

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 384 006**

51 Int. Cl.:
C12N 15/10 (2006.01)
C07K 16/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **08719320 .7**
96 Fecha de presentación: **17.03.2008**
97 Número de publicación de la solicitud: **2134841**
97 Fecha de publicación de la solicitud: **23.12.2009**

54 Título: **Métodos para producir anticuerpos scFv activos y bibliotecas para los mismos**

30 Prioridad:
15.03.2007 US 894947 P

45 Fecha de publicación de la mención BOPI:
28.06.2012

45 Fecha de la publicación del folleto de la patente:
28.06.2012

73 Titular/es:
INSERM
101, rue de Tolbiac
75654 Paris Cédex 13, FR;
UNIVERSITE LOUIS PASTEUR DE STRASBOURG
y
CENTRE NATIONAL DE LA RECHERCHE
SCIENTIFIQUE (CNRS)

72 Inventor/es:
MARTINEAU, Pierre Emile Ulysse y
WEISS, Etienne

74 Agente/Representante:
Carvajal y Urquijo, Isabel

ES 2 384 006 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 para producir anticuerpos scFv activos y bibliotecas para los mismos

Campo

5 Esta descripción se refiere a anticuerpos de cadena sencilla recombinantes y a métodos de producción y uso de tales anticuerpos.

Antecedentes

10 Los enfoques de ingeniería genética han permitido la producción de anticuerpos recombinantes que tienen especificidades de unión específicas, estructuras de dominios específicas y otras propiedades deseables. Un tipo de anticuerpo modificado por ingeniería genética es el fragmento Fv de cadena sencilla (scFv). Los fragmentos Fv de cadena sencilla son polipéptidos modificados por ingeniería genética que contienen una región variable de cadena pesada (VH) unida a una región variable de cadena ligera (VL) mediante un ligador peptídico flexible. Cada dominio VH y VL contiene tres regiones determinantes de complementariedad (CDR). Las CDR son secuencias de aminoácidos cortas que varían enormemente entre moléculas de anticuerpo y, por tanto, son responsables de la generación de la gran diversidad de especificidad de unión de anticuerpos. La combinación de las CDR de la VH más las CDR de la VL determina la especificidad de unión de cualquier anticuerpo dado.

20 Los fragmentos Fv de cadena sencilla presentan la especificidad de unión y afinidad de unión monovalente de anticuerpos de tamaño completo y proporcionan el beneficio añadido de facilidad relativa de manipulación y expresión genéticas (porque los scFv se codifican por y se expresan a partir de una única secuencia codificante, en vez de secuencias codificantes separadas, como los anticuerpos de tamaño completo). Los fragmentos Fv de cadena sencilla y otros anticuerpos recombinantes se usan en una amplia variedad de aplicaciones, por ejemplo, en pruebas de diagnóstico médico, en investigación básica; y como tratamientos de anticuerpos terapéuticos para diversas enfermedades.

25 Los intracuerpos son moléculas de anticuerpos modificados por ingeniería genética que se expresan de manera ectópica dentro de las células. Pueden usarse intracuerpos para visualizar o modular la función de un antígeno diana dentro de células vivas. Por ejemplo, el uso de intracuerpos puede inducir una desactivación fenotípica o bien mediante la inhibición directa de la función del antígeno seleccionado como diana o bien mediante el desvío del antígeno seleccionado como diana de su ubicación intracelular normal (por ejemplo, un intracuerpo puede redirigir su antígeno diana a la maquinaria de degradación). Los intracuerpos también pueden potenciar o cambiar la función de sus antígenos diana. Para dianas proteicas, los intracuerpos pueden dirigirse a una modificación postraduccional específica o a una conformación de antígeno específica. Además, una desactivación fenotípica inducida por intracuerpos puede estar confinada a un compartimento celular específico mediante el direccionamiento de un intracuerpo al compartimento subcelular específico usando una señal de encauzamiento (por ejemplo, una señal de localización nuclear, una señal de localización mitocondrial o una señal de retención en el retículo endoplásmico). Los intracuerpos también pueden modular la función de la diana modificando la estructura oligomérica de la diana.

35 Debido a que la desactivación fenotípica por intracuerpos se basa sólo en la capacidad de unión de la molécula de anticuerpo a su diana, no es necesario expresar dentro de la célula una molécula de anticuerpo completa sino sólo su sitio de unión, que está completamente ubicado dentro de la región variable (Fv). Dadas sus ventajas de tamaño pequeño y especificidad de antígeno abarcada dentro de una única cadena polipeptídica, los scFv son el tipo más común de fragmento de anticuerpo recombinante usado para la expresión de anticuerpos intracelulares.

40 Una grave limitación en el uso de intracuerpos es que la mayoría de los scFv no pueden plegarse en las condiciones reductoras del núcleo y el citosol celular. En tales condiciones, los dos puentes disulfuro conservados de scFv se reducen, desestabilizándose e inactivándose de ese modo la actividad de unión de muchos scFv. *In vitro*, la mayoría de los scFv no pueden renaturalizarse en condiciones reductoras. Análisis estadísticos de secuencias de scFv han mostrado que menos del 1% de los scFv son lo suficientemente estables como para expresarse y ser activos en ausencia de formación de enlaces disulfuro. Además, incluso si una proteína scFv es de hecho lo suficientemente estable en su forma reducida como para expresarse y ser activa *in vivo*, otros parámetros tales como susceptibilidad a proteasas o cinética de plegamiento pueden influir también en el destino *in vivo* final del intracuerpo y por tanto son críticos para la actividad y expresión de intracuerpos final.

50 Para obtener un intracuerpo activo, los enfoques actuales implican a menudo dos etapas sucesivas. En primer lugar, se identifica un panel de anticuerpos scFv o Fab que se unen específicamente a un antígeno de interés (por ejemplo, examinando una biblioteca de presentación en fago). En segundo lugar, los anticuerpos que se unen específicamente se someten a prueba para determinar su capacidad para unirse y/o inhibir el antígeno diana *in vivo*. Debido a que menos del 1% de los scFv son potencialmente útiles como intracuerpos (debido a que no se expresan y/o no pueden plegarse apropiadamente en las condiciones reductoras que existen dentro de una célula), la

identificación de scFv individual que puede usarse como intracuerpo requiere el aislamiento de más de 100 clones de scFv, un número que es improbable que se obtenga en la mayoría de los casos.

5 La solicitud WO 00/54057 da a conocer un método para determinar la capacidad de una inmunoglobulina para unirse a una antígeno diana en un entorno intracelular. El método dado a conocer consiste en proporcionar i) una inmunoglobulina intracelular asociada con una primera molécula, y ii) un antígeno diana intracelular asociado con una segunda molécula, luego evaluar la interacción intracelular entre la inmunoglobulina frente a la diana en el entorno intracelular, sabiendo que cuando las moléculas primera y segunda se llevan a interacción estable mediante la unión de la inmunoglobulina a la diana en el entorno intracelular, se genera una señal. Para evaluar la posibilidad de realizar este método en células de mamífero, se ha llevado a cabo una prueba con una primera proteína de fusión que comprende beta-galactosidasa, y una segunda proteína de fusión que comprende el scFvR4 anti-beta-galactosidasa.

La solicitud WO 01/48017 da a conocer un método para aislar un scFv con un entramado que es estable y soluble en un entorno reductor. El método dado a conocer se realiza usando un "ensayo de dos híbridos" que comprende las siguientes etapas:

15 a) se genera una biblioteca de scFv con CDR constantes y entramados variados mediante la mutación de al menos una región codificante de entramado de la secuencia de ADN de un scFv frente a un antígeno conocido y mediante la introducción de tales mutaciones en vectores de expresión adecuados,

b) se transforman células huésped que pueden expresar un antígeno conocido específico y que sólo sobreviven en presencia de interacción antígeno-scFv con dicha biblioteca de scFv,

20 c) se cultivan las células huésped así transformadas en condiciones adecuadas para expresar el antígeno y el scFv y que permiten la supervivencia celular sólo en presencia de interacción antígeno-scFv,

d) se aísla el scFv expresado en las células supervivientes y que tiene un entramado definido que es estable y soluble en el entorno reductor.

25 Der Maur Adrian *et al.* (Methods, 34: 215-224, 2004) y Der Maur Adrian *et al.* (Journal of Biological Chemistry, 277(47): 45075-45085, 2002) dan a conocer métodos para aislar un scFv con un entramado que es estable y soluble en un entorno reductor.

Der Maur Adrian *et al.* (Methods, 34: 215-224, 2004) enseña en la tabla 1 que de un total de 10^7 clones examinados, sólo un 0,1% permiten el crecimiento celular en condiciones selectivas, y sólo un 60% de estos clones que permiten el crecimiento celular en condiciones selectivas muestran expresión de lacZ dependiente de scFv.

30 Laden *et al.* (Research in Microbiology, 153: 469-474, 2002) dieron a conocer un mutante de un fragmento de anticuerpo de cadena sencilla anti-beta-galactosidasa humano, denominado scFv13R4, y muestran que scFv13R4 puede plegarse apropiadamente en el citoplasma de *Escherichia coli* y es activo *in vivo* en levaduras puesto que puede interaccionar con beta-galactosidasa en un ensayo de dos híbridos.

35 Jung y plückthun (protein Engineering, 10(8): 959-966, 1997) da a conocer un método para mejorar *in vivo* el plegado y la estabilidad de un fragmento de anticuerpo Fv de cadena sencilla usando injerto de CDR. El método se realiza con el entramado del anticuerpo 4D5 humanizado.

40 La solicitud WO 99/45959 da a conocer un motivo de entramado de anticuerpo humanizado en el que dicho entramado se ha seleccionado de una biblioteca humana basándose en la comparación con un anticuerpo murino, y la cadena pesada está codificada por el gen de VH de K5B8, y la cadena ligera está codificada por el gen de VL TR1.6.

Sin embargo, ninguno de los documentos de la técnica anterior mencionados anteriormente da a conocer una biblioteca de anticuerpos que comprenda al menos 10^6 clones de anticuerpos scFv únicos en la que cada clon codifique para una secuencia de entramado idéntica a un entramado codificado por el clon de anticuerpo scFv13R4 (SEQ ID NO: 32).

45 En vista de las dificultades anteriores en la producción e identificación de anticuerpos que pueden usarse como intracuerpos para su uso en aplicaciones médicas y de investigación, lo que se necesita son métodos más eficaces de producción y selección de anticuerpos que puedan usarse como intracuerpos.

Sumario

Los inventores han descrito en el presente documento una biblioteca de anticuerpos que incluye al menos aproximadamente 10^6 clones de scFv únicos, en la que al menos aproximadamente el 20% de los clones de scFv codifican para un anticuerpo que puede unirse específicamente de manera detectable a un antígeno diana dentro de una célula cuando el anticuerpo codificado por el clon de scFv se expresa dentro de la célula.

5 Los inventores también han descrito en el presente documento una biblioteca de anticuerpos, incluyendo la biblioteca de anticuerpos al menos aproximadamente 10^6 clones de anticuerpos scFv únicos, en la que al menos aproximadamente el 20% de los clones de anticuerpos scFv pueden expresarse dentro de una célula de *E. coli* para producir anticuerpo soluble a un nivel de al menos aproximadamente 5 mg por litro de células de *E. coli*, en la que las células de *E. coli* se han hecho crecer hasta una DO_{600nm} de aproximadamente 5.

10 En un aspecto, la invención descrita en el presente documento es una biblioteca de anticuerpos que incluye al menos aproximadamente 10^6 clones de anticuerpos scFv únicos, en la que cada clon de anticuerpo scFv único codifica para un anticuerpo scFv único que comprende al menos una de una secuencia VH CDR3 única y una secuencia VL CDR3 única, y en la que los clones de anticuerpos scFv únicos codifican para una secuencia de entramado sustancialmente idéntica a una secuencia de entramado codificada por scFv13R4 (SEQ ID NO: 32).

15 En la biblioteca de anticuerpos anterior de la invención, los clones de anticuerpos scFv únicos pueden codificar para anticuerpos scFv que incluyen una secuencia VH CDR3 única.

En la biblioteca de anticuerpos anterior de la invención, los clones de anticuerpos scFv únicos pueden codificar para anticuerpos scFv que incluyen una secuencia VH CDR3 única.

20 En la biblioteca de anticuerpos anterior de la invención, los clones de anticuerpos scFv únicos pueden codificar para anticuerpos scFv que incluyen una secuencia VL CDR3 única.

En la biblioteca de anticuerpos anterior de la invención los clones de anticuerpos scFv únicos pueden codificar para anticuerpos scFv que incluyen una secuencia VH CDR3 única y una secuencia VL CDR3 única.

25 Los inventores han descrito en el presente documento un anticuerpo scFv que puede expresarse como una proteína sustancialmente soluble en condiciones reductoras, en las que el anticuerpo scFv se aísla de la biblioteca descrita anteriormente en el tercer aspecto. El anticuerpo scFv puede unirse específicamente a un antígeno diana en condiciones reductoras.

Los inventores han descrito en el presente documento un método de producción de un anticuerpo scFv, que incluye la expresión del anticuerpo scFv descrito anteriormente en el cuarto aspecto dentro de una célula, produciendo de ese modo el anticuerpo scFv. El método puede incluir purificar adicionalmente el anticuerpo scFv de la célula.

30 En un segundo aspecto de la invención, se describe en el presente documento un método para preparar una biblioteca de anticuerpos scFv enriquecida en clones de anticuerpos scFv que pueden expresarse dentro de una célula, que incluye: a) proporcionar una primera colección de clones de anticuerpos scFv, en el que la primera colección comprende clones que comprenden una secuencia única dentro de un bucle CDR3 de VH, en el que la primera colección se ha enriquecido en clones de anticuerpos scFv que pueden expresarse de manera detectable cuando se introducen en una célula; b) proporcionar una segunda colección de clones de anticuerpos scFv, en el que la segunda colección comprende clones que comprenden una secuencia única dentro de un bucle CDR3 de VL, en el que la segunda colección se ha enriquecido en clones de anticuerpos scFv que pueden expresarse de manera detectable cuando se introducen en una célula; c) unir dominios VH de clones de anticuerpos scFv de la primera colección con dominio VL de clones de anticuerpos scFv de la segunda colección para obtener una tercera colección de clones de anticuerpos scFv, en el que la tercera colección contiene clones de anticuerpos scFv que comprenden una secuencia única dentro del bucle CDR3 de VH y una secuencia única dentro del bucle CDR3 de VL, preparando de ese modo la biblioteca de anticuerpos scFv enriquecida en clones de anticuerpos scFv que pueden expresarse dentro de una célula.

45 En el método anterior para preparar una biblioteca de anticuerpos scFv, la primera colección incluye clones de anticuerpos scFv que contienen una secuencia VL idéntica a una secuencia VL de scfv13R4 (SEQ ID NO:33) y la segunda colección puede incluir clones de anticuerpos scFv que contienen una secuencia VH idéntica a una secuencia VH de scFv13R4 (SEQ ID NO:33).

50 En el método anterior para preparar una biblioteca de anticuerpos scFv, la primera colección puede incluir clones de anticuerpos scFv que incluyen secuencias CDR1 y CDR2 idénticas en el dominio VH y la segunda colección puede incluir clones de anticuerpos scFv que incluyen secuencias CDR1 y CDR2 idénticas en el dominio VL.

En un tercer aspecto de la invención, se describe en el presente documento una biblioteca de anticuerpos producida mediante el método descrito anteriormente en el sexto aspecto.

Los inventores también describieron en el presente documento un anticuerpo seleccionado de una biblioteca de anticuerpos producida mediante el método descrito anteriormente en el segundo aspecto.

En un noveno aspecto, la invención caracteriza un método para construir una biblioteca de anticuerpos que incluye:
 5 a) seleccionar un entramado de anticuerpo scFv; b) introducir diversidad de secuencia en una región CDR3 VH del entramado de anticuerpo scFv para generar una primera biblioteca que incluye clones de anticuerpos scFv que incluyen una región CDR3 VH única; c) introducir diversidad de secuencia en una región CDR3 VL del entramado de anticuerpo scFv para generar una segunda biblioteca que incluye clones de anticuerpos scFv que incluyen una región CDR3 VL única; d) eliminar, de la primera biblioteca, clones que no expresan de manera detectable el anticuerpo scFv; e) eliminar, de la segunda biblioteca, clones que no expresan de manera detectable el anticuerpo scFv; y f) recombinar las bibliotecas primera y segunda para generar una biblioteca final que comprende clones de anticuerpos scFv que comprenden una región CDR3 VH única y una región CDR3 VL única, construyendo de ese modo la biblioteca de anticuerpos.

En el método anterior, el scFv es scFv13R4.

Es un objeto de la presente invención proporcionar una biblioteca de anticuerpos novedosa para el aislamiento de scFv expresados en el citoplasma que pueden usarse como intracuerpos.

Es otro objeto de la presente invención proporcionar una biblioteca de anticuerpos novedosa basados en un entramado individual y optimizados para expresión intracelular.

Un objeto adicional de la presente invención es proporcionar métodos novedosos de construcción y validación de una biblioteca de anticuerpos novedosa para el aislamiento de scFv expresados en el citoplasma que pueden usarse como intracuerpos.

Otro objeto de la presente invención es proporcionar métodos novedosos de construcción y validación de una biblioteca de anticuerpos novedosa basados en un entramado individual y optimizados para expresión intracelular.

Todavía otro objeto de la presente invención es proporcionar métodos novedosos de uso de una biblioteca de anticuerpos con el fin de producir scFv altamente expresados que pueden usarse como intracuerpos.

Aún otro objeto de la presente invención es proporcionar métodos novedosos de uso de una biblioteca de anticuerpos con el fin de producir scFv basados en un entramado individual y optimizados para expresión intracelular.

Estos y otros objetos, características y ventajas de la presente invención resultarán evidentes tras la revisión de la siguiente descripción detallada de las realizaciones dadas a conocer.

Breve descripción de los dibujos

Figura 1. (a) Resumen esquemático de las etapas seguidas durante la construcción de la biblioteca. Las etapas críticas son: introducción de bucles CDR3 adaptados en un entramado de scFv humano único; eliminación de clones no expresados mediante fusión con la enzima CAT y selección en placas CAM; recombinación de las 13 bibliotecas de VH y 5 de VL, y presentación en fago. (b) Resumen de los bucles CDR3 recogidos en la base de datos. (c) Distribución de los aminoácidos en cada posición de las CDR3 VH de 5 aminoácidos de longitud de 55 anticuerpos humanos reordenados.

Figura 2. Representación esquemática de la distribución de longitud de CDR3. Distribución de las longitudes de CDR3 en la base de datos y en 118 clones secuenciados de la biblioteca.

Figura 3. Análisis de inmunotransferencia de tipo Western de veinte clones de la biblioteca que se clonaron en un vector de expresión citoplasmática y se expresaron en *E. coli* bajo el control del promotor de T7. Se clonaron veinte clones de la biblioteca en un vector de expresión citoplasmática y se expresaron en *E. coli* bajo el control del promotor de T7. Se prepararon extractos solubles, se separaron mediante SDS-PAGE y se analizaron mediante tinción con Coomassie (a) o inmunotransferencia de tipo Western (b) usando 9E10 y un anticuerpo anti-IgG de ratón conjugado con fosfatasa alcalina (sustrato BCIP/NBT). Cada carril correspondía a 2×10^7 (2×10 a la 7ª) células. La flecha a la izquierda indica la posición del scFv.

Figura 4. Micrografías fluorescentes seleccionadas de la expresión de scFv escogidos al azar en células HeLa. Se transfectaron las células con constructos de scFv-EGFP tal como se indica. 13R4 y 1F4 representan los controles positivo y negativo, respectivamente. A las 24 h tras la transfección, se fijaron las células y se visualizaron bajo un microscopio fluorescente con el conjunto de filtro de isotiocianato de fluoresceína. Las micrografías representan campos típicos que contienen un número de células similar en cada caso. Aumentos: x 400.

