

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 376 454**

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

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **01271932 .4**
96 Fecha de presentación: **21.12.2001**
97 Número de publicación de la solicitud: **1345967**
97 Fecha de publicación de la solicitud: **24.09.2003**

54 Título: **LIBRERÍAS DE EXPRESIÓN EN FAGO DE FRAGMENTOS VH HUMANOS.**

30 Prioridad:
22.12.2000 US 258031 P

45 Fecha de publicación de la mención BOPI:
14.03.2012

45 Fecha de la publicación del folleto de la patente:
14.03.2012

73 Titular/es:
**CAROLE GRAD, LEGAL REPRESENTATIVE OF
HOWARD KAPLAN
18 HILLHOUSE ROAD
WINNIPEG, MANITOBA R2V 2V9, CA;
JAMSHID TANHA;
JOCELYN ENTWISTLE;
SARAN NARANG;
MICHAEL DAN y
COLIN R. MACKENZIE**

72 Inventor/es:
**TAHNA, Jamshid;
ENTWISTLE, Jocelyn;
NARANG, Saran;
DAN, Michael y
MACKENZIE, Colin, R.**

74 Agente/Representante:
Lehmann Novo, Isabel

ES 2 376 454 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

Librerías de expresión en fago de fragmentos V_H humanos

5 Campo de la invención

La presente invención se refiere a bibliotecas combinatorias que incluyen bibliotecas de expresión en fago que expresan fragmentos de unión con un único dominio de cadena pesada, que tienen las características de solubilidad preferidas.

10

Antecedentes de la invención

Los desarrollos en las tecnologías de genomaniplulación de anticuerpos e ingeniería genética han hecho posible la generación de formas de fragmentos de anticuerpos recombinantes las cuales, en muchos casos, son sustitutos funcionales de moléculas de inmunoglobulina intactas más grandes. Los fragmentos de anticuerpos con un solo dominio pesado ("dAb") han sido el tema de varios reportes en la bibliografía de patentes y científica. La bibliografía publica esfuerzos por generar bibliotecas de expresión en fago de dichos fragmentos para la selección de fagos por afinidad frente a un ligando diana.

15

20

La patente de los Estados Unidos Nº 5,702,892 ('892) da a conocer una biblioteca de expresión en fago construida en un vector de expresión derivado de M13, en la cual un fago recombinante de la biblioteca contiene un polinucleótido que codifica una proteína de fusión que comprende una proteína de recubrimiento del fago y un fragmento de unión de cadena pesada de inmunoglobulina. El fragmento de unión a la cadena pesada se extiende desde una posición secuencia arriba de CDR1 hasta una posición secuencia abajo de CDR3. La patente '892 describe que la secuencia de ADN que codifica la región CDR3 y/o la región CDR1 puede ser variada al azar de modo que la población de fagos exprese una serie de dominios de unión de cadena pesada potenciales, para la selección por afinidad frente al ligando diana. La patente de los Estados Unidos Nº 5,759,808 da a conocer una biblioteca de expresión en fago que comprende una población de fagos basada en una variación al azar de una secuencia de ADNc obtenida de linfocitos de camélidos inmunizados previamente con antígenos diana. Los anticuerpos de cadena pesada de camélidos existen naturalmente, en una composición de aproximadamente 45%, como dímeros de cadena pesada. Los anticuerpos de cadena pesada específicos para un antígeno diana pueden ser generados mediante inmunización de un miembro de la especie camélida con el antígeno diana (véase Lauwereys et al. (1998) *The EMBO J.* 17, 3512-3520).

25

30

35

Hamers-Casterman et al. (1993) *Nature* 363, 446-448 informan que los anticuerpos de cadena pesada de los camélidos son naturalmente más hidrófilos en los residuos de aminoácidos de las posiciones 44, 45 y 47 (sistema de numeración de Kabat), en FR2, la cual corresponde a la superficie donde ellos normalmente toman contacto con el dominio VL. Otra característica destacable de un V_H de camélido es que generalmente tiene una CDR3 comparativamente más larga con una alta incidencia de cisteínas y por lo tanto puede formar, a través de cisteínas apareadas en CDR1 y CDR3, bucles expuestos, que son más manejables para la unión dentro de cavidades como el sitio activo de las enzimas y los anticuerpos (Desmyter et al. (1996) *Nat. Struct. Biol.* Vol. 3, Nº 9, p. 803). Sin embargo, se ha cuestionado si anticuerpos de un solo dominio con las afinidades deseadas pueden ser generados con dichas configuraciones en ausencia de una inmunización previa, es decir con una biblioteca naive (no inmune) (Lauwereys et al. (1998) *supra*).

40

45

Davies y Riechmann (*Protein Engineering* vol. 9, Nº 6, pp. 531-537, 1996) describen un procedimiento de camelización en dos pasos de dominios V_H humanos. Los puntos de fusión de los dominios V_H se determinaron en experimentos de termodesnaturalización.

50

La presente invención da a conocer avances en la tecnología relacionada con la creación de bibliotecas que contienen proteínas análogas a la inmunoglobulina que se unen específicamente a ligandos diana por ej. antígenos.

Resumen de la invención

55

La presente invención se define en las reivindicaciones.

60

De acuerdo con un aspecto, la invención apunta a bibliotecas combinatorias que contienen variantes de una molécula de unión al ligando V_H parental, donde dicha molécula de unión al ligando parental comprende un fragmento V_H de inmunoglobulina que contiene al menos las regiones marco (FR) de un dominio V_H de inmunoglobulina y donde dichas variantes contienen al menos las regiones FR de dicho dominio V_H de inmunoglobulina, y difieren de dicha molécula de unión al ligando parental en los residuos de aminoácidos que constituyen parte de al menos una de las CDR de dicha molécula de unión al ligando parental, donde dicho dominio V_H de inmunoglobulina se describe en la SEC. ID Nº: 87 u 88.

65

Además de una biblioteca de moléculas de ácido nucleico, el término biblioteca de expresión se entiende que incluye específicamente un fago, viral, bacteriano u otra biblioteca de expresión en superficie celular, una biblioteca de

expresión en ribosoma o cualquier otro sistema de expresión de ácido nucleico funcional que permita la expresión de productos para ser cribados.

De acuerdo con otro aspecto, la invención apunta a un método para generar una biblioteca de expresión de V_H que tenga una diversidad de secuencias CDR, donde dicho método comprende:

(a) proporcionar vectores de expresión que comprendan una secuencia de ácido nucleico que codifique un V_H , donde dicho V_H se define como la variante definida en la reivindicación 1; (b) introducir mediante mutagénesis una variación en una secuencia de ácido nucleico de CDR; y (c) recuperar una biblioteca de expresión que tenga una diversidad de actividades de unión, donde dicho V_H es un V_H humano.

En este documento también se describe una biblioteca de expresión que expresa polipéptidos que comprenden V_s , donde dicha biblioteca comprende un conjunto de regiones marco que tienen una diversidad de secuencias CDR, donde dicha biblioteca tiene una diversidad de actividades de unión, donde dichas regiones marco contienen uno, dos, tres, cuatro o cinco aminoácidos de un grupo que comprende, 6-A, 23-A, 82a-N, 93-A y 108-Q. También se describe un método para identificar a un polipéptido que contiene un V_H , que se une a un ligando diana, donde dicho método comprende los pasos de cribar miembros polipeptídicos de esta biblioteca de expresión recién mencionada con respecto a su capacidad para unirse al ligando diana; e identificar al menos un miembro polipeptídico que se una al ligando diana.

También se describe un polipéptido que contiene un V_H , donde dicho polipéptido deriva de una biblioteca de acuerdo con cualquiera de los párrafos precedentes.

También se describe una población de variantes de al menos una molécula de unión al ligando V_H (dAb), donde dicha molécula de unión al ligando V_H parental comprende un fragmento de unión V_H de inmunoglobulina que comprende, al menos en gran parte, al menos las regiones marco (FR) del fragmento V_H de inmunoglobulina descrito en la figura 2, donde dichas variantes comprenden al menos en gran parte, las regiones FR del fragmento V_H de inmunoglobulina descrito en la figura 2, que incluye al menos uno de (y preferentemente todos) los residuos de aminoácidos, 6, 23, 82a, 93 y 108, y que difieren de dicha molécula de unión al ligando parental en los residuos de aminoácidos que constituyen al menos parte de al menos una de las CDR de dicha molécula de unión al ligando parental. Dicha población de variantes puede estar constituida por una o más bibliotecas combinatorias de dichas variantes, por ejemplo, micromatrices multigénicas (genochips), bibliotecas de expresión en fago, bibliotecas de expresión en ribosoma, etc.

Se debe comprender que las variantes pueden (aunque no necesariamente) formar parte de otra estructura o molécula, por ejemplo en el caso de la expresión en fago, parte de la proteína de recubrimiento del fago. Concordantemente, el término variante se usa ampliamente para referirse a las variantes de la molécula esencial (una molécula que se une a un ligando) cuando forma parte de otra estructura o moléculas (por ej. como en la expresión en fago o la expresión en ribosoma) o cuando es independiente de cualquiera de dichas combinaciones, por ej. en el caso de micromatrices antigénicas cuyos miembros pueden no estar asociados a moléculas/estructuras de soporte individuales.

De acuerdo con la invención, dicha molécula de unión al ligando V_H parental y dichas variantes comprenden cada uno de los residuos de aminoácidos 6, 23, 82a, 93 y 108, descritos en la figura 2 y preferentemente cada uno de los residuos de aminoácidos 6, 23, 74, 82a, 83, 84, 93 y 108 descritos en la figura 3. Dichas moléculas de unión al ligando V_H parentales y dichas variantes incluyen la totalidad de las regiones FR de la molécula de unión al ligando V_H parental descrita en la figura 2 o 3, opcionalmente sujetas a una o más adiciones, deleciones y/o sustituciones, (preferentemente sustituciones de aminoácidos conservadoras) las cuales, si no mejoran, preferentemente al menos no afectan adversamente las propiedades de solubilidad de dichas variantes. Opcionalmente, dichas moléculas de unión al ligando V_H parentales y dichas variantes incluyen además, la totalidad de las CDR1 y CDR2 descritas en la figura 2 o 3, opcionalmente sujetas a una o más adiciones, deleciones y/o sustituciones (preferentemente sustituciones de aminoácidos conservadoras) las cuales, si no mejoran, preferentemente al menos no afectan adversamente las propiedades de solubilidad de dichas variantes. En una realización particularmente preferida, dicha molécula de unión al ligando V_H parental y dichas variantes, comprenden además, la totalidad de CDR3 de la molécula de unión al ligando V_H parental descrita en la figura 2 con excepción de los aminoácidos 100i a 100n, que están aleatorizados para crear dicha población de variantes sujetas sólo a una o más adiciones, deleciones y/o sustituciones, (preferentemente sustituciones de aminoácidos conservadoras) las cuales, si no mejoran, preferentemente al menos no afectan adversamente las propiedades de solubilidad de dichas variantes.

Los nucleótidos que corresponden a los residuos de aminoácidos 3 a 16 pueden ser optimizados para eliminar un sitio de recombinación putativo.

Para los propósitos de este documento, se debe comprender que un efecto adverso sobre las propiedades de solubilidad se debe evaluar en términos del porcentaje de dímero y formas agregadas superiores, en comparación con el monómero, según se determina por cromatografía de exclusión por tamaño (es decir, las áreas respectivas

bajo los picos que representan el monómero, el dímero y las formas agregadas superiores, según se ilustra en la figura 5).

Además se describe una molécula de unión al ligando o una variante, que se ha identificado que se une a un ligando diana mediante cribado de una biblioteca combinatoria de la invención para una o más moléculas de unión al ligando que reconocen específicamente dicho ligando diana. Se describe dicha cualquier molécula de unión a dicho ligando específica que deriva de dicha biblioteca combinatoria de la invención. En este documento se entiende que dicha molécula de unión al ligando específica se puede obtener directamente de dicha biblioteca o se puede derivar indirectamente, por ejemplo, a través de otra genomanipulación de anticuerpos u otros pasos de modificación (por ej. creación de fragmentos, derivados, una segunda biblioteca, etc.) utilizando una molécula de unión al ligando obtenida directa o indirectamente de dicha biblioteca. También se entiende que la invención excluye moléculas de unión al ligando conocidas. En una realización de este aspecto de la invención el ligando diana es un antígeno cancerígeno.

Las figuras 2, 3 y 4 describen variaciones, descritas más detalladamente a continuación, del fragmento de unión al ligando V_H de inmunoglobulina y/o el constructo de ácido nucleico descrito en la figura 1. La figura 1 describe un fragmento de unión V_H de inmunoglobulina parental tipo silvestre, derivado del anticuerpo monoclonal humano BT32/A6 (de aquí en adelante denominado "A6") parcialmente descrito en la patente de los Estados Unidos N° 5,639,863 (de aquí en adelante denominada patente '863). Según se describe en nuestra solicitud en tramitación PCT/CA00/01027, A6 tiene la solubilidad y otras características preferidas que las tornan buenas para la creación de bibliotecas, incluidas bibliotecas naïve, de diversos tipos de fragmentos de inmunoglobulina humana que incluyen los dAb como se describe en más detalladamente a continuación. Consecuentemente, A6, y en particular, como se describe más detalladamente a continuación, al menos una parte sustancial de las regiones marco (FR) del fragmento V_H de A6 descrito en la figura 2, solo o en combinación con características de sus CDR3, proporcionan un punto de partida útil, en la forma de una molécula de unión al ligando parental, para la aleatorización o la aleatorización parcial de residuos de aminoácidos que tienden a desempeñar un papel predominante en la unión al ligando, a saber las regiones CDR de la cadena pesada y particularmente la CDR3 de la cadena pesada. Como se describe más detalladamente a continuación, los cambios en el ácido nucleico (eliminación del sitio de recombinación) en comparación con A6 tipo silvestre (Figura 1) reflejados en la figura 4 se pueden incorporar en las figuras 2 y 3, respectivamente.

La biblioteca combinatoria de la invención se puede generar mediante expresión en fago. Concordantemente se describe una biblioteca de expresión en fago que expresa una diversidad de variantes diferentes de una molécula de unión al ligando V_H parental, donde dicha molécula de unión al ligando parental comprende un fragmento de unión V_H de inmunoglobulina que tiene, al menos en gran parte, al menos las regiones FR del fragmento V_H de inmunoglobulina descrito en la figura 3 que incluye al menos los residuos de aminoácidos 6, 23, 74, 82a, 83, 84, 93 y 108, y donde dichas variantes son codificadas por secuencias de ácido nucleico que difieren de la secuencia de ácido nucleico que codifica dicha molécula de unión al ligando parental en una subsecuencia (al menos una) que codifica al menos parte de una de las CDR de dicha molécula de unión al ligando parental, preferentemente la CDR3, por lo cual dicha diversidad de variantes comprende al menos, en gran parte, las regiones FR del fragmento V_H de inmunoglobulina descrito en la figura 3, que incluye al menos uno de dichos residuos de aminoácidos y se diferencian, al menos en parte, por variaciones de aminoácidos codificadas por variaciones en dicha subsecuencia.

Además de preservar las regiones FR de A6, la molécula de unión al ligando parental basada en A6 puede comprender (y por consiguiente preservar entre los miembros de la biblioteca), en gran parte (sujeta al menos a la aleatorización parcial de regiones seleccionadas de una de las CDR (preferentemente la CDR3)), para crear diversidad de unión dentro de la biblioteca, una o más de las regiones CDR del fragmento V_H de A6. Al menos la longitud de la CDR3 de V_H tipo silvestre (23 aminoácidos) y preferentemente también los elementos de su composición aminoacídica, pueden ser preservados o al menos parcialmente preservados (aproximadamente 16 a 23 aminoácidos y más particularmente 18 a 23 aminoácidos). Opcionalmente la CDR3 también puede ser extendida en aproximadamente 1 a 10 residuos. La biblioteca puede tener opcionalmente representación de moléculas de unión que tienen CDR3s de diversas longitudes.

Se sabe que una molécula dAb, debido a la eliminación de su cadena ligera asociada, tiende, en la mayoría, sino en todos los casos, a agregarse, en diversos grados debido a la naturaleza "pegajosa" de la interfase VL. Esta pegajosidad se puede atribuir, al menos en parte, a la naturaleza hidrófoba de los residuos de V_H en esa interfase. Esta pegajosidad da como resultado una sustancial formación de dímero y/o multímero que puede reducir, en conjunto, las características de solubilidad de los miembros de la biblioteca. Concordantemente, los residuos de aminoácidos de V_H de A6 en la interfase VL pueden ser sustituidos por residuos que tiendan a minimizar la formación de agregados, por ejemplo, aminoácidos hidrófilos, incluidas una o más de las sustituciones resaltadas reflejadas en la figura 8, en comparación con la figura 1.

Como se describe más detalladamente a continuación, dichas sustituciones pueden no ser fijas en toda la población de la biblioteca, sino que son introducidas mediante aleatorización o aleatorización parcial de diversos residuos de aminoácidos de V_H de A6, incluidos particularmente los residuos de FR, entre los residuos de la interfase. (véase por

ejemplo, Padlan et al "Anatomy of the Antibody Molecule" Molecular Immunology Vol. 31, p 169-217, Tabla 25 para la catalogación y discusión relacionada de estos residuos).

5 Particularmente en el caso de bibliotecas de mayor tamaño, por ejemplo las generadas por expresión en ribosoma, es posible introducir diversidad en una o más regiones FR que contengan residuos de aminoácidos identificados en este documento como importantes para mejorar las propiedades de solubilidad, además de una o más regiones que afectan la especificidad de dAb, por ej. CDR3 y CDR1 de V_H.

10 También se describe que las regiones FR, aparte de, o además de, las modificaciones a la interfase VL (FR2) se pueden modificar mediante al menos aleatorización parcial, por ejemplo, una o ambas de FR1 (uno o más de los residuos 4 a 21) y FR4 (uno o más de los residuos 100a a 113) para mejorar, en conjunto, las características de solubilidad de los miembros de la biblioteca (por ejemplo, sesgando al menos algunos y preferentemente todos de uno o ambos de esos conjuntos de residuos [al menos 70% o más], preferentemente 90% a favor de la constitución aminoacídica parental para lograr un 10% de aleatorización).

15 En el caso de los fragmentos dAb de A6, se encontró que los sucesos de recombinación dentro de la secuencia de ácido nucleico que codifica el fragmento de unión V_H tienden a causar deleciones que producen moléculas más cortas, con posible compromiso de las características de unión. Por lo tanto, se describe que las secuencias de ácido nucleico que promueven dichos sucesos de recombinación (en los sitios de recombinación putativos) pueden ser sustituidas, para oponerse a esta tendencia, preferentemente de manera que no provoquen un cambio de aminoácidos. Esos cambios se pueden incorporar, por ejemplo, en las secuencias de nucleótidos que codifican la molécula de unión al ligando V_H parental descrita en la figura 2 o 3.

20 También se describe una población heterogénea de paquetes genéticos (por ej. fago) que tienen una proteína de superficie externa genéticamente determinada, donde los paquetes genéticos presentan colectivamente una pluralidad de diferentes fragmentos de unión al ligando V_H, preferentemente humanos (es decir con una identidad sustancial, preferentemente de al menos 80% de homología con las regiones marco y otras regiones conservadas humanas), donde cada paquete genético incluye un constructo de ácido nucleico que codifica una proteína de fusión que comprende al menos una porción de la proteína de superficie externa y una variante de al menos un fragmento de unión al ligando parental, soluble, derivado preferentemente de, o que tenga una parte sustancial de, las regiones FR de la secuencia de aminoácidos identificada en una de las figuras 2 o 3 (o una secuencia al menos 80%, preferentemente 85 a 100%, más preferentemente 90 a 100%, homóloga (% de identidad) a ella), que incluya al menos uno y preferentemente todos los residuos de aminoácidos 6, 23, 82a, 93 y 108 descritos en la figura 2 o 3, donde de los fragmentos de unión al ligando V_H variantes se extienden preferentemente desde una posición secuencia arriba de una CDR1 de cadena pesada de inmunoglobulina a una posición secuencia abajo de CDR3 (preferentemente incluyendo sustancialmente la totalidad de FR1 y/o FR4), y donde al menos parte de una CDR, preferentemente la CDR3, es una variante generada aleatoriamente de una CDR de dicho fragmento de unión al ligando V_H parental y donde la proteína de fusión se expresa preferentemente en ausencia de una cadena ligera de inmunoglobulina por lo cual los fragmento de unión al ligando V_H variantes están, en conjunto, mejor adaptados para ser, o son más capaces de ser, expresados como proteínas solubles. También son consideradas por la invención diversas combinaciones de 2, 3 y 4 sustituciones en comparación con A6 tipo silvestre en los residuos de aminoácidos 6, 23, 82a, 93 y 108.

