



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 340 499**

51 Int. Cl.:

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: **05004938 .6**

96 Fecha de presentación : **05.06.2002**

97 Número de publicación de la solicitud: **1604688**

97 Fecha de publicación de la solicitud: **14.12.2005**

54

Título: **ARNm de antígeno tumoral estabilizado con un contenido de G/C aumentado.**

30

Prioridad: **05.06.2001 DE 101 27 283**

45

Fecha de publicación de la mención BOPI:
04.06.2010

45

Fecha de la publicación del folleto de la patente:
04.06.2010

73

Titular/es: **Curevac GmbH
Paul-Ehrlich-Str. 15
72076 Tübingen, DE**

72

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

74

Agente: **Aznárez Urbieto, Pablo**

ES 2 340 499 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 de antígeno tumoral estabilizado con un contenido de G/C aumentado.

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, así como a este ARNm estabilizado. La composición farmacéutica según la invención es especialmente adecuada como vacuna.

10 La terapia genética y la vacunación genética son procedimientos médicos moleculares cuya aplicación en la terapia y la prevención de enfermedades tendrán considerables efectos sobre la práctica médica. Ambos procedimientos se basan en la introducción de ácidos nucleicos en las 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 para 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 las células, por ejemplo transfección con fosfato cálcico, transfección con polipreno, fusión de protoplastos, electroporación, microinyección y lipofección, donde especialmente la lipofección resultó ser un procedimiento adecuado.

20 Otro método especialmente propuesto en los procedimientos de vacunación genética es la utilización de ADN-virus 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 de efecto en la terapia génica así como los virales introducidos debido a posibles eventos de 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 para 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 eventos 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 cuando, debido a la integración del ADN extraño, se modifica un gen decisivo para la regulación del crecimiento celular. Por esta razón, cuando se utilizan ADN-virus como agentes terapéuticos genéticos y vacunas no se puede excluir un riesgo de desarrollo 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 o 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.

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

35 Al contrario que el ADN, la utilización de ARN como agente terapéutico genético o vacuna se puede considerar considerablemente más segura. En especial, 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 anti-ARN. Por esta razón, el ARN puede considerarse como una molécula opcional para los procedimientos de la terapia médica molecular.

40 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 al utilizar el ARN es la transferencia segura, eficiente específicamente en cuanto a la célula o bien el tejido, del ácido nucleico. Debido a que normalmente el ARN resulta muy inestable en solución, con los procedimientos tradicionales que se utilizan para el ADN hasta la fecha el ARN no ha podido utilizarse o se ha utilizado con una gran ineficacia como agente terapéutico o vacuna.

45 En cuanto a la inestabilidad son responsables las enzimas de desintegración del ARN, las llamadas 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. A este respecto, 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 200 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 decisivo para ello una interacción entre el ARN y las proteínas. Por ejemplo, se describió hace poco un "mRNA-Surveillance-System" (sistema de control de ARNm) (Hellerin y Parker, *Amun. Rev. Genet* 1999, 33: 229

ES 2 340 499 T3

a 260) en el que se conocen, debido a determinadas interacciones de proteínas “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 es llevado a cabo por exonucleasas.

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

10 En este contexto, la WO 98/34640 describe moléculas de ADN, pero sólo sintéticas, que codifican HIV gag y modificaciones de HIV gag. Las moléculas dadas a conocer en la WO 98/34640 se pueden utilizar como vacunas de polinucleótido que posibilitan una profilaxis inmunológica efectiva contra el HIV por estimulación de anticuerpos neutralizadores e inmunidad por mediación celular. Sin embargo, la WO 98/34640 no describe ningún tipo de aumento o maximización del contenido de G/C de las moléculas de ADN sintéticas codificadoras, ni ninguna molécula de ARNm modificada y utilizada en este sentido.

15 La WO 97/48370 también describe moléculas de ADN sintéticas que codifican un péptido o proteína, en particular HIV env y también modificaciones de HIV env. Pero la WO 97/48370 tampoco describe ningún tipo de aumento o maximización del contenido de G/C de las moléculas de ADN sintéticas codificadoras, ni ninguna modificación de las moléculas de ARNm utilizadas en este sentido.

20 En la EP-A-1083232 se propone para 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. También se da a conocer la utilización de estos complejos para la terapia contra el cáncer, en cuyo caso el ARNm codifica antígenos tumorales.

25 La WO 99/14346 describe otros procedimientos para la estabilización del ARNm. En particular, se proponen modificaciones del ARN 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, en especial la utilización de análogos de nucleótidos, así como grupos de bloqueo 5' y 3', una mayor longitud de la cola poli-A, así como la formación de complejos del ARNm con medios de estabilización y combinaciones de las medidas indicadas.

30 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 transiente” (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.

35 Bieler y Wagner (en: Schleaf (Hrsg.), 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 HIV-1 en el que se modifican los codones frente a la secuencia del tipo salvaje (utilización alternativa de codones, ingl. “codon usage”) de manera que se corresponda con 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 del 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 una 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 sobre la eficacia de la traducción es pequeño.

40 Por tanto, el objetivo de la presente invención consiste en proporcionar un nuevo sistema para la vacunación genética que elimina las desventajas relacionadas con las características de las sustancias terapéuticas y vacunas de ADN y aumenta la efectividad de las sustancias terapéuticas basadas en especies de ARN.

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

50 En particular, se propone un ARNm modificado de acuerdo con la invención, así como un compuesto farmacéutico que contiene un ARNm modificado de este tipo junto con un excipiente y/o vehículo farmacéuticamente compatible, donde el ARNm según la invención modificado codifica al menos un péptido o un polipéptido de antígeno tumoral, mostrando la secuencia del ARNm, especialmente en el campo que codifica al menos un péptido o un polipéptido, en comparación al ARNm del tipo salvaje, la siguiente modificación.

55 El contenido en G/C del campo que codifica el péptido o polipéptido del ARNm según la invención modificado es mayor que el contenido en G/C del campo de codificación del ARNm de tipo salvaje que codifica el péptido o el polipéptido, permaneciendo la secuencia del aminoácido codificada sin modificar frente al tipo salvaje.

ES 2 340 499 T3

Esta modificación se basa sobre 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(uridina). Por lo tanto, según invención se varían los codones frente al ARNm del tipo salvaje manteniendo la secuencia traducida del aminoácido, de forma que contienen una mayor cantidad de nucleótidos de 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, ingl. "codon usage").

Dependiendo del aminoácido a codificar por el ARNm modificado según la invención son factibles diferentes posibilidades para la modificación de la secuencia de ARNm frente a la secuencia del 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 CGC), 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 de 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 CCG;
- 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 de 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;
- 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 terminación UAA puede modificarse a UAG ó UGA;
- el codón para Cys puede modificarse de UGU a UGC;
- 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;
- 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.

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

Por supuesto, las sustituciones arriba indicadas pueden utilizarse 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

ES 2 340 499 T3

ACC (o ACG). Preferentemente, sin embargo, se utilizan combinaciones de las posibilidades anteriores de sustitución, como por ejemplo:

- 5 - 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);
- 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;
- 10 - 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 CCG);
- 15 - 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;
- 20 - 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); etc.
- 25

Preferentemente el contenido de G/C del campo del ARNm según la invención modificado que codifica el péptido o bien el polipéptido se incrementa en como mínimo un 7%, con especial preferencia en como mínimo un 15%, en particular en como mínimo un 20%, con respecto al contenido de G/C del campo codificado del ARNm de tipo salvaje que codifica el polipéptido. En este contexto es especialmente preferente aumentar al máximo el contenido de G/C del ARNm modificado, especialmente en el campo que codifica como mínimo un péptido o polipéptido, en comparación con la secuencia del tipo salvaje.

Otras modificaciones posibles no correspondientes a la invención del ARNm contenido en la composición farmacéutica caracterizada en la presente invención se basan 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. Si, por tanto, en una secuencia de ARN aparecen en mayor cantidad los llamados codones “raros”, el ARNm correspondiente se traduce claramente peor que en el caso para los ARNt relativamente “frecuentes”, donde existen codones de codificación.

Así, en el ARNm modificado (incluido en la composición farmacéutica), el campo que codifica el péptido o el polipéptido se puede modificar frente al correspondiente campo del ARNm de tipo salvaje de tal forma que como mínimo 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.

Mediante esta modificación no correspondiente a la invención se modifican las secuencias de ARNm de manera que se intercalan codones para los cuales están disponibles ARNt frecuentes.

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

Mediante esta modificación no correspondiente a la invención, de acuerdo con la invención se pueden cambiar todos los codones de la secuencia del 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 y que lleva, en cada caso, el mismo aminoácido que el ARNt relativamente raro.

Según la invención tiene especial preferencia vincular el contenido de G/C secuencial, especialmente 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. Este tipo de realización preferente proporciona un ARNm traducido y estabilizado de manera especialmente eficiente para la composición farmacéutica según la invención.

En las secuencias de ARNm eucariótico existen elementos de secuencia desestabilizantes (DSE) en los que se enlazan proteínas 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 como mínimo un péptido o un polipéptido, una o varias modificaciones frente al correspondiente campo del ARNm de tipo salvaje, de manera que no está incluido ningún elemento de la secuencia

ES 2 340 499 T3

desestabilizadora. Naturalmente, también es preferente eliminar del ARNm según la invención los DSE eventualmente presentes en los campos no traducidos (UTR 3' y/ó 5').

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

15 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 bacterianos antígenos) se puede sintetizar, por ejemplo, todo el ARNm de forma química utilizando técnicas estándar.

20 Sin embargo, preferentemente 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).

25 Por tanto, en este procedimiento se transcribe *in vitro* una molécula de ADN correspondiente para la producción del ARNm según la invención. 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 terminación para la transcripción *in vitro*. La molécula de ADN que constituye la matriz del constructo del ARN a producir, se puede obtener mediante un crecimiento fermentativo y subsiguiente aislamiento como parte de un plásmido replicable en bacterias. Como plásmidos adecuados para la presente invención se pueden citar, por ejemplo, los plásmidos pT7T (número de acceso al banco de genes U26404; Lai y col., Development 1995, 121: 2349 a 2360), la serie pGEM[®], 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 Griffen und Griffen (Editores), PCR Technology (tecnología de PCR): Current Innovation, CRC Press, Boca Raton, FL, 2001.

35 Así, es posible clonar, utilizando oligonucleótidos de ADN sintéticos cortos, conteniendo en las intersecciones que se producen transiciones cortas de un solo ramal, o mediante genes producidos por síntesis química según un método molecular-biológico conocido 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 del plásmido, en el que puede estar presente en copia simple o múltiple, mediante digestión con endonucleasas de restricción.

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

45 Según otra forma de realización preferente de la presente invención, el ARNm según la invención 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.

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

55 Además, es posible incluir en el ARNm según la invención modificado uno o varios de los llamados IRES (ingl. "internal ribosomal entry side" - sitio de entrada ribosomal interno). Así, un IRES puede actuar como única posición de enlace para ribosomas, sin embargo, también puede servir para proporcionar un ARNm que codifica varios péptidos o polipéptidos a traducir, de forma separada, por los ribosomas ("ARN multicistrónico"). Ejemplos de secuencias IRES que se pueden utilizar según la invención son aquellas de picornavirus (por ejemplo FMDV), virus de la peste (CFFV), virus de la poliomielitis (PV), virus de la encefalomiocarditis (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 virus de la parálisis de Cricket (CrPV).

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

65 Estas secuencias de estabilización pueden tener una homología de secuencia al 100% con las secuencias naturales que se presentan en los virus, bacterias y eucariotas; sin embargo, también pueden ser parcial o completamente de naturaleza sintética. Como ejemplo de secuencias de estabilización que se pueden utilizar en la presente invención se pueden citar las secuencias no traducidas (UTR) del gen de la globina β , por ejemplo de *homo sapiens* ó de *xenopus laevis*. Otro ejemplo de secuencia de estabilización tiene la fórmula general (C/U)CCAN_xCCC(U/A)Py_xUC(C/U)CC

ES 2 340 499 T3

contenida en el UTR 3' del ARNm muy 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 combinación con otras secuencias de estabilización conocidas por el técnico en la materia.

5 Para una mayor estabilización del ARNm según la invención modificado se prefiere, además, que éste tenga como mínimo un nucleótido análogo de procedencia natural. Esto se debe al hecho de que las enzimas que desintegran el ARN y existen en las células reconocen como sustrato, de preferencia, los nucleótidos de procedencia natural. Por tanto, mediante la introducción de análogos de nucleótidos es posible dificultar la desintegración del ARN, donde
10 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.

En un listado de ningún modo limitativo se pueden nombrar como ejemplos de análogos de nucleótidos a utilizar según invención fosfo-amidatos, fosfo-tioatos, nucleótidos de péptidos, metil-fosfonatos, 7-desazaguanosina, 5-metil-citosina e inosina. La producción de tales análogos es conocida por el técnico en la materia, por ejemplo de las patentes
15 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 traducidas y no traducidas del ARNm según la invención modificado.

Además, es posible mejorar la transferencia efectiva del ARNm según la invención 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 la combinación 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-
20 1083232.
25

Un posible campo de aplicación de la presente invención es la vacunación, es decir la utilización del ARNm según la invención modificado para la vacunación o la utilización de la composición farmacéutica como vacuna o la utilización según la invención del ARNm según la invención 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 (tumoral), en el presente caso de la información genética para el antígeno en forma del ARN modificado que codifica el antígeno. El ARNm según la invención 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, con lo cual se estimula una reacción inmunológica dirigida contra este polipéptido o péptido antígeno. En caso de la utilización
30 como vacuna genética para el tratamiento del cáncer, la reacción inmunológica se consigue por la introducción de la información genética para antígenos del tumor, especialmente proteínas que se expresan, exclusivamente, en células cancerosas, administrando una composición farmacéutica según la invención que contiene un ARNm según la invención que codifica un antígeno de tal tipo de cáncer. Debido a ello, se expresan el o los antígenos del cáncer en el organismo, provocándose una reacción inmunológica dirigida efectivamente contra las células cancerosas.
40

En su utilización como vacuna, la composición farmacéutica según la invención tiene su aplicación especialmente en el tratamiento de enfermedades cancerosas (donde el ARNm modificado codifica un antígeno superficial específico tumoral (TSSA)), por ejemplo para el tratamiento del melanoma maligno, carcinoma de colon, linfoma, sarcoma, carcinoma de pulmón de células pequeñas, blastomas, etc. Ejemplos específicos de antígenos tumorales son, entre otros, 707-AP, AFP, ART-4, BAGE, β -catenina/m, Bcr-abl, CAMEL, CAP-1, CASP-8, CDC27/m, CDK4/m, CEA, CT, Cyp-B, DAM, ELF2M, ETV6-AML1, G250, GAGE, GnT-V, Gp100, HAGE, HER-2/neu, HLA-A*0201-R170I, HPV-E7, HSP70-2M, HAST-2, hTERT (o hTRT), iCE, KIAA0205, LAGE, LDLR/FUr, MAGE, MART-1/Melan-A, MC 1 R, Miosina/m, MUC1, MUM-1, -2, -3, NASE-A, NY-ESO-1, p190 minor bcr-abl, Pml/RAR α , PRAME, PSA, PSM, RAGE, RU1 ó RU2, SAGE, SART-1 ó SART-3, TEL/AML1, TPI/m, TRP-1, TRP-2, TRP-2/INT2 y WT1.
50 Además, preferentemente se utilizan según la invención ARNm que codifican polipéptidos, donde los polipéptidos se tratan de poliepitopes, por ejemplo de los antígenos arriba mencionados, especialmente antígenos de superficie de células tumorales, preferentemente formas de proteínas segregadas.

