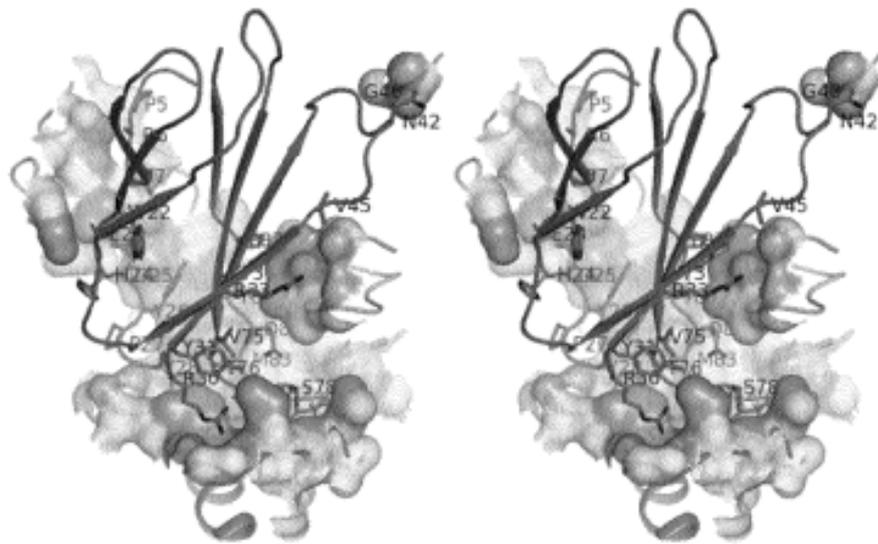


Figura 17

C.



D.

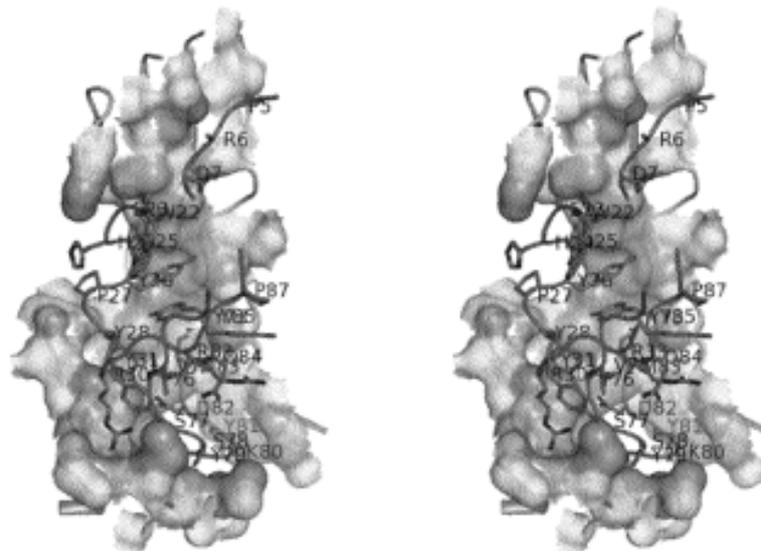
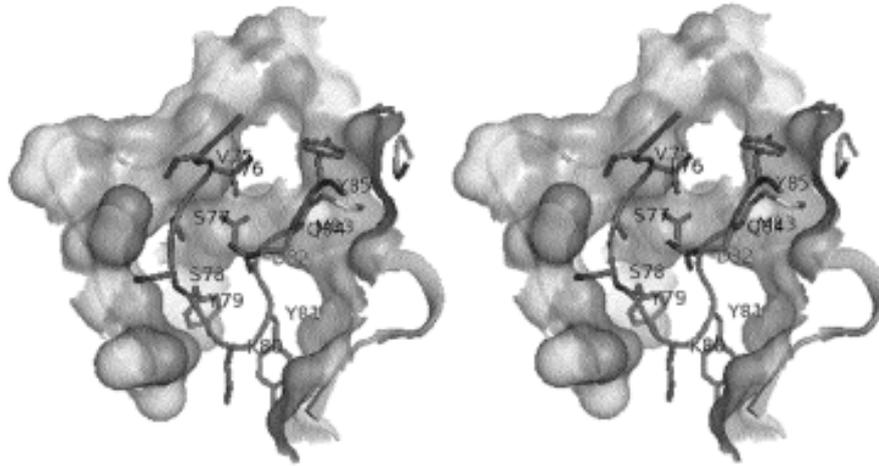


