(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 399 617**

(51) Int. Cl.:

C07K 16/18 (2006.01)
C12N 15/09 (2006.01)
A61K 48/00 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **21.05.2003** **E 03732434 (0)**

(97) Fecha y número de publicación de la concesión europea: **23.01.2013** **EP 1506236**

(54) Título: **Regiones de marco de inmunoglobulina que demuestran estabilidad potenciada en el entorno intracelulcar y métodos de identificación de las mismas**

(30) Prioridad:

22.05.2002 US 382649 P
03.01.2003 US 438256 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
02.04.2013

(73) Titular/es:

ESBATECH - A NOVARTIS COMPANY LLC (100.0%)
WAGISTRASSE 21
8952 SCHLIEREN, CH

(72) Inventor/es:

TISSOT, KATHRIN;
EWERT, STEFAN;
AUF DER MAUR, ADRIAN;
BARBERIS, ALCIDE y
ESCHER, DOMINIK

(74) Agente/Representante:

MILTENYI, Peter

ES 2 399 617 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

Regiones de marco de inmunoglobulina que demuestran estabilidad potenciada en el entorno intracelular y métodos de identificación de las mismas

Campo de la invención

La invención se refiere a la química de proteínas, la biología molecular y la inmunología.

Antecedentes de la técnica relacionada

Los anticuerpos pueden reconocer y dirigirse a casi cualquier molécula con alta especificidad y afinidad. Se ha aprovechado esta característica para convertir estas proteínas naturales en poderosas herramientas para aplicaciones de diagnóstico y terapéuticas. Los avances en la tecnología de ADN recombinante han facilitado la manipulación, clonación y expresión de genes de anticuerpos en una amplia variedad de células no linfoides (Skerra, 1988; Martineau, 1998; Verma, 1998). Se han construido varios fragmentos de anticuerpo diferentes para adaptarse de la mejor manera a las diversas aplicaciones. La entidad más pequeña que retiene la capacidad de unión al antígeno completo de la inmunoglobulina parental completa es el fragmento Fv de cadena sencilla (scFv) (Bird, 1988). Este fragmento de anticuerpo comprende las regiones variables de las cadenas pesada y ligera unidas por un enlazador peptídico flexible, lo que permite la expresión de la proteína de un único gen.

Los fragmentos de anticuerpo tienen varias ventajas importantes en comparación con la molécula de inmunoglobulina completa. Debido a su menor tamaño, se facilita la expresión y se potencia la producci en una variedad de células huésped de expresión, tales como células de *E. coli* (Plückthun, 1996). Además, los fragmentos de anticuerpo permiten una penetración mejorada en tumores en aplicaciones *in vivo* (Yokota, 1992) y pueden unirse covalentemente a diversas moléculas efectoras para enfoques terapéuticos.

Los anticuerpos que se producen de manera natural, que se secretan por células plasmáticas, han evolucionado para funcionar en un entorno oxidante, extracelular. Para obtener su estructura plegada, funcional, generalmente requieren la formación de puentes disulfuro dentro de los dominios independientes, que son cruciales para la estabilidad del plegamiento de la inmunoglobulina. A diferencia de los anticuerpos de longitud completa, los fragmentos de anticuerpo scFv o Fab pueden, en principio, expresarse funcionalmente en un entorno reductor en el interior de cualquier célula y dirigirse a cualquier compartimento para dirigirse a proteínas intracelulares y, por tanto, evocar efectos biológicos específicos (Biocca, 1991). En efecto, algunos fragmentos de anticuerpo de cadena sencilla intracelulares, que se denominan intracuerpos, se han aplicado satisfactoriamente para modular la función de proteínas diana intracelulares en diferentes sistemas biológicos. Por tanto, se ha demostrado resistencia frente a infecciones virales en biotecnología vegetal (Tavladoraki, 1993; Benvenuto, 1995), se ha mostrado la unión de intracuerpos a proteínas del VIH (Rondon, 1997) y se ha descrito la unión a productos oncogénicos (Biocca, 1993; Cochet, 1998; Lener, 2000). Además, los anticuerpos intracelulares prometen ser una herramienta valiosa en la caracterización de la función de un amplio número de genes ahora identificados a través de la secuenciación del genoma humano (Richardson, 1995; Marasco, 1997). Por ejemplo, pueden usarse en un enfoque de genómica funcional para bloquear o modular la actividad de proteínas recién identificadas, contribuyendo de ese modo a la comprensión de sus funciones. Finalmente, los intracuerpos tienen posibles aplicaciones de diagnóstico y terapéuticas, por ejemplo en prácticas de terapia génica.

A pesar de estas grandes perspectivas, la generación de intracuerpos funcionales está limitada todavía por su inestabilidad e insolubilidad o propensión a agregarse. El entorno reductor del citoplasma impide la formación de los puentes disulfuro intracatenarios conservados, haciendo por tanto que un alto porcentaje de fragmentos de anticuerpo sean inestables y, como consecuencia, no funcionales en el interior de la célula (Biocca, 1995; Proba, 1997). La estabilidad y la solubilidad de fragmentos de anticuerpo representan, por tanto, un obstáculo importante para la aplicación de intracuerpos como posibles moduladores de la función de la proteína *in vivo*. Hasta la fecha, no pueden realizarse predicciones sobre los requisitos de secuencia que hacen que un fragmento de anticuerpo sea funcional en un entorno intracelular.

Existe, por tanto, una necesidad de fragmentos de anticuerpo que funcionen bien en una amplia gama de diferentes tipos celulares y puedan usarse, por tanto, como regiones de marco para diversas especificidades de unión. Tales regiones de marco pueden usarse para construir bibliotecas para selección intracelular o pueden servir como aceptor para las partes de unión de un anticuerpo existente.

Además de ser adecuados de forma única para aplicaciones intracelulares, tales fragmentos de anticuerpo o anticuerpos completos basados en regiones de marco de dominios variables muy estables también tienen una ventaja distintiva con respecto a los demás anticuerpos en numerosas aplicaciones extracelulares e *in vitro*. Cuando se producen tales regiones de marco en un entorno oxidante, pueden formarse sus puentes disulfuro, potenciando adicionalmente su estabilidad y haciéndolos sumamente resistentes frente a la agregación y la degradación de proteasas. La semivida *in vivo* (y por tanto la resistencia frente a la agregación y la degradación por proteasas séricas) es, además de la afinidad y la especificidad, el factor individual más importante para el éxito de los anticuerpos en aplicaciones terapéuticas o de diagnóstico (Willuda, 1999). La semivida de fragmentos de anticuerpo puede aumentarse adicionalmente a través de la unión covalente de moléculas de polímero tales como

polietilenglicol (PEG) (Weir, 2002). Las moléculas estables de este tipo representan un avance significativo en el uso de anticuerpos, especialmente, pero no exclusivamente, cuando no se desea la funcionalidad Fc.

La gran importancia práctica de las bibliotecas de fragmentos de anticuerpo ha motivado la investigación en esta área. Winter (documento EP 0368684) ha proporcionado la clonación inicial y la expresión de genes de regiones variables de anticuerpo. Partiendo de estos genes, se crearon grandes bibliotecas de anticuerpos que tienen alta diversidad tanto en las regiones determinantes de complementariedad (CDR) así como en las regiones de marco. Winter no da a conocer, sin embargo, la utilidad de diferentes regiones de marco para la construcción de bibliotecas.

La enseñanza de Plückthun (documento EP 0859841), por otra parte, ha tratado de mejorar el diseño de bibliotecas limitando las regiones de marco a un número definido de secuencias consenso sintéticas. Los esfuerzos en ingeniería de proteínas que implican la introducción de una gran cantidad de mutaciones diseñadas racionalmente han sugerido previamente mutaciones hacia la secuencia consenso respectiva como un medio adecuado para la mejora de la estabilidad de dominios de inmunoglobulina variables aislados (Ohage 1999; Ohage 1999 y el documento US 5.854.027, incorporados como referencia al presente documento).

Plückthun (documento EP 0859841) da a conocer métodos para la optimización adicional de las afinidades de unión en estas secuencias consenso. La patente de Plückthun también reconoce el continuo aumento en el conocimiento referente a los anticuerpos y por consiguiente tiene como objetivo incluir tales hallazgos futuros en el diseño de bibliotecas. Sin embargo, no se sugieren posibles mejoras adicionales de las regiones de marco consenso sintéticas.

Las enseñanzas de Winter, Plückthun y otros (por ejemplo Soderlind, documento WO 0175091) han intentado, por tanto, crear grandes bibliotecas de anticuerpos centrándose en la alta diversidad en las CDR para la selección y aplicación de los scFv seleccionados en condiciones oxidantes. Todas estas bibliotecas no están optimizadas, sin embargo, para aplicaciones intracelulares y, por tanto, no son útiles para la selección y las aplicaciones en un entorno reductor, u otras condiciones que establecen requisitos especiales sobre la estabilidad y la solubilidad del fragmento de anticuerpo expresado.

Las cualidades requeridas para que los fragmentos de anticuerpo funcionen bien en un entorno reductor, por ejemplo el citoplasma de células procariotas y eucariotas, no están claras. La aplicación de anticuerpos intracelulares o "intracuerpos" está por tanto actualmente limitada por su comportamiento impredecible en condiciones reductoras, lo que puede afectar sus propiedades de estabilidad y solubilidad (Biocca, 1995; Wörn, 2000). Las presentes solicitudes de patente (documentos EP1040201, EP1166121 y W00200729) y publicaciones (Visintin, 1999) referentes a la selección intracelular para intracuerpos se centran en la tecnología de selección pero no dan a conocer secuencias de anticuerpo específicas que sean funcionales en células eucariotas, en particular en levaduras, y, por tanto, útiles para la construcción de bibliotecas en este contexto.

Visintin y Tse han descrito independientemente el aislamiento de una denominada secuencia consenso intracelular (ICS) (Visintin, 2002; Tse, 2002). Esta secuencia se derivó de varias secuencias que se habían aislado de una selección de interacción-antígeno-anticuerpo en levaduras. La entrada en la selección intracelular estuvo, sin embargo, fuertemente sesgada debido a la selección mediante presentación en fago anterior. Por tanto, todas menos una de las secuencias de entrada pertenecían al subgrupo VH 3 en el caso de Visintin *et al.* La secuencia consenso ICS publicada es completamente idéntica a la secuencia consenso para el subgrupo VH 3 humano descrito por Knappik (2000) y el documento EP0859841. 60 de los 62 aminoácidos de la ICS también son idénticos a la secuencia consenso del dominio VH humano general que se propuso por Steipe como base para la construcción de dominios variables con estabilidad potenciada (patente estadounidense n.° 6.262.238, incorporada como referencia al presente documento). Estos trabajos se basaron, a su vez, en colecciones de secuencias previas (es decir, Kabat, 1991 y definiciones de subgrupos de dominios variables y determinantes estructurales (Tomlinson, 1992; Williams, 1996; Chothia, 1989 y Chothia, 1987). Sin embargo, debido a que la entrada en la selección de intracuerpos estaba tan fuertemente sesgada (es decir, en el caso de Visintin *et al.* todos menos uno de los dominios VH eran VH3), el aislamiento de secuencias de VH3 a partir de la selección intracelular no es particularmente sorprendente. Debido al fuerte sesgo de su biblioteca de entrada, el trabajo de Tse *et al.* y Visintin *et al.* no proporciona una evaluación rigurosa del repertorio de dominios variables humanos tal como se proporcionaría mediante una investigación no sesgada y tal como se requiere para identificar las regiones de marco de intracuerpo adecuadas presentes en el repertorio humano.

Hemos descrito previamente un sistema, que permite la selección de intracuerpos estables y solubles en levaduras, independiente de su especificidad de unión a antígenos (Auf der Maur (2001), documento W00148017). Este enfoque permite la selección eficaz de bibliotecas de scFv y el aislamiento de regiones de marco específicas, que son estables y solubles en el entorno reductor de la célula de levadura. El objetivo sigue siendo aislar realmente secuencias de región de marco y usar los patrones en una primera etapa para predecir qué tipos de secuencia serían los más estables en el entorno reductor y en una segunda etapa identificar mediante análisis, recombinación y experimentos *in vivo* e *in vitro* adicionales la secuencia óptima.

Breve sumario de la invención

La presente invención compensa el eslabón perdido en el campo de la generación de anticuerpos. Proporciona

secuencias de región de marco de dominios variables de anticuerpo con características superiores con respecto a la estabilidad y la solubilidad. Éstas son características cruciales para muchas aplicaciones relevantes, tales como en diagnóstico, terapia o investigación. Estas regiones de marco pueden usarse para el injerto de especificidades de unión existentes o para la generación de bibliotecas de anticuerpos con alta estabilidad y solubilidad.

Se usaron bibliotecas de scFv para el aislamiento de regiones de marco que son estables y solubles en el entorno reductor de la célula de levadura. El funcionamiento de las regiones de marco aisladas se ha caracterizado posteriormente en líneas celulares humanas y en experimentos *in vitro*. Las regiones de marco descritas pueden servir directamente como estructuras principales aceptoras para especificidades de unión existentes o para construir bibliotecas de CDR mediante la aleatorización de uno o más de los bucles hipervariables para su uso en entornos reductores o que de otro modo constituyen un reto. Las secuencias de dominios variables aisladas se han analizado adicionalmente mediante alineamiento para identificar familias de secuencias preferidas. A partir de esas familias de secuencias de dominios variables preferidas, se eligieron secuencias óptimas basándose en un análisis estructural que excluye las secuencias que contienen residuos de región de marco que alteran el plegamiento de la inmunoglobulina. Los candidatos de secuencia de dominios variables identificados se recombinaron posteriormente en todas las variaciones posibles y se seleccionaron las combinaciones óptimas de dominios variables de la cadena ligera y pesada mediante análisis de su funcionamiento en células de levaduras, de mamíferos e *in vitro*.

Estos scFv optimizados y sus regiones de marco de dominios variables constituyentes, así como otros fragmentos de anticuerpo o anticuerpos completos derivados de los mismos, son ideales como, por ejemplo, estructuras principales aceptoras para especificidades de unión existentes o para la construcción de bibliotecas de CDR mediante aleatorización de uno o más de los bucles hipervariables para su uso en entornos reductores o que de otro modo constituyen un reto. Los anticuerpos adecuados para aplicaciones intracelulares son por definición más estables y solubles. Por consiguiente, su uso también será ventajoso en aplicaciones fuera del entorno intracelular.

La invención proporciona composiciones que comprenden regiones de marco de dominios variables de anticuerpo y fragmentos de anticuerpo Fv de cadena sencilla (scFv) que pueden incorporarse en diversos fragmentos de anticuerpo o anticuerpos completos. Se proporcionan clases de fragmentos de dominios variables de anticuerpo que son las más estables y solubles y por tanto las más adecuadas para aplicaciones intracelulares. También se proporcionan secuencias de región de marco específicas de dominios variables de anticuerpo y fragmentos de anticuerpo scFv que muestran el mejor funcionamiento en ensayos intracelulares. La invención también proporciona secuencias de región de marco específicas de dominios variables de anticuerpo y combinaciones sintéticas de dominios variables de la cadena ligera y pesada en fragmentos scFv que son, por ejemplo, óptimas para aplicaciones intracelulares y muestran un funcionamiento óptimo *in vitro* con respecto a la estabilidad y la solubilidad.

La invención proporciona reactivos de región de marco de cadena sencilla que tienen las estructuras generales:

NH_2-VL-enlazador-VH-COOH o

NH_2-VH-enlazador-VL-COOH.

En otra realización de la invención, la región de marco de cadena sencilla puede fusionarse a un segundo resto de proteína para producir un constructo de fusión de la estructura general:

NH_2-VL-enlazador-VH-segunda proteína-COOH

NH_2-segunda proteína-VL-enlazador-VH-COOH.

Puede invertirse la orientación de las regiones VH y VL en estos constructos de fusión.

En otra realización de la invención, los dominios variables pueden incorporarse en un fragmento Fab, que puede fusionarse adicionalmente a un segundo resto de proteína para producir constructos de fusión de la estructura general:

NH_2-VH-CH-segunda proteína-COOH y NH_2-VL-CL-COOH

La segunda proteína puede fusionarse al extremo o bien N-terminal o bien C-terminal de o bien la cadena pesada o bien la ligera.

En una realización preferida, la segunda proteína del constructo de fusión de región de marco de Fab o de cadena sencilla es una proteína que proporciona una lectura para ensayos intracelulares, o bien directamente o bien mediante activación transcripcional.

Otro objeto de la invención es proporcionar clases de región de marco de dominios variables de anticuerpo y secuencias de dominios variables y scFv que son adecuados para el injerto de los bucles hipervariables de anticuerpo existentes, por ejemplo, con el fin de obtener anticuerpos que son funcionales en un entorno reductor o de otro modo desafiante.

Otro objeto de la invención es proporcionar clases de región de marco de dominios variables de anticuerpo y secuencias de dominios variables y scFv que, por ejemplo, a través de aleatorización de uno o más de los bucles hipervariables de tales regiones de marco, son adecuados para la creación de bibliotecas para su uso en un entorno reductor o de otro modo desafiante.

Otro objeto de la invención es el uso de las secuencias descritas en la identificación de secuencias consenso y residuos conservados.

Los anticuerpos o fragmentos de anticuerpo que resultan del uso de las regiones de marco descritas pueden usarse como reactivos en la validación de diana y en terapia, prevención y diagnóstico de enfermedades de seres humanos, animales y plantas. Los anticuerpos pueden usarse en forma de proteína o ADN que codifica para tal proteína y no se limitan a aplicaciones intracelulares.

Breve descripción de los dibujos

La figura 1 muestra el resultado de una selección de "control de calidad" típica en levadura sometida a ensayo mediante la activación de la expresión de lacZ (véase, por ejemplo, el ejemplo 1). Se identificaron los clones positivos, seleccionados (negro) en varias selecciones diferentes y las secuencias correspondientes de los clones positivos pueden hallarse en las tablas 1 y 2. Las secuencias seleccionadas se comparan con el control positivo, el injerto de lambda muy estable (gris oscuro).

La figura 2 muestra el funcionamiento de las regiones de marco aisladas de una selección de "control de calidad" típica en levadura (negro) en la línea celular humana Hela, sometida a ensayo mediante la activación de la expresión de luciferasa en comparación con el injerto de lambda muy estable (gris oscuro). El control positivo Gal4-VP16 (blanco) proporciona el máximo nivel posible de activación transcripcional en el sistema. Se ha corregido la actividad luciferasa para la eficacia de transfección.

La figura 3 muestra el funcionamiento *in vivo* de las combinaciones de regiones de marco superiores sometidas a ensayo en levadura mediante la activación de la expresión de lacZ. Las secuencias de región de marco (negro) se comparan con el control positivo (el injerto de lambda muy estable (gris oscuro)). La numeración de las regiones de marco es tal como se describe en la tabla 5.