Figura 5. Selección de elementos de unión frente a cinco proteínas purificadas. (a) Se sometieron a prueba $2,5 \times 10^{10}$ fagos de cada ronda de selección mediante ELISA frente a su antígeno respectivo y se revelaron usando un anticuerpo monoclonal conjugado con HRP anti-M13 (Pharmacia). R0 es la biblioteca no seleccionada y Rn la reserva obtenida tras la ronda n^a de selección. Se sometió a prueba la especificidad en BSA para las 3^a rondas de selección (A450 nm -0,1 - 0,3). (b) Se sometieron a prueba fagos monoclonales de cada ronda de selección frente a proteína GST:Syk mediante ELISA como en (a). Porcentaje de scFv activos solubles (absorbancia > 0,1) seleccionados tras tres rondas frente a la proteína indicada y expresados o bien en el periplasma (gris) o bien en el citoplasma (negro).

Figura 6. Fotomicrografías de la expresión de scFV anti-histonas fusionados a EGFP en células HeLa. Se transfectaron las células y se trataron tal como se indica en la leyenda de la figura 4. Las fotografías representan células típicas transfectadas con scFv13R4 y tres clones anti-histonas representativos (2, 5 y 10). D, tinción con DAPI (azul) combinada con la señal de GFP.

Figura 7. Representación gráfica de la expresión de la actividad de scFv seleccionados en condiciones reductoras. Se prepararon extractos de células que expresan los scFv en el citoplasma como en la figura 3 en presencia o ausencia de DTT 10 mM. Se realizó ELISA en presencia o ausencia de DTT, por consiguiente. Eje x: cantidad de extracto por pocillo. Eje y: señal de ELISA a 450 nm.

Figura 8. Fotomicrografías de la doble tinción de células HeLa usando un anticuerpo y mediante inmunofluorescencia. Se transfectaron células HeLa con el clon 5 anti-histonas (figura 6) fusionado al dsRed-monómero de GFP. Se revelaron los microtúbulos en las células transfectadas mediante IF usando el scFv anti-tubulina 2F12C (tabla 3). Se observaron las células a la longitud de onda apropiada para visualizar: (A) scFv 2F12C (microtúbulos solos); (B) intracuerpo de clon 5 (histonas solas); (C) 2F12C y con 5 (tanto microtúbulos como histonas); (D) como C tinción del núcleo con DAPI.

Descripción detallada de la invención

Se proporcionan en el presente documento métodos para construir una biblioteca de scFv enriquecida en scFv que pueden expresarse dentro de una célula procariota o eucariota. Los scFv mantienen su estructura en las condiciones reductoras presentes dentro de una célula, conservan su capacidad para unirse específicamente a un antígeno diana dentro de una célula, y por tanto pueden usarse como intracuerpos en diversos enfoques terapéuticos y de investigación. Debido a que los métodos descritos en el presente documento pueden usarse para producir bibliotecas de scFv que son estables en condiciones reductoras, estas bibliotecas también son útiles para producir scFv que pueden expresarse por y purificarse a partir de células procariotas o eucariotas (por ejemplo, lisando las células y purificando el scFv deseado mediante técnicas bien conocidas). Tales scFv purificados son útiles como anticuerpos para su uso en investigación, pruebas de diagnóstico y para terapias contra enfermedades.

Una estrategia mejorada para estabilizar scFv que van a usarse *in vivo* es construir una biblioteca de scFv adaptada para expresión intracelular. Idealmente, una biblioteca de este tipo debe contener sólo scFv que puedan plegarse en condiciones reductoras tales como las encontradas en el citoplasma de una célula. Otra estrategia es construir una biblioteca de scFv basándose en un entramado de anticuerpo optimizado individual para la selección de intracuerpos.

Se describen en el presente documento métodos para construir una biblioteca de anticuerpos novedosa que se basa en un entramado de anticuerpo optimizado individual y que está adaptado para expresión intracelular. A través de evolución molecular, se obtuvo un scFv humano denominado scFv13R4, que se expresa a altos niveles en el citoplasma de *E. coli*. Además, este scFv se expresa altamente, es soluble y presenta actividad de unión específica frente a un antígeno diana en levadura y células de mamífero. Este scFv es muy estable *in vitro* y puede renaturalizarse en presencia de un agente reductor. Además, el análisis de su cinética de plegamiento mostró que se pliega más rápido que el scFv original y se agrega más lentamente *in vitro*.

La biblioteca de scFv humana basada en el entramado del scFv13R4 optimizado contiene más de un billón de clones, lo que es más que las bibliotecas previas de 10^7 a 10^8 . La diversidad de la presente biblioteca de scFv humana es mucho mayor que la de las bibliotecas previas porque se diseñó la presente biblioteca para que codificase para bucles CDR3 de VH y VL de diversas longitudes diferentes. Además, se usó una mezcla sesgada de secuencias oligonucleotídicas degeneradas para codificar para bucles CDR3 que imitan CDR3 humana. Usando bucles CDR3 optimizados y etapas de filtración para eliminar los clones no expresados, se expulsaron de la biblioteca scFv no expresados y se conservaron los scFv expresados citoplasmáticamente sin comprometer la diversidad de los clones, tal como se confirmó a través de pruebas con varias proteínas usadas como antígenos. Al contrario que bibliotecas de scFv descritas anteriormente, la mayoría de los scFv en la biblioteca se expresan bien en *E. coli* y en el citoplasma de mamíferos.

Este nuevo enfoque para construir bibliotecas de scFv permite la selección fácil y a gran escala de intracuerpos

funcionales. Por ejemplo, se han aislado de la biblioteca varios elementos de unión fuerte frente a diferentes proteínas, incluyendo las proteína cinasas Syk y Aurora-A, el dímero de tubulina $\alpha\beta$, las proteínas E6 del virus del papiloma, las histonas núcleo, gankirina y MAPK11-14. Algunos de los scFv seleccionados se expresan a un nivel excepcionalmente alto en el citoplasma bacteriano, permitiendo la purificación de 1 mg o más de scFv activo a partir de sólo 20 ml de cultivo. Además, tras tres rondas de selección frente a histonas núcleo como antígeno diana, más de la mitad de los scFv seleccionados eran activos cuando se expresaban *in vivo* en células humanas y se localizaban esencialmente en el núcleo. Este nuevo tipo de biblioteca, los métodos de creación y uso de tales bibliotecas y los anticuerpos aislados a partir de tales bibliotecas son útiles no sólo para la selección sencilla y a gran escala de intracuerpos funcionales sino también para la expresión y purificación de scFv altamente expresados que pueden usarse en numerosas aplicaciones biotecnológicas, de diagnóstico y terapéuticas.

Intracuerpos

Los intracuerpos son moléculas de anticuerpo modificadas por ingeniería genética que se expresan de manera ectópica dentro de las células. Pueden usarse intracuerpos para visualizar o inhibir un antígeno seleccionado como diana en células vivas, y por tanto encuentran uso en diversas aplicaciones de investigación y médicas (por ejemplo, de diagnóstico y terapéuticas). Sin embargo, la tecnología de intracuerpos se ha visto limitada por el hecho de que menos del 1% de los scFv en una biblioteca de expresión de anticuerpos típica son los suficientemente estables como para expresarse y/o ser activos *in vivo*, porque el entorno intracelular reduce los dos puentes disulfuro conservados que la gran mayoría de los scFv requiere para su estabilidad. Se describen en el presente documento métodos de producción de bibliotecas de scFv que están enormemente enriquecidas en scFv que son estables en condiciones reductoras y por tanto son adecuados para su uso como intracuerpos. Los intracuerpos pueden usarse para diversos enfoques de investigación, de diagnóstico y terapéuticos que emplean intracuerpos.

En la mayoría de los casos, la obtención de intracuerpos eficaces requiere actualmente dos etapas sucesivas. En primer lugar, debe aislarse un panel de anticuerpos frente al antígeno diana. Debido a la disponibilidad de bibliotecas de anticuerpos vírgenes de muy alta calidad presentados en fago, esta etapa se logra ahora fácilmente mediante presentación en fago y puede automatizarse con el fin de aislar elementos de unión frente a varias proteínas en paralelo. En una segunda etapa, estos fragmentos de anticuerpo (scFv o Fab) deben someterse a prueba *in vivo* para determinar su capacidad para inhibir su diana. Sin embargo, la mayoría de los scFv no se pliegan apropiadamente en las condiciones reductoras encontradas en el citosol y el núcleo de la célula donde se ubican la mayoría de las dianas interesantes. Esto puede dar como resultado scFv agregados e inactivos, que no pueden interaccionar con su diana. Puesto que menos del 1% de los scFv son eficaces como intracuerpos, conseguir un elemento de unión individual frente a una proteína requiere el aislamiento de 100 scFv diferentes, un número que es improbable que se obtenga en la mayoría de los casos. Esto hace que el procedimiento de identificación de intracuerpos a partir de bibliotecas de scFv regulares sea un procedimiento difícil incluso cuando el examen se realiza *in vivo* usando sistemas de dos híbridos o equivalentes. Además, esta baja proporción de scFv activos en las bibliotecas actuales da como resultado una disminución de 100 veces en el repertorio potencial examinado, lo que hace que el aislamiento de intracuerpos frente a diferentes epítomos de la misma proteína sea improbable.

Bibliotecas

Se describen en el presente documento bibliotecas de presentación en fago de scFv optimizados para expresión intracelular y métodos novedosos de construcción y uso de la biblioteca. La biblioteca se construye sobre un entramado de anticuerpo individual de un scFv original que se seleccionó debido a su actividad mejorada dentro del citoplasma. El scFv original es muy estable, tiene un plegamiento y una cinética de agregación favorables y se expresa a muy altos niveles en todos los tipos celulares sometidos a ensayo. Tener un entramado individual para la construcción de una biblioteca debe permitir niveles de expresión más comparables entre clones puesto que la mayoría de sus secuencias se conservan.

Debido a que las secuencias de CDR también desempeñan un papel en el plegamiento y la expresión de scFv, sin embargo, se cree que el nivel de expresión de los clones presentaría todavía algo de variabilidad. Para minimizar estas diferencias, se introdujo variabilidad sólo dentro de los bucles CDR3 porque estos bucles son los más variables en anticuerpos y por tanto es más probable que sean altamente tolerantes a variaciones de secuencia y longitud. La introducción de variabilidad en la CDR3 es suficiente para generar anticuerpos contra la mayoría de proteínas. Además, se sesgaron cuidadosamente las frecuencias de los aminoácidos en estos bucles de modo que se recupera la distribución observada entre secuencias humanas naturales. Cuando se compararon los niveles de expresión de clones seleccionados al azar en el citoplasma, a pesar de algunas claras diferencias, una alta proporción de los mismos se expresaban correctamente tanto en *E. coli* como en células de mamíferos. Esta proporción de scFv expresados intracelularmente es mucho mayor que en bibliotecas descritas anteriormente.

También pudieron aislarse elementos de unión (es decir, anticuerpos que se unen específicamente a un antígeno diana seleccionado) frente a cinco proteínas diferentes (Aurora-A, GST:Syk, tubulina, histonas y proteína E6 del virus del papiloma HPV16). En estudios posteriores, también se han aislado elementos de unión frente a nuevas dianas incluyendo gankirina y MAPK11-14. Para MAPK11-14, las cuatro proteínas implicadas son las cuatro

isoformas de la MAP cinasa p38. Estas proteínas son muy homólogas (~60-74% de identidad). En todos los casos, pudieron aislarse scFv específicos para la isoforma usada para la selección. Esto subraya la especificidad de los scFv que pueden aislarse a partir de esta biblioteca.

5 Una preocupación frecuente cuando se construyen bibliotecas de scFv es la optimización simultánea del tamaño y la diversidad de la biblioteca. Generalmente, el tamaño de una biblioteca de este tipo está limitado por la eficacia de transformación a aproximadamente 10^{10} clones. Dado este número limitado de clones, es por tanto de gran importancia evitar duplicados o scFv no expresados. Para solucionar este problema, se seleccionó en primer lugar un entramado de anticuerpo que ya estaba optimizado para expresión intracelular, y entonces se usó un procedimiento de dos etapas para optimizar adicionalmente la biblioteca.

10 En primer lugar, se construyeron 18 bibliotecas “pequeñas” para cada longitud de CDR3 (13 longitudes de CDR3 VH y 5 longitudes de CDR3 VL) que van a usarse para construir la biblioteca de scFv (véase el ejemplo 1, en el presente documento más adelante). Se preparó cada una de estas 18 bibliotecas reemplazando la CDR3 correspondiente del scFv13R4 original por la CDR3 aleatorizada. Entonces se fusionó la biblioteca resultante con una y sólo una CDR3 aleatorizada con el gen que codifica para cloranfenicol acetiltransferasa (CAT), que se usó como marcador seleccionable. Tras la transformación de cada biblioteca de CDR3 en *E. coli*, se sembraron en placa las bibliotecas en medio que contenía cloranfenicol (CAM) para seleccionar las fusiones CDR3-CAT que se expresaban en el citoplasma de *E. coli*. Esta etapa redujo la diversidad de cada biblioteca en aproximadamente el 10-30%.

20 Aunque se usa el gen de CAT en los ejemplos proporcionados en el presente documento, un experto habitual en la técnica entiende que pueden fusionarse ácidos nucleicos que codifican para cualquier marcador seleccionable apropiado con ácidos nucleicos que codifican para CDR3 y usarse en los métodos proporcionados en el presente documento para enriquecer bibliotecas de scFv para scFv que se expresan en células procariontas o eucariotas. Por ejemplo, los marcadores seleccionables para su uso en células bacterianas incluyen, pero no se limitan a, el gen de resistencia a kanamicina. Los ejemplos de marcadores seleccionables para su uso en células de mamífero incluyen pero no se limitan a, por ejemplo, el gen de resistencia a neomicina, el gen de resistencia a puromicina y el gen de resistencia a higromicina. Los ejemplos de marcadores seleccionables para células de levadura incluyen URA3, HIS3 y purE. Además, pueden usarse marcadores tales como proteína fluorescente verde (GFP) y derivados de la misma, beta-galactosidasa, luciferasa u otros marcadores luminiscentes, fluorométricos y/o colorimétricos, en todos los tipos de células, por ejemplo, con clasificación celular activada por fluorescencia (FACS), para enriquecer bibliotecas de scFv para clones que codifican para scFv que pueden expresarse. Un experto habitual en la técnica entenderá cómo usar las enseñanzas en el presente documento, junto con lo que se conoce en la técnica, para seleccionar un marcador seleccionable apropiado y un esquema de selección apropiado para construir las bibliotecas de scFv descritas en el presente documento.

35 En una segunda etapa, se usó PCR para ensamblar aleatoriamente las bibliotecas de VH y VL seleccionadas para generar la diversidad final, basándose en la hipótesis de que si se expresaba un clon de scFv recién generado que contenía una nueva secuencia de CDR3 en su región VH más la VL original de scFv13R4, y si se expresaba un clon de scFv recién generado que contenía una nueva secuencia de CDR3 en su región VL más la región VH original de scFv 13R4, entonces también se expresaría un clon de scFv recombinado que contenía la nueva secuencia de CDR3 en su región VH más la nueva secuencia de CDR3 en su región VL, dando como resultado de ese modo una biblioteca que contenía sólo scFv expresados. Este era el caso puesto que se expresaron al menos parcialmente 19 de 20 clones seleccionados al azar en una forma soluble en el citoplasma de *E. coli*. De manera importante, puesto que esta etapa de selección se realizó durante la construcción de la biblioteca, la diversidad de la biblioteca final estaba limitada sólo por la transformación final. Esta etapa de recombinación final, al generar una alta diversidad, garantizó que todos los clones eran únicos en la biblioteca final. Conjuntamente, este enfoque dio como resultado una biblioteca de 1,5 billones de scFv expresados y diferentes.

45 El uso satisfactorio de scFv como intracuerpos a gran escala requiere que la biblioteca cumpla varios puntos esenciales. En primer lugar, el scFv debe ser fácil de aislar. Este es el caso de los métodos descritos en el presente documento, puesto que no sólo podían aislarse elementos de unión frente a todas las proteínas sometidas a prueba, sino que también un único ciclo de selección era suficiente para conseguir casi el 100% de los elementos de unión. Esta tasa de enriquecimiento muy alta se debe presumiblemente tanto a la alta calidad de la biblioteca sesgada que contiene sólo scFv bien expresados como al uso del fago auxiliar sensible a tripsina KM13. Esto es de gran importancia puesto que puede ser posible usar un único ciclo de selección antes de las pruebas *in vivo* de los scFv como intracuerpos, lo que permite una mejor diversidad de los epítopos seleccionados como diana. Además, esta alta tasa de enriquecimiento también reduce en una gran suma la cantidad de antígeno purificado necesaria para las etapas de selección. Pudo realizarse la selección en placas de microtitulación con tan sólo 1 µg de proteína por pocillo. Puesto que no son necesarias la selección y confirmación adicionales de la actividad de unión mediante ELISA debido a la muy alta proporción de elementos de unión, es posible ahora aislar buenos intracuerpos con una cantidad muy pequeña de antígeno. En segundo lugar, el scFv debe poder plegarse en todos los compartimentos celulares, particularmente en los reductores tales como el citoplasma y el núcleo. De nuevo, este es el caso de la biblioteca de scFv descrita en el presente documento, puesto que más del 80% de los clones sometidos a prueba son al menos parcialmente solubles en la célula. Además, se ha mostrado que pueden obtenerse buenos elementos

de unión citoplasmáticos a partir de los scFv seleccionados en fagos en *E. coli* y en células eucariotas.

Se proporcionan en el presente documento bibliotecas de anticuerpos que comprenden al menos aproximadamente: 10^6 , 5×10^6 , 10^7 , 5×10^7 , 10^8 , 5×10^8 , 10^9 , $1,5 \times 10^9$, 5×10^9 , 10^{10} , 5×10^{10} , 10^{11} , 5×10^{11} , 10^{12} , 10^{13} , 10^{14} , 10^{15} ó 10^{16} clones de scFv únicos, en las que al menos aproximadamente: el 20%, el 30%, el 40%, el 50%, el 60%, el 70%, el 80%, el 85% o más de los clones de scFv codifican para un anticuerpos que puede unirse específicamente de manera detectable a un antígeno diana cuando el anticuerpo codificado por el clon de scFv se expresa dentro de una célula (por ejemplo, como un intracuerpo). El scFv descrito en el presente documento puede expresarse para unirse específicamente de manera detectable a un antígeno diana en cualquier célula procariota o eucariota en la que sería deseable unirse específicamente de manera detectable a un antígeno diana (por ejemplo, pero sin limitarse a, una célula bacteriana, una célula de levadura, una célula de insecto, una célula de anfibio, una célula de ave, una célula de mamífero y similar).

Por "se une específicamente" quiere decirse que el anticuerpo scFv se une preferentemente a su antígeno diana en vez de a otro antígeno. Por "se une específicamente de manera detectable" quiere decirse que puede observarse la unión específica del anticuerpo scFv a su antígeno diana, por ejemplo, pero sin limitarse a un cambio fenotípico en la célula en la que el anticuerpo scFv se une a su antígeno diana, o detección de la interacción entre el anticuerpo scFv y su antígeno diana (por ejemplo, colocalización del anticuerpo scFv y su antígeno diana dentro de una célula).

Usos a modo de ejemplo

Aunque las bibliotecas descritas en el presente documento se han optimizado para el aislamiento de intracuerpos, las bibliotecas pueden usarse también para seleccionar scFv para aplicaciones de diagnóstico y terapéuticas. Además, debido a que se diseñó la diversidad de CDR3 usando secuencias humanas expresadas, los scFv presentes en la biblioteca son completamente humanos y no deben inducir una respuesta de anticuerpos anti-scFv en pacientes.

Las bibliotecas descritas en el presente documento y los anticuerpos producidos por las mismas son útiles no sólo para identificar intracuerpos funcionales sino también para el aislamiento de scFv altamente expresados que podrían usarse en numerosas aplicaciones biotecnológicas y terapéuticas. Por ejemplo, la biblioteca y los anticuerpos producidos por las mismas pueden tener usos relacionados con o que incluyen, pero sin limitarse a, estrategias de administración génica de agentes terapéuticos, herramientas de descubrimiento de fármacos, contrarrestar la aglutinación de moléculas diana no deseadas, combatir estados patológicos relacionados con trastornos de proteínas plegadas de manera errónea, y unión, neutralización o modificación de la función de una diana relacionada con cáncer.

