

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 824 525**

51 Int. Cl.:

A61K 39/12 (2006.01)

C07K 14/005 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **14.09.2017 PCT/EP2017/073141**

87 Fecha y número de publicación internacional: **22.03.2018 WO18050747**

96 Fecha de presentación y número de la solicitud europea: **14.09.2017 E 17769027 (8)**

97 Fecha y número de publicación de la concesión europea: **22.07.2020 EP 3512543**

54 Título: **Mutaciones estabilizantes de trímeros de proteínas de la envoltura del VIH**

30 Prioridad:

15.09.2016 EP 16188866

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

12.05.2021

73 Titular/es:

**JANSSEN VACCINES & PREVENTION B.V.
(100.0%)
Archimedesweg 4
2333 CN Leiden, NL**

72 Inventor/es:

**RUTTEN, LUCY;
TRUAN, DAPHNÉ;
STROKAPPE, NIKA, MINDY y
LANGEDIJK, JOHANNES, PETRUS, MARIA**

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 824 525 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

Mutaciones estabilizantes de trímeros de proteínas de la envoltura del VIH

ANTECEDENTES DE LA INVENCION

5 El virus de la inmunodeficiencia humana (VIH) afecta a millones de personas en todo el mundo y la prevención del VIH mediante una vacuna eficaz sigue siendo una prioridad fundamental, incluso en la época del uso generalizado del tratamiento antirretroviral. La diversidad antigénica entre diferentes cepas y clados del virus VIH dificulta el desarrollo de vacunas con una eficacia amplia. El VIH-1 es la cepa más común y patogénica del virus, donde más de un 90% de los casos de VIH/SIDA deriva de una infección por el grupo M del VIH-1. El grupo M se subdivide adicionalmente en clados o subtipos, siendo más grande el clado C. Idealmente, una vacuna eficaz sería capaz de suscitar tanto respuestas celulares potentes como anticuerpos ampliamente neutralizantes capaces de neutralizar cepas de VIH-1 de diferentes clados.

10 La espícula de la proteína de la envoltura (Env) sobre la superficie del VIH está compuesta por un trímero de heterodímeros de las glicoproteínas gp120 y gp41 (Figura 1A). La proteína precursora gp160 es clivada por furina en la gp120, que es la cabeza de la espícula y contiene al sitio de unión al receptor de CD4 así como los bucles hipervariables grandes (V1 y V5), y gp41, que es el tallo anclado a membrana de la espícula de la proteína de la envoltura. Al igual que otras proteínas fusogénicas de clase I, gp41 contiene un péptido de fusión (FP) N-terminal, un dominio transmembrana (TM) C-terminal y un dominio citoplasmático. La fusión de membranas entre las membranas del VIH y de las células blanco requiere de una serie de cambios de conformación en la proteína de la envoltura. Se pueden desarrollar vacunas contra el VIH basado en la proteína de la envoltura.

15 Sin embargo, hay varios factores que dificultan el desarrollo de una vacuna contra el VIH basada en la proteína de la envoltura, que incluyen la gran variabilidad genética del VIH-1, el recubrimiento de carbohidratos denso de la proteína de la envoltura y la naturaleza relativamente dinámica y lábil de la estructura de la espícula de la proteína de la envoltura. La proteína de la envoltura de tipo salvaje es inestable debido a su función. Por ende, a veces se introducen modificaciones estabilizantes en la estructura de la envoltura para generar candidatos de vacunas. La proteína de la envoltura es un blanco de anticuerpos neutralizantes y está altamente glicosilada, lo cual reduce la inmunogenicidad porque oculta los epitopes de la proteína. Todos los anticuerpos ampliamente neutralizantes (bNAbs) se adaptan a estos glicanos.

20 Para el desarrollo de vacunas, se prefiere usar proteínas de la envoltura capaces de inducir bNAbs. Sin embargo, la mayoría de bNAbs solamente reconocen la conformación de la proteína de la envoltura nativa antes que sufra algún cambio de conformación. Por ello, el desarrollo de una proteína de la envoltura estable en su conformación compacta y cerrada tipo nativa, minimizando al mismo tiempo la presentación de epitopes no nativos y por consiguiente no neutralizantes, podría mejorar la eficiencia de generación de tales bNAbs. Los esfuerzos previos para producir una vacuna contra el VIH se han centrado en el desarrollo de vacunas que contienen el ectodominio de prefusión de la proteína de la envoltura trimérica del VIH, gp140. La proteína gp140 no tiene dominios transmembrana (TM) y citoplasmáticos, pero a diferencia de gp120, puede formar estructuras triméricas. Aún más, estos esfuerzos previos se han centrado fundamentalmente en el clado A. Sin embargo, la amplitud de la respuesta del anticuerpo neutralizante que fue inducida aún es limitada. Por lo tanto, también sería beneficioso si se pudiera disponer de trímeros de la envoltura nativa estabilizados contra múltiples clados del VIH.

25 Para más de dos décadas se ha intentado desarrollar una proteína de la envoltura estable en su conformación trimérica de prefusión, pero se ha logrado un éxito limitado para producir trímeros solubles, estables de la proteína de la envoltura capaces de inducir una respuesta de anticuerpo ampliamente neutralizante. Por ejemplo, se han introducido las denominadas mutaciones SOSIP (501C, 605C y 559P) en la secuencia de la proteína de la envoltura para mejorar la formación de una fracción del trímero gp140 soluble (Sanders et al., (2002), *J. Virol.* 76(17): 8875-89). Las denominadas mutaciones SOSIP incluyen residuos cisteína en las posiciones 501 y 605, y un residuo prolina en la posición 559 de acuerdo con la numeración en la gp160 de la forma aislada HXB2 del VIH-1, que es el esquema de numeración convencional usado en este campo. La introducción de los dos residuos cisteína en las posiciones 501 y 605, que se encuentran próximas entre sí en la estructura tridimensional de la proteína da como resultado un puente disulfuro. Se han descrito y utilizado proteínas de la envoltura mutantes SOSIP, tales como BG505_SOSIP y B41_SOSIP (proteínas de la envoltura de las cepas BG505 y B41 del VIH (es decir 9032-08.A1.4685) con mutaciones SOSIP), en estudios de vacunas y se ha mostrado que inducen Abs neutralizantes autólogos tier 2 (Pugach et al. 2015, *J Virol*, 89: 3380-3395; Sanders et al., *Science* (2015), 349(6224): 139-140).

30 Sin embargo, aún cuando las denominadas mutaciones SOSIP pueden estabilizar la forma del trímero de la proteína de la envoltura, la fracción del trímero de dichos mutantes SOSIP habitualmente se encuentra por debajo del 10%, todavía se producen grandes cantidades de monómeros y agregados. Hasta el mutante SOSIP, BG505_SOSIP, que es uno de las proteínas de la envoltura mutantes SOSIP más promisorias conocidas hasta la fecha en términos de su capacidad para estabilizar la forma trimérica, típicamente permite obtener tan solo hasta un 25% de la forma trimérica (Julien et al., *Proc.*

Nat. Acad. Sci. (2015), 112(38), 11947-52). Además, en esta fracción trimérica, los trímeros no son completamente estables ya que respiran por el ápice. Por lo tanto, además de las mutaciones SOSIP, se han diseñado varias sustituciones adicionales, tales como E64K, A316W y 201C-433C, para estabilizar el ápice y prevenir que respire (de Taeye et al., *Cell* (2015), 163(7), 1702-15; Kwon et al., (2015) *Nat. Struct. Mol. Biol.* 22(7) 522-31; WO 2016/037154).

- 5 Por lo tanto, existe la necesidad de trímeros estabilizados de las proteínas de la envoltura del VIH con un porcentaje de formación de trímeros mejorado, un rendimiento de trímeros mejorado y/o una estabilidad de trímeros mejorada. Preferiblemente, dichos trímeros estabilizados de las proteínas de la envoltura del VIH también presentarán una buena unión con anticuerpos ampliamente neutralizantes (bNAbs), y una unión relativamente limitada con Abs que no son ampliamente neutralizantes (no bNAbs). Un objeto de la invención consiste en proporcionar proteínas Env del VIH que
10 tengan porcentajes de trímeros mejorados y preferiblemente también rendimientos de trímeros mejorados.

RESUMEN BREVE DE LA INVENCION

La invención se relaciona con proteínas de la envoltura del VIH recombinantes de diferentes clados que presentan un porcentaje mejorado de formación de trímeros y/o rendimientos de trímeros mejorados en comparación con los trímeros de la envoltura del VIH descritos previamente. Se optimiza el plegamiento de Env, se reparan las características específicas de la cepa y se estabilizan las regiones de la conformación que son importantes para el proceso de fusión mediante las mutaciones que se describen en la presente. Esto provee un abordaje universal para optimizar el plegamiento y la estabilidad de los trímeros de la envoltura del VIH-1 cerrados de prefusión. Los trímeros de la Env VIH estables y bien plegados resultantes son de utilidad para una inmunización, por ejemplo, para aumentar la probabilidad de inducir anticuerpos ampliamente neutralizantes y reducir la inducción de anticuerpos no neutralizantes y débilmente neutralizantes tras la administración de los trímeros Env VIH recombinantes. La invención también se relaciona con moléculas de ácidos nucleicos aisladas y vectores que codifican las proteínas de la envoltura del VIH recombinantes, con células que comprenden los mismos y con composiciones de la proteína de la envoltura del VIH recombinante, la molécula de ácido nucleico, el vector y/o las células.

En un aspecto general, la invención se relaciona con proteínas de la envoltura del virus de inmunodeficiencia humana (VIH) recombinantes que tienen residuos de aminoácidos particulares en las posiciones identificadas en la secuencia de la proteína de la envoltura que estabilizan la formación de trímeros.

En determinadas formas de realización, una proteína de la envoltura (Env) del VIH recombinante de la invención comprende la secuencia de aminoácidos de una proteína Env del VIH que tiene los residuos de aminoácidos indicados en por lo menos dos de las posiciones indicadas que se seleccionan del grupo que consiste en:

- 30 (i) Phe, Leu, Met o Trp en la posición 651;
(ii) Phe, Ile, Met o Trp en la posición 655;
(iii) Asn o Gln en la posición 535;
(iv) Val, Ile o Ala en la posición 589;
(v) Phe o Trp en la posición 573;
35 (vi) Ile en la posición 204; y
(vii) Phe, Met o Ile en la posición 647,

en donde la numeración de las posiciones es acorde con la numeración en la gp160 de la forma aislada HXB2 del VIH-1. En algunas formas de realización preferidas, el residuo de aminoácido indicado en la posición 651 es Phe; el residuo de aminoácido indicado en la posición 655 es Ile; el residuo de aminoácido indicado en la posición 535 es Asn; y/o el residuo de aminoácido indicado en la posición 573 es Phe.

En determinadas formas de realización, una proteína Env del VIH recombinante de la invención comprende la secuencia de aminoácidos de una proteína Env del VIH y una sustitución de aminoácido por el residuo de aminoácido indicado en por lo menos una de las posiciones indicadas que se seleccionan del grupo que consiste en:

- 45 (i) Phe, Leu, Met o Trp en la posición 651;
(ii) Phe, Ile, Met o Trp en la posición 655;
(iii) Asn o Gln en la posición 535;

- (iv) Val, Ile o Ala en la posición 589;
- (v) Phe o Trp en la posición 573;
- (vi) Ile en la posición 204; y
- (vii) Phe, Met o Ile en la posición 647,

5 en donde la proteína Env del VIH se selecciona del grupo que consiste en:

- (1) una secuencia de aminoácidos consenso de la Env del VIH, por ejemplo del clado C (por ejemplo, que comprende la secuencia de aminoácidos de la SEQ ID N°: 2 o 3) o del clado B (por ejemplo que comprende la secuencia de aminoácidos de la SEQ ID N°: 4 o 5);
- (2) una proteína Env del VIH sintética, por ejemplo que comprende (a): la secuencia de aminoácidos de la SEQ ID N°: 6, o (b): SEQ ID N°: 6 con una mutación de Glu por Arg en la posición 166 o (c): (a) o (b) con una mutación de los aminoácidos en las posiciones 501 y 605 por residuos Cys y una mutación del aminoácido en la posición 559 por un residuo Pro, o (d): (a), (b) o (c) que tienen una mutación adicional en el sitio de clivaje de furina, por ejemplo el reemplazo de los aminoácidos en las posiciones 508-511 por RRRRRR (SEQ ID N°: 10), o (e) SEQ ID N°: 7, o (f) una secuencia Env en mosaico tal como una Env que comprende la secuencia de aminoácidos de la SEQ ID N°: 8 o 9; y
- (3) una proteína Env del VIH parental, que preferiblemente es una proteína Env del VIH de tipo salvaje, preferiblemente del clado C, que comprende por lo menos una mutación de reparación en un residuo de aminoácido presente en la posición correspondiente a una frecuencia menor que 7.5%, preferiblemente menor que 2%, de secuencias Env del VIH en una colección de por lo menos 100, preferiblemente por lo menos 1000, preferiblemente por lo menos 10000, secuencias Env del VIH de tipo salvaje, en donde la mutación de reparación es una sustitución por un residuo de aminoácido presente en la posición correspondiente a una frecuencia de por lo menos 10% de las secuencias Env del VIH en dicha colección y preferiblemente la mutación de reparación es una sustitución por el residuo de aminoácido presente en la posición correspondiente con mayor frecuencia en dicha colección;

25 y la numeración de las posiciones es acorde con la numeración en la gp160 de la forma aislada HXB2 del VIH-1. En algunas formas de realización preferidas, el residuo de aminoácido indicado en la posición 651 es Phe; el residuo de aminoácido indicado en la posición 655 es Ile; el residuo de aminoácido indicado en la posición 535 es Asn; y/o el residuo de aminoácido indicado en la posición 573 es Phe.

30 En determinadas formas de realización, una proteína Env del VIH recombinante de la invención comprende la secuencia de aminoácidos de una proteína Env del VIH y una sustitución de aminoácido por el residuo de aminoácido indicado en por lo menos una de las posiciones indicadas que se seleccionan del grupo que consiste en:

- (i) Phe, Leu, Met o Trp en la posición 651;
- (ii) Phe, Ile, Met o Trp en la posición 655;
- (iii) Asn o Gln en la posición 535;
- (iv) Val, Ile o Ala en la posición 589;
- (v) Phe o Trp en la posición 573;
- (vi) Ile en la posición 204; y
- (vii) Phe, Met o Ile en la posición 647,

40 en donde la proteína Env del VIH es una proteína Env del VIH mutante SOSIP que comprende por lo menos una mutación que resulta en los residuos de aminoácidos indicados en las posiciones indicadas que se seleccionan del grupo que consiste en:

- (a) Cys en las posiciones 501 y 605;
- (b) Pro en la posición 559 y
- (c) Cys en las posiciones 501 y 605 y Pro en la posición 559; y

45 la numeración de las posiciones es acorde con la numeración en la gp160 de la forma aislada HXB2 del VIH-1. En algunas formas de realización preferidas, el residuo de aminoácido indicado en la posición 651 es Phe; el residuo de aminoácido

indicado en la posición 655 es Ile; el residuo de aminoácido indicado en la posición 535 es Asn; y el residuo de aminoácido indicado en la posición 573 es Phe.

En otras formas de realización, una proteína Env del VIH recombinante de la invención comprende además un residuo de aminoácido indicado en por lo menos una de las posiciones indicadas que se seleccionan del grupo que consiste en:

- 5 (viii) Gln, Glu, Ile, Met, Val, Trp o Phe, preferiblemente Gln o Glu, en la posición 588;
- (ix) Lys en la posición 64 o Arg en la posición 66 o Lys en la posición 64 y Arg en la posición 66;
- (x) Trp en la posición 316;
- (xi) Cys en ambas posiciones 201 y 433;
- 10 (xii) Pro en la posición 556 o Pro en la posición 558 o Pro en las posiciones 556 y 558; (xiii) reemplazo del bucle en las posiciones de aminoácido 548-568 (bucle HR1) por un bucle de 7-10 aminoácidos, preferiblemente un bucle de 8 aminoácidos, por ejemplo que tiene una secuencia seleccionada entre cualquiera de las (SEQ ID N°: 12-17);
- (xiv) Gly en la posición 568 o Gly en la posición 569 o Gly en la posición 636 o Gly en ambas posiciones 568 y 636 o Gly en ambas posiciones 569 y 636; y/o
- 15 (xv) Tyr en la posición 302 o Arg en la posición 519 o Arg en la posición 520 o Tyr en la posición 302 y Arg en la posición 519 o Tyr en la posición 302 y Arg en la posición 520 o Tyr en la posición 302 y Arg en ambas posiciones 519 y 520.

20 En determinadas formas de realización, una proteína Env del VIH recombinante de la invención comprende además una mutación en la secuencia de clivaje de furina de la proteína Env del VIH, tal como un reemplazo en las posiciones 508-511 por RRRRRR (SEQ ID N°: 10).

En una forma de realización, la proteína Env del VIH recombinante es una proteína gp140.

En otra forma de realización, la proteína Env del VIH recombinante es una proteína gp160.

En determinadas formas de realización, la proteína Env del VIH recombinante está truncada en la región citoplasmática, por ejemplo después de 7 aminoácidos en la región citoplasmática.

25 En otro aspecto general, la invención se relaciona con un complejo trimérico que comprende un oligómero no covalente de tres de cualquiera de las proteínas Env del VIH recombinantes que se describen en la presente.

Otro aspecto general de la invención comprende proporcionar un método para mejorar el plegamiento y la estabilidad (medidos como un aumento del porcentaje de trímeros y/o del rendimiento de trímeros) de una proteína Env del VIH parental, donde dicho método comprende reparar la secuencia de aminoácidos de la proteína Env del VIH parental mediante la introducción de por lo menos una mutación de reparación, preferiblemente por lo menos 3 mutaciones de reparación en la proteína Env del VIH parental, en donde una mutación de reparación es una sustitución de aminoácido en un residuo de aminoácido que está presente en la posición correspondiente a una frecuencia menor que 7.5%, preferiblemente menor que 2%, de las secuencias Env del VIH en una colección de por lo menos 100, preferiblemente por lo menos 500, preferiblemente por lo menos 1000, preferiblemente por lo menos 10000, secuencias Env del VIH de tipo salvaje, en donde la sustitución es por un residuo de aminoácido que está presente en la posición correspondiente a una frecuencia de por lo menos 10% de las secuencias Env del VIH en dicha colección y preferiblemente la sustitución es por el residuo de aminoácido que está presente en la posición correspondiente con mayor frecuencia en dicha colección. La invención también provee una proteína Env del VIH reparada que se puede obtener mediante dicho método de la invención para mejorar el plegamiento y la estabilidad (medidos como un porcentaje de trímeros y/o el rendimiento de trímeros) de una proteína Env del VIH. La invención también provee una composición farmacéutica que comprende dicha proteína Env del VIH reparada. La invención también provee un método para producir una proteína Env del VIH, que comprende el método para reparar la proteína Env del VIH que se describe en la presente, y expresar un ácido nucleico que codifica la proteína Env del VIH estabilizada reparada en una célula huésped recombinante.

30

35

40

45 En otro aspecto general, la invención se relaciona con una proteína Env del VIH recombinante que comprende una secuencia de aminoácidos que es por lo menos 95%, 96%, 97%, 98%, 99% o 100% idéntica a la SEQ ID N°: 2, en donde preferiblemente no se tienen en cuenta las posiciones 204, 535, 573, 589, 647, 651 y 655, y preferiblemente tampoco las posiciones 64, 66, 201, 316, 433, 501, 508-511, 556, 558, 559, 588, 548-568 y 605 al determinar el % de identidad, y en donde la numeración es acorde con la numeración en la gp160 de la forma aislada HXB2 del VIH-1. En determinadas formas de realización del mismo, la proteína Env del VIH recombinante comprende una secuencia de aminoácidos que es

por lo menos 98%, 99% o 100% idéntica a la SEQ ID N°: 3, en donde preferiblemente no se tienen en cuenta las posiciones 204, 535, 573, 589, 647, 651 y 655, y preferiblemente tampoco las posiciones 64, 66, 201, 316, 433, 508-511, 556, 558, 588 y 548-568 al determinar el % de identidad, y en donde la numeración es acorde con la numeración en la gp160 de la forma aislada HXB2 del VIH-1.

5 En otro aspecto general, la invención se relaciona con una proteína Env del VIH recombinante que comprende una secuencia de aminoácidos que es por lo menos 95%, 96%, 97%, 98%, 99% o 100% idéntica a la SEQ ID N°: 4, en donde preferiblemente no se tienen en cuenta las posiciones 204, 535, 573, 589, 647, 651 y 655, y preferiblemente tampoco las posiciones 64, 66, 201, 316, 433, 501, 508-511, 556, 558, 559, 588, 548-568 y 605 al determinar el % de identidad, y en donde la numeración es acorde con la numeración en la gp160 de la forma aislada HXB2 del VIH-1. En determinadas formas de realización del mismo, la proteína Env del VIH recombinante comprende una secuencia de aminoácidos que es por lo menos 98%, 99% o 100% idéntica a la SEQ ID N°: 5, en donde preferiblemente no se tienen en cuenta las posiciones 204, 535, 573, 589, 647, 651 y 655, y preferiblemente tampoco las posiciones 64, 66, 201, 316, 433, 508-511, 556, 558, 588 y 548-568 al determinar el % de identidad, y en donde la numeración es acorde con la numeración en la gp160 de la forma aislada HXB2 del VIH-1.

10 15 En estos aspectos y formas de realización, uno o más de los aminoácidos en las posiciones indicadas que no se tienen en cuenta para determinar el % de identidad, preferiblemente se seleccionan entre los aminoácidos indicados como preferidos en la presente, por ejemplo Ile en la posición 204; Phe, Ala, Leu o Trp en la posición 651; etc. (véase las Tablas 1 y 2 más adelante).

20 25 En otro aspecto general, la invención se relaciona con una proteína Env del VIH recombinante que comprende una secuencia de aminoácidos que es por lo menos 95%, 96%, 97%, 98%, 99% o 100% idéntica a cualquiera de las SEQ ID N°: 2, 3, 4, 5, 20, 22, 24, 26, 27, 28, 29, 30, 31 o 32, en donde las SEQ ID N°: 20, 22, 24, 26, 27, 28, 29, 30, 31 o 32 son particularmente preferidas. En este aspecto, preferiblemente no se tienen en cuenta las posiciones 204, 535, 573, 589, 647, 651, 655 y 658 y preferiblemente tampoco las posiciones 64, 66, 201, 316, 433, 508-511, 556, 558, 588 y 548-568 al determinar el % de identidad, y en donde la numeración es acorde con la numeración en la gp160 de la forma aislada HXB2 del VIH-1. Además, en este aspecto, uno o más de los aminoácidos en las posiciones indicadas que no se tienen en cuenta para determinar el % de identidad, preferiblemente se seleccionan entre los aminoácidos indicados como preferidos en la presente en (i)-(vii) de la Tabla 1, (viii)-(xv) de la Tabla 2 y/o (xvi) de la Tabla 1, por ejemplo Ile en la posición 204; Phe, Leu, Met o Trp en la posición 651; etc.

30 En otro aspecto general, la invención se relaciona con una partícula, preferiblemente un liposoma o una nanopartícula, por ejemplo una nanopartícula autoensamblante, que expresa sobre su superficie una proteína Env del VIH recombinante de la invención.

35 En otro aspecto general, la invención se relaciona con una molécula de ácido nucleico aislada que codifica una proteína Env del VIH recombinante de la invención y con vectores que comprenden la molécula de ácido nucleico aislada ligada operativamente a un promotor. En una forma de realización, el vector es un vector viral. En otra forma de realización, el vector es un vector de expresión. En una forma de realización preferida, el vector viral es un vector de adenovirus.

Otro aspecto general se relaciona con una célula huésped que comprende la molécula de ácido nucleico aislada o un vector que codifica la proteína Env del VIH recombinante de la invención. Dichas células huésped se pueden usar para la producción de proteínas recombinantes, la expresión de proteínas recombinantes o la producción de partículas virales.

40 Otro aspecto general se relaciona con métodos para producir una proteína Env del VIH recombinante, que comprende cultivar una célula huésped que comprende una molécula de ácido nucleico aislada o un vector que codifica la proteína Env del VIH recombinante de la invención en condiciones adecuadas para la producción de la proteína Env del VIH recombinante.

45 Aún otro aspecto general se relaciona con una composición que comprende una proteína Env del VIH recombinante, un complejo trimérico, una molécula de ácido nucleico aislada, un vector o una célula huésped que se describen en la presente, y un vehículo farmacéuticamente aceptable.

BREVE DESCRIPCIÓN DE LAS FIGURAS

El resumen anterior, así como también la siguiente descripción detallada de la invención, se entenderán mejor cuando se lean junto con los dibujos adjuntos. Se comprenderá que la invención no se limita a las formas de realización específicas que se muestran en los dibujos.

50 En las figuras:

Las Figuras 1A y 1B muestran una representación esquemática de la estructura de las proteínas de la envoltura (Env) del VIH;

En la Figura 1A se muestra una proteína Env del VIH de longitud completa; y

5 En la Figura 1B se muestra una proteína Env del VIH soluble que contiene las denominadas mutaciones SOSIP y un truncamiento C-terminal que comienza en el residuo 664 de acuerdo con la numeración en la gp160 de la forma aislada HXB2 del VIH-1 (secuencia SOSIP.644);

10 En las Figuras 2A y 2B se muestra el porcentaje de formación de trímeros (Figura 2A) y el rendimiento de trímeros (Figura 2B) para las proteínas Env del VIH recombinantes de acuerdo con las formas de realización de la invención medidos mediante un ensayo AlphaLISA como se describe en el Ejemplo 3; las proteínas Env del VIH recombinantes evaluadas contenían una sustitución de aminoácido simple, doble o triple introducidas en el esqueleto de la secuencia consenso de la Env ConC_SOSIP del VIH del clado C (SEQ ID N°: 3); el porcentaje de trímeros y el rendimiento de trímeros se determinaron sobre la base de la unión del anticuerpo monoclonal específico del trímero (mAb) PGT145 contra cada una de las proteínas Env del VIH recombinantes; se compara el rendimiento de trímeros y el porcentaje de formación de trímeros para cada una de las proteínas Env del VIH recombinantes de la invención con los de una proteína de la envoltura que tiene el esqueleto de la secuencia ConC_SOSIP sin ninguna mutación estabilizante de trímeros adicional descrita en la presente;

20 En la Figura 3 se muestran los cromatogramas de la cromatografía de exclusión por tamaño con análisis por dispersión de luz multiangular (SEC-MALS) de las proteínas Env del VIH recombinantes de acuerdo con las formas de realización de la invención; las proteínas Env del VIH recombinantes evaluados contenían una sola sustitución de aminoácido introducida en el esqueleto de la secuencia consenso de la Env ConC_SOSIP del VIH, clado C (SEQ ID N°: 3), y se purificaron mediante cromatografía de afinidad con lectina como se describe en el Ejemplo 2; el análisis SEC-MALS se condujo como se describe en el Ejemplo 3; se indica el pico correspondiente a la forma trimérica en cada uno de los cromatogramas;

25 En la Figura 4 se muestra la estabilidad térmica de las proteínas Env del VIH recombinantes de acuerdo con las formas de realización de la invención informada como el porcentaje de trímeros remanente después del tratamiento con calor; las proteínas Env del VIH recombinantes evaluadas contenían una sustitución de aminoácidos simple, doble o triple introducida en el esqueleto de la secuencia consenso de la Env ConC_SOSIP del VIH, clado C (SEQ ID N°: 3); las proteínas Env del VIH recombinantes fueron sometidas a tratamiento con calor y se determinó el porcentaje de trímeros remanente después del tratamiento con calor mediante un ensayo AlphaLISA como se describe en el Ejemplo 4; también se muestra la estabilidad térmica de la proteína de la envoltura cuya secuencia del esqueleto ConC_SOSIP no tiene ninguna mutación estabilizante de trímeros adicional;

30 En las Figuras 5A-5B se muestra el porcentaje de formación de trímeros (Figura 5A) y el rendimiento de trímeros (Figura 5B) de proteínas Env del VIH recombinantes que tienen una sola sustitución de aminoácido introducida en el esqueleto de la secuencia consenso de la Env ConB_SOSIP del VIH, clado B (SEQ ID N°: 5) de acuerdo con las formas de realización de la invención en comparación con la de la proteína de la envoltura cuya secuencia del esqueleto ConB_SOSIP no tiene ninguna mutación estabilizante de trímeros adicional de la invención como se describe en el Ejemplo 5; el rendimiento de trímeros y el porcentaje de formación de trímeros se midieron mediante un ensayo AlphaLISA;

40 En las Figuras 6A-6B se muestra el porcentaje de formación de trímeros y el rendimiento de trímeros para proteínas Env del VIH recombinantes que tienen sustituciones de aminoácidos introducidas en el esqueleto de la secuencia sintética de la proteína de la envoltura del VIH DS_sC4_SOSIP_E166R de acuerdo con las formas de realización de la invención como se describe en el Ejemplo 6; el porcentaje de formación de trímeros y el rendimiento de trímeros se midieron mediante un ensayo AlphaLISA.

45 En la Figura 7 se muestran los cromatogramas de la cromatografía de exclusión por tamaño con análisis por dispersión de luz multiangular (SEC-MALS) de proteínas Env del VIH recombinantes de acuerdo con las formas de realización de la invención; las proteínas Env del VIH recombinantes evaluadas contenían la mutación única K655I y en cada variante siguiente se introdujo una mutación adicional en la secuencia consenso del esqueleto de la Env ConC_SOSIP del VIH, clado C (SEQ ID N°: 3), y se purificaron mediante cromatografía de afinidad con lectina como se describe en el Ejemplo 2; el análisis SEC-MALS se condujo como se describe en el Ejemplo 3; los picos que representan los monómeros de gp140 en ConC_SOSIP están indicados con un cuadrado sombreado, a la derecha del pico del trímero. En el panel inferior se muestra un acercamiento de la parte inferior del gráfico, de modo que se puede observar que cada mutación adicional causa una caída adicional en la altura del pico del monómero de gp140.

50 En las Figuras 8A-8B se muestra el porcentaje de formación de trímeros (Figura 8A) y el rendimiento de trímeros (Figura 8B) de BG505_SOSIP (derivada de una cepa del clado A de tipo salvaje) que tiene una sola sustitución de aminoácido y combinaciones de sustituciones introducidas de acuerdo con las formas de realización de la invención en comparación con las de la proteína de la envoltura cuya secuencia del esqueleto BG505_SOSIP sin ninguna mutación estabilizante de

trímeros adicional de la invención se describe en el Ejemplo 9. El rendimiento de trímeros y el porcentaje de formación de trímeros se midieron mediante un ensayo AlphaLISA.

5 En la Figura 9 se muestran los cromatogramas de una cromatografía de exclusión por tamaño con análisis por dispersión de luz multiangular (SEC-MALS) de proteínas Env del VIH recombinantes de acuerdo con las formas de realización de la invención; el análisis SEC-MALS se condujo con el sobrenadante de cultivo de células transfectadas con Env. El pico correspondiente a la forma trimérica eluye entre 7 y 7.5 minutos. La línea gris oscuro es BG505_SOSIP (derivada de una cepa del clado A de tipo salvaje) y la línea gris claro es BG505_SOSIP con las sustituciones L556P, K655I, M535N, N651F, D589V, K588E.

10 En las Figuras 10A-10B se muestra el rendimiento de trímeros para variantes de C97ZA_SOSIP, que se describen en el Ejemplo 10. Rendimiento de trímeros de C97ZA con tres sustituciones estabilizantes (L556P, T651F y M535N) (Figuras 10A y B). En la Figura 10B, la secuencia Env se optimizó aún más mediante mutaciones adicionales (21 mutaciones adicionales) que se agregaron para reparar la secuencia de Env C97ZA de acuerdo con el marco conceptual que se describe en la Figura 12 y mediante la introducción de sustituciones estabilizantes adicionales (K655I, D589V, A204I y K588E). El rendimiento de trímeros y el porcentaje de formación de trímeros se miden mediante un ensayo AlphaLISA. Las señales se normalizaron a la señal de ConC_SOSIP que se definió como 1. PNGS es un sitio potencial de N-glicosilación.

15 Figura 11. Rendimiento de trímeros de la Env de la cepa DU422 del VIH-1 con cuatro sustituciones estabilizantes (véase el Ejemplo 11 por más detalles). Todos los números se normalizaron a ConC_SOSIP (no se muestra) que se definió como 1.

20 Figura 12. Concepto universal para reparar la secuencia Env del VIH-1 ilustrada para la cepa C97ZA. El residuo con la mayor frecuencia de apariciones (denominado 'residuo consenso' en la presente) en la base de datos total del VIH-1 (barras superiores) y los residuos de la cepa C97ZA (barras inferiores) se clasifican según los porcentajes de baja a alta aparición en la posición de residuo C97ZA. Las posiciones en la secuencia C97ZA a sustituir por el residuo consenso se seleccionaron en base a los siguientes criterios: Posiciones con un residuo C97ZA que aparece en menos de un 2% de las secuencias de la base de datos de Env (barras negras). Posiciones con un residuo C97ZA que aparece entre 2% y 7.5% de las secuencias de la base de datos de Env y que están ocultas o parcialmente ocultas (barras gris oscuro). Posiciones que quedan expuestas y que son hidrofóbicas en C97ZA y los residuos consenso hidrofílicos (dos barras grises más claras) y una posición que es un residuo consenso (S234N) de un potencial sitio de N-glicosilación (PNGS).

25 Figura 13. Trímeros ENV_SOSIP del VIH cerrados de prefusión por reparación de secuencia y estabilización mutacional. Señales AlphaLISA en el sobrenadante de cultivos celulares de todas las variantes SOSIP normalizadas a ConC_SOSIP para los anticuerpos ampliamente neutralizantes.

30 Figura 14. Perfil SEC analítico de variantes de Env_SOSIP control (esqueleto SOSIP), variantes de Env reparadas de acuerdo con el concepto que se describe en el Ejemplo 12 y la Figura 12 y variantes de Env con sustituciones estabilizantes adicional de acuerdo con la tabla 3 usando sobrenadantes de cultivos celulares después de la transfección. Se sustrajo la señal falsa de los sobrenadantes de los cultivos celulares de todos los perfiles. Los picos de trímeros están indicados con un *.

35 Figura 15. Rendimiento de trímeros de variantes de Env ConC del VIH-1 sin las modificaciones estabilizantes SOSIP.

40 Figura 16. Rendimiento de trímeros (A) y porcentaje de trímeros (B) de las ConC_SOSIP con mutaciones en las posiciones 589, 647, 651 y 655 por metioninas. Todos los números se normalizaron a ConC_SOSIP (no se muestra) que se definió como 1. Se muestra una barra de error en el extremo derecho de las barras.

En las Figuras 17A-17D se muestra el porcentaje de formación de trímeros (Figura 17A, B para diferentes experimentos) y el rendimiento de trímeros (Figura 17C, D para diferentes experimentos) de las proteínas Env del VIH recombinantes con las mutaciones indicadas como se describe en el Ejemplo 15, medidos mediante un ensayo AlphaLISA.

45 En la Figura 18 se muestran los cromatogramas SEC-MALS de las proteínas Env del VIH recombinantes con las mutaciones indicadas, como se describe en el Ejemplo 15.

DESCRIPCIÓN DETALLADA DE LA INVENCION

50 La discusión de documentos, actas, materiales, dispositivos, artículos o similares que se ha incluido en la presente memoria descriptiva tiene el objetivo de proporcionar un contexto a la invención. Tal discusión no es una admisión de que cualquiera o la totalidad de estas materias forman parte de la técnica anterior con respecto a cualesquiera invenciones divulgadas o reivindicadas.

A menos que se especifique lo contrario, todos los términos científicos y técnicos que se utilizan en la presente tienen el mismo significado que les adjudica normalmente un experto en la materia a la cual esta invención pertenece. Por el contrario, ciertos términos utilizados en la presente tienen los significados que se exponen en la memoria descriptiva. Cabe señalar que en las reivindicaciones adjuntas, las formas en singular "un", "uno/a" y "el/la", tal como se utilizan en la presente, también incluyen referentes en plural a menos que el contexto indique claramente lo contrario.

A menos que se defina de otra manera, se considera que cualquiera de los valores numéricos, tal como una concentración o un rango de concentraciones, que se describe en la presente, están modificados en todos los casos por el término "aproximadamente". Por ende, un valor numérico típicamente incluye $\pm 10\%$ del valor indicado. Según se usa en la presente, el uso de un rango numérico incluye expresamente todos los subrangos posibles, todos los valores numéricos individuales dentro de dicho rango, inclusive los números enteros dentro de dichos rangos y fracciones de los valores a menos que el contexto claramente indique lo contrario.

En toda la divulgación se hace referencia a aminoácidos. Hay veinte aminoácidos naturales, así como muchos aminoácidos no naturales. Cada aminoácido conocido, que incluye aminoácidos naturales y no naturales, tiene un nombre completo, un código abreviado de una letra y un código abreviado de tres letras, todos los cuales son bien conocidos por los especialistas en el arte. Por ejemplo, los códigos abreviados de tres y una letra utilizados para los veinte aminoácidos naturales son: alanina (Ala; A), arginina (Arg; R), ácido aspártico (Asp; D), asparagina (Asn; N), cisteína (Cys; C), glicina (Gly; G), ácido glutámico (Glu; E), glutamina (Gln; Q), histidina (His; H), isoleucina (Ile; I), leucina (Leu; L), lisina (Lys; K), metionina (Met; M), fenilalanina (Phe; F), prolina (Pro; P), serina (Ser; S), treonina (Thr; T), triptofano (Trp; W), tirosina (Tyr; Y) y valina (Val; V). Los aminoácidos se pueden indicar por su nombre completo, el código abreviado de una letra o el código abreviado de tres letras.

A menos que el contexto claramente indique lo contrario, la numeración de las posiciones en la secuencia de aminoácidos de una proteína de la envoltura del VIH según se usa en la presente es acorde con la numeración en la gp160 de la forma aislada HXB2 del VIH-1 como se describe, por ejemplo, en Korber et al. (Human Retroviruses and AIDS 1998: A Compilation and Analysis of Nucleic Acid and Amino Acid Sequences. Korber et al., Eds. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, N. Mex.). La numeración acorde con HXB2 es considerada convencional en el campo de las proteínas Env del VIH. La gp160 de la forma aislada HXB2 del VIH-1 tiene la secuencia de aminoácidos que se muestra en la SEQ ID N°: 1. Se puede usar el alineamiento de una secuencia Env del VIH de interés con esta secuencia para hallar la numeración de aminoácidos correspondiente en la secuencia de interés.

Las frases "comprende la secuencia de aminoácidos de una proteína Env del VIH que tiene el residuo de aminoácido indicado en por lo menos una de las posiciones indicadas que se seleccionan del grupo que consiste en" y "comprende uno o más de los siguientes (residuos de aminoácidos)" se utilizan indistintamente en la presente.

El término "porcentaje (%) de identidad de secuencia" o "% de identidad" describe el número de coincidencias ("aciertos") de aminoácidos idénticos de dos o más secuencias de aminoácidos alineadas en comparación con el número de residuos de aminoácidos que conforma la longitud global de la secuencias de aminoácidos. En otros términos, el uso del alineamiento, para dos o más secuencias, permite determinar el porcentaje de residuos de aminoácidos que son iguales (por ejemplo, un 95%, 97% o 98% de identidad), cuando las secuencias se comparan y alinean para una correspondencia máxima medida usando un algoritmo de comparación de secuencias conocido en el arte o cuando se alinean manualmente y se inspeccionan visualmente. Por consiguiente, las secuencias que se comparan para determinar la identidad de secuencia pueden diferir en una o más sustituciones, adiciones o supresiones de aminoácidos. Los programas adecuados para alinear secuencias de proteína son conocidos por el especialista. El porcentaje identidad de secuencia de secuencias de proteínas se pueden determinar, por ejemplo, con programas tales como CLUSTALW, Clustal Omega, FASTA o BLAST, por ejemplo, usando el algoritmo NCBI BLAST (Altschul SF, et al (1997), Nucleic Acids Res. 25:3389-3402).

Una 'colección de secuencias Env del VIH', según se usa en la presente, es una colección de una cantidad representativa (por ejemplo, de por lo menos 100 o 500 o 1000, o más) de secuencias aleatorias de proteínas Env del VIH de tipo salvaje, que pueden ser del mismo clado (por ejemplo, el clado C) o de clados diferentes (por ejemplo, los clados A, B, C, etc). Las colecciones adecuadas de dichas secuencias se encuentran disponibles en bases de datos, o se pueden extraer subcolecciones de las mismas, por ejemplo las Base de datos de Secuencias del VIH (Los Alamos National Laboratory). Una colección tal comprende preferiblemente por lo menos 100 secuencias de proteínas Env del VIH, 1000 secuencias de proteínas Env del VIH, por lo menos 10000 secuencias de proteínas Env del VIH, por lo menos 50000 secuencias de proteínas Env del VIH, y puede contener más de 90000 secuencias de proteínas Env del VIH.

Una 'posición correspondiente' en una proteína Env del VIH se refiere a la posición del residuo de aminoácido cuando se alinean por lo menos dos secuencias Env del VIH. A menos que se indique de otra manera, la numeración de la posición del aminoácido para este fin es acorde con la numeración en la gp160 de la forma aislada HXB2 del VIH-1, como es habitual en el campo.

Una 'mutación estabilizante', según se usa en la presente, es una mutación que se describe en la presente en cualquiera de las entradas (i)-(vii) o (xvi), de la Tabla 1, o (viii)-(xv) de la Tabla 2, que aumenta el porcentaje de trímeros y/o el rendimiento de trímeros (que se pueden medir, por ejemplo, de acuerdo con los ensayos AlphaLISA o SEC-MALS que se describen en la presente) de una proteína Env del VIH en comparación con una molécula parental cuando la mutación se introduce mediante sustitución del aminoácido correspondiente en dicha molécula parental. Los aminoácidos que resultan de dichas mutaciones estabilizantes típicamente se encuentran raramente, a veces nunca, en las proteínas Env de formas aisladas del VIH de tipo salvaje.

Una 'mutación de reparación', según se usa en la presente, es una sustitución de un residuo de aminoácido en una proteína Env del VIH parental, donde dicho residuo de aminoácido está presente en menos del 7.5%, preferiblemente menos del 2%, en la posición correspondiente en una colección de secuencias de proteínas Env del VIH, en donde la sustitución es por un aminoácido que está presente en la posición correspondiente en dicha colección con más frecuencia, por ejemplo en por lo menos el 10% de las proteínas Env del VIH en dicha colección, y preferiblemente es por un aminoácido que está presente en la posición correspondiente en dicha colección en por lo menos el 20% de las proteínas Env del VIH o es el aminoácido más frecuente en la posición correspondiente en dicha colección. Por consiguiente, los aminoácidos que resultan de dichas mutaciones de reparación se encuentran típicamente en un porcentaje relativamente alto de las proteínas Env de formas aisladas del VIH de tipo salvaje, y en varios casos puede ser igual que aquellos en la posición correspondiente en las secuencias consenso de Env del VIH.

Una secuencia Env del VIH 'reparada y estabilizada', según se usa en la presente, típicamente contiene por lo menos una mutación de reparación y por lo menos una mutación estabilizante, preferiblemente múltiples mutaciones de reparación y múltiple mutaciones estabilizantes en comparación con la secuencia Env del VIH parental.

Los términos 'natural' o 'de tipo salvaje' se utilizan indistintamente en la presente con referencia a cepas del VIH (o las proteínas Env de las mismas), y se refieren a cepas del VIH (o proteínas Env de las mismas) presentes en la naturaleza, por ejemplo tal como en pacientes infectados con VIH.

La invención se relaciona en general con proteínas de la envoltura (Env) del VIH recombinantes que comprenden algunas sustituciones de aminoácidos en las posiciones indicadas en la secuencia de la proteína de la envoltura que estabilizan la forma del trímero de la proteína de la envoltura. La introducción de uno o más de las sustituciones de aminoácidos identificadas de la invención en la secuencia de una proteína de la envoltura del VIH puede dar como resultado un mayor porcentaje de formación de trímeros y/o un mayor rendimiento de trímeros. Esto se puede medir, por ejemplo, usando anticuerpos específicos del trímero, temperatura de fusión, cromatografía de exclusión por tamaño y unión a anticuerpos que se unen a una proteína Env plegada correctamente (trimérico estable) o, como alternativa, mal plegada (inestable o no trimérico), y se considera que un mayor porcentaje de trímeros y/o rendimiento de trímeros es indicativo de una proteína Env estable, nativa, correctamente plegada.

El virus de la inmunodeficiencia humana (VIH) es un miembro del género *Lentivirinae*, que es parte de la familia *Retroviridae*. Dos especies de VIH infectan a seres humanos: VIH-1 y VIH-2. VIH-1 es la cepa más común del virus VIH y se sabe que es más patógena que VIH-2. Según se usa en la presente, los términos "virus de inmunodeficiencia humana" y "VIH" se refieren, pero en un sentido no taxativo, al VIH-1 y VIH-2. En formas de realización preferidas, VIH se refiere al VIH-1.

El VIH se clasifica en múltiples clados con un grado elevado de divergencia genética. La expresión "clado de VIH" o "subtipo de VIH", tal como se utiliza en la presente, se refiere a virus de la inmunodeficiencia humana relacionados clasificados de acuerdo con su grado de similitud genética. El grupo más grande de formas aisladas del VIH-1 se denomina Grupo M (cepas principales) y consiste en por lo menos diez clados, A a J.

En un aspecto general, la invención se relaciona con una proteína de la envoltura (Env) del VIH recombinante. El término "recombinante" cuando se usa con referencia a una proteína se refiere a una proteína que es producida mediante una técnica recombinante o mediante síntesis química *in vitro*. De acuerdo con las formas de realización de la invención, una proteína "recombinante" tiene una secuencia de aminoácidos artificial por cuanto contiene por lo menos un elemento de secuencia (por ejemplo, sustitución, supresión, adición de aminoácidos, reemplazo de secuencia, etc.) que no se encuentra en la correspondiente secuencia natural. Preferiblemente, una proteína "recombinante" es una proteína de la envoltura del VIH no natural que se optimiza para inducir una respuesta inmunológica o para producir inmunidad contra una o más cepas naturales del VIH.

Los términos "proteína de la envoltura del VIH," "Env del VIH" y "proteína Env del VIH" se refieren a una proteína, o un fragmento o derivado de la misma, que en la naturaleza se expresa sobre la envoltura del virión VIH y permite que un VIH busque y se una a la membrana plasmática de células infectadas con el VIH. Los términos "envoltura" y "Env" se utilizan indistintamente en toda la divulgación. El gen *env* del VIH codifica la proteína precursora gp160, que es clivada proteolíticamente en las dos glicoproteínas maduras de la envoltura, gp120 y gp41. La reacción de clivaje está mediada

por una proteasa de la célula huésped, furina (o proteasas tipo furina), en un motivo de secuencia altamente conservado en los precursores de glicoproteínas de la envoltura retrovirales. Más específicamente, la gp160 se trimeriza en (gp160)₃ y luego sufre un clivaje en las dos glicoproteínas maduras asociadas de manera no covalente, gp120 y gp41. La entrada al virus está mediada posteriormente por un trímero de los heterodímeros gp120/gp41. La Gp120 es el fragmento de unión al receptor, y se une al receptor CD4 (y al correceptor) sobre una célula blanco que comprende dicho receptor, tal como, por ejemplo, una célula T auxiliar. Gp41, que está unido a gp120 de manera no covalente, es el fragmento de fusión y proporciona la segunda etapa por la cual VIH entra en la célula. Originalmente, la Gp41 está oculta dentro de la envoltura viral, pero cuando gp120 se une a un receptor y correceptor CD4, la gp120 cambia su conformación causando la exposición de gp41, donde puede asistir en la fusión con la célula huésped. Gp140 es el ectodominio de gp160.

