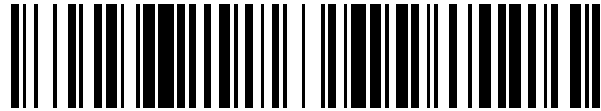


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 471 670**

21 Número de solicitud: 201231819

51 Int. Cl.:

C12N 15/11 (2006.01)

12

SOLICITUD DE PATENTE

A1

22 Fecha de presentación:

23.11.2012

43 Fecha de publicación de la solicitud:

26.06.2014

56 Se remite a la solicitud internacional:

PCT/ES2013/070809

71 Solicitantes:

**CONSEJO SUPERIOR DE INVESTIGACIONES
CIENTÍFICAS (CSIC) (60.0%)**

SERRANO, 117

28006 MADRID ES y

INSTITUTO NACIONAL DE TÉCNICA

AEROSPACIAL "ESTEBAN TERRADAS" (40.0%)

72 Inventor/es:

BERZAL HERRANZ, Alfredo;

BRIONES LLORENTE, Carlos;

SÁNCHEZ-LUQUE, Francisco José;

CUEVAS MANRUBIA, Susanna y

STICH, Michael

74 Agente/Representante:

UNGRÍA LÓPEZ, Javier

54 Título: **MOLÉCULAS INHIBIDORAS DEL VIRUS DE LA INMUNODEFICIENCIA HUMANA TIPO 1 (VIH-1), PROCEDIMIENTO DE OBTENCIÓN Y SUS APLICACIONES.**

57 Resumen:

La presente invención se refiere a un aptámero cuya estructura comprende al menos una secuencia de nucleótidos Guanina-Guanina-Citosina—Adenina-Adenina/Guanina-Guanina-Guanina-Adenina, que es capaz de unirse específicamente a la región 5'UTR del genoma del virus de la inmunodeficiencia humana tipo 1 (VIH-1), proporcionando el procedimiento de obtención de aptámeros con dicha secuencia mediante combinación de técnicas experimentales de selección in vitro de ácidos nucleicos con técnicas computacionales de optimización de secuencia. Asimismo, la invención protege una construcción genética de DNA para sintetizar dichos aptámeros, preferentemente de RNA.

La invención hace también referencia a los diferentes usos del aptámero anterior, incluyendo su uso como molécula biosensora para detectar y/o cuantificar VIH-1, como inhibidor de la producción de partículas virales de VIH-1, así como su aplicación en medicina, protegiendo también un método de tratamiento de una enfermedad causada por VIH-1, y una composición farmacéutica que comprenda dicho aptámero.

ES 2 471 670 A1

**MOLÉCULAS INHIBIDORAS DEL VIRUS DE LA INMUNODEFICIENCIA HUMANA TIPO 1 (VIH-1),
PROCEDIMIENTO DE OBTENCIÓN Y SUS APLICACIONES**

DESCRIPCIÓN

Sector de la técnica

5 La presente invención se refiere al campo de los aptámeros (moléculas de ácidos nucleicos obtenidas por selección o evolución *in vitro*, que poseen la capacidad de unirse específicamente a una diana molecular determinada) y su uso para el tratamiento de las infecciones producidas por el virus de la inmunodeficiencia humana de tipo 1 (VIH-1). Por lo tanto, la presente invención se enmarca en el campo de la biotecnología, y de forma específica en el campo del desarrollo y la elaboración de composiciones farmacéuticas para el tratamiento de enfermedades producidas por el VIH-1. La presente invención podría también derivar en un producto de interés en el campo de la terapia génica, así como en el desarrollo de sistemas de diagnóstico de VIH-1.

Estado de la técnica

15 El VIH es el agente causal del síndrome de inmunodeficiencia humana adquirida (SIDA), una enfermedad que afecta al sistema inmunitario de los seres humanos y provoca una inactivación del mismo. De las dos especies diferentes de VIH, la denominada "virus de la inmunodeficiencia humana de tipo 1 (VIH-1)" es más virulenta que la "tipo 2" (VIH-2), y de hecho es la más extendida en el mundo. La infección por VIH se ha calificado ya como pandemia. Según datos de la Organización Mundial de la Salud, en el año 2010 alrededor de 34 millones de personas en el mundo eran VIH-positivas, se detectaron alrededor de 2.7 millones de nuevas infecciones y hubo aproximadamente 1.8 millones de muertes como consecuencia de la infección por VIH (http://www.who.int/hiv/data/2011_epi_core_en.png). Del total de personas VIH-positivas, sólo 6,6 millones estaban recibiendo tratamiento antirretroviral.

25 El VIH es un retrovirus y se caracteriza, como todos los miembros de la familia *Retroviridae*, por una alternancia entre fases con genoma de RNA y DNA en su ciclo viral. El RNA genómico del VIH-1 presenta tres regiones codificantes principales denominadas *gag*, *pol* y *env* (en orden 5' a 3') flanqueadas por dos regiones no traducibles o UTRs (del inglés *untranslated regions*). El gen *gag* codifica las proteínas de la nucleocápside viral y la matriz (denominadas p17, p24 y p7); *pol* codifica enzimas para el procesamiento del material genético viral y de sus productos génicos (transcriptasa reversa o RT, proteasa o PR, e integrasa o IN); y *env* codifica las glicoproteínas gp120 (superficial) y gp41 (transmembrana), responsables de la infección y tropismo viral. El VIH dispone, además, de una serie de genes accesorios (*tat*, *rev*, *nef*, *vpr*, *vif* y *vpu*) típicos de lentivirus y que ejercen papeles reguladores posibilitando, entre otras cosas, la infección de células no proliferativas. La 5'UTR del RNA viral resulta especialmente relevante por presentar una serie de dominios estructurales altamente conservados entre los diferentes aislados virales. Estos dominios intervienen activamente en procesos clave como la transcripción, la transcripción reversa, el procesamiento de los RNA mensajeros o el empaquetamiento del RNA genómico viral en las partículas infectivas. La 5'UTR puede plegarse en dos conformaciones alternativas, llamadas LDI y BMH (Huthoff and Berkhout, 2001). Un hecho de relevancia para la presente invención es que en función del plegamiento adoptado por la 5'UTR, se exponen o no otros dominios estructurales genómicos que desempeñan funciones esenciales para el virus, como es el caso de los dominios Poli-A o DIS. Esta transición estructural puede desempeñar un papel en la regulación de procesos virales.

40 La terapia antirretroviral de gran actividad (TARGA, en inglés HAART) se ha revelado como el tratamiento más efectivo de la enfermedad: no se consigue su cura pero la reducción de la carga viral es evidente en un alto porcentaje de los pacientes infectados, lo que permite cierta restauración de la función inmunitaria. Además, se ha probado que la TARGA reduce la propagación del virus desde el individuo enfermo tratado y previene el contagio a los individuos sanos tratados en situación de riesgo. Esta terapia consiste en la combinación de dos o más fármacos antirretrovirales, los cuales son compuestos químicos de diferente naturaleza generalmente encaminados a inhibir las enzimas virales (RT, PR o IN), o bien a interferir con la fase de adhesión entre las proteínas de la cubierta del virus a los receptores de su célula diana (principalmente linfocitos T y macrófagos). Anualmente se publican las correspondientes recomendaciones de tratamiento, con la indicación de cuáles de las combinaciones de fármacos son más adecuadas en cada caso (<http://aidsinfo.nih.gov/guidelines>). Compañías farmacéuticas de la envergadura de Roche, Merck o GlaxoSmithKline distribuyen actualmente varios de estos medicamentos.

55 Otro campo de investigación relevante para la presente invención es el de los aptámeros, que son moléculas de ácidos nucleicos obtenidas por selección o evolución *in vitro* y que poseen la capacidad de unirse específicamente a una diana molecular determinada. En 1990 se comunicó en dos artículos independientes la puesta a punto de la tecnología de selección *in vitro* de RNA, en la cual se utilizaba como presión selectiva la capacidad del RNA para unirse específicamente a una molécula diana. El proceso, basado en ciclos de selección

y amplificación, se denominó “Systematic Evolution of Ligands by EXponential Enrichment” o SELEX (Tuerk and Gold, 1990), y se acuñó el término “aptámero” (de la palabra latina *aptus*, que significa encajar) para denominar a las moléculas de RNA generadas por este método (Ellington and Szostak, 1990).

En la actualidad, el punto de partida de un experimento de selección *in vitro* de ácidos nucleicos (RNA o DNA) es una población numerosa de moléculas de DNA (típicamente, de 10^{12} a 10^{16}) con una región central (por lo general, de entre 20 y 100 nucleótidos, nt, de longitud) con secuencia aleatoria o altamente mutagenizada. Dicha región central está flanqueada por dos zonas de unión a cebadores específicos (de entre 12 y 20 nt de longitud cada una de ellas) cuya secuencia es conocida y permitirá las reacciones enzimáticas requeridas en cada ronda: transcripción reversa (RT) que realiza el paso de RNA a DNA, amplificación del DNA por la reacción en cadena de la polimerasa (PCR), y transcripción *in vitro* (IVT) del DNA amplificado a RNA. La selección se introduce en el proceso permitiendo que pasen a la siguiente ronda únicamente las moléculas de ácido nucleico capaces de unirse (aptámeros) a su diana inmovilizada en un soporte (típicamente, columnas de afinidad o nanopartículas de distinto tipo), o bien las moléculas capaces de realizar una determinada función bioquímica. En este último caso, las moléculas con capacidad catalítica obtenidas se denominan “ribozimas” si son de RNA, o bien “desoxirribozimas” o “DNAzimas” si son de DNA. Las moléculas no funcionales se eliminan en cada ronda, mientras que las subpoblaciones funcionales pueden ser analizadas mediante clonaje y secuenciación. Las sucesivas rondas de selección y amplificación (generalmente, entre 6 y 15) aumentan la relación entre los ácidos nucleicos activos e inactivos presentes en la población, hasta que ésta llega a estar enriquecida en moléculas que poseen el fenotipo (capacidad de unión a una diana, o bien actividad bioquímica) buscado. Todo el proceso se desarrolla *in vitro*, y las variables experimentales pueden ser controladas en detalle. Posteriormente se puso a punto el proceso de “evolución *in vitro*”, que se distingue del de “selección *in vitro*” porque en el de evolución se introduce continuamente una fuente de variabilidad genética en la población mediante mutación (generalmente, por PCR mutagénica) y/o recombinación (Joyce, 2004). Con ello, la evolución *in vitro* permite generar nuevas moléculas en cada ronda, mientras que por selección *in vitro* en cada ronda se trabaja con una subpoblación activa a partir de la existente en la población anterior (teniendo como única fuente de variabilidad la tasa de mutación inherente a las enzimas utilizadas).

Empleando estos sistemas se han obtenido numerosos aptámeros de RNA y de DNA, cuya afinidad por su diana en ocasiones supera la de los anticuerpos por sus antígenos. Las dianas reconocidas por los aptámeros cubren un amplio rango de tamaño y complejidad molecular, desde iones hasta células completas, pasando por moléculas pequeñas (como aminoácidos, nucleótidos, antibióticos o metabolitos), péptidos, proteínas, ácidos nucleicos, agregados macromoleculares, virus u orgánulos celulares. Una ventaja adicional de los aptámeros es que su producción se realiza *in vitro* (y por tanto no se requiere el uso de animales de experimentación como sí ocurre con los anticuerpos) y que los ácidos nucleicos resultantes pueden ser modificados para aumentar su estabilidad o su resistencia a degradación por nucleasas. Este aumento en la estabilidad de los aptámeros se puede lograr mediante la adición de determinados grupos químicos (entre ellos flúor, metilo o metoxi) a diferentes posiciones de la base nitrogenada o el azúcar del nucleótido, o bien reemplazando algunos de los nucleótidos del aptámero por otras unidades propias de moléculas análogas de ácidos nucleicos (descritas más adelante). De esta forma, en ninguno de los casos se altera la secuencia nucleotídica del aptámero ni su funcionalidad, pero sí se logra un aumento de la estabilidad del mismo. Además, al tratarse de ácidos nucleicos los aptámeros se pueden amplificar enzimáticamente, y también son susceptibles de modificación para incorporar distintos tipos de marcaje (radiactivo, fluorescente u otros) y diversos grupos funcionales. La estrategia experimental de selección de aptámeros *in vitro* impone un tamaño mínimo de la molécula, que viene determinado por la necesidad de incorporar secuencias fijas flanqueando la región aleatoria central, y que resultan necesarias para la unión de los oligonucleótidos cebadores requeridos durante el proceso de amplificación. Esto ha conllevado tradicionalmente la necesidad de ensayar de forma experimental la reducción del tamaño de la molécula, eliminando partes prescindibles y llegando así a identificar, en el mejor de los casos, el dominio mínimo con capacidad aptamérica. Como se muestra más adelante, en la presente invención aportamos un aspecto novedoso en este sentido, al combinar los procedimientos clásicos de selección de ácidos nucleicos *in vitro* con un análisis bioinformático, el cual permite identificar dominios potencialmente activos sin necesidad de realizar un acortamiento del aptámero experimentalmente.

Por último, los aptámeros se diseñan para unirse específicamente y con alta afinidad a su molécula diana pudiendo, según el caso, inducir cambios conformacionales en la misma como consecuencia de dicha unión. También pueden unirse covalentemente a ribozimas o DNA-zimas para generar “aptámeros catalíticos” o “aptazimas” capaces de realizar determinadas funciones bioquímicas una vez unidos a su diana (por ejemplo, cortar la diana en el caso de que ésta sea otra molécula de RNA). Gracias a todo ello, durante las dos últimas décadas los aptámeros han ido cobrando popularidad en el ámbito de la biotecnología y la biomedicina, y en la actualidad estamos asistiendo a un crecimiento constante en el número de aplicaciones de los aptámeros en diagnóstico (utilizando aptámeros en disolución o bien inmovilizados sobre distintos tipos de biosensores) y en terapia (Klussmann, 2006).

En el campo de la terapia, una de las aplicaciones de los aptámeros ha sido su desarrollo contra dianas de RNA virales. Los RNAs genómicos virales, fundamentalmente los de cadena sencilla (ssRNA), adquieren distintos

grados de complejidad estructural y funcional en sus entornos biológicos. En particular, algunos de sus dominios funcionales más invariables en cuanto a secuencia y estables en cuanto a estructura pueden ser usados como dianas para generar aptámeros contra ellos. Estos aptámeros se han revelado como una potente herramienta antiviral frente a virus de RNA como el de la hepatitis C (Kikuchi *et al.*, 2005; Romero-López *et al.*, 2005) o el propio VIH-1 (Kolb *et al.*, 2006). En el caso del VIH-1, el dominio TAR presente en la 5'UTR (responsable de la activación de la transcripción y altamente conservado) es una diana muy atractiva para la utilización de aptámeros seleccionados contra ella (Duconge and Toulme, 1999; Watrin *et al.*, 2009). El aptámero de RNA "R06", de 24 nucleótidos, seleccionado contra el dominio TAR aislado, ha sido modificado posteriormente para su localización nucleolar y ha mostrado en cultivos celulares una inhibición de la replicación del VIH-1 superior al 90% cuando es expresado intracelularmente (Kolb *et al.*, 2006). Sin embargo, aunque se han desarrollado aptámeros contra dominios o subdominios concretos de la 5'UTR del VIH-1, el proceso de obtención no ha contemplado el hecho de que el correcto ensamblaje y exposición de los mismos en el RNA de la 5'UTR está sujeto al equilibrio existente entre los dos conformeros descritos para dicha región (BMH y LDI; Huthoff and Berkhout, 2001) y a las posibles interacciones con los subdominios vecinos de la 5'UTR. Estos fenómenos podrían estar afectando a la capacidad de unión de cualquier aptámero, incluso los que se comporten de forma óptima contra un subdominio aislado de la 5'UTR.

Por tanto se pone en evidencia la necesidad de diseñar aptámeros contra dominios funcionales del genoma del VIH-1 en un entorno más aproximado al existente en la molécula del RNA genómico. En la presente invención se han obtenido aptámeros contra uno o varios subdominios concretos de la 5'UTR del VIH-1 mediante un proceso de selección *in vitro* en el que la diana utilizada ha sido la 5'UTR completa. Esta aproximación supone una ventaja importante frente a las publicaciones y patentes previas, ya que los aptámeros obtenidos en la presente invención reconocen zonas de unión accesibles en el plegamiento completo de la 5'UTR en uno de dos conformeros posibles, en ambos, o en la interfaz cinética del equilibrio entre ellos. Así los aptámeros que se han seleccionado son capaces de unirse sin verse afectados por los impedimentos estéricos ocasionados por las regiones adyacentes y/o los demás subdominios de dicha 5'UTR y en las condiciones cinéticas necesarias para ello. Otra ventaja de estos aptámeros frente a los descritos por otros autores radica en que, tal como se ha indicado previamente, la aproximación combinada entre experimentación y simulación computacional (abordaje *in vitro / in silico*) ha permitido definir el aptámero mínimo activo sin necesidad de multiplicar los ensayos experimentales.

Breve Descripción de la Invención

La presente invención se refiere al campo de los aptámeros, moléculas de ácidos nucleicos obtenidas por selección o evolución *in vitro* que poseen la capacidad de unirse específicamente a una diana molecular determinada. En concreto, los aptámeros a los que se refiere esta invención son moléculas de RNA, de DNA o de análogos de ácidos nucleicos, que se unen a determinadas regiones genómicas del virus de la inmunodeficiencia humana de tipo 1 (VIH-1), más en concreto a dominios de RNA presentes en la región 5' no traducible (5'UTR) del genoma viral, y que previamente no se habían descrito como dianas de unión de aptámeros. De hecho, en la presente invención se ha utilizado como diana una molécula de RNA que contiene los primeros 308 nt del genoma de VIH-1, un fragmento notablemente más largo que los utilizados hasta ahora en ensayos de selección de aptámeros frente al VIH-1.

Dicha diana contiene los primeros 308 nt (SEQ ID No: 1) del total de 336 nt (SEQ ID No: 2) que constituyen la 5'UTR del RNA genómico del VIH-1, siendo los primeros 290 nt comunes (SEQ ID No: 3) al RNA genómico y la mayoría de los RNAs subgenómicos del virus (todos estos datos numéricos se aplican en el caso de la cepa NL4-3, con número de acceso en la base de datos GenBank: AF324493). Aunque en la 5'UTR se han caracterizado elementos estructurales aislados (como el TAR, frente al cual se han aislado los aptámeros descritos en la sección sobre el Estado de la técnica), tal como se ha indicado la estructura del propio TAR y de los demás elementos está influenciada por la presencia del resto de la 5'UTR. En función del plegamiento adoptado por la 5'UTR (completa, y también en el caso de utilizar sus primeros 308 nt) se exponen o no distintos dominios estructurales que desempeñan funciones esenciales para el virus, como son los dominios Poli-A o DIS. En función del tamaño del fragmento de la 5'UTR utilizado se puede favorecer o no la existencia de uno de los conformeros, y en todo caso fomentar la estabilidad de uno respecto del otro desplazando el equilibrio entre ambos hacia el de estructura más estable. En este caso, el uso de un fragmento de 308 nt de la 5'UTR del VIH-1 como diana para la selección de aptámeros es una situación más cercana a la existente en el genoma viral completo, por lo que el sistema seguido en la presente invención, a diferencia del utilizado en los casos previamente publicados, permite identificar elementos estructurales en la molécula diana que son inexistentes cuando se utilizan moléculas pequeñas derivadas de la 5'UTR. Así, los aptámeros obtenidos experimentalmente en esta invención resultaron estar específicamente dirigidos al dominio Poli-A, que se mostró como la diana más eficiente dentro de la 5'UTR del VIH-1. La frecuencia con la que aparecen aptámeros que contienen una secuencia de unión complementaria a esta diana Poli-A es mucho más alta que la de los aptámeros con secuencia complementaria a motivos en TAR (cuya secuencia y estructura se habían descrito como eficientes para la unión de ácidos nucleicos tanto aptámeros como RNAs *antisense*, Kolb *et al.*, 2006) o en DIS (dominio seleccionado por la naturaleza para desempeñar una función que implica la unión a otra molécula de RNA

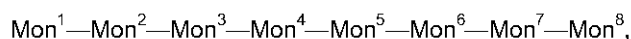
genómico viral). Así, la presente invención se basa en la demostración experimental de que los aptámeros pueden dirigirse a una nueva e imprevista diana dentro de la 5'UTR cuando su longitud es suficientemente extensa: el sitio Poli-A.

El método seguido para obtener estos aptámeros conlleva el uso de las estrategias clásicas de selección de ácidos nucleicos *in vitro*, en combinación con un novedoso análisis bioinformático o *in silico* que permite la identificación de nuevas moléculas con capacidad de unión a la diana. Los aptámeros desarrollados poseen como secuencia común a todos ellos el octanucleótido 5'-GGCAAGGA-3' o bien 5'-GGCAGGGA-3', en ambos casos formando un lazo terminal monocatenario que cierra un tallo de cadena doble de al menos tres pares de nucleótidos. Los aptámeros que contienen estas secuencias pueden poseer regiones flanqueantes tanto en su extremo 5' como 3', preferentemente con longitudes comprendidas entre 1 y 75 nucleótidos, cualquiera que sea la estructura que forman dichas regiones flanqueantes y la funcionalidad adicional que aporten las mismas. En particular, en la presente invención se han desarrollado aptámeros de una longitud total de 64 nucleótidos, todos los cuales contienen los octanucleótidos indicados en diferentes posiciones dentro de su secuencia. Además, gracias a la citada aproximación combinada *in vitro* / *in silico*, se ha logrado definir una molécula concreta de RNA de tan sólo 16 nucleótidos, con secuencia 5'-CCCCGGCAAGGAGGGG-3', que mantiene la estructura en tallo-lazo indicada y que ha demostrado poseer una funcionalidad similar a la de los aptámeros de 64 nucleótidos de longitud. Se aprovecha la capacidad de las moléculas desarrolladas para unirse a la 5'UTR del VIH-1, y preferentemente al dominio Poli-A a través del octanucleótido común a todas ellas, tanto en ensayos *in vitro* como en cultivo celular. La unión de estos aptámeros al genoma del VIH-1 produce la inhibición de la replicación del virus en cultivo celular, lo que se determina mediante ensayos específicos. Es interesante destacar el hecho de que la diana de unión del octanucleótido presente en los aptámeros al dominio Poli-A se encuentra conservada en todo el espectro de secuencias de aislados virales analizados (Fig. 1). Por tanto, es esperable que los aptámeros de la presente invención, que se unen específicamente al dominio Poli-A, serán activos frente a aislados del VIH-1 correspondientes a los grupos M (y dentro de él sus distintos subtipos y formas recombinantes), N, O y P.

Descripción de la invención

La presente invención se basa en que los inventores han identificado, gracias a la combinación de sistemas de selección de ácidos nucleicos *in vitro* con un novedoso análisis bioinformático o *in silico*, una serie de nuevos aptámeros, que contienen un motivo común mínimo y se unen específicamente a regiones genómicas del VIH-1, más en concreto a dominios de RNA presentes en la 5'UTR del genoma viral. Dicha unión entre el aptámero y el genoma de VIH-1 inhibe la replicación viral. Por tanto, estas moléculas son útiles para la elaboración de nuevas composiciones farmacéuticas para el tratamiento de infecciones provocadas por este virus.

Un primer aspecto de la presente invención se refiere a un aptámero, o cualquiera de sus derivados y enantiómeros, capaz de unirse específicamente a zonas accesibles en el plegamiento completo de la región 5'-UTR del genoma del virus de la inmunodeficiencia humana VIH tipo 1 (VIH-1), resultando útil por ejemplo para la inhibición de la replicación del VIH-1, que comprende una estructura de secuencia:



donde Mon¹ a Mon⁸ son monómeros de un ácido nucleico o de un análogo químico de ácido nucleico.

Según la presente invención, el término "monómero de ácido nucleico" se refiere a un nucleótido en el que su base nucleotídica es adenina (A), guanina (G), citosina (C), timina (T) o uracilo (U). Según la presente invención, el término "monómero de un análogo químico de ácido nucleico" se refiere a la unidad mínima que al polimerizar compone un análogo químico de ácido nucleico, conteniendo cada una de estas unidades mínimas una base nitrogenada (A, G, C, T o U) covalentemente unida al componente del esqueleto análogo de la ribosa-fosfato o la deoxirribosa-fosfato propias de los ácidos nucleicos. En la presente descripción, el término "análogo químico de ácido nucleico" puede encontrarse igualmente referido por términos sinónimos como son "molécula análoga de ácido nucleico", "análogo de ácido nucleico" o "análogo sintético de ácido nucleico".

Así, en los aptámeros de la presente invención la base nitrogenada de Mon¹ es una G, la base nitrogenada de Mon² es una G, la base nitrogenada de Mon³ es una C, la base nitrogenada de Mon⁴ es una A, la base nitrogenada de Mon⁵ es una A o una G, la base nitrogenada de Mon⁶ es una G, la base nitrogenada de Mon⁷ es una G y la base nitrogenada de Mon⁸ es una A. Con ello, la secuencia de los aptámeros de la presente invención puede también representarse como GGCAAGGA o bien GGCAGGGA, definida en base a la presencia de bases nitrogenadas capaz de producir la unión específica a la región 5'UTR del genoma de VIH-1.