45 También se describe que, sesgando la constitución aminoacídica, preferentemente aminoácido por aminoácido, a favor de la constitución aminoacídica tipo silvestre o parental, incluso porciones de la molécula de unión al ligando parental que son aleatorizadas a favor de generar variabilidad en los fragmentos de unión variantes, pueden ser genomanipuladas para mantener las características de solubilidad favorables del dominio de unión parental. Una porción del constructo que codifica al menos parte de la CDR3 puede ser sesgada o sesgada parcialmente a favor de la constitución aminoacídica parental.

50 Se describe que el fragmento de unión V_H parental puede tener naturalmente una CDR3 larga que se puede manejar para formar bucles expuestos para la unión dentro de cavidades. En una realización muy preferida, el fragmento de unión al ligando V_H parental se construye en un marco humano, o se adapta de, o se puede adaptar a, un marco humano.

55 La región de unión preferida de las variantes (correspondiente a la parte aleatorizada o parcialmente aleatorizada de la CDR3) se puede ubicar en la región carboxi terminal de la CDR3.

60 En resumen, de acuerdo con la invención, la secuencia de aminoácidos identificada en la figura 2, que incluye preferentemente al menos parte de la CDR3, proporciona la constitución aminoacídica preferida de las diversas moléculas de unión al ligando parentales preferidas, de modo que una población de moléculas de unión al ligando de cadena pesada variantes, construidas en este marco de aminoácidos, esté en conjunto, mejor adaptada para ser, o sea más capaz de ser, expresada como proteínas solubles.

65 Generalmente, la importancia de los residuos de aminoácidos 6, 23, 82a, 93 y 108, que se muestran en la figura 2, particularmente cuando se combinan los cinco, es que se pueden usar para aumentar significativamente las

propiedades de solubilidad de un fragmento de unión al ligando V_H parental, preferentemente uno, como A6, que tiene propiedades de solubilidad útiles para comenzar con el, a efectos de producir una biblioteca de variantes de dAb para la selección por afinidad frente a un ligando diana, donde dichas variantes en su conjunto están, mejor adaptadas para ser, o son más capaces de ser, expresadas como proteínas solubles.

- 5 Breve descripción de las figuras
- La invención se describirá ahora con referencia a las figuras, donde:
- 10 La figura 1 es un diagrama de una secuencia que muestra la molécula de unión al ligando V_H de A6 a partir de la cual construimos las moléculas de unión al ligando V_H parentales de acuerdo con la invención.
- La figura 2 es un diagrama de una secuencia que muestra una molécula de unión al ligando V_H parental preferida (A6V_H-L1) de acuerdo con la invención, que incorpora las sustituciones de aminoácidos E6A, S23A, S82aN, 93A y T108Q.
- 15 La figura 3 es un diagrama de una secuencia que muestra la molécula de unión al ligando A6V_H-L2 de acuerdo con la invención, que incorpora 3 sustituciones de aminoácidos más en las posiciones 74, 83 y 84.
- 20 La figura 4 es un diagrama de secuencias que muestra la molécula de unión al ligando A6V_H (codificada por A6 chi (-)) de acuerdo con invención, en la cual los ácido nucleico correspondientes a los aminoácidos 3 a 16 de A6 fueron modificados para eliminar un sitio de recombinación putativo, dejando la constitución aminoacídica de A6 inalterada.
- La figura 5 es un cromatograma de exclusión por tamaño de A6V_H-L1 dAb luego de la purificación IMAC que compara A6V_H-L1 (gráfico A) con A6V_H tipo silvestre (gráfico B) y que muestra una proporción mucho mayor de monómero (representada por el pico más grande). El gráfico C representa una versión camelizada de A6 con sustituciones de los aminoácidos de las posiciones 44, 45, 47 y 94 como se describe en este documento y en la bibliografía.
- 25 La figura 6 es un cromatograma de exclusión por tamaño de Yst9.1-L3-9, que se usó para la unión frente a Yst9.1 scFv, descrita en la figura 7.
- La figura 7 es un sensorgrama que muestra la unión de Yst9.1-L3-9 10 μM a Yst9.1 scFv inmovilizado. Los datos de de superficie control (BSA) fueron sustraídos de los datos de superficie activa (Yst9.1) de este sensorgrama. Las constantes de velocidad se calcularon mediante ajuste separado de los datos de asociación (se muestra el ajuste) y disociación a un modelo de interacción 1:1.
- 35 La figura 8 es un diagrama de secuencias que muestra las sustituciones de aminoácidos para la camelización opcional de acuerdo con una realización de la invención.
- 40 La figura 9 es una representación diagramática del vector, SJFI, utilizado para crear el vector en el cual se puede clonar la biblioteca.
- La figura 10 es un diagrama de una secuencia que describe la estrategia de construcción de la biblioteca A6V_H-L1.
- 45 La figura 11 es un diagrama de una secuencia que describe la estrategia de construcción de la biblioteca A6V_H-L1a.

Descripción detallada

- 50 Se describe una población de paquetes genéticos que tienen una proteína de superficie externa genéticamente determinada que incluye paquetes genéticos que presentan colectivamente numerosas de moléculas de unión al ligando diferentes en asociación con la proteína de superficie externa, donde cada paquete incluye un constructo de ácido nucleico que codifica una proteína de fusión la cual es al menos una porción de la proteína de superficie externa y una variante de al menos una molécula de unión al ligando parental soluble, derivada de, o que tiene la
- 55 secuencia de, aminoácidos identificada en la figura 2 (o una secuencia preferentemente al menos 80% homóloga a sus regiones marco y regiones conservadas), la cual incluye al menos uno y preferentemente todos los aminoácidos 6, 23 82a, 94 y 108 que se muestran en la figura 2 (o sustituciones conservadoras de esos residuos las cuales, si no mejoran, al menos no afectan adversamente las propiedades de solubilidad de dicha molécula de unión al ligando V_H parental), donde al menos parte del constructo, preferentemente que incluya al menos parte de la CDR3 identificada
- 60 la figura 2, codifica, o es sesgada a favor de codificar, la constitución aminoacídica del fragmento de unión al ligando parental de modo que los numerosos dominios de unión al ligando diferentes estén en conjunto, mejor adaptados para ser, o sean más capaces de ser, expresados como proteínas solubles. Las moléculas de unión al ligando V_H variantes se caracterizan preferentemente por una CDR3 que tiene de 16 a 33 aminoácidos.
- 65 El paquete genético replicable puede ser un fago recombinante y la población heterogénea de los paquetes genéticos replicables constituye colectivamente una biblioteca de expresión en fago.

5 La interfase VL puede ser genomanipulada para evitar aminoácidos hidrófobos. En otra realización, la interfase VL es genomanipulada para incluir aminoácidos, que forman interacciones débiles. Se describe que la molécula de unión al ligando parental puede tener una interfase VL de tipo camélido o que al menos uno de los aminoácidos de la interfase VL es aleatorizado o parcialmente aleatorizado en la construcción de la biblioteca.

Los fragmentos de unión V_H potenciales incluyen todas las regiones de FR1 a FR4. Se prevén deleciones parciales dentro de CDR2.

10 CDR3s de una diversidad de longitudes diferentes entre 16 y 33 aminoácidos pueden estar representadas predominantemente entre los fragmentos de unión V_H potenciales. CDR3s de una diversidad de longitudes diferentes, entre 18 y 28 aminoácidos, o entre 20 y 25, o entre 18 y 23 aminoácidos pueden estar representadas predominantemente en la biblioteca. De acuerdo con la invención, el fragmento de unión al ligando V_H parental es el fragmento de unión al ligando V_H parental identificado en la figura 2 que tiene una CDR3 de 23 aminoácidos de longitud.

15 También se prevé la existencia aleatoria, sesgada o fija de características dadas a conocer en la bibliografía de camélidos, por ejemplo cisteínas apareables en CDR1 y CDR3 (opcional) y/o la sustitución de aminoácidos hidrófilos, al menos uno de las posiciones 44, 45 y 47 y preferentemente también las posiciones 93 y 94 (sistema de numeración de Kabat).

20 Una secuencia parcial del anticuerpo BT32/A6 (A6) se da a conocer en la patente de los Estados Unidos N° 5,639,863. La secuencia completa se proporciona en la figura 1.

25 En la figura 1, las regiones CDR están delimitadas. Los números de los residuos de aminoácidos en la figura 1 y en toda la divulgación se refieren al sistema de numeración de Kabat (Kabat et al. 1991, Sequences of Proteins of Immunological Interest, publication N° 91-3242, U.S. Public Health Services, NIH, Bethesda MD) excepto en el listado de secuencias y donde se indique explícitamente o se implique de otra manera. La figura 1 corresponde a la SEC. ID N° 1 (ácido nucleico) SEC. ID N° 2 (aminoácido). La figura 1 delimita y etiqueta las regiones CDR1 (correspondiente a SEC. ID N° 3), CDR2 (SEC. ID N° 4) y CDR3 (SEC. ID N° 5).

30 El marco de A6 proporciona las características de solubilidad preferidas para crear bibliotecas de dAb. La expresión "características de solubilidad preferidas", según se usa en este documento se refiere a al menos una de las varias, a menudo correlacionadas, características que incluyen buen rendimiento, expresión como un producto soluble (por oposición a cuerpos de inclusión) dentro del periplasma del organismo hospedador, por ej. *Escherichia coli*, y una reducida tendencia a dimerizarse y formar otros agregados.

35 El término "V_H" se usa para referirse al dominio de cadena pesada variable y a sus variantes preparadas mediante técnicas de rutina en el área, incluidas sus formas conocidas, formas truncadas y de proteínas de fusión, donde dichas variantes retienen sus características de unión esenciales. Este término se usa de forma intercambiable con las expresiones como fragmento de unión V_H, fragmento de unión al ligando V_H, fragmento dAb y fragmento V_H de inmunoglobulina, a los cuales se les asigna el mismo significado.

40 El término "subsecuencia" se usa para referirse a un subconjunto de nucleótido dentro de una secuencia de ácido nucleico de referencia, que incluye un único nucleótido y nucleótidos consecutivos y no consecutivos dentro de una secuencia de más de dos nucleótidos consecutivos.

45 Los términos "polipéptido", "péptido" y "proteína", a menos que el contexto implique algo diferente, se usan en este documento de manera intercambiable, para referirse a polímeros de residuos de aminoácidos de cualquier longitud.

50 La expresión "biblioteca combinatoria" se usa para referirse a un conjunto de moléculas, que pertenecen típicamente a una clase definida (estrechamente o ampliamente) que comprende una cantidad sustancial de variantes potencialmente útiles, donde las variaciones en la molécula representan un conjunto completo o parcial de permutaciones o combinaciones de al menos algunos elementos constitutivos de una molécula de referencia, que es típicamente un molde o molécula "parental", o simplemente la clase en sí misma. A efectos de clarificar, en el caso de polipéptidos y ácidos nucleicos, los elementos constitutivos son aminoácidos y bases de ácido nucleico, respectivamente.

55 Según se usa en este documento, la frase "en gran parte" se refiere a las variaciones en comparación con una molécula de referencia que no deterioran significativamente la "funcionalidad" de esa molécula. En el caso de la molécula de unión al ligando parental y sus variantes, funcionalidad se refiere principalmente a las características de solubilidad y de unión de la molécula. Dichas variaciones (es decir la molécula de referencia en gran parte) se pueden analizar sistemáticamente para evaluar su impacto. En el caso de las regiones marco, en contraposición con las regiones CDR, debido a la conservación sustancial de los residuos de aminoácidos del marco, una gran parte del marco se referiría preferentemente al menos a 80% de identidad de los residuos de aminoácidos y más preferentemente de 85 a 100% de identidad, y aún más preferentemente al menos a 90% de identidad de los

residuos de aminoácidos. Sin embargo, se debe comprender que cada uno de los porcentajes anteriores puede flexibilizarse para descontar los casos en los que la ausencia de identidad en un residuo dado, se debe a una sustitución de aminoácidos conservadora bien reconocida, o cuando observa una clase particular de funcionalidad, por ej. hidrófila, si la sustitución es con un residuo de la misma clase. En el caso de los residuos de CDR, esos números podrían flexibilizarse aún considerablemente más, e incluye 50 a 100% de homología, y todos los porcentajes incrementales y todos los rangos de porcentajes entre esos valores. La expresión "en gran parte", en referencia a las porciones de las regiones marco y CDR, también prevé la posibilidad de adiciones y deleciones en esas regiones que preferentemente no impacten las características de solubilidad y de unión de la molécula de unión al ligando en cuestión.

La expresión "fragmento de unión al ligando" se usa ampliamente para definir todo o cualquier parte de un anticuerpo que sea capaz de unirse específicamente a cualquier ligando, en el sentido más amplio del término ligando.

Un fragmento de unión al ligando de dominio pesado humano basado en A6 es muy adecuado para el desarrollo de una biblioteca combinatoria (opcionalmente una biblioteca de expresión en fago) que se use para generar fragmentos de unión solubles que sean útiles para el diagnóstico y la terapia en humanos (debido a la limitada respuesta HAMA). Esas bibliotecas de expresión en fago se usan para generar selectivamente sondas moleculares que interactúen específicamente con un ligando, incluidas pero no exclusivamente, moléculas y macromoléculas naturales y sintéticas y se pueden usar *in vitro* (es decir, un diagnóstico) e *in vivo* (es decir, diagnóstico y/o terapéutica) como indicadores, inhibidores y agentes inmunológicos. Los tipos de moléculas y macromoléculas naturales y sintéticas incluyen, pero no exclusivamente: anticuerpos y sus fragmentos, enzimas; receptores celulares; proteínas, polipéptidos, péptidos; polinucleótidos, oligonucleótidos; carbohidratos como polisacáridos, oligosacáridos, sacáridos; lípidos; moléculas orgánicas e inorgánicas como antibióticos, esteroides, hormonas, plaguicidas, herbicidas, tinturas y polímeros.

Los anticuerpos convencionales como los que se encuentran en las especies humana y murina están compuestos por dos cadenas ligeras idénticas y dos cadenas pesadas idénticas. Los sitios de combinación de esos anticuerpos se forman por asociación de los dominios variables de ambas cadenas. Esa asociación es mediada por interacciones hidrófobas en la interfase. Estudios estructurales y bioquímicos demostraron que el dominio variable de la cadena pesada (V_H) proporciona la mayoría de los residuos que entran en contacto con el antígeno (Padlan, 1994) (Chothia & Lesk, 1987) (Chothia, Novotny, et al., 1985). Este descubrimiento ha sido la base para el desarrollo de anticuerpos de un único dominio pesado (dAb) -fragmentos de unión al antígeno recombinantes que constan sólo de V_H (Ward, Gussow, et al., 1989) (Cai & Garen, 1996). Sin embargo, en ausencia de sus VL asociados, se encontró que los V_H eran insolubles, presumiblemente debido a la interfase VL hidrófoba expuesta (Ward, Gussow, et al., 1989). Los anticuerpos de cadena pesada, encontrados en camélidos (Hamers, Atarhouch, et al., 1993) (Sheriff & Constantine), carecen de cadenas ligeras y como resultado tienen dominios variables que reflejan la ausencia de un VL asociado. Los anticuerpos de dominio único derivados de esos anticuerpos son muy solubles y la base estructural de la solubilidad ha sido parcialmente elucidada. En primer lugar, los residuos de la interfase humano/murino conservados, como Val37, Gly44, Leu45 y Trp47 son reemplazados generalmente en los anticuerpos de cadena pesada por tirosina o fenilalanina, glutamato, arginina o cisteína, y glicina, respectivamente. Esas mutaciones aumentan la hidrofiliidad de la interfase VL ya sea por sustituciones no polares a polares o, de manera más sutil, por inducción de cambios conformacionales locales (Desmyter, Transue, et al., 1996) (Spinelli, Frenken, et al.). Esta explicación es respaldada por experimentos en los cuales un V_H insoluble humano se tornó soluble mediante introducción de las mutaciones mencionadas precedentemente en las posiciones 44, 45 y 47 (Davies & Riechmann, 1994). En segundo lugar, en las estructuras resueltas de dos dAbs de camello, las CDR3 se pliegan hacia atrás en la superficie de V_H , enmascarando un área superficial importante de la interfase VL (Desmyter, Transue, et al., 1996) (Decanniere, Desmyter, et al., 1999).

Otras varias características de los V_HH son dignas de mención. Una es la frecuente aparición de los residuos de cisteína en CDR1 y CDR3 (Muyldermans, Atarhouch, et al., 1994) (Lauwereys, Arbabi, et al., 1998) (Vu, Ghahroudi, et al., 1997). Si bien la ubicación de la cisteína de CDR1 está típicamente fijada en la posición 33, la de la cisteína de CDR3 varía. Esos dos residuos forman una unión disulfuro entre CDR1 y CDR3 (Desmyter, Transue, et al., 1996) (Davies & Riechmann, 1996). En la estructura cristalina de un complejo dAb-lisozima, la unión disulfuro imparte rigidez al bucle de CDR3 que se extiende fuera del sitio de combinación y penetra profundamente en el sitio activo de la lisozima (Desmyter, Transue, et al., 1996). Una segunda característica es la longitud promedio más larga de V_HH CDR3, en comparación con los V_HH humano o murino (Muyldermans, Atarhouch, et al., 1994). Una CDR3 más larga, que es una característica de A6, aumenta la superficie de unión al antígeno y, en cierta medida, compensa la ausencia de superficie de unión al antígeno provista por VL en los anticuerpos convencionales (Desmyter, Transue, et al., 1996). Una tercera característica es la ausencia de la unión salina de CDR3 que está típicamente presente en los anticuerpos convencionales y que es formada por los residuos de arginina o lisina en la posición 94 y aspartato en la posición 101 (Desmyter, Transue, et al., 1996) (Muyldermans, Atarhouch, et al., 1994) (Spinelli, Frenken, et al., 1996) (Davies & Riechmann, 1996) (Chothia & Lesk, 1987) (Morea, Tramontano, et al., 1998).

Como fragmentos de unión al antígeno, los dAb son una alternativa atractiva a los scFv porque son de mucho menor tamaño y por el hecho de que demuestran afinidades comparables con las demostradas por los scFv (Ward,

Gussow, et al., 1989) (Spinelli, Frenken, et al., 1996) (Lauwereys, Arbabi, et al., 1998) (Davies & Riechmann, 1995) (Arbabi, Desmyter, et al., 1997) (Reiter, Schuck, et al., 1999). El menor tamaño es una ventaja en las aplicaciones que requieren penetración del tejido y una depuración sanguínea rápida. Las moléculas más pequeñas también ofrecen una tremenda ventaja en términos de los estudios estructurales (Davies & Riechmann, 1994) (Constantine, Goldfarb, et al., 1992 (Constantine, Goldfarb, et al., 1993).

La construcción de bibliotecas de anticuerpo en fago es mucho más simple y más eficaz si se usan anticuerpos de un único dominio (dab) en vez de Fabs o Fvs de una sola cadena. Se puede introducir aleatorización en un porcentaje mucho mayor de posiciones de CDR sin exceder el tamaño práctico de la biblioteca. También se evita el problema de reorganizar los apareamientos VL-V_H originales. Bibliotecas de camélidos de dAb en fago construidas a partir del repertorio de V_HH de camellos inmunizados con antígenos diana, se han comportado bien (Arbabi, Desmyter, et al., 1997) (Lauwereys, Arbabi, et al., 1998) (Decanniere, Desmyter, et al., 1999). Sin embargo, la construcción de bibliotecas a partir de camellos inmunizados presenta problemas evidentes. Además, la naturaleza no humano de productos de esas bibliotecas limita su utilidad. Bibliotecas de dAb sintéticos (Davies & Riechmann, 1995) (Reiter, Schuck, et al.1999), particularmente las basadas en un marco V_H humano, mitigan esos problemas.

Como se describe en detalle en el ejemplo 1, se construyó una biblioteca de expresión en fago de dAb, de acuerdo con la invención aleatorizando las posiciones 100i - 100n del fragmento de unión al ligando V_H parental identificado en la figura 2 (A6V_H-L1).

El tamaño real de la biblioteca fue de 2.4 x 10⁸ con 6.4 x 10⁷ secuencias posibles. Las secuencias de 27 clones tomados al azar de la biblioteca sin seleccionar por afinidad confirmaron la integridad de la biblioteca (Tabla 1).

Tabla 1. Secuencias de la región aleatorizada (100i-n) de clones tomados al azar de la biblioteca A6V_H-L1a.

1.	DQFTHS	15.	CVRGAE
2.	SSMYGN	16.	SPSLAA
3.	IKMQQN	17.	HASGRS
4.	SVDARD	18.	GYMCSL
5.	VSRFGA	19.	HNKDLA
6.	GLGSPK	20.	LADLYM
7.	IDAKWA	21.	WRRRAHE
8.	VSRFGA	22.	SDLFAR
9.	HCLPDG	23.	VSRFGA
10.	RWR?VP	24.	RYRHST
11.	VSRFGA	25.	ARLAGP
12.	LECEGC	26.	SYRPYL
13.	RNVGAL	27.	VVLGNS
14.	RRSDYL		

Como se muestra en la figura 5, este fragmento de unión al ligando V_H potencial que se muestra tiene un porcentaje sustancialmente mejor de monómero en comparación con A6 tipo silvestre.