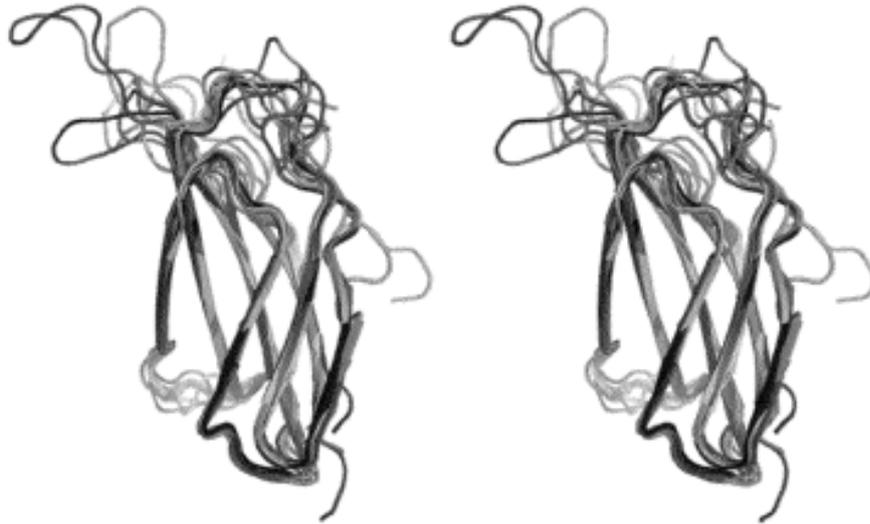
Figura 18

E.

