

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 928 475**

51 Int. Cl.:

**C12N 15/10**

(2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **14.09.2017** **PCT/US2017/051674**

87 Fecha y número de publicación internacional: **22.03.2018** **WO18053209**

96 Fecha de presentación y número de la solicitud europea: **14.09.2017** **E 17781233 (6)**

97 Fecha y número de publicación de la concesión europea: **06.07.2022** **EP 3512944**

54 Título: **Composiciones de ARN de alta pureza y métodos para su preparación**

30 Prioridad:

**14.09.2016 US 201662394711 P**

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

**18.11.2022**

73 Titular/es:

**MODERNATX, INC. (100.0%)  
200 Technology Square  
Cambridge, MA 02139, US**

72 Inventor/es:

**HOGE, STEPHEN;  
ISSA, WILLIAM;  
MIRACCO, EDWARD J.;  
NELSON, JENNIFER;  
RABIDEAU, AMY, E. y  
BUTORA, GABOR**

74 Agente/Representante:

**FERNÁNDEZ POU, Felipe**

ES 2 928 475 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

## DESCRIPCIÓN

Composiciones de ARN de alta pureza y métodos para su preparación

## 5 Antecedentes de la invención

La capacidad de diseñar, sintetizar y administrar un ácido nucleico, por ejemplo, un ácido ribonucleico (ARN) por ejemplo, un ARN mensajero (ARNm) dentro de una célula, ha proporcionado avances en los campos de los agentes terapéuticos, de diagnóstico, reactivos y para ensayos biológicos. Se están realizando muchos avances en el proceso de traducción intracelular del ácido nucleico y la producción de al menos un péptido o polipéptido codificado de interés.

El ARNm tiene un potencial terapéutico inmenso ya que los agentes terapéuticos de ARNm pueden expresar de manera transitoria esencialmente cualquier proteína deseada y al mismo tiempo evita los efectos adversos de los enfoques de administración de ácido nucleico basado en ADN y viral. Las células de mamífero, en particular, las células humanas, sin embargo, contienen sensores de ácidos nucleicos que incluyen ARN como parte del sistema inmunitario innato - y es conveniente evitar dicha detección y respuesta inmunitaria cuando se desarrollan agentes terapéuticos de ARNm.

En teoría, los ARNm producidos mediante síntesis química son prometedores como agentes terapéuticos de ARNm, sin embargo, la mayoría de la investigación en esta importante área terapéutica a la fecha se ha enfocado en el ARNm transcrito *in vitro* (IVT, por sus siglas en inglés), dado que este proceso enzimático facilita la producción de ARN largos, en el orden de 1-2 o más kB, la longitud estándar de la mayoría de las moléculas de ARNm.

Los trabajos recientes muestran que la incorporación de nucleósidos modificados, en particular, pseudouridina, redujo la activación inmunitaria innata y aumentó la traducción de ARNm, pero la inducción residual de interferones tipo I (IFN) y citocinas proinflamatorias se mantiene (Kariko et ál. (2005) *Immunity* 23(2):165-75). Hubo una evolución en cuanto a la identificación de contaminantes en ARN de IVT modificado con nucleósidos que identifican ARN de cadena doble (ARNcd) como al menos parcialmente responsable de la activación inmunitaria innata. La eliminación de dichos contaminantes mediante cromatografía líquida de alto rendimiento (HPLC) dio como resultado menores niveles de citocinas inflamatorias e IFN y, a su vez, mayores niveles de expresión en células primarias (Kariko et ál. (2011) *Nuc. Acids Res.* 39:e142). Cabe destacar que los ARNm no modificados aún indujeron niveles altos de secreción de citocinas aunque se tradujeron mejor después de la purificación por HPLC.

WO 2013/102203 describe un método de tratamiento con RNAsa III utilizado para retirar el ARNcd de ARNm IVT para la transfección repetida o continua en células humanas o animales, en particular, para reprogramar las células de un estado de diferenciación a otro. El método pretende producir preparaciones que tienen niveles reducidos de ARNcd y niveles aumentados de ARNcs, tal como lo demuestran niveles mayores de factores de reprogramación y menos toxicidad para las células. Sin embargo, dichos métodos no son compatibles para usar en la preparación de ARNm para uso terapéutico, en particular, para uso terapéutico humano. Se sabe que la RNAsa III digiere ARNcs así como también ARNcd y al tratar de retirar los contaminantes de ARNcd, la integridad del producto de ARNcs se ve necesariamente comprometida. Por lo tanto, existe la necesidad de un mejor entendimiento de la naturaleza de los contaminantes en las preparaciones de ARNm generadas mediante IVT, con el fin de controlar mejor los niveles y la naturaleza de los contaminantes en las preparaciones de IVT. Existe además una necesidad de métodos mejorados para preparar ARNm para uso terapéutico y para composiciones de alta pureza producidas de acuerdo con dichos métodos.

## 50 Compendio de la invención

La invención implica el descubrimiento de métodos novedosos para la síntesis de ARN *in vitro* y productos relacionados. Las transcripciones de ARN producidas mediante los métodos descritos en la presente tienen propiedades mejoradas que dan como resultado composiciones superiores desde el punto de vista cualitativo y cuantitativo que comprenden dichas transcripciones de ARN. Las transcripciones de ARN producidas mediante los métodos descritos en la presente tienen propiedades mejoradas particularmente importantes para ARNm, ARNinc, y otros usos terapéuticos y de diagnóstico de ARN, tales como mejor silenciamiento inmunitario y mejores perfiles de seguridad.

En particular, las composiciones de ARN IVT producidas por los métodos de la invención, están sustancialmente libres de determinados contaminantes no deseados asociados de manera rutinaria con el proceso de IVT. Cabe destacar, sin embargo, que los métodos de la invención llegan a las composiciones de ARNm adecuadas para el uso terapéutico al controlar la naturaleza y los niveles de contaminantes producidos en la reacción de IVT, es decir, los contaminantes no se producen en la reacción inicial, en contraste con los métodos descritos en la técnica que intentan retirar los contaminantes una vez producidos. Sin limitarse a la teoría, se cree que evitar la producción de contaminantes no deseados en la reacción de IVT desde el inicio

proporciona composiciones mejoradas que tienen mayor pureza y potencia, medible, por ejemplo, en términos de un aumento de traducción de ARNm intacto de longitud completa en la composición.

5 Basándose en la descripción contenida en el presente documento, la presente invención proporciona un método de preparación de ARN que comprende

- 10 (a) formar una mezcla de reacción que comprenda una plantilla de ADN, una polimerasa de ARN seleccionada de la polimerasa T7 o la polimerasa T3, y NTP que incluyen adenosina trifosfato (ATP), citidina trifosfato (CTP), uridina trifosfato (UTP), guanosina trifosfato (GTP) y opcionalmente guanosina difosfato (GDP) o un análogo de cada nucleótido respectivo, y un tampón que contiene magnesio, y  
(b) incubar la mezcla de reacción en condiciones tales que el ARN se transcriba, produciendo así una preparación que comprenda ARN transcrito *in vitro* (IVT),

en donde:

- 15 1) la concentración de GTP es al menos 2 veces mayor que la concentración de cada uno de ATP, CTP y UTP;  
2) la mezcla de reacción comprende además guanosina difosfato (GDP) y en donde la concentración de GDP es al menos 2 veces mayor que la concentración de cada uno de ATP, CTP y UTP; o  
20 3) la mezcla de reacción comprende además GDP, y en donde la relación entre la concentración de GTP más GDP y la concentración de cada uno de ATP, CTP y UTP es de al menos 2:1.

La presente invención, y algunas realizaciones preferidas de la misma, se exponen en las reivindicaciones adjuntas.

25 El experto en la técnica entenderá que los productos de transcripción de complemento inverso no pretendidos o no deseados generados en una reacción de IVT pueden tener complementariedad no solo con la transcripción de ARN que es el producto pretendido o deseado de la reacción de IVT (por ejemplo, un ARNm, ARNinc, u otro ARN con una longitud mayor que 50 nucleótidos pretendido para uso terapéutico) sino que también puede  
30 tener complementariedad con una hebra de la plantilla de ADN a partir de la cual se produce la transcripción de ARN pretendida o deseada. Sin pretender limitarse a la teoría, se cree que determinados productos de transcripción no pretendidos o no deseados que contaminan las composiciones de ARN IVT que se reducen o eliminan de acuerdo con los procesos novedosos de la presente invención se transcriben a partir del producto de transcripción de ARN pretendido o deseado, existe la posibilidad de que determinados productos de  
35 transcripción no pretendidos o no deseados que contaminan las composiciones de ARN IVT puedan transcribirse a partir de la plantilla de ADN usada en las reacciones de IVT. La última suposición, aunque es posible, no parece ser comprobada por los datos presentados en la presente que demuestran productos de transcripción de complemento inverso predominantemente complementarios a la UTR 5' y/o cola de poliA de transcripciones de ARNm, mientras que los productos de transcripción de complemento inverso  
40 complementarios a partes (por ejemplo, elementos de secuencia) de una plantilla de ADN no presente en ARNm transcrito disminuyen de manera significativa y en algunos casos no son detectables.

El experto en la técnica entenderá que los contaminantes de ARN de una determinada estructura y/o longitud son bastante propensos a estimular respuestas inmunitarias no deseadas, por ejemplo, los contaminantes de  
45 ARN de al menos 15 o al menos 20 o al menos 25 nucleótidos de longitud, en particular, los contaminantes de ARN que son de naturaleza de cadena doble (ARNcd). La eliminación de dichos contaminantes es posible usando determinadas metodologías reconocidas en la técnica (por ejemplo, etapas de métodos o procesos enzimáticos y/o de purificación). Sin embargo, cada uno de dichos procesos o etapas de purificación adicional en la generación de, por ejemplo, ARNm, ARNinc u otro ARN mayor que 50 nucleótidos de longitud pretendido  
50 para uso terapéutico, introduce la posibilidad de menor fidelidad del producto pretendido, por ejemplo, al someter el producto de reacción de IVT directa a (1) condiciones enzimáticas (por ejemplo, tratamiento de RNasa que produce fragmentos de ARN) y/o (2) alta temperatura, condiciones de solvente no fisiológico (por ejemplo, condiciones de cromatografía HPLC o RP) lo que puede comprometer la calidad del producto de ARN en el proceso de intentar degradar o retirar contaminantes.

55 Cada una de las limitaciones de la invención puede comprender varias realizaciones de la invención.

#### Breve descripción de las figuras

60 No se pretende que las figuras adjuntas estén dibujadas a escala. En las figuras, cada componente idéntico o casi idéntico que se ilustra en varias figuras se representa con el mismo número. A efectos de claridad, no todos componentes pueden etiquetarse en cada figura. En las figuras:

La Figura 1 es una gráfica que ilustra los resultados de un ensayo de IFN- $\beta$  que selecciona variantes de química hEPO, nLuc y controles de vehículo así como también ARN-1 de modelo cortos en fibroblastos BJ.  
65 La Figura 2 muestra los resultados de un análisis LCMS de una transcripción de modelo corta. El modelo

demuestra que las especies fallidas están presentes en las tres químicas. La traza superior muestra el ARN-1 de modelo corto no modificado, la traza central muestra el ARN de modelo corto con todas las uridinas modificadas en pseudouridina y todas las citidinas modificadas con 5'O-metilo, y la traza inferior muestra el ARN1 de modelo corto con algunos residuos de uridina y citidina modificados.

5 La Figura 3 muestra los resultados de un análisis LCMS de una transcripción de modelo. El modelo demuestra los perfiles de impureza del ARN-4 de modelo y hEPO preparados mediante IVT.

La Figura 4 muestra que T7 se puede utilizar para realizar la transcripción de ARN con plantilla de ARN en ausencia de plantilla de ADN que, tras el tratamiento, confiere un producto inmunoestimulador.

10 Las Figuras 5A y 5B muestran el impacto de fase inversa (RP) e IVT con un exceso de GTP en la cantidad del sustrato de RNasa III. Tanto el proceso alfa como la purificación de RP reducen el sustrato de RIII. Se muestra un efecto aditivo de combinar ambos. La Figura 5A muestra un análisis de electroforesis capilar de material hEPO G5 tratado con RNasa III. La Figura 5B muestra un análisis de electroforesis capilar de material hEPO G0 tratado con RNasa III.

La Figura 6 muestra los datos de transfección de la expresión de proteína de hEPO e IFN-β.

15 La Figura 7 es un análisis de electroforesis capilar de una transcripción corta usando diferentes procesos y tratada con RNasa III. Los datos muestran el efecto de los ARN modelo tratados con RNasa III.

Las Figuras 8A y 8B muestran los resultados del método de pureza RP-IP. La Figura 8A muestra el ARN-4 de modelo sometido a tratamiento con RNasa III seguido de IVT usando el método equimolar y la Figura 8B muestra el ARN-4 de modelo sometido a tratamiento con RNasa III seguido de IVT con un exceso de GTP.

20 La Figura 9 es un fraccionamiento de RP de hEPO tratado con y sin RNasa III.

Las Figuras 10A-10D muestra los datos del analizador de fragmentos de RNasa III con fracción de hEPO después de equimolar (Figuras 10A, 10B, y 10C). El hEPO se modificó de manera que sus bases de uridina fueran 1-metilpseudouridina. El tratamiento con RNasa III no mostró diferencias de pureza considerables usando el proceso con exceso de GTP. Con equimolar hay un sustrato considerable. La Figura 10D muestra un análisis de IFNbeta in vitro de hEPO EQ G5 no tratado o después del tratamiento con RNasa III en condiciones equimolares.

25 Las Figuras 11A-11D muestran análisis de electroforesis capilar de hEPO Alfa fraccionado por RP +/- tratamiento con RIII. Las Figuras 11A, 11B y 11C muestran los efectos de IVT con un exceso de GTP en las reacciones. La Figura 11D muestra la IVT con exceso de GTP, lo que dio como resultado ninguna respuesta de IFN.

La Figura 12 muestra los resultados de un ensayo ELISA J2 anti-ARNcd.

La Figura 13 muestra que el ARNcd se elimina mediante el tratamiento con RNasa III.

La Figura 14 muestra los resultados de un estudio de caracterización de IVT, que ilustran que la IVT con un exceso de GTP es menos sensible a picos de citocinas inducidos por baja temperatura.

35 La Figura 15 muestra los resultados de nucleasa P1 del estudio de caracterización de IVT.

Las Figuras 16A a 16B muestra un análisis de impureza mediante LCMS de IVT basada en ARN en diferentes químicas usando G5 en proceso Equimolar (Figura 16A) y proceso alfa (Figura 16B).

La Figura 17 muestra IFN-β en fibroblastos BJ en condiciones diferentes de IVT.

La Figura 18 muestra que la vacuna no puede colocar casquetes en el ARNcd.

40 La Figura 19 muestra los efectos del tratamiento con CIP en especies de ARNcd diferentes.

La Figura 20 muestra datos de pureza FA de los experimentos *in vivo*.

La Figura 21 muestra la inducción de IFN-β en fibroblastos BJ.

La Figura 22 muestra la expresión *in vivo* de hEPO.

Las Figuras 23A a 23D muestran datos de citocinas Luminex de los experimentos *in vivo*.

45 La Figura 24 muestra frecuencias de activación de células B *in vivo*.

La Figura 25 muestra los resultados de un ensayo de IFN-β, que analiza ARNcd cortos. El ensayo es un análisis in vitro de oligos cortos trifosforilados en 5'.

La Figura 26 muestra los resultados de un ensayo de IFN-β, que analiza ARNcd de 20mer y poliU/A.

La Figura 27 muestra un análisis del excedente en 3' con respecto a la respuesta de IFN-β.

50 La Figura 28 muestra los resultados de un ensayo de citocina que analiza estándares de ARNcd con excedente en 5', dúplex perfecto, y excedente en 3' de longitudes variables.

La Figura 29 es una gráfica que ilustra el análisis in vitro de especies de poliU.

La Figura 30 es una gráfica que ilustra el análisis in vitro de estándares de oligos de ARNcs.

55 La Figura 31 es una gráfica que ilustra el análisis in vitro de estándares de oligos de ARNcd con diferentes funcionalidades en 5'.

La Figura 32 es una gráfica que demuestra que la fosfatasa no puede desfosforilar ARNcd.

La Figura 33 es una gráfica que demuestra la respuesta a la dosis de impurezas de ARNcs (IFNbeta en fibroblastos BJ).

60 La Figura 34 es una gráfica que demuestra la respuesta a la dosis de impurezas de ARNcd (IFNbeta en fibroblastos BJ).

La Figura 35 es una gráfica que demuestra la respuesta de IFNbeta para nucleótido modificado en 5' en estándares de oligo directos.

La Figura 36 es una gráfica que demuestra la respuesta de IFNbeta para nucleótido modificado en 5' en estándares de oligo de complemento inverso.

65 La Figura 37 es una gráfica que demuestra la respuesta de IFNbeta para ARNcd funcionalizado con hidroxilo en 5'.

La Figura 38 es una gráfica que muestra que el proceso alfa genera más OH (limpio) que el proceso equimolar. La Figura 39 es una gráfica que muestra ARN<sub>cd</sub> calculado para 1 µg de ARN<sub>m</sub>. La Figura 40 es un esquema que muestra un proceso de transcripción in vitro (IVT) tradicional y tipos de impurezas formadas.

5

### Descripción detallada

Con el fin de mejorar los métodos para la fabricación de polímeros que codifican proteínas, se han desarrollado nuevos métodos para generar ARN. Se ha descubierto que se pueden realizar cambios a un proceso de transcripción in vitro para producir una preparación de ARN que tiene propiedades enormemente diferentes al ARN producido usando un proceso de transcripción in vitro (IVT) tradicional. Las preparaciones de ARN producidas de acuerdo con los métodos de la invención (también denominadas en la presente composiciones de ARN IVT) tienen propiedades que permiten la producción de composiciones superiores desde el punto de vista cualitativo y cuantitativo que comprenden dichas transcripciones de ARN. Incluso cuando se acoplan a procesos de purificación extensos, el ARN producido usando métodos de IVT tradicionales es distinto desde el punto de vista cualitativo y cuantitativo de las preparaciones de ARN producidas por los métodos de la invención. Por ejemplo, las preparaciones de ARN producidas por los métodos de la invención (y las composiciones que las comprenden) son menos inmunogénicas en comparación con preparaciones de ARN (y las composiciones que las comprenden) elaboradas usando IVT tradicional. Las preparaciones de ARN producidas de acuerdo con los métodos de la invención (también denominadas en la presente composiciones de ARN IVT) además tienen propiedades que permiten la producción de proteínas superiores desde el punto de vista cualitativo y cuantitativo, por ejemplo, cuando se traducen. Por ejemplo, la proteína generada a partir de preparaciones de ARN producidas por los métodos de la invención es menos inmunogénica en comparación con preparaciones de ARN elaboradas usando IVT tradicional.

25

De manera adicional, se producen mayores niveles de expresión de proteínas con mayor pureza a partir de las preparaciones de ARN descritas en la presente. Aunque no se limita a un mecanismo, se cree que los niveles de expresión de proteínas sustanciales son un resultado de la alta integridad del ARN<sub>m</sub> en las muestras purificadas. Si bien algunos procedimientos de purificación pueden retirar de manera eficaz un nivel de contaminantes mediante la degradación de esos contaminantes, la integridad del producto farmacéutico se ve afectada de manera negativa. Por ejemplo, se afirma en la técnica anterior que la digestión con RNasa de muestras de ARN<sub>m</sub> es útil para eliminar los contaminantes de ARN. Sin embargo, la digestión de RNasa también reduce la integridad del ARN<sub>m</sub> al degradar partes de la transcripción de longitud completa producida por la reacción de IVT. A diferencia de la IVT/procesos de purificación de la técnica previa, la integridad del ARN<sub>m</sub> usando los métodos de la invención es bastante alta dado que los métodos producen muy pocas transcripciones de cadena doble o ninguna, lo que requeriría la eliminación usando procedimientos tales como la digestión de RNasa.

El ARN producido mediante los procesos descritos en la presente es cualquier ARN con una longitud mayor que 30 nucleótidos que se puede utilizar para propósitos terapéuticos o de diagnóstico. En algunas realizaciones, el ARN es un ARN de más de 40, 50, 60, 75, 100, 200, 300, 400, 500, o 1000 nucleótidos de longitud. En algunas realizaciones, el ARN es un ARN de más de 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10.000, 11.000 o 12.000 nucleótidos de longitud. En algunas realizaciones, el ARN es ARN<sub>m</sub>. En algunas realizaciones, el ARN es un ARN de alrededor de 500 a alrededor de 4000 nucleótidos de longitud, 1000 alrededor de a alrededor de 2000 nucleótidos de longitud, 750 alrededor de a alrededor de 1800 nucleótidos de longitud, alrededor de 1500 a alrededor de 3000 nucleótidos de longitud, alrededor de 4000 a alrededor de 7000 nucleótidos de longitud, o alrededor de 6000 a alrededor de 12.000 nucleótidos de longitud. El ARN<sub>m</sub> puede estar modificado o no modificado. En otras realizaciones, el ARN es uno o más de los siguientes: ARN<sub>m</sub>, ARN<sub>m</sub> modificado, ARN no modificado, ARN<sub>inc</sub>, ARN de autorreplicación, ARN circular, ARN guía CRISPR.

Las reacciones de IVT tradicionales se realizan al incubar una plantilla de ADN con una ARN polimerasa y cantidades equimolares de trifosfatos de nucleótidos, que incluyen GTP, ATP, CTP, y UTP en un amortiguador de transcripción. Una transcripción de ARN que tiene un trifosfato de guanosina en el extremo 5' se produce a partir de esta reacción. Estas reacciones también dan como resultado la producción de una cantidad de impurezas tales como ARN de cadena doble y de cadena simple que son inmunoestimuladores y pueden tener un impacto aditivo. Los métodos de la invención que evitan la formación de complementos inversos evitan el reconocimiento inmunitario de ambas especies. En algunas realizaciones, los métodos de la invención dan como resultado la producción de ARN que tiene actividad de células T reducida de manera significativa en comparación con una preparación de ARN elaborada usando métodos de la técnica previa con NTP equimolares. La técnica previa intenta eliminar estos componentes no deseados usando una serie de etapas de purificación posteriores. Dichos métodos de purificación no son convenientes ya que implican tiempo y recursos adicionales y también dan como recursos la incorporación de solventes orgánicos residuales en el producto final, lo que no es conveniente para un producto farmacéutico. Es una tarea intensiva en cuanto al trabajo y el capital aumentar procesos como la cromatografía de fase inversa (RP): utilizando, por ejemplo, instalaciones a prueba de explosiones, columnas de HPLC y sistemas de purificación clasificados para alta

presión, alta temperatura, solventes inflamables, etc. El aumento y el rendimiento para la fabricación a gran escala están limitados por estos factores. La purificación posterior también es necesaria para eliminar el par de iones de alquilamonio utilizado en el proceso de RP. Por el contrario, los métodos descritos en la presente incluso mejoran los métodos utilizados actualmente (por ejemplo, RP). Una carga inferior de impurezas da lugar a mayor recuperación de purificación de ARN de longitud completa sin contaminantes inductores de citocinas, por ejemplo, mayor calidad de los materiales en el inicio. Una ventaja adicional de los procesos de IVT modificados de la invención, cuando se usa RNasa III como una purificación preparativa, es que dado que hay menos sustrato de RNasa III, se generan menos productos de escisión inertes/extraños (esos que se degradan pero no se traducen) mediante el tratamiento con RNasa III. Si solo hay cantidades mínimas del sustrato de ARNcd/RNasa III, aunque pueda ser silenciosa en citocinas, hay más producto de ARN intacto final (casquete intacto/ORF/PolíA) presente capaz de traducir proteínas. Esto da lugar a una carga reducida para cualquier purificación posterior.

Se descubrió de manera sorprendente, de acuerdo con aspectos de la invención, que la manipulación de uno o más de los parámetros de reacción en la reacción de IVT produce una preparación de ARN de ARN altamente funcional sin uno o más contaminantes no deseados producidos usando los procesos de la técnica previa. Un parámetro en la reacción de IVT que se puede manipular es la cantidad relativa de un nucleótido o análogo de nucleótido en comparación con uno o más nucleótidos o análogos de nucleótido distintos en la mezcla de reacción (por ejemplo, concentración o cantidades de nucleótidos dispares). Por ejemplo, la reacción de IVT puede incluir un exceso de nucleótidos, por ejemplo, monofosfato de nucleótido, difosfato de nucleótido o trifosfato de nucleótido y/o un exceso de análogos de nucleótido y/o análogos de nucleósido. Los métodos de la invención producen un producto de alto rendimiento que es significativamente más puro que los productos producidos mediante métodos de IVT tradicionales.

Los análogos de nucleótidos son compuestos que tienen la estructura general de un nucleótido o son similares de manera estructural a un nucleótido o parte de este. En particular, los análogos de nucleótidos son nucleótidos que contienen, por ejemplo, un análogo de la parte de ácido nucleico, parte de azúcar y/o grupos fosfato del nucleótido. Los nucleótidos incluyen, por ejemplo, monofosfatos de nucleótido, difosfatos de nucleótido, y trifosfatos de nucleótido. Un análogo de nucleótido, tal como se usa en la presente, es similar de manera estructural a un nucleótido o parte de este pero no tiene la estructura de nucleótido típica (nucleobase-ribosa-fosfato). Los análogos de nucleósidos son compuestos que tienen la estructura general de un nucleósido o son similares de manera estructural a un nucleósido o parte de este. En particular, los análogos de nucleósidos son nucleósidos que contienen, por ejemplo, un análogo de la parte de ácido nucleico y/o parte de azúcar del nucleósido.

Un trifosfato de nucleótido, tal como se usa en la presente, se refiere a una molécula que incluye una nucleobase unida a una ribosa (es decir, nucleósido) y tres fosfatos (es decir, nucleótido). Un difosfato de nucleótido se refiere a la misma molécula, pero que tiene dos restos de fosfato. Un monofosfato de nucleótido se refiere a la misma molécula, pero que tiene un resto de fosfato. El monofosfato de nucleótido, difosfato de nucleótido y trifosfato de nucleótido a menudo se denominan en la presente NMP, NDP y NTP, respectivamente. La N en NMP, NDP y NTP hace referencia a cualquier nucleótido, que incluye nucleótidos de origen natural, nucleótidos sintéticos, y nucleótidos modificados. Por lo tanto, los términos NDP y NTP se refieren a difosfatos de nucleótido y trifosfatos de nucleótido, respectivamente, que tienen cualquier nucleótido de origen natural, sintético o modificado en su interior.

Los difosfatos de nucleótido naturales incluyen al menos difosfato de adenosina (ADP), difosfato de guanosina (GDP), difosfato de citidina (CDP), y difosfato de uridina (UDP). Los trifosfatos de nucleótido naturales incluyen al menos trifosfato de adenosina (ATP), trifosfato de guanosina (GTP), trifosfato de citidina (CTP), trifosfato de 5-metiluridina (m5UTP) y trifosfato de uridina (UTP). En algunas realizaciones, el NDP y/o NTP están modificados. Por ejemplo, el NDP o NTP modificados pueden tener un mango para facilitar la purificación y aislamiento.

Los trifosfatos de nucleótido se agregan a la hebra de ARN mediante una polimerasa tal como T7 polimerasa. Los difosfatos y monofosfatos de nucleótido, por el contrario pueden iniciar la reacción (por ejemplo, servir como el primer monómero transcrito) pero no se incorporarán dentro de la hebra mediante T7 polimerasa (por ejemplo, no se incorporarán en ningún lugar en la hebra). En algunos casos, los difosfatos de nucleótido, tal como GDP, pueden incorporarse como el primer monómero. Por ejemplo, si T7 se inicia con GDP y produce un 5'GDP, se puede generar un ARN funcional. El ARN iniciado con 5' GDP aún es un sustrato para la enzima que coloca casquetes de vacuna. Cuando un exceso de NMP, tal como GMP, se utiliza en la reacción, la pureza puede mejorarse al ligar un casquete, dado que el producto transcrito con 5' PO<sub>4</sub> es un sustrato para ligasas) (por ejemplo, ligasas de ADN/ARN).

Los análogos de nucleótido útiles en la invención son estructuralmente similares a nucleótidos o partes de estos pero, por ejemplo, no son polimerizables mediante T7. Los análogos de nucleótido/nucleósido tal como se usan en la presente (que incluyen análogos de C, T, A, U, G, dC, dT, dA, dU, o dG) incluyen, por ejemplo, análogos de nucleótido antivirales, análogos de fosfato (solubles o inmovilizados, hidrolizables o no

hidrolizables), dinucleótido, trinucleótido, tetranucleótido, por ejemplo, un análogo de casquete, o un precursor/sustrato para colocación enzimática de casquetes (vacuna, o ligasa), un nucleótido etiquetado con un grupo funcional para facilitar la ligación/conjugación del casquete o resto 5' (IRES), un nucleótido etiquetado con un 5' PO<sub>4</sub> para facilitar la ligación del casquete o resto 5', o un nucleótido etiquetado con un grupo funcional/grupo protector que pueda escindirse de forma química o enzimática. Los análogos de nucleótido/nucleósido antivirales incluyen, de modo no taxativo, Ganciclovir, Entecavir, Telbivudine, Vidarabine y Cidofovir.

#### Condiciones de reacción de IVT

En aspectos de ejemplo, los métodos de la invención implican la producción de ARN a través de una reacción de IVT. La IVT es un método reconocido en la técnica utilizado para generar polinucleótidos sintéticos *in vitro*. El ARN transcrito *in vitro* (IVT) puede modificarse genéticamente para que exprese de manera transitoria proteínas al parecerse estructuralmente al ARN natural. Sin embargo, hay desafíos inherentes de esta clase de fármacos, particularmente relacionados con el control de la eficacia de traducción y la inmunogenicidad del ARN IVT. En particular, el ARN IVT produce efectos inmunitarios innatos no deseados y hay procedimientos de purificación muy rigurosos mediante HPLC que se aplican típicamente como una etapa de purificación de ARN final adicional. La eliminación de cantidades menores de fragmentos cortos de ARN de cadena doble es críticamente importante para lograr esta respuesta inmunitaria más reducida.

La reacción típica usada en la técnica previa proporciona un producto de alta fidelidad y alto rendimiento razonable. Sin embargo, el producto tiene un nivel de referencia de contaminantes, de los cuales solo se pueden eliminar algunos usando métodos de purificación de rutina. La reacción de IVT típicamente incluye lo siguiente: una ARN polimerasa, por ejemplo, una ARN polimerasa T7 a una concentración final de, por ejemplo, 1000-12000 U/ml, por ejemplo, 7000 U/ml; la plantilla de ADN a una concentración final de, por ejemplo, 10-70 nM, por ejemplo, 40 nM; nucleótidos (NTP) a una concentración final de, por ejemplo, 0.5-10 mM, por ejemplo, 7.5 mM cada uno; magnesio a una concentración final de, por ejemplo, 12-60 mM, por ejemplo, acetato de magnesio a 40 mM; un amortiguador tal como, por ejemplo, HEPES o Tris a un pH de, por ejemplo, 7-8.5, por ejemplo, HCl de Tris 40 mM, pH 8. En algunas realizaciones se puede incluir ditioneitol (DTT) 5 mM y/o espermidina 1 mM. En algunas realizaciones, se incluye un inhibidor de RNasa en la reacción de IVT para garantizar que no haya degradación inducida por RNasa durante la reacción de transcripción. Por ejemplo, se puede utilizar el inhibidor de RNasa murino a una concentración final de 1000 U/ml. En algunas realizaciones se incluye una pirofosfatasa en la reacción de IVT para escindir la pirofosfatasa inorgánica generada después de cada incorporación de nucleótidos en dos unidades de fosfato inorgánico. Esto garantiza que el magnesio permanezca en solución y no precipite como pirofosfato de magnesio. Por ejemplo, se puede utilizar una pirofosfatasa inorgánica de *E. coli* a una concentración final de 1 U/ml.

Una reacción de transcripción *in vitro* típica incluye lo siguiente:

40	1	Plantilla de ADNc	1,0 µg
	2	10x solución amortiguadora de transcripción (Tris-HCl 400 mM pH 8.0, MgCl <sub>2</sub> 190 mM, TCEP o DTT 50 mM, espermidina 10 mM)	2,0 µl
	3	NTP a la medida (cada uno 25 mM)	7,2 µl
	4	Inhibidor RNasa	20 U
45	5	ARN polimerasa T7	3000 U
	6	dH <sub>2</sub> O	hasta 20,0 µl y
	7	Incubación a 37 °C durante 1h-5h.	

La mezcla de IVT bruta se puede almacenar a 4 °C durante 4-12 horas. Luego se utiliza una unidad de DNasa libre de RNasa para digerir la plantilla original. Luego de 15 minutos de incubación a 37 °C, se purifica el ARN usando técnicas de purificación tal como resina dT, HPLC de fase inversa, el kit MEGACLEAR™ de Ambion (Austin, TX) conforme a las instrucciones del fabricante. La reacción de IVT de ejemplo no es taxativa en términos de componentes o cantidades de componentes usados.

De manera similar a los métodos tradicionales, el ARN producido por los métodos de la invención también se puede producir al formar una mezcla de reacción que comprende una plantilla de ADN, y uno o más NTP tales como ATP, CTP, UTP, GTP y un amortiguador. Luego la reacción se incuba en condiciones tales que el ARN se transcriba. Sin embargo, los métodos de la invención implican el sorprendente hallazgo de que la presencia de una cantidad en exceso de uno o más nucleótidos y/o análogos de nucleótidos puede tener un impacto significativo en el producto final. Los métodos de la invención se pueden utilizar para producir producto de alta calidad sin impurezas no pretendidas o no deseadas y sin afectar la eficacia de la reacción.

Los métodos de IVT de la invención implican una modificación en la cantidad (por ejemplo, cantidad molar) de nucleótidos en la mezcla de reacción. En algunos aspectos, se puede agregar uno o más nucleótidos y/o uno o más análogos de nucleótido en exceso a la mezcla de reacción. Un exceso de nucleótidos y/o análogos de nucleótido es cualquier cantidad mayor que la cantidad de uno o más de los otros nucleótidos tales como los

NTP en la mezcla de reacción. Por ejemplo, un exceso de un nucleótido y/o análogo de nucleótido puede ser una cantidad mayor que la cantidad de cada uno o al menos uno de los otros NTP individuales en la mezcla de reacción o puede hacer referencia a una cantidad mayor que cantidades equimolares de los otros NTP. La presente invención se expone en la reivindicación 1. En una realización de la invención, la concentración de GTP es al menos 2 veces mayor que la concentración de cada uno de ATP, CTP y UTP.

En algunas realizaciones, el GTP está presente en la mezcla en una relación de al menos 2:1, al menos 3:1, al menos 4:1, al menos 5:1, al menos 6:1, al menos 7:1, al menos 8:1, al menos 9:1, al menos 10:1, al menos 11:1, al menos 12:1, al menos 13:1, al menos 14:1, o al menos 15:1 con respecto a la concentración de cualquiera de los ATP, CTP, o UTP. La relación de GTP con respecto a otros NTP puede ser de alrededor de 2:1 a alrededor de 3:1, de alrededor de 2.5:1 a alrededor de 3.5:1, de alrededor de 3:1 a alrededor de 4:1, de alrededor de 3.5:1 a alrededor de 4.5:1, de alrededor de 4:1 a alrededor de 5:1, de alrededor de 4.5:1 a alrededor de 5.5:1, de alrededor de 5:1 a alrededor de 6:1, de alrededor de 5.5:1 a alrededor de 6.5:1, de alrededor de 6:1 a 7:1, de alrededor de 6.5:1 a alrededor de 7.5:1, de alrededor de 7:1 a alrededor de 8:1, de alrededor de 7.5:1 a alrededor de 8.5:1, de alrededor de 8:1 a alrededor de 9:1, de alrededor de 8.5:1 a alrededor de 9.5:1, y de alrededor de 9:1 a alrededor de 10:1. En una realización, la relación de concentración de GTP con respecto a la concentración de ATP, CTP, y UTP puede ser 2:1, 4:1, y 4:1, respectivamente. En otra realización, la relación de concentración de GTP con respecto a la concentración de ATP, CTP, y UTP puede ser 3:1, 6:1, y 6:1, respectivamente.

La presente invención se expone en la reivindicación 1. En una realización de la presente invención, la mezcla de reacción también comprende GDP y la concentración de GDP es al menos 2 veces mayor que la concentración de cada uno de ATP, CTP y UTP

En algunas realizaciones, el límite superior del exceso de nucleótido o análogo de nucleótido en la mezcla de reacción está regido por el límite de solubilidad.

En algunas realizaciones, los NTP son NTP de sal. Por ejemplo, los NTP pueden ser NTP de amonio, tris NTP, NTP de litio, NTP de potasio, o NTP de sodio.

La presente invención se expone en la reivindicación 1. En una realización de la presente invención, la mezcla de reacción comprende además GDP y la relación entre la concentración de GTP más GDP y la concentración de cada uno de ATP, CTP y UTP es de al menos 2:1.

En la invención, el NTP y NDP son, por ejemplo, GTP y GDP, respectivamente, y pueden estar presentes en la mezcla en una concentración al menos 6 veces (x), 7x, 8x, 9x, 10x, 11x, 12x, 13x, 14x, 15x, 15x-100x, 10x-90x, 10x-80x, 10x-70x o incluso mayor que la cantidad de cualquiera de ATP, CTP, o UTP en la mezcla de reacción. En realizaciones de ejemplo, la mezcla combinada de NTP y NDP puede agregarse a una relación molar, por ejemplo, 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 11:1, 12:1, 13:1, 14:1, 15:1, o más en la mezcla de reacción. La relación de GTP y GDP con respecto a otros NTP puede ser de alrededor de 2:1 a alrededor de 3:1, de alrededor de 2.5:1 a alrededor de 3.5:1, de alrededor de 3:1 a alrededor de 4:1, de alrededor de 3.5:1 a alrededor de 4.5:1, de alrededor de 4:1 a alrededor de 5:1, de alrededor de 4.5:1 a alrededor de 5.5:1, y de alrededor de 5:1 a alrededor de 6:1. En una realización, la relación de concentración de GTP con respecto a la concentración de ATP, CTP, y UTP puede ser 3:1, 6:1, y 6:1, respectivamente.

En algunas realizaciones, la relación de GTP con respecto a GDP se considera con respecto a la relación de nucleótido de purina con respecto a nucleótido de pirimidina (Pu:Py) en la mezcla de reacción. En algunas realizaciones, las relaciones de GTP:GDP con respecto a Pu:Py son 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 11:1, 12:1, 13:1, 14:1, 15:1, o más en la mezcla de reacción.

En algunas realizaciones, el amortiguador contiene fosfato. La concentración eficaz de fosfato es al menos 150 mM, al menos 160 mM, al menos 170 mM, al menos 180 mM, al menos 190 mM, al menos 200 mM, al menos 210 mM, al menos 220 mM, o al menos 230 mM de fosfato. En una realización, la concentración eficaz de fosfato es 180 mM. En otra realización, la concentración eficaz de fosfato es 195 mM.

En la invención, el amortiguador contiene magnesio. El amortiguador puede tener una relación de la concentración de ATP más CTP más UTP más GTP y opcionalmente, GDP con respecto a la concentración molar de  $Mg^{2+}$  de al menos 1,0, al menos 1,1, al menos 1,2, al menos 1,25, al menos 1,3, al menos 1,4, al menos 1,5, al menos 1,6, al menos 1,7, al menos 1,75, al menos 1,8, y al menos 1,85 o 3. En otras realizaciones, la relación es 1,0, 1,1, 1,2, 1,25, 1,3, 1,4, 1,5, 1,6, 1,7, 1,75, 1,8, 1,85, 2, 2,1, 2,2, 2,3, 2,4, 2,5, 2,6, 2,7, 2,8, 2,9, o 3 o cualquier intervalo de estas variables. En una realización, la relación es 1,5. En otra realización, la relación es 1,88. En una realización, la relación es 3.

La reacción de IVT (mezcla de reacción) incluye una ARN polimerasa, es decir, T7 o T3. En algunas realizaciones, la polimerasa, por ejemplo, T7 polimerasa se incluye a una concentración mayor que 5 U/μl, mayor que 10 U/μl, mayor que 20 U/μl, mayor que 50 U/μl, o mayor que 100 U/μl. En algunas realizaciones, la



concentración de polimerasa, por ejemplo, T7 polimerasa, se encuentra en un intervalo de desde alrededor de 1 a alrededor de 250 U/μl de la mezcla de reacción, por ejemplo, de alrededor de 1 a alrededor de 100 U/μl o de alrededor de 100 a alrededor de 250 U/μl. En algunas realizaciones, la concentración de T7 polimerasa se encuentra en un intervalo de desde alrededor de 30 a alrededor de 60 U/μl, de alrededor de 60 a alrededor de 80 U/μl, de alrededor de 80 a alrededor de 100 U/μl, de alrededor de 100 a alrededor de 150 U/μl o desde alrededor de 150 a alrededor de 200 U/μl. En algunas realizaciones, la polimerasa, por ejemplo, T7 polimerasa se incluye a una concentración de 7, 14, 25, 50, 75, o 140.

Tal como se usa en la presente, una plantilla de ADN se refiere a una plantilla de polinucleótidos para ARN polimerasa. La plantilla de ADN útil de acuerdo con los métodos descritos en la presente incluye, en algunas realizaciones, un gen de interés que codifica, por ejemplo, un polipéptido de interés. La plantilla de ADN en algunas realizaciones incluye un promotor de ARN polimerasa, por ejemplo, un promotor de T7 ubicado en 5' y unido de manera operativa al gen de interés y opcionalmente una secuencia que codifica una cola de poli A ubicada en 3' en el gen de interés.

Las ARN polimerasas conocidas en la técnica pueden utilizarse en los métodos de la presente invención. Las ARN polimerasas incluyen una ARN polimerasa de fagos, es decir, una ARN polimerasa T7 o una ARN polimerasa T3. Como ejemplo no taxativo, la ARN polimerasa puede estar modificada para que exhiba una mejor capacidad para incorporar un trifosfato de nucleótido modificado en 2' en comparación con una ARN polimerasa no modificada.

Tal como se usa en la presente, "gen de interés" se refiere a un polinucleótido que codifica un polinucleótido o proteína de interés. Dependiendo del contexto, el gen de interés se refiere a un ácido ribonucleico, por ejemplo, un gen de interés en una plantilla de ADN que puede transcribirse en una transcripción de ARN, o un ácido ribonucleico, por ejemplo, un gen de interés en una transcripción de ARN que puede traducirse para producir el polipéptido codificado de interés in vitro, in vivo, in situ o ex vivo. Un polipéptido de interés incluye, de modo no taxativo, elementos biológicos, anticuerpos, vacunas, proteínas o péptidos terapéuticos, etc.

Una "transcripción de ARN" se refiere a un ácido ribonucleico producido mediante una reacción de IVT usando una plantilla de ADN y una ARN polimerasa. En algunas realizaciones, la transcripción de ARN es un ARNm y típicamente incluyen la secuencia de codificación para un gen de interés y una cola de poli A. La transcripción de ARN incluye un ARNm. La transcripción de ARN puede incluir modificaciones, por ejemplo, nucleótidos modificados. Tal como se usa en la presente, el término transcripción de ARN incluye y se usa de manera intercambiable con ARNm, ARNm modificado, "ARNmm" o ARNm modificado, y construcción principal.

### Pureza

El ARN producido de acuerdo con los métodos de la invención es sorpresivamente puro y de alta integridad. Tiene menos contaminantes que las preparaciones de ARN producidas de acuerdo con métodos de IVT tradicionales. En algunas realizaciones, tiene menos contaminantes inmunitarios que las preparaciones de ARN producidas de acuerdo con métodos de IVT tradicionales. Los contaminantes son fragmentos de ARN producidos por la reacción diferentes al ARN deseado. En algunas realizaciones, los contaminantes de fragmento de ARN son productos de transcripción de complemento inverso y/o contaminantes de ARN inductores de citocinas. En otras realizaciones, los contaminantes de fragmento de ARN son ARN de cadena doble o contaminantes inmunogénicos.

En algunas realizaciones, las preparaciones de ARN producidas por los métodos de la invención, tienen menos contaminantes que las preparaciones de ARN producidas de acuerdo con métodos de IVT tradicionales. En algunas realizaciones, las preparaciones de ARN de la invención tienen al menos 10 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % menos contaminantes que las preparaciones de ARN producidas de acuerdo con métodos de IVT tradicionales. En otras realizaciones, las preparaciones de ARN producidas por los métodos de la invención están sustancialmente libres de contaminantes. En otras realizaciones, las preparaciones de ARN de la invención están 100 % libres de contaminantes.

Por lo tanto, la invención en algunos aspectos implica la preparación de un ARN IVT que está sustancialmente libre de producto de transcripción de complemento inverso sin la necesidad de etapas de purificación adicionales. Tal como se usa en la presente, el término "producto de transcripción de complemento inverso" se refiere a una molécula de ARN que resulta de la transcripción con plantilla de ARN. En algunas realizaciones, el producto de transcripción de complemento inverso puede ser un producto de transcripción con plantilla de ARN. Sin pretender limitarse a la teoría, se cree que el producto de complemento inverso es predominantemente o en su totalidad producto de transcripción con plantilla de ARN. Si el producto de complemento inverso estuviera compuesto por producto de transcripción con plantilla de ADN, el producto incluiría, por ejemplo, secuencias de nucleótidos complementarias a la región del promotor T7 de la plantilla de ADN. Los productos de complemento inverso caracterizados a la fecha están predominantemente libres de secuencias complementarias, por ejemplo, a la región del promotor T7. En algunas realizaciones, el producto

de transcripción de complemento inverso es un ARN de cadena doble (ARNcd), que puede incluir una hebra que codifica una secuencia que es un complemento inverso de al menos una parte del ARN IVT. En otras realizaciones, el producto de transcripción de complemento inverso puede ser un ARNcd con una hebra que comprende una secuencia de poliU. Ya sea la hebra de complemento inverso que codifica el polipéptido de interés o la hebra que codifica la secuencia de poliU puede iniciar con 5' trifosfato (5'ppp). El producto de transcripción con plantilla de ARN puede incluir un complemento inverso del extremo 5' del ARN IVT y/o un complemento inverso del extremo 3' del ARN IVT. Además, el complemento inverso del extremo 5' del ARN IVT puede ser complementario a la totalidad o a una parte de una UTR 5' del ARN IVT. El complemento inverso puede comprender una secuencia complementaria a los primeros 10-15, los primeros 5-15, los primeros 5-20, los primeros 10-20, los primeros 15-25 nucleótidos de la UTR 5'. Asimismo, el complemento inverso del extremo 3' del ARN IVT puede ser complementario a la totalidad o a una parte de la cola de poliA del ARN IVT. El producto de transcripción de complemento inverso puede someterse a plantilla en cualquier parte del ARN y, por lo tanto, puede ser de cualquier tamaño o complementario a cualquier ubicación en la plantilla. Por ejemplo, el producto de complemento inverso puede ser un 5-mer 10-mer 15-mer 20-mer 25-mer 40-mer 50-mer 60-mer 70-mer, 100-mer, 200-mer, etc., hasta la longitud completa del producto pretendido o deseado.

La presente descripción presenta composiciones que comprenden un ARN IVT y un excipiente farmacéuticamente aceptable sustancialmente libre de producto de transcripción de complemento inverso. En algunas realizaciones, el ARN IVT que está sustancialmente libre de producto de transcripción de complemento inverso incluye producto de transcripción de complemento inverso que conforma menos de alrededor de 10 %, 9 %, 8 %, 7 %, 6 %, 5 %, 4,5 %, 4 %, 3,5 %, 3 %, 2,5 %, 2 %, 1,5 %, 1 %, 0,9 %, 0,8 %, 0,7 %, 0,6 % o 0,55 %, 0,5 %, 0,45 %, 0,4 %, 0,35 %, 0,3 %, 0,025 %, 0,2 %, 0,15 %, 0,1 %, 0,05 %, 0,01 %, 0,005 % o 0,001 % de la masa del ARN total. La masa de la composición de ARN se puede determinar mediante cualquier medio conocido en la técnica. Los ejemplos de métodos para determinar la masa del ARN incluyen cromatografía líquida y espectrometría de masas.

Sin pretender limitarse a la teoría, se cree que en algunas realizaciones de la invención los contaminantes son complementos inversos de cadena simple unidos a una población de ARN IVT que forman una estructura de cadena doble en el contexto de ARN más largos. T7 puede formar una plantilla de un ARN naciente (hebra de sentido) fallido así como un ARN naciente de producto de longitud completa. La transcripción con plantilla de ARN (antisentido), una vez transcrita (que inicia predominantemente con pppC, pppU, y pppA), supuestamente permanece asociada al ARN sentido naciente. De manera alternativa, si se forman dichas construcciones pueden ser intrínsecamente silenciosas. La asociación de las 2 hebras de manera eficaz es ARNcd con 5'ppp. La presencia de 5'ppp en una o más de las hebras hibridadas, hace que la estructura sea inmunoestimuladora. Incluso cuando la transcripción con plantilla de ARN antisentido esté disociada del ARN, la presencia de ARNcs con pppC, pppU, y pppA aún es inductora de citocinas. Dado que los métodos de la invención producen ARN sin ARNcd tal como se ve en ELISA J2 y tratamiento con RNasa III, los productos no asumirán una estructura de ARN grande de longitud completa. Es probable que el ARN se pliegue de manera similar si se transcribe con el proceso equimolar o el proceso de IVT de la invención.

En algunas realizaciones, las preparaciones de ARN producidas por los métodos de la invención están sustancialmente libres de contaminante de ARN inductor de citocinas. Tal como se usa en la presente, el término "contaminante inductor de citocinas" se refiere a una molécula de ARN que induce la generación de citocinas, por ejemplo, interferón tipo I (inducción de IFN $\alpha/\beta$ ), por ejemplo, tal como se determina en un ensayo de inducción de citocinas basado en células, por ejemplo, tal como se determinar en un ensayo de fibroblastos BJ/IFNbeta y/o ensayo Luminex tal como se describe en los ejemplos de trabajo de la presente memoria descriptiva. En aspectos de ejemplo de la invención, el término contaminante "inductor de citocinas" se refiere a una molécula de ARN que induce la inducción de citocinas y que tiene naturaleza sustancialmente de cadena doble.

