



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 276 770**

51 Int. Cl.:

C12Q 1/00 (2006.01)

C12Q 1/70 (2006.01)

C12Q 1/68 (2006.01)

G01N 33/53 (2006.01)

C07K 14/78 (2006.01)

G01N 33/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **01913159 .8**

86 Fecha de presentación : **28.02.2001**

87 Número de publicación de la solicitud: **1266025**

87 Fecha de publicación de la solicitud: **18.12.2002**

54

Título: **Andamiajes proteínicos para miméticos de anticuerpos y otras proteínas de unión.**

30

Prioridad: **29.02.2000 US 515260**

45

Fecha de publicación de la mención BOPI:
01.07.2007

45

Fecha de la publicación del folleto de la patente:
01.07.2007

73

Titular/es: **Adnexus Therapeutics, Inc.**
100 Beaver Street
Waltham, Massachusetts 02453, US

72

Inventor/es: **Lipovsek, Dasa;**
Wagner, Richard, W. y
Kuimelis, Robert, G.

74

Agente: **Curell Suñol, Marcelino**

ES 2 276 770 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

Andamiajes proteínicos para miméticos de anticuerpos y otras proteínas de unión.

5 Antecedentes de la invención

La presente invención se refiere a andamiajes proteínicos útiles, por ejemplo, para la generación de productos que tienen nuevas características de unión.

10 Las proteínas que tienen estructuras tridimensionales relativamente definidas, habitualmente denominadas como andamiajes proteínicos, se pueden usar como reactivos para el diseño de productos de ingeniería genética. Estos andamiajes contienen típicamente una o más regiones que son susceptibles de una variación de secuencia específica o al azar, y tal aleatorización de secuencia se lleva a cabo a menudo para producir librerías de proteínas a partir de las cuales se pueden seleccionar productos deseados. Un área particular en la que dichos andamiajes son útiles es el campo del diseño de anticuerpos.

Se ha intentado un número de enfoques previos para la manipulación del sistema inmunitario de mamíferos para obtener reactivos o fármacos. Estos han incluido inyectar animales con antígenos de interés para obtener mezclas de anticuerpos policlonales que reaccionan contra antígenos específicos, la producción de anticuerpos monoclonales en cultivo de células de hibridomas (Koehler y Milstein, *Nature* 256:495, 1975), la modificación de anticuerpos monoclonales existentes para obtener nuevas propiedades de reconocimiento, o propiedades optimizadas, la creación de nuevos fragmentos de anticuerpos con características de unión deseables, y la aleatorización de anticuerpos de cadena sencilla (creados conectando las regiones variables de las cadenas pesada y ligera de moléculas de anticuerpos con un ligador peptídico flexible) seguido de la selección de la unión al antígeno mediante presentación de fago (Clackson *et al.*, *Nature* 352:624, 1991).

Además, se han propuesto varios andamiajes proteínicos no inmunoglobulínicos para obtener proteínas con nuevas propiedades de unión. Por ejemplo, se ha diseñado un andamiaje de "minicuerpo", que está relacionado con el plegamiento inmunoglobulínico, suprimiendo tres hebras beta de un dominio variable de cadena pesada de un anticuerpo monoclonal (Tramontano *et al.*, *J. Mol. Recognit.* 7:9, 1994). Esta proteína incluye 61 restos, y se puede usar para presentar dos bucles hipervariables. Estos dos bucles se han aleatorizado, y se han seleccionado productos para la unión a antígenos, pero hasta ahora el armazón parece que tiene una utilidad algo limitada debido a problemas de solubilidad. Otro armazón usado para presentar bucles ha sido tendamistat, un sándwich de láminas beta de seis hebras, de 74 restos, mantenido todo junto mediante dos enlaces de disulfuro (McConnell y Hoess, *J. Mol. Biol.* 250:460, 1995). Este andamiaje incluye tres bucles, pero, hasta la fecha, sólo dos de estos bucles se han examinado para determinar el potencial de aleatorización.

Se han estudiado otras proteínas como armazones, y se han usado para presentar restos aleatorizados sobre superficies helicoidales alfa (Nord *et al.*, *Nat. Biotechnol.* 15:772, 1997; Nord *et al.*, *Protein Eng.* 8:601, 1995), bucles entre hélices alfa en haces de hélices alfa (Ku y Schultz, *Proc. Natl. Acad. Sci. USA* 92:6552, 1995), y bucles restringidos por puentes de disulfuro, tales como los de los inhibidores de proteasa pequeños (Markland *et al.*, *Biochemistry* 35:8045, 1996; Markland *et al.*, *Biochemistry* 35:8058, 1996; Rottgen y Collins, *Gene* 164:243, 1995; Wang *et al.*, *J. Biol. Chem.* 270:12250, 1995).

También se ha demostrado (Koide *et al.*, *J. Mol. Biol.* 284:1141, 1998) el potencial del dominio de tipo III de fibronectina como un andamiaje para manipulación mediante ingeniería de nuevas proteínas de unión.

Sumario de la invención

50 La presente invención proporciona una nueva familia de proteínas capaz de evolucionar para unirse a cualquier compuesto de interés. Estas proteínas, que utilizan un andamiaje de fibronectina o de tipo fibronectina, funcionan de manera característica como anticuerpos naturales o manipulados mediante ingeniería (esto es, anticuerpos policlonales, monoclonales, o de cadena sencilla), y, además, presentan ventajas estructurales. Específicamente, la estructura de estos miméticos de anticuerpos se ha diseñado para un plegamiento, estabilidad y solubilidad óptimos, incluso en condiciones que normalmente conducen a la pérdida de estructura y función en los anticuerpos.

Estos miméticos de anticuerpos se pueden utilizar con el fin de diseñar proteínas que son capaces de unirse a virtualmente cualquier compuesto (por ejemplo, a cualquier proteína) de interés. En particular, las moléculas a base de fibronectina descritas aquí se pueden usar como andamiajes que se someten a evolución dirigida, diseñada para aleatorizar uno o más de los tres bucles de fibronectina que son análogos a las regiones que determinan la complementariedad (CDR) de una región variable de anticuerpo. Tal enfoque de evolución dirigida da como resultado la producción de moléculas de tipo anticuerpo con afinidades elevadas por antígenos de interés. Además, los andamiajes descritos aquí se pueden usar para presentar bucles expuestos definidos (por ejemplo, bucles aleatorizados previamente y seleccionados en base a la unión al antígeno), a fin de dirigir la evolución de moléculas que se unen a tales bucles introducidos. Una selección de este tipo se puede llevar a cabo para identificar moléculas de reconocimiento para cualquier bucle de tipo CDR individual, o, como alternativa, para el reconocimiento de dos o de los tres bucles de tipo CDR combinados en un epítipo no lineal.

ES 2 276 770 T3

En consecuencia, la presente invención se refiere a una proteína que incluye un dominio de tipo III de fibronectina que tiene por lo menos un bucle aleatorizado, estando la proteína caracterizada por su capacidad para unirse a un compuesto que no está unido mediante la correspondiente fibronectina de origen natural.

5 En unas formas de realización preferidas, el dominio de tipo III de fibronectina es un dominio de tipo III de fibronectina de mamífero (por ejemplo, un ser humano); y la proteína incluye el décimo módulo del dominio de tipo III de fibronectina (¹⁰F_n3). En tales proteínas, la unión al compuesto está mediada preferentemente mediante uno, dos o tres bucles de ¹⁰F_n3. En otras formas de realización preferidas, el segundo bucle de ¹⁰F_n3 se puede extender en longitud con relación al módulo de origen natural, o el ¹⁰F_n3 puede carecer de un motivo de unión a integrina. En estas moléculas, el motivo de unión a integrina se puede sustituir por una secuencia de aminoácidos en la que una secuencia de aminoácido básico-aminoácido neutro-aminoácido ácido (en la dirección N-terminal hacia C-terminal) sustituye el motivo de unión a integrina; una secuencia preferida es serina-glicina-glutamato. En otra realización preferida, las proteínas que contienen el dominio de tipo III de fibronectina de la invención carecen de enlaces de disulfuro.

15 Cualquiera de las proteínas que contienen el dominio de tipo III de fibronectina descritas en la presente memoria se puede formular como parte de una proteína de fusión (por ejemplo, una proteína de fusión que incluye además un dominio de F_c inmunoglobulínico, una proteína de complemento, una proteína de toxina, o una proteína de albúmina). Además, cualquiera de las proteínas de dominio de tipo III de fibronectina puede estar covalentemente unida a un ácido nucleico (por ejemplo, una ARN), y el ácido nucleico puede codificar la proteína. Además, la proteína puede ser un multímero, o, particularmente si carece de un motivo de unión a integrina, se puede formular en un vehículo fisiológicamente aceptable.

La presente invención también se refiere a proteínas que incluyen un dominio de tipo III de fibronectina que tiene por lo menos una mutación en la secuencia de la lámina β que cambia la estructura del andamiaje. Nuevamente, estas proteínas se caracterizan por su capacidad de unión a compuestos que no están unidos mediante la correspondiente fibronectina de origen natural.

Además, cualquiera de los andamiajes de fibronectina de la invención se puede inmovilizar sobre un soporte sólido (por ejemplo, una perla o un chip), y estos andamiajes se pueden disponer en cualquier configuración sobre el soporte sólido, incluyendo una colección.

En un aspecto relacionado, la invención se refiere además a ácidos nucleicos que codifican cualquiera de las proteínas de la invención. En unas formas de realización preferidas, el ácido nucleico es ADN o ARN.

35 En otro aspecto relacionado, la invención también se refiere a un método para generar una proteína que incluye un dominio de tipo III de fibronectina, y que es farmacéuticamente aceptable para un mamífero, que implica eliminar el dominio de unión a integrina de dicho dominio de tipo III de fibronectina. Este método se puede aplicar a cualquiera de las proteínas que contienen dominios de tipo III de fibronectina descritas anteriormente, y es particularmente útil para generar proteínas para aplicaciones terapéuticas humanas. La invención también se refiere a tales proteínas que contienen el dominio de tipo III de fibronectina, que carecen de los dominios de unión a integrina.

En todavía otros aspectos relacionados, la invención se refiere a métodos de identificación sistemática que se pueden usar para obtener o desarrollar proteínas de tipo III de fibronectina aleatorizadas, capaces de unirse a compuestos de interés, o para obtener o desarrollar compuestos (por ejemplo, proteínas) capaces de unirse a una proteína particular que contiene un motivo de tipo III de fibronectina aleatorizado. Además, la invención se refiere a procedimientos de identificación sistemática que combinan estos dos métodos, en cualquier orden, para obtener compuestos o proteínas de interés.

En particular, el primer método de identificación sistemática, útil para el aislamiento o identificación de proteínas aleatorizadas de interés, implica: (a) poner en contacto el compuesto con una proteína candidata, incluyendo la proteína candidata un dominio de tipo III de fibronectina que tiene por lo menos un bucle aleatorizado, llevándose a cabo la puesta en contacto en condiciones que permitan la formación del complejo de compuesto-proteína; y (b) obtener, a partir del complejo, la proteína que se une al compuesto.

55 El segundo método de identificación sistemática, para aislar o identificar un compuesto que se une a una proteína que tiene un dominio de tipo III de fibronectina aleatorizado, implica: (a) poner en contacto la proteína con un compuesto candidato, llevándose a cabo la puesta en contacto en condiciones que permitan la formación del complejo de compuesto-proteína; y (b) obtener, a partir del complejo, los compuestos que se unen a la proteína.

60 En unas formas de realización preferidas, los métodos implican además aleatorizar por lo menos un bucle del dominio de tipo III de fibronectina de la proteína obtenida en la etapa (b), y repetir las etapas (a) y (b) usando la proteína aleatorizada adicional, o modificando el compuesto obtenido en la etapa (b), y repitiendo las etapas (a) y (b) usando el compuesto modificado adicional. Además, el compuesto es preferentemente una proteína, y el dominio de tipo III de fibronectina es preferentemente un dominio de tipo III de fibronectina de mamífero (por ejemplo, un ser humano). En otras formas de realización preferidas, la proteína incluye el décimo módulo del dominio de tipo III de fibronectina (¹⁰F_n3), y la unión está mediada por uno, dos o tres bucles de ¹⁰F_n3. Además, el segundo bucle de ¹⁰F_n3 se puede extender en longitud con relación al módulo de origen natural, o ¹⁰F_n3 puede carecer de un motivo de unión a integrina. Nuevamente, como se describe anteriormente, el motivo de unión a integrina se puede sustituir por una

ES 2 276 770 T3

secuencia de aminoácidos en la que una secuencia de aminoácidos básicos-aminoácidos neutros-aminoácidos ácidos (en la dirección N-terminal hacia C-terminal) sustituye al motivo de unión a integrina; una secuencia preferida es serina-glicina-glutamato.

5 Los métodos de selección descritos en la presente memoria se pueden llevar a cabo usando cualquier proteína que contenga un dominio de tipo III de fibronectina. Por ejemplo, la proteína que contiene el dominio de tipo III de fibronectina puede carecer de enlaces de disulfuro, o se puede formular como parte de una proteína de fusión (por ejemplo, una proteína de fusión que incluye además un dominio F_c inmunoglobulínico, una proteína de complemento, una proteína de toxina, o una proteína de albúmina). Además, las selecciones se pueden llevar a cabo usando las
10 proteínas del dominio de tipo III de fibronectina enlazadas covalentemente a ácidos nucleicos (por ejemplo, ARN o cualquier ácido nucleico que codifique la proteína). Además, las selecciones se pueden llevar a cabo usando múltiplos proteínicos que contienen el dominio de fibronectina.

15 Preferentemente, las selecciones implican la inmovilización de la diana de unión sobre un soporte sólido. Los soportes sólidos preferidos incluyen columnas (por ejemplo, columnas de afinidad, tales como columnas de agarosa) o microchips.

Además, la invención se refiere a métodos de diagnóstico que emplean las proteínas de andamiaje de fibronectina de la invención. Tales métodos de diagnóstico se pueden llevar a cabo sobre una muestra (por ejemplo, una muestra biológica) para detectar un analito, o para detectar simultáneamente muchos analitos diferentes en la muestra. El método puede emplear cualquiera de las moléculas de andamiaje descritas aquí. Preferentemente, el método implica
20 (a) poner en contacto la muestra con una proteína que se une al analito del compuesto, y que incluye un dominio de tipo III de fibronectina que tiene por lo menos un bucle aleatorizado, llevándose a cabo la puesta en contacto en condiciones que permitan la formación del complejo de compuesto-proteína; y (b) detectar el complejo, y por lo tanto el compuesto en la muestra.
25

En unas formas de realización preferidas, la proteína se inmoviliza sobre un soporte sólido (por ejemplo, un chip o una perla), y se puede inmovilizar como parte de un conjunto. La proteína se puede unir covalentemente a un ácido nucleico, preferentemente un ácido nucleico tal como ARN, que codifica la proteína. Además, el compuesto es a
30 menudo una proteína, pero también puede ser cualquier otro analito en una muestra. La detección se puede lograr mediante cualquier técnica estándar, incluyendo, sin limitación, la radiografía, la detección mediante fluorescencia, la espectroscopía de masas, o la resonancia plasmónica de superficie.

Como se utiliza en la presente memoria, mediante “dominio de tipo III de fibronectina” se hace referencia a un
35 dominio que tiene 7 u 8 hebras beta que están distribuidas entre dos láminas beta, que se empaquetan ellas mismas entre sí para formar el núcleo de la proteína, y que contienen además bucles que conectan a las hebras beta entre sí y están expuestas al disolvente. Hay por lo menos tres de tales bucles en cada borde del sándwich de láminas beta, en el que el borde es la frontera de la proteína perpendicular a la dirección de las hebras beta. Preferentemente, un dominio de tipo III de fibronectina incluye una secuencia que muestra por lo menos 30% de identidad de aminoácidos,
40 y preferentemente por lo menos 50% de identidad de aminoácidos, con la secuencia que codifica la estructura del dominio de ¹⁰F_n3 denominado como “1ttg” (ID = “1ttg” (un ttg)), disponible de la Base de Datos de Proteínas. La identidad de secuencias citada en esta definición se determina mediante el programa de Homology, disponible de Molecular Simulation (San Diego, CA). La invención incluye además polímeros de las moléculas relacionadas con ¹⁰F_n3, que son una extensión del uso de la estructura monómera, ya sea o no que las subunidades de la poliproteína sean idénticas o diferentes en secuencia.
45

Mediante “fibronectina de origen natural” se hace referencia a cualquier proteína de fibronectina que está codificada por un organismo vivo.

50 Mediante “aleatorizado” se hace referencia a una o más alteraciones de aminoácidos con relación a una secuencia molde.

Mediante una “proteína” se hace referencia a cualquier secuencia de dos o más aminoácidos, independientemente de la longitud, modificación post-traducciona, o función. Aquí, “proteína” y “péptido” se usan de forma intercambiable.
55

Mediante “ARN” se hace referencia a una secuencia de dos o más ribonucleótidos enlazados covalentemente, de origen natural o modificados. Un ejemplo de un ARN modificado incluido en este término es ARN de fosforotioato.

60 Mediante “ADN” se hace referencia a una secuencia de dos o más desoxirribonucleótidos enlazados covalentemente, de origen natural o modificados.

Mediante un “ácido nucleico” se hace referencia a cualquiera de dos o más nucleotídicos o análogos de nucleotídicos o derivados enlazados covalentemente. Como se utiliza en la presente memoria, esta expresión incluye, sin limitación, ADN, ARN, y PNA.
65

Mediante “farmacéuticamente aceptable” se hace referencia a un compuesto o proteína que se puede administrar a un animal (por ejemplo, un mamífero) sin consecuencias médicas adversas significativas.

Mediante “vehículo fisiológicamente aceptable” se hace referencia a un vehículo que no tenga un impacto pernicioso significativo sobre el hospedante tratado, y que retiene las propiedades terapéuticas del compuesto con el que se administra. Un vehículo ejemplar, fisiológicamente aceptable, es disolución salina fisiológica. Otros vehículos fisiológicamente aceptables y sus formulaciones son conocidos por el experto en la técnica, y se describen, por ejemplo, en *Remington's Pharmaceutical Sciences*, (18ª edición), ed. A. Gennaro, 1990, Mack Publishing Company, Easton, PA.

Mediante “seleccionar” se hace referencia a dividir sustancialmente una molécula de otras moléculas en una población. Como se utiliza en la presente memoria, una etapa de “selección” proporciona por lo menos un enriquecimiento de dos veces, preferentemente 30 veces, más preferentemente 100 veces, y lo más preferible 1000 veces, de una molécula deseada con relación a moléculas indeseadas en una población, tras la etapa de selección. Una etapa de selección se puede repetir cualquier número de veces, y en un enfoque dado se pueden combinar diferentes tipos de etapas de selección.

Mediante “pareja de unión”, como se utiliza en la presente memoria, se hace referencia a cualquier molécula que tenga una afinidad específica, covalente o no covalente, por una porción de un compuesto deseado (por ejemplo, proteína) de interés. Los ejemplos de parejas de unión incluyen, sin limitación, miembros de pares de antígeno/anticuerpo, pares de proteína/inhibidor, pares de receptor/ligando (por ejemplo, pares de receptor/ligando de superficie celular, tales como los pares de receptor de hormona/hormona peptídica), pares de enzima/sustrato (por ejemplo, pares de quinasa/sustrato), pares de lectina/hidrato de carbono, agregados oligoméricos o heterooligoméricos de proteínas, pares de proteína de unión a ADN/sitio de unión a ADN, pares de ARN/proteína, y dúplex de ácidos nucleicos, heterodúplex, o hebras ligadas, así como cualquier molécula que sea capaz de formar uno o más enlaces covalentes o no covalentes (por ejemplo, enlaces de disulfuro) con cualquier porción de otra molécula (por ejemplo, un compuesto o proteína).