La biblioteca A6V_H-L1 se seleccionó por afinidad frente a anti-FLAG IgG M-2, H11 scFv y Brucella scFv.

Resultados de M-2

Se realizaron tres rondas de selección por afinidad. Se secuenciaron 21 clones de las rondas 2 y 3 (Tabla 2). El motivo reconocido por M-2 es XYKXXD. Como observamos previamente con la biblioteca A6V_H camelizada, el epítipo se posicionó preferencialmente en el extremo C-terminal de la región aleatorizada, con el residuo D de la secuencia FDI no aleatorizada formando parte del epítipo. Como se muestra en la tabla, las secuencias de los ligadores de M-2 identificadas mediante ELISA del fago reproducen razonablemente el motivo reconocido por M-2.

Tabla 2. Secuencias de los ligadores de M-2 identificadas por ELISA del fago.

Secuencia	Nº de clones
(A) constructos eliminados	
EVQLQAS-----SGYYEDDYRLFDIWGQGTQVTSS	6
EVQLQASGGGLVQP-----GYDSSGYKDLDFRFDIWGQGTQVTSS	1
EVQLQASGGGLVQPGGSLRLS-----RLKVEYYDSSYYGDHYKWFIDIWGQGTQVTSS	1
(B) constructos completos (sólo CDR3)	
DRLKVEYYDSSGYRNEYKEFDI	1
DRLKVEYYDSSGYVDEYKSFDI	1
DRLKVEYYDSSGYAGRYKDFDI	1
DRLKVEYYDSSGYRSDYKRFDI	1
DRLKVEYYDSSGYASHYKDFDI	1
DRLKVEYYDSSGYVDGYKDFDI	1
DRLKVEYYDSSGYTADYKMFDI	1
DRLKVEYYDSSGYDMDYKTFDI	1
DRLKVEYYDSSGYKS?YKSFDI	1
DRLKVEYYDSSGYDYKSQDFDI	1
DRLKVEYYDSSGYKNWDSTFDI	1
DRLKVEYYDSSGYKDWSSFDI	1
DRLKVEYYDSSGYKDGDSFFDI	1

Resultados de H11

5 Se realizaron tres rondas de selección por afinidad en cada caso. No se observaron secuencias eliminadas para ningún antígeno.

Se secuenciaron 21 ligadores de H11. La secuencia FSSP está presente en 13 de los 21 ligadores de H11 (Tabla 3).

10 Tabla 3. Secuencias de CDR3 de los ligadores de H11 identificadas por ELISA del fago. La región aleatorizada se muestra en negritas

CDR3	Nº de clones
(A) H11 scFv	
DRLKVEYYDSSG--Y S FSSPFDI	7
DRLKVEYYDSSG Y YDFSSPFDI	3
DRLKVEYYDSSG Y NLFSSPFDI	1
DRLKVEYYDSSG Y SEFSSPFDI	1
DRLKVEYYDSSG Y SDFSSPFDI	1
DRLKVEYYDSSG Y TDMSWEFDI	1
DRLKVEYYDSSG Y DLGSWEFDI	1

CDR3	Nº de clones
(A) H11 scFv	
DRLKVEYYDSSGYYDYVSWDFDI	1
DRLKVEYYDSSGYYDWGSWTFDI	1
DRLKVEYYDSSGYYDG?TWDFDI	1
DRLKVEYYDSSGYYWEGSGLFDI	1
DRLKVEYYDSSGYYIWYSGLFDI	1
DRLKVEYYDSSGYYSSWASAFDI	1

Brucella scFv (Yst9.1)

5 Se seleccionaron para expresión los clones obtenidos por selección por afinidad frente a Yst9.1 scFv. Dos clones tuvieron la secuencia VSRFGA. El VSRFGA dAb se designó Yst9.1-L3-9.

El clon de VSRFGA se expresó con buen rendimiento (aproximadamente 18 mg/litro de cultivo bacteriano) y se caracterizó por BIACORE.

10 Unión de Yst9.1-L3-9 dAb a Yst9.1 scFv inmovilizado

Luego de la purificación IMAC, el dAb se purificó por cromatografía de exclusión por tamaño en Superdex 75 antes de un análisis BIACORE. Dio un perfil de elución semejante al de la molécula A6V_H-L1 parental (Figura 6).

15 Para la determinación de la cinética de unión y las afinidades, el dAb purificado se inyectó sobre superficies activas (1800 RU de Yst9.1) y de referencia (9400 RU de BSA) para generar el sensorgrama y las constantes presentadas en la figura 7. La superficie de Yst9.1 se regeneró con HCl 10 mM (6 segundos de contacto). Todas las proteínas se inmovilizaron en chips CM5 por acoplamiento con amina. También se usó una superficie de 2000 RU de H11 scFv como referencia. El dAb no se unió a BSA. Los datos sustraídos, usando la superficie BSA como referencia, para la
20 unión de dAb a Yst9.1 scFv se ajustan bastante bien a un modelo de interacción 1:1. Se calculó una KD de 117 nM.

Se describe que el fragmento de unión al ligando parental tiene sustituciones de aminoácidos adicionales en la interfase VL que reducen la tendencia a la agregación atribuible a la "pegajosidad" del V_H dAb en esta interfase. Por ejemplo las sustituciones en las posiciones 44, 45 y 47 en comparación con A6V_H tipo silvestre (Figura 1) se ilustra
25 en la figura 8.

El fragmento de unión al ligando parental puede tener una CDR3 larga.

30 La biblioteca se puede construir usando la molécula de unión al ligando V_H parental A6V_H-L1 como molde y preserva la longitud total de esta CDR3 y, además, al menos una de las posiciones 44, 45, 47 y 94 (y preferentemente todas) es alterada, (preferentemente incluidas 44 o 45) a los residuos tipo camélido.

35 Davies y Riechmann (1995) construyeron una biblioteca de dAb camelizada mediante aleatorización de los residuos de aminoácidos de CDR3 pero la biblioteca fue diez veces más grande y produjo dAbs anti-hapteno con constantes de disociación en el rango de 100-400 nM. Sin embargo, los dAb anti-proteína aislados tuvieron una afinidad débil (Davies & Riechmann, 1995) (Davies & Riechmann, 1996). Por consiguiente, una biblioteca de menor tamaño como la construida aquí puede contener por lo tanto sólo dAbs anti-proteína débiles. El aislamiento de dichos dAb sería difícil con expresión monovalente (Lowman, Bass, et al., 1991). En un formato de vector de fago los dAb se expresan en 3-5 copias y por consiguiente existe potencial para avidéz que aumenta la probabilidad de aislar
40 ligadores débiles (Nissim, Hoogenboom, et al., 1994).

Otras opciones de la estrategia de aleatorización de las bibliotecas de la invención, se describen a continuación.

45 Se describe un método para aumentar la probabilidad de que los fragmentos de unión expresados en la biblioteca tengan características que se aproximen a las características de solubilidad deseadas encontradas en el fragmento de unión tipo silvestre. Durante la construcción de la biblioteca, se agregan nucleótidos de la región variable en una adición en pasos y se selecciona una proporción de nucleótidos que es sesgada a favor de producir aminoácidos que reflejen el ADN de la especie silvestre o parental.

Por lo tanto, un método para sesgar una biblioteca a favor de obtener los porcentajes elegidos de residuos de aminoácidos tipo silvestre, se logra creando sustituciones de residuos mediante el uso de diferentes niveles de agregado de cantidades conocidas de los diversos dNTP como se describe a continuación. Cuando se crea una biblioteca de fago, la aleatorización de los aminoácidos se logra a menudo mediante síntesis de ADN. Se aparea un cebador al lado del ADN que codifica la región variable y se agregan nucleótidos al azar para sintetizar las regiones variables aleatorizadas. Corrientemente, en el paso de síntesis del ADN utilizado para producir la región variable de la biblioteca del fago, se usa una proporción de nucleótidos de 1:1:1:1, que genera una región variable totalmente aleatoria. Con el método presente, durante la síntesis de la región variable, la probabilidad de lograr afinidad u otras características deseables encontradas en el tipo silvestre es de la manera siguiente. En cada paso de adición de un nucleótido a la región variable del ADN, se selecciona una proporción de dNTP que se sesga a favor de producir aminoácidos que reflejen el ADN de la especie parental (tipo silvestre).

La tabla 4 registra gráficamente los residuos de aminoácidos particulares o las secuencias de residuos y los tipos preferidos de sustituciones de aminoácidos de acuerdo con los diversos ejemplos de la invención que van a ser definidos más adelante. La selección de aminoácidos para aleatorización o aleatorización parcial se basa en adoptar uno o más de diversos métodos que incluyen uno o más de los siguientes:

1. reconocimiento universal de los aminoácidos tipo silvestre a través de un sesgado de base ancha a favor de los aminoácidos tipo silvestre en una o más regiones de interés (aproximadamente 10% a 90%) para mantener las características de la molécula de unión al ligando V_H parental;
2. reconocimiento selectivo de aminoácidos que son importantes para mantener como tipo silvestre, a través de un sesgado (aproximadamente 10-100%) para mantener regiones conservadas o estratégicas de residuos de aminoácidos de la molécula de unión al ligando V_H parental; y
3. reconocimiento de aminoácidos seleccionados como importantes para la interacción intermolecular y sesgado de esos aminoácidos al tipo silvestre y a aminoácidos del mismo tipo.

Tabla 4

Residuo de aminoácido N°	Descripción de diversas constituciones aminoacídicas preferidas
a. Al menos uno de 100a-100h, preferentemente en cada posición de 100a-100h	Aleatorizar;
	Al menos aproximadamente 10% de sesgado a favor de los aminoácidos tipo silvestre;
	Al menos aproximadamente 50% de sesgado a favor de los aminoácidos tipo silvestre;
	Al menos aproximadamente 90% de sesgado a favor de los aminoácidos tipo silvestre;
	Aleatorizar, pero sesgar 100f al tipo silvestre (aproximadamente 10-100%)
b. Al menos un aminoácido de: 100a-100b y 100g-100h preferentemente en cada posición de 100a-100b y 100g-100h	Aleatorizar;
	Aleatorizar con sesgo al tipo silvestre (aproximadamente 10-100%), preferentemente al menos aproximadamente 50% tipo silvestre, alternativamente al menos aproximadamente 90% de aminoácidos tipo silvestre;
	Aleatorizar con sesgo a uno de los aminoácidos seleccionado del grupo que consiste en tirosina, histidina, glutamina, asparragina, lisina, ácido aspártico y ácido glutámico (aproximadamente 10-100%)
c. Al menos uno de 100b-100g, preferentemente en cada posición de 100b-100g	Aleatorizar;
	Eliminar;
d. 100a-100h	Adiciones al azar de hasta 10 aminoácidos;
	Deleciones al azar de hasta 7 aminoácidos;
e. 95-100o	Aleatorizar;
	Adiciones al azar de hasta 10 aminoácidos;
	Deleciones al azar de hasta 7 aminoácidos;
f. Al menos uno de 95-100, preferentemente en	Aleatorizar;

Residuo de aminoácido N°	Descripción de diversas constituciones aminoacídicas preferidas
cada posición de 95-100	Aleatorizar con sesgo al tipo silvestre (aproximadamente 10-100%), preferentemente al menos aproximadamente 50% tipo silvestre, o preferentemente al menos aproximadamente 90% de aminoácidos tipo silvestre;
	Sin variación (el cebador abarca esta región)
g. 101-102 aminoácidos conservados	Sin variación (el cebador abarca esta región)
	N/A
i. Al menos un aminoácido de: 100a-100b, 100g-100h y 100l-100o, preferentemente en cada posición de 100a-100b, 100g-100h y 100l-100o	Aleatorizar con sesgo al tipo silvestre (aproximadamente 10-100%), preferentemente al menos aproximadamente 50% tipo silvestre, más preferentemente al menos aproximadamente 90% de aminoácidos tipo silvestre;
	Aleatorizar con sesgo a uno de los aminoácidos seleccionado del grupo que consiste en tirosina, histidina, glutamina, asparragina, lisina, ácido aspártico y ácido glutámico (aproximadamente 10-100%);
	Sesgar a aminoácidos aromáticos (10-100%)
j. 95-100h	Aleatorizar pero mantener todos los aminoácidos consecutivos 5-10 como tipo silvestre
k. 100a-100o	Aleatorizar pero mantener todos los aminoácidos consecutivos 5-10 como tipo silvestre

5 A menos que se implique necesariamente algo diferente como resultado de consideraciones logísticas, se debe comprender que las diversas opciones relacionadas con la elección de aminoácidos para sustitución al azar, sesgada o fija (especificada en la columna 1) así como las diversas opciones relacionadas con los tipos de sustitución (columna 2) no se excluyen mutuamente. Además se prevén las diversas permutaciones y combinaciones de dichas sustituciones. Por ejemplo, las sustituciones mencionadas en la fila a. (cualquiera o más aminoácidos y preferentemente todos los aminoácidos de los residuos 100a-100h) N° 3 (al menos aproximadamente 50% de aminoácidos tipo silvestre) se pueden combinar con la fila b. (cualquiera o más y preferentemente todos los residuos de aminoácidos 100a, 100b, 100g y 100h) N° 2 (por ejemplo, al menos aproximadamente 90% de aminoácidos tipo silvestre) de modo que, por ejemplo, cualquiera de los 3 aminoácidos en 100a - 100h se sesgue a favor del tipo silvestre en aproximadamente 50% de los fragmentos de unión al ligando V_H variantes y 100a y 100b se sesguen a favor del tipo silvestre en 90% de los fragmentos de unión potenciales.

15 Por implicancia necesaria los tres aminoácidos que son sesgados a favor del tipo silvestre no son residuos 100a ni 100b, pero pueden ser otros tres residuos cualesquiera. Además, se debe comprender que cada una de las diversas opciones descritas en este documento se dan a conocer, excepto en la medida en que sea logísticamente imposible, con referencia a cada uno de los diversos aspectos y definiciones de la invención. Por otra parte, se debe comprender que las frases como “al menos aproximadamente 10%, o aproximadamente 10-100%” pretenden especificar una preferencia por cada una de las unidades de porcentaje entre aproximadamente 7 y 100% que son prácticamente alcanzables por el diseño de cebador oligonucleótido y amplificación por PCR descritos más adelante, así como otras técnicas de PCR bien conocidas y técnicas de mutación controlada descritas en el área, y las variaciones rutinarias de dichas técnicas. Por el mismo motivo, frases como “al menos 80%” pretenden especificar una preferencia por cada una de las unidades de porcentaje entre 80% y 100%. Se debe comprender que sesgar un porcentaje menos de 100% implica, a menos que se implique o indique algo diferente, que el porcentaje remanente está totalmente aleatorizado. Además, se debe comprender, por ejemplo, que 90% de sesgado a favor de los aminoácidos tipo silvestre en una posición de aminoácido dada, debe ser aproximada mediante control de las cantidades porcentuales de cada uno de los tres nucleótidos pertinentes (de modo que, por ejemplo, el producto de las probabilidades de aparición de los tres nucleótidos deseados en secuencia en la cadena en crecimiento sea de 90%) de modo de proporcionar 90% del triplete o los tripletes de codificación correctos y un total de 10% de tripletes de codificación al azar, teniendo en cuenta la degeneración del código genético (por ejemplo si dos tripletes de codificación diferentes resultan en un aminoácido dado, entonces la suma de las probabilidades de lograr esos dos tripletes deberá ser igual a 90%). Esto se lleva a cabo preferentemente aminoácido por aminoácido, de manera que, por ejemplo, la probabilidad de lograr dos y tres aminoácidos tipo silvestre en secuencia, en el caso de 90% de sesgado es 0.81 y 0.73, respectivamente, etc. Se debe entender que este alto nivel de sesgado puede ser adecuado sólo para parte de la secuencia de codificación dentro de la cual se introduce variabilidad y que niveles mayores de sesgado son aceptables, cuando por ejemplo prácticamente todos los aminoácidos de una CDR3 larga son sesgados, como se da a conocer en una de las realizaciones este documento.

Concordantemente existe un equilibrio a ser abatido entre una biblioteca diversa grande y el sesgado a características multifactorial es como la solubilidad. No obstante se prevé que la biblioteca producida puede ser una biblioteca combinada en la cual varias bibliotecas, cada una con diferentes grados de sesgado al tipo silvestre, por ejemplo, 60%, 50%, 40% y 30%, se combinan para obtener tanto la variabilidad como la similitud deseadas. El fragmento de unión parental preferido puede ser genomanipulado para maximizar las características deseadas (por ej. solubilidad, interacción intermolecular) y después hacerlo el sujeto de las bibliotecas con diversos grados de sesgado. En esta conexión, la biblioteca podría sesgarse para ser rica en aminoácidos que sean muy solubles. Se debe comprender que ambos brazos (mitades) del bucle más largo preferido que forman las CDR3 se pueden sesgar a aminoácidos que son favorecidos por la interacción intermolecular, preferentemente aminoácidos cargados, a fin de proporcionar un método para generar, además del tamaño del bucle, estructuras de bucle diferentes. Este sesgo puede ser introducido sistemáticamente o reducido sistemáticamente mediante aleatorización, en bibliotecas combinadas colaboradoras que tienen diversos grados de sesgado.

Con respecto a la aplicación de estos métodos al V_H parental, preferentemente, las CDR3 de una diversidad de longitudes diferentes entre 16 y 33 aminoácidos están representadas predominantemente entre los fragmentos de unión al ligando V_H variantes. Preferentemente las CDR3 de una variedad de longitudes diferentes, entre 18 y 25 aminoácidos, o entre 18 y 23 aminoácidos, están representadas predominantemente en la biblioteca. Aunque el término "predominante" implica corrientemente una representación mayoritaria de los fragmentos de unión al ligando V_H variantes con la CDR3 larga especificada, la invención también prevé una representación menos sustancial, especialmente dentro de una biblioteca de tamaño razonablemente grande ($>10^7$). Preferentemente, los fragmentos de unión al ligando V_H variantes con la CDR3 larga especificada tienen una representación mayoritaria dentro de la biblioteca y más preferentemente una representación aún mayor o exclusiva.

Opcionalmente, la molécula de unión al ligando V_H parental se reduce en tamaño y la molécula de unión al ligando V_H parental se modifica opcionalmente mediante delección de una porción de la CDR2. En otra realización, las CDR3 de la misma longitud que la de la molécula de unión al ligando V_H parental, están predominantemente o exclusivamente representadas en los fragmentos de unión al ligando V_H variantes.

En otro aspecto, la región CDR3 es específicamente retenida junto con elementos de la secuencia humana de otras regiones que confieren características de solubilidad favorables, para crear una biblioteca de expresión en fago con características de solubilidad favorables, preferentemente cuando se la compara con los fragmentos de unión al ligando V_H variantes que tienen regiones hipervariables completamente aleatorizadas (particularmente CDR3). En particular, los inventores de la presente encontraron que las características de solubilidad favorables de una molécula de unión al ligando V_H parental se pueden mantener en la población de fragmentos de unión al ligando V_H variantes en el curso de la aleatorización de las regiones hipervariables mediante sesgado, de todos o de los residuos de aminoácidos seleccionados, al tipo silvestre y/o sesgado a favor de los residuos de aminoácidos que favorecen ciertas, o varios tipos de, interacciones moleculares. Esto se lleva a cabo respectivamente aumentando las cantidades porcentuales de bases nucleótidas que representan aminoácidos tipo silvestre y/o aminoácidos que proporcionan interacciones intermoleculares favorables durante el procedimiento de aleatorización por ej. mutagénesis sitio-dirigida por PCR.

Por lo tanto, los fragmentos de unión al ligando V_H variantes que tienen CDR3 relativamente largas de longitudes variables se producen insertando al azar o parcialmente al azar diferentes cantidades de tripletes de nucleótidos en cualquier parte de una porción aleatorizada del marco V_H parental. Se sintetizan cebadores de la longitud y la composición de nucleótidos deseadas seguido de amplificación por PCR. La aleatorización deseada se puede alcanzar sesgando la composición de nucleótidos del cebador. La producción de expresiones de ligadores variantes de CDR3 larga también se puede llevar a cabo combinando varias bibliotecas de fragmentos de unión al ligando V_H variantes que tienen CDR3s aleatorizadas o parcialmente aleatorizadas de longitudes uniformes respectivas diferentes. Esas estrategias no se excluyen mutuamente.

