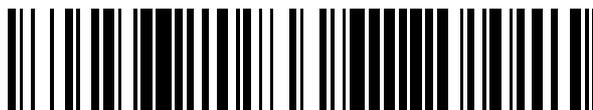


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 566 737**

51 Int. Cl.:

**C12N 15/09** (2006.01)

**A61K 39/395** (2006.01)

**C12N 15/13** (2006.01)

**C07K 16/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **25.06.2008 E 08757278 (0)**

97 Fecha y número de publicación de la concesión europea: **23.03.2016 EP 2158315**

54 Título: **Métodos de modificación de anticuerpos y anticuerpos modificados con propiedades funcionales mejoradas**

30 Prioridad:

**25.06.2007 US 937112 P**  
**12.03.2008 US 69056**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**15.04.2016**

73 Titular/es:

**ESBATECH, AN ALCON BIOMEDICAL  
RESEARCH UNIT LLC (100.0%)  
Wagistrasse 21  
8952 Schlieren, CH**

72 Inventor/es:

**URECH, DAVID y  
BORRAS, LEONARDO**

74 Agente/Representante:

**MILTENYI, Peter**

**ES 2 566 737 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

Métodos de modificación de anticuerpos y anticuerpos modificados con propiedades funcionales mejoradas

Esta solicitud reivindica prioridad sobre la solicitud provisional estadounidense con n.º de serie 60/937.112, titulada "Sequence Based Engineering and Optimization of Single Chain Antibodies", presentada el 25 de junio de 2007. Esta solicitud también reivindica prioridad sobre la solicitud provisional estadounidense con n.º de serie 61/069.056, titulada "Methods of Modifying Antibodies, and Modified Antibodies with Improved Functional Properties", presentada el 12 de marzo de 2008.

## Antecedentes de la invención

Los anticuerpos han demostrado ser agentes terapéuticos muy eficaces y satisfactorios en el tratamiento de cáncer, enfermedades autoinmunitarias y otros trastornos. Aunque normalmente se han usado clínicamente anticuerpos de longitud completa, existen varias ventajas que puede proporcionar el uso de un fragmento de anticuerpo, tal como aumento de la penetración en el tejido, ausencia de la función efectora de Fc combinada con la capacidad para añadir otras funciones efectoras y la probabilidad de menos efectos secundarios sistémicos que resultan de una semivida *in vivo* más corta de manera sistémica. Las propiedades farmacocinéticas de fragmentos de anticuerpo indican que pueden ser particularmente muy adecuados para enfoques terapéuticos locales. Además, los fragmentos de anticuerpo pueden ser más fáciles de producir que los anticuerpos de longitud completa en determinados sistemas de expresión.

Un tipo de fragmento de anticuerpo es un anticuerpo de cadena sencilla (scFv), que se compone de un dominio variable de cadena pesada ( $V_H$ ) conjugado con un dominio variable de cadena ligera ( $V_L$ ) mediante una secuencia de ligador. Por tanto, los scFv carecen de todos los dominios de región constante de anticuerpo y los residuos de aminoácido de la anterior superficie de contacto de dominio variable/constante (residuos de superficie de contacto) quedan expuestos al disolvente. Puede prepararse un scFv a partir de un anticuerpo de longitud completa (por ejemplo, molécula de IgG) a través de técnicas de modificación mediante ingeniería genética recombinante establecidas. Sin embargo, la transformación de un anticuerpo de longitud completa para dar un scFv a menudo da como resultado estabilidad y solubilidad escasas de la proteína, bajos rendimientos de producción y una alta tendencia a agregarse, lo que aumenta el riesgo de inmunogenicidad.

Por consiguiente, se han realizado intentos para mejorar propiedades tales como solubilidad y estabilidad de los scFv. Por ejemplo, Nieba, L. *et al.* (Prot. Eng. (1997) 10:435-444) seleccionaron tres residuos de aminoácido que se sabía que eran residuos de superficie de contacto y los mutaron. Observaron un aumento de la expresión periplásmica de los scFv mutados en bacterias, así como una disminución de la tasa de agregación inducida térmicamente, aunque la estabilidad termodinámica y la solubilidad no se vieron alteradas significativamente. También se han notificado otros estudios en los que se llevó a cabo mutagénesis dirigida al sitio en residuos de aminoácido particulares dentro del scFv (véanse por ejemplo, Tan, P.H. *et al.* (1988) Biophys. J. 75:1473-1482; Wörn, A. y Plückthun, A. (1998) Biochem. 37:13120-13127; Wörn, A. y Plückthun, A. (1999) Biochem. 38:8739-8750). En estos diversos estudios, se eligieron los residuos de aminoácido seleccionados para la mutagénesis basándose en sus posiciones conocidas dentro de la estructura de scFv (por ejemplo, a partir de estudios de modelado molecular).

En otro enfoque, se injertaron las regiones determinantes de complementariedad (CDR) de un scFv muy escasamente expresado en las regiones de marco de un scFv que había demostrado tener propiedades favorables (Jung, S. y Plückthun, A. (1997) Prot. Eng. 10:959-966). El scFv resultante mostró una expresión soluble y estabilidad termodinámica mejoradas.

El progreso en la modificación mediante ingeniería genética de scFv para mejorar sus propiedades funcionales se revisa, por ejemplo, en Wörn, A. y Plückthun, A. (2001) J. Mol. Biol. 305:989-1010. Sin embargo, todavía son necesarios nuevos enfoques que permitan el diseño racional de scFv con propiedades funcionales superiores, en enfoques particulares que ayuden al experto en la técnica en la selección de residuos de aminoácido potencialmente problemáticos para la modificación mediante ingeniería genética. Además, todavía son necesarios métodos de modificación mediante ingeniería genética de scFv, y otros tipos de anticuerpos, para conferir de ese modo propiedades funcionales mejoradas, tales como aumento de las propiedades de estabilidad y/o solubilidad.

## Sumario de la invención

Esta invención proporciona métodos de modificación mediante ingeniería genética de agentes de unión inmunológica, tales como anticuerpos scFv, basándose en el análisis de secuencia de regiones de marco de scFv solubles, estables que permitió la identificación de aminoácidos dentro de una secuencia de scFv que son potencialmente problemáticos para la estabilidad y/o solubilidad y la identificación de sustituciones de residuos de aminoácido preferidas en tales posiciones de aminoácido. Por tanto, pueden seleccionarse residuos de aminoácido identificados según los métodos de la invención para la mutación y pueden prepararse agentes de unión inmunológica modificados mediante ingeniería genética, tales como scFv, que se han mutado y pueden examinarse para determinar propiedades funcionales mejoradas tales como estabilidad y/o solubilidad. La invención proporciona, y demuestra el beneficio de, un enfoque "de consenso funcional" para identificar sustituciones de aminoácidos

preferidas dentro de regiones de marco de scFv basándose en el uso de una base de datos de secuencias de scFv seleccionadas funcionalmente.

5 Por consiguiente, la invención proporciona métodos de modificación mediante ingeniería genética de agentes de unión inmunológica (por ejemplo, scFv) mutando posiciones de aminoácido de regiones de marco particulares a residuos de aminoácido especificados identificados usando el enfoque de "consenso funcional" descrito en el presente documento. Todavía más, la invención proporciona armazones de región de marco de scFv, diseñados basándose en el enfoque de "consenso funcional" descrito en el presente documento, que pueden usarse como la secuencia de región de marco en la que pueden insertarse secuencias de CDR de interés para crear un agente de unión inmunológica, por ejemplo, scFv, frente a un antígeno diana de interés.

10 Preferiblemente, el agente de unión inmunológica usado en, o producido mediante, los métodos de modificación mediante ingeniería genética de la invención es un scFv, pero también pueden modificarse mediante ingeniería genética otros agentes de unión inmunológica, tales como inmunoglobulinas de longitud completa, fragmentos Fab, anticuerpos de un solo dominio (por ejemplo, los Dab) y Nanobodies según el método. La invención también abarca agentes de unión inmunológica preparados según el método de modificación mediante ingeniería genética, así como composiciones que comprenden los agentes de unión inmunológica y un portador farmacéuticamente aceptable.

15 En un aspecto, la invención proporciona un método de modificación mediante ingeniería genética de un agente de unión inmunológica, comprendiendo el agente de unión inmunológica (i) una región variable de cadena pesada, o fragmento de la misma, comprendiendo la región variable de cadena pesada residuos de región de marco de  $V_H$  y/o (ii) una región variable de cadena ligera, o fragmento de la misma, comprendiendo la región variable de cadena ligera residuos de región de marco de  $V_L$ , comprendiendo el método:

20 A) seleccionar una o más posiciones de aminoácido dentro de los residuos de región de marco de  $V_H$ , los residuos de región de marco de  $V_L$  o los residuos de región de marco de  $V_H$  y  $V_L$  para la mutación; y

25 B) mutar la una o más posiciones de aminoácido seleccionadas para la mutación, en el que la una o más posiciones de aminoácido seleccionadas para la mutación, y el/los residuo(s) de aminoácido(s) insertado(s) en la(s) posición/posiciones seleccionada(s) se describen en más detalle a continuación.

La numeración de las posiciones de aminoácido expuesta a continuación usa el sistema de numeración de AHO; las posiciones correspondientes usando el sistema de numeración de Kabat se describen además en el presente documento y las tablas de conversión para los sistemas de numeración de AHO y de Kabat se exponen en el ejemplo 1. Los residuos de aminoácido se exponen usando el código de abreviatura de una letra convencional.

30 En una realización, en la que si la una o más posiciones de aminoácido seleccionadas para la mutación son de una región variable de cadena pesada, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

- (a) Q o E en la posición de aminoácido 1;
- (b) Q o E en la posición de aminoácido 6;
- 35 (c) T, S o A en la posición de aminoácido 7, más preferiblemente T o A, incluso más preferiblemente T;
- (d) A, T, P, V o D, más preferiblemente T, P, V o D, en la posición de aminoácido 10,
- (e) L o V, más preferiblemente L, en la posición de aminoácido 12,
- (f) V, R, Q, M o K, más preferiblemente V, R, Q o M en la posición de aminoácido 13;
- 40 (g) R, M, E, Q o K, más preferiblemente R, M, E o Q, incluso más preferiblemente R o E, en la posición de aminoácido 14;
- (h) L o V, más preferiblemente L, en la posición de aminoácido 19;
- (i) R, T, K o N, más preferiblemente R, T o N, incluso más preferiblemente N, en la posición de aminoácido 20;
- (j) I, F, L o V, más preferiblemente I, F o L, incluso más preferiblemente I o L, en la posición de aminoácido 21;
- (k) R o K, más preferiblemente K, en la posición de aminoácido 45;
- 45 (l) T, P, V, A o R, más preferiblemente T, P, V o R, incluso más preferiblemente R, en la posición de aminoácido 47;
- (m) K, Q, H o E, más preferiblemente K, H o E, incluso más preferiblemente K, en la posición de aminoácido 50;
- (n) M o I, más preferiblemente I, en la posición de aminoácido 55;

## ES 2 566 737 T3

- (o) K o R, más preferiblemente K, en la posición de aminoácido 77;
- (p) A, V, L o I, más preferiblemente A, L o I, incluso más preferiblemente A, en la posición de aminoácido 78;
- (q) E, R, T o A, más preferiblemente E, T o A, incluso más preferiblemente E, en la posición de aminoácido 82;
- (r) T, S, I o L, más preferiblemente T, S o L, incluso más preferiblemente T, en la posición de aminoácido 86;
- 5 (s) D, S, N o G, más preferiblemente D, N o G, incluso más preferiblemente N, en la posición de aminoácido 87;
- (t) A, V, L o F, más preferiblemente A, V o F, incluso más preferiblemente V, en la posición de aminoácido 89;
- (u) F, S, H, D o Y, más preferiblemente F, S, H o D, en la posición de aminoácido 90;
- (v) D, Q o E, más preferiblemente D o Q, incluso más preferiblemente D, en la posición de aminoácido 92;
- 10 (w) G, N, T o S, más preferiblemente G, N o T, incluso más preferiblemente G, en la posición de aminoácido 95;
- (x) T, A, P, F o S, más preferiblemente T, A, P o F, incluso más preferiblemente F, en la posición de aminoácido 98;
- (y) R, Q, V, I, M, F o L, más preferiblemente R, Q, I, M, F o L, incluso más preferiblemente Y o L, en la posición de aminoácido 103; y
- 15 (z) N, S o A, más preferiblemente N o S, incluso más preferiblemente N, en la posición de aminoácido 107.

En otra realización, en la que si la una o más posiciones de aminoácido seleccionadas para la mutación son de una región variable de cadena ligera, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

- 20 (aa) Q, D, L, E, S o I, más preferiblemente L, E, S o I, incluso más preferiblemente L o E, en la posición de aminoácido 1;
- (bb) S, A, Y, I, P o T, más preferiblemente A, Y, I, P o T, incluso más preferiblemente P o T en la posición de aminoácido 2;
- (cc) Q, V, T o I, más preferiblemente V, T o I, incluso más preferiblemente V o T, en la posición de aminoácido 3;
- 25 (dd) V, L, I o M, más preferiblemente V o L, en la posición de aminoácido 4;
- (ee) S, E o P, más preferiblemente S o E, incluso más preferiblemente S, en la posición de aminoácido 7;
- (ff) T o I, más preferiblemente I, en la posición de aminoácido 10;
- (gg) A o V, más preferiblemente A, en la posición de aminoácido 11;
- (hh) S o Y, más preferiblemente Y, en la posición de aminoácido 12;
- 30 (ii) T, S o A, más preferiblemente T o S, incluso más preferiblemente T, en la posición de aminoácido 14;
- (jj) S o R, más preferiblemente S, en la posición de aminoácido 18;
- (kk) T o A, más preferiblemente A, en la posición de aminoácido 20;
- (ll) R o Q, más preferiblemente Q, en la posición de aminoácido 24;
- (mm) H o Q, más preferiblemente H, en la posición de aminoácido 46;
- 35 (nn) K, R o I, más preferiblemente R o I, incluso más preferiblemente R, en la posición de aminoácido 47;
- (oo) R, Q, K, E, T o M, más preferiblemente Q, K, E, T o M, en la posición de aminoácido 50;
- (pp) K, T, S, N, Q o P, más preferiblemente T, S, N, Q o P, en la posición de aminoácido 53;
- (qq) I o M, más preferiblemente M, en la posición de aminoácido 56;
- (rr) H, S, F o Y, más preferiblemente H, S o F, en la posición de aminoácido 57;
- 40 (ss) I, V o T, más preferiblemente V o T, R, incluso más preferiblemente T, en la posición de aminoácido 74;

## ES 2 566 737 T3

(tt) R, Q o K, más preferiblemente R o Q, incluso más preferiblemente R, en la posición de aminoácido 82;

(uu) L o F, más preferiblemente F, en la posición de aminoácido 91;

(vv) G, D, T o A, más preferiblemente G, D o T, incluso más preferiblemente T, en la posición de aminoácido 92;

5 (xx) S o N, más preferiblemente N, en la posición de aminoácido 94;

(yy) F, Y o S, más preferiblemente Y o S, incluso más preferiblemente S, en la posición de aminoácido 101; y

(zz) D, F, H, E, L, A, T, V, S, G o I, más preferiblemente H, E, L, A, T, V, S, G o I, incluso más preferiblemente A o V, en la posición de aminoácido 103.

10 En una realización, la región variable de cadena pesada, o fragmento de la misma, es de la familia de VH3 y, por tanto, en la que si la una o más posiciones de aminoácido seleccionadas para la mutación son de la región variable de cadena pesada de la familia de VH3, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

(i) E o Q en la posición de aminoácido 1, más preferiblemente Q;

(ii) E o Q en la posición de aminoácido 6, más preferiblemente Q;

15 (iii) T, S o A en la posición de aminoácido 7, más preferiblemente T o A, incluso más preferiblemente T;

(iv) A, V, L o F en la posición de aminoácido 89, más preferiblemente A, V o F, incluso más preferiblemente V; y

(v) R, Q, V, I, L, M o F en la posición de aminoácido 103, más preferiblemente R, Q, I, L, M o F, incluso más preferiblemente L.

20 En otra realización, la región variable de cadena pesada, o fragmento de la misma, es de la familia de VH1a y, por tanto, en la que si la una o más posiciones de aminoácido seleccionadas para la mutación son de la región variable de cadena pesada de la familia de VH1a, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

(i) E o Q en la posición de aminoácido 1, más preferiblemente E;

(ii) E o Q en la posición de aminoácido 6, más preferiblemente E;

25 (iii) L o V en la posición de aminoácido 12, más preferiblemente L;

(iv) M o K en la posición de aminoácido 13, más preferiblemente M;

(v) E, Q o K en la posición de aminoácido 14, más preferiblemente E o Q, incluso más preferiblemente E;

(vi) L o V en la posición de aminoácido 19, más preferiblemente L;

(vii) I o V en la posición de aminoácido 21, más preferiblemente I;

30 (viii) F, S, H, D o Y en la posición de aminoácido 90, más preferiblemente F, S, H o D;

(ix) D, Q o E en la posición de aminoácido 92, más preferiblemente D o Q, incluso más preferiblemente D;

(x) G, N, T o S en la posición de aminoácido 95, más preferiblemente G, N o T, incluso más preferiblemente G; y

35 (xi) T, A, P, F o S en la posición de aminoácido 98, más preferiblemente T, A, P o F, incluso más preferiblemente F.

En otra realización, la región variable de cadena pesada, o fragmento de la misma, es de la familia de VH1b y, por tanto, en la que si la una o más posiciones de aminoácido seleccionadas para la mutación son de la región variable de cadena pesada de la familia de VH1b, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

40 (i) E o Q en la posición de aminoácido 1, más preferiblemente E;

(ii) A, T, P, V o D en la posición de aminoácido 10, más preferiblemente T, P, V o D;

(iii) L o V en la posición de aminoácido 12, más preferiblemente L;

(iv) K, V, R, Q o M en la posición de aminoácido 13, más preferiblemente V, R, Q o M;

- (v) E, K, R o M en la posición de aminoácido 14, más preferiblemente E, R o M, incluso más preferiblemente R;
- (vi) R, T, K o N en la posición de aminoácido 20, más preferiblemente R, T o N, incluso más preferiblemente N;
- (vii) I, F, V o L en la posición de aminoácido 21, más preferiblemente I, F o L, incluso más preferiblemente L;
- (viii) R o K en la posición de aminoácido 45, más preferiblemente K;
- 5 (ix) T, P, V, A, R en la posición de aminoácido 47, más preferiblemente T, P, V o R, incluso más preferiblemente R;
- (x) K, Q, H o E en la posición de aminoácido 50, más preferiblemente K, H o E, incluso más preferiblemente K;
- (xi) M o I en la posición de aminoácido 55; más preferiblemente I;
- (xii) K o R en la posición de aminoácido 77, más preferiblemente K;
- 10 (xiii) A, V, L o I en la posición de aminoácido 78, más preferiblemente A, L o I, incluso más preferiblemente A;
- (xiv) E, R, T o A en la posición de aminoácido 82, más preferiblemente E, T o A, incluso más preferiblemente E;
- (xv) T, S, I o L en la posición de aminoácido 86, más preferiblemente T, S o L, incluso más preferiblemente T;
- (xvi) D, S, N o G en la posición de aminoácido 87, más preferiblemente D, N o G, incluso más preferiblemente N; y
- 15 (xvii) N, S o A en la posición de aminoácido 107, más preferiblemente N o S, incluso más preferiblemente N.

En otra realización, la región variable de cadena ligera, o fragmento de la misma, es de la familia de V $\kappa$ 1 y, por tanto, en la que si la una o más posiciones de aminoácido seleccionadas para la mutación son de la región variable de cadena ligera de la familia de V $\kappa$ 1, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

- 20 (i) D, E o I en la posición de aminoácido 1, más preferiblemente E o I, incluso más preferiblemente E;
- (ii) Q, V o I en la posición de aminoácido 3, más preferiblemente V o I, incluso más preferiblemente V;
- (iii) V, L, I o M en la posición de aminoácido 4, más preferiblemente V, L o I, incluso más preferiblemente L;
- (iv) R o Q en la posición de aminoácido 24, más preferiblemente Q;
- (v) K, R o I en la posición de aminoácido 47, más preferiblemente R o I, incluso más preferiblemente R;
- 25 (vi) K, R, E, T, M o Q en la posición de aminoácido 50, más preferiblemente K, E, T, M o Q;
- (vii) H, S, F o Y en la posición de aminoácido 57, más preferiblemente H, S o F, incluso más preferiblemente S;
- (viii) L o F en la posición de aminoácido 91, más preferiblemente F; y
- (ix) T, V, S, G o I, más preferiblemente V, S, G o I, incluso más preferiblemente V, en la posición de aminoácido 103.

30 En otra realización, la región variable de cadena ligera, o fragmento de la misma, es de la familia de V $\kappa$ 3 y, por tanto, en la que si la una o más posiciones de aminoácido seleccionadas para la mutación son de la región variable de cadena ligera de la familia de V $\kappa$ 3, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

- (i) I o T en la posición de aminoácido 2, más preferiblemente T;
- 35 (ii) V o T en la posición de aminoácido 3, más preferiblemente T;
- (iii) T o I en la posición de aminoácido 10, más preferiblemente I;
- (iv) S o Y en la posición de aminoácido 12, más preferiblemente Y;
- (v) S o R en la posición de aminoácido 18, más preferiblemente S;
- (vi) T o A en la posición de aminoácido 20, más preferiblemente A;
- 40 (vii) I o M en la posición de aminoácido 56, más preferiblemente M;

- (viii) I, V o T en la posición de aminoácido 74, más preferiblemente V o T, incluso más preferiblemente T;
- (ix) S o N en la posición de aminoácido 94, más preferiblemente N;
- (x) F, Y o S en la posición de aminoácido 101, más preferiblemente Y o S, incluso más preferiblemente S; y
- (xi) V, L o A en la posición de aminoácido 103, más preferiblemente L o A, incluso más preferiblemente A.

5 En otra realización, la región variable de cadena ligera, o fragmento de la misma, es de la familia de V $\lambda$ 1 y, por tanto, en la que si la una o más posiciones de aminoácido seleccionadas para la mutación son de la región variable de cadena ligera de la familia de V $\lambda$ 1, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

- (i) L, Q, S o E en la posición de aminoácido 1, más preferiblemente L, S o E, incluso más preferiblemente L;
- 10 (ii) S, A, P, I o Y en la posición de aminoácido 2, más preferiblemente A, P, I o Y, incluso más preferiblemente P;
- (iii) V, M o L en la posición de aminoácido 4, más preferiblemente V o M, incluso más preferiblemente V;
- (iv) S, E o P en la posición de aminoácido 7, más preferiblemente S o E, incluso más preferiblemente S;
- (v) A o V en la posición de aminoácido 11, más preferiblemente A;
- 15 (vi) T, S o A en la posición de aminoácido 14, más preferiblemente T o S, incluso más preferiblemente T;
- (vii) H o Q en la posición de aminoácido 46, más preferiblemente H;
- (viii) K, T, S, N, Q o P en la posición de aminoácido 53, más preferiblemente T, S, N, Q o P;
- (ix) R, Q o K en la posición de aminoácido 82, más preferiblemente R o Q, incluso más preferiblemente R;
- (x) G, T, D o A en la posición de aminoácido 92, más preferiblemente G, T o D, incluso más preferiblemente T; y
- 20 (xi) D V T, H o E en la posición de aminoácido 103, más preferiblemente V, T, H o E, incluso más preferiblemente V.

En otra realización, la mutación comprende además una o más (preferiblemente todas) sustituciones de cadena pesada seleccionadas del grupo que consiste en:

- 25 (i) serina (S) en la posición de aminoácido 12 usando AHo o Kabat;
- (ii) serina (S) en la posición de aminoácido 103 usando la numeración de AHo (posición de aminoácido 85 usando la numeración de Kabat); y
- (iii) serina (S) o treonina (T) en la posición de aminoácido 144 usando la numeración de AHo (posición de aminoácido 103 usando la numeración de Kabat).

30 En otro aspecto, la invención proporciona armazones de región de marco de anticuerpo aislados (por ejemplo, armazones de scFv). Por ejemplo, en diversas realizaciones, la invención proporciona un armazón de región de marco de cadena pesada aislado que comprende una secuencia de aminoácidos tal como se muestra en la figura 9 (SEQ ID NO:1), la figura 10 (SEQ ID NO:2) o la figura 11 (SEQ ID NO:3). En otra realización a modo de ejemplo, la invención proporciona un armazón de región de marco de cadena ligera aislado que comprende una secuencia de aminoácidos tal como se muestra en la figura 12 (SEQ ID NO:4), la figura 13 (SEQ ID NO:5) o la figura 14 (SEQ ID NO:6). Tales armazones pueden usarse para modificar mediante ingeniería genética agentes de unión inmunológica, tales como anticuerpos scFv. Por consiguiente, en otro aspecto, la invención proporciona un método de modificación mediante ingeniería genética de un agente de unión inmunológica, comprendiendo el agente de unión inmunológica secuencias de CDR1, CDR2 y CDR3 de cadena pesada y/o ligera, comprendiendo el método insertar las secuencias de CDR1, CDR2 y CDR3 de cadena pesada y/o ligera, respectivamente, en un armazón de región de marco de cadena pesada. En determinadas realizaciones a modo de ejemplo, el armazón de región de marco de cadena pesada comprende una secuencia de aminoácidos tal como se muestra en la figura 9 (SEQ ID NO:1), la figura 10 (SEQ ID NO:2), la figura 11 (SEQ ID NO:3), SEQ ID NO:7, SEQ ID NO:8 o SEQ ID NO:9. En una realización preferida, el armazón de región de marco de cadena pesada comprende una secuencia de aminoácidos tal como se muestra en la figura 9 (SEQ ID NO:1). En otra realización preferida, el armazón de región de marco de cadena pesada comprende una secuencia de aminoácidos tal como se muestra en la figura 10 (SEQ ID NO:2). En otra realización preferida, el armazón de región de marco de cadena pesada comprende una secuencia de aminoácidos tal como se muestra en la figura 11 (SEQ ID NO:3). En otra realización preferida, el armazón de región de marco de cadena pesada comprende una secuencia de aminoácidos de SEQ ID NO:7. En otra realización preferida, el

armazón de región de marco de cadena pesada comprende una secuencia de aminoácidos de SEQ ID NO:8. En  
 aún otra realización preferida, el armazón de región de marco de cadena pesada comprende una secuencia de  
 aminoácidos de SEQ ID NO:9. En otras realizaciones a modo de ejemplo, el armazón de región de marco de cadena  
 ligera comprende una secuencia de aminoácidos tal como se muestra en la figura 12 (SEQ ID NO:4), la figura 13  
 5 (SEQ ID NO:5), la figura 14 (SEQ ID NO:6), SEQ ID NO:10, SEQ ID NO:11 o SEQ ID NO:12. En una realización  
 preferida, el armazón de región de marco de cadena ligera comprende una secuencia de aminoácidos tal como se  
 muestra en la figura 11 (SEQ ID NO:4). En otra realización preferida, el armazón de región de marco de cadena  
 ligera comprende una secuencia de aminoácidos tal como se muestra en la figura 12 (SEQ ID NO:5). En otra  
 10 realización preferida, el armazón de región de marco de cadena ligera comprende una secuencia de aminoácidos tal  
 como se muestra en la figura 13 (SEQ ID NO:6). En otra realización preferida, el armazón de región de marco de  
 cadena ligera comprende una secuencia de aminoácidos tal como se muestra en SEQ ID NO: 10. En otra realización  
 preferida, el armazón de región de marco de cadena ligera comprende una secuencia de aminoácidos tal como se  
 muestra en SEQ ID NO:11. En aún otra realización preferida, el armazón de región de marco de cadena ligera  
 15 comprende una secuencia de aminoácidos tal como se muestra en SEQ ID NO: 12. Preferiblemente, el agente de  
 unión inmunológica es un anticuerpo scFv, aunque otros agentes de unión inmunológica tal como se describen en el  
 presente documento (por ejemplo, anticuerpos de longitud completa, Fab, Dab o Nanobodies) pueden modificarse  
 mediante ingeniería genética según los métodos de la invención. La invención también proporciona composiciones  
 de agente de unión inmunológica, tales como anticuerpos scFv, modificados mediante ingeniería genética según los  
 métodos de la invención.

20 **Breve descripción de las figuras**

La figura 1 es un esquema de diagrama de flujo que resume análisis basados en secuencia generales de scFv  
 según los métodos de la invención.

En una primera etapa, se proporciona una secuencia de un scFv que va a mejorarse en cuanto a solubilidad y  
 25 estabilidad (recuadro 1), que se compara posteriormente con bases de datos de secuencias de anticuerpos  
 (recuadro 2), tales como bases de datos de secuencias de líneas germinales de código abierto (por ejemplo, Vbase,  
 IMGt; recuadro 3) bases de datos de secuencias de anticuerpos maduros de código abierto (por ejemplo, KDB;  
 recuadro 4) o bases de datos de fragmentos de scFv solubles y estables totalmente humanos (por ejemplo, QC;  
 recuadro 5).

Aplicar una base de datos de secuencias de líneas germinales de código abierto tal como la descrita en el recuadro  
 30 3 permite la identificación de posiciones altamente conservadas que se seleccionaron durante la evolución y se cree,  
 por tanto, que contribuyen a la estabilidad de dominios variables en el contexto de anticuerpos de longitud completa  
 (recuadro 3'). Una comparación frente a una base de datos de secuencias de anticuerpos maduros de código abierto  
 4 permite la identificación de patrones que representan mejoras de estabilidad, solubilidad y/o unión que son  
 35 independientes de las CDR respectivas (recuadro 4'). Además, la comparación frente a una base de datos de  
 fragmentos de scFv solubles y estables totalmente humanos (recuadro 5) conduce a la identificación de residuos  
 críticos para la estabilidad y/o solubilidad, específicamente en el formato de scFv, así como a la identificación de  
 patrones que representan mejoras de estabilidad, solubilidad y/o unión independientes de las CDR respectivas,  
 específicamente en el formato de scFv, por ejemplo combinaciones de VL y VH (recuadro 5').

En una siguiente etapa (recuadro 6), se realiza una sustitución de residuos críticos por el aminoácido adecuado más  
 40 frecuente identificado en la base de datos respectiva.

Finalmente (recuadro 7), puede realizarse mutagénesis al azar o sesgada de residuos críticos y el examen posterior  
 para determinar una estabilidad y/o solubilidad mejoradas en el sistema QC en levadura. Los mutantes pueden  
 someterse de nuevo al procedimiento mencionado anteriormente (flecha al recuadro 2).

La figura 2 es un esquema de diagrama de flujo de un método de múltiples etapas a modo de ejemplo para el  
 45 análisis basado en secuencia de scFv.

En una primera etapa (recuadro 1), se determina la frecuencia de cada residuo en la región de marco comparando la  
 aparición de aminoácidos diferentes en cada posición, basándose en resultados proporcionados por herramientas  
 bioinformáticas. En una segunda etapa, se define un grado de conservación en cada posición, por ejemplo usando el  
 índice de Simpson con la fórmula  $D = \sum (n_i - 1) / (N(N - 1))$ . En una tercera etapa, se determina la mejor sustitución que  
 50 minimiza la energía libre total (por ejemplo, aplicando la ley de Boltzmann:  $\Delta \Delta G_{th} = -RT \ln(f_{parental} / f_{consenso})$ .  
 Finalmente (etapa 4), se determina el papel de las posibles mutaciones de estabilización. Para este fin, pueden  
 tenerse en cuenta factores tales como interacciones locales y no locales, residuos canónicos, superficies de  
 contacto, grado de exposición y propensión a giros  $\beta$ .

La figura 3 es un diagrama esquemático de un sistema de control de calidad (QC, *Quality Control*) a modo de  
 55 ejemplo para la selección de scFv estables y solubles en levadura. Con este sistema, se seleccionan células  
 huésped que pueden expresar scFv estables y solubles en un entorno reductor debido a la presencia de un  
 constructo indicador inducible cuya expresión depende de la presencia de una proteína de fusión scFv-AD-Gal11p  
 estable y soluble. La interacción de la proteína de fusión con Gal4 (1-100) forma un factor de transcripción funcional

que activa la expresión de un marcador seleccionable (véase la figura 3A). scFv inestables y/o insolubles no pueden formar un factor de transcripción funcional e inducir la expresión del marcador seleccionable y, por tanto, se excluyen de la selección (figura 3B). Los scFv seleccionados pueden obtener un plegamiento de proteína estable y soluble, incluso en condiciones reductoras, en las que no se pliegan los enlaces disulfuro, mientras que los scFv inestables y/o insolubles tienden a desplegarse, agregarse y/o degradarse. En condiciones oxidantes, los scFv seleccionados todavía revelan características de solubilidad y estabilidad superiores.

La figura 4 es un diagrama esquemático de otro sistema de control de calidad (QC) a modo de ejemplo. El concepto global para seleccionar scFv soluble y estable es el mismo descrito para la figura 3, sin embargo en esta versión, el scFv se fusiona directamente a un factor de transcripción funcional que comprende un dominio de activación (AD) y un dominio de unión a ADN (DBD). La figura 4A representa un scFv soluble y estable a modo de ejemplo que, cuando se fusiona a un factor de transcripción funcional, no dificulta la transcripción de un marcador seleccionable. En cambio, la figura 4B representa el escenario mediante el cual un scFv inestable se fusiona al factor de transcripción dando lugar a un constructo de fusión no funcional que no puede activar la transcripción del marcador seleccionable.

La figura 5 es un diagrama esquemático del análisis de variabilidad en posiciones de región de marco (FW) particulares dentro de secuencias de líneas germinales nativas antes de la mutación somática (figura 5A) y en las posiciones de FW correspondientes dentro de secuencias de anticuerpos maduros tras la mutación somática seleccionadas en el sistema QC (figura 5B). Pueden asignarse diferentes valores de variabilidad a las posiciones de FW respectivas (por ejemplo, residuos de región de marco altamente variables ("hvFR")) dentro de las secuencias de líneas germinales y QC (es decir, valores de "G" y "Q", respectivamente). Si  $G > Q$  para una posición particular, existe un número restringido de residuos de FW estables adecuados en esa posición. Si  $G < Q$  para una posición particular, esto puede indicar que el residuo se ha seleccionado de manera natural por su solubilidad y estabilidad óptimas.

La figura 6 representa el perfil de desnaturalización observado para variantes de ESBA105 tras estrés inducido térmicamente en un intervalo de temperaturas de desde 25 hasta 95°C. Las variantes de ESBA-105 que tienen retromutaciones a residuos consenso de línea germinal (V3Q, R47K o V103T) se indican mediante líneas discontinuas. Las variantes que comprenden sustituciones preferidas identificadas mediante los métodos de la invención (QC11.2, QC15.2 y QC23.2) se indican mediante líneas continuas.

La figura 7 representa una comparación de la estabilidad térmica para un conjunto de variantes de ESBA105 que comprenden o bien retromutaciones consenso (S-2, D-2, D-3), una mutación a alanina (D-1) o bien un residuo QC (QC7.1, QC11.2, QC15.2, QC23.2). Se proporciona estabilidad térmica de cada variante (en unidades arbitrarias de desplegamiento).

La figura 8 representa el perfil de desnaturalización observado para variantes de ESBA212 tras estrés inducido térmicamente en un intervalo de temperaturas de desde 25 hasta 95°C. Las variantes de ESBA-212 que tienen retromutaciones a residuos consenso de línea germinal (V3Q o R47K) se indican mediante líneas discontinuas. La molécula de ESBA212 original se indica mediante una línea continua.

La figura 9 ilustra el armazón de región de marco de scFv para la familia de VH1a. La primera fila muestra la numeración de región variable de cadena pesada usando el sistema de Kabat. La segunda fila muestra la numeración de región variable de cadena pesada usando el sistema de AHo. La tercera fila muestra la secuencia del armazón de región de marco de scFv (SEQ ID NO:1), en la que en aquellas posiciones marcadas como "X", la posición puede estar ocupada por cualquiera de los residuos de aminoácido enumerados debajo de la "X". Las posiciones marcadas como "x" y las regiones marcadas como CDR1 H1, CDR H2 y CDR H3 pueden estar ocupadas por cualquier aminoácido.

La figura 10 ilustra el armazón de región de marco de scFv para la familia de VH1b. La primera fila muestra la numeración de región variable de cadena pesada usando el sistema de Kabat. La segunda fila muestra la numeración de región variable de cadena pesada usando el sistema de AHo. La tercera fila muestra la secuencia del armazón de región de marco de scFv (SEQ ID NO:2), en la que en aquellas posiciones marcadas como "X", la posición puede estar ocupada por cualquiera de los residuos de aminoácido enumerados debajo de la "X". Las posiciones marcadas como "x" y las regiones marcadas como CDR1 H1, CDR H2 y CDR H3 pueden estar ocupadas por cualquier aminoácido.

La figura 11 ilustra el armazón de región de marco de scFv para la familia de VH3. La primera fila muestra la numeración de región variable de cadena pesada usando el sistema de Kabat. La segunda fila muestra la numeración de región variable de cadena pesada usando el sistema de AHo. La tercera fila muestra la secuencia del armazón de región de marco de scFv (SEQ ID NO:3), en la que en aquellas posiciones marcadas como "X", la posición puede estar ocupada por cualquiera de los residuos de aminoácido enumerados debajo de la "X". Las posiciones marcadas como "x" y las regiones marcadas como CDR1 H1, CDR H2 y CDR H3 pueden estar ocupadas por cualquier aminoácido.

La figura 12 ilustra el armazón de región de marco de scFv para la familia de V $\kappa$ 1. La primera fila muestra la

5 numeración de región variable de cadena ligera usando el sistema de Kabat. La segunda fila muestra la numeración de región variable de cadena ligera usando el sistema de AHO. La tercera fila muestra la secuencia del armazón de región de marco de cadena ligera de scFv (SEQ ID NO:4), en la que en aquellas posiciones marcadas como "X", la posición puede estar ocupada por cualquiera de los residuos de aminoácido enumerados debajo de la "X". Las posiciones marcadas como "." y las regiones marcadas como CDR1 L1, CDR L2 y CDR L3 pueden estar ocupadas por cualquier aminoácido.

10 La figura 13 ilustra el armazón de región de marco de scFv para la familia de V<sub>K</sub>3. La primera fila muestra la numeración de región variable de cadena ligera usando el sistema de Kabat. La segunda fila muestra la numeración de región variable de cadena ligera usando el sistema de AHO. La tercera fila muestra la secuencia del armazón de región de marco de cadena ligera de scFv (SEQ ID NO:5), en la que en aquellas posiciones marcadas como "X", la posición puede estar ocupada por cualquiera de los residuos de aminoácido enumerados debajo de la "X". Las posiciones marcadas como "." y las regiones marcadas como CDR1 L1, CDR L2 y CDR L3 pueden estar ocupadas por cualquier aminoácido.

15 La figura 14 ilustra el armazón de región de marco de scFv para la familia de VL1. La primera fila muestra la numeración de región variable de cadena ligera usando el sistema de Kabat. La segunda fila muestra la numeración de región variable de cadena ligera usando el sistema de AHO. La tercera fila muestra la secuencia del armazón de región de marco de cadena ligera de scFv, en la que en aquellas posiciones marcadas como "X", la posición puede estar ocupada por cualquiera de los residuos de aminoácido enumerados debajo de la "X". Las posiciones marcadas como "." y las regiones marcadas como CDR1 L1, CDR L2 y CDR L3 pueden estar ocupadas por cualquier aminoácido. En determinadas realizaciones preferidas, las posiciones de AHO 58 y 67-72 dentro de CDR L1 están ocupadas por los siguientes residuos respectivos: D y NNQRPS.

20 La figura 15 representa las curvas de solubilidad de precipitación de PEG de ESBA105 silvestre y variantes de solubilidad del mismo.

25 La figura 16 representa los perfiles de desnaturalización térmica para ESBA105 silvestre y variantes de solubilidad del mismo tal como se miden tras exposición térmica en un amplio intervalo de temperaturas (25-96°C).

La figura 17 representa un gel de SDS-PAGE que muestra el comportamiento de degradación de diversos mutantes de solubilidad de ESBA105 después de dos semanas de incubación en condiciones de estrés térmico.

**Descripción detallada de la invención**

30 La invención se refiere a métodos para la modificación mediante ingeniería genética basada en secuencia y la optimización de propiedades de agentes de unión inmunológica, y en particular propiedades de scFv, incluyendo pero sin limitarse a estabilidad, solubilidad y/o afinidad. Más específicamente, la presente invención da a conocer métodos para optimizar anticuerpos scFv usando análisis de secuencia de anticuerpos para identificar posiciones de aminoácido dentro de un scFv que van a mutarse para mejorar de ese modo una o más propiedades físicas del scFv. La invención también se refiere a agentes de unión inmunológica modificados mediante ingeniería genética, por ejemplo, scFv, producidos o que pueden obtenerse según los métodos de la invención.

35 La invención se basa, al menos en parte, en el análisis de la frecuencia de aminoácidos en cada posición de región de marco de cadena pesada y ligera en múltiples bases de datos de secuencias de anticuerpos. En particular, el análisis de frecuencia de bases de datos de secuencias de anticuerpos (por ejemplo, bases de datos de secuencias de anticuerpos de líneas germinales o bases de datos de anticuerpos maduros, por ejemplo, la base de datos Kabat) se ha comparado con el análisis de frecuencia de una base de datos de secuencias de scFv que se han seleccionado por tener propiedades funcionales deseadas. Asignando un grado de variabilidad a cada posición de región de marco (por ejemplo, usando el índice de Simpson) y comparando el grado de variabilidad en cada posición de región de marco dentro de los diferentes tipos de bases de datos de secuencias de anticuerpos, ha sido posible ahora identificar posiciones de región de marco de importancia para las propiedades funcionales (por ejemplo, estabilidad, solubilidad) de un scFv. Esto permite ahora definir un "consenso funcional" a las posiciones de aminoácido de región de marco, en el que se han identificado posiciones de región de marco que son o bien más o bien menos tolerantes a la variabilidad que las posiciones correspondientes en secuencias de inmunoglobulina (por ejemplo, secuencias de inmunoglobulinas de líneas germinales o maduras). Por tanto, la invención proporciona, y demuestra el beneficio de, un enfoque de "consenso funcional" basándose en el uso de una base de datos de 40 secuencias de scFv seleccionadas funcionalmente. Todavía más, la invención proporciona métodos de modificación mediante ingeniería genética de agentes de unión inmunológica (por ejemplo, scFv) mutando posiciones de aminoácido de regiones de marco particulares identificadas usando el enfoque de "consenso funcional" descrito en el presente documento.

55 Para que la invención pueda entenderse más fácilmente, en primer lugar se definen ciertos términos. A menos que se definan de otro modo, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que entiende comúnmente un experto habitual en la técnica a la que pertenece esta invención. Aunque pueden usarse materiales y métodos similares o equivalentes a los descritos en el presente documento en la práctica o las pruebas de la invención, a continuación se describen materiales y métodos adecuados. Todas las

publicaciones, solicitudes de patente, patentes, y otras referencias mencionadas en el presente documento se incorporan como referencia en su totalidad. En caso de conflicto, prevalente la presente memoria descriptiva, incluyendo las definiciones. Además, los materiales, métodos y ejemplos son únicamente ilustrativos y no pretenden ser limitativos.

5 El término "anticuerpo" tal como se usa en el presente documento es un sinónimo de "inmunoglobulina". Los anticuerpos según la presente invención pueden ser inmunoglobulinas completas o fragmentos de las mismas, que comprenden al menos un dominio variable de una inmunoglobulina, tal como dominios variables individuales, Fv (Skerra A. y Plückerthun, A. (1988) *Science* 240:1038-41), scFv (Bird, R.E. *et al.* (1988) *Science* 242:423-26; Huston, J.S. *et al.* (1988; *Proc. Natl. Acad. Sci. USA* 85:5879-83), Fab, (Fab')<sub>2</sub> u otros fragmentos que conoce bien un experto en la técnica.

El término "región de marco de anticuerpo" tal como se usa en el presente documento se refiere a la parte del dominio variable, o bien VL o bien VH, que sirve como armazón para los bucles de unión a antígeno de este dominio variable (Kabat, E.A. *et al.*, (1991) *Sequences of proteins of immunological interest*. Publicación del NIH 91-3242).

15 El término "CDR de anticuerpo" tal como se usa en el presente documento se refiere a las regiones determinantes de complementariedad del anticuerpo que consisten en los bucles de unión a antígeno tal como se definieron por Kabat E.A. *et al.*, (1991) (*Sequences of proteins of immunological interest*. Publicación del NIH 91-3242). Cada uno de los dos dominios variables de un fragmento Fv de anticuerpo contiene, por ejemplo, tres CDR.

