



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 356 934**

51 Int. Cl.:

C07K 14/11 (2006.01)

C07K 14/47 (2006.01)

A61K 48/00 (2006.01)

A61K 39/00 (2006.01)

A61K 39/145 (2006.01)

A61P 31/16 (2006.01)

A61P 35/04 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **07015835 .7**

96 Fecha de presentación : **05.06.2002**

97 Número de publicación de la solicitud: **1857122**

97 Fecha de publicación de la solicitud: **21.11.2007**

54

Título: **ARNm estabilizado con un contenido de G/C aumentado que codifica para un antígeno viral.**

30

Prioridad: **05.06.2001 DE 101 27 283**

45

Fecha de publicación de la mención BOPI:
14.04.2011

45

Fecha de la publicación del folleto de la patente:
14.04.2011

73

Titular/es: **CUREVAC GmbH**
Paul-Ehrlich-Strasse 15
72076 Tübingen, DE

72

Inventor/es: **Von der Mülbe, Florian;**
Pascolo, Steve y
Hoerr, Ingmar

74

Agente: **Aznárez Urbieto, Pablo**

ES 2 356 934 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

ARNm estabilizado con un contenido de g/c aumentado que codifica para un antígeno viral.

5 La presente invención se refiere a una composición farmacéutica que contiene un ARNm estabilizado por modificaciones de secuencia en el campo traducido y optimizado para la traducción. La composición farmacéutica según la invención es especialmente adecuada como vacuna contra enfermedades infecciosas virales. Además, se describe un procedimiento para la determinación de modificaciones de secuencia que sirven para la estabilización y la optimización de la traducción de ARNm.

10 La terapia genética y la vacunación genética son procedimientos médicos moleculares cuya aplicación en la terapia y la prevención de enfermedades tendrá considerables efectos sobre la práctica médica. Ambos procedimientos se basan en la introducción de ácidos nucleicos en células o en tejidos del paciente, así como en el subsiguiente procesamiento de la información codificada por los ácidos nucleicos introducidos, es decir la expresión de los polipéptidos deseados.

15 El método habitual de los procedimientos existentes hasta la fecha en la terapia genética y la vacunación genética es la utilización de ADN para introducir la información genética requerida en la célula. En este contexto se han desarrollado diferentes procedimientos para la introducción de ADN en células, por ejemplo transfección de fosfato cálcico, transfección de polipreno, fusión de protoplastos, electroporación, microinyección y lipofección, donde especialmente la lipofección resultó ser un procedimiento adecuado.

20 Otro procedimiento que se ha propuesto especialmente en procedimientos de vacunación genética es la utilización de virus de ADN como vehículos para el ADN. Tales virus tienen la ventaja de que, debido a sus características infecciosas, se puede conseguir un coeficiente de transfección muy alto. Los virus utilizados se modifican genéticamente de manera que, en la célula transfectada, no se forman partículas infecciosas activas. Sin embargo, a pesar de esta medida de precaución, no se puede excluir un cierto riesgo de propagación incontrolada de los genes tanto de efecto terapéutico genético como virales debido a posibles incidencias en la recombinación.

25 Normalmente, el ADN introducido en la célula se integra en cierta medida en el genoma de la célula transfectada. Por un lado, este fenómeno puede tener un efecto deseado, ya que se puede conseguir así una actividad de larga duración del ADN introducido. Por otro lado, la integración en el genoma trae consigo un considerable riesgo en la terapia genética. Así, por ejemplo, se puede producir una inserción del ADN incorporado en un gen intacto, lo que representa una mutación que obstaculiza la función del gen endógeno o, incluso, la anula por completo. Debido a tales incidencias de integración, se pueden anular, por un lado, sistemas enzimáticos vitales para la célula y, por otro lado, también existe el peligro de una transformación de la célula así modificada en un estado de degeneración si, por la integración del ADN extraño, se modifica un gen decisivo para la regulación del crecimiento celular. Por esta razón, cuando se emplean virus de ADN como agentes terapéuticos genéticos y vacunas no se puede excluir un cierto riesgo de formación de cáncer. En este contexto también se ha de tener en cuenta que para la expresión efectiva de los genes introducidos en la célula, los vehículos de ADN correspondientes, también contienen un fuerte promotor, por ejemplo el promotor de CMV viral. La integración de tales promotores en el genoma de la célula tratada puede conducir a modificaciones no deseadas de la regulación de la expresión genética en la célula.

35 Otra desventaja de la utilización del ADN como agente terapéutico genético y vacuna es la inducción de anticuerpos anti ADN patógenos en el paciente, provocando una reacción inmunológica posiblemente letal.

40 Al contrario que con el ADN, la utilización de ARN como agente terapéutico genético o vacuna se puede considerar considerablemente más segura. Especialmente, el ARN no presenta el peligro de integrarse de forma estable en el genoma de la célula transfectada. Además, no son necesarias secuencias virales como promotores para una transcripción efectiva. Adicionalmente, el ARN se desintegra *in vivo* de forma considerablemente más sencilla. Quizás debido al período de semidesintegración relativamente corto del ARN frente al ADN en el sistema sanguíneo, no se han detectado hasta la fecha anticuerpos de anti ARN. Por esta razón, el ARN puede considerarse la mejor opción de molécula para procedimientos de terapia médica molecular.

45 No obstante, los procedimientos médicos que se basan en sistemas de expresión de ARN todavía requieren, antes de una aplicación más amplia, la solución de algunos problemas básicos. Uno de los problemas de utilizar ARN es la transferencia segura, eficiente específicamente en cuanto a la célula o el tejido, del ácido nucleico. Debido a que el ARN normalmente es muy inestable en forma de solución, por los procedimientos tradicionales que se utilizan para el ADN, hasta la fecha el ARN no ha podido utilizarse o se ha utilizado con escasa eficacia como agente terapéutico o vacuna.

50 En cuanto a la inestabilidad son responsables enzimas de desintegración de ARN, las denominadas RNasas (ribonucleasas). Incluso las impurezas más pequeñas de ribonucleasas bastan para desintegrar por completo el ARN en solución. La desintegración natural del ARNm en el citoplasma celular está sujeta a una regulación muy fina. En este contexto, se conocen diversos mecanismos. Así, para un ARNm funcional tiene una importancia decisiva la estructura terminal. En el extremo 5' se encuentra la llamada estructura "cap" (caperuza) (un nucleótido de guanosina modificado) y en el extremo 3' una secuencia de hasta 100 nucleótidos de adenosina (la llamada cola poli A). A través de estas estructuras se reconoce el ARN como ARNm y se regula la desintegración. Además, existen otros procesos que

estabilizan o desestabilizan el ARN. Muchos de estos procesos todavía no son conocidos, sin embargo, con frecuencia parece decisiva para ello una interacción entre el ARN y las proteínas. Por ejemplo, se describió recientemente un "mRNA-Surveillance-System" (sistema de control de ARNm) (Hellerin y Parker, *Amun. Rev. Genet.* 1999, 33: 229 a 260) en el que se conocen, debido a determinadas interacciones de proteínas de "feedback" (retroalimentación) en el citosol, ARNm incompletos o "nonsense" (sin sentido) a los que se proporciona acceso para la desintegración, donde una parte principal de este proceso la llevan a cabo exonucleasas.

En la técnica actual se han propuesto algunas medidas para aumentar la estabilidad del ARN y, por ello, hacer posible su utilización como sustancia terapéutica genética o vacuna de ARN.

En la EP-A-1083232 se propone para la solución del problema arriba enunciado de la inestabilidad del ARN *ex vivo* un procedimiento para la introducción de ARN, especialmente ARNm, en células y organismos en los que el ARN existe en forma de un complejo con una proteína o un péptido catiónico.

La WO 99/14346 describe otros procedimientos para la estabilización de ARNm. Especialmente se proponen modificaciones de ARNm que estabilizan las especies de ARNm frente a la desintegración por las RNasas. Tales modificaciones afectan, por un lado, a la estabilización por modificaciones de secuencia, especialmente reducción del contenido de C y/o U mediante la eliminación o sustitución de bases. Por otro lado, se proponen modificaciones químicas, especialmente la utilización de análogos de nucleótidos, así como grupos de bloqueo de 5' y 3', una mayor longitud de la cola de poli-A, así como la formación de complejos de ARNm con medios de estabilización y combinaciones de las medidas indicadas.

En las patentes de Estados Unidos US 5.580.859 y US 6.214.804 se describen, entre otros dentro del marco de la "terapia genética transitoria" (TGT), vacunas y sustancias terapéuticas de ARNm. Se describen diferentes medidas para aumentar la eficacia de la traducción y la estabilidad del ARNm, medidas que se refieren, sobre todo, a regiones de secuencias no traducidas.

Bieler y Wagner (en: Schlee (Eds.), *Plasmids for Therapy and Vaccination*, capítulo 9, páginas 147 a 168, Wiley-VCH, Weinheim, 2001) informan sobre la utilización de genes sintéticos en conexión con métodos terapéuticos genéticos utilizando vacunas de ADN y vectores lentivirales. Se describe la construcción de un gen *gag* sintético derivado de VIH-1 en el que los codones se modificaron frente a la secuencia de tipo salvaje (utilización alternativa de codones, "codon usage") de manera que correspondían a la utilización de codones que se puede encontrar en genes altamente expresados de mamíferos. Así se reduce especialmente el contenido de A/T frente a la secuencia de tipo salvaje. En particular, los autores observaron un mayor coeficiente de expresión del gen *gag* sintético en las células transfectadas. Además, se observó en ratones una mayor formación de anticuerpos contra la proteína *gag* en el caso de ratones inmunizados con el constructo de ADN sintético y también una mayor liberación de citoquinas *in vitro* con células transfectadas del bazo de ratones. Finalmente, pudo observarse la inducción de una reacción inmunológica citotóxica en ratones inmunizados con el plásmido de expresión *gag*. Los autores de este artículo atribuyen las mejores características de esta vacuna de ADN esencialmente a una modificación del transporte núcleo-citoplasmático del ARNm expresado de la vacuna de ADN, modificación provocada por la utilización optimizada de codones. Por el contrario, los autores consideran que el efecto de la utilización modificada de codones en la eficacia de traducción es pequeño.

El objetivo de la presente invención consiste, por tanto, en proporcionar un nuevo sistema para la vacunación genética que suprima las desventajas relacionadas con las características de las vacunas de ADN y aumente la efectividad de las sustancias terapéuticas basadas en especies de ARN.

Este objetivo se alcanza en las formas de realización caracterizadas en las reivindicaciones de la presente invención.

En particular, se propone un ARNm modificado así como una composición farmacéutica que contiene al menos un ARNm modificado de este tipo junto con un excipiente y/o vehículo farmacéuticamente compatible, donde el ARNm modificado codifica como mínimo un péptido o polipéptido viral antigénico, donde la secuencia del ARNm, especialmente en el campo que codifica como mínimo un péptido o polipéptido, muestra, frente al ARNm del tipo salvaje, las siguientes modificaciones, que pueden existir bien alternativamente o bien en combinación.

Por un lado, el contenido de G/C del campo que codifica el péptido o el polipéptido del ARNm modificado es mayor que el contenido de G/C del campo de codificación del ARNm del tipo salvaje que codifica tal péptido o polipéptido, manteniendo la secuencia del aminoácido codificada sin modificación frente al tipo salvaje.

Esta modificación se basa en el hecho de que, para una traducción eficiente de un ARNm, el desarrollo de la secuencia del campo del ARNm a traducir es esencial. Aquí juegan un papel importante la composición y la secuencia de los diferentes nucleótidos. Especialmente, las secuencias con un mayor contenido de G(guanosina)/C(citosina) son más estables que las secuencias con un mayor contenido de A(adenosina)/U(uracilo). Por tanto, se varían los codones frente al ARNm del tipo salvaje según la invención manteniendo la secuencia traducida del aminoácido, de forma que contienen una mayor cantidad de nucleótidos G/C. Debido a que varios codones codifican uno y el mismo aminoácido (degeneración del código genético), es posible determinar los codones más favorables para la estabilidad (utilización alternativa de codones, "codon usage").

5 Dependiendo del aminoácido a codificar por el ARNm modificado son factibles diferentes posibilidades para la modificación de la secuencia de ARNm frente a la secuencia de tipo salvaje. En el caso de aminoácidos, codificados por codones que contienen exclusivamente nucleótidos G ó C, no es necesaria ninguna modificación del codón. Así, los codones para Pro (CCC ó CCG), Arg (CGC ó CGG), Ala (GCC ó GCG) y Gly (GGC ó GGG) no requieren ninguna modificación, ya que no existen ni A ni U.

En los siguientes casos se modifican los codones que contienen nucleótidos de A y/o U mediante la sustitución por otros codones que codifican los mismos aminoácidos pero no contienen ningún A y/o U. Ejemplos son:

- los codones para Pro pueden modificarse de CCU ó CCA a CCC ó CCG;
- los codones para Arg pueden modificarse de CGU ó CGA ó AGA ó AGG a CGC ó CGG;
- 10 - los codones para Ala pueden modificarse de GCU ó GCA a GCC ó GCG;
- los codones para Gly pueden modificarse de GGU ó GGA a GGC ó GGG.

Aunque, en otros casos, los nucleótidos A o U no pueden eliminarse de los codones, sin embargo es posible reducir el contenido de A y U mediante la utilización de codones que contienen menos nucleótidos A y/o U. Por ejemplo:

- Los codones para Phe pueden modificarse de UUU a UUC;
- 15 - los codones para Leu pueden modificarse de UUA, CUU ó CUA a CUC ó CUG;
- los codones para Ser pueden modificarse de UCU ó UCA ó AGU a UCC, UCG ó AGC;
- el codón para Tyr puede modificarse de UAU a UAC;
- el codón de parada UAA puede modificarse en UAG ó UGA;
- el codón para Cys puede modificarse de UGU a UGC;
- 20 - el codón His puede modificarse de CAU a CAC;
- el codón para Gln puede modificarse de CAA a CAG;
- los codones para Ile pueden modificarse de AUU ó AUA a AUC;
- los codones para Thr pueden modificarse de ACU ó ACA a ACC ó ACG;
- el codón para Asn puede modificarse de AAU a AAC;
- 25 - el codón para Lys puede modificarse de AAA a AAG;
- los codones para Val pueden modificarse de GUU ó GUA a GUC ó GUG;
- el codón para Asp puede modificarse de GAU a GAC;
- el codón para Glu puede modificarse de GAA a GAG.

30 En el caso de los codones para Met (AUG) y Trp (UGG), por el contrario, no existe ninguna posibilidad de modificación de la secuencia.

35 Las sustituciones arriba indicadas pueden utilizarse naturalmente por separado, pero también en todas las combinaciones posibles para aumentar el contenido de G/C del ARNm modificado frente a la secuencia original. Así, por ejemplo, se pueden modificar todos los codones para Thr que se presentan en la secuencia (del tipo salvaje) original a ACC (o ACG). De preferencia, sin embargo, se utilizan combinaciones de las posibilidades anteriores de sustitución, por ejemplo:

Sustitución de todos los codones que codifican Thr en la secuencia original por ACC (ó ACG) y sustitución de todos los codones que codifican originalmente Ser por UCC (ó UCG ó AGC);

40 sustitución de todos los codones que codifican Ile en la secuencia original por AUC y sustitución de todos los codones que codifican originalmente Lys por AAG y sustitución de todos los codones que codifican originalmente Tyr por UAC;

sustitución de todos los codones que codifican Val en la secuencia original por GUC (ó GUG) y sustitución de todos los codones que codifican originalmente Glu por GAG y sustitución de todos los codones que codifican originalmente Ala por GCC (ó GCG) y sustitución de todos los codones que codifican originalmente Arg por CGC (ó CGG);

sustitución de todos los codones que codifican Val en la secuencia original por GUC (ó GUG) y sustitución de todos los codones que codifican originalmente Glu por GAG y sustitución de todos los codones que codifican originalmente Ala por GCC (ó GCG) y sustitución de todos los codones que codifican originalmente Gly por GGC (ó GGG) y sustitución de todos los codones que codifican originalmente Asn por AAC;

5 sustitución de todos los codones que codifican Val en la secuencia original por GUC (ó GUG) y sustitución de todos los codones que codifican originalmente Phe por UUC y sustitución de todos los codones que codifican originalmente Cys por UCG y sustitución de todos los codones que codifican originalmente Leu por CUG (ó CUC) y sustitución de todos los codones que codifican originalmente Gln por CAG y sustitución de todos los codones que codifican originalmente Pro por CCC (ó CCG);

10 etc.

Preferentemente el contenido de G/C del campo del ARNm modificado que codifica el péptido o el polipéptido se incrementa en como mínimo un 7 %, con especial preferencia en como mínimo un 15%, en particular en como mínimo un 20% el contenido de G/C del campo codificado del ARNm del tipo salvaje que codifica el polipéptido.

15 En este contexto es especialmente preferente aumentar al máximo el contenido de G/C del ARNm modificado, especialmente en el campo que codifica al menos un péptido o un polipéptido, en comparación con la secuencia del tipo salvaje.

20 La otra modificación según invención del ARNm contenido en la composición farmacéutica caracterizada en la presente invención se basa en el conocimiento de que la eficacia de traducción también queda determinada por una frecuencia diferente en la presencia de los ARNt en las células. Por tanto, cuando en una secuencia de ARN existen en mayor cantidad los llamados codones "raros", el ARNm correspondiente se traduce claramente peor que en caso de los ARNt relativamente "frecuente" para los que existen codones de codificación.