Por ejemplo, los intracuerpos descritos en el presente documento pueden usarse para modular la actividad de la tirosina cinasa syk y otras proteínas implicadas en trastornos alérgicos (véase, por ejemplo, el documento WO2005106481; véase también Ulanova *et al.*, Expert Opin Ther Targets 2005, 9:901-921. Las MAP cinasas (MAPK) son mediadores clave de la proliferación celular y a menudo se seleccionan como diana para la inhibición en la terapia contra el cáncer (véase, por ejemplo, Sebolt-Leopold JS y Herrera, R, Nat Rev Cancer, 4:937-947 (2004)) Otras dianas interesantes son microtúbulos y proteínas asociadas (véase Jordan MA y Wilson, L, Nat Rev Cancer, 4:253-265 (2004)).

Pueden usarse inmunocuerpos para tratar o prevenir enfermedades en plantas valiosas comercialmente, tales como cultivos, por ejemplo, usando los métodos descritos en Villani, ME *et al.* Immunomodulation of cucumber mosaic virus infection by antibodies selected in vitro from a stable single-framework phage display library, Plant Molecular Biology 58(3):305 (2005).

Los intracuerpos descritos en el presente documento pueden usarse para tratar o prevenir infecciones en células humanas o animales. Por ejemplo, pueden usarse intracuerpos para tratar, mejorar o prevenir la infección de células por el virus de la inmunodeficiencia humana, usando métodos tal como se describe, por ejemplo, en Swan, CH *et al.* T-cell protection and enrichment through lentiviral CCR5 gene delivery, Gene Ther. 13:1480-1492.

Los intracuerpos descritos en el presente documento pueden usarse para seleccionar como diana proteínas implicadas en trastornos neurológicos, por ejemplo, tal como se describe, por ejemplo, en Miller, TW *et al.* (A human single-chain Fv intrabody preferentially targets amino-terminal Huntingtin's fragments in striatal models of Huntington's disease, Neurobiol Dis. 19:47-56 (2005)) y Miller y Messer (Intrabody applications in neurological disorders: progress and future prospects, Mol. Ther. 12:394-401 (2005).

Los intracuerpos descritos en el presente documento pueden usarse para terapia contra el cáncer seleccionando como diana una proteína implicada en cáncer como oncoproteínas, tal como se describe, por ejemplo, en Williams, BR y Zhu, Z (Intrabody-based approaches to cancer therapy: status and prospects, Current med Chem 13:1473-1480 (2006) y Doorbar J. y Griffin H. (2007) Intrabody strategies for the treatment of human papillomavirus-

associated disease. *Expert Opin. Biol. Ther.* 7(5), 677-689.

Los intracuerpos descritos en el presente documento pueden usarse para tratar infecciones (por ejemplo pero sin limitarse a virus de Epstein-Barr) seleccionando como diana una proteína expresada por un agente infeccioso, por ejemplo, tal como se describe en Fang CY *et al* (Modulation of Epstein-Barr Virus Latent Membrane Protein 1 Activity by Intrabodies, *Intervirology* 50:254-263 (2007).

Tal como se describe en el presente documento, puede administrarse un intracuerpo a una célula administrando a la célula un vector de expresión que codifica para el intracuerpo de interés. Se conocen bien en la técnica vectores de expresión que son adecuados para la expresión de intracuerpos. Puede lograrse la administración de vectores de expresión que codifican para los intracuerpos descritos en el presente documento mediante uno cualquiera de numerosos enfoques bien conocidos, por ejemplo, pero sin limitarse a, transferencia directa de los ácidos nucleicos, en un plásmido o vector de expresión viral, solo o en combinación con portadores tales como liposomas catiónicos. Tales vectores de expresión (que contienen secuencias promotoras y potenciadoras adecuadas para expresar una secuencia codificante operativamente unida cuando el vector de expresión se introduce en una célula) y métodos para preparar, usar y administrar tales vectores a células se conocen bien en la técnica y pueden adaptarse fácilmente para su uso para administrar intracuerpos a células.

Los vectores pueden ser cualquier construcción de nucleótidos usada para administrar genes a células, por ejemplo un plásmido o vector viral, tal como un vector retroviral que puede empaquetar un genoma retroviral recombinante (véase por ejemplo, Pastan *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 85:4486, 1988; Miller *et al.*, *Mol. Cell. Biol.* 6:2895, 1986). El retrovirus recombinante puede usarse entonces para infectar y administrar de ese modo un ácido nucleico de la invención a las células infectadas. El método exacto de introducción del ácido nucleico alterado en células de mamífero no se limita, por supuesto, al uso de vectores retrovirales. Otras técnicas están ampliamente disponibles para este procedimiento incluyendo el uso de vectores adenovirales (Mitani *et al.*, *Hum. Gene Ther.* 5:941-948, 1994), vectores de virus adenoasociados (VAA) (Goodman *et al.*, *Blood* 84:1492-1500, 1994), vectores lentivirales (Naidini *et al.*, *Science* 272:263-267, 1996), vectores retrovirales pseudotipados (Agrawal *et al.*, *Exper. Hematol.* 24:738-747, 1996). También pueden usarse técnicas de transducción física, tales como administración de liposomas y mecanismos de endocitosis mediada por receptor y otros (véase, por ejemplo, Schwartzenberger *et al.*, *Blood* 87:472-478, 1996). Esta invención puede usarse conjuntamente con cualquiera de estos u otros métodos de transferencia génica usados comúnmente. Se describen medios apropiados para la transfección, incluyendo vectores virales, transfectantes químicos o métodos fisicoquímicos tales como electroporación y difusión directa de ADN por, por ejemplo, Wolff, J. A., *et al.*, *Science*, 247, 1465-1468, (1990); y Wolff, J. A. *Nature*, 352, 815-818, (1991).

Pueden administrarse intracuerpos permeables en células (transcuerpos) a células fusionando un anticuerpo scFv con un dominio de transducción de proteínas (PTD) que permite que el intracuerpo permeable en células cruce la membrana celular y entre en la célula, por ejemplo, según los métodos descritos en Heng, BC y Cao, T (Making cell-permeable antibodies (Transbody) through fusion of protein transduction domains (PTD) with single chain variable fragment (scFv) antibodies: potential advantages over antibodies expressed within the intracellular environment (Intrabody), *Med Hypotheses* 64:1105-1108 (2005)). Alternativamente, el scFv podría mezclarse con lípidos catiónicos y administrarse eficazmente a células, tal como se muestra con anticuerpos completos (Courtète J., Sibling A.P., Zeder-Lutz G., Dalkara D., Zuber G. & Weiss E. (2007) Suppression of cervical carcinoma cell growth by intracytoplasmic co-delivery of anti-oncoprotein E6 antibody and siRNA. *Mol. Cancer Ther.* 6, 1728-36). Tales intracuerpos permeables en células pueden usarse en cultivo celular (por ejemplo, para fines de investigación) y para fines de diagnóstico (por ejemplo, para detectar un virus o microorganismo en una muestra de células de un ser humano, animal o planta que se sospecha que alberga un agente infeccioso. Tales intracuerpos permeables en células pueden administrarse a animales de investigación (por ejemplo pero sin limitarse a administración sistémica, por ejemplo, mediante administración por vía intravenosa de un intracuerpo a un animal de investigación, o mediante administración tópica, por ejemplo) para modular la actividad de un antígeno diana particular y alterar de ese modo un fenotipo en el animal. Tales intracuerpos permeables en células pueden administrarse también a pacientes animales humanos o no humanos para tratar o prevenir una enfermedad o infección tal como se describió anteriormente. Por ejemplo, pueden administrarse intracuerpos permeables en células por vía intravenosa, por vía tópica, por vía oral o mediante otros métodos bien conocidos, tal como apreciará un experto habitual en la técnica.

Uso de los presentes anticuerpos scFv para la producción de anticuerpos a gran escala

Debido a que los clones de anticuerpos scFv codifican para anticuerpos que se pliegan en condiciones reductoras, las bibliotecas descritas en el presente documento también son útiles para identificar anticuerpos scFv que pueden producirse en grandes cantidades expresando los anticuerpos scFv en células (por ejemplo, células procariotas tales como células bacterianas, o células eucariotas tales como células de levadura, células de insecto, células de mamífero o similares) y aislando los anticuerpos scFv a partir de las células. Los anticuerpos scFv de cadena sencilla para los que sería deseable purificar grandes cantidades de anticuerpo incluyen pero no se limitan a, por ejemplo, anticuerpos scFv que son útiles en investigación de laboratorio, para pruebas de diagnóstico médico, para pruebas de diagnóstico comerciales u otros tipos de pruebas de diagnóstico (por ejemplo, para detectar

microorganismos contaminantes en alimentos o agua potable), y para productos terapéuticos de anticuerpos (por ejemplo, para tratar cáncer o para tratar enfermedades infecciosas u otros tipos de enfermedades en las que puede usarse un anticuerpo para tratar la enfermedad).

Se proporcionan en el presente documento bibliotecas de anticuerpos para el aislamiento de anticuerpos scFv expresados, comprendiendo la biblioteca de anticuerpos al menos aproximadamente: 10^6 , 5×10^6 , 10^7 , 5×10^7 , 10^8 , 5×10^8 , 10^9 , $1,5 \times 10^9$, 5×10^9 , 10^{10} , 5×10^{10} , 10^{11} , 5×10^{11} , 10^{12} , 10^{13} , 10^{14} , 10^{15} ó 10^{16} clones de anticuerpos scFv únicos, en las que al menos aproximadamente: el 20%, el 25%, el 30%, el 35%, el 40%, el 50%, el 55%, el 60%, el 65%, el 70%, el 75%, el 80% o más del 80% de los clones de anticuerpos scFv pueden expresarse dentro de una célula para producir anticuerpo soluble a un nivel de al menos aproximadamente: 5 mg, 8 mg, 10 mg, 15 mg, 20 mg, 25 mg, 30 mg, 35 mg, 40 mg, 45 mg, 50 mg o más de 50 mg por litro de células en un frasco hechas crecer hasta una DO_{600nm} de aproximadamente 5.

Hay muchos enfoques bien aceptados para expresar y purificar proteínas tales como anticuerpos scFv expresados a partir de células y un experto habitual en la técnica entenderá cómo seleccionar el tipo más apropiado de sistema de expresión celular a partir del cual expresar y purificar anticuerpos scFv. Se conocen en la técnica muchos manuales que describen métodos para la expresión de proteínas. Véase, por ejemplo, *Current Protocols in Molecular Biology*, John Wiley and Sons, 2006.

Definiciones

Por "secuencia única" quiere decirse que, dentro de una colección de clones de anticuerpos scFv, hay al menos 10^6 clones que contienen una secuencia de CDR3 que es diferente de las secuencias de CDR3 de otros clones dentro de la colección de clones de anticuerpos scFv. Con referencia a un clon de anticuerpo scFv o una biblioteca de tales clones, una "secuencia única" es diferente de la secuencia presente en la posición correspondiente dentro de scFv13R4.

Los términos "entramado de anticuerpo" y "secuencia de entramado" se refieren a cualquier parte no única de los clones de anticuerpos scFv, por ejemplo, cualquier parte del anticuerpo scFv que no es una región VH CDR3 única y/o una región VL CDR3 única. Tal como se menciona en el presente documento, el entramado de anticuerpo o secuencia de entramado es la de scFv13R4, el clon de scFv original en el que se basan las presentes bibliotecas.

Por "sustancialmente idéntica" quiere decirse que dos o más secuencias de aminoácidos que están comparándose son al menos el 98%, el 98,5%, el 99% o el 99,5% iguales, o que dos o más secuencias de ácido nucleico que están comparándose codifican para secuencias de aminoácido que son al menos el 98%, el 98,5%, el 99% o el 99,5% iguales.

Por "secuencia común" quiere decirse una secuencia de nucleótidos o aminoácidos que se comparte entre clones de scFv.

Por "clon de anticuerpo scFv" quiere decirse una molécula de ácido nucleico que codifica para una especie individual de anticuerpo scFv que comprende una secuencia única dentro del dominio CDR3 VH, el dominio CDR3 VL o dentro de ambos dominios CDR3 VH y VL.

Por "antígeno diana" quiere decirse un antígeno al que se une preferentemente un anticuerpo particular, en comparación con la unión de ese anticuerpo particular a otro antígeno que no es un antígeno diana.

Por "se une específicamente" quiere decirse que un anticuerpo se une con fuerza y preferentemente a un antígeno diana particular, en comparación con la unión del anticuerpo a otros antígenos.

Por "expresado de manera ectópica" quiere decirse que la expresión de un anticuerpo de interés no se expresa de manera natural dentro de un tipo particular de células en el que el anticuerpo está expresándose, es decir, el anticuerpo se expresa dentro de la célula porque se ha introducido en la célula un constructo de expresión que codifica para el anticuerpo.

Por "aislado" o "purificado" quiere decirse que un anticuerpo scFv se ha separado sustancialmente, producido aparte de o purificado de otros componentes biológicos en la células en la que se ha producido, es decir, se ha separado sustancialmente de otros componentes celulares, tales como otras proteínas celulares, ADN, ARN, lípidos y similares. El término "aislado" o "purificado" no requiere pureza absoluta; más bien, está previsto como un término relativo. Preferiblemente, un anticuerpo scFv se purifica o se aísla de otros componentes celulares de manera que el anticuerpo scFv representa al menos: el 25%, el 30%, el 40%, el 50%, el 60%, el 70%, el 80%, el 90%, el 95% o más del contenido total de la preparación de anticuerpos scFv.

Por anticuerpo "soluble" quiere decirse que una molécula de anticuerpo no está en forma de agregado. Un

anticuerpo soluble tiene la capacidad para unirse específicamente a su antígeno diana.

Por “sustancialmente soluble” quiere decirse que al menos el 20%, por ejemplo, al menos el 25%, el 30%, el 40%, el 50%, el 60%, el 70%, el 80% o más de un anticuerpo scFv tal como se describe en el presente documento está plegado apropiadamente y por tanto puede unirse específicamente a su antígeno diana.

5 Ejemplo 1: Construcción de una biblioteca de scFv optimizada para expresión intracelular

MATERIALES

Cepas bacterianas, productos químicos y enzimas

Se describieron anteriormente medios LB y 2xYT (Miller, J.H. A Short Course in Bacterial Genetics: a Laboratory Manual and Handbook for Escherichia coli and Related Bacteria. Cold Spring Harbor Laboratory Press; 1992). La cepa C-Max5F' (Bio-rad laboratories) es *E. coli* K-12, [F' *lacI^f Tn10*] ϕ 80d*lacZ* Δ M15 Δ (*lacZYA-argF*)U169 *recA1 endA1 hsdR17* (*r_k⁻ m_k⁺*) *phoA supE44 λ - thi-1 gyrA96 relA1*. MC1061 (ATCC n.º 53338) es *E. coli* K-12, F- λ - *hsdR2 hsdM+ hsdS+ mcrA mcrB1 araD139 Δ (ara-leu)7696 Δ (lacIPOZY)X74 galE15 galUgalK16 rpsL thi*. La cepa no supresora HB2151 es *E. coli* K-12 [F'*proA⁺B⁺ lacI^f lacZ* Δ M15] *ara Δ (lac-pro) thi*. Los productos químicos se adquirieron de Sigma. Las enzimas de restricción y la *Taq* polimerasa clonada eran de Fermentas. Se adquirieron las ProofStart y *Pfu* ADN polimerasas respectivamente de Qiagen y Promega. Se purificaron ADN de plásmido, ADN separado por PCR y agarosa usando los kits Macherey-Nagel Nucleospin.

Oligonucleótidos

Se sintetizaron dieciocho oligonucleótidos escogidos usados para introducir bucles CDR3 degenerados y se purificaron usando cromatografía de líquidos de alta resolución (HPLC) por IBA GmbH (Goettingen, Alemania). Las secuencias de los 18 oligonucleótidos escogidos es tal como sigue.

H3_n = bucle CDR3 de VH de n aminoácidos de longitud; K3_n = bucle CDR3 de VL kappa (κ) de n aminoácidos de longitud; L3_n = bucle CDR3 de VL lambda (λ) de n aminoácidos de longitud. Para las posiciones degeneradas, los porcentajes de las 4 bases se facilitan como N (A/C/G/T). Las proporciones de cada nucleótido usado en cada posición escogida (degenerada) se muestra en la tabla I a continuación.

25 H3_5:

AGGGTGCCTCTGCCCCANNNNNNNNNNNNNNNNTCTCACACAGTAATAAACAGCCG (SEQ ID NO: 1);

H3_6:

AGGGTGCCTCTGCCCCANNNNNNNNNNNNNNNNNNNNNNTCTCACACAGTAATAAACAGC
CG (SEQ ID NO: 2);

H3_7:

AGGGTGCCTCTGCCCCANNNNNNNNNNNNNNNNNNNNNNNNTCTCACACAGTAATAAAC
AGCCG (SEQ ID NO: 3);

H3_8:

AGGGTGCCTCTGCCCCANNNNNNNNNNNNNNNNNNNNNNNNTCTCACACAGTAAT
AAACAGCCG (SEQ ID NO: 4);

H3_9:

AGGGTGCCTCTGCCCCANNNNNNNNNNNNNNNNNNNNNNNNTCTCACACAGT
AATAAACAGCCG (SEQ ID NO: 5);

30

H3_10:

AGGGTGCCTCTGCCCCANNTCTCACAC
AGTAATAAACAGCCG (SEQ ID NO: 6);

H3_11:

AGGGTGCCTCTGCCCCANNTCTCA
CACAGTAATAAACAGCCG (SEQ ID NO: 7);

H3_12:

AGGGTGCCTCTGCCCCANNN
TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 8);

H3_13:

AGGGTGCCTCTGCCCCANNN
NTCTCACACAGTAATAAACAGCCG (SEQ ID NO: 9);

5

H3_14:

AGGGTGCCTCTGCCCCANNN
NNNNTCTCACACAGTAATAAACAGCCG (SEQ ID NO: 10);

H3_15:

AGGGTGCCTCTGCCCCANNN
NNNNNNNTCTCACACAGTAATAAACAGCCG (SEQ ID NO: 11);

H3_16:

AGGGTGCCTCTGCCCCANNN
NNNNNNNNNTCTCACACAGTAATAAACAGCCG (SEQ ID NO: 12);

H3_17:

AGGGTGCCTCTGCCCCANNN
NNNNNNNNNNNTCTCACACAGTAATAAACAGCCG (SEQ ID NO: 13);

K3_9:

GGACGAGGCTGATTACTGCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTTCGGC
GGAGGGACCAAG (SEQ ID NO: 14);

10

K3_10:

GGACGAGGCTGATTATTACTGCNNNTTC
GGCGGAGGGACCAAG (SEQ ID NO: 15);

L3_9:

GGACGAGGCTGATTATTACTGCNNNTTCGGC
GGAGGGACCAAG (SEQ ID NO: 16);

L3_10:

GGACGAGGCTGATTATTACTGCNNNTTC
GGCGGAGGGACCAAG (SEQ ID NO: 17);

L3_11:

GGACGAGGCTGATTATTACTGCNN
TTCGGCGGAGGGACCAAG (SEQ ID NO: 18).

