



(22) Date de dépôt/Filing Date: 2002/06/05
(41) Mise à la disp. pub./Open to Public Insp.: 2002/12/12
(62) Demande originale/Original Application: 2 457 959
(30) Priorité/Priority: 2001/06/05 (DE101 27 283.9)

(51) Cl.Int./Int.Cl. *C12N 15/33* (2006.01),
A61K 39/12 (2006.01), *A61P 31/12* (2006.01),
A61P 37/04 (2006.01), *C12N 15/38* (2006.01),
C12N 15/40 (2006.01), *C12N 15/44* (2006.01),
C12N 15/49 (2006.01), *C12N 15/51* (2006.01),
C12N 15/67 (2006.01)

(71) Demandeur/Applicant:
CUREVAC GMBH, DE

(72) Inventeurs/Inventors:
HOERR, INGMAR, DE;
PASCOLO, STEVE, DE;
VON DER MULBE, FLORIAN, DD

(74) Agent: KIRBY EADES GALE BAKER

(54) Titre : COMPOSITION PHARMACEUTIQUE CONTENANT UN ARNM STABILISE ET OPTIMISE EN VUE DE LA
TRADUCTION DE SES REGIONS CODANTES
(54) Title: PHARMACEUTICAL COMPOSITION CONTAINING A STABILISED MRNA OPTIMISED FOR TRANSLATION
IN ITS CODING REGIONS

(57) **Abrégé/Abstract:**

The present invention relates to a pharmaceutical composition containing an mRNA that is stabilised by sequence modifications in the translated region and is optimised for the translation. The pharmaceutical composition according to the invention is particularly suitable as an inoculating agent as well as a therapeutic agent for tissue regeneration. In addition a process is described for determining sequence modifications that serve for the stabilisation and translation optimisation of mRNA.



Abstract

5

The present invention relates to a pharmaceutical composition containing an mRNA that is stabilised by sequence modifications in the translated region and is optimised for the translation. The pharmaceutical composition according to the invention is particularly suitable as an inoculating agent as well as a therapeutic agent for tissue regeneration. In addition a process is described for determining sequence modifications that serve for the stabilisation and translation optimisation of mRNA.

10

15

DEMANDES OU BREVETS VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVETS
COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 2

NOTE: Pour les tomes additionels, veuillez contacter le Bureau Canadien des Brevets.

JUMBO APPLICATIONS / PATENTS

THIS SECTION OF THE APPLICATION / PATENT CONTAINS MORE
THAN ONE VOLUME.

THIS IS VOLUME 1 OF 2

NOTE: For additional volumes please contact the Canadian Patent Office.

**Pharmaceutical Composition containing a stabilised mRNA
optimised for translation in its Coding Regions**

5 This application is a divisional of Canadian Patent
Application 2,457,959, which is a national phase application
derived from International Patent Application
PCT/EP2002/006180 filed June 5, 2002, which entered the
10 national phase on December 3, 2003.

The present invention relates to a pharmaceutical
composition containing an mRNA that is stabilised by
15 sequence modifications in the translated region and is
optimised for the translation. The pharmaceutical
composition according to the invention is suitable in
particular as an inoculating agents and also as a
therapeutic agent for tissue regeneration. Furthermore a
20 process for determining sequence modifications that serve
for the stabilisation and optimisation of the translation
of mRNA is disclosed.

Gene therapy and genetic vaccination are molecular medicine
25 processes whose use in the treatment and prevention of
diseases will have considerable effects on medical
practice. Both processes are based on the incorporation of
nucleic acids into a patient's cells or tissue as well as
on the subsequent processing of the information coded by
30 the incorporated nucleic acids, i.e. the expression of the
desired polypeptides.

The conventional procedure involved in previous processes
of gene therapy and genetic vaccination is the use of DNA
35 in order to incorporate the required genetic information
into the cell. In this connection various processes for
the incorporation of DNA into cells have been developed,

such as for example calcium phosphate transfection,
polybrene transfection, protoplast fusion, electroporation,
microinjection and lipofection, in which connection
lipofection in particular has proved to be a suitable
5 process.

A further process that has been suggested in particular in
the case of genetic vaccination processes is the use of DNA
viruses as DNA vehicle. Such viruses have the advantage
10 that, on account of their infectious properties, a very
high transfection rate can be achieved. The viruses used
are genetically altered so that no functional infectious
particles are formed in the transfected cell. Despite this
precautionary measure however a certain risk of the
15 uncontrolled propagation of the introduced genetically
therapeutically active as well as viral genes cannot be
excluded on account of possible recombination events.

Normally the DNA incorporated into the cell is integrated
20 to a certain extent into the genome of the transfected
cell. On the one hand this phenomenon can exert a
desirable effect, since in this way a long-lasting action
of the introduced DNA can be achieved. On the other hand
the integration into the genome brings with it a
25 significant risk for the gene therapy. Thus, this may for
example result in an insertion of the incorporated DNA into
an intact gene, which produces a mutation that interferes
with or indeed completely switches off the function of the
endogenous gene. As a result of such integration events
30 enzyme systems that are important for cellular life may on
the one hand be switched off, while on the other hand there
is also the danger of a transformation of the thus altered
cell into a degenerate state if, due to the integration of

the foreign DNA, a gene that is decisive for regulating cell growth is altered. Accordingly, when using DNA viruses as gene therapeutic agents and vaccines a carcinogenic risk cannot be excluded. In this connection
5 it should also be borne in mind that, in order to achieve the effective expression of the genes incorporated into the cell, the corresponding DNA vehicles contain a strong promoter, for example the viral CMV promoter. The integration of such promoters into the genome of the
10 treated cell may lead to undesirable changes in the regulation of the gene expression in the cell.

A further disadvantage of the use of DNA as gene therapeutic agents and vaccines is the induction of
15 pathogenic anti-DNA antibodies in the patient, resulting in a possible fatal immune response.

In contrast to DNA the use of RNA as a gene therapeutic agent or vaccine is regarded as being significantly safer.
20 In particular RNA does not bring with it the danger of becoming stably integrated into the genome of the transfected cell. In addition, no viral sequences such as promoters are necessary for the effective transcription. Over and above this, RNA is degraded significantly more
25 readily *in vivo*. Indeed, on account of the relatively short half-life of RNA in the blood circulation compared to DNA, up to now no anti-RNA antibodies have been detected. For this reason RNA may be regarded as the molecule of choice for molecular medicine therapeutic processes.

30

However, some basic problems still have to be solved before medical processes based on RNA expression systems can be widely employed. One of the problems in the use of RNA is

the reliable, cell-specific and tissue-specific efficient transfer of the nucleic acid. Since RNA is normally found to be very unstable in solution, up to now RNA could not be used or used only very inefficiently as a therapeutic agent or inoculating agent in the conventional processes that are used for DNA.

Enzymes that break down RNA, so-called RNAases (ribonucleases), are responsible for the instability. Even 10 minute contamination by ribonucleases is sufficient to break down RNA completely in solution. The natural decomposition of mRNA in the cytoplasm of cells is extremely finely regulated. Several mechanisms are known in this respect. Thus, for a functional mRNA the terminal 15 structure is of decisive importance. The so-called "cap structure" (a modified guanosine nucleotide) is located at the 5' end and a sequence of up to 200 adenosine nucleotides (the so-called poly-A tail) is located at the 3' end. The RNA is recognised as mRNA by these structures and 20 the decomposition is regulated. In addition there are further processes that stabilise or destabilise RNA. Many of these processes are still unknown, though often an interaction between the RNA and proteins appears to be decisive in this regard. For example, an mRNA surveillance 25 system has recently been described (Hellerin and Parker, Annu. Rev. Genet. 1999, 33: 229 to 260), in which incomplete or nonsense mRNA is recognised by specific feedback protein interactions in the cytosol and is made accessible to decomposition, a main part of these processes 30 being completed by exonucleases.

Certain measures have been proposed in the prior art in order to improve the stability of RNA and thereby permit its use as a gene therapeutic agent or RNA vaccine.

5 In EP-A-1083232 a process for the incorporation of RNA, in particular mRNA, into cells and organisms has been proposed in order to solve the aforementioned problem of the instability of RNA ex vivo, in which the RNA is present in the form of a complex with a cationic peptide or protein.

10

WO 99/14346 describes further processes for stabilising mRNA. In particular modifications of the mRNA are proposed that stabilise the mRNA species against decomposition by RNases. Such modifications involve on the one hand
15 stabilisation by sequence modifications, in particular reduction of the C content and/or U content by base elimination or base substitution. On the other hand chemical modifications are proposed, in particular the use of nucleotide analogues, as well as 5' and 3' blocking
20 groups, an increased length of the poly-A tail as well as the complexing of the mRNA with stabilising agents, and combinations of the aforementioned measures.

In US patents US 5,580,859 and US 6,214,804 mRNA vaccines
25 and mRNA therapeutic agents are disclosed *inter alia* within the scope of "transient gene therapy" (TGT). Various measures are described for raising the translation efficiency and the mRNA stability, that relate in particular to the non-translated sequence regions.

30

Bieler and Wagner (in: Schleef (Ed.), Plasmids for Therapy and Vaccination, Chapter 9, pp. 147 to 168, Wiley-VCH, Weinheim, 2001) report on the use of synthetic genes in

combination with gene therapy methods employing DNA vaccines and lentiviral vectors. The construction of a synthetic *gag*-gene derived from HIV-1 is described, in which the codons have been modified with respect to the wild type sequence (alternative codon usage) in such a way that it corresponded to the use of codons that can be found in highly expressed mammalian genes. In this way in particular the A/T content compared to the wild type sequence was reduced. The authors find in particular an increased rate of expression of the synthetic *gag* gene in transfected cells. Furthermore, in mice an increased antibody formation against the *gag* protein was observed in the case of mice immunised with the synthetic DNA construct and also an increased cytokine release in vitro in the case of transfected spleen cells of mice. Finally, an induction of a cytotoxic immune response in mice immunised with the *gag* expression plasmid was also found. The authors of this article attribute the improved properties of their DNA vaccine basically to a change of the nucleocytoplasmic transport of the mRNA expressed by the DNA vaccine, brought about by the optimised codon usage. On the other hand the authors maintain that the effect of the altered codon usage on the translation efficiency is only slight.

The object of the present invention is accordingly to provide a new system for gene therapy and genetic vaccination that overcomes the disadvantages associated with the properties of DNA therapeutic agents and DNA vaccines and that increases the effectiveness of therapeutic agents based on RNA species.

This object is achieved by the embodiments of the present invention characterised herein.

6a

Certain exemplary embodiments provide a pharmaceutical composition comprising at least one modified mRNA that codes for at least one tumour antigenic peptide or polypeptide in combination with a pharmaceutically compatible carrier and/or vehicle, characterised in that the G/C content of the region of the modified mRNA coding for the peptide or polypeptide is larger than the G/C content of the coding region of the wild type mRNA coding for the peptide or polypeptide and the coded amino acid sequence is unchanged compared to the wild type mRNA.

Certain exemplary embodiments further provide a modified mRNA coding for at least one tumour antigenic peptide or polypeptide, characterised in that the G/C content of the region of the modified mRNA coding for the peptide or polypeptide is increased compared to the G/C content of the coding region of the wild type mRNA coding for the peptide or polypeptide, and the encoded amino acid sequence is unchanged as compared to the wild type mRNA.

Certain exemplary embodiments further provide a pharmaceutical composition comprising at least modified mRNA that codes for at least one antigenic viral peptide or polypeptide in combination with a pharmaceutically compatible carrier and/or vehicle, characterised in that the G/C content of the region of the modified mRNA coding for the peptide or polypeptide is larger than the G/C content of the coding region of the wild type mRNA coding for the peptide or polypeptide and the coded amino acid sequence is unchanged compared to the wild type mRNA.

6b

Certain exemplary embodiments further provide a modified mRNA coding for at least one antigenic viral peptide or polypeptide, characterised in that the G/C content of the region of the modified mRNA coding for the peptide or

5 polypeptide is increased compared to the G/C content of the coding region of the wild type mRNA coding for the peptide or polypeptide, and the encoded amino acid sequence is unchanged as compared to the wild type mRNA.

In particular a modified mRNA as well as a pharmaceutical composition containing at least an mRNA modified in this way in combination with a pharmaceutically compatible carrier and/or vehicle are provided, in which the modified mRNA codes for at least one biologically active or antigenic peptide or polypeptide, wherein the sequence of the mRNA, in particular in the region coding for the at least one peptide or polypeptide, comprises the following modifications compared to the wild type mRNA, which may be present either individually or in combination.

On the one hand the G/C content of the region of the modified mRNA coding for the peptide or polypeptide is larger than the G/C content of the coding region of the wild type mRNA coding for the peptide or polypeptide, the coded amino acid sequence being unchanged compared to the wild type.

This modification is based on the fact that, for an efficient translation of an mRNA, the sequence of the region of the mRNA to be translated is essential. In this connection the composition and the sequence of the various nucleotides play an important rôle. In particular sequences with an increased G(guanosine)/C(cytosine) content are more stable than sequences with an increased A(adenosine)/U(uracil) content. Accordingly, according to the invention the codons are varied compared to the wild type mRNA, while maintaining the translated amino acid sequence, so that they contain increased amounts of G/C nucleotides. Since several codons code for one and the same amino acid (degeneration of the genetic code), the

codons most favourable for the stability can be determined (alternative codon usage).

Depending on the amino acid to be coded by the modified
5 mRNA, various possibilities for modifying the mRNA sequence compared to the wild type sequence are feasible. In the case of amino acids that are coded by codons that contain exclusively G or C nucleotides, no modification of the codon is necessary. Thus, the codons for Pro (CCC or CCG),
10 Arg (CGC or CGG), Ala (GCC or GCG) and Gly (GGC or GGG) do not require any alteration since no A or U is present.