De acuerdo con las formas de realización de la invención, una "proteína de la envoltura (Env) del VIH" puede comprender una proteína gp160 o gp140, o combinaciones, fusiones, truncamientos o derivados de la misma. Por ejemplo, una "proteína de la envoltura del VIH" puede incluir una proteína gp120 asociada de manera no covalente con una proteína gp41. Una "proteína de la envoltura del VIH" también puede ser una proteína de la envoltura del VIH truncada que incluye, pero en un sentido no taxativo, proteínas de la envoltura que comprenden un truncamiento C-terminal en el ectodominio (es decir, el dominio que se extiende dentro del espacio extracelular), un truncamiento en la gp41, tal como un truncamiento en el ectodominio de gp41, en el dominio transmembrana de gp41, o un truncamiento en el dominio citoplasmático de gp41. Una proteína de la envoltura del VIH can también puede ser una gp140, correspondiente al ectodominio gp160, o una versión extendida o truncada de gp140. La expresión de las proteínas gp140 ya fue descrita en varias publicaciones (por ejemplo Zhang et al., 2001; Sanders et al., 2002; Harris et al., 2011), y la proteína también se puede obtener de proveedores de servicio, como diferentes variantes basados, por ejemplo, en cepas diferentes del VIH. Una proteína gp140 de acuerdo con la invención puede tener una mutación en el sitio de clivaje de modo tal que el dominio gp120 y el ectodominio gp41 no se clivan y unen de manera covalente o, como alternativa, el dominio gp120 y el ectodominio gp41 se puede clivar y unir de manera covalente, por ejemplo mediante un puente disulfuro (tal como, por ejemplo, en las variantes SOSIP). Una "proteína de la envoltura del VIH" puede además ser un derivado de una proteína de la envoltura de VIH de origen natural que tiene mutaciones en la secuencia, p. ej., en los sitios de escisión de furina, y/o las denominadas mutaciones SOSIP. Una proteína de la envoltura del VIH de acuerdo con la invención también puede tener un sitio de clivaje de modo que el ectodominio gp41 y gp120 no se pueden unir de manera covalente.

En formas de realización preferidas de la invención, la proteína Env del VIH es una proteína gp140 o una proteína gp160, y más preferiblemente es una proteína gp140. En otras formas de realización preferidas de la proteína Env está truncada, por ejemplo, por supresión de los residuos después del 7º residuo de la región citoplasmática en comparación con una proteína Env natural.

De acuerdo con las formas de realización de la invención, una "proteína de la envoltura del VIH" puede ser un trímero o un monómero, y preferiblemente es un trímero. El trímero puede ser un homotrímero (por ejemplo, trímeros que comprenden tres unidades de polipéptidos idénticos) o un heterotrímero (por ejemplo, trímeros que comprenden tres unidades de polipéptidos que nos son todos idénticos). Preferiblemente, el trímero es un homotrímero. En el caso de una gp140 o gp160 clivada, se trata de un trímero de unidades de polipéptidos que son dímeros gp120-gp41, y en el caso en que los tres dímeros son iguales, se considera como un homotrímero.

Una "proteína de la envoltura del VIH" puede ser una proteína soluble o una proteína unida a membrana. Las proteínas de la envoltura unidas a membrana comprenden típicamente un dominio transmembrana, tal como en la proteína de la envoltura del VIH de longitud completa que comprende un dominio transmembrana (TM) como se muestra en la Figura 1A. Las proteínas unida a membrana pueden tener un dominio citoplasmático, pero no requieren de un dominio citoplasmático para su unión a membrana. Las proteínas de la envoltura solubles comprenden por lo menos una supresión parcial o completa del dominio transmembrana. Por ejemplo, el extremo C-terminal de una proteína de la envoltura del VIH de longitud completa se puede trincar para suprimir el dominio transmembrana, produciendo de esa manera una proteína soluble, como se muestra en la Figura 1B. Sin embargo, la proteína de la envoltura del VIH aún puede ser soluble con truncamientos más cortos y posiciones de truncamiento alternativos a las que se muestran en la Figura 1B. El truncamiento se puede efectuar en varias posiciones, y los ejemplos no taxativos comprenden después del aminoácido 664, 655, 683, etc. todos los cuales dan como resultado una proteína soluble. Una proteína Env unida a membrana de acuerdo con la invención puede comprender un dominio C-terminal completo o parcial (por ejemplo, por supresión parcial del dominio C-terminal citoplasmático, por ejemplo, en determinadas formas de realización después del 7º residuo de la región citoplasmática) en comparación con una proteína Env nativa.

Típicamente hay un péptido señal en el Extremo N-terminal de la proteína Env del VIH cuando se expresa, pero es clivada por una peptidasa de señal y por consiguiente no está presente en la proteína madura. El péptido señal se puede intercambiar con otras secuencias señal, y en la presente se proveen algunos ejemplos no taxativos de péptidos señal en las SEQ ID N°: 11, 18, 33 y 34.

De acuerdo con las formas de realización de la invención, la proteína de la envoltura del VIH, por ejemplo, gp160 o gp140, puede derivar de una secuencia de una proteína de la envoltura del VIH de cualquier clado del VIH (o 'subtipo'), por ejemplo, el clado A, clado B, clado C, clado D, clado E, clado F, clado G, clado H, etc. o combinaciones de los mismos (tal como en las 'formas recombinantes en circulación' o CRF que derivan de la recombinación entre virus de diferentes subtipos, por ejemplo, BC, AE, AG, BE, BF, ADG, etc). La proteína de la envoltura de la secuencia del VIH puede ser una secuencia natural, una secuencia en mosaico, una secuencia consenso, una secuencia sintética o cualquier derivado o fragmento de la misma. Una "secuencia en mosaico" contiene múltiples epitopes derivados de por lo menos tres secuencias de la envoltura del VIH de uno o más clados del VIH, y se pueden diseñar mediante algoritmos que optimizar la cobertura de los epitopes de células T. Los ejemplos de secuencias de proteínas de la envoltura del VIH en mosaico incluyen las que se describen, por ejemplo, en Barouch et al, *Nat Med* 2010, 16: 319-323; y WO 2010/059732, tal como por ejemplo las que se muestran en las SEQ ID N°: 8 y 9. Según se usa en la presente, una "secuencia consenso" significa una secuencia de aminoácidos artificial basada en un alineamiento de secuencias de aminoácidos de proteínas homólogas, por ejemplo, determinada mediante un alineamiento (por ejemplo, usando Clustal Omega) de las secuencias de aminoácidos de proteínas homólogas. Es el orden calculado de los residuos de aminoácidos más frecuentes, presentes en cada posición en un alineamiento de secuencia, basado en las secuencias de Env de por lo menos 1000 formas aisladas naturales del VIH. Una "secuencia sintética" es una proteína de la envoltura del VIH no natural que se optimiza para inducir una respuesta inmunológica o para producir inmunidad contra más de una cepa natural del VIH. Las proteínas de la envoltura del VIH en mosaico son ejemplos no taxativos de proteínas de la envoltura del VIH sintéticas. En formas de realización preferidas de la invención, la proteína Env del VIH es una proteína Env consenso, o una proteína Env sintética, que presenta por lo menos uno de los aminoácidos indicados en las posiciones indicadas en (i)-(vii) de acuerdo con la invención. Se prefieren en particular las proteínas Env consenso que tienen por lo menos uno, preferiblemente por lo menos dos de los residuos de aminoácidos indicados en las posiciones indicadas en (i)-(vii) de acuerdo con la invención, preferiblemente que tienen otras mutaciones SOSIP y/o del sitio de clivaje de furina como se describirá más adelante.

En determinadas formas de realización de la invención, una proteína de la envoltura del VIH, ya sea una secuencia natural, una secuencia en mosaico, una secuencia consenso, una secuencia sintética etc., comprende mutaciones de secuencia adicionales por ejemplo, en el sitio de clivaje de furinas, y/o las denominadas mutaciones SOSIP.

En algunas formas de realización de la invención, una proteína de la envoltura del VIH es una "proteína Env del VIH SOSIP mutante." Las denominadas mutaciones SOSIP son mutaciones estabilizantes de trímeros que incluyen las 'mutaciones SOS' (residuos Cys en las posiciones 501 y 605, que dan como resultado la introducción de un posible puente disulfuro entre los residuos cisteína recién creados) y la 'mutación IP' (residuo Pro en la posición 559). De acuerdo con las formas de realización de la invención, una proteína Env SOSIP mutante comprende por lo menos una mutación seleccionada del grupo que consiste en Cys en las posiciones 501 y 605; Pro en la posición 559; y preferiblemente Cys en las posiciones 501 y 605 y Pro en la posición 559. La proteína Env del VIH SOSIP mutante también pueden comprender otras mutaciones de secuencia, por ejemplo, en el sitio de clivaje de furina. Además, en determinadas formas de realización es posible agregar otras mutaciones de modo tal que la proteína Env comprenderá Pro en la posición 556 o la posición 558 o en las posiciones 556 y 558, que en la presente se encontró que no solo pueden actuar como alternativas de Pro en la posición 559 en una variante SOSIP, sino también como mutaciones adicionales que podrían mejorar aún más la formación de trímeros de una variante SOSIP que ya tiene Pro en la posición 559.

En algunas formas de realización preferidas de la invención, una proteína Env del VIH SOSIP mutante comprende Cys en las posiciones 501 y 605, y Pro en la posición 559.

En determinadas formas de realización, una proteína de la envoltura del VIH de la invención comprende además una mutación en el sitio de clivaje de furina. La mutación en la secuencia de clivaje de furina puede ser una sustitución, supresión, inserción o reemplazo de aminoácido de una secuencia por otra, o el reemplazo por una secuencia de aminoácidos conectora. Preferiblemente, en la presente invención, la mutación del sitio de clivaje de furina se puede usar para optimizar el sitio de clivaje, para así mejorar el clivaje de furina ante el tipo salvaje, por ejemplo mediante el reemplazo de la secuencia en los residuos 508-511 por RRRRRR (SEQ ID N°: 10) [es decir, el reemplazo de una secuencia de aminoácidos típica (por ejemplo, EK) en las posiciones 509-510 por cuatro residuos arginina (es decir, dos reemplazos y dos adiciones), en tanto en las posiciones 508 y 511, ya hay residuos arginina en la mayoría de las proteínas Env del VIH, por lo cual típicamente no es necesario reemplazarlos, pero dado que el resultado final en la literatura a menudo se conoce como una secuencia de aminoácidos RRRRRR, en la presente se conservó esta nomenclatura]. Se conocen otras mutaciones que mejoran el clivaje de furina y también se pueden utilizar. Como alternativa, es posible reemplazar el sitio de clivaje de furina por un conector, no que ya no será necesario el clivaje de furina pero la proteína adoptará una conformación tipo nativa (por ejemplo, como se describe en (Sharma et al, 2015) y (Georgiev et al, 2015)).

En formas de realización particulares de la invención, una proteína de la envoltura del VIH de la invención también comprende las mutaciones denominadas SOSIP (preferiblemente Cys en las posiciones 501 y 605, y Pro en la posición 559) y una mutación de secuencia en el sitio de clivaje de furina, preferiblemente un reemplazo de la secuencia en los residuos 508-511 por RRRRRR (SEQ ID N°: 10). En algunas formas de realización preferidas, la Env del VIH comprende

tanto las mutaciones SOSIP y del sitio de clivaje de furina indicadas, y además también comprende un residuo Pro en la posición 556 o 558, más preferiblemente en ambas posiciones 556 y 558.

En formas de realización preferidas de la invención, la secuencia de aminoácidos de la proteína de la envoltura del VIH es una secuencia consenso, tal como una envoltura consenso del clado C del VIH o una envoltura consenso del clado B del VIH. En una forma de realización particularmente preferida, la secuencia de aminoácidos de la proteína de la envoltura del VIH es una envoltura consenso del clado C del VIH.

Los ejemplos de proteínas de la envoltura del VIH que se pueden usar en la invención incluyen la envoltura consenso del clado C del VIH (SEQ ID N°: 2) y la envoltura consenso del clado B del VIH (SEQ ID N°: 4). Estas secuencias consensos de la envoltura del clado C y del clado B del VIH pueden comprender mutaciones adicionales que, por ejemplo, mejoran la estabilidad y/o formación de trímeros, tales como por ejemplo las denominadas mutaciones SOSIP y/o una mutación de secuencia en el sitio de clivaje de furina como se describió antes, tal como por ejemplo en la secuencia ConC_SOSIP que se muestra en la SEQ ID N°: 3 y la secuencia ConB_SOSIP que se muestra en la SEQ ID N°: 5.

Otros ejemplos no taxativos de secuencias preferidas de la proteína de la envoltura del VIH que se pueden usar en la invención (como 'antecedente' o molécula 'parental', donde luego se introduce una o más de las mutaciones de la invención) incluyen proteínas Env del VIH sintéticas, por ejemplo, que comprenden la secuencia de aminoácidos de la SEQ ID N°: 6 o la SEQ ID N°: 6 con una mutación de Glu por Arg en la posición 166, donde cualquiera opcionalmente comprende además otras mutaciones SOSIP y/o en el sitio de clivaje de furina como se describió antes. Otro ejemplo no taxativo es la SEQ ID N°: 7. Otros ejemplos no taxativos son proteínas de la envoltura del VIH en mosaico, tales como aquellas que tienen la secuencia de aminoácidos de la SEQ ID N°: 8 o 9.

En determinadas formas de realización, la molécula parental es una proteína Env del VIH de tipo salvaje, en donde uno o preferiblemente más aminoácidos fueron reparados de acuerdo con los métodos que se describen en la presente. Dichas moléculas parentales comprenden por lo menos una mutación de reparación en un residuo de aminoácido que está presente en la posición correspondiente a una frecuencia menor que 7.5%, preferiblemente menor que 2%, de las secuencias Env del VIH en una colección de por lo menos 100, preferiblemente por lo menos 500, preferiblemente por lo menos 1000, preferiblemente por lo menos 10000, preferiblemente por lo menos 20000, secuencias Env del VIH de tipo salvaje, en donde la mutación de reparación es una sustitución por un residuo de aminoácido que está presente en la posición correspondiente a una frecuencia de por lo menos 10% en las secuencias Env del VIH en dicha colección. Preferiblemente dicha sustitución es por un residuo de aminoácido que está presente en la posición correspondiente a una frecuencia de por lo menos 15%, por lo menos 20%, por lo menos 25%, en las secuencias Env del VIH en dicha colección. Preferiblemente, dicha sustitución es por el residuo de aminoácido que está presente en la posición correspondiente con mayor frecuencia en dicha colección. En algunas formas de realización preferidas, dichas moléculas parentales comprenden por lo menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, o por lo menos 20 de dichas mutaciones de reparación. Preferiblemente por lo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% o todos los residuos de aminoácidos presentes en las posiciones correspondientes a una frecuencia menor que el 2% en las secuencias Env del VIH en dicha colección son reparadas en la molécula parental en comparación con la proteína Env de tipo salvaje. En determinadas formas de realización, por lo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% o todos los residuos de aminoácidos presentes en las posiciones correspondientes a una frecuencia menor que el 7,5% en las secuencias Env del VIH en dicha colección son reparadas en la molécula parental en comparación con la proteína Env de tipo salvaje. En determinadas formas de realización, la proteína Env del VIH de tipo salvaje es de una cepa del clado A, B o C, preferiblemente de una cepa del clado C. Como resultado de estas mutaciones de reparación, la molécula parental mostrará mayor semejanza con una secuencia consenso Env del VIH que la cepa original de tipo salvaje, de aquí que el residuo de aminoácido reparado a veces se denomina 'aminoácido consenso' o 'residuo consenso' en la presente. El resultado de esta actividad de reparación es una gran mejora de las propiedades de la molécula parental resultante con respecto al plegamiento, la trimerización, expresión y/o estabilidad, y la molécula resultante se denomina 'proteína Env reparada' en la presente. La adición de las mutaciones estabilizantes de la invención (por ejemplo uno o más de (i)-(vii) (Tabla 1) y/o (xvi) (Tabla 1) y/o, opcionalmente, (viii)-(xv) (Tabla 2), en estas moléculas parentales conduce a una mejora aún mayor de uno o más entre porcentaje de trímeros, rendimiento de trímeros, estabilidad, unión a anticuerpos ampliamente neutralizantes, plegamiento, y las moléculas resultantes que derivan de las proteínas Env del VIH de tipo salvaje se conocen como 'proteína Env reparada y estabilizada' en la presente. Será evidente para el especialista que la introducción de las mutaciones estabilizantes en realidad cambiará un poco la secuencia resultante con respecto a una secuencia consenso, de modo que el resultado neto de propiedades muy mejoradas de las moléculas Env del VIH reparadas y estabilizadas se basa en dos conceptos completamente diferentes.

Las mutaciones que dan como resultado los aminoácidos indicados en las posiciones (i)-(vii) de acuerdo con la presente invención también se pueden usar en las proteínas Env del VIH en donde no hay mutaciones SOSIP (por ejemplo en las secuencias consenso Env o en las proteínas Env de formas aisladas de tipo salvaje del VIH) y es probable que también mejoren la trimerización de las mismas, ya que las mutaciones de la invención son independientes de las mutaciones SOSIP, y además se mostró que funcionan en varios esqueletos diferentes de la proteína Env del VIH. Es más, en la

presente se muestra que las mutaciones de acuerdo con la invención pueden funcionar en la ausencia de las mutaciones SOS así como en ausencia de la mutación IP para mejorar las propiedades de trimerización de la Env del VIH.

5 Una proteína de la envoltura del VIH recombinante de acuerdo con las formas de realización de la invención comprende una proteína de la envoltura del VIH que tiene algunos residuos de aminoácidos en las posiciones especificadas en la secuencia de aminoácidos de una proteína de la envoltura del VIH. En particular, se identificaron siete posiciones en la proteína de la envoltura, así como los residuos de aminoácidos particulares que serían deseables en cada una de las posiciones identificadas. Las posiciones identificadas en la secuencia de la proteína de la envoltura incluyen (i) la posición 651, (ii) la posición 655, (iii) la posición 535, (iv) la posición 589, (v) la posición 573, (vi) la posición 204 y (vii) la posición 647, en donde la numeración de las posiciones es acorde con la numeración en la gp160 de la forma aislada HXB2 del VIH-1. Una proteína Env del VIH de acuerdo con la invención comprende los residuos de aminoácidos especificados en por lo menos una de las posiciones indicadas (i)-(vii), preferiblemente en por lo menos dos de las posiciones indicadas (i)-(vii), más preferiblemente en por lo menos tres de las posiciones indicadas (i)-(vii). Los residuos de aminoácidos particulares que serían deseables en cada una de las posiciones identificadas de acuerdo con las formas de realización de la invención se muestran en la Tabla 1. Las posiciones preferidas de estas opciones son (i), (ii), (iii), (iv), (vi) y/o (vii). Las posiciones particularmente preferidas entre estas opciones son (i), (ii), (iii), (iv) y/o (vii). Una opción preferida adicional es (xvi), mencionada con más detalle más adelante en la presente.

Tabla 1: Aminoácidos deseables en las posiciones indicadas en las proteínas Env del VIH recombinantes de acuerdo con las formas de realización de la invención

Nº	Posición ¹	Residuo de aminoácido deseable
(i)	651	Phe, Leu, Met o Trp (preferiblemente Phe)
(ii)	655	Phe, Ile, Met o Trp (preferiblemente Ile)
(iii)	535	Asn o Gln (preferiblemente Asn)
(iv)	589	Val, Ile o Ala (preferiblemente Val o Ile, más preferiblemente Val)
(v)	573	Phe o Trp (preferiblemente Phe)
(vi)	204	Ile
(vii)	647	Phe, Met o Ile (preferiblemente Phe)
(xvi)	658	Val, Ile, Phe, Met, Ala o Leu (preferiblemente Val o Ile, más preferiblemente Val)

¹ De acuerdo con la numeración en la gp160 de la forma aislada HXB2 del VIH-1

20 La secuencia de aminoácidos de la proteína de la envoltura del VIH en la cual se introduce una o más sustituciones de aminoácido deseables (o aminoácidos indicados) en dichas una o más posiciones indicadas se introducen se conoce como "secuencia del esqueleto de la envoltura del VIH" o "secuencia de la envoltura del VIH parental". Por ejemplo, si la posición 651 en la secuencia ConC_SOSIP de la SEQ ID N°: 3 está mutada por Phe, entonces se considera que la secuencia ConC_SOSIP es la secuencia del "esqueleto" o "parental". Se puede usar cualquier proteína de la envoltura del VIH como secuencia de "esqueleto" o "parental" en la cual se puede introducir una nueva mutación estabilizante de

acuerdo con una forma de realización de la invención, ya sea sola o en combinación con otras mutaciones, tales como las denominadas mutaciones SOSIP y/o mutaciones en el sitio de clivaje de furina. Los ejemplos no taxativos de la proteína Env del VIH que se podrían usar como esqueleto incluyen la proteína Env del VIH de una forma aislada natural del VIH, una proteína Env del VIH sintética o una proteína Env del VIH consenso y, en algunos ejemplos no taxativos, incluyen las que comprenden la SEQ ID N°: 2, SEQ ID N°: 3, SEQ ID N°: 4, SEQ ID N°: 5, SEQ ID N°: 6, SEQ ID N°: 7, SEQ ID N°: 8 o la SEQ ID N°: 9.

De acuerdo con las formas de realización de la invención, la proteína de la envoltura del VIH puede comprender el residuo de aminoácido indicado en por lo menos una de las posiciones indicadas que se seleccionan del grupo que consiste en las posiciones 651, 655, 535, 589, 573, 204 y 647, tal como el residuo de aminoácido indicado en uno, dos, tres, cuatro, cinco, seis o siete posiciones. Preferiblemente, la proteína de la envoltura del VIH está sustituida en una, dos o tres de las posiciones indicadas, y más preferiblemente la proteína de la envoltura del VIH está sustituida en por lo menos dos de las posiciones indicadas. Aún más preferiblemente, la proteína Env del VIH está sustituida en tres de las posiciones indicadas, cuatro de las posiciones indicadas, cinco de las posiciones indicadas, seis de las posiciones indicadas o las siete posiciones indicadas. Preferiblemente, la proteína de la envoltura del VIH contiene los residuos de aminoácidos indicados en por lo menos dos de las posiciones indicadas. Más preferiblemente, la proteína de la envoltura del VIH contiene los residuos de aminoácidos indicados en tres de las posiciones indicadas. En otras formas de realización preferidas, la proteína de la envoltura del VIH contiene los residuos de aminoácidos indicados en cuatro, cinco, seis o las siete posiciones indicadas.

Las formas de realización de las proteínas Env del VIH que comprenden los aminoácidos indicados en múltiple posiciones son (posiciones numeradas de acuerdo con la numeración en la gp160 de la forma aislada HXB2 del VIH-1 seguido por el código de aminoácido de una letra para el residuo presente dicha la posición, las posiciones en una forma de realización de la proteína Env del VIH separado por comas [por ejemplo una forma de realización de una proteína Env que comprende Phe en la posición 651 e Ile en la posición 655 se describe como 651F, 655I], en tanto diferentes formas de realización (es decir, diferentes proteínas Env del VIH) se dividen por punto y coma) incluyen las siguientes.

Para las proteínas Env con los aminoácidos indicados en dos posiciones: 651F, 655I; 651F, 655F; 651F, 655M; 651F, 655W; 651F, 535N; 651F, 535Q; 651F, 589V; 651F, 589I; 651F, 589A; 651F, 573F; 651F, 573W; 651F, 204I; 651F, 647F; 651F, 647I; 651F, 647M; 651L, 655I; 651L, 655F; 651L, 655M; 651L, 655W; 651L, 535N; 651L, 535Q; 651L, 589V; 651L, 589I; 651L, 589A; 651L, 573F; 651L, 573W; 651L, 204I; 651L, 647F; 651L, 647I; 651L, 647M; 651M, 655I; 651M, 655F; 651M, 655A; 651M, 655L; 651M, 655W; 651M, 535N; 651M, 535Q; 651M, 589V; 651M, 589I; 651M, 589A; 651M, 573F; 651M, 573W; 651M, 204I; 651M, 647F; 651M, 647I; 651M, 647M; 651W, 655I; 651W, 655F; 651W, 655M; 651W, 655W; 651W, 535N; 651W, 535Q; 651W, 589V; 651W, 589I; 651W, 589A; 651W, 573F; 651W, 573W; 651W, 204I; 651W, 647F; 651W, 647I; 651W, 647M; 655I, 535N; 655I, 535Q; 655I, 589V; 655I, 589I; 655I, 589A; 655I, 573F; 655I, 573W; 655I, 204I; 655I, 647F; 655I, 647I; 655I, 647M; 655F, 535N; 655F, 535Q; 655F, 589V; 655F, 589I; 655F, 589A; 655F, 573F; 655F, 573W; 655F, 204I; 655F, 647F; 655F, 647I; 655F, 647M; 655M, 535N; 655M, 535Q; 655M, 589V; 655M, 589I; 655M, 589A; 655M, 573F; 655M, 573W; 655M, 204I; 655M, 647F; 655M, 647I; 655M, 647M; 655W, 535N; 655W, 535Q; 655W, 589V; 655W, 589I; 655W, 589A; 655W, 573F; 655W, 573W; 655W, 204I; 655W, 647F; 655W, 647I; 655W, 647M; 535N, 589V; 535N, 589I; 535N, 589A; 535N, 573F; 535N, 573W; 535N, 204I; 535N, 647F; 535N, 647I; 535N, 647M; 535Q, 589V; 535Q, 589I; 535Q, 589A; 535Q, 573F; 535Q, 573W; 535Q, 204I; 535Q, 647F; 535Q, 647I; 535Q, 647M; 589V, 573F; 589V, 573W; 589V, 204I; 589V, 647F; 589V, 647I; 589V, 647M; 589I, 573F; 589I, 573W; 589I, 204I; 589I, 647F; 589I, 647I; 589I, 647M; 589A, 573F; 589A, 573W; 589A, 204I; 589A, 647F; 589A, 647I; 589A, 647M; 573F, 204I; 573F, 647F; 573F, 647I; 573F, 647M; 573W, 204I; 573W, 647F; 573W, 647I; 573W, 647M; 204I, 647F; 204I, 647I; 201I, 647M. Cada una de estas formas de realización pueden estar presentes de acuerdo con la invención en cualquier secuencia de la Env del VIH, tal como una forma aislada de tipo salvaje, o una proteína Env del VIH mutante SOSIP o una proteína Env del VIH consenso o una proteína Env del VIH sintética. Cada una de dichas formas de realización se puede combinar con uno de los aminoácidos preferidos de acuerdo con la invención en una tercera posición entre una de las demás posiciones indicadas de (i)-(vii) de acuerdo con la invención. Dichas formas de realización, que tienen residuos de aminoácidos preferidos en tres posiciones de las posiciones indicadas (i)-(vii), se pueden combinar con uno de los aminoácidos preferidos en una cuarta posición de una de las demás posiciones indicadas de (i)-(vii) de acuerdo con la invención. Dichas formas de realización, que tienen residuos de aminoácidos preferidos en cuatro posiciones de las posiciones indicadas (i)-(vii), se pueden combinar con uno de los aminoácidos preferidos en una quinta posición de una de las demás posiciones indicadas de (i)-(vii) de acuerdo con la invención. Dichas formas de realización, que tienen residuos de aminoácidos preferidos en cinco posiciones de las posiciones indicadas (i)-(vii), se pueden combinar con uno de los aminoácidos preferidos en una sexta posición de una de las demás posiciones indicadas de (i)-(vii) de acuerdo con la invención. Dichas formas de realización, que tienen residuos de aminoácidos preferidos en seis posiciones de las posiciones indicadas (i)-(vii) se pueden combinar con uno de los aminoácidos preferidos en una séptima posición de una de las otras posiciones indicadas de (i)-(vii) de acuerdo con la invención, de modo tal que la proteína Env tiene un aminoácido preferido de acuerdo con la invención en las siete posiciones (i)-(vii) de acuerdo con la invención. Cualquiera de estas formas de realización adicionales que tienen aminoácidos preferidos de acuerdo con la invención en tres, cuatro, cinco, seis o siete de las posiciones (v)-(vii) de la

invención, pueden estar presentes en cualquier proteína Env del VIH, tal como de una forma aislada de tipo salvaje, una variante SOSIP, una proteína Env del VIH consenso, una proteína Env del VIH sintética y semejantes.

Las proteínas Env preferidas de acuerdo con la invención con los aminoácidos indicados en dos posiciones: 651F, 655I; 651F, 535N; 651F, 589V; 651F, 589I; 651F, 573F; 651F, 204I; 651F, 647F; 655I, 535N; 655I, 589V; 655I, 589I; 655I, 573F; 655I, 204I; 655I, 647F; 535N, 589V; 535N, 589I; 535N, 573F; 535N, 204I; 535N, 647F; 589V, 573F; 589V, 204I; 589V, 647F; 589I, 573F; 589I, 204I; 589I, 647F; 573F, 204I; 573F, 647F; 204I, 647F. Las proteínas Env particularmente preferidas que tienen aminoácidos preferidos en por lo menos dos posiciones de acuerdo con la invención incluyen: 651F, 655I; 655I, 535N; 655I, 589V; 535N, 589V; 535N, 647F.

Algunas proteínas Env del VIH preferidas que tienen residuos de aminoácidos preferidos en tres posiciones son: 651F, 655I, 535N; 651F, 589V, 535N; 651F, 589I, 535N; 651F, 573F, 535N; 651F, 204I, 535N; 651F, 647F, 535N; 655I, 589V, 535N; 655I, 589I, 535N; 655I, 573F, 535N; 655I, 204I, 535N; 655I, 647F, 535N; 589V, 573F, 535N; 589V, 204I, 535N; 589V, 647F, 535N; 589I, 573F, 535N; 589I, 204I, 535N; 589I, 647F, 535N; 573F, 204I, 535N; 573F, 647F, 535N; 204I, 535N; 651F, 589V; 651F, 204I, 589V; 651F, 647F, 589V; 655I, 573F, 589V; 655I, 647F, 589V; 655I, 204I, 589V; 655I, 647F, 589V; 573F, 204I, 589V; 573F, 647F, 589V; 204I, 647F, 589V; 651F, 655I, 589I; 651F, 573F, 589I; 651F, 204I, 589I; 651F, 647F, 589I; 655I, 573F, 589I; 655I, 204I, 589I; 655I, 647F, 589I; 573F, 204I, 589I; 573F, 647F, 589I; 204I, 647F, 589I; 651F, 655I, 573F; 651F, 204I, 573F; 651F, 647F, 573F; 655I, 204I, 573F; 655I, 647F, 573F; 204I, 647F, 573F; 651F, 655I, 204I; 651F, 647F, 204I; 655I, 647F, 204I; 651F, 655I, 647F; 655I, 651F, 647F; 655I, 651F, 535N; 655I, 589V, 573F; 655I, 589V, 204I. Las proteínas Env particularmente preferidas que tienen aminoácidos preferidos en por lo menos tres posiciones de acuerdo con la invención incluyen: 651F, 655I, 535N; 655I, 589V, 535N; 655I, 573F, 589V; 655I, 204I, 589V; 651F, 655I, 647F.

Algunas proteínas Env del VIH preferidas que tienen residuos de aminoácidos preferidos en cuatro posiciones son: 651F, 655I, 535N, 589V; 651F, 655I, 535N, 573F; 651F, 655I, 589V, 573F; 651F, 535N, 589V, 573F; 655I, 535N, 589V, 573F; 651F, 655I, 535N, 204I; 651F, 655I, 589V, 204I; 651F, 535N, 589V, 204I; 655I, 535N, 589V, 204I; 651F, 655I, 535N, 589V, 204I; 651F, 535N, 589V, 573F, 204I; 655I, 535N, 589V, 573F, 204I; 535N, 589V, 573F, 204I; 651F, 655I, 535N, 647F; 651F, 655I, 589V, 647F; 651F, 535N, 589V, 647F; 655I, 535N, 589V, 647F; 651F, 655I, 573F, 647F; 651F, 535N, 573F, 647F; 655I, 535N, 573F, 647F; 651F, 589V, 573F, 647F; 655I, 589V, 573F, 647F; 535N, 589V, 573F, 647F; 651F, 655I, 204I, 647F; 651F, 535N, 204I, 647F; 655I, 535N, 204I, 647F; 651F, 589V, 204I, 647F; 655I, 589V, 204I, 647F; 535N, 589V, 204I, 647F; 651F, 573F, 204I, 647F; 655I, 573F, 204I, 647F; 535N, 573F, 204I, 647F; y 589V, 573F, 204I, 647F.

Algunos ejemplos de proteínas Env del VIH preferidas que tienen residuos de aminoácidos preferidos en por lo menos cuatro posiciones incluyen: 651F, 655I, 647F, 1535N; 651F, 655I, 573F, 589V. Un ejemplo preferido de una proteína Env del VIH que comprende los residuos de aminoácidos indicados en por lo menos cuatro posiciones comprende 535N, 589V, 651F, 655I. Los ejemplos no taxativos de dichas proteínas Env del VIH se proveen en las SEQ ID N°: 20, 22, 24, 26, 27, 28, 29, 30, 31 y 32. Preferiblemente, dicha proteína Env del VIH es una proteína Env del clado C del VIH o una proteína Env del clado A del VIH, más preferiblemente una proteína Env del clado C del VIH. En determinadas formas de realización, dicha proteína Env del VIH también comprende 588E, es decir comprende por lo menos 535N, 588E, 589V, 651F, 655I. Los ejemplos no taxativos de dicha proteína Env del VIH se proveen en las SEQ ID N°: 20, 24, 26, 27, 28, 29, 30, 31 y 32. En determinadas formas de realización, dicha Env del VIH también comprende 556P, es decir comprende por lo menos 535N, 556P, 589V, 651F, 655I o por lo menos 535N, 556P, 588E, 589V, 651F, 655I. Los ejemplos no taxativos de dicha proteína Env del VIH se proveen en las SEQ ID N°: 22, 24, 26, 27, 29, 30, 31 y 32.

En una forma de realización, una proteína Env del VIH recombinante de acuerdo con la invención comprende la secuencia de aminoácidos de una proteína Env del VIH que tiene los residuos de aminoácidos indicados en por lo menos dos de las posiciones indicadas que se seleccionan del grupo que consiste en:

(i) Phe, Leu, Met o Trp en la posición 651;

(ii) Phe, Ile, Met o Trp en la posición 655;

(iii) Asn o Gln en la posición 535;

(iv) Val, Ile o Ala en la posición 589;

(v) Phe o Trp en la posición 573;

(vi) Ile en la posición 204; y

(vii) Phe, Met o Ile en la posición 647.

Por ejemplo, la proteína Env del VIH recombinante puede tener uno entre Phe, Leu, Met o Trp en la posición 651, y Asn o Gln en la posición 535, opcionalmente, residuos de aminoácidos indicados adicionales en las posiciones indicadas adicionales. Preferiblemente, por lo menos uno de los aminoácidos de (i)-(vii) se introduce en la proteína Env del VIH recombinante mediante sustitución de aminoácidos. Por ejemplo, la proteína Env del VIH recombinante se puede producir a partir de una proteína Env del VIH que no contiene ninguno o solamente uno de los residuos de aminoácidos de (i)-(vii) de modo que todos o uno o más de dichos por lo menos dos residuos de aminoácidos indicados están introducidos en la proteína Env del VIH recombinante mediante sustitución de aminoácidos.

5

En determinadas formas de realización, la proteína Env del VIH recombinante de la invención también comprende (viii) Gln, Glu, Ile, Met, Val, Trp o Phe en la posición 588, siendo preferidos Gln o Glu.

10

La secuencia de aminoácidos de la proteína Env del VIH en la cual se introducen las sustituciones descritas previamente puede ser de cualquier proteína Env del VIH conocida en el arte en vista de la presente divulgación, tal como, por ejemplo una secuencia natural del clado A, clado B, clado C, etc. del VIH; una secuencia en mosaico; una secuencia consenso, por ejemplo, una secuencia consenso del clado B o clado C; una secuencia sintética; o cualquier derivado o fragmento de la misma. En determinadas formas de realización de la invención, la secuencia de aminoácidos de la proteína Env del VIH comprende mutaciones adicionales, tales como, por ejemplo, las denominadas mutaciones SOSIP y/o una mutación en el sitio de clivaje de furina.

15

En una forma de realización particular, la proteína del esqueleto de la Env del VIH es una proteína Env del VIH mutante SOSIP que comprende por lo menos una mutación seleccionada del grupo que consiste en Cys en las posiciones 501 y 605; Pro en la posición 559. En una forma de realización preferida, la proteína Env del VIH mutante SOSIP comprende Cys en las posiciones 501 y 605, y Pro en la posición 559. De acuerdo con esta forma de realización, una proteína Env del VIH recombinante comprende la secuencia de aminoácidos de la proteína Env del VIH mutante SOSIP y una sustitución de aminoácido por el residuo de aminoácido indicado en por lo menos una de las posiciones indicadas que se seleccionan del grupo que consiste en:

20

(i) Phe, Leu, Met o Trp en la posición 651;

25

(ii) Phe, Ile, Met o Trp en la posición 655;

(iii) Asn o Gln en la posición 535;

(iv) Val, Ile o Ala en la posición 589;

(v) Phe o Trp en la posición 573;

(vi) Ile en la posición 204; y

30

(vii) Phe, Met o Ile en la posición 647.

La proteína Env del VIH mutante SOSIP puede comprender además una mutación en el sitio de clivaje de furina, tal como un reemplazo en las posiciones 608-511 por la SEQ ID N°: 10.

En otra forma de realización, una proteína Env del VIH recombinante de acuerdo con la invención comprende la secuencia de aminoácidos de una proteína Env del VIH y una sustitución de aminoácido por el residuo de aminoácido indicado en por lo menos una de las posiciones indicadas que se seleccionan del grupo que consiste en:

35

(i) Phe, Leu, Met o Trp en la posición 651;

(ii) Phe, Ile, Met o Trp en la posición 655;

(iii) Asn o Gln en la posición 535;

(iv) Val, Ile o Ala en la posición 589;

40

(v) Phe o Trp en la posición 573;

(vi) Ile en la posición 204; y

(vii) Phe, Met o Ile en la posición 647,

(vi) Ile en la posición 204; y

(vii) Phe, Met o Ile en la posición 647,

en donde la proteína Env del VIH se selecciona del grupo que consiste en:

(1) una secuencia consenso de la Env del VIH, tal como una secuencia consenso del clado C o clado B, por ejemplo, que comprende la secuencia de aminoácidos de la SEQ ID N°: 2, 3, 4 ó 5;

5 (2) una proteína Env del VIH sintética, por ejemplo que comprende la secuencia de aminoácidos de (a) la SEQ ID N°:6; (b) la SEQ ID N°: 6 con una mutación de Glu por Arg en la posición 166; (c) la SEQ ID N°: 7; (d) la SEQ ID N°: 8 ó 9, donde (a), (b) o (d) opcionalmente tienen otras mutaciones SOSIP y/o del sitio de clivaje de furina como se describió precedentemente.

10 Preferiblemente, la proteína Env del VIH recombinante comprende la secuencia de aminoácidos de una proteína Env del VIH y una sustitución de aminoácido por el residuo de aminoácido indicado en por lo menos dos de las posiciones indicadas que se seleccionan del grupo que consiste en (i)-(vii) anteriores, tal como dos posiciones o tres posiciones. Sin embargo, la proteína Env del VIH recombinante puede comprender una sustitución de aminoácido por el residuo de aminoácido indicado en una o más de las posiciones indicadas, tal como uno, dos, tres, cuatro, cinco, seis o siete de las posiciones indicadas.

15 En una forma de realización particular, la proteína del esqueleto de la Env del VIH es una Env del VIH consenso del clado C que comprende una secuencia de aminoácidos que es por lo menos 95%, 96%, 97%, 98%, 99% o 100% idéntica a la secuencia de aminoácidos de la SEQ ID N°: 2. Preferiblemente, la secuencia consenso del clado C del VIH de la SEQ ID N°: 2 además comprende las denominadas mutaciones SOSIP, es decir, Cys en las posiciones 501 y 605, y Pro en la posición 559, y más preferiblemente comprende además las denominadas mutaciones SOSIP y una mutación en el sitio de clivaje de furina, tal como por ejemplo un reemplazo en las posiciones 508-511 por la SEQ ID N°: 10. En una forma de realización particularmente preferida, la proteína del esqueleto de la Env del VIH comprende la secuencia que se muestra en la SEQ ID N°: 3, o una secuencia que es por lo menos 95% idéntica a la misma, preferiblemente en donde se reemplazan los aminoácidos en las posiciones 501, 559, 605 y 508-511 por la SEQ ID N°: 10, no están mutadas en comparación con la SEQ ID N°: 3.

25 En otra forma de realización particular, la proteína del esqueleto de la Env del VIH es una Env del VIH consenso del clado B que comprende una secuencia de aminoácidos que es por lo menos 95%, 96%, 97%, 98%, 99% o 100% idéntica a la secuencia de aminoácidos de la SEQ ID N°: 4. Preferiblemente, la secuencia consenso del clado B del VIH de la SEQ ID N°: 4 además comprende las denominadas mutaciones SOSIP, es decir, Cys en las posiciones 501 y 605, y Pro en la posición 559, y más preferiblemente comprende además las denominadas mutaciones SOSIP y una mutación en el sitio de clivaje de furina, tal como por ejemplo un reemplazo en las posiciones 508-511 por la SEQ ID N°: 10. En una forma de realización particularmente preferida, la proteína del esqueleto de la Env del VIH comprende la secuencia que se muestra en la SEQ ID N°: 5, o una secuencia que es por lo menos 95% idéntica a la misma, preferiblemente en donde se reemplazan los aminoácidos en las posiciones 501, 559, 605 y 508-511 por la SEQ ID N°: 10, no están mutadas en comparación con la SEQ ID N°: 5.

35 En aún otra forma de realización particular, la proteína del esqueleto de la Env del VIH es una proteína Env del VIH sintética, por ejemplo, que comprende la secuencia de aminoácidos de (a) la SEQ ID N°: 6; (b) la SEQ ID N°: 6 con una mutación de Glu por Arg en la posición 166; (c) la SEQ ID N°: 7; o (d) la SEQ ID N°: 8 o 9, donde (a) (b) o (d) opcionalmente tienen otras mutaciones SOSIP (501C, 605C, 559P) y/o del sitio de clivaje de furina (508-511RRRRRR) como se describió precedentemente.

40 En aún otras formas de realización particulares, la proteína del esqueleto de la Env del VIH es una proteína Env del VIH de un virus del clado A o clado C del VIH de tipo salvaje, que opcionalmente comprende mutaciones para reparar la secuencia de acuerdo con los métodos que se describen en la presente.

45 Los ejemplos de combinaciones de dos posiciones en la proteína Env del VIH que se pueden sustituir simultáneamente incluyen los residuos 535,589; 535,647; y 589,655; tales como, por ejemplo, en los mutantes dobles I535N, D589V; I535N, E647F; y D589V, K655I. Otros mutantes dobles incluyen K655I, I535N; N651F, K655I; y K655I, I573F. Un ejemplo de una combinación de tres posiciones en la proteína Env del VIH que se pueden sustituir simultáneamente incluye 535,589,655, tal como por ejemplo en el mutante triple I535N, D589V, K655I. Otros mutantes triples incluyen K655I, D589V, I573F; y K655I, N651F, I535N.

50 En determinadas formas de realización de la invención, una proteína Env del VIH recombinante de acuerdo con la invención puede comprender además un residuo de aminoácido indicado (por ejemplo mediante sustitución) e una o más posiciones adicionales indicadas seleccionadas del grupo que consiste en las posiciones (viii) 588, (ix) 64 o 66, (x) 316, (xi) 201/433, (xii) 556 o 558 o 556 y 558, (xiii) 548-568, (xiv) 568, 569 y 636, o (xv) 302, 519 o 520, como se muestra más adelante en la Tabla 2. Los inventores de la presente encontraron que algunas de estas sustituciones de aminoácidos (por ejemplo (viii)) combinan muy bien con las (combinaciones de) mutaciones (i)-(vii) de acuerdo con la invención como se describió precedentemente. Otras de estas sustituciones de aminoácidos se han informado con anterioridad en la

literatura. Por ejemplo, De Taeye et al. *Cell* (2015) 163(7), 1702-15) describieron una proteína de la envoltura del VIH que tiene una mutación doble E64K y T316W, y una proteína Env del VIH que tiene una mutación 66R; y WO 2016/037154 y Kwon et al. (*Nat. Struct. Mol. Biol.* (2015) 22(7) 522-31) describieron una proteína de la envoltura del VIH que tiene una sustitución disulfuro I204C, A433C; y Guenaga et al. (*Immunity* (2017) 46, 792-803) describieron una proteína de la envoltura del VIH que tiene una sustitución triple L568G, T569G o N636G, y N302Y, F519R, L520R. Sin embargo, según el mejor conocimiento de los inventores, estas mutaciones descritas previamente no se describieron en combinación con cualquiera de las nuevas sustituciones que se describen en la presente, por ejemplo, las sustituciones enumeradas en la Tabla 1. Estas mutaciones de aminoácidos en combinación con las sustituciones de aminoácidos de la invención pueden aumentar aún más el rendimiento de trímeros y/o el porcentaje de formación de trímeros. Estas sustituciones de aminoácidos se pueden introducir en cualquiera de las proteínas Env del VIH recombinantes que se describe en la presente además de una sustitución por el residuo de aminoácido indicado en una o más de las posiciones indicadas como se describe en la Tabla 1.

Tabla 2: Posiciones de sustitución de aminoácido y residuos de sustitución adicionales

<u>Nº</u>	<u>Posición¹</u>	<u>Residuo de aminoácido indicado</u>
(viii)	588	Gln, Glu, Ile, Met, Val, Trp o Phe (preferiblemente Gln o Glu)
(ix)	64 o 66	Lys en la posición 64; o Arg en la posición 66
(x)	316	Trp
(xi)	201 y 433	Cys en ambas posiciones
(xii)	556 o 558 o 556 y 558	Pro en cualquiera o ambas posiciones
(xiii)	548-568 (bucle HR1)	Reemplazo por un bucle más corto y menos flexible de 7-10 aminoácidos, preferiblemente un bucle de 8 aminoácidos, por ejemplo que tiene una secuencia seleccionada entre cualquiera de las (SEQ ID N°: 12-17)
(xiv)	568, 569, 636	Gly en cualquiera de estas posiciones, o Gly en ambas posiciones 568 y 636, o Gly en ambas posiciones 569 y 636
(xv)	302, 519, 520	Tyr en la posición 302, o Arg en la posición 519, o Arg en la posición 520; o Tyr en la posición 302 y Arg en la posición 519; o Tyr en la posición 302 y Arg en la posición 520; o Tyr en la posición 302 y Arg en ambas posiciones 519 y 520

¹ De acuerdo con la numeración en la gp160 de la forma aislada HXB2 del VIH-1

Las sustituciones identificadas en las posiciones indicadas de la presente invención [(i)-(vii), véase, por ejemplo, la Tabla 1] raramente o nunca están presentes en las secuencias naturales, no se encuentran combinadas en las secuencias de proteínas Env del VIH informadas previamente y no se había sugerido con anterioridad que dieran como resultado una trimerización mejorada de la proteína Env del VIH, un rendimiento de trímeros mejorado y/o una mayor estabilidad del trímero. Las mutaciones (ix)-(xi) en la Tabla 2 (que fueron informados previamente por otros) se encuentran todas en la región gp120, con la cual se une el anticuerpo específico del trímero, PGT145. Estas mutaciones mantienen al trímero cerrado por el ápice (que se encuentra en la parte superior de la molécula). Las sustituciones (xii) y (xiii) están todas en la HR1 de gp41. Excepto por la posición 204, las mutaciones de la presente invención que se muestran en la Tabla 1 están todas en la región gp41 (en la parte inferior de la molécula), pero fuera de la región HR1. Claramente, las mutaciones descritas previamente no sugerían introducir las mutaciones de la presente invención, mucho menos los efectos sorprendentes de las mismas sobre la formación de trímeros con un ápice cerrado medida mediante unión a PGT145. Además de las mutaciones puntuales (viii)-(xii) de la Tabla 2, también es posible reemplazar el bucle HR1 de la proteína Env (residuos de aminoácidos 548-568 en una secuencia de tipo salvaje, con la numeración de acuerdo con gp160 de la forma aislada HXB2) por un bucle más corto y menos flexible de 7-10 aminoácidos, preferiblemente un bucle de 8

aminoácidos, por ejemplo, que tenga una secuencia seleccionada entre cualquiera de las (SEQ ID N°: 12-17), véase por ejemplo Kong et al (Nat Commun. junio de 2016, 28:7:12040. doi: 10.1038/ncomms12040) donde se describe el reemplazo del bucle HR1 por estos bucles más cortos. Dicha variante de la Env, que además tiene los residuos de aminoácidos indicados en por lo menos una, y preferiblemente en por lo menos dos, de las posiciones indicadas (i)-(vii) de acuerdo con la invención, también es una forma de realización de la invención. Las mutaciones enumeradas en (viii)-(xiii) se pueden agregar, en determinadas formas de realización de la invención, a las proteínas Env del VIH de la invención, es decir que tienen uno o más de los aminoácidos indicados en las posiciones (i)-(vii). Además, se pueden efectuar combinaciones dentro de los grupos (viii)-(xiii), siendo un ejemplo no taxativo una combinación de las mutaciones (además de por lo menos una mutación de (i)-(vii)) en (viii) y (xii) (por ejemplo I535N, A558P, K588E). Algunos ejemplos no taxativos de mutantes dobles que se obtuvieron con un antecedente de la Env del VIH con mutaciones SOSIP y combinación de mutaciones en por lo menos una de las posiciones (i)-(vii) y en por lo menos una de las posiciones (viii)-(xiii) incluyen: 535,588; 588,589; 655,588; 558,535; y 655,556; tal como por ejemplo I535N, K588E; 588Q, D589V; K655I, K588E; A558P, I535N; y K655I, L556P. Algunos ejemplos no taxativos de dichos mutantes triples incluyen 558,535,588; 558,535,589; 558,535,655; y 558,535,651, tal como por ejemplo A558P, I535N, K588E; A558P, I535, D589V; A558P, I535N, K655I; y A558P, I535N, N651F.