5 **Tabla 1**

* Para las posiciones degeneradas, los porcentajes de las 4 bases se facilitan como N(A/C/T)

H3_5

AGGGTGCCTCTGCCCCA N(40/5/50/5) N(5/5/5/85) N(60/5/25/10) N(40/10/45/5)
N(15/30/5/50) N(5/75/10/10) N(30/35/30/5) N(50/20/10/20)N(60/20/5/15)
N(30/25/40/5) N(15/20/25/40) N(35/40/5/20) N(25/50/20/5) N(10/50/10/30)
N(5/55/35/5) TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 1)

H3_6

AGGGTGCCTCTGCCCCA N(55/5/35/5) N(10/5/5/80) N(75/15/5/5) N(45/5/45/5)
N(5/10/10/75) N(5/85/5/5) N(45/35/15/5) N(55/20/10/15)N(55/20/10/15)
N(40/30/25/5) N(10/40/15/35) N(25/45/5/25) N(45/30/20/5) N(15/40/20/25)
N(10/35/30/25) N(20/15/60/5) N(10/60/10/20) N(5/70/10/15)
TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 2)

H3_7

AGGGTGCCTCTGCCCCA N(35/5/55/5) N(15/5/15/65) N(65/15/10/10)
N(55/5/35/5) N(5/5/5/85) N(5/85/5/5) N(50/5/40/5) N(75/5/10/10) N(60/10/20/10)
N(10/45/40/5) N(15/35/25/25) N(25/55/5/15) N(35/35/25/5) N(10/45/15/30)
N(25/45/15/15) N(25/25/25/25) N(20/40/15/25)N(10/35/25/30) N(40/15/40/5)
N(5/50/15/30) N(5/80/10/5) TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 3)

H3_8

AGGGTGCCTCTGCCCCA N(55/5/35/5) N(10/5/10/75) N(75/10/5/10) N(50/5/40/5)
N(5/5/5/85) N(5/85/5/5) N(45/5/45/5) N(75/5/10/10) N(65/5/20/10) N(25/45/25/5)
N(10/25/40/25) N(30/45/5/20) N(55/10/30/5) N(5/55/15/25) N(5/55/15/25)
N(35/35/25/5) N(15/45/20/20) N(15/40/10/35) N(30/40/25/5) N(20/35/20/25)
N(15/40/20/25) N(65/5/20/10) N(15/45/5/35) N(5/70/15/10)
TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 4)

H3_9

AGGGTGCCTCTGCCCCA N(40/5/50/5) N(10/5/10/75) N(75/5/15/5) N(40/5/50/5)

N(5/5/5/85) N(5/85/5/5) N(45/5/45/5) N(85/5/5/5) N(70/5/15/10) N(55/25/15/5)
N(10/25/35/30) N(35/45/15/5) N(10/25/60/5) N(10/35/20/35) N(20/45/15/20)
N(30/30/35/5) N(15/40/20/25) N(20/40/15/25) N(65/20/10/5) N(20/40/20/20)
N(5/45/15/35) N(15/30/50/5) N(20/45/20/15) N(5/40/25/30) N(30/25/40/5)
N(10/40/10/40) N(5/60/20/15) TCTCACACAGTAATAAACAGCCG (SEQ ID NO:
5)

H3_10

AGGGTGCCTCTGCCCCA N(50/5/40/5) N(20/5/10/65) N(70/10/15/5) N(60/5/30/5)
N(5/5/5/85) N(5/85/5/5) N(35/5/55/5) N(85/5/5/5) N(75/5/15/5) N(30/35/30/5)
N(10/20/45/25) N(30/45/15/10) N(30/35/30/5) N(5/30/20/45) N(30/40/10/20)
N(40/35/20/5) N(15/35/25/25) N(25/35/15/25) N(50/30/15/5) N(10/45/20/25)
N(20/50/10/20) N(25/20/50/5) N(20/40/20/20) N(5/50/15/30) N(20/20/40/20)
N(20/35/20/25) N(5/35/30/30) N(5/30/60/5) N(15/35/10/40) N(5/60/15/20)
TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 6)

H3_11

AGGGTGCCTCTGCCCCA N(55/5/35/5) N(10/5/10/75) N(80/5/10/5) N(60/5/30/5)
N(5/5/5/85) N(5/85/5/5) N(45/5/45/5) N(85/5/5/5) N(75/10/10/5) N(40/30/25/5)
N(10/20/35/35) N(35/35/20/10) N(40/30/25/5) N(10/25/20/45) N(35/35/15/15)
N(35/30/30/5) N(15/30/25/30) N(35/30/15/20) N(20/35/40/5) N(20/30/25/25)
N(30/40/10/20) N(5/35/50/10) N(20/30/30/20) N(30/40/15/15) N(35/35/25/5)
N(20/35/20/25) N(20/40/15/25) N(45/25/10/20) N(20/35/20/25) N(20/30/25/25)
N(45/25/25/5) N(15/35/10/40) N(5/60/20/15) TCTCACACAGTAATAAACAGCCG
(SEQ ID NO: 7)

H3_12

AGGGTGCCTCTGCCCCA N(45/5/45/5) N(15/5/15/65) N(70/5/15/10) N(40/5/50/5)
N(5/5/5/85) N(5/85/5/5) N(45/5/45/5) N(85/5/5/5) N(75/5/15/5) N(10/30/55/5)
N(10/30/25/35) N(40/40/15/5) N(25/25/45/5) N(10/25/20/45) N(35/35/10/20)
N(40/40/15/5) N(20/20/35/25) N(40/30/15/15) N(10/40/45/5) N(20/25/30/25)
N(35/35/15/15) N(30/30/35/5) N(15/40/20/25) N(20/45/5/30) N(15/25/40/20)
N(20/30/25/25) N(20/40/15/25) N(15/25/40/20) N(20/35/20/25) N(25/40/15/20)
N(50/30/15/5) N(20/45/20/15) N(5/35/35/25) N(45/25/25/5) N(15/35/10/40)
N(5/65/15/15) TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 8)

H3_13

AGGGTGCCTCTGCCCCA N(45/10/40/5) N(15/5/15/65) N(75/5/15/5) N(35/5/55/5)

N(5/5/5/85) N(5/85/5/5) N(45/5/45/5) N(85/5/5/5) N(80/5/10/5) N(35/25/35/5)
N(10/30/25/35) N(40/45/10/5) N(15/25/55/5) N(10/25/10/55) N(35/35/10/20)
N(55/30/10/5) N(15/20/25/40) N(45/25/10/20) N(25/30/40/5) N(20/30/20/30)
N(30/35/15/20) N(15/25/55/5) N(15/35/20/30) N(25/45/10/20) N(10/45/40/5)
N(20/25/25/30) N(30/40/10/20)N(30/25/30/15) N(20/25/25/30) N(35/35/15/15)
N(25/30/30/15) N(20/30/20/30) N(25/35/20/20) N(40/30/25/5) N(20/45/20/15)
N(5/35/35/25) N(45/25/25/5) N(10/35/15/40) N(5/65/20/10)
TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 9)

H3_14

AGGGTGCCTCTGCCCA N(50/5/40/5) N(20/5/15/60) N(60/10/25/5) N(35/5/55/5)
N(5/5/5/85) N(5/85/5/5) N(30/20/45/5) N(85/5/5/5) N(80/5/5/10) N(30/35/30/5)
N(10/35/20/35) N(40/40/15/5) N(50/25/20/5) N(10/25/10/55) N(40/30/10/20)
N(20/20/55/5) N(15/20/20/45) N(45/25/15/15) N(15/25/55/5) N(15/30/25/30)
N(35/30/15/20) N(40/30/25/5) N(15/25/25/35) N(35/35/10/20) N(20/30/40/10)
N(15/30/25/30) N(25/40/10/25)N(50/25/15/10) N(15/35/25/25) N(35/35/10/20)
N(15/25/45/15) N(15/30/25/30) N(35/35/10/20) N(15/30/40/15) N(20/30/20/30)
N(30/40/10/20) N(30/30/15/25) N(20/35/25/20) N(20/30/30/20) N(30/25/40/5)
N(10/30/15/45) N(5/70/15/10) TCTCACACAGTAATAAACAGCCG (SEQ ID NO:
10)

H3_15

AGGGTGCCTCTGCCCA N(30/5/45/20) N(45/5/5/45) N(45/35/5/15) N(45/5/45/5)
N(5/5/5/85) N(5/85/5/5) N(45/20/30/5) N(85/5/5/5) N(85/5/5/5) N(20/35/40/5)
N(5/40/25/30) N(35/50/10/5) N(15/5/75/5) N(10/10/20/60) N(50/30/10/10)
N(25/10/55/10) N(15/20/20/45) N(50/25/15/10) N(35/30/30/5) N(15/25/25/35)
N(50/20/15/15) N(35/5/35/25) N(10/30/30/30) N(35/35/10/20) N(20/40/35/5)
N(20/20/35/25) N(35/35/15/15)N(45/35/15/5) N(15/30/25/30) N(35/35/10/20)
N(35/35/25/5) N(20/30/20/30) N(30/35/10/25) N(30/25/40/5) N(15/30/20/35)
N(35/30/15/20) N(35/25/35/5) N(15/35/20/30) N(25/35/25/15) N(5/30/45/20)
N(20/35/20/25) N(15/30/30/25) N(45/25/25/5) N(15/35/10/40) N(5/70/15/10)
TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 11)

H3_16

AGGGTGCCTCTGCCCA N(35/10/30/25) N(65/5/5/25) N(25/55/5/15)
N(45/5/45/5) N(5/5/5/85) N(5/85/5/5) N(35/30/30/5) N(85/5/5/5) N(75/5/5/15)
N(15/35/45/5) N(5/45/20/30) N(35/55/5/5) N(50/5/40/5) N(10/5/15/70)
N(60/20/10/10) N(65/5/25/5) N(10/10/25/55) N(55/15/15/15) N(30/5/60/5)
N(10/15/30/45) N(55/15/15/15) N(25/25/45/5) N(20/30/15/35) N(40/35/10/15)
N(40/20/30/10) N(20/30/25/25) N(35/30/10/25)N(45/30/15/10) N(15/35/25/25)

N(30/35/10/25) N(35/30/25/10) N(20/25/30/25) N(35/35/10/20) N(30/20/30/20)
N(20/35/20/25) N(30/40/10/20) N(45/10/25/20) N(20/25/20/35) N(30/35/15/20)
N(45/30/20/5) N(15/35/25/25) N(25/35/20/20) N(40/15/5/40) N(20/35/20/25)
N(15/30/30/25) N(30/30/35/5) N(15/30/15/40) N(5/75/15/5)
TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 12)

H3_17

AGGGTGCCTCTGCCCA N(45/10/20/25) N(70/5/5/20) N(15/60/5/20)
N(40/5/50/5) N(5/5/5/85) N(5/85/5/5) N(25/20/50/5) N(85/5/5/5) N(80/5/5/10)
N(10/45/40/5) N(5/50/20/25) N(30/60/5/5) N(5/5/85/5) N(5/5/15/75) N(65/20/5/10)
N(45/5/45/5) N(10/10/20/60) N(55/10/20/15) N(45/5/45/5) N(15/15/20/50)
N(60/15/15/10) N(60/20/15/5) N(10/30/20/40) N(45/25/15/15) N(15/35/45/5)
N(15/15/35/35) N(40/30/15/15) N(30/25/40/5) N(20/30/25/25) N(30/35/10/25)
N(20/30/40/10) N(15/35/25/25) N(30/35/10/25) N(40/30/25/5) N(20/30/25/25)
N(40/35/10/15) N(35/20/40/5) N(20/25/20/35) N(40/35/10/15) N(35/20/40/5)
N(20/20/20/40) N(35/35/15/15) N(30/20/30/20) N(20/30/20/30) N(25/30/25/20)
N(35/35/25/5) N(25/35/20/20) N(5/35/40/20) N(20/30/45/5) N(15/30/10/45)
N(5/70/15/10) TCTCACACAGTAATAAACAGCCG (SEQ ID NO: 13)

K3_9

GGACGAGGCTGATTACTGC N(5/85/5/5) N(85/5/5/5) N(5/5/85/5)
N(5/85/5/5) N(85/5/5/5) N(40/5/50/5) N(5/5/10/80) N(65/25/5/5) N(5/50/5/40)
N(30/5/25/40) N(60/10/20/10) N(5/40/10/45) N(80/5/10/5) N(30/15/50/5)
N(5/60/5/30) N(35/5/10/50) N(30/40/10/20) N(5/15/45/35) N(5/85/5/5) N(5/85/5/5)
N(5/5/85/5) N(5/40/5/50) N(30/10/25/35) N(5/35/50/10) N(85/5/5/5) N(5/85/5/5)
N(5/50/30/15) TTCGGCGGAGGGACCAAG (SEQ ID NO: 14)

K3_10

GGACGAGGCTGATTACTGC N(5/85/5/5) N(85/5/5/5) N(5/5/85/5)
N(5/85/5/5) N(85/5/5/5) N(40/5/50/5) N(5/5/5/85) N(80/10/5/5) N(25/20/5/50)
N(35/5/40/20) N(45/5/45/5) N(5/30/15/50) N(75/10/10/5) N(40/5/50/5) N(5/45/5/45)
N(20/5/5/70) N(10/45/40/5) N(5/5/85/5) N(5/85/5/5) N(5/85/5/5) N(5/5/85/5)
N(5/85/5/5) N(5/65/15/15) N(5/5/85/5) N(15/5/10/70) N(20/5/30/45) N(5/30/50/15)
N(85/5/5/5) N(5/85/5/5) N(5/40/5/50) TTCGGCGGAGGGACCAAG (SEQ ID NO:
15)

L3_9

GGACGAGGCTGATTACTGC N(5/85/5/5) N(80/5/5/10) N(5/5/85/5)
N(20/5/30/45) N(5/75/5/15) N(85/5/5/5) N(5/5/5/85) N(35/5/55/5) N(5/10/80/5)
N(5/10/80/5) N(80/10/5/5) N(5/60/5/30) N(65/5/25/5) N(15/10/70/5) N(5/65/5/25)
N(70/5/20/5) N(20/10/65/5) N(5/60/5/30) N(65/10/20/5) N(15/45/20/20)
N(5/40/10/45) N(15/5/50/30) N(10/15/25/50) N(5/20/70/5) N(10/5/80/5) N(5/5/5/85)
N(5/40/10/45)TTCGGCGGAGGGACCAAG (SEQ ID NO: 16)

L3_10

GGACGAGGCTGATTACTGC N(45/5/10/40) N(15/15/60/10) N(10/50/5/35)
N(5/5/5/85) N(5/85/5/5) N(45/5/45/5) N(5/5/5/85) N(85/5/5/5) N(5/50/5/40)
N(25/5/65/5) N(45/45/5/5) N(5/45/5/45) N(55/5/35/5) N(10/10/75/5) N(5/40/5/50)
N(85/5/5/5) N(15/10/70/5) N(5/50/5/40) N(70/5/20/5) N(30/20/45/5) N(5/75/5/15)
N(70/15/10/5) N(25/45/15/15) N(5/35/5/55) N(5/15/35/45) N(15/5/30/50)
N(5/15/75/5) N(10/5/80/5) N(5/5/5/85) N(5/50/10/35) TTCGGCGGAGGGACCAAG
(SEQ ID NO: 17)

L3_11

GGACGAGGCTGATTACTGC N(10/55/25/10) N(65/20/10/5) N(5/10/80/5)
N(20/5/20/55) N(5/85/5/5) N(5/5/85/5) N(5/5/5/85) N(30/5/60/5) N(5/5/85/5)
N(5/5/85/5) N(85/5/5/5) N(5/45/5/45) N(75/5/15/5) N(20/10/65/5) N(5/60/5/30)
N(85/5/5/5) N(5/5/85/5) N(5/35/5/55) N(5/5/5/85) N(5/30/10/55) N(85/5/5/5)
N(75/5/15/5) N(45/20/30/5) N(5/20/5/70) N(5/30/50/15) N(25/25/35/15)
N(5/45/15/35) N(10/5/50/35) N(20/5/25/50) N(5/20/60/15) N(5/5/85/5) N(5/5/5/85)
N(70/5/20/5)TTCGGCGGAGGGACCAAG (SEQ ID NO: 18)

Otros oligonucleótidos usados para construir las bibliotecas se sintetizaron mediante MWG y tienen las siguientes secuencias:

- T7.back CCGGATATAGTTCCTCCTTT (SEQ ID NO: 19);
- 5 T7.for CTGCTAACCAGTAAGGCAAC (SEQ ID NO: 20);
- M13rev-49 GAGCGGATAACAATTCACACAGG (SEQ ID NO: 21);
- M13uni-43 AGGGTTTTCCAGTCACGACGTT (SEQ ID NO: 22);
- scFvCAT.rev AACGGTGGTATATCCAGTGA (SEQ ID NO: 23);
- scFvCAT2.rev CGGTGGTATATCCAGTGATTTTT (SEQ ID NO: 24);
- 10 PliaisonH3 TGGGGCAGAGGCACCCT (SEQ ID NO: 25);
- PliaisonH3.back AGGGTGCCTCTGCCCCA (SEQ ID NO: 26);
- PliaisonL3 GCAGTAATAATCAGCCTCGTCC (SEQ ID NO: 27).

Plásmidos

Se usó el vector de fagémido pCANTAB6 (McCafferty J *et al.*, Appl Biochem Biotechnol 1994, 47:157-171) para la

fusión N-terminal de fragmentos de *NcoI/NotI*-scFv a la proteína de cubierta minoritaria pIII del fago filamentoso M13. Este fagémido se deriva de pUC 119 y contiene las siguientes secuencias en el siguiente orden: un promotor *lac*, la secuencia líder *pelB*, los sitios *NcoI* y *NotI* para la clonación de scFv, una etiqueta His6 y una c-myc reconocida por el anticuerpo monoclonal 9E10, un codón ámbar y la secuencia génica de pIII.

- 5 Para la expresión citoplasmática de los scFv en *E. coli*, se usó el plásmido pET23NN. Este plásmido se deriva de pET23d(+) (Novagen) y contiene un promotor de T7, seguido por un sitio *NcoI* que contiene el iniciador ATG, un sitio *NotI* seguido por una etiqueta c-myc y una His6.

El plásmido pscFv CAT se deriva de pTrc99A y contiene un promotor *tac*, seguido por un sitio *NcoI* que contiene el iniciador ATG de un scFv fuera de marca, un sitio *NotI* seguido por el gen de CAT. Cuando se inserta un scFv dentro de los sitios *NcoI-NotI*, el scFv se expresa como una fusión con la proteína CAT. Se realizó la construcción tal como sigue: en primer lugar, se eliminó el sitio *BstEII* único de pTrc99A (A13038) mediante digestión seguido por relleno del saliente en 5' para formar extremos romos, y ligamiento. Se digirió el plásmido resultante con *NcoI* y *NotI*, y se purificó el fragmento de 4210 pb (fragmento I). En segundo lugar, se eliminó el sitio *NcoI* único del plásmido pACYCI 84 (X06403) ubicado dentro del gen de CAT mediante mutagénesis dirigida al sitio cambiando el codón de Thr 172 de ACC a ACG. Entonces, se amplificó el gen de CAT mediante PCR usando CAT-NotI.for (TAAGGCGGCCGCAATGGAGAAAAAATCACTG; SEQ ID NO: 28) y CAT-HindIII.back (ACTGCCTTAAAAAGC**TTACGCC**; SEQ ID NO: 29). En las secuencias oligonucleotídicas, los sitios de restricción introducidos están subrayados y el comienzo y el final del gen de CAT están en negrita y cursiva. Se digirió el fragmento de PCR de 660 pb PCR mediante *NotI* y *HindIII*, y se purificó (fragmento II). En tercer lugar, se purificó un fragmento de scFv13E6 *NcoI-NotI* de 750 pb (una versión injertada del scFv13R4 que contenía los bucles CDR de un anticuerpo monoclonal anti-E6. Philibert *et al.*, por publicar) (fragmento III). En cuarto lugar, se ligaron los tres fragmentos I, II y III para dar el plásmido pscFvCAT. Finalmente, se introdujo una delección interna de 165 pb en el scFv eliminando el fragmento entre los dos sitios *PstI* del gen. El plásmido resultante, denominado pscFvCAT, es Amp^R y CAM^S puesto que la delección del fragmento de *PstI* dio como resultado un desplazamiento del marco en el scFv.

El plásmido p513-EGFP es un derivado de pSG5 (Green, S., *et al.* Nucleic Acids Res 1988, 16:369) y alberga la región codificante de EGFP (Clontech, Inc.) bajo el control del promotor temprano de SV40. Los constructos p513-scFv-EGFP corresponden a fusiones en marco de la región codificante de scFv y EGFP con un ligador de 10 residuos.