In the following cases the codons that contain A and/or U nucleotides are altered by substituting other codons that
15 code for the same amino acids, but do not contain A and/or U. Examples include:

the codons for Pro may be changed from CCU or CCA to CCC or CCG;

the codons for Arg may be changed from CGU or CGA or AGA or
20 AGG to CGC or CGG;

the codons for Ala may be changed from GCU or GCA to GCC or GCG;

the codons for Gly may be changed from GGU or GGA to GGC or GGG.

25

In other cases, although A and/or U nucleotides may not be eliminated from the codons, it is however possible to reduce the A and U content by using codons that contain fewer A and/or U nucleotides. For example:

30 the codons for Phe may be changed from UUU to UUC;

the codons for Leu may be changed from UUA, CUU or CUA to CUC or CUG;

- the codons for Ser may be changed from UCU or UCA or AGU to UCC, UCG or AGC;
- the codon for Tyr may be changed from UAU to UAC;
- the stop codon UAA may be changed to UAG or UGA;
- 5 the codon for Cys may be changed from UGU to UGC;
- the codon for His may be changed from CAU to CAC;
- the codon for Gln may be changed from CAA to CAG;
- the codons for Ile may be changed from AUU or AUA to AUC;
- the codons for Thr may be changed from ACU or ACA to ACC or
- 10 ACG;
- the codon for Asn may be changed from AAU to AAC;
- the codon for Lys may be changed from AAA to AAG;
- the codons for Val may be changed from GUU or GUA to GUC or GUG;
- 15 the codon for Asp may be changed from GAU to GAC;
- the codon for Glu may be changed from GAA to GAG.

In the case of the codons for Met (AUG) and Trp (UGG) there is however no possibility of modifying the sequence.

20

- The substitutions listed above may obviously be used individually but also in all possible combinations in order to increase the G/C content ratio of the modified mRNA compared to the original sequence. Thus for example all
- 25 codons for Thr occurring in the original (wild type) sequence can be altered to ACC (or ACG). Preferably however combinations of the substitution possibilities given above are employed, for example:
- substitution of all codons coding in the original sequence
- 30 for Thr to ACC (or ACG) and substitution of all codons coding originally for Ser to UCC (or UCG or AGC);
- substitution of all codons coding in the original sequence for Ile to AUC and substitution of all codons coding

originally for Lys to AAG and substitution of all codons coding originally for Tyr to UAC; substitution of all codons coding in the original sequence for Val to GUC (or GUG) and substitution of all codons coding originally for Glu to GAG and substitution of all codons coding originally for Ala to GCC (or GCG) and substitution of all codons coding originally for Arg to CGC (or CCG); substitution of all codons coding in the original sequence for Val to GUC (or GUG) and substitution of all codons coding originally for Glu to GAG and substitution of all codons coding originally for Ala to GCC (or GCG) and substitution of all codons coding originally for Gly to GGC (or GGG) and substitution of all codons coding originally for Asn to AAC; substitution of all codons coding in the original sequence for Val to GUC (or GUG) and substitution of all codons coding originally for Phe to UUC and substitution of all codons coding originally for Cys to UGC and substitution of all codons coding originally for Leu to CUG (or CUC) and substitution of all codons coding originally for Gln to CAG and substitution of all codons coding originally for Pro to CCC (or CCG); etc.

Preferably the G/C content of the region of the modified mRNA coding for the peptide or polypeptide is increased by at least 7% points, more preferably by at least 15% points, and particularly preferably by at least 20% points compared to the G/C content of the coded region of the wild type mRNA coding for the polypeptide.

In this connection it is particularly preferred to increase by the maximum possible amount the G/C content of the

modified mRNA, in particular in the region coding for the at least one peptide or polypeptide, compared to the wild type sequence.

5 The further modification according to the invention of the mRNA contained in the pharmaceutical composition characterised in the present invention is based on the knowledge that the translation efficiency is also determined by a different frequency in the occurrence of
10 tRNAs in cells. If therefore so-called "rare" codons are frequently present in an RNA sequence, then the corresponding mRNA is translated significantly worse than in the case when codons coding for relatively "frequent" tRNAs are present.

15

Thus, according to the invention in the modified mRNA (which is contained in the pharmaceutical composition) the region coding for the peptide or polypeptide is changed compared to the corresponding region of the wild type mRNA
20 in such a way that at least one codon of the wild type sequence that codes for a relatively rare tRNA in the cell is exchanged for a codon that codes for a relatively frequent tRNA in the cell that carries the same amino acid as the relatively rare tRNA.

25

Through this modification the RNA sequences are modified so that codons are inserted that are available for the frequently occurring tRNAs.

30 Which tRNAs occur relatively frequently in the cell and which on the other hand occur relatively seldom is known to the person skilled in the art; see for example Akashi, Curr. Opin. Genet. Dev. 2001, 11(6): 660-666.

By means of this modification, according to the invention all codons of the wild type sequence that code for a relatively rare tRNA in the cell may in each case be
5 exchanged for a codon that codes for a relatively frequent tRNA in the cell that in each case carries the same amino acid as the relatively rare tRNA.

According to the invention it is particularly preferred to
10 couple the sequential G/C fraction that according to the invention is increased, in particular maximally increased, in the modified mRNA, with the "frequent" codons, without changing the amino acid sequence of the peptide or polypeptide (one or more) coded by the coding region of the
15 mRNA. This preferred embodiment provides a particularly efficiently translated and stabilised mRNA, for example for the pharmaceutical composition according to the invention.

In the sequences of eukaryotic mRNAs there are
20 destabilising sequence elements (DSE) to which signal proteins bind and regulate the enzymatic decomposition of the mRNA *in vivo*. Accordingly, for the further stabilisation of the modified mRNA contained in the pharmaceutical composition according to the invention one
25 or more changes compared to the corresponding region of the wild type mRNA are carried out if necessary in the region coding for the at least one peptide or polypeptide, so that no destabilising sequence elements are contained.

Obviously it is also preferred according to the invention
30 to eliminate from the mRNA DSE possibly present in the non-translated regions (3' and/or 5' UTR).

Such destabilising sequences are for example AU-rich sequence ("AURES") that occur in 3'-UTR sections of numerous unstable mRNAs (Caput et al., Proc. Natl. Acad. Sci. USA 1986, 83: 1670 to 1674). The RNA molecules
5 contained in the pharmaceutical composition according to the invention are therefore preferably altered compared to the wild type mRNA so that they do not have any such destabilising sequences. This also applies to those sequence motifs that are recognised by possible
10 endonucleases, for example the sequence GAACAAG, that is contained in the 3'UTR segment of the gene coding for the transferin receptor (Binder et al., EMBO J. 1994, 13: 1969 to 1980). These sequence motifs too are preferably eliminated in the modified mRNA of the pharmaceutical
15 composition according to the invention.

Various methods are known to the person skilled in the art that are suitable for the substitution of codons in the modified mRNA according to the invention. In the case of
20 relatively short coding regions (that code for biologically active or antigenic peptides), for example the whole mRNA may be chemically synthesised using standard techniques.

Preferably however base substitutions are introduced using
25 a DNA matrix for the production of modified mRNA with the aid of techniques employed in current targeted mutagenesis; see Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, 3rd Edition, Cold Spring Harbor, NY, 2001.

30

In this method, a corresponding DNA molecule is therefore transcribed in vitro for the production of the mRNA. This DNA matrix has a suitable promoter, for example a T7 or SP6

promoter, for the *in vitro* transcription, followed by the desired nucleotide sequence for the mRNA to be produced and a termination signal for the *in vitro* transcription. According to the invention the DNA molecule that forms the matrix of the RNA construct to be produced is prepared by fermentative replication and subsequent isolation as part of a plasmid replicable in bacteria. As plasmids suitable for the present invention there may for example be mentioned the plasmids pT7Ts (GeneBank Accession No. U26404; Lai et al., Development 1995, 121: 2349 to 2360), the pGEM[®] series, for example pGEM[®]-1 (GeneBank Accession No. X65300; from Promega) and pSP64 (GeneBank-Accession No. X65327); see also Mezei and Storts, Purification of PCR Products, in: Griffin and Griffin (Eds.), PCR Technology: Current Innovation, CRC Press, Boca Raton, FL, 2001.

Thus, by using short synthetic DNA oligonucleotides that comprise short single-strand transitions at the corresponding cleavage sites, or by means of genes produced by chemical synthesis, the desired nucleotide sequence can be cloned into a suitable plasmid by molecular biology methods known to the person skilled in the art (see Maniatis et al., above). The DNA molecule is then excised from the plasmid, in which it may be present as a single or multiple copy, by digestion with restriction endonucleases.

The modified mRNA that is contained in the pharmaceutical composition according to the invention may furthermore have a 5' cap structure (a modified guanosine nucleotide). As examples of cap structures there may be mentioned m⁷G(5')ppp (5' (A,G(5')ppp(5'))A and G(5')ppp(5')G.

According to a further preferred embodiment of the present invention the modified mRNA contains a poly-A tail of at least 50 nucleotides, preferably at least 70 nucleotides, more preferably at least 100 nucleotides and particularly
5 preferably at least 200 nucleotides.

For an efficient translation of the mRNA an effective binding of the ribosomes to the ribosome binding site (Kozak sequence: GCCGCCACCAUGG, the AUG forms the start
10 codon) is furthermore necessary. In this regard it has been established that an increased A/U content around this site permits a more efficient ribosome binding to the mRNA.

In addition it is possible to introduce one or more
15 so-called IRES ("internal ribosomal entry site") into the modified mRNA. An IRES may thus act as the sole ribosome binding site, but may however also serve for the provision of an mRNA that codes several peptides or polypeptides that are to be translated independently of one another by the
20 ribosomes ("multicistronic mRNA"). Examples of IRES sequences that can be used according to the invention are those from picornaviruses (e.g. FMDV), pest viruses (CFFV), polio viruses (PV), encephalomyocarditis viruses (ECMV), foot-and-mouth disease viruses (FMDV), hepatitis C viruses
25 (HCV), classical swine fever viruses (CSFV), murine leukoma virus (MLV), simian immune deficiency viruses (SIV) or cricket paralysis viruses (CrPV).

According to a further preferred embodiment of the present
30 invention the modified mRNA contains in the 5' non-translated and/or 3' non-translated regions stabilisation

sequences that are capable of increasing the half-life of the mRNA in the cytosol.

These stabilisation sequences may exhibit a 100% sequence
5 homology with naturally occurring sequences that are present in viruses, bacteria and eukaryotic cells, but may however also be partly or completely synthetic. As an example of stabilising sequences that may be used in the present invention, the non-translated sequences (UTR) of
10 the β -globin gene, for example of *Homo sapiens* or *Xenopus laevis*, may be mentioned. Another example of a stabilisation sequence has the general formula
(C/U)CCAN_xCCC(U/A)Py_xUC(C/U)CC, which is contained in the 3'UTR of the very stable mRNA that codes for α -globin,
15 α -(I)-collagen, 15-lipoxygenase or for tyrosine hydroxylase (c.f. Holcik et al., Proc. Natl. Acad. Sci. USA 1997, 94: 2410 to 2414). Obviously such stabilisation sequences may be used individually or in combination with one another as well as in combination with other stabilisation sequences
20 known to the person skilled in the art.

For the further stabilisation of the modified mRNA it is also preferred if this contains at least one analogue of naturally occurring nucleotides. That is based on the fact
25 that the RNA-decomposing enzymes present in the cells recognise as substrate preferably naturally occurring nucleotides. By insertion of nucleotide analogues the RNA decomposition can therefore be made more difficult, whereby the effect on the translation efficiency when inserting
30 these analogues, in particular into the coding region of the mRNA, may have a positive or negative effect on the translation efficiency.

In a by no means exhaustive list the following may be mentioned as examples of nucleotide analogues that can be used according to the invention: phosphorus amidates, phosphorus thioates, peptide nucleotides, methylphosphonates, 7-deazaguanosine, 5-methylcytosine and inosine. The preparation of such analogues is known to the person skilled in the art, for example from US patents 4,373,071, US 4,401,796, US 4,415,732, US 4,458,066, US 4,500,707, US 4,668,777, US 4,973,679, US 5,047,524, US 5,132,418, US 5,153,319, US 5,262,530 and 5,700,642.

According to the invention such analogues may be present in non-translated and translated regions of the modified mRNA.

Furthermore the effective transfer of the modified mRNA into the cells to be treated or into the organism to be treated may be improved if the modified mRNA is associated with a cationic peptide or protein or is bound thereto. In particular in this connection the use of protamine as polycationic, nucleic acid-binding protein is particularly effective. In addition it is also possible to use other cationic peptides or proteins such as poly-L-lysine or histones. This procedure for the stabilisation of the modified mRNA is described in EP-A-1083232, whose relevant disclosure is included in full in the present invention.

25

In the gene therapy use of the pharmaceutical composition according to the invention (or in the use of the mRNA for gene therapy or in the use of the mRNA for the preparation of a pharmaceutical composition for gene therapy) the modified mRNA codes for at least one biologically active peptide or polypeptide that is not formed or is only insufficiently or defectively formed in the patient to be treated. Accordingly, examples of polypeptides coded by

30

the mRNA according to the invention include dystrophin, the chloride channel, which is defectively altered in cystic fibrosis, enzymes that are lacking or defective in metabolic disorders such as phenylketonuria, galactosaemia, homocystinuria, adenosine deaminase deficiency, etc., enzymes that are involved in the synthesis of neurotransmitters such as dopamine, norepinephrine and GABA, in particular tyrosine hydroxylase and DOPA decarboxylase, α -1-antitrypsin, etc. Besides this the pharmaceutical composition may be used for the release of cell surface receptors and molecules that bind to such receptors if the modified mRNA contained in the pharmaceutical composition codes for such a biologically active protein or peptide. Examples of such proteins that act in an extracellular manner or that bind to cell surface receptors include for example tissue plasminogen activator (TPA), growth hormones, insulin, interferons, granulocyte-macrophage colony stimulating factor (GM-CFS), erythropoietin (EPO), etc. By choosing suitable growth factors the pharmaceutical composition of the present invention may for example be used for tissue regeneration. In this way diseases that are characterised by a tissue degeneration, for example neurodegenerative diseases such as Alzheimer's disease, Parkinson's disease, etc. and other degenerative conditions, for example arthrosis, can be treated. In these cases the modified mRNA, in particular that contained in the pharmaceutical composition of the present invention, preferably codes for growth factors from the TGF- β family, in which in particular EGF, FGF, PDGF, BMP, GDNF, BDNF, GDF and neurotrophic factors such as NGF, neutrophines, etc., may be mentioned.