20 El término "anticuerpo de cadena sencilla" o "scFv" pretende referirse a una molécula que comprende una región variable de cadena pesada (V<sub>H</sub>) de anticuerpo y una región variable de cadena ligera (V<sub>L</sub>) de anticuerpo conectadas por un ligador. Tales moléculas de scFv pueden tener las estructuras generales: NH<sub>2</sub>-V<sub>L</sub>-ligador-V<sub>H</sub>-COOH o NH<sub>2</sub>-V<sub>H</sub>-ligador-V<sub>L</sub>-COOH.

25 Tal como se usa en el presente documento, "identidad" se refiere a la coincidencia de secuencia entre dos polipéptidos, moléculas o entre dos ácidos nucleicos. Cuando una posición en ambas de las dos secuencias comparadas está ocupada por la misma subunidad monomérica de aminoácido o base (por ejemplo, si una posición en cada una de las dos moléculas de ADN está ocupada por adenina, o una posición en cada uno de dos polipéptidos está ocupada por una lisina), entonces las moléculas respectivas son idénticas en esa posición. El "porcentaje de identidad" entre dos secuencias es una función del número de posiciones coincidentes compartidas por las dos secuencias dividido entre el número de posiciones comparadas x 100. Por ejemplo, si coinciden 6 de 10 de las posiciones en dos secuencias, entonces las dos secuencias tienen una identidad del 60%. A modo de ejemplo, las secuencias de ADN CTGACT y CAGGTT comparten una identidad del 50% (coinciden 3 de las 6 posiciones totales). Generalmente, se realiza una comparación cuando se alinean dos secuencias para dar una máxima identidad. Tal alineación puede proporcionarse usando, por ejemplo, el método de Needleman *et al.* (1970) *J Mol. Biol.* 48: 443-453, implementado de manera conveniente mediante programas informáticos tales como el programa Align (DNASTar, Inc.).

35 Secuencias "similares" son aquellas que, cuando se alinean, comparten residuos de aminoácido idénticos y similares, en los que residuos similares son sustituciones conservativas por residuos de aminoácido correspondientes en una secuencia de referencia alineada. A este respecto, una "sustitución conservativa" de un residuo en una secuencia de referencia es una sustitución por un residuo que es física o funcionalmente similar al residuo de referencia correspondiente, por ejemplo, que tiene un tamaño, forma, carga eléctrica, propiedades químicas similares, incluyendo la capacidad para formar enlaces covalentes o de hidrógeno, o similares. Por tanto, una secuencia "modificada con una sustitución conservativa" es una que difiere de una secuencia de referencia o una secuencia silvestre en la que están presentes una o más sustituciones conservativas. El "porcentaje de similitud" entre dos secuencias es una función del número de posiciones que contienen residuos coincidentes o sustituciones conservativas compartidas por las dos secuencias dividido entre el número de posiciones comparadas x 100. Por ejemplo, si coinciden 6 de 10 de las posiciones en dos secuencias y 2 de 10 posiciones contienen sustituciones conservativas, entonces las dos secuencias tienen una similitud positiva del 80%.

50 "Secuencia consenso de aminoácidos" tal como se usa en el presente documento se refiere a una secuencia de aminoácidos que puede generarse usando una matriz de al menos dos, y preferiblemente más, secuencias de aminoácidos alineadas, y permitiendo huecos en la alineación, de manera que es posible determinar el residuo de aminoácido más frecuente en cada posición. La secuencia consenso es aquella secuencia que comprende los aminoácidos que están representados más frecuentemente en cada posición. En el caso de que dos o más aminoácidos estén igualmente representados en una única posición, la secuencia consenso incluye ambos o todos de esos aminoácidos.

55 La secuencia de aminoácidos de una proteína puede analizarse a diversos niveles. Por ejemplo, puede presentarse conservación o variabilidad a nivel de un solo residuo, a nivel de múltiples residuos, múltiples residuos con huecos, etc. Los residuos pueden presentar conservación del residuo idéntico o pueden conservarse a nivel de clase. Los ejemplos de clases de aminoácido incluyen grupos R polares pero no cargados (serina, treonina, asparagina y glutamina); grupos R cargados positivamente (lisina, arginina e histidina); grupos R cargados negativamente (ácido glutámico y ácido aspártico); grupos R hidrófobos (alanina, isoleucina, leucina, metionina, fenilalanina, triptófano,

valina y tirosina); y aminoácidos especiales (cisteína, glicina y prolina). Un experto en la técnica conoce otras clases y pueden definirse usando determinaciones estructurales u otros datos para evaluar la capacidad de sustitución. En ese sentido, un aminoácido sustituible puede referirse a cualquier aminoácido que puede sustituirse y mantener la conservación funcional en esa posición.

5 Tal como se usa en el presente documento, cuando una secuencia de aminoácidos (por ejemplo, una primera secuencia de  $V_H$  o  $V_L$ ) se alinea con una o más secuencias de aminoácidos adicionales (por ejemplo, una o más secuencias de  $V_H$  o  $V_L$  en una base de datos), puede compararse una posición de aminoácido en una secuencia (por ejemplo, la primera secuencia de  $V_H$  o  $V_L$ ) con una "posición correspondiente" en la una o más secuencias de aminoácidos adicionales. Tal como se usa en el presente documento, la "posición correspondiente" representa la  
10 posición equivalente en la(s) secuencia(s) que está(n) comparándose cuando las secuencias están óptimamente alineadas, es decir, cuando las secuencias están alineadas para lograr el mayor porcentaje de identidad o porcentaje de similitud.

Tal como se usa en el presente documento, el término "base de datos de anticuerpos" se refiere a una colección de  
15 dos o más secuencias de aminoácidos de anticuerpos (una "multiplicidad" de secuencias), y normalmente se refiere a una colección de decenas, cientos o incluso miles de secuencias de aminoácidos de anticuerpos. Una base de datos de anticuerpos puede almacenar secuencias de aminoácidos de, por ejemplo, una colección de regiones  $V_H$  de anticuerpo, regiones  $V_L$  de anticuerpo o ambas, o puede almacenar una colección de secuencias de scFv que se compone de regiones  $V_H$  y  $V_L$ . Preferiblemente, la base de datos se almacena en un medio fijo, en el que pueden realizarse búsquedas, tal como en un ordenador dentro de un programa informático en el que pueden realizarse  
20 búsquedas. En una realización, la base de datos de anticuerpos es una base de datos que comprende o que consiste en secuencias de anticuerpos de líneas germinales. En otra realización, la base de datos de anticuerpos es una base de datos que comprende o que consiste en secuencias de anticuerpos maduros (es decir, expresados) (por ejemplo, una base de datos Kabat de secuencias de anticuerpos maduros, por ejemplo, una base de datos KBD). En aún otra realización, la base de datos de anticuerpos comprende o consiste en secuencias seleccionadas funcionalmente (por ejemplo, secuencias seleccionadas a partir de un ensayo QC).  
25

El término "agente de unión inmunológica" se refiere a una molécula que contiene la totalidad o una parte del sitio de unión a antígeno de un anticuerpo, por ejemplo, la totalidad o parte del dominio variable de cadena pesada y/o ligera, de manera que el agente de unión inmunológica reconoce específicamente un antígeno diana. Los ejemplos no limitativos de agentes de unión inmunológica incluyen moléculas de inmunoglobulina de longitud completa y scFv, así como fragmentos de anticuerpo, incluyendo pero sin limitarse a (i) un fragmento Fab, un fragmento monovalente que consiste en los dominios  $V_L$ ,  $V_H$ ,  $C_L$  y  $C_H1$ ; (ii) un fragmento  $F(ab')_2$ , un fragmento bivalente que comprende dos fragmentos Fab unidos mediante un puente disulfuro en la región bisagra; (iii) un fragmento  $Fab'$ , que es esencialmente un Fab con parte de la región bisagra (véase, FUNDAMENTAL IMMUNOLOGY (Paul ed., 3ª ed. 1993); (iv) un fragmento Fd que consiste en los dominios  $V_H$  y  $C_H1$ ; (v) un fragmento Fv que consiste en los dominios  $V_L$  y  $V_H$  de un solo brazo de un anticuerpo, (vi) un anticuerpo de un solo dominio tal como un fragmento Dab (Ward *et al.*, (1989) Nature 341:544-546), que consiste en un dominio  $V_H$  o  $V_L$ , un anticuerpo de camélido (véanse Hamers-Casterman, *et al.*, Nature 363:446-448 (1993) y Dumoulin, *et al.*, Protein Science 11:500-515 (2002)) o de tiburón (por ejemplo, Nanobodies® de Ig-NAR de tiburón; y (vii) un Nanobody, una región variable de cadena pesada que contiene un solo dominio variable y dos dominios constantes.  
30  
35

Tal como se usa en el presente documento, el término "propiedad funcional" es una propiedad de un polipéptido (por ejemplo, un agente de unión inmunológica) para el que se desea una mejora (por ejemplo, con relación a un polipéptido convencional) y/o resulta ventajoso para un experto en la técnica, por ejemplo, para mejorar las propiedades de fabricación o la eficacia terapéutica del polipéptido. En una realización, la propiedad funcional es estabilidad mejorada (por ejemplo, estabilidad térmica). En otra realización, la propiedad funcional es solubilidad mejorada (por ejemplo, en condiciones celulares). En aún otra realización, la propiedad funcional es la no agregación. En todavía otra realización, la propiedad funcional es una mejora en la expresión (por ejemplo, en una célula procarionota). En aún otra realización, la propiedad funcional es una mejora en el rendimiento de repliegamiento tras un proceso de purificación de cuerpos de inclusión. En determinadas realizaciones, la propiedad funcional no es una mejora en la afinidad de unión a antígeno.  
40  
45

#### 50 Análisis basado en secuencia de los scFv

La invención proporciona métodos para analizar una secuencia de scFv que permiten la identificación de posiciones de aminoácido dentro de la secuencia de scFv que va a seleccionarse para la mutación. Las posiciones de aminoácido seleccionadas para la mutación son las que se predice que influyen en propiedades funcionales del scFv, tales como solubilidad, estabilidad y/o unión a antígeno, en las que se predice que la mutación en tales  
55 posiciones mejora el desempeño del scFv. Por tanto, la invención permite una modificación mediante ingeniería genética más centrada de los scFv para optimizar su desempeño que simplemente mutar al azar posiciones de aminoácido dentro de la secuencia de scFv.

Determinados aspectos del análisis basado en secuencia de secuencias de scFv se esquematizan en el diagrama de flujo de la figura 1. Tal como se muestra en esta figura, se compara la secuencia de un scFv que va a optimizarse con las secuencias en una o más bases de datos de anticuerpos, incluyendo una base de datos de anticuerpos que  
60

se compone de secuencias de scFv seleccionadas como que son estables y solubles. Esto puede permitir la identificación de residuos críticos para la estabilidad y/o solubilidad específicamente en el formato de scFv, así como la identificación de patrones que representan mejoras en la estabilidad, solubilidad y/o unión independientemente de las CDR respectivas, específicamente en el formato de scFv (por ejemplo, combinaciones de V<sub>L</sub> y V<sub>H</sub>). Una vez que se han identificado los residuos críticos, pueden sustituirse, por ejemplo, por el aminoácido más frecuente adecuado identificado en la base de datos respectiva y/o mediante mutagénesis al azar o sesgada.

Por tanto, en un aspecto, la invención se refiere a un método de identificación de una posición de aminoácido para la mutación en un anticuerpo de cadena sencilla (scFv), teniendo el scFv secuencias de aminoácidos de V<sub>H</sub> y V<sub>L</sub>, comprendiendo el método:

- 10 a) introducir las secuencias de aminoácidos de V<sub>H</sub>, V<sub>L</sub> o V<sub>H</sub> y V<sub>L</sub> de scFv en una base de datos que comprende una multiplicidad de secuencias de aminoácidos de V<sub>H</sub>, V<sub>L</sub> o V<sub>H</sub> y V<sub>L</sub> de anticuerpo de manera que las secuencias de aminoácidos de V<sub>H</sub>, V<sub>L</sub> o V<sub>H</sub> y V<sub>L</sub> de scFv se alinean con las secuencias de aminoácidos de V<sub>H</sub>, V<sub>L</sub> o V<sub>H</sub> y V<sub>L</sub> de anticuerpo de la base de datos;
- 15 b) comparar una posición de aminoácido dentro de la secuencia de aminoácidos de V<sub>H</sub> o V<sub>L</sub> de scFv con una posición correspondiente dentro de las secuencias de aminoácidos de V<sub>H</sub> o V<sub>L</sub> de anticuerpo de la base de datos;
- c) determinar si la posición de aminoácido dentro de la secuencia de aminoácidos de V<sub>H</sub> o V<sub>L</sub> de scFv está ocupada por un residuo de aminoácido que está conservado en la posición correspondiente dentro de las secuencias de aminoácidos de V<sub>H</sub> o V<sub>L</sub> de anticuerpo de la base de datos; e
- 20 d) identificar la posición de aminoácido dentro de la secuencia de aminoácidos de V<sub>H</sub> o V<sub>L</sub> de scFv como una posición de aminoácido para la mutación cuando la posición de aminoácido está ocupada por un residuo de aminoácido que no está conservado en la posición correspondiente dentro de las secuencias de aminoácidos de V<sub>H</sub> o V<sub>L</sub> de anticuerpo de la base de datos.

Por tanto, en el método de la invención, la secuencia de un scFv de interés (es decir, la secuencia de V<sub>H</sub>, V<sub>L</sub> o ambos) se compara con las secuencias de una base de datos de anticuerpos y se determina si una posición de aminoácido en el scFv de interés está ocupada por un residuo de aminoácido que está "conservado" en la posición correspondiente de las secuencias en la base de datos. Si la posición de aminoácido de la secuencia de scFv está ocupada por un residuo de aminoácido que no está "conservado" en la posición correspondiente dentro de las secuencias de la base de datos, esa posición de aminoácido del scFv se elige para la mutación. Preferiblemente, la posición de aminoácido que se analiza es una posición de aminoácido de región de marco dentro del scFv de interés. Incluso más preferiblemente, puede analizarse cada posición de aminoácido de región de marco dentro del scFv de interés. En una realización alternativa, pueden analizarse una o más posiciones de aminoácido dentro de una o más CDR del scFv de interés. En aún otra realización, puede analizarse cada posición de aminoácido con el scFv de interés.

Para determinar si un residuo de aminoácido está "conservado" en una posición de aminoácido particular dentro de las secuencias de la base de datos de anticuerpos (por ejemplo, una posición de región de marco), puede calcularse el grado de conservación en la posición particular. Existe una variedad de diferentes modos conocidos en la técnica con los que puede cuantificarse la diversidad de aminoácidos en una posición dada, todos los cuales pueden aplicarse a los métodos de la presente invención. Preferiblemente, el grado de conservación se calcula usando el índice de diversidad de Simpson, que es una medida de la diversidad. Tiene en cuenta el número de aminoácidos presentes en cada posición, así como la abundancia relativa de cada aminoácido. El índice de Simpson (S.I.) representa la probabilidad de que dos secuencias de anticuerpos seleccionadas al azar contengan el mismo aminoácido en determinadas posiciones. El índice de Simpson tiene en cuenta dos factores principales al medir la conservación, riqueza y regularidad. Tal como se usa en el presente documento, "riqueza" es una medida del número de diferentes tipos de aminoácidos presentes en una posición particular (es decir, el número de diferentes residuos de aminoácido representados en la base de datos en esa posición es una medida de riqueza). Tal como se usa en el presente documento, "regularidad" es una medida de la abundancia de cada uno de los aminoácidos presentes en la posición particular (es decir, la frecuencia con la que aparecen los residuos de aminoácido en esa posición dentro de las secuencias de la base de datos es una medida de regularidad).

Aunque puede usarse la riqueza de residuos como una medida por sí misma para examinar el grado de conservación en una posición particular, no tiene en cuenta la frecuencia relativa de cada residuo de aminoácido presente en una determinada posición. Esto da mucho peso a aquellos residuos de aminoácido que aparecen de manera muy infrecuente en una posición particular dentro de las secuencias de una base de datos como también a aquellos residuos que aparecen de manera muy frecuente en la misma posición. La regularidad es una medida de la abundancia relativa de los diferentes aminoácidos que componen la riqueza de una posición. El índice de Simpson tiene en cuenta ambas, riqueza y regularidad, y por tanto es un modo preferido de cuantificar el grado de conservación según la presente invención. En particular, residuos poco frecuentes en posiciones muy conservadas se consideran potencialmente problemáticos y, por tanto, pueden elegirse para la mutación.

La fórmula para el índice de Simpson es  $D = \sum n_i (n_i - 1) / (N(N - 1))$ , en la que N es el número total de secuencias en la

consulta (por ejemplo, en la base de datos) y  $n_i$  es la frecuencia de cada residuo de aminoácido en la posición que está analizándose. La frecuencia de un acontecimiento de aminoácido ( $i$ ) en la base de datos es el número ( $n_i$ ) de veces que apareció el aminoácido en la base de datos. Los propios recuentos  $n_i$  se facilitan en frecuencias relativas, lo que significa que se normalizan según el número total de acontecimientos. Cuando se produce la máxima diversidad, el valor de S.I. es cero y cuando se produce la mínima diversidad, el valor de S.I. es 1. Por tanto, el intervalo de S.I. es 0-1, con una relación inversa entre la diversidad y el valor del índice.

En la figura 2 se describe en más detalle un diagrama de flujo que resume las múltiples etapas para el análisis de posiciones de aminoácido de región de marco dentro de las secuencias de la base de datos.

Por consiguiente, en una realización preferida del método descrito anteriormente, a la posición correspondiente dentro de la secuencia de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos se le asigna un grado de conservación usando el índice de Simpson. El valor de S.I. de esa posición correspondiente puede usarse como indicador de la conservación de esa posición.

En otras realizaciones, se usan alineaciones de confianza (es decir alineaciones de secuencias para las que se considera similitud de la estructura proteica) de secuencias de anticuerpos estrechamente relacionadas en la presente invención para generar matrices de abundancia relativa de aminoácidos y grado de conservación de posiciones determinadas. Estas matrices se diseñan para su uso en comparaciones en bases de datos de anticuerpo-anticuerpo. Se calcula la frecuencia observada de cada residuo y se compara con las frecuencias esperadas (que son esencialmente las frecuencias de cada residuo en los conjuntos de datos para cada posición).

El análisis de un anticuerpo scFv dado con el método descrito proporciona información sobre mutaciones admisibles biológicamente y residuos poco habituales en determinadas posiciones en el anticuerpo scFv dado y permite la predicción de posibles debilidades dentro de su región de marco. Puede usarse la rutina para modificar mediante ingeniería genética sustituciones de aminoácido que ajustan "de manera óptima" un conjunto de datos de aminoácido-frecuencia, usando el valor de S.I. y la frecuencia relativa como criterio.

El análisis basado en secuencia descrito anteriormente puede aplicarse a la región  $V_H$  del scFv, a la región  $V_L$  del scFv, o a ambas. Por tanto, en una realización, se introduce la secuencia de aminoácidos de  $V_H$  de scFv en la base de datos y se alinea con secuencias de aminoácidos de  $V_H$  de anticuerpo de la base de datos. En otra realización, se introduce la secuencia de aminoácidos de  $V_L$  de scFv en la base de datos y se alinea con secuencias de aminoácidos de  $V_L$  de anticuerpo de la base de datos. En aún otra realización, se introducen las secuencias de aminoácidos de  $V_H$  y  $V_L$  de scFv en la base de datos y se alinean con secuencias de aminoácidos de  $V_H$  y  $V_L$  de anticuerpo de la base de datos. Los algoritmos para alinear una secuencia con una colección de otras secuencias en una base de datos están bien establecidos en la técnica. Las secuencias se alinean de manera que se logre el mayor porcentaje de identidad o similitud entre las secuencias.

Los métodos de la invención pueden usarse para analizar una posición de aminoácido de interés dentro de una secuencia de scFv o, más preferiblemente, pueden usarse para analizar múltiples posiciones de aminoácido de interés. Por tanto, en la etapa b) del método descrito anteriormente, pueden compararse múltiples posiciones de aminoácido dentro de la secuencia de aminoácidos de  $V_H$  o  $V_L$  de scFv con posiciones correspondientes dentro de las secuencias de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos. Posiciones preferidas para analizarse son posiciones de región de marco dentro de las secuencias de  $V_H$  y/o  $V_L$  del scFv (por ejemplo, puede analizarse cada posición de región de marco de  $V_H$  y  $V_L$ ). Adicional o alternativamente, pueden analizarse una o más posiciones dentro de una o más CDR del scFv (aunque puede no preferirse mutar posiciones de aminoácido con las CDR, puesto que es más probable que las mutaciones dentro de las CDR afecten a la actividad de unión a antígeno que las mutaciones dentro de las regiones de marco). Todavía más, los métodos de la invención permiten el análisis de cada posición de aminoácido dentro de las secuencias de aminoácidos de  $V_H$ ,  $V_L$  o  $V_H$  y  $V_L$  de scFv.

En los métodos de la invención, la secuencia de un scFv de interés puede compararse con las secuencias dentro de una o más de una variedad de diferentes tipos de bases de datos de secuencias de anticuerpos. Por ejemplo, en una realización, las secuencias de aminoácidos de  $V_H$ ,  $V_L$  o  $V_H$  y  $V_L$  de anticuerpo de la base de datos son secuencias de aminoácidos de  $V_H$ ,  $V_L$  o  $V_H$  y  $V_L$  de anticuerpo de línea germinal. En otra realización, las secuencias de aminoácidos de  $V_H$ ,  $V_L$  o  $V_H$  y  $V_L$  de anticuerpo de la base de datos son secuencias de aminoácidos de  $V_H$ ,  $V_L$  o  $V_H$  y  $V_L$  de anticuerpo de afinidad maduro reorganizadas. En aún otra realización preferida, las secuencias de aminoácidos de  $V_H$ ,  $V_L$  o  $V_H$  y  $V_L$  de anticuerpo de la base de datos son secuencias de aminoácidos de  $V_H$ ,  $V_L$  o  $V_H$  y  $V_L$  de anticuerpo scFv seleccionadas por tener al menos una propiedad funcional deseable, tal como estabilidad de scFv o solubilidad de scFv (comentado adicionalmente a continuación).

La información de secuencias de anticuerpos puede obtenerse, recopilarse y/o generarse a partir de alineaciones de secuencias de secuencias de líneas germinales o de cualquier otra secuencia de anticuerpo que se produzca en la naturaleza. Las fuentes de secuencias pueden incluir, pero sin limitarse a, una o más de las siguientes bases de datos:

- La base de datos Kabat ([immuno.bme.nwu.edu](http://immuno.bme.nwu.edu)) (desde octubre de 2007); Johnson & Wu (2001) Nucleic Acids Res. 29: 205-206; Johnson & Wu (2000) Nucleic Acids Res. 28: 214-218). Están disponibles los datos sin

procesar de 2000 por FTP en los EE.UU. y duplicado en R.U.

- Kabatman contiene una base de datos que permite al usuario buscar la secuencia de Kabat para determinar características poco habituales de secuencia y permite al usuario hallar asignaciones canónicas para las CDR en una secuencia de anticuerpo específica.
- 5 • El sitio web AAAAA (www.bioc.unizh.ch/antibody/ (desde octubre de 2007)), una página sobre anticuerpos preparada por Annemarie Honegger que proporciona información de secuencia y datos estructurales sobre anticuerpos.
- ABG: Directorio de estructuras 3D de anticuerpos - El directorio, creado por el Grupo sobre Anticuerpos (ABG), permite al usuario acceder a las estructuras de anticuerpos recopiladas en el Banco de Datos de Proteínas (PDB). En el directorio, cada entrada del PDB tiene un hipervínculo a la fuente original para facilitar la recuperación de la información completa
- 10 • ABG: Directorios de genes de líneas germinales de los segmentos de líneas germinales de VH y VK de ratón, parte de la página web del Grupo sobre Anticuerpos en el Instituto de Biotecnología, UNAM (Universidad Nacional de México)
- 15 • IMGT®, el sistema de información internacional ImMunoGeneTics® - creado en 1989 por Marie-Paule Lefranc (Université Montpellier II, CNRS), IMGT es un recurso de conocimiento integrado especializado en inmunoglobulinas, receptores de células T y proteínas relacionadas del sistema inmunitario para el ser humano y otras especies de vertebrados. IMGT consiste en bases de datos de secuencias (IMGT/LIGM-DB, una base de datos exhaustiva de IG y TR de ser humano y otros vertebrados, con traducción para secuencias completamente anotadas, IMGT/MHC-DB, IMGT/PRIMER-DB), una base de datos genómica (IMGT/GENE-DB), una base de datos estructural (IMGT/3Dstructure-DB), un recurso web (IMGT Repertoire) (IMGT, el sistema de información internacional ImMunoGeneTics®; imgt.cines.fr (desde octubre de 2007); Lefranc *et al.* (1999) *Nucleic Acids Res.* 27: 209-212; Ruiz *et al.* (2000) *Nucleic Acids Res.* 28: 219-221; Lefranc *et al.* (2001) *Nucleic Acids Res.* 29: 207-209; Lefranc *et al.* (2003) *Nucleic Acids Res.* 31: 307-310).
- 20 • V BASE - un directorio exhaustivo de todas las secuencias de región variable de líneas germinales humanas recopiladas a partir de más de mil secuencias publicadas, incluyendo las que se encuentran en las versiones actuales de las bibliotecas de datos Genbank y EMBL.
- 25

En una realización preferida, la información de secuencias de anticuerpos se obtiene de una biblioteca de scFv que tiene regiones de marco definidas que se han seleccionado para una estabilidad y solubilidad potenciadas en un entorno reductor. Más específicamente, se ha descrito un sistema de control de calidad (QC) en levadura (véanse por ejemplo, la publicación PCT WO2001/48017; las solicitudes estadounidenses n.ºs 2001/0024831 y US 2003/0096306; las patentes estadounidenses n.ºs 7.258.985 y 7.258.986) que permite la selección intracelular de regiones de marco de scFv con una estabilidad y solubilidad potenciadas en un entorno reductor. En este sistema, se transforma una biblioteca de scFv en células huésped que pueden expresar un antígeno conocido específico y que sólo sobreviven en presencia de la interacción antígeno-scFv. Las células huésped transformadas se cultivan en condiciones adecuadas para la expresión del antígeno y el scFv y que permiten la supervivencia celular sólo en presencia de la interacción antígeno-scFv. Por tanto, pueden aislarse los scFv expresados en las células supervivientes y que tienen regiones de marco definidas que son estables y solubles en un entorno reductor. Por consiguiente, puede usarse el sistema QC para examinar una gran biblioteca de scFv para aislar de ese modo aquellos scFv preferidos que tienen regiones de marco que son estables y solubles en un entorno reductor y las secuencias de aquellos scFv seleccionados pueden recopilarse en una base de datos de secuencias de scFv. Entonces puede usarse una base de datos de scFv de este tipo para fines de comparación con otras secuencias de scFv de interés usando los métodos de la presente invención. Secuencias de región de marco de scFv preferidas que se han seleccionado y definido previamente usando el sistema QC se describen en más detalle en la publicación PCT WO2003/097697 y la solicitud estadounidense n.º 20060035320.

Se conocen en la técnica variantes del sistema QC original. En una realización a modo de ejemplo, que se ilustra esquemáticamente en la figura 3, se fusiona una biblioteca de scFv al dominio de activación (AD) del factor de transcripción de levadura Gal4, que se fusiona a su vez a una parte de la denominada proteína Gal11p (11p). El constructo de fusión scFv-AD-Gal11p se transforma entonces en células huésped que expresan los primeros 100 aminoácidos de Gal4 y, por tanto, contienen el dominio de unión a ADN de Gal4 (DBD; Gal4 (1-100)). Gal11p es una mutación puntual que se sabe que se une directamente a Gal4 (1-100) (véase Barberis *et al.*, *Cell*, 81: 359 (1995)). Las células huésped transformadas se cultivan en condiciones que son adecuadas para la expresión de la proteína de fusión de scFv y que permiten la supervivencia celular sólo en el caso de que la proteína de fusión de scFv sea lo suficientemente estable y soluble como para interactuar con Gal4 (1-100) y formar de ese modo un factor de transcripción funcional que contiene un AD unido a un DBD (figura 3A). Por tanto, pueden aislarse los scFv expresados en las células supervivientes y que tienen regiones de marco definidas que son estables y solubles en un entorno reductor. Una descripción adicional de este sistema QC a modo de ejemplo se describe en Auf der Maur *et al.*, *Methods*, 34: 215-224 (2004).

En otra realización a modo de ejemplo, en la figura 4 se representa un sistema QC empleado en los métodos de la invención. En esta versión del sistema QC, el scFv o la biblioteca de scFv se fusiona directamente a un factor de transcripción funcional y se expresa en una cepa de levadura que contiene un marcador seleccionable. El marcador seleccionable sólo se activará en presencia de una fusión funcional de scFv-factor de transcripción, lo que significa que es necesario que el constructo en su totalidad sea estable y soluble (figura 4A). En el caso de que el scFv sea inestable, formará agregados y eventualmente se degradará, provocando también de ese modo la degradación del factor de transcripción fusionado al mismo de modo que ya no puede activar la expresión del marcador seleccionable (véase la figura 4B).

En los métodos de la invención, la secuencia de un scFv de interés puede compararse con todas las secuencias dentro de una base de datos de anticuerpos o, alternativamente, sólo una parte seleccionada de las secuencias en la base de datos puede usarse para fines de comparación. Es decir, la base de datos puede estar limitada, o restringida, sólo a aquellas secuencias que tienen un alto porcentaje de similitud o identidad con el scFv de interés. Por tanto, en una realización del método de la invención, la base de datos es una base de datos restringida en la que sólo aquellas secuencias de aminoácidos de  $V_H$ ,  $V_L$  o  $V_H$  y  $V_L$  de anticuerpo que tienen alta similitud con las secuencias de aminoácidos de  $V_H$ ,  $V_L$  o  $V_H$  y  $V_L$  de anticuerpo scFv están incluidas en la base de datos.

Una vez que se introduce la secuencia de scFv de interés en la base de datos y se compara con las secuencias de anticuerpos dentro de la base de datos, se analiza la información de secuencia para proporcionar información sobre la frecuencia y variabilidad de aminoácidos de una posición dada y predecir posiciones de aminoácido potencialmente problemáticas, en posiciones de aminoácido particulares potencialmente problemáticas dentro de la región de marco del scFv. Tal información también puede usarse para diseñar mutaciones que mejoren las propiedades del scFv. Por ejemplo, puede mejorarse la solubilidad del anticuerpo reemplazando residuos hidrófobos expuestos al disolvente por residuos hidrófilos que se producen por lo demás de manera frecuente en esta posición.

En el método de la invención, existen varios posibles tipos de residuos de aminoácido que pueden estar "conservados" en una posición particular dentro de las secuencias de anticuerpos de la base de datos. Por ejemplo, un residuo de aminoácido particular puede encontrarse en esa posición a una frecuencia muy alta, indicando que este residuo de aminoácido particular se prefiere en esa posición particular. Por consiguiente, en una realización del método, en la etapa c), el residuo de aminoácido que está conservado en la posición correspondiente dentro de las secuencias de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos es el residuo de aminoácido que está con la mayor frecuencia en esa posición dentro de las secuencias de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos. En otras realizaciones, la posición puede estar "conservada" con un tipo o clase particular de residuo de aminoácido (es decir, la posición no está ocupada sólo preferentemente por un único residuo de aminoácido particular, sino que más bien está ocupada preferentemente por varios residuos de aminoácido diferentes cada uno de los cuales es del mismo tipo o clase de residuo). Por ejemplo, en la etapa c), la posición correspondiente dentro de las secuencias de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos puede estar conservada con: (i) residuos de aminoácido hidrófobos, (ii) residuos de aminoácido hidrófilos, (iii) residuos de aminoácido que pueden formar un enlace de hidrógeno o (iv) residuos de aminoácido que tienen propensión a formar una lámina  $\beta$ .

En la etapa d) del método, se identifica una posición de aminoácido dentro de la secuencia de aminoácidos de  $V_H$  o  $V_L$  de scFv como una posición de aminoácido para la mutación cuando la posición de aminoácido está ocupada por un residuo de aminoácido que no está conservado en la posición correspondiente dentro de las secuencias de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos. Existen varias situaciones posibles que identificarían una posición de aminoácido como ocupada por un residuo de aminoácido que no está "conservada" y, por tanto, como potencialmente problemática. Por ejemplo, si la posición de aminoácido correspondiente dentro de la base de datos está conservada con un residuo hidrófobo y la posición en el scFv está ocupada por un residuo hidrófilo, esta posición podría ser potencialmente problemática en el scFv y la posición puede seleccionarse para la mutación. Asimismo, si la posición de aminoácido correspondiente dentro de la base de datos está conservada con un residuo hidrófilo y la posición en el scFv está ocupada por un residuo hidrófobo, esta posición podría ser potencialmente problemática en el scFv y la posición puede seleccionarse para la mutación. En todavía otros casos, si la posición de aminoácido correspondiente dentro de la base de datos está conservada con residuos de aminoácido que pueden formar un enlace de hidrógeno o que tienen propensión a formar una lámina  $\beta$ , y la posición en el scFv está ocupada por un residuo que no puede formar un enlace de hidrógeno o no tiene propensión a formar una lámina  $\beta$ , respectivamente, esta posición podría ser potencialmente problemática en el scFv y la posición puede seleccionarse para la mutación.

En una realización preferida, los métodos descritos en la presente invención pueden usarse solos o en combinación para crear listas combinatorias de sustituciones de aminoácidos para mejorar la estabilidad y/o solubilidad de fragmentos de anticuerpo de cadena sencilla.

#### Análisis de la covarianza

La invención también se refiere a métodos para analizar la covarianza dentro de la secuencia de un scFv en comparación con secuencias de anticuerpos dentro de una base de datos. Residuos que presentan covarianza pueden ser, por ejemplo, (i) un residuo en una región de marco (FR) y un residuo en una CDR; (ii) un residuo en una CDR y un residuo en otra CDR; o (iii) un residuo en el dominio  $V_H$  y un residuo en el  $V_L$ . Residuos que interactúan

entre sí en la estructura terciaria del anticuerpo pueden presentar covarianza de manera que residuos de aminoácido preferido pueden estar conservados en ambas posiciones del par covariante y si un residuo se altera, debe alterarse el otro residuo también para mantener la estructura del anticuerpo. Se conocen en la técnica métodos para realizar un análisis de la covarianza en un conjunto de secuencias de aminoácidos. Por ejemplo, Choulier, L. *et al.* (2000) *Protein* 41:475-484 describen aplicar un análisis de la covarianza a alineaciones de secuencias de  $V_H$  y  $V_L$  de líneas germinales humanas y de ratón.

Puede combinarse un análisis de la covarianza con el método descrito anteriormente para analizar posiciones de aminoácido conservadas (etapas a-d) en el método anterior), de manera que el método comprende además las etapas de:

- 10 e) llevar a cabo un análisis de la covarianza en la secuencia de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos para identificar un par covariante de posiciones de aminoácido;
- f) comparar el par covariante de posiciones de aminoácido con posiciones correspondientes dentro de la secuencia de aminoácidos de  $V_H$  o  $V_L$  de scFv;
- 15 g) determinar si las posiciones correspondientes dentro de la secuencia de aminoácidos de  $V_H$  o  $V_L$  de scFv están ocupadas por residuos de aminoácido que están conservados en el par covariante de posiciones de aminoácido dentro de las secuencias de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos; e
- 20 h) identificar una o ambas de las posiciones correspondientes dentro de la secuencia de aminoácidos de  $V_H$  o  $V_L$  de scFv como una posición de aminoácido para la mutación cuando una o ambas de las posiciones correspondientes dentro del scFv está ocupada por un residuo de aminoácido que no está conservado en el par covariante de posiciones de aminoácido dentro de las secuencias de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos.

Adicional o alternativamente, puede realizarse un análisis de la covarianza por sí mismo, de manera que la invención proporciona un método que comprende las etapas:

- a) llevar a cabo un análisis de la covarianza en secuencias de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de una base de datos para identificar un par covariante de posiciones de aminoácido;
- 25 b) comparar el par covariante de posiciones de aminoácido con posiciones correspondientes dentro de una secuencia de aminoácidos de  $V_H$  o  $V_L$  de scFv;
- c) determinar si las posiciones correspondientes dentro de la secuencia de aminoácidos de  $V_H$  o  $V_L$  de scFv están ocupadas por residuos de aminoácido que están conservados en el par covariante de posiciones de aminoácido dentro de las secuencias de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos; e
- 30 d) identificar una o ambas de las posiciones correspondientes dentro de la secuencia de aminoácidos de  $V_H$  o  $V_L$  de scFv como una posición de aminoácido para la mutación cuando una o ambas de las posiciones correspondientes dentro del scFv está ocupada por un residuo de aminoácido que no está conservado en el par covariante de posiciones de aminoácido dentro de las secuencias de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos.

35 Los métodos de análisis de la covarianza de la invención pueden usarse para analizar un par covariante, o más de un par covariante. Por tanto, en una realización del método, se identifican múltiples pares covariantes de posiciones de aminoácido dentro de la secuencia de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos y se comparan con las posiciones correspondientes dentro de la secuencia de aminoácidos de  $V_H$  o  $V_L$  de scFv.

40 El método puede comprender además mutar una o ambas de las posiciones correspondientes dentro del scFv que están ocupadas por un residuo de aminoácido que no está conservado en el par covariante de posiciones de aminoácido dentro de las secuencias de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos. En una realización, una de las posiciones correspondientes dentro del scFv que está ocupada por un residuo de aminoácido que no está conservado en el par covariante de posiciones de aminoácido se sustituye por un residuo de aminoácido que está con la mayor frecuencia en la posición de aminoácido del par covariante. En otra realización, ambas posiciones correspondientes dentro del scFv que están ocupadas por residuos de aminoácido que no están conservados en el par covariante de posiciones de aminoácido se sustituyen por residuos de aminoácido que están con la mayor frecuencia en las posiciones de aminoácido del par covariante.

#### Modelado molecular

50 Los métodos basados en secuencia de la invención para analizar los scFv para detectar residuos potencialmente problemáticos pueden combinarse con otros métodos conocidos en la técnica para analizar relaciones de estructura/función de anticuerpos. Por ejemplo, en una realización preferida, los métodos analíticos basados en secuencia de la invención se combinan con modelado molecular para identificar residuos potencialmente problemáticos adicionales. Los métodos y el software para el modelado por ordenador de estructuras de anticuerpos, incluyendo estructuras de scFv, están establecidos en la técnica y pueden combinarse con los métodos basados en secuencia de la invención. Por tanto, en otra realización, los métodos basados en secuencia descritos

anteriormente tal como se exponen en las etapas a) - d) comprenden además las etapas de:

e) someter las secuencias de aminoácidos de  $V_H$ ,  $V_L$  o  $V_H$  y  $V_L$  de scFv a modelado molecular; e

f) identificar al menos una posición de aminoácido adicional dentro de las secuencias de aminoácidos de  $V_H$ ,  $V_L$  o  $V_H$  y  $V_L$  de scFv para la mutación.

- 5 El método puede comprender además mutar la al menos una posición de aminoácido adicional dentro de las secuencias de aminoácidos de  $V_H$ ,  $V_L$  o  $V_H$  y  $V_L$  de scFv identificadas para la mutación mediante modelado molecular.

Análisis de “consenso funcional” frente a “consenso convencional”

10 En una realización particularmente preferida, se compara el grado de variabilidad en una o más posiciones de región de marco entre una primera base de datos de secuencias de anticuerpos (por ejemplo, una(s) base(s) de datos de líneas germinales (por ejemplo, Vbase y/o IMGT) o una base de datos de anticuerpos maduros (por ejemplo, KBD) y una segunda base de datos de scFv seleccionada por tener una o más propiedades deseables, por ejemplo, una base de datos de scFv seleccionada mediante examen QC en levadura, es decir, una base de datos QC. Tal como se ilustra en la figura 5, puede asignarse un valor de variabilidad (por ejemplo, el valor del índice de Simpson) a

15 posiciones de región de marco dentro de la primera base de datos (por ejemplo, de líneas germinales), denominado valores de “G” en la figura 5, y puede asignarse un valor de variabilidad (por ejemplo, el valor del índice de Simpson) a las posiciones de región de marco correspondientes dentro de la segunda base de datos (por ejemplo, base de datos QC), denominado valores de “Q” en la figura 5. Cuando el valor de G es mayor que el valor de Q en una posición particular (es decir, más variabilidad en las secuencias de líneas germinales en esa posición que en las

20 secuencias de scFv seleccionadas), esto indica que existe un número restringido de residuos de aminoácido de región de marco de scFv estables en esa posición, residuos de aminoácido de región de marco de scFv estables que pueden ser adecuados para su uso con cualquier CDR. Alternativamente, cuando el valor de G es menor que el valor de Q en una posición particular (es decir, más variabilidad en las secuencias de scFv seleccionadas en esa posición que en las secuencias de líneas germinales), esto indica que esta posición particular es más tolerante a la

25 variabilidad en el scFv y, por tanto, puede representar una posición en la que sustituciones de aminoácidos pueden optimizar la estabilidad y/o solubilidad del scFv. La tabla A presenta una tabla resumen del número de posiciones de aminoácido, y residuos de región de marco altamente variables (hvFR), en los que o bien G es mayor que Q o bien G es menor que Q. Tal como se indica en la tabla A, la variabilidad en el número total de aminoácidos (N.º de aa) y en residuos de región de marco altamente variables (hvFR) aumentó significativamente entre línea germinal y QC-

30 FW. Las secuencias que se analizaron para generar la tabla A eran aproximadamente 90 secuencias de scFv que se seleccionaron usando el ensayo QC (tal como se describe en el documento WO03097697; denominado en el presente documento “Q”) y todas las secuencias de  $V_H$  y  $V_L$  de líneas germinales recuperadas de <http://www.bioc.unizh.ch/antibody/Sequences/index.html> en octubre de 2007 (denominado en el presente documento “G”). Para el análisis de la tabla A, no se agruparon los dominios  $V_H$  y  $V_L$  según su subtipo.

35 Tabla A: Tabla resumen

	N.º de aa	G<Q (n.º de casos)	G>Q (n.º de casos)	X/Y	n.º de hvFR (Simpson < 0,4)	G<Q (n.º de casos)	G>Q (n.º de casos)	X/Y
$V_L$	108	61	11	5,5	16	13	3	4,3
$V_H$	116	50	18	2,8	27	22	5	4,4

En vista de lo anterior, en aún otro aspecto, la invención proporciona un método de identificación de una o más posiciones de aminoácido de región de marco para la mutación en un anticuerpo de cadena sencilla (scFv), teniendo el scFv secuencias de aminoácidos de  $V_H$  y  $V_L$ , comprendiendo el método:

- 40 a) proporcionar una primera base de datos de secuencias de aminoácidos de  $V_H$ ,  $V_L$  o  $V_H$  y  $V_L$  (por ejemplo, secuencias de líneas germinales y/o anticuerpos maduros);
- b) proporcionar una segunda base de datos de las secuencias de aminoácidos de  $V_H$ ,  $V_L$  o  $V_H$  y  $V_L$  de anticuerpo scFv seleccionadas por tener al menos una propiedad funcional deseable;
- c) determinar la variabilidad de aminoácidos en cada posición de región de marco de la primera base de datos y en cada posición de región de marco de la segunda base de datos;
- 45 d) identificar una o más posiciones de región de marco en la que el grado de variabilidad de aminoácidos difiere entre la primera base de datos y la segunda base de datos para identificar de ese modo una o más posiciones de aminoácido de región de marco para la mutación en un anticuerpo de cadena sencilla (scFv).

Preferiblemente, la variabilidad de aminoácidos en cada posición de región de marco se determina mediante la asignación de un grado de conservación usando el índice de Simpson. En una realización, se identifica la una o más

50 posiciones de aminoácido de región de marco para la mutación basándose en la una o más posiciones de

aminoácido de región de marco que tienen un menor valor del índice de Simpson en la segunda base de datos (scFv) en comparación con la primera base de datos. En otra realización, se identifica la una o más posiciones de aminoácido de región de marco para la mutación basándose en la una o más posiciones de aminoácido de región de marco que tienen un mayor valor del índice de Simpson en la segunda base de datos en comparación con la primera base de datos.

En los ejemplos 2 y 3 a continuación se describen en mayor detalle los análisis de variabilidad y la identificación de residuos para la mutación, para tres familias de V<sub>H</sub> humano y tres familias de V<sub>L</sub> humano.