- 30 Se amplificaron las regiones codificantes de scFv con los cebadores oligonucleotídicos

5'-ACTGATAAGCTTGCCACCATGGCCGAGGTGC (SEQ ID NO: 30)

y 5'-TTGATTACTAGTGAGTTTTTGTCTGCGGCC (SEQ ID NO: 31)

y se insertaron en el vector p51 3-EGFP digerido con *HindIII-SpeI*.

Entramado de anticuerpo optimizado

- 35 Para maximizar la eficacia de la biblioteca de scFv, la construcción de la biblioteca comenzó con la selección de un entramado de anticuerpo optimizado individual para la selección de intracuerpos. A través de evolución molecular, se obtuvo un scFv humano denominado scFv13R4, que se expresa a altos niveles en el citoplasma de *E. coli*. Este scFv también se expresa y tiene una conformación activa y soluble en células tanto de levadura como de mamífero. Este scFv es muy estable *in vitro* y puede renaturalizarse en presencia de un agente reductor. Además, el análisis de su cinética de plegamiento mostró que se pliega más rápido que el scFv original y se agrega más lentamente *in vitro*. Las mutaciones aisladas se ubican principalmente en el dominio VH y parecen ser altamente específicas para este scFv particular puesto que no pueden transferirse a un dominio VH muy homólogo. Las secuencias de nucleótidos y aminoácidos de scFv13R4 se muestran a continuación.

Secuencia de nucleótidos y aminoácidos de scFv13R4 (de los nucleótidos 1 a 819)

```

-----|-----|-----|-----|-----|-----|
1 atggccgaggtgcagctggtggagctctgggggaagcctggtcaagcctggggggtccctg 60
1 M A E V Q L V E S G G S L V K P G G S L 20

-----|-----|-----|-----|-----|-----|
61 agactctcctgtgcagcctctggattcaccttcagtaactatagcatgaactgggtccgc 120
21 R L S C A A S G F T F S N Y S M N W V R 40

-----|-----|-----|-----|-----|-----|
121 caggctccaggaaggggctggagtgatctcatccattagtggtagtagtagatacata 180
41 Q A P G K G L E W I S S I S G S S R Y I 60

-----|-----|-----|-----|-----|-----|
181 tactacgcagacttcgtgaagggccgattcaccatctccagagacaacgccacgaactca 240
61 Y Y A D F V K G R F T I S R D N A T N S 80

-----|-----|-----|-----|-----|-----|
241 ctgtacctgcaaatgaacagcctgagagccgaggacacggctgtttattactgtgtgaga 300
81 L Y L Q M N S L R A E D T A V Y Y C V R 100

-----|-----|-----|-----|-----|-----|
301 tccagtattacgatttttgggtggcggtatggacgtctggggcagaggcaccctggtcacc 360
101 S S I T I F G G G M D V W G R G T L V T 120

-----|-----|-----|-----|-----|-----|
361 gtctcctcaggtggaggcggttcaggcggaggtggcagcggcggtggcggatcgagctct 420
121 V S S G G G G S G G G G S G G G G S Q S 140

-----|-----|-----|-----|-----|-----|
421 gtgctgactcagcctgcctccgtgtctgggtctcctggacagtcgatcaccatctcctgc 480
141 V L T Q P A S V S G S P G Q S I T I S C 160

-----|-----|-----|-----|-----|-----|
481 gctggaaccagcagtgacgttggtggttataactatgtctcctggtaccaacaacacca 540
161 A G T S S D V G G Y N Y V S W Y Q Q H P 180

-----|-----|-----|-----|-----|-----|
541 ggcaaagcccccactcatgatttatgaggacagtaagcggcctcaggggtttctaat 600
181 G K A P K L M I Y E D S K R P S G V S N 200

-----|-----|-----|-----|-----|-----|
601 cgcttctctggctccaagtctggcaacacggcctccctgacaatctctgggctccaggt 660
201 R F S G S K S G N T A S L T I S G L Q A 220

-----|-----|-----|-----|-----|-----|
661 gaggacgaggctgattactgcagctcatatacaaccaggagcactcgagttttcggc 720
221 E D E A D Y Y C S S Y T T R S T R V F G 240

-----|-----|-----|-----|-----|-----|
721 ggagggaccaagctggccgtcctaggtgcgggccgagaacaaaaactcatctcagaagag 780
241 G G T K L A V L G A A A E Q K L I S E E 260

-----|-----|-----|-----|-----|
781 gatctgaatggggccgcacatcaccatcatcaccattaa 819 (SEQ ID NO: 32)
261 D L N G A A H H H H H H * 272 (SEQ ID NO: 33)

```

MÉTODOS

Base de datos de secuencias de CDR3

Se usó la versión 5 (agosto de 1992) de la base de datos Kabat (Johnson, G., y Wu, T.T.: Kabat Database and its applications: 30 years after the first variability plot. *Nucleic Acids Res* 2000, 28:214-218). Este conjunto de datos contenía 44990 secuencias. En primer lugar, se extrajeron 4643 secuencias VH humanas que no eran un pseudogen y no estaban humanizadas. Entonces se extrajeron secuencias H3 de esta base de datos, teniendo en cuenta en primer lugar la secuencia de nucleótidos cuando estaba presente, luego la secuencia de aminoácidos. Finalmente, se mantuvieron las 3469 secuencias H3 completas que contenían sólo los 20 aminoácidos regulares, entre las que 2703 eran únicas. Se siguió el mismo procedimiento para las cadenas ligeras λ y κ , respectivamente, dando como resultado 1044 y 1291 secuencias de las cuales 775 y 828 eran únicas.

También se extrajeron secuencias de CDR3 de la base de datos IMGT/LIGM-DB tal como existía el 27 de noviembre de 2003 (Giudicelli, V., *et al.* IMGT/LIGM-DB, the IMGT(R) comprehensive database of immunoglobulin and T cell receptor nucleotide sequences. *Nucleic Acids Res* 2006, 34:D781-784). Sólo se consideraron los genes "productivos/regulares/humanos/ADNc+ARN/reordenados". Se obtuvieron 5179 secuencias H3, 1432 K3 y 1131 L3, de las cuales 4323 secuencias H3, 974 K3 y 812 L3 eran únicas.

Se recogieron también 127 secuencias de anticuerpos humanos adicionales del banco de datos de proteínas (Berman, H.M., *et al.* The Protein Data Bank. *Nucleic Acids Res* 2000, 28:235-242.). Para esto, se usó el archivo de 510 secuencias ya recopiladas por Andrew Martin el 19 de agosto de 2003 (Allcorn, L.C., y Martin, A.C. R. SACS-self-maintaining database of antibody crystal structure information. *Bioinformatics* 2002, 18:175-181).

20 Diseño de oligonucleótidos escogidos

Al sesgar las representaciones de los aminoácidos, se calcularon mezclas optimizadas de los nucleótidos en cada una de las tres posiciones de codones tal como se describió anteriormente (Wang, W., y Saven, J.G. Designing gene libraries from protein profiles for combinatorial protein experiments. *Nucleic Acids Res* 2002, 30:e120; Park, S., *et al.* Progress in the development and application of computational methods for probabilistic protein design. *Comput Chem Eng* 2005, 29: 407-421.). Se limitó la terminación prematura de las secuencias de proteínas imponiendo un límite superior de 0,05 en la probabilidad de producir un codón de terminación. Para las 34 posiciones que no recuperaron satisfactoriamente las probabilidades deseadas de los aminoácidos, se realizó una segunda optimización con el mismo método, pero sin restricción en la frecuencia de codones de terminación. Para la síntesis de oligonucleótidos, se redondearon las frecuencias calculadas en incrementos del 5% tal como sigue: todas las frecuencias entre el 0% y el 5% se redondearon al 5%; otras frecuencias se redondearon al 5% más cercano; si la suma resultante era superior al 100%, se eliminó un 5% de la frecuencia de aminoácidos redondeada mayor del 5% para la que la diferencia entre la frecuencia redondeada y la diana era máxima y se repitió el procedimiento hasta que la suma era del 100%; si la suma era inferior al 100%, se añadió un 5% a la frecuencia redondeada inferior al 95%, para la que la diferencia entre la frecuencia redondeada y la diana era máxima y se repitió el procedimiento hasta que la suma era del 100%.

Construcción de bibliotecas de VH y VL

Se introdujeron secuencias de CDR3 variables en scFv13R4 mediante ensamblaje por PCR usando una polimerasa de corrección de lectura de arranque en caliente (ProofStart, Qiagen) usando como plásmido molde pAB 1-scFv13R4p (Martineau, P., y Betton, J.M. in vitro folding and thermodynamic stability of an antibody fragment selected in vivo for high expression levels in *Escherichia coli* cytoplasm. *J Mol Biol* 1999, 292:921-929). Para introducir bucles H3 al azar, se obtuvo el extremo 5' del gen con la secuencia H3 al azar con los oligonucleótidos M13rev-49 y uno de los 13 oligonucleótidos degenerados, y se obtuvo el extremo 3' con PliaisonH3 y M13uni-43 (ambos durante 20 ciclos a 55°C). Se ensamblaron por tanto las dos bandas purificadas mediante PCR (30 ciclos, 55°C) usando M13rev-49 y M 13uni-43. Se purificó la PCR resultante usando un kit comercial (Nucleospin, Macherey-Nagel), se digirió durante 16 horas a 37°C con las enzimas *NcoI* y *NotI*, y entonces se purificó en un gel. Se siguió el mismo procedimiento para introducir bucles L3 y K3 al azar excepto porque los pares de cebadores usados fueron M13rev-49/PliaisonL3 para el extremo 5' y uno de los oligonucleótidos degenerados que codificaba para el bucle L3/K3 (K3_n o L3_n) con M13uni-43 para la parte 3' del gen.

Se ligó cada banda digerida durante 16 horas a 16°C con 1 μ g de *NcoI-NotI*, se digirió y se purificó pscFv Δ CAT en 100 μ l usando 10 unidades Weiss de ADN ligasa de T4. Se inactivó con calor el ligamiento y se purificó usando un kit comercial (Nucleospin). Entonces se electroporó el ligamiento en 300 μ l de células competentes MC 1061 (Sidhu, S.S., *et al.* Phage display for selection of novel binding peptides. *Methods Enzymol* 2000, 328:333-363) y se sembraron en placa en una placa cuadrada de 600 cm² de LB que contenía 100 μ g/ml de ampicilina, entonces se incubaron durante 16 horas a 37°C. Se rasparon las 18 bibliotecas (13 VH y 5 VL) en 10 ml de LB con glicerol al 10%, e inmediatamente se sembraron en placa 10⁹ bacterias en una placa cuadrada de 600 cm² de LB que contenía

100 µg/ml de ampicilina, IPTG 1 mM y 30 µg/ml de CAM, entonces se incubaron durante 16 horas a 37°C. Entonces se rasparon las 18 bibliotecas en 10 ml de LB con glicerol al 10% y se congelaron a -80°C. Se usó una alícuota para preparar ADN para el ensamblaje de bibliotecas.

Ensamblaje de bibliotecas

- 5 Se amplificaron las 13 bibliotecas de VH usando los cebadores M13rev-49/PliaisonH3.back usando Pfu polimerasa, y se amplificaron las 5 bibliotecas de VL usando scFvCAT.rev/H3Liaison (30 ciclos, 55°C). Se purificaron en primer lugar 18 bandas de PCR, entonces se cuantificaron cuidadosamente en gel usando el software ImageJ. Se agruparon las 13 bandas de VH en cantidades proporcionales a su frecuencia en H3 humano. Esta mezcla se denominó VHpool. Se agruparon las 2 κ de VL con el fin de obtener un 75% de bucles de 9 aminoácidos de longitud y un 25% de bucles de 10 aminoácidos de longitud. Se agruparon las bandas λ de VL para obtener un 30% de los bucles de 9 aminoácidos de longitud, un 30% de los bucles de 10 y un 40% de los bucles de 11. Finalmente, se agruparon las mezclas de κ y λ con el fin de conseguir un 50% de cada clase en la mezcla final denominada VLpool.

- 15 Se ensamblaron VHpool y VLpool mediante PCR usando *Taq* ADN polimerasa y los cebadores M13rev-49/scFvCAT2.rev en 500 µl (30 ciclos, 55°C). Se digirió sucesivamente la PCR con 20 unidades de *Nco*I y *Not*I durante al menos 6 h cada una, se purificó y entonces se cuantificó en gel. Se digirieron sucesivamente 50 µg de vector pCANTAB6 con 80 unidades de *Nco*I y *Not*I durante al menos 6 h cada una, se purificó y entonces se cuantificó en gel. Se ligaron 5 µg de pCANTAB6 linealizado con una cantidad molar igual de inserto (0,84 µg) en 500 µl a 16°C usando 50 unidades Weiss de ADN ligasa de T4. Se inactivó con calor el ligamiento y se purificó usando un kit comercial (Nucleospin). Entonces se electroporó el ligamiento purificado en 10 X 300 µl de células competentes CMax5αF⁺ (Sidhu, S.S., *et al.* Phage display for selection of novel binding peptides. *Methods Enzymol* 2000, 328:333-363), y se sembraron en placa en diez placas cuadradas de 600 cm² de LB que contenían un 1% de glucosa y 100 µg/ml de ampicilina. Tras la incubación durante 16 h a 37°C, se rasparon las células en 2xYT que contenía un 10% de glicerol y se mantuvieron congeladas a -80°C en alícuotas correspondientes a veinte veces la diversidad.

25 Antígenos

- Aurora-A es una proteína etiquetada con His y se produjo en *E. coli*. Se expresó GST:Syk en *E. coli*. (Dauvillier, S., *et al.* Intracellular single-chain variable fragments directed to the Src homology 2 domains of Syk partially inhibit Fc epsilon RI signaling in the RBL-2H3 cell line. *J Immunol* 2002, 169:2274-2283). Se expresó la proteína E6 del virus del papiloma HPV16 en la cianobacteria *Anabaena* (Desplancq *et al.*, por publicar). Se adquirieron histonas (una mezcla de H2a, H2b, H3 y H4) de Sigma (tipo II-AS. N.º H7755). Se purificó tubulina a partir de cerebro de cerdo (Williams, R.C.J., y Lee, J.C. Preparation of tubulin from brain. *Methods Enzymol* 1982, 85 (Pt B):376-385).

Rescate y examen de bibliotecas

- 35 Se realizó el rescate de bibliotecas esencialmente tal como se describió anteriormente usando un fago auxiliar sensible a tripsina (Kristensen, P., y Winter, G. Proteolytic selection for protein folding using filamentous bacteriophages. *Fold Des* 1998, 3:321-328). En resumen, se inoculó una alícuota de la biblioteca correspondiente a un exceso de 10 a 20 veces con respecto a la diversidad ($2-3 \times 10^{10}$ bacterias) en 1000 ml de 2xYT que contenía 100 µg/ml de ampicilina y un 1% de glucosa y se hizo crecer con agitación a 37°C hasta que la DO_{600nm} era de 0,7. Se infectaron 200 ml ($\sim 3 \times 10^{10}$ células) con 5×10^{11} fagos auxiliares KM 13 (Kristensen, P., y Winter, G. Proteolytic selection for protein folding using filamentous bacteriophages. *Fold Des* 1998, 3:321-328) y se incubaron sin agitación durante 30 min. a 37°C. Se sedimentaron las células, se resuspendieron en 1000 ml de 2xYT que contenía ampicilina 100 µg/ml y kanamicina 25 µg/ml y se incubaron durante la noche con agitación vigorosa a 30°C. Se precipitó el sobrenadante que contenía fagos dos veces añadiendo 1/5º del volumen de PEG-8000 al 20%, NaCl 2,5 M, y se resuspendió en PBS complementado con un 15% de glicerol. Se almacenaron alícuotas que contenían $10^{11}-10^{12}$ fagos a -80°C.

- 45 Para seleccionar elementos de unión, se recubrió una placa de 96 pocillos Nunc Maxisorp con 100 µl de antígenos purificados. Para la primera ronda, se usó una concentración de antígeno de 10-100 µg/ml. Para rondas posteriores, se usó una concentración de antígeno de 1-10 µg/ml. Se lavó la placa 3 veces con PBS que contenía un 0,1% de Tween20 (PBST) y se saturó durante 2 horas a temperatura ambiente con PBS que contenía leche desnatada al 2% (MBPS). Se añadieron $10^{11}-10^{12}$ fagos por pocillo en MPBS al 2% y se incubó durante 2 horas a temperatura ambiente. Se lavó la placa 20 veces (primera ronda) o 40 veces (2ª y 3ª rondas) con PBST, y luego se lavó 3 veces con PBS. Se eliminó el exceso de PBS, y se eluyeron los fagos añadiendo 100 µl de trietilamina 100 mM durante 10 minutos a temperatura ambiente. Se neutralizó la suspensión de fagos eluida con 50 µl de Tris-HCl 1 M pH 7,4, luego se digirió 15 minutos a temperatura ambiente con tripsina añadiendo 1,5 µl de CaCl₂ 0,1 M y 15 µl de tripsina tratada con TPCK 10 mg/ml (Sigma). Se infectó 1 ml de una cepa Cmax5αF⁺ en crecimiento exponencial a 37°C en 2xYT con 40 µl de fagos tratados con tripsina, se incubaron 30 min. a 37°C sin agitación, entonces se sembraron en

placa en una placa Petri redonda de 15 cm (LB, ampicilina 100 µg/ml, glucosa al 1%). Tras la incubación durante la noche a 37°C, se recuperaron las bacterias de la placa y se usaron para preparar una nueva reserva de fagos usando el fago auxiliar KM13. Se usaron 10^{11} - 10^{12} fagos de esta reserva durante la siguiente ronda de selección.

Selección periplásmica y citoplasmática

- 5 Para la selección periplásmica, se usaron fagos de la ronda 3 para infectar la cepa no supresora HB2151. Se sometieron a prueba clones individuales para detectar la expresión de scFv mediante ELISA en placas de microtitulación de 96 pocillos recubiertas con antígeno tal como se describe (Harrison, J.L., *et al.* Screening of phage antibody libraries. *Methods Enzymol* 1996, 267:83-109). Para la selección citoplasmática, se preparó plásmido a partir del conjunto de bacterias de la 2ª o 3ª ronda de selección, se digirió con las enzimas *NcoI* y *NotI* y se clonó la banda de 750 pb en el plásmido pET23NN desfosforilado y digerido con *NcoI-NotI*. Se transformó el ligamiento en C-Max5αF^r, y se sembraron en placa las células LB con ampicilina 100 µg/ml y se incubaron durante 16 horas a 37°C. Se rasparon las células, y se preparó ADN de plásmido y se usó para transformar BL21 (DE3) pLysS químicamente competente. Se hicieron crecer clones individuales en una placa de microtitulación de 96 pocillos que contenía 100 µl de 2xYT, 100 µg/ml de ampicilina con agitación vigorosa a 37°C hasta que la DO_{600nm} alcanzó 0,6. Se añadió IPTG hasta 0,4 mM final y se incubó la placa de microtitulación durante 16 horas a 24-30°C con agitación vigorosa en una atmósfera humidificada. Tras la centrifugación, se resuspendieron las células en 100 µl de Tris-HCl 50 mM pH 7,5, EDTA 5 mM, se congelaron/descongelaron y se incubaron 1 hora en hielo. Se añadió MgCl₂ hasta 10 mM y se digirió el ADN con 10 µg/ml de DNaseI. Se usaron 5-20 µl en un ELISA en una placa de microtitulación de 96 pocillos recubierta con antígeno (Nunc Maxisorp). Se realizó el revelado usando anticuerpo monoclonal 9E10 seguido por un anticuerpo anti-IgG de ratón conjugado con HRP.

Purificación de scFv

Se purificaron scFv clonados en el plásmido pET23NN a partir del citoplasma de BL21(DE3) pLysS y se purificaron en una columna Ni-NTA tal como se describe para el scFv13R4 original (Martineau, P., *et al.* Expression of an antibody fragment at high levels in the bacterial cytoplasm. *J Mol Biol* 1998,280:117-127.).