A further area of application of the present invention is vaccination, i.e. the use of the modified mRNA for inoculation or the use of the pharmaceutical composition as an inoculating agent or the use of the modified mRNA for the preparation of the pharmaceutical composition for inoculation purposes. Vaccination is based on introducing an antigen, in the present case the genetic information for the antigen in the form of the modified mRNA coding for the antigen, into the organism, in particular into the cell.

10 The modified mRNA contained in the pharmaceutical composition is translated into the antigen, i.e. the polypeptide or antigenic peptide coded by the modified mRNA is expressed, whereby an immune response directed against this polypeptide or antigenic peptide is stimulated. In

15 vaccination against a pathogenic organism, i.e. a virus, a bacterium or a protozoological organism, a surface antigen of such an organism is therefore used for the vaccination with the aid of the pharmaceutical composition according to the invention containing the modified mRNA coding for the

20 said surface antigen. In case of use as a genetic vaccine for treating cancer, the immune response is achieved by incorporating the genetic information for tumour antigens, in particular proteins that are expressed exclusively on cancer cells, in which a pharmaceutical composition

25 according to the invention is administered that contains an mRNA coding for such a cancer antigen. In this way the cancer antigen(s) is/are expressed in the organism, whereby an immune response is produced that is directed effectively against the cancer cells.

30

In its use as a vaccine the pharmaceutical composition according to the invention is suitable in particular for the treatment of cancers (in which the modified mRNA codes

for a tumour-specific surface antigen (TSSA)), for example for treating malignant melanoma, colon carcinoma, lymphomas, sarcomas, small-cell lung carcinomas, blastomas, etc. Specific examples of tumour antigens include, *inter alia*, 707-AP, AFP, ART-4, BAGE, β -catenin/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 (or hTRT), iCE, KIAA0205, LAGE, LDLR/FUT, MAGE, MART-1/melan-A, MC1R, myosin/m, MUC1, MUM-1, -2, -3, NA88-A, NY-ESO-1, p190 minor bcr-abl, Pml/RAR α , PRAME, PSA, PSM, RAGE, RU1 or RU2, SAGE, SART-1 or SART-3, TEL/AML1, TPI/m, TRP-1, TRP-2, TRP-2/INT2 and WT1. In addition to this the pharmaceutical composition according to the invention is used to treat infectious diseases (for example viral infectious diseases such as AIDS (HIV), hepatitis A, B or C, herpes, herpes zoster (chicken pox), German measles (rubella virus), yellow fever, dengue fever etc. (flavi viruses), 'flu (influenza viruses), haemorrhagic infectious diseases (Marburg or Ebola viruses), bacterial infectious diseases such as Legionnaires' disease (legionella), gastric ulcer (helicobacter), cholera (vibrio), *E.coli* infections, staphylococcal infections, salmonella infections or streptococcal infections (tetanus), or protozoological infectious diseases (malaria, sleeping sickness, leishmaniasis, toxoplasmosis, i.e. infections caused by plasmodium, trypanosomes, leishmania and toxoplasma). Preferably also in the case of infectious diseases the corresponding surface antigens with the strongest antigenic potential are coded by the modified mRNA. With the aforementioned genes of pathogenic germs or organisms, in particular in the case of viral genes, this is typically a secreted form of a surface antigen.

Moreover, according to the invention mRNAs preferably coding for polypeptides are employed, these polypeptides being polypeptides, for example of the aforementioned antigens, in particular surface antigens of pathogenic
5 germs or organisms, or tumour cells, preferably secreted protein forms.

In addition to this, the modified mRNA according to the invention may contain, besides the antigenic or the
10 genetically therapeutically active peptide or polypeptide, also at least one further functional section that codes for example for a cytokine promoting the immune response (monokine, lymphokine, interleukin or chemokine, such as IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9,
15 IL-10, IL-12, INF- α , INF- γ , GM-CSF, LT- α or growth factors such as hGH).

Furthermore, in order to increase the immunogenicity the pharmaceutical composition according to the invention may
20 contain one or more adjuvants. The term "adjuvant" is understood in this context to denote any chemical or biological compound that promotes a specific immune response. Various mechanisms may be involved in this connection, depending on the various types of adjuvants.
25 For example, compounds that promote through dendritic cells (DC) an endocytosis of the modified mRNA contained in the pharmaceutical composition form a first class of usable adjuvants. Other compounds that permit the maturation of the DC, for example lipopolysaccharides, TNF- α or CD40
30 ligand, comprise a further class of suitable adjuvants. In general each agent influencing in the manner of a "danger signal" the immune system (LPS, GP96, oligonucleotides with the CpG motif) or cytokines such as GM-CSF, may be used as

adjuvant, which enables an immune response against an antigen that is coded by the modified mRNA to be increased and/or specifically influenced. The aforementioned cytokines are particularly preferred in this connection.

5 Further known adjuvants include aluminium hydroxide, Freund's adjuvant as well as the aforementioned stabilising cationic peptides or polypeptides such as protamine. In addition to this lipopeptides such as Pam3Cys are also particularly suitable for use as adjuvants in the

10 pharmaceutical composition of the present invention; see Deres et al, Nature 1989, 342: 561-564.

The pharmaceutical composition according to the invention contains, apart from the modified mRNA, also a

15 pharmaceutically compatible carrier and/or a pharmaceutically compatible vehicle. Appropriate methods for achieving a suitable formulation and preparation of the pharmaceutical composition according to the invention are described in "Remington's Pharmaceutical Sciences" (Mack

20 Pub. Co., Easton, PA, 1980), all of which form part of the disclosure of the present invention. For parenteral administration suitable carriers include for example sterile water, sterile saline solutions, polyalkylene glycols, hydrogenated naphthalene and in particular

25 biocompatible lactide polymers, lactide/glycolide copolymers or polyoxyethylene/polyoxypropylene copolymers. Compositions according to the invention may contain fillers or substances such as lactose, mannitol, substances for the covalent coupling of polymers such as for example

30 polyethylene glycol to inhibitors according to the invention, complexing with metal ions or incorporation of materials in or on special preparations of polymer compound, such as for example polylactate, polyglycolic

acid, hydrogel or on liposomes, microemulsions, microcells, unilamellar or multilamellar vesicles, erythrocyte fragments or spheroplasts. The respective modifications of the compositions are chosen depending on the physical
5 behaviour, for example with regard to solubility, stability, bioavailability or degradability. Controlled or constant release of the active component according to the invention in the composition includes formulations based on lipophilic depôt substances (for example fatty acids, waxes
10 or oils). Coatings of substances or compositions according to the invention containing such substances, namely coatings with polymers (for example poloxamers or poloxamines), are also disclosed within the scope of the present invention. Moreover substances or compositions
15 according to the invention may contain protective coatings, for example protease inhibitors or permeability enhancers. Preferred carriers are typically aqueous carrier materials, in which water for injection (WFI) or water buffered with phosphate, citrate or acetate, etc., is used, and the pH is
20 typically adjusted to 5.0 to 8.0, preferably 6.0 to 7.0. The carrier or the vehicle will in addition preferably contain salt constituents, for example sodium chloride, potassium chloride or other components that for example make the solution isotonic. In addition the carrier or the
25 vehicle may contain, besides the aforementioned constituents, additional components such as human serum albumin (HSA), polysorbate 80, sugars or amino acids.

The nature and manner of the administration and the dosage
30 of the pharmaceutical composition according to the invention depend on the medical condition to be treated and its stage of progression, as well as on the body weight, age and sex of the patient.

The concentration of the modified mRNA in such formulations may therefore vary within a wide range from 1 µg to 100 mg/ml. The pharmaceutical composition according to the invention is preferably administered parenterally, for example intravenously, intraarterially, subcutaneously or intramuscularly to the patient. It is also possible to administer the pharmaceutical composition topically or orally.

10

The invention thus also provides a method for the treatment of the aforementioned medical conditions or an inoculation method for the prevention of the aforementioned conditions, which comprises the administration of the pharmaceutical composition according to the invention to a patient, in particular a human patient.

In addition a method is also provided that serves for the determination of the modified sequence of the mRNA contained in the pharmaceutical composition according to the invention. In this connection, according to the invention the adaptation of the RNA sequences is carried out with two different optimisation objectives: on the one hand with the largest possible G/C content, and on the other hand taking into account as far as possible the frequency of the tRNAs according to codon usage. In the first step of the process a virtual translation of an arbitrary RNA (or DNA) sequence is carried out in order to generate the corresponding amino acid sequence. Starting from the amino acid sequence, a virtual reverse translation is performed that provides, on account of the degenerated genetic code, possible choices for the corresponding codons. Depending on the required optimisation or

modification, corresponding selection lists and optimisation algorithms are used for choosing the suitable codons. The algorithms are executed on a computer, normally with the aid of suitable software. Thus, the
5 optimised mRNA sequence is generated and can be output for example with the aid of a suitable display device and compared with the original (wild type) sequence. The same also applies as regards the frequency of the individual nucleotides. The changes compared to the original
10 nucleotide sequence are preferably emphasised. Furthermore, according to a preferred embodiment stable sequences known in nature are read in, which can form the basis for an RNA stabilised according to natural sequence motifs. A secondary structure analysis may also be
15 provided that can analyse, on the basis of structural calculations, stabilising and destabilising properties or regions of the RNA.

In the Figures:

20

Fig. 1 shows wild type sequences and modified sequences for the influenza matrix protein.

Fig. 1A shows the wild type gene and Fig. 1B shows the amino acid sequence derived therefrom (1-letter code).

25 Fig. 1C shows a gene sequence coding for the influenza matrix protein, whose G/C content is raised compared to the wild type sequence. Fig. 1D shows the sequence of a gene that codes for a secreted form of the influenza matrix protein (including an N-terminal signal sequence), the G/C
30 content of the sequence being raised compared to the wild type sequence. Fig. 1E shows an mRNA coding for the influenza matrix protein that contains stabilising sequences compared to the wild type mRNA. Fig. 1F shows an

mRNA coding for the influenza matrix protein that in addition to stabilising sequences also contains an increased G/C content.

Fig. 1G likewise shows a modified mRNA that codes for the
5 secreted form of the influenza matrix protein and comprises, compared to the wild type, stabilising sequences and a raised G/C content. In Fig. 1A and Figs. 1C to 1G the start and stop codons are shown in bold type. Nucleotides that are changed compared to the wild type
10 sequence of Fig. 1A are shown in capital letters in 1C to 1G.

Fig. 2 shows wild type sequences and modified sequences according to the invention that code for the tumour antigen
15 MAGE1.

Fig. 2A shows the sequence of the wild type gene and Fig. 2B shows the amino acid sequence derived therefrom (3-letter code). Fig. 2C shows a modified mRNA coding for MAGE1, whose G/C content is raised compared to the wild
20 type. Fig. 2D shows the sequence of a modified mRNA coding for MAGE1, in which the codon usage has been optimised as frequently as possible in the tRNA present in the cell, with regard to the coding. Start and stop codons are shown in each case in bold type.

25

The following examples describe the invention in more detail without however restricting the scope thereof.

Example 1

30

As an exemplary embodiment of the process for determining the sequence of a modified mRNA according to the invention, a computer program was established that modifies, with the

aid of the genetic code or its degenerative nature, the nucleotide sequence of an arbitrary mRNA in such a way that a maximum G/C content in conjunction with the use of codons that code for tRNAs occurring as frequently as possible in the cell, is obtained, in which the amino acid sequence coded by the modified mRNA is identical compared to the unmodified sequence. Alternatively only the G/C content or only the codon usage compared to the original sequence may also be modified.

10

The source code in Visual Basic 6.0 (program development environment employed: Microsoft Visual Studio Enterprise 6.0 with Servicepack 3) is given in the Annexe..

15 **Example 2**

An RNA construct with a sequence of the lac-Z gene from *E.coli* optimised as regards stabilisation and translation efficiency was produced with the aid of the computer program of Example 1. A G/C content of 69% (compared to the wild type sequence of 51%; c.f. Kalnins et al., EMBO J. 1983, 2(4): 593-597) could be achieved in this way. Through the synthesis of overlapping oligonucleotides that comprise the modified sequence, the optimised sequence was produced according to methods known in the prior art. The terminal oligonucleotides have the following restriction cleavage sites: at the 5' end an *EcoRV* cleavage site, and at the 3' end a *BglIII* cleavage site. The modified lacZ sequence was incorporated into the plasmid pT7Ts (GeneBank Accession No. U 26404; c.f. Lai et al., see above) by digestion with *EcoRV/BglIII*. pT7Ts contains as non-translated regions sequences from the β -globin gene of *Xenopus laevis*, in each case, at 5' and 3'. The plasmid

30

was cleaved with the aforementioned restriction enzymes before the insertion of the modified lacZ sequence.

The pT7Ts-lac-Z construct was propagated in bacteria and
5 purified by phenol-chloroform extraction. 2 µg of the construct were transcribed *in vitro* by means of methods known to the person skilled in the art, whereby the modified mRNA was produced.

10 Example 3

The gene for the influenza matrix protein (wild type sequence, see Fig. 1A, derived amino acid sequence Fig. 1B) was optimised with the aid of the computer program
15 according to the invention of Example 1. The G/C-rich sequence variant shown in Fig. 1C was thereby formed. A G/C-rich sequence coding for the secreted form of the influenza matrix protein and which codes for an N-terminal signal sequence was also determined (see Fig. 1D). The
20 secreted form of the influenza matrix protein has the advantage of an increased immunogenicity compared to the non-secreted form.