Sin pretender limitarse a la teoría, se cree que las moléculas de ARN de cadena doble que son el resultado de transcripción de polimerasa aberrante, por ejemplo, transcripción con plantilla del ARN deseado producido en una reacción de IVT, inducen citocinas mediante la activación de una respuesta inmunitaria innata similar a la respuesta inmunitaria antiviral natural e incluye dos tipos de receptores de reconocimiento de patógenos (PRR): los receptores tipo Toll (TLR) y los receptores parecido a RIG-I (RLR), por ejemplo, receptor tipo toll 3 (TLR3), así como también helicasas de ARN, por ejemplo, RIG-I y MDA5. Los ejemplos de otras moléculas inductoras de citocinas incluyen sustratos de RNasa III. Un sustrato de RNasa III, tal como se usa en la presente, se refiere a una molécula de ARN de cadena doble que es susceptible a escisión mediante una enzima RNasa III

En algunas realizaciones, el contaminante de ARN inductor de citocinas puede ser un ARN de cadena doble con una secuencia inversa complementaria al ARN IVT o una secuencia de poliU. El complemento inverso del ARN IVT o la secuencia de poliU puede iniciar con un 5'ppp.

El contaminante de ARN inductor de citocinas puede incluir un complemento inverso del extremo 5' del ARN IVT y/o un complemento inverso del extremo 3' del ARN IVT. Además, el complemento inverso del extremo 5' del ARN IVT puede ser complementario a la totalidad o a una parte de una UTR 5' del ARN IVT. El complemento

inverso puede comprender una secuencia complementaria a los primeros 10-15, los primeros 5-15, los primeros 5-20, los primeros 10-20, los primeros 15-25 nucleótidos de la UTR 5'. En algunas realizaciones, el complemento inverso puede comprender una secuencia complementaria a un intervalo de 1-20, 1-30, 1-40, 1-50, 1-60, 1-70, 1-80, 1-90, 1-100, 1-200, 1-300, 1-400, 1-500, 1-600, 1-700, 1-800, 1-900, 1-1000, 1-2000, 1-2500, o 1-3000 nucleótidos de longitud dentro de la UTR 5'. En otras realizaciones, el complemento inverso puede comprender una secuencia complementaria a un intervalo de 10-20, 10-30, 10-40, 10-50, 1-60, 10-70, 10-80, 10-90, 10-100, 10-200, 10-300, 10-400, 10-500, 10-600, 10-700, 10-800, 10-900, 10-1000, 10-2000, 10-2500, o 10-3000 nucleótidos de longitud dentro de la UTR 5'. En otras realizaciones, el complemento inverso puede comprender una secuencia complementaria a un intervalo de 20-25, 20-30, 20-40, 20-50, 20-60, 20-70, 20-80, 20-90, 20-100, 20-200, 20-300, 20-400, 20-500, 20-600, 20-700, 20-800, 20-900, 20-1000, 20-2000, 20-2500, o 20-3000 nucleótidos de longitud dentro de la UTR 5'. Asimismo, el complemento inverso del extremo 3' del ARN IVT puede ser complementario a la totalidad o a una parte de la cola de poliA del ARN IVT. El complemento inverso puede comprender una secuencia complementaria a los primeros 10-15, los primeros 5-15, los primeros 5-20, los primeros 10-20, los primeros 15-25 nucleótidos de la UTR 3'. En algunas realizaciones, el complemento inverso puede comprender una secuencia complementaria a un intervalo de 1-20, 1-30, 1-40, 1-50, 1-60, 1-70, 1-80, 1-90, 1-100, 1-200, 1-300, 1-400, 1-500, 1-600, 1-700, 1-800, 1-900, 1-1000, 1-2000, 1-2500, 1-3000, o 1-longitud completa o tamaño máximo de los nucleótidos de ARN de longitud dentro de la UTR 3'. En otras realizaciones, el complemento inverso puede comprender una secuencia complementaria a un intervalo de 10-20, 10-30, 10-40, 10-50, 1-60, 10-70, 10-80, 10-90, 10-100, 10-200, 10-300, 10-400, 10-500, 10-600, 10-700, 10-800, 10-900, 10-1000, 10-2000, 10-2500, o 10-3000 nucleótidos de longitud dentro de la UTR 3'. En otras realizaciones, el complemento inverso puede comprender una secuencia complementaria a un intervalo de 20-25, 20-30, 20-40, 20-50, 20-60, 20-70, 20-80, 20-90, 20-100, 20-200, 20-300, 20-400, 20-500, 20-600, 20-700, 20-800, 20-900, 20-1000, 20-2000, 20-2500, o 20-3000 nucleótidos de longitud dentro de la UTR 3'.

La presente descripción incluye una composición que comprende un ARN IVT y un excipiente farmacéuticamente aceptable sustancialmente libre de contaminante de ARN inductor de citocinas. En algunas realizaciones, el contaminante de ARN inductor de citocinas conforma menos de 0,5 %, 0,45 %, 0,4 %, 0,35 %, 0,3 %, 0,25 %, 0,2 %, 0,15 %, 0,1 %, 0,05 %, 0,01 %, 0,005 % o 0,001 % de la masa del ARN. La masa de la composición de ARN se puede determinar mediante cualquier medio conocido en la técnica. Los ejemplos incluyen cromatografía líquida y espectrometría de masas.

El ARNcd del contaminante tal como el contaminante de ARN inductor de citocinas y/o el producto de transcripción de complemento inverso puede tener una longitud de 20 a 50 nucleótidos. En otras realizaciones, el ARNcd puede tener una longitud de 20-25, 25-30, 30-35, 35-40, 40-45, 45-50, 50-55, 55-60, 60-65, 65-70, 70-75, 75-80, 80-85, 85-90, 90-95, 95-100, 100-110, 110-120, 120-130, 130-140, 140-150, 150-160, 160-170, 170-180, 180-190, 190-200, 200-225, 225-250, 250-275, 275-300, 300-325, 325-350, 350-375, 375-400, 400-425, 425-450, 450-475, 475-500, 500-550, 550-600, 600-650, 650-700, 700-750, 750-800, 800-850, 850-900, 900-950, y 950-1000 nucleótidos. En algunas realizaciones, la masa del ARNcd es mayor que 40 pares de base y conforma menos de alrededor de 0,5 % de la composición de ARN.

Las hebras del contaminante pueden tener extremos 5'ppp. En algunas realizaciones, las hebras contaminantes pueden tener una abundancia menor de pppA, pppC, y pppU, en comparación con ARN producido mediante proceso equimolar. En otras realizaciones, las hebras contaminantes pueden tener relaciones menores de pppA:pppG, pppC:pppG, y/o pppU:pppG, en comparación con el proceso equimolar. Los pppNTP pueden detectarse mediante LC-MS después de la digestión total de nucleasas, por ejemplo, tratamiento con nucleasa P1. La nucleasa P1 digiere ARN y ADN en nucleótidos simples. Las únicas especies de trifosfato que deberían estar presentes son para los nucleótidos iniciadores. Si no se forman productos de transcripción con plantilla de ARN, 5'PPPG es el único trifosfato que debería estar presente dado que este es el único sitio de iniciación objetivo. La presencia y abundancia de 5'pppA, 5'pppC, y/o 5'pppU tal como se detecta mediante LC/MS después de la digestión de nucleasa P1 indican transcripciones de ARN con plantilla de ARN.

Además de tener menos impurezas, particularmente impurezas de cadena doble, las composiciones de ARN IVT tienen una alta proporción de transcripción de ARN funcional de longitud completa con respecto a otras especies de ARN en la composición, particularmente en comparación con composiciones de ARN purificadas tradicionales producidas usando métodos de IVT combinados con etapas de purificación tales como cromatografía de fase inversa o tratamiento con RNasa III. En algunas realizaciones, más de alrededor de 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, 99,5 % o 99,8 % de la masa del ARN comprende transcripciones de longitud completa de cadena simple. Además de las transcripciones de longitud completa de cadena simple, la composición de ARN IVT puede incluir otras especies de ARN de cadena simple tales como transcripciones de ARN parcial de cadena simple en orientación sentido, que incluyen transcripciones fallidas. La composición de ARN IVT, sin embargo, está sustancialmente libre de fragmentos insensibles a RNasa III.

Las composiciones de ARN descritas en la presente pueden incluir otros componentes además de la

transcripción de ARN de longitud completa, por ejemplo, transcripciones truncadas y/o transcripciones mal formadas. Por ejemplo, el ARN puede incluir transcripciones de diferentes longitudes, por ejemplo, más cortas o más largas que la transcripción de longitud completa. Por lo tanto, en algunas realizaciones, la preparación de ARN producida por los métodos de la invención incluye transcripciones truncadas y/o fallidas. La ARN polimerasa se une a un promotor de ADN y sintetiza las transcripciones de ARNm cortas. Tal como se usa en la presente, el término "transcripción truncada" se refiere a transcripciones que tienen identidad con respecto al ARN IVT, pero que tienen longitud insuficiente y no presentan todos los elementos necesarios para codificar el polipéptido de interés (por ejemplo, poli A). En determinados casos, las transcripciones truncadas se liberan antes de que el complejo de transcripción abandone el promotor, denominadas transcripciones fallidas. Tal como se usa en la presente, el término "transcripción fallida" se refiere a transcripciones que tienen identidad con respecto al ARN IVT, pero que tienen longitud insuficiente y no presentan todos los elementos necesarios para codificar el polipéptido de interés (por ejemplo, poli A), que generalmente tienen una longitud de 15 nucleótidos o menor. En aspectos de ejemplo de la invención, las transcripciones truncadas y/o fallidas están presentes y no son inductoras de citocinas. En una realización, las transcripciones truncadas y/o fallidas se eliminan de la muestra. En algunas realizaciones, las transcripciones truncadas tienen una longitud de 100 nucleótidos o menor.

También se ha determinado que los métodos de la presente invención producen composiciones que tienen menor heterogeneidad en 3' o heterogeneidad en extremo 3', también denominada en la presente mayor homogeneidad 3', u homogeneidad de extremo 3'. Los inventores de la presente determinaron que las condiciones de reacción de IVT equimolar tradicionales pueden producir transcripciones que terminan en residuo 3' diferentes (por ejemplo, transcripción que no termina de manera uniforme). Se desarrolló un ensayo presentado en los Ejemplos de Trabajo para detectar la heterogeneidad en extremo 3' provocada por las reacciones de IVT tradicionales (el ensayo distingue entre nucleobases que no son A que se producen en el extremo 3' de una transcripción de prueba particular). Cabe destacar que los métodos de la presente invención producen transcripciones que tienen un menor grado de heterogeneidad en extremo 3' (o extremos 3' más homogéneos). Por ejemplo, las transcripciones producidas de acuerdo con reacciones de IVT tradicionales (por ejemplo, reacciones equimolares) pueden producir composiciones en las cuales más de 50 % de las transcripciones (opcionalmente más de 60 %, más de 70 %, más de 75 %, más de 80 % o más) tienen extremos diferentes, mientras que las transcripciones producidas de acuerdo con reacciones de IVT de la invención (por ejemplo, reacciones alfa) pueden producir composiciones en las cuales menos de 50 % de las transcripciones, es decir, más de 50 % de las transcripciones tienen los mismos extremos, es decir, terminan en la misma nucleobase (por ejemplo, con respecto a la plantilla de ADN) (opcionalmente menos de 40 %, menos de 30 %, menos de 25 %, menos de 20% o menos) tienen extremos diferentes).

Las transcripciones truncadas dentro de la población de transcripciones de ARN parcial de cadena simple pueden incluir un intervalo de tamaños. Por ejemplo, en algunas realizaciones, al menos el 80 % de la población de transcripciones truncadas tiene una longitud de 100 nucleótidos o menor. En otras realizaciones, al menos el 50 %, 60 %, 70 %, 85 %, 90 %, 95 %, 98 % o 100 % de la población de transcripciones truncadas tiene una longitud de 100 nucleótidos o menor.

La población de transcripciones de ARN de cadena simple dentro de las composiciones de ARN IVT descritas en la presente típicamente está libre o sustancialmente libre de fragmentos insensibles a RNAsa III. Un "fragmento insensible a RNAsa III" tal como se usa en la presente se refiere a transcripciones de cadena simple que tienen identidad con respecto al ARN IVT (orientación sentido), pero que tienen longitud insuficiente y no presentan todos los elementos necesarios para codificar el polipéptido de interés (que tienen menos nucleótidos que las transcripciones de longitud completa) y donde el fragmento se produce mediante escisión enzimática, en particular RNAsa III. La producción de fragmentos insensibles a RNAsa III puede dar como resultado, por ejemplo, en un proceso de IVT tradicional (tal como se ilustra en la Figura 40) acoplado con una digestión de RNAsa III.

Tal como se muestra en la Figura 40 una primera etapa en un proceso de purificación de IVT/RNAsa III tradicional implica una reacción de transcripción que utiliza plantilla de ADNcd lineal, concentraciones equimolares de NTP y ARN polimerasa en presencia de  $Mg^{2+}$ . La reacción produce una población mezclada de transcripciones truncadas/fallidas de cadena simple, transcripción de ARN de longitud completa, transcripciones mal formadas e impurezas de complemento inverso. Las impurezas de complemento inverso pueden unirse a algunas impurezas del ARN de cadena simple o a otras impurezas, por ejemplo, transcripciones truncadas, que producen ARN de cadena doble y/o ARN que tiene regiones tanto de cadena doble como de cadena simple. Se ha postulado en la técnica que la RNAsa III se puede utilizar para degradar el ARN de cadena doble a partir de composiciones IVT, lo que lo elimina de manera eficaz de la composición. Sin embargo, la RNAsa III también puede degradar regiones de cadena doble de transcripción de ARN de longitud completa y/o transcripciones mal formadas (por ejemplo, regiones de cadena doble que son el resultado de la unión de los complementos inversos dentro de regiones de cola de poliA), lo que deja fragmentos de cadena simple que tienen una longitud menor que las transcripciones de ARN de longitud completa. Estos fragmentos de cadena simple son los fragmentos insensibles a RNAsa III descritos en la presente. Como resultado de esta degradación de RNAsa se pierden cantidades significativas de

transcripciones de longitud completa generadas durante el proceso de IVT, lo que provoca una pérdida significativa de la integridad del producto. Estas composiciones tienen una capacidad significativamente inferior de expresar proteína cuando se administra a una célula o sujeto.

- 5 Los fragmentos insensibles a RNasa III generados después del tratamiento con RNasa III de productos generados de acuerdo con métodos tales como los ilustrados en la Figura 40 pueden incluir un intervalo de tamaños. Por ejemplo, en algunas realizaciones, al menos 80% de la población de transcripciones fallidas tiene una longitud mayor que 100 nucleótidos. En otras realizaciones, al menos 50%, 60%, 70%, 85%, 90%, 95%, 98% o 100% de la población de fragmentos insensibles a RNasa III tiene una longitud mayor que 100

10 nucleótidos.

- Sin pretender limitarse a la teoría, se cree que la eliminación de determinadas especies o contaminantes, por ejemplo, especies o contaminantes de ARNcd, es importante en la preparación de composiciones de ARN IVT para uso terapéutico. Por el contrario, no se cree que la presencia de transcripciones truncadas y/o fallidas residuales en composiciones de ARN IVT sea necesaria; no se cree que dichas especies induzcan citocinas no deseadas y/o una respuesta inmunitaria innata al ARN IVT. En otras realizaciones, la preparación de ARN está sustancialmente libre de transcripciones truncadas o fallidas.

- Aunque las transcripciones truncadas/fallidas pueden estar presentes en las composiciones de ARN IVT producidas por los métodos de la invención, los fragmentos insensibles a RNasa III no están presentes en las composiciones de ARN IVT porque la composición no se trata con RNasa III. Si bien tanto las transcripciones truncadas como los fragmentos insensibles a RNasa III tienen un intervalo de tamaños o longitudes, la longitud promedio de las transcripciones truncadas es menor que la longitud promedio de los fragmentos insensibles a RNasa III. Como tales, cuando la composición comprende una población de transcripciones de ARN parcial de cadena simple en una orientación sentido y más de 80 % de la población de transcripciones de ARN parcial de cadena simple en una orientación sentido tiene una longitud de nucleótidos de 100 nucleótidos o menor. En algunas realizaciones, más de 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de la población de transcripciones de ARN parcial de cadena simple en una orientación sentido tiene una longitud de nucleótidos de 100 nucleótidos o menor. En otras realizaciones, más del 50%, 55%, 60%, 65%, 70%, 75%, 85% u 88% de la población de transcripciones de ARN parcial de cadena simple en una orientación sentido tiene una longitud de nucleótidos de 100 nucleótidos o menor.

- En algunas realizaciones, la preparación de ARN es una composición farmacéutica con un portador farmacéuticamente aceptable. En otras realizaciones, la preparación de ARN es un producto de reacción (por ejemplo, producto de reacción de IVT) que todavía no ha sido sometido a técnicas de purificación adicionales. La preparación de ARN puede incluir una cantidad de otros componentes además del ARN. El producto de reacción, sin embargo, está sustancialmente libre de producto de transcripción de complemento inverso y/o contaminantes de ARN inductores de citocinas.

#### 40 Ensayos

- La cantidad de contaminante, que incluyen producto de transcripción de complemento inverso y/o contaminante de ARN inductores de citocinas puede determinarse mediante métodos conocidos en la técnica. En la técnica se conocen muchos métodos para determinar la pureza de una muestra de ácido nucleico. Los métodos de ejemplo incluyen, de modo no taxativo, los siguientes: cromatografía líquida de alto rendimiento (tal como cromatografía de fase inversa, cromatografía de exclusión por tamaño), electroforesis en gel, y ensayos de traducción para evaluar la calidad y la pureza de la producción de ácido nucleico. La calidad de la preparación de ARN también se puede determinar usando el sistema de electroforesis basado en chip de bioanalizador. La eficacia in vitro se puede analizar, por ejemplo, al transfectar la transcripción de ARN en una línea celular humana, por ejemplo, HeLA, PBMC, fibroblastos BJ, Hek 293). La expresión de proteína del polipéptido de interés puede cuantificarse usando métodos tales como ensayo por inmunoadsorción ligado a enzimas (ELISA), análisis de transferencia Western, o citometría de flujo.

- Se han utilizado una variedad de métodos para detectar y/o cuantificar el ARNcd usando anticuerpos específicos de ARNcd. Estos incluyen ELISA, por ejemplo, ELISA tipo sándwich (Schönborn et ál. (1991) Nucleic Acids Res 19:2993-3000), transferencias puntiformes (para la cuantificación, análisis de especificidad) (Kariko et ál. (2011) Nucleic Acids Res 39:e142), e inmunoprecipitación/inmunotransferencia. En aspectos de ejemplo de la invención, los contaminantes se pueden reconocer usando un ELISA. Se puede utilizar un ensayo K1/J2 o K2/J2 para determinar la abundancia de contaminantes de ARNcd en una muestra. Un ELISA de ejemplo es un ELISA tipo sándwich tal como se muestra a continuación. Bloqueo: Las placas de microtitulación están previamente recubiertas con proteína, por ejemplo, 0.4 µg/pocillo de proteína A 4 °C durante la noche. Los sitios de unión libre se saturan con albúmina de suero bovino (BSA) (por ejemplo, 2%) en amortiguador (por ejemplo, PBS) y luego las placas se lavan con amortiguador (por ejemplo, PBS) y se almacenan a 4 °C. El anticuerpo monoclonal J2 específico de ARNcd (IgG2a) se inmoviliza en la capa de proteína A mediante incubación del sobrenadante de hibridoma (por ejemplo, 100µl por pocillo a 4 °C durante la noche). Las placas se lavan varias veces con amortiguador, por ejemplo, PBS más Tween 20 (por ejemplo, 0.5% (v/v) Tween 20)

y se agregan muestras de ácido nucleico en el amortiguador (por ejemplo, amortiguador TE, 37 °C, 2 h). Después de lavar como se indicó anteriormente, el ácido nucleico unido (es decir, antígenos J2) se identifica mediante incubación sucesiva con el sobrenadante de hibridoma diluido (por ejemplo, 1:2) del anticuerpo monoclonal de IgM K2 específico de ARNcd seguido de un anticuerpo secundario conjugado con fosfatasa alcalina (por ejemplo, IgM antirratón de cabra diluido 1:5000). Ambas etapas de incubación se realizan a 37 °C durante ~1- 2 h. El lavado, incubación de sustrato y lectura de absorción se realizan según métodos reconocidos en la técnica.

Un ensayo similar usando transferencias puntiformes se describe en Kariko et ál., Nuc. Acids Res. 2011; 39(21):e142. El ensayo se realiza mediante análisis de transferencia del ARN (200 ng) en membranas Nytran supercargadas, donde se seca y se bloquea con leche deshidratada sin grasa al 5% en amortiguador TBS-T (Tris-HCl 50 mM, NaCl 150 mM, Tween-20 al 0.05%, pH 7.4). Luego la muestra se incuba con un anticuerpo monoclonal (IgG) K1 o J2 específico de ARNcd durante una hora. Las membranas pueden lavarse con TBS-T e incubarse con un anticuerpo policlonal anticabra conjugado con HRP, por ejemplo. Las membranas se lavan nuevamente, y la señal se detecta usando TMB. La señal se desarrolla con la adición de TMB. El ensayo es útil para detectar dúplex de ARNcd con una longitud mayor que 40 pares de bases.

También se pueden utilizar ensayos de citocinas para detectar contaminantes de ARN. En la técnica se conocen muchos ensayos de citocinas. Los ensayos pueden analizar la inducción de cualquier citocina asociada a productos de IVT impuros. Estas citocinas incluyen, por ejemplo, interleucinas (IL), interferencia (IFN) alfa, beta, y gamma, y TNF. En una realización, se puede utilizar un ensayo basado en células IFN-β. Sus resultados han demostrado que se correlaciona con la presencia del sustrato de RNAsaIII tal como se detectó mediante LC-MS. Se pueden utilizar otros ensayos de citocinas basados en células, tal como, por ejemplo, IL o ensayos multiplex de citocinas.

En ensayos de expresión de BJF IFN-beta y hEPO de ejemplo, se siembran células de fibroblastos BJ (ATCC) en una placa de cultivo celular. Aproximadamente 24 horas después de sembrar, las células se transfectan con ARNm usando lipofectamina u otro agente de administración. Después de la transfección, se recolectan los sobrenadantes y se mide la expresión de IFN usando el kit ELISA de IFN-beta humano, de alta sensibilidad según las instrucciones del fabricante (PBL Assay Science). Brevemente, se mide IFN-β humano en sobrenadantes celulares mediante ensayo de inmunoabsorción ligado a enzimas (ELISA) indirecto. Las placas previamente recubiertas se incuban con sobrenadantes celulares, luego se lavan para retirar el material no unido específicamente. La expresión de IFN-β se analiza al incubar los pocillos con anticuerpo anti- IFN-β seguido de un anticuerpo secundario conjugado con peroxidasa de rábano picante (HRP). Tetrametilbencidina (TMB) es el sustrato de HRP usado para la detección. Los niveles de Epo humana se miden usando el kit ELISA de Epo Humana (Thermo Fisher).

En ensayos Luminex de ejemplo, se recolecta suero de ratones para evaluar los niveles de citocina usando tecnología de ensayo de selección Luminex (R&D Systems). Brevemente, los anticuerpos específicos de analitos se recubren previamente en perlas codificadas por color. Las perlas, estándares y muestras se colocan mediante pipeta en los pocillos y los anticuerpos inmovilizados se unen a los analitos de interés. Después de retirar mediante lavado cualquier sustancia no unida, se agrega a cada pocillo un cóctel de anticuerpos biotinilados específico para los analitos de interés. Después de un lavado para retirar cualquier anticuerpo biotinilado no unido, se agrega a cada pocillo un conjugado de estreptavidina-ficoeritrina (Streptavidin-PE), que se une a los anticuerpos de detección biotinilados. Un lavado final elimina Streptavidin-PE no unida y las perlas se vuelven a suspender en amortiguador y se leen usando un analizador Luminex (R&D Systems). Un primer láser es específico de perlas y determina qué analito se detecta. Un segundo láser determina la magnitud de la señal derivada de PE que se encuentra en proporción directa con la cantidad de analito unido.

#### Ensayos sustitutos para determinar la pureza

En aspectos de ejemplo de la descripción, puede ser conveniente determinar la pureza mediante el uso de un ensayo sustituto que sea susceptible de detección altamente cualitativa y/o cuantitativa de productos y/o impurezas. Por consiguiente, la descripción contempla la determinación de la pureza de una composición de ARN, por ejemplo, un ARN IVT, producido por un método de IVT determinado, al determinar la pureza de un ARN sustituto (por ejemplo, un ARN modelo) producido mediante condiciones idénticas. De esta manera, se puede determinar la pureza indirectamente a través de métodos de detección altamente cualitativos y/o cuantitativos en un sistema sustituto, esta determinación de pureza se correlaciona con la pureza de un ARN IVT producido en un sistema de producción. Además, se pueden utilizar un ensayo tipo sustituto o de reconstitución para determinar la cantidad y la identidad de contaminantes en una preparación de ARN dada indirectamente. En algunos casos puede ser difícil detectar niveles bajos de contaminantes o contaminantes que tienen propiedades estructurales similares a las transcripciones de ARN. Se pueden utilizar sistemas de reconstitución para analizar la actividad biológica tal como la actividad inmunoestimuladora, por ejemplo, actividad de citocinas asociada a contaminantes al agregar de nuevo los contaminantes putativos que no están en las preparaciones de ARN producidas por los métodos de la invención en comparación con la actividad biológica mediante composiciones de ARN producidas por métodos de IVT tradicionales. La reconstitución de

las preparaciones de ARN puras producidas por los métodos de la invención con contaminantes putativos puede demostrar la falta de contaminantes en las preparaciones de ARN puras.

De manera adicional, se pueden utilizar ARN modelos (ARNm sustitutos). En las mismas condiciones de IVT utilizadas para producir el ARN IVT, se produce un ARN modelo de una plantilla de ADN que codifica el ARN modelo. Un ARN modelo o ARNm sustituto, tal como se usa en la presente, se refiere a una transcripción de ARN no codificante que consiste solo en la UTR 5', UTR 3', y una cola de poliA. También se puede utilizar un ARN modelo corto. Un ARN modelo corto es una versión más corta del ARN modelo (solo consiste en UTR5' y una cola de poliA más corta (A20)). La cantidad de producto de transcripción de complemento inverso o especie inductora de citocinas en la composición se determina mediante LC-MS u otro método analítico, dado que la cantidad del ARN modelo indica la cantidad de producto de transcripción de complemento inverso o especie inductora de citocinas en la composición. La cantidad y naturaleza de los contaminantes detectados en el ensayo se correlaciona y predice la cantidad y naturaleza de los contaminantes que se obtendrán usando las condiciones de reacción de IVT idénticas para generar ARNm de longitud completa.

La preparación de ARN producida por los métodos de la invención es una preparación de alta calidad. En algunas realizaciones, la preparación de ARN que resulta directamente de un proceso de IVT se puede utilizar directamente como un reactivo de investigación o un reactivo de diagnóstico o terapéutico sin purificación adicional. En algunas realizaciones, la preparación de ARN puede someterse a una o más etapas de purificación. Por ejemplo, la preparación de ARN puede purificarse a partir de ARN truncado, plantilla de ADN, y enzimas residuales usando cromatografía oligo dT. Un ensayo de cromatografía oligo dT de ejemplo implica empacar politimidina Sepharose de 20 mer (3 ml) en una columna SPE de 5 ml en un colector al vacío de extracción de fase sólida. La transcripción de ARN se aplica a la columna, seguido de lavado y elución. La transcripción de ARN purificada con oligo dT se diafiltra en agua y se concentra hasta 1.22 mg/ml usando filtros giratorios MWCO de 100 kDa Amicon (EMD Millipore). El ARN se recupera y la concentración se puede determinar usando un bioanalizador de electroforesis en gel.

El análisis de la preparación de ARN para determinar la pureza y calidad de la muestra se puede realizar antes o después de la colocación de casquetes. De manera alternativa, el análisis puede realizarse antes o después de la purificación de afinidad basada en captura de poli A. En otra realización, el análisis se puede realizar antes o después de etapas de purificación adicionales opcionales, por ejemplo, cromatografía de intercambio de aniones y similares.

La espectrometría de masas comprende un amplio intervalo de técnicas para identificar y caracterizar compuestos en mezclas. Se pueden utilizar diferentes tipos de enfoques basados en espectrometría de masas para analizar una muestra para determinar su composición. El análisis de espectrometría de masas implica convertir una muestra que se analiza en múltiples iones mediante un proceso de ionización. Cada uno de los iones resultantes, cuando se colocan en un campo de fuerza, se mueve en el campo a lo largo de una trayectoria tal que su aceleración es inversamente proporcional a su relación de masa con respecto a carga. Por lo tanto, se produce un espectro de masa de una molécula que presenta una gráfica de abundancias relativas de iones precursores en función de sus relaciones de masa con respecto a carga. Cuando se utiliza una etapa posterior de espectrometría de masas, tal como espectrometría de masas en tándem, para analizar adicionalmente la muestra al someter los iones precursores a mayor energía, cada ion precursor puede experimentar disociación en fragmentos denominados iones de producto. Los fragmentos resultantes se pueden utilizar para proporcionar información acerca de la naturaleza y la estructura de su molécula precursora.

La espectrometría de masas MALDI-TOF (desorción/ionización láser asistida por matriz con analizador de tiempo de vuelo) proporciona la determinación espectrométrica de la masa de analitos con poca ionización o fácilmente fragmentados de baja volatilidad al incorporarlos en una matriz de material que absorbe luz y medir el peso de la molécula a medida que se ioniza y se hace volar mediante volatilización. Se aplican combinaciones de campos magnéticos y eléctricos en la muestra para hacer que el material ionizado se mueva dependiendo de la masa y carga individual de la molécula. La patente estadounidense n.º 6,043,031, otorgada a Koster et ál., describe un método de ejemplo para identificar mutaciones de base simple dentro del ADN usando MALDI-TOF y otros métodos de espectrometría de masas.

La HPLC (cromatografía líquida de alto rendimiento) se utiliza para la separación analítica de biopolímeros, en función de las propiedades de los biopolímeros. Se puede utilizar HPLC para separar secuencias de ácido nucleico en función de la carga de tamaño, y composición base. Una secuencia de ácido nucleico que tiene un par de bases de diferencia con otro ácido nucleico puede separarse usando HPLC. Por lo tanto, las muestras de ácido nucleico que son idénticas excepto por un solo nucleótido pueden separarse de manera diferencial usando HPLC para identificar la presencia o ausencia de un fragmento de ácido nucleico particular. Preferentemente, la HPLC es HPLC-UV.

En algunas realizaciones, el ARN puede purificarse sin usar una etapa de dsRNasa. Por ejemplo, es posible no utilizar RNasaIII. La composición puede producirse mediante un proceso que no utiliza una etapa de purificación de cromatografía de fase inversa. En una realización, la composición de ARN puede producirse

sin usar una purificación de cromatografía líquida de alto rendimiento (HPLC). Por lo tanto, la composición está libre de contaminantes o reactivos orgánicos residuales asociados a los procesos de purificación tradicionales.

El término “puro” tal como se usa en la presente tiene solo los agentes activos de ácido nucleico objetivo de modo que se reduce o elimina la presencia de ácidos nucleicos no relacionados, es decir, impurezas o contaminantes, que incluyen fragmentos de ARN. Por ejemplo, una muestra de ARN purificada incluye uno o más ácidos nucleicos objetivo o de prueba pero preferentemente está sustancialmente libre de otros ácidos nucleicos detectables mediante métodos descritos. Tal como se usa en la presente, el término “sustancialmente libre” se usa de manera operativa, en el contexto de la evaluación analítica del material. Preferentemente, el material purificado está sustancialmente libre de una o más impurezas o contaminantes que incluyen los productos de transcripción de complemento inverso y/o contaminante de ARN inductor de citocinas descritos en la presente y, por ejemplo, es al menos 90%, 91%, 92%, 93%, 94%, 95%, 96%, o 97% puro; más preferentemente, al menos 98% puro, y aún más preferentemente al menos 99% puro. En algunas realizaciones, una muestra de ARN pura está compuesta por 100% de los ARN objetivo o de prueba y no incluye otro ARN.

En algunas realizaciones, se utiliza electroforesis capilar (CE) para separar el ARN. Se aplica un campo eléctrico a la muestra de modo que pase a través de una solución de electrolitos, tal como una solución amortiguadora acuosa, hacia un vial de destino a través de un capilar. Los analitos migran de manera diferente en función de la movilidad electroforética y se detectan a la salida del capilar. Los datos de salida se registran y luego se visualizan como un electroferograma. Se puede utilizar junto con espectrometría de masas para determinar la identidad de los componentes de la muestra. El sistema de electroforesis capilar está completamente automatizado en el FRAGMENT ANALYZER™, que puede detectar diferencias tan pequeñas como tres pares de bases.

En algunas realizaciones, el analizador de fragmentos (FA) se puede utilizar para cuantificar y purificar el ARN. El analizador de fragmentos automatiza la electroforesis capilar y HPLC.

En algunas realizaciones, las moléculas de ARN son moléculas de ARNm. Tal como se usa en la presente, el término “ARN mensajero” (ARNm) se refiere a cualquier polinucleótido que codifica al menos un péptido o polipéptido de interés y que puede traducirse para producir el péptido polipéptido codificado de interés *in vitro*, *in vivo*, *in situ* o *ex vivo*. Un ARNm ha sido transcrito a partir de una secuencia de ADN mediante una enzima de ARN polimerasa, e interactúa con un ribosoma para sintetizar información genética codificada por el ADN. Generalmente, el ARNm se clasifica en dos subclases: ARNm-pre y ARNm maduro. El ARNm precursor (ARNm-pre) es ARNm que ha sido transcrito mediante ARN polimerasa pero no ha experimentado ningún procesamiento postranscripcional (por ejemplo, colocación de casquetes en 5', empalme, edición, y poliadenilación). El ARNm maduro ha sido modificado mediante procesamiento postranscripcional (por ejemplo, empalmado para eliminar intrones y poliadenilado) y puede interactuar con ribosomas para realizar la síntesis de proteínas. El ARNm puede aislarse de tejidos o células mediante una variedad de métodos. Por ejemplo, se puede realizar una extracción de ARN total en células o un lisado celular y el ARN total extraído resultante se puede purificar (por ejemplo, en una columna que comprende perlas de oligo-dT) para obtener ARNm extraído.

De manera alternativa, el ARNm se puede sintetizar en un ambiente libre de células, por ejemplo, mediante transcripción *in vitro* (IVT). Una “plantilla de transcripción *in vitro*” tal como se usa en la presente, se refiere a un ácido desoxirribonucleico (ADN) adecuado para usar en una reacción de IVT para la producción de ARN mensajero (ARNm). En algunas realizaciones, una plantilla de IVT codifica una región no traducida 5', contiene un marco abierto de lectura y codifica una región no traducida 3' y una cola poliA. La composición y la longitud de una secuencia de nucleótidos particular de una plantilla de IVT dependerán del ARNm de interés codificado por la plantilla.

Una “región no traducida (UTR) 5'” hace referencia a una región de un ARNm que se encuentra directamente antes (es decir, en dirección 5') con respecto al codón de partida (es decir, el primer codón de un transcrito de ARNm traducido por un ribosoma) que no codifica una proteína o péptido.

Una “región no traducida (UTR) 3'” hace referencia a una región de un ARNm que se encuentra directamente después (es decir, en dirección 3') con respecto al codón de terminación (es decir, el codón de una transcripción de ARNm que indica una terminación de traducción) que no codifica una proteína o péptido.

Un “marco de lectura abierto” es una extensión continua de ADN que comienza en un codón de partida (por ejemplo, metionina (ATG)) y que finaliza en un codón de terminación (por ejemplo, TAA, TAG o TGA) y codifica una proteína o péptido.

Una cola poliA es una región del ARNm que se encuentra después, por ejemplo, directamente después (es decir, en dirección 3'), con respecto a la UTR 3' que contiene monofosfatos de adenosina consecutivos múltiples. Una cola de poliA puede contener 10 a 300 monofosfatos de adenosina. Por ejemplo, una cola de



poliA puede contener 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 350, 400, 450, 500, 550, o 600 monofosfatos de adenosina. En algunas realizaciones, una cola de poliA contiene 50 a 250 monofosfatos de adenosina. En una instancia biológica relevante (por ejemplo, en células, *in vivo*, etc.), la cola de poliA protege al ARNm de la degradación enzimática, por ejemplo, en el citoplasma y ayuda con la finalización de la transcripción, la exportación del ARNm desde el núcleo y la traducción.

Por lo tanto, en algunas realizaciones el polinucleótido puede comprender (a) una primera región de nucleótido unidos que codifican un polipéptido de interés; (b) una primera región terminal ubicada en 5' con respecto a dicha primera región que comprende una región no traducida (UTR) 5'; (c) una segunda región terminal ubicada en 3' con respecto a dicha primera región; y (d) una región de cola. Los términos polinucleótido y ácido nucleico se utilizan de manera intercambiable en la presente.

En algunas realizaciones, el polinucleótido incluye de alrededor de 1 a alrededor de 3000, 10 a alrededor de 3000, 20 a alrededor de 3000, 30 a alrededor de 3000, 40 a alrededor de 3000, 50 a alrededor de 3000, 100 a alrededor de 3000, 200 a alrededor de 3000 nucleótidos (por ejemplo, de 200 a 500, de 200 a 1000, de 200 a 1500, de 200 a 3000, de 500 a 1000, de 500 a 1500, de 500 a 2000, de 500 a 3000, de 1,000 a 1500, de 1000 a 2000, de 1000 a 3000, de 1500 a 3000, y de 2000 a 3000).

El ARN IVT puede funcionar como ARN, pero se distinguen del ARN de tipo salvaje en sus características de diseño estructural y/o funcional que son útiles para superar los problemas existentes de producción usando agentes terapéuticos con base en ácidos nucleicos. Por ejemplo, el ARN IVT puede modificarse de manera estructural o modificarse de manera química. Tal como se usa en la presente, una modificación "estructural" es una en la cual dos o más nucleótidos enlazados se insertan, eliminan, duplican, invierten o aleatorizan en un polinucleótido sin modificación química significativa a los nucleótidos en sí. Dado que los enlaces químicos se romperán necesariamente y se reformarán para efectuar una modificación estructural, las modificaciones estructurales tienen naturaleza química y, por lo tanto, son modificaciones químicas. Sin embargo, las modificaciones estructurales darán como resultado una secuencia de nucleótidos diferente. Por ejemplo, el polinucleótido "ATCG" puede modificarse químicamente en "AT-5meC-G". El mismo polinucleótido puede modificarse estructuralmente de "ATCG" en "ATCCCG". Aquí, se ha insertado el dinucleótido "CC", lo que da como resultado una modificación estructural del polinucleótido.

El ADNc que codifica los polinucleótidos descritos en la presente se pueden transcribir usando un sistema de transcripción *in vitro* (IVT). El sistema típicamente comprende un amortiguador de transcripción, trifosfatos de nucleótido (NTP), un inhibidor de RNasa y una polimerasa. Los NTP pueden fabricarse de manera interna, pueden seleccionarse de un proveedor, o pueden sintetizarse tal como se describe en la presente. Los NTP se pueden seleccionar, de modo no taxativo, de los descritos en la presente que incluyen NTP naturales y no naturales (modificados). La polimerasa puede seleccionarse, de modo no taxativo, de ARN polimerasa T7, ARN polimerasa T3 y polimerasas mutantes tales como, de modo no taxativo, polimerasas que pueden incorporar polinucleótidos (por ejemplo, ácidos nucleicos modificados).

#### ARN modificados químicamente

Por lo tanto, en un aspecto de ejemplo, los polinucleótidos pueden incluir al menos una modificación química. Los polinucleótidos pueden incluir diversas sustituciones y/o inserciones de los polinucleótidos nativos o de origen natural. Tal como se usa en la presente, los términos "modificación química" o, según corresponda, "químicamente modificado" hacen referencia a la modificación respecto a ribo o desoxirribonucleósidos de adenosina (A), guanosina (G), uridina (U), timidina (T) o citidina (C) en uno o más de su posición, patrón, porcentaje o población. En general, en la presente estos términos no pretenden hacer referencia a las modificaciones de ribonucleótidos en restos de casquete de ARN de extremo 5' de origen natural.

Las modificaciones pueden ser varias modificaciones distintas. En algunas realizaciones, las regiones pueden contener una, dos o más (opcionalmente diferentes) modificaciones de nucleótidos o nucleósidos. En algunas realizaciones, un polinucleótido modificado introducido en una célula puede exhibir degradación en la célula, en comparación con un polinucleótido no modificado.

Las modificaciones de los polinucleótidos incluyen, de modo no taxativo, los enumerados en detalle más adelante. El polinucleótido puede comprender modificaciones que son de origen natural, de origen no natural o el polinucleótido puede comprender modificaciones tanto de origen natural como de origen no natural.

Los polinucleótidos pueden incluir cualquier modificación útil, tal como al azúcar, la nucleobase, o el enlace internucleótido (por ejemplo, a un fosfato enlazador/a un enlace fosfodiéster/a la estructura fosfodiéster). Uno o más átomos de una nucleobase de pirimidina puede reemplazarse o sustituirse con amino opcionalmente sustituido, tiol opcionalmente sustituido, alquilo opcionalmente sustituido (por ejemplo, metilo o etilo), o halo (por ejemplo, cloro o fluoro). En determinadas realizaciones, las modificaciones (por ejemplo, una o más modificaciones) están presentes en cada uno del azúcar y el enlace internucleótido. Las modificaciones pueden

ser modificaciones de ácidos ribonucleicos (ARN) en ácidos desoxirribonucleicos (ADN), ácidos nucleicos de treosa (ANT), ácidos nucleicos de glicol (ANG), ácidos nucleicos peptídicos (ANP), ácidos nucleicos bloqueados (ANB) o híbridos de estos. En la presente se describen modificaciones adicionales.

- 5 Se pueden introducir nucleótidos modificados no naturales en polinucleótidos durante la síntesis o después de la síntesis de las cadenas para alcanzar las funciones o propiedades deseadas. Las modificaciones pueden estar en un enlace internucleótido, las bases purina o pirimidina, o azúcares. La modificación puede introducirse en el extremo de una cadena o en cualquier otro lugar de la cadena; con síntesis química o con una enzima de polimerasa. Cualquiera de las regiones de los polinucleótidos puede modificarse químicamente.

10

La presente descripción proporciona polinucleótidos compuestos por nucleósidos y nucleótidos modificados o no modificados y combinaciones de estos. Tal como se describe en la presente, "nucleósido" se define como un compuesto que contiene una molécula de azúcar (por ejemplo, una pentosa o ribosa) o un derivado de esta, en combinación con una base orgánica (por ejemplo, una purina o pirimidina) o un derivado de esta (que también se denomina en la presente "nucleobase"). Tal como se describe en la presente, "nucleótido" se define como un nucleósido que incluye un grupo fosfato. Los nucleótidos modificados pueden sintetizarse mediante cualquier método útil, tal como se describe en la presente (por ejemplo, de forma química, enzimática o recombinante, para incluir uno o más nucleótidos no naturales o modificados). Los polinucleótidos pueden comprender una región o regiones de nucleótidos enlazados. Tales regiones pueden tener enlaces de estructura variable. Los enlaces pueden ser enlaces fosfodiéster estándares, en cuyo caso los polinucleótidos comprenderían regiones de nucleótidos. Es posible incorporar cualquier combinación de base/azúcar o enlazador en los polinucleótidos utilizados en el método de la invención.

20

- En algunas realizaciones, un ARN producido por los métodos de la invención incluye una combinación de una o más de las nucleobases modificadas mencionadas anteriormente (por ejemplo, una combinación de 2, 3 o 4 de las nucleobases modificadas mencionadas anteriormente).

25

Las modificaciones de ácidos nucleicos que son útiles en la presente invención incluyen, de modo no taxativo, los indicados en la Tabla a continuación.

