

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization

International Bureau

(43) International Publication Date
3 May 2012 (03.05.2012)



(10) International Publication Number
WO 2012/056449 A3

(51) International Patent Classification:
C07H 21/00 (2006.01)

(21) International Application Number:
PCT/IL2011/000827

(22) International Filing Date:
27 October 2011 (27.10.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PCT/IL2010/000895
28 October 2010 (28.10.2010) IL

(71) Applicant (for all designated States except US): **NAN-ODOC LTD.** [IL/IL]; 10 Plaut Street, 76706 Rehovot (IL).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **ABITBOL, Guy** [IL/IL]; 12 Hakonhia Street, 75437 Rishon Lezion (IL).

(74) Agents: **WEBB, Cynthia** et al.; Webb & Co., P.O. Box 2189, 76121 Rehovot (IL).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,

HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))
- with sequence listing part of description (Rule 5.2(a))

(88) Date of publication of the international search report:
19 July 2012

(54) Title: COMPOSITIONS AND METHODS FOR SPECIFIC CLEAVAGE OF EXOGENOUS RNA IN A CELL

(57) Abstract: There are provided compositions for cleaving an exogenous RNA of interest only in the presence of an endogenous signal RNA sequence, thereby activating expression of a polynucleotide of interest only in the presence of the endogenous signal RNA sequence. There are provided methods for the preparation of the composition and uses thereof in treatment and diagnosis of various conditions and disorders, for example by selectively activating expression of a toxin only in specific target cell populations.



WO 2012/056449 A3

COMPOSITIONS AND METHODS FOR SPECIFIC CLEAVAGE OF EXOGENOUS RNA IN A CELL

FIELD OF THE INVENTION

The present invention relates to compositions for cleaving an exogenous RNA of
5 interest only in the presence of an endogenous signal RNA sequence, thereby activating
expression of a polynucleotide of interest only in the presence of the endogenous signal
RNA sequence. The invention further relates to uses of the compositions in treatment and
diagnosis of various conditions and disorders, as exemplified by selectively activating
expression of a toxin only in target cell populations.

10 BACKGROUND OF THE INVENTION

RNA interference (RNAi) is a phenomenon in which dsRNA, composed of sense
RNA and antisense RNA homologous to a certain region of a target gene effects cleavage of
the homologous region of the target gene transcript, thereby inhibiting expression of the
gene. In mammals, the dsRNA should be shorter than 31 base pairs to avoid induction of an
15 interferon response that can cause cell death by apoptosis. RNAi technology is based on a
natural mechanism that utilizes microRNAs (miRNAs) to regulate posttranscriptional gene
expression [1]. miRNAs are very small RNA molecules of about 21 nucleotides in length
that appear to be derived from 70-90 nucleotide precursors that form a predicted RNA stem-
loop structure. miRNAs are expressed in organisms as diverse as nematodes, fruit flies,
20 humans and plants.

In mammals, miRNAs are generally transcribed by RNA polymerase II and the
resulting primary transcripts (pri-miRNAs) contain local stem-loop structures that are
cleaved by the Drosha-DGCR8 complex. The product of this cleavage is one or more (in
case of clusters) precursor miRNA (pre-miRNA). Pre-miRNAs are usually 70-90
25 nucleotides long with a strong stem-loop structure, and they usually contain 2 nucleotides
overhang at the 3' end [2]. The pre-miRNA is transported to the cytoplasm by Exportin-5. In
the cytoplasm, the Dicer enzyme, which is an endoribonuclease of the RNase III family,
recognizes the stem in the pre-miRNA as dsRNA and cleaves and releases a 21 bp dsRNA
(miRNA duplex) from the 3' and 5' end of the pre-miRNA. The two strands of the duplex
30 are separated from each other by the Dicer-TRBP complex and the strand that has

thermodynamically weaker 5' end is incorporated into the RNA induced silencing complex (RISC) [3]. This strand is the mature miRNA. The opposite strand, which is not incorporated into RISC is called miRNA* strand and it is degraded [1]. The mature miRNA guides RISC to a target site within mRNAs. If the target site is near perfect complementarity to the mature miRNA, the mRNA will be cleaved at a position that is located about 10 nucleotides upstream from the 3' end of the target site [3]. After the cleavage, the RISC-mature miRNA strand complex is recycled for another round of activity [4]. If the target site has lower complementarity to the mature miRNA the mRNA will not be cleaved at the target site but the translation of the mRNA will be suppressed. Although about 530 miRNAs have been identified so far in humans, it is estimated that vertebrate genomes encode up to 1,000 unique miRNAs, which are predicted to regulate expression of at least 30% of the genes [5]. See FIG. 1.

The two portions of the mRNA cleaved by the RISC-mature miRNA strand complex in mammalian cells can be detected easily by Northern analysis [6]. Two RNA transcripts of about 23 nucleotides in length, which have a complementary region of about 19 nucleotides in length at the 5' end, are hybridized with each other in the mammalian cell and are capable of directing target specific RNA interference [7]. Sequence and structural features of double stranded (ds)RNA molecules required to mediate target-specific nucleic acid modifications such as RNA-interference and/or DNA methylation are disclosed, for example, in US 7,078,196 and US 7,055,704. A dsRNA 52 nucleotides long that further comprises 20 nucleotides long ssRNA at one of the 3' ends is a substrate for a Dicer only at the blunt end [8]. In mammals, Risc is coupled to Dicer [9]. While RNA polymerase III U6 promoter is a very strong promoter for transcribing small RNA (sRNA), RNA polymerase II CMV promoter is a strong promoter for transcribing protein-coding genes.

In mammalian cells, addition of a cap (7-methylguanosine cap) to the 5' end of a mRNA, increases the translation of the mRNA by 35-50 fold. Further, addition of a poly(A) tail to the 3' end of the mRNA increases the translation of the mRNA by 114-155- fold [10]. The poly(A) tail in mammalian cells increases the functional mRNA half-life by 2.6-fold and the cap increases the functional mRNA half-life by 1.7-fold [10]. The human HIST1H2AC (H2ac) gene encodes a member of the histone H2A family. Transcripts from this gene lack poly(A) tails but instead contain a palindromic termination element (5'-

GGCUCUUUUCAGAGCC-3') that forms a conserved stem-loop structure at the 3'-UTR, which plays an important role in mRNA processing and stability [11].

Ribosome inactivating proteins (RIPs) are protein toxins that are of plant or microbial origin. RIPs inhibit protein synthesis by inactivating ribosomes. Recent studies suggest that RIPs are also capable of inducing cell death by apoptosis. Type II RIPs contain a toxic A-chain and a lectin like subunit (B-chain) linked together by a disulfide bond. The B chain is catalytically inactive, but serves to mediate entry of the A-B protein complex into the cytosol. Ricin, Abrin and Diphtheria toxin are very potent Type II RIPs. It has been reported that a single molecule of Ricin or Abrin reaching the cytosol can kill the cell [12, 13]. In addition, a single molecule of Diphtheria toxin fragment A introduced into a cell can kill the cell [14].

According to the WHO (world health organization) in 2006 there were about 39.5 million people with HIV worldwide. Many viruses, including HIV exhibit a dormant or latent phase, during which little or no protein synthesis is conducted. The viral infection is essentially invisible to the immune system during such phases. Current antiviral treatment regimens are largely ineffective at eliminating cellular reservoirs of latent viruses [15]. Viruses may be oncogenic due to an oncogene in their genome. Retroviruses may also be oncogenic due to integration at a site which truncates a gene or which places a gene under control of the strong viral cis-acting regulatory element.

According to the American Cancer Society, 7.6 million people died from cancer in the world during 2007. The nature of and basic approaches to cancer treatment are constantly changing. Some approaches to cancer treatment, such as radiotherapy, surgery and inhibition of angiogenesis, are not useful against many small metastases. Other approaches to cancer treatment, such as inhibition of cell division and destroying dividing cells have no specificity and thus may cause harmful side effects that can even kill the patient. Further approaches for cancer treatment such as induction of differentiation of tumor tissues, inhibition of oncogenes, virus that contains ligands against membrane receptor protein that unique to cancer cells, manipulations of the immune system and immunotoxin therapy, have a narrow therapeutic index and usually are not sufficiently effective. Yet other approaches to cancer treatment using tumor suppressor genes and using toxins under a promoter that is uniquely activated in cancer cells have a narrow therapeutic

index, a great potential for causing harmful side effects and usually are not sufficiently effective.

Many viruses that cause cancer are capable of causing latent infection. KSHV (Kaposi sarcoma-associated herpesvirus) causes Kaposi's sarcoma cancer; SV40 (Simian vacuolating virus 40) has the potential to cause tumors, but most often persists as a latent infection; and EBV (Epstein-Barr virus) causes Burkitt's lymphoma, nasopharyngeal carcinoma and EBV-associated gastric carcinomas.

On average, each tumor contains mutations in about 90 protein-coding genes [16]. Each tumor is initiated from a single founder cell [38]. It is most probable that at least one of these mutant genes is transcribed to mRNA. Therefore, it is highly probable that each cell of a specific tumor or each cell that is infected by a specific virus includes an RNA molecule, which comprises a specific RNA sequence (signal sequence) that is unique to the mutated or infected cell and that is not present in other normal cells of the same organism. The signal sequence can be from viral origin or from the mutated gene, that is unique to the specific tumor.

Various methods have been developed to identify any specific sequence that is unique to a specific tumor. These methods include, for example, DNA microarray, Tilling (Targeting Induced Local Lesions In Genomes) and large-scale sequencing of cancer genomes. Furthermore, the identification of this signal sequence is predicted to be even simpler thanks to the Cancer Genome Atlas (performed by the NIH), which was launched on 13.12.2005 and has been cataloguing all the genetic mutations responsible for cancer.

Compositions for the selective killing of only those cells that contain a specific signal sequence, have been proposed. One approach, developed by Intronn Company, is to build an inactive Toxin that is activated by trans-splicing between the inactive Toxin to the signal sequence [17, 18 and 19]. However, this approach has several inherent problems: The first problem is that the RNA molecule that comprises the signal sequence must be present in the cell at very high copy number, since trans-splicing events are very rare. The second problem is that in most cases this approach is not suitable for a signal sequence that is of cancer origin, since in cancer, mutations spread over a short region. The third problem is that the trans-splicing events can also occur at random and may thus cause harmful side effects. The fourth problem is that the RNA molecule that comprises the signal sequence must include an intron at a very specific site. Another approach, which won the 2004 World

Technology Award in Biotechnology, suggested using small dsDNA, ssDNA, hairpin DNA and restriction enzyme, however this approach can work only in cell extracts under very unique and not under physiological conditions in living cells [20]. Other approach, such as disclosed, for example, in WO 07/00068 are directed to a gene vector and comprising a miRNA sequence target and its use to prevent or reduce expression of transgene in a cell which comprises a corresponding miRNA. Also disclosed, for example, in WO 2010/055413, is a gene vector adapted for transient expression of a transgene in a peripheral organ cell comprising a regulatory sequence operably linked to a transgene wherein the regulatory sequence prevents or reduces expression of said transgene in hematopoietic lineage cells.

There is therefore a need for developing new compositions that are capable of selectively kill only cells that contain a signal sequence, wherein the compositions should be potent, reliable and specific as compared to compositions used in the prior art. Since that the development of these compositions can be a very complex multi-step process there is also a need for developing compositions for activating genes of interest in cells, only in the presence of a signal sequence, and for cleaving exogenous RNA of interest only in the presence of a signal sequence.

SUMMARY OF THE INVENTION

The present invention provides compositions and methods for selectively cleaving an exogenous RNA of interest in response to the presence of an endogenous signal RNA in a cell. The exogenous RNA of interest is encoded by the composition. The endogenous signal RNA is an RNA molecule which comprises a predetermined signal sequence that is a sequence of 18-25 nucleotides long. Subsequent to specific cleavage of the exogenous RNA in the presence of the endogenous signal sequence transcription of a polynucleotide of interest may be activated. The polynucleotide activated may encode a toxin thereby providing means to kill target cell populations selectively.

The compositions of the invention comprise or, encode:

(a) an exogenous RNA of interest which is an RNA sequence that comprises a specific sequence that is of sufficient complementarity to the predetermined signal sequence.

(b) a functional RNA that is capable of effecting the cleavage of the endogenous signal RNA at the 5' or 3' end of the predetermined signal sequence; and

(c) a carrier RNA that is an RNA molecule that is capable of binding to the cleaved signal RNA portion that comprises the predetermined signal sequence at the predetermined signal sequence end, in such a way that the RNA duplex that formed is 14-31 nucleotides long and it comprises 3' or 5' overhang of 0-5 nucleotides, such that the RNA duplex is a substrate for a Dicer.

Thus, following introduction of the composition into a cell comprising the endogenous signal RNA, the functional RNA effects the cleavage of the endogenous signal RNA at the 5' or 3' end of the predetermined signal sequence and then the carrier RNA is hybridized to the cleaved signal RNA portion comprising the predetermined signal sequence at the predetermined signal sequence end and directs the processing of the predetermined signal sequence by Dicer and Risc and then the Risc-signal sequence complex directs the cleavage of the exogenous RNA of interest at a specific target/cleavage site.

The Dicer or Risc processing may involve additional proteins. In another embodiment of the invention, the carrier RNA may also be generated from a second exogenous RNA molecule. The predetermined signal sequence may be selected from, but is not limited to: a viral RNA sequence, and a sequence that is unique to neoplastic cells. The functional RNA may be selected from, but is not limited to: microRNA (miRNA), lariat-form RNA, short-hairpin RNA (shRNA), siRNA expression domain, ribozyme, or the like. In another embodiment of the invention, the composition of the invention may also comprise or encode an additional functional RNA that is capable of effecting the cleavage of the endogenous signal RNA at the opposite end of the predetermined signal sequence to that cleaved by the first functional RNA. In specific embodiments, the carrier RNA that is encoded by the composition may be driven by a polymerase I based promoter or polymerase III based promoter.

In one embodiment of the invention, the exogenous RNA of interest may further comprise:

(a) a sequence encoding an exogenous protein of interest; and

(b) an inhibitory sequence that is capable of inhibiting the expression of the exogenous protein of interest;

such that the specific target/cleavage site is located between the inhibitory sequence and the sequence encoding the exogenous protein of interest, whereby, following introduction of the

composition into a cell comprising the endogenous signal RNA, the exogenous RNA of interest may be transcribed and cleaved at the specific cleavage/target site so that the inhibitory sequence is detached from the sequence encoding the exogenous protein of interest and the exogenous protein of interest is capable of being expressed.

5 The exogenous protein of interest may be selected from, but is not limited to: the protein toxins Ricin, Abrin, Diphtheria toxin, fusion protein comprising protein toxins, and the like, or combinations thereof. A single molecule of any one of these may be sufficient to kill the cell in which any of these molecules is expressed. The inhibitory sequence can be located downstream or upstream from the specific target/cleavage site. The inhibitory
10 sequence that is located upstream from the specific target/cleavage site may be, but is not limited to a plurality of initiation codons, wherein each of the initiation codons is located within a Kozak consensus sequence, or any other translation initiation motif, wherein each of the initiation codons and the sequence encoding the protein of interest are not in the same reading frame. Thus these initiation codons will cause suppression of the expression of the
15 protein of interest prior to cleavage.

In other embodiments, the predetermined signal sequence may be located at the 5' or 3' end of the endogenous signal RNA and the composition does not necessarily encode the functional RNA.

20 According to some embodiments, the components of the composition may be encoded by the same or different polynucleotide molecules. In some embodiments, one or more components of the composition may be on the same RNA molecule.

In additional specific embodiments, the present invention provides a composition for expressing an exogenous protein of interest only in the presence of an endogenous signal RNA in a cell, the exogenous protein of interest being encoded from the composition, the
25 endogenous signal RNA being an RNA molecule which comprises a predetermined signal sequence, the predetermined signal sequence being a predetermined sequence that is at least 18 nucleotides in length and the composition comprising one or more polynucleotide molecules that comprise:

30 (a) one or more polynucleotide sequence(s) encoding a functional RNA that is capable of effecting the cleavage, directly or indirectly, of the endogenous signal RNA at a predetermined cleavage site, wherein the predetermined cleavage site is the 3' end of the predetermined signal sequence; and

(b) a polynucleotide sequence encoding an exogenous RNA of interest molecule which consists essentially of:

(1) a first sequence which is of sufficient complementarity to an edge sequence to hybridize therewith, the edge sequence being located 0-5 nucleotides upstream from the predetermined cleavage site and extending upstream in the signal RNA, wherein the first sequence comprises one or more initiation codon(s), wherein each of the initiation codons consists essentially of 5'-AUG-3'; and

(2) a second sequence upstream to the first sequence, wherein the second sequence is a predetermined sequence that is 0-5 nucleotides in length; and

(3) a third sequence downstream from the first sequence, wherein the third sequence is 0-7000 nucleotides in length; and

wherein the exogenous RNA of interest molecule comprises a sequence encoding an exogenous protein of interest at least 21 nucleotides downstream from the 5' end of said exogenous RNA of interest molecule; whereby following introduction of the composition into a cell comprising the endogenous signal RNA, the functional RNA effects the cleavage, directly or indirectly, of the endogenous signal RNA at the 3' end of the predetermined signal sequence and thereby the exogenous RNA of interest molecule is hybridized to the edge sequence at the cleaved endogenous signal RNA and may direct the predetermined signal sequence to a Dicer processing that may cleave the exogenous RNA of interest molecule, whereby each of the initiation codon(s) is detached from the sequence encoding the exogenous protein of interest and the exogenous protein of interest is capable of being expressed.

In another embodiment, the edge sequence may be 25-30 nucleotides in length and may be located 2 nucleotides upstream from the predetermined cleavage site and extends upstream in the endogenous signal RNA, wherein the second sequence is 0 nucleotides in length. In yet another embodiment of the invention, each of the initiation codon(s) may be located 0-21 nucleotides downstream from the 5' end of the exogenous RNA of interest molecule, such that each of the initiation codon(s) and the sequence encoding the exogenous protein of interest are not in the same reading frame. In further embodiment of the invention, at least one of the initiation codon(s) may be located within a Kozak consensus sequence or any other translation initiation motif/element. The functional RNA may be selected from, but is not limited to: microRNA (miRNA), short-hairpin RNA (shRNA),

small-interfering RNA (siRNA) and/or ribozyme. The exogenous protein of interest may be, for example, but is not limited to Diphtheria toxin A chain, RIP protein, and any other protein toxin.

According to some embodiments, the compositions of the invention may be used in various methods and applications, such as, for example, but not limited to: regulation of gene expression, targeted cell death, treatment of a disease or a condition including, for example, proliferative disorders (such as cancer), infectious diseases, and the like, diagnosis of a disease or a condition, formation of transgenic organisms, suicide gene therapy, and the like.

According to some embodiments, there is provided a composition comprising one or more polynucleotides for directing specific cleavage of an exogenous RNA of interest at a specific target site, the cleavage taking place only in the presence of an endogenous signal RNA in a cell, the endogenous signal RNA being an RNA molecule which comprises a signal sequence, the signal sequence being any predetermined sequence of from 18 to 25 nucleotides in length, whereby introduction of said composition into a cell comprising said endogenous signal RNA, directs the cleavage of said exogenous RNA of interest at the specific target site that is located within a specific sequence, which is of sufficient complementarity to hybridize with the predetermined signal sequence.

In some embodiments, the one or more polynucleotides may comprise a first polynucleotide sequence encoding said exogenous RNA of interest; a second polynucleotide sequence encoding a functional RNA capable of mediating the cleavage of the endogenous signal RNA at a predetermined cleavage site; and a third polynucleotide sequence encoding a carrier RNA.

In some embodiments, the carrier RNA is an RNA molecule that is at least about 18 nucleotides in length and is consisting essentially of: a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, said edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides downstream from said predetermined cleavage site and extends downstream in said endogenous signal RNA; a second sequence downstream from said first sequence, wherein said second sequence is a random sequence that is 0-5 nucleotides in length; a third sequence upstream from said first sequence, wherein said third sequence is 0-7000 nucleotides in length; and the predetermined cleavage site is the 5' end of said

predetermined signal sequence. In some embodiments, the edge sequence is 23-28 nucleotides in length and is located starting from the predetermined cleavage site to about 23-28 nucleotides downstream, wherein the second sequence is 2 nucleotides in length and wherein said third sequence is 0 nucleotides in length.

5 In some embodiments, the carrier RNA is an RNA molecule that is at least about 18 nucleotides in length and is consisting essentially of: a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, said edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides upstream from said predetermined cleavage site and extends upstream in said
10 endogenous signal RNA; a second sequence upstream from the first sequence, wherein said second sequence is a random sequence that is 0-5 nucleotides in length; a third sequence downstream from the first sequence, wherein said third sequence is 0-7000 nucleotides in length; and the predetermined cleavage site is the 3' end of said predetermined signal sequence. In some embodiments, the edge sequence is 25-30 nucleotides in length and is
15 located 2 nucleotides upstream from the predetermined cleavage site and extends upstream in said endogenous signal RNA, wherein said second sequence is 0 nucleotides in length and wherein said third sequence is 0 nucleotides in length.

In some embodiments, the carrier RNA may be processed from a polynucleotide sequence comprising a carrier sequence that is at least about 18 nucleotides in length, said
20 carrier sequence consisting essentially of: a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, said edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides downstream from said predetermined cleavage site and extends downstream in said endogenous signal RNA; a second sequence downstream from said first sequence, wherein said second
25 sequence is a random sequence that is 0-5 nucleotides in length; and a third sequence upstream from said first sequence, wherein said third sequence is 0-7000 nucleotides in length; wherein the polynucleotide sequence is cleaved within the cell at a carrier cleavage site that is a 3' end of said carrier sequence; wherein the cleavage at the carrier cleavage site is effected by a functional nucleic acid which is encoded by a fourth polynucleotide
30 sequence of the composition; and wherein the predetermined cleavage site is the 5' end of said predetermined signal sequence.

In additional embodiments, the carrier RNA may be processed from a polynucleotide sequence comprising a carrier sequence that is at least about 18 nucleotides in length, said carrier sequence consisting essentially of: a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, said edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides upstream from said predetermined cleavage site and extends upstream in said endogenous signal RNA; a second sequence upstream from the first sequence, wherein said second sequence is a random sequence that is 0-5 nucleotides in length; and a third sequence downstream from said first sequence, wherein said third sequence is 0-7000 nucleotides in length; wherein the polynucleotide sequence is cleaved within the cell at a carrier cleavage site that is 5' end of said carrier sequence; the cleavage at the carrier cleavage site is effected by a functional nucleic acid which is encoded by a fourth polynucleotide sequence of the composition; and the predetermined cleavage site is the 3' end of said predetermined signal sequence.

According to some embodiments, the endogenous signal RNA is a cellular mRNA, viral RNA, or both. In further embodiments, the predetermined signal sequence is unique to neoplastic cells, viral infected cells, or both.

According to some embodiments, sufficient complementarity is at least 30% complementarity. In further embodiments, sufficient complementarity is at least 90%.

According to some embodiments, the one or more polynucleotide may comprise one or more DNA molecules, one or more RNA molecules or combinations thereof.

In some embodiments, the functional RNA may be selected from the group consisting of: microRNA (miRNA), lariat-form RNA, short-hairpin RNA (shRNA), siRNA expression domain, antisense RNA, double-stranded RNA (dsRNA), small-interfering RNA (siRNA) and ribozyme.

According to further embodiments, the exogenous RNA of interest may further comprise a sequence encoding an exogenous protein of interest; and an inhibitory sequence that is capable of inhibiting the expression of the exogenous protein of interest; wherein the specific target site is located between the inhibitory sequence and the sequence encoding the exogenous protein of interest, whereby following introduction of said composition into a cell comprising the endogenous signal RNA, the exogenous RNA of interest is transcribed and cleaved at the specific target site, whereby the inhibitory sequence is detached from the

sequence encoding the exogenous protein of interest and the exogenous protein of interest is capable of being expressed.

In some embodiments, the exogenous protein of interest is a toxin. In some embodiments, the exogenous protein of interest is selected from the group consisting of:
5 Ricin, Ricin A chain, Abrin, Abrin A chain, Diphtheria toxin A chain and modified forms thereof. In further embodiments, the exogenous protein of interest is selected from the group consisting of: alpha toxin, saporin, maize RIP, barley RIP, wheat RIP, corn RIP, rye RIP, flax RIP, Shiga toxin, Shiga-like RIP, momordin, thymidine kinase, pokeweed antiviral protein, gelonin, *Pseudomonas* exotoxin, *Pseudomonas* exotoxin A, *Escherichia coli*
10 cytosine deaminase and modified forms thereof.

According to additional embodiments, the inhibitory sequence in the exogenous RNA of interest sequence is located upstream from the specific target site. In some embodiments, the inhibitory sequence comprise one or more initiation codons, wherein each of the initiation codons and the sequence encoding the exogenous protein of interest are not
15 in the same reading frame, and wherein said inhibitory sequence, directly or indirectly, reduces the efficiency of translation of said exogenous protein of interest. In some embodiments the one or more initiation codons is consisting essentially of 5'-AUG-3'

In further embodiments, the exogenous RNA of interest may further comprise a stop codon that is located between the initiation codon and the start codon of the sequence
20 encoding the exogenous protein of interest, wherein the stop codon and the initiation codon are in the same reading frame. The stop codon may be selected from the group consisting of: 5'-UAA-3', 5'-UAG-3' and 5'-UGA-3'.

According to additional embodiments, the inhibitory sequence may further comprise a nucleotide sequence downstream from the initiation codon, wherein said nucleotide
25 sequence and said initiation codon are in the same reading frame, and wherein the nucleotide sequence encodes a sorting signal for subcellular localization. The subcellular localization may be selected from the group consisting of: mitochondria, nucleus, endosome, lysosome, peroxisome and endoplasmic reticulum (ER).

According to further embodiments, the inhibitory sequence may further comprise a
30 nucleotide sequence downstream from the initiation codon, wherein the nucleotide sequence and the initiation codon are in the same reading frame; and wherein said nucleotide sequence encodes a protein degradation signal.

According to additional embodiments, the inhibitory sequence may further comprise a nucleotide sequence downstream from the initiation codon, wherein the nucleotide sequence and the initiation codon are in the same reading frame; wherein said nucleotide sequence and said sequence encoding the exogenous protein of interest are in the same reading frame; and wherein said nucleotide sequence encodes an amino acid sequence; whereby when the amino acid sequence is fused to the exogenous protein of interest the biological function of the exogenous protein of interest is inhibited.

According to further embodiments, the RNA of interest may further comprise a stop codon downstream from the initiation codon, wherein the stop codon and the initiation codon are in the same reading frame and wherein the exogenous RNA of interest further comprises an intron downstream from the stop codon, whereby the exogenous RNA of interest is a target for nonsense-mediated decay (NMD).

In further embodiments, the inhibitory sequence may be located downstream from the sequence encoding the exogenous protein of interest and the inhibitory sequence comprises an RNA localization signal for subcellular localization or an endogenous miRNA binding site.

In some embodiments, the inhibitory sequence may be located upstream from the sequence encoding the exogenous protein of interest, wherein the inhibitory sequence is capable of forming a secondary structure, having a folding free energy of lower than -30 kcal/mol, whereby said secondary structure is sufficient to block scanning ribosomes from reaching the start codon of said exogenous protein of interest.

In some embodiments, the exogenous RNA of interest may further comprise an internal ribosome entry site (IRES) sequence downstream from the specific cleavage site and upstream from the sequence encoding the exogenous protein of interest, wherein the IRES sequence is more functional within the cleaved exogenous RNA of interest than within the intact exogenous RNA of interest.

In some embodiments, the exogenous RNA of interest may comprise a nucleotide sequence immediately upstream from the sequence encoding the exogenous protein of interest, wherein the nucleotide sequence comprises an internal ribosome entry site (IRES) sequence, which increases the efficiency of translation of said exogenous protein of interest in the cleaved exogenous RNA of interest.

In some embodiments, the RNA of interest may further comprise a nucleotide sequence comprising a cytoplasmic polyadenylation element, located immediately downstream from said sequence encoding the exogenous protein of interest, wherein said cytoplasmic polyadenylation element increases the efficiency of translation of said exogenous protein of interest in the cleaved exogenous RNA of interest.

According to some embodiments, the composition may further comprise an additional polynucleotide sequence that encodes for an additional RNA molecule, said additional RNA molecule comprises at the 3' end a nucleotide sequence that is capable of binding to a sequence that is located upstream of said specific target site and downstream from the sequence encoding the exogenous protein of interest, wherein the additional RNA molecule, directly or indirectly, increases the efficiency of translation of said exogenous protein of interest in the cleaved exogenous RNA of interest.

According to some embodiments, the composition may further comprise an additional polynucleotide sequence that encodes a cleaving component that is capable of effecting the cleavage, of said exogenous RNA of interest at a position that is located upstream from the inhibitory sequence, wherein said cleaving component(s) is selected from the group consisting of: a) a nucleic acid sequence that is located within said exogenous RNA of interest, wherein said nucleic acid sequence is selected from the group consisting of: a) endonuclease recognition site, endogenous miRNA binding site, cis acting ribozyme and miRNA sequence, wherein said nucleic acid sequence, directly or indirectly, reduces the efficiency of translation of said exogenous protein of interest in the exogenous RNA of interest; and b) an inhibitory RNA, wherein said inhibitory RNA is selected from the group consisting of: microRNA (miRNA), lariat-form RNA, short-hairpin RNA (shRNA), siRNA expression domain, antisense RNA, double-stranded RNA (dsRNA), small-interfering RNA (siRNA) and ribozyme; wherein said inhibitory RNA, directly or indirectly, reduces the efficiency of translation of said exogenous protein of interest in said exogenous RNA of interest.

In further embodiments, the specific sequence is a plurality of specific sequences and the specific target site is a plurality of specific target sites.

In some embodiments, the exogenous RNA of interest and the functional RNA are capable of being located on the same or different polynucleotide molecules. In some

embodiments, the exogenous RNA of interest, the functional RNA and the functional nucleic acid are capable of being located on one or more polynucleotide molecules.

According to further embodiments, the one or more polynucleotides of the composition may be integrated into the cell genome.

5 In some embodiments, the cell may be selected from a group consisting of: human cell, animal cell, cultured cell and plant cell. In some embodiments, the cell may be present in an organism.

According to some embodiments, there is further provided a composition comprising one or more polynucleotides for directing specific expression of an exogenous protein of
10 interest in a cell, wherein the exogenous protein of interest is expressed only in the presence of an endogenous signal RNA in a cell, the endogenous signal RNA being an RNA molecule which comprises a signal sequence, the signal sequence being any predetermined sequence of from 18 to 25 nucleotides in length, whereby introduction of said composition into a cell comprising said endogenous signal RNA directs the cleavage of an exogenous
15 RNA of interest at a specific target site that is located within a specific sequence, which is of sufficient complementarity to hybridize with the predetermined signal sequence, wherein only after the cleavage of said exogenous RNA of interest in the cell, the exogenous protein of interest, which is encoded by said cleaved exogenous RNA of interest is capable of being expressed in the cell. In some embodiments, the one or more polynucleotides includes a
20 first polynucleotide sequence encoding said exogenous RNA of interest; a second polynucleotide sequence encoding a functional RNA capable of mediating the cleavage of the endogenous signal RNA at a predetermined cleavage site; and a third polynucleotide sequence encoding a carrier RNA.

According to some embodiments, there is provided a method for killing a specific
25 cell population, which comprises an endogenous signal RNA, the method comprises: introducing the cell with a composition comprising one or more polynucleotides for directing specific cleavage of an exogenous RNA of interest at a specific target site that is located within a specific sequence, which is of sufficient complementarity to hybridize with the endogenous signal RNA, the endogenous signal RNA being an RNA molecule which
30 comprises a signal sequence, the signal sequence being any predetermined sequence of from 18 to 25 nucleotides in length; and wherein the cleavage of the exogenous RNA of interest in the cell, enables the expression of an exogenous protein of interest, capable of killing the

cell population. The one or more polynucleotides comprises: a first polynucleotide sequence encoding said exogenous RNA of interest; a second polynucleotide sequence encoding a functional RNA capable of mediating the cleavage of the endogenous signal RNA at a predetermined cleavage site; and a third polynucleotide sequence encoding a carrier RNA. In some embodiments, the endogenous signal RNA is a cellular mRNA, viral RNA, or both. The cell population may be selected from a group consisting of: human cell, animal cell, cultured cell and plant cell. In some embodiments, the cell population is a neoplastic cell population. In some embodiments, the cell population is present in an organism.

Objects and advantages of the present invention will be clear from the description that follows.

BRIEF DESCRIPTION OF THE FIGURES

The following figures are offered by way of illustration and not by way of limitation.

FIG. 1 is a general scheme of a model for biogenesis and activity of microRNAs (miRNAs) in a cell.

FIG. 2 is a schematic drawing showing an example for cleaving exogenous RNA of interest in response to the presence of an endogenous signal RNA in a cell, according to some embodiments. In this exemplary embodiment, the composition encodes for: a carrier RNA of 27 nucleotides; an exogenous RNA of interest that comprises a specific sequence which is complementary to a predetermined signal sequence of the endogenous signal RNA; and a functional RNA which is shRNA that is capable of effecting the cleavage of the endogenous signal RNA at the 5' end of the predetermined signal sequence.

FIG. 3 is a schematic drawing showing an example for cleaving exogenous RNA of interest in response to the presence of an endogenous signal RNA in a cell, according to some embodiments. In this example, the composition of the invention encodes for: a carrier RNA of 27 nucleotides, an exogenous RNA of interest that comprises a specific sequence which is complementary to the predetermined signal sequence of the endogenous signal RNA; and a functional RNA which is shRNA that is capable of effecting the cleavage of the endogenous signal RNA at the 3' end of the predetermined signal sequence.

FIG. 4 is a schematic drawing showing an example for cleaving exogenous RNA of interest in response to the presence of an endogenous signal RNA in a cell, according to some embodiments. In this example, the composition of the invention encodes for: an exogenous RNA of interest that comprises a specific sequence which is complementary to the predetermined signal sequence of the endogenous signal RNA, a functional RNA which is shRNA that is capable of effecting the cleavage of the endogenous signal RNA at the 5' end of the predetermined signal sequence, a carrier sequence that is of 27 nucleotides long and a functional nucleic acid which is cis acting ribozyme that is capable of effecting the cleavage of the carrier RNA at the 3' end of the carrier sequence.

FIG. 5 is a schematic drawing showing an example for cleaving exogenous RNA of interest in response to the presence of an endogenous signal RNA in a cell, according to some embodiments. In this example, the composition of the invention encodes for: an exogenous RNA of interest that comprises a specific sequence which is complementary to the predetermined signal sequence of the endogenous signal RNA, a functional RNA which is shRNA that is capable of effecting the cleavage of the endogenous signal RNA at the 3' end of the predetermined signal sequence, a carrier sequence that is of 27 nucleotides long and a functional nucleic acid which is cis acting ribozyme that is capable of effecting the cleavage of the carrier RNA at the 5' end of the carrier sequence.

FIG. 6A is a schematic drawing showing an example for inhibitory RNA that is capable of effecting the cleavage of the endogenous signal RNA at the 5' end of the predetermined signal sequence, according to some embodiments.

FIG. 6B is a schematic drawing showing an example for inhibitory RNA that is capable of effecting the cleavage of the endogenous signal RNA at the 3' end of the predetermined signal sequence, according to some embodiments.

FIG. 7A is a schematic drawing showing an example for inhibitory RNA that is capable of effecting the cleavage of the carrier RNA at the 3' end of the carrier sequence, according to some embodiments.

FIG. 7B is a schematic drawing showing an example for inhibitory RNA that is capable of effecting the cleavage of the carrier RNA at the 5' end of the carrier sequence, according to some embodiments.

FIG. 8A is a schematic drawing showing an example for inhibitory RNA which, according to some embodiments, is an RNA duplex that may be a substrate for Dicer.

FIG. 8B is a schematic drawing showing an example for inhibitory RNA which, according to some embodiments, is an RNA duplex that may be a substrate for Dicer.

- 5 FIG. 9A is a schematic drawing showing an example, according to some embodiments, for hammerhead-type ribozyme (SEQ ID NO. 89) that is capable of effecting the cleavage of the endogenous signal RNA or the carrier RNA at the predetermined cleavage site or at the carrier cleavage site of, respectively.

- 10 FIG. 9B is a schematic drawing showing an example, according to some embodiments, for hairpin-type ribozyme that is capable of effecting the cleavage of the endogenous signal RNA or the carrier RNA at the predetermined cleavage site or at the carrier cleavage site, respectively. The exemplary hairpin-type ribozyme is composed of SEQ ID NO. 90, preceded by a sequence complementary to the target RNA, the tetra-nucleotide AAGA (SEQ ID NO. 114) and an additional sequence complementary to the target RNA (at the 5' end of the ribozyme).

FIG. 10 is a schematic drawing showing an example, according to some embodiments, of a functional nucleic acid that is the very efficient cis-acting hammerhead ribozyme - snorbozyme (SEQ ID NO. 91) [22], which is capable of effecting the cleavage of the carrier RNA at the 3' end of the carrier sequence.

- 20 FIG. 11 is a schematic drawing showing an example, according to some embodiments, of a functional nucleic acid that is the very efficient cis-acting hammerhead ribozyme - N117 (SEQ ID NO. 92) [23] which is capable of effecting the cleavage of the carrier RNA (SEQ ID NO. 93) at the 5' end of the carrier sequence.

- 25 FIG. 12A is a schematic drawing showing an example, according to some embodiments, of a functional nucleic acid that is an endonuclease recognition site or an endogenous miRNA binding site, such that the functional nucleic acid is capable of effecting the cleavage of the carrier RNA at the 3' end of the carrier sequence.

- 30 FIG. 12B is a schematic drawing showing an example, according to some embodiments, of a functional nucleic acid that is an endonuclease recognition site or an endogenous miRNA binding site, such that the functional nucleic acid is capable of effecting the cleavage of the carrier RNA at the 5' end of the carrier sequence.

FIG. 12C is a schematic drawing showing an example, according to some embodiments, of a functional nucleic acid that is a miRNA sequence, such that the miRNA sequence is capable of effecting the cleavage of the carrier RNA at the 3' end of the carrier sequence.

FIG. 12D is a schematic drawing showing an example, according to some embodiments, of a functional nucleic acid that is a miRNA sequence, such that the miRNA sequence is capable of affecting the cleavage of the carrier RNA at the 5' end of the carrier sequence.

FIG. 13A is a schematic drawing showing an example, according to some embodiments, of a functional nucleic acid which is capable of forming stem loop structure with the carrier sequence, such that the stem loop structure is capable of effecting the cleavage of the carrier RNA at the 3' end of the carrier sequence.

FIG. 13B is a schematic drawing showing an example, according to some embodiments, for functional nucleic acid which is capable of forming stem loop structure with the carrier sequence, such that the stem loop structure is capable of effecting the cleavage of the carrier RNA at the 5' end of the carrier sequence.

FIG. 14A is a schematic drawing showing an example, according to some embodiments, for functional nucleic acid that has a stem loop structure, such that the loop comprises the carrier sequence and such that when the stem loop structure is processed by Drosha and Dicer, the carrier sequence is detached from the stem loop structure and the siRNA duplex thus formed is the functional RNA, which is then capable of effecting the cleavage of the endogenous signal RNA at the predetermined cleavage site.

FIG. 14B is a schematic drawing showing an example, according to some embodiments, of a functional nucleic acid that has a stem loop structure, such that the loop comprises the carrier sequence and such that the expression of the stem loop structure is driven by polymerase I or III based promoter and such that when the stem loop structure is processed by Dicer the carrier sequence is detached from the stem loop structure and the siRNA duplex thus formed is the functional RNA which is capable of effecting the cleavage of the endogenous signal RNA at the predetermined cleavage site.

FIG. 15A is a schematic drawing showing an example, according to some embodiments, of a carrier sequence that is located in the same RNA duplex with the functional RNA, such that the double strand region is located downstream from the carrier sequence and such that when the double strand region is processed by Dicer, the carrier sequence is detached from

the RNA duplex and the siRNA duplex thus formed is the functional RNA and is capable of effecting the cleavage of the endogenous signal RNA at the predetermined cleavage site.

FIG. 15B is a schematic drawing showing an example, according to some embodiments, for a carrier sequence that is located in the same RNA duplex with the functional RNA, such that the double strand region is located upstream from the carrier sequence and such that when the double strand region is processed by Dicer the carrier sequence is detached from the RNA duplex and the siRNA duplex thus formed is the functional RNA which is capable of effecting the cleavage of the endogenous signal RNA at the predetermined cleavage site.

FIG. 16A is a schematic drawing showing an example, according to some embodiments, for a carrier RNA that is located in the same RNA duplex with the functional RNA, such that the double strand region is located at the 5' end of the carrier RNA and such that when the double strand region is processed by Dicer, the sequence that is located at the 3' end of the carrier RNA is detached from the RNA duplex and the siRNA duplex thus formed is the functional RNA, which is capable of effecting the cleavage of the endogenous signal RNA at the predetermined cleavage site.

FIG. 16B is a schematic drawing showing an example, according to some embodiments, of a carrier RNA that is located in the same RNA duplex with the functional RNA, such that the double strand region is located at the 3' end of the carrier RNA and such that when the double strand region is processed by Dicer, the sequence that is located at the 5' end of the carrier RNA is detached from the RNA duplex and the siRNA duplex thus formed is the functional RNA, which is capable of effecting the cleavage of the endogenous signal RNA at the predetermined cleavage site.

FIG. 17A is a schematic drawing showing an example, according to some embodiments, for the carrier sequence that is located in the same RNA duplex with the functional RNA, such that the double strand region is located upstream from the carrier sequence and such that when the double strand region is processed by Dicer, the siRNA duplex that is formed is the functional RNA, which is capable of effecting the cleavage of the endogenous signal RNA at the predetermined cleavage site.

FIG. 17B is a schematic drawing showing an example, according to some embodiments, of a carrier sequence that is located in the same RNA duplex with the functional RNA, such that the double strand region is located downstream from the carrier sequence and such that when the double strand region is processed by Dicer, the siRNA duplex that is formed is the

functional RNA, which is capable of effecting the cleavage of the endogenous signal RNA at the predetermined cleavage site.

FIG. 18A is a schematic drawing showing an example, according to some embodiments, of a carrier sequence that is located in the same RNA duplex with the functional nucleic acid, such that the double strand region is located upstream from the carrier sequence and such that when the double strand region is processed by Dicer, the siRNA duplex that is formed is the functional nucleic acid which is capable of effecting the cleavage of the carrier RNA at the carrier cleavage site.

FIG. 18B is a schematic drawing showing an example, according to some embodiments, of a carrier sequence that is located in the same RNA duplex with the functional nucleic acid, such that the double strand region is located downstream from the carrier sequence and such that when the double strand region is processed by Dicer, the siRNA duplex that is formed is the functional nucleic acid, which and is capable of effecting the cleavage of the carrier RNA at the carrier cleavage site.

FIG. 19A is a schematic drawing illustrating an example, according to some embodiments, of a carrier sequence that is located in the same RNA duplex with the functional nucleic acid and with the functional RNA, such that the double strand region is located upstream from the carrier sequence, and such that when the double strand region is processed by Dicer, the siRNA duplexes that are formed are the functional nucleic acid and the functional RNA.

FIG. 19B is a schematic drawing illustrating an example, according to some embodiments, of a carrier sequence that is located in the same RNA duplex with the functional nucleic acid and with the functional RNA, such that the double strand region is located downstream from the carrier sequence and such that when the double strand region is processed by Dicer the siRNA duplexes that are formed are the functional nucleic acid and the functional RNA.

FIG. 20A is a schematic drawing showing an example, according to some embodiments, of a carrier RNA that comprises 3 contiguous carrier sequences downstream from the carrier sequence, such that the functional nucleic acid is capable of effecting the cleavage of the carrier RNA at the 3' end of the carrier sequence.

FIG. 20B is a schematic drawing showing an example, according to some embodiments, for carrier RNA that comprises 3 contiguous carrier sequences upstream from the carrier

sequence, such that the functional nucleic acid is capable of effecting the cleavage of the carrier RNA at the 5' end of the carrier sequence.

FIG. 21A is a schematic drawing showing an example, according to some embodiments, for polynucleotide molecule(s) of the composition that, in addition to the functional RNA that
5 cleaves the 5' end of the predetermined signal sequence, further transcribes an additional functional RNA that cleaves the 3' end of the predetermined signal sequence.

FIG. 21B is a schematic drawing showing an example, according to some embodiments, for polynucleotide molecule(s) of the composition that, in addition to the functional RNA that
10 cleaves the 3' end of the predetermined signal sequence, further transcribes an additional functional RNA that cleaves the 5' end of the predetermined signal sequence.

FIG. 22A is a schematic drawing showing an example, according to some embodiments, of the schematic structure of an exogenous RNA of interest that is activated by its cleavage, such that the specific sequence is located downstream from the inhibitory sequence and upstream from the sequence encoding the exogenous protein of interest.

15 FIG. 22B is a schematic drawing showing an example, according to some embodiments, of the schematic structure of an exogenous RNA of interest that is activated by its cleavage, such that the specific sequence is located upstream from the inhibitory sequence and downstream from the sequence encoding the exogenous protein of interest.

FIG. 23A is a schematic drawing showing an example, according to some embodiments, for
20 inhibitory sequence that is located upstream from the specific target/cleavage site of the exogenous RNA of interest, and comprises an AUG that is not in the same reading frame with the sequence encoding exogenous protein of interest.

FIG. 23B is a schematic drawing showing an example, according to some embodiments, for
25 inhibitory sequence that is located upstream from the specific target/cleavage site of the exogenous RNA of interest, and comprises a Kozak consensus sequence (5'-ACCAUGG-3' – SEQ ID NO. 25) that is not in the same reading frame with the sequence encoding exogenous protein of interest.

FIG. 23C is a schematic drawing showing an example, according to some embodiments, for
30 inhibitory sequence that is located upstream from the specific target/cleavage site of the exogenous RNA of interest, and comprises 2 Kozak consensus sequences that are not in the same reading frame with the sequence encoding exogenous protein of interest.

FIG. 24A is a schematic drawing showing an example, according to some embodiments, for inhibitory sequence that is located upstream from the specific target/cleavage site of the exogenous RNA of interest, and comprises an AUG and a downstream stop codon that are in the same reading frame.

- 5 FIG. 24B is a schematic drawing showing an example, according to some embodiments, for inhibitory sequence that is located upstream from the specific target/cleavage site of the exogenous RNA of interest, and comprises an AUG and a downstream: sorting signal for subcellular localization or protein degradation signal.

- 10 FIG. 24C is a schematic drawing showing an example, according to some embodiments, for inhibitory sequence that is located upstream from the specific target/cleavage site of the exogenous RNA of interest, and comprises an AUG and a downstream sequence encoding amino acids that are capable of inhibiting the biological function of the downstream protein of interest.

- 15 FIG. 24D is a schematic drawing showing an example, according to some embodiments, for inhibitory sequence that is located upstream from the specific target/cleavage site of the exogenous RNA of interest, and comprises an AUG, a downstream stop codon that is in the same reading frame with the AUG and a downstream intron, such that the exogenous RNA of interest is a target for nonsense-mediated decay (NMD).

- 20 FIG. 25A is a schematic drawing showing an example, according to some embodiments, for inhibitory sequence that is located upstream from the specific target/cleavage site of the exogenous RNA of interest and comprises a binding site for translation repressor.

- 25 FIG. 25B is a schematic drawing showing an example, according to some embodiments, for inhibitory sequence that is located upstream from the specific target/cleavage site of the exogenous RNA of interest and comprises an RNA localization signal for subcellular localization.

FIG. 25C is a schematic drawing showing an example, according to some embodiments, for inhibitory sequence that is located upstream from the specific target/cleavage site of the exogenous RNA of interest and comprises an RNA destabilizing element that is an AU-rich element or an endonuclease recognition site.

FIG. 25D is a schematic drawing showing an example, according to some embodiments, for inhibitory sequence that is located upstream from the specific target/cleavage site of the exogenous RNA of interest, and comprises a secondary structure.

FIG. 26 is a schematic drawing showing an example, according to some embodiments, for the activation of the exogenous RNA of interest by its cleavage, such that the inhibitory sequence creates a secondary structure that blocks translation and such that the cleavage creates an IRES (Internal ribosome entry site).

FIG. 27A is a schematic drawing showing an example, according to some embodiments, of additional structure that may increase the efficiency of translation of the exogenous RNA of interest, that is cleaved at the 5' end, wherein the additional structure is an IRES (Internal ribosome entry site).

FIG. 27B is a schematic drawing showing an example, according to some embodiments, of additional structure that may increase the efficiency of translation of the exogenous RNA of interest that is cleaved at the 5' end, wherein the additional structure is a stem loop structure.

FIG. 27C is a schematic drawing showing an example, according to some embodiments, of additional structure that may increase the efficiency of translation of the exogenous RNA of interest, that is cleaved at the 5' end, wherein the additional structure is a cytoplasmic polyadenylation element.

FIG. 27D is a schematic drawing showing an example, according to some embodiments, of additional structures that may increase the efficiency of translation of the exogenous RNA of interest, which is cleaved at the 5' end, wherein the additional structures are nucleotide sequences that are capable of binding to each other and by this force the exogenous RNA of interest to form a circular structure, particularly when the exogenous RNA of interest is cleaved at the specific target/cleavage site.

FIG. 28A is a schematic drawing showing an example, according to some embodiments, of additional structure that may increase the efficiency of translation of the exogenous RNA of interest, that is cleaved at the 5' end, such that the additional structure is a polypeptide that is encoded from the composition of the invention, wherein the polypeptide is capable of binding to the poly-A tail of the exogenous RNA of interest, and to a sequence within the exogenous RNA of interest of the invention and by this force the exogenous RNA of interest

to form a circular structure, particularly when the exogenous RNA of interest is cleaved at the specific target/cleavage site.

FIG. 28B is a schematic drawing illustrating an example, according to some embodiments, of additional structure that may reduce the efficiency of translation of the intact exogenous RNA of interest, such that the additional structure is a cis acting ribozyme that removes the CAP structure from the intact exogenous RNA of interest.

FIG. 29A is a schematic drawing illustrating an example, according to some embodiments, of inhibitory sequence that is located downstream from the specific target/cleavage site and comprises an intron, such that the exogenous RNA of interest is a target for nonsense-mediated decay (NMD).

FIG. 29B is a schematic drawing illustrating an example, according to some embodiments, of inhibitory sequence that is located downstream from the specific target/cleavage site and comprises a binding site for translation repressor.

FIG. 29C is a schematic drawing illustrating an example, according to some embodiments, of inhibitory sequence that is located downstream from the specific target/cleavage site and comprises an RNA localization signal for subcellular localization.

FIG. 29D is a schematic drawing illustrating an example, according to some embodiments, of an inhibitory sequence that is located downstream from the specific target/cleavage site and comprises an RNA destabilizing element that is an AU-rich element or an endonuclease recognition site.

FIG. 29E is a schematic drawing illustrating an example, according to some embodiments, of an inhibitory sequence that is located downstream from the specific target/cleavage site and comprises a secondary structure.

FIG. 30A is a schematic drawing illustrating an example, according to some embodiments, for inhibitory sequence that is located downstream from the sequence encoding exogenous protein of interest, such that the inhibitory sequence creates a secondary structure that may block translation.

FIG. 30B is a schematic drawing illustrating an example, according to some embodiments, of additional structure that may increase the efficiency of translation of the exogenous RNA of interest that is cleaved at the 3' end, such that the additional structure is IRES (Internal ribosome entry site).

FIG. 30C is a schematic drawing illustrating an example, according to some embodiments, of additional structure that may increase the efficiency of translation of the exogenous RNA of interest, that is cleaved at the 3' end, such that the additional structure is a stem loop structure.

- 5 FIG. 30D is a schematic drawing illustrating an example, according to some embodiments, of additional structure that may increase the efficiency of translation of the exogenous RNA of interest that is cleaved at the 3' end, such that the additional structure is a cytoplasmic polyadenylation element.

- 10 FIG. 31A is a schematic drawing illustrating an example, according to some embodiments, of additional structures that may increase the efficiency of translation of the exogenous RNA of interest, that is cleaved at the 3' end, such that the additional structures are nucleotide sequences that are capable of binding to each other and consequently force the exogenous RNA of interest to form a circular structure, particularly when the exogenous RNA of interest is cleaved at the specific target/cleavage site.

- 15 FIG. 31B is a schematic drawing illustrating an example, according to some embodiments, of additional structure that may increase the efficiency of translation of the exogenous RNA of interest, that is cleaved at the 3' end, such that the additional structure is a polypeptide that is encoded from the composition, wherein the polypeptide is capable of binding to the CAP and to a sequence within the exogenous RNA of interest and consequently force the
20 exogenous RNA of interest to form a circular structure, particularly when the exogenous RNA of interest is cleaved at the specific target/cleavage site.

- FIG. 31C is a schematic drawing illustrating an example, according to some embodiments, of additional structure that may increase the efficiency of translation of the exogenous RNA of interest that is cleaved at the 3' end, such that the additional structure is an additional
25 RNA molecule that is encoded from the composition and is capable of binding to the exogenous RNA of interest and consequently provide it with a poly-A tail, particularly when the exogenous RNA of interest is cleaved at the specific target/cleavage site.

- FIG. 31D is a schematic drawing illustrating an example, according to some embodiments, of additional structure that may reduce the efficiency of translation of the intact exogenous
30 RNA of interest, such that the additional structure is a cis acting ribozyme that removes the poly-A from the intact exogenous RNA of interest.

FIG. 32A is a schematic drawing illustrating an example, according to some embodiments, of the structure of the exogenous RNA of interest, which comprises 2 specific sequences, such that the inhibitory sequence is located upstream from the specific target/cleavage sites.

FIG. 32B is a schematic drawing illustrating an example, according to some embodiments, of the structure of the exogenous RNA of interest that comprises 2 specific sequences, such that the inhibitory sequence is located downstream from the specific target/cleavage sites.

FIG. 32C is a schematic drawing illustrating an example, according to some embodiments, of an exogenous RNA of interest, which comprises a sequence encoding an exogenous protein of interest, between 2 sequences that are complementary to the predetermined signal sequence and 2 inhibitory sequences, one at the 5' end and other at the 3' end of the exogenous RNA of interest.

FIG. 33 is a schematic drawing illustrating an example, according to some embodiments, for expressing exogenous protein of interest in response to the presence of an endogenous signal RNA in a cell. The composition includes polynucleotide molecule/s that encode for an exogenous RNA of interest molecule that comprises a first sequence of 27 nucleotides at the 5' end that is 100% complementary to the predetermined signal sequence and to a sequence that is upstream from the predetermined signal sequence, the first sequence also comprises an 5'-AUG-3' sequence that is not in the same reading frame with the downstream sequence encoding the exogenous protein of interest and the composition further encodes a functional RNA which is shRNA that is capable of effecting the cleavage of the endogenous signal RNA at the 3' end of the predetermined signal sequence.

FIG. 34 is a schematic drawing illustrating an example for expressing an exogenous protein of interest in response to the presence of an endogenous signal RNA in a cell, according to some embodiments. The composition includes polynucleotide molecule/s that encode for an exogenous RNA of interest molecule that comprises at the 5' end a miRNA that is capable of effecting the cleavage of the endogenous signal RNA at the 3' end of the predetermined signal sequence and a first sequence of 27 nucleotides that is complementary to the predetermined signal sequence and to the sequence that is upstream from the predetermined signal sequence, the first sequence also comprises an 5'-AUG-3' sequence that is not in the same reading frame with the downstream sequence encoding the exogenous protein of interest and such that the 5'-AUG-3' is located within a Kozak consensus sequence.

FIG. 35 is a schematic drawing illustrating an example, according to some embodiments, for expressing an exogenous protein of interest in response to the presence of an endogenous signal RNA in a cell. The composition encodes for an exogenous RNA of interest molecule that comprises at the 5' end a first strand of siRNA, such that composition further transcribes the second strand of the siRNA by polymerase I or III based promoter. The first sequence of the exogenous RNA of interest molecule is 27 nucleotides in length and is complementary to the predetermined signal sequence and to the sequence that is upstream from the predetermined signal sequence, the first sequence also comprises an 5'-AUG-3' sequence that is not in the same reading frame with the downstream sequence encoding the exogenous protein of interest, such that the 5'-AUG-3' is located within a Kozak consensus sequence.

FIG. 36A is a schematic drawing illustrating an example, according to some embodiments, of an exogenous RNA of interest that comprises a cis acting ribozyme at the 5' end, which removes the CAP structure from the exogenous RNA of interest. This removal reduces the efficiency of translation of the exogenous protein of interest in the intact exogenous RNA of interest molecule.

FIG. 36B is a schematic drawing illustrating an exemplary exogenous RNA of interest that comprises two nucleotide sequences that are capable of binding to each other and by this force the exogenous RNA of interest to form a circular structure that increases the efficiency of translation of the protein of interest particularly in the cleaved RNA of interest.

FIG. 37A is a schematic drawing illustrating an example, according to some embodiments, for cleaving exogenous RNA of interest in the presence of an endogenous signal RNA in a cell. The composition encodes for: a carrier RNA of 27 nucleotides and an exogenous RNA of interest that comprises a specific sequence which is complementary to the predetermined signal sequence.

FIG. 37B is a schematic drawing illustrating an example, according to some embodiments, for cleaving exogenous RNA of interest in the presence of an endogenous signal RNA in a cell. The composition encodes for: an exogenous RNA of interest that comprises a specific sequence which is complementary to the predetermined signal sequence; a carrier sequence that is of 27 nucleotides long and a functional nucleic acid which is a cis acting ribozyme that is capable of effecting the cleavage of the carrier RNA sequence at the 3' end of the carrier sequence.

FIG. 38A is a schematic drawing illustrating an example, according to some embodiments, for cleaving exogenous RNA of interest in the presence of an endogenous signal RNA in a cell. The composition of the invention encodes for a carrier RNA of 27 nucleotides and an exogenous RNA of interest that comprises a specific sequence which is complementary to the predetermined signal sequence.

FIG. 38B is a schematic drawing illustrating an example, according to some embodiments, for cleaving exogenous RNA of interest in the presence of an endogenous signal RNA in a cell. The composition encodes for: an exogenous RNA of interest that includes a specific sequence which is 100% complementary to the predetermined signal sequence, a carrier sequence that is of 27 nucleotides long and a functional nucleic acid which is a cis acting ribozyme that is capable of effecting the cleavage of the carrier RNA sequence at the 5' end of the carrier sequence.

FIG. 39A is a schematic drawing illustrating an example, according to some embodiments, of an exogenous RNA of interest having its inhibitory sequence located downstream from the specific target/cleavage site and the inhibitory sequence is capable of inhibiting the function of an RNA localization signal for subcellular localization.

FIG. 39B is a schematic drawing illustrating an example, according to some embodiments, of an exogenous RNA of interest having its inhibitory sequence located upstream from the specific target/cleavage site and it's the inhibitory sequence is capable of inhibiting the function of an RNA localization signal for subcellular localization.

FIG. 39C is a schematic drawing illustrating an example, according to some embodiments, of an exogenous RNA of interest, having its inhibitory sequence located upstream from the specific target/cleavage site, and comprises an AUG and a downstream sequence that encodes for amino acids that are capable of inhibiting the function of the sorting signal for subcellular localization of the exogenous protein of interest, encoded from the exogenous protein of interest.

FIG. 39D is a schematic drawing illustrating an example, according to some embodiments, of inhibitory sequence that is located downstream from the specific sequence, such that the exogenous RNA of interest does not comprise a stop codon downstream from the start codon of the sequence encoding the exogenous protein of interest, and such that the inhibitory sequence encodes an amino acid sequence that is capable of inhibiting the

cleavage of a peptide sequence that is encoded upstream, wherein the peptide sequence is capable of being cleaved by a protease in a mammalian cell.

FIG. 40 is a schematic drawing illustrating an example for using the composition of the invention to kill cancer cells of a specific patient, according to some embodiments.

- 5 FIG. 41 is a schematic drawing illustrating an example for using the composition of the invention to kill cancer cells of Burkitt's lymphomas, Hodgkin's lymphomas, gastric carcinoma and nasopharyngeal carcinoma, which are latently infected with EBV by using the LMP1 mRNA as the endogenous signal RNA, , according to some embodiments.

- 10 FIG. 42 is a schematic drawing illustrating an example for using the composition of the invention to kill HIV-1 infected cells, according to some embodiments.

FIG. 43 is a schematic drawing illustrating an example for using the composition of the invention to kill HSV-1 infected cells, according to some embodiments.

FIG. 44 is a schematic drawing illustrating an example for using the composition of the invention to kill cancer cells of a specific patient, according to some embodiments.

15 DETAILED DESCRIPTION OF THE INVENTION

- In the following detailed description of the invention when a reference term, such as: said, the, the last, the previous and the former; is used it refers to the exact term that is mentioned above (e.g. wherein said "The nucleic acid sequence" it refers to the nucleic acid sequence that is mentioned above and does not refer to the nucleotide sequence that is mentioned above). Furthermore, in the following detailed description of the invention each embodiment that refers to other embodiments is defined with them as a separate unit.

The following are terms which are used throughout the description and which should be understood in accordance with the various embodiments to mean as follows:

- 25 As referred to herein, the terms "polynucleotide molecules", "oligonucleotide", "polynucleotide", "nucleic acid" and "nucleotide" sequences may interchangeably be used herein. The terms are directed to polymers of deoxyribonucleotides (DNA), ribonucleotides (RNA), and modified forms thereof in the form of a separate fragment or as a component of a larger construct, linear or branched, single stranded, double stranded, triple stranded, or hybrids thereof. The term also encompasses RNA/DNA hybrids. The polynucleotides may
30 comprise sense and antisense oligonucleotide or polynucleotide sequences of DNA or RNA.

The DNA or RNA molecules may be, for example, but are not limited to: complementary DNA (cDNA), genomic DNA, synthesized DNA, recombinant DNA, or a hybrid thereof or an RNA molecule such as, for example, mRNA, shRNA, siRNA, miRNA, and the like. Accordingly, as used herein, the terms "polynucleotide molecules", "oligonucleotide",
5 "polynucleotide", "nucleic acid" and "nucleotide" sequences are meant to refer to both DNA and RNA molecules. The terms further include oligonucleotides composed of naturally occurring bases, sugars, and covalent internucleoside linkages, as well as oligonucleotides having non-naturally occurring portions, which function similarly to respective naturally occurring portions.

10 The terms "polypeptide," "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical analogue of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers.

As referred to herein, the term "complementarity" is directed to base pairing between
15 strands of nucleic acids. As known in the art, each strand of a nucleic acid may be complementary to another strand in that the base pairs between the strands are non-covalently connected via two or three hydrogen bonds. Two nucleotides on opposite complementary nucleic acid strands that are connected by hydrogen bonds are called a base pair. According to the Watson-Crick DNA base pairing, adenine (A) forms a base pair with
20 thymine (T) and guanine (G) with cytosine (C). In RNA, thymine is replaced by uracil (U). The degree of complementarity between two strands of nucleic acid may vary, according to the number (or percentage) of nucleotides that form base pairs between the strands. For example, "100% complementarity" indicates that all the nucleotides in each strand form base pairs with the complement strand. For example, "95% complementarity" indicates that
25 95% of the nucleotides in each strand form base pair with the complement strand. The term sufficient complementarity may include any percentage of complementarity from about 30% to about 100%.

The term "construct", as used herein refers to an artificially assembled or isolated nucleic acid molecule which may be one or more nucleic acid sequences, wherein the
30 nucleic acid sequences may comprise coding sequences (that is, sequence which encodes an end product), regulatory sequences, non-coding sequences, or any combination thereof. The

term construct encompasses, for example, vector but should not be seen as being limited thereto.

"Expression vector" refers to vectors that have the ability to incorporate and express heterologous nucleic acid fragments (such as, for example, DNA), in a foreign cell. In other words, an expression vector comprises nucleic acid sequences/fragments (such as DNA, mRNA, tRNA, rRNA), capable of being transcribed. Many prokaryotic and eukaryotic expression vectors are known and/or commercially available. Selection of appropriate expression vectors is within the knowledge of those having skill in the art.

The terms "Upstream" and "Downstream", as used herein refers to a relative position in a nucleotide sequence, such as, for example, a DNA sequence or an RNA sequence. As well known, a nucleotide sequence has a 5' end and a 3' end, so called for the carbons on the sugar (deoxyribose or ribose) ring of the nucleotide backbone. Hence, relative to the position on the nucleotide sequence, the term downstream relates to the region towards the 3' end of the sequence. The term upstream relates to the region towards the 5' end of the strand.

The terms "promoter element", "promoter" or "promoter sequence" as used herein, refer to a nucleotide sequence that is generally located at the 5' end (that is, precedes, located upstream) of the coding sequence and functions as a switch, activating the expression of a coding sequence. If the coding sequence is activated, it is said to be transcribed. Transcription generally involves the synthesis of an RNA molecule (such as, for example, a mRNA) from a coding sequence. The promoter, therefore, serves as a transcriptional regulatory element and also provides a site for initiation of transcription of the coding sequence into mRNA. Promoters may be derived in their entirety from a native source, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic nucleotide segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental conditions, or at various expression levels. Promoters which cause a gene to be expressed in most cell types at most times are commonly referred to as "constitutive promoters". Promoters that control gene expression in a specific tissue are called "tissue specific promoters".

As referred to herein, the terms "RNA of interest", "exogenous RNA of interest", and "ROI" may interchangeably be used. The terms refer to a nucleotide sequence which is introduced into a target cell and may encode for an RNA molecule within the target cell.

As referred to herein, the terms "protein of interest", "exogenous protein of interest", and "POI" may interchangeably be used. The terms refer to a peptide sequence which is translated from the exogenous RNA of interest. In some embodiments, the peptide sequence can be one or more separate proteins or a fusion protein.

As referred to herein, the terms "signal RNA" and "endogenous signal RNA" may interchangeably be used. The terms refer to an intracellular RNA molecule/sequence which comprises a predetermined signal sequence. The endogenous signal RNA molecule may be encoded by the genome of the cell, and/or from a foreign genome residing within the cell, such as, for example, from a virus residing within the cell. In some embodiments, the endogenous signal RNA is a mature mRNA molecule. In some embodiments, the endogenous signal RNA is a viral RNA. The signal RNA is present within the target cell prior to introduction of an exogenous RNA of interest into the cell.

As referred to herein, the terms "predetermined signal sequence" and "signal sequence" may interchangeably be used.

As referred to herein, the terms "predetermined cleavage site" and "an additional cleavage site", refer to a cleavage site within the sequence of the endogenous signal RNA.

As referred to herein, the terms "specific target site", "specific cleavage site" and "specific target/cleavage sites" may interchangeably be used. The terms relate to one or more cleavage sites within the sequence of the exogenous RNA of interest.

The term "expression", as used herein, refers to the production of a desired end-product molecule in a target cell. The end-product molecule may be, for example an RNA molecule (such as, for example, a mRNA molecule, siRNA molecule, and the like); a peptide or a protein; and the like; or combinations thereof.

As referred to herein, the term, "Open Reading Frame" ("ORF") is directed to a coding region which contains a start codon and a stop codon.

As referred to herein, the term "Kozak sequence" is well known in the art and is directed to a sequence on an mRNA molecule that is recognized by the ribosome as the translational start site. The terms "Kozak consensus sequence", "Kozak consensus" or

"Kozak sequence", is a sequence which occurs on eukaryotic mRNA and has the consensus (gcc)gccRccAUGG (SEQ ID NO. 24), where R is a purine (adenine or guanine), three bases upstream of the start codon (AUG), which is followed by another 'G'. In some embodiments, the Kozak sequence has the sequence RNNAUGG, wherein N is any nucleotide of A, G, C or U (SEQ ID NO. 112).

As used herein, the terms "introducing" and "transfection" may interchangeably be used and refer to the transfer of molecules, such as, for example, nucleic acids, polynucleotide molecules, vectors, and the like into a target cell(s), and more specifically into the interior of a membrane-enclosed space of a target cell(s). The molecules can be "introduced" into the target cell(s) by any means known to those of skill in the art, for example as taught by Sambrook et al. Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York (2001), the contents of which are incorporated by reference herein. Means of "introducing" molecules into a cell include, for example, but are not limited to: heat shock, calcium phosphate transfection, PEI transfection, electroporation, lipofection, transfection reagent(s), viral-mediated transfer, and the like, or combinations thereof. The transfection of the cell may be performed on any type of cell, of any origin, such as, for example, human cells, animal cells, plant cells, and the like. The cells may be, for example, but not limited to: isolated cells, tissue cultured cells, cell lines, cells present within an organism body, and the like.

The term "Kill" with respect to a cell/cell population is directed to include any type of manipulation that will lead to the death of that cell/cell population.

As referred to herein, the term "Treating a disease" or "treating a condition" is directed to administering a composition, which includes at least one reagent (which may be, for example, one or more polynucleotide molecules, one or more expression vectors, one or more substance/ingredient, and the like), effective to ameliorate symptoms associated with a disease, to lessen the severity or cure the disease, or to prevent the disease from occurring. Administration may be any administration route.

The terms "Detection, "Diagnosis" refer to methods of detection of a disease, symptom, disorder, pathological or normal condition; classifying a disease, symptom, disorder, pathological condition; determining a severity of a disease, symptom, disorder, pathological condition; monitoring disease, symptom, disorder, pathological condition progression; forecasting an outcome and/or prospects of recovery thereof.

1. STRUCTURE OF A COMPOSITION OF THE INVENTION, ACCORDING TO SOME EMBODIMENTS

According to some embodiments of the present invention, there is provided a composition for directing cleavage of exogenous RNA of interest in response to the presence of an endogenous signal RNA in a cell. The exogenous RNA of interest is encoded from the composition. The endogenous signal RNA is an RNA molecule which comprises a predetermined signal sequence, such that the predetermined signal sequence is a random sequence of from 18 to 25 nucleotides in length.

In one embodiment of the invention, the composition comprises one or more polynucleotide molecules that comprise:

- (a) a polynucleotide sequence encoding the exogenous RNA of interest, such that the exogenous RNA of interest is an RNA sequence that comprises a specific sequence which is of sufficient complementarity to the predetermined signal sequence to direct target-specific RNA interference;
- (b) one or more polynucleotide sequence(s) encoding a functional RNA that is capable of effecting the cleavage, directly or indirectly, of the endogenous signal RNA at a predetermined cleavage site, such that the predetermined cleavage site is the 5' end of the predetermined signal sequence; and
- (c) a polynucleotide sequence encoding a carrier RNA which is an RNA molecule that is at least about 18 nucleotides in length and is consisting essentially of:
 - (1) a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, the edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides downstream from the predetermined cleavage site and extends downstream in the endogenous signal RNA;
 - (2) a second sequence downstream from the first sequence, such that the second sequence is a random sequence that is 0-5 nucleotides in length; and
 - (3) a third sequence upstream from the first sequence, such that the third sequence is 0-7000 nucleotides in length.

Thus, following introduction of the composition into a cell comprising the endogenous signal RNA, the functional RNA effects the cleavage, directly or indirectly, of

the endogenous signal RNA at the 5' end of the predetermined signal sequence and then the carrier RNA is hybridized to the edge sequence at the cleaved endogenous signal RNA and directs the processing of the predetermined signal sequence and then the processed predetermined signal sequence directs the cleavage of the exogenous RNA of interest at a specific target/cleavage site that is located within the specific sequence. For example, see FIG. 2.

In some embodiments, the composition comprises one or more polynucleotide molecules that comprise:

- (a) a polynucleotide sequence encoding the exogenous RNA of interest, such that the exogenous RNA of interest is an RNA sequence that comprises a specific sequence which is of sufficient complementarity to the predetermined signal sequence to direct target-specific RNA interference;
- (b) one or more polynucleotide sequence(s) encoding a functional RNA that is capable of effecting the cleavage, directly or indirectly, of the endogenous signal RNA at a predetermined cleavage site, such that the predetermined cleavage site is the 3' end of the predetermined signal sequence; and
- (c) a polynucleotide sequence encoding a carrier RNA which is an RNA molecule that is at least about 18 nucleotides in length and is consisting essentially of:
 - (1) a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, the edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides upstream from the predetermined cleavage site and extends upstream in the endogenous signal RNA;
 - (2) a second sequence upstream from the first sequence, such that the second sequence is a random sequence that is 0-5 nucleotides in length; and
 - (3) a third sequence downstream from the first sequence, such that the third sequence is 0-7000 nucleotides in length.

Thus, following introduction of the composition into a cell comprising the endogenous signal RNA, the functional RNA effects the cleavage, directly or indirectly, of the endogenous signal RNA at the 3' end of the predetermined signal sequence and then the carrier RNA is hybridized to the edge sequence at the cleaved endogenous signal RNA and directs the processing of the predetermined signal sequence and then the processed signal

sequence directs the cleavage of the exogenous RNA of interest at a specific target/cleavage site that is located within the specific sequence. For example, see FIG. 3.

