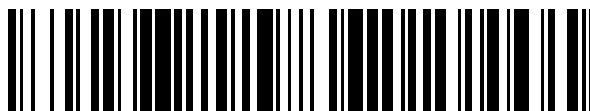


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 486 669**

51 Int. Cl.:

C12N 15/49 (2006.01)

C07K 14/16 (2006.01)

A61K 39/21 (2006.01)

A61K 39/295 (2006.01)

C12N 15/62 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **03.08.2005 E 10178939 (4)**

97 Fecha y número de publicación de la concesión europea: **11.06.2014 EP 2280073**

54 Título: **Vacunas para la prevención y tratamiento de infección por VIH**

30 Prioridad:

05.08.2004 GB 0417494

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

19.08.2014

73 Titular/es:

GLAXOSMITHKLINE BIOLOGICALS SA (100.0%)

Rue de l'Institut 89

1330 Rixensart, BE

72 Inventor/es:

ABRECHT, HELGE;

DELCHAMBRE, MARTINE;

MARCHAND, MARTINE;

MATHY, NATHALIE LOUISE;

PERMANNE, PHILIPPE JEAN GERVAIS GHISLAIN y

VOSS, GERALD HERMANN

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 486 669 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

Vacunas para la prevención y tratamiento de infección por VIH

Campo de la invención

5 La presente invención se refiere a construcciones de polipéptidos de VIH novedosas, y a composiciones farmacéuticas que las comprenden. Se describe también procedimientos para su fabricación y su uso en medicina. La invención también se refiere a polinucleótidos que codifican los polipéptidos. En particular, se describen en la presente invención proteínas de fusión que comprenden Nef de VIH-1 y Gag de VIH-1 o fragmentos de los mismos, y a polinucleótidos que los codifican. De forma más particular, se describen proteínas de fusión que comprenden proteínas Nef de VIH-1, Pol de VIH-1 y Gag de VIH-1 o fragmentos de los mismos y polinucleótidos que los codifican.

10 El VIH-1 es la causa primaria del síndrome de inmunodeficiencia adquirida (SIDA) que está catalogado como uno de los principales problemas de salud en el mundo. Hay una necesidad de una vacuna para la prevención y/o tratamiento de la infección por VIH.

Antecedente de la invención

15 El VIH-1 es un virus de ARN de la familia de los retrovirales. El genoma del VIH codifica al menos nueve proteínas que se dividen en tres clases: las proteínas estructurales principales Gag, Pol y Env, las proteínas regulatorias Tat y Rev, y las proteínas accesorias Vpu, Vpr, Vif y Nef. El genoma del VIH muestra la organización 5'LTR-gag-pol-env-LTR3' de todos los retrovirus.

20 La glicoproteína gp120 de envoltura del VIH es la proteína viral que se usa para la adhesión a la célula huésped. Esta adhesión está mediada por la unión de dos moléculas de superficie de las células T de ayuda y macrófagos, conocidos como CD4 y uno de los dos receptores de la quemoquina CCR-5 o CXCR-4. La proteína gp120 se expresa en primer lugar como una molécula precursora mayor (gp160), que luego se escinde post-translacionalmente para dar gp120 y gp41. La proteína gp120 está retenida sobre la superficie del virión mediante conexión a la molécula gp41, que está insertada en la membrana viral.

25 La proteína gp120 es la diana principal de los anticuerpos de neutralización, pero desafortunadamente la mayoría de las regiones inmunogénicas de las proteínas (bucle V3) son también las partes más variables de la proteína. Por tanto, se cree que el uso de gp120 (o su precursor gp160) como un antígeno para vacuna para educir anticuerpos de neutralización va a ser de uso limitado para una vacuna ampliamente protectora. La proteína gp120 también contiene epítomos que son reconocidos por linfocitos T citotóxicos (CTL). Estas células efectoras son capaces de eliminar las células infectadas con virus, y por tanto constituyen un segundo mecanismo inmunológico antiviral principal. Al contrario que las regiones diana de anticuerpos de neutralización algunos epítomos de CTL parecen ser conservados relativamente entre diferentes cepas de VIH. Por esta razón gp120 y gp160 pueden ser componentes antigénicos útiles en vacunas que ayudan a educir respuestas inmunes mediadas por células (particularmente CTL).

30 Proteínas de no envoltura de VIH-1 incluyen, por ejemplo, proteínas estructurales internas tales como los productos de genes gag y pol y otras proteínas no estructurales tales como Rev, Nef, Vif y Tat (Green y col., New England J. Med, 324, 5, 308 y siguientes (1991) y Bryant y col. (Ed. Pizzo), Pediatr. Infect. Dis. J., 11, 5, 390 y siguientes (1992).

La Nef del VIH es una proteína temprana, esto es, es expresada pronto en infección y en ausencia de proteína estructural.

40 El gen de Nef codifica pronto una proteína del VIH accesorias que ha demostrado poseer varias actividades. Por ejemplo, se sabe que la proteína Nef provoca la baja regulación de CD4, receptor del VIH, y moléculas de clase I MHC de la superficie celular, aunque está en debate la importancia biológica de estas funciones. De forma adicional, Nef interactúa con la ruta de señal de las células T e induce un estado activo, que en cambio puede promover expresión génica más eficiente. Algunos aislados del VIH presentan mutaciones en esta región, lo que hace que no codifiquen proteína funcional y se vean seriamente comprometidos en su replicación y patogénesis in vivo.

45 El gen de Gag se transla como una poliproteína precursora que se escinde mediante proteasa para dar productos que incluyen la proteína matriz (p17), la cápsida (p24), la nucleocápsida (p9), p6 y dos péptidos espaciales, p2 y p1.

50 El gen Gag da lugar a la proteína precursora Gag de 55 kilodalton (kD), también denominada p55, que se expresa a partir del ARNm viral no empalmado. Durante la translación el término N de p55 es miristoilado, activando su asociación con el aspecto citoplasmático de las membranas celulares. La poliproteína Gag asociada a la membrana recluta dos copias del ARN genómico viral con otras proteínas virales y celulares que activa la germinación de la partícula viral desde la superficie de una célula infectada. Tras la germinación se escinde la p55 mediante la proteasa codificada viralmente (un producto del gen de pol) durante el proceso de maduración viral en cuatro proteínas más pequeñas designadas MA (matriz [p17]), CA (cápsida [p24]), NC (nucleocápsida [p9]), y p6.

55 Además de las 3 proteínas Gag principales, todos los precursores de Gag contienen otras regiones distintas, que se escinden y permanecen en el virión como péptidos de varios tamaños. Estas proteínas tienen diferentes papeles, por ejemplo, la proteína p2 tiene un papel propuesto en la regulación de la actividad de la proteasa y contribuye a la cronología correcta del procesamiento proteolítico.

60 El polipéptido p17 (MA) se deriva del extremo miristoilado, de terminal N de la p55. La mayoría de las moléculas de MA permanecen adheridas a la superficie interna de la bicapa lipídica del virión, estabilizando la partícula. Se recluta un subconjunto de MA dentro de las capas más profundas del virión, donde comienza a formar parte del complejo

que acompaña al ADN viral al núcleo. Estas moléculas de MA facilitan el transporte nuclear del genoma viral debido a que es reconocida una señal cariofílica en MA por parte de la maquinaria de importación nuclear celular. Este fenómeno permite al VIH infectar las células que no se dividen, una propiedad inusual para un retrovirus.

5 La proteína p24 (CA) forma el núcleo cónico de las partículas virales. Se ha demostrado que la ciclofilina A interactúa con la región p24 de la p55 llevando a su incorporación a las partículas de VIH. La interacción entre Gag y ciclofilina A es esencial debido a que la disrupción de esta interacción por parte de la ciclosporina A inhibe la replicación viral.

10 La región NC de Gag es responsable del reconocimiento específico de la denominada señal de empaquetamiento del VIH. La señal de empaquetamiento consiste en cuatro estructuras tallo-bucle localizadas cerca del extremo 5' del ARN viral, y es suficiente para mediar la incorporación de un ARN heterólogo en los viriones de VIH-1. La NC se une a la señal de empaquetamiento a través de interacciones mediadas por dos motivos dedo de cinc. La NC también facilita la transcripción inversa.

15 La región polipeptídica p6 media las interacciones entre p55 Gag y la proteína accesoria Vpr, llevando a la incorporación de Vpr en viriones de ensamblado. La región p6 también contiene un denominado dominio tardío que se requiere para la liberación eficiente de viriones de germinación desde una célula infectada.

20 El gen de Pol codifica dos proteínas que contienen las dos actividades necesitadas por el virus en la infección precoz, la RT y la proteína integrasa necesarias para la integración del ADN viral en el ADN de la célula. El producto primario de Pol se escinde mediante la proteasa del virión para dar el péptido RT de terminal amino que contiene necesariamente actividades para la síntesis de ADN (actividad de ADN polimerasa dependiente de ARN y ADN así como también una función de ARNasa H) y proteína integrasa de terminal carboxi. La RT del VIH es un heterodímero de RT (p66) de longitud completa y un producto de escisión (p51) al que le falta el dominio de ARNasa H de terminal carboxi.

25 RT es una de las proteínas más altamente conservadas codificada por el genoma retroviral. Dos actividades principales de RT son la ADN Pol y la ribonucleasa H. La actividad de ADN Pol de RT usa ARN y ADN como plantados de forma intercambiable y como todas las ADN polimerasas conocidas es incapaz de iniciar la síntesis de ADN de nuevo, pero requiere una molécula pre-existente para servir como un cebador (ARN).

30 La actividad de la ARNasa H inherente en todas las proteínas RT juega el papel esencial, ya en la replicación, de eliminación del genoma de ARN cuando tiene lugar la síntesis de ADN. Degrada de forma selectiva el ARN de todas las moléculas híbridas ARN-ADN. Estructuralmente la polimerasa y la ribo H ocupan dominios no solapados, separados con la Pol cubriendo las dos terceras partes de los amino de la Pol.

La subunidad catalítica p66 está plegada en 5 subdominios distintos. El terminal amino 23 de estos tiene la parte con actividad de RT. El terminal carboxi de estos es el dominio ARNasa H.

35 El documento WO 03/025003 describe construcciones de ADN que codifican p17/24 Gag, Nef y RT del VIH-1, en los que las secuencias de ADN pueden estar optimizadas con codón para parecerse a genes humanos altamente expresados. Estas construcciones son útiles en vacunas de ADN.

40 Se han sugerido proteínas de fusión que contienen múltiples antígenos de VIH como candidatos a vacunas para VIH, por ejemplo, la fusión Nef-Tat como se describe en el documento WO 99/16884. Sin embargo, las proteínas de fusión no son de producción directa; puede haber dificultades en la expresión de las mismas debido a que no se corresponden con proteínas nativas. Puede haber dificultades a nivel de transcripción, o posteriormente. Estas tampoco son de formulación directa en una composición farmacéuticamente aceptable. De forma notable, la mayoría de los enfoques de vacunas para VIH que implican múltiples antígenos condensados conjuntamente, son enfoques de ADN o vector vivo más que proteínas de fusión de polipéptidos.

Sumario de la invención

45 La presente invención proporciona construcciones nuevas para uso en vacunas para la profilaxis y tratamiento de infecciones por VIH y SIDA.

En un aspecto la invención proporciona una composición que comprende:

50 a) un polipéptido que comprende p17 Gag o un fragmento o derivado inmunogénico del mismo, p51 RT o un fragmento o derivado inmunogénico de la misma, Nef o un fragmento o derivado inmunogénico de la misma y p24 Gag o un fragmento o derivado inmunogénico del mismo, en los que hay al menos un antígeno de VIH o fragmento inmunogénico entre p17 Gag y p24 Gag; y

b) una proteína env de VIH o fragmento inmunogénico o derivado de la misma,

en la que los fragmentos o derivados inmunogénicos son capaces de aumentar una respuesta inmune frente al antígeno nativo.

55 En un aspecto adicional la invención proporciona una composición que comprende un polinucleótido o polinucleótidos que codifican:

a) un polipéptido que comprende p17 Gag o un fragmento o derivado inmunogénico del mismo, p51 RT o un fragmento o derivado inmunogénico de la misma, Nef o un fragmento o derivado inmunogénico de la misma y p24 Gag o un fragmento o derivado inmunogénico del mismo, en los que hay al menos un antígeno de VIH o fragmento inmunogénico entre p17 Gag y p24 Gag; y

b) una proteína env de VIH o fragmento inmunogénico o derivado de la misma,

en la que los fragmentos o derivados inmunogénicos son capaces de aumentar una respuesta inmune frente al antígeno nativo.

5 También se proporciona una composición farmacéutica que comprende una composición como se describió anteriormente junto con un vehículo a adyuvante farmacéuticamente aceptable.

En las construcciones y composiciones de acuerdo con la invención como se describen en la presente invención, la Nef es preferiblemente una Nef de longitud completa.

En las construcciones de acuerdo con la invención la p17 Gag y p24 Gag son preferiblemente p17 y p24 de longitud completa respectivamente.

10 En un caso el polipéptido comprende tanto p17 como p24 Gag o fragmentos inmunogénicos de los mismos. En una construcción de este tipo el componente p24 Gag y el componente p17 Gag están separados por al menos un antígeno de VIH adicional o fragmento inmunogénico, tal como Nef y/o RT o fragmentos inmunogénicos o derivados de las mismas.

15 De forma alternativa p17 o p24 Gag se pueden proporcionar por separado. Por tanto también se describe una composición que comprende (i) un polipéptido que comprende Nef o un fragmento o derivado inmunogénico de la misma y p17 Gag o un fragmento o derivado inmunogénico del mismo, y (ii) p24 Gag o un fragmento o derivado inmunogénico del mismo; o (i) un polipéptido que comprende Nef o un fragmento o derivado inmunogénico de la misma y p24 Gag o un fragmento o derivado inmunogénico del mismo, y (ii) p17 Gag o un fragmento o derivado inmunogénico del mismo.

20 En otro caso la construcción de polipéptido comprende además Pol o un derivado de Pol tal como RT o un fragmento o derivado inmunogénico de la misma. Fragmentos particulares de RT que son fragmentos en los que la RT está truncada en el término C, preferiblemente de modo que les falta el dominio ARNasa H de terminal carboxi. Un fragmento de este tipo al que le falta el dominio ARNasa H de terminal carboxi es el fragmento p51 descrito en la presente invención.

25 Preferiblemente la RT o fragmento inmunogénico en las proteínas de fusión descritas en la presente invención es p66 RT o p51 RT.

30 El componente RT de la proteína de fusión o composición comprende de forma opcional una mutación en la posición 592, o mutación equivalente en cepas distintas a HXB2, de modo que la metionina se elimina mediante mutación en otro residuo, por ejemplo, lisina. El fin de esta mutación es eliminar un sitio que sirve como un sitio de inicio interno en los sistemas de expresión procarióticos.

El componente RT por tanto, o de forma alternativa, comprende una mutación para eliminar la actividad del enzima (transcriptasa inversa). Por tanto K231 puede estar presente en lugar de W.

35 En las proteínas de fusión que comprenden p24 y RT, puede ser preferible que la p24 preceda a la RT en la construcción debido a que cuando los antígenos se expresan sólo en *E. coli* se observa mejor expresión de p24 que de RT.

Construcciones preferidas incluyen las siguientes:

1. p24 – RT – Nef – p17
2. p24 – RT* – Nef – p17
3. p24 – p51RT – Nef – p17
- 40 4. p24 – p51RT* - Nef – p17
5. p17-p51RT-Nef
6. p17-p51RT*-Nef
7. Nef-p17
8. Nef-p17 con conector
- 45 9. p17-Nef
10. p17 – Nef con conector

* representa mutación de metionina₅₉₂ en RT a lisina

50 El conector incluido en las construcciones enumeradas anteriormente puede ser cualquier secuencia de aminoácidos corta para disminuir las interacciones potenciales entre los dos participantes de la fusión que los conecta. El conector puede ser, por ejemplo, de 4 a 10 aminoácidos de longitud. Por ejemplo, puede ser una secuencia de 6 aminoácidos tal como la secuencia GSGGGP descrita en la presente invención en los ejemplos.

En otro caso se proporciona una proteína de fusión de antígenos de VIH que comprende al menos cuatro antígenos de VIH o fragmentos inmunogénicos, en los que cuatro antígenos o fragmentos son o se derivan de Nef, Pol y Gag. Preferiblemente Gag está presente como dos componentes separados que están separados por al menos otro antígeno en la fusión. Preferiblemente la Nef es Nef de longitud completa. Preferiblemente la Pol es p66 o p51RT. Preferiblemente la Gag es p17 y p24 Gag. Otros rasgos preferidos y propiedades de los componentes del antígeno de la fusión son como se describen en la presente invención.

Aspectos preferidos son las cuatro fusiones de componentes que ya se enumeraron anteriormente:

1. p24 – RT – Nef – p17
2. p24 – RT* – Nef – p17
3. p24 – p51RT – Nef – p17
4. p24 – p51RT* - Nef – p17

El término “derivado de” o “derivado” respecto a los antígenos de VIH incluidos en la invención significa que los antígenos pueden haber sido alterados de modo limitado en comparación con sus contrapartes originales. Esto incluye mutaciones puntuales que pueden cambiar las propiedades de la proteína, por ejemplo, mediante la mejora de la expresión en sistemas procarióticos o eliminación de actividad no deseable incluyendo actividad del enzima no deseable. Las mutaciones puntuales descritas en la presente invención para la RT se diseñan para alcanzar estos hechos. Sin embargo, los antígenos deben permanecer suficientemente similares a los antígenos nativos de modo que mantengan las propiedades antigénicas deseables en una vacuna y así permanecer capaces de dar una respuesta inmune frente al antígeno nativo. De lugar o no un derivado determinado una respuesta inmune de este tipo se puede medir con un ensayo inmunológico adecuado tal como ELISA (para respuestas de anticuerpos) o citometría de flujo usando tinto adecuado para marcadores celulares y citoquinas (para respuestas celulares).

Las construcciones polipeptídicas de antígenos de VIH de acuerdo con la invención son capaces de ser expresadas en sistemas *in vitro* incluyendo sistemas procarióticos tales como *E. coli*. De forma ventajosa, estas se pueden purificar mediante procedimientos de purificación convencionales.

Las fusiones descritas en la presente invención son preferiblemente solubles cuando se expresan en un sistema de expresión seleccionado, esto es, están presentes en una cantidad sustancial en el sobrenadante de un extracto bruto del sistema de expresión. La presencia de la proteína de fusión en el extracto bruto se puede medir mediante medios convencionales tales como elución en un gel SDS, tinto con Coomassie y comprobación de la banda apropiada mediante medida densitométrica. Las proteínas de fusión de acuerdo con la invención son preferiblemente solubles al menos en el 50%, más preferiblemente solubles al menos en el 70%, lo más preferiblemente solubles en el 90% o más medido mediante las técnicas descritas en la presente invención en los ejemplos. Se conocen técnicas para mejorar la solubilidad de las proteínas expresadas recombinantemente, por ejemplo, en sistemas de expresión procarióticos se mejora la solubilidad mediante disminución de la temperatura a la que se induce la expresión génica.

Las proteínas de fusión descritas en la presente invención se pueden purificar. En particular se pueden purificar mientras permanecen solubles o significativamente solubles.

Fragmentos inmunogénicos como se describen en la presente invención contendrán al menos un epítipo del antígeno y mostrará antigenicidad por VIH y son capaces de dar una respuesta inmune cuando estaban presentes en una construcción adecuada, tal como por ejemplo cuando estaban condensados con otros antígenos de VIH o se presentaba en un vehículo, siendo dirigida la respuesta inmune contra el antígeno nativo. De forma típica los fragmentos inmunogénicos contienen al menos 20, preferiblemente 50, más preferiblemente 100 aminoácidos contiguos al antígeno de VIH.

También se describen en la presente invención polinucleótidos que codifican los polipéptidos descritos.

Se pueden usar polinucleótidos descritos en la presente invención como vacunas de polinucleótidos. Los polinucleótidos pueden estar presentes dentro de cualquiera de una variedad de sistemas de liberación conocidos por los especialistas en la técnica, incluyendo sistemas de expresión de ácido nucleico tales como ADN plásmido, sistemas de expresión bacterianos y virales. Son bien conocidas en la técnica numerosas técnicas de liberación de genes, tales como las descritas por Rolland, Crit. Rev. Therap. Drug Carrier Systems 15:143-198, 1998 y referencias citadas en ese documento. Sistemas de expresión de ácido nucleico apropiados contienen las secuencias de ADN necesarias para la expresión en el paciente (tal como un promotor adecuado y señal de terminación). Cuando el sistema de expresión es un microorganismo vivo recombinante, tal como un virus o bacteria, el gen de interés se puede insertar dentro del genoma del virus recombinante vivo o bacteria. La inoculación y la infección *in vivo* con este vector vivo llevará a la expresión *in vivo* del antígeno e inducción de respuestas inmunes. Virus y bacterias usadas para este fin son, por ejemplo: poxvirus (por ejemplo, vaccinia, fowlpox, canaripox, poxvirus modificados, por ejemplo, virus Ankara modificado (MVA)), alfavirus (virus Sindbis, virus Semliki Forest, virus de la encefalitis equina venezolana), flavivirus (virus de la fiebre amarilla, virus del dengue, virus de la encefalitis japonesa), adenovirus, virus adeno-asociados, picornavirus (poliovirus, rinovirus), virus de herpes (virus varicella zoster, etc), morbilivirus (e.g. rubeola), Listeria, Salmonella, Shigella, Neisseria, BCG. Estos virus y bacterias pueden ser virulentos, o ser atenuados de varias formas con el fin de obtener vacunas vivas. Tales vacunas vivas también forman parte de la invención.

Un vector de rubéola preferido para uso como un vector vivo de acuerdo con la invención es la cepa de Schwartz o una cepa derivada de la misma.

Un adenovirus preferido para uso como un vector vivo es un adenovirus humano sero-prevalente inferior tal como Ad5 o Ad35 o un adenovirus de origen no humano tal como un adenovirus de primate no humano tal como un adenovirus de simio. Tal adenovirus humanos sero-prevalentes inferiores o similares tendrán menos de 60, de forma típica menos del 50%, de sero-prevalencia en la población. Preferiblemente, los vectores son defectivos en la replicación. De forma típica estos virus contienen una delección en E1 y se pueden criar en líneas celulares que se transforman con un gen E1. Los adenovirus de simio preferidos son virus aislados de chimpancés. En particular C68 (también conocido como Pan9) (véase la patente de Estados Unidos nº 6083716) y se prefieren para uso en la presente invención Pan 5, 6 y Pan 7 (documento WO 03/046124). Estos vectores se pueden manipular para insertar un polinucleótido heterólogo de acuerdo con la invención de modo que se pueden expresar los polipéptidos de acuerdo con la invención. El uso, formulación y fabricación de tales vectores adenovirales recombinantes se describe detalladamente en el documento WO 03/046142.

Por tanto, se puede proporcionar Nef, p17 y p24 Gag y RT de una vacuna preferida de acuerdo con la invención en forma de un polinucleótido que codifica el polipéptido deseado.

Se pueden usar los polinucleótidos de acuerdo con la invención para expresar los polipéptidos codificados en un sistema de expresión seleccionado. Al menos uno de los antígenos de VIH, por ejemplo, la RT, se puede codificar mediante una secuencia optimizada con codón en el polinucleótido, esto quiere decir que la secuencia ha sido optimizada para la expresión en un sistema de expresión recombinante seleccionado tal como *E. coli*.

En otro caso se proporciona un polipéptido p51 RT o derivado del mismo o un polinucleótido que lo codifica, preferiblemente optimizado con codón para la expresión en un sistema de expresión adecuado, en particular un sistema procariótico tal como *E. coli*.

El polipéptido p51 RT o polinucleótido se puede usar sólo, o en combinación con una construcción de polipéptido o polinucleótido de acuerdo con la invención. Por tanto, en un aspecto más la invención proporciona una composición que comprende (i) un polipéptido que comprende Nef o un fragmento que contiene epítipo de Nef y p17 Gag y/o p24 Gag, en los que cuando ambos p17 y p24 Gag están presentes hay al menos un antígeno de VIH o fragmento inmunogénico entre ellos y (ii) un polipéptido p51 RT. También se describen polinucleótidos que codifican estos.

De acuerdo con este caso (i) se puede seleccionar por ejemplo de:

1. Nef - p17
2. Nef - p17 con conector
3. p17 - Nef
4. p17 - Nef con conector

Preferiblemente Nef es Nef de longitud completa. Preferiblemente p17 es p17 de longitud completa.

Los polipéptidos y polinucleótidos descritos en la presente invención se pueden combinar con otros antígenos o polinucleótidos que codifican otros antígenos. En particular, esto puede incluir proteínas env del VIH o fragmentos o derivados de las mismas. Las formas preferidas de env son gp120, gp140 y gp160. La env puede ser, por ejemplo, la proteína de envoltura descrita en el documento WO 00/07631 de un clon de envoltura de clado B de VIH-1 conocido como R2, o un fragmento o derivado del mismo. Por tanto, la invención proporciona además una composición que comprende cualquiera de los polipéptidos o composiciones de polipéptido de acuerdo con la invención, junto con una proteína env del VIH o fragmento o derivado de la misma. De forma similar la invención proporciona una composición que comprende un polinucleótido o polinucleótidos que codifican un polipéptido o polipéptidos de acuerdo con la invención y un polinucleótido que codifica una proteína env del VIH o fragmento o derivado de la misma.

También se describen procedimientos de preparación de los polipéptidos descritos en la presente invención, tal procedimiento comprende la expresión de un polinucleótido que codifica el polipéptido en un sistema de expresión adecuado, en particular un sistema procariótico tal como *E. coli* y reconversión del polipéptido expresado. Preferiblemente se induce la expresión a una temperatura baja, que es una temperatura por debajo de 37^o, para promover la solubilidad del polipéptido.

También se describe un procedimiento para la purificación de un polipéptido como se describe en la presente invención, tal procedimiento comprende:

- i). proporcionar una composición que comprende el polipéptido no purificado;
- ii). someter la composición a al menos dos etapas cromatográficas;
- iii). de forma opcional carboxiamidación del polipéptido;
- iv) llevar a cabo una etapa de intercambio en tampón para proporcionar la proteína en un tampón adecuado para una formulación farmacéutica.

La carboxiamidación se puede llevar a cabo entre las dos etapas cromatográficas. La etapa de carboxiamidación se puede llevar a cabo usando yodoacetamida.

En un ejemplo, el procedimiento no usa más de dos etapas cromatográficas.

También se describen composiciones farmacéuticas y composiciones inmunogénicas y vacunas que comprenden los polipéptidos y polinucleótidos de acuerdo con la invención, en combinación con un adyuvante o vehículo farmacéuticamente aceptable.

5 Las vacunas descritas en la presente invención se pueden usar para inmunización profiláctica o terapéutica contra VIH.

También se describe el uso de los polipéptidos y composiciones de polipéptido y los polinucleótidos y composiciones de polinucleótido como se describen en la presente invención, en la fabricación de una vacuna para inmunización profiláctica o terapéutica contra VIH.

10 La vacuna descrita en la presente invención contendrá una cantidad inmunoprotectora o inmunoterapéutica del polipéptido y/o antígenos de polinucleótido y se puede preparar mediante técnicas convencionales.

15 La preparación de vacuna se describe en general en *New Trends and Developments in Vaccines*, editado por Voller y col., University Park Press, Baltimore, Maryland, EEUU 1978. La encapsulación dentro de liposomas se describe, por ejemplo, por parte de Fullerton, patente de Estados Unidos 4.235.877. La conjugación de proteínas en macromoléculas se describe, por ejemplo, por parte de Likhite, patente de Estados Unidos 4.372.945 y por parte de Armor y col., patente de Estados Unidos 4.474.757.

20 La cantidad de proteína en la dosis de vacuna se selecciona como una cantidad que induce una respuesta inmunoprotectora sin efectos secundarios adversos significativos en vacunas típicas. Tal cantidad variará dependiendo de qué inmunógeno específico se use y del régimen de vacunación que se seleccione. En general se espera que cada dosis comprenda de 1 a 1000 µg de cada proteína, preferiblemente de 2 a 200 µg, lo más preferiblemente de 4 a 40 µg de la fusión de polipéptido. Se puede averiguar una cantidad óptima para una vacuna determinada mediante estudios convencionales que incluyen la observación de títulos de anticuerpo y otras respuestas inmunes en sujetos. Tras una vacunación inicial, los sujetos pueden recibir un refuerzo en aproximadamente 4 semanas, y una segunda inmunización de refuerzo a continuación.

25 Las proteínas de la presente invención están preferiblemente adyuvadas en la formulación de vacuna de la invención. Se describen adyuvantes en general en *Vaccine Design – the Subunit and Adjuvant Approach*, editado por Powell and Newman, Plenum Press, Nueva York, 1995.

Adyuvantes adecuados incluyen una sal de aluminio tal como hidróxido de aluminio o fosfato de aluminio, pero también pueden ser una sal de calcio, hierro o cinc, o pueden ser una suspensión insoluble de tirosina acilada, o azúcares acilados, polisacáridos derivatizados catiónicamente o aniónicamente, o polifosfazenos.

30 En la formulación de la invención se prefiere que la composición adyuvante incluya una respuesta Th1 preferencial. Sin embargo, se entenderá que no se excluyan otras respuestas, incluyendo otras respuestas humorales.

35 Se genera una respuesta inmune a un antígeno a través de la interacción del antígeno con las células del sistema inmune. La respuesta inmune resultante se puede distinguir ampliamente en dos categorías extremas, siendo respuestas inmunes humorales o mediadas por células (caracterizadas tradicionalmente por mecanismos efectores de anticuerpo y celulares de protección respectivamente). Estas categorías de respuesta se han denominado respuestas de tipo Th1 (respuesta mediada por célula), y respuestas inmunes de tipo Th2 (respuestas humorales).

40 Las respuestas inmunes de tipo Th1 extremas se pueden caracterizar por la generación de linfocitos T citotóxicos restringidos por el haplotipo, específicos del antígeno, y respuestas de células asesinas naturales. En ratones, las respuestas de tipo Th1 se caracterizan frecuentemente por la generación de anticuerpos del subtipo IgG2a, mientras que en el humano estas corresponden a anticuerpos de tipo IgG1. Las respuestas inmunes de tipo Th2 se caracterizan por la generación de un amplio intervalo de isotipos de inmunoglobulina incluyendo en ratones IgG1, IgA e IgM.

45 Se puede considerar que la fuerza impulsora detrás del desarrollo de estos dos tipos de respuestas inmunes son las citoquinas, un número de mensajeros proteicos identificados que sirven para ayudar a las células del sistema inmunológico y dirigir la respuesta inmune eventual a una respuesta Th1 o Th2. Por tanto, niveles elevados de citoquinas de tipo Th1 tienden a favorecer la inducción de respuestas inmunes mediadas por células respecto al antígeno administrado, mientras que niveles elevados de citoquinas de tipo Th2 tienden a favorecer la inducción de respuestas inmunes humorales respecto al antígeno.

50 Es importante recordar que la distinción de respuestas inmunes de tipo Th1 y Th2 no es absoluta. En realidad un individuo soportará una respuesta inmune que se describe que es predominantemente Th1 o predominantemente Th2. Sin embargo, es frecuentemente conveniente considerar las familias de citoquinas en términos de lo descrito en clones de células T CD4+ve murinas por parte de Mosmann y Coffman (*Mosmann, T.R. and Coffman, R.L. (1989) TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties. Annual Review of Immunology, 7, páginas 145 a 173*). Tradicionalmente las respuestas de tipo Th1 están asociadas con la producción de las citoquinas INF-γ e IL-2 por parte de linfocitos T. Otras citoquinas frecuentemente asociadas directamente con la inducción de respuestas inmunes de tipo Th1 no son producidas por células T, tales como IL-12. Por el contrario las respuestas de tipo Th2 están asociadas con la secreción de IL-4, IL-5, IL-6, IL-10 y factor de necrosis tumoral β (TNF-β).

60 Se sabe que ciertos adyuvantes de vacuna son particularmente adecuados para la estimulación de respuestas de citoquina de tipo Th1 o Th2. Tradicionalmente los mejores indicadores del equilibrio Th1:Th2 de la respuesta inmune tras una vacunación o infección incluye la medida directa de la producción de citoquinas Th1 o Th2 por parte de linfocitos T *in vitro* tras reestimulación con antígeno, y/o la medida de la relación IgG1:IgG2a de respuestas de anticuerpo específicas de antígeno.

Por tanto, un adyuvante de tipo Th1 es uno que estimula poblaciones de células T aisladas para producir altos niveles de citoquinas de tipo Th1 cuando se re-estimulan con antígeno *in vitro*, e induce respuestas de inmunoglobulina específicas del antígeno asociadas con isotipo de tipo Th1.

5 Inmunoestimulantes de tipo Th1 preferidos que se pueden formular para producir adyuvantes adecuados para uso en la presente invención incluyen los siguientes sin restringirse a estos. Monofosforil lípido A, en particular el monofosforil lípido A 3-de-O-acilado (3D-MPL), es un inmunoestimulante de tipo Th1 preferido para uso en la invención. 3D-MPL es un adyuvante bien conocido fabricado por Ribí Immunochem, Montana. Químicamente es suministrado frecuentemente como una mezcla de monofosforil lípido A 3-de-O-acilado con cadenas aciladas en 4, 5
10 ó 6. Se puede purificar y preparar mediante los procedimientos dados a conocer en el documento GB 2122204B, tal referencia también describe la preparación de difosforil lípido A, y variantes 3-O-deaciladas del mismo. Se han descrito otros lipopolisacáridos purificados y sintéticos (documentos US 6.005.099 y EP 0729473 B1; Hilgers *y col.*, 1986, *Int.Arch.Allergy.Immunol.*, 79(4):392-6; Hilgers *y col.*, 1987, *Immunology*, 60(1):141-6; y EP 0549074 B1). Una forma preferida de 3D-MPL es en la forma de una formulación particulada que tiene un tamaño de partícula pequeño inferior a 0,2 µm de diámetro, y su procedimiento de fabricación se describe en el documento EP 0689454.

15 Las saponinas son también inmunoestimulantes Th1 preferidos de acuerdo con la invención. Las saponinas son adyuvantes bien conocidos y se dan a conocer en: Lacaille-Dubois, M y Wagner H. (1996. A review of the biological and pharmacological activities of saponins. *Phytomedicine* volumen 2, páginas 363-386). Por ejemplo, se describen Quil A (derivado de la corteza del árbol sudamericano Quillaja Saponaria Molina), y fracciones del mismo, en el documento US 5.057.540 y en "Saponins as vaccine adjuvants", Kensil, C. R., *Crit Rev Ther Drug Carrier Syst*, 1996, 12 (1-2):1-55; y en el documento EP 0362279 B1. Se han descrito saponinas hemolíticas QS21 y QS17 (fracciones de Quil A purificadas con HPLC) como potentes adyuvantes sistémicos, y se describe el procedimiento de su producción en la patente de Estados Unidos número 5.057.540 y en el documento EP 0362279 B1. También se describe en estas referencias el uso de QS7 (una fracción no hemolítica de Quil-A) que actúa como un potente adyuvante para vacunas sistémicas. El uso de QS21 se describe además por parte de Kensil *y col.* (1991. *J. Immunology* volumen 146, 431-437). Se conocen también combinaciones de QS21 y polisorbato o ciclodextrina (documento WO 99/10008). Los sistemas adyuvantes particulados que comprenden fracciones de QuilA, tales como QS21 y QS7 se describen en los documentos WO 96/33739 y WO 96/11711. Un sistema de este tipo es conocido como un Iscorn y puede contener una o más saponinas.

30 Otro inmunoestimulante preferido es un oligonucleótido inmunoestimulador que contiene dinucleótidos CpG no metilados ("CpG"). CpG es una abreviatura de motivos de dinucleótido de citosin-guanosina presentes en ADN. CpG es conocido en la técnica como un adyuvante cuando se administra tanto por vía sistémica como mucosal (documentos WO 96/02555, EP 468520, Davis *y col.*, *J.Immunol*, 1998, 160(2):870-876; McCluskie y Davis, *J.Immunol.*, 1998, 161(9):4463-6). Históricamente se observó que la fracción de ADN de BCG podría ejercer un efecto anti-tumor. En otros estudios, se demostró que oligonucleótidos sintéticos derivados de secuencias de genes de BCG son capaces de inducir efectos inmunoestimuladores (tanto *in vitro* como *in vivo*). Los autores de estos estudios concluyeron que ciertas secuencias palindrómicas, incluyendo un motivo CG central, portaban esta actividad. El papel central del motivo CG en la inmunoestimulación se elucidaba más tarde en una publicación de Krieg, *Nature* 374, páginas 546 1995. Análisis detallados han mostrado que el motivo CG tiene que estar en un cierto contexto de secuencia, y que esas secuencias son comunes en ADN bacteriano pero que son raras en ADN de vertebrados. La secuencia inmunoestimuladora es frecuentemente: Purina, Purina, C, G, pirimidina, pirimidina; en la que el motivo CG no está metilado, pero se conocen otras secuencias CpG no metiladas por ser inmunoestimuladoras y se pueden usar en la presente invención.

45 En ciertas combinaciones de los seis nucleótidos está presente una secuencia palindrómica. Pueden estar presentes en el mismo oligonucleótido varios de estos motivos, bien como repeticiones de un motivo o una combinación de distintos motivos. La presencia de una o varias de estas secuencias inmunoestimuladoras que contienen oligonucleótidos puede activar distintos subconjuntos inmunes, incluyendo células asesinas naturales (que producen interferona γ y tienen actividad citolítica) y macrófagos (Wooldrige *y col.* volumen 89 (número 8), 1977). Otras secuencias que contienen CpG no metilado que no tienen esta secuencia de consenso se ha demostrado también ahora que son inmunomoduladoras.

50 CpG cuando se formula en vacunas, se administra por lo general en solución libre junto con antígeno libre (documento WO 96/02555; McCluskie y Davis, véase anteriormente) o conjuntado covalentemente con un antígeno (documento WO 98/16247), o formularse con un vehículo tal como hidróxido de aluminio ((antígeno de superficie de la hepatitis) Davis *y col.* véase anteriormente; Brazolot-Millan *y col.*, *Proc.Natl.Acad.Sci.*, EEUU, 1998, 95(26), 15553-8).

55 Tales inmunoestimulantes como se describen anteriormente, se pueden formular junto con vehículos, tales como por ejemplo liposomas, emulsiones aceite en agua y/o sales metálicas, incluyendo sales de aluminio (tal como hidróxido de aluminio). Por ejemplo, se puede formular 3D-MPL con hidróxido de aluminio (documento EP 0689454) o en emulsiones aceite en agua (documento WO 95/17210); se puede formular QS21 de forma ventajosa con liposomas que contienen colesterol (documento WO 96/33739), emulsiones aceite en agua (documento WO 95/17210) o alumbre (documento WO 98/15287); CpG se puede formular con alumbre (Davis *y col.* véase anteriormente; Brazolot-Millan véase anteriormente) o con otros vehículos catiónicos.