25 Así, según la invención, en el ARNm modificado (incluido en la composición farmacéutica), el campo que codifica el péptido o el polipéptido, se modifica frente al correspondiente campo del ARNm de tipo salvaje de tal forma que al menos un codón de la secuencia del tipo salvaje que codifica un ARNt relativamente raro en la célula se cambia por un codón que codifica un ARNt relativamente frecuente en la célula que lleva el mismo aminoácido que el ARNt relativamente raro.

Gracias a esta modificación se modifican las secuencias de ARN de manera que se intercalan codones para los cuales están disponibles ARNt's frecuentes.

30 Es conocido cuáles son los ARNt que se presentan con relativa frecuencia en la célula y cuáles son aquellos que, frente a éstos, son relativamente raros por el técnico en la materia; véase por ejemplo Akashi, Curr. Opin. Genet. Dev. 2001, 11(6): 660-666.

Mediante esta modificación se pueden cambiar según la invención todos los codones de la secuencia de tipo salvaje que codifican un ARNt relativamente raro en la célula, por, en cada caso, un codón que codifica un ARNt relativamente frecuente en la célula que lleva, en cada caso, el mismo aminoácido que el ARNt relativamente raro.

35 Según la invención tiene especial preferencia vincular el contenido de G/C secuencial, especialmente al máximo, incrementado según la invención en el ARNm modificado con los codones "frecuentes" sin modificar la secuencia del aminoácido del péptido o el polipéptido (uno o varios) codificado por el campo de codificación del ARNm. Esta forma de realización preferente proporciona un ARNm traducido y estabilizado de manera especialmente eficiente, por ejemplo para la composición farmacéutica según la invención.

40 En las secuencia de ARNm eucariótico existen elementos de secuencia desestabilizadores (DSE) en los que se enlazan proteínas de señal que regulan la desintegración enzimática del ARNm *in vivo*. Por esta razón, para una mayor estabilización del ARNm modificado incluido en la composición farmacéutica según la invención, se realizan, en caso dado, para el campo de codificación de al menos un péptido o un polipéptido, una o varias modificaciones frente al correspondiente campo del ARNm del tipo salvaje de manera que se incluya ningún elemento de secuencia desestabilizadora. Naturalmente, según la invención también es preferente eliminar del ARNm los DSE eventualmente presentes en los campos no traducidos (UTR de 3' y/ó 5').

45 Tales secuencias desestabilizadoras son, por ejemplo, secuencias ricas en AU ("AURES") presentes en segmentos UTR 3' de múltiples ARNm inestables (Caput y col., Proc. Natl. Acad. Sci. EE.UU. 1986, 83: 1670 a 1674). Las moléculas de ARN incluidas en la composición farmacéutica según la invención, se modifican, por tanto y preferentemente, de tal forma frente al ARNm del tipo salvaje que no contienen ninguna de tales secuencias desestabilizadoras. Esto es aplicable también para tales motivos de secuencia que son reconocidos por posibles endonucleasas, por ejemplo la secuencia GAACAAG incluida en el segmento 3' UTR del gen que codifica el receptor de transferina (Binder y col., EMBO J. 1994, 13: 1969 a 1980). Preferentemente también estos motivos de secuencia se eliminan en el ARNm modificado de la composición farmacéutica según la invención.

El técnico en la materia conoce diferentes procedimientos adecuados para la sustitución de codones en el ARNm modificado según la invención. En el caso de campos de codificación más cortos (que codifican péptidos de efecto biológico o antígenos) se puede sintetizar, por ejemplo, todo el ARNm de forma química utilizando técnicas estándar.

5 Preferentemente, sin embargo se introducen sustituciones de bases utilizando una matriz de ADN para la producción del ARNm modificado con ayuda de técnicas de mutagénesis precisa usuales; Maniatis y col., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, 3ª edición, Cold Spring Harbor, NY, 2001.

10 Por tanto, en este procedimiento se transcribe *in vitro* una correspondiente molécula de ADN para la producción del ARNm. Esta matriz de ADN tiene un promotor adecuado, por ejemplo un promotor de T7 ó SP6, para la transcripción *in vitro*, al que siguen la secuencia de nucleótidos deseada para el ARNm a producir y una señal de parada para la transcripción *in vitro*. Según la invención, se produce la molécula del ADN que constituye la matriz del constructo del ARN a producir mediante crecimiento fermentativo y subsiguiente aislamiento como parte de un plásmido replicable en bacterias. Como plásmidos adecuados para la presente invención se pueden nombrar, por ejemplo, los plásmidos pT7Ts (número de acceso al banco de genes U26404; Lai y col., Development 1995, 121: 2349 a 2360), serie pGEM®,
15 por ejemplo pGEM®-1 (número de acceso al banco de genes X65300; de Promega) y pSP64 (número de acceso al banco de genes X65327); véase también Mezei y Storts, Purification of PCR Products, en: Griffin und Griffin (Eds.), PCR Technology: Current Innovation, CRC Press, Boca Raton, FL, 2001.

20 Así, es posible clonar utilizando oligonucleótidos de ADN sintéticos cortos, que tienen en las intersecciones que se producen transiciones cortas de un solo ramal, o genes producidos por síntesis química según métodos moleculares-biológicos, conocidos por el técnico en la materia, la secuencia deseada de nucleótidos en un plásmido adecuado (véase Maniatis y col., s.o.). Después se recorta la molécula de ADN del plásmido, en el que puede estar presente en copia simple o múltiple, mediante digestión con endonucleasas de restricción.

25 El ARNm modificado incluido en la composición farmacéutica según la invención puede tener, además, una estructura "Cap" en 5' (un nucleótido modificado de guanosina). Como ejemplo de estructuras "Cap" pueden nombrarse m7G(5')ppp(5'(A,G(5')ppp(5')A y G(5')ppp(5')G.

Según otra forma de realización preferente de la presente invención, el ARNm modificado contiene una cola poliA de como mínimo 50 nucleótidos, preferentemente de como mínimo 70 nucleótidos, con especial preferencia de como mínimo 100 nucleótidos, en particular de como mínimo 200 nucleótidos.

30 Para una traducción eficiente del ARNm es necesario, además, un enlace efectivo de los ribosomas en la posición de enlace ribosómica (secuencia Kozak: GCCGCCACCAUGG, con AUG formado por el codón de inicio). En este sentido se ha detectado que un mayor contenido en A/U alrededor de esta posición hace posible un enlace más eficiente de los ribosomas con el ARNm.

35 Además, es posible incluir en el ARNm modificado uno o más de los llamados "IRES" (ingl. "internal ribosomal entry side" - sitio de entrada ribosomal interno). Un IRES puede actuar así como única posición de enlace para ribosomas, sin embargo, también puede servir para proporcionar un ARNm que codifica varios péptidos o polipéptidos a traducir por separado por los ribosomas ("ARN multicistrónico"). Ejemplos de secuencias IRES que se pueden utilizar según la invención son aquellos de picornavirus (por ejemplo FMDV), virus de la peste (CFFV), virus de la poliomieltis (PV), virus de la encefalo-miocarditis (ECMV), virus de la fiebre aftosa (FMDV), virus de la hepatitis C (HCV), virus clásicos de la fiebre porcina (CSFV), virus de leucoma murino (MLV), virus de la inmunodeficiencia del simio (SIV) o
40 virus de la parálisis de cricket (CrPV).

Según otra forma de realización preferente de la presente invención, el ARNm modificado tiene, en las zonas no traducidas 5' y/o 3', secuencias de estabilización que son capaces de aumentar el período de semidesintegración del ARNm en el citosol.

45 Estas secuencias de estabilización pueden tener una homología de secuencia del 100% con las secuencias de procedencia natural que se presentan en virus, bacterias y eucariotas; sin embargo, también pueden ser parcial o completamente de naturaleza sintética. Como ejemplo, para secuencias de estabilización que se pueden utilizar en la presente invención se pueden nombrar las secuencias no traducidas (UTR) del gen de globina β , por ejemplo de *homo sapiens* ó *xenopus laevis*. Otro ejemplo de secuencia de estabilización tiene la fórmula general (C/U)CCAN_xCCC(U/A)Py_xUC(C/U)CC contenida en el UTR 3' del ARNm altamente estable que codifica α -globina, α -(1)-colágeno, 15-lipoxigenasa o tirosina-hidroxilasa (véase Holcik y col., Proc. Natl. Acad. Sci. USA 1997, 94: 2410 a 2414). Naturalmente, tales secuencias de estabilización pueden utilizarse por separado o en combinación entre sí o también en
50 combinación con otras secuencias de estabilización conocidas por el técnico en la materia.

55 Para una mayor estabilización del ARNm modificado se prefiere, además, que éste tenga como mínimo un nucleótido análogo de procedencia natural. Esto se debe a que las enzimas que desintegran el ARN y existen en las células reconocen como sustrato preferente los nucleótidos de procedencia natural. Mediante la introducción de análogos de nucleótidos, por tanto, es posible dificultar la desintegración del ARN, donde la acción sobre la eficacia de

traducción al introducir estos análogos, especialmente en la zona de codificación del ARNm, puede tener un efecto positivo o negativo sobre la eficacia de traducción.

5 Se pueden citar, de modo no limitativo, como ejemplos de análogos de nucleótidos a utilizar según la invención, fosfo-amidatos, fosfo-tioatos, nucleótidos de péptidos, metil-fosfonatos, 7-desazaguanosina, 5-metilcitosina e inosina. La producción de tales análogos es conocida por el técnico en la materia, por ejemplo de las patentes US 4.373.071, 4.401.796, 4.415.732, 4.458.066, 4.500.707, 4.668.777, 4.973.679, 5.047.524, 5.132.418, 5.153.319, 5.262.530 y 5.700.642. Según la invención, tales análogos pueden estar presentes en zonas no traducidas y traducidas del ARNm modificado.

10 Además, es posible mejorar la transferencia efectiva del ARNm modificado a las células a tratar o al organismo a tratar mediante la asociación del ARNm modificado con una proteína o un péptido catiónico o una unión con los mismos. Aquí es especialmente efectiva la utilización de protamina como proteína policatiónica de enlace del ácido nucleico. Además, también es posible la utilización de otros péptidos o proteínas catiónicos, como poli-L-lisina o histonas. Este método para la estabilización del ARNm modificado se encuentra descrito en la EP-A-1083232.

15 Otro campo de aplicación de la presente invención es la vacunación, es decir la utilización del ARNm modificado para la vacunación o la utilización de la composición farmacéutica como vacuna o la utilización del ARNm modificado para la producción de la composición farmacéutica para la vacunación. La vacunación se basa en la introducción en el organismo, especialmente en la célula, de un antígeno viral, en el presente caso de la información genética para el antígeno en forma del ARNm modificado que codifica el antígeno. El ARNm modificado incluido en la composición farmacéutica se traduce en el antígeno, es decir se expresa el polipéptido o el péptido antígeno codificado por el ARNm modificado, debido a lo cual se estimula una reacción inmunológica dirigida contra este polipéptido o péptido antígeno. Con la vacunación contra un germen patológico, por ejemplo un virus, se utiliza por tanto un antígeno de superficie de un germen de este tipo para la vacunación con ayuda de la composición farmacéutica según la invención, que contiene el ARNm modificado que codifica el antígeno de superficie.

25 La composición farmacéutica según la invención se utiliza, además, contra enfermedades infecciosas virales, como SIDA (VIH), hepatitis A, B ó C, herpes, herpes Zoster (varicelas), rubéola (virus de rubéola), fiebre amarilla, dengue etc. (flavi-virus), gripe (virus de influenza), enfermedades infecciosas hemorrágicas (virus Marburg o Ebola). Preferentemente se codifican por el ARNm modificado, también en el caso de enfermedades infecciosas, los correspondientes antígenos de superficie con el mayor potencial antígeno. En los genes mencionados de gérmenes virales patógenos, esto es típicamente una forma segregada de un antígeno de superficie. Además, preferentemente se utilizan según la invención ARNm que codifican polipéptidos, donde se trata en los polipéptidos de polipeptidos, por ejemplo de los antígenos arriba mencionados, especialmente antígenos de superficie de gérmenes virales patógenos, de preferencia formas de proteína segregadas.

35 Además, el ARNm modificado según la invención también puede contener, aparte del péptido o el polipéptido antígeno o de efecto genético terapéutico al menos otro segmento funcional que codifica, por ejemplo, una citoquina que promueve la reacción inmunológica (monoquina, linfoquina, interleuquina ó quimioquina, como IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-12, INF- α , INF- γ , GM-CFS, LT- α o factores de crecimiento como hGH.

40 Además, para aumentar la inmunogenicidad, la composición farmacéutica según la invención puede contener uno o más adyuvantes. Bajo "adyuvante" se ha de entender aquí cualquier compuesto químico o biológico que favorece una reacción inmunológica específica. Dependiendo de los diferentes tipos de adyuvantes pueden tenerse en cuenta aquí diferentes mecanismos. Por ejemplo, los compuestos que promueven una endocitosis por células dendríticas (DC) en el ARNm modificado incluido en la composición farmacéutica forman una primera clase de adyuvantes a utilizar. Otros compuestos que permiten la maduración de las DC, por ejemplo lipopolisacáridos, TNF- α ó CD40-Ligando, son otra clase de adyuvantes adecuados. En general, se puede utilizar como adyuvante cualquier agente que actúe sobre el sistema inmunológico del tipo de "señal de peligro" (LPS, GP96, oligonucleótidos con el motivo CpG) o citoquinas, como GM-CFS, que permiten reforzar la reacción inmunológica contra un antígeno codificado por el ARNm modificado y/o ejercer una influencia precisa sobre esta reacción. Son especialmente preferentes las citoquinas arriba mencionadas. Otros adyuvantes conocidos son hidróxido de aluminio, adyuvante de Freund, así como los péptidos o polipéptidos catiónicos de estabilización arriba mencionados, como protamina. Además, son especialmente adecuados lipopéptidos como Pam3Cys, para ser utilizados como adyuvantes en la composición farmacéutica de la presente invención, véase Deres y col., Nature 1989, 342: 561-564.

55 La composición farmacéutica según la invención contiene, además del ARNm modificado, un excipiente farmacéuticamente compatible y/o un vehículo farmacéuticamente compatible. Los correspondientes métodos para la formulación adecuada y para la producción de la composición farmacéutica según la invención se describen en "Remington's Pharmaceutical Sciences" (Mack Pub. Co, Easton, PA, 1980), cuyo contenido completo forma parte integrante de la descripción de la presente invención. Para la administración parenteral se pueden utilizar como excipientes, por ejemplo, agua esterilizada, soluciones de sal común esterilizadas, polialquilenglicoles, naftaleno hidrogenado y especialmente polímeros lactida biocompatibles, copolímeros de lactida/glicolida o copolímeros polioxietileno/polioxipropileno. Las composiciones según la invención pueden contener sustancias de carga o sustancias como lactosa, manitol, sustancias para la unión covalente de polímeros, por ejemplo de polietilenglicol, en inhibidores según la invención, para la formación de complejos con iones metálicos o para la inclusión de materiales en o sobre

preparados especiales de compuestos polímeros, por ejemplo polilactato, ácido poliglicólico, hidrogel o en liposomas, microemulsión, micelas, vesículas unilaminares o multilaminares, fragmentos de eritrocitos o esferoplastos. Las correspondientes formas de realización de las composiciones se seleccionan en función del comportamiento físico, por ejemplo con vistas a la solubilidad, la estabilidad, la biodisponibilidad o la posibilidad de desintegración. Una liberación controlada o constante del componente del principio activo según la invención en la composición incluye formulaciones basadas en depósitos lipofílicos (por ejemplo ácidos grasos, ceras o aceites). Dentro del marco de la presente invención también se describen recubrimientos de sustancias o composiciones según la invención que contienen tales sustancias, es decir recubrimientos con polímeros (por ejemplo poloxámeros o poloxaminas). Además, las sustancias o las composiciones según la invención pueden tener recubrimientos de protección, por ejemplo inhibidores de proteasa o reforzadores de la permeabilidad. Los excipientes preferentes son, típicamente, sustancias acuosas excipientes, utilizándose agua para la inyección (WFI) o agua tamponada con fosfato, citrato o acetato etc., y ajustándose el pH normalmente a 5,0-8,0, preferentemente 6,0-7,0. El excipiente o el vehículo contiene, además y preferentemente, ingredientes salinos, por ejemplo cloruro sódico, cloruro potásico u otros componentes que, por ejemplo, convierten la solución en isotónica. Además, el excipiente o el vehículo puede contener, aparte de los ingredientes arriba indicados, componentes como seroalbúmina humana (HSA), polisorbato 80, azúcares o aminoácidos.

La forma de administración y la dosificación de la composición farmacéutica según la invención dependen de la enfermedad a tratar y de su estado de gravedad, así como también del peso corporal, de la edad y del sexo del paciente.

La concentración de ARNm modificado en tales formulaciones puede variar, por tanto, en un amplio rango, de 1 µg hasta 100 mg/ml. La composición farmacéutica según la invención se administra al paciente preferentemente de forma parenteral, por ejemplo intravenosa, intra-arterial, subcutánea, intramuscular. También es posible administrar la composición farmacéutica de forma tópica u oral.

Así, también se proporciona según la invención un procedimiento para el tratamiento de las enfermedades arriba indicadas o un procedimiento de vacunación para prevenir las enfermedades arriba mencionadas, procedimiento que comprende la administración de la composición farmacéutica según la invención a un paciente, en especial a un ser humano.