La figura 4 muestra el funcionamiento *in vivo* de las combinaciones de regiones de marco superiores sometidas a ensayo en la línea celular humana Hela mediante la activación de la expresión de luciferasa e ilustradas en comparación con el injerto de lambda muy estable (gris oscuro). El control positivo, Gal4-VP16 (blanco) proporciona el máximo nivel posible de activación transcripcional en el sistema. Se ha corregido la actividad luciferasa para la eficacia de transfección.

La figura 5 muestra el funcionamiento *in vivo* de las combinaciones de regiones de marco superiores sometidas a ensayo mediante la cantidad de proteína soluble producida en el citoplasma de la cepa de levadura *S. cerevisiae* JPY9.

La figura 6 muestra el comportamiento de expresión de combinaciones de regiones de marco seleccionadas en el periplasma de *E. coli*. La flecha indica la ubicación de la banda correspondiente a las regiones de marco de scFv.

La figura 7 muestra el funcionamiento *in vivo* de combinaciones de regiones de marco seleccionadas sometidas a ensayo en tres líneas celulares humanas (Hela, (negro), Saos-2 (gris oscuro) y HEK 293 (blanco)), mediante la activación de la expresión de luciferasa e ilustradas en comparación con el injerto de lambda muy estable. El control positivo Gal4-VP16 proporciona el máximo nivel posible de activación transcripcional en el sistema. Se ha corregido la actividad luciferasa para la eficacia de transfección.

La figura 8 representa la resistencia frente a la agregación a 37°C de combinaciones de regiones de marco seleccionadas cuantificadas mediante la cantidad de proteína monomérica presente antes y después de la incubación tal como se indica en tampón PBS. El panel A es representativo de las regiones de marco 2.4 y 5.2 y el panel B de las regiones de marco 4.4, 6.4 y 7.3.

La figura 9 representa la resistencia frente a la agregación mediante degradación por proteasas en suero humano a 37°C de combinaciones de regiones de marco seleccionadas, cuantificadas mediante la cantidad de proteína de longitud completa soluble presente antes y después de incubación prolongada.

La figura 10 muestra el funcionamiento *in vivo* de dos grupos de unión seleccionados en la región de marco novedosa 7.3 en el contexto de Fab, sometidos a ensayo en el ensayo de interacción en levadura mediante la activación de la expresión de lacZ. La expresión de las cadenas de Fab es a partir de un promotor inducible por galactosa bidireccional, en vectores o bien ars/cen o bien de 2 micrómetros. La expresión a partir del vector de Fab produce la cadena ligera de anticuerpo y una proteína de fusión VH-CH1-Gal4-AD. Los grupos de unión se dirigen contra la cinasa 1 de tipo Polo humana (hPLK1). La unión a la diana se compara con la unión no específica a un antígeno no relacionado y la unión de la región de marco no aleatorizada 7.3. Obsérvese que el scFv correspondiente que se ha incluido para referencia se expresa a partir de un promotor de actina (2 micrómetros).

La figura 11 muestra el funcionamiento *in vivo* de las regiones de marco scFv en el contexto de Fab sometidas a ensayo mediante la cantidad de proteína soluble producida en el citoplasma de la cepa de levadura JPY9. La expresión de la fusión Gal4-AD-scFv (actina/2 micrómetros) se compara con la expresión del constructo de Fab correspondiente y con la región de marco original 7.3 como Fab, ambos procedentes de dos vectores diferentes (inducible por Gal, ars/cen y 2 micrómetros). La expresión a partir del vector de Fab produce la cadena ligera de anticuerpo y una proteína de fusión VH-CH1-Gal4-AD, que se detecta en esta inmunotransferencia.

La tabla 1 muestra una alineamiento de todas las secuencias de región de marco de dominios VH seleccionadas de diversas selecciones de "control de calidad" en levadura.

La tabla 2 muestra una alineamiento de todas las secuencias de región de marco de dominios VL seleccionadas de diversas selecciones de "control de calidad" en levadura.

La tabla 3 muestra una alineamiento de secuencias recogidas aleatoriamente de la biblioteca.

La tabla 4 muestra un análisis estadístico de la frecuencia de subclases para dominios VH y VL en las secuencias aisladas con el sistema de "control de calidad". Sólo se consideraron aquellas secuencias que se encontró posteriormente que eran positivas en el ensayo cuantitativo en levadura. Las secuencias seleccionadas se comparan con la biblioteca no seleccionada tal como se determina a partir de un número limitado de secuencias aleatorias (tabla 3).

La tabla 5 muestra las secuencias usadas para la recombinación adicional y evaluación de las mejores combinaciones en scFv y sus abreviaturas (abr.), fuentes y subfamilia respectivas.

Descripción detallada de la invención

A menos que se defina de otro modo, todos los términos científicos y técnicos usados en el presente documento tienen el mismo significado que entiende comúnmente un experto habitual en la técnica a la que pertenece esta invención. Aunque pueden usarse materiales y métodos similares o equivalentes a los descritos en el presente documento, en la práctica o las pruebas de la invención, se describen a continuación materiales y métodos adecuados. Todas las publicaciones, solicitudes de patente, patentes y otras referencias mencionadas en el presente documento se incorporan como referencia en su totalidad. En caso de conflicto, dominará la presente memoria descriptiva, incluyendo las definiciones. Además, los materiales, métodos, y ejemplos son ilustrativos únicamente y no pretenden ser limitativos.

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

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

"Dominio VH" se refiere a la parte variable de la cadena pesada de una molécula de inmunoglobulina.

"Dominio VL" se refiere a la parte variable de la cadena ligera de una molécula de inmunoglobulina.

"Subtipo" VH o VL se refiere al subtipo definido por la secuencia consenso respectiva tal como se define en Knappik (2000). El término "subfamilia" o" subclase" se usa como sinónimo de "subtipo". El término "subtipo" tal como se usa

en el presente documento se refiere a secuencias que comparten un alto grado de identidad y similitud con la secuencia consenso respectiva que representa su subtipo. Se determina si una determinada secuencia de dominios variables pertenece a un "subtipo" mediante el alineamiento de la secuencia con o bien cualquier segmento de línea germinal humana conocido del dominio respectivo, o bien las secuencias consenso definidas y la posterior identificación de la mayor homología. Los métodos para determinar homologías y agrupar secuencias usando matrices de búsqueda, tales como BLOSUM (Henikoff 1992) se conocen bien por los expertos en la técnica.

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

La secuencia de aminoácidos de una proteína puede analizarse a diversos niveles. Por ejemplo, podría presentarse conservación o variabilidad a nivel de residuos individuales, nivel de múltiples residuos, múltiples residuos con huecos, etc. Los residuos podrían presentar conservación del residuo idéntico o podrían estar conservados a nivel de clase. Los ejemplos de clases de aminoácidos incluyen grupos R polares pero no cargados (serina, treonina, asparagina y glutamina); grupos R cargados positivamente (lisina, arginina e histidina); grupos R cargados negativamente (ácido glutámico y ácido aspártico); grupos R hidrófobos (alanina, isoleucina, leucina, metionina, fenilalanina, triptófano, valina y tirosina); y aminoácidos especiales (cisteína, glicina y prolina). Otras clases las conoce un experto en la técnica y pueden definirse usando determinaciones estructurales u otros datos para evaluar la capacidad de sustitución. En ese sentido, un aminoácido sustituible podría referirse a cualquier aminoácido que pudiera sustituirse y mantener la conservación funcional en esa posición.

"Secuencia consenso de polinucleótido" tal como se usa en el presente documento se refiere a una secuencia de nucleótidos, que puede generarse usando una matriz de al menos dos o preferiblemente más secuencias de ácido nucleico alineadas, y permitiendo huecos en el alineamiento, es posible determinar el residuo de nucleótido más frecuente en cada posición. La secuencia consenso es aquella secuencia que comprende los nucleótidos que se representan con la mayor frecuencia en cada posición. En el caso de que dos o más nucleótidos se representen de igual manera en una única posición, la secuencia consenso incluye ambos o todos esos nucleótidos.

"Subelemento estructural" tal como se usa en el presente documento se refiere a tramos de residuos de aminoácido dentro de una proteína o polipéptido que corresponden a una parte estructural o funcional definida de la molécula. Éstas pueden ser bucles (es decir, bucles de CDR de un anticuerpo) o cualquier otra estructura secundaria o funcional dentro de la proteína o polipéptido (es decir, dominios, hélices α, láminas β, regiones de marco de anticuerpos, etc.). Un subelemento estructural puede identificarse usando estructuras conocidas de polipéptidos similares u homólogos, o usando las matrices de secuencias de aminoácidos alineadas mencionadas anteriormente. En este caso, la variabilidad en cada posición es la base para determinar tramos de residuos de aminoácido que pertenecen a un subelemento estructural (por ejemplo, regiones hipervariables de un anticuerpo).

"Subsecuencia" tal como se usa en el presente documento se refiere a un módulo genético que codifica al menos un subelemento estructural. No es necesariamente idéntico a un subelemento estructural.

"CDR de anticuerpo" tal como se usa en el presente documento se refiere a las regiones determinantes de complementariedad del anticuerpo que consisten en bucles de unión a antígenos tal como se definen por Kabat *et al.* (1991). Cada uno de los dos dominios variables de un fragmento Fv de anticuerpo contienen, por ejemplo, tres CDR.

"Anticuerpo" tal como se usa en el presente documento es un sinónimo de "inmunoglobulina". Los anticuerpos según la presente invención pueden ser inmunoglobulinas completas o fragmentos de las mismas, que comprenden al menos un dominio variable de una inmunoglobulina, tal como dominios variables sencillos, Fv (Skerra, 1988), scFv (Bird, 1988; Huston, 1988), Fab, $(Fab')_2$ u otros fragmentos bien conocidos por un experto en la técnica.

"Región de marco de anticuerpo" tal como se usa en el presente documento se refiere a la parte del dominio variable, o bien VL o bien VH, que sirve como andamiaje para los bucles de unión a antígenos de este dominio variable (Kabat *et al.*, 1991).

Los fragmentos scFv modificados mediante ingeniería genética de manera racional han demostrado una clara correlación entre la estabilidad termodinámica de un fragmento scFv y su funcionamiento *in vivo* (Wörn, 2000; Auf der Maur, 2001). Usando un sistema desarrollado recientemente denominado "control de calidad" (Auf der Maur, 2001), se han aislado secuencias de región de marco de dominios variables de anticuerpo específicas que son adecuadas para aplicaciones intracelulares (tablas 1 y 2), se han caracterizado (figuras 1 y 2) y se han mejorado adicionalmente (figuras 3 a 9 y tabla 3). Tal como se observa en los experimentos previos, las regiones de marco que funcionan bien seleccionadas en el ensayo intracelular muestran una alta estabilidad *in vitro* tal como se demuestra por su resistencia a la agregación y degradación por proteasas a 37°C (figuras 8 y 9). Además, apareció un patrón que permite una selección de regiones de marco para aplicaciones intracelulares con una base más

general, dependiendo de su subfamilia de región de marco (tabla 4). En el presente documento, se dan a conocer secuencias de dominios variables de anticuerpo específicas útiles para aplicaciones intracelulares, así como el patrón general. Esto permite, por un lado, el uso de estas secuencias como donadores de región de marco en experimentos de injerto para obtener intracuerpos funcionales que retienen la especificidad de unión del donador de bucle. Adicionalmente, pueden construirse bibliotecas de anticuerpos usando las secuencias dadas a conocer como regiones de marco. Tales bibliotecas son adecuadas para sistemas de selección intracelulares en condiciones reductoras, tales como las que hay en células procariotas y eucariotas. Adicionalmente, las secuencias dadas a conocer pueden usarse para identificar, por ejemplo, secuencias o residuos o motivos conservados. El injerto de subelementos estructurales, por ejemplo, los de los bucles de unión de un anticuerpo (por ejemplo Jung, 1997), así como la preparación de bibliotecas de anticuerpos o fragmentos de los mismos (por ejemplo Vaughan, 1996; Knappik, 2000) se ha descrito en detalle y lo conoce bien un experto en la técnica.

Dado que las aplicaciones intracelulares exponen los fragmentos de anticuerpo a condiciones muy desfavorables (es decir temperaturas aumentadas, entorno reductor), las secuencias dadas a conocer en la presente invención han adquirido características que las hacen resistentes a las condiciones más adversas. Por tanto, cuando se comparan con secuencias "promedio", las secuencias dadas a conocer son de estabilidad y solubilidad excepcionales tal como se demuestra mediante su resistencia frente a la agregación y la degradación por proteasas (figuras 8 y 9). Estas características, junto con su excelente rendimiento de expresión hacen que las secuencias de región de marco de anticuerpo dadas a conocer sean adecuadas de forma única no sólo para uso intracelular, sino especialmente para todas las aplicaciones terapéuticas y de diagnóstico en las que una larga semivida, la robustez y la facilidad de producción son motivo de gran preocupación.

La presente invención permite el diseño de secuencias de polipéptido que comprenden al menos la parte variable de un anticuerpo que son útiles para aplicaciones en un entorno reductor o de otro modo desafiante. En una primera realización, la invención proporciona una colección de secuencias de región de marco de anticuerpo útiles para aplicaciones intracelulares (tablas 1 y 2). En una primera etapa, se selecciona una biblioteca de diversas secuencias independientemente de la afinidad de unión usando el sistema de Control de calidad en levadura. Pueden evaluarse las secuencias aisladas para determinar su funcionamiento intracelular en levadura y en células de mamífero (figuras 1 y 2).

En una realización de la invención, se analiza la colección de secuencias aisladas mediante alineamiento para identificar las subclases de dominios variables de anticuerpo y las secuencias consenso que son adecuadas para aplicaciones intracelulares.

En una realización adicional preferida de la invención, la colección de secuencias de región de marco de anticuerpo descritas anteriormente se analizan adicionalmente mediante el alineamiento de unas con otras y la agrupación en subfamilias. Todas las regiones de marco pertenecientes a un subtipo se comparan con respecto a su funcionamiento intracelular en levadura y en células de mamífero (figuras 1 y 2, como un ejemplo) y con respecto a la aparición de intercambios negativos, neutros o positivos en su secuencia de aminoácidos con relación a la secuencia consenso de subtipo respectiva. Un experto en la técnica puede distinguir entre cambios positivos, neutros y negativos basándose en el entorno estructural del residuo intercambiado particular en el dominio de inmunoglobulina. Posteriormente, se eligen secuencias de región de marco de dominios de anticuerpo variables que muestran el mejor funcionamiento intracelular y que carecen de intercambios negativos en comparación con su secuencia consenso de subtipo respectiva. Preferiblemente, se seleccionan secuencias que contienen adicionalmente intercambios de aminoácidos que se consideran positivos.

En una realización adicional preferida, los dominios variables de anticuerpo seleccionados de la cadena pesada y ligera se recombinan posteriormente en todas las combinaciones posibles para dar fragmentos scFv, con el fin de identificar las combinaciones con la mayor estabilidad y solubilidad. Para este fin, se evalúan los fragmentos scFv recombinados, novedosos para determinar su funcionamiento en condiciones reductoras en ensayos de interacción intracelulares en levadura (figura 3) y en líneas celulares de mamífero (figuras 4 y 7) y para la expresión intracelular soluble en levadura (figura 5). Se evaluaron adicionalmente combinaciones prometedoras para determinar su comportamiento en condiciones oxidantes analizando el rendimiento de expresión periplasmática en *E. coli* (figura 6), la resistencia a la agregación a temperaturas elevadas (figura 8) y la resistencia a la agregación y la degradación por proteasas tras incubación prolongada en suero humano a 37°C (figura 9). Estos datos se usan para identificar la región de marco de scFv más adecuada para cualquier aplicación específica, o bien intracelular, o bien en condiciones oxidantes.

Las secuencias de región de marco seleccionadas y optimizadas dadas a conocer en el presente documento tienen una ventaja significativa no sólo en aplicaciones intracelulares, sino en todas las aplicaciones que pueden beneficiarse de una estabilidad y/o solubilidad aumentada del scFv. Son ejemplos el almacenamiento a largo plazo a altas concentraciones requerido para aplicaciones de diagnóstico, y la semivida funcional prolongada en suero a 37°C (tal como se requiere, por ejemplo, en aplicaciones terapéuticas).

Según un aspecto de la presente invención, se proporciona una región de marco de intracuerpo que comprende una región de marco de cadena sencilla que tiene la estructura general:

NH_2-VL-enlazador-VH-COOH; o

NH_2-VH-enlazador-VL-COOH

en la que la región de marco de VH es de subtipo 1a, 1b o 3.

En otra realización, se invierte la orientación de las regiones VH y VL en la región de marco de cadena sencilla descrita anteriormente.

Según un aspecto de la presente invención, se proporciona una región de marco de intracuerpo que comprende una región de marco de cadena sencilla que tiene la estructura general:

NH_2-VL-enlazador-VH-COOH; o

NH_2-VH-enlazador-VL-COOH

en la que la región de marco de VH es de subtipo 1a, 1b o 3 y la región de marco de VL es de subtipo λ1, λ3 o κ1.

En otra realización, la invención proporciona una región de marco de cadena sencilla fusionada a un segundo resto de proteína para producir un constructo de fusión de la estructura general:

NH_2-VL-enlazador-VH-segunda proteína-COOH; o

NH_2-segunda proteína-VL-enlazador-VH-COOH

en la que la región de marco de VH es de subtipo 1a, 1b o 3 y la región de marco de VL es de subtipo λ1, λ3 o κ1.

En otra realización, puede invertirse la orientación de las regiones VH y VL en estos constructos de fusión.

En otra realización, pueden incorporarse los dominios variables en un fragmento Fab que puede fusionarse adicionalmente a un segundo resto de proteína para producir constructos de fusión de la estructura general:

NH_2-VH-CH-segunda proteína-COOH y NH_2-VL-CL-COOH

La segunda proteína puede fusionarse al extremo o bien N-terminal o bien C-terminal de o bien la cadena pesada o bien la ligera.

Tal como se da a conocer en el presente documento, existe una preferencia muy fuerte en aplicaciones intracelulares por la región de marco de VH del subtipo 3, pero también por 1a y 1b. Con respecto al dominio variable de la cadena ligera (VL), existe una clara preferencia por los números para regiones de marco del tipo kappa 1, pero también están enriquecidos lambda 1 y 3. Estos subtipos de región de marco, es decir VH 1a, 1b y 3 combinados con un dominio VL kappa 1, lambda 1 ó 3 son, por tanto, los más adecuados para uso intracelular y otras aplicaciones que requieren las propiedades de plegamiento del scFv. Por tanto, con el fin de reducir la cantidad de moléculas que no son funcionales en el entorno reductor, deben construirse preferentemente bibliotecas para sistemas de selección intracelulares a partir de una mezcla de estos subtipos de región de marco.

