



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 313 886**

51 Int. Cl.:

C07K 16/10 (2006.01) **C07K 16/42** (2006.01)
C07K 14/705 (2006.01) **C07K 14/715** (2006.01)
C07K 19/00 (2006.01) **C12N 15/62** (2006.01)
C12N 5/10 (2006.01) **A61K 39/42** (2006.01)
A61K 38/17 (2006.01) **A61P 31/18** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **00918015 .9**

96 Fecha de presentación : **16.03.2000**

97 Número de publicación de la solicitud: **1161455**

97 Fecha de publicación de la solicitud: **12.12.2001**

54

Título: **Nueva proteína quimérica para la prevención y tratamiento de infección por HIV.**

30

Prioridad: **16.03.1999 US 124681 P**

45

Fecha de publicación de la mención BOPI:
16.03.2009

45

Fecha de la publicación del folleto de la patente:
16.03.2009

73

Titular/es: **THE GOVERNMENT OF THE UNITED STATES AMERICA, as represented by THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES**
The National Institute of Health
Office of Technology Transfer
6011 Executive Boulevard, Suite 325
Rockville, Maryland 20852, US

72

Inventor/es: **Berger, Edward, A. y**
Del Castillo, Christie, M.

74

Agente: **Carpintero López, Mario**

ES 2 313 886 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

Nueva proteína quimérica para la prevención y tratamiento de infección por HIV.

5 **Campo de la invención**

La presente invención hace referencia a las proteínas útiles para la prevención y tratamiento de la infección por el virus de la inmunodeficiencia humana (VIH). Más especialmente, hace referencia a las proteínas de fusión que se unen a dos sitios de una única proteína diana, particularmente cuando un dominio de unión de la proteína de fusión se une a un sitio inducido (de la proteína diana) que está expuesto por la unión del otro dominio de unión de la proteína de fusión.

Antecedentes de la invención

15 El síndrome de la inmunodeficiencia adquirida (SIDA) es una enfermedad letal con una presencia cada vez mayor en el mundo moderno. El agente responsable de esta enfermedad, el virus de la inmunodeficiencia humana (VIH), fue identificado por primera vez en 1983. El VIH es un retrovirus linfotrópico de células T que invade y se replica en las células del sistema inmunitario, inicialmente en los linfocitos T auxiliares. La consecuente disfunción en la inmunidad de los linfocitos T da como resultado una situación inmunodeprimida. Los pacientes suelen fallecer por infecciones oportunistas asociadas producidas por virus, bacterias y hongos. Uno de los descubrimientos de laboratorio característicos sobre el SIDA es el descenso de los linfocitos T auxiliares (CD4), y en especial, se observa un descenso continuo del ratio entre los CD4 y los linfocitos T supresores CD8 a medida que avanza la enfermedad. La unión viral se efectúa inicialmente mediante la interacción de gp120, la subunidad externa de la glicoproteína de superficie del VIH (Env) con la proteína CD4 y varias moléculas correceptoras (uno de los varios receptores de quimiocina alternativos). Estas interacciones activan a continuación la subunidad de transmembrana gp41 de la glicoproteína de superficie para así provocar la fusión entre el virus y las membranas celulares. Ver *Retroviruses*, Coffin *et al.* (eds.) (1997) CSHP, New York, Cap. 11.

La infección por VIH provoca al sistema inmune humoral, aunque este, por lo general, no proporciona la protección suficiente para prevenir la infección. Env es la diana principal de los anticuerpos neutralizadores anti VIH (Wyatt *et al. Nature* 393:705-711, 1998). No obstante, Env ha evolucionado de tal manera que sus determinantes neutralizadores relativamente invariantes están protegidos frente al sistema inmune humoral. Los anticuerpos para esas regiones se generan por tanto a una frecuencia baja, y sus actividades neutralizadoras *in vivo* son normalmente débiles. Algunas regiones variables (por ejemplo, el bucle V3) son dianas para los anticuerpos neutralizantes potentes, pero estas suelen estar restringidas a un número limitado de cepas del VIH (en otras palabras, no tienen una reacción cruzada amplia). Se puede acceder a una lista de varios epítomos antigénicos de gp120 y a unas definiciones consensuadas de las regiones constantes y variables de gp120 en la solicitud PCT publicada PCT/US98/02766 (número de publicación WO 98/36087) y en Coffin *et al.* (eds.) (1997) CSHP, New York, Cap. 12.

Un anticuerpo monoclonal neutralizante (MAb) con una actividad de reactividad cruzada potente y amplia tendría un valor potencial alto en los protocolos destinados a la prevención de la infección por VIH antes o inmediatamente después de la exposición, por ejemplo en la transmisión neonatal, en la profilaxis después de la exposición, y como un inhibidor tóxico. Dicho MAb podría ser útil asimismo para tratar la infección crónica (D'Souza *et al. J Infect. Dis.* 175:1056-1062, 1997). Sin embargo, tan sólo se han descrito un puñado de MAb con las actividades neutralizantes de reactividad cruzada amplia deseadas. Dada la limitación de la potencia y la reactividad cruzada de estas moléculas, incluso las tres candidatas más prometedoras tienen un valor clínico cuestionable (D'Souza *et al.*, 1997).

Se están llevando a cabo unas investigaciones más extensas para aportar unas visiones inmunológicas o farmacológicas que permitan controlar la infección por VIH (Coffin *et al.*, 1997, Cap. 12). La interacción específica entre gp120 y CD4 ya se ha explotado en diversas investigaciones para encontrar un posible tratamiento para la infección por VIH. Ver, por ejemplo, la patente estadounidense n.º 5.817.767; Capon *et al.*, *Nature* 337:525-531, 1989. En ella se generó un fragmento soluble de CD4 (sCD4) que comprende el primer y segundo dominios de esta proteína (D1D2), y esta molécula interactúa especialmente con gp120, actuando en esencia como un señuelo molecular. Se ha demostrado que sCD4 bloquea la expansión del VIH entre las células cultivadas (Moore *et al.*, *Science* 250: 1139-1142, 1990). Sin embargo, los ensayos clínicos con sCD4 resultaron no concluyentes a los efectos sobre la carga viral humana (Schooley *et al.*, *Ann. Internal Med.* 112:247-253, 1990; Kahn *et al.*, *Ann. Internal Med.* 112:254-261, 1990). Los estudios siguientes indicaron que, al contrario que en las cepas del VIH adaptadas en el laboratorio, las cepas aisladas obtenidas directamente de pacientes infectados (cepas aisladas primarias) eran resistentes a la neutralización por sCD4 (Darr *et al.*, *Proc. Natl. Acad. Sci.* 87:6574-6578, 1990).

En otro estudio, los investigadores generaron una molécula semejante a un anticuerpo mediante la fusión de la porción de unión de CD4 a la región constante (Fc) de una cadena pesada de IgG humana (ver por ejemplo, Capon *et al.*, *Nature* 337:525-531, 1989; y Byrn *et al.*, *Nature* 344: 667-670, 1990). Esta molécula, denominada inmunoadhesina CD4, explota las funciones innatas de los Fc de la inmunoglobulina, como su capacidad para fijar el complemento, su capacidad para mediar en la citotoxicidad dependiente de los anticuerpos y su transferencia a través de la barrera placentaria. Existen unos inconvenientes significativos frente al uso de los receptores Fc asociados con CD4 debido a que dicho constructo puede desencadenar que el VIH ataque a las células con receptor Fc (como los macrófagos), y esto puede provocar un aumento de la transmisión del VIH-1 a través de la barrera placentaria.

También se ha creado una molécula recombinante complementaria, cuya porción de unión de CD4 está fusionada con la región Fv de un anticuerpo dirigido al CD3 humano; esta molécula "Janusina" puede hacer que los linfocitos T citotóxicos se vuelvan a dirigir contra las células infectadas por el VIH (Trauneker *et al.*, *Embo J.* 10:3655-3659, 1991; Trauneker *et al.*, *Int. J. Cancer: Supp.* 7:51-52, 1992). La janusina ha demostrado inhibir la fusión celular con el VIH cuando se administra *in vitro* con un anticuerpo neutralizante bien para gp41, bien para el bucle V3 de gp120 (Allaway *et al.*, *AIDS Res. Hum. Retroviruses* 9:581-587, 1993; patente estadounidense n.º 5.817.767). Este sistema es inherentemente complicado e ineficaz, dado que hay que coadministrar múltiples moléculas al sujeto.

Esta invención está dirigida a las proteínas que muestran fallos básicos en las técnicas precedentes.

WO-A-9747318 hace referencia a los usos de un receptor de quimiocina para inhibir la infección por VIH-1. En *Science*, volumen 280, 1998, páginas 1833 a 1834 se hace referencia a la determinación de la estructura atómica de la proteína gp120. En *Science*, Vol. 280, 1998, páginas 1949 a 1953 se hace referencia a la estructura de glicoproteína gp120 del VIH y a su implicación en la unión con el receptor de la quimiocina. Idziorek and Klatzmann, *AIDS Res. And Human Retroviruses*, Vol. 7, 1991, páginas 529-536, hace referencia a la construcción de moléculas quiméricas basadas en CD4 por entrecruzamiento químico.

Resumen de la invención

La presente invención como se define en las reivindicaciones saca partido del descubrimiento de que las actividades neutralizantes de los MAb frente a ciertos determinantes altamente constantes de la región de unión con el correceptor de gp120 únicamente se revelan cuando CD4 se une por primera vez a gp120 (como en un ensayo de fusión de sCD4 activado). Aunque algunos MAb para epítomos inducidos por CD4 (por ejemplo, los MAb humanos 17b y 48d Thali *et al.*, *J. Virol.* 67:3978-3988, 1993) tienen una fuerte reacción cruzada con los Env de subtipos genéticos de diversos VIH (Clades), y estos epítomos neutralizantes únicamente se exponen brevemente *in vivo*, y por tanto tienen unas dianas escasas para la unión con anticuerpos clínicamente protectores.

Los inventores han superado estas dificultades creando una proteína de fusión que contiene un fragmento de CD4 unido mediante un ligando a una cadena humana única de Fv dirigida contra un epítomo neutralizante inducido (por ejemplo, un epítomo neutralizante inducido por CD4) sobre gp120, por ejemplo, un determinante de unión a un correceptor de gp120. La unión a CD4 expone a los determinantes de gp120 altamente constantes implicados en la unión al correceptor; así, la proteína de fusión proporcionada tendrá las propiedades de un anticuerpo neutralizante con amplia reactividad cruzada y muy potente, con una alta actividad *in vivo* y sin proteínas de diana no deseables mediante Fc. Cuando la proteína de fusión deriva sustancialmente de proteínas humanas tiene una inmunogenicidad y toxicidad mínimas en los humanos. Dicho agente tiene un gran valor en la prevención de la infección durante o inmediatamente después de la exposición al VIH (transmisión madre/hijo, profilaxis tras la exposición, inhibidor tópico), y también en el tratamiento de la infección crónica.

Por consiguiente, un primer modelo de la presente invención es una proteína de fusión biespecífica neutralizante capaz de unirse a dos sitios de una proteína diana. Esta proteína tiene dos dominios de unión diferentes: un dominio de CD4 de inducción a la unión y un dominio de unión inducida, ligados funcionalmente por un ligando péptido. Las moléculas de ácido nucleico que codifican dichas proteínas de fusión son unos aspectos más de esta invención. También incluidos en la presentación se encuentran los análogos, derivados o miméticos proteicos de dichas proteínas de fusión biespecíficas neutralizantes. La disposición de los dominios de inducción a la unión y de unión inducida no necesitan organizarse en una secuencia de unión; el dominio de unión aminoproximal o carboxiproximal de la proteína de fusión puede ser el dominio de unión inducida o el de inducción a la unión.

El ligando de esta invención es de tal longitud y tiene una estructura secundaria tal que el ligando permite que el segundo dominio de unión esté en proximidad de unión con el epítomo inducido de la proteína diana cuando el primer dominio de unión está vinculado al sitio de unión de la proteína diana. El ligando, por ejemplo, puede ser bastante flexible. Los ligandos de entre 25-100 angstroms (Å) o de unos 15-100 residuos de aminoácidos de longitud son ejemplos de ligandos de una longitud suficiente para mantener un segundo dominio de unión en proximidad de unión al epítomo inducido. En SEC. ID N.º 1 se representan unos ejemplos específicos de ligandos que incluirán una o más repeticiones de una secuencia de aminoácidos. Por ejemplo, la invención abarca las proteínas de fusión biespecíficas cuando los dos dominios de unión están funcionalmente unidos a la secuencia de aminoácidos representada en SEC. ID N.º 2.

Las dianas para las proteínas de fusión biespecíficas de conformidad con esta invención incluyen las proteínas de la superficie viral. Por ejemplo, las proteínas de la superficie viral del virus de la inmunodeficiencia humana (VIH) son dianas para la invención presentada. En un modelo específico de la invención, la diana de la proteína de la superficie viral es gp120.

En otros aspectos de la invención, el primer dominio de unión es capaz de unirse a un sitio de inducción de la proteína diana, exponiendo así un epítomo inducido. El primer dominio de unión es CD4 o un fragmento del mismo. De forma alternativa, dicho primer dominio de unión se presenta como una porción de unión de una región variable de una cadena de anticuerpos pesada o ligera. El primer dominio de unión puede incluir, por ejemplo, un dominio de unión a un anticuerpo, un Fv de cadena sencilla (SCFv) o fragmentos de unión del mismo.

ES 2 313 886 T3

El segundo dominio de unión, que es capaz de formar un complejo neutralizante con un epítipo inducido de la proteína diana, puede ser, por ejemplo, un anticuerpo o fragmentos del mismo, como la región variable Fv, Fab o dominio de unión al antígeno de un anticuerpo. Otro ejemplo del segundo dominio de unión de la proteína de fusión es un Fv de cadena sencilla modificada genéticamente (SCFv).

En algunos ejemplos en particular en los que la gp120 del VIH es la diana, y el sitio de inducción es el sitio de unión de CD4 a gp120, el epítipo inducido puede ser un determinante de unión a un correceptor de gp120. Por consiguiente, los aspectos de dicha invención incluyen proteínas en las que el primer dominio de unión se une a gp120 de tal modo que ocasiona un cambio conformacional en la gp120 en complejo que expone el segundo dominio de unión. El primer dominio de unión puede derivar de una molécula CD4, e incluye CD4 y fragmentos solubles del mismo (sCD4, es decir, D1, D1D2 y otros fragmentos similares), y proteínas que copian la actividad biológica de una molécula CD4 en unión con el sitio de inducción de gp120. En otro modelo de la invención, el primer dominio de la proteína de fusión biespecífica dirigida contra gp120 deriva de un anticuerpo anti idiotípico CD4, o anticuerpos que copian a CD4 en los epítopos de exposición.

El segundo dominio de la proteína de fusión biespecífica dirigida contra gp120, que se une a un epítipo inducido por unión del primer dominio de fusión, puede elegirse entre dominios y fragmentos de proteínas que se unen a dichos epítopos inducidos de CD4. Los anticuerpos dirigidos a los epítopos inducidos, al igual que el correceptor del VIH (es decir, un receptor de quimiocina), las copias del correceptor del VIH, y los fragmentos de las proteínas del correceptor del VIH son ejemplos de fuentes para el segundo dominio de unión de una proteína de fusión biespecífica de gp120 diana de esta invención. Los ejemplos de receptores de quimiocina con actividad de correceptor del VIH incluyen a CXCR4, CCR5, CCR2B y CCR3. Los anticuerpos neutralizantes, entre los que se incluyen 17b y 48d, son ejemplos de anticuerpos. Las proteínas de fusión en las que el segundo dominio es un Fv de cadena sencilla modificado genéticamente (SCFv) derivada de dicho anticuerpo neutralizante también se incluyen en esta invención.

Un modelo en especial de esta invención es una proteína de fusión biespecífica recombinante funcional capaz de unirse a dos sitios de gp120, donde el dominio de inducción a la unión es sCD4; el dominio de unión inducida es SCFv (17b); y estos dos dominios están unidos mediante un ligando de una longitud suficiente para mantener el SCFv (17b) en proximidad de unión con un epítipo SCFv (17b) cuando sCD4 está unido a gp120. Una proteína biespecífica prototípica tiene la secuencia de aminoácido que figura en SEC. ID N.º 3. Las moléculas de ácido nucleico que codifican dicha proteína de fusión también están incluidas en esta invención; la molécula de ácido nucleico prototípica tiene la secuencia que figura en SEC. ID N.º 4. Los vectores y células que conforman esta molécula de ácido nucleico también están incluidas en la presente invención, como lo son los animales y plantas transgénicos que expresan la molécula de ácido nucleico.

Esta invención también ofrece métodos para la producción de proteínas de fusión biespecíficas recombinantes funcionales capaces de unirse a dos sitios de una proteína diana. Dicha proteína puede producirse en una célula procariota o eucariota (es decir, células de levadura, insectos y mamíferos), por ejemplo mediante transformación o transfección de dicha célula con una molécula de ácido nucleico recombinante formada por una secuencia que codifique una proteína de fusión biespecífica presentada. Dichas células transformadas pueden ser cultivadas en condiciones tales que provoquen la producción de la proteína de fusión, que así se recupera a través de métodos de purificación de proteínas. La proteína puede incluir una etiqueta molecular como una etiqueta de seis-histidina para facilitar su recuperación. En unos modelos en particular, la proteína tiene una etiqueta de hexa-histidina (hexa-his) y un sitio de corte de la trombina. La célula eucariota de este aspecto de la invención no es una célula embrionaria humana.

La invención también presenta métodos *in vitro* para la inactivación de una proteína diana, por ejemplo, una proteína gp120, al poner en contacto a la proteína diana con una proteína de fusión conforme a la presente invención. Cuando la proteína diana es una gp120, este método implica el contacto de gp120 con una proteína de fusión biespecífica dirigida contra gp120, por ejemplo, sCD4-SCFv(17b). Las proteínas, conforme a la presente invención, también pueden utilizarse para neutralizar un virus de inmunodeficiencia humana poniendo en contacto el virus de inmunodeficiencia humana con una proteína de fusión dirigida contra gp120, en virtud de esta invención. La unión de una proteína viral o recombinante gp120 con CD4 soluble o linfocito CD4 también puede bloquearse y/o evitarse mediante la puesta en contacto de la proteína gp120 con la proteína de fusión dirigida contra gp120. En cualesquiera de dichos métodos, también puede utilizarse una proteína variante, análoga o mimética de la proteína de fusión tal y como se ha mencionado en este documento.