Además, el ARNm modificado según la invención puede también contener, además del péptido o el polipéptido antígeno, como mínimo 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 , LNF- γ , GM-CFS, LT- α o factores de crecimiento como hGH.
55

Además, para aumentar la inmunogenicidad, la composición farmacéutica según la invención puede contener uno o varios adyuvantes. Como "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 según la invención 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 actúa 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 una reacción inmunológica contra un antígeno codificado por el ARNm según la invención modificado y/o ejercer una influencia precisa sobre esta reacción.
60
65

Son especialmente preferentes las citoquinas arriba mencionadas. Otros adyuvantes conocidos son hidróxido de aluminio, adyuvante de Freud, así como los péptidos o polipéptidos catiónicos de estabilización arriba mencionados, como la protamina. Además, son especialmente adecuados los lipopéptidos, como Pam3Cys, para ser utilizados como adyuvante en la composición farmacéutica de la presente invención, véase Deres y col., Nature 1989, 342: 561-564.

5 La composición farmacéutica según la invención contiene, además del ARNm según la invención modificado, un excipiente farmacéuticamente compatible y/o un vehículo farmacéuticamente compatible. Los correspondientes métodos para una formulación adecuada y la producción de la composición farmacéutica según la invención se describen en "Remington's Pharmaceutical Sciences" (Mack Pub. Co, Easton, PA, 1980). Para la administración parenteral se pueden utilizar como excipiente, por ejemplo, agua esterilizada, soluciones de sal común esterilizadas, polialquilen-
10 glicoles, naftaleno hidrogenado y, especialmente, polímeros lactida biocompatibles, copolímeros lactida/glicólido 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 los polímeros, por ejemplo de polietilenglicol como inhibidores según la invención, formación de complejo con iones de metal o inclusión de materiales
15 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 de 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 revestimientos con polímeros (por ejemplo poloxámeros o poloxaminas). Además, las sustancias o las composiciones según la invención pueden presentar recubrimientos protectores, por ejemplo inhibidores de proteasa o reforzadores de la permeabilidad. Típicamente, los excipientes preferentes sustancias
20 acuosas de excipientes, utilizándose agua para la inyección (WFI) o agua, tamponadas con fosfato, citrato o acetato, etc., y ajustándose el pH típicamente a 5,0 a 8,0, preferentemente de 6,0 a 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, azúcar o
30 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, como también del peso corporal, de la edad y del sexo del paciente.

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

40 Así, también se puede proporcionar por ejemplo 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, especialmente a un ser humano.

45 Además, también se puede emplear un procedimiento que sirve para averiguar la secuencia modificada del ARNm contenido en la composición farmacéutica según la invención. En este caso se puede llevar a cabo, por ejemplo, la adaptación de las secuencias de ARN con dos metas diferentes de optimización: por un lado, uno con un contenido de G/C lo más alto posible, por el otro lado, teniendo en cuenta la mejor frecuencia posible de los ARNt's según el Codon Usage. En el primer paso del procedimiento se realiza 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 de selección para los correspondientes codones debido al código genético degenerado. Según la optimización o modificación requerida, se utilizan para la selección del codón apropiado listas correspondientes de selección y algoritmos de optimización. La realización de los algoritmos tiene lugar, de forma típica, con ayuda de un software apropiado, en una computadora. Así, se obtiene la secuencia optimizada de ARNm y puede editarse, por ejemplo, con ayuda del 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 nucleótidos se destacan aquí especialmente. 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 los motivos naturales de secuencias.
50 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 campos del ARN.

Las figuras muestran:

65 Figura 1: secuencias modificadas y del tipo salvaje no correspondientes a la invención para la proteína matriz de la gripe. La Figura 1A muestra el gen del 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 matriz de la gripe, secuencia génica 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 matriz de la gripe (incluida una secuencia señal N-terminal) 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 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 no correspondiente a la invención que codifica la proteína 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 no correspondiente a la invención modificado que codifica la forma segregada de la proteína 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 han destacado en negrilla los codones de inicio y de 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 no correspondientes a la invención 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 alta frecuencia 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 modifica la secuencia de nucleótidos de un ARNm cualquiera con ayuda del código genético o bien su naturaleza degenerativa, de manera que resulta en un contenido máximo de G/C en relación con la utilización de codones que codifican los ARNt's que se presentan con mayor frecuencia en la célula, donde la secuencia de aminoácidos codificada por el ARNm modificado es idéntica 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% (si se compara con la secuencia del tipo salvaje del 51% (véase Kalnins y col., EMBOJ. 1983, 2(4): 593-597). Por la síntesis de oligonucleótidos solapados que tienen la secuencia modificada, se obtuvo la secuencia optimizada según procedimientos conocidos en la técnica actual. Los oligonucleótidos terminales tenían las siguientes interfaces de restricción. En el extremo 5' un interfaz EcoRV- así como en el extremo 3' un interfaz Bg/II. Mediante digestión con EcoRV/Bg/II 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 la 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 con fenol-cloroformo. Se transcribieron *in vitro* 2 μ g del constructo por procedimientos conocidos por el técnico en la materia, debido a lo cual 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 matriz de la gripe (secuencia del tipo salvaje, véase la 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 matriz de la gripe y codificaba una secuencia señal N-terminal (véase la Figura 1D). La forma segregada de la proteína matriz de la gripe tiene la ventaja de que tiene una mayor inmunogenicidad si se compara con la forma no segregada.

Partiendo de las secuencias optimizadas se diseñaron las respectivas moléculas de ARNm no correspondientes a la invención. El ARNm optimizado en cuanto al contenido de G/C y la utilización de codones para la proteína 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 de 5' o 3' del ARNm de la β -globina de *Xenopus laevis*; véase el vector pT7T 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 codifica la forma segregada de la proteína matriz de la gripe y se equipó con las secuencias de estabilización mencionadas (véase la Figura 1G).

Ejemplo 4

Con ayuda del programa de ordenador del Ejemplo 1 se modificó el ARNm que codifica el antígeno tumoral
 5 MAGe1. Así se averiguó la secuencia representada en la Figura 2C, que tiene un contenido de G/C mayor (351 G,
 291 C) en un 24% si se compara con la secuencia de tipo salvaje (275 G, 244 G). Además, se mejoró la secuencia del
 tipo salvaje, por la utilización alternativa de codones, en cuanto a la eficacia de traducción, debido a la codificación de
 los ARNt que aparecen 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%.

10 Anexo

Texto fuente de un ejemplo de programa de computador

Curevac_Genetic_Controls con los siguientes módulos

15

Nombre: Cureva_Genetic_Controls.vbp

Código:

```

20  Type=Control
    Reference="G(00020410-0000-0000-C000-000000000046)#2.0#..1.\WINDOWS\system32\STDOLE2.TLB#OLE Automation
    Object={0D452EE1-E08F-101A-852E-02608C4D0BB4}# 2.0# 0; FM20.DLL
    Object={F9043C88-F6F2-101A-A3C9-08002B2F49FB}# 1.2# 0; COMDLG32.COX
25  UserControl=Curevac_Amino.ctl
    Startup="(None)"
    HelpFile=""
    Title="Curevac Genetic Controls"
    Command32=""
30  Name="Curevac_Genetic_Controls"
    HelpContextID="0"
    CompatibleMode="1"
    MajorVer=1
    MinorVer=0
    RevisionVer=0
    AutoIncrementVer=0
    ServerSupportFiles=0
40  VersionComments="the RNA people"
    VersionCompanyName="CureVac GmbH"
    VersionFileDescription="Controls for Handling of Nucleotids"
    VersionLegalCopyright="by Christian Khump"
    VersionLegalTrademarks="Curevac Genetic Controls(tm)"
45  VersionProductName="Curevac Genetic Controls"
    CompilationType=0
    OptimizationType=0
    FavorPentiumPro(tm)=0
    CodeViewDebugInfo=0
    NoAliasing=0
    BoundsCheck=0
55  OverflowCheck=0
    FIPointCheck=0
    FDIVCheck=0
    UnroundedFP=0
    StartMode=1
    Unattended=0
    Retained=0
    ThreadPerObject=0
60  MaxNumberOfThreads=1
    ThreadingModel=1
    DebugStartupOption=1
65

```

ES 2 340 499 T3

DebugStartupComponent=Curevac_Amino

Nombre: Curevac_amino.ctf

5 Código:

```
VERSION 5.00
Object = "{0D452EE1-E08F-101A-852E-02608C4D0BB4}# 2.0# 0"; "FM20.DLL"
10 Begin VB.UserControl Curevac_Amino
    CanGetFocus = 0 False
    ClientHeight = 690
    ClientLeft = 0
    ClientTop = 0
    ClientWidth = 1200
    ScaleHeight = 690
    ScaleWidth = 1200
    15 Begin VB.Line linLower
        X1 = 0
        X2 = 1080
        Y1 = 600
        Y2 = 600
    End
    Begin VB.Line linUpper
    20 End
        X1 = 120
        X2 = 1080
        Y1 = 0
        Y2 = 0
    End
    25 Begin MSForms.Label lblAminoAcid
        DragMode = 1 'Automatic
        Height = 255
        Left = 120
        TabIndex = 1
        Top = 360
        Width = 975
        Size = "1720;450"
        SpecialEffect = 1
        FontName = "Lucida Sans Unicode"
        FontEffects = 1073741826
        FontHeight = 165
        30 FontCharSet = 0
        FontPitchAndFamily = 2
        ParagraphAlign = 3
    End
    35 Begin MSForms.Label lblTriplet
        DragMode = 1 'Automatic
        Height = 255
        Left = 120
    40
    45
    50
    55
    60
    65
```

```

5      TabIndex      = 0
      Top           = 120
      Width        = 975
      Size         = "1720;450"
      SpecialEffect = 1
      FontHeight   = 165
10     FontCharSet  = 0
      FontPitchAndFamily = 2
      ParagraphAlign = 3
      End
15     End
      Attribute VB_Name = "Curevac_Amino"
      Attribute VB_GlobalNameSpace = False
      Attribute VB_Creatable = True
20     Attribute VB_PredeclaredId = False
      Attribute VB_Exposed = True
      Option Explicit

25     Private msAAShortcut As String
      Private msAAName As String
      Private msBestTriplet As String
      Private msSecondBest As String
30     Private msThirdBest As String
      Private msTriplet As String
      Private msBackColor As Long
      Private mbShowOriginal As Boolean

35

      Public Enum EnumAminoAcid
          amaGlycin
          amaAlanin
40         amaValin
          amaLeucin
          amaIsoLeucin
          amaPhenylalanin
          amaTyrosin
          amaTryptophan
          amaAsparaginAcid
50         amaAsparagin
          amaGlutaminAcid
          amaGlutamin
          amaSerin
          amaThreonin
55         amaCystein
          amaMethionin
          amaProlin
          amaHistidin
          amaLysin
60
65

```

```
amaStop
amaStart
End Enum
```

5

```
Private Sub Main()
```

```
    Call UserControl1.Resize
```

10

```
End Sub
```

```
Public Sub Translation(ByVal sTriplet As String)
```

15

```
    msTriplet = sTriplet
```

```
    Select Case sTriplet
```

20

```
        Case "GGU", "GGC", "GGA", "GGG"
```

```
            msAAShortcut = "GLY"
```

```
        Case "GCU", "GOC", "GCA", "GOG"
```

```
            msAAShortcut = "ALA"
```

25

```
        Case "GUU", "GUC", "GUA", "GUG"
```

```
            msAAShortcut = "VAL"
```

```
        Case "UUA", "UUG", "CUU", "CUA", "CUG", "CUC"
```

```
            msAAShortcut = "LEU"
```

30

```
        Case "AUA", "AUU", "AUC"
```

```
            msAAShortcut = "ILE"
```

```
        Case "UUU", "UUC"
```

```
            msAAShortcut = "PHE"
```

35

```
        Case "UAU", "UAC"
```

```
            msAAShortcut = "TYR"
```

```
        Case "UGG"
```

```
            msAAShortcut = "TRP"
```

40

```
        Case "GAU", "GAC"
```

```
            msAAShortcut = "ASP"
```

```
        Case "AAU", "AAC"
```

```
            msAAShortcut = "ASN"
```

45

```
        Case "GAA", "GAG"
```

```
            msAAShortcut = "GLU"
```

```
        Case "CAA", "CAG"
```

```
            msAAShortcut = "GLN"
```

50

```
        Case "AGU", "AGC", "UCA", "UCU", "UOG", "UOC"
```

```
            msAAShortcut = "SER"
```

```
        Case "ACA", "ACU", "AOG", "AOC"
```

```
            msAAShortcut = "THR"
```

55

```
        Case "UGU", "UGC"
```

```
            msAAShortcut = "CYS"
```

```
        Case "AUG"
```

```
            msAAShortcut = "MET"
```

60

```
        Case "OCA", "OCU", "OOG", "OOC"
```

65

```

    msAAShortcut = "PRO"
    Case "CAU", "CAC"
    msAAShortcut = "HIS"
5   Case "AGA", "AGG", "CGA", "CGU", "CGG", "CGC"
    msAAShortcut = "ARG"
    Case "AAA", "AAG"
10  msAAShortcut = "LYS"
    Case "UAA", "UAG", "UGA"
    msAAShortcut = "STP"
    End Select
15  Call Synthesis(msAAShortcut)
    Call Show

End Sub

20  Public Sub Synthesis(ByVal sShortCut As String)

    msAAShortcut = sShortCut 'nur falls die Sub extern angesprochen wird

25  Select Case msAAShortcut
    Case "GLY"
    msAAName = "Glycin"
    msBestTriplet = "GGC"
30  msSecondBest = "GGG"
    msBackColor = 8421631
    Case "ALA"
    msAAName = "Alanin"
35  msBestTriplet = "GCG"
    msSecondBest = "GOC"
    msBackColor = 8454143
    Case "VAL"
40  msAAName = "Valin"
    msBestTriplet = "GUC"
    msSecondBest = "GUG"
    msBackColor = 8454016
45  Case "LEU"
    msAAName = "Leucin"
    msBestTriplet = "CUG"
    msSecondBest = "CUC"
50  msBackColor = 16777088
    Case "ILE"
    msAAName = "Isoleucin"
55  msBestTriplet = "AUC"
    msBackColor = 12615935
    Case "PHE"
    msAAName = "Phenylalanin"
60  msBestTriplet = "UUC"
    msBackColor = 255