60 Se prefieren también combinaciones de inmunoestimulantes, en particular una combinación de un monofosforil lípido A y un derivado de saponina (documentos WO 94/00153; WO 95/17210; WO 96/33739; WO 98/56414; WO 99/12565; WO 99/11241), más particularmente la combinación de QS21 y 3D-MPL como se describe en el documento WO 94/00153. De forma alternativa, una combinación de CpG más una saponina tal como QS21 también forma un adyuvante potente para uso en la presente invención. De forma alternativa, la saponina se puede formular en un liposoma o en un Iscorn y combinarse con un oligonucleótido inmunoestimulador.

5 Por tanto, sistemas adyuvantes adecuados incluyen, por ejemplo, una combinación de monofosforil lípido A, preferiblemente 3D-MPL, junto con una sal de aluminio. Un sistema mejorado incluye la combinación de un monofosforil lípido A y un derivado de saponina, particularmente la combinación de QS21 y 3D-MPL como se describió en el documento WO 94/00153, o una composición menos reactogénica donde el QS21 se desactiva en liposomas que contienen colesterol (DQ) como se describe en el documento WO 96/33739. Esta combinación puede comprender adicionalmente un oligonucleótido inmunoestimulador.

Se describe una formulación adyuvante particularmente potente que involucra QS21, 3D-MPL y tocoferol en una emulsión aceite en agua en el documento WO 95/17210 y es otra formulación preferida para uso en la invención.

Otra formulación preferida comprende un oligonucleótido CpG sólo o junto con una sal de aluminio.

10 En un caso adicional de la presente invención se proporciona un procedimiento para la fabricación de una formulación de vacuna como se describe en la presente invención, en el que el procedimiento comprende mezcla de un polipéptido de acuerdo con la invención con un adyuvante adecuado.

Las combinaciones adyuvantes particularmente preferidas para uso en las formulaciones de acuerdo con la invención son como sigue:

- 15 i) 3D-MPL + QS21 en un liposoma
- ii) Alumbre + 3D-MPL
- iii) Alumbre + QS21 en un liposoma + 3D-MPL
- iv) Alumbre + CpG
- v) 3D-MPL + QS21 + emulsión aceite en agua
- 20 vi) CpG

La administración de la composición farmacéutica puede tomar la forma de una o de más de una dosis individual, por ejemplo, como dosis repetidas de la misma composición que contiene polipéptido, o en un régimen de vacunación "inducción-refuerzo" heterólogo. Un régimen inducción-refuerzo heterólogo usa administración de diferentes formas de vacuna en la inducción y el refuerzo, cada una de las cuales puede incluir por sí misma dos o más administraciones. La composición de inducción y la composición de refuerzo tendrán al menos un antígeno en común, aunque no es necesariamente una forma idéntica del antígeno, esta puede ser una forma diferente del mismo antígeno.

30 Se pueden llevar a cabo inmunizaciones de inducción refuerzo de acuerdo con la invención con una combinación de proteína y formulaciones basadas en ADN. Una estrategia de este tipo se considera que es efectiva en la inducción de amplias respuestas inmunes. Vacunas de proteína adyuvadas inducen principalmente anticuerpos y respuestas inmunes de células T de ayuda, mientras la liberación de ADN como un plásmido o un vector vivo induce fuertes respuestas de linfocitos T citotóxicos (CTL). Por tanto, la combinación de proteína y vacunación con ADN proporcionará una amplia variedad de respuestas inmunes. Esto es particularmente relevante en el contexto de VIH, ya que se cree que ambos anticuerpos de neutralización y CTL son importantes para la defensa inmune contra VIH.

35 De acuerdo con la invención un calendario para la vacunación puede comprender la administración secuencial ("inducción-refuerzo") o simultánea de antígenos de de polipéptido y ADN que codifica los polipéptidos. El ADN se puede liberar como ADN desnudo tal como ADN plásmido o en la forma de un vector vivo recombinante, por ejemplo, un vector poxvirus, un vector adenovirus, un vector de virus de la rubeola, o cualquier otro vector vivo adecuado. Se pueden inyectar antígenos de proteína una o varias veces seguido de una o varias administraciones de ADN, se puede usar ADN en primer lugar durante una o más administraciones seguidas de una o más inmunizaciones con proteína.

45 Un ejemplo particular de inmunización por inducción-refuerzo de acuerdo con la invención incluye la inducción con ADN en la forma de un vector vivo recombinante tal como un vector de poxvirus modificado, en la inducción de respuestas inmunes amplias. Vacunas de proteína adyuvadas induce fuertes respuestas de linfocitos T citotóxicos (CTL). Por tanto, la combinación de proteína y vacunación con ADN proporcionará una amplia variedad de respuestas inmunes. Esto es particularmente relevante en el contexto de VIH, ya que se cree que ambos anticuerpos de neutralización y CTL son importantes para la defensa inmune contra VIH.

50 De acuerdo con la invención un calendario para la vacunación puede comprender la administración secuencial ("inducción-refuerzo") o simultánea de antígenos de de polipéptido y ADN que codifica los polipéptidos. El ADN se puede liberar como ADN desnudo tal como ADN plásmido o en la forma de un vector vivo recombinante, por ejemplo, un vector poxvirus, un vector adenovirus, un vector de virus de la rubeola, o cualquier otro vector vivo adecuado. Se pueden inyectar antígenos de proteína una o varias veces seguido de una o varias administraciones de ADN, se puede usar ADN en primer lugar durante una o más administraciones seguidas de una o más inmunizaciones con proteína.

55 Un ejemplo particular de inmunización por inducción-refuerzo de acuerdo con la invención incluye la inducción con ADN en la forma de un vector vivo recombinante tal como un vector poxvirus modificado, por ejemplo, virus Ankara modificado (MVA) o un alfavirus, por ejemplo, virus de encefalitis equina venezolana, o un vector adenovirus, o un vector de virus de la rubeola, seguido de refuerzo con una proteína, preferiblemente una proteína adyuvada.

Se describe en la presente invención un kit farmacéutico que comprende:

- a) una composición que comprende un polipéptido que comprende Nef o un fragmento o derivado inmunogénico de la misma y p17 y/o p24 Gag o un fragmento o derivado inmunogénico de los mismos, en la que cuando ambas p17 y p24 Gag están presentes hay al menos un antígeno de VIH o fragmento o derivado inmunogénico entre ellos, junto con un excipiente farmacéuticamente aceptable; y
- 5 b) una composición que comprende un polinucleótido que codifica una o más de Nef y Gag o un fragmento o derivado inmunogénico de Nef o Gag que contiene un epítipo de Nef o Gag presente en el polipéptido de a), junto con un excipiente farmacéuticamente aceptable.

Preferiblemente el polipéptido de a) comprende además RT o un fragmento o derivado inmunogénico de la misma tal como p51RT.

10 En un caso alternativo el kit farmacéutico comprende:

- a) una composición que comprende un polinucleótido que comprende Nef o un fragmento o derivado inmunogénico de la misma y p17 y/o p24 Gag o un fragmento o derivado inmunogénico de los mismos, en la que cuando ambas p17 y p24 Gag están presentes hay al menos un antígeno de VIH o fragmento o derivado inmunogénico entre ellos, junto con un excipiente farmacéuticamente aceptable; y
- 15 b) una composición que comprende un polipéptido que comprende uno o más de Nef o Gag o un fragmento o derivado inmunogénico de Nef o Gag que contiene un epítipo de Nef o Gag presente en el polipéptido de a), junto con un excipiente farmacéuticamente aceptable.

Preferiblemente el polinucleótido de a) codifica un polipéptido que comprende además RT o un fragmento o derivado inmunogénico de la misma tal como p51RT.

20 Polipéptidos y polinucleótidos preferidos para uso en un kit para inducción/refuerzo de acuerdo con la invención son polipéptidos y polinucleótidos como se describen en la presente invención. Por tanto, el componente proteico de un enfoque inducción/refuerzo de tipo proteína/ADN puede ser cualquiera de las proteínas de fusión preferidas descritas en la presente invención. Igualmente, el componente ADN puede ser un polinucleótido que codifica cualquiera de las proteínas preferidas.

25 Por tanto, por ejemplo, las fusiones p24 – RT – Nef – p17, p24 – RT* – Nef – p17, p24 – p51RT – Nef – p17, p24 – p51RT* – Nef – p17, p17 – p51RT – nef o p17 – p51RT* – Nef o cualquiera de las fusiones de p17 – Nef como se describen en la presente invención se pueden proporcionar en un kit de inducción refuerzo, en el que la composición de inducción comprende la proteína de fusión y la composición de refuerzo comprende un polinucleótido que codifica la proteína de fusión, o la composición de inducción comprende el polinucleótido y la composición de inducción comprende la proteína de fusión.

30

Tanto la composición de inducción como la composición de refuerzo se pueden liberar en más de una dosis. Adicionalmente las dosis de inducción y de refuerzo iniciales pueden ser seguidas de dosis adicionales que se pueden alternar para dar lugar a, por ejemplo, un cebado con plásmido de ADN / refuerzo con proteína / dosis de plásmido de ADN adicional / otras dosis de proteína.

35 Por optimización de codón se entiende que la secuencia de polinucleótidos está optimizada para asemejarse al uso de codón de genes en el sistema de expresión deseado, por ejemplo, un sistema procariótico tal como *E. coli*. En particular, el uso del codón en la secuencia está optimizado para asemejarse al de los genes de *E. coli* expresados en gran extensión

40 La finalidad de optimización del codón para expresión en un sistema recombinante de acuerdo con la invención es doble: mejorar los niveles de expresión del producto recombinante y hacer los productos de expresión más homogéneos (obtener un modelo de expresión más homogéneo). La homogenidad mejorada significa que hay menos productos de expresión irrelevantes tales como truncados. La adaptación de uso del codón a la expresión de *E.coli* puede eliminar también las secuencias de "desplazamiento de marco" putativas así como también la terminación prematura y/o sitios de iniciación internos.

45 El código de ADN tiene 4 letras (A, T, C y G) y usa estas para deletrear tres letras "codones" que representan los aminoácidos de las proteínas codificadas en un gen del organismo. La secuencia lineal de codones a lo largo de la molécula de ADN se translada en la secuencia lineal de aminoácidos en la(s) proteína(s) codificadas por estos genes. El código está altamente degenerado, con 61 codones que codifican para los 20 aminoácidos naturales y 3 codones que representan señales "de parada". Por tanto, la mayoría de los aminoácidos están codificados por más de un codón – de hecho varios están codificados por cuatro o más codones diferentes.

50

Donde esté disponible más de un codón para codificar para un aminoácido dado, se ha observado que los patrones de uso de codón de organismos son altamente no aleatorios. Diferentes especies muestran una preferencia diferente en su selección de codón y, además, la utilización de los codones puede ser marcadamente diferente en una única especie entre genes que se expresan a niveles superiores e inferiores. Esta preferencia es diferente en virus, plantas, bacterias y células de mamíferos, y algunas especies muestran una preferencia más fuerte lejos de una selección de codón aleatoria que otros. Por ejemplo, los humanos y otros mamíferos no son menos fuertemente preferidos que ciertas bacterias o virus. Por estas razones, hay una probabilidad significativa de que un gen viral de un virus de mamífero expresado en *E. coli*, o un gen extraño o recombinante expresado en células de mamífero tengan una distribución inapropiada de codones para expresión eficiente. Se cree que la presencia en una secuencia de ADN heteróloga de bucles de codones o una abundancia de codones que se observan raramente en el huésped en cuya expresión se van a producir, es predictiva de bajos niveles de expresión heteróloga en ese huésped.

55

60

En los polinucleótidos de la presente invención el modelo de uso del codón puede estar alterado del típico de virus

de inmunodeficiencia humana para representar más estrechamente la preferencia codónica del organismo diana, por ejemplo *E. coli*.

Hay una variedad de programas disponibles públicamente útiles para la optimización con codón, por ejemplo "CalcGene" (Hale y Thompson, Protein Expression and Purification 12: 185-189 (1998)).

5 Ejemplos

Ejemplo 1: Construcción y expresión de F4 de fusión p24 – RT – Nef – p17 y F4 optimizada con codón (co) del VIH-1

1. F4 no optimizada con codón

10 Se expresaron proteínas gag p24 de (proteína cápsida) y p17 (proteína matriz), la transcriptasa inversa y Nef de VIH-1 en cepa B834 de *E. coli*. (B834 (DE3) es un pariente auxótrofo de la metionina de BL21 (DE3)), bajo control del promotor del bacteriófago T7 (sistema de expresión pET).

Estos se expresaron como una proteína de fusión simple que contiene la secuencia completa de las cuatro proteínas. La secuencia que codifica p24 madura proviene del clon molecular de BH10 del VIH-1, secuencia p17 madura y gen de RT de HXB2 y gen de Nef de aislado de BRU.

15 Tras inducción, las células recombinantes expresaron niveles significativos de la fusión p24-RT-Nef-p17 que ascendieron hasta el 10% de la proteína total.

Cuando se cultivaron células e indujeron a 22° C, la proteína de fusión p24-RT-Nef-p17 se confinó principalmente en la fracción soluble de los lisados bacterianos (incluso tras congelación/descongelación). Cuando se cultivaron a 30° C, en torno al 30% de la proteína recombinante estaba asociada con la fracción insoluble.

20 La proteína de fusión p24-RT-Nef-p17 está hecha de hasta 1136 aminoácidos con una masa molecular de aproximadamente 129 kDa. La proteína de longitud completa migra hasta aproximadamente 130 kDa en gel SDS. La proteína tiene un punto isoeléctrico teórico (pI) de 7,96 basado en su secuencia de aminoácidos, confirmado por electroforesis en gel 2D.

Detalles del plásmido recombinante:

25 nombre: pRIT15436 (o nombre de laboratorio pET28b/p24-RT-Nef-p17)

vector huésped: pET28b replicón: colE1

selección: kanamycin

promotor: T7

inserción: gen de fusión de p24-RT-Nef-p17.

30 *Detalles de la proteína recombinante:*

Proteína de fusión p24-RT-Nef-p17: 1136 aminoácidos

Término N - p24: 232a.a. - bisagra:2a.a. - RT: 562a.a. -bisagra:2a.a. - Nef: 206a.a. - P17: 132a.a. – Término C

Secuencias de nucleótidos y aminoácidos:

Secuencia de nucleótidos

LISTA DE SECUENCIAS

atggttatcgtgcagaacatccaggggcaaatggtacatcaggccatcacctagaact
ttaatgcatgggtaaaagtagtagaagagaaggctttcagcccagaagtaatacccatg
ttttcagcattatcagaaggagccaccccacaagatttaaacaccatgctaaacacagtg
gggggacatcaagcagccatgcaaatgttaaagagaccatcaatgaggaagctgcagaa
tgggatagagtacatccagtgcagggcctattgcaccaggccagatgagagaacca
aggggaagtgacatagcaggaactactagtagtacccttcaggaacaaataggatggatgaca
aataatccacctatcccagtaggagaaattataaaagatggataatcctgggattaat
aaaatagtaagaatgtatagccctaccagcattctggacataagacaaggaccaaagaa
ccttttagagactatgtagaccggttctataaaaactctaagagccgagcaagcttcacag
gaggtaaaaaatggatgacagaaaccttggttggtccaaaatgccaaccagattgtaag
actattttaaagcattgggaccagcggctacactagaagaaatgatgacagcatgtcag
ggagtaggaggaccggccataaggcaagagttttgcatatgggccccattagccctat
tgagactgtgtcagtaaaattaaagccaggaatggatggcccaaaagttaaacaatggcc
attgacagaagaaaaataaaagcattagtagaaatttgtacagagatggaaaaggaagg
gaaaatttcaaaaattgggcctgaaaatccatacaataactccagtatttgccataaagaa
aaaagacagtactaaatggagaaaattagtagatttcagagaacttaataagagaactca
agacttctgggaagttcaattaggaataccacatcccgcaggggttaaaaaagaaaaatc
agtaacagtactggatgtgggtgatgcataatttttcagttcccttagatgaagacttcag
gaaatatactgcatttaccatacctagtataaacaatgagacaccagggattagatatca
gtacaatgtgcttccacagggatggaaaggatcaccagcaatattccaaagtagcatgac
aaaaatcttagagccttttagaaaacaaaatccagacatagttatctatcaatacatgga
tgatttgtatgtaggatctgacttagaaatagggcagcatagaacaaaaatagaggagct
gagacaacatctgttgaggtggggacttaccacaccagacaaaaaacatcagaaagaacc
tccattccttaaaatgggttatgaactccatcctgataaatggacagtacagcctatagt

gctgccagaaaaagacagctggactgtcaatgacatacagaagttagtggggaaattgaa
 ttgggcaagtcagatttaccagggtataaagtaaggcaattatgtaaactccttagagg
 aaccaaagcactaacagaagtaatacactaacagaagaagcagagctagaactggcaga
 aaacagagagattctaaaagaaccagtacatggagtgtattatgacccatcaaaagactt
 aatagcagaaatacagaagcaggggcaaggccaatggacatatcaaatttatcaagagcc
 atttaaaaatctgaaaacaggaaaatgatgcaagaatgaggggtgccacactaatgatgt
 aaaacaattaacagaggcagtgcaaaaaataaccacagaaagcatagtaatatggggaaa
 gactcctaaatttaaactgccatacaaaaaggaaacatgggaaacatgggtggacagagta
 ttggcaagccacctggattcctgagtgaggagtgtgttaatacccctccttagtgaaatt
 atgggtaccagttagagaaagaacccatagtaggagcagaaaccttctatgtagatggggc
 agctaacaggggagactaaattaggaaaagcaggatgttactaatagaggaagacaaaa
 agttgtcacccctaactgacacaacaaatcagaagactgagttacaagcaatttatctagc
 tttgcaggattcgggattagaagtaaacatagtaacagactcacaatatgcattaggaat
 cattcaagcacaaccagatcaaagtgaatcagagttagtcaatcaataatagagcagtt
 aataaaaaaggaaaaggtctatctggcatgggtaccagcacacaaggaattggaggaaa
 tgaacaagtagataaattagtcagtgctggaatcaggaaagtgctagctatgggtggca
agtggtcaaaaagtagtggtggatggcctactgtaagggaaagaatgagacgagctg
agccagcagcagatgggggtgggagcagcatctcgagacctggaaaaacatggagcaatca
caagtagcaatacagcagctaccaatgctgcttgtgcctggctagaagcacaagaggagg
aggaggtgggttttccagtcacacctcaggtaccttaagaccaatgact
tacaaggcagctgtagatcttagccactttttaaagaaaaggggggactggaagggcta
atcactcccaacgaagacaagatatccttgatctgtggatctaccacacacaaggctac
ttccctgattggcagaactacacaccagggccaggggtcagatatccactgacctttgga
tgggtgctacaagctagtaccagttgagccagataaggtagaagaggccaataaaggagag
aacaccagcttgttacacctgtgagcctgcatggaaatggatgacctgagagagaagtg
ttagagtgagggtttgacagccgcttagcatttcatcacgtggcccagagagctgcatccg
gagtaactcaagaactgcaggcctatgggtgagagagcgtcagtattaagcgggggaga
 attagatcgatgggaaaaaattcgggttaaggccaggggaaagaaaaaataaattaaa
 acatatagtatgggcaagcaggagctagaacgattcgcagttaatcctggcctgttaga
 aacatcagaaggctgtagacaaatactgggacagctacaacctccttcagacaggatc
 agaagaacttagatcattatataatacagtagcaacctctattgtgtgcatcaaaggat
 agagataaaagacaccaaggaagctttagacaagatagaggaagagcaaaacaaaagtaa
 gaaaaaagcacagcaagcagcagctgacacaggacacagcaatcaggtcagccaaaatta
 ctaa [Nº ID SEC: 1]

La secuencia p24 está en negrita

La secuencia de Nef está subrayada

5 Cajas: nucleotidos introducidos mediante construcción genética

Secuencia de aminoácidos

MVIVQNIQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATP	50
QDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVHPVHAGPIAPGQMREP	100
RGSDIAGTTSTLQEQIGWMTNPPPIVGEIYKRWIILGLNKIVRMYSPTS	150
ILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQANPDCK	200
TILKALGPAATLEEMMTACQGVGGPGHKARVLMGPISPIETVSVKLPKG	250
MDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKK	300
KDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAY	350
FSVPLDEDFRKYTAFTIP SINNETPGIRYQYNVLPQGWKGPSAIFQSSMT	400
KILEPFRKQNPDI VIYQYMDDLIVGSDLEIGQHRTKIEELRQHLLRWGLT	450
TPDKKHQKEPPFLKMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLN	500
WASQIYPGIKVRQLCKLLRGTKALTEVIPLTEEALELAENREILKEPVH	550
GVYYDPSKDLIAEIQKQGQGWTYQIYQEPFKNLKTGKYARMRGAHTNDV	600
KQLTEAVQKITTESIVIWGKTPKFKLP IQKETWETWWTEYWQATWIPEWE	650
FVNTPLVLKWLWYQLEKEPIVGAETFYVDGAANRETCLGKAGYVTNRGRQK	700
VVTLTDTTNQKTELQAIYLALQDSGLEVNI VTDSQYALGIIQAQPDQSES	750
ELVNQIIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKVLAMGGK	800
WSKSSVVGWPTVRERMRAEPAADGVGAASRDLEKHGAITSSNTAATNAA	850
CAWLEAQEEEEVGFVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQ	900
RRQDILDLDWIYHTQGYFPDWQNYTPGPGVRYPLTFGWICYKLVPEPKVE	950
EANKGENTSLHHPVSLHGMDDPEREVLEWRFD SRLAFHHVARELHPEYFK	1000
NCRPMGARASVLSGGELDRWEKIRLRPGGK KKYK LKHIVWASRELERFAV	1050
NPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQR IEIKD	1100
TKEALDKIEEEQNKSKKKAQQAAADTGHSNQVSQNY	1136

[Nº ID SEC: 2]

Secuencia de P24: aminoácidos 1-232 (en negrita)

Secuencia de RT: aminoácidos 235-795

5 **Secuencia de Nef: aminoácidos 798-1002**

Secuencia de P17: aminoácidos 1005-1136

Cajas: aminoácidos introducidos mediante construcción genética

K (Lisine): en lugar de triptópano (W). Mutación introducida para eliminar la actividad enzimática.

Expresión de la proteína recombinante:

10 En el plásmido pET, el gen diana (p24-RT-Nef-p17) está bajo control del promotor del bacteriófago T7 fuerte. Este promotor no es reconocido por la ARN polimerasa de E.coli y es dependiente de una fuente de T7 ARN polimerasa en la célula huésped. La célula huésped B834 (DE3) contiene una copia cromosomal del gen de T7 ARN polimerasa bajo control de lacUV5 y la expresión es inducida por la adición de IPTG al cultivo bacteriano.

15 Se cultivaron los pre-cultivos, en matraces de agitación, a 37° C hasta fase semi-logarítmica (A620:0,6) y luego se almacenaron a 4° C durante la noche (para evitar cultivos en fase estacionaria). Se cultivaron los cultivos en medio LBT suplementado con glucosa al 1% y 50 µg/ml de canamicina. La adición de glucosa al medio de crecimiento tiene la ventaja de reducir la expresión de proteína recombinante basal (evitando la depresión mediada por AMPc del promotor lacUV5).

20 Se usaron diez ml de cultivos almacenados durante la noche a 4° C para inocular 200 ml de medio LBT (sin glucosa) que contiene canamicina. Se cultivaron los cultivos a 30° C y 22° C y cuando la densidad óptica (OD) a 620 alcanzó 0,6, se añadió IPTG (1mM final). Se incubaron los cultivos durante 3, 5 y 18 horas (durante la noche). Se recogieron muestras antes y después de 3, 5 y 18 horas de inducción.

La preparación del extracto fue como sigue:

Se suspendieron agregados celulares en tampón* de ruptura (a una densidad óptica (O.D.) teórica de 10) y se trituró

mediante cuatro pasadas en prensa francesa (a 20.000 psi o 1250 bar). Se centrifugaron los extractos brutos (T) a 20.000g durante 30 minutos para separar las fracciones solubles (S) e insolubles (P).

*Tampón de ruptura: Tris-HCL 50 mM pH 8,0, EDTA 1mM, DTT + cóctel de inhibidores de proteasa 1mM (completo/Boehringer).

5 Análisis por SDS-PAGE y Transferencia western:

Se experimentaron fracciones correspondientes a agregado insoluble (P), sobrenadante (S) y extracto bruto (T) en SDS-PAGE al 10% en condiciones reductoras. Se detectó p24-RT-Nef-p17 recombinante mediante tinto de azul de Coomassie en Transferencia western (WB).

10 Tintado con Coomassie: la proteína p24-RT-Nef-p17 aparece como: una banda a \pm 130 kDa (ajustando con PM calculado)

PM teórico: 128,970 Daltons

PM aparente: 130 kDa

Análisis por Transferencia western:

Reactivos = - anticuerpo monoclonal para RT (p66/p51)

15 Adquirido a ABI (Advanced Biotechnologies) dilución: 1/5000

- Anticuerpo anti-ratón conjugado con fosfatasa alcalina dilución: 1/7500

Nivel de expresión: - banda específica de p24-RT-Nef-p17 muy fuerte tras inducción de 20 horas a 22° C, representando hasta el 10% de la proteína total (véase figura 1A).

“Solubilidad” de la proteína recombinante:

20 Extractos celulares “frescos” (fracciones T,S,P): con crecimiento/inducción a 22° C/20 horas, casi toda la proteína de fusión p24-RT-Nef-p17 se recupera en la fracción soluble del extracto celular (figura 1A). Con crecimiento/inducción a 30° C/20 horas, se asocia aproximadamente el 30% de la proteína p24-RT-Nef-p17 con la fracción insoluble (figura 1A).

“Congelación/descongelación” (fracciones S2, P2):

25 La fracción soluble (S1) (inducción de 20 horas a 22° C) se conservó a -20° C. Se congeló y se centrifugó a 20.000g/30 minutos: S2 y P2 (resuspendidas en volumen 1/10).

Tampón de ruptura con DTT: casi toda la proteína de fusión p24-RT-Nef-p17 permanece soluble (sólo precipitó de 1 a 5 %) (véase figura 1B)

Tampón de ruptura sin DTT: de 85 a 90 % de p24-RT-Nef-p17 permanece soluble (figura 1B)

30 Figuras:

Figura 1A – Tintado de Coomassie y transferencia western.

Figura 1B – ensayo de solubilidad de p24-RT-Nef-p17

La proteína F4 se purificó usando el procedimiento de purificación I en el ejemplo 7.

35 Las condiciones de crecimiento celular y de inducción y preparación de extractos celulares para los ejemplos que siguen son como se describe en el ejemplo 1 a menos que se especifiquen otras condiciones (por ejemplo, temperatura, composición del tampón de ruptura).

2. F4 optimizada con codón

40 Se optimiza con codón la siguiente secuencia de polinucleótidos de modo que el uso del codón se asemeja al uso del codón en un gen ampliamente expresado en *E.coli*. La secuencia de aminoácidos es idéntica a la dada anteriormente para F4 no optimizado con codón.

Secuencia de nucleótidos para F4co:

atggtcattgttcagaacatacagggccaaatgggtccaccaggcaattagtccgcgaact
cttaatgcatgggtgaaggtcgtggaggaaaaggcattctccccggaggcattccgatg
tttctgcgctatctgagggcgcaacgccgcaagaccttaataccatgcttaacacggta
ggcgggcaccaagccgctatgcaaagtctaaaagagactataaacgaagaggccgccgaa
tgggatcgagtgcacccgggtgcacgccggcccaattgcaccaggccagatgcgcgagccg
cgccgggtctgatattgcaggaactacgtctacccttcaggagcagattgggtggatgact
aacaatccaccaatcccggtcggagagatctataagagggtggatcatactgggactaac
aagatagtcgcgatgtattctccgacttctatactggatatacgccaaggccaaaggag
ccgttcagggactatgtcgaccgattctataagacccttcgcgcagagcaggcatcccag
gaggtcaaaaattggatgacagaaactcttttgggtgcagaatgcgaatccggattgtaaa
acaattttaaaggctctaggaccggccgcaacgctagaagagatgatgacggcttgtcag
ggagtccggtggaccggggcataaagcccgcgtctta^{ca}catgggcccgatatctccgat
agaaacagtttcgggtcaagcttaaaccagggatggatgggtccaaaggccaagcagtggcc
gctaaccggaagagaagattaaggcgctcgtagagatttgtactgaaatggagaaggaagg
caagataagcaagatcgggcccagagaaccgtaacaatacaccggtatttgcaataaagaa
aaaggattcaacaaaatggcgaaagcttgtagattttagggaaactaaacaagcgaacca
agacttttgggaagtccaactagggatcccacatccagccggtctaaagaagaagaaatc
ggtcacagtcctggatgtaggagacgcataattttagtgtaccgcttgatgaggacttccg
aaagtatactgcggttactataccgagcataaacaatgaaacgccaggcattcgctatca
gtacaacgtgctcccgcagggtggaaggggtctccggcgatatttcagagctgtatgac
aaaaatacttgaaccattccgaaagcagaatccggatattgtaatttaccaatacatgga
cgatctctatgtgggtcggatctagaaattggggcagcatcgactaagattgaggaact
gaggcaacatctgcttcgatggggcctcactactcccgacaagaagcaccagaaggagcc

gccgttcctaaagatgggctacgagcttcatccggacaagtggacagtacagccgatagt
gctgcccgaaggattcttggaccgtaaattgatattcagaaactagtcggcaagcttaa
ctgggctctcagatttaccaggcattaaggtccgacagctttgcaagctactgagggg
aactaaggctctaacagaggtcatccattaacggaggaagcagagcttgagctggcaga
gaatcgcgaaattcttaaggagccggtgcacgggtataactacgaccctccaaggacct
tatagccgagatccagaagcaggggcagggccaatggacgtaccagatatcaagaacc
gtttaagaatctgaagactgggaagtacgcgcgcatgagcaggggctcataactaatgatgt
aaagcaacttacggaagcagtaaaaagattactactgagctctatgtgatatggggcaa
gaccccaaagttcaagctgccatacagaaggaaacatgggaaacatggtggactgaata
ttggcaagctacctggattccagaatgggaatttgtcaacacgcccacttgttaagct
ttggtagcagcttgaaaaggagccgatagtaggggcagagaccttctatgtcgatggcgc
cgcgaatcgcgaaacgaagctaggcaaggcgggatagctgactaataggggcccga
ggcgtaccccttacggataccaccaatcagaagactgaactacaagcatttaccttgc
acttcaggatagtgccctagaggtcaacatagtcacggactctcaatatgcgcttggcat
tattcaagcgcagccagatcaaagcgaagcagcttgtaaaccaataatagaacagct
tataaagaagagaaggtatatctggcctgggtccccgctcacaagggaaattggcggcaa
tgagcaagtggacaagctagtcagcgtgggattcgcaaggttcttcgatgggggta
agtggtctaagtctagcgtagtcggctggccgacagtcgagcgcgcatgagcgcgcg
aaccagccgagatggcgtgggggcagcgtctagggatctggagaagcaggggctataa
cttccagtaaacagggcggcagcaacgcccgcgcatgagcgtggttagaagccaagaagagg
aagaagtgggtttccggtaactccccaggtgcccgttaaggccgatgacc
tataaggcagcgggtggatcttctcacttcttaaggagaaagggggctggagggtta
atcacagccagagggcagcaggtatcttctgatctgtggattaccataaccaggggtac
ttccggactggcagaattacccccggggccagggcgtgagctatcccctgacttccggg
tggtgctacaaactagtcccagtggaaccgacaaggtcgaagaggctaataagggcgag
aacacttctcttctcaccggtaagcctgcacgggatggatgaccagaacgagaggtt
ctagaatggagggttcgactctcgacttgcgttccatcacgtagcagcgcgagctgcatcca
gaatattcaagaactgdcgccaatgggcgccagggccagtgacttagtgggcgaga
actagatcgatgggaaaagatacgcctacgcccgggggcaagaagaagtaaacgttaa
gcacattgtgtgggctctcgcaacttgagcgattcgagtgatccaggcctgcttga
gacgagtgaaggctgtaggcaattctggggcagctacagccgagcctacagactggcag
cgaggagcttcgtagtctttataataaccgtcgcgactctctactgcgttcatcaacgaat
tgaaataaaggataactaaagaggcccttgataaaaattgaggaggaacagaataagtcgaa
aaagaaggccagcagggccgcccgcacaccgggcacagcaaccaggtgtccaaaacta
ctaa

[Nº ID SEC: 3]

La secuencia de p24 está en negrita

La secuencia de Nef está subrayada

5 Cajas: nucleótidos introducidos mediante construcción genética

Los procedimientos usados en relación con F4 no optimizado con codón se aplicaron para la secuencia optimizada con codón.

Detalles del plásmido recombinante:

nombre: pRIT15513 (nombre de laboratorio: pET28b/p24-RT-Nef -p17)

10 vector huésped: pET28b

replicón: colE1

selección: canamicina
 promotor: T7
 inserción: gen de fusión p24-RT-Nef-p17, optimizado con codón

5 Se expresó el gen optimizado con codón de F4 en células de *E.coli* BLR(DE3), un derivado de recA⁻ de la cepa B834(DE3). La mutación de RecA evita la producción putativa de fagos lambda.

Se cultivaron pre-cultivos, en matraces de agitación, a 37° C hasta fase semi-logarítmica (A₆₂₀:0,6) y se almacenaron luego a 4 °C durante la noche (para evitar cultivos en fase estacionaria).

10 Se cultivaron cultivos en medio LBT suplementado con glucosa al 1% y canamicina 50 µg/ml. La adición de glucosa al medio de crecimiento tiene la ventaja de reducir la expresión de proteína recombinante basal (evitando depresión mediada por AMPc del promotor lacUV5).

Se usaron diez ml de cultivos almacenados durante la noche a 4° C para inocular 200 ml de medio LBT (sin glucosa) que contiene canamicina. Se cultivaron cultivos a 37° C y cuando la densidad óptica (OD) a 260 alcanzó 0,6, se añadió IPTG (1 mM final). Se incubaron cultivos durante 19 horas más (durante la noche), a 22° C. Se recogieron muestras antes y a las 19 horas de la inducción.

15 La preparación de extracto fue como sigue:

Se resuspendieron agregados de células en tampón de muestra (a una densidad óptica (OD) teórica de 10), se hirvió y se cargaron directamente en SDS-PAGE.

Análisis SDS-PAGE y Transferencia western:

Se experimentaron muestras de extractos brutos en SDS-PAGE al 10% en condiciones reductoras.

20 Se detecta proteína recombinante p24-RT-Nef -p17 mediante tinto con azul de Coomassie (figura 2) y en Transferencia western.

Tintado con Coomassie: la proteína p24-RT-Nef-p17 aparece como: una banda a ± 130 kDa (ajustada con PM calculado)

PM teórico: 128.967 Daltons

25 PM aparente: 130 kDa

Análisis Transferencia western:

Reactivos = - Anti RT policlonal de conejo (PO3L16 de conejo) dilución: 1/10.000

- Anti Nef-Tat policlonal de conejo (388 de conejo) dilución 1/10.000

30 - Anticuerpo anti-conejo conjugado con fosfatasa alcalina dilución: 1/7500

Tras inducción a 22° C durante 19 horas, las células recombinantes BLR(DE3) expresaron la fusión F4 a un nivel muy alto que varía de 10 a 15% de la proteína total.

35 En comparación con F4 del gen nativo, el perfil de producto recombinante de F4 del gen optimizado con codón se encuentra ligeramente simplificado. La principal banda relacionada con F4 a 60 kDa desapareció, así como también bandas menores por debajo (véase la figura 2). En comparación con la cepa recombinante B834(DE3) que expresa F4, la cepa BLR(DE3) que produce F4co tiene las siguientes ventajas: mayor producción de proteína de longitud completa F4, modelo de banda menos complejo del producto recombinante.

Ejemplo 2: Construcción y expresión de P51 RT (RT optimizada con codón, truncada)

40 La región RT/p66 entre los aminoácidos 428-448 es susceptible para las proteasas de *E.coli*. La construcción de P51 termina en Leu 427 dando lugar a la eliminación del dominio ARNasa H (véase alineación de la secuencia de RT en la figura 3).

También se eliminaron las secuencias "que desplazan el marco de lectura" de *E. coli* putativas en la secuencia de gen nativo de RT (mediante optimización con codón del gen p51).

Diseño/construcción del gen sintético p51:

45 La secuencia del gen p51 sintético se diseñó de acuerdo con el uso de codón en *E.coli*. Por tanto, este se optimizó con codón de modo que el uso de codón se asemeja al uso de codón en un gen altamente expresado en *E.coli*. Se construyó el gen sintético como sigue: se ensamblaron 32 oligonucleótidos en una PCR (reacción en cadena de la polimerasa) de etapa única. En una segunda PCR se amplificó el ensamblaje de longitud completa usando los cebadores de extremo y el producto de PCR resultante se clonó en plásmido intermedio pGEM-T. Tras corrección de errores puntuales introducidos durante la síntesis del gen, el gen sintético p51 se clonó en plásmido de expresión pET29a. Este plásmido recombinante se usó para transformar células B834 (DE3).

Características de la proteína recombinante:

Secuencia de nucleótido de P51 RT

```

atgagtaggtccgatctctccgatagaaacagtttcgggtcaagcttaaaccagggatg 60
gatgggtccaaagggtcaagcagtgggccgctaacgggaagagaagattaaggcgctcgtagag 120
atgtgtactgaaatggagaaggaaggcaagataagcaagatcgggccagagaaccggtac 180
aatacaccggtatgtgcaataaagaagaaggattcaacaaaatggcgaaagcttgtagat 240
tttagggaactaaacaagcgaacccaagacttttggaagtccaactaggtatcccacat 300
ccagccgggtctaaagaagaagaaatcgggtcacagtcctggatgtaggagacgcataat 360
agtgtaccgcttgatgaggacttccgaaagtatactgctttactataccgagcataaac 420
aatgaaacgccaggcatctcgtatcagtaaacgtgctcccgcagggtggaaggggtct 480

ccggcgatatttcagagctctatgacaaaaatacttgaaccattccgaaagcagaatccg 540
gatattgtaatttaccatacatggacgatctctatgtgggctcggatctagaaattggg 600
cagcatcgcactaagattgaggaactgaggcaacatctgcttcgatggggcctcactact 660
cccgacaagaagcaccagaaggagccgccttcttaagatgggctacgagcttcatccg 720
gacaagtggacagtacagccgatagtgtgcccgaaaaggattcttggaccgtaaatgat 780
atcagaaactagtcggcaagcttaactgggcctctcagatttaccaggcattaaggctc 840
cgacagctttgcaagctactgaggggaactaaggctctaacagaggtcatcccattaacg 900
gaggaagcagagcttgagctggcagagaatcgcgaaattcttaaggagccgggtgcacggg 960
gtatactacgacctccaaggaccttatagccgagatccagaagcaggggcagggccaa 1020
tggacgtaccagatatacaagaaccgtttaagaatctgaagactgggaagtacgcgcgc 1080
atgagggggctcataactaatgatgtaagcaacttacggaagcagtacaaaagattact 1140
actgagctctatgtgatatggggcaagaccccaaaagttcaagctgcccatacagaaggaa 1200
acatgggaaacatggtggactgaatattggcaagctacctggattccagaatgggaattt 1260
gtcaacacgccgcgctggtaaaactgaggcctgctagctaa 1302
    
```

[Nº ID SEC: 4]

5 **Cajas: aminoácidos introducidos mediante construcción genética**

Secuencia de aminoácido:

```

MSTGPISPIETVSVKLPKGMGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPY 60
NTPVFAIKKDKSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDYF 120
SVPLDEDFRKYTAFTIP SINNETPGIRYQYNVLPQGWKSPAI FQSSMTKILEPFRKQNP 180
DIVIYQYMDDLIVGSDLEIGQHRTKIEELRQHLLRWGLTTPDKKHQKEPPFLKMGYELHP 240
DKWTVQPIVLPKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGT KALTEVIPLT 300
EEAELELAENREILKEPVHGVYDPSKDLIAEIQKQGQGWTYQIYQEPFKNLKTGKYAR 360
MRGAHTNDVKQLTEAVQKITTESIVIWKTPKFKLPIQKETWETWTEYQATWIPWEF 420
VNTPLVKL RPAS 433
    
```

[Nº ID SEC: 5]

Cajas: aminoácidos introducidas mediante construcción genética.