Las proteínas de la presente invención pueden utilizarse para inhibir la replicación del virus o la infectividad en un sujeto mediante administración al sujeto de una cantidad de proteína de fusión (por ejemplo la proteína de fusión sCD4-SCFv(17b)), o una proteína variante, análoga o mimética de la misma, suficiente para inhibir la infectividad o replicación del virus VIH. La proteína de fusión puede administrarse como composición farmacéutica y suministrarse de forma terapéutica a una persona infectada con el VIH, o de forma profiláctica para ayudar a evitar la infección en una persona que haya estado expuesta al virus o que esté en un alto riesgo de exposición. Las proteínas de esta invención también pueden administrarse en combinación con otro compuesto para el tratamiento o prevención de una infección por VIH, como una transcriptasa reversa (TR) del VIH, un inhibidor de la integrasa o proteasa, otro anticuerpo neutralizante del VIH1 o una toxina dirigida contra Env. El otro medicamento puede ser un agente antiviral del VIH, un agente anti infectivo del VIH y/o un inmunomodulador o combinaciones de lo anterior.

Breve descripción de las figuras

La Figura 1 es un gráfico de barras que ilustra la fusión relativa del VIH-1 mediante Env en presencia (+) o ausencia (-) de CD4 soluble entre células efectoras que expresan Env (Ba-L) y células diana que expresan CCR5 (correceptor), pero no CD4 (receptor primario).

La Figura 2 es un gráfico que muestra que el anticuerpo 17b no inhibe la fusión del VIH-1 mediante Env en el ensayo convencional (cuadrados blancos: CXCR4 y CD4 sobre la célula diana) pero inhibe con fuerza la fusión celular en el ensayo activado con sCD4 (círculos negros: únicamente CXCR4 sobre la célula diana, teniendo en cuenta sCD4). Los experimentos adicionales indican que este fenómeno ocurre con diversos Env que usan CXCR4 o CCR5, y que 17b tiene una amplia actividad de reacción cruzada con los Env aislados del VIH-1 genéticamente diversos.

La Figura 3 es un diagrama esquemático del constructo genético CD4-SCFv (17b). El constructo genético codifica sCD4 (D1D2 más la secuencia señal del CD4 N-terminal nativo), seguido por el ligando L1 (Gly₄Ser)₇, que vincula el 17b CSFv (V_H vinculado a V_L mediante el ligando L2 (Gly₄Ser)₃), seguido por el sitio de corte de la trombina y la etiqueta hexa-his. Existe un sitio BamH I en medio del L1 para activar la producción de constructos de distintas longitudes.

La Figura 4 es un diseño de los mecanismos de unión de sCD4-SCFv(17b) a gp120, y la neutralización resultante de la función Env del VIH (fusión e infección).

Las Figs. 4A, 4B y 4C representan la interacción propuesta del VIH (mediante gp120) con el receptor de la superficie celular CD4 y el correceptor CCR5, y el inicio de la fusión (mediante gp41). La interacción entre gp120 y CD4 (Fig. 4A) provoca un cambio en la forma de gp120 (Fig. 4B), lo que permite la interacción entre gp120 y CCR5 (Fig. 4B). Esto provoca un cambio de conformación en gp41 (Fig. 4C) y desemboca en la fusión. El anticuerpo (por ejemplo Mab 17b) se une levemente al epítipo expuesto temporalmente con gp120 (Fig. 4B), y así se produce tan sólo una neutralización leve de la fusión o la infección.

Las Figs. 4D y 4E ilustran un mecanismo propuesto de neutralización o fusión de sCD4-SCFv(17b). En presencia de la proteína de fusión quimérica biespecífica, el dominio de sCD4 puede unirse a gp120 e inducir un cambio de conformación en esta proteína suficiente para permitir la unión de SCFv(17b) (Fig. 4D). Esto bloquea de forma eficaz la fusión entre el VIH y la infección en la célula diana.

Lista de secuencias

El ácido nucleico y las secuencias de aminoácidos que aparecen en la lista de secuencias siguiente se muestran utilizando las abreviaturas estándar de una letra para bases de nucleótidos, y el código de tres letras para los aminoácidos. Únicamente se muestra una cadena de cada secuencia de ácido nucleico, pero la cadena complementaria se supone incluida por toda referencia en la cepa mostrada.

SEC. ID. N.º 1: muestra la repetición básica para un polipéptido ligando.

SEC. ID. N.º 2: muestra un ligando polipéptido de siete repeticiones.

SEC. ID. N.º 3: muestra la secuencia de aminoácidos de la proteína quimérica CD4-SCFv(17b).

SEC. ID. N.º 4: muestra la secuencia de ácido nucleico de CD4-SCFv(17b).

SEC. ID. N.º 5: muestra el par de oligonucleótidos sintéticos utilizados para formar la segunda mitad del sitio Stu I cerca del extremo 3' de CD4 y para producir un saliente de Spe I en el extremo 3' de un constructo intermedio (un sitio que debe ser destruido tras la unión en pCB-3); las secuencias de oligonucleótidos reconstruyen el resto del segundo dominio de CD4 (a través de ser₁₈₃), y codifica una secuencia de aminoácidos que incluyen ala₁₈₂ser₁₈₃ de CD4 D2 más un ligando intermedio de 37 residuos (gly₄ser)₆gly₄thr₂ser, seguido directamente por la secuencia de terminación translacional universal (UTS).

SEC. ID. N.º 6: muestra la secuencia de péptidos codificada por las secuencias de nucleótidos de SEC. ID N.º 5.

SEC. ID. N.º 7: muestra el iniciador (5') anverso utilizado para amplificar y adjuntar la secuencia 17b SCFv a la secuencia de ligando CD4 en pCD2. La cursiva muestra la región del iniciador que se solapa con 17b.

SEC. ID. N.º 8 muestra la secuencia de aminoácidos codificada por el iniciador de oligonucleótidos de SEC. ID N.º 7. Esta secuencia incluye los residuos GlySer en la tercera repetición (Gly₄Ser) dentro de L1 (codificado por el sitio BamH I), seguido por las restantes cuatro repeticiones (Gly₄Ser), seguido por los primeros diez residuos de 17b SCFv (en cursiva).

SEC. ID. N.º 9 muestra el iniciador 3' utilizado para amplificar y adjuntar la secuencia 17b SCFv más el sitio de corte de la trombina y la etiqueta hexa-his a la secuencia del ligando CD4 en pCD2.

SEC. ID N.º 10 muestra el péptido codificado por las secuencias de nucleótidos de SEQ ID N.º 9.

Descripción detallada de la invención**I. Abreviaturas y Definiciones**5 **A. Abreviaturas**

VIH: virus de la inmunodeficiencia humana.

gp120: la subunidad externa del complejo de la glicoproteína de superficie del VIH.

10

Env: el complejo de glicoproteína de superficie del VIH.

MAB: anticuerpo monoclonal.

15

Fv: "fragmento variable" del anticuerpo, la región variable de un anticuerpo.

SCFv: anticuerpo de cadena sencilla de regiones variables.

20 **B. Definiciones**

A menos que se especifique lo contrario, los términos técnicos se utilizan de conformidad con la interpretación convencional. Las definiciones de los términos comunes en biología molecular pueden encontrarse en Benjamin Lewin, *Genes V*, Oxford University Press, 1994 (ISBN 0-19-854287-9); Kendrew *et al.* (eds.), *The Encyclopedia of Molecular Biology*, Blackwell Science Ltd., 1994 (ISBN 0-632-02182-9); y Robert A. Meyers (ed.), *Molecular Biology and Biotechnology: a Comprehensive Desk Reference*, VCH Publishers, Inc., 1995 (ISBN 1-56081-569-8).

25

Para facilitar la revisión de los distintos modelos de la invención, se ofrecen las siguientes definiciones de los términos:

30

Animal: Organismos vertebrados multicelulares vivientes, una categoría que incluye, por ejemplo, mamíferos y aves. El término mamífero incluye tanto a los mamíferos humanos como a los no humanos. De forma similar, el término "sujeto" incluye tanto a sujetos humanos como veterinarios.

Proteína de fusión biespecífica: Proteínas que tienen al menos dos dominios fusionados entre sí, cada uno con una región de unión capaz de formar un complejo específico con una proteína diana. Por lo general, los dos dominios están genéticamente fusionados; en aquellas moléculas de ácido nucleico que codifican cada dominio de la proteína están unidas funcionalmente, por ejemplo por un oligonucleótido ligando, y así producen una molécula sencilla de ácido nucleico que codifica la fusión. El producto traducido de dicha molécula de ácido nucleico que codifica la fusión es la proteína de fusión biespecífica.

40

Las dos regiones de unión de una proteína biespecífica pueden asociarse con dos determinantes o epítomos de unión diferentes en una única molécula diana. Un dominio de unión puede unirse primero a dicha diana y así inducir un cambio de conformación en la diana de forma que la unión del segundo dominio de unión a la diana esté habilitada, facilitada o de alguna manera aumentada en cuanto a su afinidad. En dicho caso, el dominio que primero se une a la diana puede denominarse como el dominio de inducción a la unión, mientras que el otro dominio que se une en segundo lugar es el dominio de unión inducida. Estos dominios de las proteínas de fusión no necesitan organizarse en secuencias de unión; el dominio de unión aminoproximal de la proteína de fusión puede ser tanto el dominio de unión inducida como el de inducción a la unión; y esto mismo se aplica al dominio de unión carboxiproximal.

50

Las proteínas de fusión biespecíficas pueden también etiquetarse de conformidad con la proteína diana a la que se unen y neutralizan. Por ejemplo, una proteína de fusión biespecífica conforme a la presente invención que se une a dos sitios específicos de la proteína gp120 del VIH puede denominarse como la proteína de fusión biespecífica dirigida contra gp120.

55

CD4: Un grupo con factor de diferenciación 4, una proteína de superficie de las células T que intermedia en la interacción con la molécula MHC de clase II. CD4 también sirve como el primer sitio receptor para el VIH en las células T durante la infección del VIH.

Las moléculas derivadas del CD4 incluyen fragmentos de CD4, generados bien por digestión química (es decir, enzimática) o mediante métodos de modificación genética. Dichos fragmentos pueden ser uno o más dominios completos de proteína CD4 (por ejemplo, dominios extracelulares D1, D2, D3 y D4), tal y como se define en la literatura sobre inmunología, o una porción de uno o más de estos dominios bien definidos. Por ejemplo, una molécula de unión o un dominio de unión derivado de CD4 contaría con una porción suficiente de la proteína CD4 que mediaría en una interacción específica y funcional entre el fragmento que se une y un sitio de unión viral o nativo de CD4. Uno de estos fragmentos de unión incluye tanto el dominio extracelular D1 como el D2 de CD4 (CD4 D1D2), aunque los fragmentos más pequeños también pueden facilitar una unión específica y funcional similar a CD4. El sitio de unión de gp120 ha sido asignado a D1 de CD4.

65

El término “moléculas derivadas de CD4” también incluye los análogos (moléculas orgánicas no proteicas), derivados (moléculas de proteína funcionalizadas químicamente obtenidas a partir de las secuencias de proteína presentadas) o miméticos (químicos similares tridimensionales) de la estructura nativa de CD4, así como variantes de secuencias de proteínas o alelos genéticos que mantienen la capacidad de unirse funcionalmente a una molécula diana.

5 *Cambio de conformación inducido por CD4:* Un cambio inducido en la conformación tridimensional de la proteína de interacción gp120 cuando CD4 interactúa específicamente con gp120 para formar un complejo. Una característica de dicho cambio es la exposición de al menos un epítipo inducido en la molécula de interacción sobre gp120. Un epítipo inducido por dicho cambio se denomina un epítipo inducido por CD4. Dicho epítipo inducido por CD4 puede, por ejemplo, incluir epítopos de gp120 en o cerca de la región de unión al correceptor de la proteína.

15 Además de la unión a CD4, la unión de otras moléculas puede inducir la exposición de los epítopos inducidos en gp120. Estas otras moléculas de inducción se consideran similares a CD4 en términos de capacidad de inducir epítopos, debido a que exponen epítopos congruentes con o equivalentes a los epítopos de inducción expuestos tras la unión de CD4 nativo. Estas otras moléculas de inducción incluyen, entre otros, fragmentos de CD4, por ejemplo, sCD4, o un fragmento que contenga los dominios D1 o D1 y D2 de CD4 nativo. Una lectina específica para manosa (SC) también puede servir para exponer un epítipo inducido por CD4 (ver patente estadounidense n.º 5.843.454), al igual que algún MAb anti gp120.

20 *Complejo (en complejo):* Dos proteínas, o fragmentos o derivados de las mismas, se dice que forman un complejo cuando se asocian perceptiblemente entre sí de una forma específica. Dicha asociación puede percibirse en cualesquiera de diversas formas, ya sea directa o indirectamente. Los métodos directos pueden incluir la comigración en condiciones de fraccionamiento no desnaturizantes, por ejemplo. Las percepciones indirectas de la asociación dependerán de los efectos secundarios provocados por la asociación de dos proteínas o dominios de proteínas. Por ejemplo, la formación de un complejo entre una proteína y un anticuerpo puede demostrarse por la inhibición específica para anticuerpos de algunas funciones de la proteína diana. En caso de gp120, la formación de un complejo entre gp120 y un anticuerpo neutralizante para dicha proteína puede percibirse por la determinación del grado con que el anticuerpo inhibe la fusión celular dependiente de gp120 o del grado de infección por el VIH. Los ensayos sobre la inhibición de la fusión celular y la infectividad se tratarán en este documento más adelante.

30 *Exposición de un epítipo inducido:* El proceso por el que dos proteínas interactúan de forma específica para formar un complejo (un complejo de inducción), por el que se provoca un cambio en la conformación de al menos una de las dos proteínas (la proteína diana) de forma que al menos un epítipo que previamente era poco accesible (un epítipo inducido) se hace accesible para la interacción molecular. La formación de dicho complejo de inducción normalmente provoca la exposición de más un epítipo inducido, y cada uno de ellos puede volverse accesible para la interacción intramolecular.

40 *Correceptor del VIH:* Una proteína de superficie celular distinta a CD4 implicada en la interacción del virus VIH y su acceso subsiguiente al interior de una célula diana. Estas proteínas también son denominadas como correceptores de fusión para el VIH. Los ejemplos de dichas proteínas de correceptor incluyen, por ejemplo, miembros de la familia de receptores de quimiocina (por ejemplo CXCR4, CCR5, CCR3 y CCR2B).

45 Las proteínas del correceptor del VIH interactúan con los determinantes de unión al correceptor de gp120. Por lo general se cree que algunos de estos determinantes están expuestos a gp120 únicamente tras la interacción específica de gp120 con CD4, y por el consiguiente cambio de conformación inducido por CD4 en gp120 interactuante. Así, algunos determinantes de unión de correceptores del VIH son, o se solapan con, epítopos inducidos por CD4.

50 La neutralización de gp120 puede conseguirse por la unión específica de proteínas neutralizantes o fragmentos o dominios de proteína a uno o más determinantes de unión de correceptores de gp120, bloqueando así la interacción entre la gp120 en complejo y el correceptor nativo.

55 *Capacidad de neutralización del VIH:* La capacidad perceptible de una molécula para inhibir la infectividad por el virus VIH, ya sea *in vivo* o *in vitro*. Esta materia está repleta de métodos para percibir la capacidad de neutralización de diversas moléculas. Las técnicas incluyen ensayos *in vitro* basados en células mononucleares de sangre periférica (PBMC) (D'Souza *et al.*, 1997); percepción de una vinculación a un virión (Mondor *et al.*, *J. Viral.* 72:3623-3634, 1998); captación de colorante rojo neutro y ensayos de captura de antígenos (patente estadounidense n.º 5.695.927); ensayo de fusión celular de gen indicador basado en vacunas (Nussbaum *et al.*, *J. Virol.* 68:5411-5422, 1994) (ensayos habituales y activados por sCD4); ensayos de infección productiva (percepción del antígeno gag p24 o síntesis RT) (Karn, *HIV: a practical approach*. Oxford Univ. Press, Cambridge, 1995); y ensayos de reducción de títulos de la infección (Karn, 1995).

65 Además, la interacción física entre gp120 y CD4 u otras moléculas similares a CD4 puede examinarse mediante diversos métodos. Ver, por ejemplo, la patente estadounidense n.º 5.843.454 (percepción de los cambios de conformación de gp120 en la unión de varias proteínas por la liberación del virus y susceptibilidad de gp120 al corte mediante trombina del bucle V3). De forma alternativa, puede valorarse la capacidad de la molécula similar a CD4 para luchar por la unión a gp120 tanto con CD4 nativo o un anticuerpo que reconozca el sitio de unión de CD4 en gp120 (CD4BS). Esto permite el cálculo de afinidades de unión relativas mediante las técnicas habituales.

ES 2 313 886 T3

La invención también incluye análogos, derivados o miméticos de las proteínas de fusión presentadas, y que tienen capacidad de neutralización del VIH. Dichas moléculas pueden analizarse para buscarles la capacidad neutralizante del VIH mediante el ensayo de una proteína similar a la proteína de fusión presentada, en la que tiene una o más sustituciones constantes de aminoácidos, o análogos, derivados o miméticos de los mismos, y determinando si la proteína similar, análoga o mimética ofrece una neutralización del VIH. La capacidad de neutralización del VIH y la afinidad de unión a gp120 de estos compuestos derivados pueden evaluarse mediante los métodos conocidos, incluidos los tratados en esta solicitud.