Análisis de enriquecimiento / exclusión

En otro aspecto, la invención proporciona métodos para seleccionar sustituciones de residuos de aminoácido preferidas (o, alternativamente, excluir sustituciones de aminoácidos particulares) en una posición de región de marco de interés dentro de un agente de unión inmunológica (por ejemplo, para mejorar una propiedad funcional tal como estabilidad y/o solubilidad). Los métodos de la invención comparan la frecuencia de un residuo de aminoácido en una posición de región de marco de interés en una primera base de datos de secuencias de anticuerpos (por ejemplo, base(s) de datos de líneas germinales tales como Vbase y/o IMGT o, más preferiblemente, una base de datos de anticuerpos maduros tal como la base de datos Kabat (KBD)) con la frecuencia del residuo de aminoácido en una posición de aminoácido correspondiente en una segunda base de datos de scFv seleccionada por tener una o más propiedades deseables, por ejemplo, una base de datos de scFv seleccionada mediante examen QC en levadura, por ejemplo, una base de datos QC.

Tal como se describe en detalle en el ejemplo 4 a continuación, pueden agruparse secuencias de anticuerpos (por ejemplo, secuencias de VH o VL) de la primera base de datos (por ejemplo, una base de datos de secuencias de anticuerpos maduros) según su subtipo de familia de Kabat (por ejemplo, Vh1b, VH3, etc.). Dentro de cada subtipo de secuencia (es decir, subfamilia), se determina la frecuencia de cada residuo de aminoácido (por ejemplo, A, V, etc.) en cada posición de aminoácido como porcentaje de todas las secuencias analizadas de ese subtipo. Se hace lo mismo para todas las secuencias de la segunda base de datos (por ejemplo, una base de datos de scFv seleccionada por tener una o más propiedades deseables, por ejemplo, mediante examen QC). Para cada subtipo, los porcentajes resultantes (frecuencias relativas) para cada residuo de aminoácido en una posición particular se comparan entre la primera y segunda bases de datos. Cuando la frecuencia relativa de un residuo de aminoácido determinado aumenta en la segunda base de datos (por ejemplo, una base de datos QC) con relación a la primera base de datos (por ejemplo, base de datos Kabat), esto indica que el residuo respectivo se selecciona favorablemente (es decir, un "residuo enriquecido") y confiere propiedades favorables a la secuencia. A la inversa, cuando la frecuencia relativa del residuo de aminoácido disminuye en la segunda base de datos con relación a la primera base de datos, esto indica que el residuo respectivo está desfavorecido (es decir, un "residuo excluido"). Por consiguiente, los residuos enriquecidos son residuos preferidos para mejorar las propiedades funcionales (por ejemplo, estabilidad y/o solubilidad) de un agente de unión inmunológica, mientras que se evitan preferiblemente los residuos excluidos.

En vista de lo anterior, en una realización, la invención proporciona un método de identificación de un residuo de aminoácido preferido para la sustitución en un agente de unión inmunológica, comprendiendo el método:

a) proporcionar una primera base de datos de secuencias de aminoácidos de V<sub>H</sub> o V<sub>L</sub> agrupadas (por ejemplo, secuencias de líneas germinales y/o anticuerpos maduros agrupadas según su subtipo de familia de Kabat);

b) proporcionar una segunda base de datos secuencias de aminoácidos de V<sub>H</sub> o V<sub>L</sub> de anticuerpo scFv agrupadas seleccionadas por tener al menos una propiedad funcional deseable (por ejemplo, según el ensayo QC);

c) determinar la frecuencia de aminoácidos para un residuo de aminoácido en una posición de región de marco de la primera base de datos y en una posición de región de marco correspondiente de la segunda base de datos;

d) identificar el residuo de aminoácido como residuo de aminoácido preferido para la sustitución en una posición de aminoácido correspondiente del agente de unión inmunológica cuando el residuo de aminoácido aparece a una mayor frecuencia en la segunda base de datos con relación a la primera base de datos (es decir, un residuo enriquecido).

Puede cuantificarse el enriquecimiento de un residuo de aminoácido en la segunda base de datos (scFv) (por ejemplo, una base de datos QC). Por ejemplo, puede determinarse la razón entre la frecuencia relativa de un residuo dentro de la segunda base de datos (RF2) y la frecuencia relativa de un residuo dentro de la primera base de datos (RF1). Esta razón (RF2:RF1) puede denominarse "factor de enriquecimiento" (EF). Por consiguiente, en determinadas realizaciones, el residuo de aminoácido en la etapa (d) se identifica si la razón de la frecuencia relativa del residuo de aminoácido entre la primera y segunda bases de datos (en el presente documento, el "factor de enriquecimiento") es de al menos 1 (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9 ó 10). En una realización preferida, el factor de enriquecimiento es mayor de aproximadamente 1,0 (por ejemplo 1,0, 1,1, 1,2, 1,3, 1,4 ó 1,5). En aún otra realización preferida, el factor de enriquecimiento es de aproximadamente 4,0 a aproximadamente 6,0 (por ejemplo, 4,0, 4,1, 4,2, 4,3, 4,4, 4,5, 4,6, 4,7, 4,8, 4,9, 5,0, 5,1, 5,2, 5,3, 5,4, 5,5, 5,6, 5,7, 5,8, 5,9 ó 6,0). En otra realización, el factor de enriquecimiento es de aproximadamente 6,0 a aproximadamente 8,0 (por ejemplo, 6,0, 6,1, 6,2, 6,3, 6,4,

6,5, 6,6, 6,7, 6,8, 6,9, 7,0, 7,1, 7,3, 7,4, 7,5, 7,6, 7,7, 7,8, 7,9 u 8,0). En otras realizaciones, el factor de enriquecimiento es mayor de 10 (por ejemplo, 10, 100, 1000,  $10^4$ ,  $10^5$ ,  $10^6$ ,  $10^7$ ,  $10^8$ ,  $10^9$  o más). En determinadas realizaciones, pueden lograrse factores de enriquecimiento infinitos.

5 En otra realización, la invención proporciona un método de identificación de un residuo de aminoácido que va a excluirse en una posición particular de un agente de unión inmunológica, comprendiendo el método:

a) proporcionar una primera base de datos de secuencias de aminoácidos de  $V_H$  o  $V_L$  agrupadas (por ejemplo, secuencias de líneas germinales y/o anticuerpos maduros agrupadas según su subtipo de familia de Kabat);

b) proporcionar una segunda base de datos de secuencias de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo scFv agrupadas seleccionadas por tener al menos una propiedad funcional deseable (por ejemplo, según el ensayo QC);

10 c) determinar la frecuencia de aminoácidos para un residuo de aminoácido en una posición de región de marco de la primera base de datos y en una posición de región de marco correspondiente de la segunda base de datos;

d) identificar el residuo de aminoácido como residuo de aminoácido desfavorecido para la sustitución en la posición de aminoácido correspondiente del agente de unión inmunológica cuando el residuo de aminoácido aparece a una menor frecuencia en la segunda base de datos con relación a la primera base de datos, en el que dicho tipo de residuo de aminoácido es un residuo de aminoácido desfavorecido (es decir, un residuo excluido). En determinadas realizaciones preferidas, se identifica el residuo de aminoácido desfavorecido en la etapa (d) anteriormente si el factor de enriquecimiento (EF) es menor de 1.

#### Mutación de los scFv

20 En los métodos de la invención, una vez que se han identificado una o más posiciones de aminoácido dentro de un scFv como potencialmente problemáticas con respecto a las propiedades funcionales del scFv, el método puede comprender además mutar estas una o más posiciones de aminoácido dentro de la secuencia de aminoácidos de  $V_H$  o  $V_L$  de scFv. Por ejemplo, una posición de aminoácido identificada para la mutación puede sustituirse por un residuo de aminoácido que está conservado o enriquecido en la posición correspondiente dentro de las secuencias de aminoácidos  $V_H$  o  $V_L$  de anticuerpo de la base de datos.

25 Una posición de aminoácido identificada para la mutación puede mutarse usando uno de varios métodos de mutagénesis posibles bien establecidos en la técnica. Por ejemplo, puede usarse mutagénesis dirigida al sitio para realizar una sustitución de aminoácido particular en la posición de aminoácido de interés. También puede usarse mutagénesis dirigida al sitio para crear un conjunto de scFv mutados en los que se ha introducido un repertorio limitado de sustituciones de aminoácidos en la posición de aminoácido de interés.

30 Adicional o alternativamente, la(s) posición/posiciones de aminoácido identificada(s) para la mutación puede(n) mutarse mediante mutagénesis al azar o mediante mutagénesis sesgada para generar una biblioteca de scFv mutados, seguido por el examen de la biblioteca de scFv mutados y la selección de los scFv, preferiblemente selección de los scFv que tienen al menos una propiedad funcional mejorada. En una realización preferida, la biblioteca se examina usando un sistema de control de calidad (sistema QC) en levadura (descrito en mayor detalle anteriormente), que permite la selección de regiones de marco de scFv que tienen una estabilidad y/o solubilidad potenciadas en un entorno reductor.

35 Se han descrito en la técnica otras tecnologías de selección adecuadas para examinar bibliotecas de scFv, incluyendo pero sin limitarse a tecnologías de presentación tales como presentación en fago, presentación en ribosoma y presentación en levadura (Jung *et al.* (1999) J. Mol. Biol. 294: 163- 180; Wu *et al.* (1999) J. Mol. Biol. 294: 151 - 162; Schier *et al.* (1996) J. Mol. Biol. 255: 28-43).

40 En una realización, se sustituye una posición de aminoácido identificada para la mutación por un residuo de aminoácido que es el enriquecido más significativamente en la posición correspondiente dentro de las secuencias de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos. En otra realización, la posición correspondiente dentro de las secuencias de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos está conservada con residuos de aminoácido hidrófobos y la posición de aminoácido identificada para la mutación dentro del scFv se sustituye por un residuo de aminoácido hidrófobo que es el enriquecido más significativamente en la posición correspondiente dentro de las secuencias de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos. En aún otra realización, la posición correspondiente dentro de las secuencias de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos está conservada con residuos de aminoácido hidrófilos y la posición de aminoácido identificada para la mutación dentro del scFv se sustituye por un residuo de aminoácido hidrófilo que es el enriquecido más significativamente en la posición correspondiente dentro de las secuencias de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos. En aún otra realización, la posición correspondiente dentro de las secuencias de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos está conservada con residuos de aminoácido pueden formar un enlace de hidrógeno y la posición de aminoácido identificada para la mutación dentro del scFv se sustituye por un residuo de aminoácido que puede formar un enlace de hidrógeno que es el enriquecido más significativamente en la posición correspondiente dentro de las secuencias de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos. En todavía otra realización, la posición correspondiente dentro de las secuencias de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos

está conservada con residuos de aminoácido que tienen propensión a formar una lámina  $\beta$  y la posición de aminoácido identificada para la mutación dentro del scFv se sustituye por un residuo de aminoácido que tiene propensión a formar una lámina  $\beta$  que es el enriquecido más significativamente en la posición correspondiente dentro de las secuencias de aminoácidos de  $V_H$  o  $V_L$  de anticuerpo de la base de datos.

- 5 En una realización, se selecciona la mejor sustitución que minimiza la energía libre total como la mutación que va a realizarse en la(s) posición/posiciones de aminoácido de interés. La mejor sustitución que minimiza la energía libre total puede determinarse usando la ley de Boltzmann. La fórmula para la ley de Boltzmann es  $\Delta\Delta G_{th} = RT \ln(f_{parental}/f_{consenso})$ .

10 El papel desempeñado por mutaciones potencialmente estabilizantes puede determinarse además mediante el examen, por ejemplo, de interacciones locales y no locales, residuos canónicos, superficies de contacto, grado de exposición y propensión a giros  $\beta$ . Pueden aplicarse métodos de modelado molecular conocidos en la técnica, por ejemplo, en el examen adicional del papel desempeñado por mutaciones potencialmente estabilizantes. También pueden usarse métodos de modelado molecular para seleccionar sustituciones de aminoácidos “de ajuste óptimo” si está considerándose un panel de posibles sustituciones.

15 Dependiendo de la posición de aminoácido particular, puede justificarse un análisis adicional. Por ejemplo, pueden estar implicados residuos en la interacción entre la cadena pesada y la ligera o pueden interaccionar con otros residuos a través de puentes salinos o enlaces de H. En estos casos, podría requerirse un especial análisis. En otra realización de presente invención, un residuo potencialmente problemático para la estabilidad puede cambiarse a uno que sea compatible con su homólogo en un par covariante. Alternativamente, el residuo homólogo puede mutarse para que sea compatible con el aminoácido identificado inicialmente como problemático.

#### Optimización de la solubilidad

25 Residuos potencialmente problemáticos para la solubilidad en un anticuerpo scFv incluyen aminoácidos hidrófobos que están expuestos al disolvente en un scFv, pero que en el contexto de un anticuerpo de longitud completa estarían incluidos en la superficie de contacto entre dominios variables y constantes. En un scFv modificado mediante ingeniería genética, que carece de los dominios constantes, residuos hidrófobos que participaban en las interacciones entre los dominios variables y constantes quedan expuestos al disolvente (véase por ejemplo, Nieba *et al* (1997) *Protein Eng.* 10: 435-44). Estos residuos en la superficie del scFv tienden a provocar agregación y, por tanto, problemas de solubilidad.

30 Se han descrito varias estrategias para reemplazar aminoácidos hidrófobos que están expuestos al disolvente en anticuerpos scFv. Tal como conocen bien los expertos en la técnica, la modificación de residuos en determinadas posiciones afecta a las propiedades biofísicas de los anticuerpos como estabilidad, solubilidad y afinidad. En muchos casos, estas propiedades están interrelacionadas, lo que significa que el cambio de un solo aminoácido puede afectar a varias de las propiedades mencionadas anteriormente. Por tanto, la mutación de residuos hidrófobos expuestos al disolvente de manera no conservativa puede provocar una disminución de la estabilidad y/o pérdida de afinidad por su antígeno.

35 Otros enfoques intentan resolver los problemas de solubilidad mediante el uso exhaustivo de tecnologías de presentación de proteínas y/o esfuerzos de examen. Sin embargo, tales métodos llevan mucho tiempo, a menudo no pueden producir proteína soluble o dan como resultado una menor estabilidad o reducción de la afinidad del anticuerpo. En la presente invención, se dan a conocer métodos para diseñar mutaciones de residuos hidrófobos expuestos al disolvente a residuos con una mayor hidrofiliidad usando un análisis basado en secuencia. Los residuos potencialmente problemáticos pueden reemplazarse eligiendo el aminoácido hidrófilo representado con la mayor frecuencia en posiciones definidas. Si se encuentra que un residuo interacciona con cualquier otro residuo en el anticuerpo, el residuo potencialmente problemático puede mutarse, no al residuo más frecuente sino a uno que es compatible con el segundo aminoácido del par covariante. Alternativamente, un segundo aminoácido del par covariante también puede mutarse para restaurar la combinación de aminoácidos. Además, el porcentaje de similitud entre secuencias puede tenerse en cuenta para ayudar a encontrar una combinación óptima de dos aminoácidos interrelacionados.

40 Se identifican aminoácidos hidrófobos en la superficie del scFv usando varios enfoques, incluyendo pero sin limitarse a enfoques basados en la exposición al disolvente, información experimental e información de secuencia, así como modelado molecular.

45 En una realización de esta invención, se mejora la solubilidad reemplazando residuos hidrófobos expuestos en la superficie del anticuerpo scFv por los residuos hidrófilos más frecuentes presentes en estas posiciones en bases de datos. Este fundamento se basa en el hecho de que es probable que no sean problemáticos los residuos que aparecen con frecuencia. Tal como apreciarán los expertos en la técnica, las sustituciones conservativas tienen habitualmente un pequeño efecto en la desestabilización de la molécula, mientras que las sustituciones no conservativas podrían ser perjudiciales para las propiedades funcionales del scFv.

A veces, residuos hidrófobos en la superficie del anticuerpo pueden estar implicados en la interacción entre la

cadena pesada y la ligera o pueden interactuar con otros residuos a través de puentes salinos o enlaces de H. En estos casos, podría requerirse un especial análisis. En otra realización de la presente invención, los residuos potencialmente problemáticos para la solubilidad pueden mutarse no al residuo más frecuente sino a uno compatible con el par covariante o puede realizarse una segunda mutación para restaurar la combinación de aminoácidos covariantes.

5 Pueden usarse métodos adicionales para diseñar mutaciones en posiciones hidrófobas expuestas al disolvente. En otra realización de esta invención, se dan a conocer métodos que emplean la restricción de la base de datos a aquellas secuencias que revelan la mayor similitud con el scFv que va a modificarse (comentado adicionalmente en lo anterior). Aplicando una base de datos de referencia restringida de este tipo, se diseña la mutación de manera que se ajuste de forma óptima en el contexto de secuencia específico del anticuerpo que va a optimizarse. En esta situación, el residuo hidrófilo elegido puede estar de hecho escasamente representado en su posición respectiva en comparación con un mayor número de secuencias (es decir, la base de datos no restringida).

#### Optimización de la estabilidad

15 Los fragmentos de anticuerpo de cadena sencilla contienen un ligador peptídico que une covalentemente los dominios variables de cadena ligera y pesada. Aunque un ligador de este tipo es eficaz para evitar tener los dominios variables separados, y de ese modo hace que el scFv sea superior con respecto al fragmento Fv, el fragmento scFv todavía es más propenso al desplegamiento y la agregación en comparación con un fragmento Fab o con un anticuerpo de longitud completa, en ambos de los cuales V<sub>H</sub> y V<sub>L</sub> sólo están unidos indirectamente mediante los dominios constantes.

20 Otro problema común en los scFv es la exposición de residuos hidrófobos en la superficie del scFv que conduce a agregación intermolecular. Además, a veces mutaciones somáticas adquiridas durante el proceso de maduración por afinidad colocan residuos hidrófilos en el núcleo de la lámina β. Tales mutaciones pueden tolerarse bien en el formato de IgG o incluso en un fragmento Fab pero en un scFv esto contribuye claramente a la desestabilización y el consiguiente desplegamiento.

25 Los factores conocidos que contribuyen a la desestabilización de scFv incluyen: residuos hidrófobos expuestos al disolvente en la superficie del anticuerpo scFv; residuos hidrófilos poco habituales incluidos en el núcleo de la proteína, así como residuos hidrófilos presentes en la superficie de contacto hidrófoba entre las cadenas pesadas y las ligeras. Además, se sabe que las interacciones de empaquetamiento de van der Waals entre residuos apolares en el núcleo desempeñan un papel importante en la estabilidad de proteínas (Monsellier E. y Bedouelle H. (2006) J. Mol. Biol. 362:580-93, Tan *et al.* (1998) Biophys. J. 75:1473-82; Wörn A. y Plückthun A. (1998) Biochemistry 37:13120-7).

30 Por tanto, en una realización, para aumentar la estabilidad de anticuerpos scFv, se identifican aminoácidos poco habituales y/o desfavorables en posiciones muy conservadas y se mutan a aminoácidos que son más comunes en estas posiciones conservadas. Tales aminoácidos poco habituales y/o desfavorables incluyen: (i) residuos hidrófobos expuestos al disolvente en la superficie del anticuerpo scFv; (ii) residuos hidrófilos poco habituales incluidos en el núcleo de la proteína; (iii) residuos hidrófilos presentes en la superficie de contacto hidrófoba entre las cadenas pesadas y las ligeras; y (iv) residuos que perturban la superficie de contacto VH/VL mediante impedimento estérico.

35 Por tanto, en una realización de esta invención, puede lograrse un aumento de la estabilidad mediante la sustitución de aminoácidos que están escasamente representados en sus posiciones por aminoácidos que aparecen con la mayor frecuencia en estas posiciones. La frecuencia de aparición proporciona generalmente una indicación de aceptación biológica.

40 Pueden estar implicados residuos en la interacción entre la cadena pesada y la ligera o pueden interactuar con otros residuos a través de puentes salinos, enlaces de H o enlaces disulfuro. En estos casos, podría requerirse un especial análisis. En otra realización de presente invención, un residuo potencialmente problemático para la estabilidad puede cambiarse a uno que sea compatible con su homólogo en un par covariante. Alternativamente, el residuo homólogo puede mutarse para que sea compatible con el aminoácido identificado inicialmente como problemático.

45 Pueden usarse métodos adicionales para diseñar mutaciones para mejorar la estabilidad. En otra realización de esta invención, se dan a conocer métodos que emplean la restricción de la base de datos a aquellas secuencias que revelan la mayor similitud con el scFv que va a modificarse (comentado adicionalmente en lo anterior). Aplicando una base de datos de referencia restringida de este tipo, se diseña la mutación de manera que se ajuste de forma óptima en el contexto de secuencia específico del anticuerpo que va a optimizarse. La mutación usa el aminoácido más frecuente que está presente en el subconjunto seleccionado de secuencias de bases de datos. En esta situación, el residuo elegido puede estar de hecho escasamente representado en su posición respectiva en comparación con un mayor número de secuencias (es decir, la base de datos no restringida).

#### Composiciones y formulaciones de scFv

Otro aspecto de la invención se refiere a una composición de scFv preparada según los métodos de la invención. Por tanto, la invención proporciona composiciones de scFv modificado mediante ingeniería genética en las que se han introducido una o más mutaciones en la secuencia de aminoácidos, en comparación con un scFv original de interés, en las que la(s) mutación/mutaciones se ha(n) introducido en una(s) posición/posiciones que se predice que influye(n) en una o más propiedades biológicas, tales como estabilidad o solubilidad, en particular una o más posiciones de región de marco. En una realización, el scFv se ha modificado mediante ingeniería genética para que contenga una posición de aminoácido mutada (por ejemplo, una posición de región de marco). En otras realizaciones, el scFv se ha modificado mediante ingeniería genética para que contenga dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez o más de diez posiciones de aminoácido mutadas (por ejemplo, posiciones de región de marco).

Otro aspecto de la invención se refiere a formulaciones farmacéuticas de las composiciones de scFv de la invención. Tales formulaciones comprenden normalmente la composición de scFv y un portador farmacéuticamente aceptable. Tal como se usa en el presente documento, "portador farmacéuticamente aceptable" incluye todos y cada uno de los disolventes, medios de dispersión, recubrimientos, agentes antibacterianos y antifúngicos, agentes isotónicos y de retardo de la absorción, y similares que son fisiológicamente compatibles. Preferiblemente, el portador es adecuado para, por ejemplo, administración intravenosa, intramuscular, subcutánea, parenteral, espinal, epidérmica (por ejemplo, mediante inyección o infusión) o tópica (por ejemplo, a los ojos o la piel). Dependiendo de la vía de administración, el scFv puede recubrirse con un material para proteger el compuesto frente a la acción de ácidos y otras condiciones naturales que puede inactivar el compuesto.

Los compuestos farmacéuticos de la invención pueden incluir una o más sales farmacéuticamente aceptables. Una "sal farmacéuticamente aceptable" se refiere a una sal que conserva la actividad biológica deseada del compuesto original y no confiere ningún efecto toxicológico no deseado (véase por ejemplo, Berge, S. M., *et al.* (1977) J. Pharm. Sci. 66:1-19). Los ejemplos de tales sales incluyen sales de adición de ácido y sales de adición de base. Las sales de adición de ácido incluyen las derivadas de ácidos inorgánicos no tóxicos, tales como clorhídrico, nítrico, fosfórico, sulfúrico, bromhídrico, yodhídrico, fosforoso y similares, así como de ácidos orgánicos no tóxicos tales como ácidos mono y dicarboxílicos alifáticos, ácidos alcanóicos sustituidos con fenilo, ácidos hidroxialcanóicos, ácidos aromáticos, ácidos sulfónicos alifáticos y aromáticos y similares. Las sales de adición de base incluyen las derivadas de metales alcalinotérreos, tales como sodio, potasio, magnesio, calcio y similares, así como de aminas orgánicas no tóxicas, tales como N,N'-dibenciletildiamina, N-metilglucamina, cloroprocaína, colina, dietanolamina, etilendiamina, procaína y similares.

Una composición farmacéutica de la invención también puede incluir un antioxidante farmacéuticamente aceptable. Los ejemplos de antioxidantes farmacéuticamente aceptables incluyen: (1) antioxidantes solubles en agua, tales como ácido ascórbico, clorhidrato de cisteína, bisulfato de sodio, metabisulfito de sodio, sulfito de sodio y similares; (2) antioxidantes solubles en aceite, tales como palmitato de ascorbilo, hidroxianisol butilado (BHA), hidroxitolueno butilado (BHT), lecitina, galato de propilo, alfa-tocoferol, y similares; y (3) agentes quelantes de metales, tales como ácido cítrico, ácido etilendiaminatetraacético (EDTA), sorbitol, ácido tartárico, ácido fosfórico, y similares.

Los ejemplos de portadores acuosos y no acuosos adecuados que pueden emplearse en las composiciones farmacéuticas de la invención incluyen agua, etanol, polioles (tales como glicerol, propilenglicol, polietilenglicol, y similares), y mezclas adecuadas de los mismos, aceites vegetales, tales como aceite de oliva, y ésteres orgánicos inyectables, tales como oleato de etilo. Puede mantenerse una fluidez apropiada, por ejemplo, mediante el uso de materiales de recubrimiento, tales como lecitina, mediante el mantenimiento del tamaño de partícula requerido en el caso de dispersiones, y mediante el uso de tensioactivos.

Estas composiciones también pueden contener adyuvantes tales como conservantes, agentes humectantes, agentes emulsionantes y agentes dispersantes. Puede garantizarse que se impide la presencia de microorganismos tanto mediante procedimientos de esterilización, citados anteriormente, como mediante la inclusión de diversos agentes antibacterianos y antifúngicos, por ejemplo, parabeno, clorobutanol, fenol, ácido sórbico, y similares. También puede ser deseable incluir agentes isotónicos, tales como azúcares, cloruro de sodio, y similares en las composiciones. Además, puede provocarse la absorción prolongada de la forma farmacéutica inyectable mediante la inclusión de agentes que retardan la absorción tales como monoestearato de aluminio y gelatina.

Los portadores farmacéuticamente aceptables incluyen disoluciones o dispersiones acuosas estériles y polvos estériles para la preparación extemporánea de disoluciones o dispersiones inyectables estériles. En la técnica se conoce el uso de tales medios y agentes para sustancias farmacéuticamente activas. Excepto en la medida en que cualquier medio o agente convencional sea incompatible con el compuesto activo, se contempla el uso de los mismos en las composiciones farmacéuticas de la invención. También pueden incorporarse compuestos activos complementarios en las composiciones.

Normalmente, las composiciones terapéuticas deben ser estériles y estables en las condiciones de fabricación y almacenamiento. La composición puede formularse como una disolución, microemulsión, liposoma, u otra estructura ordenada adecuada para una alta concentración de fármaco. El portador puede ser un disolvente o medio de dispersión que contiene, por ejemplo, agua, etanol, polioliol (por ejemplo, glicerol, propilenglicol y polietilenglicol líquido, y similares), y mezclas adecuadas de los mismos. Puede mantenerse la fluidez apropiada, por ejemplo,

mediante el uso de un recubrimiento tal como lecitina, mediante el mantenimiento del tamaño de partícula requerido en el caso de una dispersión y mediante el uso de tensioactivos. En muchos casos, será preferible incluir agentes isotónicos, por ejemplo, azúcares, polialcoholes tales como manitol, sorbitol, o cloruro de sodio en la composición. Puede provocarse la absorción prolongada de las composiciones inyectables mediante la inclusión en la composición de un agente que retarda la absorción, por ejemplo, sales de monoestearato y gelatina.

Pueden prepararse disoluciones inyectables estériles mediante la incorporación del compuesto activo en la cantidad requerida en un disolvente apropiado con uno o una combinación de componentes enumerados anteriormente, según se requiera, seguido por esterilización por microfiltración. Generalmente, se preparan dispersiones mediante la incorporación del compuesto activo en un vehículo estéril que contiene un medio de dispersión básico y los demás componentes requeridos de los enumerados anteriormente. En el caso de polvos estériles para la preparación de disoluciones inyectables estériles, los métodos de preparación preferidos son secado a vacío y secado por congelación (liofilización) que producen un polvo del principio activo más cualquier componente deseado adicional a partir de una disolución esterilizada previamente por filtración de los mismos.

La cantidad de principio activo que puede combinarse con un material portador para producir una forma de dosificación individual dependerá del sujeto que esté tratándose, y del modo de administración particular. La cantidad de principio activo que puede combinarse con un material portador para producir una forma de dosificación individual será generalmente aquella cantidad de la composición que produce un efecto terapéutico. Generalmente, con respecto al cien por ciento, esta cantidad oscilará entre aproximadamente el 0,01 por ciento y aproximadamente el noventa y nueve por ciento de principio activo, preferiblemente entre aproximadamente el 0,1 por ciento y aproximadamente el 70 por ciento, lo más preferiblemente entre aproximadamente el 1 por ciento y aproximadamente el 30 por ciento de principio activo en combinación con un portador farmacéuticamente aceptable.

Se ajustan regímenes de dosificación para proporcionar la respuesta óptima deseada (por ejemplo, una respuesta terapéutica). Por ejemplo, puede administrarse un único bolo, pueden administrarse varias dosis divididas a lo largo del tiempo o puede reducirse o aumentarse de manera proporcional la dosis según se indique por las exigencias de la situación terapéutica. Resulta especialmente ventajoso formular composiciones parenterales en forma unitaria de dosificación para facilidad de administración y uniformidad de dosificación. Forma unitaria de dosificación tal como se usa en el presente documento se refiere a unidades diferenciadas físicamente adecuadas como dosificaciones unitarias para los sujetos que van a tratarse; cada unidad contiene una cantidad predeterminada de compuesto activo calculada para producir el efecto terapéutico deseado en asociación con el portador farmacéutico requerido. La especificación para las formas unitarias de dosificación de la invención está dictada por y depende directamente de (a) las características únicas del compuesto activo y el efecto terapéutico particular que ha de lograrse, y (b) las limitaciones inherentes en la técnica de la combinación de un compuesto activo de este tipo para el tratamiento de sensibilidad en individuos.

Modificación mediante ingeniería genética de agentes de unión inmunológica basándose en un enfoque de "consenso funcional"

Tal como se describe en detalle en los ejemplos 2 y 3, el enfoque de "consenso funcional" descrito en el presente documento, en el que se usa una base de datos de secuencias de scFv seleccionadas por sus propiedades mejoradas para analizar la variabilidad de posiciones de región de marco, permite la identificación de posiciones de aminoácido que son o bien más o bien menos tolerantes a la variabilidad en comparación con variabilidad en estas mismas posiciones en bases de datos de líneas germinales. Tal como se describe en detalle en los ejemplos 5 y 6, la retromutación de determinadas posiciones de aminoácido dentro de un scFv de muestra al residuo consenso de línea germinal tiene un efecto o bien neutro o bien perjudicial, mientras que variantes de scFv que contienen residuos de "consenso funcional" presentan un aumento de la estabilidad térmica en comparación con la molécula de scFv silvestre. Por consiguiente, las posiciones de región de marco identificadas en el presente documento a través del enfoque de consenso funcional son posiciones preferidas para la modificación de scFv para alterar, y preferiblemente mejorar, las propiedades funcionales del scFv. Tal como se expone en las tablas 3-8 en el ejemplo 3, se han identificado las siguientes posiciones de región de marco como posiciones preferidas para la modificación en las secuencias de  $V_H$  o  $V_L$  indicadas (la numeración usada a continuación es el sistema de numeración de AHo; en las tablas 1 y 2 en el ejemplo 1 se exponen las tablas de conversión para convertir la numeración de AHo al sistema de numeración de Kabat):

VH3: posiciones de aminoácido 1, 6, 7, 89 y 103;

VH1a: posiciones de aminoácido 1, 6, 12, 13, 14, 19, 21, 90, 92, 95 y 98;

VH1b: posiciones de aminoácido 1, 10, 12, 13, 14, 20, 21, 45, 47, 50, 55, 77, 78, 82, 86, 87 y 107;

Vκ1: posiciones de aminoácido 1, 3, 4, 24, 47, 50, 57, 91 y 103;

Vκ3: 2, 3, 10, 12, 18, 20, 56, 74, 94, 101 y 103; y

Vλ1: 1, 2, 4, 7, 11, 14, 46, 53, 82, 92 y 103.

Por consiguiente, una o más de estas posiciones de aminoácido puede seleccionarse para la modificación mediante ingeniería genética en agentes de unión inmunológica, tales como moléculas de scFv, para producir de ese modo formas variantes (es decir, mutadas) de los agentes de unión inmunológica. Por tanto, en aún otro aspecto, la invención proporciona un método de modificación mediante ingeniería genética de un agente de unión inmunológica, comprendiendo el método:

- 5
- a) seleccionar una o más posiciones de aminoácido dentro del agente de unión inmunológica para la mutación; y
- b) mutar la una o más posiciones de aminoácido seleccionadas para la mutación,
- en el que la una o más posiciones de aminoácido seleccionadas para la mutación se seleccionan del grupo que consiste en:
- 10 (i) posiciones de aminoácido 1, 6, 7, 89 y 103 de VH3 usando la numeración de AHo (posiciones de aminoácido 1, 6, 7, 78 y 89 usando la numeración de Kabat);
- (ii) posiciones de aminoácido 1, 6, 12, 13, 14, 19, 21, 90, 92, 95 y 98 de VH1a usando la numeración de AHo (posiciones de aminoácido 1, 6, 11, 12, 13, 18, 20, 79, 81, 82b y 84 usando la numeración de Kabat);
- 15 (iii) posiciones de aminoácido 1, 10, 12, 13, 14, 20, 21, 45, 47, 50, 55, 77, 78, 82, 86, 87 y 107 de VH1b usando la numeración de AHo (posiciones de aminoácido 1, 9, 11, 12, 13, 19, 20, 38, 40, 43, 48, 66, 67, 71, 75, 76 y 93 usando la numeración de Kabat);
- (iv) posiciones de aminoácido 1, 3, 4, 24, 47, 50, 57, 91 y 103 de Vκ1 usando la numeración de AHo (posiciones de aminoácido 1, 3, 4, 24, 39, 42, 49, 73 y 85 usando la numeración de Kabat);
- 20 (v) posiciones de aminoácido 2, 3, 10, 12, 18, 20, 56, 74, 94, 101 y 103 de Vκ3 usando la numeración de AHo (posiciones de aminoácido 2, 3, 10, 12, 18, 20, 48, 58, 76, 83 y 85 usando la numeración de Kabat); y
- (vi) posiciones de aminoácido 1, 2, 4, 7, 11, 14, 46, 53, 82, 92 y 103 de Vλ1 usando la numeración de AHo (posiciones de aminoácido 1, 2, 4, 7, 11, 14, 38, 45, 66, 74 y 85 usando la numeración de Kabat).

25 En una realización preferida, la una o más posiciones de aminoácido seleccionadas para la mutación se seleccionan del grupo que consiste en las posiciones de aminoácido 1, 6, 7, 89 y 103 de VH3 usando la numeración de AHo (posiciones de aminoácido 1, 6, 7, 78 y 89 usando la numeración de Kabat).

En otra realización preferida, la una o más posiciones de aminoácido seleccionadas para la mutación se seleccionan del grupo que consiste en las posiciones de aminoácido 1, 6, 12, 13, 14, 19, 21, 90, 92, 95 y 98 de VH1a usando la numeración de AHo (posiciones de aminoácido 1, 6, 11, 12, 13, 18, 20, 79, 81, 82b y 84 usando la numeración de Kabat).

30 En otra realización preferida, la una o más posiciones de aminoácido seleccionadas para la mutación se seleccionan del grupo que consiste en las posiciones de aminoácido 1, 10, 12, 13, 14, 20, 21, 45, 47, 50, 55, 77, 78, 82, 86, 87 y 107 de VH1b usando la numeración de AHo (posiciones de aminoácido 1, 9, 11, 12, 13, 19, 20, 38, 40, 43, 48, 66, 67, 71, 75, 76 y 93 usando la numeración de Kabat).

35 En otra realización preferida, la una o más posiciones de aminoácido seleccionadas para la mutación se seleccionan del grupo que consiste en las posiciones de aminoácido 1, 3, 4, 24, 47, 50, 57, 91 y 103 de Vκ1 usando la numeración de AHo (posiciones de aminoácido 1, 3, 4, 24, 39, 42, 49, 73 y 85 usando la numeración de Kabat).

40 En otra realización preferida, la una o más posiciones de aminoácido seleccionadas para la mutación se seleccionan del grupo que consiste en las posiciones de aminoácido 2, 3, 10, 12, 18, 20, 56, 74, 94, 101 y 103 de Vκ3 usando la numeración de AHo (posiciones de aminoácido 2, 3, 10, 12, 18, 20, 48, 58, 76, 83 y 85 usando la numeración de Kabat).

En otra realización preferida, una o más posiciones de aminoácido seleccionadas para la mutación se seleccionan del grupo que consiste en las posiciones de aminoácido 1, 2, 4, 7, 11, 14, 46, 53, 82, 92 y 103 de Vλ1 usando la numeración de AHo (posiciones de aminoácido 1, 2, 4, 7, 11, 14, 38, 45, 66, 74 y 85 usando la numeración de Kabat).

45 En diversas realizaciones, se seleccionan para la mutación una, dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez, once, doce, trece, catorce, quince, dieciséis, diecisiete, dieciocho, diecinueve, veinte o más de veinte de las posiciones de aminoácido descritas anteriormente.

50 Preferiblemente, el agente de unión inmunológica es un scFv, pero también pueden modificarse mediante ingeniería genética según el método otros agentes de unión inmunológica, tales como inmunoglobulinas de longitud completa, fragmentos Fab o cualquier otro tipo de agente de unión inmunológica descrito en el presente documento (por ejemplo, Dab o Nanobodies). La invención también abarca agentes de unión inmunológica preparados según el método de modificación mediante ingeniería genética, así como composiciones que comprenden los agentes de

unión inmunológica y un portador farmacéuticamente aceptable. Además, la presente invención se extiende además a agentes de unión inmunológica que se modificaron mediante ingeniería genética según uno cualquiera de los métodos dados a conocer en el presente documento y se producen a escala comercial.

5 En determinadas realizaciones a modo de ejemplo, un agente de unión inmunológica modificado mediante ingeniería genética según el método de la invención es un agente de unión inmunológica reconocido en la técnica que se une a un antígeno diana de importancia terapéutica o un agente de unión inmunológica que comprende regiones variables (regiones VL y/o VL) o una o más CDR (por ejemplo, CDRL1, CDRL2, CDRL3, CDRH1, CDRH2 y/o CDRH3) derivadas del agente de unión inmunológica de importancia terapéutica. Por ejemplo, agentes de unión inmunológica aprobados actualmente por la FDA u otras autoridades reguladoras pueden modificarse mediante ingeniería genética según los métodos de la invención. Más específicamente, estos agentes de unión inmunológica a modo de ejemplo incluyen, pero no se limitan a, anticuerpos anti-CD3 tales como muromonab (Orthoclone® OKT3; Johnson&Johnson, Brunswick, NJ; véase Arakawa *et al.* J. Biochem, (1996) 120:657-662; Kung y Goldstein *et al.*, Science (1979), 206: 347- 349), anticuerpos anti-CD 11 tales como efalizumab (Raptiva®, Genentech, South San Francisco, CA), anticuerpos anti-CD20 tales como rituximab (Rituxan®/Mabthera®, Genentech, South San Francisco, CA), tositumomab (Bexxar®, GlaxoSmithKline, Londres) o ibritumomab (Zevalin®, Biogen Idec, Cambridge MA) (véanse las patentes estadounidenses n.ºs 5.736.137; 6.455.043; y 6.682.734), anticuerpos anti-CD25 (IL2R $\alpha$ ) tales como daclizumab (Zenapax®, Roche, Basilea, Suiza) o basiliximab (Simulect®, Novartis, Basilea, Suiza), anticuerpos anti-CD33 tales como gemtuzumab (Mylotarg®, Wyeth, Madison, NJ, véanse las patentes estadounidenses n.ºs 5.714.350 y 6.350.861), anticuerpos anti-CD52 tales como alemtuzumab (Campath®, Millennium Pharmaceuticals, Cambridge, MA), anticuerpos anti-Gp11b/gIIa tales como abciximab (ReoPro®, Centocor, Horsham, PA), anticuerpos anti-TNF $\alpha$  tales como infliximab (Remicade®, Centocor, Horsham, PA) o adalimumab (Humira®, Abbott, Abbott Park, IL, véase la patente estadounidense n.º 6.258.562), anticuerpos anti-IgE tales como omalizumab (Xolair®, Genentech, South San Francisco, CA), anticuerpos anti-VSR tales como palivizumab (Synagis®, Medimmune, Gaithersburg, M, véase la patente estadounidense n.º 5.824.307), anticuerpos anti-EpCAM tales como edrecolomab (Panorex®, Centocor), anticuerpos anti-EGFR tales como cetuximab (Erbix®, Imclone Systems, Nueva York, NY) o panitumumab (Vectibix®, Amgen, Thousand Oaks, CA), anticuerpos anti-HER2/neu tales como trastuzumab (Herceptin®, Genentech), anticuerpos anti-integrina  $\alpha$ 4 tales como natalizumab (Tysabri®, BiogenIdec), anticuerpos anti-C5 tales como eculizumab (Soliris®, Alexion Pharmaceuticals, Cheshire, CT) y anticuerpos anti-VEGF tales como bevacizumab (Avastin®, Genentech, véase la patente estadounidense n.º 6.884.879) o ranibizumab (Lucentis®, Genentech).

A pesar de lo anterior, en diversas realizaciones, se excluye el uso de determinados agentes de unión inmunológica en los métodos de modificación mediante ingeniería genética de la invención y/o se excluyen de ser la composición de agente de unión inmunológica producida mediante los métodos de modificación mediante ingeniería genética. Por ejemplo, en diversas realizaciones, existe la condición de que el agente de unión inmunológica no sea ninguno de los anticuerpos scFv, o variantes de los mismos, tal como se dan a conocer en las publicaciones PCT WO2006/131013 y WO2008/006235, tales como ESBA105 o variantes del mismo que se dan a conocer en las publicaciones PCT WO2006/131013 y WO2008/006235.

En otras realizaciones diversas, si el agente de unión inmunológica que va a modificarse mediante ingeniería genética según los métodos descritos anteriormente es cualquiera de los anticuerpos scFv, o variantes de los mismos, dados a conocer en las publicaciones PCT WO2006/131013 o WO2008/006235, entonces puede existir la condición de que la lista de posibles posiciones de aminoácido que pueden seleccionarse para la sustitución según el método de modificación mediante ingeniería genética no incluya alguna o ninguna de las siguientes posiciones de aminoácido: posición de AHo 4 (Kabat 4) de V $\kappa$ 1 o V $\lambda$ 1; posición de AHo 101 (Kabat 83) de V $\kappa$ 3; posición de AHo 12 (Kabat 11) de VH1a o VH1b; posición de AHo 50 (Kabat 43) de VH1b; posición de AHo 77 (Kabat 66) para VH1b; posición de AHo 78 (Kabat 67) para VH1b; posición de AHo 82 (Kabat 71) para VH1b; posición de AHo 86 (Kabat 75) para VH1b; posición de AHo 87 (Kabat 76) para VH1b; posición de AHo 89 (Kabat 78) para VH3; posición de AHo 90 (Kabat 79) para VH1a; y/o posición de AHo 107 (Kabat 93) para VH1b.

En todavía otras realizaciones diversas, para cualquier agente de unión inmunológica que va a modificarse mediante ingeniería genética según los métodos descritos anteriormente, y/o cualquier agente de unión inmunológica producido según los métodos descritos anteriormente, puede existir la condición de que la lista de posibles posiciones de aminoácido que pueden seleccionarse para la sustitución según el método de modificación mediante ingeniería genética no incluya alguna o ninguna de las siguientes posiciones de aminoácido: posición de AHo 4 (Kabat 4) de V $\kappa$ 1 o V $\lambda$ 1; posición de AHo 101 (Kabat 83) de V $\kappa$ 3; posición de AHo 12 (Kabat 11) de VH1a o VH1b; posición de AHo 50 (Kabat 43) de VH1b; posición de AHo 77 (Kabat 66) para VH1b; posición de AHo 78 (Kabat 67) para VH1b; posición de AHo 82 (Kabat 71) para VH1b; posición de AHo 86 (Kabat 75) para VH1b; posición de AHo 87 (Kabat 76) para VH1b; posición de AHo 89 (Kabat 78) para VH3; posición de AHo 90 (Kabat 79) para VH1a; y/o posición de AHo 107 (Kabat 93) para VH1b.