In additional embodiment of the invention, the composition comprises one or more polynucleotide molecules that comprise:

- 5 (a) a polynucleotide sequence encoding the exogenous RNA of interest, such that the exogenous RNA of interest is an RNA sequence that comprises a specific sequence which is of sufficient complementarity to the predetermined signal sequence to direct target-specific RNA interference;
- (b) one or more polynucleotide sequence(s) encoding a functional RNA that is capable of
10 effecting the cleavage, directly or indirectly, of the endogenous signal RNA at a predetermined cleavage site, such that the predetermined cleavage site is the 5' end of the predetermined signal sequence;
- (c) a polynucleotide sequence encoding an RNA carrier sequence comprising a carrier sequence that is at least about 18 nucleotides in length and is consisting essentially of:
 - 15 (1) a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, the edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides downstream from the predetermined cleavage site and extends downstream in the endogenous signal RNA;
 - (2) a second sequence downstream from the first sequence, such that the second
20 sequence is a random sequence that is 0-5 nucleotides in length; and
 - (3) a third sequence upstream from the first sequence, such that the third sequence is 0-7000 nucleotides in length; and
- (d) one or more polynucleotide sequence(s) encoding a functional nucleic acid that is
25 capable of effecting the cleavage, directly or indirectly, of the carrier RNA sequence a carrier cleavage site, such that the carrier cleavage site is the 3' end of the carrier sequence.

Thus, following introduction of the composition into a cell comprising the endogenous signal RNA, the functional RNA affects the cleavage, directly or indirectly, of the endogenous signal RNA at the 5' end of the predetermined signal sequence and the
30 functional nucleic acid effects the cleavage, directly or indirectly, of the carrier RNA at the

3' end of the carrier sequence. The cleaved carrier RNA sequence is hybridized to the edge sequence at the cleaved endogenous signal RNA and directs the processing of the predetermined signal sequence. Then, the processed signal sequence may direct the cleavage of the exogenous RNA of interest at a specific target/cleavage site that is located within the specific sequence. For example, see FIG. 4.

In some embodiments, the composition comprises one or more polynucleotide molecules that comprise:

- (a) a polynucleotide sequence encoding the exogenous RNA of interest, such that the exogenous RNA of interest is an RNA sequence that comprises a specific sequence which is of sufficient complementarity to the predetermined signal sequence to direct target-specific RNA interference;
- (b) one or more polynucleotide sequence(s) encoding a functional RNA that is capable of effecting the cleavage, directly or indirectly, of the endogenous signal RNA at a predetermined cleavage site, such that the predetermined cleavage site is the 3' end of the predetermined signal sequence;
- (c) a polynucleotide sequence encoding a carrier RNA sequence comprising a carrier sequence that is at least about 18 nucleotides in length and is consisting essentially of:
 - (1) a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, the edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides upstream from the predetermined cleavage site and extends upstream in the endogenous signal RNA;
 - (2) a second sequence upstream from the first sequence, such that the second sequence is a random sequence that is 0-5 nucleotides in length; and
 - (3) a third sequence downstream from the first sequence, such that the third sequence is 0-7000 nucleotides in length; and
- (d) one or more polynucleotide sequence(s) encoding a functional nucleic acid that is capable of effecting the cleavage, directly or indirectly, of the carrier RNA sequence at a carrier cleavage site, such that the carrier cleavage site is the 5' end of the carrier sequence.

Thus, following introduction of the composition into a cell comprising the endogenous signal RNA, the functional RNA effects the cleavage, directly or indirectly, of the endogenous signal RNA at the 3' end of the predetermined signal sequence and the functional nucleic acid effects the cleavage, directly or indirectly, of the carrier RNA sequence at the 5' end of the carrier sequence. The cleaved carrier RNA sequence may then hybridize to the edge sequence at the cleaved endogenous signal RNA and direct the processing of the predetermined signal sequence. Then, the processed predetermined signal sequence may direct the cleavage of the exogenous RNA of interest at a specific target/cleavage site that is located within the specific sequence. For example, see FIG. 5.

According to some embodiments, the predetermined signal sequence may be chosen due to its presence within specific target cells, thereby providing a mechanism for targeting the cleavage of the exogenous RNA of interest in selected cells. The specific target cells may be any type of cells. For example, the specific target cells may be such cells as, but not limited to: benign or malignant neoplasms. On average, each tumor comprises mutations in 90 protein-coding genes [16]. Each tumor is initiated from a single founder cell [38], thus it is most probable that at least one of these mutant genes is transcribed into mRNA. The specific cells may also include, but are not limited to viral infected cells. Specificity may be achieved by modification of the sequences that encode the functional RNA, carrier RNA and/or the specific sequence in the exogenous RNA of interest.

In a co-pending application, which is directed to the activation of gene of interest in a cell expressing a specific miRNA, the predetermined signal sequence may comprise an endogenous miRNA.

According to some embodiments, the predetermined signal sequence of the present invention does not include an endogenous cellular miRNA molecule or any other type of endogenous RNA molecule (such as, for example, shRNA, ribozyme, stRNA, and the like), that is able to direct or effect cleavage of an RNA molecule within the cell.

In some embodiments, the predetermined signal sequence cannot induce/effect cleavage in the absence of one or more components of the composition of the invention.

Various methods have been developed to identify a predetermined signal sequence that is unique to specific cells. These methods include DNA microarray, Tilling (Targeting Induced Local Lesions In Genomes) and large-scale sequencing of cancer cells genomes. Furthermore the identification of the predetermined signal sequence is predicted to be even

simpler thanks to the Cancer Genome Atlas (NIH project), which was launched at 13.12.2005 and has been cataloguing all the genetic mutations responsible for cancer.

It has been reported that in mammal cells, the removal of the poly(A) tail reduces the functional mRNA half-life only by 2.6-fold and the removal of the cap reduces the functional mRNA half-life only by 1.7-fold [10]. It has also been reported that two portions of mRNA that has been cleaved by RISC-RNA complex in a cell can be easily detected by Northern analysis [6].

According to some embodiments, the carrier RNA/sequence of embodiments of the may be hybridized to the cleaved endogenous signal RNA portion that includes the predetermined signal sequence. It has been reported that in a cell, two RNA transcripts of about 23 nucleotides in length that have a complementary region of about 19 nucleotides in length at the 5' end are hybridized to each other and are capable of directing target specific RNA interference [7].

According to further embodiments, the duplex that comprises the carrier RNA and the cleaved endogenous signal RNA portion that includes the predetermined signal sequence may be a substrate for Dicer and thereafter for Risc. It has been reported that a dsRNA of 52 nucleotides long that further comprises 20 nucleotides long ssRNA at one of the 3' ends is a substrate for a Dicer at the blunt end [8]. It has also been reported that in mammalian cells, Risc is coupled to Dicer [9].

20 2. STRUCTURE OF THE FUNCTIONAL RNA AND THE FUNCTIONAL NUCLEIC ACID

This section describes various embodiments of the structure of the functional RNA and functional nucleic acid of the composition of the invention. This is illustrated, for example, in FIGs. 2, 3, 4, 5.

25 In another embodiment of the invention, the functional RNA, described in previous embodiments above (section 1) is:

(i) an inhibitory RNA comprising a sequence of from 18 to 25 nucleotides in length which is of sufficient complementarity to a target sequence for the inhibitory RNA to direct cleavage of the endogenous signal RNA at the predetermined cleavage site via, for example, RNA interference, such that the target sequence is a sequence of from 18 to 25

nucleotides in length that is located in a region within the endogenous signal RNA, such that the region is located from about 25 nucleotides downstream from the predetermined cleavage site to about 25 nucleotides upstream from the predetermined cleavage site; or a ribozyme capable of binding to the region of (i) and effecting the cleavage of the endogenous signal RNA at the predetermined cleavage site).

In one embodiment, the region of (i) that is described in the former embodiment may be located from about 11 nucleotides downstream from the predetermined cleavage site to about 12 nucleotides upstream from the predetermined cleavage site. In one embodiment, the region of (i) that is located from about 10 nucleotides downstream from the predetermined cleavage site to about 11 nucleotides upstream from the predetermined cleavage site. For example, see FIG 6A, 6B.

In another embodiment of the invention, the functional nucleic acid, described above (section 1) is:

- (i) an inhibitory RNA comprising a sequence of from 18 to 25 nucleotides in length which is of sufficient complementarity to a target sequence for the inhibitory RNA to direct cleavage of the carrier RNA sequence at the carrier cleavage site via, for example, RNA interference, whereby the target sequence is a sequence of from 18 to 25 nucleotides in length that is located in a region within the carrier RNA and wherein the region is located from about 25 nucleotides downstream from the carrier cleavage site to about 25 nucleotides upstream from the carrier cleavage site; or
- (ii) a ribozyme capable of binding to the region of (i) and effecting the cleavage of the carrier RNA at the carrier cleavage site. In one embodiment, the region of (i) that is described in the former embodiment is located from about 11 nucleotides downstream from the carrier cleavage site to about 12 nucleotides upstream from the carrier cleavage site. In another embodiment, the region of (i) is located from about 10 nucleotides downstream from the carrier cleavage site to about 11 nucleotides upstream from the carrier cleavage site. For example, see FIG 7A, 7B.

According to some embodiments, the inhibitory RNA of (i), described above may be, for example, but is not limited to: antisense RNA, double-stranded RNA (dsRNA) and/or small-interfering RNA (siRNA). In some embodiments, the inhibitory RNA of (i), may be, for example, but not limited to: microRNA (miRNA), lariat-form RNA, short-hairpin RNA (shRNA) and/or siRNA expression domain.

According to additional embodiments, the inhibitory RNA of (i) that comprises:

- (a) a first RNA molecule that comprises a nucleic acid sequence at the 5' or 3' end, such that the nucleic acid sequence is of sufficient complementarity to the target sequence of (i) to direct target-specific RNA interference; and
- 5 (b) a second RNA molecule that comprises a nucleotide sequence that is capable of binding to the nucleic acid sequence, such that the nucleotide sequence is 18-25 nucleotides in length and such that the nucleotide sequence is located at the 5' or 3' end of the second RNA molecule.

Thus, the first and second RNA molecule form 3'-overhang or 5'-overhang of 0-5
10 nucleotides on the active end of the duplex formed when each of the first and second RNA molecules is hybridized with the other whereby such that the active end of the duplex formed is the end that comprises the nucleic acid sequence and the nucleotide sequence.

In another embodiment, the first RNA molecule that is described in the former embodiment is about 25 to 30 nucleotides long and the second RNA molecule is about 25 to
15 30 nucleotides long, such that the first and second RNA molecules form 3'-overhang of 2 nucleotides on the active end of the duplex formed when each of the first and second RNA molecules is hybridized with the other and such that the duplex may be a substrate for a Dicer. For example, see FIG 8A, 8B.

In some embodiments, the ribozyme of (ii) that is described above, may be, for
20 example, but is not limited to: hammerhead-type ribozyme, hairpin ribozyme and/or tetrahymena-type ribozyme.

In additional embodiments, the ribozyme of (ii) is a hammerhead-type ribozyme [21] that comprises at the 3' end a first sequence of 7 nucleotides in length that is complementary to a sequence that is located 26 nucleotides upstream from the 3' end of the region of (i) and
25 extends upstream in the region of (i), furthermore the hammerhead-type ribozyme comprises at the 5' end a second sequence of 7 nucleotides in length that is complementary to a sequence that is located 18 nucleotides upstream from the 3' end of the region of (i) and extends upstream in the region of (i) [21]. For example, see FIG 9A.

In another embodiment, the ribozyme of (ii) is a hairpin ribozyme [21] that
30 comprises at the 5' end a nucleic acid sequence of 16 nucleotides in length, such that the nucleic acid sequence comprises at the 5' end a sequence of 8 nucleotides in length that is

complementary to a sequence that is located 28 nucleotides downstream from the 5' end of the region of (i) and extends downstream in the region of (i) and such that the nucleic acid sequence comprises at the 3' end a sequence of 4 nucleotides in length that is complementary to a sequence that is located 26 nucleotides upstream from the 3' end of the region of (i) and extends upstream in the region of (i) [21]. For example, see FIG 9B.

According to some embodiments, and without wishing to bound to theory or mechanism, the use of a ribozyme will not retain/use up, and consequently dilute, the cellular the components of the RNA interference pathway.

In another embodiment, the functional nucleic acid, described in embodiments in section 1 is a cis acting ribozyme that is located within the carrier RNA sequence and effects the cleavage of the carrier RNA at the carrier cleavage site. In some embodiments, the cis acting ribozyme may be, for example, but is not limited to the very efficient cis-acting hammerhead ribozyme: snorbozyme [22] and/or N117 [23]. For example, see FIG 10, 11.

According to some embodiments, and without wishing to bound to theory or mechanism, in by using a cis acting ribozyme, the carrier sequence that comprises it may be cleaved by itself [22], which may yield preferred results.

In another embodiment, the functional nucleic acid described in the embodiments in section 1 is an endonuclease recognition site or an endogenous miRNA binding site, such that the functional nucleic acid is located within the carrier RNA and is capable of effecting the cleavage, directly or indirectly, of the carrier RNA at the carrier cleavage site. For example, see FIG 12A, 12B.

3. STRUCTURE OF A FUNCTIONAL NUCLEIC ACID THAT HAS A STEM LOOP STRUCTURE

The functional nucleic acid that is described in embodiments of section 1 may be, for example, but is not limited to, a stem loop structure or miRNA structure, whereby the functional nucleic acid directs the cleavage of the carrier sequence at the carrier cleavage site. This section describes embodiments of the structure of this functional nucleic acid that has a stem loop structure or miRNA structure.

In some embodiments, the functional nucleic acid described in embodiments in section 1 is a miRNA sequence that is located within the carrier RNA sequence, such that following introduction of the composition into a cell, the miRNA sequence is processed, such that the processing of the miRNA sequence is capable of effecting the cleavage, directly or indirectly, of the carrier RNA at the carrier cleavage site and such that the processing of the miRNA sequence comprises Drosha processing. In one embodiment, the miRNA sequence that is described in the former embodiment comprises a sequence corresponding to a naturally occurring miRNA, or a sequence substantially identical thereto. For example, see FIG 12C, 12D.

In another embodiment, the functional nucleic acid, described in section 1 has a nucleotide sequence that is located immediately upstream from the 5' end of the carrier sequence in the carrier RNA sequence, such that the third sequence is 0 nucleotides in length and such that the nucleotide sequence is capable of binding to the carrier sequence, whereby the carrier sequence and the nucleotide sequence are capable of forming a stem loop structure that is a substrate for a Drosha. Following introduction of the composition into a cell, the stem loop structure may be processed and the processing of the stem loop structure is capable of affecting the cleavage, directly or indirectly, of the carrier RNA at the 3' end of the carrier sequence. In another embodiment, the stem loop structure that is described in the former embodiment is a maximum of about 150 nucleotides long and the processed stem loop structure is not a substrate for Dicer. For example, see FIG 13A.

In another embodiment, the functional nucleic acid, described in section 1 has a nucleotide sequence that is located immediately downstream from the 3' end of the carrier sequence in the carrier RNA, such that the third sequence is 0 nucleotides in length and such that the nucleotide sequence is capable of binding to the carrier sequence. The carrier sequence and the nucleotide sequence are capable of forming a stem loop structure that is a substrate for Drosha. Following introduction of the composition into a cell, the stem loop structure may be processed, and the processing of the stem loop structure is capable of effecting the cleavage, directly or indirectly, of the carrier RNA at the 5' end of the carrier sequence. In another embodiment, the stem loop structure that is described in the former embodiment has a maximum of about 150 nucleotides long and the processed stem loop structure is not a substrate for Dicer. For example, see FIG 13B.

According to some embodiments, and without wishing to bound to theory or mechanism, the use of miRNA sequence or stem loop structure may provide enhanced results since the carrier sequence may be cleaved independently by Drosha. In another embodiment, the functional nucleic acid described in embodiments of section 1 comprises:

- 5 (i) a first nucleotide sequence that is located immediately upstream from the 5' end of the carrier sequence in the carrier RNA, such that the third sequence is 0-50 nucleotides in length; and
- (ii) a second nucleotide sequence that is located immediately downstream from the 3' end of the carrier sequence, such that the second nucleotide sequence is capable of binding to
- 10 the first nucleotide sequence, such that the second nucleotide sequence and the first nucleotide sequence and the carrier sequence are capable of forming a stem loop structure.

Whereby, following introduction of the composition into a cell, the stem loop structure is processed, wherein the processing of the stem loop structure is capable of

15 effecting the cleavage, directly or indirectly, of the carrier RNA at the carrier cleavage site and such that the processing of the stem loop structure is capable of forming one or more RNA duplex(es). The processing of the stem loop structure may include, for example, Dicer processing and the RNA duplex(es) may be siRNA duplex(es) and/or miRNA duplex(es).

In additional embodiments, the functional RNA described in the former embodiment

20 is a nucleic acid sequence of from 18 to 25 nucleotides in length which is of sufficient complementarity to a target sequence to direct target-specific RNA interference, such that the target sequence is a sequence of from 18 to 25 nucleotides in length that is located in a region within the endogenous signal RNA. The region is located from about 25 nucleotides downstream from the predetermined cleavage site to about 25 nucleotides upstream from

25 the predetermined cleavage site, such that the nucleic acid sequence is located within the first nucleotide sequence or within the second nucleotide sequence. Following introduction of the composition into a cell, at least one RNA duplex from the one or more RNA duplex(es) comprises the nucleic acid sequence and the RNA duplex that comprises the nucleic acid sequence directs the cleavage of the endogenous signal RNA at the

30 predetermined cleavage site via RNA interference. For example, see FIG. 14A.

In another embodiment, the first nucleotide sequence or the second nucleotide sequence described in the former embodiment is the nucleic acid sequence, such that the

carrier RNA sequence is consisting essentially of: the first nucleotide sequence and the second nucleotide sequence and the carrier sequence. The first nucleotide sequence is 18-25 nucleotides in length and the second nucleotide sequence is 18-25 nucleotides in length. The stem loop structure forms 3'-overhang of 2 nucleotides and may be a substrate for a Dicer and such that the expression of the carrier RNA polynucleotide sequence is driven by
5 polymerase I based promoter or polymerase III based promoter. For example, see FIG. 14B.

In some embodiments, the region described in any of the previous 2 embodiments is located from about 11 nucleotides downstream from the predetermined cleavage site to about 12 nucleotides upstream from the predetermined cleavage site. In another
10 embodiment, the region described in the former embodiment is located from about 10 nucleotides downstream from the predetermined cleavage site to about 11 nucleotides upstream from the predetermined cleavage site.

According to some embodiments, and without wishing to bound to theory or mechanism, the use of functional RNA and a carrier sequence that are located in the same
15 RNA molecule may require less transcriptional units, which may yield advantageous results. Additional advantage of the proximity of the functional RNA and the carrier sequence is that they are synthesized in the same location in the cell at the same time and at a constant ratio.

4. STRUCTURE OF A CARRIER SEQUENCE/RNA AND FUNCTIONAL 20 RNA/NUCLEIC ACID THAT ARE LOCATED IN THE SAME RNA DUPLEX

This section describes various embodiments for the structure of the composition of the invention, described in section 1, wherein the carrier RNA and/or carrier sequence are located in the same RNA duplex together with the functional RNA or with the functional nucleic acid.

25 In another embodiment of the invention, the functional nucleic acid described in embodiments of section 1 may comprise:

- (i) a nucleotide sequence that is located immediately downstream from the 3' end of the carrier sequence in carrier RNA, such that the carrier RNA is consisting essentially of the carrier sequence and the nucleotide sequence; and (ii) one or more RNA molecule(s) that
30 are capable of binding to the nucleotide sequence; such that the nucleotide sequence and at least one of the RNA molecules form 3'-overhang or 5'-overhang of 0-5 nucleotides on

one end of the duplex, which is formed when each of the nucleotide sequence and the at least one RNA molecule is hybridized to the other. Following introduction of the composition into a cell, the nucleotide sequence and the one or more RNA molecule(s) are hybridized with each other and the nucleotide sequence is processed, such that the processing of the nucleotide sequence is capable of effecting the cleavage, directly or indirectly, carrier RNA at the carrier cleavage site and such that the processing of the nucleotide sequence is capable of forming one or more RNA duplex(es). The processing of the nucleotide sequence may include, for example, Dicer processing and the RNA duplex(es) may be siRNA duplex(es) and/or miRNA duplex(es).

In another embodiment of the invention, the functional nucleic acid, described in embodiments in section 1 may comprise:

(i) a nucleotide sequence that is located immediately upstream from the 5' end of the carrier sequence in the carrier RNA, such that the carrier RNA is consisting essentially of: the carrier sequence and the nucleotide sequence; and (ii) one or more RNA molecule(s) that are capable of binding to the nucleotide sequence; such that the nucleotide sequence and at least one of the RNA molecules from the one or more RNA molecule(s) form 3'-overhang or 5'-overhang of 0-5 nucleotides on one end of the duplex formed when each of the nucleotide sequence and the one RNA molecule is hybridized with the other. Following introduction of the composition into a cell, the nucleotide sequence and the one or more RNA molecule(s) are hybridized with the other and the nucleotide sequence is processed. The processing of the nucleotide sequence may be capable of effecting the cleavage, directly or indirectly, of the carrier RNA at the carrier cleavage site and the processing of the nucleotide sequence is capable of forming one or more RNA duplex(es), such that the processing of the nucleotide sequence may include, for example, Dicer processing and the RNA duplex(es) may be siRNA duplex(es) and/or miRNA duplex(es).

In some embodiments, the functional RNA described in any of the previous 2 embodiments is a nucleic acid sequence of from 18 to 25 nucleotides in length which is of sufficient complementarity to a target sequence to direct target-specific RNA interference. The target sequence is a sequence of from 18 to 25 nucleotides in length that is located in a region within the endogenous signal RNA, such that the region is located from about 25 nucleotides downstream from the predetermined cleavage site to about 25 nucleotides upstream from the predetermined cleavage site, such that the nucleic acid sequence is located within the nucleotide sequence or within at least one RNA molecule from the one or

more RNA molecule(s). Following introduction of the composition into a cell, at least one RNA duplex of the one or more RNA duplex(es) comprises the nucleic acid sequence, and the RNA duplex that comprises the nucleic acid sequence directs the cleavage of the endogenous signal RNA at the predetermined cleavage site via, for example, RNA interference.

In one embodiment, the region described in the former embodiment is located from about 11 nucleotides downstream from the predetermined cleavage site to about 12 nucleotides upstream from the predetermined cleavage site. In another embodiment, the one or more RNA molecule(s) that are described in any of the previous 2 embodiments is one RNA molecule, consisting essentially of the nucleic acid sequence, such that the nucleotide sequence is 18-25 nucleotides in length and the one RNA molecule is 18-25 nucleotides in length. The nucleotide sequence and the one RNA molecule form 3'-overhang of 2 nucleotides on one end of the duplex formed when each of the nucleotide sequence and the one RNA molecule is hybridized with the other, the expressions of the carrier RNA and the one RNA molecule are driven by polymerase I or III based promoter. For example, see FIG. 15A and 15B.

According to some embodiments, and without wishing to bound to theory or mechanism, when the functional RNA and the carrier sequence are located in the same RNA duplex, the carrier sequence may bring the functional RNA into proximity with the predetermined signal sequence of the endogenous signal RNA and may further bring also the components of the RNA interference pathway (for example, Dicer and Risc) into proximity with the predetermined signal sequence.

In another embodiment of the invention, the functional RNA described in embodiments in section 1 comprises:

- (i) a nucleotide sequence that is located at the 5' end of the carrier RNA; and (ii) one or more RNA molecule(s) that are capable of binding to the nucleotide sequence; such that the nucleotide sequence and at least one RNA molecule form 3'-overhang or 5'-overhang of 0-5 nucleotides on one end of the duplex, which is formed when each of the nucleotide sequence and the at least one RNA molecule is hybridized with the other. The nucleotide sequence or at least one RNA molecule include a nucleic acid sequence of from 18 to 25 nucleotides in length that is of sufficient complementarity to a target sequence to direct target-specific RNA interference, such that the target sequence is a sequence of from 18

to 25 nucleotides in length that is located in a region within the endogenous signal RNA, such that the region is located from about 25 nucleotides downstream from the predetermined cleavage site to about 25 nucleotides upstream from the predetermined cleavage site. Following introduction of the composition into a cell, the nucleotide sequence and the one or more RNA molecule(s) may be hybridized with the other and the nucleotide sequence may be processed, such that the processing of the nucleotide sequence is capable of forming one or more RNA duplex(es). The processing of the nucleotide sequence may include Dicer processing and the RNA duplex(es) may be siRNA duplex(es) and/or miRNA duplex(es), such that at least one RNA duplex from the one or more RNA duplex(es) comprises the nucleic acid sequence and such that the RNA duplex that comprises the nucleic acid sequence directs the cleavage of the endogenous signal RNA at the predetermined cleavage site via RNA interference.

In another embodiment of the invention, the functional RNA described in embodiments in section 1 may comprise:

- (i) a nucleotide sequence that is located at the 3' end of the carrier RNA; and (ii) one or more RNA molecule(s) that are capable of binding to the nucleotide sequence; such that the nucleotide sequence and at least one RNA molecule form 3'-overhang or 5'-overhang of 0-5 nucleotides on one end of the duplex which is formed when each of the nucleotide sequence and the one RNA molecule is hybridized with the other. The nucleotide sequence or at least one RNA molecule from the one or more RNA molecule(s) comprises a nucleic acid sequence of from 18 to 25 nucleotides in length that is of sufficient complementarity to a target sequence to direct target-specific RNA interference, such that the target sequence is a sequence of from 18 to 25 nucleotides in length that is located in a region within the endogenous signal RNA, wherein the region is located from about 25 nucleotides downstream from the predetermined cleavage site to about 25 nucleotides upstream from the predetermined cleavage site. Following introduction of the composition into a cell, the nucleotide sequence and the one or more RNA molecule(s) are hybridized with each other and the nucleotide sequence is processed, such that the processing of the nucleotide sequence is capable of forming one or more RNA duplex(es). The processing of the nucleotide sequence may include, for example, Dicer processing and the RNA duplex(es) may be siRNA duplex(es) and/or miRNA duplex(es), such that at least one RNA duplex from the one or more RNA duplex(es) comprises the nucleic acid sequence and such that the RNA duplex that comprises the nucleic acid sequence

directs the cleavage of the endogenous signal RNA at the predetermined cleavage site via RNA interference.

In additional embodiment, the region that is described in any of the previous 2 embodiments is located from about 11 nucleotides downstream from the predetermined cleavage site to about 12 nucleotides upstream from the predetermined cleavage site. In another embodiment, the one or more RNA molecule(s) that is described in any of the previous 3 embodiments is one RNA molecule, such that the nucleotide sequence or the one RNA molecule is consisting essentially of the nucleic acid sequence. The nucleotide sequence is 18-25 nucleotides in length and the one RNA molecule is 18-25 nucleotides in length, such that the nucleotide sequence and the RNA molecule form 3'-overhang of 2 nucleotides on one end of the duplex, which is formed when each of the nucleotide sequence and the one RNA molecule is hybridized with the other. In some embodiments, the expression of the carrier RNA and the one RNA molecule are driven by polymerase I or III based promoter. For example, see FIG. 16A and 16B.

According to some embodiments, and without wishing to bound to theory or mechanism, when the functional RNA and the carrier RNA are located in the same RNA duplex, the carrier RNA may bring the functional RNA into proximity with the predetermined signal sequence of the endogenous signal RNA and by this may also bring the components of the RNA interference pathway (for example, Dicer and Risc) into proximity with the predetermined signal sequence.

In another embodiment of the invention, the functional RNA described in embodiments in section 1 may comprise:

(i) a nucleotide sequence that is located at the 5' end of the carrier RNA; and; (ii) one or more RNA molecule(s) that are capable of binding to the nucleotide sequence; such that the nucleotide sequence and at least one RNA molecule form 3'-overhang or 5'-overhang of 0-5 nucleotides on one end of the duplex formed, when each of the nucleotide sequence and the RNA molecule is hybridized with each other. The nucleotide sequence or at least one RNA molecule may comprise a nucleic acid sequence of from 18 to 25 nucleotides in length, such that the nucleic acid sequence is of sufficient complementarity to a target sequence to direct target-specific RNA interference. The target sequence is a sequence of from 18 to 25 nucleotides in length that is located in a region within the endogenous signal RNA, such that the region is located from about 25 nucleotides

downstream from the predetermined cleavage site to about 25 nucleotides upstream from the predetermined cleavage site. Following introduction of the composition into a cell, the nucleotide sequence and the RNA molecule are hybridized with each other and the nucleotide sequence is processed, such that the processing of the nucleotide sequence is capable of forming one or more RNA duplex(es). The processing of the nucleotide sequence may include, for example, Dicer processing and the RNA duplex(es) may be siRNA duplex(es) and/or miRNA duplex(es), such that at least one RNA duplex from the one or more RNA duplex(es) includes the nucleic acid sequence and such that the RNA duplex that comprises the nucleic acid sequence directs the cleavage of the endogenous signal RNA at the predetermined cleavage site via RNA interference.

In another embodiment of the invention, the functional RNA described in embodiments in section 1 may comprise:

(i) a nucleotide sequence that is located at the 3' end of the carrier RNA sequence; and;
(ii) one or more RNA molecule(s) that are capable of binding to the nucleotide sequence; such that the nucleotide sequence and at least one RNA molecule form 3'-overhang or 5'-overhang of 0-5 nucleotides on one end of the duplex, which is formed when each of the nucleotide sequence and the one RNA molecule is hybridized with each other. The nucleotide sequence or the at least one RNA molecule may comprise a nucleic acid sequence of from 18 to 25 nucleotides in length, such that the nucleic acid sequence is of sufficient complementarity to a target sequence to direct target-specific RNA interference. The target sequence is a sequence of from 18 to 25 nucleotides in length that is located in a region within the endogenous signal RNA, such that the region is located from about 25 nucleotides downstream from the predetermined cleavage site to about 25 nucleotides upstream from the predetermined cleavage site. Following introduction of the composition into a cell, the nucleotide sequence and the one or more RNA molecule(s) are hybridized with each other and the nucleotide sequence is processed, such that the processing of the nucleotide sequence is capable of forming one or more RNA duplex(es). The processing of the nucleotide sequence may include, for example, Dicer processing and the RNA duplex(es) may be siRNA duplex(es) and/or miRNA duplex(es), such that at least one RNA duplex from the one or more RNA duplex(es) may comprise the nucleic acid sequence and such that the RNA duplex that comprises the nucleic acid sequence may direct the cleavage of the endogenous signal RNA at the predetermined cleavage site via RNA interference.

In another embodiment, the region that described in any of the previous two embodiments may be located from about 11 nucleotides downstream from the predetermined cleavage site to about 12 nucleotides upstream from the predetermined cleavage site. In another embodiment, the one or more RNA molecule(s) that are described
5 in any of the previous 3 embodiments is one RNA molecule, such that the nucleotide sequence or the one RNA molecule is consisting essentially of the nucleic acid sequence. The nucleotide sequence may be 18-25 nucleotides in length and the one RNA molecule is 18-25 nucleotides in length. The nucleotide sequence and the one RNA molecule form 3'-overhang of 2 nucleotides on one end of the duplex, which is formed when each of the
10 nucleotide sequence and the one RNA molecule is hybridized with each other. The expression of the carrier RNA and the one RNA molecule are driven by polymerase I or III based promoter. For example, see FIG. 17A and 17B.

In some embodiments, the functional nucleic acid described in embodiments in section 1 may comprise:

- 15 (i) a nucleotide sequence that is located at the 5' end of the carrier RNA sequence; and;
- (ii) one or more RNA molecule(s) that are capable of binding to the nucleotide sequence; such that the nucleotide sequence and one RNA molecule from the one or more RNA molecule(s) form 3'-overhang or 5'-overhang of 0-5 nucleotides on one end of the duplex, which is formed when each of the nucleotide sequence and the one RNA molecule is
20 hybridized with the other. Such that the nucleotide sequence or at least one RNA molecule from the one or more RNA molecule(s) comprises a nucleic acid sequence of from 18 to 25 nucleotides in length, such that the nucleic acid sequence is of sufficient complementarity to a target sequence to direct target-specific RNA interference, such that the target sequence is a sequence of from 18 to 25 nucleotides in length that is located in a
25 region within the carrier RNA and such that the region is located from about 25 nucleotides downstream from the carrier cleavage site to about 25 nucleotides upstream from the carrier cleavage site. Such that following introduction of the composition into a cell, the nucleotide sequence and the one or more RNA molecule(s) are hybridized with the other and the nucleotide sequence is processed, such that the processing of the
30 nucleotide sequence is capable of forming one or more RNA duplex(es). The processing of the nucleotide sequence may include, for example, Dicer processing and the RNA duplex(es) may be siRNA duplex(es) and/or miRNA duplex(es), such that at least one

RNA duplex from the one or more RNA duplex(es) comprises the nucleic acid sequence and such that the RNA duplex that comprises the nucleic acid sequence directs the cleavage of the carrier RNA at the carrier cleavage site via RNA interference.

5 In another embodiment of the invention, the functional nucleic acid described in embodiments in section 1 comprises:

- (i) a nucleotide sequence that is located at the 3' end of the carrier RNA sequence); and
 - (ii) one or more RNA molecule(s) that are capable of binding to the nucleotide sequence, such that the nucleotide sequence and one RNA molecule from the one or more RNA molecule(s) form 3'-overhang or 5'-overhang of 0-5 nucleotides on one end of the duplex
- 10 formed when each of the nucleotide sequence and the one RNA molecule is hybridized with the other.

Such that the nucleotide sequence or at least one RNA molecule from the one or more RNA molecule(s) comprises a nucleic acid sequence of from 18 to 25 nucleotides in length, wherein the nucleic acid sequence is of sufficient complementarity to a target

15 sequence to direct target-specific RNA interference, such that the target sequence is a sequence of from 18 to 25 nucleotides in length that is located in a region within the carrier RNA sequence and such that the region is located from about 25 nucleotides downstream from the carrier cleavage site to about 25 nucleotides upstream from the carrier cleavage site. Following introduction of the composition into a cell, the nucleotide sequence and the

20 one or more RNA molecule(s) are hybridized with each other and the nucleotide sequence is processed, such that the processing of the nucleotide sequence is capable of forming one or more RNA duplex(es). The processing of the nucleotide sequence may include, for example, Dicer processing and the RNA duplex(es) may comprise siRNA duplex(es) and/or miRNA duplex(es), such that at least one RNA duplex from the one or more RNA duplex(es)

25 comprises the nucleic acid sequence and such that the RNA duplex that comprises the nucleic acid sequence directs the cleavage of the carrier RNA at the carrier cleavage site via RNA interference.

In another embodiment, the region described in any of the previous 2 embodiments may be located from about 11 nucleotides downstream from the carrier cleavage site to

30 about 12 nucleotides upstream from the carrier cleavage site. In another embodiment, the one or more RNA molecule(s) described in any of the previous 3 embodiments is one RNA molecule, such that the nucleotide sequence or the one RNA molecule is consisting

essentially of the nucleic acid sequence. The nucleotide sequence is 18-25 nucleotides in length, such that the one RNA molecule is 18-25 nucleotides in length, and the nucleotide sequence and the one RNA molecule form 3'-overhang of 2 nucleotides on one end of the duplex formed, when each of the nucleotide sequence and the one RNA molecule is
5 hybridized with the other. The expression of the carrier RNA and the one RNA molecule may be driven by polymerase I or III based promoter. For example, see FIG. 18A and 18B.

In another embodiment, the functional RNA is a specific nucleotide sequence of from 18 to 25 nucleotides in length which is of sufficient complementarity to a specific target sequence to direct target-specific RNA interference. The specific target sequence is a
10 sequence of from 18 to 25 nucleotides in length that is located in a specific region within the endogenous signal RNA, such that the specific region is located from about 25 nucleotides downstream from the predetermined cleavage site to about 25 nucleotides upstream from the predetermined cleavage site. The specific nucleotide sequence is located within the nucleotide sequence or within at least one RNA molecule from the one or more RNA
15 molecule(s). Following introduction of the composition into a cell, at least one RNA duplex from the one or more RNA duplex(es) include the specific nucleotide sequence and such that the RNA duplex that comprises the specific nucleotide sequence directs the cleavage of the endogenous signal RNA at the predetermined cleavage site via RNA interference. For example, see FIG. 19A and 19B.

20 5 STRUCTURE OF A CARRIER RNA SEQUENCE THAT COMPRISES AT LEAST 3 CONTIGUOUS CARRIER SEQUENCES

This section describes the structure of the carrier RNA described in section 1, such that the carrier RNA comprises at least 3 contiguous carrier sequences.

In some embodiments of the invention, the carrier RNA described in the
25 embodiments in section 1 may further comprise at least 2 contiguous carrier sequences immediately downstream from the described carrier sequence. For example, see FIG. 20A.

In an additional embodiment, the carrier RNA described in embodiments of section 1 may further comprise 100 contiguous carrier sequences (that may be identical or different) immediately downstream from the carrier sequence described therein, such that the edge
30 sequence is 23-28 nucleotides in length and is located from the predetermined cleavage site to about 23-28 nucleotides downstream; the second sequence is 2 nucleotides in length; the

third sequence is 0 nucleotides in length and the expression of the polynucleotide sequence the carrier RNA is driven by CMV-IE promoter.

In further embodiments, the carrier RNA described in the embodiments in section 1 may further comprise at least 2 contiguous carrier sequences immediately upstream from the carrier sequence. For example, see FIG. 20B.

In another embodiment, the carrier RNA, described in embodiments in section 1 may further comprise 100 contiguous carrier sequences, that may be identical or different, immediately upstream from the carrier sequence, such that the edge sequence is 25-30 nucleotides in length and is located 2 nucleotides upstream from the predetermined cleavage site and extends upstream in the endogenous signal RNA; the second sequence is 0 nucleotides in length; the third sequence is 0 nucleotides in length; and such that the expression of the polynucleotide sequence the carrier RNA is driven by CMV-IE promoter.

According to some embodiments, and without wishing to bound to theory or mechanism, advantageous results may be obtained when the functional nucleic acid is, for example, microRNA (miRNA), lariat-form RNA, short-hairpin RNA (shRNA), siRNA expression domain, small-interfering RNA (siRNA) and/or trans acting ribozyme, since with such functional nucleic acids, many carrier sequences can be generated from one carrier RNA and from one functional nucleic acid.

6. STRUCTURE OF POLYNUCLEOTIDE MOLECULE(S) THAT MAY FURTHER TRANSCRIBE AN ADDITIONAL FUNCTIONAL RNA THAT MAY CLEAVE THE PREDETERMINED SIGNAL SEQUENCE AT THE OPPOSITE SIDE, WHICH IS NOT CLEAVED

This section describes embodiments of the structure of the polynucleotide molecule(s) (such as, for example, DNA molecules), described in embodiments of section 1, such that the polynucleotide molecule(s) together further transcribe an additional functional RNA that is capable of effecting the cleavage of the endogenous signal RNA at the opposite end of the predetermined signal sequence, which is not cleaved.

In some embodiments of the invention, the polynucleotide molecule(s) described in some embodiments of section 1 further comprise a polynucleotide sequence encoding an additional functional RNA that is capable of effecting the cleavage, directly or indirectly, of the endogenous signal RNA at an additional cleavage site, such that the additional cleavage

site may be located 0-1000 nucleotides downstream from the 3' end of the predetermined signal sequence. In one embodiment, the additional cleavage site may be located 0-5 nucleotides downstream from the 3' end of the predetermined signal sequence. For example, see FIG. 21A.

5 In additional embodiments, the polynucleotide molecule(s) described in some embodiments in section 1 may further comprise a polynucleotide sequence encoding an additional functional RNA that is capable of effecting the cleavage, directly or indirectly, of the endogenous signal RNA at an additional cleavage site, such that the additional cleavage site may be located 0-1000 nucleotides upstream from the 5' end of the predetermined signal
10 sequence. In one embodiment, the additional cleavage site may be located 0-5 nucleotides upstream from the 5' end of the predetermined signal sequence. For example, see FIG. 21B.

According to some embodiments, and without wishing to bound to theory or mechanism, the previous 4 embodiments may be advantageous since in these embodiments the predetermined signal sequence may be cleaved at both of its ends and thus with the
15 carrier RNA/sequence it may be a better substrate for endogenous enzymes, such as, for example, Dicer and/or Risc.

7. STRUCTURE OF AN EXOGENOUS RNA OF INTEREST

In some embodiments of the invention, the exogenous RNA of interest described in embodiments in section 1 may further comprise:

20 a sequence encoding an exogenous protein of interest; and
an inhibitory sequence that is capable of inhibiting the expression of the exogenous protein of interest;

whereby the specific target/cleavage site is located between the inhibitory sequence and the sequence encoding the exogenous protein of interest. Following introduction of the
25 composition into a cell comprising the endogenous signal RNA, the exogenous RNA of interest is transcribed and cleaved at the specific target/cleavage site whereby the inhibitory sequence is detached from the sequence encoding the exogenous protein of interest and the exogenous protein of interest is capable of being expressed. For example, see FIG. 22A and 22B. Accordingly, cleaving of the exogenous RNA of interest may lead to the expression of
30 an active exogenous protein of interest within the cell.

In some embodiments, the exogenous RNA of interest molecule may further comprise a carrier RNA sequence and/or a Functional RNA sequence.

As known in the art, mRNAs without cap or poly A tail are still capable of translating proteins. In mammalian cells, an addition of a cap increases the translation of an mRNA by 35-50 fold and an addition of a poly(A) tail increases the translation of an mRNA by 114-155- fold [10]. The poly(A) tail in mammalian cells increases the functional mRNA half-life only by 2.6-fold and the cap increases the functional mRNA half-life only by 1.7-fold [10].

Some proteins may be biologically active even at a concentration of one protein per cell. It has been reported that a single protein of Ricin or Abrin reaching the cytosol can kill that cell [12, 13]. In addition, a single protein of Diphtheria toxin fragment A introduced into a cell can kill the cell [14]. The exogenous protein of interest of the invention can be any protein or peptide. For example, in some embodiments, the exogenous protein of interest may be any type of toxin (such as, for example, Ricin, Abrin, Diphtheria toxin (DTA), botulinum toxin); an enzyme; a reporter gene; a structural gene, and the like. In some embodiments, the exogenous protein of interest may be a polypeptide which is a fusion product of two proteins, that may be have a cleavage site there between, allowing the separation of the two proteins within the cell. For example, the exogenous protein of interest may be a fusion protein of Ricin and DTA, whereby cleavage of the fusion protein by, for example, a specific protease, can result in the formation of separate DTA and Ricin proteins in the cell. In some embodiment, the exogenous protein of interest may include two separate proteins, that may be expressed by the composition. For example, the exogenous RNA of interest may encode for two separate exogenous proteins of interest, such as, for example, Ricin and DTA.

7.1. STRUCTURE OF AN EXOGENOUS RNA OF INTEREST HAVING AN INHIBITORY SEQUENCE LOCATED UPSTREAM FROM THE SPECIFIC TARGET/CLEAVAGE SITE

7.1.1. STRUCTURE OF THE INHIBITORY SEQUENCE THAT IS LOCATED UPSTREAM FROM THE SPECIFIC TARGET/CLEAVAGE SITE

The inhibitory sequence in the exogenous RNA of interest described in embodiments of Section 7 may be located upstream or downstream from the specific target/cleavage site.

This section describes the structure of the inhibitory sequence that is located upstream from the specific target/cleavage site in the exogenous RNA of interest, according to some embodiments. For example, see FIG. 22A.

In some embodiments of the invention, the inhibitory sequence that is located upstream from the specific target/cleavage site and that is described in embodiments in section 7 may comprise, for example, but is not limited to an initiation codon, whereby the initiation codon and the sequence encoding for the exogenous protein of interest are not in the same reading frame, such that the initiation codon causes a frameshift mutation to the protein of interest that is encoded downstream. For example, see FIG. 23A. In one embodiment, the initiation codon is located within a Kozak consensus sequence. In addition, modified Kozak consensus sequences that maintain the ability to function as initiator of translation may be also used. In some embodiments, any initiator of translation element may be used. For example, see FIG. 23B.

For example, the Kozak consensus sequence in human is 5'-ACCAUGG-3' (SEQ ID NO. 25) and the initiation codon is 5'-AUG-3'.

In another embodiment of the invention, the inhibitory sequence that is located upstream from the specific target/cleavage site and that is described in section 7 comprises a plurality of initiation codons, whereby each of the initiation codons and the sequence encoding exogenous protein of interest are not in the same reading frame, such that the initiation codons cause a frameshift mutation to the exogenous protein of interest that is encoded downstream. In addition, each of the initiation codons is located within a Kozak consensus sequence or a modified Kozak consensus sequences that maintain the ability to function as initiator of translation. For example, see FIG. 23C.

In some embodiments, the initiation codon may be located within or may comprise one or more TISU motifs. A TISU (Translation Initiator of Short 5'UTR) motif is distinguished from a Kozak consensus in its unique ability to direct efficient and accurate translation initiation from mRNAs with a very short 5'UTR. [39]. In other embodiment of the invention, the inhibitory sequence that is located upstream from the specific target/cleavage site and that is described in the specific embodiment in section 7 comprises an initiation codon and the exogenous RNA of interest further comprises a stop codon located between the initiation codon and the start codon of the sequence encoding the exogenous protein of interest, such that the stop codon and the initiation codon are in the

same reading frame. Such a structure creates an upstream open reading frame (uORF) that reduces the efficiency of translation of the downstream sequence encoding protein of interest. For example, see FIG. 24A. In some embodiments, the stop codon may be, for example, 5'-UAA-3' or 5'-UAG-3' or 5'-UGA-3'.

5 In some embodiments, strong stems and loops may be located downstream to upstream ORF(s) at a location that is upstream or downstream to the target sequence for the miRNA (cleavage site). The creation of such stems and loops may aid in conditions, wherein despite having reached a stop codon, the small subunit of the ribosome does not detach from the mRNA continue to scan the mRNA. The small subunit of the ribosome is
10 not capable of opening strong RNA secondary structures. Additionally, when these stems and loops are located downstream to the target sequence they may also block the degradation of the cleaved mRNA which may be performed, for example, by XRN1 exonuclease.

In another embodiment of the invention, the inhibitory sequence that is located
15 upstream from the specific target/cleavage site and that is described in the specific embodiment in section 7 comprises an initiation codon and a nucleotide sequence downstream from the initiation codon that encodes a sorting/localization/targeting signal for subcellular localization, such that the nucleotide sequence and the initiation codon are in the same reading frame and such that the subcellular localization of the protein of interest
20 inhibits its biological function. The sorting/localization signal for the subcellular localization includes, but is not limited to sorting localization signal for mitochondria, nucleus, endosome, lysosome, peroxisome, ER, or any subcellular localization or organelle. The sorting signal for the subcellular localization may be selected from, for example, but is not limited to: a peroxisomal targeting signal 2 [(R/K)(L/V/I)X5(Q/H)(L/A)] (SEQ ID NO.
25 26) or H2N----RLRVLSGHL (SEQ ID NO. 27) (of human alkyl dihydroxyacetonephosphate synthase) [30]. For example, see FIG. 24B.

In another embodiment of the invention, the inhibitory sequence that is located upstream from the specific target/cleavage site and that is described in the specific embodiment in section 7 comprises an initiation codon and a nucleotide sequence
30 downstream from the initiation codon that encodes a protein degradation signal, such that the nucleotide sequence and the initiation codon are in the same reading frame. The protein

degradation signal includes, but is not limited to ubiquitin degradation signal. For example, see FIG. 24B.

In other embodiments of the invention, the inhibitory sequence that is located upstream from the specific target/cleavage site and that is described in section 7 is designed to comprise an initiation codon and a nucleotide sequence downstream from the initiation codon that is in the same reading frame with the initiation codon and with the sequence encoding the exogenous protein of interest, such that when the amino acid sequence, which is encoded by the nucleotide sequence, is fused to the protein of interest the biological function of the protein of interest is inhibited. For example, see FIG. 24C.

In another embodiment of the invention, the inhibitory sequence that is located upstream from the specific target/cleavage site and that is described in section 7 comprises an initiation codon, and the exogenous RNA of interest further comprises a stop codon downstream from the initiation codon, such that the stop codon and the initiation codon are in the same reading frame. In addition, the exogenous RNA of interest may further comprise an intron downstream from the stop codon, such that the exogenous RNA of interest is a target for nonsense-mediated decay (NMD) that degrades the exogenous RNA of interest. For example, see FIG. 24D.

In another embodiment of the invention, the inhibitory sequence that is located upstream from the specific target/cleavage site and that is described in section 7 comprises a sequence that is capable of binding to a translation repressor protein, such that the translation repressor protein is an endogenous translation repressor protein or is encoded from the composition and such that the translation repressor protein, directly or indirectly, reduces the efficiency of translation of the protein of interest within the exogenous RNA of interest [26]. The sequence that is capable of binding to a translation repressor protein includes, for example, but is not limited to a sequence that binds the smaug repressor protein (5'-UGGAGCAGAGGCUCUGGCAGCUUUUGCAGCG-3') (SEQ ID NO. 28) [27]. For example, see FIG. 25A.

In other embodiments of the invention, the inhibitory sequence that is located upstream from the specific target/cleavage site and that is described in section 7 comprises an RNA localization signal for subcellular localization (including cotranslational import) or an endogenous miRNA binding site, such that the subcellular localization of the exogenous RNA of interest of the invention inhibits the translation of the protein of interest and

decreases the exogenous RNA of interest half-life. The RNA localization signal may be, for example, but is not limited to RNA localization signal for: myelinating periphery, mitochondria, myelin compartment, leading edge of the lamella or Perinuclear cytoplasm [24]. For example, the RNA localization signal for myelinating periphery is 5'-GCCAAGGAGCCAGAGAGCAUG-3' (SEQ ID NO. 29) or 5'-GCCAAGGAGCC-3' (SEQ ID NO. 30) [29]. For example, see FIG. 25B.

In another embodiment of the invention, the inhibitory sequence that is located upstream from the specific target/cleavage site and that is described in section 7 comprises an RNA destabilizing element that stimulates degradation of the exogenous RNA of interest, such that the RNA destabilizing element is an AU-rich element (ARE) or an endonuclease recognition site. The AU-rich element may be, for example, but is not limited to AU-rich elements that are at least about 35 nucleotides long. The AU-rich element may be, for example, but is not limited to: 5'-AUUUA-3' (SEQ ID NO. 31), 5'-UUAUUUA(U/A)(U/A)-3' (SEQ ID NO. 32) or 5'-AUUU-3' (SEQ ID NO. 33) [28]. For example, see FIG. 25C.

In another embodiment of the invention, the inhibitory sequence that is located upstream from the specific target/cleavage site and that is described in section 7 comprises a sequence that is capable of forming a secondary structure that reduces the efficiency of translation of the downstream exogenous protein of interest. In one embodiment of the invention, and without wishing to bound to theory or mechanism, the free energy of folding of the secondary structure that described in the former embodiment may be lower than -30 kcal/mol (for example, -50 kcal/mol, -80 kcal/mol) and thus the secondary structure is sufficient to block scanning ribosomes to reach the start codon of the downstream protein of interest. For example, see FIG. 25D.

In further embodiment of the invention, the inhibitory sequence that is located upstream from the specific target/cleavage site and that is described in section 7 comprises a sequence immediately upstream from the specific target/cleavage site that is capable of binding to the nucleotide sequence that is located immediately downstream from the specific target/cleavage site for the formation of a secondary structure, such that the secondary structure, directly or indirectly, reduces the efficiency of translation of the downstream exogenous protein of interest.

In some embodiments, the free energy of the folding of the secondary structure that is described in the former embodiment may be lower than -30 kcal/mol (for example, -50

kcal/mol and -80 kcal/mol) and thus this secondary structure is sufficient to block scanning ribosomes to reach the start codon of the protein of interest. In another embodiment, the specific target/cleavage site is located within the single stranded region or within the loop region in the secondary structure that is described in the former embodiment, such that the single stranded region or the loop region includes, but is not limited to region that is at least about 15 nucleotides long. In another embodiment, the exogenous RNA of interest, described in the former embodiment comprises an internal ribosome entry site (IRES) sequence downstream from the specific target/cleavage site and upstream from the sequence encoding protein of interest, such that the IRES sequence is more functional within the cleaved exogenous RNA of interest than within the intact exogenous RNA of interest. In other embodiment, at least part of the IRES sequence is located within the nucleotide sequence that is located immediately downstream from the specific target/cleavage site. For example, see FIG. 26.

In some embodiments, the IRES sequence includes, for example, but is not limited to a picornavirus IRES, a foot-and-mouth disease virus IRES, an encephalomyocarditis virus IRES, a hepatitis A virus IRES, a hepatitis C virus IRES, a human rhinovirus IRES, a poliovirus IRES, a swine vesicular disease virus IRES, a turnip mosaic potyvirus IRES, a human fibroblast growth factor 2 mRNA IRES, a pestivirus IRES, a Leishmania RNA virus IRES, a Moloney murine leukemia virus IRES a human rhinovirus 14 IRES, an aphthovirus IRES, a human immunoglobulin heavy chain binding protein mRNA IRES, a Drosophila Antennapedia mRNA IRES, a human fibroblast growth factor 2 mRNA IRES, a hepatitis G virus IRES, a tobamovirus IRES, a vascular endothelial growth factor mRNA IRES, a Coxsackie B group virus IRES, a c-myc protooncogene mRNA IRES, a human MYT2 mRNA IRES, a human parechovirus type 1 virus IRES, a human parechovirus type 2 virus IRES, a eukaryotic initiation factor 4G1 mRNA IRES, a Plautia stali intestine virus IRES, a Theiler's murine encephalomyelitis virus IRES, a bovine enterovirus IRES, a connexin 43 mRNA IRES, a homeodomain protein Gtx mRNA IRES, an AML1 transcription factor mRNA IRES, an NF-kappa B repressing factor mRNA IRES, an X-linked inhibitor of apoptosis mRNA IRES, a cricket paralysis virus RNA IRES, a p58(PITSLRE) protein kinase mRNA IRES, an ornithine decarboxylase mRNA IRES, a connexin-32 mRNA IRES, a bovine viral diarrhea virus IRES, an insulin-like growth factor I receptor mRNA IRES, a human immunodeficiency virus type 1 gag gene IRES, a classical swine fever virus IRES, a Kaposi's sarcoma-associated herpes virus IRES, a short IRES selected from a library of

random oligonucleotides, a Jembrana disease virus IRES, an apoptotic protease-activating factor 1 mRNA IRES, a Rhopalosiphum padi virus IRES, a cationic amino acid transporter mRNA IRES, a human insulin-like growth factor II leader 2 mRNA IRES, a giardavirus IRES, a Smad5 mRNA IRES, a porcine teschovirus-1 talfan IRES, a Drosophila Hairless mRNA IRES, an hSNM1 mRNA IRES, a Cbfa1/Runx2 mRNA IRES, an Epstein-Barr virus IRES, a hibiscus chlorotic ringspot virus IRES, a rat pituitary vasopressin V1b receptor mRNA IRES and/or a human hsp70 mRNA IRES.

7.1.2. ADDITIONAL STRUCTURES THAT MAY INCREASE THE EFFICIENCY OF TRANSLATION OF THE EXOGENOUS RNA OF INTEREST THAT IS CLEAVED AT THE SPECIFIC TARGET/CLEAVAGE SITE AT THE 5' END

This section describes further embodiments of additional structures of the composition of the invention that are described in embodiments in section 7, such that the additional structures may increase the efficiency of translation of the cleaved exogenous RNA of interest, wherein the cleaved exogenous RNA of interest is cleaved at the specific target/cleavage site at the 5' end.

In some embodiments, the exogenous RNA of interest described in section 7 may comprise, for example, a sequence that comprises a unique internal ribosome entry site (IRES) sequence immediately upstream from the sequence encoding the exogenous protein of interest, such that the unique IRES sequence increases the efficiency of translation of the protein of interest in the cleaved exogenous RNA of interest. For example, see FIG. 27A.

In another embodiment of the invention, the exogenous RNA of interest that is described in section 7 may comprise a unique nucleotide sequence immediately downstream from the sequence encoding the protein of interest, such that the unique nucleotide sequence comprises a unique stem loop structure and such that the unique stem loop structure, directly or indirectly, increases the efficiency of translation of the protein of interest and the cleaved exogenous RNA of interest half-life. The unique stem loop structure may be, for example, but is not limited to the conserved stem loop structure of the human histone gene 3'-UTR or a functional derivative thereof. The conserved stem loop structure of the human histone gene 3'-UTR is 5'-GGCUCUUUUCAGAGCC-3' (SEQ ID NO. 34). For example, see FIG. 27B.

In another embodiment of the invention, the exogenous RNA of interest that is described in the specific embodiment in section 7 may comprise a unique nucleotide

sequence immediately downstream from the sequence encoding the protein of interest, such that the unique nucleotide sequence comprises a cytoplasmic polyadenylation element that, directly or indirectly, increases the efficiency of translation of the protein of interest and the half-life of the cleaved exogenous RNA of interest. The cytoplasmic polyadenylation element may be, for example, but is not limited to 5'-UUUUAU-3' (SEQ ID NO. 35), 5'-UUUUUAU-3' (SEQ ID NO. 36), 5'-UUUUAAU-3' (SEQ ID NO. 37), 5'-UUUUUUAU-3' (SEQ ID NO. 38), 5'-UUUUAUU-3' (SEQ ID NO. 39) or 5'-UUUUUAUAAAG-3' (SEQ ID NO. 40) [25]. In some embodiments, the composition of the invention may also include, for example, a polynucleotide sequence that encodes a human cytoplasmic polyadenylation element binding protein (hCPEB), and/or a homologue thereof, for expressing hCPEB in any cell. For example, see FIG. 27C.

In another embodiment of the invention, the exogenous RNA of interest that is described in section 7 comprises a unique nucleotide sequence that is located downstream from the specific target/cleavage site and upstream from the sequence encoding the exogenous protein of interest, such that the unique nucleotide sequence is capable of binding to a sequence that is located downstream from the sequence encoding protein of interest. Without wishing to be bound to theory or mechanism, in such embodiment, the cleaved exogenous RNA of interest may create a circular structure that increases the efficiency of translation of the protein of interest in the cleaved exogenous RNA of interest. For example, see FIG. 27D.

In another embodiment of the invention, the exogenous RNA of interest that is described in section 7 comprises a unique nucleotide sequence that is located downstream from the specific target/cleavage site and upstream from the sequence encoding protein of interest. The unique nucleotide sequence may be capable of binding to a unique polypeptide that is, directly or indirectly, capable of binding to the poly(A) tail in the cleaved exogenous RNA of interest, and the unique polypeptide may be encoded from the composition of the invention. Without wishing to bound to theory or mechanism, in such embodiment the unique polypeptide and the cleaved exogenous RNA of interest may create a circular structure that increases the efficiency of translation of the protein of interest in the cleaved exogenous RNA of interest. For example, see FIG. 28A.

7.1.3. ADDITIONAL STRUCTURES THAT MAY REDUCE THE EFFICIENCY OF TRANSLATION OF THE INTACT EXOGENOUS RNA OF INTEREST

This section describes further embodiments of additional structures of the composition of the invention, described in section 7, such that these additional structures may reduce the efficiency of translation of the exogenous RNA of interest of the invention before it is cleaved (that is, an intact exogenous RNA of interest).

5 In some embodiments, the composition that is described in section 7 may further comprise a particular cleaving component(s) that is capable of effecting the cleavage, directly or indirectly, of the exogenous RNA of interest of the invention at a position that is located upstream from the inhibitory sequence, wherein the inhibitory sequence is located upstream from the specific target/cleavage site. In some embodiments, the particular
10 cleaving component(s) may comprise:

- (a) a particular nucleic acid sequence that is located within the exogenous RNA of interest, such that the particular nucleic acid sequence may be, for example, but not limited to: endonuclease recognition site, endogenous miRNA binding site, cis acting ribozyme and/or miRNA sequence; or
- 15 (b) a particular inhibitory RNA that is encoded from the composition of the invention, such that the particular inhibitory RNA may be, for example, but is not limited to: microRNA (miRNA), lariat-form RNA, short-hairpin RNA (shRNA), siRNA expression domain, antisense RNA, double-stranded RNA (dsRNA), small-interfering RNA (siRNA) and/or ribozyme.

20 Without wishing to bound to theory or mechanism, in such embodiment the particular cleaving component(s) may remove the CAP structure from the intact exogenous RNA of interest of the invention for reducing the efficiency of translation of the protein of interest in the intact exogenous RNA of interest. For example, see FIG. 28B.

In some embodiments, a vpg recognition sequence may be introduced, such that
25 upon cleave, the 5' cleaved end contains a vpg recognition sequence. To the vpg recognition sequence a VPG protein may bind, thereby replacing the CAP. The vpg protein may be encoded by the composition of the invention or by the first ORF of the inhibitory sequence.

According to some embodiments, and without wishing to bound to theory or mechanism, the use of cis acting ribozyme may be advantageous because the exogenous
30 RNA of interest that comprises it may be cleaved by itself [22]. The cis acting ribozyme may be, for example, but is not limited to cis-acting hammerhead ribozymes: snorbozyme [22] or N117 [23].

7.2. STRUCTURE OF THE EXOGENOUS RNA OF INTEREST HAVING ITS INHIBITORY SEQUENCE LOCATED DOWNSTREAM FROM THE SPECIFIC TARGET/CLEAVAGE SITE

7.2.1. STRUCTURE OF THE INHIBITORY SEQUENCE THAT IS LOCATED 5 DOWNSTREAM FROM THE SPECIFIC TARGET/CLEAVAGE SITE

The inhibitory sequence in the exogenous RNA of interest that is described in embodiments of section 7 can be located upstream or downstream from the specific target/cleavage site. In some embodiments, the inhibitory sequence may be located downstream from the specific target/cleavage site in the exogenous RNA of interest. For
10 example, see FIG. 22B.

In another embodiment of the invention, the inhibitory sequence that is located downstream from the specific target/cleavage site that is described in section 7 may be, for example, but is not limited to an intron. Such that the exogenous RNA of interest is a target for nonsense-mediated decay (NMD) that degrades the exogenous RNA of interest [31]. For
15 example, see FIG. 29A.

In other embodiment of the invention, the inhibitory sequence that is located downstream from the specific target/cleavage site and that is described in section 7 includes a sequence that is capable of binding to a translation repressor protein, such that the translation repressor protein is an endogenous translation repressor protein or is encoded
20 from the composition and such that the translation repressor protein, directly or indirectly, reduces the efficiency of translation of the protein of interest within the exogenous RNA of interest [26]. The sequence that is capable of binding to a translation repressor protein may be, for example, the binding sequence of smaug repressor protein (5'-UGGAGCAGAGGCUCUGGCAGCUUUUGCAGCG-3') (SEQ ID NO. 28) [27]. For
25 example, see FIG. 29B.

In another embodiment of the invention, the inhibitory sequence that is located downstream from the specific target/cleavage site and that is described in section 7 comprises an RNA localization signal for subcellular localization (including cotranslational import) or an endogenous miRNA binding site, such that the subcellular localization of the
30 exogenous RNA of interest of the invention inhibits the translation of the exogenous protein of interest and decreases the exogenous RNA of interest half-life. The RNA localization

signal may be, for example, but is not limited to RNA localization signal for: myelinating periphery, mitochondria, myelin compartment, leading edge of the lamella and/or Perinuclear cytoplasm [24]. The RNA localization signal may be, for example, RNA localization signal for myelinating periphery 5'-GCCAAGGAGCCAGAGAGCAUG-3' (SEQ ID NO. 29) or 5'-GCCAAGGAGCC-3' (SEQ ID NO. 30) [29]. For example, see FIG. 29C.

In another embodiment of the invention, the inhibitory sequence that is located downstream from the specific target/cleavage site and that is described in section 7 may be, for example, an RNA destabilizing element that stimulates degradation of the exogenous RNA of interest, such that the RNA destabilizing element is an AU-rich element (ARE) or an endonuclease recognition site. The AU-rich element may be, for example, AU-rich elements that are at least about 35 nucleotides long. The AU-rich element may be, for example, 5'-AUUUA-3' (SEQ ID NO. 31), 5'-UUAUUUA(U/A)(U/A)-3' (SEQ ID NO. 32) or 5'-AUUU-3' (SEQ ID NO. 33) [28]. For example, see FIG. 29D.

In another embodiment of the invention, without wishing to be bound to theory or mechanism, the inhibitory sequence that is located downstream from the specific target/cleavage site and that is described in section 7 comprises a sequence that is capable of forming a secondary structure that reduces the efficiency of translation of the upstream protein of interest. For example, see FIG. 29E.

In another embodiment of the invention the inhibitory sequence that is located downstream from the specific target/cleavage site and that is described in section 7 comprises a sequence immediately downstream from the specific target/cleavage site that is capable of binding to the nucleotide sequence that is located immediately upstream from the specific target/cleavage site for the formation of a secondary structure, such that the secondary structure, directly or indirectly, reduces the efficiency of translation of the upstream protein of interest. In some embodiments, the free energy of the folding of the secondary structure that is described in the former embodiment may be lower than -30 kcal/mol (for example, -50 kcal/mol, -80 kcal/mol) and hence this secondary structure may be sufficient to block scanning ribosomes from reaching the stop codon of the protein of interest. In another embodiment, the specific target/cleavage site is located within the single stranded region or within the loop region in the secondary structure that is described in the former embodiment, such that the single stranded region or the loop region include, for

example, but is not limited to a region that is at least about 15 nucleotides long. For example, see FIG. 30A.

7.2.2. ADDITIONAL STRUCTURES THAT MAY INCREASE THE EFFICIENCY OF TRANSLATION OF THE EXOGENOUS RNA OF INTEREST WHICH IS CLEAVED AT THE SPECIFIC TARGET/CLEAVAGE SITE AT THE 3' END

This section describes further embodiments of additional structures of the composition of the invention described in various embodiments of section 7, such that these additional structures may increase the efficiency of translation of the cleaved exogenous RNA of interest, wherein the cleaved exogenous RNA of interest is cleaved at the specific target/cleavage site at the 3' end.

According to some embodiments, the exogenous RNA of interest of the invention that is described in section 7 may comprise a sequence that comprises a unique internal ribosome entry site (IRES) sequence immediately upstream from the sequence encoding protein of interest, such that the unique IRES sequence may increase the efficiency of translation of the protein of interest in the cleaved exogenous RNA of interest. For example, see FIG. 30B.

In another embodiment of the invention, the exogenous RNA of interest that is described in section 7 may comprise a unique nucleotide sequence immediately downstream from the sequence encoding protein of interest, such that the unique nucleotide sequence comprises a unique stem loop structure and such that the unique stem loop structure, directly or indirectly, increases the efficiency of translation of the protein of interest and the cleaved exogenous RNA of interest half-life. The unique stem loop structure may include such structures as, but not limited to: a conserved stem loop structure of the human histone gene 3'-UTR or a functional derivative thereof. For example, the conserved stem loop structure of the human histone gene 3'-UTR is 5'-GGCUCUUUUCAGAGCC-3' (SEQ ID NO. 34). For example, see FIG. 30C.

In another embodiment, the exogenous RNA of interest that is described in section 7 may comprise a unique nucleotide sequence immediately downstream from the sequence encoding protein of interest, such that the unique nucleotide sequence comprises a cytoplasmic polyadenylation element that, directly or indirectly, may increase the efficiency of translation of the protein of interest and the cleaved exogenous RNA of interest half-life.

The cytoplasmic polyadenylation element may be selected from such elements as, but not limited to: 5'-UUUUUAU-3'(SEQ ID NO. 35), 5'-UUUUUAU-3'(SEQ ID NO. 36), 5'-UUUUAAU-3'(SEQ ID NO. 37), 5'-UUUUUUAU-3'(SEQ ID NO. 38), 5'-UUUUAAU-3'(SEQ ID NO. 39) or 5'-UUUUUAUAAAG-3' (SEQ ID NO. 40) [25]. The composition of the invention may also comprise, for example, a polynucleotide sequence that encodes a human cytoplasmic polyadenylation element binding protein (hCPEB), or a homologue thereof for expressing hCPEB in any cell. For example, see FIG. 30D.

In additional embodiment of the invention, the exogenous RNA of interest that is described in section 7 may comprise a unique nucleotide sequence that is located upstream from the specific target/cleavage site and downstream from the sequence encoding protein of interest, such that the unique nucleotide sequence is capable of binding to a sequence that is located upstream from the sequence encoding protein of interest. In such embodiment and without wishing to be bound to theory or mechanism, the cleaved exogenous RNA of interest may create a circular structure that may increase the efficiency of translation of the protein of interest in the cleaved exogenous RNA of interest. For example, see FIG. 31A.

In another embodiment of the invention, the exogenous RNA of interest that is described in section 7 may comprise a unique nucleotide sequence that is located upstream from the specific target/cleavage site and downstream from the sequence encoding protein of interest, the unique nucleotide sequence may be capable of binding to a unique polypeptide that is, directly or indirectly, capable of binding to the CAP structure in the cleaved exogenous RNA of interest, wherein the unique polypeptide is encoded from the composition of the invention. In this embodiment and without wishing to be bound to theory or mechanism, the unique polypeptide and the cleaved exogenous RNA of interest may create a circular structure that may increase the efficiency of translation of the protein of interest in the cleaved exogenous RNA of interest. For example, see FIG. 31B.

In another embodiment, the composition of the invention that is described in section 7 may comprise an additional polynucleotide sequence, which may encode an additional RNA molecule that comprises at the 3' end a nucleotide sequence that is capable of binding to a sequence that is located upstream from the specific target/cleavage site and downstream from the sequence encoding protein of interest, such that the expression of the additional polynucleotide sequence is driven by polymerase II based promoter. In such embodiment and without wishing to be bound to theory or mechanism, the additional RNA molecule is

capable of binding to the cleaved exogenous RNA of interest and provide him poly-A which may increase the efficiency of translation of the exogenous protein of interest in the cleaved exogenous RNA of interest. For example, see FIG. 31C.

5 7.2.3. ADDITIONAL STRUCTURES THAT MAY REDUCE THE EFFICIENCY OF TRANSLATION OF THE INTACT EXOGENOUS RNA OF INTEREST

This section describes further embodiments of additional structures of the composition of the invention that is described in embodiments of section 7, such that these additional structures reduce the efficiency of translation of the exogenous RNA of interest of the invention before it is cleaved.

10 In some embodiments of the invention, the composition that is described in section 7 may comprise a particular cleaving component(s) that is capable of effecting the cleavage, directly or indirectly, of the exogenous RNA of interest of embodiments of the invention at a position that is located downstream from the inhibitory sequence, wherein the inhibitory sequence is located downstream from the specific target/cleavage site. The particular
15 cleaving component(s) is:

- (a) a particular nucleic acid sequence that is located within the exogenous RNA of interest, such that the particular nucleic acid sequence may be, for example: endonuclease recognition site, endogenous miRNA binding site, cis acting ribozyme, miRNA sequence and the like; or
- 20 (b) a particular inhibitory RNA that is encoded from the composition of the invention, such that the particular inhibitory RNA may be, for example, microRNA (miRNA), lariat-form RNA, short-hairpin RNA (shRNA), siRNA expression domain, antisense RNA, double-stranded RNA (dsRNA), small-interfering RNA (siRNA) or ribozyme.

In this embodiment and without wishing to be bound to theory or mechanism, the
25 particular cleaving component(s) may remove the poly-A from the intact exogenous RNA of interest of the invention and may thus reduce the efficiency of translation of the protein of interest in the intact exogenous RNA of interest. For example, see FIG. 31D.

In some embodiments, and without wishing to be bound to theory or mechanism, the use of cis acting ribozyme may be advantageous because the exogenous RNA of interest

that comprises it may be cleaved by itself [22]. The cis acting ribozyme may be, for example, cis-acting hammerhead ribozymes: snorbozyme [22] or N117 [23].

8. DETAILED DESCRIPTION OF EXEMPLARY EMBODIMENTS OF THE INVENTION

5 According to some embodiments, as detailed below, an exogenous protein of interest may be expressed in response to the presence of an endogenous signal RNA in a cell without the involvement of Risc (RNA-induced silencing complex) mechanism.