Otros ejemplos no taxativos de las combinaciones de acuerdo con la invención incluyen 655I, 573F, 589V, 588E; 651F, 655I, 573F, 589V, 588E; 651F, 655I, 573F, 589V, 588E, 535N; 651F, 655I, 573F, 589V, 588E, 535N, 204I; 651F, 655I, 556P; 651F, 535N, 556P; 651F, 589V, 556P; 651F, 589I, 556P; 651F, 573F, 556P; 651F, 204I, 556P; 651F, 647F, 556P; 655I, 535N, 556P; 655I, 589V, 556P; 655I, 589I, 556P; 655I, 573F, 556P; 655I, 204I, 556P; 655I, 647F, 556P; 535N, 589V, 556P; 535N, 589I, 556P; 535N, 573F, 556P; 535N, 204I, 556P; 535N, 647F, 556P; 589V, 573F, 556P; 589V, 204I, 556P; 589V, 647F, 556P; 589I, 573F, 556P; 589I, 204I, 556P; 589I, 647F, 556P; 573F, 204I, 556P; 573F, 647F, 556P; 651F, 655I, 558P; 651F, 535N, 558P; 651F, 589V, 558P; 651F, 589I, 558P; 651F, 573F, 558P; 651F, 204I, 558P; 651F, 647F, 558P; 655I, 535N, 558P; 655I, 589V, 558P; 655I, 589I, 558P; 655I, 573F, 558P; 655I, 204I, 558P; 655I, 647F, 558P; 535N, 589V, 558P; 535N, 589I, 558P; 535N, 573F, 558P; 535N, 204I, 558P; 535N, 647F, 558P; 589V, 573F, 558P; 589V, 204I, 558P; 589V, 647F, 558P; 589I, 573F, 558P; 589I, 204I, 558P; 589I, 647F, 558P; 573F, 204I, 558P; 573F, 647F, 558P; 655I, 589V, 535N, 556P; 651F, 655I, 535N, 573F; 556P, 651F; 556P, 651F, 655I, 535N; 655I, 535N, 573F, 589V; 556P, 651F, 655I, 535N, 573F, 589V, 204I; 556P, 651F, 655I, 535N, 573F, 589V, 204I, 588Q; 556P, 651F, 655I, 535N, 573F, 589V, 204I, 588Q, 647F; 556P, 651F, 535N, 573F; 556P, 651F, 535N, 573F, 589V; 556P, 651F, 535N, 573F, 589V, 204I; 556P, 651F, 535N, 573F, 589V, 204I, 588Q; 556P, 651F, 535N, 573F, 589V, 204I, 588Q, 647F. Nuevamente, cualquiera de estas formas de realización pueden ser en cualquier proteína Env del VIH, por ejemplo una forma aislada de tipo salvaje, una Env consenso, una proteína Env sintética, una proteína Env mutante SOSIP, una forma aislada de tipo salvaje que contiene mutaciones de reparación de acuerdo con el concepto que se describe en la presente, etc. Algunas combinaciones preferidas de acuerdo con la invención incluyen 655I,589V,573F,651F,588E,535N,204I; 556P, 655I, 535N, 573F, 589V, 204I, 588Q; 204I,535N,556P,588E,589V,651F,655I; 535N,556P,589V,651F,655I; y 535N,556P,588E,589V,651F,655I.

En algunas formas de realización preferidas, la proteína Env del VIH comprende una secuencia que es por lo menos 95% idéntica, preferiblemente por lo menos 96%, 97%, 98%, 99% idéntica, preferiblemente 100% idéntica, a cualquiera de las SEQ ID N°: 20, 22, 24, 26, 27, 28, 29, 30, 31 y 32. Para determinar el % de identidad, preferiblemente no se tienen en cuenta las posiciones (i)-(xv) de las Tablas 1 y 2, y preferiblemente tampoco las posiciones 501, 559 y 605. Preferiblemente, los residuos de aminoácidos en aquellas posiciones son los que se encuentran en las secuencias de la SEQ ID N°: 20, 22, 24, 26, 27, 28, 29, 30, 31 o 32, respectivamente.

En determinadas formas de realización, la proteína Env del VIH de la invención comprende además: (xvi) un residuo de aminoácido seleccionado entre Val, Ile, Phe, Met, Ala o Leu en la posición 658. Preferiblemente, el aminoácido en la posición 658 es Val o Ile, más preferiblemente Val. Se encontró que esto aumentaba fuertemente el porcentaje de trímeros y el rendimiento de trímeros de la proteína Env, ya sea sola o en combinación con mutaciones seleccionadas entre (i)-(vii) de la Tabla 1 y/o (viii)-(xv) de la Tabla 2 descritas en la presente.

De acuerdo con las formas de realización de la invención, una proteína Env del VIH recombinante comprende por lo menos uno entre (a) un porcentaje mejorado de formación de trímeros y (b) un rendimiento de trímeros mejorado en comparación con una proteína Env del VIH que no contiene los residuos de aminoácidos indicados en una o más de las posiciones 651, 655, 535, 589, 573, 204, and 647 como se muestra en la Tabla 1.

Según se usa en la presente un "porcentaje mejorado de formación de trímeros" significa que se forma un mayor porcentaje de trímeros cuando la secuencia del esqueleto de la proteína de la envoltura del VIH contiene una o más de las sustituciones de aminoácidos de la invención en comparación con el porcentaje de trímeros que se forma cuando la secuencia del esqueleto de la secuencia de la envoltura del VIH no contiene dichas sustituciones de aminoácidos. Según se usa en la presente un "rendimiento de trímeros mejorado" significa que se obtiene una mayor cantidad total de la forma

trimérica de la proteína de la envoltura cuando la secuencia del esqueleto de la proteína de la envoltura del VIH contiene una o más de las sustituciones de aminoácidos de la invención en comparación con la cantidad total de la forma trimérica de la proteína de la envoltura que se obtiene cuando la secuencia del esqueleto de la secuencia de la envoltura del VIH no contiene dichas sustituciones de aminoácidos.

5 Las formación de trímeros se puede medir mediante un ensayo de unión a anticuerpos usando anticuerpos que se unen específicamente a la forma trimérica de la proteína Env del VIH. Los ejemplos de anticuerpos específicos del trímero que se pueden usar para detectar la forma trimérica incluyen, pero en un sentido no taxativo, los anticuerpos monoclonales (mAbs) PGT145, PGDM1400, PG16 y PGT151. Preferiblemente, el anticuerpo específico del trímero es el mAb PGT145. Se puede emplear cualquier ensayo de unión a anticuerpos conocido en el arte en vista de la presente divulgación para
10 medir el porcentaje de formación de trímeros de una proteína Env del VIH recombinante de la invención, tal como ELISA, AlphaLISA, etc.

En una forma de realización particular, la formación de trímeros se mide mediante un AlphaLISA. El AlphaLISA es un ensayo de proximidad basado en esferas en el cual se transfieren moléculas de oxígeno singulete, generadas por irradiación con gran energía de esferas donantes, a esferasceptoras que se encuentran dentro de una distancia de
15 aproximadamente 200 nm con respecto a las esferas donantes. La transferencia de moléculas de oxígeno singulete a las esferasceptoras inicia una serie de cascadas de reacciones químicas que resultan en una señal quimioluminiscente que luego puede ser detectada (Eglen et al. *Curr. Chem. Genomics*, 2008, 25(1): 2-10). Por ejemplo, después se pueden incubar las proteínas de la envoltura del VIH recombinantes marcadas con una marca Flag-His con un mAb específico de trímeros, esferas donantes conjugadas con el anticuerpo que se une al mAb específico de trímeros, esferas donantes
20 conjugadas con níquel, esferasceptoras conjugadas con un anticuerpo anti-His y esferasceptoras conjugadas con un anticuerpo anti-Flag. La cantidad de trímeros formada se puede determinar midiendo la señal quimioluminiscente generada a partir de un par de esferas donantes conjugadas con el anticuerpo que se une al mAb específico de trímeros y las esferasceptoras conjugadas al anticuerpo anti-His. La cantidad total de proteína de la envoltura del VIH expresada se puede determinar midiendo la señal quimioluminiscente generada por el par de esferas donantes conjugadas con níquel y esferas
25ceptoras conjugadas con anti-Flag. Por ejemplo, la cantidad de trímeros y las proteínas de la envoltura totales expresadas se pueden medir mediante un ensayo AlphaLISA como se describe con detalle en el Ejemplo 3. El porcentaje de formación de trímeros se puede calcular dividiendo la cantidad de trímeros formados por la cantidad total de proteínas de la envoltura expresadas.

La cantidad de trímeros formados y la cantidad total de proteínas de la envoltura expresadas también se pueden determinar usando técnicas cromatográficas que pueden separar la forma trimérica de las demás formas de la proteína de la envoltura
30 del VIH, por ejemplo, la forma monomérica. Los ejemplos de las técnicas que se pueden utilizar incluyen, pero en un sentido no taxativo, cromatografía de exclusión por tamaño con dispersión de luz multiangular (SEC-MALS). De acuerdo con determinadas formas de realización, el porcentaje de formación de trímeros se determina usando SEC-MALS. De acuerdo con determinadas formas de realización, el rendimiento de trímeros se determina usando SEC-MALS.

35 Ácidos nucleicos, vectores y células

En otro aspecto general, la invención provee una molécula de ácido nucleico que codifica una proteína Env del VIH recombinante de acuerdo con la invención, y un vector que comprende la molécula de ácido nucleico. Las moléculas de ácido nucleico de la invención pueden estar en forma de ARN o en forma de ADN obtenido por clonación o producido
40 sintéticamente. El ADN puede ser bicatenario o monocatenario. El ADN puede comprender, por ejemplo, ADNc, ADN genómico o combinaciones de los mismos. Las moléculas de ácidos nucleicos y los vectores se pueden usar para la producción de proteínas recombinantes, la expresión de la proteína en una célula huésped o la producción de partículas virales.

De acuerdo con las formas de realización de la invención, el ácido nucleico que codifica la proteína de la envoltura del VIH recombinante está ligado operativamente a un promotor, lo que significa que el ácido nucleico se encuentra bajo el control
45 de un promotor. El promotor puede ser un promotor homólogo (es decir, obtenido a partir de la misma fuente genética que el vector) o un promotor heterólogo (es decir, obtenido a partir de un vector o una fuente genética diferentes). Los ejemplos de promotores adecuados incluyen el promotor inmediato temprano del citomegalovirus humano (hCMV IE, o abreviado como "CMV") y el promotor del virus del sarcoma de Rous (RSV). Preferiblemente, el promotor está ubicado en dirección 5' respecto al ácido nucleico dentro de un casete de expresión.

De acuerdo con las formas de realización de la invención, un vector puede ser un vector de expresión. Los vectores de expresión incluyen, pero en un sentido no taxativo, vectores para la expresión de las proteínas recombinantes y vectores para suministrar el ácido nucleico en un sujeto para la expresión en un tejido del sujeto, tal como un vector viral. Los ejemplos de vectores virales adecuados para su uso con la invención incluyen, pero en un sentido no taxativo, vectores adenovirales, vectores de virus adeno-asociados, vectores de virus pox, vectores de vaccinia Ankara modificados (MVA),
55 vectores de virus entéricos, vectores del virus de la encefalitis equina venezolana, vectores del virus de Semliki Forest,

vectores del virus en mosaico del tabaco, vectores lentivirales, etc. El vector también puede ser un vector no viral. Los ejemplos de vectores no virales incluyen, sin carácter limitante, plásmidos, cromosomas artificiales bacterianos, cromosomas artificiales de levaduras, bacteriofagos, etc.

5 En determinadas formas de realización de la invención, el vector es un vector de adenovirus, por ejemplo, un vector de adenovirus recombinante. Un vector de adenovirus recombinante puede derivar, por ejemplo, de un adenovirus humano (HAdV, o AdHu) o de un adenovirus de simio tal como un adenovirus de chimpancé o gorila (ChAd, AdCh o SAdV) o un adenovirus rhesus (rhAd). Preferiblemente, un vector de adenovirus es un vector de adenovirus humano recombinante, por ejemplo, un adenovirus humano recombinante del serotipo 26 o cualquier adenovirus humano recombinante de los serotipos 5, 4, 35, 7, 48, etc. En otras formas de realización, un vector de adenovirus es un vector rhAd, por ejemplo
10 rhAd51, rhAd52 o rhAd53.

La preparación de vectores adenovirales recombinantes es muy conocida en la técnica. Por ejemplo, la preparación de vectores de adenovirus 26 recombinantes se describe, por ejemplo, en WO 2007/104792 y en Abbink *et al.*, (2007) *Virology* 81(9): 4654-63. Los ejemplos de secuencias genómicas del adenovirus 26 se pueden consultar en GenBank, Acceso EF 153474 y en la SEQ ID N°: 1 de WO 2007/104792. Los ejemplos de secuencias genómicas de rhAd51, rhAd52 y rhAd53 se proveen en US 2015/0291935.
15

De acuerdo con las formas de realización de la invención, los vectores que se describen en la presente pueden expresar y/o codificar cualquiera de las proteínas Env del VIH recombinantes que se describen en la presente. Habida cuenta de la degeneración del código genético, el experto es completamente consciente de que se pueden diseñar varias secuencias de ácido nucleico que codifiquen la misma proteína, de acuerdo con métodos totalmente corrientes en la técnica.
20 Opcionalmente, se pueden optimizar los codones del ácido nucleico que codifica la proteína Env del VIH recombinante de la invención para asegurar una expresión apropiada en la célula huésped (*por ejemplo*, células bacterianas o de mamífero). La optimización de codones es una tecnología ampliamente aplicada en la técnica.

La invención también provee células, preferiblemente células aisladas, que comprenden cualquiera de las moléculas de ácidos nucleicos y los vectores que se describen en la presente. Las células se pueden usar, por ejemplo, para la producción de proteínas recombinantes o para la producción de partículas virales.
25

Por consiguiente, las formas de realización de la invención también se relacionan con un método para elaborar una proteína Env del VIH recombinante. El método comprende transfectar una célula huésped con un vector de expresión que comprende un ácido nucleico que codifica una proteína Env del VIH recombinante de acuerdo con una forma de realización de la invención ligado operativamente a un promotor, cultivar la célula transfectada en condiciones adecuadas para la expresión de la proteína Env del VIH recombinante y, opcionalmente, purificar o aislar la proteína Env del VIH recombinante expresada en la célula. La proteína Env del VIH recombinante se puede aislar o recolectar de la célula mediante cualquier método conocido en el arte, que incluye cromatografía de afinidad, cromatografía de exclusión por tamaño, etc. Las técnicas usadas para expresar las proteínas recombinantes son bien conocidas por un especialista en el arte en vista de la presente divulgación. La proteína Env del VIH recombinante expresada también se puede estudiar sin purificar o aislar la proteína expresada, por ejemplo, mediante análisis del sobrenadante de células transfectadas con un vector de expresión que codifica la proteína Env del VIH recombinante y cultivadas en condiciones adecuadas para la expresión de la proteína Env del VIH.
30
35

En una forma de realización preferida, la proteína Env del VIH recombinante expresada se purifica en condiciones que permitan la asociación de la proteína para poder formar el complejo trimérico estabilizado. Por ejemplo, las células de mamífero transfectadas con un vector de expresión que codifica la proteína Env del VIH recombinante ligada operativamente a un promotor (por ejemplo el promotor CMV) se pueden cultivar a 33-39 °C, por ejemplo, a 37°C, y 2-12% de CO₂, por ejemplo 8% de CO₂. La expresión también se puede realizar en sistemas de expresión alternativos, tales como células de insecto o células de levadura, todos los cuales son convencionales en el arte. La proteína Env del VIH expresada se puede aislar luego del cultivo celular, por ejemplo, mediante cromatografía de afinidad con lectina, que permite unir glicoproteínas. La proteína Env del VIH unida a la columna se puede eluir con manopiranosido. La proteína Env del VIH eluida de la columna se puede someter a pasos de purificación adicionales, tal como cromatografía de exclusión por tamaño, según necesidad, para eliminar todo contaminante residual, por ejemplo, contaminantes celulares, y también agregados de la Env, monómeros de gp140 y monómeros de gp120. También se podrían utilizar métodos de purificación alternativos, y los ejemplos no taxativos de los mismos incluyen cromatografía de afinidad de anticuerpos, selección negativa con no bNAbs, purificación con anti-marcas u otros métodos cromatográficos tal como la cromatografía de intercambio iónico, etc, así como otros métodos conocidos en el arte, a fin de aislar la proteína Env del VIH expresada.
40
45
50

Las moléculas de ácidos nucleicos y los vectores de expresión que codifican las proteínas Env del VIH recombinantes de la invención se puede obtener mediante cualquier método conocido en el arte en vista de la presente divulgación. Por ejemplo, el ácido nucleico que codifica la proteína Env del VIH recombinante se puede preparar mediante introducción de por lo menos una de las sustituciones de aminoácidos en las posiciones indicadas en la secuencia del esqueleto de la
55

5 envoltura del VIH usando tecnología de ingeniería genética y técnicas de biología molecular, por ejemplo, mutagénesis dirigida al sitio, reacción en cadena de la polimerasa (PCR), etc., que son bien conocidos por los especialistas en el arte. Luego se puede introducir o "clonar" la molécula de ácido nucleico en un vector de expresión utilizando también técnicas de biología molecular estándar. La proteína de la envoltura del VIH recombinante se puede expresar posteriormente a partir del vector de expresión en una célula huésped y seguidamente se puede purificar la proteína expresada del cultivo celular mediante cualquier método conocido en el arte en vista de la presente divulgación.

Complejo trimérico

10 En otro aspecto general, la invención se relaciona con un complejo trimérico que comprende un oligómero no covalente de tres de las proteínas Env del VIH recombinantes acordes con la invención. El complejo trimérico puede comprender cualquiera de las proteínas Env del VIH recombinantes que se describen en la presente. Preferiblemente, el complejo trimérico comprende tres monómeros idénticos (o heterodímeros idénticos si se cliva gp140) de las proteínas Env del VIH recombinantes de acuerdo con la invención. El complejo trimérico se puede separar de las otras formas de la proteína de la envoltura del VIH, tal como la forma monomérica, o el complejo trimérico puede estar presente junto con otras formas de la proteína de la envoltura del VIH, tal como la forma monomérica.

Composiciones y métodos

15 En otro aspecto general, la invención se relaciona con una composición que comprende una proteína Env del VIH recombinante, un complejo trimérico, un ácido nucleico aislado, un vector o una célula huésped, y un vehículo farmacéuticamente aceptable. La composición puede comprender cualquiera de las proteínas Env del VIH recombinantes, complejos triméricos, moléculas de ácido nucleico aisladas, vectores o células huésped que se describen en la presente.

20 Un vehículo puede incluir uno o más excipientes farmacéuticamente aceptables tales como aglutinantes, desintegrantes, agentes para el aumento de volumen, agentes de suspensión, agentes emulsionantes, agentes humectantes, lubricantes, saborizantes, edulcorantes, conservantes, tintes, solubilizantes y recubrimientos. La naturaleza precisa del vehículo o de otro material dependerá de la ruta de administración, por ejemplo, las rutas intramuscular, intradérmica, subcutánea, oral, intravenosa, cutánea, intramucosa (por ejemplo, del intestino), intranasal o intraperitoneal. Para los preparados líquidos inyectables, por ejemplo, suspensiones y soluciones, los vehículos y aditivos adecuados incluyen agua, glicoles, aceites, alcoholes, conservantes, agentes colorantes y similares. Para los preparados sólidos orales, por ejemplo, polvos, cápsulas, comprimidos oblongos, cápsulas de gelatina y comprimidos, los vehículos y aditivos adecuados incluyen almidones, azúcares, diluyentes, agentes para la granulación, lubricantes, aglutinantes, agentes desintegrantes y similares. Para las mezclas inhalatorias/nebulizados nasales, la suspensión/solución acuosa puede comprender como vehículos y aditivos adecuados agua, glicoles, aceites, emolientes, estabilizantes, agentes humectantes, conservantes, compuestos aromáticos, sabores y similares.

35 Las composiciones de la invención se pueden formular en cualquier materia adecuada para su administración a un sujeto para facilitar la administración y mejorar la eficacia, incluidas, sin carácter limitante, la administración oral (enteral) e inyecciones parenterales. Las inyecciones parenterales incluyen una inyección intravenosa o infusión, una inyección subcutánea, una inyección intradérmica y una inyección intramuscular. Las composiciones de la invención también se pueden formular para otras vías de administración incluidas la administración transmucosal, ocular, rectal, con implantes que actúan a largo plazo, sublingual, debajo de la lengua, desde la mucosa oral para evitar la circulación portal, por inhalación o intranasal.

40 Las formas de realización de la invención también se relacionan con métodos para elaborar la composición. De acuerdo con las formas de realización de la invención, un método para producir una composición comprende mezclar una proteína Env del VIH recombinante, un complejo trimérico, un ácido nucleico aislado, un vector o una célula huésped de la invención con uno o más vehículos farmacéuticamente aceptables. El experto en la materia está familiarizado con las técnicas habituales utilizadas para preparar tales composiciones.

45 Se han utilizado con anterioridad antígenos del VIH (por ejemplo, proteínas o fragmentos de las mismas derivadas de los productos genéticos *gag*, *pol* y/o *env* del VIH) y vectores, tales como vectores virales, que expresan los antígenos del VIH en composiciones inmunogénicas y vacunas para vacunar a un sujeto contra una infección por el VIH o para generar una respuesta inmunológica contra una infección por el VIH en un sujeto. Según se usa en la presente, un "sujeto" se refiere a cualquier animal, preferiblemente un mamífero, más preferiblemente un humano, a quien se le administrará o a quien se le ha administrado una composición inmunogénica de acuerdo con las formas de realización de la invención. El término "mamífero", tal como se utiliza en la presente, engloba cualquier mamífero. Los ejemplos de mamíferos incluyen, pero en un sentido no taxativo, ratones, ratas, conejos, cobayos, monos, humanos, etc., preferiblemente un humano. Las proteínas Env del VIH recombinantes de la invención también se pueden usar como antígenos para inducir una respuesta inmunológica contra el virus de inmunodeficiencia humana (VIH) en un sujeto que lo necesita. La respuesta inmunológica puede ser contra uno o más clados del VIH, tales como el clado A, clado B, clado C, etc. Las composiciones pueden

comprender un vector a partir del cual se expresa la proteína Env del VIH recombinante o la composición puede comprender una proteína Env del VIH recombinante aislada de acuerdo con una forma de realización de la invención.

Por ejemplo, las composiciones que comprenden una proteína del VIH recombinante, o un complejo trimérico de la misma, se pueden administrar a un sujeto que lo necesita para inducir una respuesta inmunológica contra una infección por VIH en el sujeto. La composición que comprende un vector, tal como un vector de adenovirus, que codifica una proteína Env del VIH recombinante de la invención, en donde la proteína Env del VIH recombinante es expresada por el vector, también se puede administrar a un sujeto que lo necesita para inducir una respuesta inmunológica contra una infección por VIH en el sujeto. Los métodos que se describen en la presente también incluyen administrar una composición de la invención en combinación con uno o más antígenos del VIH adicionales (por ejemplo, proteínas o fragmentos de los mismos derivados de los productos genéticos *gag*, *pol*, y/o *env* del VIH) que preferiblemente se expresan a partir de uno o más vectores, tales como vectores de adenovirus o vectores MVA, que incluyen métodos de imprimación y refuerzo de una respuesta inmunológica.

En determinadas formas de realización, la proteína Env del VIH se puede expresar sobre una partícula, tal como un liposoma, una partícula tipo virus (VLP), una nanopartícula, un virosoma o exosoma, opcionalmente en combinación con adyuvantes endógenos y/o exógenos. Cuando se comparan con la proteína Env soluble o monomérica por sí sola, dichas partículas presentan típicamente una mayor eficacia de presentación de antígenos in vivo.

Los ejemplos de VLP que expresan la proteína Env del VIH se pueden preparar, por ejemplo, por coexpresión de la proteína Env del VIH con proteínas virales autoensamblantes tal como el core Gag del VIH u otras proteínas Gag retrovirales. Las VLP se asemejan a virus, pero no son infecciosas porque no contienen un material genético viral. La expresión de proteínas estructurales virales, tal como la envoltura o cápside, puede dar como resultado el autoensamblaje de las VLP. Las VLP son bien conocidas por el especialista, y su uso en vacunas se describe, por ejemplo, en (Kushnir et al, 2012).

En algunas formas de realización preferidas, la partícula es un liposoma. Un liposoma es una vesícula esférica que tiene por lo menos una bicapa lipídica. Las proteínas de la Env del VIH triméricas, por ejemplo, se pueden acoplar de manera no covalente a dichos liposomas mediante interacciones electrostáticas, por ejemplo por adición de una marca His al extremo C-terminal de la Env del VIH trimérica e incorporación de un átomo quelante bivalente tal como Ni^{2+} o Co^{2+} en el grupo de cabeza de los lípidos derivatizados en el liposoma. En algunos ejemplos de formas de realización no taxativas, el liposoma comprende 1,2-distearoil-sn-glicero-3-fosfocolina (DSPC), colesterol y la sal de níquel o cobalto de 1,2-dioleoil-sn-glicero-3-[(N-(ácido 5-amino-1-carboxipentil)iminodiacético)succinilo] (DGS-NTA(Ni^{2+}) o DGS-NTA(Co^{2+})) a una relación molar de 60:36:4. En formas de realización preferidas, las proteínas de la Env del VIH triméricas se acoplan de manera covalente a la superficie liposómica, por ejemplo por medio de un grupo funcional maleimida integrado en la superficie del liposoma. En determinados ejemplos de formas de realización no taxativas de las mismas, el liposoma comprende los lípidos DSPC, colesterol y 1,2-dipalmitoil-sn-glicero-3-fosfoetanolamina-N-[4-(p-maleimidometil)ciclohexan-carboxamida] a una relación molar de 54:30:16. La proteína Env del VIH se puede acoplar a la misma, por ejemplo, a través de una cisteína C-terminal agregada en la proteína Env del VIH. Las variantes acopladas de manera covalente son más estables, generan títulos elevados de IgG específico del antígeno y los epitopes en la 'base' antigénicamente menos relevante de la Env trimérica quedan enmascaradas. Los métodos para preparar los trímeros de la Env del VIH acoplados a liposomas, así como su caracterización, son conocidos y se han descrito, por ejemplo, en (Bale et al, 2017). La invención también provee una proteína Env del VIH de la invención fusionada y/o expresada sobre un liposoma.

En determinadas formas de realización, se fusiona una proteína Env del VIH de la invención a partículas autoensamblantes o se expresa sobre nanopartículas. Las nanopartículas de antígenos son ensamblajes de polipéptidos que presentan múltiples copias de antígenos, por ejemplo, la proteína Env del VIH de la presente invención, que dan como resultado múltiples sitios de unión (avidéz) y pueden proporcionar una mayor inmunogenicidad y estabilidad del antígeno. La preparación y el uso de nanopartículas proteicas autoensamblantes para usar en vacunas son bien conocidos por el especialista, véase, por ejemplo (Zhao et al, 2014), (López-Sagasetta et al, 2016). A modo de ejemplos no taxativos, las nanopartículas que se pueden autoensamblar se pueden basar en ferritina, bacterioferritina o DPS. Las nanopartículas de DPS que expresan proteínas sobre su superficie se describe, por ejemplo, en WO2011/082087. La descripción de los antígenos trimérico del VIH-1 sobre dichas partículas se han descrito, por ejemplo, en (He et al, 2016). Otras nanopartículas proteicas autoensamblantes, así como su preparación, se divulgan, por ejemplo, en WO 2014/124301 y US 2016/0122392. La invención también provee una proteína Env del VIH de la invención fusionada y/o expresada sobre una nanopartícula autoensamblante. La invención también provee composiciones que comprenden VLP, liposomas o nanopartículas autoensamblantes de acuerdo con la invención.

En determinadas formas de realización, se incluye un adyuvante en una composición de la invención o se coadministra con una composición de la invención. El uso de adyuvantes es opcional, y puede mejorar aún más las respuestas inmunológicas cuando la composición se utiliza en vacunaciones. Los adyuvantes adecuados para la coadministración o

5 inclusión en las composiciones de acuerdo con la invención preferiblemente serán aquellas que son potencialmente seguros, bien tolerados y eficaces en las personas. Dichos adyuvantes son bien conocidos por el especialista, y los ejemplos no taxativos incluyen QS-21, Detox-PC, MPL-SE, MoGM-CSF, TiterMax-G, CRL-1005, GERBU, TERamida, PSC97B, Adjumer, PG-026, GSK-I, GcMAF, B-aletina, MPC-026, Adjuvax, CpG ODN, Betafectina, sales de aluminio tal como fosfato de aluminio (por ejemplo AdjuPhos) o hidróxido de aluminio, y MF59.

10 Otros aspectos de la invención se relacionan con proteínas de la envoltura del VIH recombinantes que comprenden una secuencia de aminoácidos que es por lo menos 95%, 96%, 97%, 98%, 99% o 100% idéntica a la secuencia de aminoácidos de la SEQ ID N°: 2 o SEQ ID N°: 4, que representan secuencias de la envoltura del VIH consenso del clado C y consenso del clado B, respectivamente. Estas secuencias consenso no se han observado en ninguna secuencia natural, y por
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80
 85
 90
 95
 100
 105
 110
 115
 120
 125
 130
 135
 140
 145
 150
 155
 160
 165
 170
 175
 180
 185
 190
 195
 200
 205
 210
 215
 220
 225
 230
 235
 240
 245
 250
 255
 260
 265
 270
 275
 280
 285
 290
 295
 300
 305
 310
 315
 320
 325
 330
 335
 340
 345
 350
 355
 360
 365
 370
 375
 380
 385
 390
 395
 400
 405
 410
 415
 420
 425
 430
 435
 440
 445
 450
 455
 460
 465
 470
 475
 480
 485
 490
 495
 500
 505
 510
 515
 520
 525
 530
 535
 540
 545
 550
 555
 560
 565
 570
 575
 580
 585
 590
 595
 600
 605
 610
 615
 620
 625
 630
 635
 640
 645
 650
 655
 660
 665
 670
 675
 680
 685
 690
 695
 700
 705
 710
 715
 720
 725
 730
 735
 740
 745
 750
 755
 760
 765
 770
 775
 780
 785
 790
 795
 800
 805
 810
 815
 820
 825
 830
 835
 840
 845
 850
 855
 860
 865
 870
 875
 880
 885
 890
 895
 900
 905
 910
 915
 920
 925
 930
 935
 940
 945
 950
 955
 960
 965
 970
 975
 980
 985
 990
 995

FORMAS DE REALIZACIÓN

25 La forma de realización 1 es una proteína Env del VIH recombinante que comprende la secuencia de aminoácidos de una proteína Env del VIH que tiene los residuos de aminoácidos indicados en por lo menos dos de las posiciones indicadas que se seleccionan del grupo que consiste en:

- (i) Phe, Leu, Met o Trp, preferiblemente Phe, en la posición 651;
- (ii) Phe, Ile, Met o Trp, preferiblemente Ile, en la posición 655;
- 30 (iii) Asn o Gln, preferiblemente Asn, en la posición 535;
- (iv) Val, Ile o Ala, preferiblemente Val o Ile, en la posición 589;
- (v) Phe o Trp, preferiblemente Phe, en la posición 573;
- (vi) Ile en la posición 204; y
- (vii) Phe, Met o Ile, preferiblemente Phe, en la posición 647,

35 en donde la numeración de las posiciones es acorde con la numeración en la gp160 de la forma aislada HXB2 del VIH-1.

La forma de realización 2 es una proteína Env del VIH recombinante, que comprende la secuencia de aminoácidos de una proteína Env del VIH y una sustitución de aminoácido por el residuo de aminoácido indicado en por lo menos una de las posiciones indicadas que se seleccionan del grupo que consiste en:

- 40 (i) Phe, Leu, Met o Trp, preferiblemente Phe, en la posición 651;
- (ii) Phe, Ile, Met o Trp, preferiblemente Ile, en la posición 655;
- (iii) Asn o Gln, preferiblemente Asn, en la posición 535;
- (iv) Val, Ile o Ala, preferiblemente Val o Ile, en la posición 589;
- (v) Phe o Trp, preferiblemente Phe, en la posición 573;
- 45 (vi) Ile en la posición 204; y
- (vii) Phe, Met o Ile, preferiblemente Phe, en la posición 647,

en donde la proteína Env del VIH se selecciona del grupo que consiste en:

- 5 (1) una proteína Env del VIH que tiene una secuencia consenso, por ejemplo del clado C o del clado B, por ejemplo que comprende la secuencia de aminoácidos de la (SEQ ID N°: 2, 3, 4 o 5); o
- 5 (2) una proteína Env del VIH sintética, por ejemplo que comprende la secuencia de aminoácidos de (a) la SEQ ID N°: 6; (b) la SEQ ID N°: 6 con una mutación de Glu por Arg en la posición 166; (c) la SEQ ID N°: 7; o (d) la SEQ ID N°: 8 o SEQ ID N°: 9), en donde (a), (b) o (d) opcionalmente pueden comprender además mutaciones SOSIP (Cys en las posiciones 501 y 605 y Pro en la posición 559) y/o del sitio de clivaje de furina (por ejemplo la SEQ ID N°: 10 que reemplaza los aminoácidos 508-511); o
- 10 (3) una proteína Env del VIH de tipo salvaje, preferiblemente del clado C, que comprende por lo menos una mutación de reparación en un residuo de aminoácido que está presente en la posición correspondiente a una frecuencia menor que 7.5%, preferiblemente menor que 2%, de las secuencias Env del VIH en una colección de por lo menos 100, preferiblemente por lo menos 1000, preferiblemente por lo menos 10000, secuencias Env del VIH de tipo salvaje, en donde la mutación de reparación es una sustitución por un residuo de aminoácido que está presente en la posición correspondiente a una frecuencia de por lo menos 10% en las secuencias Env del VIH en dicha colección y preferiblemente la mutación de reparación es una sustitución por el residuo de aminoácido que está presente en la posición correspondiente con mayor frecuencia en dicha colección; y

en donde la numeración de las posiciones es acorde con la numeración en la gp160 de la forma aislada HXB2 del VIH-1.

- 20 La forma de realización 3 es una proteína Env del VIH recombinante, que comprende la secuencia de aminoácidos de una proteína Env del VIH y una sustitución de aminoácido por el residuo de aminoácido indicado en por lo menos una de las posiciones indicadas que se seleccionan del grupo que consiste en:

- (i) Phe, Leu, Met o Trp, preferiblemente Phe, en la posición 651;
- (ii) Phe, Ile, Met o Trp, preferiblemente Ile, en la posición 655;
- 25 (iii) Asn o Gln, preferiblemente Asn, en la posición 535;
- (iv) Val, Ile o Ala, preferiblemente Val o Ile, en la posición 589;
- (v) Phe o Trp, preferiblemente Phe, en la posición 573;
- (vi) Ile en la posición 204; y
- (vii) Phe, Met o Ile, preferiblemente Phe, en la posición 647,

- 30 en donde la proteína Env del VIH es una proteína Env del VIH mutante SOSIP que comprende por lo menos una mutación seleccionada del grupo que consiste en:

- (a) Cys en las posiciones 501 y 605;
- (b) Pro en la posición 559;
- (c) Cys en las posiciones 501 y 605 y Pro en la posición 559; y

- 35 la numeración de las posiciones es acorde con la numeración en la gp160 de la forma aislada HXB2 del VIH-1.

La forma de realización 4 es la proteína Env del VIH recombinante de la forma de realización 2, que comprende los residuos de aminoácidos indicados en por lo menos dos de las posiciones indicadas que se seleccionan del grupo que consiste en (i) a (vii).

- 40 La forma de realización 5 es la proteína Env del VIH recombinante de la forma de realización 3, que comprende los residuos de aminoácidos indicados en por lo menos dos de las posiciones indicadas que se seleccionan del grupo que consiste en (i) a (vii).

La forma de realización 6 es la proteína Env del VIH recombinante de cualquiera de las formas de realización 1, 2 y 4, que además comprende Cys en las posiciones 501 y 605 o Pro en la posición 559, preferiblemente Cys en las posiciones 501 y 605 y Pro en la posición 559.

La forma de realización 7 es la proteína Env del VIH recombinante de cualquiera de las formas de realización 1 a 6, que comprende los residuos de aminoácidos indicados en por lo menos tres de las posiciones indicadas que se seleccionan del grupo que consiste en (i) a (vii).

5 La forma de realización 8 es la proteína Env del VIH recombinante de cualquiera de las formas de realización 1 a 6, que comprende los residuos de aminoácidos indicados en por lo menos cuatro de las posiciones indicadas que se seleccionan del grupo que consiste en (i) a (vii).

La forma de realización 9 es la proteína Env del VIH recombinante de cualquiera de las formas de realización 1 a 6, que comprende los residuos de aminoácidos indicados en por lo menos cinco de las posiciones indicadas que se seleccionan del grupo que consiste en (i) a (vii).

10 La forma de realización 10 es la proteína Env del VIH recombinante de cualquiera de las formas de realización 1 a 6, que comprende los residuos de aminoácidos indicados en por lo menos seis de las posiciones indicadas que se seleccionan del grupo que consiste en (i) a (vii).

15 La forma de realización 11 es la proteína Env del VIH recombinante de cualquiera de las formas de realización 1 a 6, que comprende los residuos de aminoácidos indicados en siete de las posiciones indicadas que se seleccionan del grupo que consiste en (i) a (vii).

20 La forma de realización 12 es la proteína Env del VIH recombinante de cualquiera de las formas de realización 1, 4 y 5, en donde dichas por lo menos dos posiciones y residuos indicados son una combinación que se selecciona del grupo que consiste en: 651F, 655I; 651F, 535N; 651F, 589V; 651F, 589I; 651F, 573F; 651F, 204I; 651F, 647F; 655I, 535N; 655I, 589V; 655I, 589I; 655I, 573F; 655I, 204I; 655I, 647F; 535N, 589V; 535N, 589I; 535N, 573F; 535N, 204I; 535N, 647F; 589V, 573F; 589V, 204I; 589V, 647F; 589I, 573F; 589I, 204I; 589I, 647F; 573F, 204I; 573F, 647F; y 204I, 647F.

25 La forma de realización 13 es la proteína Env del VIH recombinante de la forma de realización 7, en donde dichas por lo menos tres posiciones y residuos indicados son una combinación que se selecciona del grupo que consiste en: 651F, 655I, 535N; 651F, 589V, 535N; 651F, 589I, 535N; 651F, 573F, 535N; 651F, 204I, 535N; 651F, 647F, 535N; 655I, 589V, 535N; 655I, 589I, 535N; 655I, 573F, 535N; 655I, 204I, 535N; 655I, 647F, 535N; 589V, 573F, 535N; 589I, 573F, 535N; 589I, 204I, 535N; 589I, 647F, 535N; 573F, 204I, 535N; 573F, 647F, 535N; 204I, 647F, 535N; 651F, 655I, 589V; 651F, 573F, 589V; 651F, 204I, 589V; 651F, 647F, 589V; 655I, 573F, 589V; 655I, 204I, 589V; 655I, 647F, 589V; 573F, 204I, 589V; 573F, 647F, 589V; 204I, 647F, 589V; 651F, 655I, 589I; 651F, 573F, 589I; 651F, 204I, 589I; 651F, 647F, 589I; 655I, 573F, 589I; 655I, 204I, 589I; 655I, 647F, 589I; 573F, 204I, 589I; 573F, 647F, 589I; 204I, 647F, 589I; 651F, 655I, 573F; 651F, 204I, 573F; 651F, 647F, 573F; 655I, 204I, 573F; 655I, 647F, 573F; 204I, 647F, 573F; 651F, 655I, 204I; 651F, 647F, 204I; 655I, 647F, 204I; 651F, 655I, 647F; 655I, 651F, 647F; 655I, 651F, 535N; 655I, 589V, 573F; y 655I, 589V, 204I.

30 La forma de realización 14 es la proteína Env del VIH recombinante de la forma de realización 8, en donde dichas por lo menos cuatro posiciones y residuos indicados son una combinación que se selecciona del grupo que consiste en: 651F, 655I, 535N, 589V, 573F; 651F, 655I, 535N, 573F; 651F, 655I, 589V, 573F; 651F, 535N, 589V, 573F; 655I, 535N, 589V, 573F; 651F, 655I, 535N, 204I; 651F, 655I, 589V, 204I; 651F, 535N, 589V, 204I; 655I, 535N, 589V, 204I; 651F, 655I, 573F, 204I; 651F, 535N, 573F, 204I; 655I, 535N, 573F, 204I; 651F, 589V, 573F, 204I; 655I, 589V, 573F, 204I; 535N, 589V, 573F, 204I; 651F, 655I, 535N, 647F; 651F, 655I, 589V, 647F; 651F, 535N, 589V, 647F; 655I, 535N, 589V, 647F; 651F, 655I, 573F, 647F; 651F, 535N, 573F, 647F; 655I, 535N, 573F, 647F; 651F, 589V, 573F, 647F; 655I, 589V, 573F, 647F; 535N, 589V, 573F, 647F; 651F, 655I, 204I, 647F; 651F, 535N, 204I, 647F; 655I, 535N, 204I, 647F; 651F, 589V, 204I, 647F; 655I, 589V, 204I, 647F; 535N, 589V, 204I, 647F; y 589V, 573F, 204I, 647F.

35 La forma de realización 15 es la proteína Env del VIH recombinante de la forma de realización 9, en donde dichas por lo menos cinco posiciones y residuos indicados son una combinación que se selecciona del grupo que consiste en: 651F, 655I, 535N, 589V, 573F; 651F, 655I, 535N, 589V, 204I; 651F, 655I, 535N, 573F, 204I; 651F, 655I, 589V, 573F, 204I; 655I, 535N, 589V, 573F, 204I; 651F, 655I, 535N, 589V, 647F; 651F, 655I, 535N, 573F, 647F; 651F, 655I, 589V, 573F, 647F; 651F, 535N, 589V, 573F, 647F; 655I, 535N, 589V, 573F, 647F; 651F, 655I, 535N, 204I, 647F; 651F, 655I, 589V, 204I, 647F; 651F, 535N, 589V, 204I, 647F; 655I, 535N, 589V, 204I, 647F; 651F, 655I, 573F, 204I, 647F; 651F, 535N, 573F, 204I, 647F; 655I, 535N, 573F, 204I, 647F; 651F, 589V, 573F, 204I, 647F; 655I, 589V, 573F, 204I, 647F; y 535N, 589V, 573F, 204I, 647F.

50 La forma de realización 16 es la proteína Env del VIH recombinante de la forma de realización 10, en donde dichas por lo menos seis posiciones y residuos indicados son una combinación que se selecciona del grupo que consiste en: 651F, 655I, 535N, 589V, 573F, 204I; 651F, 655I, 535N, 589V, 573F, 647F; 651F, 655I, 535N, 589V, 204I, 647F; 651F, 655I, 535N, 573F, 204I, 647F; 651F, 655I, 589V, 573F, 204I, 647F; 655I, 535N, 589V, 573F, 204I, 647F; y 655I, 535N, 589V, 573F, 204I, 647F.

La forma de realización 17 es la proteína Env del VIH recombinante de cualquiera de las formas de realización 1 a 16, que además comprende una sustitución de aminoácido por el residuo de aminoácido indicado en por lo menos una de las posiciones indicadas que se seleccionan del grupo que consiste en:

- (viii) Gln, Glu, Ile, Met, Val, Trp o Phe, preferiblemente Gln o Glu, en la posición 588;
- 5 (ix) Lys en la posición 64 o Arg en la posición 66 o ambos Lys en la posición 64 y Arg en la posición 66;
- (x) Trp en la posición 316;
- (xi) Cys en ambas posiciones 201 y 433;
- (xii) Pro en la posición 556 o 558 o en ambas posiciones 556 y 558; y
- 10 (xiii) reemplazo del bucle en las posiciones de aminoácido 548-568 (bucle HR1) por un bucle de 7-10 aminoácidos, preferiblemente un bucle de 8 aminoácidos, por ejemplo que tiene una secuencia seleccionada entre cualquiera de las (SEQ ID N°: 12-17);
- (xiv) Gly en la posición 568 o Gly en la posición 569 o Gly en la posición 636 o Gly en ambas posiciones 568 y 636 o Gly en ambas posiciones 569 y 636; y/o
- 15 (xv) Tyr en la posición 302 o Arg en la posición 519 o Arg en la posición 520 o Tyr en la posición 302 y Arg en la posición 519 o Tyr en la posición 302 y Arg en la posición 520 o Tyr en la posición 302 y Arg en ambas posiciones 519 y 520,

en donde la numeración de las posiciones es acorde con la numeración en la gp160 de la forma aislada HXB2 del VIH-1.

La forma de realización 18 es la proteína Env del VIH recombinante de cualquiera de las formas de realización 1 a 17, que además comprende una mutación en una secuencia de clivaje de furina de la proteína Env del VIH.

- 20 La forma de realización 19 es la proteína Env del VIH recombinante de la forma de realización 18, en donde la mutación en el sitio de clivaje de furina es un reemplazo en las posiciones 508-511 por RRRRRR (SEQ ID N°: 10).

La forma de realización 20 es la proteína Env del VIH recombinante de cualquiera de las formas de realización 1 a 19, que es una gp140 o gp160.

- 25 La forma de realización 21 es la proteína Env del VIH recombinante de cualquiera de las formas de realización 1 a 20, en donde la proteína Env del VIH recombinante presenta por lo menos uno entre un aumento del porcentaje de formación de trímeros y un rendimiento de trímeros mejorado en comparación con una proteína Env del VIH que no comprende dichos uno o más de los residuos de aminoácidos indicados en las posiciones indicadas que se seleccionan del grupo que consiste en (i) a (vii).

- 30 La forma de realización 22 es la proteína Env del VIH recombinante de la forma de realización 21, en donde la formación de trímeros se mide mediante cromatografía de exclusión por tamaño con dispersión de luz multiangular (SEC-MALS).

La forma de realización 23 es la proteína Env del VIH recombinante de cualquiera de las formas de realización 1 a 22, que además comprende un residuo de aminoácido seleccionado entre Val, Ile, Phe, Met, Ala o Leu, preferiblemente Val o Ile, más preferiblemente Val, en la posición 658.

- 35 La forma de realización 24 es la proteína Env del VIH recombinante de cualquiera de las formas de realización 1 a 23, que comprende la combinación de aminoácidos seleccionada del grupo que consiste en:

(a) 655I, 589V, 573F, 651F, 588E, 535N, 204I;

(b) 556P, 655I, 535N, 573F, 589V, 204I, 588Q;

(c) 204I, 535N, 556P, 588E, 589V, 651F, 655I;

(d) 535N, 556P, 589V, 651F, 655I; y

- 40 (e) 535N, 556P, 588E, 589V, 651F, 655I.

La forma de realización 25 es la proteína Env del VIH recombinante de cualquiera de las formas de realización 1 a 24, que comprende una secuencia de aminoácidos que es por lo menos 95%, 96%, 97%, 98%, 99% idéntica o 100% idéntica a

cualquiera de las SEQ ID N°: 3, 5, 20, 22, 24, 26, 27, 28, 29, 30, 31 o 32, preferiblemente por lo menos 98% idéntica a cualquiera de las SEQ ID N°: 20, 22, 24, 26, 27, 28, 29, 30, 31 o 32.

La forma de realización 26 es un complejo trimérico que comprende un oligómero no covalente de tres de las proteínas Env del VIH recombinantes de cualquiera de las formas de realización 1 a 25.