#### Mutación de agentes de unión inmunológica en posiciones a modo de ejemplo y preferidas

Tal como se describe en detalle en el ejemplo 7, el enfoque de consenso funcional descrito en el presente documento se ha usado satisfactoriamente para identificar sustituciones de residuos de aminoácido particulares que

están enriquecidos en la base de datos de scFv ("QC") seleccionada. Por ejemplo, las tablas 13-18 en el ejemplo 7 enumeran sustituciones de aminoácidos a modo de ejemplo y preferidas en posiciones de aminoácido definidas dentro de regiones de marco de la familia de VH3, VH1a, VH1b, Vκ1, Vκ3 o Vλ1. Las sustituciones a modo de ejemplo incluyen el residuo consenso identificado a partir del análisis de las bases de datos de líneas germinales (IMGT y Vbase) y anticuerpos maduros (KDB), así como los residuos de aminoácido identificados como que están enriquecidos preferentemente en la base de datos de región de marco de scFv (QC) seleccionada. La sustitución más preferida identificada es el residuo que presenta el mayor enriquecimiento en esa posición en la base de datos de región de marco de scFv (QC) seleccionada.

Por consiguiente, la invención proporciona métodos de modificación mediante ingeniería genética en los que se introducen una o más sustituciones de aminoácidos especificadas en un agente de unión inmunológica, tal como un anticuerpo scFv. Tales sustituciones pueden llevarse a cabo usando métodos de biología molecular convencionales, tales como mutagénesis dirigida al sitio, mutagénesis mediada por PCR y similares.

En una realización, la invención proporciona un método de modificación mediante ingeniería genética de un agente de unión inmunológica, tal como un anticuerpo scFv, en el que se realizan una o más sustituciones de aminoácidos en una o más posiciones de aminoácido, en el que el residuo de aminoácido que se usa para la sustitución en el agente de unión inmunológica se selecciona de los residuos de aminoácido a modo de ejemplo y preferidos identificados en las tablas 13-18 en el presente documento. Por tanto, la invención proporciona un método de modificación mediante ingeniería genética de un agente de unión inmunológica, comprendiendo el agente de unión inmunológica (i) una región variable de cadena pesada, o fragmento de la misma, de la familia de VH3, VH1a o VH1b, comprendiendo la región variable de cadena pesada residuos de región de marco de V<sub>H</sub> o (ii) una región variable de cadena ligera, o fragmento de la misma, de la familia de V<sub>κ</sub> 1, V<sub>κ</sub>3 o V<sub>λ</sub>1, comprendiendo la región variable de cadena ligera residuos de región de marco de V<sub>L</sub>, comprendiendo el método:

A) seleccionar una o más posiciones de aminoácido dentro de los residuos de región de marco de V<sub>H</sub>, los residuos de región de marco de V<sub>L</sub> o los residuos de región de marco de V<sub>H</sub> y V<sub>L</sub> para la mutación; y

B) mutar la una o más posiciones de aminoácido seleccionadas para la mutación,

a) en el que si la una o más posiciones de aminoácido seleccionadas para la mutación son de una región variable de cadena pesada de la familia de VH3, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

(i) ácido glutámico (E) o glutamina (Q) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;

(ii) ácido glutámico (E) o glutamina (Q) en la posición de aminoácido 6 usando el sistema de numeración de AHo o de Kabat;

(iii) treonina (T), serina (S) o alanina (A) en la posición de aminoácido 7 usando el sistema de numeración de AHo o de Kabat;

(iv) alanina (A), valina (V), leucina (L) o fenilalanina (F) en la posición de aminoácido 89 usando el sistema de numeración de AHo (posición de aminoácido 78 usando el sistema de numeración de Kabat); y

(v) arginina (R), glutamina (Q), valina (V), isoleucina (I), leucina (L), metionina (M) o fenilalanina (F) en la posición de aminoácido 103 usando el sistema de numeración de AHo (posición de aminoácido 89 usando la numeración de Kabat);

b) en el que si la una o más posiciones de aminoácido seleccionadas para la mutación son de una región variable de cadena pesada de la familia de VH1a, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

(i) ácido glutámico (E) o glutamina (Q) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;

(ii) ácido glutámico (E) o glutamina (Q) en la posición de aminoácido 6 usando el sistema de numeración de AHo o de Kabat;

(iii) leucina (L) o valina (V) en la posición de aminoácido 12 usando el sistema de numeración de AHo (posición de aminoácido 11 usando el sistema de numeración de Kabat);

(iv) metionina (M) o lisina (K) en la posición de aminoácido 13 usando el sistema de numeración de AHo (posición de aminoácido 12 usando el sistema de numeración de Kabat);

(v) ácido glutámico (E), glutamina (Q) o lisina (K) en la posición de aminoácido 14 usando el sistema de numeración de AHo (posición de aminoácido 13 usando el sistema de numeración de Kabat);

- (vi) leucina (L) o valina (V) en la posición de aminoácido 19 usando el sistema de numeración de AHo (posición de aminoácido 18 usando el sistema de numeración de Kabat);
- (vii) isoleucina (I) o valina (V) en la posición de aminoácido 21 usando el sistema de numeración de AHo (posición de aminoácido 20 usando el sistema de numeración de Kabat);
- 5 (viii) fenilalanina (F), serina (S), histidina (H), ácido aspártico (D) o tirosina (Y) en la posición de aminoácido 90 usando el sistema de numeración de AHo (posición de aminoácido 79 usando el sistema de numeración de Kabat);
- (ix) ácido aspártico (D), glutamina (Q) o ácido glutámico (E) en la posición de aminoácido 92 usando el sistema de numeración de AHo (posición de aminoácido 81 usando el sistema de numeración de Kabat);
- 10 (x) glicina (G), asparagina (N), treonina (T) o serina (S) en la posición de aminoácido 95 usando el sistema de numeración de AHo (posición de aminoácido 82b usando el sistema de numeración de Kabat); y
- (xi) treonina (T), alanina (A), prolina (P), fenilalanina (F) o serina (S) en la posición de aminoácido 98 usando la numeración de AHo (posición de aminoácido 84 usando la numeración de Kabat);
- c) en el que si la una o más posiciones de aminoácido seleccionadas para la mutación son de una región variable de cadena pesada de la familia de VH1b, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:
- 15 (i) ácido glutámico (E) o glutamina (Q) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;
- (ii) alanina (A), treonina (T), prolina (P), valina (V) o ácido aspártico (D) en la posición de aminoácido 10 usando el sistema de numeración de AHo (posición de aminoácido 9 usando el sistema de numeración de Kabat);
- 20 (iii) leucina (L) o valina (V) en la posición de aminoácido 12 usando el sistema de numeración de AHo (posición de aminoácido 11 usando el sistema de numeración de Kabat);
- (iv) lisina (K), valina (V), arginina (R), glutamina (Q) o metionina (M) en la posición de aminoácido 13 usando el sistema de numeración de AHo (posición de aminoácido 12 usando el sistema de numeración de Kabat);
- 25 (v) ácido glutámico (E), lisina (K), arginina (R) o metionina (M) en la posición de aminoácido 14 usando el sistema de numeración de AHo (posición de aminoácido 13 usando el sistema de numeración de Kabat);
- (vi) arginina (R), treonina (T), lisina (K) o asparagina (N) en la posición de aminoácido 20 usando el sistema de numeración de AHo (posición de aminoácido 19 usando el sistema de numeración de Kabat);
- 30 (vii) isoleucina (I), fenilalanina (F), valina (V) o leucina (L) en la posición de aminoácido 21 usando el sistema de numeración de AHo (posición de aminoácido 20 usando el sistema de numeración de Kabat);
- (viii) arginina (R) o lisina (K) en la posición de aminoácido 45 usando el sistema de numeración de AHo (posición de aminoácido 38 usando el sistema de numeración de Kabat);
- (ix) treonina (T), prolina (P), valina (V), alanina (A) o arginina (R) en la posición de aminoácido 47 usando el sistema de numeración de AHo (posición de aminoácido 40 usando el sistema de numeración de Kabat);
- 35 (x) lisina (K), glutamina (Q), histidina (H) o ácido glutámico (E) en la posición de aminoácido 50 usando el sistema de numeración de AHo (posición de aminoácido 43 usando el sistema de numeración de Kabat);
- (xi) metionina (M) o isoleucina (I) en la posición de aminoácido 55 usando la numeración de AHo (posición de aminoácido 48 usando la numeración de Kabat);
- 40 (xii) lisina (K) o arginina (R) en la posición de aminoácido 77 usando la numeración de AHo (posición de aminoácido 66 usando la numeración de Kabat);
- (xiii) alanina (A), valina (V), leucina (L) o isoleucina (I) en la posición de aminoácido 78 usando el sistema de numeración de AHo (posición de aminoácido 67 usando el sistema de numeración de Kabat);
- (xiv) ácido glutámico (E), arginina (R), treonina (T) o alanina (A) en la posición de aminoácido 82 usando el sistema de numeración de AHo (posición de aminoácido 71 usando el sistema de numeración de Kabat);
- 45 (xv) treonina (T), serina (S), isoleucina (I) o leucina (L) en la posición de aminoácido 86 usando el sistema de numeración de AHo (posición de aminoácido 75 usando el sistema de numeración de Kabat);
- (xvi) ácido aspártico (D), serina (S), asparagina (N) o glicina (G) en la posición de aminoácido 87 usando el sistema de numeración de AHo (posición de aminoácido 76 usando el sistema de numeración de Kabat); y

(xvii) asparagina (N), serina (S) o alanina (A) en la posición de aminoácido 107 usando el sistema de numeración de AHO (posición de aminoácido 93 usando el sistema de numeración de Kabat);

d) en el que si la una o más posiciones de aminoácido seleccionadas para la mutación son de una región variable de cadena ligera de la familia de V<sub>κ</sub>1, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

5 (i) ácido aspártico (D), ácido glutámico (E) o isoleucina (I) en la posición de aminoácido 1 usando el sistema de numeración de AHO o de Kabat;

(ii) glutamina (Q), valina (V) o isoleucina (I) en la posición de aminoácido 3 usando el sistema de numeración de AHO o de Kabat;

10 (iii) valina (V), leucina (L), isoleucina (I) o metionina (M) en la posición de aminoácido 4 usando el sistema de numeración de AHO o de Kabat;

(iv) arginina (R) o glutamina (Q) en la posición de aminoácido 24 usando el sistema de numeración de AHO o de Kabat;

15 (v) lisina (K), arginina (R) o isoleucina (I) en la posición de aminoácido 47 usando el sistema de numeración de AHO (posición de aminoácido 39 usando el sistema de numeración de Kabat);

(vi) lisina (K), arginina (R), ácido glutámico (E) treonina (T), metionina (M) o glutamina (Q) en la posición de aminoácido 50 usando el sistema de numeración de AHO (posición de aminoácido 42 usando el sistema de numeración de Kabat);

20 (vii) histidina (H), serina (S), fenilalanina (F) o tirosina (Y) en la posición de aminoácido 57 usando el sistema de numeración de AHO (posición de aminoácido 49 usando el sistema de numeración de Kabat);

(viii) leucina (L) o fenilalanina (F) en la posición de aminoácido 91 usando el sistema de numeración de AHO (posición de aminoácido 73 usando el sistema de numeración de Kabat); y

(ix) treonina (T), valina (V), serina (S), glicina (G) o isoleucina (I) en la posición de aminoácido 103 usando el sistema de numeración de AHO (posición de aminoácido 85 usando el sistema de numeración de Kabat);

25 e) en el que si la una o más posiciones de aminoácido seleccionadas para la mutación son de una región variable de cadena ligera de la familia de V<sub>κ</sub>3, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

(i) isoleucina (I) o treonina (T) en la posición de aminoácido 2 usando el sistema de numeración de AHO o de Kabat;

30 (ii) valina (V) o treonina (T) en la posición de aminoácido 3 usando el sistema de numeración de AHO o de Kabat;

(iii) treonina (T) o isoleucina (I) en la posición de aminoácido 10 usando el sistema de numeración de AHO o de Kabat;

35 (iv) serina (S) o tirosina (Y) en la posición de aminoácido 12 usando el sistema de numeración de AHO o de Kabat;

(v) serina (S) o arginina (R) en la posición de aminoácido 18 usando el sistema de numeración de AHO o de Kabat;

(vi) treonina (T) o alanina (A) en la posición de aminoácido 20 usando el sistema de numeración de AHO o de Kabat;

40 (vii) isoleucina (I) o metionina (M) en la posición de aminoácido 56 usando el sistema de numeración de AHO (posición de aminoácido 48 usando el sistema de numeración de Kabat);

(viii) isoleucina (I), valina (V) o treonina (T) en la posición de aminoácido 74 usando el sistema de numeración de AHO (posición de aminoácido 58 usando el sistema de numeración de Kabat);

45 (ix) serina (S) o asparagina (N) en la posición de aminoácido 94 usando el sistema de numeración de AHO (posición de aminoácido 76 usando el sistema de numeración de Kabat);

(x) fenilalanina (F), tirosina (Y) o serina (S) en la posición de aminoácido 101 usando el sistema de numeración de AHO (posición de aminoácido 83 usando el sistema de numeración de Kabat); y

(xi) valina (V), leucina (L) o alanina (A) en la posición de aminoácido 103 usando la numeración de AHO

(posición de aminoácido 85 usando la numeración de Kabat); y

f) en el que si la una o más posiciones de aminoácido seleccionadas para la mutación son de una región variable de cadena ligera de la familia de  $V\lambda 1$ , la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

- 5 (i) leucina (L), glutamina (Q), serina (S) o ácido glutámico (E) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;
- (ii) serina (S), alanina (A), prolina (P), isoleucina (I) o tirosina (Y) en la posición de aminoácido 2 usando el sistema de numeración de AHo o de Kabat;
- 10 (iii) valina (V), metionina (M) o leucina (L) en la posición de aminoácido 4 usando el sistema de numeración de AHo o de Kabat;
- (iv) serina (S), ácido glutámico (E), prolina (P) en la posición de aminoácido 7 usando el sistema de numeración de AHo o de Kabat;
- (v) alanina (A) o valina (V) en la posición de aminoácido 11 usando el sistema de numeración de AHo o de Kabat;
- 15 (vi) treonina (T), serina (S) o alanina (A) en la posición de aminoácido 14 usando el sistema de numeración de AHo o de Kabat;
- (vii) histidina (H) o glutamina (Q) en la posición de aminoácido 46 usando el sistema de numeración de AHo (posición de aminoácido 38 usando el sistema de numeración de Kabat);
- 20 (viii) lisina (K), treonina (T), serina (S), asparagina (N), glutamina (Q) o prolina (P) en la posición de aminoácido 53 usando el sistema de numeración de AHo (posición de aminoácido 45 usando el sistema de numeración de Kabat);
- (ix) arginina (R), glutamina (Q) o lisina (K) en la posición de aminoácido 82 usando el sistema de numeración de AHo (posición de aminoácido 66 usando el sistema de numeración de Kabat);
- 25 (x) glicina (G), treonina (T), ácido aspártico (D), alanina (A) en la posición de aminoácido 92 usando el sistema de numeración de AHo (posición de aminoácido 74 usando el sistema de numeración de Kabat); y
- (xi) ácido aspártico (D), valina (V), treonina (T), histidina (H) o ácido glutámico (E) en la posición de aminoácido 103 usando la numeración de AHo (posición de aminoácido 85 usando la numeración de Kabat).

30 En una realización preferida, el agente de unión inmunológica es un anticuerpo scFv. En otras realizaciones, el agente de unión inmunológica es, por ejemplo, una inmunoglobulina de longitud completa, Dab, Nanobody o un fragmento Fab.

La invención también abarca agentes de unión inmunológica preparados según el método descrito anteriormente. Preferiblemente, el agente de unión inmunológica es un anticuerpo scFv. En otras realizaciones, el agente de unión inmunológica es, por ejemplo, una inmunoglobulina de longitud completa, Dab, Nanobody o un fragmento Fab. La invención también abarca composiciones farmacéuticas que comprenden el/los agente(s) de unión inmunológica mencionado(s) anteriormente y un portador farmacéuticamente aceptable.

35 En otra realización, la invención proporciona un método de modificación mediante ingeniería genética de un agente de unión inmunológica, tal como un anticuerpo scFv, en el que se realizan una o más sustituciones de aminoácidos en una o más posiciones de aminoácido, en el que el residuo de aminoácido que se usa para la sustitución en el agente de unión inmunológica se selecciona de los residuos de aminoácido a modo de ejemplo y preferidos identificados en las tablas 13-18 en el presente documento, pero que no incluyen el residuo de aminoácido consenso identificado a partir del análisis de las bases de datos de líneas germinales (IMGT y Vbase) y anticuerpos maduros (KDB). Es decir, las sustituciones se seleccionan de los residuos de aminoácido que presentan enriquecimiento en la base de datos de scFV (QC) seleccionada. Por tanto, en esta realización, la invención proporciona un método de modificación mediante ingeniería genética de un agente de unión inmunológica, comprendiendo el agente de unión

40 inmunológica (i) una región variable de cadena pesada, o fragmento de la misma, de la familia de VH3, VH1a o VH1b, comprendiendo la región variable de cadena pesada residuos de región de marco de  $V_H$  o (ii) una región variable de cadena ligera, o fragmento de la misma, de la familia de  $V_{\kappa 1}$ ,  $V_{\kappa 3}$  o  $V\lambda 1$ , comprendiendo la región variable de cadena ligera  $V_L$  residuos de región de marco, comprendiendo el método:

45 A) seleccionar una o más posiciones de aminoácido dentro de los residuos de región de marco de  $V_H$ , los residuos de región de marco de  $V_L$  o los residuos de región de marco de  $V_H$  y  $V_L$  para la mutación; y

B) mutar la una o más posiciones de aminoácido seleccionadas para la mutación,

a) en el que si la una o más posiciones de aminoácido seleccionadas para la mutación son de una región variable de cadena pesada de la familia de VH3, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

- (i) glutamina (Q) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;
- 5 (ii) glutamina (Q) en la posición de aminoácido 6 usando el sistema de numeración de AHo o de Kabat;
- (iii) treonina (T) o alanina (A) en la posición de aminoácido 7 usando el sistema de numeración de AHo o de Kabat;
- (iv) alanina (A), valina (V) o fenilalanina (F) en la posición de aminoácido 89 usando el sistema de numeración de AHo (posición de aminoácido 78 usando el sistema de numeración de Kabat); y
- 10 (v) arginina (R), glutamina (Q), isoleucina (I), leucina (L), metionina (M) o fenilalanina (F) en la posición de aminoácido 103 usando el sistema de numeración de AHo (posición de aminoácido 89 usando la numeración de Kabat);

b) en el que si la una o más posiciones de aminoácido seleccionadas para la mutación son de una región variable de cadena pesada de la familia de VH1a, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

- (i) ácido glutámico (E) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;
- (ii) ácido glutámico (E) en la posición de aminoácido 6 usando el sistema de numeración de AHo o de Kabat;
- (iii) leucina (L) en la posición de aminoácido 12 usando el sistema de numeración de AHo (posición de aminoácido 11 usando el sistema de numeración de Kabat);
- 20 (iv) metionina (M) en la posición de aminoácido 13 usando el sistema de numeración de AHo (posición de aminoácido 12 usando el sistema de numeración de Kabat);
- (v) ácido glutámico (E) o glutamina (Q) en la posición de aminoácido 14 usando el sistema de numeración de AHo (posición de aminoácido 13 usando el sistema de numeración de Kabat);
- 25 (vi) leucina (L) en la posición de aminoácido 19 usando el sistema de numeración de AHo (posición de aminoácido 18 usando el sistema de numeración de Kabat);
- (vii) isoleucina (I) en la posición de aminoácido 21 usando el sistema de numeración de AHo (posición de aminoácido 20 usando el sistema de numeración de Kabat);
- (viii) fenilalanina (F), serina (S), histidina (H) o ácido aspártico (D) en la posición de aminoácido 90 usando el sistema de numeración de AHo (posición de aminoácido 79 usando el sistema de numeración de Kabat);
- 30 (ix) ácido aspártico (D) o glutamina (Q) en la posición de aminoácido 92 usando el sistema de numeración de AHo (posición de aminoácido 81 usando el sistema de numeración de Kabat);
- (x) glicina (G), asparagina (N) o treonina (T) en la posición de aminoácido 95 usando el sistema de numeración de AHo (posición de aminoácido 82b usando el sistema de numeración de Kabat); y
- 35 (xi) treonina (T), alanina (A), prolina (P) o fenilalanina (F) en la posición de aminoácido 98 usando la numeración de AHo (posición de aminoácido 84 usando la numeración de Kabat);

c) en el que si la una o más posiciones de aminoácido seleccionadas para la mutación son de una región variable de cadena pesada de la familia de VH1b, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

- (i) ácido glutámico (E) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;
- 40 (ii) treonina (T), prolina (P), valina (V) o ácido aspártico (D) en la posición de aminoácido 10 usando el sistema de numeración de AHo (posición de aminoácido 9 usando el sistema de numeración de Kabat);
- (iii) leucina (L) en la posición de aminoácido 12 usando el sistema de numeración de AHo (posición de aminoácido 11 usando el sistema de numeración de Kabat);
- 45 (iv) valina (V), arginina (R), glutamina (Q) o metionina (M) en la posición de aminoácido 13 usando el sistema de numeración de AHo (posición de aminoácido 12 usando el sistema de numeración de Kabat);
- (v) ácido glutámico (E), arginina (R) o metionina (M) en la posición de aminoácido 14 usando el sistema de numeración de AHo (posición de aminoácido 13 usando el sistema de numeración de Kabat);

- (vi) arginina (R), treonina (T) o asparagina (N) en la posición de aminoácido 20 usando el sistema de numeración de AHo (posición de aminoácido 19 usando el sistema de numeración de Kabat);
- (vii) isoleucina (I), fenilalanina (F) o leucina (L) en la posición de aminoácido 21 usando el sistema de numeración de AHo (posición de aminoácido 20 usando el sistema de numeración de Kabat);
- 5 (viii) lisina (K) en la posición de aminoácido 45 usando el sistema de numeración de AHo (posición de aminoácido 38 usando el sistema de numeración de Kabat);
- (ix) treonina (T), prolina (P), valina (V) o arginina (R) en la posición de aminoácido 47 usando el sistema de numeración de AHo (posición de aminoácido 40 usando el sistema de numeración de Kabat);
- 10 (x) lisina (K), histidina (H) o ácido glutámico (E) en la posición de aminoácido 50 usando el sistema de numeración de AHo (posición de aminoácido 43 usando el sistema de numeración de Kabat);
- (xi) isoleucina (I) en la posición de aminoácido 55 usando la numeración de AHo (posición de aminoácido 48 usando la numeración de Kabat);
- (xii) lisina (K) en la posición de aminoácido 77 usando la numeración de AHo (posición de aminoácido 66 usando la numeración de Kabat);
- 15 (xiii) alanina (A), leucina (L) o isoleucina (I) en la posición de aminoácido 78 usando el sistema de numeración de AHo (posición de aminoácido 67 usando el sistema de numeración de Kabat);
- (xiv) ácido glutámico (E), treonina (T) o alanina (A) en la posición de aminoácido 82 usando el sistema de numeración de AHo (posición de aminoácido 71 usando el sistema de numeración de Kabat);
- 20 (xv) treonina (T), serina (S) o leucina (L) en la posición de aminoácido 86 usando el sistema de numeración de AHo (posición de aminoácido 75 usando el sistema de numeración de Kabat);
- (xvi) ácido aspártico (D), asparagina (N) o glicina (G) en la posición de aminoácido 87 usando el sistema de numeración de AHo (posición de aminoácido 76 usando el sistema de numeración de Kabat); y
- (xvii) asparagina (N) o serina (S) en la posición de aminoácido 107 usando el sistema de numeración de AHo (posición de aminoácido 93 usando el sistema de numeración de Kabat);
- 25 d) en el que si la una o más posiciones de aminoácido seleccionadas para la mutación son de una región variable de cadena ligera de la familia de  $V_{\kappa 1}$ , la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:
- (i) ácido glutámico (E) o isoleucina (I) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;
- 30 (ii) valina (V) o isoleucina (I) en la posición de aminoácido 3 usando el sistema de numeración de AHo o de Kabat;
- (iii) valina (V), leucina (L) o isoleucina (I) en la posición de aminoácido 4 usando el sistema de numeración de AHo o de Kabat;
- (iv) glutamina (Q) en la posición de aminoácido 24 usando el sistema de numeración de AHo o de Kabat;
- 35 (v) arginina (R) o isoleucina (I) en la posición de aminoácido 47 usando el sistema de numeración de AHo (posición de aminoácido 39 usando el sistema de numeración de Kabat);
- (vi) lisina (K), ácido glutámico (E) treonina (T), metionina (M) o glutamina (Q) en la posición de aminoácido 50 usando el sistema de numeración de AHo (posición de aminoácido 42 usando el sistema de numeración de Kabat);
- 40 (vii) histidina (H), serina (S) o fenilalanina (F) en la posición de aminoácido 57 usando el sistema de numeración de AHo (posición de aminoácido 49 usando el sistema de numeración de Kabat);
- (viii) fenilalanina (F) en la posición de aminoácido 91 usando el sistema de numeración de AHo (posición de aminoácido 73 usando el sistema de numeración de Kabat); y
- 45 (ix) valina (V), serina (S), glicina (G), isoleucina (I) en la posición de aminoácido 103 usando el sistema de numeración de AHo (posición de aminoácido 85 usando el sistema de numeración de Kabat);
- e) en el que si la una o más posiciones de aminoácido seleccionadas para la mutación son de una región variable de cadena ligera de la familia de  $V_{\kappa 3}$ , la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

- (i) treonina (T) en la posición de aminoácido 2 usando el sistema de numeración de AHo o de Kabat;
- (ii) treonina (T) en la posición de aminoácido 3 usando el sistema de numeración de AHo o de Kabat;
- (iii) isoleucina (I) en la posición de aminoácido 10 usando el sistema de numeración de AHo o de Kabat;
- (iv) tirosina (Y) en la posición de aminoácido 12 usando el sistema de numeración de AHo o de Kabat;
- 5 (v) serina (S) en la posición de aminoácido 18 usando el sistema de numeración de AHo o de Kabat;
- (vi) alanina (A) en la posición de aminoácido 20 usando el sistema de numeración de AHo o de Kabat;
- (vii) metionina (M) en la posición de aminoácido 56 usando el sistema de numeración de AHo (posición de aminoácido 48 usando el sistema de numeración de Kabat);
- 10 (viii) valina (V) o treonina (T) en la posición de aminoácido 74 usando el sistema de numeración de AHo (posición de aminoácido 58 usando el sistema de numeración de Kabat);
- (ix) asparagina (N) en la posición de aminoácido 94 usando el sistema de numeración de AHo (posición de aminoácido 76 usando el sistema de numeración de Kabat);
- (x) tirosina (Y) o serina (S) en la posición de aminoácido 101 usando el sistema de numeración de AHo (posición de aminoácido 83 usando el sistema de numeración de Kabat); y
- 15 (xi) leucina (L) o alanina (A) en la posición de aminoácido 103 usando la numeración de AHo (posición de aminoácido 85 usando la numeración de Kabat); y

f) en el que si la una o más posiciones de aminoácido seleccionadas para la mutación son de una región variable de cadena ligera de la familia de V $\lambda$ 1, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

- 20 (i) leucina (L), serina (S) o ácido glutámico (E) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;
- (ii) alanina (A), prolina (P), isoleucina (I) o tirosina (Y) en la posición de aminoácido 2 usando el sistema de numeración de AHo o de Kabat;
- 25 (iii) valina (V) o metionina (M) en la posición de aminoácido 4 usando el sistema de numeración de AHo o de Kabat;
- (iv) serina (S) o ácido glutámico (E) en la posición de aminoácido 7 usando el sistema de numeración de AHo o de Kabat;
- (v) alanina (A) en la posición de aminoácido 11 usando el sistema de numeración de AHo o de Kabat;
- 30 (vi) treonina (T) o serina (S) en la posición de aminoácido 14 usando el sistema de numeración de AHo o de Kabat;
- (vii) histidina (H) en la posición de aminoácido 46 usando el sistema de numeración de AHo (posición de aminoácido 38 usando el sistema de numeración de Kabat);
- 35 (viii) treonina (T), serina (S), asparagina (N), glutamina (Q) o prolina (P) en la posición de aminoácido 53 usando el sistema de numeración de AHo (posición de aminoácido 45 usando el sistema de numeración de Kabat);
- (ix) arginina (R) o glutamina (Q) en la posición de aminoácido 82 usando el sistema de numeración de AHo (posición de aminoácido 66 usando el sistema de numeración de Kabat);
- (x) glicina (G), treonina (T) o ácido aspártico (D) en la posición de aminoácido 92 usando el sistema de numeración de AHo (posición de aminoácido 74 usando el sistema de numeración de Kabat); y
- 40 (xi) valina (V), treonina (T), histidina (H) o ácido glutámico (E) en la posición de aminoácido 103 usando la numeración de AHo (posición de aminoácido 85 usando la numeración de Kabat).

En una realización preferida, el agente de unión inmunológica es un anticuerpo scFv. En otras realizaciones, el agente de unión inmunológica es, por ejemplo, una inmunoglobulina de longitud completa, Dab, Nanobody o un fragmento Fab.

- 45 La invención también abarca agentes de unión inmunológica preparados según el método descrito anteriormente. Preferiblemente, el agente de unión inmunológica es un anticuerpo scFv. En otras realizaciones, el agente de unión

inmunológica es, por ejemplo, una inmunoglobulina de longitud completa, Dab, Nanobody o un fragmento Fab. La invención también abarca composiciones farmacéuticas que comprenden el/los agente(s) de unión inmunológica mencionado(s) anteriormente y un portador farmacéuticamente aceptable.

5 En aún otra realización, la invención proporciona un método de modificación mediante ingeniería genética de un agente de unión inmunológica, tal como un anticuerpo scFv, en el que se realizan una o más sustituciones de aminoácidos en una o más posiciones de aminoácido, en el que el residuo de aminoácido que se usa para la sustitución en el agente de unión inmunológica se selecciona de los residuos de aminoácido preferidos identificados en las tablas 13-18 en el presente documento (es decir, que no incluyen el residuo de aminoácido consenso identificado a partir del análisis de las bases de datos de líneas germinales (IMGT y Vbase) y anticuerpos maduros (KDB) o los residuos menos enriquecidos de la base de datos de scFv seleccionada). Es decir, las sustituciones se seleccionan sólo de aquellos residuos de aminoácido que presentan el mayor enriquecimiento en la base de datos de scFV (QC) seleccionada. Por tanto, en esta realización, la invención proporciona un método de modificación mediante ingeniería genética de un agente de unión inmunológica, comprendiendo el agente de unión inmunológica (i) una región variable de cadena pesada, o fragmento de la misma, de la familia de VH3, VH1a o VH1b, comprendiendo la región variable de cadena pesada residuos de región de marco de V<sub>H</sub> o (ii) una región variable de cadena ligera, o fragmento de la misma, de la familia de V<sub>κ</sub>1, V<sub>κ</sub>3 o V<sub>λ</sub>1, comprendiendo la región variable de cadena ligera residuos de región de marco de V<sub>L</sub>, comprendiendo el método:

A) seleccionar una o más posiciones de aminoácido dentro de los residuos de región de marco de V<sub>H</sub>, los residuos de región de marco de V<sub>L</sub> o los residuos de región de marco de V<sub>H</sub> y V<sub>L</sub> para la mutación; y

20 B) mutar la una o más posiciones de aminoácido seleccionadas para la mutación,

a) en el que si la una o más posiciones de aminoácido seleccionadas para la mutación son de una región variable de cadena pesada de la familia de VH3, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

(i) glutamina (Q) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;

25 (ii) glutamina (Q) en la posición de aminoácido 6 usando el sistema de numeración de AHo o de Kabat;

(iii) treonina (T) en la posición de aminoácido 7 usando el sistema de numeración de AHo o de Kabat;

(iv) valina (V) en la posición de aminoácido 89 usando el sistema de numeración de AHo (posición de aminoácido 78 usando el sistema de numeración de Kabat); y

30 (v) leucina (L) en la posición de aminoácido 103 usando el sistema de numeración de AHo (posición de aminoácido 89 usando la numeración de Kabat);

b) en el que si la una o más posiciones de aminoácido seleccionadas para la mutación son de una región variable de cadena pesada de la familia de VH1a, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

(i) ácido glutámico (E) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;

35 (ii) ácido glutámico (E) en la posición de aminoácido 6 usando el sistema de numeración de AHo o de Kabat;

(iii) leucina (L) en la posición de aminoácido 12 usando el sistema de numeración de AHo (posición de aminoácido 11 usando el sistema de numeración de Kabat);

(iv) metionina (M) en la posición de aminoácido 13 usando el sistema de numeración de AHo (posición de aminoácido 12 usando el sistema de numeración de Kabat);

40 (v) ácido glutámico (E) en la posición de aminoácido 14 usando el sistema de numeración de AHo (posición de aminoácido 13 usando el sistema de numeración de Kabat);

(vi) leucina (L) en la posición de aminoácido 19 usando el sistema de numeración de AHo (posición de aminoácido 18 usando el sistema de numeración de Kabat);

45 (vii) isoleucina (I) en la posición de aminoácido 21 usando el sistema de numeración de AHo (posición de aminoácido 20 usando el sistema de numeración de Kabat);

(viii) fenilalanina (F), serina (S), histidina (H) o ácido aspártico (D) en la posición de aminoácido 90 usando el sistema de numeración de AHo (posición de aminoácido 79 usando el sistema de numeración de Kabat);

(ix) ácido aspártico (D) en la posición de aminoácido 92 usando el sistema de numeración de AHo (posición de aminoácido 81 usando el sistema de numeración de Kabat);

(x) glicina (G) en la posición de aminoácido 95 usando el sistema de numeración de AHo (posición de aminoácido 82b usando el sistema de numeración de Kabat); y

(xi) fenilalanina (F) en la posición de aminoácido 98 usando la numeración de AHo (posición de aminoácido 84 usando la numeración de Kabat);

5 c) en el que si la una o más posiciones de aminoácido seleccionadas para la mutación son de una región variable de cadena pesada de la familia de VH1b, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

(i) ácido glutámico (E) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;

10 (ii) treonina (T), prolina (P), valina (V) o ácido aspártico (D) en la posición de aminoácido 10 usando el sistema de numeración de AHo (posición de aminoácido 9 usando el sistema de numeración de Kabat);

(iii) leucina (L) en la posición de aminoácido 12 usando el sistema de numeración de AHo (posición de aminoácido 11 usando el sistema de numeración de Kabat);

(iv) valina (V), arginina (R), glutamina (Q) o metionina (M) en la posición de aminoácido 13 usando el sistema de numeración de AHo (posición de aminoácido 12 usando el sistema de numeración de Kabat);

15 (v) arginina (R) en la posición de aminoácido 14 usando el sistema de numeración de AHo (posición de aminoácido 13 usando el sistema de numeración de Kabat);

(vi) asparagina (N) en la posición de aminoácido 20 usando el sistema de numeración de AHo (posición de aminoácido 19 usando el sistema de numeración de Kabat);

20 (vii) leucina (L) en la posición de aminoácido 21 usando el sistema de numeración de AHo (posición de aminoácido 20 usando el sistema de numeración de Kabat);

(viii) lisina (K) en la posición de aminoácido 45 usando el sistema de numeración de AHo (posición de aminoácido 38 usando el sistema de numeración de Kabat);

(ix) arginina (R) en la posición de aminoácido 47 usando el sistema de numeración de AHo (posición de aminoácido 40 usando el sistema de numeración de Kabat);

25 (x) lisina (K) en la posición de aminoácido 50 usando el sistema de numeración de AHo (posición de aminoácido 43 usando el sistema de numeración de Kabat);

(xi) isoleucina (I) en la posición de aminoácido 55 usando la numeración de AHo (posición de aminoácido 48 usando la numeración de Kabat);

30 (xii) lisina (K) en la posición de aminoácido 77 usando la numeración de AHo (posición de aminoácido 66 usando la numeración de Kabat);

(xiii) alanina (A) en la posición de aminoácido 78 usando el sistema de numeración de AHo (posición de aminoácido 67 usando el sistema de numeración de Kabat);

(xiv) ácido glutámico (E) en la posición de aminoácido 82 usando el sistema de numeración de AHo (posición de aminoácido 71 usando el sistema de numeración de Kabat);

35 (xv) treonina (T) en la posición de aminoácido 86 usando el sistema de numeración de AHo (posición de aminoácido 75 usando el sistema de numeración de Kabat);

(xvi) asparagina (N) en la posición de aminoácido 87 usando el sistema de numeración de AHo (posición de aminoácido 76 usando el sistema de numeración de Kabat); y

40 (xvii) asparagina (N) en la posición de aminoácido 107 usando el sistema de numeración de AHo (posición de aminoácido 93 usando el sistema de numeración de Kabat);

d) en el que si la una o más posiciones de aminoácido seleccionadas para la mutación son de una región variable de cadena ligera de la familia de V $\kappa$ 1, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

(i) ácido glutámico (E) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;

45 (ii) valina (V) en la posición de aminoácido 3 usando el sistema de numeración de AHo o de Kabat;

(iii) leucina (L) en la posición de aminoácido 4 usando el sistema de numeración de AHo o de Kabat;

- (iv) glutamina (Q) en la posición de aminoácido 24 usando el sistema de numeración de AHo o de Kabat;
  - (v) arginina (R) en la posición de aminoácido 47 usando el sistema de numeración de AHo (posición de aminoácido 39 usando el sistema de numeración de Kabat);
  - 5 (vi) lisina (K), ácido glutámico (E) treonina (T), metionina (M) o glutamina (Q) en la posición de aminoácido 50 usando el sistema de numeración de AHo (posición de aminoácido 42 usando el sistema de numeración de Kabat);
  - (vii) serina (S) en la posición de aminoácido 57 usando el sistema de numeración de AHo (posición de aminoácido 49 usando el sistema de numeración de Kabat);
  - 10 (viii) fenilalanina (F) en la posición de aminoácido 91 usando el sistema de numeración de AHo (posición de aminoácido 73 usando el sistema de numeración de Kabat); y
  - (ix) valina (V) en la posición de aminoácido 103 usando el sistema de numeración de AHo (posición de aminoácido 85 usando el sistema de numeración de Kabat);
- e) en el que si la una o más posiciones de aminoácido seleccionadas para la mutación son de una región variable de cadena ligera de la familia de  $V_{\kappa 3}$ , la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:
- 15 (i) treonina (T) en la posición de aminoácido 2 usando el sistema de numeración de AHo o de Kabat;
  - (ii) treonina (T) en la posición de aminoácido 3 usando el sistema de numeración de AHo o de Kabat;
  - (iii) isoleucina (I) en la posición de aminoácido 10 usando el sistema de numeración de AHo o de Kabat;
  - (iv) tirosina (Y) en la posición de aminoácido 12 usando el sistema de numeración de AHo o de Kabat;
  - 20 (v) serina (S) en la posición de aminoácido 18 usando el sistema de numeración de AHo o de Kabat;
  - (vi) alanina (A) en la posición de aminoácido 20 usando el sistema de numeración de AHo o de Kabat;
  - (vii) metionina (M) en la posición de aminoácido 56 usando el sistema de numeración de AHo (posición de aminoácido 48 usando el sistema de numeración de Kabat);
  - 25 (viii) treonina (T) en la posición de aminoácido 74 usando el sistema de numeración de AHo (posición de aminoácido 58 usando el sistema de numeración de Kabat);
  - (ix) asparagina (N) en la posición de aminoácido 94 usando el sistema de numeración de AHo (posición de aminoácido 76 usando el sistema de numeración de Kabat);
  - (x) serina (S) en la posición de aminoácido 101 usando el sistema de numeración de AHo (posición de aminoácido 83 usando el sistema de numeración de Kabat); y
  - 30 (xi) alanina (A) en la posición de aminoácido 103 usando la numeración de AHo (posición de aminoácido 85 usando la numeración de Kabat); y
- f) en el que si la una o más posiciones de aminoácido seleccionadas para la mutación son de una región variable de cadena ligera de la familia de  $V_{\lambda 1}$ , la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:
- 35 (i) leucina (L) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;
  - (ii) prolina (P) en la posición de aminoácido 2 usando el sistema de numeración de AHo o de Kabat;
  - (iii) valina (V) en la posición de aminoácido 4 usando el sistema de numeración de AHo o de Kabat;
  - (iv) serina (S) en la posición de aminoácido 7 usando el sistema de numeración de AHo o de Kabat;
  - (v) alanina (A) en la posición de aminoácido 11 usando el sistema de numeración de AHo o de Kabat;
  - 40 (vi) treonina (T) en la posición de aminoácido 14 usando el sistema de numeración de AHo o de Kabat;
  - (vii) histidina (H) en la posición de aminoácido 46 usando el sistema de numeración de AHo (posición de aminoácido 38 usando el sistema de numeración de Kabat);
  - (viii) treonina (T), serina (S), asparagina (N), glutamina (Q) o prolina (P) en la posición de aminoácido 53 usando el sistema de numeración de AHo (posición de aminoácido 45 usando el sistema de numeración de Kabat);
  - 45

(ix) arginina (R) en la posición de aminoácido 82 usando el sistema de numeración de AHo (posición de aminoácido 66 usando el sistema de numeración de Kabat);

(x) treonina (T) en la posición de aminoácido 92 usando el sistema de numeración de AHo (posición de aminoácido 74 usando el sistema de numeración de Kabat); y

5 (xi) valina (V) en la posición de aminoácido 103 usando la numeración de AHo (posición de aminoácido 85 usando la numeración de Kabat).

En una realización preferida, el agente de unión inmunológica es un anticuerpo scFv. En otras realizaciones, el agente de unión inmunológica es, por ejemplo, una inmunoglobulina de longitud completa, Dab, Nanobody o un fragmento Fab.

10 La invención también abarca agentes de unión inmunológica preparados según el método descrito anteriormente. Preferiblemente, el agente de unión inmunológica es un anticuerpo scFv. En otras realizaciones, el agente de unión inmunológica es, por ejemplo, una inmunoglobulina de longitud completa, Dab, Nanobody o un fragmento Fab. La invención también abarca composiciones farmacéuticas que comprenden el/los agente(s) de unión inmunológica mencionado(s) anteriormente y un portador farmacéuticamente aceptable.

15 Aunque los diversos métodos de modificación mediante ingeniería genética expuestos anteriormente en esta subsección proporcionan un listado de todas las sustituciones a modo de ejemplo y preferidas definidas en las tablas 13-18 en el presente documento para las familias de VH3, VH1a, VH1b, Vκ1, Vκ3 y Vλ1, respectivamente, debe entenderse que la invención abarca métodos en los que sólo se realizan una o unas pocas sustituciones de aminoácidos en una región variable seleccionada de VH3, VH1a, VH1b, Vκ1, Vκ3 y Vλ1, así como métodos en los  
20 que se realizan una, unas pocas o muchas sustituciones de aminoácidos en una o más regiones variables seleccionadas de una familia de VH3, VH1a, VH1b, Vκ1, Vκ3 o Vλ1, tal como en una región variable de cadena pesada seleccionada de una familia de VH3, VH1a o VH1b y una región variable de cadena ligera seleccionada de una familia de Vκ1, Vκ3 o Vλ1 en un agente de unión inmunológica que comprende una región variable de cadena pesada y una de cadena ligera (por ejemplo, un scFv). Es decir, todas y cada una de las posibles combinaciones de  
25 sustituciones seleccionadas de las sustituciones a modo de ejemplo y preferidas definidas en las tablas 13-18 pretenden estar abarcadas por los métodos de modificación mediante ingeniería genética, y los agentes de unión inmunológica resultantes preparados según esos métodos.

Por ejemplo, en diversas realizaciones, el método comprende realizar una, dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez o más de diez de las sustituciones de aminoácidos especificadas en una región variable de cadena  
30 pesada seleccionada de una región variable de la familia de VH3, VH1a o VH1b. En otras realizaciones diversas, el método comprende realizar una, dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez o más de diez de las sustituciones de aminoácidos especificadas en una región variable de cadena ligera seleccionada de una región variable de la familia de Vκ1, Vκ3 o Vλ1.

A pesar de lo anterior, en diversas realizaciones, se excluye el uso de determinados agentes de unión inmunológica en los métodos de modificación mediante ingeniería genética de la invención y/o se excluyen de ser la composición de agente de unión inmunológica producida mediante los métodos de modificación mediante ingeniería genética. Por  
35 ejemplo, en diversas realizaciones, existe la condición de que el agente de unión inmunológica no sea ninguno de los anticuerpos scFv, o variantes de los mismos, tal como se da a conocer en las publicaciones PCT WO2006/131013 y WO2008/006235, tal como ESBA105 o variantes del mismo que se dan a conocer en las  
40 publicaciones PCT WO2006/131013 y WO2008/006235, incorporándose expresamente el contenido de cada una de ellas en el presente documento como referencia.