En una realización preferida, el dominio VH de los fragmentos de anticuerpo de la invención es del subgrupo 1a, 1b o 3.

En una realización preferida, el dominio VL de los fragmentos de anticuerpo de la invención es del subgrupo kappa 1, lambda 1 ó 3.

En una realización preferida, los fragmentos de anticuerpo usados como regiones de marco se seleccionan del grupo que consiste en: 1.1, 2.1, 3.1, 4.1, 5.1, 1.2, 2.2, 3.2, 4.2, 5.2, 1.3, 2.3, 3.3, 4.3, 5.3, 7.3, 1.4, 2.4, 3.4, 4.4, 5.4 y 6.4 tal como se describe en la tabla 5.

En una realización de la invención, se identifican al menos dos y preferiblemente más regiones de marco y luego se analizan. Puede establecerse una base de datos de las secuencias de proteína en la que se alinean unas con otras las secuencias de proteína. El alineamiento puede usarse entonces para definir, por ejemplo, residuos, subelementos, subsecuencia o subgrupos de secuencias de región de marco que muestran un alto grado de similitud tanto en la secuencia como, si esa información está disponible, en la disposición estructural.

La longitud de los subelementos oscila preferiblemente, pero no exclusivamente entre 1 aminoácido (tal como un residuo en el sitio activo de una enzima o un residuo determinante de la estructura) y 150 aminoácidos (por ejemplo, dominios de proteína completos). Lo más preferiblemente, la longitud oscila entre 3 y 25 aminoácidos, tal como se encuentra de la manera más común en los bucles de CDR de anticuerpo.

En otra realización, se sintetizan secuencias de ácido nucleico consenso, que se predicen a partir del análisis. Esto puede lograrse mediante uno cualquiera de los diversos métodos bien conocidos por el profesional experto en la técnica, por ejemplo, mediante enfoques de síntesis génica total o basados en PCR.

En otra realización, las secuencias de ácido nucleico se clonan en un vector. El vector podría ser un vector de secuenciación, un vector de expresión o un vector de presentación (por ejemplo, presentación en fago), que se conocen todos por los expertos en la técnica. Un vector podría comprender una secuencia de ácido nucleico, o dos o más secuencias nucleicas, o bien en un operón diferente o bien en el mismo. En el último caso, podrían clonarse o bien por separado o bien como secuencias contiguas.

En una realización, los polipéptidos tienen un patrón de aminoácidos característico de una especie particular. Esto puede lograrse, por ejemplo, deduciendo las secuencias consenso de una colección de proteínas homólogas de sólo una especie, lo más preferiblemente de una colección de proteínas humanas.

Una realización adicional de la presente invención se refiere a proteínas de fusión proporcionando una secuencia de ADN que codifica tanto para el polipéptido, tal como se describió anteriormente, así como para un resto adicional.

En realizaciones adicionales, la invención proporciona secuencias de ácido nucleico, vectores que contienen las secuencias de ácido nucleico, células huésped que contienen los vectores, y polipéptidos que pueden obtenerse según los métodos descritos en el presente documento.

En una realización adicional, la invención proporciona la síntesis o colocación de otro modo de sitios de restricción en el extremo de las secuencias de ácido nucleico de la invención, permitiendo que se clonen en vectores adecuados.

En una realización adicional preferida, la invención proporciona sistemas de vector que son compatibles con las secuencias de ácido nucleico que codifican para los polipéptidos. Los vectores comprenden sitios de restricción, que serían, por ejemplo, únicos dentro del sistema de vector y esencialmente únicos con respecto a los sitios de restricción incorporados en las secuencias de ácido nucleico que codifican para los polipéptidos, excepto por ejemplo los sitios de restricción necesarios para clonar las secuencias de ácido nucleico en el vector.

En otra realización, la invención proporciona un kit, que comprende una o más de la lista de secuencias de ácido nucleico, vectores recombinantes, polipéptidos y vectores según los métodos descritos anteriormente, y, por ejemplo, células huésped adecuadas para producir los polipéptidos.

Todas las realizaciones anteriores de la presente invención pueden efectuarse usando técnicas convencionales de biología molecular conocidas por un experto en la técnica.

En otra realización, la secuencia de ácido nucleico es cualquier secuencia que puede codificar para los polipéptidos de la invención.

En otra realización, los ácidos nucleicos de la invención se usan en terapia génica.

En otra realización, la región de marco de cadena sencilla es una variante de una cualquiera de las secuencias 1.1, 2.1, 3.1, 4.1, 5.1, 1.2, 2.2, 3.2, 4.2, 5.2, 1.3, 2.3, 3.3, 4.3, 5.3, 7.3, 1.4, 2.4, 3.4, 4.4, 5.4, 6.4 (tabla 5), en las que "variante" tal como se usa en el presente documento se refiere a una secuencia que presenta una identidad del 90% o mayor, mientras que mantiene una estabilidad potenciada.

En otra realización, la región de marco de cadena sencilla es un derivado de una cualquiera de secuencias 1.1, 2.1, 3.1, 4.1, 5.1, 1.2, 2.2, 3.2, 4.2, 5.2, 1.3, 2.3, 3.3, 4.3, 5.3, 7.3, 1.4, 2.4, 3.4, 4.4, 5.4, 6.4 (tabla 5) en la que "derivado" tal como se usa en el presente documento se refiere a una secuencia que mantiene sólo aquellos aminoácidos que son críticos para la función y la estabilidad de la molécula. Los intercambios neutros o positivos aislados en la región de marco tal como se describen en el ejemplo 3, no se considera que sean un cambio relevante en las regiones de marco de anticuerpo de la presente invención.

En una realización preferida de la invención, la región de marco de cadena sencilla se fusiona a una segunda proteína, en la que esa proteína proporciona una lectura para ensayos intracelulares. La lectura puede ser o bien directa, por ejemplo en forma de una fusión a una proteína detectable, por ejemplo GFP (proteína fluorescente verde), proteína fluorescente azul potenciada, proteína fluorescente amarilla potenciada, proteína fluorescente cian potenciada que puede observarse mediante fluorescencia, u otras parejas de fusión con diferentes métodos de detección. Alternativamente, puede lograrse una lectura a través de la activación transcripcional de un gen indicador, en el que la pareja de fusión en la proteína de fusión de scFv es o bien un activador transcripcional, tal como el dominio de activación Gal4, o una proteína de unión a ADN, tal como el dominio de unión a ADN de Gal4 o LexA, que activa la transcripción de un gen indicador de una enzima, tal como β-galactosidasa, luciferasa, α-galactosidasa, β-glucuronidasa, cloranfenicol acetiltransferasa y otras, que proporcionan a su vez una lectura. Las proteínas de fusión, que proporcionan una lectura las conoce bien un experto en la técnica.

Otra realización de la invención es un anticuerpo que comprende una región de marco descrita en el presente documento.

Otra realización de la invención es el uso del anticuerpo de la presente invención.

Una realización adicional preferida de la invención es el uso de las clases de región de marco descritas anteriormente de dominios variables de anticuerpo y secuencias de dominios variables y scFv para el injerto de bucles hipervariables de anticuerpos existentes, con el fin de obtener anticuerpos que son funcionales en un entorno reductor o de otro modo desafiante.

Otra realización adicional preferida de la invención es el uso de las clases de región de marco descritas de dominios variables de anticuerpo y secuencias de dominios variables y scFv, por ejemplo a través de aleatorización de uno o más de los bucles hipervariables de tales regiones de marco, para la creación de bibliotecas para aplicaciones en un entorno reductor o de otro modo desafiante.

Tal como resultaría evidente para un experto habitual en la técnica, las moléculas de la invención descritas en el presente documento pueden usarse en aplicaciones de diagnóstico y terapéuticas, validación de dianas y terapia génica.

La invención puede ilustrarse mediante los siguientes ejemplos, que no pretenden limitar el alcance de la invención en modo alguno.

Bibliografía

Agatep, R., Kirkpatrick, D. L., Parchaliuk, R. A., Woods y Gietz, R. D. (1998). "Transformation of Saccharomyces cerevisiae by lithium acetate/singlestranded carrier DNA/polyethylene glycol protocol." Technical Tips Online (http ://tto. trends. com).

Auf der Maur, A., Escher, D. y Barberis, A. (2001). "Antigen-independent selection of stable intracellular single-chain antibodies." FEBS Lett 508: 407-412.

Benvenuto, E. y Tavladoraki, P. (1995). "Immunotherapy of plant viral diseases." Trends Microbiol 3 (7): 272-5.

Biocca, S. y Cattaneo, A. (1995) "Intracellular immunization: antibody targeting to subcellular compartments." Trends Cell Biol 5: 248-252.

Biocca, S., Di Luzio, A., Werge, T. y Cattaneo, A. (1991). "Intracellular immunization: expression of antibody domains in the cytoplasm and in the nucleus of mammalian cells." Cytotechnology 5 Suppl 1: 49-50.

Biocca, S., Pierandrei-Amaldi, P. y Cattaneo, A. (1993). "Intracellular expression of anti-p21ras single chain Fv fragments inhibits meiotic maturation of xenopus oocytes." Biochem Biophys Res Commun 197 (2): 422-7.

Biocca, S., Ruberti, F., Tafani, M., Pierandrei-Amaldi, P. y Cattaneo, A. (1995). "Redox state of single chain Fv fragments targeted to the endoplasmic reticulum, cytosol and mitochondria." Bio/Technology 13 (10): 1110-5.

Bird, R. E., Hardman, K. D., Jacobson, J. W., Johnson, S., Kaufman, B. M., Lee, S. M., Lee, T., Pope, S. H., Riordan, G. S. y Whitlow, M. (1988). "Single-chain antigen-binding proteins." Science 242 (4877): 423-6.

Chothia, C. y Lesk, A. M. (1987). "Canonical structures for the hypervariable regions of immunoglobulins."J. Mol. Biol. 196 (4): 901-17.

Chothia, C., Lesk, A. M., Tramontano, A., Levitt, M., Smith-Gill, S. J., Air, G., Sheriff, S., Padlan, E. A., Davies, D., Tulip, W. R., Colman, W. R., Spinelli, S., Alzari, P. M. y Poljak, R. J. (1989). "Conformations of immunoglobulin hypervariable regions." J. Mol. Biol. 342: 877-883.

Johnson, G., Kabat, E. A. y Wu, T. T. (1996). kabat database of sequences of proteins of immunological interest. WEIR'S Handbook of experimental Immunology I Immunochemistry and Molecular Immunology. Cambridge, MA, Blackwell Science Inv: 6.1-6.21.

Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S. y Foeller, C. (1991). Sequences of proteins of immunological interest. NIH Publication 91-3242. Washington, DC, US Department of health and human services.

Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., Hoess, A., Wolle, J., Plückthun, A. y Virnekas, B. (2000). "Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides."J Mol Biol 296 (1): 57-86.

Cochet, 0., Kenigsberg, M., Delumeau, I., Virone-Oddos, A., Multon, M. C., Fridman, W. H., Schweighoffer, F., Teillaud, J. L. y Tocque, B. (1998). "Intracellular expression of an antibody fragment-neutralizing p21 ras promotes tumor regression." Cancer Res 58 (6): 1170-6.

Corpet, F. (1988). "Multalin. Multiple sequence alignment with hierarchical clustering." Nucleic Acids Res 16 (22): 10881-10890.

Cox, J. P., Tomlinson, I. M. y Winter, G. (1994). "A directory of human germ-line V kappa segments reveals a strong bias in their usage." Eur J Immunol 24 : 827-836.

de Haard, H. J., van Neer, N., Reurs, A., Hufton, S. E., Roovers, R. C., Henderik, P., de Bruine, A. P., Arends, J. -W. y Hoogenboom, H. R. (1999). "a large non-immunized human Fab fragment phage library that permits rapid isolation and kinetic analysis of high affinity antibodies." J. Biol. Chem. 274: 18218-18230.

Escher, D., Bodmer-Glavas, M., Barberis, A. y Schaffner, W. (2000). "Conservation of glutamine-rich transactivation function between yeast and humans."Mol Cell Biol 20 (8): 2774-82.

Escher, D. y Schaffner, W. (1996). "Improved "activator trap" method for the isolation of transcriptional activation domains from random DNA fragments."Biotechniques 21 (5): 848-54.

Krebber, A., Bornhauser, S., Burmester, J., Honegger, A., Willuda, J., Bosshard, H. R. y Plückthun, A. (1997). "Reliable cloning of functional variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system." J Immunol Methods 201: 35-55.

Lener, M., Horn, I. R., Cardinale, A., Messina, S., Nielsen, U. B., Rybak, S. M., Hoogenboom, H. R., Cattaneo, A. y Biocca, S. (2000). "Diverting a protein from its cellular location by intracellular antibodies. The case of p2lRas." Eur J Biochem 267 (4): 1196-205.

Marasco, W. A. (1997). "Intrabodies: turning the humoral immune system outside in for intracellular immunization." Gene Ther 4 (1): 11-5.

Martineau, P., Jones, P. y Winter, G. (1998). "Expression of an antibody fragment at high levels in the bacterial cytoplasm."J. Mol. Biol. 280 (1): 117-27.

Ohage, E. y Steipe, B. (1999). "Intrabody construction and expression. I. The critical role of VL domain stability."J Mol Biol 291 (5): 1119-28.

Ohage, E. C., Wirtz, P., Barnikow, J. y Steipe, B. (1999). "Intrabody construction and expression. II. A synthetic catalytic Fv fragment."J Mol Biol 291 (5): 1129-34.

Plückthun, A., Krebber, A., Krebber, C., Horn, U., Kupfer, U., Wenderoth, R., Nieba, L., Proba, K. y Riesenberg, D. (1996). Producing antibodies in Escherichia coli: from PCR to fermentation. Antibody Engineering, a practical approach. Oxford, Oxford University Press.

Proba, K., Honegger, A. y Plückthun, A. (1997). "A natural antibody missing a cysteine in VH: consequences for thermodynamic stability and folding." J. Mol. Biol. 265 (2): 161-72.

Richardson, J. H. y Marasco, W. A. (1995). "Intracellular antibodies: development and therapeutic potential." Trends Biotechnol 13 (8): 306-10.

Rondon, I. J. y Marasco, W. A. (1997). "Intracellular antibodies (intrabodies) for gene therapy of infectious diseases." Annu Rev Microbiol 51: 257-83.

Skerra, A. y Plückthun, A. (1988). "Assembly of a functional immunoglobulin Fv fragment in Escherichia coli." Science 240 (4855): 1038-41.

Tavladoraki, P., Benvenuto, E., Trinca, S., De Martinis, D., Cattaneo, A. y Galeffi, P. (1993). "Transgenic plants expressing a functional single-chain Fv antibody are specifically protected from virus attack." Nature 366 (6454): 469-72.

Tomlinson, I. M., Walter, G., Marks, J. D., Llewelyn, M. B. y Winter, G. (1992). "The repertoire of human germline VH sequences revealy about 50 groups of VH segments with different hypervariable loops." J Mol Biol 227: 776-798.

Tse, E., Lobato, M. N., Forster, A., Tanaka, T., Chung, G. y Rabbitts, T. (2002). "Intracellular antibody capture technology: application to selection of intracellular antibodies recognising the BCR-ABL oncogenic protein." J. Mol Biol 317 ((1)): 85-94.

Verma, R., Boleti, E. y George, A. J. (1998). "Antibody engineering: comparison of bacterial, yeast, insect and mammalian expression systems." J Immunol Methods 216 (1-2): 165-81.

Visintin, M., Settanni, G., Maritan, A., Graziosi, S., Marks, J. D. y Cattaneo, A. (2002). "The intracellular antibody capture technology (IACT): towards a consensus sequence for intracellular antibodies."J. Mol. Biol. 317 ((I)): 73-83.

Visintin, M., Tse, E., Axelson, H., Rabbitts, T. H. y Cattaneo, A. (1999). "Selection of antibodies for intracellular function using a two-hybrid in vivo system." Proc Natl Acad Sci U S A 96 (21): 11723-8.

Welschhof, M., Ternesss, P., Kolbinger, F., Zewe, M., Dubel, S., Dorsam, H., Hain, C., Finger, M., Jung, M., Moldenhauer, G., Hayashi, N., Little, M. y Opelz, G. (1995). "Amino acid sequence based PCR primers for amplification of rearranged human heavy and light chain immunoglobulin variable region genes." J Immunol Methods 179: 203-214.

Williams, S. C., Frippiat, J. P., Tomlinson, I. M., Ignatovic, 0., Lefranc, M. P. y Winter, G. (1996). "Sequence and evolution of the human germline V-lambda repertoire." J Mol Biol 264: 220-232.

Wörn, A., Auf der Maur, A., Escher, D., Honegger, A., Barberis, A. y Plückthun, A. (2000). "Correlation between in vitro stability and in vivo performance of anti-GCN4 intrabodies as cytoplasmic inhibitors." J. Biol. Chem. 275 (4): 2795-803.

Yokota, T., Milenic, D. E., Whitlow, M. y Schlom, J. (1992). "Rapid tumor penetration of a single-chain Fv and comparison with other immunoglobulin forms." Cancer Res 52 (12): 3402-8.

Auf der Maur, A., Zahnd, C., Fischer, F., Spinelli, S., Honegger, A., Cambillau, C., Escher, D., Plückthun, A. y Barberis, A. (2002). "Direct in vivo screening of intrabody libraries constructed on a highly stable single-chain framework." J Biol Chem 277 (47): 45075-45085.

Gietz, R. D. y Sugino, A. (1988). "New yeast-Escherichia coli shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites" Gene 74: 527-534.

Willuda, J., Honegger, A.; Waibel, P., Schubiger, A. Stahel, R., Zangmeister-Wittke, U. y Plückthun, A. (1999). "High thermal stability is essential for tumor targeting of antibody fragments: engineering of a humanized anti-epithelial glycoprotein-2 (epithelial cell adhesion molecule) single-chain Fv fragment" Cancer Research 59: 5758-5767.

Wörn, A. y Plückthun, A. (1999). "Different equilibrium stability behavior of scFv fragments: identification, classification, and improvement by protein engineering" Biochemistry 38: 8739-8750.

La invención se ilustra adicionalmente en los siguientes ejemplos no limitativos.

Ejemplo 1

Selección de regiones de marco de intracuerpo a través de la selección de una biblioteca humana en el sistema de "control de calidad" en levadura

La selección con el sistema de "control de calidad" para determinar regiones de marco estables se realizó esencialmente tal como se describe en detalle por Auf der Maur (documento WO0148017, Auf der Maur 2001, cada uno incorporado como referencia al presente documento).