Mediante un “soporte sólido” se hace referencia a, sin limitación, cualquier columna (o material de columna), perla, tubo de ensayo, cápsula de microtitulación, partícula sólida (por ejemplo, agarosa o Sefarosa), microchip (por ejemplo, silicio, vidrio de silicio, o chip de oro), o membrana (por ejemplo, la membrana de un liposoma o vesícula), a los que un andamiaje de fibronectina o un complejo de afinidad se puede unir, ya sea directa o indirectamente (por ejemplo, a través de otros intermedios de parejas de unión, tales como otros anticuerpos o Proteína A), o en los que se puede embeber un andamiaje de fibronectina o un complejo de afinidad (por ejemplo, a través de un receptor o canal).

La presente invención proporciona un número de ventajas. Por ejemplo, como se describe con más detalle a continuación, el presente mimético de anticuerpo muestra propiedades biofísicas mejoradas, tales como estabilidad bajo condiciones reductoras, y solubilidad a concentraciones elevadas. Además, estas moléculas se pueden expresar fácilmente y plegar en sistemas procariontas, tales como *E. coli*, en sistemas eucariotas, tales como levadura, y en sistemas de traducción *in vitro*, tales como el sistema de lisado de reticulocitos de conejo. Además, estas moléculas son extremadamente susceptibles a técnicas de maduración de afinidad, que implican múltiples ciclos de selección, incluyendo la selección *in vitro* usando tecnología de fusión de ARN-proteína (Roberts y Szostak, Proc. Natl. Acad. Sci USA 94:12297, 1997; Szostak *et al.*, documento WO 00/47775, Szostak *et al.* documento WO 98/31700), presentación de fago (véase, por ejemplo, Smith y Petrenko, Chem. Rev. 97:317, 1997), y sistemas de presentación de levaduras (véase, por ejemplo, Boder y Wittrup, Nature Biotech. 15:553, 1997).

Otras características y ventajas de la presente invención se pondrán de manifiesto a partir de la siguiente descripción detallada de la misma, y a partir de las reivindicaciones.

Breve descripción de los dibujos

La Figura 1 es una fotografía que muestra una comparación entre las estructuras de regiones variables de cadena pesada de un anticuerpo, procedentes del camello (azul oscuro) y de la llama (azul claro), en cada una de las dos orientaciones.

La Figura 2 es una fotografía que muestra una comparación entre las estructuras de la región variable de cadena pesada del anticuerpo de camello (azul oscuro), la región variable de cadena pesada del anticuerpo de llama (azul claro), y un módulo de número 10 de tipo III de fibronectina (¹⁰F_n3) (amarillo).

La Figura 3 es una fotografía que muestra un módulo de número 10 de tipo III de fibronectina (¹⁰F_n3), correspondiendo los bucles a los bucles de unión a antígeno en cadenas pesadas de IgG destacadas en rojo.

La Figura 4 es una gráfica que ilustra una alineación de secuencias entre un dominio de proteína de tipo III de fibronectina y dominios proteínicos relacionados.

La Figura 5 es una fotografía que muestra las similitudes estructurales entre un dominio de ¹⁰F_n3 y 15 proteínas relacionadas, incluyendo fibronectinas, tenascinas, colágenos, y undulina. En esta fotografía, las regiones están marcadas según lo siguiente: constante, azul oscuro; conservada, azul claro; neutra, blanca; variable, rojo; y el motivo de unión a integrina RGD (variable), amarillo.

La Figura 6 es una fotografía que muestra modelos de llenado de espacios de los módulos 9 y 10 de fibronectina III, en cada una de las dos orientaciones diferentes. Están marcados los dos módulos y el bucle de unión a integrina

ES 2 276 770 T3

(RGD). En esta figura, azul indica restos positivamente cargados, rojo indica restos negativamente cargados, y blanco indica restos sin cargar.

La Figura 7 es una fotografía que muestra modelos de llenado de espacios de los módulos 7-10 de fibronectina III, en cada una de las tres orientaciones diferentes. Están marcados los cuatro módulos. En esta figura, azul indica restos positivamente cargados, rojo indica restos negativamente cargados, y blanco indica restos sin cargar.

La Figura 8 es una fotografía que ilustra la formación, bajo diferentes condiciones salinas, de fusiones de ARN-proteína que incluyen dominios de tipo III de fibronectina.

La Figura 9 es una serie de fotografías que ilustran la selección de fusiones de ARN-proteína que contienen el dominio de tipo III de fibronectina, según se mide mediante análisis de señales por PCR.

La Figura 10 es una gráfica que ilustra un incremento en el porcentaje de unión a TNF- α durante las selecciones descritas aquí, así como una comparación entre selecciones de fusiones de ARN-proteína y de proteínas libres.

La Figura 11 es una serie de representaciones esquemáticas que muestran IgG, $^{10}\text{Fn3}$, Fn-CH, -CH₂-CH₃, y Fn-CH₂-CH₃ (a favor de las agujas del reloj desde la parte superior izquierda).

La Figura 12 es una fotografía que muestra un modelo molecular de Fn-CH₁-CH₂-CH₃ basado en estructuras tridimensionales conocidas de IgG (cristalografía mediante rayos X) y $^{10}\text{Fn3}$ (RMN y cristalografía mediante rayos X).

La Figura 13 es una gráfica que muestra el transcurso de tiempo de una selección de fusión de ácido nucleico-proteína a base de $^{10}\text{Fn3}$ ejemplar de ligantes de TNF- α . Se muestran la proporción de conjunto de fusión de ácido nucleico-proteína (diamantes en blanco) y el conjunto de proteína libre (círculos en blanco) que se unen a TNF- α -Sefarosa, y la proporción de conjunto de proteína libre (círculos en negro) que se unen a Sefarosa sin derivatizar.

Las Figuras 14 y 15 son gráficas que ilustran la unión a TNF- α mediante ligantes de Fn de TNF- α . En particular, estas figuras muestran datos de espectros de masas obtenidos a partir de un chip de fusión de $^{10}\text{Fn3}$ y un chip que no es de fusión, respectivamente.

Las Figuras 16 y 17 son la fosforoimagen y el barrido de fluorescencia, respectivamente, de un conjunto de $^{10}\text{Fn3}$, que ilustra la unión a TNF- α .

Descripción detallada

Los nuevos miméticos de anticuerpos descritos aquí se han diseñado para ser superiores tanto a los fragmentos derivados de anticuerpos como a los armazones que no son anticuerpos, por ejemplo aquellos armazones descritos anteriormente.

La ventaja principal de estos miméticos de anticuerpos con respecto a los fragmentos de anticuerpos es estructural. Estos andamiajes derivan de módulos completos, estables, estructurales solubles, encontrados en proteínas de fluidos orgánicos humanos. En consecuencia, muestran un mejor plegamiento y propiedades de termoestabilidad que los fragmentos de anticuerpos, cuya creación implica la eliminación de partes del plegamiento natural del anticuerpo, exponiendo a menudo a restos de aminoácidos que, en un anticuerpo intacto, estarían enterrados en un entorno hidrófobo, tal como una interfaz entre los dominios variables y constantes. La exposición de tales residuos hidrófobos a disolvente aumenta la probabilidad de agregación.

Además, los miméticos de anticuerpos descritos aquí no tienen enlaces de disulfuro, que, según se sabe, retardan o evitan el plegamiento apropiado de fragmentos de anticuerpos en ciertas condiciones. Puesto que los andamiajes presentes no descansan en los disulfuros para la estabilidad del plegamiento natural, son estables en condiciones reductoras, a diferencia de anticuerpos y sus fragmentos, que se desenredan con la ruptura del enlace de disulfuro.

Además, estos andamiajes a base de fibronectina proporcionan las ventajas funcionales de moléculas de anticuerpos. En particular, a pesar de hecho de que el módulo de $^{10}\text{Fn3}$ no es una inmunoglobulina, su plegamiento global está próximo al de la región variable de la cadena pesada de IgG (Figura 2), haciendo posible presentar los tres bucles de fibronectina análogos a las CDR en orientaciones relativas similares a aquellas de los anticuerpos naturales. Debido a esta estructura, los presentes miméticos de anticuerpos presentan propiedades de unión a antígeno que son similares en naturaleza y afinidad a aquellas de los anticuerpos, y se pueden emplear *in vitro* una aleatorización de los bucles y una estrategia de barajado que es similar al proceso de maduración de afinidad de anticuerpos *in vivo*.

Ahora se describen a continuación andamiajes ejemplares a base de fibronectina, y su uso para identificar, seleccionar y desarrollar nuevas proteínas de unión, así como sus ligandos diana. Estos ejemplos se proporcionan con el fin de ilustrar, y no limitar, la invención.

Motivo estructural de ¹⁰F_n3

Los miméticos de anticuerpos de la presente invención se basan en la estructura de un módulo de fibronectina de tipo III (Fn3), un dominio habitual encontrado en proteínas estructurales y de la sangre de los mamíferos. Este dominio aparece más de 400 veces en la base de datos de secuencias proteínicas, y se ha estimado que aparece en un 2% de las proteínas secuenciadas hasta la fecha, incluyendo fibronectinas, tenascina, proteínas citoesqueléticas intracelulares, y enzimas procariotas (Bork y Doolittle, Proc. Natl. Acad. Sci. USA 89:8990, 1992; Bork *et al.*, Nature Biotech. 15:553, 1997; Meinke *et al.*, J. Bacteriol. 175:1910, 1993; Watanabe *et al.*, J. Biol. Chem. 265:15659, 1990). En particular, estos andamiajes incluyen, como moldes, el décimo módulo de Fn3 humano (¹⁰F_n3), que comprende 94 restos de aminoácidos. El plegamiento global de este dominio está estrechamente relacionado con el del fragmento de anticuerpo funcional más pequeño, la región variable de la cadena pesada, que comprende toda la unidad de reconocimiento de antígeno en IgG de camello y de llama (Figura 1, 2). Las principales diferencias entre los dominios del camello y de la llama y el dominio de ¹⁰F_n3 son que (i) ¹⁰F_n3 tiene unas pocas hebras beta menos (siete frente a nueve), y (ii) las dos láminas beta, empaquetadas una frente a otra, están conectadas mediante un puente de disulfuro en los dominios de camello y de llama, pero no en ¹⁰F_n3.

Los tres bucles de ¹⁰F_n3 que corresponden a los bucles de unión a antígeno de la cadena pesada de IgG corren entre los restos de aminoácidos 21-31, 51-56 y 76-88 (Figura 3). La longitud del primer y del tercer bucle, 11 y 12 restos, respectivamente, cae dentro del intervalo de los bucles de reconocimiento de antígeno correspondientes, encontrados en cadenas pesadas de anticuerpos, esto es, los restos 10-12 y 3-25, respectivamente. En consecuencia, una vez aleatorizados y seleccionados en busca de la elevada afinidad por el antígeno, estos dos bucles tienen contactos con los antígenos equivalentes a los contactos de los bucles correspondientes en anticuerpos.

Por el contrario, el segundo bucle de ¹⁰F_n3 tiene sólo 6 restos de longitud, mientras que el bucle correspondiente en las cadenas pesadas de anticuerpos oscila desde 16 a 19 restos. Por lo tanto, para optimizar la unión al antígeno, el segundo bucle de ¹⁰F_n3 se extiende preferentemente en 10-13 restos (además de ser aleatorizado), para obtener la flexibilidad y afinidad en la unión al antígeno más elevadas posibles. De hecho, en general, las longitudes así como las secuencias de los bucles de tipo CDR de los miméticos de anticuerpos se pueden aleatorizar durante la maduración de afinidad *in vitro* o *in vivo* (como se describe con más detalle a continuación).

El décimo dominio de tipo III de fibronectina humana, ¹⁰F_n3, se repliega rápidamente incluso a baja temperatura; su conformación de cadena principal se ha recuperado en 1 segundo a 5°C. La estabilidad termodinámica de ¹⁰F_n3 es elevada ($\Delta G_U = 24$ kJ/mol = 5,7 kcal/mol), lo que se correlaciona con su elevada temperatura de fusión de 110°C.

Uno de los papeles fisiológicos de ¹⁰F_n3 es una subunidad de fibronectina, una glicoproteína que existe en una forma soluble en fluidos corporales, y en una forma insoluble en la matriz extracelular (Dickinson *et al.*, J. Mol. Biol. 236:1079, 1994). Un monómero de fibronectina de 220-250 kD contiene 12 módulos de tipo I, dos módulos de tipo II, y 17 módulos de tipo III de fibronectina (Potts y Campbell, Curr. Opin. Cell Biol. 6:648, 1994). Los diferentes módulos de tipo III están implicados en la unión de fibronectina a integrinas, heparina, y sulfato de condroitina. Se encontró que ¹⁰F_n3 media la adhesión celular a través de un motivo Arg-Gly-Asp (RGD) de unión a integrina en uno de sus bucles expuestos. Se ha demostrado que motivos RGD similares están implicados en la unión a integrinas mediante otras proteínas, tales como fibrinógeno, el factor de von Willebrand, y vitronectina (Hynes *et al.*, Cell 69:11, 1992). No se ha descrito ningún otro papel de unión a la matriz o a las células para ¹⁰F_n3.

La observación de que ¹⁰F_n3 tiene sólo una actividad ligeramente más adhesiva que un péptido corto que contiene RGD es consistente con la conclusión de que la actividad de unión a células de ¹⁰F_n3 está localizada en el péptido de RGD en lugar de estar distribuida en toda la estructura de ¹⁰F_n3 (Baron *et al.*, Biochemistry 31:2068, 1992). El hecho de que ¹⁰F_n3 sin el motivo de RGD no se una probablemente a otras proteínas plasmáticas ni a la matriz extracelular hace a ¹⁰F_n3 un andamiaje útil para sustituir al anticuerpo. Además, la presencia de ¹⁰F_n3 en fibrinógeno natural en el torrente sanguíneo sugiere que el propio ¹⁰F_n3 probablemente no sea inmunógeno en el organismo de origen.

Además, se ha determinado que el armazón de ¹⁰F_n3 presenta secuencias de bucles expuestos que toleran la aleatorización, facilitando la generación de diversos conjuntos de miméticos de anticuerpos. Esta determinación se realizó examinando la flexibilidad de la secuencia de ¹⁰F_n3. En particular, la secuencia de ¹⁰F_n3 humano se alineó con las secuencias de fibronectinas procedentes de otras fuentes, así como también con secuencias de proteínas relacionadas (Figura 4), y los resultados de este alineamiento se cartografiaron sobre la estructura tridimensional del dominio ¹⁰F_n3 humano (Figura 5). Este alineamiento reveló que la mayoría de los restos conservados se encuentran en el núcleo del sándwich de láminas beta, mientras que los restos muy variables están localizados a lo largo de los extremos de las láminas beta, incluyendo los términos N y C, sobre las caras, accesibles al disolvente, de ambas láminas beta, y sobre los tres bucles accesibles al disolvente que sirven como los bucles hipervariables para la maduración de afinidad del mimético de anticuerpo. A la vista de estos resultados, es improbable que la aleatorización de estos tres bucles tenga un efecto adverso sobre el plegamiento global o sobre la estabilidad del propio armazón de ¹⁰F_n3.

Para la secuencia de ¹⁰F_n3 humano, este análisis indica que, como mínimo, los aminoácidos 1-9, 44-50, 61-54, 82-94 (bordes de las láminas beta); 19, 21, 30-46 (par), 79-65 (impar) (caras accesibles al disolvente de ambas láminas beta); 21-31, 51-56, 76-88 (bucles accesibles al disolvente, de tipo CDR); y 14-16 y 36-45 (otros bucles accesibles al disolvente y giros beta), se pueden aleatorizar para desarrollar proteínas de unión a compuestos nuevas o mejoradas. Además, como se expone anteriormente, las alteraciones en las longitudes de uno o más bucles expuestos al disolvente

también se pueden incluir en tales métodos de evolución directos. Como alternativa, también se pueden usar cambios en las secuencias de la lámina β para desarrollar nuevas proteínas. Estas mutaciones cambian el andamiaje y de ese modo alteran indirectamente la estructura o estructuras de los bucles. Si se acepta el enfoque, las mutaciones no deberían saturar la secuencia, sino más bien se deberían de introducir unas pocas mutaciones. Preferentemente, no se deben de introducir, mediante este enfoque, más de 10 cambios de aminoácidos, y, más preferentemente, más de 3 cambios de aminoácidos, a las secuencias de las láminas β .

Fusiones de fibronectina

Los miméticos de anticuerpos descritos aquí se pueden fusionar a otros dominios proteínicos. Por ejemplo, estos miméticos se pueden integrar con la respuesta inmunitaria humana fusionando la región constante de una IgG (F_c) con un módulo $^{10}\text{Fn3}$, preferentemente a través del término C de $^{10}\text{Fn3}$. La F_c en tal molécula de fusión de $^{10}\text{Fn3}$ activa el componente de complemento de la respuesta inmunitaria, y aumenta el valor terapéutico del mimético de anticuerpo. De forma similar, se puede usar una fusión entre $^{10}\text{Fn3}$ y una proteína de complemento, tal como Clq, para seleccionar células como dianas, y se puede usar una fusión entre $^{10}\text{Fn3}$ y una toxina para destruir específicamente células que posean un antígeno particular. Además, $^{10}\text{Fn3}$, en cualquier forma, se puede fusionar con albúmina para aumentar su semivida en el torrente sanguíneo, y su penetración en los tejidos. Cualquiera de estas fusiones se puede generar mediante técnicas estándares, por ejemplo mediante expresión de la proteína de fusión a partir de un gen de fusión recombinante construido usando secuencias génicas disponibles públicamente.

Multímeros del andamiaje de fibronectina

Además de los monómeros de fibronectina, cualquiera de los constructos de fibronectina descritos aquí se puede generar como dímeros o multímeros de miméticos de anticuerpos a base de $^{10}\text{Fn3}$, como medio para incrementar la valencia y de este modo la avidéz de la unión al antígeno. Tales multímeros se pueden generar a través de unión covalente entre módulos de $^{10}\text{Fn3}$ individuales, por ejemplo imitando la unión al término C hasta N de $^8\text{Fn3}$ - $^9\text{Fn3}$ - $^{10}\text{Fn3}$, o imitando dímeros de anticuerpos que se mantienen juntos a través de sus regiones constantes. Se puede explotar un constructo de $^{10}\text{Fn3}$ para diseñar dímeros del esquema general de $^{10}\text{Fn3}$ -Fc::Fc- $^{10}\text{Fn3}$. Los enlaces manipulados mediante ingeniería en la interfaz de Fc::Fc pueden ser covalentes o no covalentes. Además, se pueden usar parejas dimerizantes o multimerizantes distintas de Fc en híbridos de $^{10}\text{Fn3}$ para crear tales estructuras de órdenes superiores.