En otras realizaciones diversas, si el agente de unión inmunológica que va a modificarse mediante ingeniería genética según los métodos descritos anteriormente es cualquiera de los anticuerpos scFv, o variantes de los  
45 mismos, dados a conocer en las publicaciones PCT WO2006/131013 o WO2008/006235, entonces puede existir la condición de que la lista de posibles posiciones de aminoácido que pueden seleccionarse para la sustitución según el método de modificación mediante ingeniería genética no incluya alguna o ninguna de las siguientes posiciones de aminoácido: posición de AHo 4 (Kabat 4) de Vκ1 o Vλ1; posición de AHo 101 (Kabat 83) de Vκ3; posición de AHo 12 (Kabat 11) de VH1a o VH1b; posición de AHo 50 (Kabat 43) de VH1b; posición de AHo 77 (Kabat 66) para VH1b; posición de AHo 78 (Kabat 67) para VH1b; posición de AHo 82 (Kabat 71) para VH1b; posición de AHo 86 (Kabat  
50 75) para VH1b; posición de AHo 87 (Kabat 76) para VH1b; posición de AHo 89 (Kabat 78) para VH3; posición de AHo 90 (Kabat 79) para VH1a; y/o posición de AHo 107 (Kabat 93) para VH1b.

En todavía otras realizaciones diversas, para cualquier agente de unión inmunológica que va a modificarse mediante ingeniería genética según los métodos descritos anteriormente, y/o cualquier agente de unión inmunológica producido según los métodos descritos anteriormente, puede existir la condición de que la lista de posibles  
55 posiciones de aminoácido que pueden seleccionarse para la sustitución según el método de modificación mediante ingeniería genética no incluya alguna o ninguna de las siguientes posiciones de aminoácido: posición de AHo 4 (Kabat 4) de Vκ1 o Vλ1; posición de AHo 101 (Kabat 83) de Vκ3; posición de AHo 12 (Kabat 11) de VH1a o VH1b; posición de AHo 50 (Kabat 43) de VH1b; posición de AHo 77 (Kabat 66) para VH1b; posición de AHo 78 (Kabat 67)

para VH1b; posición de AHO 82 (Kabat 71) para VH1b; posición de AHO 86 (Kabat 75) para VH1b; posición de AHO 87 (Kabat 76) para VH1b; posición de AHO 89 (Kabat 78) para VH3; posición de AHO 90 (Kabat 79) para VH1a; y/o posición de AHO 107 (Kabat 93) para VH1b.

#### Armazones de región de marco

5 Tal como se describe en detalle en el ejemplo 8, el enfoque de consenso funcional descrito en el presente documento se ha usado satisfactoriamente para diseñar secuencias de armazón de región de marco que incorporan las sustituciones de aminoácidos a modo de ejemplo y preferidas identificadas para posiciones de aminoácido particulares con familias de región variable. En estos armazones, las regiones CDR no se especifican; más bien, tales secuencias de armazón pueden usarse como “moldes” en las que pueden insertarse secuencias de CDR  
 10 (CDRL1, CDRL2, CDRL3, CDRH1, CDRH2 y/o CDRH3) para crear regiones variables que es probable que presenten propiedades de estabilidad y/o solubilidad deseables debido a las sustituciones de aminoácidos a modo de ejemplo o preferidas incorporadas en el armazón, basándose en las secuencias de scFv seleccionadas (seleccionadas basándose en sus propiedades de estabilidad y/o solubilidad deseables). Por ejemplo, una secuencia de armazón de región de marco de cadena pesada para la familia de VH1a se expone en la figura 9 (SEQ ID NO:1), una secuencia de armazón de región de marco de cadena pesada para la familia de VH1b se expone en la figura 10 (SEQ ID NO:2) una secuencia de armazón de región de marco de cadena pesada para la familia de VH3 se expone en la figura 11 (SEQ ID NO:3), una secuencia de armazón de región de marco de cadena ligera para la familia V $\kappa$ 1 se expone en la figura 12 (SEQ ID NO:4), una secuencia de armazón de región de marco de cadena ligera para la familia V $\kappa$ 3 se expone en la figura 13 (SEQ ID NO:5) y una secuencia de armazón de región de marco de cadena ligera para la familia V $\lambda$ 1 se expone en la figura 14 (SEQ ID NO:6).

Por consiguiente, en otro aspecto, la invención proporciona un método de modificación mediante ingeniería genética de un agente de unión inmunológica, comprendiendo el agente de unión inmunológica secuencias de CDR1, CDR2 y CDR3 de cadena pesada, comprendiendo el método insertar las secuencias de CDR1, CDR2 y CDR3 de cadena pesada en un armazón de región de marco de cadena pesada, comprendiendo el armazón de región de marco de cadena pesada una secuencia de aminoácidos tal como se muestra en la figura 9 (SEQ ID NO:1), la figura 10 (SEQ ID NO:2) o la figura 11 (SEQ ID NO:3). En una realización, el armazón de región de marco de cadena pesada comprende una secuencia de aminoácidos tal como se muestra en la figura 9 (SEQ ID NO:1). En otra realización, el armazón de región de marco de cadena pesada comprende una secuencia de aminoácidos tal como se muestra en la figura 10 (SEQ ID NO:2). En aún otra realización, el armazón de región de marco de cadena pesada comprende una secuencia de aminoácidos tal como se muestra en la figura 11 (SEQ ID NO:3).

Adicional o alternativamente, la invención proporciona un método de modificación mediante ingeniería genética de un agente de unión inmunológica, comprendiendo el agente de unión inmunológica secuencias de CDR1, CDR2 y CDR3 de cadena ligera, comprendiendo el método insertar las secuencias de CDR1, CDR2 y CDR3 de cadena ligera en un armazón de región de marco de cadena ligera, comprendiendo el armazón de región de marco de cadena ligera una secuencia de aminoácidos tal como se muestra en la figura 12 (SEQ ID NO:4), la figura 13 (SEQ ID NO:5) o la figura 14 (SEQ ID NO:6). En una realización, el armazón de región de marco de cadena ligera comprende una secuencia de aminoácidos tal como se muestra en la figura 12 (SEQ ID NO:4). En otra realización, el armazón de región de marco de cadena ligera comprende una secuencia de aminoácidos tal como se muestra en la figura 13 (SEQ ID NO:5). En aún otra realización, el armazón de región de marco de cadena ligera comprende una secuencia de aminoácidos tal como se muestra en la figura 14 (SEQ ID NO:6).

Preferiblemente, el agente de unión inmunológica modificado mediante ingeniería genética según el método es un anticuerpo scFv, aunque también pueden modificarse mediante ingeniería genética según el método otros agentes de unión inmunológica, tales como inmunoglobulinas de longitud completa y fragmentos Fab. En determinadas realizaciones a modo de ejemplo, una o más de las CDR (por ejemplo, CDRL1, CDRL2, CDRL3, CDRH1, CDRH2 y/o CDRH3) se derivan de cualquiera de los agentes de unión inmunológica de importancia terapéutica comentados anteriormente. Las CDR pueden insertarse en los armazones de región de marco usando técnicas de biología molecular convencionales.

La invención también abarca agentes de unión inmunológica modificados mediante ingeniería genética según el método descrito anteriormente usando armazones de región de marco. Preferiblemente, el agente de unión inmunológica es un anticuerpo scFv, aunque también están abarcados otros agentes de unión inmunológica, tales como inmunoglobulinas de longitud completa, Dab, Nanobodies y fragmentos Fab. También están abarcadas composiciones farmacéuticas que comprenden tales agentes de unión inmunológica y un portador farmacéuticamente aceptable.

En aún otro aspecto, la invención proporciona un armazón de región de marco de cadena pesada aislado que comprende una secuencia de aminoácidos tal como se muestra en la figura 9, la figura 10 o la figura 11. Tales armazones de región de marco de cadena pesada pueden prepararse usando técnicas de biología molecular convencionales.

A pesar de lo anterior, en diversas realizaciones, puede excluirse el uso de determinadas secuencias de armazón de región de marco en los métodos de modificación mediante ingeniería genética basados en armazón de la invención

y/o se excluyen de ser la composición de agente de unión inmunológica producida mediante los métodos de modificación mediante ingeniería genética basados en armazón. Por ejemplo, en diversas realizaciones, existe la condición de que la secuencia del armazón de región de marco no sea ninguna de las secuencias de región de marco de scFv dadas a conocer en la publicación PCT WO2001/048017, la publicación PCT WO2003/097697, la publicación de patente estadounidense n.º 20010024831 y/o la publicación de patente estadounidense US 20030096306, incorporándose expresamente el contenido de cada una de ellas en el presente documento como referencia.

En otras realizaciones diversas de los métodos de modificación mediante ingeniería genética basados en armazón descritos anteriormente, o agentes de unión inmunológica que resultan de los mismos, puede existir la condición de que determinadas posiciones de aminoácido mostradas en las figuras 9, 10 u 11 como variables (es decir, mostradas como "X", con la lista de posibles residuos de aminoácido para esa posición enumerados a continuación de la "X") pueden restringirse de ser variables. Por ejemplo, en determinadas realizaciones, existe la condición de que cualquiera o todas de las siguientes posiciones de aminoácido pueden limitarse a sólo el residuo de aminoácido que se enumera en primer lugar a continuación de la "X", o que se enumera en segundo lugar a continuación de la "X", o (cuando esté presente) que se enumera en tercer lugar a continuación de la "X", o (cuando esté presente) que se enumera en cuarto lugar a continuación de la "X" o (cuando esté presente) que se enumera en quinto lugar a continuación de la "X" o (cuando esté presente) que se enumera en sexto lugar a continuación de la "X": posición de AHo 12 (Kabat 11) de VH1a o VH1b; posición de AHo 50 (Kabat 43) de VH1b; posición de AHo 77 (Kabat 66) para VH1b; posición de AHo 78 (Kabat 67) para VH1b; posición de AHo 82 (Kabat 71) para VH1b; posición de AHo 86 (Kabat 75) para VH1b; posición de AHo 87 (Kabat 76) para VH1b; posición de AHo 89 (Kabat 78) para VH3; posición de AHo 90 (Kabat 79) para VH1a; y/o posición de AHo 107 (Kabat 93) para VH1b.

#### Otras realizaciones

Se entiende que la invención también incluye cualquiera de las metodologías, referencias y/o composiciones expuestas en los Apéndices (A-C) de la solicitud de patente estadounidense provisional con n.º de serie 60/905.365 y los Apéndices (A-1) de la solicitud de patente estadounidense provisional con n.º de serie 60/937.112, incluyendo, pero sin limitarse a, bases de datos identificadas, bioinformática, métodos de manipulación e interpretación de datos *in silico*, ensayos funcionales, secuencias preferidas, posiciones / alteraciones de residuo(s) preferidas, identificación y selección de región de marco, alteraciones de región de marco, alineación e integración de CDR y alteraciones/mutaciones preferidas.

Puede hallarse información adicional referente a estas metodologías y composiciones en los documentos U.S.S.N. 60/819.378; y 60/899.907, y la publicación PCT WO2008/006235, titulada "scFv Antibodies Which Pass Epithelial And/Or Endothelial Layers" presentada en julio de 2006 y el 6 de febrero de 2007 respectivamente; el documento WO06131013A2 titulado "Stable And Soluble Antibodies Inhibiting TNF $\alpha$ " presentado el 6 de junio de 2006; el documento EP1506236A2 titulado "Immunoglobulin Frameworks Which Demonstrate Enhanced Stability In The Intracellular Environment And Methods of Identifying Same" presentado el 21 de mayo de 2003; el documento EP1479694A2 titulado "Intrabodies ScFv with defined framework that is stable in a reducing environment" presentado el 18 de diciembre de 2000; el documento EP1242457B1 titulado "Intrabodies With Defined Framework That Is Stable In A Reducing Environment And Applications Thereof" presentado el 18 de diciembre de 2000; el documento WO03097697A2 titulado "Immunoglobulin Frameworks Which Demonstrate Enhanced Stability In The Intracellular Environment And Methods of Identifying Same" presentado el 21 de mayo de 2003; y el documento WO0148017A1 titulado "Intrabodies With Defined Framework That Is Stable In A Reducing Environment And Applications Thereof" presentado el 18 de diciembre de 2000; y Honegger *et al.*, J. Mol. Biol. 309:657-670 (2001).

Además, se entiende que la invención también incluye metodologías y composiciones adecuadas para el descubrimiento y/o la mejora de otros formatos de anticuerpo, por ejemplo, anticuerpos de longitud completa o fragmentos de los mismos, por ejemplo Fab, Dab, y similares. Por consiguiente, los principios y residuos identificados en el presente documento como adecuados para la selección o alteración para lograr propiedades biofísicas y/o terapéuticas deseadas que pueden aplicarse a una amplia gama de agentes de unión inmunológica. En una realización, anticuerpos terapéuticamente relevantes, por ejemplo, anticuerpos aprobados por la FDA, se mejoran mediante la modificación de una o más posiciones de residuo tal como se da a conocer en el presente documento.

Sin embargo, la invención no se limita a la modificación mediante ingeniería genética de agentes de unión inmunológica. Por ejemplo, un experto en la técnica reconocerá que los métodos de la invención pueden aplicarse a la modificación mediante ingeniería genética de otras moléculas de unión, distintas de inmunoglobulina, incluyendo, pero sin limitarse a, moléculas de unión a fibronectina tales como adnectinas (véase el documento WO01/64942 y las patentes estadounidenses n.ºs 6.673.901, 6.703.199, 7.078.490 y 7.119.171), Affibodies (véanse por ejemplo, las patentes estadounidenses 6.740.734 y 6.602.977 y en el documento WO00/63243), anticualinas (también conocidas como lipocalinas) (véanse los documentos WO99/16873 y WO05/019254), proteínas de dominio A (véanse los documentos WO02/088171 y WO04/044011) y proteínas con repeticiones de anquirina tales como darpinas o proteínas con repeticiones de leucina (véanse los documentos WO02/20565 y WO06/083275).

La presente divulgación se ilustra adicionalmente mediante los siguientes ejemplos, que no deben considerarse

como limitativos adicionalmente. El contenido de todas las figuras y todas las referencias, patentes y solicitudes de patente publicadas citadas en toda esta solicitud se incorporan expresamente en el presente documento como referencia en su totalidad.

EJEMPLO 1: Sistemas de numeración de posiciones de anticuerpo

- 5 En este ejemplo se proporcionan tablas de conversión para dos sistemas de numeración diferentes usados para identificar posiciones de residuo de aminoácido en regiones variables de cadena pesada y ligera de anticuerpo. El sistema de numeración de Kabat se describe adicionalmente en Kabat *et al.* (Kabat, E. A., *et al.* (1991) *Sequences of Proteins of Immunological Interest*, quinta edición, Departamento de Salud y Servicios Humanos de los EE.UU., publicación del NIH n.º 91-3242). El sistema de numeración de AHo se describe adicionalmente en Honegger, A. y Plückthun, A. (2001) *J Mol. Biol.* 309:657-670.
- 10

Numeración de región variable de cadena pesada

Tabla 1: Tabla de conversión para las posiciones de residuo en el dominio variable de cadena pesada

Kabat	AHo	Kabat	AHo	Kabat	AHo
1	1	44	51	87	101
2	2	45	52	88	102
3	3	46	53	89	103
4	4	47	54	90	104
5	5	48	55	91	105
6	6	49	56	92	106
7	7	50	57	93	107
*	8	51	58	94	108
8	9	52	59	95	109
9	10	52a	60	96	110
10	11	52b	61	97	111
11	12	52c	62	98	112
12	13	*	63	99	113
13	14	53	64	100	114
14	15	54	65	100a	115
15	16	55	66	100b	116
16	17	56	67	100c	117
17	18	57	68	100d	118
18	19	58	69	100e	119
19	20	59	70	100f	120
20	21	60	71	100g	121
21	22	61	72	100h	122
22	23	62	73	100i	123
23	24	63	74	*	124
24	25	64	75	*	125
25	26	65	76	*	126
26	27	66	77	*	127
*	28	67	78	*	128
27	29	68	79	*	129
28	30	69	80	*	130
29	31	70	81	*	131
30	32	71	82	*	132
31	33	72	83	*	133
32	34	73	84	*	134
33	35	74	85	*	135
34	36	75	86	*	136
35	37	76	87	101	137
35a	38	77	88	102	138

35b	39	78	89	103	139
*	40	79	90	104	140
*	41	80	91	105	141
*	42	81	92	106	142
36	43	82	93	107	143
37	44	82a	94	108	144
38	45	82b	95	109	145
39	46	82b	96	110	146
40	47	83	97	111	147
41	48	84	98	112	148
42	49	85	99	113	149
43	50	86	100		

Columna 1, posición de residuo en el sistema de numeración de Kabat. Columna 2, número correspondiente en el sistema de numeración de AHo para la posición indicada en la columna 1. Columna 3, posición de residuo en el sistema de numeración de Kabat. Columna 4, número correspondiente en el sistema de numeración de AHo para la posición indicada en la columna 3. Columna 5, posición de residuo en el sistema de numeración de Kabat. Columna 6, número correspondiente en el sistema de numeración de AHo para la posición indicada en la columna 5

Numeración de región variable de cadena ligera

5 Tabla 2: Tabla de conversión para las posiciones de residuo en el dominio variable de cadena ligera

Kabat	AHo	Kabat	AHo	Kabat	AHo
1	1	43	51	83	101
2	2	44	52	84	102
3	3	45	53	85	103
4	4	46	54	86	104
5	5	47	55	87	105
6	6	48	56	88	106
7	7	49	57	89	107
8	8	50	58	90	108
9	9	*	59	91	109
10	10	*	60	92	110
11	11	*	61	93	111
12	12	*	62	94	112
13	13	*	63	95	113
14	14	*	64	95a	114
15	15	*	65	95b	115
16	16	*	66	95c	116
17	17	51	67	95d	117
18	18	52	68	95e	118
19	19	53	69	95f	119
20	20	54	70	*	120
21	21	55	71	*	121
22	22	56	72	*	122
23	23	57	73	*	123
24	24	58	74	*	124
25	25	59	75	*	125
26	26	60	76	*	126

27	27	61	77	*	127
*	28	62	78	*	128
27a	29	63	79	*	129
27b	30	64	80	*	130
27c	31	65	81	*	131
27d	32	66	82	*	132
27e	33	67	83	*	133
27f	34	68	84	*	134
*	35	*	85	*	135
28	36	*	86	*	136
29	37	69	87	96	137
30	38	70	88	97	138
31	39	71	89	98	139
32	40	72	90	99	140
33	41	73	91	100	141
34	42	74	92	101	142
35	43	75	93	102	143
36	44	76	94	103	144
37	45	77	95	104	145
38	46	78	96	105	146
39	47	79	97	106	147
40	48	80	98	107	148
41	49	81	99	108	149
42	50	82	100		

Columna 1, posición de residuo en el sistema de numeración de Kabat. Columna 2, número correspondiente en el sistema de numeración de AHo para la posición indicada en la columna 1. Columna 3, posición de residuo en el sistema de numeración de Kabat. Columna 4, número correspondiente en el sistema de numeración de AHo para la posición indicada en la columna 3. Columna 5, posición de residuo en el sistema de numeración de Kabat. Columna 6, número correspondiente en el sistema de numeración de AHo para la posición indicada en la columna 5

**EJEMPLO 2:** Análisis basado en secuencia de secuencias de scFv

En este ejemplo se describe en detalle el análisis basado en secuencia de secuencias de scFv. En la figura 1 se muestra un diagrama de flujo que resume el procedimiento del análisis.

5 **Recopilación y alineación de secuencias de inmunoglobulinas humanas**

Se recopilaron secuencias de dominios variables de anticuerpos maduros humanos y líneas germinales de diferentes bases de datos y se introdujeron en una base de datos personalizada como secuencias de aminoácidos de una letra. Se alinearon las secuencias de anticuerpos usando una implementación en EXCEL del algoritmo de alineación de secuencias de Needleman-Wunsch (Needleman *et al.*, J Mol Biol., 48(3):443-53 (1970)). Entonces se subdividió la base de datos en cuatro alineamientos diferentes (según la fuente de datos original) para facilitar el análisis y la comparación posteriores, tal como sigue:

VBase: Secuencias de líneas germinales humanas

IMGT: Secuencias de líneas germinales humanas

Base de datos KDB: Anticuerpos maduros

15 **Base de datos QC: Base de datos interna de ESBATech que comprende regiones de marco de scFv seleccionado que se seleccionaron mediante examen de control de calidad**

El sistema de examen QC, y secuencias de región de marco de scFv que tienen propiedades funcionales deseables seleccionadas a partir del mismo, se describen adicionalmente, por ejemplo, en la publicación PCT WO2001/48017; la solicitud estadounidense n.º 20010024831; el documento US 20030096306; las patentes estadounidenses n.ºs 7.258.985 y 7.258.986; la publicación PCT WO2003/097697 y la solicitud estadounidense n.º 20060035320.

20 La introducción de huecos y la nomenclatura de posiciones de residuo se realizaron siguiendo el sistema de numeración de AHo para dominio variable de inmunoglobulina (Honegger, A. y Plückthun, A. (2001) J Mol. Biol. 309:657-670). Posteriormente, se identificaron regiones de marco y regiones CDR según Kabat *et al.* (Kabat, E. A., *et al.* (1991) Sequences of Proteins of Immunological Interest, quinta edición, Departamento de Salud y Servicios Humanos de los EE.UU., publicación del NIH n.º 91-3242). Se descartaron las secuencias en la base de datos KDB completas en menos del 70% o que contenían múltiples residuos indeterminados en las regiones de marco. También

se excluyeron las secuencias con una identidad de más del 95% con cualquier otra secuencia dentro de la base de datos para evitar el ruido aleatorio en el análisis.

#### Asignación de secuencias a subgrupos

5 Se clasificaron las secuencias de anticuerpos en familias diferenciadas mediante agrupación de los anticuerpos según métodos de clasificación basados en la homología de secuencia (Tomlinson, I.M. *et al* (1992) J. Mol. Biol. 227:776-798; Williams, S.C. y Winter, G. (1993) Eur. J. Immunol. 23:1456-1461; Cox, J.P. *et al*. (1994) Eur. J. Immunol. 24:827-836). Se restringió el porcentaje de homología con la región consenso de la familia al 70% de similitud. En casos en los que las secuencias mostraron conflictos entre dos o más familias de líneas germinales diferentes, o el porcentaje de homología fue inferior al 70% (con cualquier familia), se determinó el homólogo de línea germinal más próximo, se analizaron la longitud de las CDR, las clases canónicas y los residuos de subtipo caracterizadores en detalle para asignar correctamente la familia.

#### Análisis estadístico

15 Una vez que se definieron las agrupaciones de familia, se realizaron análisis estadísticos para coincidencias identificadas en el "examen de control de calidad ("QC")" (tal examen QC se describe en detalle en la publicación PCT WO2003/097697). Sólo fueron posibles análisis para las familias más representadas (VH3, VH1a, VH1b, Vκ1, Vκ3 y Vλ1) puesto que es necesario un número mínimo de secuencias para el análisis. Se calcularon las frecuencias de residuos,  $f_i(r)$ , para cada posición,  $i$ , mediante el número de veces que se observó el tipo de residuo particular dentro del conjunto de datos dividido entre el número total de secuencias. Se calculó la entropía posicional,  $N(i)$ , como medida de la variabilidad de la posición de cada residuo (Shenkin, P. S. *et al*. (1991) Proteins 11:297-313; Larson, S.M. y Davidson, A.R. (2000) Protein Sci. 9:2170-2180; Demarest, S.J. *et al*. (2004) J. Mol Biol. 335:41-48) usando el índice de Simpson que es una medida matemática de la diversidad en un sistema que proporciona más información sobre la composición de aminoácidos que simplemente la riqueza. Se calculó el grado de diversidad para cada posición,  $i$ , teniendo en cuenta el número de aminoácidos diferentes presentes, así como la abundancia relativa de cada residuo.

$$25 \quad D = \frac{\sum_{i=1}^r n(n-1)}{N(N-1)}$$

donde: D es el índice de Simpson, N es el número total de aminoácidos, r es el número de aminoácidos diferentes presentes en cada posición y n es el número de residuos de un tipo de aminoácido particular.

30 Se examinó la base de datos QC de las regiones de marco de Fv seleccionadas (seleccionadas mediante el examen QC) usando diferentes criterios para definir las características únicas. Se usaron los diferentes alineamientos en la base de datos de secuencias para definir el grado de variabilidad de posiciones de residuo dentro de las regiones de marco de Fv y para identificar posiciones tolerantes a la variación no comunes en la naturaleza que están presentes en las regiones de marco de Fv seleccionadas. Se definió como umbral una diferencia en las puntuaciones de entropía posicional igual o de más del 10%. Se seleccionaron posiciones adicionales si el residuo en una posición dada estaba ocupado por un aminoácido observado de manera poco frecuente en los demás alineamientos de secuencia, es decir, observado de manera poco frecuente en las bases de datos de líneas germinales (VBase y IMGT) y la base de datos KDB. Si se encontró que el comportamiento de un residuo era ciertamente diferente, (con baja representación o no representado en cualquiera de los alineamientos de secuencia), se definió la posición de residuo como única.

40 El fundamento subyacente a la identificación de características únicas de las secuencias de región de marco de Fv seleccionadas son las propiedades superiores probadas de las regiones de marco y el uso potencial de estos hallazgos para una formación de armazón mejorada. Se supuso que posiciones altamente conservadas en la naturaleza que muestran un determinado grado de variabilidad en las regiones de marco seleccionadas deben tolerar la mutagénesis al azar y presentar un aumento de la probabilidad de hallar aminoácidos alternativos superior al residuo nativo en un formato de scFv. Además, una preferencia pronunciada por un aminoácido poco común es una indicación de selección natural hacia determinado residuo. Basándose en estas dos directrices estadísticas, se eligieron diferentes residuos dentro de las cadenas pesada y ligera como o bien posiciones flotantes (tolerantes a la variabilidad) o bien sustituciones preferidas (residuos poco habituales).

#### EJEMPLO 3: Identificación de posiciones de residuo tolerantes a la variabilidad y poco habituales

50 Usando el enfoque de análisis de scFv basado en secuencia descrito anteriormente en el ejemplo 2, se analizaron tres familias de región variable de cadena pesada (VH3, VH1a y VH1b) y tres familias de región variable de cadena ligera (Vκ1, Vκ3 y Vλ1) para identificar posiciones de aminoácido tolerantes a la variabilidad. En particular, se determinó el grado de diversidad, calculado usando el índice de Simpson, para cada posición de aminoácido para secuencias dentro de cuatro bases de datos diferentes, Vbase, IMGT, KDB y QC (scFv seleccionados), tal como se describió anteriormente. Se identificaron las posiciones de residuo de aminoácido tolerantes a las variantes y poco habituales basándose en diferencias en los valores del índice de Simpson en aquellas posiciones para las bases de datos de líneas germinales IMGT y Vbase en comparación con la base de datos de scFv seleccionado QC.

Adicionalmente, para las posiciones de interés identificadas, se identificó el residuo consenso de línea germinal y se determinó la frecuencia de ese residuo consenso en las bases de datos QC y KDB.

5 A continuación en las tablas 3, 4 y 5, respectivamente, se muestran los resultados del análisis de variabilidad para las familias de región variable de cadena pesada VH3, VH1a y VH1b. Para cada tabla, las columnas son tal como sigue: columna 1: posición de residuo de aminoácido usando el sistema de numeración de AHO (puede obtenerse la conversión al sistema de numeración de Kabat usando la tabla de conversión expuesta como tabla 1 en el ejemplo 1); columnas 2 a 5: diversidad calculada para cada alineamiento de anticuerpo en la base de datos para la posición de residuo indicada en la columna 1; columna 6: residuo consenso de la familia de línea germinal correspondiente y KDB; columna 7: frecuencia relativa de residuos en la base de datos KDB para el residuo consenso en la columna 6; 10 y columna 8: frecuencia relativa de residuos en la base de datos de scFv seleccionado QC para el residuo consenso en la columna 6.

Tabla 3: Análisis de variabilidad de residuos y frecuencias correspondientes del aminoácido consenso identificado en la línea germinal para la familia de VH3.

Posición de residuo	Línea germinal de IMGT	Línea germinal de Vbase	scFv seleccionado de QC	Sec. de KDB	Residuo consenso	f(cons KDB)	f(cons QC)
1	0,68	0,65	0,50	0,53	E	66,67	53,57
6	1,00	1,00	0,57	0,86	E	92,56	68,97
7	1,00	0,91	0,65	0,93	S	96,33	77,59
89	0,86	0,83	0,55	0,71	L	84,06	70,18
103	0,73	0,76	0,38	0,76	V	86,85	55,36

15 Tabla 4: Análisis de variabilidad de residuos y frecuencias correspondientes del aminoácido consenso identificado en la línea germinal para la familia de VH1a.

Posición de residuo	Línea germinal de IMGT	Línea germinal de Vbase	scFv seleccionado de QC	Sec. de KDB	Residuo consenso	f(cons KDB)	f(cons QC)
1	0,82	0,83	0,62	0,77	Q	86,60	75,00
6	1,00	1,00	0,51	0,74	Q	84,31	58,30
12	1,00	1,00	0,72	0,93	V	96,29	83,30
13	1,00	1,00	0,72	0,86	K	92,59	83,30
14	1,00	1,00	0,60	0,93	K	96,29	75,00
19	1,00	1,00	0,72	1,00	V	100,00	83,30
21	0,83	0,83	0,72	0,96	V	98,14	83,30
90	1,00	1,00	0,47	0,89	Y	94,44	66,60
92	0,83	1,00	0,60	0,93	E	96,29	75,00
95	0,83	0,83	0,49	0,70	S	83,33	66,60
98	1,00	1,00	0,39	0,83	S	90,74	38,30

Tabla 5: Análisis de variabilidad de residuos y frecuencias correspondientes del aminoácido consenso identificado en la línea germinal para la familia de VH1b.

Posición de residuo	Línea germinal de IMGT	Línea germinal de Vbase	scFv seleccionado de QC	Sec. de KDB	Residuo consenso	f(cons KDB)	f(cons QC)
1	0,82	0,83	0,58	0,92	Q	95,65	70,59
10	0,82	0,83	0,52	0,73	A	85,00	70,59
12	1,00	1,00	0,64	0,86	V	92,59	76,47
13	1,00	1,00	0,52	0,86	K	92,59	70,59
14	1,00	1,00	0,54	0,88	K	93,83	70,59
20	1,00	1,00	0,61	0,86	K	92,59	76,47
21	0,83	0,83	0,47	0,84	V	91,36	64,71
45	0,70	0,83	0,64	0,90	R	95,06	76,47
47	0,83	1,00	0,31	0,95	A	97,53	47,06
50	0,70	0,70	0,48	0,76	Q	86,42	64,71
55	0,83	0,83	0,64	0,82	M	90,12	76,47
77	1,00	1,00	0,64	1,00	R	100,00	76,47
78	0,83	1,00	0,32	0,76	A	86,42	47,06
82	0,45	0,39	0,25	0,36	R	55,56	29,41
86	0,45	0,45	0,37	0,27	I	24,69	17,65
87	0,57	0,70	0,30	0,53	S	70,37	25,00
107	1,00	1,00	0,60	0,90	A	95,00	75,00

A continuación en las tablas 6, 7 y 8, respectivamente, se muestran los resultados del análisis de variabilidad para las familias de región variable de cadena ligera  $V_{\kappa 1}$ ,  $V_{\kappa 3}$  y  $V_{\lambda 1}$ . Para cada tabla, las columnas son tal como sigue: columna 1: residuo de posición de aminoácido usando el sistema de numeración de AHO (puede obtenerse la conversión al sistema de numeración de Kabat usando la tabla de conversión expuesta como tabla 1 en el ejemplo 1); columnas 2 a 5: diversidad calculada para cada alineamiento de anticuerpo en la base de datos para la posición de residuo indicada en la columna 1; columna 6: residuo consenso de la familia de línea germinal correspondiente y KDB; columna 7: frecuencia relativa de residuos en la base de datos KDB para el residuo consenso en la columna 6; y columna 8: frecuencia relativa de residuos en la base de datos de scFv seleccionado QC para el residuo consenso en la columna 6.

- 5
- 10 Tabla 6: Análisis de variabilidad de residuos y frecuencias correspondientes del aminoácido consenso identificado en la línea germinal para la familia de  $V_{\kappa 1}$ .

Posición de residuo	Línea germinal de IMGT	Línea germinal de Vbase	scFv seleccionado de QC	Sec. de KDB	Residuo consenso	f(cons KDB)	f(cons QC)
1	0,52	0,47	0,61	0,68	D	81,5	23,3
3	0,76	0,72	0,66	0,55	Q	72,0	18,6
4	0,65	0,73	0,57	0,62	M	76,0	23,3
24	0,69	0,72	0,64	0,74	R	85,3	76,7
47	1,00	1,00	0,69	0,88	K	94,0	81,4
50	1,00	1,00	0,60	0,79	R	89,0	76,7
57	1,00	1,00	0,58	0,79	Y	88,6	74,4
91	0,83	0,81	0,70	0,77	L	86,6	81,4
103	0,91	1,00	0,67	0,90	T	81,4	95,7

Tabla 7: Análisis de variabilidad de residuos y frecuencias correspondientes del aminoácido consenso identificado en la línea germinal para la familia de  $V_{\kappa 3}$ .

Posición de residuo	Línea germinal de IMGT	Línea germinal de Vbase	scFv seleccionado de QC	Sec. de KDB	Residuo consenso	f(cons KDB)	f(cons QC)
2	1,00	1,00	0,72	0,69	I	82,47	83,33
3	1,00	1,00	0,72	0,64	V	77,93	83,33
10	1,00	1,00	0,72	0,93	T	96,19	83,33
12	1,00	1,00	0,72	0,98	S	98,84	83,33
18	1,00	1,00	0,72	0,92	R	95,86	83,33
20	1,00	1,00	0,68	0,95	T	97,30	66,67
56	1,00	1,00	0,72	0,91	I	95,31	83,33
74	1,00	1,00	0,50	0,86	I	92,61	66,67
94	1,00	1,00	0,72	0,82	S	90,29	83,33
101	1,00	1,00	0,50	0,91	F	95,14	66,67
103	1,00	1,00	0,50	0,82	F	90,47	66,67

- 15 Tabla 8: Análisis de variabilidad de residuos y frecuencias correspondientes del aminoácido consenso identificado en la línea germinal para la familia de  $V_{\lambda 1}$ .

Posición de residuo	Línea germinal de IMGT	Línea germinal de Vbase	scFv seleccionado de QC	Sec. de KDB	Residuo consenso	f(cons KDB)	f(cons QC)
1	1,00	1,00	0,45	0,70	Q	81,10	62,50
2	1,00	1,00	0,27	0,73	S	85,13	37,50
4	1,00	1,00	0,60	0,85	L	92,00	75,00
7	1,00	1,00	0,77	0,99	P	99,32	87,50
11	0,59	0,52	0,53	0,51	V	59,88	37,50
14	0,59	0,52	0,49	0,51	A	59,95	31,25
46	1,00	1,00	0,70	0,80	Q	89,00	81,25
53	1,00	1,00	0,49	0,90	K	94,63	68,75
82	1,00	1,00	0,60	0,90	K	94,88	75,00
92	0,59	0,68	0,51	0,54	A	69,82	68,75
103	1,00	1,00	0,50	0,86	D	92,84	68,75

Tal como se expuso en las tablas 3-8 anteriormente, se encontró que un subconjunto de posiciones de residuo en las regiones de marco de scFv seleccionado del sistema QC estaban fuertemente sesgadas hacia determinados residuos no presentes o con escasa representación en las líneas germinales (VBase y IMGT) y en anticuerpos maduros (KDB), lo que sugirió que la estabilidad de scFv puede mejorarse de manera racional basándose en las

características únicas de las secuencias de región de marco seleccionadas en el sistema de examen en levadura de control de calidad.

**EJEMPLO 4:** Selección de residuos preferidos

5 Para seleccionar sustituciones de residuos de aminoácido preferidas (o, alternativamente, excluir residuos de aminoácido) en una posición de aminoácido particular que se sabe que mejora las propiedades funcionales (por ejemplo, estabilidad y/o solubilidad) de un scFv, se agruparon secuencias de VH y VL de la base de datos Kabat de secuencias de anticuerpos madurados según su subtipo de familia (por ejemplo, VH1b, VH3, etc.). Dentro de cada subfamilia de secuencias, se determinó la frecuencia de cada residuo de aminoácido en cada posición de aminoácido como porcentaje de todas las secuencias analizadas de un grupo de subtipos. Se hizo lo mismo para 10 todas las secuencias de la base de datos QC que consiste en anticuerpos que se preseleccionaron por una estabilidad y/o solubilidad potenciadas mediante el denominado sistema QC. Para cada subtipo, se compararon los porcentajes resultantes (frecuencias relativas) para cada residuo de aminoácido obtenido para las secuencias de Kabat y para las secuencias QC en cada posición correspondiente. En el caso de que la frecuencia relativa de un determinado residuo de aminoácido aumentase en la base de datos QC con relación a la base de datos Kabat, el residuo respectivo se consideró un residuo preferido en la posición dada para mejorar la estabilidad y/o solubilidad de un scFv. A la inversa, en el caso de que la frecuencia relativa de un determinado residuo de aminoácido disminuyera en la base de datos QC en comparación con la base de datos Kabat, el residuo respectivo se consideró 15 desfavorable en esa posición en el contexto de un formato de scFv.

20 La tabla 9 representa un análisis de la frecuencia de residuos a modo de ejemplo en la posición de aminoácido H78 (numeración de AHo; posición de Kabat H67) para el subtipo VH1b en las diferentes bases de datos. Las columnas en la tabla 9 son tal como sigue: columna 1: tipo de residuo; columna 2: frecuencia de residuos en la base de datos de líneas germinales IMGT; columna 3: frecuencia de residuos en la base de datos de líneas germinales Vbase; columna 4: frecuencia de residuos en una base de datos QC; columna 5: frecuencia de residuos en una base de datos Kabat.

25 Tabla 9: Frecuencia relativa de residuos en la posición 78 (numeración de AHo) para el subtipo VH1b en dos bases de datos de líneas germinales, una base de datos QC y una base de datos Kabat de anticuerpos maduros.

Residuo	IMGT_germ	Vbase_germ	Base de datos QC	KDB_VH1B
D				
E				
K				
R				
H				
T				
S				
N				
Q				
G				
A			24	2
C				
P				
V	91	100	47	86
I			18	1
L			12	
M				
F	9			10
Y				
W				
Consenso	V	V	V	V
% de concordancia	91	100	47	86
N.º de sec.*	11	11	17	81

\* Número de secuencias recopiladas para el análisis de la frecuencia de residuos

30 En la base de datos QC, se observó un residuo de alanina (A) a una frecuencia del 24%, un factor de 12 por encima de la frecuencia del 2% observada para el mismo residuo en una base de datos de anticuerpos maduros Kabat (KDB\_VH1B). Por consiguiente, se considera que un residuo de alanina en la posición H78 (numeración de AHo) es un residuo preferido en esa posición para potenciar las propiedades funcionales (por ejemplo, estabilidad y/o solubilidad) de un scFv. En cambio, se observó un residuo de valina (V) en la base de datos QC a una frecuencia relativa del 47%, mucho menor que la frecuencia del 86% observada en la base de datos de anticuerpos maduros Kabat y la frecuencia de más del 90% observada para el mismo residuo en bases de datos de líneas germinales (el 91% en IMGT-germ y el 100% en Vbase-germ). Por tanto, se consideró que un residuo de valina (V) era un residuo 35 desfavorable en la posición H78 en el contexto de un formato de scFv.

**EJEMPLO 5:** Comparación de Variantes de scFv ESBA105 a partir de dos enfoques diferentes

En este ejemplo se comparó la estabilidad de variantes de scFv preparadas mediante dos enfoques diferentes. El anticuerpo scFv parental era ESBA105, que se ha descrito previamente (véanse, por ejemplo, las publicaciones PCT WO2006/131013 y WO2008/006235). Se seleccionó un conjunto de variantes de ESBA105 usando el sistema de examen en levadura de control de calidad ("variantes QC"), variantes que también se han descrito previamente (véanse, por ejemplo, las publicaciones PCT WO2006/131013 y WO2008/006235). Se preparó el otro conjunto de variantes mediante la retromutación de determinadas posiciones de aminoácido a la secuencia consenso de línea germinal preferida identificada mediante el análisis de secuencia descrito en los ejemplos 2 y 3 anteriormente. Se seleccionaron las retromutaciones mediante la búsqueda dentro de las secuencias de aminoácidos de posiciones que estuvieran conservadas en la secuencia de línea germinal pero que contuvieran un aminoácido poco habitual o de baja frecuencia en el scFv seleccionado (denominado enfoque de modificación mediante ingeniería genética de residuo consenso de línea germinal).

Se sometieron a prueba todas las variantes para determinar la estabilidad sometiendo las moléculas a un estrés térmico inducido. Mediante la exposición a un amplio intervalo de temperaturas (25-95°C) fue posible determinar puntos medios aproximados de las transiciones de desplegamiento térmico (TM) para cada variante. Se realizaron mediciones de estabilidad térmica para las moléculas silvestres y las variantes con la espectroscopía ATR-FTIR en la que se guió la luz IR a través de un interferómetro. La señal medida es el interferograma, realizando una transformada de Fourier con esta señal, el espectro final es idéntico al de la espectroscopía infrarroja convencional (dispersiva).

A continuación en la tabla 10 se resumen los resultados de desplegamiento térmico y se representan gráficamente en la figura 6. Las columnas en la tabla 10 son tal como sigue: columna 1: variantes de ESBA105; columna 2: dominio que contiene la mutación; columna 3: mutación/mutaciones en numeración de AHO; columna 4: puntos medios de TM calculados a partir de las curvas de desplegamiento térmico en la figura 6; columna 5: actividad relativa en comparación con el ESBA105 parental; columna 5: estrategia de mutagénesis para la variante especificada en la columna 1.

Tabla 10: Comparación de variantes de ESBA105 a partir de dos enfoques diferentes y su contribución a la estabilidad global medida en FTIR (puntos medios calculados para las transiciones de desplegamiento térmico).

Variante	Dominio	Mutación	TM °C	Actividad de unión	Descripción
E105			61,53		Molécula parental
ESBA105_QC11.2	VH	F78L	66,26	1	Variante QC
ESBA105_QC15.2	VH	K50R, F78I	65,47	1	Variante QC
ESBA105_QC23.2	VH	F78L	66,53	1	Variante QC
ESBA105_VL R47K	VL	R47K	62,4	0,9	retromutada a residuo consenso
ESBA105_VL V103T	VL	V103T	60,7	1	retromutada a residuo consenso
ESBA105_VL V3Q	VL	V3Q	61,9	1,2	retromutada a residuo consenso

En comparación con las variantes QC, las retromutaciones al residuo consenso de línea germinal tuvieron un efecto negativo o nulo sobre la estabilidad térmica y actividad de ESBA105. Por tanto, estos resultados contradicen el enfoque de modificación mediante ingeniería genética de residuo consenso que se ha usado por otros para mejorar la estabilidad en diferentes anticuerpos y formatos (véanse, por ejemplo, Steipe, B *et al* (1994) J. Mol. Biol. 240:188-192; Ohage, E. y Steipe, B. (1999) J Mol. Biol. 291:1119-1128; Knappik, A. *et al.* (2000) J. Mol. Biol. 296:57-86, Ewert, S. *et al.* (2003) Biochemistry 42:1517-1528; y Monsellier, E. y Bedouelle, H. (2006) J. Mol. Biol. 362:580-593).

En un experimento independiente, se compararon las variantes QC anteriores (QC11.2, QC15.2 y QC23.2) y una variante QC adicional (QC7.1) con un segundo conjunto de variantes que tenían o bien retromutaciones consenso (S-2, D-2 y D-3) o bien retromutación a alanina (D-1) (véase la tabla 11). En la tabla 11 se indican la identidad del residuo en posiciones de región de marco seleccionadas y en la figura 7 se representa la estabilidad térmica medida (en unidades arbitrarias de desplegamiento).

Tabla 11: Se proporciona la identidad de los residuos de región de marco en posiciones de región de marco seleccionadas de variantes de ESBA105 que comprenden o bien retromutaciones consenso (S-2, D-2, D-3), una retromutación a alanina (D-1) o bien un residuo QC (QC7.1, QC11.2, QC15.2, QC23.2). Los residuos que difieren de los del anticuerpo parental ESBA105 se representan en cursiva y negrita. Se proporcionan las posiciones de aminoácido en la numeración de Kabat.

	Superficie de contacto VL-CL	Superficie de contacto VH-CH	Bucle exterior	Bucle exterior	Superficie de contacto VH-CH
	L83	H43	H67	H69	H78
Original	V	K	F	F	V
QC7.1	<b>E</b>	K	F	F	<b>A</b>
QC11.2	V	K	<b>L</b>	F	V
QC15.2	V	<b>R</b>	<b>I</b>	F	V
QC23.2	V	K	<b>L</b>	F	V
S-2	V	K	<b>V</b>	F	V
D-1	V	K	<b>A</b>	F	V
D-2	V	K	<b>V</b>	<b>L</b>	V
D-3	V	K	F	<b>L</b>	V

Aunque algunas variantes consenso (S-2 y D-1) presentaron una notable potenciación de la estabilidad térmica, esta potenciación fue menor que la potenciación de la estabilidad térmica lograda por cada una de las cuatro variantes QC.