According to some specific embodiments, there is provided a composition for expressing an exogenous protein of interest in response to the presence of endogenous
10 signal RNA in a cell, the exogenous protein of interest is encoded from the composition, the endogenous signal RNA is an RNA molecule which comprises a predetermined signal sequence, the predetermined signal sequence is a predetermined sequence that is at least 18 nucleotides in length and the composition may comprise one or more polynucleotide molecule(s), such as, for example, DNA molecules, that comprise:

15 (a) one or more polynucleotide sequence(s) encoding a functional RNA that is capable of effecting the cleavage, directly or indirectly, of the endogenous signal RNA at a predetermined cleavage site, such that the predetermined cleavage site is the 3' end of the predetermined signal sequence; and

(b) a polynucleotide sequence encoding an exogenous RNA of interest molecule, which is
20 an RNA molecule that is consisting essentially of:

(1) a first sequence which is of sufficient complementarity to an edge sequence to hybridize therewith, the edge sequence is located 0-5 nucleotides upstream from the predetermined cleavage site and extends upstream in the endogenous signal RNA, such that the first sequence comprises one or more initiation codon(s), such that each of the
25 initiation codons is consisting essentially of 5'-AUG-3'; and

(2) a second sequence upstream from the first sequence, such that the second sequence is a random sequence that is 0-5 nucleotides in length; and

(3) a third sequence downstream from the first sequence, such that the third sequence is 0-7000 nucleotides in length; and

such that the exogenous RNA of interest molecule comprises a sequence encoding exogenous protein of interest at least 21 nucleotides downstream from the 5' end of the exogenous RNA of interest molecule; and such that, following introduction of the composition into a cell comprising the endogenous signal RNA, the functional RNA effects the cleavage, directly or indirectly, of the endogenous signal RNA at the 3' end of the predetermined signal sequence. Thereby, the exogenous RNA of interest molecule is hybridized to the edge sequence at the cleaved endogenous signal RNA and directs the predetermined signal sequence to Dicer processing that may cleave the exogenous RNA of interest molecule, such that each of the initiation codon(s) is detached from the sequence encoding the exogenous protein of interest and the exogenous protein of interest is capable of being expressed. For example, see FIG. 33.

In some embodiments of the invention, the edge sequence is 25-30 nucleotides in length and is located 2 nucleotides upstream from the predetermined cleavage site and extends upstream in the endogenous signal RNA and such that the second sequence is 0 nucleotides in length, such as shown, for example in FIG. 33.

In another embodiment of the invention, at least one of the initiation codon(s) is located within a Kozak consensus sequence, such as, for example, Kozak consensus sequence 5'-ACCAUGG-3' (SEQ ID NO. 25), demonstrated for example in FIG. 34.

In additional embodiment of the invention, each of the initiation codon(s) is located 0-21 nucleotides downstream from the 5' end of the exogenous RNA of interest molecule, such that each of the initiation codon(s) and the sequence encoding the exogenous protein of interest are not in the same reading frame. In further embodiment of the invention, at least one of the initiation codon(s) described above is located within a Kozak consensus sequence such as, for example, Kozak consensus sequence 5'-ACCAUGG-3' (SEQ ID NO. 25).

In some embodiments, the functional RNA may be, for example, but is not limited to: microRNA (miRNA), lariat-form RNA, short-hairpin RNA (shRNA), siRNA expression domain, antisense RNA, double-stranded RNA (dsRNA), small-interfering RNA (siRNA), ribozyme, or combinations thereof.

In some exemplary embodiments, the functional RNA may be, for example, a microRNA (miRNA), such that the miRNA and the exogenous RNA of interest molecule are capable of being located on the same or different RNA molecules. In some

embodiments, the miRNA may be located upstream from the second sequence at the exogenous RNA of interest molecule, as demonstrated, for example, in FIG. 34.

According to some embodiments, and without wishing to be bound to theory or mechanism, the previous embodiment may be advantageous since in such embodiment, the CAP structure may be removed from the exogenous RNA of interest molecule and since in
5 this embodiment the composition encodes for only one RNA molecule.

In another exemplary embodiment, the functional RNA may be, for example, a small-interfering RNA (siRNA), such that one RNA strand of the siRNA is located at the 5' end of the exogenous RNA of interest molecule and the other strand of the siRNA is
10 transcribed from the composition by, for example, polymerase I or polymerase III based promoter(s) and such that following introduction of the composition into a cell, both of the siRNA strands are hybridized and detached from the exogenous RNA of interest molecule, for example, by Dicer. This is demonstrated, for example, in FIG. 35.

According to some embodiments, and without wishing to be bound to theory or
15 mechanism, the previous embodiment may be advantageous since in such embodiment the functional RNA and the exogenous RNA of interest molecule are located in the same RNA duplex, thus the exogenous RNA of interest molecule brings the functional RNA into proximity with the predetermined signal sequence of the endogenous signal RNA and by this may also bring also the components of the RNA interference pathway (such as, for
20 example, Dicer) into proximity with the predetermined signal sequence. Another advantage includes the removal of the CAP structure from the exogenous RNA of interest molecule by Dicer.

In another embodiment of the invention, the exogenous RNA of interest molecule may further comprise a nucleotide sequence located upstream from the sequence encoding
25 the protein of interest and downstream from each of the initiation codons, such that this nucleotide sequence is of sufficient complementarity to the predetermined signal sequence or to the sequence that is located at the 5' end of the exogenous RNA of interest molecule, and is able to direct target-specific RNA interference. For example, the Risc processing that follows the Dicer processing can be used for activating more exogenous RNA of interest
30 molecule molecules.

In another embodiment of the invention, the composition further comprises one or more polynucleotide sequence(s) encoding a functional nucleic acid that is capable of

effecting the cleavage, directly or indirectly, of the exogenous RNA of interest molecule upstream from the second sequence, such that the functional nucleic acid is:

- 5 (a) a specific nucleic acid sequence that is located within the exogenous RNA of interest molecule, such that the specific nucleic acid sequence is: endonuclease recognition site, endogenous miRNA binding site, cis acting ribozyme or miRNA sequence; or
- (b) an inhibitory RNA that is encoded from a DNA molecule(s), such that the inhibitory RNA is: microRNA (miRNA), lariat-form RNA, short-hairpin RNA (shRNA), siRNA expression domain, antisense RNA, double-stranded RNA (dsRNA), small-interfering RNA (siRNA) or ribozyme.

10 In this embodiment, the functional nucleic acid may remove the CAP structure from the intact exogenous RNA of interest for reducing the efficiency of translation of the exogenous protein of interest from the non-cleaved (intact) exogenous RNA of interest. For example, see FIG. 36A.

In another embodiment, the third sequence described above, includes a nucleotide sequence upstream from the sequence encoding protein of interest, such that the nucleotide sequence is capable of binding to a sequence that is located downstream from the sequence encoding protein of interest. In this embodiment the cleaved exogenous RNA of interest molecule creates a circular structure that may increase the efficiency of translation of the protein of interest in the cleaved exogenous RNA of interest molecule. For example, see

20 FIG. 36B.

In another embodiment of the invention, the polynucleotide molecule(s) that is described above together further comprise a polynucleotide sequence encoding an additional functional RNA that is capable of effecting the cleavage, directly or indirectly, of the endogenous signal RNA at an additional cleavage site, such that the additional cleavage site is located 0-1000 nucleotides upstream from the 5' end of the predetermined signal sequence. For example, see FIG. 21B. In another embodiment, the additional cleavage site that described in the former embodiment is located 0-5 nucleotides upstream from the 5' end of the predetermined signal sequence. For example, see FIG. 21B.

9. DESCRIPTION OF ADDITIONAL EMBODIMENTS OF THE INVENTION

This section describes additional embodiments of the invention that are directed to the cleavage of the exogenous RNA of interest in response to the presence of an endogenous signal RNA in a cell, without the cleaving the endogenous signal RNA. Such embodiments may be useful, for example, for endogenous signal RNA of a viral origin.

5 According to some embodiments, there is provided a composition for cleaving an exogenous RNA of interest in response to the presence of an endogenous signal RNA in a cell, the exogenous RNA of interest is encoded from the composition, the endogenous signal RNA is an RNA molecule which comprises a predetermined signal sequence at the 5' end, the predetermined signal sequence is a predetermined sequence of from 18 to 25 nucleotides
10 in length, the composition comprises one or more polynucleotides molecules (such as, for example, DNA and/or RNA molecules), that comprise:

(a) a polynucleotide sequence encoding the exogenous RNA of interest, such that the exogenous RNA of interest is an RNA sequence that comprises a specific sequence which is of sufficient complementarity to the predetermined signal sequence to direct, for
15 example, target-specific RNA interference;

(b) a polynucleotide sequence encoding a carrier RNA, such that expression of the polynucleotide sequence the carrier RNA sequence is driven by a promoter selected from the group consisting of: polymerase I based promoter and polymerase III based promoter, such that the carrier RNA is an RNA molecule that is at least about 18 nucleotides in
20 length and is consisting essentially of:

(1) a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, the edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides downstream from the 5' end of the endogenous signal RNA and extends downstream in the endogenous signal RNA;

25 (2) a second sequence downstream from the first sequence, such that the second sequence is a random sequence that is 0-5 nucleotides in length; and

(3) a third sequence upstream from the first sequence, such that the third sequence is 0-7000 nucleotides in length; and

whereby, following introduction of the composition into a cell comprising the
30 endogenous signal RNA, the carrier RNA is hybridized to the edge sequence and directs the processing of the predetermined signal sequence, and then the processed predetermined

signal sequence may direct the cleavage of the exogenous RNA of interest at a specific target (cleavage) site that is located within the specific sequence. For example, see FIG. 37A.

In additional embodiments, there is provided a composition for cleaving exogenous RNA of interest in response to the presence of an endogenous signal RNA in a cell, the exogenous RNA of interest is encoded from the composition, the endogenous signal RNA is an RNA molecule which comprises a predetermined signal sequence at the 5' end, the signal sequence is a predetermined sequence of from 18 to 25 nucleotides in length, the composition comprises one or more polynucleotide molecules (such as, for example, DNA molecules and/or RNA molecule), the polynucleotide molecules together comprise:

- (a) a polynucleotide sequence encoding the exogenous RNA of interest, such that the exogenous RNA of interest is an RNA sequence that comprises a specific sequence which is of sufficient complementarity to the predetermined signal sequence to direct cleavage, for example, by target-specific RNA interference;
- (b) a polynucleotide sequence encoding an RNA sequence that comprises a carrier sequence that is at least about 18 nucleotides in length, the carrier sequence consisting essentially of:
 - (1) a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, the edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides downstream from the 5' end of the endogenous signal RNA and extends downstream in the endogenous signal RNA;
 - (2) a second sequence downstream from the first sequence, such that the second sequence is a random sequence that is 0-5 nucleotides in length; and
 - (3) a third sequence upstream from the first sequence, such that the third sequence is 0-7000 nucleotides in length; and
- (c) one or more polynucleotide sequence(s) encoding a functional nucleic acid that is capable of effecting the cleavage, directly or indirectly, of the carrier RNA sequence at a carrier cleavage site, such that the carrier cleavage site is the 3' end of the carrier sequence;

whereby, following introduction of the composition into a cell comprising the endogenous signal RNA, the functional nucleic acid effects the cleavage, directly or

indirectly, of the carrier RNA sequence at the 3' end of the carrier sequence and then the cleaved carrier sequence is hybridized to the edge sequence and directs the processing of the predetermined signal sequence and then the processed predetermined signal sequence directs the cleavage of the exogenous RNA of interest at a specific cleavage/target site that
5 is located within the specific sequence. For example, see FIG. 37B.

In some embodiments of the invention, the edge sequence, described above is 23-29 nucleotides in length and may be located from the 5' end of the endogenous signal RNA to about 23-29 nucleotides downstream, such that the second sequence may be 2 nucleotides in length and such that the third sequence may be 0 nucleotides in length. For example, see
10 FIG. 37A, 37B.

In an additional embodiment, there is provided a composition for cleaving exogenous RNA of interest in response to the presence of an endogenous signal RNA in a cell, the exogenous RNA of interest is encoded from the composition, the endogenous signal RNA is an RNA molecule which comprises a predetermined signal sequence at the 3' end,
15 the predetermined signal sequence is a random sequence of from 18 to 25 nucleotides in length, the composition comprises one or more polynucleotide molecules (such as, for example, DNA or RNA molecules), the polynucleotide molecules together comprise:

(a) a polynucleotide sequence encoding the exogenous RNA of interest, such that the exogenous RNA of interest is an RNA sequence that comprises a specific sequence which
20 is of sufficient complementarity to the predetermined signal sequence to direct cleavage, for example, by target-specific RNA interference;

(b) a polynucleotide sequence encoding a carrier RNA, such that expression of the carrier RNA is driven by a promoter selected from the group consisting of: polymerase I based promoter and polymerase III based promoter, the carrier RNA is an RNA molecule that is
25 at least about 18 nucleotides in length and is consisting essentially of:

- (1) a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, the edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides upstream from the 3' end of the endogenous signal RNA and extends upstream in the endogenous signal RNA;
- 30 (2) a second sequence upstream from the first sequence, such that the second sequence is a random sequence that is 0-5 nucleotides in length; and

(3) a third sequence downstream from the first sequence, such that the third sequence is 0-7000 nucleotides in length; and

whereby, following introduction of the composition into a cell comprising the endogenous signal RNA, the carrier RNA is hybridized to the edge sequence and directs the processing of the predetermined signal sequence and then the processed predetermined signal sequence directs the cleavage of the exogenous RNA of interest at a specific cleavage/target site that is located within the specific sequence. For example, see FIG. 38A.

According to further embodiments, there is provided a composition for cleaving exogenous RNA of interest in response to the presence of an endogenous signal RNA in a cell, the exogenous RNA of interest is encoded from the composition, the endogenous signal RNA is an RNA molecule which comprises a predetermined signal sequence at the 3' end, the predetermined signal sequence is a random/predetermined sequence of from 18 to 25 nucleotides in length, the composition comprises one or more polynucleotide molecules (such as, for example, DNA molecules and/or RNA molecules), the polynucleotide molecules together comprise:

(a) a polynucleotide sequence encoding the exogenous RNA of interest, such that the exogenous RNA of interest is an RNA sequence that comprises a specific sequence which is of sufficient complementarity to the predetermined signal sequence to direct cleavage, for example, by target-specific RNA interference;

(b) a polynucleotide sequence encoding an RNA sequence that comprises a carrier sequence that is at least about 18 nucleotides in length, the carrier sequence consisting essentially of:

(1) a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, the edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides upstream from the 3' end of the endogenous signal RNA and extends upstream in the endogenous signal RNA;

(2) a second sequence upstream from the first sequence, such that the second sequence is a random sequence that is 0-5 nucleotides in length; and

(3) a third sequence downstream from the first sequence, such that the third sequence is 0-7000 nucleotides in length; and

(c) one or more polynucleotide sequence(s) encoding a functional nucleic acid that is capable of effecting the cleavage, directly or indirectly, of the carrier RNA sequence at a carrier cleavage site, such that the carrier cleavage site is the 5' end of the carrier sequence;

5 whereby, following introduction of the composition into a cell comprising the endogenous signal RNA, the functional nucleic acid effects the cleavage, directly or indirectly, of the carrier RNA sequence at the 5' end of the carrier sequence and then the cleaved carrier sequence is hybridized to the edge sequence and directs the processing of the predetermined signal sequence and then the processed predetermined signal sequence
10 directs the cleavage of the exogenous RNA of interest at a specific cleavage/target site that is located within the specific sequence. For example, see FIG. 38B.

 According to some exemplary embodiments, the edge sequence described above is about 25-30 nucleotides in length and may be located 2 nucleotides upstream from the 3' end of the endogenous signal RNA and extends upstream in the endogenous signal RNA,
15 such that the second sequence may be 0 nucleotides in length and such that the third sequence is 0 nucleotides in length. For example, see FIG. 38A, 38B.

 According to some embodiments, the functional nucleic acid described above is:

(a) a specific nucleic acid sequence that is located within the carrier RNA sequence and such that the specific nucleic acid sequence is, for example, endonuclease recognition
20 site, endogenous miRNA binding site, cis acting ribozyme, a miRNA sequence, and the like, or combinations thereof; or

(b) an inhibitory RNA that is encoded from the polynucleotide molecule(s), such that the inhibitory RNA is, for example, microRNA (miRNA), lariat-form RNA, short-hairpin RNA (shRNA), siRNA expression domain, antisense RNA, double-stranded RNA
25 (dsRNA), small-interfering RNA (siRNA), ribozyme, and the like, or combinations thereof. For example, see FIG. 37B, 38B.

 According to some embodiments, the exogenous RNA of interest described above is located at the third sequence.

 According to further embodiments, the exogenous RNA of interest described above
30 may further comprise:

(a) a sequence encoding an exogenous protein of interest; and

(b) an inhibitory sequence that is capable of inhibiting the expression of the exogenous protein of interest;

such that the specific target/cleavage site is located between the inhibitory sequence and the sequence encoding the protein of interest, such that, following introduction of the composition into a cell comprising the endogenous signal RNA, the exogenous RNA of interest is transcribed and cleaved at the specific target/cleavage site so that the inhibitory sequence is detached from the sequence encoding the protein of interest and the protein of interest is capable of being expressed.

In another embodiment, the inhibitory sequence described above, may be located upstream from the specific target/cleavage site, such that the inhibitory sequence comprises a plurality of initiation codons, such that each of the initiation codons and the sequence encoding the exogenous protein of interest are not in the same reading frame, such that each of the initiation codons is consisting essentially of 5'-AUG-3', such that at least one of the initiation codons is located within a Kozak sequence.

10. ADDITIONAL EMBODIMENTS OF THE INVENTION

This section defines and describes further embodiments of the composition of the invention that are described in any of the previous embodiments in any of the previous sections.

The endogenous signal RNA may be, for example, but is not limited to: viral RNA, cellular RNA, such as, for example, mRNA, and the like, that comprises the predetermined signal sequence. The predetermined signal sequence may be, for example, signal sequence that is unique to neoplastic cells, signal sequence that is from viral origin, and the like, or combinations thereof. In some embodiments, the predetermined signal sequence does not comprise any other type of an endogenous RNA molecule (such as, for example, miRNA, shRNA, ribozyme, stRNA, and the like), that is able to direct or effect cleavage of an RNA molecule within the cell.

According to some embodiments, the cell that may be used in embodiments of the invention may be any type of cell from any origin, such as, for example, but not limited to: mammalian cell, avian cell, plant cell, human cell, animal cell, and the like. The cell may be a cultured cell (primary cell or a cell line), or any cell that is present in an organism or a plant.

According to some embodiments, and without wishing to be bound to theory or mechanism, the duplex that is formed when the carrier RNA/sequence is hybridized to the cleaved endogenous signal RNA, such as described, for example in section 1, may be a substrate for a Dicer.

5 In some embodiments, the edge sequence that is described in embodiments in section 1 is 23-28 nucleotides in length and is located from the predetermined cleavage site to about 23-28 nucleotides downstream, such that the second sequence is 2 nucleotides in length and such that the third sequence is 0 nucleotides in length.

10 In another embodiment, the edge sequence that is described in embodiments in section 1 is 25-30 nucleotides in length and is located 2 nucleotides upstream from the predetermined cleavage site and extends upstream in the endogenous signal RNA, such that the second sequence is 0 nucleotides in length and such that the third sequence is 0 nucleotides in length.

15 In additional embodiments, the carrier RNA or the carrier sequence that is described in section 1 or 9 above, may be designed such that the duplex that is formed when the predetermined signal sequence is cleaved, for example, by Dicer, is thermodynamically weaker at the 5' end of the predetermined signal sequence than at the 3' end of the predetermined signal sequence. Such that the strand that is loaded into Risc is the strand that comprises the predetermined signal sequence.

20 The term "sufficient complementarity" may include, but is not limited to: being capable of binding, or at least partially complementary. In some embodiments, the term sufficient complementarity is in the range of about 30-100%. For example, in some embodiments, the term sufficient complementarity is at least about 30% complementarity. For example, in some embodiments, the term sufficient complementarity is at least about 50% complementarity. For example, in some embodiments, the term sufficient complementarity is at least about 70% complementarity. For example, in some embodiments, the term sufficient complementarity is at least about 90% complementarity. For example, in some embodiments, the term sufficient complementarity is about 100% complementarity.

30 In one embodiment of the invention, the expression of the carrier RNA polynucleotide sequence that is described in section 1 may be driven by polymerase I based promoter or polymerase III based promoter. In some embodiments, the expression of the

carrier RNA polynucleotide sequence described in section 1 may be driven by a promoter that may be, but is not limited to: RNA polymerase III 5S promoter, U6 promoter, adenovirus VA1 promoter, Vault promoter, H1 promoter, telomerase RNA or tRNA gene promoter or a functional derivative thereof.

5 The exogenous protein of interest that is described in any of sections 7, 8 or 9 may be any type of protein or peptide. In some embodiments, the exogenous protein of interest may be, for example, but is not limited to: alpha toxin, saporin, maize RIP, barley RIP, wheat RIP, corn RIP, rye RIP, flax RIP, Shiga toxin, Shiga-like RIP, momordin, pokeweed antiviral protein, gelonin, Pseudomonas exotoxin, Pseudomonas exotoxin A or modified
10 forms thereof. In some embodiments, the exogenous protein of interest may be, for example, but is not limited to: Ricin A chain, Abrin A chain, Diphtheria toxin fragment A or modified forms thereof. The exogenous protein of interest may be, for example, but is not limited to, an enzyme (such as, for example, Luciferase), a fluorescent protein, a structural protein, and the like.

15 In some embodiment of the invention, the exogenous protein of interest may be a toxin that may also affect neighboring cells. This toxin may be, for example, but is not limited to: the complete form of: Ricin, Abrin, Diphtheria toxin or modified forms thereof. In another embodiment of the invention, the exogenous protein of interest may be an enzyme that its product can kill also the neighboring cells. Such as enzyme may be, for
20 example, but is not limited to: HSV1 thymidine kinase, such that the composition of the invention further comprises the prodrug – ganciclovir; or Escherichia coli cytosine deaminase, such that the composition of the invention further comprises the prodrug - 5-fluorocytosine (5-FC).

 In another embodiment of the invention, the exogenous RNA of interest or the
25 intermediate RNA that is described in any of sections 7, 8 or 9 is encoded from a viral vector and the exogenous protein of interest is a product of gene that is necessary for the viral vector reproduction, such that the viral vector reproduces in response to the presence of the endogenous signal RNA in a cell and kills the cell during the process of reproduction. This viral vector may also be, for example, but is not limited to a gene that is capable of
30 stopping the viral vector reproduction when a specific molecule is present in the cell (for example, TetR-VP16 / Doxycycline). Hence, when the viral vector is presumed to accumulate enough mutations for reproduction in cells that do not comprise the endogenous

signal RNA, the specific molecule can be administered for stopping all the viral vectors reproduction in the body and then after the degradation of most of the viral vectors in the body cells new viral vectors can be administered again. This viral vector may also comprise, a gene that is capable of killing the cell when a specific prodrug is present (e.g. thymidine kinase / ganciclovir), such that when the viral vector is presumed to accumulate enough mutations for reproduction in cells that do not comprise the endogenous signal RNA the specific prodrug can be administered for killing all the viral vectors in the body and then new viral vectors can be administered again.

In another embodiment, the RNA molecule(s) that are encoded from the compositions of the invention are encoded from a viral vector that is capable of being reproduced in a way that may kill the cell during the process of reproduction. Such that the predetermined signal sequence is not present in, for example, cancer cells, and is present in most of the healthy or nonmetastatic tumourigenic cells of the body of a specific patient and such that the exogenous protein of interest that is described in any of sections 7, 8 or 9 is a toxin that may be, for example, but is not limited to: Ricin A chain, Abrin A chain, Diphtheria toxin fragment A or modified forms thereof. Such that when the viral vector enters a healthy or nonmetastatic tumourigenic cell it may kill the cell and when the viral vector enters a cancer cell it kills the cancer cell during the process of the viral vector reproduction, thus the major concentration of the viral vector is present in the cancer area of the body. This viral vector may also comprise, for example, a gene that is capable of stopping the viral vector reproduction when a specific molecule is present in the cell (e.g. TetR-VP16 / Doxycycline). Such that when the viral vector is presumed to get enough mutations for reproduction in cells that comprise the endogenous signal RNA the specific molecule can be administered for stopping all the viral vectors reproduction in the body and then after the degradation of most of the viral vectors in the body cells new viral vectors can be administered again. This viral vector may also comprise, for example, a gene that is capable of killing the cell when a specific prodrug is present (e.g. thymidine kinase / ganciclovir), such that when the viral vector is presumed to get enough mutations for reproduction in cells that comprise the endogenous signal RNA the specific prodrug can be administered for killing all the viral vectors in the body and then new viral vectors can be administered again.

In another embodiment of the invention, the specific sequence that is located within the exogenous RNA of interest that is described in section 1 or 9 is a plurality of specific sequences and the specific target/cleavage site is a plurality of specific target/cleavage sites. Such that for the exogenous RNA of interest that is described in section 7, wherein said
5 "upstream from the specific cleavage site" also includes upstream from all the cleavage sites and wherein said "downstream from the cleavage site" also includes downstream from all the specific cleavage sites. For example, see FIG. 32A, 32B.

In another embodiment, the specific sequence that is located within the exogenous RNA of interest that is described in section 1 or 9 is one or more specific sequence(s) and
10 the specific target/cleavage site is one or more specific target/cleavage site(s) and the exogenous RNA of interest further comprises: a sequence encoding exogenous protein of interest downstream from the specific target/cleavage site(s), one or more unique sequence(s), such that each of the unique sequence(s) is of sufficient complementarity to the predetermined signal sequence to direct target-specific RNA interference, such that each of
15 the unique sequence(s) is located downstream from the sequence encoding the exogenous protein of interest and 2 inhibitory sequences one at the 5' end of the exogenous RNA of interest and other at the 3' end of the exogenous RNA of interest, such that each of the inhibitory sequences is capable of inhibiting the expression of the exogenous protein of interest. Such that when the endogenous signal RNA is present in a cell, the two inhibitory
20 sequences are detached from the sequence encoding the exogenous protein of interest and the exogenous protein of interest is capable of being expressed. For example, see FIG. 32C.

In another embodiment of the invention, the polynucleotide molecule(s) (such as DNA molecules and/or RNA molecules) of the composition that are described in any of sections 1, 8 or 9 may further comprise a polynucleotide sequence encoding Dicer, or a
25 homologue thereof.

In another embodiment of the invention, the polynucleotide molecule(s) of the composition that are described in section 1 or 9 together further comprise a polynucleotide sequence encoding one or more RISC components, or a homologue thereof.

In another embodiment of the invention, the polynucleotide molecule(s) of the
30 composition that are described in any of sections 1, 8 or 9 may further comprise a polynucleotide sequence encoding one or more RNA molecules that are capable of

unwinding the secondary structure of the endogenous signal RNA at the predetermined signal sequence.

In another embodiment of the invention, the polynucleotide molecule(s) of the composition that are described in any of sections 1, 8 or 9 may further comprise a polynucleotide sequence encoding a special functional RNA that is capable of inhibiting the expression, directly or indirectly, of an endogenous exonuclease. The special functional RNA may be, for example, but is not limited to: microRNA (miRNA), lariat-form RNA, short-hairpin RNA (shRNA), siRNA expression domain, antisense RNA, double-stranded RNA (dsRNA), small-interfering RNA (siRNA) or ribozyme.

The inhibitory sequence that is described in any of the embodiments above may be a sequence or a part of a sequence such that when it is detached from the sequence encoding the exogenous protein of interest, the exogenous protein of interest is capable of being expressed, and when it is not detached from the sequence encoding the exogenous protein of interest, it is capable of inhibiting the expression of the exogenous protein of interest, when it is within its specific context in the exogenous RNA of interest. Such that the inhibitory sequence may also be, only a part of any of the inhibitory sequences that described above within its specific context. For example, instead of an inhibitory sequence that is an out of reading frame 5'-AUG-3', the inhibitory sequence may also be only the A or the 5'-AU-3' part in the context of --UG-3' or --G-3' respectively (in other words, the exogenous RNA of interest comprises an out of reading frame 5'-AUG-3' at the 5' end, however the sequence that will be detached is only the 5'-AU-3' part).

In another embodiment, the carrier RNA that is described in section 1 can also be 14-18 nucleotides long.

In another embodiment, the first sequence and the edge sequence that are described in any of sections 1, 8 or 9 can also be 29-200 nucleotides long, as long as the duplex that is formed when they are hybridized does not activate the PKR in the cell.

In additional embodiment, the cells that are described in any of the previous embodiments in any of the previous sections to which the composition of the invention is inserted/introduced, may further be, cells extract or in vitro mixture that comprises cellular proteins (such as, for example, Dicer, Risc).

In another embodiment of the invention, the exogenous RNA of interest that is described in section 7 may further comprise an RNA localization signal for subcellular localization (including cotranslational import) between the specific target/cleavage site and the sequence encoding for the exogenous protein of interest, such that the inhibitory
5 sequence is capable of inhibiting the function of the RNA localization signal for subcellular localization such that the subcellular localization of the exogenous RNA of interest is necessary for the proper expression of the protein of interest. For example, see FIG. 39A, 39B.

In another embodiment of the invention, the inhibitory sequence that is described in
10 section 7 comprises an initiation codon upstream from the specific target/cleavage site, such that the initiation codon is consisting essentially of 5'-AUG-3', such that the inhibitory sequence further comprises a nucleotide sequence encoding an amino acid sequence immediately downstream from the initiation codon, such that the nucleotide sequence and the sequence encoding the exogenous protein of interest are in the same reading frame, such
15 that the amino acid sequence is capable of inhibiting the function of the sorting signal for subcellular localization of the exogenous protein of interest and such that the subcellular localization of the exogenous protein of interest is necessary for its proper expression. For example, see FIG. 39C.

In another embodiment of the invention, the exogenous RNA of interest that is
20 described in section 7 does not comprise a stop codon downstream from the start codon of the sequence encoding the exogenous protein of interest, such that the inhibitory sequence is located downstream from the sequence encoding the exogenous protein of interest within the exogenous RNA of interest, such that the inhibitory sequence and the sequence encoding the exogenous protein of interest are in the same reading frame, such that the inhibitory
25 sequence encodes an amino acid sequence that is selected from the group consisting of:

- (a) an amino acid sequence that is capable of inhibiting the function/activity of the exogenous protein of interest;
- (b) an amino acid sequence that is a sorting signal for subcellular localization;
- (c) an amino acid sequence that is a protein degradation signal;
- 30 (d) an amino acid sequence that is capable of inhibiting the function of the sorting signal for subcellular localization of the protein of interest; and

(e) an amino acid sequence that is capable of inhibiting the cleavage of a peptide sequence that is encoded by a nucleotide sequence that is located between the specific target/cleavage site and the start codon of the sequence encoding the exogenous protein of interest, such that the nucleotide sequence and the sequence encoding the exogenous protein of interest are in the same reading frame and such that the peptide sequence is capable of being cleaved by a protease in a mammalian cell. It has been previously reported that in the human cell during translation of truncated mRNA without stop codon(s), the ribosome stalls at the terminal codon and the cognate tRNA molecule remains bound to the polypeptide chain and to the ribosome, however, it is possible for a peptidyl-tRNA species, in the midst of translation, to be processed by the endoplasmic reticulum signal peptidase [37]. For example, see FIG. 39D

11. PREPARATION OF THE COMPOSITION OF THE INVENTION

In one embodiment of the invention, the exogenous RNA of interest and the functional RNA, that are described in embodiments in section 1 are capable of being located on the same or different RNA molecules.

In another embodiment of the invention, the exogenous RNA of interest and/or the functional RNA, that are described in embodiments in section 1 are capable of being located within a third sequence.

In another embodiment of the invention, the exogenous RNA of interest, the functional RNA, the carrier RNA and the functional nucleic acid, that are described in embodiments in section 1 are capable of being located on one or more RNA molecules.

In some embodiments of the invention, the one or more polynucleotide molecule(s) described in any of the previous embodiments in any of the previous sections comprises one or more DNA molecule. In some embodiments, the one or more DNA molecule(s) are present in one or more DNA vectors (such as, for example, expression vectors), and/or viral vectors.

The polynucleotide molecule(s) (such as DNA molecules and/or RNA molecules) of the composition of the invention may be recombinantly engineered, by any of the methods known in the art, into a variety of host vector systems that may also provide for replication of the polynucleotide molecule(s) in large scale and which contain the necessary elements for directing the transcription of the RNA molecule(s) that are encoded from the

composition of the invention. The use of such vectors to transfect target cells in the patient may result in the transcription of sufficient amounts of the RNA molecule(s) that are encoded from the composition of the invention. For example, a vector can be introduced in vivo such that it is taken up by a cell and directs the transcription of these RNA molecule(s).

5 Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired RNA molecule(s) that encoded from the composition of the invention. Such vectors can be constructed by recombinant DNA technology methods well known in the art or can be prepared by any method known in the art for the synthesis of DNA molecules.

10 The recombinant polynucleotide constructs (such as, for example, recombinant DNA constructs), that encode for the RNA molecule(s) which are encoded from the composition of the invention may be, for example, a plasmid, vector, viral construct, or others known in the art, used for replication and expression in the appropriate target cell (which may be, for example, mammalian cells). Expression of these RNA molecule(s) can be regulated by any
15 promoter known in the art to act in the target cell (such as, for example, mammalian cells, which include, for example, human cells). Such promoters can be inducible or constitutive. Such promoters include, for example, but are not limited to: the SV40 early promoter region, the promoter contained in the 3' long terminal repeat of Rous sarcoma virus, the herpes thymidine kinase promoter, the regulatory sequences of the metallothionein gene, the
20 viral CMV promoter, the human chorionic gonadotropin-beta promoter, and the like. In some embodiments, the promoter may be an RNA Polymerase I promoter (i.e., a promoter that is recognized by RNA Pol. I), such as, for example, the promoter of ribosomal DNA (rDNA) gene. In such embodiments, the termination signal of the exogenous RNA of interest molecule may be a RNA Pol. I termination signal or a RNA polymerase II
25 termination signal (such as, for example, a polyA signal). Any type of plasmid, cosmid, YAC or viral vector can be used to prepare the recombinant polynucleotide constructs which can be introduced directly into the target tissue/cell site. Alternatively, viral vectors can be used which selectively infect the desired target cell.

For the formation of transgenic organism that is resistant to viral infection, it is
30 desirable that the vector that encodes for the RNA molecule(s) that are encoded from the composition of the invention will have a selectable marker. A number of selection systems can be used, including but not limited to selection for expression of the herpes simplex virus

thymidine kinase, hypoxanthine-guanine phosphoribosyltransferase and adenine phosphoribosyl transferase protein in tk-, hgp^{rt}- or ap^{rt}-deficient cells, respectively. Also, anti-metabolic resistance can be used as the basis of selection for dihydrofolate transferase (dhfr), which confers resistance to methotrexate; xanthine-guanine phosphoribosyl transferase (gpt), which confers resistance to mycophenolic acid; neomycin (neo), which
5 confers resistance to aminoglycoside G-418; and hygromycin B phosphotransferase (hygro) which confers resistance to hygromycin.

Vectors for use in the practice of the invention include any eukaryotic expression vectors. In some embodiments of the invention, the RNA molecule(s) that are encoded from
10 the composition of the invention are encoded by a viral expression vector. The viral expression vector may be, for example, but is not limited to those belonging to a family of: Herpesviridae, Poxviridae, Adenoviridae, Papillomaviridae, Parvoviridae, Hepadnaviridae, Retroviridae, Reoviridae, Filoviridae, Paramyxoviridae, Pneumoviridae, Rhabdoviridae, Orthomyxoviridae, Bunyaviridae, Hantaviridae, Picornaviridae, Caliciviridae, Togaviridae,
15 Flaviviridae, Arenaviridae, Coronaviridae, or Hepaciviridae. The viral expression vector may also include, but is not limited to an adenoviral vector that its cellular tropism has been modified by the replacement of the adenovirus terminal knob domain of the fiber protein (HI loop), which is exposed at the fiber surface.

In another embodiment of the invention, the composition of the invention may
20 comprise the RNA molecule(s) that is encoded from this composition, or derivatives or modified versions thereof, single-stranded or double-stranded. These RNA molecule(s) that are encoded from the composition of the invention may be, for example, but are not limited to deoxyribonucleotides, ribonucleosides, phosphodiester linkages, modified linkages or bases other than the five biologically occurring bases (adenine, guanine, thymine, cytosine
25 and uracil).

The RNA molecule(s) that are encoded from the composition of the invention can be prepared by any method known in the art for the synthesis of RNA molecules. For example, these RNA molecule(s) may be chemically synthesized using commercially available reagents and synthesizers by methods that are well known in the art. Alternatively, these
30 RNA molecule(s) can be generated by in vitro and in vivo transcription of DNA sequences that encoding these RNA molecule(s). Such DNA sequences can be incorporated into a wide variety of vectors which incorporate suitable RNA polymerase promoters such as the

T7 or SP6 polymerase promoters. These RNA molecule(s) may be produced in high yield via in vitro transcription using plasmids such as SPS65. In addition, RNA amplification methods such as Q-beta amplification can be utilized to produce these RNA molecule(s).

The polynucleotide molecules, such as, the DNA molecule(s) and/or the RNA molecule(s), and/or the RNA molecules encoded by the composition can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, in order to improve stability of the molecule, hybridization, transport into the cell, etc. In addition, modifications can be made to reduce susceptibility to nuclease degradation. The polynucleotide molecule(s) of the composition of the invention and/or the RNA molecules encoded by the composition may include any other appended groups such as, for example, peptides (for example, for targeting host cell receptors in vivo), or agents facilitating transport across the cell membrane or the blood-brain barrier, hybridization-triggered cleavage agents or intercalating agents. Various other well known modifications can be introduced as a means of increasing intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences of ribo- or deoxy-nucleotides to the 5' and/or 3' ends of the molecule. In some circumstances where increased stability is desired, nucleic acids having modified internucleoside linkages such as 2'-O-methylation may be preferred. Nucleic acids containing modified internucleoside linkages may be synthesized using reagents and methods that are well known in the art.

The polynucleotide molecule(s) of the composition and/or the RNA molecule(s) that encoded from the composition of the invention may be purified by any suitable means, as are well known in the art (for example, reverse phase chromatography or gel electrophoresis).

Cells that produce viral vectors that together encode the RNA molecule(s) that are encoded from the composition of the invention, can also be used for transplantation in a patient for continuous treatment. These cells can further carry a specific gene that can kill them if a specific molecule is introduced to the patient's circulating system (for example: HSV1 Thymidine kinase / Ganciclovir).

In one embodiment, each one of the RNA molecule(s) that are encoded from the composition of the invention can be an RNA molecule or a reproducing RNA molecule. Such that the reproducing RNA molecule is an RNA molecule that comprises a sequence that is complementary to any of these RNA molecule(s) such that the reproducing RNA

molecule is capable of being replicated in the cell for the formation of any of these RNA molecule(s).

In another embodiment, each of the RNA molecule(s) that are encoded from the composition of the invention can be prepared from various types, including, but are not limited to: synthetic RNA, synthetic RNA with modified bases, RNA that is produced by in vitro transcription, DNA molecule that encodes the RNA molecule, vector or viral vector that encodes the RNA molecule or DNA with modified bases that encodes the RNA molecule. For example, the functional RNA can be a synthetic siRNA while the exogenous RNA of interest can be encoded from a viral vector and while the Carrier RNA can be encoded from a plasmid.

12. USES AND ADMINISTRATION OF THE COMPOSITION OF THE INVENTION

The composition of the present invention may be used in various applications including, but not limited to: regulation of gene expression, targeted cell death, treatment, and/or prevention of various diseases and health related conditions (such as, for example, proliferative disorders (for example, cancer), infectious diseases and the like), diagnosis of various health related conditions, formation of transgenic organisms, suicide gene therapy, and the like. In one exemplary embodiment, the composition of the present invention can be used to activate toxic gene in cells that comprise viral RNA, in order to kill these cells. In another exemplary embodiment, the composition of the present invention can be used to activate toxic gene in cells that include an endogenous mRNA which comprises a predetermined signal sequence that is unique to cancer cells, for the targeted and specific killing of these cells.

According to some embodiments, there is thus provided a method for killing a specific cell/cell population, wherein the cell population comprises an endogenous signal RNA, comprising a predetermined signal sequence, which is unique and specific for these cells; the method includes introducing the cells with the composition of the invention, wherein the composition comprises one or more polynucleotides for directing the specific cleavage of an exogenous RNA of interest at a specific target site that is located within a specific sequence, which is of sufficient complementarity to hybridize with the predetermined signal sequence, wherein the cleavage of the exogenous RNA of interest leads to the expression of an exogenous protein of interest, capable of killing the cells.

According to some embodiments, the exogenous protein of interest may be selected from, but not limited to: any type of protein that can damage the cell function and as a result lead to the death of the cell. The protein may be selected from such types of proteins as, but not limited to: toxins, cell growth inhibitors, modulators of cellular growth, inhibitors of cellular signaling pathways, modulators of cellular signaling pathways, modulators of cell permeability, modulators of cellular processes, and the like.

According to some embodiments, and without wishing to be bound to theory or mechanism, the composition and methods of the present invention may provide a specific and targeted "all or none" response in a cell. In other words, compositions and methods of the present invention are such that the exogenous RNA of interest is cleaved (and consequently, the protein of interest is expressed and activated) only in those target cells, which include a specific endogenous signal RNA, whereas cells that do not include the endogenous signal RNA will not be effected by the composition of the invention. The composition and methods of the present invention may thus provide enhanced safety and control, since no leakiness in the expression of the exogenous protein of interest is observed in cells which do not include the endogenous signal RNA, which comprises the predetermined signal sequence.

In further embodiments, the composition of the present invention can be used to activate reporter gene in the presence of viral RNA for the diagnosis of viral infection diseases. In another embodiment, the composition of the present invention may be used to stably transfect cells for the formation of transgenic organism that is resistant to viral infection. In another embodiment, the composition of the present invention can be used to stably transfect cells for the formation of transgenic organism that is able to activate reporter gene in the presence of viral RNA for the diagnosis of viral infection diseases. In yet another embodiment, the composition of the present invention can be used to monitor in real time the changes in RNA sequence in the cell.

Various delivery systems and methods are well known in the art, which can be used to transfer/introduce/transfect the composition of the invention into target cells. The delivery systems and methods include, for example, use of various transfecting agents, encapsulation in liposomes, microparticles, microcapsules, recombinant cells that are capable of expressing the composition, receptor-mediated endocytosis, construction of the composition of the invention as part of a viral vector or other vector, viral vectors that are

capable of being reproduced without killing the cell during the process of reproduction and that comprise the composition of the invention, viral vectors that are not capable of reproduction and that comprise the composition of the invention, injection of cells that produce viral vectors that comprise the composition of the invention, injection of DNA, electroporation, calcium phosphate mediated transfection, and the like, or any other suitable delivery system known or to be developed in the future.

In some embodiments, the present invention also provides for pharmaceutical compositions comprising an effective amount of the composition of the invention, and a pharmaceutically acceptable carrier. The term "Pharmaceutically acceptable" means approved by a regulatory agency of the Federal or a state government or listed in the U.S. Pharmacopeia or other generally recognized pharmacopeia for use in animals, and more particularly in humans. The term "Carrier" in the phrase "Pharmaceutically acceptable carrier" refers to a diluent, adjuvant, excipient, or vehicle with which the therapeutic is administered.

In some embodiments, the pharmaceutical compositions of the invention may be administered locally to the target area in need of treatment. This may be achieved by, for example, and not limited to: local infusion during surgery, topical application, e.g. in conjunction with a wound dressing after surgery, by injection, by means of a catheter, by means of a suppository, or by means of an implant, said implant being of a porous, non-porous, or gelatinous material, including membranes, such as sialastic membranes, or fibers. The local administration may be also achieved by control release drug delivery systems, such as nanoparticles, matrices such as controlled-release polymers or hydrogels.

In some embodiments, the composition of the invention can be administered in amounts which are effective to produce the desired effect in the targeted cell. Effective dosages of the composition of the invention can be determined through procedures well known to those in the art which address such parameters as biological half-life, bioavailability and toxicity. The amount of the composition of the invention which is effective depend on the nature of the disease or disorder being treated, and may be determined by standard clinical techniques. In addition, in vitro assays may optionally be employed to help identify optimal dosage ranges. The administered means may also include, but are not limited to permanent or continuous injection of the composition of the invention to the patient blood stream.

According to some embodiments, the present invention also provides for pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention optionally associated with such container(s) can be a notice in the form prescribed by a governmental agency
5 regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human or animal administration.

EXAMPLES

10 The following examples are offered by way of illustration and not by way of limitation and are examples of embodiments of the present invention.

EXAMPLE 1: use of the composition of the invention to kill the cancer cells of a specific patient

15 According to the American Cancer Society, 7.6 million people died from cancer in the world during 2007.

In this example the composition of the invention is designed to specifically kill the cancer cells of a specific patient. The first step in the designing of the composition of the invention for a specific patient is to identify the predetermined signal sequence, which is a sequence of 18-25 nucleotides long of an endogenous RNA molecule that is present in the
20 cancer cells of this specific patient, such that the predetermined signal sequence is not present in any endogenous RNA molecule in the healthy or nonmetastatic tumourigenic cells of the body of the specific patient. Therefore, the predetermined signal sequence is an RNA sequence of a gene that is mutated in the cancer cells. On average each tumor comprises mutations in 90 protein-coding genes [16] and each tumor initiated from a single
25 founder cell [38], therefore there is a need to identify only one of them that it is transcribed into RNA molecule in the cancer cells.

Various methods can be used for the identification of this predetermined signal sequence; these methods include, but are not limited to DNA microarray, Tilling (Targeting Induced Local Lesions in Genomes) and large-scale sequencing of cancer genomes.
30 Furthermore the identification of the predetermined signal sequence can utilize the Cancer

Genome Atlas project that has been cataloguing all the genetic mutations responsible for cancer by their genes.

In this example, the predetermined signal sequence that is unique to the cancer cells of the specific patient is: 5'-UAUUAUUAUCUUGGCCGCCCCG-3' (SEQ ID NO. 41) and is located in an endogenous mRNA (SEQ ID NO. 42). Therefore, in this example the composition of the invention is designed to kill cells that comprise mRNA that comprises the sequence 5'-UAUUAUUAUCUUGGCCGCCCCG-3' (SEQ ID NO. 41). The functional RNA in this example (SEQ ID NO. 43) is shRNA that is designed to effect the cleavage of the 5' end of the predetermined signal sequence. The sequence of the cleaved shRNA portion formed after processing by Dicer that hybridizes with the endogenous mRNA is set forth as SEQ ID NO. 44. The functional RNA is transcribed under the control of the very strong U6 promoter of RNA polymerase III. The G at the 5' end and the UU at the 3' end of the shRNA are necessary for the transcription by U6 promoter of RNA polymerase III. For example, see FIG. 40. It has been reported that in the cell the functional half-life of each of the two portions of a cleaved mRNA is reduced from the intact mRNA only by 2.6-1.7 fold [10]. It has also been reported that two portions of an mRNA that has been cleaved by RISC-RNA complex in a cell can be easily detected by Northern analysis [6].

The carrier RNA in this example is designed to be transcribed under the control of the very strong U6 promoter of RNA polymerase III and is designed to include the sequence: 3'-UUAUAAUAAUAGAACCGCGGGCGGUG-5' (SEQ ID NO. 45), the G at the 5' end and the UU at the 3' end of the carrier RNA are necessary for the transcription by U6 promoter of RNA polymerase III. In the cell, the carrier RNA is hybridized to the cleaved mRNA portion that includes the predetermined signal sequence and after processing by Dicer (SEQ ID NO. 46), the duplex that is formed is thermodynamically weaker at the 5' end of the predetermined signal sequence, thus the predetermined signal sequence is the strand that is loaded into Risc [3]. For example, see FIG. 40. The sequence of the cleaved carrier RNA portion formed after processing by Dicer is set forth as SEQ ID NO. 47.

It has been reported that in the cell, two RNA transcripts of about 23 nucleotides in length that have a complementary region of about 19 nucleotides in length at the 5' end are hybridized with each other and are capable of directing target specific RNA interference [7]. It has also been reported that a dsRNA of 52 nucleotides long that further comprises 20

nucleotides long ssRNA at one of the 3' ends is a substrate for a Dicer at the blunt end [8]. Furthermore it has also been reported that in mammal Risc is coupled to Dicer [9].

The specific sequence in the exogenous RNA of interest of this example is designed to comprise the sequence: 5'-CGGGCGGCCAAGAUAAUAAUA-3' (SEQ ID NO. 48) that is 100% complementary to the predetermined signal sequence. For example, see FIG. 40. The exogenous RNA of interest is also designed to comprise a sequence encoding Diphtheria toxin fragment A (DT-A) downstream from the specific sequence. The exogenous RNA of interest is designed to be transcribed under the control of the strong viral CMV promoter. For example, see FIG. 40.

It has been reported that a single molecule of Diphtheria toxin fragment A introduced into a cell can kill the cell [14] and in mammal cells, the removal of a cap reduces translation of mRNA by 35-50 fold and reduces the functional mRNA half-life only by 1.7-fold [10].

The exogenous RNA of interest is also designed to comprise an inhibitory sequence upstream from the specific sequence, the inhibitory sequence comprises 3 initiation codons that 2 of them are located within the human Kozak consensus sequence: 5'-ACCAUGG-3' (SEQ ID NO. 25) and each one of them is not in the same reading frame with the start codon of DT-A. For example, see FIG. 40.

The exogenous RNA of interest is also designed to comprise the very efficient cis-acting hammerhead ribozyme – N117 [23] at the 5' end for reducing the efficiency of translation of the exogenous RNA of interest of the invention before it is cleaved. The cis-acting hammerhead ribozyme – N117 also comprises 2 initiation codons however each one of them is not in the same reading frame with the start codon of DT-A. For example, see FIG. 40. The entire sequence of the exogenous RNA of this example is set forth as SEQ ID NO. 49.

In this example, the functional RNA, the carrier RNA and the exogenous RNA of interest are transcribed by a viral vector. For example, see FIG. 40.

Such that, in the cell, the viral vector transcribes: the functional RNA, the carrier RNA and the exogenous RNA of interest. The cis acting ribozyme N117 in the exogenous RNA of interest removes the CAP structure from the 5' end for reducing any translation by the exogenous RNA of interest and the out of reading frame initiation codons prevent

translation of DT-A. The functional RNA (shRNA) effects the cleavage of the 5' end of the predetermined signal sequence. The carrier RNA is hybridized to the cleaved mRNA portion that comprises the predetermined signal sequence and the predetermined signal sequence is processed by Dicer and loaded into Risc. The Risc-signal sequence complex
5 cleaves the exogenous RNA of interest at the specific sequence and the out of reading frame initiation codons are detached, so that DT-A is expressed at least one time, which enough to cause cell death. For example, see FIG. 40. The sequence of the cleaved exogenous RNA of this example is set forth as SEQ ID NO. 50.

10 **EXAMPLE 2:** use of the composition of the invention to kill EBV-associated gastric carcinomas cancer cells, nasopharyngeal carcinoma cancer cells, Burkitt's lymphoma cancer cells and Hodgkin's lymphoma cancer cells

Epstein-Barr virus (EBV) is a ubiquitous human gammaherpesvirus that establishes life-long latent infections in B lymphocytes following the primary infection. EBV infects
15 the majority of the population worldwide and has been implicated in the pathogenesis of several human malignancies including Burkitt's and Hodgkin's lymphomas, gastric carcinoma and nasopharyngeal carcinoma (NPC) [32]. EBV infection is mainly characterized by the expression of latent genes including EBNA1, LMP1, LMP2 and EBER [32]. LMP1 (latent membrane protein 1) was the first EBV latent gene found to be able to
20 transform cell lines and alter the phenotype of cells due to its oncogenic potential [32]. In human epithelial cells, LMP1 alters many functional properties that are involved in tumor progression and invasions [32].

In this example the composition of the invention is designed to kill cancer cells of Burkitt's lymphomas, Hodgkin's lymphomas, gastric carcinoma and nasopharyngeal
25 carcinoma, which are latently infected with EBV, by using the LMP1 mRNA as the endogenous signal RNA and by using the sequence: 5'-CUCUGUCCACUUGGAGCCCUU-3' (SEQ ID NO. 51 - nucleotides 269-289 of LMP1 mRNA) as the predetermined signal sequence. For example, see FIG. 41. Nucleotides 255-304 of LMP1 mRNA are also shown in the figure and set forth as SEQ ID NO. 52. This
30 predetermined signal sequence is chosen because it is located in a region that does not have RNA secondary structure and because it is located in a region that has been shown to be a

good target for siRNA [33]. Furthermore, this predetermined signal sequence is also chosen because its cleavage creates a relatively short RNA molecule of 289 nucleotides long.

In this example, the carrier sequence and the functional RNA are located in the same RNA duplex that is hybridized in the cell, such that the double strand region is located at the 5' end of the carrier sequence and such that when the double strand region is processed by Dicer, the carrier sequence is detached from the RNA duplex and the siRNA duplex that is formed is the functional RNA and is capable of effecting the cleavage of the mRNA of LMP-1 at the 3' end of the predetermined signal sequence. The 2 strands of the RNA duplex are:

10 UUCUCUGGAAGAGACAGGUGAACCUCGGGAACCUCGGGAAACAUAUGAGG-5' (SEQ ID NO. 53) and 5'-GGAGCCCUUUGUAUACUCCUU-3' (SEQ ID NO. 54). The 2 strands of the RNA duplex are transcribed under the control of the very strong U6 promoter of RNA polymerase III, thus their 5' end is G and their 3' end is UU. For example, see FIG. 41. The sequence of the cleaved strand (after Dicer processing) capable of hybridizing to the mRNA of LMP-1 and affecting its cleavage at the 3' end of the predetermined signal sequence is set forth as SEQ ID NO. 55.

When the mRNA of LMP-1 is cleaved at the 3' end of the predetermined signal sequence, the carrier sequence (SEQ ID NO. 56) directs the predetermined signal sequence to Dicer processing and the duplex that is formed is thermodynamically weaker at the 5' end of the predetermined signal sequence, thus the predetermined signal sequence will be the strand that will be loaded into Risc [3]. For example, see FIG. 41. The sequence of the second strand, namely the cleaved carrier sequence after processing by Dicer is set forth as SEQ ID NO. 57.

The specific sequence in the exogenous RNA of interest of the example is designed to comprise the sequence: 3'-GAGACAGGUGAACCUCGGGAA-5' (SEQ ID NO. 58) that is 100% complementary to the predetermined signal sequence. The exogenous RNA of interest is also designed to comprise a sequence encoding Diphtheria toxin (DT) downstream from the specific sequence. The exogenous RNA of interest is designed to be transcribed under the control of the strong viral CMV promoter. The exogenous RNA of interest is also designed to comprise an inhibitory sequence upstream from the specific sequence. The inhibitory sequence comprises 2 initiation codons that are located within the human Kozak consensus sequence: 5'-ACCAUGG-3' (SEQ ID NO. 25) and each one of

them is not in the same reading frame with the start codon of DT. The exogenous RNA of interest is also designed to comprise the very efficient cis-acting hammerhead ribozyme – snorbozyme [22] at the 5' end for reducing the efficiency of translation of the exogenous RNA of interest of the invention before it is cleaved. The cis-acting hammerhead ribozyme – snorbozyme also comprises 2 initiation codons however each one of them is not in the same reading frame with the start codon of DT. The exogenous RNA of interest is also designed to comprise a nucleotide sequence of 23 nucleotides downstream from the specific sequence and upstream from the sequence encoding DT, such that the nucleotide sequence is capable of binding to a sequence of 23 nucleotides that is located downstream from the sequence encoding DT, such that the exogenous RNA of interest forms a circular structure that increases the efficiency of translation of DT particularly when the exogenous RNA of interest is cleaved. For example, see FIG.41. The entire sequence of the exogenous RNA of this example is set forth as SEQ ID NO. 59.

In this example, the two strands of the RNA duplex and the exogenous RNA of interest are transcribed by a viral vector (see FIG. 41). Such that, in the cell, the viral vector transcribes: the two strands of the RNA duplex and the exogenous RNA of interest. The cis acting ribozyme, snorbozyme, in the exogenous RNA of interest removes the CAP structure from the 5' end for reducing any translation by the exogenous RNA of interest and the out of reading frame initiation codons prevent translation of DT. The two strands of the RNA duplex are hybridized with each other and with the predetermined signal sequence at the LMP-1 mRNA, the double strand region of the RNA duplex is cleaved by Dicer and forms the functional RNA that is siRNA and the carrier sequence. The siRNA cleaves the predetermined signal sequence at the 3' end and the carrier sequence directs the cleaved predetermined signal sequence to Dicer processing. The processed predetermined signal sequence is loaded into Risc and then the Risc-signal sequence complex cleaves the exogenous RNA of interest at the specific sequence and the out of reading frame initiation codons are detached, so that DT is capable of being expressed. The sequence of the cleaved exogenous RNA of this example is set forth as SEQ ID NO. 60. The RNA portion that comprises the sequence encoding DT forms a circular structure that increases DT translation for killing the cancer cell and the neighboring cells. For example, see FIG. 41.

In this example, the functional RNA and the carrier sequence are located in the same RNA duplex, thus the carrier sequence may bring the functional RNA into proximity with

the predetermined signal sequence and by this may also bring the components of the RNA interference pathway (e.g. Dicer and Risc) into proximity with the predetermined signal sequence.

5 **EXAMPLE 3:** use of the composition of the invention to kill HIV-1 infected cells

According to the World Health Organization in 2006 there were about 39.5 million people with HIV worldwide. According to current estimates of the Joint United Nations Program on HIV and AIDS, HIV is set to infect 90 million people in Africa. HIV (Human immunodeficiency virus) can lead to the acquired immunodeficiency syndrome (AIDS).
10 Two species of HIV infect humans: HIV-1 and HIV-2. HIV-1 is more virulent, relatively easily transmitted, and is the cause of the majority of HIV infections globally. HIV-2 is less transmittable than HIV-1 and is largely confined to West Africa.

Many viruses, including HIV, exhibit a dormant or latent phase, during which little or no protein synthesis is conducted. The viral infection is essentially invisible to the
15 immune system during such phases. Current antiviral treatment regimens are largely ineffective at eliminating cellular reservoirs of latent viruses [15].

In this example, the composition of the invention is designed to kill HIV-1 infected cells by using the HIV-1 mRNA as the endogenous signal RNA and by using the sequence: 5'-UACCAAUGCUGCUUGUGCCUG-3' (SEQ ID NO. 61 - nucleotides 8492-8512 of
20 HIV-1 mRNA) as the predetermined signal sequence. For example, see FIG. 42. Nucleotides 8477-8527 of HIV-1 mRNA are also shown in the figure and set forth as SEQ ID NO. 62. This predetermined signal sequence is chosen because it is located in a region that does not include an RNA secondary structure and because it is located in a region that has been shown to be a good target for siRNA [34].

25 The exogenous RNA of interest of this example is transcribed under the control of the strong viral CMV promoter and is designed to comprise 2 specific sequences, such that each one of them is: 3'-AUGGUUACGACGAACACGGAC-5' (SEQ ID NO. 63) that is 100% complementary to the predetermined signal sequence. The exogenous RNA of interest is also designed to comprise a sequence encoding Diphtheria toxin fragment A (DT-
30 A) between the 2 specific sequences. In mammal cells single molecule of Diphtheria toxin fragment A introduced into a cell can kill the cell [14]. The exogenous RNA of interest is

also designed to comprise two inhibitory sequences one at the 5' end and other at the 3' end. The inhibitory sequence that is located at the 5' end of the exogenous RNA of interest is designed to include 3 initiation codons, such that one of them is located within the human Kozak consensus sequence: 5'-ACCAUGG-3' (SEQ ID NO. 25), such that each one of them is not in the same reading frame with the start codon of DT-A and such that all the 3 initiation codons are in the same reading frame. The inhibitory sequence that is located at the 5' end of the exogenous RNA of interest also comprises a nucleotide sequence downstream from the 3 initiation codons and upstream from the 2 specific sequences, such that the nucleotide sequence is in the same reading frame with the 3 initiation codons and such that the nucleotide sequence encodes for a sorting signal for the subcellular localization that is the Peroxisomal targeting signal 2 of the human alkyl dihydroxyacetonephosphate synthase (H₂N---RLRVLSGHL – SEQ ID NO. 27) [30]. In mammal cells proteins that bear a sorting signal for the subcellular localization can be localized to the subcellular localization while they are being translated with their mRNA. For example, see FIG.42.

The inhibitory sequence that is located at the 3' end of the exogenous RNA of interest is designed to include an intron downstream from the 2 specific sequences, such that the exogenous RNA of interest is a target for nonsense-mediated decay (NMD) that degrades RNA molecule that comprises an intron downstream from the coding sequence [31]. The intron comprises 2 artificial microRNAs that are designed to affect the cleavage of the predetermined signal sequence at the 5' end and at the 3' end (SEQ ID NOs. 64 and 65, respectively) [35]. The inhibitory sequence that is located at the 3' end of the exogenous RNA of interest also comprises an AU-rich element (ARE) at the 3' end that stimulates degradation of the exogenous RNA of interest. The AU-rich elements is 47 nucleotides long and it comprises the sequences: 5'-AUUUA-3' (SEQ ID NO. 31) and 5'-UUAUUA(U/A)(U/A)-3'(SEQ ID NO. 32) [28]. For example, see FIG. 42. The entire sequence of the exogenous RNA of this example is composed of SEQ ID NO. 66, SEQ ID NO. 113 and an intron comprising the two artificial microRNAs described above in between. The carrier RNA in this example, is designed to be transcribed under the control of the very strong U6 promoter of RNA polymerase III and is designed to include the sequence: 3'-UUAUGGUUACGACGAACACGG-5' (SEQ ID NO. 67), the G at the 5' end and the UU at the 3' end of the carrier RNA are necessary for the transcription by U6 promoter of RNA polymerase III. In the cell the carrier RNA of the invention may be hybridized to the cleaved predetermined signal sequence and the duplex that is formed is

thermodynamically weaker at the 5' end of the predetermined signal sequence, thus the predetermined signal sequence is the strand that is loaded into Risc [3]. For example, see FIG. 42.

In this example, the carrier RNA and the exogenous RNA of interest are transcribed
 5 by a viral vector. Such that, in the cell, the viral vector transcribe: the carrier RNA and the
 exogenous RNA of interest. The out of reading frame initiation codons prevent translation
 of DT-A, the Peroxisomal targeting signal 2 sends the erroneous protein and the exogenous
 RNA of interest to the peroxisome. The intron targets the exogenous RNA of interest to
 degradation by the nonsense-mediated decay (NMD) and the AU-rich element also
 10 stimulates degradation of the exogenous RNA of interest. In the presence of the HIV-1
 mRNA in the cell the two artificial microRNAs cleave the predetermined signal sequence at
 the 5' end and at the 3' end and the carrier RNA is hybridized to the cleaved predetermined
 signal sequence, and the signal sequence may be loaded into Risc. Then the Risc-signal
 sequence complex may cleave the exogenous RNA of interest at the two specific sequences
 15 and all the inhibitory sequences are detached, so that DT-A is expressed at least one time,
 which enough to cause cell death. For example, see FIG .42. The sequence of the cleaved
 exogenous RNA of this example is set forth as SEQ ID NO. 68.

In this example, the predetermined signal sequence is cleaved from both of its ends
 and thus with the carrier RNA it is a better substrate for Dicer or Risc.

20 The viral vector may also encode transcriptional factors that are capable of
 enhancing the transcription of HIV-1 mRNA in HIV-1 infected cell (e.g. NF- κ B). The viral
 vector may also encode genes that are capable of preventing new HIV-1 particles production
 (e.g. Rev, which prevents HIV-1 mRNA splicing).

25 **EXAMPLE 4:** use of the composition of the invention to kill hsv-1 infected cells

Many viruses, including HSV-1 (herpes simplex virus-1) exhibit a dormant or latent
 phase, during which no protein synthesis is conducted. The viral infection is essentially
 invisible to the immune system during such phases. Current antiviral treatment regimens are
 largely ineffective at eliminating cellular reservoirs of latent viruses [15]. The latency-
 30 associated transcript (*LAT*) of herpes simplex virus-1 (HSV-1) is the only viral gene that is
 expressed during latent infection in neurons. *LAT* inhibits apoptosis and maintains latency

by promoting the survival of infected neurons. No protein product has been attributed to the *LAT* gene.

In this example, the composition of the invention is designed to kill HSV-1 infected cells by using the latency-associated transcript (*LAT*) as the endogenous signal RNA and by
 5 using the sequence: 5'-AAGCGCCGGCCGGCCGCUGGU-3' (SEQ ID NO. 69 - nucleotides 108-128 of the latency-associated transcript - *LAT* of HSV-1) as the predetermined signal sequence. For example, see FIG. 43. Nucleotides 101-140 of HSV-1 *LAT* mRNA are also shown in the figure and set forth as SEQ ID NO. 70. This predetermined signal sequence is chosen because its cleavage creates a relatively short RNA
 10 molecule of 128 nucleotides long. For example, see FIG. 43.

In this example, the carrier sequence and the functional RNA are located in the same stem loop structure (SEQ ID NO. 71) that is transcribed by the RNA polymerase III U6 promoter. Such that when the stem loop structure is processed by Dicer, the carrier sequence (SEQ ID NO. 72) is detached from the stem loop structure and the siRNA duplex that is
 15 formed is the functional RNA, which is capable of effecting the cleavage of *LAT* at the 3' end of the predetermined signal sequence. The sequences of the strands of the siRNA duplex that is formed are set forth as SEQ ID NOs. 73 and 74. The sequence of the cleaved *LAT* portion that comprises the predetermined signal sequence is set forth as SEQ ID NO. 75. The G at the 5' end and the UU at the 3' end of the stem loop structure are necessary for
 20 the transcription by U6 promoter of RNA polymerase III.

In the cell, the carrier sequence is hybridized to the cleaved *LAT* portion that comprises the predetermined signal sequence and after processing by Dicer, the duplex that is formed is thermodynamically weaker at the 5' end of the predetermined signal sequence, thus the predetermined signal sequence will be the strand that will be loaded into Risc [3].
 25 For example, see FIG. 43.

The exogenous RNA of interest of this example is transcribed under the control of the strong viral CMV promoter and is designed to comprise 2 specific sequences, such that each one of them is: 5'-ACCAGCGGCCGGCCGGCGCUU-3' (SEQ ID NO. 76) that is 100% complementary to the predetermined signal sequence. The exogenous RNA of
 30 interest is also designed to comprise a sequence encoding Diphtheria toxin (DT) between the 2 specific sequences. The exogenous RNA of interest is also designed to comprise 2 inhibitory sequences one at the 5' end and other at the 3' end of the exogenous RNA of

interest. The inhibitory sequence that is located at the 5' end of the exogenous RNA of interest is designed to include 3 initiation codons, such that 2 of them are located within the human Kozak consensus sequence: 5'-ACCAUGG-3' (SEQ ID NO. 25), such that each one of them is not in the same reading frame with the start codon of DT. The inhibitory sequence that is located at the 3' end of the exogenous RNA of interest is designed to comprise the translational repressor smaug recognition elements (SRE): 5'-UGGAGCAGAGGCUCUGGCAGCUUUUGCAGCG-3' (SEQ ID NO. 28) downstream from the 2 specific sequences. For example, see FIG. 43. Smaug 1 is encoded in human chromosome 14 and is capable of repressing translation of SRE-containing messengers [26, 27]. Murine Smaug 1 is expressed in the brain and is abundant in synaptoneurosomes, a subcellular region where translation is tightly regulated by synaptic stimulation [26]. The inhibitory sequence that is located at the 3' end of the exogenous RNA of interest also comprises an RNA localization signal for myelinating periphery (A2RE - Nuclear Ribonucleoprotein A2 Response Element): 5'-GCCAAGGAGCCAGAGAGCAUG-3' (SEQ ID NO. 29) at the 3' end [29]. For example, see FIG. 43. A2RE is a *cis*-acting sequence that is located at the 3'-untranslated region of MBP (Myelin basic protein) mRNA and is sufficient and necessary for MBP mRNA transport to the myelinating periphery of oligodendrocytes [29]. The hnRNP (Heterogeneous Nuclear Ribonucleoprotein) A2 binds the A2RE and mediates transport of MBP [29].

The exogenous RNA of interest of this example also comprises a cytoplasmic polyadenylation element (CPE) immediately downstream from the sequence encoding DT. The CPE comprises the sequence 5'-UUUUUUAUU-3' (SEQ ID NO. 38) immediately downstream from the sequence encoding DT and the sequence 5'-UUUUAUU-3' (SEQ ID NO. 39) 91 nucleotides downstream from the sequence encoding DT [25]. For example, see FIG. 43. In mammals, CPEB (cytoplasmic polyadenylation element binding protein) is present in the dendritic layer of the hippocampus (the portion of the brain that is responsible for long-term memory) [36]. In the synapto-dendritic compartment of mammalian hippocampal neurons, CPEB appears to stimulate the translation of α -CaMKII mRNA, which comprises CPE, by polyadenylation-induced translation [36]. The entire sequence of the exogenous RNA of this example is set forth as SEQ ID NO. 77.

In this example the exogenous RNA of interest and the stem loop structure are transcribed by a viral vector. Such that after the transcription of the exogenous RNA of

interest and the stem loop structure, the out of reading frame initiation codons prevent translation of DT, the Smaug1 (translational repressor) binds to the smaug recognition elements (SRE) and inhibits DT translation and the hnRNP A2 binds the A2RE and mediates the transport of the RNA molecule to the myelinating periphery. Correspondingly
5 the stem loop structure is processed by Dicer such that the carrier sequence is detached from the stem loop structure and the siRNA duplex that is formed is the functional RNA, and then the functional RNA effects the cleavage of LAT at the 3' end of the predetermined signal sequence. Then the carrier sequence is hybridized to the LAT portion that comprises the predetermined signal sequence and the predetermined signal sequence is processed by Dicer
10 and loaded into Risc. Then, the Risc-signal sequence complex cleaves the exogenous RNA of interest at the 2 specific sequences and all the inhibitory sequences are detached, so that the CPEB (cytoplasmic polyadenylation element binding protein) binds to the CPE and stimulates the extension of the poly-A tail in the cleaved exogenous RNA of interest, such that DT is capable of being expressed and kills the cell and the neighboring cells. For
15 example, see FIG. 43. The sequence of the cleaved exogenous RNA of this example is set forth as SEQ ID NO. 78.

In this example, the functional RNA and the carrier sequence are located in the same RNA molecule, which requires less transcriptional units. The major advantage of this proximity of the functional RNA and the carrier sequence is that they are created in the
20 same place in the cell and in the same time and also at a constant ratio.

EXAMPLE 5: use of the composition of the invention to kill cancer cells of a specific patient.

In this example the composition of the invention is designed to kill the cancer cells
25 of a specific patient.

As described in Example 1 above, the first step in the designing of the composition of the invention for a specific patient is to identify the predetermined signal sequence, which is a sequence of 18-25 nucleotides long of an RNA molecule that present in the cancer cells of this specific patient, such that the predetermined signal sequence is not present in any
30 RNA molecule in the healthy or nonmetastatic tumourigenic cells of the body of this specific patient. Therefore, the predetermined signal sequence is an RNA sequence of a gene that is mutated in the cancer cells. On average each tumor comprises mutations in 90

protein-coding genes [16] and each tumor initiated from a single founder cell [38], therefore there is a need to identify only one of them that it is transcribed into an RNA molecule in the cancer cells. Various methods can be used for the identification of this predetermined signal sequence; these methods include, but are not limited to DNA microarray, Tilling
5 (Targeting Induced Local Lesions in Genomes) and large-scale sequencing of cancer genomes. Furthermore the identification of the predetermined signal sequence can utilize the Cancer Genome Atlas project that has been cataloguing all the genetic mutations responsible for cancer by their genes.

In this example, the predetermined signal sequence that is unique to the cancer cells
10 of the specific patient is: 5'-AAUUAAGUUUAUGAACGGGUC-3' (SEQ ID NO. 79) and is located in an endogenous mRNA. Therefore, in this example the composition of the invention is designed to kill cells that comprise endogenous mRNA (as the endogenous signal RNA) that comprises the sequence 5'-AAUUAAGUUUAUGAACGGGUC-3' (SEQ ID NO. 79). An exemplary endogenous mRNA comprising said predetermined signal
15 sequence is shown in FIG. 44 and set forth as SEQ ID NO. 80.

The functional RNA in this example is Rz-B, a hammerhead-type ribozyme (SEQ ID NO. 81) [21] that is designed to effect the cleavage of the 3' end of the predetermined signal sequence. The sequence of the exemplary endogenous mRNA comprising the predetermined signal sequence after cleavage is set forth as SEQ ID NO. 82. The hammerhead-type
20 ribozyme Rz-B is transcribed under the control of the very strong U6 promoter of RNA polymerase III. The G at the 5' end and the UU at the 3' end of the hammerhead-type ribozyme Rz-B are necessary for the transcription by U6 promoter of RNA polymerase III. For example, see FIG. 44. It has been reported that in the cell the functional half-life of each of the two portions of a cleaved mRNA is reduced from the intact mRNA only by 2.6-1.7
25 fold [10]. It has also been reported that two portions of an mRNA that has been cleaved by RISC-RNA complex in a cell can be easily detected by Northern analysis [6].