Los términos adicionales siguientes se usan en este documento como sigue, a menos que el contexto implique lógicamente algo diferente:

"Sesgar", "sesgado a favor de" y formas relacionadas de esos términos pretenden generalmente referirse a la ponderación en el curso de la introducción de una variación en la molécula de unión al ligando parental.

"Homólogo" u "homología" según se usa en este documento se refiere a la "identidad" o "similitud" según se usa en el área, y significa relaciones entre dos o más secuencias de polinucleótidos o aminoácidos, según se determina por comparación de las secuencias. En el área, identidad también significa el grado de relación de la secuencia entre secuencias de polinucleótidos, como puede ser el caso, según se determina por la correspondencia entre series de dichas secuencias. Se pueden calcular fácilmente tanto la identidad como la similitud (Lesk, A. M., ed., Computational Molecular Biology, Oxford University Press, Nueva York, 1988; Smith, D. W., ed., Biocomputing: Informatics and Genome Projects, Academic Press, Nueva York, 1993; Griffin, A. M., and Griffin, H. G., eds., Computer Analysis of Sequence Data, Part I, Humana Press, Nueva Jersey, 1994; von Heinje, G., Sequence Analysis in Molecular Biology, Academic Press, 1987; y Gribskov, M. y Devereux, J., eds., Sequence Analysis Primer, M Stockton Press, New York, 1991). Si bien existen varios métodos para medir la identidad y la similitud

entre dos secuencias de polinucleótidos, ambos términos son bien conocidos por los expertos (von Heinje, G., 1987; Gribskov, M. y Devereux, J., 1991; y Carillo, H., y Lipman, D., 1988). Los métodos empleados comúnmente para determinar la identidad o la similitud entre secuencias incluyen, pero no exclusivamente, los publicados en Carillo, H., y Lipman, D. (1988, SIAM J. Applied Math., 48:1073). Los métodos para determinar la identidad y la similitud están codificados en programas informáticos. Los métodos de programas informáticos preferidos para determinar la identidad y la similitud entre dos secuencias incluyen, pero no exclusivamente, el paquete de programas GCG (Devereux, J., et al. (1984), Nucleic Acids Research 12(1): 387), BLASTP, BLASTN, y FASTA (Atschul, S. F. et al. (1990), J.Molec. Biol. 215: 403). "Porcentaje de homología" o "% homólogo" o los términos relacionados incluyen tanto las interpretaciones como los métodos de cálculo siguientes: 1) un porcentaje aproximado de la secuencia de referencia en términos del número de residuos comunes (por ej. 80% de 11 se entiende que es una aproximación en la medida en que la aplicación del porcentaje no produzca un número de unidades de residuos, en cuyo caso tanto el número inmediatamente superior como el inmediatamente inferior, 9 y 8 respectivamente, se considera que deben estar cubiertos); 2) el porcentaje de fragmentos de unión teóricamente alcanzable que puede tener la secuencia completa de tipos silvestre, que se calcula como un producto de las probabilidades de que el aminoácido de tipos silvestre aparezca en una posición de aminoácido dada.

Regiones "conservadas" se refiere a las regiones que se encuentran comúnmente en al menos otros anticuerpos del mismo tipo o en al menos la misma especie de mamífero.

"Tipo silvestre" se refiere al fragmento de unión parental, que puede ser una variante del natural o el fragmento de unión al ligando A6 V_H parental nativo, dependiendo del contexto.

"En pasos" se refiere a la adición de, por ejemplo, ácidos nucleicos, de tal manera que la cantidad de ácidos nucleicos agregada en cada paso sea rigurosamente controlada, generalmente un ácido nucleico por vez.

"Extenderse o abarcar" no impide las delecciones o adiciones dentro del fragmento de unión V_H parental que no sean antagónicas a la operación de la invención.

"Tipo camélido" se refiere específicamente a una o más características de la interfase VL del camélido.

"Soluble" incluye el significado generalmente atribuido en el área e incluye sin limitaciones (basada en los fenómenos correlacionados con la solubilidad) las cantidades relativas de proteína recombinante plegada naturalmente liberadas desde la célula.

"Porcentaje de sesgo" o "% de fragmentos de unión" (o "sesgado 10-100%", etc.) se refiere al sesgado aminoácido por aminoácido (aunque otras técnicas para lograr el mismo efecto pueden resultar evidentes para los expertos). De manera similar, la especificación de que existen aminoácidos de tipo silvestre en una posición o serie de posiciones especificadas, por ejemplo, en al menos aproximadamente 50% de los fragmentos de unión potenciales pretende significar tanto que se busca el 50% de sesgo en una posición dada o que un total de 50% de los tripletes de nucleótidos correctos está representado.

"Aproximadamente" en referencia a los porcentajes pretende tomar en cuenta la baja de diversos fragmentos de unión al ligando V_H variantes deseados, la asunción de que el resultado probabilístico no se alcanzará en la práctica y que cierta variación en los métodos para obtener los resultados especificados se considera adecuada. El término 50% en referencia a un número impar de residuos de aminoácidos significa que se hace referencia a uno más o uno menos de la mitad de los aminoácidos.

La práctica de la presente invención emplea, a menos que se indique algo diferente, técnicas convencionales de biología molecular (inclusive técnicas de recombinación), microbiología, biología celular, bioquímica e inmunología, que forman parte del conocimiento de los expertos. Dichas técnicas se explican completamente en la bibliografía, como, "Molecular Cloning: A Laboratory Manual", segunda edición (Sambrook et al., 1989); "Oligonucleotide Synthesis" (M.J. Gait, ed., 1984); "Animal Cell Culture (R.I. Freshney, ed, 1987); "Methods in Enzymology" (Academic Press, Inc.); "Handbook of Experimental Immunology" (D.M. Wei & C.C. Blackwell, eds.); "Gene Transfer Vectors for Mammalian Cells" (J.M. Miller & M.P. Calos, cds., 1987); "Current Protocols in Molecular Biology (F.M. Ausubel et al., eds., 1987); "PCR: The Polymerase Chain Reaction", (Mullis et al., eds., 1994); "Current Protocols in Immunology" (J.E. Coligan et al., eds., 1991). Esas técnicas se pueden aplicar a la producción de los polinucleótidos y polipéptidos de la invención, y, como tales, pueden ser tenidas en cuenta en la elaboración y la práctica de la invención. Las técnicas particularmente útiles para las realizaciones particulares se tratarán en las secciones siguientes.

Las técnicas de recombinación genética han permitido la clonación y la expresión de anticuerpos, sus fragmentos funcionales y los antígenos reconocidos. Esos anticuerpos genomanipulados proporcionan nuevos métodos de producción y nuevas modalidades de tratamiento. Por ejemplo, se expresaron fragmentos de inmunoglobulinas funcionales en bacterias y, en semillas y plantas de tabaco transgénicas. Skerra (1993) Curr.Opin. Immunol. 5:256:262; Fiedler and Conrad (1995) Bio/Technology 13:1090-1093; Zhang et al. (1993) Cancer Res. 55:3384-3591;

Ma et al. (1995) Science 268:916; y para una reseña de anticuerpos sintéticos, véase Barbas (1995) Nature Med. 1:836-839.

5 Las secuencias de nucleótidos se pueden aislar, amplificar y procesar mediante técnicas de recombinación estándares. Dichas técnicas corrientes en el área incluyen la digestión con nucleasas de restricción y la amplificación por la reacción en cadena de la polimerasa (PCR), o una combinación de ellas adecuada.

La tecnología de PCR se describe en las patentes de los Estados Unidos N° 4,683,195; 4,800,159; 4,754,065; y 4,683,202, así como en PCR: The Polymerase Chain Reaction, Mullis et al., eds., Birkaswer Press, Boston (1994).

10 Además de los métodos de PCR específicos para sesgar los residuos de aminoácidos de A6 tipo silvestre detallados más adelante, es posible producir múltiples cebadores oligonucleótidos diferentes que consten de los residuos de aminoácidos especificados (uno o más) de la molécula A6 tipo silvestre (por ej. residuos de CDR3), mezclando éstos en concentraciones adecuadas con un cebador oligonucleótido completamente aleatorizado (por ej. CDR3) y sometiendo la mezcla de cebadores oligonucleótidos a PCR. Esto dará como resultado una población de biblioteca de fago sesgada a la elección de cada uno (es decir las cantidades de los cebadores selectivamente aleatorizados y totalmente aleatorizados en la mezcla determinará el porcentaje de cada representación de CDR3 en la biblioteca).

20 Los polinucleótidos que comprenden una secuencia deseada se pueden insertar en un vector adecuado y el vector a su vez se puede introducir en una célula hospedadora adecuada para replicar y amplificar. Los polinucleótidos se pueden introducir en células hospedadoras por cualquier medio conocido en el área. Las células se transforman introduciendo un polinucleótido exógeno por absorción directa, endocitosis, transfección, apareamiento-f o electroporación. Una vez introducido, el polinucleótido exógeno se puede mantener dentro de la célula como un vector no integrado (como un plásmido) o integrado en la célula hospedadora mediante métodos estándares. Véase, por ej., Sambrook et al. (1989). El ARN también se puede obtener de una célula hospedadora transformada, o se puede obtener directamente del ADN usando una ARN polimerasa dependiente de ADN.

30 Los vectores de clonación y expresión adecuados incluyen todos los conocidos en el área, por ejemplo, lo que se usan en sistemas de expresión en bacterias, mamíferos, levaduras e insectos. Los vectores específicos y las células huésped adecuadas son conocidos en el área y no se describen en detalle en este documento. Véase, por ej. Gaceta y Ramji, Vectors, John Wiley & Sons (1994).

35 Las técnicas de expresión en fago se describen en general o se mencionan en algunas de las referencias generales precedentes, así como en las patentes de los Estados Unidos N° 4,593,002; 5,403,484; 5,837,500; 5,571,698; 5,750,373; 5,821,047; 5,223,409 y 5,702,892. "Phage Display of Peptides and Proteins", (Kay, Brian K. et al., 1996); "Methods in Enzymology", Vol. 267 (Abelson, John N., 1996); "Immunology Methods Manual", (Lefkovits, Ivan, 1997); "Antibody phage display technology and its applications", (Hoogenboom, Hennie R. et al., 1998). Immunotechnology 4 p.1-20; Cesareni G et al. Phage displayed peptide libraries. Comb Chem High Throughput Screen. 1999 Feb;2(1):1-17; Yip, YL et al. Epitope discovery using monoclonal antibodies and phage peptide libraries. Comb Chem High Throughput Screen. 1999 Jun;2(3):125-38; Rodi DJ et al. Phage-display technology-finding a needle in a vast molecular haystack. Curr Opin Biotechnol. 1999 Feb; 10(1):87-93.

45 Generalmente, el ADN que codifica millones de variantes de un fragmento de unión parental puede ser clonado por lotes en el genoma del fago como una fusión al gen que codifica una de las proteínas de recubrimiento del fago (pIII, pVI o pVIII). Una vez expresada, la fusión de la proteína de recubrimiento se incorporará en nuevas partículas del fago nueva que son ensambladas en la bacteria. La expresión del producto de fusión y su subsiguiente incorporación en el recubrimiento del fago maduro produce el ligando que se presenta la superficie del fago, mientras su material genético reside dentro de la partícula del fago. Esta conexión entre el genotipo y el fenotipo del ligando permite el enriquecimiento de un fago específico, por ej. usando selección en un objetivo inmovilizado. El fago que expresa un ligando pertinente será retenido, en tanto el fago no adherente se eliminará por lavado. El fago unido se puede recuperar de la superficie, reinfectar en una bacteria y volver a cultivar para el enriquecimiento adicional, y eventualmente para el análisis de unión. El éxito de la expresión en fago del ligando depende de la combinación de este método de expresión y enriquecimiento, con la síntesis de grandes repertorios combinatorios en el fago.

55 Si bien el uso del fago se describe como una realización para la producción de bibliotecas de expresión y selección de fragmentos de unión particulares, se debe comprender que se puede usar cualquier paquete genético adecuado para la producción de bibliotecas de la invención. Dichos paquetes genéticos adecuados incluyen células, esporas y virus (véase la patente de los Estados Unidos N° 5,571,698), o cualquier otro paquete genético replicable adecuado. Con respecto a los métodos a base de células, otro método popular de presentación de una biblioteca es el sistema de dos híbridos (Feilds and Stemglanz, 1994, Trends in Genetics 10:286-292). Los expertos en el área entenderán que los sistemas *in vitro* (que no son a base de células) pueden ser igualmente aplicables a los métodos de la presente invención, por ejemplo expresión en ribosoma (Hanes et al., 1998) o fusión de ARN-péptido (Mattheakis et al., 1994, Proc Natl Acad Sci USA 91:9022-26; Hanes et al., 1999, Curr Top Microbiol Immunol 243:107-22).

65 La expresión en ribosoma es una técnica bien documentada y que puede ser útil para generar bibliotecas. Este método completamente *in vitro* permitió tener bibliotecas con una diversidad $>10^{12}$. En ese método, se expresa un

péptido en la superficie de un ribosoma que lo está traduciendo. En resumen, una biblioteca de moléculas de ARNm (podríamos comenzar con A6) es traducida por un sistema de traducción *in vitro* hacia el extremo 3', de modo que el ribosoma no se desprende. La proteína emerge del ribosoma de modo tal que se puede plegar pero no se desprende. En algunos casos, hay un paso de plegado adicional en un ambiente oxidante (importante para las proteínas con enlaces disulfuro). El complejo entero de la proteína plegada, el ribosoma y el ARNm, que es estable por varios días, se puede después seleccionar por afinidad frente a un ligando que es reconocido por la proteína traducida. (Por ejemplo, la proteína traducida podría ser un anticuerpo y el ligando su antígeno). Después el ARNm se puede amplificar mediante transcripción inversa y PCR. Esta técnica ha sido utilizado con éxito para generar fragmentos del anticuerpo scFv con gran afinidad por su objetivo. Se hace referencia a Hanes, J., Jermutus, L., Weber-Bornhauser, S., Bosshard, H.R. & Pluckthun, A. Ribosome display efficiently selects and evolves high-affinity antibodies *in vitro* from immune libraries. *Proc. Natl. Acad. Sci. USA* 95, 14130-14135 (1998); Schaffitzel, C., Hanes, J., Jermutus, L. & Pluckthun, A. Ribosome display: an *in vitro* method for selection and evolution of antibodies from libraries. *Journal of Immunological Methods* 231, 119-135 (1999); He, M. et al. Selection of a human anti-progesterone antibody fragment from a transgenic mouse library by ARM ribosome display. *Journal of Immunological Methods* 231, 105-117 (1999); Roberts, R.W. Totally *in vitro* protein selection using mRNA-protein fusions and ribosome display. *Current Opinion in Chemical Biology* 3, 268-273 (1999); Williams, C. Biotechnology match making: screening orphan ligands and receptors. *Current Opinion in Biotechnology* 11, 42-46 (2000); Mattheakis, L.C., Bhatt, R.R. & Dower, W.J. An *in vitro* polysome display system for identifying ligands from very large peptide libraries. *Proc. Natl. Acad. Sci. USA* 91, 9022-9026 (1994).

Construcción de la biblioteca A6V_H-L1

Ejemplo 1

Se construyó una biblioteca de expresión en fago de dAb empleando la porción V_H de A6 como molde de inicio y sustituciones de aminoácidos en las posiciones 6, 23, 82a, 93 y 108, como se muestra en la tabla 5 más adelante. Además, se introdujeron mutaciones silenciosas en codones para los aminoácidos 3-16 (FR1) para eliminar un sitio de recombinación putativo. Todas las mutaciones se introdujeron mediante empalmado por superposición de extensión-PCR (SOE). El V_H modificado se usó como un andamio para la aleatorización completa en las posiciones 100i-100n ubicadas en el extremo carboxi terminal de la CDR3 de A6.

Tabla 5. Molde de A6V_H-L1

Posición	Posición e identidad del aminoácido				
	6	23	82a	93	108
A6V _H tipo silvestre	E	S	S	V	T
A6V _H modificado	A	A	N	A	Q

Eliminación del sitio de recombinación putativo

Se cambiaron los codones para los aminoácidos 3 a 16, que rodean el sitio de recombinación. Utilizando los pares de cebadores Chi.R/FP y Chi.F/RP (Tabla 6) y el plásmido pSJF-A6V_H como molde, se sintetizaron fragmentos 5' y 3' por PCR en un volumen total de 50 µl que contenía 10 pmol de cada cebador, cada una de las 4 dNTP 2 mM, tampón 1X y 2.5 unidades de AmpliTaq™ ADN polimerasa (Perkin Elmer). El protocolo de PCR consistió en un paso de desnaturalización inicial a 94 °C durante 3 min, seguido de 30 ciclos a 94 °C durante 30 segundos, 55 °C durante 30 s, 72 °C durante 1 min y un paso de extensión final a 72 °C durante 10 min. Los dos fragmentos se purificaron en gel usando el juego Q1Aquick Gel Extraction™ (Q1AGEN) y se ensambló un constructo más grande desde los fragmentos 5' y 3' realizando un empalme por superposición de extensión (SOE) PCR usando RP y FP. En resumen, el vial de reacción que contenía los fragmentos 5' y 3', cada uno de las 4 dNTP 200 µM, 5 µl de tampón 10X (NEB), y 2 unidades de Vent ADN polimerasa (NEB) se sometió a 7 ciclos de 1 min a 94 °C y 2.5 min a 72 °C. Para amplificar el constructo ensamblado, se agregaron los cebadores RP y FP a una concentración final de 1 pmol/µl y la mezcla se sometió a 30 ciclos de 1 min a 94 °C, 30 segundos a 55 °C y 1 min a 72 °C. El producto amplificado se purificó (Q1Aquick PCR Purification™ kit) y se usó como molde para los pasos de mutagénesis. La secuenciación reveló que las secuencias de recombinación putativas habían sido eliminadas.

Modificación de las posiciones 6, 23, 82a, 93 y 108

Empleando SOE, se mutaron inicialmente las posiciones 23, 82a y 93 (Tabla 5). Usando el molde Chi, se sintetizaron tres fragmentos por PCR usando los pares de cebadores A6V_H.S23A.R/A6V_H.S82aN/V93A.F (fragmento 1), A6V_H.S23A.F/RP (fragmento 2) y A6V_H.V93A.R/FP (fragmento 3) (Tabla 6). Los primeros dos fragmentos se ensamblaron por SOE usando RP/A6V_H.S82aN/V93A y el fragmento resultante se empalmó al tercer fragmento usando los cebadores RP/FP. La secuenciación reveló las mutaciones deseadas en las posiciones 23, 82a y 93. Este producto se usó a su vez como molde para la ronda final de experimentos SOE para alterar las posiciones 6 y

108. En primer lugar, se sintetizaron dos fragmentos usando los pares de cebadores A6V_H.E6A.R/A6V_H.T108Q.F y A6V_H.E6A.F/RP seguido de SOE usando los cebadores RP/A6V_H.T108Q.F. El constructo final se purificó, se digirió con *Bam*HI y *Eco*RI, se volvió a purificar y se ligó al plásmido de expresión restringida por *Bam*HI/*Eco*RI. Después se analizaron las colonias para identificar los clones que contenían el A6V_H mutado.

5

Tabla 6. Cebadores utilizados para la eliminación del sitio chi y la mutagénesis de A6V_H.