5 Por consiguiente, los resultados en el presente documento demuestran que la presión de selección aplicada en el "sistema de selección en levadura de control de calidad" produce una subpoblación de armazones que sí que contienen características comunes rara vez observadas en la naturaleza (aunque todavía humanas) y presumiblemente responsables de las propiedades biofísicas superiores de estas regiones de marco. Mediante la exposición a 60°C de diferentes variantes de ESBA105, fue posible volver a confirmar las propiedades superiores de las sustituciones preferidas identificadas en la base de datos de región de marco de scFv seleccionado. Por tanto, se ha demostrado que el enfoque de "consenso funcional" descrito en el presente documento basándose en las secuencias de scFv seleccionadas obtenidas a partir del sistema de examen en levadura QC produce variantes de scFv que tienen estabilidad térmica superior a la de variantes preparadas usando el enfoque de consenso de línea germinal.

#### EJEMPLO 6: Variantes de scFv ESBA212

15 En este ejemplo se comparó la estabilidad de variantes de consenso de línea germinal de un anticuerpo scFv (ESBA212) con una especificidad de unión diferente de ESBA105. Se prepararon todas las variantes de ESBA212 mediante la retromutación de determinadas posiciones de aminoácido a la secuencia consenso de línea germinal preferida identificada mediante el análisis de secuencia descrito en los ejemplos 2 y 3 anteriormente. Se seleccionaron las retromutaciones mediante la búsqueda dentro de las secuencias de aminoácidos de posiciones que estuvieran conservadas en la secuencia de línea germinal pero que contuvieran un aminoácido poco habitual o de baja frecuencia en el scFv seleccionado (denominado enfoque de modificación mediante ingeniería genética de residuo consenso de línea germinal). Como en el ejemplo 5, se sometieron a prueba todas las variantes para determinar la estabilidad sometiendo las moléculas a un estrés térmico inducido.

25 A continuación en la tabla 11 se resumen los resultados de desplegamiento térmico para las variantes de ESBA212 y se representan gráficamente en la figura 8. Las columnas en la tabla 11 son tal como sigue: columna 1: variantes de ESBA212; columna 2: dominio que contiene la mutación; columna 3: mutación/mutaciones en numeración de AHo; columna 4: puntos medios de TM calculados a partir de las curvas de desplegamiento térmico en la figura 7; columna 5: actividad relativa en comparación con el ESBA212 parental; columna 5: estrategia de mutagénesis para la variante especificada en la columna 1.

30 Tabla 12: Comparación de variantes de ESBA212 retromutadas al residuo consenso de línea germinal y su contribución a la estabilidad global medida en FTIR (puntos medios calculados para las transiciones de desplegamiento térmico).

Variante	Dominio	Mutación	TM °C	Actividad de unión	Descripción
ESBA212			63,66		Molécula parental
ESBA212_VL R47K	VL	R47K	59,94	2,8	retromutada a residuo consenso
ESBA212_VL V3Q	VL	V3Q	63,6	1,1	retromutada a residuo consenso

35 Tal como se observa para el anticuerpo scFv ESBA no relacionado, retromutaciones al residuo consenso de línea germinal tuvieron un efecto negativo o nulo sobre la estabilidad térmica y actividad de ESBA212. Por tanto, estos resultados sirven para destacar adicionalmente lo inadecuado de los enfoques basados en consenso convencionales. Estas deficiencias pueden abordarse mediante el empleo de la metodología de consenso funcional de la invención.

**EJEMPLO 7:** Sustituciones de aminoácidos a modo de ejemplo y preferidas en posiciones de región de marco de scFv identificadas

Usando el enfoque de análisis de scFv basado en secuencia descrito anteriormente en los ejemplos 2, 3 y 4, fue posible identificar sustituciones de aminoácidos a modo de ejemplo y preferidas en posiciones de residuo de aminoácido dentro de las regiones de marco de scFv en la base de datos de scFv seleccionado QC que presentaban diferencias en la variabilidad en comparación con la bases de datos de líneas germinales. Se realizó este análisis mediante la determinación de la frecuencia de cada uno de los veinte aminoácidos en cada posición de región de marco particular de interés dentro de las dos bases de datos de líneas germinales (IMGT y Vbase), la base de datos de scFv seleccionado QC y la base de datos de anticuerpos maduros (KDB), tal como se describió en el ejemplo 4 para la posición de AHo 78 (posición de Kabat 67) para la familia de cadena pesada de VH1b como ejemplo representativo. Se identificaron sustituciones de aminoácidos a modo de ejemplo y preferidas para tres familias de región variable de cadena pesada, VH3, VH1a y VH1b, y para tres familias de región variable de cadena ligera, V $\kappa$ 1, V $\kappa$ 3 y V $\lambda$ 1.

A continuación en las tablas 13-18 se resumen los resultados. Para cada tabla, la columna uno muestra la posición de residuo usando el sistema de numeración de AHo, la columna dos muestra el residuo consenso de línea germinal, la columna tres muestra las sustituciones a modo de ejemplo encontradas en las regiones de marco de scFv seleccionado de QC, la columna 4 muestra el residuo preferido encontrado en las regiones de marco de scFv seleccionado de QC y las columnas 5 a 8 muestran la frecuencia relativa de residuos en las cuatro bases de datos diferentes para la sustitución preferida (mostrada en la columna 4) en la posición de residuo indicada en la columna 1.

Tabla 13: Sustituciones de aminoácidos a modo de ejemplo y preferidas de posiciones de residuo identificadas como características únicas de las regiones de marco de scFv seleccionado de QC para la familia de VH3.

Posición de residuo	Residuo consenso	Sustituciones	Sustitución preferida	Línea germinal de IMGT	Línea germinal de VBase	scFv seleccionado de QC	Sec. de KDB
1	E	E, Q	Q	15,38	22,73	46,43	28,13
6	E	E, Q	Q	0,00	0,00	31,03	6,98
7	S	T, S, A	T	0,00	4,55	20,69	0,46
89	L	A, V, L, F	V	0,00	0,00	22,81	6,37
103	V	R, Q, V, I, L, M, F	L	11,54	13,64	25,00	9,96

Tabla 14: Sustituciones de aminoácidos a modo de ejemplo y preferidas de posiciones de residuo identificadas como características únicas de las regiones de marco de scFv seleccionado de QC para la familia de VH1a.

Posición de residuo	Residuo consenso	Sustituciones	Sustitución preferida	Línea germinal de IMGT	Línea germinal de VBase	scFv seleccionado de QC	Sec. de KDB
1	Q	E, Q	E	10,00	9,09	25,00	0,00
6	Q	E, Q	E	0,00	0,00	41,67	15,69
12	V	L, V	L	0,00	0,00	16,67	0,00
13	K	M, K	M	0,00	0,00	16,67	0,00
14	K	E, Q, K	E	0,00	0,00	16,67	1,85
19	V	L, V	L	0,00	0,00	16,67	0,00
21	V	I, V	I	9,09	9,09	16,67	0,00
90	Y	F, S, H, D, Y	Nd				
92	E	D, Q, E	D	9,09	0,00	16,67	1,85
95	S	G, N, T, S	G	0,00	0,00	16,67	7,41
98	S	T, A, P, F, S	F	0,00	0,00	16,67	1,85

ES 2 566 737 T3

Tabla 15: Sustituciones de aminoácidos a modo de ejemplo y preferidas de posiciones de residuo identificadas como características únicas de las regiones de marco de scFv seleccionado de QC para la familia de VH1b.

Posición de residuo	Residuo consenso	Sustituciones	Sustitución preferida	Línea germinal de IMGT	Línea germinal de VBase	scFv seleccionado de QC	Sec. de KDB
1	Q	Q, E	E	10,00	9,09	29,41	1,45
10	A	A, T, P, V, D	T	0,00	0,00	11,76	2,50
12	V	V, L	L	0,00	0,00	23,53	7,41
13	K	K, V, R, Q, M	V	0,00	0,00	11,76	0,00
14	K	E, K, R, M	R	0,00	0,00	17,65	2,47
20	K	R, T, K, N	N	0,00	0,00	11,76	0,00
21	V	I, F, V, L	L	0,00	0,00	17,65	2,47
45	R	R, K	K	0,00	0,00	23,53	0,00
47	A	T, P, V, A, R	R	0,00	0,00	23,53	0,00
50	Q	K, Q, H, E	K	18,18	18,18	23,53	2,47
55	M	M, I	I	9,09	9,09	23,53	3,70
77	R	K, R	K	0,00	0,00	23,53	0,00
78	V	A, V, L, I	A	0,00	0,00	23,53	2,47
82	R	E, R, T, A	E CONS	9,09	9,09	29,41	1,23
86	I	T, S, I, L	T CONS	63,64	63,64	52,94	29,63
87	S	D, S, N, G	N CONS	0,00	0,00	37,50	18,52
107	A	N, S, A	N	0,00	0,00	18,75	0,00

Tabla 16: Sustituciones de aminoácidos a modo de ejemplo y preferidas de posiciones de residuo identificadas como características únicas de las regiones de marco de scFv seleccionado de QC para la familia de Vκ1.

Posición de residuo	Residuo consenso	Sustituciones	Sustitución preferida	Línea germinal de IMGT	Línea germinal de VBase	scFv seleccionado de QC	Sec. de KDB
1	D	D, E, I	E	0%	0%	74%	10%
3	Q	Q, V, I	V	0%	0%	79%	8%
4	M	V, L, I, M	L	23%	16%	72%	21%
24	R	R, Q	Q	9%	11%	23%	11%
47	K	K, R, I	R	0%	0%	16%	2%
50	R	K, R, E, T, M, Q	nd				
57	Y	H, S, F, Y	S	0%	0%	14%	5%
91	L	L, F	F	9%	11%	19%	12%
103	T	V, S, G, I	V	0%	0%	9%	1%

Tabla 17: Sustituciones de aminoácidos a modo de ejemplo y preferidas de posiciones de residuo identificadas como características únicas de las regiones de marco de scFv seleccionado de QC para la familia de Vκ3.

Posición de residuo	Residuo consenso	Sustituciones	Sustitución preferida	Línea germinal de IMGT	Línea germinal de VBase	scFv seleccionado de QC	Sec. de KDB
2	I	I, T	T	0%	0%	17%	1%
3	V	V, T	T	0%	0%	17%	0%
10	T	T, I	I	0%	0%	17%	1%
12	S	S, Y	Y	0%	0%	17%	0%
18	R	S, R	S	0%	0%	17%	1%
20	T	T, A	A	0%	0%	17%	1%
56	I	I, M	M	0%	0%	17%	2%
74	I	I, V, T	T	0%	0%	17%	1%
94	S	S, N	N	0%	0%	17%	3%
101	F	F, Y, S	S	0%	0%	17%	2%
103	F	F, L, A	A	0%	0%	17%	0%

ES 2 566 737 T3

Tabla 18: Sustituciones de aminoácidos a modo de ejemplo y preferidas de posiciones de residuo identificadas como características únicas de las regiones de marco de scFv seleccionado de QC para la familia de Vλ1.

Posición de residuo	Residuo consenso	Sustituciones	Sustitución preferida	Línea germinal de IMGT	Línea germinal de VBase	scFv seleccionado de QC	Sec. de KDB
1	Q	L, Q, S, E	L	0,00	0,00	18,75	0,79
2	S	S, A, P, I, Y	P	0,00	0,00	31,25	0,37
4	L	V, M, L	V	0,00	0,00	18,75	5,45
7	P	S, E, P	S	0,00	0,00	6,25	0,68
11	V	A, V	A	28,57	40,00	62,50	38,95
14	A	T, S, A	T	28,57	40,00	62,50	38,22
46	Q	H, Q	H	0,00	0,00	18,75	9,21
53	K	K, T, S, N, Q, P	nd				
82	K	R, Q, K	R	0,00	0,00	18,75	3,32
92	A	G, T, D, A	T	0,00	0,00	12,50	0,51
103	D	D, V, T, H, E	V	0,00	0,00	12,50	0,26

5 Tal como se demuestra mediante los resultados mostrados en las tablas 13-18, se encontró que un subconjunto de posiciones de residuo en las regiones de marco de scFv seleccionado de QC estaban fuertemente sesgadas hacia determinados residuos no presentes o con escasa representación en las secuencias de línea germinal y en secuencias de anticuerpos maduros y, por tanto, aparentemente no usados en el formato de Ig o fragmentos derivados. Por tanto, las sustituciones a modo de ejemplo y preferidas identificadas en las regiones de marco de scFv seleccionado de QC representan residuos de aminoácido que es probable que contribuyan a las propiedades funcionales deseables (por ejemplo, estabilidad, solubilidad) presentadas por las regiones de marco de scFv seleccionado de QC.

**EJEMPLO 8:** Armazones de región de marco de scFv basándose en consenso funcional

10 Basándose en las sustituciones de aminoácidos a modo de ejemplo y preferidas identificadas en el ejemplo 7, se diseñaron armazones de región de marco de scFv basándose en el enfoque de consenso funcional descrito en el presente documento. En estos armazones de región de marco de scFv, las secuencias de CDR1, CDR2 y CDR3 no están definidas, puesto que estos armazones representan secuencias de región de marco en las que puede insertarse esencialmente cualquier secuencia de CDR1, CDR2 y CDR3. Además, en los armazones de región de marco de scFv, se permite que aquellas posiciones de aminoácido que se han identificado como que son propensas a variabilidad (tal como se expone en las tablas del ejemplo 7) estén ocupadas por cualquiera de las sustituciones de aminoácidos a modo de ejemplo o preferidas identificadas para esa posición.

20 En las figuras 9-11 se representan armazones de región de marco de cadena pesada (véanse también las tablas 19-21). Por tanto, para la familia de VH1a, el armazón de región de marco de scFv se ilustra en la figura 9 (véase también la tabla 19). Para la familia de VH1b, el armazón de región de marco de scFv se ilustra en la figura 10 (véase también la tabla 20). Para la familia de VH3, el scFv región de marco se ilustra en la figura 11 (véase también la tabla 21). Para las alineaciones en cada una de estas figuras, la primera fila muestra la numeración de región variable de cadena pesada usando el sistema de Kabat y la segunda fila muestra la numeración de región variable de cadena pesada usando el sistema de AHo. La tercera fila muestra la secuencia del armazón de región de marco de scFv, en la que en aquellas posiciones marcadas como "X", la posición puede estar ocupada por cualquiera de los residuos de aminoácido enumerados debajo de la "X". Además, las posiciones marcadas como "x" (es decir, 26, 27, 28, 29 de Kabat y 27, 29, 30, 31 de AHo en las figuras) y las regiones marcadas como CDR pueden estar ocupadas por cualquier aminoácido. Para las posiciones variables marcadas como "X", el primer residuo de aminoácido enumerado debajo de la "X" representa el residuo consenso de línea germinal, el segundo residuo de aminoácido enumerado debajo de la "X" representa la sustitución de aminoácido preferida en esa posición y los residuos de aminoácido adicionales enumerados debajo de la "X" (si los hubiera) representan otras sustituciones de aminoácidos a modo de ejemplo en esa posición.

Tabla 19: Armazón de región de marco de cadena pesada de la familia de VH1

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												
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												
X	V	Q	L	V	X	S		G	A	E	X	X	X	P	G	S	S	X	K	X	S	C	K	A	S												
Q					Q						V	K	K				V			V																	
E					E						L	M	E					L		I																	
													Q																								
26	*	27	28	29	30	31	32	33	34	35	35a	35b	*	*	*	36	37	38	39	40	41	42	43	44	45												
27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52												
x	x	x	x	x												W	V	R	Q	A	P	G	Q	G	L												
CDR H1												CDR H2												R		V		T									
46	47	48	49	50	51	52	52a	52b	52c	*	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68											
53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79											
E	W	M	G																																		
69	70	71	72	73	74	75	76	77	78	79	80	81	82	82a	82b	82b	83	84	85	86	87	88	89	90	91												
80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105												
I	T	A	D	E	S	T	S	T	A	X	M	X	L	S	X	L	R	X	E	D	T	A	V	Y	Y												
										Y	E				S			S																			
										F	D				G			F																			
										S	Q				T			T																			
										H					N			A																			
										D								P																			
92	93	94	95	96	97	98	99	100	100a	100b	100c	100d	100e	100f	100g	100h	100i	*	*	*	*	*	*	*	*												
106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128															
CDR H3												CDR H3												126		127		128									
C	A	R																																			
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*											
129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154												
										W	G	Q	G	T	L	V	T	V	T	V	S	S	S	S	S												
CDR H3												CDR H3												146		147		148		149		150		151		152	



Tabla 21: Armazón de región de marco de cadena pesada de la familia de VH3

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										
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										
X	V	Q	L	V	X	X	G	P	G	L	V	K	P	S	E	T	L	R	L	S	C	A	A	S											
E					E	S																													
Q					Q	T																													
						A																													
26	*	27	28	29	30	31	32	33	34	35	35a	35b	*	*	36	37	38	39	40	41	42	43	44	45											
27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52										
G		x	x	x											W	V	R	Q	A	P	G	K	G	L											
												CDR H1																							
46	47	48	49	50	51	52	52a	52b	52c	*	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68									
53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79									
E	W	V	S																																
												CDR H2																							
69	70	71	72	73	74	75	76	77	78	79	80	81	82	82a	82b	83	84	85	86	87	88	89	90	91											
80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105										
I	S	R	D	N	S	K	N	T	X	Y	L	Q	M	N	S	L	R	A	E	D	T	A	X	Y	Y										
									L																										
									V																										
									A																										
									F																										
92	93	94	95	96	97	98	99	100	100a	100b	100c	100d	100e	100f	100g	100h	100i	100j	*	*	*	*	*	*	*										
106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128													
C	A	R																																	
												CDR H3																							
*	*	*	*	*	*	*	*	*	*	101	102	103	104	105	106	107	108	109	110	111	112	113													
129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149															
										W	G	Q	G	T	L	V	T	V	T	V	S	S	S	S	S										
												CDR H3																							

En las figuras 12-14 se representan armazones de región de marco de cadena ligera (véanse también las tablas 22-24). Para la familia de  $V_{\kappa 1}$ , el armazón de región de marco de scFv se ilustra en la figura 12 (véase también la tabla 22). Para la familia de  $V_{\kappa 3}$ , el armazón de región de marco de scFv se ilustra en la figura 13 (véase también la tabla 23). Para la familia de  $V_{\lambda 1}$ , el scFv región de marco se ilustra en la figura 14 (véase también la tabla 24). Para las alineaciones en cada una de estas figuras, la primera fila muestra la numeración de región variable de cadena ligera usando el sistema de Kabat y la segunda fila muestra la numeración de región variable de cadena ligera usando el sistema de AHo. La tercera fila muestra la secuencia del armazón de región de marco de scFv, en la que en aquellas posiciones marcadas como "X", la posición puede estar ocupada por cualquiera de los residuos de aminoácido enumerados debajo de la "X." Además, las posiciones de región de marco marcadas como "." y las regiones marcadas como CDR pueden estar ocupadas por cualquier aminoácido.



Tabla 23: Armazón de región de marco de cadena ligera de la familia de Vκ3

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	27a	27b	27c	27d	27e
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	29	30	31	32	33
E	X	X	L	T	Q	S	P	G	X	L	X	L	S	P	G	E	X	A	X	L	S	C	R								
T	T								I	Y							S	A	A												
I	V								T	S							R	T													

27f	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66
34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66							
								W	Y	Q	Q	K	P	G	Q	A	P	R	L	L	M	Y																	

51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81								
67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99						
							G	I	P	D	R	F	S	G	S	G	S	G	T	D	F	T	L	T	I	N	R	L	E	P	E							
								T																														
								V																														

82	83	84	85	86	87	88	89	90	91	92	93	94	95	95a	95b	95c	95d	95e	95f	96	97	98	99	100	101	102	103	104	105	106	107	108						
100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132						
D	S	A	A	Y	Y	C																																
	Y																																					
	F																																					

133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149
			96	97	98	99	100	101	102	103	104	105	106	107	108	
						F	G	G	T	K	L	E	I	K	R	



**EJEMPLO 9:** Generación de scFv con solubilidad mejorada

En este ejemplo se usó un enfoque basado en modelado estructural y análisis de secuencia para identificar mutaciones en regiones de marco de scFv que dan como resultado solubilidad mejorada.

**a) Análisis estructural**

- 5 Se modeló la estructura 3D del scFv ESBA105 usando el servidor automatizado de modelado de homología estructural de proteínas, accesible mediante el servidor web ExpASy. Se analizó la estructura según la superficie relativa accesible al disolvente (rSAS) y se clasificaron los residuos tal como sigue: (1) expuestos para residuos que muestran una rSAS  $\geq 50\%$ ; y (2) parcialmente expuestos para residuos con  $50\% \leq rSAS < 25\%$ . Los residuos hidrófobos con una rSAS  $\geq 25\%$  se consideraron parches hidrófobos. Para validar el área accesible al disolvente de cada parche hidrófobo encontrado, se realizaron cálculos a partir de 27 archivos de PDB con alta homología con ESBA105 y una resolución mayor de 2,7 Å. Se calcularon la rSAS promedio y la desviación estándar para los parches hidrófobos y se examinaron en detalle para cada uno de ellos (véase la tabla 25).

Tabla 25: Evaluación de los parches hidrófobos.

Residuo	Dominio	% de superficie expuesta al disolvente	% de STDE	rSAS	Variabilidad de secuencia	Superficie de contacto VH/antígeno	Superficie de contacto VH/VL	Superficie de contacto VH/CH
2	VH	23,06	19,26	10-25%	10-25%	> 0-20%	> 0-20%	0
4	VH	0,66	1,26	0-10%	0-10%	0		0
5	VH	61,85	12,96	50-75%	10-25%	0	> 0-20%	0
12	VH	70,27	9,17	50-75%	10-25%	0	0	60-80%
103	VH	35,85	5,85	25-50%	10-25%	0	> 0-2%	> 0-2%
144	VH	62,17	7,82	50-75%	10-25%	0	0	> 0-2%
15	VL	49,59	9,77	25-50%	10-25%	0	0	0
147	VL	31,19	23,32	25-50%	10-25%	0	0	60-80%

Columna 1: posición de residuo en el sistema de numeración de AHo. Columna 2, dominio para la posición indicada en la columna 1. Columna 3, cálculos de área accesible al disolvente promedio a partir de 27 archivos de PDB. Columna 4, desviaciones estándar de la columna 3. Columnas 5 a 9, papel estructural de los parches hidrófobos recuperados de AHo.

- 15 La mayor parte de los parches hidrófobos identificados en ESBA105 correspondían a la superficie de contacto dominio variable-constante (VH/CH). Esto se correlacionó con hallazgos previos de residuos hidrófobos expuestos al disolvente en un formato de scFv (Nieba *et al.*, 1997). Dos de los parches hidrófobos (VH 2 y VH 5) también contribuían a la interacción VL-VH y, por tanto, se excluyeron del análisis posterior.

**b) Diseño de mutaciones de solubilidad**

- 20 Se recuperaron un total de 122 secuencias de VL y 137 de VH de la página web sobre anticuerpos de Annemarie Honegger ([www.bioc.uzh.ch/antibody](http://www.bioc.uzh.ch/antibody)). Las secuencias correspondían originariamente a 393 estructuras de anticuerpo en formato de Fv o Fab extraídas del Banco de Datos de Proteínas (PDB) ([www.rcsb.org/pdb/home/home.do](http://www.rcsb.org/pdb/home/home.do)). Se usaron las secuencias para el análisis independientemente de la especie o el subgrupo para aumentar la probabilidad de encontrar aminoácidos alternativos con mayor hidrofiliidad que el residuo nativo. Se excluyeron las secuencias que tenían más del 95% de identidad con cualquier otra secuencia dentro de la base de datos para reducir el sesgo. Se alinearon las secuencias y se analizaron para determinar la frecuencia de residuos. Se aplicaron algoritmos y herramientas de análisis de secuencia para identificar y seleccionar mutaciones hidrófilas para perturbar los parches hidrófobos en ESBA105. Se alinearon las secuencias siguiendo el sistema de numeración de AHo para dominio variable de inmunoglobulina (Honegger y Plückthun 2001). Se restringió el análisis a las regiones de marco.

- 30 Se calculó la frecuencia de residuos,  $f(r)$ , para cada posición,  $i$ , en la base de datos personalizada mediante el número de veces que se observa un residuo particular dentro del conjunto de datos dividido entre el número total de secuencias. En una primera etapa, se calculó la frecuencia de aparición de los diferentes aminoácidos para cada parche hidrófobo. Se analizó la frecuencia de residuos para cada parche hidrófobo identificado en ESBA105 a partir de la base de datos personalizada descrita anteriormente. La tabla 26 notifica la frecuencia de residuos en los parches hidrófobos dividida entre la totalidad de los residuos presentes en la base de datos.

Tabla 26. Frecuencia de residuos de 259 secuencias de anticuerpos maduros en un formato de scFv o Fab para los parches hidrófobos identificados en ESBA105

Residuo	VH 4	VH 12	VH 103	VH 144	VL 15	VL 147
A	0,23046215	0	0	0	3,8647343	0,176821923
C	0	0	0	0	0	0
D	0	0	0	0	0	0
E	0	0	0	0	0	0
F	0,483091787	0	0,483091787	0	0	0
G	0	0	0	0	0	0
H	0	0	0	0	0	0
I	0	2,415458937	9,661835749	0	5,314009662	70,38834951
K	0	0	0	0	0	0
L	96,61835749	89,85507246	7,246376812	27,0531401	45,89371981	15,53398058
M	0	0	10,62801932	1,93236715	0	0,970873786
N	0	0	0	0	0	0
P	0,966183575	0	0	0,966183575	21,73913043	0,485436893
Q	0	0	0	0,483091787	0	0
R	0	0	7,246376812	0	0	0
S	0	0,966183575	0	18,84057971	0	0
T	0	0	15,4589372	50,72463768	0,966183575	0
V	1,93236715	6,763285024	49,27536232	0	22,22222222	12,62135922
W	0	0	0	0	0	0
Y	0	0	0	0	0	0

Columna 1: tipo de residuo. Columnas 2 a 5, frecuencia relativa de residuos para los parches hidrófobos en la cadena pesada. Columnas 6 y 7, frecuencia relativa de residuos para los parches hidrófobos en la cadena ligera.

5 En la segunda etapa, se usó la frecuencia de residuos hidrófilos en los parches hidrófobos para diseñar las mutaciones de solubilidad mediante la selección del residuo hidrófilo más abundante en cada parche hidrófobo. La tabla 27 notifica los mutantes de solubilidad identificados usando este enfoque. Se calculó la hidrofobicidad de los residuos parentales y mutantes como hidrofobicidad promedio de valores publicados en varios artículos y expresados en función del nivel de exposición de la cadena lateral al disolvente.

Tabla 27. Diferentes mutaciones de solubilidad introducidas en ESBA105 para perturbar los parches hidrófobos

Residuo	Dominio	% de superficie expuesta al disolvente	Residuo parental	Hidrofobicidad del residuo parental	Mutación de solubilidad	Hidrofobicidad de mutaciones
4	VH	0,66	L	85,2	A	42,7
12	VH	70,27	V	73,2	S	28
103	VH	35,85	V	73,2	T	32,8
144*	VH	62,17	V	73,2	S	28
15	VL	49,59	V	73,2	T	32,8
147	VL	31,19	L	85,2	A	42,7

\*El parche hidrófobo en la posición 144 se intercambió no por el residuo hidrófilo más abundante en la base de datos sino por Ser puesto que esta ya estaba contenida en el donador de CDR de ESBA105.

Columna 1, posición de residuo en el sistema de numeración de AHo. Columna 2, dominio para la posición indicada en la columna 1. Columna 3, cálculos de área accesible al disolvente promedio a partir de 27 archivos de PDB. Columna 4, residuos parentales en ESBA105. Columna 5, hidrofobicidades promedio de la columna 4, recuperadas de AHo. Columnas 6, residuo hidrófilo más abundante en la posición indicada en la columna 1. Hidrofobicidad promedio de la columna 6 recuperada de AHo.

c) Pruebas de variantes de ESBA105 de solubilidad

10 Se introdujeron las mutaciones de solubilidad solas o en múltiples combinaciones y se sometieron a prueba para determinar los patrones de rendimiento de repliegamiento, expresión, actividad y estabilidad y agregación. La tabla 28 muestra las diversas combinaciones de mutaciones de solubilidad introducidas en cada variante optimizada de ESBA105 basándose en la posible contribución a la solubilidad y el nivel de riesgo de que la mutación altere la unión a antígeno.

15

Tabla 28: Diseño de variantes de solubilidad para ESBA105.

Residuo de superficie hidrófobo	Dominio	Residuo parental	Mutantes**			
			Opt 1_0	Opt 0_2	Opt 1_2	Opt 2_4
15	VL	V	X		X	X
147*	VL	V				X
4*	VH	L				X
12	VH	V		X	X	X
103*	VH	V				X
144	VH	L		X	X	X

\*Sometido a prueba por separado en una segunda tanda

\*\*El guión bajo separa el número de mutaciones contenidas en la cadena ligera y en la pesada, respectivamente.

Columna 1, posición de residuo en el sistema de numeración de AHo. Columna 2, dominio para la posición indicada en la columna 1. Columna 3, residuo parental en ESBA105 en los diferentes parches hidrófobos. Columna 4, diferentes variantes que contienen mutaciones de solubilidad en las posiciones indicadas.

i. Mediciones de solubilidad

5 Se determinaron las solubilidades máximas de ESBA105 y variantes midiendo la concentración de proteína en los sobrenadantes de mezclas de PEG-proteína centrifugadas. Se mezcló 1:1 una concentración inicial de 20 mg/ml con disoluciones de PEG que oscilaban entre el 30 y el 50% de saturación. Se eligieron estas condiciones basándose en el perfil de solubilidad observado para el ESBA105 silvestre después la determinación empírica de la dependencia lineal de Log S frente a la concentración de Peg (% p/v). En la figura 15 se representan curvas de solubilidad de varios ejemplos de ESBA105 variante que presentaban una solubilidad superior. En la tabla 29A también se proporciona una lista completa de valores de solubilidad.

Tabla 29. Actividad y solubilidad máxima estimada de los mutantes en comparación con el ESBA105 parental.

Molécula	E105	E105 Opt 1_0	E105 Opt 0_2	E105 Opt 1_2	E105 VH V103T	E105 VL V147A
Ordenada en el origen	1,956	2,228	2,179	2,163	2,223	2,047
Solubilidad máxima	90,36	169,04	151,01	145,55	167,11	111,43
Actividad con relación a ESBA105	1	1,4	1,5	1,5	1,2	2

10 ii. Mediciones de estabilidad térmica

15 Se realizaron mediciones de estabilidad térmica para el ESBA105 parental y seguimientos de la solubilidad usando espectroscopía ATR-FTIR. Se sometieron las moléculas a una exposición térmica a un amplio intervalo de temperaturas (de 25 a 95°C). Se obtuvo el perfil de desnaturalización mediante la aplicación de una transformada de Fourier a las señales de interferograma (véase la figura 16). Se usaron los perfiles de desnaturalización para aproximar los puntos medios de las transiciones de desplegamiento térmico (TM) para cada variante de ESBA105 aplicando el modelo sigmoidal de Boltzmann (tabla 30).

Tabla 30: Puntos medios de las transiciones de desplegamiento térmico (TM) para cada variante de solubilidad.

	ESBA105	E105 Opt 1.0	E105 Opt 1.2	E105 Opt 0.2	E105 VH V103T	E105 VL V147A
Boltzmann sigmoidal						
Valores de ajuste óptimo						
INFERIOR	0,3604	-0,405	0,7032	0,4516	0,4691	-0,6873
SUPERIOR	100,4	99,3	98,84	99,04	99,2	99,16
V50	61,53	59,91	59,39	60,86	62,08	55,89
PENDIENTE	2,935	2,886	3,117	2,667	2,682	3,551
Desviación estándar						
INFERIOR	0,5206	0,3471	0,6652	0,4953	0,3938	0,4754
SUPERIOR	0,5361	0,3266	0,6116	0,4891	0,4167	0,3714
V50	0,1047	0,06658	0,1328	0,0949	0,07811	0,0919
PENDIENTE	0,09039	0,05744	0,1146	0,08199	0,06751	0,08235
Intervalos de confianza del 95%						
INFERIOR	de -0,7432 a 1,464	de -1,141 a 0,3309	de -0,7071 a 2,114	de -0,5984 a 1,502	de -0,3658 a 1,304	de -1,695 a 0,3206
SUPERIOR	de 99,25 a 101,5	de 98,61 a 99,99	de 97,54 a 100,1	de 98,01 a 100,1	de 98,32 a 100,1	de 98,38 a 99,95
V50	de 61,31 a 61,75	de 59,77 a 60,06	de 59,11 a 59,67	de 60,66 a 61,06	de 61,91 a 62,24	de 55,70 a 56,09
PENDIENTE	de 2,743 a 3,127	de 2,764 a 3,007	de 2,874 a 3,360	de 2,494 a 2,841	de 2,539 a 2,825	de 3,376 a 3,725
Bondad del ajuste						
Grados de libertad	16	16	16	16	16	16
R <sup>2</sup>	0,9993	0,9997	0,999	0,9994	0,9996	0,9996
Suma absoluta de cuadrados	26,18	10,8	37,2	24	16,14	15,11
Sy.x	1,279	0,8217	1,525	1,225	1,004	0,9719

iii. Mediciones de agregación

5 También se analizaron ESBA105 y sus variantes de solubilidad en una prueba con dependencia temporal para evaluar el comportamiento de degradación y agregación. Para este fin, se incubaron proteínas solubles (20 mg/ml) a una temperatura elevada (40°C) en tampones fosfato a pH 6,5. Se mantuvieron muestras de control a -80°C. Se analizaron las muestras después de un periodo de incubación de dos semanas para determinar la degradación (SDS-PAGE) y agregación (SEC). Esto permitió descartar variantes que eran propensas a degradación (véase la figura 17) o que presentaban una tendencia a formar agregados solubles o insolubles (véase la tabla 31).

Tabla 31: Mediciones de agregación insoluble.

Proteína	Pérdida de proteína (agregados insolubles)
ESBA105	1,14%
ESBA105 Opt 1_0	8,17%
ESBA105 Opt 0_2	4,45%
ESBA105 Opt 1_2	46,60%
ESBA105 VH V103T	-1,95%

iv. Expresión y repliegamiento de variantes de solubilidad

10 También se sometieron a prueba los mutantes de solubilidad para determinar la expresión y el rendimiento de repliegamiento con relación a la molécula de ESBA105 parental. En la tabla 31 se muestran los resultados de estos estudios.

Tabla 31. Expresión y plegamiento de variantes de solubilidad.

	Residuo de superficie hidrófobo							Expresión con relación a ESBA105	Rendimiento de plegamiento, mg/l
	VH				VL				
ESBA105	L4	V12	V103	L144	V15	F52	V147	1,0	34
Opt 1_0					T			1,15	12,5
Opt 0_2		S		S				1,10	35
Opt 1_2		S		S	T			0,96	44
Opt 2_4	A	S	T	S	T		A	1,20	no puede producirse
VH L4A								1,0	no puede producirse
VH V103T			T					1,1	55
VL V147A							A	1,2	20

Aunque todos los mutantes de solubilidad hidrófilos presentaban una solubilidad mejorada en comparación con la molécula de ESBA105 parental, sólo algunas de estas moléculas se presentaron como adecuadas para otras propiedades biofísicas. Por ejemplo, muchas variantes tenían una estabilidad térmica y/o rendimiento de plegamiento reducidos con relación a la molécula de ESBA105 parental. En particular, el reemplazo hidrófilo en la posición VL147 disminuyó gravemente la estabilidad. Por tanto, se combinaron mutaciones de solubilidad que no afectaron significativamente a la estabilidad térmica y se sometieron a estrés térmico adicional para confirmar sus propiedades.

Tres mutantes que contenían una combinación de cuatro mutaciones de solubilidad diferentes (Opt1.0, Opt0.2 y VH:V103T) mejoraron significativamente la solubilidad de ESBA105 sin afectar a la reproducibilidad, actividad o estabilidad térmica. Sin embargo, un mutante que tenía las mutaciones combinadas de Opt1.0 y Opt0.2 en ESBA105 (Opt 1\_2) presentaba un aumento de la cantidad de agregados insolubles después de la incubación durante 2 semanas a 40°C (véase la tabla 29). Esto podría explicarse por el papel de Val en la posición VL 15 en un giro de lámina beta, puesto que Val tiene la mayor propensión a formar láminas beta de todos los aminoácidos. Este resultado demostró que se tolera una mutación de solubilidad individual en la posición VL 15, pero no en combinación con mutantes de solubilidad que perturban otros parches hidrófobos. Por tanto, se seleccionaron las mutaciones contenidas en Opt0\_2 y VH:V103T como las de mejor desempeño para mejorar propiedades de solubilidad de moléculas de scFv.

#### EJEMPLO 10: Generación de scFv con solubilidad y estabilidad potenciadas

Se optimizaron adicionalmente las variantes de ESBA105 identificadas mediante diseño de solubilidad, mediante sustitución con mutaciones estabilizantes identificadas mediante el ensayo de control de calidad (QC). Se crearon un total de 4 constructos que contenían entre 1 y 3 de las mutaciones de solubilidad identificadas en el ejemplo 9 anteriormente, en combinación con todas las mutaciones estabilizantes encontradas en QC 7.1 y 15.2 (es decir, D31N y V83E en el dominio VL y V78A, K43 y F67L en el dominio VH). Todos los constructos optimizados produjeron más proteína soluble que un scFv silvestre (véase la tabla 33). El mejor constructo presentaba de manera sistemática un aumento de más de 2 veces de la solubilidad con respecto al silvestre. No se vieron afectadas significativamente ni la actividad ni la estabilidad de las moléculas de scFv por la combinación de mutaciones estabilizantes y potenciadoras de la solubilidad.

Tabla 33: ScFv con una solubilidad y estabilidad optimizadas

Proteína	Mutaciones de VL/VH	T <sub>m</sub> de FTIR (°C)	Solubilidad de PEG (mg/ml)	Actividad con relación a E105	kD
QC7.1D-N-15.2	VL: D31N; V83E VH: V78A; K43R; F67L	69,0	90	1,7	9,06 x 10 <sup>-10</sup>
QC7.1D-N-15.2 VH V103T	VL: D31N; V83E VH: V78A; K43R; F67L; V103T	68,9	106	1,5	8,79 x 10 <sup>-10</sup>
QC7.1D-N-15.2 Opt 0_2	VL: D31N; V83E VH: V12S; V78A; K43R; F67L; L144S	66,6	121	1,2	8,12 x 10 <sup>-10</sup>
QC7.1D-N-15.2 VH V103T Opt 0_2	VL: D31N; V83E VH: V12S; V78A; K43R; F67L; V103T; L144S	67,3	186	1,5	1,34 x 10 <sup>-9</sup>

Se usaron los valores de solubilidad para las 4 variantes para deconvolucionar la contribución de cada mutación a la solubilidad del scFv. Todas las mutaciones parecieron contribuir a la solubilidad del scFv de manera aditiva aunque varios de estos residuos estén relativamente próximos entre sí tanto en la secuencia primaria como dentro de la estructura 3D. El análisis indicó que una combinación de tres mutaciones potenciadoras de la solubilidad en el dominio VH (V12S, L144S, V103T (o V103S)) representaba ~60% de la solubilidad de scFv. Puesto que los parches hidrófobos están conservados en los dominios variables de todos los agentes de unión inmunológica, esta

combinación óptima de mutaciones puede usarse para mejorar la solubilidad de prácticamente cualquier scFv u otra molécula de agente de unión inmunológica.