- 5 La forma de realización 27 es una partícula, por ejemplo un liposoma o una nanopartícula, por ejemplo una nanopartícula autoensamblante, que expresa la proteína Env del VIH recombinante de cualquiera de las formas de realización 1-25 o el complejo trimérico de la forma de realización 26.

La forma de realización 28 es una molécula de ácido nucleico aislada que codifica una proteína Env del VIH recombinante de cualquiera de las formas de realización 1 a 25.

- 10 La forma de realización 29 es un vector que comprende la molécula de ácido nucleico aislada de la forma de realización 28 ligada operativamente a un promotor.

La forma de realización 30 es el vector de la forma de realización 29, que es un vector de adenovirus.

La forma de realización 31 es una célula huésped que comprende la molécula de ácido nucleico aislada de la forma de realización 28 o el vector de la forma de realización 29 o 30.

- 15 La forma de realización 32 es un método para producir una proteína Env del VIH recombinante, que comprende cultivar la célula huésped de la forma de realización 31 en condiciones adecuadas para la producción de la proteína Env del VIH recombinante.

- 20 La forma de realización 33 es un método para producir una proteína Env del VIH recombinante que comprende obtener un vector de expresión que comprende el ácido nucleico aislado de la forma de realización 28 ligado operativamente a un promotor; transfectar una célula con el vector de expresión; cultivar la célula transfectada en condiciones adecuadas para la expresión de la proteína Env del VIH recombinante; y purificar la proteína Env del VIH recombinante en condiciones que permitan la formación de un complejo trimérico estabilizado.

- 25 La forma de realización 34 es un método para producir una proteína Env del VIH recombinante de acuerdo con cualquiera de las formas de realización 1 a 25, que comprende introducir por lo menos una sustitución de aminoácido que da como resultado el residuo de aminoácido indicado en una posición seleccionada del grupo que consiste en (i)-(vii), es una secuencia del esqueleto de la proteína de la envoltura del VIH.

La forma de realización 35 es el método de acuerdo con forma de realización 34, en donde se introduce una secuencia de nucleótidos que codifica la sustitución de aminoácido en un ácido nucleico que codifica la secuencia del esqueleto de la proteína de la envoltura del VIH.

- 30 La forma de realización 36 es el método de las formas de realización 34 o 35, en donde la secuencia del esqueleto de la proteína de la envoltura del VIH se selecciona del grupo que consiste en: SEQ ID N°: 2, la SEQ ID N°: 3, la SEQ ID N°: 4, la SEQ ID N°: 5, la SEQ ID N°: 6, la SEQ ID N°: 7, la SEQ ID N°: 8, la SEQ ID N°: 9, la SEQ ID N°: 6 que tiene una mutación de Glu por Arg en la posición 166; SEQ ID N°: 6 que tiene Cys en las posiciones 501 y 605 y Pro en la posición 559 y/o que tiene la SEQ ID N°: 10 que reemplaza los aminoácidos 508-511; SEQ ID N°: 6 que tiene una mutación de Glu por Arg en la posición 166, que además tiene Cys en las posiciones 501 y 605 y Pro en la posición 559 y/o que tiene la SEQ ID N°: 10 que reemplaza los aminoácidos 508-511; SEQ ID N°: 8 que tiene Cys en las posiciones 501 y 605 y Pro en la posición 559 y/o que tiene la SEQ ID N°: 10 que reemplaza los aminoácidos 508-511; SEQ ID N°: 9 que tiene Cys en las posiciones 501 y 605 y Pro en la posición 559 y/o que tiene la SEQ ID N°: 10 que reemplaza los aminoácidos 508-511; y una proteína Env del VIH de tipo salvaje que comprende mutaciones que dan como resultado por lo menos (a), (b) o (c), preferiblemente por lo menos dos de (a), (b) y (c), más preferiblemente (a), (b) y (c) de los siguientes: (a) Cys en las posiciones 501 y 506 y Pro en la posición 559,

(b) de la SEQ ID N°: 10 que reemplaza los aminoácidos 508-511, y/o

- 45 (c) por lo menos una mutación de reparación en un residuo de aminoácido que está presente en la posición correspondiente a una frecuencia menor que 7.5%, preferiblemente menor que 2%, de las secuencias Env del VIH en una colección de por lo menos 100, preferiblemente por lo menos 1000, preferiblemente por lo menos 10000, secuencias Env del VIH de tipo salvaje, en donde la mutación de reparación es una sustitución por un residuo de aminoácido que está presente en la posición correspondiente a una frecuencia de por lo menos 10% de las secuencias Env del VIH en dicha colección y preferiblemente la mutación de reparación es una sustitución por el residuo de aminoácido que está presente en la posición correspondiente con mayor frecuencia en dicha colección.

- La forma de realización 37 es una composición que comprende la proteína Env del VIH recombinante de cualquiera de las formas de realización 1 a 25, el complejo trimérico de la forma de realización 26, la partícula de la forma de realización 27, la molécula de ácido nucleico aislada de la forma de realización 28, el vector de la forma de realización 29 o 30, o la célula huésped de la forma de realización 31, y un vehículo farmacéuticamente aceptable.
- 5 La forma de realización 38 es una composición de la forma de realización 37, que además comprende un adyuvante.
- La forma de realización 39 es un método para producir la composición de la forma de realización 37, que comprende mezclar la proteína Env del VIH recombinante, el complejo trimérico, la partícula, el ácido nucleico aislado, el vector o la célula huésped con uno o más vehículos farmacéuticamente aceptables.
- 10 La forma de realización 40 es un método para vacunar a un sujeto contra una infección por VIH que comprende administrar al sujeto una composición que comprende la proteína de la envoltura del VIH recombinante de cualquiera de las formas de realización 1 a 25, el complejo trimérico de la forma de realización 26, la partícula de la forma de realización 27 o el vector de la forma de realización 29 o 30.
- 15 La forma de realización 41 es un método para producir una respuesta inmunológica contra una infección por VIH en un sujeto que lo necesita, que comprende administrar al sujeto una composición que comprende la proteína de la envoltura del VIH recombinante de cualquiera de las formas de realización 1 a 25, el complejo trimérico de la forma de realización 26, la partícula de la forma de realización 27 o el vector de la forma de realización 29 o 30.
- La forma de realización 42 es una proteína Env del VIH recombinante que comprende la secuencia de aminoácidos de la SEQ ID N°: 2, o una secuencia que es por lo menos 95% idéntica a la misma.
- 20 La forma de realización 43 es una proteína Env del VIH recombinante que comprende la secuencia de aminoácidos de la SEQ ID N°: 3, o una secuencia que es por lo menos 95% idéntica a la misma.
- La forma de realización 44 es una proteína Env del VIH recombinante que comprende la secuencia de aminoácidos de la SEQ ID N°: 4, o una secuencia que es por lo menos 95% idéntica a la misma.
- La forma de realización 45 es una proteína Env del VIH recombinante que comprende la secuencia de aminoácidos de la SEQ ID N°: 5, o una secuencia que es por lo menos 95% idéntica a la misma.
- 25 La forma de realización 46 es una molécula de ácido nucleico aislada que codifica la proteína Env del VIH recombinante de cualquiera de las formas de realización 42 a 45.
- La forma de realización 47 es un vector que comprende la molécula de ácido nucleico aislada de la forma de realización 46 ligada operativamente a un promotor.
- La forma de realización 48 es el vector de la forma de realización 47, que es un vector de adenovirus.
- 30 La forma de realización 49 es una célula huésped que comprende la molécula de ácido nucleico aislada de la forma de realización 46 o el vector de la forma de realización 47 o 48.
- La forma de realización 50 es una composición que comprende la proteína Env del VIH recombinante de cualquiera de las formas de realización 42 a 45, la molécula de ácido nucleico aislada de la forma de realización 46, el vector de la forma de realización 47 o 48 o la célula huésped de la forma de realización 49, y un vehículo farmacéuticamente aceptable.
- 35 La forma de realización 51 es un método para mejorar el porcentaje de trímeros y/o el rendimiento de trímeros (que representa el plegamiento y la estabilidad) de una proteína Env del VIH parental, donde dicho método comprende reparar la secuencia de aminoácidos de la proteína Env del VIH parental mediante la introducción de por lo menos una mutación de reparación, preferiblemente por lo menos 3 mutaciones de reparación en la proteína Env del VIH parental, en donde una mutación de reparación es una sustitución de aminoácido en un residuo de aminoácido presente en la posición correspondiente a una frecuencia menor que 7.5%, preferiblemente menor que 2%, de las secuencias Env del VIH en una colección de por lo menos 100, preferiblemente por lo menos 500, preferiblemente por lo menos 1000, preferiblemente por lo menos 10000, secuencias Env del VIH de tipo salvaje, en donde la sustitución es por un residuo de aminoácido presente en la posición correspondiente a una frecuencia de por lo menos 10% de las secuencias Env del VIH en dicha colección y preferiblemente la sustitución es por el residuo de aminoácido presente en la posición correspondiente con mayor frecuencia en dicha colección.
- 40
- 45 La forma de realización 52 es el método de la forma de realización 51, en donde se repara por lo menos 50%, preferiblemente por lo menos 80%, de los residuos de aminoácidos en la proteína Env del VIH parental presentes en las posiciones correspondientes a una frecuencia menor que 7.5% de las secuencias Env del VIH en dicha colección.

La forma de realización 53 es el método de la forma de realización 51, en donde se repara por lo menos 50%, preferiblemente por lo menos 80%, de los residuos de aminoácidos en la proteína Env del VIH parental presentes en las posiciones correspondientes a una frecuencia menor que 2% de las secuencias Env del VIH en dicha colección.

5 La forma de realización 54 es el método de cualquiera de las formas de realización 51 a 53, en donde la proteína Env del VIH parental es del clado C.

La forma de realización 55 es el método de cualquiera de las formas de realización 51 a 54, en donde la proteína Env del VIH parental es una proteína Env del VIH de tipo salvaje.

La forma de realización 56 es el método de cualquiera de las formas de realización 51 a 54, en donde la proteína Env del VIH parental comprende uno o más de los siguientes:

10 (a) Cys en las posiciones 501 y 506 y Pro en la posición 559;

(b) una mutación en una secuencia de clivaje de furina de la proteína Env del VIH, por ejemplo de la SEQ ID N°: 10 que reemplaza los aminoácidos 508-511;

(c) Phe en la posición 651;

(d) Ile en la posición 655;

15 (e) Asn en la posición 535;

(f) Val en la posición 589;

(g) Phe en la posición 573;

(h) Ile en la posición 204;

(i) Phe en la posición 647;

20 (j) Val en la posición 658;

(k) Gln o Glu en la posición 588; y/o

(l) Pro en la posición 556, 558 o 556 y 558.

La forma de realización 57 es una proteína Env del VIH recombinante que se puede obtener mediante el método de cualquiera de las formas de realización 51 a 56.

25 EJEMPLOS

Ejemplo 1: Generación de secuencias consenso de la envoltura clado C y clado B del VIH

Secuencia consenso de la envoltura del clado C del VIH

30 Se desarrolló una secuencia consenso de la proteína de la envoltura (Env) del clado C del VIH como la secuencia del esqueleto para estudiar los efectos de varias mutaciones sobre la formación de trímeros de las proteínas Env del VIH. Se descargó un alineamiento de secuencias de 3,434 secuencias de la proteína de la envoltura de formas aisladas virales del VIH conocidas de la base de datos Los Alamos (<http://www.hiv.lanl.gov/content/index>). Entre las 3,434 secuencias, solamente se seleccionaron 1,252 secuencias del clado C para generar la secuencia consenso de la proteína Env del clado C del VIH. En las posiciones donde no se pudo identificar claramente un residuo consenso basado en el

35 alineamiento, se usó la secuencia consenso para identificar las secuencias de tipo salvajes más próximas mediante una búsqueda BLAST. Luego se seleccionó el residuo consenso de estas posiciones como el aminoácido en las secuencias de tipo salvaje más próximas identificadas a partir de la búsqueda BLAST. La secuencia consenso de la Env del clado C del VIH se muestra en la SEQ ID N°: 2. Las dos secuencias con la mayor homología con la SEQ ID N°: 2 usando BLAST fueron las secuencias con los números de Genbank ADM30337.1 y ADM30340.1, ambas con un 90% de identidad de secuencia con la SEQ ID N°: 2.

40 La secuencia consenso de la Env del clado C del VIH se modificó adicionalmente mediante introducción de las denominadas mutaciones SOSIP, que incluyen residuos cisteína en las posiciones 501 y 605 y un residuo prolina en la posición 559, así como optimización del sitio de clivaje de furina mediante reemplazo del sitio de furina en los residuos 508-511 por 6 residuos arginina. Además, se mutó Val en la posición 295 por Asn (V295N) para crear un sitio de glicosilación N-ligado presente en la mayoría de las cepas del VIH y que puede mejorar la unión a determinados

anticuerpos usados en algunos experimentos. Adicionalmente, se truncó el extremo C-terminal en el residuo 664, dando como resultado una secuencia que codifica una proteína gp140 del VIH soluble. Todas las posiciones de sustitución/modificación que se describieron antes son relativas a la numeración en la gp160 de la forma aislada HXB2 del VIH-1. La secuencia resultante de la gp140 del VIH, denominada "ConC_SOSIP", se muestra en la (SEQ ID N°: 3). La secuencia ConC_SOSIP se usó como la secuencia de esqueleto o parental de la envoltura del VIH en la cual se introdujeron mutaciones adicionales, por ejemplo, sustituciones de aminoácidos simples y dobles, para producir proteínas Env del VIH recombinantes de acuerdo con las formas de realización de la invención.

Secuencia consenso de la envoltura del clado B del VIH

Se generó una secuencia consenso de la Env del clado B del VIH usando un procedimiento similar al que se describió antes para generar la secuencia consenso de la Env del clado C del VIH. La secuencia consenso del clado B se generó usando 1,708 secuencias de la proteína de la envoltura del clado B de formas aisladas virales del clado B conocidas. La secuencia consenso de la Env del clado B del VIH se muestra en la SEQ ID N°: 4.

La secuencia consenso de la Env del clado B del VIH se modificó adicionalmente mediante introducción de las denominadas mutaciones SOSIP, optimización del sitio de clivaje de furina por reemplazo del sitio de furina por 6 residuos arginina, y truncamiento del extremo C-terminal en el residuo 664, como se describió antes, dando como resultado una secuencia que codifica una secuencia consenso de la gp140 del clado B del VIH soluble. La secuencia resultante de la proteína Env de la gp140 del VIH, denominada "ConB_SOSIP", se muestra en la (SEQ ID N°: 5).

Se encontró sorprendentemente que las moléculas basadas en el consenso tenían niveles de expresión mejoradas ante las moléculas basadas en formas aisladas naturales, y además ya tenían niveles de trimerización mejoradas. Por ende, las moléculas de las SEQ ID N°: 2-5 ya presentan propiedades sorprendentemente ventajosas.

Ejemplo 2: Expresión y purificación de la proteína Env del VIH recombinante

Las proteínas Env del VIH recombinantes se expresaron y purificaron como proteínas de la gp140 solubles. Se introdujeron mutaciones simples (sustituciones de aminoácidos) y combinaciones de las mismas (por ejemplo, mutaciones dobles y triples) en la secuencia consenso de esqueleto ConC_SOSIP para generar una serie de variantes de la proteína Env del VIH recombinante.

Generación y expresión de construcciones y variantes de la Env de la gp140 del VIH

El ADN que codifica la secuencia consenso de la Env del clado C del VIH, ConC_SOSIP, indicado en la SEQ ID N°: 3 se sintetizó y se optimizaron los codones en GenScript (Piscataway, NJ 08854) o Gene Art (Life Technologies, Carlsbad, CA). La secuencia de codones optimizados se clonó luego en el vector pcDNA2004 para generar una construcción de la Env gp140 del clado C del VIH, que se usó como la secuencia de esqueleto de la envoltura del VIH para introducir mutaciones adicionales. Se introdujeron mutaciones en la secuencia de esqueleto ConC_SOSIP mediante mutagénesis dirigida al sitio y reacción en cadena de la polimerasa (PCR) aplicadas a la construcción de la Env gp140 del clado C del VIH, pcDNA2004. Se transfectaron transitoriamente células HEK-Expi293F o células HEK293F con un 90% del vector pcDNA2004 que codifica la secuencia ConC_SOSIP o una variante de la misma y 10% de un vector pcDNA2004 que codifica la furina proteasa (furina-pcDNA2004) de acuerdo con las instrucciones del proveedor. Las células transfectadas se cultivaron durante 5 días a 37 °C y 10% de CO₂. Los sobrenadantes del cultivo se centrifugaron durante 10 minutos a 1250 x g. A continuación, el sobrenadante centrifugado se filtró bajo condiciones estériles usando un filtro de vacío de 0.22 µm y se guardó a 4 °C hasta su uso posterior.

Para las expresiones en un formato de 96 pocillos las células se cultivaron durante 3 días a 37 °C y 10% de CO₂. Se mezclaron 4 µl de Optimem (medio de cultivo) con 4 µl de ADN 100 ng/µl y se agregaron 8 µl de mezcla Expi293F (Optimem 54 µl/ml) e incubaron durante 20 minutos. Subsiguientemente, se agregaron 200 µl/pocillo de células Expi293F a 2.5 x 10⁶ células/ml. Se cosechó el sobrenadante del cultivo y se centrifugó durante 5 minutos a 300 g para separar células y desechos celulares. A continuación, el sobrenadante centrifugado se filtró bajo condiciones estériles usando un filtro de vacío de 0.22 µm y se guardó a 4 °C hasta su uso posterior.

Purificación de la proteína Env gp140 del VIH

La proteína Env gp140 VIH expresada a partir del vector pcDNA2004 se purificó de acuerdo con un protocolo de purificación de dos pasos usando una columna de lectina de *Galantus nivalis* (Vectorlabs, AL-1243) para la purificación inicial y una columna (GE) Superdex200 Increase en un paso subsiguiente para eliminar los contaminantes residuales. Para el paso inicial usando la columna de lectina de *Galantus nivalis*, se diluyó el sobrenadante de cultivo con solución amortiguadora (Tris 40 mM, NaCl 500 mM, pH 7.5) y se hizo pasar sobre una columna de agarosa lectina CV Tricorn 10-50 de 4 ml a una velocidad de 4 ml por minuto. A continuación, la columna se lavó con cuatro volúmenes de columna de solución amortiguadora (Tris 40 mM, NaCl 500 mM, pH 7.5) y se eluyó con cuatro volúmenes de columna de Tris 40 mM,

NaCl 500 mM y manopiranosido 1 M, pH 7.5 con un flujo ascendente de 1.6 ml/min, lo que significa que se cambió la dirección del flujo desde abajo hacia arriba para aumentar la velocidad de elución de la proteína de la envoltura y disminuir el volumen de elución. El eluato se concentró usando un concentrador centrífugo (50K, Amicon Ultra, Millipore).

5 La proteína Env gp140 VIH se purificó adicionalmente sobre una columna Superdex200 usando Tris 50 mM, NaCl 150 mM, pH 7.4 como solución amortiguadora de corrida. El segundo pico que eluyó de la columna contenía la proteína Env gp140 VIH. Se agruparon las fracciones que contienen a este pico y se confirmó la identidad del pico como proteína Env gp140 VIH usando transferencia Western y SDS-PAGE, y/o análisis SEC-MALS. La concentración de la proteína Env gp140 VIH purificada se determinó midiendo la densidad óptica a 280 nm y la proteína Env gp140 VIH purificada se guardó a 4 °C hasta su uso posterior.

10 Análisis por SDS-PAGE y transferencia Western

Se analizaron muestras sobrenadantes de cultivos celulares que contienen la proteína Env gp140 VIH expresada y se purificó la proteína Env gp140 VIH sobre geles de Bis-Tris NuPAGE 4-12% (p/v), MOPS 1X (Life Technologies) bajo condiciones reductoras o no reductoras y se transfirieron usando la tecnología iBlot (Life Technologies). Todos los procedimientos se condujeron de acuerdo con las instrucciones del proveedor. Para el análisis de pureza, los geles se tiñeron con el colorante para proteínas Krypton Infrared (Thermo Scientific) o el colorante para proteínas SYPRO Rubi (Bio-Rad). Para el análisis de transferencia Western, las membranas se evaluaron con un anticuerpo anti-marca de 6x-Histidina (anti-His-HRP). Los geles y las membranas de transferencia se escanearon con un instrumento Odyssey (Li-Cor) y las imágenes se analizaron usando el software Odyssey 3.0 (Li-Cor).

20 Obtención de imágenes de la formación de trímeros de la Env del VIH mediante microscopía electrónica con tinción negativa

Se usó microscopía electrónica con tinción negativa (NS-EM) para obtener imágenes de los trímeros de la proteína de la envoltura que tienen la secuencia de esqueleto ConC_SOSIP, que se purificaron usando lectina de *Galanthus nivalis* seguido por cromatografía de exclusión por tamaño, y se condujo como se describe en Julien et al. 2015 (*Proc. Natl. Acad. Sci.* (2015) 112(38) 11947-52). Las muestras de los trímeros se diluyeron hasta 0.01-0.5 mg/ml en solución amortiguadora salina-Tris (TBS), pH 7.4, y se adhirieron sobre una rejilla 200 Cu recubierta con carbono (EMS CF200-Cu) sobre la cual se había realizado una descarga luminiscente al aire a 2*10⁻¹ mbar, 25 mA, 30 segundos, justo antes del uso. Subsiguientemente, se aplicó una gota de 3 µl drop de la muestra de trímeros diluida a la rejilla durante 1 min seguido por transferencia con papel de filtro (Whatman N° 1 o 4). Las rejillas se secaron durante un minuto, luego se tiñeron con 3 µl de acetato de uranilo (UAc) al 2.3% durante 60 segundos. Los datos se recolectaron usando un microscopio electrónico FEI Tecnai F20 que opera a 120 keV, con un aumento de 25,000x que dio como resultado un tamaño de pixel de 4.68 Å en el plano del espécimen. Las imágenes se adquirieron con un ultra escáner Gatan BM.

Casi todas las partículas en las imágenes (de este material enriquecido en trímeros) eran trímeros cerrados bien formados (no se muestran los datos).

35 **Ejemplo 3: Selección de variantes Env gp140 VIH recombinantes por el rendimiento de trímeros y el porcentaje de formación de trímeros**

Se seleccionaron las variantes de la proteína Env del VIH recombinante generadas en el Ejemplo 2 por la formación de trímeros a fin de identificar aquellas mutaciones que mejoraban el porcentaje de trímero formados y/o por el rendimiento de trímeros mejorado con relación a la secuencia de esqueleto ConC_SOSIP. La selección de alto rendimiento del porcentaje de trímeros y rendimientos de trímeros se condujo usando un ensayo AlphaLISA para evaluar la unión de un panel de anticuerpos ampliamente neutralizantes contra el VIH (bNABs) y no bNABs a las proteínas Env del VIH recombinantes. Los resultados del ensayo AlphaLISA se confirmaron mediante cromatografía de exclusión por tamaño y dispersión de luz multiangular (SEC-MALS).

Análisis del ensayo AlphaLISA®

45 Se midió la expresión total de la proteína Env gp140 VIH y la cantidad total de trímeros nativos correctamente plegados en 200 variantes de gp140 VIH con sustituciones de aminoácidos simples introducidas en la secuencia ConC_SOSIP generada como se describe en el Ejemplo 2, en los sobrenadantes de cultivos celulares mediante un ensayo AlphaLISA. También se evaluaron las variantes de gp140 VIH que contienen mutaciones dobles y triples. La proteína Env del VIH que comprende la secuencia ConC_SOSIP sin ninguna mutación adicional se evaluó a efectos comparativos.

50 Para el análisis se usaron los siguientes anticuerpos monoclonales (mAbs), entre otros: mAb PGT145, mAb PGDM1400, mAb PG16, mAb PGT151, mAb 35O22, mAb PGT128, mAb PG9, mAb F105, mAb B6, mAb 447-52d, mAb 14e y mAb 17b. Los mAbs 447-52D (AB014), PG9 (AB015) y PG16 (AB016) se obtuvieron de Polimun Scientific Immunobiologische Forschung GmbH (Klosterneuburg, Austria). El anticuerpo no neutralizante b6 se obtuvo de Dennis R. Burton (The Scripps

- Research Institute, La Jolla, CA) y el anticuerpo no neutralizante 14e se obtuvo de James E. Robinson (Tulane University, Nueva Orleans, LA). Para los mAbs PGT145 (PDB: 3U1S), PGDM1400 (PDB: 4RQQ), PGT151 (PDB: 4NUG), 35O22 (PDB: 4TVP), F105 (PDB: 1U6A), PGT128 (PDB: 3TYG) y 17b (PDB: 4RQS) se clonaron los ácidos nucleicos que codifican las secuencias publicadas en un vector de expresión y se produjeron para la evaluación de las proteínas Env del VIH. Con la excepción de los mAbs F105, B6, 447-52d, 14e y 17b, los anticuerpos usados para el análisis son anticuerpos ampliamente neutralizantes (bNAbs). Los bNAbs pueden neutralizar múltiples cepas virales VIH. Entre los bNAbs, PGT145, PGDM1400 y PG16 son conectores del ápice y son específicos del trímero. PGT151 también es específico del trímero, pero se une en la interfase de dos protómeros de gp120 y gp41, y es dependiente del clivaje. La unión de los no bNAbs es indicativa de un plegamiento incorrecto o de una conformación de trímero abierto.
- 5 El plegamiento de la proteína también se evaluó midiendo la unión de las variantes de la proteína Env gp140 VIH soluble a un anticuerpo (mAb 17b) conocido por unirse al sitio de unión al correceptor de la proteína de la envoltura del VIH, que solamente queda expuesto después de la unión de CD4 (no se muestran los datos). En particular, se usó el receptor de CD4 soluble (sCD4) en combinación con el mAb 17 para evaluar el cambio de conformación inducido por CD4. La unión del mAb 17b a la variante de la proteína Env gp140 VIH sin unión previa de CD4 a la proteína de la envoltura es indicativo de una proteína de la envoltura parcialmente desplegada o predisparada (es decir, una Env inestable que adopta la conformación "abierta" en la ausencia de unión a CD4).
- 10 Para el ensayo AlphaLISA, se prepararon construcciones de Env gp140 VIH en el vector pcDNA2004 que contiene un conector seguido por una marca de marca de sortasa A seguido por una marca Flag seguido por un conector flexible (G₄S)₇ y termina con una marca His, (la secuencia de la marca, que se ubicó por el extremo C-terminal de la proteína Env del VIH, se provee en la SEQ ID N°: 19). Las construcciones de Env gp140 VIH se expresaron en células HEK-Expi293, que se cultivaron durante tres días en placas de 96 pocillos (200 µl/pocillo). Se diluyeron sobrenadantes crudos 120 veces en solución amortiguadora de AlphaLISA (PBS + Tween-20 0.05% + BSA 0.5 mg/ml). Para los ensayos basados en mAb 17b, los sobrenadantes se diluyeron 12 veces. A continuación, se transfirieron 10 µl de cada dilución a una placa de 96 pocillos y se mezclaron con 40 µl de esferas aceptoras, esferas donantes y uno de los mAbs enumerados precedentemente. Las esferas donantes se conjugaron con ProtA (Cat N°: AS102M, Lote N° 1831829, Perkin Elmer), que se une al mAb. Las esferas aceptoras se conjugaron con un anticuerpo anti-His (Cat N°: AL128M, Perkin Elmer), que se une a la marca His de la construcción. Para la cuantificación del rendimiento total de proteínas, que incluyen todas las formas de la proteína de la envoltura, se usó una combinación de esferas donantes conjugadas con níquel (Cat N°: AS101M, Perkin Elmer) para la detección de la marca His junto con esferas aceptoras conjugadas con un anticuerpo anti-Flag (Cat N°: AL112R, Perkin Elmer) para la detección de la marca Flag. En las pruebas donde se usa el mAb 17b en combinación con sCD4-His, se usó una combinación de esferas donantes ProtA y esferas aceptoras anti-Flag (no se muestran datos). Se mezcló una muestra con esferas donantes y aceptoras para detectar la formación de trímeros, y una segunda muestra de la misma variante Env se mezcló con esferas donantes conjugadas con níquel y esferas aceptoras conjugadas con anti-Flag para medir la cantidad total de proteínas expresadas (es decir, el rendimiento total de proteínas).
- 20 La mezcla del sobrenadante que contiene la proteína Env gp140 VIH expresada, se incubaron el mAb, las esferas donantes y esferas aceptoras a temperatura ambiente durante 2 horas sin agitación. Subsiguientemente, se midió la señal quimioluminiscente con un instrumento lector de placas Synergy NEO (BioTek). Se sustrajo la señal basal promedio atribuida a células transfectadas simuladas de las cuentas AlphaLISA medidas para cada una de las variantes de Env gp140 VIH. A continuación, el juego de datos completo se dividió por la señal medida para la proteína Env del VIH que tiene la secuencia de esqueleto ConC_SOSIP para normalizar la señal de cada una de las variantes de Env gp140 VIH evaluadas al esqueleto. Los datos de unión de cada una de las variantes de Env gp140 VIH al mAb específico de trímero PGT145 se usaron para determinar el porcentaje de formación de trímeros y el rendimiento de trímeros de cada una de las variantes. Se usó la unión a los otros mAbs para evaluar el patrón de unión general de las variantes de la Env del VIH a los bNAbs y no bNAbs (no se muestra).
- 25 El porcentaje de formación de trímeros para cada una de las variantes de la Env del VIH se calculó dividiendo la señal quimioluminiscente normalizada obtenida de la mezcla de muestras de variantes de la Env del VIH, el mAb PGT145, las esferas donantes conjugadas con ProtA y esferas aceptoras conjugadas con anti-His o la señal quimioluminiscente normalizada obtenida de la mezcla de muestras de la variante de la Env del VIH, los esferas donantes conjugadas con anti-His y esferas aceptoras conjugadas con anti-Flag.
- 30 El rendimiento de trímeros para cada una de las variantes de la Env del VIH se determinó con relación al rendimiento de trímeros para la proteína Env del VIH que comprende la secuencia de esqueleto ConC_SOSIP sin ninguna mutación adicional. La señal quimioluminiscente normalizada obtenida de la unión del mAb PGT145 a la proteína de la envoltura ConC_SOSIP se definió en 1, y la señal quimioluminiscente normalizada obtenida de la unión del mAb PGT145 a cada una de las proteínas gp140 VIH se normalizó para este valor.
- 35 Resultados del análisis del ensayo AlphaLISA: Porcentaje de trímeros y rendimientos de trímeros

El porcentaje de formación de trímeros determinado mediante el ensayo AlphaLISA para varias sustituciones de aminoácidos simples, dobles y triples de la lista de (i)-(vii) en la Tabla 1 precedente en la secuencia de esqueleto ConC_SOSIP se muestra en la Figura 2A. Entre las aproximadamente 200 variantes de Env gp140 VIH que contienen sustituciones de aminoácidos simples evaluadas, siete se identificaron posiciones de sustitución para las cuales el porcentaje de trímeros formados aumentó en por lo menos 25% con relación al porcentaje de trímeros formados para la secuencia de esqueleto ConC_SOSIP sin ninguna sustitución de aminoácido adicional.

Los resultados que se muestran en la Figura 2A demuestran que las siete posiciones de sustitución preferidas para las cuales se observó un aumento significativo en el porcentaje de formación de trímeros incluyen N651, K655, I535, D589, I573, A204 y E647 de acuerdo con la numeración en la gp160 de la forma aislada HXB2 del VIH-1. En particular, las sustituciones de aminoácidos simples que dieron como resultado el mejor porcentaje de formación de trímeros incluyeron N651F, K655I(F/W) (aunque también hubo un experimento en el cual K655F no se tradujo en una mejora), I535N, D589V(/A), I573F, A204I, E647F. Algunas mutaciones que se evaluaron en combinación con varias de estas mutaciones incluyeron K588Q/E, I556P y A558P, y éstas mejoraron aún más el porcentaje de trímeros de mutantes con aminoácidos preferidos en las posiciones de la invención (i)-(vii) de la Tabla 1) en este experimento.

Todas las sustituciones dobles evaluadas en este experimento mostraron un mayor porcentaje de formación de trímeros que las correspondientes sustituciones simples, y todas las sustituciones triples evaluadas mostraron un mayor porcentaje de formación de trímeros que las correspondientes mutaciones simples y dobles (Figura 2A). Estos resultados impredecibles y sorprendentes indican que estas mutaciones podían mostrar una forma de sinergia en estos experimentos con respecto a la trimerización de la proteína de la envoltura.

Además del mayor porcentaje de formación de trímeros, también resulta deseable un mayor rendimiento de trímeros. Por ello, también se determinó el rendimiento de trímeros de las variantes de gp140 VIH que contiene mutaciones simples, dobles y triples en la secuencia de esqueleto ConC_SOSIP mediante el ensayo AlphaLISA. Los resultados se muestran en la Figura 2B. La mayoría de las variantes de gp140 VIH que contienen mutaciones simples (con excepción de I535N, D589A y D589I), mostró un mayor rendimiento de trímeros que la proteína de la envoltura ConC_SOSIP. Sin embargo, el análisis SEC-MALS más preciso del mutante I535N, que se describirá abajo, mostró un aumento en el rendimiento de trímeros. Además, mutaciones adicionales en combinación con I535N, tal como D589V, dio como resultado el mismo rendimiento de trímeros que el observado para la proteína de la envoltura que comprende dicha sustitución adicional particular en ausencia de la mutación I535N. El rendimiento de trímeros de las variantes con mutaciones dobles también aumentó, donde cada una de las variantes de mutación simple mostró un mayor rendimiento de trímeros que la proteína de la envoltura ConC_SOSIP (Figura 2B).

También se evaluó el porcentaje de formación de trímeros de las variantes de gp140 VIH con mutaciones dobles en el esqueleto ConC_SOSIP como se había descrito previamente en la literatura, que incluyen la sustitución doble E64K, T316W descrita por (De Taeye et al., *supra*), y la sustitución doble disulfuro I204C, A433C descrita por (Kwon et al., *supra*). Las sustitución doble E64K, T316W dio como resultado un menor porcentaje de formación de trímeros que la proteína de la envoltura ConC_SOSIP, es decir, un 15% (no se muestran los datos). Aunque la sustitución doble disulfuro I204C, A433C aumentó el porcentaje de trímeros a un 43% (no se muestran los datos), las nuevas sustituciones dobles que se describen en la presente, tales como I535N/K588E, K588Q/D589V, K655I/K588E, I535N/D589V, I535N/E647F, D589V/K655I, e I535N/K655I (FIG. 2A) dieron como resultado un porcentaje aún mayor de formación de trímeros en el experimento AlphaLISA.

También se introdujeron mutaciones adicionales (prolina en los residuos 558 y/o 556) en el esqueleto ConC_SOSIP, y se midió el porcentaje de formación de trímeros y de rendimiento de trímeros de estas proteínas Env gp140 VIH. Tanto las sustituciones simples de Pro en la posición 558 o 556, y la sustitución doble de prolina en ambas posiciones 556 y 558 además de las mutaciones SOSIP ya contenidas en el esqueleto ConC_SOSIP (es decir, Cys en las posiciones 501 y 605, y Pro en la posición 559) aumentaron el porcentaje de formación de trímeros y el rendimiento de trímeros (no se muestran los datos). De hecho, la introducción de una o más de las nuevas sustituciones de aminoácidos estabilizantes de la invención en el esqueleto ConC_SOSIP que además comprende residuos Pro en las posiciones 558 y/o 556 también aumenta el porcentaje de formación de trímeros y/o el rendimiento de trímeros (por ejemplo Figura 2A, por ejemplo A558P/I535N, K655I/L556P, y varios mutantes triples que incluyen la mutación A558P).

Los datos de unión de las variantes de Env gp140 VIH a los otros bNAbs y no bNAbs demostraron que la mayoría de las mutaciones simples, dobles y triples evaluadas que aumentaban el rendimiento de trímeros y el porcentaje de formación de trímeros, tales como los indicados en las Figuras 2A y 2B, también mostraron una mayor unión a los bNAbs, y la misma o menor unión a los no bNAbs con relación a la cantidad de unión observada a los bNAbs y no bNAbs para la proteína de la envoltura del VIH que comprende la secuencia de esqueleto ConC_SOSIP (no se muestran los datos). Para el desarrollo de vacunas, se prefiere una mayor unión a los bNAbs y una menor unión a los no bNAbs. Por consiguiente, los datos demuestran que las proteínas de la envoltura del VIH que comprenden las sustituciones de aminoácidos en las posiciones

(i)-(vii) indicadas en la Tabla 1 precedente tienen propiedades deseables con respecto a los patrones de unión a los anticuerpos ampliamente neutralizantes y no ampliamente neutralizantes.

Análisis SEC-MALS

El análisis SEC-MALS también se usó para verificar el rendimiento de trímeros y el porcentaje de formación de trímeros de las variantes de gp140 VIH seleccionadas usando el ensayo AlphaLISA. Las variantes de gp140 VIH se expresaron en cultivos a escala de 30 ml y se purificaron mediante aplicación de sobrenadantes libre de células sobre 200 µl de esferas de lectina de *Galanthus nivalis* (Vectorlab Cat N° AL-1243) en columnas de flujo gravimétrico Poliprep (Biorad Cat N° 731-1550). Las esferas se lavaron con 2 ml de solución amortiguadora de unión (Tris 40 mM, NaCl 500 mM, pH 7.4). Las proteínas se eluyeron usando 250-500 µl de Tris 40 mM, NaCl 500 mM, manopiranosido 1 M, pH 7.4. Se usó un sistema de cromatografía líquida de alta presión (Agilent Technologies) y el instrumento MiniDAWN TREOS (Wyatt) acoplado a un detector de índice de refracción Optilab T-REX (Wyatt) en el experimento SEC-MALS. En total se aplicaron ya sea 100 µl de elución de lectina o aproximadamente 30 µg de proteína a una columna TSK-Gel G3000SWxl (Tosoh Bioscience) equilibrada en solución amortiguadora de corrida (fosfato de sodio 150 mM, NaCl 50 mM, pH 7.0) a razón de 1 ml/min. Los datos se analizaron usando el paquete de software Astra 6, y los cálculos de peso molecular derivaron de la señal del índice de refracción.

Los cromatogramas SEC-MALS de la proteína de la envoltura ConC_SOSIP y de las variantes de gp140 VIH que contienen mutaciones simples se muestran en la Figura 3. En general, los resultados obtenidos del análisis SEC-MALS eran comparables y consistentes con los resultados obtenidos con el análisis AlphaLISA. El cromatograma de la proteína de la envoltura ConC_SOSIP tiene cuatro picos principales, donde el segundo pico que eluyó aproximadamente a los 7.3 minutos es el pico del trímero. Se determinó que la proteína de la envoltura ConC_SOSIP era aproximadamente un 27% trimérica. La formación de agregados y monómeros indica que hay algo de mal plegamiento e inestabilidad asociados con la proteína Env gp140 VIH que comprende la secuencia consenso ConC_SOSIP. Como lo demuestran los cromatogramas que se muestran en la Figura 3, todas las sustituciones simples dieron como resultado un pico de trímero relativamente mayor en comparación con el pico del trímero correspondiente a la proteína de la envoltura ConC_SOSIP, lo que indica que el rendimiento de trímeros había aumentado para cada una de las variantes de gp140 VIH.

En su conjunto, los resultados demuestran que las sustituciones de aminoácidos identificados en (i)-(vii) de la Tabla 1 en la presente proporcionan proteínas Env del VIH recombinantes con un mayor porcentaje de formación de trímeros y/o un rendimiento de trímeros mejorado. En particular, las variantes de la proteína Env del VIH que comprenden múltiples sustituciones en las posiciones identificadas de (i)-(vii) de la Tabla 1, tales como combinaciones de dos o más de las mutaciones identificadas, típicamente mostraron un rendimiento de trímeros aún más mejorado y/o un mayor porcentaje de formación de trímeros ante las variantes de la proteína Env del VIH que sólo comprende una mutación simple, que muestra un posible efecto sinérgico de las mutaciones combinadas (i)-(vii) de la Tabla 1. Según el mejor conocimiento de los inventores, no se ha informado de la presencia de estas combinaciones de sustituciones de aminoácidos en las secuencias de la proteína de la envoltura del VIH natural, y por lo tanto se considera que todas las combinaciones (entre (i)-(vii)) son combinaciones nuevas de mutaciones estabilizantes del trímero. Las proteínas de la envoltura del VIH que tienen un mayor porcentaje de formación de trímeros, tales como las proteínas de la envoltura del VIH recombinantes de la invención, son ventajosas desde la perspectiva de elaboración, tal como para vacunas, porque la purificación y eliminación de la proteína de la envoltura presente en la preparación en la conformación no nativa indeseadas será menos necesarias. Además, un mayor rendimiento de expresión total del trímero es ventajoso para la elaboración de una vacuna.

Ejemplo 4: Estabilidad de las proteínas de la envoltura del VIH triméricas

La estabilidad térmica de las proteínas Env del VIH recombinantes de acuerdo con las formas de realización de la invención se evaluó mediante AlphaLISA y calorimetría de barrido diferencial (DSC).

Mediciones de la estabilidad térmica usando AlphaLISA

La estabilidad al calor se evaluó midiendo la pérdida de trímeros intactos tras el tratamiento con calor basado en la unión al mAb específico del trímero, PGT145. El sobrenadante crudo (20 µl) se calentó a 60 °C durante 1 hora. A continuación, se centrifugaron las muestras a la velocidad máxima durante cinco minutos para eliminar los agregados. El ensayo AlphaLISA se realizó como se describió antes en el Ejemplo 3.

Los resultados se muestran en la Figura 4, y los datos se informan como el porcentaje de trímeros que quedaron intactos después del tratamiento con calor. A partir de los resultados, se puede observar que la mayoría de las proteínas Env gp140 VIH recombinantes mutantes simples de la invención evaluadas mostraban una mayor estabilidad al calor que la proteína de la envoltura ConC_SOSIP. También se encontró que las proteínas de la envoltura del VIH que comprenden las sustituciones dobles y triples estabilizantes de trímeros que se identifican en la presente y que fueron evaluadas tienen una mayor estabilidad al calor que la proteína de la envoltura ConC_SOSIP.

Mediciones de la estabilidad térmica usando DSC

5 Se determinó la temperatura de fusión (T_m) de las variantes de Env gp140 VIH mediante DSC usando un sistema de DSC capilar MicroCal. Cada medición se efectuó con una temperatura inicial de 20 °C y una temperatura final de 110 °C a una velocidad de barrido de 100 °C/hora. Se usó una muestra de proteínas a una concentración de 0.5 mg/ml (400 μ l) para cada medición. Los datos se analizaron usando el software Origin J. (herramienta de análisis MicroCal VP).

10 Se determinó que la temperatura de fusión (T_m) de la proteína de la envoltura ConC_SOSIP medida usando DSC era de 69.8 °C y la temperatura de inicio de la fusión era de 60.1 °C. La T_m medida para la proteína de la envoltura ConC_SOSIP era mayor que la correspondiente a la proteína de la envoltura BG505_SOSIP (proteína de la envoltura del VIH de la forma aislada viral BG505 que comprende las denominadas mutaciones SOSIP), con una T_m informada de 67.0 °C (Kwon et al, 2015). Esto indica que una proteína de la envoltura del VIH que comprende la secuencia de esqueleto ConC_SOSIP tiene propiedades más favorables con respecto a la estabilidad al calor que otra secuencia de la envoltura del VIH conocida con mutaciones estabilizantes del trímero.

15 Se midió la T_m del mutante K655I de ConC_SOSIP y era de 72.3 °C y la temperatura de inicio de la fusión era de 63.7 °C, que es aún más alta que la T_m de la proteína de la envoltura ConC_SOSIP. Se midió la T_m del mutante A558P, N651F, I535N de ConC_SOSIP y era de 77.29 °C con una temperatura de inicio de 74.87 °C. Por consiguiente, los resultados de la DSC confirman los resultados de estabilidad al calor determinados con el ensayo AlphaLISA.

20 En su conjunto, los resultados demuestran que las proteínas Env del VIH que comprenden por lo menos una de las sustituciones de aminoácidos que se describen en la presente típicamente tienen una mayor estabilidad al calor que las proteínas de la envoltura que carecen de dichas mutaciones. Los resultados también demuestran que todas las variantes de la proteína Env del VIH de sustitución doble presentaban una mayor estabilidad al calor que la proteína de la envoltura ConC_SOSIP. Las variantes de la proteína Env del VIH de sustitución triple también eran más estables que la proteína de la envoltura ConC_SOSIP.

Ejemplo 5: Variantes de proteínas de la envoltura del VIH recombinantes basadas en una secuencia consenso de una proteína de la envoltura del clado B

25 Se generaron proteínas Env del VIH recombinantes de acuerdo con las formas de realización de la invención que comprenden una sola sustitución de aminoácido (I535N, D589V, N651F o K655I) introducida en la secuencia consenso ConB_SOSIP del clado B (SEQ ID N°: 5) y se purificaron como se describe en el Ejemplo 2. El rendimiento de trímeros y el porcentaje de formación de trímeros se midieron mediante un ensayo AlphaLISA como se describe en el Ejemplo 3.

30 Los resultados se muestran en la Figura 5A (porcentaje de formación de trímeros) y en la Figura 5B (rendimiento de trímeros). Los valores informados son relativos al valor medido para la proteína de la envoltura ConB_SOSIP, que se definió en 1 tanto para el porcentaje de formación de trímeros como para el rendimiento de trímeros. Los resultados muestran que todas las mutaciones evaluadas aumentaron el porcentaje de formación de trímeros. El rendimiento de trímeros era aproximadamente igual o mayor con relación a la proteína de la envoltura ConB_SOSIP para todas las mutaciones evaluadas.

35 Estos resultados demuestran que estas mutaciones también tenían un efecto estabilizante sobre la proteína de la envoltura, por ejemplo, un rendimiento de trímeros mejorado, un mayor porcentaje de formación de trímeros, etc., cuando se introducen en una secuencia de esqueleto de la proteína de la envoltura del VIH diferente, en este caso una secuencia consenso derivada del clado B.

Ejemplo 6: Variantes de proteínas de la envoltura del VIH recombinantes basadas en una secuencia sintética de una proteína de la envoltura

40 Se prepararon proteínas Env del VIH recombinantes de acuerdo con las formas de realización de la invención que comprenden sustituciones de aminoácidos introducidas en una proteína de la envoltura del VIH sintética (denominada 'DS_sC4_SOSIP_E166R') que comprende la secuencia que se muestra en la SEQ ID N°: 7 y se purificaron como se describe en el Ejemplo 2. La proteína de la envoltura del VIH sintética, DS_sC4_SOSIP_E166R, comprende las denominadas mutaciones SOSIP (Cys en los residuos 501 y 605, y Pro en el residuo 559), Cys en los residuos 201 y 433 que dan como resultado la introducción de un enlace disulfuro (DS) y Arg en la posición 166 para estabilizar el ápice. Además, la proteína está truncada en la posición 655. El porcentaje de formación de trímeros y el rendimiento de trímeros se midieron mediante un ensayo AlphaLISA como se describe en el Ejemplo 3.

50 Los resultados se muestran en la Figura 6, donde se compara el porcentaje de formación de trímeros de cada una de las variantes evaluadas con el porcentaje de formación de trímeros (Fig 6A) y el rendimiento de trímeros (Fig 6B) del esqueleto DS_sC4_SOSIP_E166R. Se observó un mayor porcentaje de formación de trímeros para cada una de las variantes evaluadas en comparación con la secuencia del esqueleto.

Además de E166R, se cambiaron algunos otros aminoácidos que aparecen raramente por algunos más prevalentes en la posición correspondiente en una colección de proteínas Env del VIH de tipo salvaje (A114Q, E117K, T375S y I434M), para 'repara' la proteína de acuerdo con un marco que se explicará con mayor detalle más adelante en el Ejemplo 12 y en la Figura 12. En esta proteína 'reparada', las mutaciones estabilizantes A204I, y K655I mejoran aún más la sC4_SOSIP (Fig.13).

Los resultados de este Ejemplo son coherentes con los del Ejemplo 5 en demostrar que las mutaciones descritas en la presente también tienen un efecto estabilizante sobre la proteína de la envoltura, por ejemplo, un mayor porcentaje de formación de trímeros y/o un rendimiento de trímeros mejorado, etc., cuando se introducen en secuencias de esqueleto de la proteína de la envoltura del VIH diferentes, en este caso una secuencia Env no consenso, sintética.