30

Nombre	Símbolo	Base	De origen natural
2-metiltio-N6-(cis-hidroxiisopentenil)adenosina	ms2i6A	A	SÍ
2-metiltio-N6-metiladenosina	ms2m6A	A	SÍ
2-metiltio-N6-treonil carbamoiladenosina	ms2t6A	A	SÍ
N6-glicinilcarbamoiladenosina	g6A	A	SÍ
N6-isopenteniladenosina	i6A	A	SÍ
N6-metiladenosina	m6A	A	SÍ
N6-treonilcarbamoiladenosina	t6A	A	SÍ
1,2'-O-dimetiladenosina	m1Am	A	SÍ
1-metiladenosina	m1A	A	SÍ
2'-O-metiladenosina	Am	A	SÍ
2'-O-ribosiladenosina (fosfato)	Ar(p)	A	SÍ
2-metiladenosina	m2A	A	SÍ
2-metiltio-N6 isopenteniladenosina	ms2i6A	A	SÍ
2-metiltio-N6-hidroxinorvalil carbamoiladenosina	ms2hn6A	A	SÍ
2'-O-metiladenosina	m6A	A	SÍ
2'-O-ribosiladenosina (fosfato)	Ar(p)	A	SÍ
Isopenteniladenosina	lga	A	SÍ
N6-(cis-hidroxiisopentenil)adenosina	io6A	A	SÍ
N6,2'-O-dimetiladenosina	m6Am	A	SÍ
N6,2'-O-dimetiladenosina	m6Am	A	SÍ
N6,N6,2'-O-trimetiladenosina	m62Am	A	SÍ
N6,N6-dimetiladenosina	m62A	A	SÍ

(continuación)

Nombre	Símbolo	Base	De origen natural
N6-acetiladenosina	ac6A	A	SÍ
N6-hidroxinorvalilcarbamoiladenosina	hn6A	A	SÍ
N6-metil-N6-treonilcarbamoiladenosina	m6t6A	A	SÍ
2-metiladenosina	m2A	A	SÍ
2-metiltio-N6-isopenteniladenosina	ms2i6A	A	SÍ
7-deaza-adenosina	--	A	NO
N1-metil-adenosina	--	A	NO
N6, N6 (dimetil)adenina	--	A	NO
N6-cis-hidroxi-isopentenil-adenosina	--	A	NO
$\alpha$ -tio-adenosina	--	A	NO
2 (amino)adenina	--	A	NO
2 (aminopropil)adenina	--	A	NO
2 (metiltio) N6 (isopentenil)adenina	--	A	NO
2-(alquil)adenina	--	A	NO
2-(aminoalquil)adenina	--	A	NO
2-(aminopropil)adenina	--	A	NO
2-(halo)adenina	--	A	NO
2-(halo)adenina	--	A	NO
2-(propil)adenina	--	A	NO
2'-Amino-2'-desoxi-ATP	--	A	NO
2'-Azido-2'-desoxi-ATP	--	A	NO
2'-Desoxi-2'-a-aminoadenosina TP	--	A	NO
2'-Desoxi-2'-a-aminoadenosina TP	--	A	NO
6(alquil)adenina	--	A	NO
6(metil)adenina	--	A	NO
6-(alquil)adenina	--	A	NO
6-(metil)adenina	--	A	NO
7(deaza)adenina	--	A	NO
8(alquenil)adenina	--	A	NO
8(alquinil)adenina	--	A	NO
8 (amino)adenina	--	A	NO
8(tioalquil)adenina	--	A	NO
8-(alquenil)adenina	--	A	NO
8-(alquil)adenina	--	A	NO
8-(alquinil)adenina	--	A	NO
8-(amino)adenina	--	A	NO
8-(halo)adenina	--	A	NO
8-(hidroxil)adenina	--	A	NO
8-(tioalquil)adenina	--	A	NO
8-(tiol)adenina	--	A	NO
8-azido-adenosina	--	A	NO
aza adenina	--	A	NO
deaza adenina	--	A	NO

(continuación)

Nombre	Símbolo	Base	De origen natural
N6 (metil)adenina	--	A	NO
N6-(isopentil)adenina	--	A	NO
7-deaza-8-aza-adenosina	--	A	NO
7-metiladenina	--	A	NO
1-Deazaadenosina TP	--	A	NO
2'Fluoro-N6-Bz-desoxiadenosina TP	--	A	NO
2'-OMe-2-Amino-ATP	--	A	NO
2'O-metil-N6-Bz-desoxiadenosina TP	--	A	NO
2'-a-Etiniladenosina TP	--	A	NO
2-aminoadenina	--	A	NO
2-Aminoadenosina TP	--	A	NO
2-Amino-ATP	--	A	NO
2'-a-Trifluorometiladenosina TP	--	A	NO
2-Azidoadenosina TP	--	A	NO
2'-b-Etiniladenosina TP	--	A	NO
2-Bromoadenosina TP	--	A	NO
2'-b-Trifluorometiladenosina TP	--	A	NO
2-Cloroadenosina TP	--	A	NO
2'-Desoxi-2',2'-difluoroadenosina TP	--	A	NO
2'-Desoxi-2'-a-mercaptopadenosina TP	--	A	NO
2'-Desoxi-2'-a-tiometoxiadenosina TP	--	A	NO
2'-Desoxi-2'-b-aminoadenosina TP	--	A	NO
2'-Desoxi-2'-b-azidoadenosina TP	--	A	NO
2'-Desoxi-2'-b-bromoadenosina TP	--	A	NO
2'-Desoxi-2'-b-cloroadenosina TP	--	A	NO
2'-Desoxi-2'-b-fluoroadenosina TP	--	A	NO
2'-Desoxi-2'-b-yodoadenosina TP	--	A	NO
2'-Desoxi-2'-b-mercaptopadenosina TP	--	A	NO
2'-Desoxi-2'-b-tiometoxiadenosina TP	--	A	NO
2-Fluoroadenosina TP	--	A	NO
2-Yodoadenosina TP	--	A	NO
2-Mercaptopadenosina TP	--	A	NO
2-metoxi-adenina	--	A	NO
2-metiltio-adenina	--	A	NO
2-Trifluorometiladenosina TP	--	A	NO
3-Deaza-3-bromoadenosina TP	--	A	NO
3-Deaza-3-cloroadenosina TP	--	A	NO
3-Deaza-3-fluoroadenosina TP	--	A	NO
3-Deaza-3-yodoadenosina TP	--	A	NO
3-Deazaadenosina TP	--	A	NO
4'-Azidoadenosina TP	--	A	NO
4'-adenosina carbocíclica TP	--	A	NO
4'-Etiniladenosina TP	--	A	NO

(continuación)

Nombre	Símbolo	Base	De origen natural
5'-Homo-adenosina TP	--	A	NO
8-Aza-ATP	--	A	NO
8-bromo-adenosina TP	--	A	NO
8-Trifluorometiladenosina TP	--	A	NO
9-Deazaadenosina TP	--	A	NO
2-aminopurina	--	A/G	NO
7-deaza-2,6-diaminopurina	--	A/G	NO
7-deaza-8-aza-2,6-diaminopurina	--	A/G	NO
7-deaza-8-aza-2-aminopurina	--	A/G	NO
2,6-diaminopurina	--	A/G	NO
7-deaza-8-aza-adenina, 7-deaza-2-aminopurina	--	A/G	NO
2-tiocitidina	s2C	C	SÍ
3-metilcitidina	m3C	C	SÍ
5-formilcitidina	f5C	C	SÍ
5-hidroximetilcitidina	hm5C	C	SÍ
5-metilcitidina	m5C	C	SÍ
N4-acetilcitidina	ac4C	C	SÍ
2'-O-metilcitidina	Cm	C	SÍ
2'-O-metilcitidina	Cm	C	SÍ
5,2'-O-dimetilcitidina	m5 Cm	C	SÍ
5-formil-2'-O-metilcitidina	f5Cm	C	SÍ
Lisidina	k2C	C	SÍ
N4,2'-O-dimetilcitidina	m4Cm	C	SÍ
N4-acetil-2'-O-metilcitidina	ac4Cm	C	SÍ
N4,N4-Dimetil-2'-OMe-Citidina TP	m4C	C	SÍ
4-metilcitidina	--	C	NO
5-aza-citidina	--	C	NO
Pseudo-iso-citidina	--	C	NO
pirrolo-citidina	--	C	NO
$\alpha$ -tio-citidina	--	C	NO
2-(tio)citosina	--	C	NO
2'-Amino-2'-desoxi-CTP	--	C	NO
2'-Azido-2'-desoxi-CTP	--	C	NO
2'-Desoxi-2'-a-aminocitidina TP	--	C	NO
2'-Desoxi-2'-a-azidocitidina TP	--	C	NO
3 (deaza) 5 (aza)citosina	--	C	NO
3(metil)citosina	--	C	NO
3-(alquil)citosina	--	C	NO
3-(deaza) 5 (aza)citosina	--	C	NO
3-(metil)citidina	--	C	NO
4,2'-O-dimetilcitidina	--	C	NO
5(halo)citosina	--	C	NO
5(metil)citosina	--	C	NO

(continuación)

Nombre	Símbolo	Base	De origen natural
5(propinil)citosina	--	C	NO
5 (trifluorometil)citosina	--	C	NO
5-(alquil)citosina	--	C	NO
5-(alquinil)citosina	--	C	NO
5-(halo)citosina	--	C	NO
5-(propinil)citosina	--	C	NO
5-(trifluorometil)citosina	--	C	NO
5-bromo-citidina	--	C	NO
5-yodo-citidina	--	C	NO
5-propinil citosina	--	C	NO
6-(azo)citosina	--	C	NO
6-aza-citidina	--	C	NO
aza citosina	--	C	NO
deaza citosina	--	C	NO
N4 (acetil)citosina	--	C	NO
1-metil-1-deaza-pseudoisocitidina	--	C	NO
1-metil-pseudoisocitidina	--	C	NO
2-metoxi-5-metil-citidina	--	C	NO
2-metoxi-citidina	--	C	NO
2-tio-5-metil-citidina	--	C	NO
4-metoxi-1-metil-pseudoisocitidina	--	C	NO
4-metoxi-pseudoisocitidina	--	C	NO
4-tio-1-metil-1-deaza-pseudoisocitidina	--	C	NO
4-tio-1-metil-pseudoisocitidina	--	C	NO
4-tio-pseudoisocitidina	--	C	NO
5-aza-zebularina	--	C	NO
5-metil-zebularina	--	C	NO
pirrolo-pseudoisocitidina	--	C	NO
Zebularina	--	C	NO
(E)-5-(2-Bromo-vinil)citidina TP	--	C	NO
clorhidrato de 2,2'-anhidro-citidina TP	--	C	NO
2'Fluor-N4-Bz-citidina TP	--	C	NO
2'Fluoro-N4-Acetil-citidina TP	--	C	NO
2'-O-Metil-N4-Acetil-citidina TP	--	C	NO
2'O-metil-N4-Bz-citidina TP	--	C	NO
2'-a-Etinilcitidina TP	--	C	NO
2'-a-Trifluorometilcitidina TP	--	C	NO
2'-b-Etinilcitidina TP	--	C	NO
2'-b-Trifluorometilcitidina TP	--	C	NO
2'-Desoxi-2',2'-difluorocitidina TP	--	C	NO
2'-Desoxi-2'-a-mercaptopcitidina TP	--	C	NO
2'-Desoxi-2'-a-tiometoxicitidina TP	--	C	NO
2'-Desoxi-2'-b-aminocitidina TP	--	C	NO

(continuación)

Nombre	Símbolo	Base	De origen natural
2'-Desoxi-2'-b-azidocitidina TP	--	C	NO
2'-Desoxi-2'-b-bromocitidina TP	--	C	NO
2'-Desoxi-2'-b-clorocitidina TP	--	C	NO
2'-Desoxi-2'-b-fluorocitidina TP	--	C	NO
2'-Desoxi-2'-b-yodocitidina TP	--	C	NO
2'-Desoxi-2'-b-mercaptopcitidina TP	--	C	NO
2'-Desoxi-2'-b-tiometoxicitidina TP	--	C	NO
2'-O-Metil-5-(1-propinil)citidina TP	--	C	NO
3'-Etilcitidina TP	--	C	NO
4'-Azidocitidina TP	--	C	NO
4'-citidina carbocíclica TP	--	C	NO
4'-Etilcitidina TP	--	C	NO
5-(1-Propinil)ara-citidina TP	--	C	NO
5-(2-Cloro-fenil)-2-tiocitidina TP	--	C	NO
5-(4-Amino-fenil)-2-tiocitidina TP	--	C	NO
5-Aminoalil-CTP	--	C	NO
5-Cianocitidina TP	--	C	NO
5-Etilara-citidina TP	--	C	NO
5-Etilcitidina TP	--	C	NO
5'-Homo-citidina TP	--	C	NO
5-Metoxicitidina TP	--	C	NO
5-Trifluorometil-Citidina TP	--	C	NO
N4-Amino-citidina TP	--	C	NO
N4-Benzoil-citidina TP	--	C	NO
Pseudoisocitidina	--	C	NO
7-metilguanosina	m7G	G	SÍ
N2,2'-O-dimetilguanosina	m2Gm	G	SÍ
N2-metilguanosina	m2G	G	SÍ
Wyosina	imG	G	SÍ
1,2'-O-dimetilguanosina	m1Gm	G	SÍ
1-metilguanosina	m1G	G	SÍ
2'-O-metilguanosina	Gm	G	SÍ
2'-O-ribosilguanosina (fosfato)	Gr(p)	G	SÍ
2'-O-metilguanosina	Gm	G	SÍ
2'-O-ribosilguanosina (fosfato)	Gr(p)	G	SÍ
7-aminometil-7-deazaguanosina	preQ1	G	SÍ
7-ciano-7-deazaguanosina	preQ0	G	SÍ
Arqueosina	G+	G	SÍ
Metilwyosina	mimG	G	SÍ
N2,7-dimetilguanosina	m2,7G	G	SÍ
N2,N2,2'-O-trimetilguanosina	m22Gm	G	SÍ
N2,N2,7-trimetilguanosina	m2,2,7G	G	SÍ
N2,N2-dimetilguanosina	m22G	G	SÍ

(continuación)

Nombre	Símbolo	Base	De origen natural
N2,7,2'-O-trimetilguanosina	m2,7Gm	G	SÍ
6-tio-guanosina	--	G	NO
7-deaza-guanosina	--	G	NO
8-oxo-guanosina	--	G	NO
N1-metil-guanosina	--	G	NO
$\alpha$ -tio-guanosina	--	G	NO
2(propil)guanina	--	G	NO
2-(alquil)guanina	--	G	NO
2'-Amino-2'-desoxi-GTP	--	G	NO
2'-Azido-2'-desoxi-GTP	--	G	NO
2'-Desoxi-2'-a-aminoguanosina TP	--	G	NO
2'-Desoxi-2'-a-azidoguanosina TP	--	G	NO
6(metil)guanina	--	G	NO
6-(alquil)guanina	--	G	NO
6-(metil)guanina	--	G	NO
6-metil-guanosina	--	G	NO
7(alquil)guanina	--	G	NO
7(deaza)guanina	--	G	NO
7(metil)guanina	--	G	NO
7-(alquil)guanina	--	G	NO
7-(deaza)guanina	--	G	NO
7-(metil)guanina	--	G	NO
8(alquil)guanina	--	G	NO
8(alquinil)guanina	--	G	NO
8(halo)guanina	--	G	NO
8(tioalquil)guanina	--	G	NO
8(alquenil)guanina	--	G	NO
8-(alquil)guanina	--	G	NO
8-(alquinil)guanina	--	G	NO
8(amino)guanina	--	G	NO
8(halo)guanina	--	G	NO
8(hidroxil)guanina	--	G	NO
8-(tioalquil)guanina	--	G	NO
8-(tiol)guanina	--	G	NO
aza guanina	--	G	NO
deaza guanina	--	G	NO
N (metil)guanina	--	G	NO
N-(metil)guanina	--	G	NO
1-metil-6-tio-guanosina	--	G	NO
6-metoxi-guanosina	--	G	NO
6-tio-7-deaza-8-aza-guanosina	--	G	NO
6-tio-7-deaza-guanosina	--	G	NO
6-tio-7-metil-guanosina	--	G	NO



(continuación)

Nombre	Símbolo	Base	De origen natural
7-deaza-8-aza-guanosina	--	G	NO
7-metil-8-oxo-guanosina	--	G	NO
N2,N2-dimetil-6-tio-guanosina	--	G	NO
N2-metil-6-tio-guanosina	--	G	NO
1-Me-GTP	--	G	NO
2'Fluoro-N2-isobutil-guanosina TP	--	G	NO
2'O-metil-N2-isobutil-guanosina TP	--	G	NO
2'-a-Etinilguanosina TP	--	G	NO
2'-a-Trifluorometilguanosina TP	--	G	NO
2'-b-Etinilguanosina TP	--	G	NO
2'-b-Trifluorometilguanosina TP	--	G	NO
2'-Desoxi-2',2'-difluoroguanosina TP	--	G	NO
2'-Desoxi-2'-a-mercaptoguanosina TP	--	G	NO
2'-Desoxi-2'-a-tiometoxiguanosina TP	--	G	NO
2'-Desoxi-2'-b-aminoguanosina TP	--	G	NO
2'-Desoxi-2'-b-azidoguanosina TP	--	G	NO
2'-Desoxi-2'-b-bromoguanosina TP	--	G	NO
2'-Desoxi-2'-b-cloroguanosina TP	--	G	NO
2'-Desoxi-2'-b-fluoroguanosina TP	--	G	NO
2'-Desoxi-2'-b-yodoguanosina TP	--	G	NO
2'-Desoxi-2'-b-mercaptoguanosina TP	--	G	NO
2'-Desoxi-2'-b-tiometoxiguanosina TP	--	G	NO
4'-Azidoguanosina TP	--	G	NO
4'-guanosina carbocíclica TP	--	G	NO
4'-Etinilguanosina TP	--	G	NO
5'-Homo-guanosina TP	--	G	NO
8-bromo-guanosina TP	--	G	NO
9-Deazaguanosina TP	--	G	NO
N2-isobutil-guanosina TP	--	G	NO
1-metilinosina	m1I	I	SÍ
Inosina	I	I	SÍ
1,2'-O-dimetilinosina	m1Im	I	SÍ
2'-O-metilinosina	Im	I	SÍ
7-metilinosina		I	NO
2'-O-metilinosina	Im	I	SÍ
Epoxiqueuosina	oQ	Q	SÍ
galactosil-queuosina	galQ	Q	SÍ
Manosilqueuosina	manQ	Q	SÍ
Queuosina	Q	Q	SÍ
aliamino-timidina	--	T	NO
aza timidina	--	T	NO
deaza timidina	--	T	NO
desoxi-timidina	--	T	NO

(continuación)

Nombre	Símbolo	Base	De origen natural
2'-O-metiluridina	--	U	SÍ
2-tiouridina	s2U	U	SÍ
3-metiluridina	m3U	U	SÍ
5-carboximetiluridina	cm5U	U	SÍ
5-hidroxiuridina	ho5U	U	SÍ
5-metiluridina	m5U	U	SÍ
5-taurinometil-2-tiouridina	tm5s2U	U	SÍ
5-taurinometiluridina	tm5U	U	SÍ
Dihidrouridina	D	U	SÍ
Pseudouridina	Ψ	U	SÍ
(3-(3-amino-3-carboxipropil)uridina	acp3U	U	SÍ
1-metil-3-(3-amino-5-carboxipropil)pseudouridina	m1acp3Ψ	U	SÍ
1-metilpseduouridina	m1Ψ	U	SÍ
1-metil-pseudouridina	--	U	SÍ
2'-O-metiluridina	Um	U	SÍ
2'-O-metilpseudouridina	Ψm	U	SÍ
2'-O-metiluridina	Um	U	SÍ
2-tio-2'-O-metiluridina	s2Um	U	SÍ
3-(3-amino-3-carboxipropil)uridina	acp3U	U	SÍ
3,2'-O-dimetiluridina	m3Um	U	SÍ
3-Metil-pseudo-Uridina TP	--	U	SÍ
4-tiouridina	s4U	U	SÍ
5-(carboxihidroximetil)uridina	chm5U	U	SÍ
éster metílico de 5-(carboxihidroximetil)uridina	mchm5U	U	SÍ
5,2'-O-dimetiluridina	m5Um	U	SÍ
5,6-dihidro-uridina	--	U	SÍ
5-aminometil-2-tiouridina	nm5s2U	U	SÍ
5-carbamoilmetil-2'-O-metiluridina	ncm5Um	U	SÍ
5-carbamoilmetiluridina	ncm5U	U	SÍ
5-carboxihidroximetiluridina	--	U	SÍ
éster metílico de 5-carboxihidroximetiluridina	--	U	SÍ
5-carboximetilaminometil-2'-O-metiluridina	cmnm5Um	U	SÍ
5-carboximetilaminometil-2-tiouridina	cmnm5s2U	U	SÍ
5-carboximetilaminometil-2-tiouridina	--	U	SÍ
5-carboximetilaminometiluridina	cmnm5U	U	SÍ
5-carboximetilaminometiluridina	--	U	SÍ
5-Carbamoilmetiluridina TP	--	U	SÍ
5-metoxicarbonilmetil-2'-O-metiluridina	mcm5Um	U	SÍ
5-metoxicarbonilmetil-2-tiouridina	mcm5s2U	U	SÍ
5-metoxicarbonilmetiluridina	mcm5U	U	SÍ
5-metoxiuridina	mo5U	U	SÍ
5-metil-2-tiouridina	m5s2U	U	SÍ
5-metilaminometil-2-selenouridina	mnm5se2U	U	SÍ

(continuación)

Nombre	Símbolo	Base	De origen natural
5-metilaminometil-2-tiouridina	mnms2U	U	SÍ
5-metilaminometiluridina	mnmsU	U	SÍ
5-Metildihidrouridina	--	U	SÍ
ácido 5-oxiacético- Uridina TP	--	U	SÍ
éster metílico del ácido 5-oxiacético-Uridina TP	--	U	SÍ
N1-metil-pseudo-uridina	--	U	SÍ
ácido uridina 5-oxiacético	cmo5U	U	SÍ
éster metílico del ácido uridina 5-oxiacético	mcmo5U	U	SÍ
3-(3-Amino-3-carboxipropil)-Uridina TP	--	U	SÍ
5-(iso-Pentenilaminometil)- 2-tiouridina TP	--	U	SÍ
5-(iso-Pentenilaminometil)-2'-O-metiluridina TP	--	U	SÍ
5-(iso-Pentenilaminometil)uridina TP	--	U	SÍ
5-propinil uracilo	--	U	NO
$\alpha$ -tio-uridina	--	U	NO
1 (aminoalquilamino-carboniletilenil)-2(tio)-pseudouracilo	--	U	NO
1 (aminoalquilaminocarboniletilenil)-2,4-(ditio)pseudouracilo	--	U	NO
1 (aminoalquilaminocarboniletilenil)-4(tio)pseudouracilo	--	U	NO
1 (aminoalquilaminocarboniletilenil)-pseudouracilo	--	U	NO
1 (aminocarboniletilenil)-2(tio)-pseudouracilo	--	U	NO
1 (aminocarboniletilenil)-2,4-(ditio)pseudouracilo	--	U	NO
1 (aminocarboniletilenil)-4(tio)pseudouracilo	--	U	NO
1 (aminocarboniletilenil)-pseudouracilo	--	U	NO
2(tio)-pseudouracilo sustituido en 1	--	U	NO
2,4-(ditio)pseudouracilo sustituido en 1	--	U	NO
4 (tio)pseudouracilo sustituido en 1	--	U	NO
pseudouracilo sustituido en 1	--	U	NO
1-(aminoalquilamino-carboniletilenil)-2(tio)-pseudouracilo	--	U	NO
1-Metil-3-(3-amino-3-carboxipropil)pseudouridina TP	--	U	NO
1-Metil-3-(3-amino-3-carboxipropil)pseudo-UTP	--	U	NO
1-Metil-pseudo-UTP	--	U	NO
2 (tio)pseudouracilo	--	U	NO
2' desoxi uridina	--	U	NO
2' fluorouridina	--	U	NO
2-(tio)uracilo	--	U	NO
2,4-(ditio)psuedouracilo	--	U	NO
2' metil, 2' amino, 2' azido, 2' fluoro-guanosina	--	U	NO
2'-Amino-2'-desoxi-UTP	--	U	NO
2'-Azido-2'-desoxi-UTP	--	U	NO
2'-Azido-desoxiuridina TP	--	U	NO

2'-O-metilpseudouridina	--	U	NO
2' desoxi uridina	2' dU	U	NO

(continuación)

Nombre	Símbolo	Base	De origen natural
2' fluorouridina	--	U	NO
2'-Desoxi-2'-a-aminouridina TP	--	U	NO
2'-Desoxi-2'-a-azidouridina TP	--	U	NO
2-metilpseudouridina	m3Ψ	U	NO
3(3-amino-3-carboxipropil)uracilo	--	U	NO
4 (tio)pseudouracilo	--	U	NO
4-(tio)pseudouracilo	--	U	NO
4-(tio)uracilo	--	U	NO
4-tiouracilo	--	U	NO
5 (1,3-diazolo-1-alquil)uracilo	--	U	NO
5 (2-aminopropil)uracilo	--	U	NO
5 (aminoalquil)uracilo	--	U	NO
5 (dimetilaminoalquil)uracilo	--	U	NO
5 (guanidinioalquil)uracilo	--	U	NO
5 (metoxicarbonilmetil)-2-(tio)uracilo	--	U	NO
5 (metoxicarbonil-metil)uracilo	--	U	NO
5 (metil) 2 (tio)uracilo	--	U	NO
5 (metil) 2,4 (ditio)uracilo	--	U	NO
5 (metil) 4 (tio)uracilo	--	U	NO
5 (metilaminometil)-2 (tio)uracilo	--	U	NO
5 (metilaminometil)-2,4 (ditio)uracilo	--	U	NO
5 (metilaminometil)-4 (tio)uracilo	--	U	NO
5 (propinil)uracilo	--	U	NO
5 (trifluorometil)uracilo	--	U	NO
5-(2-aminopropil)uracilo	--	U	NO
5-(alquil)-2-(tio)pseudouracilo	--	U	NO
5-(alquil)-2,4 (ditio)pseudouracilo	--	U	NO
5-(Alquil)-4 (tio)pseudouracilo	--	U	NO
5-(alquil)pseudouracilo	--	U	NO
5-(alquil)uracilo	--	U	NO
5-(alquinil)uracilo	--	U	NO
5-(alilamino)uracilo	--	U	NO
5-(cianoalquil)uracilo	--	U	NO
5-(dialquilaminoalquil)uracilo	--	U	NO
5-(dimetilaminoalquil)uracilo	--	U	NO
5-(guanidinioalquil)uracilo	--	U	NO
5-(halo)uracilo	--	U	NO
5-(1,3-diazolo-1-alquil)uracilo	--	U	NO
5-(metoxi)uracilo	--	U	NO
5-(metoxicarbonilmetil)-2-(tio)uracilo	--	U	NO
5-(metoxicarbonil-metil)uracilo	--	U	NO

5-(metil) 2 (tio)uracilo	--	U	NO
5-(metil) 2,4 (ditio)uracilo	--	U	NO

(continuación)

Nombre	Símbolo	Base	De origen natural
5-(metil) 4 (tio)uracilo	--	U	NO
5-(metil)-2-(tio)pseudouracilo	--	U	NO
5-(metil)-2,4 (ditio)pseudouracilo	--	U	NO
5-(Metil)-4 (tio)pseudouracilo	--	U	NO
5-(metil)pseudouracilo	--	U	NO
5-(metilaminometil)-2 (tio)uracilo	--	U	NO
5-(metilaminometil)-2,4 (ditio)uracilo	--	U	NO
5-(metilaminometil)-4-(tio)uracilo	--	U	NO
5-(propinil)uracilo	--	U	NO
5-(trifluorometil)uracilo	--	U	NO
5-aminoalil-uridina	--	U	NO
5-bromo-uridina	--	U	NO
5-yodo-uridina	--	U	NO
5-uracilo	--	U	NO
6 (azo)uracilo	--	U	NO
6-(azo)uracilo	--	U	NO
6-aza-uridina	--	U	NO
aliamino-uracilo	--	U	NO
aza uracilo	--	U	NO
deaza uracilo	--	U	NO
N3 (metil)uracilo	--	U	NO
Ácido pseudo-UTP-1-2-etanoico	--	U	NO
Pseudouracilo	--	U	NO
4-Tio-pseudo-UTP	--	U	NO
1-carboximetil-pseudouridina	--	U	NO
1-metil-1-deaza-pseudouridina	--	U	NO
1-propinil-uridina	--	U	NO
1-taurinometil-1-metil-uridina	--	U	NO
1-taurinometil-4-tio-uridina	--	U	NO
1-taurinometil-pseudouridina	--	U	NO
2-metoxi-4-tio-pseudouridina	--	U	NO
2-tio-1-metil-1-deaza-pseudouridina	--	U	NO
2-tio-1-metil-pseudouridina	--	U	NO
2-tio-5-aza-uridina	--	U	NO
2-tio-dihidropseudouridina	--	U	NO
2-tio-dihidrouridina	--	U	NO
2-tio-pseudouridina	--	U	NO
4-metoxi-2-tio-pseudouridina	--	U	NO
4-metoxi-pseudouridina	--	U	NO
4-tio-1-metil-pseudouridina	--	U	NO

4-tio-pseudouridina	--	U	NO
5-aza-uridina	--	U	NO
Dihidropseudouridina	--	U	NO

(continuación)

Nombre	Símbolo	Base	De origen natural
(±)1-(2-Hidroxipropil)pseudouridina TP	--	U	NO
(2R)-1-(2-Hidroxipropil)pseudouridina TP	--	U	NO
(2S)-1-(2-Hidroxipropil)pseudouridina TP	--	U	NO
(E)-5-(2-Bromo-vinil)ara-uridina TP	--	U	NO
(E)-5-(2-Bromo-vinil)uridina TP	--	U	NO
(Z)-5-(2-Bromo-vinil)ara-uridina TP	--	U	NO
(Z)-5-(2-Bromo-vinil)uridina TP	--	U	NO
1-(2,2,2-Trifluoroetil)-pseudo-UTP	--	U	NO
1-(2,2,3,3,3-Pentafluoropropil)pseudouridina TP	--	U	NO
1-(2,2-Dietoxietil)pseudouridina TP	--	U	NO
1-(2,4,6-Trimetilbencil)pseudouridina TP	--	U	NO
1-(2,4,6-Trimetil-bencil)pseudo-UTP	--	U	NO
1-(2,4,6-Trimetil-fenil)pseudo-UTP	--	U	NO
1-(2-Amino-2-carboxietil)pseudo-UTP	--	U	NO
1-(2-Amino-etil)pseudo-UTP	--	U	NO
1-(2-Hidroxietil)pseudouridina TP	--	U	NO
1-(2-Metoxietil)pseudouridina TP	--	U	NO
1-(3,4-Bis-trifluorometoxibencil)pseudouridina TP	--	U	NO
1-(3,4-Dimetoxibencil)pseudouridina TP	--	U	NO
1-(3-Amino-3-carboxipropil)pseudo-UTP	--	U	NO
1-(3-Amino-propil)pseudo-UTP	--	U	NO
1-(3-Ciclopropil-prop-2-inil)pseudouridina TP	--	U	NO
1-(4-Amino-4-carboxibutil)pseudo-UTP	--	U	NO
1-(4-Amino-bencil)pseudo-UTP	--	U	NO
1-(4-Amino-butil)pseudo-UTP	--	U	NO
1-(4-Amino-fenil)pseudo-UTP	--	U	NO
1-(4-Azidobencil)pseudouridina TP	--	U	NO
1-(4-Bromobencil)pseudouridina TP	--	U	NO
1-(4-Clorobencil)pseudouridina TP	--	U	NO
1-(4-Fluorobencil)pseudouridina TP	--	U	NO
1-(4-Yodobencil)pseudouridina TP	--	U	NO
1-(4-Metanosulfonilbencil)pseudouridina TP	--	U	NO
1-(4-Metoxibencil)pseudouridina TP	--	U	NO
1-(4-Metoxi-bencil)pseudo-UTP	--	U	NO
1-(4-Metoxi-fenil)pseudo-UTP	--	U	NO
1-(4-Metilbencil)pseudouridina TP	--	U	NO
1-(4-Metil-bencil)pseudo-UTP	--	U	NO
1-(4-Nitrobencil)pseudouridina TP	--	U	NO
1-(4-Nitro-bencil)pseudo-UTP	--	U	NO
1-(4-Nitro-fenil)pseudo-UTP	--	U	NO

1-(4-Tiometoxibencil)pseudouridina TP	--	U	NO
1-(4-Trifluorometoxibencil)pseudouridina TP	--	U	NO
1-(4-Trifluorometilbencil)pseudouridina TP	--	U	NO

(continuación)

Nombre	Símbolo	Base	De origen natural
1-(5-Amino-pentil)pseudo-UTP	--	U	NO
1-(6-Amino-hexil)pseudo-UTP	--	U	NO
1,6-Dimetil-pseudo-UTP	--	U	NO
1-[3-(2-{2-[2-(2-Aminoetoxi)-etoxi]-etoxi}-propionil)pseudouridina TP	--	U	NO
1-[3-[2-(2-Aminoetoxi)-etoxi]-propionil]pseudouridina TP	--	U	NO
1-Acetilpseudouridina TP	--	U	NO
1-Alquil-6-(1-propinil)-pseudo-UTP	--	U	NO
1-Alquil-6-(2-propinil)-pseudo-UTP	--	U	NO
1-Alquil-6-alil-pseudo-UTP	--	U	NO
1-Alquil-6-etinil-pseudo-UTP	--	U	NO
1-Alquil-6-homoalil-pseudo-UTP	--	U	NO
1-Alquil-6-vinil-pseudo-UTP	--	U	NO
1-Alilpseudouridina TP	--	U	NO
1-Aminometil-pseudo-UTP	--	U	NO
1-Benzoilpseudouridina TP	--	U	NO
1-Benciloximetilpseudouridina TP	--	U	NO
1-Bencil-pseudo-UTP	--	U	NO
1-Biotinil-PEG2-pseudouridina TP	--	U	NO
1-Biotinilpseudouridina TP	--	U	NO
1-Butil-pseudo-UTP	--	U	NO
1-Cianometilpseudouridina TP	--	U	NO
1-Ciclobutilmetil-pseudo-UTP	--	U	NO
1-Ciclobutil-pseudo-UTP	--	U	NO
1-Cicloheptilmetil-pseudo-UTP	--	U	NO
1-Cicloheptil-pseudo-UTP	--	U	NO
1-Ciclohexilmetil-pseudo-UTP	--	U	NO
1-Ciclohexil-pseudo-UTP	--	U	NO
1-Ciclooctilmetil-pseudo-UTP	--	U	NO
1-Ciclooctil-pseudo-UTP	--	U	NO
1-Ciclopentilmetil-pseudo-UTP	--	U	NO
1-Ciclopentil-pseudo-UTP	--	U	NO
1-Ciclopropilmetil-pseudo-UTP	--	U	NO
1-Ciclopropil-pseudo-UTP	--	U	NO
1-Etil-pseudo-UTP	--	U	NO
1-Hexil-pseudo-UTP	--	U	NO
1-Homoalilpseudouridina TP	--	U	NO
1-Hidroximetilpseudouridina TP	--	U	NO
1-iso-propil-pseudo-UTP	--	U	NO
1-Me-2-tio-pseudo-UTP	--	U	NO
1-Me-4-tio-pseudo-UTP	--	U	NO

1-Me-alfa-tio-pseudo-UTP	--	U	NO
1-Metanosulfonilmetilpseudouridina TP	--	U	NO
1-Metoximetilpseudouridina TP	--	U	NO

(continuación)

Nombre	Símbolo	Base	De origen natural
1-Metil-6-(2,2,2-Trifluoroetil)pseudo-UTP	--	U	NO
1-Metil-6-(4-morfolino)-pseudo-UTP	--	U	NO
1-Metil-6-(4-tiomorfolino)-pseudo-UTP	--	U	NO
1-Metil-6-(fenil sustituido)pseudo-UTP	--	U	NO
1-Metil-6-amino-pseudo-UTP	--	U	NO
1-Metil-6-azido-pseudo-UTP	--	U	NO
1-Metil-6-bromo-pseudo-UTP	--	U	NO
1-Metil-6-butil-pseudo-UTP	--	U	NO
1-Metil-6-cloro-pseudo-UTP	--	U	NO
1-Metil-6-ciano-pseudo-UTP	--	U	NO
1-Metil-6-dimetilamino-pseudo-UTP	--	U	NO
1-Metil-6-etoxi-pseudo-UTP	--	U	NO
1-Metil-6-etilcarboxilato-pseudo-UTP	--	U	NO
1-Metil-6-etil-pseudo-UTP	--	U	NO
1-Metil-6-fluoro-pseudo-UTP	--	U	NO
1-Metil-6-formil-pseudo-UTP	--	U	NO
1-Metil-6-hidroxiamino-pseudo-UTP	--	U	NO
1-Metil-6-hidroxi-pseudo-UTP	--	U	NO
1-Metil-6-yodo-pseudo-UTP	--	U	NO
1-Metil-6-iso-propil-pseudo-UTP	--	U	NO
1-Metil-6-metoxi-pseudo-UTP	--	U	NO
1-Metil-6-metilamino-pseudo-UTP	--	U	NO
1-Metil-6-fenil-pseudo-UTP	--	U	NO
1-Metil-6-propil-pseudo-UTP	--	U	NO
1-Metil-6-terc-butil-pseudo-UTP	--	U	NO
1-Metil-6-trifluorometoxi-pseudo-UTP	--	U	NO
1-Metil-6-trifluorometil-pseudo-UTP	--	U	NO
1-Morfolinometilpseudouridina TP	--	U	NO
1-Pentil-pseudo-UTP	--	U	NO
1-Fenil-pseudo-UTP	--	U	NO
1-Pivaloilpseudouridina TP	--	U	NO
1-Propargilpseudouridina TP	--	U	NO
1-Propil-pseudo-UTP	--	U	NO
1-propinil-pseudouridina	--	U	NO
1-p-tolil-pseudo-UTP	--	U	NO
1-terc-Butil-pseudo-UTP	--	U	NO
1-Tiometoximetilpseudouridina TP	--	U	NO
1-Tiomorfolinometilpseudouridina TP	--	U	NO
1-Trifluoroacetilpseudouridina TP	--	U	NO
1-Trifluorometil-pseudo-UTP	--	U	NO



1-Vinilpseudouridina TP	--	U	NO
2,2'-anhidro-uridina TP	--	U	NO
2'-bromo-desoxiuridina TP	--	U	NO

(continuación)

Nombre	Símbolo	Base	De origen natural
2'-F-5-Metil-2'-desoxi-UTP	--	U	NO
2'-OMe-5-Me-UTP	--	U	NO
2'-OMe-pseudo-UTP	--	U	NO
2'-a-Etiniluridina TP	--	U	NO
2'-a-Trifluorometiluridina TP	--	U	NO
2'-b-Etiniluridina TP	--	U	NO
2'-b-Trifluorometiluridina TP	--	U	NO
2'-Desoxi-2',2'-difluorouridina TP	--	U	NO
2'-Desoxi-2'-a-mercaptopuridina TP	--	U	NO
2'-Desoxi-2'-a-tiometoxiuridina TP	--	U	NO
2'-Desoxi-2'-b-aminouridina TP	--	U	NO
2'-Desoxi-2'-b-azidouridina TP	--	U	NO
2'-Desoxi-2'-b-bromouridina TP	--	U	NO
2'-Desoxi-2'-b-clorouridina TP	--	U	NO
2'-Desoxi-2'-b-fluorouridina TP	--	U	NO
2'-Desoxi-2'-b-yodouridina TP	--	U	NO
2'-Desoxi-2'-b-mercaptopuridina TP	--	U	NO
2'-Desoxi-2'-b-tiometoxiuridina TP	--	U	NO
2-metoxi-4-tio-uridina	--	U	NO
2-metoxiuridina	--	U	NO
2'-O-Metil-5-(1-propinil)uridina TP	--	U	NO
3-Alquil-pseudo-UTP	--	U	NO
4'-Azidouridina TP	--	U	NO
4'-uridina carbocíclica TP	--	U	NO
4'-Etiniluridina TP	--	U	NO
5-(1-Propinil)ara-uridina TP	--	U	NO
5-(2-Furanil)uridina TP	--	U	NO
5-Cianouridina TP	--	U	NO
5-Dimetilaminouridina TP	--	U	NO
5'-Homo-uridina TP	--	U	NO
5-yodo-2'-fluoro-desoxiuridina TP	--	U	NO
5-Feniletiniluridina TP	--	U	NO
5-Trideuterometil-6-deuterouridina TP	--	U	NO
5-Trifluorometil-Uridina TP	--	U	NO
5-Vinilarauridina TP	--	U	NO
6-(2,2,2-Trifluoroetil)-pseudo-UTP	--	U	NO
6-(4-Morfolino)-pseudo-UTP	--	U	NO
6-(4-Tiomorfolino)-pseudo-UTP	--	U	NO
6-(Fenil sustituido)-pseudo-UTP	--	U	NO
6-Amino-pseudo-UTP	--	U	NO

6-Azido-pseudo-UTP	--	U	NO
6-Bromo-pseudo-UTP	--	U	NO
6-Butil-pseudo-UTP	--	U	NO

(continuación)

Nombre	Símbolo	Base	De origen natural
6-Cloro-pseudo-UTP	--	U	NO
6-Ciano-pseudo-UTP	--	U	NO
6-Dimetilamino-pseudo-UTP	--	U	NO
6-Etoxi-pseudo-UTP	--	U	NO
6-Etilcarboxilato-pseudo-UTP	--	U	NO
6-Etil-pseudo-UTP	--	U	NO
6-Fluoro-pseudo-UTP	--	U	NO
6-Formil-pseudo-UTP	--	U	NO
6-Hidroxiamino-pseudo-UTP	--	U	NO
6-Hidroxipseudo-UTP	--	U	NO
6-Yodo-pseudo-UTP	--	U	NO
6-iso-Propil-pseudo-UTP	--	U	NO
6-Metoxi-pseudo-UTP	--	U	NO
6-Metilamino-pseudo-UTP	--	U	NO
6-Metil-pseudo-UTP	--	U	NO
6-Fenil-pseudo-UTP	--	U	NO
6-Fenil-pseudo-UTP	--	U	NO
6-Propil-pseudo-UTP	--	U	NO
6-terc-Butil-pseudo-UTP	--	U	NO
6-Trifluorometoxi-pseudo-UTP	--	U	NO
6-Trifluorometil-pseudo-UTP	--	U	NO
Alfa-tio-pseudo-UTP	--	U	NO
Pseudouridina ácido 1-(4-metilbencenosulfónico) TP	--	U	NO
Pseudouridina ácido 1-(4-metilbenzoico) TP	--	U	NO
Pseudouridina TP ácido 1-[3-(2-etoxi)]propiónico	--	U	NO
Pseudouridina TP ácido 1-[3-{2-(2-[2-(2-etoxi)-etoxi]-etoxi)-etoxi}]propiónico	--	U	NO
Pseudouridina TP 1-[3-{2-(2-[2-{2-(2-etoxi)-etoxi]-etoxi)-etoxi}-etoxi}]propiónico	--	U	NO
Pseudouridina TP ácido 1-[3-{2-(2-[2-(2-etoxi)-etoxi]-etoxi)]propiónico	--	U	NO
Pseudouridina TP ácido 1-[3-{2-(2-etoxi)-etoxi}] propiónico	--	U	NO
Pseudouridina TP ácido 1-metilfosfónico	--	U	NO
Pseudouridina TP éster dietílico del ácido 1-metilfosfónico	--	U	NO
Ácido pseudo-UTP-N1-3-propiónico	--	U	NO
Ácido pseudo-UTP-N1-4-butanoico	--	U	NO
Ácido pseudo-UTP-N1-5-pentanoico	--	U	NO
Ácido pseudo-UTP-N1-6-hexanoico	--	U	NO
Ácido pseudo-UTP-N1-7-heptanoico	--	U	NO
Ácido pseudo-UTP-N1-metil-p-benzoico	--	U	NO

Ácido pseudo-UTP-N1-p-benzoico	--	U	NO
Wybutosina	yW	W	SÍ
Hidroxiwybutosina	OHyW	W	SÍ
Isowyosina	imG2	W	SÍ

(continuación)

Nombre	Símbolo	Base	De origen natural
Peroxiwybutosina	o2yW	W	SÍ
Hidroxiwybutosina sin modificar	OHyW*	W	SÍ
4-demetilwyosina	imG-14	W	SÍ

En algunas realizaciones, la nucleobase modificada es pseudouridina ( $\psi$ ), N1-metilpseudouridina ( $m^1\psi$ ), 2-tiouridina, 4'-tiouridina, 5-metilcitosina, 2-tio-1-metil-1-deaza-pseudouridina, 2-tio-1-metil-pseudouridina, 2-tio-5-aza-uridina, 2-tio-dihidropseudouridina, 2-tio-dihidrouridina, 2-tio-pseudouridina, 4-metoxi-2-tio-pseudouridina, 4-metoxi-pseudouridina, 4-tio-1-metil-pseudouridina, 4-tio-pseudouridina, 5-aza-uridina, dihidropseudouridina, 5-metiluridina, o 2'-O-metiluridina. En algunas realizaciones, un ARN producido por los métodos de la invención incluye una combinación de una o más de las nucleobases modificadas mencionadas anteriormente (por ejemplo, una combinación de 2, 3 o 4 de las nucleobases modificadas mencionadas anteriormente).

A<producido por los métodos de la invención>A

En algunas realizaciones, la nucleobase modificada es 1-metil-pseudouridina ( $m^1\psi$ ), 5-metoxi-uridina ( $mo^5U$ ), 5-metil-citidina ( $m^5C$ ), pseudouridina ( $\psi$ ),  $\alpha$ -tio-guanosina, o  $\alpha$ -tio-adenosina. En algunas realizaciones, un ARNm producido por los métodos de la invención incluye una combinación de una o más de las nucleobases modificadas mencionadas anteriormente (por ejemplo, una combinación de 2, 3 o 4 de las nucleobases modificadas mencionadas anteriormente).

En algunas realizaciones, el ARN comprende pseudouridina ( $\psi$ ) y 5-metil-citidina ( $m^5C$ ). En algunas realizaciones, el ARN comprende 1-metil-pseudouridina ( $m^1\psi$ ). En algunas realizaciones, el ARN comprende 1-metil-pseudouridina ( $m^1\psi$ ) y 5-metil-citidina ( $m^5C$ ). En algunas realizaciones, el ARN comprende 2-tiouridina ( $s^2U$ ). En algunas realizaciones, el ARN comprende 2-tiouridina y 5-metil-citidina ( $m^5C$ ). En algunas realizaciones, el ARN comprende 5-metoxi-uridina ( $mo^5U$ ). En algunas realizaciones, el ARN comprende 5-metoxi-uridina ( $mo^5U$ ) y 5-metil-citidina ( $m^5C$ ). En algunas realizaciones, el ARN comprende 2'-O-metil uridina. En algunas realizaciones, el ARN comprende 2'-O-metil uridina y 5-metil-citidina ( $m^5C$ ). En algunas realizaciones, el ARN comprende N6-metil-adenosina ( $m^6A$ ). En algunas realizaciones, el ARN comprende N6-metil-adenosina ( $m^6A$ ) y 5-metil-citidina ( $m^5C$ ).

En determinadas realizaciones, un ARN producido por los métodos de la invención se modifica de manera uniforme (es decir, completamente modificado, modificado a través de toda la secuencia) para una modificación particular. Por ejemplo, es posible modificar de manera uniforme un ARN con 5-metil-citidina ( $m^5C$ ), lo que significa que todos los residuos citosina en la secuencia de ARN se reemplazan con 5-metil-citidina ( $m^5C$ ). De manera similar, los ARN producidos por los métodos de la invención se pueden modificar de manera uniforme para cualquier tipo de residuo nucleótido presente en la secuencia mediante el reemplazo con un residuo modificado tal como los indicados anteriormente.

En algunas realizaciones, la nucleobase modificada es una citosina modificada. Las nucleobases, nucleósidos y nucleótidos de ejemplo que tienen una citosina modificada incluyen N4-acetil-citidina ( $ac4C$ ), 5-metil-citidina ( $m^5C$ ), 5-halo-citidina (por ejemplo, 5-yodo-citidina), 5-hidroximetil-citidina ( $hm5C$ ), 1-metil-pseudoisocitidina, 2-tio-citidina ( $s2C$ ), 2-tio-5-metil-citidina.

En algunas realizaciones, la nucleobase modificada es una uridina modificada. Las nucleobases, nucleósidos y nucleótidos de ejemplo que tienen una uridina modificada incluyen 5-cianouridina o 4' tiouridina.

En algunas realizaciones, la nucleobase modificada es una adenina modificada. Las nucleobases, nucleósidos y nucleótidos de ejemplo que tienen una adenina modificada incluyen 7-desaza-adenina, A<producido por los métodos de la invención>A, 1-metil-adenosina ( $m1A$ ), 2-metil-adenina ( $m2A$ ), N6-metil-adenosina ( $m6A$ ) y 2,6-diaminopurina.

En algunas realizaciones, la nucleobase modificada es una guanina modificada. Las nucleobases, nucleósidos y nucleótidos de ejemplo que tienen una guanina modificada incluyen inosina ( $I$ ), 1-metil-inosina ( $m1I$ ), wyosina ( $imG$ ), metilwyosina ( $mimG$ ), 7-desaza-guanosina, 7-ciano-7-desaza-guanosina ( $preQ0$ ), 7-aminometil-7-desaza-guanosina ( $preQ1$ ), 7-metil-guanosina ( $m7G$ ), 1-metil-guanosina ( $m1G$ ), 8-oxo-guanosina, 7-metil-8-

oxo-guanosina.

En una realización, los polinucleótidos producidos por los métodos de la presente invención, pueden tener una modificación química uniforme de la totalidad o cualquiera del mismo tipo de nucleótido o una población de modificaciones producidas por una mera titulación hacia abajo de la misma modificación de partida en la totalidad o cualquiera del mismo tipo de nucleótido, o un porcentaje medido de una modificación química de cualquiera del mismo tipo de nucleótido pero con incorporación aleatoria, tal como cuando todas las uridinas se reemplazan por un análogo de uridina, por ejemplo, pseudouridina. En otra realización, los polinucleótidos pueden tener una modificación química uniforme de dos, tres o cuatro del mismo tipo de nucleótido a través del polinucleótido entero (tal como todas las uridinas y todas las citosina, etc. se modifican de la misma manera). Cuando los polinucleótidos producidos por los métodos de la invención de la invención se modifican químicamente y/o estructuralmente los polinucleótidos pueden denominarse "polinucleótidos modificados".

Generalmente, la longitud del polinucleótido IVT (por ejemplo, ARN IVT) que codifica un polipéptido de interés es mayor que alrededor de 30 nucleótidos de longitud (por ejemplo, al menos o mayor que alrededor de 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2500, y 3000, 4000, 5000, 6000, 7.000, 8000, 9000, 10.000, 20.000, 30.000, 40.000, 50.000, 60.000, 70.000, 80.000, 90.000 o hasta 100.000 nucleótidos inclusive).

En algunas realizaciones, el polinucleótido IVT (por ejemplo, ARN IVT) incluye de alrededor de 30 a alrededor de 100.000 nucleótidos (por ejemplo, de 30 a 50, de 30 a 100, de 30 a 250, de 30 a 500, de 30 a 1000, de 30 a 1500, de 30 a 3000, de 30 a 5000, de 30 a 7000, de 30 a 10.000, de 30 a 25.000, de 30 a 50.000, de 30 a 70.000, de 100 a 250, de 100 a 500, de 100 a 1000, de 100 a 1500, de 100 a 3000, de 100 a 5000, de 100 a 7000, de 100 a 10.000, de 100 a 25.000, de 100 a 50.000, de 100 a 70.000, de 100 a 100.000, de 500 a 1000, de 500 a 1500, de 500 a 2000, de 500 a 3000, de 500 a 5000, de 500 a 7000, de 500 a 10.000, de 500 a 25.000, de 500 a 50.000, de 500 a 70.000, de 500 a 100.000, de 1000 a 1500, de 1000 a 2000, de 1000 a 3000, de 1000 a 5000, de 1000 a 7000, de 1000 a 10.000, de 1000 a 25.000, de 1000 a 50.000, de 1000 a 70.000, de 1000 a 100.000, de 1500 a 3000, de 1500 a 5000, de 1500 a 7000, de 1500 a 10.000, de 1500 a 25.000, de 1500 a 50.000, de 1500 a 70.000, de 1500 a 100.000, de 2000 a 3000, de 2000 a 5000, de 2000 a 7000, de 2000 a 10.000, de 2000 a 25.000, de 2000 a 50.000, de 2000 a 70.000, y de 2000 a 100.000).

En algunas realizaciones, un ácido nucleico tal como se describe en la presente es un polinucleótido quimérico. Los polinucleótidos quiméricos o construcciones de ARN mantienen una organización modular similar a los polinucleótidos IVT, pero los polinucleótidos quiméricos comprenden una o más modificaciones o alteraciones estructurales y/o químicas que imparten propiedades útiles al polinucleótido. Como tales, los polinucleótidos quiméricos, que son moléculas de ARN modificadas de la presente invención, producidos por los métodos de la invención, se denominan "ARN modificado quimérico" o "ARN quimérico". Los polinucleótidos quiméricos tienen partes o regiones que difieren en tamaño y/o patrón de modificación química, posición de modificación química, porcentaje de modificación química o población de modificación química y combinaciones de las anteriores.

#### Polipéptidos de interés

En algunas realizaciones de la invención, el ARN es uno o más de los siguientes: ARNm, ARNm modificado, ARN no modificado, ARN Inc, ARN de autorreplicación, ARN circular, ARN guía CRISPR, y similares. En realizaciones, el ARN es ARN que codifica un polipéptido tal como ARNm o ARN de autorreplicación.

Las composiciones de ARN altamente puras pueden utilizarse para producir polipéptidos de interés, por ejemplo, proteínas terapéuticas, antígeno de vacuna, y similares. En algunas realizaciones, los ácidos nucleicos son ARN terapéuticos. Tal como se usa en la presente, el término "ARN terapéutico" hace referencia a un ARNm que codifica una proteína terapéutica. Las proteínas terapéuticas median una variedad de efectos en una célula hospedadora o un sujeto con el fin de tratar una enfermedad o mejorar los signos y síntomas de una enfermedad. Por ejemplo, una proteína terapéutica puede reemplazar una proteína que es deficiente o anormal, aumentar la función de una proteína endógena, proporcionar una función novedosa a una célula (por ejemplo, inhibir o activar un activador celular endógena, o actuar como agente de administración para otro compuesto terapéutico (por ejemplo, un conjugado de anticuerpo-fármaco). El ARNm terapéutico puede ser útil para el tratamiento de las siguientes enfermedades y afecciones: infecciones bacterianas, infecciones virales, infecciones parasitarias, trastornos de proliferación celular, trastornos genéticos, y trastornos autoinmunitarios.

Por lo tanto, los polinucleótidos producidos por los métodos de la invención se pueden utilizar como agentes terapéuticos o profilácticos. Se proporcionan para usar en la medicina. Por ejemplo, el ARN descrito en la presente se puede administrar a un sujeto, donde los polinucleótidos se traducen in vivo para producir un péptido terapéutico. Se proporcionan composiciones, métodos, kits, y reactivos para el diagnóstico, tratamiento o prevención de una enfermedad o afección en seres humanos y otros mamíferos. Los agentes terapéuticos

activos incluyen los polinucleótidos, células que contienen polinucleótidos o polipéptidos traducidos a partir de polinucleótidos.