**Lista de secuencias**

- <110> David Urech, Leonard Borrás
- 5 <120> Métodos de modificación de anticuerpos, y anticuerpos modificados con propiedades funcionales mejoradas
- <130> P106700PC00
- <160> 12
- <170> PatentIn versión 3.5
- <210> 1
- 10 <211> 147
- <212> PRT
- <213> Secuencia artificial
- <220>
- <223> Figura 9 – cadena pesada de la familia de VH1
- 15 <220>
- <221> misc\_feature
- <222> (1)..(1)
- <223> Xaa es Gln o Glu
- <220>
- 20 <221> misc\_feature
- <222> (6)..(6)
- <223> Xaa es Gln o Glu
- <220>
- <221> misc\_feature
- 25 <222> (11)..(11)
- <223> Xaa es Val o Leu
- <220>
- <221> misc\_feature
- <222> (12)..(12)
- 30 <223> Xaa es Lys o Met
- <220>
- <221> misc\_feature
- <222> (13)..(13)
- <223> Xaa es Lys, Glu o Gln
- 35 <220>
- <221> misc\_feature
- <222> (18)..(18)

<223> Xaa es Val o Leu  
<220>  
<221> misc\_feature  
<222> (20)..(20)

5 <223> Xaa es Val o Ile  
<220>  
<221> misc\_feature  
<222> (26)..(40)  
<223> Xaa puede ser cualquier aminoácido que se produzca de manera natural

10 <220>  
<221> misc\_feature  
<222> (30)..(40)  
<223> CDR H1  
<220>

15 <221> misc\_feature  
<222> (55)..(74)  
<223> Xaa puede ser cualquier aminoácido que se produzca de manera natural  
<220>  
<221> misc\_feature

20 <222> (55)..(74)  
<223> CDR H2  
<220>  
<221> misc\_feature  
<222> (88)..(88)

25 <223> Xaa es Tyr, Phe, Ser, His o Asp  
<220>  
<221> misc\_feature  
<222> (90)..(90)  
<223> Xaa es Glu, Asp, Gln

30 <220>  
<221> misc\_feature  
<222> (93)..(93)  
<223> Xaa es Ser, Gly, Thr, Asn  
<220>

35 <221> misc\_feature  
<222> (96)..(96)  
<223> Xaa es Ser, Phe, Thr, Ala, Pro

<220>

<221> misc\_feature

<222> (107)..(136)

<223> Xaa puede ser cualquier aminoácido que se produzca de manera natural

5 <220>

<221> misc\_feature

<222> (107)..(136)

<223> CDR H3

<400> 1

Xaa Val Gln Leu Val Xaa Ser Gly Ala Glu Xaa Xaa Xaa Pro Gly Ser  
1 5 10 15

Ser Xaa Lys Xaa Ser Cys Lys Ala Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
20 25 30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Trp Val Arg Gln Ala Pro Gly Gln  
35 40 45

Gly Leu Glu Trp Met Gly Xaa  
50 55 60

Xaa Arg Val Thr Ile Thr Ala  
65 70 75 80

Asp Glu Ser Thr Ser Thr Ala Xaa Met Xaa Leu Ser Xaa Leu Arg Xaa  
85 90 95

Glu Asp Thr Ala Val Tyr Tyr Cys Ala Arg Xaa Xaa Xaa Xaa Xaa Xaa  
100 105 110

10

Xaa  
115 120 125

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Trp Gly Gln Gly Thr Leu Val Thr  
130 135 140

Val Ser Ser  
145

<210> 2

<211> 147

<212> PRT

15 <213> Secuencia artificial

<220>

<223> Figura 10 – cadena pesada de la familia de VH1B

<220>

- <221> misc\_feature  
<222> (1)..(1)  
<223> Xaa es Gln o Glu  
<220>
- 5 <221> misc\_feature  
<222> (9)..(9)  
<223> Xaa es Ala, Thr, Asp, Pro o Val  
<220>  
<221> misc\_feature
- 10 <222> (11)..(11)  
<223> Xaa es Val o Leu  
<220>  
<221> misc\_feature  
<222> (12)..(12)
- 15 <223> Xaa es Lys, Val, Arg, Gln, Met  
<220>  
<221> misc\_feature  
<222> (13)..(13)  
<223> Xaa es Lys, Arg, Glu o Met
- 20 <220>  
<221> misc\_feature  
<222> (19)..(19)  
<223> Xaa es Lys, Asn, Arg o Thr  
<220>
- 25 <221> misc\_feature  
<222> (20)..(20)  
<223> Xaa es Val, Leu, Ile o Phe  
<220>  
<221> misc\_feature
- 30 <222> (26)..(40)  
<223> Xaa puede ser cualquier aminoácido que se produzca de manera natural  
<220>  
<221> misc\_feature  
<222> (30)..(40)
- 35 <223> CDR H1  
<220>  
<221> misc\_feature

- <222> (43)..(43)  
 <223> Xaa es Arg o Lys  
 <220>  
 <221> misc\_feature
- 5 <222> (45)..(45)  
 <223> Xaa es Ala, Arg, Thr, Val o Pro  
 <220>  
 <221> misc\_feature  
 <222> (48)..(48)
- 10 <223> Xaa es Gln, Lys, Glu o His  
 <220>  
 <221> misc\_feature  
 <222> (53)..(53)  
 <223> Xaa es Met o Ile
- 15 <220>  
 <221> misc\_feature  
 <222> (55)..(74)  
 <223> Xaa puede ser cualquier aminoácido que se produzca de manera natural  
 <220>
- 20 <221> misc\_feature  
 <222> (55)..(74)  
 <223> CDR H2  
 <220>  
 <221> misc\_feature
- 25 <222> (75)..(75)  
 <223> Xaa es Arg o Lys  
 <220>  
 <221> misc\_feature  
 <222> (76)..(76)
- 30 <223> Xaa es Val, Ala, Ile o Leu  
 <220>  
 <221> misc\_feature  
 <222> (105)..(105)  
 <223> Xaa es Ala, Asn o Ser
- 35 <220>  
 <221> misc\_feature  
 <222> (107)..(136)

ES 2 566 737 T3

<223> Xaa puede ser cualquier aminoácido que se produzca de manera natural

<220>

<221> misc\_feature

<222> (107)..(136)

5 <223> CDR H3

<400> 2

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

<210> 3

10 <211> 147

<212> PRT

<213> Desconocida

<220>

<223> Figura 11 – cadena pesada de la familia de VH3

15 <220>

<221> misc\_feature

<222> (1)..(1)

<223> Xaa es Gln o Glu

<220>

- <221> misc\_feature  
<222> (6)..(6)  
<223> Xaa es Glu o Gln  
<220>
- 5 <221> misc\_feature  
<222> (7)..(7)  
<223> Xaa es Ser, Thr o Ala  
<220>  
<221> misc\_feature
- 10 <222> (27)..(40)  
<223> Xaa puede ser cualquier aminoácido que se produzca de manera natural  
<220>  
<221> misc\_feature  
<222> (30)..(40)
- 15 <223> CDR H1  
<220>  
<221> misc\_feature  
<222> (55)..(74)  
<223> Xaa puede ser cualquier aminoácido que se produzca de manera natural
- 20 <220>  
<221> misc\_feature  
<222> (55)..(74)  
<223> CDR H2  
<220>
- 25 <221> misc\_feature  
<222> (87)..(87)  
<223> Xaa es Leu, Val, Ala o Phe  
<220>  
<221> misc\_feature
- 30 <222> (101)..(101)  
<223> Xaa es Val, Leu, Ile, Met, Phe, Arg o Gln  
<220>  
<221> misc\_feature  
<222> (107)..(136)
- 35 <223> Xaa puede ser cualquier aminoácido que se produzca de manera natural  
<220>  
<221> misc\_feature

ES 2 566 737 T3

<222> (107)..(136)

<223> CDR H3

<400> 3

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

5 <210> 4

<211> 149

<212> PRT

<213> Desconocida

<220>

10 <223> Figura 12 – cadena pesada de la familia de Vk1

<220>

<221> misc\_feature

<222> (1)..(1)

<223> Xaa es Glu, Asp o Ile

15 <220>

<221> misc\_feature

<222> (3)..(3)

- <223> Xaa es Val, Gln o Ile  
 <220>  
 <221> misc\_feature  
 <222> (4)..(4)
- 5 <223> Xaa es Leu, Met, Val o Ile  
 <220>  
 <221> misc\_feature  
 <222> (24)..(24)  
 <223> Xaa es Gln o Arg
- 10 <220>  
 <221> misc\_feature  
 <222> (25)..(42)  
 <223> CDR L1; Xaa puede ser cualquier aminoácido que se produzca de manera natural  
 <220>
- 15 <221> misc\_feature  
 <222> (47)..(47)  
 <223> Xaa es Arg, Lys o Ile  
 <220>  
 <221> misc\_feature
- 20 <222> (50)..(50)  
 <223> Xaa es Lys, Arg, Glu, Thr, Met o Gln  
 <220>  
 <221> misc\_feature  
 <222> (57)..(57)
- 25 <223> Xaa es Ser, Tyr, Phe o His  
 <220>  
 <221> misc\_feature  
 <222> (58)..(72)  
 <223> CDR L2; Xaa puede ser cualquier aminoácido que se produzca de manera natural
- 30 <220>  
 <221> misc\_feature  
 <222> (85)..(86)  
 <223> Xaa puede ser cualquier aminoácido que se produzca de manera natural  
 <220>
- 35 <221> misc\_feature  
 <222> (91)..(91)  
 <223> Xaa es Phe o Leu

<220>

<221> misc\_feature

<222> (103)..(103)

<223> Xaa es Val, Thr, Ser, Gly o Ile

5 <220>

<221> misc\_feature

<222> (107)..(138)

<223> CDR L3; Xaa puede ser cualquier aminoácido que se produzca de manera natural

<400> 4

Asp Arg Val Thr Ile Thr Cys Xaa  
20 25 30

Xaa Trp Tyr Gln Gln Xaa Pro  
35 40 45

Gly Xaa Ala Pro Lys Leu Leu Ile Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
50 55 60

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gly Val Pro Ser Arg Phe Ser Gly  
65 70 75 80

Ser Gly Ser Gly Xaa Xaa Thr Asp Phe Thr Xaa Thr Ile Ser Ser Leu  
85 90 95

Gln Pro Glu Asp Phe Ala Xaa Tyr Tyr Cys Xaa Xaa Xaa Xaa Xaa Xaa  
100 105 110

Xaa  
115 120 125

Xaa Phe Gly Gln Gly Thr Lys  
130 135 140

10 Val Glu Ile Lys Arg  
145

<210> 5

<211> 149

<212> PRT

<213> Desconocida

15 <220>

<223> Figura 13 – cadena ligera de la familia de Vk3

<220>

<221> misc\_feature

<222> (2)..(2)

20 <223> Xaa es Thr o Ile

- <220>  
<221> misc\_feature  
<222> (3)..(3)  
<223> Xaa es Thr o Val
- 5 <220>  
<221> misc\_feature  
<222> (10)..(10)  
<223> Xaa es Ile o Thr  
<220>
- 10 <221> misc\_feature  
<222> (12)..(12)  
<223> Xaa es Tyr o Ser  
<220>  
<221> misc\_feature
- 15 <222> (18)..(18)  
<223> Xaa es Ser o Arg  
<220>  
<221> misc\_feature  
<222> (20)..(20)
- 20 <223> Xaa es Ala o Thr  
<220>  
<221> misc\_feature  
<222> (25)..(42)  
<223> CDR L1; Xaa puede ser cualquier aminoácido que se produzca de manera natural
- 25 <220>  
<221> misc\_feature  
<222> (56)..(56)  
<223> Xaa es Met o Ile  
<220>
- 30 <221> misc\_feature  
<222> (58)..(72)  
<223> CDR L2; Xaa puede ser cualquier aminoácido que se produzca de manera natural  
<220>  
<221> misc\_feature
- 35 <222> (74)..(74)  
<223> Xaa es Thr, Val o Ile  
<220>

ES 2 566 737 T3

<221> misc\_feature

<222> (85)..(86)

<223> Xaa puede ser cualquier aminoácido que se produzca de manera natural

<220>

5 <221> misc\_feature

<222> (94)..(94)

<223> Xaa es Asn o Ser

<220>

<221> misc\_feature

10 <222> (101)..(101)

<223> Xaa es Ser, Tyr o Phe

<220>

<221> misc\_feature

<222> (103)..(103)

15 <223> Xaa es Ala, Leu o Val

<220>

<221> misc\_feature

<222> (107)..(138)

<223> CDR L3; Xaa puede ser cualquier aminoácido que se produzca de manera natural

20 <400> 5

Glu Xaa Xaa Leu Thr Gln Ser Pro Gly Xaa Leu Xaa Leu Ser Pro Gly  
 1 5 10 15

Glu Xaa Ala Xaa Leu Ser Cys Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
 20 25 30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Trp Tyr Gln Gln Lys Pro  
 35 40 45

Gly Gln Ala Pro Arg Leu Leu Xaa Tyr Xaa Xaa Xaa Xaa Xaa Xaa

ES 2 566 737 T3

50 55 60

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gly Xaa Pro Asp Arg Phe Ser Gly  
65 70 75 80

Ser Gly Ser Gly Xaa Xaa Thr Asp Phe Thr Leu Thr Ile Xaa Arg Leu  
85 90 95

Glu Pro Glu Asp Xaa Ala Xaa Tyr Tyr Cys Xaa Xaa Xaa Xaa Xaa Xaa  
100 105 110

Xaa  
115 120 125

Xaa Phe Gly Gly Gly Thr Lys  
130 135 140

Leu Glu Ile Lys Arg  
145

<210> 6

<211> 149

<212> PRT

5 <213> Desconocida

<220>

<223> Figura 14 – cadena ligera de la familia de VL1

<220>

<221> misc\_feature

10 <222> (1)..(1)

<223> Xaa es Leu, Gln, Ser o Glu

<220>

<221> misc\_feature

<222> (2)..(2)

15 <223> Xaa es Ser, Ala, Pro, Ile o Thr

<220>

<221> misc\_feature

<222> (4)..(4)

<223> Xaa es Val, Leu o Met

20 <220>

<221> misc\_feature

<222> (7)..(7)

<223> Xaa es Ser, Glu o Pro

<220>

25 <221> misc\_feature

- <222> (8)..(8)  
<223> Xaa puede ser cualquier aminoácido que se produzca de manera natural  
<220>  
<221> misc\_feature
- 5 <222> (11)..(11)  
<223> Xaa es Ala o Val  
<220>  
<221> misc\_feature  
<222> (14)..(14)
- 10 <223> Xaa es Thr, Ser o Ala  
<220>  
<221> misc\_feature  
<222> (25)..(42)  
<223> CDR L1; Xaa puede ser cualquier aminoácido que se produzca de manera natural
- 15 <220>  
<221> misc\_feature  
<222> (46)..(46)  
<223> Xaa es His o Gln  
<220>
- 20 <221> misc\_feature  
<222> (53)..(53)  
<223> Xaa es Thr, Lys, Ser, Asn, Gln o Pro  
<222> (58)..(72)  
<223> CDR L2; Xaa puede ser cualquier aminoácido que se produzca de manera natural
- 25 <220>  
<221> misc\_feature  
<222> (82)..(82)  
<223> Xaa es Arg, Gln o Lys  
<220>
- 30 <221> misc\_feature  
<222> (85)..(86)  
<223> Xaa puede ser cualquier aminoácido que se produzca de manera natural  
<220>  
<221> misc\_feature
- 35 <222> (92)..(92)  
<223> Xaa es Thr, Gly, Asp o Ala  
<220>

<221> misc\_feature

<222> (103)..(103)

<223> Xaa es Val, Asp, Thr, His o Glu

<220>

5 <221> misc\_feature

<222> (107)..(138)

<223> CDR L3; Xaa puede ser cualquier aminoácido que se produzca de manera natural

<400> 6

Xaa Xaa Val Xaa Thr Gln Xaa Xaa Pro Ser Xaa Ser Gly Xaa Pro Gly  
1 5 10 15

Gln Arg Val Thr Ile Ser Cys Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
20 25 30

Xaa Trp Tyr Gln Xaa Leu Pro  
35 40 45

Gly Thr Ala Pro Xaa Leu Leu Ile Tyr Xaa Xaa Xaa Xaa Xaa Xaa  
50 55 60

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gly Val Pro Asp Arg Phe Ser Gly  
65 70 75 80

Ser Xaa Ser Gly Xaa Xaa Thr Ser Ala Ser Leu Xaa Ile Ser Gly Leu  
85 90 95

Gln Ser Glu Asp Glu Ala Xaa Tyr Tyr Cys Xaa Xaa Xaa Xaa Xaa Xaa  
100 105 110

Xaa  
115 120 125

Xaa Phe Gly Gly Gly Thr Lys  
130 135 140

Leu Thr Val Leu Gly  
145

10

<210> 7

<211> 236

<212> PRT

<213> Desconocida

15 <220>

<223> Figura 9 – cadena pesada de la familia de VH1

<220>

<221> misc\_feature

<222> (1)..(1)

- <223> Xaa es Gln o Glu  
 <220>  
 <221> misc\_feature  
 <222> (6)..(6)
- 5 <223> Xaa es Gln o Glu  
 <220>  
 <221> misc\_feature  
 <222> (11)..(11)  
 <223> Xaa es Val o Leu
- 10 <220>  
 <221> misc\_feature  
 <222> (12)..(12)  
 <223> Xaa es Lys o Met  
 <220>
- 15 <221> misc\_feature  
 <222> (13)..(13)  
 <223> Xaa es Lys, Glu o Gln  
 <220>  
 <221> misc\_feature
- 20 <222> (18)..(18)  
 <223> Xaa es Val o Leu  
 <220>  
 <221> misc\_feature  
 <222> (20)..(20)
- 25 <223> Xaa es Val o Ile  
 <220>  
 <221> misc\_feature  
 <222> (26)..(29)  
 <223> Xaa puede ser cualquier aminoácido que se produzca de manera natural
- 30 <220>  
 <221> misc\_feature  
 <222> (30)..(79)  
 <223> CDR H1; al menos 3 y hasta 50 aminoácidos pueden estar presentes o ausentes; si están presentes, Xaa puede ser cualquier aminoácido que se produzca de manera natural
- 35 <220>  
 <221> misc\_feature  
 <222> (94)..(143)

ES 2 566 737 T3

<223> CDR H2; al menos 3 y hasta 50 aminoácidos pueden estar presentes o ausentes; si están presentes, Xaa puede ser cualquier aminoácido que se produzca de manera natural

<220>

<221> misc\_feature

5 <222> (157)..(157)

<223> Xaa es Tyr, Phe, Ser, His o Asp

<220>

<221> misc\_feature

<222> (159)..(159)

10 <223> Xaa es Glu, Asp, Gln

<220>

<221> misc\_feature

<222> (162)..(162)

<223> Xaa es Ser, Gly, Thr, Asn

15 <220>

<221> misc\_feature

<222> (165)..(165)

<223> Xaa es Ser, Phe, Thr, Ala, Pro

<220>

20 <221> misc\_feature

<222> (176)..(225)

<223> CDR H3; al menos 3 y hasta 50 aminoácidos pueden estar presentes o ausentes; si están presentes, Xaa puede ser cualquier aminoácido que se produzca de manera natural

<400> 7

Xaa Val Gln Leu Val Xaa Ser Gly Ala Glu Xaa Xaa Xaa Pro Gly Ser  
1 5 10 15

Ser Xaa Lys Xaa Ser Cys Lys Ala Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
20 25 30

Xaa  
35 40 45

Xaa  
50 55 60

Xaa Trp  
65 70 75 80

Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met Gly Xaa Xaa Xaa  
85 90 95

25 Xaa Xaa



- <221> misc\_feature  
 <222> (12)..(12)  
 <223> Xaa es Lys, Val, Arg, Gln, Met  
 <220>
- 5 <221> misc\_feature  
 <222> (13)..(13)  
 <223> Xaa es Lys, Arg, Glu o Met  
 <220>
- <221> misc\_feature  
 10 <222> (19)..(19)  
 <223> Xaa es Lys, Asn, Arg o Thr  
 <220>
- <221> misc\_feature  
 <222> (20)..(20)  
 15 <223> Xaa es Val, Leu, Ile o Phe  
 <220>
- <221> misc\_feature  
 <222> (26)..(29)  
 <223> Xaa puede ser cualquier aminoácido que se produzca de manera natural  
 20 <220>
- <221> misc\_feature  
 <222> (30)..(79)  
 <223> CDR H1; al menos 3 y hasta 50 aminoácidos pueden estar presentes o ausentes; si están presentes, Xaa puede ser cualquier aminoácido que se produzca de manera natural  
 25 <220>
- <221> misc\_feature  
 <222> (82)..(82)  
 <223> Xaa es Arg o Lys  
 <220>
- <221> misc\_feature  
 30 <222> (84)..(84)  
 <223> Xaa es Ala, Arg, Thr, Val o Pro  
 <220>
- <221> misc\_feature  
 35 <222> (87)..(87)  
 <223> Xaa es Gln, Lys, Glu o His  
 <220>

- <221> misc\_feature  
 <222> (92)..(92)  
 <223> Xaa es Met o Ile  
 <220>
- 5 <221> misc\_feature  
 <222> (94)..(143)  
 <223> CDR H2; al menos 3 y hasta 50 aminoácidos pueden estar presentes o ausentes; si están presentes, Xaa puede ser cualquier aminoácido que se produzca de manera natural  
 <220>
- 10 <221> misc\_feature  
 <222> (144)..(144)  
 <223> Xaa es Arg o Lys  
 <220>  
 <221> misc\_feature
- 15 <222> (145)..(145)  
 <223> Xaa es Val, Ala, Ile o Leu  
 <220>  
 <221> misc\_feature  
 <222> (149)..(149)
- 20 <223> Xaa es Glu, Arg, Thr, Ala  
 <220>  
 <221> misc\_feature  
 <222> (153)..(153)  
 <223> Xaa es Thr, Ser, Ile, Leu
- 25 <220>  
 <221> misc\_feature  
 <222> (154)..(154)  
 <223> Xaa es Asp, Ser, Asn, Gly  
 <220>
- 30 <221> misc\_feature  
 <222> (174)..(174)  
 <223> Xaa es Ala, Asn o Ser  
 <220>  
 <221> misc\_feature
- 35 <222> (176)..(225)  
 <223> CDR H3; al menos 3 y hasta 50 aminoácidos pueden estar presentes o ausentes; si están presentes, Xaa puede ser cualquier aminoácido que se produzca de manera natural

ES 2 566 737 T3

<400> 8

Xaa Val Gln Leu Val Gln Ser Gly Xaa Glu Xaa Xaa Xaa Pro Gly Ala  
 1 5 10 15

Ser Val Xaa Xaa Ser Cys Lys Ala Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
 20 25 30

Xaa  
 35 40 45

Xaa  
 50 55 60

Xaa Trp  
 65 70 75 80

Val Xaa Gln Xaa Pro Gly Xaa Gly Leu Glu Trp Xaa Gly Xaa Xaa Xaa  
 85 90 95

Xaa  
 100 105 110

Xaa  
 115 120 125

Xaa  
 130 135 140

Xaa Thr Met Thr Xaa Asp Thr Ser Xaa Xaa Thr Ala Tyr Met Glu Leu  
 145 150 155 160 165

Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys Xaa Arg Xaa  
 165 170 175

Xaa  
 180 185 190

Xaa  
 195 200 205

Xaa  
 210 215 220

Xaa Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser  
 225 230 235

<210> 9

5 <211> 236

<212> PRT

<213> Desconocida

<220>

- <223> Figura 11 – cadena pesada de la familia de VH3
- <220>
- <221> misc\_feature
- <222> (1)..(1)
- 5 <223> Xaa es Glu o Gln
- <220>
- <221> misc\_feature
- <222> (6)..(6)
- <223> Xaa es Glu o Gln
- 10 <220>
- <221> misc\_feature
- <222> (7)..(7)
- <223> Xaa es Ser, Thr o Ala
- <220>
- 15 <221> misc\_feature
- <222> (27)..(29)
- <223> Xaa puede ser cualquier aminoácido que se produzca de manera natural
- <220>
- <221> misc\_feature
- 20 <222> (30)..(79)
- <223> CDR H1; al menos 3 y hasta 50 aminoácidos pueden estar presentes o ausentes; si están presentes, Xaa puede ser cualquier aminoácido que se produzca de manera natural
- <220>
- <221> misc\_feature
- 25 <222> (94)..(143)
- <223> CDR H2; al menos 3 y hasta 50 aminoácidos pueden estar presentes o ausentes; si están presentes, Xaa puede ser cualquier aminoácido que se produzca de manera natural
- <220>
- <221> misc\_feature
- 30 <222> (156)..(156)
- <223> Xaa es Leu, Val, Ala o Phe
- <220>
- <221> misc\_feature
- <222> (170)..(170)
- 35 <223> Xaa es Val, Leu, Ile, Met, Phe, Arg o Gln
- <220>
- <221> misc\_feature

ES 2 566 737 T3

<222> (176)..(225)

<223> CDR H3; al menos 3 y hasta 50 aminoácidos pueden estar presentes o ausentes; si están presentes, Xaa puede ser cualquier aminoácido que se produzca de manera natural

<400> 9

Xaa Val Gln Leu Val Xaa Xaa Gly Pro Gly Leu Val Lys Pro Ser Glu  
1 5 10 15

Thr Leu Arg Leu Ser Cys Ala Ala Ser Gly Xaa Xaa Xaa Xaa Xaa Xaa  
20 25 30

Xaa  
35 40 45

Xaa  
50 55 60

Xaa Trp  
65 70 75 80

Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val Ser Xaa Xaa Xaa  
85 90 95

Xaa  
100 105 110

Xaa  
115 120 125

Xaa Arg  
130 135 140

Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Xaa Tyr Leu Gln Met  
145 150 155 160

Asn Ser Leu Arg Ala Glu Asp Thr Ala Xaa Tyr Tyr Cys Ala Arg Xaa  
165 170 175

Xaa  
180 185 190

Xaa  
195 200 205

Xaa  
210 215 220

Xaa Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser  
225 230 235

5

<210> 10

- <211> 234  
 <212> PRT  
 <213> Desconocida  
 <220>
- 5 <223> Figura 12 – cadena pesada de la familia de Vk1  
 <220>  
 <221> misc\_feature  
 <222> (1)..(1)  
 <223> Xaa es Glu, Asp o Ile
- 10 <220>  
 <221> misc\_feature  
 <222> (3)..(3)  
 <223> Xaa es Val, Gln o Ile  
 <220>
- 15 <221> misc\_feature  
 <222> (4)..(4)  
 <223> Xaa es Leu, Met, Val o Ile  
 <220>  
 <221> misc\_feature
- 20 <222> (24)..(24)  
 <223> Xaa es Gln o Arg  
 <220>  
 <221> misc\_feature  
 <222> (25)..(74)
- 25 <223> CDR L1; al menos 3 y hasta 50 aminoácidos pueden estar presentes o ausentes; si están presentes, Xaa puede ser cualquier aminoácido que se produzca de manera natural  
 <220>  
 <221> misc\_feature  
 <222> (79)..(79)
- 30 <223> Xaa es Arg, Lys o Ile  
 <220>  
 <221> misc\_feature  
 <222> (82)..(82)  
 <223> Xaa es Lys, Arg, Glu, Thr, Met o Gln
- 35 <220>  
 <221> misc\_feature  
 <222> (89)..(89)

- <223> Xaa es Ser, Tyr, Phe o His  
<220>  
<221> misc\_feature  
<222> (90)..(139)
- 5 <223> CDR L2; al menos 3 y hasta 50 aminoácidos pueden estar presentes o ausentes; si están presentes, Xaa puede ser cualquier aminoácido que se produzca de manera natural <220>  
<221> misc\_feature  
<222> (152)..(153)  
<223> Xaa puede ser cualquier aminoácido que se produzca de manera natural
- 10 <220>  
<221> misc\_feature  
<222> (158)..(158)  
<223> Xaa es Phe o Leu  
<220>
- 15 <221> misc\_feature  
<222> (170)..(170)  
<223> Xaa es Val, Thr, Ser, Gly o Ile  
<220>  
<221> misc\_feature
- 20 <222> (174)..(223)  
<223> CDR L3; al menos 3 y hasta 50 aminoácidos pueden estar presentes o ausentes; si están presentes, Xaa puede ser cualquier aminoácido que se produzca de manera natural  
<400> 10

ES 2 566 737 T3

Xaa Ile Xaa Xaa Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly  
 1 5 10 15

Asp Arg Val Thr Ile Thr Cys Xaa  
 20 25 30

Xaa  
 35 40 45

Xaa  
 50 55 60

Xaa Trp Tyr Gln Gln Xaa Pro  
 65 70 75 80

Gly Xaa Ala Pro Lys Leu Leu Ile Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
 85 90 95

Xaa  
 100 105 110

Xaa  
 115 120 125

Xaa Gly Val Pro Ser Arg  
 130 135 140

Phe Ser Gly Ser Gly Ser Gly Xaa Xaa Thr Asp Phe Thr Xaa Thr Ile  
 145 150 155 160

Ser Ser Leu Gln Pro Glu Asp Phe Ala Xaa Tyr Tyr Cys Xaa Xaa Xaa  
 165 170 175

Xaa  
 180 185 190

Xaa  
 195 200 205

Xaa Phe  
 210 215 220

Gly Gln Gly Thr Lys Val Glu Ile Lys Arg  
 225 230

<210> 11

<211> 234

<212> PRT

5 <213> Desconocida

<220>

<223> Figura 13 – cadena ligera de la familia de Vk3

- <220>  
 <221> misc\_feature  
 <222> (2)..(2)  
 <223> Xaa es Thr o Ile
- 5 <220>  
 <221> misc\_feature  
 <222> (3)..(3)  
 <223> Xaa es Thr o Val
- <220>
- 10 <221> misc\_feature  
 <222> (10)..(10)  
 <223> Xaa es Ile o Thr
- <220>
- <221> misc\_feature
- 15 <222> (12)..(12)  
 <223> Xaa es Tyr o Ser
- <220>
- <221> misc\_feature  
 <222> (18)..(18)
- 20 <223> Xaa es Ser o Arg
- <220>
- <221> misc\_feature  
 <222> (20)..(20)  
 <223> Xaa es Ala o Thr
- 25 <220>  
 <221> misc\_feature  
 <222> (25)..(74)  
 <223> CDR L1; al menos 3 y hasta 50 aminoácidos pueden estar presentes o ausentes; si están presentes, Xaa puede ser cualquier aminoácido que se produzca de manera natural
- 30 <220>  
 <221> misc\_feature  
 <222> (88)..(88)  
 <223> Xaa es Met o Ile
- <220>
- 35 <221> misc\_feature  
 <222> (90)..(139)  
 <223> CDR L2; al menos 3 y hasta 50 aminoácidos pueden estar presentes o ausentes; si están presentes, Xaa

puede ser cualquier aminoácido que se produzca de manera natural

<220>

<221> misc\_feature

<222> (141)..(141)

5 <223> Xaa es Thr, Val o Ile

<220>

<221> misc\_feature

<222> (152)..(153)

<223> Xaa puede ser cualquier aminoácido que se produzca de manera natural

10 <220>

<221> misc\_feature

<222> (161)..(161)

<223> Xaa es Asn o Ser

<220>

15 <221> misc\_feature

<222> (168)..(168)

<223> Xaa es Ser, Tyr o Phe

<220>

<221> misc\_feature

20 <222> (170)..(170)

<223> Xaa es Ala, Leu o Val

<220>

<221> misc\_feature

<222> (174)..(223)

25 <223> CDR L3; al menos 3 y hasta 50 aminoácidos pueden estar presentes o ausentes; si están presentes, Xaa puede ser cualquier aminoácido que se produzca de manera natural

<400> 11

ES 2 566 737 T3

Glu Xaa Xaa Leu Thr Gln Ser Pro Gly Xaa Leu Xaa Leu Ser Pro Gly  
 1 5 10 15

Glu Xaa Ala Xaa Leu Ser Cys Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
 20 25 30

Xaa  
 35 40 45

Xaa  
 50 55 60

Xaa Trp Tyr Gln Gln Lys Pro  
 65 70 75 80

Gly Gln Ala Pro Arg Leu Leu Xaa Tyr Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
 85 90 95

Xaa  
 100 105 110

Xaa  
 115 120 125

Xaa Gly Xaa Pro Asp Arg  
 130 135 140

Phe Ser Gly Ser Gly Ser Gly Xaa Xaa Thr Asp Phe Thr Leu Thr Ile  
 145 150 155 160

Xaa Arg Leu Glu Pro Glu Asp Xaa Ala Xaa Tyr Tyr Cys Xaa Xaa Xaa  
 165 170 175

Xaa  
 180 185 190

Xaa  
 195 200 205

Xaa Phe  
 210 215 220

Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg  
 225 230

<210> 12

<211> 234

5 <212> PRT

<213> Desconocida

<220>

<223> Figura 14 – cadena ligera de la familia de VL1

<220>

10 <221> misc\_feature

- <222> (1)..(1)  
<223> Xaa es Leu, Gln, Ser o Glu  
<220>  
<221> misc\_feature
- 5 <222> (2)..(2)  
<223> Xaa es Ser, Ala, Pro, Ile o Thr  
<220>  
<221> misc\_feature  
<222> (4)..(4)
- 10 <223> Xaa es Val, Leu o Met  
<220>  
<221> misc\_feature  
<222> (7)..(7)  
<223> Xaa es Ser, Glu o Pro
- 15 <220>  
<221> misc\_feature  
<222> (8)..(8)  
<223> Xaa puede ser cualquier aminoácido que se produzca de manera natural  
<220>
- 20 <221> misc\_feature  
<222> (11)..(11)  
<223> Xaa es Ala o Val  
<220>  
<221> misc\_feature
- 25 <222> (14)..(14)  
<223> Xaa es Thr, Ser o Ala  
<220>  
<221> misc\_feature  
<222> (25)..(74)
- 30 <223> CDR L1; al menos 3 y hasta 50 aminoácidos pueden estar presentes o ausentes; si están presentes, Xaa puede ser cualquier aminoácido que se produzca de manera natural  
<220>  
<221> misc\_feature  
<222> (78)..(78)
- 35 <223> Xaa es His o Gln  
<220>  
<221> misc\_feature

<222> (85)..(85)  
 <223> Xaa es Thr, Lys, Ser, Asn, Gln o Pro  
 <222> (90)..(139)  
 5 <223> CDR L2; al menos 3 y hasta 50 aminoácidos pueden estar presentes o ausentes; si están presentes, Xaa puede ser cualquier aminoácido que se produzca de manera natural  
 <220>  
 <221> misc\_feature  
 <222> (149)..(149)  
 <223> Xaa es Arg, Gln o Lys  
 10 <220>  
 <221> misc\_feature  
 <222> (152)..(153)  
 <223> Xaa puede ser cualquier aminoácido que se produzca de manera natural  
 <220>  
 15 <221> misc\_feature  
 <222> (159)..(159)  
 <223> Xaa es Thr, Gly, Asp o Ala  
 <220>  
 <221> misc\_feature  
 20 <222> (170)..(170)  
 <223> Xaa es Val, Asp, Thr, His o Glu  
 <220>  
 <221> misc\_feature  
 <222> (174)..(223)  
 25 <223> CDR L3; al menos 3 y hasta 50 aminoácidos pueden estar presentes o ausentes; si están presentes, Xaa puede ser cualquier aminoácido que se produzca de manera natural  
 <400> 12

ES 2 566 737 T3

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

**REIVINDICACIONES**

1. Método de modificación mediante ingeniería genética de un agente de unión inmunológica, comprendiendo el agente de unión inmunológica (i) una región variable de cadena pesada, o fragmento de la misma, comprendiendo la región variable de cadena pesada residuos de región de marco de V<sub>H</sub> y/o (ii) una región variable de cadena ligera, o fragmento de la misma, comprendiendo la región variable de cadena ligera residuos de región de marco de V<sub>L</sub>, comprendiendo el método:
- 5
- A) seleccionar una o más posiciones de aminoácido dentro de los residuos de región de marco de V<sub>H</sub>, los residuos de región de marco de V<sub>L</sub> o los residuos de región de marco de V<sub>H</sub> y V<sub>L</sub> para la mutación; y
- B) mutar la una o más posiciones de aminoácido seleccionadas para la mutación,
- 10
- a) en el que si la una o más posiciones de aminoácido de cadena pesada seleccionadas para la mutación son de una región variable de cadena pesada de la familia de VH3 humana, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:
- (i) glutamina (Q) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;
- (ii) glutamina (Q) en la posición de aminoácido 6 usando el sistema de numeración de AHo o de Kabat;
- 15
- (iii) treonina (T) o alanina (A) en la posición de aminoácido 7 usando el sistema de numeración de AHo o de Kabat;
- (iv) alanina (A), valina (V) o fenilalanina (F) en la posición de aminoácido 89 usando el sistema de numeración de AHo (posición de aminoácido 78 usando el sistema de numeración de Kabat); y
- 20
- (v) arginina (R), glutamina (Q), isoleucina (I), leucina (L), metionina (M) o fenilalanina (F) en la posición de aminoácido 103 usando el sistema de numeración de AHo (posición de aminoácido 89 usando la numeración de Kabat);
- b) en el que si la una o más posiciones de aminoácido de cadena pesada seleccionadas para la mutación son de una región variable de cadena pesada de la familia de VH1a humana, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:
- 25
- (i) ácido glutámico (E) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;
- (ii) ácido glutámico (E) en la posición de aminoácido 6 usando el sistema de numeración de AHo o de Kabat;
- 30
- (iii) leucina (L) en la posición de aminoácido 12 usando el sistema de numeración de AHo (posición de aminoácido 11 usando el sistema de numeración de Kabat);
- (iv) metionina (M) en la posición de aminoácido 13 usando el sistema de numeración de AHo (posición de aminoácido 12 usando el sistema de numeración de Kabat);
- 35
- (v) ácido glutámico (E) o glutamina (Q) en la posición de aminoácido 14 usando el sistema de numeración de AHo (posición de aminoácido 13 usando el sistema de numeración de Kabat);
- (vi) leucina (L) en la posición de aminoácido 19 usando el sistema de numeración de AHo (posición de aminoácido 18 usando el sistema de numeración de Kabat);
- (vii) isoleucina (I) en la posición de aminoácido 21 usando el sistema de numeración de AHo (posición de aminoácido 20 usando el sistema de numeración de Kabat);
- 40
- (viii) fenilalanina (F), serina (S), histidina (H) o ácido aspártico (D) en la posición de aminoácido 90 usando el sistema de numeración de AHo (posición de aminoácido 79 usando el sistema de numeración de Kabat);
- (ix) ácido aspártico (D) o glutamina (Q) en la posición de aminoácido 92 usando el sistema de numeración de AHo (posición de aminoácido 81 usando el sistema de numeración de Kabat);
- 45
- (x) glicina (G), asparagina (N) o treonina (T) en la posición de aminoácido 95 usando el sistema de numeración de AHo (posición de aminoácido 82b usando el sistema de numeración de Kabat); y
- (xi) treonina (T), alanina (A), prolina (P) o fenilalanina (F) en la posición de aminoácido 98 usando la numeración de AHo (posición de aminoácido 84 usando la numeración de Kabat);

c) en el que si la una o más posiciones de aminoácido de cadena pesada seleccionadas para la mutación son de una región variable de cadena pesada de la familia de VH1b humana, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

- 5 (i) ácido glutámico (E) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;
- (ii) treonina (T), prolina (P), valina (V) o ácido aspártico (D) en la posición de aminoácido 10 usando el sistema de numeración de AHo (posición de aminoácido 9 usando el sistema de numeración de Kabat);
- 10 (iii) leucina (L) en la posición de aminoácido 12 usando el sistema de numeración de AHo (posición de aminoácido 11 usando el sistema de numeración de Kabat);
- (iv) valina (V), arginina (R), glutamina (Q) o metionina (M) en la posición de aminoácido 13 usando el sistema de numeración de AHo (posición de aminoácido 12 usando el sistema de numeración de Kabat);
- 15 (v) ácido glutámico (E), arginina (R) o metionina (M) en la posición de aminoácido 14 usando el sistema de numeración de AHo (posición de aminoácido 13 usando el sistema de numeración de Kabat);
- (vi) arginina (R), treonina (T) o asparagina (N) en la posición de aminoácido 20 usando el sistema de numeración de AHo (posición de aminoácido 19 usando el sistema de numeración de Kabat);
- 20 (vii) isoleucina (I), fenilalanina (F) o leucina (L) en la posición de aminoácido 21 usando el sistema de numeración de AHo (posición de aminoácido 20 usando el sistema de numeración de Kabat);
- (viii) lisina (K) en la posición de aminoácido 45 usando el sistema de numeración de AHo (posición de aminoácido 38 usando el sistema de numeración de Kabat);
- 25 (ix) treonina (T), prolina (P), valina (V) o arginina (R) en la posición de aminoácido 47 usando el sistema de numeración de AHo (posición de aminoácido 40 usando el sistema de numeración de Kabat);
- (x) lisina (K), histidina (H) o ácido glutámico (E) en la posición de aminoácido 50 usando el sistema de numeración de AHo (posición de aminoácido 43 usando el sistema de numeración de Kabat);
- (xi) isoleucina (I) en la posición de aminoácido 55 usando la numeración de AHo (posición de aminoácido 48 usando la numeración de Kabat);
- 30 (xii) lisina (K) en la posición de aminoácido 77 usando la numeración de AHo (posición de aminoácido 66 usando la numeración de Kabat);
- (xiii) alanina (A), leucina (L) o isoleucina (I) en la posición de aminoácido 78 usando el sistema de numeración de AHo (posición de aminoácido 67 usando el sistema de numeración de Kabat);
- 35 (xiv) ácido glutámico (E), treonina (T) o alanina (A) en la posición de aminoácido 82 usando el sistema de numeración de AHo (posición de aminoácido 71 usando el sistema de numeración de Kabat);
- (xv) treonina (T), serina (S) o leucina (L) en la posición de aminoácido 86 usando el sistema de numeración de AHo (posición de aminoácido 75 usando el sistema de numeración de Kabat);
- 40 (xvi) ácido aspártico (D), asparagina (N) o glicina (G) en la posición de aminoácido 87 usando el sistema de numeración de AHo (posición de aminoácido 76 usando el sistema de numeración de Kabat); y
- (xvii) asparagina (N) o serina (S) en la posición de aminoácido 107 usando el sistema de numeración de AHo (posición de aminoácido 93 usando el sistema de numeración de Kabat);

d) en el que si la una o más posiciones de aminoácido de cadena ligera seleccionadas para la mutación son de una región variable de cadena ligera de la familia de Vκ1 humana, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

- 45 (i) ácido glutámico (E) o isoleucina (I) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;
- (ii) valina (V) o isoleucina (I) en la posición de aminoácido 3 usando el sistema de numeración de AHo o de Kabat;

- (iii) valina (V), leucina (L) o isoleucina (I) en la posición de aminoácido 4 usando el sistema de numeración de AHo o de Kabat;
- (iv) glutamina (Q) en la posición de aminoácido 24 usando el sistema de numeración de AHo o de Kabat;
- 5 (v) arginina (R) o isoleucina (I) en la posición de aminoácido 47 usando el sistema de numeración de AHo (posición de aminoácido 39 usando el sistema de numeración de Kabat);
- (vi) arginina (R), ácido glutámico (E) treonina (T), metionina (M) o glutamina (Q) en la posición de aminoácido 50 usando el sistema de numeración de AHo (posición de aminoácido 42 usando el sistema de numeración de Kabat);
- 10 (vii) histidina (H), serina (S) o fenilalanina (F) en la posición de aminoácido 57 usando el sistema de numeración de AHo (posición de aminoácido 49 usando el sistema de numeración de Kabat);
- (viii) fenilalanina (F) en la posición de aminoácido 91 usando el sistema de numeración de AHo (posición de aminoácido 73 usando el sistema de numeración de Kabat); y
- 15 (ix) valina (V), serina (S), glicina (G) o isoleucina (I) en la posición de aminoácido 103 usando el sistema de numeración de AHo (posición de aminoácido 85 usando el sistema de numeración de Kabat);
- e) en el que si la una o más posiciones de aminoácido de cadena ligera seleccionadas para la mutación son de una región variable de cadena ligera de la familia de V $\kappa$ 3 humana, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:
- 20 (i) treonina (T) en la posición de aminoácido 2 usando el sistema de numeración de AHo o de Kabat;
- (ii) treonina (T) en la posición de aminoácido 3 usando el sistema de numeración de AHo o de Kabat;
- (iii) isoleucina (I) en la posición de aminoácido 10 usando el sistema de numeración de AHo o de Kabat;
- (iv) tirosina (Y) en la posición de aminoácido 12 usando el sistema de numeración de AHo o de Kabat;
- 25 (v) serina (S) en la posición de aminoácido 18 usando el sistema de numeración de AHo o de Kabat;
- (vi) alanina (A) en la posición de aminoácido 20 usando el sistema de numeración de AHo o de Kabat;
- (vii) metionina (M) en la posición de aminoácido 56 usando el sistema de numeración de AHo (posición de aminoácido 48 usando el sistema de numeración de Kabat);
- 30 (viii) valina (V) o treonina (T) en la posición de aminoácido 74 usando el sistema de numeración de AHo (posición de aminoácido 58 usando el sistema de numeración de Kabat);
- (ix) asparagina (N) en la posición de aminoácido 94 usando el sistema de numeración de AHo (posición de aminoácido 76 usando el sistema de numeración de Kabat);
- (x) tirosina (Y) o serina (S) en la posición de aminoácido 101 usando el sistema de numeración de AHo (posición de aminoácido 83 usando el sistema de numeración de Kabat); y
- 35 (xi) leucina (L) o alanina (A) en la posición de aminoácido 103 usando la numeración de AHo (posición de aminoácido 85 usando la numeración de Kabat);
- f) en el que si la una o más posiciones de aminoácido de cadena ligera seleccionadas para la mutación son de una región variable de cadena ligera de la familia de V $\lambda$ 1 humana, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:
- 40 (i) leucina (L), serina (S) o ácido glutámico (E) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;
- (ii) alanina (A), prolina (P), isoleucina (I) o tirosina (Y) en la posición de aminoácido 2 usando el sistema de numeración de AHo o de Kabat;
- 45 (iii) valina (V) o metionina (M) en la posición de aminoácido 4 usando el sistema de numeración de AHo o de Kabat;
- (iv) ácido glutámico (E) en la posición de aminoácido 7 usando el sistema de numeración de AHo o de Kabat;

- (v) alanina (A) en la posición de aminoácido 11 usando el sistema de numeración de AHo o de Kabat;
- (vi) treonina (T) o serina (S) en la posición de aminoácido 14 usando el sistema de numeración de AHo o de Kabat;
- 5 (vii) histidina (H) en la posición de aminoácido 46 usando el sistema de numeración de AHo (posición de aminoácido 38 usando el sistema de numeración de Kabat);
- (viii) treonina (T), serina (S), asparagina (N), glutamina (Q) o prolina (P) en la posición de aminoácido 53 usando el sistema de numeración de AHo (posición de aminoácido 45 usando el sistema de numeración de Kabat);
- 10 (ix) arginina (R) o glutamina (Q) en la posición de aminoácido 82 usando el sistema de numeración de AHo (posición de aminoácido 66 usando el sistema de numeración de Kabat);
- (x) glicina (G), treonina (T) o ácido aspártico (D) en la posición de aminoácido 92 usando el sistema de numeración de AHo (posición de aminoácido 74 usando el sistema de numeración de Kabat); y
- (xi) valina (V), treonina (T), histidina (H) o ácido glutámico (E) en la posición de aminoácido 103 usando la numeración de AHo (posición de aminoácido 85 usando la numeración de Kabat).
- 15 2. Método según la reivindicación 1, en el que la mutación comprende además una o más sustituciones de cadena pesada seleccionadas del grupo que consiste en:
- (i) serina (S) en la posición de aminoácido 12 usando AHo o Kabat;
- (ii) serina (S) en la posición de aminoácido 103 usando la numeración de AHo (posición de aminoácido 85 usando la numeración de Kabat); y
- 20 (iii) serina (S) o treonina (T) en la posición de aminoácido 144 usando la numeración de AHo (posición de aminoácido 103 usando la numeración de Kabat).
3. Método según la reivindicación 1, en el que la mutación comprende además las siguientes sustituciones de cadena pesada:
- (i) serina (S) en la posición de aminoácido 12 usando AHo o Kabat;
- 25 (ii) serina (S) en la posición de aminoácido 103 usando la numeración de AHo (posición de aminoácido 85 usando la numeración de Kabat); y
- (iii) serina (S) o treonina (T) en la posición de aminoácido 144 usando la numeración de AHo (posición de aminoácido 103 usando la numeración de Kabat).
- 30 4. Método según una cualquiera de las reivindicaciones 1-3, en el que el agente de unión inmunológica es un anticuerpo scFv, una inmunoglobulina de longitud completa, un fragmento Fab, un Dab o un Nanobody.
5. Método según la reivindicación 1, en el que si la una o más posiciones de aminoácido de cadena pesada seleccionadas para la mutación son de una región variable de cadena pesada de la familia de VH3, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:
- (i) glutamina (Q) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;
- 35 (ii) glutamina (Q) en la posición de aminoácido 6 usando el sistema de numeración de AHo o de Kabat;
- (iii) treonina (T) en la posición de aminoácido 7 usando el sistema de numeración de AHo o de Kabat;
- (iv) valina (V) en la posición de aminoácido 89 usando el sistema de numeración de AHo (posición de aminoácido 78 usando el sistema de numeración de Kabat); y
- 40 (v) leucina (L) en la posición de aminoácido 103 usando el sistema de numeración de AHo (posición de aminoácido 89 usando la numeración de Kabat).
6. Método según la reivindicación 1, en el que si la una o más posiciones de aminoácido de cadena pesada seleccionadas para la mutación son de una región variable de cadena pesada de la familia de VH1a, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:
- (i) ácido glutámico (E) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;
- 45 (ii) ácido glutámico (E) en la posición de aminoácido 6 usando el sistema de numeración de AHo o de Kabat;