Composición inyectable: Una composición fluida aceptable desde el punto de vista farmacéutico que consta de al menos un principio activo, es decir, una proteína de fusión biespecífica. El principio activo suele estar disuelto o en suspensión en un vehículo aceptable fisiológicamente, y la composición puede además constar de cantidades menores de una o más sustancias auxiliares no tóxicas, como agentes emulsionantes, conservantes y agentes tamponadores de pH y similares. Dichas composiciones inyectables que son útiles para su uso con las proteínas de fusión de esta invención son habituales; sus formulaciones son bien conocidas entre los expertos en la materia.

Aislado: Un componente biológico “aislado” (como una molécula de ácido nucleico, proteína u orgánulo) es aquel que ha sido sustancialmente separado o purificado respecto de otros componentes biológicos en la célula del organismo en que el componente se da naturalmente, por ejemplo, otros ARN o ADN cromosómico o extracromosómico, proteínas y orgánulos. Los ácidos nucleicos y proteínas que han sido “aislados” pueden ser ácidos nucleicos y proteínas purificados mediante métodos de purificación habituales. El término también abarca los ácidos nucleicos y proteínas preparados por expresión recombinante en una célula hospedadora, así como los ácidos nucleicos sintetizados químicamente.

Anticuerpos neutralizantes: Se dice que un anticuerpo neutraliza una función biológica cuando es capaz de unirse específicamente a una proteína diana de tal forma que inhibe las funciones biológicas consiguientes de dicha proteína diana. Por lo general, toda proteína que puede efectuar esta actividad de bloqueo específico se considera una proteína neutralizante; por tanto, los anticuerpos son un tipo específico de proteínas neutralizantes. El complejo formado por la unión de una proteína neutralizante a una proteína diana se denomina complejo neutralizante.

Los anticuerpos que se unen a los virus y bacterias, y así evitan la unión de estos patógenos a las células hospedadoras diana, se dice que neutralizan al patógeno. Así, los anticuerpos que se unen a las proteínas del VIH y reducen perceptiblemente la capacidad del virus para unirse o acceder al interior de las células diana (es decir, células T o macrófagos) son anticuerpos neutralizantes del VIH. Por lo general, los anticuerpos neutralizantes del VIH pueden dividirse en varias clases diferentes en función de la región de la superficie viral de la proteína a la que se une el anticuerpo. Entre las amplias clases de dichos anticuerpos se encuentran los anticuerpos anti gp41 y anti gp120. Existen diversas regiones antigénicas en la proteína gp120 que cuentan con epítomos para la generación natural o de laboratorio de los anticuerpos neutralizantes del VIH (Ver WO 98/36087). Los anticuerpos neutralizantes de amplia reactividad cruzada suelen interactuar con regiones relativamente no variantes de Env.

Una fuente principal de anticuerpos neutralizantes es la sangre periférica de los pacientes infectados con el virus del VIH. Dichos aislados primarios pueden clonarse y/o inmortalizarse mediante las técnicas habituales. Además del aislamiento de los anticuerpos neutralizantes de generación natural, existen unos procedimientos destinados específicamente a producirlos de forma artificial y son bien conocidos entre los expertos en la materia. Ver patentes estadounidenses n.º 5.843.454, 5.695.927, 5.643.756 y 5.013.548, por ejemplo.

Ligando: Un péptido, normalmente de entre dos y 150 residuos de aminoácidos en longitud, que sirve para unir dos dominios de proteína en una proteína de fusión multidominio. Se pueden encontrar ejemplos de ligandos específicos, por ejemplo, en Hennecke *et al* (*Protein Eng.* 11:405-410, 1998); y en las patentes estadounidenses n.º 5.767.260 y 5.856.456.

Según los dominios que se van a unir, y según su posible función en la proteína de fusión, los ligandos pueden ser de entre dos y 150 aminoácidos de longitud, aunque estos límites únicamente se citan a modo de referencia. La longitud del ligando influye en la tendencia de las proteínas de fusión para formar agregados multiméricos no específicos (Alfthan *et al.*, 1998 *Protein Eng.* 8:725-731, 1998). Así, los ligandos más cortos tienden a fomentar la multimerización, mientras que los ligandos más largos tienden a mantener proteínas de fusión monoméricas. Los agregados también pueden minimizarse mediante el uso de secuencias de ligandos específicas, tal y como se muestra en la patente estadounidense n.º 5.856.456.

Los ligandos pueden ser repetitivos o no repetitivos. Un ligando repetitivo clásico utilizado en la producción de una cadena sencilla de Fv (SCFV) es el ligando (Gly₄Ser)₃ (o (GGGS)₃ o (G₄S)₃). Más recientemente se han producido ligandos no repetitivos, y los métodos para la generación aleatoria de dichos ligandos es bien conocida (Hennecke *et al.*, *Protein Eng.* II: 405-410, 1998). Además, los ligandos pueden elegirse para que cuenten con caracteres más o menos secundarios (por ejemplo, carácter helicoidal, patente estadounidense n.º 5.637.481) en función de la conformación deseada en la proteína de fusión final. Cuantos más caracteres secundarios tenga un ligando, más restringida será la estructura de la proteína de fusión final. Por tanto, los ligandos sustancialmente flexibles que no gozan de una gran estructura secundaria permiten la flexión de la proteína de fusión en el ligando.

Un ligando es capaz de retener un dominio de unión de una proteína en proximidad de unión de un sitio de unión diana cuando el ligando tiene una longitud y flexibilidad suficiente para permitir una interacción específica entre el dominio de unión y el sitio de unión diana. En caso de las proteínas de fusión biespecífica de esta invención, un ligando que mantiene la proximidad de unión permite la unión secuencial con la diana del dominio de inducción a la unión de la proteína de fusión en primer lugar, y, a continuación, del dominio de unión inducida. Un ligando que mantiene los dominios de una proteína de fusión biespecífica en proximidad de unión con una diana puede ser considerado un ligando operativo o funcional, dado que tiene relación con dicha proteína de fusión biespecífica.

Oligonucleótido: Una secuencia lineal de polinucleótidos de entre seis y 300 bases de nucleótidos de longitud.

Operablemente unido: Una primera secuencia de ácido nucleico está operablemente unida a una segunda secuencia de ácido nucleico cuando la primera secuencia de ácido nucleico está situada en una relación funcional con la segunda secuencia de ácido nucleico. Por ejemplo, un promotor está operablemente unido a una secuencia de codificación si el promotor afecta a la transcripción o expresión de la secuencia de codificación. Por lo general, las secuencias de ADN operablemente unidas son contiguas, y, cuando es necesario unir dos regiones de codificación de proteínas, están en el mismo marco de lectura.

ORF (open reading frame o marco de lectura abierto): Una serie de tripletes de nucleótidos (codones) de codificación para aminoácidos sin codones de terminación internos. Estas secuencias suelen poder traducirse en un péptido.

Parenteral: Administrado fuera del intestino, es decir, no administrado a través del tracto alimenticio. Normalmente, las formulaciones parenterales son aquellas que se administran mediante cualquier modo posible, excepto la ingesta. Este término hace especial referencia a las inyecciones, ya sean administradas por vía intravenosa, intratecal, intramuscular, intraperitoneal o subcutánea, y varias aplicaciones de superficie, entre las que se incluye la aplicación intranasal, intradérmica y tópica, por ejemplo.

Vehículos aceptables farmacéuticamente: Los vehículos aceptables farmacéuticamente útiles para esta invención son los convencionales. *Remington's Pharmaceutical Sciences*, by E. W. Martin, Mack Publishing Co., Easton, PA, 15th Edition (1975), describe las composiciones y formulaciones adecuadas para la administración farmacéutica de las proteínas de fusión presentadas en este documento.

Por lo general, la naturaleza del portador dependerá del modo especial de administración que se vaya a emplear. Por ejemplo, las formulaciones parenterales suelen ser fluidos inyectables que incluyen fluidos aceptables farmacéutica y fisiológicamente, como el agua, el suero fisiológico, soluciones salinas equilibradas, dextrosa acuosa, glicerol o similar, que actúen como un vehículo. Para las composiciones sólidas (como formas de comprimidos, píldoras, tabletas o cápsulas), pueden incluirse los portadores sólidos convencionales no tóxicos, por ejemplo sustancias farmacéuticas como el manitol, lactosa, almidón o el estearato de magnesio. Además de los vehículos biológicamente neutrales, las composiciones farmacéuticas que se administren pueden contener cantidades menores de sustancias auxiliares no tóxicas, como agentes humectantes o conservantes, y agentes tamponadores de pH y similares, por ejemplo, acetato de sodio o monolaurato de sorbitano.

Purificado: El término purificado no requiere una pureza absoluta; de hecho, debe entenderse como un término relativo. Así, por ejemplo, un preparado purificado de una proteína de fusión es un preparado en el que la proteína de fusión está más enriquecida que lo que puede estar dicha proteína en su entorno generativo, por ejemplo, dentro de una célula o en una cámara de reacción bioquímica. En algunos modelos, una preparación de una proteína de fusión está purificada de forma que la proteína de fusión representa al menos el 50% del contenido total de proteína del preparado.

Recombinante: Una molécula de ácido nucleico recombinante es aquella que tiene una secuencia que no se da de forma natural o que tiene una secuencia creada por una combinación artificial de dos segmentos de secuencia separados de alguna otra manera. Esta combinación artificial puede lograrse mediante síntesis química o, de forma más común, mediante manipulación artificial de segmentos aislados de ácidos nucleicos, es decir, mediante técnicas de modificación genética.

De forma similar, una proteína recombinante es aquella codificada por una molécula de ácido nucleico recombinante.

Identidad de la secuencia: La similitud entre dos secuencias de ácido nucleico o entre dos secuencias de aminoácidos se expresa en términos de similitud entre las secuencias, y a ello nos referimos como identidad de la secuencia. La identidad de la secuencia suele valorarse en términos de porcentajes de identidad (o de similitud u homología); cuanto mayor es el porcentaje, más similar serán las dos secuencias. Los homólogos de la proteína de fusión biespecífica tendrán un grado relativamente alto de identidad de secuencia cuando se alineen mediante los métodos habituales.

Los métodos de alineación de secuencias para su comparación son bien conocidos entre los expertos en la materia. Existen varios programas y algoritmos de alineación descritos en: Smith and Waterman (*Adv. Appl. Math.* 2: 482, 1981); Needleman and Wunsch (*J. Mol. Biol.* 48: 443-453, 1970); Pearson and Lipman (*Proc. Natl. Acad. Sci., USA* 85:2444-2448, 1988); Higgins and Sharp (*Gene*, 73:237-244, 1988); Higgins and Sharp (*CABIOS* 5: 151-153, 1989); Corpet *et al.* (*Nuc. Acids Res.* 16: 10881-10890, 1988); Huang *et al.* (*Comp. Appls. Biosci.* 8: 155-165, 1992); y

ES 2 313 886 T3

Pearson *et al.* (*Methods in Molecular Biology* 24: 307-331,1994). Altschul *et al.* (*Nature Genet.*, 6:119-129,1994) presenta una consideración detallada de los métodos de alineación de secuencias y cálculos de homología.

Las herramientas de alineación ALIGN (Myers and Miller, *CABIOS* 4:11-17, 1989) o LFASTA (Pearson and Lipman, *Proc. Natl. Acad. Sci., USA* 85:2444-2448, 1988) pueden utilizarse para ejecutar las comparaciones de secuencias (Internet Program © 1996, W. R. Pearson and the University of Virginia, "fasta20u63" version2.0u63, fecha de lanzamiento diciembre 1996). ALIGN compara secuencias enteras una frente a otra, mientras que LFASTA compara regiones de similitud local. Estas herramientas de alineaciones y sus tutoriales respectivos están disponibles en Internet en <http://biology.ncsa.uiuc.edu>.

Los ortólogos de las proteínas de fusión biespecíficas presentadas suelen estar caracterizadas por la posesión de más de un 75% de identidad de secuencia contada en una alineación completa con la secuencia de aminoácidos de una proteína de fusión biespecífica utilizando ALIGN con los parámetros definidos por defecto.

La herramienta de búsqueda de alineación total (BLAST o *Basic Local Alignment Search Tool*) de NCBI (Altschul *et al.*, *J Mol Biol.* 1990 215:403-410, 1990) está disponible en diversas fuentes, entre las que se incluye el Centro Nacional de Información sobre Biotecnología (NCBI, Bethesda, Maryland) y en Internet, para su uso en conexión con los programas de análisis de secuencias blastp, blastn, blastx, tblastn y tblastx. Se puede acceder a través de <http://www.ncbi.nlm.nih.gov/BLAST/>. Se puede acceder a una descripción de cómo determinar la identidad de la secuencia utilizando el programa que se puede encontrar en http://www.ncbi.nlm.nih.gov/BLAST/blast_help.html.

Para comparaciones de secuencias de aminoácidos de más de 30 aminoácidos, se utiliza la función "Blast 2 sequences" mediante la matriz por defecto BLOSUM62 configurada para los parámetros por defecto (coste de existencia de penalización de II, y coste de penalización por residuo de I). Cuando se alinean péptidos cortos (de menos de unos 30 aminoácidos) la alineación debe efectuarse utilizando la función "Blast 2 sequences", utilizando la matriz PAM30 configurada como parámetros por defecto (penalización 9 abierta, extensión de penalizaciones 1). Las proteínas con mayor similitud a las secuencias de referencia se mostrarán aumentando el porcentaje de identidades cuando se evalúe con este método, con una identidad de secuencia de al menos 90%, al menos 92%, al menos 94%, al menos 95%, al menos 97%, al menos 98% o al menos 99%. Asimismo, la identidad de la secuencia puede compararse a lo largo de la longitud completa de uno o ambos dominios de unión de las proteínas de fusión presentadas. En este caso, las identidades del porcentaje serán esencialmente similares a aquellas tratadas en la identidad de secuencia de longitud total.

Cuando, para la identidad de secuencia, se compara un fragmento significativamente menor que la secuencia completa, los homólogos suelen tener al menos un 80% de identidad de secuencia en ventanas cortas de entre 10 y 20 aminoácidos, y pueden tener identidades de secuencia de al menos 85%, al menos 90%, al menos 95% o al menos 99%, según su similitud con la secuencia de referencia. La identidad de secuencia sobre dichas ventanas cortas puede determinarse mediante LFASTA; los métodos se describen en <http://biology.ncsa.uiuc.edu>. Cualquier experto en la materia apreciará que estos registros de identidad de secuencia se facilitan únicamente a modo de referencia; es totalmente posible que puedan obtenerse homólogos extremadamente importantes que se salgan de los registros facilitados. La presente invención aporta no sólo los homólogos péptidos que se describen anteriormente, sino también las moléculas de ácido nucleico que codifican a dichos homólogos.

Una indicación alternativa de que dos moléculas de ácido nucleico están estrechamente vinculadas es que las dos moléculas se hibridan la una a la otra en condiciones estrictas. Las condiciones estrictas dependen de las secuencias, y son diferentes en parámetros ambientales diferentes. Por lo general, las condiciones estrictas se seleccionan entre 5°C y los 20°C menos que el punto de fusión termal (T_m) para la secuencia específica en una fuerza iónica y pH definidos. El T_m es la temperatura (en fuerza iónica y pH definidos) a la que el 50% de la secuencia meta hibrida con una sonda perfectamente encajada. Las condiciones para la hibridación del ácido nucleico y el cálculo de rigurosidad pueden encontrarse en Sambrook *et al.* (In *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, New York, 1989), y Tijssen (*Laboratory Techniques in Biochemistry and Molecular Biology Part I*, Ch. 2, Elsevier, New York, 1993). Las moléculas de ácido nucleico que hibridan en condiciones rigurosas a las secuencias de proteínas de fusión biespecíficas presentadas, hibridarán por lo general a una sonda basada bien en la secuencia completa de codificación de una proteína de fusión, un dominio de unión completo y otras porciones seleccionadas de la secuencia de codificación en condiciones de lavado de 0,2 x SSC, 0,1% SDS a 65°C.

Las secuencias de ácido nucleico que no muestran un alto grado de identidad pueden, no obstante, codificar secuencias de aminoácidos similares debido a la degeneración del código genético. Se entiende que los cambios en la secuencia de ácido nucleico pueden efectuarse mediante esta degeneración para producir múltiples secuencias de ácido nucleico, y cada una codificará sustancialmente la misma proteína.

Agente de unión específica: Un agente que une de forma sustancial únicamente a un agente definido. Así, un agente de unión específica de gp120 se une únicamente de forma sustancial con la proteína gp120. Tal y como se utiliza en este documento, "agente de unión específica de gp120" incluye a los anticuerpos anti gp120 y otros agentes que se unen sustancialmente únicamente a una proteína gp120.

Los anticuerpos anti gp120 pueden producirse con los procedimientos habituales descritos en un gran número de textos, entre los que se incluye Harlow and Lane (*Using Antibodies, A Laboratory Manual*, CSHL, New York, 1999,

ISBN 0-87969-544-7). Asimismo, algunas técnicas pueden mejorar la producción de los anticuerpos neutralizantes (patentes estadounidenses n.º 5.843.454, 5.695.927, 5.643.756 y 5.013.548). La determinación de que un agente en especial se une sustancialmente únicamente a una proteína gp120 puede efectuarse mediante el uso o la adaptación de los procedimientos rutinarios. Un ensayo adecuado *in vitro* hace uso del procedimiento Western Blot (descrito en muchos textos habituales, entre los que se incluye Harlow and Lane, 1999). La técnica Western Blot puede utilizarse para determinar que un agente de unión a una proteína determinada, como el anticuerpo monoclonal anti gp120, se une de forma sustancial únicamente a la proteína MSG. Los anticuerpos para gp120 son bien conocidos entre los expertos en la materia.

Los fragmentos más cortos de anticuerpos también pueden servir como agentes de unión específica. Por ejemplo, los Fab, Fv y cadenas sencillas de Fv (SCFv) que se unen a gp120 serían agentes de unión específica de gp120.