Los plásmidos para la expresión de los constructos de fusión de scFv para la selección en levadura se derivaron de pESBA-Act (Wörn, 2000). Contiene el gen de levadura TRP1 para la selección por transformación en *S. cerevisiae* y el origen de replicación de 2 micrómetros para garantizar altos números de copias. Además tiene un promotor de actina constitutivo para una fuerte expresión y la secuencia de terminación de la transcripción GAL11, separados por un sitio de clonación múltiple. Para la manipulación en sistemas bacterianos, también tiene un origen de replicación bacteriano y el gen de resistencia a amp.

El dominio de activación Gal4 (aminoácidos del AD 768-881) se amplificó originariamente mediante PCR usando pGAD424 (Clontech) como molde con cebadores que incluían la señal de localización nuclear del antígeno T de SV40 T N-terminal al Gal4-AD. Los fragmentos de ADN que codifican para los aminoácidos 263-352 de Gal11P se amplificaron mediante PCR y se clonaron en el marco, N-terminal al constructo SV40-NLS-Gal4-AD. La biblioteca de scFv humano, amplificada a partir de ADN de células de bazo humanas tal como se describe en otra parte (Welschhof, 1995; Krebber, 1997; de Haard, 1999), se clonó en el marco, N-terminal a este constructo de fusión mediante sitios SfiI, y en la orientación V_L-enlazador-V_H en la que el enlazador tiene la secuencia $(GGGS)_4$. Por tanto, la expresión produce una proteína de fusión de la estructura general scFv-Gal11p-SV40 NLS-Gal4AD.

Se llevó a cabo la selección en la cepa de levadura *S. cerevisiae* YDE172 (*MATα ura3-52 leu2Δ1 trp1d63 his3Δ200 lys2Δ 385 gal4Δ 11*) (Auf der Maur, 2001), que se derivó de la cepa JPY9 (Escher, 2000) integrando los genes indicadores LacZ y HIS3 orientados de manera divergente bajo el control de UAS_G natural de las secuencias reguladoras Gal1-GAL10 en el locus *his3Δ200*. La activación transcripcional del sistema indicador está mediada por el resto Gal4-AD del constructo de fusión de scFv, tras la interacción específica de su resto Gal11p con el dominio de unión a ADN Gal4 (DBD, aminoácidos 1-100). El Gal4-DBD se proporciona mediante la expresión de un segundo plásmido, pMP83. Contiene el gen de levadura LEU2 para la selección por transformación en *S. cerevisiae* y el origen de replicación ARS CEN. Además, tiene un promotor de actina constitutivo para una fuerte expresión y la secuencia de terminación de la transcripción GAL11. Para la manipulación en sistemas bacterianos, también tiene un origen de replicación bacteriano y el gen de resistencia a amp.

Para la selección, la cepa de levadura *S. cerevisiae* YDE172 se cotransformó con una biblioteca de scFv como constructo de fusión en el vector pESBA-Act2 mientras que el vector pMP83 proporcionó Gal4-DBD. Se usó un protocolo de transformación con acetato de litio convencional (Agatep, 1998). Tras la transformación, se sembraron las células en placas deficientes (-Trp/-Leu/-His) que contienen 3-aminotriazol 80 mM. Se recogieron las colonias tras una incubación de 3 días a 30°C y se volvieron a sembrar en cultivos en estría en placas deficientes (-Trp/-Leu/-His) que contenían 3-aminotriazol 80 mM. Se sometieron a prueba las que volvieron a crecer para determinar la

expresión de LacZ-mediante el desarrollo de color azul en un ensayo de filtro con placas que contenían el sustrato X-Gal. Se tomaron los clones positivos para análisis adicional que implicaba el aislamiento del plásmido que porta scFv de levadura, la transformación en *E. coli* DH5α, el aislamiento de plásmido de colonias individuales de *E. coli* y nueva transformación en la cepa de levadura *S. cerevisiae* YDE172 recién preparada para el ensayo tal como se describe a continuación. Se realizaron todos los métodos según procedimientos convencionales, bien conocidos por un experto habitual en la técnica.

Además, se usó un procedimiento de selección modificado en el que el scFv se fusionó directamente tanto a un dominio de unión a ADN (aminoácidos de LexA 1-202) y un dominio de activación (Gal4, aminoácidos 768-881) para producir un constructo de fusión de la siguiente estructura: scFv-LexA-NLS-Gal4AD. Los plásmidos para la expresión de los constructos de fusión de scFv para la selección en levadura se derivaron de pESBA-Act2. Contiene el gen de levadura TRP1 para la selección por transformación en *S. cerevisiae* y el origen de replicación de 2 micrómetros para garantizar altos números de copias. Además, tiene un promotor de actina constitutivo (para una fuerte expresión) y la secuencia de terminación de la transcripción GAL11 separados por un sitio de clonación múltiple. Para la manipulación en sistemas bacterianos, también tiene un origen de replicación bacteriano y el gen de resistencia a amp.

Se llevó a cabo la selección en la cepa de levadura *S. cerevisiae* ImmunaLHB (*MATα ura3-52 leu2Δ1 trp1d63 his3Δ200 lys2Δ1 385*) que se derivó de la cepa JPY5 integrando los genes indicadores LacZ e HIS3 orientados de manera divergente bajo el control de un promotor bidireccional con seis sitios de unión a LexA (plásmido indicador de integración pDE200, Escher 2000) en el locus *his3Δ200* e integrando el gen indicador LEU2 bajo el control de un promotor con ocho sitios de unión a LexA (derivado de EGY48) en el locus *leu2Δ1*. La activación transcripcional del sistema indicador está mediada por el resto Gal4-AD del constructo de fusión de scFv. Se llevó a cabo la selección esencialmente tal como se describió anteriormente usando medio deficiente (-Trp/-Leu/-His) y concentraciones de 3-aminotriazol de hasta 40 mM.

Ejemplo 2

Evaluación del funcionamiento *in vivo*

a) en levadura

Para el análisis cuantitativo del funcionamiento de las regiones de marco seleccionadas en levadura (figuras 1 y 3), se transformó la cepa de *S. cerevisiae* Immuna LHB con los scFv aislados como constructos de fusión LexA-Gal4-AD en el vector pESBA-Act2 siguiendo un protocolo de transformación con acetato de litio convencional (Agatep, 1998). Tras la transformación, se sembraron las células en placas deficientes (-Trp). Se inocularon cultivos de 2 ml durante la noche en medio deficiente (-Trp) por duplicado a partir de cultivos en estría que contenían varias colonias y se hicieron crecer a 30°C. Se diluyeron los cultivos en 1 ml de medio deficiente (-Trp) hasta una densidad óptica a 600 nm (DO600) de 0,7. Entonces se hicieron crecer a 30°C durante 2 h. Para el ensayo, se tomaron 100 µl de cultivo celular, se mezclaron con 900 µl de tampón, 45 µl de cloroformo y 30 µl de SDS al 0,1%, se agitaron en vórtex y se incubaron a temperatura ambiente durante 5 minutos. Se inició el desarrollo de color mediante la adición de 0,2 ml de ONPG (4 mg/ml) y se detuvo con 0,5 ml de Na_2CO_3 (1 M). Se calculó la actividad teniendo en cuenta la DO600 del cultivo de ensayo, así como el tiempo de incubación del desarrollo de color y el volumen de cultivo usado.

Se secuenciaron los clones que eran al menos iguales o mejores que el control positivo (el injerto de lambda muy estable descrito anteriormente (Wörn, 2000; Auf der Maur, 2001)) para identificar el subtipo de región de marco (definiciones de subtipos de región de marco según Tomlinson, (1992), Cox, (1994) y Williams, (1996)). La secuenciación reveló una preferencia asombrosa por determinados subtipos de región de marco. Para el dominio variable de cadena pesada (VH), nunca se encontraron los subtipos de región de marco 2 y 6 y el 4 estaba notablemente reducido entre los clones positivos. Corregido para el funcionamiento de las secuencias aisladas en el ensayo intracelular en levadura, existe una fuerte preferencia por la región de marco de VH del subtipo 3, pero también por 1a y 1b en aplicaciones intracelulares. Con respecto al dominio variable de cadena ligera (VL); existe una clara preferencia por las regiones de marco de los subtipos kappa 1, lambda 1 y lambda 3 (tabla 4).

Estos subtipos de región de marco, es decir, VH 1a, 1b y 3 combinados con un dominio VL kappa 1, lambda 1 y lambda 3 son, por tanto, los más adecuados para uso intracelular y otras aplicaciones con requisitos rigurosos referentes a las propiedades de plegamiento del scFv. Las bibliotecas para los sistemas de selección intracelulares deben construirse, por ejemplo, preferentemente a partir de una mezcla de estos subtipos de región de marco únicamente, para reducir la cantidad moléculas que no son funcionales en el entorno reductor.

b) en células de mamífero

Se usó la línea celular Hela para el análisis cuantitativo del funcionamiento de las regiones de marco seleccionadas en células humanas (figuras 2, 4 y 7). Se proporcionó el gen indicador de luciferasa a partir de un plásmido indicador contransfectado pGL3 (Promega) que contenía la luciferasa bajo el control de la UAS de Gal4 natural. Los vectores de expresión de mamífero usados para la transfección transitoria contienen Gal4 (1-147) fusionado en el extremo C-

terminal a VP16-AD bajo el control de un promotor de CMV. Se clonaron en el marco los scFv aislados, C-terminal a una fusión Gal4 (1-147)-VP16 para producir una proteína de fusión Gal4 (1-147)-VP16-scFv tras la expresión. Se cultivaron las células en DMEM complementado con FCS al 2,5% y l-glutamina 2 mM. Se llevaron a cabo transfecciones transitorias según el protocolo de Polyfect (Qiagen) en placas de cultivo tisular de 60 mm usando 0,01-0,1 μg del vector que contenía el constructo de scFv, 0,5 μg de un plásmido de expresión Gal4 (1-147)-VP16-scFv dirigido por el promotor de CMV y 0,5 μg de un vector de expresión de lacZ como referencia para la eficacia de transfección. Se recogieron las células 24-48 horas tras la transfección, se resuspendieron en 1000 μl de tampón y se lisaron mediante tres ciclos de congelación-descongelación. Se centrifugó el lisado celular y se sometió a ensayo el sobrenadante para determinar la actividad luciferasa usando disolución de ensayo de luciferasa (Promega) y para determinar la actividad de LacZ según el protocolo convencional. Se corrigió la actividad luciferasa obtenida, con la actividad de LacZ para tener en cuenta la variación en la eficacia de transfección.

Ejemplo 3

Alineamiento múltiple y análisis de la comparación de secuencias

Para dilucidar el patrón general de secuencias de región de marco adecuadas para aplicaciones intracelulares, se aislaron todos los clones positivos (es decir, los que crecen en condiciones selectivas en el sistema de control de calidad) y se secuenció la parte que codifica para los scFv. Posteriormente, se dividieron las secuencias de scFv en sus componentes de cadena ligera y pesada para permitir el alineamiento de los dominios respectivos (tablas 1 y 2) según el esquema de numeración ajustado estructural de dominios de inmunoglobulina por Honegger (2001).

Para permitir la evaluación de los datos obtenidos, se generó un alineamiento que representaba la biblioteca no seleccionada (tabla 3). Con el fin de obtener secuencias no seleccionadas, se transformó la biblioteca en células de *E. coli* que no expresan los genes de scFv y se recogieron los clones al azar para el aislamiento de plásmidos y la secuenciación de la secuencia de scFv. La biblioteca cubre el repertorio de anticuerpos humanos tal como se esperaba y por tanto no tiene sesgo hacia subgrupos específicos, distintos a los esperados por el patrón de expresión encontrado generalmente en seres humanos.

Se agruparon las secuencias de VH y VL según su subgrupo. Se destacaron los cambios en la secuencia consenso específica de subgrupo. Un experto en la técnica puede distinguir entre cambios positivos, neutros y negativos basándose en el entorno estructural del residuo intercambiado particular (por ejemplo Honegger, 2001). Un intercambio de un residuo perteneciente a un grupo particular de aminoácidos por un residuo del mismo grupo se valida en general como un intercambio neutro. Un intercambio de un residuo perteneciente al grupo de aminoácidos hidrófobos que apuntan al núcleo hidrófobo de la proteína por un aminoácido del grupo de aminoácidos polares pero no cargados o cargados positiva o negativamente sería enormemente desfavorable debido a que sitios donadores/aceptores de hidrógeno no satisfechos alteran el empaquetamiento compacto del núcleo hidrófobo. Por tanto, un cambio de este tipo se considera negativo. Un intercambio de un residuo perteneciente al grupo de aminoácidos polares pero no cargados en la superficie del dominio de inmunoglobulina por un aminoácido del grupo de aminoácidos cargados positiva o negativamente es sumamente favorable ya que aumenta la solubilidad de la proteína. Por tanto, un cambio de este tipo se valida positivamente, mientras que el intercambio de un residuo polar a uno hidrófobo es sumamente desfavorable ya que disminuye la solubilidad de la proteína y por tanto se valida negativamente. En las posiciones con un ángulo phi positivo conservado, un intercambio de cualquier aminoácido por glicina se valida positivamente mientras que un intercambio de glicina por cualquier aminoácido se valida negativamente porque la glicina es el único aminoácido que puede formar un ángulo phi positivo. La pérdida de un puente salino conservado entre las posiciones 45-53, 45-100, 77-100 y 108-137 debido a un intercambio de un aminoácido del grupo de residuos cargados positiva o negativamente por un aminoácido no cargado da como resultado una estabilidad termodinámica disminuida y, por tanto, se considera negativa.

Finalmente, se eligieron 7 dominios VL y 4 dominios VH que se seleccionaron preferentemente durante el control de calidad (es decir, mostrando los intercambios menos negativos y más positivos de la secuencia consenso y que cubren los subgrupos) y que muestran cada uno un alto funcionamiento *in vivo* en levadura. Las secuencias se resumen en la tabla 5 e incluyen dos Vκ1 (k I 27 (1. x) y k III 25 (2. x)), dos Vκ3 (k IV 103 (3. x) y k IV135 (5. x)), una Vλ1 (k IV 107 (4. x)), dos Vλ3 (a33 (7. x) y a43 (6. x)), una VH1b (a33 (x. 3)) y tres VH3 (una fwl0 (x. 2), a43 (x. 4) y a44 (x. 1)). Estos dominios VL y VH se intercambiaron proporcionando 22 combinaciones novedosas en el formato de scFv (1.1, 2.1, 3.1, 4.1, 5.1, 1.2, 2.2, 3.2, 4.2, 5.2, 1.3, 2.3, 3.3, 4.3, 5.3, 7.3, 1.4, 2.4, 3.4, 4.4, 5.4, 6.4).

Ejemplo 4

Evaluación del funcionamiento *in vivo* de dominios intercambiados

a) Funcionamiento en un ensayo intracelular en levadura y células de mamífero

Se sometieron a prueba las 22 combinaciones para determinar su funcionamiento *in vivo* en levadura y células de mamífero tal como se describe en el ejemplo 2 (figuras 3 y 4).

b) Expresión de proteína soluble en condiciones reductoras en levadura

Para comparar los rendimientos de proteína soluble tras la expresión en condiciones reductoras, se expresaron las regiones de marco seleccionadas como una fusión con Gal4 AD en el citoplasma de la levadura *S. cerevisiae*. Los constructos de fusión en el vector pESBA-Act2 tenían la estructura general Gal4 AD-scFv. Se transformaron tal como se describió anteriormente en la cepa de levadura *S. cerevisiae* JPY9 y se sembraron en placas deficientes, -Trp.

Se inocularon 5 ml de cultivos durante la noche en medio deficiente (-Trp) a partir de cultivos en estría que contenían varias colonias y se hicieron crecer a 30°C. Se diluyeron los cultivos en 50 ml de medio deficiente (-Trp) hasta una densidad óptica a 600 nm (DO600) de 0,5. Se hicieron crecer a 30°C durante 5 h. Para el extracto de células nativas, se recogieron mediante centrifugación 2,5 ml de cultivo celular normalizado a una DO600 de 3, se congelaron en nitrógeno líquido y posteriormente se resuspendieron en 75 µl de Y-PER (Pierce) que contenía inhibidor de proteasas (PMSF). Se agitó en vórtex brevemente el sedimento celular resuspendido y se incubó (con ligera agitación) a 20°C durante 20 min. Se sedimentó el material insoluble y agregado a velocidad máxima en una centrífuga Eppendorf a 4°C durante 10 min. Se mezcló el sobrenadante con colorante de carga, se calentó hasta 100°C durante 5 min. y se separó en una SDS-PAGE al 12%. Se visualizaron los constructos de fusión Gal4 AD-scFv solubles mediante inmunotransferencia de tipo Western mediante la detección del resto Gal4 con un anticuerpo de ratón monoclonal anti-Gal4AD (Santa Cruz Biotechnology) como anticuerpo primario y un conjugado anti-ratón-peroxidasa (Sigma) como anticuerpo secundario y usando un sustrato quimioluminiscente (Pierce) (figura 5). Los procedimientos de SDS-PAGE e inmunotransferencia de tipo Western los conoce bien un experto habitual en la técnica.

c) Comportamiento de expresión en el periplasma de *E. coli*

Para la evaluación del comportamiento de expresión periplasmática en *E. coli* (figura 6), se clonaron regiones de marco de scFv aisladas en un vector bacteriano que albergaba el gen de resistencia a cam (catR) y el gen represor lac I (Krebber, 1997), con una secuencia líder pelB N-terminal y una cola de his C-terminal bajo el control del operador/promotor lac. Se transformaron *E. coli* JM83 competentes con estos plásmidos. Se inocularon 50 ml de medio dYT que contenía 35 mg/l de cloranfenicol en matraces de agitación, 1:40 con un cultivo durante la noche y se incubaron a 30°C. Se indujeron las células a una DO600 de 0,8 con IPTG 1 mM y se recogieron tras 3 horas de inducción mediante centrifugación. Se resuspendió el sedimento en Tris 50 mM, pH 7.5, NaCl 500 mM y se normalizó a una DO600 de 10. Se analizaron muestras de cada fragmento scFv o bien directamente (extracto total) o bien tras sonicación seguido por centrifugación (fracción soluble) mediante SDS-PAGE. Se estimó la cantidad de proteína soluble a partir del gel con tinción de Coomassie.