Corresponding mRNA molecules were designed starting from
25 the optimised sequences. The mRNA for the influenza matrix protein, optimised as regards G/C content and codon usage, was additionally provided with stabilising sequences in the 5' region and 3' region (the stabilisation sequences derive from the 5'-UTRs and 3'-UTRs of the β -globin-mRNA of *Xenopus*
30 *laevis*; c.f. pT7Ts-Vektor in Lai et al., see above) (see Figs. 1E and 1F). The mRNA coding for the secreted form of the influenza matrix protein was likewise also sequence

optimised in the translated region and provided with the
aforementioned stabilising sequences (see Fig. 1G).

5

Example 4

10 The mRNA coding for the tumour antigen MAGE1 was modified
with the aid of the computer program of Example 1. The
sequence shown in Fig. 2C was determined in this way, and
has a 24% higher G/C content (351 G, 291 C) compared to the
wild type sequence (275 G, 244 G). In addition, by means
15 of alternative codon usage, the wild type sequence was
improved as regards the translation efficiency on the basis
of the coding of tRNAs occurring more frequently in the
cell (see Fig. 2D). The G/C content was likewise raised by
24% by the alternative codon usage.

20

Appendix: Source text of a computer program according to the invention

5 Curevac_Genetic_Controls with the following modules

Name: Curevac_Genetic_Controls.vbp

Code:

```

Type=Control
10 Reference=*G{00020430-0000-0000-C000-000000000046}#2.0#0#..\WINNT\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.OCX
UserControl=Curevac_Amino.ctl
Startup="(None)"
15 HelpFile=""
Title="Curevac Genetic Controls"
Command32=""
Name="Curevac_Genetic_Controls"
HelpContextID="0"
20 CompatibleMode="1"
MajorVer=1
MinorVer=0
RevisionVer=0
AutoIncrementVer=0
25 ServerSupportFiles=0
VersionComments="the RNA people"
VersionCompanyName="CureVac GmbH"
VersionFileDescription="Controls for Handling of Nucleotids"
VersionLegalCopyright="by Christian Klump"
30 VersionLegalTrademarks="Curevac Genetic Controls(tm)"
VersionProductName="Curevac Genetic Controls"
CompilationType=0
OptimizationType=0
FavorPentiumPro(tm)=0
35 CodeViewDebugInfo=0
NoAliasing=0
BoundsCheck=0
OverflowCheck=0
FlPointCheck=0
40 FDIVCheck=0
UnroundedFP=0
StartMode=1
Unattended=0
Retained=0
45 ThreadPerObject=0
MaxNumberOfThreads=1
ThreadingModel=1

```

DebugStartupOption=1
 DebugStartupComponent=Curevac_Amino

```

5      Name: Curevac_Amino.ctl
      Code:
      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
15      ClientWidth     = 1200
          ScaleHeight    = 690
          ScaleWidth     = 1200
          Begin VB.Line linLower
              X1          = 0
20      X2              = 1080
              Y1          = 600
              Y2          = 600
          End
          Begin VB.Line linUpper
25      X1              = 120
              X2          = 1080
              Y1          = 0
              Y2          = 0
          End
30     Begin MSForms.Label lblAminoAcid
          DragMode       = 1 'Automatic
          Height         = 255
          Left           = 120
          TabIndex       = 1
35     Top             = 360
          Width          = 975
          Size           = "1720;450"
          SpecialEffect  = 1
          FontName       = "Lucida Sans Unicode"
40     FontEffects     = 1073741826
          FontHeight     = 165
          FontCharSet    = 0
          FontPitchAndFamily= 2
          ParagraphAlign = 3
45     End
          Begin MSForms.Label lblTriplet
          DragMode       = 1 'Automatic
          Height         = 255
          Left           = 120
  
```

```

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

      Private msAAShortcut As String
20     Private msAAName As String
      Private msBestTriplett As String
      Private msSecondBest As String
      Private msThirdBest As String
      Private msTriplett As String
25     Private msBackColor As Long
      Private mbShowOriginal As Boolean

      Public Enum enuAminoAcid
30         amaGlycin
            amaAlanin
            amaValin
            amaLeucin
            amaIsoLeucin
35         amaPhenylalanin
            amaTyrosin
            amaTryptophan
            amaAsparaginAcid
            amaAsparagin
40         amaGlutaminAcid
            amaGlutamin
            amaSerin
            amaThreonin
            amaCystein
45         amaMethionin
            amaProlin
            amaHistidin
            amaLysin
            amaStop

```

```

    amaStart
End Enum

Private Sub Main()
5
    Call UserControl_Resize

End Sub

10 Public Sub Translation(ByVal sTriplet As String)

    msTriplet = sTriplet

    Select Case sTriplet
15      Case "GGU", "GGC", "GGA", "GGG"
        msAAShortcut = "GLY"
      Case "GCU", "GCC", "GCA", "GCG"
        msAAShortcut = "ALA"
      Case "GUU", "GUC", "GUA", "GUG"
20      msAAShortcut = "VAL"
      Case "UUA", "UUG", "CUU", "CUA", "CUG", "CUC"
        msAAShortcut = "LEU"
      Case "AUA", "AUU", "AUC"
        msAAShortcut = "ILE"
25      Case "UUU", "UUC"
        msAAShortcut = "PHE"
      Case "UAU", "UAC"
        msAAShortcut = "TYR"
      Case "UGG"
30      msAAShortcut = "TRP"
      Case "GAU", "GAC"
        msAAShortcut = "ASP"
      Case "AAU", "AAC"
        msAAShortcut = "ASN"
35      Case "GAA", "GAG"
        msAAShortcut = "GLU"
      Case "CAA", "CAG"
        msAAShortcut = "GLN"
      Case "AGU", "AGC", "UCA", "UCU", "UCG", "UCC"
40      msAAShortcut = "SER"
      Case "ACA", "ACU", "ACG", "ACC"
        msAAShortcut = "THR"
      Case "UGU", "UGC"
        msAAShortcut = "CYS"
45      Case "AUG"
        msAAShortcut = "MET"
      Case "CCA", "CCU", "CCG", "CCC"
        msAAShortcut = "PRO"
      Case "CAU", "CAC"

```

```

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

    End Sub

    Public Sub Synthesis(ByVal sShortCut As String)
15
        msAAShortcut = sShortCut 'nur falls die Sub extern angesprochen wird

        Select Case msAAShortcut
            Case "GLY"
20             msAAName = "Glycin"
                msBestTriplett = "GGC"
                msSecondBest = "GGG"
                msBackColor = 8421631
            Case "ALA"
25             msAAName = "Alanin"
                msBestTriplett = "GCG"
                msSecondBest = "GCC"
                msBackColor = 8454143
            Case "VAL"
30             msAAName = "Valin"
                msBestTriplett = "GUC"
                msSecondBest = "GUG"
                msBackColor = 8454016
            Case "LEU"
35             msAAName = "Leucin"
                msBestTriplett = "CUG"
                msSecondBest = "CUC"
                msBackColor = 16777088
            Case "ILE"
40             msAAName = "Isoleucin"
                msBestTriplett = "AUC"
                msBackColor = 12615935
            Case "PHE"
                msAAName = "Phenylalanin"
45             msBestTriplett = "UUC"
                msBackColor = 255
            Case "TYR"
                msAAName = "Tyrosin"
                msBestTriplett = "UAC"

```

```
msBackColor = 4210816
Case "TRP"
  msAAName = "Tryptophan"
  msBestTriplett = "UGG"
5  msBackColor = 4227327
Case "ASP"
  msAAName = "Asparaginsäure"
  msBestTriplett = "GAC"
  msBackColor = 8388863
10 Case "ASN"
  msAAName = "Asparagin"
  msBestTriplett = "AAC"
  msBackColor = 4227200
Case "GLU"
15  msAAName = "Glutaminsäure"
  msBestTriplett = "GAG"
  msBackColor = 32768
Case "GLN"
  msAAName = "Glutamin"
20  msBestTriplett = "GGC"
  msBackColor = 8421440
Case "SER"
  msAAName = "Serin"
  msBestTriplett = "AGC"
25  msSecondBest = "UCG"
  msThirdBest = "UCC"
  msBackColor = 16512
Case "THR"
  msAAName = "Threonin"
30  msBestTriplett = "ACG"
  msBackColor = 16711680
Case "CYS"
  msAAName = "Cystein"
  msBestTriplett = "UGC"
35  msBackColor = 6932960
Case "MET"
  msAAName = "Methionin"
  msBestTriplett = "AUG"
  msBackColor = 10417643
40 Case "PRO"
  msAAName = "Prolin"
  msBestTriplett = "CCG"
  msSecondBest = "CCC"
  msBackColor = 12898746
45 Case "HIS"
  msAAName = "Histidin"
  msBestTriplett = "CAC"
  msBackColor = 12898746
Case "ARG"
```

```

        msAAName = "Arginin"
        msBestTriplett = "CGC"
        msSecondBest = "CGG"
        msBackColor = 6174925
5      Case "LYS"
        msAAName = "Lysin"
        msBestTriplett = "AAG"
        msBackColor = 14141641
      Case "STP"
10      msAAName = "Stop"
        msBestTriplett = "UGA"
        msSecondBest = "UAG"
        msBackColor = 11332093
      End Select
15    End Sub

    Public Sub Show()

20      lblAminoAcid.Caption = msAAShortcut
      If mbShowOriginal = True Then
        lblTriplett.Caption = msTriplett
      Else
        lblTriplett.Caption = msBestTriplett
25      End If
      lblAminoAcid.BackColor = msBackColor
      lblTriplett.BackColor = msBackColor

    End Sub
30    Public Function ChooseBestGC(sShortCut As String)

      Select Case sShortCut
        Case "GLY"
35          msAAName = "Glycin"
          msBestTriplett = "GGC"
          msSecondBest = "GGG"
          msBackColor = 8421631
        Case "ALA"
40          msAAName = "Alanin"
          msBestTriplett = "GCG"
          msSecondBest = "GCC"
          msBackColor = 8454143
        Case "VAL"
45          msAAName = "Valin"
          msBestTriplett = "GUC"
          msSecondBest = "GUG"
          msBackColor = 8454016
        Case "LEU"

```

```

msAAName = "Leucin"
msBestTriplett = "CUG"
msSecondBest = "CUC"
msBackColor = 16777088
5 Case "ILE"
  msAAName = "Isoleucin"
  msBestTriplett = "AUC"
  msBackColor = 12615935
10 Case "PHE"
  msAAName = "Phenylalanin"
  msBestTriplett = "UUC"
  msBackColor = 255
  Case "TYR"
15   msAAName = "Tyrosin"
   msBestTriplett = "UAC"
   msBackColor = 4210816
  Case "TRP"
20   msAAName = "Tryptophan"
   msBestTriplett = "UGG"
   msBackColor = 4227327
  Case "ASP"
   msAAName = "Asparaginsäure"
   msBestTriplett = "GAC"
   msBackColor = 8388863
25 Case "ASN"
  msAAName = "Asparagin"
  msBestTriplett = "AAC"
  msBackColor = 4227200
  Case "GLU"
30   msAAName = "Glutaminsäure"
   msBestTriplett = "GAG"
   msBackColor = 16711808
  Case "GLN"
   msAAName = "Glutamin"
35   msBestTriplett = "GGC"
   msBackColor = 12632256
  Case "SER"
   msAAName = "Serin"
   msBestTriplett = "AGC"
40   msSecondBest = "UCG"
   msThirdBest = "UCC"
   msBackColor = 8421504
  Case "THR"
   msAAName = "Threonin"
45   msBestTriplett = "ACG"
   msBackColor = &HD4C0FF
  Case "CYS"
   msAAName = "Cystein"
   msBestTriplett = "UGC"

```

```

        msBackColor = &HD5C0FF
    Case "MET"
        msAAName = "Methionin"
        msBestTriplett = "AUG"
5       msBackColor = &HD6C0FF
    Case "PRO"
        msAAName = "Prolin"
        msBestTriplett = "CCG"
        msSecondBest = "CCC"
10      msBackColor = &HD7C0FF
    Case "HIS"
        msAAName = "Histidin"
        msBestTriplett = "CAC"
        msBackColor = &HD8C0FF
15      Case "ARG"
            msAAName = "Arginin"
            msBestTriplett = "CGC"
            msSecondBest = "CGG"
            msBackColor = &HD9C0FF
20      Case "LYS"
            msAAName = "Lysin"
            msBestTriplett = "AAG"
            msBackColor = &HE0C0FF
        Case "STP"
25      msAAName = "Stop"
            msBestTriplett = "UGA"
            msSecondBest = "UAG"
            msBackColor = &HE1C0FF
    End Select
30
End Function

35
Public Property Let ShowOriginal(ByVal bNewValue As Boolean)
    mbShowOriginal = bNewValue
End Property

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

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

Public Property Get AAName() As String
    AAName = msAAName

```

```

End Property

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

Public Property Get SecondBest() As String
    SecondBest = msSecondBest
End Property
10 Public Property Get ThirdBest() As String
    ThirdBest = msThirdBest
End Property

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

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

Private Sub UserControl_Resize()
25 Dim lOuterWidth As Long
    Dim lDistance As Long

    lDistance = 30
30 lOuterWidth = ScaleWidth + lDistance

    With lblTriplet
35 .Top = 0
        .Left = lDistance / 2
        .Width = ScaleWidth - lDistance
        .Height = ScaleHeight / 2 + 30
    End With
40 With lblAminoAcid
        .Top = ScaleHeight / 2 - 30
        .Left = lDistance / 2
        .Width = ScaleWidth - lDistance
        .Height = ScaleHeight / 2
45 End With
    With linUpper
        .X1 = ScaleLeft
        .Y1 = 0
        .X2 = ScaleWidth