En el ámbito de la presente invención, VIH-1 se refiere a un virus infeccioso humano con tropismo preferente por las células del sistema inmunitario (preferentemente linfocitos T CD4⁺, macrófagos y microglía) y con potencial para producir una inmunodeficiencia en los individuos infectados por él. Sus partículas virales productivas han de

contener dos moléculas de RNA de polaridad positiva de longitud genómica (aprox. 9.7 kb) con una región codificante formada por nueve genes, tres de ellos estructurales (*gag*, *pol* y *env*) y los otros seis reguladores (*tat*, *rev*, *nef*, *vpu*, *vif*, *vpr* y *vpx*), flanqueada por regiones no traducibles denominadas 5'UTR y 3'UTR según su posición. Los genes estructurales deben codificar las proteínas estructurales del virión (cápside y proteínas estructurales presentes en la envuelta lipídica) y las enzimas para la transcripción reversa del material genético (transcriptasa reversa), la integración del DNA copia en el genoma celular (integrasa) y la maduración proteolítica de los productos génicos de la expresión del virus (proteasa). El ciclo de vida del virus debe alternar entre una fase de virión con material genético de RNA y una fase intracelular con material genético de DNA.

En la taxonomía oficial de la especie VIH-1 se han definido hasta el momento cuatro grupos: M (que a su vez contiene 9 subtipos: A –con subsubtipos A1 y A2–, B, C, D, F –subsubtipos F1 y F2–, G, H, J y K; y al menos 49 formas recombinantes entre distintos subtipos), N (con 5 subtipos: I, II, III, IV y V), O y P. Dentro de cada subtipo se distinguen cepas, variantes, y aislados. La elevada tasa de mutación del VIH-1 hace que exista una notable variabilidad entre las secuencias de los aislados de los pacientes, incluso dentro de un mismo individuo infectado. En cualquier caso, dicha variabilidad inter- e intra-aislado queda agrupada dentro de la especie VIH-1, puesto que los virus que forman parte de ella poseen un linaje replicativo propio, ocupan un nicho ecológico particular, y se encuentran a menor distancia genética (es decir, poseen mayor similitud de secuencia) entre sí que con los aislados pertenecientes a la siguiente especie más cercana (en este caso VIH-2, la segunda especie que infecta a humanos dentro del género *Lentivirus*). Por ello, cuando se utiliza el término VIH-1 en este documento, se hace referencia a cualquier virus o población de virus cuya secuencia o secuencias poseen mayor similitud de secuencia con el total de los aislados secuenciados hasta el momento para VIH-1 que con los aislados correspondientes a cualquier otra especie de virus. En el ámbito de la presente invención, y sin que sirva de limitación, ejemplos de VIH-1 son los que se indican en la Figura 1 u otros conocidos por el experto medio en la materia como los que se describen en el documento "Kuiken C., *et al.* 2011" incluido aquí como referencia.

Generalmente, el VIH-1 posee una elevada conservación de secuencia en la 5'UTR de su genoma (Kuiken C., *et al.* 2011). En el ámbito de la presente invención, la 5'UTR se refiere a la región 5' no traducible de los diferentes RNAs virales, y que generalmente presenta un elevado grado de estructuración. La región 5'UTR también es conocida en el estado de la técnica por región "5'-*leader*" o "5'-líder" del RNA mensajero de VIH-1. Igualmente, la 5'UTR contiene una parte de la secuencia repetida terminal larga o LTR (del Inglés *Long Terminal Repeat*), que está presente en ambos extremos del DNA genómico viral, concretamente las regiones R (repetida, véase región entre posiciones 1 a 98, ambas inclusive, en SEQ ID No: 1, SEQ ID No: 2 y SEQ ID No: 3) y U5 (única en 5', véase región entre posiciones 99 a 180, ambas inclusive, en SEQ ID No: 1, SEQ ID No: 2 y SEQ ID No: 3). En este sentido, es destacable que la región R se encuentra también en el extremo 3' de todos los RNAs virales, y contiene en su secuencia la zona de unión prevista para los aptámeros protegidos en la presente invención (véase región entre posiciones 81 y 86, ambas inclusive, en SEQ ID No: 1, SEQ ID No: 2 y SEQ ID No: 3).

La 5'UTR es diferente en las diferentes familias de RNAs virales subgenómicos y en el RNA genómico no procesado, siendo la región comprendida por los primeros 290 nt común a todos ellos. Preferentemente, en la presente invención la 5'UTR hace referencia a los primeros 308 nt del RNA genómico viral no procesado, que contiene dicha región común y 18 nucleótidos más, necesarios para completar el dominio SL (interrumpido por el *splicing* en los RNAs subgenómicos). La 5'UTR de 308 nucleótidos, denominada aquí como UTR₃₀₈, contiene los dominios estructurales previamente descritos: TAR, Poli-A, PBS, DIS (véase, respectivamente, regiones entre posiciones 1 a 58, 59 a 105, 117 a 237 y 244 a 278, posiciones inicial y final inclusive, en SEQ ID No: 1, SEQ ID No: 2 y SEQ ID No: 3) y SL (véase región entre posiciones 283 a 301, ambas inclusive, en SEQ ID No: 1 y SEQ ID No: 2). No obstante, es necesario precisar que, tal como se ha mencionado anteriormente, la variabilidad entre los aislados virales hace que estos datos numéricos deban tomarse como orientativos, pues han sido calculados usando la secuencia de la cepa de laboratorio NL4-3. Sin que sirva de limitación, ejemplos de regiones 5'UTR de VIH-1 según la invención son las secuencias de nucleótidos de las regiones 5'UTR conocidas por el experto medio en el campo como las que se pueden encontrar en el documento "Kuiken C., *et al.* 2011" incluido aquí como referencia.

En el ámbito de la presente invención, el término "aptámero" se refiere a una molécula de RNA, DNA o cualquier análogo químico de ácido nucleico que comprende una secuencia de una sucesión de bases nitrogenadas, donde dicha sucesión de bases nitrogenadas es responsable de su actividad independientemente del esqueleto polimérico al que se encuentran ancladas (que determinará si es un ácido nucleico natural como el DNA o el RNA, o un análogo químico de los mismos). Como se ha indicado anteriormente, el principio de acción de cualquier análogo de ácido nucleico es, en el contexto de esta invención, exactamente el mismo que el del DNA o RNA, y está definido por la secuencia de las bases nitrogenadas que contiene. No obstante, estos análogos generalmente poseen mayor estabilidad que los ácidos nucleicos naturales. De hecho, en distintas áreas de la biotecnología ha resultado muy útil el uso de moléculas análogas de los ácidos nucleicos naturales para diversas aplicaciones terapéuticas y de diagnóstico (revisiones en: Geyer *et al.*, 2003; Briones y Moreno, 2012). En el estado de la técnica actual se conocen diversos tipos de análogos químicos de ácidos nucleicos, como por ejemplo son, y sin que sirva de limitación, los definidos en la presente descripción.

Dentro de dichos análogos de ácidos nucleicos, es particularmente interesante el caso de los “ácidos nucleicos cerrados” o *locked nucleic acids* (abreviados como LNA), caracterizados por ser polímeros de nucleótidos de β-D-ribofuranosa modificada con un conector metileno 2'-O, 4'-C. Ese tipo de enlace restringe o “cierra” la ribofuranosa en su conformación 3'-endo. El LNA forma híbridos con el DNA y con el RNA mucho más estables que los propios dímeros DNA-DNA, DNA-RNA o RNA-RNA. Otra característica interesante de los LNAs es que nucleótidos con esta conformación pueden unirse a nucleótidos convencionales de DNA o RNA, formando moléculas híbridas en las que se promueve la interacción con otro ácido nucleico a través de los lugares ocupados por los monómeros de LNA.

Son también muy útiles los análogos de ácidos nucleicos denominados “ácidos nucleicos peptídicos” o *peptide nucleic acids* (PNA). El PNA es un polímero de naturaleza peptídica, a diferencia del esqueleto de azúcares y fosfatos característico de los ácidos nucleicos naturales. El esqueleto del PNA está formado por unidades de N-(2-aminoetil)-glicina unidas por enlaces peptídicos, es aquiral, eléctricamente neutro y carece de átomos de fósforo. Las bases nucleotídicas púricas (A y G) y pirimidínicas (C y T) están unidas al esqueleto del PNA mediante conectores metilencarbonilo en una conformación tal que permiten la interacción mediante puentes de hidrógeno con las bases nucleotídicas de los ácidos nucleicos naturales, del mismo modo a como ocurre entre éstas de forma natural en el DNA y el RNA. El PNA forma dímeros extraordinariamente estables tanto con DNA como con RNA. Relacionado con el PNA, recientemente se ha sintetizado un nuevo análogo de ácidos nucleicos denominado “ácido nucleico peptídico tioéster” o *thioester PNA* (tPNA).

Otros análogos de ácidos nucleicos son los siguientes: i) “ácidos nucleicos derivados de glicerol” *glycerol-derived nucleic acids* (GNA), cuyo esqueleto está compuesto por unidades acíclicas de glicerol unidas por enlaces fosfodiéster; ii) *pyranosyl-RNA* (p-RNA), que contiene anillos de seis átomos de tipo β-D-ribopiranosido en lugar de ribofuranosilo como el RNA, unidas por enlaces fosfodiéster 4'→2'; iii) “ácidos nucleicos treósicos” o *threose nucleic acid* (TNA), basado en unidades de α-L-treofuranosilo unidas por enlaces fosfodiéster 3'→2'; iv) “oligo-morfolinofosforoamidatos” o *phosphoroamidate morpholino oligomers* (PMO), cuyos monómeros están formados por anillos de morfolino unidos entre sí por enlaces fosfordiamidato.

Por tanto, en los análogos químicos de ácidos nucleicos, cada una de las unidades monoméricas que los componen es una seleccionada del siguiente grupo:

- β-D-ribofuranosido modificado con un conector metileno 2'-O, 4'-C, con una base nitrogenada unida mediante enlace N-glucosídico entre el nitrógeno 9 de las purinas o el 1 de las pirimidinas y el carbono 1', en el caso de análogos LNA; dicho monómero es susceptible de ser polimerizado mediante enlace fosfodiéster 5'→3';
- N-(2-aminoetil)-glicina, con una base nitrogenada unida mediante un conector metilencarbonilo entre el nitrógeno 9 de las purinas o el 1 de las pirimidinas y el grupo amino del esqueleto de la glicina, en caso de análogos PNA; dicho monómero es susceptible de ser polimerizado mediante enlace peptídico entre el carbono del grupo carboxilo del esqueleto de la glicina y el nitrógeno terminal del grupo aminoetileno;
- Aminoácido cisteína, con una base nitrogenada unida mediante un conector metilencarbonilo entre el nitrógeno 9 de las purinas o el 1 de las pirimidinas y el azufre del esqueleto de cisteína, en caso de análogos tPNA; dicho monómero es susceptible de ser polimerizado mediante enlace peptídico convencional;
- 1,2,3-propanotriol (también llamado glicerol), con una base nitrogenada unida mediante enlace N-glucosídico entre el nitrógeno 9 de las purinas o el 1 de las pirimidinas y el carbono 1', en caso de análogos GNA; dicho monómero es susceptible de ser polimerizado mediante enlace fosfodiéster 3'→2';
- β-D-ribopiranosido, con una base nitrogenada unida mediante enlace N-glucosídico entre el nitrógeno 9 de las purinas o el 1 de las pirimidinas y el carbono 1', en caso de análogos p-RNA; dicho monómero es susceptible de ser polimerizada mediante enlace fosfodiéster 4'→2';
- α-L-treofuranosido, con una base nitrogenada unida mediante enlace N-glucosídico entre el nitrógeno 9 de las purinas o el 1 de las pirimidinas y el carbono 1', en caso de análogos TNA; dicho monómero es susceptible de ser polimerizado mediante enlace fosfodiéster 3'→2';
- Tetrahidro-1,4-oxazina (también llamado morfolino), con la base nitrogenada unida mediante un enlace N-glucosídico entre el nitrógeno 9 de las purinas o el 1 de las pirimidinas y el carbono 1' del anillo de morfolino, en caso de análogos PMO; dicho monómero es susceptible de ser polimerizado mediante conectores fosforoamidato unidos por enlaces amida con el nitrógeno 4' y éster con el oxígeno 1'.

Hasta el momento, la síntesis de gran parte de los análogos químicos de ácidos nucleicos se lleva a cabo por procedimientos de síntesis química similares al utilizado para la síntesis química de ácidos nucleicos convencionales como, a modo de ejemplo y sin que sirva de limitación, la síntesis en fase sólida utilizando como

5 Adicionalmente, se ha descrito la posibilidad de polimerizar ciertos análogos de ácidos nucleicos mediante sistemas enzimáticos que utilizan variantes modificadas genéticamente de las polimerasas naturales (Pinheiro *et al.*, 2012).

En el ámbito de la presente invención, los derivados del aptámero se refieren a aptámeros de RNA o DNA, donde al menos uno de los monómeros de ácido nucleico de dicho aptámero presenta una modificación química.

10 El experto medio en el campo puede encontrar que la introducción de modificaciones químicas que no alteren la secuencia nucleotídica ni la disposición espacial de las bases nitrogenadas no modifica la funcionalidad de la molécula de los aptámeros descritos en la presente invención, que depende fundamentalmente del principio de complementariedad de secuencia entre el aptámero y su molécula diana. Sin embargo, tales modificaciones sí tienen un efecto positivo ya que aumentan la estabilidad y la vida media del aptámero. El aumento en la

15 estabilidad redundante en un incremento en la eficiencia global del aptámero, ya que su mayor resistencia a la degradación conlleva un aumento en la disponibilidad intracelular del inhibidor. A modo de ejemplo y sin que sirva de limitación, esta modificación puede ser útil para aumentar la resistencia a la degradación por enzimas nucleasas (RNAsas o DNAsas) presentes en los medios biológicos. Las modificaciones en los nucleótidos que producen tal aumento de resistencia a la degradación enzimática comprenden un sustituyente seleccionado, sin

20 que sirva de limitación, de la siguiente relación: un átomo de flúor, un grupo amino, un grupo metilo, un grupo metoxi y cualquier combinación de los mismos. También se logra un aumento en la resistencia de los aptámeros a la degradación por exonucleasas por medio del bloqueo de sus extremos 5' o 3' empleando diversos tipos de moléculas o sustituyentes, como por ejemplo introduciendo como sustituyente una molécula que bloquee el grupo fosfato del extremo 5' del ácido nucleico, o bien el grupo hidroxilo de su extremo 3'.

25 Según se entiende la invención, las modificaciones en los nucleótidos para producir resistencia a la degradación enzimática del aptámero que comprenden átomos de flúor se refieren a la sustitución de al menos uno de los grupos químicos de la molécula del nucleótido por un grupo químico que contenga al menos un átomo de flúor, como ocurre y sin que sirva de limitación en los (2'-deoxi-2'-fluoro)-nucleótidos.

30 Según se entiende la invención, las modificaciones en los nucleótidos para producir resistencia a la degradación enzimática del aptámero que comprenden un grupo amino se refieren a la sustitución de al menos uno de los grupos químicos de la molécula del nucleótido por un grupo químico que contenga al menos un grupo amino, como ocurre y sin que sirva de limitación en los 2'-O-(3-aminopropil)-nucleótidos.

35 Según se entiende la invención, las modificaciones en los nucleótidos para producir resistencia a la degradación enzimática del aptámero que comprenden un grupo metilo se refieren a la sustitución de al menos uno de los grupos químicos de la molécula del nucleótido por un grupo químico que contenga al menos un grupo metilo, como ocurre y sin que sirva de limitación en los 2'-O-metil-nucleótidos.

40 Según se entiende la invención, las modificaciones en los nucleótidos para producir resistencia a la degradación enzimática del aptámero que comprenden un grupo metoxi se refieren a la sustitución de al menos uno de los grupos químicos de la molécula del nucleótido por un grupo químico que contenga al menos un grupo metoxietilo, como ocurre y sin que sirva de limitación en los 2'-O-metoxietil-nucleótidos.

45 Según se entiende la invención, las modificaciones en los nucleótidos para aumentar la resistencia del aptámero a la degradación por exonucleasas mediante bloqueo del grupo fosfato del extremo 5' se refieren a la unión covalente de un átomo o una molécula orgánica o inorgánica a dicho grupo fosfato, como puede ser y sin que sirva de limitación la unión mediante un enlace éster de una molécula de α,α' -[4,12-dioxo-6-[[[5-(fosfonooxi)pentil]amino]carbonil]-3,13-dioxa-5,11-diaza-1,15-pentadecanedil]bis[ω -metoxipoli(oxi-1,2-etanedil)].

Según se entiende la invención, las modificaciones en los nucleótidos para aumentar la resistencia del aptámero a la degradación por exonucleasas mediante bloqueo del grupo hidroxilo del extremo 3' se refieren a la unión covalente de un átomo o una molécula orgánica o inorgánica a dicho grupo 3'-hidroxilo, como puede ser y sin que sirva de limitación la unión mediante un conector fosfodiéster 3'→3 de un deoxirribonucleósido de timina.

50 Asimismo, otras modificaciones de los aptámeros contempladas en la presente invención son aquellas en las que al menos uno de los monómeros de ácido nucleico en aptámeros de RNA es sustituido por un monómero de LNA (descrito anteriormente), resultando así una molécula análoga a dicho RNA, que puede ser también opcionalmente modificado, como se ha explicado, con átomos de flúor o grupos amino, metilo o metoxi. Por otra parte, con objeto de aumentar la resistencia de los aptámeros a degradación por nucleasas, también resultan de

55 utilidad las formas enantioméricas de los aptámeros de RNA o DNA, y que pueden presentar opcionalmente alguna de las modificaciones descritas aquí. De este modo, la presente invención contempla también la

protección de los enantiómeros de cualquiera de los aptámeros de RNA o DNA descritos en la presente invención. Hasta el momento, la síntesis de estos ácidos nucleicos modificados se realiza mediante la inclusión del precursor monomérico modificado en la fase apropiada de la síntesis química, o mediante la incorporación directa del precursor nucleotídico modificado durante una reacción enzimática de síntesis, bien por la tolerancia de la enzima polimerasa a la incorporación de nucleótidos modificados o mediante la utilización de enzimas artificialmente modificadas para tal fin (revisiones en Klussmann, 2006; Lauridsen *et al.* 2012).

En una realización preferida del aptámero o de cualquiera de sus derivados y enantiómeros definidos anteriormente en este aspecto de la invención, además de la estructura con secuencia $\text{Mon}^1\text{—Mon}^2\text{—Mon}^3\text{—Mon}^4\text{—Mon}^5\text{—Mon}^6\text{—Mon}^7\text{—Mon}^8$, comprende también una región flanqueante previa unida al monómero Mon^1 y una región flanqueante posterior unida al monómero Mon^8 de dicha estructura. Según la invención, cada región flanqueante comprende una secuencia cuya longitud es de al menos 4 monómeros de ácido nucleico y/o de análogo químico de ácido nucleico, preferentemente de entre 4 y 75 monómeros, donde las bases nitrogenadas de al menos 3 de los 4 monómeros (preferentemente los 4 monómeros) de la región flanqueante previa más próximos a Mon^1 , contiguos o no, se aparean con las bases nitrogenadas de al menos 3 de los 4 primeros monómeros (preferentemente los 4 monómeros) de la región flanqueante posterior más próximos a Mon^8 , de tal manera que la estructura de secuencia $\text{Mon}^1\text{—Mon}^2\text{—Mon}^3\text{—Mon}^4\text{—Mon}^5\text{—Mon}^6\text{—Mon}^7\text{—Mon}^8$ forma un lazo terminal monocatenario que cierra un tallo de cadena doble formado por las regiones flanqueantes con dicho nucleótidos apareados. Esto quiere decir que, cuando el aptámero comprende ambas regiones flanqueantes con cuatro monómeros, la estructura del aptámero comprende una estructura de secuencia $\text{Mon}^{a4}\text{—Mon}^{a3}\text{—Mon}^{a2}\text{—Mon}^{a1}\text{—Mon}^1\text{—Mon}^2\text{—Mon}^3\text{—Mon}^4\text{—Mon}^5\text{—Mon}^6\text{—Mon}^7\text{—Mon}^8\text{—Mon}^{b1}\text{—Mon}^{b2}\text{—Mon}^{b3}\text{—Mon}^{b4}$, donde Mon^{a1} a Mon^{a4} son los monómeros de la región previa unida a Mon^1 , y Mon^{b1} a Mon^{b4} son los monómeros de la región posterior unida a Mon^8 . Análogamente, en el caso en que el aptámero tenga secuencias flanqueantes de entre 4 y 75 monómeros, dicho aptámero comprende una estructura de secuencia $\text{Mon}^{a(i+4)}\text{—}(\text{Mon}^{ax})_i\text{—Mon}^{a3}\text{—Mon}^{a2}\text{—Mon}^{a1}\text{—Mon}^1\text{—Mon}^2\text{—Mon}^3\text{—Mon}^4\text{—Mon}^5\text{—Mon}^6\text{—Mon}^7\text{—Mon}^8\text{—Mon}^{b1}\text{—Mon}^{b2}\text{—Mon}^{b3}\text{—}(\text{Mon}^{bx})_j\text{—Mon}^{b(j+4)}$, donde i y j se seleccionan independientemente entre un valor comprendidos entre 0 y 71, donde Mon^{a1} , Mon^{a2} , Mon^{a3} , Mon^{ax} y $\text{Mon}^{a(i+4)}$ son los monómeros de ácido nucleico o de análogo químico de ácido nucleico de la región flanqueante previa y Mon^{b1} , Mon^{b2} , Mon^{b3} , Mon^{bx} y $\text{Mon}^{b(j+4)}$ son los monómeros de ácido nucleico o de análogo químico de ácido nucleico de la región flanqueante posterior. Y además, el fragmento $\text{Mon}^{a(i+4)}\text{—}(\text{Mon}^{ax})_i\text{—Mon}^{a3}\text{—Mon}^{a2}\text{—Mon}^{a1}$ con el fragmento $\text{Mon}^{b1}\text{—Mon}^{b2}\text{—Mon}^{b3}\text{—}(\text{Mon}^{bx})_j\text{—Mon}^{b(j+4)}$ forma un tallo de cadena doble por apareamiento de al menos 3 de los monómeros del grupo compuesto por: Mon^{a1} , Mon^{a2} , Mon^{a3} , Mon^{ax} y $\text{Mon}^{a(i+4)}$, con al menos 3 de los monómeros del grupo compuesto por: Mon^{b1} , Mon^{b2} , Mon^{b3} , Mon^{bx} y $\text{Mon}^{b(j+4)}$.

Según la presente invención, las reglas de complementariedad a las que están sujetos estos apareamientos respectivos de Mon^{a1} a Mon^{a4} con Mon^{b1} a Mon^{b4} son: A con T o U; C con G; G con C o U; T con A; U con A o G.

Las regiones flanqueantes, además de servir para formar la estructura anterior al lazo que cierra el tallo de cadena doble, pueden estar definidas por secuencias que sirvan para estabilizar y dirigir el aptámero a un compartimento subcelular, como por ejemplo y sin carácter limitante, mediante el uso de horquillas del snRNA U6, tales como SEQ ID No: 4 y/o SEQ ID No: 5, de forma que el aptámero puede localizarse en el núcleo celular. En ese sentido, tales horquillas no son sino un ejemplo de las "regiones flanqueantes" descritas anteriormente, sea cual sea la estructura que forman y la funcionalidad adicional que aporten las mismas. Otro ejemplo posible sería el descrito para la localización nucleolar de un aptámero dirigido contra TAR mediante la incorporación de sendas regiones terminales del snRNA U16 (como se describe en Kolb *et al.*, 2006).

Según la invención, la longitud de la región flanqueante puede variar en función del tamaño de las regiones (sean estas funcionales o no, y de origen natural o artificial), que se desee añadir al aptámero para conferirle cualquier propiedad adicional. A modo de ejemplo y sin carácter limitante, en la presente invención, para conseguir una localización nuclear de los aptámeros al tiempo que se aumenta su resistencia a la degradación por exonucleasas, se procedió a añadir las horquillas flanqueantes del snRNA U6, pues se conoce que confieren dichas propiedades. Concretamente estas horquillas, en 5' y 3' del aptámero tienen longitudes de 25 y 24 nucleótidos respectivamente.

En otra realización preferida, el aptámero, o cualquiera de sus derivados y enantiómeros, comprende una estructura de secuencia $\text{Mon}^{a4}\text{—Mon}^{a3}\text{—Mon}^{a2}\text{—Mon}^{a1}\text{—Mon}^1\text{—Mon}^2\text{—Mon}^3\text{—Mon}^4\text{—Mon}^5\text{—Mon}^6\text{—Mon}^7\text{—Mon}^8\text{—Mon}^{b1}\text{—Mon}^{b2}\text{—Mon}^{b3}\text{—Mon}^{b4}$, donde Mon^{a1} a Mon^{a4} y Mon^{b1} a Mon^{b4} son monómeros de un ácido nucleico o de un análogo químico de ácido nucleico de tal manera que:

- i. la bases nitrogenadas de Mon^{a1} a Mon^{a4} se seleccionan independientemente del siguiente grupo: una A, una C, una G, una T y un U (donde Mon^{a3} es preferentemente A, G, T o U, y/o donde Mon^{a4} es preferentemente C, T o U.);
- ii. la base nitrogenada de Mon^{b1} es complementaria a la base nitrogenada de Mon^{a1} ;

- iii. la base nitrogenada de $\text{Mon}^{\text{b}2}$ es complementaria a la base nitrogenada de $\text{Mon}^{\text{a}2}$;
- iv. la base nitrogenada de $\text{Mon}^{\text{b}3}$ es complementaria a la base nitrogenada de $\text{Mon}^{\text{a}3}$;
- v. la base nitrogenada de $\text{Mon}^{\text{b}4}$ es complementaria a la base nitrogenada de $\text{Mon}^{\text{a}4}$;

5 y donde Mon^1 a Mon^8 son los monómeros como se definieron anteriormente. En esta realización, la región flanqueante previa comprende la secuencia de monómeros $\text{Mon}^{\text{a}4}$ — $\text{Mon}^{\text{a}3}$ — $\text{Mon}^{\text{a}2}$ — $\text{Mon}^{\text{a}1}$ y la región flanqueante posterior comprende la secuencia de monómeros $\text{Mon}^{\text{b}1}$ — $\text{Mon}^{\text{b}2}$ — $\text{Mon}^{\text{b}3}$ — $\text{Mon}^{\text{b}4}$, y además el monómero $\text{Mon}^{\text{a}1}$ se aparea o hibrida con $\text{Mon}^{\text{b}1}$, el monómero $\text{Mon}^{\text{a}2}$ se aparea con $\text{Mon}^{\text{b}2}$, el monómero $\text{Mon}^{\text{a}3}$ se aparea con $\text{Mon}^{\text{b}3}$ y el monómero $\text{Mon}^{\text{a}4}$ se aparea con $\text{Mon}^{\text{b}4}$.