- 5 Los polinucleótidos pueden inducirse para traducción en una célula, tejido u organismo. Dicha traducción puede producirse in vivo, ex vivo, en cultivo, o in vitro. Se pone en contacto una célula, tejido u organismo con una cantidad eficaz de una composición que contiene un polinucleótido que contiene los polinucleótidos de ARN.

- 10 Se proporciona una "cantidad eficaz" de los polinucleótidos en función, al menos en parte, del tejido objetivo, el tipo de células objetivo, las vías de administración, las características físicas del polinucleótido (por ejemplo, tamaño y extensión de los nucleósidos modificados) y otros componentes de los polinucleótidos y otros factores determinantes. En general, una cantidad eficaz de los polinucleótidos proporciona una producción de péptidos inducida o reforzada en la célula, preferentemente más eficaz que una composición que contiene un polinucleótido no modificado correspondiente que codifica el mismo péptido. Es posible demostrar la producción aumentada de péptidos mediante el aumento de transfección celular, aumento de traducción de proteínas del polinucleótido, reducción de degradación de ácidos nucleicos (tal como se demuestra, por ejemplo, mediante el aumento de la duración de la traducción proteica a partir de un polinucleótido modificado) o alteración de producción de péptidos en la célula hospedadora.

- 20 El ARN producido por los métodos de la invención puede diseñarse para que codifique polipéptidos de interés que se seleccionan de varias categorías objetivo que incluyen, de modo no taxativo, elementos biológicos, anticuerpos, vacunas, péptidos o proteínas terapéuticas, péptidos que penetran células, proteínas secretadas, proteínas de membrana plasmática, proteína citoplasmáticas o citoesqueléticas, proteínas unidas a la membrana intracelular, proteínas nucleares, proteínas asociadas a enfermedad humana, restos de direccionamiento o aquellas proteínas codificadas por el genoma humano para las cuales no se ha identificado una indicación terapéutica pero que, sin embargo, tienen utilidad en áreas de investigación y descubrimiento. "Proteína terapéutica" se refiere a una proteína que, cuando se administra a una célula, tiene un efecto terapéutico, diagnóstico y/o profiláctico, y/o provoca un efecto farmacológico y/o biológico deseado.

- 30 El ARN descrito en la presente puede codificar uno o más elementos biológicos. Tal como se usa en la presente, un "elemento biológico" es una molécula basada en polipéptidos producida mediante los métodos proporcionados en la presente y que se pueden utilizar para tratar, curar, mitigar, prevenir, o diagnosticar una afección médica o enfermedad grave o potencialmente mortal. Los elementos biológicos incluyen, de modo no taxativo, extractos alergénicos (por ejemplo, para pruebas y vacunas de alergia), componentes sanguíneos, productos de terapia génica, tejido humano o productos celulares usados en trasplantes, vacunas, anticuerpos monoclonales, citocinas, factores de crecimiento, enzimas, trombolíticos, e inmunomoduladores, entre otros.

- 40 Uno o más elementos biológicos que se comercializan o que están en desarrollo actualmente pueden codificarse por el ARN producido por los métodos de la invención. Sin pretender limitarse a la teoría, se cree que la incorporación de los polinucleótidos codificantes de un elemento biológico conocido en el ARN producido por los métodos de la invención dará como resultado una eficacia terapéutica mejorada debido, al menos en parte, a la especificidad, pureza y/o selectividad de los diseños de construcción.

- 45 El ARN descrito en la presente puede codificar uno o más anticuerpos o fragmentos de estos. El término "anticuerpo" incluye anticuerpos monoclonales (que incluyen anticuerpos de longitud completa que tienen una región Fc de inmunoglobulina), composiciones de anticuerpos con especificidad poliepitópica, anticuerpos multiespecíficos (por ejemplo, anticuerpos biespecíficos, diacuerpos y moléculas de cadena simple), así como fragmentos de anticuerpos. En la presente, el término "inmunoglobulina" (Ig) se utiliza de forma intercambiable con "anticuerpo". Tal como se usa en la presente, el término "anticuerpo monoclonal" hace referencia a un anticuerpo obtenido de una población de anticuerpos sustancialmente homogéneos, es decir, los anticuerpos individuales que comprenden la población son idénticos con la excepción de posibles mutaciones de origen natural y/o modificaciones postraduccionales (por ejemplo, isomerizaciones, amidaciones) que pueden estar presentes en cantidades mínimas. Los anticuerpos monoclonales son muy específicos y se dirigen a un único punto antigénico.

- 55 Los anticuerpos monoclonales de la presente incluyen específicamente anticuerpos "quiméricos" (inmunoglobulinas) en los cuales una parte de la cadena pesada y/o ligera es idéntica u homóloga a secuencias correspondientes en anticuerpos derivados de una especie particular o que pertenecen a una clase o subclase de anticuerpo particular, mientras que el resto de la o las cadenas son idénticas u homólogas a las secuencias correspondientes en los anticuerpos derivados de otra especie o que pertenecen a otra clase o subclase de anticuerpo, así como también fragmentos de dichos anticuerpos, siempre y cuando exhiban la actividad biológica deseada. Los anticuerpos quiméricos de interés en la presente incluyen, de modo no taxativo, anticuerpos "primatizados" que comprenden secuencias de unión al antígeno de dominio variable derivadas de un primate no humano (por ejemplo, monos del viejo mundo, simios, etc.) y secuencias de región constante humana.

- 65 Un "fragmento de anticuerpo" comprende una parte de un anticuerpo intacto, preferentemente, la región

variable y/o de unión al antígeno del anticuerpo intacto. Los ejemplos de fragmentos de anticuerpo incluyen fragmentos Fab, Fab', F(ab')<sub>2</sub> y Fv; diacuerpos; anticuerpos lineales; nanocuerpos; moléculas de anticuerpos de cadena simple; y anticuerpos multiespecíficos formados a partir de fragmentos de anticuerpo.

- 5 Cualquiera de las cinco clases de inmunoglobulinas, IgA, IgD, IgE, IgG e IgM, puede estar codificada por el ARN producido por los métodos de la invención, que incluye las cadenas pesadas designadas alfa, delta, épsilon, gamma y mu, respectivamente. También se incluyen las secuencias de polinucleótidos que codifican las subclases, gamma y mu. Por lo tanto, cualquiera de las subclases puede estar codificada en parte o en su totalidad e incluye las siguientes subclases: IgG1, IgG2, IgG3, IgG4, IgA1 e IgA2. De acuerdo con la presente invención, uno o más anticuerpos o fragmentos que se comercializan o que están en desarrollo actualmente pueden codificarse por el ARN producido por los métodos de la invención.

- 15 Los anticuerpos codificados por el ARN producido por los métodos de la invención pueden utilizarse para tratar afecciones o enfermedades en muchas áreas terapéuticas, tales como, de modo no taxativo, hematología, cardiovascular, del SNC, intoxicación (que incluye antivenenos), dermatología, endocrinología, gastrointestinal, imagenología médica, musculoesquelética, oncología, inmunología, respiratoria, sensorial y antiinfecciosa.

- 20 En una realización, el ARN descrito en la presente puede codificar anticuerpos monoclonales y/o variantes de estos. Las variantes de los anticuerpos también pueden incluir, de modo no taxativo, variantes de sustitución, sustitución conservadora de aminoácidos, variantes de inserción, variantes de eliminación y/o derivados covalentes. En una realización, el ARN descrito en la presente puede codificar una región Fc de inmunoglobulina. En otra realización, el ARN puede codificar una región Fc de inmunoglobulina variante.

- 25 El ARNm descrito en la presente puede codificar uno o más antígenos de vacuna. Tal como se usa en la presente, un "antígeno de vacuna" es una preparación biológica que mejora la inmunidad a una enfermedad o agente infeccioso particular. De acuerdo con la presente invención, uno o más antígenos de vacuna que se comercializan o que están en desarrollo actualmente pueden codificarse por el ARN producido por los métodos de la invención.

- 30 Los antígenos de vacuna codificados por el ARN producido por los métodos de la invención se pueden utilizar para tratar afecciones o enfermedades en muchas áreas terapéuticas tales como, de modo no taxativo, cáncer, alergia y enfermedad infecciosa. En algunas realizaciones, las vacunas contra el cáncer pueden ser vacunas contra el cáncer personalizadas en forma de concatémicos o ARN individuales que codifican epítomos peptídicos o una combinación de estos.

- 35 El ARN producido por los métodos de la invención se puede diseñar para que codifique uno o más péptidos antimicrobianos (PAM) o péptidos antivirales (PAV). Los PAMP y PAV han sido aislados y descritos de una amplia variedad de animales tales como, de modo no taxativo, microorganismos, invertebrados, plantas, anfibios, aves, peces, y mamíferos. Los polipéptidos antimicrobianos descritos en la presente pueden bloquear la fusión celular y/o entrada viral mediante uno o más virus con envoltura (por ejemplo, VIH, VHC). Por ejemplo, el polipéptido antimicrobiano puede comprender o consistir en un péptido sintético que corresponde a una región, por ejemplo, una secuencia consecutiva de al menos alrededor de 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, o 60 aminoácidos de la subunidad transmembrana de una proteína con envoltura viral, por ejemplo, VIH-1 gp120 o gp41. Las secuencias de nucleótidos de aminoácido de VIH-1 gp120 o gp41 se describen, por ejemplo, en Kuiken et ál., (2008). "HIV Sequence Compendium," Los Alamos National Laboratory.

- 50 En algunas realizaciones, el polipéptido antimicrobiano puede tener al menos alrededor de 75%, 80%, 85%, 90%, 95%, 100% de homología de secuencia con la secuencia de proteína viral correspondiente. En algunas realizaciones, el polipéptido antimicrobiano puede tener al menos alrededor de 75%, 80%, 85%, 90%, 95%, o 100% de homología de secuencia con la secuencia de proteína viral correspondiente.

- 55 En otras realizaciones, el polipéptido antimicrobiano puede comprender o consistir en un péptido sintético que corresponde a una región, por ejemplo, una secuencia consecutiva de al menos alrededor de 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, o 60 aminoácidos del dominio de unión de una proteína de unión a la cápside. En algunas realizaciones, el polipéptido antimicrobiano puede tener al menos alrededor de 75%, 80%, 85%, 90%, 95%, o 100% de homología de secuencia con la secuencia correspondiente de la proteína de unión a la cápside.

- 60 Los polipéptidos antimicrobianos descritos en la presente pueden bloquear la dimerización de proteasa e inhibir la escisión de proproteínas virales (por ejemplo, procesamiento Gag-pol de VIH) en proteínas funcionales, lo que evita la liberación de uno o más virus con envoltura (por ejemplo, VIH, VHC). En algunas realizaciones, el polipéptido antimicrobiano puede tener al menos alrededor de 75%, 80%, 85%, 90%, 95%, 100% de homología de secuencia con la secuencia de proteína viral correspondiente.

- 65 En otras realizaciones, el polipéptido antimicrobiano puede comprender o consistir en un péptido sintético que corresponde a una región, por ejemplo, una secuencia consecutiva de al menos alrededor de 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, o 60 aminoácidos del dominio de unión de una proteína de unión a la proteasa. En

algunas realizaciones, el polipéptido antimicrobiano puede tener al menos alrededor de 75%, 80%, 85%, 90%, 95%, o 100% de homología de secuencia con la secuencia correspondiente de la proteína de unión a la proteasa.

- 5 Una lista no taxativa de enfermedades infecciosas que pueden tratar los antígenos de vacuna de ARN o péptidos antimicrobianos se presenta a continuación: virus de inmunodeficiencia humana (VIH), VIH que produce una infección micobacteriana, caquexia relacionada con SIDA, infección por Citomegalovirus relacionada con SIDA, nefropatía asociada con VIH, lipodistrofia, meningitis criptocócica relacionada con SIDA, neutropenia relacionada con SIDA, infecciones de *Pneumocystis jirovecii* (*Pneumocystis carinii*), toxoplasmosis
- 10 relacionada con SIDA, hepatitis A, B, C, D o E, herpes, herpes zóster (varicela), sarampión alemán (virus de rubeola), fiebre amarilla, dengue, etc. (flavivirus), (virus de influenza), enfermedades infecciosas hemorrágicas (virus de Marburgo o Ébola), enfermedades infecciosas bacterianas tales como la enfermedad del legionario (*Legionella*), úlcera gástrica (*Helicobacter*), cólera (*Vibrio*), infecciones de *E. coli*, infecciones estafilocócicas, infecciones de salmonela o infecciones estreptocócicas, tétano (*Clostridium tetani*), enfermedades infecciosas
- 15 protozoarias (malaria, tripanosomiasis africana, leishmaniasis, toxoplasmosis, es decir, infecciones provocadas por *Plasmodium*, *Trypanosoma*, *Leishmania* y *Toxoplasma*), difteria, lepra, sarampión, pertussis, rabia, tétano, tuberculosis, tifoidea, varicela, infecciones diarreicas tales como amebiasis, diarrea asociada con *Clostridium difficile* (CDAD), criptosporidiosis, Giardiasis, Ciclosporiasis y gastroenteritis por rotavirus, encefalitis tal como encefalitis japonesa, encefalitis equina de Western y encefalitis transmitida por garrapatas (TBE), enfermedades
- 20 fúngicas de la piel tales como candidiasis, onicomicosis, *Tinea capitis*/tiña del cuero cabelludo, *Tinea corporis*/tiña corporal, *Tinea cruris*/tiña inguinal, esporotricosis y *Tinea pedis*/pie de atleta, Meningitis tal como *Haemophilus influenzae* de tipo b (Hib), Meningitis, infecciones meningocócicas virales e infección neumocócica, enfermedades tropicales olvidadas tales como fiebre hemorrágica argentina, Leishmaniasis, infecciones por nemátodos/lombrices intestinales, infección por virus del río Ross y enfermedad del virus del Nilo occidental (WNV), ETS que no son de VIH tales como Tricomoniasis, infecciones del virus del papiloma humano (HPV), enfermedades de clamidia de transmisión sexual, cancroide y sífilis, infecciones bacterianas no sépticas tales como celulitis, enfermedad de Lyme, infección de SARM, *Pseudomonas*, infecciones estafilocócicas, fiebre
- 25 botanosa, Leptospirosis, fiebre reumática, botulismo, enfermedad de *Rickettsia* y mastoiditis, infecciones parasíticas tales como cisticercosis, equinococosis, infecciones por trematodos, triquinosis, babesiosis, Hipodermiasis, botriocefalosis y Tripanosomiasis, infecciones respiratorias tales como infección por adenovirus, infecciones por aspergilosis, influenza aviar (H5N1), influenza, infecciones de RSV, síndrome respiratorio agudo grave (SRAG), sinusitis, Legionelosis, Coccidioidomicosis e influenza porcina (H1N1), sepsis tal como bacteremia, sepsis/choque séptico, sepsis en infantes prematuros, infección del tracto urinario tal como infecciones vaginales (bacterianas), infecciones vaginales (fúngicas) e infección gonocócica,
- 35 enfermedades cutáneas virales tales como infecciones del parvovirus B19, verrugas, herpes genital, herpes orofacial, culebrilla, infecciones del oído interno, citomegalovirus congénito, enfermedades transmitidas por alimentos tales como brucelosis (especie *Brucella*), *Clostridium perfringens* (toxina Epsilon), *E. Coli* O157:H7 (*Escherichia coli*), Salmonelosis (especie *Salmonella*), Shigelosis (*Shigella*), Vibriosis y Listeriosis, bioterrorismo y posibles enfermedades epidémicas tales como fiebre hemorrágica del ébola, fiebre de Lassa,
- 40 fiebre hemorrágica de Marburgo, plaga, enfermedad del virus de Antrax Nipah, Hantavirus, viruela, muermo (*Burkholderia mallei*), Melioidosis (*Burkholderia pseudomallei*), Psitacosis (*Chlamydia psittaci*), fiebre Q (*Coxiella burnetii*), Tularemia (*Fancisella tularensis*), rubéola, paperas y polio.

- 45 El ARN descrito en la presente puede codificar uno o más péptidos o proteínas terapéuticas validadas o “en prueba”. Uno o más péptidos o proteínas terapéuticas que se comercializan o que están en desarrollo actualmente pueden codificarse por el ARN de la presente invención. Los péptidos o proteínas terapéuticas codificados por el ARN producido por los métodos de la invención pueden utilizarse para tratar afecciones o enfermedades en muchas áreas terapéuticas, tales como, de modo no taxativo, hematología, cardiovascular, del SNC, intoxicación (que incluye antivenenos), dermatología, endocrinología, genética, genitourinaria,
- 50 gastrointestinal, musculoesquelética, oncología, e inmunología, respiratoria, sensorial y antiinfecciosa.

- El ARN producido por los métodos de la invención puede codificar uno o más polipéptidos que penetran células. Tal como se usa en la presente, “polipéptido que penetra células” o CPP se refiere a un polipéptido que puede facilitar la absorción celular de moléculas. Un polipéptido que penetra células puede contener una o más
- 55 etiquetas detectables. Los polipéptidos pueden estar parcialmente etiquetados o completamente etiquetados en su totalidad. El ARN puede codificar la etiqueta detectable completamente, parcialmente o no codificarla en absoluto. El péptido que penetra células también puede incluir una secuencia de señal. Tal como se usa en la presente, una “secuencia de señal” se refiere a una secuencia de residuos de aminoácidos unidos en el extremo amino de una proteína naciente durante la traducción de proteínas. La secuencia de señal se puede
- 60 utilizar para señalar la secreción del polipéptido que penetra células.

- En una realización, el ARN también puede codificar una proteína de fusión. La proteína de fusión se puede crear mediante unión operativa de una proteína cargada con una proteína terapéutica. Tal como se usa en la presente, “unido de manera operativa” se refiere a la proteína terapéutica y la proteína cargada conectadas de
- 65 manera tal que permitan la expresión del complejo cuando se introduce en la célula. Tal como se usa en la presente, “proteína cargada” se refiere a una proteína que lleva una carga eléctrica positiva, negativa o neutra

en general. Preferentemente, la proteína terapéutica puede estar unida de manera covalente a la proteína cargada en la formación de la proteína de fusión. La relación de la carga de superficie con respecto a los aminoácidos totales o de superficie puede ser aproximadamente 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 o 0.9.

- 5 El polipéptido que penetra células codificado por el ARN puede formar un complejo después de haberse traducido. El complejo puede comprender una proteína cargada unida, por ejemplo, unida de manera covalente, al polipéptido que penetra células.

- 10 En una realización, el polipéptido que penetra células puede comprender un primer dominio y un segundo dominio. El primer dominio puede comprender un polipéptido supercargado. El segundo dominio puede comprender un compañero de unión a proteínas. Tal como se usa en la presente, "compañero de unión a proteínas" incluye, de modo no taxativo, anticuerpos y fragmentos funcionales de estos, proteínas de andamiaje, o péptidos. El polipéptido que penetra células puede comprender además un compañero de unión intracelular para el compañero de unión a proteínas. El polipéptido que penetra células puede secretarse de una célula donde se puede introducir el ARN. El polipéptido que penetra células también puede penetrar la primera célula.

- 20 En una realización, el ARN puede codificar un polipéptido que penetra células que comprende un compañero de unión a proteínas. El compañero de unión a proteínas puede incluir, de modo no taxativo, un anticuerpo, un anticuerpo supercargado o un fragmento funcional. El ARN puede introducirse en la célula donde se introduce un polipéptido que penetra células que comprende el compañero de unión a proteínas.

- 25 Las células humanas y otras células eucariotas se subdividen mediante membranas en compartimientos funcionalmente distintos. Cada compartimiento unido a la membrana, u orgánulo, contiene diferentes proteínas esenciales para la función del orgánulo. La célula usa "señales de clasificación" que son motivos de aminoácidos ubicados dentro de la proteína, para direccionar proteínas a orgánulos celulares particulares. Un tipo de señal de clasificación, llamada secuencia de señal, péptido señal, o secuencia líder, dirige una clase de proteínas a un orgánulo llamado retículo endoplasmático (RE).

- 30 Las proteínas direccionadas al ER mediante una secuencia de señal pueden liberarse en el espacio extracelular como una proteína secretada. De manera similar, las proteínas que residen en la membrana celular también pueden secretarse en el espacio extracelular mediante escisión proteolítica de un "enlazador" que sostiene la proteína a la membrana. Sin pretender limitarse a la teoría, las moléculas de la presente invención se pueden utilizar para explotar el tránsito celular descrito anteriormente. Como tal, en algunas realizaciones de la invención, se proporciona ARN para expresar una proteína secretada. En una realización, esta se puede utilizar en la fabricación de grandes cantidades de productos génicos humanos valiosos.

- 40 En algunas realizaciones de la invención, se proporciona ARN para expresar una proteína de la membrana plasmática.

- En algunas realizaciones de la invención, se proporciona ARN para expresar una proteína citoplasmática o citoesquelética.

- 45 En algunas realizaciones de la invención, se proporciona ARN para expresar una proteína unida a la membrana intracelular.

En algunas realizaciones de la invención, se proporciona ARN para expresar una proteína nuclear.

- 50 En algunas realizaciones de la invención, se proporciona ARN para expresar una proteína asociada a una enfermedad humana.

- 55 El ARN puede tener una secuencia de nucleótidos de un ARN nativo o de origen natural o que codifica un péptido nativo o de origen natural. De manera alternativa, el ARN puede tener una secuencia de nucleótidos que tiene un porcentaje de identidad con la secuencia de nucleótidos de un ARN o nativo o de origen natural o el ARNm puede tener una secuencia de nucleótidos que codifica un péptido que tiene un porcentaje de identidad con la secuencia de nucleótidos de un péptido nativo o de origen natural. Tal como se conoce en la técnica, el término "identidad" se refiere a una relación entre las secuencias de dos o más péptidos, determinada mediante la comparación de las secuencias. En la técnica, la identidad también hace referencia al grado de relación entre péptidos según se determina por la cantidad de coincidencias entre hebras de dos o más residuos de aminoácidos. La identidad mide el porcentaje de coincidencias idénticas entre la menor de dos o más secuencias con alineaciones de espacios (si los hubiera) conforme a un modelo matemático o programa informático particular (es decir, "algoritmos"). Es posible calcular la identidad de péptidos relacionados mediante métodos conocidos. Dichos métodos incluyen, de modo no taxativo, los descritos en Computational Molecular Biology, Lesk, A. M., ed., Oxford University Press, Nueva York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W., ed., Academic Press, Nueva York, 1993; Computer Analysis of Sequence Data, Parte 1, Griffin, A. M., y Griffin, H. G., eds., Humana Press, Nueva Jersey, 1994; Sequence



Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; Sequence Analysis Primer, Gribskov, M. y Devereux, J., eds., M. Stockton Press, Nueva York, 1991; y Carillo et ál., SIAM J. Applied Math. 48, 1073 (1988).

- 5 Por lo tanto, en algunas realizaciones, los péptidos codificados por los ARN son variantes de polipéptidos que pueden tener la misma actividad o una actividad similar que el polipéptido de referencia. De manera alternativa, la variante puede tener una actividad alterada (por ejemplo, mayor o menor) con respecto a un polipéptido de referencia. En general, las variantes de un polinucleótido o polipéptido particular tendrán al menos 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, pero menos de 100 % de identidad de secuencia con dicho polinucleótido o polipéptido de referencia particular, según se determina mediante parámetros y programas de alineación de secuencia descritos en la presente y conocidos por los expertos en la técnica. Dichas herramientas de alineación incluyen las del paquete BLAST (Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, y David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.) Se describen otras herramientas en la presente, específicamente en la definición de "identidad". Los parámetros por defecto en el algoritmo BLAST incluyen, por ejemplo, un umbral esperado de 10, tamaño de palabra de 28, resultados de coincidencia/no coincidencia 1, -2, costos de espacio Lineal. Se puede aplicar cualquier filtro así como también una selección para repeticiones específicas de especies, por ejemplo, Homo sapiens.

- 20 De acuerdo con la presente invención, los polinucleótidos incluyen ARN para codificar uno o más polipéptidos de interés o fragmentos de estos. Un polipéptido de interés puede incluir, de modo no taxativo, polipéptidos enteros, múltiples polipéptidos o fragmentos de polipéptidos. Tal como se usa en la presente, el término "polipéptidos de interés" hacen referencia a cualquier polipéptido que se selecciona para codificarse en la construcción primaria producida por los métodos de la invención. Tal como se usa en la presente, "polipéptido" significa un polímero de residuos de aminoácidos (naturales o no naturales) enlazados juntos a menudo por enlaces peptídicos. El término, tal como se usa en la presente, se refiere a proteínas, polipéptidos y péptidos de cualquier tamaño, estructura o función. En algunos casos, el polipéptido codificado tiene menos de alrededor de 50 aminoácidos y, por lo tanto, se denomina un péptido. Si el polipéptido es un péptido, tendrá una longitud de al menos alrededor de 2, 3, 4 o al menos 5 residuos de aminoácidos. Por lo tanto, los polipéptidos incluyen productos génicos, polipéptidos de origen natural, polipéptidos sintéticos, homólogos, ortólogos, parálogos, fragmentos y otros equivalentes, variantes y análogos de los que anteceden. Un polipéptido puede ser una única molécula o puede ser un complejo multimolecular, tal como un dímero, trímero o tetramero. Estos también pueden comprender polipéptidos de cadena simple o de cadena múltiple, tales como anticuerpos o insulina, y pueden estar asociados o unidos. Las uniones disulfuro más comunes se encuentran en los polipéptidos de cadena múltiple. El término polipéptido también puede aplicarse a polímeros de aminoácidos, en los cuales uno o más residuos de aminoácidos son un análogo químico artificial de un aminoácido correspondiente de origen natural.

- 40 El término "variante de polipéptido" se refiere a moléculas que difieren en su secuencia de aminoácidos de una secuencia de referencia o natural. Las variantes de secuencia de aminoácidos pueden poseer sustituciones, eliminaciones y/o inserciones en determinadas posiciones dentro de la secuencia de aminoácidos, en comparación con una secuencia de referencia o natural. Normalmente, las variantes tendrán al menos alrededor de 50 % de identidad con una secuencia de referencia o natural y, preferentemente, serán idénticas en al menos alrededor de 80 %, más preferentemente al menos alrededor de 90 % a una secuencia de referencia o natural.

- Tal como se usa en la presente, el término "mimético variante" es uno que contiene uno o más aminoácidos que imitarían una secuencia activada. Por ejemplo, el glutamato puede servir como un mimético de fosfotreonina y/o fosfoserina. De manera alternativa, los miméticos variantes pueden producir la desactivación o un producto inactivado que contiene el mimético, por ejemplo, la fenilalanina puede actuar como una sustitución inactivante de la tirosina, o la alanina puede actuar como una sustitución inactivamente de la serina.

- 55 La presente descripción contempla varios tipos de composiciones que están basados en polipéptidos, inclusive variantes y derivados. Estos incluyen variantes y derivados de sustitución, inserción, eliminación y covalentes. El término "derivado" se usa como sinónimo del término "variante", pero generalmente se refiere a una molécula que ha sido modificada y/o cambiada de cualquier manera en comparación con una molécula de referencia o molécula inicial.

- 60 Como tal, el ARN que codifica polipéptidos contiene sustituciones, inserciones y/o adiciones, eliminaciones y modificaciones covalentes con respecto a las secuencias de referencia, en particular las secuencias de polipéptidos descritas en la presente están incluidas dentro del alcance de la presente invención. Por ejemplo, los marcadores de secuencia o aminoácidos, tales como una o más lisinas, pueden agregarse a las secuencias de péptidos (por ejemplo, en los extremos N o C). Pueden utilizarse marcadores de secuencia para la purificación o localización de péptidos. Pueden utilizarse lisinas para aumentar la solubilidad de péptidos o para

permitir la biotilación. De manera alternativa, los residuos de aminoácidos ubicados en las regiones de los extremos amino y carboxi de la secuencia de aminoácidos de un péptido o proteína pueden eliminarse de manera opcional para proporcionar secuencias truncadas. Determinados aminoácidos (por ejemplo, residuos del extremo C o extremo N) pueden eliminarse de manera alternativa, dependiendo del uso de la secuencia, por ejemplo, como expresión de la secuencia como parte de una secuencia mayor que sea soluble, o que esté unida a un soporte sólido.

Las "variantes de sustitución", cuando se refieren a polipéptido, son las que tienen eliminado al menos un residuo de aminoácidos en una secuencia natural o inicial y un aminoácido diferente insertado en su lugar en la misma posición. Las sustituciones pueden ser simples, donde solo se sustituyó una molécula de aminoácido, o pueden ser múltiples, cuando se sustituyeron dos o más aminoácidos en la misma molécula.

Tal como se usa en la presente, el término "sustitución de aminoácido conservadora" se refiere a la sustitución de un aminoácido que está presente normalmente en la secuencia con un aminoácido diferente de tamaño, carga o polaridad similar. Los ejemplos de sustituciones conservadoras incluyen la sustitución de un residuo no polar (hidrofóbico), tal como isoleucina, valina y leucina, con otro residuo no polar. De manera similar, los ejemplos de sustituciones conservadoras incluyen la sustitución de un residuo polar (hidrofílico) por otro, tal como entre arginina y lisina, entre glutamina y asparagina, y entre glicina y serina. De manera adicional, la sustitución de un residuo básico, tal como lisina, arginina o histidina, por otro, o la sustitución de un residuo ácido, tal como ácido aspártico o ácido glutámico por otro residuo ácido, son ejemplos adicionales de sustituciones conservadoras. Los ejemplos de sustituciones no conservadoras incluyen la sustitución de un residuo aminoácido no polar (hidrofóbico), tal como isoleucina, valina, leucina, alanina, metionina, con un residuo polar (hidrofílico), tal como cisteína, glutamina, ácido glutámico o lisina, y/o un residuo polar con un residuo no polar.

"Variantes de inserción", cuando se refiere a polipéptidos, son las que tienen uno o más aminoácidos insertados de manera inmediatamente adyacente a un aminoácido en una posición particular en una secuencia natural o inicial. "Inmediatamente adyacente" a un aminoácido significa conectado al grupo funcional alfa-carboxi o alfa-amino del aminoácido.

"Variaciones de eliminación", cuando se refiere a polipéptidos, son las que tienen uno o más aminoácidos eliminados en la secuencia de aminoácidos natural o inicial. De manera común, las variantes de eliminación tendrán uno o más aminoácidos eliminados en una región particular de la molécula.

"Derivados covalentes", cuando se refiere a polipéptidos, incluyen modificaciones de una proteína natural o inicial con un agente de derivación proteínico o no proteínico orgánico, y/o modificaciones postraduccionales. Las modificaciones covalentes se introducen tradicionalmente mediante la reacción de residuos de aminoácidos dirigidos de la proteína con un agente de derivación orgánico que es capaz de reaccionar con residuos terminales o cadenas laterales seleccionadas, o mediante mecanismos de aprovechamiento de modificaciones postraduccionales que funcionan en células hospedadoras recombinantes seleccionadas. Los derivados covalentes resultantes son útiles en programas dirigidos a identificar residuos importantes para verificar la actividad biológica, para inmunoensayos, o para la preparación de anticuerpos contra proteínas para la purificación por inmunoafinidad de la glicoproteína recombinante. Dichas modificaciones se encuentran dentro del conocimiento de la técnica y se llevan a cabo sin experimentación innecesaria.

Determinadas modificaciones postraduccionales son el resultado de la acción células hospedadoras recombinantes sobre el polipéptido expresado. Los residuos de glutaminilo y asparaginilo a menudo se desamidán de forma postraduccional y se convierten en los residuos aspartilo y glutamilo correspondientes. De manera alternativa, estos residuos se desamidán en condiciones ligeramente ácidas. Cualquier forma de estos residuos puede estar presente en los polipéptidos producidos de acuerdo con la presente descripción.

Otras modificaciones postraduccionales incluyen hidroxilación de prolina y lisina, fosforilación de grupos hidroxilo de residuos de serilo y treonilo, metilación de los grupos alfa-amino de cadenas laterales de histidina, lisina y arginina (T. E. Creighton, *Proteins: Structure and Molecular Properties*, W.H. Freeman & Co., San Francisco, páginas 79-86 (1983)).

Tal como se usa en la presente al hacer referencia a polipéptidos, el término "dominio" se refiere a un motivo de un polipéptido que tenga una o más características o propiedades funcionales o estructurales identificables (por ejemplo, capacidad de unión, servir como sitio para las interacciones entre proteínas).

Tal como se usa en la presente en referencia a polipéptidos, el término "sitio", en lo referente a casos basados en aminoácidos, se usa como sinónimo de "residuo aminoácido" y "cadena lateral de aminoácido". Un sitio representa una posición dentro de un péptido o polipéptido que puede modificarse, manipularse, alterarse, derivarse o variarse dentro de las moléculas basadas en polipéptidos de la presente descripción.

Tal como se usa en la presente los términos "extremos" o "extremo", en referencia a polipéptidos, se refieren a una extremidad de un péptido o polipéptido. Dicha extremidad no se limita únicamente al sitio inicial o final del péptido o polipéptido, sino que puede incluir aminoácidos adicionales en las regiones de los extremos. Las moléculas basadas en polipéptidos de la presente descripción pueden caracterizarse como moléculas que tienen tanto un extremo N (terminado en un aminoácido con un grupo amino libre (NH<sub>2</sub>)) como un extremo C (terminado en un aminoácido con un grupo carboxilo libre (COOH)). En algunos casos, las proteínas de la descripción están compuestas de cadenas de polipéptidos que están unidas mediante enlaces disulfuro o mediante fuerzas no covalentes (multímeros, oligómeros). Estos tipos de proteínas tendrán múltiples extremos N y C. De manera alternativa, se pueden modificar los extremos de los polipéptidos de manera que comiencen o terminen, según sea el caso, con un resto no basado en polipéptidos, tal como un conjugado orgánico.

#### Formulaciones/composiciones farmacéuticas

La presente descripción proporciona polinucleótidos y composiciones farmacéuticas de estos opcionalmente en combinación con uno o más excipientes farmacéuticamente aceptables. Las composiciones farmacéuticas pueden comprender opcionalmente una o más sustancias activas adicionales, por ejemplo, sustancias terapéuticamente y/o profilácticamente activas. Las composiciones farmacéuticas de la descripción pueden ser estériles y/o libres de pirógenos. Es posible encontrar consideraciones generales de formulación y/o producción de agentes farmacéuticos, por ejemplo, en *Remington: The Science and Practice of Pharmacy* 21.<sup>a</sup> ed., Lippincott Williams & Wilkins, 2005.

En algunos casos, las composiciones se administran a humanos, sujetos o pacientes humanos. A los efectos de la presente descripción, la frase "ingrediente activo" generalmente se refiere a polinucleótidos, por ejemplo, ARNm que codifica polinucleótidos que se administrará tal como se describe en la presente.

Las formulaciones de las composiciones farmacéuticas descritas en la presente pueden prepararse mediante cualquier método conocido o que se desarrolle posteriormente en la técnica de la farmacología. En general, tales métodos de preparación incluyen la etapa de asociación del ingrediente activo con un excipiente y/o uno o más ingredientes accesorios y luego, si fuera necesario y/o deseable, la división, formación y/o empaque del producto en una unidad de dosis única o de dosis múltiples deseada.

A<producidos por los métodos de la invención>A

Las cantidades relativas del ingrediente activo, el excipiente farmacéuticamente aceptable y/o cualquier ingrediente adicional en una composición farmacéutica de acuerdo con la descripción variarán conforme a la identidad, tamaño y/o afección del sujeto tratado y también conforme a la vía por la que se administrará la composición. A modo de ejemplo, la composición puede comprender entre 0,1 % y 100 %, por ejemplo, entre 0,5 y 50 %, entre 1-30 %, entre 5-80 %, al menos 80 % (p/p) de ingrediente activo.

Los polinucleótidos A<>A pueden formularse utilizando uno o más excipientes para: (1) aumentar la estabilidad, (2) aumentar la transfección de células, (3) permitir la liberación sostenida o demorada (por ejemplo, a partir de una formulación de depósito), (4) alterar la biodistribución (por ejemplo, objetivo para tipos de células o tejidos específicos), (5) aumentar la traducción de proteínas codificadas in vivo y (6) alterar el perfil de liberación de proteínas codificadas in vivo. Además de los excipientes tradicionales, tales como cualquier solvente, medios de dispersión, diluyentes u otros vehículos líquidos, auxiliares de dispersión o suspensión, agentes tensioactivos, agentes isotónicos, agentes espesantes o emulsionantes, conservantes, los excipientes de la presente invención pueden incluir, de modo no taxativo, lipídicos, liposomas, nanopartículas lipídicas, polímeros, lipoplejos, nanopartículas de cápside, péptidos, proteínas, células transfectadas con polinucleótidos, hialuronidasa, miméticos de nanopartículas y combinaciones de estos.

En una realización, los polinucleótidos pueden formularse en un complejo de policationes-lípidos. Es posible lograr la formación del complejo de policationes-lípidos mediante métodos conocidos en la técnica. A modo de ejemplo no taxativo, el polication puede incluir un polipéptido o péptido catiónico, tal como, de modo no taxativo, polilisina, poliornitina y/o poliarginina. En otro caso, los polinucleótidos pueden formularse en un complejo de policationes-lípidos que puede incluir además un lípido no catiónico tal como, de modo no taxativo, colesterol o fosfatidiletanolamina de dioleilo (DOPE).

Una formulación de liposomas puede verse afectada, de modo no taxativo, por la selección del componente lipídico catiónico, el grado de saturación lipídica catiónica, la naturaleza de la pegilación, la relación de todos los componentes y los parámetros biofísicos tales como el tamaño. En un ejemplo de Semple et ál. (Semple et ál. *Nature Biotech.* 2010 28:172-176), la formulación de liposoma se compone de 57,1 % de lípido catiónico, 7,1 % de dipalmitoilfosfatidilcolina, 34,3 % de colesterol y 1,4 % de PEG-c-DMA. Como ejemplo adicional, la alteración de la composición del lípido catiónico puede administrar ARNip en forma más eficaz a diversas células que presentan antígenos (Basha et ál. *Mol Ther.* 2011 19:2186-2200). En algunas realizaciones, las formulaciones de liposomas pueden comprender de alrededor de 35 a alrededor de 45 % de lípido catiónico, de alrededor de 40 % a alrededor de 50 % de lípido catiónico, de alrededor de 50 % a alrededor de 60 % de

lípidos catiónicos y/o de alrededor de 55 % a alrededor de 65 % de lípidos catiónicos. En algunas realizaciones, la relación de lípidos con respecto a ARN en liposomas puede ser de alrededor de 5:1 a alrededor de 20:1, de alrededor de 10:1 a alrededor de 25:1, de alrededor de 15:1 a alrededor de 30:1 y/o al menos 30:1.

- 5 En algunas realizaciones, la relación de PEG en las formulaciones de nanopartículas lipídicas (LNP) puede aumentar o reducirse y/o la longitud de la cadena de carbono del lípido con PEG puede modificarse de C14 a C18 para alterar la farmacocinética y/o la biodistribución de las formulaciones de LNP. Como ejemplo no taxativo, las formulaciones de LNP pueden contener de alrededor de 0,5 % a alrededor de 3,0 %, de alrededor de 1,0 % a alrededor de 3,5 %, de alrededor de 1,5 % a alrededor de 4,0 %, de alrededor de 2,0 % a alrededor de 4,5 %, de alrededor de 2,5 % a alrededor de 5,0 % y/o de alrededor de 3,0 % a alrededor de 6,0 % de la relación molar lipídica de PEG-c-DOMG (R-3-[(ω-metoxi-poli(etilenglicol)2000)carbamoil])-1,2-dimiristiloxipropil-3-amina) (también denominada en la presente PEG-DOMG) en comparación con el lípido catiónico, DSPC y colesterol. En otra realización, es posible reemplazar PEG-c-DOMG con un lípido con PEG tal como, de modo no taxativo, PEG-DSG (1,2-distearoil-sn-glicerol, metoxipoli(etilenglicol)), PEG-DMG (1,2-dimiristiloxil-sn-glicerol) y/o PEG-DPG (1,2-dipalmitoil-sn-glicerol, metoxipoli(etilenglicol)). Puede seleccionarse el lípido catiónico de cualquier lípido conocido en la técnica tal como, de modo no taxativo, DLin-MC3-DMA, DLin-DMA, C12-200 y DLin-KC2-DMA.

- 20 En una realización, el polinucleótido se formula en una nanopartícula que puede comprender al menos un lípido. El lípido se puede seleccionar, de modo no taxativo, de DLin-DMA, Dlin-K-DMA, 98N12-5, C12-200, DLin-MC3-DMA, DLin-KC2-DMA, DODMA, PLGA, PEG, PEG-DMG, lípidos pegilados y lípidos de aminoalcoholes. En otro aspecto, el lípido puede ser un lípido catiónico tal como, de modo no taxativo, DLin-DMA, DLin-D-DMA, DLin-MC3-DMA, DLin-KC2-DMA, DODMA y lípidos de aminoalcoholes. El lípido catiónico de aminoalcohol puede ser cualquiera de los lípidos descritos y/o producidos mediante los métodos descritos en la publicación de patente estadounidense n.º US20130150625. Como ejemplo no taxativo, el lípido catiónico puede ser 2-amino-3-[(9Z,12Z)-octadeca-9,12-dien-1-iloxi]-2-[(9Z,2Z)-octadeca-9,12-dien-1-iloxi]metil}propan-1-ol (Compuesto 1 en US20130150625), 2-amino-3-[(9Z)-octadec-9-en-1-iloxi]-2-[(9Z)-octadec-9-en-1-iloxi]metil}propan-1-ol (Compuesto 2 en US20130150625), 2-amino-3-[(9Z,12Z)-octadeca-9,12-dien-1-iloxi]-2-[(9Z,12Z)-octadeca-9,12-dien-1-iloxi]metil}propan-1-ol (Compuesto 3 en US20130150625) y 2-(dimetilamino)-3-[(9Z,12Z)-octadeca-9,12-dien-1-iloxi]-2-[(9Z,12Z)-octadeca-9,12-dien-1-iloxi]metil}propan-1-ol (Compuesto 4 en US20130150625), o cualquier sal o estereoisómero farmacéuticamente aceptable de estos.

- Usualmente, las formulaciones de nanopartículas lipídicas comprenden un lípido, en particular, un lípido catiónico ionizable, por ejemplo, 2,2-dilinoil-4-dimetilaminoetil-[1,3]-dioxolano (Dlin-KC2-DMA), dilinoil-metil-4-dimetilaminobutirato (Dlin-MC3-DMA) o di((Z)-non-2-en-1-il) 9-((4-(dimetilamino)butanoil)oxi)heptadecanodioato (L319), y además comprende un lípido neutro, un esteroide y una molécula capaz de reducir la agregación de partículas, por ejemplo, un lípido PEG o modificado por PEG.

- 40 En una realización, una formulación de nanopartículas lipídicas consiste esencialmente en (i) al menos un lípido seleccionado del grupo que consiste en 2,2-dilinoil-4-dimetilaminoetil-[1,3]-dioxolano (Dlin-KC2-DMA), dilinoil-metil-4-dimetilaminobutirato (Dlin-MC3-DMA) y 9-((4-(dimetilamino)butanoil)oxi)heptadecanodioato de di((Z)-non-2-en-1-il) (L319), (ii) un lípido neutro seleccionado de DSPC, DPPC, POPC, DOPE y SM, (iii) un esteroide, por ejemplo, colesterol, y (iv) un lípido PEG, por ejemplo, PEG-DMG o PEG-cDMA, en una relación molar de 20-60 % lípido catiónico: 5-25 % de lípido neutro: 25-55 % de esteroide, 0,5-15 % de lípido PEG.

- 45 En una realización, la formulación incluye de alrededor de 25 % a alrededor de 75 % en base molar de un lípido catiónico seleccionado de 2,2-dilinoil-4-dimetilaminoetil-[1,3]-dioxolano (Dlin-KC2-DMA), dilinoil-metil-4-dimetilaminobutirato (Dlin-MC3-DMA) y di((Z)-non-2-en-1-il) 9-((4-(dimetilamino)butanoil)oxi)heptadecanodioato (L319), por ejemplo, de alrededor de 35 a alrededor de 65 %, de alrededor de 45 a alrededor de 65 %, alrededor de 60 %, alrededor de 57,5 %, alrededor de 50 % o alrededor de 40 % en base molar.

- 55 En una realización, la formulación incluye de alrededor de 0,5 % a alrededor de 15 % en base molar del lípido neutro, por ejemplo, de alrededor de 3 a alrededor de 12 %, de alrededor de 5 a alrededor de 10 % o alrededor de 15 %, alrededor de 10 % o alrededor de 7,5 % en base molar. Los lípidos neutros de ejemplo incluyen, de modo no taxativo, DSPC, POPC, DPPC, DOPE y SM. En una realización, la formulación incluye de alrededor de 5 % a alrededor de 50 % en base molar del esteroide (por ejemplo, alrededor de 15 a alrededor de 45 %, alrededor de 20 a alrededor de 40 %, alrededor de 40 %, alrededor de 38,5 %, alrededor de 35 %, o alrededor de 31 % en base molar. Un esteroide de ejemplo es el colesterol. En una realización, la formulación incluye de alrededor de 0,5 % a alrededor de 20 % en base molar del PEG o lípido modificado por PEG (por ejemplo, alrededor de 0,5 a alrededor de 10 %, alrededor de 0,5 a alrededor de 5 %, alrededor de 1,5 %, alrededor de 0,5 %, alrededor de 1,5 %, alrededor de 3,5 %, o alrededor de 5 % en base molar. En una realización, el PEG o lípido modificado por PEG comprende una molécula de PEG de peso molecular promedio de 2000 Da. En otras realizaciones, el PEG o lípido modificado por PEG comprende una molécula de PEG de peso molecular promedio menor de 2000, por ejemplo, alrededor de 1500 Da, alrededor de 1000 Da o alrededor de 500 Da. Los lípidos modificados por PEG de ejemplo incluyen, de modo no taxativo, PEG-distearoil glicerol (PEG-

DMG) (también denominado en la presente PEG-C14 o C14-PEG), PEG-cDMA.

En una realización, las formulaciones incluyen 25-75 % de un lípido catiónico seleccionado de 2,2-dilinoil-4-dimetilaminoetil-[1,3]-dioxolano (DLin-KC2-DMA), dilinoil-metil-4-dimetilaminobutirato (DLin-MC3-DMA) y di((Z)-non-2-en-1-il) 9-((4-(dimetilamino)butanoil)oxi)heptadecanodioato (L319), 0,5-15 % del lípido neutro, 5-50 % del esteroil y 0,5-20 % del PEG o lípido modificado por PEG en una base molar.

En una realización, las formulaciones incluyen 35-65% de un lípido catiónico seleccionado de 2,2-dilinoil-4-dimetilaminoetil-[1,3]-dioxolano (DLin-KC2-DMA), dilinoil-metil-4-dimetilaminobutirato (DLin-MC3-DMA) y di((Z)-non-2-en-1-il) 9-((4-(dimetilamino)butanoil)oxi)heptadecanodioato (L319), 3-12% del lípido neutro, 15-45% del esteroil y 0.5-10% del PEG o lípido modificado por PEG en una base molar.

En una realización, las formulaciones incluyen 45-65% de un lípido catiónico seleccionado de 2,2-dilinoil-4-dimetilaminoetil-[1,3]-dioxolano (DLin-KC2-DMA), dilinoil-metil-4-dimetilaminobutirato (DLin-MC3-DMA) y di((Z)-non-2-en-1-il) 9-((4-(dimetilamino)butanoil)oxi)heptadecanodioato (L319), 5-10% del lípido neutro, 25-40% del esteroil y 0.5-10% del PEG o lípido modificado por PEG en una base molar.

En una realización, las formulaciones incluyen alrededor de 60 % de un lípido catiónico seleccionado de 2,2-dilinoil-4-dimetilaminoetil-[1,3]-dioxolano (DLin-KC2-DMA), dilinoil-metil-4-dimetilaminobutirato (DLin-MC3-DMA) y di((Z)-non-2-en-1-il) 9-((4-(dimetilamino)butanoil)oxi)heptadecanodioato (L319), alrededor de 7,5 % del lípido neutro, alrededor de 31 % del esteroil y alrededor de 1,5 % del PEG o lípido modificado por PEG en una base molar.

En una realización, las formulaciones de las invenciones incluyen alrededor de 50% de un lípido catiónico seleccionado de 2,2-dilinoil-4-dimetilaminoetil-[1,3]-dioxolano (DLin-KC2-DMA), dilinoil-metil-4-dimetilaminobutirato (DLin-MC3-DMA) y di((Z)-non-2-en-1-il) 9-((4-(dimetilamino)butanoil)oxi)heptadecanodioato (L319), alrededor de 10% del lípido neutro, alrededor de 38,5 % del esteroil y alrededor de 1,5 % del PEG o lípido modificado por PEG en una base molar.

En una realización, las formulaciones incluyen alrededor de 50 % de un lípido catiónico seleccionado de 2,2-dilinoil-4-dimetilaminoetil-[1,3]-dioxolano (DLin-KC2-DMA), dilinoil-metil-4-dimetilaminobutirato (DLin-MC3-DMA) y di((Z)-non-2-en-1-il) 9-((4-(dimetilamino)butanoil)oxi)heptadecanodioato (L319), alrededor de 10 % del lípido neutro, alrededor de 35 % del esteroil, alrededor de 4,5 % o alrededor de 5 % del PEG o lípido modificado por PEG, y alrededor de 0,5 % del lípido de direccionamiento en una base molar.

En una realización, las formulaciones incluyen alrededor de 40% de un lípido catiónico seleccionado de 2,2-dilinoil-4-dimetilaminoetil-[1,3]-dioxolano (DLin-KC2-DMA), dilinoil-metil-4-dimetilaminobutirato (DLin-MC3-DMA) y di((Z)-non-2-en-1-il) 9-((4-(dimetilamino)butanoil)oxi)heptadecanodioato (L319), alrededor de 15% del lípido neutro, alrededor de 40% del esteroil y alrededor de 5% del PEG o lípido modificado por PEG en una base molar.