Chi.R
CAATTACAAGAAAGTGGTGGCGGACTGGTGCAACCAGGAGGTTCCCTGAGACTC
Chi.F ,
ACTTTCTTGTAATTGGACCTCGGCCTGCGC
A6VH.S23A.R
CTCTCCTGTGCTGCCTCTGGA
A6VH.S23A.F
TCCAGAGGCAGCACAGGAGAG
A6VH.S82aN/V93A.F
CGCACAGTAATACACAGCCGTGTCTCAGCTCTCAGACTGTTTCATTTGAAGATA
A6VH.V93A.R
GTGTATTACTGTGCGAAAGACAGG
A6VH.E6A.R
CAATTACAAGCTAGTGGTGGC
A6VH.E6A.F
GCCACCACTAGCTTGTAAATTG
A6VH.T108Q.F
TATGGATCCTGAGGAGACGGTGACCTGTGTCCCTTGGCC
A6VH.Apall.R
CATGACCACAGTGACAGGAGGTCCAATTACAAGCTAGA
A6VH.NOT.T108Q
CGATTCTGCGGCCGCTGAGGAGACGGTGACCTGTGTCCCTTGGCCCCAGATATC
RP
GCGGATAACAATTTACACAGGAA
FP
CGCCAGGGTTTTCCAGTCACGAC

Clonación, expresión y evaluación de la formación de dímero/multímero

10 El A6V_H modificado, denominado A6V_H-L1 se clonó en un vector para la expresión en *E. coli* usando *Eco*RI y *Bam*H1. Se inocularon 30 ml de LB que contenían 100 µg/ml de ampicilina con una sola colonia que albergaba pSJF2-dAb y el cultivo se agitó a 240 rpm a 37 °C durante toda la noche. A la mañana siguiente todo el cultivo de la

15 noche anterior se usó para inocular 1 litro de medio M9 complementado con 5 µg/ml de vitamina B1, 0.4% de ácidos casamino y 100 µg/ml de ampicilina. El cultivo se agitó a temperatura ambiente durante 30 h a 160 rpm y a continuación se complementó con 100 ml de medio de inducción 10x y 100 µl de isopropiltio-β-D- galactósido 1 M. El cultivo se agitó durante otras 60 h, la fracción periplásmica se extrajo por el método de shock osmótico (Anand et al., 1991), y la presencia de dAb en el extracto se detectó mediante inmunotransferencia de tipo Western (MacKenzie 1994). La fracción periplásmica se dializó extensamente en tampón HEPES (N-[2-hidroxietil]pirazina -N' [2-ácido etanosulfónico]) 10 mM de pH 7.0 y NaCl 500 mM. La presencia de la etiqueta His₅ C-terminal de dAb permitió la

20 purificación de la proteína en un paso por cromatografía de afinidad con metal inmovilizado usando una columna HiTrap Chelating™ (Pharmacia). La columna de 5 ml se cargó con Ni²⁺ aplicando 30 ml de una solución de 5 mg/ml de NiCl₂.6H₂O y después se lavó con 15 ml de agua desionizada. La purificación se llevó a cabo como se describe (MacKenzie, 1994) excepto que el tampón de inicio fue tampón HEPES 10 mM, imidazol 10 mM, NaCl 500 mM, pH 7.0, y la proteína unida se eluyó con un gradiente de imidazol 10-500 mM. La pureza de la proteína se determinó

25

mediante SDS-PAGE (Laemmli). Para detectar la presencia de dímero/múltiplo de dAb en la preparación de proteína, se realizó una cromatografía de filtración en gel usando Superdex75 (Pharmacia) según se describe (Deng et al., 1995).

5 Construcción de la biblioteca

Se generó un producto de PCR usando A6V_H-L1 como molde y los cebadores RP y A6V_H.Rndm100i-n.F, CCTTGGCCCCAGATATCAA6[(A/C)NN]GTAATAACCACTACTATC. El último cebador se degeneró en la región que codifica los residuos 100i-n. Una segunda PCR empleo 180 pmol del producto anterior y 100 pmol de cada uno de los cebadores A6V_H.Apal.R, CATGACCACAGTGCACAGGAGGTCCAATTACAAGCTAG, y A6V_H.NotT108Q.F, CGATTCTGCGGCCGCTGAGGAGACGGTGACCTGTGTCCCTTGGCCCCAGATATC.

Los cebadores del segundo conjunto son complementarios de los extremos 5' y 3' de lo genes de dAb e incorporan los sitios de restricción *ApalI* y *NotI* (secuencias subrayadas) al final de los genes amplificados. Los productos amplificados se purificaron, se cortaron secuencialmente con las endonucleasas de restricción *ApalI* y *NotI*, se volvieron a purificar, y se ligaron al vector de expresión del fago fd-tet tratado con *ApalI/NotI*. Luego de esto, 1.5 µg del producto ligado desalinizado se mezclaron con 40 µl de la cepa TG1 de *E. coli* competente y las células se transformaron por electroporación. La determinación de la transformación, la amplificación y purificación de la biblioteca del fago y del tamaño de la biblioteca se realizó como se describe a continuación. La estrategia de aleatorización se describe en la figura 11.

Determinación del tamaño de la biblioteca.

Para determinar el tamaño de la biblioteca, inmediatamente a continuación de la transformación y después de la adición del medio SOC, una pequeña alícuota de las células electroporadas se diluyó serialmente en células TG1 en multiplicación exponencial. Doscientos µl de las células diluidas se mezclaron con 3 ml de top-agarosa a 50 °C y se vertieron inmediatamente en placas 2 x YT precalentadas a 37 °C. Las placas se incubaron durante toda la noche a 37 °C y la cantidad de placas se usó para determinar el tamaño de la biblioteca.

30 Selección por afinidad

La selección por afinidad (inmunoadsorción) se realizó usando tiras de 8 pocillos Nunc-Immuno MaxiSorp™ (Nunc). En resumen, los pocillos se recubrieron durante toda la noche agregando 150 µl de antígeno 100 µg/ml en PBS. A la mañana siguiente, se enjuagaron tres veces con PBS y a continuación se bloquearon con 400 µl de PBS al 2% (p/v) leche descremada (al 2% en MPBS) a 37 °C durante 2 h. Los pocillos se enjuagaron como antes y se agregaron 10¹² unidades de transducción del fago en MPBS al 2%. La mezcla se incubó a temperatura ambiente durante 1.5 h después de lo cual se eliminó el fago sin unirse del sobrenadante. Los pocillos se enjuagaron 10 veces con PBS-Tween 20 al 0,1% (v/v) y 10 veces con PBS para eliminar el detergente. El fago unido se eluyó agregando 200 µl de trietilamina 100 mM recién preparada, subiendo y bajando el contenido varias veces con una pipeta e incubando la mezcla a temperatura ambiente durante 10 min. El fago eluido se transfirió a un tubo que contenía 100 µl de Tris 1 M-HCl, pH 7.4 y se agitó con un equipo tipo vórtex para neutralizar la trietilamina. Luego de esto, se infectaron 10 ml de un cultivo de TG1 en multiplicación exponencial, con 150 µl del fago eluido incubando la mezcla a 37 °C durante 30 min. Se usaron diluciones seriadas de las células infectadas para determinar el título del fago eluido según se describe en la sección anterior. El resto de las células infectadas se voltearon y se resuspendieron en 900 µl de 2 x YT. Las células se mezclaron en alícuotas de 300 µl con 3 ml de top-agarosa y el fago se propagó en las placas durante toda la noche a 37 °C. A la mañana siguiente el fago se purificó, se le determinó el título, y un total de 10¹¹ unidades de transducción del fago se usaron para otras rondas de selección.

Ejemplo 2

También se construyó una biblioteca (A6V_H-L1a) en la cual se eliminaron los residuos 100i y 100n y se aleatorizaron los residuos 95-100h con 50% de sesgado a favor de la constitución aminoacídica parental. El procedimiento para la construcción de A6V_H-L1a fue idéntico al de la A6V_H-L1 excepto porque el cebador de aleatorización utilizado en el primer paso de la PCR fue A6V_H.95-100hRndm CCCTTGGCCCCAGATATCAA14 (A/C)NN]TTTCGCACAGTA ATACAC. La secuencia del A6V_H para esta biblioteca se muestra en la figura 12.

Ejemplo 3

También se construyó A6V_H-L2, una molécula de unión al ligando V_H parental en la cual cada una de las posiciones 6, 23, 74, 82a, 83, 84, 93 y 108 se modificó como se muestra en la figura 3 en comparación con la A6V_H nativa (Figura 1). Se introdujeron mutaciones en las posiciones 74, 83 y 84 con un único cebador mutagénico, A6V_H.S74A.F, (5'CATTGGAAG-ATACAGAGTGTCTTGGCATTGTCTCT3'), el cual cuando se usó junto con RP (véase el ejemplo 1 antes) generó un fragmento N-terminal. Un segundo cebador, A6V_H.R83K/A84P.R, (5'TATCTTCAAATGAACAGTCTGAAAC-CAGAGGACACGGCT3'), junto con FP (véase el ejemplo 1 antes) generó un fragmento C-terminal. Los dos fragmentos se unieron mediante SOE, usando RP y FP como cebadores, para dar

el gen que codifica a A6V_H-L2. Esta molécula mostró niveles ligeramente mejores de producción de monómero cuando se expresó y se evaluó por cromatografía de exclusión por tamaño en comparación con A6V_H-L1.

5 Tabla 7. Identidad de los aminoácidos en las posiciones 6, 23, 74, 82a, 83, 84, 93 y 108 para A6V_H y las mutantes de A6V_H-3

Identidad de los aminoácidos								
Kabat #	6	23	74	82a	83	84	93	108
A6VH	E	S	S	S	R	A	V	T
A6VH-L1	A	A	S	N	R	A	A	Q
A6VH-L2	A	A	A	N	K	P	A	Q

Ejemplo 4 - Resonancia de plasmón superficial

10 Los estudios de unión se realizaron usando BIACORE Upgrade (Biacore Inc., Piscataway, NJ) según se describe (Jönsson et al., 1991). Se inmovilizaron aproximadamente 14 000 RU de anti-FLAG M2 IgG o IgG de control en chips sensores CM5 por acoplamiento con amina. Se pasaron anticuerpos de un solo dominio sobre las superficies de los chips sensores en tampón HEPES 10 mM de pH 7.4, NaCl 150 mM, EDTA 3.4 mM, P-20 al 0.005% (Biacore Inc.) a 25 °C y a una velocidad de flujo de 5 µl/min. Los datos del sensorgrama se analizaron usando el paquete informático BIAevaluation 3.0 (Biacore Inc.).

Ejemplo 5 - Enzimoinmunoanálisis de absorción (ELISA)

20 Se recubrieron placas Nunc-Immuno MaxiSorp™ (Nunc) durante toda la noche a 4 °C con 150 µl de 3B1 scFv o BSA in PBS 10 µg/ml. El contenido se eliminó y las placas se golpearon sobre una toalla de papel para eliminar todo el líquido remanente en los pocillos. Los pocillos se bloquearon agregando 300 µl de MPBS al 2% e incubándolos durante 2 h a 37 °C. El contenido de los pocillos se vació como antes, se agregaron 100 µl de dAb en fago purificado en MPBS al 2%, y los pocillos se incubaron a temperatura ambiente durante 1.5 h. El contenido se volvió a vaciar y los pocillos se lavaron 5 veces con PBS-Tween 20 al 0.05% (v/v) y a continuación se secaron con una toalla de papel para eliminar cualquier resto de tampón de lavado. Se agregaron 100 µl de la disolución recomendada de HRP/conjugado de anticuerpo monoclonal Anti-M13 (Amersham Pharmacia Biotech) en MPBS al 2% y los pocillos se incubaron a temperatura ambiente durante 1 h. Los pocillos se lavaron seis veces como antes y la unión de dAb al antígeno se detectó coloriméricamente agregando 100 µl de mezclas iguales de sustrato de peroxidasa TMB y H₂O₂ (Kirkegaard and Perry Laboratories, Gaithersburg, MD, USA) a temperatura ambiente por varios minutos. La reacción se detuvo agregando 100 µl de H₃PO₄ 1 M y se midió la A₄₅₀ con un lector de placas DYNATECH MR5000 (Dynatech Laboratories, Chantilly, VA, Estados Unidos).

Ejemplo 6 - Introducción de una variación genética en la secuencia correspondiente a la región CDR3 de la cadena pesada de A6 - Residuos aleatorizados

35 Se prepararon oligonucleótidos que contenían regiones CDR3 mutadas aleatoriamente en un sintetizador de ADN Applied Biosystems 394 según se describe antes.

1. Producción de 23 residuos aleatorizados (CDR3 1-23):

40 Se usa la fórmula del anticodón [(A/C)NN] que produce una reducción en el uso posible de codones de 64 a 32 y reduce la cantidad de codones de terminación posibles. Por consiguiente, la posición uno comprende sólo A y C en la mezcla de reacción sintética. Para la aleatorización completa de las segunda y tercera posiciones de los codones la mezcla de dNTP comprende 25% de cada una de A,G,C y T.

45 El cebador de aleatorización oligonucleótido 3' se diseñó de modo que los últimos 15 nucleótidos del marco 3 y los primeros 17 nucleótidos del marco 4 se mantuvieron constantes para la hibridación. Los nucleótidos que codifican los aminoácidos intervinientes, a saber los aminoácidos 1-23 de la región CDR3 se aleatorizaron usando el cebador siguiente: 5' (GTTGTCCCTTGCCCCA n[(A/C)NN]TTTCACACAGTAATA) 3' (Donde n = 23, hebra antisentido).

50 Usando 50% de A y 50% de C para la posición del primer nucleótido para cada triplete anticodón y 25% de cada una de A, C, G y T para la segunda y tercera posiciones de nucleótidos para n = 23, se logró la aleatorización completa de los 23 aminoácidos de la CDR3 de A6.

2. Síntesis de CDRs que contienen 15 a 23 residuos

55 Los cebadores se adaptan reduciendo n a 15-23 en la fórmula del cebador anterior manteniendo simultáneamente constantes los nucleótidos flanqueantes.

3. Para la síntesis de CDR3s que contienen 24-33 residuos

5 Los cebadores se adaptarán para aumentar n a 24-23 en la fórmula del cebador anterior manteniendo simultáneamente constantes los nucleótidos flanqueantes.

Ejemplo 7 - Sesgado de aleatorización selectiva para 50% de homología con la tirosina parental

10 Para alcanzar aproximadamente 50% de homología con el tipo silvestre en cualquier posición en la región A6 dAb CDR3 durante la síntesis antisentido usando el sintetizador de ADN, se usará el ejemplo siguiente. En el caso de la tirosina, que es codificada por TAC o TAT (hebra antisentido GTA o ATA) se agregarán cantidades conocidas de los nucleótidos de la manera siguiente para la hebra antisentido.

15 Posición del nucleótido del primer anticodón: Se agregan 80% de A y 20% de C a la solución de dNTP, y no se agregan G ni T para reducir la degeneración del codón.
 Posición del nucleótido del segundo anticodón: 80% de T y aproximadamente 6.67% de C, 6.67% de A y 6.67% de G.
 Posición del nucleótido del tercer anticodón la mezcla: 80% de A y aproximadamente 6.67% de T y 6.67% de G y 6.7% de C.

20 La probabilidad calculada de la tirosina sería por lo tanto $0.8 \times 0.8 \times 0.8 \times 100\% = 51.2\%$. Por consiguiente aproximadamente 51% de las cadenas de la biblioteca contendrán tirosina de A6 tipo silvestre en la posición especificada.

25 Ejemplo 8 - Sesgado de aleatorización selectiva para 50% de homología con la serina parental

Usando la misma estrategia para alcanzar aproximadamente 50% de homología con la serina tipo silvestre en una o más posiciones, el ejemplo siguiente es útil.

30 Usando sólo A y/o C en la posición del primer anticodón el aminoácido serina podría tener dos codones que son AGT, TCT y TCG (ACT, AGA y CGA antisentido, respectivamente). Los niveles de agregado de cantidades conocidas de nucleótido serían los siguientes:

35 Posición del nucleótido del primer anticodón: 50% de A y 50% de C.
 Posición del nucleótido del segundo anticodón: 35.35% de C, 35.35% de G, 14.65% de A y 14.65% de T
 Posición del nucleótido del tercer anticodón: 35.35% de A, 35.35% de T, 14.65% de C y 14.65% de G.

La probabilidad de producir serina para cualquier fragmento dado, usando esta estrategia es $(1 \times [0.3535 + 0.3535] \times [0.3535 + 0.3535] \times 100\% = 50\%$. Por lo tanto, aproximadamente 50% de las cadenas tendrán una serina en la posición seleccionada.

40 Ejemplo 9 - Sesgado de aleatorización selectiva para 50% de homología con la serina parental

45 Para alcanzar aproximadamente 10% de homología con el tipo silvestre en cualquier posición en la región A6 dAb CDR3 durante la síntesis antisentido usando el sintetizador de ADN, se puede usar el ejemplo siguiente. En el caso de la tirosina, que es codificada por TAC o TAT (hebra antisentido GTA o ATA) se agregarán cantidades conocidas de los nucleótidos de la manera siguiente para la hebra antisentido.

50 Posición del nucleótido del primer anticodón: Se agregan 47% de A y 53% de C; no se agregan G y T para reducir la degeneración del codón.
 Posición del nucleótido del segundo anticodón: 47% de T y aproximadamente 17.67% de C, 17.67% de A y 17.67% de G.
 Posición del nucleótido del tercer anticodón: 47% de A y aproximadamente 17.67% de T y 17.67% de G y 17.67% de C.

55 La probabilidad calculada de la tirosina es por lo tanto $0.47 \times 0.47 \times 0.47 \times 100\% = 10.4\%$. Por consiguiente aproximadamente 10% de las cadenas de la biblioteca contendrán tirosina de A6 tipo silvestre en la posición especificada.

Ejemplo 10 - Sesgado de aleatorización selectiva para 50% de homología con la serina parental

60 Para alcanzar aproximadamente 90% de homología con los aminoácidos tipo silvestre en cualquier posición en la región A6 dAb CDR3 durante la síntesis antisentido usando el sintetizador de ADN, se usará el ejemplo siguiente. En el caso de la tirosina, que es codificada por TAC o TAT (hebra antisentido GTA o ATA) se agregarán cantidades conocidas de los nucleótidos de la manera siguiente:

65 Posición del nucleótido del primer anticodón: Se agregan 97% de A y 3% de C; no se agregan G y T para reducir la degeneración del codón. Por esta razón, sólo se usan A y C en la posición del primer anticodón para los 20 aminoácidos que existen naturalmente.

ES 2 376 454 T3

Posición del nucleótido del segundo anticodón: 97% de T y aproximadamente 1% de C, 1% de A y 1% de G.

Posición del nucleótido del tercer anticodón: 97% de A y aproximadamente 1% de T y 1% de G y 1% de C.

- 5 La probabilidad calculada de la tirosina será por lo tanto $0.97 \times 0.97 \times 0.97 \times 100\% = 91.3\%$. Por consiguiente aproximadamente 90% de las cadenas de la biblioteca contendrán tirosina de A6 tipo silvestre en la posición especificada.
- 10 Se pueden calcular y aplicar al sintetizador de ADN, usando los métodos de los ejemplos anteriores, de aproximadamente 10% a aproximadamente 90% de representación de aminoácidos tipo silvestre en uno o más residuos de aminoácidos especificados en la CDR3 de A6.