Además, también se proporciona un procedimiento útil para averiguar la secuencia modificada del ARNm contenido en la composición farmacéutica según la invención. en este caso, según la invención se lleva a cabo la adaptación de las secuencias de ARN con dos metas diferentes de optimización. Por un lado, una con un contenido de G/C lo más alto posible, por otro lado, teniendo en cuenta la frecuencia mejor posible de los ARNt's según Codon Usage. En el primer paso del procedimiento se lleva a cabo una traducción virtual de una secuencia cualquiera de ARN (o ADN) con el fin de generar la correspondiente secuencia de aminoácidos. Partiendo de la secuencia de aminoácidos se realiza una traducción virtual inversa, que proporciona las posibilidades para los correspondientes codones debido al código genético degenerado. Según la optimización o la modificación requerida, se utilizan para la selección del codón apropiado listas correspondientes de selección y algoritmos de optimización. Típicamente la realización de los algoritmos se lleva a cabo con ayuda de un software apropiado en una computadora. Así, se obtiene la secuencia optimizada de ARNm y puede editarse, por ejemplo, con ayuda de un correspondiente dispositivo de visualización, y compararse con la secuencia original (del tipo salvaje). Lo mismo es aplicable también para la frecuencia de los distintos nucleótidos. Las modificaciones frente a la secuencia original de los nucleótidos se destacan aquí de forma preferente. Además, según una forma de realización preferente, se dan entrada por lectura a secuencias estables conocidas en la naturaleza que pueden proporcionar la base para un ARN estabilizado según motivos naturales de secuencias. También se puede prever un análisis de estructura secundario que puede analizar, con ayuda de cálculos de estructura, características de estabilización y de desestabilización o bien campos del ARN.

Las figuras muestran:

Figura 1: secuencias del tipo salvaje y modificadas para la proteína de matriz de la gripe. La figura 1a muestra el gen de tipo salvaje y la figura 1B muestra la secuencia de aminoácidos derivada del mismo (código de una letra). La figura 1C muestra una secuencia del gen que codifica la proteína de matriz de la gripe, secuencia cuyo contenido de G/C está incrementado frente a la secuencia del tipo salvaje. La figura 1D muestra la secuencia de un gen que codifica una forma segregada de la proteína de matriz de la gripe (incluida una secuencia de señal terminal de N) donde el contenido de G/C de la secuencia frente a la secuencia del tipo salvaje está incrementado. La figura 1E muestra un ARNm que codifica la proteína de matriz de la gripe, ARNm que contiene secuencias de estabilización si se compara con el ARNm del tipo salvaje. La figura 1F muestra un ARNm que codifica la proteína de matriz de la gripe, ARNm que tiene, además de las secuencias de estabilización, un mayor contenido de G/C. La figura 1G muestra también un ARNm modificado que codifica la forma segregada de la proteína de matriz de la gripe y tiene, si se compara con el tipo salvaje, secuencias de estabilización y un mayor contenido de G/C. En la figura 1A y las figuras 1C a 1G se destacan en negrilla los codones de inicio y parada. Los nucleótidos modificados frente a la secuencia del tipo salvaje de la figura 1A están representados en las figuras 1C a 1G con mayúsculas.

Figura 2: secuencias del tipo salvaje y secuencias modificadas según la invención que codifican el antígeno del tumor MAGE1. La figura 2A muestra la secuencia del gen del tipo salvaje y la figura 2B la secuencia de aminoácidos derivada del mismo (código de tres letras). La figura 2C muestra un ARNm modificado que codifica el MAGE1, cuyo contenido de G/C está incrementado frente al tipo salvaje. La figura 2D muestra la secuencia de un ARNm modificado que codifica el MAGE1, ARNm en el que se optimizó la utilización del codón en cuanto a la codificación de ARNt que se presentan con la mayor frecuencia posible en la célula. Los codones de inicio y parada están marcados en negrilla.

Los siguientes ejemplos explican más en detalle la presente invención sin limitarla.

Ejemplo 1

Como forma de realización a modo de ejemplo del procedimiento para averiguar la secuencia de un ARNm modificado según la invención se preparó un programa de ordenador que modificaba la secuencia de nucleótidos de un ARNm cualquiera con ayuda del código genético o de su naturaleza degenerativa, de manera que resultara en un contenido máximo de G/C en relación con la utilización de codones que codifican ARNt's que se presentan con la mayor frecuencia posible en la célula, siendo la secuencia de aminoácidos codificada por el ARNm modificado idéntica frente a la secuencia sin modificar. Alternativamente, también se puede modificar solamente el contenido de G/C o solamente la utilización de codones frente a la secuencia original.

El código fuente en Visual Basic 6.0 (entorno de desarrollo utilizado: Microsoft Visual Studio Enterprise, 6.0 con Servicepack 3) está indicado en el anexo.

Ejemplo 2

Con ayuda del programa de ordenador del Ejemplo 1 se preparó, a partir de *E. coli*, un constructo de ARN con una secuencia del gen lac-Z optimizada en cuanto a la estabilización y la eficacia de traducción. Así se pudo conseguir un contenido de G/C del 69% (en comparación con la secuencia del tipo salvaje del 51%; véase Kalnins y col., EMBOJ. 1983, 2(4): 593-597). Mediante la síntesis de los oligonucleótidos solapados que tienen la secuencia modificada, se obtuvo la secuencia optimizada según procedimientos conocidos en la técnica. Los oligonucleótidos terminales tenían las siguientes interfaces de restricción: en el extremo 5' una interfaz *EcoRV*- así como en el extremo 3' una interfaz *BglII*. Mediante la digestión con *EcoRV/BglII* se transformó la secuencia modificada lacZ en el plásmido pT7Ts (número de acceso del banco de genes U 26404; véase Lai y col., s.o.). pT7Ts contiene como regiones no traducidas secuencias del gen de globina β de *Xenopus levis* en cada caso en 5' y 3'. El plásmido se cortó antes de introducir la secuencia modificada de lacZ con las enzimas de restricción mencionadas.

El constructo de pT7Ts-lac-Z se multiplicó en bacterias y se purificó por extracción de fenol-cloroformo. Se transcribieron *in vitro* 2 μ g del constructo por procedimientos conocidos por el técnico en la materia, con lo que se obtuvo el ARNm modificado.

Ejemplo 3

Con ayuda del programa de ordenador según la invención del Ejemplo 1, se optimizó el gen para la proteína de matriz de la gripe (secuencia de tipo salvaje, véase figura 1A, secuencia de aminoácidos derivada figura 1B). Con ello se obtuvo la variante de secuencia rica en G/C representada en la figura 1C. También se averiguó una secuencia rica en G/C que codificaba la forma segregada de la proteína de matriz de la gripe y codificaba una secuencia de señal terminal de N (véase la figura 1D). La forma segregada de la proteína de matriz de la gripe tiene la ventaja de su mayor inmunogenicidad en comparación con la forma no segregada.

Partiendo de las secuencias optimizadas se diseñaron las moléculas de ARNm correspondientes. El ARNm optimizado en cuanto al contenido de G/C y la utilización de codones para la proteína de matriz de la gripe se equipó adicionalmente con secuencias de estabilización en el sitio 5' y 3' (las secuencias de estabilización provienen de los UTRs 5' o 3' del ARNm de β -globina de *Xenopus laevis*; véase vector pT7Ts en Lai y col., s.o.) (véanse las figuras 1E y 1F). Igualmente también se optimizó la secuencia en la zona traducida del ARNm que codificaba la forma segregada de la proteína de matriz de la gripe y se equipó con las secuencias de estabilización mencionadas (véase la figura 1G).

Ejemplo Comparativo 4

Con ayuda del programa de ordenador del Ejemplo 1 se modificó el ARNm que codificaba el antígeno tumoral MA-GE1. Así se averiguó la secuencia representada en la figura 2C que tiene un contenido en G/C mayor (351 G, 291 C) en un 24% si se compara con la secuencia de tipo salvaje (275 G, 244 C). Además, se mejoró la secuencia de tipo salvaje por utilización alternativa de codones en cuanto a la eficacia de traducción, debido a la codificación de los ARNt que existen con mayor frecuencia en la célula (véase la figura 2D). Mediante la utilización alternativa de codones también se incrementó el contenido de G/C en un 24%.

Anexo: texto fuente de un programa de ordenador según la invención

Curevac_Genetic_Controls con los siguientes módulos

Type=Control

Reference=*G{00020430-0000-0000-C000-000000000046}#2.0#0#.\.\WINNTSystem32\STB#OLE Automation

5 Object={0D452EE1-E08F-101A-852E-02608CAD0BB4}#2.0#0; FM20.DLL

Object={0D452EE1.E08F-101A-852E.02608C4D0BB4}# 2.0# 0; FM20.DLL

Object={F9043C88-F6F2-101A-A3C9-08002B2F49FB}# 1.2# 0; COMDLG32.OCX

UserControl=Curevac_Amino.ctl

Startup="(None)"

10 HelpFile=""

Title ="Curevac Genetic Controls"

Command32=" "

Name="Curevac_Genetic_Controls"

HelpContextID="0"

15 CompatibleMode="1"

MajorVer=1

MinorVer=0

RevisionVer=0

AutoIncrementVer=0

20 ServerSupportFiles =0

VersionComments="the RNA people"

VersionCompanyName="CureVac GmbH"

VersionFileDescription-"Controls for Handling of Nucleotids"

VersionLegalCopyright="byChristian Klump"

25 VersionLegalTrademarks="Guevac Genetic Controls(tm)"

VersionProductName="Curevac Genetic Controls"

CompilationType=0

OptimizationType=0

FavorPentiumPro(tm)=0

30 CodeViewDebugInfo=0

NoAliasing=0

BoundsCheck=0

OverflowCheck=0

FIPointCheck=0

35 FDIVCheck=0

UnroundedFP=0

StartMode=1

```
Unattended=0
Retained=0
ThreadPerObject=0
MaxNumberOfThreads=1
5 ThreadingModel=1
DebugStartupOption=1
DebugStartupComponent=Curevac_Amino
Name: Curevac_Amino.ctl
Code:
10 VERSION 5.00
Object = "{0D452EE1-E08F-101A-852E-02608C4D0BB4}# 2.0# 0"; "FM20.DLL"
Begin VB.UserControl Curevac_Amino
CanGetFocus = 0 Talse
ClientHeight = 690
15 ClientLeft = 0
ClientTop = 0
ClientWidth = 1200
ScaleHeight = 690
ScaleWidth = 1200
20 Begin VB.Line linLower
X1 = 0
X2 = 1080
Y1 = 600
Y2 = 600
25 End
Begin VB.Line linUpper
X1 = 120
X2 = 1080
Y1 = 0
30 Y2 = 0
End
Begin MSForms.Label lblAminoAcid
DragMode = 1 'Automatic
Height = 255
35 Left = 120
TabIndex = 1
Top = 360
```

```
Width = 975
Size - "1720;450"
SpecialEffect = 1
FontName = "Lucida Sans Unicode"
5 FontEffects = 1073741826
FontHeight = 165
FontCharSet = 0
FontPitchAndFamily- 2
ParagraphAlign = 3
10 End
Begin MSForms.Label lblTriplett
DragMode = 1 'Automatic
Height = 255
Left = 120
15 TabIndex = 0
Top = 120
Width = 975
Size = "1720;450"
SpecialEffect = 1
20 FontHeight = 165
FontChatSet = 0
FontPitchAndFamily= 2
ParagraphAlign = 3
End
25 End
Attribute VB_Name = "Curevac_Amino"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = True
Attribute VB_PredeclaredId = False
30 Attribute VB_Exposed = True
Option Explicit
Private msAAShortcut As String
Private msAAName As String
Private msBestTriplett As String
35 Private msSecondBest As String
Private msThirdBest As String
Private msTriplett As String
```

```

Private msBackColor As Long
Private mbShowOriginal As Boolean
Public Enum enuAminoAcid
amaGlycin
5  amaAlanin
amaValin
amaLeucin
amalsoLeucin
amaPhenylalanin
10 amaTyrosin
amaTryptophan
amaAsparaginAcid
amaAsparagin
armGlutaminAcid
15 amaGlutamin
amaSerin
amaThreonin
amaCystein
amaMethionin
20 amaProlin
amaHistidin
amaLysin
amaStop
amaStart
25 End Enum
Private Sub Main()
Call UserControl_Resize
End Sub
Public Sub Translation(ByVal sTriplet As String)
30 msTriplet = sTriplet
Select Case sTriplet
Case "GGU", "GGC", "GGA", "GGG"
msAAShortcut = "GLY"
Case "GCU", "GCC", "GCA", "GCG"
35 msAAShortcut = "ALA"
Case "GUU", "GUC", "GUA", "GUG"
msAAShortcut = "VAL"

```

```

Case "UUC", "UUG", "CUU", "CUA", "CUG", "CUC"
msAAShortcut = "LEU"
Case "AUA", "AUU", "AUC"
msAAShortcut = "ILE"
5 Case "UUU", "UUC"
nwAAShortcut = "PHE"
Case "UAU", "UAC"
msAAShortcut = "TYR"
Case "UGG"
10 msAAShortcut = "TRP"
Case "GAU", "GAC"
msAAShortcut = "ASP"
Case "AAU", "AAC"
msAAShortcut = "ASN"
15 Case "GAA", "GAG"
msAAShortcut = "GLU"
Case "CAA", "CAG"
msAAShortcut = "GLN"
Case "AGU", "AGC", "UCA", "UCU", "UCG", "UCC"
20 msAAShortcut = "SER"
Case "ACA", "AGU", "ACG", "AOC"
msAAShortcut = "THR"
Case "UGU", "UGC"
msAAShortcut = "CYS"
25 Case "AUG"
msAAShortcut = "MET"
Case "CCU", "COG", "COC"
msAAShortcut = "PRO"
Case "CAU", "CAC"
30 msAAShortcut = "HIS"
Case "AGA", "AGG", "CGA", "CGU", "CGG", "CGC"
msAAShortcut = "ARG"
Case "AAA", "AAG"
msAAShortcut = "LYS"
35 Case "UAA", "UAG", "UGA"
msAAShortcut = "STP"
End Select

```

```
Call Synthesis(msAAShortcut)
Call Show
End Sub
Public Sub Synthesis(ByVal sShortCut As String)
5  msAAShortcut = sShortCut 'sólo dirigida a Sub externo'
  Select Case msAAshorcut
    Case "GLY"
      msAAName ="Glycin"
      msBestTriplett = "GGC"
10  msSecondBest = "GGG"
      msBackColor=8421631
      Case "ALA"
        msAAName ="Alanin"
        msBestTriplett = "GCG"
15  msSecondBest ="GCC"
      msBackColor = 8454143
      Case "VAL"
        msAAName = "Valin"
        msBestTriplett = "GUC"
20  msSecondBest = "GUG"
      msBackColor = 8454016
      Case "LEU"
        msAAName = "Leucin"
        msBestTriplett = "CUG"
25  msSecondBest = "CUC"
      msBackColor = 16777088
      Case "ILE"
        msAAName ="Isoleucin"
        msBestTriplett = "AUC"
30  msBackColor = 12615935
      Case "PHE"
        msAAName = "Phenylalanin"
        msBestTriplett = "UUC"
        msBackColor = 255
35  Case "TYP"
        msAAName = "Tyrosin"
        msBestTriplett = "UAC"
```

msBackColor = 4210816
Case "TRP"
msAAName = "Tryptophan"
msBestTriplet = "UGG"

5 msBackColor = 4227327
Case "ASP"
msAAName = "Asparaginspure"
msBestTriplet " "GAC"
msBackColor = 8388863

10 Case "ASN"
msAAName = "Asparagin"
msBestTriplet ="AAC"
msBackColor = 4227200
Case"GLU"

15 msAAName = "Glutaminspure"
msBestTriplet = "GAG"
msBackColor = 32768
Case "GLN"
msAAName = "Glutamin"

20 msBestTriplet = "GGC"
msBackColor = 8421440
Case "SER"
msAAName = "Serin"
msBestTriplet = "AGC"

25 msSecondBest = "UCG"
msThirdBest = "UCC"
msBackColor= 16512
Case "THR"
msAAName = "Threonin"

30 msBestTriplet = "ACG"
msBackColor =16711680
Case "CYS"
msAAName = "Cystein"
msBestTriplet = "UGC"

35 msBackColor = 6932960
Case "MET"
msAAName = "Methionin"

```

msBestTriplett= "AUG"
msBackColor = 10417643
Case "PRO"
msAAName = "Prolin"
5  msBestTriplett= "CCG"
msSecondBest = "CCC"
msBackColor = 12898746
Case "HIS"
msAAName ="Histidin"
10  msBestTriplett = "CAC"
msBackColor = 12898746
Case "ARG"
msAAName = "Arginin"
rmBestTriplett = "CGC"
15  msSecondBest = "CGG"
msBackColor = 6174925
Case "LYS"
msAAName = "Lysin"
msBestTriplett = "AAG"
20  msBackColor =14141641
Case "STP"
msAAName = "Stop"
msBestTriplett - "UGA"
msSecondBest = "UAG"
25  msBackColor =11332093
End Select
End Sub
Public Sub Show()
lblAminoAcid.Caption = msAAShortcut
30  If mbShowOriginal = True Then
lblTriplett.Caption = msTriplett
Else
lblTriplett.Caption = msBestTriplett
End If
35  lblAminoAcid.BackColor = msBackColor
lblTriplett.BackColor= msBackColor
End Sub