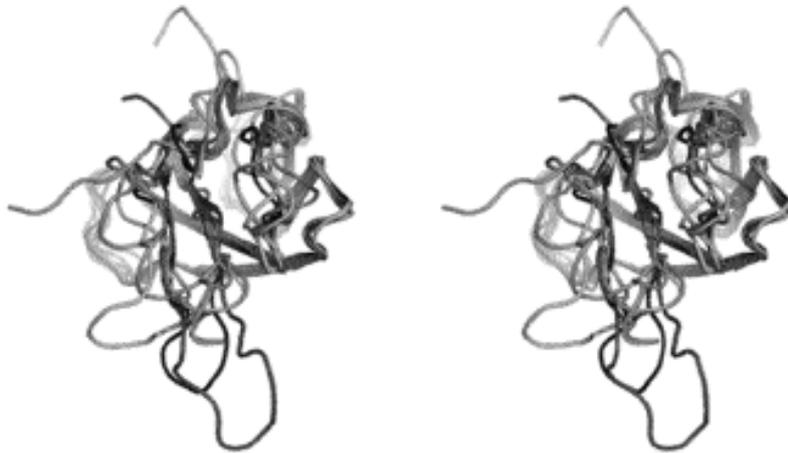


**Figura 18 (continuación)**

**B.**



**C.**



**Figura 19**

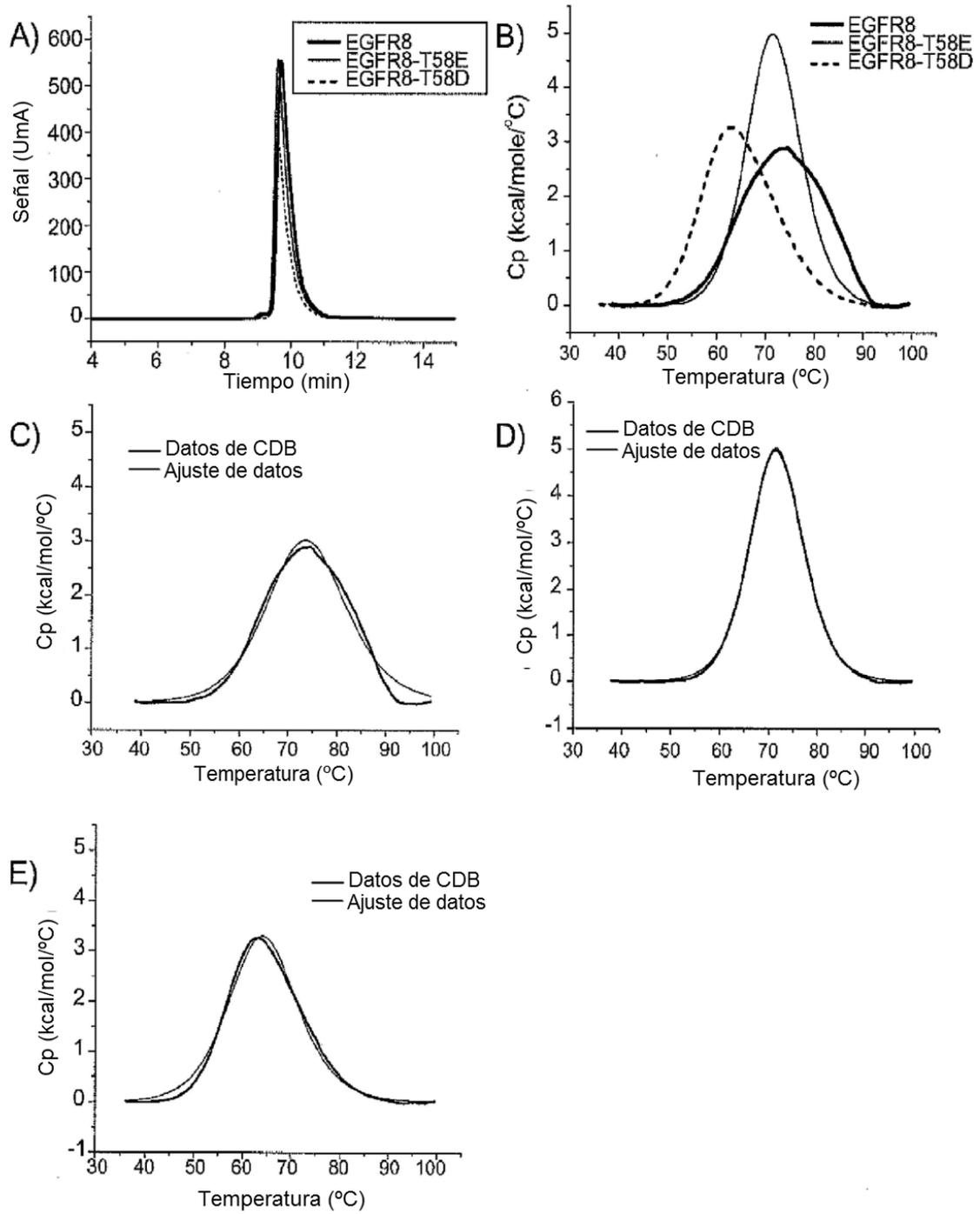


Figura 20

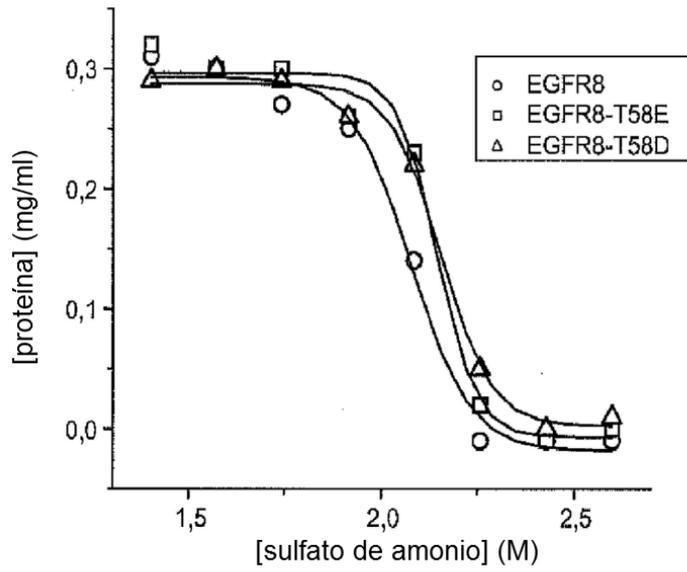


Figura 21

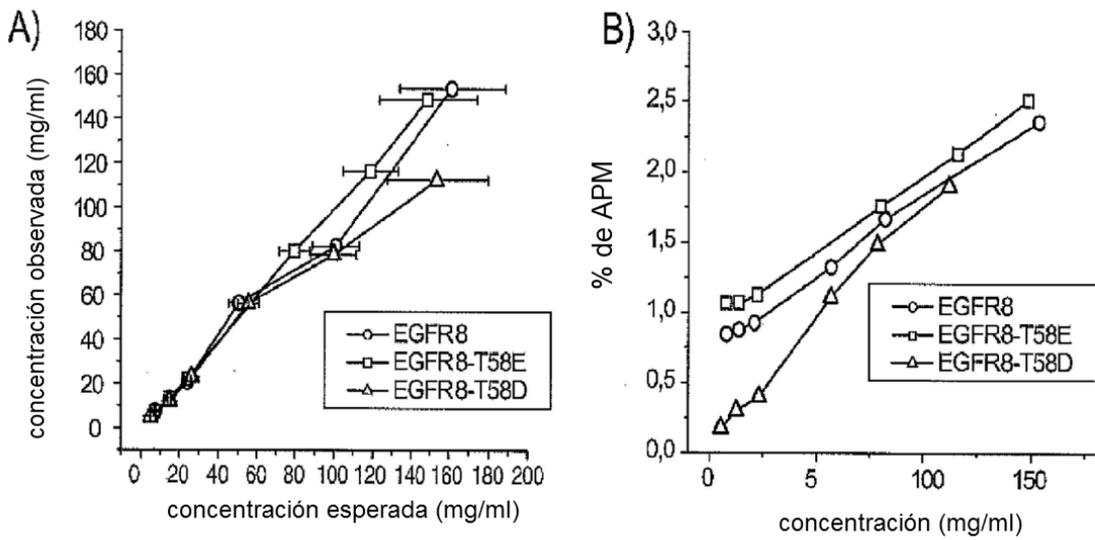