10 En una realización preferida del aptámero, o de cualquiera de sus derivados y enantiómeros, los monómeros que lo constituyen son nucleótidos, pudiendo ser ribonucleótidos (aptámero de RNA) o desoxirribonucleótidos (aptámero de DNA). En una realización más preferida de la anterior, el aptámero, es una molécula de RNA.

15 En otra realización preferida del aptámero, o de cualquiera de sus derivados y enantiómeros, cuando la estructura Mon^1 — Mon^2 — Mon^3 — Mon^4 — Mon^5 — Mon^6 — Mon^7 — Mon^8 está constituida por monómeros de ácidos nucleicos, dicha estructura consiste en un octanucleótido de secuencia 5'—GGCARGGA—3', donde R representa un nucleótido seleccionado entre un nucleótido A o un nucleótido G.

20 En otra realización preferida del aptámero, o de cualquiera de sus derivados y enantiómeros, cuando la estructura Mon^1 — Mon^2 — Mon^3 — Mon^4 — Mon^5 — Mon^6 — Mon^7 — Mon^8 está constituida por monómeros de ácidos nucleicos, la estructura $\text{Mon}^{\text{a}4}$ — $\text{Mon}^{\text{a}3}$ — $\text{Mon}^{\text{a}2}$ — $\text{Mon}^{\text{a}1}$ — Mon^1 — Mon^2 — Mon^3 — Mon^4 — Mon^5 — Mon^6 — Mon^7 — Mon^8 — $\text{Mon}^{\text{b}1}$ — $\text{Mon}^{\text{b}2}$ — $\text{Mon}^{\text{b}3}$ — $\text{Mon}^{\text{b}4}$ consiste en un polinucleótido de secuencia seleccionada entre SEQ ID No: 6 y SEQ ID No: 7. SEQ ID No: 6 y SEQ ID No: 7 corresponden a ácidos nucleicos con la misma secuencia de bases nitrogenadas, siendo SEQ ID No: 6 una molécula de DNA y SEQ ID No: 7 una molécula de RNA. La secuencia de SEQ ID No: 6 es 5'-N¹N²N³N⁴GGCARGGAN⁵N⁶N⁷N⁸-3', donde N¹, N², N³ y N⁴ se seleccionan independientemente entre A, C, G o T (y preferentemente N³ es A, G o T, y/o N⁴ es T o C, SEQ ID No: 8); R se selecciona independientemente entre A y G; N⁵ es el nucleótido complementario a N⁴; N⁶ es el nucleótido complementario a N³; N⁷ es el nucleótido complementario a N²; y N⁸ es el nucleótido complementario a N¹. La secuencia de SEQ ID No: 7 es 5'-N¹N²N³N⁴GGCARGGAN⁵N⁶N⁷N⁸-3', donde N¹, N², N³ y N⁴ se seleccionan independientemente entre A, C, G o U (y preferentemente N³ es A, G o U, y/o N⁴ es U o C, SEQ ID No: 9); R se selecciona entre A y G; N⁵ es el nucleótido complementario a N⁴; N⁶ es el nucleótido complementario a N³; N⁷ es el nucleótido complementario a N²; N⁸ es el nucleótido complementario a N¹. Las reglas de complementariedad a las que están sujetos estos apareamientos son: A con T o con U, C con G, G con C o con U, T con A, U con A o con G.

35 En otra realización preferida del aptámero, o de cualquiera de sus derivados y enantiómeros, cuando la estructura Mon^1 — Mon^2 — Mon^3 — Mon^4 — Mon^5 — Mon^6 — Mon^7 — Mon^8 o la estructura $\text{Mon}^{\text{a}4}$ — $\text{Mon}^{\text{a}3}$ — $\text{Mon}^{\text{a}2}$ — $\text{Mon}^{\text{a}1}$ — Mon^1 — Mon^2 — Mon^3 — Mon^4 — Mon^5 — Mon^6 — Mon^7 — Mon^8 — $\text{Mon}^{\text{b}1}$ — $\text{Mon}^{\text{b}2}$ — $\text{Mon}^{\text{b}3}$ — $\text{Mon}^{\text{b}4}$ está constituida por monómeros de ácidos nucleicos, el aptámero o su derivado comprende una secuencia nucleotídica seleccionada del grupo compuesto por: SEQ ID No: 10 (IX03-1), SEQ ID No: 11 (IX24-1), SEQ ID No: 12 (IX36-1), SEQ ID No: 13 (X02-1), SEQ ID No: 14 (X04-2), SEQ ID No: 15 (X09-1), SEQ ID No: 16 (X10-7), SEQ ID No: 17 (X13-1), SEQ ID No: 18 (X36-1), SEQ ID No: 19 (X41-2), SEQ ID No: 20 (XI1-17), SEQ ID No: 21 (XI21-7), SEQ ID No: 22 (XI23-3), SEQ ID No: 23 (XI141-2), SEQ ID No: 24 (XI13-1), SEQ ID No: 25 (XI101-1), SEQ ID No: 26 (XI149-1), SEQ ID No: 27 (XI107-1), SEQ ID No: 28 (XI129-1), SEQ ID No: 29 (XIV22-23), SEQ ID No: 30 (XIV26-6), SEQ ID No: 31 (XIV1-2), SEQ ID No: 32 (XIV32-1), SEQ ID No: 33 (XIV5-1), SEQ ID No: 34 (XIV12-1), SEQ ID No: 35 [RNA16(+)], SEQ ID No: 36 [RNA16(+)]var-G], SEQ ID No: 37 (L-XIV22-23) y SEQ ID No: 38 (L-XIV26-6). En otra realización preferida como la anterior, la secuencia nucleotídica del aptámero o de su derivado se selecciona del consenso SEQ ID No: 39 [RNA16(+)] genérico], esto es SEQ ID No: 35 [RNA16(+)] o SEQ ID No: 36 [RNA16(+)] var-G], y más preferentemente es SEQ ID No: 35 [RNA16(+)].

A lo largo de la presente descripción, los aptámeros definidos en este aspecto de la invención, o cualquiera de sus derivados y enantiómeros, pueden aparecer igualmente referidos como "aptámeros de la invención".

50 Un segundo aspecto de la invención se refiere a una construcción genética de DNA útil para la síntesis por transcripción, bien *in vitro* (IVT) o bien intracelular, de un aptámero de RNA de la invención, caracterizado porque comprende una secuencia de nucleótidos de DNA que codifica la secuencia de nucleótidos del aptámero de RNA.

Un tercer aspecto de la invención se refiere a un procedimiento de obtención de un aptámero de la invención, que comprende la combinación de técnicas experimentales de selección *in vitro* de ácidos nucleicos con técnicas computacionales de optimización que permiten analizar tanto las secuencias generadas por selección *in vitro*

como sus correspondientes estructuras secundarias. Sin que sirva de limitación, un ejemplo del procedimiento utilizado se detalla más adelante en los Ejemplos 1 y 2. En estos casos, las técnicas experimentales consisten en:

- 5 i. Generación de un molde de DNA para la obtención por transcripción *in vitro* de una población de moléculas de RNA. Este molde de DNA bicatenario se obtuvo por hibridación y extensión enzimática de dos oligonucleótidos de DNA parcialmente complementarios diseñados al efecto. Estos oligonucleótidos se obtuvieron comercialmente y fueron generados por síntesis química de DNA. Uno de dichos oligonucleótidos contiene la secuencia del promotor de la RNA polimerasa del fago T7 en su extremo 5', seguida por una región de 25 nt de secuencia aleatoria. Un ejemplo de oligonucleótidos adecuados es el par 5'EcoRIK (SEQ ID No. 40) y 3'RANDOMK (SEQ ID No. 41).
- 10 ii. Obtención de una población de moléculas de RNA consistentes en una región de secuencia aleatoria y 25 nt de longitud flanqueada por sendas regiones anterior y posterior de secuencia constante y conocida, de 10 y 29 nt de longitud respectivamente. Esta población se generó por IVT usando como molde el obtenido en el paso "i";
- 15 iii. Selección de aptámeros contra la molécula diana de RNA que contiene los primeros 308 nt de la 5'UTR del VIH-1 (SEQ ID No: 1), inmovilizada en un soporte sólido de sefarosa-estreptavidina, controlando variables como temperatura de unión, tiempo de unión y relación molecular aptámero:diana;
- iv. Recuperación de los aptámeros unidos a la molécula diana;
- v. Transcripción reversa de las moléculas de RNA correspondientes a los aptámeros seleccionados mediante reacción enzimática, utilizando la actividad transcriptasa reversa de la enzima Tth procedente de *Thermus thermophilus*. De esta forma se obtiene una población de moléculas de DNA copia (cDNA) de cadena simple, de secuencia complementaria a la de las moléculas de RNA de partida;
- 20 vi. Amplificación del cDNA obtenido en el paso "v" mediante PCR, usando como cebadores oligonucleótidos sintéticos de DNA y la actividad DNA polimerasa DNA-dependiente también presente en la enzima Tth. Uno de ellos debe contener la secuencia del promotor de la RNA polimerasa del fago T7 de manera que ésta quede incluida en el producto amplificado en una posición equivalente a la que presenta en el molde de DNA generado en el paso "i";
- 25 vii. Clonaje del producto de la PCR del paso "vi" en un vector plasmídico comercial, utilizando bacterias competentes de cepas de *Escherichia coli*;
- 30 viii. Repetición sucesiva de ciclos o rondas experimentales definidos por los pasos "ii" a "vi", siendo preferida dicha repetición de 14 rondas como por ejemplo se describe en el caso de la presente invención;
- 35 viii. Secuenciación de un mínimo de 30 clones moleculares producto del paso "vii" de cada ronda de selección.

Las técnicas computacionales utilizadas a continuación de las experimentales consisten en:

- 35 ix. Alineamiento de las secuencias correspondientes a todos los aptámeros obtenidos en el proceso de selección, determinación de sus longitudes y de las repeticiones de secuencias producidas en cada ronda, o entre rondas diferentes;
- x. Cuantificación de la diferencia o "distancia" existente entre la secuencia de todas las moléculas, medida como el número de mutaciones entre ellas (distancia de Hamming);
- 40 xi. Análisis de las secuencias mediante alineamiento múltiple y programas de "clustering" basados en distancias nucleotídicas, y caracterización de grupos y subgrupos de secuencias que presentan mayor similitud de secuencia;
- xii. Plegamiento *in silico* de todas las moléculas de RNA generadas durante el proceso de selección en su estructura de mínima energía o "MFE structure" empleando para ello el programa RNAfold del Vienna RNA package;
- 45 xiii. Representación de las estructuras en la notación dot-bracket (en la que los nucleótidos desapareados se indican por puntos, y los que forman parte de pares de bases mediante paréntesis) y alineamiento de dichas estructuras representadas en tal notación;

- xiv. Cuantificación de la distancia existente entre las estructuras en la notación dot-bracket, empleando para ello tres aproximaciones complementarias denominadas "Hamming", "Base pair" y "Tree edit", tal como se describe en detalle en el Ejemplo 2 de la presente invención;
- 5 xv. Análisis de las estructuras alineadas mediante alineamiento múltiple y programas de "clustering" basados en las distancias obtenidas, y caracterización de grupos y subgrupos de secuencias que presentan mayor similitud estructural;
- xvi. Análisis de todas las estructuras, y en especial de las que forman parte de los grupos puestos en evidencia en el punto xv, en busca de motivos estructurales comunes;
- 10 xvii. Análisis de las secuencias presentes en las moléculas que comparten dichos motivos estructurales comunes;
- xviii. Caracterización del motivo estructural mínimo (o "aptámero mínimo") compartido por los aptámeros detectados en el punto xvii, y análisis de la posibilidad de diseñar moléculas de RNA sintéticas que garanticen la presencia de las secuencias detectadas en las regiones desapareadas de dichos dominios, haciendo más estable las regiones que forman sus tallos bicatenarios;
- 15 xix. Análisis de las posibles interacciones entre los aptámeros seleccionados y la diana de RNA utilizada durante el proceso de selección, y también de la interacción entre el "aptámero mínimo" y la diana, estudiando tanto la posición de la diana en la que se produce dicha interacción como la estabilidad termodinámica de cada apareamiento aptámero-diana;
- 20 xx. Diseño de una molécula utilizable como "control negativo" de interacción, cuya estructura sea la misma que la presente en el "aptámero mínimo" pero cuya secuencia en su región desapareada tenga la menor probabilidad de interacción con el RNA diana.

En esencia, y de manera general, las técnicas experimentales de selección *in vitro* de ácidos nucleicos para la obtención de los aptámeros de la invención comprenden:

- 25 a. Obtener por transcripción *in vitro* moléculas de RNA de secuencia aleatoria, de al menos 25 nucleótidos de longitud, flanqueada por regiones anterior y posterior de secuencia constante y conocida preferentemente por transcripción de un molde de DNA bicatenario obtenido previamente por hibridación y extensión del par de oligonucleótidos 5'EcoRIK (SEQ ID No. 40) y 3'RANDOMK (SEQ ID No. 41);
- 30 b. Seleccionar y aislar las moléculas de RNA obtenidas en la etapa anterior con mayor afinidad por la molécula diana correspondiente a la región 5'UTR del VIH-1 (SEQ ID No 1). Preferentemente este proceso se consigue por purificación en columna de afinidad con dicha molécula inmovilizada en un soporte sólido de sefarosa-estreptavidina;
- c. Obtener y amplificar un cDNA obtenido a partir de las moléculas de RNA aisladas en la etapa anterior, preferentemente con el par de oligonucleótidos 5'EcoRIK (SEQ ID No: 40) y 3'XhoIK (SEQ ID No: 42).
- 35 d. Secuenciar el cDNA amplificado, preferentemente por clonación y secuenciación de clones;
- e. Repetir sucesivos ciclos definidos por las etapas "a-d", utilizando como molde de transcripción de la etapa "a" el producto amplificado en la etapa "c", para obtener en la población resultante un enriquecimiento progresivo de moléculas con mayor capacidad de unión a la diana.

En cuanto a las técnicas computacionales de optimización de las secuencias de aptámeros generadas por las técnicas selección *in vitro* anteriores, básicamente comprende:

- 40 f. Clasificar las secuencias de cDNA secuenciadas en los distintos ciclos "a-d" repetidos, comprendiendo cuantificación de la diferencia nucleotídica existente entre dichas secuencias, alineamiento múltiple y clústering o agrupamiento de secuencias de dichos cDNA, plegamiento de la secuencia de RNA derivada del cDNA a la estructura de mínima energía, caracterización del motivo estructural mínimo entre secuencias y análisis de interacciones entre secuencias de DNA y la diana de RNA 5'UTR del VIH-1 (SEQ ID No: 1). En realizaciones preferidas, esta clasificación de secuencias comprende las etapas ix a xx anteriormente definidas.
- 45

Un cuarto aspecto de la invención hace referencia al uso de un aptámero de la invención como molécula con capacidad de unirse específicamente a la región 5'UTR del genoma del VIH-1. Los aptámeros de la invención

tienen por tanto utilidad como agente de interferencia de una función y/o actividad de la región 5'UTR del genoma del VIH-1.

5 En el ámbito de la invención, el término "función y/o actividad de la región 5'UTR" se refiere al conjunto de procesos fisiológicos en los que está implicada la región 5'UTR completa, alguna de sus subregiones (R o U5) o alguno de sus elementos funcionales (TAR, Poly-A, PBS, DIS, SD y Psi). Entre estas funciones destacan la activación de la transcripción, la poliadenilación de los RNAs virales genómicos y subgenómicos, el reclutamiento del cebador para la transcripción reversa, el cambio de cadena molde durante la transcripción reversa, la dimerización de los RNAs genómicos y su empaquetamiento en las partículas virales, el *splicing* o procesamiento diferencial de diferentes poblaciones de RNA genómico y la traducción de los RNAs virales.

10 Un quinto aspecto de la invención hace referencia al uso de un aptámero de la invención como inhibidor de la replicación del VIH-1 en un cultivo celular.

Un sexto aspecto de la invención hace referencia al uso de un aptámero de la invención para detectar, caracterizar y/o cuantificar VIH-1 *in vitro*. En este caso, la unión del aptámero a su diana indicaría la presencia de VIH-1, preferentemente de aislados con la diana de unión del aptámero conservada, pudiendo ser directamente 15 relacionable la cantidad de aptámero unido con la cantidad de unidades del genoma de VIH-1. Igualmente, en otra aplicación relacionada la localización del aptámero una vez interiorizado en una célula susceptible de estar infectada puede indicar la localización del genoma de VIH-1 (o de fragmentos subgenómicos del mismo que mantengan su 5'UTR) mediante la utilización de técnicas de hibridación fluorescente *in situ* (FISH, de *fluorescence in situ hybridization*), inmunofluorescencia, inmunohistoquímica o similares. En estos casos, el 20 aptámero no modificado puede ser detectable y cuantificable mediante técnicas de transcripción reversa y PCR cuantitativa. El aptámero puede también ser modificado para conferirle alguna propiedad detectable directa o indirectamente, permitiendo igualmente su localización en una muestra biológica. Así, el aptámero puede detectarse, cuantificarse y localizarse directamente si ha sido sintetizado utilizando radioisótopos incorporados a los componentes estructurales del ácido nucleico o análogo de ácido nucleico. La detección y cuantificación se 25 podría realizar mediante instrumentos de medición directa (por ejemplo, contador Geiger Müller) o indirecta (por ejemplo, contador de centelleo), y la localización mediante radiografía o sistemas de almacenamiento de fosforescencia. Adicionalmente, el aptámero puede detectarse, cuantificarse y localizarse indirectamente mediante la detección, cuantificación y localización de un componente covalentemente unido al aptámero a tal efecto, que podría consistir, sin que este listado limite las posibilidades, en: un fluoróforo detectable por su 30 emisión de radiación de fluorescencia en un rango de longitud de onda determinada tras una excitación específica mediante radiación en un rango de longitud de onda diferente; biotina o avidina, moléculas que se pueden unir a la proteína estreptavidina convenientemente marcada; un compuesto orgánico o inorgánico susceptible de ser detectado por anticuerpos específicos y detectable por técnicas tales como inmunofluorescencia o inmunohistoquímica; una molécula con capacidad enzimática detectable porque catalice 35 específicamente alguna reacción colorimétrica o luminiscente. Finalmente, el aptámero podría detectarse indirectamente mediante la detección, por las técnicas arriba mencionadas, de otro ácido nucleico con capacidad para hibridar específicamente con él.

Un séptimo aspecto de la invención hace referencia al uso de un aptámero de la invención como molécula biosensora para la detección y/o cuantificación de VIH-1 en una muestra biológica, como por ejemplo, y sin 40 limitarse a, plasma, sangre, semen, saliva, tejidos, preparaciones histológicas, extractos celulares u otros. En estos casos, el aptámero podría ser detectado, cuantificado y localizado de forma similar a como se describe en el anterior aspecto de la invención.

Un octavo aspecto de la invención hace referencia al uso de un aptámero de la invención para la fabricación de una composición farmacéutica.

45 En un noveno aspecto, la invención se refiere a dicha composición farmacéutica. La composición farmacéutica según la presente invención comprende al menos un aptámero, o cualquiera de sus derivados y enantiómeros. Dicha composición puede ser igualmente referida en la presente descripción como "composición farmacéutica de la invención" o "composición de la invención", y puede comprender simultáneamente más de un aptámero de cualquiera de los definidos en el primer aspecto de esta invención. Adicionalmente, puede comprender otras 50 moléculas terapéuticas o potenciadoras, junto con, opcionalmente, uno o más adyuvantes y/o vehículos farmacéuticamente aceptables.

De acuerdo con la presente invención, las composiciones farmacéuticas de la invención se pueden administrar a humanos o se pueden usar en medicina veterinaria, particularmente con otros mamíferos. La administración se llevará a cabo con una "dosis terapéuticamente eficaz", siendo suficiente para que demuestre un beneficio para 55 el paciente, por lo general relacionado con la disminución de la capacidad replicativa del virus dentro del organismo infectado. Tal beneficio puede suponer la mejora de al menos un síntoma relacionado con la infección por VIH-1. La prescripción del tratamiento, es decir, las decisiones sobre las dosis, periodicidad, duración, etc., recaerá bajo la responsabilidad del médico general o del especialista que atienda al paciente infectado.

En el ámbito de la presente invención, el término "dosis terapéuticamente eficaz" se refiere a una cantidad (o cantidad por unidad de masa del individuo al que se le administra) de un fármaco o compuesto que ocasiona un efecto terapéutico detectable en un individuo o un grupo de individuos a los que les es administrada, ocasionando mínimos efectos secundarios o tóxicos. El término "dosis terapéuticamente eficaz-50" o "dosis terapéuticamente eficaz-95" incluye un valor estadístico en el que el efecto terapéutico debe ser detectable en el 50% o el 95% de los individuos a los que les es administrada. En lo relativo a los efectos tóxicos del fármaco o compuesto, es preferible que la dosis terapéutica eficaz no ocasione ninguno. Sin embargo, aunque en ocasiones puedan darse efectos tóxicos, puede llegarse a un compromiso en el que se considera que éstos son preferibles al desarrollo normal de la enfermedad, infección o dolencia sin la administración del fármaco o compuesto, y pueden a su vez ser tratados mediante terapias adicionales. Tal es el caso de los fármacos actuales en el tratamiento contra el VIH.

Las composiciones farmacéuticas, de acuerdo con la presente invención, pueden comprender, además del ingrediente activo, uno o más adyuvantes y/o vehículos farmacéuticamente aceptables y conocidos para los expertos en la materia. Tales materiales deben no ser tóxicos y, además, no interferir con la eficacia del ingrediente activo (el aptámero, en el caso de la presente invención). La naturaleza precisa del transportador o del material coadyuvante dependerá de la vía de administración, que puede ser, por ejemplo y sin que sirva de criterio de exclusión, oral, por inhalación, por absorción cutánea o mucosa, o mediante inyección, por ejemplo, cutánea, subcutánea o intravenosa. Son especialmente preferidas aquellas composiciones farmacéuticas en las que la molécula activa se encuentra en una forma habitual para hacerla llegar al interior de una célula (en este caso, una célula infectada por el VIH-1), como por ejemplo en el interior de una vesícula formada por ácidos grasos, de un liposoma formado por lípidos, de una nanoestructura formada por proteínas u otros componentes moleculares para tales fines conocidos por los expertos en el campo. Las formas farmacéuticas para la administración oral pueden ser en forma de tabletas, cápsulas, polvo o disolución líquida. Una tableta puede contener un transportador sólido tal como gelatina o un adyuvante. Las composiciones farmacéuticas líquidas generalmente comprenden un transportador líquido tal como agua, aceites animales, vegetales o sintéticos. Las soluciones salinas fisiológicas pueden contener dextrosa u otra solución de sacarido o de glicol tal como etilenglicol, propilenglicol o polietilenglicol. Para la inyección intravenosa, cutánea o subcutánea, o bien inyección en el lugar de la dolencia, el principio activo estará en forma de una solución acuosa parenteralmente aceptable, que sea apirógena y que tenga un pH apropiado, estabilidad y tonicidad. Aquellos que sean expertos en la técnica podrán preparar soluciones apropiadas utilizando, por ejemplo, vehículos isotónicos tal como una inyección salina, inyección con solución de Ringer e inyección con solución de Ringer con lactato. Cuando se requieran, se podrán incluir conservantes, estabilizadores, sistemas tampón, antioxidantes y/o otros aditivos.

Otro aspecto de la invención hace referencia al uso en medicina de los aptámeros de la invención, y por tanto también al uso de al menos un aptámero de la invención para la fabricación de una composición farmacéutica ya sea para el tratamiento y/o prevención de una enfermedad (preferentemente una enfermedad causada por el VIH, tales como una infección por VIH-1 o el SIDA) o para inhibir la replicación del VIH en un paciente infectado por el mismo. Debe entenderse que este aspecto de la invención también protege un método de tratamiento y/o prevención de una enfermedad producida por el VIH (preferentemente por VIH-1) en un sujeto, que comprende administrar una cantidad terapéuticamente efectiva de un aptámero de la invención.

Descripción de las figuras

Figura 1.- Alineamiento de las secuencias correspondientes al dominio Poli-A presente en la 5'UTR de diferentes aislados virales de VIH-1 y SIV de chimpancé. En la figura se muestra una adaptación de un alineamiento del compendio de secuencias publicado en Kuiken C., *et al.* (2011) que comprende la secuencia del dominio Poli-A. En la parte superior se muestra un consenso de secuencia del aislado de referencia B.FR.83.HXB2 y debajo se alinean las secuencias de los diferentes aislados de VIH-1 (correspondientes a los grupos M -con distintos subtipos y formas recombinantes-, N, O y P del virus) de los que se tiene registro de la secuencia del dominio Poli-A, así como, a modo comparativo, diferentes aislados del virus de la inmunodeficiencia de simios (SIV) que afecta al chimpancé. Los puntos (.) significan que no hay registro de esa posición en el fragmento secuenciado del aislado; los guiones (-) significan que la posición secuenciada en el aislado tiene la misma base nucleotídica que en la secuencia del aislado de referencia; cualquier modificación de la secuencia de un aislado con respecto a la secuencia del aislado de referencia se indica con la letra (A, C, G o T) que representa al nucleótido correspondiente. En el recuadro se enmarca la secuencia de unión del octanucleótido 5'-GGCARGGA-3' presente en los aptámeros descritos en esta invención (más detalles en el texto, y en la figura 8).