Cantidad terapéuticamente eficaz de una proteína de fusión biespecífica: Una cantidad de proteína de fusión biespecífica suficiente para conseguir un efecto deseado en un sujeto tratado. Por ejemplo, podría ser la cantidad necesaria para inhibir la proliferación viral o para neutralizar de forma perceptible la enfermedad con los mecanismos de unión del organismo. Por lo general, esta cantidad será suficiente para inhibir de forma perceptible la replicación o capacidad de infección del virus (por ejemplo, el VIH).

Una cantidad eficaz de proteína de fusión biespecífica puede administrarse en una única dosis o en varias dosis, por ejemplo, a diario, durante el transcurso de un tratamiento. Sin embargo, la cantidad eficaz de proteína de fusión dependerá de la proteína de fusión aplicada, del sujeto tratado, de la gravedad y tipo de infección y de la forma de administración de la proteína de fusión. Por ejemplo, una cantidad terapéuticamente eficaz de proteína de fusión puede variar entre 0,01 mg/kg de peso corporal y 1 g/kg de peso corporal.

Las proteínas de fusión presentadas en la presente invención tienen una igual aplicación en tratamientos médicos y veterinarios. Por tanto, el término general “sujeto tratado” se entiende que incluye a todos los animales (es decir, humanos, simios, perros, gatos, caballos y vacas) que están o pueden estar infectados con un virus u otro organismo causante de enfermedad susceptible de ser neutralizado con una proteína de fusión biespecífica.

Transformado: Una célula transformada es una célula en la que se ha introducido una molécula de ácido nucleico mediante técnicas de biología molecular. Tal y como se utiliza en este documento, el término transformación abarca todas las técnicas por las que una molécula de ácido nucleico puede introducirse en dicha célula, mediante transfección con vectores virales, transformación con vectores plásmidos e introducción de ADN desnudo mediante electroporación, lipofección y aceleración de partículas.

Vector: Una molécula de ácido nucleico introducida en una célula hospedadora y que produce de esta manera una célula hospedadora transformada. Un vector puede incluir secuencias de ácido nucleico que le permiten replicarse en una célula hospedadora, como origen de una replicación. Un vector puede también incluir un o más genes marcadores seleccionables y otros elementos genéticos bien conocidos entre los expertos en la materia.

II. Construcción, Expresión y Purificación de las Proteínas de Fusión Biespecíficas

A. Selección de dominios del componente

La presente invención, de conformidad con las reivindicaciones, presenta en términos generales una proteína de fusión biespecífica que se une a dos sitios diferentes de una proteína diana. De este modo, toda proteína diana con dos sitios de unión diferentes es un ejemplo de diana para una proteína de fusión biespecífica. Las dianas en especial tienen proteínas en las que uno de los dos sitios de unión (el sitio de unión inducido) está expuesto/inducido por la unión de la proteína de fusión a un primer sitio de unión (el sitio de inducción a la unión) de la diana. La elección de los dominios de unión de la proteína para la incorporación de la proteína de fusión biespecífica presentada será establecida por la proteína diana elegida. La elección del ligando también estará influida por la proteína diana y por los sitios de unión elegidos. Por lo general, el ligando utilizado en una fusión biespecífica será de una longitud y tendrá un carácter secundario tal que mantendrá el dominio de unión inducida dentro de la proximidad de unión del sitio de unión inducido de la proteína diana una vez que el dominio de inducción a la unión de la proteína de fusión haya formado un complejo específico con la diana.

En algunos modelos, la proteína diana es una glicoproteína de superficie del VIH, por ejemplo la gp120 del VIH-1. En algunos de estos y otros modelos, el sitio de inducción a la unión es el sitio de unión CD4 de gp120. Así, el dominio de inducción a la unión de la proteína de fusión biespecífica presentada puede ser un fragmento de unión de CD4, por ejemplo, sCD4. De forma alternativa, cualquier otra molécula que interactúe de forma específica con gp120 de tal manera que se exponga a uno o más epítopos inducidos también serviría como la fuente de un dominio de proteína de inducción a la unión. Los fragmentos específicos utilizados para construir la proteína de fusión deberían elegirse de forma que la conformación de la fusión final provoque una unión funcional y de inducción a gp120; esto se puede someter a ensayo bien de forma directa (es decir, con mediciones de afinidad), o de forma indirecta (es decir, ensayos de neutralización).

Las copias de CD4 que no derivan de CD4 se presentan como fuentes de dominios de inducción a la unión de las proteínas de fusión presentadas. Por ejemplo, una lectina específica para manosa (SC) puede servir para inducir

cambios de conformación inducida de CD4 (ver patente estadounidense n.º 5.843.454). De forma alternativa, también pueden utilizarse los anticuerpos que se unen al sitio de unión de CD4 u otro epítipo de gp120, y así inducen un cambio de conformación similar a CD4 en la proteína en complejo.

5 Los análogos de CD4 no péptidos también pueden utilizarse en esta presentación, por ejemplo, análogos estructurales orgánicos o inorgánicos de los dominios de interacción con gp120 de la molécula CD4.

10 Los dominios de unión inducida de una proteína de fusión dirigida contra gp120 incluirán anticuerpos (o fragmentos del mismo) que reconocen epítopos inducidos de la gp120 en complejo. En algunos modelos, dichos anticuerpos tienen una alta reactividad cruzada frente a diversos aislados de VIH-1. Los epítopos inducidos incluyen todos aquellos a los que nos referimos como epítopos inducidos de CD4 (CD4i) y, en especial, aquellos que solapan los determinantes de unión a correceptores de gp120. Los anticuerpos monoclonales neutralizantes previamente identificados pueden utilizarse, e incluyen, entre otros, los anticuerpos monoclonales humanos 17b, 48d y CG10.

15 De igual modo, los dominios de unión inducida de las moléculas quiméricas presentadas pueden ser moléculas no péptidos, por ejemplo análogos estructurales orgánicos o no orgánicos de SCFv(17b).

20 Además de los anticuerpos que se unen a epítopos inducidos de gp120, otras fuentes de dominios de unión inducida incluyen fragmentos de correceptores que interactúan de forma específica con un/os dominio/s de unión a correceptor de gp120.

25 La construcción de una proteína de fusión biespecífica específica para gp120 puede verse ayudada por la revisión de la estructura cristalográfica de rayos X del complejo ternario que contiene el núcleo de gp120, un fragmento de dos dominios de CD4 (D1D2), y un Fab de un MAb (17) humano con alta reactividad cruzada dirigidos contra los determinantes de unión al correceptor de gp120 (Kwong *et al.*, *Nature* 393:648-659, 1998). Los exámenes informáticos de las coordenadas estructurales de este complejo ternario a través de FRODO (Jones *et al. Meth. Enzymol.* 115:157-171, 1985; Jones, *J. Appl. Cryst.* 11:268-272, 1978; Pflugrath *et al. Methods and Applications in Crystallography*, páginas 407-420, Clarendon Press, Oxford), han revelado las elecciones para la construcción de la proteína quimérica. La distancia más corta entre los extremos libres de CD4 y el 17b Fab es de 56 Å, es decir, desde el extremo C-terminal libre del fragmento D1D2 sCD4 al extremo N-terminal de la cadena pesada de 17b Fab. Un ligando que conecte estos extremos estaría esencialmente libre de impedimentos estéricos entre CD4 y el extremo N-terminal de la cadena ligera 17b. También se podrían hacer unas conexiones posibles entre el extremo N-terminal de CD4 y el extremo C-terminal de las cadenas pesadas o ligeras 17b; dichas conexiones requerirían ligandos de unos 65 Å y de unos 86 Å, respectivamente. En las últimas dos conexiones se necesita un ligando para eludir otras porciones del complejo, incluidos los bucles variables voluminosos.

B. Ensamblaje

40 Es bien conocida la construcción de moléculas quiméricas, en especial de proteínas de fusión, a partir de dominios de proteínas conocidas. Por lo general, una molécula de ácido nucleico que codifica los dominios de proteínas deseados se acoplan mediante técnicas de modificación genética habituales para crear un único oligonucleótido de fusión operablemente unido. Las técnicas de biología molecular pueden encontrarse en Sambrook *et al.* (In *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, New York, 1989). En los siguientes documentos de patentes pueden encontrarse ejemplos específicos de proteínas de multidominio modificadas genéticamente, en especial, aquellas basadas en moléculas de la superfamilia de la inmunoglobulina, unidas por varios ligandos:

Patente estadounidense n.º 5.856.456 (“ligando para polipéptidos de fusión ligados”);

50 Patente estadounidense n.º 5.696.237 (“Proteína de fusión de toxina anticuerpo recombinante”);

Patente estadounidense n.º 5.767.260 (“Proteínas de fusión de unión con antígenos”);

Patente estadounidense n.º 5.587.455 (“Agente citotóxico contra la infección específica por virus”) y

55 WO 98/36087 (“Tolerancia inmunológica a los epítopos del VIH”).

Los análogos no péptidos que sirven como dominios de inducción a la unión o de unión inducida de la invención pueden unirse al dominio opuesto de las moléculas quiméricas mediante las técnicas de unión química conocidas, entre las que se incluye la unión cruzada química. Los elementos de unión cruzada son bien conocidos, y se pueden encontrar ejemplos de moléculas utilizadas para la unión cruzada, por ejemplo, en la patente estadounidense n.º 6.027.890 (“Métodos y composiciones para mejorar la sensibilidad en el análisis de los ensayos biológicos”).

C. Expresión

65 Cualquier experto en la materia entenderá que existen miles de formas para expresar una proteína recombinante para que pueda ser purificada a continuación. Por lo general, un vector de expresión que tenga una secuencia de ácido nucleico que codifique la proteína deseada se transformará en un microorganismo para la expresión. Dichos microorganismos pueden ser procariontes (bacterias) o eucariontes (como la levadura). Un ejemplo de las especies de

bacterias que pueden utilizarse es la *Escherichia coli* (*E. coli*), que se ha utilizado muy ampliamente como sistema de expresión experimental en el laboratorio. Un sistema de expresión eucariota puede utilizarse cuando la proteína de interés requiere unas modificaciones post-traduccionales específicas para eucariotas, como la glicosilación. Así, la proteína puede expresarse con un sistema de expresión viral (como el virus Vaccinia).

5

La proteína también puede expresarse en cultivo de tejido celular animal, y dicho sistema puede utilizarse cuando las modificaciones de proteínas específicas de animales son deseables o requeridas en la proteína recombinante.

El vector de expresión puede contar con una secuencia que codifique un péptido diana, posicionado de tal forma que se fusione con la secuencia de codificación de la proteína de fusión biespecífica. Esto permite que la proteína de fusión biespecífica se dirija contra un lugar subcelular o extracelular específico. Se conocen varios péptidos de diana procariotas y eucariotas, así como moléculas de ácido nucleico que codifican dichos péptidos. En un sistema de expresión procariota, podría usarse una secuencia de señal para segregarse la nueva proteína sintetizada. En un sistema de expresión eucariota, el péptido diana especificaría el hecho de ser diana de la proteína híbrida en uno o más compartimentos subcelulares específicos, o sería segregado de la célula, según el tipo de péptido elegido. A través del uso de una secuencia de señal de secreción eucariota, la proteína de fusión biespecífica puede expresarse en un animal transgénico (como una vaca, cerdo u oveja) de tal modo que la proteína se segregue a la leche del animal.

Los vectores aptos para la transformación estable de células cultivables son bien conocidos. Por lo general, dichos vectores cuentan con un sitio de clonación múltiple para la inserción de una molécula de ácido nucleico clonada, de forma que se efectúe bajo el control de transcripción de las secuencias reguladoras en 3' y 5'. Además, los vectores de transformación incluyen uno o más marcadores seleccionables; para la transformación de las bacterias, este suele ser un gen resistente a antibióticos. Dichos vectores de transformación suelen contar también con una región reguladora de promotor (es decir, una región reguladora que controla la expresión inducible o constitutiva), un sitio de inicio de la transcripción, un sitio de unión a un ribosoma, una señal de proceso de RNA y un sitio de terminación de la transcripción, cada uno de ellos situado en relación con el sitio de clonación múltiple. Para la producción de grandes cantidades de proteínas recombinantes, puede utilizarse un promotor inducible. Esto permite la producción selectiva de la proteína recombinante, y permite unos niveles más altos de producción que los promotores constitutivos, y permite asimismo la producción de proteínas recombinantes que pueden ser tóxicas para la célula de expresión, si se expresa constitutivamente.

Además de estas pautas generales, en el mercado existen kits de purificación/expresión de proteínas. Ver, por ejemplo, el sistema de expresión QIAexpress™ de QIAGEN (Chatsworth, California) y varios sistemas de expresión de INVITROGEN (Carlsbad, California). Según los datos facilitados por los fabricantes, dichos kits podrán utilizarse para la producción y purificación de las proteínas de fusión biespecíficas presentadas.

35

D. Purificación

Cualquier experto en la materia entenderá que existen miles de formas de purificar polipéptidos recombinantes, y que los métodos habituales de purificación de proteínas pueden utilizarse para purificar las proteínas de fusión biespecíficas presentadas. Dichos métodos incluyen, por ejemplo, métodos de cromatografía de proteínas con intercambio de iones, filtrado de gel, HPLC, cromatografía de afinidad a los anticuerpos monoclonales y aislamiento de cuerpos de inclusión de proteínas no solubles tras una sobreproducción. Asimismo, las etiquetas de afinidad para la purificación, por ejemplo, una secuencia de seis-histidina, pueden fusionarse de forma recombinante a la proteína y utilizarse para facilitar la purificación del polipéptido. Un sitio proteolítico específico, por ejemplo un sitio de digestión específico para la trombina, puede manipularse en la proteína entre la etiqueta y la fusión en sí misma para facilitar la eliminación de la etiqueta tras la purificación.

Los kits que existen en el mercado para la purificación/expresión de la proteína ofrecen unos protocolos a medida para la purificación de proteínas creada según cada sistema. Ver, por ejemplo, el sistema de expresión QIAexpress™ de QIAGEN (Chatsworth, California) y varios sistemas de expresión facilitados por INVITROGEN (Carlsbad, California). Cuando se utiliza un kit comercial para producir una proteína de fusión biespecífica, el protocolo de purificación del fabricante es un protocolo especialmente revelado para la purificación de dicha proteína. Por ejemplo, las proteínas expresadas con una etiqueta hexa-his aminoterminal pueden purificarse mediante la unión a una matriz de cromatografía de afinidad del metal del ácido nitrilotriacético-níquel (Ni-NTA) (*The QIAexpressionist*, QIAGEN, 1997).

De forma alternativa, las especificidades de unión, de los dominios de unión primero o segundo, o de ambos, de la proteína de fusión presentada pueden explotarse para facilitar la purificación específica de las proteínas. Un método de ejecución de dicha purificación específica sería la cromatografía de columna mediante resina de columna a la que se vincula la molécula diana o un epítipo, fragmento o dominio de la molécula diana.

Si la proteína de fusión biespecífica se produce en forma secretada, es decir, secretada en la leche de un animal transgénico, la purificación será del fluido secretado. De forma alternativa, la purificación puede ser innecesaria si la proteína de fusión puede aplicarse directamente al sujeto en el fluido secretado (como la leche).

65

III. Variación de la Proteína de Fusión Biespecífica

A. Variantes de secuencia

5 Las características de unión, y, por tanto, la actividad neutralizante de las proteínas de fusión presentadas en este documento, no recaen en la secuencia de aminoácidos exacta, sino en la estructura tridimensional inherente en las secuencias de aminoácidos codificadas por las secuencias de ADN. Es posible recrear las características de unión de alguna de estas proteínas o dominios de proteínas de la presente invención recreando la estructura tridimensional sin tener que recrear necesariamente la secuencia exacta de aminoácidos. Esto puede conseguirse diseñando una secuencia de ácido nucleico que codifique la estructura tridimensional, pero que difiera, por ejemplo por razón de la redundancia, del código genético. De forma similar, la secuencia de ADN también puede variarse mientras que siga produciendo una proteína neutralizante funcional.

15 Las proteínas de unión variantes biespecíficas neutralizantes incluyen proteínas que difieren en su secuencia de aminoácidos de la secuencia presentada, pero que comparten una homología significativa en cuanto a la estructura de la secuencia con cualquiera de las proteínas aportadas. La variación puede darse en un dominio único de la proteína de fusión (es decir, el primero o segundo dominio de unión, o el ligando). La variación también puede ocurrir en más de uno de estos dominios en cualquier proteína variante en especial. Dichos variantes pueden producirse manipulando la secuencia de nucleótidos de, por ejemplo, una secuencia de codificación CD4-SCFv(17b) mediante los procedimientos habituales, entre los que se incluye la mutagénesis dirigida al sitio o PCR. Las modificaciones más simples abarcan la sustitución de uno o más aminoácidos para aquellos aminoácidos que tienen unas propiedades bioquímicas similares. Es probable que estas sustituciones llamadas conservadoras tengan un impacto mínimo sobre la actividad de la proteína resultante, especialmente cuando se efectúan fuera del sitio de unión de cada dominio. La Tabla 1 muestra los aminoácidos que pueden sustituirse por un aminoácido original en una proteína, y que se consideran sustituciones conservadoras.

TABLA 1

Residuo original	Sustitución conservadora
Ala	ser
Arg	lys
Asn	gln; his
Asp	glu
Cys	ser
Gln	asn
Glu	asp
Gly	pro
His	asn; gln
Ile	leu; val
Leu	ile; val
Lys	arg; gln; glu
Met	leu; ile
Phe	met; leu; tyr
Ser	thr
Thr	ser
Trp	tyr
Tyr	trp; phe
Val	ile; leu

Se pueden obtener más cambios sustanciales en la estructura de la proteína mediante selección de las sustituciones de aminoácidos menos conservadoras que las que figuran en la Tabla 1. Dichos cambios incluyen residuos cambiantes que difieren de forma más significativa en su efecto para mantener la estructura del polipéptido (es decir, conformación helicoidal o de hoja plegada) cerca de la sustitución, carga o hidrofobicidad de la molécula en el sitio diana, o el volumen de una cadena lateral específica. Por lo general, se espera que las siguientes sustituciones produzcan los mayores cambios en las propiedades de la proteína: (a) un residuo hidrofílico (como serilo o treonilo) se sustituye por un residuo hidrofóbico (como leucilo, isoleucilo, fenilalanina, valilo o alanilo); (b) una cisteína o prolina se sustituye por cualquier otro residuo; (c) un residuo con una cadena lateral electropositiva (como lisilo, arginilo o histadilo) se sustituye por un residuo electronegativo (como glutamilo o aspartilo); o (d) un residuo con una cadena lateral voluminosa (como fenilalanina) se sustituye por uno sin cadena lateral (como glicina).