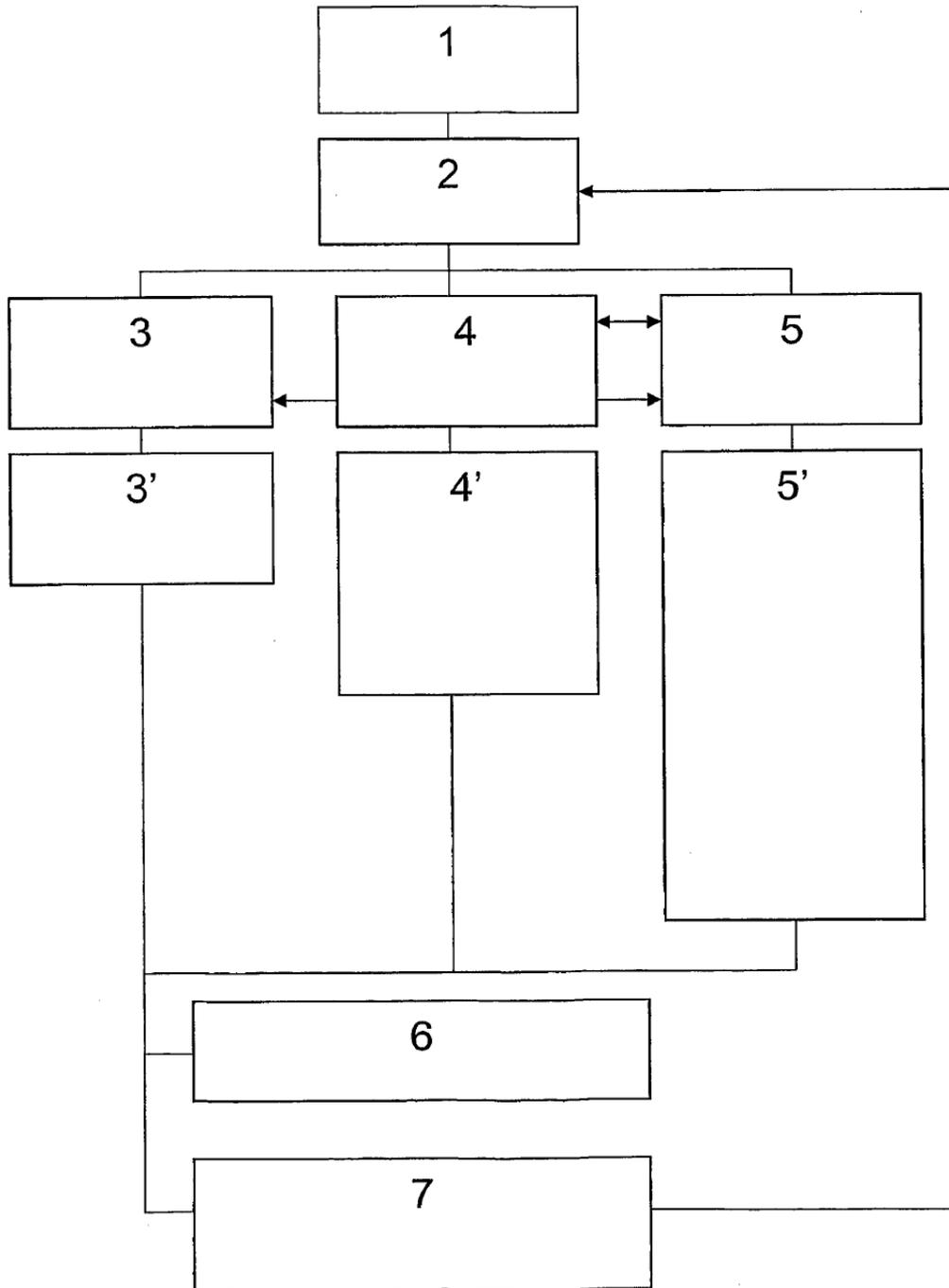
- (iii) leucina (L) en la posición de aminoácido 12 usando el sistema de numeración de AHo (posición de aminoácido 11 usando el sistema de numeración de Kabat);
- (iv) metionina (M) en la posición de aminoácido 13 usando el sistema de numeración de AHo (posición de aminoácido 12 usando el sistema de numeración de Kabat);
- 5 (v) ácido glutámico (E) en la posición de aminoácido 14 usando el sistema de numeración de AHo (posición de aminoácido 13 usando el sistema de numeración de Kabat);
- (vi) leucina (L) en la posición de aminoácido 19 usando el sistema de numeración de AHo (posición de aminoácido 18 usando el sistema de numeración de Kabat);
- 10 (vii) isoleucina (I) en la posición de aminoácido 21 usando el sistema de numeración de AHo (posición de aminoácido 20 usando el sistema de numeración de Kabat);
- (viii) fenilalanina (F), serina (S), histidina (H) o ácido aspártico (D) en la posición de aminoácido 90 usando el sistema de numeración de AHo (posición de aminoácido 79 usando el sistema de numeración de Kabat);
- (ix) ácido aspártico (D) en la posición de aminoácido 92 usando el sistema de numeración de AHo (posición de aminoácido 81 usando el sistema de numeración de Kabat);
- 15 (x) glicina (G) en la posición de aminoácido 95 usando el sistema de numeración de AHo (posición de aminoácido 82b usando el sistema de numeración de Kabat); y
- (xi) fenilalanina (F) en la posición de aminoácido 98 usando la numeración de AHo (posición de aminoácido 84 usando la numeración de Kabat).
7. Método según la reivindicación 1, en el que si la una o más posiciones de aminoácido de cadena pesada seleccionadas para la mutación son de una región variable de cadena pesada de la familia de VH1b, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:
- 20 (i) ácido glutámico (E) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;
- (ii) treonina (T) en la posición de aminoácido 10 usando el sistema de numeración de AHo (posición de aminoácido 9 usando el sistema de numeración de Kabat);
- 25 (iii) leucina (L) en la posición de aminoácido 12 usando el sistema de numeración de AHo (posición de aminoácido 11 usando el sistema de numeración de Kabat);
- (iv) valina (V) en la posición de aminoácido 13 usando el sistema de numeración de AHo (posición de aminoácido 12 usando el sistema de numeración de Kabat);
- 30 (v) arginina (R) en la posición de aminoácido 14 usando el sistema de numeración de AHo (posición de aminoácido 13 usando el sistema de numeración de Kabat);
- (vi) asparagina (N) en la posición de aminoácido 20 usando el sistema de numeración de AHo (posición de aminoácido 19 usando el sistema de numeración de Kabat);
- (vii) leucina (L) en la posición de aminoácido 21 usando el sistema de numeración de AHo (posición de aminoácido 20 usando el sistema de numeración de Kabat);
- 35 (viii) lisina (K) en la posición de aminoácido 45 usando el sistema de numeración de AHo (posición de aminoácido 38 usando el sistema de numeración de Kabat);
- (ix) arginina (R) en la posición de aminoácido 47 usando el sistema de numeración de AHo (posición de aminoácido 40 usando el sistema de numeración de Kabat);
- 40 (x) lisina (K) en la posición de aminoácido 50 usando el sistema de numeración de AHo (posición de aminoácido 43 usando el sistema de numeración de Kabat);
- (xi) isoleucina (I) en la posición de aminoácido 55 usando la numeración de AHo (posición de aminoácido 48 usando la numeración de Kabat);
- (xii) lisina (K) en la posición de aminoácido 77 usando la numeración de AHo (posición de aminoácido 66 usando la numeración de Kabat);
- 45 (xiii) alanina (A) en la posición de aminoácido 78 usando el sistema de numeración de AHo (posición de aminoácido 67 usando el sistema de numeración de Kabat);
- (xiv) ácido glutámico (E) en la posición de aminoácido 82 usando el sistema de numeración de AHo

- (posición de aminoácido 71 usando el sistema de numeración de Kabat);
- (xv) treonina (T) en la posición de aminoácido 86 usando el sistema de numeración de AHo (posición de aminoácido 75 usando el sistema de numeración de Kabat);
- 5 (xvi) asparagina (N) en la posición de aminoácido 87 usando el sistema de numeración de AHo (posición de aminoácido 76 usando el sistema de numeración de Kabat); y
- (xvii) asparagina (N) en la posición de aminoácido 107 usando el sistema de numeración de AHo (posición de aminoácido 93 usando el sistema de numeración de Kabat).
8. Método según la reivindicación 1, en el que si la una o más posiciones de aminoácido de cadena ligera seleccionadas para la mutación son de una región variable de cadena ligera de la familia de V<sub>K</sub>1, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:
- 10 (i) ácido glutámico (E) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;
- (ii) valina (V) en la posición de aminoácido 3 usando el sistema de numeración de AHo o de Kabat;
- (iii) leucina (L) en la posición de aminoácido 4 usando el sistema de numeración de AHo o de Kabat;
- (iv) glutamina (Q) en la posición de aminoácido 24 usando el sistema de numeración de AHo o de Kabat;
- 15 (v) arginina (R) en la posición de aminoácido 47 usando el sistema de numeración de AHo (posición de aminoácido 39 usando el sistema de numeración de Kabat);
- (vi) arginina (R), ácido glutámico (E), treonina (T), metionina (M) o glutamina (Q) en la posición de aminoácido 50 usando el sistema de numeración de AHo (posición de aminoácido 42 usando el sistema de numeración de Kabat);
- 20 (vii) serina (S) en la posición de aminoácido 57 usando el sistema de numeración de AHo (posición de aminoácido 49 usando el sistema de numeración de Kabat);
- (viii) fenilalanina (F) en la posición de aminoácido 91 usando el sistema de numeración de AHo (posición de aminoácido 73 usando el sistema de numeración de Kabat); y
- 25 (ix) valina (V) en la posición de aminoácido 103 usando el sistema de numeración de AHo (posición de aminoácido 85 usando el sistema de numeración de Kabat).
9. Método según la reivindicación 1, en el que si la una o más posiciones de aminoácido de cadena ligera seleccionadas para la mutación son de una región variable de cadena ligera de la familia de V<sub>K</sub>3, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:
- 30 (i) treonina (T) en la posición de aminoácido 2 usando el sistema de numeración de AHo o de Kabat;
- (ii) treonina (T) en la posición de aminoácido 3 usando el sistema de numeración de AHo o de Kabat;
- (iii) isoleucina (I) en la posición de aminoácido 10 usando el sistema de numeración de AHo o de Kabat;
- (iv) tirosina (Y) en la posición de aminoácido 12 usando el sistema de numeración de AHo o de Kabat;
- (v) serina (S) en la posición de aminoácido 18 usando el sistema de numeración de AHo o de Kabat;
- (vi) alanina (A) en la posición de aminoácido 20 usando el sistema de numeración de AHo o de Kabat;
- 35 (vii) metionina (M) en la posición de aminoácido 56 usando el sistema de numeración de AHo (posición de aminoácido 48 usando el sistema de numeración de Kabat);
- (viii) treonina (T) en la posición de aminoácido 74 usando el sistema de numeración de AHo (posición de aminoácido 58 usando el sistema de numeración de Kabat);
- 40 (ix) asparagina (N) en la posición de aminoácido 94 usando el sistema de numeración de AHo (posición de aminoácido 76 usando el sistema de numeración de Kabat);
- (x) serina (S) en la posición de aminoácido 101 usando el sistema de numeración de AHo (posición de aminoácido 83 usando el sistema de numeración de Kabat); y
- (xi) alanina (A) en la posición de aminoácido 103 usando la numeración de AHo (posición de aminoácido 85 usando la numeración de Kabat).
- 45 10. Método según la reivindicación 1, en el que si la una o más posiciones de aminoácido de cadena ligera

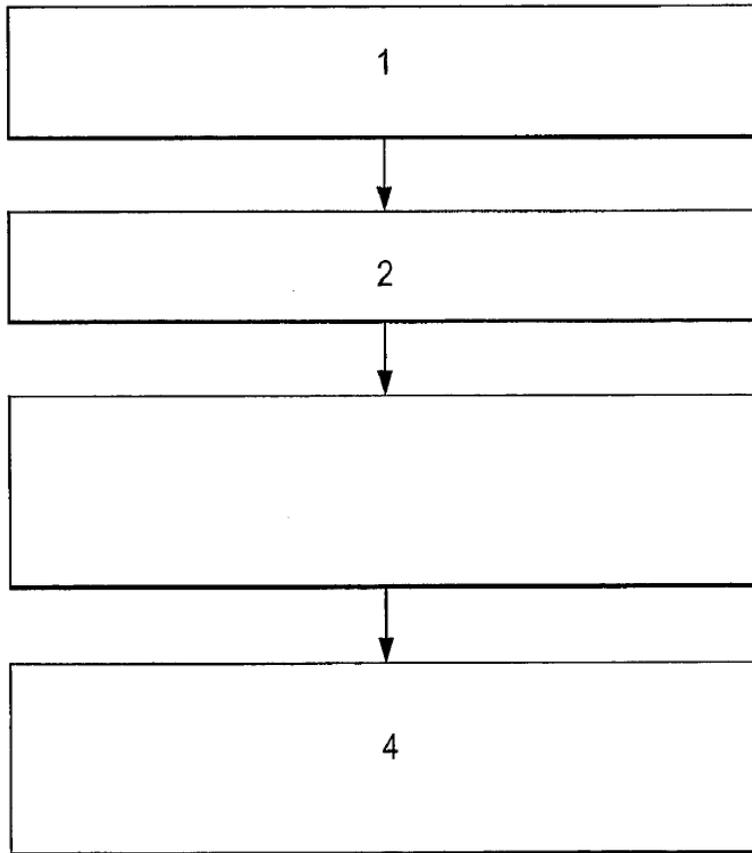
seleccionadas para la mutación son de una región variable de cadena ligera de la familia de V $\lambda$ 1, la mutación comprende una o más sustituciones seleccionadas del grupo que consiste en:

- (i) leucina (L) en la posición de aminoácido 1 usando el sistema de numeración de AHo o de Kabat;
- (ii) prolina (P) en la posición de aminoácido 2 usando el sistema de numeración de AHo o de Kabat;
- 5 (iii) valina (V) en la posición de aminoácido 4 usando el sistema de numeración de AHo o de Kabat;
- (iv) serina (S) en la posición de aminoácido 7 usando el sistema de numeración de AHo o de Kabat;
- (v) alanina (A) en la posición de aminoácido 11 usando el sistema de numeración de AHo o de Kabat;
- (vi) treonina (T) en la posición de aminoácido 14 usando el sistema de numeración de AHo o de Kabat;
- 10 (vii) histidina (H) en la posición de aminoácido 46 usando el sistema de numeración de AHo (posición de aminoácido 38 usando el sistema de numeración de Kabat);
- (viii) treonina (T), serina (S), asparagina (N), glutamina (Q) o prolina (P) en la posición de aminoácido 53 usando el sistema de numeración de AHo (posición de aminoácido 45 usando el sistema de numeración de Kabat);
- 15 (ix) arginina (R) en la posición de aminoácido 82 usando el sistema de numeración de AHo (posición de aminoácido 66 usando el sistema de numeración de Kabat);
- (x) treonina (T) en la posición de aminoácido 92 usando el sistema de numeración de AHo (posición de aminoácido 74 usando el sistema de numeración de Kabat); y
- (xi) valina (V) en la posición de aminoácido 103 usando la numeración de AHo (posición de aminoácido 85 usando la numeración de Kabat).
- 20 11. Método según una cualquiera de las reivindicaciones 5-10, en el que la mutación comprende además una o más sustituciones de cadena pesada seleccionadas del grupo que consiste en:
  - (i) serina (S) en la posición de aminoácido 12 usando AHo o Kabat;
  - (ii) serina (S) en la posición de aminoácido 103 usando la numeración de AHo (posición de aminoácido 85 usando la numeración de Kabat); y
  - 25 (iii) serina (S) o treonina (T) en la posición de aminoácido 144 usando la numeración de AHo (posición de aminoácido 103 usando la numeración de Kabat).
- 12. Método según una cualquiera de las reivindicaciones 5-10, en el que la mutación comprende además las siguientes sustituciones de cadena pesada:
  - (i) serina (S) en la posición de aminoácido 12 usando AHo o Kabat;
  - 30 (ii) serina (S) en la posición de aminoácido 103 usando la numeración de AHo (posición de aminoácido 85 usando la numeración de Kabat); y
  - (iii) serina (S) o treonina (T) en la posición de aminoácido 144 usando la numeración de AHo (posición de aminoácido 103 usando la numeración de Kabat).
- 13. Método según una cualquiera de las reivindicaciones 8-12, en el que el agente de unión inmunológica es un anticuerpo scFv, una inmunoglobulina de longitud completa, un fragmento Fab, un Dab o un Nanobody.
- 35 14. Método de modificación mediante ingeniería genética de un agente de unión inmunológica, comprendiendo el agente de unión inmunológica secuencias de CDR1, CDR2 y CDR3 de cadena pesada, comprendiendo el método insertar las secuencias de CDR1, CDR2 y CDR3 de cadena pesada en un armazón de región de marco de cadena pesada, comprendiendo el armazón de región de marco de cadena pesada una secuencia de aminoácidos tal como se muestra en la figura 9 (SEQ ID NO:1), la figura 10 (SEQ ID NO:2), la figura 11 (SEQ ID NO:3), SEQ ID NO: 7, SEQ ID NO:8 o SEQ ID NO:9, con la condición de que la secuencia de aminoácidos no sea la secuencia consenso de línea germinal.
- 40 15. Método según la reivindicación 14, en el que el armazón de región de marco de cadena pesada comprende una secuencia de aminoácidos tal como se muestra en la figura 9 (SEQ ID NO:1).
- 45 16. Método según la reivindicación 14, en el que el armazón de región de marco de cadena pesada comprende una secuencia de aminoácidos tal como se muestra en la figura 10 (SEQ ID NO:2).

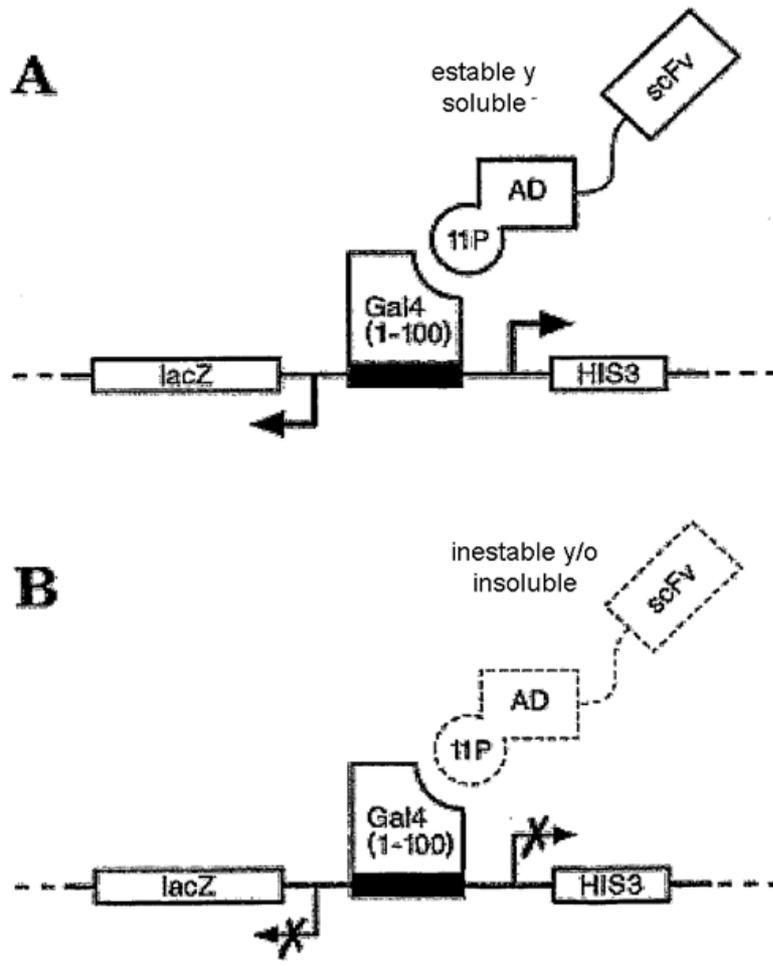
17. Método según la reivindicación 14, en el que el armazón de región de marco de cadena pesada comprende una secuencia de aminoácidos tal como se muestra en la figura 11 (SEQ ID NO:3).
- 5 18. Método de modificación mediante ingeniería genética de un agente de unión inmunológica, comprendiendo el agente de unión inmunológica secuencias de CDR1, CDR2 y CDR3 de cadena ligera, comprendiendo el método insertar las secuencias de CDR1, CDR2 y CDR3 de cadena ligera en un armazón de región de marco de cadena ligera, comprendiendo el armazón de región de marco de cadena ligera una secuencia de aminoácidos tal como se muestra en la figura 12 (SEQ ID NO:4), la figura 13 (SEQ ID NO:5), la figura 14 (SEQ ID NO:6), SEQ ID NO: 10, SEQ ID NO:11 o SEQ ID NO:12, con la condición de que la secuencia de aminoácidos no sea la secuencia consenso de línea germinal.
- 10 19. Método según la reivindicación 18, en el que el armazón de región de marco de cadena ligera comprende una secuencia de aminoácidos tal como se muestra en la figura 12 (SEQ ID NO:4).
20. Método según la reivindicación 18, en el que el armazón de región de marco de cadena ligera comprende una secuencia de aminoácidos tal como se muestra en la figura 13 (SEQ ID NO:5).
- 15 21. Método según la reivindicación 18, en el que el armazón de región de marco de cadena ligera comprende una secuencia de aminoácidos tal como se muestra en la figura 14 (SEQ ID NO:6).
22. Método según una cualquiera de las reivindicaciones 18-21, en el que el agente de unión inmunológica es un anticuerpo scFv.
- 20 23. Armazón de región de marco de cadena pesada aislado que comprende una secuencia de aminoácidos tal como se muestra en la figura 9 (SEQ ID NO:1), la figura 10 (SEQ ID NO:2) o la figura 11 (SEQ ID NO:3), con la condición de que la secuencia de aminoácidos no sea la secuencia consenso de línea germinal.
24. Armazón de región de marco de cadena ligera aislado que comprende una secuencia de aminoácidos tal como se muestra en la figura 12 (SEQ ID NO:4), la figura 13 (SEQ ID NO:5) o la figura 14 (SEQ ID NO:6), con la condición de que la secuencia de aminoácidos no sea la secuencia consenso de línea germinal.



*Fig. 1*

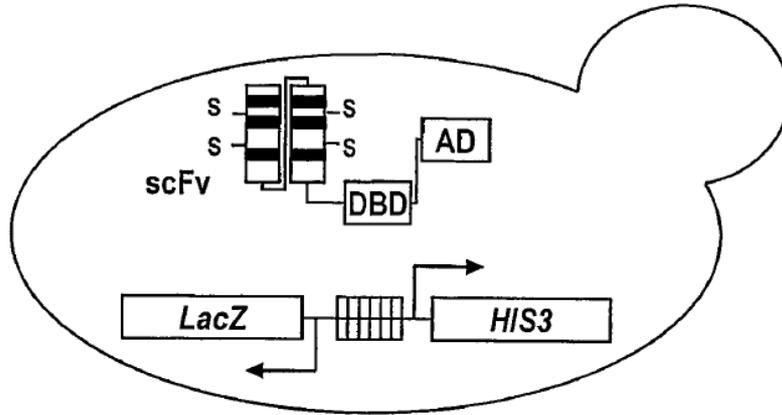


*Fig. 2*

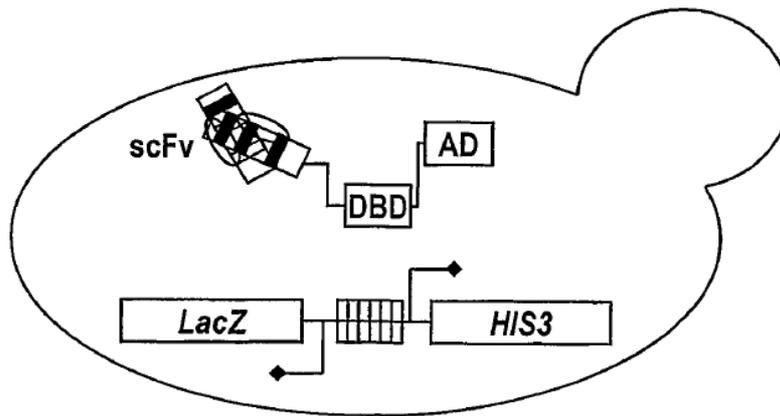


*Fig. 3*

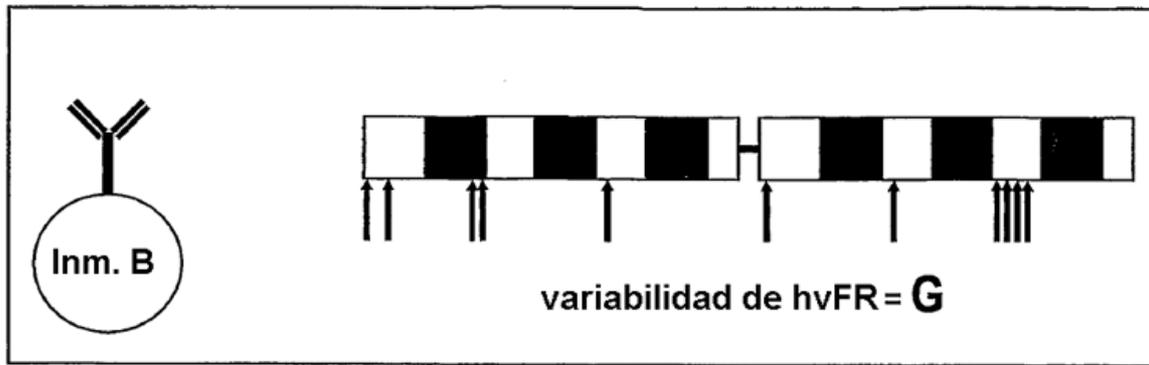
**A**



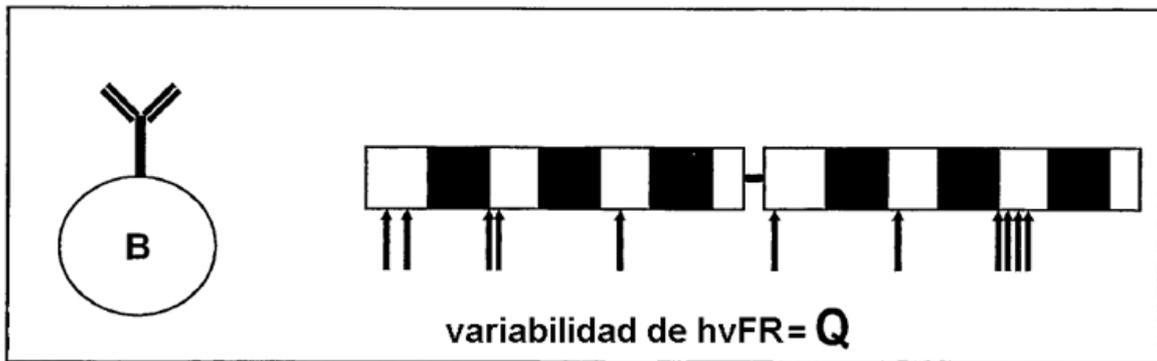
**B**



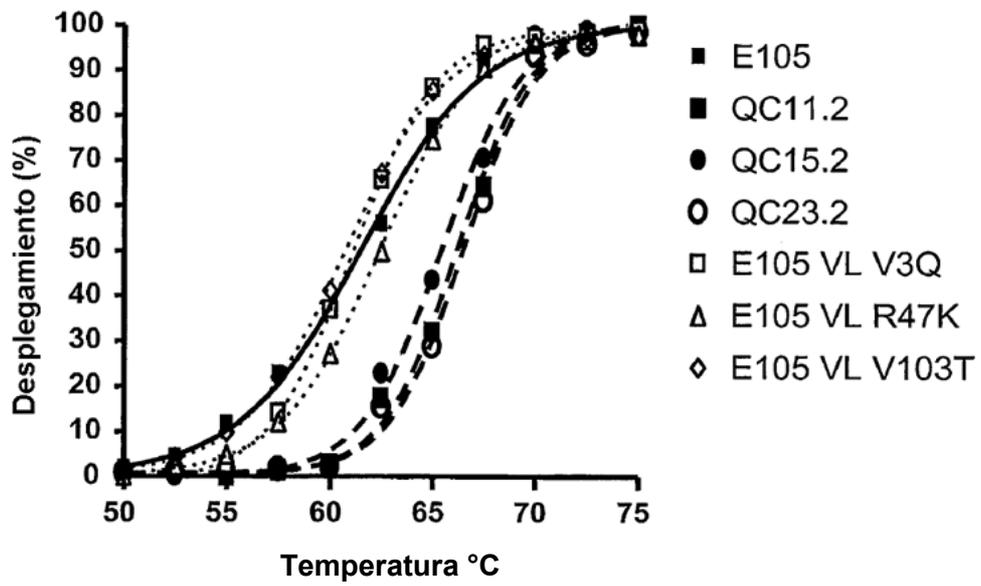
*Fig. 4*



*Fig. 5A*

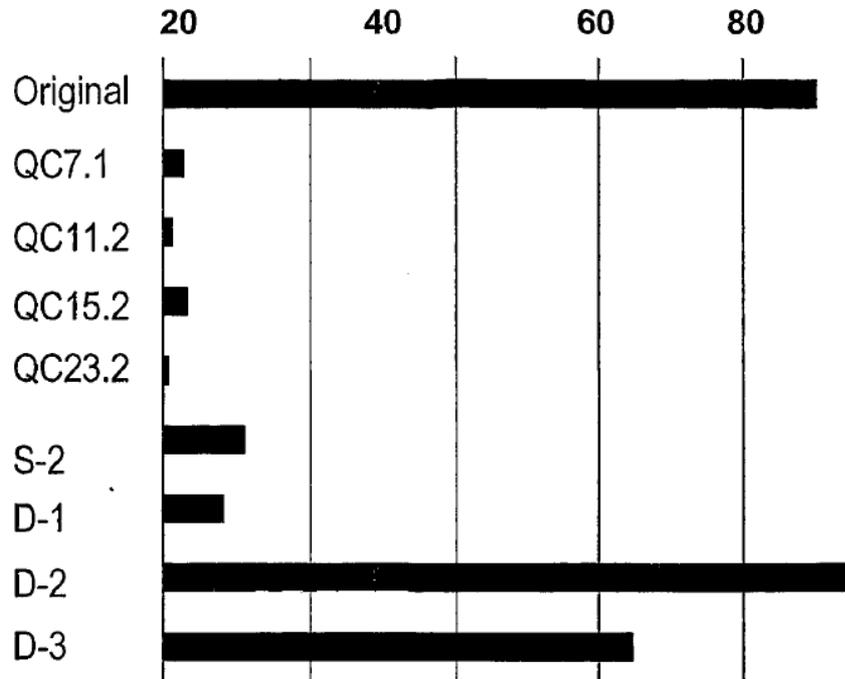


*Fig. 5B*

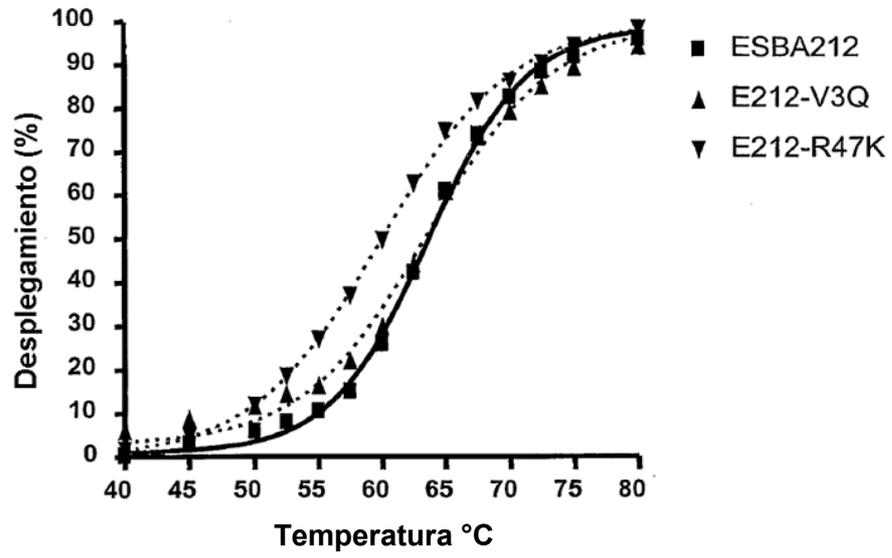


*Fig. 6*

unidades arbitrarias de despliegamiento [% del máx.]



*Fig. 7*



*Fig. 8*

Armazón de región de marco de cadena pesada de la familia de VH1

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	
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	
X	V	Q	L	V	X	S		G	A	E	X	X	X	P	G	S	S	X	K	X	S	C	K	A	S	
Q					Q						V	K	K					V		V						
E					E						L	M	E					L		I						
													Q													
26	*	27	28	29	30	31	32	33	34	35	35a	35b	*	*	*	36	37	38	39	40	41	42	43	44	45	
27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	
x		x	x	x												W	V	R	Q	A	P	G	Q	G	L	
46	47	48	49	50	51	52	52a	52b	52c	*	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68
53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
E	W	M	G																							
69	70	71	72	73	74	75	76	77	78	79	80	81	82	82a	82b	82b	83	84	85	86	87	88	89	90	91	
80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	
I	T	A	D	E	S	T	S	T	A	X	M	X	L	S	X	L	R	X	E	D	T	A	V	Y	Y	
										Y	E	E	S	S	S			S								
										F		D		G				F								
										S		Q		T				T								
										H				N				A								
										D								P								
92	93	94	95	96	97	98	99	100	100a	100b	100c	100d	100e	100f	100g	100h	100i	100j	*	*	*	*	*	*	*	
106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128				
C	A	R																								
*	*	*	*	*	*	*	*	*	*	101	102	103	104	105	106	107	108	109	110	111	112	113				
129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149						
										W	G	Q	G	T	L	V	T	V	T	V	S	S				

Fig. 9

Armazón de región de marco de cadena pesada de la familia de VH1B

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
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
X	V	Q	L	V	Q	S	G	X	E	X	X	X	X	P	G	A	S	V	X	X	S	C	K	A	S
Q								A		V	K	K							K	V					
E								T		L	V	R							N	L					
								D				R	E						R	I					
								P		Q	M								T	F					
								V		M															

26	*	27	28	29	30	31	32	33	34	35	35a	35b	*	*	*	36	37	38	39	40	41	42	43	44	45
27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52
x		x	x	x												W	V	X	Q	X	P	G	X	G	L

CDR H1

46	47	48	49	50	51	52	52a	52b	52c	*	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68
53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
E	W	X	G																R		A					
																			K		R					
																				T	V					

CDR H2

69	70	71	72	73	74	75	76	77	78	79	80	81	82	82a	82b	83	84	85	86	87	88	89	90	91	
80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105
M	T	E	D	T	S	T	N	T	A	Y	M	E	L	S	S	L	R	S	E	D	T	A	V	Y	Y

92	93	94	95	96	97	98	99	100	100a	100b	100c	100d	100e	100f	100g	100h	100i	*	*	*	*	*	*	*	*
106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128			
C	X	R																							
		A																							
		N																							
		S																							

CDR H3

129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	
*	*	*	*	*	*	*	*	*	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119

CDR H3

Fig. 10

Armazón de región de marco de cadena pesada de la familia de VH3

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
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
X	V	Q	L	V	X	X	X	G	P	G	L	V	K	P	S	E	T	L	R	L	S	C	A	A	S
E					E	S																			
Q					Q	T																			
					A																				

26	*	27	28	29	30	31	32	33	34	35	35a	35b	*	*	*	36	37	38	39	40	41	42	43	44	45
27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52
G	x	x	x	x												W	V	R	Q	A	P	G	K	G	L

46	47	48	49	50	51	52	52a	52b	52c	*	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	
53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	
E	W	V	V	S																					R	F	T

CDR H2																											
69	70	71	72	73	74	75	76	77	78	79	80	81	82	82a	82b	82b	83	84	85	86	87	88	89	90	91		
80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105		
I	S	R	D	N	S	K	N	T	X	Y	L	Q	M	N	S	L	R	A	E	D	T	A	X	Y	Y		
								L															V				
								V															L				
								A															I				
								F															M				
																							F				
																							R				
																							Q				

92	93	94	95	96	97	98	99	100	100a	100b	100c	100d	100e	100f	100g	100h	100i	*	*	*	*	*	*	*	*	*
106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128				
C	A	R																								

CDR H3																											
*	*	*	*	*	*	*	*	*	101	102	103	104	105	106	107	108	109	110	111	112	113						
129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149							
										W	G	Q	G	G	T	L	V	T	V	V	S	S					

Fig. 11





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	27a	27b	27c	27d	27e	
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
X	X	V	X	T	Q	X	.	P	S	X	S	G	X	P	G	Q	R	V	T	I	S	C	S									
L	S	V	V			S			A			T																				
Q	A	L	L			E		V				S																				
S	P	M				P						A																				
E	I																															
		Y																														

CDR L1

27f	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50									
34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66
									W	Y	Q	X	L	P	G	T	A	P	X	L	L	I	Y	D								
											H								T													
											Q								K													
																			S													
																			N													
																			Q													
																			P													

CDR L1

51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81		
67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99
N	N	Q	R	P	S	G	V	P	D	R	F	S	G	S	X	S	G	.	T	S	A	S	L	X	I	S	G	L	Q	S	E	
															R									T								
															Q									G								
															K									D								
																								A								

CDR L2

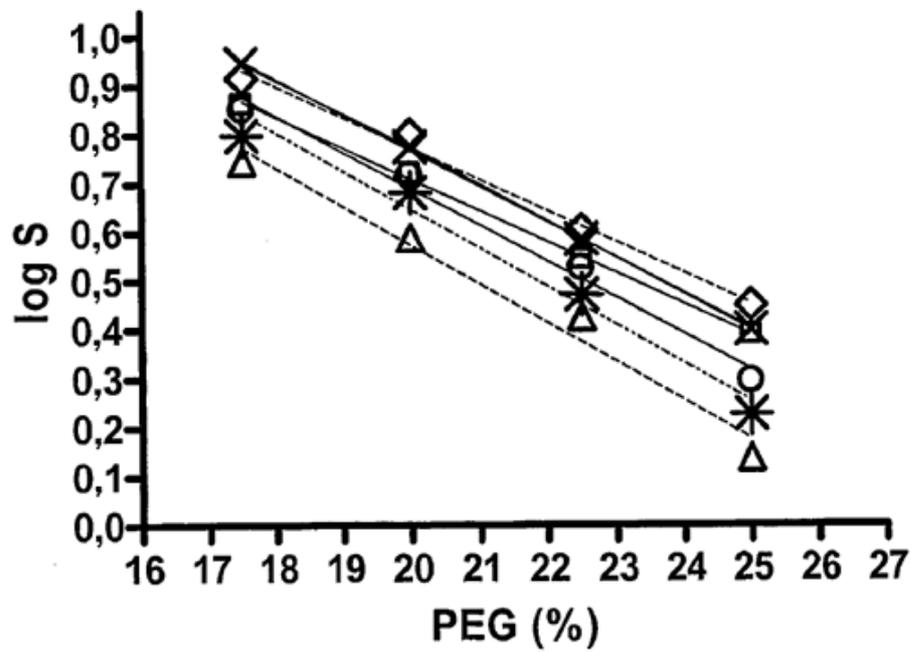
82	83	84	85	86	87	88	89	90	91	92	93	94	95	95a	95b	95c	95d	95e	95f													
100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132
D	E	A	X	Y	Y	C																										
			V																													
			D																													
			T																													
			H																													
			E																													

CDR L3

				96	97	98	99	100	101	102	103	104	105	106	107	108
133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149
				F	G	G	G	T	K	L	L	T	V	L	G	

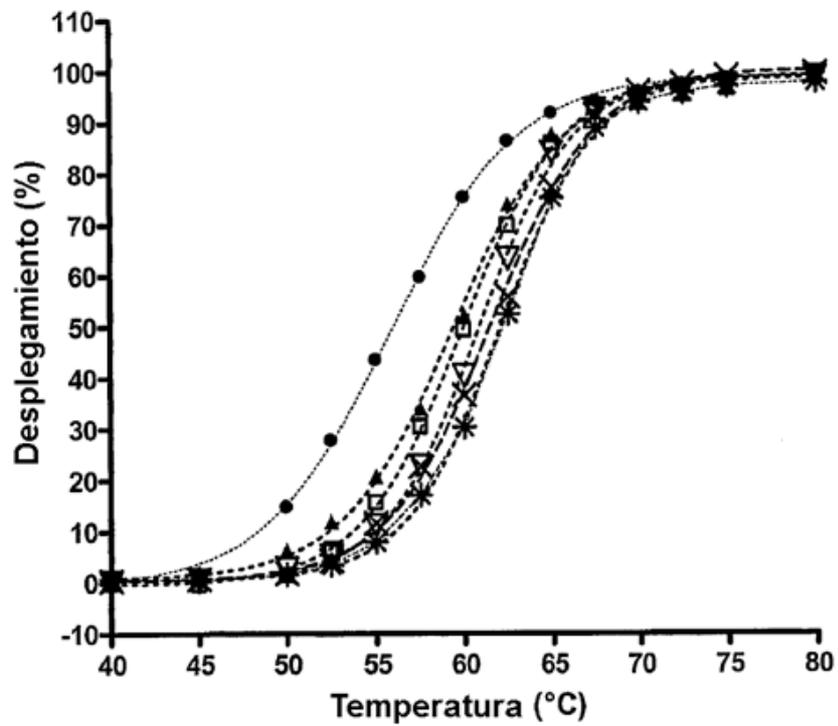
CDR L3

Fig. 14



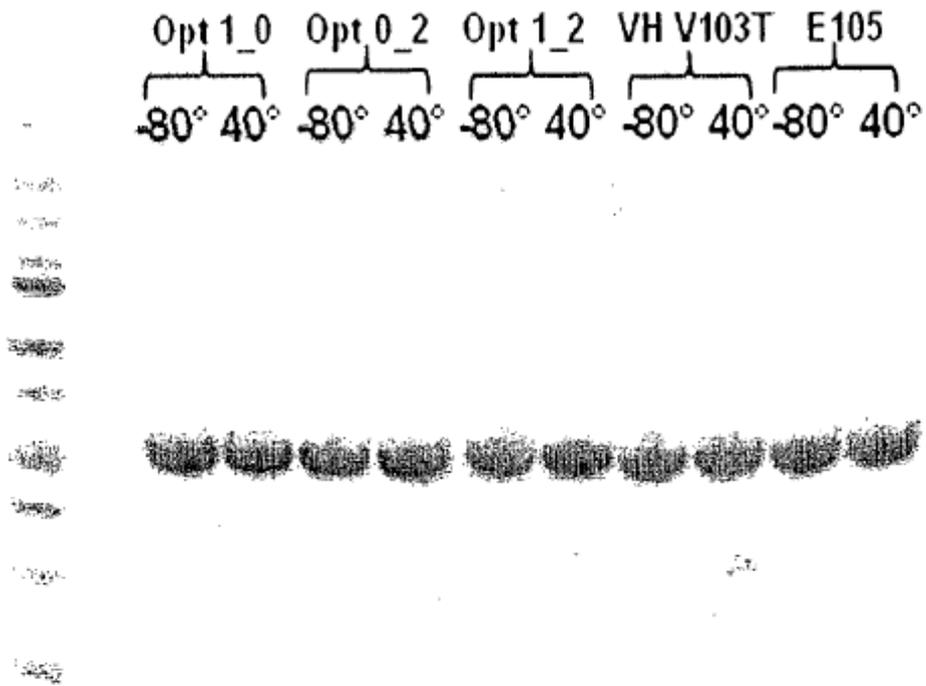
- E105
- × E105 Opt 1\_0
- E105 Opt 0\_2
- △ E105 Opt 1\_2
- \* E105 VH V103T
- ◇ E105 VL V147A

*Fig. 15*



- X ESBA105
- E105 Opt1.0
- ▲ E105 Opt1.2
- ▽ E105 Opt0.2
- \* E105 VH V103T
- E105 VL V147A

*Fig. 16*



*Fig. 17*