```

```
Public Function ChooseBestGC(sShortCut As String)
Select Case sShortCut
Case "GLY"
msAAName = "Glycin"
5 msBestTriplett ="GGC"
msSecondBest = "GGG"
msBackColor = 8421631
Case "ALA"
msAAName = "Alanin"
10 msBestTriplett - "GCC"
msSecondBest = "GCC"
msBackColor = 8454143
Case "VAL"
msAAName = "Valin"
15 msBestTriplett = "GUC"
msSecondBest = "GUG"
msBackColor = 8454016
Case "LEU"
msAAName = "Leucin"
20 msBestTriplett = "CUG"
msSecondBest = "CUC"
msBackColor = 16777088
Case "ILE"
msAAName= "Isoleucin"
25 msBestTriplett = "AUC"
msBackColor = 12615935
Case "PHE"
msAAName = "Phenylalanin"
msBestTriplett= "UUC"
30 msBackColor = 255
Case "TYR"
msAAName = "TyroSin"
msBestTriplett= "UAC"
msBackColor = 4210816
35 Case "TRP"
msAAName = "tryptophan"
msBestTriplett - "UGG"
```

msBackColor= 4227327
Case "ASP"
msAAName = "Asparaginspure"
msBestTriplett = "GAC"

5 msBackColor = 8388863
Case "ASN"
msAAName = "Asparagin"
msBestTriplett = "AAC"
msBackColor = 4227200

10 Case "GLU"
msAAName = "Glutaminspure"
msBestTriplett = "GAG"
msBackColor = 16711808
Case "GLN"

15 msAAName = "Glutamin"
msBestTriplett= "GGC"
msBackColor = 12632256
Case "SER"
msAAName = "Serin"

20 msBestTriplett= "AGC"
msSecondBest = "UCG"
msThirdBest = "UCC"
msBackColor= 8421504
Case "THR"

25 msAAName = "Threonin"
msBestTriplett= "ACG"
msBackColor = &HD400FF
Case "CYS"
msAAName = "Cystein"

30 msBestTriplett = "UGC"
msBackColor = &HD500DFF
Case "MET"
msAAName = "Methionin"
msBestTriplett = "AUG"

35 msBackColor = &HD600FF
Case "PRO"
msAAName = "Prolin"

```

msBestTriplet = "CCG"
msSecondBest = "CCC"
msBackColor = &HD700FF
Case "HIS"
5  msAAName = "Histidin"
   msBestTriplet = "CAC"
   msBackColor = &HD8C0FF
   Case "ARG"
    msAAName = "Argining"
10  msBestTriplet = "CGC"
    msSecondBest = "CGG"
    nuBackColor = &HD900FF
    msAAName = "Lysin"
    msBestTriplet = "AAG"
15  msBackColor = &HE0C0DFF
    Case "STP"
    msAAName = "Stop"
    msBestTriplet = "UGA"
    msSecondBest = "UAG"
20  msBackColor = &HE1C0FF
    End Select
    End Function

Public Property Let ShowOriginal(ByVal bNewValue As Boolean)
mbShowOriginal = bNewValue
25 End Property

Public Property Get ShowOriginal() As Boolean
ShowOriginal = mbShowOriginal
End Property

Public Property Get AAShortcut() As String
30 AAShortcut = msAAShortcut
End Property

Public Property Get AAName() As String
AAName = msAAName
End Property
35 Public Property Get BestTriplet() As String
BestTriplet = msBestTriplet
End Property

```

```

Public Property Get SecondBest() As String
SecondBest = msSecondBest
End Property
Public Property Get ThirdBest() As String
5 ThirdBest =msThirdBest
End Property
Public Property Get BackColor() As String
BackColor = msBackColor
End Property
10 Private Property Get Triplet() As String
Triplet = msTriplet
End Property
Private Sub UserControl_Resize()
Dim IOuterWidth As Long
15 Dim IDistance As Long
IDistance =30
IOuterWidth = Scale Width + IDistance
With lblTriplet.
.Top = 0
20 .Left= IDistance / 2
.Width = ScaleWidth - IDistance
.Height = ScaleHeight / 2 + 30
End With
With lblAminoAcid
25 .Top = ScaleHeight / 2 - 30 .Left = IDistance / 2
.Width = ScaleWidth - IDistance
Height = ScaleHeight / 2
End With
With linUpper
30 .X1 = ScaleLeft
.Y1 =0
.X2 = Schale Width
.Y2 =0
End With
35 With linLower
.X1 = ScaleLeft
.Y1= ScaleHeight - 20

```

```

.X2 = Scale Width
.Y2 = ScaleHeight - 20
End With
End Sub
5  Curevac_RNA_Optimizer con los siguientes módulos
Name: Curevac_RNA_Optimizer.vbp
Code:
Type=Exe
Reference=*G{00020430-0000-0000-C000-
10  000000000046)# 2.0# 0# ..\..\WINNT\System32\STDOLE2.TLB# OLE Automation
Object-{F9043C88-F6F2-101A-A3C9-08002B2F49FB)# 1.2#0; COMDLG32.OCX
Module= basMain; basMain.bas
Form=mdiMain.frm
Object=*\ACurevac_Genetic_Controls.vbp
15  Object={38911DA0-E448-11D0-84A3-00DD01104159)# 1.1#0; COMCT332.OCX
Object={3B7C8863-D78F-101B-B9B5-04021C009402)# 1.2#0; RICHTX32.OCX
Form=frmOutPut.frm
Form=frmStatistics.frm
Module =basPublics; basPublics.bas
20  Form=frmOptions.frm
Form=frmInput.frm
Form=frmDisplay.frm
Form=frmAbout.frm
Startup="mdiMain"
25  HelpFile=""
Title="RNA-Optimizer"
Command32=""
Name="RNA_Optimizer"
HelpContextID="0"
30  CompatibleMode="0"
MajorVer=1
MinorVer=0
RevisionVer=0
AutoIncrementVer=0
35  ServerSupportFiles =0
VersionComments="the RNA people"
VersionCompanyName="CureVac GmbH"

```

```

VersionFileDescription="Application for Optimization of RNA"
VersionLegalCopyright="by Christian Klump"
VersionLegalTrademarks="Curevac RNA-Optimizer(tm)"
VersionProductName="Curevac RNA-Optimizer"
5  CompilationType =0
   OptimizationType =0
   FavorPentiumPro(tm) =0
   CodeViewDebugInfo=0
   NoAliasing=0
10  BoundsCheck=0
   OverflowCheck=0
   FIPointCheck=0
   FDIVCheck=0
   UnroundedFP=0
15  StartMode=0
   Unattended=0
   Retained=0
   ThreadPerObject=0
   MaxNumberOfTreads=1
20  DebugStartupOption=0
   Name: basMain.bas
   Code:
   Attribute VB_Name = "basMain"
   Option Explicit
25  Private Type udtAminoAcid
   sShortcut As String
   sName As String
   sBestTriplet As String
   sSecondBest As String
30  sThirdBest As String
   End Type
   Private mastrAminoAcids() As udtAminoAcid
   Public Sub RebuildChain(bShowOriginal As Boolean)
   Dim iLoop As Integer
35  For iLoop =1 To glAminoCounter
   With frmDisplay.Curevac_Amino1(iLoop)
   .ShowOriginal = bShowOriginal

```

```

.Show
End With
Next iLoop
End Sub
5 Public Sub ResetPublics()
gsOriginalCode=""
gsFormattedCode = " "
glCodeLenght = ""
End Sub
10 Public Sub BuildAminoChain(ByVal sTempCode As String, bGraphical As Boolean)
Dim IVerticalOffset As Long
Dim IHorizontalOffset As Long
Dim ITop As Long
Dim asTriplets() As String
15 Dim sAminoChain As String
Dim Loop As Long
If bGraphical = True Then
ITop = frmDisplay.optNucleotids(0).Height + 40
End If
20 ReDim mastrAminoAcids(glAminoCounter)
asTriplets = Split(sTempCode)
For ILoop =0 To glAminoCounter- 1
If bGraphical = True Then
Load frmDisplay.Curevac_Amino1(ILoop +1)
25 With frmDisplay.Curevac_Amino1(1Loop + 1)
Call. Translation(asTriplets(ILoop))
.Left = ILoop* Width - IHorizontalOffset
.Top = IVerticalOffset + ITop
If .Left + 2*.Width > frmDisplay.ScaleWidth Then
30 IVerticalOffset = IVerticalOffset + .Height + 150
IHorizontalOffset - ILoop* With + .Width
End If
sAminoChain = sAminoChain + .aaShortcut + ","
.Visible - True
35 End With
Else
mastrAminoAcids(ILoop) = Translation(asTriplets(ILoop))

```

```

sTest = sTest + mastrAminoAcids(iLoop).sName + ";"
End If
Next
If bGraphical = True Then
5 frmDisplayShow
Else
Call frmOutPut.Show
End If
Debug.Print sAminoChain
10 End Sub

Private Function Translation(ByVal sTriplet As String) As udtAminoAcid
Dim stiTemp As udtAminoAcid
If mdiMain.optKindOfOptimize(0) = True Then
'Optimization for Most GC
15 Select Case sTriplet
Case "GGU", "GGC", "GGA", "GGG"
stiTemp.sShortcut = "GLY"
stiTemp.sName = "Glycin"
stiTemp.sBestTriplet = "GGC"
20 stiTemp.sSecondBest = "GGG"
Case "GCU", "GCC", "GCA", "GCG"
stiTemp.sShortcut = "ALA"
stiTemp.sName = "Alanin"
stiTemp.sBestTriplet = "GCG"
25 stiTemp.sSecondBest = "GCC"
Case "GUU", "GUC", "GUA", "GUG"
stiTemp.sShortcut = "VAL"
stiTemp.sName = "Valin"
stiTemp.sBestTriplet = "GUC"
30 stiTemp.sSecondBest = "GUG"
Case "UUA", "UUG", "CUU", "CUA", "CUG", "CUC"
stiTemp.sShortcut = "LEU"
stiTemp.sName = "Leucin"
stiTemp.sBestTriplet = "CUG"
35 stiTemp.sSecondBest = "CUC"
Case "AUA", "AUU", "AUC"
stiTemp.sShortcut = "ILE"

```

```

stiTemp.sName = "Isoleucin"
stiTemp.sBestTriplett = "AUC"
Case "UUU", "UUC"
stiTemp.sShortcut = "PHE"
5  stiTemp.sName= "Phenylalanin"
stiTemp.sBestTriplett = "UUC"
Case "UAU", "UAC"
stiTemp.sShortcut = "TYR"
stiTemp.sName = "Tyrosin"
10 stiTemp.sBestTriplett = "UAC"
Case "UGG"
stiTemp.Shortcut = "TRP"
stiTemp.sName = "Tryptophan"
stiTemp.sBestTriplett = "UGG"
15 Case "GAU", "GAC"
stiTemp.sShortcut = "ASP"
stiTemp.sName = "Asparaginspure"
stiTemp.sBestTriplett = "GAC"
Case "AAU", "AAC"
20 stiTemp.sShortcut = "ASN"
stiTemp.sName = "Asparagin"
stiTemp.sBestTriplett = "AAC"
Case "GAA", "GAG"
stiTemp.sShortcut = "GLU"
25 stiTemp.sName = "Glutaminspure"
stiTemp.sBestTriplett = "GAG"
Case "CAA", "CAG"
stiTemp.sShortcut = "GLN"
stiTemp.sName = "Glutamin"
30 stiTemp.sBestTriplett = "CAG"
Case "AGU", "AGC", "UCA", "UCU", "UCG", "UCC"
stiTemp.sShortcut = "SER"
stiTemp.sName = "Serin"
stiTemp.sBestTriplett = "AGC"
35 stiTemp.sSecondBest = "UCG"
stiTemp.sThirdBest = "UCC"
Case "ACA", "ACU", "ACG", "ACC"

```

```

stiTemp.sShortcut = "THR"
stiTemp.sName = "Threonin"
stiTemp.sBestTriplett = "ACG"
Case "UGU", "UGC"
5  stiTemp.sShortcut = "CYS"
stiTemp.sName = "Cystein"
stiTemp.sBestTriplett = "UGC"
Case "AUG"
stiTemp.sShorcut = "MET"
10 stiTemp.sName = "Methionin"
stiTemp.sBestTriplett = "AUG"
Case "CCA", "OCU", "CCG", "CCC"
stiTemp.sShortcut = "PRO"
stiTemp.sName = "Prolin"
15 stiTemp.sBestTriplett = "CCG"
stiTemp.sSecondBest = "CCC"
Case "CAU", "CAC"
stiTemp.Shortcut = "HIS"
stiTemp.sName = "Histidin"
20 stiTemp.sBestTriplett = "CAC"
Case "AGA", "AGG", "CGA", "CGU", "CGG", "CGC"
stiTemp.sShortcut = "ARG"
stiTemp.sName = "Arginin"
stiTemp.sBestTriplett = "CGC"
25 stiTemp.sSecondBest = "CGG"
Case "AAA", "AAG"
stiTemp.sShorccut = "LYS"
stiTemp.sName = "Lysin"
stiTemp.sBestTriplett = "AAG"
30 Case "UAA", "UAG", "UGA"
stiTemp.sShortcut = "STP"
stiTemp.sName = "Stop"
stiTemp.sBestTriplett = "UGA"
stiTemp.sSecondBest = "UAG"
35 End Select
Translation = stiTemp
Else 'Optimization for best frequency

```

```

Select Case sTriplet
Case "GGU", "GGC", "GGA", "GGG"
stiTemp.sShortcut = "GLY"
stiTemp.sName = "Glycin"
5 stiTemp.sBestTriplet = "GGC"
stiTemp.sSecondBest = "GGU"
Case "GCU", "GCC", "GCA", "GCG"
stiTemp.sShortcut = "ALA"
stiTemp.sName = "Alanin"
10 stiTemp.sBestTriplet = "GCC"
stiTemp.sSecondBest = "GCU"
Case "GUU", "GUC", "GUA", "GUG"
stiTemp.sShortcut = "VAL"
stiTemp.sName = "Valin"
15 stiTemp.sBestTriplet = "GUG"
stiTemp.sSecondBest = "GUC"
Case "UUA", "UUG", "CUU", "CUA", "CUG", "CUC"
stiTemp.sShortcut = "LEU"
stiTemp.sName = "Leucin"
20 stiTemp.sBestTriplet = "CUG"
stiTemp.sSecondBest = "CUC"
Case "AUA", "AUU", "AUC"
stiTemp.sShortcut = "ILE"
stiTemp.sName = "Isoleucin"
25 stiTemp.sBestTriplet = "AUC"
Case "UUU", "UUC"
stiTemp.sShortcut = "PHE"
stiTemp.sName = "Phenylalanin"
stiTemp.sBestTriplet = "UUC"
30 Case "UAU", "UAC"
stiTemp.sShortcut = "TYR"
stiTemp.sName = "Tyrosin"
stiTemp.sBestTriplet = "UAC"
Case "UGG"
35 stiTemp.sShortcut = "TRP"
stiTemp.sName = "Tryptophan"
stiTemp.sBestTriplet = "UGG"

```

Case "GAU", "GAC"
 stiTemp.sShortcut = "ASP"
 stiTemp.sName = "Asparaginspure"
 stiTemp.sBestTriplett = "GAC"

5 Case "AAU", "AAC"
 stiTemp.sShortcut = "ASN"
 stiTemp.sName = "Asparagin"
 stiTemp.sBestTriplett = "AAC"

Case "GAA", "GAG"

10 stiTemp.sShortcut = "GLU"
 stiTemp.sName = "Glutaminspure"
 stiTemp.sBestTriplett = "GAG"

Case "CAA", "CAG"

15 stiTemp.sShortcut = "GLN"
 stiTemp.sName = "Glutamin"
 stiTemp.sBestTriplett = "CAG"

Case "AGU", "AGC", "UCA", "UCU", "UCG", "UCC"

stiTemp.sShortcut = "SER"
 stiTemp.sName = "Serin"

20 stiTemp.sBestTriplett = "AGC"
 stiTemp.sSecondBest = "UCC"
 stiTemp.sThirdBest = "UCU"

Case "ACA", "ACU", "ACG", "ACC"

25 stiTemp.sShortcut = "THR"
 stiTemp.sName = "Threonin"
 stiTemp.sBestTriplett = "ACC"

Case "UGU", "UGC"

stiTemp.sShortcut = "CYS"
 stiTemp.sName = "Cystein"

30 stiTemp.sBestTriplett = "UGC"

Case "AUG"

stiTemp.sShortcut = "MET"
 stiTemp.sName = "Methionin"
 stiTemp.sBestTriplett = "AUG"

35 Case "CCA", "CCU", "CCG", "CCC"

stiTemp.sShortcut = "PRO"
 stiTemp.sName = "Prolin"

```

stiTemp.sBestTriplett = "CCC"
stiTemp.sSecondBest = "CCU"
Case "CAU", "CAC"
stiTemp.sShortcut = "HIS"
5 stiTemp.sName = "Histidin"
stiTemp.sBestTriplett = "CAC"
Case "AGA", "AGG", "CGA", "CGU", "CGG", "CGC"
stiTemp.sShorcut = "ARG"
stiTemp.sName = "Arginin"
10 stiTemp.sBestTriplett = "CGC"
stiTemp.sSecondBest = "AGG"
Case "AAA", "AAG"
stiTemp.sShortcut = "LYS"
stiTemp.sName = "Lysin"
15 stiTemp.sBestTriplett = "AAG"
Case "UAA", "UAG", "UGA"
stiTemp.sShortcut = "STP"
stiTemp.sName = "Stop"
stiTemp.sBestTriplett = "UGA"
20 stiTemp.sSecondBest = "UAG"
End Select
Translation = stiTemp
End If
End Function
25 Public Function ReverseTranscription() As String
Dim sTempCode As String
Dim iLoop As Integer
For iLoop = 0 To glAminoCounter - 1
sTempCode = sTempCode + mastrAminoAcids(iLoop).sBestTriplett
30 Next
ReverseTranscription = sTempCode
End Function
Public Sub CountBases(ByVal sTempCode As String)
Dim lLoop As Long
35 Dim sFormattedCode As String
Dim sActualBase As String
Dim lCodeLength As Long