El dominio de unión variante o las secuencias de codificación de las proteínas de fusión puede producirse mediante las técnicas de mutagénesis de ADN habituales, por ejemplo, la mutagénesis del iniciador M13. Se puede acceder a los detalles de estas técnicas en Sambrook (In *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, New York, 1989), Cap. 15. Por el uso de dichas técnicas, pueden crearse unos variantes que difieran en menor medida de las secuencias de codificación de la proteína de fusión biespecífica presentada en este documento. Las moléculas de ADN y las secuencias de nucleótidos que derivan de las presentadas específicamente en este documento y que difieren de las presentadas por la supresión, adición o sustitución de nucleótidos, mientras sigan codificando a una proteína que se une dos veces a gp120, neutralizando así la capacidad de infección del virus VIH, estarán incluidos en la presente invención. En su forma más sencilla, dichos variantes pueden diferir de las secuencias presentadas por la alteración de la región codificadora para encajar con la parcialidad en el uso del codón del organismo en particular en el cual se introducirá la molécula.

De forma alternativa, la región codificadora puede alterarse aprovechando la degeneración del código genético para alterar la secuencia de codificación, de forma que mientras la secuencia de nucleótidos esté sustancialmente alterada, no obstante, codificará una proteína con una secuencia de aminoácidos sustancialmente similar a las secuencias de fusión presentadas. Por ejemplo, el residuo 18 de los aminoácidos de la proteína CD4-SCFv(17b) (tras el corte de la secuencia de la señal N-terminal) es alanina. El triplete codón del nucleótido GCT codifica este residuo de alanina. Debido a la degeneración del código genético, tres otros tripletes codones de nucleótidos (GCG, GCC y GCA) también codifican la alanina. De ese modo, la secuencia de nucleótidos de la secuencia de codificación de CD4-SCFv (17b) presentada podría modificarse en esta posición con cualquiera de estos tres codones alternativos sin afectar a la composición de aminoácidos o a las características de la proteína codificada. En función de la degeneración del código genético, las moléculas variantes de ADN pueden derivar del cADN y de las secuencias genéticas presentadas en este documento que utilizan técnicas de mutagénesis de ADN, tal y como se describe anteriormente o por síntesis de secuencias de ADN. De este modo, esta invención también abarca las secuencias de ácido nucleico que codifican una proteína de fusión biespecífica neutralizante, pero que varían de las secuencias de ácido nucleico presentadas en virtud de la degeneración del código genético.

B. Modificaciones de péptidos

La presente invención incluye moléculas activas biológicamente que copian la acción de las proteínas de fusión biespecíficas de la presente invención y neutralizan específicamente el Env del VIH. Las proteínas de la invención incluyen modelos sintéticos de proteínas que se generan de forma natural descritas en este documento, así como análogos (moléculas orgánicas no péptidos), derivados (moléculas de proteínas funcionalizadas químicamente y obtenidas empezando con las secuencias de péptidos presentadas) y variantes (homólogos) de estas proteínas que se unen específicamente con y neutralizan la gp120 del VIH. Cada proteína de la invención se forma de una secuencia de aminoácidos, que pueden ser aminoácidos L y/o D, generados de forma natural o de alguna otra manera.

Las proteínas pueden ser modificadas por una gran variedad de técnicas químicas para producir derivados con esencialmente la misma actividad que las proteínas no modificadas, y, opcionalmente, con otras propiedades deseadas. Por ejemplo, los grupos de ácido carboxílico de la proteína, ya sea una cadena carboxi terminal o lateral pueden aparecer en forma de sal de un catión aceptable por la práctica farmacéutica, o esterificados en forma de ester C₁-C₁₆, o convertidos a una amida de fórmula NR₁R₂, donde R₁ y R₂ son independientemente H o C₁-C₁₆ alquilo, o combinados para formar un anillo heterocíclico, como un anillo integrado por 5 o 6 miembros. Los grupos aminos de la proteína, ya sea una cadena amino terminal o lateral, pueden aparecer en forma de ácido de adición de sal aceptado por la práctica farmacéutica, como HCl, HBr, ácido acético, benzoico, tolueno, sulfónico, maleico, tartárico y otras sales orgánicas, o puede ser modificado a C₁-C₁₆ alquilo o diaquilo amino, o convertido posteriormente a un amido.

Los grupos hidroxilos de las cadenas laterales de la proteína pueden convertirse a alcoxi C₁-C₁₆ o a ester C₁-C₁₆, usando las técnicas conocidas entre los expertos en la materia. Los anillos fenil y felónicos de las cadenas laterales de la proteína pueden sustituirse por uno o más átomos de halógenos, como la fluorina, clorina, bromina o yodina, o con alquilo C₁-C₁₆, ácidos carboxílicos y ésteres de los mismos, o amidas de dichos ácidos carboxílicos. Los grupos de metileno de las cadenas laterales de la proteína pueden extenderse hasta alquilenos C₁-C₁₆ homólogos. Los tioles pueden quedar protegidos con cualquier grupo protector bien conocido, como los grupos acetamidas. Los expertos en la materia reconocerán los métodos de introducción de estructuras cíclicas en las proteínas de esta invención para seleccionar y aportar unos límites de conformación a la estructura que da como resultado una mejor estabilidad.

También puede ser ventajoso introducir uno o más enlaces disulfuro para conectar los marcos de las cadenas pesadas y ligeras en el dominio SCFv. Esta modificación suele mejorar la estabilidad y afinidad de los SCFV (Reiter *et al.*, *Protein Engineering* 7:697-704, 1994). También aquí, la estructura de cristal de rayos X que contiene 17 Fab (Kwong *et al.*, *Nature* 393:648-659, 1998) puede utilizarse para evaluar los sitios óptimos para la creación de los residuos de cisteína de las cadenas pesadas y ligeras.

Los modelos peptidomiméticos y organomiméticos también se encuentran dentro del alcance de la presente invención, donde la disposición tridimensional de los constituyentes químicos de dichos peptidomiméticos y organomiméticos copian la disposición tridimensional de la estructura de la proteína y las cadenas laterales de aminoácidos componentes de la proteína de fusión neutralizante biespecífica, resultante de dichos peptidomiméticos y organomiméticos de las proteínas de esta invención con una capacidad neutralizante mejorada o valorable. Para las aplicaciones informáticas de modelado por ordenador, un farmacóforo es una definición idealizada, tridimensional de los requisitos estructurales para la actividad biológica. Los peptidomiméticos y organomiméticos pueden diseñarse para encajar con cada farmacóforo con los programas de modelado por ordenador actuales (mediante el uso de un diseño de fármaco asistido por ordenador o CAAD). Ver Walters, "Computer-Assisted Modeling of Drugs", en Klegernian & Groves, eds., 1993, *Pharmaceutical Biotechnology*, Interpharm Press: Buffalo Grove, IL, pp. 165-174 y *Principles of Pharmacology* Munson (ed.) 1995, Cap. 102, para las descripciones de las técnicas utilizadas en CADD. Los miméticos preparados mediante dichas técnicas que producen proteínas de fusión neutralizantes también se incluyen en el marco de la invención.

C. Variación de la longitud del dominio

Se apreciará que los dominios de las proteínas de la presente invención pueden combinarse para producir moléculas de proteínas de fusión sin unir necesariamente los componentes en el mismo lugar. Se cree que es posible usar fragmentos más cortos o más largos de cada dominio del componente vinculados por un ligando funcional. Por ejemplo, cualquier componente que se una con unos 10 residuos de aminoácidos del residuo especificado, y que siga facilitando un fragmento de unión funcional, abarca el mismo dominio. Sin embargo, pueden utilizarse los dominios de longitud sustancialmente más larga o más corta. Por ejemplo, en ciertos modelos, la proteína puede inducir una secuencia líder más un CD4 de cuatro dominios (D1-D4, residuos de aminoácidos 1-372), o sólo el primer dominio de CD4 (residuos D1 1-113).

IV. Actividad de las proteínas de fusión

Es importante evaluar la actividad química, física y biológica de las proteínas de fusión biespecíficas presentadas. Entre otros usos, dichos ensayos permiten la optimización de los dominios elegidos, así como la optimización de la longitud y conformación de los ligandos utilizados para conectarlos. Las moléculas de control deberían incluirse en cada ensayo; normalmente, dichos ensayos incluirán cada dominio por separado, así como los dos dominios como moléculas independientes mezcladas en la reacción, por ejemplo en un ratio molar 1:1. En caso de una proteína de fusión biespecífica CD4-SCFv(17b), dichos controles también incluirán a sCD4 y SCFv(17b), por ejemplo.

A. Proteína de fusión para una proteína diana

La afinidad de la proteína de fusión con la proteína diana puede determinarse mediante el uso de diversas técnicas. Por ejemplo, los análisis sobre coimmunoprecipitación con proteínas etiquetadas metabólicamente pueden utilizarse para determinar la unión de las proteínas sCD4-SCFv, como sCD4-SCFv(17b) con gp120 del VIH-1 soluble mediante MAb anti gp120 que no interfieran con la unión de CD4 (como MAb D47, que se une a V3) o un anticuerpo policlonal con el extremo C-término de gp120. ELISA también puede utilizarse para examinar las características de unión de cada dominio de la quimera.

B. Ensayos de neutralización

Pueden utilizarse diversos ensayos para evaluar la capacidad de las proteínas de fusión presentadas para inhibir la función de la proteína diana. Los componentes individuales de la proteína de fusión servirán como controles. Por lo general, los ensayos serán específicos para la proteína de fusión/diana. Por ejemplo, muchos análisis funcionales pueden probar la capacidad de las fusiones sCD4-SCFv para neutralizar la Env del VIH. Es especialmente ventajoso el uso de las Env de diversas cepas del VIH-1 para probar la amplitud de la inhibición (capacidad de neutralización) de cada proteína de fusión para diferentes subtipos genéticos y diferentes fenotipos del VIH-1 (es decir, el uso del correceptor). Además, es ventajoso para probar dichas proteínas de fusión en los ensayos habituales y activados con sCD4 para la fusión celular mediante Env. Para tales experimentos, los ejemplos de controles son los conocidos MAb neutralizantes del VIH-1, y los MAB frente a epítomos inducidos de CD4 sobre gp120. También deberá probarse la posible inhibición sinérgica con otros MAB neutralizantes conocidos de amplia reactividad cruzada (como b12, 2F5, F105, 2G12).

En caso de proteínas de fusión dirigidas contra gp120, el ensayo de fusión celular del gen indicador basado en Vaccinia puede utilizarse para evaluar la inhibición de la fusión (Nussbaum *et al.*, *J. Virol.* 68:5411-5422, 1994). Una población de células de tejido de cultivo (como BS-C-1, HeLa o NIH 3T3) que expresan uniformemente la unión codificada con el virus vaccinia y las glicoproteínas de superficie virales que median en la fusión, se mezcla con otra población que expresa el/los correspondiente/s receptor/es celular/es. En el caso de las fusiones sCD4-SCFv, donde la

proteína diana es la gp120 del VIH-1, una población celular expresa la Env del VIH-1, mientras que la otra expresa necesariamente los receptores del VIH-1 (es decir, CD4 y un receptor de quimiocina). El citoplasma de cualquiera de las poblaciones celulares también contiene ADN polimerasa del bacteriófago T7 codificado con el virus vaccinia; el citoplasma de la otra contiene un plásmido transfectado con el gen *lacZ* de *E. coli* ligado al promotor del T7. Tras la mezcla de las dos poblaciones, la fusión celular da como resultado la activación del gen *lacZ* mediante la introducción de ARN polimerasa del T7 en proximidad con el promotor *lacZ* del transfectado T7 en el citoplasma de las células fusionadas. La actividad resultante β -galactosidasa (β -gal) es proporcional a la cantidad de fusión que se da y que puede valorarse mediante un ensayo colorimétrico de lisatos de células detergentes o tinción *in situ*. La actividad neutralizante de la fusión celular de las proteínas de fusión biespecíficas se valora por tanto mediante la evaluación de su inhibición de producción de β -gal.

Las fusiones dirigidas contra gp120 (como sCD4-SCFv) también pueden someterse a prueba de su capacidad de bloqueo de la infección por VIH-1 mediante ensayos de única etapa (como el uso de líneas celulares de indicador, Vodicka *et al.*, *Virology* 233:193-198, 1997). Las células diana que expresan CD4 y un correceptor específico, y que contienen un gen indicador de *lacZ* ligado al LRT (repetición de terminal larga) del VIH-1, están infectadas con cepas específicas del VIH (Vodicka, 1997). La integración de un provirus del VIH en estas células conduce a la producción del transactivador viral, Tat, que a continuación ejecuta la expresión del gen β -gal a través de la interacción del LTR. La actividad de sCD4-SCFv se valora por su inhibición de producción de células positivas de β -gal (tintadas de azul con X-gal), que es proporcional a su capacidad de bloquear la infección por el VIH-1.

V. Incorporación de proteínas de fusión biespecíficas en composiciones farmacéuticas

Las composiciones farmacéuticas que comprenden al menos una proteína de fusión biespecífica tal y como se describe en este documento como un ingrediente activo estarán formuladas normalmente con un vehículo sólido o líquido, en función del modo particular de administración elegido. Los vehículos y excipientes útiles aceptables farmacéuticamente para esta invención son los convencionales. Por ejemplo, las formulaciones parenterales suelen ser fluidos inyectables que son vehículos fluidos aceptables farmacéuticos y fisiológicamente, como el agua, el suero fisiológico, otras soluciones salinas equilibradas, dextrosa acuosa, glicerol o similar. Los excipientes que pueden incluirse son, por ejemplo, otras proteínas, como la albúmina de suero humano o las preparaciones de plasma. Si se desea, las composiciones farmacéuticas que se administrarán pueden contener también cantidades menores de sustancias auxiliares no tóxicas, como agentes humectantes o conservantes, y agentes tamponadores de pH y similares, por ejemplo, acetato de sodio o monolaurato de sorbitano.

También pueden incluirse otros agentes farmacéuticos y medicinales, por ejemplo los derivados de nucleósido (como AZT) o inhibidores de la proteasa. Asimismo, puede resultar ventajoso incluir otros inhibidores de fusión, como uno o más anticuerpos neutralizantes.

La forma de dosificación de la composición farmacéutica se determinará por el modo de administración elegido. Por ejemplo, se pueden utilizar formulaciones tópicas u orales, además de los fluidos inyectables. Los preparados tópicos pueden incluir colirios, ungüentos, esprais y similares. Las formulaciones orales pueden ser líquidas (como jarabes, soluciones o suspensiones), o sólidas (como comprimidos, píldoras, tabletas o cápsulas). Para las composiciones orales, los portadores sólidos no tóxicos convencionales pueden incluir sustancias farmacéuticas como el manitol, lactosa, almidón o el estearato de magnesio. Los métodos actuales de preparación de dichas formas de dosificación son bien conocidos, o lo serán, entre los expertos en la materia.

Las composiciones farmacéuticas que comprenden una proteína de fusión biespecífica pueden formularse en formas de dosificación unitarias, adecuadas para la administración individual de dosis precisas. Una dosis unitaria posible contiene aproximadamente 100 μ g de proteína. La cantidad de componente activo administrado dependerá del sujeto tratado, de la gravedad del problema y de la forma de administración, y se deja bajo el criterio del especialista que lo prescriba. Dentro de estas limitaciones, la formulación que se administrará contendrá una cantidad del componente/s activo/s en una suma eficaz para lograr el efecto deseado en el sujeto que se está tratando.

VI. Uso clínico de las proteínas de fusión biespecíficas

La potente actividad neutralizante frente a los virus mostrada por las proteínas de fusión biespecíficas presentadas hace de ellas un elemento útil para el tratamiento de infecciones virales en humanos y otros sujetos animales. Los posibles virus susceptibles son, entre otros, los virus de la inmunodeficiencia, como el VIH o similar, o virus relacionados en simios y otros animales. Además, otros sistemas virales o microbianos implicados en la interacción de un primer sitio de inducción a la unión y un segundo sitio de unión inducida de una única proteína también serán susceptibles de neutralización mediante las proteínas de fusión biespecífica de la presente invención. Las proteínas de fusión biespecíficas presentadas en este documento también pueden utilizarse en la detección altamente sensible o purificación de la proteína diana.

Las proteínas de fusión biespecíficas de esta invención pueden administrarse a humanos u otros animales en cuyas células sean eficaces, en diversas formas, ya sea tópica, oral, intravenosa, intramuscular, intraperitoneal, intranasal, intradérmica, intratecal y subcutánea. El modo de administración en especial, así como el programa de las dosis, lo seleccionará el especialista clínico correspondiente, teniendo en cuenta los factores específicos del caso (es decir, el sujeto en cuestión, la enfermedad y el grado de la enfermedad, así como si el tratamiento será profiláctico o posterior

a la infección). El tratamiento estará formado por dosis diarias o multidiarias de las proteínas de fusión biespecíficas durante un periodo de unos pocos días, meses o incluso años.

Si el tratamiento es a través de la administración directa de células que expresan la proteína de fusión biespecífica al sujeto, dichas células (como células madre hematopoyéticas o pluripotenciales transgénicas o células B) pueden administrarse en una dosis de entre unas 10^6 y 10^{10} células en una o varias ocasiones. El número de células dependerá del paciente, así como la proteína de fusión y las células elegidas para expresar la proteína.