```

65

Case "TYR"

msAAName = "Tyrosin"

msBestTriplet = "UAC"

msBackColor = 4210816

Case "TRP"

msAAName = "Tryptophan"

msBestTriplet = "UGG"

msBackColor = 4227327

Case "ASP"

msAAName = "Asparaginsäure"

msBestTriplet = "GAC"

msBackColor = 8388863

Case "ASN"

msAAName = "Asparagin"

msBestTriplet = "AAC"

msBackColor = 4227200

Case "GLU"

msAAName = "Glutaminsäure"

msBestTriplet = "GAG"

msBackColor = 32768

Case "GLN"

msAAName = "Glutamin"

msBestTriplet = "GGC"

msBackColor = 8421440

Case "SER"

msAAName = "Serin"

msBestTriplet = "AGC"

msSecondBest = "UOG"

msThirdBest = "UOC"

msBackColor = 16512

Case "THR"

msAAName = "Threonin"

msBestTriplet = "ACG"

msBackColor = 16711680

Case "CYS"

msAAName = "Cystein"

msBestTriplet = "UGC"

msBackColor = 6932960

Case "MET"

msAAName = "Methionin"

msBestTriplet = "AUG"

msBackColor = 10417643

Case "PRO"

msAAName = "Prolin"

msBestTriplet = "OOG"

msSecondBest = "OOC"

msBackColor = 12898746

Case "HIS"

5

10

15

20

25

30

35

40

45

50

55

60

65

```

msAAName = "Histidin"
msBestTriplet = "CAC"
msBackColor = 12898746
5 Case "ARG"
msAAName = "Arginin"
msBestTriplet = "OGC"
10 msSecondBest = "OGG"
msBackColor = 6174925
Case "LYS"
msAAName = "Lysin"
15 msBestTriplet = "AAG"
msBackColor = 14141641
Case "STP"
msAAName = "Stop"
20 msBestTriplet = "UGA"
msSecondBest = "UAG"
msBackColor = 11332093
End Select

25 End Sub

Public Sub Show()

30 lblAminoAcid.Caption = msAAShortcut
If mbShowOriginal = True Then
    lblTriplet.Caption = msTriplet
Else
35 lblTriplet.Caption = msBestTriplet
End If
lblAminoAcid.BackColor = msBackColor
40 lblTriplet.BackColor = msBackColor

End Sub

45 Public Function ChooseBestGC(sShortCut As String)

Select Case sShortCut
Case "GLY"
50 msAAName = "Glycin"
msBestTriplet = "GGC"
msSecondBest = "GGG"
msBackColor = 8421631
Case "ALA"
55 msAAName = "Alanin"
msBestTriplet = "GOG"
msSecondBest = "GOC"
msBackColor = 8454143
60 Case "VAL"

```

65

5
msAAName = "Valin"
msBestTriplet = "GUC"
msSecondBest = "GUG"
msBackColor = 8454016
Case "LEU"
10
msAAName = "Leucin"
msBestTriplet = "CUG"
msSecondBest = "CUC"
msBackColor = 16777088
Case "ILE"
15
msAAName = "Isoleucin"
msBestTriplet = "AUC"
msBackColor = 12615935
Case "PHE"
20
msAAName = "Phenylalanin"
msBestTriplet = "UUC"
msBackColor = 255
Case "TYR"
25
msAAName = "Tyrosin"
msBestTriplet = "UAC"
msBackColor = 4210816
Case "TRP"
30
msAAName = "Tryptophan"
msBestTriplet = "UGG"
msBackColor = 4227327
Case "ASP"
35
msAAName = "Asparaginsäure"
msBestTriplet = "GAC"
msBackColor = 8388863
Case "ASN"
40
msAAName = "Asparagin"
msBestTriplet = "AAC"
msBackColor = 4227200
Case "GLU"
45
msAAName = "Glutaminsäure"
msBestTriplet = "GAG"
msBackColor = 16711808
Case "GLN"
50
msAAName = "Glutamin"
msBestTriplet = "GGC"
msBackColor = 12632256
Case "SER"
55
msAAName = "Serin"
msBestTriplet = "AGC"
msSecondBest = "UOG"
msThirdBest = "UOC"
60
msBackColor = 8421504
Case "THR"

65

```

msAAName = "Threonin"
msBestTriplet = "ACG"
msBackColor = &dHD4COFF
5 Case "CYS"
msAAName = "Cystein"
msBestTriplet = "UGC"
msBackColor = &dHD5COFF
10 Case "MET"
msAAName = "Methionin"
msBestTriplet = "AUG"
msBackColor = &dHD6COFF
15 Case "PRO"
msAAName = "Prolin"
msBestTriplet = "OOG"
msSecondBest = "OOC"
msBackColor = &dHD7COFF
20 Case "HIS"
msAAName = "Histidin"
msBestTriplet = "CAC"
msBackColor = &dHD8COFF
25 Case "ARG"
msAAName = "Arginin"
msBestTriplet = "OGC"
msSecondBest = "OGG"
msBackColor = &dHD9COFF
30 Case "LYS"
msAAName = "Lysin"
msBestTriplet = "AAG"
msBackColor = &dHE0COFF
35 Case "STP"
msAAName = "Stop"
msBestTriplet = "UGA"
msSecondBest = "UAG"
msBackColor = &dHE1COFF
40 End Select
45
End Function
50
Public Property Let ShowOriginal(ByVal bNewValue As Boolean)
mbShowOriginal = bNewValue
55 End Property
Public Property Get ShowOriginal() As Boolean
ShowOriginal = mbShowOriginal
60 End Property
65

```

```

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

```

5

```

Public Property Get AAName() As String
  AAName = msAAName
End Property

```

10

```

Public Property Get BestTriplet() As String
  BestTriplet = msBestTriplet
End Property

```

15

```

Public Property Get SecondBest() As String
  SecondBest = msSecondBest
End Property

```

20

```

Public Property Get ThirdBest() As String
  ThirdBest = msThirdBest
End Property

```

25

```

Public Property Get BackColor() As String
  BackColor = msBackColor
End Property

```

30

```

Private Property Get Triplet() As String
  Triplet = msTriplet
End Property

```

35

```

Private Sub UserControl_Resize()

```

40

```

  Dim lOuterWidth As Long
  Dim lDistance As Long

```

45

```

    lDistance = 30

```

```

    lOuterWidth = ScaleWidth + lDistance

```

50

```

    With lblTriplet

```

```

      .Top = 0

```

```

      .Left = lDistance / 2

```

```

      .Width = ScaleWidth - lDistance

```

55

```

      .Height = ScaleHeight / 2 + 30

```

```

    End With

```

```

    With lblAminoAcid

```

60

```

      .Top = ScaleHeight / 2 - 30

```

65

ES 2 340 499 T3

```
5      .Left = IDistance / 2
      .Width = ScaleWidth - IDistance
      .Height = ScaleHeight / 2
      End With
      With linUpper
10         .X1 = ScaleLeft
         .Y1 = 0
         .X2 = ScaleWidth
         .Y2 = 0
      End With
      With linLower
15         .X1 = ScaleLeft
         .Y1 = ScaleHeight - 20
         .X2 = ScaleWidth
         .Y2 = ScaleHeight - 20
      End With
20
      End Sub
25
```

Curevac_RNA_Optimizer con los siguientes módulos

30 Nombre: Curevac_RNA-Optimizer.vbp

Código:

```
35      Type=Exe
      Reference=*\G{00020430-0000-0000-C000-
      000000000046}# 2.0# 0# ..\..\WINNT\System32\STDOLE2.TLB# OLE Automation
40      Object={F9043C88-F6F2-101A-A3C9-08002B2F49FB}# 1.2# 0; OOMDLG32.OCX
      Module=basMain; basMain.bas
      Form=mdiMain.frm
      Object=*\ACurevac_Genetic_Controls.vbp
45      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
50      Module=basPublics; basPublics.bas
      Form=frmOptions.frm
      Form=frmInput.frm
      Form=frmDisplay.frm
      Form=frmAbout.frm
55      Startup="mdiMain"
      HelpFile=""
      Title="RNA-Optimizer"
      Command32=""
60      Name="RNA_Optimizer"
      HelpContextID="0"
65
```

```

CompatibleMode="0"
MajorVer=1
MinorVer=0
RevisionVer=0
AutoIncrementVer=0
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"
CompilationType=0
OptimizationType=0
FavorPentiumPro(tm)=0
CodeViewDebugInfo=0
NoAliasing=0
BoundsCheck=0
OverflowCheck=0
FPointCheck=0
FDIVCheck=0
UnroundedFP=0
StartMode=0
Unattended=0
Retained=0
ThreadPerObject=0
MaxNumberOfThreads=1
DebugStartupOption=0

Name: basMain.bas
Code:
Attribute VB_Name = "basMain"
Option Explicit

Private Type udtAminoAcid
    sShortcut As String
    sName As String
    sBestTriplet As String
    sSecondBest As String
    sThirdBest As String
End Type

Private mastrAminoAcids() As udtAminoAcid

Public Sub RebuildChain(bShowOriginal As Boolean)

Dim iLoop As Integer

```

65

```

For iLoop = 1 To glAminoCounter
  With frmDisplay.Curevac_Amino1(iLoop)
    .ShowOriginal = bShowOriginal
5     Show
    End With
  Next iLoop

10 End Sub

Public Sub ResetPublics()
  gsOriginalCode = ""
  gsFormattedCode = ""
  glCodeLength = ""
15 End Sub

20

Public Sub BuildAminoChain(ByVal sTempCode As String, bGraphical As Boolean)

  Dim lVerticalOffset As Long
  Dim lHorizontalOffset As Long
  Dim lTop As Long

  Dim asTriplets() As String
  Dim sAminoChain As String
  Dim lLoop As Long

  If bGraphical = True Then
35     lTop = frmDisplay.optNucleotids(0).Height + 40
  End If
  ReDim mastrAminoAcids(glAminoCounter)
  asTriplets = Split(sTempCode)
40  For lLoop = 0 To glAminoCounter - 1
    If bGraphical = True Then
      Load frmDisplay.Curevac_Amino1(lLoop + 1)
      With frmDisplay.Curevac_Amino1(lLoop + 1)
45         Call .Translation(asTriplets(lLoop))

        .Left = lLoop * .Width - lHorizontalOffset
        .Top = lVerticalOffset + lTop
50     If .Left + 2 * .Width > frmDisplay.ScaleWidth Then
        lVerticalOffset = lVerticalOffset + .Height + 150
        lHorizontalOffset = lLoop * .Width + .Width
55     End If

        sAminoChain = sAminoChain + aaShortcut + ","
        .Visible = True
60
65

```

```

    End With
Else
5     mstrAminoAcids(iLoop) = Translation(asTriplets(iLoop))
    sTest = sTest + mstrAminoAcids(iLoop).sName + ","
End If
Next
10    If bGraphical = True Then
        frmDisplay.Show
    Else
        Call frmOutPut.Show
15    End If
    Debug.Print sAminoChain
End Sub

20 Private Function Translation(ByVal sTriplet As String) As udtAminoAcid

    Dim stTemp As udtAminoAcid

25    If mdiMain.optKindOfOptimize(0) = True Then
        Optimization for Most GC
        Select Case sTriplet
30            Case "GGU", "GGC", "GGA", "GGG"
                stTemp.sShortcut = "GLY"
                stTemp.sName = "Glycin"
                stTemp.sBestTriplet = "GGC"
                stTemp.sSecondBest = "GGG"
35            Case "GCU", "GCC", "GCA", "GCG"
                stTemp.sShortcut = "ALA"
                stTemp.sName = "Alanin"
                stTemp.sBestTriplet = "GCG"
                stTemp.sSecondBest = "GCC"
40            Case "GUU", "GUC", "GUA", "GUG"
                stTemp.sShortcut = "VAL"
                stTemp.sName = "Valin"
                stTemp.sBestTriplet = "GUC"
                stTemp.sSecondBest = "GUG"
45            Case "UUA", "UUG", "CUU", "CUA", "CUG", "CUC"
                stTemp.sShortcut = "LEU"
                stTemp.sName = "Leucin"
                stTemp.sBestTriplet = "CUG"
                stTemp.sSecondBest = "CUC"
50            Case "AUA", "AUU", "AUC"
                stTemp.sShortcut = "ILE"
                stTemp.sName = "Isoleucin"
                stTemp.sBestTriplet = "AUC"
55            Case "UUU", "UUC"
                stTemp.sShortcut = "PHE"
                stTemp.sName = "Phenylalanin"
60
65

```

```

    stTemp.sBestTriplet = "UUC"
Case "UAU", "UAC"
    stTemp.sShortcut = "TYR"
    stTemp.sName = "Tyrosin"
    stTemp.sBestTriplet = "UAC"
Case "UGG"
    stTemp.sShortcut = "TRP"
    stTemp.sName = "Tryptophan"
    stTemp.sBestTriplet = "UGG"
Case "GAU", "GAC"
    stTemp.sShortcut = "ASP"
    stTemp.sName = "Asparaginsäure"
    stTemp.sBestTriplet = "GAC"
Case "AAU", "AAC"
    stTemp.sShortcut = "ASN"
    stTemp.sName = "Asparagin"
    stTemp.sBestTriplet = "AAC"
Case "GAA", "GAG"
    stTemp.sShortcut = "GLU"
    stTemp.sName = "Glutaminsäure"
    stTemp.sBestTriplet = "GAG"
Case "CAA", "CAG"
    stTemp.sShortcut = "GLN"
    stTemp.sName = "Glutamin"
    stTemp.sBestTriplet = "CAG"
Case "AGU", "AGC", "UCA", "UCU", "UOG", "UCC"
    stTemp.sShortcut = "SER"
    stTemp.sName = "Serin"
    stTemp.sBestTriplet = "AGC"
    stTemp.sSecondBest = "UOG"
    stTemp.sThirdBest = "UCC"
Case "ACA", "ACU", "AOC", "AOC"
    stTemp.sShortcut = "THR"
    stTemp.sName = "Threonin"
    stTemp.sBestTriplet = "AOC"
Case "UGU", "UGC"
    stTemp.sShortcut = "CYS"
    stTemp.sName = "Cystein"
    stTemp.sBestTriplet = "UGC"
Case "AUG"
    stTemp.sShortcut = "MET"
    stTemp.sName = "Methionin"
    stTemp.sBestTriplet = "AUG"
Case "CCA", "CCU", "COG", "CCC"
    stTemp.sShortcut = "PRO"
    stTemp.sName = "Prolin"
    stTemp.sBestTriplet = "COG"
    stTemp.sSecondBest = "CCC"