En ejemplos particulares, se pueden generar multímeros enlazados covalentemente construyendo genes de fusión que codifican el multímero, o, como alternativa, manipulando codones para restos de cisteína en secuencias monoméricas, y permitiendo que se produzca la formación de enlaces de disulfuro entre los productos de expresión. Los multímeros enlazados no covalentemente también se pueden generar mediante una variedad de técnicas. Éstas incluyen la introducción, en secuencias monoméricas, de codones que corresponden a los restos cargados positiva y/o negativamente, y permitiendo que se produzcan interacciones entre estos restos en los productos de expresión (y por lo tanto entre los monómeros). Este enfoque se puede simplificar aprovechando la ventaja de los restos cargados presentes naturalmente en una subunidad monómera, por ejemplo los restos cargados negativamente de fibronectina. Otro medio para generar miméticos de anticuerpos enlazados no covalentemente es introducir, en el gen del monómero (por ejemplo, en los términos amino o carboxi), las secuencias codificantes de proteínas o dominios proteínicos que se sabe que interaccionan. Tales proteínas o dominios proteínicos incluyen motivos de espiral-espiral, motivos de cremallera de leucina, y cualquiera de las numerosas subunidades proteínicas (o sus fragmentos) que se sabe que dirigen la formación de dímeros o de multímeros de órdenes superiores.

Moléculas de tipo fibronectina

Aunque $^{10}\text{Fn3}$ representa un andamiaje preferido para la generación de miméticos de anticuerpos, en las moléculas descritas aquí se puede sustituir $^{10}\text{Fn3}$ por otras moléculas. Éstas incluyen, sin limitación, los módulos $^1\text{Fn3}$ - $^9\text{Fn3}$ y $^{11}\text{Fn3}$ - $^{17}\text{Fn3}$ de fibronectina humana, así como módulos de Fn3 relacionados procedentes de animales no humanos y de procariotas. Además, también se pueden usar módulos de Fn3 procedentes de otras proteínas con homología de secuencias con respecto a $^{10}\text{Fn3}$, tales como tenascinas e undulinas. Los módulos procedentes de diferentes organismos y proteínas progenitoras pueden ser los más apropiados para diferentes aplicaciones; por ejemplo, a la hora de diseñar un mimético de anticuerpo, puede ser tremendamente deseable generar esa proteína a partir de una fibronectina o de una molécula de tipo fibronectina que sea natural para el organismo para el que se pretende una molécula terapéutica o de diagnóstico.

Desarrollo dirigido de proteínas de unión a base de andamiajes

Los miméticos de anticuerpos descritos aquí se pueden usar en cualquier técnica para desarrollar proteínas de unión nuevas o mejoradas. En un ejemplo particular, la diana de unión se inmoviliza sobre un soporte sólido, tal como una resina de columna o un pocillo de placa de microtitulación, y la diana se pone en contacto con una librería de proteínas de unión candidatas a base de andamiajes. Tal librería puede consistir en clones de $^{10}\text{Fn3}$ construidos a partir del andamiaje de $^{10}\text{Fn3}$ de tipo natural, a través de la aleatorización de la secuencia y/o la longitud de los bucles de tipo CDR de $^{10}\text{Fn3}$. Si se desea, esta librería puede ser una librería de fusión de ARN-proteína generada, por ejemplo, mediante las técnicas descritas en Szostak *et al.*, documento WO 00/47775; Szostak *et al.*, documento WO 98/31700; y Roberts y Szostak, Proc. Natl. Acad. Sci. USA (1997) vol. 94, p. 12297-12302. Como alternativa, puede ser una

ES 2 276 770 T3

librería de ADN-proteína (por ejemplo, como la descrita en Lohse, DNA-Protein Fusions and Uses Thereof, U.S.S.N. 60/110.549, U.S.S.N. 09/459.190, y US 99/28472). La librería de fusión se incubaba con la diana inmovilizada, el soporte se lava para eliminar los ligantes no específicos, y los ligantes más fuertemente unidos se eluyen en condiciones muy restrictivas, y se someten a PCR para recuperar la información de secuencia, o para crear una nueva librería de ligantes que se puede usar para repetir el proceso de selección, con o sin mutagénesis posterior de la secuencia. Se puede realizar un número de rondas de selección hasta que se obtengan ligantes de afinidad suficiente por el antígeno.

En un ejemplo particular, el andamiaje de ¹⁰F_n3 se puede usar como la diana de selección. Por ejemplo, si se requiere una proteína que se una a una secuencia peptídica específica presentada en un bucle de diez restos, se construye un único clon de ¹⁰F_n3 en el que uno de sus bucles se ha ajustado a la longitud de diez y a la secuencia deseada. El nuevo clon se expresa *in vivo* y se purifica, y después se inmoviliza sobre un soporte sólido. Después, se permite que una librería de fusión de ARN-proteína, a base de un andamiaje apropiado, interactúe con el soporte, el cual se lava después, y las moléculas deseadas se eluyen y se vuelven a seleccionar como se describe anteriormente.

De forma similar, el andamiaje de ¹⁰F_n3 se puede usar para encontrar proteínas naturales que interactúan con la secuencia peptídica presentada en el bucle de ¹⁰F_n3. La proteína de ¹⁰F_n3 se inmoviliza como se describe anteriormente, y se identifica una librería de fusión de ARN-proteína para la búsqueda de ligantes para el bucle presentado. Los ligantes se enriquecen a través de múltiples rondas de selección, y se identifican mediante secuenciación de ADN.

Además, en los enfoques anteriores, aunque las librerías de ARN-proteína representan librerías ejemplares para el desarrollo directo, en los métodos de selección de la invención se puede usar cualquier tipo de librería a base de andamiajes.

Uso

El mimético de anticuerpo descrito aquí se puede desarrollar para unirse a cualquier antígeno de interés. Estas proteínas tienen propiedades termodinámicas superiores a aquellas de los anticuerpos naturales, y se pueden desarrollar rápidamente *in vitro*. En consecuencia, estos miméticos de anticuerpos se pueden emplear en lugar de los anticuerpos en todas las áreas en las que se usan los anticuerpos, incluyendo en los campos de investigación, terapéutico y de diagnóstico. Además, debido a que estos andamiajes presentan propiedades de solubilidad y estabilidad superiores a los anticuerpos, los miméticos de anticuerpos descritos aquí también se pueden usar en condiciones que destruirían o inactivarían moléculas de anticuerpos. Finalmente, debido a que los andamiajes de la presente invención se pueden desarrollar para unirse virtualmente a cualquier compuesto, estas moléculas proporcionan proteínas de unión completamente nuevas que también encuentran uso en las áreas de investigación, de diagnóstico y terapéutica.

Resultados experimentales

Las moléculas de andamiajes experimentales descritas aquí se generaron y ensayaron, por ejemplo, en protocolos de selección, según lo siguiente.

Construcción de librerías

Se construyó una librería de complejos a partir de tres fragmentos, cada uno de los cuales contenía un área aleatorizada que corresponde a un bucle de tipo CDR. Los fragmentos se denominaron BC, DE y FG, basándose en los nombres de los bucles de tipo CDR-H contenidos en ellos; además de ¹⁰F_n3 y una secuencia aleatorizada, cada uno de los fragmentos contenía tramos que codifican un dominio de His₆ N-terminal o una etiqueta de péptido FLAG C-terminal. En cada unión entre dos fragmentos (es decir, entre los fragmentos BC y DE, o entre los fragmentos DE y FG), cada fragmento de ADN contenía secuencias de reconocimiento para la endonucleasa de restricción EarI de tipo IIS. Esta enzima de restricción permitió el corte y empalme junto de fragmentos adyacentes a la vez que elimina todas las secuencias extrañas, que no son ¹⁰F_n3. También permite un mezclamiento de tipo recombinación de los tres fragmentos de ¹⁰F_n3 entre ciclos de mutagénesis y selección.

Cada fragmento se ensambló a partir de dos oligonucleótidos que solapan, que primero se hibridaron, después se extendieron para formar la forma de ADN bicatenario del fragmento. Los oligonucleótidos que se usaron para construir y procesar los tres fragmentos se enumeran más abajo; las especies "Top" y "Bot", para cada fragmento, son los oligonucleótidos que contenían toda la secuencia codificante de ¹⁰F_n3. En estas denominaciones de los oligonucleótidos, "N" indica A, T, C o G, y "S" indica C o G.

HfnLBCTop(His):

5'- GG AAT TCC TAA TAC GAC TCA CTA TAG GGA CAA TTA CTA TTT ACA
ATT ACA ATG CAT CAC CAT CAC CAT CAC GTT TCT GAT GTT CCG AGG
GAC CTG GAA GTT GTT GCT GCG ACC CCC ACC AGC-3' (SEQ ID NO: 1)

ES 2 276 770 T3

HfnLBCTop (un término N alternativo):

5' GG AAT TCC TAA TAC GAC TCA CTA TAG GGA CAA TTA CTA TTT ACA
ATT ACA ATG GTT TCT GAT GTT CCG AGG GAC CTG GAA GTT GTT GCT
GCG ACC CCC ACC AGC-3' (SEQ ID NO: 2)

HFnLBCBot-flag8:

5'-AGC GGA TGC CTT GTC GTC GTC GTC CTT GTA GTC GCT CTT CCC TGT
TTC TCC GTA AGT GAT CCT GTA ATA TCT (SNN)7 CCA GCT GAT CAG TAG
GCT GGT GGG GGT CGC AGC -3' (SEQ ID NO: 3)

HFnBC3'-flag8:

5'-AGC GGA TGC CTT GTC GTC GTC GTC CTT GTA GTC GCT CTT CCC TGT
TTC TCC GTA AGT GAT CC-3' (SEQ ID NO: 4)

HFnLDETop:

5'- GG AAT TCC TAA TAC GAC TCA CTA TAG GGA CAA TTA CTA TTT ACA
ATT ACA ATG CAT CAC CAT CAC CAT CAC CTC TTC ACA GGA GGA AAT
AGC CCT GTC C-3' (SEQ ID NO: 5)

HFnLDEBot-flag8:

5'-AGC GGA TGC CTT GTC GTC GTC GTC CTT GTA GTC GCT CTT CGT ATA
ATC AAC TCC AGG TTT AAG GCC GCT GAT GGT AGC TGT (SNN)4 AGG CAC
AGT GAA CTC CTG GAC AGG GCT ATT TCC TCC TGT -3' (SEQ ID NO: 6)

HFnDE3'-flag8:

5'-AGC GGA TGC CTT GTC GTC GTC GTC CTT GTA GTC GCT CTT CGT ATA
ATC AAC TCC AGG TTT AAG G-3' (SEQ ID NO: 7)

HFnLFGTop:

5'- GG AAT TCC TAA TAC GAC TCA CTA TAG GGA CAA TTA CTA TTT ACA
ATT ACA ATG CAT CAC CAT CAC CAT CAC CTC TTC TAT ACC ATC ACT
GTG TAT GCT GTC-3' (SEQ ID NO: 8)

ES 2 276 770 T3

HF_nLFGBot-flag8:

5' **AGC GGA TGC CTT GTC GTC GTC GTC CTT GTA GTC TGT TCG GTA ATT**
AAT GGA AAT TGG (SNN)₁₀ AGT GAC AGC ATA CAC AGT GAT GGT ATA -3'
(SEQ ID NO: 9)

HF_nFG3'-glag8:

5' **AGC GGA TGC CTT GTC GTC GTC GTC CTT GTA GTC TGT TCG GTA ATT**
AAT GGA AAT TGG -3' (SEQ ID NO: 10)

T7TMV (introduce el promotor T7 y la región no traducida TMV necesarios para la traducción *in vitro*):

5' **GCG TAA TAC GAC TCA CTA TAG GGA CAA TTA CTA TTT ACA**
ATT ACA-3' (SEQ ID NO: 11)

ASAflag8:

5'-AGC GGA TGC CTT GTC GTC GTC GTC CTT GTA GTC-3' (SEC ID n°: 12)

Unispl-s (oligonucleótido de tablilla usado para ligar ARNm al ligador que contiene puromicina, descrito por Roberts *et al.*, 1997, más arriba):

5'-TTTTTTTTTTNAGCGCATCC-3' (SEC ID n°: 13)

A 18---2PEG (ligador de ADN-puromicina):

5'-(A)₁₈(PEG)₂CCPur (SEC ID n°: 14)

Los pares de oligonucleótidos (500 pmoles de cada uno) se hibridaron en 100 μ l de 10 mM de Tris 7,5, 50 mM de NaCl durante 10 minutos a 85°C, seguido de un enfriamiento lento (0,5-1 hora) a temperatura ambiente. Los fragmentos hibridados con salientes monocatenarios se extendieron entonces usando 100 U de Klenow (New England Biolabs, Beverly, MA) por cada alcuota de 100 μ l de oligos hibridados, y el tampón estaba hecho de 838,5 μ l de H₂O, 9 μ l de 1 M de Tris 7,5, 5 μ l de 1 M de MgCl₂, 20 μ l de 10 mM de dNTP, y 7,5 μ l de 1M de DTT. Las reacciones de extensión transcurrieron durante 1 hora a 25°C.

A continuación, cada uno de los fragmentos de ADN bicatenario se transformó en una fusión de ARN-proteína usando la técnica desarrollada por Szostak *et al.*, documento WO 00/47775, Szostak *et al.*, documento WO 98/31700, y Roberts y Szostak, Proc. Natl. Acad. Sci. USA (1997) vol. 94, p. 12297-12302. De forma resumida, los fragmentos se transcribieron usando un kit de transcripción *in vitro* Ambion, MEGAshortscript (Ambion, Austin, TX), y la fusión resultante de ARNm-ADN-puromicina-proteína se purificó usando Oligo (dT) celulosa, y se sintetizó una hebra de ADN complementaria usando transcriptasa inversa y los cebadores de RT descritos anteriormente (Unisplint-S o flagASA), siguiendo las instrucciones del fabricante.

La fusión de ARN-proteína obtenida para cada fragmento se purificó después sobre la resina apropiada a su etiqueta de purificación peptídica, es decir, sobre Ni-NTA agarosa para la etiqueta de His₆ y M2 agarosa para la etiqueta de FLAG, siguiendo el procedimiento recomendado por el fabricante. El componente de ADNc de las fusiones de ARN-proteína que se unen a la etiqueta se amplificó mediante PCR usando perlas de PCR Ready-to-Go de Pharmacia, 10 pmoles de cebadores de PCR 5' y 3', y el siguiente programa de PCR (Pharmacia, Piscataway, NJ): Etapa 1: 95°C durante 3 minutos; Etapa 2: 95°C durante 30 segundos, 58/62°C durante 30 segundos, 72°C durante 1 minuto, 20/25/30 ciclos, según se requiera; Etapa 3: 72°C durante 5 minutos; Etapa 4: 4°C hasta el final.

ES 2 276 770 T3

El ADN amplificado resultante se escindió mediante 5 U de EarI (New England Biolabs) por 1 μ g de ADN; la reacción tuvo lugar en tampón de ADN ligasa de T4 (New England Biolabs) a 37°C durante 1 hora, y le siguió una incubación a 70°C durante 15 minutos para inactivar EarI. Se combinaron cantidades iguales de los fragmentos de ADN BC, DE y FG, y se ligaron para formar un gen de ¹⁰F_n3 de longitud total con bucles aleatorizados. La ligación requirió 10 U de EarI reciente (New England Biolabs) y 20 U de T4 ADN ligasa (Promega, Madison, WI), y tomó 1 hora a 37°C.

Se obtuvieron tres librerías diferentes de la manera descrita anteriormente. Cada una contenía la forma del bucle FG con 10 restos aleatorizados. Los bucles BC y DE de la primera librería tenían la secuencia de ¹⁰F_n3 de tipo natural; un bucle BC con 7 restos aleatorizados y un bucle DE de tipo natural formaron la segunda librería; y un bucle BC con 7 restos aleatorizados y un bucle DE con 4 restos aleatorizados formaron la tercera librería. La complejidad del bucle FG en cada una de estas tres librerías fue 10¹³; los otros dos bucles aleatorizados proporcionaron el potencial para que una complejidad demasiado grande fuese muestreada en un laboratorio.

Las tres librerías construidas se combinaron en una librería maestra a fin de simplificar el proceso de selección; se esperaba que la propia unión a la diana seleccionase la librería más adecuada para una exposición particular. Las fusiones de ARN-proteína se obtuvieron a partir de la librería maestra siguiendo el procedimiento general descrito en Szostak *et al.*, documento WO 00/47775 Szostak *et al.*, documento WO 98/31700; y Roberts y Szostak, Proc. Natl. Acad. Sci. USA (1997) vol. 94, p. 12297-12302 (Figura 8).

Selecciones de las fusiones

La librería maestra en forma de fusión de ARN-proteína se sometió a selección para la unión a TNF- α . Se emplearon dos protocolos: uno en el que la diana se inmovilizó sobre una columna de agarosa, y uno en el que la diana se inmovilizó sobre un chip BIACORE. En primer lugar, una intensa optimización de las condiciones para minimizar los ligantes de fondo de la columna de agarosa produjo las condiciones favorables del tampón de 50 mM de HEPES pH 7,4, 0,02% de Triton, 100 μ g/ml de ADN de esperma de salmón cizallado. En este tampón, la unión no específica de la fusión de ARN de ¹⁰F_n3 a TNF- α Sefarosa fue 0,3%. El fondo de unión no específica de la ARN-ADN de ¹⁰F_n3 a TNF- α Sefarosa fue de 0,1%. Durante cada ronda de selección sobre TNF- α Sefarosa, se preincubó en primer lugar la librería Profusion™ durante una hora con Sefarosa no derivatizada, para eliminar cualesquiera ligantes no específicos que queden; el resultado del paso de caudal de este preaclaramiento se incubó durante otra hora con TNF- α Sefarosa.

Durante cada ronda de selección sobre TNF- α Sefarosa, la librería de fusión de ARN-proteína se preincubó en primer lugar durante una hora con Sefarosa no derivatizada, para eliminar cualesquiera ligantes no específicos que queden. El resultado del paso de flujo a partir de este preaclaramiento se incubó durante otra hora con TNF- α Sefarosa. La TNF- α Sefarosa se lavó durante 3-30 minutos.

Después de cada selección, el ADNc procedente de la fusión de ARN-proteína que se había eluido a partir del soporte sólido con 0,3 M de NaOH o 0,1 M de KOH se amplificó mediante PCR; una banda de ADN del tamaño esperado persistió a través de múltiples rondas de selección (Figura 9); se observaron resultados similares en los dos protocolos de selección alternativos, y, en la Figura 9, sólo se muestran los datos procedentes de la selección en la columna de agarosa.

En las primeras siete rondas, la unión de las fusiones de ARN-proteína de la librería a la diana permanecieron bajas; por el contrario, cuando la proteína libre se tradujo a partir de conjuntos de ADN en diferentes etapas de la selección, la proporción de la especie de unión en la columna aumentó significativamente entre rondas (Figura 10). Se pueden llevar a cabo selecciones similares con cualquier otra diana de especie de unión (por ejemplo, IL-1 e IL-3).

Estudios con animales

¹⁰F_n3 de tipo natural contiene un motivo tripeptídico de unión a integrina, arginina 78-glicina 79-aspartato 80 (el motivo "RGD") en la punta del bucle FG. A fin de evitar la unión a la integrina y una respuesta inflamatoria potencial basada en este tripéptido *in vivo*, se generó una forma mutante de ¹⁰F_n3 que contenía una secuencia inerte, serina 78-glicina 79-glutamato 80 (el "mutante SGE"), una secuencia que se encontró en el dominio ¹¹F_n3 de tipo natural, estrechamente relacionado. Este mutante SGE se expresó como una proteína libre, marcada con His₆ N-terminalmente, en *E. coli*, y se purificó hasta homogeneidad sobre una columna de quelato metálico, seguido de una columna de exclusión por tamaños.