The carrier sequence of this example is: 5'-CCCGUUCAUAAACUUAUUAACCGGUC-3' (SEQ ID NO. 83) and 103 contiguous carrier sequences are located in an RNA sequence that is transcribed under the control of the
30 strong viral CMV promoter. Such that Rz-A, a hammerhead-type ribozyme (SEQ ID NO. 84) [21], is designed to effect the cleavage of the 3' end of the carrier sequence that is located at the 5' end of the RNA sequence. The hammerhead-type ribozyme Rz-A is

transcribed under the control of the very strong U6 promoter of RNA polymerase III. The G at the 5' end and the UU at the 3' end of the hammerhead-type ribozyme Rz-A are necessary for the transcription by U6 promoter of RNA polymerase III. For example, see FIG. 44.

5 In the cell, the hammerhead-type ribozyme Rz-A detaches up to 101 perfect carrier sequences from 1 RNA sequence. The detached carrier sequence is hybridized to the cleaved mRNA portion that comprises the predetermined signal sequence and after Dicer processing the duplex that is formed is thermodynamically weaker at the 5' end of the predetermined signal sequence, thus the predetermined signal sequence will be the strand
10 that will be loaded into Risc [3]. For example, see FIG.44. The sequence of the second strand of the duplex that is formed, namely the cleaved carrier sequence after processing by Dicer, is set forth as SEQ ID NO. 85.

The specific sequence in the exogenous RNA of interest of the example is designed to comprise the sequence: 3'-UUA AUUCAA AUACUUGCCCAG-5' (SEQ ID NO. 86) that
15 is 100% complementary to the predetermined signal sequence. The exogenous RNA of interest is also designed to comprise a sequence encoding Diphtheria toxin (DT) downstream from the specific sequence. The exogenous RNA of interest is designed to be transcribed under the control of the strong viral CMV promoter. The exogenous RNA of interest is also designed to comprise an inhibitory sequence upstream from the specific
20 sequence. The inhibitory sequence comprises 3 initiation codons that 2 of them are located within the human Kozak consensus sequence: 5'-ACCAUGG-3' (SEQ ID NO. 25) and each one of them is not in the same reading frame with the start codon of DT. The exogenous RNA of interest of the invention also comprises the palindromic termination element (PTE) from the human HIST1H2AC (H2ac) gene 3'UTR (5'-GGCUCUUUCAGAGCC-3' – SEQ
25 ID NO. 34)) downstream from the sequence encoding DT. For example, see FIG. 44. The PTE plays an important role in mRNA processing and stability [11]. Transcripts from HIST1H2AC gene lack poly(A) tails and are still stable, due to the PTE. The entire sequence of the exogenous RNA of this example is set forth as SEQ ID NO. 87.

In this example the exogenous RNA of interest, the hammerhead-type ribozyme Rz-
30 B / Rz-A and the RNA sequence that comprising 103 carrier sequences are transcribed by a viral vector. Such that in the cell, the viral vector transcribes: the exogenous RNA of interest, the hammerhead-type ribozyme Rz-B / Rz-A and the RNA sequence that

comprising 103 carrier sequences. The out of reading frame initiation codons prevent translation of DT. The hammerhead-type ribozyme Rz-B cleaves the 3' end of the predetermined signal sequence. The hammerhead-type ribozyme Rz-A detaches up to 101 perfect carrier sequences from 1 RNA sequence. The detached carrier sequence is
5 hybridized to the cleaved mRNA portion that comprises the predetermined signal sequence and the predetermined signal sequence is processed by Dicer and loaded into Risc. The Risc-signal sequence complex cleaves the exogenous RNA of interest at the specific sequence and the out of reading frame initiation codons are detached and the palindromic termination element stabilizes the cleaved exogenous RNA of interest and protects it from
10 degradation, so that the DT is capable of being expressed and kills the cell and the neighboring cells population. For example, see FIG. 44. The sequence of the cleaved exogenous RNA of this example is set forth as SEQ ID NO. 88.

EXAMPLE 6 – Specific cellular expression of an exogenous protein of interest encoded by an exogenous RNA of interest
15

General protocol for experiments described herein below:

The day before transfection, about 120,000 of T293 cells per well were seeded in 24 well plate, at the day of transfection each well was co-transfected with:

1. Renila/luciferase plasmid - 170ng of plasmid expressing Renilla luciferase gene & firefly luciferase gene (plasmid E11, Psv40-INTRON-MCS-RLuc---Phsvtk-Fluc, SEQ ID NO: 22 or plasmid E65, Psv40-INTRON-Tsp-TD1-TLacZ-RLuc-PTS-60ATG---Phsvtk-FLuc, SEQ ID NO. 23).
20
2. Tested plasmid = 30ng of tested plasmid (as detailed below).
3. siRNA+ or siRNA- = 10 pmole of siRNA double stranded molecule that can induce
25 cleavage (siRNA+) or does not induce cleavage (siRNA-) of the mRNA encoded by the tested plasmid. (detailed below).

The transfection was performed using lipofectamine 2000 transfection reagent (Invitrogen) according to manufacturer protocol. 48 hrs post transfection the Renilla luciferase gene expression was measured using the dual luciferase reported assay kit (Promega) and
30 luminometer (glomax 20/20 promega), and the relative light units (RLU) were determined.

The tested plasmid may be any type of the following plasmids:

Negative control = Plasmid that does not encode for a diphtheria toxin (DTA);

Positive control = Plasmid that constitutively encodes for diphtheria toxin (DTA);

Test plasmid = plasmid of the composition of the invention, i.e. plasmid comprising target sites for siRNA+ between an inhibitory sequence and a downstream sequence encoding for diphtheria toxin (DTA). For the test plasmid, when the co-transfected siRNA+ cleaves the inhibitory sequence of the test plasmid, the diphtheria toxin is capable of being expressed and kills the cells in which it is expressed, thereby – reducing Renilla expression and overall measurement of RLU.

- 10 The tested plasmid was tested with 2 different siRNAs+ and with 2 different siRNAs-, separately, and each in triplicate.

The results are calculated as follows:

- 15 Fold of Activation = Average of measured RLU (Relative light unit) in the presence of each of the 2 siRNA- with the test plasmid (6 wells) divided by the average of RLU using one of the siRNA+ with the test plasmid (3 wells).

Fold of leakage = Average of RLU using all the siRNAs-/+ with the negative control plasmid divided by the Average of RLU using each of the 2 siRNA- with the test plasmid.

siRNA+/-RLU = Average of measured RLU in the presence of one co-transfected siRNA+ or the presence of two co-transfected siRNA-, independently.

- 20 The plasmids were constructed using common and known methods practiced in the art of molecular biology. The backbone vectors for the constructed plasmids described herein below are: psiCHECK™-2 Vectors (promega, Cat. No. C8021) or pcmv6-A-GFP (OriGene, Cat. No. PS100026). The appended name of each plasmid indicates sequences which are comprised within the plasmid sequence, as further detailed below, with respect to
- 25 the test plasmids.

siRNA sequences:

1. RL Duplex (Dharmacon, Cat. No. P-002070-01-20) (SEQ ID NO. 65 (sense strand) and SEQ ID 66 (anti sense strand)).

2. GFPDuplex II (Dharmacon, Cat. No. P-002048-02-20), (SEQ ID NO. 67 (sense strand) and SEQ ID NO.68 (anti sense strand)).
3. siRNA – Control (Sigma, Cat. No., VC30002 000010), (SEQ ID NO. 69 (sense strand) and SEQ ID NO.70, (anti sense strand)).
- 5 4. Anti β Gal siRNA-1 ((target site: TlacZ (SEQ ID NO. 71), Dharmacon, Cat. No. P-002070-01-20) (SEQ ID NO. 72 (sense strand) and SEQ ID NO. 73 (antisense strand)).
5. Luciferase GL3 Duplex ((target site: Tfluc (SEQ ID NO. 74), Dharmacon, Cat. No. D-001400-01-20), (SEQ ID NO. 75 (sense strand) and SEQ ID NO. 76 (antisense strand)).
6. GFPDuplex I ((target site: TD1, (SEQ ID NO. 77), Dharmacon, Cat. No. P-002048-01-20), (SEQ ID NO. 78 (sense strand) and SEQ ID NO. 79 (antisense strand)).
- 10 7. TCTL (target site: TCTL (SEQ ID NO. 80), Dharmacon, SEQ ID NO. 81 (sense strand) and SEQ ID NO. 82 (anti sense strand)).

In each experiment, the siRNA that has target site in the test plasmid is used as siRNA+, and the other siRNAs that do not have a corresponding target site in the tested plasmid was used as siRNA-.

15

Negative control plasmids:

1. E34 (SEQ ID NO. 10) - Pcmv-4ORF⁺-TD1-Tfluc---Psv40-TGFP.
2. E71 (SEQ ID. NO. 17) - Psv40-INTRON-4ORF⁺---Phsvtk-Fluc.
3. E38 - 3CARz-4S&L. The insert of E38 (SEQ ID. NO. 19) was ligated into a PMK shuttle vector (GeneArt) at *pacI* and *XhoI* restriction sites.
- 20

Positive control plasmids:

1. E28 (SEQ ID. NO. 11) - Pcmv-Tfluc-TD1-cDTAWT---Psv40-TGFP.
2. E20 (SEQ ID. NO. 12) - Pcmv-nsDTA---Psv40-TGFP
3. E70 (SEQ ID. NO. 13) - Psv40-INTRON-cDTAWT---Phsvtk-Fluc
- 25 4. E3 (SEQ ID. NO. 14) - Pcmv-KDTA---Psv40-TGFP
5. E89 (SEQ ID. NO. 15) - Pcmv---DT⁺A---Psv40-TGFP
6. E110 (SEQ ID. NO. 16) - Pcmv-D5⁺TA---Psv40-TGFP
7. E4 (SEQ ID. NO. 18) - Pcmv-KDTA---Psv40-Hygro

8. E10 (SEQ ID. NO. 20) - Pef1-DTA24---ZEO::GFP-Pcmv

9. E143 (SEQ ID. NO. 21) - 3PolyA-Prpl19-cDTAWT---Phsvtk-Fluc

Test plasmids

1. **E80** (SEQ ID. NO. 1) - Pcmv-4ORF[^]-TD1-Tfluc-S-cDTAWT---Psv40-TGFP (pCMV
5 promoter (nts. 420-938 of SEQ ID NO. 1); 4ORF[^]= Inhibitory sequence composed of: 9
TISU sequences and 57 kozak sequences, with 57, 57, 36, 36, 21, 21, 21, and 21nt between
adjacent ATG codons, in 4 consecutive ORFs (nt 1027-3547 of SEQ ID NO. 1). The first
ORF (nt. 1031-1651 of SEQ ID NO. 1) is 621nt & is translated from TISU (nt. 1027-1038
of SEQ ID NO. 1), and the next 3ORF[^] (nt. 1662-2996, nt. 2306-2941 and nt 2951-3547 of
10 SEQ ID NO. 1) are translated from Kozak sequence, The last ORF (nt 2951-3547 of SEQ
ID NO. 1) stops before the coding sequence of the wild type DTA (cDTAwT= wt DTA
coding sequence, without promoter/splicing/termination/polyA sites and with kozak
sequence (nt 3568-4155 of SEQ ID NO. 1); followed by TGFP coding sequence under the
control of the SV40 promoter)). The plasmid further comprises target sites TD1 (SEQ ID
15 NO. 77) and Tfluc (SEQ ID NO. 74).

2. **E54** (SEQ ID. NO. 2) - Pcmv-4CARZ-PTS-60ATG[^]-3ORF[^]-TD1-Tfluc-incDTAWT---
Psv40-TGFP (pCMV promoter (nucleotides (nt.) 420-938 of SEQ ID NO. 2); 4CAR = 4 Cis
Acting Ribozyme (nt. 1013-1373 of SEQ ID NO. 2); PTS = Peroxisomal targeting signal
(nt. 1420-1500 of SEQ ID NO. 2); 60ATG[^] = 61 ATG, 46 in Kozak sequence with 53nt
20 between almost every 2 ATG (nt. 1534-4554 of SEQ ID NO. 2) and with stop codons
inside the DTA coding sequence (nt. 6745-7332 of SEQ ID NO. 2); TGFP coding sequence
(nt. 8452-9143 of SEQ ID NO. 2) under the control of the psv40 promoter (nt. 8092-8399 of
SEQ ID. NO. 2)). The plasmid further comprises target sites TD1 (SEQ ID NO. 77) and
Tfluc (SEQ ID NO. 74).

25 3. **E113** (SEQ ID. NO. 3) - Pcmv-4ORF[^]-TD1-Tfluc-PK-D5[^]TA---Psv40-TGFP (pCMV
promoter (nts. 420-938 of SEQ ID NO. 3); 4ORF[^] (nt. 1027-3547 of SEQ ID NO. 3); PK =
pseudoknot – stem and loop, such that the 6nt of the loop are hybridized to the start codon
of DTA (nt 3561-3611 of SEQ ID No. 3); 5[^] = 5 human introns (nts. 3712-3801, 3856-
3960, 4066-4173, 4380-4519 and 4617-4783 of SEQ ID NO. 3) that are located within the
30 coding sequence of the DTA (nts. 3609-3806 of SEQ ID NO. 3) and contain T-rich
sequences for terminating RNA Polymerase 1 and/or 3 transcription, the introns are
embedded in cDTAwT coding sequence; TGFP coding sequence (nts 5906- 6597 of SEQ ID

NO. 3) under the control of the psv40 promoter (nts. 5546-5853 of SEQ ID NO. 3)). The plasmid further comprises target sites TD1 (SEQ ID NO. 77) and Tfluc (SEQ ID NO. 74).

4. **E91** (SEQ ID NO. 4) - Pcmv-4ORF[^]-TD1-Tfluc-DT[^]A---Psv40-TGFP (pCMV promoter (nts. 420-938 of SEQ ID NO. 4), 4ORF[^] (nt. 1027-3507 of SEQ ID NO. 4); DT[^]A - kozak DTA with an intron from Human Collagen 16A1 gene and without promoter/splicing/polyA signal (nt. 3520-4444 of SEQ ID NO. 4); TGFP coding sequence (nt. 5544-6235 of SEQ ID NO. 4) under the control of pSV40 promoter (nt. 5184-5491) The plasmid further comprises target sites TD1 (SEQ ID NO. 77) and Tfluc (SEQ ID NO. 74).
5. **E112** (SEQ ID NO. 5) - Pcmv-4ORF[^]-2xTLacZinINTRON-8X[TCTL+TD1]-PK-D5[^]TA---Psv40-TGFP (pCMV promoter (nts. 420-938 of SEQ ID NO. 5), 4ORF[^] (nt. 1027-3436 of SEQ ID NO. 5); 2xTLacZinINTRON = 2 target of TLacZ in the intron of the commercial plasmid pSELECT-GFPzeo-LacZ (nt. 3438-3638 of SEQ ID NO. 5); 8X[TCTL+TD1] (nt. 3647-4052 of SEQ ID NO. 5); PK = pseudoknot - stem and loop, such that the 6nt of the loop are hybridized to the start codon of DTA (nt 4059-4109 of SEQ ID No. 5); 5[^] = 5 human introns (nts. 4210-4299, 4354-4458, 4564-4671, 4878-5017 and 5115-5281 of SEQ ID NO. 5) that are located within the coding sequence of the DTA (nt. 4107-5304 of SEQ ID NO. 5) and contain T-rich sequences for terminating RNA Polymerase 1 and/or 3 transcription, the introns are embedded in a cDTAwt coding sequence; TGFP coding sequence (nt 6404- 7095 of SEQ ID NO. 5) under the control of the psv40 promoter (nts. 6044- 6351 of SEQ ID NO. 5)). The plasmid further comprises 8 copies of target sites TD1 (SEQ ID NO. 77), TCTL (SEQ ID NO. 80) and 2 copies of TLacZ (SEQ ID NO. 71).

6. **E87** (SEQ ID NO. 6) - Pcmv-4ORF[^]-TD1-3TLacZ-Tctl-BGlob-25G-XRN1S&L-DT[^]A---Psv40-TGFP (pCMV promoter (nts. 420-938 of SEQ ID NO. 6); 4ORF[^] (nt. 1027-3430 of SEQ ID NO. 6); BGlob = beta globin 5' truncated end that is capped (nt. 3577-3655 of SEQ ID NO 6). 25G = a stretch of 25 consecutive G nucleotides (nt. 3660-3684 of SEQ ID NO. 6) that can block/interfere with XRN exoribonuclease enzyme; XRN1S&L = stem and loop structure of the yellow fever virus 3'UTR that can block XRN1 exoribonuclease (nt. 3687-3767of SEQ ID NO. 6). DT[^]A = kozak DTA with an intron from Human Collagen 16A1 gene and without promoter/splicing/polyA signal (nt. 3787-4711of SEQ ID NO. 6); TGFP coding sequence (nt 6404- 7095 of SEQ ID NO. 6) under the control of the psv40 promoter

(nts. 5811-6502 of SEQ ID NO. 6)). The plasmid further comprises TD1 (SEQ ID NO. 77), 3 copies of TLacZ (SEQ ID NO. 71) and TCTL target sites (SEQ ID NO. 80).

7. **E123** (SEQ ID. NO. 7) - Psv40-INTRON-4ORF[^]-3X[TD1-TLacZ]-4PTE-SV40intron-HBB-DTA---Phsvtk-Fluc (pSV40 promoter (nt. 7-419 of SEQ ID NO. 7), 4ORF[^]=9 TISU sequences and 57 kozak sequences, with 57, 57, 36, 36, 21, 21, 21, and 21nt between adjacent ATG codons, in 4 consecutive ORFs (nt 722-2387 of SEQ ID NO. 7); 4PTE = 4 kinds of the stem and loop structures of the Palindromic termination element (nt. 3318-3473 of SEQ ID NO. 7). SV40intron = SV40 small t antigen intron (nt. 3505-3596 of SEQ ID NO. 7); HBB = hemoglobin beta mRNA without ATG and including its first intron (nt. 3627-4406 of SEQ ID NO. 7); cDTAwt coding sequence (nt. 4431-5014 of SEQ ID NO. 7); HSKVK promoter (nt. 5106-5858 of SEQ ID NO. 7) and firefly luciferase coding sequence (nt. 5894-7546 of SEQ ID NO. 7). The plasmid further comprises 3 copies of TD1 (SEQ ID NO. 77) and TLacZ target sites (SEQ ID NO. 71).

8. **E30** (SEQ ID. NO. 8) - Pcmv-4ORF[^]-TD1-Tfluc-incDTAWT---Psv40-TGFP (pCMV promoter (nts. 420-938 of SEQ ID NO. 8); 4ORF[^]=9 TISU sequences and 57 kozak sequences, with 57, 57, 36, 36, 21, 21, 21, and 21nt between adjacent ATG codons, in 4 consecutive ORFs (nt 1027-3547 of SEQ ID NO. 8). The first ORF (nt. 1031-1651 of SEQ ID NO. 8) is translated from TISU (nt. 1027-1038 of SEQ ID NO. 8), and the next 3ORF[^] (nt. 1662-2996, nt. 2306-2941 and nt 2951-3547 of SEQ ID NO. 8) are translated from Kozak sequence, The last ORF (nt 2951-3516 of SEQ ID NO. 8) stops inside the coding sequence of the wild type DTA (cDTAwt= wt DTA coding region, without promoter/splicing/termination/polyA sites and with kozak sequence (nt 3568-4155 of SEQ ID NO. 8); followed by TGFP coding sequence under the control of the SV40 promoter)). The plasmid further comprises target sites TD1 (SEQ ID NO. 77) and Tfluc (SEQ ID NO. 74).

9. **E142** (SEQ ID. NO. 9) - 3PolyA-Prp119-4ORF[^]-TD1-Tfluc-S-cDTAWT---Phsvtk-Fluc. 3PolyA = HSV poly A, SV40 poly A, synthetic poly A (nt. 60-247 of SEQ ID NO. 9); Prp119 = promoter of RPL19 (ribosomal protein L19) taken with its first intron (nt. 248-1941 of SEQ ID NO. 9); 4ORF[^]=9 TISU sequences and 57 kozak sequences, with 57, 57, 36, 36, 21, 21, 21, and 21nt between adjacent ATG codons, in 4 consecutive ORFs (nt 1948-4366 of SEQ ID NO. 9); coding sequence of the wild type DTA (nt. 4457-5044 of SEQ ID NO. 9); HSKVK promoter (nt. 5136-5888 of SEQ ID NO. 9) and firefly luciferase coding

sequence (nt. 5924-7576 of SEQ ID. NO. 9). The plasmid further comprises target sites TD1 (SEQ ID NO. 77) and Tfluc (SEQ ID NO. 74).

Results:

The results are presented in following tables 1-5 and 6A-C. The results show the RLU measured in cells transfected with the indicated plasmids and siRNA molecules under various experimental conditions. The siRNA+ molecules used are the siRNA molecules that can bind their corresponding target sequence(s) within the tested plasmid.

Table 1:

Tested plasmid	Fold of Activation	Fold of leakage	RLU in the presence of siRNA+	RLU in the presence of siRNA-
E34 (SEQ ID NO. 10) - Pcmv-4ORF [^] -TD1-Tfluc---Psv40-TGFP				93M
E28 (SEQ ID NO. 11) - Pcmv-Tfluc-TD1-cDTAWT---Psv40-TGFP				35K
E20 (SEQ ID NO. 12) - Pcmv-nsDTA---Psv40-TGFP				52K
E70 (SEQ ID NO. 13) - Psv40-INTRON-cDTAWT--Phsvtk-Fluc				249K
E54 (SEQ ID. NO. 2) - Pcmv-4CARZ-PTS-60ATG [^] -3ORF [^] -TD1-Tfluc-incDTAWT---Psv40-TGFP	4	5.1	4.4M	18M

10 **Table 2:**

Tested plasmid	Fold of Activation	Fold of leakage	RLU in the presence of siRNA+	RLU in the presence of siRNA-
E34 (SEQ ID NO. 10)- Pcmv-4ORF [^] -TD1-Tfluc---Psv40-TGFP				33M
E28 (SEQ ID NO. 11) - Pcmv-Tfluc-TD1-cDTAWT---Psv40-TGFP				33K
E3 (SEQ ID NO. 14) - Pcmv-KDTA---Psv40-TGFP				45K
E89 (SEQ ID NO. 15) - Pcmv---DT [^] A---Psv40-TGFP				16K
E110 (SEQ ID NO. 16) - Pcmv-D5 [^] TA---Psv40-TGFP				21K
E113 (SEQ ID. NO. 3) - Pcmv-4ORF [^] -TD1-Tfluc-PK-D5 [^] TA---Psv40-TGFP	6	15	367K	2.2M
E80 (SEQ ID. NO. 1) - Pcmv-4ORF [^] -TD1-Tfluc-S-cDTAWT---Psv40-TGFP	5.2	15	427K	2.2M
E91 (SEQ ID. NO. 4) - Pcmv-4ORF [^] -TD1-Tfluc-DT [^] A---Psv40 TGFP	4.73	15	467K	2.2M

E112 (SEQ ID. NO. 5) - Pcmv-4ORF ⁺ -2xTLacZinlNTRON-8X[TCTL+TD1]-PK-D5 ⁺ TA--Psv40-TGFP	4.25	18.3	425K	1.8M
E87 (SEQ ID. NO. 6) - Pcmv-4ORF ⁺ -TD1-3TLacZ-TctI-BGlob-25G-XRN1S&L-DT ⁺ A---Psv40-TGFP	4.15	22	364K	1.5M

Table 3

Tested plasmid	Fold of Activation	Fold of leakage	RLU in the presence of siRNA ⁺	RLU in the presence of siRNA ⁻
E71 (SEQ ID NO. 17) - Psv40-INTRON-4ORF ⁺ ---Phsvtk-Fluc				22.5M
E70 (SEQ ID NO. 3) - Psv40-INTRON-cDTAWT--Phsvtk-Fluc				819K
E123 (SEQ ID. NO. 7) - Psv40-INTRON-4ORF ⁺ -3X[TD1-TLacZ]-4PTE-SV40intron-HBB-DTA---Phsvtk-Fluc	3.37	1.8	3.7M	12.5M

Table 4:

Tested plasmid	Fold of Activation	Fold of leakage	RLU in the presence of siRNA ⁺	RLU in the presence of siRNA ⁻
E34 (SEQ ID NO. 10) - Pcmv-4ORF ⁺ -TD1-Tfluc---Psv40-TGFP				35M
E3 (SEQ ID NO. 14) - Pcmv-KDTA---Psv40-TGFP				47K
E4 (SEQ ID NO. 18) - Pcmv-KDTA---Psv40-Hygro				54K
E30 (SEQ ID. NO. 8) - Pcmv-4ORF ⁺ -TD1-Tfluc-incDTAWT---Psv40-TGFP	2.96	10.9	1.1M	3.2M

5

Table 5:

Tested plasmid	Fold of Activation	Fold of leakage	RLU in the presence of siRNA ⁺	RLU in the presence of siRNA ⁻
E38 (SEQ ID NO. 19) - 3CARz-4S&L				137M
E10 (SEQ ID NO. 20) - PefI-DTA24---ZEO::GFP-Pcmv				55K
E143 (SEQ ID NO. 21) - 3PolyA-Prp119-cDTAWT--Phsvtk-Fluc				132K
E142 (SEQ ID. NO. 9) - 3PolyA-Prp119-4ORF ⁺ -TD1-Tfluc-S-cDTAWT---Phsvtk-Fluc	2.53	5.9	9.1M	23M

Table 6A

Experiment number		1	2	3	4	5	6	7
-------------------	--	---	---	---	---	---	---	---

Number of 293HEK cells per well (24 well plate)	#293cells	135K	180K	150K	120K	150K	120K	90K
Hours post transfection	hrPT	5hr	9hr	48hr	48hr	48hr	48hr	48hr
co-transfection of Renilla expressing plasmid [ng]	REN	E11[170]	E11[170]	E11[170]	E11[170]	E11[170]	E11[170]	E11[170]
co-transfection of siRNA+ or siRNA-: [pico mole]	siRNA	[10]	[10]	[10]	[10]	[10]	[10]	[10]
co-transfection of one of the test plasmids below [ng]: / Results shown below for each plasmid are RLU measured under the indicated experimental condition	↓ / RLU	[30]	[30]	[30]	[30]	[30]	[30]	[30]
Co transfection of a Plasmid comprising the sequence: Pcmv-Tfluc-TD1-cDTAWT---Psv40-TGFP.	E28 (SEQ ID NO. 11)	8.38K	37.89K	81.5K	33K	30.6K	9.8K	7.59K
Co transfection of Plasmid comprising the sequence: Pcmv-4ORF ^Δ -TD1-Tfluc---Psv40-TGFP	E34 (SEQ ID NO. 10)	161K	8.8M	83M	33M	40M	23M	11M
Co transfection of Plasmid comprising the sequence: Pcmv-4ORF ^Δ -TD1-Tfluc-S-cDTAWT---Psv40-TGFP + co-transfected with siRNA-	E80 (SEQ ID. NO. 1)	110K	4.15M	7.17M	2.2M	4.33M	2.3M	1.1M
Co transfection of Plasmid comprising the	E80 (SEQ ID.	33K*	1.35M*	3M*	427K*	1.65M*	800K	354K

sequence Pcmv-4ORF ⁺ - TD1-Tfluc-S- cDTAWT--- Psv40-TGFP co-transfected with siRNA+	NO. 1)							
Fold of activation = RLU measured in the presence of siRNA- divided by RLU measured in the presence of siRNA+	si-/si+	3.33	3	2.4	5.1	2.6	2.87	3.1
Fold of leakiness = E34 (SEQ ID NO. / E80- {smaller than 1 = 0 leakage}	E34 /E80-	1.46	2.12	11.57	15	9.23	10	10

Table 6B

Experiment number		8	9	10	11	12	13	14
Number of 293HEK cells per well (24 well plate)	#293cells	100K	120K	120K	100K	100K	100K	125K
Hours post transfection	hrPT	72hr	48hr	48hr	48hr	48hr	48hr	48hr
co-transfection of Renilla expressing plasmid [ng]	REN	E11[195]	E65[15]	E11[170]	E11[170]	E11[140]	E11[110]	E11[170]
co-transfection of siRNA+ or siRNA-: [pico mole]	siRNA	[10]	[10]	[5.5]	[10]	[10]	[10]	[10]
co-transfection of one of the test plasmids below [ng]: / results shown are RLU under the indicated experimental condition	↓ / RLU	[5]	[30]**	[30]	[30]	[60]	[90]	[30]
Co transfection of Plasmid comprising the	E28 (SEQ ID NO. 11)	128K	2.43K					

sequence: Pcmv-Tfluc-TD1-cDTAWT---Psv40-TGFP								
Co transfection of Plasmid comprising the sequence: Pcmv-4ORF ⁺ -TD1-Tfluc---Psv40-TGFP	E34 (SEQ ID NO. 10)	117M	1.1M	97M				
Co transfection of Plasmid comprising the sequence: Pcmv-4ORF ⁺ -TD1-Tfluc-S-cDTAWT---Psv40-TGFP + co-transfected with siRNA-	E80 (SEQ ID NO. 1)	14M	65K	10.3M	4.9M	2.4M	1.4M	7.2M
Co transfection of Plasmid comprising the sequence Pcmv-4ORF ⁺ -TD1-Tfluc-S-cDTAWT---Psv40-TGFP co-transfected with siRNA+	E80 (SEQ ID NO. 1)	2.69M*	18K*	2.7M*	1.2M*	586K*	347K*	2.1M*
Fold of activation = RLU measured in the presence of siRNA- divided by RLU measured in the presence of siRNA+	si-/si+	5.2	3.6	3.8	4	4.1	4	3.4
Fold of leakiness = E34 (SEQ ID NO. / E80- {smaller than 1 = 0 leakage})	E34 /E80-	8.35	16.92	9.41				

Table 6C

Experiment number		15	16	17	18	19	20	21
Number of 293HEK cells per well (24 well plate)	#293cells	125K	125K	100K	100K	100K	100K	200K

Hours post transfection	hrPT	48hr	48hr	72hr	72hr	72hr	72hr	24hr
co-transfection of Renilla expressing plasmid [ng]	REN	E11[140]	E11[110]	E11[150]	E11[150]	E11[750]	E11[750]	E11[170]
co-transfection of siRNA- or siRNA-: [pico mole]	siRNA	[10]	[10]	[10]	[15]	[10]	[15]	[10]
co-transfection of one of the test plasmids below [ng]: / results shown are RLU under the indicated experimental condition	↓ / RLU	[60]	[90]	[50]	[50]	[50]	[50]	[30]
Co transfection of Plasmid comprising the sequence: Pcmv-Tfluc-TD1-cDTAWT---Psv40-TGFP	E28 (SEQ ID NO. 11)							97K
Co transfection of Plasmid comprising the sequence: Pcmv-4ORF ^Δ -TD1-Tfluc---Psv40-TGFP	E34 (SEQ ID NO. 10)							10.7M
Co transfection of Plasmid comprising the sequence: Pcmv-4ORF ^Δ -TD1-Tfluc-S-cDTAWT---Psv40-TGFP + co-transfected with siRNA-	E80 (SEQ ID. NO. 1)	3.16M	1.76M	3.67M	4.3M	13.3M	13.3M	4.2M
Co transfection of Plasmid comprising the sequence Pcmv-4ORF ^Δ -TD1-Tfluc-S-cDTAWT---Psv40-TGFP co-transfected with siRNA+	E80 (SEQ ID. NO. 1)	950K*	573K*	1.4M*	1.4M*	5.8M*	6.1M*	2.1M*

Fold of activation = RLU measured in the presence of siRNA- divided by RLU measured in the presence of siRNA+	si-/si+	3.32	3	2.6	3	2.3	2.18	2
Fold of leakiness = E34 (SEQ ID NO. /E80- {smaller than 1 = 0 leakage})	E34 /E80-							2.54

With respect to Table 6A-6C:

*= Indicate that the 2 siRNA+ show significant activation;

** = co-transfected also with 155ng of plasmid E38 (SEQ ID NO. 19).

5

The results presented above in Tables 1-5 and 6A-6C clearly show that in the presence of an siRNA molecule(s) capable of inducing cleavage of the exogenous RNA of interest, the exogenous protein of interest (DTA) is expressed which, in turn results in increased cell death. The increased cell death results in reduced overall RLU measurements in the well, since less cells are expressing/producing the luciferase gene. The results demonstrate that indeed, only in cells which comprise a specific siRNA, the exogenous protein of interest (DTA in this example) is expressed, since only in these cells, cleavage of the exogenous RNA of interest at the cleavage site is induced, thereby allowing expression of the exogenous protein of interest in the cells.

15

REFERENCES

1. T. Dalmay, 2008, MicroRNAs and cancer, Journal of Internal Medicine 263, 4: 366–375.
- 20 2. Y Zeng, 2006, Principles of micro-RNA production and maturation, Oncogene 25: 6156–6162.
3. BM Engels and G Hutvagner, 2006, Principles and effects of microRNA-mediated post-transcriptional gene regulation, Oncogene 25: 6163–6169.

4. Benjamin Haley & Phillip D Zamore, 2004, Kinetic analysis of the RNAi enzyme complex, *Nature Structural & Molecular Biology* 11: 599 - 606.
5. William CS Cho, 2007, OncomiRs: the discovery and progress of microRNAs in cancers. *Molecular Cancer* 6: 60.
- 5 6. Yan Zeng, Rui Yi and Bryan R. Cullen, 2003, MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. *Proc Natl Acad Sci U S A* 100, 17: 9779-9784.
7. Wadhwa R, Kaul SC, Miyagishi M, Taira K, 2004, Know-how of RNA interference and its applications in research and therapy. *Mutat Res* 567, 1: 71-84.
- 10 8. Sayda M. Elbashir, Winfried Lendeckel and Thomas Tuschl, 2001, RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes & Dev* 15, 2: 188-200.
9. Richard I. Gregory, Thimmaiah P. Chendrimada, Neil Cooch, Ramin Shiekhattar, 2005, Human RISC Couples MicroRNA Biogenesis and Posttranscriptional Gene Silencing. *Cell* 123: 631-640.
- 15 10. Gallie DR, 1991, The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. *Genes Dev* 5, 11: 2108-16.
11. Entrez Gene: HIST1H2AC histone cluster 1, H2ac.
12. Lord MJ, Jolliffe NA, Marsden CJ, et al, 2003, Ricin Mechanisms of Cytotoxicity, *Toxicol Rev* 22, 1: 53-64.
- 20 13. Jeen-Kuan Chen, Chih-Hung Hung, Yen-Chywan Liaw and Jung-Yaw Lin, 1997, Identification of amino acid residues of Abrin-a A chain is essential for catalysis and reassociation with Abrin-a B chain by site-directed mutagenesis, *Protein Engineering* 10: 827-833.
14. Yamaizumi, M, Mekada, E, Uchida, T. and Okada Y, 1978, One molecule of Diphtheria toxin fragment A introduced into a cell can kill the cell, *cell* 15, 1: 245-50.
- 25 15. Weinberg M.S and Morris K.V, 2006, Are viral-encoded microRNAs mediating latent HIV-1 infection?, *DNA Cell Biol* 25: 223- 231.
16. Velculescu VE et al, 2006, The Consensus Coding Sequences of Human Breast and Colorectal Cancers, *Science* 314, 5797: 268 - 274.

17. Intronn, Inc.
18. M. Puttaraju¹, Sharon F. Jamison, S. Gary Mansfield, Mariano A. Garcia-Blanco and Lloyd G. Mitchell, 1999, Spliceosome-mediated RNA trans-splicing as a tool for gene therapy. *Nature Biotechnology* 17: 246 – 252.
- 5 19. Song MS, Lee SW, 2006, Cancer-selective induction of cytotoxicity by tissue-specific expression of targeted trans-splicing ribozyme, *FEBS Lett* 580, 21: 5033-43.
20. Yaakov Benenson, Binyamin Gil, Uri Ben-Dor, Rivka Adar & Ehud Shapiro, 2004, An autonomous molecular computer for logical control of gene expression, *Nature* 429, 6990: 423-9.
- 10 21. Sano M, Kato Y, Taira K, Functional gene-discovery systems based on libraries of hammerhead and hairpin ribozymes and short hairpin RNAs, 2005, *Mol Biosyst.* 1:27-35.
22. Maurille J. Fournier et al, 1999, A small nucleolar RNA: ribozyme hybrid cleaves a nucleolar RNA target *in vivo* with near-perfect efficiency, *PNAS* 96, 12: 6609-6614.
23. Laising Yen, Jennifer Svendsen, Jeng-Shin Lee, John T. Gray, Maxime Magnier,
15 Takashi Baba, Robert J. D'Amato & Richard C. Mulligan, 2004, Exogenous control of mammalian gene expression through modulation of RNA self-cleavage, *Nature* 431, 471-476.
24. Chabanon Herve and Ian Mickleburgh, 2004, Zipcodes and postage stamps: mRNA localisation signals and their *trans*-acting binding proteins, *Briefings in Functional*
20 *Genomics and Proteomics* 3:240-256.
25. Wu L, Wells D, Tay J, Mendis D, Abbott MA, Barnitt A, Quinlan E, Heynen A, Fallon JR, Richter JD, 1998, CPEB-mediated cytoplasmic polyadenylation and the regulation of experience-dependent translation of alpha-CaMKII mRNA at synapses, *Neuron* 21, 5: 936-8.
- 25 26. Maria V. Baez and Graciela L. Boccaccio, Career investigator of the Consejo Nacional de Investigaciones Cientificas y Tecnologicas, 2005, Mammalian Smaug Is a Translational Repressor That Forms Cytoplasmic Foci Similar to Stress Granules, *J. Biol. Chem* 280, 52: 43131-43140.

27. C A Smibert, J E Wilson, K Kcrr, and P M Macdonald, 1996, smaug protein represses translation of unlocalized nanos mRNA in the Drosophila embryo, GENES & DEVELOPMENT 10:2600-2609.
28. Carine Barreau, Luc Paillard and H. Beverley Osborne, 2006, AU-rich elements and associated factors: are there unifying principles?, Nucleic Acids Research 33, 22: 7138-7150.
29. Trent P. Munro, Rebecca J. Magee, Grahame J. Kidd, John H. Carson, Elisa Barbarese, Lisa M. Smith, and Ross Smith, 1999, Mutational Analysis of a Heterogeneous Nuclear Ribonucleoprotein A2 Response Element for RNA Trafficking, J Biol Chem, 274, 48: 34389-34395.
30. Suresh subramani, 1998, Components Involved in Peroxisome Import, Biogenesis, Proliferation, Turnover, and Movement, PHYSIOLOGICAL REVIEWS 78: 171-188.
31. Isken O, Maquat LE, 2007, Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function, Genes Dev 21, 15:1833-56.
32. Zheng H, Li LL, Hu DS, Deng XY, Cao Y, 2007, Role of Epstein-Barr virus encoded latent membrane protein 1 in the carcinogenesis of nasopharyngeal carcinoma, Cell Mol Immunol, 3: 185-96.
33. Mei YP, Zhou JM, Wang Y, Huang H, Deng R, Feng GK, Zeng YX, Zhu XF, 2007, Silencing of LMP1 induces cell cycle arrest and enhances chemosensitivity through inhibition of AKT signaling pathway in EBV-positive nasopharyngeal carcinoma cells, Cell Cycle 6, 11: 1379-85.
34. Jacque JM, Triques K, Stevenson M, 2002, Modulation of HIV-1 replication by RNA interference, Nature 418, 6896: 435-8.
35. Zeng Y, Wagner EJ, Cullen BR, 2002, Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells, Mol Cell 6: 1327-33.
36. Joel D. Richter, 2001, Think globally, translate locally: What mitotic spindles and neuronal synapses have in common, Proc Natl Acad Sci U S A 98, 13: 7069-7071.

37. Michael S. Wollenberg and Sanford M. Simon, 2004, Signal Sequence Cleavage of Peptidyl-tRNA Prior to Release from the Ribosome and Translocon, *J. Biol. Chem.* 279: 24919-24922.
38. Dan Frumkin, Adam Wasserstrom, Shalev Itzkovitz, Tomer Stern, Alon Harmelin, Raya
5 Eilam, Gideon Rechavi and Ehud Shapiro, 2008, Cell Lineage Analysis of a Mouse Tumor, *Cancer Research* 68: 5924-5931.
39. Elfakess R, Dikstein R. (2008). A translation initiation element specific to mRNAs with very short 5'UTR that also regulates transcription. *PloS One*. 2008 Aug 28;3(8):e3094.

CLAIMS

1. A composition comprising one or more polynucleotides for directing specific cleavage of an exogenous RNA of interest at a specific target site, the cleavage taking place only
5 in the presence of an endogenous signal RNA in a cell, the endogenous signal RNA being an RNA molecule which comprises a signal sequence, the signal sequence being any predetermined sequence of from 18 to 25 nucleotides in length, whereby introduction of said composition into a cell comprising said endogenous signal RNA, directs the cleavage of said exogenous RNA of interest at the specific target site
10 that is located within a specific sequence, which is of sufficient complementarity to hybridize with the predetermined signal sequence.
2. The composition of claim 1 wherein said one or more polynucleotides comprises:
a first polynucleotide sequence encoding said exogenous RNA of interest;
a second polynucleotide sequence encoding a functional RNA capable of mediating the
15 cleavage of the endogenous signal RNA at a predetermined cleavage site; and
a third polynucleotide sequence encoding a carrier RNA.
3. The composition of claim 2, wherein said carrier RNA is an RNA molecule that is at least about 18 nucleotides in length and is consisting essentially of:
(1) a first sequence of from 14 to 31 nucleotides in length which is of sufficient
20 complementarity to an edge sequence to hybridize therewith, said edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides downstream from said predetermined cleavage site and extends downstream in said endogenous signal RNA;
(2) a second sequence downstream from said first sequence, wherein said second sequence is a random sequence that is 0-5 nucleotides in length;
25 (3) a third sequence upstream from said first sequence, wherein said third sequence is 0-7000 nucleotides in length; and
wherein said predetermined cleavage site is the 5' end of said predetermined signal sequence.
4. The composition of claim 2, wherein said carrier RNA comprises an RNA molecule that
30 is at least about 18 nucleotides in length and is consisting essentially of:
(1) a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, said edge sequence is 14-

31 nucleotides in length and is located 0-5 nucleotides upstream from said predetermined cleavage site and extends upstream in said endogenous signal RNA;

(2) a second sequence upstream from the first sequence, wherein said second sequence is a random sequence that is 0-5 nucleotides in length;

5 (3) a third sequence downstream from the first sequence, wherein said third sequence is 0-7000 nucleotides in length; and
wherein said predetermined cleavage site is the 3' end of said predetermined signal sequence.

10 5. The composition of claim 2 wherein said carrier RNA is processed from a polynucleotide sequence comprising a carrier sequence that is at least about 18 nucleotides in length, said carrier sequence consisting essentially of:

(1) a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, said edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides downstream from said predetermined cleavage site and extends downstream in said endogenous signal RNA;

15 (2) a second sequence downstream from said first sequence, wherein said second sequence is a random sequence that is 0-5 nucleotides in length; and

(3) a third sequence upstream from said first sequence, wherein said third sequence is 0-7000 nucleotides in length;

20 wherein said polynucleotide sequence is cleaved within the cell at a carrier cleavage site that is a 3' end of said carrier sequence;

wherein the cleavage at the carrier cleavage site is effected by a functional nucleic acid which is encoded by a fourth polynucleotide sequence of the composition; and

25 wherein the predetermined cleavage site is the 5' end of said predetermined signal sequence.

6. The composition of claim 2 wherein said carrier RNA is processed from a polynucleotide sequence comprising a carrier sequence that is at least about 18 nucleotides in length, said carrier sequence consisting essentially of:

30 (1) a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, said edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides upstream from said predetermined cleavage site and extends upstream in said endogenous signal RNA;

(2) a second sequence upstream from the first sequence, wherein said second sequence is a random sequence that is 0-5 nucleotides in length; and

(3) a third sequence downstream from said first sequence, wherein said third sequence is 0-7000 nucleotides in length;

5 wherein said polynucleotide sequence is cleaved within the cell at a carrier cleavage site that is 5' end of said carrier sequence;

wherein the cleavage at the carrier cleavage site is effected by a functional nucleic acid which is encoded by a fourth polynucleotide sequence of the composition; and

10 wherein said predetermined cleavage site is the 3' end of said predetermined signal sequence.

7. The composition of claim 3 or 5, wherein said edge sequence is 23-28 nucleotides in length and is located from the predetermined cleavage site to about 23-28 nucleotides downstream, wherein said second sequence is 2 nucleotides in length and wherein said third sequence is 0 nucleotides in length.

15 8. The composition of claim 4 or 6, wherein said edge sequence is 25-30 nucleotides in length and is located 2 nucleotides upstream from the predetermined cleavage site and extends upstream in said endogenous signal RNA, wherein said second sequence is 0 nucleotides in length and wherein said third sequence is 0 nucleotides in length.

20 9. The composition of claim 1, wherein said endogenous signal RNA is a cellular mRNA, viral RNA, or both.

10. The composition of claim 1, wherein said predetermined signal sequence is unique to neoplastic cells, viral infected cells, or both.

11. The composition of claim 1, wherein sufficient complementarity is at least 30% complementarity.

25 12. The composition of claim 1, wherein said sufficient complementarity is at least 90%.

13. The composition of claim 1, wherein said one or more polynucleotides comprises one or more DNA molecules, one or more RNA molecules or combinations thereof.

30 14. The composition of claim 2, wherein said functional RNA is selected from the group consisting of: microRNA (miRNA), lariat-form RNA, short-hairpin RNA (shRNA), siRNA expression domain, antisense RNA, double-stranded RNA (dsRNA), small-interfering RNA (siRNA) and ribozyme.

15. The composition of claim 1, wherein said exogenous RNA of interest further comprises:
(a) a sequence encoding an exogenous protein of interest; and

(b) an inhibitory sequence that is capable of inhibiting the expression of the exogenous protein of interest;

wherein said specific target site is located between the inhibitory sequence and the sequence encoding an exogenous protein of interest, whereby following introduction of said composition into a cell comprising the endogenous signal RNA, said exogenous RNA of interest is transcribed and cleaved at said specific target site whereby said inhibitory sequence is detached from said sequence encoding the exogenous protein of interest and the exogenous protein of interest is capable of being expressed.

16. The composition of claim 15, wherein said exogenous protein of interest is selected from the group consisting of: Ricin, Ricin A chain, Abrin, Abrin A chain, Diphtheria toxin A chain and modified forms thereof.

17. The composition of claim 15, wherein said exogenous protein of interest is selected from the group consisting of: alpha toxin, saporin, maize RIP, barley RIP, wheat RIP, corn RIP, rye RIP, flax RIP, Shiga toxin, Shiga-like RIP, momordin, thymidine kinase, pokeweed antiviral protein, gelonin, Pseudomonas exotoxin, Pseudomonas exotoxin A, Escherichia coli cytosine deaminase and modified forms thereof.

18. The composition of claim 15, wherein said inhibitory sequence is located upstream from said specific target site and wherein the inhibitory sequence reduces the efficiency of translation of said exogenous protein of interest.

19. The composition of claim 18, wherein said inhibitory sequence comprises a plurality of initiation codons, wherein each of said initiation codons and the sequence encoding the exogenous protein of interest are not in the same reading frame.

20. The composition of claim 19, wherein said exogenous RNA of interest further comprises a stop codon that is located between said initiation codon and the start codon of said sequence encoding the exogenous protein of interest, wherein the stop codon and the initiation codon are in the same reading frame.

21. The composition of claim 19, wherein said inhibitory sequence further comprises a nucleotide sequence downstream from the initiation codon, wherein said nucleotide sequence and said initiation codon are in the same reading frame, and wherein the nucleotide sequence encodes a sorting signal for subcellular localization.

22. The composition of claim 21, wherein said subcellular localization is selected from the group consisting of: mitochondria, nucleus, endosome, lysosome, peroxisome and endoplasmic reticulum (ER).

23. The composition of claim 19, wherein said inhibitory sequence further comprises a nucleotide sequence downstream from the initiation codon, wherein said nucleotide sequence and said initiation codon are in the same reading frame; and wherein said nucleotide sequence encodes a protein degradation signal.
24. The composition of claim 19, wherein said inhibitory sequence further comprises a nucleotide sequence downstream from the initiation codon, wherein said nucleotide sequence and said initiation codon are in the same reading frame; wherein said nucleotide sequence and said sequence encoding the exogenous protein of interest are in the same reading frame, wherein said nucleotide sequence encodes an amino acid sequence, whereby when the amino acid sequence is fused to the exogenous protein of interest the biological function of the exogenous protein of interest is inhibited.
25. The composition of claim 19, wherein said exogenous RNA of interest further comprises a stop codon downstream from said initiation codon, wherein said stop codon and said initiation codon are in the same reading frame and wherein said exogenous RNA of interest further comprises an intron downstream from the stop codon, whereby the exogenous RNA of interest is a target for nonsense-mediated decay (NMD).
26. The composition of claim 15, wherein said inhibitory sequence is located downstream from the sequence encoding the exogenous protein of interest and wherein said inhibitory sequence comprises an RNA localization signal for subcellular localization or an endogenous miRNA binding site.
27. The composition of claim 15, wherein said exogenous RNA of interest comprises an internal ribosome entry site (IRES) sequence downstream from the specific cleavage site and upstream from the sequence encoding the exogenous protein of interest, wherein said IRES sequence is more functional within the cleaved exogenous RNA of interest than within the intact exogenous RNA of interest.
28. The composition of claim 15, wherein said exogenous RNA of interest further comprises a nucleotide sequence immediately upstream from said sequence encoding the exogenous protein of interest, wherein said nucleotide sequence comprises an internal ribosome entry site (IRES) sequence, capable of increasing the efficiency of translation of said exogenous protein of interest in the cleaved exogenous RNA of interest.
29. The composition of claim 15, wherein said exogenous RNA of interest further comprises a nucleotide sequence comprising a cytoplasmic polyadenylation element, located immediately downstream from said sequence encoding the exogenous protein of

interest, wherein said cytoplasmic polyadenylation element increases the efficiency of translation of said exogenous protein of interest in the cleaved exogenous RNA of interest.

30. The composition of claim 15, wherein said composition further comprises an additional
5 polynucleotide sequence that encodes an additional RNA molecule, said additional RNA molecule comprises at the 3' end a nucleotide sequence that is capable of binding to a sequence that is located upstream of said specific target site and downstream from the sequence encoding the exogenous protein of interest, wherein said additional RNA molecule, increases the efficiency of translation of said exogenous protein of interest in
10 the cleaved exogenous RNA of interest.

31. The composition of claim 15, wherein said composition further comprise an additional polynucleotide sequence that encodes a cleaving component that is capable of effecting the cleavage, of said exogenous RNA of interest at a position that is located upstream from the inhibitory sequence, wherein said cleaving component(s) is selected from the
15 group consisting of:

(a) a nucleic acid sequence that is located within said exogenous RNA of interest, wherein said nucleic acid sequence is selected from the group consisting of:

endonuclease recognition site, endogenous miRNA binding site, cis acting ribozyme and miRNA sequence, wherein said nucleic acid sequence, reduces the efficiency of

20 translation of said exogenous protein of interest in the exogenous RNA of interest; and

(b) an inhibitory RNA, wherein said inhibitory RNA is selected from the group consisting of: microRNA (miRNA), lariat-form RNA, short-hairpin RNA (shRNA),

siRNA expression domain, antisense RNA, double-stranded RNA (dsRNA), small-interfering RNA (siRNA) and ribozyme, wherein said inhibitory RNA, reduces the

25 efficiency of translation of said exogenous protein of interest in said exogenous RNA of interest.

32. The composition of claim 15, wherein the specific sequence is a plurality of specific sequences and wherein said specific target site is a plurality of specific target sites.

33. The composition of claim 15, wherein said inhibitory sequence is located upstream from
30 the sequence encoding the exogenous protein of interest, wherein the inhibitory sequence is capable of forming a secondary structure having a folding free energy of lower than -30 kcal/mol, whereby said secondary structure is sufficient to block scanning ribosomes from reaching the start codon of said exogenous protein of interest.

34. The composition of claim 2, wherein said exogenous RNA of interest and said functional RNA are capable of being located on the same or different polynucleotide molecules.
35. The composition of claim 5 or 6, wherein said exogenous RNA of interest, the functional RNA and the functional nucleic acid are capable of being located on one or more polynucleotide molecules.
36. The composition of claim 1, wherein said one or more polynucleotides are integrated into the cell genome.
37. The composition of claim 1, wherein said cell is selected from a group consisting of: human cell, animal cell, cultured cell and plant cell.
38. The composition of claim 1, wherein said cell is present in an organism
39. A composition comprising one or more polynucleotides for directing specific expression of an exogenous protein of interest in a cell, wherein the exogenous protein of interest is expressed only in the presence of an endogenous signal RNA in a cell, the endogenous signal RNA being an RNA molecule which comprises a signal sequence, the signal sequence being any predetermined sequence of from 18 to 25 nucleotides in length, whereby introduction of said composition into a cell comprising said endogenous signal RNA, directs the cleavage of an exogenous RNA of interest at a specific target site that is located within a specific sequence, which is of sufficient complementarity to hybridize with the predetermined signal sequence, wherein only after the cleavage of said exogenous RNA of interest in the cell, the exogenous protein of interest, which is encoded by said cleaved exogenous RNA of interest is capable of being expressed in the cell.
40. The composition of claim 39 wherein said one or more polynucleotides comprises:
a first polynucleotide sequence encoding said exogenous RNA of interest;
a second polynucleotide sequence encoding a functional RNA capable of mediating the cleavage of the endogenous signal RNA at a predetermined cleavage site; and
a third polynucleotide sequence encoding a carrier RNA.
41. The composition of claim 40, wherein said carrier RNA is an RNA molecule that is at least about 18 nucleotides in length and is consisting essentially of:
(1) a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, said edge sequence is 14-

31 nucleotides in length and is located 0-5 nucleotides downstream from said predetermined cleavage site and extends downstream in said endogenous signal RNA;

(2) a second sequence downstream from said first sequence, wherein said second sequence is a random sequence that is 0-5 nucleotides in length; and

5 (3) a third sequence upstream from said first sequence, wherein said third sequence is 0-7000 nucleotides in length.

42. The composition of claim 40, wherein said carrier RNA comprises an RNA molecule that is at least about 18 nucleotides in length and is consisting essentially of:

10 (1) a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, said edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides upstream from said predetermined cleavage site and extends upstream in said endogenous signal RNA;

(2) a second sequence upstream from the first sequence, wherein said second sequence is a random sequence that is 0-5 nucleotides in length; and

15 (3) a third sequence downstream from the first sequence, wherein said third sequence is 0-7000 nucleotides in length.

43. The composition of claim 40 wherein said carrier RNA is processed from a polynucleotide sequence comprising a carrier sequence that is at least about 18 nucleotides in length, said carrier sequence consisting essentially of:

20 (1) a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, said edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides downstream from said predetermined cleavage site and extends downstream in said endogenous signal RNA;

25 (2) a second sequence downstream from said first sequence, wherein said second sequence is a random sequence that is 0-5 nucleotides in length; and

(3) a third sequence upstream from said first sequence, wherein said third sequence is 0-7000 nucleotides in length;

wherein said polynucleotide sequence is cleaved within the cell at a carrier cleavage site that is a 3' end of said carrier sequence; and

30 wherein the cleavage at the carrier cleavage site is effected by a functional nucleic acid which is encoded by a fourth polynucleotide sequence of the composition.

44. The composition of claim 40 wherein said carrier RNA is processed from a polynucleotide sequence comprising a carrier sequence that is at least about 18 nucleotides in length, said carrier sequence consisting essentially of:

- 5 (1) a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, said edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides upstream from said predetermined cleavage site and extends upstream in said endogenous signal RNA;
- (2) a second sequence upstream from the first sequence, wherein said second sequence is a random sequence that is 0-5 nucleotides in length; and
- 10 (3) a third sequence downstream from said first sequence, wherein said third sequence is 0-7000 nucleotides in length;

wherein said polynucleotide sequence is cleaved within the cell at a carrier cleavage site that is 5' end of said carrier sequence; and

15 wherein the cleavage at the carrier cleavage site is effected by a functional nucleic acid which is encoded by a fourth polynucleotide sequence of the composition.

45. The composition of claim 39, wherein said exogenous RNA of interest further comprises:

- a) a sequence encoding the exogenous protein of interest; and
- a) an inhibitory sequence that is capable of inhibiting the expression of the exogenous
- 20 protein of interest;

wherein said specific target site is located between the inhibitory sequence and the sequence encoding the exogenous protein of interest, whereby, following introduction of said composition into a cell comprising the endogenous signal RNA, said exogenous RNA of interest is transcribed and cleaved at said specific target site whereby the

25 inhibitory sequence is detached from said sequence encoding the exogenous protein of interest and the exogenous protein of interest is capable of being expressed.

46. The composition of claim 39, wherein said endogenous signal RNA is a cellular mRNA, viral RNA, or both.

47. The composition of claim 39, wherein said predetermined signal sequence is unique to

30 neoplastic cells, viral infected cells, or both.

48. The composition of claim 39, wherein sufficient complementarity is at least 30% complementarity.

49. The composition of claim 39, wherein said sufficient complementarity is at least 90%.

50. The composition of claim 39, wherein said one or more polynucleotides comprises one or more DNA molecules, one or more RNA molecules or combinations thereof.
51. The composition of claim 39, wherein said exogenous protein of interest is selected from the group consisting of: Ricin, Ricin A chain, Abrin, Abrin A chain, Diphtheria toxin A chain, alpha toxin, saporin, maize RIP, barley RIP, wheat RIP, corn RIP, rye RIP, flax RIP, Shiga toxin, Shiga-like RIP, momordin, thymidine kinase, pokeweed antiviral protein, gelonin, Pseudomonas exotoxin, Pseudomonas exotoxin A, Escherichia coli cytosine deaminase and modified forms thereof.
52. The composition of claim 39, wherein said functional RNA is selected from the group consisting of: microRNA (miRNA), lariat-form RNA, short-hairpin RNA (shRNA), siRNA expression domain, antisense RNA, double-stranded RNA (dsRNA), small-interfering RNA (siRNA) and ribozyme.
53. The composition of claim 42, wherein said cell is selected from a group consisting of: human cell, animal cell, cultured cell and plant cell.
54. The composition of claim 42, wherein said cell is present in an organism.
55. A method for killing a specific cell population which comprises an endogenous signal RNA, the method comprises: introducing the cell population with a composition comprising one or more polynucleotides for directing specific cleavage of an exogenous RNA of interest at a specific target site that is located within a specific sequence, which is of sufficient complementarity to hybridize with the endogenous signal RNA, the endogenous signal RNA being an RNA molecule which comprises a signal sequence, the signal sequence being any predetermined sequence of from 18 to 25 nucleotides in length; and wherein the cleavage of the exogenous RNA of interest in the cell population, allows the expression of an exogenous protein of interest, capable of killing the cell population.
56. The method of claim 55 wherein said one or more polynucleotides comprises:
a first polynucleotide sequence encoding said exogenous RNA of interest;
a second polynucleotide sequence encoding a functional RNA capable of mediating the cleavage of the endogenous signal RNA at a predetermined cleavage site; and
a third polynucleotide sequence encoding a carrier RNA.
57. The method of claim 56, wherein said carrier RNA is an RNA molecule that is at least about 18 nucleotides in length and is consisting essentially of:

- (1) a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, said edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides downstream from said predetermined cleavage site and extends downstream in said endogenous signal RNA;
- 5 (2) a second sequence downstream from said first sequence, wherein said second sequence is a random sequence that is 0-5 nucleotides in length; and
- (3) a third sequence upstream from said first sequence, wherein said third sequence is 0-7000 nucleotides in length.

10 58. The method of claim 56, wherein said carrier RNA comprises an RNA molecule that is at least about 18 nucleotides in length and is consisting essentially of:

- (1) a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, said edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides upstream from said predetermined cleavage site and extends upstream in said endogenous signal RNA;
- 15 (2) a second sequence upstream from the first sequence, wherein said second sequence is a random sequence that is 0-5 nucleotides in length; and
- (3) a third sequence downstream from the first sequence, wherein said third sequence is 0-7000 nucleotides in length.

20 59. The method of claim 56 wherein said carrier RNA is processed from a polynucleotide sequence comprising a carrier sequence that is at least about 18 nucleotides in length, said carrier sequence consisting essentially of:

- (1) a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, said edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides downstream from said predetermined cleavage site and extends downstream in said endogenous signal RNA;
- 25 (2) a second sequence downstream from said first sequence, wherein said second sequence is a random sequence that is 0-5 nucleotides in length; and
- (3) a third sequence upstream from said first sequence, wherein said third sequence is 0-7000 nucleotides in length;

30 wherein said polynucleotide sequence is cleaved within the cell at a carrier cleavage site that is a 3' end of said carrier sequence; and

wherein the cleavage at the carrier cleavage site is effected by a functional nucleic acid which is encoded by a fourth polynucleotide sequence of the composition.

60. The method of claim 56 wherein said carrier RNA is processed from a polynucleotide sequence comprising a carrier sequence that is at least about 18 nucleotides in length, said carrier sequence consisting essentially of:

(1) a first sequence of from 14 to 31 nucleotides in length which is of sufficient complementarity to an edge sequence to hybridize therewith, said edge sequence is 14-31 nucleotides in length and is located 0-5 nucleotides upstream from said predetermined cleavage site and extends upstream in said endogenous signal RNA;

(2) a second sequence upstream from the first sequence, wherein said second sequence is a random sequence that is 0-5 nucleotides in length; and

(3) a third sequence downstream from said first sequence, wherein said third sequence is 0-7000 nucleotides in length;

wherein said polynucleotide sequence is cleaved within the cell at a carrier cleavage site that is 5' end of said carrier sequence; and

wherein the cleavage at the carrier cleavage site is effected by a functional nucleic acid which is encoded by a fourth polynucleotide sequence of the composition.

61. The method of claim 55, wherein said endogenous signal RNA is a cellular mRNA, viral RNA, or both.

62. The method of claim 55, wherein sufficient complementarity is at least 30% complementarity.

63. The method of claim 55, wherein said sufficient complementarity is at least 90%.

64. The method of claim 55, wherein said polynucleotide comprises one or more DNA molecules, one or more RNA molecules or combinations thereof.

65. The method of claim 55, wherein said functional RNA is selected from the group consisting of: microRNA (miRNA), lariat-form RNA, short-hairpin RNA (shRNA), siRNA expression domain, antisense RNA, double-stranded RNA (dsRNA), small-interfering RNA (siRNA) and ribozyme.

66. The method of claim 55, wherein said exogenous protein of interest is selected from the group consisting of: Ricin, Ricin A chain, Abrin, Abrin A chain, Diphtheria toxin A chain, alpha toxin, saporin, maize RIP, barley RIP, wheat RIP, corn RIP, rye RIP, flax RIP, Shiga toxin, Shiga-like RIP, momordin, thymidine kinase, pokeweed antiviral protein, gelonin, Pseudomonas exotoxin, Pseudomonas exotoxin A, Escherichia coli cytosine deaminase and modified forms thereof.

67. The method of claim 55, wherein said cell population is selected from a group consisting of: human cell, animal cell, cultured cell and plant cell.
68. The method claim 55, wherein said cell population is a neoplastic cell population.
69. The method of claim 55, wherein said cell population is present in an organism.