```

```

ICodeLength = Len(sTempCode)
For ILoop =1 To CLng(ICodeLength)
sActualBase = Mid(sTempCode, ILoop, 1)
Select Case sActualBase
5  Case "A", "a"
    gIOptAdenin = gIOptAdenin + 1
    Case "U", "u", "T", "t"
    gIOptThymin = gIOptThymin + 1
    Case "G", "g"
10  gIOptGuanin = gIOptGuanin + 1
    Case "C", "c"
    gIOptCytosin = gIOptCytosin + 1
End Select
Next ILoop
15  End Sub

Name: basPublics.bas
Code:
Attribute VB_Name = "basPublics"
Option Explicit
20  Public gsOriginalCode As String
    Public gsFormattedCode As String
    Public gICodeLength As Long
    Public gIAminoCounter As Long
    Public gIAdenin As Long
25  Public gIThymin As Long
    Public gIGuanin As Long
    Public gICytosin As Long
    Public gIOptAdenin As Long
    Public gIOptThymin As Long
30  Public gIOptGuanin As Long
    Public gIOptCytosin As Long
    Public gbSequenceChanged As Boolean

Name: mdiMain.frm
Code:
35  VERSION 5.00

Object = "{38911DA0-E448-11D0-84A3-00DD01104159}# 1.1#0"; "COMCT332.OCX"
Begin VB.MDIForm mdiMain

```

```
BackColor = &H00FFFFFF&
Caption = "Curevac_RNA_Analyzer"
ClientHeight = 8745
ClientLeft = 165
5 ClientTop = 735
ClientWidth = 12255
Icon = "mdiMain.frx":0000
LinkTopic = "MDIForm1"
Picture = "mdiMain.frx":058A
10 StartUpPosition = 3 'Windows Default
WindowState = 2 'Maximized
Begin ComCtl3.CoolBar cbarMain
Align = 1 'Align Top
Height = 885
15 Left = 0
TabIndex = 0
Top = 0
Width = 12255
_ExtentX = 21616
20 _ExtentY = 1561
BandCount = 1
BandBorders = 0 'False
_CBWidth = 12255
_CBHeight = 885
25 _Version = "6.7.8988"
MinHeight1 = 825
Width1 = 4995
FixedBackground1= 0 'False
NewRow1 = 0 'False
30 Begin VB.OptionButton optKindOfOptimize
Caption = "Best Frequency"
Height = 255
Index = 1
Left = 4920
35 TabIndex = 7
Top = 480
Value = -1 True
```

```
Width = 1575
End
Begin VB.OptionButton optKindOfOptimize
Caption = "Best GC"
5 Height = 255
Index = 0
Left = 4920
TabIndex = 6
Top = 120
10 Width = 1575
End
Begin VB.CommandButton cmdShowInput
Caption = "Input"
Height = 765
15 Left = 0
Picture = "mdiMain.frx":0C2A
Style = 1 'Graphical
TabIndex = 5
Top = 0
20 Width = 840
End
Begin VB.CommandButton cmdShowOptions
Caption = "Option"
Enabled = 0 'False
25 Height = 760
Left = 3480
Picture = "mdiMain.frx":0F34
Style = 1 'Graphical
TabIndex = 4
30 Top = 0
Width = 735
End
Begin VB.CommandButton cmdShowStatistics
Caption = "Statistics"
35 Enabled = 0 'False
Height = 760
Left = 2640
```

```
Picture = "mdiMain.fx":131E
Style = 1 'Graphical
TabIndex = 3
Top = 0
5 Width = 735
End
Begin VB.CommandButton cmdShowDisplay
Caption = "Display"
Enabled = 0 'False
10 Height = 760
Left = 1800
Picture = "mdiMain.frx":1628
Style = 1 'Graphical
TabIndex = 2
15 Top = 0
Width = 735
End
Begin VB.CommandButton cmdShowOutput
Caption = "Output"
20 Enabled = 0 'False
Height = 760
Left = 960
Picture = "mdiMain.frx":20E2
Style = 1 'Graphical
25 TabIndex = 1
Top = 0
Width = 735
End
End
30 Begin VB.Menu mnuMain
Caption = "&Input..."
Index = 0
End
Begin VB.Menu mnuMain
35 Caption = "&Results"
Index = 1
Begin VB.Menu mnuResults
```

```
Caption = "&Output..."
Enabled = 0 'False
Index = 0
End
5 Begin VB.Menu mnuResults
Caption = "&Display..."
Enabled = 0 'False
Index = 1
End
10 Begin VB.Menu mnuResults
Caption = "&Statistics..."
Enabled = 0 'False
Index = 2
End
15 End
Begin VB.Menu mnuMain
Caption = "E&xtras"
Index = 4
Begin VB.Menu mnuExtras
20 Caption = "&Language"
Index = 0
Begin VB.Menu mnuLanguage
Caption = "English"
Checked = -1 True
25 Index = 0
End
Begin VB.Menu mnuLanguage
Caption = "&German"
Enabled = 0 'False
30 Index = 1
End
Begin VB.Menu mnuLanguage
Caption = "&French"
Enabled = 0 'False
35 Index = 2
End
End
End
```

```
Begin VB.Menu mnuExtras
Caption = "&Options..."
Enabled = 0 'False
Index = 1
5 End
Begin VB.Menu mnuExtras
Caption = "-"
Index = 2
End
10 Begin VB.Menu mnuExtras
Caption = "&About"
Index = 3
End
End
15 Begin VB.Menu mnuMain
Caption = "&Windows"
Index = 5
WindowList = -1 True
Begin VB.Menu mnuWindows
20 Caption = "Tile &Horizontally"
Index = 0
End
Begin VB.Menu nmuWindows
Caption = "Tile &Vertically"
25 Index = 1
End
Begin VB.Menu mnuWindows
Caption = "&Cascade"
Index = 2
30 End
End
Begin VB.Menu mnuMain
Caption = "&Exit"
Index = 6
35 End
End
Attribute VB_Name = "mdiMain"
```

```
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = False
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
5 Option Explicit
Private Sub cmdShowDisplay_Click()
Call frmDisplay.Show
End Sub
Private Sub cmdShowInput_Click()
10 Call frmInput.Show
End Sub
Private Sub cmdShowOptions_Click()
Call frmOptions.Show
End Sub
15 Private Sub cmdShowOutput_Click()
Call frmOutPut.Show
End Sub
Private Sub cmdShowStatistics_Click()
Call frmStatistics.Show
20 End Sub
Private Sub MDIForm_Load()
Call InitControls
Call frmInput.Show
End Sub
25 Private Sub InitControls()
Dim INextLeft As Long
Const ISpace As Long = 60
Const IButtonWidth As = 850
Const IButtonHeight As Long = 770
30 INextLeft = ISpace + IButtonWidth
With cmdShowInput
.Top = ISpace
.Left = ISpace
.Width = IButtonWidth
35 .Height = IButtonHeight
End With
With cmdShowOutput
```

```

.Top = ISpace
.Left = INextLeft + ISpace
.Width = IButtonWidth
.Height = IButtonHeight
5  End With
With cmdShowDisplay
.Top = ISpace
.Left = INextLeft * 2 + ISpace
.Width = IButtonWidth
10 .Height = IButtonHeight
End With
With cmdShowStatistics
.Top = ISpace
.Left = INextLeft * 3 + ISpace
15 .Width = IButtonWidth
.Height = IButtonHeight
End With
With cmdShowOptions
.Top = ISpace
20 .Left = INextLeft * 4 + ISpace
.Width = IButtonWidth
.Height = IButtonHeight
End With
cbarMain.Height = IButtonHeight + 2 * ISpace
25 End Sub
Private Sub mnuMain_Click(Index As Integer)
Select Case Index
Case 0
Call frmInput.Show
30 Case 6
Unload Me
End Select
End Sub
Private Sub mnuResults_Click(Index As Integer)
35 Select Case Index
Case 0 'Output
Call frmOutPut.Show

```

```

Case 1
Call frmDisplay.Show
Case 2 'Statistische Auswertung
Call frmStatistics.Show
5 End Select
End Sub
Private Sub mnuExtras_Click(Index As Integer)
Select Case Index
Case 0 'Output
10 Case 1
Call frmDisplay.Show
Case 3
Call frmAbout.Show(vbModal)
End Select
15 End Sub
Private Sub mnuWindows_Click(Index As Integer)
Select Case Index
Case 0 'Horizontal
Me.Arrange (vbTileHorizontal)
20 Case 1 'Vertikal
Me.Arrange (vbTileVertical)
Case 2 'Kaskadieren
Me.Arrange (vbCascade)
End Select
25 End Sub
Name: frmInput.frm
Code:
VERSION 5.00
Object = "{F9043C88-F6F2-101A-A3C9-08002B2F49FB}# 1.2# 0"; "COMDLG32.OCX"
30 Object = "{3B7C8863-D78F-101B-B9B5-04021C009402}# 1.2#0"; "RICHTX32.OCX"
Begin VB.Form frmInput
Caption = "Input"
ClientHeight = 5730
ClientLeft = 60
35 ClientTop = 345
ClientWidth = 13200
Icon = "frmInput.frx":0000

```

```
LinkTopic = "Form1"
MDIChild = -1 True
ScaleHeight = 5730
ScaleWidth = 13200
5 WindowState = 2 'Maximized
Begin RichTextLib.RichTextBox txtFormatted
Height = 1815
Left = 120
TabIndex = 3
10 Top = 360
With = 12735
_ExtentX = 22463
_ExtentY = 3201
_Version = 393217
15 Bordetstyle = 0
Enabled = -1 True
ScrollBars = 2
DisableNoScroll = -1 True
TextRTF = $"frmInput.frx":030A
20 BeginPropertyFont {0BE35203-8F91-11CE-9DE3-00AA004BB851}
Name = "Fixedsys"
Size = 9
Charset = 0
Weight = 400
25 Underline = 0 'False
Italic = 0 'False
Strikethrough = 0 'False
EndProperty
End
30 Begin VB.OptionButton optSequence
Caption = "Formatted"
Height = 255
Index = 1
Left = 2760
35 Style = 1 'Graphical
TabIndex = 2
Top = 0
```

```
Value = -1 True
Width = 855
End
Begin VB.OptionButton optSequence
5  Caption = "Original"
    Height = 255
    Index = 0
    Left = 1920
    Style = 1 'Graphical
10  TabIndex = 1
    Top = 0
    Width = 855
    End
    Begin VB.CommandButton cmdLoad
15  Caption = "Load Sequence"
        Height = 285
        Left = 0
        TabIndex = 0
        Top = 0
20  Width = 1695
        End
        Begin MSComDlg.CommonDialog CommonDialog1
            Left = 3840
            Top = -120
25  _ExtentX = 847
            _ExtentY = 847
            _Version = 393216
            End
            End
30  Attribute VB_Name = "frmInput"
        Attribute VB_GlobalNameSpace = False
        Attribute VB_Creatable = False
        Attribute VB_PredeclaredId = True
        Attribute VB_Exposed = False
35  Option Explicit
        Private Sub Form_Activate()
            Me.ZOrder
```

```

Call InitControls
End Sub
Public Sub LoadSequence()
Dim sFile As String
5 Dim sTempCode As String
On Error Go To ErrorHandler
With CommonDialog1
.FileName = App.Path & "\SampleRNA.txt"
.CancelError = True
10 Call.ShowOpen
End With
If Len(CommonDialog1.FileName) >0. Then
gbSequenceChanged = True
gsFormattedCbde=""
15 gsOriginalCode = ""
sFile = CommonDialog1.FileName
Open sFile For Input As # 1
Input # 1, gsOriginalCode
Close # 1
20 Call FormatText
Call FillSequence
mdiMain.cmdShowDisplay.Enabled = True
mdiMain.cmdShowOutput.Enabled = True
mdiMain.cmdShowStacistics.Enabled = True
25 End If
Exit Sub
ErrorHandler:
Dim IAnswer As Long
If ErrNumber <>cdlCancel Then
30 IAnswer = MsgBox("Datei konnte nicht geladen werden", vbRetryCancel, "Dateiproblem")
If IAnswer = vbRetry Then
Resume
End If
End If
35 End Sub
If IAnswer = vbRetry Then
Resume

```

```

End If
End If
End Sub
Private Sub FormatText()
5  Dim iLoop As Integer
   Dim sActualBase As String
   glCodeLength = Len(gsOriginalCode)
   For iLoop = 0 To glCodeLength - 1
   sActualBase = Mid(gsOriginalCode, iLoop +1,1)
10  If iLoop Mod 3 = 0 Then
   gsFormattedCode = gsFormattedCode + " "
   glAminoCounter = glAminoCounter + 1
   End If
   Select Case sActualBase
15  Case "A", "a"
   gsformattedCode = gsFormattedCode + "A"
   glAdenin = glAdenin + 1
   Case "U", "u", "T", "t"
   gsformattedCode = gsformattedCode + "U"
20  glThymin = glThymin + 1
   Case "G", "g"
   gsFormattedCode = gsFormattedCode + "G"
   glGuanin = glGuanin + 1
   Case "C", "c"
25  gsFormattedCode = gsFormattedCode + "C"
   glCytosin = glCytosin + 1
   End Select
   Next iLoop
   gsFormattedCode = Trim(gsFormattedCode)
30  End Sub
Private Sub FillSequence()
   If optSequence(0).Value = True Then
   txtFormatted.Text = gsOriginalCode
   Else
35  txtFormatted.Text = gsFormattedCode
   End If
End Sub

```

```

Private Sub cmdShowOutput_Click()
Call frmOutPut.Show
End Sub
Private Sub cmdLoad_Click()
5 Call LoadSequence
End Sub
Private Sub InkControls()
Const ISpace As Long = 40
Const IButtonHeight As Long = 285
10 With cmdLoad
.Top = ISpace
.Height = IButtonHeight
.Left = ISpace
End With
15 With optSequence(0)
.Top = ISpace
.Height = IButtonHeight
.Left = cmdLoad.Left + cmdLoad.Width + ISpace + 200
End With
20 With optSequence(1)
.Top = ISpace
.Height = IButtonHeight
.Left = optSequence(0).Left + optSequence(0).Width
End With
25 With txtFormatted
.Top = cmdLoad.Height + 2 * ISpace
.Left = ISpace
.Width = Me.ScaleWidth
.Height = Me.ScaleHeight - txtFomatted.Top
30 End With
End Sub
Private Sub optSequence_Click(Index As Integer)
Call FillSequence
End Sub
35 Name: frmOutPut.frm
Code:
VERSION 5.00

```

```
Object = "{3B7C8863-D78F-101B-B9B5-04021C009402}# 1.2# 0"; "RICHTX32.OCX"
Begin VB.Form frmOutPut
Caption = "OutPut"
ClientHeight = 9120
5 ClientLeft = 60
ClientTop = 345
ClientWidth = 10215
Icon = "frmOutPut.frx":0000
LinkTopic = "Form1"
10 MDIChild = -1 True
ScaleHeight = 9120
ScaleWidth = 10215
Begin RichTextLib.RichTextBox txtOptimized
Height = 2655
15 Left = 0
TabIndex = 0
Top = 480
Width = 9855
_ExtemX = 17383
20 _ExtentY = 4683
_Version = 393217
Enabled = -1 True
ScrollBars = 2
DisableNoScroll = -1 True
25 TextRTF = $"frmOutPut.frx":0442
BeginPropertyFont {0BE35203-8F91-11CE-9DE3-00AA004BB851}
Name = "Fixedsys"
Size = 9
Charset = 0
30 Weight = 400
Underline = 0 'False
Italic = 0 'False
Strikethrough = 0 'False.
EndProperty
35 End
End
Attribute VB_Name = "frmOutPut"
```

```
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = False
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
5 Option Explicit
Private Sub Form_Activate()
Me.ZOrder
Call InitControls
If gbSequenceChanged = True Then
10 Call BuildAminoChain(gsFormattedCode, Falte)
End If
gbSequenceChanged = False
txtOptimized.Text = LCase(ReverseTranscription)
End Sub
15 Private Sub InitControls()
Const ISpace As Long = 40
With txtOptimized
.Top = ISpace
.Left = ISpace
20 .Width = Me.ScaleWidth + ISpace
.Height = Me.ScaleHeight + ISpace
End With
End Sub
Name: frmDisplay.frm
25 Code:
VERSION 5.00
Object = "**\ACurevac_Genetic_Controls.vbp"
Begin VB.Form frmDisplay
Caption = "Display"
30 ClientHeight = 8205
ClientLeft = 60
ClientTop = 345
ClientWidth = 8745
Icon = "frmDisplay.frx":0000
35 LinkTopic = "Form1"
MDIChild = -1 True
ScaleHeight = 8205
```

```
ScaleWidth = 8745
ShowInTaskbar = 0 'False
Begin VB.OptionButton optNucleotids
Caption = "Original"
5 Height = 255
Index = 0
Left = 0
Style = 1 'Graphical
TabIndex = 1
10 Top = 0
Width = 855
End
Begin VB.OptionButton optNucleotids
Caption = "Optimized"
15 Height = 255
Index = 1
Left = 840
Style = 1 'Graphical
TabIndex = 0
20 Top = 0
Value = -1 True
Width = 855
End
Begin Curevac_Genetic_Controls.Curevac_Amino Curevac_Amino1
25 Height = 495
Index = 0
Left = 240
Top = 360
Visible = 0 'False
30 Width 735
_ExtentX = 1296
_ExtentY = 873
End
End
35 Attribute VB_Name = "frmDisplay"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = False
```

```

Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
Option Explicit
Private Sub Form_Activate()
5  Me.ZOrder
   If gbSequenceChanged = True Then
   Call BuildAminoChain(gsFormattedCode, True)
   End If
   gbSequenceChanged = False
10 End Sub
   Private Sub optNucleotids_Click(Index As Integer)
   Dim bShowOriginal As Boolean
   If Index = 0 Then 'Original Clicked
   bShowOriginal = True
15 Else 'Optimized Clicked
   bShowOriginal = False
   End If
   Call RebuildChain(bShowOriginal)
   End Sub
20 Name: frmStatistics.frm
   Code:
   VERSION 5.00
   Begin VB.Form frmStatistics
   Caption = "Statistics"
25 ClientHeight = 6450
   ClientLeft = 60
   ClientTop = 345
   ClientWidth = 8595
   Icon = "frmStatistics.frx":0000
30 LinkTopic = "formy1"
   MDIChild = -1 True
   ScaleHeight = 6450
   ScaleWidth = 8595
   Begin VB.Frame Frame1
35 Caption = "original"
   Height = 2655
   Left = 120