Ejemplo 5

Evaluación detallada de 5 combinaciones con propiedades superiores para uso extracelular

Se eligieron cinco combinaciones como ejemplos que muestran un buen funcionamiento en ensayos intracelulares tanto en levadura como en mamífero, producen proteína soluble durante la expresión en levadura y *E. coli*, y cubren los subgrupos que se seleccionaron preferentemente durante el control de calidad (2.4, 4.4, 5.2, 6.4 y 7.3, véase la tabla 5 para más detalles). Se analizaron estas combinaciones en mayor detalle para evaluar adicionalmente su uso en condiciones reductoras, así como oxidantes.

a) Funcionamiento en un ensayo intracelular en diferentes células de mamífero

Se llevó a cabo el análisis cuantitativo del funcionamiento de las cinco combinaciones en células humanas usando células Hela y además usando la línea celular de osteosarcoma humano Saos-2 y la línea celular de riñón embrionario humano HEK293 tal como se realiza en el ejemplo 2 (figura 7).

b) Funcionamiento *in vitro*

Expresión y purificación

Para la evaluación del funcionamiento *in vitro*, se expresaron las cinco combinaciones superiores en el periplasma de *E. coli* (figura 6). Se inoculo la cantidad de 0,1 l de medio dYT que contenía cloranfenicol 35 mg/l en matraces de agitación, 1:40 con un cultivo durante la noche y se incubaron a 30°C. Se indujeron las células a una DO550 de 1,5 con IPTG 1 mM y se recogieron tras 2 horas de inducción mediante centrifugación. Para la purificación de los scFv, se resuspendió el sedimento celular y se lisó mediante sonicación. Tras la centrifugación en SS34 a 20 krpm, 4°C durante 30 minutos, se aplicó el sobrenadante a una columna de afinidad Ni-MC (HiTrap™ Chelating HP, 1 ml, Amersham Pharmacia) a pH 7,5 y se eluyó con imidazol 200 mM usando un sistema Äkta Basic de Amersham Pharmacia. La pureza de los fragmentos scFv era mayor al 98% tal como se determinó mediante SDS-PAGE (datos no mostrados). Se determinó la concentración de la proteína purificada usando el coeficiente de extinción calculado a 280 nm. Se normalizó el rendimiento de proteína purificada soluble a un volumen de cultivo de 1 l con una DO600 de 10 y se varió desde 8 hasta más de 55 mg.

Resistencia a la agregación

Se ha mostrado que la resistencia frente a la agregación se correlaciona con estabilidad termodinámica (Wörn, 1999) *in vitro* y la eficacia de la localización tumoral en un modelo de tumor xenoinjertado en ratones (Willuda, 1999). Con el fin de someter a prueba la estabilidad, la resistencia a la agregación y la reversibilidad de desplegamiento, se mantuvieron muestras de 200 µl de las proteínas purificadas a concentraciones de 6 µM en Tris 50 mM, pH 7,5, NaCl 100 mM o bien durante 4 días a 4°C o 4 días a 37°C o bien durante 3 días a 4°C seguido por una incubación de 15 ó 60 minutos a 100°C, enfriamiento lento hasta temperatura ambiente y una incubación durante la noche a 4°C. Se analizó posteriormente el estado oligomético de cada muestra en una columna de filtración en gel equilibrada con Tris 50 mM, pH 7,5, NaCl 100 mM para estimar la cantidad de material agregado frente a monomérico (figura 8). Se inyectaron las proteínas en una columna Superdex-75 (Amersham Pharmacia) en un volumen de 100 µl y una velocidad de flujo de 1 ml/min en un sistema Äkta Basic (Amersham Pharmacia).

Resistencia a la degradación por proteasas

Para determinar la estabilidad de las regiones de marco aisladas frente a la degradación por proteasas, un parámetro que es importante para aplicaciones terapéuticas, se incubaron las regiones de marco purificadas en suero humano a 37°C (figura 9).

La proteína de scFv con cola de his purificada (véase anteriormente) a una concentración de 50 µM se diluyó diez veces con suero humano hasta una concentración final de 5 µM en suero al 90%. Entonces, o bien se incubaron las muestras a 37°C durante o bien 3 días o bien 1 día, o bien se tomaron directamente para la carga. Antes de la carga, se sedimentó el material insoluble y agregado a máxima velocidad en una centrífuga Eppendorf a 4°C durante 10 min. Se diluyó el sobrenadante seis veces con un colorante de carga para reducir la cantidad de suero cargado en el gel, se calentó hasta 100°C durante 5 min. y se separó en una SDS-PAGE al 12%. Se visualizaron los fragmentos scFv con cola de his solubles mediante inmunotransferencia de tipo Western mediante la detección de la cola de his con un anticuerpo de ratón monoclonal anti-his (Qiagen) como anticuerpo primario y un conjugado anti-ratón-peroxidasa (Sigma) como anticuerpo secundario y usando un sustrato quimioluminiscente (Pierce). Los procedimientos de SDS-PAGE e inmunotransferencia de tipo Western los conoce bien un experto habitual en la técnica.

Ejemplo 6

Selección de grupos de unión de antígenos a través de la selección de una biblioteca de CDR aleatorizada en la región de marco 7.3 en el sistema de selección por interacción en levadura

La selección con el sistema de interacción para grupos de unión de antígenos se realizó esencialmente tal como se describió anteriormente en detalle (Auf der Maur, 2002).

Los plásmidos para la expresión de los constructos de fusión de scFv para la selección en levadura se derivaron de pESBA-Act2. Contiene el marcador nutricional de levadura TRP1 y el origen de replicación de 2 micrómetros. Además tiene un promotor de actina constitutivo para una fuerte expresión y la secuencia de terminación de la transcripción GAL11, separados por un sitio de clonación múltiple. Para la manipulación en sistemas bacterianos, también tiene un origen de replicación bacteriano y el gen de resistencia a amp.

El dominio de activación Gal4 (aminoácidos de AD 768-881) se amplificó originariamente mediante PCR usando pGAD424 (Clontech) como molde con cebadores que incluían la señal de localización nuclear del antígeno T de SV40 N-terminal al Gal4-AD. Se obtuvo la biblioteca de scFv mediante amplificación por PCR de la región de marco de scFv 7.3 usando cebadores que aleatorizan 7 aminoácidos dentro de la CDR3 de VH. Se clonó el producto de PCR resultante en la región de marco 7.3, presente en el vector en la orientación V_L-enlazador-V_H, como una fusión C-terminal a Gal4-AD. Por tanto, la expresión produce una proteína de fusión de la estructura general Gal4-AD-scFv.

Se llevó a cabo la selección en la cepa de levadura *S. cerevisiae* Immuna LHB (*MATα ura3-52 leu2Δ1 trp1d63 his3Δ200 lys2Δ 385*). Se derivó de la cepa JPY5 integrando los genes indicadores LacZ y HIS3 orientados de manera divergente bajo el control de un promotor bidireccional con seis sitios de unión a LexA (plásmido indicador de integración pDE200, Escher 2000) en el locus *his3Δ200* e integrando el gen indicador LEU2 bajo el control de un promotor con ocho sitios de unión a LexA (derivado de EGY48) en el locus *leu2Δ1*.

La activación transcripcional del sistema indicador está mediada por el resto Gal4-AD del constructo de fusión de scFv, tras la interacción específica de su resto scFv con el resto antigénico de la proteína de fusión cebo. La proteína de fusión cebo consiste en el dominio cinasa de la cinasa 1 de tipo polo humana (hPlk1-KD) fusionada de manera C-terminal a la proteína LexA de unión a ADN. Se amplificó por PCR el dominio cinasa (aminoácidos 2-332) a partir de un ADNc de hPlk1 usando el cebador en 5' 5'-tgctctagaagt gctgcagtgactgcag-3' (Seq. Id. No. 12) y el cebador en 3' 5'-ggttgtcgacttacaggctgctgggagcaatcg-3' (Seq. Id. No. 13). Se clonó el producto de PCR resultante de manera C-terminal a LexA mediante XbaI y SalI en el vector cebo. El vector cebo contiene el marcador nutricional URA3 y un origen de replicación Ars Cen. La expresión de la proteína de fusión cebo está dirigida por un promotor de actina activo de manera constitutiva. Se termina la transcripción mediante la secuencia de terminación GAL11. El vector cebo también porta un origen de replicación bacteriano y el gen de resistencia a amp para la propagación en sistemas bacterianos.

Para la selección, la cepa de levadura *S. cerevisiae* Immuna LHB se cotransformó con una biblioteca de scFv como fusión a Gal4-AD en el vector pESBA-Act2 y el vector cebo que proporciona la fusión LexA-hPLK1-KD siguiendo un protocolo de transformación con acetato de litio convencional (Agatep, 1998). Tras la transformación, se sembraron las células en placas deficientes (-Trp/-Leu/-Ura). Se recogieron las colonias tras incubación durante de 3 a 5 días a 30°C y se volvieron a sembrar en cultivos en estría en placas deficientes (-Trp/-Leu/-Ura). Se sometieron a prueba las que volvieron a crecer para determinar la expresión de lacZ mediante el desarrollo de color azul en un ensayo de filtro en placas que contenían el sustrato XGal. Se tomaron los clones positivos para análisis adicional que implicaba el aislamiento del plásmido que porta scFv de la levadura, la transformación en *E. coli* DH5α, el aislamiento del plásmido de colonias individuales de *E. coli*, la secuenciación y la nueva transformación en la cepa de levadura *S. cerevisiae* Immuna LHB recién preparada para el ensayo tal como se describe a continuación. Se realizaron todos los métodos según procedimientos convencionales, bien conocidos por un experto habitual en la técnica.

Ejemplo 7

Evaluación del funcionamiento *in vivo* de constructos de Fab derivados de regiones de marco de scFv novedosas

Para evaluar el efecto beneficioso del uso de regiones de marco estables de dominios variables sobre diferentes formatos de anticuerpo, se construyeron vectores de expresión de Fab para su uso en la selección por interacción en levadura.

a) Constructos de Fab para la selección intracelular en levadura

Se construyeron dos vectores de expresión diferentes para permitir diferentes niveles de expresión. Los vectores se basan en estructuras principales de o bien yEplac 112 (2 micrómetros) o bien yCplac22 (ars/cen) (Gietz, 1988). Ambos contienen el marcador nutricional de levadura TRP1, un promotor de Gal1/Gal10 bidireccional, inducible, un origen de replicación bacteriano y el gen de resistencia a amp para la manipulación en sistemas bacterianos. En un sentido, se clonó el dominio VH de la región de marco 7.3 de manera N-terminal al dominio CH1 de IgG1 que incluye la cisteína C-terminal, seguido por un enlazador y el dominio de activación Gal4 (aminoácidos de AD 768-881) que incluye el antígeno T de SV40. Por el otro lado, se clonó el dominio VL de la región de marco 7.3 de manera N-terminal al dominio CL (lambda) que incluye la cisteína C-terminal. Los terminadores son el terminador Gal11 en el lado de la cadena pesada y el terminador de ciclina 1 en el lado de la cadena ligera.

b) Funcionamiento en un ensayo intracelular en levadura

Para el análisis cuantitativo del funcionamiento de los grupos de unión a antígenos en formato de scFv y Fab en levadura (figuras 1 y 3), la cepa de *S. cerevisiae* Immuna LHB se cotransformó con los scFv aislados como constructos de fusión Gal4-AD en el vector pESBA-Act2 y el vector cebo que contiene la fusión LexA-hPLK1-KD siguiendo un protocolo de transformación con acetato de litio convencional (Agatep, 1998). Tras la transformación, se sembraron las células en placas deficientes (-Trp, -Ura, Glc). Se inocularon cultivos de 2 ml durante la noche en medio deficiente (-Trp, -Ura, Glc) por duplicado a partir de cultivos en estría que contenían varias colonias y se hicieron crecer a 30°C. Se diluyeron los cultivos en 1 ml de medio deficiente (-Trp, -Ura, Gal) hasta una densidad óptica a 600 nm (DO600) de 0,7. Se hicieron crecer a 30°C durante 5 h. Se llevó a cabo el ensayo tal como se describió anteriormente.

c) Expresión de proteína soluble en condiciones reductoras en levadura

Para comparar los rendimientos de proteína soluble tras la expresión en condiciones reductoras, se expresaron los constructos de scFv y Fab, junto con el vector cebo hPLK1-KD, tal como se describió anteriormente, en el citoplasma de levadura *S. cerevisiae*. Se transformaron tal como se describió anteriormente en la cepa de levadura YDE173 y se sembraron en placas deficientes -Trp, -Ura, que contenían glucosa.

Se inocularon cultivos de 5 ml durante la noche en medio deficiente (-Trp, -Ura, Glc) a partir de cultivos en estría que contenían varias colonias y se hicieron crecer a 30°C. Se diluyeron los cultivos con YPAG hasta una densidad óptica a 600 nm (DO600) de 0,5. Se hicieron crecer a 30°C durante 7,5 h. Para el extracto de células nativas, se recogieron mediante centrifugación 2,5 ml de cultivo celular normalizado hasta una DO600 de 3, se congelaron en nitrógeno líquido y posteriormente se resuspendieron en 75 µl de Y-PER (Pierce). Se agitó en vórtex brevemente el sedimento celular resuspendido y se incubó ligeramente con agitación a 20°C durante 20 min. Posteriormente, se sedimentó el material insoluble y agregado a máxima velocidad en una centrífuga Eppendorf a 4°C durante 10 min. Se mezcló el sobrenadante con colorante de carga, se calentó hasta 100°C durante 5 min. y se separó en una SDS-PAGE al 12%. Se visualizaron la fusión Gal4-AD-scFv soluble y la parte de cadena pesada de Fab fusionado a Gal4-AD mediante inmunotransferencia de tipo Western mediante la detección del resto Gal4 con un anticuerpo de ratón monoclonal anti-Gal4-AD (Santa Cruz Biotechnology) como anticuerpo primario y un conjugado anti-ratón-peroxidasa (Sigma) como anticuerpo secundario y usando un sustrato quimioluminiscente (Pierce) (figura 11). SDS- Los procedimientos de SDS-PAGE e inmunotransferencia de tipo Western los conoce bien un experto habitual en la técnica.

Lista de secuencias

<110> ESBATECH AG

<120> Regiones de marco de inmunoglobulina que demuestran estabilidad potenciada en el entorno intracelular y métodos de identificación de las mismas

<130> 08420PC

<150> Documento US 60/382,649

<151> 22-05-2002

<150> US 60/438,246

<151> 03-01-2003

<160> 13

<170> PatentIn versión 3.1

<210> 1

<211> 107

<212> PRT

<213> Secuencia artificial

<220>

<223> región de marco de anticuerpo

<400> 1

```
Glu Ile Val Met Thr Gln Ser Pro Ser Thr Leu Ser Ala Ser Val Gly
1               5                   10                  15

Asp Arg Val Ile Ile Thr Cys Arg Ala Ser Gln Ser Ile Ser Ser Trp
            20                  25                  30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
        35                  40                  45

Tyr Lys Ala Ser Ser Leu Glu Ser Gly Val Pro Ser Arg Phe Ser Gly
    50                  55                  60

Ser Gly Ser Gly Ala Glu Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65                  70                  75                  80

Asp Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Tyr Lys Ser Tyr Trp Thr
                85                  90                  95

Phe Gly Gln Gly Thr Lys Leu Thr Val Leu Gly
            100                 105
```

<210> 2

<211> 108

<212> PRT

<213> Secuencia artificial

<220>

<223> región de marco de anticuerpo

<400> 2

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

Asp Arg Val Thr Leu Thr Cys Arg Ala Ser Gln Gly Ile Arg Asn Glu
20 25 30

Leu Ala Trp Tyr Gln Gln Arg Pro Gly Lys Ala Pro Lys Arg Leu Ile
35 40 45

Tyr Ala Gly Ser Ile Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly
50 55 60

Ser Gly Ser Gly Thr Glu Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65 70 75 80

Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln Tyr Tyr Ser Leu Pro Tyr
85 90 95

Met Phe Gly Gln Gly Thr Lys Val Asp Ile Lys Arg
100 105

<210> 3

<211> 109

<212> PRT

<213> Secuencia artificial

<220>

<223> región de marco de anticuerpo

<400> 3

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

Glu Ser Ala Ala Leu Ser Cys Arg Ala Ser Gln Gly Val Ser Thr Asn
20 25 30

Val Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu Ile
35 40 45

Tyr Gly Ala Thr Thr Arg Ala Ser Gly Val Pro Ala Arg Phe Ser Gly
50 55 60

Ser Gly Ser Gly Thr Glu Phe Thr Leu Thr Ile Asn Ser Leu Gln Ser
65 70 75 80

Glu Asp Phe Ala Ala Tyr Tyr Cys Gln Gln Tyr Lys His Trp Pro Pro
85 90 95

Trp Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg
100 105

<210> 4

<211> 111

<212> PRT

<213> Secuencia artificial

<220>

<223> región de marco de anticuerpo

<400> 4

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

Lys Val Thr Ile Ser Cys Ser Gly Ser Thr Ser Asn Ile Gly Asp Asn
20 25 30

Tyr Val Ser Trp Tyr Gln Gln Leu Pro Gly Thr Ala Pro Gln Leu Leu
35 40 45

Ile Tyr Asp Asn Thr Lys Arg Pro Ser Gly Ile Pro Asp Arg Phe Ser
50 55 60

Gly Ser Lys Ser Gly Thr Ser Ala Thr Leu Gly Ile Thr Gly Leu Gln
65 70 75 80

Thr Gly Asp Glu Ala Asp Tyr Tyr Cys Gly Thr Trp Asp Ser Ser Leu
85 90 95

Ser Gly Val Val Phe Gly Gly Gly Thr Lys Leu Thr Val Leu Gly
100 105 110

<210> 5

<211> 108

<212> PRT

<213> Secuencia artificial

<220>

<223> región de marco de anticuerpo

<400> 5

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

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Thr Leu Thr His Tyr
20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu Ile
35 40 45

Tyr Asp Thr Ser Lys Arg Ala Thr Gly Val Pro Ala Arg Phe Ser Gly
50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Glu Pro
65 70 75 80

Glu Asp Ser Ala Leu Tyr Tyr Cys Gln Gln Arg Asn Ser Trp Pro His
85 90 95

Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg
100 105

<210> 6

<211> 109

<212> PRT

<213> Secuencia artificial

<220>

<223> región de marco de anticuerpo

<400> 6

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

Thr Ala Thr Val Thr Cys Gly Gly Asn Asn Ile Gly Ser Lys Ser Val
20 25 30

His Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Val Leu Val Val Tyr
35 40 45

Asp Asp Ser Asp Arg Pro Ser Gly Ile Pro Glu Arg Phe Ser Gly Ser
50 55 60

Asn Ser Gly Asn Thr Ala Thr Leu Thr Ile Arg Arg Val Glu Ala Gly
65 70 75 80

Asp Glu Ala Asp Tyr Tyr Cys Gln Val Trp Asp Ser Ser Ser Asp His
85 90 95

Asn Val Phe Gly Ser Gly Thr Lys Val Glu Ile Lys Arg
100 105

<210> 7

<211> 109

<212> PRT

<213> Secuencia artificial

<220>

<223> región de marco de anticuerpo

<400> 7

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

Thr Ala Arg Ile Ser Cys Gly Gly Asn Asn Ile Glu Thr Ile Ser Val
20 25 30

His Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Val Leu Val Val Ser
35 40 45