En particular, la secuencia de ADN que codifica His₆-¹⁰F_n3(SGE) se clonó en el vector de expresión pET9a, y se transformó en células BL21 DE3 pLysS. El cultivo se hizo crecer entonces en caldo LB que contiene 50 μ g/ml de canamicina a 37°C, con agitación, hasta A₅₆₀ = 1,0, y después se indujo con 0,4 mM de IPTG. El cultivo inducido se incubó adicionalmente, en las mismas condiciones, toda la noche (14-18 horas); las bacterias se recuperaron mediante centrifugación estándar, a baja velocidad. El pelete celular se resuspendió en 1/50 del volumen de cultivo original de tampón de lisis (50 mM de Tris 8,0, 0,5 M de NaCl, 5% de glicerol, 0,05% de Triton X-100, y 1 mM de PMSF), y las células se lisaron haciendo pasar la pasta resultante a través de un microfluidizador M110-EH de Microfluidics Corporation, tres veces. El lisado se aclaró mediante centrifugación, y el sobrenadante se filtró a través de un filtro de 0,45 μ m, seguido de la filtración a través de un filtro de 0,2 μ m. Se cargaron 100 ml del lisado aclarado sobre una columna de cobalto Talon de 5 ml (Clontech, Palo Alto, CA), lavada con 70 ml de tampón de lisis, y se eluyó con

ES 2 276 770 T3

un gradiente lineal de 0-30 mM de imidazol en tampón de lisis. El caudal a través de la columna durante todas las etapas fue de 1 ml/min. La proteína eluida se concentró 10 veces mediante diálisis (corte de MW = 3.500) frente a 15.000-20.000 PEG. La muestra resultante se dializó en tampón 1 (tampón de lisis sin glicerol), después se cargó, 5 ml cada vez, sobre una columna de exclusión por tamaños de 16 x 60 mm Sephacryl 100, equilibrada en tampón 1. La columna operó a 0,8 ml/min., en tampón 1; se reunieron todas las fracciones que contenían una proteína del peso molecular MW esperado, se concentraron 10X como se describe anteriormente, y después se dializaron en PBS. Se acopló toxicón (MA) para realizar identificaciones de endotoxinas y estudios con animales sobre la muestra resultante.

En estos estudios con animales, los niveles de endotoxina en las muestras examinadas hasta la fecha han estado por debajo del nivel de detección del ensayo. En un estudio toxicológico preliminar, se inyectó esta proteína en dos ratones a la dosis terapéutica estimada 100X de 2,6 mg/ratón. Los animales sobrevivieron las dos semanas del estudio sin efectos de enfermedad aparentes. Estos resultados sugieren que ¹⁰F_{n3} se puede incorporar de forma segura en un fármaco IV.

Constructos alternativos para uso *in vivo*

Para prolongar la semivida del dominio de ¹⁰F_{n3} de 8 kD, también se ha construido una molécula más grande que imita a los anticuerpos naturales. Esta molécula de ¹⁰F_{n3}-F_c contiene los dominios -CH₁-CH₂-CH₃ (Figura 11) o -CH₂-CH₃ de la región constante de IgG del hospedante; en estos constructos, el dominio de ¹⁰F_{n3} se injerta sobre el término N en lugar del dominio de V_H de IgG (Figuras 11 y 12). Es de esperar que tales constructos de tipo anticuerpo mejoren la farmacocinética de la proteína, así como su capacidad para aprovechar la respuesta inmunitaria natural.

A fin de construir la forma murina del clon de ¹⁰F_{n3}-CH₁-CH₂-CH₃, primero se amplificó la región -CH₁-CH₂-CH₃ de una librería de ADNc de hígado y bazo de ratón (Clontech), y después se ligó en el vector pET25b. Los cebadores usados en la clonación fueron 5' Fc Nest y 3' 5 Fc Nest, y los cebadores usados para injertar los sitios de restricción apropiados sobre los extremos del inserto recuperado fueron 5' FC HIII y 3' FC Nhe:

| | | |
|---------------|--|--------------------|
| 5' Fc Nest 5' | GCG GCA GGG TTT GCT TAC TGG GGC CAA GGG 3' | (SEC ID nº: 15); |
| 3' Fc Nest 5' | GGG AGG GGT GGA GGT AGG TCA CAG TCC 3' | (SEC ID nº: 16); |
| 3' Fc Nhe 5' | TTT GCT AGC TTT ACC AGG AGA GTG GGA GGC 3' | (SEC ID nº: 17); y |
| 5' Fc HIII 5' | AAA AAG CTT GCC AAA ACG ACA CCC CCA TCT GTC 3' | (SEC ID nº: 18). |

Se usó una PCR adicional para eliminar la región CH₁ de este clon, y crear la parte Fc del clon ¹⁰F_{n3}-CH₂-CH₃ más corto. La secuencia codificante de ¹⁰F_{n3} se corta y empalma sobre el extremo 5' de cada clon; se puede usar el ¹⁰F_{n3} de tipo natural clonado a partir de la misma librería de ADNc de bazo de ratón, o un ¹⁰F_{n3} modificado, obtenido mediante mutagénesis o aleatorización de las moléculas. Los oligonucleótidos usados en la clonación del ¹⁰F_{n3} murino de tipo natural fueron:

Mo5PCR-NdeI

5' CATATGGTTTCTGATATCCGAGAGATCTGGAG 3' (SEC ID nº: 19);

Mo5PCR-His-NdeI (para un término N alternativo con la etiqueta de purificación His)

5' CAT ATG CAT CAC CAT CAC CAT CAG GTT TCT GAT ATT CCG AGA G 3' (SEC ID nº: 20); y

Mo3PCR-EcoRI

5' GAATTCCTATGTTTTATAATTGATGGAAAC 3' (SEC ID nº: 21).

Los equivalentes humanos de los clones se construyeron usando la misma estrategia con secuencias oligonucleotídicas humanas.

Andamiajes de ¹⁰F_{n3} en aplicaciones de chips proteínicos

La adecuabilidad del andamiaje de ¹⁰F_{n3} para aplicaciones de chips proteínicos es la consecuencia de (1) su capacidad para soportar muchas funciones de unión que se pueden seleccionar rápidamente en la mesa de trabajo o en una instalación automatizada, y (2) sus propiedades biofísicas superiores.

Las propiedades de unión versátiles de ¹⁰F_{n3} son una función de los bucles presentados por el plegamiento de sándwich beta de tipo inmunoglobulínico de ¹⁰F_{n3}. Como se explica anteriormente, estos bucles son similares a las regiones determinantes de la complementariedad de los dominios variables de anticuerpos, y pueden cooperar de forma similar a aquellos bucles de anticuerpos a fin de unirse a antígenos. En el sistema de la presente, los bucles de ¹⁰F_{n3} BC (restos 21-30), DE (restos 51-56), y FG (restos 76-87), se aleatorizan en secuencia, en longitud, o tanto en secuencia

ES 2 276 770 T3

como en longitud, a fin de generar diversas librerías de fusiones de ARNm-¹⁰F_n3. Los ligantes en tales librerías se enriquecen entonces basándose en su afinidad por una diana inmovilizada o marcada, hasta que se genera una pequeña población de ligantes de alta afinidad. También se pueden emplear una PCR y recombinación con tendencia al error para facilitar la maduración por afinidad de los ligantes seleccionados. Debido a los protocolos de selección y de maduración por afinidad rápidos y eficaces, se pueden seleccionar en un corto período de tiempo ligantes para un gran número de dianas.

Como un andamiaje para ligantes a ser inmovilizados sobre chips proteínicos, el dominio de ¹⁰F_n3 tiene la ventaja, con respecto a los fragmentos de anticuerpos y anticuerpos monocatenarios, de ser más pequeño y más fácil de manipular. Por ejemplo, a diferencia de los andamiajes monocatenarios o de los dominios variables aislados de anticuerpos, que pueden variar ampliamente en su estabilidad y solubilidad, y que requieren un entorno oxidante para conservar sus enlaces de disulfuro estructuralmente esenciales, ¹⁰F_n3 es extremadamente estable, con una temperatura de fusión de 110°C, y una solubilidad a una concentración > 16 mg/ml. El andamiaje de ¹⁰F_n3 también no contiene disulfuros ni cisteínas libres; en consecuencia, es insensible al potencial redox de su entorno. Una ventaja adicional de ¹⁰F_n3 es que sus bucles de unión a antígeno y su término N están en el extremo del sándwich de beta, opuesto al término C; de este modo, la unión de un andamiaje de ¹⁰F_n3 a un chip mediante su término C alinea los bucles de unión a antígenos, permitiendo su accesibilidad más elevada a la disolución que se ensaya. Puesto que ¹⁰F_n3 es un dominio único de sólo 94 restos de aminoácidos, también es posible inmovilizarlo sobre una superficie de chip a una densidad mayor que la usada para anticuerpos monocatenarios, con sus aproximadamente 250 restos. Además, la hidrofilia del andamiaje de ¹⁰F_n3, que se refleja en la elevada solubilidad de su dominio, conduce a una unión de fondo de ¹⁰F_n3 menor que la media a una superficie de un chip.

La estabilidad del andamiaje de ¹⁰F_n3 así como su adecuabilidad para la información de librería y selección de ligantes es probable que sean compartidas por una gran clase de tipo F_n3 de dominios proteínicos con un plegamiento de tipo inmunoglobulínico, tales como los dominios de tenascina, N-cadherina, E-cadherina, ICAM, titina, GCSF-R, receptor de citoquina, inhibidor de glicosidasa, y cromoproteína antibiótica. Las características claves compartidas por todos los citados dominios son un armazón estable proporcionado por dos láminas beta, que están empaquetadas una contra la otra y que están conectadas mediante por lo menos tres bucles accesibles a disolvente, por borde de la lámina; tales bucles se pueden aleatorizar para generar una librería de ligantes potenciales sin interrumpir la estructura del armazón (como se describe anteriormente).

Inmovilización de ligantes de armazones de fibronectina (ligantes de Fn)

Para inmovilizar ligantes de Fn a una superficie de un chip, se puede utilizar un número de técnicas ejemplares. Por ejemplo, los ligantes de Fn se pueden inmovilizar como fusiones de ARN-proteína mediante hibridación de Watson-Crick del resto de ARN de la fusión a un ADN complementario base inmovilizado sobre la superficie del chip (como se describe, por ejemplo, en Addressable Protein Arrays, documentos U.S.S.N. 60/080.686; U.S.S.N. 09/282.734; y WO 99/51773). Como alternativa, los ligantes de Fn se pueden inmovilizar como proteínas libres directamente sobre una superficie de un chip. Para la deposición de los ligantes de Fn sobre la superficie del chip, se pueden usar dispositivos manuales así como también robóticos. Se pueden usar robots que depositan puntos, para la deposición de ligantes de Fn con una densidad elevada en un formato de disposición en conjunto (por ejemplo, mediante el método de of Lueking *et al.*, *Anal Biochem.*, 15 de mayo 1999; 270(1):103-11). También se pueden utilizar diferentes métodos para anclar el ligante de Fn sobre la superficie del chip. Se puede usar un número de procedimientos de inmovilización estándares, incluyendo los descritos en *Methods in Enzymology* (K. Mosbach y B. Danielsson, eds.), volúmenes 135 y 136, Academic Press, Orlando, Florida, 1987; Nilsson *et al.*, *Protein Expr. Purif.* octubre de 1997; 11(1):1-16; y las referencias allí. La inmovilización orientada de ligantes de Fn puede ayudar a incrementar la capacidad de unión de los ligantes de Fn unidos a chips. Los enfoques ejemplares para lograr el acoplamiento orientado se describen en Lu *et al.*, *The Analyst* (1996), vol. 121, p. 29R-32R; y Turkova, *J Chromatogr B Biomed Sci App.* 5 de febrero de 1999; 722(1-2):11-31. Además, también se puede aplicar cualquiera de los métodos descritos aquí para anclar ligantes de Fn a superficies de chip a la inmovilización de ligantes de Fn sobre perlas, o sobre otros soportes.

Captura y detección de la proteína diana

Se pueden usar poblaciones seleccionadas de ligantes de Fn para la detección y/o cuantificación de dianas de análisis, por ejemplo en muestras tales como muestras biológicas. Para llevar a cabo este tipo de ensayo de diagnóstico, se inmovilizan ligantes de Fn seleccionados, para dianas de interés, sobre un soporte apropiado para formar chips proteínicos de múltiples características. A continuación, se aplica una muestra al chip, y se identifican los componentes de la muestra que se asocian con los ligantes de Fn basándose en la especificidad por la diana de los ligantes inmovilizados. Usando esta técnica, se pueden identificar o cuantificar simultáneamente uno o más componentes en una muestra (por ejemplo, con un medio para llevar a cabo un perfil de la muestra).

Los métodos de detección de diana permiten medir los niveles de dianas proteínicas unidas, e incluyen, sin limitación, radiografía, barrido de fluorescencia, espectroscopía de masas (MS), y resonancia plasmónica de superficie (SPR). La autorradiografía, que usa un sistema de formación de fosforoimágenes (Molecular Dynamics, Sunnyvale, CA), se puede usar para la detección y cuantificación de la proteína diana que se ha marcado radioactivamente, por ejemplo usando ³⁵S-metionina. El barrido de fluorescencia, usando un escáner láser (véase más abajo), se puede usar para la detección y cuantificación de dianas marcadas fluorescentemente. Como alternativa, el barrido de fluorescencia se puede usar para la detección de ligandos marcados fluorescentemente que se unen ellos mismos a la proteína diana

(por ejemplo, anticuerpos específicos de dianas marcados fluorescentemente, o estreptavidina marcada fluorescentemente que se une a la biotina diana, como se describe más abajo).

5 Se puede usar la espectroscopía de masas para detectar e identificar dianas unidas basándose en su masa molecular. La desorción de la proteína diana unida se puede lograr con asistencia mediante láser directamente a partir de la superficie del chip, como se describe más abajo. La detección de las masas también permite determinaciones, basándose en la masa molecular, de modificaciones de la diana, incluyendo modificaciones post-traduccionales como la fosforilación o glicosilación. La resonancia plasmónica de superficie se puede usar para la cuantificación de dianas proteínicas unidas, en la que el o los ligandos de Fn se inmovilizan sobre una superficie de oro adecuada (por ejemplo, como se
10 obtiene de Biacore, Suecia).

A continuación se describen esquemas ejemplares para seleccionar ligandos de Fn (en este caso, ligandos de Fn específicos para la proteína, TNF- α), y el uso de esas poblaciones seleccionadas para la detección sobre chips. Este ejemplo se proporciona con fines ilustrativos de la invención, y no se debe de interpretar como limitante.
15

Selección de ligantes de TNF- α basándose en andamiajes de ¹⁰F_{n3}

En un uso ejemplar para la selección de andamiajes de fibronectina sobre chips, se realizó una selección basada en ¹⁰F_{n3} frente a TNF- α , usando una librería de variantes de ¹⁰F_{n3} humanas con bucles aleatorizados BC, DE, y FG. La librería se construyó a partir de tres fragmentos de ADN, cada uno de los cuales contenía secuencias nucleotídicas que codificaban aproximadamente un tercio de ¹⁰F_{n3} humano, incluyendo uno de los bucles aleatorizados. Las secuencias de ADN que codificaron los restos del bucle enumerados anteriormente se reconstruyeron mediante síntesis oligonucleotídica, de forma que los codones para los restos de interés se sustituyeron por (NNS)_n, en el que N representa cualquiera de los cuatro desoxirribonucleótidos (A, C, G o T), y S representa C o G. El término C de cada fragmento
20 contenía la secuencia para la etiqueta de purificación FLAG.

Una vez extendido mediante Klenow, cada fragmento de ADN se transcribió, y el transcrito se ligó a un ligador de ADN que contiene puromicina, y se tradujo *in vitro*, como se describe por Szostak *et al.* (Roberts y Szostak, Proc. Natl. Acad. Sci USA 94:12297, 1997; Szostak *et al.*, documento WO 00/47775; Szostak *et al.*, documento WO 98/31700),
30 para generar una fusión de ARNm-péptido, que entonces se transcribió de forma inversa en una fusión de ADN-ARNm-péptido. La unión del péptido etiquetado con FLAG a agarosa M2 separó moléculas de fusión de longitud total de aquellas que contienen desplazamientos del marco o codones de parada superfluos; el ADN asociado con la fusión de longitud completa purificada se amplificó mediante PCR, después los tres fragmentos de ADN se cortaron mediante endonucleasa de restricción Ear I, y se ligaron para formar el molde de longitud completa. El molde se transcribió, y el transcrito se ligó a ligadores de ADN que contienen puromicina, y se tradujo para generar una librería de fusión de ¹⁰F_{n3}-ARN-proteína, que entonces se transcribió de forma inversa para producir la librería de fusión de
35 ADN-ARNm-péptido, que se usó subsiguientemente en la selección.

La selección de ligantes de TNF- α tuvo lugar en 50 mM de HEPES, pH 7,4, 0,02% de Triton-X, 0,1 mg/ml de ADN de esperma de salmón. La librería de fusión de ARN-proteína se incubó con TNF- α inmovilizado sobre Sefarosa; después de lavar, el ADN asociado con los ligantes más fuertemente unidos se eluyó con 0,1 M de KOH, se amplificó mediante PCR, se transcribió, y el transcrito se ligó, se tradujo y se transcribió de forma inversa en el material de partida para la siguiente ronda de selección.
40

Se realizaron diez rondas de tal selección (como se muestra en la Figura 13); dieron como resultado un conjunto de fusión de ARN-proteína que se une a TNF- α -Sefarosa con el Kd medio aparente de 120 nM. Los componentes clonales específicos del conjunto que se caracterizaron mostraron una unión a TNF- α en el intervalo de 50-500 nM.
45

Inmovilización del ligante de Fn, captura de la proteína diana, y detección mediante MALDI-TOF

50 Como una primera etapa para la inmovilización de los ligantes de Fn a una superficie de un chip, se preparó una sonda de captura oligonucleotídica con un sintetizador de ADN automatizado (PE BioSystems Expedite 8909), usando el enfoque de fosforamidito en un soporte sólido. Todos los reactivos se obtuvieron de Glen Research. La síntesis se inició con un soporte sólido que contiene un enlace de disulfuro para proporcionar eventualmente una funcionalidad tiólica 3'-terminal. Los primeros cuatro monómeros a añadir fueron unidades de óxido de hexaetileno, seguido de 20 monómeros T. El grupo DMT 5'-terminal no se eliminó. La sonda de captura se escindió del soporte sólido, y se desprotegió con hidróxido de amonio, se concentró hasta sequedad en una centrifuga de vacío, y se purificó mediante HPLC de fase inversa usando un gradiente de acetonitrilo en tampón de acetato de trietilamonio. Se recogieron fracciones apropiadas a partir de la HPLC, se evaporaron hasta sequedad en una centrifuga de vacío, y el grupo DMT 5'-terminal se eliminó mediante tratamiento con 80% de AcOH durante 30 minutos. El ácido se eliminó por evaporación, y el oligonucleótido se trató entonces con 100 mM de DTT durante 30 minutos, para romper el enlace de disulfuro. El DTT se eliminó mediante extracción repetida con EtOAc. El oligonucleótido se precipitó mediante etanol a partir de la capa acuosa que queda, y se comprobó su pureza mediante HPLC de fase inversa.
60

65 La sonda de captura 3'-tiólica se ajustó hasta 250 μ M en tampón 1X PBS desgasificado, y se aplicó como una gotita única (75 μ l) a un chip revestido con oro de 9 x 9 mm (Biacore) en una cámara inundada de argón que contiene una pequeña cantidad de agua. Después de 18 horas a temperatura ambiente, la disolución de la sonda de captura se eliminó, y el chip funcionalizado se lavó con 50 ml de tampón 1X PBS (2 x durante 15 minutos cada vez) con

ES 2 276 770 T3

agitación suave, y después se aclaró con 50 ml de agua (2 x durante 15 minutos cada vez) de la misma manera. El líquido que queda se eliminó con cuidado, y los chips funcionalizados se usaron inmediatamente, o se almacenaron a 4°C en argón.