```

```

TabIndex = 9
Top = 720
Width = 2175
Begin VB.Label lblAdenin
5  BorderStyle = 1 'Fixed Single
   Height = 495
   Left = 840
   TabIndex = 17
   Top = 240
10  Width = 1215
   End
   Begin VB.Label lblCytosin
   BorderStyle = 1 'Fixed Single
   Height = 495
15  Left = 840
   TabIndex = 16
   Top = 2040
   Width = 1215
   End
20  Begin VB.Label lblGuanin
   BorderStyle = 1 'Fixed Single
   Height = 495
   Left = 840
   TabIndex = 15
25  Top = 1440
   Width = 1215
   End
   Begin VB.Label lblThymin
   BorderStyle = 1 'Fixed Single
30  Height = 495
   Left = 840
   TabIndex = 14
   Top = 840
   Width = 1215
35  End
   Begin VB.Label Label14
   Caption = "Cytosin"

```

```
Height = 255
Left = 120
TabIndex = 13
Top = 2160
5 Width = 615
End
Begin VB.Label Label2
Caption = "Thymin"
Height = 255
10 Left = 120
TabIndex = 12
Top = 960
Width = 615
End
15 Begin VB.Label Label3
Caption = "Guanin"
Height = 255
Left = 120
TabIndex = 11
20 Top = 1560
Width = 615
End
Begin VB.Label Label1
Caption = "Adenin"
25 Height = 255
Left = 120
TabIndex = 10
Top = 360
Width = 615 End
30 End
Begin VB.Frame Frame2
Caption = "optimized"
Height = 2655
Left = 2520
35 TabIndex = 0
Top = 720
Width = 2175
```

```
Begin VB.Label Label4
Caption = "adenin"
Height = 255
Left = 120
5  TabIndex = 8
Top = 360
Width = 615
End
Begin VB.Label Label6
10  Caption = "Cytosin"
Height = 255
Left = 120
TabIndex = 7
Top = 2160
15  Width = 615
End
Begin VB.Label Label8
Caption = "Guanin"
Height = 255
20  Left = 120
TabIndex = 6
Top = 1560
Width = 615
End
25  Begin VB.Label Label12
Caption = "Thymin"
Height = 255
Left = 120
TabIndex = 5
30  Top = 960
Width = 615
End
Begin VB.Label lblOptAdenin
BorderStyle = 1 'Fixed Single
35  Height = 495
Left = 840
TabIndex = 4
```

```
Top = 240
Width = 1215
End
Begin VB.Label lblOptThymin
5  BorderStyle = 1 'Fixed Single
Height = 495
Left = 840
TabIndex = 3
Top = 840
10 Width = 1215
End
Begin VB.Label lblOptGuanin
BorderStyle = 1 'Fixed Single
Height = 495
15 Left = 840
TabIndex = 2
Top = 1440
Width = 1215
End
20 Begin VB.Label lblOptCytosin
BorderStyle = 1 'Fixed Single
Height = 495
Left = 840
TabIndex = 1
25 Top = 2040
Width = 1215
End
End
Begin VB.Label lblBases
30 BorderStyle = 1 'Fixed Single
Height = 495
Left = 1320
TabIndex = 19
Top = 120
35 Width = 1215
End
Begin VB.Label Label5
```

```
Caption = "Sum of bases"
Height = 255
Left = 120
TabIndex = 18
5 Top = 240
Width = 1095
End
End
Attribute VB_Name = "frmStatistics"
10 Attribute VB_GlobalNameSpace = False
Attribute VB_CReatable = False
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
Option Explicit
15 Private Sub Form_Activate()
Me.ZOrder
lblBases.Caption = CStr(glAminoCounter * 3)
lblAdenin.Caption = CStr(glAdenin)
lblThymin.Caption = CStr(glThymin)
20 lblGuanin.Caption = CStr(glGuanin)
lblCytosin.Caption = CStr(glCytosin)
Call CountBases(frmOutPut.txtOptimized.Text)
lblCptAdeninoCaption = CStr(glOptAdenin)
lblOptThymin.Caption = CStr(glOptThymin)
25 lblOptGuanin.Caption = CStr(glOptGuanin)
lblOptCytosin.Caption = CStr(glOptCytosin)
End Sub
Name: frmOptions.frm
Code:
30 VERSION 5.00
Begin VB.Form frmOptions
Caption = "Options"
ClientHeight = 3915
ClientLeft = 60
35 ClientTop = 345
ClientWidth = 5760
Icon = "frmOptions.frx":0000
```

```
LinkTopic = "Form1"
MDIChild = -1 True
ScaleHeight = 3915
ScaleWidth = 5760
5 End
Attribute VB_Name = "frmOptions"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = False
Attribute VB_PredeclaredId = True
10 Attribute VB_Exposed = False
Option Explicit
Name: frmAbout.frm
Code:
VERSION 5.00
15 Begin VB.Form frmAbout
BorderStyle = 3 'Fixed Dialog
Caption = "About Curevac RNA-Analyzer"
ClientHeight = 2490
ClientLeft = 2340
20 ClientTop = 1935
ClientWidth = 6195
ClipControls = 0 'False
LinkTopic = "Form2"
MaxButton = 0 False
25 MmButton = 0 'False
ScaleHeight = 1718.642
ScaleMode = 0 'User
ScaleWidth = 5817.425
ShowInTaskbar = 0 'False
30 Begin VB.PictureBox picIcon
AutoSize = -1 True
ClipControls = 0 False
Height = 1215
Left = 120
35 Picture = "fmtAbout.frx":0000
ScaleHeight = 811.195
ScaleMode = 0 'User
```

```
ScaleWidth = 2401.98
TabIndex = 1
Top = 120
Width = 3480
5 End
Begin VB.CommandButton cmdOK
Cancel = -1 True
Caption = "OK"
Default = -1 True
10 Height = 345
Left = 2520
TabIndex = 0
Top = 2040
Width = 1260
15 End
Begin VB.Label lblCompany
Caption = "lblCompany"
Height = 255
Left = 3720
20 TabIndex = 5
Top = 1200
Width = 2295
End
Begin VB.Line Line1
25 BorderColor = &H00808080&
BorderStyle = 6 'Inside Solid
Index = 1
X1 = 112.686
X2 = 5746.996
30 Y1 = 1325218
Y2 = 1325218
End
Begin VB.Label lblDescription
Caption = "App Description"
35 ForeColor = &H00000000&
Height = 330
Left - 120
```

```
    TabIndex = 2
    Top = 1560
    Width = 5925
    End
5  Begin VB.Label lblTitle
    Caption = "Application Title"
    ForeColor = &H00000000&
    Height = 360
    Left = 3720
10  TabIndex = 3
    Top = 240
    Width = 2325
    End
    Begin VB.Line Line1
15  BorderColor = &H00FFFFFF&
    BorderWidth = 2
    Index = 0
    X1 = 112.686
    X2 = 5746.996
20  Y1 = 1325218
    Y2 = 1325218
    End
    Begin VB.Label lblVersion
    Caption = "Version"
25  Height = 225
    Left = 3720
    TabIndex = 4
    Top = 720
    Width = 2325
30  End
    End
    Attribute VB_Name = "frmAbout"
    Attribute VB_GlobalNameSpace = False
    Attribute VB_Creatable = False
35  Attribute VB_PredeclaredId = True
    Attribute VB_Exposed = False
    Option Explicit
```

```

Private Sub cmdOK_Click()
Unload Me
End Sub
Private Sub Form_Load()
5 Me.Caption = "About " & App.Title
lblVersion.Caption = Version " & App.Major & "." & App.Minor & "." & App.Revision
lblTitle.Caption = App.Title
lblDescription.Caption = App.FileDescription
lblCompany.Caption = App.CompanyName
10 End Sub

```

LISTADO DE SECUENCIAS

```

<110> CureVac GmbH
<120> Composición farmacéutica que contiene un ARNm estabilizado y optimizado para la traducción en sus campos
de codificación.
15 <130> CU01P003WOEPT4
<140> PCT/EP02/06180
<141> 2002-06-05
<160> 13
<170> PatentIn Ver. 2.1
20 <210> 1
<211> 774
<212> ADN
<213> Influenza virus
<220>
25 <223> Influenza-Matrix: Gen del tipo salvaje (para comparación)
<220>
<223> Codón de inicio: atg (Nucleótidos 11 a 13), codón de parada: tga (Nucleótidos 767 a 769)
agatctaaag atgagtcctc taaccgaggt cgaaacgtac gttctctcta tcatcccgtc 60
aggccccctc aaagccgaga tcgcacagag acttgaagat gtctttgcag ggaagaacac 120
cgatcttgag gttctcatgg aatggctaaa gacaagacca atcctgtcac ctctgactaa 180
ggggatttta ggatttgtgt tcacgctcac cgtgcccagt gagcgaggac tgcagcgtag 240
acgctttgtc caaaatgcc ttaatgggaa cggggatcca aataacatgg acaaagcagt 300
taaaactgtat aggaagctca agagggagat aacattccat ggggccaag aaatctcact 360
cagttattct gctggtgcac ttgccagttg tatgggcctc atatacaaca ggatgggggc 420
tgtgaccact gaagtggcat ttggcctggt atgtgcaacc tgtgaacaga ttgctgactc 480
ccagcatcgg tctcataggc aaatggtgac aacaaccaac ccactaatca gacatgagaa 540
cagaatgggt ttagccagca ctacagctaa ggctatggag caaatggctg gatcagagtga 600
gcaagcagca gaggccatgg aggttgctag tcaggctagg caaatgggtgc aagcgatgag 660
aaccattggg actcatccta gctccagtg cgggtctgaaa aatgatcttc ttgaaaattt 720
gcaggcctat cagaaaacgaa tgggggtgca gatgcaacgg ttcaagtgaa ctag 774
<210> 2
30 <211> 252

```

<212> PRT

<213> virus de la gripe

>400> 2

```

Met Ser Leu Leu Thr Glu Val Glu Thr Tyr Val Leu Ser Ile Ile Pro
  1          5          10          15
Ser Gly Pro Leu Lys Ala Glu Ile Ala Gln Arg Leu Glu Asp Val Phe
      20          25          30
Ala Gly Lys Asn Thr Asp Leu Glu Val Leu Met Glu Trp Leu Lys Thr
      35          40          45
Arg Pro Ile Leu Ser Pro Leu Thr Lys Gly Ile Leu Gly Phe Val Phe
      50          55          60
Thr Leu Thr Val Pro Ser Glu Arg Gly Leu Gln Arg Arg Arg Phe Val
  65          70          75          80
Gln Asn Ala Leu Asn Gly Asn Gly Asp Pro Asn Asn Met Asp Lys Ala
      85          90          95
Val Lys Leu Tyr Arg Lys Leu Lys Arg Glu Ile Thr Phe His Gly Ala
      100          105          110
Lys Glu Ile Ser Leu Ser Tyr Ser Ala Gly Ala Leu Ala Ser Cys Met
      115          120          125
Gly Leu Ile Tyr Asn Arg Met Gly Ala Val Thr Thr Glu Val Ala Phe
      130          135          140
Gly Leu Val Cys Ala Thr Cys Glu Gln Ile Ala Asp Ser Gln His Arg
      145          150          155          160
Ser His Arg Gln Met Val Thr Thr Thr Asn Pro Leu Ile Arg His Glu
      165          170          175
Asn Arg Met Val Leu Ala Ser Thr Thr Ala Lys Ala Met Glu Gln Met
      180          185          190
Ala Gly Ser Ser Glu Gln Ala Ala Glu Ala Met Glu Val Ala Ser Gln
      195          200          205
Ala Arg Gln Met Val Gln Ala Met Arg Thr Ile Gly Thr His Pro Ser
      210          215          220
Ser Ser Ala Gly Leu Lys Asn Asp Leu Leu Glu Asn Leu Gln Ala Tyr
      225          230          235          240
Gln Lys Arg Met Gly Val Gln Met Gln Arg Phe Lys
      245          250

```

5

<210> 3

<211> 775

<212> ADN

<213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Gripe-Matrix: Gen con un mayor contenido de G/C

<220>

<223> Codón de inicio: atg (Nucleótidos 11 a 13), codón de parada: tag (Nucleótidos 767 a 769)

5 <400> 3

```

agatctaaag atgagcctgc tgaccgaggt ggagacctac gtgctgagca tcatccccag 60
cgccccctg aaggccgaga tgcgccagag gctggaggac gtgttcgccg gcaagaacac 120
cgacctggag gtgctgatgg agtggctgaa gaccaggccc atcctgagcc ccctgaccaa 180
gggcatcctg ggcttcgtgt tcaccctgac cgtgcccagc gagcgcggcc tgcagcgccc 240
ccgcttcgtg cagaacgccc tgaacggcaa cggcgacccc aacaacatgg acaaggccgt 300
gaagctgtac aggaagctga agagggagat caccttccac ggcgccaagg agatcagcct 360
gagctacagc gccggcgccc tggccagctg catgggctgt atctacaaca ggatgggctc 420
cgtgaccacc gaggtggcct tcggcctggt gtgcgccacc tgcgagcaga tcgccgacag 480
ccagcaccgc agccacaggc agatggtgac caccaccaac cccctgatca ggcacgagaa 540
caggatggtg ctggccagca ccaccgcaa ggccatggag cagatggccg gcagcagcga 600
gcaggccgcc gaggccatgg aggtggccag ccaggccagg cagatggtgc aggccatgag 660
gaccatcggc acccacccca gcagcagcgc cggcctgaag aacgacctgc tggagaacct 720
gcaggcctac cagaagcgca tgggcgtgca gatgcagcgc tcaagtgaac ctagt 775

```

<210> 4

<211> 844

<212> ADN

10 <213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Gripe-Matrix: Gen para forma segregada (con secuencia de señal N-terminal) con un mayor contenido de G/C

<220>

15 <223> Codón de inicio: atg (Nucleótidos 11 a 13), codón de parada tga (Nucleótidos 836 a 838)

<400> 4

```

agatctaaag atggccgtca tggcccccg caccctggtg ctgctgctga gcggcgccct 60
ggcctgacc cagacctggg ctagcctgct gaccgaggtg gagacctacg tgctgagcat 120
catccccagc ggccccctga aggccgagat cgcccagagg ctggaggacg tgttcgccgg 180
caagaacacc gacctggagg tgctgatgga gtggctgaag accaggcca tcctgagccc 240
cctgaccaag ggcatcctgg gcttcgtgtt caccctgacc gtgccagcgc agcgcggcct 300
gcagcgccgc cgcttcgtgc agaacgcctt gaacggcaac ggcgaccca acaacatgga 360
caaggccgtg aagctgtaca ggaagctgaa gagggagatc accttccagc gcgccaagga 420
gatcagcctg agctacagcg ccggcgccct ggccagctgc atgggcctga tctacaacag 480
gatgggcccgtg gtgaccaccg aggtggcctt cggcctggtg tgcgccacct gcgagcagat 540
cgccgacagc cagcaccgca gccacaggca gatggtgacc accaccaacc ccctgatcag 600
gcacgagaac aggatggtgc tggccagcac caccgccaag gccatggagc agatggccgg 660
cagcagcgag caggccgccc aggccatgga ggtggccagc caggccaggc agatggtgca 720
ggccatgagg accatcggca cccaccccag cagcagcgcg ggcctgaaga acgacctgct 780
ggagaacctg caggccctacc agaagcgcac gggcgtgagc atgcagcgc tcaagtgaac 840
tagt 844

```

<210> 5

<211> 942

20 <212> ARN

<213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Gripe-Matrix: ARNm con secuencias de estabilización

<220>

<223> Las secuencias de estabilización provienen de los UTRs de 5' o 3' del ARNm de β -globina de *Xenopus laevis*

<220>

5 <223> Codón de inicio: aug (Nucleótidos 56 a 58), Codón de parada: uga (Nucleótidos 812 a 814)