En la patente estadounidense n.º 5.529.774 se presenta una estrategia general para la transferencia de genes a células donantes. Por lo general, un gen que codifica una proteína con los efectos terapéuticos deseados se clona en un vector de expresión viral, y el vector se introduce a continuación en el organismo diana. El virus infecta las células y produce la secuencia de la proteína *in vivo* cuando tiene el efecto terapéutico deseado. Ver, por ejemplo, Zabner *et al.*, *Cell* 75:207-216, 1993. Como alternativa a la adición de las secuencias que codifican la proteína de fusión biespecífica o una proteína homóloga en el ADN de un virus, también es posible introducir dicho gen en el ADN somático de las células infectadas o no infectadas mediante los métodos bien conocidos entre los expertos en la materia (Sambrook *et al.*, In *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, New York, 1989). Estos métodos pueden utilizarse para introducir las proteínas de fusión presentadas en este documento en las células humanas para aportar una resistencia a largo plazo contra la infección por el VIH-1 o el SIDA. Por ejemplo, la terapia génica puede utilizarse para segregar la proteína en las superficies mucosas, como la mucosa vaginal, rectal u oral.

Las proteínas de fusión biespecíficas dirigidas contra gp120 del VIH-1, por ejemplo sCD4-SCFv(17b), son especialmente útiles para la prevención o infección durante o inmediatamente después de la exposición al VIH (como la transmisión madre/hijo, profilaxis tras la exposición e inhibidor tópico). En dichos casos, se administrarán una o más dosis de la proteína de fusión biespecífica antes o poco después del momento desencadenante. Para evitar o mejorar la transmisión madre/hijo de la infección viral, por ejemplo, puede ser beneficioso administrar la proteína de fusión biespecífica dirigida contra gp120 a la madre de forma intermitente durante el embarazo, y/o inmediatamente antes o durante el parto, y/o directamente al neonato inmediatamente tras su nacimiento. Los tratamientos profilácticos tras la exposición pueden ser especialmente beneficiosos cuando ha habido una exposición accidental (por ejemplo, una exposición accidental médica), entre lo que se incluye un pinchazo con una aguja contaminada o la exposición médica a sangre u otro fluido contaminado con el VIH.

La presente invención también incluye combinaciones de proteínas de fusión biespecíficas quiméricas con uno o más agentes útiles en el tratamiento de la enfermedad, como la enfermedad por VIH. Por ejemplo, los compuestos de esta invención pueden administrarse, ya sea antes o después de la exposición al virus, en combinación con dosis eficaces de otros antivirales, inmunomoduladores, antiinfecciosos y/o vacunas. El término "administración en combinación" se refiere tanto a la administración simultánea como secuencial de los agentes activos.

Algunos ejemplos de agentes antivirales que pueden utilizarse en combinación con las proteínas de fusión biespecíficas quiméricas de la invención son: AL-721 (de Ethigen, Los Ángeles, California), beta interferón humano recombinante (de Triton Biosciences, Alameda, California), Acemannan (de Carrington Labs, Irving, Texas), gangiclovir (de Syntex, Palo Alto, California), didehidrodeoxitimidina o d4T (de Bristol-Myers-Squibb), EL10 (de Elan Corp., Gainesville, Georgia), dideoxicitidina o ddC (de Hoffman-LaRoche), Novapren (de Novaferon labs, Inc., Akron, Ohio), zidovudina o AZT (de Burroughs Wellcome), ribaririn (de Viratek, Costa Mesa, California), alfa interferón y aciclovir (de Burroughs Wellcome), Indinavir (de Merck & Co.), 3TC (de Glaxo Wellcome), Ritonavir (de Abbott), Saquinavir (de Hoffmann-LaRoche), y otros.

Algunos ejemplos de inmunomoduladores que pueden utilizarse en combinación con las proteínas de fusión biespecíficas quiméricas de la invención son AS-101 (Wyeth-Ayerst Labs.), bropirimina (Upjohn), gamma interferón (Genentech), GM-CSF (Genetics Institute), IL-2 (Cetus o Hoffman-LaRoche), globulina inmune humana (Cutter Biological), IMREG (de Imreg, Nueva Orleans, Luisiana), SK&F106528, y TNF (Genentech).

Algunos ejemplos de antiinfecciosos en los que se pueden utilizar las proteínas de fusión biespecíficas quiméricas pueden ser clindamicina con primaquina (de Upjohn, para el tratamiento de *Pneumocystis pneumonia*), fluconazol (de Pfizer, para el tratamiento de meningitis criptocócica o candidiasis), nistatina, pentamidina, trimetoprim-sulfametoxazol, y muchos otros.

Las terapias de combinación, por supuesto, no se limitan a las citadas en estos ejemplos, pero incluyen toda composición para el tratamiento de la enfermedad por VIH (incluido el tratamiento contra el SIDA).

VII. Kits

Las proteínas quiméricas presentadas en este documento pueden suministrarse en forma de kit para el uso en prevención y/o para el tratamiento de enfermedades (como la infección por el VIH y el SIDA). En dichos kits, se facilitará una cantidad clínica eficaz de una o más proteínas de fusión biespecíficas quiméricas, en uno o más contenedores. Las proteínas de fusión biespecíficas quiméricas pueden suministrarse suspendidas en solución acuosa o en comprimidos liofilizados, por ejemplo. En algunos modelos, las proteínas quiméricas se facilitarán en forma de composición farmacéutica.

ES 2 313 886 T3

Los kits, conforme a esta invención, también pueden ir acompañados de instrucciones, normalmente por escrito, para ayudar al usuario que está tratándose la enfermedad (como una infección por VIH o el SIDA) con una proteína de fusión biespecífica quimérica. Opcionalmente, dichas instrucciones pueden facilitarse en un soporte informático.

- 5 Los contenedores en los que se suministran las proteínas pueden ser contenedores convencionales capaces de albergar la forma suministrada, por ejemplo, tubos de microfuga, ampollas o botellas. En algunas aplicaciones, las proteínas quiméricas pueden suministrarse en cantidades de único uso con la dosis preparada en tubos, desechables habituales o contenedores equivalentes.
- 10 La cantidad de proteína de fusión biespecífica quimérica suministrada en el kit puede ser toda cantidad adecuada, en función, por ejemplo, del mercado al que se dirige dicho producto. Por ejemplo, si el kit está adaptado para la investigación o el uso clínico, la cantidad de cada proteína quimérica suministrada podría ser una cantidad suficiente para varios tratamientos.
- 15 De conformidad con esta invención, algunos kits, también incluirán uno o más agentes útiles para el tratamiento de la enfermedad, como la enfermedad por VIH. Por ejemplo, dichos kits pueden incluir una o más dosis eficaces de otros antivirales, inmunomoduladores, antiinfecciosos y/o vacunas.

Ejemplo 1

20 *Construcción de una secuencia de codificación de CD4-SCFv(17b)*

Una proteína de fusión dirigida contra gp120, sCD4.SCFv(17b), se construye ligando el extremo C-terminal de CD4 (D1D2, 183 residuos aminoácidos) al extremo N-terminal de la cadena pesada de 17b SCFv, que contiene la cadena pesada en su extremo N-terminal, ligada a través de su extremo C-terminal a la cadena ligera (ver diagrama esquemático del constructo, Fig. 3). El ADN de 17b SCFv se obtuvo de R. Wyatt and J. Sodroski, Dana Farber Cancer Institute, Boston, Massachusetts. La hibridoma productor de 17b-MAb se obtuvo de J. Robinson, Tulane University.

30 Los ligandos se eligieron para tener una longitud y flexibilidad suficientes para conectar los segmentos de proteína deseada sin inducir una torsión inaceptable. Para el SCFv, se eligió la secuencia de 15 residuos de aminoácidos (Gly₄Ser)₃, (L2 designado), que se utilizó de forma satisfactoria para la producción de SCFv. Esta secuencia aporta una flexibilidad excelente con una agregación mínima. El ligando entre el extremo C-terminal de CD4 y el N-terminal del SCFv (L1 designado, SEC. ID N.º 2), es de siete repeticiones de la misma secuencia Gly₄Ser. Las estimaciones conservadoras indican que este ligando de 35 residuos aminoácidos es lo suficientemente largo como para permitir que CD4 y SCFv se unan simultáneamente a sus respectivos sitios de unión en gp120. En la Fig. 3 se muestra un esquema del constructo genético. Se ha introducido un único sitio de restricción BamH I dentro de L1 para permitir la producción de constructos con ligandos más cortos o más largos, y especialmente para aportar controles negativos (ligandos muy cortos, no permitiendo así que tanto las porciones de CD4 como de SCFv de una misma molécula se unan simultáneamente a sus sitios de unión respectivos en gp120).

40 El plásmido iniciador de CD4 es pCB3, que contiene la longitud completa de CD4 cADN (incluida su secuencia de señal en 5' natural) en el plásmido de expresión de vaccinia pSC59 (Broder & Berger, *J. Virol.* 67:913-926, 1993). Este plásmido fue digerido con Stu I, que corta cerca del extremo del segundo dominio de sCD4, y con Spe I, que corta el vector debajo de la inserción de CD4 y deja un saliente en 5'.

45 Los oligonucleótidos sintéticos (SEC. ID N.º 11) se anillaron juntos para recapitular el extremo 5' de la segunda mitad del sitio Stu I (CCT) y las siguientes dos bases (CC) del cADN de CD4, y para producir un saliente Spe I en el extremo 3' (este sitio se destruirá una vez ligado a pCB-3). La secuencia de oligonucleótidos reconstruye el resto del segundo dominio de CD4 (a través de ser₁₈₃), y codifica al ligando intermedio de 37 aminoácidos (gly₄)ser₆gly₄thr₂ser, seguido directamente por la secuencia de terminación translacional universal (UTS) (SEC. ID N.º 6). Se ha incluido deliberadamente un sitio BamH en el ligando cerca del extremo de la tercera repetición (gly₄ser) para permitir la consiguiente unión de 17b SCFv con la secuencia exacta de L1, y para permitir la modificación de la longitud del ligando. El plásmido intermedio resultante es el designado pCD 1. Este constructo está confirmado por el análisis de secuencia de ADN efectuado con las técnicas habituales. Para facilitar los procedimientos siguientes, la secuencia del ligando de sCD4 se volvió a clonar en el derivado pSC59 sin un sitio BamH I, formando el plásmido intermedio pCD2.

60 El plásmido iniciador 17b que contiene cADN de 17b SCFv en un vector plásmido (*pmt del 0*) fue donado por el Dr. Richard Wyatt (Dana Farber Cancer Institute, Boston, Massachusetts). El cADN de SCFv se construyó con la cadena pesada en el segmento 5' y la cadena ligera en el segmento 3', ligados a través del ADN que codifica el ligando L2 (gly₄ser)₃. El constructo de 17b SCFv tiene una secuencia de señal TPA en el extremo 5', y unas secuencias que corresponden a un sitio de corte de la trombina y a una etiqueta hexa-his (para facilitar la purificación) en el extremo 3', seguido por un codón de parada. También puede producirse un constructo sin el sitio de corte de la trombina y la etiqueta hexa-his.

65 La tecnología PCR se utilizó para unir la secuencia de 17b SCFv a la secuencia del ligando de CD4 en pCD2. Los iniciadores adecuados se representan en las SEC. ID N.º 7 y 9. El iniciador (5') anverso (SEC. ID N.º 7) contiene un sitio BamH I cerca del extremo 5' (precedido por un saliente), seguido por nucleótidos que reconstruyen la tercera (gly₄ser) repetición y la cuarta repetición adicional (gly₄ser); todo esto, seguido de nucleótidos que corresponden

exactamente con el inicio de la cadena pesada 17b (excluyendo la secuencia de señal 5' que comienza con CAG GTG). El iniciador 3' (SEC. ID N.º 9) empieza con unos sitios de restricción conveniente para clonarse en pCD2 (Spe I y otros), seguidos de nucleótidos exactamente complementarios al extremo 3' de la secuencia de 17b SCFv en pmt del 0 (codón de parada, etiqueta hexa-his y sitio de corte de la trombina).

Estos iniciadores se utilizan para iniciar el vector plásmido que contiene la secuencia 17b SCFv de *pmt del 0*, y el producto PCR resultante digerido por BamH I más una enzima de restricción que corta el extremo 3' opuesto (es decir, Spe I). Este fragmento digerido se clona a la fuerza en pCD2, que ha sido digerido con las mismas enzimas (BamH I y Spe I). El plásmido resultante (designado aquí como pCD3) contiene el origen del constructo final sCD4-SCFv(17b) (con el sitio de corte de la trombina y la etiqueta hexa-his) abajo del promotor de vaccinia precoz/tardío, fuerte, sintético en pSC59. Existe unos sitios únicos y convenientes de restricción en cada lado de la secuencia sCD4-SCFv para futuros posibles pasos de clonación.

El cADN de 17b SCFv (incluida la secuencia de señal 5') también ha sido eliminada del vector *pmt del 0* por la digestión del enzima de restricción o PCR, y clonado en el plásmido de expresión de vaccinia pSC59 para aportar un constructo de control.

Ejemplo 2

Expresión y purificación de la proteína de fusión CD4-SCFv(17b)

A. Expresión

Para pequeñas cantidades de expresión de proteína, puede utilizarse la tecnología de expresión de vaccinia para producir el sCD4-SCFv(17b) (así como la proteína de control 17b SCFv). El plásmido que contiene el constructo en el plásmido de expresión de vaccinia pSC59 se utiliza para producir un recombinante de vaccinia, mediante el uso de la tecnología habitual. Para dicha expresión, las células adecuadas (HeLa, BSC-I, etc.) se infectan con el virus de vaccinia recombinante; tras una incubación de entre 24 y 36 horas a 37°C, la proteína recombinante se presenta en el sobrenadante de cultivo. Pueden llevarse a cabo estudios iniciales bioquímicos y funcionales con el sobrenadante no fraccionado; cuando sea necesario, la proteína sCD4-SCFv puede ser purificada (ver más adelante). Se pueden llevar a cabo experimentos iniciales a pequeña escala con pequeñas cantidades de material (5-20 microgramos, obtenidos a partir de 1-5 x 10⁷ células). La preparación puede intensificarse; para producción a gran escala, puede ser ventajoso utilizar las tecnologías de mayor rendimiento para la expresión de proteínas recombinantes (como baculovirus, levadura o *E. coli*).

Se analizó la expresión para el producto de la proteína secretada pCD1 (los primeros dos dominios de CD4 a través de ser₁₈₃ más el ligando de 37 aminoácidos). Las células BSC-1 se transfectaron con pCD1 y se infectaron con un virus de vaccinia en estado natural, y a continuación se incubaron durante la noche a 37°C. Los sobrenadantes se analizaron con análisis Western (inmunoblot) utilizando anticuerpos contra CD4. Tal y como se esperaba, la proteína codificada por pCD1 migró ligeramente más lenta que el sCD4 de dos dominios estándar purificado (Upjohn- Pharmacia, Kalamazoo, Michigan).

La proteína de fusión de longitud completa pCD3 sCD4-SCFv(17b) (sCD4-17b) se ha expresado y probado de forma similar, y el dominio 17b SCFv (clonado en pSC59) puede examinarse del mismo modo. La proteína de fusión sCD4-17b (al menos un fragmento de lo secretado) tenía el tamaño molecular esperado (aproximadamente 55 kDa) cuando se analizó con SDS PAGE y la técnica Western blot. La proteína reaccionó con fuerza con los anticuerpos contra CD4 o la etiqueta hexa-his, confirmando la presencia de estas porciones de N-terminal y C-terminal, así como el marco de lectura correcto.

B. Purificación

La proteína de fusión expresada como se ha construido anteriormente con una etiqueta hexa-his amino terminal se purificó utilizando esta etiqueta molecular. La etiqueta permite la unión específica y la purificación de la proteína de fusión mediante la unión a la matriz de cromatografía de afinidad del metal del ácido nitrilotriacético-níquel (Ni-NT A) (ver, por ejemplo, *The QIAexpressionist*, QIAGEN, 1997). Se utilizó una etiqueta hexa-his en los presentes ejemplos.

Los métodos de purificación alternativa incluyen una combinación de HPLC y cromatografía de columna líquida convencional (filtrado de gel, intercambio de iones, enfoque isoeléctrico).

C. Caracterización primaria

Para probar la unión de gp120 al dominio 17b de la proteína de fusión sCD4-17b, se cubrieron primero unas placas ELISA de 96 pocillos con el 13B8.2 anti CD4 MAb (Beckman Coulter, Chaska, MN, Catálogo n.º IM0398), cuyo epítipo de CD4 solapa los determinantes implicados en la unión a gp120. Las placas se incubaron bien con sCD4-17b, bien con una solución tampón de control. Cuando la proteína quimérica se capturó de esta manera, la fracción 17b siguió estando disponible para unir gp120 en complejo a sCD4; sin embargo, la fracción de sCD4 no se une

libremente a gp120, dado que se capturó en la placa por el anti CD4 MAb que bloquea el sitio de unión. Las placas se incubaron con gp120 (aislado IIIB Ratner *et al.*, *Nature* 313:277-284, 1985) en complejo con sCD4. La unión de gp120 fue detectada por un antisuero anti gp120 policlonal, seguida por un anti IgI de conejo conjugado con peroxidasa de rábano. Las placas se lavaron y se incubaron con sustrato ABTS, y el producto oxidado se cuantificó mediante la medición de la absorción a 450 nm. Los resultados indican una unión específica: los valores de la absorción fueron de 0,15 con la proteína quimérica sCD4-17b, comparados con el 0,05 con la solución tampón de control.

Para probar la funcionalidad de la región sCD4 de la proteína quimérica, las placas ELISA se cubrieron primero con una etiqueta anti his MAb (QIAGEN Inc., Valencia, California, Catálogo n.º 34670), y a continuación se incubaron con la proteína quimérica purificada o con una solución tampón de control. Con la proteína quimérica capturada de esta forma, la fracción sCD4 estaba disponible para unirse a gp120 libre; no obstante, la fracción 17b no pudo unirse a gp120, que no estaba en forma compleja con sCD4. Las placas se incubaron con gp120 libre, y se detectó la unión de la forma antedicha. Los resultados indican una unión específica: los valores de la absorción fueron de 0,46 con sCD4-17b, comparados con el 0,05 con la solución tampón de control. Así, los ensayos ELISA confirmaron las propiedades de unión funcional esperadas para cada fracción de la proteína quimérica: la fracción 17b se ligó al complejo gp120/sCD4, y la fracción CD4 se ligó a gp120 libre.