5 Se trató aproximadamente 1 pmol de conjunto de fusión de ¹⁰F_n3 procedente de la décima ronda de selección de TNF- α (anteriormente) con ARNasa A durante varias horas, se ajustó hasta 5X SSC en 70 μ l, y se aplicó a un chip de oro funcionalizado procedente de lo anterior, como una única gotita. Se usó un dispositivo de cesto de 50 μ l de volumen para cerrar herméticamente la mezcla de fusión con el chip funcionalizado, y el aparato se hizo girar de forma continua a 4°C. Después de 18 horas, el aparato se abrió, el chip de oro se lavó con 50 ml de 5X SSC
10 durante 10 minutos, con agitación suave. El exceso de líquido se eliminó con cuidado de la superficie del chip, y el chip se pasivó con una disolución de bloqueo (1X TBS + 0,02% de Tween-20 + 0,25% de BSA) durante 10 minutos a 4°C. El exceso de líquido se eliminó con cuidado, y la disolución que contiene 500 μ g/ml de TNF- α en la misma disolución de bloqueo de la composición se aplicó al chip como una única gotita, y se incubó a 4°C durante dos horas con mezclamiento ocasional de la gotita vía Pipetman. Después de la eliminación de la disolución de unión, el chip
15 se lavó durante 5 minutos a 4°C con agitación suave (50 ml de 1X TBS + 0,02% de Tween-20), y entonces se secó a temperatura ambiente. Se preparó un segundo chip exactamente como se describe anteriormente, excepto que no se añadió la fusión a la mezcla de hibridación.

Después, se aplicó uniformemente una matriz de MALDI-TOF (15 mg/ml de ácido 3,5-dimetoxi-4-hidroxibenzoico en 1:1 de etanol/10% de ácido fórmico en agua) a los chips de oro, con un robot de 3 ejes de alta precisión (MicroGrid, BioRobotics). Se usó una herramienta de 16 púas para transferir la matriz desde una placa de microtitulación de 384 pocillos a los chips, produciendo características de diámetro de 200 micrómetros con un campo de 600 micrómetros. Los ajustes del instrumento del espectrómetro de masas mediante MALDI-TOF (Voyager DE, PerSeptive Biosystems) fueron los siguientes: voltaje acelerador = 25 k, voltaje de la cuadrícula = 92%, voltaje del alambre guía = 0,05%, retraso = 200 on, potencia del láser = 2400, puerta de masa baja = 1500, iones negativos = apagado. Los chips de oro se colocaron individualmente en una etapa de muestra de MALDI modificada para mantener el nivel del chip igual al nivel de la etapa, permitiendo de este modo la distancia apropiada de vuelo. Se usaron el monitor de vídeo del instrumento y el sistema de control de movimiento para dirigir el haz de láser a características individuales de la matriz.
20

Las Figuras 14 y 15 muestran los espectros de masas del chip de fusión de ¹⁰F_n3 y el chip de no fusión, respectivamente. En cada caso, se analizó un número pequeño de características de 200 micrómetros para recoger los espectros, pero la Figura 15 requirió significativamente más adquisiciones. La señal a 17,5 kDa corresponde al monómero de TNF- α .
25

35 *Inmovilización del ligante de Fn, captura de la proteína diana, y detección mediante fluorescencia*

Se trataron portaobjetos de microscopio de vidrio, previamente limpios, de 1 x 3 pulgadas (2,54 x 7,66 cm) (Gold-seal, #3010) con Nanostrip (Cyantek) durante 15 minutos, con NaOH acuoso al 10% a 70°C durante 3 minutos, y con HCl acuoso al 1% durante 1 minuto, y se aclararon a conciencia con agua desionizada después de cada reactivo. Los portaobjetos se secaron entonces en un secador de vacío sobre sulfato de calcio anhidro durante varias horas. Se preparó una disolución al 1% de aminopropiltrimetoxisilano en 95% de acetona/5% de agua, y se dejó hidrolizar durante 5 minutos. Los portaobjetos de vidrio se sumergieron en la disolución silánica hidrolizada durante 5 minutos, con agitación suave. El exceso de silano se eliminó sometiendo los portaobjetos a diez lavados de 5 minutos, usando porciones recientes de 95% de acetona/5% de agua para cada lavado, con agitación suave. Los portaobjetos se curaron entonces calentando a 110°C durante 20 minutos. Los portaobjetos tratados con silano se sumergieron en una disolución al 0-2%, recientemente preparada, de 1,4-diisotiocianato de fenileno en 90% de DMF/10% de piridina, durante dos horas, con agitación suave. Los portaobjetos se lavaron secuencialmente con 90% de DMF/10% de piridina, con metanol y con acetona. Después de secar al aire, los portaobjetos funcionalizados se almacenaron a 0°C en un secador a vacío sobre sulfato de calcio anhidro. Se obtuvieron resultados similares con portaobjetos comerciales reactivos con aminas (3-D Link, Surmodics).
40

Se prepararon sondas de captura oligonucleotídicas con un sintetizador de ADN automatizado (PE BioSystems Expedite 8909), usando la química de fosforamidito convencional. Todos los reactivos provenían de Glen Research. La síntesis se inició con un soporte sólido que contiene una funcionalidad amino protegida ortogonalmente, con lo que la amina 3'-terminal no está enmascarada hasta la etapa de desprotección final. Los primeros cuatro monómeros a añadir fueron unidades de óxido de hexaetileno, seguido de los monómeros A, G, C y T estándares. Todas las oligosecuencias de captura se escindieron del soporte sólido y se desprotegeron con hidróxido de amonio, se concentraron hasta sequedad, se hicieron precipitar en etanol, y se purificaron mediante HPLC de fase inversa usando un gradiente de acetonitrilo en tampón de acetato de trietilamonio. Se recogieron fracciones apropiadas procedentes de la HPLC, se evaporaron hasta sequedad en una centrifuga de vacío, y después se coevaporaron con una porción de agua. Los oligos de captura purificados, marcados con amina, se ajustaron hasta una concentración de 250 μ M en 50 mM de tampón de carbonato de sodio (pH 9,0) que contiene 10% de glicerol. Las sondas se colocaron como manchas sobre una superficie de vidrio reactiva con amina, en posiciones definidas en un patrón de disposición 5 x 5 x 6, con un robot de 3 ejes (MicroGrid, BioRobotics). Se usó una herramienta con 16 agujas para transferir el líquido desde las placas de microtitulación de 384 pocillos, produciendo características de 200 micrómetros con un nivel de 600 micrómetros. Cada subcuadrícula de 24 características representa una única sonda de captura (es decir, 24 puntos duplicados). La disposición del conjunto se incubó a temperatura ambiente en un entorno saturado de humedad, durante 12-18
45
50
55
60
65

ES 2 276 770 T3

horas. La reacción de adhesión se terminó sumergiendo los chips en hidróxido amónico acuoso al 2%, durante cinco minutos, con agitación suave, seguido del aclarado con agua destilada (3X durante 5 minutos cada vez). La disposición del conjunto se empapó finalmente en disolución 10X PBS durante 30 minutos a temperatura ambiente, y después se aclaró nuevamente durante 5 minutos en agua destilada.

5

Se identificaron las secuencias específicas y termodinámicamente isoenergéticas a lo largo del ARNm de ¹⁰F_n3, para que sirvan como puntos de captura para el autoensamblaje y el anclaje de la proteína de ¹⁰F_n3. El programa de ordenador HybSimulator v4.0 (Advanced Gene Computing Technology, Inc.) facilitó la identificación y el análisis de sondas potenciales de captura. Se escogieron seis sondas de captura únicas, y se imprimieron sobre el chip, tres de las cuales son complementarias a las regiones comunes del ARNm del conjunto de fusión de ¹⁰F_n3 (CP3', CP5', y CPflag). Las tres secuencias que quedan (CPneg1, CPneg2, y CPneg3) no son complementarias, y funcionan en parte como controles negativos. Cada una de las sondas de captura presenta un término 3'-amino y cuatro unidades espaciadoras de óxido de hexaetileno, como se describe anteriormente. Lo siguiente es una lista de las secuencias de sondas de captura que se emplearon (5'-3'):

15

| | | |
|---------|----------------------|-----------------|
| CP3': | TGTAATAGTAATTGTCCC | (SEC ID n°: 22) |
| CP5': | TTTTTTTTTTTTTTTTTTTT | (SEC ID n°: 23) |
| CPneg1: | CCTGTAGGTGTCCAT | (SEC ID n°: 24) |
| CPflag: | CATCGTCCTTGTAGTC | (SEC ID n°: 25) |
| CPneg2: | CGTCGTAGGGTA | (SEC ID n°: 26) |
| CPneg3: | CAGGTCTTCTCAGAGA | (SEC ID n°: 27) |

20

Se ajustó aproximadamente 1 pmol de conjunto de fusión de ¹⁰F_n3 procedente de la décima ronda de selección de TNF- α hasta 5X SSC que contiene 0,02% de Tween-20 y 2 mM de complejo de vanadil ribonucleótido, en un volumen total de 350 μ l. Todo el volumen se aplicó a la microcolección en un dispositivo de cesto de 400 μ l, y el montaje se hizo girar continuamente durante 18 horas a temperatura ambiente. Después de la hibridación, el portaobjetos se lavó secuencialmente con porciones de 500 ml agitadas de 5X SSC, 2,5X SSC y 1X SSC durante 5 minutos cada vez. Las trazas de líquido se eliminaron mediante centrifugación, y se dejó que el portaobjetos se secase al aire.

30

Se recogió TNF- α humano recombinante (500 μ g, liofilizado, de Prepro Tech) en 230 μ l de 1X PBS, y se dializó frente a 700 ml de 1X PBS agitado, a 4°C durante 18 horas en una unidad de microdialización (corte de peso molecular 3.500, Pierce). El TNF- α dializado se trató con reactivo de biotilación NHS-LC-LC ligado a EZ (20 μ g, Pierce) durante 2 horas a 0°C, y se dializó nuevamente frente a 700 ml de 1X PBS agitado, a 4°C durante 18 horas en una unidad microdializadora (corte de peso molecular 3.500, Pierce). El conjugado resultante se analizó mediante espectrometría de masas de MALDI-TOF, y se encontró que estaba casi completamente funcionalizado con un único resto de biotina.

35

Cada uno de los siguientes procesos se realizó a 4°C con agitación continua o con mezclamiento. La superficie del microconjunto de proteínas se pasivó mediante tratamiento con 1X TBS que contiene 0,02% de Tween-20 y 0,2% de BSA (200 μ l) durante 60 minutos. El TNF- α biotilado (concentración de 100 nM obtenida en el tampón de pasivación) se puso en contacto con el microconjunto durante 120 minutos. El microconjunto se lavó con 1X TBS que contiene 0,02% de Tween-20 (3X 50 ml, 5 minutos cada lavado). La estreptavidina marcada fluorescentemente (2,5 μ g/ml de conjugado de Alexa 546-estreptavidina, de Molecular Probes, obtenido en el tampón de pasivación) se puso en contacto con el microconjunto durante 60 minutos. El microconjunto se lavó con 1X TBS que contiene 0,02% de Tween-20 (2X 50 ml, 5 minutos cada lavado), seguido de un aclarado de 3 minutos con 1X TBS. Las trazas de líquido se eliminaron por centrifugación, y el portaobjetos se dejó secar al aire a temperatura ambiente.

45

Se realizó el barrido mediante láser de fluorescencia con un sistema GSI Lumonics ScanAtray 5000, usando una resolución de píxel de 10 μ M, y longitudes de onda preestablecidas de excitación y de emisión para el colorante de Alexa 546. El análisis de fosforoimágenes se realizó con un sistema Molecular Dynamics Storm. El tiempo de exposición fue de 48 horas con contacto directo entre la microcolección y la pantalla de almacenamiento de fósforo. El barrido de fosforoimágenes se realizó en el ajuste de resolución de 50 μ M, y los datos se extrajeron con un programa de ordenador ImageQuant v.4.3.

50

55

Las Figuras 16 y 17 son el barrido de fosforoimágenes y de fluorescencia, respectivamente, del mismo conjunto. La fosforoimagen muestra cuándo se hibridó la fusión de ¹⁰F_n3 basándose en la señal de ³⁵S-metionina. El barrido de fluorescencia muestra cuándo se unió el TNF- α marcado.

60

Otras formas de realización

Otras formas de realización se encuentran comprendidas en las reivindicaciones.

65

REIVINDICACIONES

- 5 1. Conjunto de proteínas inmovilizadas sobre un soporte sólido, comprendiendo cada una de dichas proteínas un dominio de tipo III de fibronectina que tiene por lo menos un bucle aleatorizado, por lo menos una lámina β aleatorizada, o una combinación de los mismos, y estando **caracterizado** porque es capaz de unirse a un compuesto que no está unido mediante una fibronectina correspondiente de origen natural.
- 10 2. Conjunto según la reivindicación 1, en el que dicho dominio de tipo III de fibronectina es un dominio de tipo III de fibronectina de mamífero.
3. Conjunto según la reivindicación 2, en el que dicho dominio de tipo III de fibronectina es un dominio de tipo III de fibronectina humano.
- 15 4. Conjunto según la reivindicación 1, en el que cada una de dichas proteínas comprende el décimo módulo de dicho dominio de tipo III de fibronectina ($^{10}\text{Fn3}$).
5. Conjunto según la reivindicación 4, en el que la aleatorización de uno, dos o tres bucles da como resultado la unión de la proteína a dicho compuesto.
- 20 6. Conjunto según la reivindicación 5, en el que la aleatorización de los dos bucles da como resultado la unión de la proteína a dicho compuesto.
7. Conjunto según la reivindicación 6, en el que la aleatorización de los tres bucles da como resultado la unión de la proteína a dicho compuesto.
- 25 8. Conjunto según la reivindicación 4, en el que dicho $^{10}\text{Fn3}$ carece de un motivo de unión a integrina.
9. Conjunto según la reivindicación 1, en el que cada una de dichas proteínas carece de enlaces de disulfuro.
- 30 10. Conjunto según la reivindicación 1, en el que cada una de dichas proteínas es un monómero o un dímero.
11. Conjunto según la reivindicación 1, en el que cada una de dichas proteínas está unida covalentemente a un ácido nucleico.
- 35 12. Conjunto según la reivindicación 11, en el que dicho ácido nucleico codifica la proteína unida covalentemente.
13. Conjunto según la reivindicación 12, en el que dicho ácido nucleico es ARN.
- 40 14. Conjunto según la reivindicación 1, en el que dicho soporte sólido es un chip.
15. Método para identificar una proteína que se une a un compuesto, comprendiendo dicho método:
- 45 (a) poner en contacto dicho compuesto con un conjunto de proteínas candidatas inmovilizadas sobre un soporte sólido, comprendiendo cada una de dichas proteínas candidatas un dominio de tipo III de fibronectina que tiene por lo menos un bucle aleatorizado, una lámina β aleatorizada, o una combinación de los mismos, llevándose a cabo dicho contacto en condiciones que permiten la formación del complejo de compuesto-proteína; y
- 50 (b) identificar, a partir de dicho complejo, una proteína que se une a dicho compuesto.
16. Método para identificar un compuesto que se une a una proteína, comprendiendo dicha proteína un dominio de tipo III de fibronectina que tiene por lo menos un bucle aleatorizado, por lo menos una lámina β aleatorizada, o una combinación de los mismos, comprendiendo dicho método:
- 55 (a) poner en contacto un conjunto de proteínas inmovilizadas sobre un soporte sólido con un compuesto candidato, comprendiendo cada una de dichas proteínas un dominio de tipo III de fibronectina que tiene por lo menos un bucle aleatorizado, una lámina β aleatorizada, o una combinación de los mismos, llevándose a cabo dicho contacto en condiciones que permiten la formación del complejo de compuesto-proteína; y
- 60 (b) identificar, a partir de dicho complejo, un compuesto que se une a una proteína del conjunto.
17. Método según la reivindicación 15, comprendiendo asimismo dicho método las etapas siguientes:
- 65 (c) aleatorizar además una proteína que se une a dicho compuesto en la etapa (b);
- (d) formar un conjunto sobre un soporte sólido con las proteínas aleatorizadas adicionales de la etapa (c); y

ES 2 276 770 T3

- (e) repetir las etapas (a) y (b) usando, en la etapa (a), el conjunto de proteínas aleatorizadas adicionales como dicho conjunto de proteínas candidatas.

18. Método según la reivindicación 16, comprendiendo asimismo dicho método las etapas siguientes:

- (c) modificar el compuesto que se une a dicha proteína en la etapa (b); y
- (d) repetir las etapas (a) y (b) usando, en la etapa (a), dicho compuesto modificado adicional como dicho compuesto candidato.

19. Método según la reivindicación 15 ó 16, en el que dicho soporte sólido es un chip.

20. Método para detectar un compuesto en una muestra, comprendiendo dicho método:

- (a) poner en contacto una muestra con una proteína que se une a dicho compuesto, y que comprende un dominio de tipo III de fibronectina que tiene por lo menos tres bucles aleatorizados, llevándose a cabo dicho contacto en condiciones que permiten la formación del complejo de compuesto-proteína; y
- (b) detectar dicho complejo, detectando de ese modo dicho compuesto en dicha muestra,

en el que dicha proteína se inmoviliza sobre un soporte sólido, y en el que dicha proteína se inmoviliza sobre dicho soporte sólido como parte de un conjunto.

21. Método según la reivindicación 20, en el que dicho soporte sólido es una perla o un chip.

22. Método según la reivindicación 15, 16 ó 20, en el que dicho compuesto es una proteína.

23. Método según la reivindicación 15, 16 ó 20, en el que dicho dominio de tipo III de fibronectina es un dominio de tipo III de fibronectina de mamífero.

24. Método según la reivindicación 23, en el que dicho dominio de tipo III de fibronectina es un dominio de tipo III de fibronectina humano.

25. Método según la reivindicación 15, 16 ó 20, en el que cada una de dichas proteínas comprende el décimo módulo de dicho dominio de tipo III de fibronectina (¹⁰Fn3).

26. Método según la reivindicación 25, en el que la aleatorización de los tres bucles da como resultado la unión de dicha proteína a dicho compuesto.

27. Método según la reivindicación 25, en el que dicho ¹⁰Fn3 carece de un motivo de unión a integrina.

28. Método según la reivindicación 15, 16 ó 20, en el que cada una de dichas proteínas está unida covalentemente a un ácido nucleico.