```

40

```

        .Y2 = 0
    End With
    With linLower
        .X1 = ScaleLeft
5       .Y1 = ScaleHeight - 20
        .X2 = ScaleWidth
        .Y2 = ScaleHeight - 20
    End With

10    End Sub

15    Curevac_RNA_Optimizer with following modules

    Name: Curevac_RNA_Optimizer.vbp
    Code:
    Type=Exe
20    Reference=*\G{00020430-0000-0000-C000-
    000000000046}#2.0#0#..\..\WINNT\System32\STDOLE2.TLB#OLE Automation
    Object={F9043C88-F6F2-101A-A3C9-08002B2F49FB}#1.2#0; COMDLG32.OCX
    Module=basMain; basMain.bas
    Form=mdiMain.frm
25    Object=*\ACurevac_Genetic_Controls.vbp
    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
30    Module=basPublics; basPublics.bas
    Form=frmOptions.frm
    Form=frmInput.frm
    Form=frmDisplay.frm
    Form=frmAbout.frm
35    Startup="mdiMain"
    HelpFile=""
    Title="RNA-Optimizer"
    Command32=""
    Name="RNA_Optimizer"
40    HelpContextID="0"
    CompatibleMode="0"
    MajorVer=1
    MinorVer=0
    RevisionVer=0
45    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
5 OptimizationType=0
FavorPentiumPro(tm)=0
CodeViewDebugInfo=0
NoAliasing=0
BoundsCheck=0
10 OverflowCheck=0
FIPointCheck=0
FDIVCheck=0
UnroundedFP=0
StartMode=0
15 Unattended=0
Retained=0
ThreadPerObject=0
MaxNumberOfThreads=1
DebugStartupOption=0
20

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

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

35 Private mstrAminoAcids() As udtAminoAcid

Public Sub RebuildChain(bShowOriginal As Boolean)

Dim iLoop As Integer
40 For iLoop = 1 To glAminoCounter
    With frmDisplay.Curevac_Amino1(iLoop)
        .ShowOriginal = bShowOriginal
        .Show
45 End With
    Next iLoop

End Sub

```

```

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

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

10 Dim IVerticalOffset As Long
    Dim IHorizontalOffset As Long
    Dim ITop As Long

    Dim asTripletts() As String
15 Dim sAminoChain As String
    Dim lLoop As Long

    If bGraphical = True Then
        ITop = frmDisplay.optNucleotids(0).Height + 40
20 End If
    ReDim mastrAminoAcids(glAminoCounter)
    asTripletts = Split(sTempCode)
    For lLoop = 0 To glAminoCounter - 1
        If bGraphical = True Then
25 Load frmDisplay.Curevac_Amino1(lLoop + 1)
            With frmDisplay.Curevac_Amino1(lLoop + 1)

                Call .Translation(asTripletts(lLoop))

30 .Left = lLoop * .Width - IHorizontalOffset
                .Top = IVerticalOffset + ITop
                If .Left + 2 * .Width > frmDisplay.ScaleWidth Then
                    IVerticalOffset = IVerticalOffset + .Height + 150
                    IHorizontalOffset = lLoop * .Width + .Width
35 End If

                sAminoChain = sAminoChain + .aaShortcut + ","
                .Visible = True
            End With
40 Else
                mastrAminoAcids(lLoop) = Translation(asTripletts(lLoop))
                sTest = sTest + mastrAminoAcids(iLoop).sName + ","
            End If
        Next
45 If bGraphical = True Then
            frmDisplay.Show
        Else
            Call frmOutPut.Show
        End If
    End Sub

```

```

    Debug.Print sAminoChain
End Sub

```

```

Private Function Translation(ByVal sTriplet As String) As udtAminoAcid
5
    Dim stiTemp As udtAminoAcid

    If mdiMain.optKindOfOptimize(0) = True Then
        'Optimization for Most GC
10    Select Case sTriplet
        Case "GGU", "GGC", "GGA", "GGG"
            stiTemp.sShortcut = "GLY"
            stiTemp.sName = "Glycin"
            stiTemp.sBestTriplet = "GGC"
15            stiTemp.sSecondBest = "GGG"
        Case "GCU", "GCC", "GCA", "GCG"
            stiTemp.sShortcut = "ALA"
            stiTemp.sName = "Alanin"
            stiTemp.sBestTriplet = "GCG"
20            stiTemp.sSecondBest = "GCC"
        Case "GUU", "GUC", "GUA", "GUG"
            stiTemp.sShortcut = "VAL"
            stiTemp.sName = "Valin"
            stiTemp.sBestTriplet = "GUC"
25            stiTemp.sSecondBest = "GUG"
        Case "UUA", "UUG", "CUU", "CUA", "CUG", "CUC"
            stiTemp.sShortcut = "LEU"
            stiTemp.sName = "Leucin"
            stiTemp.sBestTriplet = "CUG"
30            stiTemp.sSecondBest = "CUC"
        Case "AUA", "AUU", "AUC"
            stiTemp.sShortcut = "ILE"
            stiTemp.sName = "Isoleucin"
            stiTemp.sBestTriplet = "AUC"
35            stiTemp.sSecondBest = "AUC"
        Case "UUU", "UUC"
            stiTemp.sShortcut = "PHE"
            stiTemp.sName = "Phenylalanin"
            stiTemp.sBestTriplet = "UUC"
        Case "UAU", "UAC"
40            stiTemp.sShortcut = "TYR"
            stiTemp.sName = "Tyrosin"
            stiTemp.sBestTriplet = "UAC"
        Case "UGG"
            stiTemp.sShortcut = "TRP"
45            stiTemp.sName = "Tryptophan"
            stiTemp.sBestTriplet = "UGG"
        Case "GAU", "GAC"
            stiTemp.sShortcut = "ASP"
            stiTemp.sName = "Asparaginsäure"

```

```

        stiTemp.sBestTriplett = "GAC"
    Case "AAU", "AAC"
        stiTemp.sShortcut = "ASN"
        stiTemp.sName = "Asparagin"
5       stiTemp.sBestTriplett = "AAC"
    Case "GAA", "GAG"
        stiTemp.sShortcut = "GLU"
        stiTemp.sName = "Glutaminsäure"
        stiTemp.sBestTriplett = "GAG"
10      Case "CAA", "CAG"
        stiTemp.sShortcut = "GLN"
        stiTemp.sName = "Glutamin"
        stiTemp.sBestTriplett = "CAG"
    Case "AGU", "AGC", "UCA", "UCU", "UCG", "UCC"
15      stiTemp.sShortcut = "SER"
        stiTemp.sName = "Serin"
        stiTemp.sBestTriplett = "AGC"
        stiTemp.sSecondBest = "UCG"
        stiTemp.sThirdBest = "UCC"
20      Case "ACA", "ACU", "ACG", "ACC"
        stiTemp.sShortcut = "THR"
        stiTemp.sName = "Threonin"
        stiTemp.sBestTriplett = "ACG"
    Case "UGU", "UGC"
25      stiTemp.sShortcut = "CYS"
        stiTemp.sName = "Cystein"
        stiTemp.sBestTriplett = "UGC"
    Case "AUG"
        stiTemp.sShortcut = "MET"
30      stiTemp.sName = "Methionin"
        stiTemp.sBestTriplett = "AUG"
    Case "CCA", "CCU", "CCG", "CCC"
        stiTemp.sShortcut = "PRO"
        stiTemp.sName = "Prolin"
35      stiTemp.sBestTriplett = "CCG"
        stiTemp.sSecondBest = "CCC"
    Case "CAU", "CAC"
        stiTemp.sShortcut = "HIS"
        stiTemp.sName = "Histidin"
40      stiTemp.sBestTriplett = "CAC"
    Case "AGA", "AGG", "CGA", "CGU", "CGG", "CGC"
        stiTemp.sShortcut = "ARG"
        stiTemp.sName = "Arginin"
        stiTemp.sBestTriplett = "CGC"
45      stiTemp.sSecondBest = "CGG"
    Case "AAA", "AAG"
        stiTemp.sShortcut = "LYS"
        stiTemp.sName = "Lysin"
        stiTemp.sBestTriplett = "AAG"

```

```

    Case "UAA", "UAG", "UGA"
        stiTemp.sShortcut = "STP"
        stiTemp.sName = "Stop"
        stiTemp.sBestTriplett = "UGA"
5       stiTemp.sSecondBest = "UAG"
    End Select
    Translation = stiTemp
Else 'Optimization for best frequency
    Select Case sTriplett
10      Case "GGU", "GGC", "GGA", "GGG"
            stiTemp.sShortcut = "GLY"
            stiTemp.sName = "Glycin"
            stiTemp.sBestTriplett = "GGC"
            stiTemp.sSecondBest = "GGU"
15      Case "GCU", "GCC", "GCA", "GCG"
            stiTemp.sShortcut = "ALA"
            stiTemp.sName = "Alanin"
            stiTemp.sBestTriplett = "GCC"
            stiTemp.sSecondBest = "GCU"
20      Case "GUU", "GUC", "GUA", "GUG"
            stiTemp.sShortcut = "VAL"
            stiTemp.sName = "Valin"
            stiTemp.sBestTriplett = "GUG"
            stiTemp.sSecondBest = "GUC"
25      Case "UUA", "UUG", "CUU", "CUA", "CUG", "CUC"
            stiTemp.sShortcut = "LEU"
            stiTemp.sName = "Leucin"
            stiTemp.sBestTriplett = "CUG"
            stiTemp.sSecondBest = "CUC"
30      Case "AUA", "AUU", "AUC"
            stiTemp.sShortcut = "ILE"
            stiTemp.sName = "Isoleucin"
            stiTemp.sBestTriplett = "AUC"
            Case "UUU", "UUC"
35             stiTemp.sShortcut = "PHE"
                stiTemp.sName = "Phenylalanin"
                stiTemp.sBestTriplett = "UUC"
            Case "UAU", "UAC"
40             stiTemp.sShortcut = "TYR"
                stiTemp.sName = "Tyrosin"
                stiTemp.sBestTriplett = "UAC"
            Case "UGG"
45             stiTemp.sShortcut = "TRP"
                stiTemp.sName = "Tryptophan"
                stiTemp.sBestTriplett = "UGG"
            Case "GAU", "GAC"
                stiTemp.sShortcut = "ASP"
                stiTemp.sName = "Asparaginsäure"
                stiTemp.sBestTriplett = "GAC"

```

```

Case "AAU", "AAC"
    stiTemp.sShortcut = "ASN"
    stiTemp.sName = "Asparagin"
    stiTemp.sBestTriplett = "AAC"
5 Case "GAA", "GAG"
    stiTemp.sShortcut = "GLU"
    stiTemp.sName = "Glutaminsäure"
    stiTemp.sBestTriplett = "GAG"
10 Case "CAA", "CAG"
    stiTemp.sShortcut = "GLN"
    stiTemp.sName = "Glutamin"
    stiTemp.sBestTriplett = "CAG"
Case "AGU", "AGC", "UCA", "UCU", "UCG", "UCC"
15 stiTemp.sShortcut = "SER"
    stiTemp.sName = "Serin"
    stiTemp.sBestTriplett = "AGC"
    stiTemp.sSecondBest = "UCC"
    stiTemp.sThirdBest = "UCU"
Case "ACA", "ACU", "ACG", "ACC"
20 stiTemp.sShortcut = "THR"
    stiTemp.sName = "Threonin"
    stiTemp.sBestTriplett = "ACC"
Case "UGU", "UGC"
    stiTemp.sShortcut = "CYS"
25 stiTemp.sName = "Cystein"
    stiTemp.sBestTriplett = "UGC"
Case "AUG"
    stiTemp.sShortcut = "MET"
    stiTemp.sName = "Methionin"
30 stiTemp.sBestTriplett = "AUG"
Case "CCA", "CCU", "CCG", "CCC"
    stiTemp.sShortcut = "PRO"
    stiTemp.sName = "Prolin"
    stiTemp.sBestTriplett = "CCC"
35 stiTemp.sSecondBest = "CCU"
Case "CAU", "CAC"
    stiTemp.sShortcut = "HIS"
    stiTemp.sName = "Histidin"
    stiTemp.sBestTriplett = "CAC"
40 Case "AGA", "AGG", "CGA", "CGU", "CGG", "CGC"
    stiTemp.sShortcut = "ARG"
    stiTemp.sName = "Arginin"
    stiTemp.sBestTriplett = "CGC"
    stiTemp.sSecondBest = "AGG"
45 Case "AAA", "AAG"
    stiTemp.sShortcut = "LYS"
    stiTemp.sName = "Lysin"
    stiTemp.sBestTriplett = "AAG"
Case "UAA", "UAG", "UGA"

```

47

```

        stiTemp.sShortcut = "STP"
        stiTemp.sName = "Stop"
        stiTemp.sBestTriplett = "UGA"
        stiTemp.sSecondBest = "UAG"
5      End Select
      Translation = stiTemp
    End If

  End Function

10  Public Function ReverseTranscription() As String
    Dim sTempCode As String
    Dim iLoop As Integer
    For iLoop = 0 To glAminoCounter - 1
15      sTempCode = sTempCode + mastrAminoAcids(iLoop).sBestTriplett
    Next
    ReverseTranscription = sTempCode
  End Function

20  Public Sub CountBases(ByVal sTempCode As String)

    Dim lLoop As Long
    Dim sFormattedCode As String
    Dim sActualBase As String
25  Dim lCodeLength As Long

    lCodeLength = Len(sTempCode)
    For lLoop = 1 To CLng(lCodeLength)
      sActualBase = Mid(sTempCode, lLoop, 1)
30      Select Case sActualBase
        Case "A", "a"
          glOptAdenin = glOptAdenin + 1
        Case "U", "u", "T", "t"
          glOptThymin = glOptThymin + 1
35      Case "G", "g"
          glOptGuanin = glOptGuanin + 1
        Case "C", "c"
          glOptCytosin = glOptCytosin + 1
      End Select
40    Next lLoop
  End Sub

45  Name: basPublics.bas
  Code:
  Attribute VB_Name = "basPublics"
  Option Explicit