10 **K (Lisina): en lugar de triptófano (W). Mutación introducida para eliminar la actividad del enzima.**

Longitud, peso molecular, punto isoeléctrico (IP):

433 AA, PM: 50,3 kDa,, IP: 9,08

Expresión de p51 en células B834(DE3):

5 Se evaluaron el nivel de expresión de P51 y solubilidad de proteína recombinante, en paralelo con la cepa de producción RT/p66.

Nivel de expresión de p51:

Condiciones de inducción: se cultivaron/inducieron células a 37° C (+IPTG 1mM), durante 5 horas.

Tampón de ruptura: Tris/HCl 50 mM, pH:7,5, EDTA 1mM, +/- DTT 1mM.

Análisis Transferencia western:

10 **Reactivos:**

- anti RT policlonal de conejo (PO3L16 de conejo) (dilución: 1/10.000)

- anticuerpo anti-conejo conjugado con fosfatasa alcalina (dilución: 1/7500)

Se experimentaron fracciones celulares correspondientes a extractos brutos (T), agregados insolubles (P) y sobrenadante (S) en SDS-PAGE al 10% en condiciones reductoras.

15 Como se ilustró en el gel tintado con Coomassie y Transferencia western (figura 4) se observó una expresión muy alta de P51 (de 15 a 20% de la proteína total), mayor que la observada para P66.

Para ambas proteínas p51 y p66 (después de 5 horas de inducción a 37° C), el 80% de los productos recombinantes se recuperaron en la fracción soluble (S1) de extractos celulares (véase la figura 4). Cuando se expresaron a 30° C, se asociaron el 99% de proteínas recombinantes con la fracción soluble (datos no mostrados).

20 El modelo de Transferencia western para p51 fue multibanda, pero menos complejo que el observado para P66.

Ensayo de solubilidad

Ensayo de solubilidad: congelación/descongelación de la fracción soluble (S1) (inducción de 5 horas, 37° C) preparada en condiciones reductoras (tampón de ruptura con DTT) y no reductoras. Tras descongelación se centrifugaron las muestras de S1 a 20.000g/30 minutos, generando S2 y P2 (p2 se resuspende en volumen 1/10).

25 Después de la congelación/descongelación de fracciones solubles (S1), preparadas en condiciones reductoras, así como en condiciones no reductoras, se recuperan aún el 99% de p51 y p66 en la fracción soluble (S2). Sólo se encuentra un 1% en el precipitado (P2). Esto se muestra en la figura 5.

Ejemplo 3: Construcción y expresión de p17-Nef y Nef-p17 con o sin conector

30 Se construyeron proteínas de fusión dobles con y sin conectores. Los conectores ayudaron a disminuir las interacciones potenciales entre los dos partícipes de la fusión y son como siguen:

Nef-[GSGGGP]-P17 y p17-[GSGGGP]-Nef

Construcción de plásmidos recombinantes:

- **Vector de expresión pET29a/Nef-p17:**

35 Se amplificó el gen de fusión Nef-p17 mediante PCR a partir del plásmido recombinante de F4. El producto de PCR se clonó en el vector de clonación pGEM-T intermedio y a continuación en el vector de expresión pET29a.

- **Vector de expresión pET28b/p17-Nef:**

Se amplificó el gen de Nef mediante PCR a partir del plásmido recombinante de F4. Se clonó el producto de PCR en el vector de clonación pGEM-T intermedio y a continuación en el vector de expresión pET28b/p17, como un terminal C en la fusión de marco con el gen p17.

- 40 • **Vector de expresión pET29a/Nef-conector-p17 y pET28b/p17-conector-Nef:**

Se insertó un fragmento de ADN de 18 bp que codifica para el conector hexapeptídico (GSGGGP) entre los partícipes de la fusión Nef y p17, mediante mutagénesis dirigida al sitio (usando el "Sistema de mutagénesis dirigida al sitio GeneTailor", Invitrogen).

Características de la proteína recombinante:

- 45 • **Longitud, peso molecular, punto isoeléctrico (IP) Nef-p17 (denominado NP):** 340 AA, PM: 38,5 kDa, IP:7,48

Nef-GSGGGP-P17 (denominado NLP): 346 AA, PM:38,9 kDa, IP: 7,48

p17-Nef (denominado PN): 342 AA, PM: 38,7 kDa, IP: 7,19

p17-GSGGGP-Nef (denominado PLN): 348 AA, PM: 39,1k Da, IP: 7,19

- **Secuencias de aminoácido y secuencias de polinucleótido:**

5 Secuencia de nucleótido de Nef-p17

```

Atgggtggcaagtgggtcaaaaagtagtgtgggtggatggcctactgtaagggaaagaatg 60
Agacgagctgagccagcagcagatgggggtgggagcagcatctcgagacctggaaaaacat 120
Ggagcaatcacaagtagcaatacagcagctaccaatgctgcttgtgctggctagaagca 180
Caagaggaggaggaggtgggttttccagtcacacctcaggtaccttaagaccaatgact 240
Tacaaggcagctgtagatcttagccactttttaaagaaaaggggggactggaagggcta 300
Attcactcccaacgaagacaagatatccttgatctgtggatctaccacacacaaggctac 360
Ttccctgattggcagaactacacaccagggccaggggtcagatatccactgacctttgga 420
Tggtgctacaagctagtaccagttgagccagataaggtagaagaggccaataaaggagag 480
Aacaccagcttgttacacctgtgagcctgcatggaatggatgacctgagagagaagtg 540
Ttagagtggaggtttgacagccgctagcatttcatcacgtggccccgagagctgcatccg 600
Gagtacttcaagaactgcaggcctatgggtgagagagcgtcagtat taagcgggggagaa 660
Ttagatcgatgggaaaaaatcgggttaaggccaggggaaagaaaaaatataaattaaaa 720
Cataatagtagtgggcaagcagggagctagaacgatcgcagttaatcctggcctgttagaa 780
Acatcagaaggctgtagacaaatactgggacagctacaacctcccttcagacaggatca 840
Gaagaacttagatcattatataatacagtagcaacctctattgtgtgcatcaaaggata 900
Gagataaaagacaccaaggaagctttagacaagatagaggaagagcaaaacaaaagtaag 960
Aaaaaagcacagcaagcagcagctgacacaggacacagcaatcaggtcagccaaaattac 1020
Taa 1023
    
```

[Nº ID SEC: 6]

Nef-p17 (NP)

```

MGGKWSKSSVVGWPTVRERMRAEPAADGVGAASRDLEKHGAITSSNTAATNAACAWLEA 60
QEEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRQDILDWYHTQGY 120
FPDWQNYTPGPGVRYPLTFGWICYKLVPEPDKVEEANKGENTSLHPVSLHGMDPPERV 180

LEWRFD SRLAFHHVARELHPEYFKNCRPMGARASVLSGGELDRWEKIRLRPGGKKKYK 240
HIWASRELERFAVNPGLLETSEGCRQILGQLQPSLQTSSEELRSLYNTVATLYCVHQRI 300
EIKDTKEALDKIEEEQNKSKKKAQQAADTGHNSQVSNY 340
    
```

10 [Nº ID SEC:7]

Caja: aminoácidos introducidos mediante construcción genética.

La secuencia Nef está en negrita.

Secuencia de nucleótidos de P17-Nef:

AtgggtgCGagagcgtcagtattaagcgggggagaattagatcgatgggaaaaaatcgg 60
 Ttaaggccaggggaaagaaaaatataaattaaaacatatagtagtgggcaagcaggag 120
 Ctagaacgattcgcagttaatcctggcctgttagaaaacatcagaaggctgtagacaaata 180
 Ctgggacagctacaacccatcccttcagacaggatcagaagaacttagatcattatataat 240
 Acagtagcaaccctctattgtgtgcatcaaaggatagagataaaagacaccaaggaagct 300
 Ttagacaagatagaggaagagcaaaacaaaagtaagaaaaagcacagcaagcagcagct 360
 Gacacaggacacagcaatcaggtcagccaaaattacctcgacaggcctatgggtggcaag 420
 Tggtcaaaaagtagtgggtggatggcctactgtaagggaaagaatgagacgagctgag 480
 Ccagcagcagatgggggtgggagcagcatctcgagacctggaaaaacatggagcaatcaca 540
 Agtagcaatacagcagctaccaatgctgcttgtgacctggctagaagcacaagaggaggag 600
 Gagggtgggttttccagtcacacctcaggtacctttaagaccaatgacttacaaggcagct 660
 Gtagatcttagccactttttaaagaaaaggggggactggaagggttaattcactcccaa 720
 Cgaagacaagatatccttgatctgtggatctaccacacacaaggctacttcctgattgg 780
 Cagaactacacaccagggccaggggtcagatatccactgacctttggatgggtgctacaag 840
 Ctagtaccagttgagccagataaggtagaagaggccaataaaggagagaacaccagcttg 900
 Ttacacctgtgagcctgcatggaatggatgacctgagagagaagtgttagagtggagg 960
 Tttgacagccgctagcatttcatcacgtggcccagagctgcatccggagtacttcaag 1020
 Aactgctaa 1029

[Nº ID SEC: 8]

P17-Nef (PN)

MGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQI 60
LGQLQPSLQTGSEELRSLYNTVATLYCVHQRIEIKDTKEALDKIEEEQNKSKKKAQQAAA 120
DTGHSNQVSONY**LDRP****MGGKWSKSSVVGWPTVRERMRAEPAADGVGAASRDLEKHGAIT 180**
SSNTAATNAACAWLEAQEEEEVGFVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQ 240
RRQDILDWYHTQGYFPDWQNYTPGGVRYPLTFGWYKLVPEPKVVEANKGENTSL 300
LHPVSLHGMDDPEREVLEWRFD SRLAFHHVARELHPEYFKNC 342

5 [Nº ID SEC: 9]

Caja: aminoácidos introducidos mediante construcción genética.

La secuencia de p17 está en negrita.

Secuencia de nucleótidos de Nef-conector-p17:

Atgggtggcaagtgggtcaaaaagtagtggttggatggcctactgtaagggaaagaatg 60
 Agacgagctgagccagcagcagatgggggtgggagcagcatctcgagacctggaaaaacat 120
 Ggagcaatcacaagtagcaatacagcagctaccaatgctgcttgtgcctggctagaagca 180
 Caagaggaggaggagggtgggttttccagtcacacctcaggtacctttaagaccaatgact 240
 Tacaaggcagctgtagatcttagccactttttaaagaaaaggggggactggaagggtta 300
 Attcactcccaacgaagacaagatctcttgatctgtggatctaccacacacaaggctac 360
 Ttccctgattggcagaactacacaccagggccaggggtcagatatccactgacctttgga 420
 Tgggtgctacaagctagtaccagttgagccagataaggtagaagaggccaataaaggagag 480
 Aacaccagcttgttacaccctgtgagcctgcatggaatggatgaccctgagagagaagtg 540
 Ttagagtggaggttgacagccgctagcatttcatcacgtggcccgagagctgcatccg 600
 Gagtacttcaagaactgcaggcctggatccggtggcggccctatgggtgagagagcgtca 660
 Gtattaagcgggggagaattagatcgatgggaaaaaattcgggtaaggccagggggaaag 720
 Aaaaaatataaattaaaacatatagtatgggcaagcagggagctagaacgattcgcagtt 780
 Aatcctggcctgttagaaacatcagaaggctgttagacaaatactgggacagctacaacca 840
 Tcccttcagacaggatcagaagaacttagatcattatataatacagtagcaaccctctat 900
 Tgtgtgcatcaaaggatagagataaaagacaccaaggaagctttagacaagatagaggaa 960
 Gagcaaaacaaaagtaagaaaaaagcacagcaagcagcagctgacacaggacacagcaat 1020
 Caggtcagccaaaattactaa 1041

[Nº ID SEC: 10]

Nef-conector-p17 (NLP)

MGGKWSKSSVVGWPTVRERMRAEPAADGVGAASRDLEKHGAI TSSNTAATNAACAWLEA 60
 QEEEEVGFVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRQDILDWYHTQGY 120
 FPDWQNYTPGPGVRYPLTFGWICYKLVPEPDKVEEANKGENTSLLHPVSLHGMDDPEREV 180
 LEWRFDSRLAFHHVARELHPEYFKNCRP GSGGGPMGARASVLSGGELDRWEKIRLRPGGK 240
 KKYKLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLY 300
 CVHQRIEIKDTEALDKIEEEQNKSKKKAQQAADTGHNSQVSNY 346

5 [Nº ID SEC: 11]

Conector hexapeptídico

Caja: aminoácidos introducidos mediante construcción genética

Secuencia de nucleótido de P17-conector-Nef:

Atgggtgagagagcgtcagtattaagcgggggagaattagatcgatgggaaaaaattcgg 60
 Ttaaggccagggggaaagaaaaatataaattaaaacatatagtatgggcaagcagggag 120
 Ctagaacgattcgcagttaatcctggcctgttagaaacatcagaaggctgtagacaaata 180
 Ctgggacagctacaaccatcccttcagacaggatcagaagaacttagatcattatataat 240
 Acagtagcaaccctctattgtgtgcatcaaaggatagagataaaagacaccaaggaagct 300

Ttagacaagatagaggaagagcaaaacaaaagtaagaaaaagcacagcaagcagcagct 360
 Gacacaggacacagcaatcaggtcagccaaaattacctcgacaggcctggatccggtggc 420
 Ggtcctatgggtggcaagtgggtcaaaaagtagtgggtggatggcctactgtaagggaa 480
 Agaatgagacgagctgagccagcagcagatgggggtgggagcagcatctcgagacctggaa 540
 Aaocatggagcaatcacaagtagcaatacagcagctaccaatgctgcttgtgcctggcta 600
 Gaagcacaagaggaggaggaggtgggtttccagtcacacctcaggtaccttaagacca 660
 Atgacttacaaggcagctgtagatcttagccactttttaaagaaaaggggggactggaa 720
 Gggctaattcactcccaacgaagacaagatatccttgatctgtggatctaccacacacaa 780
 Ggctacttccttgattggcagaactacacaccagggccaggggtcagatatccactgacc 840
 Tttggatgggtgctacaagctagtaccagttgagccagataaggtagaagaggccaataaa 900
 Ggagagaacaccagcttgttacacctgtgagcctgcatggaatggatgacctgagaga 960
 Gaagtgttagagtggaggttgacagccgcctagcatttcatcacgtggccccgagagctg 1020
 Catccggagtacttcaagaactgctaa 1047

[Nº ID SEC: 12]

P17-conector-Nef (PLN)

MGARASVLSGGELDRWEKIRLRPGGKKKYKCLKHIVWASRELERFAVNPGLLETSEGCRQI 60
 LGQLQPSLQTGSEELRSLYNTVATLYCVHQRIEIKDTKEALDKIEEQNKSKKKAQAAA 120
 DTGHSNQVSQNYLDRP GSGGGFPMGGKWSKSSVVGWPTVRRMRRAEPAADGVGAASRDLE 180
 KHGAITSSNTAATNAACAWLEAQEEEEVGFVPTPOVPLRPMTYKAAVDLSHFLKEKGGLE 240
 GLIHSQRRQDILDLDWIYHTQGYFPDWQNYTPGPGVRYPLTFGWICYKLVPEPDKVEEANK 300
 GENTSLHHPVSLHGMDDPEREVLEWRFDSRLAFHHVARELHPEYFKNC 348

5 [Nº ID SEC:13]

Conector hexapeptídico

Caja: aminoácidos introducidos mediante construcción genética.

Expresión comparativa de Nef-p17, fusiones p17-Nef, con y sin conectores:

10 Las cuatro cepas recombinantes se indujeron a 30° C durante 3 horas, en paralelo con cepas que producen F4 y Nef. Se prepararon extractos brutos y se analizaron mediante gel tintado con Coomassie y Transferencia westernting.

Análisis Transferencia western:

- Reactivos:*
- anti RT policlonal de conejo (PO3L16 de conejo) (dilución: 1/10.000)
 - anticuerpo anti-conejo cojugado con fosfatasa alcalina (dilución: 1/7500)

15 Como se ilustra en la figura 6, las fusiones Nef-p17 y p17-Nef, con y sin conector, se expresan a un alto nivel (10% de las proteínas totales).

En el Transferencia western: las cuatro construcciones de fusión dobles presentan un modelo multi-banda, pero menos complejo que el observado para F4. Cuando se expresan solas, las proteínas Nef y p17 presentan patrones de banda simples.

20 Se analizaron adicionalmente cepas que expresan fusiones Nef-p17 (NP) y p17-Nef (PN) sin péptido conector (ensayos de solubilidad, véase a continuación).

Ensayo de solubilidad de Nef-p17 y p17-Nef:

Se indujeron proteínas Nef-p17 and p17-Nef en paralelo con cepas que producen F4 y Nef.

Condiciones de inducción: crecimiento/inducción de células a 30 °C (+IPTG 1 mM), durante 3 horas.

25 *Tampón de ruptura:* Tris/HCl 50 mM pH: 8, NaCl 50 mM, EDTA 1 mM

Extractos celulares frescos:

Se prepararon extractos celulares (en condiciones no reductoras) y fracciones que corresponden a extractos brutos (T), agregado insoluble (P) y se analizaron los sobrenadantes (S1) en gel tintado con Coomassie y Transferencia

western.

Como se ilustra en la figura 7 en gel tintado con Coomassie y Transferencia western, casi todas las Nef-p17, p17-Nef, así como también proteínas Nef se recuperan en la fracción soluble (S) de extractos celulares. Para la construcción F4: del 5 a 10% de proteína recombinante ya se recuperaba en la fracción de agregado.

5 **Conclusiones:**

Todas las construcciones de fusión dobles ensayadas son altamente expresadas (> 10% de proteína total). Las proteínas de fusión P17-Nef y Nef-p17 son más solubles que F4. Ambas presentan un modelo de WB menos complejo.

Ejemplo 4: Construcción y expresión de p24-RT*-Nef-p17 (F4*)

10 F4* es una versión mutada de la fusión F4 (p24-RT/p66-Nef-p17) donde se sustituye la metionina en la posición 592 por una lisina. Esta metionina es un sitio de "inicio" transcripcional interno putativo, ya que soportó secuenciación del terminal N llevada a cabo en una muestra de eluido en Q-sepharosa del experimento de purificación de F4. Incluso, la banda pequeña relacionada con F4 principal a 62 kDa presente en la muestra eluida en Q comienza en la metionina 592.

15 La metionina se reemplaza por una lisina: RMR → RKR. El motivo RKR está presente de forma natural en secuencias de RT de clado A.

Se evaluó el impacto de esta mutación en epítomos CD4-CD8:

- se pierde un epítomo CTL de HLA-A3 (A* 3002), pero están presentes otros 9 epítomos HLA-A3 en la secuencia de RT.
- 20 - no se identificó epítomo de ayuda en esta región.

Características de proteína recombinante:

Término N - **p24: 232a.a.** - **bisagra:2a.a.** - **RT: 562a.a.** - **bisagra:2a.a.** - **Nef: 206a.a.** -
 - **P17: 132a.a.** - **Término C**

- **Longitud, peso molecular, punto isoeléctrico (IP):**

25 1136 AA, 129 kDa, IP: 8,07

- **Secuencia de nucleótidos:**

**atgggttatcgtgcagaacatccaggggcaaatggtacatcaggccatatcacctagaact
 ttaaatgcatgggtaaaagtagtagagaaggaaggcttcagccagaagtaataccatg
 ttttcagcattatcagaaggagccaccccacaagatttaaacaccatgctaaacacagtg**

gggggacatcaagcagccatgcaaagtgttaaagagaccatcaatgaggaagctgcagaa
 tgggatagagtacatccagtgcatgcagggcctatgcaccaggccagatgagagaacca
 aggggaagtgcacatagcaggaactactagtaccctcaggaacaaataggatggatgaca
 aataatccacctatcccagtaggagaaattataaaagatggataatcctgggattaat
 aaaatagtaagaatgtatagccctaccagcattctggacataagacaaggaccaaagaa
 ccttttagagactatgttagaccggttctataaaactcctaagagccgagcaagcttcacag
 gaggtaaaaaatggatgcacagaaacctgtttgggtccaaaatgcgaaccagattgtaag
 actattttaaagcattgggaccagcggctacactagaagaaatgatgcacagcatgtcag
 ggagtaggaggaccggccataaggcaagagttttgcatatgggccccattagccctat
 tgagactgtgtcagtaaaattaaagccaggaatggatggccccaaaagttaaacaatggcc
 attgacagaagaaaaataaaagcattagttagaaatttgtacagagatggaaaaggaagg
 gaaaatttcaaaaattgggacctgaaaatccatacaatactccagatttggccataaagaa
 aaaagacagtactaaatggagaaaattagttagatttcagagaacttaataagagaactca
 agacttctgggaagttcaattaggaataccacatcccgcaggggttaaaaaagaaaaatc
 agtaacagtactggatgtgggtgatgcataatcttccagttcccttagatgaagacttcag
 gaaatatactgcatttaccatacctagtataaacaatgagacaccagggttagatatca
 gtacaatgtgcttcacagggatggaaaggatcaccagcaatattccaaagtagcatgac
 aaaaatccttagagccttttagaaaacaaaatccagacatagttatctatcaatacatgga
 tgatttgtatgtaggatctgacttagaaatagggcagcatagaacaaaaatagaggagct
 gagacaacatctgttgagggtgggacttaccacaccagacaaaaaacatcagaaagaacc
 tccattccttaaaatgggttatgaactccatcctgataaatggacagtacagcctatagt
 gctgccagaaaaagacagctggactgtcaatgacatacagaagttagtggggaaattgaa
 ttgggcaagtcagatttaccagggatataagtaaggcaattatgtaaactccttagagg
 aaccaagcactaacagaagtaataccactaacagaagaagcagagctagaactggcaga
 aaacagagagattctaaaagaaccagtagatggagtgattatgacccatcaaagactt
 aatagcagaaatacagaagcaggggcaaggccaatggacatatcaatttatcaagagcc
 atttaaaaatctgaaaacaggaaaatatgcacgtaaacgcgggtgccacactaatgatgt
 aaaacaattaacagaggcagtgcaaaaaataaccacagaaaagcatagtaatatggggaaa
 gactcctaaatttaaactgccatacaaaaggaaacatgggaaacatggtggacagagta
 ttggcaagccacctggattcctgagtgggagtttgttaatacccctccttttagtgaatt
 atgggtaccagttagagaaagaaccatagtaggagcagaaaccttctatgtagatggggc
 agctaacaggggagactaaattaggaaaagcaggatagtactaatagaggaagacaaaa
 agttgtcacccctaactgacacaacaaatcagaagactgagttacaagcaatttatctagc
 tttgcaggattcgggattagaagtaaacatagtaaacagactcacaatatgcattaggaat
 cattcaagcacaaccagatcaaagtgaatcagagttagtcaatcaataatagagcagtt
 aataaaaaaggaaaaggctctatctggcatgggtaccagcacacaaaggaattggaggaaa
 tgaacaagtagataaattagttagtgctggaatcaggaaagtgctagctatgggtggca
agtggtcaaaaagtagtggtggatggcctactgtaagggaaagaatgagacgagctg
agccagcagcagatggggtgggagcagcatctcgagacctggaaaaacatggagcaatca
caagtagcaatacagcagctaccaatgctgcttgtgcctggctagaagcacaagaggagg
aggaggtgggtttccagtcacacctcaggtaccttaagaccaatgact
tacaaggcagctgtagatcttagccactttttaaagaaaaggggggactggaagggcta
atcactcccaacgaagacaagatatccttgatctgtggatctaccacacacaaggctac

ttccctgattggcagaactacacaccagggccaggggtcagatatccactgacctttgga
tgggtgctacaagctagtaccagttgagccagataaggtagaagaggccaataaaggagag
aacaccagcttgttacaccctgtgagcctgcatggaatggatgacacctgagagagaagtg
ttagagtggaggtttgacagccgcctagcatttcacacgtggcccagagagctgcatccg
gagtactcaagaactgcagggcctatgggtgcgagagcgtcagtattaagcgggggaga
 attagatcgatgggaaaaaattcggttaaggccaggggaaagaaaaaataaaattaa
 acatatagtatgggcaagcaggagctagaacgattcgcagttaatcctggcctgttaga
 aacatcagaaggctgtagacaaatactgggacagctacaaccatcccttcagacaggatc
 agaagaacttagatcattatataatacagtagcaaccctctattgtgtgcatcaaaggat
 agagataaaagacaccaaggaagctttagacaagatagaggaagagcaaaacaaaagtaa
 gaaaaagcacagcaagcagcagctgacacaggacacagcaatcaggtcagccaaaatta
 ctaa

[Nº ID SEC: 14]

La secuencia de p24 está en negrita

La secuencia de Nef está subrayada

5 Cajas: nucleótidos introducidos mediante construcción genética

• **Secuencia de aminoácidos**

MVIVQNIQGQMVHQAI	SPRTLNAWVKVVEEKAFSPEV	IPMFSALSEGATP	50	
QDLN	TMLNTVGGHQAAMQMLKET	INEEAAEWDRVHPVHAGPI	APGQ	100
RGSDIAGTTSTLQEQIGWMTN	NPPI	PVGEIYKRWIILGLNKIVR	MYSP	150
ILDIRQGPKEPFRDYVDRFYKTL	RAEQASQEVKNWMTETLLVQ	NANPDCK	200	
TILKALGPAATLEEMMTACQGV	GGPGHKARVLM	GPISPIETVSVK	LKPG	250
MDGPKVKQWPLTEEKIKALVEI	CTEMEKEGKISKIGPENPY	NTPVFAIKK	300	
KDSTKWRKLVDFRELNKRTQDF	WEVQLGIPHPAGLKKKSV	TVLDVGDAY	350	
FSVPLDEDFRKYTAFTIP	SINNETPGIRYQYNVLPQGW	KGSPAIFQSSMT	400	
KILEPFRKQNPDIVIYQY	MDDLIVGSDLEIGQHR	TKIEELRQHLLRWGLT	450	
TPDKKHQKEPFLM	MGYELHPDKWTVQPIVL	PEKDSWTVNDIQKLVGKLN	500	
WASQIYPGIKVRQLCKLLRGT	KALTEVIPLEEEAELELAEN	REILKEPVH	550	
GVYYDPSKDLIAEIQKQGQ	QWTYQIYQEPFKNLKTGKY	ARKRGAHTNDV	600	
KQLTEAVQKITTESI	VIWIKTPKFKLP	IQKETWETWWTEYWQATWI	PEWE	650
FVNTPLVKLWYQLEKEP	IVGAETFYVDGAANRETKL	GKAGYVTNRGRQK	700	
VVTLTDTTNQKTELQAIYLAL	QDSGLEVNIVTDSQYALGI	IQAQPDQSES	750	
ELVNQII	EQLIKKEKVYLAWVPAHKG	GIGNEQVDKLVSAGIRKVL	LAMGGK	800
WSKSSVVGWPTVRERMRAE	PAADGVGAASRDLEKHGAIT	SSNTAATNAA	850	
CAWLEAQEEEEVGFVPTPQV	PLRPMTYKAAVDLSHFLKE	KGGLEGLIHSQ	900	
RRQDILDWYHTQGYFPD	WQNYTPGPGVRYPLTFGW	CYKLVPEPKVE	950	
EANKGENTSLLHPVSLHG	MDDPEREVLEWRFDSRLAF	HVARELHPEYFK	1000	
NCRP	MGARASVLSGGELDRWEK	IRLRPGGKKKYK	LKHIVWASRELERFAV	1050
NPGLLETSEGCRQILGQLQ	PSLQTGSEELRSYNTVAT	LYCVHQRIEIKD	1100	
TKEALDKIEEQNKSKKKAQ	QAAADTGHSNQVSONY		1136	

[Nº ID SEC: 15]

Secuencia de P24: aminoácidos 1-232 (en negrita)

Secuencia de RT: aminoácidos 235-795

Secuencia de Nef: aminoácidos 798-1002

5 Secuencia de P17: aminoácidos 1005-1136

Cajas: aminoácidos introducidos mediante construcción genética

K (Lisina): en lugar de metionina (codón “de comienzo” interno)

K (Lisina) ~~M~~: en lugar de triptófano (W). Mutación introducida para eliminar la actividad del enzima.

Expresión de F4* en células B834(DE3):

10 Se indujo cepa recombinante F4* a 22° C durante 18 horas, en paralelo con la construcción no mutada F4. Se prepararon extractos brutos y se analizaron mediante gel tintado con Coomassie y Transferencia westernting.

Como se ilustra en la figura 8, se expresó F4* en un alto nivel (10% de proteína total), ligeramente mayor en comparación con F4 y desapareció la banda a 62 kDa pequeña.

Análisis Transferencia western:

15 *Reactivos:*

- anti p24 pool 3 Mabs (JC13.1, JC16.1, IG8.1.1)(dilución 1/5000)

- anti RT policlonal de conejo (O3L16 de conejo) (dilución: 1/10 000)

- anti Nef-Tat policlonal de conejo (388 de conejo) (dilución n 1/10 000)

- anticuerpo anticonejo conjugado con fosfatasa alcalina (dilución: 1/7500)

20 - anticuerpo antiratón conjugado con fosfatasa alcalina (dilución: 1/7500)

Ejemplo 5: Construcción y expresión de F3 y F3* (F3 mutado)

F3 (p17-p51-Nef) y F3* (p17-p51*-Nef) en las que el sitio de iniciación de metionina interno putativo es reemplazado por lisina.

Se podrían usar las fusiones F3 y F3* en combinación con p24.

25 **Construcción de plásmidos recombinantes:**

F3: Se excitó la secuencia que codifica p51 (como fragmento de ADN Scal y Stul) a partir de plásmido de expresión pET29a/p51 y se ligó en plásmido pET28b/p17-Nef, en el sitio Stul (localizado entre p17 y el gen de Nef), como una fusión en marco con secuencias de p17 y Nef. La construcción de fusión resultante p17-p51-Nef se denomina F3.

30 **F3*:** Se consiguió mutación del sitio de iniciación de metionina interno putativo usando el “sistema de mutagénesis dirigida al sitio Gene Tailor” (Invitrogen), generando la construcción F3*.

Se usaron plásmidos F3 y F3* para transformar células B834 (DE3).

Características de proteína recombinante:

Término N p17: 134a.a. - bisagra:2a.a. - p51/p51*: 426a.a. - bisagra:2a.a. - Nef: 206a.a. **Término C**

• **Longitud, peso molecular, punto isoeléctrico (IP)** 770 AA, 88.5 kDa, IP:8,58

35 • **Secuencia de nucleótidos (para F3*)**

```
atgggtgcgagagcgtcagtattaagcgggggagaattagatcgatgggaaaaaattcgg      60
ttaaggccagggggaaagaaaaatataaattaaaacatatagtatgggcaagcaggag      120
ctagaacgattcgcagttaatcctggcctgtagaaacatcagaaggctgtagacaaata      180
ctgggacagctacaaccatccctcagacaggatcagaagaacttagatcattatataat      240
```

acagtagcaaccctctattgtgtgcatcaaaggatagagataaaagacaccaaggaagct 300
 ttagacaagatagaggaagagcaaaaacaaaagtaagaaaaaagcacagcaagcagcagct 360
 gacacaggacacagcaatcaggtcagccaaaattacctcgacaggactGGTCCGATCTCT 420
 CCGATAGAAACAGTTTCGGTCAAGCTTAAACCAGGGATGGATGGTCCAAAGGTCAAGCAG 480
 TGGCCGCTAACGGAAGAGAAGATTAAGGCGCTCGTAGAGATTTGTACTGAAATGGAGAAG 540
 GAAGGCAAGATAAGCAAGATCGGGCCAGAGAACCCGTACAATACACCCGGTATTTGCAATA 600
 AAGAAGAAGGATTCAACAAAATGGCGAAAGCTTGTAGATTTTAGGGAACTAAACAAGCGA 660
 ACCCAAGACTTTTGGGAAGTCCAAGTAGGTATCCACATCCAGCCGGTCTAAAGAAGAAG 720
 AAATCGGTACACAGTCCTGGATGTAGGAGACGCATATTTTAGTGTACCGCTTGATGAGGAC 780
 TTCCGAAAGTATACTGCGTTTACTATAACCGAGCATAAACAATGAAACGCCAGGCATTTCGC 840
 TATCAGTACAACGTGCTCCCGCAGGGCTGGAAGGGGTCTCCGGCGATATTTAGAGCTCT 900
 ATGACAAAATACTTGAACCATTCCGAAAGCAGAATCCGGATATTGTAATTTACCAATAC 960
 ATGGACGATCTCTATGTGGGCTCGGATCTAGAAATTGGGCAGCATCGCACTAAGATTGAG 1020
 GAAGTGAAGCAACATCTGCTTCGATGGGGCCTCACTACTCCCGACAAGAAGCACCAGAAG 1080
 GAGCCGCGTTCCATAAGATGGGCTACGAGCTTCATCCGGACAAGTGGACAGTACAGCCG 1140
 ATAGTGCTGCCCCGAAAAGGATTCTTGGACCGTAAATGATATTCAGAAACTAGTCGGCAAG 1200
 CTTAACTGGGCCTCTCAGATTTACCCAGGCATTAAGGTCCGACAGCTTTGCAAGCTACTG 1260
 AGGGGAACTAAGGCTCTAACAGAGGTCATCCCATTAACGGAGGAAGCAGAGCTTGAGCTG 1320
 GCAGAGAATCGCGAAATTCCTAAGGAGCCGGTGCACAGGGTATACTACGACCCCTCCAAG 1380
 GACCTTATAGCCGAGATCCAGAAGCAGGGGCAGGGCCAATGGACGTACCAGATATATCAA 1440
 GAACCGTTTAAGAATCTGAAGACTGGGAAGTACGCGCGCAAACGAGGGGCTCATACTAAT 1500
 GATGTAAGCAACTTACGGAAGCAGTACAAAAGATTACTACTGAGTCTATTGTGATATGG 1560
 GGCAAGACCCCCAAAGTTCAAGCTGCCATACAGAAGGAAACATGGGAAACATGGTGGACT 1620
 GAATATTGGCAAGCTACCTGGATTCCAGAATGGGAATTTGTCAACACGCCCGCTGGTA 1680
 AAATGaggcctATgggtggcaagtgggtcaaaaagtagtgtgggttgatggcctactgta 1740
 agggaaagaatgagacgagctgagccagcagcagatgggggtgggagcagcatctcgagac 1800
 ctggaaaaacatggagcaatcacaagtagcaatacagcagctaccaatgctgcttgtgcc 1860
 tggctagaagcacaagaggaggaggagggtgggttttccagtcacacctcaggtacctta 1920
 agaccaatgacttacaaggcagctgtagatcttagccactttttaaagaaaagggggga 1980
 ctggaagggtcaattcactcccaacgaagacaagatataccttgatctgtggatctaccac 2040
 acacaaggctacttccctgattggcagaactacacaccagggccaggggtcagatatacca 2100
 ctgaccttggatgggtgctacaagctagtaccagttgagccagataaggtagaagaggcc 2160
 aataaaggagagaacaccagcttgttacacctgtgagcctgcatggaatggatgacct 2220
 gagagagaagtgttagagtggaggtttgacagccgcttagcatttcatcacgtggcccca 2280
 gagctgcatccggagtacttcaagaactgctaa 2213

[Nº ID SEC:16]

P17: secuencia en negrita

P51: secuencia en letras mayúsculas

5 Nef: secuencia en lestras minúsculas

Cajas: nucleótidos introducidos mediante construcción genética

- **Secuencia de aminoácidos (para F3)**

MGARASVLSGGELDRWEKIRLRPGGKKKYLKHIVWASRELERFAVNPGLLETSEGCRQI 60
LGQLQPSLQGTGSEELRSLYNTVATLYCVHQRIEIKDTRKALDKIEEBEQNKSKKRAQQAAA 120
DTGHSNQVSQNYLDRTGPI SPIETVSVKLPKGMGPKVKQWPLTEEKIKALVEICTEMEK 180
EGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGI PHPAGLKKK 240
KSVTVLDVGDAYFSVPLDEDFRKYTAFTI PSINNETPGIRYQYNVLPQGWKSPAI FQSS 300
MTKILEPFRKQNPDI VIYQYMDDLVYVGS DLEIGQHRTKIEELRQHLLRWGLTTPDKKHQK 360
EPPFLKMGYELHPDKWTVQPI VLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLL 420
RGTKALTEVI PLTEEALELAENREILKEPVHGVYDPSKDLIAEIQKQGQGWTYQIYQ 480
EPFKNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGKTPKFKLPIQKETWETWWT 540
EYQWATWIPEWEFVNT PPLVKLRPMGGKWSKSSVVGWPTVRERMRRRAEPAADGVGAASRD 600
LEKHGAI TSSNTAATNAACAWLEAQEEEEVGFVTPQVPLRPMTYKAAVDLSHFLKEKGG 660
LEGLIHSQRRQDILD LWIYHTQGYFPDWQNYTPGPGVRYPLTFGWICYKLVPEVDPKVEEA 720
NKGENTSLLHPVSLHGMDPPEREVLEWRFD SRLAFHHVARELHPEYFKNC 770

[Nº ID SEC: 17]

Secuencia de P17: aminoácidos 1-134 (en negrita)

Secuencia de P51: aminoácidos 137-562

5 Secuencia de Nef: aminoácidos 565-770

Cajas: aminoácidos introducidos mediante construcción genética

Metionina 494 reemplazada por lisina (K) en construcción F3*

K(Lisina)K: en lugar de triptófano (W). Mutación introducida para eliminar la actividad del enzima.

Expresión de F3 en células B834(DE3):

10 Se evaluaron el nivel de expresión de F3 y las solubilidades de proteína recombinante en paralelo con cepas de producción de (p24-p66-Nef-p17) de F4 y p17-Nef (F2).

Condiciones de inducción: crecimiento de células a 37 °C / inducidas a 30 °C (+IPTG 1mM), durante 3 horas.