Bibliografía adicional

- 5 Anand, N. N., Dubuc, G., Phipps, J., MacKenzie, C. R., Sadowska, J., Young, N. M., Bundle, D. R., and Narang, S. A. (1991). Synthesis and expression in *Escherichia coli* of cistronic DNA encoding an antibody fragment specific for a *Salmonella* serotype B O-antigen. *Gene* 100:39-44.
- Arbabi, G. M., Desmyter, A., Wyns, L., Hamers, R., and Muyldermans, S. (1997). Selection and identification of single domain antibody fragments from camel heavy-chain antibodies. *FEBS Lett* 414:521-526.
- 10 Bodenhausen, G. and Ruben, D. J., *Chem. Phys. Lett.*, 69, 185-188, 1980.
- Cai, X. and Garen, A. (1996). A melanoma-specific VH antibody cloned from a fusion phage library of a vaccinated melanoma patient. *Proc Natl Acad Sci USA* 93:6280-6285.
- 15 Chothia, C and Lesk, A. M. (1987). Canonical structures for the hypervariable regions of immunoglobulins. *JMol Biol* 196:901-917.
- Chothia, C., Novotny, J., Brucoleri, R., and Karplus, M. (1985). Domain association in immunoglobulin molecules. The packing of variable domains. *JMol Biol* 186:651-663.
- 20 Clackson, T., Hoogenboom, H. R., Griffiths, A. D., and Winter, G. (1991). Making antibody fragments using phage display libraries. *Nature* 352:624-628.
- Constantine, K. L., Goldfarb, V., Wittekind, M., Anthony, J., Ng, S. C., and Mueller, L. (1992). Sequential ¹H and ¹⁵N NMR assignments and secondary structure of a recombinant anti-digoxin antibody VL domain. *Biochemistry* 31:5033-5043.
- 25 Constantine, K. L., Goldfarb, V., Wittekind, M., Friedrichs, M. S., Anthony, J., Ng, S. C., and Mueller, L. (1993). Aliphatic ¹H and ¹³C resonance assignments for the 26-10 antibody VL domain derived from heteronuclear multidimensional NMR spectroscopy. *J Biomol NMR* 3:41-54.
- 30 Dan, M., Earley, E. M., Griffin, M. C., Maiti, P. K., Prashar, A. K., Yuan, X. Y., Friesen, A. D., and Kaplan, H. As 1995. Human monoclonal antibody BT32/A6 and a cell cycle-independent glioma-associated surface antigen. *J. Neurosurg.* 82, 475-480.
- 35 Davies, J. and Riechmann, L. (1994). 'Camelising' human antibody fragments: NMR studies on VH domains. *FEBS Lett* 339:285-290.
- Davies, J. and Riechmann, L. (1995). Antibody VH domains as small recognition units. *Biotechnology N Y* 13:475-479.
- 40 Davies, J. and Riechmann, L. (1996b). Affinity improvement of single antibody VH domains: residues in all three hypervariable regions affect antigen binding. *Immunotechnology* 2:169-179.
- 45 Davies, J. and Riechmann, L. (1996a). Single antibody domains as small recognition units: design and in vitro antigen selection of camelized, human VH domains with improved protein stability. *Protein Eng* 9:531-537.
- Decanniere, K., Desmyter, A., Lauwereys, M., Ghahroudi, M. A., Muyldermans, S., and Wyns, L. (1999). A single-domain antibody fragment in complex with RNase A: non-canonical loop structures and nanomolar affinity using two CDR loops. *Structure* 7:361-370.
- 50 Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. and Bax, A., *J. Biomol. NMR* 6, 277-293, 1995.
- Deng, S. J., MacKenzie, C. R., Hirama, T., Brousseau, R., Lowary, T. L., Young, N. M., Bundle, D. R., and Narang, S. A. (1995). Basis for selection of improved carbohydrate-binding single-chain antibodies from synthetic gene libraries. *Proc Natl Acad Sci U S A* 92:4992-4996.
- 55 Deng, S. J., MacKenzie, C. R., Sadowska, J., Michniewicz, J., Young, N. M., Bundle, D. R., and Narang, S. A. (1994). Selection of antibody single-chain variable fragments with improved carbohydrate binding by phage display. *J Biol Chem* 269:9533-9538.
- 60 Desmyter, A., Transue, T. R., Ghahroudi, M. A., Thi, M. H., Poortmans, F., Hamers, R., Muyldermans, S., and Wyns, L. (1996). Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme [see comments]. *Nat Struct Biol* 3:803-811.
- 65

- Hamers, C. C., Atarhouch, T., Muyldermans, S., Robinson, G., Hamers, C., Songa, E. B., Bendahman, N., and Hamers, R. (1993). Naturally occurring antibodies devoid of light chains. *Nature* 363:446-448.
- 5 Harrison, J. L., Williams, S. C., Winter, G., and Nissim, A. (1996). Screening of phage antibody libraries. *Methods Enzymol* 267:83-109.
- Johnson, B. A. and Blevins, R. A., *J. Chem., Phys.*, 29, 1012-1014, 1994.
- 10 Jönsson, U., Fägerstam, L., Ivarsson, B., Johnsson, B., Karlsson, R., Lundh, K., Löfås, S., Persson, B., Roos, H., Rönnberg, I., Sjölander, S., Stenberg, E., Stahlberg, R., Urbaniczky, C., Östlin, H., and Malmqvist, M. (1991). Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology. *BioTechniques* 11, 620-627.
- 15 Knappik, A. and Pluckthun, A. (1994). An improved affinity tag based on the FLAG peptide for the detection and purification of recombinant antibody fragments. *BioTechniques* 17:754-761.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680-685.
- 20 Lauwereys, M., Arbabi, G. M., Desmyter, A., Kinne, J., Holzer, W., De Genst, E., Wyns, L., and Muyldermans, S. (1998). Potent enzyme inhibitors derived from dromedary heavy-chain antibodies. *EMBO J* 17:3512-3520.
- Lowman, H. B., Bass, S. H., Simpson, N., and Wells, J. A. (1991). Selecting high-affinity binding proteins by monovalent phage display. *Biochemistry* 30:10832-10838.
- 25 MacKenzie, R. and To, R. (1998). The role of valency in the selection of anti-carbohydrate single-chain Fvs from phage display libraries. *J. Immunol. Methods* 220, 39-49.
- 30 MacKenzie, C. R., Sharma, V., Brummell, D., Bilous, D., Dubuc, G., Sadowska, J., Young, N. M., Bundle, D. R., and Narang, S. A. (1994). Effect of C lambda-C kappa domain switching on Fab activity and yield in *Escherichia coli*: synthesis and expression of genes encoding two anti-carbohydrate Fabs. *Biotechnology N Y* 12:390-395.
- 35 MacKenzie, C.R., Hiramata, T., Deng, S.-J., Bundle, D.R., Narang, S.A., and Young, N.M. (1996). Analysis by surface plasmon resonance of the influence of valence on the ligand-binding affinity and kinetics of an anti-carbohydrate antibody. *J. Biol. Chem.* 271, 1527-1533
- McAfferty, J., Griffiths, A.D., Winter, G. and Chiswell, D.J. (1990). *Nature* 348, 552-554. Messing, J. (1983). New M13 vectors for cloning. *Methods in Enzymology* 101, 20-78.
- 40 Miceli, R. M., Degraaf, M. E., and Fischer, H. D. (1994). Two-stage selection of sequences from a random phage display library delineates both core residues and permitted structural range within an epitope. *J Immunol Methods* 167:279-287.
- 45 Morea, V., Tramontano, A., Rustici, M., Chothia, C., and Lesk, A. M. (1998). Conformations of the third hypervariable region in the VH domain of immunoglobulins. *J Mol Biol* 275:269-294.
- Muyldermans, S., Atarhouch, T., Saldanha, J., Barbosa, J. A., and Hamers, R. (1994). Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains. *Protein Eng* 7:1129-1135.
- 50 Narang, S. A., Yao, F. L., MichnieWics, J. J., Dubuc, G., Phipps, J., and Somorjai, R. L. (1987) *Protein Eng.* 1, 481-485.
- Narang, S. A., Brummell, D. A., Sharma, V., Dubuc, G., MacKenzie, C. R., MichnieWics, J. J., Sadowska, J., Anand, N., Young, N. M., and Bundle, D.R. (1991) *Nucleic Acids Res. Symp. Ser.* 24, 173-179.
- 55 Nissim, A., Hoogenboom, H. R., Tomlinson, I. M., Flynn, G., Midgley, C., Lane, D., and Winter, G. (1994). Antibody fragments from a 'single pot' phage display library as immunochemical reagents. *EMBO J* 13:692-698.
- 60 Padlan, E. A. (1994). Anatomy of the antibody molecule. *Mol Immunol* 31:169-217.
- Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR* 2, 661-665.
- 65 Reiter, Y., Shuck, P., Boyd, L. F., and Plaxin, D. (1999). An antibody single-domain phage display library of a native heavy chain variable region: isolation of functional single-domain VH molecules with a unique interface. *J Mol Biol* 290:685-698.

- Sambrook, J. Fritsch E. F. and Maniatis T (1989). "Molecular Cloning: A laboratory Manual (2nd ed.)," Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- 5 Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74:5463-5467.
- Sheriff, S. and Constantine, K. L. (1996). Redefining the minimal antigen-binding fragment [news; comment]. *Nat Struct Biol* 3:733-736.
- 10 Sklenar, V., Piotto, M., Leppik, R., and Saudek, V. (1993) *J Magn. Reson.* A102, 241-245.
- Slatter, M., Schleucher, J. and Griesinger, C., *Progr. NMR Spectrosc.*, 1999, 34, 93-158.
- 15 Spinelli, S., Frenken, L., Bourgeois, D., de Ron, L., Bos, W., Verrips, T., Anguille, C., Cambillau, C., and Tegoni, M. (1996). The crystal structure of a llama heavy chain variable domain [letter] [see comments]. *Nat Struct Biol* 3:752-757.
- Tung, W. L. and Chow, K. C. (1995). A modified medium for efficient electotransformation of *E. coli*. *Trends Genet* 11:128-129.
- 20 Vierra, O. and Messing, J. (1982). The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* 19, 259-268.
- 25 Vu, K. B., Ghahroudi, M. A., Wyns, L., and Muyldermans, S. (1997). Comparison of llama VH sequences from conventional and heavy chain antibodies. *Mol Immunol* 34:1121-1131.
- Ward, E. S., Gussow, D., Griffiths, A. D., Jones, P. T., and Winter, G. (1989). Binding activities of a repertoire of single immunoglobulin variable domains secreted from *Escherichia coli* [see comments]. *Nature* 341:544-546.
- 30 Wishart, D. S. and Sykes, B. D., *J Biomol NMR*, 4, 171-180, 1994.
- Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L. and Sykes, B. D., *J. Biomol. NMR*, 6 135-140, 1995.
- 35 Zacher, A.N., Stock, C.A., Golden, J.W. and Smith, G.P. (1980). *Gene* 9, 127-140.

LISTADO DE SECUENCIAS

- 5 <110> Kaplan, Howard
 Entwistle, Jocelyn
 Tanha, Jamshid
 Narang, Saran
 Dan, Michael
 Mackenzie, Colin R
- 10 <120> Bibliotecas de expresión en fago de fragmentos VH humanos
 <130> 33956-0043
- 15 <140> PCT/CA01/
 <141> 2001-12-21
 <150> US 60/258,031
 <151> 2000-12-22
- 20 <160> 92
 <170> PatentIn Ver. 2.1
- 25 <210> 1
 <211> 6
 <212> PRT
 <213> Secuencia artificial
- 30 <220>
 <223> Descripción de la secuencia artificial: región aleatorizada
 <400> 1
 Asp Gln Phe Thr His Ser
 1 5
- 35 <210> 2
 <211> 6
 <212> PRT
 <213> Secuencia artificial
- 40 <220>
 <223> Descripción de la secuencia artificial: región aleatorizada
 <400> 2
 Ser Ser Met Tyr Gly Asn
 1 5
- 45 <210> 3
 <211> 6
 <212> PRT
 <213> Secuencia artificial
- 50 <220>
 <223> Descripción de la secuencia artificial: región aleatorizada
 <400> 3
 Ile Lys Met Gln Gln Asn
 1 5
- 55 <210> 4
 <211> 6
 <212> PRT
 <213> Secuencia artificial
- 60 <220>
 <223> Descripción de la secuencia artificial: región aleatorizada

Ser Val Asp Ala Arg Asp
 1 5
 <400> 4
 <210> 5
 5 <211> 6
 <212> PRT
 <213> Secuencia artificial
 <220>
 10 <223> Descripción de la secuencia artificial: región aleatorizada
 <400> 5
 Val Ser Arg Phe Gly Ala
 1 5
 <210> 6
 15 <211> 6
 <212> PRT
 <213> Secuencia artificial
 <220>
 20 <223> Descripción de la secuencia artificial: región aleatorizada
 Gly Leu Gly Ser Pro Lys
 1 5
 <400> 6
 <210> 7
 25 <211> 6
 <212> PRT
 <213> Secuencia artificial
 <220>
 30 <223> Descripción de la secuencia artificial: región aleatorizada
 Ile Asp Ala Lys Trp Ala
 1 5
 <400> 7
 <210> 8
 <211> 6
 35 <212> PRT
 <213> Secuencia artificial
 <220>
 40 <223> Descripción de la secuencia artificial: región aleatorizada
 <400> 8
 Val Ser Arg Phe Gly Ala
 1 5
 <210> 9
 <211> 6
 45 <212> PRT
 <213> Secuencia artificial
 <220>
 50 <223> Descripción de la secuencia artificial: región aleatorizada
 <400> 9
 His Cys Leu Pro Asp Gly
 1 5
 <210> 10
 <211> 6
 55 <212> PRT<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: región aleatorizada

0 Arg Trp Arg Val Pro
1r r r a 5r

5 <400> 10

<210> 11

<211> 6

<212> PRT

<213> Secuencia artificial

10

<220>

<223> Descripción de la secuencia artificial: región aleatorizada

Val Ser Arg Phe Gly Ala
1 5

<400> 11

15

<210> 12

<211> 6

<212> PRT

<213> Secuencia artificial

20

<220>

<223> Descripción de la secuencia artificial: región aleatorizada

Leu Glu Cys Glu Gly Cys
1 5

<400> 12

25

<210> 13

<211> 6

<212> PRT

<213> Secuencia artificial

30

<220>

<223> Descripción de la secuencia artificial: región aleatorizada

Arg Asn Val Gly Ala Leu
1 5

<400> 13

35

<210> 14

<211> 6

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: región aleatorizada

<400> 14

Arg Arg Ser Asp Tyr Leu
1 5

45

<210> 15

<211> 6

<212>PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: región aleatorizada

Cys Val Arg Gly Ala Glu
1 5

<400> 15

55

<210> 16

<211> 6

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: región aleatorizada

Ser Pro Ser Leu Ala Ala
1 5

5 <400> 16

<210> 17

<211> 6

<212> PRT

<213> Secuencia artificial

10

<220>

<223> Descripción de la secuencia artificial: región aleatorizada

His Ala Ser Gly Arg Ser
1 5

<400> 17

15 <210> 18

<211> 6

<212> PRT<213> Secuencia artificial

<220>

20 <223> Descripción de la secuencia artificial: región aleatorizada

Gly Tyr Met Cys Ser Leu
1 5

<400> 18

<210> 19

<211> 6

25 <212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: región aleatorizada

30

His Asn Lys Asp Leu Ala
1 5

<400> 19

<210> 20

<211> 6

<212> PRT

35 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: región aleatorizada

40 400>20

Leu Ala Asp Leu Tyr Met
1 5

<210> 21

<211> 6

45 <212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: región aleatorizada

50

Trp Arg Arg Ala His Glu
1 5

<400> 21

<210> 22

<211> 6

<212> PRT

55 <213> Secuencia artificial

ES 2 376 454 T3

<220>
 <223> Descripción de la secuencia artificial: motivo

5 <400> 28
 Xaa Tyr Lys Xaa Xaa Asp
 1 5

<210> 29
 <211> 31
 <212> PRT
 10 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: ligadores de M-2

15 <400> 29
 Glu Val Gln Leu Gln Ala Ser Ser Gly Tyr Tyr Glu Asp Asp Tyr Arg
 1 5 10 15
 Leu Phe Asp Ile Trp Gly Gln Gly Thr Gln Val Thr Val Ser Ser
 20 25 30

<210> 30
 <211> 43
 <212> PRT
 20 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: ligadores de M-2

25 <400> 30
 Glu Val Gln Leu Gln Ala Ser Gly Gly Gly Leu Val Gln Pro Gly Tyr
 1 5 10 15
 Tyr Asp Ser Ser Gly Tyr Tyr Lys Asp Leu Asp Thr Arg Phe Asp Ile
 20 25 30
 Trp Gly Gln Gly Thr Gln Val Thr Val Ser Ser
 35 40

<210> 31
 <211> 53
 <212> PRT
 30 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: ligadores de M-2

35 <400> 31
 Glu Val Gln Leu Gln Ala Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
 1 5 10 15
 Ser Leu Arg Leu Ser Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Tyr
 20 25 30
 Tyr Gly Asp His Tyr Lys Trp Phe Asp Ile Trp Gly Gln Gly Thr Gln
 35 40 45
 Val Thr Val Ser Ser
 50

<210> 32
 <211> 23
 <212> PRT
 40 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: ligadores de M-2

ES 2 376 454 T3

<400> 32

Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Arg Asn
 1 5 10 15

Glu Tyr Lys Glu Phe Asp Ile
 20

<210> 33

5 <211> 23

<212> PRT

<213> Secuencia artificial

<220>

10 <223> Descripción de la secuencia artificial: ligadores de M-2

<400> 33

Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Val Asp
 1 5 10 15

Glu Tyr Lys Ser Phe Asp Ile
 20

<210> 34

15 <211> 23

<212> PRT

<213> Secuencia artificial

<220>

20 <223> Descripción de la secuencia artificial: ligadores de M-2

<400> 34

Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Ala Gly
 1 5 10 15

Arg Tyr Lys Asp Phe Asp Ile
 20

<210> 35

25 <211> 23

<212> PRT

<213> Secuencia artificial

<220>

30 <223> Descripción de la secuencia artificial: ligadores de M-2

<400> 35

Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Arg Ser
 1 5 10 15

Asp Tyr Lys Arg Phe Asp Ile
 20

<210> 36

35 <211> 23

<212> PRT

<213> Secuencia artificial

<220>

40 <223> Descripción de la secuencia artificial: ligadores de M-2

<400> 36

Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Ala Ser
 1 5 10 15

His Tyr Lys Asp Phe Asp Ile
 20

ES 2 376 454 T3

<210> 37
<211> 23
<212> PRT
<213> Secuencia artificial

5
<220>
<223> Descripción de la secuencia artificial: ligadores de M-2

<400> 37
Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Val Asp
1 5 10 15
Gly Tyr Lys Asp Phe Asp Ile
20

10
<210> 38
<211> 23
<212> PRT
<213> Secuencia artificial

15
<220>
<223> Descripción de la secuencia artificial: ligadores de M-2

<400> 38
Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Thr Ala
1 5 10 15
Asp Tyr Lys Met Phe Asp Ile
20

20
<210> 39
<211> 23
<212> PRT
<213> Secuencia artificial

25
<220>
<223> Descripción de la secuencia artificial: ligadores de M-2

<400> 39
Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Asp Met
1 5 10 15
Asp Tyr Lys Thr Phe Asp Ile
20

30
<210> 40
<211> 22
<212> PRT
<213> Secuencia artificial

35
<220>
<223> Descripción de la secuencia artificial: ligadores de M-2

<400> 40
Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Lys Ser
1 5 10 15
Tyr Lys Ser Phe Asp Ile
20

40
<210> 41
<211> 23
<212> PRT
<213> Secuencia artificial

45
<220>
<223> Descripción de la secuencia artificial: ligadores de M-2

<400> 41

ES 2 376 454 T3

Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Asp Tyr
 1 5 10 15
 Lys Ser Gln Asp Phe Asp Ile
 20
 <210> 42
 <211> 23
 <212> PRT
 5 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: ligadores de M-2
 10 <400> 42
 Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Asp Tyr
 1 5 10 15
 Lys Ser Gln Asp Phe Asp Ile
 20
 <210> 43
 <211> 23
 <212> PRT
 15 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: ligadores de M-2
 20 <400> 43
 Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Lys Asp
 1 5 10 15
 Trp Asp Ser Ser Phe Asp Ile
 20
 <210> 44
 <211> 23
 25 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: ligadores de M-2
 30 <400> 44
 Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Lys Asp
 1 5 10 15
 Gly Asp Ser Phe Phe Asp Ile
 20
 <210> 45
 <211> 21
 35 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: secuencias de CDR3
 40 de los ligadores de H11
 <400> 45
 Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Ser Phe Ser
 1 5 10 15
 Ser Pro Phe Asp Ile
 20
 <210> 46
 45 <211> 23

<212> PRT
 <213> Secuencia artificial

<220>
 5 <223> Descripción de la secuencia artificial: secuencias de CDR3
 de los ligadores de H11

<400> 46
 Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Tyr Asp
 1 5 10 15
 Phe Ser Ser Pro Phe Asp Ile
 20

10 <210> 47
 <211> 23
 <212> PRT
 <213> Secuencia artificial

15 <220>
 <223> Descripción de la secuencia artificial: secuencias de CDR3
 de los ligadores de H11

<400> 47
 Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Asn Leu
 1 5 10 15
 Phe Ser Ser Pro Phe Asp Ile
 20

20 <210> 48
 <211> 23
 <212> PRT
 <213> Secuencia artificial

25 <220>
 <223> Descripción de la secuencia artificial: secuencias de CDR3
 de los ligadores de H11

30 <400> 48
 Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Ser Glu
 1 5 10 15
 Phe Ser Ser Pro Phe Asp Ile
 20

<210> 49
 <211> 23
 <212> PRT
 35 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: secuencias de CDR3
 de los ligadores de H11

40 <400> 49
 Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Ser Asp
 1 5 10 15
 Phe Ser Ser Pro Phe Asp Ile
 20

<210> 50
 <211> 23
 <212> PRT
 45 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: secuencias de CDR3

ES 2 376 454 T3

de los ligadores de H11

<400> 50

Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Thr Asp
1 5 10 15

Met Ser Trp Glu Phe Asp Ile
20

5 <210> 51
<211> 23
<212> PRT
<213> Secuencia artificial

10 <220>
<223> Descripción de la secuencia artificial: secuencias de CDR3
de los ligadores de H11

<400> 51

Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Asp Leu
1 5 10 15

Gly Ser Trp Glu Phe Asp Ile
20

15 <210> 52
<211> 23
<212> PRT
<213> Secuencia artificial

20 <220>
<223> Descripción de la secuencia artificial: secuencias de CDR3
de los ligadores de H11

25 <400> 52

Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Asp Tyr
1 5 10 15

Val Ser Trp Glu Phe Asp Ile
20

<210> 53
<211> 23
<212> PRT
30 <213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: secuencias de CDR3
de los ligadores de H11

35 <400> 53

Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Asp Trp
1 5 10 15

Gly Ser Trp Thr Phe Asp Ile
20

<210> 54
<211> 22
40 <212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: secuencias de CDR3
de los ligadores de H11

45 <400> 54

ES 2 376 454 T3

Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Asp Gly
 1 5 10 15
 Thr Trp Asp Phe Asp Ile
 20
 <210> 55
 <211> 23
 <212> PRT
 5 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: secuencias de CDR3
 de los ligadores de H11
 10
 <400> 55
 Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Trp Glu
 1 5 10 15
 Gly Ser Gly Leu Phe Asp Ile
 20
 <210> 56
 <211> 23
 15 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: secuencias de CDR3
 de los ligadores de H11
 20
 <400> 56
 Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Ile Trp
 1 5 10 15
 Tyr Ser Gly Leu Phe Asp Ile
 20
 <210> 57
 <211> 23
 25 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: secuencias de CDR3
 de los ligadores de H11
 30
 <400> 57
 Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr Ser Ser
 1 5 10 15
 Trp Ala Ser Ala Phe Asp Ile
 20
 <210> 58
 <211> 6
 35 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: dAb designado
 Yst9.1-L3-9
 40
 <400> 58
 Val Ser Arg Phe Gly Ala
 1 5
 45 <210> 59
 <211> 54