```

```

Public gsOriginalCode As String
Public gsFormattedCode As String
Public glCodeLength As Long

5  Public glAminoCounter As Long
   Public glAdenin As Long
   Public glThymin As Long
   Public glGuanin As Long
   Public glCytosin As Long
10  Public glOptAdenin As Long
   Public glOptThymin As Long
   Public glOptGuanin As Long
   Public glOptCytosin As Long

15  Public gbSequenceChanged As Boolean

Name: mdiMain.frm
20  Code:
   VERSION 5.00
   Object = "{38911DA0-E448-11D0-84A3-00DD01104159}#1.1#0"; "COMCT332.OCX"
   Begin VB.MDIForm mdiMain
       BackColor      = &H00FFFFFF&
25  Caption          = "Curevac_RNA_Analyzer"
       ClientHeight   = 8745
       ClientLeft     = 165
       ClientTop      = 735
       ClientWidth    = 12255
30  Icon             = "mdiMain.frx":0000
       LinkTopic      = "MDIForm1"
       Picture         = "mdiMain.frx":058A
       StartupPosition = 3 'Windows Default
       WindowState    = 2 'Maximized
35  Begin ComCtl3.CoolBar cbarMain
       Align          = 1 'Align Top
       Height         = 885
       Left           = 0
       TabIndex       = 0
40  Top              = 0
       Width          = 12255
       _ExtentX       = 21616
       _ExtentY       = 1561
       BandCount      = 1
45  BandBorders      = 0 'False
       _CBWidth       = 12255
       _CBHeight      = 885
       _Version       = "6.7.8988"
       MinHeight1     = 825

```

```

Width1      = 4995
FixedBackground1= 0 'False
NewRow1     = 0 'False
Begin VB.OptionButton optKindOfOptimize
5   Caption   = "Best Frequency"
    Height    = 255
    Index     = 1
    Left      = 4920
    TabIndex  = 7
10   Top      = 480
    Value     = -1 'True
    Width     = 1575
End
Begin VB.OptionButton optKindOfOptimize
15   Caption   = "Best GC"
    Height    = 255
    Index     = 0
    Left      = 4920
    TabIndex  = 6
20   Top      = 120
    Width     = 1575
End
Begin VB.CommandButton cmdShowInput
25   Caption   = "Input"
    Height    = 765
    Left      = 0
    Picture    = "mdiMain.frx":0C2A
    Style     = 1 'Graphical
    TabIndex  = 5
30   Top      = 0
    Width     = 840
End
Begin VB.CommandButton cmdShowOptions
35   Caption   = "Options"
    Enabled   = 0 'False
    Height    = 760
    Left      = 3480
    Picture    = "mdiMain.frx":0F34
    Style     = 1 'Graphical
40   TabIndex  = 4
    Top      = 0
    Width     = 735
End
Begin VB.CommandButton cmdShowStatistics
45   Caption   = "Statistics"
    Enabled   = 0 'False
    Height    = 760
    Left      = 2640
    Picture    = "mdiMain.frx":131E

```

50

```

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

```

```

End
End
Begin VB.Menu mnuMain
  Caption      = "E&xtras"
5  Index       = 4
  Begin VB.Menu mnuExtras
    Caption     = "&Language"
    Index       = 0
    Begin VB.Menu mnuLanguage
10     Caption   = "English"
        Checked   = -1 'True
        Index     = 0
    End
  End
  Begin VB.Menu mnuLanguage
15     Caption   = "&German"
        Enabled   = 0 'False
        Index     = 1
    End
  End
  Begin VB.Menu mnuLanguage
20     Caption   = "&French"
        Enabled   = 0 'False
        Index     = 2
    End
  End
End
End
25  Begin VB.Menu mnuExtras
    Caption     = "&Options..."
    Enabled     = 0 'False
    Index       = 1
  End
End
30  Begin VB.Menu mnuExtras
    Caption     = "-"
    Index       = 2
  End
End
35  Begin VB.Menu mnuExtras
    Caption     = "&About"
    Index       = 3
  End
End
End
Begin VB.Menu mnuMain
40  Caption     = "&Windows"
    Index       = 5
    WindowList  = -1 'True
    Begin VB.Menu mnuWindows
      Caption   = "Tile &Horizontally"
45      Index     = 0
    End
    Begin VB.Menu mnuWindows
      Caption   = "Tile &Vertically"
      Index     = 1
    End
  End
End

```

```

End
Begin VB.Menu mnuWindows
    Caption      = "&Cascade"
    Index        = 2
5   End
End
Begin VB.Menu mnuMain
    Caption      = "&Exit"
    Index        = 6
10  End
End
Attribute VB_Name = "mdiMain"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = False
15  Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
Option Explicit

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

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

Private Sub cmdShowOptions_Click()
    Call frmOptions.Show
End Sub
30

Private Sub cmdShowOutput_Click()
    Call frmOutPut.Show
End Sub

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

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

Private Sub InitControls()
45    Dim lNextLeft As Long
    Const lSpace As Long = 60
    Const lButtonWidth As Long = 850
    Const lButtonHeight As Long = 770

```

```

    lNextLeft = lSpace + lButtonWidth

    With cmdShowInput
5      .Top = lSpace
        .Left = lSpace
        .Width = lButtonWidth
        .Height = lButtonHeight
    End With
10
    With cmdShowOutput
        .Top = lSpace
        .Left = lNextLeft + lSpace
        .Width = lButtonWidth
15      .Height = lButtonHeight
    End With

    With cmdShowDisplay
        .Top = lSpace
20      .Left = lNextLeft * 2 + lSpace
        .Width = lButtonWidth
        .Height = lButtonHeight
    End With

25    With cmdShowStatistics
        .Top = lSpace
        .Left = lNextLeft * 3 + lSpace
        .Width = lButtonWidth
        .Height = lButtonHeight
30    End With

    With cmdShowOptions
        .Top = lSpace
        .Left = lNextLeft * 4 + lSpace
35      .Width = lButtonWidth
        .Height = lButtonHeight
    End With

    cbarMain.Height = lButtonHeight + 2 * lSpace
40

End Sub

Private Sub mnuMain_Click(Index As Integer)
45
    Select Case Index
        Case 0
            Call frmInput.Show
        Case 6

```

```

        Unload Me
    End Select

End Sub

5 Private Sub mnuResults_Click(Index As Integer)

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

    End Sub

    Private Sub mnuExtras_Click(Index As Integer)
20      Select Case Index
        Case 0 'Output
            '
        Case 1
25      Call frmDisplay.Show
        Case 3
            Call frmAbout.Show(vbModal)
        End Select

30      End Sub

    Private Sub mnuWindows_Click(Index As Integer)

35      Select Case Index
        Case 0 'Horizontal
            Me.Arrange (vbTileHorizontal)
        Case 1 'Vertikal
            Me.Arrange (vbTileVertical)
40      Case 2 'Kaskadieren
            Me.Arrange (vbCascade)
        End Select

    End Sub

45

    Name: frmInput.frm
    Code:
    VERSION 5.00

```

Object = "{F9043C88-F6F2-101A-A3C9-08002B2F49FB}#1.2#0"; "COMDLG32.OCX"
 Object = "{3B7C8863-D78F-101B-B9B5-04021C009402}#1.2#0"; "RICHTEXT32.OCX"

Begin VB.Form frmInput

```

Caption      = "Input"
5 ClientHeight = 5730
  ClientLeft  = 60
  ClientTop   = 345
  ClientWidth = 13200
  Icon        = "frmInput.frx":0000
10 LinkTopic   = "Form1"
  MDIChild    = -1 'True
  ScaleHeight = 5730
  ScaleWidth  = 13200
  WindowState = 2 'Maximized
15 Begin RichTextLib.RichTextBox txtFormatted
  Height      = 1815
  Left        = 120
  TabIndex    = 3
  Top         = 360
20 Width       = 12735
  _ExtentX    = 22463
  _ExtentY    = 3201
  _Version    = 393217
  BorderStyle = 0
25 Enabled     = -1 'True
  ScrollBars   = 2
  DisableNoScroll = -1 'True
  TextRTF      = "$"frmInput.frx":030A
  BeginProperty Font {0BE35203-8F91-11CE-9DE3-00AA004BB851}
30   Name       = "Fixedsys"
     Size      = 9
     Charset   = 0
     Weight    = 400
     Underline = 0 'False
35   Italic     = 0 'False
     Strikethrough = 0 'False
  EndProperty
End
Begin VB.OptionButton optSequence
40  Caption     = "Formatted"
     Height    = 255
     Index     = 1
     Left      = 2760
     Style     = 1 'Graphical
45  TabIndex    = 2
     Top       = 0
     Value     = -1 'True
     Width     = 855
End

```

```

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

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

40   Public Sub LoadSequence()

        Dim sFile As String
        Dim sTempCode As String

45       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 = ""
5    gsOriginalCode = ""
    sFile = CommonDialog1.FileName
    Open sFile For Input As #1
    Input #1, gsOriginalCode
    Close #1
10    Call FormatText
    Call FillSequence

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

Exit Sub

20  Errorhandler:

    Dim lAnswer As Long

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

End Sub

Private Sub FormatText()
35    Dim iLoop As Integer
    Dim sActualBase As String

    glCodeLength = Len(gsOriginalCode)
40    For iLoop = 0 To glCodeLength - 1
        sActualBase = Mid(gsOriginalCode, iLoop + 1, 1)
        If iLoop Mod 3 = 0 Then
            gsFormattedCode = gsFormattedCode + " "
            glAminoCounter = glAminoCounter + 1
45    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"
5           gsFormattedCode = gsFormattedCode + "G"
            glGuanin = glGuanin + 1
        Case "C", "c"
            gsFormattedCode = gsFormattedCode + "C"
            glCytosin = glCytosin + 1
10        End Select
        Next iLoop
        gsFormattedCode = Trim(gsFormattedCode)

    End Sub

15 Private Sub FillSequence()
    If optSequence(0).Value = True Then
        txtFormatted.Text = gsOriginalCode
    Else
20        txtFormatted.Text = gsFormattedCode
    End If

    End Sub

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

    Private Sub cmdLoad_Click()
30        Call LoadSequence
    End Sub

    Private Sub InitControls()

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

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

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

```

```

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

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

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

20
Name: frmOutPut.frm
Code:
VERSION 5.00
Object = "{3B7C8863-D78F-101B-B9B5-04021C009402}#1.2#0"; "RICHTX32.OCX"
25 Begin VB.Form frmOutPut
    Caption      = "OutPut"
    ClientHeight = 9120
    ClientLeft   = 60
    ClientTop    = 345
30 ClientWidth  = 10215
    Icon         = "frmOutPut.frx":0000
    LinkTopic    = "Form1"
    MDIChild     = -1 'True
    ScaleHeight  = 9120
35 ScaleWidth   = 10215
    Begin RichTextLib.RichTextBox txtOptimized
        Height      = 2655
        Left        = 0
        TabIndex    = 0
40 Top          = 480
        Width       = 9855
        _ExtentX    = 17383
        _ExtentY    = 4683
        _Version    = 393217
45 Enabled      = -1 'True
        ScrollBars  = 2
        DisableNoScroll = -1 'True
        TextRTF     = $"frmOutPut.frx":0442
        BeginProperty Font {0BE35203-8F91-11CE-9DE3-00AA004BB851}

```

60

```

        Name      = "Fixedsys"
        Size       = 9
        Charset    = 0
        Weight     = 400
5      Underline  = 0 'False
        Italic     = 0 'False
        Strikethrough = 0 'False
    EndProperty
End
10 End
    Attribute VB_Name = "frmOutPut"
    Attribute VB_GlobalNameSpace = False
    Attribute VB_Creatable = False
    Attribute VB_PredeclaredId = True
15  Attribute VB_Exposed = False
    Option Explicit

    Private Sub Form_Activate()

20      Me.ZOrder
        Call InitControls
        If gbSequenceChanged = True Then
            Call BuildAminoChain(gsFormattedCode, False)
        End If
25      gbSequenceChanged = False
        txtOptimized.Text = LCase(ReverseTranscription)

    End Sub

30

    Private Sub InitControls()

        Const lSpace As Long = 40

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

45  Name: frmDisplay.frm
    Code:
    VERSION 5.00
    Object = "*\ACurevac_Genetic_Controls.vbp"
    Begin VB.Form frmDisplay

```

```

Caption      = "Display"
ClientHeight = 8205
ClientLeft   = 60
ClientTop    = 345
5 ClientWidth = 8745
Icon         = "frmDisplay.frx":0000
LinkTopic    = "Form1"
MDIChild     = -1 'True
ScaleHeight  = 8205
10 ScaleWidth = 8745
ShowInTaskbar = 0 'False
Begin VB.OptionButton optNucleotids
    Caption    = "Original"
    Height     = 255
15    Index    = 0
    Left       = 0
    Style      = 1 'Graphical
    TabIndex   = 1
    Top        = 0
20    Width    = 855
End
Begin VB.OptionButton optNucleotids
    Caption    = "Optimized"
    Height     = 255
25    Index    = 1
    Left       = 840
    Style      = 1 'Graphical
    TabIndex   = 0
    Top        = 0
30    Value    = -1 'True
    Width      = 855
End
Begin Curevac_Genetic_Controls.Curevac_Amino Curevac_Amino1
    Height     = 495
35    Index    = 0
    Left       = 240
    Top        = 360
    Visible    = 0 'False
    Width      = 735
40    _ExtentX = 1296
    _ExtentY   = 873
End
End
Attribute VB_Name = "frmDisplay"
45 Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = False
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
Option Explicit

```

```
Private Sub Form_Activate()
```

```

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

10  End Sub