29. Método según la reivindicación 28, en el que dicho ácido nucleico codifica la proteína unida covalentemente.

30. Método según la reivindicación 29, en el que dicho ácido nucleico es ARN.

31. Método según la reivindicación 15, 16 ó 20, en el que dicho complejo o dicho compuesto se detectan mediante radiografía, detección mediante fluorescencia, espectroscopía de masas, o resonancia plasmónica de superficie.

32. Conjunto según la reivindicación 1, en el que dicha proteína que contiene el dominio de tipo III de fibronectina tiene por lo menos una mutación en la lámina β .

33. Conjunto según la reivindicación 1, en el que dicha proteína que contiene el dominio de tipo III de fibronectina tiene por lo menos un bucle aleatorizado.

34. Conjunto según la reivindicación 1, en el que dicha proteína que contiene el dominio de tipo III de fibronectina tiene por lo menos tres bucles alterados.

35. Conjunto según la reivindicación 34 o método según la reivindicación 20, en los que dichos tres bucles son los bucles BC, DE, FG del dominio de tipo III de fibronectina.

36. Método según la reivindicación 15 ó 16, en el que dicha proteína que contiene el dominio de tipo III de fibronectina tiene por lo menos una mutación en la lámina β .

37. Método según la reivindicación 15 ó 16, en el que dicha proteína que contiene el dominio de tipo III de fibronectina tiene por lo menos un bucle aleatorizado.

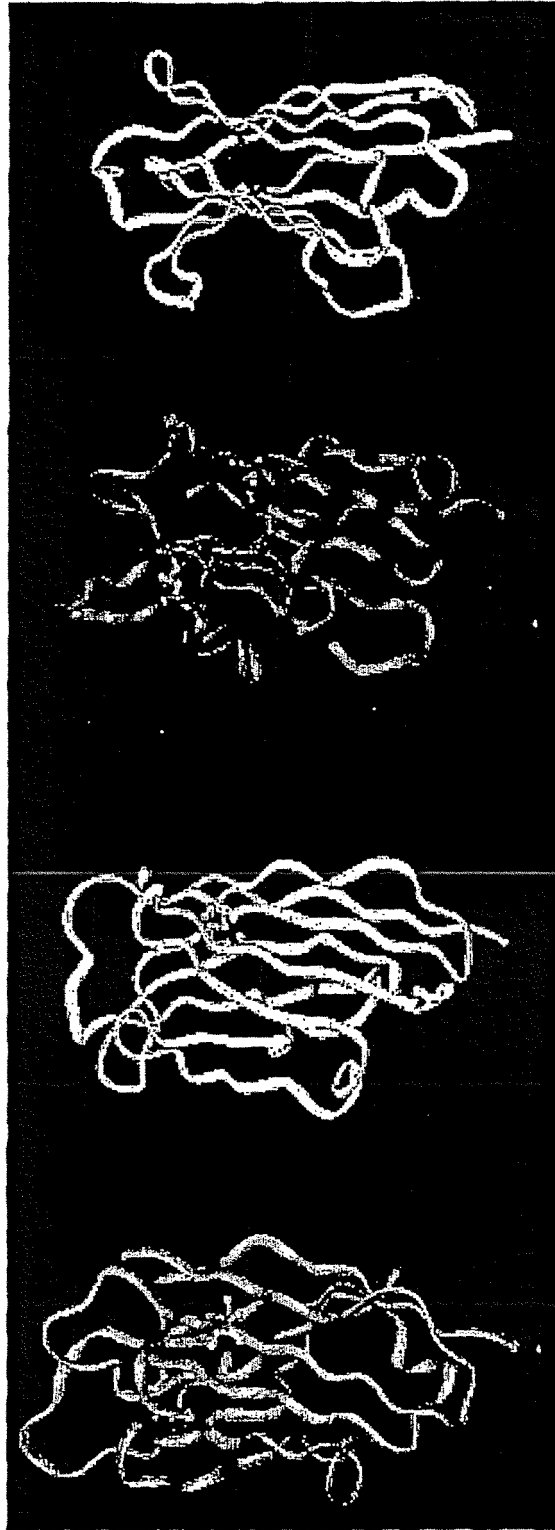


Fig. 1

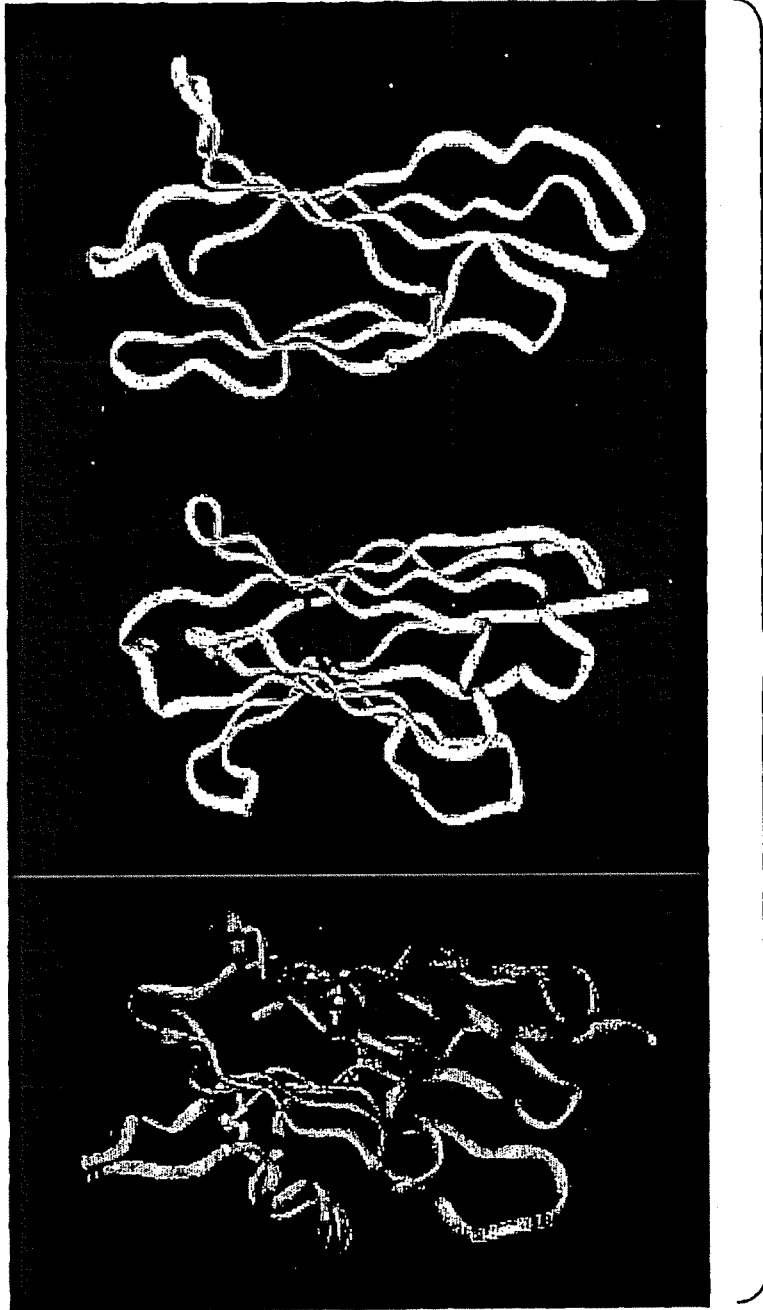


Fig. 2

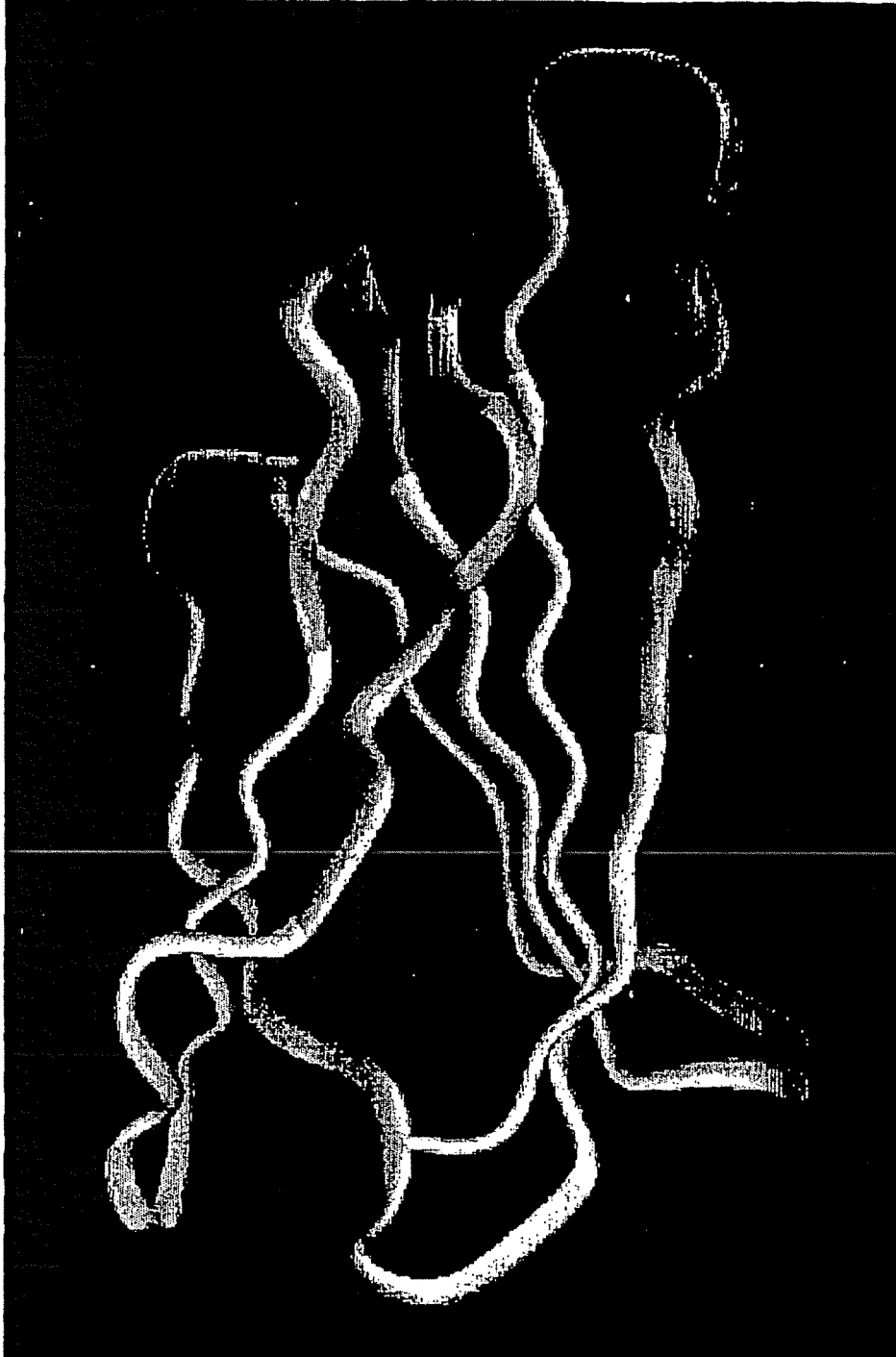


Fig. 3

| | | | | | | | | | | | | | | | | | | | | | |
|--------|------------|------------|------------|------------|------------|------------|------------|-------------|------------|----------|------|----|----|----|----|----|----|----|----|----|----|
| Hs FND | VSDVPRD-LE | 1 | 9 | 10 | 19 | 20 | 29 | 30 | 37 | 38 | 47 | 48 | 57 | 58 | 67 | 68 | 77 | 78 | 87 | 88 | 94 |
| Bt FN | VSDVPRD-LE | VIAATPTSL | ISWDAPAVTV | RYRITVYG-- | ETGGSFVQE | FVPGSKSTA | TISGLKPGVD | YITVYAVTG | RGSPASSXP | ISINVRT | 94 | | | | | | | | | | |
| Rn FN | VSDVPRD-LE | VIAATPTSL | ISWDAPAVTV | RYRITVYG-- | ETGGSFVQE | FVPGSKSTA | TISGLKPGVD | YITVYAVTG | RGSPASSXP | VSINVRT | 1510 | | | | | | | | | | |
| Mm FN | VSDVPRD-LE | VIAATPTSL | ISWDAPAVTV | RYRITVYG-- | ETGGSFVQE | FVPGSKSTA | TINAIKPGAD | YITVYAVTG | RGSPASSXP | VSINVRT | 1611 | | | | | | | | | | |
| Oc FN | VSDVPRD-LE | VIAATPTSL | ISWDAPAVTV | RYRITVYG-- | ETGGSFVQE | FVPGSKSTA | TINAIKPGAD | YITVYAVTG | RGSPASSXP | VSINVRT | 712 | | | | | | | | | | |
| Gg FN | VSDVPRD-LE | VNPSTPTSL | ISWDAPAVTV | RYRITVYG-- | ETGGSFVQE | FVPGSKSTA | TITGLKPGVD | YITVYAVTG | RGSPASSXP | VTVYTKT | 64 | | | | | | | | | | |
| Xl FN | VSDVPTD-LE | VTSSTPTLT | ISWEAPAVSV | RYRITVYS-- | QTGCHGPEKE | FVPGTSMTA | TIRGLNPGVS | YITVYAVTG | RGSPASSXP | LTIHKT | 443 | | | | | | | | | | |
| Cf FN | AIDAPSn-Lr | FLATTPnSLL | VSNDPPrArI | TGYIIKVe-- | kPGSPPrEW | pPrPGVTeA | TITGLePTE | YITQVIALKn | NQKSePLIGr | kKTDEL- | 197 | | | | | | | | | | |
| Ec FN | AIDAPSn-Lh | FLATTPnSLL | ISWQPrArI | TGYIIKVe-- | kPGSPPrEW | pPrbPGVTeA | TITGLePTE | YITQVIALKn | NQKSePLIGr | rKTdEp- | 197 | | | | | | | | | | |
| Hs TC | VS-PPKD-Ly | VTeVTeTVN | LAWDn-eMrV | TeYLIVYTP- | -THEGGLEMQ | FrVPGQPrST | IIQeLePVE | YFIRVFPALIE | NKKSIPVSAI | V----- | 686 | | | | | | | | | | |
| Ss TP | VS-PPKD-Li | VTeVTeTVN | LAWDn-eMrV | TeYLIVYTP- | -THEGGLEMQ | FrVPGQPrST | TIRLePVE | YFIRVFPALIE | NKKSIPVSAI | V----- | 686 | | | | | | | | | | |
| Mm TX | MiDPQD-Lr | VAVTPTID | LSWIrPQaev | DrFVVSXV-- | --SAGnQrVr | LeVPPeADrT | QLTdLMPGVE | YVTVTAERG | HAVSYPASIr | ANTG---- | 889 | | | | | | | | | | |
| Hs CAP | TlpVvVSLn | IYdVGTTH | VQWQp-VGGA | TGYILSYkPv | kDTEPrPK | VcLGPVNDM | QLTdLVPNTE | YAVTVQAVILh | dLTSePVTvr | e----- | 1551 | | | | | | | | | | |
| Oc C12 | TlpVvVSLN | IYdVGTTH | VQWQp-VGGA | TGYILSYkPv | kDTEPrPK | QdVklRQVTH | ----- | ----- | ----- | ----- | 322 | | | | | | | | | | |
| Gg C14 | LalPmaSDLk | LYdVShSSMR | ARWnG-VAGA | TGYMILYAPL | TEGLAAeKE | IKIGeASTel | ELDGLLENTE | YTVTVYAMF- | ----- | ----- | 508 | | | | | | | | | | |
| Hs U1 | LalPmaSDLI | LYdVtenSMR | VKWDn-VpGA | SGYLLIYAPL | TEGLAGeKE | MKIGeHTIdI | ELSGLLENTE | YTVTVYAMFG | eeASDPVTGq | e----- | 321 | | | | | | | | | | |

| | | | | | | | | | | | | | | | | | | | | | |
|-------|-------------------------------------|---------------|-------|------------------|--------|-------|-----------------------|---------|-------|-------|---------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| var. | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| cons. | P | L | V | SL | M | V | Y | I | Y | I | L | PGVD | Y | ITV | A | G | S | P | | | |
| | M | I | I | TV | A | F | V | | | | | L | I | NTE | VQL | N | R | | | | |
| | | L | F | M | I | I | L | | | | | AS | R | R | E | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | |
| CAP | Precursor de colágeno alfa | | | | | | | | | | NEGRITA | | | | | | | | | | |
| C12 | Colágeno tipo 12 | | | | | | | | | | | | | | | | | | | | |
| FND | Dominio de tipo III de fibronectina | | | | | | | | | | | | | | | | | | | | |
| FN | Fibronectina | | | | | | | | | | | | | | | | | | | | |
| TP | Precursor de tenascina | | | | | | | | | | | | | | | | | | | | |
| TC | Tenascina C | | | | | | | | | | | | | | | | | | | | |
| U1 | Undulina 1 | | | | | | | | | | | | | | | | | | | | |
| Bt | Bovis taurus | vaca | Cf | Canis familiaris | perro | Ec | Equus caballis | caballo | | | | | | | | | | | | | |
| Ss | Sus scrofa | cerdo | Hs | Homo sapiens | hombre | Oc | Oryctolagus cuniculus | conejo | | | | | | | | | | | | | |
| XI | Xenopus laevis | sapo africano | | | | | | | | | | | | | | | | | | | |