Figura 2.- Análisis de la variabilidad de secuencias, mediante programas de clustering y filogenia basados en distancias. Los grupos 1 a 3 de secuencias se marcan con rectángulos. Las secuencias que poseen el motivo conservado se marcan con círculos negros.

Figura 3.- Progresión de la distancia entre secuencias (Hamming) y estructuras (base pair) dentro de cada ronda a lo largo del proceso de selección *in vitro*. La tendencia global de los valores medios es

decreciente en ambos casos, con las excepciones de un incremento de distancias intra-ronda entre estructuras al pasar de la ronda III a la V, y de la VIII a la IX. Este incremento puede deberse a la selección simultánea de dos o más grupos cuyas diferencias mutuas en secuencia y/o estructura son mayores de las que se obtendrían en media comparando con secuencias al azar.

5 **Figura 4.- Clasificación de las estructuras a partir de sus distancias *base pair*, empleando los mismos programas que en la clasificación de secuencias de la Fig. 2.** Los grupos 1 a 3 de secuencias se marcan con rectángulos, como en la Fig. 2. Las secuencias que poseen el motivo conservado se marcan con círculos negros. Se muestran las estructuras MFE características de los grupos 1, 2a y 2b.

10 **Figura 5.- Estructuras MFE correspondientes a los aptámeros de la ronda XIV, señalando con un rectángulo el lazo 5'-GGCAAGGA-3' (5'-GGCAGGGA-3' en XIV32) que forma parte del motivo conservado con estructura en notación dot-bracket (((.....))).** El número N de repeticiones de cada secuencia se muestra como un sufijo "_N". El cuadro superior izquierdo muestra las estructuras correspondientes a las secuencias del grupo 2, el superior derecho las del grupo 1, y el inferior las otras tres secuencias de la ronda XIV que no forman parte de estos grupos. Se observa que las secuencias pertenecientes a los grupos 1 y 2 también poseen una relación estructural clara. Las energías libres de plegamiento correspondientes a cada secuencia (en Kcal/mol) son $\Delta G[XIV22]=-10,80$; $\Delta G[XIV1]=-11,80$; $\Delta G[XIV32]=-10,80$; $\Delta G[XIV12]=-10,90$; $\Delta G[XIV26]=-15,70$; $\Delta G[XIV5]=-11,10$; $\Delta G[XIV48]=-9,80$; $\Delta G[XIV37]=-7,32$; $\Delta G[XIV25]=-11,50$. Más detalles en el texto.

15 **Figura 6.- Secuencias de las rondas IX a XIV que contienen el motivo de secuencia 5'-GGCARGGA-3' con estructura en notación dot-bracket (((.....))).** Al alinear las secuencias en torno a dicho motivo se comprueba la gran degeneración en los nucleótidos que constituyen el tallo (stem) de 4 bp que cierra el lazo (loop) de secuencia conservada.

20 **Figura 7.- Estructura secundaria correspondiente a la molécula RNA16(+)** de secuencia 5'-CCCCGGCAAGGAGGGG-3' (SEQ ID No. 35), considerada como el motivo mínimo para interactuar específicamente con la 5'UTR de VIH-1. La energía libre asociada a esa MFE es de -6,50 kcal/mol, y la frecuencia de dicha MFE dentro del repertorio estructural (ED) de esta secuencia es del 91,82 %.

25 **Figura 8.- Predicción bioinformática de interacciones entre la diana UTR₃₀₈** (SEQ ID No: 1) y cuatro de las moléculas generadas en la presente invención: los aptámeros XIV22-23 (ejemplo del grupo 2a), XIV1-2 (grupo 2b), XIV26-6 (grupo 1), y la molécula RNA16(+). En (A) y (B) se muestra gráficamente la energía libre total de las posibles interacciones de regiones de la molécula UTR₃₀₈ (SEQ ID No: 1) con las cuatro moléculas estudiadas. En el eje X se representa las posiciones (nt) de las regiones de la UTR₃₀₈ y en el eje Y el valor de energía libre total de la unión. Marcados desde (a) a (f) se encuentran los seis valles que se correlacionan con posibles zonas de unión de UTR₃₀₈ a los aptámeros. En (C) se muestran, marcadas de (a) a (e) como en los paneles (A) y (B), las cinco posibles regiones de interacción de los aptámeros con la molécula UTR₃₀₈ plegada en forma BMH. La región (a), situada en el dominio Poli-A, es la que genera la interacción más estable: para ella se muestra el apareamiento con el motivo 5'-GGCAAGGA-3' (enmarcado), identificado en las cuatro moléculas estudiadas XIV22-23 (SEQ ID No: 29), XIV1-2 (SEQ ID No: 31), XIV26-6 (SEQ ID No: 30) y RNA16(+). (SEQ ID No: 35).

30 **Figura 9.- Esquema de los RNAs inhibidores L-XIV22-23 (SEQ ID No: 37) y L-XIV26-6 (SEQ ID No: 38) utilizados en los ensayos de unión *in vitro* a la región UTR₃₀₈** (SEQ ID No: 1) y en los de inhibición en cultivos celulares del VIH-1. Los aptámeros XIV22-23 y XIV26-6 se clonaron entre los sitios *KpnI* y *Apal* del vector pU6 (Sánchez-Luque *et al.*, 2010). Como RNA control se utilizó la construcción vacía (llamada L-vacío).

35 **Figura 10.- Ejemplo de unión de la molécula L-XIV22-23 (SEQ ID No: 37) a la molécula UTR₃₀₈** (SEQ ID No: 1). A la izquierda se muestra una autorradiografía del resultado de la electroforesis en gel nativo de las reacciones de unión de la molécula L-XIV22-23 (SEQ ID No: 37) a concentraciones crecientes de molécula UTR₃₀₈ (SEQ ID No: 1). Los puntos que aparecen a la derecha del gel indican los diferentes conformeros de la molécula L-XIV22-23 (SEQ ID No: 37) no unida, pues se detectan ya en el primer carril (donde no hay molécula diana). Las puntas de flecha indican complejos de unión entre ambas moléculas. La cuantificación total de los complejos de unión dividido entre la cuantificación total de complejos y molécula libre nos dan el tanto por uno ($\theta/1$) de unión. La gráfica de la derecha muestra la media y el error estándar de tres experimentos independientes. La curva a la que se ajusta la unión de esta molécula sigue el modelo de unión cooperativa con un único sitio de unión y se representa como una línea punteada.

40 **Figura 11.- Ejemplo de unión de la molécula L-XIV26-6 (SEQ ID No: 38) a la molécula UTR₃₀₈** (SEQ ID No: 1). A la izquierda se muestra una autorradiografía del resultado de la electroforesis en gel nativo de las reacciones de unión de la molécula L-XIV26-6 (SEQ ID No: 38) a concentraciones crecientes de molécula UTR₃₀₈ (SEQ ID No: 1). Los puntos que aparecen a la derecha del gel indican los diferentes conformeros de la molécula L-XIV26-6 (SEQ ID No: 38) no unida, pues se detectan ya en el primer carril (donde no hay molécula diana). Las

puntas de flecha indican complejos de unión entre ambas moléculas. La cuantificación total de los complejos de unión dividido entre la cuantificación total de complejos y molécula libre nos dan el % de unión. La gráfica de la derecha muestra la media y el error estándar de tres experimentos independientes. La curva a la que se ajusta la unión de esta molécula sigue un modelo hiperbólico y se representa como una línea punteada.

5 **Figura 12.- Ejemplo de unión de la molécula RNA16(+)** (SEQ ID No: 35) **a la molécula UTR₃₀₈** (SEQ ID No: 1). A la izquierda se muestra una autorradiografía del resultado de la electroforesis en gel nativo de las reacciones de unión de la molécula RNA16(+)

10 (SEQ ID No: 35) a concentraciones crecientes de molécula UTR₃₀₈ (SEQ ID No: 1). Los puntos que aparecen a la derecha del gel indican los diferentes conformeros de la molécula RNA16(+)

15 (SEQ ID No: 35) no unida, pues se detectan ya en el primer carril (donde no hay molécula diana). Las puntas de flecha indican complejos de unión entre ambas moléculas. La cuantificación total de los complejos de unión dividido entre la cuantificación total de complejos y molécula libre nos dan el % de unión. La gráfica de la derecha muestra la media y el error estándar de tres experimentos independientes. La curva a la que se ajusta la unión de esta molécula sigue un modelo hiperbólico y se representa como una línea punteada.

20 **Figura 13.- Ejemplo de inhibición de la replicación del VIH-1 en cultivos celulares por la molécula L-XIV22-23** (SEQ ID No: 37). En la figura se muestra el resultado de la cuantificación del antígeno viral p24 en el sobrenadante celular de cultivos de células HEK293T cotransfectados con 100 ng de pNL4-3 y 500 ng de RNA inhibidor 48 h previamente a la medida. En la gráfica se muestra la media y el error estándar en tres experimentos independientes del porcentaje de antígeno p24 respecto al registrado con el RNA control L-vacío en cada triplicado. El análisis ANOVA muestra que los resultados son estadísticamente significativos (** refleja valor $p < 0.01$).

25 **Figura 14.- Ejemplo de inhibición de la replicación del VIH-1 en cultivos celulares por la molécula L-XIV26-6** (SEQ ID No: 38). En la figura se muestra el resultado de la cuantificación del antígeno viral p24 en el sobrenadante celular de cultivos de células HEK293T cotransfectados con 100 ng de pNL4-3 y 500 ng de RNA inhibidor 48 h previamente a la medida. En la gráfica se muestra la media y el error estándar en tres experimentos independientes del porcentaje de antígeno p24 respecto al registrado con el RNA control L-vacío en cada triplicado. El análisis ANOVA muestra que los resultados son estadísticamente significativos (** refleja valor $p < 0.01$).

30 **Figura 15.- Ejemplo de inhibición de la replicación del VIH-1 en cultivos celulares por la molécula RNA16(+)** (SEQ ID No: 35). En la figura se muestra el resultado de la cuantificación del antígeno viral p24 en el sobrenadante celular de cultivos de células HEK293T cotransfectados con 100 ng de pNL4-3 y 500 ng de RNA inhibidor 48 h previamente a la medida. En la gráfica se muestra la media y el error estándar en tres experimentos independientes del porcentaje de antígeno p24 respecto al registrado con el RNA control RNA16(-) en cada triplicado. El análisis ANOVA muestra que los resultados son estadísticamente significativos (** refleja valor $p < 0.01$).

35 Ejemplos de realización de la invención

Los aptámeros desarrollados poseen como secuencia común mínima a todos ellos el octanucleótido 5'-GGCARGGA-3' (notación IUB para nt degenerados: R corresponde bien a A ó bien a G), es decir, 5'-GGCAAGGA-3' o bien 5'-GGCAGGGA-3'. En ambos casos, esta secuencia forma un lazo terminal monocatenario que cierra un tallo de cadena doble de al menos tres pares de nucleótidos. Los aptámeros que contienen estas secuencias mínimas pueden poseer regiones flanqueantes tanto en su extremo 5' como 3' con longitudes comprendidas entre 1 y 75 nucleótidos, sea cual sea la estructura que forman dichas regiones flanqueantes y la funcionalidad adicional que aporten las mismas. En particular, en la presente invención se han desarrollado aptámeros de una longitud total de 64 nucleótidos, todos los cuales contienen los octanucleótidos indicados. (Ejemplos 1 y 2).

45 En base a este conocimiento y a la citada aproximación combinada *in vitro* / *in silico*, se ha logrado definir una molécula concreta de RNA de tan sólo 16 nucleótidos, llamada "RNA16(+)", con secuencia SEQ ID No: 35 (5'-CCCCGGCAAGGAGGG-3'). Esta molécula contiene el motivo 5'-GGCAAGGA-3' en un lazo terminal que cierra un tallo muy estable de 4 pares de bases G-C (Ejemplo 3). Se ha sintetizado dicha molécula RNA16(+) y se ha sometido a los mismos ensayos que los aptámeros de 64 nucleótidos obtenidos por selección *in vitro*.

50 Se ha observado mediante análisis bioinformáticos que los aptámeros obtenidos *in vitro* y la molécula RNA16(+) se unen específicamente al dominio Poli-A de la 5'UTR de VIH-1 (Ejemplo 4).

La interacción específica entre los aptámeros seleccionados o la molécula RNA16(+) con la 5'UTR de VIH-1 se ha comprobado *in vitro* (Ejemplos 5.1 a 5.3). Además, se ha demostrado que la unión aptámero-ligando, o bien RNA16(+)-ligando, produce la inhibición de la replicación de VIH-1 en cultivo celular, lo que se determina mediante ensayos específicos (Ejemplos 5.4 a 5.6).

Ejemplo 1.- Proceso de obtención de los aptámeros obtenidos por selección *in vitro*.

Se ha realizado un proceso de selección *in vitro* de aptámeros de RNA con capacidad de unirse específicamente a la 5'UTR de VIH-1. La población de partida utilizada, denominada "Ronda 0", fue generada por transcripción *in vitro* (*in vitro* transcription o IVT) utilizando la enzima T7 RNA polimerasa y un molde sintético. Este molde consistió en un conjunto de moléculas de DNA de 84 nucleótidos (nt) de longitud obtenido por hibridación y extensión de los oligonucleótidos 5'EcoRIK (SEQ ID No: 40, 5'-GGATAATACGACTCACTATAGGGAATTCAA-3'), que incorpora el promotor de la T7 RNA polimerasa, y 3'RANDOMK (SEQ ID No: 41, 5'-GTGAGCTCGTAGTATCAGATCACTCCATNNNNNNNNNNNNNNNNNNNNNNNNNTTGAATTCCTATAGTG-3'), donde N indica cada una de las 25 posiciones aleatorias. Resulta así una población de moléculas de RNA (Ronda 0) de 64 nt de longitud que varían entre sí tan sólo en la región central de 25 nt de secuencia aleatoria, flanqueada por dos regiones de secuencia fija, a las que se unirán los oligonucleótidos cebadores o primers utilizados durante el proceso.

La molécula diana utilizada para la selección de aptámeros fue un fragmento de 308 nt correspondiente a la 5'UTR del genoma de VIH-1 clonado en el plásmido pNL4-3 VIH-1 (abreviada como UTR₃₀₈), y con secuencia SEQ ID No: 1. Dicha diana se obtuvo por IVT a partir de una molécula de DNA generada por PCR utilizando los oligonucleótidos 5'T7pNL4-3 (SEQ ID No: 43, 5'-TAATACGACTCACTATAGGGTCTCTCTGGTTAG-3') y 3'T7pNL4-3 (SEQ ID No: 44, 5'-AATTTTTGGCGTACTCACCAGT-3'). El uso como diana de un fragmento que contiene los primeros 308 nucleótidos de los 336 que posee la 5'UTR permite la selección de aptámeros frente a dicha región en un entorno estructural que representa la situación natural en el genoma del virus, lo que no se había logrado previamente por otros autores que han utilizado como diana elementos estructurales aislados de la 5'UTR, por ejemplo el elemento TAR (Ducongé y Toulmé, 1999; Kolb *et al.*, 2006; Watrin *et al.*, 2009).

El proceso de síntesis de una molécula diana de este tipo, descrito en detalle en (Romero-López *et al.*, 2005), incluyó la biotilización interna durante el proceso de IVT, por medio de la adición de UTP biotilado a la mezcla de reacción, a una concentración final de 0.106 mM. Con ello se obtuvo, de media, la inclusión de un nucleótido de UMP biotilado por cada molécula diana. La reacción de IVT se llevó a cabo a 37°C durante 2 horas, y el molde de DNA se eliminó empleando DNasa RQ1 (Promega) a 1U/μg DNA, por medio de una reacción a 37°C durante 30 minutos. El producto de transcripción fue purificado por electroforesis en gel de poliacrilamida.

La molécula UTR₃₀₈ biotilada se unió a una columna de Sefarose HiTrap™ Streptavidine HP (GE Healthcare). Para ello, todo el producto de la transcripción de la diana fue precipitado utilizando etanol y se resuspendió en 1 ml del tampón de unión recomendado por el fabricante de las columnas. La población de molécula diana fue renaturalizada mediante incubación a 64°C durante 10 minutos, seguida de otra incubación a 37°C durante 10 minutos. Finalmente, esta población fue introducida en una columna de sefarosa previamente equilibrada con el paso de 10 ml de tampón de unión. Tras esto, la columna se lavó con otros 10 ml de tampón de unión y equilibrada con 10 ml de TMN 1X (20 mM TRIS-acetato; 10 mM acetato de magnesio; 100 mM cloruro de sodio).

Para comenzar el proceso de selección *in vitro*, la población de RNA de 64 nt denominada "Ronda 0" se precipitó utilizando etanol y se resuspendió en 1 ml de tampón TMN 1X. El RNA se renaturalizó en este tampón mediante incubación a 64°C durante 10 minutos, seguida de otra incubación a 37°C durante 10 minutos. Con objeto de eliminar de la población aquellas moléculas de RNA que pudieran unirse inespecíficamente a la sefarosa presente en las columnas, se realizó una selección negativa acoplando una columna de sefarosa-estreptavidina en tandem por encima de aquella que contenía la molécula diana unida. Esta columna vacía también fue equilibrada con 10 ml de TMN 1X. El RNA renaturalizado correspondiente a la Ronda 0 se introdujo en el tandem de columnas y se pasó un segundo volumen de TMN 1X para que avanzase hasta la columna con la molécula diana inmovilizada. Tras un tiempo de incubación de 30 minutos a 25°C, la columna para selección negativa fue retirada y las moléculas no unidas a la columna con la diana se eliminaron mediante lavado a la misma temperatura con 10 ml de TMN 1X. Aquellas moléculas que se uniesen a la sefarosa o la estreptavidina y no a la diana, habrían quedado retenidas por la primera columna. Por su parte, las moléculas de RNA de la Ronda 0 unidas a la diana UTR₃₀₈ se eluyeron de la columna por desnaturalización, haciendo pasar a través de ella 10 ml de TMN 1X a 95°C. Las primeras 4 fracciones de 1 ml recuperadas se unieron y concentraron utilizando el sistema Centricon Ultracel YM-3 (Amicon Bioseparations, Millipore), y se precipitaron con etanol.

La mitad de la población de RNA unida se guardó a -80°C, y el resto se retrotranscribió a cDNA utilizando la actividad RT de la enzima Tth DNA polimerasa (Promega) y el cebador 3'XhoIK (SEQ ID No: 42, 5'-GTGAGCTCGTAGTATCAGATCACTCCAT-3'). Posteriormente, el cDNA obtenido se amplificó por PCR utilizando la misma enzima y empleando los cebadores 3'XhoIK (SEQ ID No: 42) y 5'EcoRIK (SEQ ID No: 40, incluyendo este último la secuencia del promotor de T7). El DNA amplificado se purificó con fenol, se precipitó con etanol y se utilizó como molde para una reacción de IVT con la que se obtuvo el RNA denominado "Ronda I".

Este proceso se repitió a lo largo de 14 rondas de amplificación-selección, aumentándose la presión selectiva a lo largo del proceso. Este incremento de la astringencia en la interacción aptámero-diana se logró por medio de un aumento de la temperatura de unión (que fue de 25°C para las rondas I-III y de 37°C para las IV-XIV). Por su

parte, la relación aptámero-diana fue de 1:1 para las rondas I-X y de 1000:1 para las XI-XIV. En cada ronda, la mitad del DNA amplificado se clonó en el vector pGEM-T easy (Promega), y se obtuvieron más de 30 clones moleculares que fueron secuenciados.

Ejemplo 2.- Análisis de la estructura y la secuencia de los aptámeros obtenidos por selección *in vitro*.

- 5 Se realizó un análisis de secuencia (mediante clonaje molecular y secuenciación) y estructura (por medio de programas de plegamiento *in silico*) de 299 moléculas individuales correspondientes a la Ronda 0 y las 14 rondas de selección, en la cantidad que se indica a continuación: Ronda 0, 30 seqs; Ronda I, 30; Ronda III, 31; Ronda V, 24; Ronda VIII, 28; Ronda IX, 32; Ronda X, 35; Ronda XI, 52; Ronda XIV, 37. De las 299 moléculas, 216 eran únicas, siendo las demás secuencias repetidas. Se comprobó que entre las 216 moléculas, 188 poseían la longitud esperada de 64 nt mientras que las demás presentaban deleciones de 1 o más nucleótidos en su región variable. El análisis de secuencias mediante programas de alineamiento y comparación mostró que en varios casos algunas de las secuencias estaban repetidas, tanto dentro de una misma ronda como en rondas diferentes. Las secuencias repetidas obtenidas se muestran en la Tabla 1.

15 **Tabla 1:** Secuencias repetidas presentes a lo largo del proceso de selección *in vitro*, energía libre asociada a la MFE en cada caso y grupo al que pertenecen

0 ^a	I	III	V	VIII	IX ^b	X	XI	XIV	ΔG de la MFE (Kcal/mol)	Grupo ^c
016-2									-10.70	
029-2									-8.60	
		III45B-2							-5.10	
					IX05-3	X05-12	XI73-3		-4.03	Gr. 3
					IX08-2				-11.10	
					IX36	X10-7	XI1-17	XIV22-23	-10.80	Gr. 2a
						X04-2	XI23-3		-20.20	Gr. 1
						X11-2			-6.61	
							XI21-7	XIV26-6	-15.70	Gr. 1
							XI141-2	XIV1-2	-11.80	Gr. 2b
						X41-2			-15.20	

- 20 **Notas:** a) se muestra en una misma línea las secuencias iguales presentes en distintas rondas; las secuencias repetidas dentro de cada ronda se indican con la notación "Seq-R", siendo R el número de repeticiones (por ejemplo, XIV22-23 significa que de la secuencia ejemplificada por el aptámero XIV22 hay 23 moléculas iguales en la ronda XIV); b) se resalta en negrita las secuencias que incluyen el motivo 5'-GGCARGGA-3'; c) los grupos se describen en el texto.

Se estimó la distancia entre las secuencias de cada una de las rondas, medida como número de mutaciones entre ellas (distancia de Hamming). A pesar de que la longitud total de las secuencias es de 64 nt, la distancia máxima es de 25 (región variable). Se observó que a partir de la ronda IX se comenzaron a definir grupos de secuencias claramente diferenciados, lo que indicaba que el proceso de selección estaba siendo eficiente.

- 25 Se analizaron las secuencias mediante alineamiento múltiple y programas de "clustering" basados en distancias. El resultado del clustering de secuencias se muestra en la Fig. 2, y permitió definir tres grupos de secuencias (el sistema de notación se indica en la Tabla 1):

i) Grupo 1: X04-2, X09, X13, XI23-3, XI107, XI21-7, XIV5, XIV26-6

ii) Grupo 2:

- 30 - Subgrupo 2a: IX36, X10-7, XI1-17, XI129, XI149, XIV12, XIV22-23, XIV32

- Subgrupo 2b: XI101, XI141-2, XIV1-2

iii) Grupo 3: IX05-3, X05-12, XI65, XI73-3

- 35 Con ello se observaba que los dos grupos de secuencias que llegan a la ronda XIV (Grupos 1 y 2) habían comenzado a ser seleccionados en la ronda IX. Por su parte, el Grupo 3 aparecía entre las rondas IX y XI, pero no estaba representado en la ronda XIV.

Todas las secuencias completas (64 nt) se plegaron en su estructura de mínima energía o "MFE structure" empleando el programa RNAfold del Vienna RNA package en versión 1.5, con los parámetros energéticos descritos en (Mathews *et al.*, 1999), sin permitir la presencia de pares de bases aislados, y definiendo como

longitud mínima para un tallo bicatenario o stem 2 pares de bases (bp). Esas estructuras plegadas, representadas en la notación dot-bracket (en la que los nucleótidos desapareados se indican por puntos, y los que forman parte de pares de bases mediante paréntesis) y alineadas, son susceptibles de ser comparadas entre sí y así calcular la distancia que existe entre ellas. Dicha distancia entre estructuras se midió empleando tres métodos diferentes:

- i) *Hamming*: Entre dos estructuras alineadas se compara posición a posición el símbolo presente [".", "(" o ")"], y cuando son diferentes se computa distancia 1 para esa posición. La distancia total entre las dos estructuras será la suma de las distancias para cada posición.
- ii) *Base pair*: La distancia entre dos estructuras es igual al número de pares de bases que hay que romper y formar de nuevo después para pasar de la una a la otra.
- iii) *Tree edit*: Se calcula utilizando la representación de las estructuras como árboles (donde un par se convierte en un nodo) y permitiendo cierta clase de operaciones, como inserciones, deleciones y sustituciones. Una distancia tree edit siempre es igual o mayor a una distancia base pair.

Para el conjunto de datos, los resultados obtenidos por los tres sistemas fueron equivalentes a nivel cualitativo. En la Fig. 3 se observa la progresión, a lo largo del proceso de selección *in vitro*, de la distancia entre secuencias (Hamming) y estructuras (base pair) dentro de cada ronda. Se detectó una tendencia globalmente decreciente en ambos casos, lo que indica la aparición de un sesgo en la composición que probablemente refleja la presencia de secuencias y estructuras más favorecidas para su unión a la diana UTR₃₀₈ (SEQ ID No: 1).

Se analizaron las distancias entre estructuras mediante los mismos programas de agrupamiento o clustering utilizados con las 188 secuencias. Los resultados de clasificación obtenidos (para distancias base pair entre estructuras) se muestran en la Fig. 4. La topología obtenida (equivalente a la que se genera empleando distancias de Hamming entre estructuras) mostró la presencia de los grupos ya identificados en la Fig. 2, que se compactan más respecto a los grupos o clusters de secuencias debido a la degeneración entre los espacios de secuencias y estructuras (dado que distintas secuencias pliegan en una misma estructura). Así, el Grupo 1 aparece formando un cluster compacto, con las estructuras iguales a la representada por XIV26-6 y las que quedan a distancias entre 1 y 6 de ella. En los Grupos 2a y 2b, la mayor parte de las estructuras son idénticas, si bien tres de ellas (XI101, XI129 y XIV12) mostraron plegamientos muy alejados a las demás de su grupo.