10 **Ejemplo 7: Otras combinaciones de mutaciones de la Env del VIH**

Se prepararon proteínas Env del VIH recombinantes de acuerdo con las formas de realización de la invención que comprenden sustituciones de aminoácidos introducidas en ConC_SOSIP (que comprende la secuencia que se muestra en la SEQ ID N°: 3) y se purificaron como se describe en el Ejemplo 2. El porcentaje de formación de trímeros se midió mediante un ensayo AlphaLISA como se describe en el Ejemplo 3. Subsiguientemente, se purificó una selección menor de combinaciones (las que se muestran más adelante en cursiva y adicionalmente K655I; I535N, D589V; I535N, K655I; D589V, K655I) usando lectina de *Galanthus nivalisy* se analizó el contenido de trímeros usando SEC-MALS como se describe en el Ejemplo 3. La estabilidad se midió como se describe en el Ejemplo 4.

Para este experimento se prepararon los siguientes mutantes:

- K655I, N651F;*
- 20 *K655I, N651F, E647F;*
K655I, N651F, E647F, I535N;
K655I, N651F, I535N;
K655I, I573F;
K655I, D589V, I573F;
- 25 *K655I, D589V, I573F, N651F;*
K655I, D589V, I573F, K588E;
K655I, D589V, I573F, N651F, K588E;
K655I, D589V, I573F, N651F, K588E, I535N;
K655I, D589V, I573F, N651F, K588E, I535N, A204I;
- 30 *K655I, D589V, I535N, L556P;*
K655I, D589V, I573F, N651F, K588E, L556P;
K655I, D589V, A204I;
L556P, N651F;
L556P, N651F, K655I;
- 35 *L556P, N651F, K655I, I535N;*
L556P, N651F, K655I, I535N, I573F;
L556P, N651F, K655I, I535N, I573F, D589V;
L556P, N651F, K655I, I535N, I573F, D589V, A204I;
L556P, N651F, K655I, I535N, I573F, D589V, A204I, K588Q;
- 40 *L556P, N651F, K655I, I535N, I573F, D589V, A204I, K588Q, E647F;*

L556P, N651F, I535N;

L556P, N651F, I535N, I573F;

L556P, N651F, I535N, I573F, D589V;

L556P, N651F, I535N, I573F, D589V, A204I;

5 L556P, N651F, I535N, I573F, D589V, A204I, K588Q;

L556P, N651F, I535N, I573F, D589V, A204I, K588Q, E647F;

L556P, K655I, I535N;

L556P, K655I, I535N, I573F;

L556P, K655I, I535N, I573F, D589V;

10 L556P, K655I, I535N, I573F, D589V, A204I;

L556P, K655I, I535N, I573F, D589V, A204I, K588Q;

L556P, K655I, I535N, I573F, D589V, A204I, K588Q, E647F;

L556P, N651F, I535N, I573F, D589V, A204I, K588Q con eliminación de la mutación SOS.

15 Todas las combinaciones de sustituciones evaluadas en el esqueleto ConC_SOSIP mostraron un mayor porcentaje de trímeros, un mayor rendimiento de trímeros y una mayor estabilidad de trímeros a 60 °C en comparación con el esqueleto en AlphaLISA (no se muestran los datos). El análisis SEC_MALS confirmó el mayor porcentaje de trímeros de todas las mutaciones en el esqueleto evaluadas (no se muestran los datos).

20 En el caso de un conjunto que comprende mutaciones adicionales de una en una hasta un total de nueve mutaciones, el análisis SEC-MALS mostró que con la introducción de cada mutación siguiente aumentaba la relación de trímero/monómero, a medida que disminuía la altura del pico del monómero, en tanto la altura del pico del trímero permaneció igual en el gráfico SEC (Figura 7). Entre todas las variantes evaluadas por SEC-MALS, la variante con las sustituciones L556P, N651F, K655I, I535N, I573F, D589V, A204I, K588Q, E647F mostró el porcentaje de trímeros más alto (menos monómeros de gp140 y menos monómeros de gp120), el mayor rendimiento total de proteínas y una de las mayores estabildades a la temperatura. Esto significa que estas mutaciones se pueden combinar sin pérdida de trímeros

25 en comparación con el esqueleto. Además, esto sugiere que, en general, la adición de las mutaciones que se describen en (i)-(vii) de la Tabla 1, opcionalmente combinadas con las mutaciones descritas en la Tabla 2, da como resultado una trimerización aún mayor.

30 También se evaluó una construcción con las mutaciones L556P, N651F, I535N, I573F, D589V, A204I, K588Q, en las cuales se habían eliminado las 'mutaciones SOS' (es decir los dos residuos cisteína en las posiciones 501 y 605 se habían revertido en los residuos de aminoácidos presentes originalmente en la secuencia consenso del clado C). Aunque su estabilidad a la temperatura era menor, el mutante tenía un porcentaje y rendimiento de trímeros comparables ya que su mutante correspondiente comprendía la mutación SOS. El mutante en el cual se había eliminado la mutación SOS aún ofrecía una ventaja por cuanto se unía menos a los no bNAbs que su contraparte correspondiente que contiene SOS (que comprende las mutaciones L556P, N651F, I535N, I573F, D589V, A204I, K588Q). Esto demuestra que las propiedades

35 ventajosas de la presente invención, tal como un elevado porcentaje de trimerización, también se pueden obtener con proteínas Env del VIH que no comprenden todas las mutaciones SOSIP.

Un mutante (evaluado en el esqueleto ConC_SOSIP), basado en una combinación de propiedades favorables en cuanto a nivel de expresión, formación de trímeros y unión al anticuerpo ampliamente neutralizante PGT151, comprende las siguientes mutaciones: L556P, N651F, I535N, I573F, D589V, A204I, K588Q.

40 En el antecedente ConC_SOSIP, las 9 sustituciones más exitosas eran L556P, E647F, N651F, K655I, I535N, D589V, I573F y K588E en gp41 y A204I en gp120. La combinación de estas 9 sustituciones condujo a una mayor estabilidad, contenido de trímeros y rendimiento de trímeros. Dado que la adición de L556P en esta variante con 9 sustituciones tenía un efecto relativamente limitado sobre el mayor porcentaje de trímeros, y la sustitución E647F en este contexto parecía obstaculizar la unión a PGT151, estas dos mutaciones no se usaban siempre en otras variantes, y se encontró que una

45 variante con 7 sustituciones (denominada ConC_SOSIP_7mut, a veces también denominada 'ConC_SOSIP estabilizada' o 'ConC_base' en la presente; que incluye N651F, K655I, I535N, D589V, I573F, K588E y A204I) era ligeramente más

estable (mayor temperatura de fusión) que la variante con las 9 sustituciones indicada precedentemente. La secuencia completa de esta variante (ConC_SOSIP Env estabilizada, VIH 160544) se provee en la SEQ ID N°: 20.

5 En este momento, un mutante particularmente preferido [evaluado en el esqueleto ConC_SOSIP con las siguientes mutaciones adicionales: (a) D279N, A281V, A362Q (aumentan la similitud con los virus fundadores transmitidos, como describieron otros); (b) Del139-152 (supresión de un bucle variable para reducir para probabilidad de inducir anticuerpos contra este bucle); y (c) V295N (introducción de un sitio de glicano que está presente en la mayoría de las cepas del VIH)], basado en una combinación de propiedades favorables en cuanto a nivel de expresión, formación de trímeros y unión a un anticuerpo ampliamente neutralizante, comprende las siguientes mutaciones estabilizantes de la invención: N651F, K655I, I535N, I573F, D589V, A204I, K588E. La secuencia completa de esta variante (ConC_SOSIP.v3 Env estabilizada
10 (HIV170654, ConC_SOSIP.v3)) se provee en la SEQ ID N°: 28.

En una variante adicional, se agregó una mutación K658V a esta construcción (véase también el ejemplo 15 más adelante), que mejoró aún más los resultados.

Ejemplo 8: Partículas autoensamblantes que presentan la proteína Env del VIH estabilizada

15 Se prepararon partículas autoensamblantes de DPS y ferritina que presentan proteínas Env estabilizadas de manera similar a lo descrito en (He et al, 2016). Para tal fin, se fusionó la proteína gp140 al extremo N-terminal de las partículas por medio de un conector de aminoácidos corto (por ejemplo GSG o AAAGS, pero también se pueden usar otros conectores, véase por ejemplo He et al, 2016) a nivel de ADN y la proteína de fusión se expresó en células Expi293F. Un ejemplo de una partícula que se preparó de esta manera se basó en ferritina fusionada a una proteína Env del VIH ConC_SOSIP (SEQ ID N°: 3) con las siguientes mutaciones: I535N, A558P, D589V, K655I. También se prepararon
20 partículas de ferritina con esta proteína Env que comprende una mutación V570D adicional, que según se ha informado mejoran la trimerización (Kesavardhana et al, 2014), pero se observó que esta mutación conduce a un fuerte aumento de la unión a un anticuerpo no neutralizante (17b), lo cual resulta indeseable. También se fusionó una Env con estas cinco mutaciones a dos tipos de partículas de DPS, de Helicobacter pylori y de Mycobacterium smegmatis (véase, por ejemplo, WO2011/082087 por la preparación de las partículas DPS). También se fusionó una Env con estas cinco mutaciones y
25 además un puente disulfuro que introduce la mutación doble I201C-A433C a la ferritina.

Las partículas se purificaron a partir de un sobrenadante libre de células con esferas de afinidad PGDM140 y las partículas se analizaron usando SEC-MALS con una columna TSKgel G6000PWCL. Un análisis con SEC-MALS, así como PAGE Nativa (3-12%), confirmó que se formaron partículas de los tamaños aproximadamente esperados.

30 De manera similar, también se prepararon nanopartículas de ferritina y DPS autoensamblantes que presentan una Env del VIH que comprende una secuencia ConC_SOSIP con la siguiente combinación de mutaciones: (L556P, N651F, I535N, I573F, D589V, A204I, K588Q).

Se prepararon asimismo otros liposomas y/o nanopartículas autoensamblantes que presenta otras variantes de la Env del VIH descritas en la presente, por ejemplo la Env del VIH de la SEQ ID N°: 20, 22, 24, 26, 27, 28, 29, 30, 31, o 32.

35 **Ejemplo 9: Variantes de la proteína de la envoltura del VIH recombinantes basadas en una secuencia de la proteína de la envoltura del Clado A**

Se introdujeron proteínas Env del VIH recombinantes de acuerdo con las formas de realización de la invención que comprenden sustituciones de aminoácidos simples (I535N, D589V, N651F, K655I, I573F, A204I o E647F) en una proteína de la envoltura del clado A del VIH de tipo salvaje con la modificación SOSIP (denominada 'BG505_SOSIP') como se describe en el Ejemplo 2. La proteína de la envoltura del VIH, BG505_SOSIP, comprende las denominadas mutaciones
40 SOSIP (Cys en los residuos 501 y 605, y Pro en el residuo 559), así como Cys adicionales en los residuos 201 y 433 lo que da como resultado la introducción de un enlace disulfuro (DS), y un sitio potencial de N-glicosilación en la posición 332 (mutación T332N). La proteína está truncada en la posición 664. La secuencia de BG505_SOSIP se muestra en la SEQ ID N°: 21.

45 El porcentaje de formación de trímeros y el rendimiento de trímeros se midieron mediante un ensayo AlphaLISA como se describe en el Ejemplo 3. El porcentaje de formación de trímeros y el rendimiento de trímeros de cada una de las variantes evaluadas se compararon con BG505_SOSIP. Se observó un mayor porcentaje de formación de trímeros para las sustituciones M535N, D589V, N651F o K655I en comparación con la secuencia del esqueleto (por ejemplo, la Figura 8A). La combinación, por ejemplo, de L556P, K655I y M535N mostró un rendimiento y porcentaje de trímeros aún mayor (por ejemplo, las Figuras 8A y 8B). La combinación de N651F y D589V mejoró aún más el rendimiento y el porcentaje de
50 trímeros (no se muestran los datos). Los resultados de este ejemplo para un virus del clado A son coherentes con los de los ejemplos 10 y 11 (clado C) más adelante y del ejemplo 5 (clado B), en donde las mutaciones I535N, D589V, N651F y K655I también mostraron un efecto estabilizante sobre la proteína de la envoltura derivada de cepas de tipo salvaje, por ejemplo, un mayor porcentaje de formación de trímeros y/o un rendimiento de trímeros mejorado. Claramente, estas

mutaciones de la invención también mejoran la trimerización de la Env del VIH derivada de una cepa de tipo salvaje del clado A.

En este momento, un mutante particularmente preferido (evaluado en el esqueleto BG505_SOSIP, basado en una combinación de propiedades favorables en cuanto a nivel de expresión, formación de trímeros y unión a un anticuerpo ampliamente neutralizante, es aquel que comprende las siguientes mutaciones: L556P, K655I, M535N, N651F, D589V, (véase, por ejemplo, la Figura 9 donde se muestra una formación de trímeros fuertemente mejorada de dicho mutante según un análisis SEC-MALS, y en Figura 13, donde se muestra una unión claramente mejorada de anticuerpos ampliamente neutralizantes para dicho mutante). La secuencia de esta BG505_SOSIP Env estabilizada (HIV170863) se muestra en la SEQ ID N°: 22.

10 La adición de la mutación Q658V proporcionó una pequeña mejora adicional.

Una construcción preferida adicional contiene las mutaciones L556P, K655I, M535N, N651F, D589V, así como las mutaciones 'DS' (Cys en las posiciones 201 y 433 que dan como resultado la introducción de un enlace disulfuro), R588E y Q658V. La secuencia de dicha variante (BG505_SOSIP.v2 Env, HIV171814) se provee en la SEQ ID N°: 29.

15 **Ejemplo 10: Variantes de la proteína de la envoltura del VIH recombinantes basadas en una secuencia de la proteína de la envoltura de tipo salvaje del clado C**

Se generaron proteínas Env del VIH recombinantes de acuerdo con las formas de realización de la invención que comprenden la sustitución de aminoácido simple T651F, la sustitución de aminoácidos doble T651F, M535N introducidas en una secuencia WT C97ZA_SOSIP Env (SEQ ID N°: 23) con la sustitución adicional L556P (C97ZA_SOSIP_L556P) y se expresaron como se describe en el Ejemplo 2. El rendimiento de trímeros y el porcentaje de formación de trímeros se midieron mediante un ensayo AlphaLISA como se describe en el Ejemplo 3.

Los resultados se muestran en las Figuras 10A y B. El rendimiento de trímeros de C97ZA_SOSIP_L556P_T651F_M535N son cinco veces mayor el correspondiente al esqueleto C97ZA_SOSIP.

Por consiguiente, las sustituciones L556P, T651F y M535N proporcionaban una gran mejora de C97ZA_SOSIP, pero la unión a bNAbs y el porcentaje de trímeros para esta variante derivada de un tipo salvaje del clado C aún era mucho menor que para el esqueleto ConC_SOSIP. Dado que una wt Env se puede adaptar a su huésped, posiblemente reduciendo su aptitud general, y de esa manera se puede corromper el plegamiento, la secuencia Env se 'reparó' de acuerdo con el marco conceptual que se describirá más adelante en el Ejemplo 12 y en la Figura 12. Se cambió un total de 21 residuos, para reparar la secuencia, y se agregaron tres sitios potenciales de N-glicosilación (PNGS) para llenar los denominados "huecos de glicano" (posiciones donde en por lo menos el 50% de las proteínas Env de cepas de tipo salvaje el VIH hay un sitio potencial de N-glicosilación). Las mutaciones introducidas según este marco para C97ZA_SOSIP se indican en la Tabla 3 en la columna de 'mutaciones de reparación'. La adición de la mutación estabilizante K655I divulgada en la presente aumentó aún más el porcentaje y el rendimiento de trímeros, al igual que D589V, A204I y K588E.

Estos resultados demuestran que las mutaciones T651F, M535N y K655I, D589V, A204I y K588E descritas en la presente también tenían un efecto estabilizante sobre la proteína de la envoltura, por ejemplo, un rendimiento de trímeros mejorado, un mayor porcentaje de formación de trímeros, cuando se introducen en C97ZA_SOSIP (derivada de una proteína Env de una cepa de tipo salvaje del clado C) y variantes de la misma.

En este momento, una variante particularmente preferida (evaluado en el esqueleto C97ZA_SOSIP, basado en una combinación de propiedades favorables en cuanto a nivel de expresión, formación de trímeros y unión a un anticuerpo ampliamente neutralizante, es aquel que comprende las siguientes mutaciones: Q567K (descrita anteriormente por otros); A198T, S243N, K236T, V295N (para llenar los huecos de glicano); M34L, T46K, T58A, Q171K, G172V, P179L, L183Q, I192R, N209T, M307I, Q350R, N352H, Y353F, D412N, G429E, V455T, I489V, L491I, G500K, S547G, T578A, T651N (para reparar la secuencia); V505N, E507T, T663N (sitios potenciales de N-glicosilación agregados a la base de la molécula); y A204I, M535N, L556P, K588E, D589V, T651F, K655I (mutaciones estabilizantes de la invención). Los datos de esta variante se muestran, por ejemplo, en la Figura 13, véase en particular la 'C97ZA estabilizada y reparada' en la misma), que muestra un gran aumento de la unión a un anticuerpo ampliamente neutralizante en comparación con la molécula wt C97ZA Env original. La secuencia de esta variante (C97ZA_SOSIP Env estabilizada y reparada (HIV170690)) se provee en la SEQ ID N°: 24.

La adición de la mutación K658V estabilizó esta proteína aún más.

Una variante preferida adicional incluye la mutación 'DS' y K658V, y la secuencia de esta variante (C97ZA_SOSIP.v2 Env, HIV171810) se provee en la SEQ ID N°: 30.

Ejemplo 11: Variantes de la proteína de la envoltura del VIH recombinantes basadas en otra secuencia de la proteína de la envoltura de tipo salvaje del clado C

En la proteína Env de la cepa Du422 del clado C, se introdujeron mutaciones SOSIP y se rellenaron dos huecos de glicano en la posición 295 y 386 con las mutaciones K295N y D386N. Además, se repararon algunos residuos de acuerdo con el marco conceptual que se describe en el Ejemplo 12 y en la Figura 12 (V272I, W456R, G466E y F643Y), y se introdujeron las sustituciones estabilizantes L556P, I535N, N651F y D589V. Todas las sustituciones adicionales dieron como resultado mayores rendimientos de trímeros y porcentajes de trímeros (por ejemplo, la Figura 11).

En una variante específica evaluada con estas cuatro mutaciones estabilizantes (SEQ ID N°: 25), la sustitución adicional K655I aumentó aún más el rendimiento de trímeros y el porcentaje de trímeros en un factor de 1.3 y 1.4, respectivamente (no se muestran los datos).

En este momento, una variante Du422_SOSIP Env particularmente preferida, basada en una combinación de propiedades favorables en cuanto a nivel de expresión, formación de trímeros y unión a un anticuerpo ampliamente neutralizante, es aquel que comprende las siguientes mutaciones: L556P, K655I, M535N, N651F, D589V, K588E, I201C, A433C, V272I, W456R, G466E, F643Y, D386N y K295N. La secuencia de esta variante (Du422_SOSIP Env (HIV170859) estabilizada y reparada se provee en la SEQ ID N°: 26. Los datos de esta variante se muestran, por ejemplo, en la Figura 13 (véase la Du422 estabilizada y reparada en la misma), que muestra un gran aumento de la unión a un anticuerpo ampliamente neutralizante en comparación con la molécula wt Du422 Env original.

Otra variante preferida comprende adicionalmente la mutación 'DS' y K658V, y la secuencia de esta variante (Du422_SOSIP.v1 Env, HIV171812) se provee en la SEQ ID N°: 31.

Ejemplo 12: Reparación y estabilización de varias secuencias de Env VIH-1

Dado que las secuencias wt [de tipo salvaje] de virus aislados de pacientes infectados pueden haber adquirido mutaciones desestabilizantes que impiden un plegamiento correcto, primero se repararon las secuencias wt Env del clado C, C97ZA, DU422, y el mosaico sC4.

Para buscar mutaciones no óptimas en las secuencias de tipo salvaje, se efectuó un alineamiento de todas las secuencias Env VIH-1 de la base de datos UniProt y la base de datos de VIH Los Alamos (~90.000 secuencias) y se calculó la distribución de aminoácidos para cada aminoácido. En general, se sustituyó una cantidad de aminoácidos relativamente raros en las secuencias wt Env por aminoácidos más comunes (basado en la frecuencia para la posición correspondiente en la base de datos) de acuerdo con el marco conceptual que se describe en la Figura 12.

Además, se introdujeron dos sustituciones adicionales, Y353F y Q171K, en el ápice de C97ZA_SOSIP para mejorar posiblemente la unión de los anticuerpos dirigidos al ápice, y se introdujeron sitios de glicano adicionales mediante la sustitución de D411N, K236T y V295N porque estos sitios potenciales de N-glicosilación (PNGS) estaban conservados en >50%. A continuación, se transfirieron las sustituciones estabilizantes que se describieron en los ejemplos previos a la secuencia reparada.

La ConC_SOSIP estabilizada contiene las sustituciones A204I, I535N, I573F, K588E, D589V, N651F y K655I (ConC_SOSIP estabilizada). La secuencia completa de la ConC_SOSIP estabilizada se provee en la SEQ ID N°: 20.

En la Tabla 3 se provee un resumen de algunas de las variantes de proteínas Env y sus mutaciones.

Tabla 3 Variantes de la proteína Env del VIH.

Proteínas	mutaciones de la literatura	PNGS agregados	secuencia directriz (SEQ ID NO:)	mutaciones de reparación	mutaciones estabilizantes	otras mutaciones	terminal
ConC_SOSIP		V295N	11				664
ConC_SOSIP estabilizada		V295N	11		A204I, I535N, I573F, K588E, D589V, N651F, K655I		664
BG505_SOSIP		T332N	34				664

ES 2 824 525 T3

BG505_SOSIP estabilizada		T332N	34		M535N, L556P, D589V, N651F, K655I		664
C97ZA_SOSIP	L535M, Q567K		nativa				664
C97ZA_SOSIP reparada	L535M, Q567K	A198T, S243N, K236T, V295N	11	M34L, T46K, T58A, Q171K, G172V, P179L, L183Q, I192R, N209T, M307I, Q350R, N352H, Y353F, D412N, G429E, V455T, I489V, L491I, G500K, S547G, T578A, T651N		V505N, E507T, T663N	664
C97ZA_SOSIP reparada estabilizada	y Q567K	A198T, S243N, K236T, V295N	11	M34L, T46K, T58A, Q171K, G172V, P179L, L183Q, I192R, N209T, M307I, Q350R, N352H, Y353F, D412N, G429E, V455T, I489V, L491I, G500K, S547G, T578A, T651N	A204I, M535N, L556P, K588E, D589V, T651F, K655I	V505N, E507T, T663N	664
Du422_SOSIP		D386N , K295N	11				664
Du422_SOSIP reparada		D386N , K295N	11	V272I, W456R, G466E, F643Y			664
Du422_SOSIP reparada estabilizada	y	D386N , K295N	11	V272I, W456R, G466E, F643Y	M535N, L556P, K588E, D589V, N651F, K655I		664

DS_sC4_SOSIP	I201C-A433C	V295N	33				655
DS_sC4_SOSIP reparada	I201C-A433C	V295N	33	A114Q, E117K, E166R, T375S, I434M			655
DS_sC4_SOSIP reparada y estabilizada	I201C-A433C	V295N	33	A114Q, E117K, E166R, T375S, I434M	A204I, I535N, L556P, Q588E, D589V, N651F, K655I	delta138-152 (SSNGTYNIIHNE TYK), delta191 (SEKSSSENSSE), delta 463 (GVP)	655

Tabla 3 Varias variantes de la proteína Env del VIH descritas en la presente. En la columna 'mutaciones de la literatura' se describen las mutaciones que se usaron en estas construcciones y que fueron descritas previamente por otros. En la columna 'PNGS agregados' se describen mutaciones que agregan un potencial sitio de N-glicosilación (en las posiciones donde muchas proteínas Env de tipo salvaje comprenden un sitio tal). En la columna 'secuencia directriz' se describe la secuencia directriz que se usó para la expresión si no era la secuencia directriz original (nativa). En la columna 'mutaciones de reparación' se describen las mutaciones que mejoran el plegamiento y la estabilidad (medidos como el rendimiento y el porcentaje de trímeros, basado en la unión a bNAbs) de algunas de las proteínas Env de tipo salvaje, como se describe en el Ejemplo 12 y en la Figura 12. En la columna 'mutaciones estabilizantes' se describen las mutaciones de la invención que estabilizan la proteína y mejoran la trimerización como se divulga en la presente. En la columna 'otras mutaciones' se describen mutaciones adicionales efectuadas para algunas construcciones. En la columna 'terminal' se describe la posición del último aminoácido (la numeración en toda la tabla es con respecto a la secuencia HXB2 Env).

Se evaluaron los sobrenadantes de células transfectadas transitoriamente con variantes de Env tipo salvaje (wt), reparadas y estabilizadas por la unión a varios anticuerpos específicos del trímero ampliamente neutralizantes dirigidos al ápice. Las sustituciones de reparación y en especial las sustituciones estabilizantes tuvieron un impacto importante sobre el contenido de trímeros (Figuras 13 y 14), determinado con AlphaLISA (Figura 13) y SEC-MALS (Figura 14).

La secuencia de una variante preferida de la proteína Env DS_sC4 reparada y estabilizada (DS_sC4_SOSIP Env (HIV170686) reparada y estabilizada) se provee en la SEQ ID N°: 27.

Otra variante preferida de la misma se provee en la SEQ ID N°: 32 (sC4_SOSIP.v4 Env reparada y estabilizada.)

Ejemplo 13: Las mutaciones estabilizantes de la invención funcionan en ausencia de las mutaciones SOSIP

Según se mostró en los ejemplos precedentes, las 7 mutaciones (A204I, I535N, I573F, K588E, D589V, N651F y K655I) mejoraron el rendimiento y porcentaje de trímeros en la ConC_SOSIP (que da como resultado una 'ConC_base' o 'ConC_SOSIP estabilizada' o 'ConcC_SOSIP 7mut') (por ejemplo, las Figuras 13 y 15).

Este ejemplo muestra que las diferentes mutaciones SOSIP (es decir, la mutación 'SOS': 2 sustituciones por residuos Cys en las posiciones 501 y 605; y la 'mutación IP': sustitución por el residuo Pro en la posición 559) contribuyen en una estabilización adicional, pero no son necesarios para obtener los beneficios de las mutaciones de la invención.

Se mostró que las 7 mutaciones también mejoran el rendimiento de trímeros en la denominada ConC_SOS, que no contiene la mutación estabilizante I559P (mutación 'IP'), como se muestra en la Figura 15 (compárese ConC_SOS versus ConC_SOS, 7mut). Por ende, la mutación 'IP' no es esencial para obtener los beneficios de las mutaciones que se describen en la presente. La adición de la mutación I559P dio como resultado un gran aumento, lo que muestra que la mutación 'IP' es beneficiosa en esta construcción además de las 7 mutaciones de la invención. La mutación estabilizante IP (I559P) también se podría reemplazar por A558P o L556P, donde ambas también dan como resultado un gran aumento ante la variante que carece de la mutación I559P.

Además, la ConC_IP, 7 mut, que contiene las 7 mutaciones de la invención que se describieron antes, pero carece de las mutaciones 'SOS', aún mostró un rendimiento de trímeros muy alto, lo que demuestra que las mutaciones 'SOS' tampoco son esenciales para obtener los beneficios de las mutaciones que se describen en la presente (por ejemplo, compárese

ConC_SOSIP versus ConC_IP, 7 mut), en línea con las observaciones en el Ejemplo 7. La adición de la mutación 'SOS' aumenta aún más el rendimiento de trímeros.

5 Por consiguiente, en tanto los trímeros Env que contienen las mutaciones estabilizantes que se describen en la presente pueden aprovechar una estabilización adicional con las mutaciones SOSIP, ninguna de las 3 mutaciones SOSIP es necesaria para obtener los beneficios (por ejemplo, un rendimiento de trímeros mejorado) de las mutaciones estabilizantes que se describen en la presente.

Ejemplo 14: La sustitución por metionina en las posiciones 647, 651 o 655 mejora la calidad del trímero

10 Además de las mutaciones que se describen en el Ejemplo 2, las posiciones 589, 647, 651 y 655 se sustituyeron individualmente por un residuo Met en un esqueleto ConC_SOSIP (SEQ ID N°: 3) y se evaluaron según el porcentaje y rendimiento de trimerización usando los métodos que se describieron antes. Se mostró que una Met en las posiciones 647, 651 o 655, al igual que las mutaciones que se describieron en el Ejemplo 2, mejoraron la calidad del trímero (mayor porcentaje y rendimiento de trímeros, unión aumentada a bNAb), como se puede observar en la Figura 16.

15 Por consiguiente, además de la sustitución por Phe, Ala o Trp en la posición 651, la sustitución por Met en la posición 651 también mejora la formación de trímeros; además de la sustitución por Phe, Ile o Trp en la posición 655, la sustitución por Met en la posición 655 también mejora la formación de trímeros; y además de la sustitución por Phe, o Ile en la posición 647, la sustitución por Met en la posición 647 también mejora la formación de trímeros.

Ejemplo 15: Proteína Env del VIH con una mutación estabilizante de trímeros en la posición 658.

20 Se prepararon proteínas Env del VIH recombinantes con mutaciones de sustitución en la posición 658 (la numeración es acorde con la gp160 de la forma aislada HXB2 del VIH-1), en el esqueleto ConC_SOSIP (SEQ ID N°: 3). K658 se mutó por Val, Ile, Phe, Leu, Met o Ala. Además, se obtuvieron algunos mutantes dobles, donde estas mutaciones se combinaron con una de las mutaciones estabilizantes que se describieron antes, K655I. El porcentaje de formación de trímeros se determinó mediante el ensayo AlphaLISA como se describe en el Ejemplo 3.

25 Los resultados se muestran en las Figuras 17A y B (porcentaje de trímeros, medido en diferentes experimentos, por ende dos paneles) y las Figuras 17C y D (rendimiento de trímeros, medido en diferentes experimentos, por ende dos paneles). Estos resultados demuestran que la sustitución en la posición 658 por Ile, Phe, Met, Leu, Ala o Val dio como resultado un mayor porcentaje de formación de trímeros y un rendimiento de trímeros mejorado. La sustitución por Ile en la posición 658 dio como resultado aumentos que se encuentran aproximadamente en el mismo rango que la mutación K655I (Figuras 17A, C), que era el mutante simple de mejor desempeño de las mutaciones (i)-(vii) en la Tabla 1 descritas antes (véase, por ejemplo, la Figura 2A). La sustitución por Val en la posición 658 dio como resultado mejoras aún mayores (Figuras 30 17A, C).

Los resultados también demostraron que la sustitución en la posición 658 por Ile o Val se podría combinar con la mutación K655I descrita antes, y que esto dio como resultado una mejora adicional ante cada uno de los correspondientes mutantes simples (Figuras 17A, C).

35 El mutante K658V también se evaluó usando SEC-MALS. Se hicieron crecer cultivos en 96 pocillos durante tres días al igual que en un AlphaLISA. El sobrenadante se cargó directamente en una columna SEC-MALS. Los cromatogramas obtenidos para el sobrenadante simulado (con expresión de furina) se sustrajo de los cromatogramas del sobrenadante con proteínas Env. La proteína trimérica eluyó de la columna entre los 7 y 8 minutos. Los resultados se muestran en la Figura 18, y confirman que el mutante K658V mostró una trimerización mejorada ante la proteína Env de antecedente, y ante la proteína Env mutante K655I.

40 Este ejemplo demuestra que la sustitución del aminoácido en la posición 658 en la proteína Env del VIH por Val, Ile, Phe, Met, Leu o Ala da como resultado un mayor porcentaje de trímeros y rendimiento de trímeros.

45 Se condujeron otros experimentos para medir la formación de trímeros de variantes usando AlphaLISA y/o SEC-MALS en las variantes de la Env del VIH, en donde la mutación K658V está presente en combinación con otras mutaciones de las Tablas 1 y/o 2 como se describe en la presente, así como en cepas de los clado A y B del VIH. Por ejemplo, ya se ha mostrado que la mutación 658V mejora la variante ConC_SOSIP,7mut como se describió antes (ejemplo 7), así como BG505_SOSIP con L556P, K655I, M535N, N651F, D589V, K588E (ejemplo 9), así como la C97ZA_SOSIP reparada y estabilizada (ejemplo 10).

50 Sobre la base de los resultados que se describieron precedentemente, se espera que la mutación del aminoácido en la posición 658 por un residuo Valina, Isoleucina, Fenilalanina, Leucina, Metionina o Alanina, preferiblemente por Valina, mejorará la formación de trímeros y/o el rendimiento de trímeros en diferentes antecedentes de proteínas Env del VIH.

Ejemplo 16. Inmunización con las proteínas Env del VIH estabilizadas

Se condujo un estudio de inmunización en conejo con la proteína Env soluble y con proteínas Env acopladas a liposomas. La imprimación se efectúa con la ConC_SOSIP estabilizada.v3 (SEQ ID N°: 28), y es seguida por cuatro refuerzos, cada uno con otra proteína, es decir con 1) sC4_SOSIP.v4 reparada y estabilizada (SEQ ID N°: 32); 2) C97ZA_SOSIP.v2 reparada y estabilizada (SEQ ID N°: 30); 3) Du422_SOSIP.v1 reparada y estabilizada (SEQ ID N°: 31); y 4) BG505_SOSIP.v2 estabilizada (SEQ ID N°: 29).

Se aísla suero después de sucesivas inmunizaciones, y se analiza por los anticuerpos inducidos que se unen particularmente a la conformación estable, cerrada, de prefusión de la Env (usando un ELISA), así como por la inducción de bNAbs (usando ensayos de neutralización de virus).

Los ejemplos precedentes muestran que la invención provee un abordaje universal para optimizar el plegamiento y la estabilidad de proteínas de trómeros de la envoltura cerradas, de prefusión, del VIH.

Se comprenderá que los ejemplos y las formas de realización que se describen en la presente solamente se ofrecen a efectos ilustrativos, y se podrían efectuar cambios a las formas de realización descritas antes. Por lo tanto, se comprenderá que esta invención no se limita a las formas de realización particulares divulgadas, sino que está definida por las reivindicaciones adjuntas.

LISTADO DE SECUENCIAS

SEQ ID N°: 1 gp160 de la forma aislada HXB2 del VIH-1 (secuencia señal en cursiva; los aminoácidos en las posiciones (i)-(vii) para las mutaciones de acuerdo con la invención están indicados mediante un sombreado gris)

MRVKEKYQHLWRWGWWRWGTMLLGLMLICSATEKLWVTVYYGVPVWKEATTLFCASDAKAYDTEVHNVWATHACVPT
 DPNPQEVVLNVNTENFNMWKNMVEQMHEIISLWDQSLKPCVKLTPLCVSLKCTDLKNDTNTNSSSGRMIMEKGEIKN
 CSFNISTSIRGKVQKEYAFFYKLDIIPIDNDTTSYKLTSCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNKTFNGTGPCTN
 VSTVQCTHGIRPVVSTQLLNGSLAEEEEVIRSVNFTDNAKTIIVQLNTSVEINCTRPNNNTRKRIRIQRGPGRFVFIGKIGN
 MRQAHCNISRAKWNNTLQIASKLREQFGNNKTIIFKQSSGGDPEIVTHSFNCGGEFFYCNSTQLFNSTWFNSTWSTEGS
 NNTGSDTITLPCRIKQIINMWQKVGKAMYAPPISGQIRCSSNITGLLLTRDGGNSNNESEIFRPGGGDMRDNRSELYKY
 KVKIEPLGVAPTAKRRVVQREKRAVGIGALFLGFLGAAGSTMGAASMTLTVQARQLLSGIVQQQNNLLRAIEAQQHLLQ
 LTVWGIKQLQARILAVERYLKDQQLLGIWGC SGKLICTTAVPWNASWSNKSLEQIWNHTTWMEWDREINNYTSLIHSLEE
 SQNQKEKNEQELLELDKWAASLWNWFNITNWLWYIKLFIMIVGGLRIVFAVLSIVNRVRRQGYSPFSFQTHLPTPRGPDR
 PEGIEEGGERDRDRSIRLVNGLSLALIWDLLRSLCLFSYHRLRDLILLIVTRIVELLGRRGWEALKYWWNLLQYWSQELKNS
 AVSLLNATAIAVAEGTDRVIEVVQGACRAIRHPRRIRQGLERILL

SEQ ID N°: 2 Env consenso del clado C del VIH (secuencia consenso solamente, no incluye ninguna secuencia señal, dominio transmembrana (664 es el último aminoácido), mutaciones SOSIP y/o mutaciones en el sitio de clivaje de furina; los aminoácidos en las posiciones (i)-(vii) para las mutaciones de acuerdo con la invención están indicados mediante un sombreado gris)

NLWVTVYYGVPVWKEAKTTLFCASDAKAYEKEVHNVWATHACVPTDPNPQEMVLENTENFNMWKNMVDQMHEIIS
 LWDQSLKPCVKLTPLCVTLNCTNVNVTNTNNNNMKEEMKNCSFNNTTEIRDKKQKEYALFYRLDIVPLNENSSEYRLINCN
 TSTITQACPKVSFDPIPIHYCAPAGYAILKCNKTFNGTGPCNNVSTVQCTHGIRPVVSTQLLNGSLAEEEEIIRSENLTDNA
 KTIIVHLNESVEINCTRPNNNTRKSIRIGPGQTFYATGDIIGDIRQAHCNISEAKWNKTLQRVKKKLKEHFNPNTIKFAPSSGG
 DLEITTHSFNCRGEFFYCNTSKLFNSTYNNNTSNSTITLPCRIKQIINMWQEVGRAMYAPPIAGNITCKSNITGLLLTRDGGN
 NNNNTETFRPGGGDMRDNRSELYKYKVVVEIKPLGIAPTAKRRVVEREKRRAVGIGAVFLGFLGAAGSTMGAASITLTV
 QARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGIWGC SGKLICTTAVPWNSSWSNKSQ
 EDIWDNMTWMQWDREISNYTDIYRLLLEESQNQQEKNEKDLLALD

SEQ ID N°: 3 ConC_SOSIP (secuencia consenso madura del clado C con mutaciones SOSIP y sitios de clivaje de furina (en cursiva) y truncamiento C-terminal; los aminoácidos en las posiciones (i)-(vii) para las mutaciones de acuerdo con la invención están indicados mediante un sombreado gris)

NLWVTVYYGVPVWKEAKTTLFCASDAKAYEKEVHNVWATHACVPTDPNPQEMVLENTENFNMWKNMVDQMHEIIS
 LWDQSLKPCVKLTPLCVTLNCTNVNVTNTNNNNMKEEMKNCSFNNTTEIRDKKQKEYALFYRLDIVPLNENSSEYRLINCN
 TSTITQACPKVSFDPIPIHYCAPAGYAILKCNKTFNGTGPCNNVSTVQCTHGIRPVVSTQLLNGSLAEEEEIIRSENLTDNA
 KTIIVHLNESVEINCTRPNNNTRKSIRIGPGQTFYATGDIIGDIRQAHCNISEAKWNKTLQRVKKKLKEHFNPNTIKFAPSSGG
 DLEITTHSFNCRGEFFYCNTSKLFNSTYNNNTSNSTITLPCRIKQIINMWQEVGRAMYAPPIAGNITCKSNITGLLLTRDGGN
 NNNNTETFRPGGGDMRDNRSELYKYKVVVEIKPLGIAPTCKRRVVE~~RRRRRR~~AVGIGAVFLGFLGAAGSTMGAASITLTV
 QARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGIWGC SGKLICTTAVPWNSSWSNKSQ
 SQEDIWDNMTWMQWDREISNYTDIYRLLLEESQNQQEKNEKDLLALD

SEQ ID N°: 4 Env consenso clado B del del VIH (secuencia consenso solamente, no incluye ninguna secuencia señal, dominio transmembrana (664 es el último aminoácido), mutaciones SOSIP y/o mutaciones en el sitio de clivaje de furina; los aminoácidos en las posiciones (i)-(vii) para las mutaciones de acuerdo con la invención están indicados mediante un sombreado gris)

5 AEKLVWTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVLENTENFNMWKNMVEQM HEDI
 ISLWDQSLKPCVKLTPLCVTLNCTDLNNTTNNNSSEKMEKGEIKNCSFNITTSIRDKVQKEYALFYKLDVVPIDNNNTSY
 RLISCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNDDKFFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEEEEVVIRS
 ENFTDNAKTIIVQLNESVEINCTRPNNNTRKSIHIGPGRFYATGDIIGDIRQAHCNISRTKWNNTLKQIVKKLREQFGNKTIV
 FNQSSGGDPEIVMHSFNCGGEFFYCNTTQLFNSTWNSNGTWNNTTGNDDITLPCRIKQIINMWQEVGKAMYAPPPIRGQIR
 10 CSSNITGLLLTRDGGNNNNNTTETFRPGGGDMRDNRSELYKYKVVVIEPLGVAPTKCKRRRVVQRRRRRRRAVGIGAMFL
 GFLGAAGSTMGAASITLTVQARQLLSGIVQQQNLLRAPEAQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSSG
 KLICCTAVPWNTSWSNKSLEIWDNMTWMQWEREIDNYTGLIYTLIEESQNQQEKNEQELLELD

SEQ ID N°: 5 ConB_SOSIP (secuencia consenso madura del clado B con mutaciones SOSIP y en el sitio de clivaje de furina (en cursiva) y truncamiento C-terminal; los aminoácidos en las posiciones (i)-(vii) para las mutaciones de acuerdo con la invención están indicados mediante un sombreado gris)

15 AEKLVWTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVLENTENFNMWKNMVEQM HEDI
 ISLWDQSLKPCVKLTPLCVTLNCTDLNNTTNNNSSEKMEKGEIKNCSFNITTSIRDKVQKEYALFYKLDVVPIDNNNTSY
 RLISCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNDDKFFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEEEEVVIRS
 ENFTDNAKTIIVQLNESVEINCTRPNNNTRKSIHIGPGRFYATGDIIGDIRQAHCNISRTKWNNTLKQIVKKLREQFGNKTIV
 FNQSSGGDPEIVMHSFNCGGEFFYCNTTQLFNSTWNSNGTWNNTTGNDDITLPCRIKQIINMWQEVGKAMYAPPPIRGQIR
 20 CSSNITGLLLTRDGGNNNNNTTETFRPGGGDMRDNRSELYKYKVVVIEPLGVAPTKCKRRRVVQRRRRRRRAVGIGAMFL
 GFLGAAGSTMGAASITLTVQARQLLSGIVQQQNLLRAPEAQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSSG
 KLICCTAVPWNTSWSNKSLEIWDNMTWMQWEREIDNYTGLIYTLIEESQNQQEKNEQELLELD

SEQ ID N°: 6 fragmento Mos2S Env C4 de la proteína de la envoltura del VIH sintética; los aminoácidos en las posiciones (i)-(vii) para las mutaciones de acuerdo con la invención están indicados mediante un sombreado gris)

25 MGNLWVTVYYGVPVWKDAKTTLFCASDAKAYEKEVHNVWATHACVPTDPNPQEVVLENTENFNMWKNMVDQMHED
 IISLWDASLEPCVKLTPLCVTLNCRNVRNVSSNGTYNIIHNETYKEMKNCSFNATTVVEDRKRQKVHALFYRLDIVPLDENNS
 SEKSSSENSSEYYRLINCNTSAITQACPKVSFDPIPIHYCAPAGYAILKCNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLL
 NGSLAEEEEIIRSENLTNNAKTIIVHLNETVNITCTRPNNNTRKSIIRIGPGQTFYATGDIIGDIRQAHCNLSRDGWNKTLQGVK
 30 KKLAEHFNPNTIKFAPHSGGDLEITHTFNCRGEFFYCNTSNLNFESNIERNDSIITLPCRIKQIINMWQEVGRAIYAPPIAGNI
 TCRSNITGLLLTRDGGSNNGVPNDTETFRPGGGDMRNNWRSELYKYKVVVEVKPLGVAPTEAKRRRVVEREKRAVGIGAVF
 LGILGAAGSTMGAASITLTVQARQLLSGIVQQSNLLRAIEAQHMLQLTVWGIKQLQTRVLAIERYLQDQQLLGLWGCSSG
 KLICTTAVPWNTSWSNKSQTDIWDNMTWMQWDKEIGNYTG EYRLLLEESQNQQEK

SEQ ID N°: 7 (secuencia de DS_sC4_SOSIP_E166R; los aminoácidos en las posiciones (i)-(vii) para las mutaciones de acuerdo con la invención están indicados mediante un sombreado gris)

35 MGNLWVTVYYGVPVWKDAKTTLFCASDAKAYEKEVHNVWATHACVPTDPNPQEVVLENTENFNMWKNMVDQMHED
 IISLWDASLEPCVKLTPLCVTLNCRNVRNVSSNGTYNIIHNETYKEMKNCSFNATTVVRDRKRQKVHALFYRLDIVPLDENNS
 SEKSSSENSSEYYRLINCNTSACTQACPKVSFDPIPIHYCAPAGYAILKCNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLL
 NGSLAEEEEIIRSENLTNNAKTIIVHLNETVNINCTRPNNNTRKSIIRIGPGQTFYATGDIIGDIRQAHCNLSRDGWNKTLQGVK
 40 KKLAEHFNPNTIKFAPHSGGDLEITHTFNCRGEFFYCNTSNLNFESNIERNDSIITLPCRIKQIINMWQEVGRCIYAPPIAGNI
 TCRSNITGLLLTRDGGSNNGVPNDTETFRPGGGDMRNNWRSELYKYKVVVEVKPLGVAPTECKRRRVERRRRRRRAVGIG
 AVFLGILGAAGSTMGAASITLTVQARQLLSGIVQQSNLLRAIEAQHMLQLTVWGIKQLQTRVLAIERYLQDQQLLGLWGCSSG
 CSGKLICCTAVPWNTSWSNKSQTDIWDNMTWMQWDKEIGNYTG EYRLLLEESQNQQEK

SEQ ID N°: 8 (Mos1.Env, secuencia de la proteína de la envoltura en mosaico del VIH; los aminoácidos en las posiciones (i)-(vii) para las mutaciones de acuerdo con la invención están indicados mediante un sombreado gris)

45 AGKLVWTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVLENTENFNMWKNMVEQM HEDI
 ISLWDQSLKPCVKLTPLCVTLNCTDDVRNVNNAATNTNSSWGEPMEKGEIKNCSFNITTSIRNKVQKQYALFYKLDVVPID
 NDSNNTYRLISCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNDDKFFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSL
 AEEEEVIRSENFTNNAKTIMVQLNVSEINCTRPNNNTRKSIHIGPGRFYATGDIIGDIRQAHCNISRANWNNNTLRQIVEKL
 50 GKQFGNNKTIVFNHSSGGDPEIVMHSFNCGGEFFYCNTKLFNSTWTWNNSTWNNTKRSNDTEEHITLPCRIKQIINMWQ
 EVGKAMYAPPPIRGQIRCSSNITGLLLTRDGGNDTSGTEIFRPGGGDMRDNRSELYKYKVVVIEPLGVAPTKAKRRRVVQS
 EKSAVGIGAVFLGFLGAAGSTMGAASMTLTVQARLLLSGIVQQQNLLRAIEAQHLLQLTVWGIKQLQARVLAVERYLKD
 QQLLGIWGCSSGKLICTTTPWNASWSNKSLEDKIWNMTWMEWEREINNYTSLIYTLIEESQNQQEK

SEQ ID N°: 9 (Mos2.Env, secuencia de la proteína de la envoltura en mosaico del VIH; los aminoácidos en las posiciones (i)-(vii) para las mutaciones de acuerdo con la invención están indicados mediante un sombreado gris)

5 MGNLWVTVYYGVPVWKEAKTTLFCASDAKAYEKEVHNWVATHACVPTDPNPQEMVLENTENFNMWKNDMVDQMHE
 DIIRLWDQSLKPCVKLTPLCVTLNCTNVTNTNNNNMKEEMKNCNFNTTTEIRDKKQKEYALFYRLDIVPLNENSSSEYRLINC
 10 NGSLAEEEEIIRSENLTNNAKTIIVHLNETVNITCTRPNNNTRKSIRIGPGQTFYATGDIIGDIRQAHCNLSRDGWNKTLQGVK
 KKLAEHFPNKTINFTSSSGDLEITTHSFNCRGEFFYCNTSGLFNGTYPNGTNSNSSNITLPCRIKQIINMWQEVGRAM
 YAPPIAGNITCRSNITGLLLTRDGGSNNGVPNDTETFRPGGDMRNNWRSELYKYKVVVEVKPLGVAPTEAKRRVVEREK
 RAVGIGAVFLGILGAAGSTMGAASITLTVQARQLLSGIVQQSNLLRAIEAQQHMLQLTVWGfKQLQARVLAIERYLQDQQL
 LGLWGC SGKLICTTAVPWNTSWSNKSQTDIWDNMTWMQWDKEIGNYTGEIYRLL EESQNQQEK