```

```

Case "CAU", "CAC"
  suTemp.sShortcut = "HIS"
  suTemp.sName = "Histidin"
  suTemp.sBestTriplet = "CAC"
Case "AGA", "AGG", "OGA", "OGU", "OGG", "OGC"
  suTemp.sShortcut = "ARG"
  suTemp.sName = "Arginin"
  suTemp.sBestTriplet = "OGC"
  suTemp.sSecondBest = "OGG"
Case "AAA", "AAG"
  suTemp.sShortcut = "LYS"
  suTemp.sName = "Lysin"
  suTemp.sBestTriplet = "AAG"
Case "UAA", "UAG", "UGA"
  suTemp.sShortcut = "STP"
  suTemp.sName = "Stop"
  suTemp.sBestTriplet = "UGA"
  suTemp.sSecondBest = "UAG"
End Select
Translation = suTemp
Else Optimization for best frequency
Select Case sTriplet
  Case "GGU", "GGC", "GGA", "GGG"
    suTemp.sShortcut = "GLY"
    suTemp.sName = "Glycin"
    suTemp.sBestTriplet = "GGC"
    suTemp.sSecondBest = "GGU"
  Case "GCU", "GOC", "GCA", "GOG"
    suTemp.sShortcut = "ALA"
    suTemp.sName = "Alanin"
    suTemp.sBestTriplet = "GOC"
    suTemp.sSecondBest = "GCU"
  Case "GUU", "GUC", "GUA", "GUG"
    suTemp.sShortcut = "VAL"
    suTemp.sName = "Valin"
    suTemp.sBestTriplet = "GUG"
    suTemp.sSecondBest = "GUC"
  Case "UUA", "UUG", "CUU", "CUA", "CUG", "CUC"
    suTemp.sShortcut = "LEU"
    suTemp.sName = "Leucin"
    suTemp.sBestTriplet = "CUG"
    suTemp.sSecondBest = "CUC"
  Case "AUA", "AUU", "AUC"
    suTemp.sShortcut = "ILE"
    suTemp.sName = "Isoleucin"
    suTemp.sBestTriplet = "AUC"
  Case "UUU", "UUC"
    suTemp.sShortcut = "PHE"

```

```

    stTemp.sName = "Phenylalanin"
    stTemp.sBestTriplet = "UUC"
5   Case "UAU", "UAC"
    stTemp.sShortcut = "TYR"
    stTemp.sName = "Tyrosin"
    stTemp.sBestTriplet = "UAC"
10  Case "UGG"
    stTemp.sShortcut = "TRP"
    stTemp.sName = "Tryptophan"
    stTemp.sBestTriplet = "UGG"
15  Case "GAU", "GAC"
    stTemp.sShortcut = "ASP"
    stTemp.sName = "Asparaginsäure"
    stTemp.sBestTriplet = "GAC"
20  Case "AAU", "AAC"
    stTemp.sShortcut = "ASN"
    stTemp.sName = "Asparagin"
    stTemp.sBestTriplet = "AAC"
25  Case "GAA", "GAG"
    stTemp.sShortcut = "GLU"
    stTemp.sName = "Glutaminsäure"
    stTemp.sBestTriplet = "GAG"
30  Case "CAA", "CAG"
    stTemp.sShortcut = "GLN"
    stTemp.sName = "Glutamin"
    stTemp.sBestTriplet = "CAG"
35  Case "AGU", "AGC", "UCA", "UCU", "UOG", "UOC"
    stTemp.sShortcut = "SER"
    stTemp.sName = "Serin"
    stTemp.sBestTriplet = "AGC"
    stTemp.sSecondBest = "UOC"
    stTemp.sThirdBest = "UCU"
40  Case "ACA", "ACU", "AOG", "AOC"
    stTemp.sShortcut = "THR"
    stTemp.sName = "Threonin"
    stTemp.sBestTriplet = "AOC"
45  Case "UGU", "UGC"
    stTemp.sShortcut = "CYS"
    stTemp.sName = "Cystein"
    stTemp.sBestTriplet = "UGC"
50  Case "AUG"
    stTemp.sShortcut = "MET"
    stTemp.sName = "Methionin"
    stTemp.sBestTriplet = "AUG"
55  Case "OCA", "OCU", "OOG", "OOC"
    stTemp.sShortcut = "PRO"
    stTemp.sName = "Prolin"
    stTemp.sBestTriplet = "OOC"
60
65

```

```

    stTemp.sSecondBest = "OCU"
    Case "CAU", "CAC"
5      stTemp.sShortcut = "HIS"
      stTemp.sName = "Histidin"
      stTemp.sBestTriplet = "CAC"
    Case "AGA", "AGG", "OGA", "OGU", "OGG", "OGC"
10     stTemp.sShortcut = "ARG"
      stTemp.sName = "Arginin"
      stTemp.sBestTriplet = "OGC"
      stTemp.sSecondBest = "AGG"
15     Case "AAA", "AAG"
      stTemp.sShortcut = "LYS"
      stTemp.sName = "Lysin"
      stTemp.sBestTriplet = "AAG"
20     Case "UAA", "UAG", "UGA"
      stTemp.sShortcut = "STP"
      stTemp.sName = "Stop"
      stTemp.sBestTriplet = "UGA"
25     stTemp.sSecondBest = "UAG"
    End Select
    Translation = stTemp
  End If
30
End Function

Public Function ReverseTranscription() As String
35   Dim sTempCode As String
   Dim iLoop As Integer
   For iLoop = 0 To glAminoCounter - 1
     sTempCode = sTempCode + mastrAminoAcids(iLoop).sBestTriplet
40   Next
   ReverseTranscription = sTempCode
End Function

45 Public Sub CountBases(ByVal sTempCode As String)

   Dim iLoop As Long
   Dim sFormattedCode As String
   Dim sActualBase As String
50   Dim iCodeLength As Long

   iCodeLength = Len(sTempCode)
   For iLoop = 1 To CLng(iCodeLength)
55     sActualBase = Mid(sTempCode, iLoop, 1)
     Select Case sActualBase
       Case "A", "a"
60         glOptAdenin = glOptAdenin + 1
       Case "U", "u", "T", "t"

```

```

    glOptThymin = glOptThymin + 1
    Case "G", "g"
    glOptGuanin = glOptGuanin + 1
    Case "C", "c"
    glOptCytosin = glOptCytosin + 1
    End Select
Next ILoop
End Sub

```

15 Name: basPublics.bas

Code:

Attribute VB_Name = "basPublics"

Option Explicit

Public gsOriginalCode As String

Public gsFormattedCode As String

Public glCodeLength As Long

Public glAminoCounter As Long

Public glAdenin As Long

Public glThymin As Long

Public glGuanin As Long

Public glCytosin As Long

Public glOptAdenin As Long

Public glOptThymin As Long

Public glOptGuanin As Long

Public glOptCytosin As Long

Public gbSequenceChanged As Boolean

45 Name: mdiMain.frm

Code:

VERSION 5.00

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

Begin VB.MDIForm mdiMain

BackColor = &H0FFFFFF&

Caption = "Curevac_RNA_Analyzer"

ClientHeight = 8745

ClientLeft = 165

ClientTop = 735

ClientWidth = 12255

Icon = "mdiMain.frx".0000

LinkTopic = "MDIForm1"

Picture = "mdiMain.frx".058A

```

StartUpPosition = 3 'Windows Default
WindowState = 2 'Maximized
Begin ComCtl3.CoolBar cbarMain
5   Align = 1 'Align Top
    Height = 885
    Left = 0
10  TabIndex = 0
    Top = 0
    Width = 12255
    _ExtentX = 21616
15  _ExtentY = 1561
    BandCount = 1
    BandBorders = 0 'False
    _CBWidth = 12255
20  _CBHeight = 885
    _Version = "6.7.8988"
    MinHeight1 = 825
    Width1 = 4995
    FixedBackground1 = 0 'False
25  NewRow1 = 0 'False
Begin VB.OptionButton optKindOfOptimize
    Caption = "Best Frequency"
30  Height = 255
    Index = 1
    Left = 4920
    TabIndex = 7
35  Top = 480
    Value = -1 'True
    Width = 1575
End
40  Begin VB.OptionButton optKindOfOptimize
    Caption = "Best GC"
    Height = 255
    Index = 0
45  Left = 4920
    TabIndex = 6
    Top = 120
    Width = 1575
End
50  Begin VB.CommandButton cmdShowInput
    Caption = "Input"
    Height = 765
55  Left = 0
    Picture = "mdiMain.frx".OC2A
    Style = 1 'Graphical
    TabIndex = 5
60  Top = 0
    Width = 840

```

65

```

End
Begin VB.CommandButton cmdShowOptions
5   Caption   = "Options"
    Enabled   = 0 'False
    Height    = 760
    Left      = 3480
10   Picture   = "mdlMain.frx":0F34
    Style     = 1 'Graphical
    TabIndex  = 4
    Top       = 0
    Width     = 735
End
Begin VB.CommandButton cmdShowStatistics
20   Caption   = "Statistics"
    Enabled   = 0 'False
    Height    = 760
    Left      = 2640
    Picture   = "mdlMain.frx":131E
25   Style     = 1 'Graphical
    TabIndex  = 3
    Top       = 0
    Width     = 735
End
30   Begin VB.CommandButton cmdShowDisplay
    Caption   = "Display"
    Enabled   = 0 'False
35   Height    = 760
    Left      = 1800
    Picture   = "mdlMain.frx":1628
    Style     = 1 'Graphical
40   TabIndex  = 2
    Top       = 0
    Width     = 735
End
45   Begin VB.CommandButton cmdShowOutput
    Caption   = "Output"
    Enabled   = 0 'False
    Height    = 760
50   Left      = 960
    Picture   = "mdlMain.frx":20E2
    Style     = 1 'Graphical
    TabIndex  = 1
55   Top       = 0
    Width     = 735
End
End
Begin VB.Menu mnuMain
60   Caption   = "&Input..."

```

65

```

Index      = 0
End
5  Begin VB.Menu mnuMain
Caption    = "&Results"
Index     = 1
10  Begin VB.Menu mnuResults
Caption    = "&Output..."
Enabled   = 0 'False
Index     = 0
End
15  Begin VB.Menu mnuResults
Caption    = "&Display..."
Enabled   = 0 'False
Index     = 1
20  End
25  Begin VB.Menu mnuResults
Caption    = "&Statistics..."
Enabled   = 0 'False
Index     = 2
End
End
30  Begin VB.Menu mnuMain
Caption    = "E&xtras"
Index     = 4
35  Begin VB.Menu mnuExtras
Caption    = "&Language"
Index     = 0
40  Begin VB.Menu mnuLanguage
Caption    = "English"
Checked   = -1 'True
Index     = 0
End
45  Begin VB.Menu mnuLanguage
Caption    = "&German"
Enabled   = 0 'False
Index     = 1
End
50  Begin VB.Menu mnuLanguage
Caption    = "&French"
Enabled   = 0 'False
Index     = 2
End
55  End
End
60  Begin VB.Menu mnuExtras
Caption    = "&Options..."
Enabled   = 0 'False
Index     = 1
End
65

```

```

5      Begin VB.Menu mnuExtras
        Caption   = ". ."
        Index     = 2
      End
      Begin VB.Menu mnuExtras
        Caption   = "&About"
        Index     = 3
      End
      End
      Begin VB.Menu mnuMain
        Caption   = "&Windows"
        Index     = 5
        WindowList = -1 True
      Begin VB.Menu mnuWindows
        Caption   = "Tile &Horizontally"
        Index     = 0
      End
        Begin VB.Menu mnuWindows
          Caption   = "Tile &Vertically"
          Index     = 1
        End
        Begin VB.Menu mnuWindows
          Caption   = "&Cascade"
          Index     = 2
        End
      End
      End
      Begin VB.Menu mnuMain
        Caption   = "&Exit"
        Index     = 6
      End
      End
      Attribute VB_Name = "mdiMain"
      Attribute VB_GlobalNameSpace = False
      Attribute VB_Creatable = False
      Attribute VB_PredeclaredId = True
      Attribute VB_Exposed = False
      Option Explicit

50     Private Sub cmdShowDisplay_Click()
          Call frmDisplay.Show
        End Sub

55     Private Sub cmdShowInput_Click()
          Call frmInput.Show
        End Sub

60     Private Sub cmdShowOptions_Click()
          Call frmOptions.Show
        End Sub

```

65

End Sub

5 Private Sub cmdShowOutput_Click()
 Call frmOutPut.Show
 End Sub

10 Private Sub cmdShowStatistics_Click()
 Call frmStatistics.Show
 End Sub

15 Private Sub MDIForm_Load()
 Call InitControls
 Call frmInput.Show
 End Sub

20 Private Sub InitControls()

25 Dim INextLeft As Long
 Const ISpace As Long = 60
 Const IButtonWidth As Long = 850
 Const IButtonHeight As Long = 770

30 INextLeft = ISpace + IButtonWidth

35 With cmdShowInput
 .Top = ISpace
 .Left = ISpace
 .Width = IButtonWidth
 .Height = IButtonHeight
 End With

40 With cmdShowOutput
 .Top = ISpace
 .Left = INextLeft + ISpace
 45 .Width = IButtonWidth
 .Height = IButtonHeight
 End With

50 With cmdShowDisplay
 .Top = ISpace
 .Left = INextLeft * 2 + ISpace
 55 .Width = IButtonWidth
 .Height = IButtonHeight
 End With

60 With cmdShowStatistics
 .Top = ISpace
 .Left = INextLeft * 3 + ISpace

65

```

5      .Width = IButtonWidth
      .Height = IButtonHeight
      End With

      With cmdShowOptions
10     .Top = ISpace
        .Left = INextLeft * 4 + ISpace
        .Width = IButtonWidth
        .Height = IButtonHeight
      End With

15     cbarMain.Height = IButtonHeight + 2 * ISpace

20     End Sub

      Private Sub mnuMain_Click(Index As Integer)

25         Select Case Index
            Case 0
                Call frmInput.Show
            Case 6
30             Unload Me
        End Select

      End Sub

35     End Sub

      Private Sub mnuResults_Click(Index As Integer)

40         Select Case Index
            Case 0 'Output
                Call frmOutPut.Show
            Case 1
45             Call frmDisplay.Show
            Case 2 'Statistische Auswertung
                Call frmStatistics.Show
        End Select

50     End Sub

      Private Sub mnuExtras_Click(Index As Integer)

55         Select Case Index
            Case 0 'Output

60             Case 1
                Call frmDisplay.Show
            Case 3

65

```

```

    Call frmAbout.Show(vbModal)
End Select

```

```

5 End Sub

```

```

10 Private Sub rnuWindows_Click(Index As Integer)

```

```

    Select Case Index
    Case 0 Horizontal
    Me.Arrange (vbTileHorizontal)
15 Case 1 Vertical
    Me.Arrange (vbTileVertical)
    Case 2 Kaskadieren
    Me.Arrange (vbCascade)
20 End Select

```

```

End Sub

```

```

25

```

```

Name: frmInput.frm

```

```

Code:

```

```

30 VERSION 5.00

```

```

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

```

```

Object = "{3B7C8863-D78F-101B-B9B5-040210D09402}# 1.2# 0"; "RICHTEXT32.OCX"

```

```

Begin VB.Form frmInput

```

```

35 Caption = "Input"

```

```

ClientHeight = 5730

```

```

ClientLeft = 60

```

```

ClientTop = 345

```

```

40 ClientWidth = 13200

```

```

Icon = "frmInput.frx":0000

```

```

LinkTopic = "Form1"