Al analizar conjuntamente las secuencias y sus plegamientos se observó que existe un motivo de 8 nucleótidos, con secuencia 5'-GGCAAGGA-3', que aparece en 25 secuencias diferentes de las rondas IX, X, XI y XIV (no en las anteriores), que son las marcadas con puntos azules en las Figs. 2 y 4:

- iv) Ronda IX: 03, 24, 36
- v) Ronda X: 02, 04-2, 09, 10-7, 13, 36, 41-2
- vi) Ronda XI: 1-17, 13, 21-7, 23-3, 101, 107, 108, 129, 141-2, 149
- vii) Ronda XIV: 1-2, 5, 12, 22-23, 26-6

Por tanto, ese motivo apareció en 25/216 secuencias diferentes y 86/299 secuencias totales. Entre ellas, la secuencia XI108, no fue considerada para el clustering por presentar una deleción de 1 nt en la región variable. Fue fundamental para la presente invención comprobar que en la ronda final del proceso (XIV), todas las secuencias excepto tres (XIV25, XIV37 y XIV48) presentan este motivo y pertenecen a los Grupos 1 ó 2 (Tabla 1 y Figs. 2 y 4). Por tanto, parecía fundamental la presencia de ese motivo (no descrito previamente en la literatura) para que el aptámero fuera seleccionado durante el proceso. Expresado de otra forma, dicho motivo resultaba imprescindible para la interacción aptámero-ligando. La secuencia XIV32 (Grupo 2a) poseía este mismo motivo pero con una mutación puntual (transición A→G en la posición 5 del motivo): 5'-GGCAGGGA-3'. Se comprobó que esa secuencia mutante del motivo aparecía exclusivamente en esta molécula, y no en las demás de esta u otras rondas.

Se comprobó que todas las secuencias (pertenecientes a las rondas IX, X, XI y XIV) que forman parte de los Grupos 1 y 2 poseen el motivo 5'-GGCARGGA-3' (notación IUB para nt degenerados: R corresponde bien a A ó bien a G). Las secuencias que incluyen ese motivo pero no forman parte de los Grupos 1 ó 2 son: IX03, IX24, X02, X36, X41-2, XI13. De estas seis secuencias, IX03, X36, X41-2 y XI13 formaban a su vez un cluster bien definido en la Fig. 4.

En las estructuras MFE en que plegaban las secuencias presentes en la ronda XIV (Fig. 5), cuando esta secuencia aparecía siempre estaba formando un lazo desapareado de 8 nt que cerraba un tallo de al menos

cuatro pares de bases [estructura en notación dot-bracket (((((((.....))))))]. Era destacable el hecho de que la mutación presente en XIV32 no altera la estructura de ese motivo. Los nt que forman los pares flanqueantes del tallo cerrado por ese lazo son variables. El alineamiento de las secuencias que contienen ese motivo (Fig. 6) mostró una secuencia consenso "NNDYGGCARGGARNNN" (D=A,G,U; Y=C,U; R=A,G; N=A,C,G,U).

5 El motivo conservado NNDYGGCARGGARNNN con estructura en notación dot-bracket (((((((.....)))))), presente en los aptámeros seleccionados, tenía gran interés debido a lo siguiente:

i) Su presencia era claramente creciente en las rondas IX (3 de 32 secuencias, 9.4%), X (15/35, 42.9%) XI (35/52, 67.3%) y XIV (34/37, 91.9%), y podía correlacionarse con el momento en que comenzó a aumentarse la presión selectiva en los experimentos.

10 ii) La secuencia del lazo desapareado presente en este motivo (8 nt comunes en una región variable de 25, es decir, casi un tercio de la secuencia) no podía haber sido seleccionada al azar. La probabilidad de que una secuencia de 25 nt escogida al azar contenga un motivo concreto de 8 nt es de 2 entre 10.000. Considerando que el motivo se halla en dos secuencias independientemente seleccionadas en la ronda XIV: si se preguntara con qué probabilidad se obtendría el motivo repetido escogiendo dos secuencias aleatorias al azar, la respuesta es una vez de cada 25 millones. Además, la exigencia de que esté formando un lazo al final de un tallo de 4 bp hace descender considerablemente esta probabilidad.

15 iii) En las secuencias en que aparece este motivo, su estructura común sugería que el lazo de 8 nt podría interactuar con alguna región desapareada de la molécula diana.

20 Además, a pesar de la gran diferencia entre las dos estructuras a las que se llega en la ronda XIV (Grupos 1 y 2), era destacable el hecho de que este dominio estaba presente en ambas. Por tanto, esto podría indicar que en el proceso de selección en primer lugar (ya en la ronda IX, ver Tabla 1 y Fig. 6) se seleccionaron moléculas que contenían ese motivo, y una vez fijado ese requerimiento se dio la posibilidad de que el motivo se rodeara por dos contextos estructurales diferentes: i) XIV26-6 y resto de secuencias del Grupo 1; ii) XIV22-23 y resto de secuencias del Grupo 2. Así, probablemente se estaba promocionando en primer lugar una interacción por complementariedad de secuencias entre el aptámero y la molécula diana UTR₃₀₈ (SEQ ID No: 1), y posteriormente se daría una interacción "estructural" que podía ser de dos tipos diferentes.

Ejemplo 3.- Proceso de obtención de una molécula sintética de RNA objeto de la invención a partir de los aptámeros obtenidos por selección *in vitro*.

30 En función de lo descrito en el Ejemplo 2, se decidió utilizar el análisis bioinformático realizado para diseñar una molécula sintética de RNA de 16 nt, que contuviera el motivo indicado en un lazo terminal definido por un tallo muy estable de 4 bp. Así, se diseñó y sintetizó la molécula llamada "RNA16(+)" con secuencia SEQ ID No: 35 (5'-CCCCGGCAAGGAGGGG-3'), en la que cuatro pares C-G consecutivos garantizaban la estructura en notación dot-bracket (((((((.....)))))) tal como se muestra en la Fig. 7. Si la hipótesis de trabajo era correcta, esta molécula RNA16(+) debería unirse con gran afinidad a la diana UTR₃₀₈ (SEQ ID No: 1), y eventualmente podría ser capaz de inhibir la replicación de VIH-1 con una eficiencia similar a la mostrada por los aptámeros de la Ronda XIV que contenían el motivo 5'-GGCAAGGA-3'. Un aspecto interesante y novedoso de la molécula RNA16(+), además de que tal secuencia nunca se ha descrito como inhibidora de VIH, es que se había obtenido mediante una aplicación combinada *in vitro-in silico*. De hecho, sin esta combinación entre el trabajo experimental y el análisis bioinformático detallado habría sido imposible la selección de la molécula RNA16(+), debido a las restricciones de tamaño impuestas por los procedimientos de selección *in vitro*.

Ejemplo 4.- Estudio bioinformático de la interacción aptámero-diana y molécula RNA16(+)-diana.

45 En primer lugar se realizó un estudio bioinformático para comprobar si tanto los aptámeros seleccionados en la Ronda XIV como la molécula RNA16(+) genérico (SEQ ID No: 39) poseían lugares de interacción mediados por complementariedad de bases entre lazos desapareados con la molécula diana. Para ello se utilizó el programa RNAup del Vienna RNA package en versión 1.5. Utilizando este programa, la interacción RNA-RNA se evaluó, en primer lugar, calculando la probabilidad de que un intervalo de secuencia (en este caso, el lazo de 8 nt desapareados) no se uniera a ningún lugar de la molécula diana (UTR₃₀₈, SEQ ID No: 1). A continuación, se calculó la energía de unión correspondiente a todos los casos posibles de interacción entre las dos moléculas, y se optimizó este parámetro para encontrar las interacciones más favorecidas, es decir, aquellas que generaban estructuras con un valor más negativo de energía libre. Se trabajó con los mismos parámetros escogidos para RNAfold. Adicionalmente, para explorar más detalladamente los posibles lugares de interacción de la diana, se utilizó los programas pKiss y pKnotsRG, proporcionados por el RNA Studio package de la Universidad de Bielefeld (bibiserv.techfak.uni-bielefeld.de), que permitieron evaluar, respectivamente, la presencia de motivos de tipo "kissing loop" y de pseudonudos de distintos tipos.

El resultado de este estudio bioinformático para los aptámeros XIV22-23 (SEQ ID No: 29), XIV26-6 (SEQ ID No: 30) y XIV1-2 (SEQ ID No: 31), así como con la molécula RNA16(+), se muestra en la Fig. 8. Los resultados que arroja este análisis son las posibles interacciones entre regiones de ambas moléculas (aptámero y diana UTR₃₀₈) y vienen representadas como la energía libre total de la unión entre sendas regiones de las moléculas participantes. Es, por tanto, un estudio recíproco del que sólo se muestra la parte correspondiente a la molécula UTR₃₀₈ (Figs. 8A y 8B). En las curvas de esta gráfica se observan diferentes valles trasladables a la secuencia de la diana marcada en el eje X. Aquellos marcados de la (a) a la (f) se corresponden con lugares de interacción que implican al motivo 5'-GGCAAGGA-3' en los aptámeros y la molécula RNA16(+) (SEQ ID No: 35), si bien las curvas para cada molécula no coinciden globalmente debido a que el contexto de secuencia en que se encuentra este motivo en cada aptámero es diferente. Todas estas regiones de interacción se pueden identificar en el plegamiento de tipo BMH de la molécula UTR₃₀₈ (SEQ ID No: 1) que se muestra en la Fig. 8C. A priori, la región (a) de interacción (situada en el dominio Poli-A) sería la más probable y se propone como principal responsable de la unión con el motivo 5'-GGCAAGGA-3' del aptámero de la invención basándose en tres criterios: i) el mismo valle (Figs. 8A y 8B) coincide en la interacción de la UTR₃₀₈ (SEQ ID No: 1) con las cuatro moléculas [aptámeros XIV22-23 (SEQ ID No: 29), XIV26-6 (SEQ ID No: 30), XIV1-2 (SEQ ID No: 31) y RNA16(+) (SEQ ID No: 35)]; ii) presenta una mayor estabilidad en la interacción dado que tiene una menor energía libre total de unión; y iii) en dicha interacción intervienen 6 nucleótidos contiguos en ambas moléculas.

Adicionalmente, tanto los experimentos de unión como de inhibición que implicaran a la molécula RNA16(+) (SEQ ID No: 35) requerían el análisis comparativo de una molécula útil a modo de control negativo. Como tal control negativo se necesitaba una molécula que, teniendo también 16 nt de longitud y la misma estructura en notación dot-bracket (((((((.....)))))) fuese la menos favorecida para su unión a la molécula diana UTR₃₀₈ (SEQ ID No: 1). Para ello se plegaron extensivamente *in silico* todas las moléculas de tipo 5'-CCCCNNNNNNNNGGGG-3' utilizando el programa RNAfold, y se seleccionaron aquellas que generaran la estructura en notación dot-bracket (((((((.....)))))) como la más estable. Así, de las 65.536 secuencias (4⁸) iniciales, se comprobó que 19.087 plegaban preferentemente en la estructura en notación dot-bracket (((((((.....)))))). A continuación se evaluó, mediante comparación sistemática de secuencias, cuales de ellas tenían un menor número de nucleótidos complementarios en la molécula diana UTR₃₀₈ (SEQ ID No: 1), tanto en sus regiones de cadena sencilla como en las de doble cadena, en cualquiera de las dos orientaciones y permitiendo pares G-C, A-U y G-U. Con ello se obtuvo un grupo de 112 secuencias con la restricción de que formaran un máximo de 5 posibles pares de bases entre la molécula de 16 nt y la diana, siendo consecutivos un máximo de 4 de esos pares de bases. Utilizando el programa RNAup se calculó la energía de interacción para todas estas moléculas de 16 nt, teniendo en cuenta la estructura de la diana. Los resultados mostraron energías libres de interacción entre -0.40 y -7.20 kcal/mol. Con todo ello, se escogió como control negativo, denominado "RNA16(-)", la molécula con secuencia SEQ ID No: 45 (5'-CCCCGAAAACAAGGGG-3'). Sus propiedades son las siguientes: i) posee la menor (en valor absoluto) energía libre de unión con la diana (-0.40 kcal/mol) de todas las ensayadas; ii) la interacción con la diana se establece únicamente en una región; iii) tal interacción implica únicamente 2 pares de bases: posiciones 22 y 23 de la diana UTR₃₀₈ (SEQ ID No: 1) y posiciones 5-6 de la molécula RNA16(-) (SEQ ID No: 45).

Ejemplo 5.- Ensayos de unión y ensayos funcionales de inhibición de la replicación de VIH-1 por el aptámero de la invención.

La inhibición de la replicación del VIH-1 en cultivos celulares por los diferentes aptámeros de la invención resultó ser productiva mediante la utilización de moléculas quiméricas consistentes en el aptámero correspondiente (XIV22-23 -SEQ ID No: 29- o XIV26-6 -SEQ ID No: 30-) flanqueado por las horquillas 5' y 3' del RNA pequeño nuclear U6 (U6 snRNA). Está descrito que estas horquillas confieren mayor estabilidad al RNA y le proporcionan una localización nuclear (Good *et al.*, 1997). Los ensayos de inhibición del VIH-1 así como los ensayos de unión aptámero-UTR₃₀₈ *in vitro* se realizaron con estas moléculas quiméricas. Un esquema de la estructura de dichas moléculas se muestra en la Fig. 9. El control negativo para esta inhibición consistió en la molécula de RNA formada por estas dos horquillas sin aptámero localizado entre ambas. A estos nuevos RNAs se les llamó L-XIV22-23 (SEQ ID No: 37) y L-XIV26-6 (SEQ ID No: 38), y al control negativo se le denominó L-vacío. Las moléculas de RNA16(+) (SEQ ID No: 35) y RNA16(-) (SEQ ID No: 45) fueron utilizadas para ambos ensayos sin modificación ni unión covalente a las horquillas.

Los DNAs clonados en pGEM-T easy correspondientes a los aptámeros XIV22-23 y XIV26-6 fueron utilizados como molde para PCR con los cebadores 5'KpnIC3 (SEQ ID No: 46, 5'-CGACTCGGTACCGGGAATTCAA-3') y 3'ApalC3 (SEQ ID No: 47, 5'-TCTGGGCCCGTCGAGCTCGTAGTATC-3'). Los productos de amplificación consistían en DNAs que contenían las secuencias de los aptámeros de RNA seleccionados, flanqueadas por las dianas de restricción *KpnI* y *Apal* en sus extremos 5' y 3' respectivamente. El vector de clonación pU6 (Sánchez-Luque, F.J. *et al.* 2010) y los productos de amplificación fueron digeridos con las enzimas *KpnI* y *Apal*. Ambos productos de amplificación fueron clonados en el vector digerido mediante ligación con T4 DNA ligasa (Roche) y se transformaron bacterias competentes de la cepa DH5 α de *Escherichia coli* con la mezcla de ligación para su selección y conservación.

Para los ensayos de unión, todas las moléculas fueron transcritas *in vitro* tal como se detalla en el apartado de Proceso de obtención de los aptámeros, salvo las moléculas RNA16(+) (SEQ ID No: 35) y RNA16(-) (SEQ ID No: 45), que fueron sintetizadas químicamente (IBA GmbH). El molde para la transcripción de los aptámeros fue obtenido mediante amplificación por PCR de las construcciones plasmídicas que contenían las conjugaciones aptámero-horquillas del snRNA U6 mediante los oligonucleótidos 5'T7U6 (SEQ ID No: 48: 5'-TAATACGACTCACTATAGGGGTCGTCTTCTGCACGACAT-3') y 3'ApalU6 (SEQ ID No: 49: 5'-AGCGGGCCAAAAAGCGGACCGAAGTCCGC-3'). En este caso, además, la molécula UTR₃₀₈ (SEQ ID No: 1) no fue biotinilada. Las moléculas de aptámero fueron defosforiladas utilizando fosfatasa alcalina intestinal bovina (CIAP, Roche) previo a su marcaje en 5' por fosforilación mediante T4 polinucleótido quinasa (T4PNK, New England Biolabs) y usando como fuente de carbono ³²PγATP (Perkin Elmer). Las moléculas RNA16(+) (SEQ ID No: 35) y RNA16(-) (SEQ ID No: 45) fueron directamente marcadas en 5' al encontrarse directamente defosforiladas tras la síntesis química. Todas las moléculas fueron purificadas previamente a su uso mediante electroforesis en gel de acrilamida seguida de extracción fenólica y precipitación con etanol.

Trazas de cada aptámero fueron sometidas a reacciones de unión a cantidades crecientes de molécula UTR₃₀₈ (SEQ ID No: 1). Las moléculas fueron desnaturalizadas y renaturalizadas por separado mediante incubación en TMN 1X a 65°C durante 10 minutos seguida de incubación a 37°C también durante 10 minutos. Posteriormente, aptámeros y UTR₃₀₈ (SEQ ID No: 1) fueron mezclados de modo que la concentración final de UTR₃₀₈ en cada reacción fue de 0, 2, 20, 200 y 400 nM, respectivamente. Las reacciones de unión se realizaron a 37°C durante 30 minutos y fueron detenidas por adición de un volumen de tampón de carga nativo 2X (30% glicerol (v/v), 20% TMN 5X (v/v), 0'4% xylene-cyanol (p/v), 0'4% bromofenol (p/v) y 4% tRNA (p/v)) e incubación a 4°C. Las reacciones fueron resueltas en electroforesis en geles nativos al 6% de poliacrilamida en TRIS-acético 50 mM y acetato magnésico 10 mM como tampón de gel y como tampón de electroforesis. Las electroforesis fueron realizadas a 4°C y a no más de 15 mA por gel. Los geles fueron secados por combinación de vacío y calor e incubados 10-16h con pantallas excitables por radiación (Storage phosphor screens, GE Healthcare). Estas pantallas fueron reveladas en un escáner Typhoon 9400 (GE Healthcare).

Los ensayos de inhibición del VIH-1 fueron realizados en cultivos celulares de la línea de riñón embrionario humano 293T (HEK293T). Los ensayos de inhibición se realizaron simulando un ambiente post-integrativo de la infección viral, esto es, en el momento en que el genoma viral se ha integrado en el genoma celular como DNA proviral. Para esto, las células fueron co-transfectadas con un DNA plasmídico que contenía una copia integrada de la cepa viral X4 NL4-3 y los correspondientes RNAs inhibidores. Los RNAs fueron transcritos *in vitro* y directamente purificados por extracción fenólica, y el DNA fue purificado de cultivo bacteriano mediante Plasmid Mini Kt (Qiagen). Se utilizaron placas de cultivo de 24 pocillos, y 24 horas antes de la transfección se sembraron 250.000 células en 250 µl de medio DMEM con 10% de suero fetal bovino (FBS) y 2 mM de L-glutamina. Para la transfección se diluyó 1 µl de Lipofectamine 2000 reagent (Invitrogen) en 25 µl de Opti-MEM (PAA Laboratories GmbH) y se incubó 5 minutos a temperatura ambiente. Paralelamente se diluyeron 500 ng de cada RNA inhibidor y 100 ng de pNL4-3 en otros 25 µl de Opti-MEM, se mezclaron con los anteriores y se incubaron a temperatura ambiente durante 20 minutos. A continuación se añadieron 450 µl de DMEM con 10% de FBS y 2 mM de L-glutamina y se sustituyó el medio del pocillo de cultivo por esta mezcla. La actividad viral en estos cultivos fue medida 48 horas después mediante la cuantificación de la cantidad de antígeno viral p24 en el sobrenadante celular por inmunoensayo enzimático (Greenscreen HIV-1 Ag assay, BioRad).

Ejemplo 5.1: Unión del aptámero L-XIV22-23 (SEQ ID No: 37) a la molécula diana UTR₃₀₈ (SEQ ID No: 1).

El ensayo de unión del aptámero L-XIV22-23 (SEQ ID No: 37) y el resultado de su cuantificación se muestran en la Fig. 10. La curva de unión del aptámero a la molécula diana UTR₃₀₈ se ajusta a una modelo de unión cooperativa con un único sitio de unión. Esta curva sigue la ecuación siguiente:

$$Y = B_{\max} \cdot X^h / (K_D^h + X^h),$$

donde Y es el % de unión registrado; X es la concentración de UTR₃₀₈; B_{máx} es la unión máxima predicha según la ecuación de ajuste; K_D es la constante de disociación, que se define como la concentración de UTR₃₀₈ a la que se produce la mitad de la unión máxima (valor de X para Y = B_{máx}/2); y h es el coeficiente de Hill que indica el grado de cooperatividad de modo que h < 1 indica cooperatividad negativa y h > 1 cooperatividad positiva.

Para la molécula L-XIV22-23 (SEQ ID No: 37) la unión máxima esperada es B_{máx} = 87.59±1.26%; la constante de disociación es K_D = 153.5±4.8 nM; y el coeficiente de Hill es h = 2.60±0.86. El coeficiente de ajuste a esta gráfica es R² = 0.9987.

Ejemplo 5.2: Unión del aptámero L-XIV26-6 (SEQ ID No: 38) a la molécula diana UTR₃₀₈ (SEQ ID No: 1).

El ensayo de unión del aptámero L-XIV26-6 (SEQ ID No: 38) y el resultado de su cuantificación se muestran en la Fig. 11. La curva de unión del aptámero a la molécula diana UTR₃₀₈ (SEQ ID No: 1) se ajusta a un modelo hiperbólico de unión. Esta curva sigue la ecuación siguiente:

$$5 \quad Y = B_{\max} \cdot X / (K_D + X) ,$$

donde Y es el % de unión registrado; X es la concentración de UTR₃₀₈; B_{máx} es la unión máxima predicha según la ecuación de ajuste; K_D es la constante de disociación, que se define como la concentración de UTR₃₀₈ a la que se produce la mitad de la unión máxima (X para Y = B_{máx}/2).

10 Para la molécula L-XIV26-6 (SEQ ID No: 38) la unión máxima esperada es B_{máx} ≈ 100% y la constante de disociación es K_D = 81.67±12.81 nM. El coeficiente de ajuste a esta gráfica es R² = 0.9912.

Ejemplo 5.3: Unión de la molécula RNA16(+) (SEQ ID No: 35) a la molécula diana UTR₃₀₈ (SEQ ID No: 1).

El ensayo de unión de la molécula RNA16(+) (SEQ ID No: 35) y el resultado de su cuantificación se muestran en la Fig. 12. La curva de unión del aptámero a la molécula diana UTR₃₀₈ (SEQ ID No: 1) se ajusta a un modelo hiperbólico de unión, con la misma ecuación y parámetros mostrados en el Ejemplo 5.2.

15 Para la molécula RNA16(+) (SEQ ID No: 35) la unión máxima esperada es B_{máx} ≈ 79.35±8.20% y la constante de disociación es K_D = 279.5±60.0 nM. El coeficiente de ajuste a esta gráfica es R² = 0.9908.

Ejemplo 5.4: Ensayo de inhibición de la replicación de VIH-1 producida por el aptámero L-XIV22-23 (SEQ ID No: 37).

20 La molécula L-XIV22-23 (SEQ ID No: 37) consigue una inhibición post-integrativa del VIH-1 en cultivo celular de la línea HEK293T del 76.77±7.48% comparada con la molécula control L-vacío. En la Fig. 13 se muestra la cuantificación de tres experimentos independientes. El análisis por test estadístico ANOVA muestra que la inhibición es estadísticamente significativa con un valor p < 0.01.

Ejemplo 5.5: Ensayo de inhibición de la replicación de VIH-1 producida por el aptámero L-XIV26-6 (SEQ ID No: 38).

25 La molécula L-XIV26-6 (SEQ ID No: 38) consigue una inhibición post-integrativa del VIH-1 en cultivo celular de la línea HEK293T del 80.47±2.55% comparada con la molécula control L-vacío. En la Fig. 14 se muestra la cuantificación de tres experimentos independientes. El análisis por test estadístico ANOVA muestra que la inhibición es estadísticamente significativa con un valor p < 0.01.

Ejemplo 5.6: Inhibición de la replicación de VIH-1 producida por la molécula RNA16(+) (SEQ ID No: 35).

30 La molécula RNA16(+) (SEQ ID No: 35) consigue una inhibición post-integrativa del VIH-1 en cultivo celular de la línea HEK293T del 63.34±2.14% comparada con la molécula control RNA16(-) (SEQ ID No: 45). En la Fig. 15 se muestra la cuantificación de tres experimentos independientes. El análisis por test estadístico ANOVA muestra que la inhibición es estadísticamente significativa con un valor p < 0.01.

REFERENCIAS

35 Briones C, Moreno M (2012). Applications of peptide nucleic acids (PNAs) and locked nucleic acids (LNAs) in biosensor development. *Anal Bioanal Chem* 402: 3071-3089.

Duongé F, Toulmé JJ (1999). *In vitro* selection identifies key determinants for loop-loop interactions: RNA aptamers selective for the TAR RNA element of HIV-1. *RNA* 5: 1605-1614.

40 Ellington AD, Szostak JW (1990). *In vitro* selection of RNA molecules that bind specific ligands. *Nature* 346: 818-822.

Geyer CR, Battersby TR, Benner SA (2003). Nucleobase pairing in Watson-Crick-like genetic expanded information systems. *Structure* 11: 1485-1498.

Good PD, Krikos AJ, Li SX, *et al.* (1997). Expression of small, therapeutic RNAs in human cell nuclei. *Gene Ther* 4: 45-54.