Fig. 4

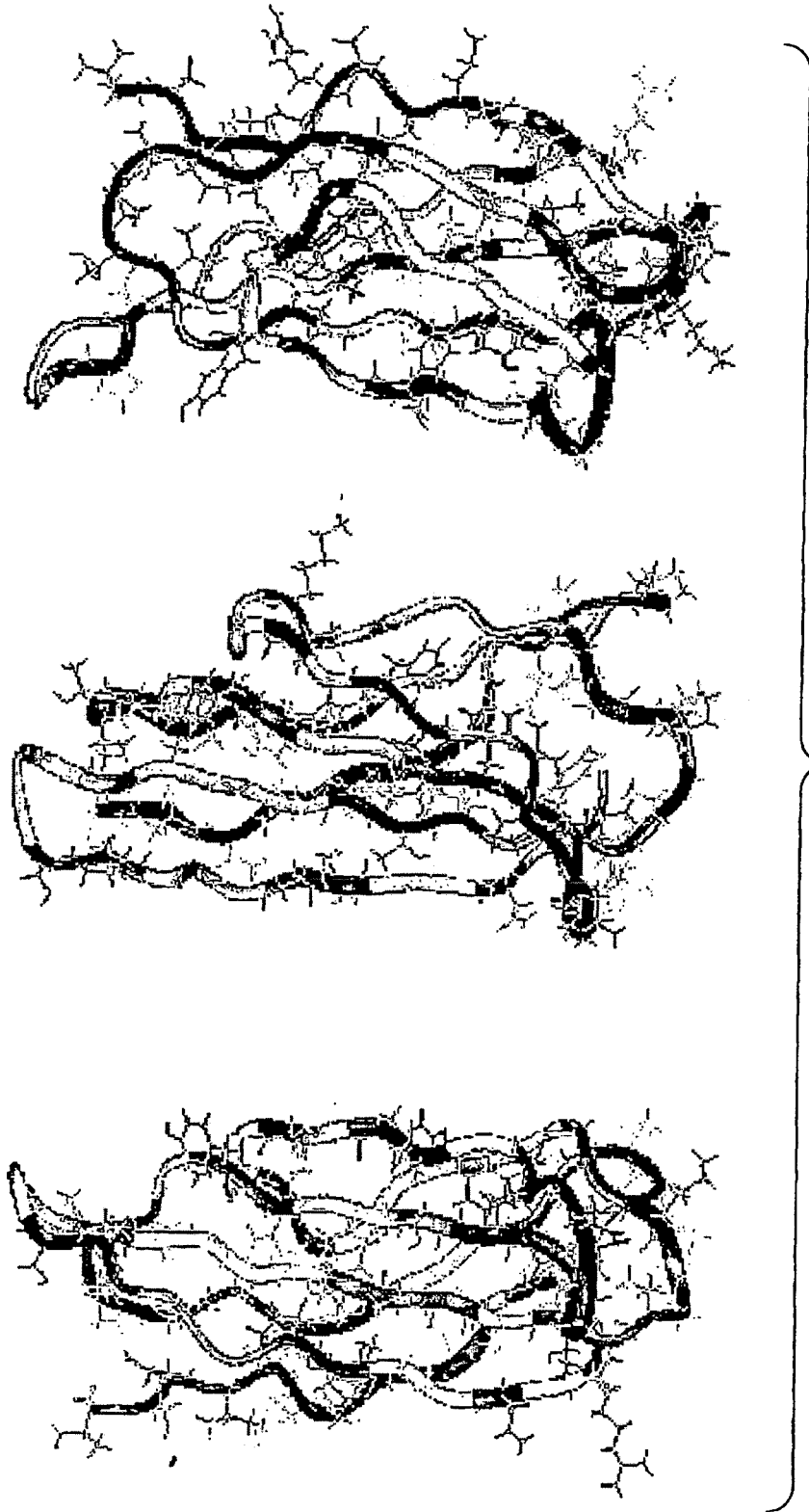
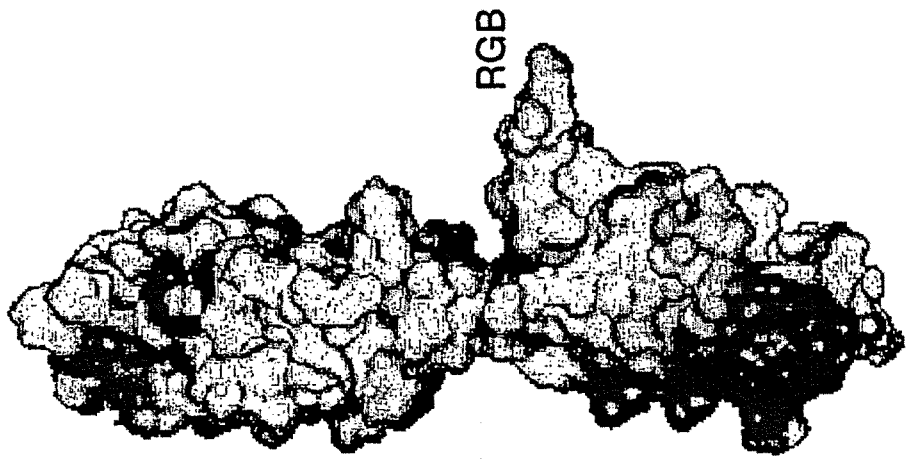
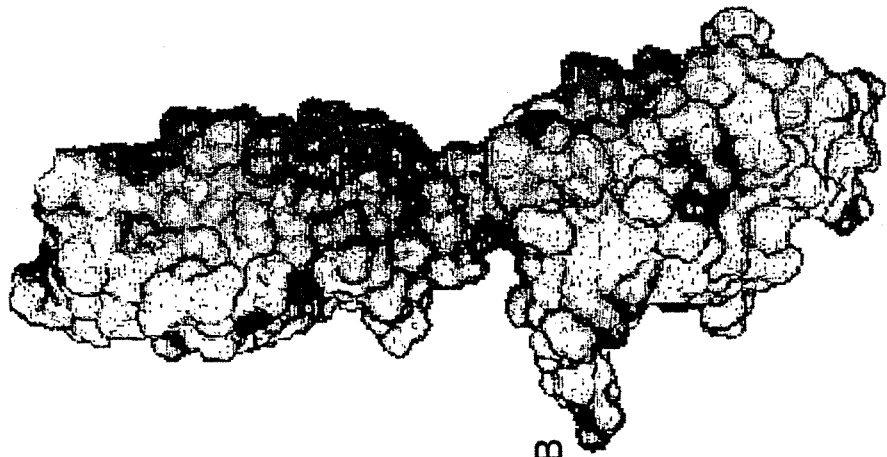