```

```

MDIChild = -1 True

```

```

45 ScaleHeight = 5730

```

```

ScaleWidth = 13200

```

```

WindowState = 2 Maximized

```

```

Begin RichTextLib.RichTextBox txtFormatted

```

```

50 Height = 1815

```

```

Left = 120

```

```

TabIndex = 3

```

```

Top = 360

```

```

55 Width = 12735

```

```

_ExtentX = 22463

```

```

_ExtentY = 3201

```

```

_Version = 393217

```

```

60 BorderStyle = 0

```

```

Enabled = -1 True

```

```

65

```

```

5      ScrollBars      = 2
      DisableNoScroll = -1 True
      TextRTF         = "$frmInput.frx".030A
      BeginProperty Font {0BE35203-8F91-11CE-9DE3-00AA004BB851}
10         Name       = "Fixedsys"
         Size        = 9
         CharSet     = 0
         Weight      = 400
         Underline   = 0 False
         Italic      = 0 False
15         Strikethrough = 0 False
      EndProperty
      End
      Begin VB.OptionButton optSequence
20         Caption    = "Formatted"
         Height     = 255
         Index      = 1
         Left       = 2760
         Style      = 1 'Graphical
25         TabIndex  = 2
         Top        = 0
         Value      = -1 True
         Width     = 855
      End
      Begin VB.OptionButton optSequence
30         Caption    = "Original"
         Height     = 255
         Index      = 0
         Left       = 1920
         Style      = 1 'Graphical
35         TabIndex  = 1
         Top        = 0
         Width     = 855
      End
      Begin VB.CommandButton cmdLoad
40         Caption    = "Load Sequence"
         Height     = 285
         Left       = 0
         TabIndex  = 0
45         Top        = 0
         Width     = 1695
      End
      End
      Begin MSComDlg.CommonDialog CommonDialog1
50         Left      = 3840
         Top       = -120
         _ExtentX  = 847
55         _ExtentY = 847
         _Version  = 393216
60
65

```

```
End
End
Attribute VB_Name = "frmInput"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = False
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
Option Explicit

Private Sub Form_Activate()
    Me.ZOrder
    Call InitControls
End Sub

Public Sub LoadSequence()

Dim sFile As String
Dim sTempCode As String

    On Error GoTo ErrorHandler
    With CommonDialog1
        .FileName = App.Path & "\SampleRNA.txt"
        .CancelError = True
        Call ShowOpen
    End With
    If Len(CommonDialog1.FileName) > 0 Then
        gbSequenceChanged = True
        gsFormattedCode = ""
        gsOriginalCode = ""
        sFile = CommonDialog1.FileName
        Open sFile For Input As # 1
        Input # 1, gsOriginalCode
        Close # 1
        Call FormatText
        Call FillSequence

        mdiMain.cmdShowDisplay.Enabled = True
        mdiMain.cmdShowOutput.Enabled = True
        mdiMain.cmdShowStatistics.Enabled = True
    End If

Exit Sub

ErrorHandler:

Dim lAnswer As Long
```

65

```

    If Err.Number <> cdlCancel Then
        lAnswer = MsgBox("Datei konnte nicht geladen werden", vbRetryCancel, "Dateiproblem")
    If lAnswer = vbRetry Then
        Resume
    End If
End If

```

```
End Sub
```

```
Private Sub FormatText()
```

```
    Dim iLoop As Integer
    Dim sActualBase As String
```

```

    glCodeLength = Len(gsOriginalCode)
    For iLoop = 0 To glCodeLength - 1
        sActualBase = Mid(gsOriginalCode, iLoop + 1, 1)
        If iLoop Mod 3 = 0 Then
            gsFormattedCode = gsFormattedCode + " "
            glAminoCounter = glAminoCounter + 1
        End If
        Select Case sActualBase
            Case "A", "a"
                gsFormattedCode = gsFormattedCode + "A"
                glAdenin = glAdenin + 1
            Case "U", "u", "T", "t"
                gsFormattedCode = gsFormattedCode + "U"
                glThymin = glThymin + 1
            Case "G", "g"
                gsFormattedCode = gsFormattedCode + "G"
                glGuanin = glGuanin + 1
            Case "C", "c"
                gsFormattedCode = gsFormattedCode + "C"
                glCytosin = glCytosin + 1
        End Select
    Next iLoop
    gsFormattedCode = Trim(gsFormattedCode)

```

```
End Sub
```

```
Private Sub FillSequence()
```

```

    If optSequence(0).Value = True Then
        txtFormatted.Text = gsOriginalCode
    Else
        txtFormatted.Text = gsFormattedCode
    End If

```

End Sub

5 Private Sub cmdShowOutput_Click()
 Call frmOutPut.Show
 End Sub

10 Private Sub cmdLoad_Click()
 Call LoadSequence
 End Sub

15 Private Sub InitControls()

Const lSpace As Long = 40
 Const lButtonHeight As Long = 285

20 With cmdLoad
 .Top = lSpace
 Height = lButtonHeight
 Left = lSpace
 End With

25 With optSequence(0)
 .Top = lSpace
 Height = lButtonHeight
 Left = cmdLoad.Left + cmdLoad.Width + lSpace + 200
 End With

30 With optSequence(1)
 .Top = lSpace
 Height = lButtonHeight
 Left = optSequence(0).Left + optSequence(0).Width
 End With

35 With txtFormatted
 .Top = cmdLoad.Height + 2 * lSpace
 Left = lSpace
 .Width = Me.ScaleWidth
 Height = Me.ScaleHeight - txtFormatted.Top
 End With

40 End Sub

45 Private Sub optSequence_Click(Index As Integer)
 Call FillSequence
 End Sub

50 Name: frmOutPut.frm

55

```

Code:
VERSION 5.00
Object = "{3B7C8863-D78F-101B-B9B5-04021CD09402}#1.2#0"; "RIGHTX32.OCX"
5
Begin VB.Form frmOutPut
    Caption       = "OutPut"
    ClientHeight  = 9120
    ClientLeft    = 60
    ClientTop     = 345
    ClientWidth   = 10215
    Icon          = "frmOutPut.frx".0000
    LinkTopic     = "Form1"
    MDIChild      = -1 True
    ScaleHeight   = 9120
    ScaleWidth    = 10215
    Begin RichTextLib.RichTextBox txtOptimized
20
        Height     = 2655
        Left       = 0
        TabIndex   = 0
        Top        = 480
        Width      = 9855
        _ExtentX   = 17383
        _ExtentY   = 4683
        _Version    = 393217
        _Enabled   = -1 True
        ScrollBars = 2
        DisableNoScroll = -1 True
        TextRTF    = "$*frmOutPut.frx".0442
    BeginProperty Font {0BE35203-8F91-11CE-9DE3-00AA004BB851}
35
        Name       = "Fixedsys"
        Size      = 9
        Charset    = 0
        Weight     = 400
        Underline  = 0 False
        Italic     = 0 False
        Strikethrough = 0 False
    EndProperty
45
    End
End
Attribute VB_Name = "frmOutPut"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = False
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
Option Explicit
55
Private Sub Form_Activate()

    Me.ZOrder
60
65

```

```

Call InitControls
If gbSequenceChanged = True Then
  Call BuildAminoChain(gsFormattedCode, False)
End If
gbSequenceChanged = False
txtOptimized.Text = LCase(ReverseTranscription)
End Sub

```

```

Private Sub InitControls()

Const lSpace As Long = 40

  With txtOptimized
    .Top = lSpace
    .Left = lSpace
    .Width = Me.ScaleWidth + lSpace
    .Height = Me.ScaleHeight + lSpace
  End With
End Sub

```

```

Name: frmDisplay.frm
Code:
VERSION 5.00
Object = "*\ACurevac_Genetic_Controls.vbp"
Begin VB.Form frmDisplay
  Caption = "Display"
  ClientHeight = 8205
  ClientLeft = 60
  ClientTop = 345
  ClientWidth = 8745
  Icon = "frmDisplay.frx":0000
  LinkTopic = "Form1"
  MDIChild = -1 'True
  ScaleHeight = 8205
  ScaleWidth = 8745
  ShowInTaskbar = 0 'False
  Begin VB.OptionButton optNucleotids
    Caption = "Original"
    Height = 255
    Index = 0
    Left = 0
    Style = 1 'Graphical
    TabIndex = 1
    Top = 0

```

```

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

40 Private Sub Form_Activate()

    Me.ZOrder
    If gbSequenceChanged = True Then
45         Call BuildAminoChain(gsFormattedCode, True)
    End If
    gbSequenceChanged = False

50 End Sub

Private Sub optNucleotids_Click(Index As Integer)
55

Dim bShowOriginal As Boolean

    If Index = 0 Then 'Original Clicked
60         bShowOriginal = True
    End If

```

65

```

Else 'Optimized Clicked
  bShowOriginal = False
End If
Call RebuildChain(bShowOriginal)

```

End Sub

```

Name: frmStatistics.frm
Code:
VERSION 5.00
Begin VB.Form frmStatistics
  Caption       = "Statistics"
  ClientHeight  = 6450
  ClientLeft    = 60
  ClientTop     = 345
  ClientWidth   = 8595
  Icon          = "frmStatistics.frm":0000
  LinkTopic     = "Form1"
  MDIChild      = -1 'True
  ScaleHeight   = 6450
  ScaleWidth    = 8595
  Begin VB.Frame Frame1
    Caption      = "original"
    Height       = 2655
    Left         = 120
    TabIndex     = 9
    Top          = 720
    Width        = 2175
    Begin VB.Label lblAdenin
      BorderStyle = 1 'Fixed Single
      Height      = 495
      Left        = 840
      TabIndex    = 17
      Top         = 240
      Width       = 1215
    End
    Begin VB.Label lblCytosin
      BorderStyle = 1 'Fixed Single
      Height      = 495
      Left        = 840
      TabIndex    = 16
      Top         = 2040
      Width       = 1215
    End
    Begin VB.Label lblGuanin
      BorderStyle = 1 'Fixed Single

```

```

5      Height      = 495
        Left       = 840
        TabIndex   = 15
        Top        = 1440
        Width      = 1215
        End
10     Begin VB.Label lblThymin
        BorderStyle = 1 Fixed Single
        Height      = 495
        Left       = 840
15     TabIndex   = 14
        Top        = 840
        Width      = 1215
        End
20     Begin VB.Label Label14
        Caption     = "Cytosin"
        Height      = 255
        Left       = 120
25     TabIndex   = 13
        Top        = 2160
        Width      = 615
        End
30     Begin VB.Label Label2
        Caption     = "Thymin"
        Height      = 255
        Left       = 120
35     TabIndex   = 12
        Top        = 960
        Width      = 615
        End
40     Begin VB.Label Label3
        Caption     = "Guanin"
        Height      = 255
        Left       = 120
45     TabIndex   = 11
        Top        = 1560
        Width      = 615
        End
50     Begin VB.Label Label1
        Caption     = "Adenin"
        Height      = 255
        Left       = 120
55     TabIndex   = 10
        Top        = 360
        Width      = 615
        End
60     End
        Begin VB.Frame Frame2

```

65

```
5      Caption    = "optimized"
      Height     = 2655
      Left       = 2520
      TabIndex   = 0
      Top        = 720
      Width      = 2175
10     Begin VB.Label Label4
      Caption    = "Adenin"
      Height     = 255
      Left       = 120
      TabIndex   = 8
      Top        = 360
      Width      = 615
15     End
      Begin VB.Label Label6
      Caption    = "Cytosin"
      Height     = 255
      Left       = 120
      TabIndex   = 7
      Top        = 2160
      Width      = 615
20     End
      Begin VB.Label Label8
      Caption    = "Guanin"
      Height     = 255
      Left       = 120
      TabIndex   = 6
      Top        = 1560
      Width      = 615
25     End
      Begin VB.Label Label12
      Caption    = "Thymin"
      Height     = 255
      Left       = 120
      TabIndex   = 5
      Top        = 960
      Width      = 615
30     End
      Begin VB.Label lblOptAdenin
      BorderStyle = 1 'Fixed Single
      Height     = 495
      Left       = 840
      TabIndex   = 4
      Top        = 240
      Width      = 1215
35     End
      Begin VB.Label lblOptThymin
      BorderStyle = 1 'Fixed Single
40     End
45
50
55
60
65
```

```
5      Height      = 495
      Left        = 840
      TabIndex    = 3
      Top         = 840
      Width       = 1215
      End
10     Begin VB.Label lblOptGuanin
      BorderStyle = 1 Fixed Single
      Height      = 495
      Left        = 840
      TabIndex    = 2
      Top         = 1440
      Width       = 1215
      End
15     Begin VB.Label lblOptCytosin
      BorderStyle = 1 Fixed Single
      Height      = 495
      Left        = 840
      TabIndex    = 1
      Top         = 2040
      Width       = 1215
      End
20     End
      End
25     Begin VB.Label lblBases
      BorderStyle = 1 Fixed Single
      Height      = 495
      Left        = 1320
      TabIndex    = 19
      Top         = 120
      Width       = 1215
      End
30     End
      End
35     Begin VB.Label Label5
      Caption     = "Sum of bases"
      Height      = 255
      Left        = 120
      TabIndex    = 18
      Top         = 240
      Width       = 1095
      End
40     End
      End
45     Attribute VB_Name = "frmStatistics"
      Attribute VB_GlobalNameSpace = False
      Attribute VB_Creatable = False
      Attribute VB_PredeclaredId = True
      Attribute VB_Exposed = False
      Option Explicit
50     End
      End
55     Private Sub Form_Activate()
60
65
```

Me.ZOrder

```

5      lblBases.Caption = CStr(glAminoCounter * 3)
      lblAdenin.Caption = CStr(glAdenin)
      lblThymin.Caption = CStr(glThymin)
      lblGuanin.Caption = CStr(glGuanin)
10     lblCytosin.Caption = CStr(glCytosin)

      Call CountBases(frmOutPut.txtOptimized.Text)
      lblOptAdenin.Caption = CStr(glOptAdenin)
      lblOptThymin.Caption = CStr(glOptThymin)
15     lblOptGuanin.Caption = CStr(glOptGuanin)
      lblOptCytosin.Caption = CStr(glOptCytosin)

20     End Sub

```

```

25     Name: frmOptions.frm
      Code:
      VERSION 5.00
      Begin VB.Form frmOptions
30         Caption       = "Options"
         ClientHeight   = 3915
         ClientLeft     = 60
         ClientTop      = 345
         ClientWidth    = 5760
35         Icon         = "frmOptions.frx":0000
         LinkTopic      = "Form1"
         MDIChild       = -1 'True
         ScaleHeight    = 3915
         ScaleWidth     = 5760
40         End
      Attribute VB_Name = "frmOptions"
      Attribute VB_GlobalNameSpace = False
45     Attribute VB_Creatable = False
      Attribute VB_PredeclaredId = True
      Attribute VB_Exposed = False
      Option Explicit
50

```

```

55     Name: frmAbout.frm
      Code:
      VERSION 5.00
      Begin VB.Form frmAbout
60         BorderStyle   = 3 Fixed Dialog