```

```
Private Sub optNucleotids_Click(Index As Integer)
```

```

15  Dim bShowOriginal As Boolean

    If Index = 0 Then 'Original Clicked
        bShowOriginal = True
    Else 'Optimized Clicked
20    bShowOriginal = False
    End If
    Call RebuildChain(bShowOriginal)

    End Sub
25

```

Name: frmStatistics.frm

Code:

```

30  VERSION 5.00
    Begin VB.Form frmStatistics
        Caption      = "Statistics"
        ClientHeight = 6450
        ClientLeft   = 60
35    ClientTop     = 345
        ClientWidth  = 8595
        Icon         = "frmStatistics.frx":0000
        LinkTopic    = "Form1"
        MDIChild     = -1 'True
40    ScaleHeight   = 6450
        ScaleWidth   = 8595
        Begin VB.Frame Frame1
            Caption    = "original"
            Height     = 2655
45    Left         = 120
            TabIndex   = 9
            Top        = 720
            Width      = 2175
            Begin VB.Label lblAdenin

```

```

        BorderStyle = 1 'Fixed Single
        Height      = 495
        Left        = 840
        TabIndex     = 17
5       Top         = 240
        Width       = 1215
    End
    Begin VB.Label lblCytosin
        BorderStyle = 1 'Fixed Single
10      Height      = 495
        Left        = 840
        TabIndex     = 16
        Top         = 2040
        Width       = 1215
15      End
    Begin VB.Label lblGuanin
        BorderStyle = 1 'Fixed Single
        Height      = 495
        Left        = 840
20      TabIndex     = 15
        Top         = 1440
        Width       = 1215
    End
    Begin VB.Label lblThymin
25      BorderStyle = 1 'Fixed Single
        Height      = 495
        Left        = 840
        TabIndex     = 14
        Top         = 840
30      Width       = 1215
    End
    Begin VB.Label Label14
        Caption      = "Cytosin"
        Height      = 255
35      Left        = 120
        TabIndex     = 13
        Top         = 2160
        Width       = 615
    End
    Begin VB.Label Label2
40      Caption      = "Thymin"
        Height      = 255
        Left        = 120
        TabIndex     = 12
        Top         = 960
45      Width       = 615
    End
    Begin VB.Label Label3
        Caption      = "Guanin"

```

```

    Height      = 255
    Left        = 120
    TabIndex    = 11
    Top         = 1560
5    Width      = 615
End
Begin VB.Label Label1
    Caption     = "Adenin"
    Height      = 255
10   Left       = 120
    TabIndex    = 10
    Top         = 360
    Width       = 615
End
15 End
Begin VB.Frame Frame2
    Caption     = "optimized"
    Height      = 2655
    Left        = 2520
20   TabIndex   = 0
    Top         = 720
    Width       = 2175
Begin VB.Label Label4
    Caption     = "Adenin"
25   Height     = 255
    Left        = 120
    TabIndex    = 8
    Top         = 360
    Width       = 615
30 End
Begin VB.Label Label6
    Caption     = "Cytosin"
    Height      = 255
    Left        = 120
35   TabIndex   = 7
    Top         = 2160
    Width       = 615
End
Begin VB.Label Label8
    Caption     = "Guanin"
40   Height     = 255
    Left        = 120
    TabIndex    = 6
    Top         = 1560
45   Width      = 615
End
Begin VB.Label Label12
    Caption     = "Thymin"
    Height      = 255

```

```

Left      = 120
TabIndex  = 5
Top       = 960
Width     = 615
5 End
Begin VB.Label lblOptAdenin
  BorderStyle = 1 'Fixed Single
  Height      = 495
  Left        = 840
10  TabIndex   = 4
  Top         = 240
  Width       = 1215
End
Begin VB.Label lblOptThymin
15  BorderStyle = 1 'Fixed Single
  Height      = 495
  Left        = 840
  TabIndex    = 3
  Top         = 840
20  Width      = 1215
End
Begin VB.Label lblOptGuanin
  BorderStyle = 1 'Fixed Single
  Height      = 495
25  Left       = 840
  TabIndex    = 2
  Top         = 1440
  Width       = 1215
End
30 Begin VB.Label lblOptCytosin
  BorderStyle = 1 'Fixed Single
  Height      = 495
  Left        = 840
  TabIndex    = 1
35  Top        = 2040
  Width       = 1215
End
End
Begin VB.Label lblBases
40  BorderStyle = 1 'Fixed Single
  Height      = 495
  Left        = 1320
  TabIndex    = 19
  Top         = 120
45  Width      = 1215
End
Begin VB.Label Label5
  Caption     = "Sum of bases"
  Height      = 255

```

66

```

        Left      = 120
        TabIndex   = 18
        Top        = 240
        Width      = 1095
5    End
End
Attribute VB_Name = "frmStatistics"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = False
10 Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
Option Explicit

Private Sub Form_Activate()
15    Me.ZOrder

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

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

30 End Sub

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

```

```

Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = False
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
5 Option Explicit

Name: frmAbout.frm
10 Code:
VERSION 5.00
Begin VB.Form frmAbout
    BorderStyle = 3 'Fixed Dialog
    Caption = "About Curevac RNA-Analyzer"
15 ClientHeight = 2490
    ClientLeft = 2340
    ClientTop = 1935
    ClientWidth = 6195
    ClipControls = 0 'False
    LinkTopic = "Form2"
20 MaxButton = 0 'False
    MinButton = 0 'False
    ScaleHeight = 1718.642
    ScaleMode = 0 'User
    ScaleWidth = 5817.425
25 ShowInTaskbar = 0 'False
    Begin VB.PictureBox picIcon
        AutoSize = -1 'True
        ClipControls = 0 'False
30 Height = 1215
        Left = 120
        Picture = "frmAbout.frx":0000
        ScaleHeight = 811.195
        ScaleMode = 0 'User
        ScaleWidth = 2401.98
35 TabIndex = 1
        Top = 120
        Width = 3480
    End
40 Begin VB.CommandButton cmdOK
    Cancel = -1 'True
    Caption = "OK"
    Default = -1 'True
    Height = 345
45 Left = 2520
    TabIndex = 0
    Top = 2040
    Width = 1260
End

```

```

Begin VB.Label lblCompany
  Caption      = "lblCompany"
  Height       = 255
  Left         = 3720
5   TabIndex    = 5
  Top          = 1200
  Width        = 2295
End
Begin VB.Line Line1
10  BorderColor = &H00808080&
  BorderStyle  = 6 'Inside Solid
  Index        = 1
  X1           = 112.686
  X2           = 5746.996
15  Y1          = 1325.218
  Y2          = 1325.218
End
Begin VB.Label lblDescription
20  Caption     = "App Description"
  ForeColor    = &H00000000&
  Height       = 330
  Left         = 120
  TabIndex     = 2
  Top          = 1560
25  Width       = 5925
End
Begin VB.Label lblTitle
  Caption      = "Application Title"
  ForeColor    = &H00000000&
30  Height      = 360
  Left         = 3720
  TabIndex     = 3
  Top          = 240
  Width        = 2325
35  End
Begin VB.Line Line1
  BorderColor  = &H00FFFFFF&
  BorderWidth  = 2
  Index        = 0
40  X1          = 112.686
  X2          = 5746.996
  Y1          = 1325.218
  Y2          = 1325.218
End
45  Begin VB.Label lblVersion
  Caption      = "Version"
  Height       = 225
  Left         = 3720
  TabIndex     = 4

```

```
        Top      = 720
        Width    = 2325
    End
End
5  Attribute VB_Name = "frmAbout"
   Attribute VB_GlobalNameSpace = False
   Attribute VB_Creatable = False
   Attribute VB_PredeclaredId = True
   Attribute VB_Exposed = False
10 Option Explicit

   Private Sub cmdOK_Click()
       Unload Me
   End Sub
15

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

25
```

DEMANDES OU BREVETS VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVETS
COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 2

NOTE: Pour les tomes additionels, veuillez contacter le Bureau Canadien des Brevets.

JUMBO APPLICATIONS / PATENTS

THIS SECTION OF THE APPLICATION / PATENT CONTAINS MORE
THAN ONE VOLUME.

THIS IS VOLUME 1 OF 2

NOTE: For additional volumes please contact the Canadian Patent Office.

Claims:

1. A pharmaceutical composition comprising at least modified mRNA that codes for at least one antigenic viral peptide or polypeptide in combination with a pharmaceutically compatible carrier and/or vehicle, characterised in that the G/C content of the region of the modified mRNA coding for the peptide or polypeptide is larger than the G/C content of the coding region of the wild type mRNA coding for the peptide or polypeptide and the coded amino acid sequence is unchanged compared to the wild type mRNA.
2. The pharmaceutical composition according to claim 1, characterised in that the G/C content of the region of the modified mRNA coding for the peptide or polypeptide is at least 7% points larger than the G/C content of the coding region of the wild type mRNA coding for the peptide or polypeptide.
3. The pharmaceutical composition according to claim 1 or 31, characterised in that all codons of the wild type sequence that code for a relatively rare tRNA in the cell are exchanged in each case for a codon that codes for a relatively common tRNA in the cell that in each case carries the same amino acid as the relatively rare tRNA.
4. The pharmaceutical composition according to claim 1 or 2, characterised in that the region of the modified mRNA coding for the peptide is altered in such a way as

to produce a maximum G/C content in conjunction with the codons that code for the relatively frequent tRNAs.

5. The pharmaceutical composition according to any one of
5 claims 1 to 4, characterised in that at least one of the region coding for the peptide, the 5' non-translated region and the 3' non-translated region of the modified mRNA is altered compared to the wild type mRNA in such a way that it does not contain any
10 destabilising sequence elements.
6. The pharmaceutical composition according to any one of
claims 1 to 5, characterised in that the modified mRNA has at least one of a 5' cap structure, or a poly-A
15 tail of at least 70 nucleotides, an IRES, a 5' stabilisation sequence, and a 3' stabilisation sequence.
7. The pharmaceutical composition according to any one of
20 claims 1 to 6, characterised in that the modified mRNA comprises at least one analogue of naturally occurring nucleotides.
8. The pharmaceutical composition according to claim 7,
25 characterised in that the analogue is selected from the group consisting of phosphorus thioates, phosphorus amidates, peptide nucleotides, methylphosphonates, 7-deazaguanosine, 5-methylcytosine and inosine.
9. The pharmaceutical composition according to any one of
30 claims 1 to 8, characterised in that the mRNA additionally codes for at least one cytokine.

10. The pharmaceutical composition according to any one of claims 1 to 8, which in addition contains at least one cytokine.
- 5 11. Use of the pharmaceutical composition according to any one of claims 1 to 8 for inoculation against viral infectious diseases.
- 10 12. Use of the pharmaceutical composition according to any one of claims 1 to 10 for inoculation against AIDS, hepatitis A, B or C, herpes, herpes zoster, rubella, dengue, haemorrhagic fever, yellow fever and influenza.
- 15 13. A modified mRNA coding for at least one antigenic viral peptide or polypeptide, characterised in that the G/C content of the region of the modified mRNA coding for the peptide or polypeptide is increased compared to the G/C content of the coding region of the wild type mRNA coding for the peptide or polypeptide, and the encoded amino acid sequence is unchanged as compared to the wild type mRNA.
- 20 14. The modified mRNA according to claim 13, characterised in that the G/C content of the region of the modified mRNA coding for the peptide or polypeptide is increased by at least 7 % points compared to the G/C content of the coding region of the wild type mRNA coding for the peptide or polypeptide.
- 25 15. The modified mRNA according to claim 13, characterised in that the G/C content of the region of the modified mRNA coding for the peptide or polypeptide is increased

by at least 15 % points compared to the G/C content of the coding region of the wild type mRNA coding for the peptide or polypeptide.

- 5 16. The modified mRNA according to claim 13, 14 or 15, characterised in that the modified mRNA comprises at least one of a 5' cap structure, a poly-A tail of at least 70 nucleotides, an IRES, a 5' stabilisation sequence, and a 3' stabilisation sequence.
- 10 17. The modified mRNA according to any one of claims 13 to 16, characterised in that the modified mRNA comprises at least one analogue of naturally occurring nucleotides.
- 15 18. The modified mRNA according to claim 17, characterised in that the analogue is selected from the group consisting of phosphorus thioates, phosphorus amidates, peptide nucleotides, methylphosphonates, 7-
- 20 deazaguanosine, 5-methylcytosine, and inosine.
19. The modified mRNA according to any one of claims 13 to 18, characterised in that the mRNA codes for a surface antigen of pathogenic viral germs.
- 25 20. The modified mRNA according to any one of claims 13 to 19, characterised in that the antigenic viral peptide or polypeptide is also a tumour antigen.
- 30 21. The modified mRNA according to any one of claims 13 to 20, characterised in that the polypeptide is a polypeptide of viral antigens.

22. The modified mRNA according to any one of claims 13 to 21, characterised in that the mRNA additionally codes for at least one cytokine.
- 5 23. The modified mRNA according to any one of claims 13 to 22, characterised in that the mRNA is associated with a cationic peptide or protein or is bound thereto.
- 10 24. The modified mRNA according to claim 23, characterised in that the cationic peptide or protein is selected from the group consisting of protamine, poly-L lysine and histones.
- 15 25. A pharmaceutical composition, characterised in that the pharmaceutical composition comprises modified mRNA according to any one of claims 13 to 24 in combination with a pharmaceutically acceptable carrier and/or vehicle.
- 20 26. The pharmaceutical composition according to claim 25, characterised in that the pharmaceutical composition comprises at least one immune response stimulating adjuvant.
- 25 27. The pharmaceutical composition according to claim 25 or 26, which in addition comprises at least one cytokine.
- 30 28. Use of a pharmaceutical composition according to any one of claims 25 to 27 or a modified mRNA according to any one of claims 13 to 24 for the preparation of a

vaccine for inoculation against viral infectious diseases.

29. Use of a pharmaceutical composition according to any
5 one of claims 25 to 27 for the preparation of a vaccine
for the treatment of at least one of: AIDS, hepatitis
A, B or C, herpes, herpes zoster, rubella, dengue,
haemorrhagic infectious diseases, yellow fever and
influenza.