SEQ ID N°: 10 (secuencia del sitio de clivaje de furina mutante)

RRRRRR

SEQ ID N°: 11 (ejemplo de una secuencia señal (por ejemplo, usada para ConC_SOSIP, y algunas de las variantes derivadas del tipo salvaje))

15 MRVGRILRNWQQWWIWGILGFWMLMICNVVG (nótese: el último VG podría ser el comienzo de la proteína madura o el final de la secuencia señal)

SEQ ID N°: 12 (ejemplo de 8 secuencias de aminoácidos que pueden reemplazar al bucle HR1)

NPDWL PDM

SEQ ID N°: 13 (ejemplo de 8 secuencias de aminoácidos que pueden reemplazar al bucle HR1)

20 GSGSGSGS

SEQ ID N°: 14 (ejemplo de 8 secuencias de aminoácidos que pueden reemplazar al bucle HR1)

DDVHPDWD

SEQ ID N°: 15 (ejemplo de 8 secuencias de aminoácidos que pueden reemplazar al bucle HR1)

RDTFALMM

25 **SEQ ID N°: 16** (ejemplo de 8 secuencias de aminoácidos que pueden reemplazar al bucle HR1)

DEEKVMD F

SEQ ID N°: 17 (ejemplo de 8 secuencias de aminoácidos que pueden reemplazar al bucle HR1)

DEDPHWDP

SEQ ID N°: 18 (ejemplo de una secuencia señal (por ejemplo, usada para ConB_SOSIP))

30 MRVKGIRKNYQHLWRWGTMLLGMLMICA

SEQ ID N°: 19 (marca usada para las construcciones gp140 VIH en el ensayo AlphaLISA)

AAALPETGGGSDYKDDDDKPGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSHHHHHH

SEQ ID N°: 20 (ConC_SOSIP estabilizada, 'ConC_SOSIP_7mut' (HIV160544))

35 NLWVTVYYGVPVWKEAKTTLFCASDAKAYEKEVHNWVATHACVPTDPNPQEMVLENTENFNMWKNDMVDQMHE
 DIIRLWDQSLKPCVKLTPLCVTLNCTNVTNTNNNNMKEEMKNCNFNTTTEIRDKKQKEYALFYRLDIVPLNENSSSEYRLINC
 40 TSTITQICPKVSFDPIPIHYCAPAGYAILKCNKTFNGTGPCNNVSTVQCTHGfKIPVSTQLLLNGSLAEEEEIIRSENLT
 DAKTIIVHLNESVEINCTRPNNNTRKSIRIGPGQTFYATGDIIGDIRQAHCNISEAKWNKTLQRVKKLKEHFPNKTIFAPSSGGD
 LEITTHSFNCRGEFFYCNTSKLFNSTYNNTTSNSTITLPCRIKQIINMWQEVGRAMYAPPIAGNITCKSNITGLLLTRDGGNN
 NNTETFRPGGDMRDNRSELYKYKVVVEIKPLGIAPTCKRRVVERRRRRAVGIGAVFLGFLGAAGSTMGAASnTLTV
 QARQLLSGIVQQSNLLRAPEAQQHMLQLTVWGfKQLQARVLAIERYLQDQQLLGLWGC SGKLICTTAVPWNTSWSNKSQ
 EDIWDNMTWMQWDREISNYTDIYRLL EESQfQQEiNEKDLLALD

SEQ ID N°: 21 (proteína Env BG505_SOSIP (HIV150673))

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWVWATHACVPTDPNPQEIHLNVTVEEFNMWKNMVEQMHTDII
 SLWDQSLKPCVKLTPLCVTLQCTNVTNITDDMRGELKNCSFNMTTELDRDKKQKVYSLFYRLDVVQINENQGNRSNNSN
 KEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEVEEVM
 5 IRSENITNNAKILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHNCVSKATWNETLGKVVQKLRKHFG
 NNTIIRFANSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYA
 PPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGMDMRDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGI
 GAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWVGKQLQARVLAVERYLRDQQLLGI
 WGCSGKLICTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQSQEQEKEQDLLALD

SEQ ID N°: 22 (proteína Env BG505_SOSIP estabilizada (HIV170863))

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWVWATHACVPTDPNPQEIHLNVTVEEFNMWKNMVEQMHTDII
 SLWDQSLKPCVKLTPLCVTLQCTNVTNITDDMRGELKNCSFNMTTELDRDKKQKVYSLFYRLDVVQINENQGNRSNNSN
 KEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEVEEVM
 15 IRSENITNNAKILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHNCVSKATWNETLGKVVQKLRKHFG
 NNTIIRFANSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYA
 PPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGMDMRDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGI
 GAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWVGKQLQARVLAVERYLRvQQLLGIW
 GCSGKLICTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQfQQEINEQDLLALD

SEQ ID N°: 23 (proteína Env wt C97ZA_SOSIP con L535M y Q567K (HIV150673))

NMWVTVYYGVPVWTDAKTTLFCASDTKAYDREHVHNWATHACVPTDPNPQEIHLNVTENFNMWKNMVDQMHEDIIS
 LWDQSLKPCVKLTPLCVTLHCTNATFKNNVTNDMNKEIRNCSFNNTTEIRDKKQGGYALFYRPDIVLLKENRNNSNNSSEYI
 LINCNASTITQACPKVNFDPPIPIHYCAPAGYAILKCNKTFSGKGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEKEIIIRSEN
 LTDNVKTIIVHLNKSVEIVCTRPNNNTRKSMRIGPGQTFYATGDIIGDIRQAYCNISGSKWNETLKRKVEKLEQENYNNNKTIK
 FAPSSGGDLEITTHSFNCRGEFFYCNTTRLFNNAETEDETITLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLVR
 25 DGGEDNKTEEIFRPGGGNMKDNWRSELYKYKVIELKPLGIAPTGcKRRVVERrrrrRAVGIgAVFLGFLGAAGSTMGAASmT
 LTVQARQLLSSIVQQQSNLLRApEAQQHMLkLTVWGIKQLQTRVLAIERYLKDDQQLLGIWGCSGKLIcCtNVPWNSSWSNK
 SQTDIWNNMTWMEWDREISNYTDTIYRLLLEDSTQEQEKEKDLLALD

SEQ ID N°: 24 (proteína Env C97ZA_SOSIP reparada y estabilizada (HIV170690))

NIWVTVYYGVPVWkDAKTTLFCASDaKAYDREHVHNWATHACVPTDPNPQEIHLNVTENFNMWKNMVDQMHEDIISL
 WDQSLKPCVKLTPLCVTLHCTNATFKNNVTNDMNKEIRNCSFNNTTEIRDKKQkvYALFYRIDIVqLKENRNNSNNSSEYrLIN
 CNtSTITQqCPKvIFDPIPIHYCAPAGYAILKCNKTFnGtGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEKEIIIRSENLTDNV
 KTIIVHLNKSVEInCTRPNNNTRKSiRIGPGQTFYATGDIIGDIRQAYCNISGSKWNETLKRKVEKLErEhfNNNKTIKFAPSSGG
 DLEITTHSFNCRGEFFYCNTTRLFNNAETEnETITLPCRIKQIINMWQeVGRAMYAPPIAGNITCKSNITGLLLtRDGGEDNKt
 EEIFRPGGGNMKDNWRSELYKYKvEiKPLGIAPTkcKRRnVtRrrrrRAVGIgAVFLGFLGAAGSTMGAASnTLTVQARQLLS
 35 gIVQQQSNLpRApEAQQHMLkLTVWGIKQLQaRVLAIERYLvQQLLGIWGCSGKLIcCtNVPWNSSWSNKsQTDIWNNMT
 WMEWDREISNYTDTIYRLLLEDSTQfQQEiNEKDLLAnD

SEQ ID N°: 25 (variante de la construcción Du422 reparada y estabilizada (HIV161818))

NLWVTVYYGVPVWKEAKTTLFCASDAKAYDKEVHNWVWATHACVPTDPNPQEIHLNVTENFNMWKNMVDQMHEDIISL
 WDQSLKPCVKLTPLCVTLNCKNVNISANANATATLNSSMNGEIKNCSFNNTTELRDKKQKVYALFYKPDVVPLNGGEHNE
 40 TGEYILINCNSSTcTQACPKVSFDPIPIHYCAPAGYAILKCNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEVEEIII
 RSENLTNNIKTIIVHLNKSVEINCTRPNNNTRKSVRIGPGQTFYATGEIIGDIREAHCNISRETWNSTLIQVKEKLEHREHYNKTIK
 FEPSSGGDLEVTTTHSFNCRGEFFYCNTTKLFNETKLFNESEYVDNKTIILPCRIKQIINMWQEVGRcMYAPPIEGNITCKSNI
 TGLLLTRDGGENSTEEVFRPGGGNMKDNWRSELYKYKvVEIKPLGVAPTCKCRKnVtRRRRRRRAVGLGAVLLGFLGAAG
 45 STMGAASnTLTVQARQLLSGIVQQQSNLpRAPEAQQHLLQLTVWGIKQLQTRVLAIERYLKvQQLLGLWGCSGKLICTAV
 PWNSSWSNKSLGDIWDNMTWMQWDREISNYTNTIYRLLLEDSTQfQQEKEKDLLAnD

SEQ ID N°: 26 (Du422_SOSIP reparada y estabilizada (HIV170859))

NLWVTVYYGVPVWKEAKTTLFCASDAKAYDKEVHNWVWATHACVPTDPNPQEIHLNVTENFNMWKNMVDQMHEDIISL
 WDQSLKPCVKLTPLCVTLNCKNVNISANANATATLNSSMNGEIKNCSFNNTTELRDKKQKVYALFYKPDVVPLNGGEHNE
 TGEYILINCNSSTITQACPKVSFDPIPIHYCAPAGYAILKCNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEVEEIII
 50 RSENLTNNIKTIIVHLNKSVEINCTRPNNNTRKSVRIGPGQTFYATGEIIGDIREAHCNISRETWNSTLIQVKEKLEHREHYNKTIK
 FEPSSGGDLEVTTTHSFNCRGEFFYCNTTKLFNETKLFNESEYVDNKTIILPCRIKQIINMWQEVGRAMYAPPIEGNITCKSNI
 TGLLLTRDGGENSTEEVFRPGGGNMKDNWRSELYKYKvVEIKPLGVAPTCKCRKVVGRRRRRRRAVGLGAVLLGFLGAAG

ES 2 824 525 T3

STMGAASnTLTVQARQLLSGIVQQQSNLpRAPEAQQHLLQLTVWGIKQLQTRVLAIERYLEvQQLLGLWGCSGKLICTAV
PWNSSWSNKS LGDIWDNM TWMQWDREISNYTNTIYRLLLEDSSQfQQEINEKDLLALD

SEQ ID N°: 27 (DS_sC4_SOSIP reparada y estabilizada (HIV170686))

5 MGNLWVTVYYGVPVWKA DAKTTLFCASDAKAYEKEVHNW WATHACVPTDPNPQEIVLGNVTENFNMWKN DMVDQM HEDIISL
IISLWDqSlkPCVKLTPLCVTLNCRNVRNVEMKNCSFNATTVvRDRKQKVHALFYRLDIVPLDENNSSYRLINCNTSAcTQiC
PKVSFDPIPIHYCAPAGYAILKCNKTFNGTGPCNNVSTVQCTHG IKPVVSTQLLLNGSLAE EEEIII RSENLTNNAKTIIVHLNE
TVNINCTRPNNNTRKSIRIGPGQTFYATGDIIGDIRQAHCNLSRDGWNKT LQG VKKKLAEHFPNKTIKFAPHSGGDLEITTHs
FNCRGEFFYCNTSNLFNESNIERNDSIITLPCRIKQIINMWQEVGRcmYAPPIAGNITCRSNITGLLLTRDGGSNNDTETFR
10 PGGGDMRNNWRSELYKYKVVEIKPLGVAPTECKRRVVERRRRRRRAVGIGAVFLGILGAAGSTMGAASnTLTVQARQLLS
GIVQQQSNLpRAPEAQQHMLQLTVWGIKQLQTRVLAIERYLEvQQLLGLWGCSGKLICTAVPWN TSWSNKSQTDIWDN
MTWMQWDKEIGNYTGEIYRLLLEESQfQQEi

SEQ ID N°: 28 (ConC_SOSIP.v3 estabilizada (HIV170654))

15 NLWVTVYYGVPVWKEAKTTLFCASDAKAYEKEVHNW WATHACVPTDPNPQEMVLENTENFNMWKN DMVDQM HEDIISL
LWDQSLKPCVKLTPLCVTLNCTNVNVTMKNCSFNNTTTEIRDKKQKEYALFYRLDIVPLNENSSEYRLINCNTSTITQICPKV
SFDPIPIHYCAPAGYAILKCNKTFNGTGPCNNVSTVQCTHG IKPVVSTQLLLNGSLAE EEEIII RSENLTNNAKTIIVHLNESVE
IVCTRPNNNTRKSIRIGPGQTFYATGDIIGDIRQAHCNISEAKWNKT LQRVKKKLEHFPNKTIKFQPSGGDLEITTHSFNC
RGEFFYCNTSKLFNSTYNNTTSNSTITLPCRIKQIINMWQEVGRAMYAPPIAGNITCKSNITGLLLTRDGGNNNNNTETFRP
GGGDMRDNRSELYKYKVVEIKPLGIAPTCKRRVVERRRRRRRAVGIGAVFLGFLGAAGSTMGAASnTLTVQARQLLSGI
20 VQQQSNLLRAPEAQQHMLQLTVWGFQQLQARVLAIERYLEVQQLLGIWGCSGKLICTAVPWNSSWSNKSQEDIWDNM
TWMQWDREISNYTDTIYRLLLEESQfQQEINEKDLLALD

SEQ ID N°: 29 (BG505_SOSIP.v2 estabilizada (HIV171814))

25 AENLWVTVYYGVPVWKAETTLFCASDAKAYETE KHNW WATHACVPTDPNPQEIHLNVT EEFNMWKNM VEQMHTDII
SLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMTTEL RDKKQKVYSLFYRLDVQINENQGNRSNNSN
KEYRLINCNTSAcTQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHG IKPVVSTQLLLNGSLAE EEEV
MIRSENITNNAKNILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSKATWNETLGKVVVKLRKHF
GNNTIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQcMYA
PPIQG VIRCVSNTGLILTRDGGSTNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGI
30 GAVFLGFLGAAGSTMGAASnTLTVQARNLLSGIVQQQSNLpRAPEAQQHLLKLTVWGIKQLQARVLAVERYLevQQLLGIW
GCSGKLICTNVPWNSSWSNRNLSEIWDNM T WLQWDKEISNYTQIIYGLLEESQfQQEINEvDLLALD

SEQ ID N°: 30 (C97ZA_SOSIP.v2 reparada y estabilizada (HIV171810))

35 NIWVTVYYGVPVWkDAKTTLFCASDaKAYDREVHNW WATHACVPTDPNPQEIVLENTENFNMWKN DMVDQM HEDIISL
WDQSLKPCVKLTPLCVTLHCTNATFKNNVTNDMNKEIRNCSFNNTTTEIRDKKQkvYALFYRIDIVqLKENRNNNSNNSEYrLIN
CNtStcTQiCPKvifDPIPIHYCAPAGYAILKCNKTFnGtGPCNNVSTVQCTHG IKPVVSTQLLLNGSLAEKEIII RSENLTDNV
KTIIVHLNKSVEInCTRPNNNTRKSIRIGPGQTFYATGDIIGDIRQAYCNISGSKWNETLKRKVEKLErEhfNNNKTIKFAPSSGG
DLEITTHSFNCRGEFFYCNTTRLFNNA TE nETITLPCRIKQIINMWQeVGRcMYAPPIAGNITCKSNITGLLLTRDGGEDNKT
EEIFRPGGGNMKDNWRSELYKYKVvEiKPLGIAPTkcKRRnVtRrrrrRAVGIGAVFLGFLGAAGSTMGAASnTLTVQARQLLS
gIVQQQSNLpRApEAQQHMLkLTVWGIKQLQaRVLAIERYLEvQQLLGIWGCSGKLIcCtNVPWNSSWSNKSQTDIWNMT
WMEWDREISNYTDTIYRLLLEDSSQfQQEINEvDLLAnD

SEQ ID N°: 31 (Du422_SOSIP.v1 reparada y estabilizada (HIV171812))

40 NLWVTVYYGVPVWKEAKTTLFCASDAKAYDKEVHNW WATHACVPTDPNPQEIVLENTENFNMWKN DMVDQM HEDIISL
WDQSLKPCVKLTPLCVTLNCKNVNISANANATATLNSSMNGEIKNCSFNNTTTEL RDKKQKVYALFYKPDVVPLNGGEHNE
TGEYILINCNSSTcTQACPKVSFDPIPIHYCAPAGYAILKCNKTFNGTGPCNNVSTVQCTHG IKPVVSTQLLLNGSLAE EEEIII
RSENLTNNIKTIIVHLNKSVEINCTRPNNNTRKSVRIGPGQTFYATGEIIGDIREAHCNISRETWNSTLIQVKEKLREHYNKTIK
45 FEPSSGGDLEVTTHSFNCRGEFFYCNTTKLFNETKLFNESEYVDNKTIILPCRIKQIINMWQEVGRcMYAPPIEGNITCKSNI
TGLLLTRDGGENSTEEVFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTCKKRKVVGRRRRRRRAVGLGAVLLGFLGAAG
STMGAASnTLTVQARQLLSGIVQQQSNLpRAPEAQQHLLQLTVWGIKQLQTRVLAIERYLEvQQLLGLWGCSGKLICTAV
PWNSSWSNKS LGDIWDNM TWMQWDREISNYTNTIYRLLLEDSSQfQQEINEvDLLALD

SEQ ID N°: 32 (sC4_SOSIP.v4 estabilizada y reparada)

50 MGNLWVTVYYGVPVWKA DAKTTLFCASDAKAYEKEVHNW WATHACVPTDPNPQEIVLGNVTENFNMWKN DMVDQM HEDIISL
IISLWDqSlkPCVKLTPLCVTLNCRNVRNVEMKNCSFNATTVvRDRKQKVHALFYRLDIVPLDENNSSYRLINCNTSAcTQiC
PKVSFDPIPIHYCAPAGYAILKCNKTFNGTGPCNNVSTVQCTHG IKPVVSTQLLLNGSLAE EEEIII RSENLTNNAKTIIVHLNE

TVNIVCTRPNNNTRKSIRIGPGQTFYATGDIIGDIRQAHCNLSRDGWNKTLQGvkkkLAEHFpNkTIKFAPHSgGDLEITThs
 FNCRGEFFYCNTSNLNFESNIERNDSIITLPCRIKQIINMWQEVGRcmYAPPIAGNITCRSNITGLLLTRDGGSNNDTETFR
 PGGGDMRNNWRSELYKYKVVVEVKPLGVAPTECKRRnVtRRRRRRRAVGIGAVFLGILGAAGSTMGAASnTLTVQARQLLS
 GIVQQSNLpRAPEAQQHMLQLTVWGIKQLQTRVLAIERYLedQQLLGLWGCSGKLIcCTAVPWNTSWSNKSQTDIWDN
 5 MTWMQWDKEIGNYTGEIYRLLLEESQfQQEi

SEQ ID N°: 33 (ejemplo de una secuencia señal (por ejemplo, usada para las variantes de DS_sC4_SOSIP)

MRVRGMLRNWQQWWIWSSLGFWMLMIYSV

SEQ ID N°: 34 (ejemplo de una secuencia señal (por ejemplo, usada para las variantes de BG505_SOSIP)

MRVMGIQRNCQHlFRWGTMILGMIIICSA

10 BIBLIOGRAFÍA

1. Sanders et al. J. Virol. (2002) 76(17), 8875-89
2. Sanders et al. Science (2015) 349(6224), 139-140
3. Julien et al. Proc. Nat. Acad. Sci. (2015) 112(38), 11947-52
4. de Taeye et al. Cell (2015) 163(7), 1702-15
- 15 5. Kwon et al. (2015) Nat. Struct. Mol. Biol. 22(7) 522-31
6. Eglen et al. Curr. Chem. Genomics, (2008) 25(1), 2-10
7. Kong et al, Nat Commun. 2016 Jun 28;7:12040. doi: 10.1038/ncomms12040
8. Julien et al. Proc. Natl. Acad. Sci. (2015) 112(38) 11947-52
9. Barouch et al, Nat Med 2010, 16: 319-323
- 20 10. WO 2010/059732
11. Solicitud de Patente Europea EP15200138.4
12. Sharma SK, et al. Cell Rep. (2015) 11(4):539-50. doi: 10.1016/j.celrep.2015.03.047.
13. Georgiev IS, et al. J Virol. (2015) 89(10):5318-29. doi: 10.1128/JVI.03451-14.
14. López-Sagaseta J, et al (2016) Computational y Struct Biotechnol J 14: 58-68.
- 25 15. Zhao L, et al (2014) Vaccine 32: 327-337
16. He L, et al (2016) Nat Commun. 2016 Jun 28;7:12041. doi: 10.1038/ncomms12041
17. WO2011082087
18. Kesavardhana A y Varadarajan R (2014) J Virol 88: 9590-9604
19. Guenaga J, et al (2015) Immunity 46: 792-803
- 30 20. Bale S, et al (2017) J. Virol. doi:10.1128/JVI.00443-17
21. Abbink et al (2007) Viro. 81(9): 4654-64
22. Altschul SF et al (1997) Nucleic Acid Res. 25: 3389-3402
23. Harris et al (2011) PNAS 108 (28): 11440-11445
24. Kushnir et al (2012) Vaccine (31): 58-83
- 35 25. WO 2007/104792
26. WO 2014/124301

- 27. US 2016/0122392
- 28. WO 2016/037154
- 29. Pugach et al., 2015, *J Virol*, 89: 3380-3395

5 **LISTA DE SECUENCIAS**

<110> Janssen Vaccines & Prevention

B.V. Rutten, Lucy

Truan, Daphné

Strokappe, Nika M.

10 Langedijk, Johannes P.M.

<120> Mutaciones estabilizantes de trímeros de proteínas de la envoltura del VIH

<130> 0276 WO 00 ORD

15

<150> EP16188866.4

<151> 2016-09-15

<160> 34

20

<170> PatentIn versión 3.5

<210> 1

<211> 856

25 <212> PRT

<213> VIH-1

<220>

<221> característica_misc

30 <223> gp160 de la forma aislada HXB2 del VIH-1

<400> 1

ES 2 824 525 T3

Met Arg Val Lys Glu Lys Tyr Gln His Leu Trp Arg Trp Gly Trp Arg
 1 5 10 15

Trp Gly Thr Met Leu Leu Gly Met Leu Met Ile Cys Ser Ala Thr Glu
 20 25 30

Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala
 35 40 45

Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu
 50 55 60

Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn
 65 70 75 80

Pro Gln Glu Val Val Leu Val Asn Val Thr Glu Asn Phe Asn Met Trp
 85 90 95

Lys Asn Asp Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp
 100 105 110

Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Ser
 115 120 125

ES 2 824 525 T3

Leu Lys Cys Thr Asp Leu Lys Asn Asp Thr Asn Thr Asn Ser Ser Ser
 130 135 140

Gly Arg Met Ile Met Glu Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn
 145 150 155 160

Ile Ser Thr Ser Ile Arg Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe
 165 170 175

Tyr Lys Leu Asp Ile Ile Pro Ile Asp Asn Asp Thr Thr Ser Tyr Lys
 180 185 190

Leu Thr Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val
 195 200 205

Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala
 210 215 220

Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr
 225 230 235 240

Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser
 245 250 255

Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val Val Ile
 260 265 270

Arg Ser Val Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu
 275 280 285

Asn Thr Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg
 290 295 300

Lys Arg Ile Arg Ile Gln Arg Gly Pro Gly Arg Ala Phe Val Thr Ile
 305 310 315 320

Gly Lys Ile Gly Asn Met Arg Gln Ala His Cys Asn Ile Ser Arg Ala
 325 330 335

Lys Trp Asn Asn Thr Leu Lys Gln Ile Ala Ser Lys Leu Arg Glu Gln
 340 345 350

Phe Gly Asn Asn Lys Thr Ile Ile Phe Lys Gln Ser Ser Gly Gly Asp
 355 360 365

Pro Glu Ile Val Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr
 370 375 380

ES 2 824 525 T3

Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Phe Asn Ser Thr Trp
 385 390 395 400
 Ser Thr Glu Gly Ser Asn Asn Thr Glu Gly Ser Asp Thr Ile Thr Leu
 405 410 415
 Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Lys Val Gly Lys
 420 425 430
 Ala Met Tyr Ala Pro Pro Ile Ser Gly Gln Ile Arg Cys Ser Ser Asn
 435 440 445
 Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Ser Asn Asn Glu
 450 455 460
 Ser Glu Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg
 465 470 475 480
 Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val
 485 490 495
 Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg Ala
 500 505 510
 Val Gly Ile Gly Ala Leu Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser
 515 520 525
 Thr Met Gly Ala Ala Ser Met Thr Leu Thr Val Gln Ala Arg Gln Leu
 530 535 540
 Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu
 545 550 555 560
 Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu
 565 570 575
 Gln Ala Arg Ile Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu
 580 585 590
 Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val
 595 600 605
 Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu Glu Gln Ile Trp Asn
 610 615 620
 His Thr Thr Trp Met Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Ser
 625 630 635 640

ES 2 824 525 T3

Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn
 645 650 655
 Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp
 660 665 670
 Phe Asn Ile Thr Asn Trp Leu Trp Tyr Ile Lys Leu Phe Ile Met Ile
 675 680 685
 Val Gly Gly Leu Val Gly Leu Arg Ile Val Phe Ala Val Leu Ser Ile
 690 695 700
 Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr His
 705 710 715 720
 Leu Pro Thr Pro Arg Gly Pro Asp Arg Pro Glu Gly Ile Glu Glu Glu
 725 730 735
 Gly Gly Glu Arg Asp Arg Asp Arg Ser Ile Arg Leu Val Asn Gly Ser
 740 745 750
 Leu Ala Leu Ile Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr
 755 760 765
 His Arg Leu Arg Asp Leu Leu Leu Ile Val Thr Arg Ile Val Glu Leu
 770 775 780
 Leu Gly Arg Arg Gly Trp Glu Ala Leu Lys Tyr Trp Trp Asn Leu Leu
 785 790 795 800
 Gln Tyr Trp Ser Gln Glu Leu Lys Asn Ser Ala Val Ser Leu Leu Asn
 805 810 815
 Ala Thr Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Val Ile Glu Val
 820 825 830
 Val Gln Gly Ala Cys Arg Ala Ile Arg His Ile Pro Arg Arg Ile Arg
 835 840 845
 Gln Gly Leu Glu Arg Ile Leu Leu
 850 855

<210> 2

<211> 615

<212> PRT

<213> Secuencia artificial

5

<220>

<223> Secuencia consenso de la Env del clado C del VIH

<400> 2

ES 2 824 525 T3

Asn Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala
 1 5 10 15

 Lys Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu
 20 25 30

 Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn
 35 40 45

 Pro Gln Glu Met Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp
 50 55 60

 Lys Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp
 65 70 75 80

 Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr
 85 90 95

 Leu Asn Cys Thr Asn Val Asn Val Thr Asn Thr Asn Asn Asn Asn Met
 100 105 110

 Lys Glu Glu Met Lys Asn Cys Ser Phe Asn Thr Thr Thr Glu Ile Arg
 115 120 125

 Asp Lys Lys Gln Lys Glu Tyr Ala Leu Phe Tyr Arg Leu Asp Ile Val
 130 135 140

 Pro Leu Asn Glu Asn Ser Ser Glu Tyr Arg Leu Ile Asn Cys Asn Thr
 145 150 155 160

 Ser Thr Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Asp Pro Ile Pro
 165 170 175

 Ile His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn
 180 185 190

 Lys Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn Val Ser Thr Val Gln
 195 200 205

 Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn
 210 215 220

 Gly Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr
 225 230 235 240

ES 2 824 525 T3

Asp Asn Ala Lys Thr Ile Ile Val His Leu Asn Glu Ser Val Glu Ile
 245 250 255
 Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly
 260 265 270
 Pro Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg
 275 280 285
 Gln Ala His Cys Asn Ile Ser Glu Ala Lys Trp Asn Lys Thr Leu Gln
 290 295 300
 Arg Val Lys Lys Lys Leu Lys Glu His Phe Pro Asn Lys Thr Ile Lys
 305 310 315 320
 Phe Ala Pro Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe
 325 330 335
 Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Lys Leu Phe Asn
 340 345 350
 Ser Thr Tyr Asn Asn Thr Thr Ser Asn Ser Thr Ile Thr Leu Pro Cys
 355 360 365
 Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala Met
 370 375 380
 Tyr Ala Pro Pro Ile Ala Gly Asn Ile Thr Cys Lys Ser Asn Ile Thr
 385 390 395 400
 Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn Asn Asn Asn Thr Glu
 405 410 415
 Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu
 420 425 430
 Leu Tyr Lys Tyr Lys Val Val Glu Ile Lys Pro Leu Gly Ile Ala Pro
 435 440 445
 Thr Lys Ala Lys Arg Arg Val Val Glu Arg Glu Lys Arg Arg Ala Val
 450 455 460
 Gly Ile Gly Ala Val Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr
 465 470 475 480
 Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu

ES 2 824 525 T3

				485					490					495			
Ser	Gly	Ile	Val	Gln	Gln	Gln	Ser	Asn	Leu	Leu	Arg	Ala	Ile	Glu	Ala		
			500					505					510				
Gln	Gln	His	Met	Leu	Gln	Leu	Thr	Val	Trp	Gly	Ile	Lys	Gln	Leu	Gln		
		515					520					525					
Ala	Arg	Val	Leu	Ala	Ile	Glu	Arg	Tyr	Leu	Lys	Asp	Gln	Gln	Leu	Leu		
	530					535					540						
Gly	Ile	Trp	Gly	Cys	Ser	Gly	Lys	Leu	Ile	Cys	Thr	Thr	Ala	Val	Pro		
545					550					555					560		
Trp	Asn	Ser	Ser	Trp	Ser	Asn	Lys	Ser	Gln	Glu	Asp	Ile	Trp	Asp	Asn		
				565					570					575			
Met	Thr	Trp	Met	Gln	Trp	Asp	Arg	Glu	Ile	Ser	Asn	Tyr	Thr	Asp	Thr		
			580					585					590				
Ile	Tyr	Arg	Leu	Leu	Glu	Glu	Ser	Gln	Asn	Gln	Gln	Glu	Lys	Asn	Glu		
		595					600					605					
Lys	Asp	Leu	Leu	Ala	Leu	Asp											
	610					615											

<210> 3

<211> 616

<212> PRT

5 <213> Secuencia artificial

<220>

<223> Secuencia ConC_SOSIP

10 <400> 3

ES 2 824 525 T3

Asn Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala
1 5 10 15

Lys Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu
20 25 30

Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn
35 40 45

Pro Gln Glu Met Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp
50 55 60

Lys Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp
65 70 75 80

ES 2 824 525 T3

Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr
 85 90 95

Leu Asn Cys Thr Asn Val Asn Val Thr Asn Thr Asn Asn Asn Asn Met
 100 105 110

Lys Glu Glu Met Lys Asn Cys Ser Phe Asn Thr Thr Thr Glu Ile Arg
 115 120 125

Asp Lys Lys Gln Lys Glu Tyr Ala Leu Phe Tyr Arg Leu Asp Ile Val
 130 135 140

Pro Leu Asn Glu Asn Ser Ser Glu Tyr Arg Leu Ile Asn Cys Asn Thr
 145 150 155 160

Ser Thr Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Asp Pro Ile Pro
 165 170 175

Ile His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn
 180 185 190

Lys Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn Val Ser Thr Val Gln
 195 200 205

Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn
 210 215 220

Gly Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr
 225 230 235 240

ES 2 824 525 T3

Asp Asn Ala Lys Thr Ile Ile Val His Leu Asn Glu Ser Val Glu Ile
245 250 255

Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly
260 265 270

Pro Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg
275 280 285

Gln Ala His Cys Asn Ile Ser Glu Ala Lys Trp Asn Lys Thr Leu Gln
290 295 300

Arg Val Lys Lys Lys Leu Lys Glu His Phe Pro Asn Lys Thr Ile Lys
305 310 315 320

Phe Ala Pro Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe

ES 2 824 525 T3

				325						330						335
Asn	Cys	Arg	Gly	Glu	Phe	Phe	Tyr	Cys	Asn	Thr	Ser	Lys	Leu	Phe	Asn	
			340					345					350			
Ser	Thr	Tyr	Asn	Asn	Thr	Thr	Ser	Asn	Ser	Thr	Ile	Thr	Leu	Pro	Cys	
		355					360					365				
Arg	Ile	Lys	Gln	Ile	Ile	Asn	Met	Trp	Gln	Glu	Val	Gly	Arg	Ala	Met	
	370					375					380					
Tyr	Ala	Pro	Pro	Ile	Ala	Gly	Asn	Ile	Thr	Cys	Lys	Ser	Asn	Ile	Thr	
385					390					395					400	
Gly	Leu	Leu	Leu	Thr	Arg	Asp	Gly	Gly	Asn	Asn	Asn	Asn	Asn	Thr	Glu	
				405					410						415	
Thr	Phe	Arg	Pro	Gly	Gly	Gly	Asp	Met	Arg	Asp	Asn	Trp	Arg	Ser	Glu	
			420					425					430			
Leu	Tyr	Lys	Tyr	Lys	Val	Val	Glu	Ile	Lys	Pro	Leu	Gly	Ile	Ala	Pro	
		435					440					445				
Thr	Lys	Cys	Lys	Arg	Arg	Val	Val	Glu	Arg	Arg	Arg	Arg	Arg	Arg	Ala	
	450					455					460					
Val	Gly	Ile	Gly	Ala	Val	Phe	Leu	Gly	Phe	Leu	Gly	Ala	Ala	Gly	Ser	
465					470					475					480	

ES 2 824 525 T3

Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu
 485 490 495

Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Pro Glu
 500 505 510

Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu
 515 520 525

Gln Ala Arg Val Leu Ala Ile Glu Arg Tyr Leu Lys Asp Gln Gln Leu
 530 535 540

Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Cys Thr Ala Val
 545 550 555 560

Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser Gln Glu Asp Ile Trp Asp
 565 570 575

Asn Met Thr Trp Met Gln Trp Asp Arg Glu Ile Ser Asn Tyr Thr Asp
 580 585 590

Thr Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn
 595 600 605

Glu Lys Asp Leu Leu Ala Leu Asp
 610 615

<210> 4

<211> 630

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Secuencia consenso de la Env del Clado B del VIH

10

<400> 4

ES 2 824 525 T3

Ala Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys
1 5 10 15

Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp
20 25 30

Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp
35 40 45

Pro Asn Pro Gln Glu Val Val Leu Glu Asn Val Thr Glu Asn Phe Asn
50 55 60

Met Trp Lys Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser
65 70 75 80

Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys
85 90 95

Val Thr Leu Asn Cys Thr Asp Leu Asn Asn Asn Thr Thr Asn Asn Asn
100 105 110

Ser Ser Ser Glu Lys Met Glu Lys Gly Glu Ile Lys Asn Cys Ser Phe
115 120 125

Asn Ile Thr Thr Ser Ile Arg Asp Lys Val Gln Lys Glu Tyr Ala Leu
130 135 140

Phe Tyr Lys Leu Asp Val Val Pro Ile Asp Asn Asn Asn Thr Ser Tyr
145 150 155 160

Arg Leu Ile Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys

ES 2 824 525 T3

Leu Thr Arg Asp Gly Gly Asn Asn Asn Asn Asn Thr Thr Glu Thr Phe
 420 425 430

Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr
 435 440 445

Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys
 450 455 460

Cys Lys Arg Arg Val Val Gln Arg Arg Arg Arg Arg Ala Val Gly
 465 470 475 480

Ile Gly Ala Met Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met
 485 490 495

Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser
 500 505 510

Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Pro Glu Ala Gln
 515 520 525

Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala
 530 535 540

Arg Val Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly
 545 550 555 560

Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Cys Thr Ala Val Pro Trp
 565 570 575

Asn Thr Ser Trp Ser Asn Lys Ser Leu Asp Glu Ile Trp Asp Asn Met
 580 585 590

Thr Trp Met Gln Trp Glu Arg Glu Ile Asp Asn Tyr Thr Gly Leu Ile
 595 600 605

Tyr Thr Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln
 610 615 620

Glu Leu Leu Glu Leu Asp
 625 630

<210> 5

<211> 630

<212> PRT

<213> Secuencia artificial

5

<220>

<223> Secuencia ConB_SOSIP

<400> 5

ES 2 824 525 T3

Ala Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys
1 5 10 15

Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp
20 25 30

Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp
35 40 45

Pro Asn Pro Gln Glu Val Val Leu Glu Asn Val Thr Glu Asn Phe Asn
50 55 60

Met Trp Lys Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser
65 70 75 80

Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys
85 90 95

Val Thr Leu Asn Cys Thr Asp Leu Asn Asn Asn Thr Thr Asn Asn Asn
100 105 110

Ser Ser Ser Glu Lys Met Glu Lys Gly Glu Ile Lys Asn Cys Ser Phe
115 120 125

Asn Ile Thr Thr Ser Ile Arg Asp Lys Val Gln Lys Glu Tyr Ala Leu
130 135 140

Phe Tyr Lys Leu Asp Val Val Pro Ile Asp Asn Asn Asn Thr Ser Tyr
145 150 155 160

Arg Leu Ile Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys
165 170 175

Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe
180 185 190

ES 2 824 525 T3

Ala Ile Leu Lys Cys Asn Asp Lys Lys Phe Asn Gly Thr Gly Pro Cys
195 200 205

Thr Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val
210 215 220

Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val Val
225 230 235 240

ES 2 824 525 T3

Ile Arg Ser Glu Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln
 245 250 255

Leu Asn Glu Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr
 260 265 270

Arg Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Ala Thr Gly
 275 280 285

Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Ile Ser Arg Thr
 290 295 300

Lys Trp Asn Asn Thr Leu Lys Gln Ile Val Lys Lys Leu Arg Glu Gln
 305 310 315 320

Phe Gly Asn Lys Thr Ile Val Phe Asn Gln Ser Ser Gly Gly Asp Pro
 325 330 335

Glu Ile Val Met His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys
 340 345 350

Asn Thr Thr Gln Leu Phe Asn Ser Thr Trp Asn Ser Asn Gly Thr Trp
 355 360 365

Asn Asn Thr Thr Gly Asn Asp Thr Ile Thr Leu Pro Cys Arg Ile Lys
 370 375 380

Gln Ile Ile Asn Met Trp Gln Glu Val Gly Lys Ala Met Tyr Ala Pro
 385 390 395 400

Pro Ile Arg Gly Gln Ile Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu
 405 410 415

ES 2 824 525 T3

Leu Thr Arg Asp Gly Gly Asn Asn Asn Asn Asn Thr Thr Glu Thr Phe
 420 425 430

Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr
 435 440 445

Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys
 450 455 460

Cys Lys Arg Arg Val Val Gln Arg Arg Arg Arg Arg Ala Val Gly
 465 470 475 480

Ile Gly Ala Met Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met
 485 490 495

Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser
 500 505 510

Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Pro Glu Ala Gln
 515 520 525

Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala
 530 535 540

Arg Val Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly
 545 550 555 560

Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Cys Thr Ala Val Pro Trp
 565 570 575

Asn Thr Ser Trp Ser Asn Lys Ser Leu Asp Glu Ile Trp Asp Asn Met
 580 585 590

Thr Trp Met Gln Trp Glu Arg Glu Ile Asp Asn Tyr Thr Gly Leu Ile
 595 600 605

Tyr Thr Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln
 610 615 620

Glu Leu Leu Glu Leu Asp
 625 630

ES 2 824 525 T3

<210> 6

<211> 625

<212> PRT

<213> Secuencia artificial

5

<220>

<223> Fragmento MoS2 Env C4 de la proteína de la envoltura del VIH sintético

<400> 6

Met Gly Asn Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys
1 5 10 15

Asp Ala Lys Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu
20 25 30

Lys Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp
35 40 45

Pro Asn Pro Gln Glu Ile Val Leu Gly Asn Val Thr Glu Asn Phe Asn
50 55 60

10

ES 2 824 525 T3

Met Trp Lys Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser
65 70 75 80

Leu Trp Asp Ala Ser Leu Glu Pro Cys Val Lys Leu Thr Pro Leu Cys
85 90 95

Val Thr Leu Asn Cys Arg Asn Val Arg Asn Val Ser Ser Asn Gly Thr
100 105 110

Tyr Asn Ile Ile His Asn Glu Thr Tyr Lys Glu Met Lys Asn Cys Ser
115 120 125

Phe Asn Ala Thr Thr Val Val Glu Asp Arg Lys Gln Lys Val His Ala
130 135 140

Leu Phe Tyr Arg Leu Asp Ile Val Pro Leu Asp Glu Asn Asn Ser Ser
145 150 155 160

Glu Lys Ser Ser Glu Asn Ser Ser Glu Tyr Tyr Arg Leu Ile Asn Cys
165 170 175

Asn Thr Ser Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Asp Pro
180 185 190

Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys
195 200 205

ES 2 824 525 T3

Asn	Asn	Lys	Thr	Phe	Asn	Gly	Thr	Gly	Pro	Cys	Asn	Asn	Val	Ser	Thr
	210					215					220				
Val	Gln	Cys	Thr	His	Gly	Ile	Lys	Pro	Val	Val	Ser	Thr	Gln	Leu	Leu
225					230					235					240
Leu	Asn	Gly	Ser	Leu	Ala	Glu	Glu	Glu	Ile	Ile	Ile	Arg	Ser	Glu	Asn
				245					250					255	
Leu	Thr	Asn	Asn	Ala	Lys	Thr	Ile	Ile	Val	His	Leu	Asn	Glu	Thr	Val
			260					265					270		
Asn	Ile	Thr	Cys	Thr	Arg	Pro	Asn	Asn	Asn	Thr	Arg	Lys	Ser	Ile	Arg
		275					280					285			
Ile	Gly	Pro	Gly	Gln	Thr	Phe	Tyr	Ala	Thr	Gly	Asp	Ile	Ile	Gly	Asp
	290					295					300				
Ile	Arg	Gln	Ala	His	Cys	Asn	Leu	Ser	Arg	Asp	Gly	Trp	Asn	Lys	Thr
305					310					315					320

ES 2 824 525 T3

Leu Gln Gly Val Lys Lys Lys Leu Ala Glu His Phe Pro Asn Lys Thr
 325 330 335

Ile Lys Phe Ala Pro His Ser Gly Gly Asp Leu Glu Ile Thr Thr His
 340 345 350

Thr Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Asn Leu
 355 360 365

Phe Asn Glu Ser Asn Ile Glu Arg Asn Asp Ser Ile Ile Thr Leu Pro
 370 375 380

Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala
 385 390 395 400

Ile Tyr Ala Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser Asn Ile
 405 410 415

Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Ser Asn Asn Gly Val Pro
 420 425 430

ES 2 824 525 T3

Asn Asp Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asn Asn
 435 440 445

Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Val Lys Pro Leu
 450 455 460

Gly Val Ala Pro Thr Glu Ala Lys Arg Arg Val Val Glu Arg Glu Lys
 465 470 475 480

Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly Ala Ala
 485 490 495

Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg
 500 505 510

Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala
 515 520 525

Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly Ile Lys
 530 535 540

Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln Asp Gln
 545 550 555 560

Gln Leu Leu Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr
 565 570 575

Ala Val Pro Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr Asp Ile
 580 585 590

Trp Asp Asn Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly Asn Tyr
 595 600 605

Thr Gly Glu Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln Gln Glu
 610 615 620

Lys
 625

ES 2 824 525 T3

<210> 7

<211> 627

<212> PRT

<213> Secuencia artificial

5

<220>

<223> Secuencia de DS_sC4_SOSIP_E166R

<400> 7

Met Gly Asn Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys
1 5 10 15

Asp Ala Lys Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu
20 25 30

Lys Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp
35 40 45

Pro Asn Pro Gln Glu Ile Val Leu Gly Asn Val Thr Glu Asn Phe Asn
50 55 60

Met Trp Lys Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser
65 70 75 80

Leu Trp Asp Ala Ser Leu Glu Pro Cys Val Lys Leu Thr Pro Leu Cys
85 90 95

Val Thr Leu Asn Cys Arg Asn Val Arg Asn Val Ser Ser Asn Gly Thr
100 105 110

Tyr Asn Ile Ile His Asn Glu Thr Tyr Lys Glu Met Lys Asn Cys Ser
115 120 125

Phe Asn Ala Thr Thr Val Val Arg Asp Arg Lys Gln Lys Val His Ala
130 135 140

10

ES 2 824 525 T3

Leu Phe Tyr Arg Leu Asp Ile Val Pro Leu Asp Glu Asn Asn Ser Ser
 145 150 155 160

Glu Lys Ser Ser Glu Asn Ser Ser Glu Tyr Tyr Arg Leu Ile Asn Cys
 165 170 175

Asn Thr Ser Ala Cys Thr Gln Ala Cys Pro Lys Val Ser Phe Asp Pro
 180 185 190

Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys
 195 200 205

Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn Val Ser Thr
 210 215 220

Val Gln Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu
 225 230 235 240

Leu Asn Gly Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn
 245 250 255

Leu Thr Asn Asn Ala Lys Thr Ile Ile Val His Leu Asn Glu Thr Val
 260 265 270

Asn Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg
 275 280 285

Ile Gly Pro Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp
 290 295 300

Ile Arg Gln Ala His Cys Asn Leu Ser Arg Asp Gly Trp Asn Lys Thr
 305 310 315 320

Leu Gln Gly Val Lys Lys Lys Leu Ala Glu His Phe Pro Asn Lys Thr
 325 330 335

Ile Lys Phe Ala Pro His Ser Gly Gly Asp Leu Glu Ile Thr Thr His
 340 345 350

Thr Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Asn Leu
 355 360 365

Phe Asn Glu Ser Asn Ile Glu Arg Asn Asp Ser Ile Ile Thr Leu Pro
 370 375 380

Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Cys
 385 390 395 400

ES 2 824 525 T3

Ile Tyr Ala Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser Asn Ile
 405 410 415

Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Ser Asn Asn Gly Val Pro
 420 425 430

Asn Asp Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asn Asn
 435 440 445

Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Val Lys Pro Leu
 450 455 460

Gly Val Ala Pro Thr Glu Cys Lys Arg Arg Val Val Glu Arg Arg Arg
 465 470 475 480

Arg Arg Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly
 485 490 495

Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln
 500 505 510

Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu
 515 520 525

Arg Ala Pro Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly
 530 535 540

Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln
 545 550 555 560

Asp Gln Gln Leu Leu Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys
 565 570 575

Cys Thr Ala Val Pro Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr
 580 585 590

Asp Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly
 595 600 605

Asn Tyr Thr Gly Glu Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln
 610 615 620

Gln Glu Lys
 625

<210> 8

<211> 631

<212> PRT

<213> Secuencia artificial

5

<220>

<223> Secuencia de Mos1.Env

<400> 8

ES 2 824 525 T3

Ala Gly Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys
1 5 10 15

Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp
20 25 30

Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp
35 40 45

Pro Asn Pro Gln Glu Val Val Leu Glu Asn Val Thr Glu Asn Phe Asn
50 55 60

Met Trp Lys Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser
65 70 75 80

Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys
85 90 95

Val Thr Leu Asn Cys Thr Asp Asp Val Arg Asn Val Thr Asn Asn Ala
100 105 110

Thr Asn Thr Asn Ser Ser Trp Gly Glu Pro Met Glu Lys Gly Glu Ile
115 120 125

Lys Asn Cys Ser Phe Asn Ile Thr Thr Ser Ile Arg Asn Lys Val Gln
130 135 140

Lys Gln Tyr Ala Leu Phe Tyr Lys Leu Asp Val Val Pro Ile Asp Asn
145 150 155 160

Asp Ser Asn Asn Thr Asn Tyr Arg Leu Ile Ser Cys Asn Thr Ser Val
165 170 175

Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His
180 185 190

Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Lys
 195 200 205

Phe Asn Gly Thr Gly Pro Cys Thr Asn Val Ser Thr Val Gln Cys Thr
 210 215 220

His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser
 225 230 235 240

Leu Ala Glu Glu Glu Val Val Ile Arg Ser Glu Asn Phe Thr Asn Asn
 245 250 255

Ala Lys Thr Ile Met Val Gln Leu Asn Val Ser Val Glu Ile Asn Cys
 260 265 270

Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile His Ile Gly Pro Gly
 275 280 285