Horvath SJ, Firca JR, Hunkapiller T, Hunkapiller MW, Hood L (1987). An automated DNA synthesizer employing deoxynucleoside 3'-phosphoramidites. *Methods Enzymol.* 1987;154:314-26.

5 Huthoff H, Berkhout B (2001). Two alternating structures of the HIV-1 leader RNA. *RNA* 7: 143-157.

Joyce GF (2004). Directed evolution of nucleic acid enzymes. *Annu Rev Biochem* 73: 791-836.

Kikuchi K, Umehara T, Fukuda K, Kuno A, Hasegawa T, Nishikawa S (2005). A hepatitis C virus (HCV) internal ribosome entry site (IRES) domain III-IV-targeted aptamer inhibits translation by binding to an apical loop of domain IIIId. *Nucleic Acids Res* 33: 683-692.

10 Kolb G, Reigadas S, Castanotto D, *et al.* (2006). Endogenous expression of an anti-TAR aptamer reduces HIV-1 replication. *RNA Biol* 3: 150-156.

Klussmann S. (ed) (2006). *The Aptamer Handbook*. Wiley-VCH, Weinheim, Germany.

Kuiken C, Brian F, Leitner T, *et al.* (2011). *HIV Sequence Compendium 2011*. Los Alamos National Laboratory, Theoretical Biology and Biophysics, Los Alamos (NM), United States.

15 Lauridsen LH, Rothnagel JA, Veedu, RN (2012). Enzymatic Recognition of 2'-Modified Ribonucleoside 5'-Triphosphates: Towards the Evolution of Versatile Aptamers. *Chem Bio Chem* 13: 19-25.

Mathews DH, Sabina J, Zuker M, Turner H (1999). Expanded Sequence Dependence of Thermodynamic Parameters Provides Robust Prediction of RNA Secondary Structure. *J Mol Biol* 288: 911-940.

20 Pinheiro VB, Taylor AI, Cozens C, Abramov M, Renders M, Zhang S, Chaput JC, Wengel J, Peak-Chew SY, McLaughlin SH, Herdewijn P, Holliger P (2012). Synthetic genetic polymers capable of heredity and evolution. *Science* 336: 341-344.

Romero-López C, Barroso-delJesus A, Puerta-Fernández E, Berzal-Herranz A (2005). Interfering with hepatitis C virus IRES activity using RNA molecules identified by a novel *in vitro* selection method. *Biol Chem* 386: 183-190.

25 Sánchez-Luque FJ, Reyes-Darias JA, Puerta-Fernández E, Berzal-Herranz A (2010). Inhibition of HIV-1 replication and dimerization interference by dual inhibitory RNAs. *Molecules* 15: 4757-4772.

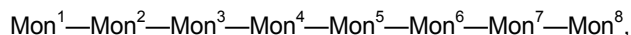
Tuerk C, Gold L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* 249: 505-510.

Watrin M, Von Pelchrzim F, Dausse E, Schroeder R, Toulmé JJ (2009). *In vitro* selection of RNA aptamers derived from a genomic human library against the TAR RNA element of HIV-1. *Biochemistry* 48: 6278-6284.

30

REIVINDICACIONES

1. Aptámero caracterizado porque comprende una estructura de secuencia



5 responsable de la unión específica a zonas accesibles en el plegamiento completo de la región 5'-UTR del genoma de virus de la inmunodeficiencia humana VIH tipo 1 (VIH-1), donde Mon¹ a Mon⁸ son monómeros de ácido nucleico (nucleótidos), siendo Mon¹ un nucleótido de guanina (G), Mon² un nucleótido de guanina (G), Mon³ un nucleótido de citosina (C), Mon⁴ un nucleótido de adenina (A), Mon⁵ un nucleótido de adenina (A) o de guanina (G), Mon⁶ un nucleótido de guanina (G), Mon⁷ un nucleótido de guanina (G) y Mon⁸ un nucleótido de adenina (A), de manera que dicha estructura consiste en un octanucleótido de secuencia 5'-GGCARGGA-3', en donde R es un nucleótido de adenina (A) o un nucleótido de guanina (G); o el enantiómero de dicho aptámero.

15 2. Aptámero según la reivindicación 1 donde la estructura además comprende una región flanqueante previa unida al nucleótido Mon¹ y una región flanqueante posterior unida al nucleótido Mon⁸ de dicha estructura, donde cada región flanqueante comprende al menos una secuencia de 4 nucleótidos, donde las bases nitrogenadas de al menos 3 de los 4 nucleótidos de la región flanqueante previa más próximos a Mon¹, contiguos o no, se aparean con las bases nitrogenadas de al menos 3 de los 4 primeros nucleótidos de la región flanqueante posterior más próximos a Mon⁸, de tal manera que la estructura del octanucleótido Mon¹—Mon²—Mon³—Mon⁴—Mon⁵—Mon⁶—Mon⁷—Mon⁸ con secuencia 5'-GGCARGGA-3' forma un lazo terminal monocatenario que cierra un tallo de cadena doble formado por las regiones flanqueantes con dichos nucleótidos apareados; o el enantiómero de dicho aptámero.

20 3. Aptámero según una cualquiera de las reivindicaciones 1 ó 2, que comprende una estructura de secuencia Mon^{a4}—Mon^{a3}—Mon^{a2}—Mon^{a1}—Mon¹—Mon²—Mon³—Mon⁴—Mon⁵—Mon⁶—Mon⁷—Mon⁸—Mon^{b1}—Mon^{b2}—Mon^{b3}—Mon^{b4}, donde Mon^{a1} a Mon^{a4} y Mon^{b1} a Mon^{b4} son nucleótidos de tal manera que:

- i. las bases nitrogenadas de Mon^{a1} a Mon^{a4} se seleccionan independientemente del siguiente grupo: A, C, G, T y U;
- 25 ii. la base nitrogenada de Mon^{b1} es complementaria a la base nitrogenada de Mon^{a1};
- iii. la base nitrogenada de Mon^{b2} es complementaria a la base nitrogenada de Mon^{a2};
- iv. la base nitrogenada de Mon^{b3} es complementaria a la base nitrogenada de Mon^{a3};
- v. la base nitrogenada de Mon^{b4} es complementaria a la base nitrogenada de Mon^{a4};

30 donde Mon¹ a Mon⁸ son los monómeros definidos en la reivindicación 1, y donde las reglas de complementariedad a las que están sujetos estos apareamientos respectivos de Mon^{a1} a Mon^{a4} con Mon^{b1} a Mon^{b4} son: A es complementaria a T o U; C es complementaria a G; G es complementaria a C o U; T es complementaria a A; U es complementaria a A o G; o el enantiómero de dicho aptámero.

4. Aptámero según una cualquiera de las reivindicaciones 1 a 3 que consiste en una molécula de RNA o de DNA; o el enantiómero de dicho aptámero.

35 5. Aptámero según la reivindicación 3 o según las reivindicaciones 3 y 4, donde la estructura Mon^{a4}—Mon^{a3}—Mon^{a2}—Mon^{a1}—Mon¹—Mon²—Mon³—Mon⁴—Mon⁵—Mon⁶—Mon⁷—Mon⁸—Mon^{b1}—Mon^{b2}—Mon^{b3}—Mon^{b4} consiste en un polinucleótido de secuencia seleccionada entre SEQ ID No: 6 y SEQ ID No: 7; o el enantiómero de dicho aptámero.

40 6. Aptámero según una cualquiera de las reivindicaciones 1 a 4, que comprende una secuencia nucleotídica seleccionada del grupo compuesto por: SEQ ID No: 10, SEQ ID No: 11, SEQ ID No: 12, SEQ ID No: 13, SEQ ID No: 14, SEQ ID No: 15, SEQ ID No: 16, SEQ ID No: 17, SEQ ID No: 18, SEQ ID No: 19, SEQ ID No: 20, SEQ ID No: 21, SEQ ID No: 22, SEQ ID No: 23, SEQ ID No: 24, SEQ ID No: 25, SEQ ID No: 26, SEQ ID No: 27, SEQ ID No: 28, SEQ ID No: 29, SEQ ID No: 30, SEQ ID No: 31, SEQ ID No: 32, SEQ ID No: 33, SEQ ID No: 34, SEQ ID No: 35, SEQ ID No: 36, SEQ ID No: 37 y SEQ ID No: 38; o el enantiómero de dicho aptámero.

45 7. Aptámero según una cualquiera de las reivindicaciones 1 a 5, que comprende una secuencia nucleotídica seleccionada del grupo compuesto por: SEQ ID No: 35 y SEQ ID No: 36; o el enantiómero de dicho aptámero.

8. Aptámero según una cualquiera de las reivindicaciones 1 a 7, donde el nucleótido del extremo 5' presenta una modificación para aumentar la resistencia a la degradación por exonucleasas mediante bloqueo del grupo fosfato del extremo 5'; o el enantiómero de dicho aptámero.

9. Aptámero según una cualquiera de las reivindicaciones 1 a 7, donde el nucleótido del extremo 3' presenta una modificación para aumentar la resistencia a la degradación por exonucleasas mediante bloqueo del grupo OH del extremo 3'; o el enantiómero de dicho aptámero.
- 5 10. Construcción genética de DNA útil para la síntesis por transcripción, *in vitro* o intracelular, de un aptámero de RNA definido en una cualquiera de las reivindicaciones 1 a 7, caracterizado porque comprende una secuencia de nucleótidos de DNA correspondiente a la secuencia de nucleótidos de dicho aptámero.
- 10 11. Procedimiento de obtención de un aptámero definido en una cualquiera de las reivindicaciones 1 a 9, caracterizado porque comprende la combinación de técnicas experimentales de selección *in vitro* de ácidos nucleicos con técnicas computacionales de optimización que permiten analizar tanto las secuencias generadas por selección *in vitro* como sus correspondientes estructuras secundarias, comprendiendo:
- a. obtener por transcripción *in vitro* moléculas de RNA de secuencia aleatoria, de al menos 25 nucleótidos de longitud, flanqueada por regiones anterior y posterior de secuencia constante y conocida;
 - b. seleccionar y aislar las moléculas de RNA obtenidas en la etapa anterior con mayor afinidad por la molécula diana correspondiente a la región 5'UTR del VIH-1;
 - 15 c. obtener y amplificar un cDNA obtenido a partir de las moléculas de RNA aisladas en la etapa anterior;
 - d. secuenciar el cDNA amplificado;
 - e. repetir sucesivos ciclos definidos por las etapas a-d utilizando el producto amplificado en la etapa c como molde para la transcripción *in vitro* de la etapa a, para obtener en la población resultante un enriquecimiento progresivo de moléculas con mayor capacidad de unión a la diana;
 - 20 f. clasificar las secuencias de cDNA secuenciadas en los distintos ciclos a-d repetidos, comprendiendo cuantificación de la diferencia nucleotídica existente entre dichas secuencias, alineamiento múltiple y clústering o agrupamiento de secuencias de dichos cDNA, plegamiento de la secuencia de RNA derivada del cDNA a la estructura de mínima energía, caracterización del motivo estructural mínimo entre secuencias y análisis de interacciones entre secuencias de DNA y la diana de RNA 5'UTR del VIH-1.
- 25 12. Uso de un aptámero o enantiómero del mismo definido en una cualquiera de las reivindicaciones 1 a 9, como agente de unión específica a la región 5'UTR del genoma de VIH-1.
13. Uso de un aptámero o enantiómero del mismo definido en una cualquiera de las reivindicaciones 1 a 9, como agente de interferencia de una función y/o actividad de la región 5'UTR del genoma de VIH-1.
- 30 14. Uso de un aptámero o enantiómero del mismo definido en una cualquiera de las reivindicaciones 1 a 9, como inhibidor de la replicación de VIH-1 en un cultivo celular.
15. Uso de un aptámero o enantiómero del mismo definido en una cualquiera de las reivindicaciones 1 a 9, para detectar, caracterizar y/o cuantificar VIH-1 *in vitro*.
16. Uso de un aptámero o enantiómero del mismo definido en una cualquiera de las reivindicaciones 1 a 9, como biosensor para la detección de VIH-1 en una muestra biológica.
- 35 17. Uso de un aptámero o enantiómero del mismo definido en una cualquiera de las reivindicaciones 1 a 9, para la fabricación de una composición farmacéutica.
18. Uso de un aptámero o enantiómero del mismo definido en una cualquiera de las reivindicaciones 1 a 9, para la fabricación de una composición farmacéutica para el tratamiento y/o prevención de una enfermedad.
- 40 19. Uso de un aptámero o enantiómero del mismo según la reivindicación 18, donde la enfermedad es una enfermedad causada por VIH-1.
20. Uso de un aptámero o enantiómero del mismo definido en una cualquiera de las reivindicaciones 1 a 9, para la fabricación de una composición farmacéutica para inhibir la replicación de VIH-1 en un paciente infectado por el mismo.
- 45 21. Composición farmacéutica caracterizada porque comprende al menos un aptámero o enantiómero del mismo definido en una cualquiera de las reivindicaciones 1 a 9, y al menos un adyuvante farmacéuticamente aceptable y/o un vehículo farmacéuticamente aceptable.

Figura 1

Señal Poli-A

CACTGCTTAAGCC.TCAATAAAG.	CTTGCC	TGAGT.GCTTCA,AGTAGTG	B.FR.83.HX82
		-A-T-	A1.IT.02.60000
			A1.RU.08.PokA1Ru
			A1.RW.93.93RW037A
		-T-	B.CN.05.05CNH8 hp3
			B.CU.99.Cu19
		T	B.DE.86.D31
			B.GA.88.OYI
			B.JP.05.DR6538
		-T-	B.KR.04.04KMH5
		T	B.NL.00.671.00T36
			B.RU.04.04RU128005
			B.TW.94.TWCY5
		-AGT-	B.US.09.N.2
		CTG--C	C.IN.03.D24
		CT-	C.ZM.02.02ZM168
			D.CD.83.ELI
	G	-A-C--C-	D.KE.97.ML415.2
			D.KR.04.04K8H8
			D.SN.90.SE365
			F1.BR.07.07BR844
			F1.R0.96.BCI.R07
GAG-C			G.BE.96.DRCBL
AG-C		-AGT-	G.CM.04.178.15
			G.CN.06.sh52
			G.CU.99.Cu74
			G.GH.03.03GH175G
			G.PT.x.PT2695
		-C-T-	H.GB.00.00GBAC4001
			J.CM.04.04CMU11421
AG-C		-A--G-	01.AE.CF.90.90CF11697
AG-C		-A--G-	01.AE.JP.x.DR0492
AG-C		-A--G-	01.AE.TH.04.BKM
			02.AG.GH.03.03GH181AG
			02.AG.LR.x.POC44951
		AT-	06.cpx.AU.96.BFP90
		CTGAGC-	08.BC.CN.06.nx2
			12.BF.AR.99.ARMA159
		-G-	26.AU.CD.02.02CD.MBTB047
			27.cpx.FR.04.04CD.FR.KZS
		-A-	32.06A1.EE.01.EE0369
	C		39.BF.BR.04.04BRRJ179
	C		40.BF.BR.05.05BRRJ055
			42.BF.LU.03.LuBF.05.03
			43.02G.SA.03.J11223
G--CG-		AGA--C	O.BE.87.ANT70
G--CG-		AGA--C	O.CM.91.MVP5180
		AGA--C	O.CM.96.96CMA102
		AGA--C	O.CM.98.98CMA104
		AGA--C	O.CM.99.99CMU4122
G--CG-		AGA--C	O.FR.92.VAU
G--CG-		AGA--C	O.SN.99.99SE.MP1299
		-A--AA--G-	N.CM.02.D300131
		-A--AA--G-	N.CM.04.04CM.1015.04
		-A--AA--AGT.G-	N.CM.06.U14842
		-A--AA--G-	N.CM.95.YBF30
		-A--AA--G-	N.CM.97.YBF106
		-AG-A-AGC-	P.CM.06.U14788
CGAAGC		-AG-A--C-	P.FR.09.RBF168
		-C-GCA-	CPZ.CM.05.SIVcpzMT145
		-TA-TTG-C-	CPZ.GA.88.GAB1
G--CG-	C	GAGAGT-T-	CPZ.TZ.00.TAN1
		-T-	CPZ.US.85.US.Marilyn