5

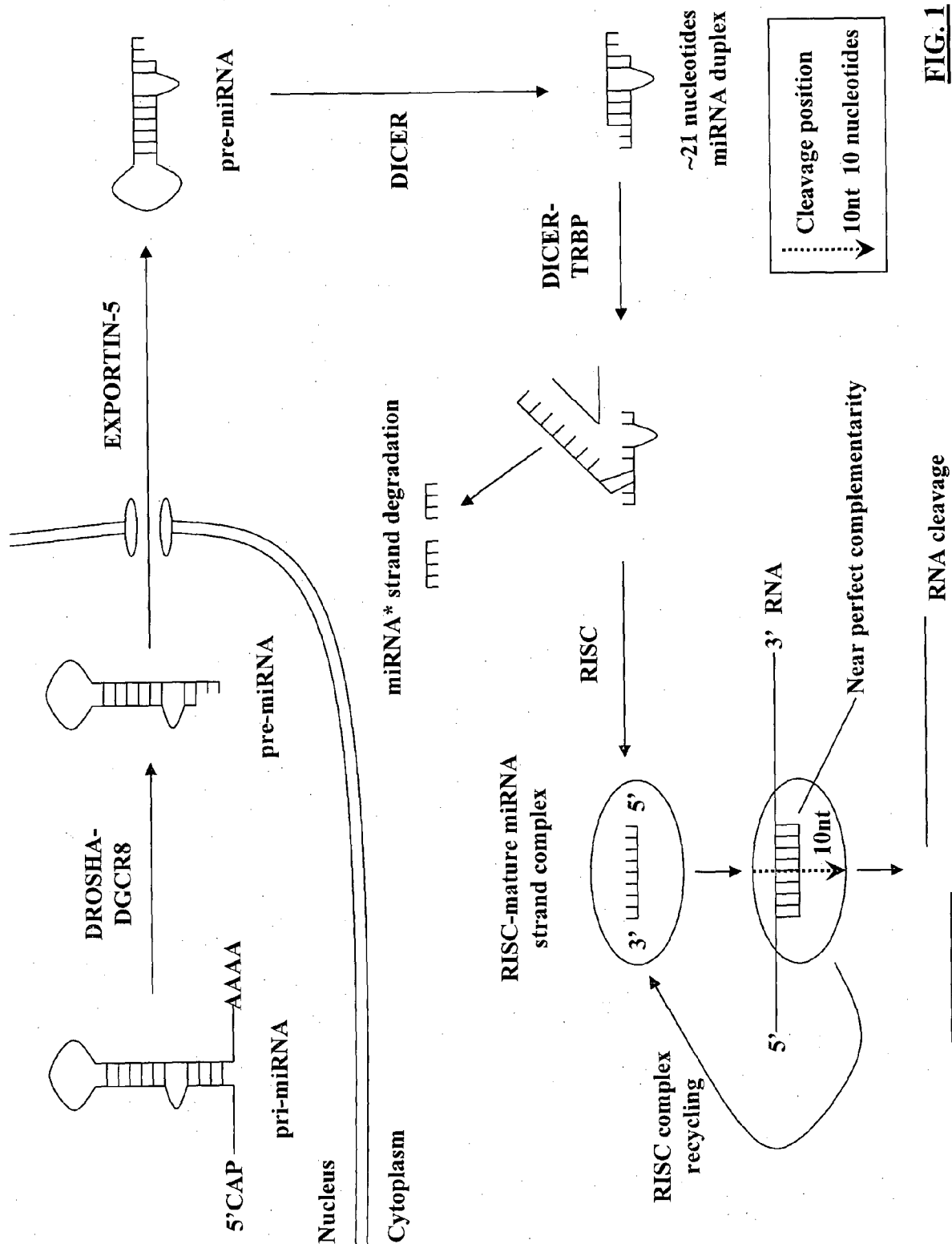


FIG. 1

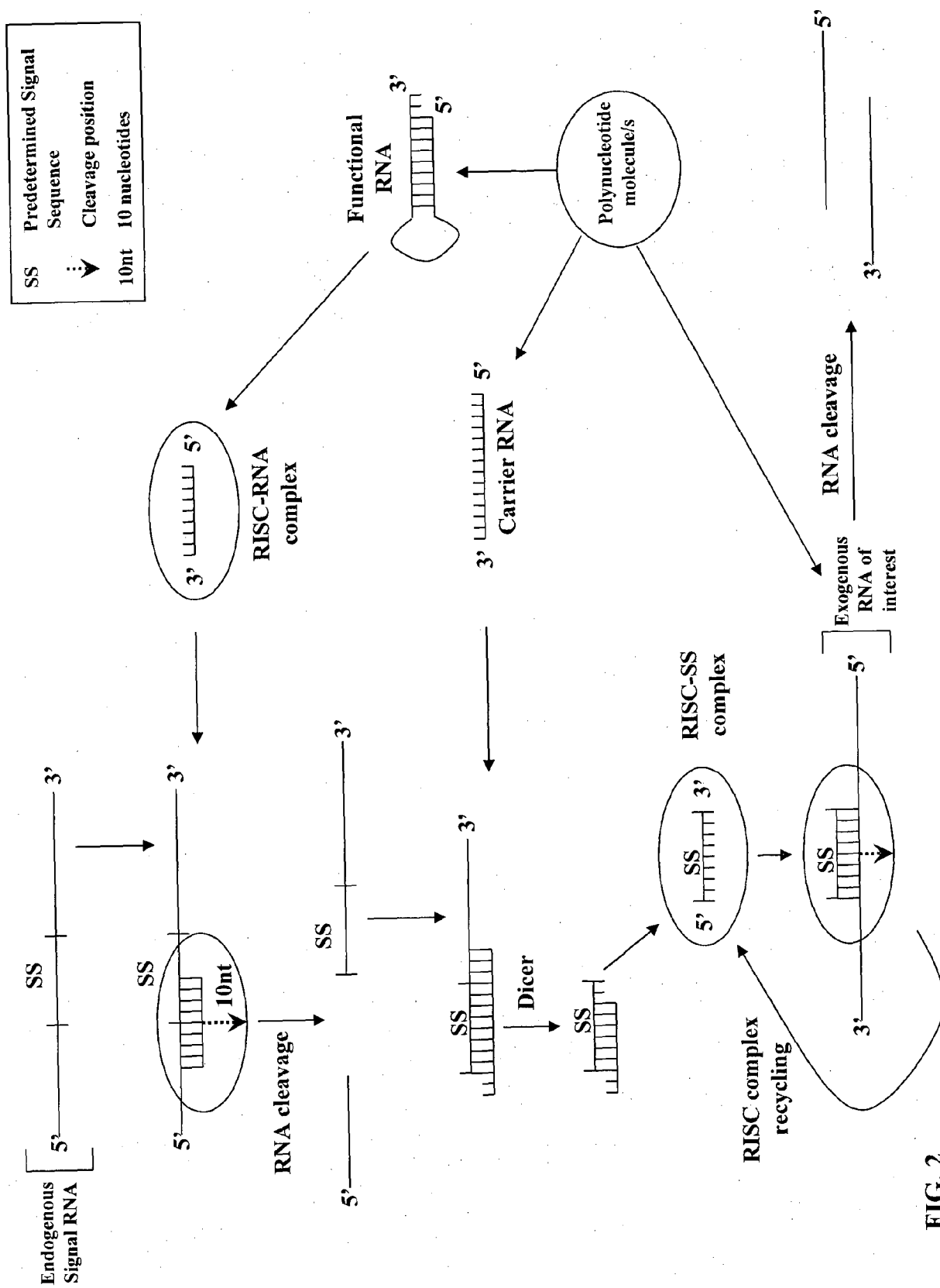


FIG. 2

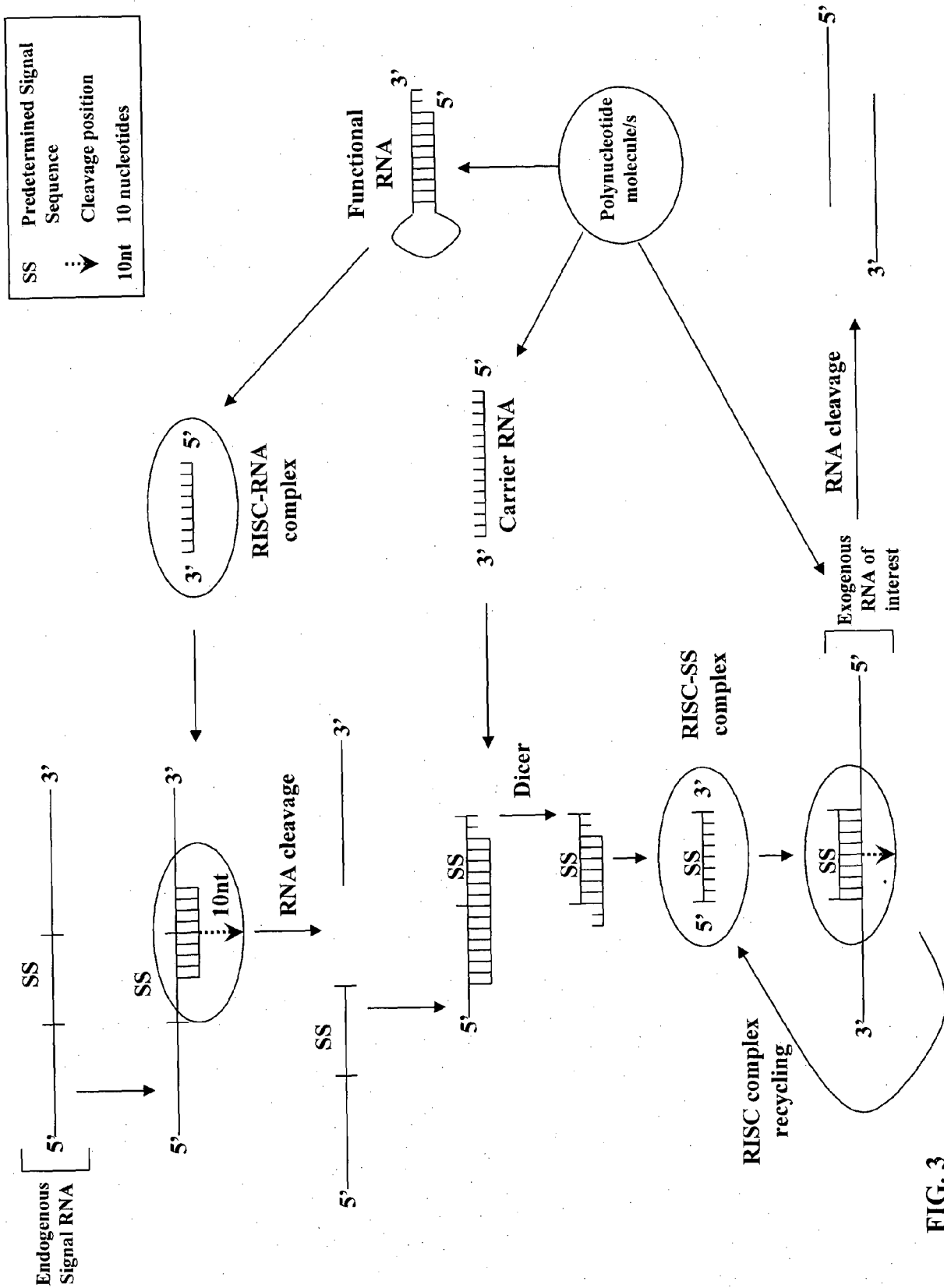


FIG. 3

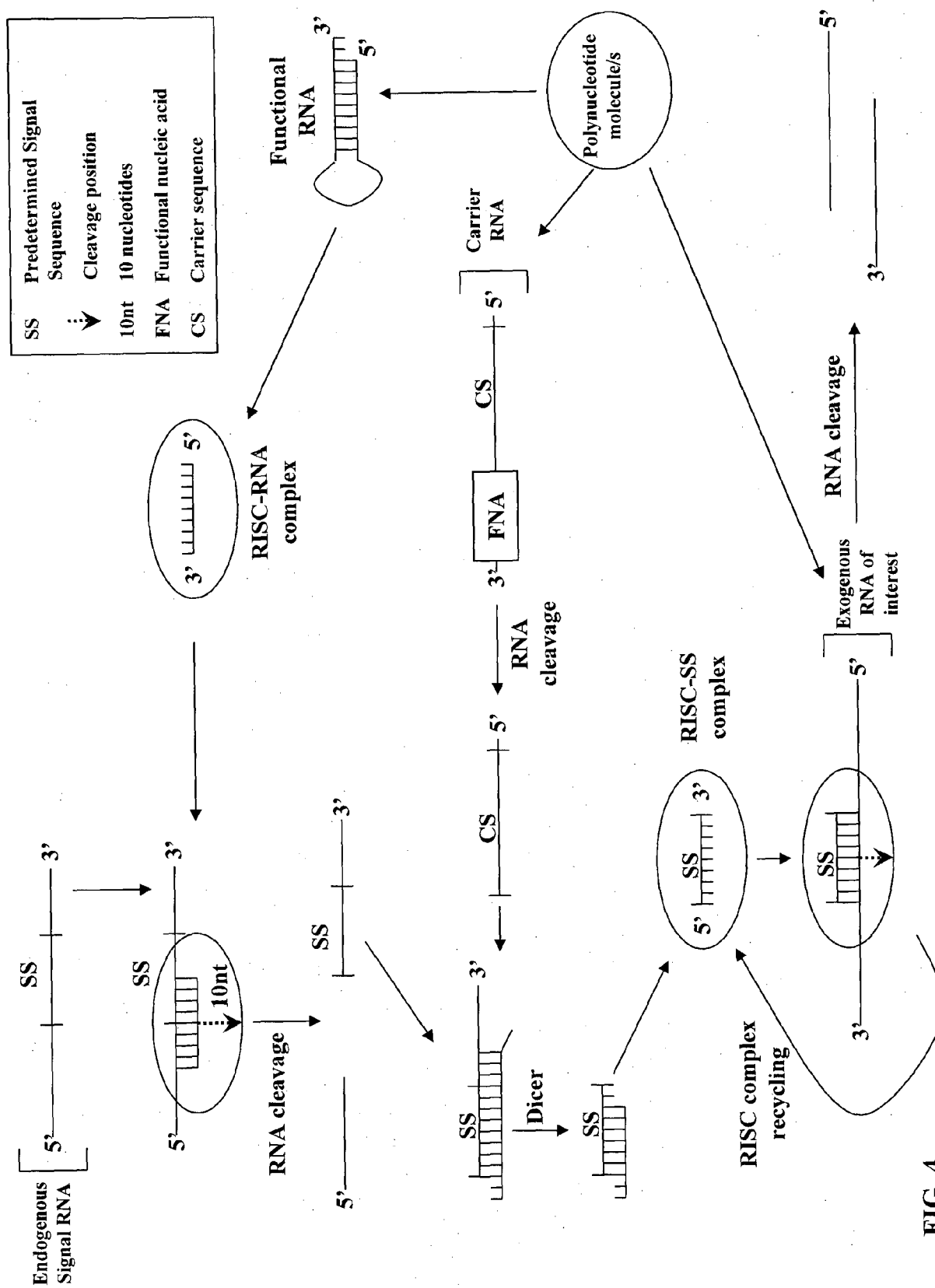


FIG. 4

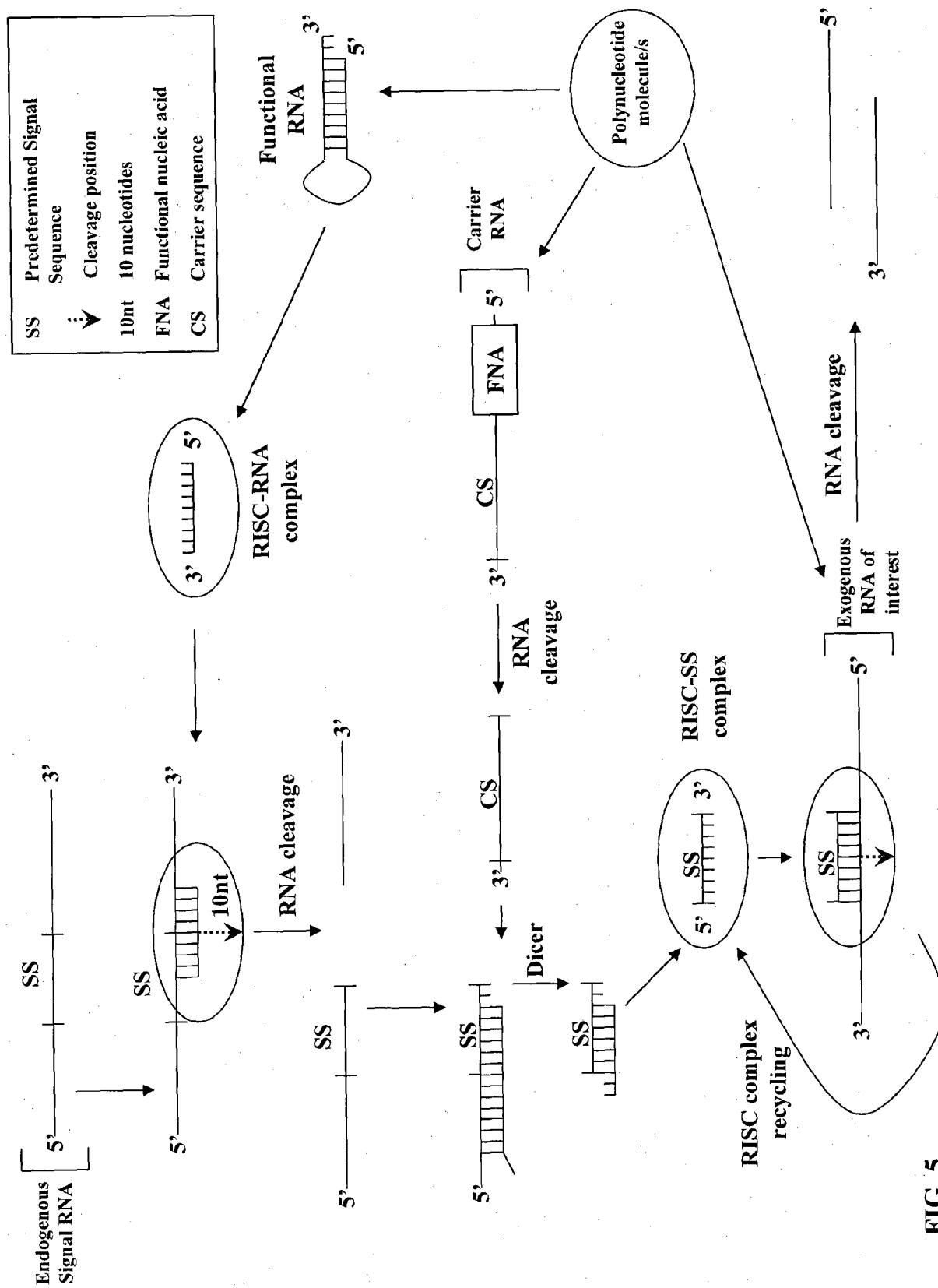
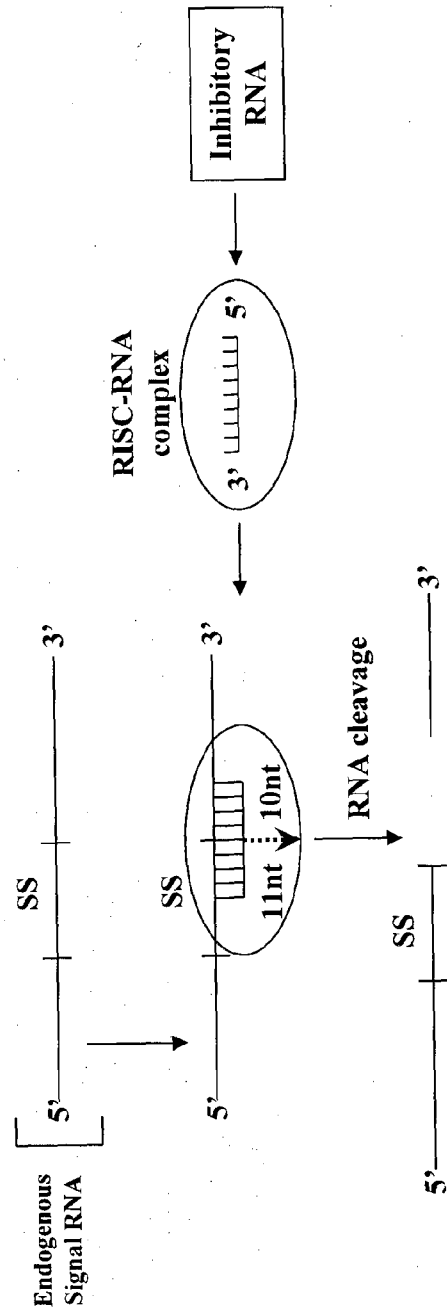
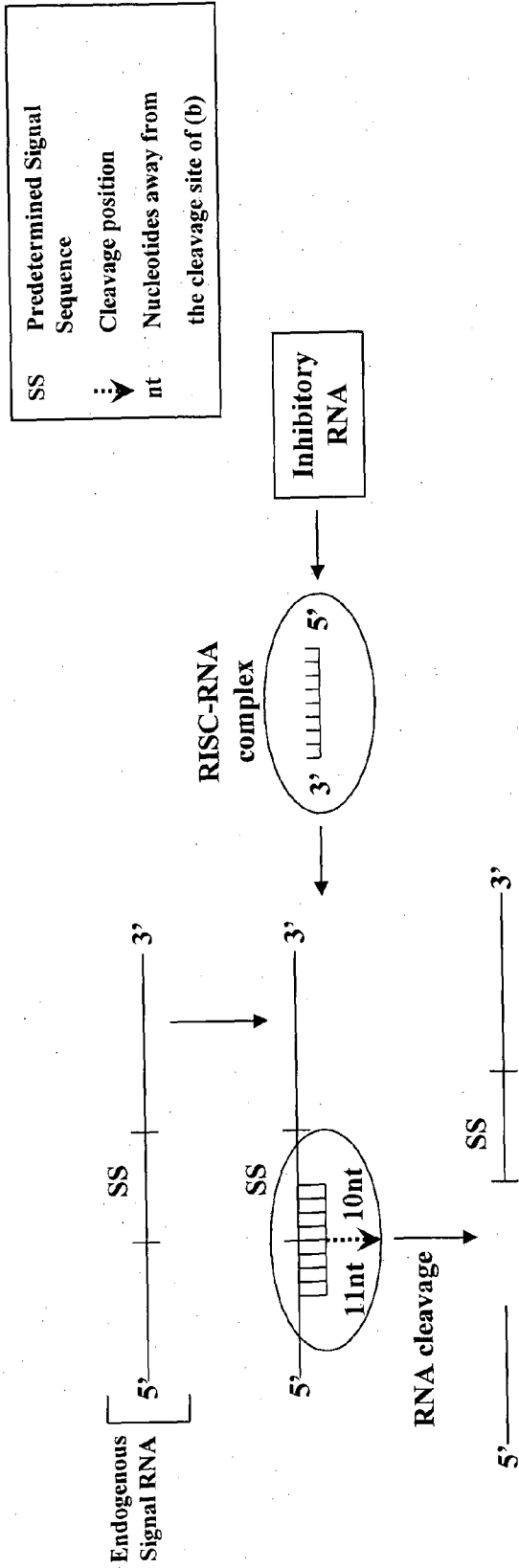


FIG. 5



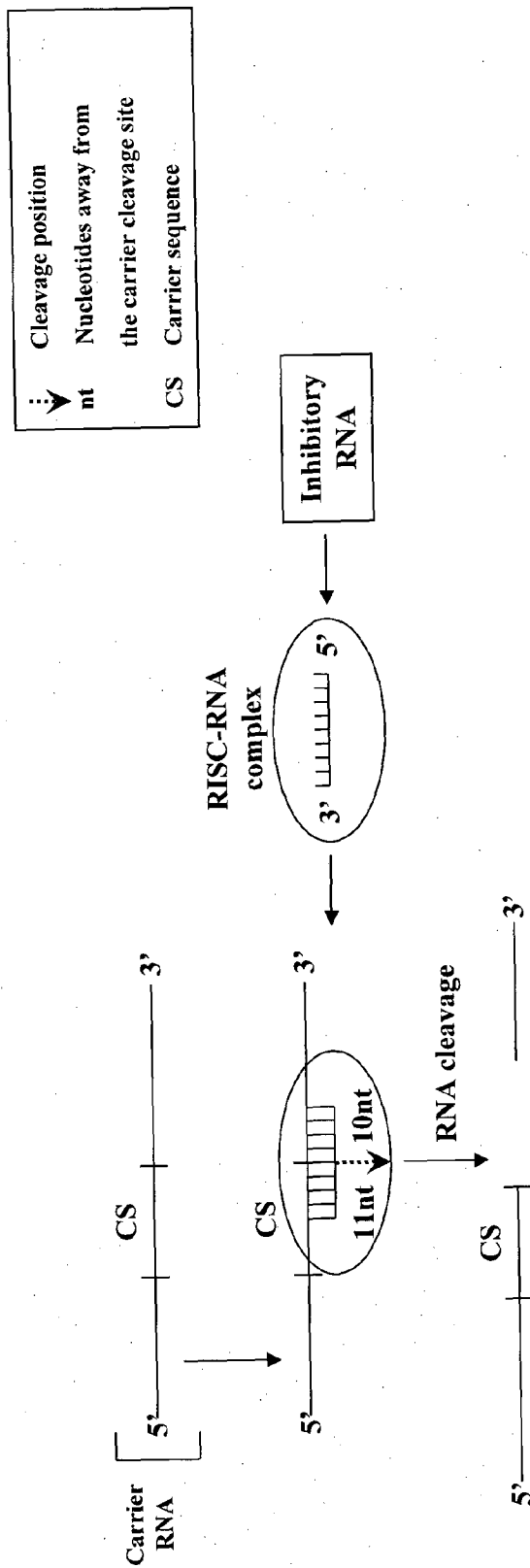


FIG. 7A

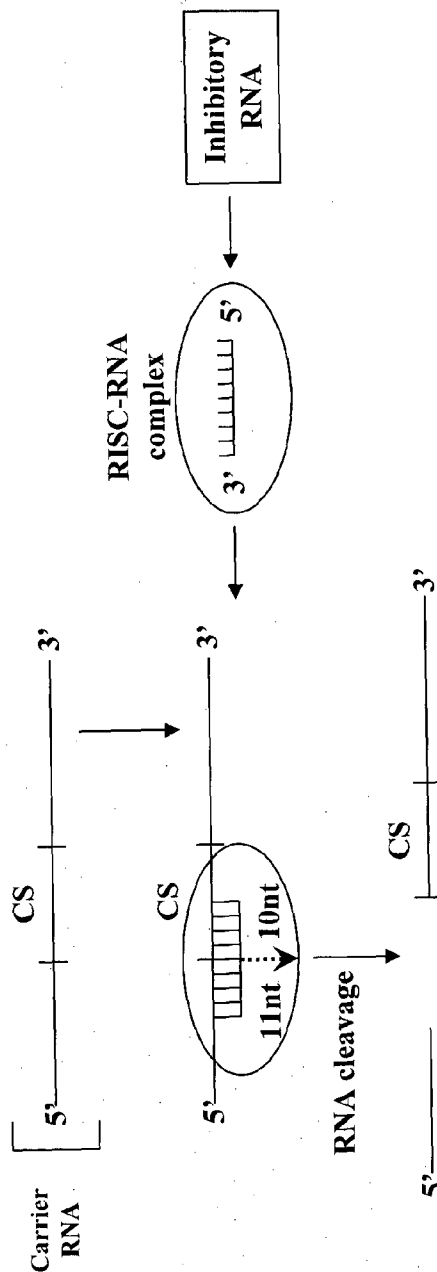


FIG. 7B

NAS Nucleic acid sequence of 21 nucleotides long that is complementary to the target sequence of (i)

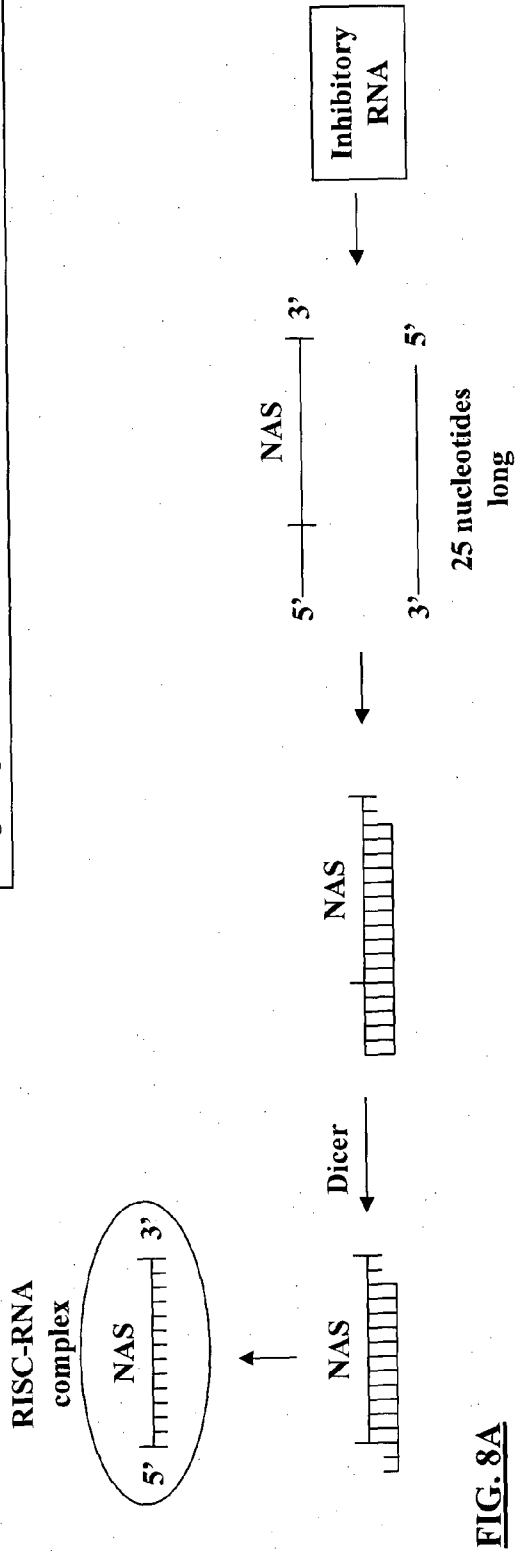


FIG. 8A

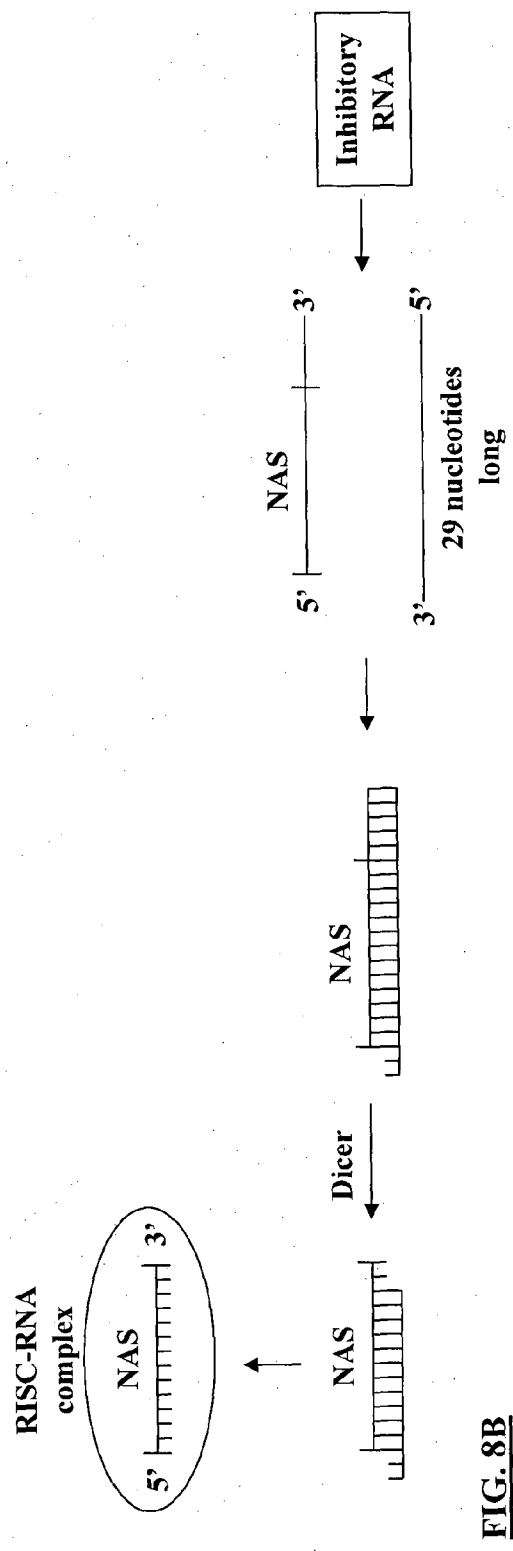
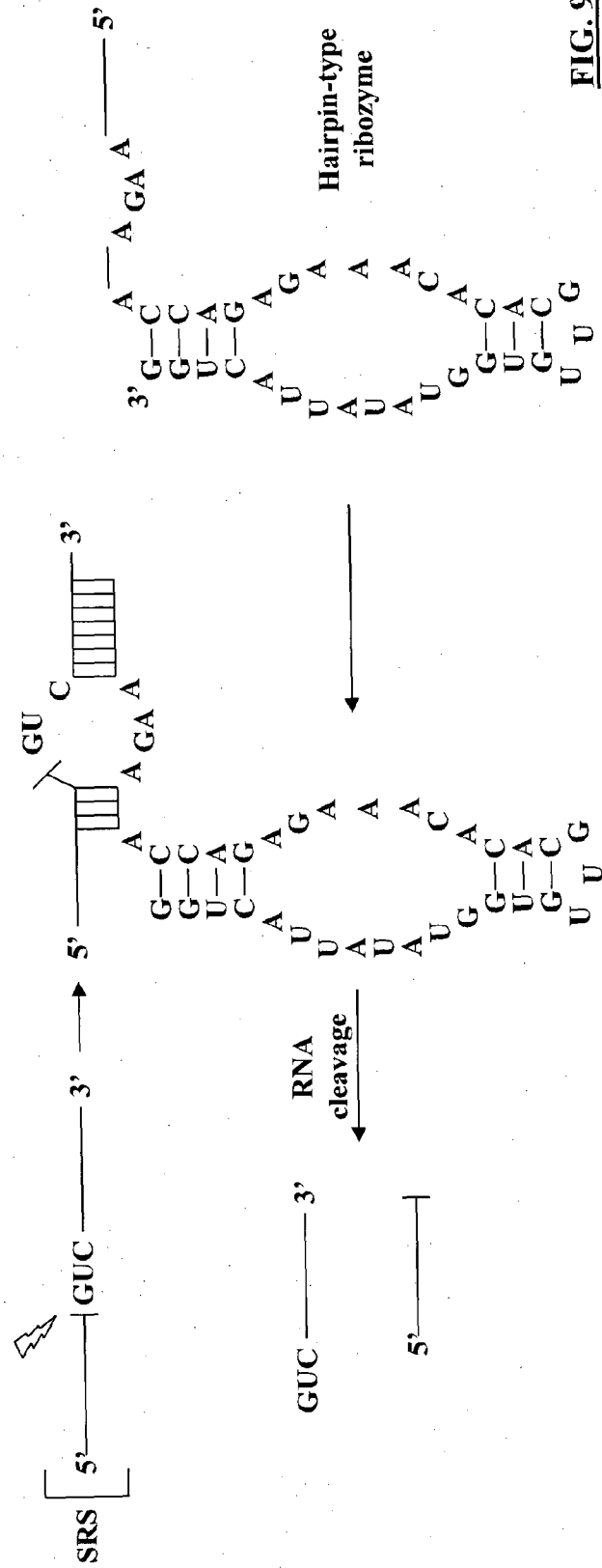
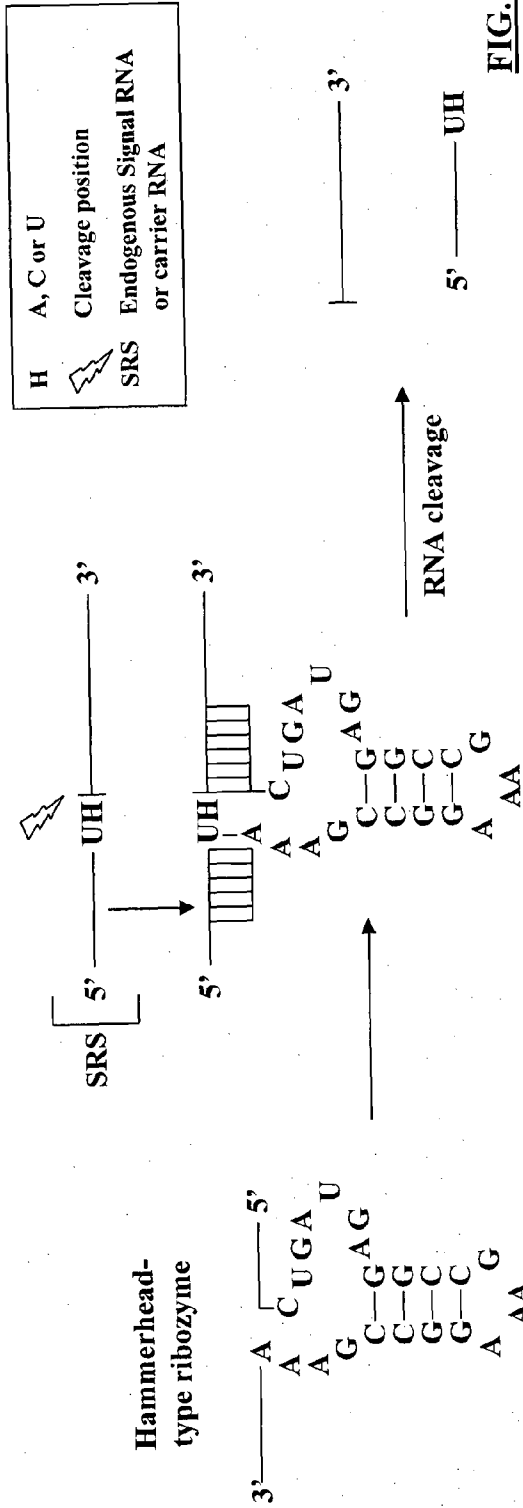


FIG. 8B



Cleavage position

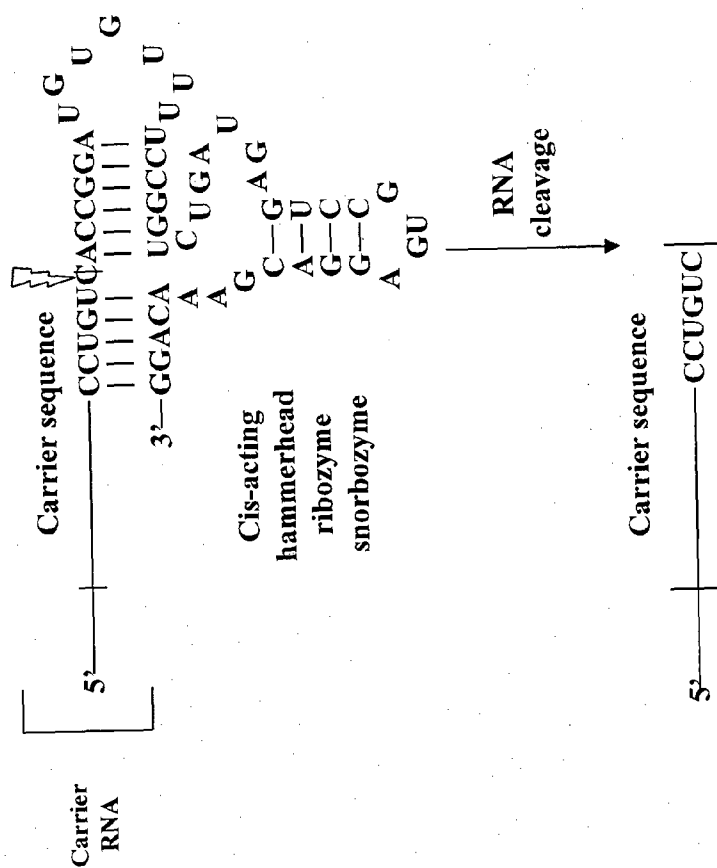


FIG. 10

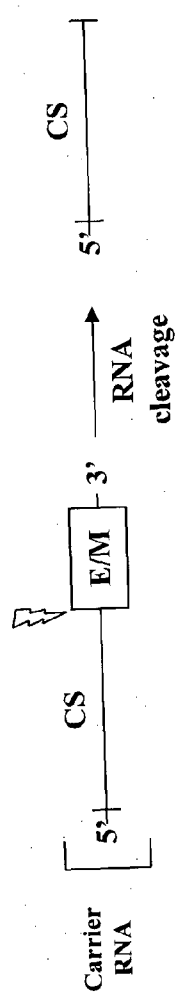
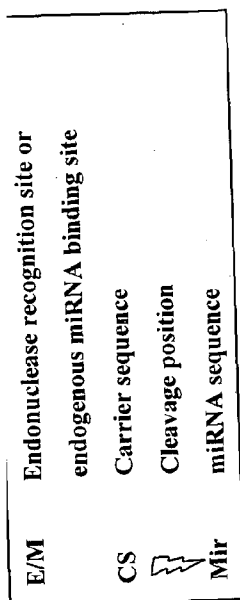


FIG. 12A

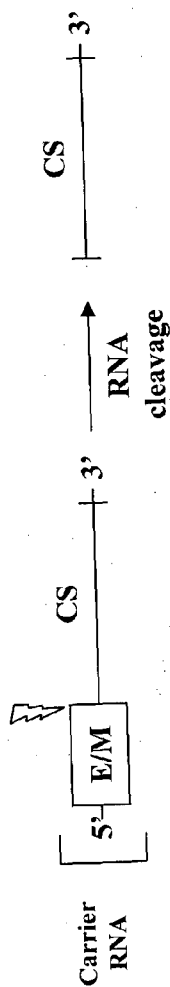


FIG. 12B

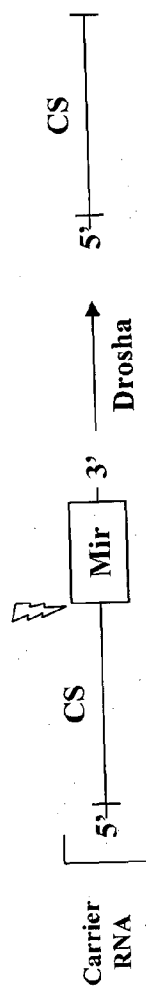


FIG. 12C

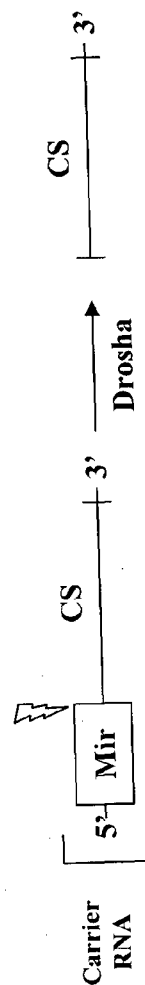
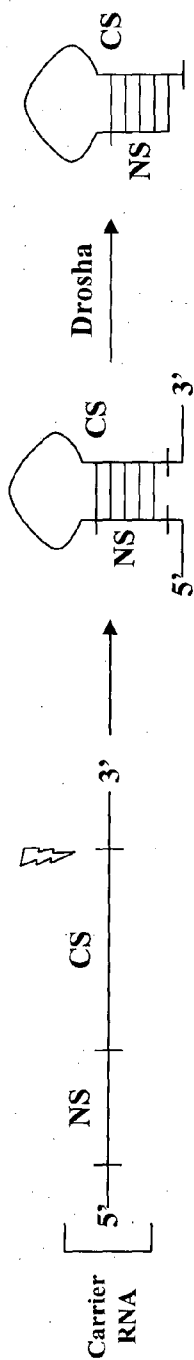
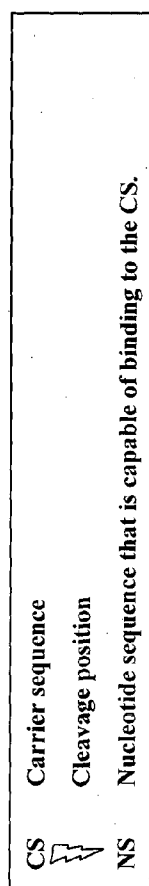
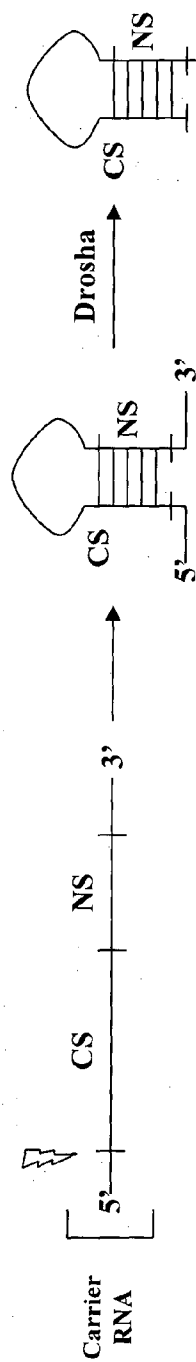


FIG. 12D

**FIG. 13A****FIG. 13B**

CS	Carrier sequence	N1, N2	Nucleotide sequences that are capable of binding to each other.
RSOC	Carrier RNA	P-I/III	Polymerase I or III based promoter.

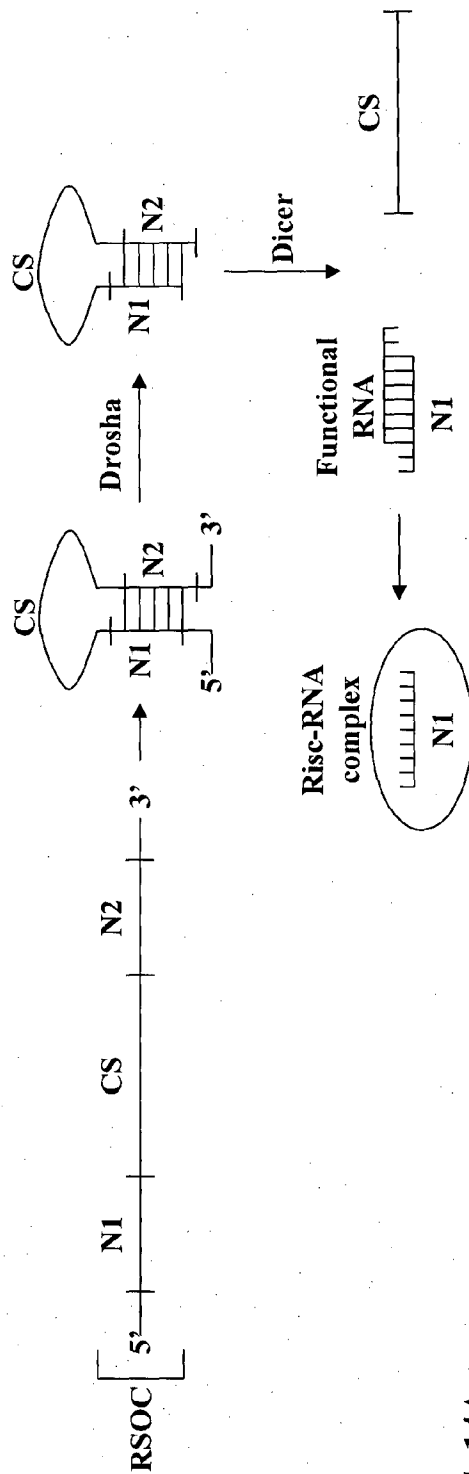


FIG. 14A

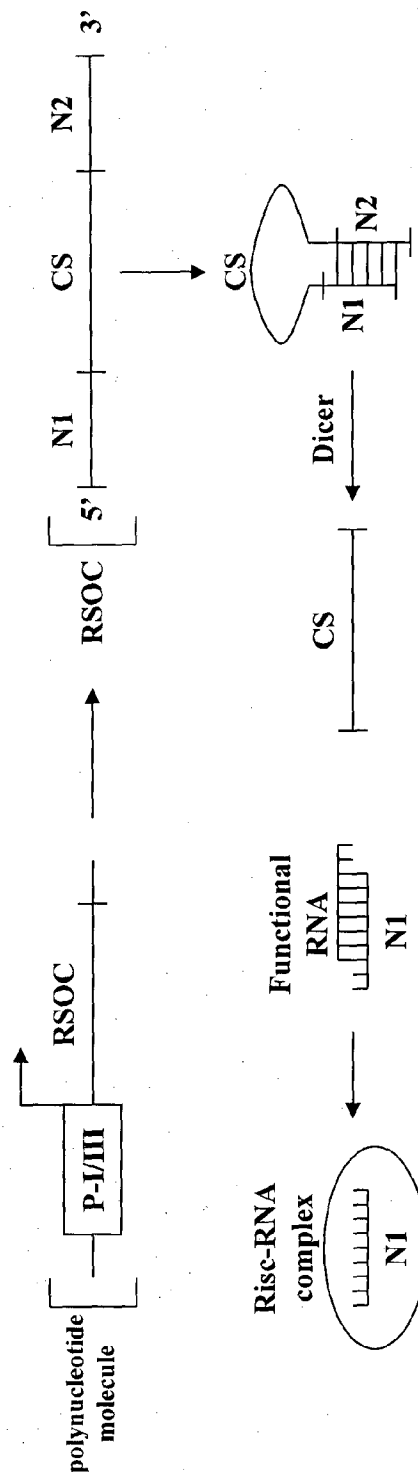


FIG. 14B

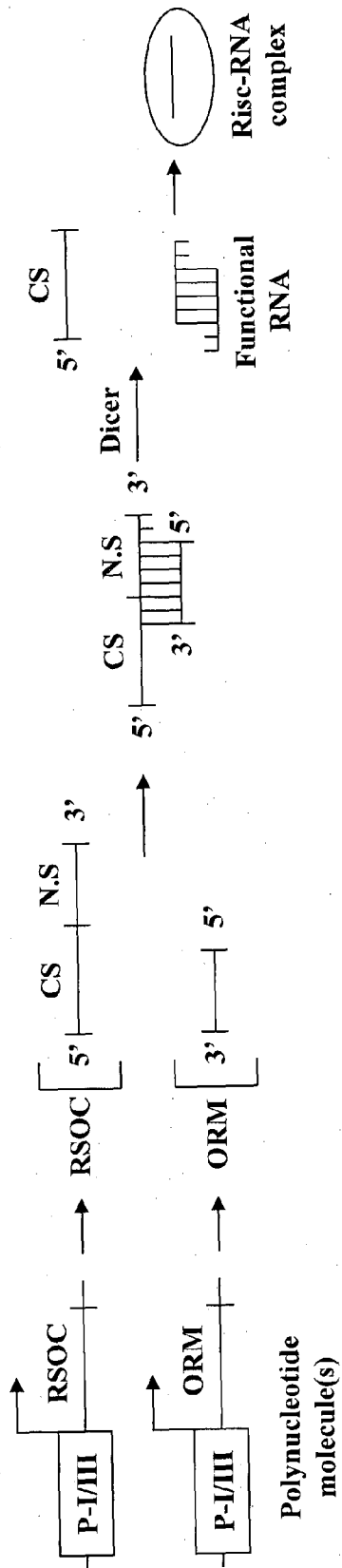
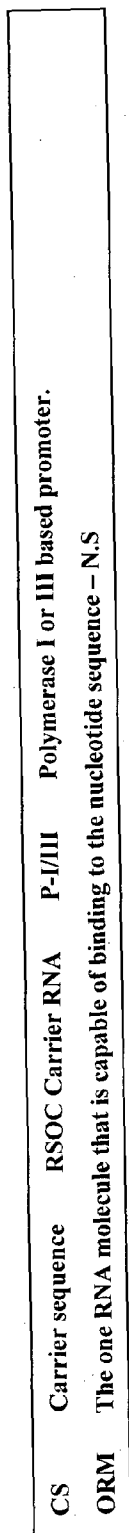


FIG. 15A

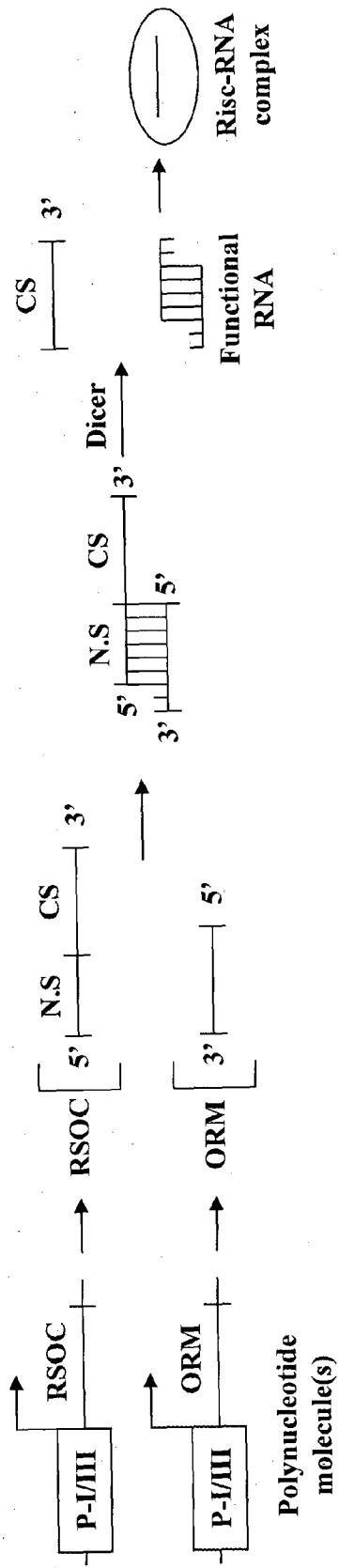


FIG. 15B

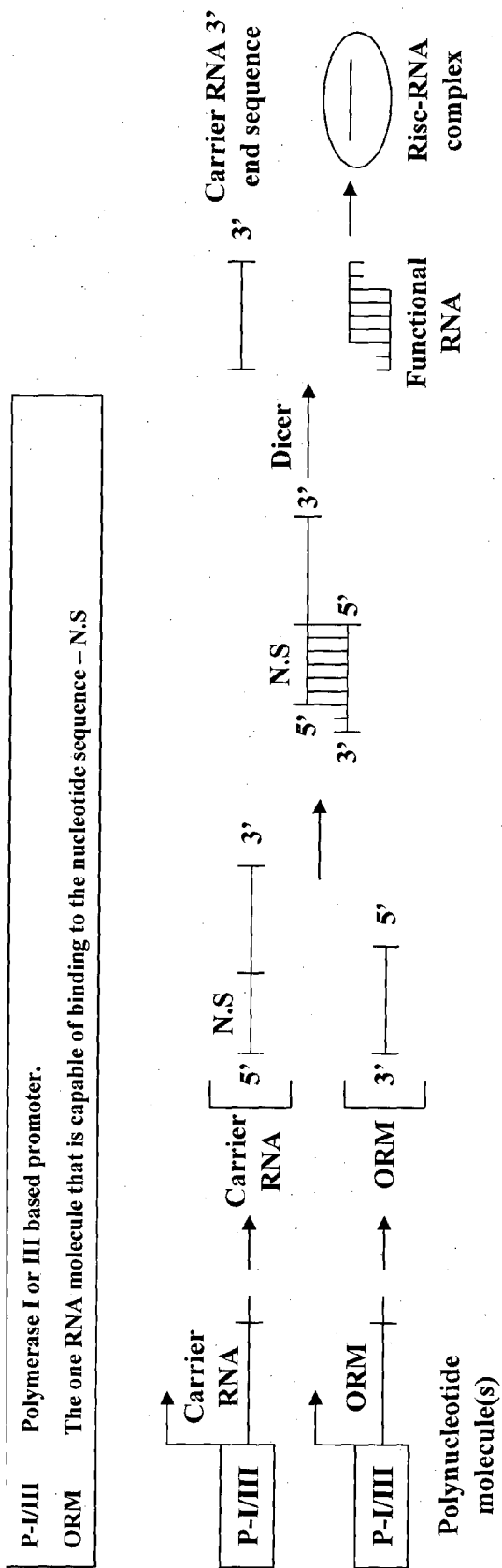


FIG. 16A

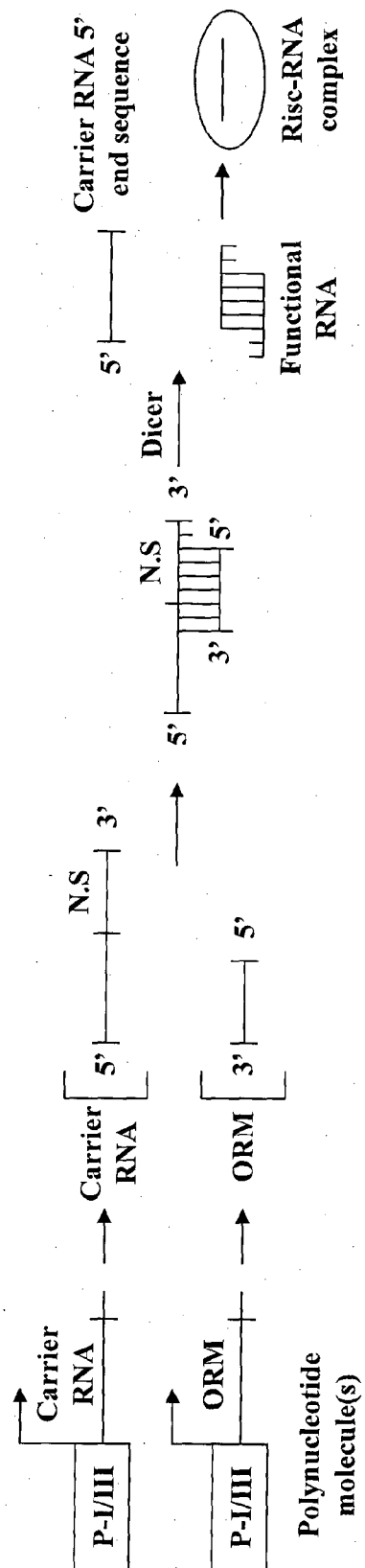
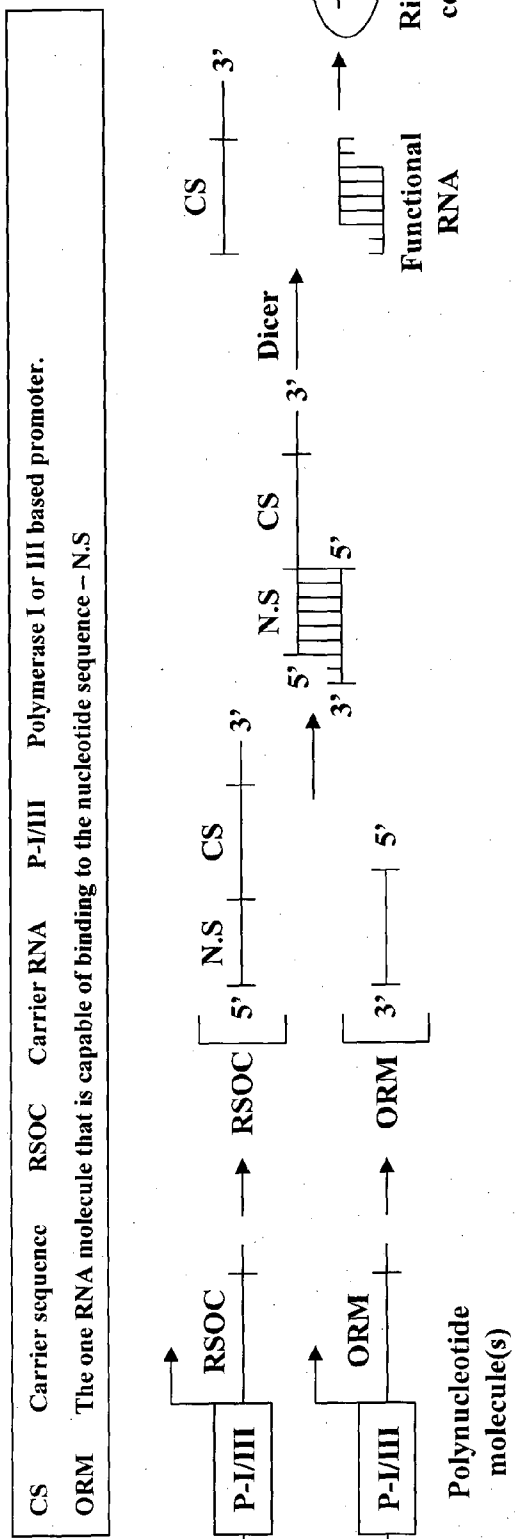
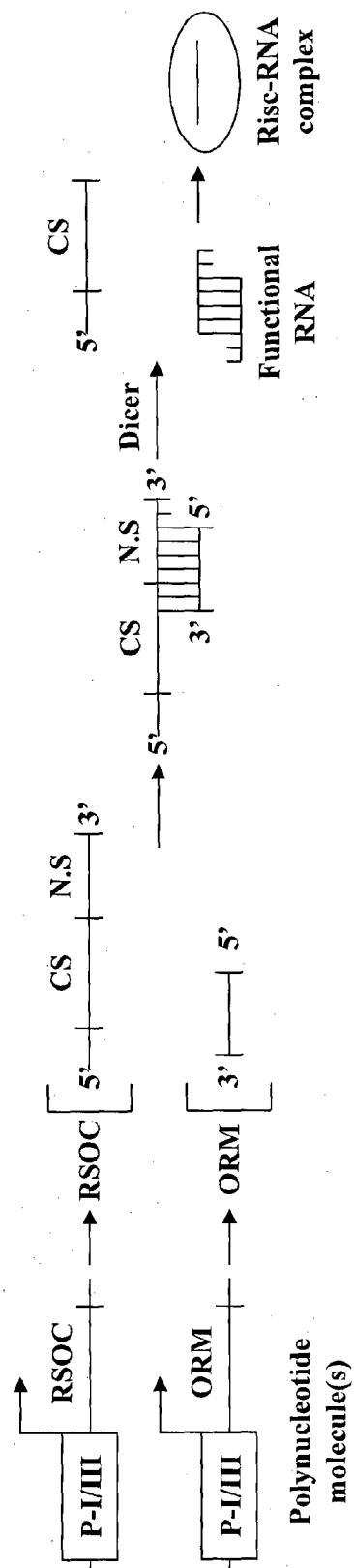


FIG. 16B

**FIG. 17A****FIG. 17B**

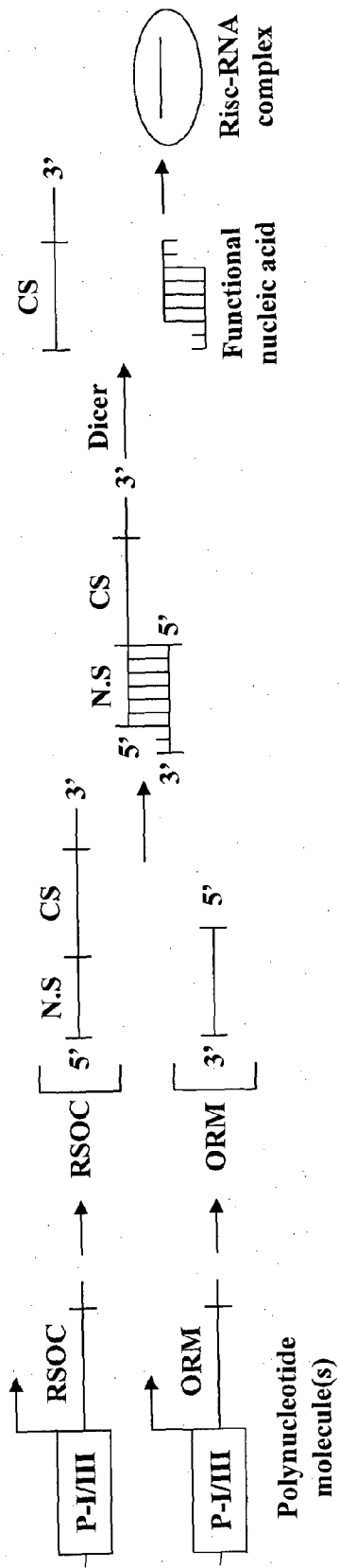
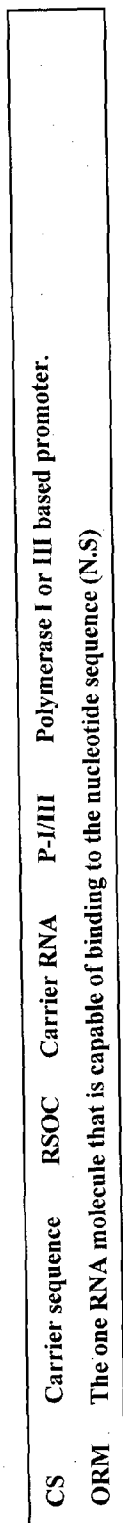


FIG. 18A

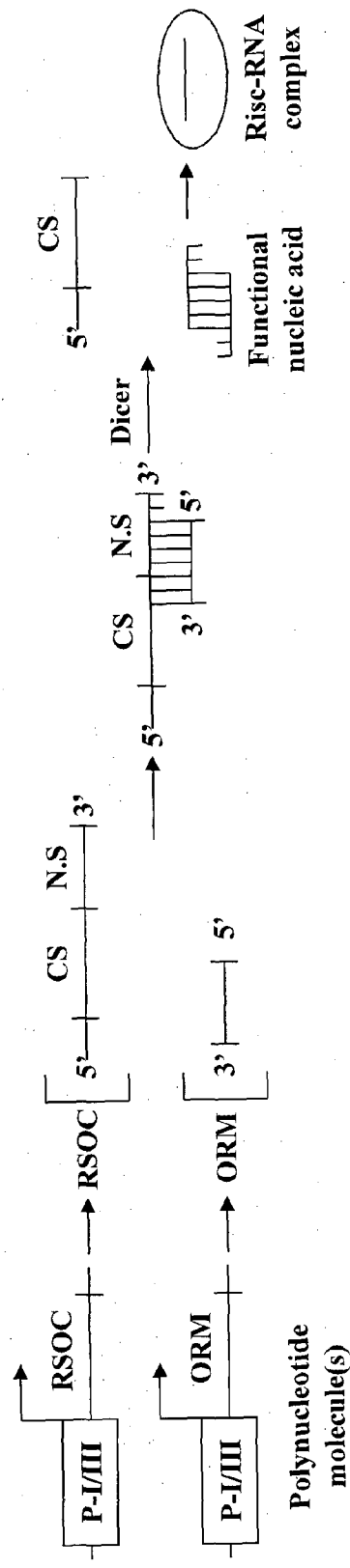


FIG. 18B

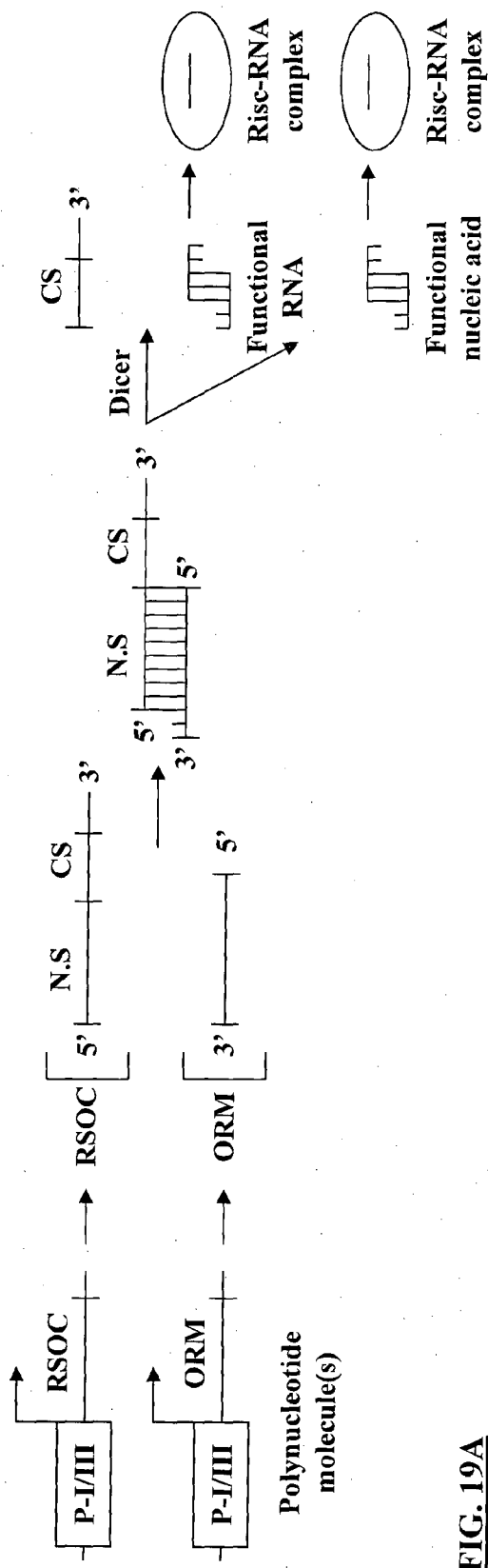
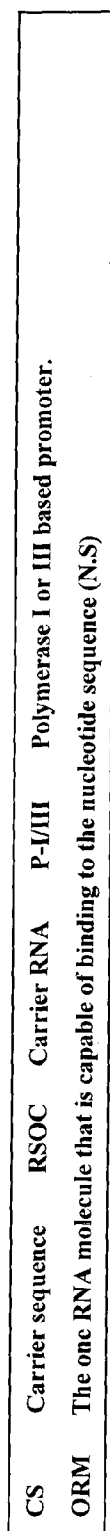


FIG. 19A

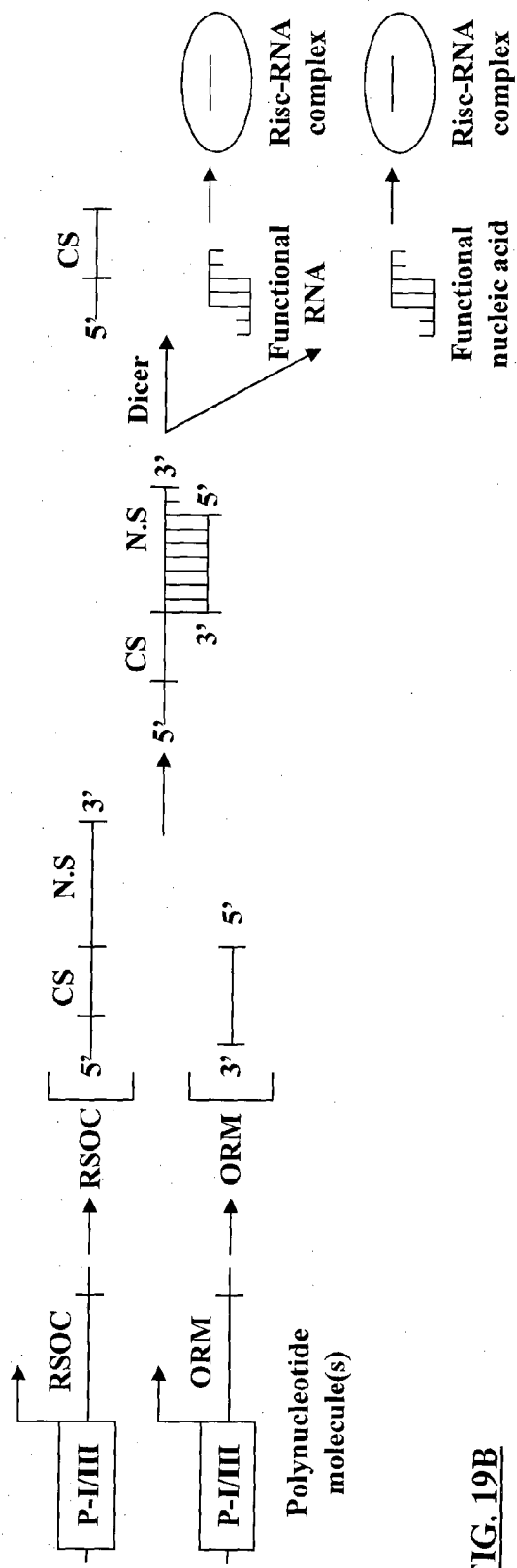


FIG. 19B

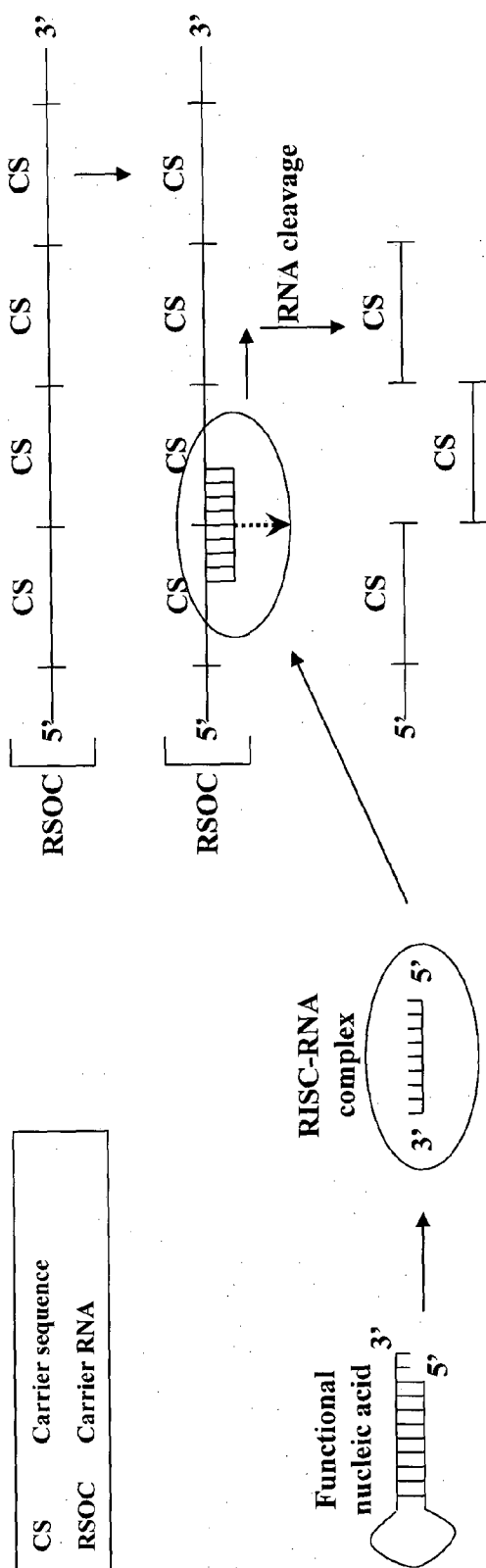


FIG. 20A

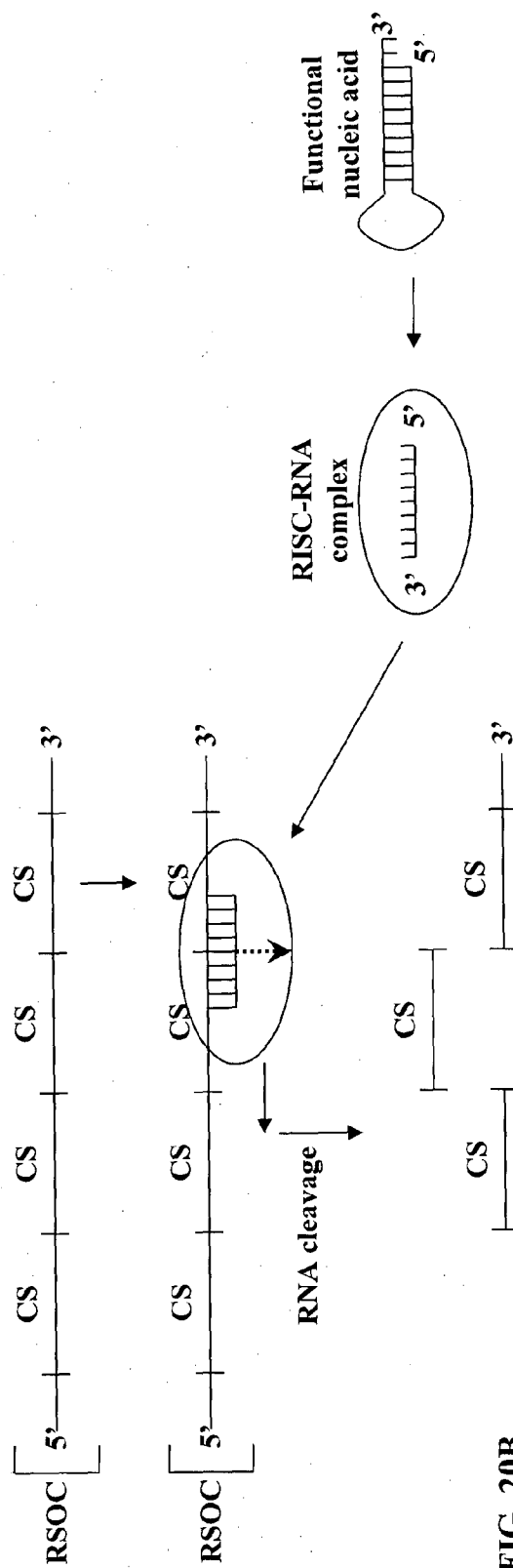


FIG. 20B

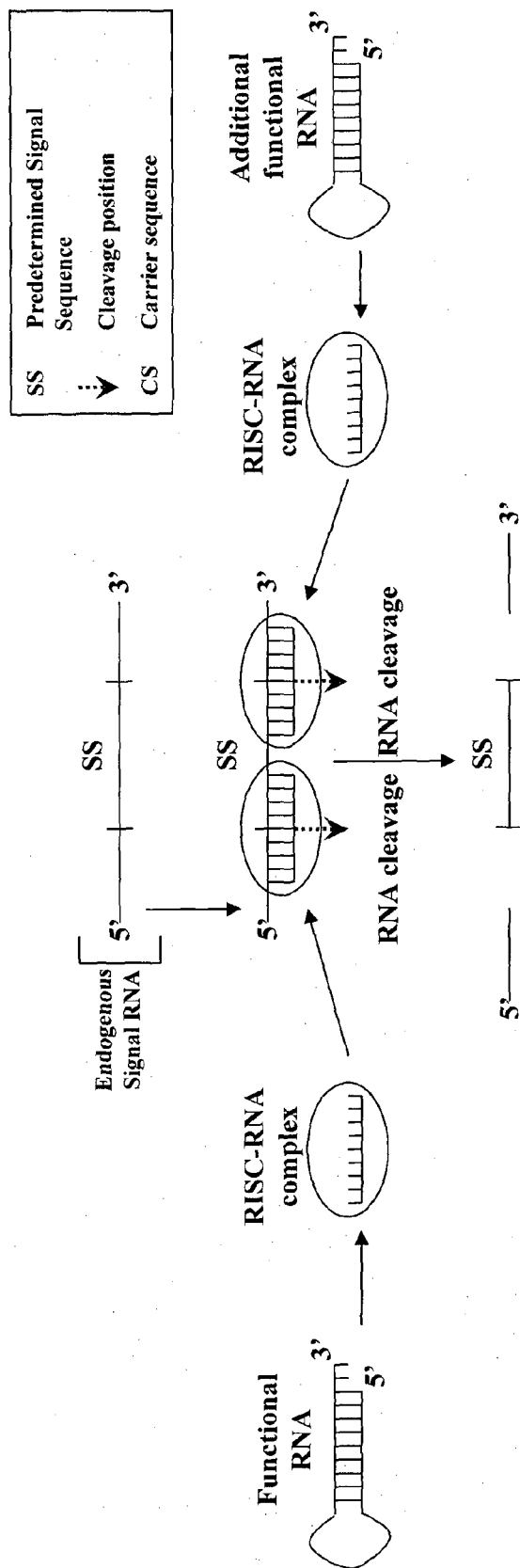


FIG. 21A

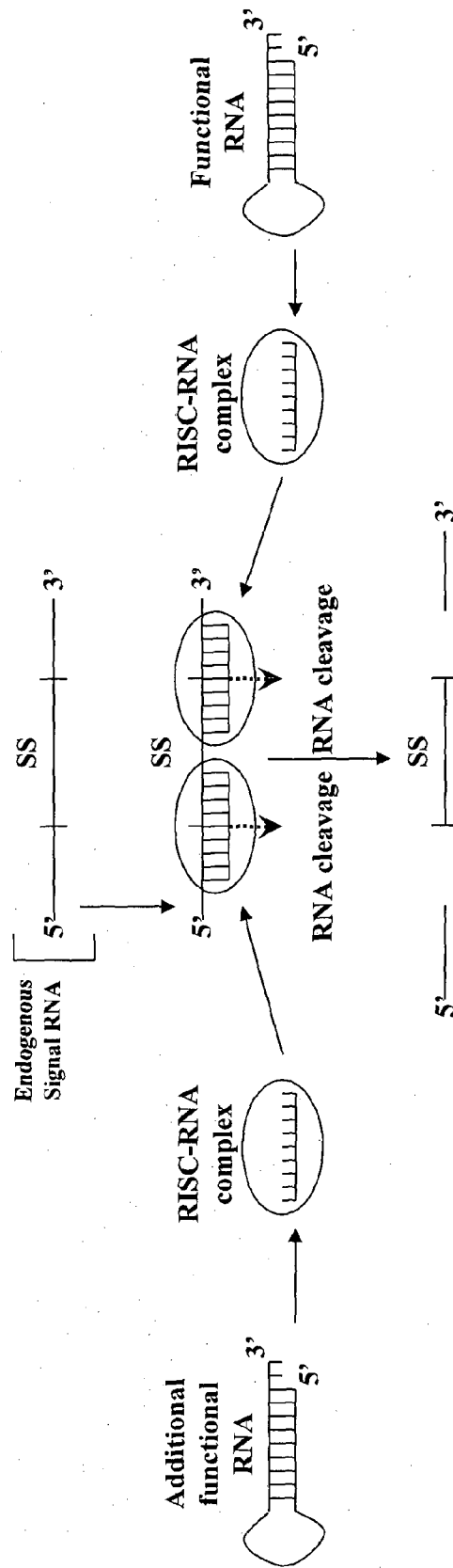
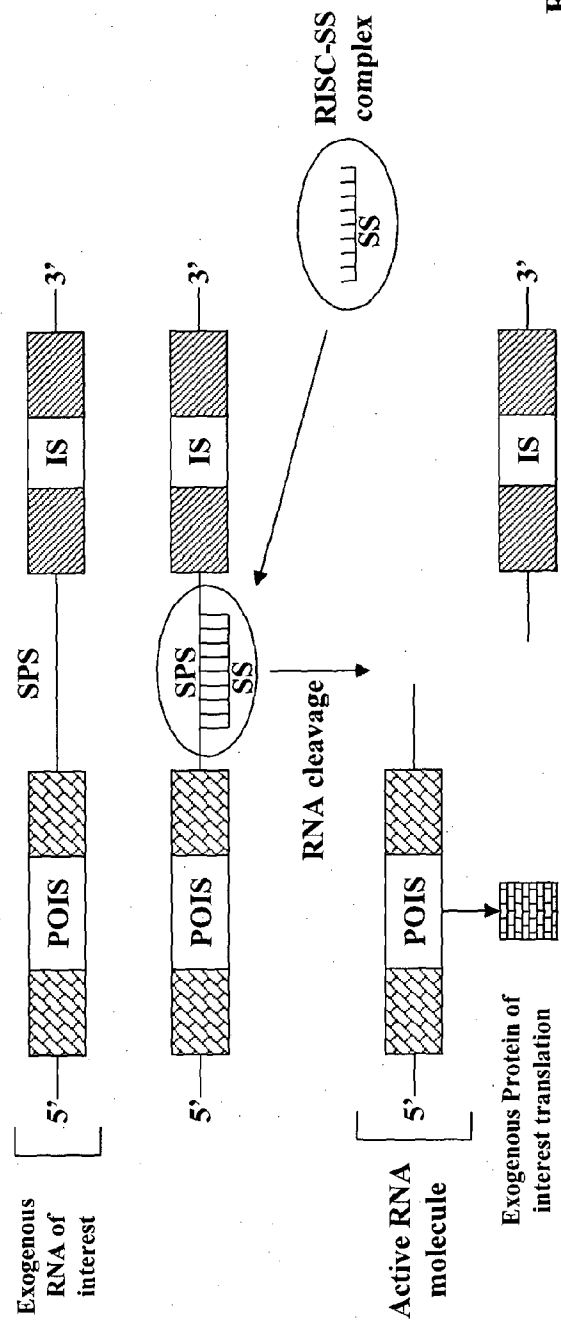
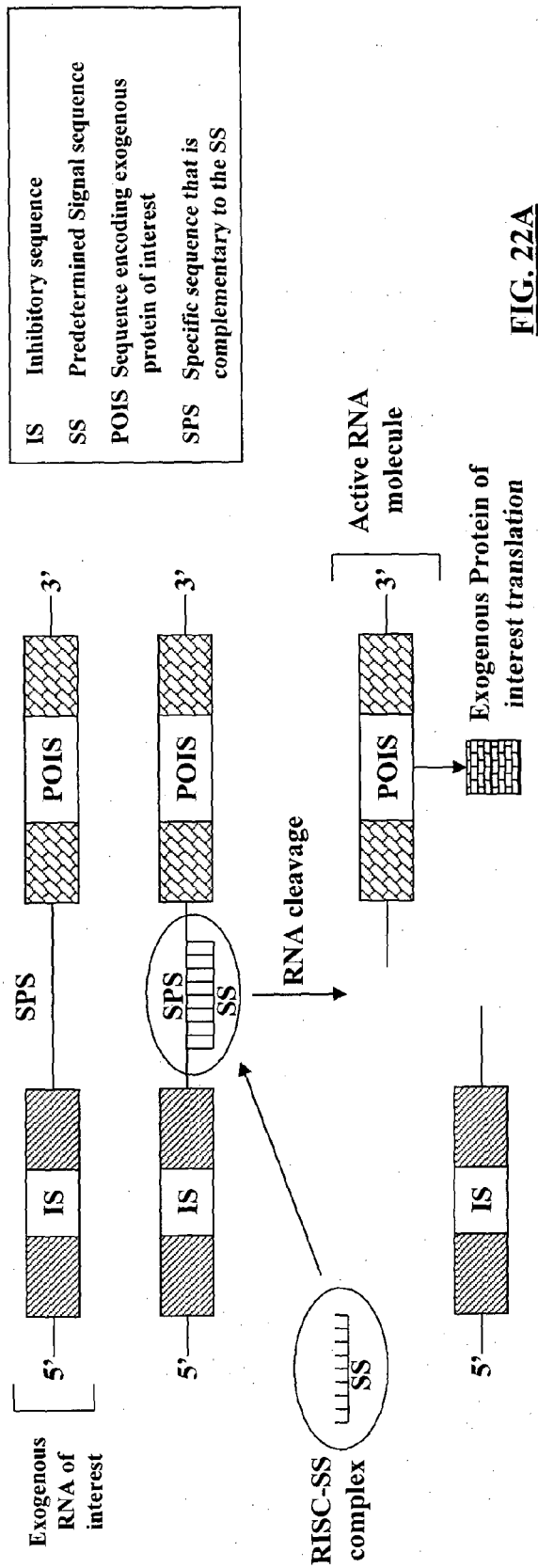