Asp Asp Ser Val Arg Pro Ser Gly Ile Pro Glu Arg Phe Ser Gly Ser
50 55 60

Asn Ser Gly Asn Thr Ala Thr Leu Thr Ile Ser Arg Val Glu Ala Gly
65 70 75 80

Asp Glu Ala Asp Tyr Tyr Cys Gln Val Trp Asp Ser Ser Ser Asp Tyr
85 90 95

Val Val Phe Gly Gly Gly Thr Lys Leu Thr Val Leu Gly
100 105

<210> 8

<211> 124

<212> PRT

<213> Secuencia artificial

<220>

<223> región de marco de anticuerpo

<400> 8

Gln Val Gln Leu Val Gln 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 Ser Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Ser Ala Ile Ser Gly Ser 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 Ala His Val Leu Arg Phe Leu Glu Trp Leu Pro Asp Ala Phe Asp

100 105 110

Ile Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
115 120

<210> 9

<211> 108

<212> PRT

<213> Secuencia artificial

<220>

<223> región de marco de anticuerpo

<400> 9

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

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Ser Ile Ser Ser Tyr
20 25 30

Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
35 40 45

Tyr Ala Ala Ser Ser Ser Gln Ser Gly Val Pro Ser Arg Phe Arg Gly
50 55 60

Ser Glu Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Asn Leu Gln Pro
65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Ser Tyr Arg Thr Pro Phe
85 90 95

Thr Phe Gly Pro Gly Thr Lys Val Glu Ile Lys Arg
100 105

<210> 10

<211> 123

<212> PRT

<213> Secuencia artificial

<220>

<223> región de marco de anticuerpo

<400> 10

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

Val Lys Val Ser Cys Thr Ala Ser Gly Tyr Ser Phe Thr Gly Tyr Phe
20 25 30

Leu His Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met Gly

35 40 45

Arg Ile Asn Pro Asp Ser Gly Asp Thr Ile Tyr Ala Gln Lys Phe Gln
50 55 60

Asp Arg Val Thr Leu Thr Arg Asp Thr Ser Ile Gly Thr Val Tyr Met
65 70 75 80

Glu Leu Thr Ser Leu Thr Ser Asp Asp Thr Ala Val Tyr Tyr Cys Ala
85 90 95

Arg Val Pro Arg Gly Thr Tyr Leu Asp Pro Trp Asp Tyr Phe Asp Tyr
100 105 110

Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
115 120

<210> 11

<211> 122

<212> PRT

<213> Secuencia artificial

<220>

<223> región de marco de anticuerpo

<400> 11

Glu Val Gln Leu Val Glu 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 Ser Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Ser Ala Ile Ser Gly Ser 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 Ala Gly Ile Ala Val Ala Gly Thr Gly Phe Asp Tyr Trp
100 105 110

Gly Gln Gly Thr Leu Val Thr Val Ser Ser
115 120

<210> 12

<211> 28

<212> ADN

<213> Secuencia artificial

<220>

<223> Cebador de PCR

<400> 12

tgctctagaa gtgctgcagt gactgcag 28

<210> 13

<211> 33

<212> ADN

<213> Secuencia artificial

<220>

<223> Cebador de PCR

<400> 13

ggttgtcgac ttacaggctg ctgggagcaa tcg 33

REIVINDICACIONES

1. Región de marco de cadena sencilla que tiene la estructura general:

NH_2-VL-enlazador-VH-COOH; o NH_2-VH-enlazador-VL-COOH,

en la que

(i) la región de marco de VH tiene la secuencia de aminoácidos J (Seq. Id. No. 10)

VQLVQSGAEVKKPGASVKVSCTASGYSFTGYFLHWVRQAPGQGLEWMGRINPDSGDT IYAQKFQDRVTLTRDTSIGTVYMELTSLTSDDTAVYYCARVPRGTYLDPWDYFDYWG QGTLVTVSS; y

(ii) la región de marco de VL tiene la secuencia de aminoácidos B (Seq. Id. No. 2)

EIVLTQSPSSLSASVGDRVTLTCRASQGIRNELAWYQQRPGKAPKRLIYAGSILQSG VPSRFSGSGSGTEFTLTISSLQPEDVAVYYCQQYYSLPYMFGQGTKVDIKR;

o una variante de las mismas que presenta una identidad del 90% o superior con la región de marco de cadena sencilla a la vez que mantiene una estabilidad potenciada según se determina mediante el ensayo del ejemplo 2.

2. Región de marco de cadena sencilla según la reivindicación 1, que es BJ.

3. Región de marco de cadena sencilla según la reivindicación 2, fusionada a un segundo resto de proteína para producir un constructo de fusión de la estructura general

NH_2-VL-enlazador-VH-segunda proteína-COOH; o NH_2-segunda proteína-VL-enlazador-VH-COOH

4. Región de marco de cadena sencilla según la reivindicación 3, en la que se invierte la orientación de las regiones VH y VL.

5. Región de marco de cadena sencilla según la reivindicación 3, en la que la segunda proteína proporciona una lectura para ensayos intracelulares.

6. Uso de al menos dos secuencias de región de marco en la identificación de una clase de residuo de región de marco conservada, en el que una de las al menos dos secuencias de región de marco es una secuencia de región de marco según la reivindicación 1.

7. Uso según la reivindicación 6, en la que la clase de residuo de región de marco conservada se selecciona del grupo que consiste en:

grupos R polares pero no cargados;

grupos R cargados positivamente;

grupos R cargados negativamente;

grupos R hidrófobos; y

aminoácidos especiales.

8. Uso de al menos dos secuencias de región de marco en la identificación de al menos una secuencia de región de marco conservada, en la que una de las al menos dos secuencias de región de marco es una secuencia de región de marco según la reivindicación 1.

9. Uso según la reivindicación 8, en el que la secuencia de región de marco conservada es de 2-5 residuos.

10. Uso según la reivindicación 8, en el que la secuencia de región de marco conservada es de 5-10 residuos.

11. Uso según la reivindicación 8, en el que la secuencia de región de marco conservada es de 10-25 residuos.

12. Uso según las reivindicaciones 8-11, en el que la secuencia de región de marco conservada tiene huecos.

13. Anticuerpo o fragmento de anticuerpo que comprende la VL y la VH de la región de marco de cadena sencilla según una cualquiera de las reivindicaciones 1-5.

14. Ácido nucleico capaz de codificar para la región de marco de cadena sencilla según una cualquiera de las reivindicaciones 1-5.

15. Vector que comprende el ácido nucleico según la reivindicación 14.

16. Célula huésped que comprende el ácido nucleico según la reivindicación 14.

17. Uso del ácido nucleico según la reivindicación 14, en terapia génica.

18. Región de marco de cadena sencilla según una cualquiera de las reivindicaciones 1-5, anticuerpo o fragmento de anticuerpo según la reivindicación 13, para su uso en la validación de dianas, aplicaciones de diagnóstico, construcciones de bibliotecas o aplicaciones terapéuticas.

19. Uso de la región de marco de cadena sencilla según una cualquiera de las reivindicaciones 1-5, como región de marco aceptora para el injerto de bucles hipervariables.

Figura 1: Funcionamiento intracelular tras la selección de control de calidad en levadura *S. cerevisiae*

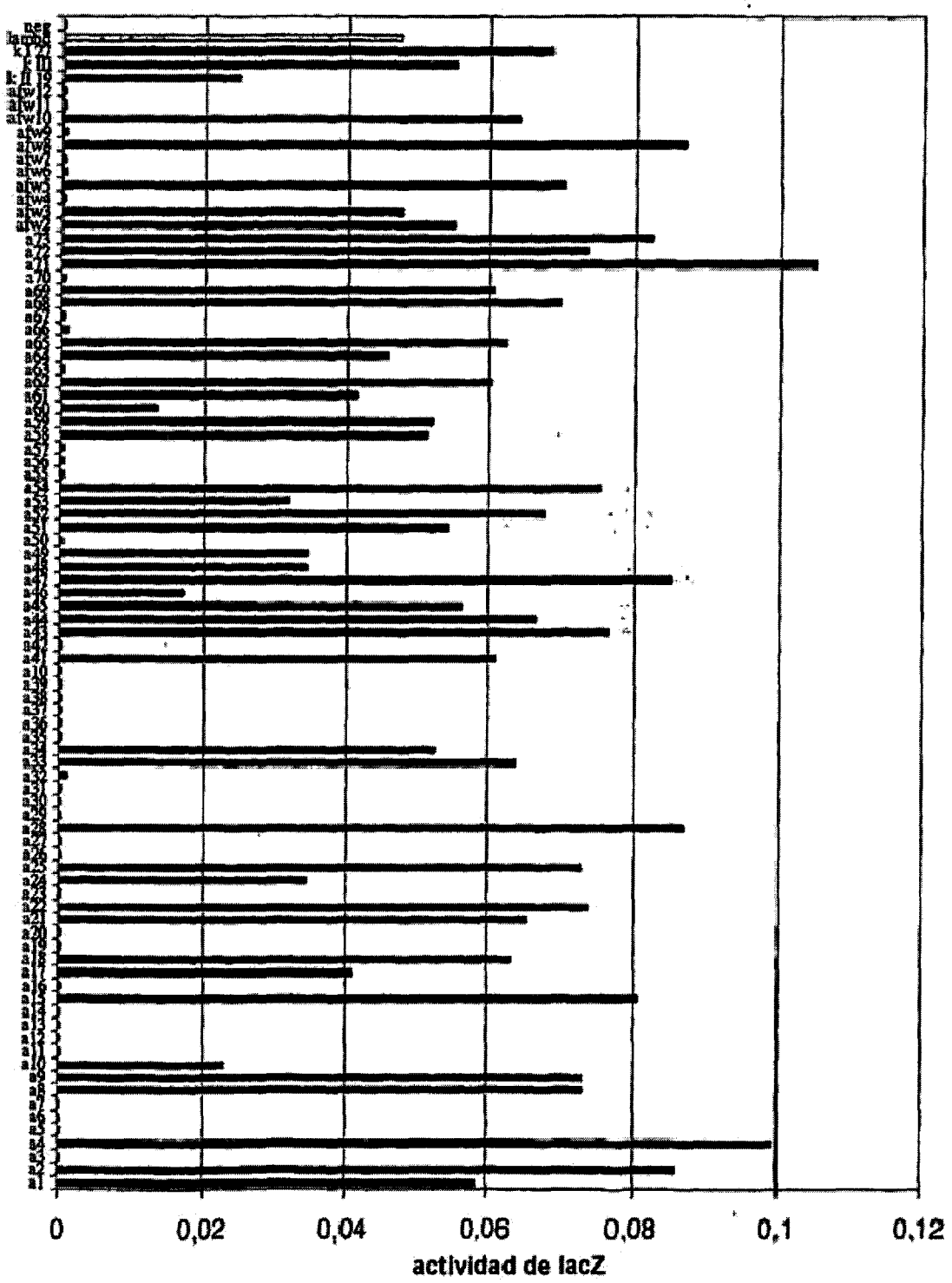


Figura 2: Funcionamiento intracelular de regiones de entramado seleccionadas en una línea celular Hela humana

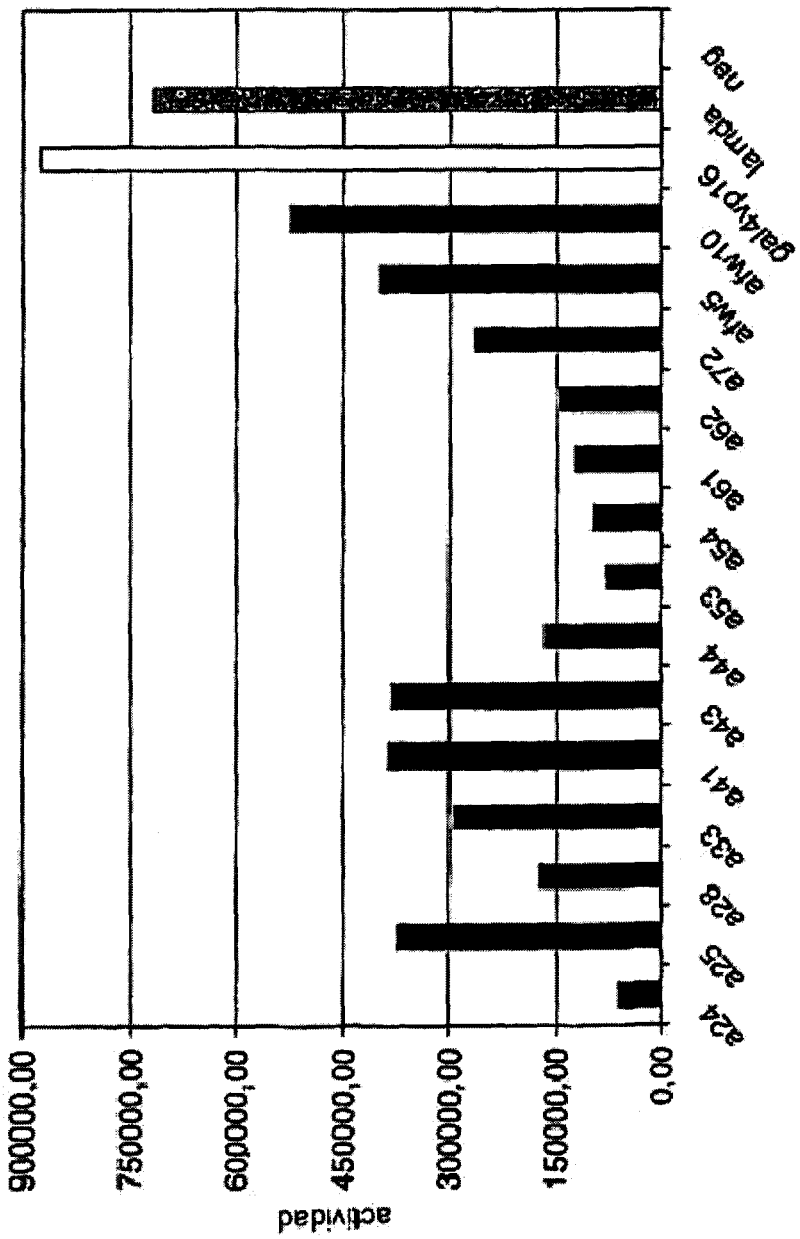


Fig 3:

Funcionamiento intracelular de regiones de entramado novedosas en levadura *S. cerevisiae*

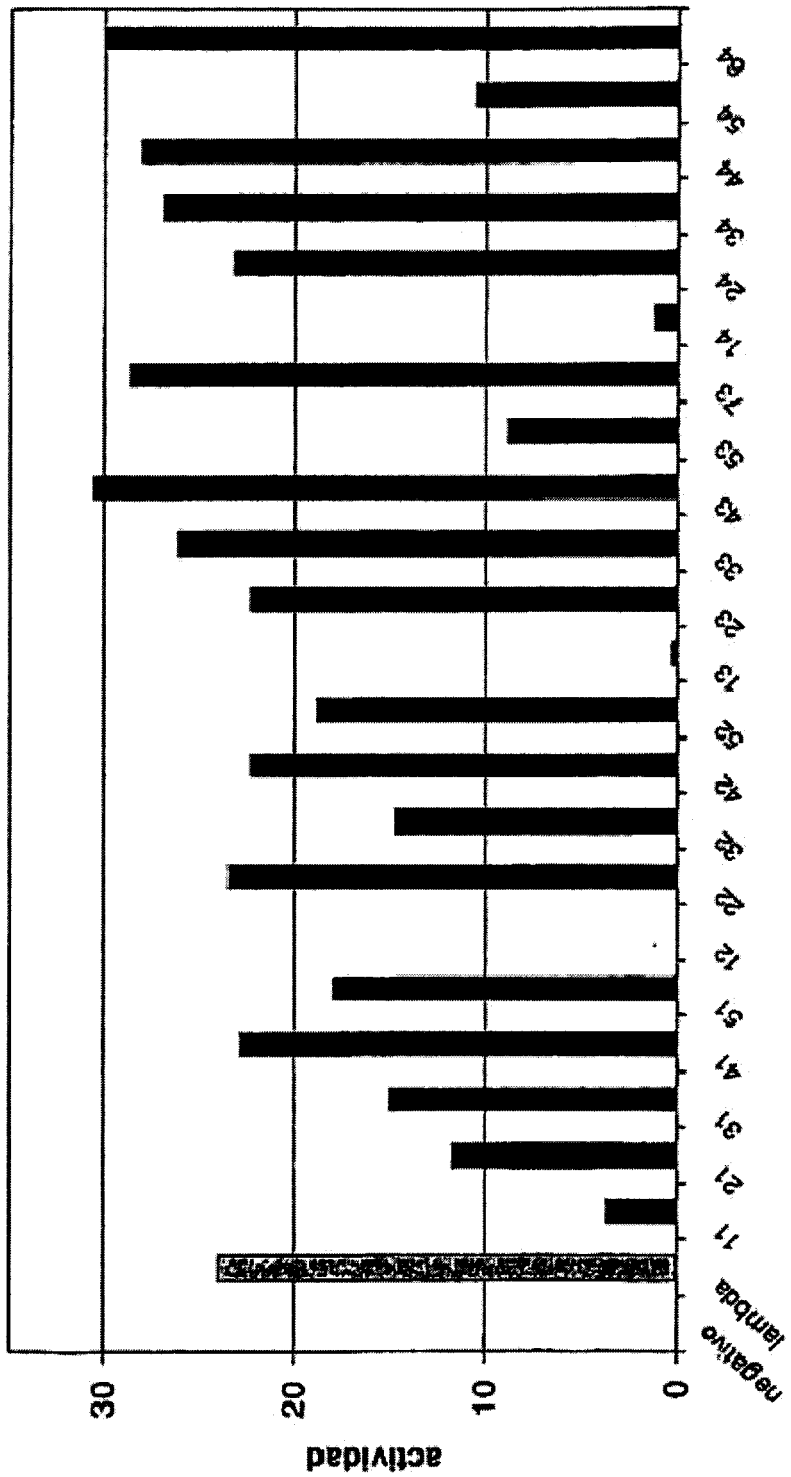
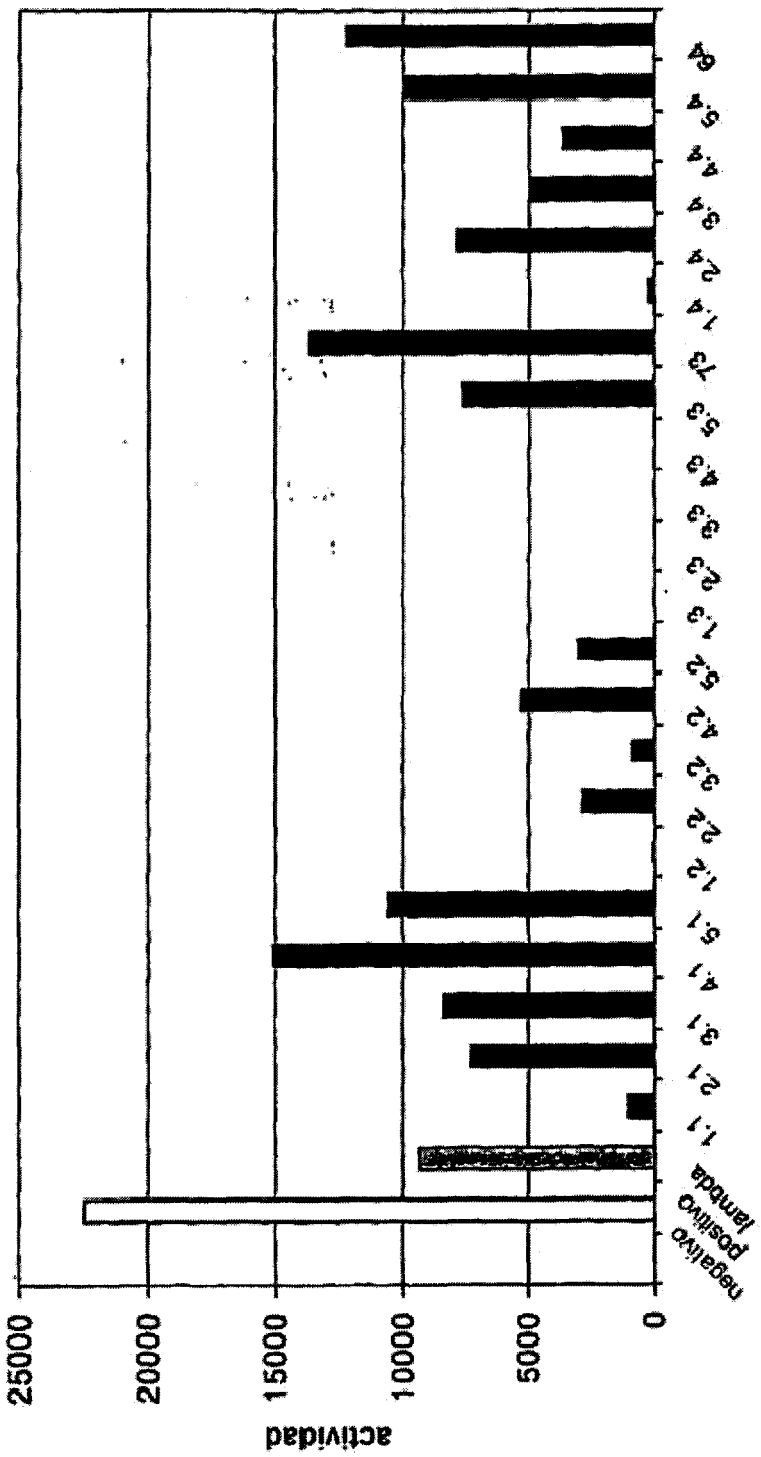