VH-1

SIV de Chimpancé

Figura 2

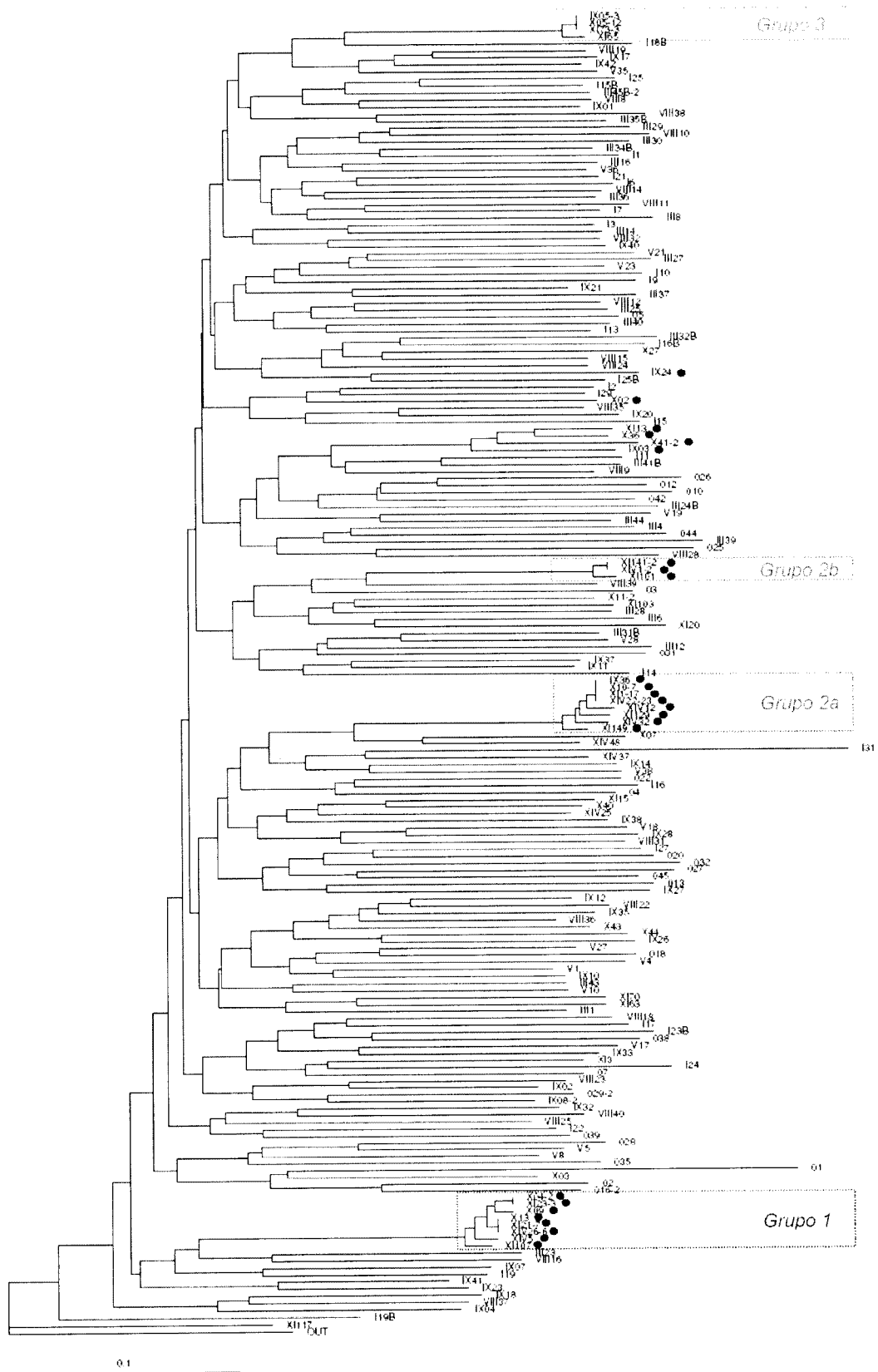


Figura 3

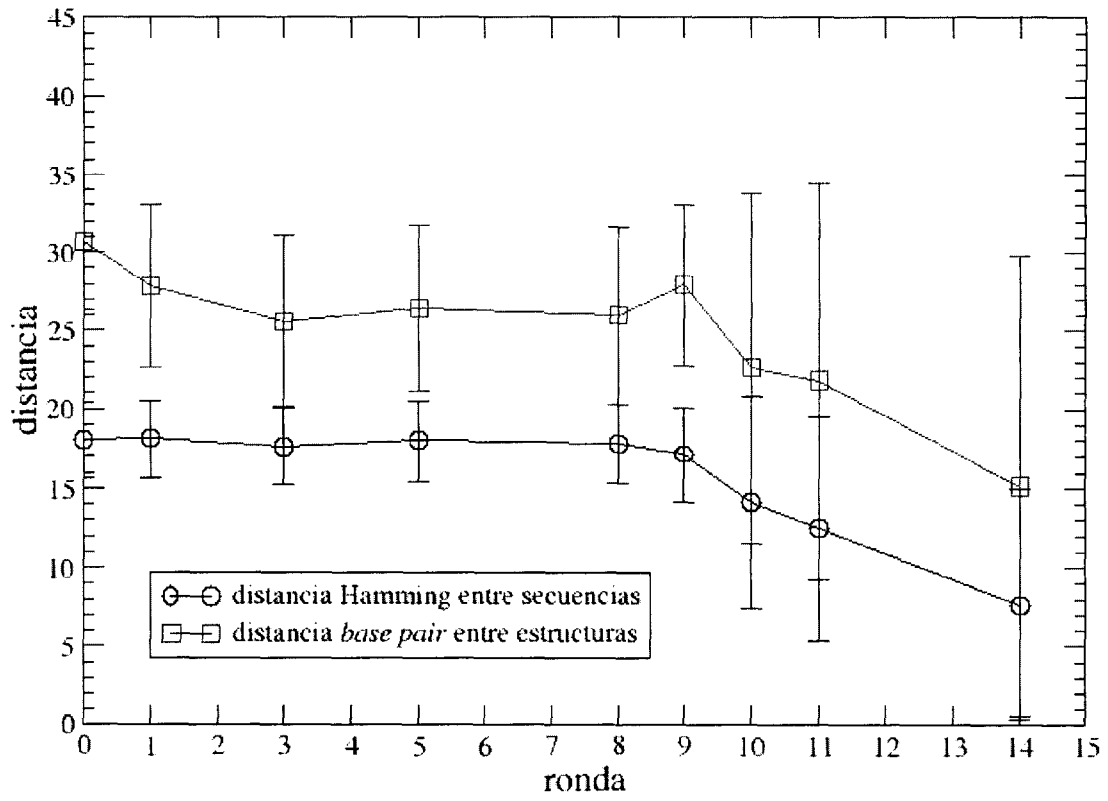


Figura 5

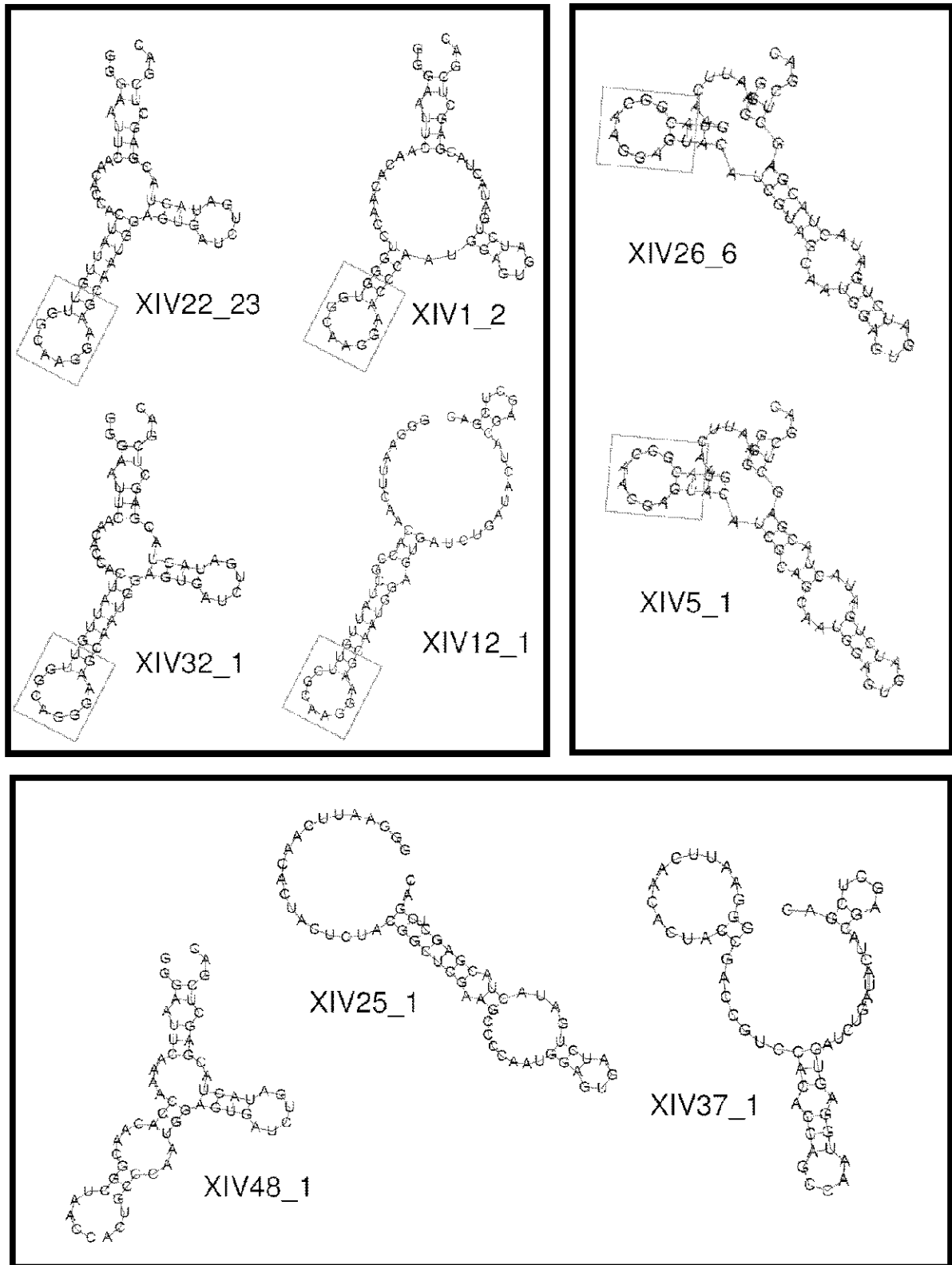


Figura 6

IX03-1 GGGAAUUCAAGACACGAACAUAGUGGCAAGGAACUAUGGAGUGAUCUGAUACUACGAGCUCGAC
 IX24-1 GGGAAUUCAACGACAGGCUCCAUGUGGCAAGGAAAUUGGAGUGAUCUGAUACUACGAGCUCGAC
 IX36-1 GGGAAUUCAACACCACUAUUGUUGGCAAGGAAGCAAUGGAGUGAUCUGAUACUACGAGCUCGAC
 X02-1 GGGAAUUCAACUUAAGCAGUGGCAAGGAACUGCAAUGGAGUGAUCUGAUACUACGAGCUCGAC
 X04-2 GGGAAUUCAAGUACGGCAAGGAGUACAUCGUAGUAAUGGAGUGAUCUGAUACUACGAGCUCGAC
 X09-1 GGGAAUUCAAGUACGGCAAGGAGUACAUCGUAGGAAUGGAGUGAUCUGAUACUACGAGCUCGAC
 X10-7 GGGAAUUCAACACCACUAUUGUUGGCAAGGAAGCAAUGGAGUGAUCUGAUACUACGAGCUCGAC
 X13-1 GGGAAUUCAAGUACGGCAAGGAGUACAUCGUAGCAAUGGAGUGAUCUGAUACUACGAGCUCGAC
 X36-1 GGGAAUUCAAGACUAUACCAUUGUGGCAAGGAACAAUGGAGUGAUCUGAUACUACGAGCUCGAC
 X41-2 GGGAAUUCAAGCAUACUGCAUCGUGGCAAGGAACGAUUGGAGUGAUCUGAUACUACGAGCUCGAC
 XI1-17 GGGAAUUCAACACCACUAUUGUUGGCAAGGAAGCAAUGGAGUGAUCUGAUACUACGAGCUCGAC
 XI21-7 GGGAAUUCAAGUACGGCAAGGAGUACAUCGUAGCAAUGGAGUGAUCUGAUACUACGAGCUCGAC
 XI23-3 GGGAAUUCAAGUACGGCAAGGAGUACAUCGUAGUAAUGGAGUGAUCUGAUACUACGAGCUCGAC
 XI141-2 GGGAAUUCAACACAACCUGGUGGCAAGGAACCCAAUGGAGUGAUCUGAUACUACGAGCUCGAC
 XI13-1 GGGAAUUCAAGAAUAGCACAUUGUGGCAAGGAACAAUGGAGUGAUCUGAUACUACGAGCUCGAC
 XI101-1 GGGAAUUCAACACUACCUGGUGGCAAGGAACCCAAUGGAGUGAUCUGAUACUACGAGCUCGAC
 XI149-1 GGGAAUUCAACACCACUAUUGUUGGCAAGGAACAAUGGAGUGAUCUGAUACUACGAGCUCGAC
 XI107-1 GGGAAUUCAAGUACGGCAAGGAGUACAUCGUAACAAUGGAGUGAUCUGAUACUACGAGCUCGAC
 XI129-1 GGGAAUUCAACACCACUAUUGUUGGCAAGGAAGUAAUGGAGUGAUCUGAUACUACGAGCUCGAC
 XIV22-23 GGGAAUUCAACACCACUAUUGUUGGCAAGGAAGCAAUGGAGUGAUCUGAUACUACGAGCUCGAC
 XIV26-6 GGGAAUUCAAGUACGGCAAGGAGUACAUCGUAGCAAUGGAGUGAUCUGAUACUACGAGCUCGAC
 XIV1-2 GGGAAUUCAACACAACCUGGUGGCAAGGAACCCAAUGGAGUGAUCUGAUACUACGAGCUCGAC
 XIV32-1 GGGAAUUCAACACCACUAUUGUUGGCAAGGAAGCAAUGGAGUGAUCUGAUACUACGAGCUCGAC
 XIV5-1 GGGAAUUCAAGUACGGCAAGGAGUACAUCGCAGCAAUGGAGUGAUCUGAUACUACGAGCUCGAC
 XIV12-1 GGGAAUUCAACCCGUAUUGUUGGCAAGGAAGCAAUGGAGUGAUCUGAUACUACGAGCUCGAC

Motivo consenso: **NNDYGGCARGGARNNN**
 (((((((.....))))))

Figura 7

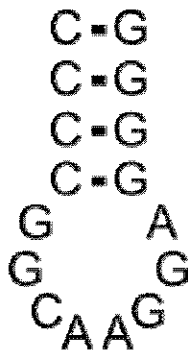
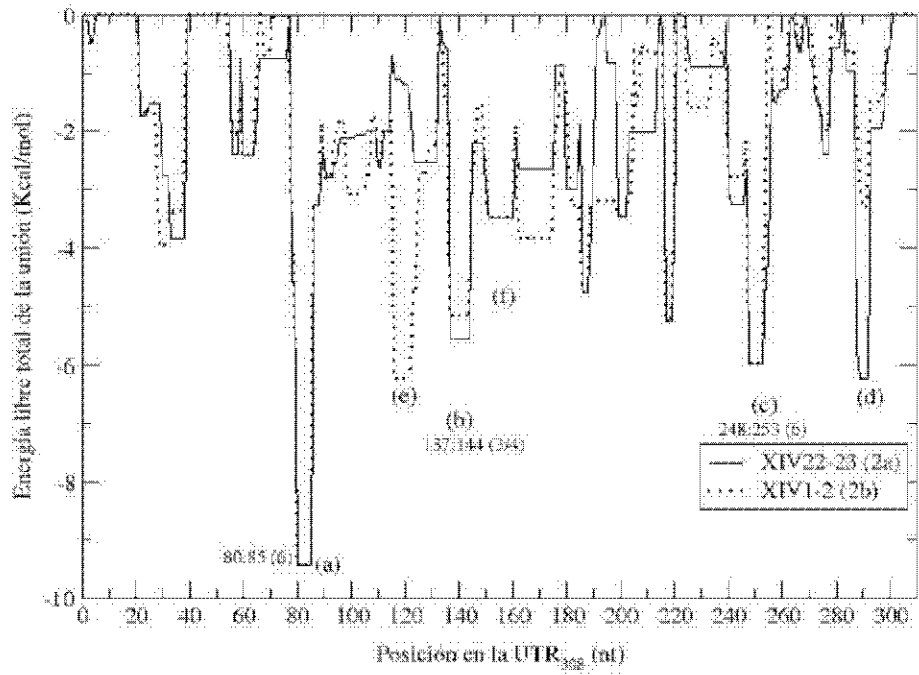


Figura 8

A)



B)

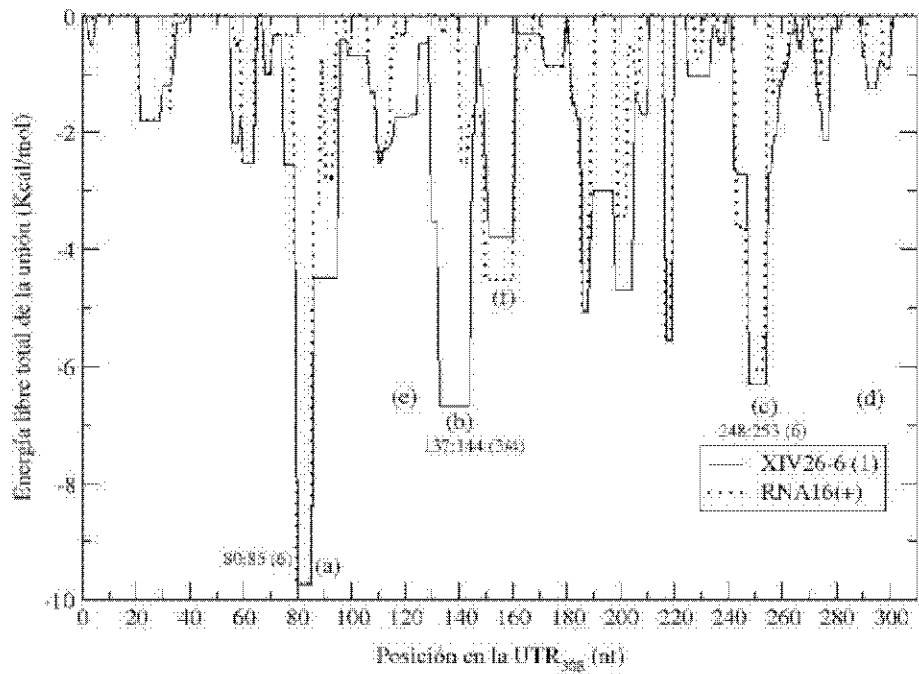


Figura 8 (cont.)

C)

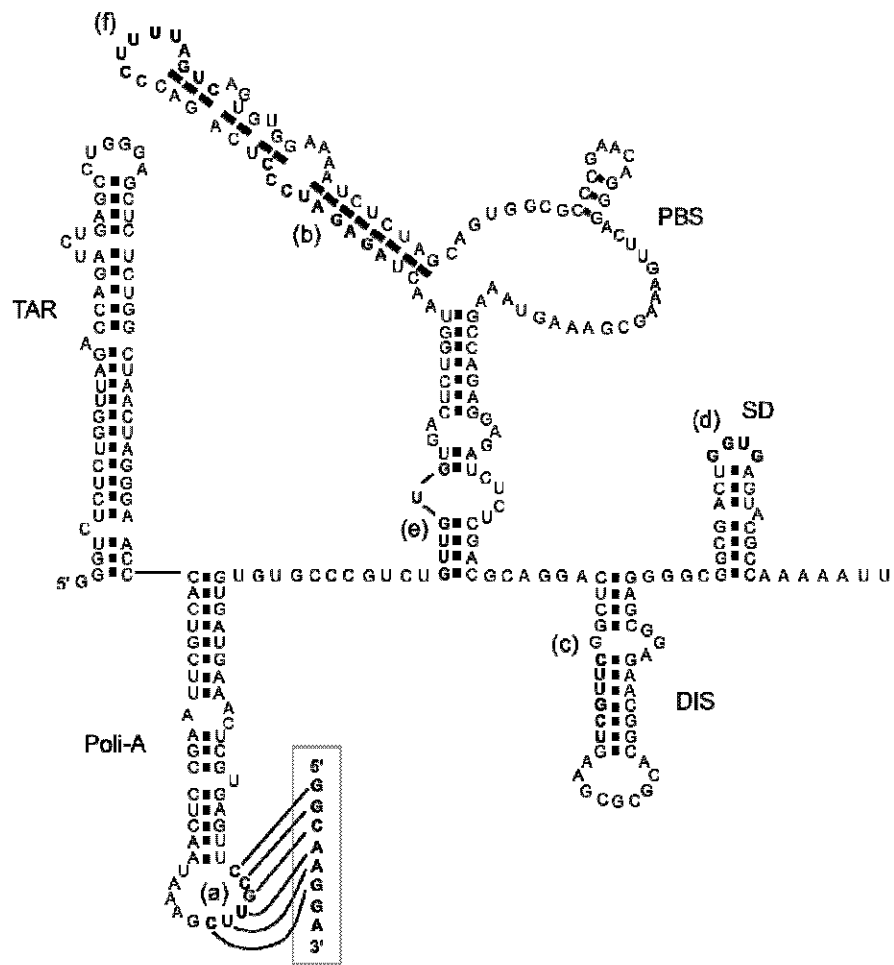


Figura 9

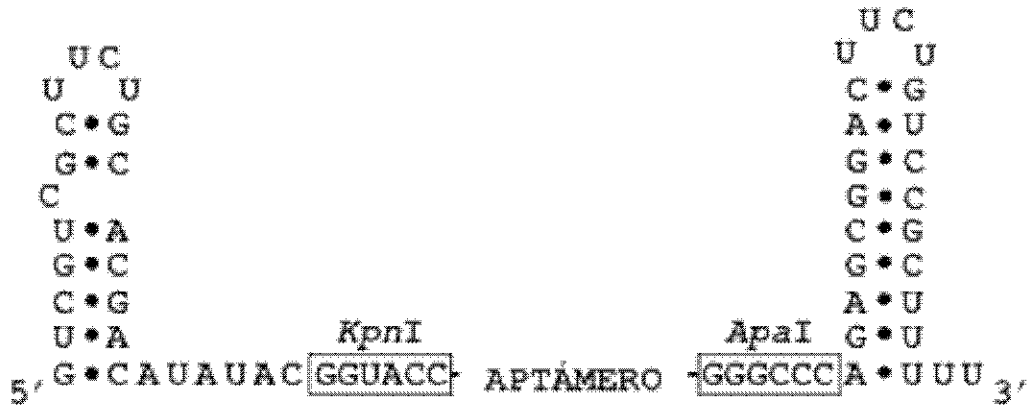


Figura 10

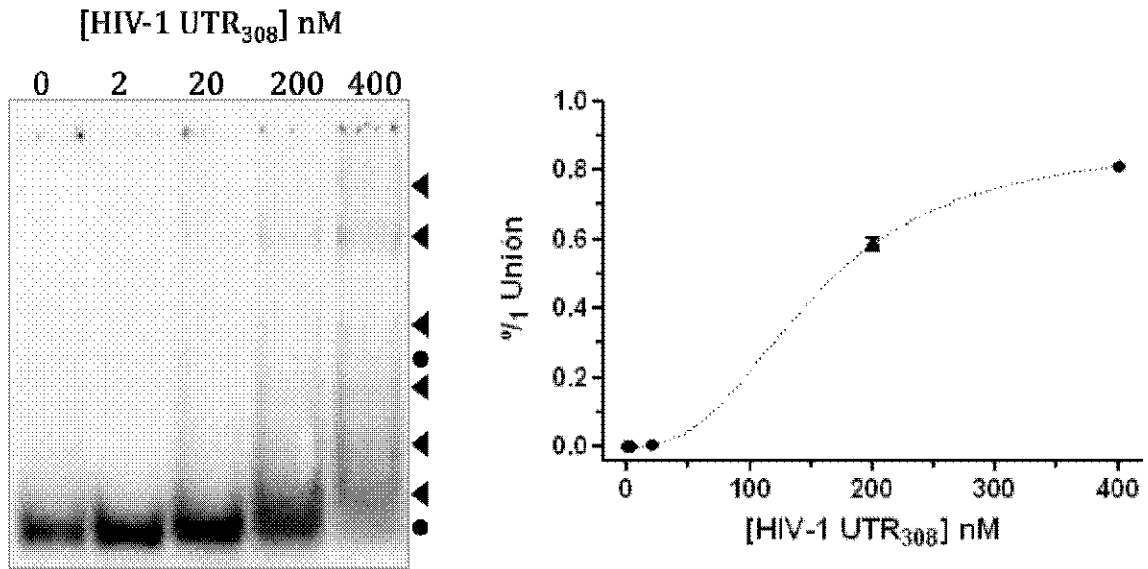


Figura 11

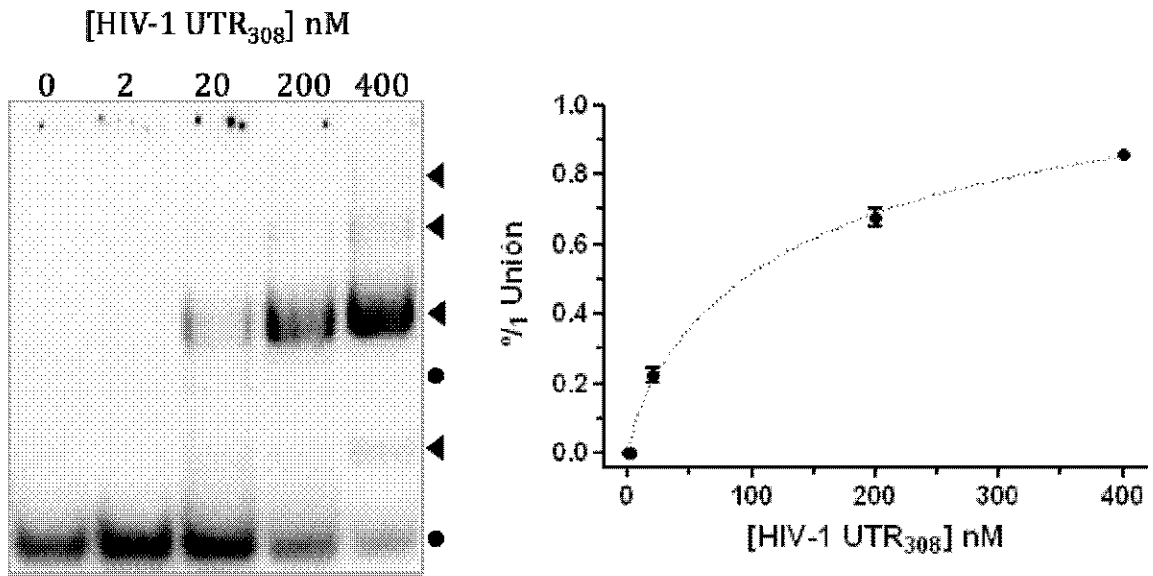


Figura 12

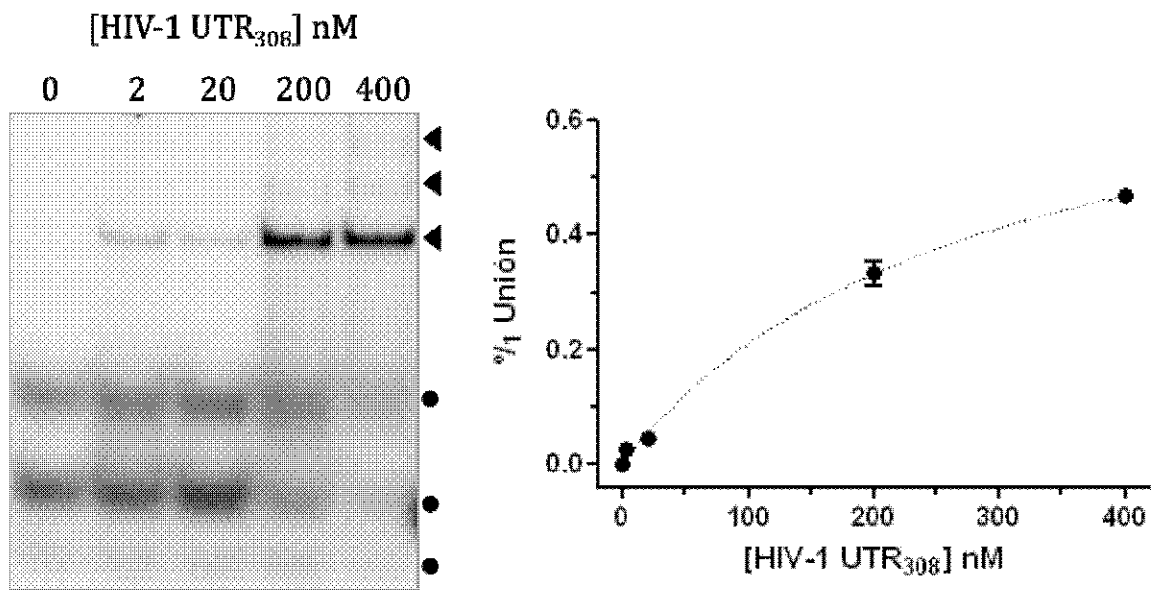


Figura 13

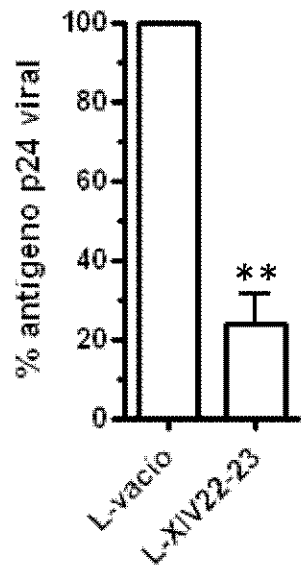


Figura 14

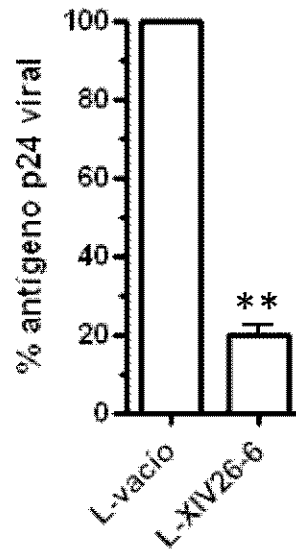
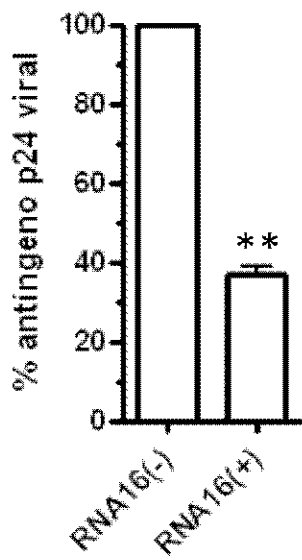


Figura 15



ES 2 471 670 A1

LISTA DE SECUENCIAS

<110> CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS
INSTITUTO NACIONAL DE TECNICA AEROSPACIAL

<120> MOLÉCULAS INHIBIDORAS DEL VIRUS DE LA INMUNODEFICIENCIA HUMANA TIPO 1
(VIH-1), PROCEDIMIENTO DE OBTENCIÓN Y SUS APLICACIONES

<160> 49

<170> BISSAP 1.0

<210> 1
<211> 308
<212> RNA
<213> HIV-1 vector pNL4-3

<220>
<221> source
<222> 1..308
<223> /mol_type="RNA"
/note="UTR308"
/organism="HIV-1 vector pNL4-3"

<220>
<221> misc_feature
<222> 1..308
<223> /note="Nucleótidos 1 al 308 del RNA genómico de la cepa NL4-3 de
VIH-1"

<220>
<221> misc_feature
<222> 1..98
<223> /note="Región R"

<220>
<221> misc_feature
<222> 1..58
<223> /note="TAR"

<220>
<221> misc_feature
<222> 59..105
<223> /note="Poli-A"

<220>
<221> misc_feature
<222> 81..86
<223> /note="Zona de unión del RNA16(+)"

<220>
<221> misc_feature
<222> 99..180
<223> /note="Región U5"

<220>
<221> misc_feature
<222> 117..237
<223> /note="PBS"

<220>
<221> misc_feature
<222> 244..278
<223> /note="DIS"

<220>
<221> misc_feature
<222> 283..301
<223> /note="SD"

<400> 1

ES 2 471 670 A1

```

gggucucucu gguuagacca gaucugagcc ugggagcucu cuggcuaacu agggaaccca      60
cugcuuaagc cucaauaaag cuugccuuga gugcucaaag uagugugugc cgcuguguug      120
ugugacucug guaacuagag aucccucaga cccuuuuagu caguguggaa aaucucuagc      180
aguggcgccc gaacagggac uugaaagcga aaguaaagcc agaggagauc ucucgacgca      240
ggacucggcu ugcugaagcg cgcacggcaa gaggcgaggg gcggcgacug gugaguacgc      300
caaaaauu                                     308

```

```

<210> 2
<211> 336
<212> RNA
<213> HIV-1 vector pNL4-3

```

```

<220>
<221> source
<222> 1..336
<223> /mol_type="RNA"
      /note="Región 5'UTR del RNA genómico del VIH-1 NL4-3"
      /organism="HIV-1 vector pNL4-3"

```

```

<220>
<221> misc_feature
<222> 1..98
<223> /note="Región R"

```

```

<220>
<221> misc_feature
<222> 1..58
<223> /note="TAR"

```

```

<220>
<221> misc_feature
<222> 59..105
<223> /note="Poli-A"

```

```

<220>
<221> misc_feature
<222> 81..86
<223> /note="Zona de unión del RNA16(+)"

```

```

<220>
<221> misc_feature
<222> 99..180
<223> /note="Región U5"

```

```

<220>
<221> misc_feature
<222> 117..237
<223> /note="PBS"

```

```

<220>
<221> misc_feature
<222> 244..278
<223> /note="DIS"

```

```

<220>
<221> misc_feature
<222> 283..301
<223> /note="SD"

```

```

<220>
<221> misc_feature
<222> 313..326
<223> /note="Psi"

```

```

<400> 2

```

ES 2 471 670 A1

```

gggucucucu gguuagacca gaucugagcc ugggagcucu cuggcuaacu agggaaccca      60
cugcuuaagc cucaauaaag cuugccuuga gugcucaaag uagugugugc cgcucuguug      120
ugugacucug guaacuagag aucccucaga cccuuuuagu caguguggaa aaucucuagc      180
aguggcgccc gaacagggac uugaaagcga aaguaaagcc agaggagauc ucucgacgca      240
ggacucggcu ugcugaagcg cgcacggcaa gaggcgaggg gcggcgacug gugaguacgc      300
caaaaauuuu gacuagcgga ggcuagaagg agagag                                  336

```

```

<210> 3
<211> 290
<212> RNA
<213> HIV-1 vector pNL4-3

```

```

<220>
<221> source
<222> 1..290
<223> /mol_type="RNA"
      /note="Región común de las 5'UTR de los RNAs genómico y
      subgenómicos del VIH-1 NL4-3"
      /organism="HIV-1 vector pNL4-3"