Arg Ala Phe Tyr Thr Ala Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala
 290 295 300

His Cys Asn Ile Ser Arg Ala Asn Trp Asn Asn Thr Leu Arg Gln Ile
 305 310 315 320

Val Glu Lys Leu Gly Lys Gln Phe Gly Asn Asn Lys Thr Ile Val Phe
 325 330 335

Asn His Ser Ser Gly Gly Asp Pro Glu Ile Val Met His Ser Phe Asn
 340 345 350

Cys Gly Gly Glu Phe Phe Tyr Cys Asn Ser Thr Lys Leu Phe Asn Ser
 355 360 365

Thr Trp Thr Trp Asn Asn Ser Thr Trp Asn Asn Thr Lys Arg Ser Asn
 370 375 380

Asp Thr Glu Glu His Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile
 385 390 395 400

ES 2 824 525 T3

Asn Met Trp Gln Glu Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg
405 410 415

Gly Gln Ile Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg
420 425 430

Asp Gly Gly Asn Asp Thr Ser Gly Thr Glu Ile Phe Arg Pro Gly Gly
435 440 445

Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val
450 455 460

Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Arg

ES 2 824 525 T3

<223> Secuencia de Mos2.Env

<400> 9

Met Gly Asn Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys
1 5 10 15

Glu Ala Lys Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu
 20 25 30

Lys Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp
 35 40 45

5

ES 2 824 525 T3

Pro Asn Pro Gln Glu Met Val Leu Glu Asn Val Thr Glu Asn Phe Asn
50 55 60

Met Trp Lys Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Arg
65 70 75 80

Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys
85 90 95

Val Thr Leu Glu Cys Arg Asn Val Arg Asn Val Ser Ser Asn Gly Thr
100 105 110

Tyr Asn Ile Ile His Asn Glu Thr Tyr Lys Glu Met Lys Asn Cys Ser
115 120 125

Phe Asn Ala Thr Thr Val Val Glu Asp Arg Lys Gln Lys Val His Ala
130 135 140

Leu Phe Tyr Arg Leu Asp Ile Val Pro Leu Asp Glu Asn Asn Ser Ser
145 150 155 160

Glu Lys Ser Ser Glu Asn Ser Ser Glu Tyr Tyr Arg Leu Ile Asn Cys
165 170 175

Asn Thr Ser Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Asp Pro
180 185 190

Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys
195 200 205

Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn Val Ser Thr
210 215 220

Val Gln Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu
225 230 235 240

Leu Asn Gly Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn
245 250 255

Leu Thr Asn Asn Ala Lys Thr Ile Ile Val His Leu Asn Glu Thr Val
260 265 270

Asn Ile Thr Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg
275 280 285

Ile Gly Pro Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp

ES 2 824 525 T3

290 295 300

Ile Arg Gln Ala His Cys Asn Leu Ser Arg Asp Gly Trp Asn Lys Thr
 305 310 315 320

Leu Gln Gly Val Lys Lys Lys Leu Ala Glu His Phe Pro Asn Lys Thr
 325 330 335

Ile Asn Phe Thr Ser Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His
 340 345 350

Ser Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly Leu
 355 360 365

Phe Asn Gly Thr Tyr Met Pro Asn Gly Thr Asn Ser Asn Ser Ser Ser
 370 375 380

Asn Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln
 385 390 395 400

Glu Val Gly Arg Ala Met Tyr Ala Pro Pro Ile Ala Gly Asn Ile Thr
 405 410 415

Cys Arg Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Ser
 420 425 430

Asn Asn Gly Val Pro Asn Asp Thr Glu Thr Phe Arg Pro Gly Gly Gly
 435 440 445

Asp Met Arg Asn Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val
 450 455 460

Glu Val Lys Pro Leu Gly Val Ala Pro Thr Glu Ala Lys Arg Arg Val
 465 470 475 480

Val Glu Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly
 485 490 495

Ile Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu
 500 505 510

Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser
 515 520 525

Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr
 530 535 540

ES 2 824 525 T3

Val Trp Gly Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg
545 550 555 560

Tyr Leu Gln Asp Gln Gln Leu Leu Gly Leu Trp Gly Cys Ser Gly Lys
565 570 575

Leu Ile Cys Thr Thr Ala Val Pro Trp Asn Thr Ser Trp Ser Asn Lys
580 585 590

Ser Gln Thr Asp Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Lys
595 600 605

Glu Ile Gly Asn Tyr Thr Gly Glu Ile Tyr Arg Leu Leu Glu Glu Ser
610 615 620

Gln Asn Gln Gln Glu Lys
625 630

<210> 10

<211> 6

<212> PRT

5 <213> Secuencia artificial

<220>

<223> secuencia mutante del sitio de clivaje de furina

10 <400> 10

Arg Arg Arg Arg Arg Arg
1 5

<210> 11

<211> 31

<212> PRT

15 <213> Secuencia artificial

<220>

<223> secuencia señal

Asp Glu Asp Pro His Trp Asp Pro
 1 5

<210> 18

<211> 29

<212> PRT

5 <213> Secuencia artificial

<220>

<223> secuencia señal

10 <400> 18

Met Arg Val Lys Gly Ile Arg Lys Asn Tyr Gln His Leu Trp Arg Trp
 1 5 10 15

Gly Thr Met Leu Leu Gly Met Leu Met Ile Cys Ser Ala
 20 25

<210> 19

<211> 61

<212> PRT

15 <213> Secuencia artificial

<220>

<223> secuencia marcadora

20 <400> 19

Ala Ala Ala Leu Pro Glu Thr Gly Gly Gly Ser Asp Tyr Lys Asp Asp
 1 5 10 15

Asp Asp Lys Pro Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
 20 25 30

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
 35 40 45

Gly Ser Gly Gly Gly Gly Ser His His His His His His
 50 55 60

<210> 20

<211> 616

<212> PRT

<213> Secuencia artificial

5

<220>

<223> proteína ConC_SOSIP Env estabilizada (HIV160544)

<400> 20

ES 2 824 525 T3

Asn Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala
 1 5 10 15
 Lys Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu
 20 25 30
 Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn
 35 40 45
 Pro Gln Glu Met Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp
 50 55 60
 Lys Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp
 65 70 75 80
 Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr
 85 90 95
 Leu Asn Cys Thr Asn Val Asn Val Thr Asn Thr Asn Asn Asn Asn Met
 100 105 110
 Lys Glu Glu Met Lys Asn Cys Ser Phe Asn Thr Thr Thr Glu Ile Arg
 115 120 125
 Asp Lys Lys Gln Lys Glu Tyr Ala Leu Phe Tyr Arg Leu Asp Ile Val
 130 135 140
 Pro Leu Asn Glu Asn Ser Ser Glu Tyr Arg Leu Ile Asn Cys Asn Thr
 145 150 155 160
 Ser Thr Ile Thr Gln Ile Cys Pro Lys Val Ser Phe Asp Pro Ile Pro
 165 170 175
 Ile His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn
 180 185 190
 Lys Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn Val Ser Thr Val Gln
 195 200 205
 Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn
 210 215 220
 Gly Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr
 225 230 235 240

ES 2 824 525 T3

Asp Asn Ala Lys Thr Ile Ile Val His Leu Asn Glu Ser Val Glu Ile
 245 250 255
 Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly
 260 265 270
 Pro Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg
 275 280 285
 Gln Ala His Cys Asn Ile Ser Glu Ala Lys Trp Asn Lys Thr Leu Gln
 290 295 300
 Arg Val Lys Lys Lys Leu Lys Glu His Phe Pro Asn Lys Thr Ile Lys
 305 310 315 320
 Phe Ala Pro Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe
 325 330 335
 Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Lys Leu Phe Asn
 340 345 350
 Ser Thr Tyr Asn Asn Thr Thr Ser Asn Ser Thr Ile Thr Leu Pro Cys
 355 360 365
 Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala Met
 370 375 380
 Tyr Ala Pro Pro Ile Ala Gly Asn Ile Thr Cys Lys Ser Asn Ile Thr
 385 390 395 400
 Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn Asn Asn Thr Glu
 405 410 415
 Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu
 420 425 430
 Leu Tyr Lys Tyr Lys Val Val Glu Ile Lys Pro Leu Gly Ile Ala Pro
 435 440 445
 Thr Lys Cys Lys Arg Arg Val Val Glu Arg Arg Arg Arg Arg Ala
 450 455 460
 Val Gly Ile Gly Ala Val Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser
 465 470 475 480
 Thr Met Gly Ala Ala Ser Asn Thr Leu Thr Val Gln Ala Arg Gln Leu

ES 2 824 525 T3

				485						490						495
Leu	Ser	Gly	Ile	Val	Gln	Gln	Gln	Ser	Asn	Leu	Leu	Arg	Ala	Pro	Glu	
			500					505					510			
Ala	Gln	Gln	His	Met	Leu	Gln	Leu	Thr	Val	Trp	Gly	Phe	Lys	Gln	Leu	
			515					520				525				
Gln	Ala	Arg	Val	Leu	Ala	Ile	Glu	Arg	Tyr	Leu	Glu	Val	Gln	Gln	Leu	
	530					535					540					
Leu	Gly	Ile	Trp	Gly	Cys	Ser	Gly	Lys	Leu	Ile	Cys	Cys	Thr	Ala	Val	
545					550					555					560	
Pro	Trp	Asn	Ser	Ser	Trp	Ser	Asn	Lys	Ser	Gln	Glu	Asp	Ile	Trp	Asp	
				565					570					575		
Asn	Met	Thr	Trp	Met	Gln	Trp	Asp	Arg	Glu	Ile	Ser	Asn	Tyr	Thr	Asp	
			580					585					590			
Thr	Ile	Tyr	Arg	Leu	Leu	Glu	Glu	Ser	Gln	Phe	Gln	Gln	Glu	Ile	Asn	
		595					600					605				
Glu	Lys	Asp	Leu	Leu	Ala	Leu	Asp									
	610						615									

<210> 21

<211> 634

<212> PRT

5 <213> Secuencia artificial

<220>

<223> Proteína BG505_SOSIP Env (HIV150673)

10 <400> 21

ES 2 824 525 T3

Ala Glu Asn Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys
1 5 10 15

Asp Ala Glu Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu
20 25 30

Thr Glu Lys His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp
35 40 45

Pro Asn Pro Gln Glu Ile His Leu Glu Asn Val Thr Glu Glu Phe Asn
50 55 60

Met Trp Lys Asn Asn Met Val Glu Gln Met His Thr Asp Ile Ile Ser
65 70 75 80

ES 2 824 525 T3

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

ES 2 824 525 T3

Asn Val Pro Trp Asn Ser Ser Trp Ser Asn Arg Asn Leu Ser Glu Ile
580 585 590

Trp Asp Asn Met Thr Trp Leu Gln Trp Asp Lys Glu Ile Ser Asn Tyr
595 600 605

Thr Gln Ile Ile Tyr Gly Leu Leu Glu Glu Ser Gln Asn Gln Gln Glu
610 615 620

Lys Asn Glu Gln Asp Leu Leu Ala Leu Asp
625 630

<210> 22

<211> 634

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Proteína BG505_SOSIP Env estabilizada (HIV170863)

10

<400> 22

ES 2 824 525 T3

Ala Glu Asn Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys
 1 5 10 15

Asp Ala Glu Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu
 20 25 30

Thr Glu Lys His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp
 35 40 45

Pro Asn Pro Gln Glu Ile His Leu Glu Asn Val Thr Glu Glu Phe Asn
 50 55 60

Met Trp Lys Asn Asn Met Val Glu Gln Met His Thr Asp Ile Ile Ser
 65 70 75 80

Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys
 85 90 95

Val Thr Leu Gln Cys Thr Asn Val Thr Asn Asn Ile Thr Asp Asp Met
 100 105 110

Arg Gly Glu Leu Lys Asn Cys Ser Phe Asn Met Thr Thr Glu Leu Arg
 115 120 125

Asp Lys Lys Gln Lys Val Tyr Ser Leu Phe Tyr Arg Leu Asp Val Val
 130 135 140

Gln Ile Asn Glu Asn Gln Gly Asn Arg Ser Asn Asn Ser Asn Lys Glu

ES 2 824 525 T3

Ala Met Tyr Ala Pro Pro Ile Gln Gly Val Ile Arg Cys Val Ser Asn
 405 410 415

Ile Thr Gly Leu Ile Leu Thr Arg Asp Gly Gly Ser Thr Asn Ser Thr
 420 425 430

Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg
 435 440 445

Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val
 450 455 460

Ala Pro Thr Arg Cys Lys Arg Arg Val Val Gly Arg Arg Arg Arg Arg
 465 470 475 480

Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Phe Leu Gly Ala Ala
 485 490 495

Gly Ser Thr Met Gly Ala Ala Ser Asn Thr Leu Thr Val Gln Ala Arg
 500 505 510

Asn Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Pro Arg Ala
 515 520 525

Pro Glu Ala Gln Gln His Leu Leu Lys Leu Thr Val Trp Gly Ile Lys
 530 535 540

Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu Arg Val Gln
 545 550 555 560

Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Cys Thr
 565 570 575

Asn Val Pro Trp Asn Ser Ser Trp Ser Asn Arg Asn Leu Ser Glu Ile
 580 585 590

Trp Asp Asn Met Thr Trp Leu Gln Trp Asp Lys Glu Ile Ser Asn Tyr
 595 600 605

Thr Gln Ile Ile Tyr Gly Leu Leu Glu Glu Ser Gln Phe Gln Gln Glu
 610 615 620

Ile Asn Glu Gln Asp Leu Leu Ala Leu Asp
 625 630

<210> 23

<211> 619

<212> PRT

<213> Secuencia artificial

5

<220>

<223> Proteína wt C97ZA_SOSIP Env con L535M y Q567K (HIV150673)

<400> 23

ES 2 824 525 T3

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

ES 2 824 525 T3

Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Lys Glu Ile Ile Ile Arg
 225 230 235 240

Ser Glu Asn Leu Thr Asp Asn Val Lys Thr Ile Ile Val His Leu Asn
 245 250 255

Lys Ser Val Glu Ile Val Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys
 260 265 270

Ser Met Arg Ile Gly Pro Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile
 275 280 285

Ile Gly Asp Ile Arg Gln Ala Tyr Cys Asn Ile Ser Gly Ser Lys Trp
 290 295 300

Asn Glu Thr Leu Lys Arg Val Lys Glu Lys Leu Gln Glu Asn Tyr Asn
 305 310 315 320

Asn Asn Lys Thr Ile Lys Phe Ala Pro Ser Ser Gly Gly Asp Leu Glu
 325 330 335

Ile Thr Thr His Ser Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn
 340 345 350

Thr Thr Arg Leu Phe Asn Asn Asn Ala Thr Glu Asp Glu Thr Ile Thr
 355 360 365

Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Gly Val Gly
 370 375 380

Arg Ala Met Tyr Ala Pro Pro Ile Ala Gly Asn Ile Thr Cys Lys Ser
 385 390 395 400

Asn Ile Thr Gly Leu Leu Leu Val Arg Asp Gly Gly Glu Asp Asn Lys
 405 410 415

Thr Glu Glu Ile Phe Arg Pro Gly Gly Gly Asn Met Lys Asp Asn Trp
 420 425 430

Arg Ser Glu Leu Tyr Lys Tyr Lys Val Ile Glu Leu Lys Pro Leu Gly
 435 440 445

Ile Ala Pro Thr Gly Cys Lys Arg Arg Val Val Glu Arg Arg Arg Arg
 450 455 460

Arg Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Phe Leu Gly Ala
 465 470 475 480

ES 2 824 525 T3

Ala Gly Ser Thr Met Gly Ala Ala Ser Met Thr Leu Thr Val Gln Ala
485 490 495

Arg Gln Leu Leu Ser Ser Ile Val Gln Gln Gln Ser Asn Leu Leu Arg
500 505 510

Ala Pro Glu Ala Gln Gln His Met Leu Lys Leu Thr Val Trp Gly Ile
515 520 525

Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Lys Asp
530 535 540

Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Cys
545 550 555 560

Thr Asn Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser Gln Thr Asp
565 570 575

Ile Trp Asn Asn Met Thr Trp Met Glu Trp Asp Arg Glu Ile Ser Asn
580 585 590

Tyr Thr Asp Thr Ile Tyr Arg Leu Leu Glu Asp Ser Gln Thr Gln Gln
595 600 605

Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp
610 615

<210> 24

<211> 619

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Proteína C97ZA_SOSIP Env reparada y estabilizada (HIV170690)

10

<400> 24

ES 2 824 525 T3

Asn Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Asp Ala
1 5 10 15

Lys Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Arg Glu
20 25 30

Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn
35 40 45

Pro Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp
50 55 60

ES 2 824 525 T3

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

ES 2 824 525 T3

Asn Asn Lys Thr Ile Lys Phe Ala Pro Ser Ser Gly Gly Asp Leu Glu
 325 330 335
 Ile Thr Thr His Ser Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn
 340 345 350
 Thr Thr Arg Leu Phe Asn Asn Asn Ala Thr Glu Asn Glu Thr Ile Thr
 355 360 365
 Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly
 370 375 380
 Arg Ala Met Tyr Ala Pro Pro Ile Ala Gly Asn Ile Thr Cys Lys Ser
 385 390 395 400
 Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Glu Asp Asn Lys
 405 410 415
 Thr Glu Glu Ile Phe Arg Pro Gly Gly Gly Asn Met Lys Asp Asn Trp
 420 425 430
 Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Ile Lys Pro Leu Gly
 435 440 445
 Ile Ala Pro Thr Lys Cys Lys Arg Arg Asn Val Thr Arg Arg Arg Arg
 450 455 460
 Arg Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Phe Leu Gly Ala
 465 470 475 480
 Ala Gly Ser Thr Met Gly Ala Ala Ser Asn Thr Leu Thr Val Gln Ala
 485 490 495
 Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Pro Arg
 500 505 510
 Ala Pro Glu Ala Gln Gln His Met Leu Lys Leu Thr Val Trp Gly Ile
 515 520 525
 Lys Gln Leu Gln Ala Arg Val Leu Ala Ile Glu Arg Tyr Leu Glu Val
 530 535 540
 Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Cys
 545 550 555 560
 Thr Asn Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser Gln Thr Asp
 565 570 575

ES 2 824 525 T3

Ile Trp Asn Asn Met Thr Trp Met Glu Trp Asp Arg Glu Ile Ser Asn
580 585 590

Tyr Thr Asp Thr Ile Tyr Arg Leu Leu Glu Asp Ser Gln Phe Gln Gln
595 600 605

Glu Ile Asn Glu Lys Asp Leu Leu Ala Asn Asp
610 615

<210> 25

<211> 628

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Variante del constructo Du422 reparado y estabilizado (HIV161818)

10

<400> 25

ES 2 824 525 T3

Asn Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala
 1 5 10 15

Lys Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Lys Glu
 20 25 30

Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn
 35 40 45

Pro Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp
 50 55 60

Lys Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp
 65 70 75 80

Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr
 85 90 95

Leu Asn Cys Lys Asn Val Asn Ile Ser Ala Asn Ala Asn Ala Thr Ala
 100 105 110

Thr Leu Asn Ser Ser Met Asn Gly Glu Ile Lys Asn Cys Ser Phe Asn
 115 120 125

Thr Thr Thr Glu Leu Arg Asp Lys Lys Gln Lys Val Tyr Ala Leu Phe
 130 135 140

Tyr Lys Pro Asp Val Val Pro Leu Asn Gly Gly Glu His Asn Glu Thr
 145 150 155 160

ES 2 824 525 T3

Gly Glu Tyr Ile Leu Ile Asn Cys Asn Ser Ser Thr Cys Thr Gln Ala
 165 170 175

Cys Pro Lys Val Ser Phe Asp Pro Ile Pro Ile His Tyr Cys Ala Pro
 180 185 190

Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr
 195 200 205

Gly Pro Cys Asn Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys
 210 215 220

Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu
 225 230 235 240

Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr Asn Asn Ile Lys Thr Ile
 245 250 255

Ile Val His Leu Asn Lys Ser Val Glu Ile Asn Cys Thr Arg Pro Asn
 260 265 270

Asn Asn Thr Arg Lys Ser Val Arg Ile Gly Pro Gly Gln Thr Phe Tyr
 275 280 285

Ala Thr Gly Glu Ile Ile Gly Asp Ile Arg Glu Ala His Cys Asn Ile
 290 295 300

Ser Arg Glu Thr Trp Asn Ser Thr Leu Ile Gln Val Lys Glu Lys Leu
 305 310 315 320

Arg Glu His Tyr Asn Lys Thr Ile Lys Phe Glu Pro Ser Ser Gly Gly
 325 330 335

Asp Leu Glu Val Thr Thr His Ser Phe Asn Cys Arg Gly Glu Phe Phe
 340 345 350

Tyr Cys Asn Thr Thr Lys Leu Phe Asn Glu Thr Lys Leu Phe Asn Glu
 355 360 365

Ser Glu Tyr Val Asp Asn Lys Thr Ile Ile Leu Pro Cys Arg Ile Lys
 370 375 380

Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Cys Met Tyr Ala Pro
 385 390 395 400

Pro Ile Glu Gly Asn Ile Thr Cys Lys Ser Asn Ile Thr Gly Leu Leu
 405 410 415

ES 2 824 525 T3

Leu Thr Arg Asp Gly Gly Glu Asn Ser Thr Glu Glu Val Phe Arg Pro
 420 425 430

Gly Gly Gly Asn Met Lys Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr
 435 440 445

Lys Val Val Glu Ile Lys Pro Leu Gly Val Ala Pro Thr Lys Cys Lys
 450 455 460

Arg Lys Asn Val Thr Arg Arg Arg Arg Arg Arg Ala Val Gly Leu Gly
 465 470 475 480

Ala Val Leu Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala
 485 490 495

Ala Ser Asn Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile
 500 505 510

Val Gln Gln Gln Ser Asn Leu Pro Arg Ala Pro Glu Ala Gln Gln His
 515 520 525

Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Thr Arg Val
 530 535 540

Leu Ala Ile Glu Arg Tyr Leu Lys Val Gln Gln Leu Leu Gly Leu Trp
 545 550 555 560

Gly Cys Ser Gly Lys Leu Ile Cys Cys Thr Ala Val Pro Trp Asn Ser
 565 570 575

Ser Trp Ser Asn Lys Ser Leu Gly Asp Ile Trp Asp Asn Met Thr Trp
 580 585 590

Met Gln Trp Asp Arg Glu Ile Ser Asn Tyr Thr Asn Thr Ile Tyr Arg
 595 600 605

Leu Leu Glu Asp Ser Gln Phe Gln Gln Glu Lys Asn Glu Lys Asp Leu
 610 615 620

Leu Ala Asn Asp
 625

<210> 26

<211> 628

<212> PRT

<213> Secuencia artificial

<220>

5 <223> Proteína Du422_SOSIP Env reparada y estabilizada (HIV170859)

<400> 26

ES 2 824 525 T3

Asn Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala
 1 5 10 15
 Lys Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Lys Glu
 20 25 30
 Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn
 35 40 45
 Pro Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp
 50 55 60
 Lys Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp
 65 70 75 80
 Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr
 85 90 95
 Leu Asn Cys Lys Asn Val Asn Ile Ser Ala Asn Ala Asn Ala Thr Ala
 100 105 110
 Thr Leu Asn Ser Ser Met Asn Gly Glu Ile Lys Asn Cys Ser Phe Asn
 115 120 125
 Thr Thr Thr Glu Leu Arg Asp Lys Lys Gln Lys Val Tyr Ala Leu Phe
 130 135 140
 Tyr Lys Pro Asp Val Val Pro Leu Asn Gly Gly Glu His Asn Glu Thr
 145 150 155 160
 Gly Glu Tyr Ile Leu Ile Asn Cys Asn Ser Ser Thr Ile Thr Gln Ala
 165 170 175
 Cys Pro Lys Val Ser Phe Asp Pro Ile Pro Ile His Tyr Cys Ala Pro
 180 185 190
 Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr
 195 200 205
 Gly Pro Cys Asn Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys
 210 215 220
 Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu
 225 230 235 240

ES 2 824 525 T3

Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr Asn Asn Ile Lys Thr Ile
 245 250 255
 Ile Val His Leu Asn Lys Ser Val Glu Ile Asn Cys Thr Arg Pro Asn
 260 265 270
 Asn Asn Thr Arg Lys Ser Val Arg Ile Gly Pro Gly Gln Thr Phe Tyr
 275 280 285
 Ala Thr Gly Glu Ile Ile Gly Asp Ile Arg Glu Ala His Cys Asn Ile
 290 295 300
 Ser Arg Glu Thr Trp Asn Ser Thr Leu Ile Gln Val Lys Glu Lys Leu
 305 310 315 320
 Arg Glu His Tyr Asn Lys Thr Ile Lys Phe Glu Pro Ser Ser Gly Gly
 325 330 335
 Asp Leu Glu Val Thr Thr His Ser Phe Asn Cys Arg Gly Glu Phe Phe
 340 345 350
 Tyr Cys Asn Thr Thr Lys Leu Phe Asn Glu Thr Lys Leu Phe Asn Glu
 355 360 365
 Ser Glu Tyr Val Asp Asn Lys Thr Ile Ile Leu Pro Cys Arg Ile Lys
 370 375 380
 Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala Met Tyr Ala Pro
 385 390 395 400
 Pro Ile Glu Gly Asn Ile Thr Cys Lys Ser Asn Ile Thr Gly Leu Leu
 405 410 415
 Leu Thr Arg Asp Gly Gly Glu Asn Ser Thr Glu Glu Val Phe Arg Pro
 420 425 430
 Gly Gly Gly Asn Met Lys Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr
 435 440 445
 Lys Val Val Glu Ile Lys Pro Leu Gly Val Ala Pro Thr Lys Cys Lys
 450 455 460
 Arg Lys Val Val Gly Arg Arg Arg Arg Arg Ala Val Gly Leu Gly
 465 470 475 480
 Ala Val Leu Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala

ES 2 824 525 T3

Met Gly Asn Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys
1 5 10 15

Asp Ala Lys Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu
20 25 30

Lys Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp
35 40 45

Pro Asn Pro Gln Glu Ile Val Leu Gly Asn Val Thr Glu Asn Phe Asn
50 55 60

ES 2 824 525 T3

Met Trp Lys Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser
65 70 75 80

Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys
85 90 95

Val Thr Leu Asn Cys Arg Asn Val Arg Asn Val Glu Met Lys Asn Cys
100 105 110

Ser Phe Asn Ala Thr Thr Val Val Arg Asp Arg Lys Gln Lys Val His
115 120 125

Ala Leu Phe Tyr Arg Leu Asp Ile Val Pro Leu Asp Glu Asn Asn Ser
130 135 140

Ser Tyr Arg Leu Ile Asn Cys Asn Thr Ser Ala Cys Thr Gln Ile Cys
145 150 155 160

Pro Lys Val Ser Phe Asp Pro Ile Pro Ile His Tyr Cys Ala Pro Ala
165 170 175

Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr Gly
180 185 190

Pro Cys Asn Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys Pro
195 200 205

Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu
210 215 220

Ile Ile Ile Arg Ser Glu Asn Leu Thr Asn Asn Ala Lys Thr Ile Ile
225 230 235 240

Val His Leu Asn Glu Thr Val Asn Ile Asn Cys Thr Arg Pro Asn Asn
245 250 255

Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro Gly Gln Thr Phe Tyr Ala
260 265 270

Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Leu Ser
275 280 285

Arg Asp Gly Trp Asn Lys Thr Leu Gln Gly Val Lys Lys Lys Leu Ala
290 295 300

Glu His Phe Pro Asn Lys Thr Ile Lys Phe Ala Pro His Ser Gly Gly

ES 2 824 525 T3

Lys Ser Gln Thr Asp Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp
565 570 575

Lys Glu Ile Gly Asn Tyr Thr Gly Glu Ile Tyr Arg Leu Leu Glu Glu
580 585 590

Ser Gln Phe Gln Gln Glu Ile
595

<210> 28

<211> 607

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Proteína ConC_SOSIP.v3 Env estabilizada (HIV170654)

10

<400> 28

ES 2 824 525 T3

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

ES 2 824 525 T3

Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Ile
 420 425 430

Lys Pro Leu Gly Ile Ala Pro Thr Lys Cys Lys Arg Arg Val Val Glu
 435 440 445

Arg Arg Arg Arg Arg Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly
 450 455 460

Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Asn Thr Leu
 465 470 475 480

Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser
 485 490 495

Asn Leu Leu Arg Ala Pro Glu Ala Gln Gln His Met Leu Gln Leu Thr
 500 505 510

Val Trp Gly Phe Lys Gln Leu Gln Ala Arg Val Leu Ala Ile Glu Arg
 515 520 525

Tyr Leu Glu Val Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys
 530 535 540

Leu Ile Cys Cys Thr Ala Val Pro Trp Asn Ser Ser Trp Ser Asn Lys
 545 550 555 560

Ser Gln Glu Asp Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Arg
 565 570 575

Glu Ile Ser Asn Tyr Thr Asp Thr Ile Tyr Arg Leu Leu Glu Glu Ser
 580 585 590

Gln Phe Gln Gln Glu Ile Asn Glu Lys Asp Leu Leu Ala Leu Asp
 595 600 605

<210> 29

<211> 634

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Proteína BG505_SOSIP.v2 Env estabilizada (HIV171814)

5

<400> 29

Ala Glu Asn Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys
1 5 10 15

Asp Ala Glu Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu

ES 2 824 525 T3

Thr Arg Lys Ser Ile Arg Ile Gly Pro Gly Gln Ala Phe Tyr Ala Thr
 275 280 285
 Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Val Ser Lys
 290 295 300
 Ala Thr Trp Asn Glu Thr Leu Gly Lys Val Val Lys Gln Leu Arg Lys
 305 310 315 320
 His Phe Gly Asn Asn Thr Ile Ile Arg Phe Ala Asn Ser Ser Gly Gly
 325 330 335
 Asp Leu Glu Val Thr Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe
 340 345 350
 Tyr Cys Asn Thr Ser Gly Leu Phe Asn Ser Thr Trp Ile Ser Asn Thr
 355 360 365
 Ser Val Gln Gly Ser Asn Ser Thr Gly Ser Asn Asp Ser Ile Thr Leu
 370 375 380
 Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Arg Ile Gly Gln
 385 390 395 400
 Cys Met Tyr Ala Pro Pro Ile Gln Gly Val Ile Arg Cys Val Ser Asn
 405 410 415
 Ile Thr Gly Leu Ile Leu Thr Arg Asp Gly Gly Ser Thr Asn Ser Thr
 420 425 430
 Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg
 435 440 445
 Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val
 450 455 460
 Ala Pro Thr Arg Cys Lys Arg Arg Val Val Gly Arg Arg Arg Arg Arg
 465 470 475 480
 Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Phe Leu Gly Ala Ala
 485 490 495
 Gly Ser Thr Met Gly Ala Ala Ser Asn Thr Leu Thr Val Gln Ala Arg
 500 505 510
 Asn Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Pro Arg Ala
 515 520 525

ES 2 824 525 T3

Pro Glu Ala Gln Gln His Leu Leu Lys Leu Thr Val Trp Gly Ile Lys
530 535 540

Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu Glu Val Gln
545 550 555 560

Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Cys Thr
565 570 575

Asn Val Pro Trp Asn Ser Ser Trp Ser Asn Arg Asn Leu Ser Glu Ile
580 585 590

Trp Asp Asn Met Thr Trp Leu Gln Trp Asp Lys Glu Ile Ser Asn Tyr
595 600 605

Thr Gln Ile Ile Tyr Gly Leu Leu Glu Glu Ser Gln Phe Gln Gln Glu
610 615 620

Ile Asn Glu Val Asp Leu Leu Ala Leu Asp
625 630

<210> 30

<211> 619

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Proteína C97ZA_SOSIP.v2 Env reparada y estabilizada (HIV171810)

10

<400> 30

ES 2 824 525 T3

Asn Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Asp Ala
 1 5 10 15

Lys Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Arg Glu
 20 25 30

Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn
 35 40 45

Pro Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp
 50 55 60

Lys Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp
 65 70 75 80

Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr
 85 90 95

ES 2 824 525 T3

Leu His Cys Thr Asn Ala Thr Phe Lys Asn Asn Val Thr Asn Asp Met
 100 105 110

Asn Lys Glu Ile Arg Asn Cys Ser Phe Asn Thr Thr Thr Glu Ile Arg
 115 120 125

Asp Lys Lys Gln Lys Val Tyr Ala Leu Phe Tyr Arg Leu Asp Ile Val
 130 135 140

Gln Leu Lys Glu Asn Arg Asn Asn Ser Asn Asn Ser Glu Tyr Arg Leu
 145 150 155 160

Ile Asn Cys Asn Thr Ser Thr Cys Thr Gln Ile Cys Pro Lys Val Thr
 165 170 175

Phe Asp Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile
 180 185 190

Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn
 195 200 205

Val Ser Thr Val Gln Cys Thr His Gly Ile Lys Pro Val Val Ser Thr
 210 215 220

Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Lys Glu Ile Ile Ile Arg
 225 230 235 240

Ser Glu Asn Leu Thr Asp Asn Val Lys Thr Ile Ile Val His Leu Asn
 245 250 255

Lys Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys
 260 265 270

Ser Ile Arg Ile Gly Pro Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile
 275 280 285

Ile Gly Asp Ile Arg Gln Ala Tyr Cys Asn Ile Ser Gly Ser Lys Trp
 290 295 300

Asn Glu Thr Leu Lys Arg Val Lys Glu Lys Leu Arg Glu His Phe Asn
 305 310 315 320

Asn Asn Lys Thr Ile Lys Phe Ala Pro Ser Ser Gly Gly Asp Leu Glu
 325 330 335

Ile Thr Thr His Ser Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn
 340 345 350

ES 2 824 525 T3

Thr Thr Arg Leu Phe Asn Asn Asn Ala Thr Glu Asn Glu Thr Ile Thr
 355 360 365
 Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly
 370 375 380
 Arg Cys Met Tyr Ala Pro Pro Ile Ala Gly Asn Ile Thr Cys Lys Ser
 385 390 395 400
 Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Glu Asp Asn Lys
 405 410 415
 Thr Glu Glu Ile Phe Arg Pro Gly Gly Gly Asn Met Lys Asp Asn Trp
 420 425 430
 Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Ile Lys Pro Leu Gly
 435 440 445
 Ile Ala Pro Thr Lys Cys Lys Arg Arg Asn Val Thr Arg Arg Arg Arg
 450 455 460
 Arg Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Phe Leu Gly Ala
 465 470 475 480
 Ala Gly Ser Thr Met Gly Ala Ala Ser Asn Thr Leu Thr Val Gln Ala
 485 490 495
 Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Pro Arg
 500 505 510
 Ala Pro Glu Ala Gln Gln His Met Leu Lys Leu Thr Val Trp Gly Ile
 515 520 525
 Lys Gln Leu Gln Ala Arg Val Leu Ala Ile Glu Arg Tyr Leu Glu Val
 530 535 540
 Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Cys
 545 550 555 560
 Thr Asn Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser Gln Thr Asp
 565 570 575
 Ile Trp Asn Asn Met Thr Trp Met Glu Trp Asp Arg Glu Ile Ser Asn
 580 585 590
 Tyr Thr Asp Thr Ile Tyr Arg Leu Leu Glu Asp Ser Gln Phe Gln Gln
 595 600 605

Glu Ile Asn Glu Val Asp Leu Leu Ala Asn Asp
610 615

<210> 31

<211> 628

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Proteína Du422_SOSIP.vl Env reparada y estabilizada (HIV171812)

10

<400> 31

ES 2 824 525 T3

Asn Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala
 1 5 10 15

Lys Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Lys Glu
 20 25 30

Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn
 35 40 45

Pro Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp
 50 55 60

Lys Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp
 65 70 75 80

Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr
 85 90 95

Leu Asn Cys Lys Asn Val Asn Ile Ser Ala Asn Ala Asn Ala Thr Ala
 100 105 110

Thr Leu Asn Ser Ser Met Asn Gly Glu Ile Lys Asn Cys Ser Phe Asn
 115 120 125

Thr Thr Thr Glu Leu Arg Asp Lys Lys Gln Lys Val Tyr Ala Leu Phe
 130 135 140

Tyr Lys Pro Asp Val Val Pro Leu Asn Gly Gly Glu His Asn Glu Thr
 145 150 155 160

Gly Glu Tyr Ile Leu Ile Asn Cys Asn Ser Ser Thr Cys Thr Gln Ala
 165 170 175

Cys Pro Lys Val Ser Phe Asp Pro Ile Pro Ile His Tyr Cys Ala Pro
 180 185 190

ES 2 824 525 T3

Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr
195 200 205

Gly Pro Cys Asn Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys
210 215 220

Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu
225 230 235 240

Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr Asn Asn Ile Lys Thr Ile
245 250 255

Ile Val His Leu Asn Lys Ser Val Glu Ile Asn Cys Thr Arg Pro Asn
260 265 270

Asn Asn Thr Arg Lys Ser Val Arg Ile Gly Pro Gly Gln Thr Phe Tyr
275 280 285

Ala Thr Gly Glu Ile Ile Gly Asp Ile Arg Glu Ala His Cys Asn Ile
290 295 300

Ser Arg Glu Thr Trp Asn Ser Thr Leu Ile Gln Val Lys Glu Lys Leu
305 310 315 320

Arg Glu His Tyr Asn Lys Thr Ile Lys Phe Glu Pro Ser Ser Gly Gly
325 330 335

Asp Leu Glu Val Thr Thr His Ser Phe Asn Cys Arg Gly Glu Phe Phe
340 345 350

Tyr Cys Asn Thr Thr Lys Leu Phe Asn Glu Thr Lys Leu Phe Asn Glu
355 360 365

Ser Glu Tyr Val Asp Asn Lys Thr Ile Ile Leu Pro Cys Arg Ile Lys
370 375 380

Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Cys Met Tyr Ala Pro
385 390 395 400

Pro Ile Glu Gly Asn Ile Thr Cys Lys Ser Asn Ile Thr Gly Leu Leu
405 410 415

Leu Thr Arg Asp Gly Gly Glu Asn Ser Thr Glu Glu Val Phe Arg Pro
420 425 430

Gly Gly Gly Asn Met Lys Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr
435 440 445

ES 2 824 525 T3

Lys Val Val Glu Ile Lys Pro Leu Gly Val Ala Pro Thr Lys Cys Lys
 450 455 460

Arg Lys Val Val Gly Arg Arg Arg Arg Arg Arg Ala Val Gly Leu Gly
 465 470 475 480

Ala Val Leu Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala
 485 490 495

Ala Ser Asn Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile
 500 505 510

Val Gln Gln Gln Ser Asn Leu Pro Arg Ala Pro Glu Ala Gln Gln His
 515 520 525

Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Thr Arg Val
 530 535 540

Leu Ala Ile Glu Arg Tyr Leu Glu Val Gln Gln Leu Leu Gly Leu Trp
 545 550 555 560

Gly Cys Ser Gly Lys Leu Ile Cys Cys Thr Ala Val Pro Trp Asn Ser
 565 570 575

Ser Trp Ser Asn Lys Ser Leu Gly Asp Ile Trp Asp Asn Met Thr Trp
 580 585 590

Met Gln Trp Asp Arg Glu Ile Ser Asn Tyr Thr Asn Thr Ile Tyr Arg
 595 600 605

Leu Leu Glu Asp Ser Gln Phe Gln Gln Glu Ile Asn Glu Val Asp Leu
 610 615 620

Leu Ala Leu Asp
 625

<210> 32

<211> 599

<212> PRT

<213> Secuencia artificial

<220>

<223> Proteína sC4_SOSIP.v4 Env estabilizada y reparada

5

<400> 32

Met	Gly	Asn	Leu	Trp	Val	Thr	Val	Tyr	Tyr	Gly	Val	Pro	Val	Trp	Lys
1				5					10					15	

ES 2 824 525 T3

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

ES 2 824 525 T3

Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Leu Ser
 275 280 285

Arg Asp Gly Trp Asn Lys Thr Leu Gln Gly Val Lys Lys Lys Leu Ala
 290 295 300

Glu His Phe Pro Asn Lys Thr Ile Lys Phe Ala Pro His Ser Gly Gly
 305 310 315 320

Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys Arg Gly Glu Phe Phe
 325 330 335

Tyr Cys Asn Thr Ser Asn Leu Phe Asn Glu Ser Asn Ile Glu Arg Asn
 340 345 350

Asp Ser Ile Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met
 355 360 365

Trp Gln Glu Val Gly Arg Cys Met Tyr Ala Pro Pro Ile Ala Gly Asn
 370 375 380

Ile Thr Cys Arg Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly
 385 390 395 400

Gly Ser Asn Asn Asn Asp Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp
 405 410 415

Met Arg Asn Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu
 420 425 430

Val Lys Pro Leu Gly Val Ala Pro Thr Glu Cys Lys Arg Arg Asn Val
 435 440 445

Thr Arg Arg Arg Arg Arg Arg Ala Val Gly Ile Gly Ala Val Phe Leu
 450 455 460

Gly Ile Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Asn Thr
 465 470 475 480

Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln
 485 490 495

Ser Asn Leu Pro Arg Ala Pro Glu Ala Gln Gln His Met Leu Gln Leu
 500 505 510

Thr Val Trp Gly Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu

ES 2 824 525 T3

515

520

525

Arg Tyr Leu Glu Asp Gln Gln Leu Leu Gly Leu Trp Gly Cys Ser Gly
 530 535 540

Lys Leu Ile Cys Cys Thr Ala Val Pro Trp Asn Thr Ser Trp Ser Asn
 545 550 555 560

Lys Ser Gln Thr Asp Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp
 565 570 575

Lys Glu Ile Gly Asn Tyr Thr Gly Glu Ile Tyr Arg Leu Leu Glu Glu
 580 585 590

Ser Gln Phe Gln Gln Glu Ile
 595

<210> 33

<211> 29

5 <212> PRT

<213> Secuencia artificial

<220>

<223> secuencia señal

10

<400> 33

Met Arg Val Arg Gly Met Leu Arg Asn Trp Gln Gln Trp Trp Ile Trp
 1 5 10 15

Ser Ser Leu Gly Phe Trp Met Leu Met Ile Tyr Ser Val
 20 25

<210> 34

<211> 29

15 <212> PRT

<213> Secuencia artificial

REIVINDICACIONES

1. Una proteína de la envoltura (Env) del virus de inmunodeficiencia humana (VIH) recombinante, que comprende dos o más de los siguientes residuos de aminoácidos:
- (i) Phe, Leu, Met o Trp, preferiblemente Phe, en la posición 651;
- 5 (ii) Phe, Ile, Met o Trp, preferiblemente Ile, en la posición 655;
- (iii) Asn o Gln, preferiblemente Asn, en la posición 535;
 - (iv) Val, Ile o Ala, preferiblemente Val, en la posición 589;
 - (v) Phe o Trp, preferiblemente Phe, en la posición 573;
 - (vi) Ile en la posición 204; y/o
- 10 (vii) Phe, Met o Ile, preferiblemente Phe, en la posición 647,
- en donde la numeración de las posiciones es acorde con la numeración en la gp160 de la forma aislada HXB2 del VIH-1.
2. Una proteína Env del VIH recombinante, que comprende uno o más de los siguientes residuos de aminoácidos:
- (i) Phe, Leu, Met o Trp, preferiblemente Phe, en la posición 651;
 - (ii) Phe, Ile, Met o Trp, preferiblemente Ile, en la posición 655;
- 15 (iii) Asn o Gln, preferiblemente Asn, en la posición 535;
- (iv) Val, Ile o Ala, preferiblemente Val, en la posición 589;
 - (v) Phe o Trp, preferiblemente Phe, en la posición 573;
 - (vi) Ile en la posición 204; y/o
 - (vii) Phe, Met o Ile, preferiblemente Phe, en la posición 647,
- 20 en donde la proteína Env del VIH se selecciona del grupo que consiste en:
- (1) una secuencia consenso de la Env del VIH, por ejemplo del clado C, por ejemplo que comprende la secuencia de aminoácidos de la SEQ ID N°: 2 o 3 o por ejemplo del clado B, por ejemplo que comprende la secuencia de aminoácidos de la SEQ ID N°: 4 o 5;
 - (2) una proteína Env del VIH sintética, por ejemplo que comprende la secuencia de aminoácidos de (a): SEQ ID N°: 6; o (b): SEQ ID N°: 6 con una mutación de Glu por Arg en la posición 166; o (c): (a) o (b) con una mutación de los aminoácidos en las posiciones 501 y 605 por residuos Cys y una mutación del aminoácido en la posición 559 por un residuo Pro; o (d): (a), (b) o (c) que tienen una mutación adicional en el sitio de clivaje de furina, por ejemplo el reemplazo de los aminoácidos en las posiciones 508-511 por RRRRRR (SEQ ID N°: 10); o (e) SEQ ID N°: 7; o (f) SEQ ID N°: 8 o SEQ ID N°: 9; y
 - (3) una proteína Env del VIH parental que preferiblemente es una proteína Env del VIH de tipo salvaje, preferiblemente del clado C, que comprende por lo menos una mutación de reparación en un residuo de aminoácido que está presente en la posición correspondiente a una frecuencia menor que 7.5%, preferiblemente menor que 2%, de las secuencias Env del VIH en una colección de por lo menos 1000, preferiblemente por lo menos 10000, secuencias Env del VIH de tipo salvaje, en donde la mutación de reparación es una sustitución por un residuo de aminoácido que está presente en la posición correspondiente a una frecuencia de por lo menos 10% de las secuencias Env del VIH en dicha colección y preferiblemente
- 30 la mutación de reparación es una sustitución por el residuo de aminoácido que está presente en la posición correspondiente con mayor frecuencia en dicha colección;
- 35
- y
- en donde la numeración de las posiciones es acorde con la numeración en la gp160 de la forma aislada HXB2 del VIH-1.
3. Una proteína Env del VIH recombinante, que comprende uno o más de los siguientes residuos de aminoácidos:
- (i) Phe, Leu, Met o Trp, preferiblemente Phe, en la posición 651;
- 40 (ii) Phe, Ile, Met o Trp, preferiblemente Ile, en la posición 655;
- (iii) Asn o Gln, preferiblemente Asn, en la posición 535;

- (iv) Val, Ile o Ala, preferiblemente Val, en la posición 589;
 - (v) Phe o Trp, preferiblemente Phe, en la posición 573;
 - (vi) Ile en la posición 204; y/o
 - (vii) Phe, Met o Ile, preferiblemente Phe, en la posición 647,
- 5 en donde la proteína Env del VIH es una proteína Env del VIH que comprende por lo menos uno de los siguientes:
- (a) Cys en las posiciones 501 y 605;
 - (b) Pro en la posición 559;
 - (c) Cys en las posiciones 501 y 605 y Pro en la posición 559; y
- la numeración de las posiciones es acorde con la numeración en la gp160 de la forma aislada HXB2 del VIH-1.
- 10 4. La proteína Env del VIH recombinante de la reivindicación 2 o 3, que comprende dos o más de los residuos de aminoácidos indicados en (i) a (vii).
5. La proteína Env del VIH recombinante de cualquiera de las reivindicaciones 1, 2 y 4, que comprende Cys en las posiciones 501 y 605 o Pro en la posición 559, preferiblemente Cys en las posiciones 501 y 605 y Pro en la posición 559.
- 15 6. La proteína Env del VIH recombinante de cualquiera de las reivindicaciones 1 a 5, que comprende tres o más de los residuos de aminoácidos indicados en (i) a (vii).
7. La proteína Env del VIH recombinante de cualquiera de las reivindicaciones 1 a 5, que comprende cuatro o más de los residuos de aminoácidos indicados en (i) a (vii).
8. La proteína Env del VIH recombinante de acuerdo con cualquiera de las reivindicaciones 1 a 7, que comprende Phe en la posición 651, Ile en la posición 655, Asn en la posición 535 y Val en la posición 589.
- 20 9. La proteína Env del VIH recombinante de cualquiera de las reivindicaciones 1 a 8, que comprende cinco o más de los residuos de aminoácidos indicados en (i) a (vii).
10. La proteína Env del VIH recombinante de cualquiera de las reivindicaciones 1 a 9, que además comprende uno o más de los siguientes:
- (viii) Gln, Glu, Ile, Met, Val, Trp o Phe, preferiblemente Gln o Glu, en la posición 588;
- 25 (ix) Lys en la posición 64 o Arg en la posición 66 o Lys en la posición 64 y Arg en la posición 66;
- (x) Trp en la posición 316;
 - (xi) Cys en ambas posiciones 201 y 433;
 - (xii) Pro en la posición 556 o 558 o en ambas posiciones 556 y 558;
- 30 (xiii) reemplazo del bucle en las posiciones de aminoácido 548-568 (bucle HR1) por un bucle de 7-10 aminoácidos, preferiblemente un bucle de 8 aminoácidos, por ejemplo que tiene una secuencia seleccionada entre cualquiera de las (SEQ ID N°: 12-17);
- (xiv) Gly en la posición 568 o Gly en la posición 569 o Gly en la posición 636 o Gly en ambas posiciones 568 y 636 o Gly en ambas posiciones 569 y 636; y/o
- 35 (xv) Tyr en la posición 302 o Arg en la posición 519 o Arg en la posición 520 o Tyr en la posición 302 y Arg en la posición 519 o Tyr en la posición 302 y Arg en la posición 520 o Tyr en la posición 302 y Arg en ambas posiciones 519 y 520.
11. La proteína Env del VIH recombinante de cualquiera de las reivindicaciones 1 a 10, que además comprende una mutación en una secuencia de clivaje de furina de la proteína Env del VIH, preferiblemente un reemplazo en las posiciones 508-511 por RRRRRR (SEQ ID N°: 10).
- 40 12. La proteína Env del VIH recombinante de cualquiera de las reivindicaciones 1 a 11, que comprende una secuencia de aminoácidos que es por lo menos 95% idéntica a cualquiera de las SEQ ID N°: 3, 5, 20, 22, 24, 26, 27, 28, 29, 30, 31, o 32.