FIG. 21B



SPS	Specific sequence that is complementary to the predetermined signal sequence
POIS	Sequence encoding exogenous protein of interest
ND3	Number of nucleotides that is not evenly divided by 3

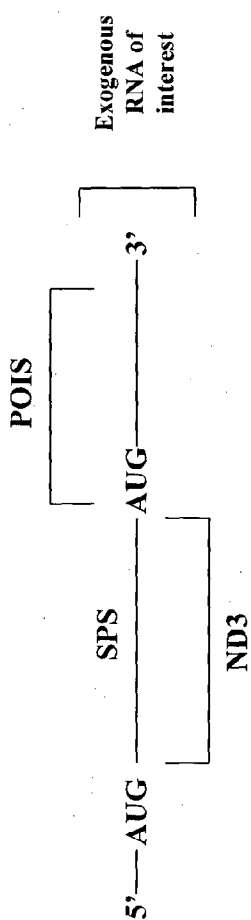


FIG. 23A

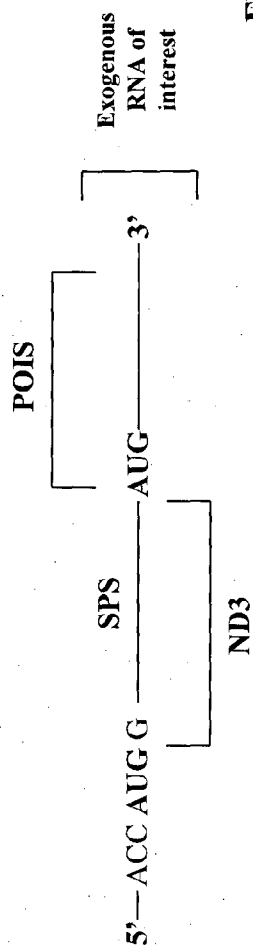


FIG. 23B

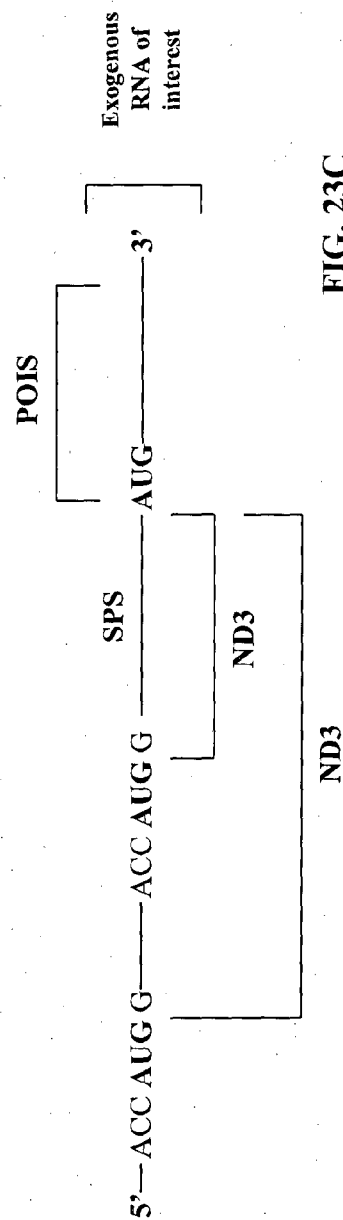
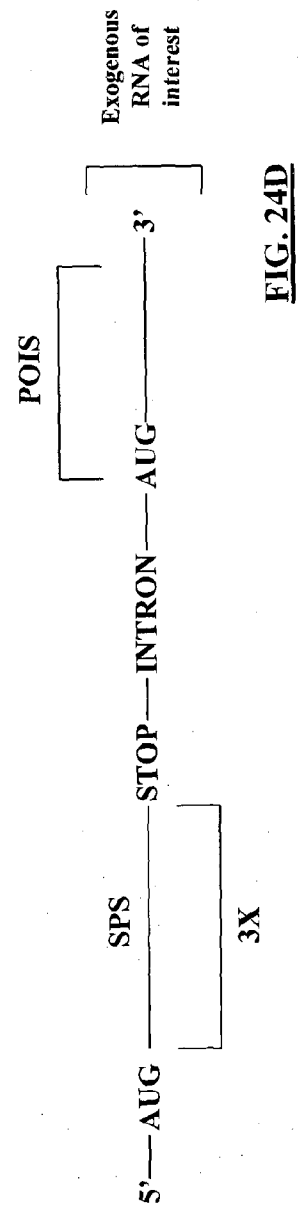
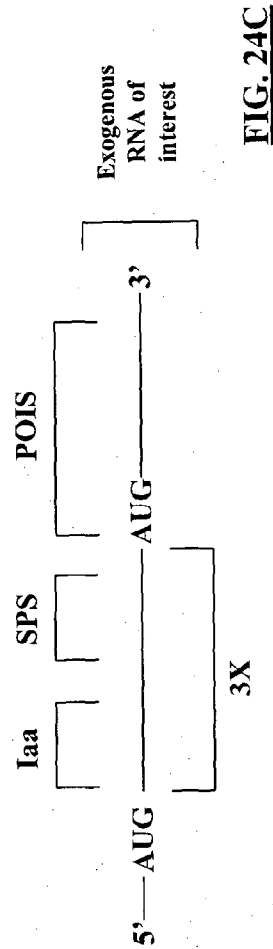
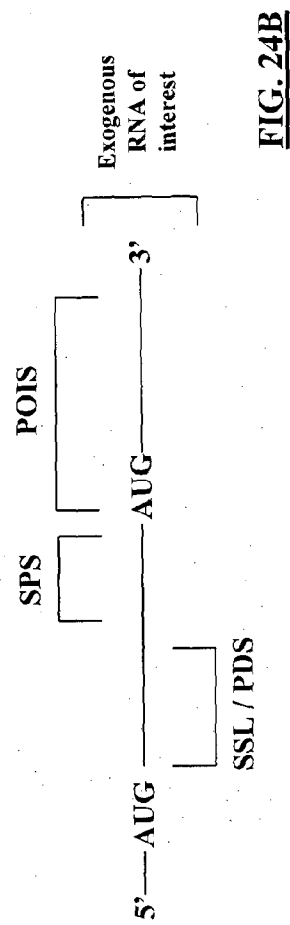
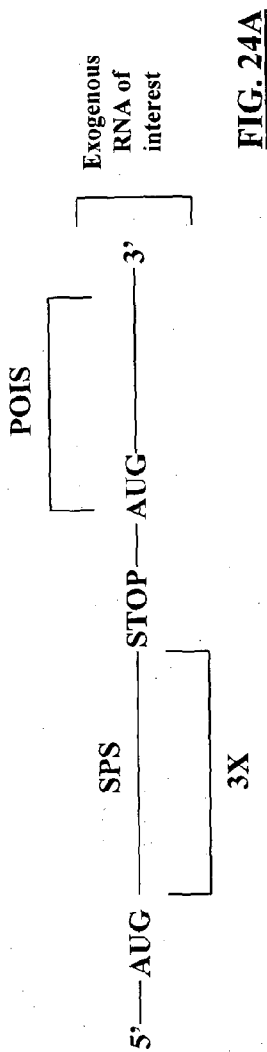


FIG. 23C

SPS	Specific sequence that is complementary to the predetermined signal sequence
POIS	Sequence encoding exogenous protein of interest
3X	Number of nucleotides that is evenly divided by 3
STOP	5'-UAA-3' or 5'-UAG-3' or 5'-UGA-3'
SSL	Sequence encoding sorting signal for subcellular localization
PDS	Sequence encoding Protein degradation signal
Iaa	Sequence encoding amino acids that are capable of inhibiting the biological function of the protein of interest



SPS	Specific sequence that is complementary to the predetermined signal sequence
POIS	Sequence encoding exogenous protein of interest
TR	Binding site for translation repressor
RLS	RNA localization signal for subcellular localization
RDE	RNA destabilizing element that is AU-rich element or endonuclease recognition site
SES	Secondary structure

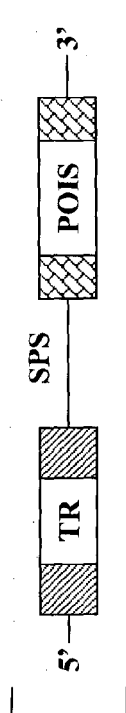


FIG. 25A

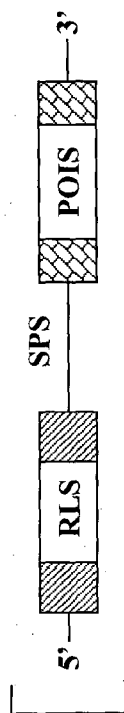


FIG. 25B

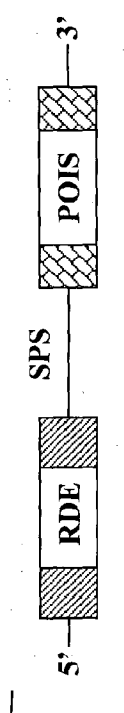


FIG. 25C

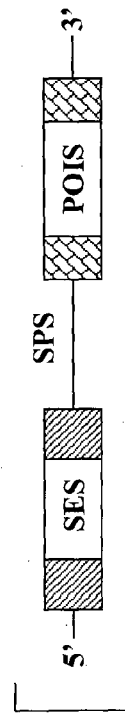
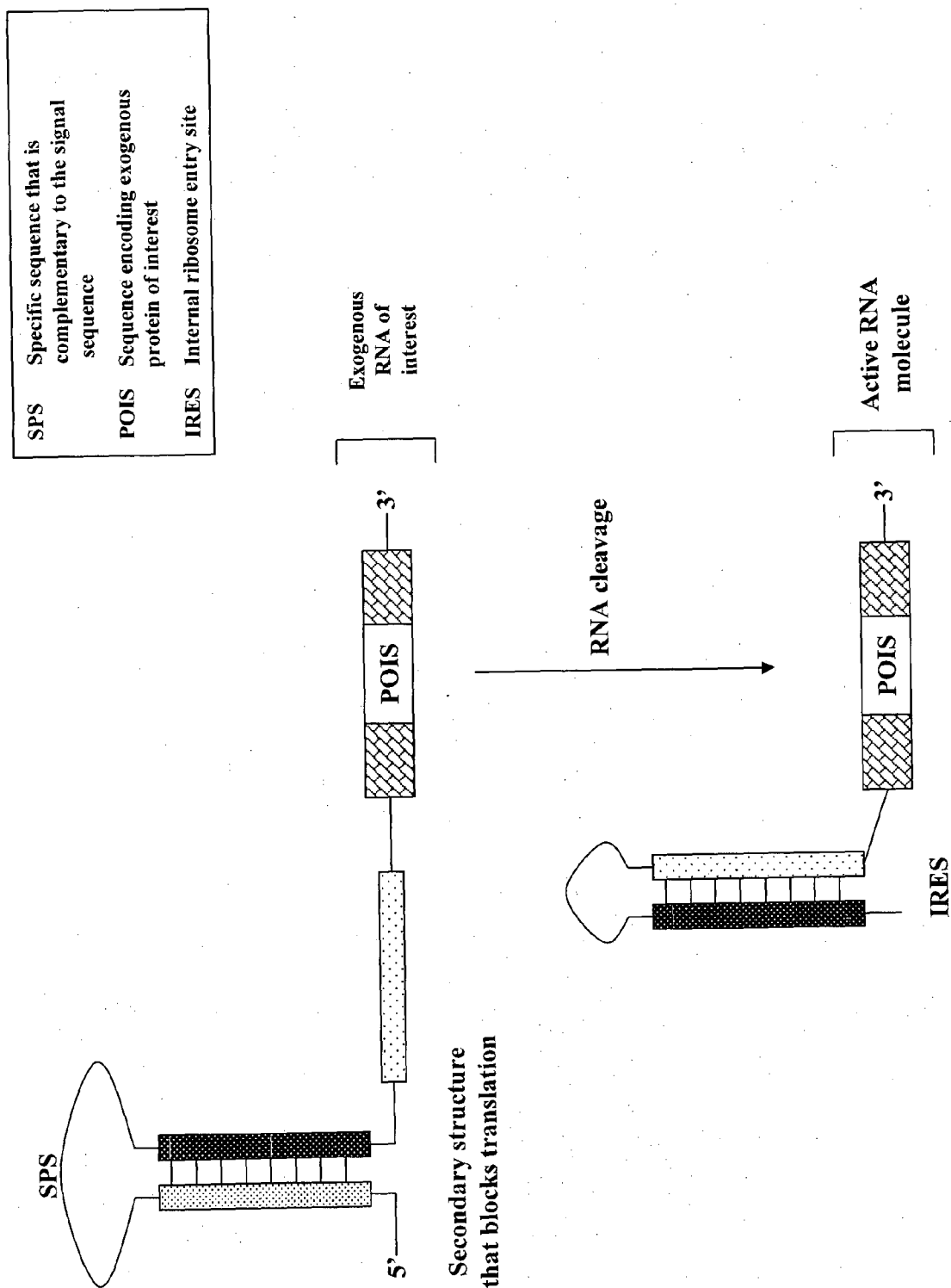


FIG. 25D

**FIG. 26**

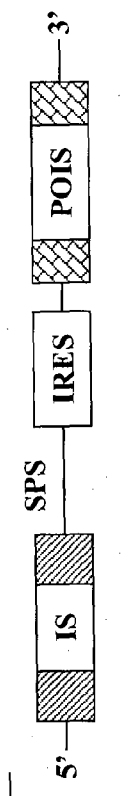
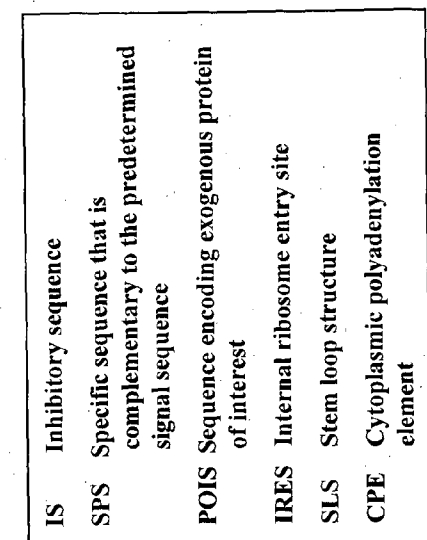


FIG. 27A

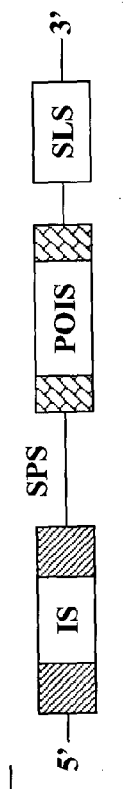


FIG. 27B

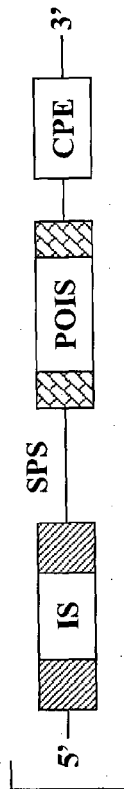


FIG. 27C

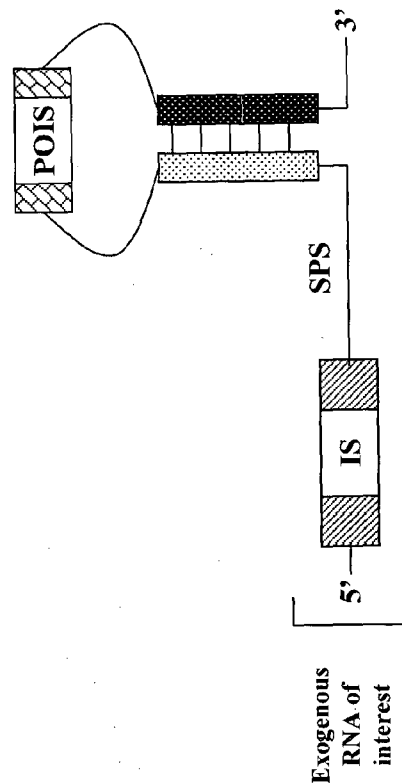


FIG. 27D

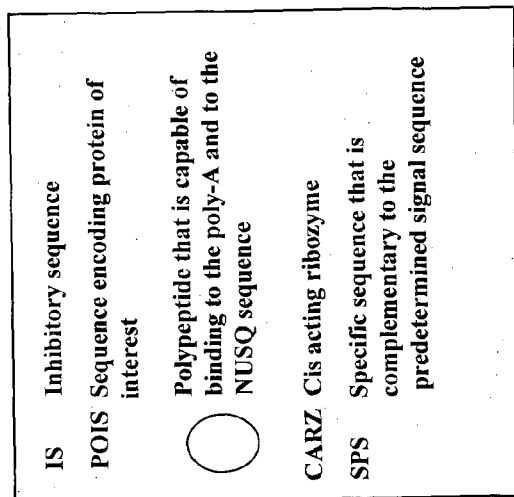


FIG. 28A

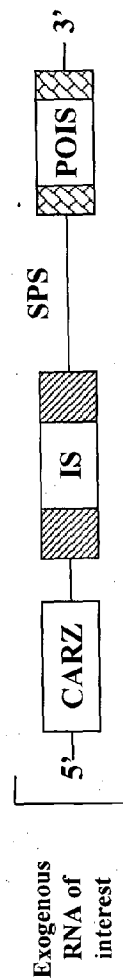
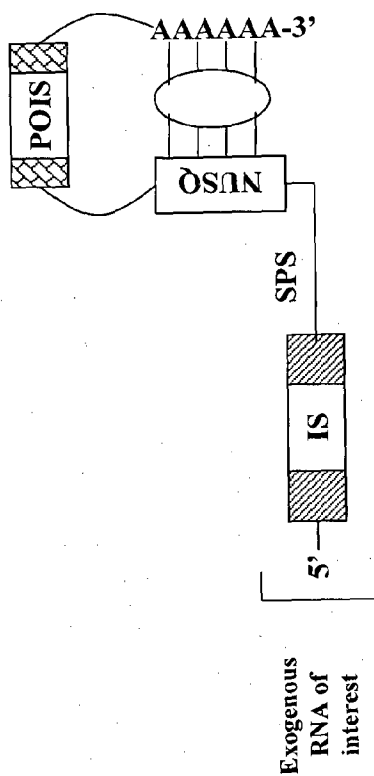


FIG. 28B

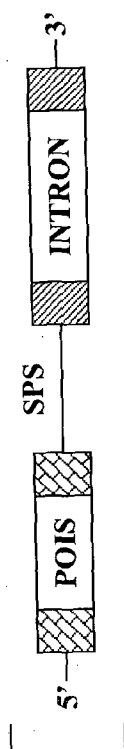
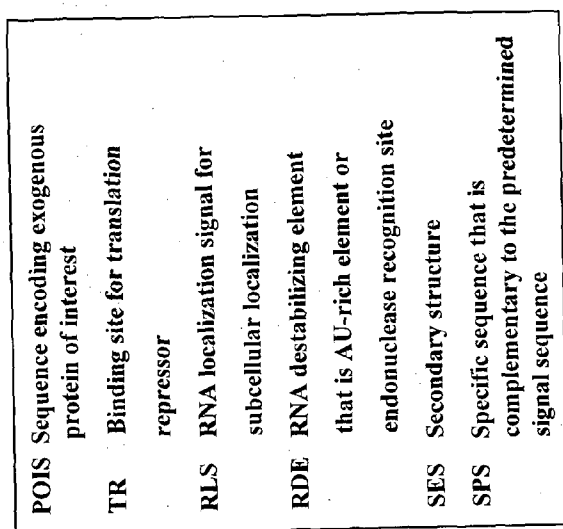


FIG. 29A

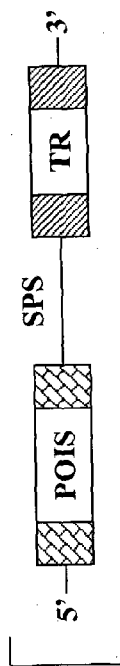


FIG. 29B

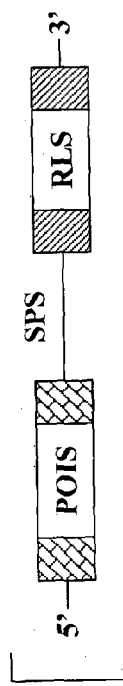


FIG. 29C

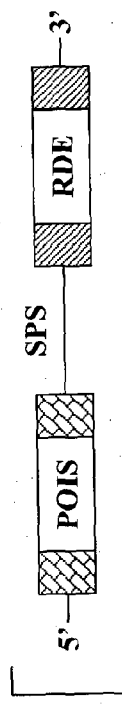


FIG. 29D

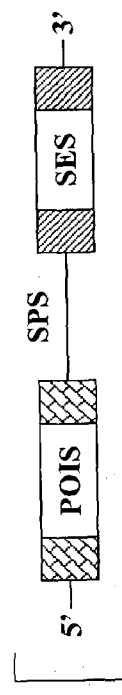
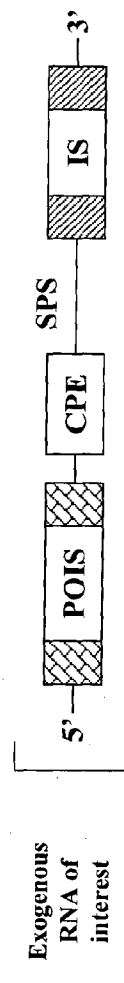
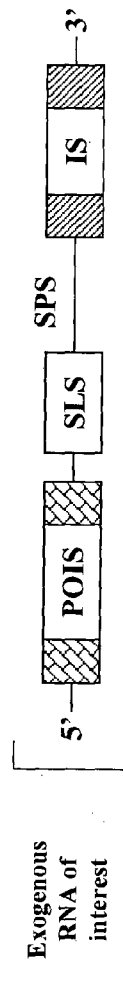
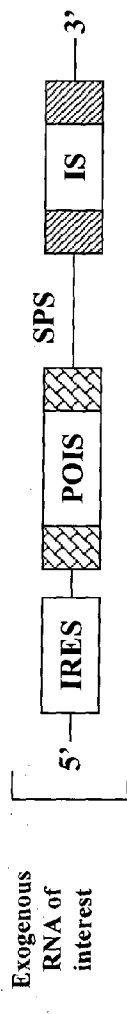
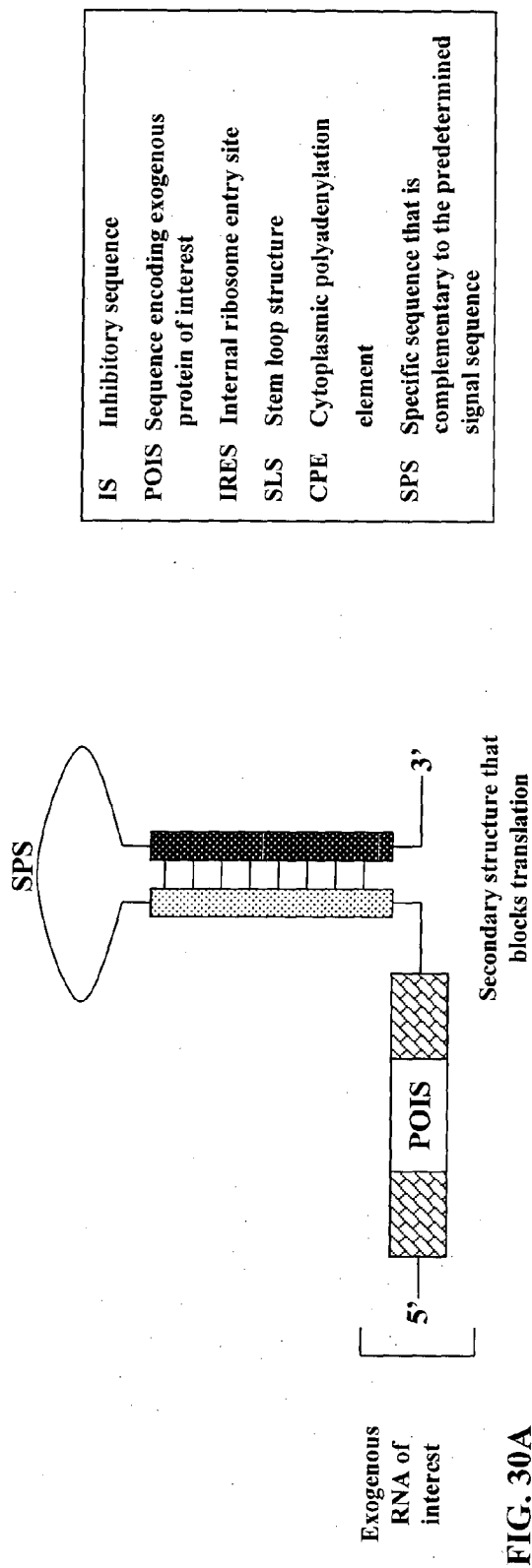
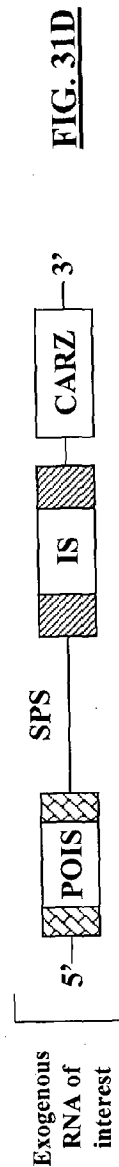
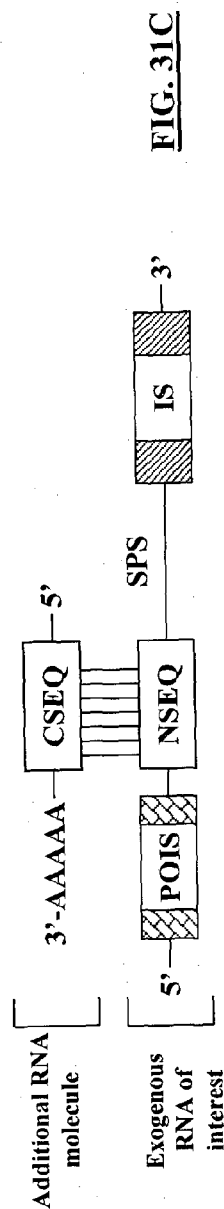
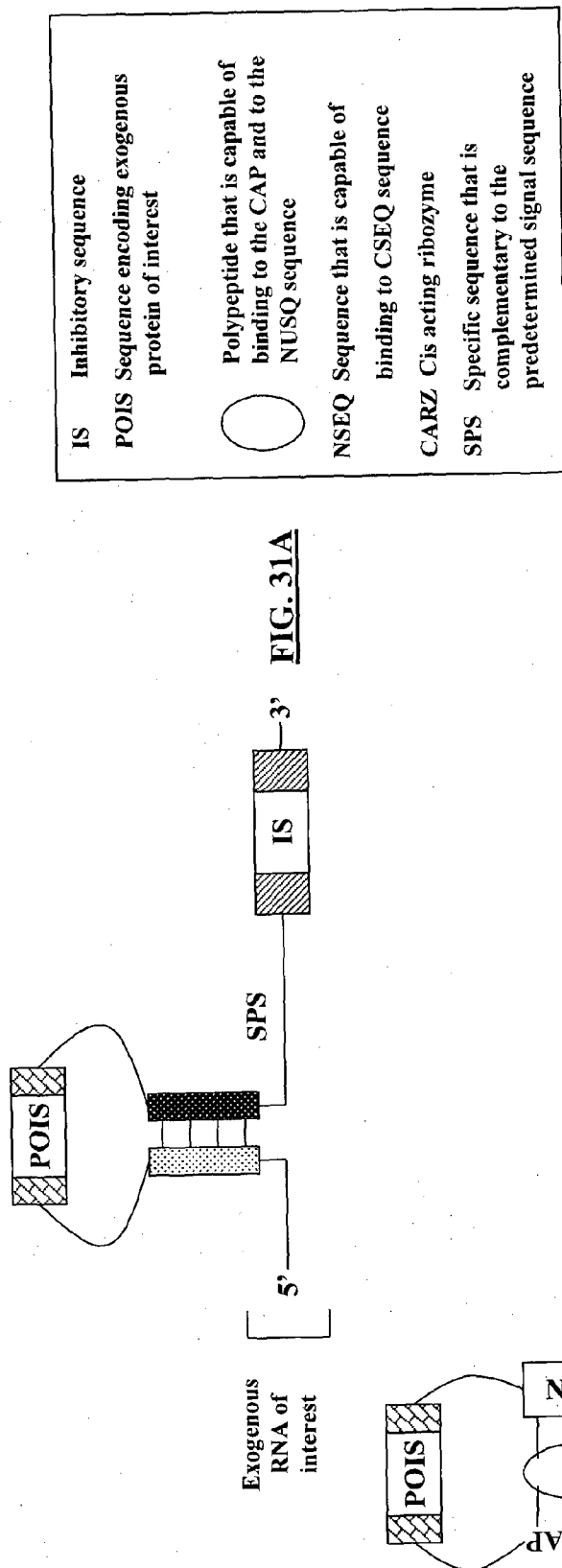


FIG. 29E





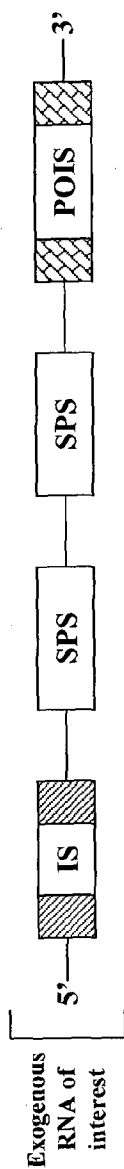
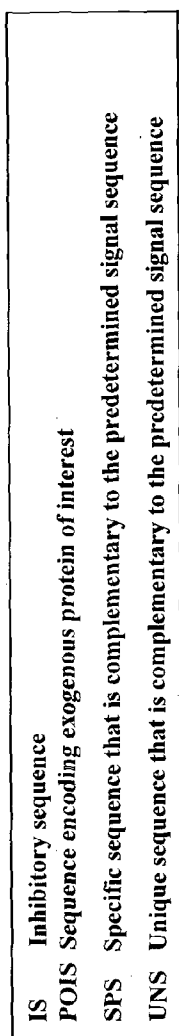


FIG. 32A

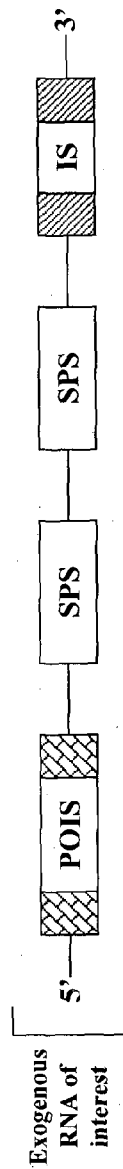


FIG. 32B

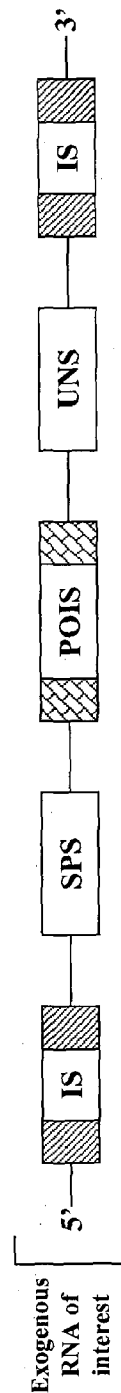


FIG. 32C

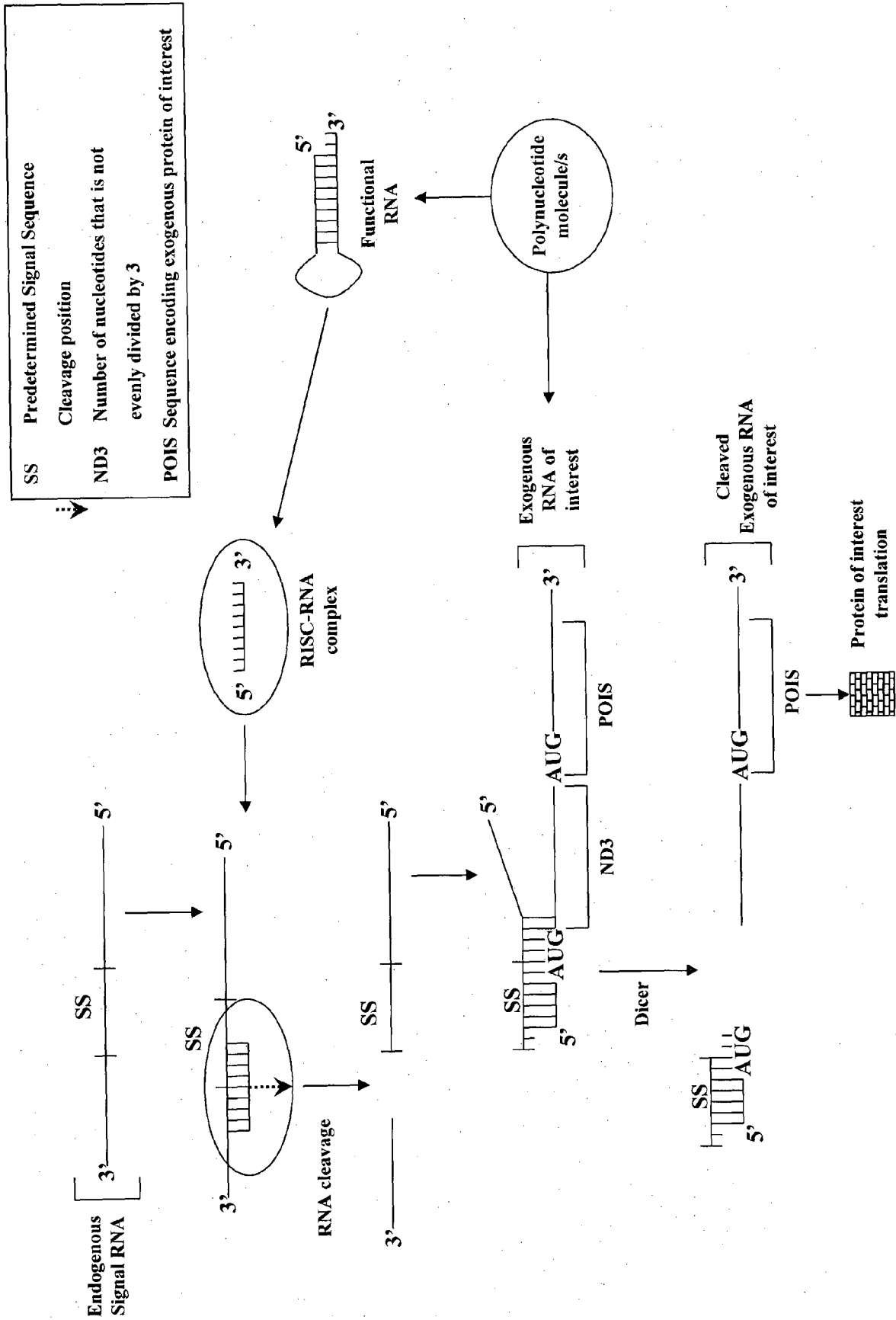


FIG. 33

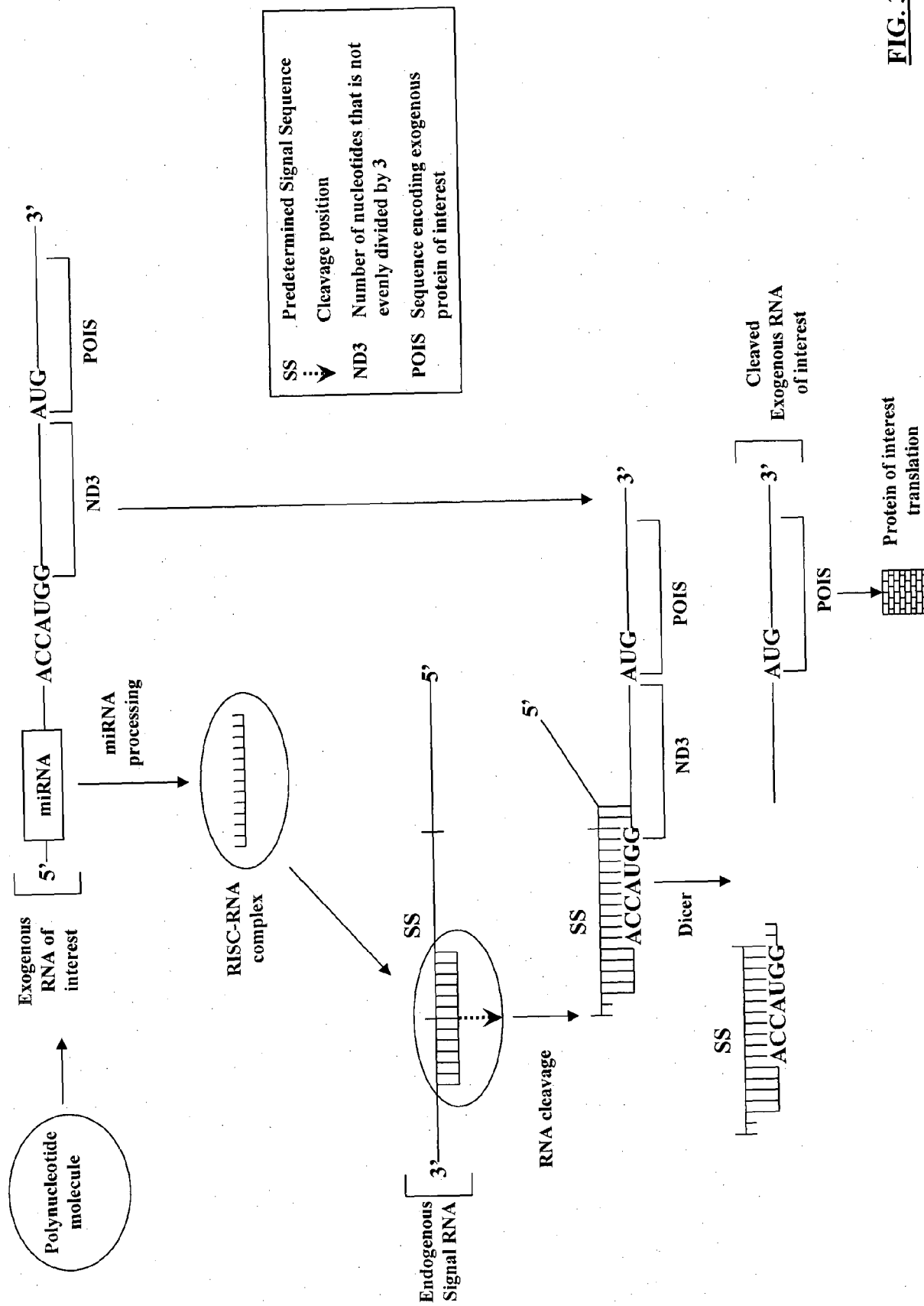


FIG. 34

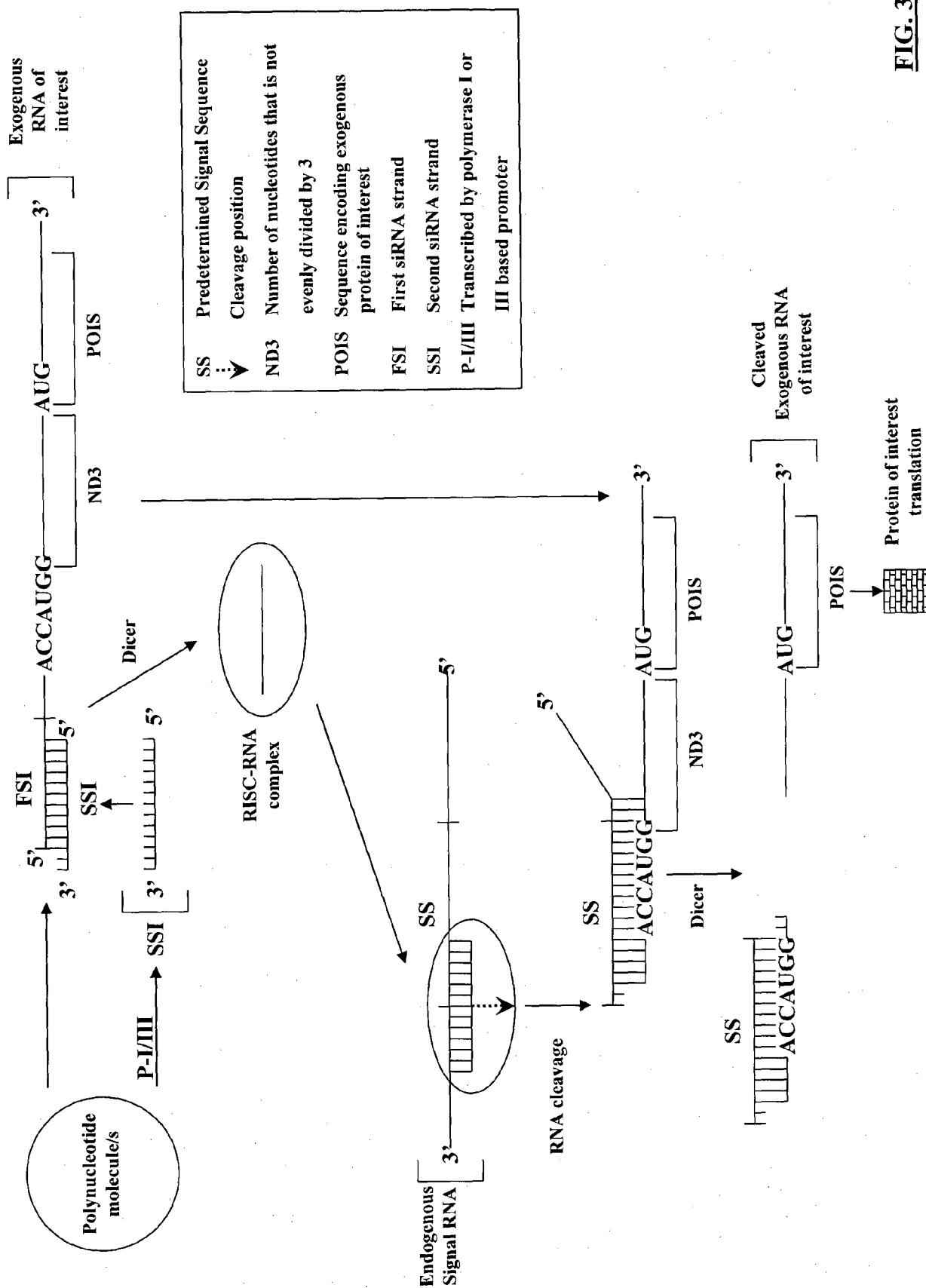


FIG. 35

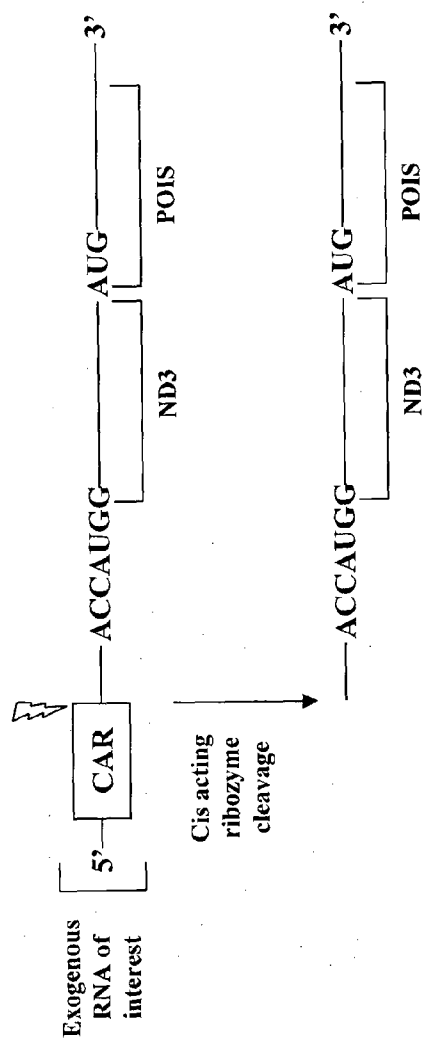
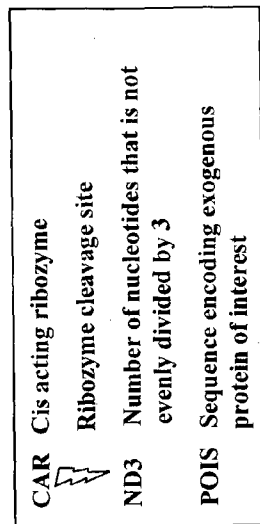


FIG. 36A

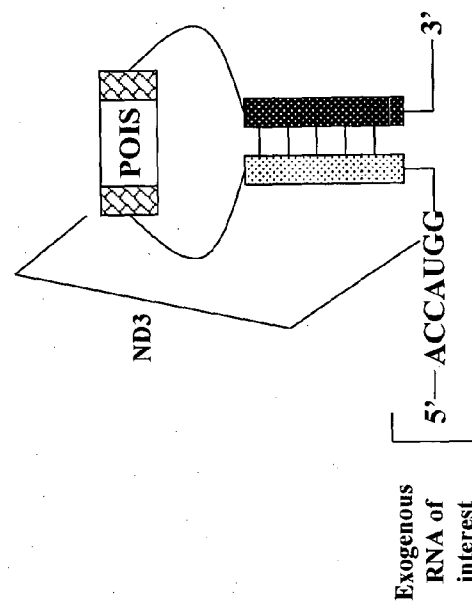
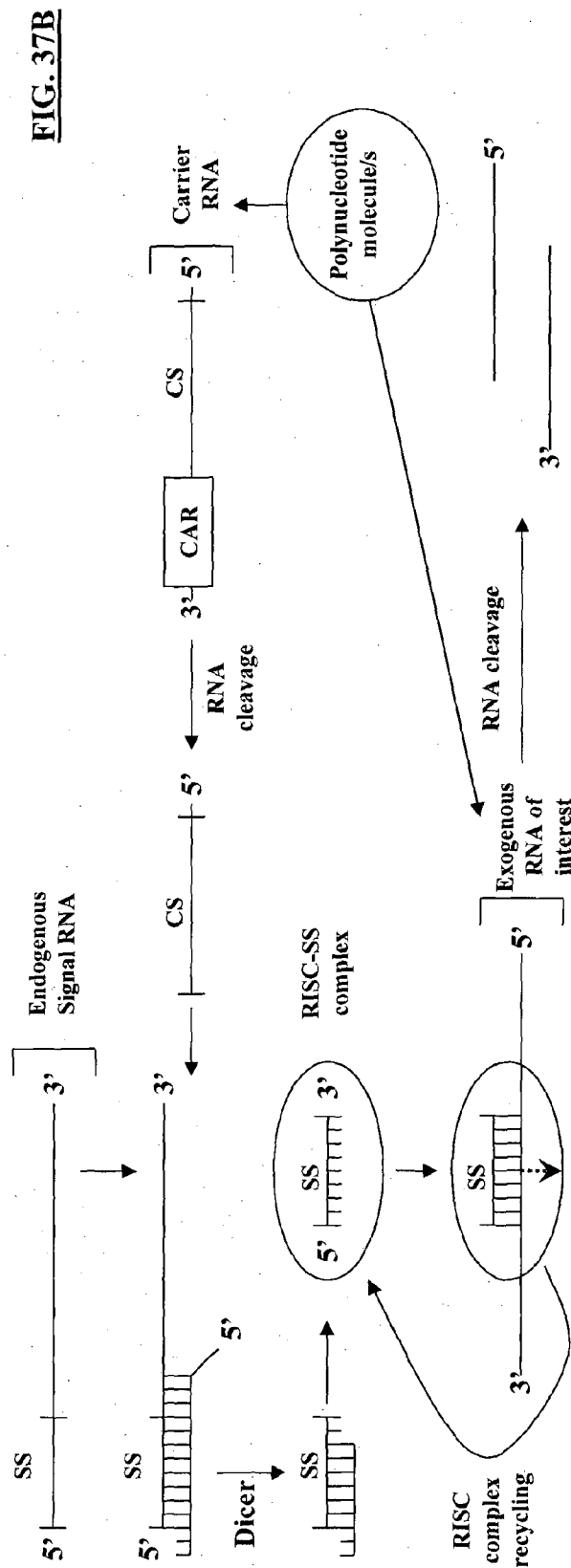
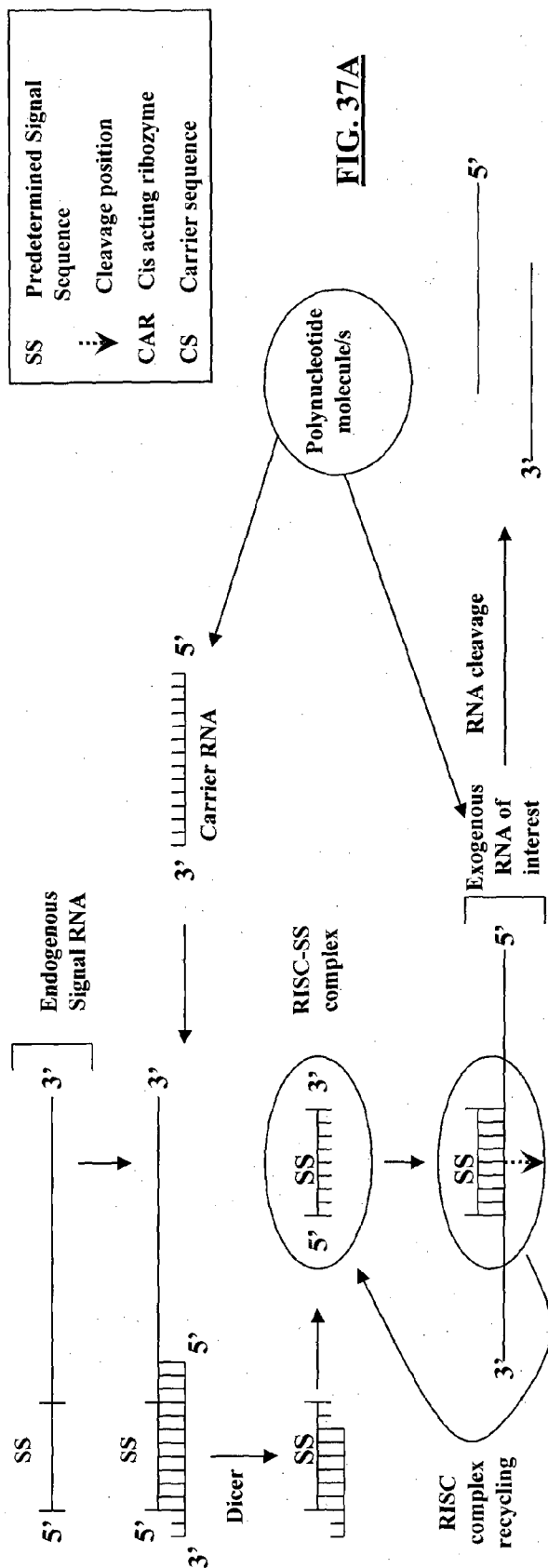
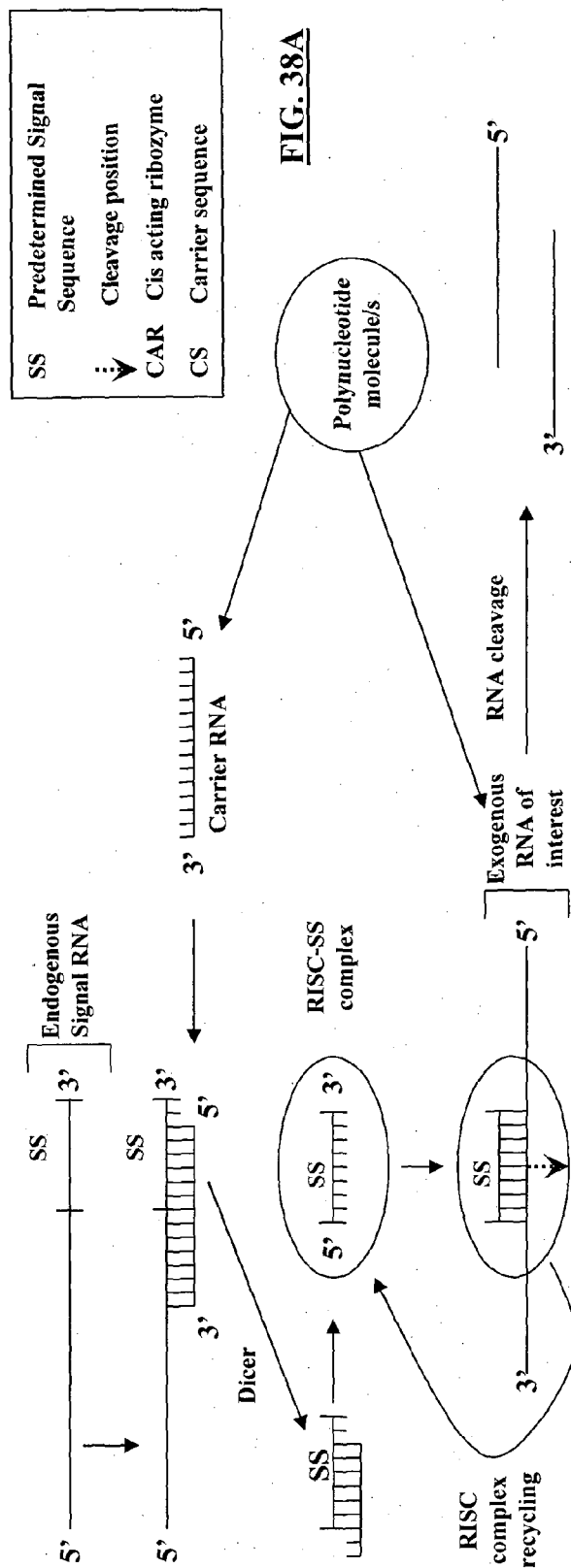
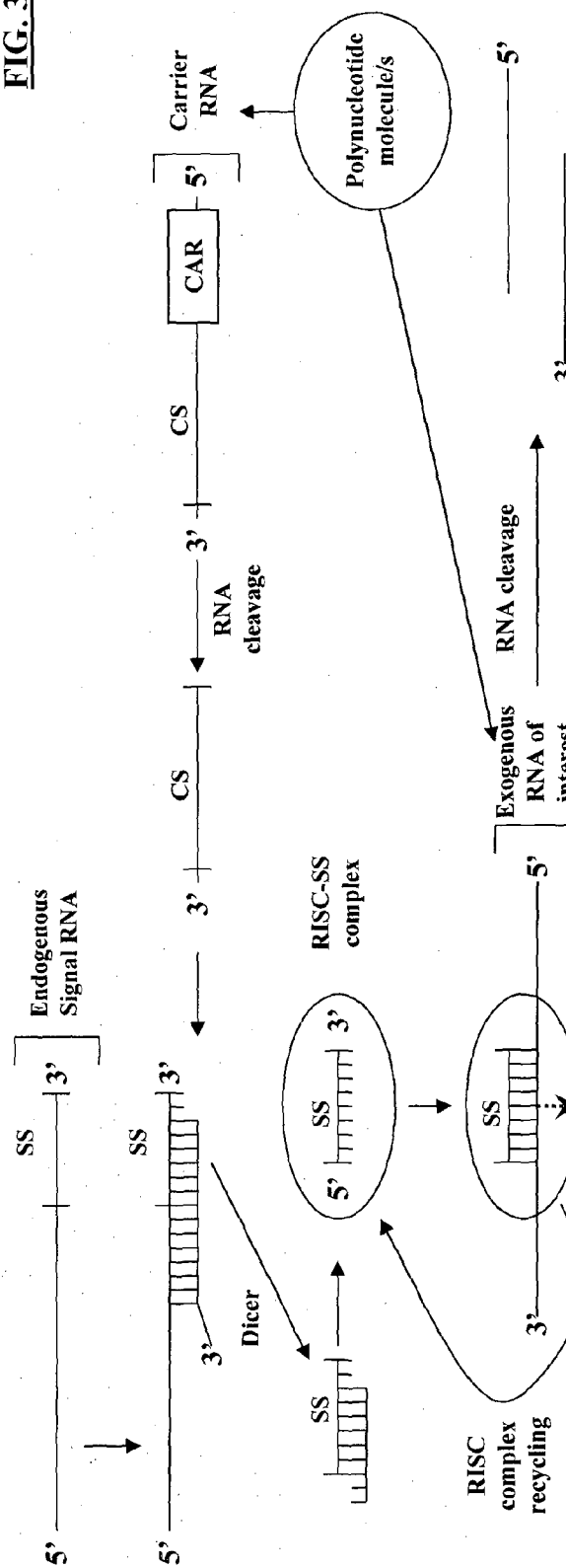


FIG. 36B



**FIG. 38B**

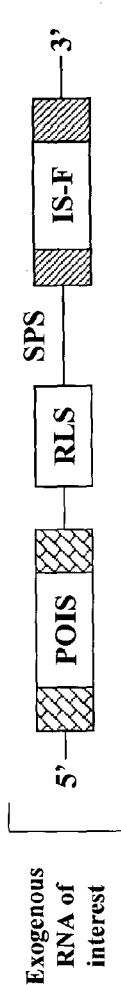


FIG. 39A

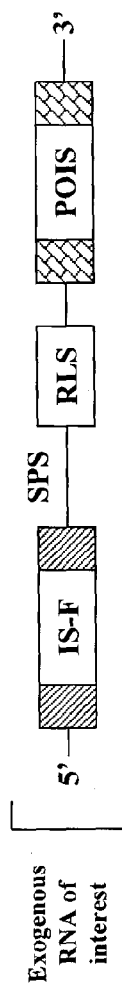


FIG. 39B

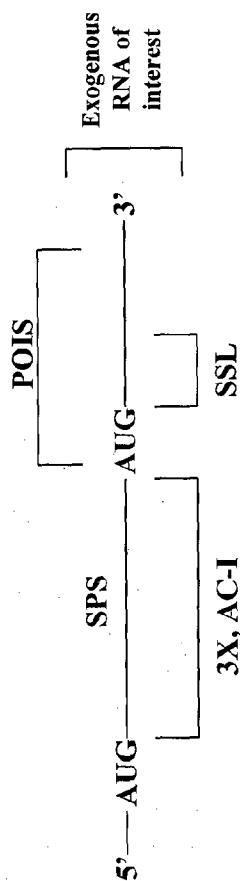


FIG. 39C

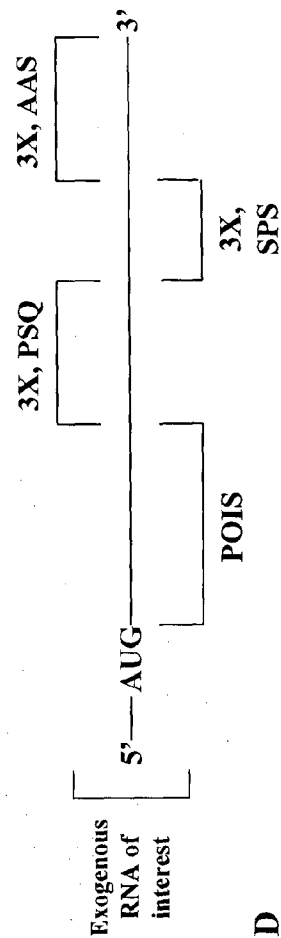
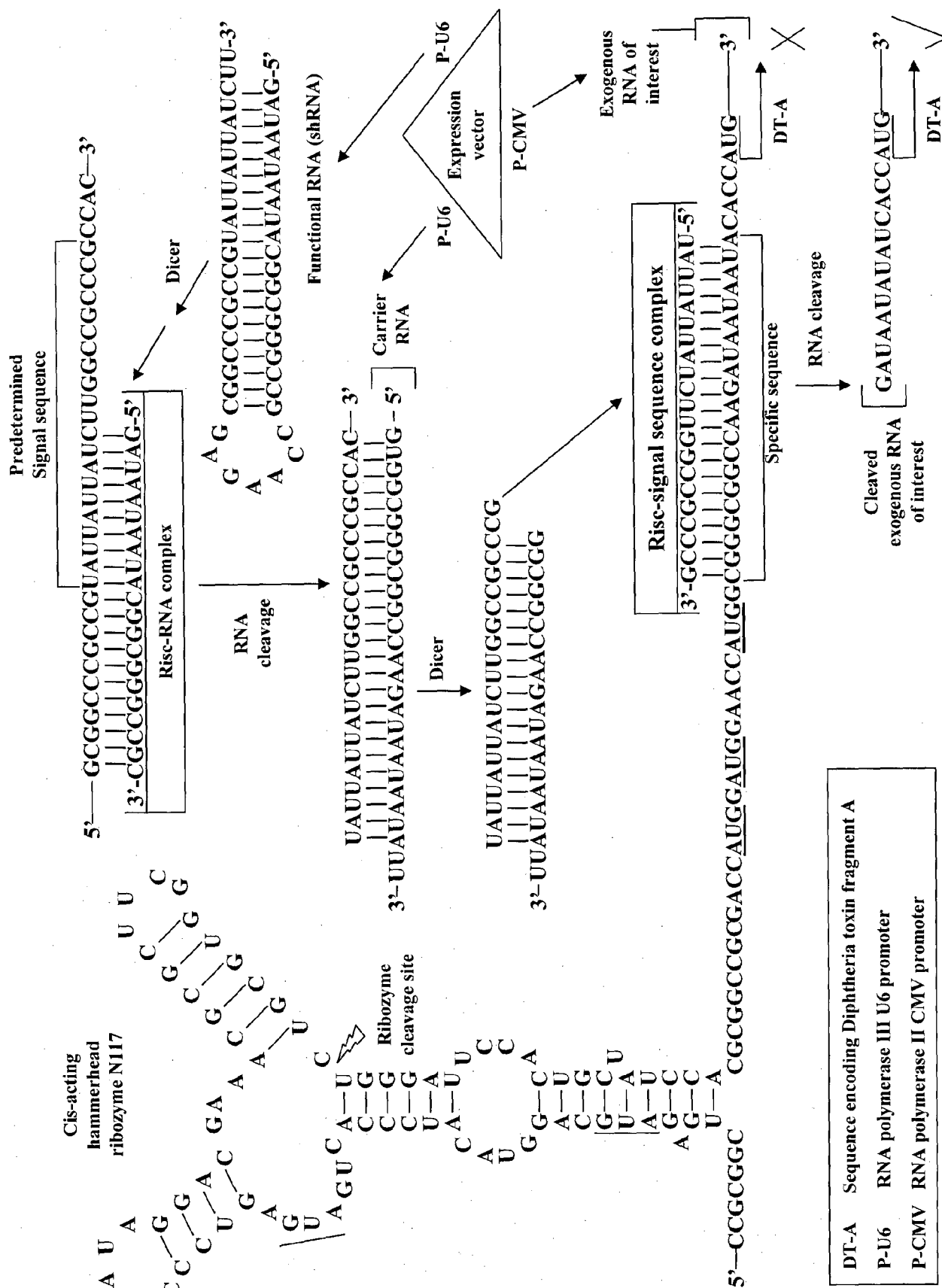


FIG. 39D

SPS	The specific sequence
POIS	Sequence encoding exogenous protein of interest
RLS	RNA localization signal for subcellular localization
IS-F	Sequence that is capable of inhibiting the RLS function
3X	Number of nucleotides that is evenly divided by 3
SSL	Sequence encoding sorting signal for subcellular localization
AC-I	Sequence encodes amino acids that are capable of inhibiting the SSL function
*	Without stop codon downstream from the start codon of POIS
PSQ	Sequence encoding peptide sequence that is capable of being cleaved by a protease
AAS	Sequence encoding amino acid sequence that is capable of inhibiting PSQ cleavage