Fig. 1A

Influenza matrix: wild type gene (for comparison)

agatctaaagatgagtccttctaaccgaggtcgaaacgtacgttctctcta
tcatcccgtcaggccccctcaaagccgagatcgcacagagacttgaagat
gtctttgcagggaagaacaccgatcttgaggttctcatggaatggctaaa
gacaagaccaatcctgtcacctctgactaaggggatttttaggatttgtgt
tcacgctcacctgcccagtgagcagaggactgcagcgtagacgctttgtc
caaatgcccttaatgggaacggggatccaaataacatggacaaagcagt
taaactgtataggaagctcaagagggagataacattccatggggccaaag
aatctcactcagttattctgctggtgcacttgccagttgtatgggcctc
atatacaacaggatgggggctgtgaccactgaagtggcatttggcctggt
atgtgcaacctgtgaacagattgctgactcccagcatcggtctcataggc
aatggtgacaacaaccaaccactaatcagacatgagaacagaatggtt
ttagccagcactacagctaaggctatggagcaaattggctggatcgagtga
gcaagcagcagaggccatggaggttgctagtcaggctaggcaaattggtgc
aagcgatgagaaccattgggactcatcctagctccagtgctggtctgaaa
aatgatcttcttgaaaatttgaggcctatcagaaacgaatgggggtgca
gatgcaacggttcaagtgaactag

Fig. 1B

Influenza matrix: protein sequence

MSLLTEVETYVLSIIPSGPLKAEIAQRLEDVFAGKNTDLEVLMEWLKTRP
ILSPLTKGILGFVFTLTVPSERGLQRRRFVQNALNGNGDPNNMDKAVKLY
RKLKREITFHGAKEISLSYSAGALASCMGLIYNRMGAVTTEVAFGLVCAT
CEQIADSQHRSHRQMVTNTNPLIRHENRMVLASTTAKAMEQMAGSSEQAA
EAMEVASQARQMVMRTIGTHPSSSAGLKNDLLENLQAYQKRMGVQMQR
FK*

2/6

Fig. 1C

Influenza matrix: gene with increased G/C content

agatctaaagatgagCctGctGaccgaggtGgaGacCtacgtGctGAGCa
 tcatcccCAGCggccccctGaaGgccgagatcgcCcagagGctGgaGgaC
 gtGttCgcCggCaagaacaccgaCctGgaggtGctGatggaGtggctGaa
 gacCagGccCatcctgAGCccCctgacCaagggCatCCTGggCttCgtgt
 tcacCctGaccgtgcccagCgagcgCggCctgcagcgCCGCcgcttCgtG
 caGaaCgcccctGaaCggCaacggCgaCccCaaCaacatggacaaGgcCgt
 GaaGctgtaCaggaagctGaagagggagatCacCttccaCggCgccaAGg
 aGatcAGCctGagCtaCAGCgcCggCgcCctGgccagCtgCatgggcctG
 atCtacaacaggatgggCgcCgtgaccacCgaGgtggcCttCggcctggt
 GtgCgcCacctgCgaGcagatCgcCgacAGCcagcaCcgCAGCcaCaggc
 aGatggtgacCacCaccaacccCctGatcagGcaCgagaacagGatggtG
 CTGgccagcacCacCgcCaaggcCatggagcaGatggcCggCAGCaGCga
 gcaGgcCgcCgaggccatggaggtGgcCagCcaggcCaggcaGatggtgc
 aGgcCatgagGaccatCggCacCcaCccCagcAGCagCgcCggCctgaaG
 aaCgaCctGctGgaGaaCCTGcaggcctaCcagaaGcgCatgggCgtgca
 gatgcaGcgCttcaagtgaactagt

Fig. 1D

Influenza matrix: gene for secreted form (with N-terminal signal sequence) with increased G/C content

AgatctaaagatgGCCGTCATGGCCCCCGCACCCCTGGTGCTGCTGCTGA
 GCGGCGCCCTGGCCCTGACCCAGACCTGGGCTagCctGctGaccgaggtG
 gaGacCtacgtGctGAGCatcatcccCAGCggccccctGaaGgccgagat
 cgcCcagagGctGgaGgaCgtGttCgcCggCaagaacaccgaCctGgagg
 tGctGatggaGtggctGaagacCagGccCatcctgAGCccCctgacCaag
 ggCatCCTGggCttCgtgttcacCctGaccgtgcccagCgagcgCggCct
 gcagcgCCGCcgcttCgtGcaGaaCgcccctGaaCggCaacggCgaCccCa
 aCaacatggacaaGgcCgtGaaGctgtaCaggaagctGaagagggagatC
 acCttccaCggCgccaAGgaGatcAGCctGagCtaCAGCgcCggCgcCct
 GgccagCtgCatgggcctGatCtacaacaggatgggCgcCgtgaccacCg
 aGgtggcCttCggcctggtGtgCgcCacctgCgaGcagatCgcCgacAGC
 cagcaCcgCAGCcaCaggcaGatggtgacCacCaccaacccCctGatcag
 GcaCgagaacagGatggtGCTGgccagcacCacCgcCaaggcCatggagc
 aGatggcCggCAGCaGCgagcaGgcCgcCgaggccatggaggtGgcCagC
 caggcCaggcaGatggtgcaGgcCatgagGaccatCggCacCcaCccCag
 cAGCagCgcCggCctgaaGaaCgaCctGctGgaGaaCCTGcaggcctaCc
 agaaGcgCatgggCgtgcagatgcaGcgCttcaagtgaactagt

3/6

Fig. 1E

Influenza matrix: mRNA with stabilisation sequences

GCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUCAACUUUGGCagauc
 uaaagau**g**agucucuaaaccgaggucgaaacguacguucucucuauc
 ccgucaggcccccuaaagccgagaucgcacagagacuugaagaugucuu
 ugcagggaagaacaccgaucugagguucuauggaauaggcuaaagaca
 gaccaauccugucaccucugacuaaggggaauuuaggauuuguguucacg
 cucaccgugcccagugagcgaggacugcagcguagacgcuuuguccaaaa
 ugcccuuaauagggaacggggauccaaauaacauggacaaagcaguuaaac
 uguauaggaagcucaagagggagauaacaauccauggggccaaagaauc
 ucacucaguuaauucugcuggugcacuugccaguuguauggggccucaua
 caacaggauugggggcugugaccacugaaguggcauuuggccugguaugug
 caaccugugaacagauugcugacuccagcaucggucucauaggcaaau
 gugacaacaaccaaccacuaaucagacauagagaacagaauugguuuagc
 cagcacuacagcuaaggcuauaggagcaaaugggcuggaucgagugagcaag
 cagcagaggccauggagguugcuagucaggcuaggcaaauggugcaagcg
 augagaaccauugggacucauccuagcuccagugcuggucugaaaauga
 ucuucugaaaaauugcaggccuaucaagaacgaauugggggugcagaugc
 aacgguucaaguga**ACUAGUGACUGACUAGCCCGCUGGGCCUCCCAACGG**
GCCCUCCUCCCUCCUUGCACCAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Fig. 1F

Influenza matrix: mRNA with increased G/C content and stabilisation sequences

GCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUCAACUUUGGCagauc
 uaaagau**g**agCcuGcuGaccgagguGgaGacCuacguGcuGAGCaucauc
 ccCAGCggcccccGaaGgccgagaucgcCagagGcuGgaGgaCguGuu
 CgcCggCaagaacaccgaCcuGgagguGcuGauggaGuggcuGaagacCa
 gGccCauccugAGCccCugacCaagggCauCCUGggCuuCguguucacC
 cuGaccgugcccagCgagcgCggCcuagcagcgCCGCcgcuuCguGcaGaa
 CgcccGaaCggCaacggCgaCccCaaCaacauggacaaGgcCguGaaGc
 uguaCaggaagcuGaagagggagauCacCuuccaCggCgccaGgaGauc
 AGCcuGagCuaCAGCgcCggCgcCcuGgccagCugCauugggcuGauCua
 caacaggauugggCgcCgugaccacCgaGguggcCuuCggccugguGugCg
 cCaccugCgaGcagauCgcCgacAGCagcaCcgCAGCcaCaggcaGaug
 gugacCacCaccaaccCcuGaucagGcaCgagaacagGaugguGCUGgc
 cagcacCacCgcCaaggcCauaggagcaGauaggcCggCAGCaGCgagcaGg
 cCgcCgaggccauggagguGgcCagCaggcCaggcaGauaggugcaGgcC
 augagGaccuauCggCacCcaCccCagcAGCagCgcCggCcuGaaGaaCga
 CcuGcuGgaGaaCCUGcaggccuaCcagaaGcgCauugggCgugcagaugc
 aGcgCuucaaguga**ACUAGUGACUGACUAGCCCGCUGGGCCUCCCAACGG**
GCCCUCCUCCCUCCUUGCACCAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Fig. 1G

Influenza matrix: mRNA coding for secreted form with increased G/C content and stabilisation sequences

[illegible]

5/6

Fig. 2A

MAGE1: wild type gene (for comparison)

catcatgtctcttgagcagaggagtctgcactgcaagcctgaggaagccc
 ttgaggcccaacaagaggccctgggctggtgtgtgtgcaggctgccacc
 tcctcctcctctcctctggtcctgggcaccctggaggaggtgccactgc
 tgggtcaacagatcctccccagagtcctcagggagcctccgcctttccca
 ctaccatcaacttcactcgacagaggcaacccagtgagggttccagcagc
 cgtgaagaggaggggccaagcacctcttgatcctggagtccttggtccg
 agcagtaatcactaagaagggtggctgatttggttggttttctgctcctca
 aatatcgagccaggagccagtcacaaaggcagaaatgctggagagtgctc
 atcaaaaattacaagcactgttttcctgagatcttcggcaaagcctctga
 gtccttgagctggtccttggtgacgtgaaggaagcagacccaccg
 gccactcctatgtccttggtcactgcctaggtctctcctatgatggcctg
 ctgggtgataatcagatcatgccaagacaggcttcctgataattgtcct
 ggtcatgattgcaatggaggcgccatgctcctgaggaggaaatctggg
 aggagctgagtgatggaggtgtatgatgggaggaggacacagtgccat
 ggggagcccaggaagctgctcacccaagatttggtgcaggaaaagtacct
 ggagtaccggcaggtgccggacagtgatcccgcacgctatgagttcctgt
 ggggtccaagggccctcgctgaaaccagctatgtgaaagtccttgagtat
 gtgatcaaggtcagtgcaagagttcgctttttcttcccatccctgcgtga
 agcagctttgagagaggaggaagaggaggagtctgagcatga

Fig. 2B

MAGE1: protein sequence

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

Fig. 2C

MAGE1: mRNA with increased G/C content

augagccuggagcagcgcagccugcacugcaagccggaggaggcgcuggaggcgcagcagga
ggcgcuggggccuggucugcguccaggcggcgacgagcagcagcagcccgcugguccugggca
cgcuggaggaggucccgacggcgggcagcacggacccgcccgcagagcccgcagggcgcgagc
gcuucccgacgacgaucaacuucacgcgccagcgccagccgagcaggggcagcagcagccg
cgaggaggaggggcccgcagcagcagcugcauccuggagagccuguuccgcgcggucaucacga
agaaggucgcggaccuggucggcuuccugcugcugaaguaccgcgcgcgcgagccggucacg
aaggcggagaugcuggagagcgucaucaagaacuacaagcacugcuucccgagaucuucgg
caaggcgagcgagagccugcagcuggucucggcaucgacgucaaggaggcggacccgacgg
gccacagcuacguccuggucacgugccuggggccugagcuacgacggccugcuggggcgacaac
cagaucaugcccgaagacgggcuuccugaucaucguccuggucaugaucgcgauggaggggcg
ccacgcgccggaggaggagaucugggaggagcugagcgucauggaggucacgacggccgcg
agcacagcgcuacggcgagccgcgcgaagcugcugacgcaggaccugguccaggagaaguac
cuggaguaccgccaggucccgacagcgacccggcgcgcuacgaguuccuguggggcccgcg
cgcgcuuggcgagacgagcuacgucaaggucuggaguacgucaucaaggucagcgcgcg
uccgcucucucucccgagccugcgcgaggcgcgcgagggaggaggaggggcgucuga
gcgugauga

Fig. 2D

MAGE1: mRNA with alternative codon usage

augagccuggagcagcgcagccugcacugcaagcccaggaggcccuggaggcccagcagga
ggcccuggggccuggugugcgugcaggccgccaccagcagcagcagccccuggugcugggca
cccuggaggaggugcccaccgcccggcagcacccgacccccccagagccccagggcgccagc
gccuuccccaccaccaucaacuucacccgccagcgccagcccagcgaggggcagcagcagccg
cgaggaggaggggcccagcaccagcugcauccuggagagccuguuccgcgcgugaucacca
agaagguggccgaccuggugggcuuccugcugcugaaguaccgcgcccgcgagcccugacc
aaggccgagaugcuggagagcgugaaucaagaacuacaagcacugcuuccccgagaucuucgg
caaggccagcgagagccugcagcuggugucggcaucgacgugaaggaggccgacccccaccg
gccacagcuacgugcuggugaccugccuggggccugagcuacgacggccugcuggggcgacaac
cagaucaugcccgaagaccggcuuccugaucaucgugcuggugaugaucgccauggaggggcg
ccacgcccccgaggaggagaucugggaggagcugagcgugauggaggugacgacggccgcg
agcacagcgccuacggcgagccccgcaagcugcugacccaggaccuggugcaggagaaguac
cuggaguaccgccaggugcccgcacagcgaccccgcccgcuaacgaguuccuguggggcccgcg
cgcccuggccgagaccagcuacgugaaggugcuggaguacgugaaucaaggugagcgcccgcg
ugcgcuucucucucccgagccugcgcgaggccgcccugcgcgagggaggaggaggggcguguga
gccugauga