- 25 Transfección de células e inmunofluorescencia

Se mantuvieron células HeLa en medio de cultivo tisular de Eagle modificado por Dulbecco (DMEM; Invitrogen) complementado con L-glutamina (2 mM), penicilina (100 UI/ml), estreptomycin (25 µg/ml) y suero de ternero feral inactivado con calor al 10% a 37°C en una atmósfera con un 5% de CO₂ humidificada. Se llevó a cabo la transfección transitoria con el reactivo de lípidos TransFectin (Bio-Rad, Hercules, CA, EE.UU.) según las instrucciones del fabricante. Se sembraron las células sobre cubreobjetos en placas de 6 pocillos a $2,5 \times 10^5$ células/pocillo el día antes de la transfección. Se mezclaron 1 µg de ADN y 2 µl de reactivo diluido en 100 µl de DMEM y se dejaron a temperatura ambiente durante 20 minutos. Se hicieron crecer las células a 37°C durante 24 horas tras la adición de la mezcla. Se visualizaron las proteínas etiquetadas con GFP expresadas tras la fijación de las células transfectadas con paraformaldehído al 4% en PBS durante 45 minutos a temperatura ambiente. Tras lavado extenso con PBS, se secaron las células y se montaron con Fluoromount-G (SouthernBiotech, Birmingham, RU). Se examinaron las células procesadas con un microscopio de fluorescencia Zeiss Axioplan equipado con una cámara Olympus DP50. Se recogieron imágenes con objetivo Zeiss 40X planneofluar y se procesaron usando Adobe Photoshop 5.5. Para la figura 8, se transfectaron HeLa con el clon 5 anti-histonas fusionado con el dsRed-monómero de GFP, se fijaron como anteriormente y se permeabilizaron con Triton X-100 (0,2%, 5 min.). Se reveló la red de microtúbulos con el scFv 2F12C (tabla 3) a 3 µg/ml usando el 9E10 anti-myc y un anticuerpo anti-IgG de ratón con Alexa Fluor 488. Se observaron las células mediante microscopía confocal (X63).

RESULTADOS

Etapas 1: Selección del entramado de anticuerpo

Con el fin de maximizar la eficacia de la biblioteca de scFv, la construcción de la biblioteca comenzó con la selección de un entramado de anticuerpo optimizado individual para la selección de intracuerpos. A través de evolución molecular, se obtuvo un scFv humano denominado scFv13R4, que se expresa a altos niveles en el citoplasma de *E. coli*. Este scFv también se expresa y tiene una conformación soluble y activa en células tanto de levadura como de mamífero. Este scFv es muy estable *in vitro* y puede renaturalizarse en presencia de un agente reductor. Además, el análisis de su cinética de plegamiento mostró que se pliega más rápido que el scFv original y se agrega más lentamente *in vitro*. Las mutaciones aisladas se ubican principalmente en el dominio VH y parecen ser altamente específicas para este scFv particular puesto que no pueden transferirse a un dominio VH muy homólogo. Se muestran a continuación las secuencias de nucleótidos y aminoácidos de scFv13R4.

Etapas 2: Introducción de diversidad en secuencias de CDR3

a) Base de datos de secuencias de CDR3 humanas

Se recopilaron secuencias de CDR3 humanas de tres fuentes principales: la base de datos Kabat (Johnson, G. y Wu, T.T. Kabat Database and its applications: 30 years after the first variability plot. *Nucleic Acids Res* 2000, 28:214-218), la base de datos IMGT database (Giudicelli, V., *et al.* IMGT/LIGM-DB, the IMGT(R) comprehensive database of immunoglobulin and T cell receptor nucleotide sequences. *Nucleic Acids Res* 2006, 34:D781-784) y el RCSB PDB (Berman, H.M., *et al.* The Protein Data Bank. *Nucleic Acids Res* 2000, 28:235-242). Tras eliminar los duplicados, la base de datos contenía 5179 secuencias únicas de CDR3 H3, 1432 de K3 y 1131 de L3. Puede observarse que la mayoría de las secuencias H3 eran únicas puesto que, por ejemplo, en la base de datos Kabat, 2368 secuencias H3 (88%) se encontraron sólo una vez entre las 2703 secuencias H3 completas. El resultado era comparable en el caso de L3 y K3 puesto que, respectivamente, el 87% y el 82% de las secuencias eran únicas en la base de datos Kabat. Esto subraya la muy alta variabilidad de las secuencias de CDR3 humanas.

Esta variabilidad, sin embargo, no se distribuye uniformemente en el bucle, y la frecuencia de cada aminoácido varía de una posición a otra y para cada longitud de bucle. Además, la distribución de aminoácidos depende del origen de la secuencia del anticuerpo. Este sesgo puede deberse a una restricción estructural como por ejemplo en el caso del antepenúltimo residuo que es frecuentemente un aspartato (D101 usando el esquema de numeración de Kabat) y desempeña un papel importante en el cambio entre la conformación extendida y doblada de la H3. En otros casos, este sesgo puede deberse sólo al número limitado de secuencias disponibles para los segmentos D y J, y pueden tolerarse aminoácidos distintos de los encontrados en anticuerpos naturales.

Para la construcción de la biblioteca, se decidió preparar CDR3 que imitan la distribución natural por dos motivos principales: i) un objetivo era generar scFv que fueran tan humanos como fuese posible para un posible uso en terapia de seres humanos; y ii) el mantenimiento de la distribución de aminoácidos natural dará como resultado más probablemente anticuerpos funcionales.

Se alinearon secuencias de CDR3 de la base de datos por longitud, y se calcularon las frecuencias para cada uno de los 20 posibles aminoácidos en cada posición y para cada longitud de bucle. En el caso de las CDR3 de cadena ligera, se analizaron las secuencias independientemente para cada clase. En el caso de las secuencias H3, esto dio como resultado 35 tablas, una para cada longitud de H3 entre 1 y 34 aminoácidos. Para un bucle de longitud n, esta tabla contenía 20n frecuencias.

b) Diseño de oligonucleótidos para codificar para bucles CDR3

Se diseñaron dieciocho oligonucleótidos para seguir la distribución de aminoácidos encontrada en la base de datos de CDR3 recopilada. Se usaron ciento noventa y dos mezclas optimizadas de los cuatro nucleótidos en cada posición del codón para coincidir tanto como fuese posible con la distribución de aminoácidos deseada. La principal ventaja de este enfoque es que sólo requiere síntesis de oligonucleótidos clásica, lo que da como resultado una mejor calidad de los oligonucleótidos. Debido a las restricciones del código genético, sin embargo, no es posible seguir de manera precisa una distribución de aminoácidos arbitraria. Además, si la biblioteca no sigue estrictamente la distribución de aminoácidos natural, esto puede introducir diversidad no natural interesante en los bucles CDR3.

Se calcularon en primer lugar mezclas optimizadas para las 249 posiciones variables para hacer coincidir la distribución con una frecuencia mínima de codones de terminación, entonces se optimizaron adicionalmente 34 posiciones que estaban demasiado distantes de la distribución diana relajando esta última restricción. Debido a la restricción del código genético, algunas posiciones no estaban perfectamente optimizadas. Por ejemplo, en la posición 3, alanina y glicina estaban subrepresentadas en la mezcla y fue necesario introducir una cantidad sustancial de algunos aminoácidos no encontrados de manera natural como cisteína con el fin de coincidir con otras frecuencias de aminoácidos. Sin embargo, hubo un buen acuerdo global entre la base de datos y las frecuencias codificadas en oligonucleótidos puesto que los aminoácidos más frecuentemente encontrados estaban representados a las tasas más altas en la biblioteca y los aminoácidos poco comunes estaban presentes habitualmente a una frecuencia baja. Las secuencias de los 18 oligonucleótidos degenerados usados para construir los bucles CDR3 se proporcionaron anteriormente.

c) Construcción de bibliotecas de bucles CDR3 para cadenas VH y VL

Se construyeron bibliotecas independientes para cada longitud de bucle CDR3. Se realizó esto independientemente para cada una de la cadena pesada y las dos clases de cadenas ligeras. Para cada biblioteca, se introdujeron bucles CDR3 al azar mediante PCR y la biblioteca resultante se clonó entonces de nuevo en el gen 13R4 de scFv, que se fusionó con el gen de CAT en el vector pscFvCAT. Esto dio como resultado bibliotecas de clones de scFv13R4 con uno y solo un bucle CDR3 aleatorizado.

Se construyó en primer lugar la biblioteca de bucles H3 de 5 aminoácidos de longitud. Se secuenciaron cuarenta y tres clones elegidos aleatoriamente. Algunas posiciones divergían de los valores esperados de frecuencias para los

20 aminoácidos aunque, en promedio, la distribución de aminoácidos en la biblioteca coincidía con la distribución esperada. Esto mostró que la calidad de los oligonucleótidos era buena y que la biblioteca resultante seguía la distribución natural de los aminoácidos en bucles H3 humanos.

Tabla 2

Diversidad de las bibliotecas de CDR3

	Diversidad inicial ^a	Fenotipo CAM ^b			Diversidad final ^c
		++	+	-	
H3-5	1,4e8	15/20	3/20	2/20	1,1e8
H3-6	2,7e7	20/20	0/20	0/20	2,7e7
H3-7	9,2e7	16/20	3/20	1/20	7,4e7
H3-8	5,0e6	14/20	4/20	2/20	3,5e6
H3-9	2,4e8	16/20	1/20	3/20	1,9e8
H3-10	1,0e7	10/20	6/20	4/20	5,0e6
H3-11	2,2e7	9/12	1/12	2/12	1,7e7
H3-12	1,0e7	15/20	2/20	3/20	7,5e6
H3-13	2,3e7	8/12	3/12	1/12	1,5e7
H3-14	1,0e7	12/20	4/20	4/20	6,0e6
H3-15	2,1e7	8/12	3/12	1/12	1,4e7
H3-16	1,0e7	12/20	2/20	6/20	6,0e6
H3-17	1,1e7	12/20	4/20	4/20	6,6e6
K3-9	3,6e6	11/16	1/16	4/16	2,5e6
K3-10	4,4e6	10/16	1/16	5/16	2,8e6
L3-9	1,7e7	11/20	6/20	3/20	9,4e6
L3-10	1,5e7	10/20	9/20	1/20	7,5e6
L3-11	1,8e7	15/20	3/20	2/20	1,4e7

(continuación)

Diversidad de las bibliotecas de CDR3

Fenotipo CAM^b

Diversidad inicial^a ++ + - Diversidad final^c

^a Diversidad inicial de la biblioteca clonada como fusión con CAT y seleccionada en ampicilina. Este es el número de clones obtenidos tras la transformación.

^b Se comprobaron entre 12 y 20 clones de la transformación en placas de CAM. Se incubaron las placas durante 16 h a 37°C y se estimó el tamaño de colonia. Las columnas ++, + y – proporcionan, respectivamente, la fracción de clones que crecieron de manera normal, las que dieron colonias minúsculas y las que no crecieron en absoluto.

^c Diversidad de las bibliotecas seleccionadas en CAM. Se estima la diversidad a partir de la columna “Diversidad inicial” y “Fenotipo CAM” suponiendo que la diversidad final es próxima a (Diversidad inicial) X (Fenotipo CAM ++). La diversidad real puede ser superior puesto que algunos de los clones indicados como + en la columna “Fenotipo CAM” pueden estar presentes en baja abundancia.

Etapa 3: Eliminación de secuencias no expresadas

5 Tal como se esperaba, no todos los clones formados a partir de la construcción de las bibliotecas de VH y VL eran funcionales por tres motivos principales: i) los oligonucleótidos usados para introducir diversidad pueden contener codones de terminación; ii) pueden introducirse codones de terminación o desplazamientos del marco mediante la PCR y las etapas de clonación; o iii) los scFv apenas se expresan en el citoplasma. Para eliminar estos clones de scFv no funcionales, se seleccionaron clones expresados fusionando el gen de scFv y la enzima acetil transferasa
10 (CAT) usando el método de Maxwell KL *et al.* (A simple in vivo assay for increased protein solubility, Protein Sci 1999; 8:1908-1911) tal como se describe a continuación.

Creación de proteínas de fusión scFV-CAT

15 En resumen, se clonaron independientemente las bibliotecas de scFv entre los sitios *NcoI* y *NotI* del vector pscFvCAT bajo el control del promotor tac y en marco con un gen de CAT en el sentido de 3'. Se expresó por tanto la proteína de fusión scFv-CAT en el citoplasma. Si un scFv no se expresaba apropiadamente debido a la inclusión de un codón de terminación o una mutación de desplazamiento del marco, o si no podía plegarse en el citoplasma, la proteína scFv-CAT o bien no se expresaría o bien no sería activa, dando como resultado un fenotipo sensible a cloranfenicol (CAM^S). Por otro lado, si el scFv se expresaba apropiadamente, la proteína scFv-CAT resultante se expresaría bien en el citoplasma, y la bacteria sería resistencia a cloranfenicol (CAM^R). Ajustando la concentración
20 de cloranfenicol (CAM), puede incluso seleccionarse la expresión de diferentes niveles de solubilidad de la proteína scFv-CAT.

25 Se sometieron a prueba diferentes concentraciones de CAM durante esta etapa de selección que oscilaban entre 15 y 200 µg/ml. A la concentración más alta de CAM, la biblioteca se enriqueció en scFv bien expresados, pero también en clones que contenían plásmidos recombinados que albergaban delecciones parciales o completas del gen de scFv. A continuación, se sembraron en placa las bibliotecas en una concentración de CAM media (30 µg/ml). Esta concentración era lo suficientemente alta como para eliminar todos los scFv no expresados o que se agregaban fuertemente, pero no dio como resultado una cantidad detectable de plásmidos que albergaban la delección de scFv.

30 Para estimar el tamaño final de las bibliotecas, se aislaron al menos al menos 12 clones de cada biblioteca inicial, que se seleccionó en ampicilina y en CAM para determinar la fracción de clones CAM^R. Algunos clones crecieron rápidamente y formaron colonias en placas de ampicilina/CAM/IPTG (tabla 2, columna ++), algunos crecieron lentamente (columna +) y algunos no crecieron en absoluto (columna -). Se estimó por tanto el tamaño de las bibliotecas de scFv expresado (seleccionadas en Amp+CAM+IPTG) como el producto del tamaño de la biblioteca original (seleccionada en ampicilina) por la frecuencia de los clones CAM^R. Los tamaños de las 18 bibliotecas se proporcionan en la última columna de la tabla 2 y oscilaban entre 2,5x10⁶ y 1,9x10⁸.

35 Etapa 4: Ensamblaje de la biblioteca seleccionada en CAM

Se ensambló la biblioteca final recombinando las 13 bibliotecas de VH seleccionadas en CAM con las 5 bibliotecas de VL seleccionadas en CAM. La diversidad teórica posible es de aproximadamente 10^{15} ($\sim 13 \text{ VH} \times 10^7 \times 5 \text{ VL} \times 10^6$). Esto es mucho más de lo que podría obtenerse mediante electroporación. Por tanto, es muy improbable obtener dos veces el mismo clon en la biblioteca final.

5 Se amplificaron las 13 bibliotecas de VH y las 5 de VL mediante PCR con una secuencia solapante de 17 nucleótidos, luego se purificaron y cuantificaron en gel de agarosa. Se agruparon entonces los fragmentos purificados de VH y VL en cantidades proporcionales a la distribución natural de las longitudes de bucles CDR3 en anticuerpos humanos (figura 3a). Finalmente, se ensamblaron las mezclas de VH y VL mediante PCR, se digirieron y se clonaron en un vector adecuado para la presentación en fago. Se electroporó la biblioteca en la cepa
10 $C_{\text{max5}\alpha\text{F}}$, dando como resultado una biblioteca de $1,5 \times 10^9$ clones que contenía un inserto de scFv, tal como se comprobó mediante PCR en 100 colonias escogidas aleatoriamente. Se ensamblaron las 18 bibliotecas seleccionadas en CAM en cantidades proporcionales a la distribución natural de longitudes de bucles CDR3 en anticuerpos humanos para formar una biblioteca final de más de un billón de clones.

15 Se secuenciaron ciento dieciocho clones para determinar las secuencias y longitudes de bucles. Casi todas las longitudes de bucles se encontraban en la biblioteca. Bucles de 11 y 16 aminoácidos de longitud estaban también subrepresentados en la biblioteca. Esto se debe presumiblemente a la mala calidad de estos oligonucleótidos tal como se muestra mediante su perfil en un bioanalizador Agilent. Longitudes de bucle que oscilaban entre 7 y 12 estaban sobrerrepresentadas en la biblioteca pero sólo en un factor de dos veces. Las longitudes de bucles de 8 y
20 17 aminoácidos, que son las más frecuentemente encontradas en anticuerpos humanos, estaban todas presentes en la biblioteca. El número de clones secuenciados era demasiado pequeño para analizar la frecuencia de los aminoácidos encontrados en los bucles CDR3. Excepto por algo de contaminación con la secuencia 13R4 de scFv original, no se encontró ninguna secuencia de CDR3 dos veces en la biblioteca.

Expresión de scFv en la célula

25 Debido al uso novedoso de la etapa de selección en CAM, se optimizaron independientemente las bibliotecas de VH y VL para la expresión en la célula. Debido a esta optimización de las bibliotecas de VH y VL, el resultado debe ser sólo proteínas scFv expresadas. Además, puesto que sólo se modifican los bucles CDR3 entre el entramado de anticuerpo scFv13R4 original y las bibliotecas de scFv resultantes, la mayoría de los residuos de la superficie de contacto entre los dos dominios se conservan entre clones. Por tanto, es probable que cualquier VH se ensamble
30 correctamente con cualquier VL y que el nivel de expresión del scFv resultante esté próximo al de ambos clones de las bibliotecas de VH y VL seleccionadas en CAM. En otras palabras, si una VH, con un bucle H3 modificado, se expresa bien cuando se fusiona al VL13R4, también se expresará bien cuando se acopla con una VL con un bucle CDR3 modificado y se selecciona como una fusión con el VH1 3R4. Se sometió a prueba esta hipótesis escogiendo clones al azar de la biblioteca final y expresándolos en el citoplasma de *E. coli* y en el citosol de mamíferos.

35 Se preparó ADN a partir de la biblioteca final y se clonaron los genes de scFv en un plásmido para la expresión citoplasmática bajo el control del promotor de T7 fuerte. Debe indicarse que los muy altos niveles de expresión obtenidos bajo tal promotor fuerte favorecen la agregación con respecto a la expresión soluble debido a la cinética de orden alto del proceso de agregación. La rigurosidad de esta prueba es por tanto alta y podría ser posible aumentar la razón de soluble frente a insoluble usando un promotor más débil. Se sometieron a prueba veinte clones en *E. coli* y 19 de ellos mostraban al menos algo de expresión soluble en el citoplasma (figura 3). Un cuarto de los
40 clones (5/20; clones 3, 10, 11, 16, 19) se expresaban a muy altos niveles puesto que los scFv eran claramente visibles en un gel teñido con Coomassie. Para obtener una visión más global de los niveles de expresión soluble en *E. coli*, se clonó la biblioteca en frente del gen de GFPuv bajo el control del promotor de T7. Si el scFv es soluble y se expresa en el citoplasma, esto debe dar como resultado actividad de proteína fluorescente verde (GFP) que puede monitorizarse directamente en un transiluminador UV. Se sometieron a prueba aproximadamente 1000 clones para detectar la presencia de actividad de GFP detectable y aproximadamente el 60% presentaba un fenotipo GFP⁺,
45 lo que de nuevo indicaba que la mayoría de los clones de scFv de la biblioteca final se expresaban correctamente en el citoplasma de *E. coli*. Estas dos pruebas demostraron que el método novedoso de construcción de una biblioteca de scFv tal como se describió anteriormente era muy satisfactorio en la generación de scFv expresados de manera citoplasmática en *E. coli*.

50 A continuación, se sometió a prueba la expresión de la biblioteca en células de mamífero. Se clonaron quince scFv en un vector de expresión de mamífero como fusiones con el gen de EGFP y bajo el control del promotor temprano de SV40, luego se transfirieron en células HeLa. Se muestran resultados típicos en la figura 4. Tres clones se expresaban a un alto nivel soluble, comparable con el del scFv13R4 original (clon 15), se encontró que 10 scFv eran principalmente solubles aunque todavía estaba presente algo de material agregado en la célula (clones 33 y 36), y 2
55 clones se acumulaban esencialmente como agregados citoplasmáticos (clon 24), tal como se observó con el scFv 1F4 anti-oncoproteína E6 derivado de hibridoma (figura 4). En conclusión, trece de los quince scFv sometidos a prueba se expresaban como proteínas solubles que podían detectarse fácilmente en el citoplasma y en el núcleo de las células transfectadas.