En una realización, las formulaciones incluyen alrededor de 57.2% de un lípido catiónico seleccionado de 2,2-dilinoil-4-dimetilaminoetil-[1,3]-dioxolano (DLin-KC2-DMA), dilinoil-metil-4-dimetilaminobutirato (DLin-MC3-DMA) y di((Z)-non-2-en-1-il) 9-((4-(dimetilamino)butanoil)oxi)heptadecanodioato (L319), alrededor de 7,1% del lípido neutro, alrededor de 34.3% del esteroil y alrededor de 1.4% del PEG o lípido modificado por PEG en una base molar.

En una realización, las formulaciones incluyen alrededor de 57,5 % de un lípido catiónico seleccionado del lípido de PEG es PEG-cDMA (PEG-cDMA se trata adicionalmente en Reyes et ál. (J. Controlled Release, 107, 276-287 (2005), cuyo contenido se incorpora a la presente en su totalidad mediante esta referencia), alrededor de 7,5 % del lípido neutro, alrededor de 31,5 % del esteroil y alrededor de 3,5 % del PEG o lípido modificado por PEG en base molar.

En realizaciones preferidas, la formulación de nanopartículas lipídicas consiste esencialmente en una mezcla lipídica en relaciones molares de 20-70 % lípido catiónico: 5-45% de lípido neutro: 20-55 % de colesterol: 0,5-15 % de lípido modificado por PEG, más preferentemente en una relación molar de alrededor de 20-60 % lípido catiónico: 5-25 % de lípido neutral: 25-55 % de colesterol: 0,5-15 % lípido modificado por PEG.

En realizaciones particulares, la relación lipídica molar es 50/10/38,5/1,5 (% molar lípido catiónico/lípido neutro, por ejemplo, DSPC/Col/lípido modificado por PEG, por ejemplo, PEG-DMG, PEG-DSG o PEG-DPG), 57,2/7,1/34,3/1,4 (% molar lípido catiónico/lípido neutro, por ejemplo, DPPC/Col/ lípido modificado por PEG, por ejemplo, PEG-cDMA), 40/15/40/5 (% molar lípido catiónico/lípido neutro, por ejemplo, DSPC/Col/ lípido modificado por PEG, por ejemplo, PEG-DMG), 50/10/35/4,5/0,5 (% molar lípido catiónico/lípido neutro, por ejemplo, DSPC/Col/lípido modificado por PEG, por ejemplo, PEG-DSG), 50/10/35/5 (lípido catiónico/lípido neutro, por ejemplo, DSPC/Col/lípido modificado por PEG, por ejemplo, PEG-DMG), 40/10/40/10 (% molar lípido catiónico/lípido neutro, por ejemplo, DSPC/Col/lípido modificado por PEG, por ejemplo, PEG-DMG o

PEG-cDMA), 35/15/40/10 (% molar lípido catiónico/lípido neutro, por ejemplo, DSPC/Col/lípido modificado por PEG, por ejemplo, PEG-DMG o PEG-cDMA) o 52/13/30/5 (% molar lípido catiónico/lípido neutro, por ejemplo, DSPC/Col/lípido modificado por PEG, por ejemplo, PEG-DMG o PEG-cDMA).

- 5 Las composiciones de nanopartículas lipídicas y métodos de ejemplo para producirlas se describen, por ejemplo, en Semple et ál. (2010) *Nat. Biotechnol.* 28:172-176; Jayarama et ál. (2012), *Angew. Chem. Int. Ed.*, 51: 8529–8533; y Maier et ál. (2013) *Molecular Therapy* 21, 1570-1578).

10 En una realización, las formulaciones de nanopartículas lipídicas descritas en la presente pueden comprender un lípido catiónico, un lípido de PEG y un lípido estructural, y opcionalmente comprenden un lípido no catiónico. Como ejemplo no taxativo, la nanopartícula lipídica puede comprender alrededor de 40-60 % de lípido catiónico, alrededor de 5-15 % de un lípido no catiónico, alrededor de 1-2 % de un lípido de PEG y alrededor de 30-50 % de un lípido estructural. Como ejemplo no taxativo adicional, la nanopartícula lipídica puede comprender alrededor de 50 % de lípido catiónico, alrededor de 10 % de un lípido no catiónico, alrededor de 1,5 % de un lípido de PEG y alrededor de 38,5 % de un lípido estructural. Incluso como otro ejemplo no taxativo, la nanopartícula lipídica puede comprender alrededor de 55 % de lípido catiónico, alrededor de 10 % de un lípido no catiónico, alrededor de 2,5 % de un lípido de PEG y alrededor de 32,5 % de un lípido estructural. En una realización, el lípido catiónico puede ser cualquier lípido catiónico descrito en la presente, tal como, de modo no taxativo, Dlin-KC2-DMA, Dlin-MC3-DMA y L319.

20 En una realización, las formulaciones de nanopartículas lipídicas descritas en la presente pueden ser nanopartículas lipídicas de 4 componentes. La nanopartícula lipídica puede comprender un lípido catiónico, un lípido no catiónico, un lípido PEG y un lípido estructural. Como ejemplo no taxativo, la nanopartícula lipídica puede comprender alrededor de 40-60 % de lípido catiónico, alrededor de 5-15 % de un lípido no catiónico, alrededor de 1-2 % de un lípido de PEG y alrededor de 30-50 % de un lípido estructural. Como ejemplo no taxativo adicional, la nanopartícula lipídica puede comprender alrededor de 50 % de lípido catiónico, alrededor de 10 % de un lípido no catiónico, alrededor de 1,5 % de un lípido de PEG y alrededor de 38,5 % de un lípido estructural. Incluso como otro ejemplo no taxativo, la nanopartícula lipídica puede comprender alrededor de 55 % de lípido catiónico, alrededor de 10 % de un lípido no catiónico, alrededor de 2,5 % de un lípido de PEG y alrededor de 32,5 % de un lípido estructural. En un caso, el lípido catiónico puede ser cualquier lípido catiónico descrito en la presente, tal como, de modo no taxativo, Dlin-KC2-DMA, Dlin-MC3-DMA y L319.

35 En una realización, las formulaciones de nanopartículas lipídicas descritas en la presente pueden comprender un lípido catiónico, un lípido no catiónico, un lípido de PEG y un lípido estructural. Como ejemplo no taxativo, la nanopartícula lipídica comprende alrededor de 50 % del lípido catiónico Dlin-KC2-DMA, alrededor de 10 % del lípido no catiónico DSPC, alrededor de 1,5 % del lípido de PEG PEG-DOMG y alrededor de 38,5 % del lípido estructural colesterol. Como ejemplo no taxativo, la nanopartícula lipídica comprende alrededor de 50 % del lípido catiónico Dlin-MC3-DMA, alrededor de 10 % del lípido no catiónico DSPC, alrededor de 1.5% del lípido de PEG PEG-DOMG y alrededor de 38,5 % del lípido estructural colesterol. Como ejemplo no taxativo, la nanopartícula lipídica comprende alrededor de 50 % del lípido catiónico Dlin-MC3-DMA, alrededor de 10 % del lípido no catiónico DSPC, alrededor de 1.5% del lípido de PEG PEG-DMG y alrededor de 38,5 % del lípido estructural colesterol. Incluso como otro ejemplo no taxativo, la nanopartícula lipídica comprende alrededor de 55 % del lípido catiónico L319, alrededor de 10 % del lípido no catiónico DSPC, alrededor de 2,5 % del lípido con PEG PEG-DMG y alrededor de 32,5 % del lípido estructural colesterol.

45 En una realización, los polinucleótidos producidos por los métodos de la invención se pueden formular en nanopartículas lipídicas con un diámetro de alrededor de 10 a alrededor de 100 nm tal como, de modo no taxativo, alrededor de 10 a alrededor de 20 nm, alrededor de 10 a alrededor de 30 nm, alrededor de 10 a alrededor de 40 nm, alrededor de 10 a alrededor de 50 nm, alrededor de 10 a alrededor de 60 nm, alrededor de 10 a alrededor de 70 nm, alrededor de 10 a alrededor de 80 nm, alrededor de 10 a alrededor de 90 nm, alrededor de 20 a alrededor de 30 nm, alrededor de 20 a alrededor de 40 nm, alrededor de 20 a alrededor de 50 nm, alrededor de 20 a alrededor de 60 nm, alrededor de 20 a alrededor de 70 nm, alrededor de 20 a alrededor de 80 nm, alrededor de 20 a alrededor de 90 nm, alrededor de 20 a alrededor de 100 nm, alrededor de 30 a alrededor de 40 nm, alrededor de 30 a alrededor de 50 nm, alrededor de 30 a alrededor de 60 nm, alrededor de 30 a alrededor de 70 nm, alrededor de 30 a alrededor de 80 nm, alrededor de 30 a alrededor de 90 nm, alrededor de 30 a alrededor de 100 nm, alrededor de 40 a alrededor de 50 nm, alrededor de 40 a alrededor de 60 nm, alrededor de 40 a alrededor de 70 nm, alrededor de 40 a alrededor de 80 nm, alrededor de 40 a alrededor de 90 nm, alrededor de 40 a alrededor de 100 nm, alrededor de 50 a alrededor de 60 nm, alrededor de 50 a alrededor de 70 nm, alrededor de 50 a alrededor de 80 nm, alrededor de 50 a alrededor de 90 nm, alrededor de 50 a alrededor de 100 nm, alrededor de 60 a alrededor de 70 nm, alrededor de 60 a alrededor de 80 nm, alrededor de 60 a alrededor de 90 nm, alrededor de 60 a alrededor de 100 nm, alrededor de 70 a alrededor de 80 nm, alrededor de 70 a alrededor de 90 nm, alrededor de 70 a alrededor de 100 nm, alrededor de 80 a alrededor de 90 nm, alrededor de 80 a alrededor de 100 nm y/o alrededor de 90 a alrededor de 100 nm.

65 En una realización, las nanopartículas lipídicas pueden tener un diámetro de alrededor de 10 a 500 nm. En una

realización, la nanopartícula lipídica puede tener un diámetro mayor de 100 nm, mayor de 150 nm, mayor de 200 nm, mayor de 250 nm, mayor de 300 nm, mayor de 350 nm, mayor de 400 nm, mayor de 450 nm, mayor de 500 nm, mayor de 550 nm, mayor de 600 nm, mayor de 650 nm, mayor de 700 nm, mayor de 750 nm, mayor de 800 nm, mayor de 850 nm, mayor de 900 nm, mayor de 950 nm o mayor de 1000 nm. En algunas realizaciones, la nanopartícula de lípido catiónico tiene un diámetro promedio de 50-150 nm. En algunas realizaciones, la nanopartícula de lípido catiónico tiene un diámetro promedio de 80-100 nm.

Las cantidades relativas del ingrediente activo, el excipiente farmacéuticamente aceptable y/o cualquier ingrediente adicional en una composición farmacéutica de acuerdo con la presente descripción pueden variar conforme a la identidad, tamaño y/o afección del sujeto que se trata y también conforme a la vía por la que se administrará la composición. Por ejemplo, la composición puede comprender entre 0,1 % y 99 % (p/p) del ingrediente activo. A modo de ejemplo, la composición puede comprender entre 0,1 % y 100 %, por ejemplo, entre 0,5 y 50 %, 1-30 %, 5-80 %, al menos 80 % (p/p) de ingrediente activo.

En una realización, las composiciones que contienen los polinucleótidos descritos en la presente y formulado en una nanopartícula lipídica que comprende MC3, colesterol, DSPC y PEG2000-DMG, el amortiguador citrato trisódico, sacarosa y agua para inyección. Como ejemplo no taxativo, la composición comprende: 2.0 mg/ml de la sustancia de fármaco, 21.8 mg/ml de MC3, 10.1 mg/ml de colesterol, 5,4 mg/ml de DSPC, 2.7 mg/ml de PEG2000-DMG, 5.16 mg/ml de citrato trisódico, 71 mg/ml de sacarosa y alrededor de 1.0 ml de agua para inyección.

La aplicación de la presente invención no se limita a los detalles de construcción y a la disposición de los componentes establecidos en la descripción que se encuentra más adelante o ilustrados en las figuras. La invención permite otras realizaciones, así como también llevarse a la práctica de varias maneras. Además, la redacción y la terminología utilizadas en la presente tienen fines descriptivos y no deberían ser interpretadas en forma taxativa. El uso de "que incluye", "que comprende" o "que tiene", "que contiene", "que implica", y variaciones estos en la presente pretende abarcar los ítems que se enumeran a continuación en la presente y sus equivalentes, así como ítems adicionales.

## 30 EJEMPLOS

### **Ejemplo 1. Elaboración de polinucleótidos**

De acuerdo con la presente invención, es posible realizar la elaboración de polinucleótidos y/o partes o regiones de estos mediante el uso de métodos indicados en WO2014/152027 presentada el 15 de marzo de 2013 titulada "Manufacturing Methods for Production of RNA Transcripts" [Métodos de elaboración para la producción de transcripciones de ARN], (número de expediente M500).

Los métodos de purificación pueden incluir los indicados en la solicitud internacional WO2014/152030 y WO2014/152031.

Es posible poner en práctica los métodos de detección y caracterización de los polinucleótidos tal como se indica en WO2014/144039.

La caracterización de los polinucleótidos de la descripción es posible mediante un procedimiento seleccionado del grupo que consiste en mapeo de polinucleótidos, secuenciación de transcriptasa inversa, análisis de la distribución de carga y detección de impurezas de ARN, en los que la caracterización comprende la determinación de la secuencia de transcripción de ARN, la determinación de la pureza de la transcripción de ARN o la determinación de la carga de heterogeneidad de la transcripción de ARN. Tales métodos se encuentran indicados, por ejemplo, en WO2014/144711 y WO2014/144767.

### **Ejemplo 2. Síntesis de polinucleótido quimérico**

#### **Introducción**

De acuerdo con la presente descripción, es posible unir o ligar dos regiones o partes de un polinucleótido quimérico mediante química de trifosfato.

De acuerdo con este método, se sintetiza químicamente una primera región o parte de 100 nucleótidos o menos con un monofosfato en 5' y desOH en el extremo 3' u OH bloqueado. Si la región es más extensa que 80 nucleótidos, es posible sintetizarla como dos hebras para su ligación.

Si la primera región o parte se sintetiza como una región o parte modificada en forma no posicional mediante transcripción in vitro (IVT), puede seguir la conversión de monofosfato en 5' con posterior formación de casquetes del extremo 3'.

Los grupos protectores de monofosfato se pueden seleccionar de cualquiera de los conocidos en la técnica.

La segunda región o parte del polinucleótido quimérico se puede sintetizar con síntesis química o métodos IVT. Los métodos IVT pueden incluir una ARN polimerasa que puede emplear un cebador como un casquete modificado. De manera alternativa, es posible sintetizar químicamente un casquete de hasta 130 nucleótidos y acoplarlo a la región o parte IVT.

Se observa que, en el caso de los métodos de ligación, la ligación con ADN ligasa T4 seguida del tratamiento con DNasa debería evitar la concatenación fácilmente.

No es necesario que la totalidad del polinucleótido quimérico sea elaborada con una estructura fosfato-azúcar. Si una de las regiones o partes codifica un polipéptido, entonces es preferible que tal región o parte comprenda una estructura fosfato-azúcar.

Luego se realiza la ligación mediante cualquier química clic, química ortoclic, solulink u otra química de bioconjugado conocida por los expertos en la técnica.

#### Vía sintética

El polinucleótido quimérico se elabora mediante una serie de segmentos de partida. Dichos segmentos incluyen:

- (a) el segmento en 5' protegido y con casquete que comprende un OH en 3' normal (SEG. 1)
- (b) el segmento de trifosfato en 5' que puede incluir la región que codifica un polipéptido y que comprende un OH en 3' normal (SEG. 2)
- (c) el segmento de monofosfato en 5' para el extremo 3' del polinucleótido quimérico (por ejemplo, la cola) que comprende cordicepina o carece de OH en 3' (SEG. 3)

Luego de la síntesis (química o IVT), el segmento 3 (SEG. 3) con cordicepina y luego con pirofosfatasa para crear el monofosfato en 5'.

Luego, el segmento 2 (SEG. 2) se liga al SEG. 3 con ARN ligasa. Entonces, el polinucleótido ligado se purifica y se trata con pirofosfatasa para escindir el difosfato. Luego, puede purificarse la construcción SEG.2-SEG. 3 tratada y el SEG. 1 se liga al extremo 5'. Se puede llevar a cabo una etapa de purificación adicional del polinucleótido quimérico.

En los casos en los que el polinucleótido quimérico codifica un polipéptido, es posible representar los segmentos unidos o ligados de la siguiente manera: UTR 5' (SEG. 1), marco de lectura abierto u ORF (SEG. 2) y UTR 3' + PoliA (SEG. 3).

Los rendimientos de cada etapa pueden ser tanto como 90-95 %.

#### **Ejemplo 3: PCR para la producción de la plantilla de ADN**

Los procedimientos de PCR para la preparación de ADNc se llevan a cabo con 2x KAPA HIFI™ HotStart ReadyMix de Kapa Biosystems (Woburn, MA). Este sistema incluye 12.5 µl de 2x KAPA ReadyMix; 0.75 µl de cebador directo (10 uM); 0.75 µl de cebador inverso (10 uM); -100 ng de plantilla de ADNc y dH<sub>2</sub>O diluido hasta 25.0 µl. Las condiciones de reacción se encuentran a 95 °C durante 5 minutos y 25 ciclos de 98 °C durante 20 segundos, luego 58 °C durante 15 segundos, luego 72 °C durante 45 segundos, luego 72 °C durante 5 minutos, luego 4 °C.

La reacción se limpia con el micro kit de PCR PURELINK™ de Invitrogen (Carlsbad, CA) conforme a las instrucciones del fabricante (hasta 5 µg). Las reacciones más grandes requerirán una limpieza con un producto de mayor capacidad. Luego de la limpieza, se cuantifica el ADNc con el NANODROP™ y se analiza mediante electroforesis en gel de agarosa para confirmar que el ADNc es del tamaño esperado. Luego, se presenta el ADNc para el análisis de secuenciación antes de proceder a la reacción de transcripción *in vitro*.

#### **Ejemplo 4. Análisis de IVT e IFN-β de ARN-1 modelo corto**

A las reacciones de IVT se les puede "añadir" nucleótidos específicos, lo que da como resultado contaminantes con respuesta mínima a citocinas y mejores rendimientos. En un proceso, aumentó la relación de GTP con respecto a otros NTP, lo que dio como resultado una mayor carga de NPT total y una mayor relación molar de NTP:Mg<sup>2+</sup>, tal como se muestra en la Tabla 1. Estos nucleótidos específicos se denominan ARN modelo corto. El objetivo de los ARN modelo corto era recapitular la misma señal de citocinas/efecto del método equimolar (técnica previa) en comparación con los métodos de la invención observados en ARN de longitud completa. La



construcción era lo suficientemente pequeña para detectar impurezas de longitud completa mediante LC/MS lo que no fue actualmente viable con ARN de longitud completa.

Tabla 1. Formulaciones de nucleótidos

	Equimolar	Alfa	GDP alfa
[GDP] mM	0	0	30
[GTP] mM	7,5	30	15
[ATP] mM	7,5	15	15
[CTP] mM	7,5	7,5	7,5
[UTP] mM	7,5	7,5	7,5
Total [Nuc] mM	30	60	75
[Mg2+] mM	40	40	40
Nuc: Mg2+	0,75	1,50	1,88
Eficaz [fosfato]	90	180	195
T7 (U/μl reacción)	7	14	14

5 Con el fin de determinar si una transcripción modelo corta preparada usando dos procesos de IVT diferentes imita la respuesta de IFNbeta (fibroblasto BJ) in vitro que se observó para transcripciones de longitud completa, se realizó un estudio independiente del proceso experimental (es decir, IVT en bruto, ultrafiltrada, o purificada con dT). Se generó un ARN sustituto, ARN-1 modelo corto, producido en las mismas condiciones de IVT que  
10 el ARN objetivo y a partir del mismo dúplex de plantilla de ADN. La respuesta a citocinas imitó la del ARN de longitud completa en un ensayo de IFN-β en fibroblastos BJ (Figura 1). La transcripción modelo corta preparada usando el proceso de IVT equimolar tiene una mayor respuesta a IFNbeta que el material preparado usando el proceso de IVT alfa. La respuesta a citocina se conservó después de la diafiltración y la purificación con dT.

15 Además, la realización de perfil del ARN-1 modelo corto (todas las uridinas modificadas en pseudouridina y todas las citidinas modificadas por 5'O-metil), mostró la presencia de complementos inversos (ARNcd) en el proceso equimolar pero no en la reacción Alfa. Se identificaron nueve especies diferentes – los picos superpuestos entre las dos especies fallidas representadas, mientras que los picos observados solamente en la formulación equimolar son especies complementarias inversas. Se realizó la selección en especies con  
20 cantidades variables de modificaciones de uridina y citidina, y se encontraron especies complementarias inversas y fallidas en las tres (Figura 2). Las transcripciones fallidas son sentido y las complementarias inversas son transcripciones antisentido. Los niveles de IFN-β también variaron entre las tres (Figura 2). El superior es G0 = Estándar A, G, C, U, el central es G5 = Estándar A,G,C y 1-metilpseudoU y el inferior es G6 = Estándar A,G,C y 5-metoxiuridina. La transcripción modelo corta preparada mediante 2 procesos diferentes en 3  
25 químicas diferentes tienen perfiles de impurezas similares mediante LCMS, a pesar de las diferencias en el ensayo de BJJ, que parece ser sensible a diferentes químicas.

Un análisis de LCMS de una transcripción modelo y transcripción intacta se preparó a 25 °C durante 6h IVT. Una transcripción corta y una transcripción de longitud completa se prepararon usando los procesos equimolar  
30 y alfa a 25 °C durante 6h (no la IVT normal, que es 37 °C 2h). Las reacciones de IVT a temperatura baja (<30 °C) producen una mayor abundancia de ARN de complemento inverso que 37 °C. Los perfiles de impureza de las muestras se analizaron mediante LCMS. Las muestras preparadas usando ambos procesos contenían productos fallidos/truncados. Las transcripciones preparadas usando el proceso equimolar contenían especies de poliU de diversas longitudes. El proceso alfa todavía mitiga la formación de complementos inversos incluso  
35 a 25 °C donde la abundancia fue mayor que las condiciones estándares usando procesos equimolares (Figura 3).

#### *Detección de ARN complementario/antisentido*

40 Las impurezas inmunoestimuladoras parecen ser impulsadas por la transcripción de ARN con plantilla de ARN, dado que T7 retira mediante polimerización el ARN transcrito naciente. El ARN complementario/antisentido resultante (ARNcd) que se generó muestra bases mezcladas, un componente de poliU, y un trifosfato en 5' (5'ppp), se que inicia con cualquier base. Se realizó la transcripción con plantilla de ARN. La hEPO G5 purificada de fase inversa (fría) se incubó a 4 mg/ml (consistente con un rendimiento típico de una reacción de  
45 transcripción in vitro), con todos los componentes de IVT, excepto la plantilla de ADN, que se determinó que estaba por debajo del umbral requerido para la transcripción in vitro mediante qPCR, para explorar el fenómeno de la transcripción con plantilla de ARN. Los materiales se analizaron mediante UPLC y ensayo de IFNbeta de BJJ in vitro. El análisis de UPLC demostró que los productos de ARN aberrantes, y más notoriamente las especies de poliU, se producen a través de transcripción con plantilla de ARN (en ausencia de la plantilla de  
50 ADN). La reacción realizada con ARN y todos los componentes de IVT fue caliente con IFNbeta, mientras que la reacción realizada sin ARN fue fría. (Figura 4).

El ARN de cadena doble se puede depurar del ARN transcrito purificado por fase inversa (RP) . Se realizó un análisis de electroforesis capilar del material de hEPO G5 tratado con RNasa III. La hEPO G5 se preparó usando condiciones ya sea equimolares o alfa y una parte se purificó mediante RP. Luego las muestras se trataron con RNasa III y se analizaron mediante electroforesis capilar. El material preparado mediante el proceso alfa contenía menos sustrato de RNasa III que el preparado por el proceso equimolar. La purificación de RP depura la mayoría del sustrato de RNasa III del material equi y alfa. El proceso alfa y la purificación de fase inversa aparentemente proporcionan una reducción sinérgica en el sustrato de RNasa III (por ejemplo, ARNcd) (Figura 5A). Se realizó un análisis de electroforesis capilar del material de hEPO G0 tratado con RNasa III. La hEPO G0 se preparó usando el proceso equimolar y una parte se purificó mediante RP. Luego las muestras se trataron con RNasa III y se analizaron mediante electroforesis capilar (azul: tratado; negro: no tratado). El material que se purificó mediante RP contenía menos sustrato de RNasa III que el material que no se purificó mediante RP (estado de la técnica actual) (Figura 5B).

La RNasa III es una nucleasa específica de ARNcd. Las preparaciones de ARN se someten a tratamiento con RNasa III durante un tiempo fijo. Los perfiles de pureza/impureza de ARN se comparan antes y después del tratamiento con RNasa III y se miden mediante HPLC o electroforesis capilar. La cantidad de producto de longitud completa degradado fue proporcional al nivel de impurezas de ARN de cadena doble presente en la preparación de ARN. Se consideró que la cantidad degradada tal como lo indicó un cambio en el tamaño aparente/tiempo de retención/ es la cantidad del sustrato de RNasa III. Las muestras sin ARNcd muestran una pureza casi idéntica antes y después del tratamiento con RNasa III. Las muestras que contenían cantidades significativas de ARNcd mostrarán la formación de producto de escisión sustancial y agotamiento del ARN de longitud completa presente en la carga de alimentación no tratada tal como se observa mediante HPLC o electroforesis capilar. (Ejemplo: si 80% del ARNm original permanece después del tratamiento con RNasa III, 20% era un sustrato para RNasa III.)

Se realizó un análisis de expresión de IFNbeta y hEPO de las muestras de las Figuras 5A y 5B. Las muestras tratadas y no tratadas se analizaron para determinar cómo el tratamiento con RNasa III afecta la respuesta de IFNbeta y la expresión de hEPO en fibroblastos BJ (in vitro). Se descubrió que el material equimolar hEPO A100 G5 se expresa de manera similar para las muestras tratadas con RIII y no tratadas. Los niveles de citocinas para la muestra de -RP tratada equi disminuyeron después del tratamiento con RIII. El material alfa hEPO A100 G5 se expresó en su totalidad de manera similar y tuvo respuesta de IFNbeta baja/nula. El material TL no se expresa. Sin embargo, el tratamiento con RIII redujo el nivel de CK para muestras + y -RP. El material G0 hEPO A100 presentó los mayores efectos del tratamiento con RNasa III. Después del tratamiento con RNasa III, las muestras purificadas + y -RP presentaron un aumento de la expresión y una disminución del nivel de IFNbeta (Figura 6).

Se realizó un análisis de electroforesis capilar de una transcripción corta usando diferentes procesos y tratada con RNasa III. La pregunta era si una transcripción modelo corta se podría utilizar para caracterizar el efecto del tratamiento con RNasa III. Una construcción de ARN sustituto se transcribió usando los procesos equimolar o alfa luego se trató con RNasa III y se analizó mediante electroforesis capilar. Aparentemente el material equimolar contiene la mayoría del sustrato de RNasa III, mientras que el material del proceso alfa no contiene ningún sustrato, según CE. Con el tratamiento con RNasa III, el pico del ARN modelo cambia 5-7 nucleótidos en el analizador de fragmentos. La IVT equimolar forma un segundo pico drástico 216 nucleótidos más corto que el pico principal, que también se observó con el ARNm que contiene ORF (Figura 7). Por lo tanto, usando ARN modelo y LC/MS, se puede caracterizar de manera precisa qué componente se cortó, y mediante deducción, qué componente quedó. El método de pureza de HPLC mostró poliU tras el aislamiento y el enriquecimiento de los productos de escisión (Figura 9). Se realizó un análisis de RP de transcripción de ARN sustituto 2 EQ tratada con RNasa III (misma construcción que la figura superior en la Figura 7...análisis alternativo). Se observaron varias especies mediante análisis de RP después del tratamiento de la transcripción de ARN sustituto con RNasa III. Dado que este método de RP era un método selectivo de colas, se presenta la hipótesis de que estos picos con elución temprana son oligos cortos y/o variantes de cola (Figura 8A). Se realizó un análisis de RP de una construcción de ARN sustituto 2 alfa transcrita usando la transcripción del proceso alfa tratada con RNasa III. No se observó ninguna evolución de picos de impureza adicionales ni cambios considerables a la pureza general en la traza de RP para el material alfa de ARN sustituto tratado con RIII. Por lo tanto, se puede concluir que el proceso alfa no produce las mismas especies de ARNcd que el proceso equimolar (Figura 8B).

#### *Abundancia de ARNcd mediante RNasa III y datos de citocina a partir de fracciones de RP*

Se realizó un fraccionamiento de RP de hEPO tratado con y sin RNasa III. El ARNm de hEPO G5 transcrito mediante ambos procesos, equimolar y alfa, se purificó mediante HPLC de fase inversa. Las fracciones se recolectaron en el gradiente de elución. Las fracciones se trataron con RNasa III, se analizaron posteriormente mediante electroforesis capilar comparando el ARN no tratado y tratado con RNasa III (superpuesto). Las fracciones de ARN también se transfectaron en fibroblastos BJ y la inducción de IFN-B se evaluó antes y después del tratamiento con RNasa III (Figura 9).

Se realizó un análisis de electroforesis capilar de hEPO fraccionada mediante RP EQ +/- tratamiento con RIII. Se trató hEPO EQ con RNasa III luego se purificó mediante RP, se fraccionó, y se analizó mediante CE (azul: tratada con RIII; negro: no tratada) Las fracciones equimolares tempranas (fracciones 1-6) que contienen impurezas de ARN indican abundancia considerable de ARNcd/sustrato de RNasa III. El ARNcd estaba enriquecido en las fracciones tempranas. Esto fue confirmado por niveles de IFN-B altos. Las fracciones equimolares 7-9 indican niveles menores de ARNcd/sustrato de RNasa III mediante CE, lo que también se confirmó mediante niveles en disminución de IFN-B. El tratamiento con RNasa III de cada fracción reduce la inducción de IFN-B a niveles basales, lo que nuevamente confirma que las impurezas de IFN-B están compuestas de ARNcd (Figura 10A-C). Análisis de IFNbeta in vitro de hEPO EQ G5 no tratado o después del tratamiento con RNasa III (Figura 10D).

Se realizó un análisis de electroforesis capilar de hEPO fraccionada mediante RP alfa +/- tratamiento con RIII. Se trató hEPO alfa con RNasa III luego se purificó mediante RP, se fraccionó, y se analizó mediante CE (azul: tratado con RIII; negro: no tratado) (Figura 11A-C). Análisis de IFNbeta in vitro de hEPO Alfa G5 no tratado o después del tratamiento con RNasa III (Figura 11D). Todas las fracciones indican niveles mínimos de ARNcd (sustrato de RNasa III) mediante electroforesis capilar dado que los electroferogramas superpuestos son prácticamente idénticos. Esto se confirmó mediante niveles basales de IFN-B tanto en fracciones tratadas como en fracciones no tratadas. El ARN transcrito con el proceso alfa no presentaba ARNcd, a pesar de que había impurezas de ARN de longitud no completa/ truncadas en las fracciones tempranas.

#### *Detección mediante ELISA de abundancia de ARNcd*

El ELISA de ARNcd J2/K2 se desarrolló para medir la abundancia de ARNcd. Las placas se recubrieron con anticuerpos monoclonales J2 (IgG) y luego se bloquearon. Luego se agregó el ARN de interés a concentraciones dadas y se incubó durante una hora. Se agregó el anticuerpo monoclonal K2 (IgM), y se agregó un anticuerpo policlonal de cabra anti-IgM conjugado con HRP. Se agregó TMB para desarrollar la señal. El ensayo detecta dúplex con una longitud mayor que 40 pares de bases. J2 puede ayudar a determinar el criterio de valoración de RIII. Se observó una inactivación de ARNcd con RP o tratamiento con RNasa III (Figura 12). El ensayo de J2 sugiere que hay considerablemente más ARNcd en el material equimolar que en el alfa. El proceso de RP retira el material de ARNcd de las muestras EQ, de manera similar al tratamiento con RIII. El ARNm de reacción alfa tuvo menos ARNcd en la carga de alimentación en comparación con el producto de IVT equimolar. También, se observó que una inactivación en materiales de RNasa III purificados con dT fue mayor que TrisRP. El ensayo también ilustró que se eliminó ARNcd mediante el tratamiento con RNasa III (Figura 13). Mientras que los niveles de ARNcd detectados por J2 varían en función de la construcción/proceso/química, el tratamiento con RNasa III aparentemente reduce la mayoría de los niveles de ARNcd al nivel de referencia. El análisis de LCMS después del tratamiento con nucleasa P1 mostró NTP adicionales presentes en ARNm de FFB, dado que los complementos inversos no iniciados con G estaban presentes en mayor abundancia en el grupo equimolar, en comparación con el grupo de reacción alfa (Tabla 2).

Tabla 2.

	pppA ng/ml	pppG ng/ml	pppC ng/ml	pppU ng/ml
<b>G5 hEPO Equimolar sin casquete</b>	<b>31</b>	<b>172</b>	<b>53</b>	<b>BQL</b>
<b>G5 hEPO Formulación de la invención Sin casquete</b>	<b>12</b>	<b>159</b>	<b>BQL</b>	<b>BQL</b>

#### *Mitigación de la transcripción con plantilla de ARN en IVT a bajas temperaturas*

Se realizó una respuesta a IFNbeta para una construcción de hEPO preparada usando diferentes procesos y químicas. Tal como se muestra en el estudio de caracterización de IVT, la reacción alfa fue menos sensible a picos de citocinas inducidos por baja temperatura (Figura 14). La IVT equimolar estándar cuando se realiza a 25 °C genera IFN-B mejorado que induce la abundancia de impurezas. Se realizó un análisis de nucleótidos totales en construcciones de hEPO preparadas usando diferentes procesos. Las condiciones del proceso alfa (GDP y GTP) mitigan IFN-B extraños que inducen la formación de impurezas cuando se realiza la IVT a 37 °C y 25 °C. La distribución de nucleótidos fue consistente en las regiones de control en el estudio de nucleasa P1 (cuadros grises) y el mismo plásmido de hEPO muestra escisión consistente (Figura 15). Con los procesos de IVT equimolares ejecutados a 25 °C, la abundancia de Uridina/ 1-metil pseudo U se encontraba en mayor abundancia que a 37 °C supuestamente debido a la evolución de Poli U, los procesos Alfa (GDP y GTP) muestran distribución de nucleótidos consistente incluso a 25 °C lo que respalda adicionalmente la mitigación de formación de impurezas. Las diferentes alturas de barra posiblemente se deban a diferentes concentraciones. El contenido de U aumentó con el tiempo con las IVT estándares a 25 °C. La composición de nucleótidos corregida de manera molar se da en las Tablas 3 y 4. G y U son valores teóricos, mientras que A y está sobrerrepresentado y C está subrepresentado. Hubo menos desviación en las condiciones de temperatura con la reacción alfa, y la mayor desviación se observó en la U para IVT estándares.

Tabla 3.

	% de A	% de G	% de C	% de U
GDP 37 °C 2hr	35,5	27,9	18,3	18,3
GDP 25 °C 6hr	35,2	28,1	18,4	18,2
GDP 25C durante la noche	34,8	27,7	19,3	18,2
GTP 37 °C 2hr	35,2	27,5	18,9	18,4
GTP 25 °C 6hr	35,2	27,5	19,1	18,3
GTP 25 °C durante la noche	35,1	27,7	19,0	18,2
G5 37 °C 2hr	35,4	26,8	19,5	18,3
G5 25 °C 6hr	33,8	27,7	18,3	20,1
G5 25 °C durante la noche	32,1	26,5	17,1	24,3
G0 37 °C 2hr	36,5	26,7	19,1	17,6
G0 25 °C 6hr	33,0	24,4	17,1	25,6
<b>Teórico</b>	<b>31,7 %</b>	<b>27,3 %</b>	<b>22,4 %</b>	<b>18,6 %</b>

Tabla 4. Desviación estándar

	A	G	C	U
GDP	0,34	0,23	0,53	0,04
GTP	0,07	0,15	0,06	0,09
G5 estándar	1,66	0,60	1,20	3,08
G0 estándar	2,53	1,65	1,43	5,61

5

*Transcripción con plantilla de ARN “forzada”*

Se realizó un análisis de impureza mediante LCMS de IVT basada en ARN en diferentes químicas usando G5 eGFP EQ. Se establecieron las IVT (en diferentes químicas) usando 4 mg/ml de ARN eGFP G5 equi y no se generan especies de poliU con plantilla de ADN a partir de la transcripción con plantilla de ARN. La transcripción con plantilla de ARN se produce independientemente de la química. El eGFP G5 purificado de fase inversa (fría) se incubó a 4 mg/ml (consistente con un rendimiento típico de una reacción de transcripción in vitro), con todos los componentes de IVT, excepto la plantilla de ADN, que se determinó que estaba por debajo del umbral requerido para la transcripción in vitro mediante qPCR, para explorar el fenómeno de la transcripción con plantilla de ARN. Los materiales se analizaron mediante LC/MS (Figura 16A (proceso Equimolar), 16B (proceso alfa)).

Se realizó un análisis de IFNbeta para productos de IVT con plantilla de ARN. El material fue el de la Figura 16. Después de la IVT basada en ARN, las muestras se analizaron in vitro para determinar si había una correlación entre el análisis de LCMS (especies de poliU) y la respuesta a IFNbeta. Las IVT basadas en ARN Alfa (A100) producen material que tiene menor respuesta a CK que el material equimolar. (G0 o G5); Lo que sugiere nuevamente que el proceso alfa mitiga la formación de producto de transcripción con plantilla de ARN. Todo el G6 estaba frío; (intrínsecamente frío a pesar de la formación de impurezas) (Figura 17).

Una vacuna no puede colocar casquetes en ARNcd, se realizó para determinar si el ARNcd puede presentar casquetes. Se hibridaron oligos de complemento inverso y directo con trifosfatos en 5', luego se sometieron al proceso de formación de casquetes usando vacuna para determinar si el ARNcd podía presentar casquetes. pppF oligo (contiene 5'pppG) puede presentar casquetes en casquete1 usando vacuna. pppRC oligo (contiene 5'pppU) no puede presentar casquetes usando vacuna. ARNcd con pppF+pppRC no puede presentar casquetes usando vacuna (Figura 18).

Se realizó una desfosforilación de ARNcd usando CIP para determinar si la fosfatasa intestinal de ternero desfosforila el ARNcd con diferentes excedentes en 5'/3'. Se hibridaron varios oligos F y RC luego se trataron con CIP y se analizaron mediante LCMS para determinar la eficacia de la desfosforilación. El ARNcd con excedentes en 5' puede desfosforilarse completamente. El ARNcd de dúplex perfecto puede desfosforilarse parcialmente. El ARNcd con excedentes en 3' no puede desfosforilarse completamente. Esto demuestra por qué la fosfatasa no fue 100% eficaz para reducir la respuesta a CK (Figura 19).

Se realizó una pureza CE de ARNm dosificados en el estudio in vivo mencionado anteriormente indicado en electroferogramas. (Figura 20). La estructura puede cumplir un rol, dado que la UTR 5' puede alterar la eficacia de la reacción alfa (Figura 20).

*Estudios in vivo*

El tratamiento con RIII mostro inactivación de citocinas distinta en ratones, tanto con especies no modificadas como con especies completamente modificadas (toda la uridina era 1-metilpseudouridina). Se puede utilizar como un método independiente para inactivar citocinas hasta niveles casi basales. A ratones hembra (n=5 por

grupo) se les inyectó una vez 0.5 mg/kg de materiales de prueba seleccionados.

Se realizó un análisis de IFNbeta de ARNm de hEPO G0 y G5 transcrito usando diferentes procesos de IVT y permutaciones de purificación para determinar cuál es la respuesta de IFNbeta para el material generado con diferentes químicas, procesos, y tratamientos (in vitro; fibroblastos BJ). Los procesos +RIII, +RP, y alfa tienen efectos similares (reducción) en la respuesta de IFNbeta in vitro en G5. G0 demuestra que alfa +RIII fue superior a equi +RIII (Figura 21).

Se realizó una expresión in vivo (ratones Balb-C) de hEPO preparada usando químicas G0 o G5, procesos equimolar o alfa, y se realizó el tratamiento con RNasa III a través de ELISA para determinar cómo afectó la expresión mediante hEPO preparada usando diversos procesos (química, proceso de IVT, tratamiento con RIII). Mismo material que las Figuras 12, 20, 21, 22. No hay diferencias significativas en la expresión de hEPO usando ARN generado mediante procesos diferentes (Figura 22).

Un panel de citocinas (luminex) para ARNm de hEPO G0 y G5 in vivo (ratones Balb-C) transcrito usando diferentes procesos de IVT y permutaciones de purificación con y sin tratamiento de RNasa III se realizó para determinar con el mismo material que las Figuras 12, 20, 21, 22. Análisis de citocinas después del tratamiento de ratones Balb-C con 0.5 mpk de ARN formulado. El material alfa no tratado tiene una respuesta de CK similar que el material EQ (G5) en IP10. El proceso alfa junto con el tratamiento con RNasa III en general parece conferir menos actividad inmunoestimuladora que el comparador equimolar (+RNasa III), especialmente en G0. El tratamiento con RIII reduce la respuesta a CK (Figura 23A-23D).

Se realizó una activación de células B de bazo de ratones Balb-C tratados con hEPO preparada usando químicas G0 o G5, procesos equimolar o alfa, y tratamiento con RNasa III, las células B activadas (CD86+ CD69+) se analizaron para determinar su respuesta después del tratamiento con hEPO preparada usando varios procesos. La activación de células B se correlaciona aproximadamente con el panel de citocinas (luminex). El tratamiento con RIII reduce la respuesta a CK, con respecto a los controles no tratados (Figura 24).

Se realizó un análisis in vitro de oligos cortos trifosforilados en 5' para determinar si los oligos cortos trifosforilados en 5' estimulan una respuesta de IFNbeta in vitro (BJF, IFNbeta). El ARNcs y ARNcd < 12 nucleótidos o pares de base no estimulan IFNbeta en BJF (Figura 25).

Se realizó un análisis in vitro de poliU y estándares de ARNcd para determinar si usando estándares de oligos trifosforilados en 5', se estimula una respuesta de IFNbeta en BJF. El ARNcd de pppF+pppRC >20 pb fue inmunoestimulador. La poliU trifosforilada en 5' no fue inmunoestimuladora como ARNcs o ARNcd < 25 nt o pb (Figura 26).

Se realizó un análisis in vitro de estándares de ARNcd con excedente en 3' para determinar si hay o no una dependencia de la longitud del excedente en 3' en la respuesta de IFNbeta para estándares de ARNcd. Cuanto más largo el excedente en 3', menor era la respuesta de IFNbeta (Figura 27).

Se realizó un análisis in vitro de estándares de ARNcd con excedente en 5', dúplex perfecto, y excedente en 3' de longitudes variables para determinar si hay o no una dependencia de longitud/identidad de excedente en la respuesta de IFNbeta. Los oligos pppRC de ARNcs están calientes. El ARNcd con excedente en 5' (y dúplex de ~ 20 pb) es menos inmunoestimulador que el dúplex perfecto. El ARNcd con excedente en 3' tiene inmunoestimulación equivalente o menos inmunoestimulación que el dúplex perfecto. Cuanto más largo el dúplex de ARNcd, mayor era la respuesta a CK (Figura 28).

Se realizó un análisis in vitro de especies de poliU para determinar qué es necesario para que las especies de poliU simulen una respuesta de IFNbeta. Los estándares de poliU de ARNcs con trifosfato en 5' son ligeramente inmunoestimuladores. Como ARNcd con OH-F30 (dúplex de 30 pb), hubo una respuesta a CK aditiva. Sin embargo, los estándares de poliU no son inmunoestimuladores en presencia de ARNm de longitud completa (A100) (Figura 29).

Se realizó un análisis in vitro de estándares de oligonucleótidos de ARNcs para determinar cuál era la respuesta de IFNbeta para los estándares de ARNcs. (F u oligos directos representan especies fallidas o truncadas; R o RC u oligos de complemento inverso representan especies de complemento inverso generadas mediante transcripción con plantilla de ARN). Los oligos directos con trifosfato en 5' no estimulan IFNbeta. Los oligos de complemento inverso con trifosfato en 5' estimulan IFNbeta a >25 nt de longitud (Figura 30).

Se realizó un análisis in vitro de estándares de oligos de ARNcd con diferentes funcionalidades en 5' para determinar si la funcionalidad en 5' (trifosfato vs hidroxilo) afecta o no la respuesta de IFNbeta. El trifosfato en 5' en oligo F e hidroxilo en 5' en oligo RC simula IFNbeta a > 20 pb. El trifosfato en 5' en oligo F y trifosfato en 5' en oligo RC simula IFNbeta a > 25 pb. El hidroxilo en 5' en oligo F y trifosfato en 5' en oligo RC simula IFNbeta a > 20 pb. Solo fue necesario un trifosfato para simular IFNbeta a >20 pb (Figura 31).

Se realizó un análisis de IFNbeta de estándares de oligo de ARNcd tratado con CIP. Los estándares de oligo de ARNcd con longitudes de excedentes variables se trataron con CIP y se analizó la respuesta de IFNbeta. Las muestras en las que no se observó desfosforilación (es decir, excedente en 5') mediante LCMS (figura 19) no estimularon IFNbeta como la muestra no tratada. Las muestras que no estaban desfosforiladas mediante CIP (es decir, dúplex perfecto y excedente en 3') no presentaron cambios en la respuesta de IFNbeta +/-CIP (Figura 32).

Se realizó una respuesta a la dosis de impurezas de ARNcs (IFNbeta en fibroblastos BJ) para determinar cuál era la dependencia de la dosis de los estándares de impureza de ARNcs in vitro (BJF, IFNbeta). Qué cantidad de un tipo de impureza es necesaria para estimular una respuesta. El ARNcs <20 mer (RC) no estaba caliente. El ARNcs >30 mer (RC) estimuló IFNbeta a alrededor de >1 ng/ul (o > 2.5 ng/transfección). Hubo una dependencia de longitud aparente con respecto a la respuesta a CK del ARNcs (RC). Cuanto más largo el oligo, mayor era la respuesta a CK (Figura 33).

Se realizó una respuesta a la dosis de impurezas de ARNcd (IFNbeta en fibroblastos BJ) para determinar la dependencia de la dosis de los estándares de impureza de ARNcd in vitro (BJF, IFNbeta). El ARNcd de 20-30 pb estimula IFNbeta a ~>1ng/ul (>2.5 ng/transfección). El ARNcd >30 pb estimula IFNbeta a ~0.1 ng/ul (>0.25 ng/transfección). Fue necesaria la dilución >1000x de ARNcd >30 pb para silenciar la respuesta de IFNbeta... lo que indica que solo unas pocas moléculas de estas impurezas fueron suficientes para estimular una respuesta. Hubo una dependencia de longitud aparente con respecto a la respuesta a CK del ARNcd... cuanto más largo el dúplex, mayor era la respuesta a CK (Figura 34).

Se realizó una respuesta de IFNbeta nucleótido modificado en 5' en estándares de oligo directos para determinar si el nucleótido en 5' (trisfosforilado) afecta o no la respuesta a CK. Cuando se agregó una A/C/U en 5' a los oligos F, se indujo una respuesta a CK. Un elemento distinto a G en 5' no cambia la respuesta a IFNbeta (todavía frío). Esto sugiere que había una dependencia de secuencia (dado que los RC están calientes) (Figura 35).

Se realizó una respuesta de IFNbeta nucleótido modificado en 5' en estándares de oligo de complemento inverso para determinar si el nucleótido en 5' (trisfosforilado) afecta o no la respuesta a CK. Cuando se agregó una G en 5' a los oligos RC, se indujo una respuesta a CK. 5' pppG en oligos RC SILENCIA la respuesta de IFNbeta. Esto sugiere que no importa cuál es la identidad de secuencia, una G en 5' silenciará la respuesta a CK (Figura 36).

Se realizó una respuesta de IFNbeta para ARNcd funcionalizado con hidroxilo en 5' para determinar cuánto afecta la funcionalidad en 5' (ppp vs OH) a la respuesta de citocinas en ARNcd. Se analizaron el dúplex 5' ppp-F/5'ppp-RC con respecto a 5'OH-F/5'OH-RC. Hubo una respuesta de IFNbeta para el ARNcd >30 pb con 5' ppp o 5' OH. Esto sugiere que el ARNcd >30 pb, sin importar cuál era la funcionalidad en 5', estimuló una respuesta de IFNbeta en BJF (Figura 37).

#### *Implicaciones*

La transcripción con plantilla de ARN como un subproducto de IVT se redujo en su mayoría y casi se eliminó con el proceso de reacción alfa. Como resultado de mitigar la formación de complementos inversos, se abordaron ambas impurezas de interés: ARNcd (con 5'ppp) y ARN que inician con elementos diferentes a pppG (pppA, pppC, y pppU).

La transcripción con plantilla de ARN se vio mejorada a temperaturas de reacción de IVT menores que 37 °C (por ejemplo, a 25 °C). El tiempo de rampa para lograr un calentamiento de 37 °C mientras se producía la reacción IVT, puede dar lugar a mayores niveles de impureza, especialmente a medida que aumento el tiempo de rampa de 25 °C a 37 °C. El proceso de reacción alfa fue más tolerante a 25 °C, ya que se detectaron especies de transcripción con plantilla de ARN en trazas a temperatura ambiente.

La RP fue más eficaz con la reacción alfa que con el proceso equimolar. Hubo una carga de impurezas menor con la reacción alfa, lo que dio lugar a separaciones mejoradas. Se puede resolver la "pureza" de manera más explícita y potencialmente puede aumentar la exposición a la carga, lo que aumenta la productividad. Un solo ciclo de RP fue adecuado para llevar una especie no modificada al punto de referencia utilizando la reacción alfa de GDP, mientras que se requieren de dos a tres ciclos de secuencia de RP para obtener dos fracciones de una especie no modificada utilizando IVT equimolar.