<400> 5

```

gcuuguucuu uuugcagaag cucagaauaa acgcucaacu uuggcagauc uaaagaugag 60
ucuucuaacc gaggucgaaa cguacguucu cucuaucac cgcucaggcc ccucuaaagc 120
cgagaucgca cagagacuug aagaugucuu ugcagggag aacaccgauc uugagguucu 180
cauggaaugg cuaaagacaa gaccaauccu gucaccucug acuaagggga uuuuaggauu 240
uguguucacg cucaccgugc ccagugagcg aggacugcag cguagacgcu uuguccaaaa 300
ugcccuuaau gggaacgggg auccaaaaua cauggacaaa gcaguuaaac uguauaggaa 360
gcucaagagg gagauaacau uccauggggc caaagaaauc ucacucaguu auucugcugg 420
ugcacuugcc aguuguauug gccucauaua caacaggaug ggggcuguga ccacugaagu 480
ggcauuugc cugguauugc caaccuguga acagauugcu gacucccagc aucggucuca 540
uaggcaaacg gugacaacaa ccaaccacu aaucagacau gagaacagaa ugguuuagc 600
cagcacuaca gcuaaggcua uggagcaauu ggcuggaucg agugagcaag cagcagaggc 660

cauggagguu gcuagucagg cuaggcaauu ggugcaagcg augagaacca uugggacuca 720
uccuagcucc agugcugguc ugaaaaauga ucuucuugaa aaauugcagg ccuaucaaga 780
acgaaugggg gugcagaugc aacgguucaa gugaacuagu gacugacuag cccgcugggc 840
cucccaacgg gccuccucc ccuccuugca ccaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 900
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa 942
    
```

<210> 6

10 <211> 942

<212> ARN

<213> Secuencia Artificial

<220>

15 <223> Descripción de Secuencia Artificial: Gripe-Matrix: ARNm con un mayor contenido de G/C y secuencias de estabilización

<220>

<223> Las secuencias de estabilización provienen de los UTRs de 5' o 3' del ARNm de β -globina de *Xenopus laevis*

<220>

<223> Codón de inicio: aug (Nucleótidos 56 a 58), Codón de parada: uga (Nucleótidos 812 a 814)

20 <400> 6

```

gcuuguucuu uuugcagaag cucagaauaa acgcucaacu uuggcagauc uaaagaugag 60
ccugcugacc gagguggaga ccuacgugcu gagcauacac cccagcggcc ccuugaaggc 120
cgagaucgca cagaggcugc aggacguguu cgccggcaag aacaccgacc uggaggugcu 180
gauggagugg cugaagacca ggcccauccu gagccccug accaagggca uccugggcuu 240
cguguucacc cugaccgugc ccagcgagcg cggccugcag cgccgccgcu ucgugcagaa 300
cgcccugaac ggcaacggcg accccaacaa cauggacaag gccgugaagc uguacaggaa 360
gcugaagagg gagaucaccu uccacggcgc caaggagauc agccugagcu acagcgccgg 420
cgcccuggcc agcugcaugg gccugaucua caacaggaug ggcgccguga ccaccgaggu 480
ggccuucggc cuggugugcg ccaccugcga gcagaucgcc gacagccagc accgcagcca 540
cagcagagc gugaccacca ccaacccccu gaucaggcac gagaacagga uggugcuggc 600
cagcaccacc gccaaaggcca uggagcagau ggccggcagc agcgagcagg ccgcccaggc 660
cauggaggug gccagccagg ccaggcagau ggugcaggcc augaggacca ucggcaccca 720
cccagcagc agcgccggcc ugaagaacga ccugcuggag aaccugcagg ccuaccagaa 780
gcgcauuggc gugcagaugc agcgcuucua gugaacuagu gacugacuag cccgcugggc 840
cucccaacgg gccuccucc ccuccuugca ccaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 900
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa 942
    
```

<210> 7

<211> 1011

<212> ARN

<213> Secuencia Artificial

5 <220>

<223> Descripción de Secuencia Artificial: Gripe-Matrix: para ARNm con mayor contenido de G/C y secuencias de estabilización que codifican la forma segregada

<220>

<223> Codón de inicio: aug (Nucleótidos 56 a 58), Codón de parada: uga (Nucleótidos 881 a 883)

10 <400> 7

```

gcuguuucuu uuugcagaag cucagaauaa acgcucaacu uuggcagauc uaaagauggc 60
cgucauggcc ccccgacccc uggugcugcu gcugagcggc gcccuggccc ugacccagac 120
cugggccagc cugcugaccg agguggagac cuacgugcug agcaucaucc ccagcggccc 180
ccugaaggcc gagaucgccc agaggcugga ggacguguuc gccggcaaga acaccgaccu 240

ggaggugcug auggaguggc ugaagaccag gcccauccug agccccuga ccaagggcau 300
ccugggcuuc guguucaccc ugaccgugcc cagcgagcgc ggccugcagc gccgcccguu 360
cgugcagaac gcccugaacg gcaacggcga cccaacaac auggacaagg ccgugaagcu 420
guacaggaag cugaagaggg agaucaccuu ccacggcgcc aaggagauca gccugagcua 480
cagcgcgggc gcccuggcca gcugcauggg ccugaucuac aacaggauug gcgcccugac 540
caccgaggug gccuucggcc uggugugcgc caccugcgag cagaucgccg acagccagca 600
ccgcagccac aggcagaugg ugaccaccac caacccccug aucaggcacg agaacaggau 660
ggugcuggcc agcaccaccg ccaaggccau ggagcagaug gccggcagca gcgagcaggc 720
cgccgaggcc auggaggugg ccagccaggc caggcagaug gugcaggcca ugaggaccu 780
cggcaccac cccagcagca gcgcccggcc gaagaacgac cugcuggaga accugcaggc 840
cuaccagaag cgcaugggcg ugcagaugca gcgcuucaag ugaacuagug acugacuagc 900
ccgcugggccc ucccaacggg ccucuccccc cuccuugcac caaaaaaaaa aaaaaaaaaa 960
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa a 1011

```

<210> 8

<211> 940

15 <212> ADN

<213> Homo Sapiens

<220>

<223> MAGE1: Gen de tipo salvaje (para comparación)

<220>

20 <223> Codón de inicio: atg (Nucleótidos 5 a 7), Codón de parada: tga (Nucleótidos 932 a 934)

<400> 8

```

catcatgtct cttgagcaga ggagtctgca ctgcaagcct gaggaagccc ttgaggccca 60
acaagaggcc ctgggcctgg tgtgtgtgca ggctgccacc tcctcctcct ctctctgtgt 120
cctgggcacc ctggaggagg tgcccactgc tgggtcaaca gatcctcccc agagtcctca 180
gggagcctcc gcctttccca ctaccatcaa ctccactcga cagaggcaac ccagtgaggg 240
ttccagcagc cgtgaagagg aggggccaaag cacctcttgt atcctggagt ccttgttccg 300
agcagtaatc actaagaagg tggctgattt ggttggtttt ctgctcctca aatatcgagc 360
cagggagcca gtcacaaagg cagaaatgct ggagagtgtc atcaaaaatt acaagcactg 420
ttttcctgag atcttcggca aagcctctga gtccttgag ctgggtcttg gcattgacgt 480
gaaggaagca gacccaccg gccactccta tgtccttgtc acctgcctag gtctctccta 540
tgatggcctg ctgggtgata atcagatcat gcccaagaca ggcttcctga taattgtcct 600
ggtcatgatt gcaatggagg gcggccatgc tcctgaggag gaaatctggg aggagctgag 660
tgtgatggag gtgatgatg ggagggagca cagtgcctat ggggagccca ggaagctgct 720
cacccaagat ttggtgcagg aaaagtacct ggagtaccgg caggtgccgg acagtgatcc 780
cgcacgctat gagttcctgt ggggtccaag ggccctcgct gaaaccagct atgtgaaagt 840
ccttgagtat gtgatcaagg tcagtgcaag agttcgcttt ttcttcccat ccctgcgtga 900
agcagctttg agagaggagg aagagggagt ctgagcatga 940

```

<210> 9

<211> 308

<212> PRT

5 <213> Homo Sapiens

<220>

<223> Antígeno de tumor MAGE1: secuencia de proteínas

<400> 9

```

Ser Leu Glu Gln Arg Ser Leu His Cys Lys Pro Glu Glu Ala Leu Glu
 1           5           10           15
Ala Gln Gln Glu Ala Leu Gly Leu Val Cys Val Gln Ala Ala Thr Ser
          20           25           30
Ser Ser Ser Pro Leu Val Leu Gly Thr Leu Glu Glu Val Pro Thr Ala
 35           40           45

```

Gly Ser Thr Asp Pro Pro Gln Ser Pro Gln Gly Ala Ser Ala Phe Pro
 50 55 60

Thr Thr Ile Asn Phe Thr Arg Gln Arg Gln Pro Ser Glu Gly Ser Ser
 65 70 75 80

Ser Arg Glu Glu Glu Gly Pro Ser Thr Ser Cys Ile Leu Glu Ser Leu
 85 90 95

Phe Arg Ala Val Ile Thr Lys Lys Val Ala Asp Leu Val Gly Phe Leu
 100 105 110

Leu Leu Lys Tyr Arg Ala Arg Glu Pro Val Thr Lys Ala Glu Met Leu
 115 120 125

Glu Ser Val Ile Lys Asn Tyr Lys His Cys Phe Pro Glu Ile Phe Gly
 130 135 140

Lys Ala Ser Glu Ser Leu Gln Leu Val Phe Gly Ile Asp Val Lys Glu
 145 150 155 160

Ala Asp Pro Thr Gly His Ser Tyr Val Leu Val Thr Cys Leu Gly Leu
 165 170 175

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

Phe Leu Ile Ile Val Leu Val Met Ile Ala Met Glu Gly Gly His Ala
 195 200 205

Pro Glu Glu Glu Ile Trp Glu Glu Leu Ser Val Met Glu Val Tyr Asp
 210 215 220

Gly Arg Glu His Ser Ala Tyr Gly Glu Pro Arg Lys Leu Leu Thr Gln
 225 230 235 240

Asp Leu Val Gln Glu Lys Tyr Leu Glu Tyr Arg Gln Val Pro Asp Ser
 245 250 255

Asp Pro Ala Arg Tyr Glu Phe Leu Trp Gly Pro Arg Ala Leu Ala Glu
 260 265 270

Thr Ser Tyr Val Lys Val Leu Glu Tyr Val Ile Lys Val Ser Ala Arg
 275 280 285

Val Arg Phe Phe Phe Pro Ser Leu Arg Glu Ala Ala Leu Arg Glu Glu
 290 295 300

Glu Glu Gly Val
 305

<210> 10

<211> 939

<212> ARN

5 <213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: MAGE1: ARNm con un mayor contenido de G/C

<220>

<223> Codón de inicio: aug (Nucleótidos 1 a 3), codón de parada: uga (Nucleótidos 937 a 939)

<400> 10

```

augagccugg agcagcgcag ccugcacugc aagccggagg aggcgcugga ggcgcagcag 60
gagggcgcugg gccuggucug cguccaggcg gcgacgagca gcagcagccc gcugguccug 120
ggcacgcugg aggagguccc gacggcgggc agcacggacc cgccgcagag cccgcagggc 180
gcgagcgcgu ucccgcgcag gaucaacuuc acgcgccagc gccagccgag cgagggcagc 240
agcagccgcu aggaggagg cccgagcacg agcugcaucc uggagagccu guuccgcgcg 300
gucaucacga agaaggucgc ggaccugguc ggcuuccugc ugcugaagua ccgcgcgcgc 360
gagccgguca cgaaggcggg gaugcuggag agcgucauca agaacuacia gcacugcuuc 420
ccggagaucu ucggcaaggc gagcgagagc cugcagcugg ucuucggcau cgacgucaag 480
gagggcgacc cgacggggcca cagcuacguc cuggucacgu gccugggccu gagcuacgac 540
ggccugcugg gcgacaacca gaucaugccc aagaccggcu uccugaucuau cguccugguc 600
augaucgcga uggaggggcg ccacgcgccg gaggaggaga ucuggggagg gcugagcguc 660
auggaggucu acgacggccg cgagcacagc agcugacggc agcccgcaa gcugcugacg 720
caggaccugg uccaggagaa guaccuggag uaccgccagg ucccggacag cgacccggcg 780
cgcuacgagu uccouguggg cccgcgcgcg cuggcggaga cgagcuacgu caagguccug 840
gaguacguca ucaaggucag cgcgcgcguc cgcuucuucu ucccagagccu gcgcgaggcg 900
gcgcugcgcg aggaggagga gggcgucuga gcgugauga 939
    
```

5 <210> 11

<211> 939

<212> ARN

<213> Secuencia Artificial

<220>

10 <223> Descripción de Secuencia Artificial: MAGE1: ARNm con utilización alternativa de codón.

<220>

<223> Codón de inicio: aug (Nucleótidos 1 a 3), codón de parada: uga (Nucleótidos 937 a 939)

<400> 11

```

augagccugg agcagcgcag ccugcacugc aagcccagg aggccugga ggcccagcag 60
gagggccugg gccuggugug cgugcaggcc gccaccagca gcagcagccc ccuggugcug 120
ggcaccugg aggaggugcc caccgccggc agcaccgacc cccccagag cccccagggc 180
gccagcgcgu ucccaccac caucaacuuc acccgccagc gccagcccag cgagggcagc 240
agcagccgcu aggaggagg ccccagcacc agcugcaucc uggagagccu guuccgcgcc 300
gugaucacca agaagguggc cgaccuggug ggcuuccugc ugcugaagua ccgcgcccgc 360
gagcccguga ccaaggccga gaugcuggag agcgugauc agaacuacia gcacugcuuc 420
cccgagaucu ucggcaaggc cagcgagagc cugcagcugg uguucggcau cgacgugaag 480
gaggccgacc ccaccggcca cagcuacgug cuggugaccu gccugggccu gagcuacgac 540
ggccugcugg gcgacaacca gaucaugccc aagaccggcu uccugaucuau cgugcuggug 600
augaucgcca uggaggggcg ccacgcccc gaggaggaga ucuggggagg gcugagcgug 660
auggaggugu acgacggccg cgagcacagc gccuacggcg agccccgaa gcugcugacc 720
caggaccugg ugcaggagaa guaccuggag uaccgccagg ugcccagacag cgaccccgcc 780
cgcuacgagu uccuguggg ccccgcgcc cuggccgaga ccagcuacgu gaaggugcug 840
gaguacguga ucaaggugag cgcccgcgug cgcuucuucu ucccagagccu gcgcgaggcc 900
ggccugcgcg aggaggagga gggcguguga gccugauga 939
    
```

15 <210> 12

<211> 7

<212> ARN

<213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Un motivo de secuencia reconocible para una endonucleasa, motivo de secuencia que está contenido en el segmento UTR 3' del gen que codifica el receptor de transferina (p. 10 de la descripción)

<400> 12

5 gaacaag7

<210> 13

<211> 13

<212> ARN

<213> Secuencia Artificial

10 <220>

<223> Descripción de Secuencia Artificial: Secuencia Kozak, sitio de enlace de ribosomas (p. 12 de la descripción)

<400> 13

gccgccacca ugg 13

REIVINDICACIONES

1. ARNm modificado que codifica como mínimo un péptido o un polipéptido viral antígeno, caracterizado porque el contenido de G/C del campo del ARNm modificado que codifica el péptido o polipéptido es mayor que el contenido de G/C del campo de codificación del ARNm del tipo salvaje que codifica el péptido o polipéptido, y la secuencia del aminoácido codificada permanece sin modificar frente al tipo salvaje.
2. ARNm modificado según la reivindicación 1, caracterizado porque el contenido de G/C del campo del ARNm modificado que codifica el péptido o polipéptido es como mínimo un 7%, preferentemente como mínimo un 15%, mayor que el contenido de G/C del campo de codificación del ARNm del tipo salvaje que codifica el péptido o polipéptido.
3. ARNm modificado según una de las reivindicaciones 1 a 2, caracterizado porque el ARNm modificado presenta una estructura cap en 5' y/o una cola poliA de como mínimo 70 nucleótidos y/o un IRES y/o una secuencia de estabilización de 5' y/o una secuencia de estabilización de 3'.
4. ARNm modificado según una de las reivindicaciones 1 a 3, caracterizado porque el ARNm modificado tiene como mínimo un nucleótido análogo de procedencia natural.
5. ARNm modificado según la reivindicación 4, caracterizado porque el análogo se selecciona de entre el grupo compuesto por fosfotioatos, fosfoamidatos, nucleótidos peptídicos, metilfosfonatos, 7-deazaguanosina, 5-metilcitosina e inosina.
6. ARNm modificado según cualquiera de las reivindicaciones 1 a 5, caracterizado porque el antígeno viral proviene de la forma segregada de un antígeno de superficie.
7. ARNm modificado según cualquiera de las reivindicaciones 1 a 5, caracterizado porque el ARNm codifica un antígeno de superficie de gérmenes virales patógenos.
8. ARNm modificado según cualquiera de las reivindicaciones 1 a 7, caracterizado porque el ARNm está asociado o unido a una proteína o un péptido catiónico.
9. ARNm modificado según la reivindicación 8, caracterizado porque la proteína o el péptido catiónico se selecciona de entre el grupo consistente en protamina, poli-L-lisina e histonas.
10. ARNm modificado según cualquiera de las reivindicaciones 1 a 9, caracterizado porque el polipéptido es un poliepitope de antígenos virales.
11. ARNm modificado según cualquiera de las reivindicaciones 1 a 10, caracterizado porque el ARNm modificado es un ARNm multicistrónico.
12. ARNm modificado según cualquiera de las reivindicaciones 1 a 11, caracterizado porque el ARNm codifica además al menos una citoquina.
13. ARNm modificado según cualquiera de las reivindicaciones 1 a 12, caracterizado porque el ARN multicistrónico presenta más de una secuencia IRES, seleccionándose las secuencias IRES en especial entre picornavirus (por ejemplo FMDV), virus de la peste (CFFV), virus de la poliomielitis (PV), virus de encefalo-miocarditis (ECMV), virus de la fiebre aftosa (FMDV), virus de la hepatitis C (HCV), virus clásicos de la fiebre porcina (CSFV), virus de leucemia murina (MLV), virus de la inmunodeficiencia del simio (SIV) o virus de la parálisis de cricket (CrPV).
14. Composición farmacéutica caracterizada porque contiene el ARNm modificado según cualquiera de las reivindicaciones 1 a 13 junto con un excipiente y/o vehículo farmacéuticamente compatible.
15. Composición farmacéutica según la reivindicación 14, caracterizada porque contiene como mínimo un adyuvante que estimula la reacción inmunológica.
16. Composición farmacéutica según cualquiera de las reivindicaciones 14 ó 15, que contiene como mínimo además una citoquina.
17. Utilización de una composición farmacéutica según cualquiera de las reivindicaciones 14 a 16 o de un ARNm modificado según cualquiera de las reivindicaciones 1 a 13 para la preparación de una vacuna para la vacunación contra enfermedades infecciosas virales.
18. Utilización de una composición farmacéutica según la reivindicación 17 para la preparación de una vacuna para la vacunación contra SIDA, hepatitis A, B ó C, herpes, herpes Zoster, rubéola, dengue, enfermedades infecciosas hemorrágicas, fiebre amarilla y gripe.