Figura 4:
Funcionamiento intracelular de regiones de entramado novedosas en una línea celular Hela de mamífero



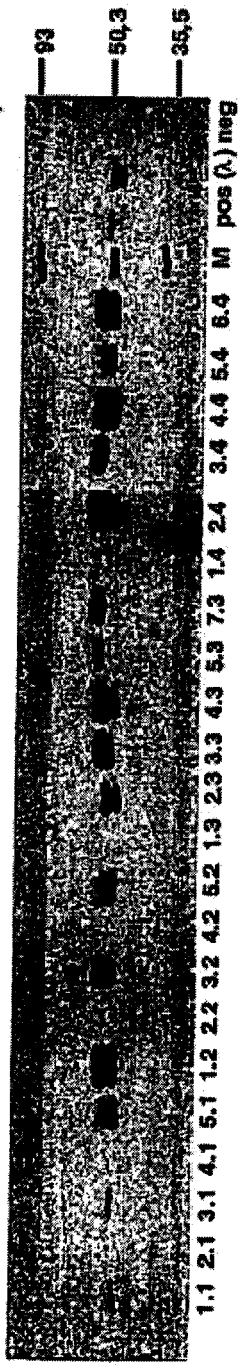


Figura 5: Expresión soluble en el citoplasma de levadura *S. cerevisiae*

Figura 6: Comportamiento de expresión en *E. coli*

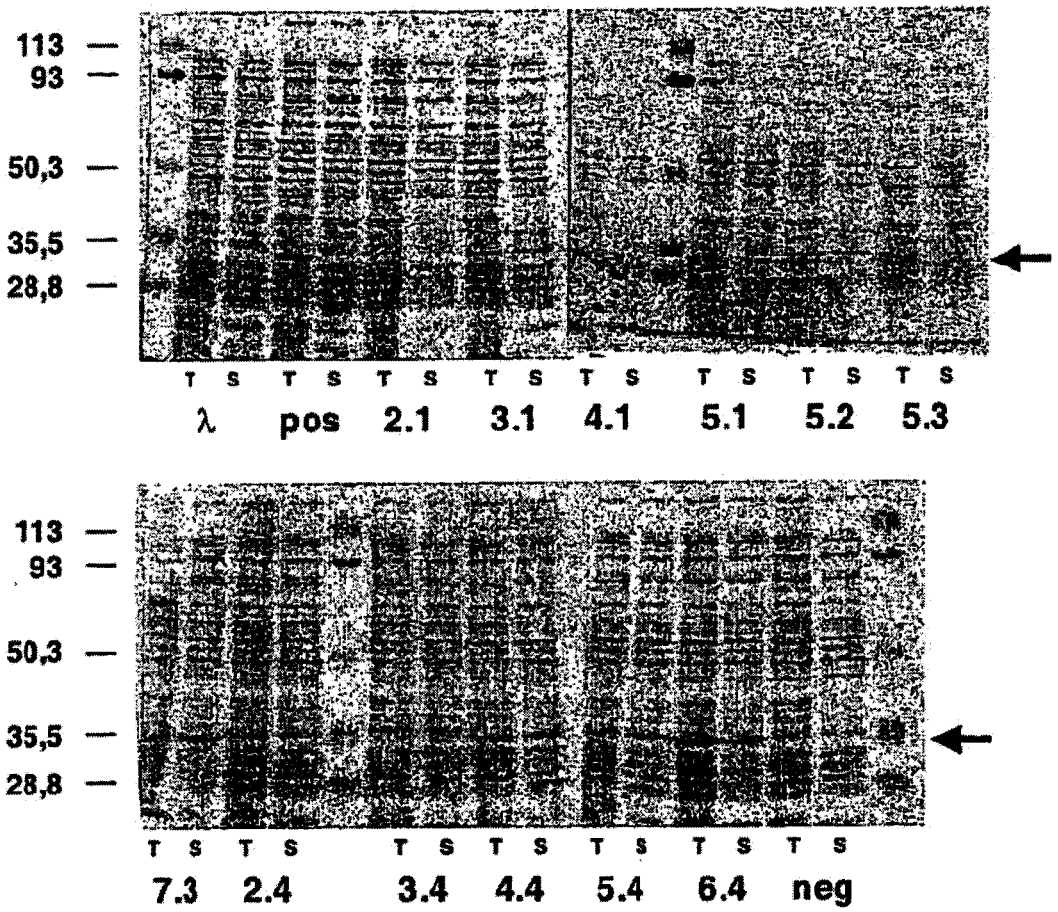


Figura 7:
Funcionamiento intracelular de regiones de entramado novedosas seleccionadas en diferentes líneas celulares de mamífero

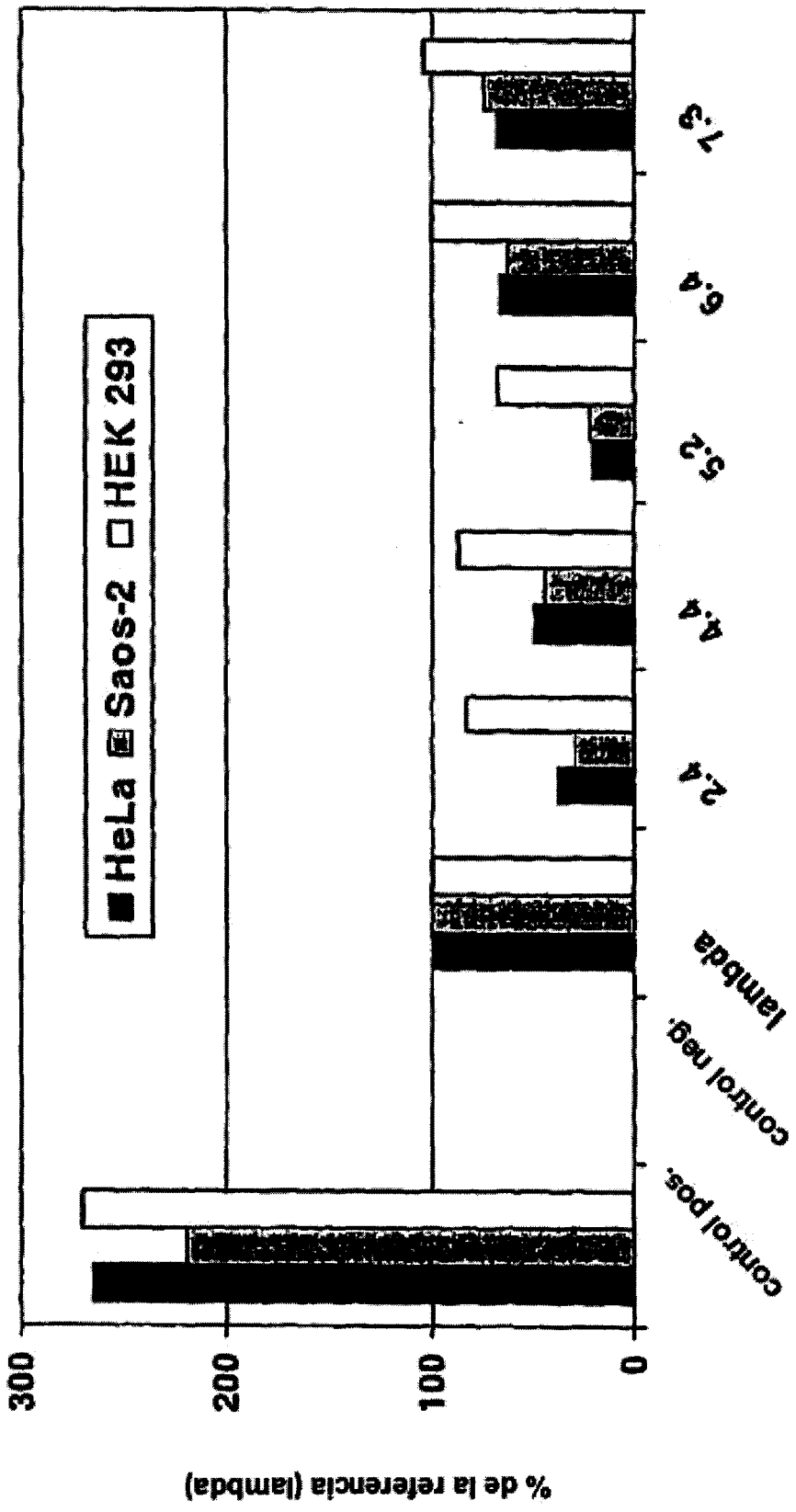


Fig. 8A: Resistencia frente a la agregación a 37°C

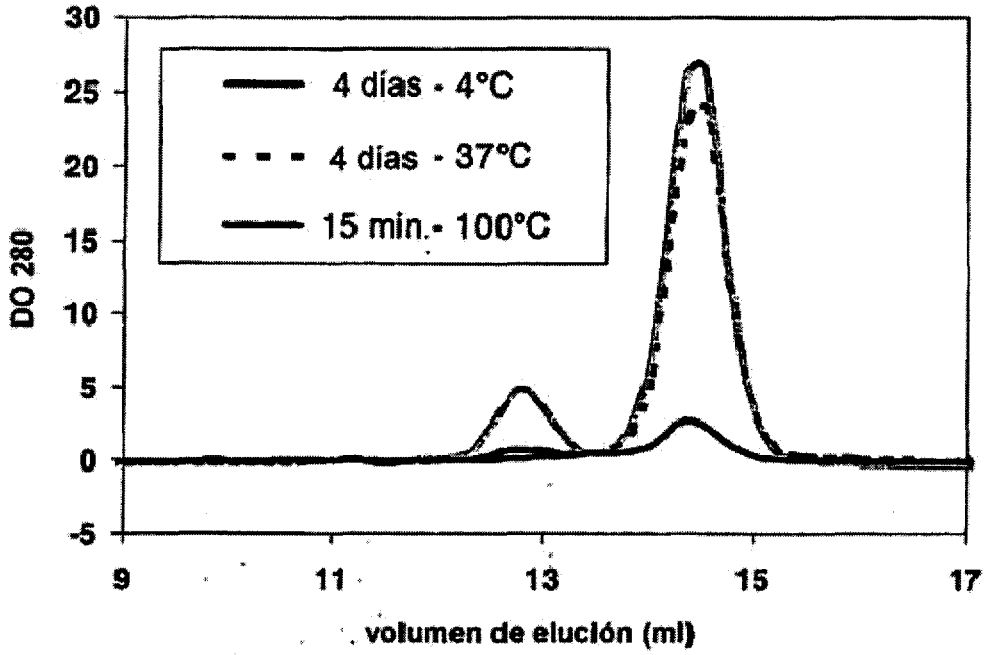


Fig 8B

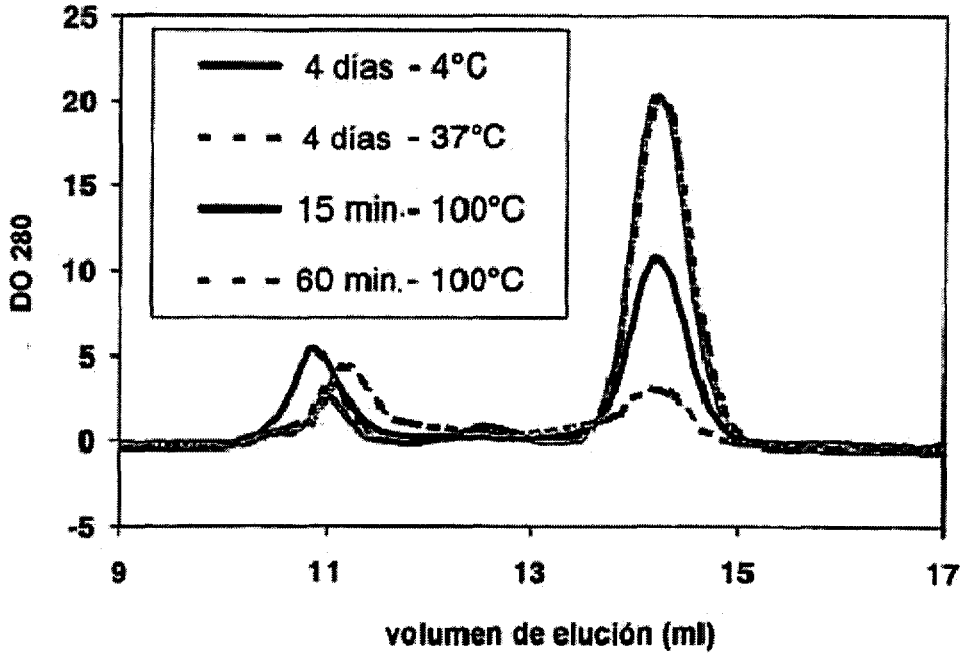


Figura 9: Resistencia frente a la agregación y la degradación por proteasas en suero humano

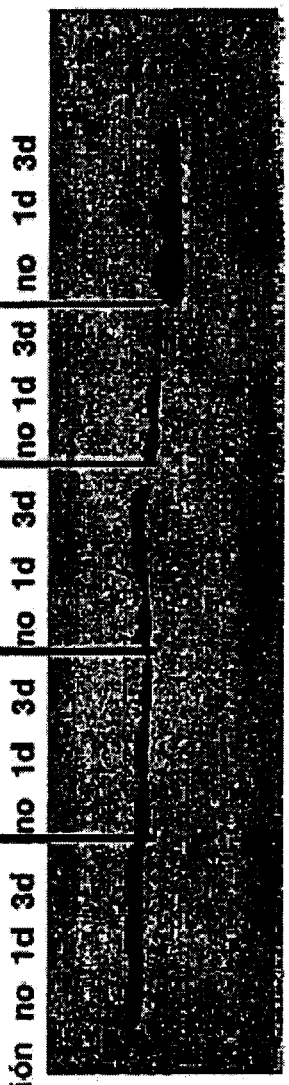


Figura 10

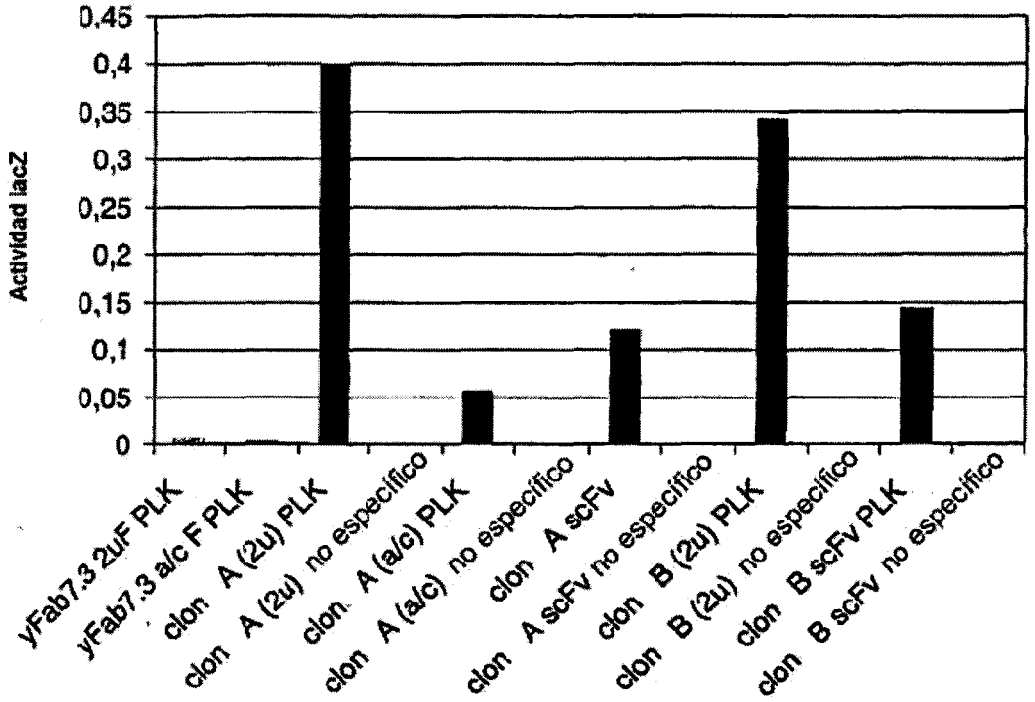

Actividad lacZ
0,45
0,4
0,35
0,3
0,25
0,2
0,15
0,1
0,05
0
yFab7.3 2u/F PLK
yFab7.3 a/c F PLK
clon A (2u) PLK
clon A (2u) no específico
clon A (a/c) PLK
clon A (a/c) no específico
clon A scFv
clon A scFv no específico
clon B (2u) PLK
clon B (2u) no específico
clon B scFv PLK
clon B scFv no específico