Tampones de ruptura:

F4 : Tris/HCl 50 mM pH:8,0, NaCl 50 mM, EDTA 1mM, +/- DTT 1mM

15 F2: Tris/HCl 50 mM pH:8,0, NaCl 50 mM, EDTA 1mM, sin DTT

F3: Tris/HCl 50 mM pH:7,5, NaCl 50 mM, EDTA 1 mM, +/- DTT 1 mM

Análisis Transferencia western:

reactivos

- anti RT policlonal de conejo (PO3L16 de conejo) (dilución: 1/10 000)

20 - anti Nef-Tat policlonal de conejo (388 de conejo) (dilución 1/10 000)

- anticuerpo anti-conejo conjugado con fosfatasa alcalina (dilution: 1/7500)

Extractos celulares “frescos”

25 Se analizaron las fracciones celulares correspondientes a extractos brutos (T), agregados insolubles (P) y sobrenadante (S) en SDS-PAGE al 10% en condiciones reductoras. Como se ilustra en la figura 9, la proteína de fusión F3 se expresa en un alto nivel (10% de proteína total). Casi toda la F3 se recupera en la fracción soluble (S) de extractos celulares, aunque de 5 a 10% de producto F4 está ya asociado con la fracción de agregado. El modelo de WB es simplificado en comparación con F4.

Expresión de F3* en células B834(DE3):

30 Se indujo la cepa recombinante F3* a 37º C durante 3 horas, construida en paralelo con F3 no mutada. Se prepararon los extractos celulares brutos y se analizaron mediante gel tintado con Coomassie y Transferencia westernting. Como se ilustra en la figura 10, se expresa la proteína de fusión F3* en un nivel muy alto (de 10 a 20%

de proteína total). Hubo un modelo de WB simplificado en comparación con F3; había desaparecido una banda muy decaída a +/- 32 kDa (detectado sólo por WB).

Ejemplo 6: Construcción y expresión de F4(p51) y F4(p51)*

Se usó RT/p51 en la construcción de fusión F4 (en lugar de RT/p66).

5 **F4(p51) = p24-p51-Nef-p17**

F4(p51)* = p24-p51*-Nef-p17 - F4(p51) mutado: sitio de iniciación de metionina interno putativo (presente en la parte de RT) reemplazado por lisina, para simplificar adicionalmente el modelo del antígeno.

Construcción de plásmidos recombinantes:

10 **F4(p51):** Se amplificó la secuencia que codifica p51 mediante PCR a partir del plásmido de expresión pET29a/p51. Se incorporaron sitios de restricción dentro de los cebadores de PCR (NdeI y StuI en el extremo 5'. AvrII en el extremo 3' de la secuencia de codificación). Se clonó el producto de PCR en plásmido intermedio pGem-T y se secuenció. Se restringió el plásmido intermedio pGem-T/p51 mediante NdeI y AvrII y se ligó el fragmento p51 en el plásmido de expresión pET28b/p24-RT/p66-Nef-p17 restringido por NdeI y NheI (dando lugar a la escisión de la secuencia de RT/p66). La ligadura se llevó a cabo mediante combinación de reacciones de digestión a concentraciones apropiadas, en presencia de ligasa de ADN de T4. Se usó el producto de ligadura para transformar las células DH5α de E.coli. Se confirmó la verificación de la inserción de p51 dentro del marco de lectura translacional correcto (en lugar de RT/p66 en la fusión f4) mediante secuenciación de ADN. La construcción de fusión resultante p24-RT/p51-Nef-p17 se denominó F4(p51).

20 **F4(p51)*:** Se consiguió la mutación del sitio de iniciación de la metionina interno putativo (presente en RT/p51) con "sistema de mutagénesis dirigido al sitio GeneTailor " (Invitrogen), generando la construcción F4(p51)*.

Se usaron plásmidos de expresión F4(p51) y F4(p51)* para transformar las células B834(DE3).

Características de proteínas recombinantes:

Término N p24: 232a.a. - bisagra:4a.a. - p51/51*: 426a.a. - bisabra:3a.a. - Nef: 206a.a. -

bisagra:2a.a. - p17: 132a.a. - **Término C**

- 25
- **Longitud, peso molecular, punto isoeléctrico (IP):** 1005 AA, 114.5 kDa, IP: 8,47
 - **Secuencia de nucleótidos (para F4(p51)*)**

```

Atgggttatcgtgcagaacatccaggggcaaatggtacatcaggccatatcacctagaact      60
Ttaaattgcatgggtaaaagtagtagaagagaaggctttcagcccagaagtaatacccatg    120
Ttttcagcattatcagaaggagccaccccacaagatttaaacaccatgctaaacacagtg    180
Gggggacatcaagcagccatgcaaatgtttaaagagaccatcaatgaggaagctgcagaa    240
Tgggatagagtacatccagtgcagggcctattgaccaggccagatgagagaacca        300
Aggggaagtgcatagcaggaactactagtacccttcaggaacaaataggatggatgaca    360
Aataatccacctatcccagtaggagaaatttataaaagatggataatcctgggattaaat    420
    
```

Aaaatagtaagaatgtatagccctaccagcattctggacataagacaaggaccaaagaa 480
Ccttttagagactatgttagaccggttctataaaactetaagagccgagcaagcttcacag 540
Gaggtaaaaaattggatgacagaaaccttggtggccaaaatgccaaccagatgtgaag 600
Actattttaaagcattgggaccagcggctacactagaagaaatgatgacagcatgtcag 660
GgagtaggaggaccggccataaggcaagagttttgCATATGaggcct**GGTCCGATCTCT** 720
CCGATAGAAACAGTTTTCGGTCAAGCTTAAACCAGGGATGGATGGTCCAAAGGTCAAGCAG 780
TGGCCGCTAACGGAAGAGAAGATTAAGGCGCTCGTAGAGATTTGTACTGAAATGGAGAAG 840
GAAGGCAAGATAAGCAAGATCGGGCCAGAGAACCCGTACAATACCCGGTATTTGCAATA 900
AAGAAGAAGGATTCAACAAAATGGCGAAAGCTTGTAGATTTTAGGGAACATAACAAGCGA 960
ACCCAAGACTTTTGGGAAGTCCAAGTACTAGGTATCCACATCCAGCCGGTCTAAAGAAGAAG 1020
AAATCGGTACAGTCTGGATGTAGGAGACGCATATTTTAGTGTACCGCTTGATGAGGAC 1080
TTCCGAAAAGTATACTGCGTTTACTATACCGAGCATAAACAATGAAACGCCAGGCATTCGC 1140
TATCAGTACAACGTGCTCCCGCAGGGCTGGAAGGGGTCTCCGGCGATATTTAGAGCTCT 1200
ATGACAAAAATACTTGAACCATTCCGAAAGCAGAATCCGGATATTGTAATTTACCAATAC 1260
ATGGACGATCTCTATGTGGGCTCGGATCTAGAAATTTGGGCAGCATCGACTAAGATTGAG 1320
GAAGTGAAGCAACATCTGCTTCGATGGGGCCTCACTACTCCCGACAAGAAGCACAGAAAG 1380
GAGCCGCCGTTTCTAAAGATGGGCTACGAGCTTCATCCGGACAAGTGGACAGTACAGCCG 1440
ATAGTGCTGCCCCGAAAAGGATTCTTGACCCTAAATGATATTCAGAAACTAGTCGGCAAG 1500
CTTAACTGGGCCTCTCAGATTTACCCAGGCATTAAGTCCGACAGCTTTGCAAGCTACTG 1560
AGGGGAACAAAGGCTCTAACAGAGGTCAATCCATTAACGGAGGAAGCAGAGCTTGAGCTG 1620
GCAGAGAATCGCGAAATTTCTTAAGGAGCCGGTGCACAGGGTATACTACGACCCCTCCAAG 1680
GACCTTATAGCCGAGATCCAGAAGCAGGGGCAGGGCCAATGGACGTACCAGATATATCAA 1740
GAACCGTTTAAAGAATCTGAAGACTGGGAAGTACGCGCGCAAACGAGGGGCTCATACTAAT 1800
GATGTAAAGCAACTTACGGAAGCAGTACAAAAGATTACTACTGAGTCTATTGTGATATGG 1860
GGCAAGACCCCAAAGTTCAAGCTGCCATACAGAAGGAAACATGGGAAACATGGTGGACT 1920
GAATATTGGCAAGCTACCTGGATTCCAGAATGGGAATTTGTCAACACGCCGCCCTGGTA 1980
AAACTGgcccctaGCT**ATG**ggtggcaagtggtcaaaaagtagtggtggatggcctact 2040
Gtaagggaaagaatgagacgagctgagccagcagcagatggggtgggagcagcatctcga 2100
Gacctggaaaaacatggagcaatcacaagtagcaatacagcagctaccaatgctgcttgt 2160
Gcctggctagaagcacaagaggaggaggaggtgggtttccagtcacacctcaggtacct 2220
Ttaagaccaatgacttacaagcagctgtagatcttagccactttttaaagaaaagggg 2280
Ggactggaagggctaattcactccaacgaagacaagatccttgatctgtggatctac 2340
Cacacacaaggctacttccctgattggcagaactacacaccagggccaggggtcagatat 2400
Ccactgacctttggatggtgctacaagctagtaccagttgagccagataaggtagaagag 2460
Gccaataaaggagagaaacaccagcttggtacaccctgtgagcctgcatggaatggatgac 2520
Cctgagagagaagtggttagagtgagggttgacagccgcttagcatttcacacgtggcc 2580
CgagagctgcatccggagtacttcaagaactgAGGCCT**ATGGGTGCGAGAGCGTCAGTA** 2640
TTAAGCGGGGGAGAATTAGATCGATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAAGAAA 2700
AAATATAAAATTA AAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAAT 2760
CCTGGCCTGTTAGAAAACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCC 2820
CTTCAGACAGGATCAGAAGAAGTCTAGATCATTATATAATACAGTAGCAACCCTCTATTGT 2880
GTGCATCAAAGGATAGAGATAAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAG 2940
CAAAAACAAAAGTAAGAAAAAAGCACAGCAAGCAGCAGCTGACACAGGACACAGCAATCAG 3000

GTCAGCCAAAATTACTaa

3018

[Nº ID SEC: 18]

P24: secuencia en negrita

P51: secuencia en letras mayúsculas

Nef: secuencia en letras minúsculas

P17: secuencia subrayada

Cajas: nucleótidos introducidos mediante construcción genética

5 • **Secuencia de aminoácidos (para F4(p51)*)**

```

MVI VQNI QGQM VHQA I SPRTL NAWV KVV E EKAF S PEVI PMFS ALSE GAT PQDL NTML NTV      60
GGHQA AMQML KET INEEAA EWDR VHPV HAGPI APGQM REPRG SDIAGTT STLQE QIGWMT      120
NNPPI PVGEI YKRWI ILGLNK IVRMYSPTS ILDIR QGPK EPFRDY VDRFYKTLRAEQASQ      180
EVKNWMTETLLVQNaNPDKTILKALGPAATLEEMMTACQGVGGPGHKARVLHMRPGPI S      240
PIETVSVK LKPGMDGPKVKQWPLTEEKI KALVEI CTEMEKEGKISKIGPENPYNTPVFAI      300
KKKDKSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKSVTVLVDVGDAYFSVPLDED      360
FRKYTAFTI PSINNETPGIRYQYNVLPQGWKGS PAIFQSSMTKILEPFRKQNPDI VIYQY      420
MDDL YVGS DLEIGQHRTKI EELRQHLLRWGLTTPDKKHQKEPPFLKMGYELHPDKWTVQP      480
IVLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGTKALTEVIPLTEEALEL      540
AENREILKEPVHGVYDPSKDLIAEIQKQGQGWTYQIYQEPFKNLKTGKYARKRGHAHTN      600
DVKQLTEAVQKITTESI VIWGKTPKFKLPIQKETWETWTEYQATWIPEWEFVNTPPLV      660
KLALAMGGKWSKSSVVGWPTVRERMRAEPAADGVGAASRDLEKHGAITSSNTAATNAAC      720
AWLEAQEEEEVGFVPTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRQDILDWIIY      780
HTQGYFPDQWNYTPGPGVRYPLTFGWCYKLVPEPDKVEEANKGENTSLLHPVSLHGMD      840
PEREVLEWRFD SRLAFHHVARELHPEYFKNDRFMGARASVLSGGELDRWEKIRLRPGGKK      900
KYK LKH I V WASRELERFAVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYC      960
VHQRIEIKDTKEALDKIEEEQNKSKKKAQQAADTGHSNQVSQNY      1005
    
```

[Nº ID SEC:19]

P24: aminoácidos 1-232

P51: aminoácidos 237-662

10 Nef: aminoácidos 666-871

P17: aminoácidos 874-1005

K (Lisina): en lugar de metionina (codón "de comienzo" interno)

K (Lisina) **K**: en lugar de triptófano (W). Mutación introducida para eliminar la actividad del enzima.

Expresión de F4(p51) en células B834(DE3):

15 Se evaluaron el nivel de expresión de F4(p51) y la solubilidad de la proteína recombinante en paralelo con la cepa que expresa F4.

Condiciones de inducción: crecimiento de células a 37 °C/inducido a 22 °C (+IPTG 1mM), durante 19 horas.

Tampón de ruptura: Tris/HCl 50 mM pH:7,5, EDTA 1 mM, DTT 1 mM

Análisis Transferencia western:

20 reactivos

- anti RT policlonal de conejo (PO3L16 de conejo) (dilución: 1/10 000)

- anti Nef-Tat policlonal de conejo (388 de conejo) (dilución 1/10 000)

- anticuerpo anti-conejo conjugado con fosfatasa alcalina (dilución: 1/7500)

Se analizaron fracciones celulares correspondientes a extractos brutos (T), agregado insoluble (P) y sobrenadante (S) en SDS-PAGE en condiciones de reducción del 10%.

Como se ilustra en la figura 11, se expresa F4(p51) a un alto nivel (10% de proteína total), similar a F4. Casi toda la F4(p51) se recupera en la fracción soluble (S) de extractos celulares. Tras detección con un reactivo anti-Nef-tat reagent, el modelo de WB para F4(p51) mostró ser simplificado (reducción de productos truncados por debajo de +/- 60kDa).

Expresión de F4(p51)* en células B834(DE3):

Se indujo la cepa recombinante F4(p51)* a 22° C durante 18 horas, en paralelo con la construcción no mutada F4(p51), F4 y F4*. Se prepararon extractos celulares brutos y se analizaron mediante gel tintado con Coomassie y Transferencia westernting. Como se ilustra en la figura 12 se observó alta expresión de fusiones F4(p51) y F4(p51)*, representando al menos el 10% de la proteína total. Patrón de WB: reducción de productos truncados por debajo de +/- 60kDa. Además, para la construcción F4(p51)*, ha desaparecido la banda de 47kDa (debido al sitio de inicio interno).

Ejemplo 7: Purificación de F4, F4(p51)* y F4* - Procedimiento de purificación I

La proteína de fusión F4, que comprende los 4 antígenos de VIH-1 p24-RT-Nef-p17, se purificó a partir de homogeneizado de células de *E. coli* de acuerdo con el procedimiento de purificación I, que comprende las siguientes etapas principales:

- Precipitación de F4 con sulfato de amonio
- Cromatografía de intercambio de cationes en SO3 Fractogel (modo positivo)
- Cromatografía de interacción hidrófoba en octil-Sepharosa (modo positivo)
- Cromatografía de intercambio aniónico en Q-sepharosa FF (modo positivo)
- Cromatografía de filtración en gel Superdex 200 en presencia de SDS
- Diálisis y concentración

De forma adicional, se purificaron la proteína de fusión F4(p51)* (RT reemplazada por p51 optimizado con codón que porta una mutación adicional Met592Lys) y la proteína F4* (F4 que porta una mutación Met592Lys adicional) usando el mismo procedimiento de purificación I.

Cuantificación de la proteína

- Se determinó la proteína total usando el ensayo de Lowry. Antes de la medida de la concentración de proteína se dializan todas las muestras durante la noche frente a PBS, se usó SDS al 0,1% para eliminar las sustancias que interfieren (urea, DTT). Se usó BSA (Pierce) como patrón.

SDS-PAGE y transferencia western

- Se prepararon muestras en tampón de muestra para SDS-PAGE en condiciones no reductoras (+/- β-mercaptoetanol) y se calentó durante 5 minutos a 95° C.
- Se separaron las proteínas en geles SDS-poliacrilamida de 4 a 20% a 200 V durante 75 minutos usando geles de Tris-glicina de Novex pre-fundidos o geles Criterion (Bio-Rad), de un 1 mm de grosor.
- Se visualizaron las proteínas con azul de Coomassie R250.
- Para las transferencias western (WB), se transfirieron las proteínas desde el gel SDS a membranes de nitrocelulosa (Bio-Rad) a 4° C durante 1,5 horas a 100 V o durante la noche a 30 V.
- Se detectó F4 usando anticuerpos monoclonales contra los distintos antígenos, anti-p24, anti-Nef-Tat, anti-RT (a veces se usó una mezcla de anti-p24 y anti Nef-Tat para detectar un número máximo de bandas de proteína).
- Se unieron anticuerpos anti-ratón o anti-conejo conjugados con fosfatasa alcalina a los anticuerpos primarios y se visualizaron bandas de proteínas usando BCIP y NBT como los sustratos.

Transferencia western de anti-E. coli

- Se separaron 5 µg de proteína (Lowry) mediante SDS-PAGE y se transfirieron a membranes de nitrocelulosa como anteriormente.
- Se detectaron proteínas de célula huésped residuales usando anticuerpos de anti -E. coli policlonales. Se visualizaron bandas de proteína con la reacción de fosfatasa alcalina como anteriormente.

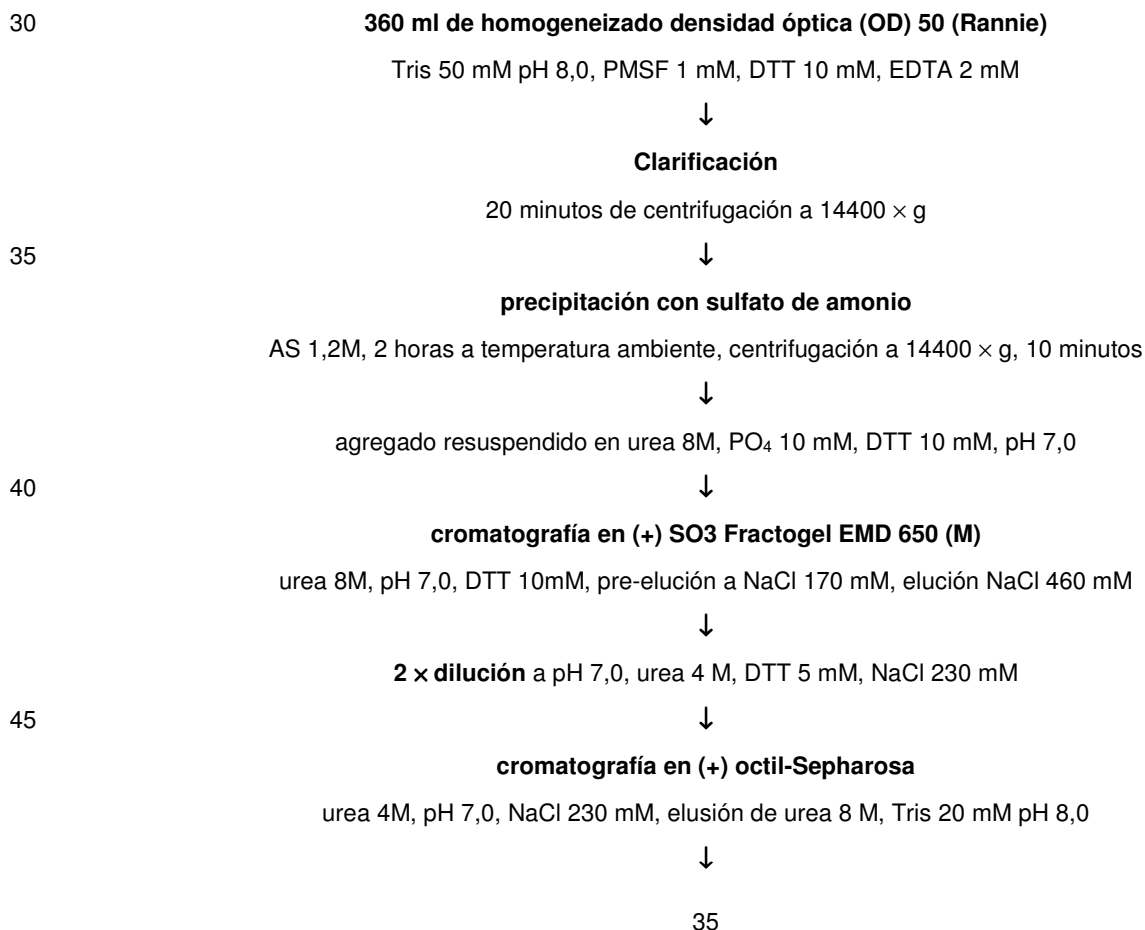
Procedimiento de purificación I

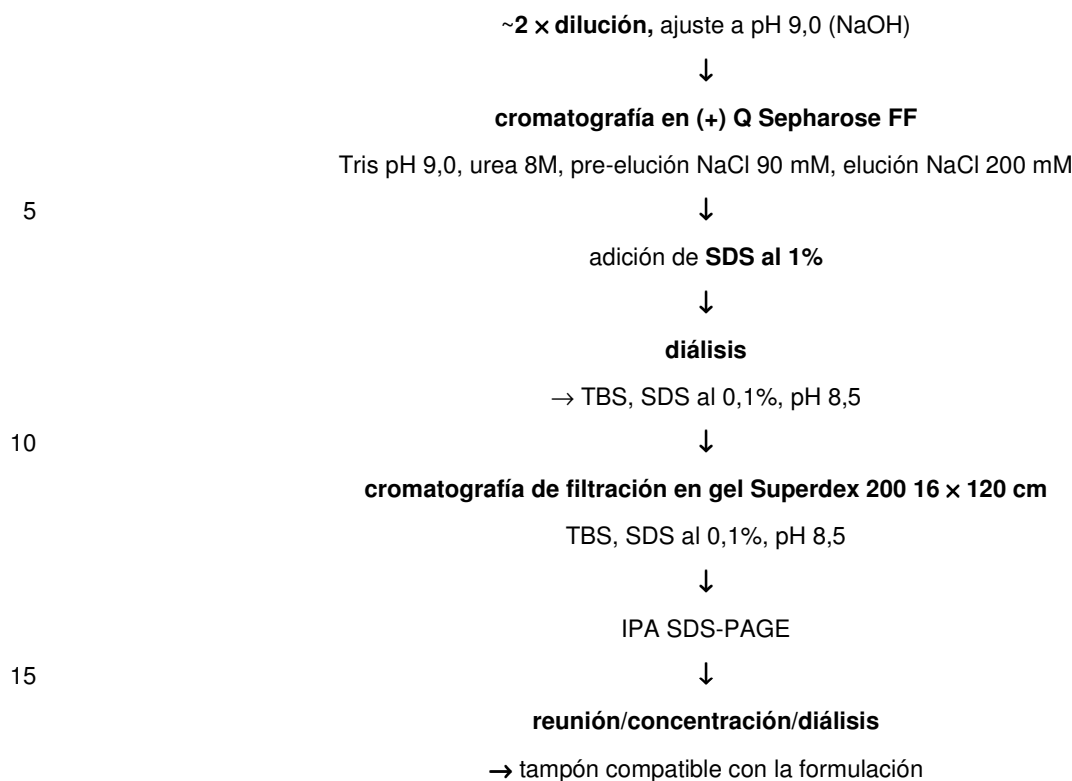
El procedimiento I comprende una precipitación con sulfato de amonio y cuatro etapas cromatográficas:

- Se homogenizaron células de E.coli en tampón Tris 50 mM a pH 8,0 en presencia de DTT 10 mM, PMSF 1 mM, EDTA 1 mM a una densidad óptica a 50 (~360 ml). Se aplicaron 2 pasadas Rannie a 1000 bar.
- Se eliminaron mediante centrifugación a 14400 × g durante 20 minutos células detriticas y el material insoluble.
- 5 ▪ Se añadió sulfato de amonio (AS) desde una solución stock 3,8M al sobrenadante clarificado hasta una concentración final de 1,2M. Se precipitaron proteínas durante ~2 horas a temperatura ambiente (RT) y luego se agregan mediante centrifugación (10 minutos a 14400 × g). Se resuspendió el agregado en urea 8M, DTT 10 mM en tampón fosfato 10 mM a pH 7,0.
- 10 ▪ Se capturó el antígeno en una columna de SO3 Fractogel (Merck) en presencia de urea 8M y DTT 10 mM a pH 7,0 en tampón fosfato. Se lavó la columna para eluir la proteína no unida seguido de una etapa de pre-elución con NaCl 170 mM para eliminar las proteínas de célula huésped unidas (HCP). Se eluyó luego F4 con NaCl 460 mM, urea 8M, DTT 10 mM en tampón fosfato a pH 7,0.
- 15 ▪ El eluido en SO3 se diluyó dos veces con tampón fosfato 10 mM, pH 7, y se cargó en una columna de octil-Sepharosa (Amersham Biosciences) en presencia de urea 4 M, DTT 1 mM, NaCl 230 mM en tampón fosfato a pH 7,0. Tras una etapa de lavado (tampón de equilibrado) se eluyó F4 unido con urea 8 M, DTT 1 mM en tampón Tris 25 mM a pH 8,0.
- 20 ▪ Se diluyó el eluido en octilo y se ajustó a pH 9,0 y se unió luego F4 a una columna de Q-sepharosa (Amersham Bioscience) en presencia de urea 8M a pH 9,0 (Tris 25 mM). Se lavó por arrastre la proteína no unida (urea 8M, Tris 25 mM a pH 9,0) y una etapa de pre-elución (NaCl 90 mM en urea 8M, Tris 25 mM, pH 9,0) eliminó las HCP y los productos de degradación de F4. Se desorbió F4 de la columna con NaCl 200 mM, urea 8M en tampón Tris a pH 9,0.
- 25 ▪ Se salpicó una alícuota de eluido en Q con SDS al 1% y se dializó frente a tampón PBS que contiene SDS al 0,1% y DTT 1 mM para eliminar la urea antes de inyectar la muestra en la columna de filtración en gel (Superdex 200 de calidad para preparación, dos columnas de 16 × 60 cm conectadas en serie). Se reunieron las fracciones relevantes tras análisis SDS-PAGE en proceso.
- Se dializaron las muestras dos veces a temperatura ambiente en membranas de diálisis (corte a 12-14 kDa) durante la noche frente a 1 l de arginina 0,5M, Tris 10 mM, glutatona 5 mM, pH 8,5.

Las etapas de purificación secuenciales se muestran en el flujograma siguiente.

Diagrama de flujo de purificación





IPA – Análisis en-proceso

Todos los tampones contienen **DTT 1 mM** si no es específica de otra forma.

20 Resultados de la purificación de F4

Seguimiento de SDS-PAGE/Transferencia western del procedimiento de purificación

La figura 13 muestra el gel SDS y la transferencia western de anti-p24/anti-Nef-Tat de las fracciones que contienen F4 recogidas durante la purificación de F4.

25 El homogeneizado de *E. coli* se muestra en la figura 3, carril 2, estimando que F4 representa aproximadamente el 10% de las proteínas totales (barridos de densidad de geles SDS tintados con azul de Coomassie). Tras centrifugación se recuperó la fracción soluble de F4 en el sobrenadante clarificado (carril 3). La etapa de precipitación con sulfato de amonio eliminó muchas impurezas (carril 4) y redujo la carga proteica para la etapa cromatográfica subsiguiente. De forma adicional, se usó urea 8M para resuspender los complejos disociados precipitados de F4 con HCP y permitió tanto la captura completa de F4 como la elución cuantitativa desde la resina SO3. El eluido en SO3 mostrado en la carril 5 estaba enriquecido considerablemente en F4 pero el modelo heterogéneo permaneció principalmente invariable. La columna de octil-Sepharosa hidrófoba eliminó principalmente las HCP de bajo peso molecular (LMW) y los productos de degradación de F4 (carril 6), con lo que se simplifica el modelo de F4. La cromatografía en Q-sepharosa simplificó adicionalmente el modelo de F4 y eliminó muchas impurezas (carril 7). Se obtuvo pureza final en términos de *E.coli* tras esta etapa. De hecho, no se detectaron proteínas de célula huésped en el eluido en Q mediante análisis transferencia western de anti-*E. coli*. La F4 purificada producida de esta forma se designó como F4Q. La columna de Superdex 200 separó los productos de degradación de F4 de bajo peso molecular (LMW) del F4 de longitud completa mejorando la homogeneidad de F4 en el eluido de Superdex 200 (carril 8). El término F4S se puede usar para designar a F4 purificada de acuerdo con el esquema completo del procedimiento I.

40 Se realizó una transferencia western de anti-*E. coli* de las mismas fracciones recogidas durante la purificación de F4. La ausencia de bandas visibles en la transferencia western de anti-*E. coli* indicó contaminación por HCP por debajo del 1% en el eluido en Q y en el eluido en Superdex.

Recuperación de F4 y proteína

45 La recuperación de F4 en cada etapa del procedimiento de purificación se estimó a partir de análisis por SDS-PAGE y transferencia western. Para estimar la recuperación de F4 en geles SDS, los volúmenes de muestra cargados en los geles SDS correspondían a los volúmenes de las distintas fracciones recogidas durante la purificación.

La **tabla 1** muestra la recuperación de proteína en las fracciones que contienen F4.

Tabla 1: recuperación de proteína en las fracciones positivas de F4 recogidas durante el procedimiento de purificación (360 ml de homogeneizado). Se determinó la concentración de proteína con el ensayo Lowry.

Etapas de purificación	Proteína (mg)	Recuperación en etapa (%)	Recuperación acumulada (%)
Homogeneizado	6500	100	100
Homogeneizado clarificado	4641	71	71,1
Precipitado con sulfato de amonio (AS) resuspendido	728	16	11
Eluido en SO3	247	34	3,8
Eluido en octil-Sepharosa	129	52	2,0
Eluido en Q-sepharosa	74	57	1,1
Superdex 200	36	49	0,6

5 La tabla muestra la cantidad de proteína en el homogeneizado y el material soluble, incluyendo F4, recuperado en el sobrenadante tras la etapa de clarificación. La etapa de precipitación con sulfato de amonio (AS) eliminó una gran cantidad de HCP y sólo se observó una ligera pérdida de F4 en el gel SDS. La cromatografía en SO3 eliminó adicionalmente muchas impurezas y el gel SDS indicó una gran recuperación de F4. Por el contrario, las recuperaciones de proteína de ~50% medida con las columnas de octil-Sepharosa y Q-sepharosa se acompañaron también de pérdidas de F4. La recuperación de proteína tras la cromatografía de filtración en gel fue de aproximadamente el 50%. El gel SDS muestra que se eliminaron muchas bandas de proteína LMW (bandas de degradación de F4), reduciendo concomitantemente la recuperación de F4.

Rendimiento en F4

15 La tabla 1 anterior muestra que se podría obtener aproximadamente 36 mg de F4 purificada a partir de 360 ml de homogeneizado a una densidad óptica (OD) de 50. Por tanto, 1 l de homogeneizado a densidad óptica (OD) de 50 debería dar aproximadamente 100 mg de F4 purificada. Debido a que se alcanzaron densidades ópticas de 70 – 90 durante el procedimiento de fermentación, el rendimiento por litro de fermento estaría, de acuerdo con lo anterior, en el intervalo de 140 a 180 mg de F4.

Resultados de la purificación de F4(p51)*

Se purificó la construcción de fusión F4(p51)* usando el procedimiento de purificación I descrito anteriormente sin modificaciones.

20 Seguimiento de SDS-PAGE/Transferencia western del procedimiento de purificación

La figura 14 muestra los análisis en gel SDS y transferencia western con anti-p24/anti-Nef-Tat de las fracciones que contienen F4(p51)* recogidas durante la purificación de F4(p51)*.

25 El gel SDS y el transferencia western demuestran que la proteína de fusión F4(p51)* se comporta globalmente de forma similar a F4 en la etapa de precipitación con sulfato de amonio así como también durante las etapas cromatográficas. F4(p51)* purificada presentaba un modelo de heterogeneidad similar a F4 purificada.

Una transferencia western de anti *E. coli* indicó que la contaminación con HCP estaba por debajo del 1% tanto en el eluido en Q como en el eluido en Superdex.

Rendimiento

30 Se perdieron aproximadamente el 25% de F4(p51)* en la fracción insoluble del homogeneizado. De forma adicional, debido a que el procedimiento de purificación no estaba adaptado a esta proteína, se observaron pérdidas en las etapas cromatográficas. Por lo tanto se redujo la recuperación global de F4(p51)* en aproximadamente 25 mg por litro de homogeneizado (densidad óptica, OD, 50). Extrapolado a 1 litro de cultivo a densidad óptica (OD) de 177, el rendimiento estaría, de acuerdo con lo anterior, en el intervalo de 85 mg de F4(p51)*.

Resultados de purificación de F4*

35 La construcción de fusión F4* se purificó usando el procedimiento de purificación I descrito anteriormente sin modificaciones.

Seguimiento de SDS-PAGE/Transferencia western del procedimiento de purificación

La figura 15 muestra el gel SDS y transferencia western con anti-p24/anti-Nef-Tat de las fracciones que contienen F4* recogidas durante la purificación de F4*.

Como con F4(p51)* también se puede señalar que F4* se comportó globalmente de forma bastante similar a F4 durante el procedimiento de purificación. Se recuperó la proteína en las fracciones esperadas como se muestra en el gel SDS y transferencia western. Una transferencia western de anti-*E. coli* demostró también la eliminación de la mayor parte de HCP ya después de la columna de Q-sepharosa.

5 Rendimiento

La recuperación global fue de aproximadamente 17 mg de F4* purificada obtenida a partir de 465 mg de homogeneizado de densidad óptica (OD) 50. Extrapolado a un 1 l de cultivo a densidad óptica (OD) de 140, el rendimiento estaría, de acuerdo con lo anterior, en el intervalo de 100 mg de F4*.

10 En resumen, se purificaron las tres proteínas de fusión F4, F4(p51)* y F4* usando el procedimiento de purificación I. El gel SDS en la figura 16 compara las tres proteínas purificadas mostrando el distinto nivel de heterogeneidad de las construcciones después de la etapa de Q-sepharosa y después de la eliminación de bandas de LMW (bajo peso molecular) mediante la columna de Superdex 200.

Ejemplo 8: Purificación de F4 y F4co (optimizada con codón) – Procedimiento de purificación II

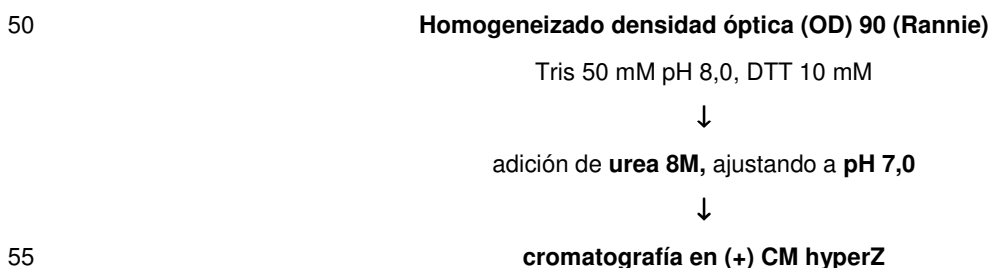
Procedimiento de purificación II

15 Se desarrolló también un procedimiento de purificación simplificado, el procedimiento II en comparación con el procedimiento I. El procedimiento II consiste únicamente en 2 etapas cromatográficas y diálisis/diafiltración final para intercambio de tampón. De forma reseñable, se introduce una columna cromatográfica CM hyperZ (BioSeptra) para reemplazar la etapa de clarificación, la precipitación con sulfato de amonio y la cromatografía en SO3 del procedimiento I (ejemplo 7). El procedimiento II se usó para purificar tanto F4 como F4 optimizada con codón completo ("F4co"). Para F4co, se desarrollaron dos formas distintas del procedimiento II, una que involucra la carboxiamidación y otra que no. La finalidad de la etapa de carboxiamidación era evitar la agregación por oxidación de la proteína. Esta carboxiamidación se lleva a cabo tras la primera etapa cromatográfica (CM hyperZ).

- 25 ▪ Se homogenizaron células de *E. coli* (que expresan F4 o F4co) en tampón Tris 50 mM a pH 8,0 en presencia de DTT 10 mM, a una densidad óptica (OD) de 90. Se aplicaron 2 pasadas por equipo Rannie a 1000 bar.
- 30 ▪ Se añadió urea 8M al homogeneizado antes de aplicación a la resina CM hyperZ (BioSeptra) equilibrada con urea 8M en tampón fosfato a pH 7. Se realizó la captura del antígeno en un modo discontinuo. Se empaquetó luego la resina en una columna, se lavaron por arrastre las proteínas no unidas con el tampón de equilibrado y se eliminaron las proteínas de célula huésped unidas (HCP) mediante una etapa de pre-elución con NaCl 120 mM. Se eluyó luego F4co con NaCl 360 mM, urea 8M, DTT 10 mM en tampón fosfato a pH 7,0.
- 35 ▪ Para controlar la agregación por oxidación de la proteína de fusión, se pueden carboxiamidar los grupos cisteína de F4co con yodoacetamida. Por tanto, se añade opcionalmente yodoacetamida 50 mM al eluido en CM hyperZ y se realizó la carboxiamidación durante 30 minutos a temperatura ambiente en la oscuridad.
- 40 ▪ El eluido en CM hyperZ se diluyó luego de forma adecuada (aproximadamente de 5 a 8 veces) y se ajustó a pH 9,0. Se unió luego F4co o F4coca a una columna de Q-sepharosa (Amersham Bioscience) en presencia de urea 8M en tampón Tris a pH 9,0. Se lavó por arrastre la proteína no unida con tampón de equilibrado y una etapa de pre-elución con NaCl 90 mM (sólo con proteína no carboxiamidada) en el mismo tampón eliminó la HCP unida. Se desorbió la F4co de la columna con NaCl 200 mM, urea 8M en tampón Tris a pH 9,0.
- 45 ▪ Se dializaron muestras dos veces a temperatura ambiente en membranas de diálisis (corte 12-14 kDa) durante la noche frente a 1 l de arginina 0,5M, tampón Tris 10 mM, glutatona 10 mM (añadido sólo a la proteína no carboxiamidada), pH 8,5. De forma alternativa, se consiguió el intercambio de tampón mediante diafiltración frente a 10 volúmenes de muestra del mismo tampón usando una membrana de flujo tangencial con corte a 30 o 50 kDa.
- 50 ▪ Finalmente, se filtró el producto dializado en condiciones de esterilidad a través de una membrana de 0,22 µm.

Se muestran las etapas de purificación secuenciales en el flujograma siguiente.

Diagrama de flujo de purificación



pH 7,0, urea 8M, DTT 10mM, pre-elución a NaCl 120 mM, elución en NaCl 360 mM



carboxiamidación opcional: adición de yodoacetamida 50 mM, 30 minutos a temperatura ambiente



5 **dilución y ajuste** a pH 9,0, urea 8M



cromatografía en (+) Q Sepharose FF

Tris pH 9,0, urea 8M, pre-elución, elución con NaCl*



10 **diálisis/diafiltración**

→ tampón fosfato, arginina 0,5M, pH 8,5 (glutaciona 10 mM)



Filtración en condiciones de esterilidad

15 Todos los tampones contenían DTT si F4co no estaba carboxiamidada y glutaciona en el volumen purificado. Se omitieron agentes reductores una vez que la proteína estaba carboxiamidada. *NaCl – para F4co esto fue NaCl 200 mM, para elución de F4coca elution fue mediante gradiente de NaCl. Esta etapa se puede optimizar adicionalmente para F4coca mediante pre-elución con NaCl 60 mM y eluyendo con NaC 100 mM; y para F4co mediante elución con NaCl 100 mM (no se necesita etapa de pre-elución).