Fig. 1A

Matriz de la gripe: Gen del tipo salvaje (para comparación)

agatctaaagatgagtccttctaaccgagggtcgaaacgtacgttctctcta
 tcatcccgtcaggccccctcaaagccgagatcgcacagagacttgaagat
 gtctttgcaggaagaacaccgatcttgagggttctcatggaatggctaaa
 gacaagaccaatcctgtcacctctgactaaggggattttaggatttgtgt
 tcacgctcaccgtgccagtgagcgaggactgcagcgtagacgctttgtc
 caaaatgcccttaatgggaacggggatccaaataacatggacaaagcagt
 taaactgtataggaagctcaagagggagataacattccatggggccaaag
 aatctcactcagttattctgctggcacttgccagttgtatgggcctc
 atatacaacaggatgggggctgtgaccactgaagtggcatttggcctggt
 atgtgcaacctgtgaacagattgctgactcccagcatcggctctcataggc
 aatggtgacaacaaccaaccactaatcagacatgagaacagaatggtt
 ttagccagcactacagctaaggctatggagcaaatggctggatcgagtga
 gcaagcagcagaggccatggagggttgctagtcaggctaggcaaatggtgc
 aagcgatgagaaccattgggactcatcctagctccagtgctggtctgaaa
 atgatcttcttgaaaatttgaggcctatcagaaacgaatgggggtgca
 gatgcaacggttcaagtgaactag

Fig. 1B

Matriz de la gripe: Secuencia de proteínas

MSLLTEVETYVLSIIPSGPLKAEIAQRLEDVFAGKNTDLEVLMEWLKTRP
 ILSPLTKGILGFVFTLTVPSEGLQRRRFVQNALNGNGDPNNMDKAVKLY
 RCLKREITFHGAKEISLSYSAGALASCMGLIYNRMGAVTTEVAFGLVCAT
 CEQIADSQHRSHRQMVTTTNPLIRHENRMVLASTTAKAMEQMAGSSEQAA
 EAMEVASQARQMVQAMRTIGTHPSSSAGLKNLLENLQAYQKRMGVQMR
 FK*

Fig. 1C

Matriz de la gripe: Gen con contenido de G/C aumentado

agatctaaagatgagCctGctGaccgaggtGgaGacCtacgtGctGAGCa
 tcatcccCAGCggccccctGaaGgccgagatcgcCcagagGctGgaGgaC
 gtGttCgcCggCaagaacaccgaCctGgaggtGctGatggaGtggctGaa
 gacCagGccCatcctgAGCccCctgacCaagggCatCCTGggCttCgtgt
 tcacCctGaccgtgcccagCgagcgCggCctgcagcgCCGCcgcttCgtG
 caGaaCgccctGaaCggCaacggCgaCccCaaCaacatggacaaGgcCgt
 GaaGctgtaCaggaagctGaagagggagatCacCttccaCggCgccaGg
 aGatcAGCctGagCtaCAGCgcCggCgcCctGgccagCtgCatgggcctG
 atCtacaacaggatgggCgcCgtgaccacCgaGgtggcCttCggcctggt
 GtgCgcCacctgCgaGcagatCgcCgacAGCcaGcaCcgCAGCcaCaggc
 aGatgggtgacCacCaccaacccCctGatcagGcaCgagaacagGatgggG
 CTGgccagcacCacCgcCaagggCatggagcaGatggcCggCAGCaGCga
 gcaGgcCgcCgaggccatggaggtGgcCagCcagggCaggcaGatgggtgc
 aGgcCatgagGaccatCggCacCcaCccCagcAGCagCgcCggCctgaaG
 aaCgaCctGctGgaGaaCCTGcagggcctaCagaaGcgCatgggCgtgca
 gatgcaGcgCttcaagtgaactagt

Fig. 1D

Matriz de la gripe: Gen para forma segregada (secuencia de señal terminal de N) con un contenido de G/C aumentado

AgatctaaagatgGCCGTcATGGCCCCCGCACCCCTGGTGCTGCTGCTGA
 GCGGCGCCCTGGCCCTGACCCAGACCTGGGCTagCctGctGaccgaggtG
 gaGacCtacgtGctGAGCatcatcccCAGCggccccctGaaGgccgagat
 cgcCcagagGctGgaGgaCgtGttCgcCggCaagaacaccgaCctGgagg
 tGctGatggaGtggctGaagacCagGccCatcctgAGCccCctgacCaag
 ggCatCCTGggCttCgtgttcacCctGaccgtgcccagCgagcgCggCct
 gcagcgCCGCcgcttCgtGcaGaaCgccctGaaCggCaacggCgaCccCa
 aCaacatggacaaGgcCgtGaaGctgtaCaggaagctGaagagggagatC
 acCttccaCggCgccaGgaGatcAGCctGagCtaCAGCgcCggCgcCct
 GgccagCtgCatgggcctGatCtacaacaggatgggCgcCgtgaccacCg
 aGgtggcCttCggcctggtGtgCgcCacctgCgaGcagatCgcCgacAGC
 cagcaCcgCAGCcaCaggcaGatgggtgacCacCaccaacccCctGatcag
 GcaCgagaacagGatgggGCTGgccagcacCacCgcCaagggCatggagc
 aGatggcCggCAGCaGCgagcaGgcCgcCgaggccatggaggtGgcCagC
 cagggCaggcaGatgggtgcaGgcCatgagGaccatCggCacCcaCccCag
 cAGCagCgcCggCctgaaGaaCgaCctGctGgaGaaCCTGcagggcctaCc
 agaaGcgCatgggCgtgcagatgcaGcgCttcaagtgaactagt

Fig. 1E

Matriz de la gripe: ARNm con secuencias de estabilización

GCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUCAACUUUGGCagauc
 uaaagaugagucucuaaaccgaggucgaaacguacguucucucuauc
 ccgucaggcccccuaaagccgagaucgcacagagacuugaagaugucuu
 ugcagggagaacaaccgaucuuagagguucucauggaauggcuaaagaca
 gaccaauccugucaccucugacuaaggggauuuuaggauuuguguuca
 cucaccgugcccagugagcggaggacugcagcguagacgcuuuguccaaa
 ugcccuuaaugggaacggggauccaaauaacauggacaaagcaguuaaac
 úguauaggaagcucaagagggagauaacauuccauggggccaaagaauc
 ucacucaguuaauucugcugggucacuugccaguuguauggggccucaua
 caacaggaugggggucugugaccacugaaguggcauuuggccugguaug
 caaccugugaacagauugcugacuccagcaucggucucauaggcaaaug
 gugacaacaaccaaccacuaaucagacaugagaacagaaugguuuuagc
 cagcacuacagcuaaaggcuauaggagcaaauggcuggaucgagugagcaag
 cagcagaggccauggagguugcuagucaggcuaggcaaauggugcaagcg
 augagaaccuuugggacucacuccuagcuccagugcuggucugaaaauga
 ucuucugaaaauuugcaggccuaucagaaacgaauggggggugcagaugc
 aacgguucaagugaACUAGUGACUGACUAGCCCGCUGGGCCUCCCAACGG
 GCCCUCUCCCUCCUUGCACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AA

Fig. 1F

Matriz de la gripe: ARNm con un contenido de G/C aumentado y con secuencias de estabilización

GCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUCAACUUUGGCagauc
 uaaagaugagCcuGcuGaccgagguGgaGacCuacguGcuGAGCaucauc
 ccCAGCggcccccUGaaGgcccagaucgcCagagGcuGgaGgaCguGuu
 CgcCggCaagaacaccgaCcuGgagguGcuGauggaGuggcuGaagacCa
 gGccCauccugAGCccCugacCaagggCauCCUGggCuuCguguucaC
 cuGaccgugcccagCgagcgcCggCcuGcagcgcCCGCcgcuuCguGcaGaa
 CgcccuGaaCggCaacggCgaCccCaaCaacauggacaaGgcCguGaaGc
 uguaCaggaagcuGaagagggagauCacCuuccaCggCgccaGgaGauc
 AGCcuGagCuaCAGCgcCggCgcCcuGgccagCugCaugggccuGauCua
 caacaggauggggCgcCgugaccacCgaGguggcCuuCggccugguGugCg
 cCaccugCgaGcagauCgcCgacAGCcagcaCcgCAGCcaCaggcaGaug
 gugacCacCaccaaccCcuGaucagGcaCgagaacagGaugguGCUGgc
 cagcacCacCgcCaaggcCauggagcaGauggcCggCAGCaGCgagcaGg
 cCgcCgaggccauggagguGgcCagCcaaggcCaggcaGaugguGcaGgcC
 augagGaccuauCggCacCcaCccCagcAGCagCgcCggCcuGaaGaaCga
 CcuGcuGgaGaaCCUGcaggccuaCcagaaGcgCaugggCgugcagaugc
 aGcgCuucaagugaACUAGUGACUGACUAGCCCGCUGGGCCUCCCAACGG
 GCCCUCUCCCUCCUUGCACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AA

Fig. 1G

Matriz de la gripe: ARNm con un contenido de G/C aumentado y con secuencias de estabilización, que codifica la forma segregada

GCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUCAACUUUGGCagauc
 uaaagaugGCCGUCAUGGCCCCCGCACCCUGGUGCUGCUGCUGAGCGGC
 GCCCUGGCCUGACCCAGACCUGGGCCagCcuGcuGaccgagguGgaGac
 CuacguGcuGAGCaucaucccCAGCggcccccGaaGgccgagaucgcCc
 agagGcuGgaGgaCguGuuCgcCggCaagaacaccgaCcuGgagguGcuG
 auggaGuggcuGaagacCagGccCauccugAGCccCugacCaagggCau
 CCUGggCuuCguguuacCcuGaccgugcccagCgagcgCggCcuGcagc
 gCCGCcgcuuCguGcaGaaCgcccuGaaCggCaacggCgaCccCaaCaac
 auggacaaGgcCguGaaGcuGuaCaggaagcuGaagagggagauCacCuu
 ccaCggCgccaGgaGaucAGCcuGagCuaCAGCgcCggCgcCcuGgcca
 gCugCaugggccuGauCuacaacaggauGGGgCgcCgugaccacCgaGgug
 gcCuuCggccugguGugCgcCaccugCgaGcagauCgcCgacAGCcagca
 CcgCAGCcaCaggcaGauggugacCacCaccáaccCcuGaucagGcaCg
 agaacagGaugguGCUGgccagcacCacCgcCaaggcCauggagcaGaug
 gcCggCAGCaGCgagcaGgcCgcCgaggccauggagguGgcCagCaggc
 CaggcaGauggugcaGgcCaugagGaccauCggCacCcaCccCagcAGCa
 gCgcCggCcuGaaGaaCgaCcuGcuGgaGaaCCUGcaggccuaCcagaaG
 cgCaugggCgugcagaugcaGcgCuucaagugaACUAGUGACUGACUAGC
 CCGCUGGGCCUCCCAACGGGCCUCCUCCCCUCCUUGCACCAAAAAAAAAA
 AA
 AAAAAAAAAA

Fig. 2A

MAGE1: Gen del tipo salvaje (para comparación)

catcatgtctcttgagcagaggagtctgcactgcaagcctgaggaagccc
 ttgaggcccaacaagaggccctgggcctggtgtgtgtgcaggctgccacc
 tcctcctcctctcctctggtcctgggcaccctggaggaggtgccactgc
 tgggtcaacagatcctccccagagtcctcaggagcctccgcctttccca
 ctacatcaacttcaactcgacagaggcaaccagtgagggttccagcagc
 cgtgaagaggaggggccaagcacctcttgtatcctggagtccttgttccg
 agcagtaatcactaagaaggtggctgatttggttggttttctgctcctca
 aatcgcagccaggagccagtcacaaaggcagaaatgctggagagtgctc
 atcaaaaattacaagcactgttttctgagatcttcggcaaagcctctga
 gtccttgcaagctggtctttggcattgacgtgaaggaagcagaccccaccg
 gccactcctatgtccttgtcacctgcctaggtctctcctatgatggcctg
 ctgggtgataatcagatcatgcccaagacaggcttcctgataattgtcct
 ggtcatgattgcaatggagggcggccatgctcctgaggaggaaatctggg
 aggagctgagtgatggaggtgtatgatgggagggagcacagtgctat
 ggggagcccaggaagctgctcacccaagatttgggtgcaggaaaagtacct
 ggagtaccggcaggtgccggacagtgatcccgcacgctatgagttcctgt
 ggggtccaagggccctcgctgaaaccagctatgtgaaagtccttgagtat
 gtgatcaaggtcagtgcaagagttcgctttttcttcccacccctgcgtga
 agcagctttgagagaggaggaagagggagctgagcatga

Fig. 2B

MAGE1: Secuencia de proteínas

SER, LEU, GLU, GLN, ARG, SER, LEU, HIS, CYS, LYS, PRO, GLU, GLU, ALA, LEU, GLU,
 ALA, GLN, GLN, GLU, ALA, LEU, GLY, LEU, VAL, CYS, VAL, GLN, ALA, ALA, THR,
 SER, SER, SER, SER, PRO, LEU, VAL, LEU, GLY, THR, LEU, GLU, GLU, VAL, PRO, TH
 R, ALA, GLY, SER, THR, ASP, PRO, PRO, GLN, SER, PRO, GLN, GLY, ALA, SER, ALA,
 PHE, PRO, THR, THR, ILE, ASN, PHE, THR, ARG, GLN, ARG, GLN, PRO, SER, GLU, GL
 Y, SER, SER, SER, ARG, GLU, GLU, GLU, GLY, PRO, SER, THR, SER, CYS, ILE, LEU,
 GLU, SER, LEU, PHE, ARG, ALA, VAL, ILE, THR, LYS, LYS, VAL, ALA, ASP, LEU, VA
 L, GLY, PHE, LEU, LEU, LEU, LYS, TYR, ARG, ALA, ARG, GLU, PRO, VAL, THR, LYS,
 ALA, GLU, MET, LEU, GLU, SER, VAL, ILE, LYS, ASN, TYR, LYS, HIS, CYS, PHE, PR
 O, GLU, ILE, PHE, GLY, LYS, ALA, SER, GLU, SER, LEU, GLN, LEU, VAL, PHE, GLY,
 ILE, ASP, VAL, LYS, GLU, ALA, ASP, PRO, THR, GLY, HIS, SER, TYR, VAL, LEU, VA
 L, THR, CYS, LEU, GLY, LEU, SER, TYR, ASP, GLY, LEU, LEU, GLY, ASP, ASN, GLN,
 ILE, MET, PRO, LYS, THR, GLY, PHE, LEU, ILE, ILE, VAL, LEU, VAL, MET, ILE, AL
 A, MET, GLU, GLY, GLY, HIS, ALA, PRO, GLU, GLU, GLU, ILE, TRP, GLU, GLU, LEU,
 SER, VAL, MET, GLU, VAL, TYR, ASP, GLY, ARG, GLU, HIS, SER, ALA, TYR, GLY, GL
 U, PRO, ARG, LYS, LEU, LEU, THR, GLN, ASP, LEU, VAL, GLN, GLU, LYS, TYR, LEU,
 GLU, TYR, ARG, GLN, VAL, PRO, ASP, SER, ASP, PRO, ALA, ARG, TYR, GLU, PHE, LE
 U, TRP, GLY, PRO, ARG, ALA, LEU, ALA, GLU, THR, SER, TYR, VAL, LYS, VAL, LEU,
 GLU, TYR, VAL, ILE, LYS, VAL, SER, ALA, ARG, VAL, ARG, PHE, PHE, PHE, PRO, SE
 R, LEU, ARG, GLU, ALA, ALA, LEU, ARG, GLU, GLU, GLU, GLU, GLY, VAL, STP -
 , ALA, STP