Figura 11

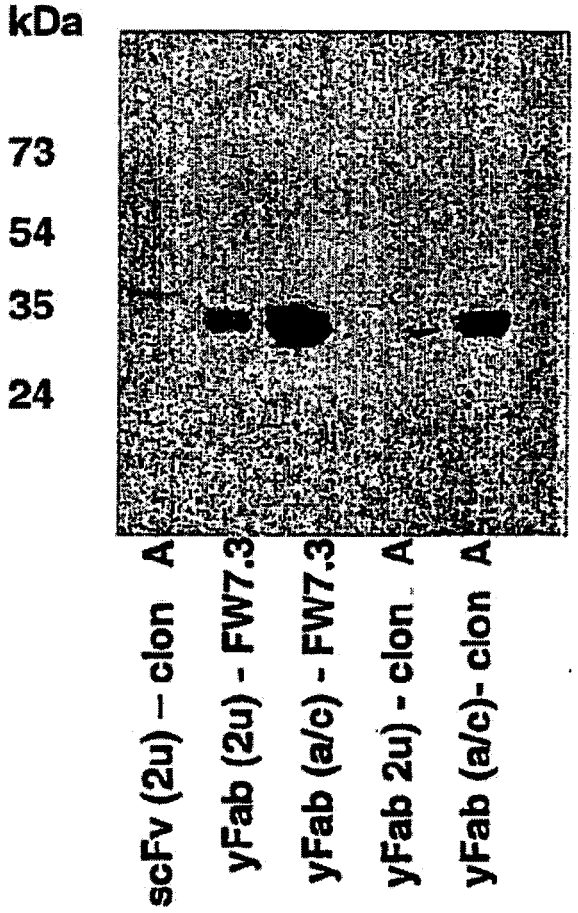

kDa
73
54
35
24
scFv (2u) – clon A
yFab (2u) - FW7.3
yFab (a/c) - FW7.3
yFab 2u) - clon A
yFab (a/c)- clon A

Tabla 1: Alineamiento de secuencias de dominio VH seleccionadas

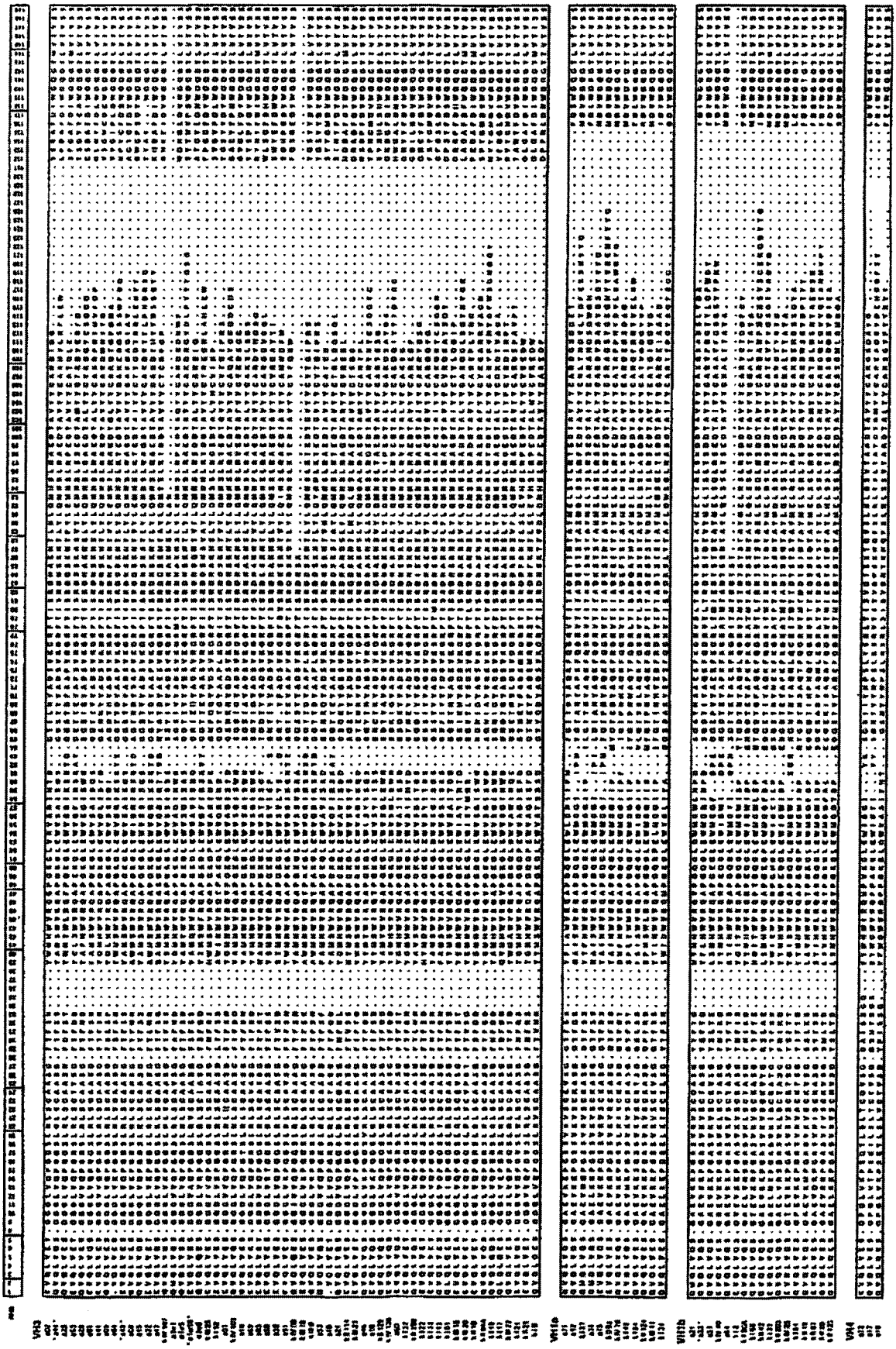


Tabla 2: Alineamiento de secuencias de dominio VL seleccionadas

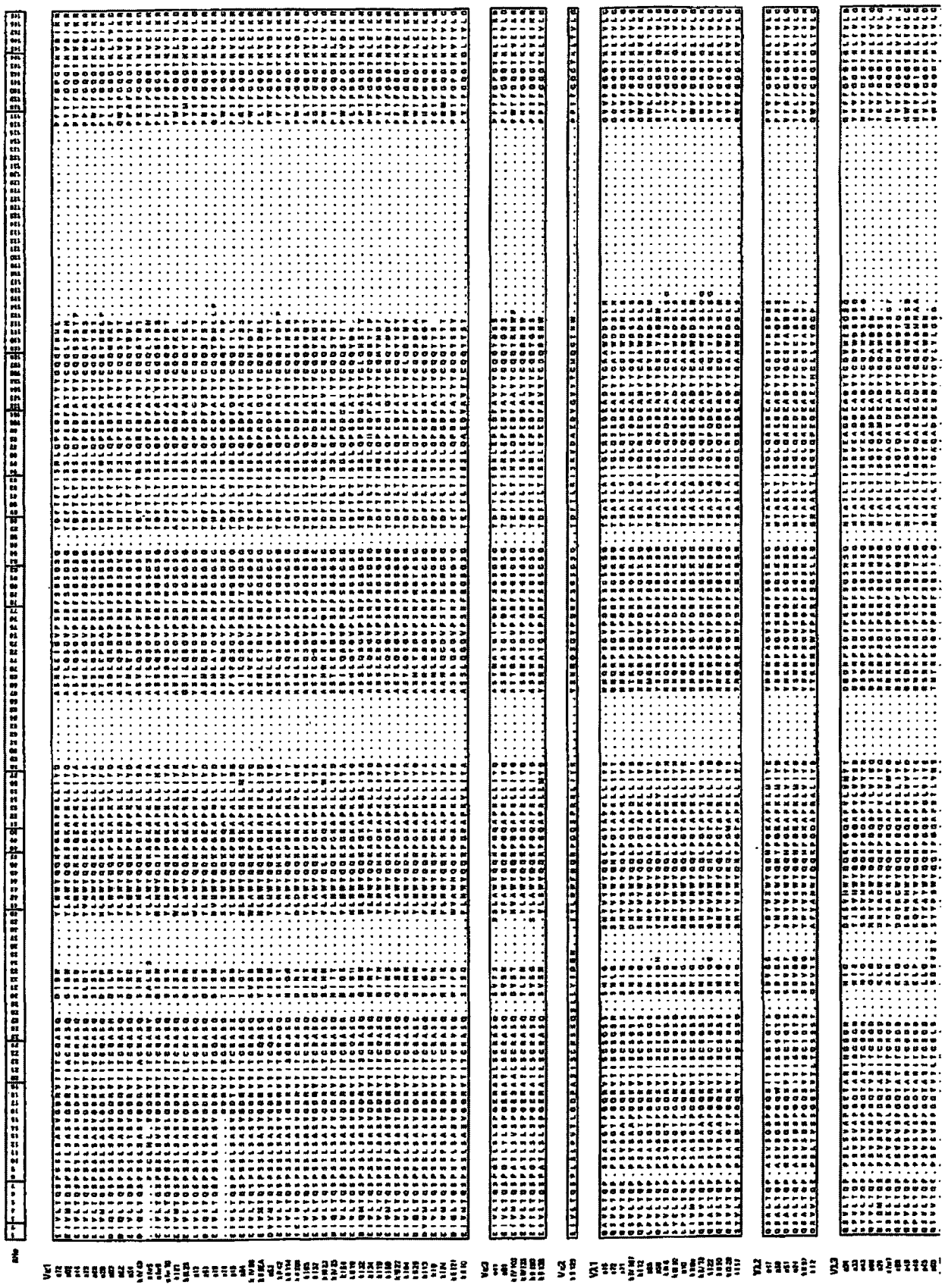


Tabla 3: Alineamiento de miembros de la biblioteca al azar

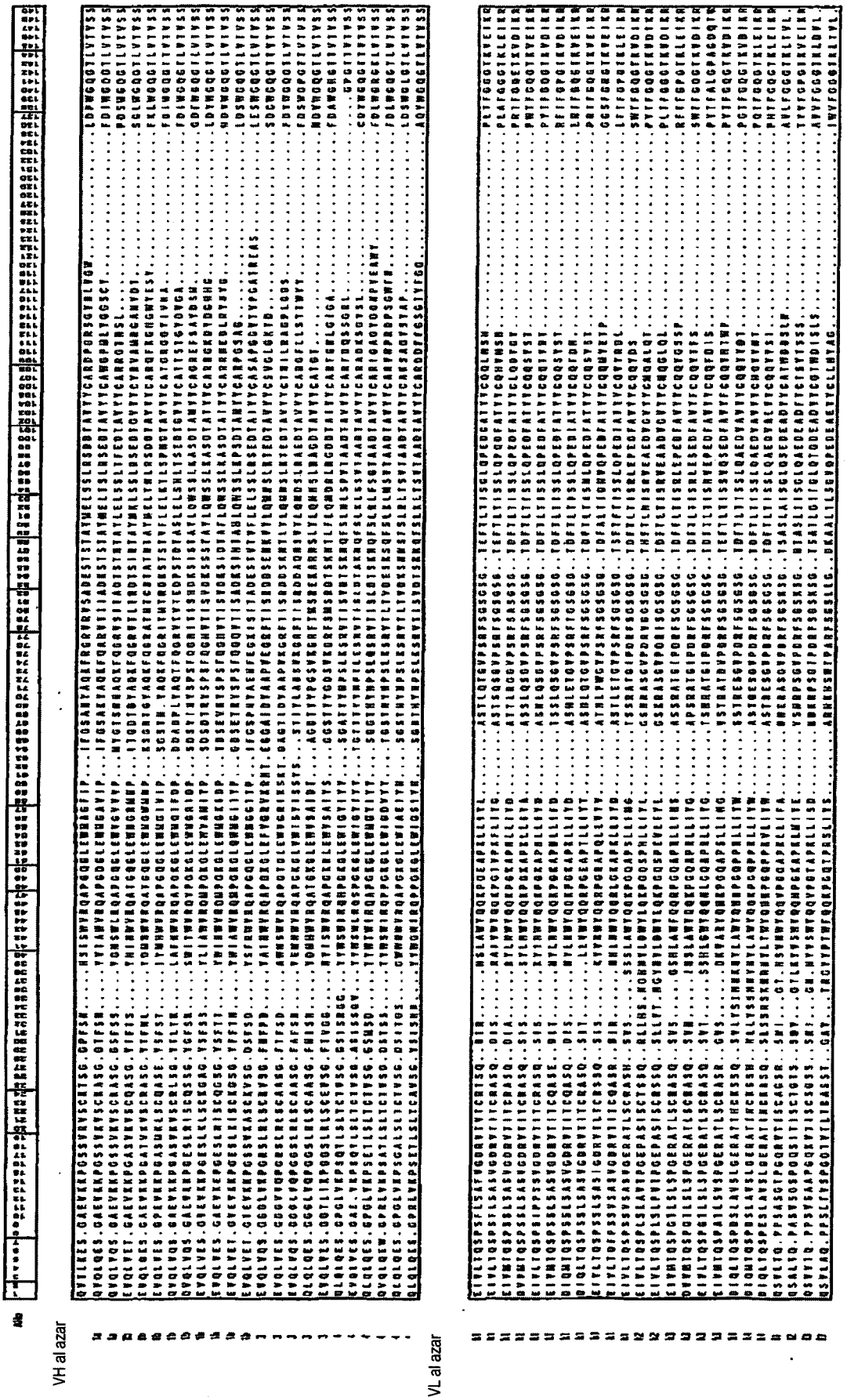


Tabla 4:
Estadística de la frecuencia de subgrupos de regiones de entramado que funcionan bien seleccionadas en el sistema de control de calidad (% de las secuencias seleccionadas totales)

Dominio VL	VL seleccionado	VL al azar
VL κ1	52%	46%
VL κ2	5%	8%
VL κ3	5%	17%
VL κ4	0	12,5%
VL λ1	19%	4%
VL λ2	5%	4%
VL λ3	14%	8%
Dominio VH	**VH seleccionado**	**VH al azar**
VH 1b	19%	39%
VH 1a	9%	13%
VH 2	0	0
VH 3	67%	22%
VH 4	5%	26%
VH 5	0	0%
VH 6	0	0%

Tabla 5: Secuencias de dominio VH y VL usadas para las combinaciones de regiones de entramado novedosas			
abr.	origen	subgrupo	secuencia de aminoácidos
1.x	K I 27	Vκ1	EIVMTQSPSTLSASVGDRVIITCRASQSISSW LAWYQQKPGKAPKLLIYKASSLESGVPSRFS GSGSGAEFTLTISSLQPDDFATYYCQQYKSY WTFGQGTKLTVLG (Seq. Id. No. 1)
2.x	K III25	Vκ1	EIVLTQSPSSLSASVGDRVTLTCRASQGIRNE LAWYQQRPGKAPKRLIYAGSILQSGVPSRFS GSGSGTEFTLTISSLQPEDVAVYYCQQYYSL PYMFGQGTKVDIKR (Seq. Id. No. 2)
3.x	K IV103	Vκ3	EIVMTQSPATLSVSPGESAALSCRASQGVST NVAWYQQKPGQAPRLLIYGATTRASGVPA RFSGSGSGTEFTLTINSLQSEDFAAYYCQQY KHWPPWTFGQGTKVEIKR (Seq. Id. No. 3)
4.x	K IV107	Vλ1	QSVLTQPPSVSAAPGQKVTISCSGSTSNIGDN YVSWYQQLPGTAPQLLIYDNTKRPSGIPDRF SGSKSGTSATLGITGLQTGDEADYYCGTWD SSLSGVVFGGGTKLTVLG (Seq. Id. No. 4)
5.x	K IV135	Vκ3	EIVLTQSPATLSLSPGERATLSCRASQTLTHY LAWYQQKPGQAPRLLIYDTSKRATGVPARF SGSGSGTDFTLTISSLEPEDSALYYCQQRNS WPHTFGGGTKLEIKR (Seq. Id. No. 5)
6.x	A43	Vλ3	SYVLTQPPSVSVAPGQTATVTCGGNNIGSKS VHWYQQKPGQAPVLVVYDDSDRPSGIPERF SGSNSGNTATLTIRRVEAGDEADYYCQVWD SSSDHNVFGSGTKVEIKR (Seq. Id. No. 6)
7.x	A33	Vλ3	LPVLTQPPSVSVAPGQTARISCGGNNIETISV HWYQQKPGQAPVLVVSDDSVRPSGIPERFS GSNSGNTATLTISRVEAGDEADYYCQVWDS SSDYVVFGGGTKLTVLG (Seq. Id. No. 7)

Tabla 5 (continuación)

x.1	A44	VH3	QVQLVQSGGGLVQPGGSLRLSCAASGFTF SSYAMSWVRQAPGKGLEWVSAISGSGGST YYADSVKGRFTISRDNSKNTLYLQMNSLR AEDTAVYYCAAHVLRFLEWLPDAFDIWG QGTLVTVSS (Seq. Id. No. 8)
x.2	Afw10	VH3	EIVLTQSPSSLSASLGDRVTITCRASQSISSY LNWYQQKPGKAPKLLIYAASSSQSGVPSR FRGSESGTDFTLTISNLQPEDFATYYCQQS YRTPFTFGPGTKVEIKR (Seq. Id. No. 9)
x.3	A33	VH1b	VQLVQSGAEVKKPGASVKVSCTASGYSFT GYFLHWVRQAPGQGLEWMGRINPDSGDTI YAQKFQDRVTLTRDTSIGTVYMELTSLTSD DTAVYYCARVPRGTYLDPWDYFDYWGQ GTLVTVSS (Seq. Id. No. 10)
x.4	A43	VH3	EVQLVESGGGLVQPGGSLRLSCAASGFTFS SYAMSWVRQAPGKGLEWVSAISGSGGST YYADSVKGRFTISRDNSKNTLYLQMNSLR AEDTAVYYCAKDAGIAVAGTGFDYWGQG TLVTVSS (Seq. Id. No. 11)