```

65

```

Caption      = "About Curevac RNA-Analyzer"
ClientHeight = 2490
ClientLeft   = 2340
ClientTop    = 1935
ClientWidth  = 6195
ClipControls = 0 False
LinkTopic    = "Form2"
MaxButton    = 0 False
MinButton    = 0 False
ScaleHeight  = 1718.642
ScaleMode    = 0 User
ScaleWidth   = 5817.425
ShowInTaskbar = 0 False
Begin VB.PictureBox picIcon
    AutoSize    = -1 True
    ClipControls = 0 False
    Height      = 1215
    Left        = 120
    Picture     = "frmAbout.frx":0000
    ScaleHeight = 811.195
    ScaleMode   = 0 User
    ScaleWidth  = 2401.98
    TabIndex    = 1
    Top         = 120
    Width       = 3480
End
Begin VB.CommandButton cmdOK
    Cancel      = -1 True
    Caption     = "OK"
    Default     = -1 True
    Height      = 345
    Left        = 2520
    TabIndex    = 0
    Top         = 2040
    Width       = 1260
End
Begin VB.Label lblCompany
    Caption     = "lblCompany"
    Height      = 255
    Left        = 3720
    TabIndex    = 5
    Top         = 1200
    Width       = 2295
End
Begin VB.Line Line1
    BorderColor = &H1D0808080&
    BorderStyle = 6 Inside Solid
    Index       = 1

```

ES 2 340 499 T3

```

X1      = 112.686
X2      = 5746.996
Y1      = 1325.218
Y2      = 1325.218
End
Begin VB.Label lblDescription
Caption  = "App Description"
ForeColor = &H100000000&
Height  = 330
Left    = 120
TabIndex = 2
Top     = 1560
Width   = 5925
End
Begin VB.Label lblTitle
Caption  = "Application Title"
ForeColor = &H100000000&
Height  = 360
Left    = 3720
TabIndex = 3
Top     = 240
Width   = 2325
End
Begin VB.Line Line1
BorderColor = &H100FFFFFF&
BorderWidth = 2
Index      = 0
X1         = 112.686
X2         = 5746.996
Y1         = 1325.218
Y2         = 1325.218
End
Begin VB.Label lblVersion
Caption  = "Version"
Height  = 225
Left    = 3720
TabIndex = 4
Top     = 720
Width   = 2325
End
End
Attribute VB_Name = "frmAbout"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = False
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
Option Explicit

```

5

10

15

20

25

30

35

40

45

50

55

60

65

```
Private Sub cmdOK_Click()  
    Unload Me  
End Sub
```

5

```
Private Sub Form_Load()  
    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  
End Sub
```

10

15

20

25

30

35

40

45

50

55

60

65

REIVINDICACIONES

1. ARNm modificado que codifica como mínimo un péptido o un polipéptido antígeno tumoral, **caracterizado** porque el contenido de G/C de la zona del ARNm modificado que codifica el péptido o el polipéptido es mayor que el contenido de G/C del campo de codificación del ARNm del tipo salvaje que codifica el péptido o el polipéptido, y porque 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 de la zona del ARNm modificado que codifica el péptido o el 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 el 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 consistente en fosforotioatos, fosforamidatos, nucleótidos de péptidos, metilfosfonatos, 7-desazaguanosina, 5-metilcitosina e inosina.
6. ARNm modificado según una de las reivindicaciones 1 a 5, **caracterizado** porque el ARNm codifica un antígeno de superficie específico de tumores.
7. ARNm modificado según una de las reivindicaciones 1 a 6, **caracterizado** porque el ARNm modificado codifica un antígeno tumoral seleccionado de entre el grupo consistente en 707-AP, AFP, ART-4, BAGE, β -catenina/m, Bcr-abl, CAMEL, CAP-1, CASP-8, CDC27/m, CDK4/m, CEA, CT, Cyp-B, DAM, ELF2M, ETV6-AML1, G250, GAGE, GnT-V, Gp100, HAGE, HER-2/neu, HLA-A*0201-R1701, HPV-E7, HSP70-2M, HAST-2, hTERT (o hTRT), iCE, KIAA0205, LAGE, LDLR/FUT, MAGE, MART-1/Melan-A, MC1R, Miosina/m, MUC1, MUM-1, -2, -3, NA88-A, NY-ESO-1, p190 minor bcr-abl, Pml/RAR α , PRAME, PSA, PSM, RAGE, RU1 ó RU2, SAGE, SART-1 ó SART-3, TEL/AML1, TPI/m, TRP-1, TRP-2, TRP-2/INT2 y WT1.
8. ARNm modificado según una de las reivindicaciones 1 a 7, **caracterizado** porque el polipéptido es un polipeptido de antígenos de tumores.
9. ARNm modificado según una de las reivindicaciones 1 a 8, **caracterizado** porque el ARNm codifica, además, como mínimo una citoquina.
10. ARNm modificado según una de las reivindicaciones 1 a 9, **caracterizado** porque el ARNm está asociado o unido a una proteína o un péptido catiónico.
11. ARNm modificado según la reivindicación 10, **caracterizado** porque la proteína o el péptido catiónico se selecciona de entre el grupo consistente en protamina, poli-L-lisina e histonas.
12. Composición farmacéutica **caracterizada** porque contiene el ARNm modificado según una de las reivindicaciones 1 a 11 junto con un excipiente y/o vehículo farmacéuticamente compatible.
13. Composición farmacéutica según la reivindicación 12, **caracterizada** porque contiene como mínimo un adyuvante que estimula la respuesta inmunológica.
14. Composición farmacéutica según una de las reivindicaciones 12 ó 13, que contiene como mínimo además una citoquina.
15. Utilización de una composición farmacéutica según una de las reivindicaciones 12 a 14 o de un ARNm modificado según una de las reivindicaciones 1 a 11 para la preparación de una vacuna para la inoculación contra el cáncer.
16. Utilización de una composición farmacéutica según las reivindicaciones 12 a 14 para la preparación de un medicamento para el tratamiento de melanomas, carcinomas del colon, linfomas, carcinomas de pulmón de células pequeñas, sarcomas ó blastomas.

Fig. 1A

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

```
agatctaaagatgagtccttctaaccgaggtcgaaacgtaegttctcteta
tcatcccgtcaggccccctcaaagccgagatcgcacagagacttgaagat
gtctttgcagggagaacaccgatcttgaggttctcatggaatggctaaa
gacaagaccaatcctgtcacctctgactaaggggattttaggatttgtgt
tcacgetcacegtgccagtgagcgaggactgcagcgtagacgctttgtc
caaaatgcccttaatgggaacggggatccaaataacatggacaaagcagt
taaactgtataggaagctcaagagggagataacattccatggggccaaag
aaatctcactcagttattctgctgggtgcacttgccagttgtatgggcctc
atatacaacaggatgggggctgtgaccactgaagtggcatttggcctggt
atgtgcaacctgtgaacagattgctgactcccagcatcggtctcataggc
aaatggtgacaacaaccaaccactaatcagacatgagaacagaatggtt
ttagccagcactacagctaaggctatggagcaaattggctggatcgagtgga
gcaagcagcagaggccatggagggttgctagtcaggctaggcaaattggtgc
aagcgatgagaaccattgggactcatcctagctccagtgctggtctgaaa
aatgatcttcttgaaaatttgcaggcctatcagaaacgaatggggggtgca
gatgcaacggttcaagtgaactag
```

Fig. 1B

Matriz de la influenza: Secuencia de proteínas

```
MSLLTEVETYVLSIIPSGPLKAEIAQRLEDVDFAGKNTDLEVLMEWLKTRP
ILSPLTKGILGFVFTLTVPSERGLQRRRFVQNALNGNGDPNNMDKAVKLY
RCLKREITFHGAKEISLSYSAGALASCMGLIYNRMGAVTTEVAFGLVCAT
CEQIADSQHRSHRQMVTTTNPLIRHENRMVLASTTAKAMEQMAGSSEQAA
EAMEVASQARQMVQAMRTIGTHPSSSAGLKNLDLENLQAYQKRMGVQMQR
FK*
```

Fig. 1C

Matriz de la influenza: Gen con un contenido de G/C aumentado

agatctaaagatgagCctGctGaccgaggtGgaGacCtacgtGctGAGCa
 tcatcccCAGCggccccctGaaGgccgagatcgcCcagagGctGgaGgaC
 gtGttCgcCggCaagaacaccgaCctGgaggtGctGatggaGtggctGaa
 gacCagGccCatcctgAGCccCctgacCaagggCatCCTGggCttCgtgt
 tcacCctGaccgtgcccagCgagcgCggCctgcagcgCCGccgcttCgtG
 caGaaCgcccctGaaCggCaacggCgaCccCaaCaacatggacaaGgcCgt
 GaaGctgtaCaggaagctGaagagggagatCacCttccaCggCgccaaGg
 aGatcAGCctGagCtaCAGCgcCggCgcCctGgccagCtgCatgggcctG
 atCtacaacaggatgggCgcCgtgaccacCgaGgtggcCttCggcctggt
 GtgCgcCacctgCgaGcagatCgcCgacAGCcagcaCcgCAGCcaCaggc
 aGatggtgacCacCaccaaccccCctGatcagGcaCgagaacagGatggtG
 CTGgccagcacCacCgcCaagggCatggagcaGatggcCggCAGCaGCga
 gcaGgcCgcCgaggccatggaggtGgcCagCagggcCaggcaGatggtgc
 aGgcCatgagGaccatCggCacCcaCccCagcAGCagCgcCggCctgaaG
 aaCgaCctGctGgaGaaCCTGcagggcctaCcaGaaGcgCatgggCgtgca
 gatgcaGcgCttcaagtgaactagt

Fig. 1D

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

AgatctaaagatgGCCGTCATGGCCCCCGCACCCCTGGTGCTGCTGCTGA
 GCGGCGCCCTGGCCCTGACCCAGACCTGGGCTagCctGctGaccgaggtG
 gaGacCtacgtGctGAGCatcatcccCAGCggccccctGaaGgccgagat
 cgcCcagagGctGgaGgaCgtGttCgcCggCaagaacaccgaCctGgaggt
 GctGatggaGtggctGaagacCagGccCatcctgAGCccCctgacCaag
 ggCatCCTGggCttCgtgttcacCctGaccgtgcccagCgagcgCggCct
 gcagcgCCGccgcttCgtGcaGaaCgcccctGaaCggCaacggCgaCccCa
 aCaacatggacaaGgcCgtGaaGctgtaCaggaagctGaagagggagatC
 acCttccaCggCgccaaGgaGatcAGCctGagCtaCAGCgcCggCgcCct
 GgccagCtgCatgggcctGatCtacaacaggatgggCgcCgtgaccacCg
 aGgtggcCttCggcctggtGtgCgcCacctgCgaGcagatCgcCgacAGC
 cagcaCcgCAGCcaCaggcaGatggtgacCacCaccaaccccCctGatcag
 GcaCgagaacagGatggtGCTGgccagcacCacCgcCaagggCatggagc
 aGatggcCggCAGCaGCgagcaGgcCgcCgaggccatggaggtGgcCagC
 cagggcCaggcaGatggtgcaGgcCatgagGaccatCggCacCcaCccCag
 cAGCagCgcCggCctgaaGaaCgaCctGctGgaGaaCCTGcagggcctaC
 agaaGcgCatgggCgtgcagatgcaGcgCttcaagtgaactagt

Fig. 1E

Matriz de la influenza: ARNm con secuencias de estabilización

GCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUCAACUUUGGCagauc
uaaagangagucuucuaaccgaggucgaaacguacguucucucuauc
ccgucaggccccucaagccgagaucgcacagagacuugaagaugucuu
ugcagggagaacaccgaucuuagguucucauggaauaggcuaaagaca
gaccaauccugucaccucugacuaaggggauuuuaggauuuguguuacg
cucaccgugcccagugagcgaggacugcagcguagacgcuuuguccaaa
ugcccuuaauugggaacggggaucuaaauacauggacaagcaguuaaac
uguauaggaagcucaagaggggagauaacauuccauggggccaaagaauc
ucacucaguuaucucgucggugcacuugccaguuguauggggccucaua
caacaggaugggggucugacacugaaguggcauuuggccugguaugug
caaccugugaacagauugcugacuccagcaucggucucauaggcaau
gugacaacaaccaacccacuaaucagacaugagaacagaaugguuuagc
cagcacuacagcuauaggcuauaggagcaauaggcuggaucgagugagcaag
cagcagaggccauggagguugcuagucaggcuaggcaauaggugcaagcg
augagaaccuugggacucauccuagcuccagugcuggucugaaaauga
ucuucuuagaaaauuugcaggccuaucaagaaacgaauugggggugcagaugc
aacgguucaagugaACUAGUGACUGACUAGCCCGUGGGCCUCCCAACGG
GCCUCCUCCCUCCUUGCACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AA

Fig. 1F

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

GCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUCAACUUUGGCagauc
uaaagaugagCcuGcuGaccgagguGgaGacCuacguGcuGAGCaucauc
ccCAGCggccccuGaaGgccgagaucgcCagagGcuGgaGgaCguGuu
CgcCggCaagaacaccgaCcuGgagguGcuGauggaGuggcuGaagacCa
gGccCauccugAGCccCcuGacCaagggCauCCUGggCuuCguguuacC
cuGaccgugcccagCgagcgCggCcuGcagcgCCGCcgcuuCguGcaGaa
CgccccuGaaCggCaacggCgaCccCaaCaacauggacaGgcCguGaaGc
uguaCaggaagcuGaagaggggagauCacCuuccaCggCgccaGgaGauc
AGCcuGagCuaCAGCgcCggCgcCcuGgccagCugCauggggccuGauCua
caacaggaugggCgcCgugaaccacCgaGguggcCuuCggccugguGugCg
cCaccugCgaGcagauCgcCgacAGCagcaCcgCAGCcaCaggcaGaug
gugacCacCaccaacccCcuGaucagGcaCgagaacagGaugguGCUGgc
cagcacCacCgcCaaggcCauggagcaGauggcCggCAGCaGCgagcaGg
cCgcCgaggccauggagguGgcCagCcaggcCaggcaGauggugcaGgcC
augagGaccuucggCacCcaCccCagcAGCagCgcCggCcuGaaGaaCga
CcuGcuGgaGaaCCUGcaggccuaCagaaGcgCaugggCgugcagaugc
aGcgCuucaagugaACUAGUGACUGACUAGCCCGUGGGCCUCCCAACGG
GCCUCCUCCCUCCUUGCACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AA