Resultados: purificación de F4co

20 La figura 17 muestra un gel SDS de fracciones que contienen F4 recogidas durante la purificación de F4co y la purificación de F4co carboxiamidado ("F4coca").

25 La resina CM hyperZ capturó completamente F4co del homogeneizado bruto (carril 1) en presencia de urea 8M y se alcanzó elución cuantitativa con NaCl 360 mM. El eluido en CM hyperZ mostrado en la carril 2 estaba enriquecido considerablemente en F4co. Después de dilución apropiada y ajuste de la muestra a pH 9, se unió F4co o F4coca a una columna de Q-sepharosa. Se eluyó luego de forma específica F4co o F4coca con NaCl 200 mM como se muestra en la carril 3. Esta cromatografía no eliminó sólo las proteínas de célula huésped que quedan sino también el ADN y endotoxinas. Para llevar el material purificado a un tampón compatible con la formulación, el eluido de Q-sepharosa se dializó frente a tampón Tris 10 mM, arginina 0,5 M, glutaciona 10 mM pH 8,5 en una membrana de diálisis con corte a 12-14 kDa. Se omitió la glutaciona con la proteína carboxiamidada.

30 La purificación de F4co y F4coca dio aproximadamente 500 mg de material purificado por l de cultivo de densidad óptica (OD) 130. Este estaba en un intervalo similar al observado previamente con el F4 no optimizado con codón.

Como se describió anteriormente, se han desarrollado dos procedimientos de purificación distintos (I y II) para purificar las distintas construcciones de F4. La figura 18 compara los distintos volúmenes purificados que se obtuvieron.

35 El gel SDS en la figura 18 ilustra claramente el distinto modelo de las dos diferentes proteínas, F4 y F4co. Mientras que F4 presentaba varias bandas (LMW) gruesas de peso molecular bajo, sólo las bandas finas fueron visibles con la F4co optimizada con codón. El procedimiento I y procedimiento II dan lugar a un modelo para F4co muy similar. El análisis de transferencia western de anti-*E. coli* confirmó la pureza de las proteínas purificadas indicando contaminación con proteína de célula huésped por debajo del 1% en todas las preparaciones.

40 **Ejemplo 9: inmunogenicidad de F4 en ratones**

Formulación:

Formulación de adyuvante 1B:

45 Para preparar la formulación adyuvante 1B, se seca en vacío una mezcla de lípido (tal como fosfatidilcolina bien de yema de huevo o sintética) y colesterol y 3 D-MPL en disolvente orgánico (o de forma alternativa en una corriente de gas inerte). Se añade luego una solución acuosa (tal como solución salina tamponada con fosfato), y se agita el recipiente hasta que todo el lípido está en suspensión. Se microfluidiza luego esta suspensión hasta que el tamaño del liposoma se reduzca hasta aproximadamente 100 nm, y luego se filtra en condiciones de esterilidad a través de un filtro de 0,2 µm. Se podría reemplazar esta etapa por extrusión o ultrasonidos.

50 Típicamente la relación de colesterol:fosfatidilcolina es 1:4 (en relación peso/peso), y se añade la solución acuosa para dar una concentración en colesterol final de 5 a 50 mg/ml.

Los liposomas tienen un tamaño definido de 100 nm y se designan como SUV (para vesículas unilamelares

pequeñas). Si esta solución se congela y se descongela repetidamente las vesículas se condensan para formar estructuras multilamelares grandes (MLV) de tamaño que varía de 500 nm a 15 µm.

Los liposomas son estables por sí mismos en el tiempo y no tienen capacidad fusogénica.

5 Se añade QS21 en solución acuosa a los liposomas para alcanzar concentraciones finales de 3 D- MPL y QS21 de 100 µg/ml.

Formulación 2A: monofosforil lípido A 3-desacilado y QS21 en una emulsión aceite en agua;

La preparación de la emulsión aceite en agua se puede realizar siguiendo el protocolo que se describe en el documento WO 95/17210. La emulsión contiene en detalle: escualeno al 5%, tocoferol al 5%, tween 80 al 2,0%; el tamaño de partícula es 180 nm.

10 Preparación de emulsión de aceite en agua (concentrado dos veces)

Se disuelve Tween 80 en solución salina tamponada con fosfato (PBS) para dar una solución al 2% en PBS. Para proporcionar 100 ml se agitan 5 g de emulsión doblemente concentrada de DL alfatocoferol y 5 ml de escualeno para mezclar completamente. Se añade 90 ml de solución PBS/Tween y se mezcla completamente. Se pasa luego la emulsión resultante a través de una jeringuilla y se microfluidiza finalmente usando un equipo microfluidics M110S. Las gotas de aceite resultante tienen un tamaño de aproximadamente 180 nm.

15 La emulsión en volumen estéril se añade a PBS para alcanzar una concentración final de 500 µl de emulsión por ml (en relación volumen/volumen). Se añade luego 3 D-MPL para alcanzar una concentración final de 100 µg. Se añade luego QS21 para alcanzar una concentración final de 100 µg por ml. Entre cada adición de componente se agita el producto intermedio durante 5 minutos.

20 F4Q no optimizado con codón, purificado de acuerdo con el procedimiento de purificación I, se diluyó en un tampón de fosfato/arginina pH 6,8. Se mezcló la dilución con dos adyuvantes concentrados diferentes (adyuvantes 2A y 1B) con el fin de obtener una formulación final de 40 µg/dosis de 500 µl de F4 en presencia de 290 (para adyuvante 2A) - 300 (para adyuvante 1B) mM de argnina, 50 µg de MPL y 50 µg de QS21. Se inyectaron en ratones 100 µl de cada formulación.

25 Se realizaron estudios de inmunogenicidad en ratones para evaluar las respuestas inmunes celulares y humorales a los cuatro antígenos encontrados dentro de F4 (p24, p17, RT y Nef).

Debido a la complejidad del antígeno F4, se inmunizaron dos veces ocho cepas de ratón, cada una con un antecedente genético diferente, en el día 0 y día 21 con 8 µg de proteína F4 adyuvantada preparada como se describió anteriormente, en un volumen de 100 µl. Se recogieron muestras de suero y bazo 14 días tras la última inmunización (día 35) para análisis de las respuestas humorales y celulares en cada uno de los cuatro componentes de F4 (p24, p17, RT y Nef), así como también F4.

35 Se caracterizaron las respuestas a anticuerpo totales mediante ELISA específico para p24, p17, RT, Nef y F4. La siguiente tabla, tabla 2, resume donde se observaron respuestas humorales específicas del antígeno en cada cepa. Los resultados indican la presencia o ausencia de anticuerpos en comparación con animales de control inmunizados con adyuvante sólo. Los resultados presentados son una compilación de dos experimentos separados pero idénticos. En la tabla, 2A se refiere al antígeno formulado con 3D-MPL y QS21 en una emulsión aceite en agua y 1B se refiere a antígeno formulado con 3D-MPL, QS21 y liposomas que contienen colesterol.

Tabla 2

Cepa de ratón	P17	P24	Nef	RT	F4
CB6F1	+/- +2A -1B	+	+	+	+
Balb/c	- +2A -1B	+	+	+	+
C3H	-	-	-	-	-
DBA	-	+	+	+	+
CBA	-	-	+/- +2A -1B	+	+
129Sv	-	+	+	+	+
B6D2F1	+/- +2A -1B	+	+	+	+

OF1	+	+	+	+	+
-----	---	---	---	---	---

+ = presencia de anticuerpos

- = ausencia de anticuerpos

5 Los ratones OF1 dieron respuestas de anticuerpo para los cuatro componentes de F4. Las respuestas observadas se muestran en la figura 19. +/- indica que la respuesta observada era débil o sólo se observó con uno de los dos adyuvantes. Por ejemplo, respuestas de p17 a ratones B6D2F1: +/- general con +2A y -1B significa que hay una respuesta con 2A (no débil) y ninguna con 1B. Respuestas p17 en ratones Balb/c: - general, con un +/- 2A y un - 1B, aquí el +/- significa que la respuesta con adyuvante 2A era débil.

10 Se caracterizaron respuestas celulares mediante tinción por citometría de flujo para expresión de CD4 y CD8, IFN γ y IL-2 (tinción de citoquina intracelular para expresión de IFN γ y IL-2), siguiendo reestimulación de células de bazo con péptidos específicos p24, p17, RT o Nef, usando colecciones de librerías de péptidos de 15 mers con solapamiento de 11 mer. Las respuestas de CD4 fueron la respuesta celular dominante observada. La siguiente tabla, tabla 3, resume dónde se observaron las respuestas de CD4+IL-2+ específicas del antígeno para cada cepa de ratón. De nuevo, esto se muestra como presencia o ausencia de una respuesta.

Tabla 3

Cepa de ratón	P17	P24	Nef	RT
CB6F1	- +/-2A -1B	+	+	+
Balb/c	-	+/- débil	+/- débil	+
C3H	+	+	-	+
DBA	+	+	+	+
CBA	+	+ +2A -1B	-	+
129Sv	+	+	-	+
B6D2F1	-	+	+	+
OF1	-	-	-	+

+ = presencia de CD4+IL-2+

- = ausencia de CD4+IL-2+

15 Ratones DBA dieron respuestas de CD4 a los cuatro componentes de F4. Las respuestas de CD4+IL-2+ y CD4+IFN γ + observadas para esta cepa de ratón se muestran en la figura 20.

20 En resumen, F4 formulada en cualquiera de las dos formulaciones de adyuvante es capaz de promover respuestas humorales y celulares a p24, p17, RT y Nef. Esto demuestra que cada región de F4 es inmunogénica en una situación *in vivo*.

También se describe en la presente invención:

1. Un polipéptido que comprende Nef o un fragmento o derivado inmunogénico de la misma, y p17 Gag y/o p24 Gag o fragmentos o derivados inmunogénicos de las mismas, en el que cuando ambos p17 y p24 Gag se encuentran presentes hay al menos un antígeno de VIH o fragmento inmunogénico entre ellos.
- 25 2. El polipéptido de acuerdo con el punto 1 que comprende además Pol o RT o un fragmento o derivado inmunogénico del mismo.
3. El polipéptido de acuerdo con el punto 1 ó el punto 2, en el que RT o fragmento inmunogénico es un fragmento en el que RT está truncada en el término C tal que le falta el dominio ARNasa H de terminal carboxi.
4. El polipéptido de acuerdo con el punto 3, en el que el fragmento de RT es el fragmento p51.
- 30 5. El polipéptido de acuerdo con alguna de las afirmaciones 2 a 4, en el que RT comprende una mutación en la posición 592 para reemplazar la metionina por otro residuo, por ejemplo, lisina.
6. El polipéptido de acuerdo con alguna de las afirmaciones 1 a 4, en el que la Nef es Nef de longitud completa.
7. Un polipéptido seleccionado entre uno de los siguientes:

1. p24 – RT – Nef – p17
2. p24 – RT* – Nef – p17
3. p24 – p51RT – Nef – p17
4. p24 – p51RT* - Nef – p17
- 5 5. p17-p51RT-Nef
6. p17-p51RT*-Nef
7. Nef-p17
8. Nef-p17 con conector
9. p17-Nef
- 10 10. p17 – Nef con conector

en los que * representa mutación de metionina₅₉₂ en RT a lisina

8. Un procedimiento para la purificación de un polipéptido de acuerdo con alguna de las afirmaciones 1 a 7, tal procedimiento comprende:
 - i) proporcionar una composición que comprende el polipéptido no purificado;
 - 15 ii) someter la composición a al menos dos etapas cromatográficas;
 - iii) de forma opcional carboxiamidación del polipéptido;
 - iv) llevar a cabo una etapa de intercambio en tampón para proporcionar la proteína en un tampón adecuado para una formulación farmacéutica.
9. El procedimiento de acuerdo con el punto 8, en el que no hay más de dos etapas cromatográficas.
- 20 10. El procedimiento de acuerdo con el punto 8 o el punto 9, en el que la carboxiamidación se lleva a cabo entre las dos etapas cromatográficas.
11. Una composición que comprende (i) un polipéptido que comprende Nef o un fragmento o derivado inmunogénico de la misma y p17 Gag o un fragmento o derivado inmunogénico del mismo, y (ii) p24 Gag o un fragmento o derivado inmunogénico del mismo; o (i) un polipéptido que comprende Nef o un fragmento o derivado inmunogénico de la misma y p24 Gag o un fragmento o derivado inmunogénico del mismo, y (ii) p17 Gag o un fragmento o derivado inmunogénico del mismo.
- 25 12. La composición de acuerdo con el punto 11 que comprende además RT o un fragmento o derivado inmunogénico de la misma.
13. Una composición que comprende (i) un polipéptido que comprende Nef o un fragmento o derivado inmunogénico de la misma y p17 y/o p24 Gag o fragmentos o derivados inmunogénicos de los mismos, en la que cuando ambas p17 y p24 Gag están presentes hay al menos un antígeno de VIH o fragmento o derivado inmunogénico entre ellos y (ii) un polipéptido p51 RT.
- 30 14. Una composición de acuerdo con una cualquiera de las afirmaciones 11 a 13, en la que (i) se selecciona de
 1. Nef - p17
 - 35 2. Nef - p17 con conector
 3. p17 - Nef
 4. p17 – Nef con conector
15. Una composición de acuerdo con alguna de las afirmaciones 12 a 14, en la que la RT o fragmento inmunogénico de la misma es un fragmento en el que RT está truncada en el término C de modo que le falta el dominio ARNasa H de terminal carboxi.
- 40 16. La composición de acuerdo con el punto 15, en la que el fragmento de RT es el fragmento p51.
17. La composición de acuerdo con alguna de las afirmaciones 11 a 16, en la que la RT comprende una mutación en la posición 592 tal que la metionina se reemplaza por otro residuo, por ejemplo, lisina.
18. La composición de acuerdo con alguna de las afirmaciones 11 a 17, en la que la Nef es Nef de longitud completa.
- 45 19. Un polipéptido que es una fusión de antígenos de VIH que comprende al menos cuatro antígenos de VIH o fragmentos o derivados inmunogénicos, en el que los cuatro antígenos o fragmentos son o se derivan de Nef, Pol y

Gag.

20. El polipéptido de acuerdo con el punto 19 en el que Gag está presente como dos componentes separados que están separados por al menos otro antígeno en la fusión.
21. El polipéptido de acuerdo con el punto 19 ó 20 en el que la Nef es Nef de longitud completa.
- 5 22. El polipéptido de acuerdo con alguna de las afirmaciones 19 a 21 en el que Pol es p66 o p51RT.
23. El polipéptido de acuerdo con alguna de las afirmaciones 19 a 22 en el que Gag es p17 y p24 Gag.
24. Un polinucleótido o polinucleótidos que codifica un polipéptido o composición de polipéptidos o proteína de fusión de acuerdo con alguna de las afirmaciones 1 a 23.
- 10 25. Un polipéptido de p51 RT o derivado del mismo o un polinucleótido que lo codifica, preferiblemente codón optimizado para expresión en *E. coli*.
26. Un polipéptido codificado por el polinucleótido de acuerdo con el punto 25.
27. Una composición farmacéutica que comprende un polipéptido o polinucleótido o composición de polipéptidos o polinucleótidos de acuerdo con cualquier afirmación previa o un polipéptido purificado de acuerdo con cualquier afirmación previa, junto con un vehículo o adyuvante farmacéuticamente aceptable.
- 15 28. Una composición farmacéutica de acuerdo con el punto 27, en la que el adyuvante es un adyuvante que induce Th1 tal como QS21 o 3D-MPL o una combinación de QS21 y 3D-MPL.

LISTADO DE SECUENCIAS

<110> GlaxoSmithKline Biologicals SA Abrecht, Helge Delchambre, Martine Marchand, Martine Mathy, Nathalie Louise Permanne, Philippe, Jean, Gervais, Ghislain voss, Gerald, Hermann

20

<120> Vacuna

<130> VB60990

<140> PCT/EP2005/008434

25

<141> 2005-08-03

<160> 19

<170> FastSEQ para windows version 4.0

30

<210> 1

<211> 3411

<212> ADN

<213> Secuencia artificial

35

<220>

<223> Fusión p24-RT-Nef-p17

<400> 1

atggttatcg	tgcagaacat	ccagggggcaa	atggtacatc	aggccatatc	acctagaact	60
ttaaattgcat	gggtaaaaagt	agtagaagag	aaggccttca	gcccagaagt	aatacccatg	120
ttttcagcat	tatcagaagg	agccacccca	caagatttaa	acaccatgct	aaacacagtg	180
gggggacatc	aagcagccat	gcaaatgtta	aaagagacca	tcaatgagga	agctgcagaa	240
tgggatagag	tacatccagt	gcatgcaggg	cctattgcac	caggccagat	gagagaacca	300
aggggaagtg	acatagcagg	aactactagt	acccttcagg	aacaaatagg	atggatgaca	360
aaataatccac	ctatcccagg	aggagaaaatt	tataaaagat	ggataatcct	gggattaaat	420
aaaatagtaa	gaatgtatag	ccctaccagc	attctggaca	taagacaagg	accaaaagaa	480
ccttttagag	actatgtaga	ccggttctat	aaaactctaa	gagccgagca	agcttcacag	540
gaggtaaaaa	attggatgac	agaaaccttg	ttggtccaaa	atgpgaacc	agattgtaag	600
actatttttaa	aagcattggg	accagcggct	acactagaag	aaatgatgac	agcatgtcag	660
ggagtaggag	gacccggcca	taaggcaaga	gttttgcata	tgggccccat	tagccctatt	720
gagactgtgt	cagtaaaatt	aaagccagga	atggatggcc	caaaagttaa	acaatggcca	780
ttgacagaag	aaaaataaaa	agcattagta	gaaatttgta	cagagatgga	aaaggaaggg	840
aaaaattca	aaattggcag	tgaaaatcca	tacaatactc	cagtatttgc	cataaagaaa	900
aaagacagta	ctaaatggag	aaaattagta	gatttcagag	aacttaataa	gagaactcaa	960
gacttctggg	aagtcaatt	aggaatacca	catcccgag	ggttaaaaaa	gaaaaaatca	1020
gtaacagtac	tggatgtggg	tgatgcata	ttttcagttc	ccttagatga	agacttcagg	1080
aaatagtag	catttaccat	acctagtata	aacaatgaga	caccagggat	tagatatcag	1140
tacaatgtgc	ttccacaggg	atggaaagga	tcaccagcaa	tattccaag	tagcatgaca	1200
aaaatcttag	agccttttag	aaaacaaaat	ccagacatag	ttatctatca	atacatggat	1260
gatttgtatg	taggatctga	cttagaaata	gggcagcata	gaacaaaaat	agaggagctg	1320
agacaacatc	tggtgaggtg	gggacttacc	acaccagaca	aaaaacatca	gaaagaacct	1380
ccattcctta	aaatgggtta	tgaactccat	cctgataaat	ggacagtaca	gcctatagtg	1440
ctgccagaaa	aagacagctg	gactgtcaat	gacatacaga	agttagtggg	gaaattgaat	1500
tgggcaagtc	agatttacc	agggattaaa	gtaaggcaat	tatgtaact	ccttagagga	1560
accaaaagc	taacagaagt	aataccacta	acagaagaag	cagagctaga	actggcagaa	1620
aacagagaga	ttctaaaaga	accagtacat	ggagtgtatt	atgaccatc	aaaagactta	1680
atagcagaaa	tacagaagca	ggggcaaggc	caatggacat	atcaaattta	tcaagagcca	1740
tttaaaaatc	tgaaaacagg	aaaatagca	agaatgaggg	gtgcccacac	taatgatgta	1800
aaacaattaa	cagagcgagt	gcaaaaaata	accacagaaa	gcatagtaat	atggggaaag	1860
actcctaaat	ttaaactgcc	catacaaaaag	gaaacatggg	aaacatggtg	gacagagtat	1920
tggcaagcca	cctggattcc	tgagtgggag	tttgtaata	cccctcctt	agtgaattta	1980
tggtagcagt	tagagaaaga	acccatagta	ggagcagaaa	ccttctatgt	agatggggca	2040
gctaacaggg	agactaaatt	agggaaagca	ggatatgtta	ctaataagag	aagacaaaaa	2100
gttgtcacc	taactgacac	aacaaatcag	aagactgagt	tacaagcaat	ttatctagct	2160
ttgcaggatt	cgggattaga	agtaaacata	gtaacagagt	cacaatatgc	attaggaatc	2220
attcaagcac	acccagatca	aagtgaatca	gagttagtca	atcaaataat	agagcagtta	2280
ataaaaaagg	aaaaggctta	tctggcatgg	gtaccagcac	acaaaggaat	tggaggaaat	2340
gaacaagtag	ataaattagt	cagtgtgga	atcaggaaag	tgctagctat	gggtggcaag	2400
tggtcaaaaa	gtagtgtgg	tggatggcct	actgtaaggg	aaagaatgag	acgagctgag	2460
ccagcagcag	atgggggtgg	agcagcatct	cgagacctgg	aaaaacatgg	agcaatcaca	2520
agtagcaata	cagcagctac	caatgctgct	tgtgcctggc	tagaagcaca	agaggaggag	2580
gaggtgggtt	ttccagtcac	acctcaggt	cctttaagac	caatgactta	caaggcagct	2640

gtagatctta	gccacttttt	aaaagaaaag	gggggactgg	aagggcta	tcactcccaa	2700
cgaagacaag	atatccttga	tctgtggatc	taccacacac	aaggctactt	ccctgattgg	2760
cagaactaca	caccagggcc	aggggtcaga	tatccactga	cctttggatg	gtgctacaag	2820
ctagtaccag	ttgagccaga	taaggtagaa	gaggccaata	aaggagagaa	caccagcttg	2880
ttacaccctg	tgagcctgca	tggaatggat	gacccctgaga	gagaagtgtt	agagtggagg	2940
tttgacagcc	gcctagcatt	tcatcacgtg	gcccagagagc	tgcatccgga	gtacttcaag	3000
aactgcaggc	ctatgggtgc	gagagcgtca	gtattaagcg	ggggagaatt	agatcgatgg	3060
gaaaaaattc	ggttaaggcc	agggggaaaag	aaaaaatata	aattaaaaca	tatagtatgg	3120
gcaagcaggg	agctagaacg	attcgcagtt	aatcctggcc	tgttagaac	atcagaaggc	3180
tgtagacaaa	tactgggaca	gctacaacca	tccttcagaa	caggatcaga	agaacttaga	3240
tcattatata	atacagtagc	aaccctctat	tgtgtgcatc	aaaggataga	gataaaagac	3300
accaaggaag	cttagacaaa	gatagaggaa	gagcaaaaaca	aaagtaagaa	aaaagcacag	3360
caagcagcag	ctgacacagg	acacagcaat	caggtcagcc	aaaattacta	a	3411

<210> 2

<211> 1136

<212> PRT

5 <213> Secuencia artificial

<220>

<223> Fusión p24-Rr-Nef-p17

<400> 2

```

Met Val Ile Val Gln Asn Ile Gln Gly Gln Met Val His Gln Ala Ile
 1      5      10      15
Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu Glu Lys Ala
      20      25      30
Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser Glu Gly Ala
      35      40      45
Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln
      50      55      60
Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu Ala Ala Glu
65      70      75      80
Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala Pro Gly Gln
      85      90      95
Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Thr Leu
      100      105      110
Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile Pro Val Gly
      115      120      125
Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg
130      135      140
Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly Pro Lys Glu
145      150      155      160
Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu Arg Ala Glu
      165      170      175
Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr Leu Leu Val
      180      185      190
Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala Leu Gly Pro
      195      200      205
Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly
210      215      220
Pro Gly His Lys Ala Arg Val Leu His Met Gly Pro Ile Ser Pro Ile
225      230      235      240
Glu Thr Val Ser Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val
      245      250      255
Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile
      260      265      270
Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu
      275      280      285
Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr
290      295      300
Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln
305      310      315      320
Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys
      325      330      335
Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser
      340      345      350
Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro
      355      360      365
Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu
      370      375      380
Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr
385      390      395      400
Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr
      405      410      415
Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln

```

5

420 425 430
 His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly
 435 440 445
 Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Lys
 450 455 460
 Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val
 465 470 475 480
 Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val
 485 490 495
 Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg
 500 505 510
 Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile
 515 520 525
 Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile
 530 535 540
 Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu
 545 550 555 560
 Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile
 565 570 575
 Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met
 580 585 590
 Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln
 595 600 605
 Lys Ile Thr Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe
 610 615 620
 Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr
 625 630 635 640
 Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro
 645 650 655
 Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala
 660 665 670
 Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly
 675 680 685
 Lys Ala Gly Tyr Val Thr Asn Arg Gly Arg Gln Lys Val Val Thr Leu
 690 695 700
 Thr Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile Tyr Leu Ala
 705 710 715 720
 Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr
 725 730 735
 Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Gln Ser Glu Ser Glu Leu
 740 745 750
 Val Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu
 755 760 765
 Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp
 770 775 780
 Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu Ala Met Gly Gly Lys
 785 790 795 800
 Trp Ser Lys Ser Ser Val Val Gly Trp Pro Thr Val Arg Glu Arg Met
 805 810 815
 Arg Arg Ala Glu Pro Ala Ala Asp Gly Val Gly Ala Ala Ser Arg Asp
 820 825 830
 Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr Ala Ala Thr Asn
 835 840 845
 Ala Ala Cys Ala Trp Leu Glu Ala Gln Glu Glu Glu Glu Val Gly Phe
 850 855 860
 Pro Val Thr Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys Ala Ala
 865 870 875 880
 Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu Glu Gly Leu
 885 890 895
 Ile His Ser Gln Arg Arg Gln Asp Ile Leu Asp Leu Trp Ile Tyr His
 900 905 910
 Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly Pro Gly
 915 920 925
 Val Arg Tyr Pro Leu Thr Phe Gly Trp Cys Tyr Lys Leu Val Pro Val
 930 935 940
 Glu Pro Asp Lys Val Glu Glu Ala Asn Lys Gly Glu Asn Thr Ser Leu
 945 950 955 960
 Leu His Pro Val Ser Leu His Gly Met Asp Asp Pro Glu Arg Glu Val
 965 970 975
 Leu Glu Trp Arg Phe Asp Ser Arg Leu Ala Phe His His Val Ala Arg
 980 985 990
 Glu Leu His Pro Glu Tyr Phe Lys Asn Cys Arg Pro Met Gly Ala Arg
 995 1000 1005
 Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp Glu Lys Ile Arg
 1010 1015 1020
 Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys Leu Lys His Ile Val Trp

```

1025          1030          1035          1040
Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro Gly Leu Leu Glu
1045          1050          1055
Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu Gln Pro Ser Leu
1060          1065          1070
Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn Thr Val Ala Thr
1075          1080          1085
Leu Tyr Cys Val His Gln Arg Ile Glu Ile Lys Asp Thr Lys Glu Ala
1090          1095          1100
Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys Lys Lys Ala Gln
1105          1110          1115          1120
Gln Ala Ala Ala Asp Thr Gly His Ser Asn Gln Val Ser Gln Asn Tyr
1125          1130          1135

```

<210> 3

<211> 3411

<212> ADN

5 <213> Secuencia artificial

<220>

<223> F4co

10 <400> 3

atgggtcattg	ttcagaacat	acagggccaa	atgggtccacc	aggcaattag	tccgcgaact	60
cttaatgcat	gggtgaaggt	cgtggaggaa	aaggcattct	ccccggagg	cattccgatg	120
ttttctgctg	tatctgaggg	cgcaacgccg	caagacctta	ataccatgct	taacacggta	180
ggcgggcacc	aagccgctat	gcaaatgcta	aaagagacta	taaacgaaga	ggccgccgaa	240
tgggatcgag	tgcacccggg	gcacgccggc	ccaattgcac	caggccagat	gcgagagccg	300
cgcggtctg	atattgcagg	aactacgtct	acccttcagg	agcagattgg	gtggatgact	360
aacaatccac	caatcccggg	cggagagatc	tataagaggt	ggatcatact	gggactaaac	420
aagatagttc	gcatgtattc	tccgacttct	atactggata	tacgccaaag	cccaaaggag	480
ccgttcaggg	actatgtcga	ccgattctat	aagacccttc	gcgcagagca	ggcatcccag	540
gaggtcaaaa	attggatgac	agaaactctt	ttgggtgcaga	atgccaatcc	ggattgtaaa	600
acaattttaa	aggctctagg	accggccgca	acgctagaag	agatgatgac	ggcttgctag	660
ggagtccgtg	gaccggggca	taaagccccg	gtcttacaca	tggggccgat	atctccgata	720
gaaacagttt	cggtaagctt	taaaccaggg	atggatggtc	caaagggtcaa	gcagtgcccg	780
ctaaccggaag	agaagattaa	ggcgtctcgt	gagatttgta	ctgaaatgga	gaaggaaggc	840
aagataagca	agatcggggc	agagaacccg	tacaatacac	cggtatttgc	aataaagaaa	900
aaggattcaa	caaaatggcg	aaagcttgta	gatttttaggg	aactaaacaa	gcgaacccaa	960
gacttttggg	aagtccaact	agggatccca	catccagccg	gtctaaagaa	gaagaaatcg	1020
gtcacagtcc	tggatgtagg	agacgcatat	tttagtgtag	cgcttgatga	ggacttccga	1080
aggtatactg	cgtttactat	accgagcata	aacaatgaaa	cgccaggcat	tcgctatcag	1140
tacaacgtgc	tcccgcaggg	ctggaagggg	tctccggcga	tatttcagag	ctgtatgaca	1200
aaaatacttg	aaccattccg	aaagcagaat	ccggatattg	taattttacca	atacatggac	1260
gatctctatg	tgggctcggg	tctagaaatt	gggcagcatc	gcactaagat	tgaggaactg	1320
aggcaacatc	tgcttcgatg	gggctcactc	actcccagca	agaagcacca	gaaggagccg	1380
ccgttcctaa	agatgggcta	cgagcttcat	ccggacaagt	ggacagtaca	gccgatagtg	1440
ctgcccgaaa	aggattcttg	gaccgtaaat	gatattcaga	aactagtcgg	caagcttaac	1500
tgggctcttc	agatttacc	aggcattaag	gtccgacagc	tttgcaagct	actgagggga	1560
actaaggctc	taacagaggt	catcccatta	acggaggaag	cagagcttga	gctggcagag	1620
aatcgcgaaa	ttcttaagga	gcccgtgcac	ggggtatact	acgacccctc	caaggacctt	1680
atagcccaga	tccagaagca	ggggcagggc	caatggacgt	accagatata	tcaagaaccg	1740
tttaagaatc	tgaagactgg	gaagtacgcg	cgcatgagag	gggctcatac	taatgatgta	1800
aagcaactta	cggaagcagt	acaaaagatt	actactgagt	ctattgtgat	atggggcaag	1860
accccaaagt	tcaagctgcc	catacagaag	gaaacatggg	aaacatgggt	gactgaaatg	1920
tggcaagcta	cctggattcc	agaatgggaa	tttgtcaaca	cgccgccact	tgtaagctt	1980
tggtaaccag	ttgaaaagga	gcccgatagta	ggggcagaga	ccttctatgt	cgatggcggc	2040
gcgaatcgcg	aaacgaagct	aggcaaggcg	ggatacgtga	ctaatagggg	ccgccaaaag	2100
gtcgtaaccc	ttacggatac	caccaatcag	aagactgaac	tacaagcgat	ttaccttgca	2160
cttcaggata	gtggcctaga	ggtcaacata	gtcacggact	ctcaatatgc	gcttggcatt	2220
attcaagcgc	agccagatca	aagcgaaagc	gagcttgtaa	accaaataat	agaacagctt	2280
ataaagaaag	agaaggtata	tctggcctgg	gtccccgctc	acaaggggat	tggcggcaat	2340
gagcaagtgg	acaagctagt	cagcgtctgg	attcgcgaag	ttcttgcgat	ggggggtaag	2400
tggctctaagt	ctagcgtagt	cggttggccg	acagtccgcg	agcgcgatgc	acgcgccgaa	2460
ccagccgcag	atggcgtggg	ggcagcgtct	agggatctgg	agaagcacgg	ggctataact	2520
tccagtaaga	cggcggcgac	gaacgccgca	tgcgcatggt	tagaagccca	agaagaggaa	2580
gaagtagggt	ttccggtaac	tccccaggtg	ccgttaaggc	cgatgacct	taaggcagcg	2640
gtggatcttt	ctcacttctt	taaggagaaa	ggggggctgg	agggcttaat	tcacagccag	2700
aggcgacagg	atattcttga	tctgtggatt	taccataacc	aggggtactt	tccggactgg	2760
cagaattaca	ccccggggcc	aggcgtgcgc	tatcccctga	cttccgggtg	gtgctacaaa	2820
ctagtcccag	tggaaccgga	caaggtcgaa	gaggctaata	agggcgagaa	cacttctctt	2880
cttcacccgg	taagcctgca	cgggatggat	gaccagaac	gagaggttct	agaatggagg	2940
ttcgactctc	gacttgcggt	ccatcacgta	gcacgcgagc	tgcattccaga	atatttcaag	3000
aactgccgcc	caatggggcg	cagggccagt	gtacttagtg	gcggagaact	agatcgatgg	3060
gaaaagatac	gcctacgccc	ggggggcaag	aagaagtaca	agcttaagca	cattgtgtgg	3120

gcctctcgcg	aacttgagcg	attcgcagtg	aatccagggc	tgcttgagac	gagtgaaggc	3180
tgtaggcaaa	ttctggggca	gctacagccg	agcctacaga	ctggcagcga	ggagcttcgt	3240
agtctttata	ataccgtcgc	gactctctac	tgcgttcatc	aacgaattga	aataaaggat	3300
actaaagagg	cccttgataa	aattgaggag	gaacagaata	agtcgaaaaa	gaaggcccag	3360
cagggccgccc	ccgacaccgg	gcacagcaac	caggtgtccc	aaaactacta	a	3411

<210> 4

<211> 1302

<212> ADN

5 <213> Secuencia artificial

<220>

<223> p51 RT

5 <400> 4

```

atgagtactg gtccgatctc tccgatagaa acagtttcgg tcaagcttaa accagggatg 60
gatgggtccaa aggtcaagca gtggccgcta acggaagaga agattaaggc gtcgtagag 120
atgtgtactg aaatggagaa ggaaggcaag ataagcaaga tcgggccaga gaaccctac 180
aatacaccgg tatttgcaat aaagaagaag gattcaacaa aatggcgaaa gctttagat 240
tttagggaac taaacaagcg aacccaagac ttttgggaag tccaactagg tatcccat 300
ccagccggtc taaagaagaa gaaatcggtc acagtcctgg atgtaggaga cgcataat 360
agtgtaccgc ttgatgagga cttccgaaag tatactgctt ttactatacc gagcataaac 420
aatgaaacgc caggcattcg ctatcagtac aacgtgctcc cgcagggtg gaaggggtct 480
ccggcgatat ttcagagctc tatgacaaaa atacttgaac cattccgaaa gcagaatccg 540
gatattgtaa tttaccaata catggacgat ctctatgtgg gctcggatct agaaattggg 600
cagcatcgca ctaagattga ggaactgagg caacatctgc ttcgatgggg cctcactact 660
cccgacaaga agcaccagaa ggagccgccg ttcctaaaga tgggctacga gcttcatccg 720
gacaagtgga cagtacagcc gatagtgtctg cccgaaaagg attcttgac cgtaaattgat 780
attcagaaac tagtcggcaa gcttaactgg gcctctcaga tttaccagg cattaaggct 840
cgacagcttt gcaagctact gaggggaact aaggctctaa cagaggatc cccattaacg 900
gaggaagcag agcttgagct ggcagagaat cgcgaaattc ttaaggagcc ggtgcacggg 960
gtatactacg acccctccaa ggaccttata gccgagatcc agaagcaggg gcagggccaa 1020
tggacgtacc agatataatca agaaccgttt aagaatctga agactgggaa gtacgcgcgc 1080
atgcgagggg ctcatactaa tgatgtaaag caacttacgg aagcagtaca aaagattact 1140
actgagtcta ttgtgatatg gggcaagacc ccaaagttca agctgcccac acagaaggaa 1200
acatgggaaa catggtggac tgaatatgg caagctacct ggattccaga atgggaattt 1260
gtcaacacgc cgccgctggt aaaactgagg cctgctagct aa 1302

```

<210> 5

<211> 433

<212> PRT

10 <213> Secuencia artificial

<220>

<223> p51 RT

15 <400> 5

Met Ser Thr Gly Pro Ile Ser Pro Ile Glu Thr Val Ser Val Lys Leu
 1 5 10 15
 Lys Pro Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu
 20 25 30
 Glu Lys Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met Glu Lys Glu
 35 40 45
 Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val
 50 55 60
 Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp
 65 70 75 80
 Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu
 85 90 95
 Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser Val Thr Val
 100 105 110
 Leu Asp Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Asp Phe
 115 120 125
 Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn Glu Thr Pro
 130 135 140
 Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Ser
 145 150 155 160
 Pro Ala Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu Pro Phe Arg
 165 170 175
 Lys Gln Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Asp Asp Leu Tyr
 180 185 190
 Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys Ile Glu Glu
 195 200 205
 Leu Arg Gln His Leu Leu Arg Trp Gly Leu Thr Thr Pro Asp Lys Lys
 210 215 220
 His Gln Lys Glu Pro Pro Phe Leu Lys Met Gly Tyr Glu Leu His Pro

225 230 235 240
 Asp Lys Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys Asp Ser Trp
 245 250 255
 Thr Val Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser
 260 265 270
 Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg
 275 280 285
 Gly Thr Lys Ala Leu Thr Glu Val Ile Pro Leu Thr Glu Glu Ala Glu
 290 295 300
 Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val His Gly
 305 310 315 320
 Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln Lys Gln
 325 330 335
 Gly Gln Gly Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn
 340 345 350
 Leu Lys Thr Gly Lys Tyr Ala Arg Met Arg Gly Ala His Thr Asn Asp
 355 360 365
 Val Lys Gln Leu Thr Glu Ala Val Gln Lys Ile Thr Thr Glu Ser Ile
 370 375 380
 Val Ile Trp Gly Lys Thr Pro Lys Phe Lys Leu Pro Ile Gln Lys Glu
 385 390 395 400
 Thr Trp Glu Thr Trp Thr Glu Tyr Trp Gln Ala Thr Trp Ile Pro
 405 410 415
 Glu Trp Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Arg Pro Ala
 420 425 430
 Ser