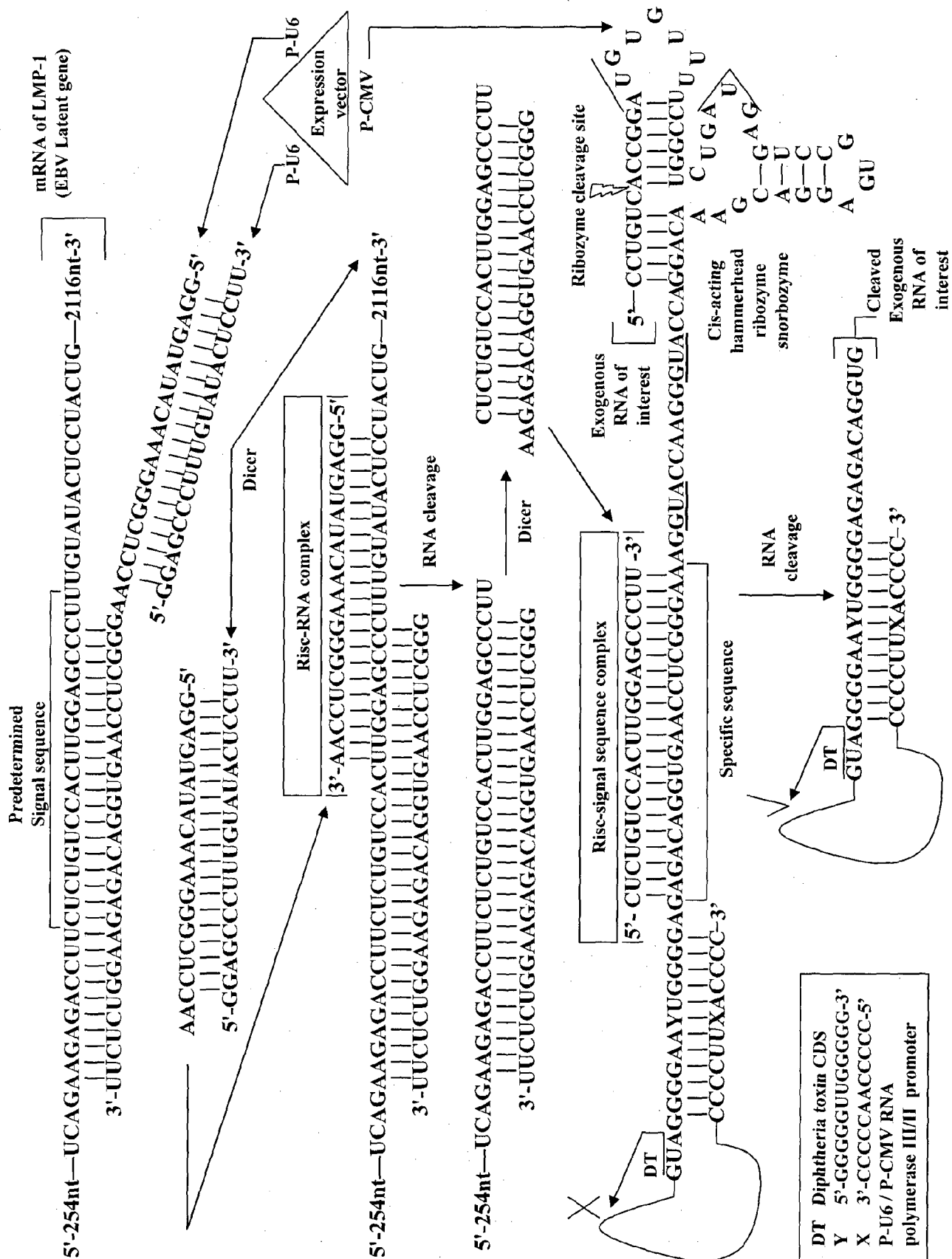
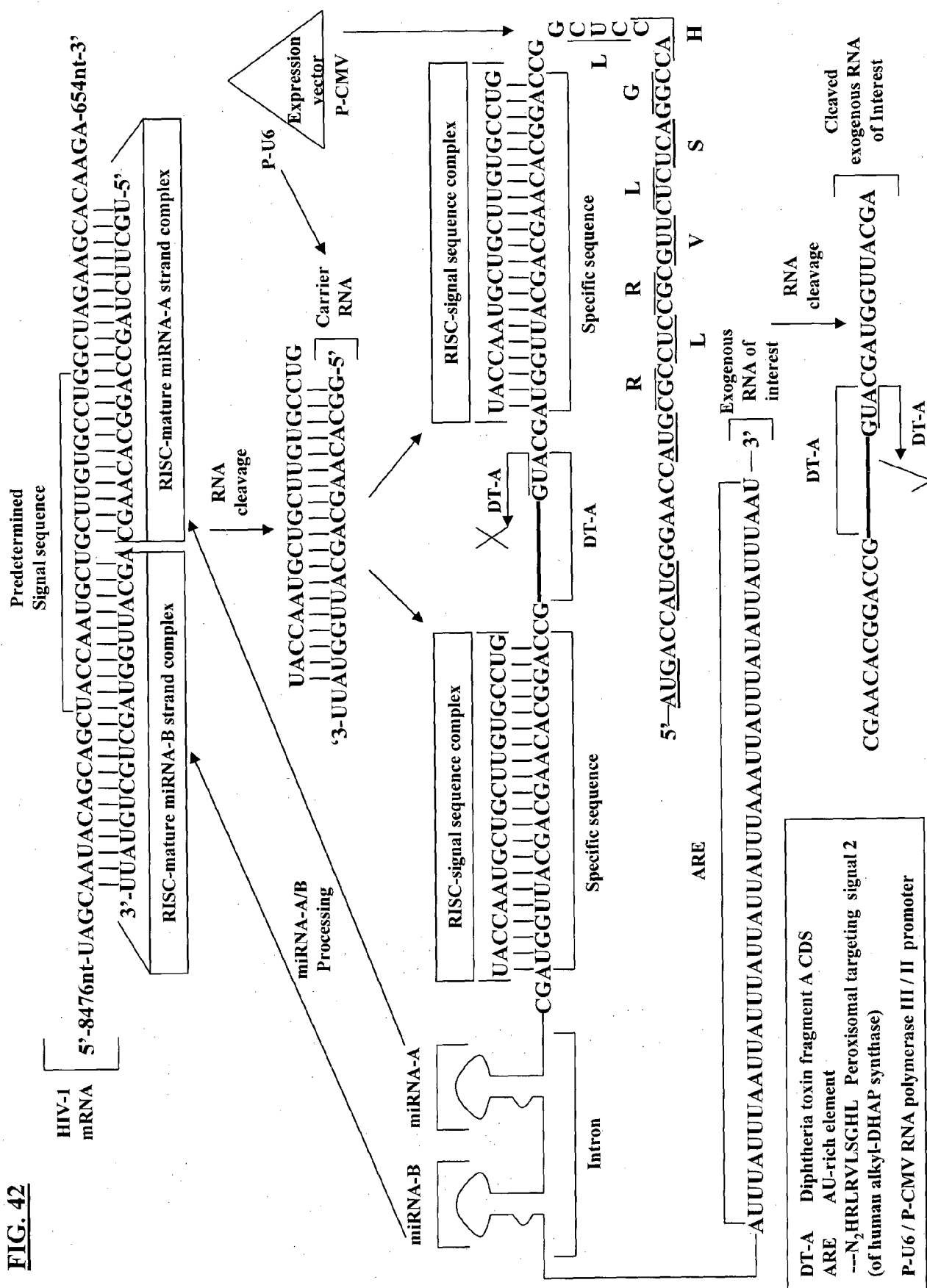
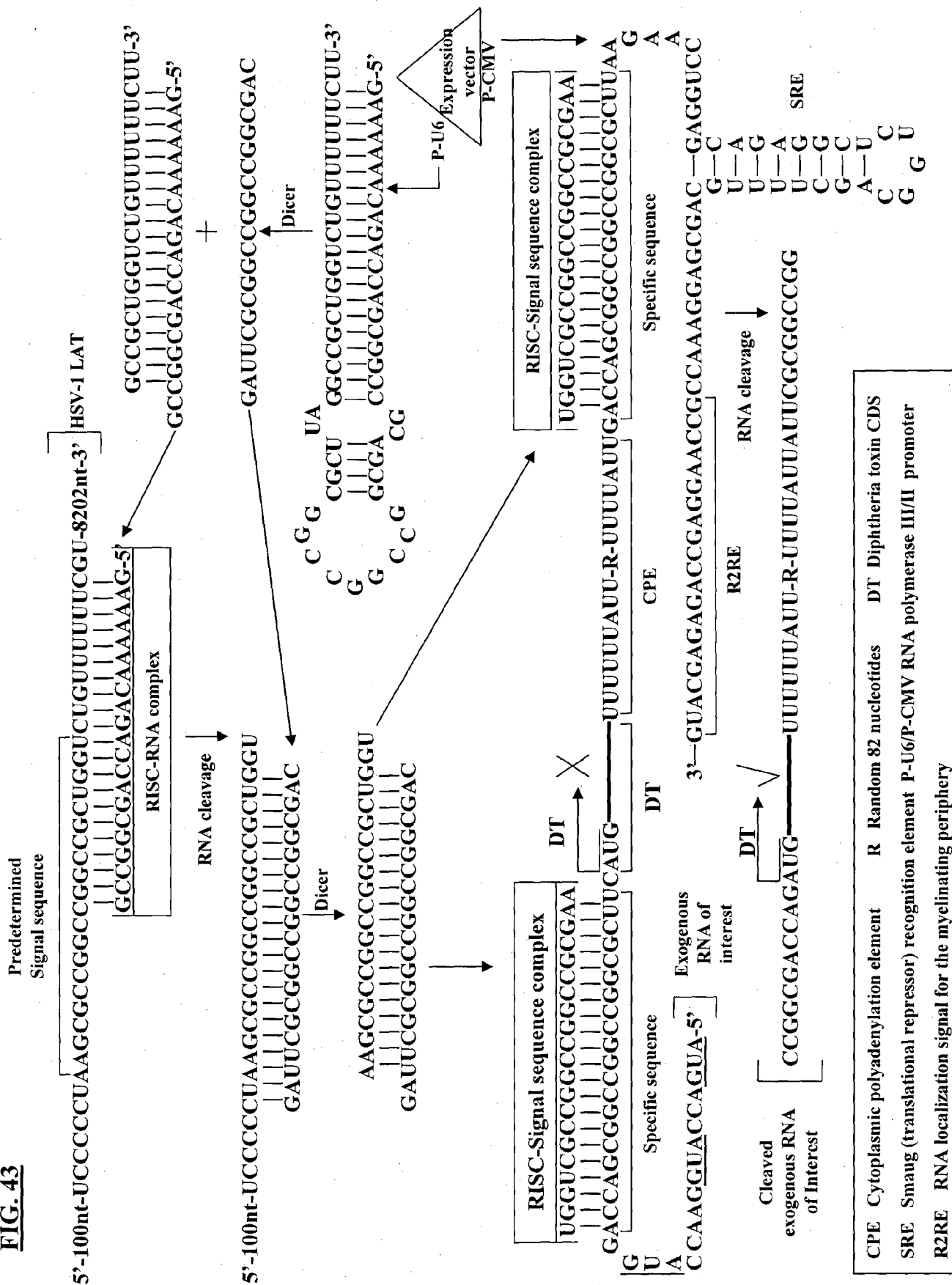
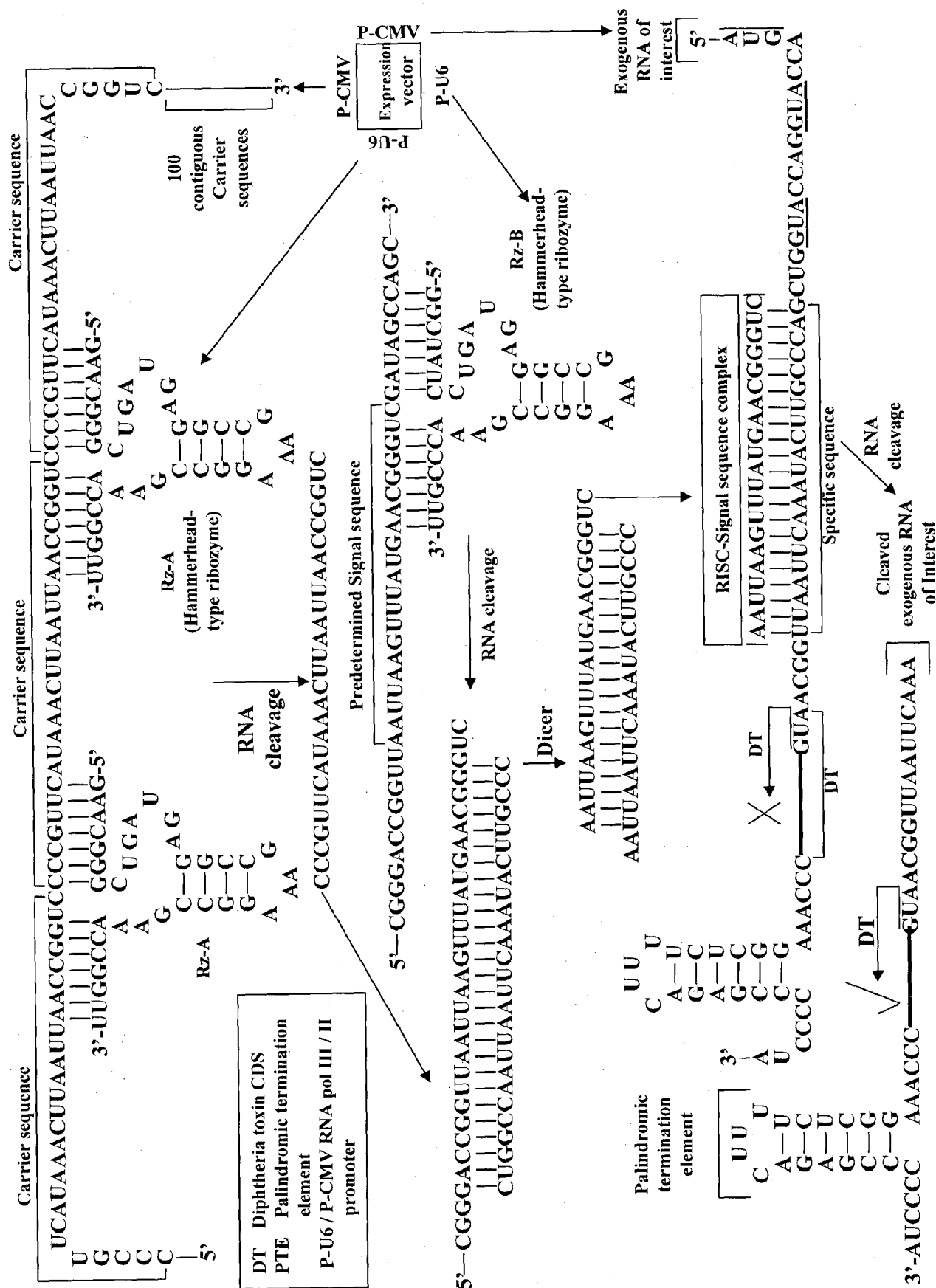


FIG. 41







NaDoc 002-PCT-1_seq listing_ST25.txt
SEQUENCE LISTING

<110> NanoDoc Ltd.

<120> COMPOSITIONS AND METHODS FOR SPECIFIC CLEAVAGE OF EXOGENOUS RNA
IN A CELL

<130> NaDoc/002 PCT-1

<150> PCT/IL2010/000895

<151> 2010-10-28

<160> 114

<170> PatentIn version 3.5

<210> 1

<211> 8905

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 1

aacaaaatat taacgcttac aatttccatt cgccattcag gctgcgcaac tgttggaag	60
ggcgatcggg ggcggcctct tcgctattac gccagctggc gaaaggggga tgtgctgcaa	120
ggcgattaag ttgggtaacg ccagggtttt cccagtcacg acgttgtaaa acgacggcca	180
gtgccaagct gatctataca ttgaatcaat attggcaatt agccatatta gtcattgggt	240
atatagcata aatcaatatt ggctattggc cattgcatac gttgtatcta taccataata	300
tgtacattta tattggctca tgtccaatat gaccgccatg ttgacattga ttattgacta	360
gttattaata gtaatcaatt acgggggtcat tagttcatag cccatatatg gagttccgcg	420
ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc cgccattga	480
cgtcaataat gacgtatggt cccatagtaa cgccaatagg gactttccat tgacgtcaat	540
gggtggagta ttacggtaa actgcccact tggcagtaca tcaagtgtat catatgccaa	600
gtccgcccc tattgacgtc aatgacggta aatggcccg cttggcattat gccagtaca	660
tgaccttacg ggactttcct acttggcagt acatctacgt attagtcac gctattacca	720
tgggtgatgcg gttttggcag tacaccaatg ggcgtggata gcggtttgac tcacggggat	780
ttccaagtct ccacccatt gacgtcaatg ggagtttggt ttggcaccaa aatcaacggg	840
actttccaaa atgtcgtaat aaccccgccc cgttgacgca aatgggcggg aggcgtgtac	900
gggtgggaggt ctatataagc agagctcgtt tagtgaaccg tcagaatttt gtaatacgac	960
tcactatagg gcggccggga attcgtcgac tggatccggg acctagctag gtagcaattg	1020
accggtcaag atggcgggca acaacaaca caacaacaac aacaacaaca acaacaaca	1080
caacaagaag atggcgggca caacaacaac aacaacaaca acaacaaca caacaacaac	1140
caacaacaag atggcgggca acaacaaca caacaacaac aacaagaaga tggcgggcaac	1200
aacaacaaca acaacaaca caaccaagat ggcggccaac aacaacaaga agatggcggc	1260
aacaacaaca accaagatgg cggccaaca caacaagaag atggcgggca caacaacaac	1320

NaDoc 002-PCT-1_seq listing_ST25.txt

caagatggcg gcacgcgtcg gtccggctag ccgtacgctc cttagcgacg aaatctactg	1380
ccccctgag agccaccatg gcttggggtc ctacgctgtg caggccaagt ttggagatta	1440
caacaaagaa ggccgccatg gtgggcacct cagctctgag cggctcatcc gccaccatgg	1500
gttgaccag cacaaactta ccagggaccg ccgccatggc cggaccagg cgtgccacca	1560
tggacaccgt gggttgcgcc gccatggtgc tctgttgag tgccaccatg gtgctcagga	1620
cctgggccgc catggaatac ctgataactg ataagccacc atgggaacag acctttggct	1680
tggagttgac gcccttgac tcaacattta cgaggccgcc atggagtcca ccccaaagat	1740
tggctttcct tggagtgaag tcaggaacat ctctgccacc atggaaaagt ttgtcatcaa	1800
gcccacgac aaggccgcca tggactttgt gttttacgcc ccacgtctca cagccaccat	1860
ggggaccctg cagctcgccg ccatggacca cgagttgtac gccaccatgg ggaagcctga	1920
caccgccgcc atggagcaga cgaaggccgc caccatggag gctgataagc tgataagccg	1980
ccatgggctg gaaacagaga agaaaaggag agaaaccgtg gagagagaga aagagcgcca	2040
ccatggcgag aaggaggagt tgttctgctg gctgcaggac tacgaggaga agacaagccg	2100
ccatgggaga gacctctcg agcagattca gaggggccac catggggagg aggagaggaa	2160
gcgggcacag gagggccgcc atggcccaga ggctgaccgc caccatggac tgcgggctaa	2220
ggggcccat gggagacagg cgggtgggcca ccatgggagc caggagcagc gccgccatgg	2280
gctacctgat aactgataag ccaccatggt ggaagaggcg cggaggcgca aggaggacga	2340
agttgaagag tggcagcaag ccgccatgga agcccaggac gacctggtca agaccaagga	2400
ggagctgcac ctggtgccgg ccaccatggc gccaccacca ccaccgtgt acgagccggc	2460
cgccatggac gtccaggaga gcttgcaaga cgagggtgcc accatggcgg gctacagcgc	2520
agccgccatg gctgacggca tccgggccac catggacgag gagaagcgtg ccgccatggc	2580
agagaagaac gaggccacca tggggcctga taagctgata agccgccatg gggcccgaga	2640
cgagaacaag aggaccaca acgacatcat ccacaacgag agccaccatg gaggccggga	2700
caagtacaag acgctgcggc agatccggca gggcaacacc agccgccatg gcgacgagtt	2760
cgaggccctg caacagccag gccaccatgg agggcagagg ggtgctcata gcgggcgctg	2820
ccgccatggc cagcttctg tctgccacca tggaagtctc ggaactcgcc gccatggcag	2880
ttcctttcga agccaccatg gcaacagaaa cattcgccgc catggaccac ctgataactg	2940
ataagccacc atggttgcaa tctgccaag caggcctgat tctcgcgatt actcgcgaat	3000
caccgccgcc atggtgctgg gagcaggact cattgaatta cggaaaacgc ctgtcaagtc	3060
tcaggccacc atggggaact ggcctgtgtc atacaagagt caggccgcca tggggaaacg	3120
tggcaggact tccatctgtg ccgccaccat ggtgtattcg aaacgagccg ccatggattt	3180
tctcatctct gccaccatgg catctttgta cattgccgcc atgggagggg tcaaaattgc	3240
caccatggtg gctgataagt tgatagtaac cgccatggtg tttcatccag tcgccaccat	3300
gggctggcag agagcagccg ccatggcagc gtcagtgggt gccaccatgg cttggatttt	3360

NaDoc 002-PCT-1_seq listing_ST25.txt

tttttttggt tttttttttt ttgtcaaca attttacaac acattgtgtc gacgagctca	3420
agcttccccg cgcgccccgg tccgtccgga ctacggcaag ctgaccctga agttcatccc	3480
aaaacttacg ctgagtactt cgatctgggt accccggatc cgtgatagta acctgatagt	3540
aacctgataa tagcagatct cgccgccatg ggagctgatg atgtggttga ttcttcgaaa	3600
tcttttgta tggaaaactt ttcttcgtac cacgggacga aacctgggtta tgtggattcc	3660
attcaaaaag gcatacaaaa gccaaaatct ggtacacaag gaaactatga cgatgattgg	3720
aaagggtttt atagtaccga caacaaatat gacgctgcgg gatactctgt ggataatgaa	3780
aaccgcgtct ctggaaaagc tggaggcgtg gtcaaagtga cgtatccagg actgacgaag	3840
gttctcgcac taaagggtgga taatgccgaa actattaaga aagagttagg tttaaagtct	3900
actgaaccgc tcatggagca agtcggaacg gaagagtta tcaaaagatt cggtgatggt	3960
gcttcgcgtg tagtgctcag ccttcccttc gctgagggga gttctagcgt tgagtacatc	4020
aacaactggg aacaggcgaa agcgttaagc gtagaacttg agattaactt tgaaaccctg	4080
ggaaaacgtg gccaaagatgc gatgtatgag tatatggctc aagcctgtgc aggaaatcgt	4140
gtcaggcgat agtgaactag tatccggaat ctagagcggc cgcactcgag gtttaaacgg	4200
ccggccgcgg tcatagctgt ttctgaaca gatcccggtt ggcattccctg tgaccctcc	4260
ccagtgcctc tcttgccctt ggaagttgcc actccagtgc ccaccagcct tgtcctaata	4320
aaattaagtt gcatcatttt gtctgactag gtgtccttct ataattattat ggggtggagg	4380
ggggtggtat ggagcaaggg gcaagttggg aagacaacct gtagggcctg cggggtctat	4440
tgggaaccaa gctggagtgc agtggcacia tcttggtca ctgcaatctc cgcctcctgg	4500
gttcaagcga ttctcctgcc tcagcctccc gagttgttgg gattccaggc atgcatgacc	4560
aggctcagct aatttttggt tttttgtag agacggggtt tcaccatatt ggccacgctg	4620
gtctccaaact cctaatttca ggtgatctac ccaccttggc ctcccaaatt gctgggatta	4680
caggcgtgaa cactgctcc ctccctgtc cttctgattt taaaataact ataccagcag	4740
gaggacgtcc agacacagca taggctacct ggccatgccc aaccggtggg acatttgagt	4800
tgcttgcttg gcactgtcct ctcatgcgtt ggggtccactc agtagatgcc tgttgaattg	4860
ggtacgcggc cagcttggct gtggaatgtg tgtcagttag ggtgtggaaa gtccccaggc	4920
tccccagcag gcagaagtat gcaaagcatg catctcaatt agtcagcaac cagggtgtgga	4980
aagtccccag gctccccagc aggcagaagt atgcaaagca tgcatctcaa ttagtcagca	5040
accatagtcc cgccccctaac tccgcccata ccgcccctaa ctccgcccag ttccgccc	5100
tctccgcccc atggctgact aatttttttt atttatgcag aggccgaggc cgcctcggcc	5160
tctgagctat tccagaagta gtgaggaggc ttttttgtag gcctaggctt ttgcaaaaag	5220
ctcccgaggc cttgtatatc cattttcggg tctgatcaag agacacgtac gaccatggag	5280
agcgacgaga gcggcctgcc cgccatggag atcgagtgcc gcatcaccgg caccctgaac	5340
ggcgtggagt tcgagctggt gggcgccgga gagggcacc cagagcaggg ccgcatgacc	5400

NaDoc 002-PCT-1_seq listing_ST25.txt

aacaagatga agagcaccaa aggcgccttg accttcagcc cctacctgct gagccacgtg	5460
atgggctacg gcttctacca cttcggcacc taccacagcg gctacgagaa ccccttcctg	5520
cacgccatca acaacggcgg ctacaccaac acccgcatcg agaagtacga ggacggcggc	5580
gtgctgcacg tgagcttcag ctaccgctac gaggccggcc gcgtgatcgg cgacttcaag	5640
gtgatgggca cgggcttccc cgaggacagc gtgatcttca ccgacaagat catccgcagc	5700
aacgccaccg tggagcacct gcaccccatg ggcgataacg atctggatgg cagcttcacc	5760
cgcaccttca gcctgcgcga cggcggctac tacagctccg ttgtggacag ccacatgcac	5820
ttcaagagcg ccatccaccc cagcatccta cagaacgggg gccccatgtt cgccttcgcg	5880
cgcgtggagg aggatcacag caacaccgag ctgggcatcg tggagtacca gcacgccttc	5940
aagaccccg atgcagatgc cggatgaaga taactgcagc gggactctgg gggttcgaaat	6000
gaccgaccaa gcgacgccc acctgccatc acgagatttc gattccaccg ccgccttcta	6060
tgaaagggtg ggcttcggaa tcgttttcgg ggacgcgggc tggatgatcc tccagcgcg	6120
ggatctcatg ctggagttct tcgcccaccc caacttgttt attgcagctt ataatggtta	6180
caaataaagc aatagcatca caaatttcac aaataaagca tttttttcac tgcattctag	6240
ttgtggtttg tccaaactca tcaatgtatc ttatcatgtc tgtataccgt cgacctctag	6300
ctagagcttg gcgtaatcat ggtcatagct gtttcctgtg tgaaattggt atccgctcac	6360
aattccacac aacatacgag ccggaagcat aaagtgtaaa gcctgggggtg cctaagagt	6420
gagctaactc acattaattg cgttgcgctc actgcccgtt ttccagtcgg gaaacctgtc	6480
gtgccagctg cattaatgaa tcggccaacg cgcggggaga ggcgggttgc gtattggg	6540
ctcttcgct tcctcgctca ctgactcgct gcgctcggtc gttcggctgc ggcgagcgg	6600
atcagctcac tcaaaggcgg taatacgggt atccacagaa tcaggggata acgcaggaaa	6660
gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc	6720
gtttttccat aggtccgccc cccctgacga gcatcaciaa aatcgacgct caagtcagag	6780
gtggcgaaac ccgacaggac tataaagata ccaggcgttt ccccttgga gtcctcctg	6840
gcgctctcct gttccgacct tgccgcttac cggatacctg tccgccttcc tcccttcgg	6900
aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcgggtg aggtcggtc	6960
ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ctttatccg	7020
taactatcgt cttgagtcca acccggttag acacgactta tcgccactgg cagcagccac	7080
tggtaacagg attagcagag cgaggatgt aggcgggtgct acagagttct tgaagtgg	7140
gcctaactac ggctacacta gaagaacagt atttggtatc tgcgctctgc tgaagccagt	7200
taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg	7260
tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc	7320
tttgatcttt tctacggggt ctgacgtca gtggaacgaa aactcacgtt aagggatttt	7380
ggatcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa aatgaagttt	7440

NaDoc 002-PCT-1_seq listing_ST25.txt

taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag	7500
tgaggcacct atctcagcga tctgtctatt tcgttcatcc atagttgcct gactccccgt	7560
cgtgtagata actacgatac gggagggcct accatctggc cccagtgtg caatgatacc	7620
gcgagacca cgctcaccgg ctccagattt atcagcaata aaccagccag ccggaagggc	7680
cgagcgcaga agtggtcctg caactttatc cgctccatc cagtctatta attggtgccg	7740
ggaagctaga gtaagtagtt cgccagttaa tagtttgccg aacgttggtg ccattgctac	7800
aggcatcgtg gtgtcacgct cgctgtttgg tatggcttca ttcagctccg gttcccaacg	7860
atcaaggcga gttacatgat ccccatgtt gtgcaaaaaa gcggttagct ccttcggtcc	7920
tccgatcgtt gtcagaagta agttggccgc agtggtatca ctcatggtta tggcagcact	7980
gcataattct cttactgtca tgccatccgt aagatgcttt tctgtgactg gtgagtactc	8040
aaccaagtca ttctgagaat agtgtagcg gcgaccgagt tgctcttgcc cggcgtcaat	8100
acgggataat accgcgccac atagcagaac tttaaaagtg ctcatcattg gaaaacgttc	8160
ttcggggcga aaactctcaa ggatcttacc gctgttgaga tccagttcga tgtaaccac	8220
tcgtgcaccc aactgatctt cagcatcttt tactttcacc agcgtttctg ggtgagcaaa	8280
aacaggaagg caaatgccc caaaaaagg aataaggcg acacggaaat gttgaatact	8340
catactcttc ctttttcaat attattgaag catttatcag ggttattgtc tcatgagcgg	8400
atacatattt gaatgtattt agaaaaataa acaaataggg gttccgcgca catttccccg	8460
aaaagtgcc cctgacgcgc cctgtagcgg cgcatthaag gcggcggtg tgggtggttac	8520
gcgcagcgtg accgctacac ttgccagcgc cctagcgcgc gctcctttcg ctttcttccc	8580
ttcctttctc gccacgttcg ccggctttcc ccgtcaagct ctaaatecgg ggctcccttt	8640
agggttccga tttagtgtt tacggcacct cgaccccaaa aaacttgatt agggtagtg	8700
ttcacgtagt gggccatcgc cctgatagac ggtttttcgc cctttgacgt tggagtccac	8760
gttctttaat agtggactct tgttccaaac tggaacaaca ctcaacccta tctcgggtcta	8820
ttcttttgat ttataaggga ttttgccgat ttcggcctat tggttaaaaa atgagctgat	8880
ttaacaaaaa tttacgcga atttt	8905

<210> 2
 <211> 12082
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 2	60
aacaaaatat taacgcttac aatttccatt cgccattcag gctgcgcaac tgttgggaag	
ggcgatcggg gcgggcctct tcgctattac gccagctggc gaaaggggga tgtgctgcaa	120
ggcgattaag ttgggtaacg ccagggtttt cccagtcacg acgttgtaaa acgacggcca	180
gtgccaaagt gatctataca ttgaatcaat attggcaatt agccatatta gtcattggtt	240

NaDoc 002-PCT-1_seq listing_ST25.txt

atatagcata aatcaatatt ggctattggc cattgcatac gttgtatcta tatcataata	300
tgtacattta tattggctca tgtccaatat gaccgccatg ttgacattga ttattgacta	360
gttattaata gtaatcaatt acgggggtcat tagttcatag cccatatatg gagttccgcg	420
ttacataact tacggtaaatt ggcccgcctg gctgaccgcc caacgacccc cgcccattga	480
cgtaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat tgacgtcaat	540
gggtggagta ttacggtaa actgcccact tggcagtaca tcaagtgtat catatgccaa	600
gtccgcccc tattgacgtc aatgacggta aatggccgc ctggcattat gccagtaca	660
tgacctacg ggactttcct acttggcagt acatctacgt attagtcac gctattacca	720
tggtgatgcg gttttggcag tacaccaatg ggcgtggata gcggttgac tcacggggat	780
ttccaagtct ccacccatt gacgtcaatg ggagtttgtt ttggcaccaa aatcaacggg	840
actttccaaa atgtcgtaat aacccgccc cgttgacgca aatgggcggt aggcgtgtac	900
ggtgggaggt ctatataagc agagctcgtt tagtgaaccg tcagaatttt gtaatacgac	960
tcactatagg gcggccggga attcgtcgac tggatccgct agctgataat agccgatacc	1020
ctgtcaccgg atgtgctttc cggctgatg agtccgtgag gacgaaacag gactgtcgca	1080
gagacaaccg ccgataccct gtcaccggat gtgctttccg gtctgatgag tccgtgagga	1140
cgaacagga ctgtcgagg gcaccgtatc aactctgaga tgcaggtaca tccagctgat	1200
gagtcccaa taggacgaaa cgcgttcgg tgcgtcctgg attccactgc tatccactat	1260
tcaccaaag agactatcaa ctctgagatg caggtagatc cagctgatga gtcccaaata	1320
ggacgaaacg cgcttcggtg cgtcctggat tccactgcta tccactattc atccgtctcc	1380
gatcgtctag agcctgacct agctgaccta gccgccacca tgcagaggct gcaggtagtg	1440
ctgggccacc tgaggggtcc ggccgattcc ggtggatgc cgcaggcagc gccttgctg	1500
agatctacta gtggtgacct agctgaccta gccgtacca tggaccctgt tgtgctgcaa	1560
aggagagact gggagaaccc tggagtcacc cagctcactg aacttaaccc tagccctgcc	1620
accatggctt ggaggaactc cgaggaagcc aggactgaca gtaacagtag gcagcgccgc	1680
catggcaacg gagagtggag gtttgcttg tttgacgcta aggatagtgt ggccaccatg	1740
gggctggagt gcgacctccc agaggcggct gacgttaaag ttagcagcag ccgcatggg	1800
cacggctacg acgcgcccat ctacagtac gtttagcacta agatcgccac catggcccct	1860
tttgtgcca ccgagaaccc gactaactgt agcagtgaca cctgccgcca tggcgagagc	1920
tggctgcaag aaggacagac taagattagc tgtgacggag ccaccatggc cttccacctc	1980
tggctgcaacg gcaggtgtaa cggtagcggg gaagacagcc gccatggctc cgagttcgac	2040
ctctctgcct tccgtaacgc tgaagataac agggccacca tgggtggtgct cagggtgtcc	2100
gacggcagct atagggctga ccgtaacatg tgccgccatg gtggcatctt cagggacgtc	2160
agcctgctta gactgagac taaccaggcc accatggctc acgttgccac gaggttcaac	2220
gacgatagca gtgaagctaa gctgggccgc catgggcaga tgtgtggaga actcagagag	2280

NaDoc 002-PCT-1_seq listing_ST25.txt

tctgacagta gcactaagag cgccaccatg ggcgagaccc aggtggcctc tggcacagct	2340
gacttttagag ctaagatcag ccgccatgga ggaggctacg ccgacagagc cacccttgag	2400
cttagcgta agaacgccac catgggggtct gccgagaccc ccaacctcta cagtaacgtt	2460
agggctgagc acagccgcca tggcacgctc atcgaagccg aagcctgcga taacgggtgac	2520
agtagagtcg ccaccatgga cggcctgctg ctgctcaacg gcaagcctaa gcttgacagt	2580
agagtcagcc gccatgggca ccatcctctg cacggacaag tcattaaggc tgagactagg	2640
gtggccacca tgggtgtcac gaagcagaac aacttcaacg ctaagagtga ctctagctac	2700
cgccgccatg gtctctggta caccctgtgc gacaggaata accttagggg tgacgaggcc	2760
accatggctg agacacacgg catggtgccc acgaataagc ttagagctga cccagccgc	2820
catggtgcca tgtccgagag agtcaccagg attaagcata gagctgagaa cgccaccatg	2880
gtcatcatct ggtctctggg caacgagtct aagcatagag ctgaccacgg ccgccatggc	2940
aggtggatca agtctgccga cccagtaga cctgagcata acgaagccac catggcagac	3000
accacagcca cagacatcat ctgtagcatt gaggctaagg tcggccgcca tgggcccctc	3060
cctgctgtgc ccaagtggag tagcactgag tgtaactctg ccaccatgga aacgagacct	3120
ctcatcctgt gcgagtatag acctgaaagt aaacccggcc gccatggctt tgccaagtac	3180
tggcaagcct tactgagta tagcactaag caagccacca tggcccgcga ctgggtggac	3240
cagtcactca ttgagtatag cgataacggc agccgccatg gtgcctacgg aggagacttt	3300
ggcgacactg agaatagcac taagtctgcc accatgggccc tgggtcttgc cgaccggact	3360
ccgcctgacc cttagactaa ggccagccgc catggacagt tcttcccgtt cacgctgtct	3420
ggtgaaacta acgatagcac agccaccatg gtcttcagac actccgacaa cgagctgctt	3480
gactgtaagg ctaggctggg ccgccatggt ctggcttctg gcgagggtgcc tctggctgag	3540
gctaatacata gaaaggccac catggaactg cccgagctgc ctgagccaga gtctgacgat	3600
aaccttaggc tcagccgcca tggggttcag cccaacgcaa cagcttggtc tgagcctagc	3660
cataactctg ccaccatgga gtggaggctg gccgagaacc tctcggttga ccttagggct	3720
aactctcgcc gccatggtca cctcacaaca tccgaaatgg agtttgacat taggcttaac	3780
aacgccacca tggagttcaa caggcagtct ggcttctgt ctgagattag gactaaagac	3840
agccgccatg gcctctctcc tctccgagac ctgttcacta gggtgagct taactgccc	3900
accatggtgt cagaggccac caggatcgac ccaaatagtt gtgaggataa gtggagccgc	3960
catggacact accaggccga ggctgccctg cttagctgtg aagctaacca gcgaatgcct	4020
ggggctctca tcaccacagc ccacgcttgt agccctgaag ctaagacagc gaatgcctgg	4080
ggcaagacct acagaatcga cggccatagg cctgaggcta accagcgaat gcctggggct	4140
gcctccgaca cacctcacc tgctaggatt gagcttaact cagcgaatgc ctggggcgca	4200
gagagggtca actggctggg tgagggtaat cataggcagc gaatgcctgg ggccacagct	4260
gcctgcttcg acacctgtga gcttaacctt agcgcagcga atgcctgggg cgtgttcctt	4320

NaDoc 002-PCT-1_seq listing_ST25.txt

tccgagaacg gccttgagtg taacactagg cagcgaatgc ctggggctca ccagtggagg	4380
ggagacttgc ctgacagtaa ctctaggcca gcgaatgcct ggggcatgga aacctctcac	4440
agacagcttg accataagga taggcagcga atgcctgggg ccgacggctt ccacatgggc	4500
attggtgaag ataactctag ctacgcgaat gcctggggcg agttcccac gcgtcggctc	4560
ggctagccgt acgctcctta gcgacgaaat ctactgcccc cctgagagcc accatggctt	4620
ggggtcctac gctgtgcagg ccaagtttgg agattacaac aaagaaggcc gccatggtgg	4680
gcacctcagc tctgagcggc tcatccgcca ccatgggttg gaccagcaca aacttaccag	4740
ggaccgccgc catggccgga cccaggcgtg ccaccatgga caccgtgggt tgcgccgcca	4800
tggtgctctg ttggagtgcc accatggtgc tcaggacctg ggccgccatg gaatacctga	4860
taactgataa gccaccatgg gaacagacct ttggcttgga gttgacgcc ttggactcaa	4920
catttacgag gccgccatgg agttcacccc aaagattggc tttccttgga gtgaaatcag	4980
gaacatctct gccaccatgg aaaagtttgt catcaagccc atcgacaagg ccgccatgga	5040
ctttgtgttt tacgccccac gtctcacagc caccatgggg accctgcagc tcgccgccat	5100
ggaccacgag ttgtacgcca ccatggggaa gcctgacacc gccgccatgg agcagacgaa	5160
ggccgccacc atggaggctg ataagctgat aagccgccat gggctggaaa cagagaagaa	5220
aaggagagaa accgtggaga gagagaaaga gcgccaccat ggcgagaagg aggagtgtt	5280
gctgcggctg caggactacg aggagaagac aagccgccat gggagagacc tctcggagca	5340
gattcagagg gccaccatg gggaggagga gaggaagcgg gcacaggagg gccgccatgg	5400
cccagaggct gaccgccacc atggactgcg ggctaaggcg cgccatggga gacaggcgg	5460
gggccaccat gggagccagg agcagcgccg ccatgggcta cctgataact gataagccac	5520
catggtggaa gaggcgcgga ggcgcaagga ggacgaagtt gaagagtggc agcaagccgc	5580
catggaagcc caggacgacc tggtaagac caaggaggag ctgcacctgg tgccggccac	5640
catggcgcca ccaccaccac ccgtgtacga gccggcgcc atggacgtcc aggagagctt	5700
gcaagacgag ggtgccacca tggcgggcta cagcgagcc gccatggctg acggcatccg	5760
ggccaccatg gacgaggaga agcgtgccgc catggcagag aagaacgagg ccaccatggg	5820
gcctgataag ctgataagcc gccatggggc ccgagacgag aacaagagga cccacaacga	5880
catcatccac aacgagagcc accatggagg ccgggacaag tacaagacgc tgcggcagat	5940
ccggcagggc aacaccagcc gccatggcga cgagttcgag gccctgcaac agccaggcca	6000
ccatggaggg cagaggggtg ctcatagcgg gcgctgccgc catggccacg cttgtgtctg	6060
ccaccatgga agtctcggaa ctcgccgcca tggcagttcc tttcgaagcc accatggcaa	6120
cagaaacatt cgccgccatg gaccacctga taactgataa gccaccatgg ttgcaatcgt	6180
gccaaagcagg cctgattctc gcgattactc gcgaatcacc gccgccatgg tgctgggagc	6240
aggactcatt gaattacgga aaacgcctgt caagtctcag gccaccatgg ggaactggcc	6300
tgtgtcatat aagagtcagg ccgccatggg gaaacgtggc aggacttcca tctgtgccgc	6360

NaDoc 002-PCT-1_seq listing_ST25.txt

caccatggtg tattcgaaac gagccgcat ggattttctc atctctgcca ccatggcatc	6420
tttgtagatt gccgcatgg gaggggtcaa aattgccacc atggtggctg ataagttgat	6480
agtaaccgcc atggtgtttc atccagtcgc caccatgggc tggcagagag cagccgcat	6540
ggcagcgtca gtggtggcca ccatggcttg gatttttttt ttgttttttt tttttttg	6600
tcaacaattt tacaacacat tgtgtcgacg agctcaagct tcccggcgcg ccccggtccg	6660
tccggactac ggcaagctga ccctgaagtt catcccaaaa cttacgctga gtacttcgat	6720
ctggtcacc cggatctcgc cgccatggga gctgatgatg tggttgattc ttcgaaatct	6780
tttgtcatgg aaaacttttc ttcgtaccac gggacgaaac ctggttatgt ggattccatt	6840
caaaaaggca taaaaagcc aaaatctggt acacaaggaa actatgacga tgattggaaa	6900
gggttttata gtaccgacaa caaatatgac gctgcgggat actctgtgga taatgaaaac	6960
ccgctctctg gaaaagctgg aggcgtggtc aaagtgcgt atccaggact gacgaaggtt	7020
ctcgactaa aggtggataa tgccgaaact attaagaaag agttaggttt aagtctcact	7080
gaaccgctca tggagcaagt cggaacggaa gagtttatca aaagattcgg tgatggtgct	7140
tcgctgtag tgctcagcct tcccttcgct gaggggagtt ctacgcttga gtacatcaac	7200
aactgggaac aggcgaaagc gttaagcgta gaacttgaga ttaactttga aaccctgga	7260
aaacgtggcc aagatgcgat gtatgagtat atggctcaag cctgtgcagg aaatcgtgct	7320
aggcgatagt gaactagtat ccggaatcta gagcggccgc actcagggtt taaacggccg	7380
gccgcggtca tagctgtttc ctgaacagat cccgggtggc atccctgtga cccctcccca	7440
gtgcctctcc tggccctgga agttgccact ccagtggcca ccagccttgt cctaataaaa	7500
ttaagttgca tcattttgct tgactagggtg tccttctata atattatggg gtggaggggg	7560
gtggtatgga gcaaggggca agttgggaag acaacctgta gggcctgcgg ggtctattgg	7620
gaaccaagct ggagtgcagt ggcacaatct tggctcactg caatctccgc ctctgggtt	7680
caagcgattc tcctgcctca gcctcccagag ttgttgggat tccaggcatg catgaccagg	7740
ctcagctaat tttgttttt ttggtagaga cgggggttca ccatattggc cagctggctc	7800
tccaactcct aatctcaggt gatctaccca ccttggcctc ccaaattgct gggattacag	7860
gcgtgaacca ctgctccctt ccctgtcctt ctgattttaa aataactata ccagcaggag	7920
gacgtccaga cacagcatag gctacctggc catgccaac cgggtgggaca tttagattgc	7980
ttgcttgga ctgtcctctc atgcgttggg tccactcagt agatgcctgt tgaattgggt	8040
acgcggccag cttggctgtg gaatgtgtgt cagttagggt gtggaaagtc cccaggctcc	8100
ccagcaggca gaagtatgca aagcatgcat ctcaattagt cagcaaccag gtgtggaaag	8160
tccccaggct cccagcagg cagaagtatg caaagcatgc atctcaatta gtcagcaacc	8220
atagtccgc ccctaactcc gccatcccg cccctaactc cgccagttc cgccattct	8280
ccgccccatg gctgactaat ttttttatt tatgcagagg ccgaggccgc ctcggcctct	8340
gagctattcc agaagtagtg aggaggcttt ttggaggcc taggcttttg caaaaagctc	8400

NaDoc 002-PCT-1_seq listing_ST25.txt

ccgggagctt gtatatccat tttcggatct gatcaagaga cacgtacgac catggagagc	8460
gacgagagcg gcctgcccgc catggagatc gagtgccgca tcaccggcac cctgaacggc	8520
gtggagttcg agctggtggg cggcggagag ggcacccccg agcagggccg catgaccaac	8580
aagatgaaga gcaccaaagg cgccctgacc ttcagcccct acctgctgag ccacgtgatg	8640
ggctacggct tctaccactt cggcacctac cccagcggct acgagaaccc cttcctgcac	8700
gccatcaaca acggcggcta caccaacacc cgcacgaga agtacgagga cggcggcgctg	8760
ctgcacgtga gcttcagcta ccgctacgag gccggcccgcg tgatcggcga cttcaagggtg	8820
atgggacacc gcttccccga ggacagcgtg atcttcaccg acaagatcat ccgcagcaac	8880
gccaccgtgg agcacctgca ccccatgggc gataacgac tggtatggcag cttcacccgc	8940
accttcagcc tgcgcgacgg cggctactac agctccgtgg tggacagcca catgcacttc	9000
aagagcgcca tccaccccag catcctacag aacggggggc ccatgttcgc cttccgcccgc	9060
gtggaggagg atcacagcaa caccgagctg ggcacgtggt agtaccagca cgccttcaag	9120
accccggtat cagatgccgg tgaagaataa ctgcagcggg actctggggt tcgaaatgac	9180
cgaccaagcg acgcccaccc tgccatcacg agatttcgat tccaccgccg ctttctatga	9240
aaggttgggc ttcggaatcg ttttccggga cggcggctgg atgatcctcc agcgcgggga	9300
tctcatgctg gagttcttcg cccaccccaa cttgtttatt gcagcttata atgggttaca	9360
ataaagcaat agcatcaca atttcacaaa taaagcattt ttttactgc attctagttg	9420
tggtttgtcc aaactcatca atgtatctta tcatgtctgt ataccgtcga cctctagcta	9480
gagcttgccg taatcatggt catagctgtt tcctgtgtga aattgttatc cgctcacaat	9540
tccacacaac atacgagccg gaagcataaa gtgtaaagcc tgggggtgcct aatgagtga	9600
ctaactcaca ttaattgcgt tgcgctcact gcccgctttc cagtcgggaa acctgtcgtg	9660
ccagctgcat taatgaatcg gccaacgcgc ggggagaggc ggtttgcgta ttgggcgctc	9720
ttccgcttcc tcgctcactg actcgtcgcg ctcggctcgtt cggctgcggc gagcggtatc	9780
agctcactca aaggcggtaa tacggttatc cacagaatca ggggataacg caggaaagaa	9840
catgtgagca aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcggt	9900
tttccatagg ctccgcccc ctgacgagca tcacaaaaat cgacgctcaa gtcagagggtg	9960
gcgaaacccg acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtcgc	10020
ctctcctgtt ccgacctgc cgtttaccgg atacctgtcc gcctttctcc cttcgggaag	10080
cgtggcgctt tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc	10140
caagctgggc tgtgtgcacg aacccccgt tcagcccgac cgctgcgcct tatccggtaa	10200
ctatcgtctt gagtccaacc cggtaaagaca cgacttatcg ccaactggcag cagccactgg	10260
taacaggatt agcagagcga ggtatgtagg cgggtgtaca gagttcttga agtggtggcc	10320
taactacggc tacactagaa gaacagtatt tggatatctg gctctgctga agccagttac	10380
cttcggaaaa agagttggta gctcttgatc cggcaaaaa accaccgctg gtagcgggtg	10440

NaDoc 002-PCT-1_seq listing_ST25.txt

```

tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt 10500
gatcttttct acggggtctg acgctcagtg gaacgaaaac tcacgttaag ggattttggt 10560
catgagatta tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaagttttta 10620
atcaatctaa agtatatatg agtaaacttg gtctgacagt taccaatgct taatcagtga 10680
ggcacctatc tcagcgatct gtctatttcg ttcattccata gttgcctgac tccccgtcgt 10740
gtagataact acgatacggg agggcttacc atctggcccc agtgctgcaa tgataccgcg 10800
agaccacgc tcaccggctc cagatttatc agcaataaac cagccagccg gaagggccga 10860
gcgcagaagt ggtcctgcaa ctttatccgc ctccatccag tctattaatt gttgccggga 10920
agctagagta agtagttcgc cagttaatag tttgcgcaac gttgttgcca ttgctacagg 10980
catcgtgggtg tcacgctcgt cgtttggtat ggcttcattc agctccggtt cccaacgatc 11040
aaggcgagtt acatgatccc ccatgttggtg caaaaaagcg gttagctcct tcggtcctcc 11100
gatcgttggtc agaagtaagt tggccgcagt gttatcactc atggttatgg cagcactgca 11160
taattctctt actgtcatgc catccgtaag atgcttttct gtgactgggtg agtactcaac 11220
caagtcattc tgagaatagt gtatgcggcg accgagttgc tcttgcccgg cgtcaatacg 11280
ggataatacc gcgccacata gcagaacttt aaaagtgtc atcattggaa aacgttcttc 11340
ggggcgaaaa ctctcaagga tcttaccgct gttgagatcc agttcgatgt aaccactcg 11400
tgcaccaaac tgatcttcag catcttttac tttcaccagc gtttctgggt gagcaaaaac 11460
aggaaggcaa aatgccgcaa aaaagggaat aaggcgaca cggaaatgtt gaatactcat 11520
actcttcctt tttcaatatt attgaagcat ttatcagggg tattgtctca tgagcggata 11580
catatttgaa tgtatttaga aaaataaaca aataggggtt ccgcgcacat tccccgaaa 11640
agtgccacct gacgcgccct gtagcggcgc attaaagcgc gcgggtgtgg tggttacgcg 11700
cagcgtgacc gctacacttg ccagcgccct agcgcgcgct ctttctgctt tcttcccttc 11760
ctttctcgcc acgttcgccg gtttccccg tcaagctcta aatcgggggc tccctttagg 11820
gttccgattt agtgctttac ggcacctga ccccaaaaaa cttgattagg gtgatggttc 11880
acgtagtggg ccatcgccct gatagcggg ttttcgccct ttgacgttg agtccacgtt 11940
ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggctctattc 12000
ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta 12060
acaaaaattt aacgcgaatt tt 12082

```

<210> 3
 <211> 9536
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 3
 aacaaaatat taacgcttac aatttccatt cgccattcag gctgcgcaac tgttggaag 60

NaDoc 002-PCT-1_seq listing_ST25.txt

ggcgatcggg	gcgggcctct	tcgctattac	gccagctggc	gaaaggggga	tgtgctgcaa	120
ggcgattaag	ttgggtaacg	ccagggtttt	cccagtcacg	acgttgtaaa	acgacggcca	180
gtgccaaagt	gatctataca	ttgaatcaat	attggcaatt	agccatatta	gtcattgggt	240
atatagcata	aatcaatatt	ggctattggc	cattgcatac	gttgatatcta	tatcataata	300
tgtacattta	tattggctca	tgtccaatat	gaccgccatg	ttgacattga	ttattgacta	360
gttattaata	gtaatcaatt	acgggggtcat	tagttcatag	cccatatatg	gagttccgcg	420
ttacataact	tacggtaaat	ggcccgcctg	gctgaccgcc	caacgacccc	cgcccattga	480
cgtcaataat	gacgtatggt	cccatagtaa	cgccaatagg	gactttccat	tgacgtcaat	540
gggtggagta	tttacggtaa	actgcccact	tggcagtaca	tcaagtgtat	catatgccaa	600
gtccgcccc	tattgacgtc	aatgacggta	aatggccgc	ctggcattat	gcccagtaca	660
tgaccttacg	ggactttcct	acttggcagt	acatctacgt	attagtcatc	gctattacca	720
tggtgatgcg	gttttggcag	tacaccaatg	ggcgtggata	gcggtttgac	tcacggggat	780
ttccaagtct	ccacccatt	gacgtcaatg	ggagtttggt	ttggcaccaa	aatcaacggg	840
actttccaaa	atgtcgtaat	aacccgcgcc	cgttgacgca	aatgggcggg	aggcgtgtac	900
gggtgggagg	ctatataagc	agagctcggt	tagtgaaccg	tcagaatttt	gtaatacgac	960
tcactatagg	gcggccggga	attcgtcgac	tggatccggg	acctagctag	gtagcaattg	1020
accggtcaag	atggcggcca	acaacaacaa	caacaacaac	aacaacaaca	acaacaacaa	1080
caacaagaag	atggcggcaa	caacaacaac	aacaacaaca	acaacaacaa	caacaacaac	1140
caacaacaag	atggcggcca	acaacaacaa	caacaacaac	aacaagaaga	tggcggcaac	1200
aacaacaaca	acaacaacaa	caaccaagat	ggcggccaac	aacaacaaga	agatggcggc	1260
aacaacaaca	accaagatgg	cggccaacaa	caacaagaag	atggcggcaa	caacaacaac	1320
caagatggcg	gcacgcgtcg	gtccggctag	ccgtacgctc	cttagcgacg	aaatctactg	1380
ccccctgag	agccaccatg	gcttggggtc	ctacgctgtg	caggccaagt	ttggagatta	1440
caacaaagaa	ggccgccatg	gtgggcacct	cagctctgag	cggctcatcc	gccaccatgg	1500
gttggaccag	cacaaactta	ccagggaccg	ccgccatggc	cggacccagg	cgtgccacca	1560
tggacaccgt	gggttgcgcc	gccatggtgc	tctgttgagg	tgccaccatg	gtgctcagga	1620
cctgggcccgc	catggaatac	ctgataactg	ataagccacc	atgggaacag	acctttggct	1680
tggagttgac	gcccttggac	tcaacattta	cgaggccgcc	atggagttca	ccccaaagat	1740
tggctttcct	tggagtgaag	tcaggaacat	ctctgccacc	atggaaaagt	ttgtcatcaa	1800
gcccacgac	aaggccgcca	tggactttgt	gttttacgcc	ccacgtctca	cagccaccat	1860
ggggaccctg	cagctcgccg	ccatggacca	cgagttgtac	gccaccatgg	ggaagcctga	1920
caccgcccgc	atggagcaga	cgaaggccgc	caccatggag	gctgataagc	tgataagccg	1980
ccatgggctg	gaaacagaga	agaaaaggag	agaaaccgtg	gagagagaga	aagagcgcca	2040
ccatggcgag	aaggaggagt	tgttgctgcg	gctgcaggac	tacgaggaga	agacaagccc	2100

NaDoc 002-PCT-1_seq listing_ST25.txt

ccatggggaga gacctctcgg agcagattca gagggggccac catggggagg aggagaggaa	2160
gcgggcacag gagggccgcc atggcccaga ggctgaccgc caccatggac tgcgggctaa	2220
gggcccgaat gggagacagg cgggtggcca ccatgggagc caggagcagc gccgccatgg	2280
gctacctgat aactgataag ccacatggt ggaagaggcg cggaggcgca aggaggacga	2340
agttgaagag tggcagcaag ccgcatgga agcccaggac gacctggtca agaccaagga	2400
ggagctgcac ctggtgccgg ccacatggc gccaccacca ccaccgtgt acgagccggc	2460
cgccatggac gtccaggaga gcttgcaaga cgagggtgcc accatggcg gctacagcgc	2520
agccgccatg gctgacggca tccgggccac catggacgag gagaagcgtg ccgccatggc	2580
agagaagaac gagggcacca tggggcctga taagctgata agccgccatg gggcccaga	2640
cgagaacaag aggaccaca acgacatcat ccacaacgag agccaccatg gaggccggga	2700
caagtacaag acgctgcggc agatccggca gggcaacacc agccgccatg gcgacgagtt	2760
cgaggccctg caacagccag gccacatgg agggcagagg ggtgctcata gcgggcgctg	2820
ccgccatggc cacgcttggt tctgccacca tggaggtctc ggaactcgcc gccatggcag	2880
ttcctttcga agccaccatg gcaacagaaa cattcgccgc catggaccac ctgataactg	2940
ataagccacc atggttgcaa tcgtgccaag caggcctgat tctcgcgatt actcgcgaat	3000
caccgccgcc atggtgctgg gagcaggact cattgaatta cggaaaacgc ctgtcaagtc	3060
tcaggccacc atggggaact ggcctgtgtc atacaagagt caggccgcca tggggaaacg	3120
tggcaggact tccatctgtg ccgccaccat ggtgtattcg aaacgagccg ccatggattt	3180
tctcatctct gccaccatgg catctttgta cattgccgcc atgggagggg tcaaaattgc	3240
caccatggtg gctgataagt tgatagtaac cgccatggtg tttcatccag tcgccaccat	3300
gggctggcag agagcagccg ccatggcagc gtcagtgggt gccaccatgg cttggatttt	3360
tttttttggt tttttttttt ttgctcaaca attttacaac acattgtgtc gacgagctca	3420
agcttcccgg cgcgccccgg tccgtccgga ctacggcaag ctgaccctga agttcatccc	3480
aaaacttacg ctgagtactt cgatctgggt accccggatc cgtgatagta acctgatagt	3540
aacctgataa tagcagatct gcagcttggg gtatcagtca cattcggtg gtacccctcc	3600
ggaagcgaat gggagccgac gatgtggtcg attcttcgaa atcttttgtc atggaaaact	3660
tttcttcgta ccacgggacg aaacctgggt atgtggattc cattcaaaaa ggtaggttta	3720
atgttcgtta gatatagttg cagcttctaa caaacatcaa aactgattat gcttaggggt	3780
tttcttttta ttttttaaca ggcatacaaa agccaaaatc tggtagacaa ggaaactacg	3840
acgacgattg gaaaggtgag gcaactcagg tgcaggactt ggactataaa cccaatggag	3900
aagatagccc ttcaacctct gtgacttttc taaagctact ttccccctt tttgccttag	3960
ggttttacag taccgacaac aaatacgacg ctgcgggata ctctgtggac aacgaaaacc	4020
cgctctctgg aaaagctgga ggcgtggtca aagtcacgta tccagggtcaa aggaaataaa	4080
tttttagaat ccatttattt gtactgaagt aaaagttcac atatgcaact tctatttaat	4140

NaDoc 002-PCT-1_seq listing_ST25.txt

aggTTAACTT cACAAACCTA tTCTGTACCA tagggCTCAC gAAAGTTCTC gCACTCAAAG	4200
tGgACAATGC cGAAACTATC aAGAAAGAGT tGgGTCTCTC tCTCACCGAA cCGCTCATGG	4260
agCAAGTCGG aACGGAAGAG tTTATCAAAA gATTCGGCGA tGgTGCTTCG cGTGTCGTGC	4320
tCAGCCTTCC cTTGCGCGAG gGgAGTTCCA gCGTCGAGTA cATCAACAAC tGgGAACAGG	4380
tATGAATGCA aTTGTTGGCA tCTTTTTTTA aAGTTATGTT tAAGATATGA agTTAAAATT	4440
attttCAAAT CTGTAGTTAG gCTAGTCATT AAAACTTTTT cCAGGTCAGA actTACGACC	4500
tGCTTTTATT tCCAAATAGG cGAAAGCGCT cAGCGTCGAA CTCGAGATCA actTCGAAAC	4560
ccGTGGAAAA cGTGGCCAAG aTGCGATGTA cGAGTATATG gCTCAAGCCT gTGCAGGTGG	4620
gCAGCTCATG agCCcAGGAG attCTGTCTT gTTTCTGTGC cTAGTGGAGT tTgTTAGTTT	4680
gCTGTGATTA gCTGGCAACG gAAACTGGAT tCATGTTGCA gAGgGTTTTT cTCATCTGGG	4740
tATTCTTGGT tttccACTTA cACTttCCCC gTCTTTTCTG tagGAAATCG tGTCAGGCGA	4800
tagTGAGCGG cCGcACTCGA gGtTTAAACG gCCgGCCGCG gTCATAGCTG tttcCTGAAC	4860
agATCCCGGG tGgCATCCCT gTGACCCCTC cCCAGTGCCT CTCCTGGCCC tGGAAGTTGC	4920
cACTCCAGTG cccACCAGCC tTGTCTAAT AAAATTAAGT tGCATCATTT tGTCTGACTA	4980
gGTGTCTTC tATAATATTA tGgGtGGAG gGgGtGGTA tGgAGCAAGG gGCAAGTTGG	5040
gaAGACAACC tGTAGgGCCT gCGgGtCTA tTgGGAACCA agCTGGAGTG cAGTGGCACA	5100
aTCTTGGCTC actGCAATCT cCGCCTCTG gGtTCAAGCG attCTCTGC CTCAGCCTCC	5160
cGAGtTgTTG gGATTCCAGG cATGCATGAC cAGGCTCAGC tAATTTTTGT tTTTTTGTA	5220
gAGACgGGGT tTcACCATAT tGgCCACGCT gGTCTCCAAC tCCTAATCTC agGTGATCTA	5280
cccACCTTGG cCTCCCAAAT tGCTgGGATT aCAGGCGTGA accACTGCTC cTTCCCTGT	5340
cTTTCTGATT tTAAATAAC tATACCAGCA gGAGGACGTC cAGACACAGC aTAGGCTACC	5400
tGgCCATGCC CAACCGGTGG gACATTGAG tTgCTTGCTT gGcACTGTCC tCTCATGCGT	5460
tGgGTCCACT cAGTAGATGC CTGTTGAATT gGgtACGCGG cCAGCTTGGC tGTGGAATGT	5520
gTGTcAGTTA gGgtGTGGAA agTCCCCAGG CTCCCCAGCA gGcAGAAGTA tGCAAAGCAT	5580
gCATCTCAAT tagTCAGCA cCAGGTGTGG AAAGTCCCCA gGCTCCCCAG cAGGCAGAAG	5640
tATGCAAAGC aTGCATCTCA attAGTCAGC aACCATAGTC cCGCCCCtaa CTCGCCCCAT	5700
ccCGCCCCta actCGCCCCA gTtCGCCCC tTtCGCCCC cATgGCTGAC tAATTTTTTT	5760
tATTTATGCA gAGGCGAGG cCGCCTCGGC CTCTGAGCTA tTCCAGAAGT agTGAGGAGG	5820
CTTTTTTGA gGCCTAGGCT tTtGAAAAA gTCCCCGGA gCTTGTATAT cCATTttCGG	5880
aTCTGATCAA gAGACAGTA cGACCATGGA gAGCGACGAG agCGGCCTGC cCGCCATGGA	5940
gATCGAGTGC cGCATCACCg gCACCTGAA cGgCGTGGAG tTCGAGCTGG tGgGCGGCGG	6000
agAGggcACC cCGAGCAGG gCCGcATGAC CAACAAGATg aAGAGCACCA aAGGCGCCCT	6060
gACCTTCAGC cCTACCTGC tGAGCCACGT gatGGGCTAC gGCTTCTACC actTCGGCAC	6120
CTACCCcAGC gGCTACGAGA aCCCTTCTCT gCAGCCATC AACAACGGCG gCTACACCAa	6180

NaDoc 002-PCT-1_seq listing_ST25.txt

cacccgcac	gagaagtacg	aggacggcgg	cgtgctgcac	gtgagcttca	gctaccgcta	6240
cgaggccggc	cgcgatgatc	gcgacttcaa	ggtgatgggc	accggcttcc	ccgaggacag	6300
cgtgatcttc	accgacaaga	tcatccgcag	caacgccacc	gtggagcacc	tgcaccccat	6360
gggcgataac	gatctggatg	gcagcttcac	ccgcaccttc	agcctgcgcg	acggcggcta	6420
ctacagctcc	gtggtggaca	gccacatgca	cttcaagagc	gccatccacc	ccagcatcct	6480
acagaacggg	ggcccatgt	tcgccttccg	ccgcgtggag	gaggatcaca	gcaacaccga	6540
gctgggcac	gtggagtacc	agcacgcctt	caagaccccc	gatgcagatg	ccggtgaaga	6600
ataactgcag	cgggactctg	gggttcgaaa	tgaccgacca	agcgacgccc	aacctgccat	6660
cacgagattt	cgattccacc	gccgccttct	atgaaagggt	gggcttcgga	atcgttttcc	6720
gggacgccgg	ctggatgatc	ctccagcgcg	gggatctcat	gctggagttc	ttcgcccacc	6780
ccaacttggt	tattgcagct	tataatgggt	acaaataaag	caatagcatc	acaaatttca	6840
caaataaagc	atTTTTTTca	ctgcattcta	gttgtggttt	gtccaaactc	atcaatgtat	6900
cttatcatgt	ctgtataccg	tcgacctcta	gctagagctt	ggcgtaatca	tggatcatagc	6960
tgtttcctgt	gtgaaattgt	tatccgctca	caattccaca	caacatacga	gccggaagca	7020
taaagtgtaa	agcctggggg	gcctaagtga	tgagctaact	cacattaatt	gcgttgcgct	7080
cactgccgc	tttccagtcg	ggaaacctgt	cgtgccagct	gcattaatga	atcggccaac	7140
gcgcggggag	aggcggtttg	cgtattgggc	gctcttccgc	ttcctcgctc	actgactcgc	7200
tgcgctcgg	cgttcggctg	cggcgagcgg	tatcagctca	ctcaaaggcg	gtaatacgg	7260
tatccacaga	atcaggggat	aacgcaggaa	agaacatgtg	agcaaaaggc	cagcaaaagg	7320
ccaggaaccg	taaaaaggcc	gcgttgctgg	cgTTTTTcca	taggctccgc	ccccctgacg	7380
agcatcacia	aatcgacgc	tcaagtcaga	ggtggcgaaa	cccgacagga	ctataaagat	7440
accaggcggt	tccccctgga	agctccctcg	tgcgctctcc	tgttccgacc	ctgccgctta	7500
ccggatacct	gtccgccttt	ctcccttcgg	gaagcgtggc	gctttctcat	agctcacgct	7560
gtaggtatct	cagttcgggtg	taggtcgctt	gctccaagct	gggctgtgtg	cacgaacccc	7620
ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	tcttgagtcc	aacccggtaa	7680
gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	gattagcaga	gcgaggatg	7740
taggcgggtg	tacagagttc	ttgaagtgg	ggcctaacta	cggctacact	agaagaacag	7800
tatttggtat	ctgcgctctg	ctgaagccag	ttaccttcgg	aaaaagagtt	ggtagctctt	7860
gatccggcaa	acaaaccacc	gctggtagcg	gtggTTTTTT	tgTTTgcaag	cagcagatta	7920
cgcgcagaaa	aaaaggatct	caagaagatc	ctttgatctt	ttctacgggg	tctgacgctc	7980
agtggaaacga	aaactcacgt	taagggattt	tggatcatgag	attatcaaaa	aggatcttca	8040
cctagatcct	tttaaattaa	aatgaagtt	ttaaataaat	ctaaagtata	tatgagtaaa	8100
cttggtctga	cagttaccaa	tgcttaatca	gtgaggcacc	tatctcagcg	atctgtctat	8160
ttcgttcatc	catagttgcc	tgactccccg	tcgtgtagat	aactacgata	cgggagggct	8220

NaDoc 002-PCT-1_seq listing_ST25.txt

taccatctgg cccagtgct gcaatgatac cgcgagaccc acgctcaccg gctccagatt	8280
tatcagcaat aaaccagcca gccggaaggc cggagcgagc aagtggctct gcaactttat	8340
ccgcctccat ccagctctatt aattggtgcc gggaagctag agtaagtagt tcgccagtta	8400
atagtttgcg caacgttggt gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg	8460
gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga tccccatgt	8520
tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt tgtcagaagt aagttggccg	8580
cagtgttatc actcatggtt atggcagcac tgcataattc tcttactgtc atgccatccg	8640
taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc	8700
ggcgaccgag ttgtcttgc ccggcgtaa tacgggataa taccgcgcca catagcagaa	8760
ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg aaaactctca aggatcttac	8820
cgctgttgag atccagttcg atgtaacca ctcgtgcacc caactgatct tcagcatctt	8880
ttactttcac cagcgtttct gggtagcaa aaacaggaag gcaaaatgcc gcaaaaaagg	8940
gaataagggc gacacggaaa tgtgaatac tcatactctt cttttttcaa tattattgaa	9000
gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata	9060
aacaaatagg ggttcgcgc acatttccc gaaaagtgcc acctgacgcg ccctgtagcg	9120
gcgcattaag cgcggcggt gtgggtgta cgcgcagcgt gaccgctaca cttgccagcg	9180
ccctagcgcc cgctccttc gtttcttcc cttcctttct cgccacgttc gccggcttcc	9240
cccgtaagc tctaaatcgg gggctccct tagggttccg atttagtgct ttacggcacc	9300
tcgaccccaa aaaacttgat tagggtagt gtacacgtag tgggcatcg ccctgataga	9360
cggtttttcg ccctttgacg ttggagtcca cgttctttta tagtggactc ttgttcaaaa	9420
ctggaacaac actcaaccct atctcggctt attcttttga ttataaaggg attttgccga	9480
tttcggccta ttgggttaaaa aatgagctga tttaacaaaa atttaacgcg aatttt	9536

<210> 4
 <211> 9174
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 4	
aacaaaatat taacgcttac aatttccatt cgccattcag gctgcgcaac tgttggaag	60
ggcgatcggg gcgggcctct tcgctattac gccagctggc gaaaggggga tgtgctgcaa	120
ggcgattaag ttgggtaacg ccagggtttt ccagtcacg acgttgtaaa acgacggcca	180
gtgccaagct gatctataca ttgaatcaat attggcaatt agccatatta gtcattgggt	240
atatagcata aatcaatatt ggctattggc cattgcatac gttgtatcta tatcataata	300
tgtacattta tattggctca tgtccaatat gaccgccatg ttgacattga ttattgacta	360
gttattaata gtaatcaatt acgggggtcat tagttcatag cccatatatg gagttccgcg	420

NaDoc 002-PCT-1_seq listing_ST25.txt

ttacataact tacggtaa	ggcccgcc	gctgaccgcc	caacgacccc	cgccattga	480
cgtcaataat gacgtatgtt	cccatagtaa	cgccaatagg	gactttccat	tgacgtcaat	540
gggtggagta ttacggtaa	actgcccact	tggcagtaca	tcaagtgtat	catatgccaa	600
gtccgcccc tattgacgtc	aatgacggta	aatggccgc	ctggcattat	gcccagtaca	660
tgaccttacg ggactttcct	acttggcagt	acatctacgt	attagtcac	gctattacca	720
tggatgatgcg gttttggcag	tacaccaatg	ggcgtggata	gcggtttgac	tcacggggat	780
ttccaagtct ccacccatt	gacgtcaatg	ggagtttgtt	ttggcaccaa	aatcaacggg	840
actttccaaa atgtcgtaat	aaccccgccc	cgttgacgca	aatggcggt	aggcgtgtac	900
ggtgggaggt ctatataagc	agagctcgtt	tagtgaaccg	tcagaat	gtaatacgac	960
tcactatagg gcggccggga	attcgtcgac	tggatccggt	acctagctag	gtagcaattg	1020
accggtcaag atggcgcca	acaacaaca	caacaacaac	aacaacaaca	acaacaaca	1080
caacaagaag atggcgcca	caacaacaac	aacaacaaca	acaacaaca	caacaacaac	1140
caacaacaag atggcgcca	acaacaaca	caacaacaac	aacaagaaga	tggcggaac	1200
aacaacaaca acaacaaca	caaccaagat	ggcgccaac	aacaacaaga	agatggcggc	1260
aacaacaaca accaagatgg	cggccaaca	caacaagaag	atggcgcca	caacaacaac	1320
caagatggcg gcacgcgtcg	gtccggctag	ccgtacgctc	cttagcgacg	aaatctactg	1380
ccccctgag agccaccatg	gcttggggtc	ctacgctgtg	caggccaagt	ttggagatta	1440
caacaaagaa ggccgcatg	gtgggcacct	cagctctgag	cggctcatcc	gccaccatgg	1500
gttgaccag cacaaactta	ccagggaccg	ccgccatggc	cggaccagg	cgtgccacca	1560
tggacaccgt gggttgcgcc	gccatggtgc	tctgttggag	tgccaccatg	gtgctcagga	1620
cctgggccgc catggaatac	ctgataactg	ataagccacc	atgggaacag	acctttggct	1680
tggagttgac gcccttggac	tcaacattta	cgaggccgcc	atggagttca	cccaaagat	1740
tggctttcct tggagtga	tcaggaacat	ctctgccacc	atggaaaagt	ttgtcatcaa	1800
gcccatcgac aaggccgcca	tggactttgt	gttttacgcc	ccacgtctca	cagccaccat	1860
ggggaccctg cagctcgccg	ccatggacca	cgagttgtac	gccaccatgg	ggaagcctga	1920
caccgccgc atggagcaga	cgaaggccgc	caccatggag	gctgataagc	tgataagccg	1980
ccatgggctg gaaacagaga	agaaaaggag	agaaaccgtg	gagagagaga	aagagcgcca	2040
ccatggcgag aaggaggagt	tggtgctgcg	gctgcaggac	tacgaggaga	agacaagccg	2100
ccatgggaga gacctctcgg	agcagattca	gaggggccac	catggggagg	aggagaggaa	2160
gcgggcacag gagggccgcc	atggcccaga	ggctgaccgc	caccatggac	tgcggtctaa	2220
gggcccgc atggagacagg	cgggtggcca	ccatgggagc	caggagcagc	gccgccatgg	2280
gctacctgat aactgataag	ccaccatggt	ggaagaggcg	cggaggcgca	aggaggacga	2340
agttgaagag tggcagcaag	ccgccatgga	agcccaggac	gacctggtca	agaccaagga	2400
ggagctgcac ctggtgccgg	ccaccatggc	gccaccacca	ccaccgtgt	acgagccggc	2460