#### **Ejemplo 5. El ARN generado con el proceso alfa tiene <40 % de transcripciones mal formadas**

Se generó ARN de hEPO usando procesos equimolar o alfa. El ARN se digirió con Rnasa T1 y el fragmento de cola se analizó mediante LCMS. Un fragmento de cola con un mP en 3' indica una transcripción mal formada. El proceso alfa generó un 3'OH (limpio) que se calculó era de 32780 Da y un 5'OH/3'mP (mal formado) que se

calculó era de 32860 Da. El proceso equimolar generó un 5'OH/3'mP (mal formado) de 32861 Da y una cantidad mucho menor de 3'OH (limpio) (Figura 38).

**Ejemplo 6. La digestión total de ARN en el proceso equimolar tiene una mayor abundancia de nucleótidos que no son GTP 5' que el ARN producido con procesos alfa**

Se generó ARN de marco de lectura abierto corto usando cuatro condiciones diferentes: 37 °C 2 horas (estándar) en comparación con 25 °C 6 horas y equimolar en comparación con alfa. Cada ARN se digirió de manera enzimática en nucleótidos simples, luego la abundancia de nucleótidos en 5' se analizó mediante LCMS (por ejemplo, pppG o GTP, pppA o ATP, etc.). Una G en 5' es el primer nucleótido con plantilla, lo que significa que si se genera una población de ARN impuro, entonces habrá una gran fracción de nucleótidos en 5' como ATP, CTP, o UTP (por ejemplo, procesos equimolares), y si se genera una población de ARN puro, entonces la mayoría de nucleótidos en 5' serían GTP (por ejemplo, procesos alfa).

**Ejemplo 7. El ARN de marco de lectura abierto corto generado mediante procesos alfa genera menos complementos inversos**

Se generó ARN de marco de lectura abierto corto en químicas G0 (de tipo salvaje) y G5 (m<sup>1</sup>Ψ) usando procesos equimolar o alfa que contenía 32P-GTP o 32P-CTP. 32P-GTP etiqueta las transcripciones fallidas y 32P-CTP etiqueta las transcripciones de complemento inverso. No hubo diferencia en perfiles fallidos o de RC entre las químicas G0 y G5. Los procesos equimolar y alfa tienen perfiles fallidos similares. El proceso equimolar genera más complementos inversos que el proceso alfa.

**Ejemplo 8. La digestión de RNasa T1 informa poblaciones de transcripciones mal formadas**

La RNasa T1 es una endonucleasa que escinde específicamente ARN después de nucleótidos de guanosina, lo que deja un hidróxido en 5' y un monofosfato en 3'. Para construcciones que contienen un nucleótido de guanosina con plantilla en el extremo 3', se puede utilizar RNasa T1 para distinguir poblaciones de transcripciones mal formadas, lo que deja un monofosfato en 3', en comparación con transcripciones limpias, que contienen un hidróxido en 3'. Se generó ARN de hEPO usando procesos equimolar o alfa, luego se digirió con RNasa T1 y se analizó mediante espectrometría de masas. El fragmento de oligo en 3' se analizó y se cuantificó por su heterogeneidad en 3'. El ARN generado con el proceso equimolar contiene aproximadamente 70-80 % de transcripciones mal formadas, mientras que el ARN generado con el proceso alfa contiene 30-40 % de transcripciones mal formadas.

**Ejemplo 9. Digestión total de ARN para determinar la pureza de la muestra**

Se generó una construcción de ARN modelo corto (sustituto ARN 1) usando procesos equimolar o alfa con incubación a 37 °C o 25 °C. Cada muestra se purificó mediante resina oligo dT para eliminar los NTP que no reaccionaron. Cada muestra se digirió de manera enzimática en nucleótidos simples usando S1 y nucleasas de benzonasa. La abundancias de nucleótidos en 5', que están trifosforiladas en el extremo 5', se analizaron mediante espectrometría de masas. Se integraron los picos de iones extraídos y los % de abundancias de NTP se indican en la Tabla 5. Dado que GTP en 5' es el primer nucleótido con plantilla, un valor alto de % de GTP indica una población de ARN pura. Asimismo, una abundancia relativamente baja de GTP indica una población de ARN impuro. Las muestras generadas usando el proceso EQ tienen >5 % de nucleótidos en 5' que no son GTP, mientras que las muestras generadas usando un proceso alfa tienen <5 % de nucleótidos en 5' que no son GTP.

Tabla 5.

	% de GTP	% de ATP	% de CTP	% de UTP
ARN 1 sustituto equimolar 37 °C	94,4	0,8	3,8	1,2
ARN 1 sustituto alfa 37 °C	97,4	0,3	2,3	0,0
ARN 1 sustituto equimolar 25°C	68,1	5,0	12,3	21,6
ARN 1 sustituto alfa 25°C	97,6	0,2	1,9	0,4

**Ejemplo 10. El ELISA de ARNcd indica la presencia de ARNcd**

Se generó una construcción de ARN hEPO usando procesos equimolar o alfa y se purificó ya sea solo mediante resina oligo dT (-RP) o con purificación de fase inversa (+RP). Se utiliza ELISA usando anticuerpos específicos de ARNcd para determinar las diferencias relativas en las purezas de ARN generados usando procesos diferentes. Los ARN generados usando el proceso equimolar contienen significativamente más ARNcd que los ARN generados con el proceso alfa (Figura 39). La purificación de fase inversa mejora las purezas de los ARN -RP.

**Ejemplo 11. El análisis en gel de secuenciación radiactiva determina la pureza de la muestra**

Se generó una construcción de ARN modelo corto (sustituto ARN 1) usando procesos equimolar o alfa. Cada reacción de IVT contenía  $^{32}\text{P}$ -GTP, que etiqueta transcripciones fallidas, o  $^{32}\text{P}$ -CTP, que etiqueta transcripciones de complemento inverso. Las muestras de ARN se analizaron mediante gel de secuenciación. En función de los datos de  $^{32}\text{P}$ -GTP, no hay diferencia en la abundancia de transcripciones fallidas en los dos procesos. En función de los datos de  $^{32}\text{P}$ -CTP, el proceso equimolar genera más complementos inversos que el proceso alfa.

**Ejemplo 12. Análisis in vivo de ARNcd dopado en ARNm**

El ARNm de hEPO purificado de fase inversa se dopó con 5 %, 0,5 % o 0,05 % p/p de ARNcd de 60 pb, que corresponde a los primeros 60 nucleótidos de la construcción de hEPO intacta. Las muestras de ARNm se formularon en MC3, luego se dosificaron IV en ratones C57BL/6 y Balb-c a 0,5 mpk. Después de 24 horas, se cosecharon los bazo y se homogeneizaron para generar suspensiones de células simples. Las células del bazo se tiñeron con marcadores de células B, luego se analizaron mediante citometría de flujo. Las poblaciones de células B activadas se analizaron en función de su expresión de marcadores CD86 y CD69. El ARNcd de 60pb tiene la siguiente secuencia:

60mer 5' UTR Epo F                    GGGAAAUAGAGAGAAAAGAGUAAGAAGAAUUAAGAGCCACCAUGGGAGUGCACG

60mer 5' UTR Epo R                    CGUGCACUCCCAUGGUGGCUCUUAUUAUUCUUCUACUCUUCUUUCUCUCUUAUUUCCC

Los grupos que recibieron ARNm de hEPO generado mediante un proceso alfa tuvieron menor activación de células B que los grupos que recibieron hEPO con ARNcd dopado. También se recolectó suero a las 6 h y se analizó mediante un panel luminex de citocinas. Las tendencias de expresión para los marcadores IP-10, IFN- $\gamma$ , e IFN- $\alpha$  del panel luminex fueron consistentes con las tendencias del análisis de activación de células B, lo que indica que el ARNcd en la muestra dopada desencadenó la vía de interferón tipo I, lo que dio lugar a la activación de células B.

**LISTADO DE SECUENCIAS**

<110> Moderna Therapeutics, Inc.

<120> COMPOSICIONES DE ARN DE ALTA PUREZA Y MÉTODOS PARA SU PREPARACIÓN

<130> M1378.70047WO00

<140> Sin asignar aún

<141> Simultáneamente con el presente documento

<150> US 62/394,711

<151> 14-09-2016

<160> 32

<170> PatentIn versión 3.5

<210> 1

<211> 12

<212> ARN

<213> Secuencia artificial

<220>

<223> Polinucleótido sintético

<400> 1

cucuuauuuc cc 12

<210> 2

<211> 49

<212> ARN

<213> Secuencia artificial



<220>  
 <223> Polinucleótido sintético

5      <220>  
          <221> misc\_feature  
          <222> (1)..(1)  
          <223> Modificado con OH

10     <220>  
          <221> misc\_feature  
          <222> (49)..(49)  
          <223> Modificado con ppp

15     <400> 2  
          ggaaauaaga gagaaaagac ccuuuuuucu cucuuuucu cucauuucu 49

20     <210> 3  
          <211> 55  
          <212> ARN  
          <213> Secuencia artificial

25     <220>  
          <223> Polinucleótido sintético

30     <220>  
          <221> misc\_feature  
          <222> (1)..(1)  
          <223> Modificado con OH

35     <220>  
          <221> misc\_feature  
          <222> (55)..(55)  
          <223> Modificado con ppp

40     <400> 3  
          gggaauaag agagaaaaga agaguccuu uauucucucu uuucuucuca uucuu 55

45     <210> 4  
          <211> 60  
          <212> ARN  
          <213> Secuencia artificial

50     <220>  
          <223> Polinucleótido sintético

55     <220>  
          <221> misc\_feature  
          <222> (1)..(1)  
          <223> Modificado con OH

60     <220>  
          <221> misc\_feature  
          <222> (60)..(60)  
          <223> Modificado con ppp

65     <400> 4  
          gggaauaag agagaaaaga agaguaagaa cccuuuuuuc ucucuuuuuc ucucuuuuuc 60

70     <210> 5  
          <211> 65  
          <212> ARN  
          <213> Secuencia artificial

75     <220>  
          <223> Polinucleótido sintético

5	<220>		
	<221> misc_feature		
	<222> (1)..(1)		
	<223> Modificado con OH		
10	<220>		
	<221> misc_feature		
	<222> (65)..(65)		
	<223> Modificado con ppp		
	<400> 5		
	gggaaaauag agagaaaaga agaguaagaa gaaauccuu uauucucucu uuucuucua	60	
	uuuuu	65	
15	<210> 6		
	<211> 70		
	<212> ARN		
	<213> Secuencia artificial		
20	<220>		
	<223> Polinucleótido sintético		
25	<220>		
	<221> misc_feature		
	<222> (1)..(1)		
	<223> Modificado con OH		
30	<220>		
	<221> misc_feature		
	<222> (70)..(70)		
	<223> Modificado con ppp		
	<400> 6		
	gggaaaauag agagaaaaga agaguaagaa gaaauauag cccuuuauuc ucucuuuucu	60	
	ucucauucuu	70	
35	<210> 7		
	<211> 12		
	<212> ARN		
	<213> Secuencia artificial		
40	<220>		
	<223> Polinucleótido sintético		
45	<220>		
	<221> misc_feature		
	<222> (1)..(1)		
	<223> Modificado con ppp		
	<400> 7		
	gggaaauag ag 12		
50	<210> 8		
	<211> 12		
	<212> ARN		
	<213> Secuencia artificial		
55	<220>		
	<223> Polinucleótido sintético		
60	<220>		
	<221> misc_feature		
	<222> (1)..(1)		
	<223> Modificado con ppp		
	<400> 8		

cucuauuuuc cc 12

5 <210> 9  
<211> 16  
<212> ARN  
<213> Secuencia artificial

10 <220>  
<223> Polinucleótido sintético

15 <220>  
<221> misc\_feature  
<222> (1)..(1)  
<223> Modificado con ppp

20 <220>  
<221> misc\_feature  
<222> (16)..(16)  
<223> Modificado con ppp

25 <400> 9  
gggaaauacc cuuuau 16

30 <210> 10  
<211> 24  
<212> ARN  
<213> Secuencia artificial

35 <220>  
<223> Polinucleótido sintético

40 <220>  
<221> misc\_feature  
<222> (1)..(1)  
<223> Modificado con ppp

45 <220>  
<221> misc\_feature  
<222> (24)..(24)  
<223> Modificado con ppp

50 <400> 10  
gggaaauaag agcccuuuau ucuc 24

55 <210> 11  
<211> 20  
<212> ARN  
<213> Secuencia artificial

60 <220>  
<223> Polinucleótido sintético

65 <220>  
<221> misc\_feature  
<222> (1)..(1)  
<223> Modificado con ppp

<220>  
<221> misc\_feature  
<222> (20)..(20)  
<223> Modificado con ppp

<400> 11  
gggaaauacc cuuuauucuc 20

<210> 12

<211> 20  
 <212> ARN  
 <213> Secuencia artificial

5      <220>  
       <223> Polinucleótido sintético

<220>  
 <221> misc\_feature  
 10      <222> (1)..(1)  
       <223> Modificado con ppp

<220>  
 <221> misc\_feature  
 15      <222> (20)..(20)  
       <223> Modificado con ppp

<400> 12  
 gggaaauaag agcccuuuau 20  
 20

<210> 13  
 <211> 31  
 <212> ARN  
 <213> Secuencia artificial  
 25

<220>  
 <223> Polinucleótido sintético

<220>  
 30      <221> misc\_feature  
       <222> (1)..(1)  
       <223> Modificado con ppp

<220>  
 35      <221> misc\_feature  
       <222> (31)..(31)  
       <223> Modificado con ppp

<400> 13  
 40      gggaaauaag agaaaagaag aguccuuua u 31

<210> 14  
 <211> 35  
 <212> ARN  
 45      <213> Secuencia artificial

<220>  
 <223> Polinucleótido sintético

<220>  
 50      <221> misc\_feature  
       <222> (1)..(1)  
       <223> Modificado con ppp

<220>  
 55      <221> misc\_feature  
       <222> (35)..(35)  
       <223> Modificado con ppp

<400> 14  
 60      gggaaauaag agaaaagaag aguccuuua uucuc 35

<210> 15  
 <211> 40  
 65      <212> ARN  
       <213> Secuencia artificial

<220>  
 <223> Polinucleótido sintético

5      <220>  
          <221> misc\_feature  
          <222> (1)..(1)  
          <223> Modificado con ppp

10     <220>  
          <221> misc\_feature  
          <222> (40)..(40)  
          <223> Modificado con ppp

15     <400> 15  
          gggaaauaag agagaaaaga cccuuuauuc ucucuuuucu 40

<210> 16  
 <211> 41  
 20     <212> ARN  
          <213> Secuencia artificial

<220>  
 <223> Polinucleótido sintético

25     <220>  
          <221> misc\_feature  
          <222> (1)..(1)  
          <223> Modificado con OH

30     <220>  
          <221> misc\_feature  
          <222> (41)..(41)  
          <223> Modificado con ppp

35     <400> 16  
          aaaaaaaaaa aaaaaaaaaa aaaaauuuuu uuuuuuuuuu u 41

40     <210> 17  
          <211> 45  
          <212> ARN  
          <213> Secuencia artificial

<220>  
 <223> Polinucleótido sintético

45     <220>  
          <221> misc\_feature  
          <222> (1)..(1)  
          <223> Modificado con OH

50     <220>  
          <221> misc\_feature  
          <222> (45)..(45)  
          <223> Modificado con ppp

55     <400> 17  
          aaaaaaaaaa aaaaaaaaaa aaaaauuuuu uuuuuuuuuu uuuuu 45

60     <210> 18  
          <211> 49  
          <212> ARN  
          <213> Secuencia artificial

65     <220>  
          <223> Polinucleótido sintético

5      <220>  
       <221> misc\_feature  
       <222> (1)..(1)  
       <223> Modificado con OH

10     <220>  
       <221> misc\_feature  
       <222> (49)..(49)  
       <223> Modificado con ppp

15     <400> 18  
       aaaaaaaaaa aaaaaaaaaa aaaaauuuuu uuuuuuuuuu uuuuuuuuu 49

20     <210> 19  
       <211> 43  
       <212> ARN  
       <213> Secuencia artificial

25     <220>  
       <221> misc\_feature  
       <222> (1)..(1)  
       <223> Modificado con ppp

30     <220>  
       <221> misc\_feature  
       <222> (43)..(43)  
       <223> Modificado con ppp

35     <400> 19  
       gggaaauaag agaaaagaag aguccuuua uucucucuuu ucu 43

40     <210> 20  
       <211> 40  
       <212> ARN  
       <213> Secuencia artificial

45     <220>  
       <221> misc\_feature  
       <222> (1)..(1)  
       <223> Modificado con ppp

50     <220>  
       <221> misc\_feature  
       <222> (40)..(40)  
       <223> Modificado con ppp

55     <400> 20  
       gggaaauaag agagaaaaga cccuuuuuuc ucucuuuuu 40

60     <210> 21  
       <211> 40  
       <212> ARN  
       <213> Secuencia artificial

65     <220>  
       <221> misc\_feature

<222> (1)..(1)  
 <223> Modificado con ppp  
  
 5      <220>  
       <221> misc\_feature  
       <222> (40)..(40)  
       <223> Modificado con OH  
  
 10     <400> 21  
       gggaaauaag agagaaaaga cccuuuauuc ucucuuuucu 40  
  
       <210> 22  
       <211> 40  
       <212> ARN  
 15     <213> Secuencia artificial  
  
       <220>  
       <223> Polinucleótido sintético  
  
 20     <220>  
       <221> misc\_feature  
       <222> (1)..(1)  
       <223> Modificado con OH  
  
 25     <220>  
       <221> misc\_feature  
       <222> (40)..(40)  
       <223> Modificado con ppp  
  
 30     <400> 22  
       gggaaauaag agagaaaaga cccuuuauuc ucucuuuucu 40  
  
       <210> 23  
       <211> 45  
 35     <212> ARN  
       <213> Secuencia artificial  
  
       <220>  
       <223> Polinucleótido sintético  
 40     <220>  
       <221> misc\_feature  
       <222> (1)..(1)  
       <223> Modificado con OH  
 45     <220>  
       <221> misc\_feature  
       <222> (45)..(45)  
       <223> Modificado con ppp  
 50     <400> 23  
       gggaaauaag agagaaaaga agagucccuu uauucucucu uuucu 45  
  
       <210> 24  
 55     <211> 50  
       <212> ARN  
       <213> Secuencia artificial  
  
       <220>  
 60     <223> Polinucleótido sintético  
  
       <220>  
       <221> misc\_feature  
       <222> (1)..(1)  
 65     <223> Modificado con OH

<220>  
 <221> misc\_feature  
 <222> (50)..(50)  
 <223> Modificado con ppp  
 5  
 <400> 24  
 gggaaauaag agagaaaaga agaguaagaa cccuuuauuc ucucuuuucu 50  
 <210> 25  
 10 <211> 55  
 <212> ARN  
 <213> Secuencia artificial  
 <220>  
 15 <223> Polinucleótido sintético  
 <220>  
 <221> misc\_feature  
 <222> (1)..(1)  
 20 <223> Modificado con OH  
 <220>  
 <221> misc\_feature  
 <222> (55)..(55)  
 25 <223> Modificado con ppp  
 <400> 25  
 gggaaauaag agagaaaaga agaguaagaa gaaauccuu uauucucucu uuucu 55  
 30 <210> 26  
 <211> 60  
 <212> ARN  
 <213> Secuencia artificial  
 35 <220>  
 <223> Polinucleótido sintético  
 <220>  
 <221> misc\_feature  
 40 <222> (1)..(1)  
 <223> Modificado con OH  
 <220>  
 <221> misc\_feature  
 45 <222> (60)..(60)  
 <223> Modificado con ppp  
 <400> 26  
 gggaaauaag agagaaaaga agaguaagaa gaaauaauag cccuuuauuc ucucuuuucu 60  
 50 <210> 27  
 <211> 67  
 <212> ARN  
 <213> Secuencia artificial  
 55 <220>  
 <223> Polinucleótido sintético  
 <220>  
 <221> misc\_feature  
 60 <222> (1)..(1)  
 <223> Modificado con OH  
 <220>  
 65 <221> misc\_feature  
 <222> (67)..(67)



	<223> Modificado con ppp	
	<400> 27 gggaaaauaag agagaaaaga agaguaagaa gaaauauaag agccaccccc uuuaauucucu	60
5	cuuuucu	67
	<210> 28 <211> 26 <212> ARN <213> Secuencia artificial	
10	<220> <223> Polinucleótido sintético	
15	<400> 28 gacucuucuu uucucucuua uuuccc 26	
	<210> 29 <211> 31 <212> ARN <213> Secuencia artificial	
20	<220> <223> Polinucleótido sintético	
25	<400> 29 guucuuaucuc uucuuuucuc ucuauuuucc c 31	
	<210> 30 <211> 36 <212> ARN <213> Secuencia artificial	
30	<220> <223> Polinucleótido sintético	
35	<400> 30 gauuucuucu uacucuucuu uucucucuua uuuccc 36	
40	<210> 31 <211> 60 <212> ARN <213> Secuencia artificial	
45	<220> <223> Polinucleótido sintético	
	<400> 31 gggaaaauaag agagaaaaga agaguaagaa gaaauauaag agccaccaug ggagugcacg 60	
50	<210> 32 <211> 60 <212> ARN <213> Secuencia artificial	
55	<220> <223> Polinucleótido sintético	
60	<400> 32 cgugcacucc caugguggcu cuuauuuuc uucuuacucu ucuuuucucu cuuauuuucc 60	

## REIVINDICACIONES

## 1. Un método de preparación de ARN que comprende

- 5 (a) formar una mezcla de reacción que comprenda una plantilla de ADN, una polimerasa de ARN seleccionada de la polimerasa T7 o la polimerasa T3, y NTP que incluyen adenosina trifosfato (ATP), citidina trifosfato (CTP), uridina trifosfato (UTP), guanosina trifosfato (GTP) y opcionalmente guanosina difosfato (GDP) o un análogo de cada nucleótido respectivo, y un tampón que contiene magnesio, y
- 10 (b) incubar la mezcla de reacción en condiciones tales que el ARN se transcriba, produciendo así una preparación que comprenda ARN transcrito *in vitro* (IVT),

en donde:

- 15 1) la concentración de GTP es al menos 2 veces mayor que la concentración de cada uno de ATP, CTP y UTP;
- 2) la mezcla de reacción comprende además guanosina difosfato (GDP) y en donde la concentración de GDP es al menos 2 veces mayor que la concentración de cada uno de ATP, CTP y UTP; o
- 20 3) la mezcla de reacción comprende además GDP, y en donde la relación entre la concentración de GTP más GDP y la concentración de cada uno de ATP, CTP y UTP es de al menos 2:1.

2. El método de la reivindicación 1, en donde la concentración de GTP es al menos 2 veces mayor que la concentración de cada uno de ATP, CTP y UTP.

25 3. El método de la reivindicación 1, en donde la relación entre la concentración de GTP y la concentración de cada uno de ATP, CTP y UTP es de al menos 3:1, al menos 4:1, al menos 5:1 o al menos 6:1, para producir el ARN.

30 4. El método de la reivindicación 1, en donde la relación entre la concentración de GTP y la concentración de ATP, CTP y UTP es de 2:1, 4:1 y 4:1, respectivamente; o en donde la relación entre la concentración de GTP y la concentración de ATP, CTP y UTP es de 3:1 6:1 y 6:1, respectivamente.

35 5. El método de la reivindicación 1, en donde la mezcla de reacción comprende GTP y GDP y en donde la relación de la concentración entre GTP más GDP y la concentración de uno cualquiera de ATP, CTP o UTP es de al menos 3:1, al menos 4:1, al menos 5:1 o al menos 6:1; o en donde la relación entre la concentración de GTP más GDP y la concentración de ATP, CTP y UTP es de 3:1, 6:1 y 6:1, respectivamente.

40 6. El método de la reivindicación 1, en donde más de alrededor del 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, 99,5 % o 99,8 % de la masa del ARN comprende transcripciones de longitud completa de cadena simple.

45 7. El método de la reivindicación 1, en donde la preparación comprende una población de transcripciones de ARN parcial de cadena simple en una orientación sentido y en donde más del 80 % de la población de transcripciones de ARN parcial de cadena simple en una orientación sentido tiene una longitud de nucleótidos de 100 nucleótidos o menor.

8. El método de la reivindicación 7, en donde (i) más del 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 %, o (ii) el 100 % de la población de transcripciones de ARN parcial de cadena simple en una orientación sentido tiene una longitud de nucleótidos de 100 nucleótidos o menor.

50 9. El método de la reivindicación 1, en donde:

- (i) menos del 0,25 % de la masa del ARN en la preparación es contaminante de ARN inductor de citocinas;
- o
- 55 (ii) menos del 0,5 % de la masa del ARN en la preparación es un producto de transcripción de complemento inverso.

10. El método de la reivindicación 9, en donde el producto de transcripción de complemento inverso es un ARNcd que comprende:

- 60 (i) una hebra que codifica una secuencia que es un complemento inverso de al menos una parte del ARN IVT; o
- (ii) una hebra que comprende una secuencia poliU.

65 11. El método de la reivindicación 10, en donde (i) la hebra que codifica la secuencia que es el complemento inverso del ARN IVT se inicia con un trifosfato en 5' (5'-PPP), o (ii) la hebra que comprende la secuencia de poliU se inicia con un trifosfato en 5' (5'-PPP).

5

12. El método de la reivindicación 9, en donde el producto de transcripción de complemento inverso comprende un complemento inverso del extremo 5' del ARN IVT y/o un complemento inverso del extremo 3' del ARN que codifica el polipéptido de interés.

13. El método de una cualquiera de las reivindicaciones 10-12, en donde el ARN IVT es un ARN que codifica un polipéptido de interés y en donde el complemento inverso del ARN que codifica el polipéptido de interés comprende una secuencia complementaria a la totalidad o a una parte de un marco de lectura abierto del ARN que codifica el polipéptido de interés.

Figura 1

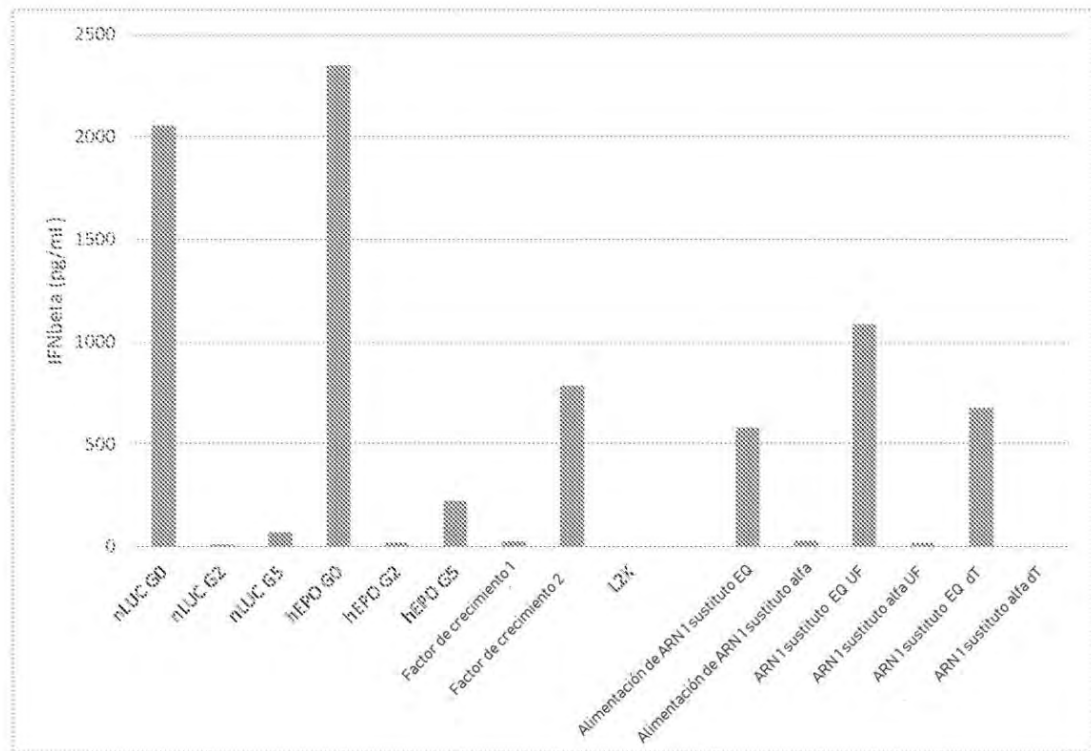


Figura 2.

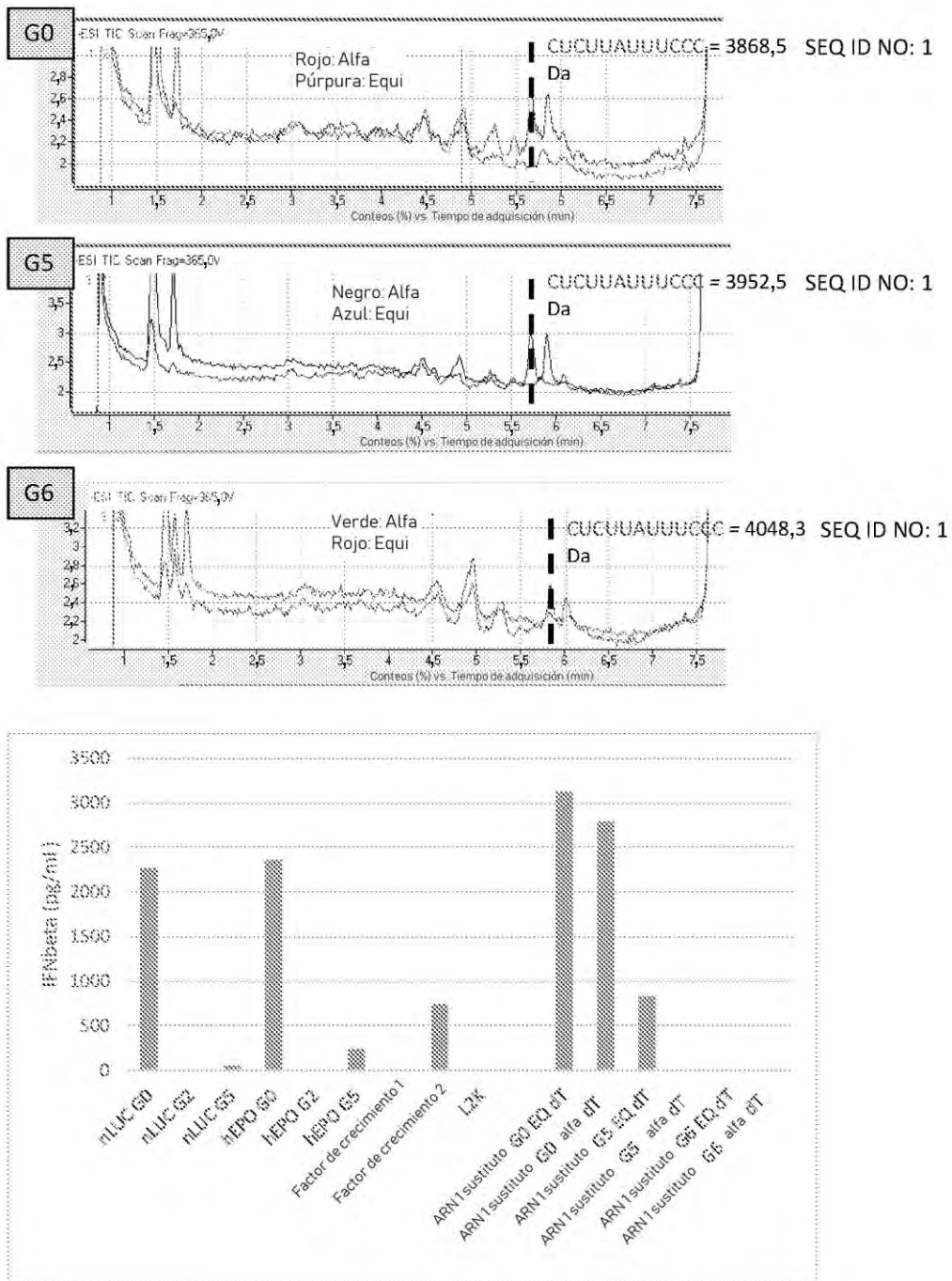


Figura 3.

ARN 2  
sustituto EQ

ARN 2  
sustituto alfa

hEPO EQ

hEPO alfa

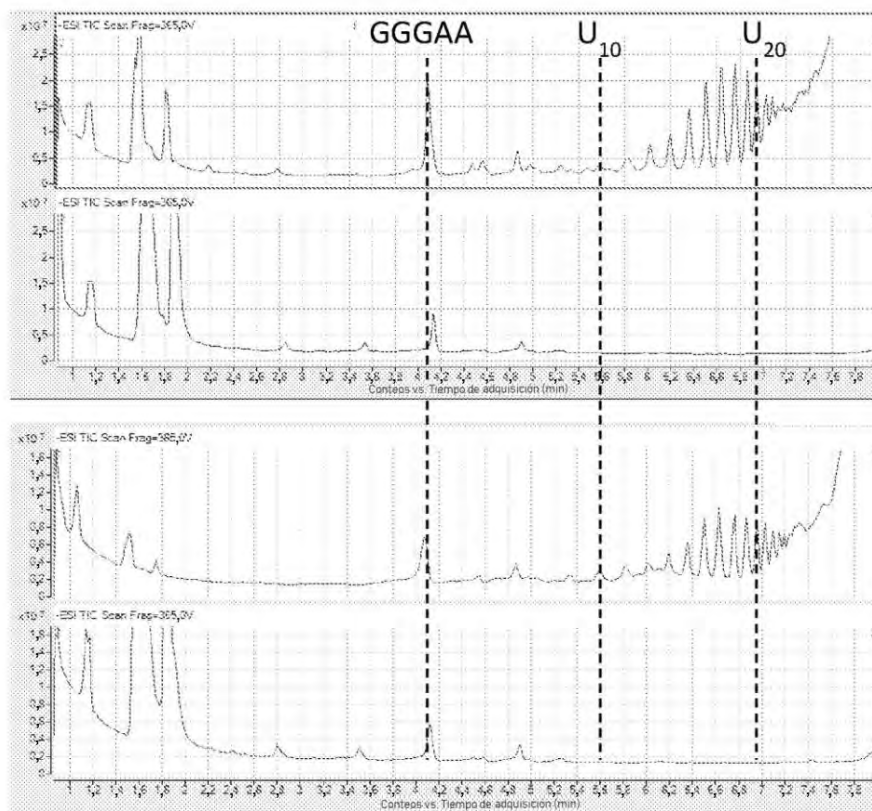


Figura 4.

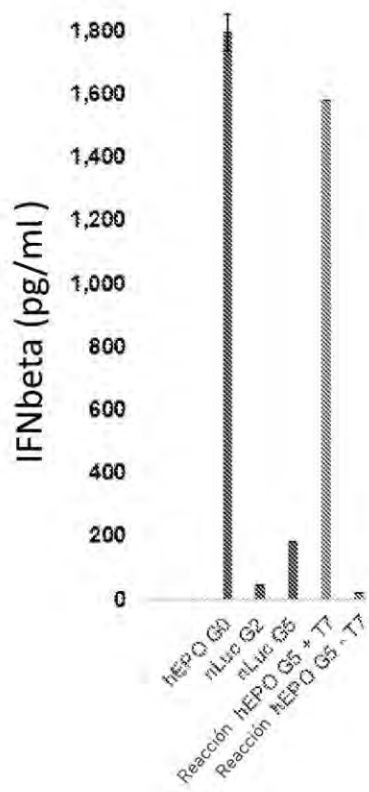
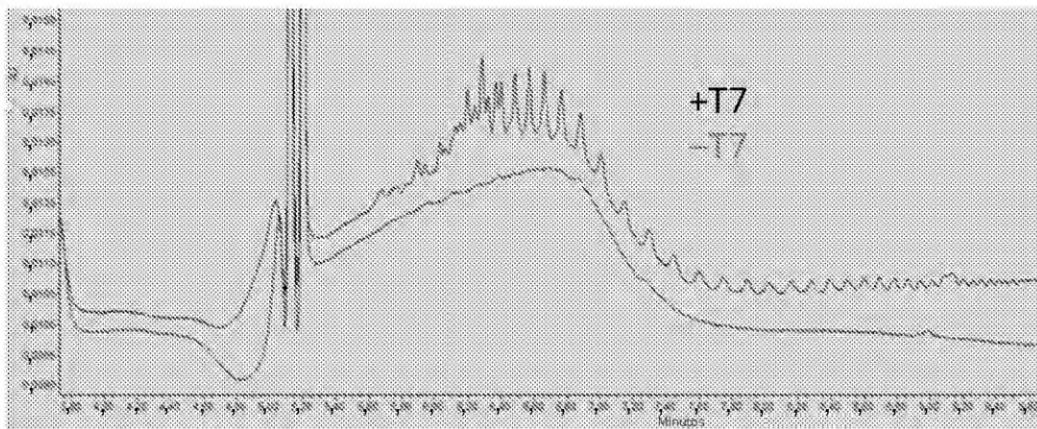


Figura 5A

La aparición de especies de MW menor en las trazas azules (+RNasa III) se correlaciona con cantidad de impureza de ARNcd en la muestra

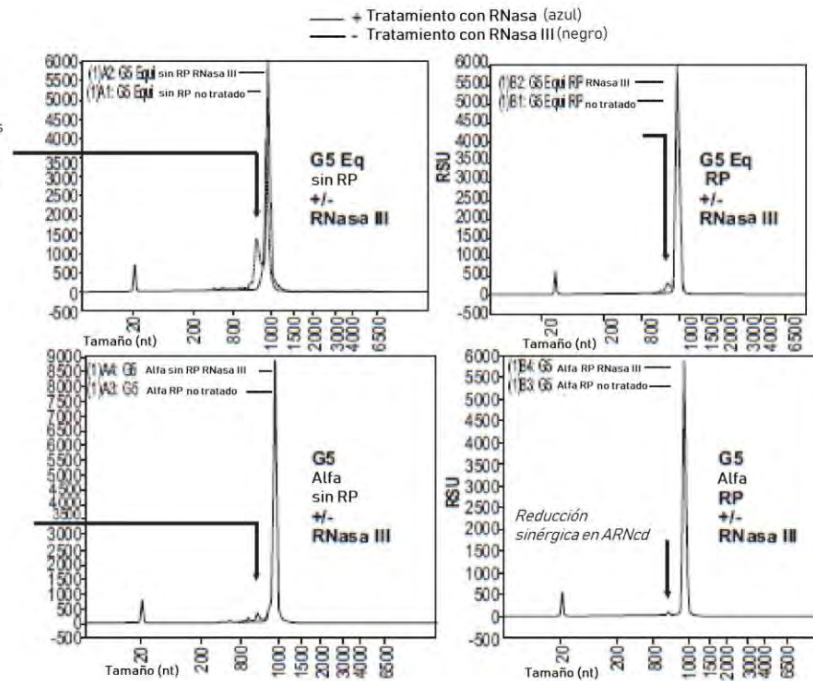




Figura 5B.

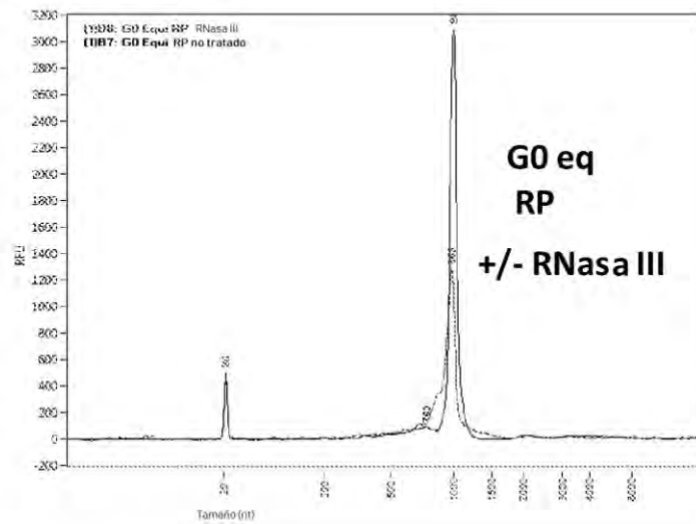
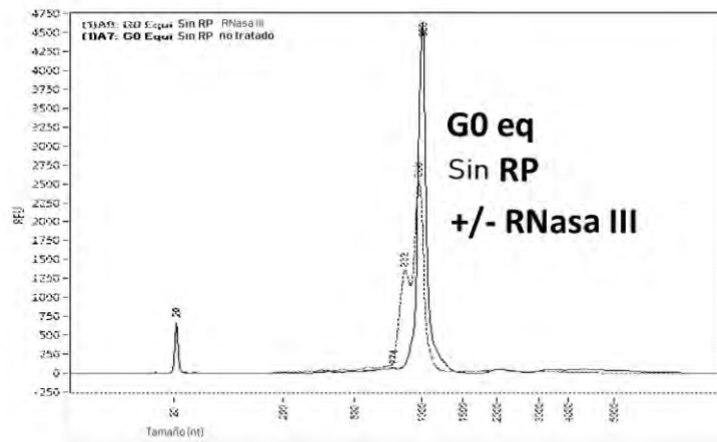


Figura 6.

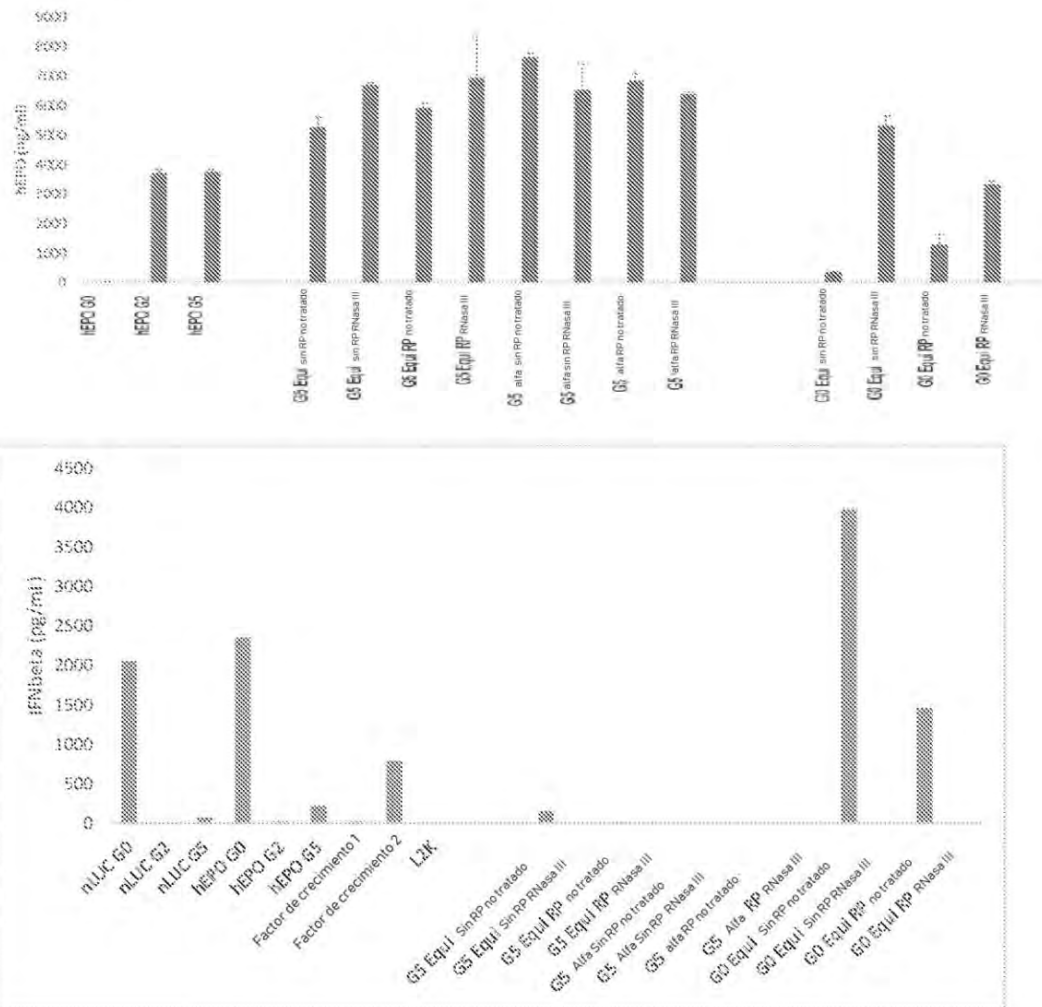


Figura 7.

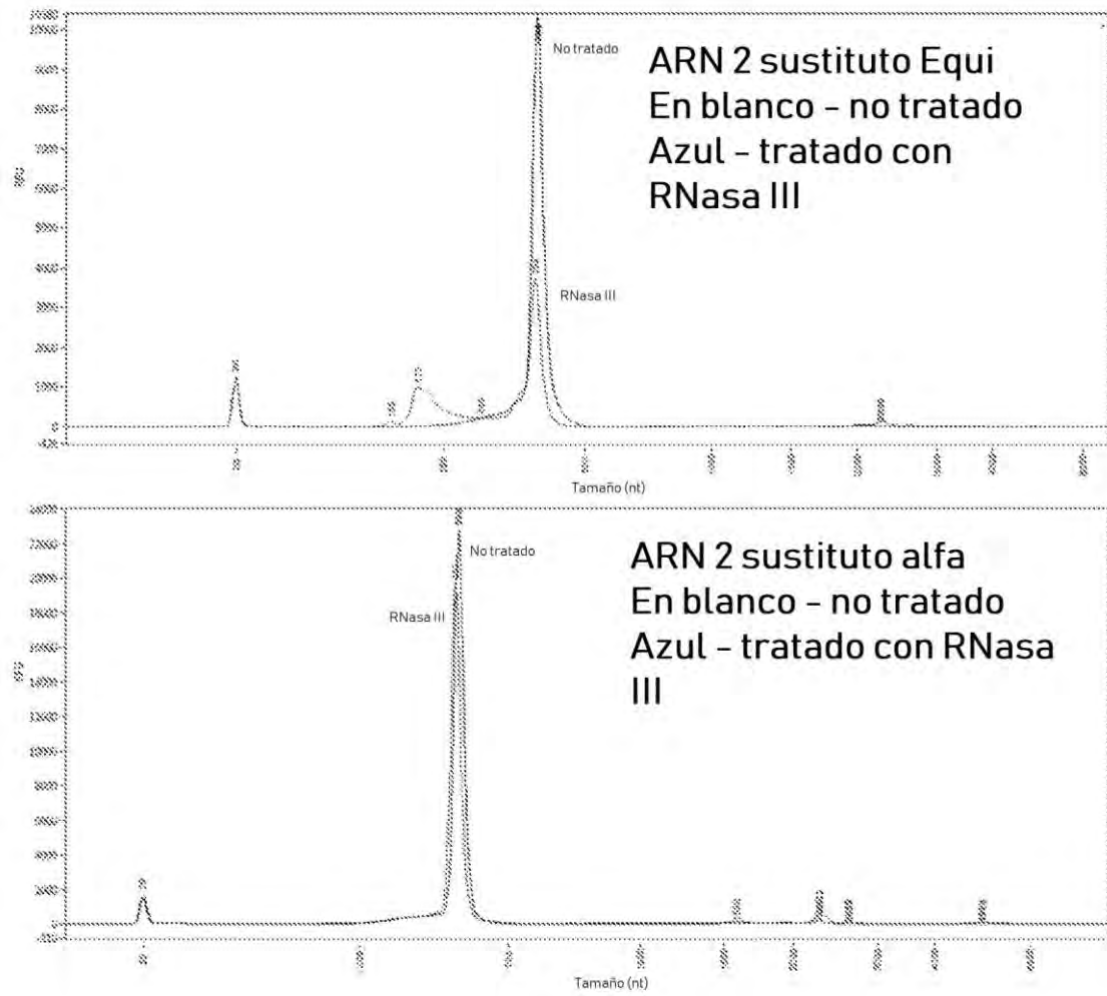


Figura 8A.

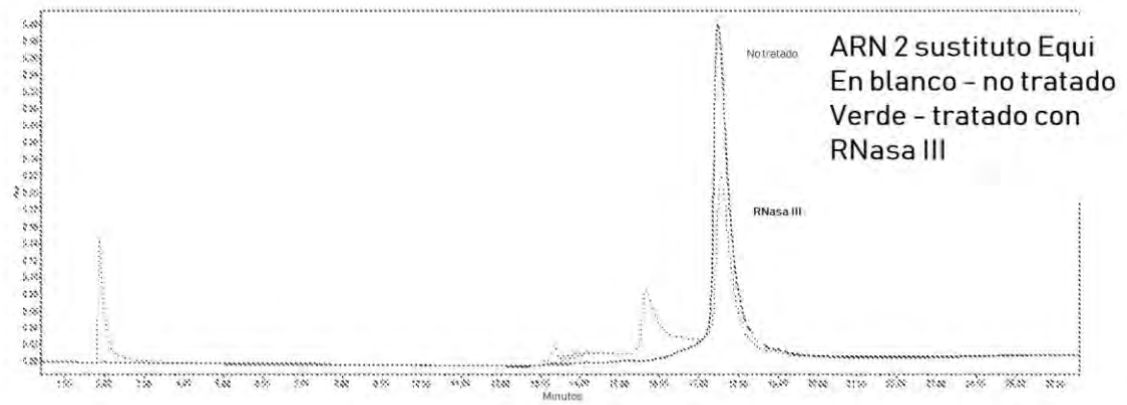


Figura 8B.