- <210> 6
- 5 <211> 1023
- <212> ADN

<213> Secuencia artificial

<220>

<223> Nef p17

5

<400> 6

```

atgggtggca agtgggtcaaa aagtagtggt gttggatggc ctactgtaag ggaaagaatg 60
agacgagctg agccagcagc agatgggggtg ggagcagcat ctcgagacct ggaaaaacat 120
ggagcaatca caagtagcaa tacagcagct accaatgctg cttgtgcctg gctagaagca 180
caagaggagg aggaggtggg tttccagtc acacctcagg tacctttaag accaatgact 240
tacaaggcag ctgtagatct tagccacttt ttaaaagaaa aggggggact ggaagggcta 300
attcactccc aacgaagaca agatatecctt gatctgtgga tctaccacac acaaggctac 360
ttccctgatt ggcagaacta cacaccaggg ccaggggtca gatatecact gacctttgga 420
tggtgctaca agctagtacc agttgagcca gataaggtag aagaggccaa taaaggagag 480
aacaccagct tgttacacc tgtagcctg catggaatgg atgaccctga gagagaagtg 540
ttagagtgga ggtttgacag ccgctagca tttcatcacg tggcccgaga gctgcatccg 600
gagtacttca agaactgcag gcctatgggt gcgagagcgt cagtattaag cgggggagaa 660
ttagatcgat gggaaaaaat tcggttaagg ccagggggaa agaaaaaata taaattaaaa 720
catatagtat gggcaagcag ggagctagaa cgattcgagc ttaatcctgg cctgttagaa 780
acatcagaag gctgtagaca aatactggga cagctacaac catcccttca gacaggatca 840
gaagaactta gatcattata taatacagta gcaaccctct attgtgtgca tcaaaggata 900
gagataaaag acaccaagga agctttagac aagatagagg aagagcaaaa caaaagtaag 960
aaaaaagcac agcaagcagc agctgacaca ggacacagca atcaggtcag ccaaattac 1020
taa 1023
    
```

<210> 7

10 <211> 340

<212> PRT

<213> Secuencia artificial

<220>

15 <223> Nef p17 (NP)

<400> 7

Met Gly Gly Lys Trp Ser Lys Ser Ser Val Val Gly Trp Pro Thr Val
 1 Arg Glu Arg Met 5 Arg Arg Ala Glu Pro Ala Ala Asp Gly Val 15 Gly Ala
 20 Ala Ser Arg Asp Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr
 35 Ala Ala Thr Asn Ala Ala Cys Ala Trp Leu Glu Ala Gln Glu Glu Glu
 50 Glu Val Gly Phe Pro Val Thr Pro Gln Val Pro 60 Leu Arg Pro Met Thr
 65 70 75 80

Tyr Lys Ala Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly
 85 Ile His Ser Gln Arg Arg Gln Asp Ile Leu Asp Leu
 100 Trp Ile Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr
 115 Pro Gly Pro Gly Val Arg Tyr Pro Leu Thr Phe Gly Trp Cys Tyr Lys
 130 Leu Val Pro Val Glu Pro Asp Lys Val Glu Glu Ala Asn Lys Gly Glu
 145 Asn Thr Ser Leu Leu His Pro Val Ser Leu His Gly Met Asp Asp Pro
 165 Glu Arg Glu Val Leu Glu Trp Arg Phe Asp Ser Arg Leu Ala Phe His
 180 His Val Ala Arg Glu Leu His Pro Glu Tyr Phe Lys Asn Cys Arg Pro
 195 Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp
 210 Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys Leu Lys
 225 His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro
 245 Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu
 260 Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn
 275 Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Ile Lys Asp
 290 Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys
 305 Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly His Ser Asn Gln Val
 325 330 335
 Ser Gln Asn Tyr
 340

<210> 8

<211> 1029

5 <212> ADN

<213> Secuencia artificial

<220>

<223> p17 Nef

10

<400> 8

ES 2 486 669 T3

```

atgggtgcga gagcgtcagt attaagcggg ggagaattag atcgatggga aaaaattcgg 60
ttaaggccag ggggaaagaa aaaatataaa ttaaaacata tagtatgggc aagcagggag 120
ctagaacgat tgcagattaa tcctggcctg ttagaacaat cagaaggctg tagacaaata 180
ctgggacagc tacaaccatc ccttcagaca ggatcagaag aacttagatc attatataat 240
acagtagcaa ccctctattg tgtgcatcaa aggatagaga taaaagacac caaggaagct 300
ttagacaaga tagaggaaga gcaaaacaaa agtaagaaaa aagcacagca agcagcagct 360
gacacaggac acagcaatca ggtcagccaa aattacctcg acaggcctat ggggtggcaag 420
tggtcacaaa gtagtgtggt tggatggcct actgtaaggg aaagaatgag acgagctgag 480
ccagcagcag atgggggtggg agcagcatct cgagacctgg aaaaacatgg agcaatcaca 540
agtagcaata cagcagctac caatgctgct tgtgcctggc tagaagcaca agaggaggag 600
gagggtgggtt ttccagtcac acctcaggta cctttaagac caatgactta caaggcagct 660
gtagatctta gccacttttt aaaagaaaag gggggactgg aagggctaatt cactcccaa 720
cgaagacaag atatccttga tctgtggatc taccacacac aaggctactt ccctgattgg 780
cagaactaca caccagggcc aggggtcaga tatccactga cctttggatg gtgctacaag 840
ctagtaccag ttgagccaga taaggtagaa gaggccaata aaggagagaa caccagcttg 900
ttacaccctg tgagcctgca tggaatggat gaccctgaga gagaagtgtt agagtggagg 960
tttgacagcc gcctagcatt tcatcacgtg gcccgagagc tgcattccgga gtacttcaag 1020
aactgctaa 1029

```

<210> 9

<211> 342

<212> PRT

5 <213> Secuencia artificial

<220>

<223> p17 Nef (PN)

10 <400> 9

```

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp
 1           5           10           15
Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys Leu Lys

```

ES 2 486 669 T3

			20					25				30						
His	Ile	Val	Trp	Ala	Ser	Arg	Glu	Leu	Glu	Arg	Phe	Ala	Val	Asn	Pro			
		35					40					45						
Gly	Leu	Leu	Glu	Thr	Ser	Glu	Gly	Cys	Arg	Gln	Ile	Leu	Gly	Gln	Leu			
	50					55					60							
Gln	Pro	Ser	Leu	Gln	Thr	Gly	Ser	Glu	Glu	Leu	Arg	Ser	Leu	Tyr	Asn			
	65				70					75					80			
Thr	Val	Ala	Thr	Leu	Tyr	Cys	Val	His	Gln	Arg	Ile	Glu	Ile	Lys	Asp			
				85					90					95				
Thr	Lys	Glu	Ala	Leu	Asp	Lys	Ile	Glu	Glu	Glu	Gln	Asn	Lys	Ser	Lys			
			100					105					110					
Lys	Lys	Ala	Gln	Gln	Ala	Ala	Ala	Asp	Thr	Gly	His	Ser	Asn	Gln	Val			
		115					120					125						
Ser	Gln	Asn	Tyr	Leu	Asp	Arg	Pro	Met	Gly	Gly	Lys	Trp	Ser	Lys	Ser			
	130					135					140							
Ser	Val	Val	Gly	Trp	Pro	Thr	Val	Arg	Glu	Arg	Met	Arg	Arg	Ala	Glu			
	145				150					155					160			
Pro	Ala	Ala	Asp	Gly	Val	Gly	Ala	Ala	Ser	Arg	Asp	Leu	Glu	Lys	His			
				165				170						175				
Gly	Ala	Ile	Thr	Ser	Ser	Asn	Thr	Ala	Ala	Thr	Asn	Ala	Ala	Cys	Ala			
			180					185					190					
Trp	Leu	Glu	Ala	Gln	Glu	Glu	Glu	Glu	Val	Gly	Phe	Pro	Val	Thr	Pro			
		195					200					205						
Gln	Val	Pro	Leu	Arg	Pro	Met	Thr	Tyr	Lys	Ala	Ala	Val	Asp	Leu	Ser			
	210					215					220							
His	Phe	Leu	Lys	Glu	Lys	Gly	Gly	Leu	Glu	Gly	Leu	Ile	His	Ser	Gln			
	225				230					235					240			
Arg	Arg	Gln	Asp	Ile	Leu	Asp	Leu	Trp	Ile	Tyr	His	Thr	Gln	Gly	Tyr			
				245					250					255				
Phe	Pro	Asp	Trp	Gln	Asn	Tyr	Thr	Pro	Gly	Pro	Gly	Val	Arg	Tyr	Pro			
			260					265					270					
Leu	Thr	Phe	Gly	Trp	Cys	Tyr	Lys	Leu	Val	Pro	Val	Glu	Pro	Asp	Lys			
		275					280					285						
Val	Glu	Glu	Ala	Asn	Lys	Gly	Glu	Asn	Thr	Ser	Leu	Leu	His	Pro	Val			
	290					295					300							
Ser	Leu	His	Gly	Met	Asp	Asp	Pro	Glu	Arg	Glu	Val	Leu	Glu	Trp	Arg			
	305				310					315					320			
Phe	Asp	Ser	Arg	Leu	Ala	Phe	His	His	Val	Ala	Arg	Glu	Leu	His	Pro			
				325					330					335				
Glu	Tyr	Phe	Lys	Asn	Cys													
			340															

<210> 10

<211> 1041

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Nef-conector-p17

10 <400> 10

ES 2 486 669 T3

```

atgggtggca agtgggtcaaa aagtagtggt gttggatggc ctactgtaag ggaaagaatg 60
agacgagctg agccagcagc agatgggggtg ggagcagcat ctcgagacct ggaaaaacat 120
ggagcaatca caagtagcaa tacagcagct accaatgctg cttgtgcctg gctagaagca 180
caagaggagg aggaggtggg tttccagtc acacctcagg tacctttaag accaatgact 240
tacaaggcag ctgtagatct tagccacttt ttaaaagaaa aggggggact ggaagggcta 300
attcactccc aacgaagaca agatatcctt gatctgtgga tctaccacac acaaggctac 360
ttccctgatt ggcagaacta cacaccaggg ccaggggtca gatatccact gacctttgga 420
tgggtgctaca agctagtacc agttgagcca gataaggtag aagaggccaa taaaggagag 480
aacaccagct tgttacacc tgtgagcctg catggaatgg atgaccctga gagagaagtg 540
ttagagtgga ggtttgacag ccgcctagca tttcatcacg tggcccgaga gctgcatccg 600
gagtacttca agaactgcag gcctggatcc ggtggcggcc ctatgggtgc gagagcgta 660
gtattaagcg ggggagaatt agatcgatgg gaaaaaattc ggttaaggcc agggggaaag 720
aaaaaatata aattaaaaca tatagtatgg gcaagcaggg agctagaacg attcgcagtt 780
aatcctggcc tgttagaaac atcagaaggc tgtagacaaa tactgggaca gctacaacca 840
tcccttcaga caggatcaga agaacttaga tcattatata atacagtagc aaccctctat 900
tgtgtgcatc aaaggataga gataaaagac accaaggaag ctttagacaa gatagaggaa 960
gagcaaaaca aaagtaagaa aaaagcacag caagcagcag ctgacacagg acacagcaat 1020
caggtcagcc aaaattacta a

```

<210> 11

<211> 346

<212> PRT

5 <213> Secuencia artificial

<220>

<223> Nef-conector-p17 (NLP)

10 <400> 11

Met Gly Gly Lys Trp Ser Lys Ser Ser Val Val Gly Trp Pro Thr Val
 1 Arg Glu Arg Met 5 Arg Arg Ala Glu Pro 10 Ala Ala Asp Gly Val Gly Ala
 20 Ala Ser Arg Asp Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr
 35 Ala Ala Thr Asn Ala Ala Cys Ala Trp Leu Glu Ala Gln Glu Glu Glu
 50 Glu Val Gly Phe Pro Val Thr Pro Gln Val Pro Leu Arg Pro Met Thr
 65 Tyr Lys Ala Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly
 85 Leu Glu Gly Leu Ile His Ser Gln Arg Arg Gln Asp Ile Leu Asp Leu
 100 Trp Ile Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr
 115 Pro Gly Pro Gly Val Arg Tyr Pro Leu Thr Phe Gly Trp Cys Tyr Lys
 130 Leu Val Pro Val Glu Pro Asp Lys Val Glu Glu Ala Asn Lys Gly Glu
 145 Asn Thr Ser Leu Leu His Pro Val Ser Leu His Gly Met Asp Asp Pro
 165 Glu Arg Glu Val Leu Glu Trp Arg Phe Asp Ser Arg Leu Ala Phe His
 180 His Val Ala Arg Glu Leu His Pro Glu Tyr Phe Lys Asn Cys Arg Pro
 195 Gly Ser Gly Gly Gly Pro Met Gly Ala Arg Ala Ser Val Leu Ser Gly
 210 Gly Glu Leu Asp Arg Trp Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys
 225 Lys Lys Tyr Lys Leu Lys His Ile Val Trp Ala Ser Arg Glu Leu Glu
 245 Arg Phe Ala Val Asn Pro Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg
 260 Gln Ile Leu Gly Gln Leu Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu
 275 Leu Arg Ser Leu Tyr Asn Thr Val Ala Thr Leu Tyr Cys Val His Gln
 290 Arg Ile Glu Ile Lys Asp Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu
 305 Glu Gln Asn Lys Ser Lys Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr
 325 Gly His Ser Asn Gln Val Ser Gln Asn Tyr
 340 345

<210> 12

<211> 1047

5 <212> ADN

<213> Secuencia artificial

<220>

<223> p17-conector-Nef

10

<400> 12

ES 2 486 669 T3

```

atgggtgcga gagcgtcagt attaagcggg ggagaattag atcgatggga aaaaattcgg 60
ttaaggccag ggggaaagaa aaaatataaa ttaaaacata tagtatgggc aagcagggag 120
ctagaacgat tcgcagttaa tcctggcctg ttagaacat cagaaggctg tagacaaata 180
ctgggacagc tacaaccatc ccttcagaca ggatcagaag aacttagatc attatataat 240
acagtagcaa ccctctattg tgtgcatcaa aggatagaga taaaagacac caaggaagct 300
ttagacaaga tagaggaaga gcaaaacaaa agtaagaaaa aagcacagca agcagcagct 360
gacacaggac acagcaatca ggtcagccaa aattacctcg acaggcctgg atccggtggc 420
ggtcctatgg gtggcaagtg gtcaaaaagt agtgtggtg gatggcctac tgtaagggaa 480
agaatgagac gagctgagcc agcagcagat ggggtgggag cagcatctcg agacctggaa 540
aaacatggag caatcacaag tagcaataca gcagctacca atgctgcttg tgcctggcta 600
gaagcacaag aggaggagga ggtgggtttt ccagtcacac ctcaggtacc ttaagacca 660
atgacttaca aggcagctgt agatcttagc cactttttaa aagaaaaggg gggactggaa 720
gggctaattc actccaacg aagacaagat atccttgatc tgtggatcta ccacacacaa 780
ggctacttcc ctgattggca gaactacaca ccagggccag gggtcagata tccactgacc 840
tttggatggg gctacaagct agtaccagtt gagccagata aggtagaaga ggccaataaa 900
ggagagaaca ccagcttggt acaccctgtg agcctgcatg gaatggatga ccctgagaga 960
gaagtgttag agtggaggtt tgacagccg ctagcatttc atcacgtggc ccgagagctg 1020

```

catccggagt acttcaagaa ctgctaa 1047

<210> 13

<211> 348

<212> PRT

5 <213> Secuencia artificial

<220>

<223> p17-conector-Nef (PLN)

10 <400> 13

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp
 1 10 15
 Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys Leu Lys
 20 25 30
 His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro
 35 40 45
 Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu
 50 55 60
 Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn
 65 70 75 80
 Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Ile Lys Asp
 85 90 95
 Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys
 100 105 110
 Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly His Ser Asn Gln Val
 115 120 125
 Ser Gln Asn Tyr Leu Asp Arg Pro Gly Ser Gly Gly Gly Pro Met Gly
 130 135 140
 Gly Lys Trp Ser Lys Ser Ser Val Val Gly Trp Pro Thr Val Arg Glu
 145 150 155 160
 Arg Met Arg Arg Ala Glu Pro Ala Ala Asp Gly Val Gly Ala Ala Ser
 165 170 175
 Arg Asp Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr Ala Ala
 180 185 190
 Thr Asn Ala Ala Cys Ala Trp Leu Glu Ala Gln Glu Glu Glu Val
 195 200 205
 Gly Phe Pro Val Thr Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys
 210 215 220
 Ala Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu Glu
 225 230 235 240
 Gly Leu Ile His Ser Gln Arg Arg Gln Asp Ile Leu Asp Leu Trp Ile
 245 250 255
 Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly
 260 265 270
 Pro Gly Val Arg Tyr Pro Leu Thr Phe Gly Trp Cys Tyr Lys Leu Val
 275 280 285
 Pro Val Glu Pro Asp Lys Val Glu Glu Ala Asn Lys Gly Glu Asn Thr
 290 295 300
 Ser Leu Leu His Pro Val Ser Leu His Gly Met Asp Asp Pro Glu Arg
 305 310 315 320
 Glu Val Leu Glu Trp Arg Phe Asp Ser Arg Leu Ala Phe His His Val
 325 330 335
 Ala Arg Glu Leu His Pro Glu Tyr Phe Lys Asn Cys
 340 345

<210> 14

<211> 3411

<212> ADN

5 <213> Secuencia artificial

<220>

<223> p24-RT*-Nef-p17 (F4*)

10 <400> 14

atggttatcg tgcagaacat ccaggggcaa atggtacatc aggccatatc acctagaact 60
 ttaaatgcat gggtaaaagt agtagaagag aaggctttca gcccagaagt aataccatg 120
 ttttcagcat tatcagaagg agccaccca caagatthaa acaccatgct aaacacagt 180
 gggggacatc aagcagccat gcaaatgtta aaagagacca tcaatgagga agctgcagaa 240
 tgggatagag tacatccagt gcatgcaggg cctattgcac caggccagat gagagaacca 300
 aggggaagtg acatagcagg aactactagt acccttcagg aacaaatagg atggatgaca 360
 aataatccac ctatcccagt aggagaaatt tataaaagat ggataatcct gggattaat 420
 aaaaatgtaa gaatgtatag ccctaccagc attctggaca taagacaagg accaaaagaa 480
 ccttttagag actatgtaga ccggttctat aaaactctaa gagccgagca agcttcacag 540
 gaggtaaaaa attggatgac agaaaacctg ttggtccaaa atgccaacc agattgtaag 600

actattttaa aagcattggg accagcggct aactagaag aatgatgac agcatgtcag 660
 ggagtaggag gacccggcca taaggcaaga gttttgcata tgggccccat tagccctatt 720
 gagactgtgt cagtaaaatt aaagccagga atggatggcc caaaagttaa acaatggcca 780
 ttgacagaag aaaaaataaa agcattagta gaaatthgtta cagagatgga aaaggaagg 840
 aaaatthcaa aaattggggc tgaaaatcca tacaatactc cagtattthgc cataaagaaa 900
 aaagcagta ctaaatggag aaaattagta gatttcagag aacttaataa gagaactcaa 960
 gacttctggg aagttcaatt aggaatacca catcccgcag ggttaaaaaa gaaaaaatca 1020
 gtaacagtac tggatgtggg tgatgcata ttttcagttc ccttagatga agacttcagg 1080
 aaatatactg catttaccat acctagtata aacaatgaga caccagggat tagatatcag 1140
 tacaatgtgc tccacaggg atggaaagga tcaccagcaa tattccaaag tagcatgaca 1200
 aaaatcttag agcctthtag aaaacaaaat ccagacatag ttatctatca atacatggat 1260
 gattgtatg taggatctga cttagaataa gggcagcata gaacaaaaat agaggagctg 1320
 agacaacatc tgttgagggt gggacttacc acaccagaca aaaaacatca gaaagaacct 1380
 ccattcctta aaatgggtta tgaactccat cctgataaat ggacagtaca gcctatagtg 1440
 ctgccagaaa aagacagctg gactgtcaat gacatacaga agttagtggt gaaattgaa 1500
 tgggcaagtc agatttacc agggattaaa gtaaggcaat tatgtaaact ccttagagga 1560
 accaaagcac taacagaagt aataccacta acagaagaag cagagctaga actggcagaa 1620
 aacagagaga ttctaagaaga accagtacat ggagtgatt atgacccatc aaaagactc 1680
 atagcagaaa tacagaagca ggggcaaggc caatggacat atcaaatthta tcaagagcca 1740
 tthaaaaatc tgaaaacagg aaaatagca cgtaaacgcg gtgccacac taatgatgta 1800
 aaacaatthaa cagaggcagt gcaaaaaata accacagaaa gcatagtaat atggggaaag 1860
 actcctaaat ttaaaactgcc catacaaaag gaaacatggg aaacatggtg gacagagtat 1920
 tggcaagcca cctggattcc tgagtgggag thgttaataa cccctccttt agtgaaata 1980
 tggtagcagt tagagaaaga acccatagta ggagcagaaa ccttctatgt agatggggca 2040
 gctaacaggg agactaaatt aggaaaagca ggatattgta ctaatagagg aagacaaaaa 2100
 gttgtcacc taactgacac aacaaatcag aagactgagt tacaagcaat thtatctagct 2160
 ttgcaggatt cgggattaga agtaaacata gtaacagact cacaatatgc attaggaatc 2220
 attcaagcac aaccagatca aagtgaatca gagttagta atcaataat agagcagtha 2280
 aaaaaaaagg aaaaggtcta tctggcatgg gtaccagcac acaaaggaat tggaggaaat 2340
 gaacaagttag ataaattag cagtgtcggg atcaggaaag tgctagctat gggtaggcaag 2400
 tggtaaaaa gtagtgtggg tggatggcct actgtaaggg aaagaatgag acgagctgag 2460
 ccagcagcag atgggggtgg agcagcatct cgagacctgg aaaaacatgg agcaatcaca 2520
 agtagcaata cagcagctac caatgctgct tgtgcctggc tagaagcaca agaggaggag 2580
 gaggtagggt tccagctcac acctcaggtta cctthtaagac caatgactta caaggcagct 2640
 gtagatctta gccactthtt aaaagaaaaag gggggactgg aagggttaat tcaactccca 2700
 cgaagacaag atatccttga tctgtggatc taccacacac aaggctactt ccctgattgg 2760
 cagaactaca caccagggcc aggggtcaga tatccactga cctthggatg gtgctacaag 2820
 ctagtaccag ttgagccaga taaggtagaa gaggccaata aaggagagaa caccagcttg 2880
 ttacaccctg tgagcctgca tggaaatgat gacctgaga gagaagtgtt agagtggagg 2940
 ttgacagcc gcctagcatt tcatcacgtg gcccagagc tgcattccgga gtacttcaag 3000
 aactgcaggc ctatgggtgc gagagcgtca gtattaagcg ggggagaatt agatcgatgg 3060
 gaaaaaatc ggtaaggcc agggggaaag aaaaaatata aattaaaaca tatagtatgg 3120
 gcaagcaggg agctagaacg attcgcagtt aatcctggcc tgttagaac atcagaaggc 3180
 ttagacaaa tactgggaca gctacaacca tcccttcaga caggatcaga agaacttaga 3240
 tcattatata atacagtac accctctat tgtgttcac aaaggataga gataaaagac 3300
 accaaggaag ctttagacaa gatagaggaa ggcacaaaca aaagtaagaa aaaagcacag 3360
 caagcagcag ctgacacagg acacagcaat caggctagcc aaaattacta a 3411

<210> 15

<211> 1136

5 <212> PRT

<213> Secuencia artificial

<220>

<223> p24-RT*-Nef-p17 (F4*)

5 <400> 15

```

Met Val Ile Val Gln Asn Ile Gln Gly Gln Met Val His Gln Ala Ile
 1          5          10          15
Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu Glu Lys Ala
          20          25          30
Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser Glu Gly Ala
          35          40          45
Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln
          50          55          60
Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu Ala Ala Glu
65          70          75          80
Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala Pro Gly Gln
          85          90          95
Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Thr Leu
          100          105          110
Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile Pro Val Gly
          115          120          125
Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg
130          135          140
Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly Pro Lys Glu

```

145 Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu Arg Ala Glu 160
 165 Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr Leu Leu Val 175
 180 Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala Leu Gly Pro 190
 195 Ala Ala Thr Leu Glu Glu Met Thr Ala Cys Gln Gly Val Gly Gly 205
 210 Pro Gly His Lys Ala Arg Val Leu His Met Gly Pro Ile Ser Pro Ile 220
 225 Glu Thr Val Ser Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val 235
 245 Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile 255
 260 Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu 270
 275 Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr 285
 290 Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln 300
 305 Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys 315
 325 Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser 330
 335 Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro 345
 350 Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu 355
 360 Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr 365
 370 Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr 375
 385 Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln 380
 390 His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly 395
 400 Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Lys 410
 415 Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val 425
 420 Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val 430
 435 Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg 445
 440 Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile 450
 455 Pro Leu Thr Glu Glu Ala Glu Val Tyr Tyr Asp Pro Ser Lys Asp Leu 460
 460 Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile 475
 470 Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Lys 480
 485 Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln 490
 495 Lys Ile Thr Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe 505
 510 Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Thr Thr Glu Tyr 515
 515 Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro 520
 520 Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala 525
 525 Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly 530
 530 Lys Ala Gly Tyr Val Thr Asn Arg Gly Arg Gln Lys Val Val Thr Leu 535
 535 Thr Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile Tyr Leu Ala 540
 540 Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr 545
 545 Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Gln Ser Glu Ser Glu Leu 550
 550 Val Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu 555

		755					760				765				
Ala	Trp	Val	Pro	Ala	His	Lys	Gly	Ile	Gly	Gly	Asn	Glu	Gln	Val	Asp
	770					775					780				
Lys	Leu	Val	Ser	Ala	Gly	Ile	Arg	Lys	Val	Leu	Ala	Met	Gly	Gly	Lys
785					790					795					800
Trp	Ser	Lys	Ser	Ser	Val	Val	Gly	Trp	Pro	Thr	Val	Arg	Glu	Arg	Met
			805						810					815	
Arg	Arg	Ala	Glu	Pro	Ala	Ala	Asp	Gly	Val	Gly	Ala	Ala	Ser	Arg	Asp
			820					825					830		
Leu	Glu	Lys	His	Gly	Ala	Ile	Thr	Ser	Ser	Asn	Thr	Ala	Ala	Thr	Asn
		835					840					845			
Ala	Ala	Cys	Ala	Trp	Leu	Glu	Ala	Gln	Glu	Glu	Glu	Glu	Val	Gly	Phe
	850					855					860				
Pro	Val	Thr	Pro	Gln	Val	Pro	Leu	Arg	Pro	Met	Thr	Tyr	Lys	Ala	Ala
865					870					875					880
Val	Asp	Leu	Ser	His	Phe	Leu	Lys	Glu	Lys	Gly	Gly	Leu	Glu	Gly	Leu
				885					890					895	
Ile	His	Ser	Gln	Arg	Arg	Gln	Asp	Ile	Leu	Asp	Leu	Trp	Ile	Tyr	His
			900					905					910		
Thr	Gln	Gly	Tyr	Phe	Pro	Asp	Trp	Gln	Asn	Tyr	Thr	Pro	Gly	Pro	Gly
		915					920					925			
Val	Arg	Tyr	Pro	Leu	Thr	Phe	Gly	Trp	Cys	Tyr	Lys	Leu	Val	Pro	Val
	930					935					940				
Glu	Pro	Asp	Lys	Val	Glu	Ala	Asn	Lys	Gly	Glu	Asn	Thr	Ser	Leu	
945					950				955						960
Leu	His	Pro	Val	Ser	Leu	His	Gly	Met	Asp	Asp	Pro	Glu	Arg	Glu	Val
				965					970					975	
Leu	Glu	Trp	Arg	Phe	Asp	Ser	Arg	Leu	Ala	Phe	His	His	Val	Ala	Arg
			980					985					990		
Glu	Leu	His	Pro	Glu	Tyr	Phe	Lys	Asn	Cys	Arg	Pro	Met	Gly	Ala	Arg
		995					1000					1005			
Ala	Ser	Val	Leu	Ser	Gly	Gly	Glu	Leu	Asp	Arg	Trp	Glu	Lys	Ile	Arg
	1010					1015					1020				
Leu	Arg	Pro	Gly	Gly	Lys	Lys	Lys	Tyr	Lys	Leu	Lys	His	Ile	Val	Trp
1025					1030					1035					1040
Ala	Ser	Arg	Glu	Leu	Glu	Arg	Phe	Ala	Val	Asn	Pro	Gly	Leu	Leu	Glu
				1045					1050					1055	
Thr	Ser	Glu	Gly	Cys	Arg	Gln	Ile	Leu	Gly	Gln	Leu	Gln	Pro	Ser	Leu
			1060					1065						1070	
Gln	Thr	Gly	Ser	Glu	Glu	Leu	Arg	Ser	Leu	Tyr	Asn	Thr	Val	Ala	Thr
		1075					1080					1085			
Leu	Tyr	Cys	Val	His	Gln	Arg	Ile	Glu	Ile	Lys	Asp	Thr	Lys	Glu	Ala
	1090					1095					1100				
Leu	Asp	Lys	Ile	Glu	Glu	Gln	Asn	Lys	Ser	Lys	Lys	Lys	Ala	Gln	
1105					1110				1115					1120	
Gln	Ala	Ala	Ala	Asp	Thr	Gly	His	Ser	Asn	Gln	Val	Ser	Gln	Asn	Tyr
				1125					1130					1135	

<210> 16

<211> 2313

<212> ADN

5 <213> Secuencia artificial

<220>

<223> p17-p51*-Nef (F3*)

10 <400> 16

atgggtgcga gagcgtcagt attaagcggg ggagaattag atcgatggga aaaaattcgg 60
 ttaaggccag ggggaaagaa aaaatataaa ttaaaacata tagtatgggc aagcagggag 120
 ctagaacgat tgcgagttaa tcctggcctg ttagaaacat cagaaggctg tagacaaata 180
 ctgggacagc tacaaccatc ccttcagaca ggatcagaag aacttagatc attatataat 240
 acagtagcaa cctctatttg tgtgcatcaa aggatagaga taaaagacac caaggaagct 300
 ttagacaaga tagaggaaga gcaaaacaaa agtaagaaaa aagcacagca agcagcagct 360
 gacacaggac acagcaatca ggtcagccaa aattacctcg acaggactgg tccgatctct 420
 ccgatagaaa cagtttcggg caagcttaaa ccagggatgg atgggccaaa ggtcaagcag 480
 tggccgctaa cggaagagaa gattaaggcg ctcgtagaga ttgtactga aatggagaag 540
 gaaggcaaga taagcaagat cgggccagag aaccctgaca atacaccggg atttgcaata 600
 aagaagaagg attcaacaaa atggcgaaaag cttgtagatt ttagggaact aaacaagcga 660
 acccaagact ttgggaagt ccaactaggg atccccatc cagccggctc aaagaagaag 720
 aaatcggctca cagtcctgga tgtaggagac gcatatttta gtgtaccgct tgatgaggac 780
 ttccgaaagt atactgctt tactataccg agcataaaca atgaaacgcc aggcattcgc 840
 tatcagtaca acgtgctccc gcagggctgg aaggggtctc cggcgatatt tcagagctct 900
 atgacaaaaa tacttgaacc attccgaaaag cagaatccgg atattgtaat ttaccaatac 960
 atggacgatc tctatgtggg ctcggatcta gaaattgggc agcatcgcac taagattgag 1020
 gaactgaggc aacatctgct tcgatggggc ctcactactc ccgacaagaa gcaccagaag 1080

gagccgccgt tcctaaagat gggctacgag cttcatccgg acaagtggac agtacagccg 1140
 atagtgtctc ccgaaaagga ttcttgacc gtaaatgata ttcagaaact agtcggcaag 1200
 cttaactggg cctctcagat ttaccaggc attaaggctc gacagctttg caagctactg 1260
 aggggaacta aggctctaac agaggctatc ccattaacgg aggaagcaga gcttgagctg 1320
 gcagagaatc gcgaaattct taaggagccg gtgcacaggg tatactacga cccctccaag 1380
 gaccttatag ccgagatcca gaagcagggg cagggccaat ggacgtacca gatatatcaa 1440
 gaaccgttta agaacttgaa gactgggaag tacgcgcgca aacgaggggc tcataactaat 1500
 gatgtaaagc aacttacgga agcagtacaa aagattacta ctgagtctat tgtgatatgg 1560
 ggcaagacc caaagttcaa gctgcccata cagaaggaaa catgggaaac atgggtggact 1620
 gaataattggc aagctacctg gattccagaa tgggaatttg tcaacacgcc gccgctggta 1680
 aaactgaggc ctatgggtgg caagtggcga aaaagttagt tgggtggatg gcctactgta 1740
 agggaaaagaa tgagacgagc tgagccagca gcagatgggg tgggagcagc atctcgagac 1800
 ctggaaaaac atggagcaat cacaagtagc aatacagcag ctaccaatgc tgcttgtgcc 1860
 tggctagaag cacaagagga ggaggaggtg ggttttccag tcacacctca ggtaccttta 1920
 agaccaatga cttacaaggc agctgtagat cttagccact ttttaaaaga aaagggggga 1980
 ctggaagggc taattcactc ccaacgaaga caagatatcc ttgatctgtg gatctaccac 2040
 acacaaggct acttccctga ttggcagaac tacacaccag ggccaggggt cagatatcca 2100
 ctgacctttg gatgggtgcta caagctagta ccagttgagc cgataaaggc agaagaggcc 2160
 aataaaggag agaacaccag cttgttacac cctgtgagcc tgcattggaat ggatgaccct 2220
 gagagagaag tgtttagagt gaggtttgac agccgcctag catttcatca cgtggcccga 2280
 gagctgcatc cggagtactt caagaactgc taa 2313

<210> 17

<211> 770

<212> PRT

5 <213> Secuencia artificial

<220>

<223> p17-p51-Nef (F3)

10 <400> 17

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp
 1 Lys Ile Arg 5 Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys 15 Leu Lys
 His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro
 35 20 40 45 50 55 60 65
 Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu
 Gln Pro Ser Leu Gln Thr 70 Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn
 80 Thr Val Ala Thr 85 Leu Tyr Cys Val His Gln Arg Ile Glu Ile Lys Asp
 90 95 Thr Lys Glu Ala 100 Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys
 105 110 Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly His Ser Asn Gln Val
 115 120 125 Ser Gln Asn Tyr Leu Asp Arg Thr Gly Pro Ile Ser Pro Ile Glu Thr
 130 135 140 145 Val Ser Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys Gln
 150 155 160 Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile Cys Thr
 165 170 175 Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro
 180 185 190 Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp
 195 200 205 Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe
 210 215 220 Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys
 225 230 235 Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Val Pro
 240 245 250 Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile
 255 260 265 Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln
 270 275 280 Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys Ile
 285 290 295 Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr Gln Tyr
 300 305 310 Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg
 315 320 325 Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly Leu Thr
 330 335 340 Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Lys Met Gly
 345 350

		355					360					365				
Tyr	Glu	Leu	His	Pro	Asp	Lys	Trp	Thr	Val	Gln	Pro	Ile	Val	Leu	Pro	
	370					375					380					
Glu	Lys	Asp	Ser	Trp	Thr	Val	Asn	Asp	Ile	Gln	Lys	Leu	Val	Gly	Lys	
385					390					395					400	
Leu	Asn	Trp	Ala	Ser	Gln	Ile	Tyr	Pro	Gly	Ile	Lys	Val	Arg	Gln	Leu	
				405					410					415		
Cys	Lys	Leu	Leu	Arg	Gly	Thr	Lys	Ala	Leu	Thr	Glu	Val	Ile	Pro	Leu	
			420					425					430			
Thr	Glu	Glu	Ala	Glu	Leu	Glu	Leu	Ala	Glu	Asn	Arg	Glu	Ile	Leu	Lys	
		435					440					445				
Glu	Pro	Val	His	Gly	Val	Tyr	Tyr	Asp	Pro	Ser	Lys	Asp	Leu	Ile	Ala	
	450					455					460					
Glu	Ile	Gln	Lys	Gln	Gly	Gln	Gly	Gln	Trp	Thr	Tyr	Gln	Ile	Tyr	Gln	
465					470					475					480	
Glu	Pro	Phe	Lys	Asn	Leu	Lys	Thr	Gly	Lys	Tyr	Ala	Arg	Met	Arg	Gly	
				485					490					495		
Ala	His	Thr	Asn	Asp	Val	Lys	Gln	Leu	Thr	Glu	Ala	Val	Gln	Lys	Ile	
			500					505					510			
Thr	Thr	Glu	Ser	Ile	Val	Ile	Trp	Gly	Lys	Thr	Pro	Lys	Phe	Lys	Leu	
		515					520					525				
Pro	Ile	Gln	Lys	Glu	Thr	Trp	Glu	Thr	Trp	Trp	Thr	Glu	Tyr	Trp	Gln	
	530					535					540					
Ala	Thr	Trp	Ile	Pro	Glu	Trp	Glu	Phe	Val	Asn	Thr	Pro	Pro	Leu	Val	
545					550					555					560	
Lys	Leu	Arg	Pro	Met	Gly	Gly	Lys	Trp	Ser	Lys	Ser	Ser	Val	Val	Gly	
				565					570					575		
Trp	Pro	Thr	Val	Arg	Glu	Arg	Met	Arg	Arg	Ala	Glu	Pro	Ala	Ala	Asp	
			580					585					590			
Gly	Val	Gly	Ala	Ala	Ser	Arg	Asp	Leu	Glu	Lys	His	Gly	Ala	Ile	Thr	
		595					600					605				
Ser	Ser	Asn	Thr	Ala	Ala	Thr	Asn	Ala	Ala	Cys	Ala	Trp	Leu	Glu	Ala	
	610					615					620					
Gln	Glu	Glu	Glu	Glu	Val	Gly	Phe	Pro	Val	Thr	Pro	Gln	Val	Pro	Leu	
625					630					635					640	
Arg	Pro	Met	Thr	Tyr	Lys	Ala	Ala	Val	Asp	Leu	Ser	His	Phe	Leu	Lys	
				645					650					655		
Glu	Lys	Gly	Gly	Leu	Glu	Gly	Leu	Ile	His	Ser	Gln	Arg	Arg	Gln	Asp	
			660					665					670			
Ile	Leu	Asp	Leu	Trp	Ile	Tyr	His	Thr	Gln	Gly	Tyr	Phe	Pro	Asp	Trp	
		675					680					685				
Gln	Asn	Tyr	Thr	Pro	Gly	Pro	Gly	Val	Arg	Tyr	Pro	Leu	Thr	Phe	Gly	
	690					695					700					
Trp	Cys	Tyr	Lys	Leu	Val	Pro	Val	Glu	Pro	Asp	Lys	Val	Glu	Glu	Ala	
705					710					715					720	
Asn	Lys	Gly	Glu	Asn	Thr	Ser	Leu	Leu	His	Pro	Val	Ser	Leu	His	Gly	
				725					730					735		
Met	Asp	Asp	Pro	Glu	Arg	Glu	Val	Leu	Glu	Trp	Arg	Phe	Asp	Ser	Arg	
			740					745					750			
Leu	Ala	Phe	His	His	Val	Ala	Arg	Glu	Leu	His	Pro	Glu	Tyr	Phe	Lys	
		755					760					765				
Asn	Cys															
	770															