Fig. 1G

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

GCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUCAACUUUGGCagauc
uaaagaugGCCGUCAUGGCCCCCCGCACCCUGGUGCUGCUGCUGAGCGGC
GCCCUGGCCUGACCCAGACCUGGGCCagCcuGcuGaccgagguGgaGac
CuacguGcuGAGCaucaucccCAGCggcccccGaaGgcccagaucgcCc
agagGcuGgaGgaCguGuuCgcCggCaagaacaccgaCcuGgagguGcuG
auggaGuggcuGaagacCagGccCauccugAGCccCcugacCaagggCau
CCUGggCuuCguguucacCcuGaccgugcccagCgagcgCggCcuqcagc
gCCGCcgcuuCguGcaGaaCgcccuGaaCggCaacggCgaCccCaaCaac
auggacaaGgcCguGaaGcuguaCaggaagcuGaagagggagauCacCuu
ccaCggCgccaaGgaGaucAGCcuGagCuaCAGCgcCggCgcCcuGgcca
gCugCaugggccuGauCuacaacaggaugggCgcCgugaccacCgaGgug
gcCuuCggccugguGugCgcCaccugCgaGcagauCgcCgacAGCcagca
CcgCAGCcaCaggcaGauggugacCacCaccaaccccCcuGaucagGcaCg
agaacagGaugguGCUGgccagcacCacCgcCaaggcCauggagcaGaug
gcCggCAGCaGCgagcaGgcCgcCgaggccauggagguGgcCagCcaggc
CaggcaGauggugcaGgcCaugagGaccauCggCacCcaCccCagcAGCa
gCgcCggCcugaaGaaCgaCcuGcuGgaGaaCCUGcaggccuaCcagaaG
cgCaugggCgugcagaugcaGcgCuucaagugaACUAGUGACUGACUAGC
CCGCUGGGCCUCCCAACGGGCCCUCCUCCCCUCCUUGCACCAAAAAAAAAA
AA
AAAAAAAAAAAA

Fig. 2A

MAGE1: Gen del tipo salvaje (para comparación)

catcatgtctcttgagcagaggagtctgcactgcaagcctgaggaagccc
 ttgaggcccaacaagaggccctgggcctggtgtgtgtgcaggctgccacc
 tctctctctctctctggtcctgggcaccctggaggaggtgccactgc
 tgggtcaacagatcctccccagagtcctcagggagcctccgcctttcca
 ctaccatcaactcactcgacagaggcaaccagtgagggtccagcagc
 cgtgaagaggagggccaagcacctcttgatcctggagtccttgttccg
 agcagtaatcactaagaaggtggctgatttggttggtttctgctcctca
 aatatcgagccaggagccagtcacaaaggcagaaatgctggagagtgc
 atcaaaaattacaagcactgttttctgagatcttcggcaaagcctctga
 gtccttgacagctggtcttggcattgacgtgaaggaagcagacccccaccg
 gccactcctatgtccttgtcacctgectaggtctctcctatgatggcctg
 ctgggtgataatcagatcatgcccaagacaggcttcttgataattgtcct
 ggtcatgattgcaatggagggcgccatgctcctgaggaggaatctggg
 aggagctgagtgatggaggtgtatgatgggagggagcacagtgcctat
 ggggagcccaggaagctgctcacccaagatttggtgcaggaaaagtacct
 ggagtaccggcaggtgccggacagtgatcccgcacgctatgagttcctgt
 ggggtccaagggccctcgctgaaaccagctatgtgaaagtccttgagtat
 gtgatcaaggtcagtgcaagagttcgcttttcttcccacccctgcgtga
 agcagctttgagagaggaggaagagggtctgagcatga

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, THR, 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, GLY, 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, VAL, 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, PRO, 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, VAL, 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, ALA, 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, GLU, 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, LEU, 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, SER, LEU, ARG, GLU, ALA, ALA, LEU, ARG, GLU, GLU, GLU, GLU, GLY, VAL, STOP, ALA, STOP

Fig. 2C

MAGE1: ARNm con un contenido de G/C aumentado

augagccuggagcagcgcagccugcacugcaagccggaggaggcgucuggaggcgagcagga
 ggcgucgggcccuggucugcguccaggcggcgacgagcagcagcagcccgcugguccugggca
 cgcuggaggaggucccgacggcggcagcacggacccgcgcagagcccgcaggggcgagc
 gcuucccgacgacgaucaacuucacgcgccagccagccgagcgaggggcagcagcagccg
 cgaggaggaggggcccagcagcagcugcauccuggagagccuguuuccgcgggucaucacga
 agaaggucgcggaccuggucggcuuccugcugcugaaguaaccgcgcgcgcgagccggucacg
 aaggcggagaugcuggagagcgucacuaagaacuacaagcacugcuucccgagauucuucgg
 caaggcgagcgagagccugcagcuggucuuucggcaucgacgucuaaggaggcggacccgacgg
 gccacagcuacguccuggucacgugccugggcccugagcuaacgacggccugcuggggcgacaac
 cagaucaugcccgaagacgggcuuccugaucaucguccuggucaugaucgcgauggagggcg
 ccacgcgccggaggaggagaucugggaggagcugagcgucuauggaggucacgacggcccg
 agcacagcgcguacggcgagcccgcgcaagcugcugacgcaggaccugguccaggagaaguac
 cuggaguaccgccaggucccgacagcagcccggcgcgcuaacgaguuccuguggggcccgcg
 cgcgcuggcggagacgagcuacgucuaaggucuccuggagucgucuaaaggucagcgcgcgcg
 uccgcuucuucuucccgagccugcgcgagggcggcgcugcgcgagggaggaggaggggcgucuga
 gcgugauga

Fig. 2D

MAGE1: ARNm con utilización alternativa de codones

augagccuggagcagcgcagccugcacugcaagcccggaggaggcccuggaggcccagcagga
 ggcccugggcccuggugugcgugcaggcccgccaccagcagcagcagccccuggugcugggca
 cccuggaggaggugcccaccgcccggcagcacccgacccccccagagccccagggcgccagc
 gccuuccccaccaccaucaaacuucaccccggcagcgcagcccagcagggcagcagcagccg
 cgaggaggaggggcccagcaccagcugcauccuggagagccuguuuccgcgccgugaucacca
 agaaggugggccgaccuggugggcuuccugcugcugaaguaaccgcgcccgagcccugagacc
 aaggccgagaugcuggagagcgugaucaagaacuacaagcacugcuuccccgagauucuucgg
 caaggccagcagagagccugcagcugguguuucggcaucgacgugaaggaggcccgacccccaccg
 gccacagcuacgugcuggugaccugccugggcccugagcuaacgacggccugcuggggcgacaac
 cagaucaugcccgaagaccggcuuccugaucaucgugcuggugaugaucgccauggagggcg
 ccacgcccccgaggaggagaucugggaggagcugagcgugauggagguguaacgacggcccg
 agcacagcgcuaacggcgagccccgcaagcugcugacccaggaccuggugcaggagaaguac
 cuggaguaccgccaggugcccagcagcagcccggcgcgcuaacgaguuccuguggggcccccg
 cgcccuggccgagaccagcuacgugaaggugcuggaguaacgugaucuaaggugagcgcgccg
 ugcgcuuucuucuuccccagccugcgcgagggcccugcgcgagggaggaggaggggcguguga
 gccugauga

ES 2 340 499 T3

LISTA DE SECUENCIAS

- <110> Von der Muelbe, Florian
Hoerr, Igmarr (solamente EE.UU.)
Pascolo, Steve (solamente EE.UU:)
- <120> Composición farmacéutica que contiene un ARNm estabilizado y optimizado para la traducción en sus campos de codificación.
- <130> CU01P003WO EPT3.
<140> PCT/EP02/06180
<141> 2002-06-05
<160> 13
<170> PatentIn Ver. 2.1
- <210> 1
<211> 774
<212> ADN
<213> Influenza virus
<220>
<223> Influenza-Matrix: Gen del tipo salvaje (para comparación)
<220>
<223> Códon de inicio: atg (Nucleótidos 11 a 13), codón de parada: tga (Nucleótidos 767 a 769)
- <400> 1
- ```
35 agatctaaag atgagttctc taaccgaggt cgaaacgtac gttctctcta tcatcccgtc 60
 aggccccctc aaagccgaga tcgcacagag acttgaagat gtctttgcag ggaagaacac 120
 cgatcttgag gttctcatgg aatggctaaa gacaagacca atcctgtcac ctctgactaa 180
 ggggatttta ggatttgyt tcacgctcac cgtgccaggt gagcgaggac tgcagcgtag 240
40 acgctttgtc caaaatgccc ttaatgggaa cggggatcca aataacatgg acaaagcagt 300
 taaactgtat aggaagctca agagggagat aacattccat ggggccaaag aaatctcact 360
 cagttattct gctggtgcac ttgccagttg tatgggcctc atatacaaca ggatgggggc 420
 tgtgaccact gaagtggcat ttggcctggt atgtgcaacc tgtgaacaga ttgctgactc 480
45 ccagcatcgg tctcataggc aatgggtgac aacaaccaac ccactaatca gacatgagaa 540
 cagaatggtt ttagccagca ctacagctaa ggctatggag caaatggctg gatcgagtga 600
 gcaagcagca gaggccatgg aggttgctag tcaggctagg caaatggctc aagcgatgag 660
50 aaccattggg actcactccta gctccagttc tggcttgaaa aatgatcttc ttgaaaattt 720
 gcaggcctat cagaaacgaa tgggggtgca gatgcaacgg ttcaagtgaa ctag 774
```
- <210> 2  
<211> 252  
<212> PRT  
<213> Influenza virus

ES 2 340 499 T3

<400> 2

5 Met Ser Leu Leu Thr Glu Val Glu Thr Tyr Val Leu Ser Ile Ile Pro  
1 5 10 15

10 Ser Gly Pro Leu Lys Ala Glu Ile Ala Gln Arg Leu Glu Asp Val Phe  
20 25 30

15 Ala Gly Lys Asn Thr Asp Leu Glu Val Leu Met Glu Trp Leu Lys Thr  
35 40 45

20 Arg Pro Ile Leu Ser Pro Leu Thr Lys Gly Ile Leu Gly Phe Val Phe  
50 55 60

25 Thr Leu Thr Val Pro Ser Glu Arg Gly Leu Gln Arg Arg Arg Phe Val  
65 70 75 80

30 Gln Asn Ala Leu Asn Gly Asn Gly Asp Pro Asn Asn Met Asp Lys Ala  
85 90 95

35 Val Lys Leu Tyr Arg Lys Leu Lys Arg Glu Ile Thr Phe His Gly Ala  
100 105 110

40 Lys Glu Ile Ser Leu Ser Tyr Ser Ala Gly Ala Leu Ala Ser Cys Met  
115 120 125

45 Gly Leu Ile Tyr Asn Arg Met Gly Ala Val Thr Thr Glu Val Ala Phe  
130 135 140

50 Gly Leu Val Cys Ala Thr Cys Glu Gln Ile Ala Asp Ser Gln His Arg  
145 150 155 160

55 Ser His Arg Gln Met Val Thr Thr Thr Asn Pro Leu Ile Arg His Glu  
165 170 175

60 Asn Arg Met Val Leu Ala Ser Thr Thr Ala Lys Ala Met Glu Gln Met  
180 185 190

65 Ala Gly Ser Ser Glu Gln Ala Ala Glu Ala Met Glu Val Ala Ser Gln  
195 200 205

70 Ala Arg Gln Met Val Gln Ala Met Arg Thr Ile Gly Thr His Pro Ser  
210 215 220

75 Ser Ser Ala Gly Leu Lys Asn Asp Leu Leu Glu Asn Leu Gln Ala Tyr  
225 230 235 240

80 Gln Lys Arg Met Gly Val Gln Met Gln Arg Phe Lys  
245 250

## ES 2 340 499 T3

<210> 3

<211> 775

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Influenza-Matrix: Gen con un mayor contenido de G/C

10

<220>

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

15 <400> 3

```
20 agatctaaag atgagcctgc tgaccgaggt ggagacctac gtgetgagca tcatccccag 60
 cggccccctg aaggccgaga tcgcccagag gctggaggac gtgttcgccg gcaagaacac 120
 cgacctggag gtgctgatgg agtggctgaa gaccaggccc atcctgagcc cctgaccaa 180
 gggcatcctg ggcttcgtgt tcacctgac cgtgcccagc gagcgcggcc tgcagcgccg 240
25 ccgcttcgtg cagaacgccc tgaacggcaa cggcgacccc aacaacatgg acaaggccgt 300
 gaagctgtac aggaagctga agagggagat caccttccac ggcgccaagg agatcagcct 360
 gagctacagc gccggcgccc tggccagctg catgggcctg atctacaaca ggatgggccc 420
 cgtgaccacc gaggtggcct tcggcctggt gtgcgccacc tgcgagcaga tcgccgacag 480
30 ccagcaccgc agccacaggc agatggtgac caccaccaac cccctgatca ggcaccgagaa 540
 caggatggtg ctggccagca ccaccgcaa ggccatggag cagatggccg gcagcagcga 600
 gcaggccgcc gagggccatgg aggtggccag ccaggccagg cagatggtgc aggccatgag 660
 gaccatcggc acccaaccca gcagcagcgc cggcctgaag aacgacctgc tggagaacct 720
35 gcaggcctac cagaagcgca tgggcgtgca gatgcagcgc ttcaagtga ctagt 775
```

40 <210> 4

<211> 844

<212> ADN

45 <213> Secuencia artificial

<220>

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

50 <220>

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

55

60

65

ES 2 340 499 T3

<400> 4

```

 agatctaaag atggccgtca tggccccccg caccctggtg ctgctgctga gcggcgccct 60
 ggccctgacc cagacctggg ctagcctgct gaccgaggtg gagacctacg tgctgagcat 120
5 catccccagc ggccccctga aggccgagat cgcccagagg ctggaggacg tgttcgccgg 180
 caagaacacc gacctggagg tgctgatgga gtggctgaag accaggccca tcctgagccc 240
 cctgaccaag ggcatectgg gcttcgtggt caccctgacc gtgccagcgg agcgcgccct 300
10 gcagcgccgc cgcttcgtgc agaacgccct gaacggcaac ggcgacccca acaacatgga 360
 caaggccgtg aagctgtaca ggaagctgaa gagggagatc accttcacg gcgccaagga 420
 gatcagcctg agctacagcg ccggcgccct ggccagctgc atgggcctga tctacaacag 480
 gatgggcgcc gtgaccaccg aggtggcctt cggcctggtg tgcgccacct gcgagcagat 540
15 cgccgacagc cagcaccgca gccacaggca gatggtgacc accaccaacc ccctgatcag 600
 gcacgagaac aggatggtgc tggccagcac caccgccaag gccatggagc agatggccgg 660
 cagcagcgag caggccgccc aggccatgga ggtggccagc caggccaggc agatggtgca 720
20 ggccatgagg accatcggca cccaccccag cagcagcgcc ggcctgaaga acgacctgct 780
 ggagaacctg caggcctacc agaagcgcac gggcgtgcag atgcagcgct tcaagtgaac 840
 tagt 844

```

<210> 5

<211> 942

<212> ARN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Influenza-Matrix: ARNm con secuencias de estabilización

<220>

<223> Las secuencias de estabilización provienen de los UTRs de 5' o 3' de la ARNm de  $\beta$ -globina de *Xenopus laevis*