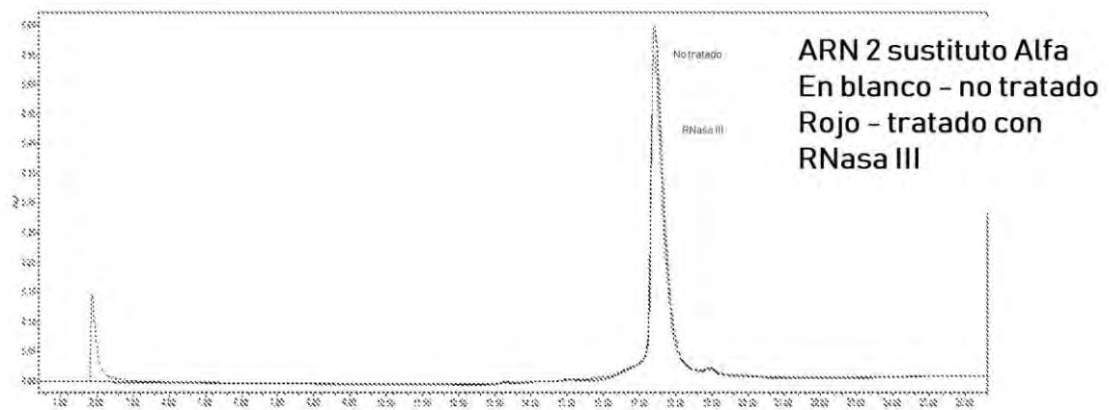


Figura 9.

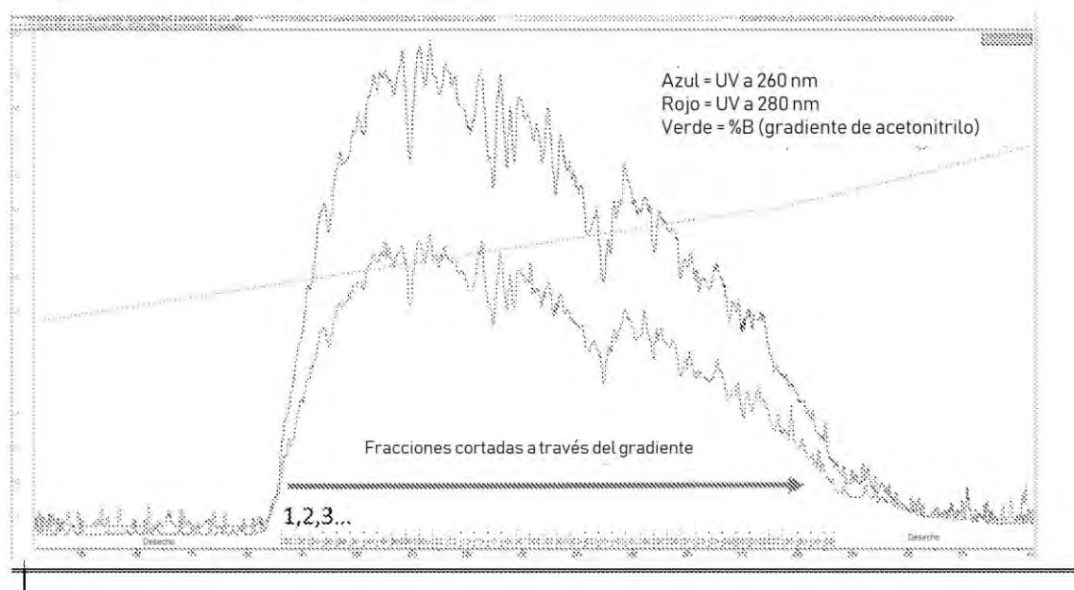


Figura 10A.

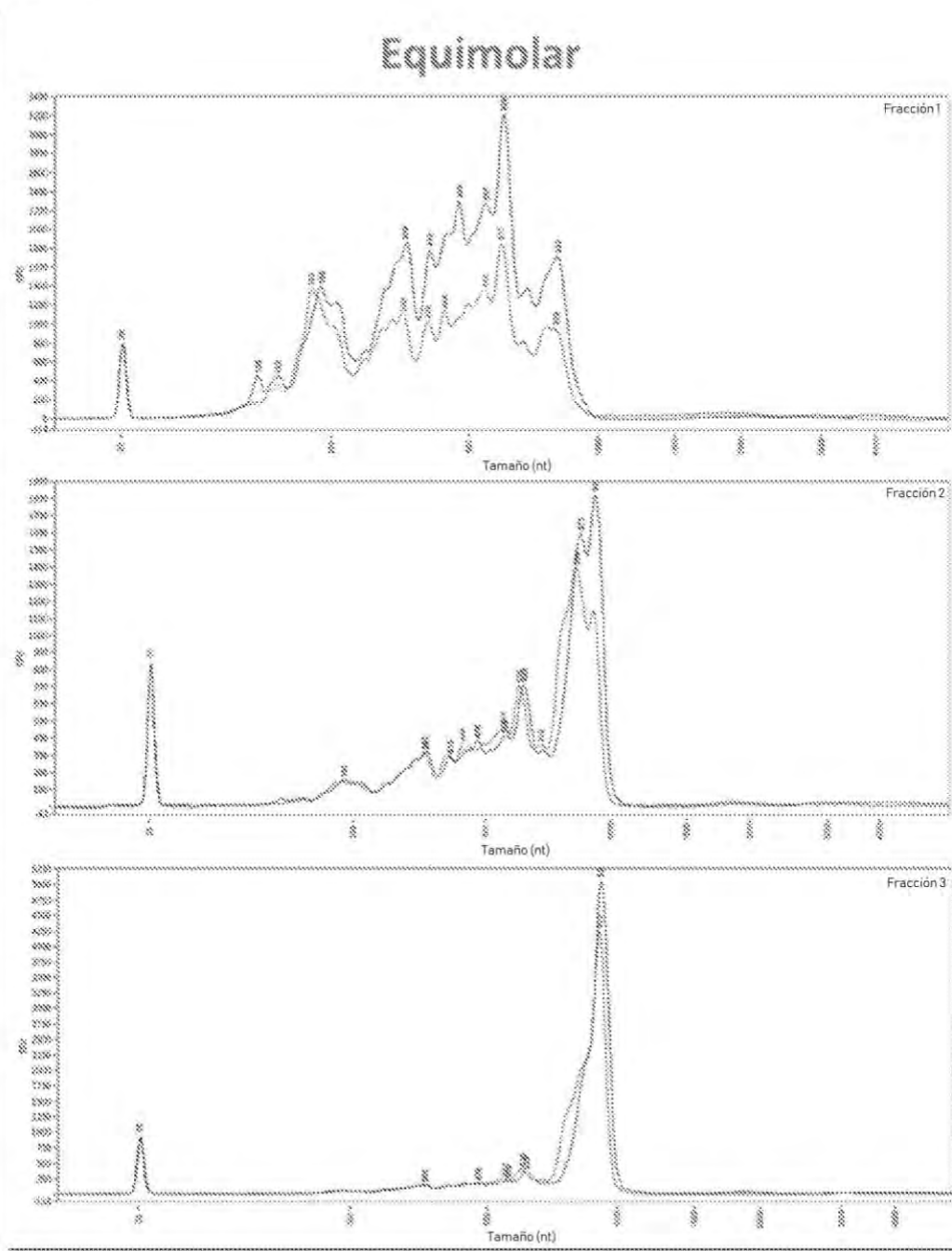


Figura 10B.

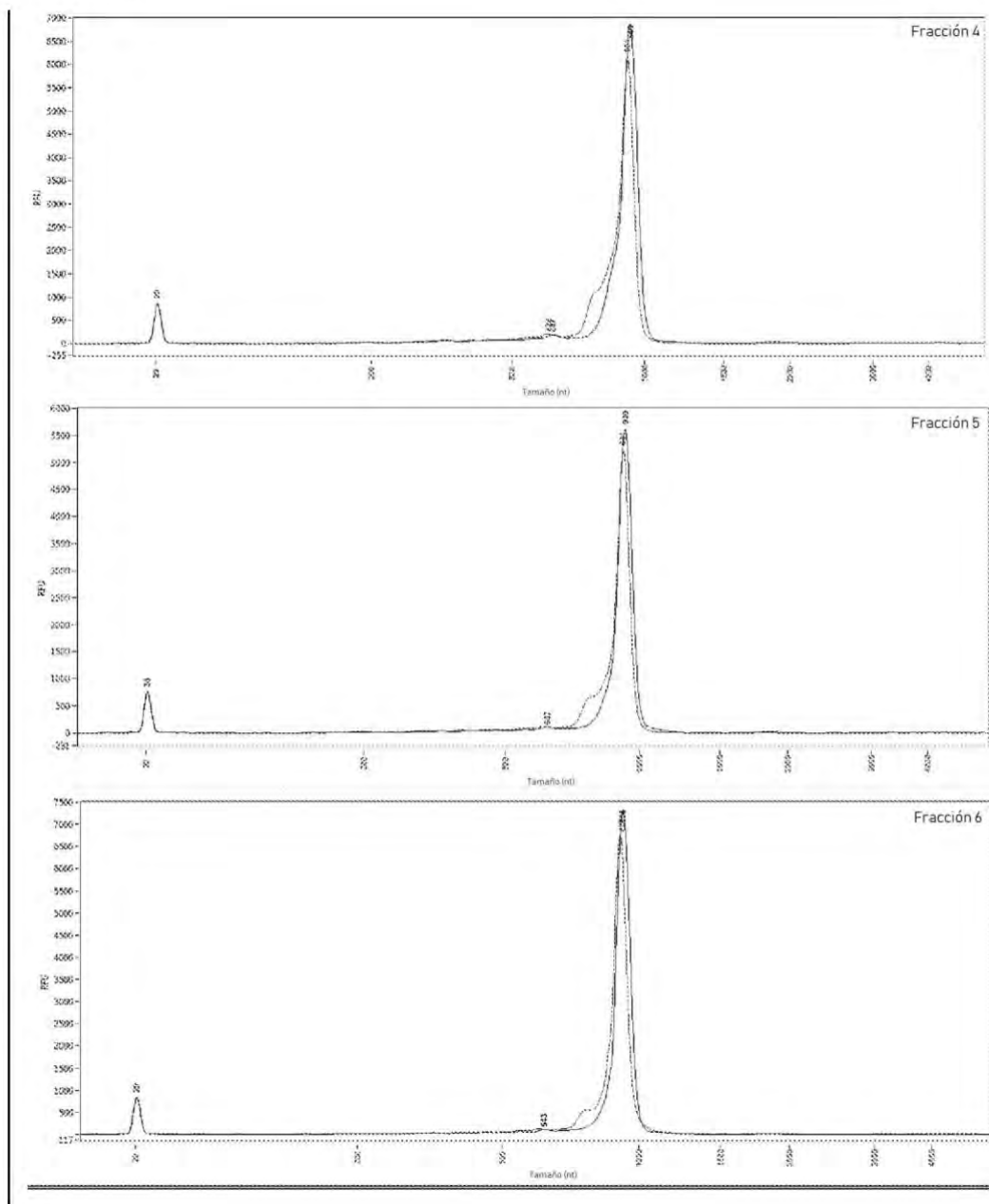




Figura 10C.

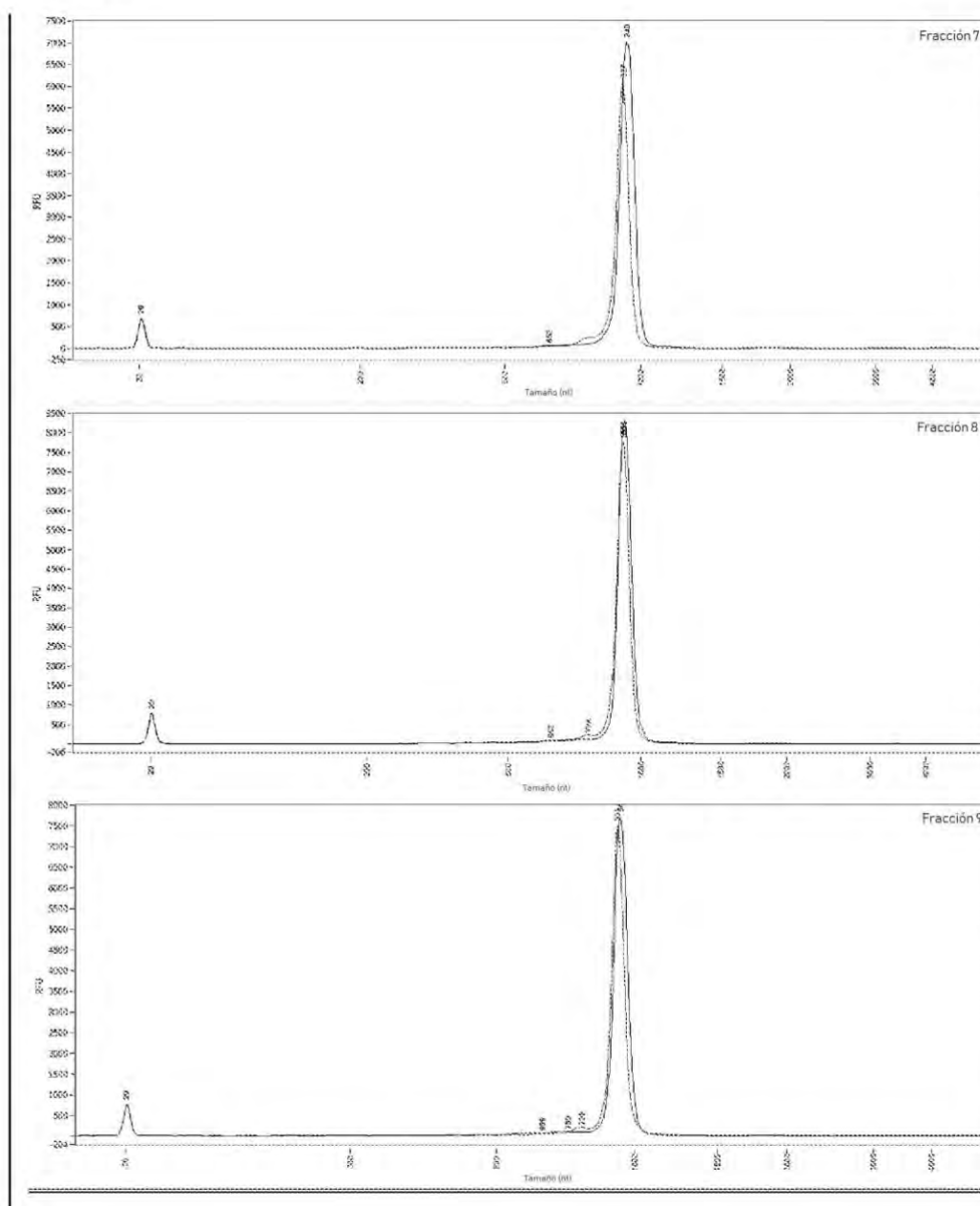


Figura 10D.

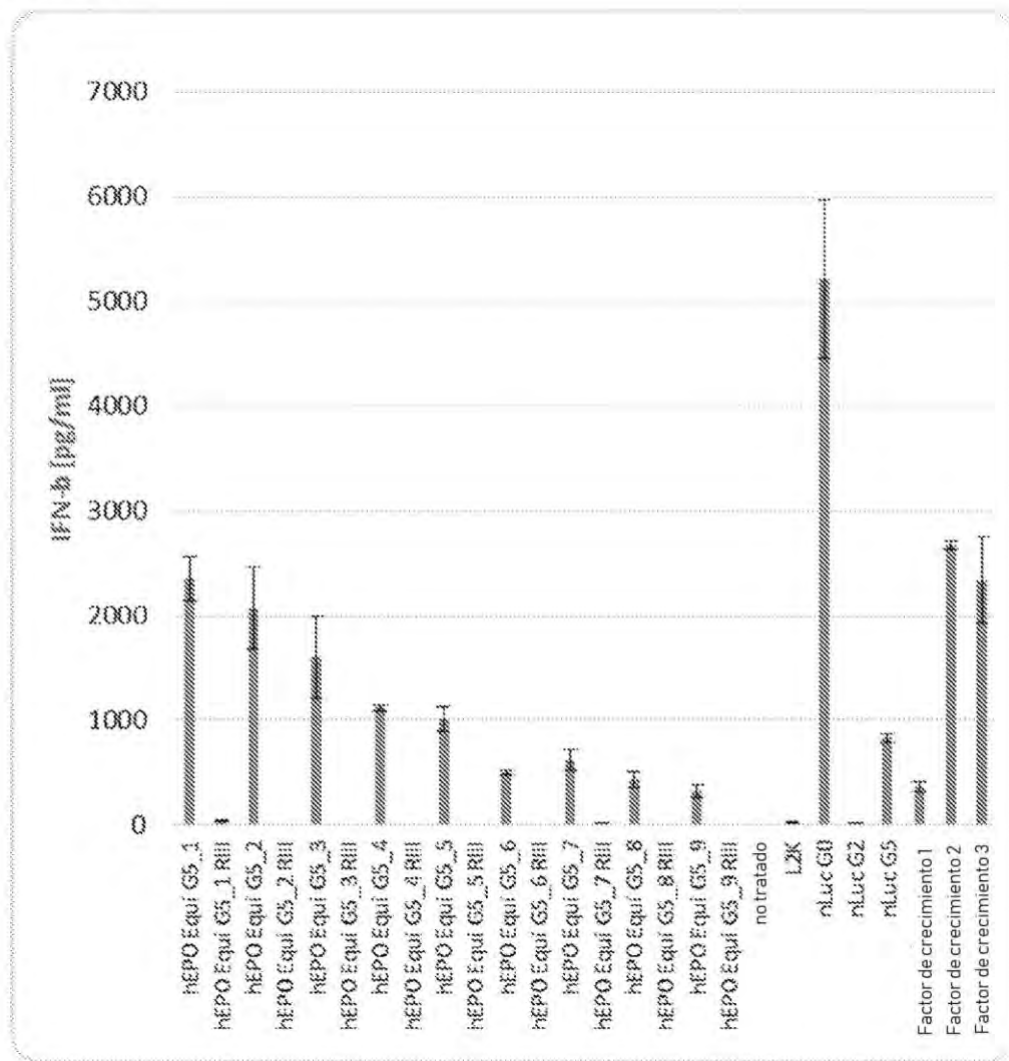


Figura 11A.

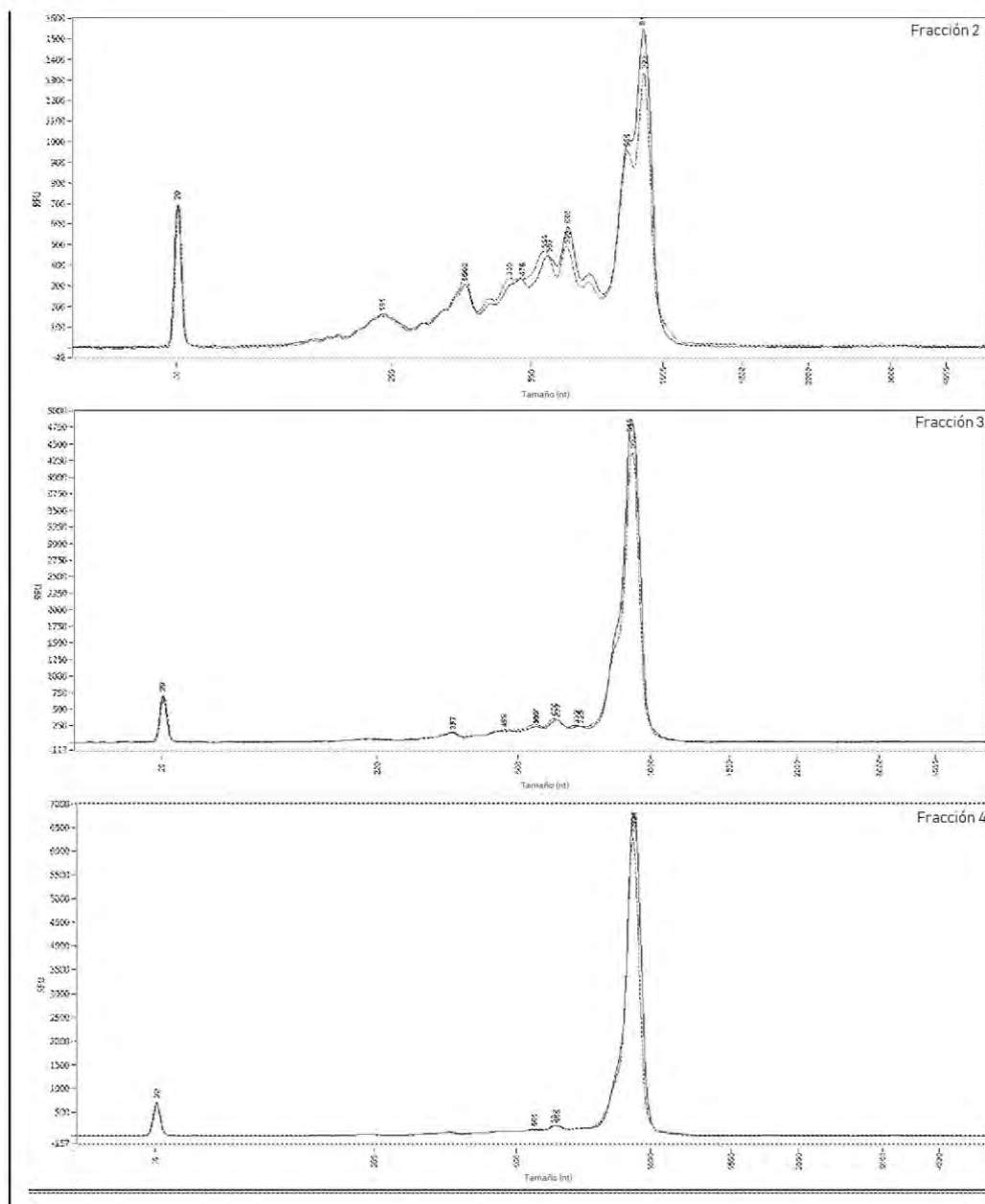


Figura 11B.

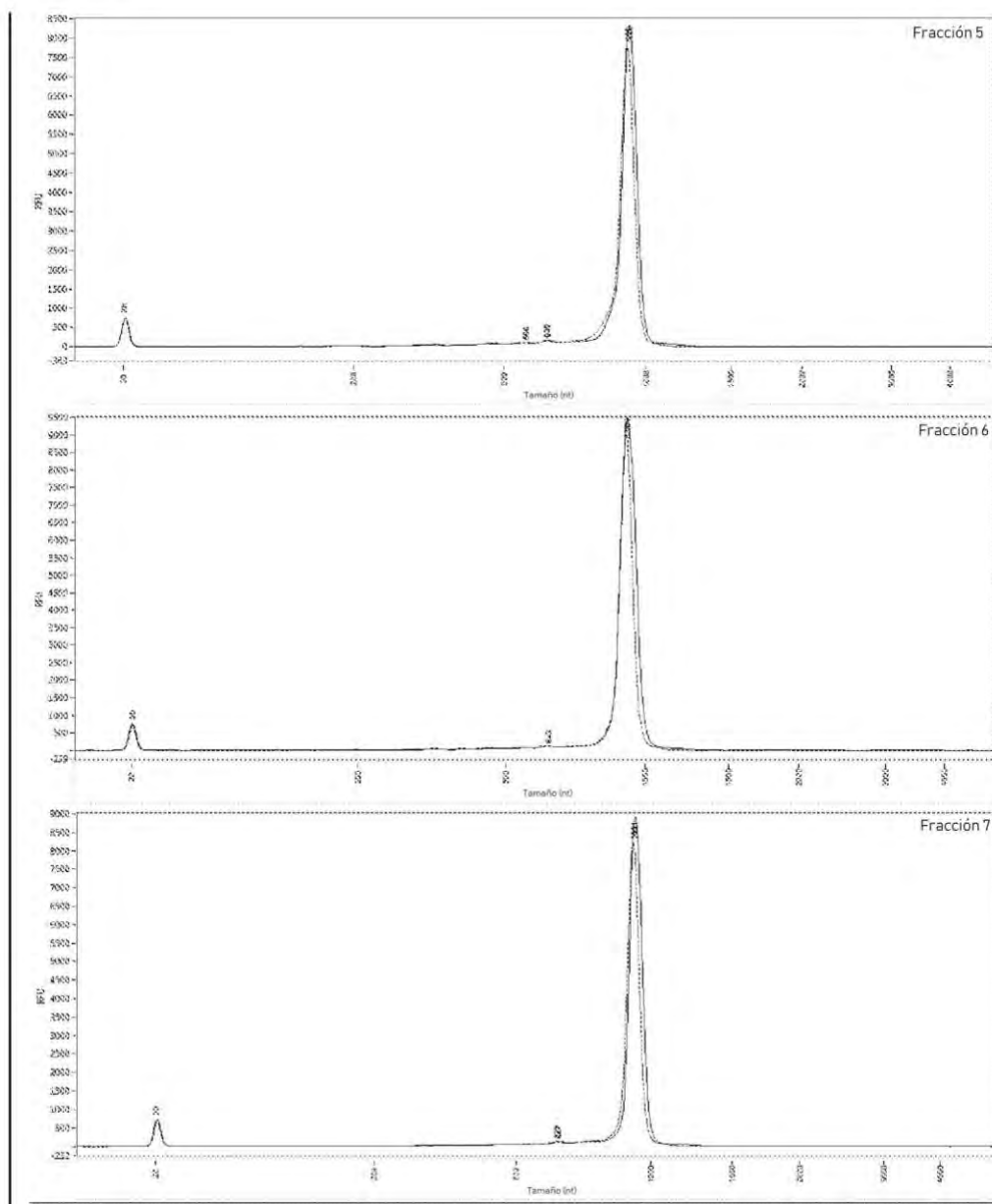


Figura 11C.

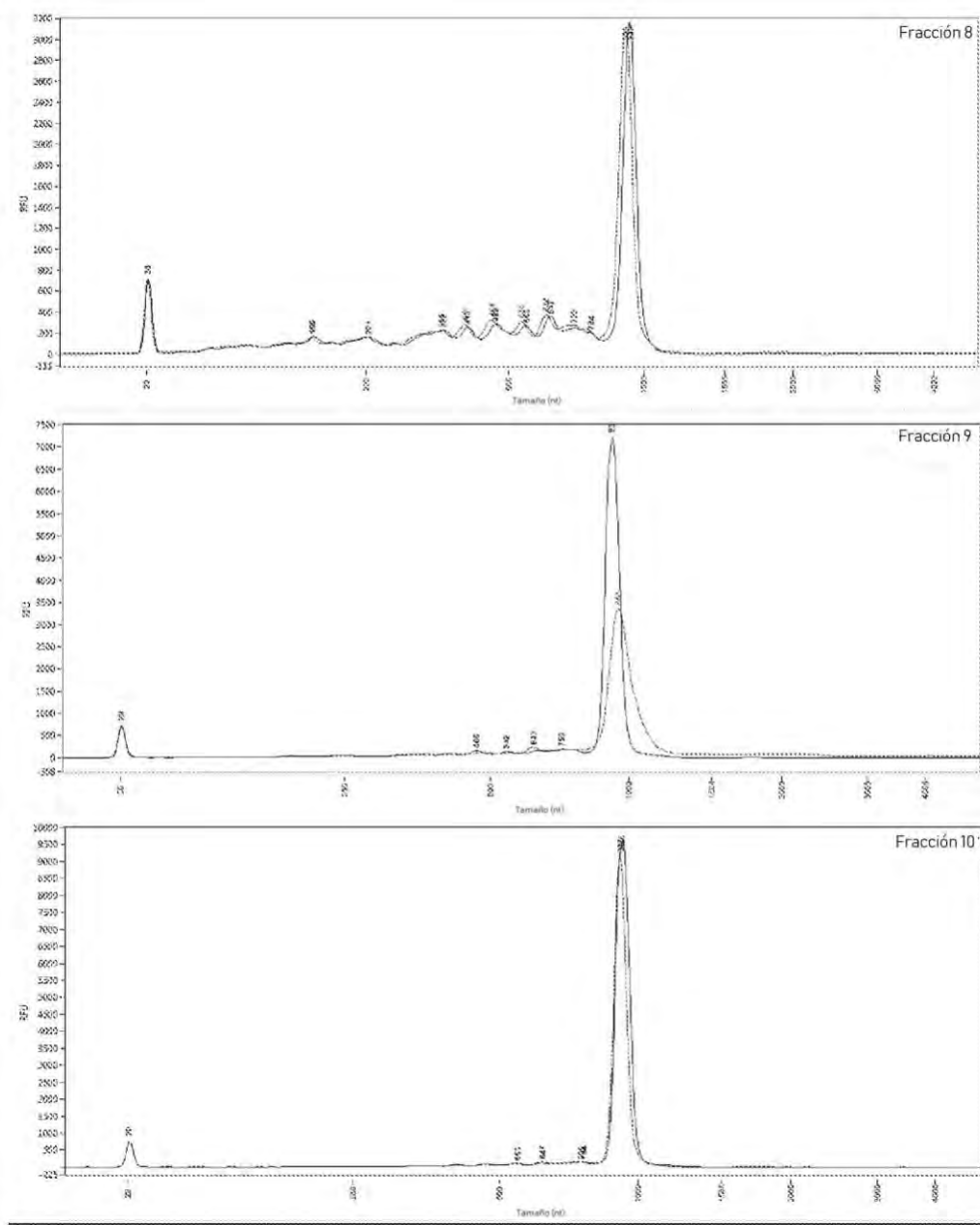


Figura 11C.

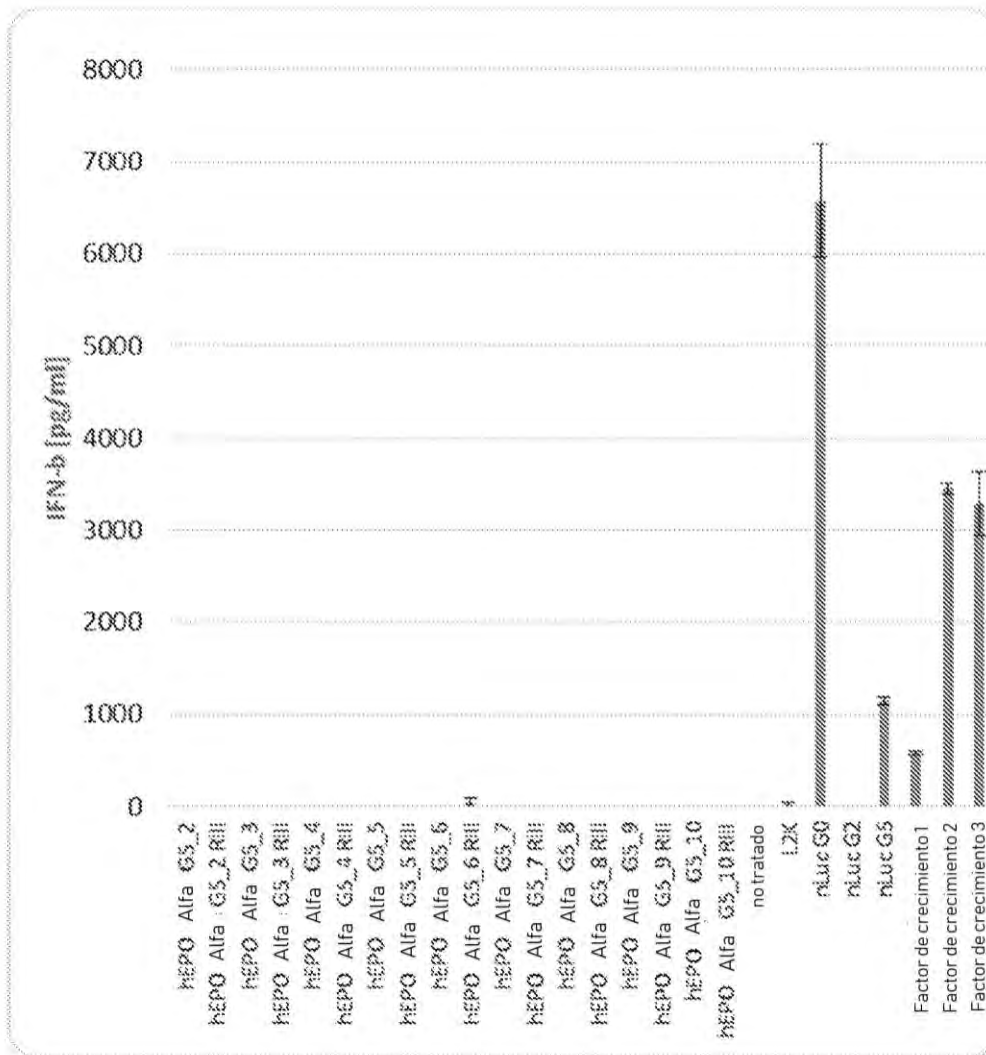


Figura 12.

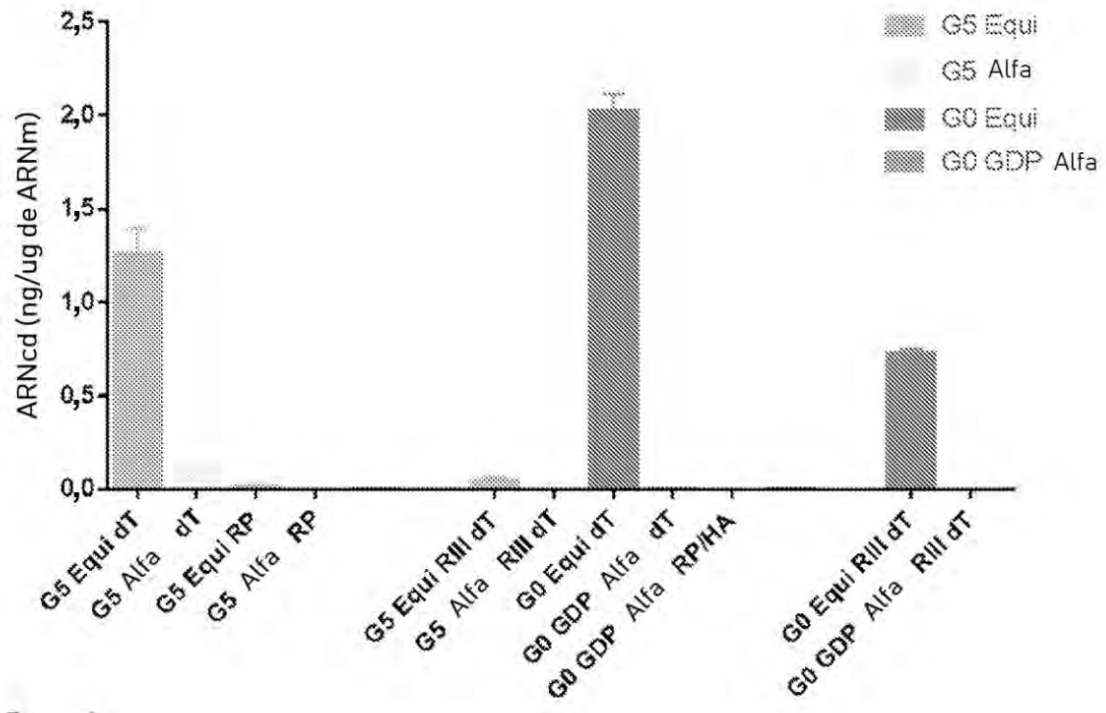


Figura 13.

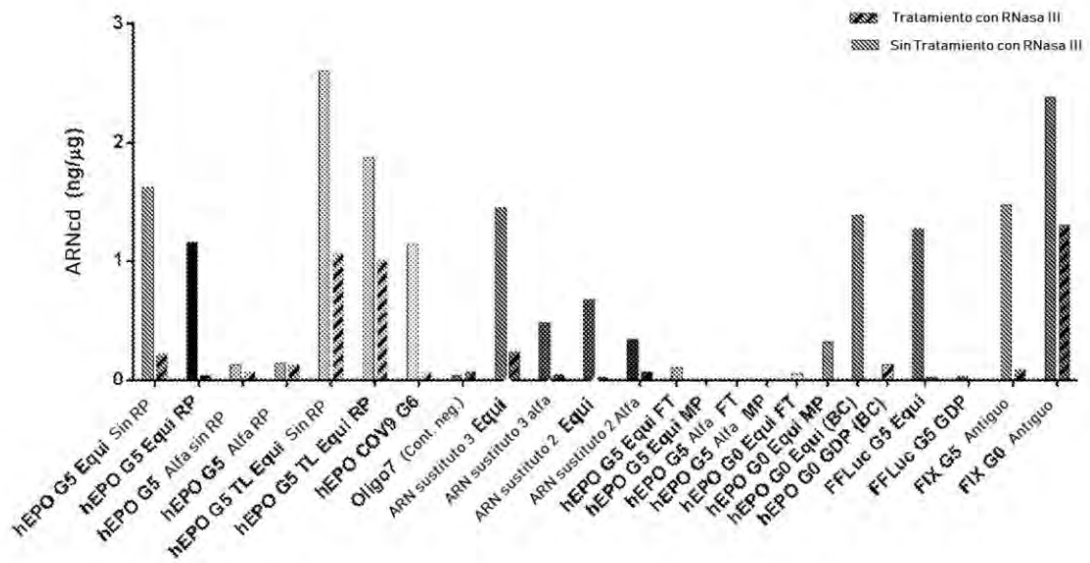




Figura 14.

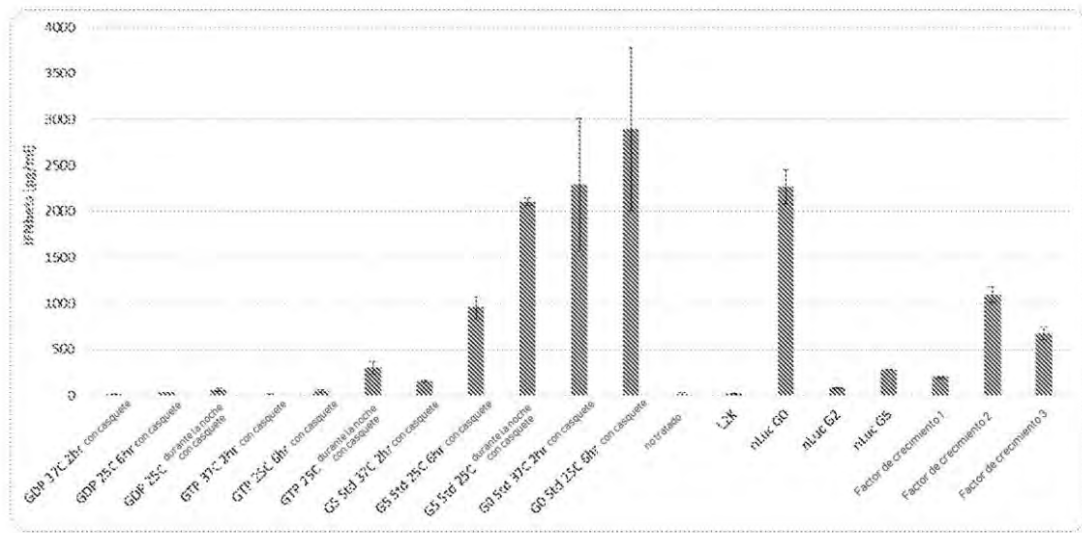


Figura 15.

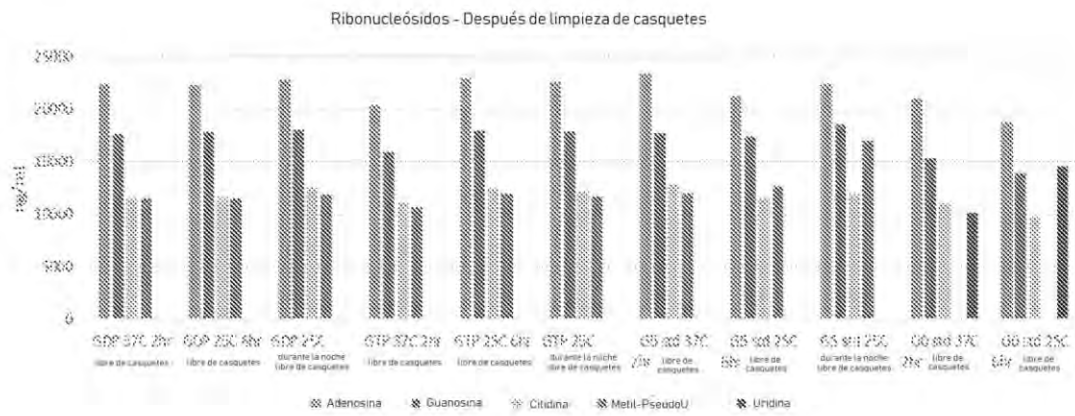


Figura 16A.

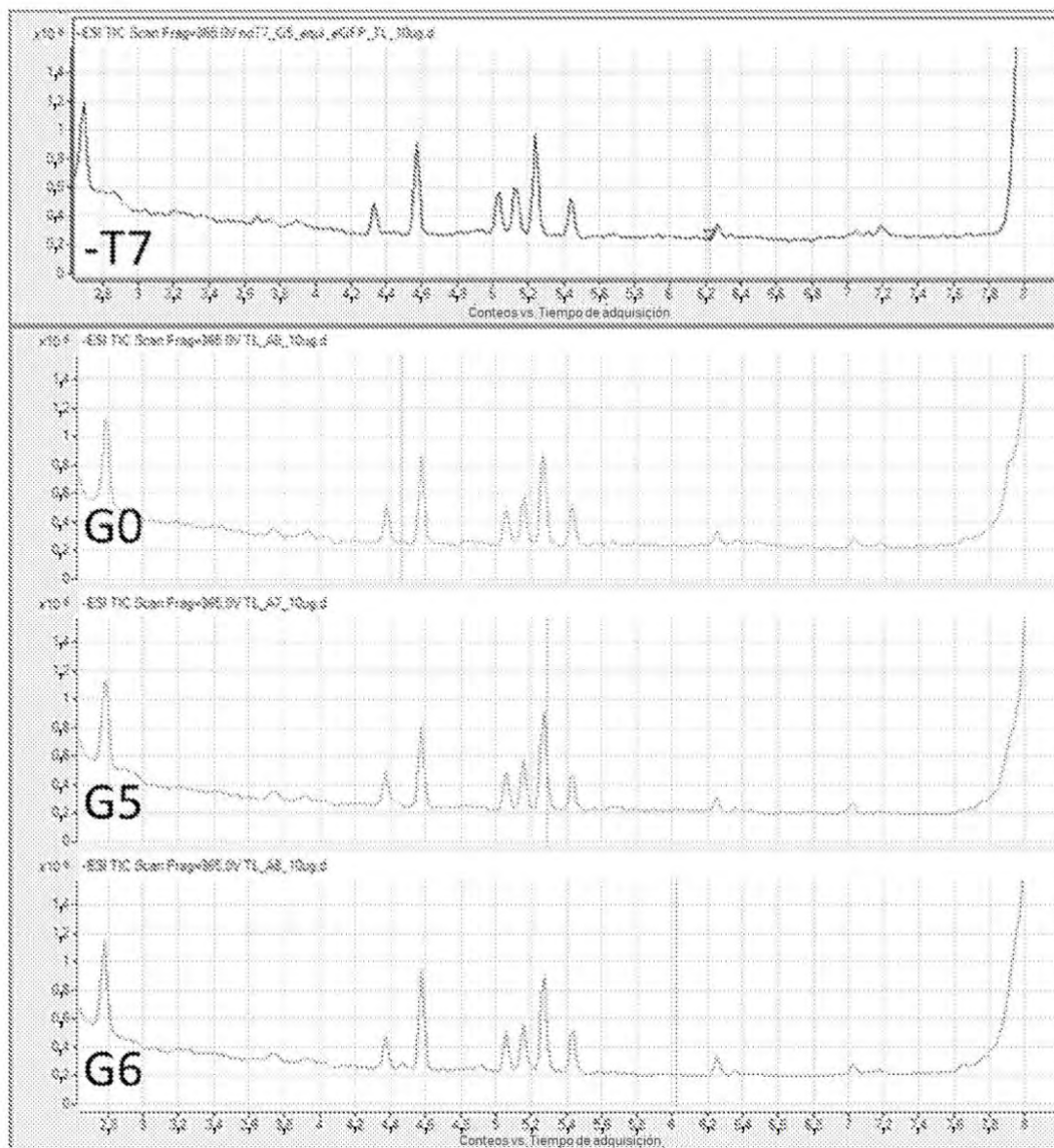
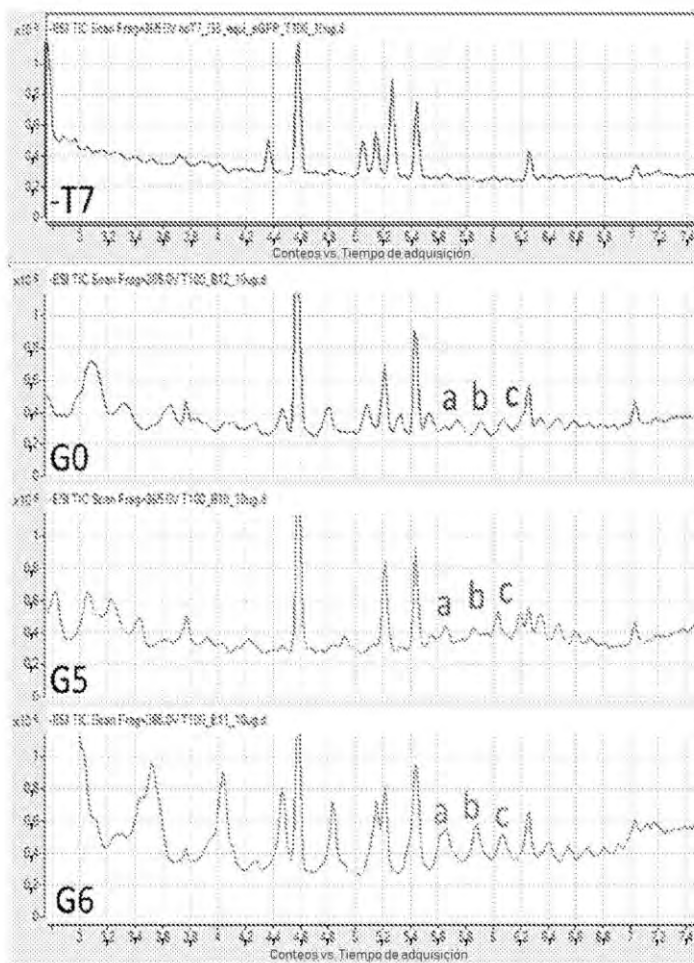


Figura 16B.



	ID
-T7	Ninguna
G0 (a)	ppp-U <sub>11</sub> -OH
G0 (b)	ppp-U <sub>14</sub> -OH
G0 (c)	ppp-U <sub>15</sub> -OH
G5 (a)	ppp-U <sub>12</sub> -OH
G5 (b)	ppp-U <sub>13</sub> -OH
G5 (c)	ppp-U <sub>14</sub> -OH
G6 (a)	ppp-U <sub>11</sub> -OH
G6 (b)	ppp-U <sub>12</sub> -OH
G6 (c)	ppp-U <sub>13</sub> -OH

Figura 17.

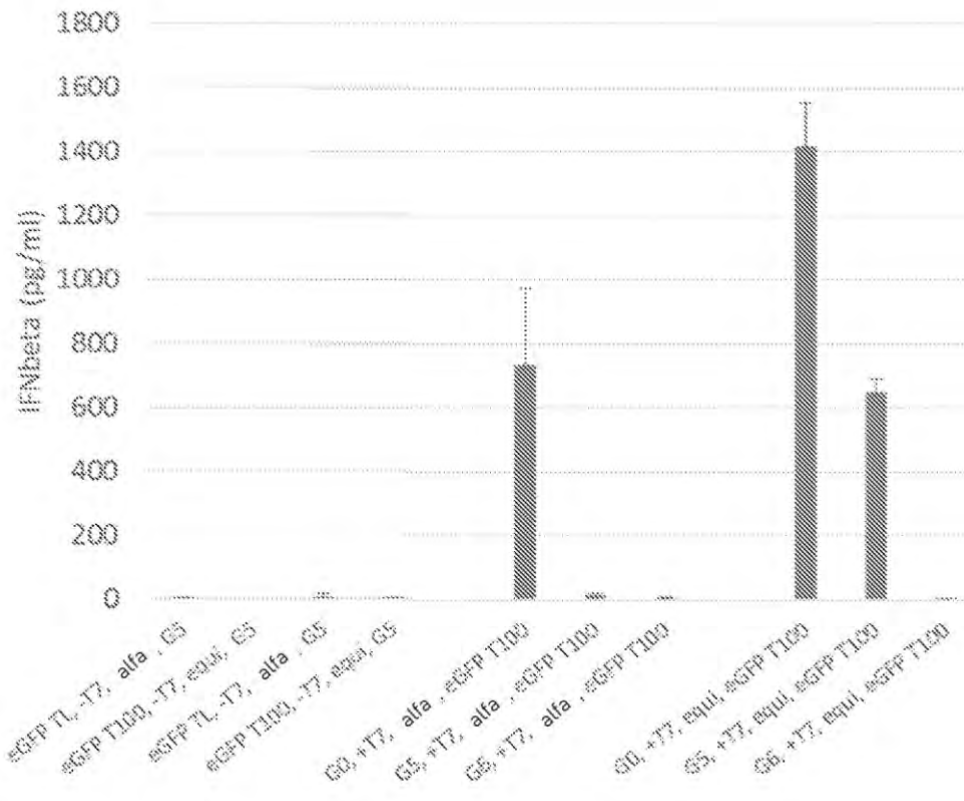
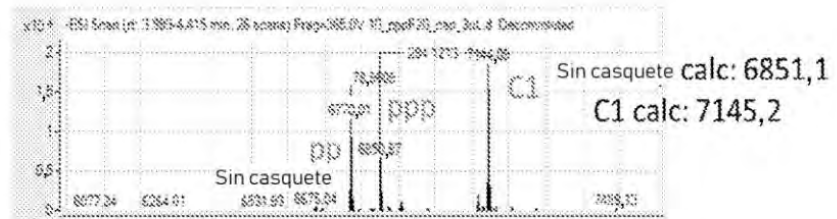


Figura 18.

pppF20



pppRC20



pppF20+pppRC20



Figura 19.