<210> 18

<211> 3000

5 <212> ADN

<213> Secuencia artificial

<220>

<223> F4(p51)*

<400> 18

atggttatcg	tgcagaacat	ccagggggcaa	atggtacatc	aggccatatc	acctagaact	60
ttaaattgcat	gggtaaaagt	agtagaagag	aaggctttca	gcccagaagt	aatacccatg	120
ttttcagcat	tatcagaagg	agccacccca	caagatttaa	acaccatgct	aaacacagtg	180
gggggacatc	aagcagccat	gcaaatgtta	aaagagacca	tcaatgagga	agctgcagaa	240
tgggatagag	tacatccagt	gcatgcaggg	cctattgcac	caggccagat	gagagaacca	300
aggggaagtg	acatagcagg	aactactagt	acccttcagg	aacaaatagg	atggatgaca	360
aaaaatccac	ctatcccagt	aggagaaaatt	tataaaagat	ggataatcct	gggattaat	420
aaaatagtaa	gaatgtatag	ccctaccagc	attctggaca	taagacaagg	accaaaagaa	480
ccttttagag	actatgtaga	ccggttctat	aaaactctaa	gagccgagca	agcttcacag	540
gaggtaaaaa	attggatgac	agaaaccttg	ttggtccaaa	atgcaacc	agattgtaag	600
actattttaa	aagcattggg	accagcggct	acactagaag	aaatgatgac	agcatgtcag	660
ggagtaggag	gacccggcca	taaggcaaga	gttttgcata	tgaggcctgg	tccgatctct	720

ccgatagaaa	cagtttcggt	caagcttaaa	ccagggatgg	atggtccaaa	ggtcaagcag	780
tggccgctaa	cggaagagaa	gattaaggcg	ctcgtagaga	tttgtactga	aatggagaag	840
gaaggcaaga	taagcaagat	cgggccagag	aacccgtaca	atacaccggt	atttgcaata	900
aagaagaagg	attcaacaaa	atggcgaag	cttgtagatt	ttaggggaact	aaacaagcga	960
accaagact	tttgggaagt	ccaactaggt	atcccacatc	cagccggtct	aaagaagaag	1020
aaatcggta	cagtcctgga	tgtaggagac	gcatatttta	gtgtaccgct	tgtaggagc	1080
ttccgaaagt	atacttcggt	tactataccg	agcataaaaca	atgaaacgcc	aggcattcgc	1140
tatcagtaca	acgtgctccc	gcagggctgg	aaggggtctc	cggcgatatt	tcagagctct	1200
atgacaaaaa	tacttgaacc	attccgaaag	cagaatccgg	atattgtaat	ttaccaatac	1260
atggacgatc	tctatgtggg	ctcggatcta	gaaattgggc	agcatcgac	taagattgag	1320
gaactgaggc	aacatctgct	tcgatggggc	ctcactactc	ccgacaagaa	gcaccagaag	1380
gagccgccgt	tcctaaagat	gggctacgag	cttcatccgg	acaagtggac	agtacagcc	1440
atagtctgc	ccgaaaagga	ttcttggacc	gtaaatgata	ttcagaaact	agtcggcaag	1500
cctaactggg	cctctcagat	ttaccaggc	attaaggctc	gacagcttg	caagctactg	1560
aggggaacta	aggctctaac	agaggtcatc	ccattaacgg	aggaagcaga	gcttgagctg	1620
gcagagaatc	gcgaaattct	taaggagccg	gtgcacaggg	tatactacga	cccctccaag	1680
gaccttatag	ccgagatcca	gaagcagggg	cagggccaat	ggacgtacca	gatataatca	1740
gaaccgttta	agaatctgaa	gactgggaag	tacgcgcgca	aacgaggggc	tcatactaat	1800
gatgtaaac	aacttacgga	agcagtacaa	aagattacta	ctgagtctat	tgtgatatgg	1860
ggcaagacc	caaagttcaa	gctgcccata	cagaaggaaa	catgggaaac	atggtggact	1920
gaatattggc	aagctacctg	gattccagaa	tgggaatttg	tcaacacgcc	gccgctggtg	1980
aaactggccc	tagctatggg	tggcaagtgg	tcaaaaagta	gtgtggttgg	atggcctact	2040
gtaagggaaa	gaatgagacg	agctgagcca	gcagcagatg	gggtgggagc	agcatctcga	2100
gacctggaaa	aacatggagc	aatcacaagt	agcaatacag	cagctaccaa	tgctgcttgt	2160
gcctggctag	aagcacaaga	ggaggaggag	gtgggttttc	cagtcacacc	tcaggtacct	2220
ttaagaccaa	tgacttacaa	ggcagctgta	gatcttagcc	actttttaa	agaaaagggg	2280
ggactggaag	ggctaattca	ctcccacga	agacaagata	tccttgatct	gtggatctac	2340
cacacacaag	gctacttccc	tgattggcag	aactacacac	cagggccagg	ggtcagatat	2400
ccactgacct	ttggatggg	ctacaagcta	gtaccagtgg	agccagataa	ggtagaagag	2460
gccaataaag	gagagaacac	cagcttggtg	caccctgtga	gcctgcatgg	aatggatgac	2520
cctgagagag	aagtgttaga	gtggaggttt	gacagccgcc	tagcatttca	tcacgtggcc	2580
cgagagctgc	atccggagta	cttcaagaac	tgcaggccta	tgggtgagag	agcgtcagta	2640
ttaagcgggg	gagaattaga	tcgatgggaa	aaaattcggg	taaggccagg	gggaaagaaa	2700
aaatataaat	taaaacatat	agtatgggca	agcagggagc	tagaacgatt	cgcagttaat	2760
cctggcctgt	tagaaacatc	agaaggctgt	agacaataac	tgggacagct	acaaccatcc	2820
cttcagacag	gatcagaaga	acttagatca	ttatataata	cagtagcaac	cctctattgt	2880
gtgcatcaaa	ggatagagat	aaaagacacc	aaggaagctt	tagacaagat	agaggaagag	2940
caaaacaaaa	gtaagaaaaa	agcacagcaa	gcagcagctg	acacaggaca	cagcaatcag	3000

5

<210> 19

<211> 1005

<212> PRT

<213> Secuencia artificial

<220>

<223> F4(p51)*

5

<400> 19

Met	Val	Ile	Val	Gln	Asn	Ile	Gln	Gly	Gln	Met	Val	His	Gln	Ala	Ile
1				5				10					15		
Ser	Pro	Arg	Thr	Leu	Asn	Ala	Trp	Val	Lys	Val	Val	Glu	Glu	Lys	Ala
			20					25				30			
Phe	Ser	Pro	Glu	Val	Ile	Pro	Met	Phe	Ser	Ala	Leu	Ser	Glu	Gly	Ala
		35					40					45			
Thr	Pro	Gln	Asp	Leu	Asn	Thr	Met	Leu	Asn	Thr	Val	Gly	Gly	His	Gln
	50					55					60				
Ala	Ala	Met	Gln	Met	Leu	Lys	Glu	Thr	Ile	Asn	Glu	Glu	Ala	Ala	Glu
65					70					75					80
Trp	Asp	Arg	Val	His	Pro	Val	His	Ala	Gly	Pro	Ile	Ala	Pro	Gly	Gln
				85					90					95	
Met	Arg	Glu	Pro	Arg	Gly	Ser	Asp	Ile	Ala	Gly	Thr	Thr	Ser	Thr	Leu
			100					105					110		
Gln	Glu	Gln	Ile	Gly	Trp	Met	Thr	Asn	Asn	Pro	Pro	Ile	Pro	Val	Gly
		115					120					125			
Glu	Ile	Tyr	Lys	Arg	Trp	Ile	Ile	Leu	Gly	Leu	Asn	Lys	Ile	Val	Arg
	130					135					140				
Met	Tyr	Ser	Pro	Thr	Ser	Ile	Leu	Asp	Ile	Arg	Gln	Gly	Pro	Lys	Glu
145					150					155					160
Pro	Phe	Arg	Asp	Tyr	Val	Asp	Arg	Phe	Tyr	Lys	Thr	Leu	Arg	Ala	Glu
				165					170					175	
Gln	Ala	Ser	Gln	Glu	Val	Lys	Asn	Trp	Met	Thr	Glu	Thr	Leu	Leu	Val
			180					185					190		
Gln	Asn	Ala	Asn	Pro	Asp	Cys	Lys	Thr	Ile	Leu	Lys	Ala	Leu	Gly	Pro
		195					200					205			
Ala	Ala	Thr	Leu	Glu	Glu	Met	Met	Thr	Ala	Cys	Gln	Gly	Val	Gly	Gly

210	Pro	Gly	His	Lys	Ala	Arg	Val	Leu	His	Met	Arg	Pro	Gly	Pro	Ile	Ser	
225	Pro	Ile	Glu	Thr	Val	Ser	Val	Lys	Leu	Lys	Pro	Gly	Met	Asp	Gly	Pro	
	Lys	Val	Lys	Gln	Trp	Pro	Leu	Thr	Glu	Glu	Lys	Ile	Lys	Ala	Leu	Val	
	Glu	Ile	Cys	Thr	Glu	Met	Glu	Lys	265	Glu	Gly	Lys	Ile	Ser	270	Ile	Gly
	Pro	Glu	Asn	Pro	Tyr	Asn	Thr	Pro	Val	Phe	Ala	Ile	Lys	Lys	Lys	Asp	
	Ser	Thr	Lys	Trp	Arg	Lys	Leu	Val	Asp	Phe	Arg	Glu	Leu	Asn	Lys	Arg	
305	Thr	Gln	Asp	Phe	Trp	Glu	Val	Gln	Leu	Gly	Ile	Pro	His	Pro	Ala	Gly	
	Leu	Lys	Lys	Lys	Lys	Ser	Val	Thr	Val	Leu	Asp	Val	Gly	Asp	Ala	Tyr	
	Phe	Ser	Val	Pro	Leu	Asp	Glu	Asp	Phe	Arg	Lys	Tyr	Thr	Ala	Phe	Thr	
	Ile	Pro	Ser	Ile	Asn	Asn	Glu	Thr	Pro	Gly	Ile	Arg	Tyr	Gln	Tyr	Asn	
	Val	Leu	Pro	Gln	Gly	Trp	Lys	Gly	Ser	Pro	Ala	Ile	Phe	Gln	Ser	Ser	
385	Met	Thr	Lys	Ile	Leu	Glu	Pro	Phe	Arg	Lys	Gln	Asn	Pro	Asp	Ile	Val	
	Ile	Tyr	Gln	Tyr	Met	Asp	Asp	Leu	Tyr	Val	Gly	Ser	Asp	Leu	Glu	Ile	
	Gly	Gln	His	Arg	Thr	Lys	Ile	Glu	Glu	Leu	Arg	Gln	His	Leu	Leu	Arg	
	Trp	Gly	Leu	Thr	Thr	Pro	Asp	Lys	Lys	His	Gln	Lys	Glu	Pro	Pro	Phe	
	Leu	Lys	Met	Gly	Tyr	Glu	Leu	His	Pro	Asp	Lys	Trp	Thr	Val	Gln	Pro	
465	Ile	Val	Leu	Pro	Glu	Lys	Asp	Ser	Trp	Thr	Val	Asn	Asp	Ile	Gln	Lys	
	Leu	Val	Gly	Lys	Leu	Asn	Trp	Ala	Ser	Gln	Ile	Tyr	Pro	Gly	Ile	Lys	
	Val	Arg	Gln	Leu	Cys	Lys	Leu	Leu	Arg	Gly	Thr	Lys	Ala	Leu	Thr	Glu	
	Val	Ile	Pro	Leu	Thr	Glu	Glu	Ala	Glu	Leu	Glu	Leu	Ala	Glu	Asn	Arg	
	Glu	Ile	Leu	Lys	Glu	Pro	Val	His	Gly	Val	Tyr	Tyr	Asp	Pro	Ser	Lys	
545	Asp	Leu	Ile	Ala	Glu	Ile	Gln	Lys	Gln	Gly	Gln	Gly	Gln	Trp	Thr	Tyr	
	Gln	Ile	Tyr	Gln	Glu	Pro	Phe	Lys	Asn	Leu	Lys	Thr	Gly	Lys	Tyr	Ala	
	Arg	Lys	Arg	Gly	Ala	His	Thr	Asn	Asp	Val	Lys	Gln	Leu	Thr	Glu	Ala	
	Val	Gln	Lys	Ile	Thr	Thr	Glu	Ser	Ile	Val	Ile	Trp	Gly	Lys	Thr	Pro	
	Lys	Phe	Lys	Leu	Pro	Ile	Gln	Lys	Glu	Thr	Trp	Glu	Thr	Trp	Trp	Thr	
625	Glu	Tyr	Trp	Gln	Ala	Thr	Trp	Ile	Pro	Glu	Trp	Glu	Phe	Val	Asn	Thr	
	Pro	Pro	Leu	Val	Lys	Leu	Ala	Leu	Ala	Met	Gly	Gly	Lys	Trp	Ser	Lys	
	Ser	Ser	Val	Val	Gly	Trp	Pro	Thr	Val	Arg	Glu	Arg	Met	Arg	Arg	Ala	
	Glu	Pro	Ala	Ala	Asp	Gly	Val	Gly	Ala	Ala	Ser	Arg	Asp	Leu	Glu	Lys	
	His	Gly	Ala	Ile	Thr	Ser	Ser	Asn	Thr	Ala	Ala	Thr	Asn	Ala	Ala	Cys	
705	Ala	Trp	Leu	Glu	Ala	Gln	Glu	Glu	Glu	Glu	Val	Gly	Phe	Pro	Val	Thr	
	Pro	Gln	Val	Pro	Leu	Arg	Pro	Met	Thr	Tyr	Lys	Ala	Ala	Val	Asp	Leu	
	Ser	His	Phe	Leu	Lys	Glu	Lys	Gly	Gly	Leu	Glu	Gly	Leu	Ile	His	Ser	
	Gln	Arg	Arg	Gln	Asp	Ile	Leu	Asp	Leu	Trp	Ile	Tyr	His	Thr	Gln	Gly	
	Tyr	Phe	Pro	Asp	Trp	Gln	Asn	Tyr	Thr	Pro	Gly	Pro	Gly	Val	Arg	Tyr	
785	Pro	Leu	Thr	Phe	Gly	Trp	Cys	Tyr	Lys	Leu	Val	Pro	Val	Glu	Pro	Asp	
	Lys	Val	Glu	Glu	Ala	Asn	Lys	Gly	Glu	Asn	Thr	Ser	Leu	Leu	His	Pro	

Val	Ser	Leu	820 His	Gly	Met	Asp	Asp	825 Pro	Glu	Arg	Glu	Val	830 Leu	Glu	Trp
Arg	Phe	Asp	835 Ser	Arg	Leu	Ala	Phe	840 His	His	Val	Ala	845 Arg	Glu	Leu	His
Pro	Glu	Tyr	Phe	Lys	Asn	Cys	Arg	Pro	Met	Gly	Ala	860 Arg	Ala	Ser	Val
865 Leu	Ser	Gly	Gly	Glu	870 Leu	Asp	Arg	Trp	Glu	Lys	Ile	Arg	Leu	Arg	880 Pro
Gly	Gly	Lys	Lys	Lys	Tyr	Lys	Leu	Lys	890 His	Ile	Val	Trp	Ala	Ser	Arg
Glu	Leu	Glu	Arg	Phe	Ala	Val	Asn	905 Pro	Gly	Leu	Leu	Glu	910 Thr	Ser	Glu
Gly	Cys	Arg	Gln	Ile	Leu	Gly	Gln	920 Leu	Gln	Pro	Ser	Leu	Gln	Thr	Gly
930 Ser	Glu	Glu	Leu	Arg	Ser	Leu	Tyr	Asn	Thr	Val	Ala	Thr	Leu	Tyr	Cys
945 Val	His	Gln	Arg	Ile	950 Glu	Ile	Lys	Asp	Thr	Lys	Glu	Ala	Leu	Asp	Lys
Ile	Glu	Glu	Glu	Gln	Asn	Lys	Ser	Lys	970 Lys	Lys	Ala	Gln	Gln	Ala	Ala
Ala	Asp	Thr	980 Gly	His	Ser	Asn	Gln	985 Val	Ser	Gln	Asn	Tyr			
			995				1000					1005			

REIVINDICACIONES

1. Una composición que comprende:
 - a) un polipéptido que comprende p17 Gag o un fragmento o derivado inmunogénico de la misma, p51 RT o un fragmento o derivado inmunogénico de la misma, Nef o un fragmento o derivado inmunogénico de la misma y p24 Gag o un fragmento o derivado inmunogénico de la misma, en los que hay al menos un antígeno de VIH o fragmento inmunogénico entre p17 Gag y p24 Gag; y
 - b) una proteína env de VIH o fragmento inmunogénico o derivado de la misma,en la que los fragmentos o derivados inmunogénicos son capaces de aumentar una respuesta inmune frente al antígeno nativo.
- 5 2. La composición de acuerdo con la reivindicación 1, en la que la proteína env de VIH es gp120. .
3. La composición de acuerdo con una cualquiera de las reivindicaciones 1 o 2, en la que el p51 RT comprende una mutación en la posición 592 para reemplazar la metionina por otro residuo, por ejemplo, lisina.
4. La composición de acuerdo con una cualquiera de las reivindicaciones 1 a 3, en la que la Nef es Nef de longitud completa.
- 15 5. Una composición que comprende un polinucleótido o polinucleótidos que codifican:
 - a) un polipéptido que comprende p17 Gag o un fragmento o derivado inmunogénico de la misma, p51 RT o un fragmento o derivado inmunogénico de la misma, Nef o un fragmento o derivado inmunogénico de la misma y p24 Gag o un fragmento o derivado inmunogénico de la misma, en los que hay al menos un antígeno de VIH o fragmento inmunogénico entre p17 Gag y p24 Gag; y
 - b) una proteína env de VIH o fragmento inmunogénico o derivado de la misma.en la que los fragmentos o derivados inmunogénicos son capaces de aumentar una respuesta inmune frente al antígeno nativo.
- 20 6. Una composición farmacéutica que comprende una composición de acuerdo con cualquier reivindicación previa, junto con un vehículo o adyuvante farmacéuticamente aceptable.
- 25 7. Una composición farmacéutica de acuerdo con la reivindicación 6, en la que el adyuvante es un adyuvante que induce Th1 tal como QS21 o 3D-MPL o una combinación de QS21 y 3D-MPL.

Figura 1A

SDS-PAGE al 10% en condiciones reductoras (F4 p24-RT-Nef-p17)

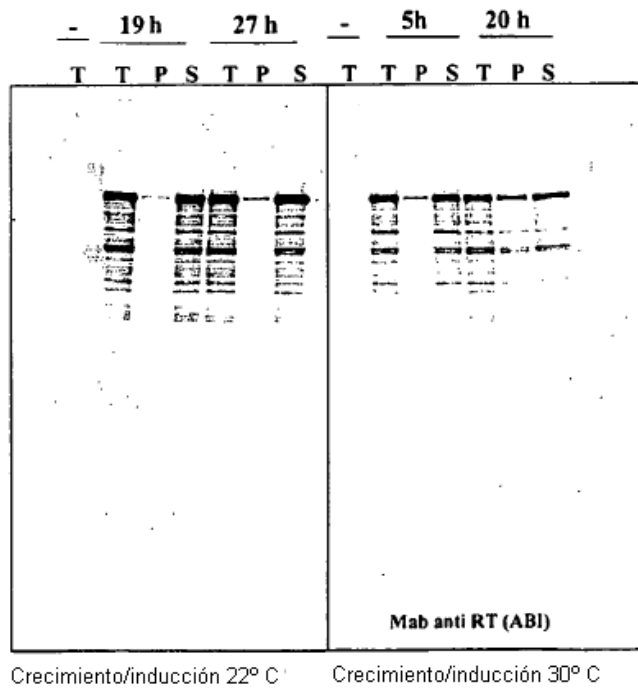
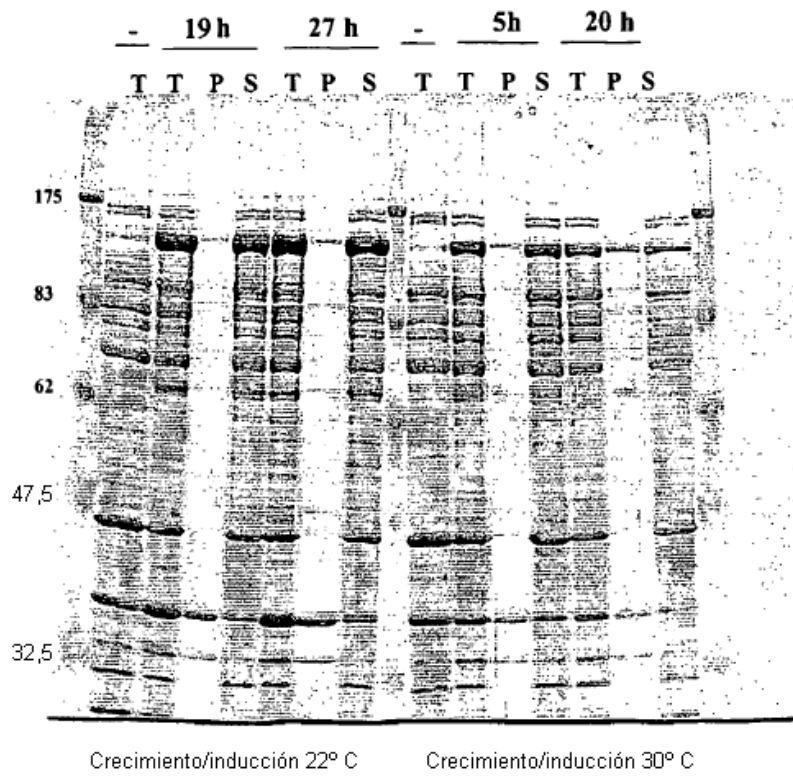
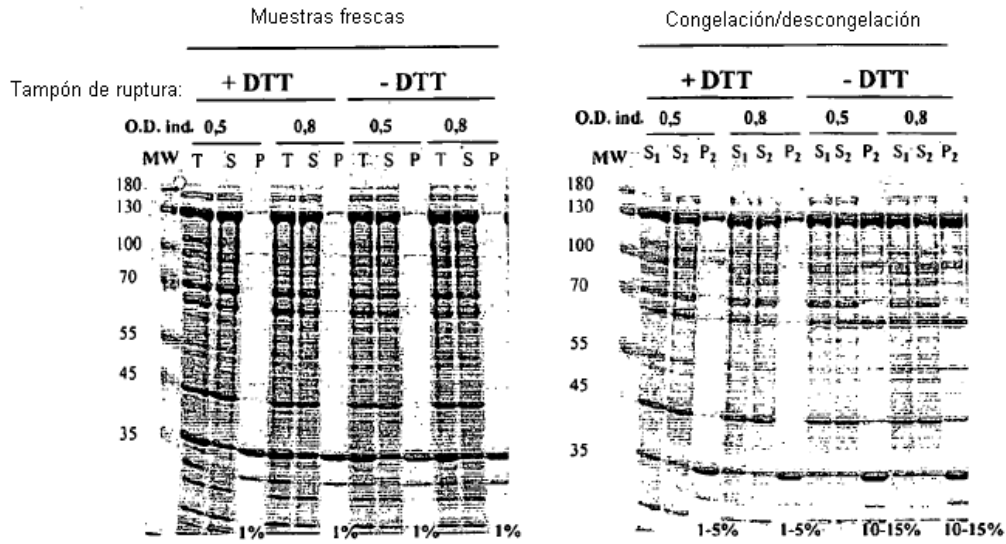


Figura 1B

Ensayo de solubilidad (F4 p24-RT-Nef-p17)

Geles reductores (SDS-PAGE al 10%)

Coomassie



Transferencia Western:

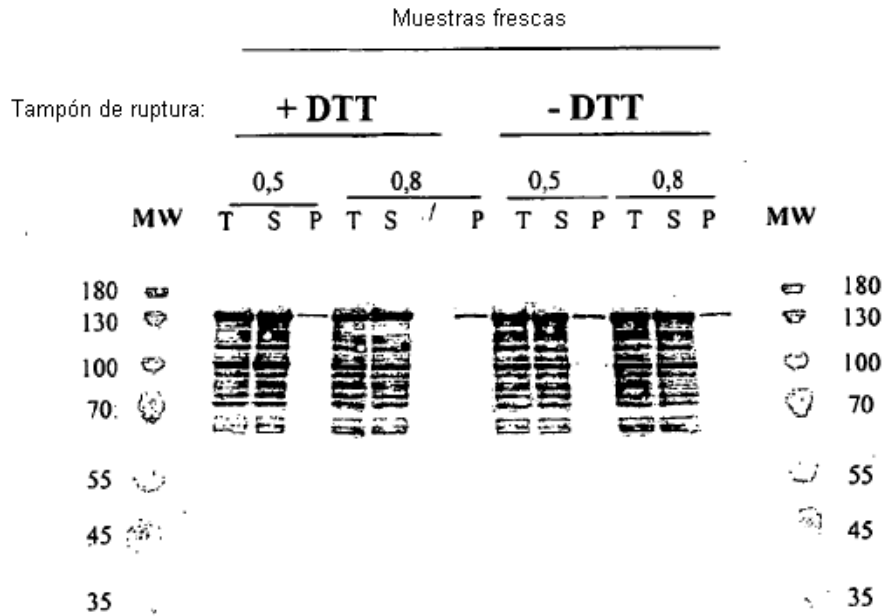
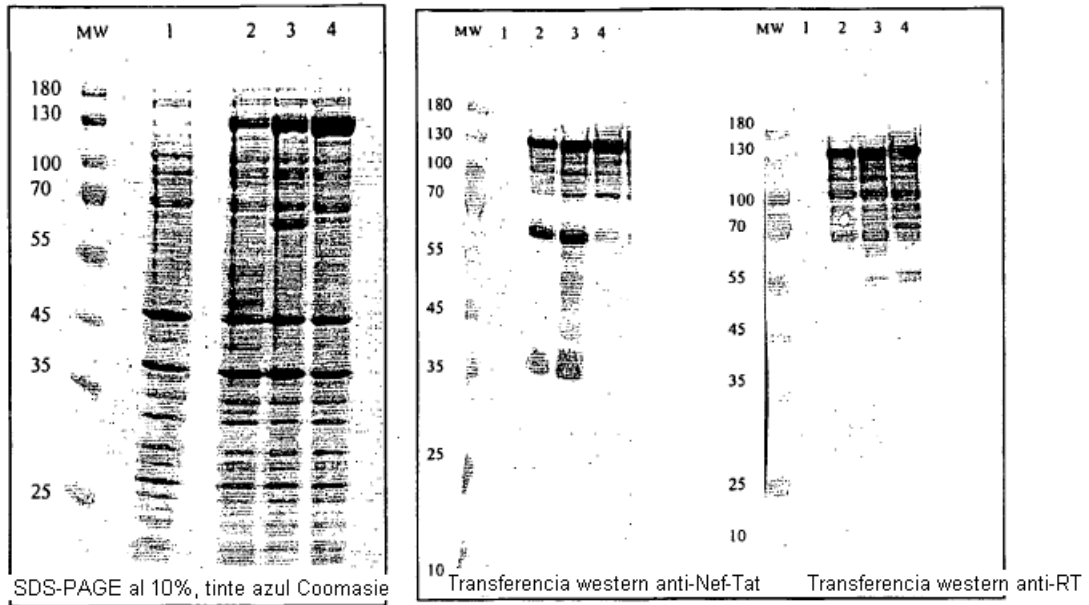


Figura 2: gel tintado con Coomassie y transferencia western para F4 optimizado con codón



1/ no inducido
 2/ B834(DE3) / F4 (gen nativo)
 3/ BLR(DE3) / F4 (gen nativo)
 4/ BLR(DE3) / F4 (gen optimizado con codón)

Figura 3

Asignación de proteínas RT: construcción GSK, HXB2, BH10

	* 20 * 40 *	
RT-GSK	: MGPISPIETVSVKLPKPGMDGPKVKQWPLTEEEKIKALVEICTEMEKEGKISKIGPE	: 55
RT-HXB2	: --PISPIETVSVKLPKPGMDGPKVKQWPLTEEEKIKALVEICTEMEKEGKISKIGPE	: 53
RT-BH10	: --PISPIETVSVKLPKPGMDGPKVKQWPLTEEEKIKALVEICTEMEKEGKISKIGPE	: 53
	60 * 80 * 100 *	
RT-GSK	: NPYNTPVFAIKKDKSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTV	: 110
RT-HXB2	: NPYNTPVFAIKKDKSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTV	: 108
RT-BH10	: NPYNTPVFAIKKDKSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTV	: 108
	120 * 140 * 160	
RT-GSK	: LDVGDAYFSVPLDEDFRKYTAFTIPSIINNETPGIRYQYNVLPQGKGS PAIFQSS	: 165
RT-HXB2	: LDVGDAYFSVPLDEDFRKYTAFTIPSIINNETPGIRYQYNVLPQGKGS PAIFQSS	: 163
RT-BH10	: LDVGDAYFSVPLDEDFRKYTAFTIPSIINNETPGIRYQYNVLPQGKGS PAIFQSS	: 163
	* 180 * 200 * 220	
RT-GSK	: MTKILEPFRKQNPDIYIQYMDLTVGSDLEIGQHRTKIEELRQHLLRWGLTTPD	: 220
RT-HXB2	: MTKILEPFRKQNPDIYIQYMDLTVGSDLEIGQHRTKIEELRQHLLRWGLTTPD	: 218
RT-BH10	: MTKILEPFRKQNPDIYIQYMDLTVGSDLEIGQHRTKIEELRQHLLRWGLTTPD	: 218
	* 240 * 260 *	
RT-GSK	: KKHQKEPPFLWNGYELHPDKWTVQPIVLPKDSWTVNDIQKLVGKLNWASQIYPG	: 275
RT-HXB2	: KKHQKEPPFLWNGYELHPDKWTVQPIVLPKDSWTVNDIQKLVGKLNWASQIYPG	: 273
RT-BH10	: KKHQKEPPFLWNGYELHPDKWTVQPIVLPKDSWTVNDIQKLVGKLNWASQIYPG	: 273
	280 * 300 * 320 *	
RT-GSK	: IKVRQLCKLLRGTALTEVIPLTEEALELAENREILKEPVHGVYDPSKDLIAE	: 330
RT-HXB2	: IKVRQLCKLLRGTALTEVIPLTEEALELAENREILKEPVHGVYDPSKDLIAE	: 328
RT-BH10	: IKVRQLCKLLRGTALTEVIPLTEEALELAENREILKEPVHGVYDPSKDLIAE	: 328
	340 * 360 * 380	
RT-GSK	: IQKQGQGWTYQIQEPPFKNLKTGKYARMGAHTNDVKQLTEAVQKITTESIVIW	: 385
RT-HXB2	: IQKQGQGWTYQIQEPPFKNLKTGKYARMGAHTNDVKQLTEAVQKITTESIVIW	: 383
RT-BH10	: IQKQGQGWTYQIQEPPFKNLKTGKYARMGAHTNDVKQLTEAVQKITTESIVIW	: 383
	* 400 * 420 * 440	
RT-GSK	: GKTPKFLPIQKETWETWTEYQATWIPWEPVNTPLVKKLWYQLEKEPIVGAE	: 440
RT-HXB2	: GKTPKFLPIQKETWETWTEYQATWIPWEPVNTPLVKKLWYQLEKEPIVGAE	: 438
RT-BH10	: GKTPKFLPIQKETWETWTEYQATWIPWEPVNTPLVKKLWYQLEKEPIVGAE	: 438
	* 460 * 480 *	
RT-GSK	: <u>TFYVDGAANRETKLKGAGYVTNRGRQKVVTLDTTNQKTELQAIYLAQDSGLEV</u>	: 495
RT-HXB2	: <u>TFYVDGAANRETKLKGAGYVTNRGRQKVVTLDTTNQKTELQAIYLAQDSGLEV</u>	: 493
RT-BH10	: <u>TFYVDGAANRETKLKGAGYVTNRGRQKVVTLDTTNQKTELQAIYLAQDSGLEV</u>	: 493
	500 * 520 * 540 *	
RT-GSK	: <u>NIVTDSQYALGIIQAQPDQSESELVNQITEQLIKKEKVYLAWVPAHKGIGGNEQV</u>	: 550
RT-HXB2	: <u>NIVTDSQYALGIIQAQPDQSESELVNQITEQLIKKEKVYLAWVPAHKGIGGNEQV</u>	: 548
RT-BH10	: <u>NIVTDSQYALGIIQAQPDQSESELVNQITEQLIKKEKVYLAWVPAHKGIGGNEQV</u>	: 548
	560	
RT-GSK	: <u>DKLVSAGIRKVL</u>	: 562
RT-HXB2	: <u>DKLVSAGI</u> ----	: 556
RT-BH10	: <u>DKLVSAGI</u> ----	: 556

K: mutación puntual introducida para eliminar actividad del enzima

5 **N**egrita y subrayado: aminoácidos que difieren de RT/HXB2

Bloque resaltado: dominio de RNasa H

Negrita: primeros aminoácidos de integrasa

Paréntesis: final de p51

Figura 4

Gel tintado con Coomasie y Transferencia western para P51 RT (optimizado con codón)

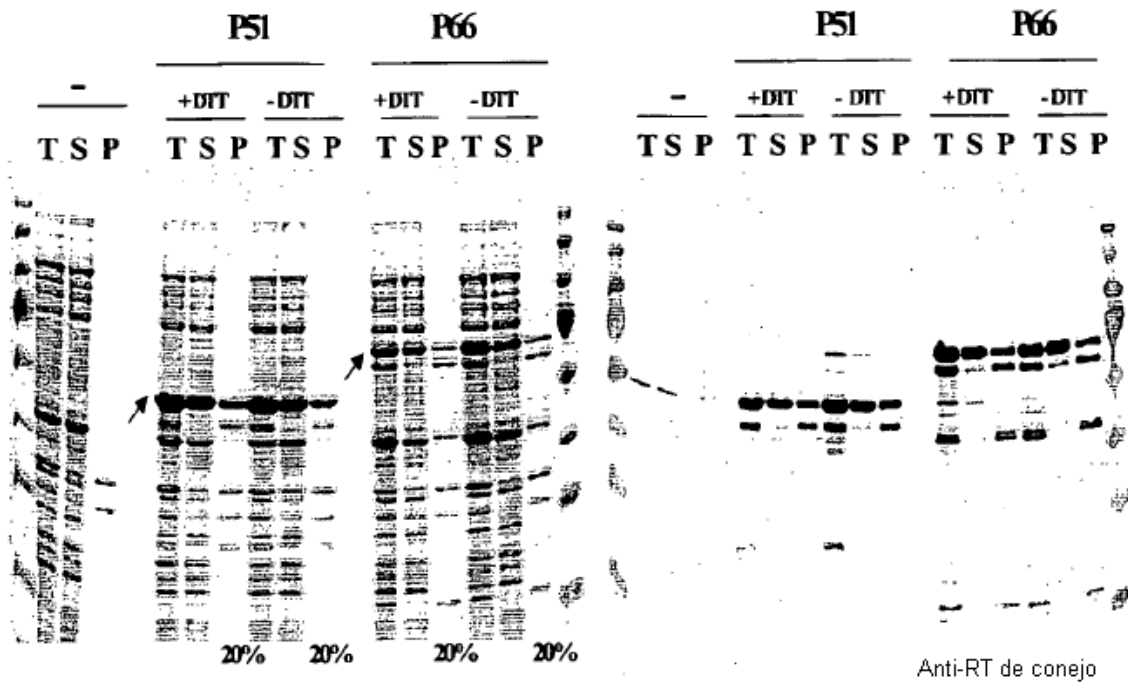
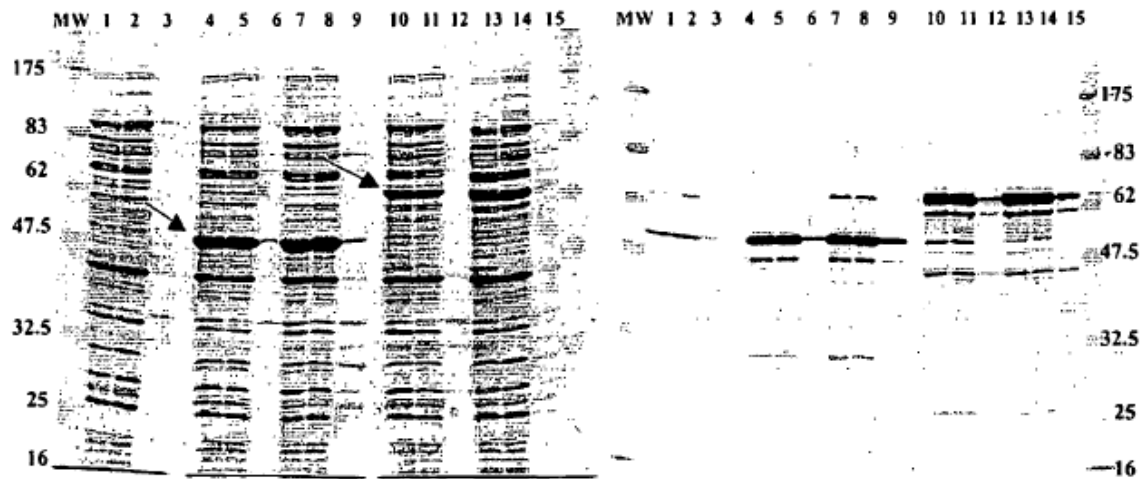


Figura 5:

Ensayo de solubilidad
RT/P51 y RT/p66

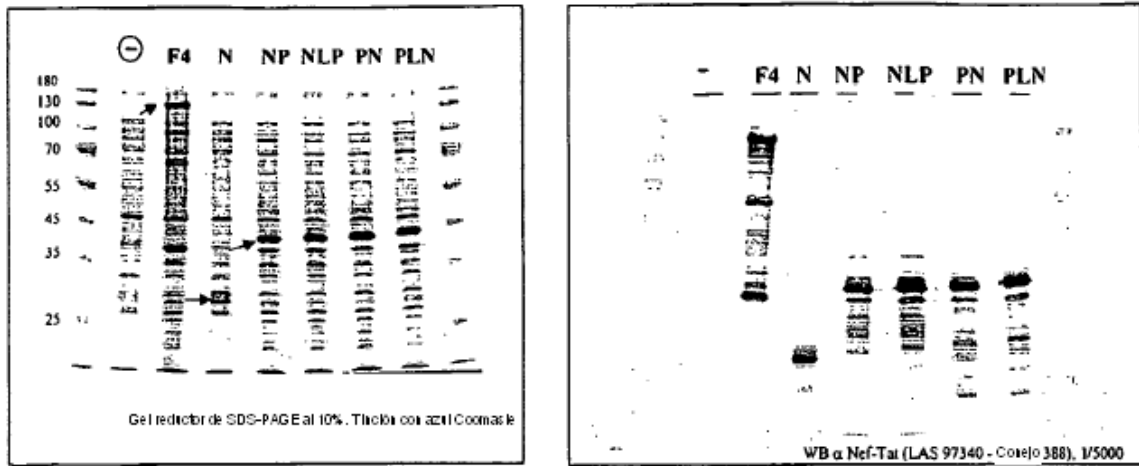
S1: fracción soluble (3 h de inducción a 30° C) conservada a -20° C,
descongelado y centrifugado a 20.000g / 30 min.

S2
P2 (resuspendido 1/10 vol.)