Conjuntamente, estos resultados mostraban que más del 85% de los clones de la biblioteca final expresaban scFv soluble en *E. coli* (16/20) y citoplasma de mamíferos (13/15), mientras que aproximadamente el 20% de ellos expresaban scFv a un muy alto nivel (5/20 en *E. coli* y 3/15 en células eucariotas). Globalmente, la mayoría de los clones se expresaban bien en las condiciones reductoras del citoplasma bacteriano y eucariota. Esto es una gran mejora con respecto a resultados obtenidos anteriormente con bibliotecas de scFv no optimizadas.

Selección de elementos de unión

Tal como se mostró anteriormente, la biblioteca contiene una muy alta proporción de clones expresados. La siguiente etapa fue seleccionar anticuerpos de la biblioteca frente a proteínas particulares. Por tanto, se usó la biblioteca de presentación en fago para seleccionar elementos de unión frente a cinco antígenos diferentes usando proteínas purificadas adsorbidas sobre placas de microtitulación. Se realizaron tres rondas de selección, y se sometieron a ensayo los fagos eluidos mediante ELISA frente a los antígenos inmovilizados. En todos los casos, se obtuvo una señal positiva tras una única ronda de selección. Esta señal aumentó fuertemente tras dos rondas y no aumentó adicionalmente durante la tercera ronda de selección. Este procedimiento de selección muy rápido se debía presumiblemente tanto a la propia biblioteca focalizada, que contiene sólo scFv expresados que dan como resultado un fondo bajo, como al uso de un fago auxiliar sensible a tripsina que disminuyó adicionalmente el nivel de fondo.

Para caracterizar el procedimiento de selección, se seleccionaron 60 clones de las tres rondas de selección frente a la fusión GST:Syk. Se usaron estos clones para preparar fagos monoclonales, que entonces se sometieron a prueba para detectar la unión al antígeno mediante ELISA. La figura 7 muestra la distribución de los valores de ELISA obtenidos para cada ronda de selección. La distribución era normal con una fuerte homogeneidad de la señal en cada ronda de selección puesto que más de la mitad de los clones mostraba una señal de ELISA dentro de 0,1 del valor pico. Durante el procedimiento de selección, la señal pico aumentó desde 0,4-0,5 tras una única ronda hasta 0,9 tras la ronda 2 y 1,0 tras la ronda 3, en buena correlación con los resultados obtenidos con los fagos policlonales. Además, tras una única ronda de selección, casi el 100% de los clones reconocían ya el antígeno. Esto mostró que el uso de la biblioteca optimizada en combinación con un fago auxiliar sensible a tripsina da como resultado una ausencia casi total de fondo durante el procedimiento de selección.

A continuación se sometió a prueba si la biblioteca contenía clones que expresaban scFv soluble en el periplasma. Se infectó la cepa HB2151 no supresora con los fagos eluidos tras la tercera ronda de selección frente a tubulina, GST:Syk y las histonas núcleo. Se prepararon extractos periplásmicos y se sometieron a prueba para detectar la actividad de unión mediante ELISA. En los tres casos, el 12-20% de los clones dio una señal fuerte con valores de absorbancia superiores a 0,5 (10 veces el fondo), y aproximadamente el 30% eran claramente positivos con un valor de absorbancia superior a 0,1 (dos veces el fondo). Estos resultados se comparaban favorablemente con los notificados con otras bibliotecas de scFv, subrayando de nuevo la alta proporción de clones bien expresados presentes en la biblioteca. Además, esto mostró que el enfoque de selección en CAM seleccionaba eficazmente constructos sin codones de terminación presentes en los oligonucleótidos. Esto es de hecho de gran importancia para aislar scFv solubles a partir de bibliotecas de presentación en fago puesto que frecuentemente se seleccionan codones de terminación ámbar en CDR durante la selección de bibliotecas sintéticas y semisintéticas.

En ambas de las caracterizaciones previas, los scFv se expresaban en condiciones oxidantes en el periplasma de *E. coli*, o bien como una fusión scFv-pIII o bien como una proteína soluble. Además, se realizó la selección en fago, de nuevo en condiciones oxidantes. Para someter a prueba si los scFv también se expresaban de hecho en el citoplasma, se subclonó el mismo conjunto de clones (ronda 3) en un vector de expresión citoplasmática bajo el control del promotor de T7 fuerte. Para cada antígeno, se sometieron a prueba noventa y cinco clones mediante ELISA para detectar la unión a su respectivo antígeno. En cada caso, el número de clones positivos era comparable o incluso mejor que en la selección periplásmica. Por ejemplo, en el caso de GST:Syk, el 80% de los clones sometidos a prueba eran positivos tras tres rondas de selección. Esto demostró que la etapa de selección periplásmica no disminuía la proporción de scFv solubles en el citoplasma. Además, cuando se usa la biblioteca optimizada tal como se describió anteriormente, no es necesario seleccionar directamente dentro del citoplasma para evitar introducir un sesgo durante la selección.

Se secuenciaron clones individuales de la 2ª y 3ª ronda de selección frente a tubulina. Se muestran las secuencias en la tabla 3.

Tabla 3

Secuencias de algunos scfv anti-tubulina

Nombre	VH CDR3		VL CDR3		frecuencia ^a	rendimiento ^b	WB ^c	IF ^d
	Secuencia	longitud	Secuencia	longitud				
<i>Ronda 2</i>								
C12C	SSITIFGGGMDV	12	HSREVHRTF	9	1/5	19		
	(SEQ ID NO: 34)		(SEQ ID NO: 35)					
E12C	SGGNTFDY	8	QQYYRKPWT	9	1/5	53		
	(SEQ ID NO: 36)		(SEQ ID NO: 37)					
F1C	GNADGGENWELFDK	14	QLYQNTLWT	9	2/5	52		
	(SEQ ID NO: 38)		(SEQ ID NO: 39)					
H6C	SSITIFGGGMDV	12	QQNWTSPLS	9	1/5	nd		
	(SEQ ID NO: 40)		(SEQ ID NO: 41)					
<i>Ronda 3</i>								
2C1C	RGRDY	5	QQYNTSPFS	9	1/6	8,6	+	-
	(SEQ ID NO: 42)		(SEQ ID NO: 43)					
2E11C	GRNVLNY	7	QQNSSSPRFT	10	2/6	8,7	+	-
	(SEQ ID NO: 44)		(SEQ ID NO: 45)					
2F12C	GRRALGN	7	QQYNTSPFS	9	1/6	45	+	+
	(SEQ ID NO: 46)		(SEQ ID NO: 47)					
2G4C	GRRALGN	7	LTWSMRSAI	9	1/6	15	+	+
	(SEQ ID NO: 48)		(SEQ ID NO: 49)					
2G9C	GRRALGN	7	LTTENSVYRLV	11	1/6	50	+	-
	(SEQ ID NO: 50)		(SEQ ID NO: 51)					

Secuencias de las CDR3 de los mejores clones positivos en un ELISA usando scFv expresados de manera citoplasmática a partir de la 2ª (5 clones) y la 3ª (6 clones) ronda de selección (tabla 3). ^a Frecuencia de aparición del scFv entre clones secuenciados de la misma ronda. ^b Rendimiento: mg de scFv purificado a partir de 1 litro de células hechas crecer en un frasco (DO₆₀₀ = 5). ^c WB: detección de tubulina en extractos de cerebro mediante inmunotransferencia de tipo Western. ^d IF: + significa que el scFv puede revelar la red de microtúbulos mediante inmunofluorescencia (figura 8). Las secuencias de los clones 2C1C, 2E11C, 2F12C, 2G4C y 2G9C se han presentado a la base de datos EMBL y sus números de registro son respectivamente AM886280, AM886281, AM886282, AM886283 and AM886284.

5 En todos los casos, los clones secuenciados eran los que proporcionaban la mejor señal en el ELISA realizado con scFv expresados de manera citoplasmática. La mayoría de los clones eran diferentes puesto que sólo un clon de la 2ª ronda y uno de la 3ª ronda se encontraron dos veces. Esto demuestra que está todavía presente una alta diversidad tras 3 rondas de selección. Se purificaron ocho de los scFv anti-tubulina mediante cromatografía de afinidad a partir del citoplasma. En todos los casos, se purificaron más de 8 mg de scFv a partir de un litro de células

hechas crecer en un frasco ($DO_{600} = 5$), y algunos scFv se expresaban a un nivel por célula comparable al nivel de expresión excepcionalmente alto notificado para un anti-HER2 en el periplasma de *E. coli*.

Funcionalidad de scFv como intracuerpos

5 Para determinar si el scFv aislado podría unirse a su diana *in vivo*, se caracterizaron los scFv anti-histona expresados en células humanas. Se clonó la tercera ronda de selección en el vector p513-EGFP y se transfectaron diez clones elegidos aleatoriamente en células HeLa. Se muestran en la figura 6 resultados típicos de las células que expresan las fusiones scFv-EGFP y observadas mediante microscopía confocal. Un scFv se expresaba como agregados citoplasmáticos. Cuatro scFv se expresaban como proteínas citoplasmáticas solubles, tal como se juzga mediante la tinción homogénea de las células, a un nivel comparable al del scFv13R4. Finalmente, la expresión de tres scFv dio lugar a una tinción más fuerte del núcleo (figura 6, clon 2) y dos scFv se localizaban exclusivamente en el núcleo (figura 6, clones 5 y 10). Puesto que estas proteínas de fusión scFv-EGFP se expresaban en el citoplasma de la célula y no contenían una señal de localización nuclear, esto sugería que podían interactuar *in vivo* con las histonas y por tanto eran activas dentro de la célula. Este análisis mostró que aproximadamente la mitad de los clones presentes tras la tercera ronda de selección frente a histonas núcleo podían unirse a su diana nuclear *in vivo*.
 10 Esto se confirmó *in vitro* mediante inmunotransferencia puntual y de tipo Western con scFv purificado. Además, la secuenciación de estos clones mostró que contenían diferentes regiones CDR3 de cadena pesada y ligera.
 15

Caracterización *in vitro* de scFv anti-tubulina

20 Para demostrar la actividad del scFv anti-tubulina en las condiciones reductoras en el citoplasma celular, se extrajeron los scFv en presencia de un agente reductor y se comparó la señal de ELISA con la obtenida con los scFv extraídos en condiciones oxidantes. Tal como se muestra en la figura 7, los cinco scFv sometidos a prueba dieron la misma señal de ELISA en ambas condiciones, lo que demuestra que los scFv conservan una actividad completa en condiciones reductoras incluso en ausencia de formación de enlaces disulfuro. Los cinco scFv podían reconocer tubulina no plegada mediante inmunotransferencia de tipo Western en extractos de cerebro y la proteína nativa en un ELISA de competición. Se sometió a prueba la capacidad de los cinco scFv para interactuar con microtúbulos en células mediante IF. Los clones 2F12C y 2G4C revelaron la red de microtúbulos en células.
 25

La figura 8 ilustra la utilidad de la biblioteca como fuente para estudios proteómicos tanto *in vitro* como *in vivo*: se transfectaron células HeLa con el clon 5 anti-histonas fusionado con una Red-GFP, y se reveló la red de microtúbulos mediante IF usando el scFv 2F12C.

30 Conjuntamente, estos resultados muestran que la biblioteca descrita en este informe es altamente diversa y funcional y permite un aislamiento rápido y fácil de intracuerpos completamente humanos activos *in vivo*.

REIVINDICACIONES

- 5 1. Biblioteca de anticuerpos que comprende al menos aproximadamente 10^8 clones de anticuerpos scFv únicos, en la que cada clon de anticuerpo scFv único codifica para un anticuerpo scFv único que comprende al menos una de una secuencia VH CDR3 única y una secuencia VL CDR3 única, y en la que los clones de anticuerpos scFv únicos codifican para una secuencia de entramado idéntica a una secuencia de entramado codificada por un clon de anticuerpo scFv13R4 de SEQ ID NO: 32.
2. Biblioteca de anticuerpos según la reivindicación 1, en la que los clones de anticuerpos scFv únicos codifican para anticuerpos scFv que comprenden una secuencia VH CDR3 única.
- 10 3. Biblioteca de anticuerpos según la reivindicación 1, en la que los clones de anticuerpos scFv únicos codifican para anticuerpos scFv que comprenden una secuencia VL CDR3 única.
4. Biblioteca de anticuerpos según la reivindicación 1, en la que los clones de anticuerpos scFv únicos codifican para anticuerpos scFv que comprenden una secuencia VH CDR3 única y una secuencia VL CDR3 única.
5. Método para preparar una biblioteca de anticuerpos scFv enriquecida en clones de anticuerpos scFv que pueden expresarse dentro de una célula, que comprende:
 - 15 a) proporcionar una primera colección de clones de anticuerpos scFv, en el que la primera colección comprende clones que comprenden una secuencia única dentro de un bucle CDR3 de VH, en el que la primera colección se ha enriquecido en clones de anticuerpos scFv que contienen una secuencia VL idéntica a la secuencia VL de un clon de anticuerpo scFv13R4 de SEQ ID IVO: 33;
 - 20 b) proporcionar una segunda colección de clones de anticuerpos scFv, en el que la segunda colección comprende clones que comprenden una secuencia única dentro de un bucle CDR3 de VL, en el que la segunda colección se ha enriquecido en clones de anticuerpos scFv que contienen una secuencia VH idéntica a la secuencia VH de un clon de anticuerpo scFv13R4 de SEQ ID NO: 33;
 - 25 c) unir dominios VH de clones de anticuerpos scFv de la primera colección con dominios VL de clones de anticuerpos scFv de la segunda colección para obtener una tercera colección de clones de anticuerpos scFv, en el que la tercera colección contiene clones de anticuerpos scFv que comprenden una secuencia única dentro del bucle CDR3 de VH y una secuencia única dentro del bucle CDR3 de VL,

preparando de ese modo la biblioteca de anticuerpos scFv enriquecida en clones de anticuerpos scFv que pueden expresarse dentro de una célula.
- 30 6. Método según la reivindicación 5, en el que la primera colección comprende clones de anticuerpos scFv que comprenden secuencias CDR1 y CDR2 idénticas en el dominio VH y en el que la segunda colección comprende clones de anticuerpos scFv que comprenden secuencias CDR1 y CDR2 idénticas en el dominio VL.
7. Biblioteca de anticuerpos producida mediante el método según la reivindicación 5.
8. Método para construir una biblioteca de anticuerpos que comprende:
 - 35 a) seleccionar el entramado de un clon de anticuerpo scFv13R4 de SEQ ID NO: 33 como entramado de anticuerpo scFv;
 - b) introducir diversidad de secuencia en una región CDR3 VH del entramado de anticuerpo scFv para generar una primera biblioteca que comprende clones de anticuerpos scFv que comprenden una región CDR3 VH única;
 - c) introducir diversidad de secuencia en una región CDR3 VL del entramado de anticuerpo scFv para generar una segunda biblioteca que comprende clones de anticuerpos scFv que comprenden una región CDR3 VL única;
 - 40 d) eliminar, de la primera biblioteca, clones que no expresan de manera detectable anticuerpo scFv;
 - e) eliminar, de la segunda biblioteca, clones que no expresan de manera detectable anticuerpo scFv; y
 - f) recombinar las bibliotecas primera y segunda para generar una biblioteca final que comprende clones de anticuerpos scFv que comprenden una región CDR3 VH única y una región CDR3 VL única;

construyendo de ese modo la biblioteca de anticuerpos.