<212> ADN
 <213> Secuencia artificial

 <220>
 5 <223> Descripción de la secuencia artificial: cebador

 <400> 59
 caattacaag aaagtgggtg cggactggtg caaccaggag gttccctgag actc 54

 10 <210> 60
 <211> 30
 <212> ADN
 <213> Secuencia artificial

 15 <220>
 <223> Descripción de la secuencia artificial: cebador

 <400> 60
 20 actttctgt aattggacct cggcctgcgc 30

 <210> 61
 <211> 21
 <212> ADN
 <213> Secuencia artificial

 25 <220>
 <223> Descripción de la secuencia artificial: cebador

 <400> 61
 30 ctctcctgtg ctgcctctgg a 21

 <210> 62
 <211> 21
 <212> ADN
 35 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: cebador

 <400> 62
 40 tccagaggca gcacaggaga g 21

 <210> 63
 <211> 54
 45 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: cebador

 <400> 63
 50 cgcacagtaa tacacagccg tgtcctcagc tctcagactg ttcattgaa gata 54

 <210> 64
 <211> 24
 <212> ADN
 <213> Secuencia artificial

 <220>
 60 <223> Descripción de la secuencia artificial: cebador

 <400> 64
 gtgtattact gtgcgaaaga cagg 24

 65 <210> 65
 <211> 21

<212> ADN
 <213> Secuencia artificial

 <220>
 5 <223> Descripción de la secuencia artificial: cebador

 <400> 65
 caattacaag ctagtgtgg c 21

 10 <210> 66
 <211> 21
 <212> ADN
 <213> Secuencia artificial

 15 <220>
 <223> Descripción de la secuencia artificial: cebador

 <400> 66
 gccaccacta gcttgaatt g 21
 20
 <210> 67Ç
 <211> 39
 <212> ADN
 <213> Secuencia artificial
 25
 <220>
 <223> Descripción de la secuencia artificial: cebador

 <400> 67
 30 tatggatcct gaggagacgg tgacctgtgt cccttgcc 39

 <210> 68
 <211> 39
 <212> ADN
 35 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: cebador

 <400> 68
 40 catgaccaca gtgcacagga ggtccaatta caagctaga 39

 <210> 69
 <211> 54
 45 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: cebador
 50
 <400> 69
 cgattctgcg gccgctgagg agacggtgac ctgtgtcct tggccccaga tatc 54

 <210> 70
 55 <211> 24
 <212> ADN
 <213> Secuencia artificial

 <220>
 60 <223> Descripción de la secuencia artificial: cebador

 <400> 70
 gcggataaca atttcacaca ggaa 24

 65 <210> 71
 <211> 24

<212> ADN
 <213> Secuencia artificial

 <220>
 5 <223> Descripción de la secuencia artificial: cebador

 <400> 71
 cgccagggtt ttcccagtca cgac 24

 10 <210> 72
 <211> 42
 <212> ADN
 <213> Secuencia artificial

 15 <220>
 <223> Descripción de la secuencia artificial: cebador

 <400> 72
 20 ccttggcccc agatatcaaa acnngtaata accactacta tc 42

 <210> 73
 <211> 38
 <212> ADN
 <213> Secuencia artificial

 25 <220>
 <223> Descripción de la secuencia artificial: cebador

 <400> 73
 30 catgaccaca gtgcacagga ggtccaatta caagctag 38

 <210> 74
 <211> 54
 <212> ADN
 35 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: cebador

 <400> 74
 40 cgattctgcg gccgctgagg agacggtgac ctgtgtccct tggccccaga tatc 54

 <210> 75
 <211> 43
 45 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: cebador

 <400> 75
 50 cccttgcccc cagatatcaa aacnnttgc cacagtaata cac 43

 <210> 76
 <211> 36
 <212> ADN
 <213> Secuencia artificial

 <220>
 60 <223> Descripción de la secuencia artificial: cebador

 <400> 76
 cattgaaga tacagagtgt tcttggcatt gtctct 36

 65 <210> 77
 <211> 39

ES 2 376 454 T3

<212> ADN
 <213> Secuencia artificial

<220>
 5 <223> Descripción de la secuencia artificial: cebador

<400> 77
 tatcttcaaa tgaacagtct gaaaccagag gacacggct 39

10 <210> 78
 <211> 37
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> Descripción de la secuencia artificial: cebador

<400> 78
 gttgtccct ggcccanac nnttcacac agtaata 37

20 <210> 79
 <211> 396
 <212> ADN
 <213> Humano

25 <400> 79
 gaggtccagc tgcaggagtc tgggggaggc ttagtccagc ctgggggggc cctgagactc 60
 tcctgttcag cctctggatt caccttcagt agctatgcta tgcactgggt cgcaccaggct 120
 ccaggaagg gactggaata tgtttcagct attagtagta atgggggtag cacatactac 180
 gcagactccg tgaagggcag attcaccatc tccagagaca attccaagaa cactctgtat 240
 cttcaaatga gcagtctgag agctgaggac acggctgtgt attactgtgt gaaagacagg 300
 ttaaaagtgg agtactatga tagtagtggt tattacgttt ctcgggtcgg tgcttttgat 360
 atctggggcc aagggacac ggtcaccgtc tcatca 396

<210> 80
 <211> 396
 30 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: A6 humano
 35 con modificaciones

<400> 80
 gaggtccagc tgcaggcttc tgggggaggc ttagtccagc ctgggggggc cctgagactc 60
 tcctgtgctg cctctggatt caccttcagt agctatgcta tgcactgggt cgcaccaggct 120
 ccaggaagg gactggaata tgtttcagct attagtagta atgggggtag cacatactac 180
 gcagactccg tgaagggcag attcaccatc tccagagaca attccaagaa cactctgtat 240
 cttcaaatga acagtctgag agctgaggac acggctgtgt attactgtgc gaaagacagg 300
 ttaaaagtgg agtactatga tagtagtggt tattacgttt ctcgggtcgg tgcttttgat 360
 atctggggcc aagggacaca ggtcaccgtc tcatca 396

<210> 81
 40 <211> 396
 <212> ADN
 <213> Secuencia artificial

<220>
 45 <223> Descripción de la secuencia artificial: A6 humano
 con modificaciones

<400> 81

ES 2 376 454 T3

```
gaggtccagc tgcaggcttc tgggggaggc ttagtccagc ctggggggtc cctgagactc 60
tcctgtgctg cctctggatt caccttcagt agctatgcta tgcactgggt cgcaggct 120
ccaggggaagg gactggaata tgtttcagct attagtagta atgggggtag cacatactac 180
gcagactccg tgaagggcag attcaccatc tccagagaca atgccaagaa cactctgtat 240
cttcaaatga acagtctgaa accagaggac acggctgtgt attactgtgc gaaagacagg 300
ttaaagtgg agtactatga tagtagtggt tattacgttt ctcggttcgg tgcttttgat 360
atctggggcc aagggacaca ggtcaccgtc tcatca 396
```

<210> 82

<211> 396

5 <222> ADN

<213> Secuencia artificial

<220>

10 <223> Descripción de la secuencia artificial: A6 humano con modificaciones

```
gaggtccaat tacaggaagg tgggtggcgga ctggtgcaac caggaggatc cctgagactc 60
tcctgttcag cctctggatt caccttcagt agctatgcta tgcactgggt cgcaggct 120
ccaggggaagg gactggaata tgtttcagct attagtagta atgggggtag cacatactac 180
gcagactccg tgaagggcag attcaccatc tccagagaca attccaagaa cactctgtat 240
cttcaaatga gcagtctgag agctgaggac acggctgtgt attactgtgt gaaagacagg 300
ttaaagtgg agtactatga tagtagtggt tattacgttt ctcggttcgg tgcttttgat 360
atctggggcc aagggacaac ggtcaccgtc tcatca 396
```

<400> 82

<210> 83

<211> 396

15 <212> ADN

<213> Secuencia artificial

<220>

20 <223> Descripción de la secuencia artificial: A6 humano con modificaciones

<400> 83

```
gaggtccagc tgcaggagtc tgggggaggc ttagtccagc ctggggggtc cctgagactc 60
tcctgttcag cctctggatt caccttcagt agctatgcta tgcactgggt cgcaggct 120
ccaggggaagg aacgtgaagg tgtttcagct attagtagta atgggggtag cacatactac 180
gcagactccg tgaagggcag attcaccatc tccagagaca attccaagaa cactctgtat 240
cttcaaatga gcagtctgag agctgaggac acggctgtgt attactgtgc agcagacagg 300
ttaaagtgg agtactatga tagtagtggt tattacgttt ctcggttcgg tgcttttgat 360
atctggggcc aagggacaac ggtcaccgtc tcatca 396
```

<210> 84

25 <211> 396

<212> ADN

<213> Secuencia artificial

<220>

30 <223> Descripción de la secuencia artificial: A6 humano con modificaciones

<400> 84

```
gaggtccaat tacaagctag tgggtggcgga ctggtgcaac caggaggttc cctgagactc 60
tcctgtgctg cctctggatt caccttcagt agctatgcta tgcactgggt cgcaggct 120
ccaggggaagg gactggaata tgtttcagct attagtagta atgggggtag cacatactac 180
gcagactccg tgaagggcag attcaccatc tccagagaca attccaagaa cactctgtat 240
cttcaaatga acagtctgag agctgaggac acggctgtgt attactgtgc gaaagacagg 300
ttaaagtgg agtactatga tagtagtggt tattacnmmn nmmmmmmmmn mnnmmttgat 360
atctggggcc aagggacaca ggtcaccgtc tcctca 396
```

35 <210> 85

<211> 378

<212> ADN

<213> Secuencia artificial

40 <220>

<223> Descripción de la secuencia artificial: A6 humano con modificaciones

ES 2 376 454 T3

<400> 85

```

gaggtccaat tacaagctag tggtagcgga ctggtgcaac caggacctc cctgagactc 60
tcctgtgctg cctctggatt caccttcagt agctatgcta tgcactgggt cgcagcaggct 120
ccaggggaag gactggaata tgtttcagct attagtagta atgggggtag cacatactac 180
gcagactccg tgaagggcag attcaccatc tccagagaca attccaagaa cactctgtat 240
cttcaaataa acagtctgag agctgaggac acggctgtgt attactgtgg gaaannmnm 300
nmnmnmnmnm nmnmnmnmnm mnmnmnmnm nmnmnmmttg atatctgggg ccaagggaca 360
caggtcaccg tctcctca                                     378
    
```