Fig. 5



9Fn3

10Fn3



RGB

Fig. 6

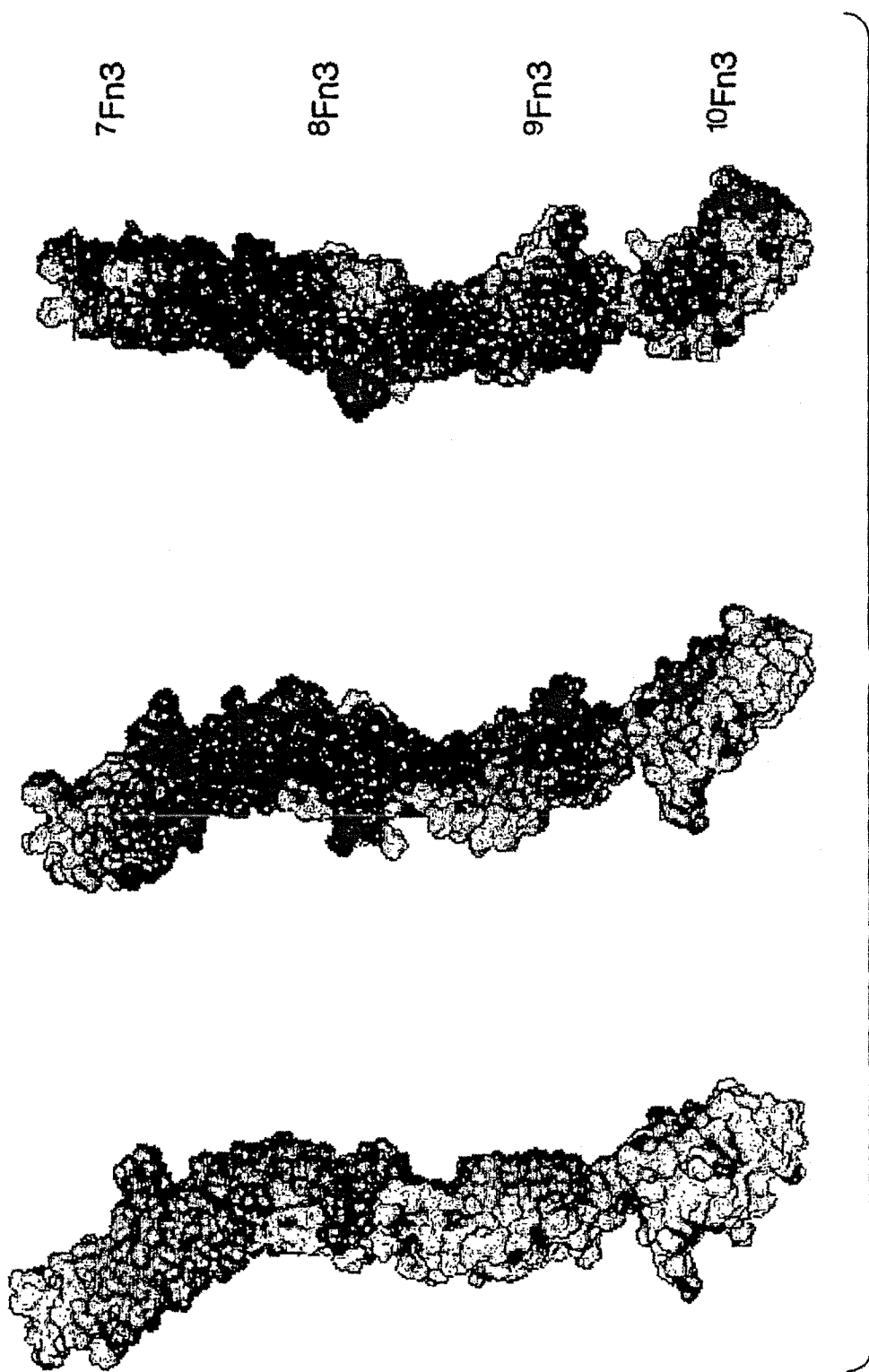


Fig. 7

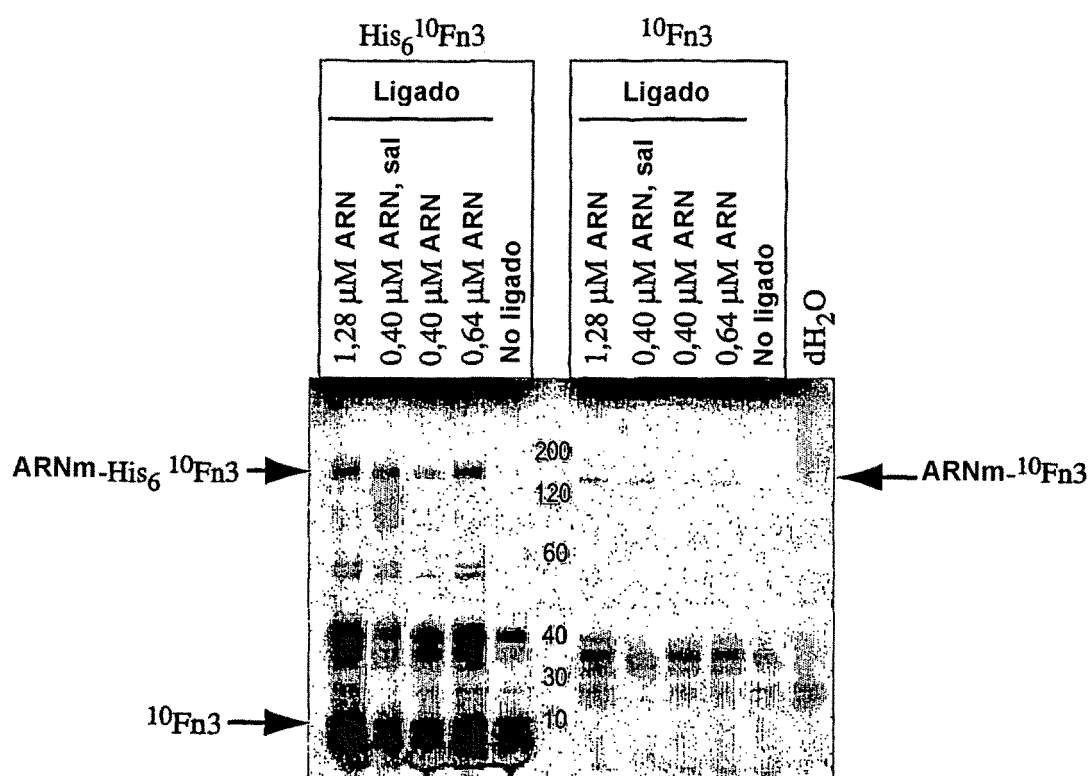


Fig. 8

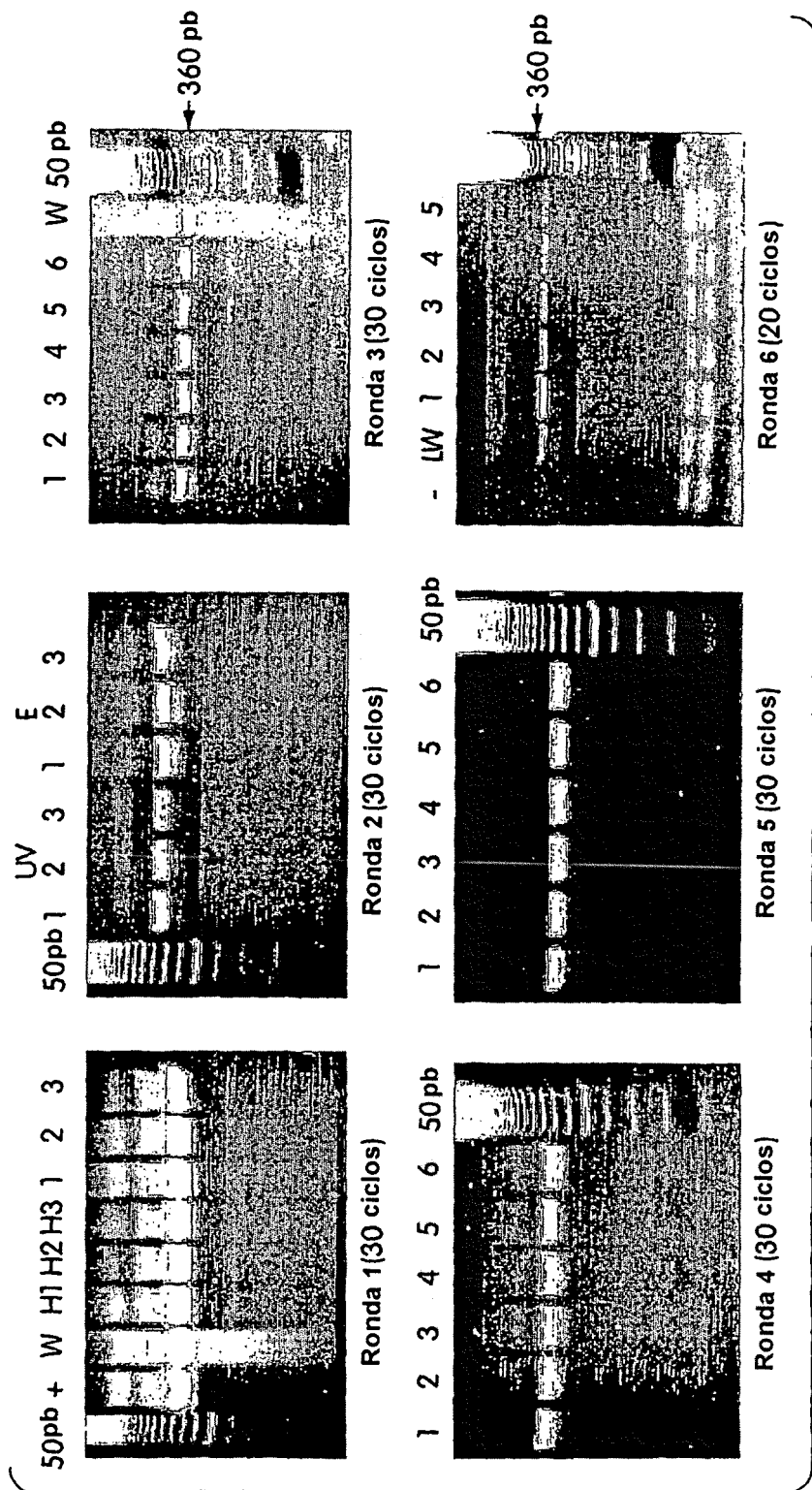


Fig. 9

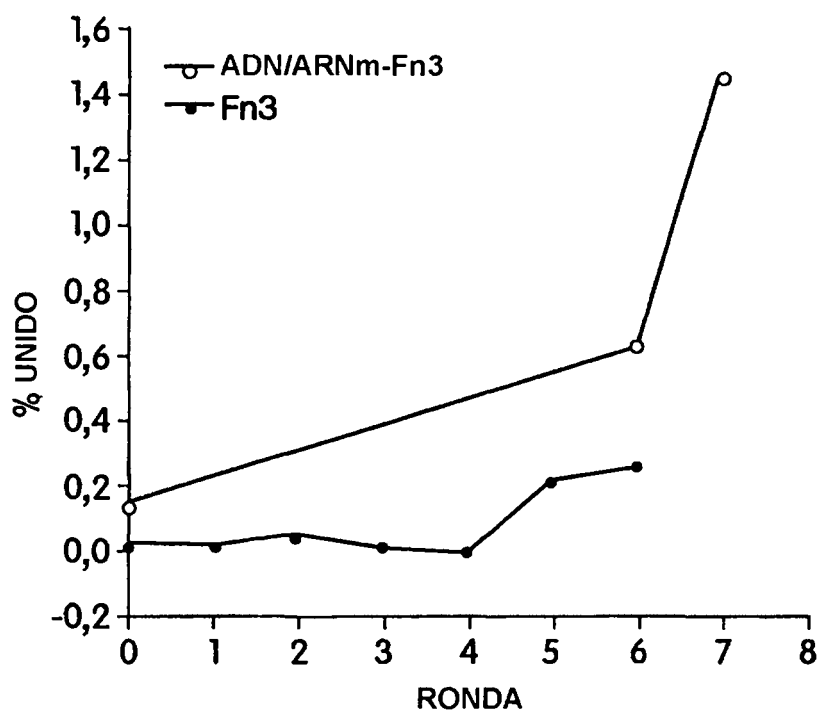


Fig. 10

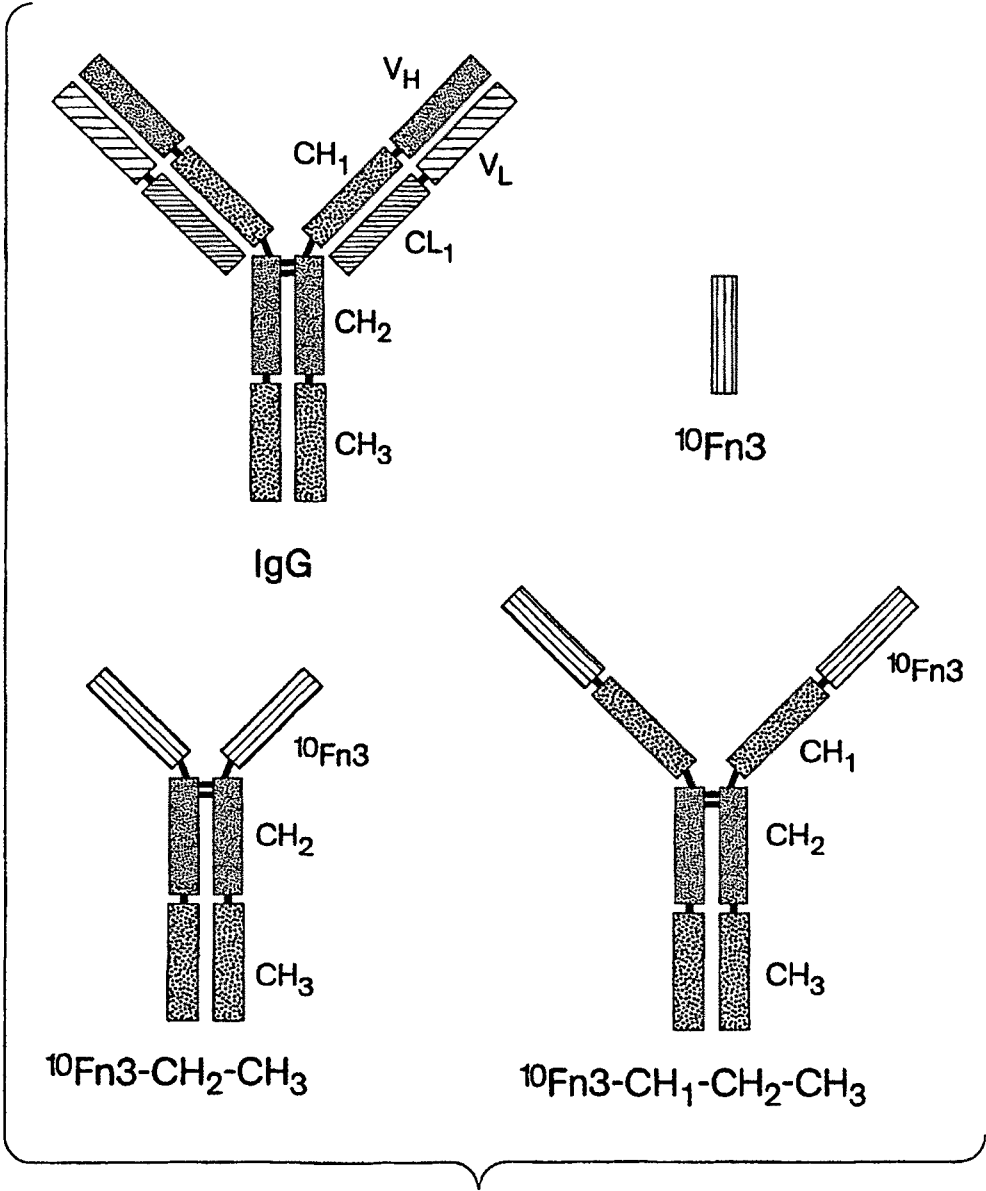


Fig. 11

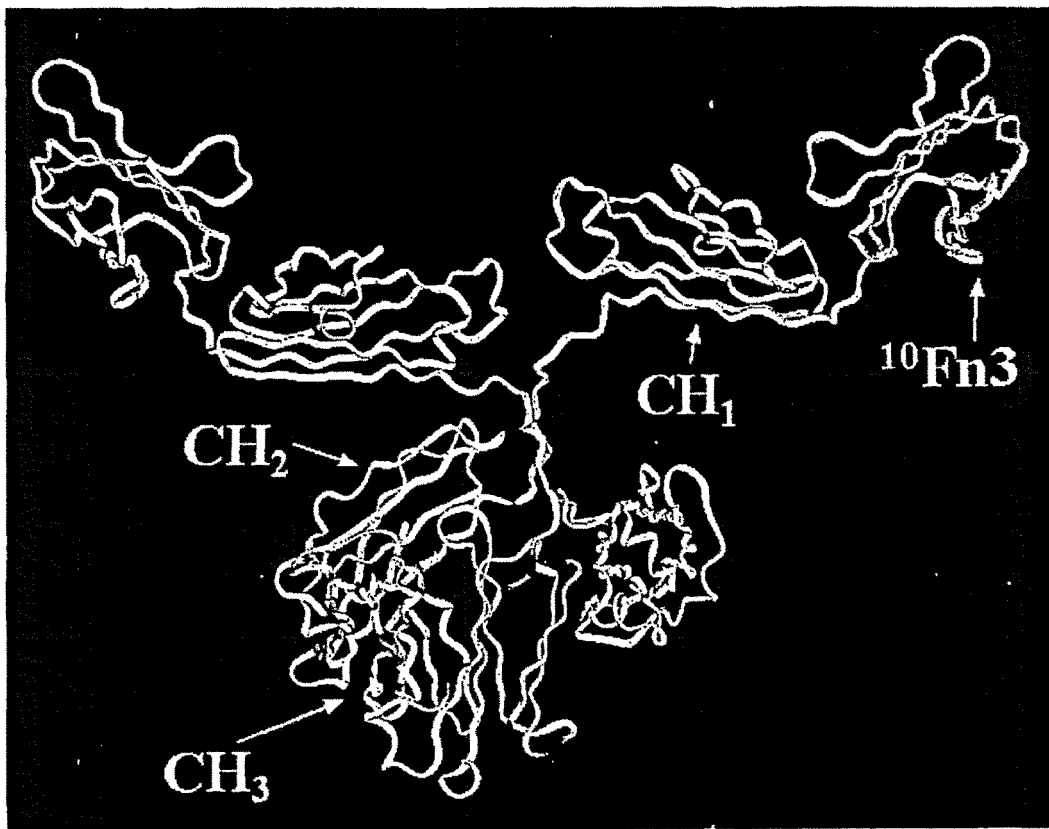


Fig. 12

FIG. 13

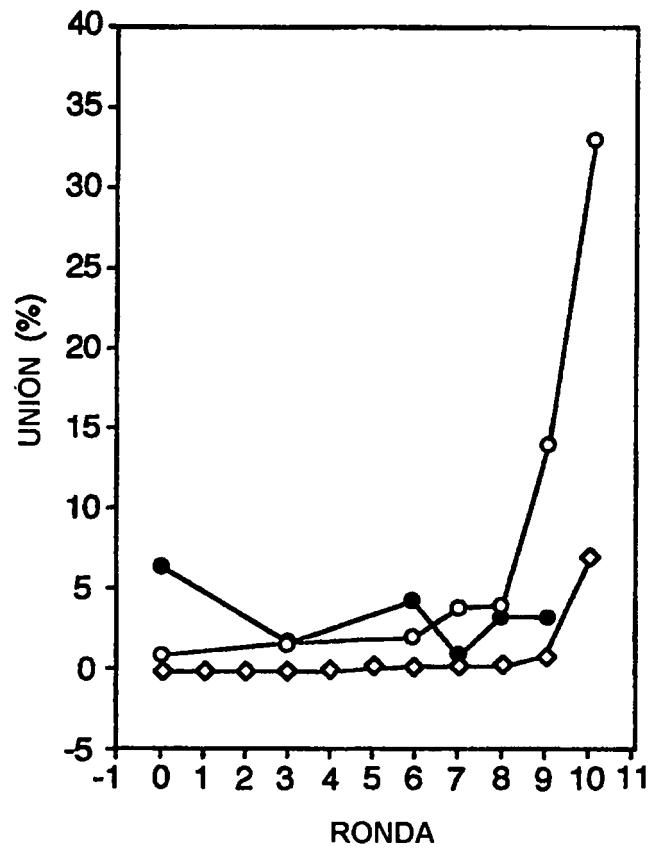


FIG. 14

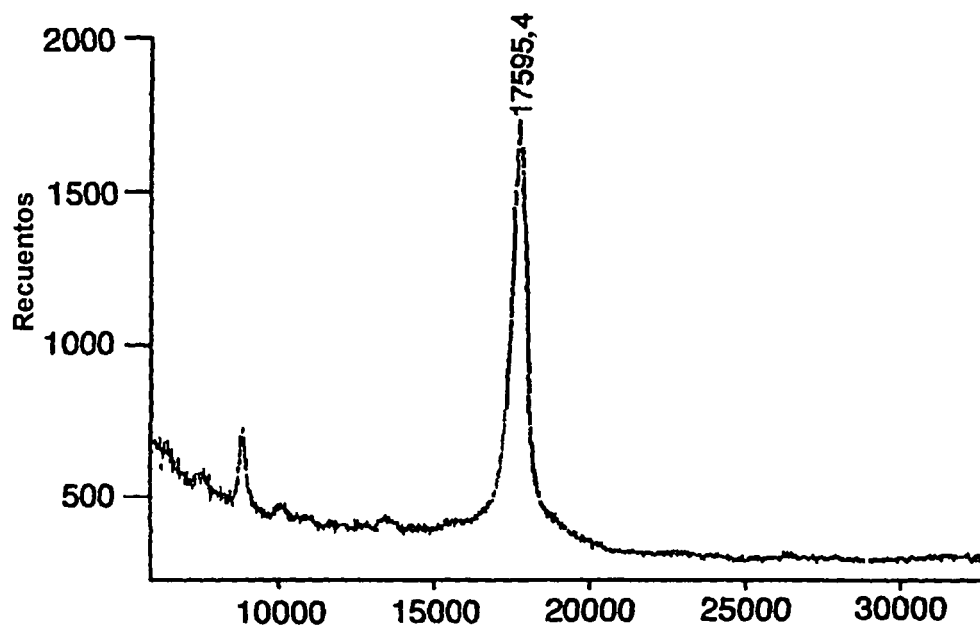


FIG. 15

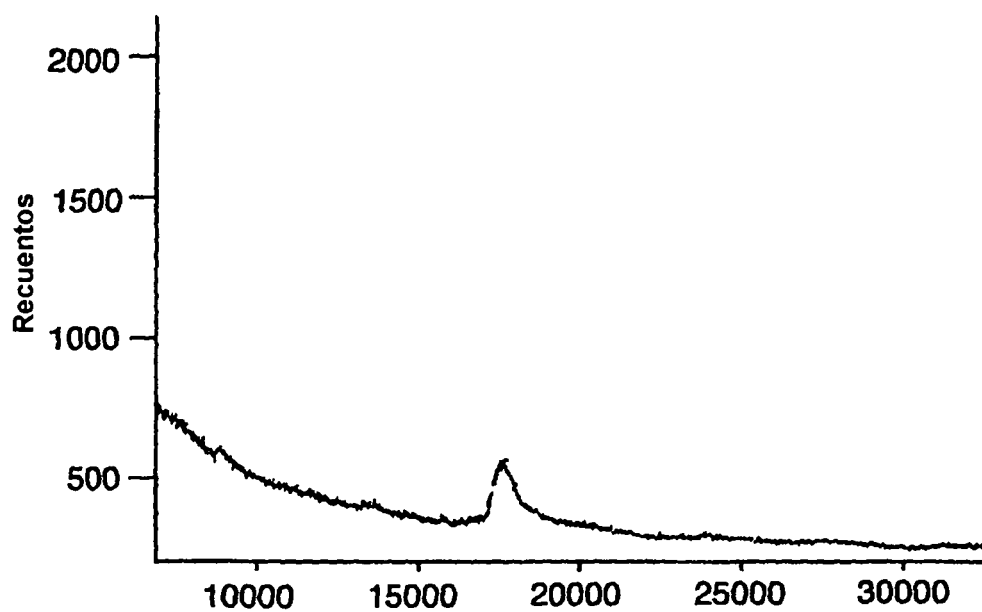


FIG. 16

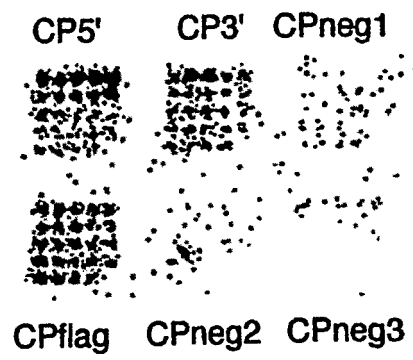
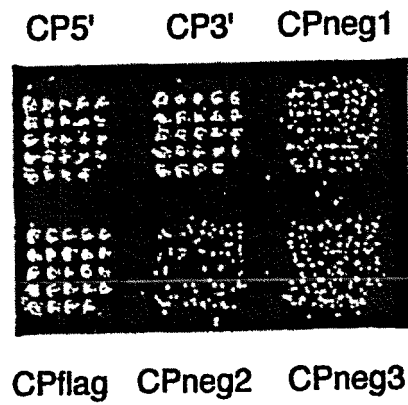


FIG. 17



ES 2 276 770 T3

LISTA DE SECUENCIAS

<110> Phylos, Inc.

5 <120> ANDAMIAJES PROTEÍNICOS PARA MIMÉTICOS DE ANTICUERPOS Y OTRAS PROTEÍNAS DE UNIÓN

10 <130> 171-125

<140> EP 01913159.8

<141> 2002-09-26

15 <150> 09/515,260

<151> 2000-02-29

20 <160> 27

<170> FastSEQ para Windows Versión 4.0

<210> 1

25 <211> 122

<212> ADN

<213> *Homo sapiens*

30

<400> 1

35 ggaattccta atacgactca ctatagggac aattactatt tacaattaca atgcatcacc 60
atcaccatca cgtttctgat gttccgaggg acctggaagt tgttgctgcg acccccacca 120
gc 122

<210> 2

40 <211> 104

<212> ADN

<213> *Homo sapiens*

45 <400> 2

50 ggaattccta atacgactca ctatagggac aattactatt tacaattaca atggtttctg 60
atgttccgag ggacctgga gttgttgctg cgacccccac cagc 104

<210> 3

55 <211> 126

<212> ADN

<213> *Homo sapiens*

60 <220>

<221> misc_feature

<222> (1)...(126)

<223> n = A, T, C o G

65

<221> misc_feature

<222> (1)...(126)

ES 2 276 770 T3

<223> s = C o G

<400> 3

5

```
agcggatgcc ttgtcgtcgt cgtccttgta gtcgctcttc cctgtttctc cgtaagtgat 60
cctgtaatat ctsnnsnnsn nsnnsnnsnn snnccagctg atcagtaggc tggtaggggt 120
cgcagc 126
```

10

<210> 4

<211> 62

<212> ADN

15

<213> *Homo sapiens*

<400> 4

20

```
agcggatgcc ttgtcgtcgt cgtccttgta gtcgctcttc cctgtttctc cgtaagtgat 60
cc 62
```

25

<210> 5

<211> 99

<212> ADN

<213> *Homo sapiens*

30

<400> 5

35

```
ggaattccta atacgactca ctatagggac aattactatt tacaattaca atgcatcacc 60
atccatca cctcttcaca ggaggaaata gccctgtcc 99
```

40

<210> 6

<211> 132

<212> ADN

<213> *Homo sapiens*

45

<220>

<221> misc_feature

<222> (1)..(132)

<223> n = A,T,C o G

50

<221> misc_feature

<222> (1)..(132)

<223> s = C o G

55

<400> 6

60

```
agcggatgcc ttgtcgtcgt cgtccttgta gtcgctcttc gtataatcaa ctccagggtt 60
aaggccgctg atggtagctg tsnnsnnsnn snnaggcaca gtgaactcct ggacagggtt 120
atttctcct gt 132
```

65

<210> 7

<211> 64

<212> ADN

ES 2 276 770 T3

<213> *Homo sapiens*

<400> 7

5

```
agcggatgcc ttgtcgtcgt cgtccttgta gtcgctcttc gtataatcaa ctccaggttt 60
aagg                                             64
```

10

<210> 8

<211> 101

<212> ADN

15

<213> *Homo sapiens*

<400> 8

20

```
ggaattccta atacgactca ctataggac aattactatt tacaattaca atgcatcacc 60
atcacatca cctcttctat accatcactg tgtatgctgt c                               101
```

25

<210> 9

<211> 114

<212> ADN

30

<213> *Homo sapiens*

<220>

<221> misc_feature

35

<222> (1)...(114)

<223> n = A, T, C o G

<221> misc_feature

40

<222> (1)..(114)

<223> s = C o G

<400> 9

45

```
agcggatgcc ttgtcgtcgt cgtccttgta gtctgttcgg taattaatgg aaattggsnn 60
snnnsnnsnns nnsnnsnnsn nsnnsnngt gacagcatac acagtgatgg tata           114
```

50

<210> 10

<211> 57

<212> ADN

55

<213> *Homo sapiens*

<400> 10

60

```
agcggatgcc ttgtcgtcgt cgtccttgta gtctgttcgg taattaatgg aaattgg
```

57

<210> 11

65

<211> 45

<212> ADN

<213> fago T7 y virus del mosaico del tabaco

ES 2 276 770 T3

| | | |
|----|---|----|
| | <400> 11 | |
| | gcgtaatacgc actcactata gggacaatta ctattacaa ttaca | 45 |
| 5 | <210> 12 | |
| | <211> 33 | |
| | <212> ADN | |
| 10 | <213> Secuencia artificial | |
| | <220> | |
| | <223> secuencia de flag | |
| 15 | <400> 12 | |
| | agcggatgcc ttgtcgtcgt cgtccttgta gtc | 33 |
| 20 | <210> 13 | |
| | <211> 19 | |
| | <212> ADN | |
| 25 | <213> Secuencia artificial | |
| | <220> | |
| | <223> oligonucleótido de tablilla | |
| 30 | <221> misc_feature | |
| | <222> (1)...(19) | |
| | <223> n = A, T, C o G | |
| 35 | <400> 13 | |
| | tttttttn agcggatgc | 19 |
| 40 | <210> 14 | |
| | <211> 20 | |
| | <212> ADN | |
| 45 | <213> Secuencia artificial | |
| | <220> | |
| | <223> oligonucleótido ligador de puromicina | |
| 50 | <400> 14 | |
| | aaaaaaaaa aaaaaaaccc | 20 |
| 55 | <210> 15 | |
| | <211> 30 | |
| | <212> ADN | |
| 60 | <213> <i>Mus musculus</i> | |
| | <400> 15 | |
| 65 | gcggcagggt ttgcttactg gggccaagg | 30 |
| | <210> 16 | |

ES 2 276 770 T3

<211> 27
<212> ADN
<213> *Mus musculus*
5 <400> 16
gggaggggtg gaggtaggtc acagtcc 27
10 <210> 17
<211> 30
<212> ADN
15 <213> *Mus musculus*
<400> 17
20 tttgctagct ttaccaggag agtgggaggc 30
<210> 18
<211> 33
25 <212> ADN
<213> *Mus musculus*
<400> 18
30 aaaaagcttg ccaaaacgac acccecatct gtc 33
<210> 19
35 <211> 33
<212> ADN
<213> *Mus musculus*
40 <400> 19
catatggttt ctgatattcc gagagatctg gag 33
45 <210> 20
<211> 43
<212> ADN
50 <213> *Mus musculus*
<400> 20
55 catatgcatc accatcacca tcacgtttct gatattccga gag 43
<210> 21
<211> 30
60 <212> ADN
<213> *Mus musculus*
<400> 21
65 gaattcctat gttttataat tgatggaaac 30
<210> 22

ES 2 276 770 T3

| | | | |
|----|---------------------------------|--|----|
| | <i><211> 19</i> | | |
| | <i><212> ADN</i> | | |
| | <i><213> Mus musculus</i> | | |
| 5 | <i><400> 22</i> | | |
| | tgtaaatagt aattgtccc | | 19 |
| 10 | <i><210> 23</i> | | |
| | <i><211> 20</i> | | |
| | <i><212> ADN</i> | | |
| 15 | <i><213> Mus musculus</i> | | |
| | <i><400> 23</i> | | |
| 20 | ttttttttt ttttttttt | | 20 |
| | <i><210> 24</i> | | |
| | <i><211> 15</i> | | |
| 25 | <i><212> ADN</i> | | |
| | <i><213> Mus musculus</i> | | |
| | <i><400> 24</i> | | |
| 30 | cctgtaegtg tccat | | 15 |
| | <i><210> 25</i> | | |
| 35 | <i><211> 16</i> | | |
| | <i><212> ADN</i> | | |
| | <i><213> Mus musculus</i> | | |
| 40 | <i><400> 25</i> | | |
| | catcgtcctt gtagtc | | 16 |
| 45 | <i><210> 26</i> | | |
| | <i><211> 13</i> | | |
| | <i><212> ADN</i> | | |
| 50 | <i><213> Mus musculus</i> | | |
| | <i><400> 26</i> | | |
| 55 | cgtcgtaggg gta | | 13 |
| | <i><210> 27</i> | | |
| | <i><211> 17</i> | | |
| 60 | <i><212> ADN</i> | | |
| | <i><213> Mus musculus</i> | | |
| | <i><400> 27</i> | | |
| 65 | caggtcttct tcagaga | | 17 |