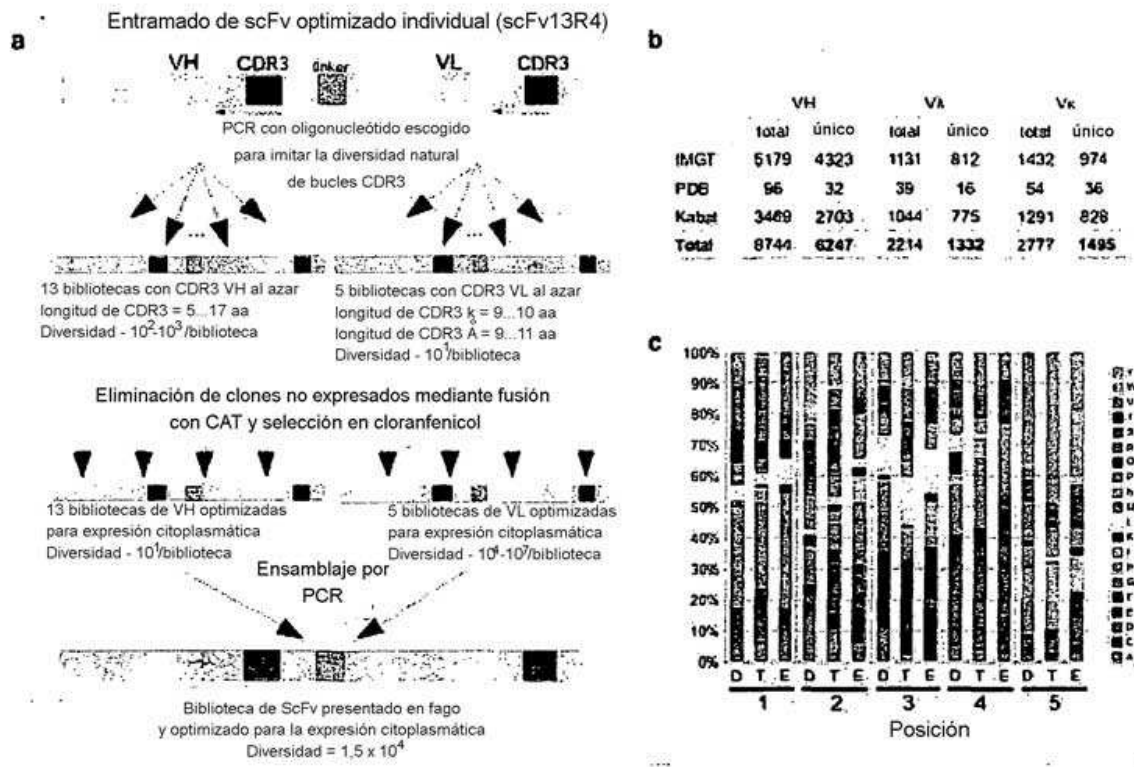


Figura 1

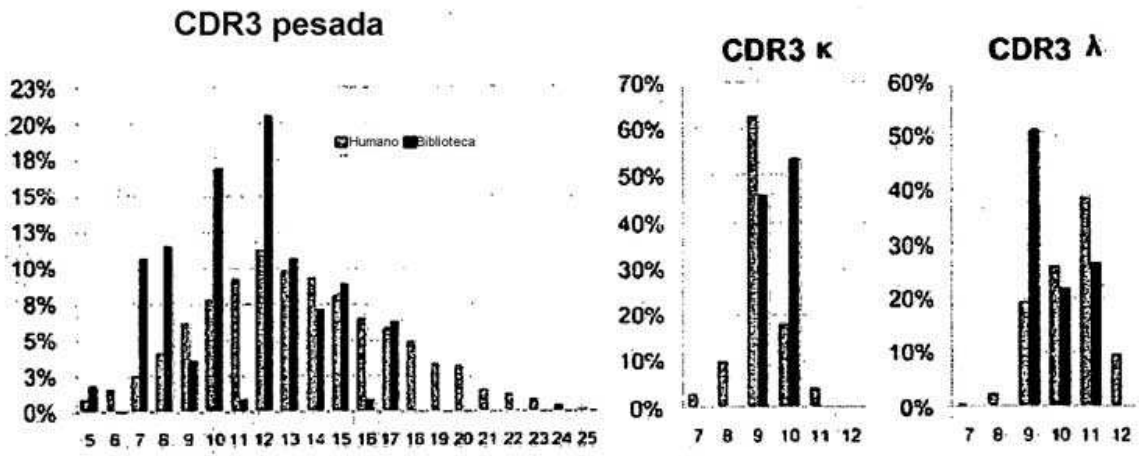


Figura 2

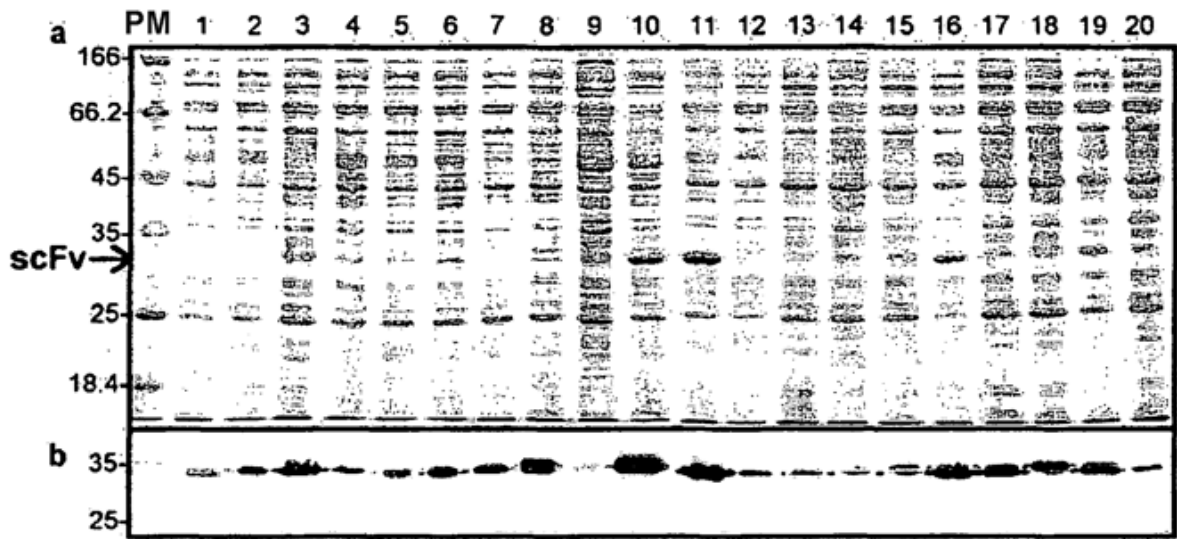


Figura 3

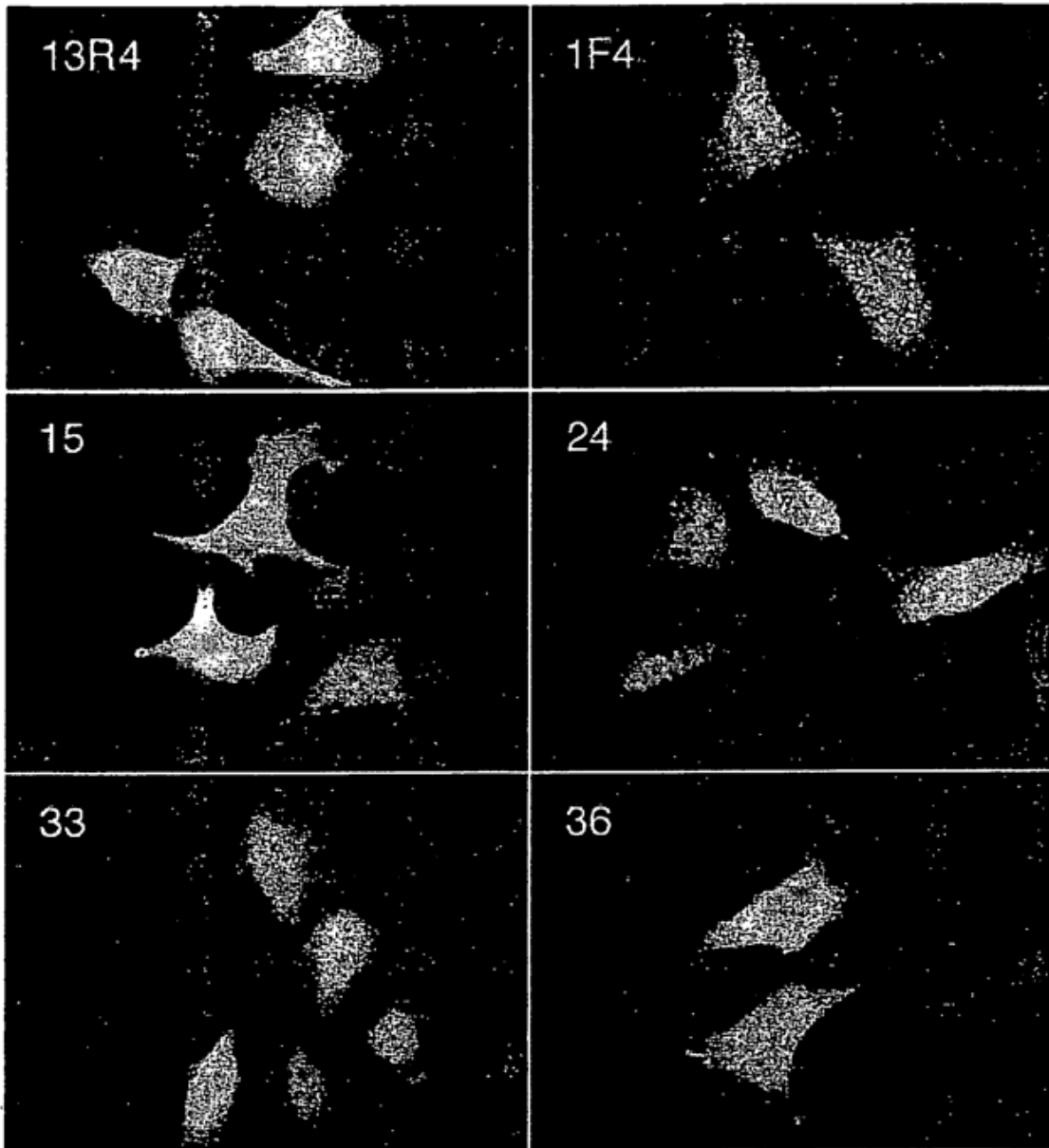


Figura 4

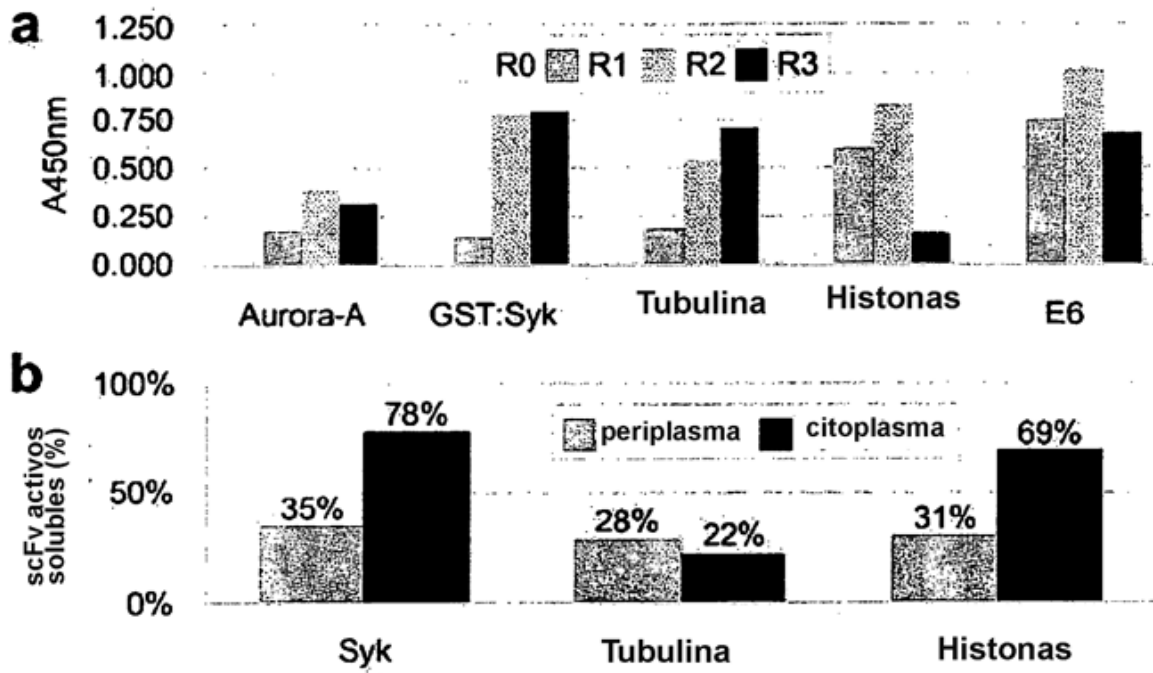


Figura 5

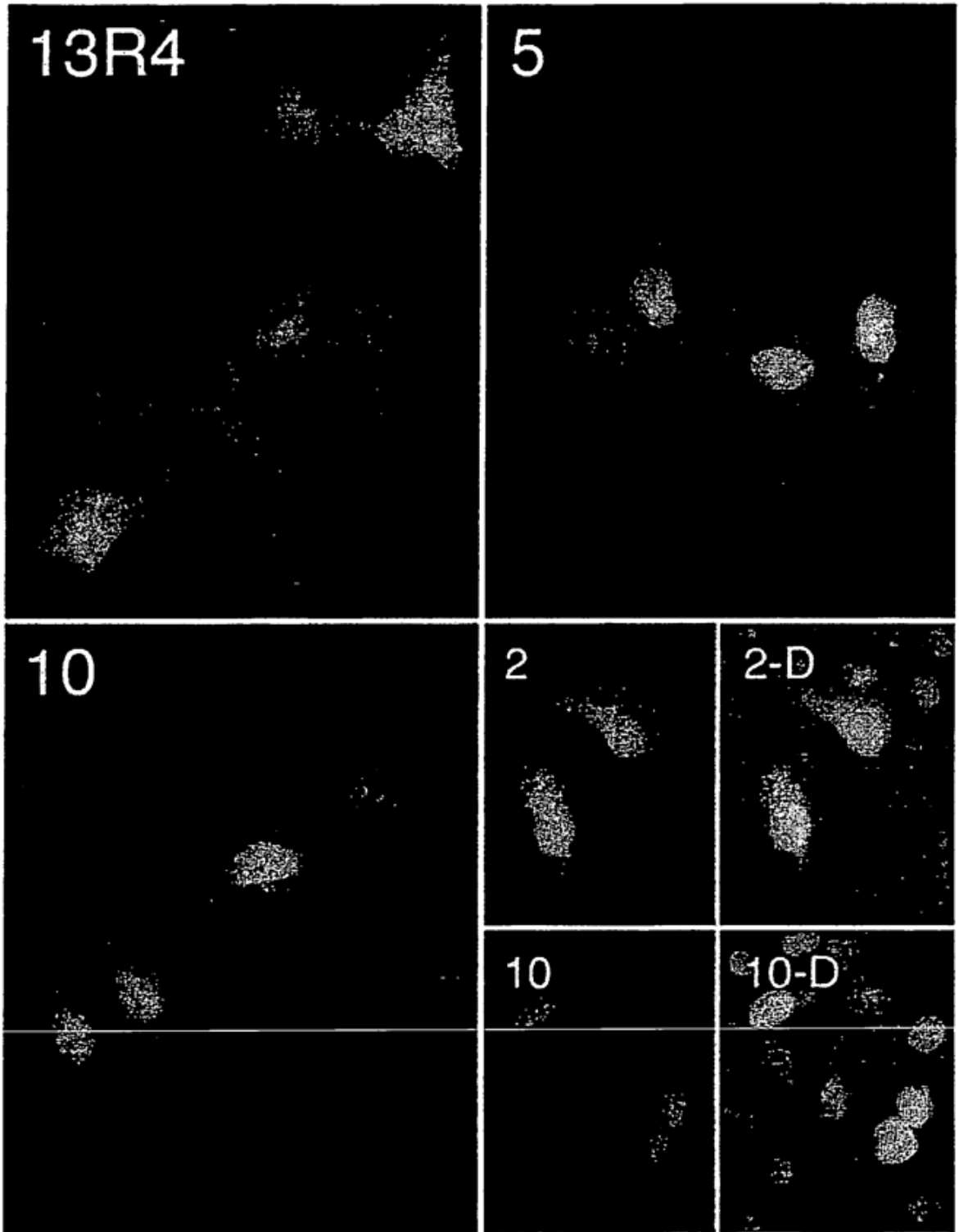


Figura 6

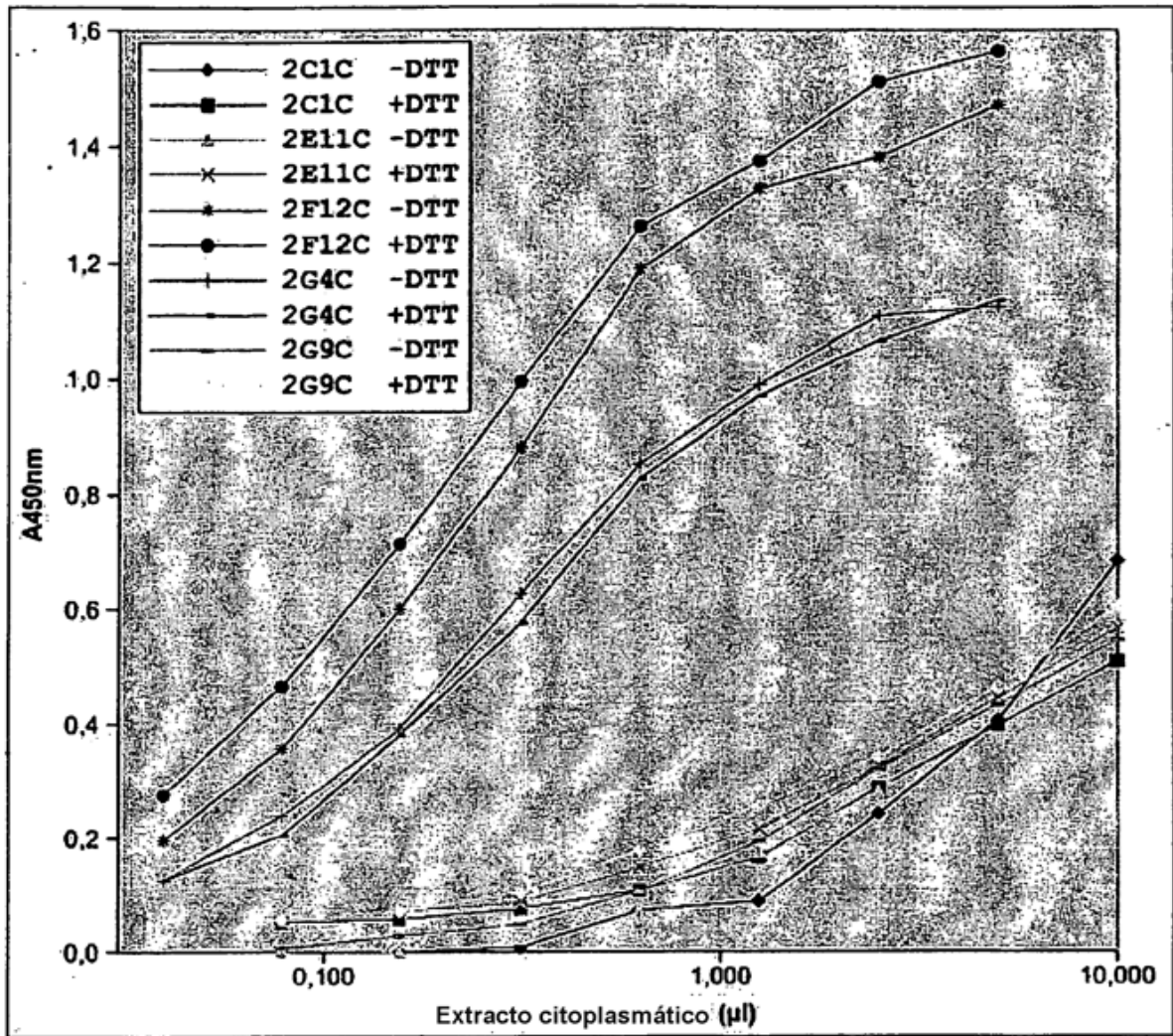


Figura 7

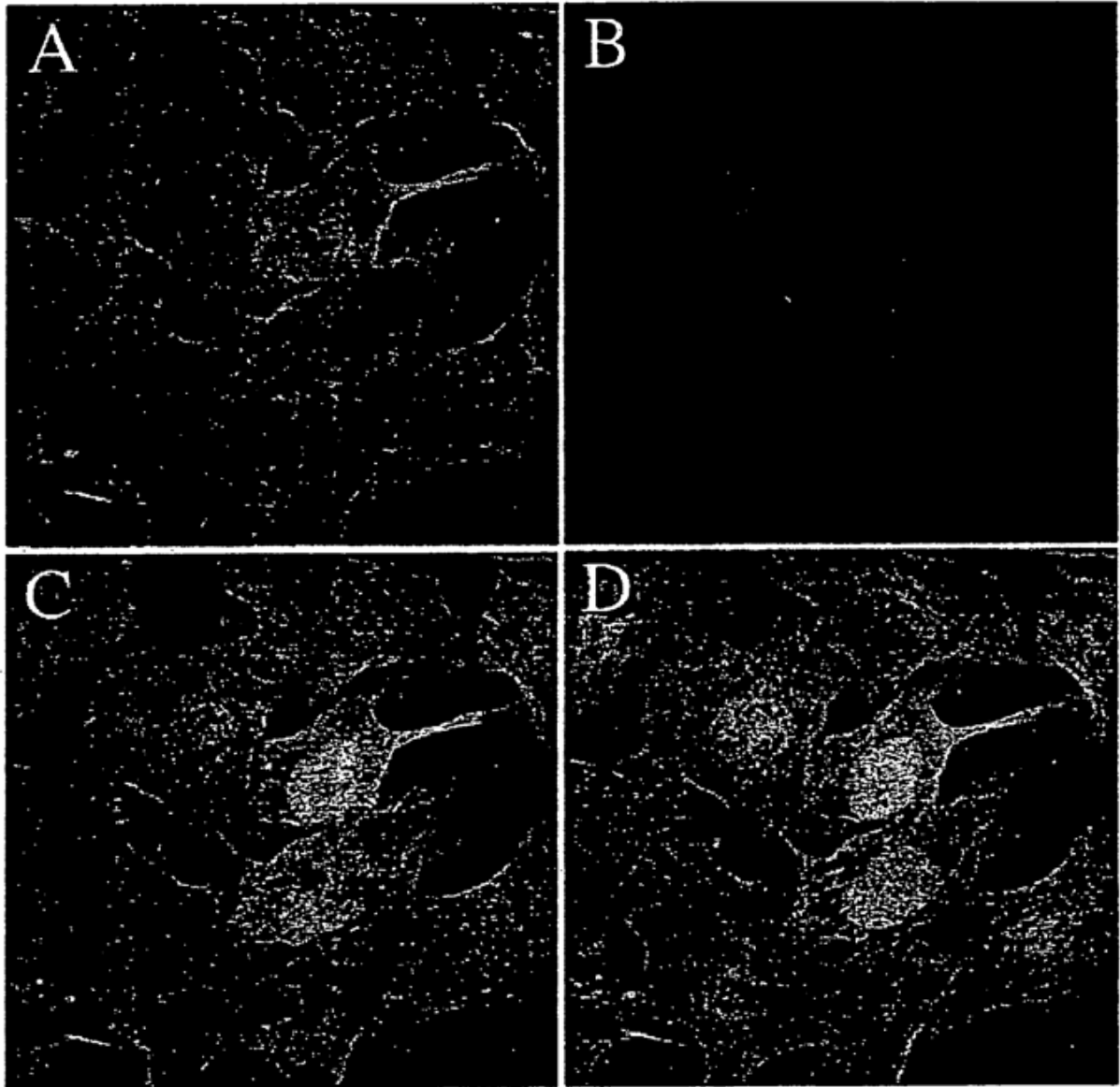


Figura 8