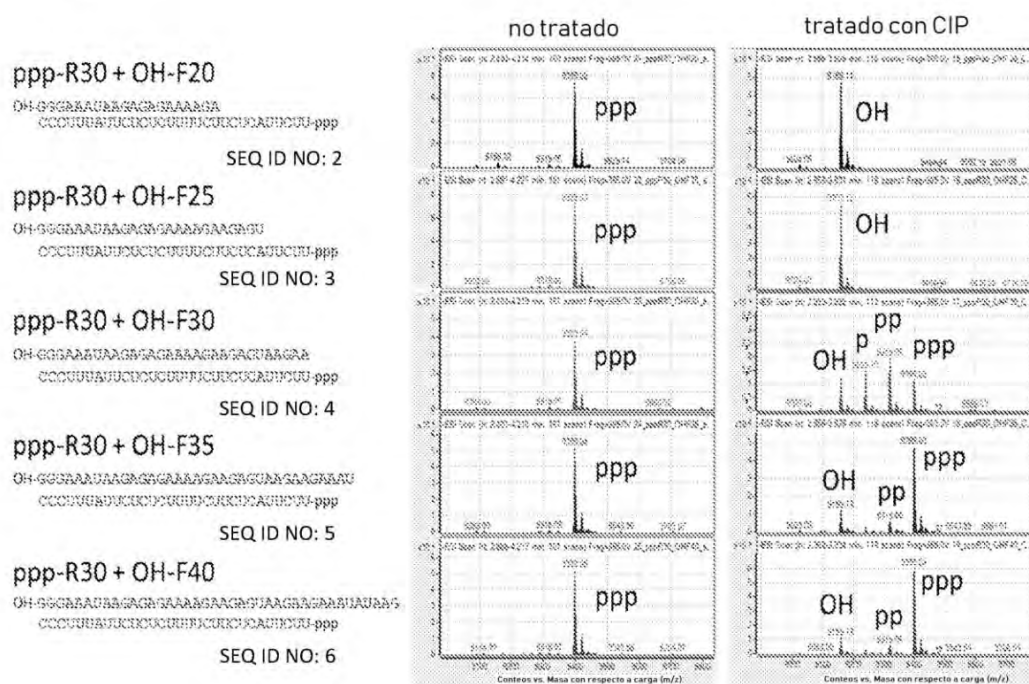


Figura 20.

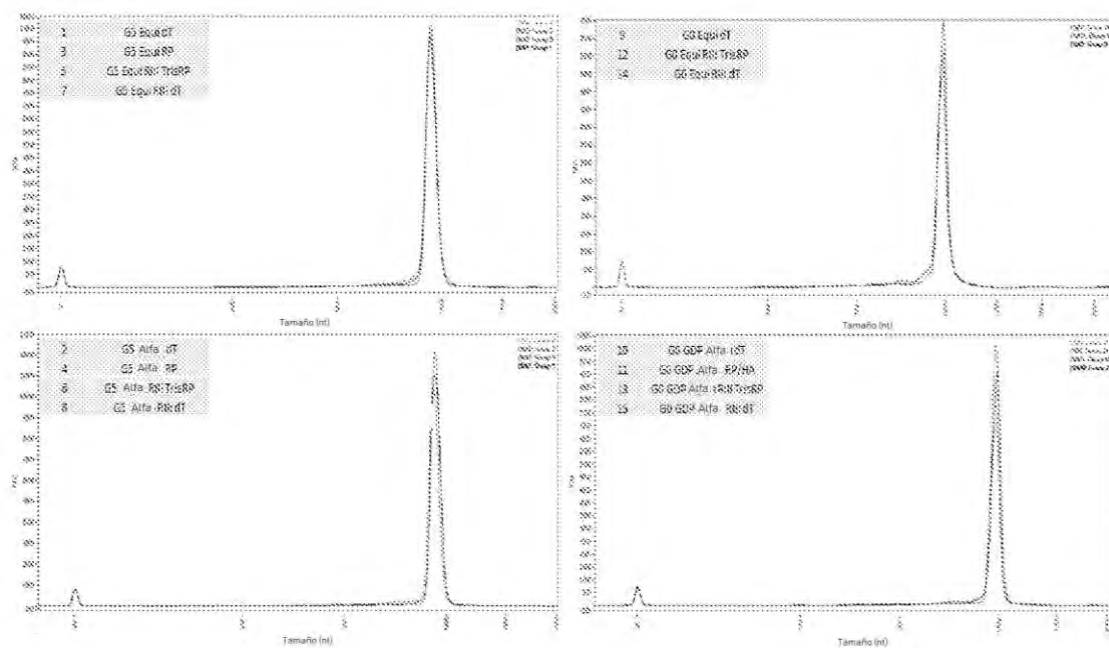




Figura 21.

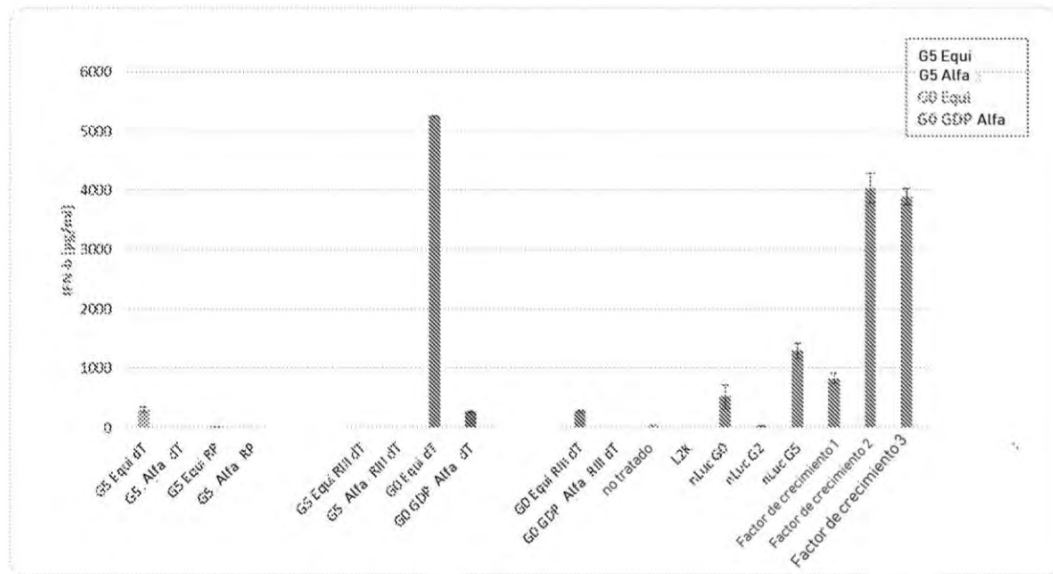


Figura 22.

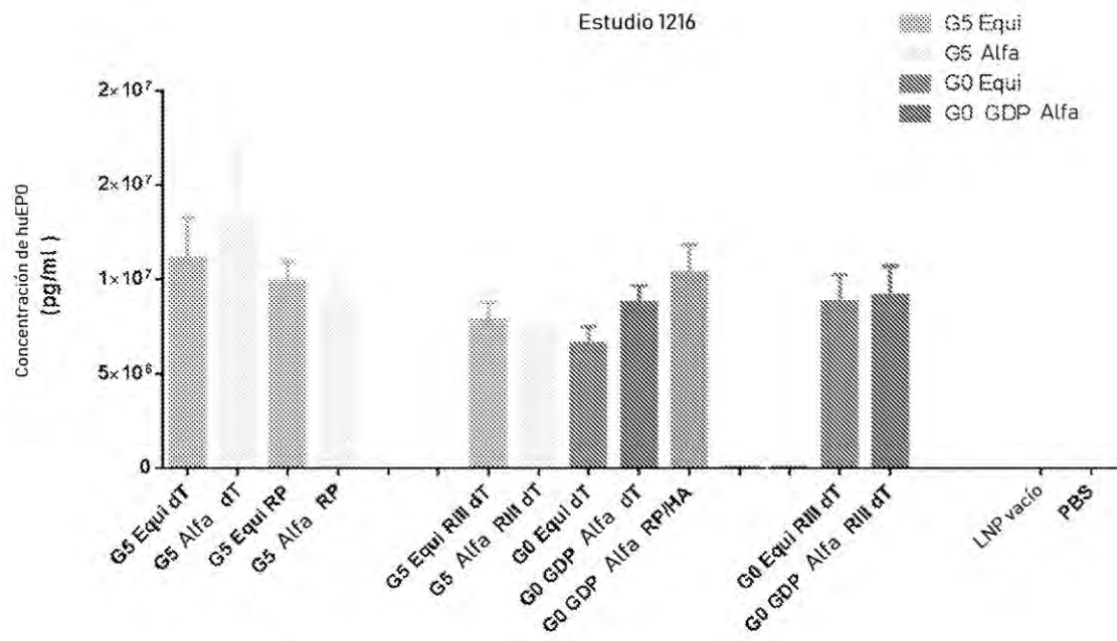


Figura 23A.

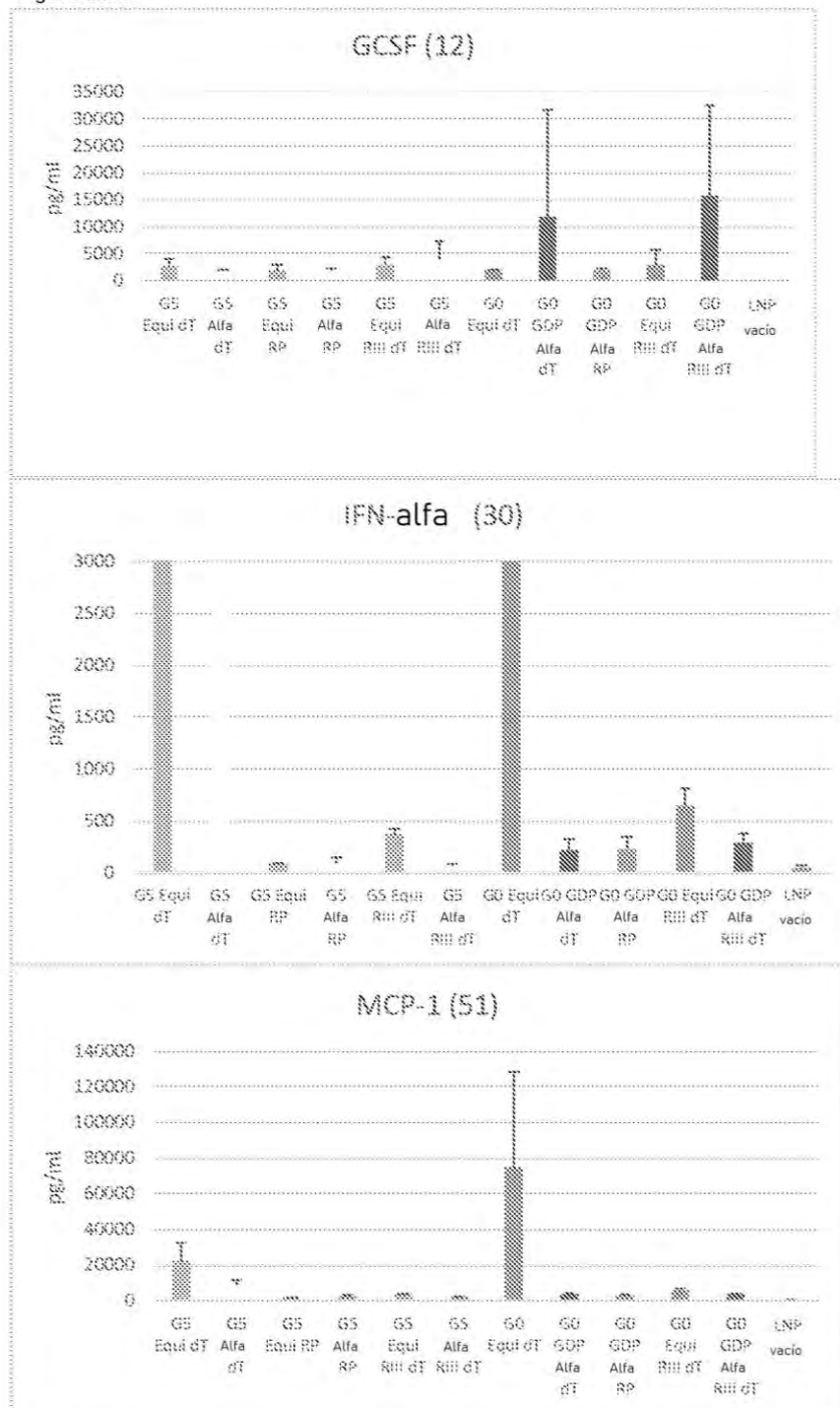


Figura 23B.

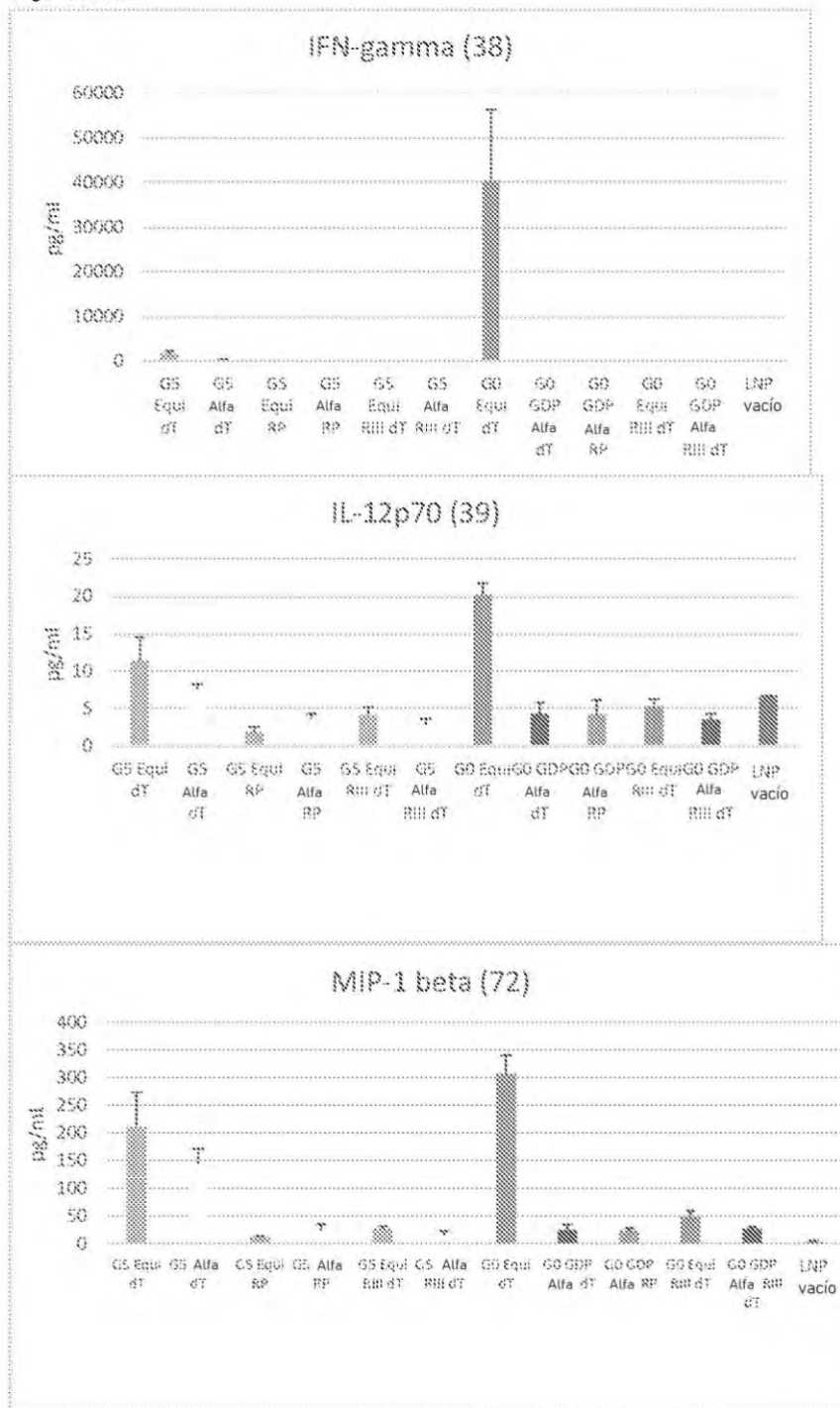


Figura 23C.

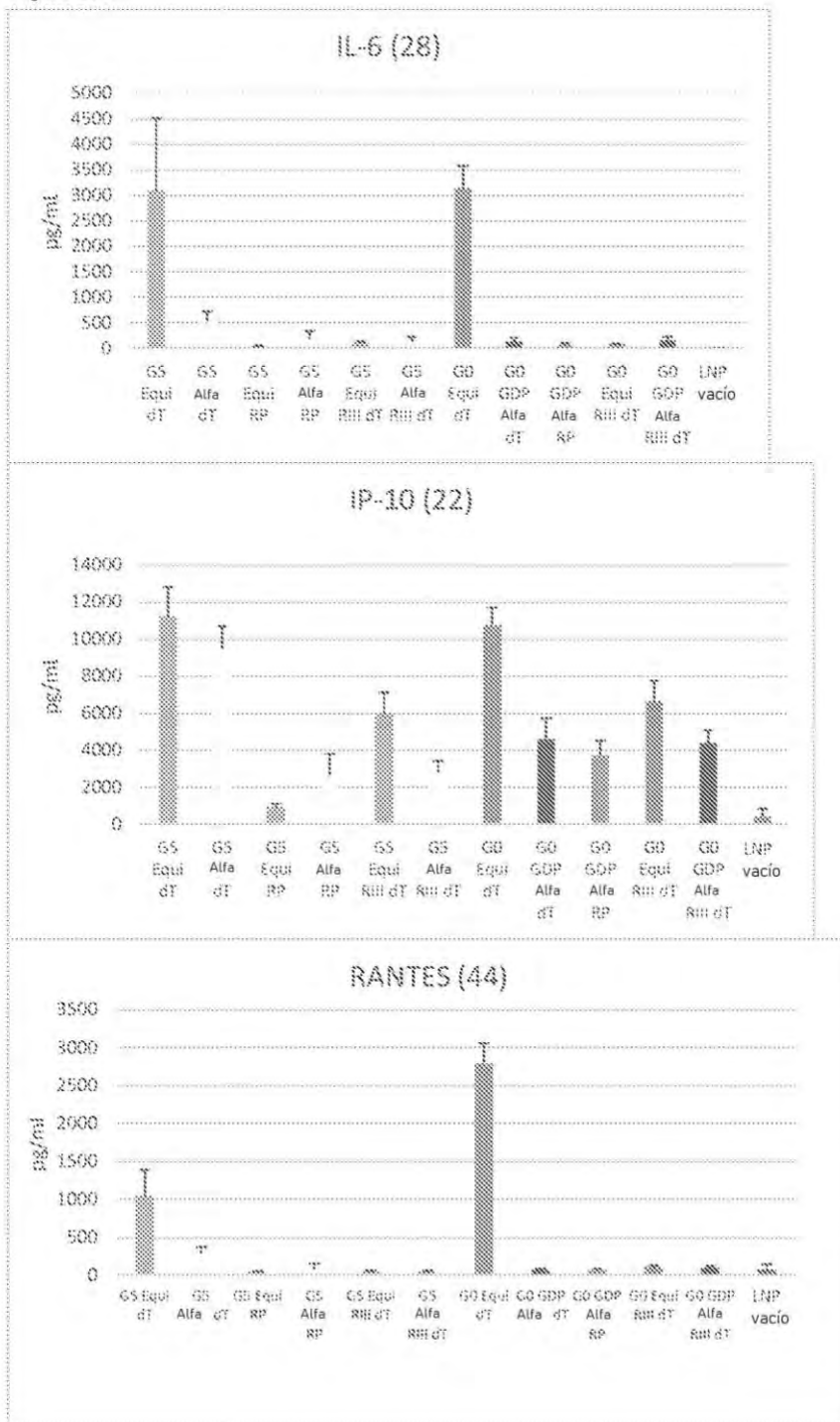


Figura 23D.

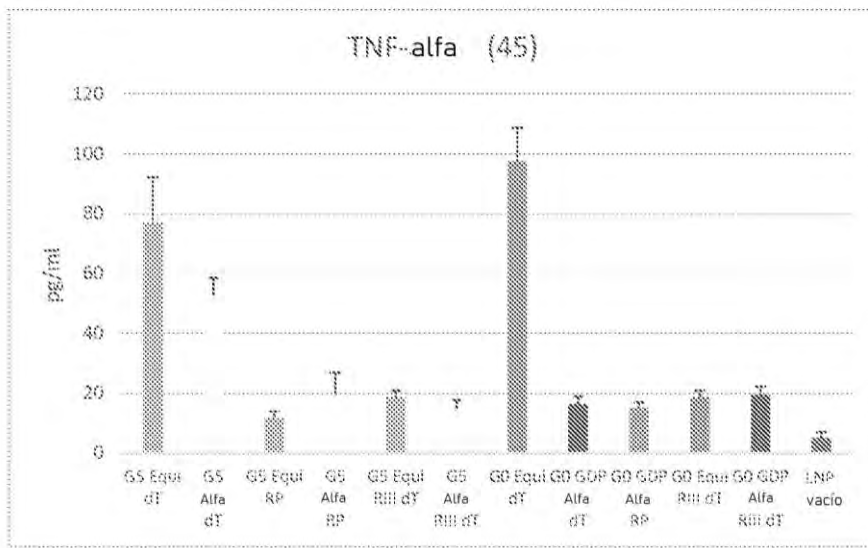


Figura 24.

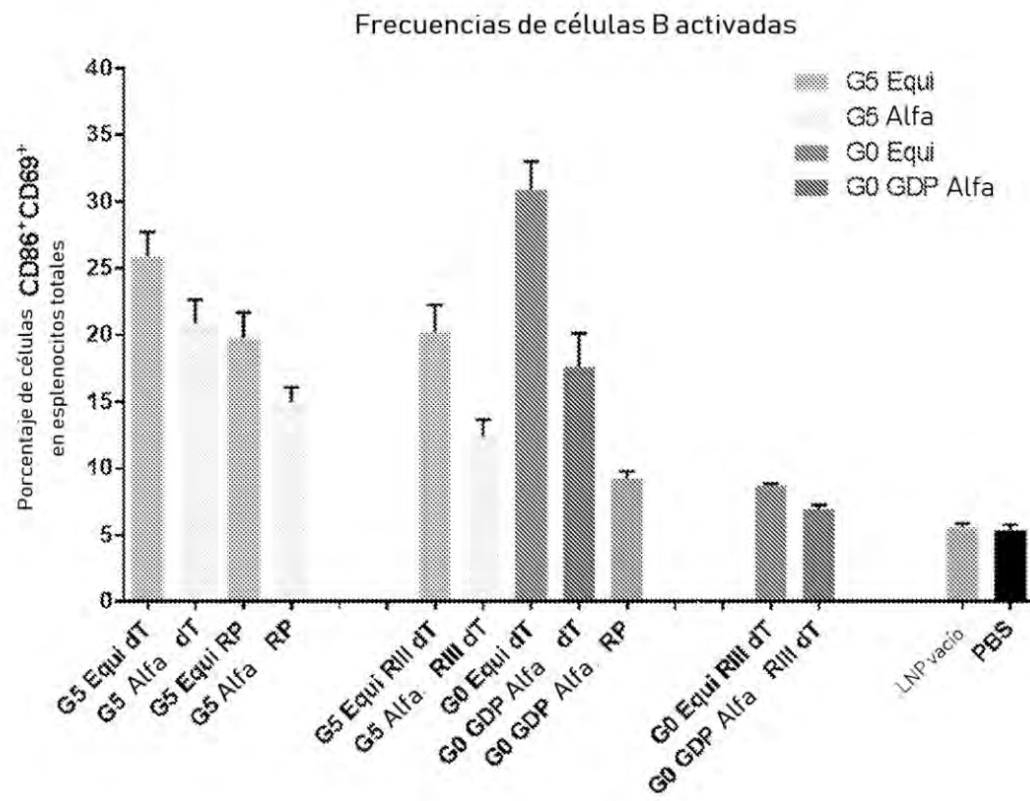


Figura 25.

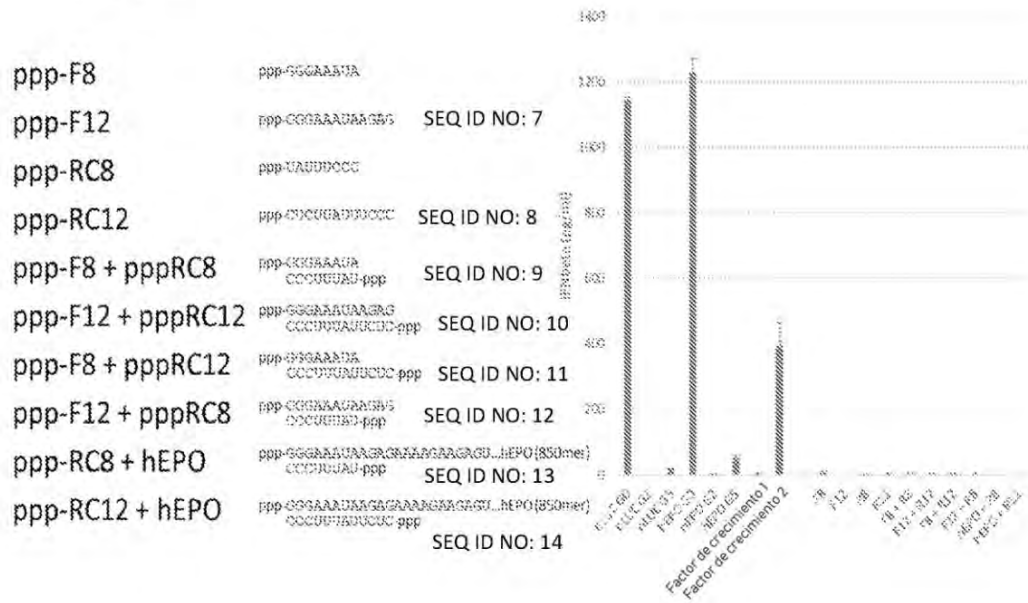




Figura 26.

ppp-F20 + ppp-RC20 (Tm 59.9C)	ppp-GCGTAAATGAGCAATGAAAGG COCUUUAAUUCUUAUUCUUUUU-ppp	SEQ ID NO: 15
ppp-U16 + OH-A25 (Tm 22.6C)	OH-AAAAAAAAAAAAAAAAAAAAAAAA UCUUUUUUUUUUUUUUU-ppp	SEQ ID NO: 16
ppp-U20 + OH-A25 (Tm 31.7C)	OH-AAAAAAAAAAAAAAAAAAAAAAAA UCUUUUUUUUUUUUUUU-ppp	SEQ ID NO: 17
ppp-U25 + OH-A25 (Tm 39C)	OH-AAAAAAAAAAAAAAAAAAAAAAAA UUUUUUUUUUUUUUUUUUUU-ppp	SEQ ID NO: 18
ppp-RC20 + C1-ARN sustituto 2	ppp-SCGTAATGAGCAATGAAAGG... (250mer) COCUUUAAUUCUUAUUCUUUUU-ppp	SEQ ID NO: 19

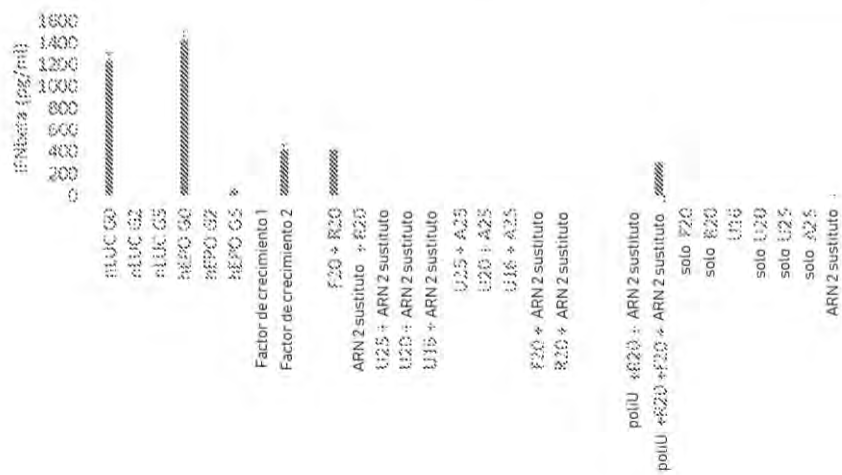


Figura 27.

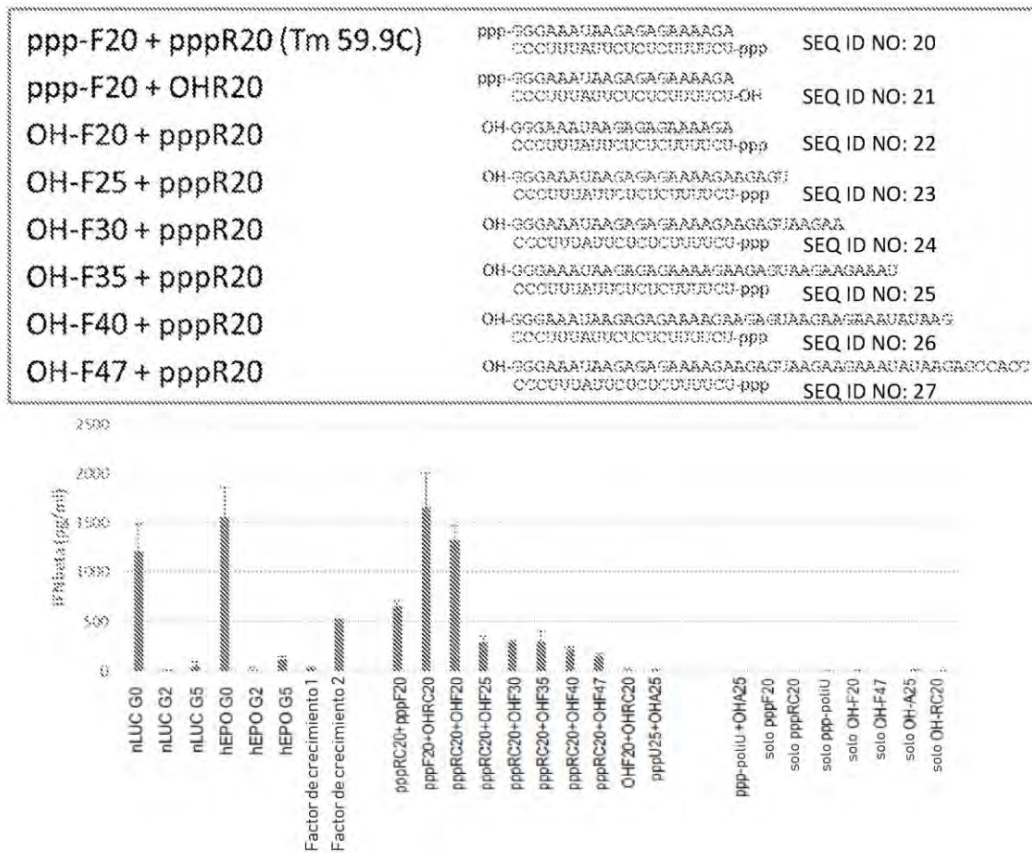


Figura 28.

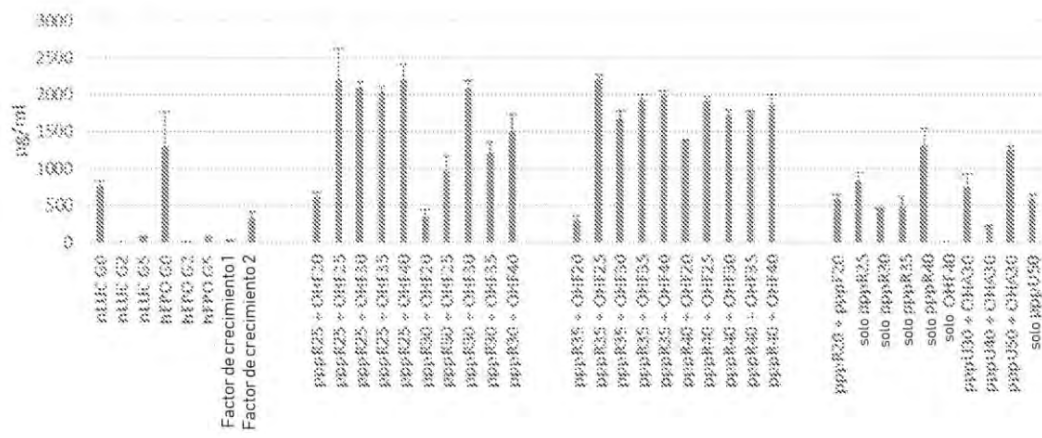


Figura 29.

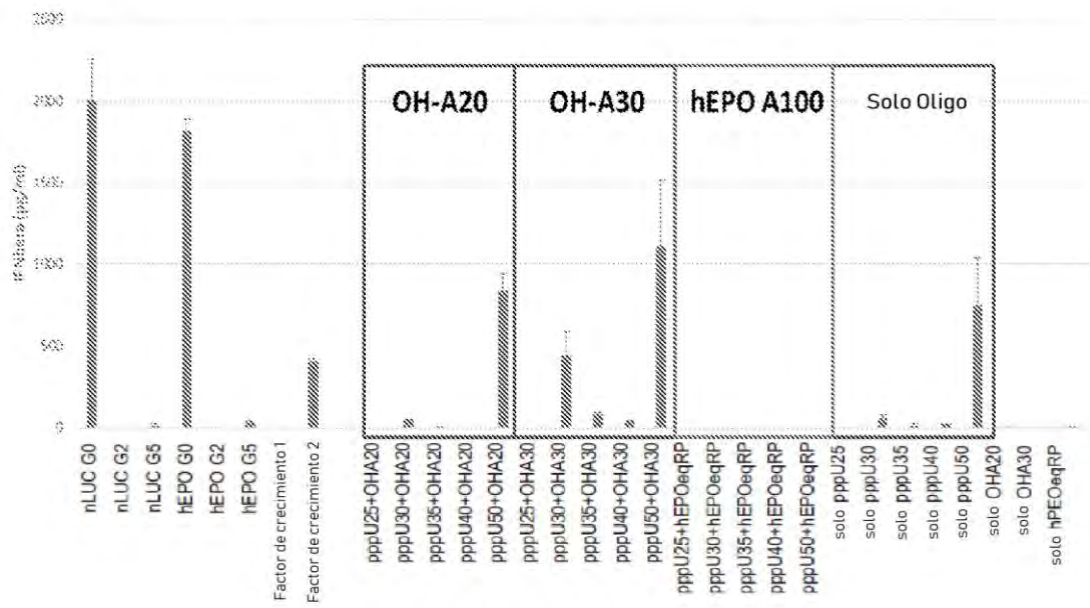


Figura 30.

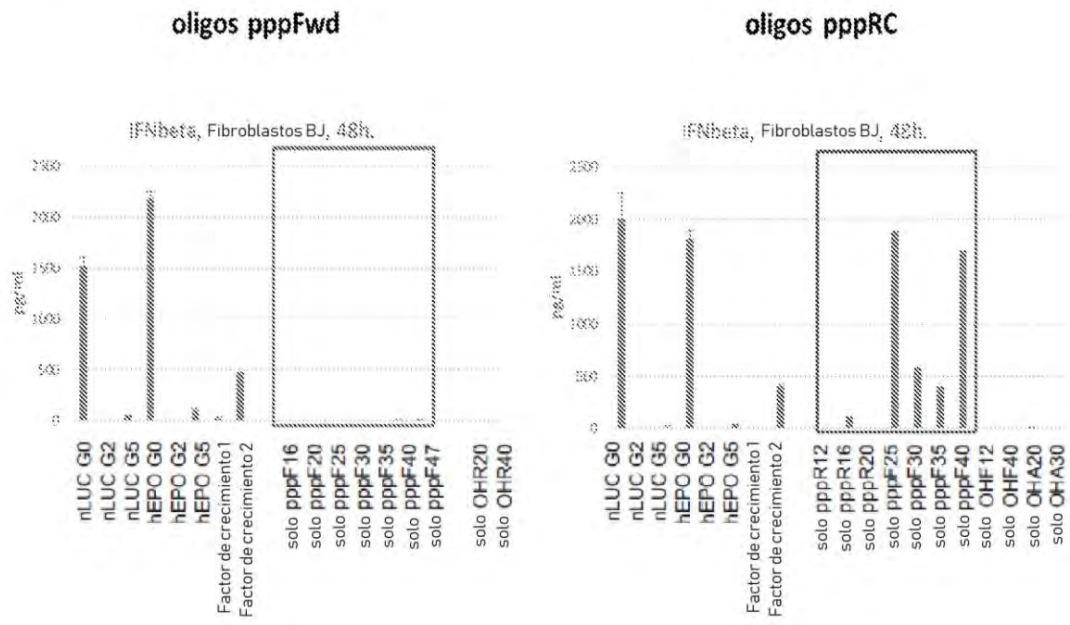


Figura 31.

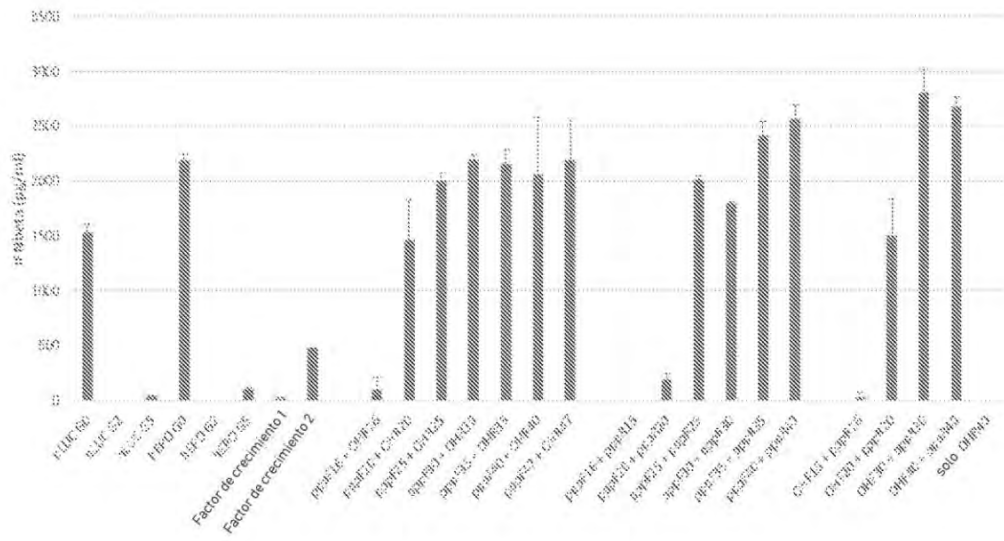




Figura 33.

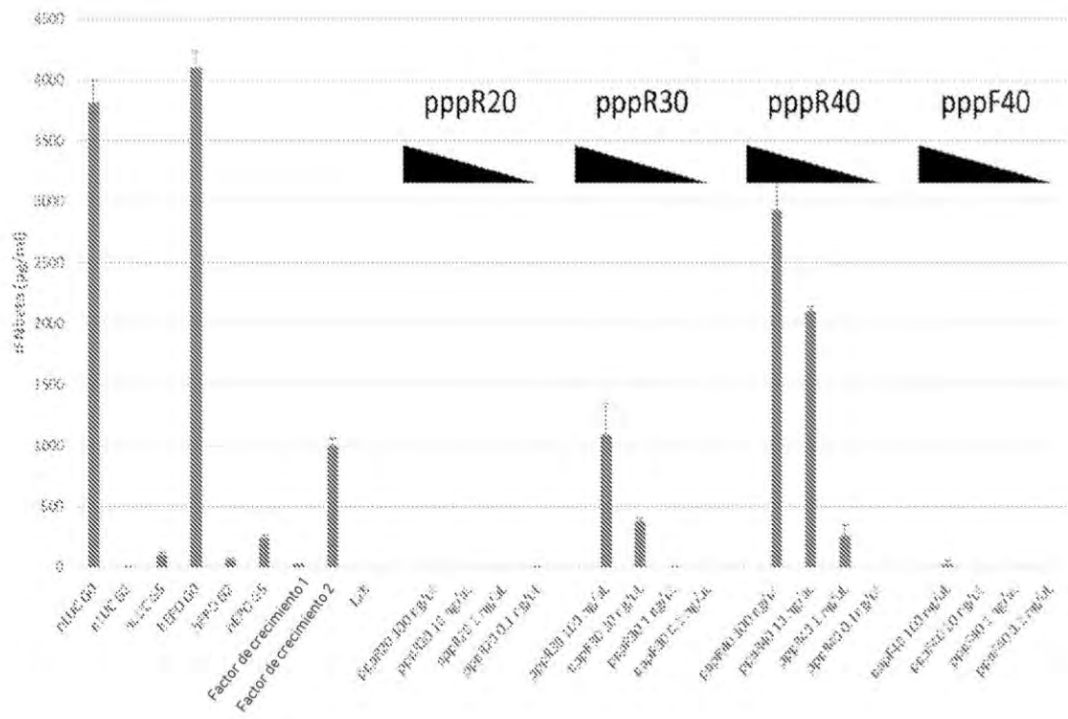




Figura 34.

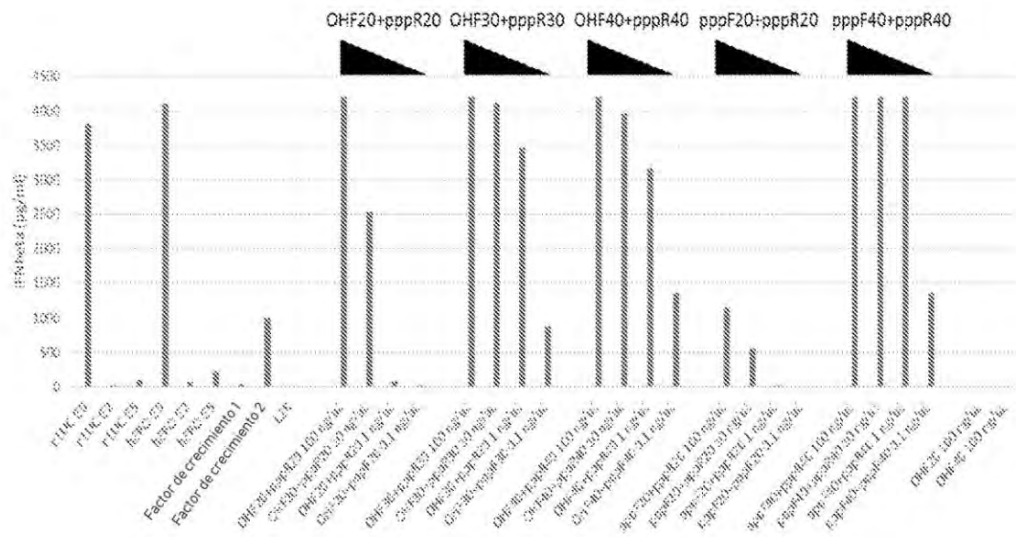
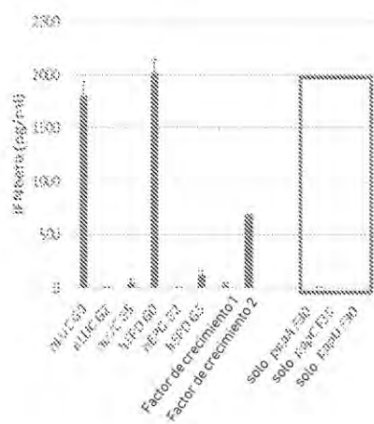
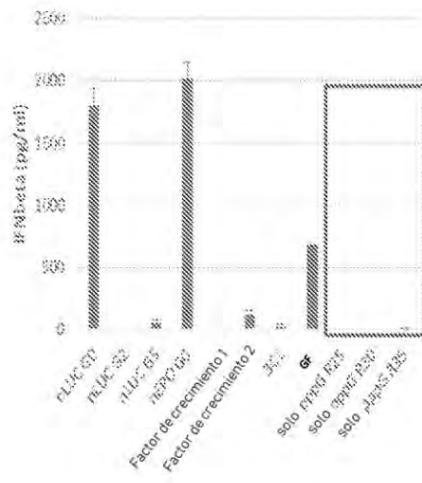


Figura 35.



Oligo	Secuencia (5' → 3')	
AF30	AGGGAAAUAAGAGAGAAAAGAAGAGUAAGAA	SEQ ID NO: 28
CF30	CGGGAAAUAAGAGAGAAAAGAAGAGUAAGAA	SEQ ID NO: 29
UF30	UGGGAAAUAAGAGAGAAAAGAAGAGUAAGAA	SEQ ID NO: 30

Figura 36.



Oligo	Secuencia (5' → 3')
GR25	GACUCUUCUUUUUCUCUCUUAUUUCCC
GR30	GUUCUUACUCUUCUUUUUCUCUCUUAUUUCCC
GR35	GAUUUCUUCUUACUCUUCUUUUUCUCUCUUAUUUCCC

Figura 37

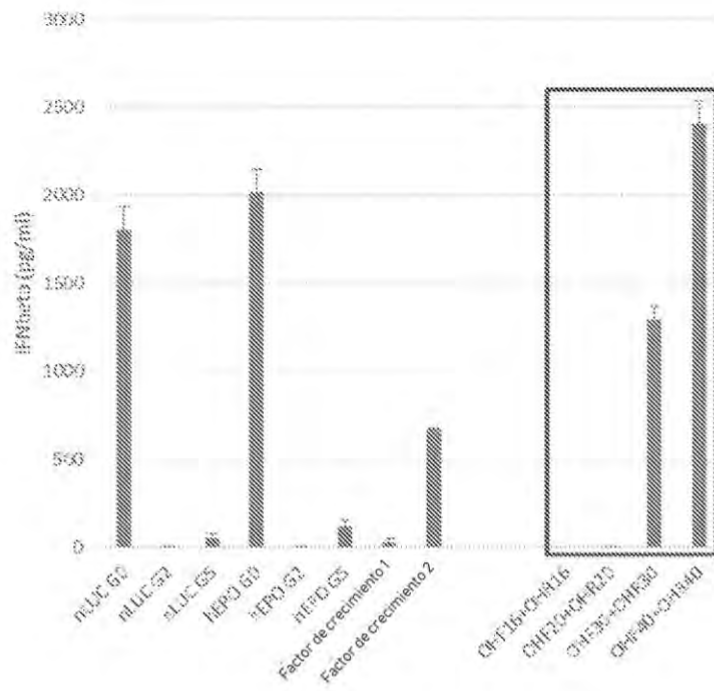


Figura 38.

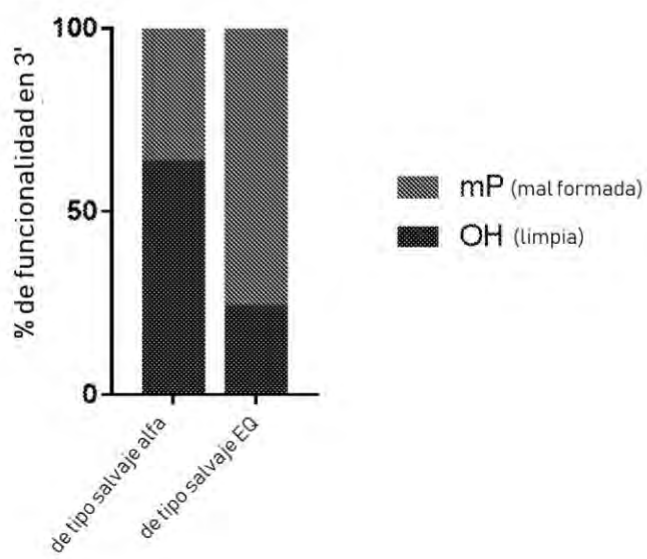


Figura 39

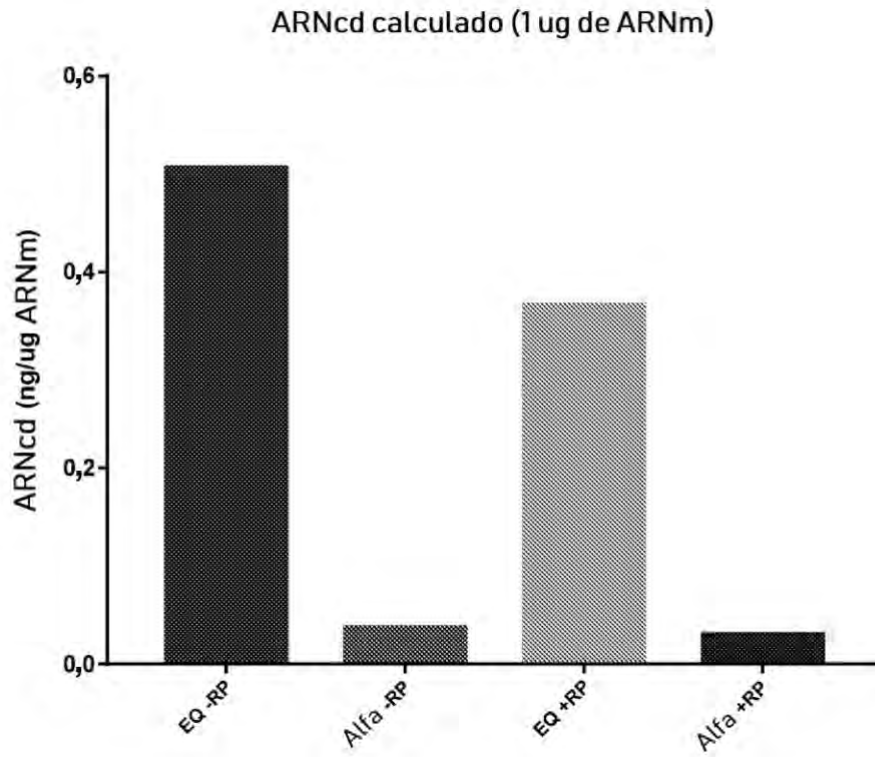


Figura 40.