1. S1	+ DTT	Control negativo	10. S1	+ DTT	P66
2. S2			11. S2		
3. P2			12. P2		
4. S1	+ DTT	P51co	13. S1	- DTT	
5. S2			14. S2		
6. P2			15. P2		
7. S1					
8. S2	- DTT				
9. P2					

Figura 6



NP=Nef-p17

PN= p17-Nef

NLP=Nef-6S666F- p17

PLN= p17-6S666F- Nef

Figura 7

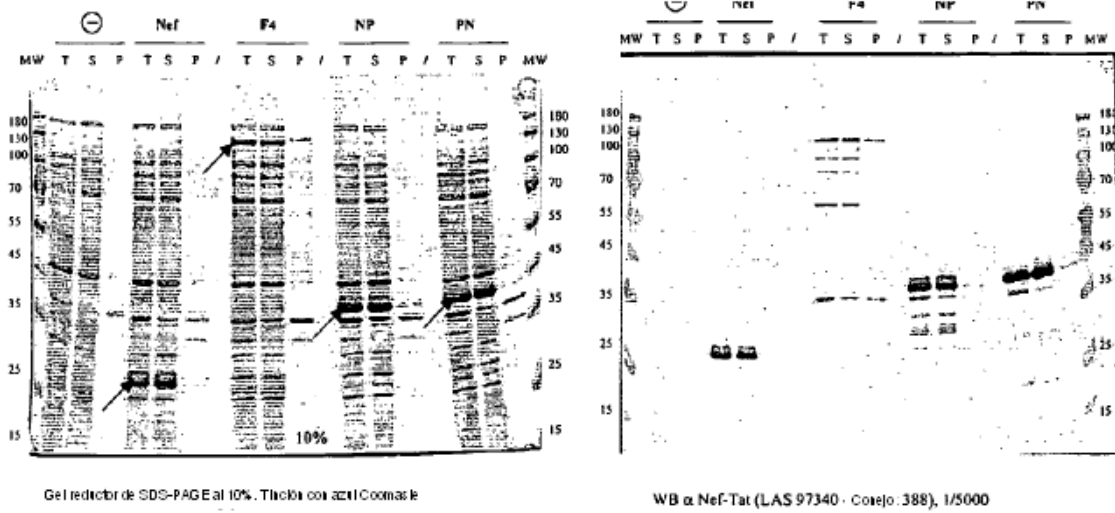
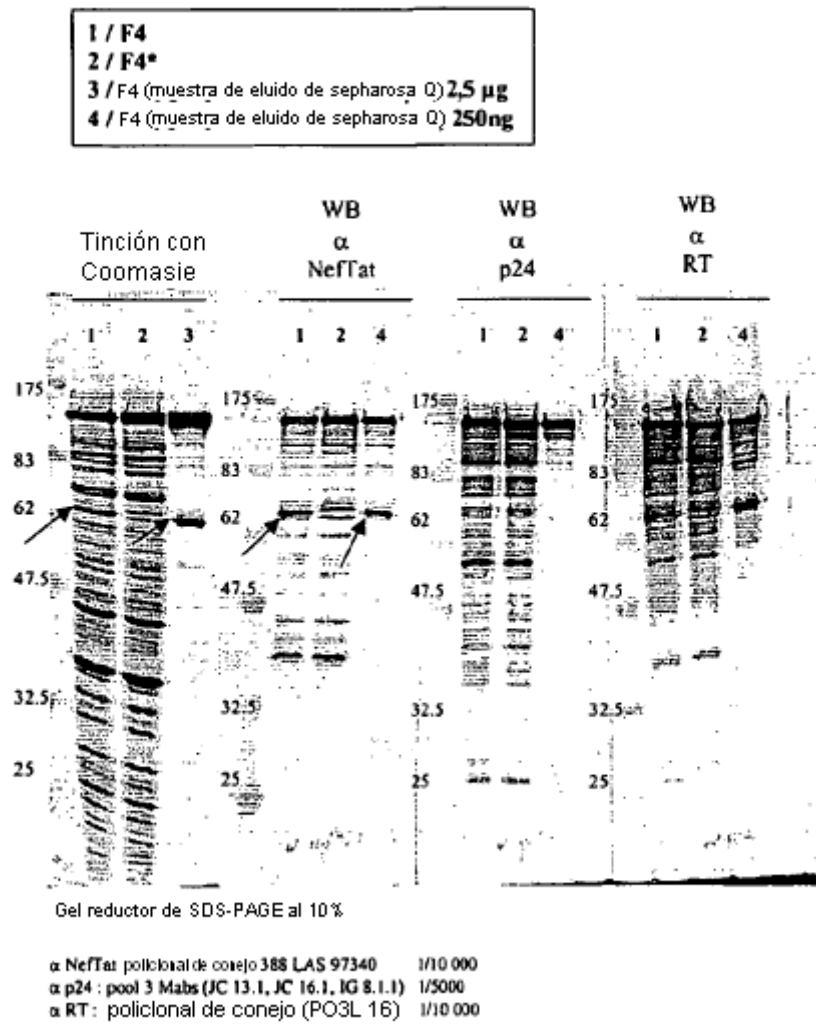


Figura 8

Inducción: 19 horas, 22° C



→ :Banda generada con el sitio "de inicio" interno de metionina

Figura 9

Condición de inducción; 3 h, 30° C

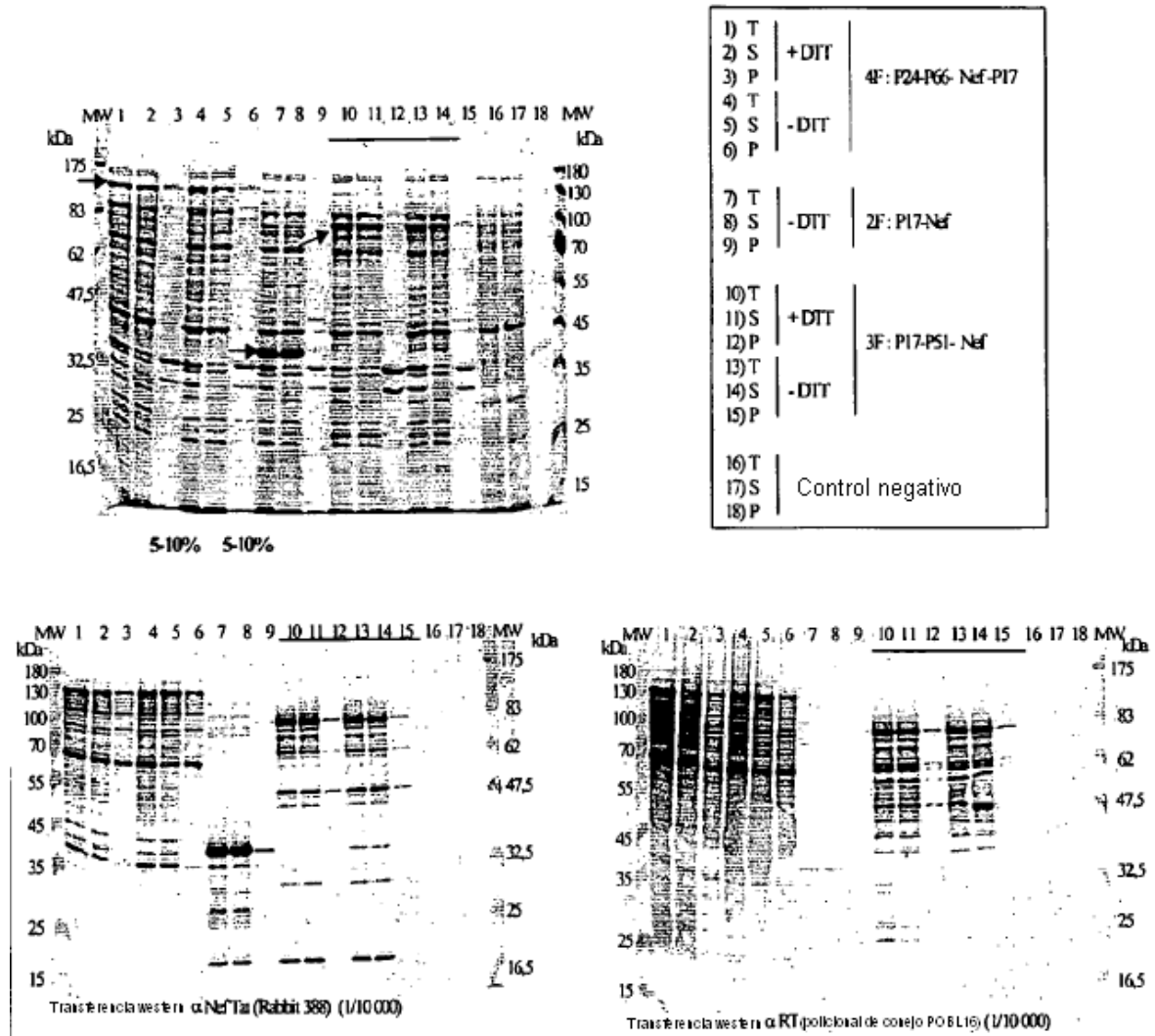


Figura 10

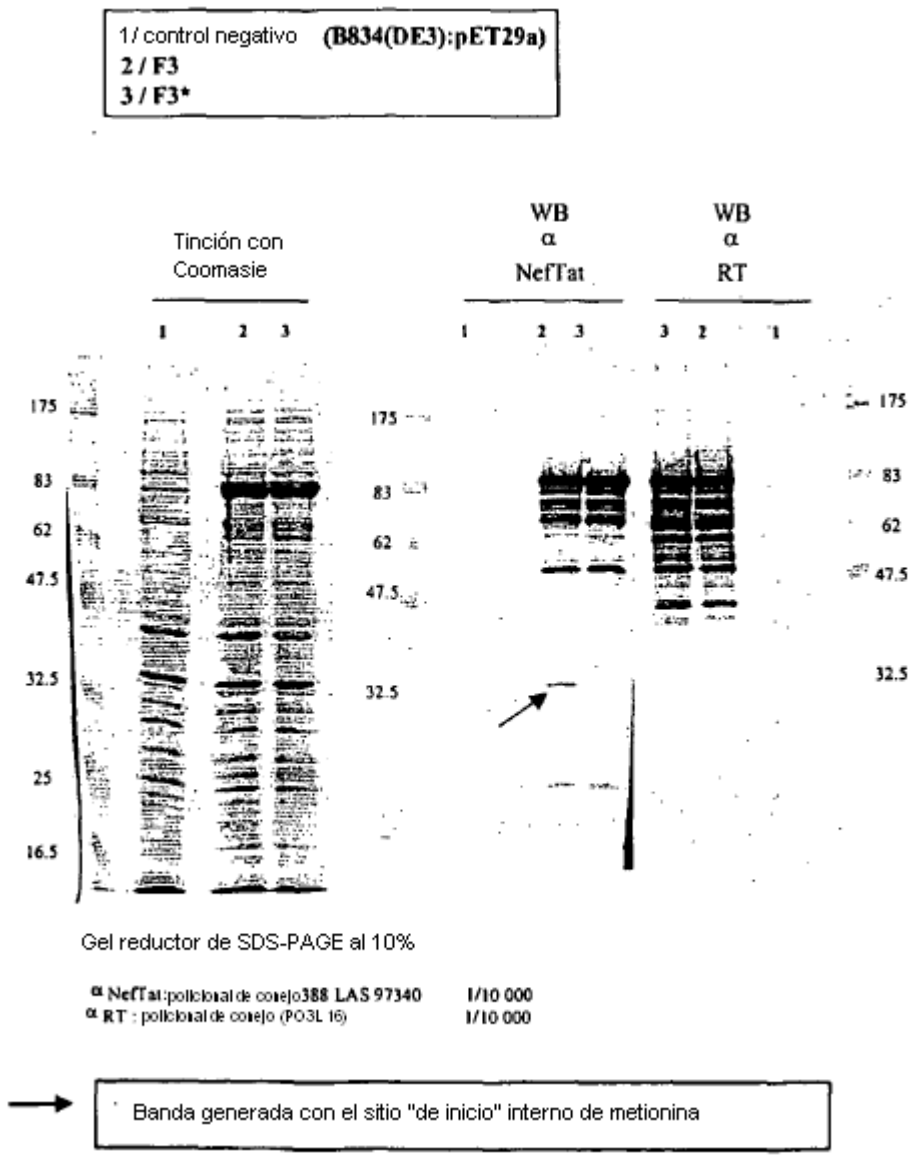


Figura 11

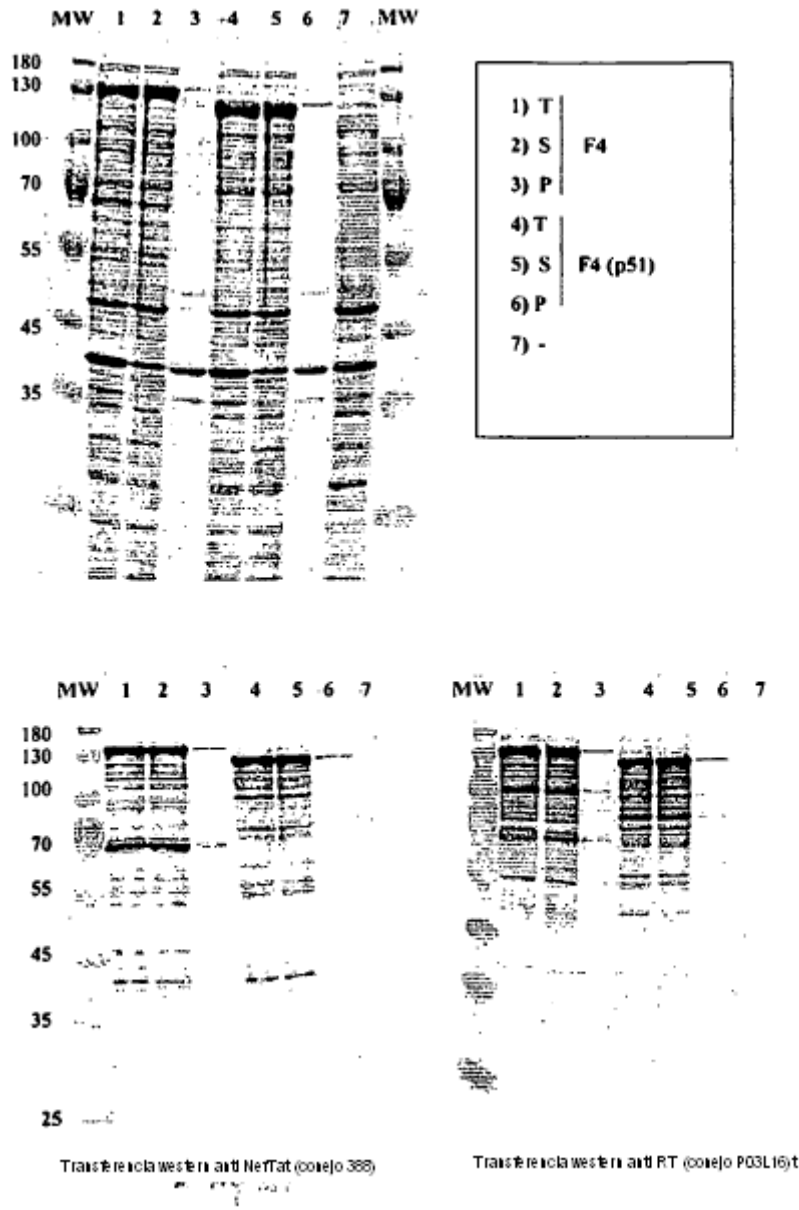
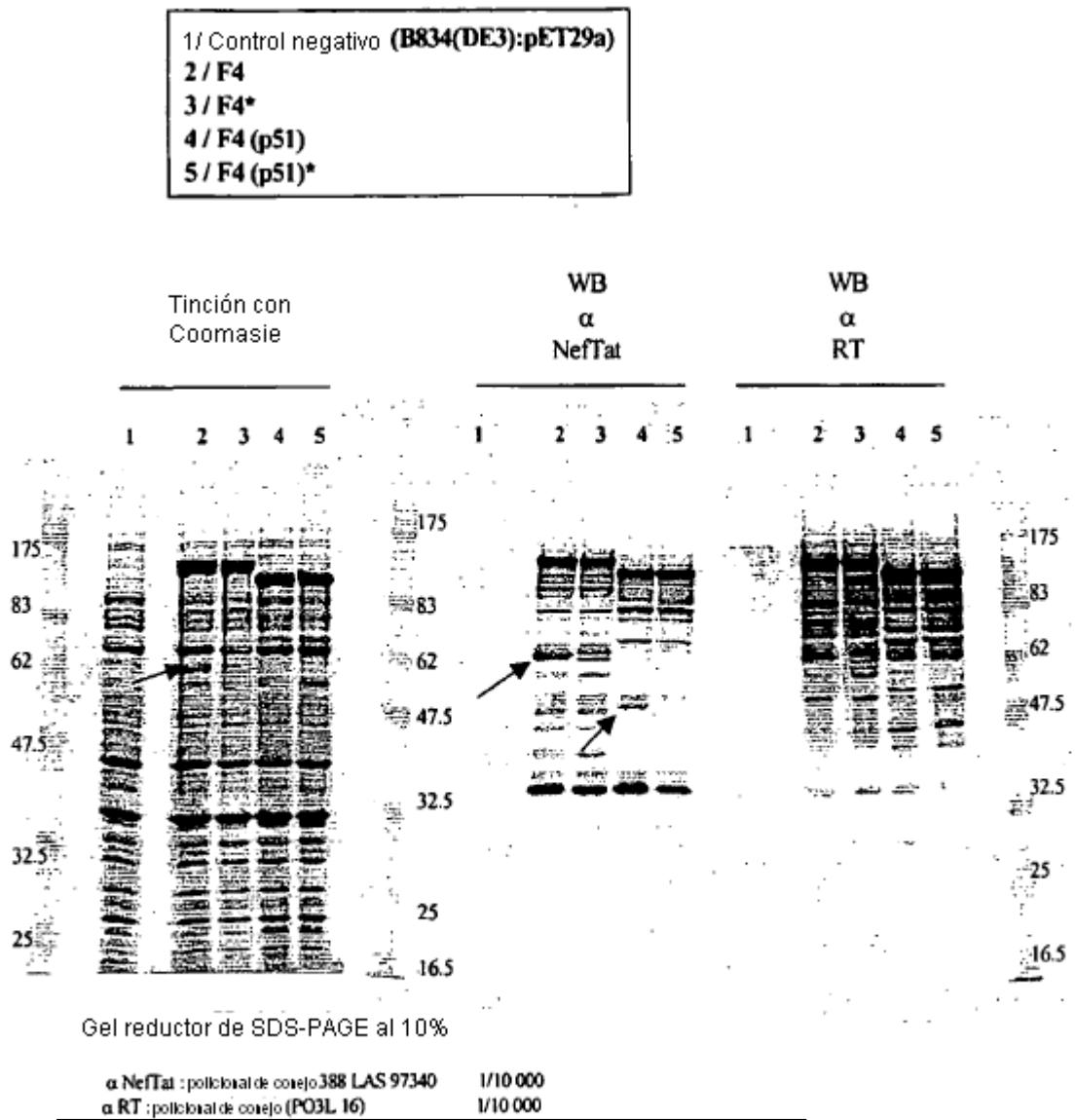


Figura 12

Inducción: 19 horas, 22° C



Banda generada por el sitio "de inicio" interno de metionina

Figura 13

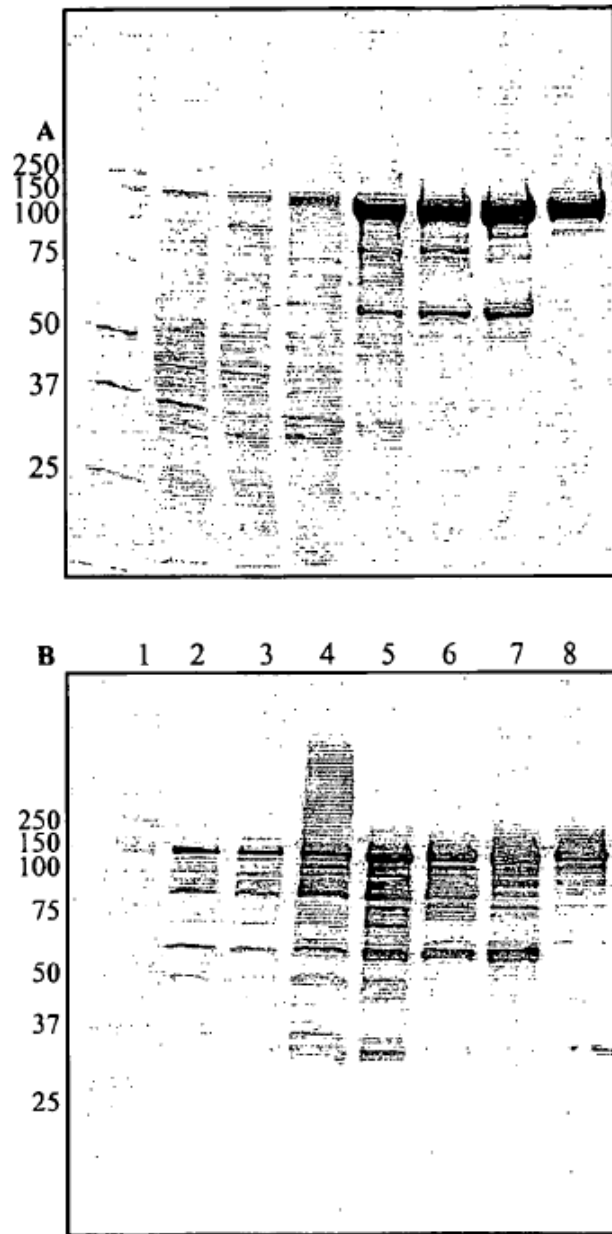


Figura 13: purificación de F4

5 A: Gel SDS al 4-20% tintado con azul Coomassie, 5 μ g de proteína/carril; B: transferencia western de anti-p24/anti Nef-Tat, 0,5 μ g de proteína/carril 1; patrón MW; carril 2, homogeneizado; carril 3, homogeneizado clarificado; carril 4, precipitado en AS resuspendido; carril 5, eluido en SO3; carril 6, eluido en octil sepharosa; carril 7, eluido en Q sepharosa; carril 8, eluido en Superdex 200.

Figura 14

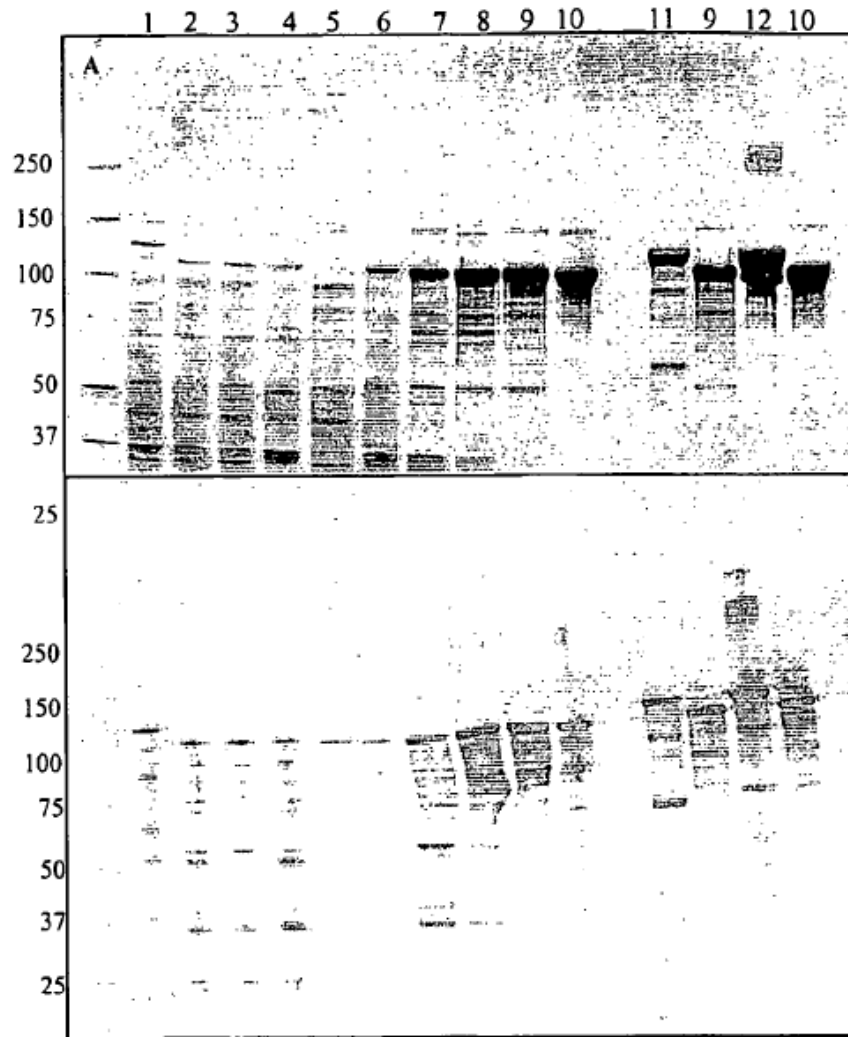


Figura 14: purificación de F4(p51)*

5 A: Gel SDS al 4-20% tintado con azul Coomassie, 5 µg de proteína/carril; B: transferencia western de anti-p24/anti Nef-Tat, 0,5 µg de proteína/carril. Carril 1, homogeneizado F4; carril 2, homogeneizado F4(51)*; carril 3, homogeneizado clarificado; carril 4, agregado de homogeneizado; carril 5, sobrenadante de precipitación en AS; carril 6, precipitado en AS resuspendido; carril 7, eluido en SO3; carril 8, eluido en octil sepharosa; carril 9, eluido en Q sepharosa; carril 10, eluido en Superdex 200; carril 11, eluido de F4 en Q; carril 12, eluido de F4 en Superdex 200.

Figura 15

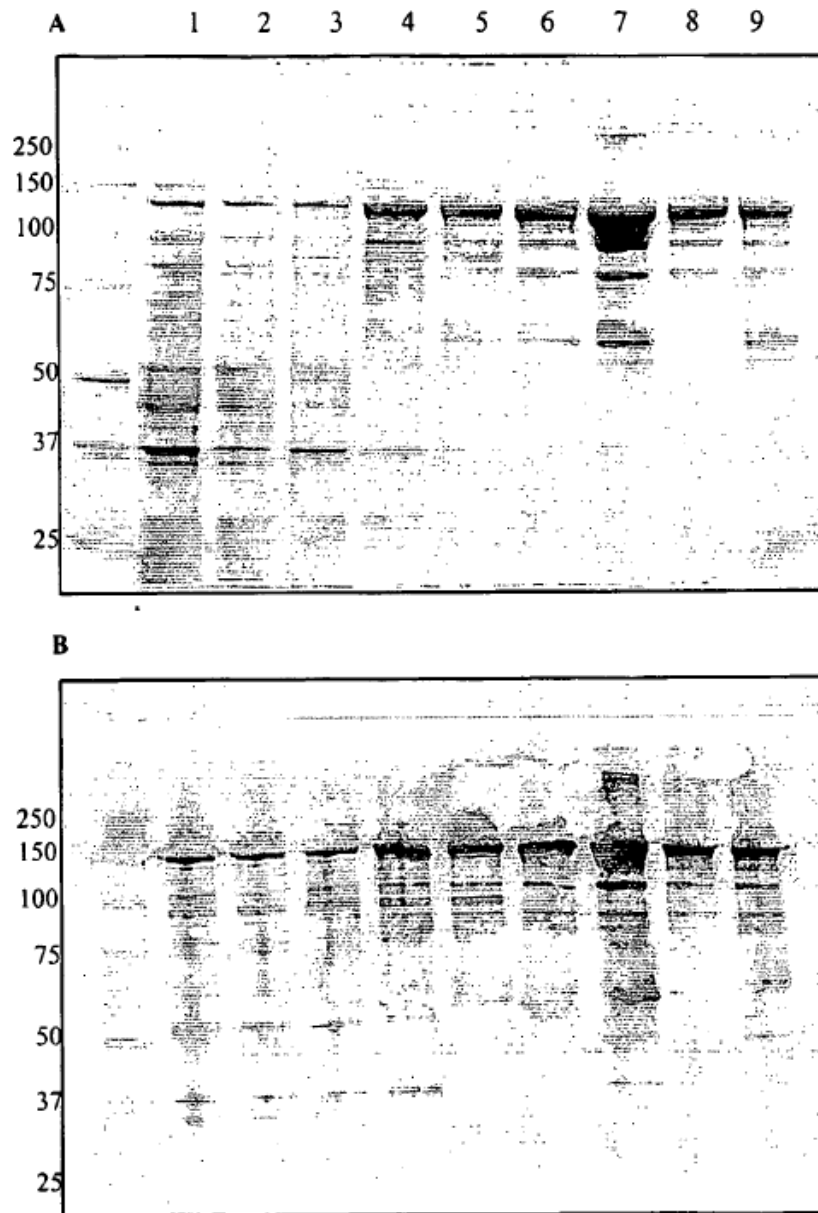
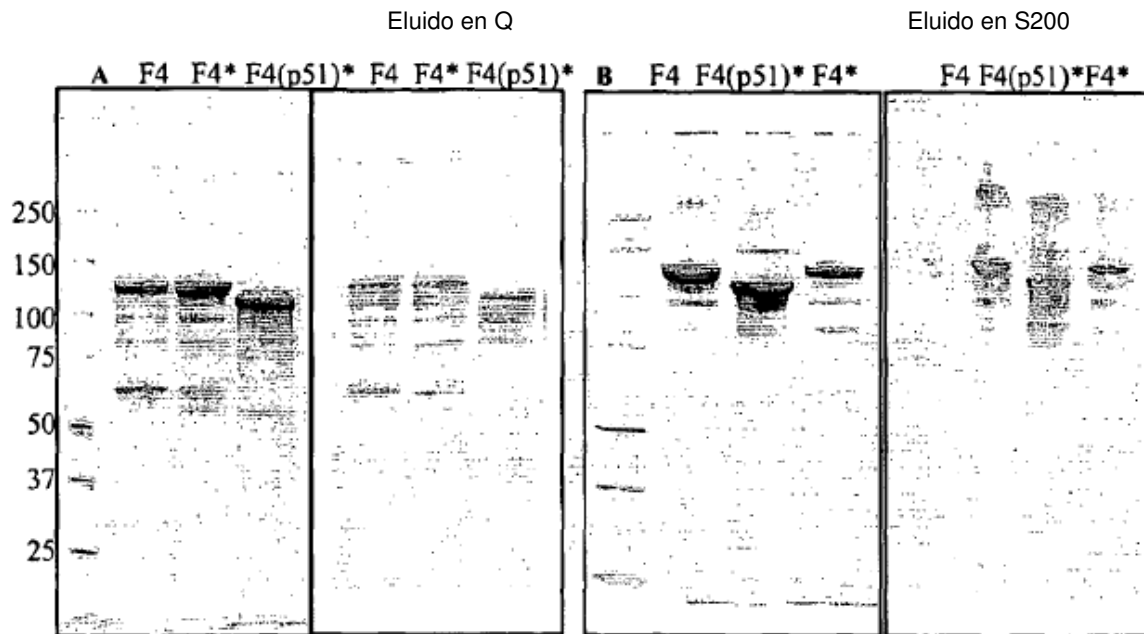


Figura 15: purificación de F4*.

5 A: Gel SDS al 4-20% tintado con azul Coomassie, 5 μ g de proteína/carril; B: transferencia western de anti-p24/anti Nef-Tat, 0,5 μ g de proteína/carril 1. Carril 1, homogeneizado; 2, homogeneizado clarificado; 3, precipitado en AS resuspendido; 4, eluido en SO3; 5, eluido en octil sepharosa; 6, eluido en Q; 7, eluido en Q sepharosa concentrado/dializado; 8, eluido en Superdex 200; 9, fracción rechazada de eluido en Superdex 200.

Figura 16



5

Figura 16: pureza de F4 vs F4* vs F4(51)*

A: eluido en Q, B: eluido en S200. Se cargaron los geles de SDS de 4-20% con 5 µg de proteína para el tinte azul Coomassie (en la izquierda) y con 0,5 µg para la transferencia western anti-p24/anti-Nef-Tat (en la derecha).

Figura 17

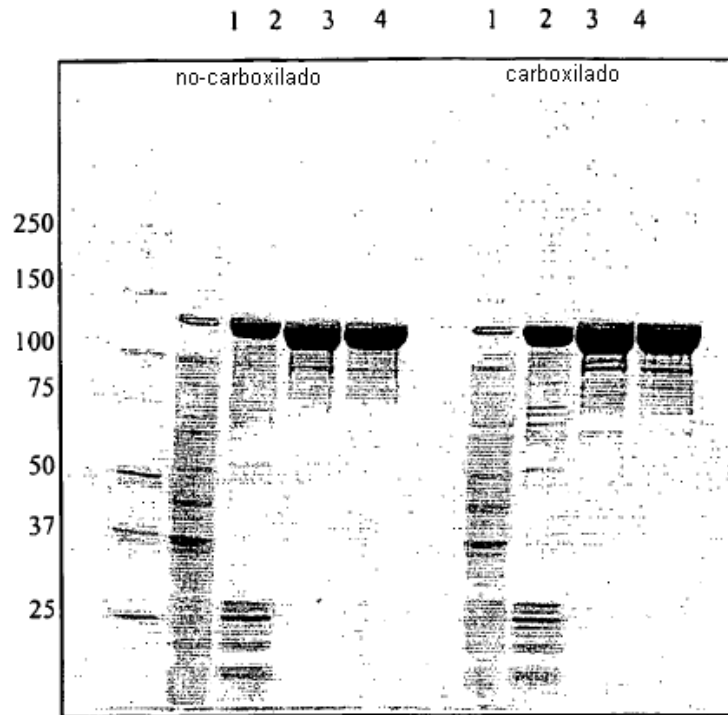
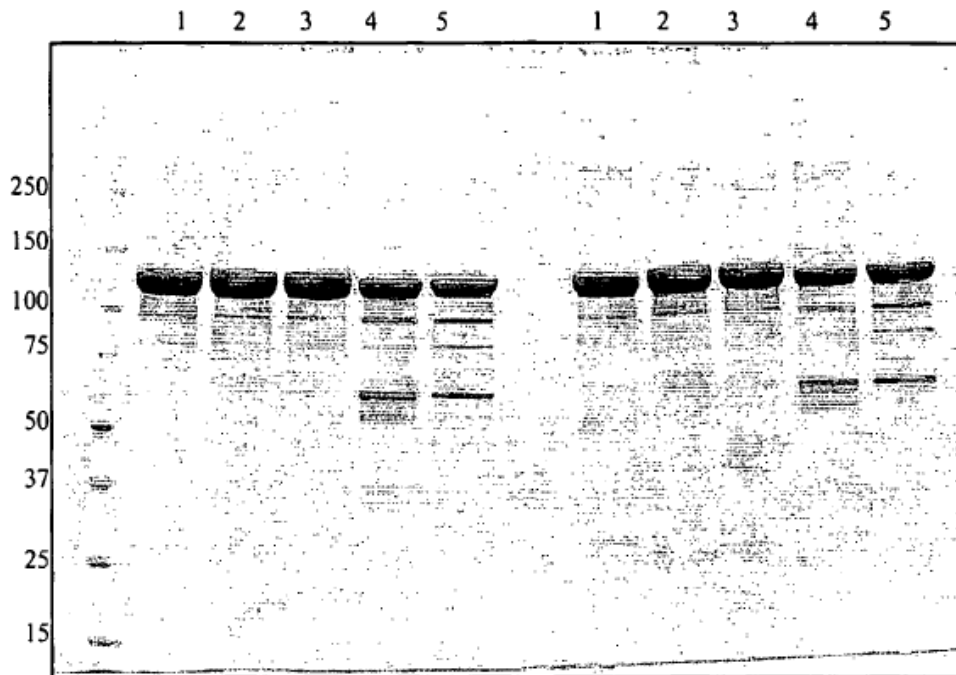


Figura 17: SDS-PAGE seguido de la purificación de F4co y F4co carboxiamidado. Se separaron 5 μ g de cada fracción recogida durante la purificación de F4co o F4coca en un gel SDS al 4-12%. Se tiñó el gel con azul Coomassie. 1: homogeneizado; 2: eluido en CM hyperZ; 3: eluido en Q sepharosa; 4: volumen purificado.

Figura 18



5 Figura 18: análisis por SDS-PAGE de F4, F4co y F4coca purificados de acuerdo con el procedimiento de purificación I o procedimiento de purificación II. Se separaron 5 μ g de cada proteína en un gel SDS al 4-12% en condiciones reductoras (izquierda) o en condiciones no reductoras (derecha). Se tiñó el gel con azul Coomassie. 1: procedimiento II – F4co; 2: procedimiento II – F4coca; 3: procedimiento I – F4coca; 4: procedimiento I – F4; 5: procedimiento I – F4 carboxiaminado.

Figura 19

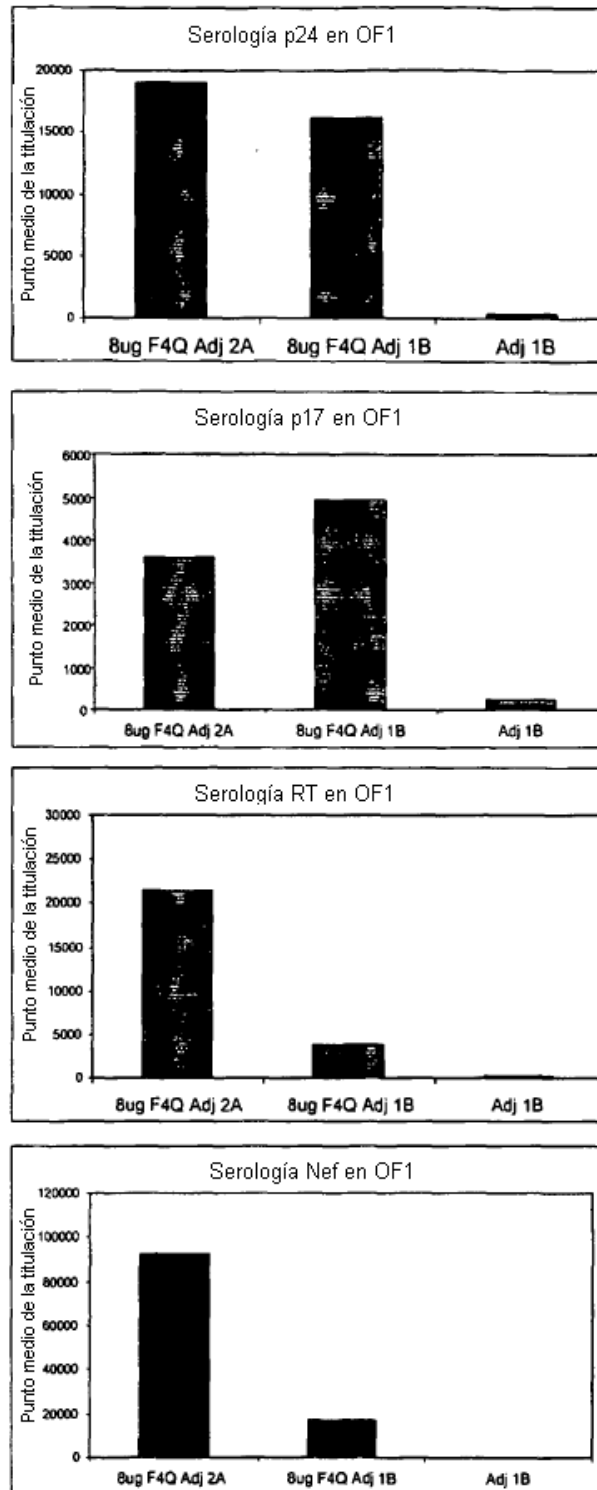


Figura 20