```

```

<220>
<221> misc_feature
<222> 1..98
<223> /note="Región R"

```

```

<220>
<221> misc_feature
<222> 1..58
<223> /note="TAR"

```

```

<220>
<221> misc_feature
<222> 59..105
<223> /note="Polí-A"

```

```

<220>
<221> misc_feature
<222> 81..86
<223> /note="Zona de unión del RNA16(+)"

```

```

<220>
<221> misc_feature
<222> 99..180
<223> /note="Región U5"

```

```

<220>
<221> misc_feature
<222> 117..237
<223> /note="PBS"

```

```

<220>
<221> misc_feature
<222> 244..278
<223> /note="DIS"

```

```

<220>
<221> misc_feature
<222> 283..290
<223> /note="Nucleótidos 1 al 8 del SD"

```

```

<400> 3
gggucucucu gguuagacca gaucugagcc ugggagcucu cuggcuaacu agggaaccca      60
cugcuuaagc cucaauaaag cuugccuuga gugcucaaag uagugugugc cgcucuguug      120

```


ES 2 471 670 A1

ugugacucug gaaacuagag auccucaga cccuuuagu caguguggaa aaucucuagc 180
 aguggcgccc gaacaggac uugaaagcga aaguaagcc agaggagauc ucucgacgca 240
 ggacucggcu ugcugaagcg cgcacggcaa gaggcgaggg gcggcgacug 290

<210> 4
 <211> 25
 <212> RNA
 <213> Homo sapiens

<220>
 <221> source
 <222> 1..25
 <223> /mol_type="RNA"
 /note="Horquilla 5' del snRNA U6"
 /organism="Homo sapiens"

<400> 4
 gucgucgcu ucgacgaca uauac 25

<210> 5
 <211> 24
 <212> RNA
 <213> Homo sapiens

<220>
 <221> source
 <222> 1..24
 <223> /mol_type="RNA"
 /note="Horquilla 3' del snRNA U6"
 /organism="Homo sapiens"

<400> 5
 agagcggacu ucuguccgcu uuuu 24

<210> 6
 <211> 16
 <212> DNA
 <213> Artificial sequence

<220>
 <221> source
 <222> 1..16
 <223> /mol_type="DNA"
 /note="Motivo consenso var-DNA"
 /organism="Artificial sequence"

<220>
 <221> misc_feature
 <222> 1
 <223> /note="Nucleótido N, se selecciona independientemente entre A, C,
 G o T"

<220>
 <221> misc_feature
 <222> 2
 <223> /note="Nucleótido N, se selecciona independientemente entre A, C,
 G o T"

<220>
 <221> misc_feature
 <222> 3
 <223> /note="Nucleótido N, se selecciona independientemente entre A, C,
 G o T"

<220>
 <221> misc_feature

ES 2 471 670 A1

<222> 4
<223> /note="Nucleótido N, se selecciona independientemente entre A, C,
G o T"

<220>
<221> misc_feature
<222> 9
<223> /note="Nucleótido R, se selecciona independientemente entre A o G"

<220>
<221> misc_feature
<222> 13
<223> /note="Nucleótido N, es el complementario al nucleótido en
posición 4"

<220>
<221> misc_feature
<222> 14
<223> /note="Nucleótido N, es el complementario al nucleótido en
posición 3"

<220>
<221> misc_feature
<222> 15
<223> /note="Nucleótido N, es el complementario al nucleótido en
posición 2"

<220>
<221> misc_feature
<222> 16
<223> /note="Nucleótido N, es el complementario al nucleótido en
posición 1"

<400> 6
nnnnggcarg gannnn

16

<210> 7
<211> 16
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..16
<223> /mol_type="RNA"
/note="Motivo consenso var-RNA"
/organism="Artificial Sequence"

<220>
<221> misc_feature
<222> 1
<223> /note="Nucleótido N, se selecciona independientemente entre A, C,
G o U"

<220>
<221> misc_feature
<222> 2
<223> /note="Nucleótido N, se selecciona independientemente entre A, C,
G o U"

<220>
<221> misc_feature
<222> 3
<223> /note="Nucleótido N, se selecciona independientemente entre A, C,
G o U"

<220>
<221> misc_feature
<222> 4

ES 2 471 670 A1

<223> /note="Nucleótido N, se selecciona independientemente entre A, C, G o U"

<220>

<221> misc_feature

<222> 9

<223> /note="Nucleótido R, se selecciona independientemente entre A o G"

<220>

<221> misc_feature

<222> 13

<223> /note="Nucleótido N, es el complementario al nucleótido en posición 4"

<220>

<221> misc_feature

<222> 14

<223> /note="Nucleótido N, es el complementario al nucleótido en posición 3"

<220>

<221> misc_feature

<222> 15

<223> /note="Nucleótido N, es el complementario al nucleótido en posición 2"

<220>

<221> misc_feature

<222> 16

<223> /note="Nucleótido N, es el complementario al nucleótido en posición 1"

<400> 7

nnnnggcarg gannnn

16

<210> 8

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<222> 1..16

<223> /mol_type="DNA"
/note="Motivo consenso var-DNA preferente"
/organism="Artificial Sequence"

<220>

<221> misc_feature

<222> 1

<223> /note="Nucleótido N, se selecciona independientemente entre A, C, G o T"

<220>

<221> misc_feature

<222> 2

<223> /note="Nucleótido N, se selecciona independientemente entre A, C, G o T"

<220>

<221> misc_feature

<222> 3

<223> /note="Nucleótido D, se selecciona independientemente entre A, G o T"

<220>

<221> misc_feature

<222> 4

<223> /note="Nucleótido Y, se selecciona independientemente entre T o C"

ES 2 471 670 A1

"

<220>
 <221> misc_feature
 <222> 9
 <223> /note="Nucleótido R, se selecciona independientemente entre A o G"

<220>
 <221> misc_feature
 <222> 13
 <223> /note="Nucleótido R, que es A cuando Y es T, o es G cuando Y es C"

<220>
 <221> misc_feature
 <222> 14
 <223> /note="Nucleótido N, que es T cuando D es A, o es C cuando D es G, o es A cuando D es T"

<220>
 <221> misc_feature
 <222> 15
 <223> /note="Nucleótido N, que es A cuando N en posición 2 es T, o es C cuando N en posición 2 es G, o es G cuando N en posición 2 es C, o es T cuando N en posición 2 es A"

<220>
 <221> misc_feature
 <222> 16
 <223> /note="Nucleótido N, que es A cuando N en posición 1 es T, o es C cuando N en posición 1 es G, o es G cuando N en posición 1 es C, o es T cuando N en posición 1 es A"

<400> 8
 ndyggcarg garnnn 16

<210> 9
 <211> 16
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <222> 1..16
 <223> /mol_type="RNA"
 /note="Motivo consenso var-RNA preferente"
 /organism="Artificial Sequence"

<220>
 <221> misc_feature
 <222> 1
 <223> /note="Nucleótido N, se selecciona independientemente entre A, C, G o U"

<220>
 <221> misc_feature
 <222> 2
 <223> /note="Nucleótido N, se selecciona independientemente entre A, C, G o U"

<220>
 <221> misc_feature
 <222> 3
 <223> /note="Nucleótido D, se selecciona independientemente entre A, G o U"

<220>
 <221> misc_feature
 <222> 4

ES 2 471 670 A1

<223> /note="Nucleótido Y, se selecciona independientemente entre U o C"

<220>

<221> misc_feature

<222> 9

<223> /note="Nucleótido R, se selecciona independientemente entre A o G"

<220>

<221> misc_feature

<222> 13

<223> /note="Nucleótido R, que se selecciona entre A o G cuando Y es U, o es G cuando Y es C"

<220>

<221> misc_feature

<222> 14

<223> /note="Nucleótido N, que es U cuando D es A, o se selecciona entre C o U cuando D es G, o se selecciona entre A o G cuando D es U"

<220>

<221> misc_feature

<222> 15

<223> /note="Nucleótido N,seleccionado entre A o G cuando N en posición 2 es U, o entre C o U cuando N en posición 2 es G, o es G cuando N en posición 2 es C, o es U cuando N en posición 2 es A o G"

<220>

<221> misc_feature

<222> 16

<223> /note="Nucleótido N,seleccionado entre A o G cuando N en posición 1 es U, o entre C o U cuando N en posición 1 es G, o es G cuando N en posición 1 es C, o es U cuando N en posición 1 es A o G"

<400> 9

nndyggcarg garnnn

16

<210> 10

<211> 64

<212> RNA

<213> Artificial sequence

<220>

<221> source

<222> 1..64

<223> /mol_type="RNA"
/note="IX03-1"
/organism="Artificial sequence"

<220>

<221> misc_feature

<222> 1..10

<223> /note="Región Constante 5'"

<220>

<221> misc_feature

<222> 11..35

<223> /note="Región Seleccionada"

<220>

<221> misc_feature

<222> 36..64

<223> /note="Región Constante 3'"

<400> 10

gggaauuca gacacgaaca uaguggcaag gaacuaugga gugaucugau acuacgagcu

60

cgac

64

ES 2 471 670 A1

```

<210> 11
<211> 64
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
      /note="IX24-1"
      /organism="Artificial Sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región Seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 11
gggaauuca cgcacaggcuc cauguggcaa ggaaaugga gugaucugau acuacgagcu      60
cgac                                                                                   64

<210> 12
<211> 64
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
      /note="IX36-1"
      /organism="Artificial Sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región Seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 12
gggaauuca caccacuauu guuggcaagg aagcaugga gugaucugau acuacgagcu      60
cgac                                                                                   64

<210> 13
<211> 64
<212> RNA
<213> Artificial Sequence

```

ES 2 471 670 A1

```

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
      /note="X02-1"
      /organism="Artificial sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región Seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 13
gggaauuca cuucaagcag uggcaaggaa cugcaaugga gugaucugau acuacgagcu      60
cgac                                                                    64

<210> 14
<211> 64
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
      /note="X04-2"
      /organism="Artificial sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región Seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 14
gggaauuca guacggcaag gaguacaucg uaguaaugga gugaucugau acuacgagcu      60
cgac                                                                    64

<210> 15
<211> 64
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"

```

ES 2 471 670 A1

```

    /note="X09-1"
    /organism="Artificial Sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 15
gggaauuca guacggcaag gaguacaucg uaggaugga gugaucugau acuacgagcu    60
cgac                                                                    64

<210> 16
<211> 64
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
    /note="X10-7"
    /organism="Artificial Sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 16
gggaauuca caccacuauu guuggcaagg aagcaugga gugaucugau acuacgagcu    60
cgac                                                                    64

<210> 17
<211> 64
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
    /note="X13-1"
    /organism="Artificial Sequence"

<220>
<221> misc_feature

```


ES 2 471 670 A1

```

<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 17
gggaauuca g uacggcaag gaguacaucg uagcaaugga gugaucugau acuacgagcu      60
cgac                                                                    64

<210> 18
<211> 64
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
      /note="X36-1"
      /organism="Artificial Sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 18
gggaauuca gacua uacca uuguggcaag gaacaaugga gugaucugau acuacgagcu      60
cgac                                                                    64

<210> 19
<211> 64
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
      /note="X41-2"
      /organism="Artificial Sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature

```

ES 2 471 670 A1

```

<222> 11..35
<223> /note="Región Seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 19
gggaauuca gcauacugca ucguggcaag gaacgaugga gugaucugau acuacgagcu      60
cgac                                                                    64

<210> 20
<211> 64
<212> RNA
<213> Artificial sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
      /note="XI1-17"
      /organism="Artificial sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región Seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 20
gggaauuca caccacuauu guuggcaagg aagcaugga gugaucugau acuacgagcu      60
cgac                                                                    64

<210> 21
<211> 64
<212> RNA
<213> Artificial sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
      /note="XI21-7"
      /organism="Artificial sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región Seleccionada"

<220>
<221> misc_feature

```

ES 2 471 670 A1

```

<222> 36..64
<223> /note="Región Constante 3'"

<400> 21
gggaauuca guacggcaag gaguacaucg uagcaaugga gugaucugau acuacgagcu      60
cgac                                                                    64

<210> 22
<211> 64
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
      /note="XI23-3"
      /organism="Artificial Sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región Seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 22
gggaauuca guacggcaag gaguacaucg uaguaaugga gugaucugau acuacgagcu      60
cgac                                                                    64

<210> 23
<211> 64
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
      /note="XI141-2"
      /organism="Artificial Sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región Seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 23
gggaauuca cacaaccugg guggcaagga acccaaugga gugaucugau acuacgagcu      60

```

```

cgac 64

<210> 24
<211> 64
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
      /note="XI13-1"
      /organism="Artificial sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región Seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 24
gggaauuca gaauagcaca uuguggcaag gaacaugga gugaucugau acuacgagcu 60

cgac 64

<210> 25
<211> 64
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
      /note="XI101-1"
      /organism="Artificial sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región Seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 25
gggaauuca cacuaccugg guggcaagga acccaugga gugaucugau acuacgagcu 60

cgac 64

<210> 26

```

ES 2 471 670 A1

```

<211> 64
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
      /note="XI149-1"
      /organism="Artificial sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región Seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 26
gggaauuca caccacuau guuggcaagg aaacaugga gugaucugau acuacgagcu      60
cgac                                                                64

<210> 27
<211> 64
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
      /note="XI107-1"
      /organism="Artificial sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región Seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 27
gggaauuca guacggcaag gaguacaucg uaacaugga gugaucugau acuacgagcu      60
cgac                                                                64

<210> 28
<211> 64
<212> RNA
<213> Artificial Sequence

<220>

```

ES 2 471 670 A1

```

<221> source
<222> 1..64
<223> /mol_type="RNA"
      /note="XI129-1"
      /organism="Artificial Sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región Seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 28
gggaauuca caccacuau guuggcaagg aaguaugga gugaucugau acuacgagcu      60
cgac                                             64

<210> 29
<211> 64
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
      /note="XIV22-23"
      /organism="Artificial Sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región Seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 29
gggaauuca caccacuau guuggcaagg aagcaugga gugaucugau acuacgagcu      60
cgac                                             64

<210> 30
<211> 64
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
      /note="XIV26-6"
      /organism="Artificial Sequence"

```

ES 2 471 670 A1

```

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 30
gggaauuca g uacggcaag gaguacaucg uagcaaugga gugaucugau acuacgagcu 60
cgac 64

<210> 31
<211> 64
<212> RNA
<213> Artificial sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
      /note="XIV1-2"
      /organism="Artificial sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 31
gggaauuca cacaaccugg guggcaagga acccaaugga gugaucugau acuacgagcu 60
cgac 64

<210> 32
<211> 64
<212> RNA
<213> Artificial sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
      /note="XIV32-1"
      /organism="Artificial sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

```

ES 2 471 670 A1

```

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región Seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 32
gggaauuca caccacuauu guuggcaggg aagcaaugga gugaucugau acuacgagcu    60
cgac                                                                    64

<210> 33
<211> 64
<212> RNA
<213> Artificial sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
      /note="XIV5-1"
      /organism="Artificial sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región Seleccionada"

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 33
gggaauuca guacggcaag gaguacaucg cagcaaugga gugaucugau acuacgagcu    60
cgac                                                                    64

<210> 34
<211> 64
<212> RNA
<213> Artificial sequence

<220>
<221> source
<222> 1..64
<223> /mol_type="RNA"
      /note="XIV12-1"
      /organism="Artificial sequence"

<220>
<221> misc_feature
<222> 1..10
<223> /note="Región Constante 5'"

<220>
<221> misc_feature
<222> 11..35
<223> /note="Región Seleccionada"

```


ES 2 471 670 A1

```

<220>
<221> misc_feature
<222> 36..64
<223> /note="Región Constante 3'"

<400> 34
gggaauuca caccgcuauu guuggcaagg aagcaaugga gugaucugau acuacgagcu      60
cgac                                                                    64

<210> 35
<211> 16
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..16
<223> /mol_type="RNA"
      /note="RNA16(+)"
      /organism="Artificial sequence"

<220>
<221> misc_feature
<222> 5..12
<223> /note="Apical Loop"

<400> 35
ccccggcaag gagggg                                                    16

<210> 36
<211> 16
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..16
<223> /mol_type="RNA"
      /note="RNA16(+) var-G"
      /organism="Artificial sequence"

<220>
<221> misc_feature
<222> 5..12
<223> /note="Apical Loop var-G"

<400> 36
ccccgacagg gagggg                                                    16

<210> 37
<211> 125
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..125
<223> /mol_type="RNA"
      /note="L-XIV22-23"
      /organism="Artificial sequence"

<220>
<221> misc_feature
<222> 1..25
<223> /note="horquilla 5' del snRNA U6"

```

ES 2 471 670 A1

```

<220>
<221> misc_feature
<222> 32..95
<223> /note="XIV22-23"

<220>
<221> misc_feature
<222> 102..125
<223> /note="horquilla 3' de1 snRNA U6"

<400> 37
gucgucgcuu cugcacgaca uauacgguac cgggaauuca acaccacuau uguuggcaag      60
gaagcaaugg agugaucuga uacuacgagc ucgacgggcc cagagcggac uucuguccgc      120
uuuuu                                          125

<210> 38
<211> 125
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..125
<223> /mol_type="RNA"
      /note="L-XIV26-6"
      /organism="Artificial sequence"

<220>
<221> misc_feature
<222> 1..25
<223> /note="horquilla 5' de1 snRNA U6"

<220>
<221> misc_feature
<222> 32..95
<223> /note="XIV26-6"

<220>
<221> misc_feature
<222> 102..125
<223> /note="horquilla 3' de1 snRNA U6"

<400> 38
gucgucgcuu cugcacgaca uauacgguac cgggaauuca aguacggcaa ggaguacauc      60
guagcaaugg agugaucuga uacuacgagc ucgacgggcc cagagcggac uucuguccgc      120
uuuuu                                          125

<210> 39
<211> 16
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..16
<223> /mol_type="RNA"
      /note="RNA16(+) genérico"
      /organism="Artificial Sequence"

<220>
<221> misc_feature
<222> 5..12
<223> /note="Apical Loop genérico"

<220>

```

ES 2 471 670 A1

```

<221> misc_feature
<222> 9
<223> /note="Nucleótido R, se selecciona entre A o G"

<400> 39
ccccggcarg gagggg                                     16

<210> 40
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..30
<223> /mol_type="DNA"
      /note="5'EcoRIK"
      /organism="Artificial Sequence"

<220>
<221> misc_feature
<222> 4..23
<223> /note="promotor de la T7 RNA polimerasa"

<220>
<221> misc_feature
<222> 23..28
<223> /note="diana de corte para la enzima EcoRI"

<400> 40
ggataatacg actcactata ggaattcaa                       30

<210> 41
<211> 71
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..71
<223> /mol_type="DNA"
      /note="3'RANDOMK"
      /organism="Artificial Sequence"

<220>
<221> misc_feature
<222> 4..9
<223> /note="diana de corte para la enzima XhoI"

<220>
<221> misc_feature
<222> 30..54
<223> /note="región de secuencia aleatoria"

<220>
<221> misc_feature
<222> 30
<223> /note="Nucleótido N, se selecciona independientemente entre A, C,
      G o T"

<220>
<221> misc_feature
<222> 31
<223> /note="Nucleótido N, se selecciona independientemente entre A, C,
      G o T"

<220>
<221> misc_feature
<222> 32

```

ES 2 471 670 A1

<223> /note="Nucleótido N, se selecciona independientemente entre A, C, G o T"

<220>

<221> misc_feature

<222> 33

<223> /note="Nucleótido N, se selecciona independientemente entre A, C, G o T"

<220>

<221> misc_feature

<222> 34

<223> /note="Nucleótido N, se selecciona independientemente entre A, C, G o T"

<220>

<221> misc_feature

<222> 35

<223> /note="Nucleótido N, se selecciona independientemente entre A, C, G o T"

<220>

<221> misc_feature

<222> 36

<223> /note="Nucleótido N, se selecciona independientemente entre A, C, G o T"

<220>

<221> misc_feature

<222> 37

<223> /note="Nucleótido N, se selecciona independientemente entre A, C, G o T"

<220>

<221> misc_feature

<222> 38

<223> /note="Nucleótido N, se selecciona independientemente entre A, C, G o T"

<220>

<221> misc_feature

<222> 39

<223> /note="Nucleótido N, se selecciona independientemente entre A, C, G o T"

<220>

<221> misc_feature

<222> 40

<223> /note="Nucleótido N, se selecciona independientemente entre A, C, G o T"

<220>

<221> misc_feature

<222> 41

<223> /note="Nucleótido N, se selecciona independientemente entre A, C, G o T"

<220>

<221> misc_feature

<222> 42

<223> /note="Nucleótido N, se selecciona independientemente entre A, C, G o T"

<220>

<221> misc_feature

<222> 43

<223> /note="Nucleótido N, se selecciona independientemente entre A, C, G o T"

<220>

<221> misc_feature

ES 2 471 670 A1

```

<222> 44
<223> /note="Nucleótido N, se selecciona independientemente entre A, C,
      G o T"

<220>
<221> misc_feature
<222> 45
<223> /note="Nucleótido N, se selecciona independientemente entre A, C,
      G o T"

<220>
<221> misc_feature
<222> 46
<223> /note="Nucleótido N, se selecciona independientemente entre A, C,
      G o T"

<220>
<221> misc_feature
<222> 47
<223> /note="Nucleótido N, se selecciona independientemente entre A, C,
      G o T"

<220>
<221> misc_feature
<222> 48
<223> /note="Nucleótido N, se selecciona independientemente entre A, C,
      G o T"

<220>
<221> misc_feature
<222> 49
<223> /note="Nucleótido N, se selecciona independientemente entre A, C,
      G o T"

<220>
<221> misc_feature
<222> 50
<223> /note="Nucleótido N, se selecciona independientemente entre A, C,
      G o T"

<220>
<221> misc_feature
<222> 51
<223> /note="Nucleótido N, se selecciona independientemente entre A, C,
      G o T"

<220>
<221> misc_feature
<222> 52
<223> /note="Nucleótido N, se selecciona independientemente entre A, C,
      G o T"

<220>
<221> misc_feature
<222> 53
<223> /note="Nucleótido N, se selecciona independientemente entre A, C,
      G o T"

<220>
<221> misc_feature
<222> 54
<223> /note="Nucleótido N, se selecciona independientemente entre A, C,
      G o T"

<220>
<221> misc_feature
<222> 57..62
<223> /note="diana de corte para la enzima EcoRI"

<400> 41
gtcgagctcg tagtatcaga tcactccath nnnnnnnnnn nnnnnnnnnn nnnnttgaat

```

tccctatagt g 71

<210> 42
 <211> 29
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> source
 <222> 1..29
 <223> /mol_type="DNA"
 /note="3'XhoIK"
 /organism="Artificial Sequence"

<220>
 <221> misc_feature
 <222> 4..9
 <223> /note="sitio de corte para la enzima XhoI"

<400> 42
 gtcgagctcg tagtatcaga tcactccat 29

<210> 43
 <211> 33
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> source
 <222> 1..33
 <223> /mol_type="DNA"
 /note="5'T7pNL43"
 /organism="Artificial Sequence"

<220>
 <221> misc_feature
 <222> 1..20
 <223> /note="promotor de la T7 RNA polimerasa"

<220>
 <221> misc_feature
 <222> 20..30
 <223> /note="nucleótidos del 1 al 11 del RNA genómico de la cepa NL4.3 de VIH-1"

<400> 43
 taatacgact cactataggg tctctctggt tag 33

<210> 44
 <211> 22
 <212> DNA
 <213> HIV-1 vector pNL4-3

<220>
 <221> source
 <222> 1..22
 <223> /mol_type="DNA"
 /note="3'T7pNL43"
 /organism="HIV-1 vector pNL4-3"

<220>
 <221> misc_feature
 <222> 1..22
 <223> /note="nucleótidos complementarios a la región comprendida entre las posiciones 307 y 288 del RNA genómico de la cepa NL4.3 de VIH-1"

ES 2 471 670 A1

<400> 44
aatttttggc gtactcacca gt 22

<210> 45
<211> 16
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..16
<223> /mol_type="RNA"
/note="RNA16(-)"
/organism="Artificial Sequence"

<400> 45
ccccgaaaac aagggg 16

<210> 46
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..22
<223> /mol_type="DNA"
/note="5'KpnIC3"
/organism="Artificial Sequence"

<220>
<221> misc_feature
<222> 7..12
<223> /note="sitio de corte para la enzima KpnI"

<400> 46
cgactcggta ccggaattc aa 22

<210> 47
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..26
<223> /mol_type="DNA"
/note="3'ApaIC3"
/organism="Artificial Sequence"

<220>
<221> misc_feature
<222> 4..9
<223> /note="sitio de corte para la enzima ApaI"

<400> 47
tctgggcccg tcgagctcgt agtatc 26

<210> 48
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<222> 1..41
<223> /mol_type="DNA"

ES 2 471 670 A1

```

    /note="5'T7U6"
    /organism="Artificial sequence"
<220>
<221> misc_feature
<222> 1..20
<223> /note="promotor de la T7 RNA polimerasa"

<220>
<221> misc_feature
<222> 21..41
<223> /note="secuencia de DNA equivalente a la de los nucleótidos 1 a
    21 de la horquilla 5' del snRNA U6 "

<400> 48
taatacgact cactataggg gtcgctcgctt ctgcacgaca t                                41

<210> 49
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<221> source
<222> 1..30
<223> /mol_type="DNA"
    /note="3' ApaIU6"
    /organism="Artificial sequence"

<220>
<221> misc_feature
<222> 4..9
<223> /note="sitio de corte para la enzima ApaI"

<220>
<221> misc_feature
<222> 10..30
<223> /note="secuencia de DNA complementaria a la de los nucleótidos 4
    a 24 de la horquilla 3' del snRNA U6"

<400> 49
agcgggccca aaaagcggac agaagtccgc                                30

```