Figura 22

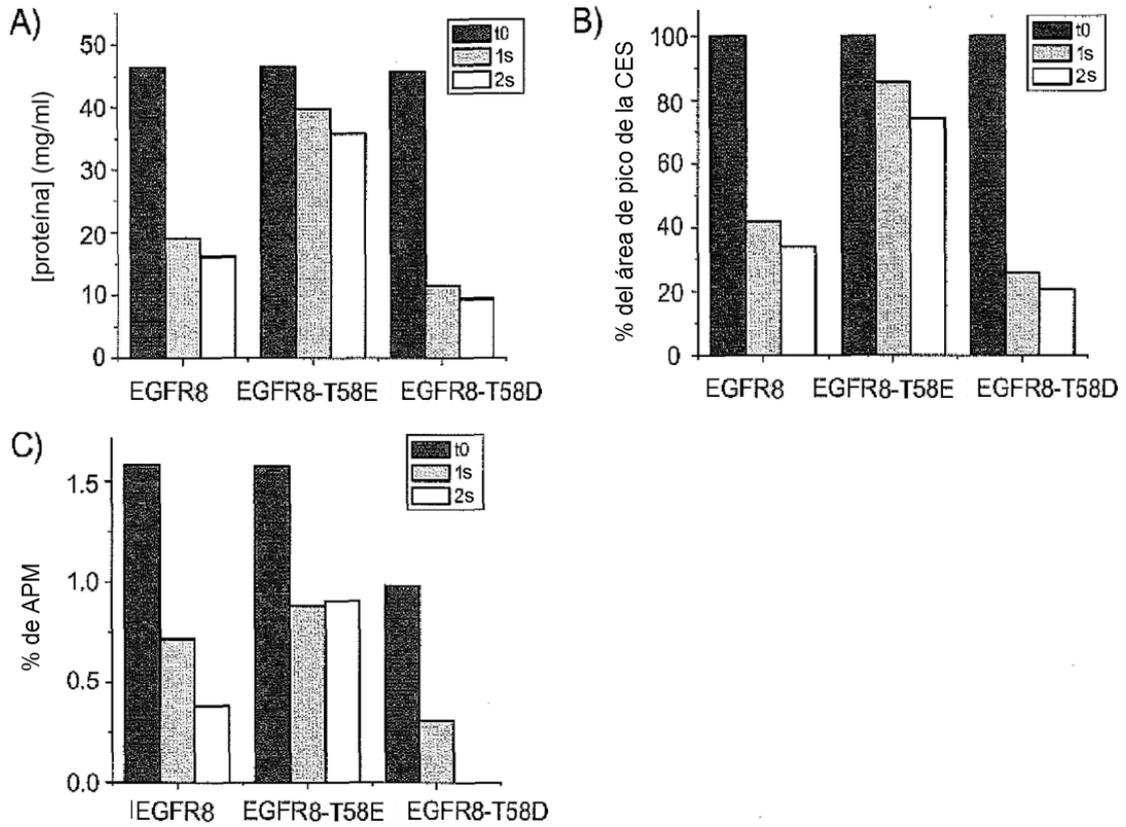


Figura 23

	1	<b>BC</b>		<b>DE</b>	60
EGFR4	(1)	VSDVPRDLEVVAATPTSLLISWYWEGLPYQYYRITYGETGGNSPVQEF		TV	RDVNTATIS
EGFR8	(1)	VSDVPRDLEVVAATPTSLLISWDSGRGSYQYYRITYGETGGNSPVQEF		TV	GPVHTATIS
Consenso	(1)	VSDVPRDLEVVAATPTSLLISW		YQYYRITYGETGGNSPVQEF	TV V TATIS
	61	<b>FG</b>			112
EGFR4	(61)	GLKPGVDYTTIVYAVTLWYNPDTHEYIYHTIPISIN		YRTEIDKPSQH	HHHHH
EGFR8	(61)	GLKPGVDYTTIVYAVTLHKPHADGPHTYHESPI		SIN	YRTEIDKPSQH
Consenso	(61)	GLKPGVDYTTIVYAVT		H YH	PI