<220>

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

<400> 5

```

 gcuuguucuu uuugcagaag cucagaauaa acgcucaacu uuggcagauc uaaagauagag 60
 ucuucuaacc gaggucgaaa cguacguucu cucuaucauc ccgucaggcc cccucaaaagc 120
45 cgagaucgca cagagacuug aagaugucuu ugcagggag aacaccgauc uugagguucu 180
 cauggaungg cuaaagacaa gaccaauccu gucaccucug acuaagggga uuuuaggauu 240
 uguguucacg cucaccgugc ccagugagcg aggacugcag cguagacgcu uuguccaaaa 300
50 ugcccuaau gggaacgggg auccaaaaua cauggacaaa gcaguuaaac uguauaggaa 360
 gcucaagagg gagauaaca uccauggggc caaagaauc ucacucaguu auucugcugy 420
 ugcacuugcc aguuguaugg gccucauaua caacaggaug ggggcuguga ccacugaagu 480
 ggcuuuuggc cugguaugug caaccuguga acagauugcu gacueccagc aucggucuca 540
55
 uaggcaaaug gugacaacaa ccaaccacu aaucagacau gagaacagaa ugguuuuagc 600
 cagcacuaca gcuaaggcua uggagcaaa uggcuggauc agugagcaag cagcagaggc 660
60 cauggagguu gcuaugcagg cuaggcaaa uggugcaagc augagaacca uugggacuca 720
 uccuagcucc agugcugguc ugaaaaauga ucuucuugaa aauuugcagg ccuaucagaa 780
 acgaaugggg gugcagaugc aacgguucaa gugaacuagu gacugacuag cccgcugggc 840
 cucccaacgg gccuccucc ccuccuugca ccaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 900
65 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa 942

```

## ES 2 340 499 T3

<210> 6

<211> 942

<212> ARN

5 <213> Secuencia artificial

<220>

10 <223> Descripción de la secuencia artificial: Influenza-Matrix: ARNm con un mayor contenido de G/C y secuencias de estabilización

<220>

15 <223> Las secuencias de estabilización provienen de los UTRs de 5' o 3' del ARNm de la  $\beta$ -globina de *Xenopus laevis*

<220>

20 <223> Codón de inicio: aug (Nucleótidos 56 a 58), Codón de parada: uga (Nucleótidos 812 a 814)

<400> 6

```

25 gcuuguucuu uuugcagaag cucagaauaa acgcucaacu uuggcagauc uaaagaugag 60
 ccugcugacc gagguggaga ccuacgugcu gagcaucauc cccagcggcc ccugaaaggc 120
 cgagaucgcc cagaggcugg aggacguguu cgccggcaag aacaccgacc uggaggugcu 180
 gauggagugg cugaagacca ggcccuaucu gagccccug accaagggca uccugggcuu 240
30 cguguucacc cugaccgugc ccagcagagc cggccugcag cgcgcgccgu ucgugcagaa 300
 cgcccugaac ggcaacggcg accccaacaa cauggacaag gccgugaagc uguacaggaa 360
 gcugaagagg gagaucaccu uccaaggcgc caaggagauc agccugagcu acagcgcggg 420
 cgcccuggcc agcugcaugg gccugaucua caacaggaug ggcgcccuga ccaccgaggu 480
35 ggccuucggc cuggugugcg ccaccugcga gcagaucgcc gacagccagc accgcagcca 540
 caggcagaug gugaccacca ccaacccccu gaucaggcac gagaacagga uggugcuggc 600
 cagcaccacc gccaaaggcca uggagcagau ggccggcagc agcgagcagg ccgccgaggc 660
40 cauggaggug gccagccagg ccaggcagau ggugcaggcc augaggacca ucggcaacca 720
 ccccaagcagc agcgccggcc ugaaagaacga ccugcuggag aaccugcagg ccuaccagaa 780
 gcgcaugggc gugcagaugc agcgcuucau gugaacuagu gacugacuag cccgcugggc 840
45 cucccaacgg gccuccucc ccuccuugca ccaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 900
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa 942

```

<210> 7

50 <211> 1011

<212> ARN

<213> Secuencia artificial

55 <220>

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

<220>

60 <223> Codón de inicio: aug (Nucleótidos 56 a 58), Codón de parada: uga (Nucleótidos 881 a 883)

65

ES 2 340 499 T3

<400> 7

```

5 gcuuguucuu uuugcagaag cucagaauaa acgcucaacu uuggcagauc uaaagauggc 60
 cgucauggcc ccccqcaccc uggugcugcu gcugagcggc gcccuiggccc ugaccagac 120
 cugggccagc cugcugaccg agguggagac cuacgugcug agcaucaucc ccagcggccc 180
 ccugaaggcc gagaucgccc agaggcugga ggacguguuc gccggcaaga acacigaccu 240
10 ggaggugcug auggaguggc ugaagaccag gcccauccug agccccuga ccaagggauc 300
 ccugggcuuc guguucaccc ugaccgugcc cagcgagcgc ggccugcagc gccgcccguu 360
 cgugcagaac gcccugaacg gcaacggcga cccaacaac auggacaagg ccgugaagcu 420
 guacagggaag cugaagaggg agaucaccuu ccacggcgc aaggagauca gccugagcua 480
15 cagcgcggc gcccuggcca gcugcauggg ccugaucuac aacaggauug gcgcccugac 540
 caccgaggug gccuucggcc uggugugcgc caccugcgag cagaucgccg acagccagca 600
 ccgcagccac aggcagaugg ugaccaccac caaccccug aucaggcacg agaacaggau 660
 ggugcuggcc agcaccaccg ccaaggccau ggagcagaug gccgycagca gcgagcaggc 720
20 cgccgaggcc auggaggugg ccagccaggc caggcagaug gugcaggcca ugaggaccau 780
 cggcaccac cccagcagca gcgcccggccu gaagaacgac cugcuggaga accugcaggc 840
 cuaccagaag cgcaugggcg ugcagaugca gcgcuucaag ugaacuagug acugacuagc 900
 ccgcugggcc uccaacggg ccucuccccc cuccuugcac caaaaaaaaa aaaaaaaaaa 960
25 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa a 1011

```

<210> 8

<211> 940

30 <212> ADN

<213> *Homo Sapiens*

<220>

35 <223> MAGE1: Gen del tipo salvaje (para comparación)

<220>

40 <223> Codón de inicio: atg (Nucleótidos 5 a 7), Codón de parada: tga (Nucleótidos 932 a 934)

<400> 8

```

45 catcatgtct cttgagcaga ggagtctgca ctgcaagcct gaggaagccc ttgaggccca 60
 acaagaggcc ctgggcctgg tgtgtgtgca ggctgccacc tcctcctcct ctcctctggt 120

 cctgggcacc ctggaggagg tgcccactgc tgggtcaaca gatcctcccc agagtccctca 180
50 gggagcctcc gcctttccca ctaccatcaa cttcactcga cagaggcaac ccagtgaggg 240
 tccagcagc cgtgaagagg aggggccaag cacctcttgt atcctggagt ccttgttccg 300
 agcagtaate actaagaagg tggctgatit ggttggtttt ctgctcctca aatatcgagc 360
55 caggagcca gtcacaaagg cagaaatgct ggagagtgtc atcaaaaatt acaagcactg 420
 ttttcctgag atcttcggca aagcctctga gtccttgag ctggtctttg gcattgacgt 480
 gaaggaagca gaccccaccg gccactccta tgtccttgtc acctgcctag gtctctocta 540
 tgatggcctg ctgggtgata atcagatcat gcccaagaca ggcttcctga taattgtcct 600
60 ggtcatgatt gcaatggagg gcggccatgc tcctgaggag gaaatctggg aggagctgag 660
 tgtgatggag gtgtatgatg ggaggagca cagtgcctat ggggagccca ggaagctgct 720
 caccacaagat ttggtgcagg aaaagtacct ggagtaccgg caggtgccgg acagtgatcc 780
65 cgcacgctat gagttcctgt ggggtccaag ggccctcgct gaaaccagct atgtgaaagt 840
 ccttgagtat gtgatcaagg tcagtgcaag agttcgctt ttcttcccat ccctgcgtga 900
 agcagctttg agagaggagg aagagggagt ctgagcatga 940

```

ES 2 340 499 T3

<210> 9

<211> 308

<212> PRT

5 <213> *Homo Sapiens*

<220>

<223> Antígeno tumoral MAGE1: secuencia de proteínas

10

<400> 9

15       **Ser Leu Glu Gln Arg Ser Leu His Cys Lys Pro Glu Glu Ala Leu Glu**  
          **1                  5                  10                  15**

20       **Ala Gln Gln Glu Ala Leu Gly Leu Val Cys Val Gln Ala Ala Thr Ser**  
                  **20                  25                  30**

25       **Ser Ser Ser Pro Leu Val Leu Gly Thr Leu Glu Glu Val Pro Thr Ala**  
                  **35                  40                  45**

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

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

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

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

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

50

55

60

65

ES 2 340 499 T3

Glu Ser Val Ile Lys Asn Tyr Lys His Cys Phe Pro Glu Ile Phe Gly  
 130 135 140  
 5 Lys Ala Ser Glu Ser Leu Gln Leu Val Phe Gly Ile Asp Val Lys Glu  
 145 150 155 160  
 10 Ala Asp Pro Thr Gly His Ser Tyr Val Leu Val Thr Cys Leu Gly Leu  
 165 170 175  
 15 Ser Tyr Asp Gly Leu Leu Gly Asp Asn Gln Ile Met Pro Lys Thr Gly  
 180 185 190  
 20 Phe Leu Ile Ile Val Leu Val Met Ile Ala Met Glu Gly Gly His Ala  
 195 200 205  
 25 Pro Glu Glu Glu Ile Trp Glu Glu Leu Ser Val Met Glu Val Tyr Asp  
 210 215 220  
 30 Gly Arg Glu His Ser Ala Tyr Gly Glu Pro Arg Lys Leu Leu Thr Gln  
 225 230 235 240  
 35 Asp Leu Val Gln Glu Lys Tyr Leu Glu Tyr Arg Gln Val Pro Asp Ser  
 245 250 255  
 40 Asp Pro Ala Arg Tyr Glu Phe Leu Trp Gly Pro Arg Ala Leu Ala Glu  
 260 265 270  
 45 Thr Ser Tyr Val Lys Val Leu Glu Tyr Val Ile Lys Val Ser Ala Arg  
 275 280 285  
 50 Val Arg Phe Phe Phe Pro Ser Leu Arg Glu Ala Ala Leu Arg Glu Glu  
 290 295 300  
 55 Glu Glu Gly Val  
 305

<210> 10

<211> 939

<212> ARN

<213> Artificial Sequence

<220>

<223> Descripción de la 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)

65

ES 2 340 499 T3

<400> 10

5        **augagccugg agcagcgcag ccugcacugc aagccggagg aggcgcugga ggcgcagcag 60**  
**gagggcuggg gccuggucug cguccaggcg gcgacgagca gcagcagccc gcugguccug 120**  
**ggcacgcugg aggagguccc gacggcgggc agcacggacc cgcgcagag cccgcagggc 180**  
**gcgagcgcgu ucccgacgac gaucaacuuc acgcgccagc gccagccgag cgagggcagc 240**  
10       **agcagcccgg aggaggaggg cccgagcacg agcugcaucc uggagagccu guuccgcgcg 300**  
**gucaucacga agaaggucgc ggaccugguc ggcuuccugc ugcugaagua ccgcgcgcgc 360**  
**gagccgguca cgaaggcggg gaugcuggag agcgucauca agaacuaca gcacugcuuc 420**  
**ccggagaucu ucggcaaggc gagcgagagc cugcagcugg ucuucggcau cgacgucaag 480**  
15       **gagggcgacc cgacgggcca cagcuacguc cuggucacgu gccugggccu gagcuacgac 540**  
**ggccugcugg gcgacaacca gaucaugccc aagacgggcu uccugaucuau cguccugguc 600**  
**augaucgcga uggaggggcg ccacgcgccg gaggaggaga ucugggagga gcugagcguc 660**  
20       **auggagguc uacgacggcc cgagcacagc gcguacggcg agccgcgcaa gcugcugacg 720**  
**caggaccugg uccaggagaa guaccuggag uaccgccagg ucccggacag cgacccggcg 780**  
**cgcuaagagu uccugugggg cccgcgcgcg cuggcggaga cgagcuacgu caagguccug 840**  
**gaguacguca ucaaggucag cgcgcgcguc cgcuuuucu ucccagccu gcgcgaggcg 900**  
25       **gcgcugcgcg aggaggaggg gggcgcugca gcgugauga** 939

<210> 11

<211> 939

30 <212> ARN

<213> Secuencia artificial

<220>

35 <223> Descripción de la secuencia artificial: MAGE1: ARNm con utilización alternativa de codón.

<220>

40 <223> Codón de inicio: aug (Nucleótidos 1 a 3), codón de parada: uga (Nucleótidos 937 a 939)

<400> 11

45       **augagccugg agcagcgcag ccugcacugc aagcccagg aggcccuugga ggcgcagcag 60**  
**gagggccugg gccuggugug cgugcaggcc gccaccagca gcagcagccc ccuggugcug 120**  
**ggcacccugg aggaggugcc caccgccggc agcaccgacc cccccagag ccccagggc 180**  
**gccagcgcgu ucccaccac caucaacuuc acccgcagc gccagcccag cgagggcagc 240**  
50       **agcagcccgg aggaggaggg cccagcacc agcugcaucc uggagagccu guuccgcgcc 300**  
**gugaucacca agaagguggc cgaccuggug ggcuuccugc ugcugaagua ccgcgcccgc 360**  
**gagcccguga ccaaggccga gaugcuggag agcgugaucaga agaacuaca gcacugcuuc 420**  
**cccgagaucu ucggcaaggc cagcgagagc cugcagcugg uguucggcau cgacgugaag 480**  
55       **gagggcgacc ccaccggcca cagcuacgug cuggugaccu gccugggccu gagcuacgac 540**  
**ggccugcugg gcgacaacca gaucaugccc aagaccggcu uccugaucuau cgugcuggug 600**  
**augaucgcca uggaggggcg ccacgcccc gaggaggaga ucugggaggga gcugagcgug 660**  
60       **auggaggugu acgacggccc cgagcacagc gccuacggcg agccccgcaa gcugcugacc 720**  
  
**caggaccugg ugcaggagaa guaccuggag uaccgccagg ugcccagacag cgaccccgcc 780**  
**cgcuaagagu uccugugggg cccccgcgc cuggccgaga ccagcuacgu gaaggugcug 840**  
65       **gaguacguga ucaaggugag cgcgcgcgug cgcuuuucu ucccagccu gcgcgaggcc 900**  
**ggccugcgcg aggaggaggg gggcgcugca gccugauga** 939

## ES 2 340 499 T3

<210> 12

<211> 7

<212> ARN

5 <213> Secuencia artificial

<220>

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

<400> 12

15 gaacaag

7

<210> 13

20 <211> 13

<212> ARN

<213> Secuencia artificial

25 <220>

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

<400> 13

30 gccgccacca ugg

13

35

40

45

50

55

60

65