13. La proteína Env del VIH recombinante de cualquiera de las reivindicaciones 1 a 12, que además comprende: (xvi) un residuo de aminoácido seleccionado entre Val, Ile, Phe, Met, Ala o Leu, preferiblemente Val o Ile, más preferiblemente Val, en la posición 658.
- 5 14. La proteína Env del VIH recombinante de cualquiera de las reivindicaciones 1 a 13, que es una proteína gp140 o gp160.
15. La proteína Env del VIH recombinante de cualquiera de las reivindicaciones 1 a 14, que es del clado C o del clado A, preferiblemente del clado C.
16. Un complejo trimérico que comprende un oligómero no covalente de tres de las proteínas Env del VIH recombinantes de cualquiera de las reivindicaciones 1 a 15.
- 10 17. Una partícula, preferiblemente un liposoma o una nanopartícula, que expresa sobre su superficie una proteína Env del VIH recombinante de cualquiera de las reivindicaciones 1 a 15 o un complejo trimérico de la reivindicación 16.
18. Una molécula de ácido nucleico aislada que codifica una proteína Env del VIH recombinante de cualquiera de las reivindicaciones 1 a 15.
- 15 19. Un vector que comprende la molécula de ácido nucleico aislada de la reivindicación 18 ligada operativamente a un promotor.
20. El vector de la reivindicación 19, que es un vector de adenovirus.
21. Una célula huésped que comprende la molécula de ácido nucleico aislada de la reivindicación 18 o el vector de la reivindicación 19 o 20.
- 20 22. Un método para producir una proteína Env del VIH recombinante, que comprende cultivar la célula huésped de la reivindicación 21 en condiciones adecuadas para la producción de la proteína Env del VIH recombinante.
23. Una composición que comprende la proteína Env del VIH recombinante de cualquiera de las reivindicaciones 1 a 15, el complejo trimérico de la reivindicación 16, la partícula de la reivindicación 17, la molécula de ácido nucleico aislada de la reivindicación 18 o el vector de la reivindicación 19 o 20 y un vehículo farmacéuticamente aceptable.
- 25 24. Un método para mejorar la formación de trímeros de una proteína Env del VIH, donde dicho método comprende sustituir uno o más residuos de aminoácidos en una proteína Env del VIH parental, en donde dichas una o más sustituciones dan como resultado uno o más de los siguientes aminoácidos:
- (i) Phe, Leu, Met o Trp, preferiblemente Phe, en la posición 651;
- (ii) Phe, Ile, Met o Trp, preferiblemente Ile, en la posición 655;
- (iii) Asn o Gln, preferiblemente Asn, en la posición 535;
- 30 (iv) Val, Ile o Ala, preferiblemente Val, en la posición 589;
- (v) Phe o Trp, preferiblemente Phe, en la posición 573;
- (vi) Ile en la posición 204; y/o
- (vii) Phe, Met o Ile, preferiblemente Phe, en la posición 647,
- en donde la numeración de las posiciones es acorde con la numeración en la gp160 de la forma aislada HXB2 del VIH-1.

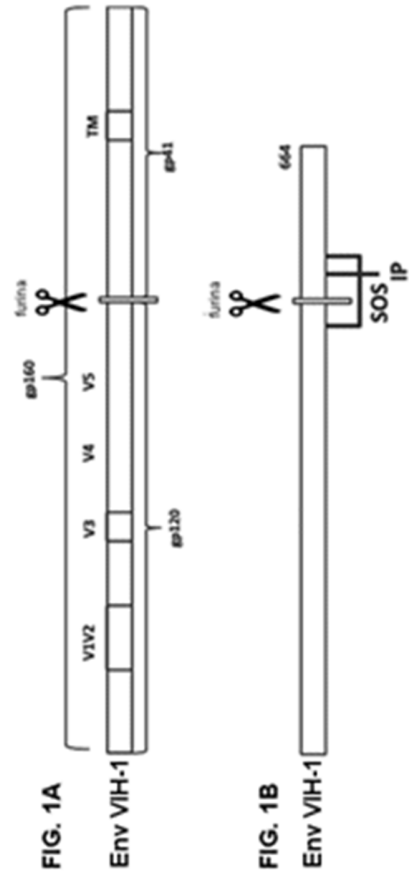


Fig. 1

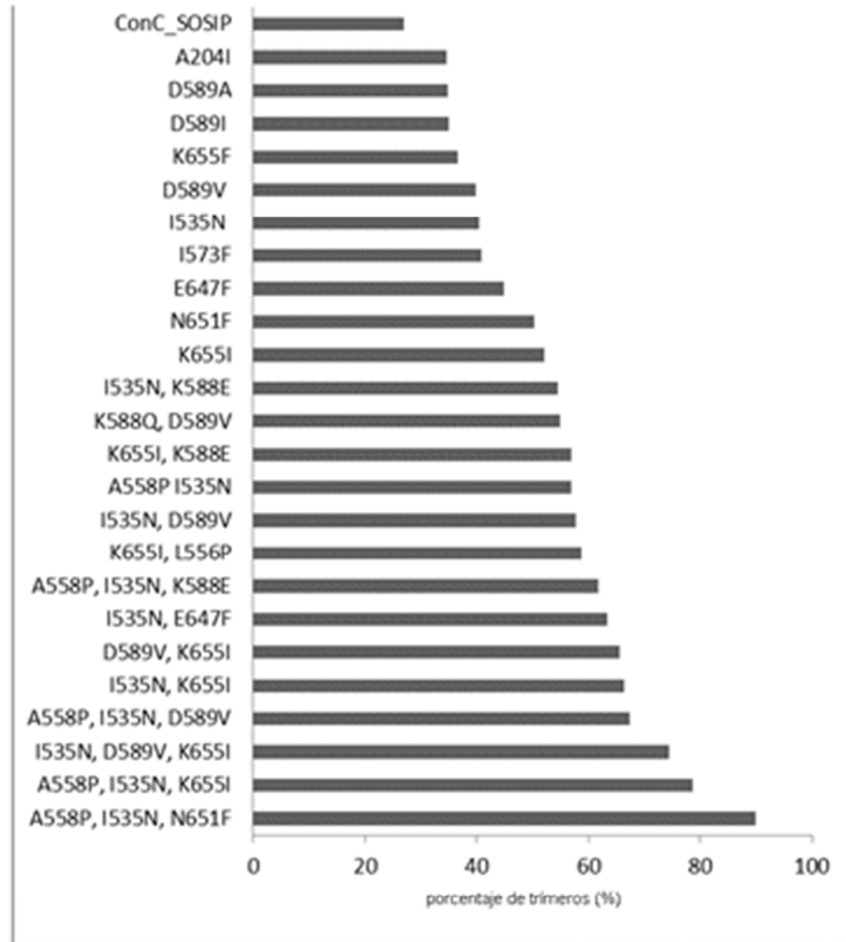


Fig. 2A

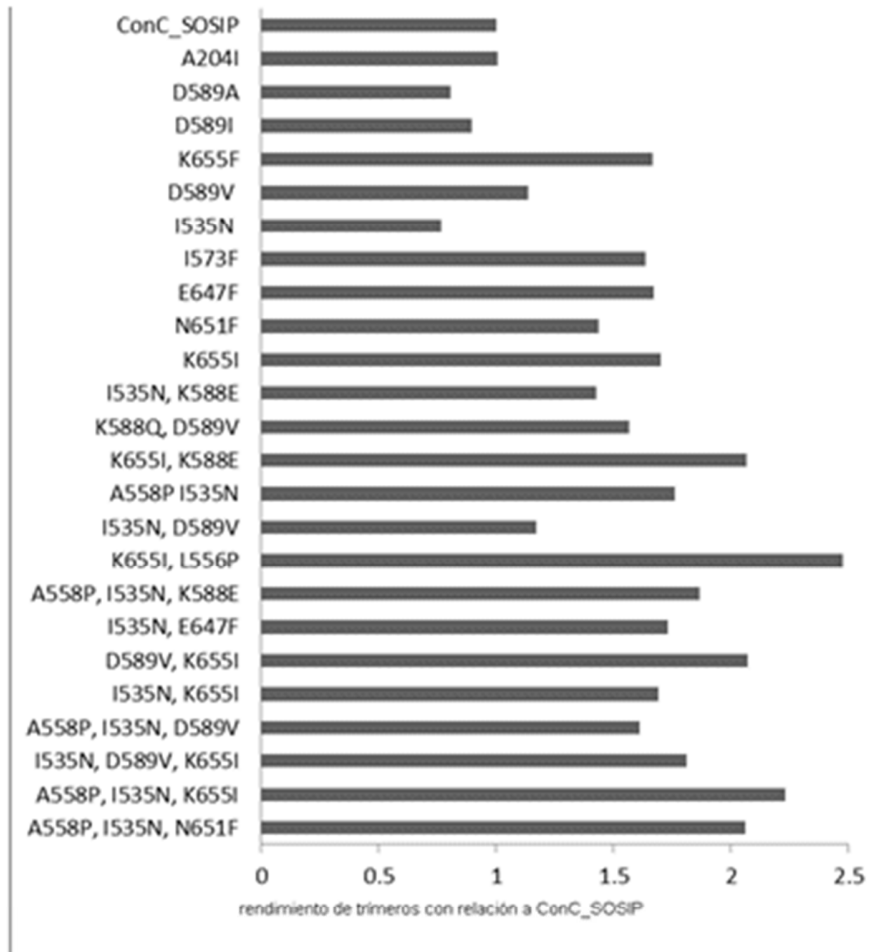


Fig. 2B

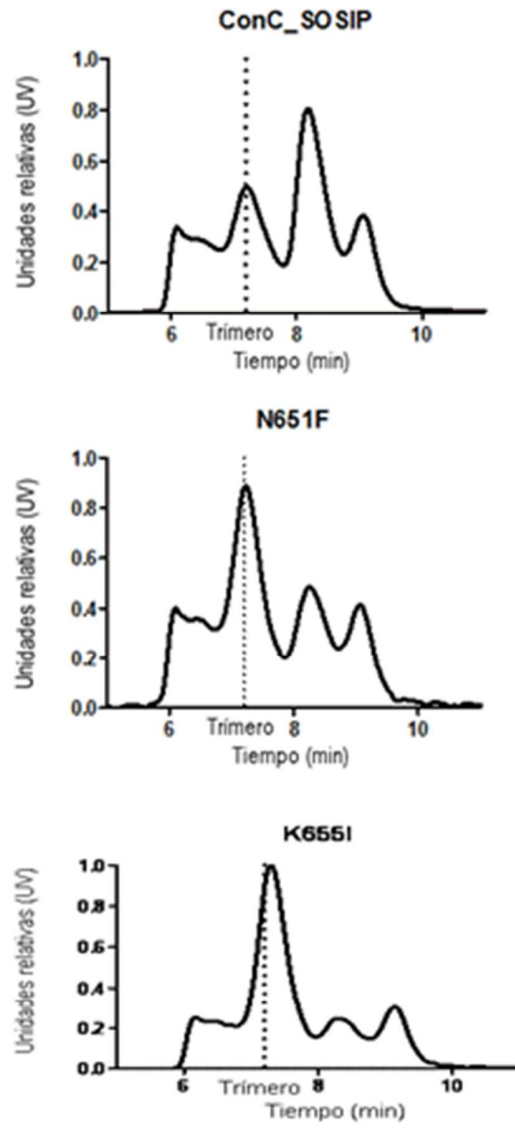


Fig. 3

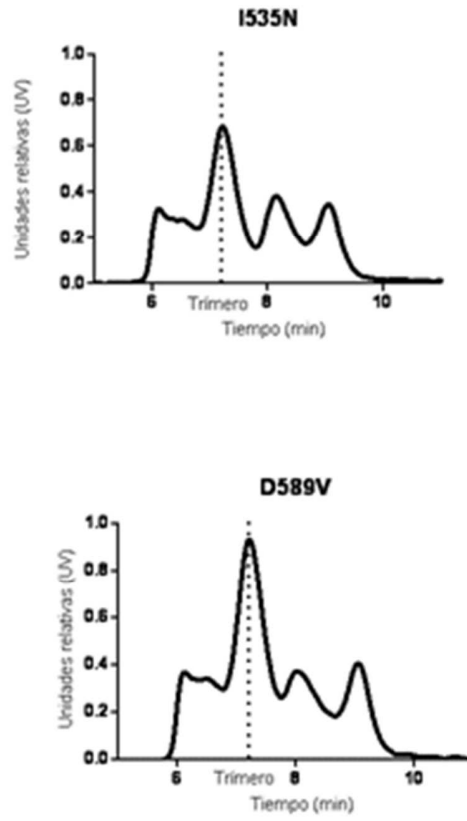


Fig. 3, continuación

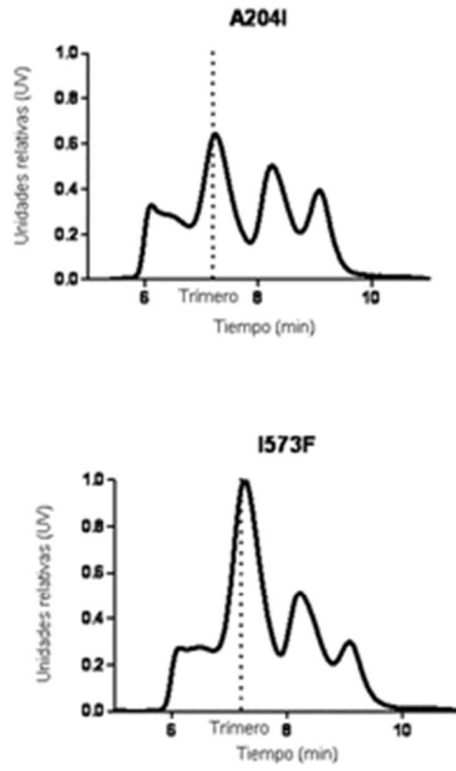


Fig. 3, continuación

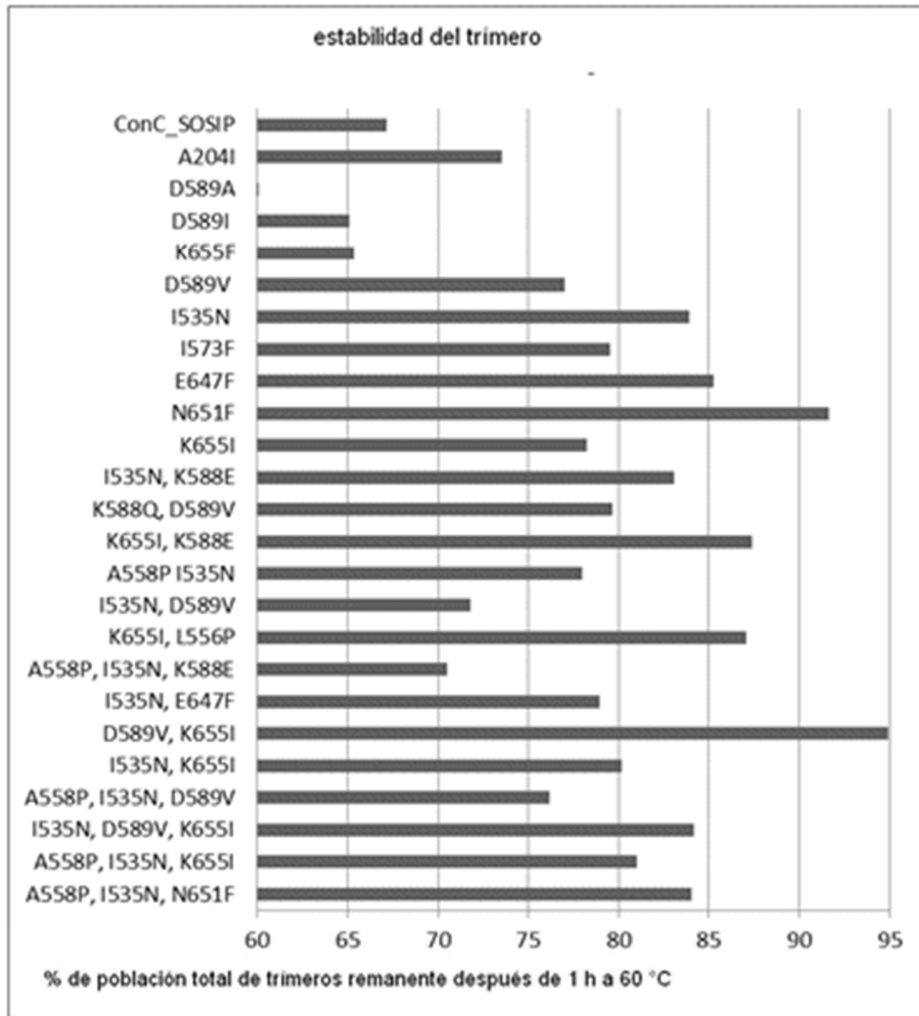


Fig. 4

FIG. 5B

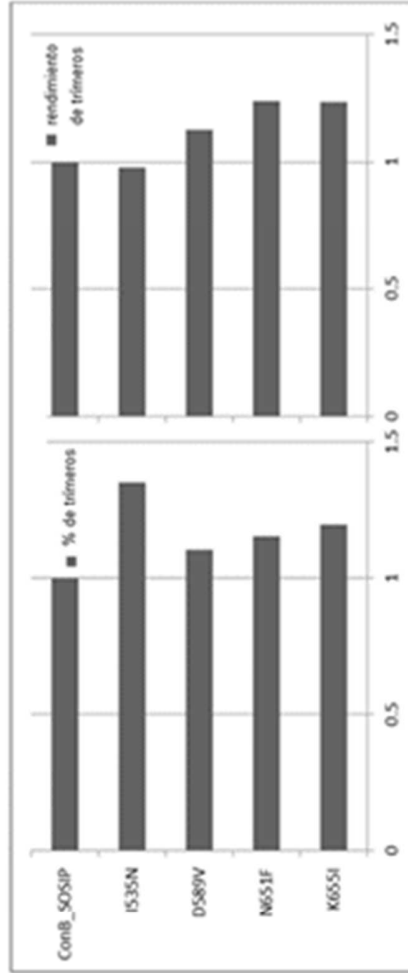


FIG. 5A

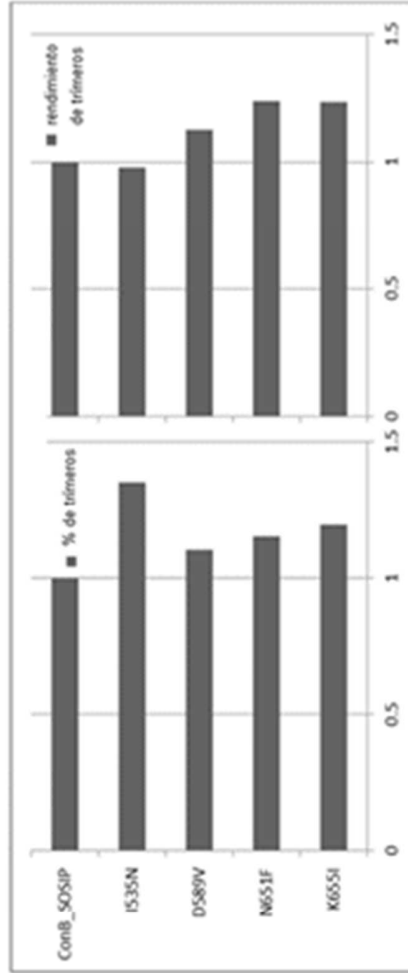


FIG. 6A

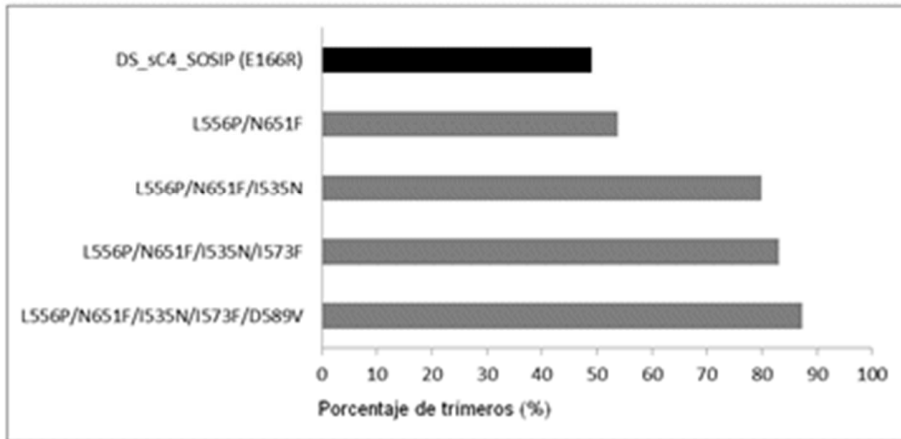
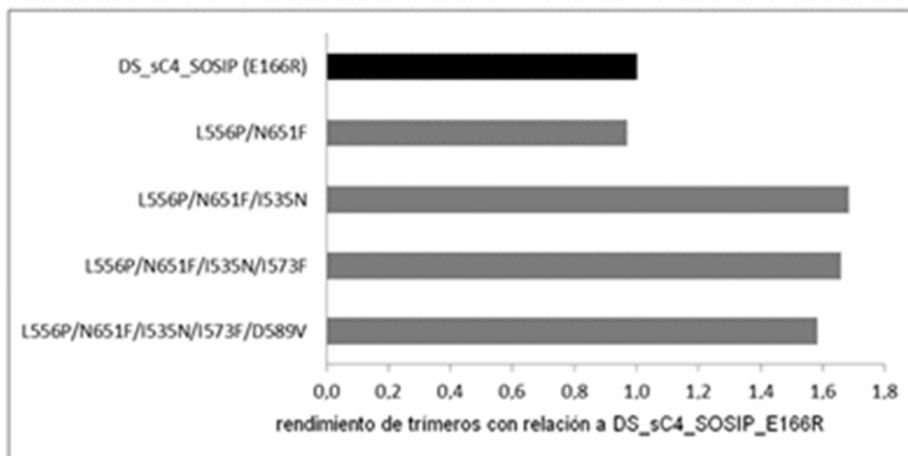


FIG. 6B



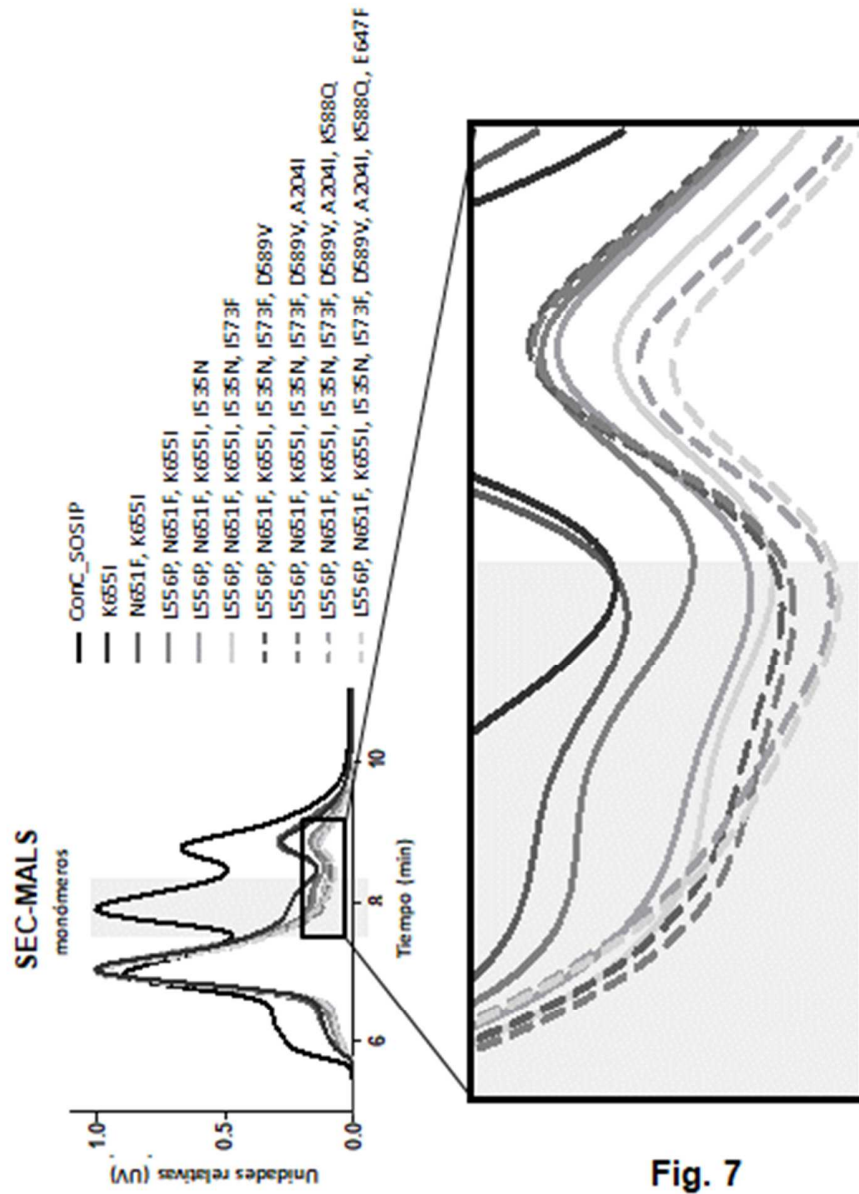


Fig. 7

FIG. 8A

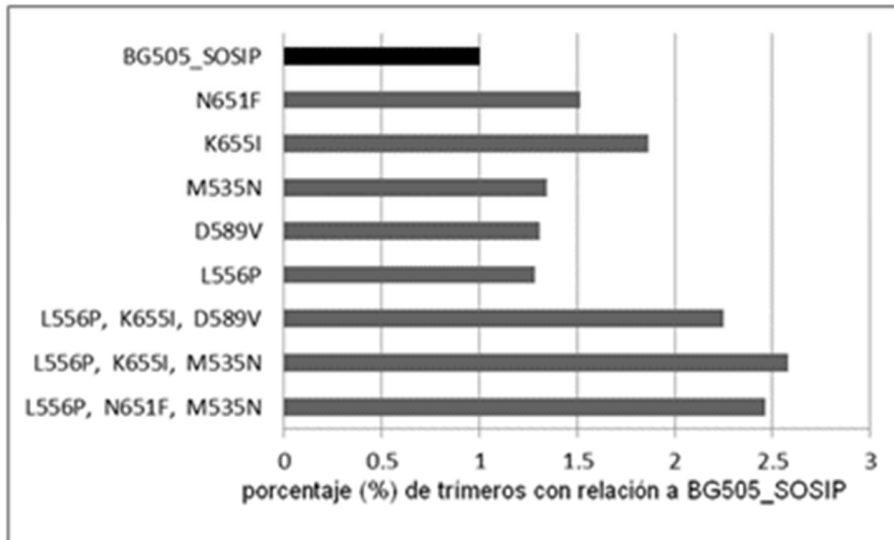


FIG. 8B

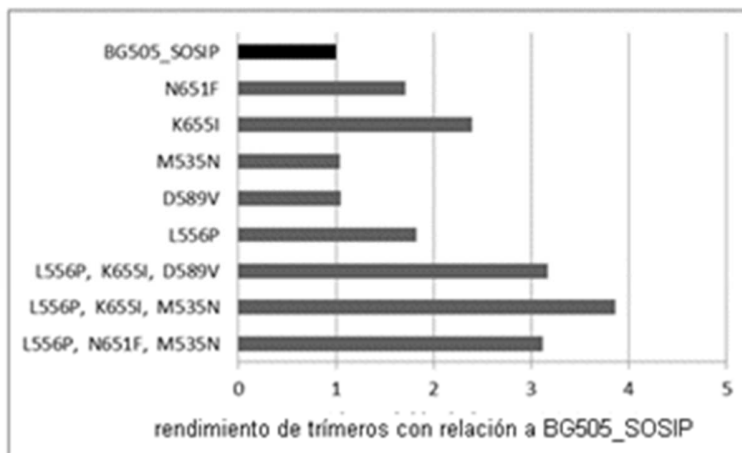


FIG. 9

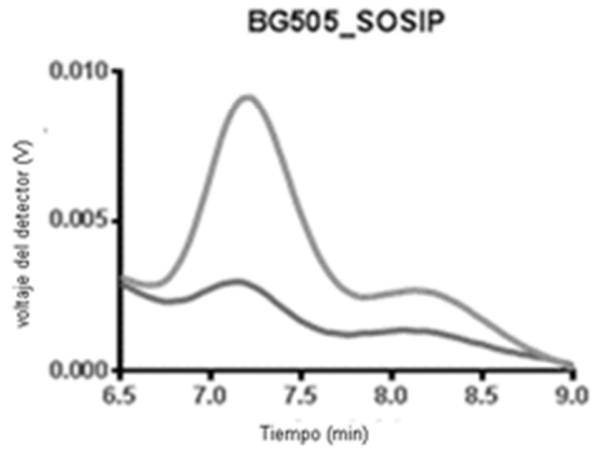


FIG. 10A

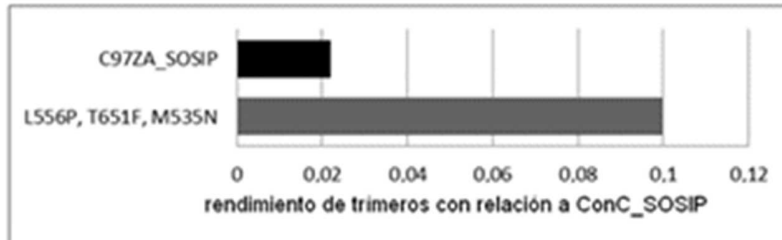


FIG. 10B

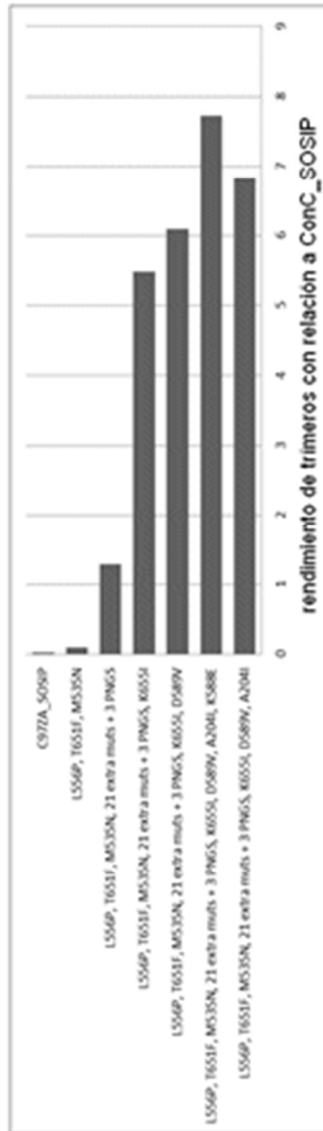


FIG. 11

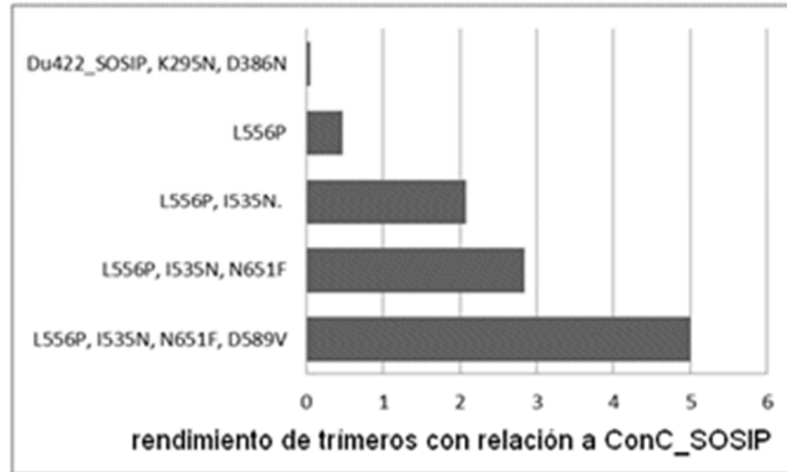


FIG. 12

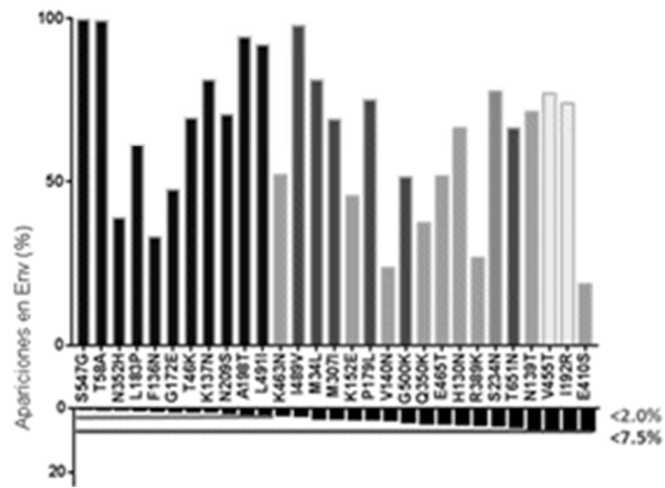


FIG. 13

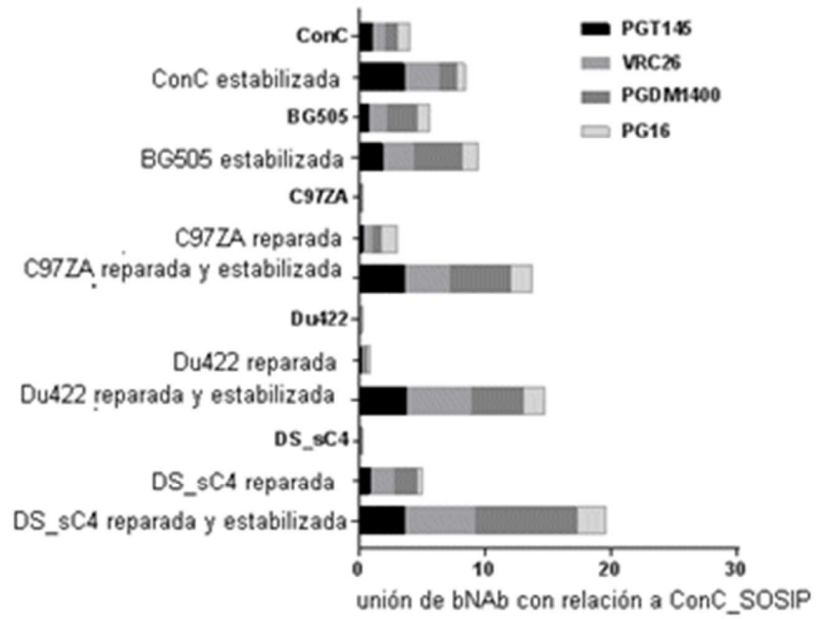


FIG. 14

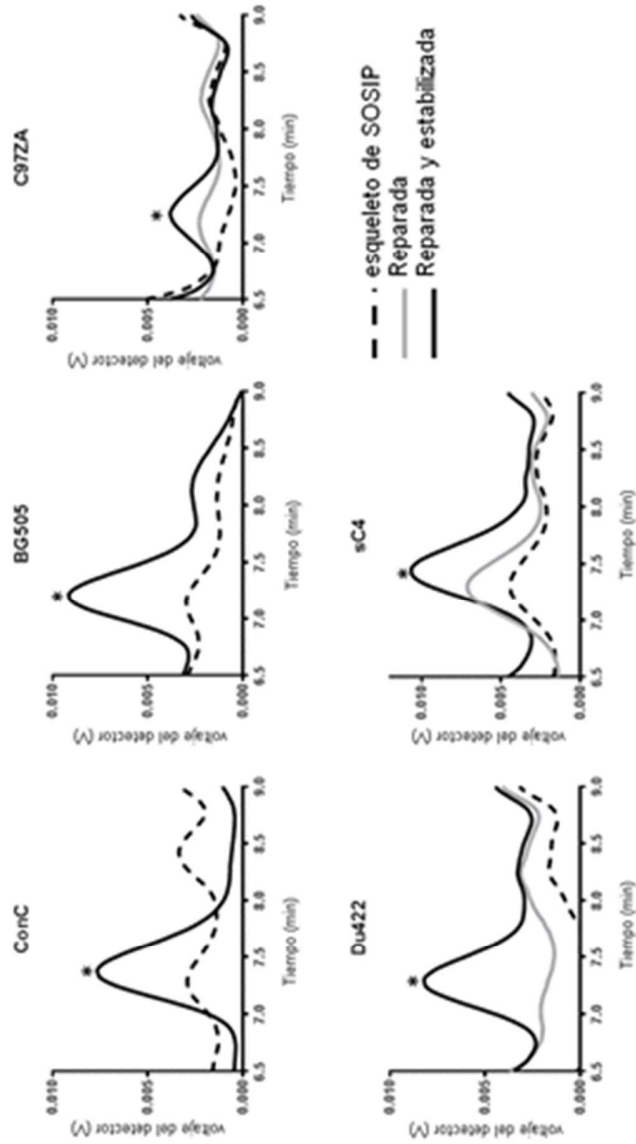


FIG. 15

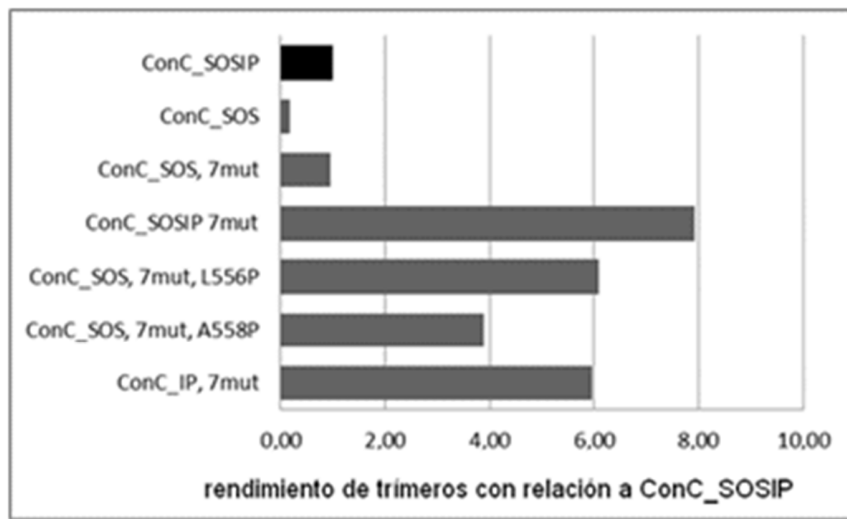


FIG. 16A

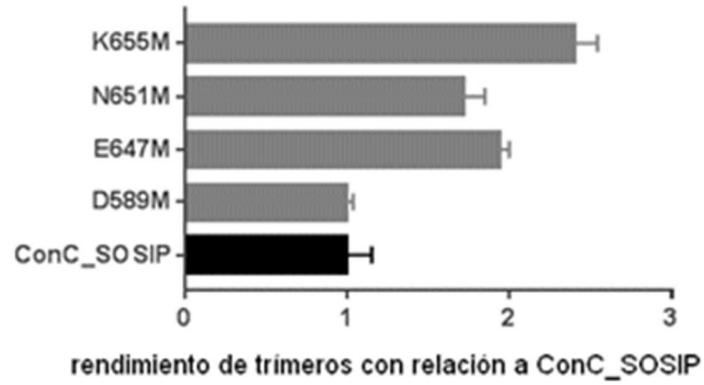


FIG. 16B

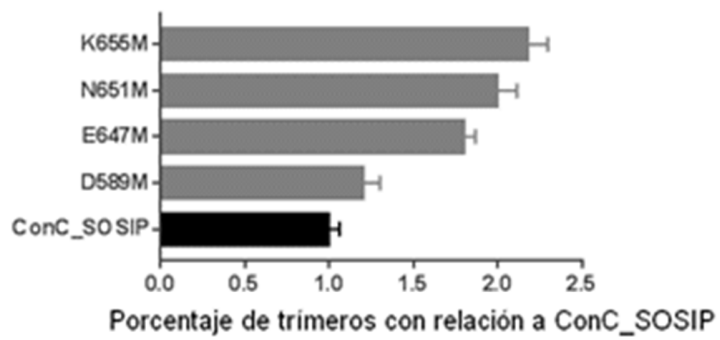


FIG. 17A

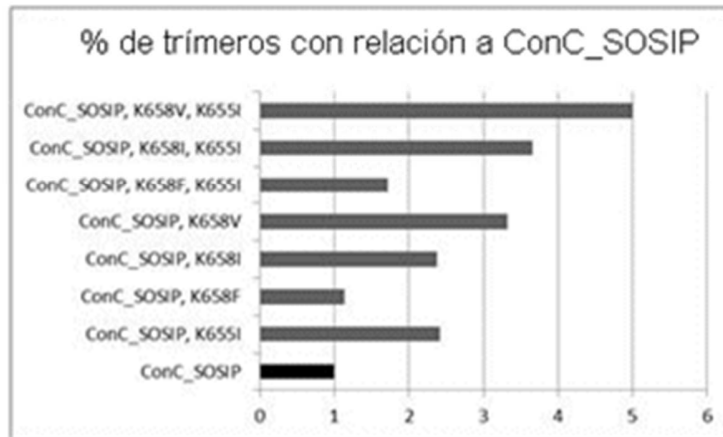


FIG. 17B

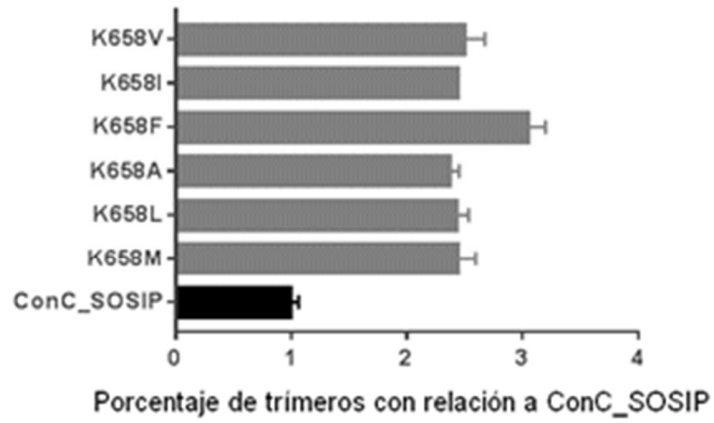


FIG. 17C

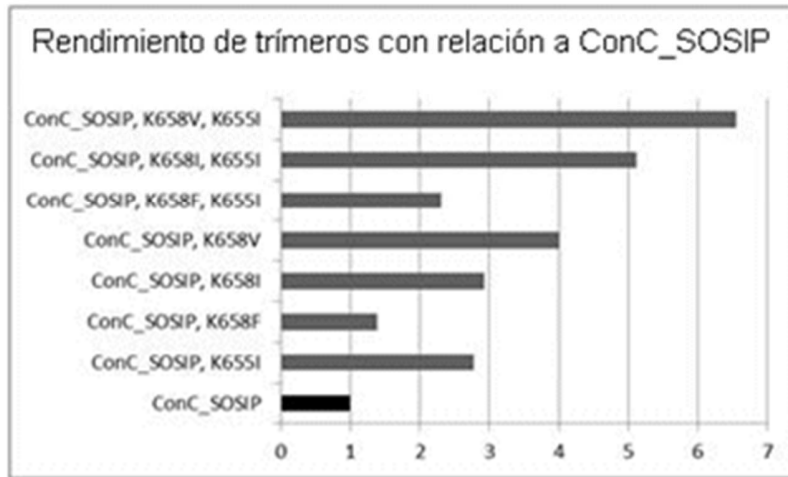
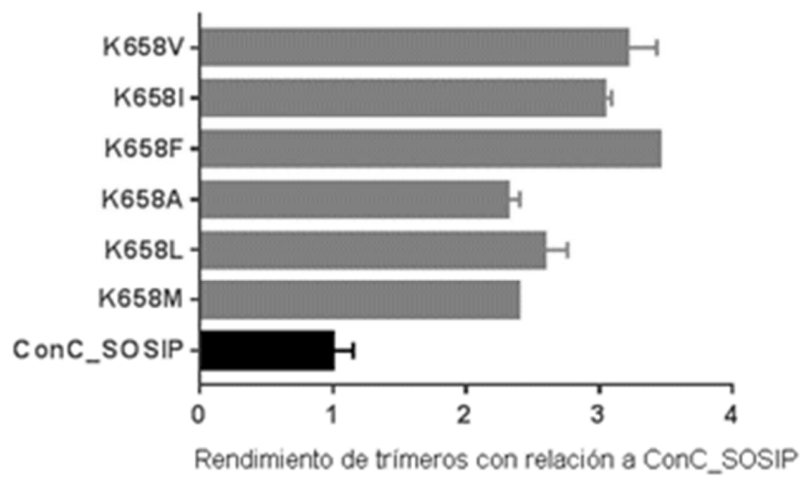


FIG. 17D



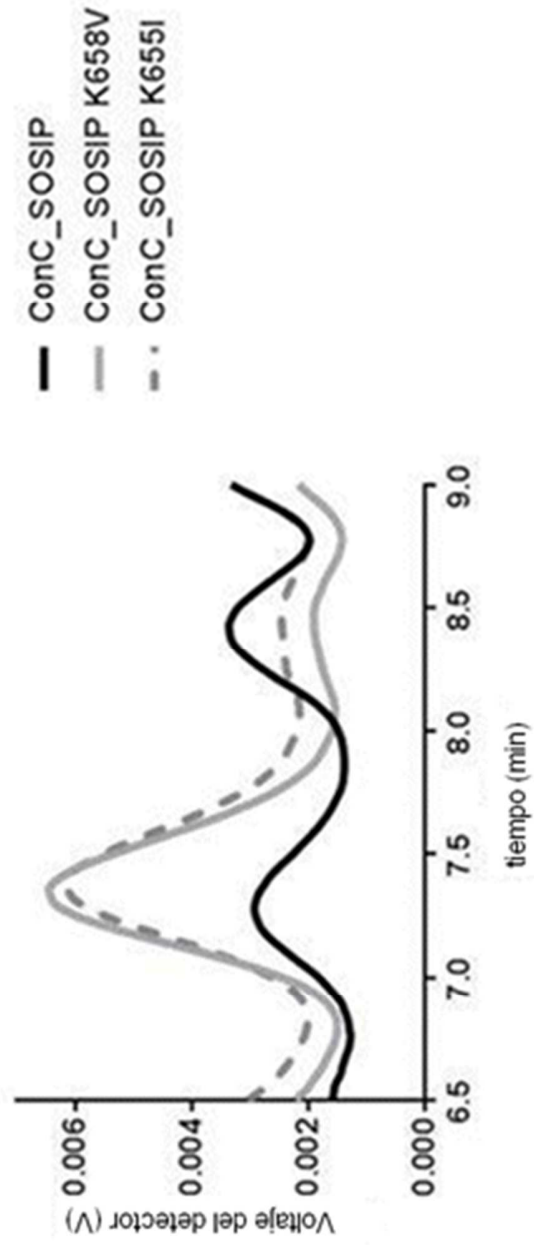


FIG. 18