Ejemplo 3

Valoraciones de neutralización de la superficie del VIH

A. Fusión celular de gen indicador basado en vaccinia

La fusión celular a través de Env activada por CD4 se valoró mediante el ensayo de gen indicador basado en vaccinia (Nussbaum *et al.*, *J. Virol.* 68:5411-5422, 1994). Para el experimento que figura en la Tabla 2 y en la Fig. 1, las células efectoras (HeLa) se transfectaron con plásmido pGINT7-β-gal (LacZ ligado al promotor T7), y, a continuación, se infectaron con recombinantes de vaccinia que codifican bien el mutante no escindido Unc Env o en estado natural (WT) SF162 Env (Broder & Berger, *Proc. Natl. Sci. USA* 92: 9004-9008, 1995). Las células diana se crearon transfectando células NIH 3T3 con plásmido pGA9-CCR5 que contenía el CCR5 cADN ligado a un promotor de vaccinia (Alkhatib *et al.*, *Science* 272:1955-1958, 1996), y, a continuación, se infectaron estas células con el virus WR de vaccinia en estado natural. Las células diana también contaban con y expresaban una T7 ARN polimerasa del bacteriófago. Previamente a los ensayos de fusión, las células transfectadas se incubaron durante la noche a 31°C para permitir la expresión de las proteínas recombinantes, y, a continuación, se lavaron.

Para cada ensayo de fusión, se prepararon mezclas de células efectoras y diana (1 x 10⁵ de cada tipo de célula por pocillo, en pocillos duplicados) en ausencia o en presencia de sCD4 (100 nM final). Tras 2,5 horas a 37°C, las células fueron lisadas con NP-40, y la actividad β-gal se cuantificó mediante procedimientos habituales (Tabla 2 y Fig. 1). Se determinó una fusión relativa (actividad específica de β-gal) a partir de las muestras duplicadas, y se calculó como WT-Unc.

TABLA 2

Ensayo de fusión de células con gen indicador basado en vaccinia utilizando CD4 soluble

	<u>B-gal total</u>	<u>Unc Env (control)</u>				<u>WT Env (SF162)</u>				<u>Fusión relativa</u> (WT-Unc)
		<u>duplicados</u>		<u>promedi</u>	<u>duplicados</u>		<u>promedi</u>			
	(Datos sin procesar)			<u>0</u>			<u>0</u>			
- sCD4	0,50	0,40	0,45	0,50	0,40	0,45	0,0			
+ sCD4	0,40	0,50	0,45	6,60	5,20	5,90	5,45			

Los ensayos de fusión basados en vaccinia pueden utilizarse para neutralizar la capacidad de las proteínas de fusión biespecíficas presentadas en este documento. La capacidad neutralizante de MAb 17b se ha demostrado que depende de la adición de CD4 soluble de la siguiente forma (ver Tabla 3 y Fig. 2) Las células efectoras se crearon mediante células HeLa coinfectoras con un recombinante de vaccinia que codificaba VIH-1 Env (LAV) (Broder & Berger, *Proc. Natl. Acad. Sci., USA* 92:9004-9008, 1995), y otra que codificaba polimerasa T7 ARN. Las células diana se crearon mediante cotransfección de células NIH 3T3 con plásmidos pYFI-fusin (Feng *et al.*, *Science* 272:872-877,

ES 2 313 886 T3

1996) que codificaban CXCR4 y pGINT7- β -gal (*LacZ* ligado al promotor T7). Las células diana fueron a continuación transfectadas con virus de vaccinia vCB-3 (que codificaban CD4, ensayo habitual) (Broder *et al.*, *Virology* 193:483-491, 1993), o WR (virus en estado natural, ensayo sCD4). Como controles en segundo plano, las células diana fueron transfectadas únicamente con pGINT7- β -gal (es decir, no con un correceptor). Las células transfectadas se incubaron durante la noche a 31°C para permitir la expresión de proteínas recombinantes, y, a continuación, se lavaron. Las células efectoras se incubaron durante 30 minutos a 37°C con la concentración indicada de MAb 17b (Tabla 3).

Para los ensayos de fusión se prepararon unas mezclas entre las células efectoras y las células diana indicadas (2 x 10⁵ de cada tipo de célula por pocillo, en pocillos duplicados); en el ensayo habitual, las células diana expresaron CXCR4 y CD4, y CD4 no soluble añadido; en el ensayo de sCD4, las células diana expresaron sólo CXCR4 y se añadió CD4 (200 nM final). Tras 2,5 horas a 37°C, las células se lisaron y se cuantificó la actividad de β -gal. Los valores en segundo plano de B-gal de control (ensayo habitual, 0,6; ensayo de sCD4, 0,2), obtenidos con células diana sin correceptor, se sustrayeron para aportar los datos presentados en la Tabla 3. Los datos representan el porcentaje de control (sin MAb) por cada ensayo.

TABLA 3

Inhibición mediada por MAb del ensayo de fusión

[17b] (μ g/ml)	Ensayo habitual B-gal	% control	Ensayo sCD4 B-gal	% control
Ninguno	42,3	100,0	11,89	100,0
0,1	39,5	93,4	13,55	113,9
0,5	43,9	103,8	4,66	39,2
1	39,8	94,1	1,68	14,1
5	50,5	119,4	0	0

La eficacia de las proteínas de fusión biespecíficas descritas en este documento para la neutralización de la fusión se prueba de modo similar, añadiendo unas cantidades diferentes de la proteína de fusión biespecífica, es decir, sCD4-SCFv(17b), al ensayo anterior. No es necesario añadir los sCD4 y SCFv(17b) exógenos y otras proteínas de unión a gp120, aunque pueden utilizarse como controles, como se ha hecho anteriormente, o para determinar las eficacias inhibitoras comparadas con la proteína de fusión biespecífica. Utilizando este ensayo, los efectos de los promedios a partir de las células de control infectadas con virus de vaccinia en estado natural WR se compararon con los promedios de las células infectadas con el virus de vaccinia recombinante que codificaba sCD4-17b. Los valores específicos de β -galactosidasa relativos eran de 23,4, con el promedio de control, y de <1 con el promedio de sCD4-SCFv. Así, la proteína sCD4-17b quimérica inhibía con fuerza la fusión celular al VIH a través de Env.

Ejemplo 4

50 *Producción a gran escala y análisis de sCD4-17b*

Para producir grandes cantidades de la proteína sCD4-17b, el constructo de ADN se ha transferido al vector plásmido pET11b (Novagen, Madison, Wisconsin, Catálogo n.º 69437-3), que es adecuado para la expresión de alto nivel inducible en *E. coli*. Este sistema implica la clonación de genes diana bajo control de una fuerte señal de transcripción del bacteriófago 17. Una vez establecidos en una célula bacteriana hospedadora sin expresión, los plásmidos se transfieren a hospedadores de expresión que contenían una copia cromosómica del gen de la polimerasa ARN T7 bajo el control de lacUV5, y la expresión de la proteína recombinante de interés (en este caso sCD4-17b) se induce por la adición de IPTG. La proteína expresada está producida a muy alto nivel, y puede constituir más del 50% de la proteína celular total en el cultivo inducido dentro de unas pocas horas después de la inducción. Los resultados de la técnica Western blot indican una expresión de alto nivel de sCD4-17b a partir del plásmido pET11b.

La proteína producida puede ser desnaturizada y renaturalizada a partir de los cuerpos de inclusión para aportar una gran cantidad de proteína funcional sCD4-17b. Esta proteína puede utilizarse en estudios *in vitro* para probar la inhibición en ensayos tanto de fusión celular mediante Env como de infección por VIH (producción p24).

Asimismo, la proteína sCD4-17b puede utilizarse para estudios *in vivo*. Un modelo *in vivo* abarca ratones SCID reconstituidos con timo más hígado humano (Pettoello-Montovani *et al.*, *J. Infect. Dis.* 177:337-346, 1998); este

ES 2 313 886 T3

sistema se utilizará para probar si sCD4-17b inhibe (y hasta qué punto), o evita, la infección aguda por VIH. Este sistema se ha utilizado de forma satisfactoria para demostrar las actividades de bloqueo potente de otros agentes anti VIH (como inhibidores de proteasa e inhibidores de transcriptasa reversos y toxinas dirigidas contra Env) (Pettoello-Montovani *et al.*, *J. Infect. Dis.*: 177:337-346, 1998).

5

Un segundo ejemplo de un sistema *in vivo* para probar la actividad de sCD4-17b implica a los monos Rhesus expuestos a los virus SVIH (virus recombinantes que contienen gag y pol SIV, más una superficie del VIH; Li *et al.*; *J. Virol.*: 69:7061-7071, 1995). Este sistema se utilizará para probar si la proteína sCD4-17b inhibe (y hasta qué punto), o previene, la infección aguda por SVIH.

10

Los efectos de sCD4-17b contra la infección crónica también se examinarán utilizando de nuevo el sistema SCID-hu/VIH-1 del ratón y el sistema SHIV del mono.

15

Tanto los sistemas de estudio *in vivo* e *in vitro* se utilizarán también para probar la potencia de la proteína sCD4-17b cuando se utiliza en combinación con otros agentes anti VIH (como RT y los inhibidores de proteasa u otros MAb neutralizantes de VIH-1).

20

Los ejemplos anteriores únicamente se facilitan a modo de ilustración. Pueden efectuarse numerosas variaciones sobre las moléculas biológicas y sobre los métodos descritos anteriormente para crear y utilizar las moléculas de fusión biespecífica capaces de unir dos sitios de una única proteína, y, especialmente, dos sitios de la proteína de superficie gp120 del VIH, y para su uso para la detección, tratamiento y prevención de la infección por el VIH. Reivindicamos todos estos temas que forman parte de las siguientes reivindicaciones.

25

30

35

40

45

50

55

60

65

REIVINDICACIONES

- 5 1. Una proteína de fusión biespecífica neutralizante que consta de un primer dominio de unión, un segundo dominio de unión y un ligando que conecta el primer dominio con el segundo dominio, y donde la proteína de fusión interactúa específicamente con una proteína que tiene dos sitios diferentes de unión (proteína diana) para formar un complejo,
- 10 donde el ligando mantiene el segundo dominio de unión en proximidad de unión con el segundo sitio de unión cuando el primer dominio de unión está vinculado al primer dominio de unión;
- 15 donde el primer dominio de unión está en CD4, que es capaz de unirse a un primer sitio de unión de la proteína diana (sitio de inducción) induciendo así un cambio de conformación en la proteína diana y exponiendo un segundo sitio de unión, haciendo así que el segundo sitio de unión sea accesible para la interacción molecular, cuyo segundo sitio de unión es un epítipo que anteriormente era poco accesible (epítipo inducido);
- 20 donde el segundo dominio de unión consta de una porción de unión de una región variable de una cadena ligera o pesada de anticuerpos capaz de formar un complejo neutralizante con el segundo sitio de unión de la proteína diana;
- y donde el complejo neutralizante está formado por la unión específica del segundo dominio de unión al segundo dominio de unión de tal forma que se inhiba el funcionamiento biológico subsiguiente de la proteína diana.
- 25 2. Una proteína de fusión conforme a la reivindicación 1, donde la proteína diana es una proteína de superficie viral de un virus.
- 30 3. Una proteína de fusión conforme a la reivindicación 2, donde el virus es el virus de inmunodeficiencia humana (VIH).
4. Una proteína de fusión conforme a la reivindicación 3, donde la proteína de superficie viral es gp120.
5. Una proteína conforme a la reivindicación 1, donde el primer dominio de unión comprende CD4D1 o CD4D1D2.
6. Una proteína conforme a la reivindicación 5, donde el primer dominio de unión es sCD4.
- 35 7. Una proteína conforme a la reivindicación 1, donde el segundo dominio de unión consta de una cadena sencilla Fv (SCFv).
8. Una proteína conforme a la reivindicación 7, donde el SCFv es SCFv17b, y donde scFv17b corresponde al aminoácido 244 hasta el aminoácido 495 de la SEC. ID N.º 3.
- 40 9. Una proteína conforme a la reivindicación 1, donde el segundo dominio de unión deriva de un anticuerpo monoclonal neutralizante.
10. Una proteína conforme a la reivindicación 4, donde el segundo dominio de unión copia una actividad biológica de una molécula correceptora del VIH en la unión a gp120.
- 45 11. Una proteína conforme a la reivindicación 4, donde el primer dominio de unión es un dominio de unión CD4 de gp120.
12. Una proteína conforme a la reivindicación 1, donde el segundo dominio de unión es capaz de unirse al menos a un epítipo altamente conservado de la región de gp120 de unión al correceptor.
- 50 13. Una proteína conforme a la reivindicación 1, donde el ligando tiene una longitud de 15-100 angstroms.
14. Una proteína conforme a la reivindicación 1, donde el ligando tiene una longitud de 10-100 angstroms de residuos de aminoácidos.
- 55 15. Una proteína conforme a la reivindicación 1, donde el ligando consta de al menos una secuencia de aminoácidos representada por SEC. ID N.º 1.
- 60 16. Una proteína conforme a la reivindicación 15, donde el ligando consta de una secuencia de aminoácidos representada por SEC. ID N.º 2.
17. Una proteína de fusión biespecífica recombinante funcional capaz de unirse a dos sitios de gp120, a saber:
- 65 a) sCD4;
- b) SCFv(17b), donde SCFv(17b) corresponde al aminoácido 244 hasta el aminoácido 495 de SEC. ID N.º 3; y

ES 2 313 886 T3

c) un ligando de una longitud suficiente para mantener a SCFv(17b) en proximidad de unión a un epítipo SCFv (17b) cuando sCD4 está vinculado a gp120.

5 18. Una proteína conforme a la reivindicación 17, donde el ligando tiene una secuencia de aminoácidos representada por SEC. ID N.º 2.

19. Una molécula aislada de ácido nucleico que codifica una proteína conforme a la reivindicación 17.

10 20. Una molécula de ácido nucleico conforme a la reivindicación 19, donde la secuencia de ácido nucleico está representada por SEC. ID N.º 4.

21. Una proteína codificada por la molécula de ácido nucleico conforme a la reivindicación 20.

15 22. Una molécula aislada de ácido nucleico con la secuencia de ácido nucleico SEC. ID N.º 4.

23. Una célula eucariota transgénica que consta de una molécula aislada de ácido nucleico conforme a la reivindicación 22, donde la célula eucariota no es una célula embrionaria humana.

20 24. Un método de producción en una célula eucariota de una proteína de fusión biespecífica recombinante funcional capaz de unirse a dos sitios de una proteína diana, que consta de las fases de:

a) transfección de la célula eucariota con una molécula de ácido nucleico recombinante que codifica la proteína de fusión biespecífica neutralizante de la reivindicación 1;

25 b) cultivo de la célula eucariota transfectada en condiciones que provoquen la producción de la proteína conforme a la reivindicación 1, y

30 c) recuperación de la proteína producida por la célula eucariota cultivada, donde la célula eucariota no es una célula embrionaria humana.

25. El método de la reivindicación 24, donde la célula eucariota es una célula mamífera.

35 26. El método de la reivindicación 24, donde la recuperación de la proteína comprende:

a) la identificación de la proteína por la presencia de una etiqueta molecular, y

40 b) la separación de la proteína con la etiqueta molecular así identificada a partir de las moléculas sin etiqueta, de forma que se recupera la proteína producida por la célula eucariota cultivada.

27. Un método *in vitro* para la inactivación de una proteína gp120, que comprende la fase de:

45 Contacto de la proteína gp120 con una proteína conforme a la reivindicación 4.

28. Un método *in vitro* para neutralizar un virus de la inmunodeficiencia humana, que comprende la fase de:

50 Contacto del virus de inmunodeficiencia humana con una proteína conforme a la reivindicación 4.

29. Un método *in vitro* para bloquear y evitar la unión de una proteína gp120 viral o recombinante a CD4 soluble o a CD4 linfocito, que comprende la fase de:

55 Contacto de la proteína gp120 con una proteína conforme a la reivindicación 4.

60 30. Una proteína conforme a la reivindicación 4 para su uso en un método de inhibición de la replicación o infección del virus del VIH en un sujeto, que comprende la administración al sujeto de una cantidad de dicha proteína suficiente para inhibir la replicación o infección del virus VIH.

31. Una composición que comprende la proteína conforme a la reivindicación 1.

65 32. Una composición farmacéutica que comprende la proteína conforme a la reivindicación 1, y un portador farmacéuticamente aceptable.

ES 2 313 886 T3

33. La proteína conforme a la reivindicación 30, donde la proteína se administra en forma de una composición farmacéutica.

5 34. Un kit, que comprende una dosis clínicamente eficaz de la proteína de fusión biespecífica neutralizante de la reivindicación 1, donde la proteína de fusión se une a una proteína diana, y es eficaz para tratar o evitar la infección por el virus de la inmunodeficiencia humana (VIH), y donde la proteína diana es una proteína de superficie del VIH.

10 35. El kit de la reivindicación 34, donde la proteína de fusión biespecífica neutralizante se suministra en forma de composición farmacéutica.

36. Una proteína conforme a la reivindicación 9, donde el anticuerpo monoclonal neutralizante se selecciona de un grupo que consta de 17b, 48b y CG10.