Figura 24

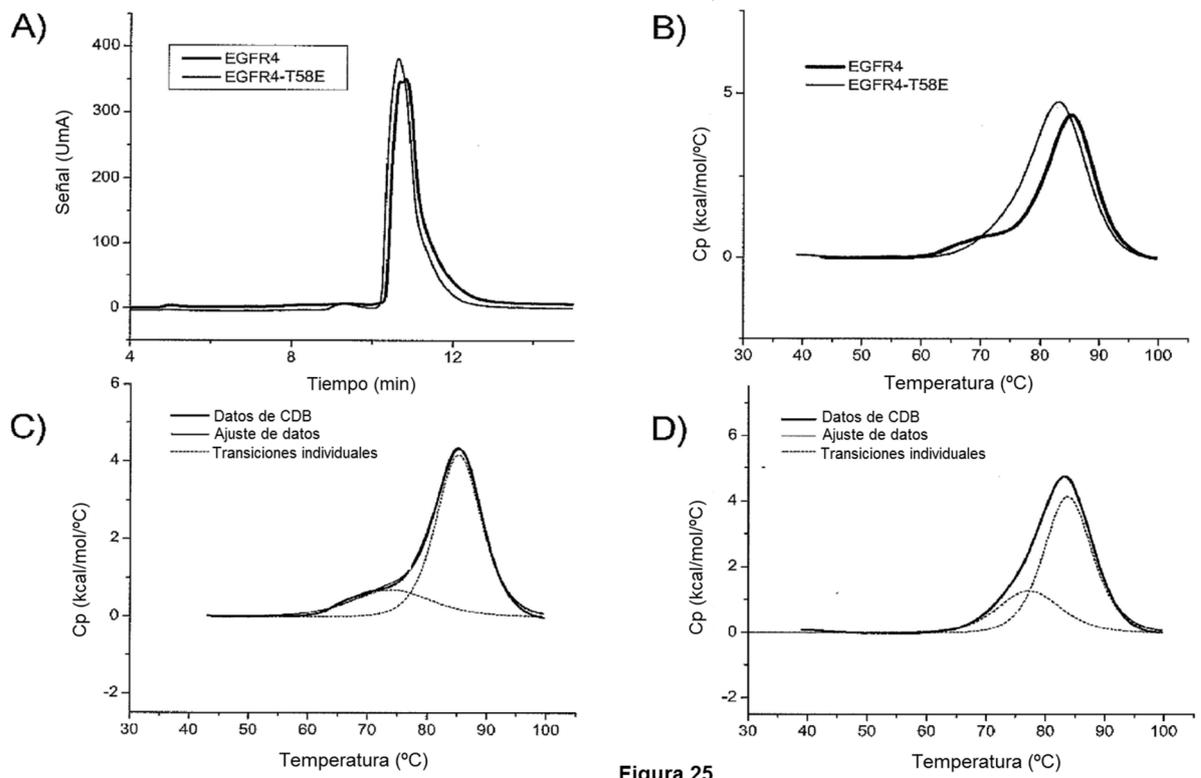


Figura 25

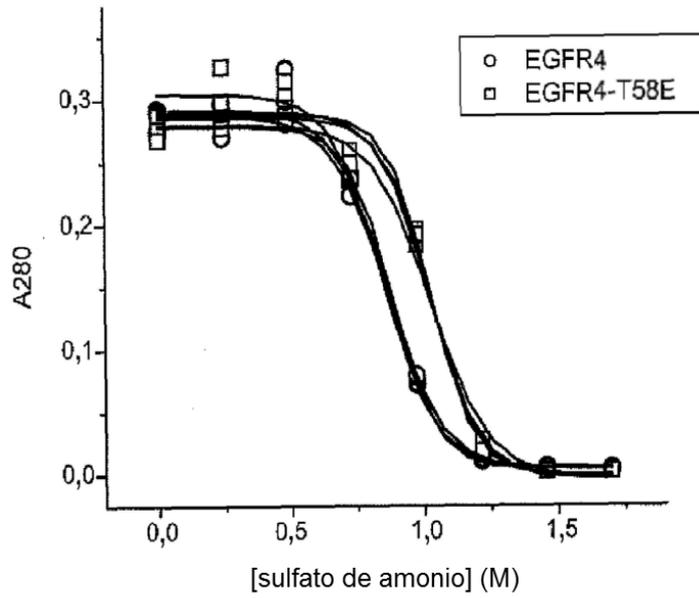


Figura 26

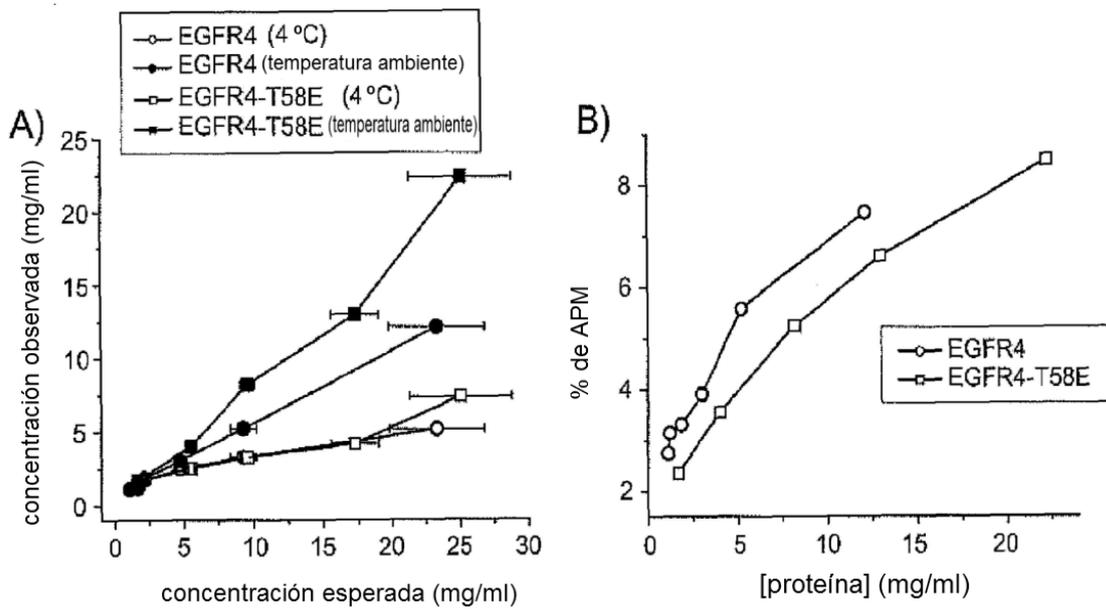


Figura 27