<210> 86

<211> 132

5 <212> PRT

<213> Humano

<400> 86

```

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

10 <210> 87

<211> 132

<212> PRT

<213> Secuencia artificial

15 <220>

<223> Descripción de la secuencia artificial: A6 humano con modificaciones

<400> 87

ES 2 376 454 T3

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

<210> 88
 <211> 132
 <212> PRT
 5 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: A6 humano
 con modificaciones

10 <400> 88
 Glu Val Gln Leu Gln Ala Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
 1 5 10 15
 Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Ser Tyr
 20 25 30
 Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Tyr Val
 35 40 45
 Ser Ala Ile Ser Ser Asn Gly Gly Ser Thr Tyr Tyr Ala Asp Ser Val
 50 55 60
 Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Thr Leu Tyr
 65 70 75 80
 Leu Gln Met Asn Ser Leu Lys Pro Glu Asp Thr Ala Val Tyr Tyr Cys
 85 90 95
 Ala Lys Asp Arg Leu Lys Val Glu Tyr Tyr Asp Ser Ser Gly Tyr Tyr
 100 105 110
 Val Ser Arg Phe Gly Ala Phe Asp Ile Trp Gly Gln Gly Thr Gln Val
 115 120 125
 Thr Val Ser Ser
 130

15 <210> 89
 <211> 132
 <212> PRT
 <213> Secuencia artificial

<220>
 20 <223> Descripción de la secuencia artificial: A6 humano
 con modificaciones

ES 2 376 454 T3

<400> 89

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

<210> 90

5 <211> 132

<212> PRT

<213> Secuencia artificial

<220>

10 <223> Descripción de la secuencia artificial: A6 humano con modificaciones

<400> 90

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

15 <210> 91

<211> 132

<212> PRT

<213> Secuencia artificial

20 <220>

<223> Descripción de la secuencia artificial: A6 humano con modificaciones

ES 2 376 454 T3

<400> 91

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

- 5 <210> 92
- <211> 126
- <212> PRT
- <213> Secuencia artificial

- 10 <220>
- <223> Descripción de la secuencia artificial: A6 humano con modificaciones

<400> 92

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

15

REIVINDICACIONES

- 5 1. Una biblioteca combinatoria que contiene variantes de una molécula de unión al ligando V_H parental, donde dicha molécula de unión al ligando parental comprende un fragmento V_H de inmunoglobulina que contiene al menos las regiones marco (FR) de un dominio V_H de inmunoglobulina y donde dichas variantes contienen al menos las regiones FR de dicho dominio V_H de inmunoglobulina, y difieren de dicha molécula de unión al ligando parental en los residuos de aminoácidos que constituyen parte de al menos una de las CDR de dicha molécula de unión al ligando parental, donde dicho dominio V_H de inmunoglobulina se describe en la SEC. ID N°: 87 u 88.
- 10 2. Una biblioteca de acuerdo con la reivindicación 1, donde dichas variantes incluyen la totalidad de la CDR1 y la CDR2 del dominio V_H de SEC. ID N°: 87 u 88.
- 15 3. Un método para generar una biblioteca de expresión de V_H que tenga una diversidad de secuencias CDR, donde dicho método comprende:
- (a) proporcionar vectores de expresión, donde dichos vectores comprenden una secuencia de ácido nucleico que codifica un V_H, dicho V_H se define como la variante definida en la reivindicación 1;
- (b) introducir mediante mutagénesis una variación en una secuencia de ácido nucleico CDR; y
- 20 (c) recuperar una biblioteca de expresión que tenga una diversidad de actividades de unión donde dicho V_H es V_H humano.
4. El método de acuerdo con la reivindicación 3, donde los V_H tienen diversidad de CDR sólo en las secuencias CDR3.
- 25 5. Un método para identificar un polipéptido que comprende un V_H que se une al ligando diana, que comprende los pasos de:
- (a) cribar miembros polipeptídicos de la biblioteca de la reivindicación 1 o 2 para determinar su capacidad para unirse preferencialmente al ligando diana; e
- 30 (b) identificar al menos un miembro polipeptídico que se una al ligando diana.
6. Un dominio de inmunoglobulina que comprende una variante de una molécula de unión al ligando V_H parental, donde dicha molécula de unión al ligando parental comprende un fragmento V_H de inmunoglobulina que contiene al menos las regiones FR del dominio V_H de inmunoglobulina de SEC. ID N°: 87 y donde dichas variantes comprenden al menos las regiones FR del dominio V_H de inmunoglobulina de SEC. ID N°: 87 y difieren de dicha molécula de unión al ligando parental en los residuos de aminoácidos que constituyen al menos parte de al menos una de las CDR de dicha molécula de unión al ligando parental.
- 35 7. Dominio de inmunoglobulina de acuerdo con la reivindicación 6, donde dichas variantes difieren de dicha molécula de unión al ligando parental en los aminoácidos que constituyen al menos parte de la CDR3.

Figuras

Figura 1

Estructura del dominio V_H del anticuerpo A6 humano

1	2	3	4	5	6	7	8	9	10	11	12	13	14
GAG	GTC	CAG	CTG	CAG	GAG	TCT	GGG	GGA	GGC	TTA	GTC	CAG	CCT
E	V	Q	L	Q	E	S	G	G	G	L	V	Q	P
15	16	17	18	19	20	21	22	23	24	25	26	27	28
GGG	GGG	TCC	CTG	AGA	CTC	TCC	TGT	TCA	GCC	TCT	GGA	TTC	ACC
G	G	S	L	R	L	S	C	S	A	S	G	F	T
29	30	31	32	33	34	35	36	37	38	39	40	41	42
TTC	AGT	AGC	TAT	GCT	ATG	CAC	TGG	GTC	CGC	CAG	GCT	CCA	GGG
F	S	S	Y	A	M	H	W	V	R	Q	A	P	G
CDR1													
43	44	45	46	47	48	49	50	51	52	a	53	54	55
AAG	GGA	CTG	GAA	TAT	GTT	TCA	GCT	ATT	AGT	AGT	AAT	GGG	GGT
K	G	L	E	Y	V	S	A	I	S	S	N	G	G
CDR2													
56	57	58	59	60	61	62	63	64	65	66	67	68	69
AGC	ACA	TAC	TAC	GCA	GAC	TCC	GTG	AAG	GGC	AGA	TTC	ACC	ATC
S	T	Y	Y	A	D	S	V	K	G	R	F	T	I
70	71	72	73	74	75	76	77	78	79	80	81	82	a
TCC	AGA	GAC	AAT	TCC	AAG	AAC	ACT	CTG	TAT	CTT	CAA	ATG	AGC
S	R	D	N	S	K	N	T	L	Y	L	Q	M	S
b	c	83	84	85	86	87	88	89	90	91	92	93	94
AGT	CTG	AGA	GCT	GAG	GAC	ACG	GCT	GTG	TAT	TAC	TGT	GTG	AAA
S	L	R	A	E	D	T	A	V	Y	Y	C	V	K

ES 2 376 454 T3

Figura 1 (continuación)

95	96	97	98	99	100	a	b	c	d	e	f	g	h
GAC	AGG	TTA	AAA	GTG	GAG	TAC	TAT	GAT	AGT	AGT	GGT	TAT	TAC
<u>D</u>	<u>R</u>	<u>L</u>	<u>K</u>	<u>V</u>	<u>E</u>	<u>Y</u>	<u>Y</u>	<u>D</u>	<u>S</u>	<u>S</u>	<u>G</u>	<u>Y</u>	<u>Y</u>
CDR3													
i	j	k	l	m	n	o	101	102	103	104	105	105	107
GTT	TCT	CGG	TTC	GGT	GCT	TTT	GAT	ATC	TGG	GGC	CAA	GGG	ACA
<u>V</u>	<u>S</u>	<u>R</u>	<u>F</u>	<u>G</u>	<u>A</u>	<u>F</u>	<u>D</u>	<u>I</u>	<u>W</u>	<u>G</u>	<u>Q</u>	<u>G</u>	<u>T</u>
108	109	110	111	112	113								
ACG	GTC	ACC	GTC	TCA	TCA								
<u>T</u>	<u>V</u>	<u>T</u>	<u>V</u>	<u>S</u>	<u>S</u>								

Figura 2

Estructura del dominio V_H del anticuerpo A6 humano

1	2	3	4	5	6	7	8	9	10	11	12	13	14
GAG	GTC	CAG	CTG	CAG	GCT	TCT	GGG	GGA	GGC	TTA	GTC	CAG	CCT
E	V	Q	L	Q	A	S	G	G	G	L	V	Q	P
15	16	17	18	19	20	21	22	23	24	25	26	27	28
GGG	GGG	TCC	CTG	AGA	CTC	TCC	TGT	GCT	GCC	TCT	GGA	TTC	ACC
G	G	S	L	R	L	S	C	A	A	S	G	F	T
29	30	31	32	33	34	35	36	37	38	39	40	41	42
TTC	AGT	AGC	TAT	GCT	ATG	CAC	TGG	GTC	CGC	CAG	GCT	CCA	GGG
F	S	<u>S</u>	<u>Y</u>	<u>A</u>	<u>M</u>	<u>H</u>	W	V	R	Q	A	P	G
CDR1													
43	44	45	46	47	48	49	50	51	52	a	53	54	55
AAG	GGA	CTG	GAA	TAT	GTT	TCA	GCT	ATT	AGT	AGT	AAT	GGG	GGT
K	G	L	E	Y	V	S	<u>A</u>	<u>I</u>	<u>S</u>	<u>S</u>	<u>N</u>	<u>G</u>	<u>G</u>
CDR2													
56	57	58	59	60	61	62	63	64	65	66	67	68	69
AGC	ACA	TAC	TAC	GCA	GAC	TCC	GTG	AAG	GGC	AGA	TTC	ACC	ATC
<u>S</u>	<u>T</u>	<u>Y</u>	<u>Y</u>	<u>A</u>	<u>D</u>	<u>S</u>	<u>V</u>	<u>K</u>	<u>G</u>	R	F	T	I
70	71	72	73	74	75	76	77	78	79	80	81	82	a
TCC	AGA	GAC	AAT	TCC	AAG	AAC	ACT	CTG	TAT	CTT	CAA	ATG	AAC
S	R	D	N	S	K	N	T	L	Y	L	Q	M	N
b	c	83	84	85	86	87	88	89	90	91	92	93	94
AGT	CTG	AGA	GCT	GAG	GAC	ACG	GCT	GTG	TAT	TAC	TGT	GCG	AAA
S	L	R	A	E	D	T	A	V	Y	Y	C	A	K

ES 2 376 454 T3

Figura 2 (continuación)

95	96	97	98	99	100	a	b	c	d	e	f	g	h
GAC	AGG	TTA	AAA	GTG	GAG	TAC	TAT	GAT	AGT	AGT	GGT	TAT	TAC
<u>D</u>	<u>R</u>	<u>L</u>	<u>K</u>	<u>V</u>	<u>E</u>	<u>Y</u>	<u>Y</u>	<u>D</u>	<u>S</u>	<u>S</u>	<u>G</u>	<u>Y</u>	<u>Y</u>
CDR3													
i	j	k	l	m	n	o	101	102	103	104	105	105	107
GTT	TCT	CGG	TTC	GGT	GCT	TTT	GAT	ATC	TGG	GGC	CAA	GGG	ACA
<u>V</u>	<u>S</u>	<u>R</u>	<u>F</u>	<u>G</u>	<u>A</u>	<u>F</u>	<u>D</u>	<u>I</u>	<u>W</u>	<u>G</u>	<u>Q</u>	<u>G</u>	<u>T</u>
108	109	110	111	112	113								
CAG	GTC	ACC	GTC	TCA	TCA								
<u>Q</u>	<u>V</u>	<u>T</u>	<u>V</u>	<u>S</u>	<u>S</u>								

Figura 3

Estructura del dominio V_H del anticuerpo A6 humano

1	2	3	4	5	6	7	8	9	10	11	12	13	14
GAG	GTC	CAG	CTG	CAG	GCT	TCT	GGG	GGG	GGC	TTA	GTC	CAG	CCT
E	V	Q	L	Q	A	S	G	G	G	L	V	Q	P
15	16	17	18	19	20	21	22	23	24	25	26	27	28
GGG	GGG	TCC	CTG	AGA	CTC	TCC	TGT	GCT	GCC	TCT	GGA	TTC	ACC
G	G	S	L	R	L	S	C	A	A	S	G	F	T
29	30	31	32	33	34	35	36	37	38	39	40	41	42
TTC	AGT	AGC	TAT	GCT	ATG	CAC	TGG	GTC	CGC	CAG	GCT	CCA	GGG
F	S	S	Y	A	M	H	W	V	R	Q	A	P	G
CDR1													
43	44	45	46	47	48	49	50	51	52	a	53	54	55
AAG	GGA	CTG	GAA	TAT	GTT	TCA	GCT	ATT	AGT	AGT	AAT	GGG	GGT
K	G	L	E	Y	V	S	A	I	S	S	N	G	G
CDR2													
56	57	58	59	60	61	62	63	64	65	66	67	68	69
AGC	ACA	TAC	TAC	GCA	GAC	TCC	GTG	AAG	GGC	AGA	TTC	ACC	ATC
S	T	Y	Y	A	D	S	V	K	G	R	F	T	I
70	71	72	73	74	75	76	77	78	79	80	81	82	a
TCC	AGA	GAC	AAT	GCC	AAG	AAC	ACT	CTG	TAT	CTT	CAA	ATG	AAC
S	R	D	N	A	K	N	T	L	Y	L	Q	M	N
b	c	83	84	85	86	87	88	89	90	91	92	93	94
AGT	CTG	AAA	CCA	GAG	GAC	ACG	GCT	GTG	TAT	TAC	TGT	GCG	AAA
S	L	K	P	E	D	T	A	V	Y	Y	C	A	K

Figura 3 (continuación)

95	96	97	98	99	100	a	b	c	d	e	f	g	h
GAC	AGG	TTA	AAA	GTG	GAG	TAC	TAT	GAT	AGT	AGT	GGT	TAT	TAC
<u>D</u>	<u>R</u>	<u>L</u>	<u>K</u>	<u>V</u>	<u>E</u>	<u>Y</u>	<u>Y</u>	<u>D</u>	<u>S</u>	<u>S</u>	<u>G</u>	<u>Y</u>	<u>Y</u>
CDR3													
i	j	k	l	m	n	o	101	102	103	104	105	105	107
GTT	TCT	CGG	TTC	GGT	GCT	TTT	GAT	ATC	TGG	GGC	CAA	GGG	ACA
<u>V</u>	<u>S</u>	<u>R</u>	<u>F</u>	<u>G</u>	<u>A</u>	<u>F</u>	<u>D</u>	<u>I</u>	<u>W</u>	<u>G</u>	<u>Q</u>	<u>G</u>	<u>T</u>
108	109	110	111	112	113								
<u>CAG</u>	<u>GTC</u>	<u>ACC</u>	<u>GTC</u>	<u>TCA</u>	<u>TCA</u>								
<u>Q</u>	<u>V</u>	<u>T</u>	<u>V</u>	<u>S</u>	<u>S</u>								

Figura 4

Estructura del dominio V_H del anticuerpo A6 humano. Los nucleótidos mutados que abarcan los residuos 7-48 para eliminar el sitio de recombinación están en negritas y subrayados

1	2	3	4	5	6	7	8	9	10	11	12	13	14
GAG	GTC	CAA	TTA	CAG	GAA	AGT	GGT	GGC	GGA	CTG	GTG	CAA	CCA
E	V	Q	L	Q	E	S	G	G	G	L	V	Q	P
15	16	17	18	19	20	21	22	23	24	25	26	27	28
GGA	GGA	TCC	CTG	AGA	CTC	TCC	TGT	TCA	GCC	TCT	GGA	TTC	ACC
G	G	S	L	R	L	S	C	S	A	S	G	F	T
29	30	31	32	33	34	35	36	37	38	39	40	41	42
TTC	AGT	AGC	TAT	GCT	ATG	CAC	TGG	GTC	CGC	CAG	GCT	CCA	GGG
F	S	S	Y	A	M	H	W	V	R	Q	A	P	G
CDR1													
43	44	45	46	47	48	49	50	51	52	a	53	54	55
AAG	GGA	CTG	GAA	TAT	GTT	TCA	GCT	ATT	AGT	AGT	AAT	GGG	GGT
K	G	L	E	Y	V	S	A	I	S	S	N	G	G
CDR2													
56	57	58	59	60	61	62	63	64	65	66	67	68	69
AGC	ACA	TAC	TAC	GCA	GAC	TCC	GTG	AAG	GGC	AGA	TTC	ACC	ATC
S	T	Y	Y	A	D	S	V	K	G	R	F	T	I
70	71	72	73	74	75	76	77	78	79	80	81	82	a
TCC	AGA	GAC	AAT	TCC	AAG	AAC	ACT	CTG	TAT	CTT	CAA	ATG	AGC
S	R	D	N	S	K	N	T	L	Y	L	Q	M	S
b	c	83	84	85	86	87	88	89	90	91	92	93	94
AGT	CTG	AGA	GCT	GAG	GAC	ACG	GCT	GTG	TAT	TAC	TGT	GTG	AAA
S	L	R	A	E	D	T	A	V	Y	Y	C	V	K

Figura 4 (continuación)

95	96	97	98	99	100	a	b	c	d	e	f	g	h
GAC	AGG	TTA	AAA	GTG	GAG	TAC	TAT	GAT	AGT	AGT	GGT	TAT	TAC
D	R	L	K	V	E	Y	Y	D	S	S	G	Y	Y
CDR3													
i	j	k	l	m	n	o	101	102	103	104	105	105	107
GTT	TCT	CGG	TTC	GGT	GCT	TTT	GAT	ATC	TGG	GGC	CAA	GGG	ACA
V	S	R	F	G	A	F	D	I	W	G	Q	G	T
108	109	110	111	112	113								
ACG	GTC	ACC	GTC	TCA	TCA								
T	V	T	V	S	S								

Figura 5

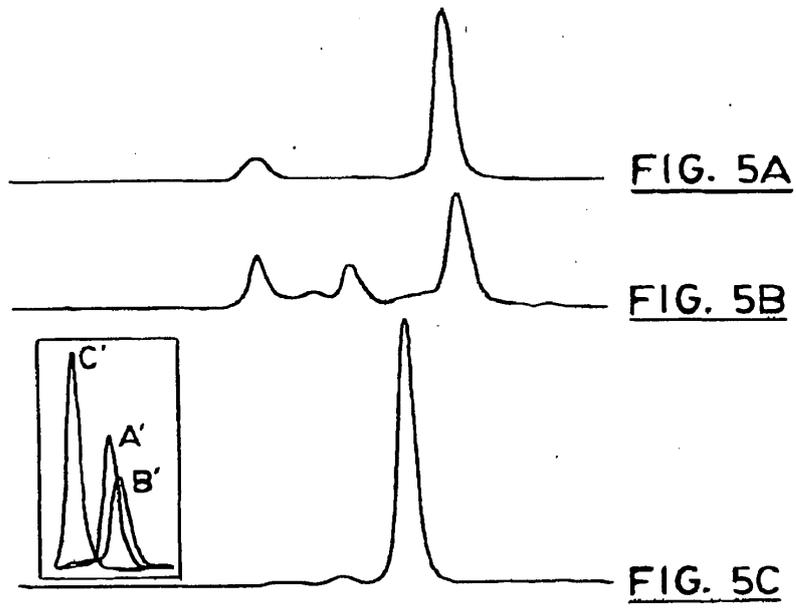


Figura 6

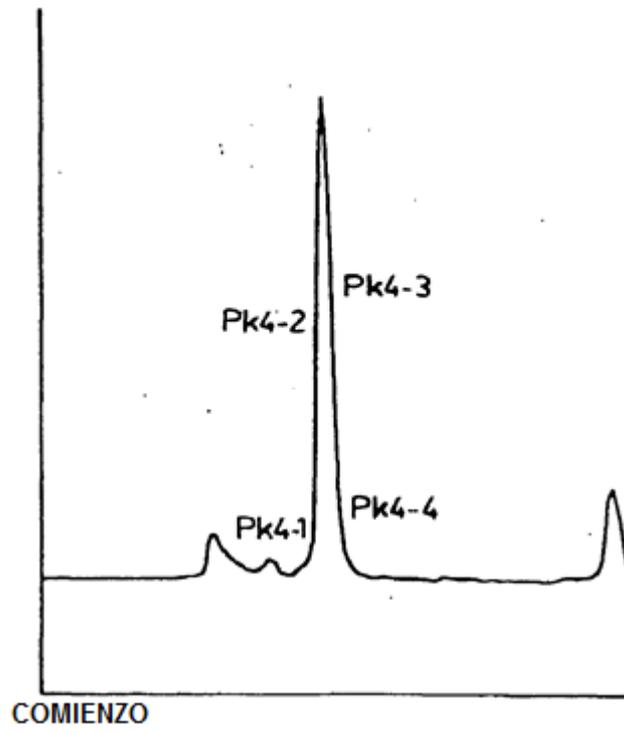
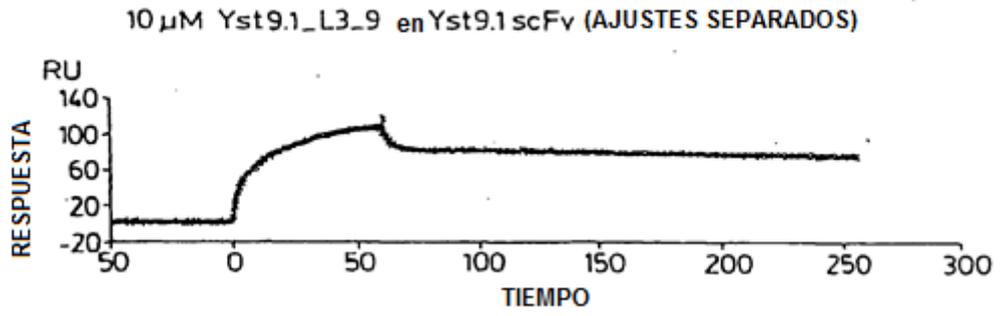


Figura 7
Figura

7

Figura 7



ES 2 376 454 T3

Figura 8

Estructura del dominio V_H modificado del anticuerpo A6 humano que muestra sustituciones en las posiciones 44, 45, 47, 93 y 94. El sitio *NheI* está subrayado

1	2	3	4	5	6	7	8	9	10	11	12	13	14
GAG	GTC	CAG	CTG	CAG	GAG	TCT	GGG	GGA	GGC	TTA	GTC	CAG	CCT
E	V	Q	L	Q	E	S	G	G	G	L	V	Q	P
15	16	17	18	19	20	21	22	23	24	25	26	27	28
GGG	GGG	TCC	CTG	AGA	CTC	TCC	TGT	TCA	GCC	TCT	GGA	TTC	ACC
G	G	S	L	R	L	S	C	S	A	S	G	F	T
29	30	31	32	33	34	35	36	37	38	39	40	41	42
TTC	AGT	AGC	TAT	GCT	ATG	CAC	TGG	GTC	CGC	CAG	GCT	CCA	GGG
F	S	<u>S</u>	<u>Y</u>	<u>A</u>	<u>M</u>	<u>H</u>	W	V	R	Q	A	P	G
CDR1													
43	44	45	46	47	48	49	50	51	52	a	53	54	55
AAG	<u>GAA</u>	<u>CGT</u>	<u>GAA</u>	<u>GGT</u>	GTT	TCA	GCT	ATT	AGT	AGT	AAT	GGG	GGT
K	E	R	E	G	V	S	<u>A</u>	<u>I</u>	<u>S</u>	<u>S</u>	<u>N</u>	<u>G</u>	<u>G</u>
CDR2													
56	57	58	59	60	61	62	63	64	65	66	67	68	69
AGC	ACA	TAC	TAC	GCA	GAC	TCC	GTG	AAG	GGC	AGA	TTC	ACC	ATC
<u>S</u>	<u>T</u>	<u>Y</u>	<u>Y</u>	<u>A</u>	<u>D</u>	<u>S</u>	<u>V</u>	<u>K</u>	<u>G</u>	R	F	T	I
70	71	72	73	74	75	76	77	78	79	80	81	82	a
TCC	AGA	GAC	AAT	TCC	AAG	AAC	ACT	CTG	TAT	CTT	CAA	ATG	AGC
S	R	D	N	S	K	N	T	L	Y	L	Q	M	S
b	c	83	84	85	86	87	88	89	90	91	92	93	94
AGT	CTG	AGA	GCT	GAG	GAC	ACG	GCT	GTG	TAT	TAC	TGT	GCA	GCA
S	L	R	A	E	D	T	A	V	Y	Y	C	A	A

ES 2 376 454 T3

Figura 8 (continuación)

95	96	97	98	99	100	a	b	c	d	e	f	g	h
GAC	AGG	TTA	AAA	GTG	GAG	TAC	TAT	GAT	AGT	AGT	GGT	TAT	TAC
<u>D</u>	<u>R</u>	<u>L</u>	<u>K</u>	<u>V</u>	<u>E</u>	<u>Y</u>	<u>Y</u>	<u>D</u>	<u>S</u>	<u>S</u>	<u>G</u>	<u>Y</u>	<u>Y</u>
CDR3													
i	j	k	l	m	n	o	101	102	103	104	105	105	107
GTT	TCT	CGG	TTC	GGT	GCT	TTT	GAT	ATC	TGG	GGC	CAA	GGG	ACA
<u>V</u>	<u>S</u>	<u>R</u>	<u>F</u>	<u>G</u>	<u>A</u>	<u>F</u>	<u>D</u>	<u>I</u>	<u>W</u>	<u>G</u>	<u>Q</u>	<u>G</u>	<u>T</u>
108	109	110	111	112	113								
ACG	GTC	ACC	GTC	TCA	TCA								
<u>T</u>	<u>V</u>	<u>T</u>	<u>V</u>	<u>S</u>	<u>S</u>								

Figura 9

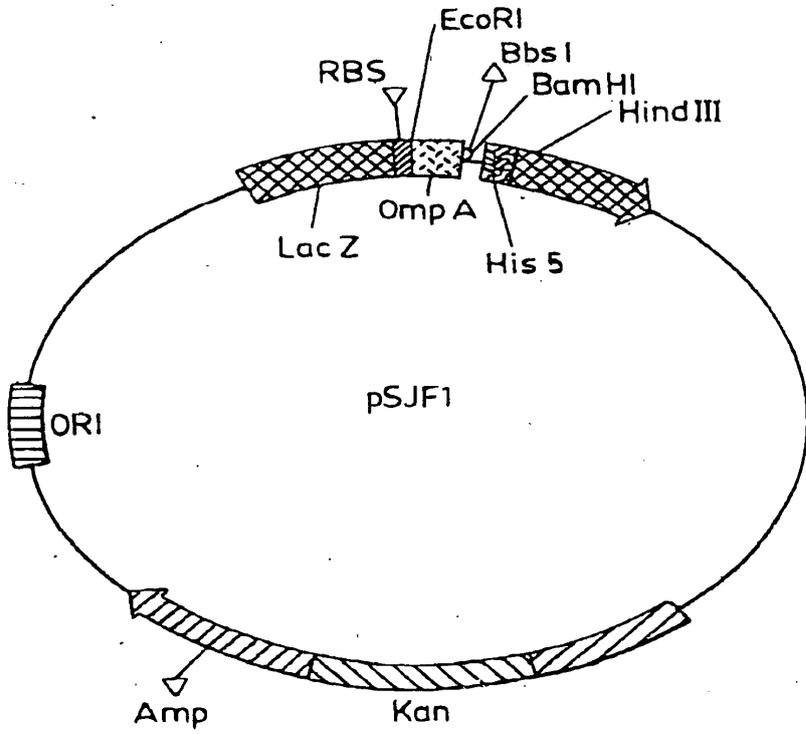


Figura 10

Secuencia A6VH para la biblioteca A6VH-L1

5 mutaciones (en engritas)
Aleatorización en 100i-n

GAG GTC CAA TTA CAA GCT AGT GGT GGC GGA CTG GTG CAA CCA GGA GGT TCC CTG
E V Q L Q A₆ S G G G L V Q P G G S L

AGA CTC TCC TGT GCT GCC TCT GGA TTC ACC TTC AGT AGC TAT GCT ATG CAC TGG GTC
R L S C A₂₃ A S G F T F S S Y A M H W V

CGC CAG GCT CCA GGG AAG GGA CTG GAA TAT GTT TCA GCT ATT AGT AGT AAT GGG
R Q A P G K G L E Y V S A I S S N G

GGT AGC ACA TAC TAC GCA GAC TCC GTG AAG GGC AGA TTC ACC ATC TCC AGA GAC
G S T Y Y A D S V K G R F T I S R D

AAT TCC AAG AAC ACT CTG TAT CTT CAA ATG AAC AGT CTG AGA GCT GAG GAC ACG
N S₇₄ K N T L Y L Q M N_{82a} S L R₈₃ A₈₄ E D T

GCT GTG TAT TAC TGT GCG AAA GAC AGG TTA AAA GTG GAG TAC TAT GAT AGT AGT
A V Y Y C A₉₃ K D R L K V E Y Y D S S

GGT TAT TAC 6 [NNA/C] TTT GAT ATC TGG GGC CAA GGG ACA CAG GTC ACC GTC TCC
G Y Y 6 [N] F D I W G Q G T Q₁₀₈ V T V S

TCA
S

Figura 11

Secuencia A6VH para la biblioteca A6VH-L1a

5 mutaciones (en engritas)

100i-n eliminados

Aleatorización en 95-100h con 50% de dopaje de bases en cada posición

```

GAG GTC CAA TTA CAA GCT AGT GGT GGC GGA CTG GTG CAA CCA GGA GGT TCC CTG
E V Q L Q A6 S G G G L V Q P G G S L
AGA CTC TCC TGT GCT GCC TCT GGA TTC ACC TTC AGT AGC TAT GCT ATG CAC TGG GTC
R L S C A23 A S G F T F S S Y A M H W V
CGC CAG GCT CCA GGG AAG GGA CTG GAA TAT GTT TCA GCT ATT AGT AGT AAT GGG
R Q A P G K G L E Y V S A I S S N G
GGT AGC ACA TAC TAC GCA GAC TCC GTG AAG GGC AGA TTC ACC ATC TCC AGA GAC
G S T Y Y A D S V K G R F T I S R D
AAT TCC AAG AAC ACT CTG TAT CTT CA1 ATG AAC AGT CTG AGA GCT GAG GAC ACG
N S71 K N T L Y L Q M N82a S L R83 A84 E D T
GCT GTG TAT TAC TGT GCG AAA 14 [50%NNA/C] TTT GAT ATC TGG GGC CAA GGG ACA
A V Y Y C A93 K 14[ N ] F D I W G Q G T
CAG GTC ACC GTC TCC TCA
Q108 V T V S S
    
```