15

20

25

30

35

40

45

50

55

60

65

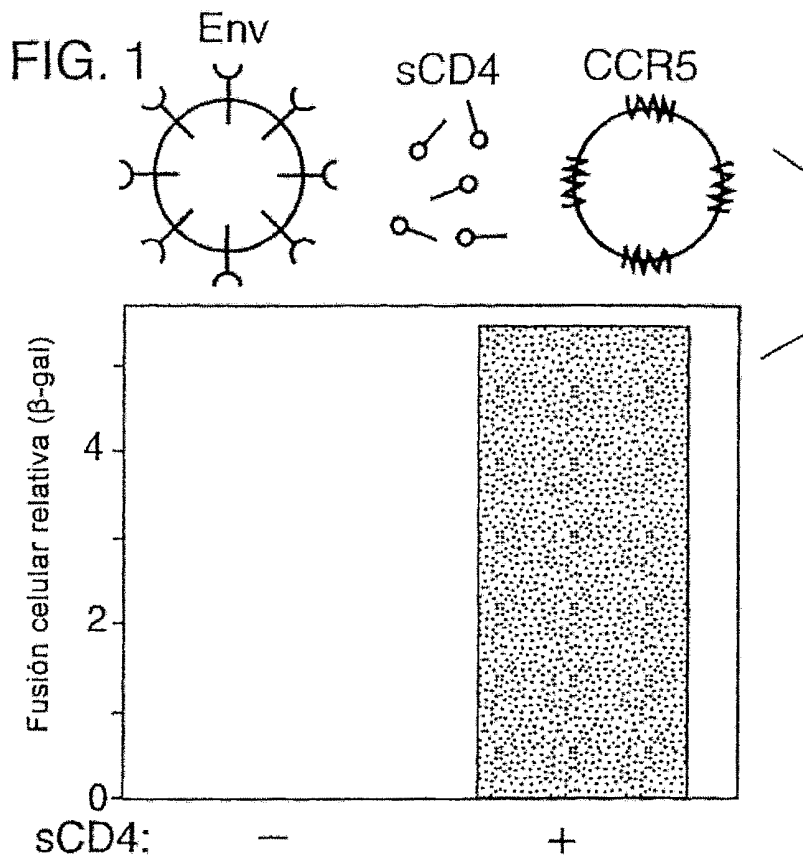


FIG. 2

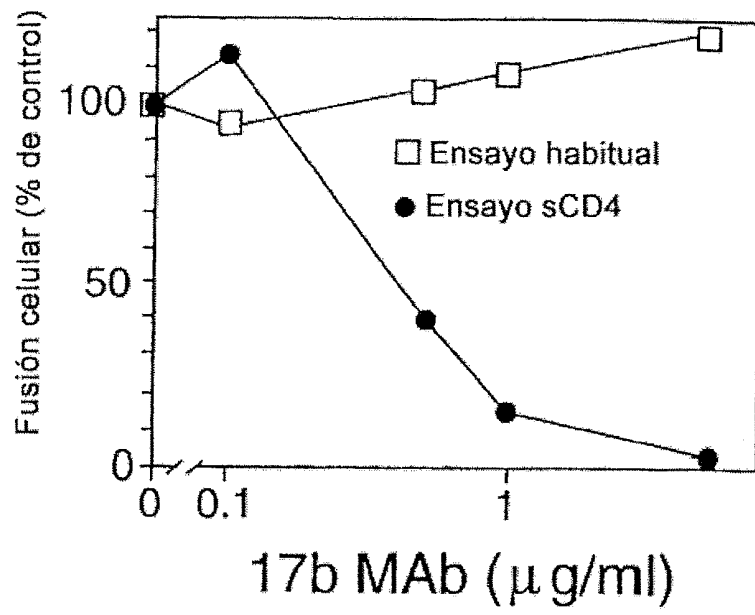
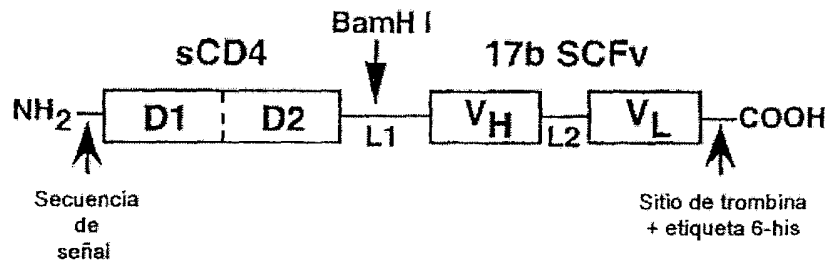
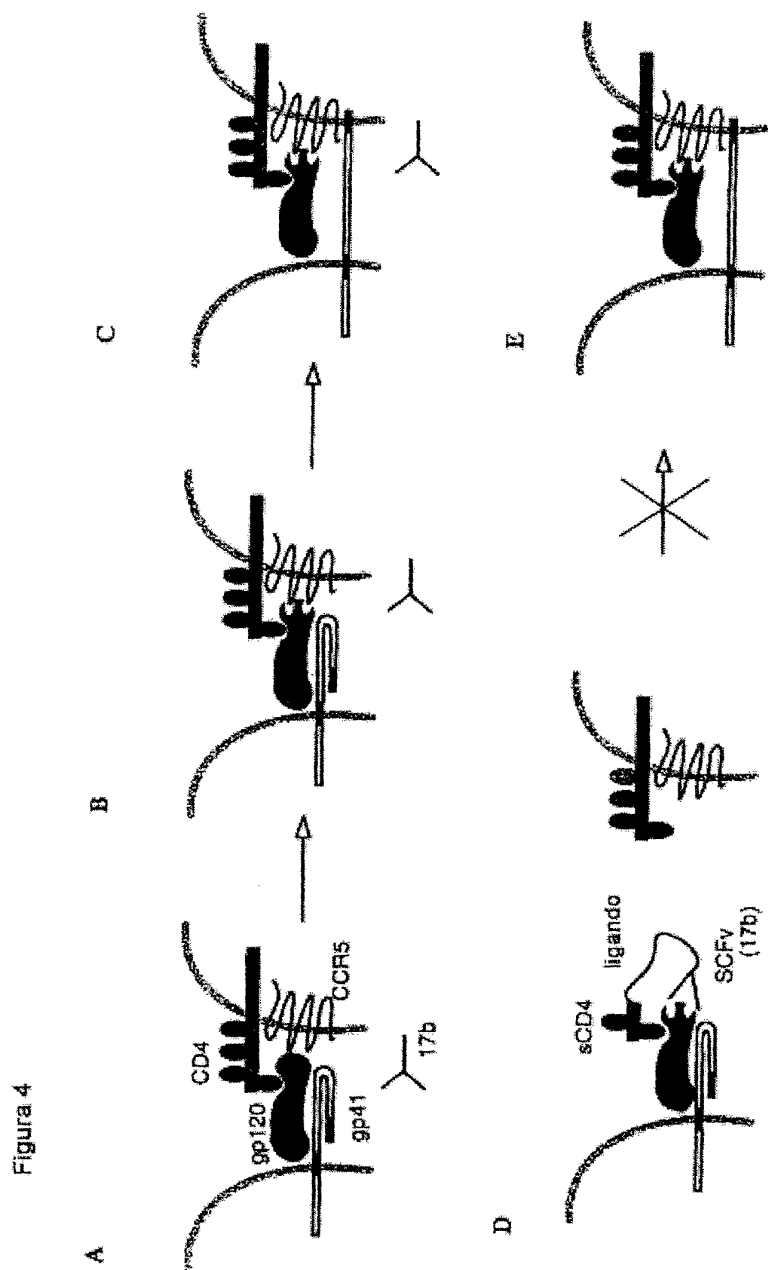


Figura 3





ES 2 313 886 T3

LISTA DE SECUENCIAS

<110> Berger, y col.

5 <120> PROTEÍNA QUIMÉRICA NOVEDOSA PARA PREVENCIÓN Y TRATAMIENTO DE INFECCIÓN POR VIH

10 <130> PCT de Berger 4239-54106

<140>

<141>

15 <150> 60/124.681

<151> 16-3-1999

20 <160> 11

<170> PatentIn version 2.1

<210> 1

25 <211> 5

<212> Proteína

<213> Secuencia Artificial

30 <220>

<223> Descripción de Secuencia Artificial: péptido de engarce

<400> 1

35

Gly Gly Gly Gly Ser
1 5

40 <210> 2

<211> 35

<212> Proteína

<213> Secuencia Artificial

45

<220>

<223> Descripción de Secuencia Artificial: engarce polipeptídico repetido siete veces

50 <400> 2

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
1 5 10 15

55

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
20 25 30

Gly Gly Ser
35

60 <210> 3

<211> 508

<212> Proteína

<213> Secuencia Artificial

65

<220>

<223> Descripción de Secuencia Artificial: CD4-scFv (17b)

ES 2 313 886 T3

<400> 3

Met Asn Arg Gly Val Pro Phe Arg His Leu Leu Leu Val Leu Gln Leu
 1 3 10 15

5 Ala Leu Leu Pro Ala Ala Thr Gln Gly Lys Lys Val Val Leu Gly Lys
 20 25 30

10 Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala Ser Gln Lys Lys Ser
 35 40 45

Ile Gln Phe His Trp Lys Asn Ser Asn Gln Ile Lys Ile Leu Gly Asn
 50 55 60

15 Gln Gly Ser Phe Leu Thr Lys Gly Pro Ser Lys Leu Asn Asp Arg Ala
 65 70 75 80

20 Asp Ser Arg Arg Ser Leu Trp Asp Gln Gly Asn Phe Pro Leu Ile Ile
 85 90 95

Lys Asn Leu Lys Ile Glu Asp Ser Asp Thr Tyr Ile Cys Glu Val Glu
 100 105 110

25 Asp Gln Lys Glu Glu Val Gln Leu Leu Val Phe Gly Leu Thr Ala Asn
 115 120 125

30 Ser Asp Thr His Leu Leu Gln Gly Gln Ser Leu Thr Leu Thr Leu Glu
 130 135 140

Ser Pro Pro Gly Ser Ser Pro Ser Val Gln Cys Arg Ser Pro Arg Gly
 145 150 155 160

35 Lys Asn Ile Gln Gly Gly Lys Thr Leu Ser Val Ser Gln Leu Glu Leu
 165 170 175

Gln Asp Ser Gly Thr Trp Thr Cys Thr Val Leu Gln Asn Gln Lys Lys
 180 185 190

40 Val Glu Phe Lys Ile Asp Ile Val Val Leu Ala Phe Gln Lys Ala Ser
 195 200 205

45 Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
 210 215 220

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
 225 230 235 240

50 Gly Gly Ser Gln Val Gln Leu Leu Glu Ser Gly Ala Glu Val Lys Lys
 245 250 255

55 Pro Gly Ser Ser Val Lys Val Ser Cys Lys Ala Ser Gly Asp Thr Phe
 260 265 270

Ile Arg Tyr Ser Phe Thr Trp Val Arg Gln Ala Pro Gly Gln Gly Leu
 275 280 285

60 Glu Trp Met Gly Arg Ile Ile Thr Ile Leu Asp Val Ala His Tyr Ala
 290 295 300

65 Pro His Leu Gln Gly Arg Val Thr Ile Thr Ala Asp Lys Ser Thr Ser
 305 310 315 320

Thr Val Tyr Leu Glu Leu Arg Asn Leu Arg Ser Asp Asp Thr Ala Val

ES 2 313 886 T3

				325						330				335			
	Tyr	Phe	Cys	Ala	Gly	Val	Tyr	Glu	Gly	Glu	Ala	Asp	Glu	Gly	Glu	Tyr	
				340					345					350			
5																	
	Asp	Asn	Asn	Gly	Phe	Leu	Lys	His	Trp	Gly	Gln	Gly	Thr	Leu	Val	Thr	
			355					360					365				
10																	
	Val	Thr	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	
		370					375						380				
15																	
	Gly	Ser	Glu	Leu	Glu	Leu	Thr	Gln	Ser	Pro	Ala	Thr	Leu	Ser	Val	Ser	
	385					390					395					400	
20																	
	Pro	Gly	Glu	Arg	Ala	Thr	Leu	Ser	Cys	Arg	Ala	Ser	Glu	Ser	Val	Ser	
				405						410					415		
25																	
	Ser	Asp	Leu	Ala	Trp	Tyr	Gln	Gln	Lys	Pro	Gly	Gln	Ala	Pro	Arg	Leu	
			420						425					430			
30																	
	Leu	Ile	Tyr	Gly	Ala	Ser	Thr	Arg	Ala	Thr	Gly	Val	Pro	Ala	Arg	Phe	
			435					440					445				
35																	
	Ser	Gly	Ser	Gly	Ser	Gly	Ala	Glu	Phe	Thr	Leu	Thr	Ile	Ser	Ser	Leu	
		450				455						460					
40																	
	Gln	Ser	Glu	Asp	Phe	Ala	Val	Tyr	Tyr	Cys	Gln	Gln	Tyr	Asn	Asn	Trp	
	465				470						475					480	
45																	
	Pro	Pro	Arg	Tyr	Thr	Phe	Gly	Gln	Gly	Thr	Arg	Leu	Glu	Ile	Lys	Leu	
					485					490					495		
50																	
	Val	Pro	Arg	Gly	Ser	Gly	His	His	His	His	His	His					
				500					505								

<210> 4

40 <211> 1440

<212> ADN

<213> Secuencia Artificial

45 <220>

<223> Descripción de Secuencia Artificial: CD4-scFv (17b)

<400> 4

50	atgaaccggg	gagtcacctt	taggcacttg	cttctgggtg	tgcaactggc	gctcctccca	60
	gcagccactc	agggaaagaa	agtgggtgctg	ggcaaaaaag	gggatacagt	ggaaactgacc	120
	tgtacagctt	cccagaagaa	gagcatacaa	ttccactgga	aaaactccaa	ccagataaag	180
	attctgggaa	atcagggctc	cttcttaact	aaaggtccat	ccaagctgaa	tgatcgcgct	240
55	gactcaagaa	gaagcctttg	ggaccaagga	aacttcccc	tgatcatcaa	gaatcttaag	300
	atagaagact	cagatactta	catctgtgaa	gtggaggacc	agaaggagga	ggtgcaattg	360
	ctagtgttcg	gattgactgc	caactctgac	acccacctgc	ttcaggggca	gagcctgacc	420
	ctgaccttgg	agagccccc	tggtagttag	ccctcagtgc	aatgtaggag	tccaaggggt	480
60	aaaaacatac	agggggggaa	gaccctctcc	gtgtctcage	tgagctcca	ggatagtggc	540
	acctggacat	gcactgtctt	gcagaaccag	aagaaggcgg	agttcaaaat	agacatcgtg	600
	gtgctagctt	tccagaaggc	ctcgggaggt	ggcggtagtg	ggggaggcgg	ttcaggcggg	660
	ggtggatccg	gtggcggagg	gtcgggcggg	ggtggaagcg	ggggtggcgg	ctccggaggc	720
	ggagggttcac	aggtgcagct	gctcgagctc	ggggctgagg	tgaagaagcc	tgggtcctcg	780
65	gtgaaggctc	cctgcaaggc	ctctggagac	accttcacaa	gatatagttt	tacctgggtg	840
	cgacagggcc	ctggacaagg	ccttgagtgg	atgggaagga	tcctcactat	ccttgatgta	900
	gcactactcg	caccgcacct	ccagggcaga	gtcacgatta	ccgcggacaa	gtccaagagc	960

ES 2 313 886 T3

```
ecagctctacc tggagctggg gaatctaaga tctgacgata cggccgtata tttctgtggg 1020
ggagtgtacc agggagagggc ggacgagggg gaatatgata ataatgggtt tctgaaacat 1080
tggggccagg gaaccctggg caaggtcacc tcagggtggg gtggctccgg aggtgggggg 1140
5 agcggtgggg ggggatctga actcaggtt acgcagttc cagccaccct gtctgtgctt 1200
ccaggggaaa gagccaccct ctctgcagg gccagtgaga gtgttagtag cgacttagcc 1260
tggtagcagc agaaaacctg ccaggctccc aggtcctca tatatgggtg atccaccagg 1320
gcdaccgggtg tcccagccag gttcagtggc agtgggtctg gggcagaatt cactctcacc 1380
atcagcagcc tgcagctcga agatcttgca gtttattact gtcagcagta caataactgg 1440
```

10 <210> 5

<211> 127

<212> ADN

15 <213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: oligonucleótido sintético

20 <400> 5

```
cctccggagg tggcggtagt gggggaggcg gttcagggcg aggtggatcc ggagggcgag 60
ggtcggggcg ggggtggaagc ggggggtggcg gctctggtgg cggaggtacc actagttaag 120
tgaagtag 127
```

25 <210> 6

<211> 39

<212> Proteína

30 <213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: péptido codificado por la SEQ ID N.º: 5

35 <400> 6

```
Ala Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
40 I 5 10 15
Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
20 25 30
Gly Gly Gly Gly Thr Thr Ser
45 35
```

<210> 7

<211> 103

50 <212> ADN

<213> Secuencia Artificial

<220>

55 <223> Descripción de Secuencia Artificial: cebador

<400> 7

```
ttttatggat ccggtggcgg agggctgggc ggggggggaa ggggggggtg cggctccgga 60
ggcggaggtt cacaggtgca gctgctcgag tctggggctg agg 103
```

<210> 8

<211> 32

65 <212> Proteína

<213> Secuencia Artificial

ES 2 313 886 T3

<220>

<223> Descripción de Secuencia Artificial: péptido codificado por SEQ ID N.º: 7

5 <400> 8

Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
1 5 10 15

10

Ser Gly Gly Gly Gly Ser Gln Val Gln Leu Leu Glu Ser Gly Ala Glu
20 25 30

<210> 9

<211> 65

15

<212> ADN

<213> Secuencia Artificial

<220>

20

<223> Descripción de Secuencia Artificial: cebador

<400> 9

25

taatttatcg atcacgtgac tagtcctagg cccgggtcaa tgatgatgat gatgatggcc 60
gctgc 65

<210> 10

<211> 8

30

<212> Proteína

<213> Secuencia Artificial

<220>

35

<223> Descripción de Secuencia Artificial: péptido codificado por SEQ ID N.º: 9

<400> 10

40

Ser Gly His His His His His His
1 5

<210> 11

45

<211> 131

<212> ADN

<213> Secuencia Artificial

50

<220>

<223> Descripción de Secuencia Artificial: oligonucleótido inverso

<400> 11

55

ctagctactc acttaactag tggtaacctcc gccacctgag ccgccacccc cgcttccacc 60
ccccgccgga cctccgctcc ccggatccac ctcggcctga aecgcctccc cactaccgcc 120
aactccggag g 131

60

65