NaDoc 002-PCT-1_seq listing_ST25.txt

cgccatggac gtccaggaga gcttgcaaga cgaggggtgcc accatggcgg gctacagcgc	2520
agccgccatg gctgacggca tccggggccac catggacgag gagaagcgtg ccgccatggc	2580
agagaagaac gaggccacca tggggcctga taagctgata agccgccatg gggcccgaga	2640
cgagaacaag aggaccaca acgacatcat ccacaacgag agccaccatg gaggccggga	2700
caagtacaag acgctgcggc agatccggca gggcaacacc agccgccatg gcgacgagtt	2760
cgaggccctg caacagccag gccaccatgg agggcagagg ggtgctcata gcgggcgctg	2820
ccgccatggc cacgcttggtg tctgccacca tggaagtctc ggaactcgcc gccatggcag	2880
ttcctttcga agccaccatg gcaacagaaa cattcgccgc catggaccac ctgataactg	2940
ataagccacc atggttgcaa tcgtgccaag caggcctgat tctcgcgatt actcgcgaat	3000
caccgccgcc atggtgctgg gagcaggact cattgaatta cggaaaacgc ctgtcaagtc	3060
tcaggccacc atggggaact ggctgtgtc atacaagagt caggccgcca tggggaaacg	3120
tggcaggact tccatctgtg ccgccaccat ggtgtattcg aaacgagccg ccatggattt	3180
tctcatctct gccaccatgg catctttgta cattgccgcc atgggagggg tcaaaattgc	3240
caccatgggtg gctgataagt tgatagtaac cgccatgggtg tttcatccag tcgccaccat	3300
gggctggcag agagcagccg ccatggcagc gtcagtgggtg gccaccatgg cttggatttt	3360
tttttttgtt tttttttttt ttgctcaaca attttacaac acattgtgtc gacgagctca	3420
agcttcccgg cgcgccccgg tccgtccgga ctacggcaag ctgaccctga agttcatccc	3480
aaaacttacg ctgagtactt cgatctggtc accggtacca tgggagccga cgatgtggtc	3540
gattcttcga aatcttttgt catggaaaac ttttcttcgt accacgggac gaaacctggt	3600
tatgtggatt ccattcaaaa aggcatacaa aagccaaaat ctggtacaca aggaaactac	3660
gacgacgatt ggaaaggggt ttacagtacc gacaacaaat acgacgctgc gggatactct	3720
gtggacaacg aaaaccgct ctctggaaaa gctggaggcg tggtaaagt cacgtatcca	3780
ggtagtctc tagccctgcc ttgacctgtc ctctcagcac ttccattagc cagctaccta	3840
cttccatcca ctcccaaact tcagggtctt gcctgcccc agaggcacag gacttagttc	3900
tgggaccagg gatcaggccg cagccctggc ctgctgttgc ttctgtcagg gacttgccct	3960
tgacccagc ctctctgacc ctgagggtct ccttggggag ctcttctgaa tttgggctgg	4020
cagatacccc acccagacca ggtctgccg tgcggcaggg ccagtggggc aggttggctg	4080
tggctgctgt gccctagtct gccctttctg acttgcaggg ctcacgaagg ttctgcact	4140
caaggtggac aatgccgaaa ctatcaagaa agagttgggt ctcagcctca ccgaaccgct	4200
catggagcaa gtcggaacgg aagagtttat caaaagattc ggtgatgggt cttcgcgtgt	4260
agtgtcagc cttcccttcg ctgaggggag ttctagcgtt gagtacatca acaactggga	4320
acaggcgaag gcgttaagcg tagaacttga gattaacttt gaaacccgtg gaaaacgtgg	4380
ccaagatgcg atgtatgagt atatggctca agcctgtgca ggaaatcgtg tcaggcgata	4440
gtgagcggcc gcactcgagg tttaaacggc cggccgcggt catagctgtt tcctgaacag	4500

NaDoc 002-PCT-1_seq listing_ST25.txt

atccccgggtg gcatccctgt gacccctccc cagtgcctct cctggccctg gaagttgcca	4560
ctccagtgcc caccagcctt gtcctaataa aattaagttg catcattttg tctgactagg	4620
tgtccttcta taatattatg ggggtggaggg ggggtggtatg gagcaagggg caagttggga	4680
agacaacctg tagggcctgc ggggtctatt ggggaaccaag ctggagtgca gtggcacaat	4740
cttggtcac tgcaatctcc gcctcctggg ttcaagcgat tctcctgcct cagcctccc	4800
agttgttggg attccaggca tgcattgacca ggctcagcta atttttgttt ttttggtaga	4860
gacgggggtt caccatattg gccacgctgg tctccaactc ctaatctcag gtgatctacc	4920
caccttggtc tcccaaattg ctgggattac aggcgtgaac cactgctccc ttcctgtcc	4980
ttctgatttt aaaataacta taccagcagg aggacgtcca gacacagcat aggctacctg	5040
gcatgcccc accggtggga catttgagtt gcttgcttgg cactgtcctc tcatgcgttg	5100
gggtccactca gtagatgcct gttgaattgg gtacgcggcc agcttggttg tggaaatgtg	5160
gtcagttagg gtgtggaaag tccccaggct ccccgagcagg cagaagtatg caaagcatgc	5220
atctcaatta gtcagcaacc aggtgtggaa agtccccagg ctccccagca ggcagaagta	5280
tgaaagcat gcatctcaat tagtcagcaa ccatagtccc gccctaact ccgcccattc	5340
cgcccctaac tccgcccagt tccgcccatt ctccgcccc tggctgacta atttttttta	5400
tttatgcaga ggccgaggcc gcctcggcct ctgagctatt ccagaagtag tgaggaggct	5460
tttttgagg cctaggcttt tgcaaaaagc tcccgggagc ttgtatatcc attttcggat	5520
ctgatcaaga gacacgtacg accatggaga gcgacgagag cggcctgccc gccatggaga	5580
tcgagtgccg catcaccggc accctgaacg gcgtggagtt cgagctggtg ggcggcggag	5640
agggcacccc cgagcagggc cgcattgacca acaagatgaa gagcaccaaa ggcggcctga	5700
ccttcagccc ctacctgtg agccacgtga tgggtctacg cttctaccac ttcggcacct	5760
accccgagcg ctacgagaac cccttcctgc acgccatcaa caacggcggc tacaccaaca	5820
cccgcattca gaagtacgag gacggcggcg tgctgcacgt gagcttcagc taccgctacg	5880
aggccggccg cgtgatcggc gacttcaagg tgatgggcac cggcttcccc gaggacagcg	5940
tgatcttcac cgacaagatc atccgcagca acgcccacgt ggagcacctg caccatgg	6000
gcgataacga tctggatggc agcttcaccc gcaccttcag cctgcgcgac ggcggctact	6060
acagctccgt ggtggacagc cacatgcact tcaagagcgc catccacccc agcatcctac	6120
agaacggggg ccccatgttc gccttcgcc gcgtggagga ggatcacagc aacaccgagc	6180
tgggcatcgt ggagtaccag cagccttca agaccccgga tgcagatgcc ggtgaagaat	6240
aactgcagcg ggactctggg gttcgaaatg accgaccaag cgacgcccac cctgccatca	6300
cgagatttcg attccaccgc cgccttctat gaaagggttg gcttcggaat cgttttccg	6360
gacgccggct ggatgatcct ccagcgcggg gatctcatgc tggagttctt cccccacccc	6420
aacttgttta ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca	6480
aataaagcat ttttttact gcattctagt tgtggtttgt ccaaactcat caatgtatct	6540

NaDoc 002-PCT-1_seq listing_ST25.txt

tatcatgtct gtataccgtc gacctctagc tagagcttgg cgtaatcatg gtcatagctg	6600
tttctgtgt gaaattgtta tccgctcaca attccacaca acatacgagc cggaagcata	6660
aagtgtaaag cctgggggtgc ctaatgagtg agctaactca cattaattgc gttgcgctca	6720
ctgcccgtt tccagtcggg aaacctgtcg tgccagctgc attaatgaat cggccaacgc	6780
gcggggagag gcggtttgcg tattgggctg tcttccgctt cctcgctcac tgactcgctg	6840
cgctcggtcg ttcggctgcg gcgagcggta tcagctcact caaaggcggg aatacggtta	6900
tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc	6960
aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc ccctgacgag	7020
catcacaaaa atcgacgtc aagtcagagg tggcgaaacc cgacaggact ataaagatac	7080
caggcgtttc cccctggaag ctccctcggt cgctctcctg ttccgaccct gccgcttacc	7140
ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcatag ctacgctgt	7200
aggtatctca gttcgggtga ggtcgctgc tccaagctgg gctgtgtgca cgaaccccc	7260
gttcagcccg accgctgctg cttatccggg aactatcgct ttgagtccaa cccggtaaga	7320
cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta	7380
ggcggtgcta cagagttctt gaagtgggtg cctaactacg gctacactag aagaacagta	7440
tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga	7500
tccggcaaac aaaccaccgc tggtagcggg ggtttttttg tttgcaagca gcagattacg	7560
cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag	7620
tggaacgaaa actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc	7680
tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact	7740
tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt	7800
cgttcatcca tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggtta	7860
ccatctggcc ccagtgtctg aatgataccg cgagaccac gctcaccggc tccagattta	7920
tcagcaataa accagccagc cggaaaggcc gagcgcagaa gtggtcctgc aactttatcc	7980
gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat	8040
agtttgcgca acgttggtgc cattgtaca ggcacgtgg tgtcacgctc gtcgtttggt	8100
atggcttcat tcagctccgg ttccaacga tcaaggcgag ttacatgatc ccccatgttg	8160
tgcaaaaaag cggttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca	8220
gtgttatcac tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta	8280
agatgctttt ctgtgactgg tgagtactca accaagtcac tctgagaata gtgtatgcgg	8340
cgaccgagtt gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact	8400
ttaaaagtgc tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg	8460
ctgttgagat ccagttcgat gtaacccact cgtgcaccca actgatcttc agcatctttt	8520
actttcacca gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga	8580

NaDoc 002-PCT-1_seq listing_ST25.txt

ataagggcga cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc	8640
atttatcagg gttattgtct catgagcgga tacatatctg aatgtattta gaaaaataaa	8700
caaatagggg ttccgcgcac atttccccga aaagtgccac ctgacgcgcc ctgtagcggc	8760
gcattaagcg cggcgggtgt ggtggttacg cgcagcgtga ccgctacact tgccagcgcc	8820
ctagcgcgcc ctcttttcgc tttcttcctt tcctttctcg ccacgttcgc cggctttccc	8880
cgtcaagctc taaatcgggg gctcccttta gggttccgat ttagtgcttt acggcacctc	8940
gacccccaaa aacttgatta gggatggtt tcacgtatg ggccatcgcc ctgatagacg	9000
gtttttcgcc ctttgacgtt ggagtcacg ttctttaata gtggactctt gttccaaact	9060
ggaacaacac tcaaccctat ctcggtctat tcttttgatt tataagggat tttgccgatt	9120
tcggcctatt ggttaaaaaa tgagctgatt taacaaaaat ttaacgcgaa tttt	9174

<210> 5
 <211> 10034
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 5	
aacaaaatat taacgcttac aatttccatt cgccattcag gctgcgcaac tgttggaag	60
ggcgatcggg gcgggcctct tcgctattac gccagctggc gaaaggggga tgtgctgcaa	120
ggcgattaag ttgggtaacg ccagggtttt ccagtcacg acgttgtaaa acgacggcca	180
gtgccaagct gatctataca ttgaatcaat attggcaatt agccatatta gtcattggtt	240
atatagcata aatcaatatt ggctattggc cattgcatac gttgtatcta tatcataata	300
tgtacattta tattggctca tgtccaatat gaccgccatg ttgacattga ttattgacta	360
gttattaata gtaatcaatt acggggtcat tagttcatag cccatatatg gagttccgcg	420
ttacataact tacggtaaat ggccgcctg gctgaccgcc caacgacccc cgcccattga	480
cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat tgacgtcaat	540
gggtggagta ttacggtaa actgcccact tggcagtaca tcaagtgtat catatgcca	600
gtccgcccc tattgacgtc aatgacggtt aatggcccgc ctggcattat gccagtcaca	660
tgaccttacg ggactttcct acttggcagt acatctacgt attagtcac gctattacca	720
tggatgatcg gttttggcag tacaccaatg ggcgtggata gcggtttgac tcacggggat	780
ttccaagtct ccacccatt gacgtcaatg ggagtttgtt ttggcaccaa aatcaacggg	840
actttccaaa atgtcgtaat aaccccgcgc cgttgacgca aatgggcggg aggcgtgtac	900
ggtgggaggt ctatataagc agagctcgtt tagtgaaccg tcagaatttt gtaatacgac	960
tcactatagg gcggccggga attcgtcgac tggatccggg acctagctag gtagcaattg	1020
accggtcaag atggcggcca acaacaaca caacaacaac aacaacaaca acaacaaca	1080
caacaagaag atggcggcaa caacaacaac aacaacaaca acaacaaca caacaacaac	1140

NaDoc 002-PCT-1_seq listing_ST25.txt

caacaacaag atggcgcca acaacaaca caacaacaac aacaagaaga tggcggaac	1200
aacaacaaca acaacaaca caaccaagat ggcgccaac aacaacaaga agatggcggc	1260
aacaacaaca accaagatgg cggccaaca caacaagaag atggcggaac caacaacaac	1320
caagatggcg gcacgcgtcg gtccggctag ccgtacgctc cttagcgacg aaatctactg	1380
ccccctgag agccaccatg gcttggggtc ctacgctgtg caggccaagt ttggagatta	1440
caacaagaa ggccgcatg gtgggcacct cagctctgag cggctcatcc gccaccatgg	1500
gttgaccag cacaactta ccagggaccg ccgcatggc cggaccagc cgtgccacca	1560
tggacaccgt gggttgcgc gccatggtgc tctgttgag tgccaccatg gtgctcagga	1620
cctgggccgc catggaatac ctgataactg ataagccacc atgggaacag acctttggct	1680
tggagttgac gcccttgac tcaacattta cgaggccgc atggagttca ccccaaagat	1740
tggctttcct tggagtgaat tcaggaacat ctctgccacc atggaaaagt ttgtcatcaa	1800
gccatcgac aaggccgcca tggactttgt gttttacgcc ccacgtctca cagccaccat	1860
ggggaccctg cagctcgccg ccatggacca cgagttgtac gccaccatgg ggaagcctga	1920
caccgccgc atggagcaga cgaaggccg caccatggag gctgataagc tgataagccg	1980
ccatgggctg gaaacagaga agaaaaggag agaaaccgtg gagagagaga aagagcgcca	2040
ccatggcgag aaggaggagt tgttgctcg gctgcaggac tacgaggaga agacaagccg	2100
ccatgggaga gacctctcg agcagattca gagggccac catggggagg aggagaggaa	2160
gcgggcacag gagggccgc atggcccaga ggctgaccgc caccatggac tgcgggctaa	2220
ggggccgcat gggagacagg cgggtggcca ccatgggagc caggagcagc gccgccatgg	2280
gctacctgat aactgataag ccaccatggt ggaagaggcg cggaggcgca aggaggacga	2340
agttgaagag tggcagcaag ccgcatgga agcccaggac gacctggtca agaccaagga	2400
ggagctgcac ctggtgccg ccaccatggc gccaccacca ccaccgtgt acgagccggc	2460
cgccatggac gtccaggaga gcttgcaaga cgagggtgcc accatggcg gctacagcg	2520
agccgccatg gctgacggca tccgggccac catggacgag gagaagcgtg ccgccatggc	2580
agagaagaac gagggcacc tggggcctga taagtgata agccgccatg gggcccaga	2640
cgagaacaag aggaccaca acgacatcat ccacaacgag agccaccatg gagggcgga	2700
caagtacaag acgtgcggc agatccggca gggcaacacc agccgccatg gcgacgagtt	2760
cgaggccctg caacagccag gccaccatgg agggcagagg ggtgctcata gcggcgctg	2820
ccgccatggc cagcttgtg tctgccacca tggaaagtctc ggaactcgcc gccatggcag	2880
ttcctttcga agccaccatg gcaacagaaa cattcgccgc catggaccac ctgataactg	2940
ataagccacc atggttgcaa tcgtgccaag caggcctgat tctcgcgatt actcggaat	3000
caccgccgc atggtgctg gagcaggact cattgaatta cggaacgc ctgtcaagtc	3060
tcaggccacc atggggaact ggcctgtgtc atacaagagt caggccgcca tggggaacg	3120
tggcaggact tccatctgtg ccgccaccat ggtgtattcg aaacgagccg ccatggattt	3180

NaDoc 002-PCT-1_seq listing_ST25.txt

tctcatctct gccaccatgg catctttgta cattgccgcc atgggagggg tcaaaattgc	3240
caccatgggtg gctgataagt tgatagtaac cgccatgggtg tttcatccag tcgccaccat	3300
gggctggcag agagcagccg ccatggcagc gtcagtgggtg gccaccatgg cttggatttt	3360
tttttttggt tttttttttt ttgctcaaca atttttacaac acattgtgtc gacgagctca	3420
agcttcccgg cgcgccctgg ctgagctgta caagggtaag tctactgactg tctatgcctg	3480
ggaaaggggtg ggcaggagat ggggcagtgc agggaaaagtg gcactatgaa cccaactaca	3540
caaatcagcg atttcaacaa caactacaca aatcagcgat ttcaattgta ctaaccttct	3600
tctctttcct ctcttgacag gaggagccat catcgcccga tatcccaatc gcttaccgat	3660
tcagaatcta cggcaagctg accctgaagt tcatcaatcg cttaccgatt cagaatccct	3720
acggcaagct gaccctgaag ttcataatc gcttaccgat tcagaatccc tacggcaagc	3780
tgaccctgaa gttcatcaat cgcttaccga ttcagaatcc ctacggcaag ctgaccctga	3840
agttcatcaa tcgcttaccg attcagaatc cctacggcaa gctgaccctg aagttcatca	3900
atcgcttacc gattcagaat ccctacggca agctgaccct gaagttcatc aatcgcttac	3960
cgattcagaa tccctacggc aagctgaccc tgaagttcat caatcgctta ccgattcaga	4020
atccctacgg caagctgacc ctgaagttca tcagatctgc agcttgggggt atcagtcaca	4080
ttcggtggtg acccctccgg aagcgaatgg gagccgacga tgtggtcgat tcttcgaaat	4140
cttttgatcat ggaaaacttt tcttcgtacc acgggacgaa acctggttat gtggattcca	4200
ttcaaaaagg taggtttaat gttcgttaga tatagttgca gcttctaaca aacatcaaaa	4260
ctgattatgc ttagggtttt tctttttatt ttttaacagg catacaaaag ccaaaatctg	4320
gtacacaagg aaactacgac gacgattgga aagggtgaggc actcaggggtg caggacttgg	4380
actataaacc caatggagaa gatagccctt caacctctgt gacttttcta aagctacttt	4440
ccccctttt tgccttaggg ttttacagta ccgacaacaa atacgacgct gcgggatact	4500
ctgtggacaa cgaaaaccgg ctctctggaa aagctggagg cgtgggtcaaa gtcacgtatc	4560
cagggtcaaag gaaataaatt tttagaatcc atttatttgt actgaagtaa aagttcacat	4620
atgcaacttc tatttaatag gttaacttca caaacctatt ctgtaccata gggctcacga	4680
aagttctcgc actcaaagtg gacaatgccg aaactatcaa gaaagagttg ggtctctctc	4740
tcaccgaacc gctcatggag caagtcggaa cggaagagtt tatcaaaaga ttcggcgatg	4800
gtgcttcgcg tgtcgtgctc agccttcctt tcgccgaggg gagttccagc gtcgagtaca	4860
tcaacaactg ggaacaggta tgaatgcaat tggtggcatc tttttttaa gttatgttta	4920
agatatgaag ttaaaattat tttcaaatct gtagttaggc tagtcattaa aactttttcc	4980
aggtcagaac ttacgacctg cttttatttc caaataggcg aaagcgctca gcgtcgaact	5040
cgagatcaac ttcgaaaccc gtggaaaacg tggccaagat gcgatgtacg agtatatggc	5100
tcaagcctgt gcagggtggc agctcatgag ccaggagat tctgtcttgt ttctgtgcct	5160
agtggagttt gttagtttgc tgtgattagc tggcaacgga aactggattc atgttgcaga	5220

NaDoc 002-PCT-1_seq listing_ST25.txt

gggtttttct catctgggta ttcttggtt tccacttaca ctttccccgt cttttctgta	5280
ggaaatcgtg tcaggcgata gtgagcggcc gactcgagg tttaaacggc cggccgcggt	5340
catagctgtt tcctgaacag atccccgggtg gcatccctgt gacccctccc cagtgcctct	5400
cctggccctg gaagttgcca ctccagtgcc caccagcctt gtcctaataa aattaagttg	5460
catcattttg tctgactagg tgtccttcta taatattatg ggggtggagg ggggtggtatg	5520
gagcaagggg caagttggga agacaacctg tagggcctgc ggggtctatt gggaaccaag	5580
ctggagtga gtggcacaat cttggctcac tgcaatctcc gcctcctggg ttcaagcgat	5640
tctcctgcct cagcctccc agttgttggg attccaggca tgcagacca ggctcagcta	5700
atttttgttt ttttggtaga gacgggggtt caccatattg gccacgctgg tctccaactc	5760
ctaattctcag gtgatctacc caccttggcc tcccaaattg ctgggattac aggcgtgaac	5820
cactgctccc ttcctgtcc ttctgatttt aaaataacta taccagcagg aggacgtcca	5880
gacacagcat aggctacctg gccatgcccc accggtggga catttgagtt gcttgcttgg	5940
cactgtcctc tcatgcgttg ggtccactca gtagatgcct gttgaattgg gtacgcggcc	6000
agcttggtg tggaatgtgt gtcagttagg gtgtggaaag tccccaggct cccagcagg	6060
cagaagtatg caaagcatgc atctcaatta gtcagcaacc aggtgtggaa agtccccagg	6120
ctccccagca ggcagaagta tgcaaagcat gcatctcaat tagtcagcaa ccatagtccc	6180
gcccctaact ccgcccattc cgcccctaac tccgcccagt tccgcccatt ctccgcccc	6240
tggctgacta atttttttta tttatgcaga ggccgaggcc gcctcggcct ctgagctatt	6300
ccagaagtag tgaggaggct tttttggagg cctaggcttt tgcaaaaagc tcccgggagc	6360
ttgtatatcc attttcgat ctgatcaaga gacacgtacg accatggaga gcgacgagag	6420
cggcctgccc gccatggaga tcgagtgccg catcaccggc accctgaacg gcgtggagtt	6480
cgagctggtg ggcggcggag agggcacccc cgagcagggc cgcatgacca acaagatgaa	6540
gagcaccaaa ggcgccctga cttcagccc ctacctgctg agccacgtga tgggctacgg	6600
ctttaccac ttcggcacct acccagcgg ctacgagaac cccttcctgc acgccatcaa	6660
caacggcggc tacaccaaca cccgcatcga gaagtacgag gacggcggcg tgctgcacgt	6720
gagcttcagc taccgctacg aggcggccg cgtgatcggc gacttcaagg tgatgggcac	6780
cggcttcccc gaggacagcg tgatcttcac cgacaagatc atccgcagca acgccaccgt	6840
ggagcacctg caccatggtg gcgataacga tctggatggc agcttcaccc gcaccttcag	6900
cctgcgcgac ggcggctact acagctccgt ggtggacagc cacatgcact tcaagagcgc	6960
catccacccc agcatcctac agaacggggg ccccatgttc gccttcgcc gcgtggagga	7020
ggatcacagc aacaccgagc tgggcatcgt ggagtaccag cacgccttca agaccccgga	7080
tgcagatgcc ggtgaagaat aactgcagcg ggactctggg gttcgaaatg accgaccaag	7140
cgacgcccc cctgccatca cgagatttcg attccaccgc cgccttctat gaaagggttg	7200
gcttcggaat cgttttccgg gacgccggct ggatgatcct ccagcgcggg gatctcatgc	7260

NaDoc 002-PCT-1_seq listing_ST25.txt

tggagttctt cgcccccccc aacttgttta ttgcagctta taatggttac aaataaagca	7320
atagcatcac aaatttcaca aataaagcat ttttttctact gcatttctagt tgtggtttgt	7380
ccaaactcat caatgtatct tatcatgtct gtataccgtc gacctctagc tagagcttgg	7440
cgtaatcatg gtcatactgt tttcctgtgt gaaattgtta tccgctcaca attccacaca	7500
acatacgagc cggaagcata aagtgtaaag cctgggggtgc ctaatgagtg agctaactca	7560
cattaattgc gttgcgctca ctgcccgtt tccagtcggg aaacctgtcg tgccagctgc	7620
attaatgaat cggccaacgc gcggggagag gcggtttgcg tattgggagc tcttccgctt	7680
cctcgctcac tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta tcagctcact	7740
caaaggcggg aatacgggta tccacagaat caggggataa cgcaggaaag aacatgtgag	7800
caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata	7860
ggctccgccc ccctgacgag catcacaaaa atcgacgctc aagtcagagg tggcgaaacc	7920
cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcggt cgctctcctg	7980
ttccgaccct gccgcttacc ggatacctgt ccgcctttct cccttcggga agcgtggcgc	8040
tttctcatag ctacagctgt aggtatctca gttcgggtgta ggtcgttcgc tccaagctgg	8100
gctgtgtgca cgaaccccc gttcagccc accgctgagc cttatccggt aactatcgtc	8160
ttgagtcaa cccggtaaga cagcacttat cgccactggc agcagccact ggtaacagga	8220
ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtgggtg cctaactacg	8280
gctacactag aagaacagta tttggtatct gcgctctgct gaagccagtt accttcggaa	8340
aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggg ggtttttttg	8400
tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt	8460
ctacggggtc tgacgctcag tggaacgaaa actcacgtta agggattttg gtcatagagat	8520
tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct	8580
aaagtatata tgagtaaact tggcttgaca gttaccaatg cttaatcagt gaggcaccta	8640
tctcagcgat ctgtctatct cgttcatcca tagttgcctg actccccgtc gtgtagataa	8700
ctacgatacg ggagggctta ccatctggcc ccagtgtgc aatgataccg cgagaccac	8760
gctcaccggc tccagattta tcagcaataa accagccagc cggaagggcc gagcgcagaa	8820
gtggtcctgc aactttatcc gcctccatcc agtctattaa ttgttgccgg gaagctagag	8880
taagtagttc gccagttaat agtttgcgca acgttggtgc cattgctaca ggcatcgtgg	8940
tgtcacgctc gtcgttttgt atggcttcat tcagctccgg tcccaacga tcaaggcgag	9000
ttacatgatc ccccatgttg tgcaaaaaag cggttagctc cttcggtcct ccgatcgttg	9060
tcagaagtaa gttggccgca gtgttatcac tcattggttat ggcagcactg cataattctc	9120
ttactgtcat gccatccgta agatgctttt ctgtgactgg tgagtactca accaagtcac	9180
tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaata cgggataata	9240
ccgcgccaca tagcagaact ttaaaagtgc tcattcattgg aaaacgttct tcggggcgaa	9300

NaDoc 002-PCT-1_seq listing_ST25.txt

```

aactctcaag gatcttaccg ctgttgagat ccagttcgat gtaaccact cgtgcaccca 9360
actgatcttc agcatctttt actttcacca gcgtttctgg gtgagcaaaa acaggaaggc 9420
aaaatgccgc aaaaaaggga ataagggcga cacggaaatg ttgaatactc atactcttcc 9480
tttttcaata ttattgaagc atttatcagg gttattgtct catgagcgga tacatatttg 9540
aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgccac 9600
ctgacgcgcc ctgtagcggc gcattaagcg cggcgggtgt ggtggttacg cgcagcgtga 9660
ccgctacact tgccagcgcc ctagcgcggc ctcttttcgc tttcttccct tcctttctcg 9720
ccacgttcgc cggctttccc cgtaagctc taaatcgggg gctcccttta gggttccgat 9780
ttagtgcttt acggcacctc gaccccaaaa aacttgatta gggtgatggt tcacgtagt 9840
ggccatcgcc ctgatagacg gtttttcgcc ctttgacgtt ggagtccacg ttctttaata 9900
gtggactctt gttccaaact ggaacaacac tcaaccctat ctcggtctat tcttttgatt 9960
tataagggat ttgccgatt tcggcctatt ggttaaaaaa tgagctgatt taacaaaaat 10020
ttaacgcgaa tttt 10034

```

```

<210> 6
<211> 9441
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Synthetic

```

```

<400> 6
aacaaaatat taacgcttac aatttccatt cgccattcag gctgcgcaac tgttggaag 60
ggcgatcggg gcgggcctct tcgctattac gccagctggc gaaaggggga tgtgctgcaa 120
ggcgattaag ttgggtaacg ccagggtttt cccagtcacg acgttgtaaa acgacggcca 180
gtgccaagct gatctataca ttgaatcaat attggcaatt agccatatta gtcattggtt 240
atatagcata aatcaatatt ggctattggc cattgcatac gttgtatcta tatcataata 300
tgtacattta tattggctca tgtccaatat gaccgccatg ttgacattga ttattgacta 360
gttattaata gtaatcaatt acgggggtcat tagttcatag cccatatatg gagttccgcg 420
ttacataact tacggtaa at ggccgcctg gctgaccgcc caacgacccc cgccattga 480
cgtcaataat gacgtatggt cccatagtaa cgccaatagg gactttccat tgacgtcaat 540
gggtggagta ttacggtaa actgcccact tggcagtaca tcaagtgtat catatgccaa 600
gtccgccccc tattgacgtc aatgacggta aatggccgc ctggcattat gccagtaca 660
tgaccttacg ggactttcct acttggcagt acatctacgt attagtcac gctattacca 720
tggatgatgcg gttttggcag tacaccaatg ggcgtggata gcggtttgac tcacggggat 780
ttccaagtct ccacccatt gacgtcaatg ggagtttgtt ttggcaccaa aatcaacggg 840
actttccaaa atgtcgaat aaccccgccc cgttgacgca aatgggcggg aggcgtgtac 900
ggtgggaggt ctatataagc agagctcggt tagtgaaccg tcagaatttt gtaatacgac 960

```

NaDoc 002-PCT-1_seq listing_ST25.txt

tcactatagg gcggccggga attcgtcgac tggatccggt acctagctag gtagcaattg	1020
accggtcaag atggcgggca acaacaaca caacaacaac aacaacaaca acaacaaca	1080
caacaagaag atggcgggca caacaacaac aacaacaaca acaacaaca caacaacaac	1140
caacaacaag atggcgggca acaacaaca caacaacaac aacaagaaga tggcgggcaac	1200
aacaacaaca acaacaaca caaccaagat ggcggccaac aacaacaaga agatggcggc	1260
aacaacaaca accaagatgg cggccaaca caacaagaag atggcgggca caacaacaac	1320
caagatggcg gcacgcgtcg gtccggctag ccgtacgctc cttagcgacg aaatctactg	1380
ccccctgag agccaccatg gcttggggtc ctacgctgtg caggccaagt ttggagatta	1440
caacaaagaa ggccgccatg gtggggcacct cagctctgag cggctcatcc gccaccatgg	1500
gttgaccag cacaactta ccagggaccg ccgccatggc cggaccagg cgtgccacca	1560
tggacaccgt gggttgcgcc gccatggtgc tctgttggag tgccaccatg gtgctcagga	1620
cctgggccgc catggaatac ctgataactg ataagccacc atgggaacag acctttggct	1680
tggagttgac gcccttggac tcaacattta cgaggccgcc atggagtcca ccccaaagat	1740
tggctttcct tggagtgaag tcaggaacat ctctgccacc atggaaaagt ttgtcatcaa	1800
gcccatcgac aaggccgcca tggactttgt gttttacgcc ccacgtctca cagccaccat	1860
ggggaccctg cagctcgccg ccatggacca cgagttgtac gccaccatgg ggaagcctga	1920
caccgccgcc atggagcaga cgaaggccgc caccatggag gctgataagc tgataagccg	1980
ccatgggctg gaaacagaga agaaaaggag agaaaccgtg gagagagaga aagagcgcca	2040
ccatggcgag aaggaggagt tgttgctgcg gctgcaggac tacgaggaga agacaagccg	2100
ccatgggaga gacctctcg agcagattca gaggggccac catggggagg aggagaggaa	2160
gcgggcacag gagggccgcc atggcccaga ggctgaccgc caccatggac tgcgggctaa	2220
gggcccgcct gggagacagg cgggtgggcca ccatgggagc caggagcagc gccgccatgg	2280
gctacctgat aactgataag ccaccatggt ggaagaggcg cggaggcgca aggaggacga	2340
agttgaagag tggcagcaag ccgccatgga agcccaggac gacctggtca agaccaagga	2400
ggagctgcac ctggtgccgg ccaccatggc gccaccacca ccaccctgt acgagccggc	2460
cgccatggac gtccaggaga gcttgcaaga cgagggtgcc accatggcgg gctacagcgc	2520
agccgccatg gctgacggca tccgggccac catggacgag gagaagcgtg ccgccatggc	2580
agagaagaac gagggcacca tggggcctga taagctgata agccgccatg gggcccagga	2640
cgagaacaag aggaccaca acgacatcat ccacaacgag agccaccatg gagggcgga	2700
caagtacaag acgctgcggc agatccggca gggcaacacc agccgccatg gcgacgagtt	2760
cgaggccctg caacagccag gccaccatgg agggcagagg ggtgctcata gcgggcgctg	2820
ccgccatggc cagcttgtg tctgccacca tggaagtctc ggaactcgcc gccatggcag	2880
ttcctttcga agccaccatg gcaacagaaa cattcgccgc catggaccac ctgataactg	2940
ataagccacc atggttgcaa tcgtgccaag caggcctgat tctcgcgatt actcgcaat	3000

NaDoc 002-PCT-1_seq listing_ST25.txt

caccgccgcc atggtgctgg gagcaggact cattgaatta cgaaaaacgc ctgtcaagtc	3060
tcaggccacc atggggaact ggcctgtgtc atacaagagt caggccgcca tggggaaacg	3120
tggcaggact tccatctgtg ccgccaccat ggtgtattcg aaacgagccg ccatggattt	3180
tctcatctct gccaccatgg catctttgta cattgccgcc atgggagggg tcaaaattgc	3240
caccatggtg gctgataagt tgatagtaac cgccatggtg tttcatccag tcgccaccat	3300
gggctggcag agagcagccg ccatggcagc gtcagtgggt gccaccatgg cttggatttt	3360
tttttttgtt tttttttttt ttgtcaaca attttacaac acattgtgtc gacgagctca	3420
agcttcccg ggcgtctacg gcaagctgac cctgaagttc atccaaaact acacaaatca	3480
gcgatttcaa caaaactaca caaatcagcg atttcaacaa caaaactaca caaatcagcg	3540
atttcaacaa aatcgcttac cgattcagaa tcgcccgggg atctgtccac tgctgttgct	3600
gttttgggca tccatcagga gaaggctcac ggcaacaaag tgctcggtgc ctttactacg	3660
gggggggggg gggggggggg ggggcccgaag ttgtcagccc agaaccacac acgagttttg	3720
ccactgggaa gctgtgatcc agtgcaggct gggacagccg acctccagcg cgcggtcacc	3780
ggtaccatgg gagccgacga tgtggtcgat tcttcgaaat cttttgtcat ggaaaacttt	3840
tcttcgtacc acgggacgaa acctgggttat gtggattcca ttcaaaaagg catacaaaag	3900
ccaaaatctg gtacacaagg aaactacgac gacgattgga aagggtttta cagtaccgac	3960
aacaaatacg acgctgcggg atactctgtg gacaacgaaa acccgctctc tggaaaagct	4020
ggaggcgtgg tcaaagtcac gtatccaggt gagtctctag ccctgccttt gcctgtcctc	4080
tcagcacttc cattagccag ctacctactt ccatccactc ccaaacttca gggctctgcc	4140
tgcccccaga ggcacaggac ttagttctgg gaccagggat caggccgcag ccctggcctg	4200
ctgttgcttc tgcagggac ttgcctttga cccagcctc tctgaccctc agggctcctt	4260
tggggagctc ttctgaattt gggctggcag ataccaccac cagaccaggt ctgccggtgc	4320
ggcagggcca gtggggcagg ttggctgtgg ctgctgtgcc ctagtctgcc ctttctgact	4380
tgcagggctc acgaaggttc tcgcactcaa ggtggacaat gccgaaacta tcaagaaaga	4440
gttgggtctc agcctcaccg aaccgctcat ggagcaagtc ggaacggaag agtttatcaa	4500
aagattcggg gatggtgctt cgcgtgtagt gctcagcctt cccttcgctg aggggagttc	4560
tagcgttgag tacatcaaca actggaaca ggcgaaagcg ttaagcgtag aacttgagat	4620
taactttgaa acccgaggaa aacgtggcca agatgcatg tatgagtata tggctcaagc	4680
ctgtgcagga aatcggtgca ggcgatagtg agcggccgca ctcgaggttt aaacggccgg	4740
ccgcggtcat agctgtttcc tgaacagatc ccgggtggca tccctgtgac ccctccccag	4800
tgctctcctt ggccctggaa gttgccactc cagtgccac cagccttgtc ctaataaaat	4860
taagtgtcat cattttgtct gactaggtgt ctttctataa tattatgggg tggagggggg	4920
tggtatggag caaggggcaa gttgggaaga caacctgtag ggcctgcggg gtctattggg	4980
aaccaagctg gagtgcagtg gcacaatctt ggctcactgc aatctccgcc tcctgggttc	5040

NaDoc 002-PCT-1_seq listing_ST25.txt

aagcgattct cctgcctcag cctccccgagt tggtgggatt ccaggcatgc atgaccaggc	5100
tcagctaatt tttgtttttt tggtagagac ggggtttcac catattggcc acgctggtct	5160
ccaactccta atctcaggtg atctaccac cttggcctcc caaattgctg ggattacagg	5220
cgtgaaccac tgctcccttc cctgtccttc tgattttaaa ataactatac cagcaggagg	5280
acgtccagac acagcatagg ctacctggcc atgcccaacc ggtgggacat ttgagttgct	5340
tgcttggcac tgtcctctca tgcgttggt ccactcagta gatgcctgtt gaattgggta	5400
cgcgccagc ttggctgtgg aatgtgtgtc agttagggtg tggaaagtc ccaggctccc	5460
cagcaggcag aagtatgcaa agcatgcac tcaattagtc agcaaccagg tgtggaaagt	5520
ccccaggctc ccagcaggc agaagtatgc aaagcatgca tctcaattag tcagcaacca	5580
tagtccccgcc cctaactccg cccatccccg ccctaactcc gccagttcc gccattctc	5640
cgccccatgg ctgactaatt ttttttattt atgcagaggc cgaggccgcc tcggcctctg	5700
agctattcca gaagtagtga ggaggctttt ttggaggcct aggcttttgc aaaaagctcc	5760
cgggagcttg tatatccatt ttcggatctg atcaagagac acgtacgacc atggagagcg	5820
acgagagcgg cctgcccgcc atggagatcg agtgccgcac caccggcacc ctgaacggcg	5880
tggagttcga gctggtgggc ggcggagagg gcacccccga gcagggccgc atgaccaaca	5940
agatgaagag caccaaaggc gccctgacct tcagccccta cctgctgagc cacgtgatgg	6000
gctacggctt ctaccacttc ggcacctacc ccagcggcta cgagaacccc ttcctgcacg	6060
ccatcaacaa cggcggctac accaacaccc gcacgcagaa gtacgaggac ggcggcgtgc	6120
tgcacgtgag cttcagctac cgctacgagg ccggccgcgt gatcggcgac ttcaagggtga	6180
tgggcaccgg cttccccgag gacagcgtga tcttcaccga caagatcatc cgcagcaacg	6240
ccaccgtgga gcacctgcac cccatgggag ataacgatct ggatggcagc ttcacccgca	6300
ccttcagcct gcgcgacggc ggctactaca gctccgtggt ggacagccac atgcacttca	6360
agagcgccat ccacccagc atcctacaga acggggggccc catgttcgcc ttccgccgcg	6420
tggaggagga tcacagcaac accgagctgg gcatcgtgga gtaccagcac gccttcaaga	6480
ccccggatgc agatgccggt gaagaataac tgcagcggga ctctgggggt cgaaatgacc	6540
gaccaagcga cgcccaacct gccatcacga gatttcgatt ccaccgccgc cttctatgaa	6600
aggttgggct tcggaatcgt tttccgggac gccggctgga tgatcctcca gcgcggggat	6660
ctcatgctgg agttcttcgc ccaccccaac ttgtttattg cagcttataa tggttacaaa	6720
taaagcaata gcatcacaaa tttcacaaat aaagcatitt tttcactgca ttctagttgt	6780
ggtttgtcca aactcatcaa tgtatcttat catgtctgta taccgtcgac ctctagctag	6840
agcttggcgt aatcatggtc atagctgttt cctgtgtgaa attgttatcc gtcacaatt	6900
ccacacaaca tacgagccgg aagcataaag tgtaaagcct ggggtgccta atgagtgagc	6960
taactcacat taattgcgtt gcgctcactg cccgctttcc agtcgggaaa cctgtcgtgc	7020
cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat tgggcgctct	7080

NaDoc 002-PCT-1_seq listing_ST25.txt

tccgcttcct cgctcactga ctcgctgctc tcggctcgttc ggctgcggcg agcggtatca	7140
gctcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc aggaagaac	7200
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt	7260
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagagggtg	7320
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgctc	7380
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cttttctccc ttcgggaagc	7440
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggc cgttcgctcc	7500
aagctgggct gtgtgcacga acccccgtt cagcccagcc gctgcgcctt atccggtaac	7560
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt	7620
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct	7680
aactacggct acactagaag aacagtatct ggtatctgct ctctgctgaa gccagttacc	7740
ttcggaataa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt	7800
ttttttgttt gcaagcagca gattacgcgc agaaaaaag gatctcaaga agatcctttg	7860
atcttttcta cggggtctga cgctcagtg aacgaaaact cacgttaagg gattttggtc	7920
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa	7980
tcaatctaaa gtatatatga gtaaacctgg tctgacagtt accaatgctt aatcagtgag	8040
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgctctg	8100
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga	8160
gaccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag	8220
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa	8280
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc	8340
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca	8400
aggcgagtta catgatcccc catgttggtc aaaaaagcgg ttagctcctt cggtcctccg	8460
atcgttggtc gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat	8520
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc	8580
aagtcattct gagaatagtg tatgcccga ccgagttgct cttgcccggc gtcaatacgg	8640
gataataacc cgccacatag cagaacttta aaagtgtcga tcattggaaa acgttcttcg	8700
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt	8760
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca	8820
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata	8880
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac	8940
atatttgaat gtatttagaa aaataaaca ataggggttc cgcgcacatt tccccgaaaa	9000
gtgccacctg acgcgccctg tagcggcgca ttaagcgcgg cgggtgtggt ggttacgcgc	9060
agcgtgaccg ctacacttgc cagcgcctta gcgcccgtc ctttcgcttt cttcccttc	9120

NaDoc 002-PCT-1_seq listing_ST25.txt

```

tttctcgcca cgttcgccgg ctttccccgt caagctctaa atcgggggct ccctttaggg 9180
ttccgattta gtgcttttacg gcacctcgac cccaaaaaac ttgattaggg tgatggttca 9240
cgtagtgggc catcgccctg atagacggtt tttcgccctt tgacgttgga gtccacgttc 9300
tttaatagtg gactcttggt ccaaactgga acaacactca accctatctc ggtctattct 9360
tttgatttat aagggtattt gccgatttcg gcctattggt taaaaaatga gctgatttaa 9420
caaaaattta acgcaattt t 9441

```

<210> 7
 <211> 9635
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

```

<400> 7
agatctgctg agcaccatgg cctgaaataa cctctgaaag aggaacttgg ttaggtacct 60
tctgaggcgg aaagaaccag ctgtggaatg tgtgtcagtt aggggtgtgga aagtccccag 120
gctccccagc aggcagaagt atgcaaagca tgcattctca ttagtcagca accaggtgtg 180
gaaagtcccc aggtcccca gcaggcagaa gtatgcaaag catgcatctc aattagtcag 240
caaccatagt cccgccccta actccgcccc tcccgcccct aactccgccc agttccgccc 300
attctccgcc ccatggctga ctaatttttt ttatttatgc agaggccgag gccgcctcgg 360
cctctgagct attccagaag tagtgaggag gctttttttg aggcctaggc ttttgcaaaa 420
agcttgattc ttctgacaca acagtctcga acttaagctg cagaagttgg tcgtgaggca 480
ctgggcaggt aagtatcaag gttacaagac aggtttaagg agaccaatag aaactgggct 540
tgtcgagaca gagaagactc ttgcgtttct gataggcacc tatttgtctt actgacatcc 600
actttgcctt tctctccaca ggtgtccact cccagttcaa ttacagctct taaggctaga 660
gtacttaata cgactcacta taggctagac gcggtacct gctaggtagc aattgaccgg 720
tcaagatggc ggccaacaac aacaacaaca acaacaaca caacaacaac aacaacaaca 780
agaagatggc ggcaacaaca acaacaaca caacaacaac aacaacaaca acaaccaaca 840
acaagatggc ggccaacaac aacaacaaca acaacaaca gaagatggcg gcaacaaca 900
caacaacaac aacaacaacc aagatggcgg ccaacaaca caagaagatg gcggcaaca 960
caacaacca gatggcggcc aacaacaaca agaagatggc ggcaacaaca acaaccaaga 1020
tggcggcacg cgtcgggtcc gctagccgta cgctccttag cgacgaaatc tactgcccc 1080
ctgagagcca ccatggcttg gggtcctacg ctgtgcaggc caagtttgga gattacaaca 1140
aagaaggccg ccatggtggt cacctcagct ctgagcggct catccgccac catgggttgg 1200
accagcaca acttaccagg gaccgccgcc atggccggac ccaggcgtgc caccatggac 1260
accgtgggtt gcgccccat ggtgctctgt tggagtgcca ccatggtgct caggacctgg 1320
gccgccatgg aatactgat aactgataag ccaccatggg aacagacctt tggcttggag 1380

```

NaDoc 002-PCT-1_seq listing_ST25.txt

ttgacgccct tggactcaac atttacgagg ccgccatgga gttcacccca aagattggct	1440
ttccttggag tgaaatcagg aacatctctg ccaccatgga aaagtttgtc atcaagccca	1500
tcgacaaggc cgccatggac tttgtgtttt acgccccacg tctcacagcc accatgggga	1560
ccctgcagct cgccgcatg gaccacgagt tgtacgccac catggggaag cctgacaccg	1620
ccgccatgga gcagacgaag gccgccacca tggaggctga taagctgata agccgcatg	1680
ggctggaaac agagaagaaa aggagagaaa ccgtggagag agagaaagag cgccaccatg	1740
gcgagaagga ggagttgttg ctgcggctgc aggactacga ggagaagaca agccgcatg	1800
ggagagacct ctcgagcag attcagaggg gccaccatgg ggaggaggag aggaagcggg	1860
cacaggaggg ccgccatggc ccagaggctg accgccacca tggactgcgg gctaagggcc	1920
gccatgggag acaggcgggtg ggccaccatg ggagccagga gcagcgccgc catgggctac	1980
ctgataactg ataagccacc atggtggaag aggcgcggag gcgcaaggag gacgaagttg	2040
aagagtggca gcaagccgcc atggaagccc aggacgacct ggtcaagacc aaggaggagc	2100
tgcacctggt gccggccacc atggcgccac caccaccacc cgtgtacgag ccggccgcca	2160
tggacgtcca ggagagcttg caagacgagg gtgccaccat ggcgggctac agcgagccg	2220
ccatggctga cggcatccgg gccaccatgg acgaggagaa gcgtgccgcc atggcagaga	2280
agaacgaggc caccatgggg cctgataagc tgataagccg ccatggggcc cgagacgaga	2340
acaagaggac ccacaacgac atcatccaca acgagagcca ccatggaggc cgggacaagt	2400
acaagacgct gcggcagatc cggcagggca acaccagccg ccatggcgac gaggctcagg	2460
ccctgcaaca gccaggccac catggagggc agaggggtgc tcatagcggg cgctgccgcc	2520
atggccacgc ttgtgtctgc caccatggaa gtctcggaac tcgccgcat ggcagttcct	2580
ttcgaagcca ccatggcaac agaaacattc gccgcatgg accacctgat aactgataag	2640
ccaccatggt tgcaatcgtg ccaagcaggc ctgattctcg cgattactcg cgaatcaccg	2700
ccgccatggt gctgggagca ggactcattg aattacggaa aacgcctgtc aagtctcagg	2760
ccaccatggg gaactggcct gtgtcataca agagtcaggc cgccatgggg aaacgtggca	2820
ggacttccat ctgtgccgcc accatggtgt attcgaaacg agccgcatg gatcttctca	2880
tctctgccac catggcatct ttgtacattg ccgccatggg aggggtcaaa attgccacca	2940
tgggtggctga taagttgata gtaaccgcca tgggtgttca tccagtcgcc accatgggct	3000
ggcagagagc agccgcatg gcagcgtcag tgggtggccac catggcttgg atttttttt	3060
ttgttttttt tttttttgct caacaatttt acaacacatt gtgtcgacga gctcgtgcgc	3120
acctacggca agctgaccct gaagttcatc caacaaaact acacaaatca gcgatttcca	3180
caacaactac ggcaagctga cctgaagtt catccaacaa aactacacaa atcagcgatt	3240
tccacaacaa ctacggcaag ctgacctga agttcatcca acaaaactac acaaatcagc	3300
gatttccagc aaggcaacca aaggctcttt ttagagccac ctttcaacgc gcaaggcaac	3360
aaaaggccct tttcagggcc acctttcaag agggcgcaag gcaaccaaag gctcttttca	3420

NaDoc 002-PCT-1_seq listing_ST25.txt

gagccacctt tcaaggcgca aggcaaccaa aggctctttt cagagcccc tttattggac	3480
aaactaccta cagagattta aagctctaag gtaaataataa aatttttaag tgtataatgt	3540
gttaaactac tgattctaata tgtttgtgta ttttagattc caacctatgg aactgatcaa	3600
tcggagcagt ggtggaatcc ctttaaacat ttgctgtctga cacaactgtg ttcactagca	3660
acctcaaaca gacaccacgg tgcattctgac tcctgaggag aagtctgccg ttactgccct	3720
gtggggcaag gtgaacgtgg acgaagtgg tgctgaggcc ctgggcagggt tggatatcaag	3780
gttacaagac aggtttaagg agaccaatag aaactgggca tgtggagaca gagaagactc	3840
ttgggtttct gataggcact gactctctct gcctattgggt ctattttccc acccttaggc	3900
tgctggtggt ctacccttgg acccagagggt tctttgagtc ctttggggat ctgtccactg	3960
ctgaagctgt tacgggcaac cctaagctga aggctcctgg caagaaagtg ctcggtgcct	4020
ttagtgatcg cctggctcac ctggacaacc tcaagggcac ctttgccacg ctgagtgagc	4080
tgcactgtga caagctgcac gtgtatcctg agaacttcag gctcctgggc aacgtgctgg	4140
tctgtgtgct ggcccatcac cttggcaaag aattcacccc accagtgcag gctgcctatc	4200
agaaagtgggt ggctggtgtg gctaacgccc tggcccaaa gtatcactaa gctcgctttc	4260
ttgctgtccg atttctatta gaggttcctt tgttccctaa gtccaactac gaaactgggg	4320
gatattctga agggccttga gcatctggat tctgcctggc gcgccggtca ccccgatcc	4380
gtgatagtaa cctgatagta acctgataat agcagatctc gccgccatgg gagctgatga	4440
tgtggttgat tcttcgaaat cttttgtcat ggaaaacttt tcttcgtacc acgggacgaa	4500
acctggttat gtggattcca ttcaaaaagg catacaaaag ccaaaatctg gtacacaagg	4560
aaactatgac gatgattgga aagggtttta tagtaccgac aacaaatatg acgtgcggg	4620
atactctgtg gataatgaaa acccgctctc tggaaaagct ggaggcgtgg tcaaagtac	4680
gtatccagga ctgacgaagg ttctcgcact aaagggtggat aatgccgaaa ctattaagaa	4740
agagttaggt ttaagtctca ctgaaccgct catggagcaa gtcggaacgg aagagtttat	4800
caaaagattc ggtgatggtg cttcgcgtgt agtgctcagc cttcccttcg ctgaggggag	4860
ttctagcgtt gagtacatca acaactggga acaggcgaaa gcgttaagcg tagaacttga	4920
gattaacttt gaaaccctg gaaaacgtgg ccaagatgcg atgtatgagt atatggctca	4980
agcctgtgca ggaaatcgtg tcaggcgata gtgaactagt atccggaatc tagagcggcc	5040
gctggccgca ataaaatatac tttattttca ttacatctgt gtgttggttt tttgtgtgag	5100
gatctaaatg agtcttcgga cctcgcgggg gccgcttaag cggtgggttag ggtttgtctg	5160
acgcgggggg agggggaagg aacgaaacac tctatttcgg aggcggctcg gggtttggtc	5220
ttggtggcca cgggcacgca gaagagcgcc gcgacacct taagcacccc cccgccctcc	5280
gtggaggcgg gggtttggtc ggcgggtggt aactggcggg ccgctgactc gggcgggtcg	5340
cgcgccccag agtgtgacct tttcggctctg ctcgcagacc cccgggcggc gccgccgcgg	5400
cggcgacggg ctcgctgggt cctaggctcc atggggaccg tatacgtgga caggctctgg	5460

NaDoc 002-PCT-1_seq listing_ST25.txt

agcatccgca cgactgcggt gatattaccg gagaccttct gcgggacgag ccgggtcacg	5520
cggtgacgc ggagcgtccg ttgggcgaca aacaccagga cggggcacag gtacactatc	5580
ttgtcacccg gaggcgcgag ggactgcagg agcttcaggg agtggcgag ctgcttcac	5640
cccgtggccc gttgctcgcg tttgctggcg gtgtccccgg aagaaatata tttgcatgtc	5700
tttagttcta tgatgacaca aacccgccc agcgtcttgt cattggcgaa ttcgaacacg	5760
cagatgcagt cggggcggcg cggccccagg tccacttcgc atattaagggt gacgcgtgtg	5820
gcctcgaaca ccgagcgacc ctgcagcgac ccgcttaaaa gcttggcatt ccggtactgt	5880
tggtaaagcc accatggccg atgctaagaa cattaagaag ggccctgctc cttctaccc	5940
tctggaggat ggcaccgctg gcgagcagct gcacaaggcc atgaagagggt atgccctggt	6000
gcctggcacc attgccttca ccgatgccca cattgagggtg gacatcacct atgccgagta	6060
cttcgagatg tctgtgcgcc tggccgaggc catgaagagg tacggcctga acaccaacca	6120
ccgcatcgtg gtgtgctctg agaactctct gcagttcttc atgccagtgc tgggcgcct	6180
gttcacgga gtggccgtgg cccctgctaa cgacatttac aacgagcgcg agctgctgaa	6240
cagcatgggc atttctcagc ctaccgtggt gtctgtgtct aagaagggcc tgcagaagat	6300
cctgaacgtg cagaagaagc tgcctatcat ccagaagatc atcatcatgg actctaagac	6360
cgactaccag ggcttcaga gcatgtacac attcgtgaca tctcatctgc ctctggctt	6420
caacgagtac gacttcgtgc cagagtcttt cgacaggac aaaaccattg ccctgatcat	6480
gaacagctct gggcttaccg gcctgcctaa gggcgtggcc ctgcctcatc gcaccgcctg	6540
tgtgcgcttc tctcagccc gcgaccctat tttcggaac cagatcatcc ccgacaccgc	6600
tattctgagc gtggtgccat tccaccacgg cttcggcatg ttcaccaccc tgggctacct	6660
gatttgccg tttcgggtgg tgctgatgta ccgcttcgag gaggagctgt tcctgcgcag	6720
cctgcaagac taaaaattc agtctgccct gctggtgcca accctgttca gcttcttcgc	6780
taagagcacc ctgatcgaca agtacgacct gtctaacctg cacgagattg cctctggcgg	6840
cgccccactg tctaaggagg tgggcgaagc cgtggccaag cgctttcatc tgccaggcat	6900
ccgccagggc tacggcctga ccgagacaac cagcgccatt ctgattaccc cagagggcga	6960
cgacaagcct ggcgcctggt gcaagggtgt gccattcttc gaggccaagg tgggtggacct	7020
ggacaccggc aagaccctgg gagtgaacca gcgcggcgag ctgtgtgtgc gcggccctat	7080
gattatgtcc ggctacgtga ataaccctga ggccacaaac gccctgatcg acaaggacgg	7140
ctggctgcac tctggcgaca ttgcctactg ggacgaggac gagcacttct tcatcgtgga	7200
ccgcctgaag tctctgatca agtacaaggg ctaccagggtg gcccagccg agctggagtc	7260
tatcctgctg cagcacccta acattttcga cgccggagtg gccggcctgc ccgacgacga	7320
tgccggcgag ctgcctgccg ccgtcgtcgt gctggaacac ggcaagacca tgaccgagaa	7380
ggagatcgtg gactatgtgg ccagccagggt gacaaccgcc aagaagctgc gcggcgaggt	7440
ggtgttcgtg gacgagggtgc ccaagggcct gaccggcaag ctggacgccc gcaagatccg	7500

NaDoc 002-PCT-1_seq listing_ST25.txt

cgagatcctg atcaaggcta agaaaggcgg caagatcgcc gtgtaataat tctagagtcg	7560
gggcgccgg cgccttcgag cagacatgat aagatacatt gatgagtttg gacaaaccac	7620
aactagaatg cagtgaataa aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt	7680
tgtaccatt ataagctgca ataaacaagt taacaacaac aattgcattc attttatgtt	7740
tcaggttcag ggggaggtgt gggaggtttt ttaaagcaag taaaacctct acaaagtgtg	7800
taaaatcgat aaggatccag gtggcacttt tcggggaaat gtgcgcggaa cccctatttg	7860
tttatttttc taaatacatt caaatatgta tccgctcatg agacaataac cctgataaat	7920
gcttcaataa tattgaaaaa ggaagagtat gagtattcaa catttccgtg tcgcccttat	7980
tccctttttt gcggcatttt gccttcctgt ttttgctcac ccagaaacgc tgggtgaaagt	8040
aaaagatgct gaagatcagt tgggtgcacg agtgggttac atcgaactgg atctcaacag	8100
cggtaaagatc cttgagagtt ttcgccccga agaacgtttt ccaatgatga gcacttttaa	8160
agttctgcta tgtggcgcgg tattatcccg tattgacgcc gggcaagagc aactcggctg	8220
ccgcatacac tattctcaga atgacttggg tgagtactca ccagtcacag aaaagcatct	8280
tacggatggc atgacagtaa gagaattatg cagtgtgccc ataaccatga gtgataacac	8340
tgcggccaac ttacttctga caacgatcgg aggaccgaag gagctaaccg cttttttgca	8400
caacatgggg gatcatgtaa ctgccttga tcgttgggaa ccggagctga atgaagccat	8460
accaaacgac gagcgtgaca ccacgatgcc thtagcaatg gcaacaacgt tgcgcaaact	8520
attaactggc gaactactta ctctagcttc ccggcaacaa ttaatagact ggatggaggc	8580
ggataaagtt gcaggaccac ttctgcgctc ggcccttcg gctggctggg ttattgctga	8640
taaatctgga gccggtgagc gtgggtctcg cggatcatt gcagcactgg ggccagatgg	8700
taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta tggatgaacg	8760
aaatagacag atcgtgaga taggtgcctc actgattaag cattggtaac tgtcagacca	8820
agtttactca tatatacttt agattgattt aaaacttcat ttttaattta aaaggatcta	8880
ggtgaagatc ctttttgata atctcatgac caaaatccct taacgtgagt tttcgttcca	8940
ctgagcgtca gaccccgtag aaaagatcaa aggatcttct tgagatcctt ttttctgcg	9000
cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca gcggtgggtt gtttgccgga	9060
tcaagagcta ccaactcttt ttccgaaggt aactggcttc agcagagcgc agatacaaaa	9120
tactgttctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc	9180
tacatacctc gctctgctaa tcctgttacc agtggctgct gccagtggcg ataagtcgtg	9240
tcttaccggg ttggactcaa gacgatagtt accggataag gcgcagcggg cgggctgaac	9300
ggggggttcg tgcacacagc ccagcttggg gcgaacgacc tacaccgaac tgagatacct	9360
acagcgtgag ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg acaggatatcc	9420
ggtaagcggc agggctcgaa caggagagcg cacgagggag cttccagggg gaaacgcctg	9480
gtatctttat agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg	9540

NaDoc 002-PCT-1_seq listing_ST25.txt

ctcgtcaggg gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacgggttcct 9600
ggccttttgc tggccttttg ctcacatggc tcgac 9635

<210> 8
<211> 8866
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 8
aacaaaatat taacgcttac aatttcatt cgccattcag gctgcgcaac tgttggaag 60
ggcgatcggg gggggcctct tcgctattac gccagctggc gaaaggggga tgtgctgcaa 120
ggcgattaag ttgggtaacg ccagggtttt cccagtcacg acgttgtaaa acgacggcca 180
gtgccaagct gatctataca ttgaatcaat attggcaatt agccatatta gtcattgggt 240
atatagcata aatcaatatt ggctattggc cattgcatac gttgtatcta taccataata 300
tgtacattta tattgggtca tgtccaatat gaccgccatg ttgacattga ttattgacta 360
gttattaata gtaatcaatt acgggggtcat tagttcatag cccatatatg gagttccgcg 420
ttacataact tacggtaaat ggcccgctg gctgaccgcc caacgacccc cgccattga 480
cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat tgacgtcaat 540
gggtggagta ttacggtaa actgcccact tggcagtaca tcaagtgtat catatgccaa 600
gtccgcccc tattgacgtc aatgacggta aatggccgc ctggcattat gccagtaga 660
tgaccttacg ggactttcct acttggcagt acatctacgt attagtcac gctattacca 720
tgggtgatgcg gttttggcag tacaccaatg ggcgtggata gcggtttgac tcacggggat 780
ttccaagtct ccacccatt gacgtcaatg ggagtttgtt ttggcaccaa aatcaacggg 840
actttccaaa atgtcgtaat aacccgccc cggtgacgca aatgggcggg aggcgtgtac 900
gggtgggagg ctatataagc agagctcgtt tagtgaaccg tcagaatttt gtaatacgac 960
tcactatagg gcggccggga attcgtcgac tggatccggg acctagctag gtagcaattg 1020
accggtcaag atggcggcca acaacaaca caacaacaac aacaacaaca acaacaaca 1080
caacaagaag atggcggcca caacaacaac aacaacaaca aacaacaaca caacaacaac 1140
caacaacaag atggcggcca acaacaaca caacaacaac aacaagaaga tggcggcaac 1200
aacaacaaca acaacaaca caaccaagat ggcggccaac aacaacaaga agatggcggc 1260
aacaacaaca accaagatgg cggccaaca caacaagaag atggcggcaa caacaacaac 1320
caagatggcg gcacgcgtcg gtccggctag ccgtacgctc cttagcgacg aaatctactg 1380
ccccctgag agccaccatg gcttgggggtc ctacgctgtg caggccaagt ttggagatta 1440
caacaaagaa ggccgcatg gtgggcacct cagctctgag cggctcatcc gccaccatgg 1500
gttgaccag cacaactta ccagggaccg ccgcatggc cggaccagg cgtgccacca 1560
tggacaccgt gggttgcgcc gccatggtgc tctgttggag tgccaccatg gtgctcagga 1620

NaDoc 002-PCT-1_seq listing_ST25.txt

cctgggcccgc catggaatac ctgataactg ataagccacc atgggaacag acctttggct	1680
tggagttgac gcccttggac tcaacattta cgaggccgcc atggagttca ccccaaagat	1740
tggctttcct tggagtgaat tcaggaacat ctctgccacc atggaaaagt ttgtcatcaa	1800
gccccatcgac aaggccgcca tggactttgt gttttacgcc ccacgtctca cagccaccat	1860
ggggaccctg cagctcgccg ccatggacca cgagttgtac gccaccatgg ggaagcctga	1920
caccgcccgc atggagcaga cgaaggccgc caccatggag gctgataagc tgataagccg	1980
ccatgggctg gaaacagaga agaaaaggag agaaaccgtg gagagagaga aagagcgcca	2040
ccatggcgag aaggaggagt tgttctgctg gctgcaggac tacgaggaga agacaagccg	2100
ccatgggaga gacctctcgg agcagattca gaggggccac catggggagg aggagaggaa	2160
gcgggcacag gagggccgcc atggcccaga ggctgaccgc caccatggac tgcgggctaa	2220
gggcccgcac ggagagacagg cgggtgggcca ccatgggagc caggagcagc gccgccatgg	2280
gctacctgat aactgataag ccaccatggt ggaagaggcg cggaggcgca aggaggacga	2340
agttgaagag tggcagcaag ccgccatgga agcccaggac gacctggtca agaccaagga	2400
ggagctgcac ctggtgccgg ccaccatggc gccaccacca ccacctgtgt acgagccggc	2460
cgccatggac gtccaggaga gcttgcaaga cgaggggtgcc accatggcgg gctacagcgc	2520
agccgccatg gctgacggca tccgggccac catggacgag gagaagcgtg ccgccatggc	2580
agagaagaac gaggccacca tggggcctga taagctgata agccgccatg gggcccgaga	2640
cgagaacaag aggaccaca acgacatcat ccacaacgag agccaccatg gaggccggga	2700
caagtacaag acgctgcggc agatccggca gggcaacacc agccgccatg gcgacgagtt	2760
cgaggccctg caacagccag gccaccatgg agggcagagg ggtgctcata gcgggcgctg	2820
ccgccatggc cacgcttctg tctgccacca tggaaagtct ggaactcgcc gccatggcag	2880
ttcctttcga agccaccatg gcaacagaaa cattcgccgc catggaccac ctgataactg	2940
ataagccacc atggttgcaa tcgtgccaa gaggcctgat tctcgcgatt actcgcgaat	3000
caccgcccgc atggtgctgg gagcaggact cattgaatta cggaaaacgc ctgtcaagtc	3060
tcaggccacc atggggaact ggcctgtgtc atacaagagt caggccgcca tggggaaacg	3120
tggcaggact tccatctgtg ccgccaccat ggtgtattcg aaacgagccg ccatggatgt	3180
tctcatctct gccaccatgg catctttgta cattgccgcc atgggagggg tcaaaattgc	3240
caccatggtg gctgataagt tgatagtaac cgccatggtg tttcatccag tcgccaccat	3300
gggctggcag agagcagccg ccatggcagc gtcagtgggt gccaccatgg cttggatgtt	3360
tttttttgtt tttttttttt ttgctcaaca attttacaac acattgtgtc gacgagctca	3420
agcttcccgg cgcgccccgg tccgtccgga ctacggcaag ctgaccctga agttcatccc	3480
aaaacttacg ctgagtactt cgatctggtc accccggatc tcgccgccat gggagctgat	3540
gatgtgggtg attcttcgaa atcttttctc atggaaaact tttcttcgta ccacgggacg	3600
aaacctggtt atgtggattc cattcaaaaa ggcatacaaa agccaaaatc tgggtacacaa	3660

NaDoc 002-PCT-1_seq listing_ST25.txt

ggaaactatg acgatgattg gaaagggttt tatagtaccg acaacaaata tgacgctgcg	3720
ggatactctg tggataatga aaacccgctc tctggaaaag ctggaggcgt ggtcaaagtg	3780
acgtatccag gactgacgaa ggttctcgca ctaaagggtg ataatgccga aactattaag	3840
aaagagttag gtttaagtct cactgaaccg ctcatggagc aagtcggaac ggaagagttt	3900
atcaaaagat tcggtgatgg tgcttcgctg gtagtgctca gccttccctt cgctgagggg	3960
agttctagcg ttgagtacat caacaactgg gaacaggcga aagcgtaag cgtagaactt	4020
gagattaact ttgaaacccg tggaaaacgt ggccaagatg cgatgtatga gtatatggct	4080
caagcctgtg caggaaatcg tgcaggcga tagtgaacta gtatccggaa tctagagcgg	4140
ccgcactcga ggtttaaacg gccggccgcg gtcatactg tttcctgaac agatcccggg	4200
tggcatccct gtgacccctc ccagtgctt ctctggccc tggaaagtgc cactccagt	4260
cccaccagcc ttgtcctaataaaaattagt tgcattctt tgtctgacta ggtgtccttc	4320
tataatatta tggggtggag gggggtggtg tggagcaagg ggcaagttgg gaagacaacc	4380
tgtagggcct gcggggtcta ttgggaacca agctggagtg cagtggcaca atcttggtc	4440
actgcaatct ccgcctcctg ggttcaagcg attctcctgc ctcagcctcc cgagttgttg	4500
ggattccagg catgcatgac caggctcagc taatttttgt ttttttggtg gagacggggt	4560
ttcaccatat tggccacgct ggtctccaac tcctaattctc aggtgatcta cccaccttg	4620
cctcccaaata tgctgggatt acaggcgtga accactgctc cttccctgt cttctgatt	4680
ttaaaataac tataccagca ggaggacgtc cagacacagc ataggctacc tggccatgcc	4740
caaccggtgg gacatttgag ttgcttgctt ggcactgtcc tctcatgctg tgggtccact	4800
cagtagatgc ctgttgatt gggtagcgg ccagcttggc tgtggaatgt gtgtcagtta	4860
gggtgtggaa agtccccagg ctccccagca ggcagaagta tgcaaagcat gcatctcaat	4920
tagtcagcaa ccaggtgtgg aaagtcccca ggctccccag caggcagaag tatgcaaagc	4980
atgcatctca attagtcagc aaccatagtc ccgcccctaa ctccgcccac cccgccccta	5040
actccgcca gttccgcca ttctccgccc catggctgac taattttttt tatttatgca	5100
gaggccgagg ccgcctcggc ctctgagcta ttccagaagt agtgaggagg cttttttgga	5160
ggcctaggct ttgcaaaaa gctcccggga gcttgatat ccattttcgg atctgatcaa	5220
gagacacgta cgaccatgga gagcgacgag agcggcctgc ccgccatgga gatcgagtgc	5280
cgcatcaccg gcaccctgaa cggcgtggag ttcgagctgg tgggcggcgg agagggcacc	5340
cccgagcagg gccgcatgac caacaagatg aagagcacca aaggcgccct gaccttcagc	5400
ccctacctgc tgagccacgt gatgggctac ggcttctacc acttcggcac ctacccagc	5460
ggctacgaga accccttcct gcacgccatc aacaacggcg gctacaccaa caccgcac	5520
gagaagtacg aggacggcgg cgtgctgcac gtgagcttca gctaccgcta cgaggccggc	5580
cgctgatcg gcgacttcaa ggtgatgggc accggcttcc ccgaggacag cgtgatcttc	5640
accgacaaga tcatccgcag caacgccacc gtggagcacc tgcaccccat gggcgataac	5700

NaDoc 002-PCT-1_seq listing_ST25.txt

gatctggatg gcagcttcac ccgcaccttc agcctgcgcg acggcggcta ctacagctcc	5760
gtgggtggaca gccacatgca cttcaagagc gccatccacc ccagcatcct acagaacggg	5820
ggcccatgt tgccttccg ccgcgtggag gaggatcaca gcaacaccga gctgggcatc	5880
gtggagtacc agcacgcctt caagaccccg gatgcagatg ccggtgaaga ataactgcag	5940
cgggactctg gggttcgaaa tgaccgacca agcgacgccc aacctgccat cacgagattt	6000
cgattccacc gccgccttct atgaaagggt gggcttcgga atcgttttcc gggacgccgg	6060
ctggatgac ctccagcgcg gggatctcat gctggagtcc ttcgcccacc ccaacttggt	6120
tattgcagct tataatgggt acaataaag caatagcatc acaaatttca caaataaagc	6180
atttttttca ctgcattcta gttgtggttt gtccaaactc atcaatgtat cttatcatgt	6240
ctgtataccg tcgacctcta gctagagctt ggcgtaatca tggcatagc tgtttcctgt	6300
gtgaaattgt tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa	6360
agcctgggggt gcctaatagag tgagctaact cacattaatt gcgttgcgct cactgcccgc	6420
ttccagtcg ggaaacctgt cgtgccagct gcattaatga atcgccaac gcgcggggag	6480
aggcggtttg cgtattgggc gctcttccgc ttctcgcgc actgactcgc tgcgctcggt	6540
cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacgggt tatccacaga	6600
atcaggggat aacgcaggaa agaactatgt agcaaaaggc cagcaaaagg ccaggaaccg	6660
taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcacia	6720
aatcgacgc tcaagtcaga ggtggcgaaa ccgcagagga ctataaagat accaggcggt	6780
tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct	6840
gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct	6900
cagttcgggt taggtcggtc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc	6960
cgaccgctgc gccttatccg gtaactatcg tcttgagtc aacccggtaa gacacgactt	7020
atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcgggtgc	7080
tacagagttc ttgaagtggg ggcctaacta cggctacact agaagaacag tatttggtat	7140
ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa	7200
acaaaccacc gctggtagcg gtgggttttt tgtttgcaag cagcagatta cgcgagaaa	7260
aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga	7320
aaactcacgt taagggtatt tggtcatgag attatcaaaa aggatcttca cctagatcct	7380
tttaaatata aatgaagtt ttaaatcaat cttaaagtata tatgagtaaa cttggtctga	7440
cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcac	7500
catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggct taccatctgg	7560
ccccagtgc gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat	7620
aaaccagcca gccggaaggg ccgagcgagc aagtggctct gcaactttat ccgcctccat	7680
ccagtctatt aattgttgcc gggaagctag agtaagtagt tcgccagtta atagtttgcg	7740

NaDoc 002-PCT-1_seq listing_ST25.txt

```

caacgttggt gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc 7800
attcagctcc ggttcccaac gatcaaggcg agttacatga tcccccatgt tgtgcaaaaa 7860
agcgggtagc tccttcggtc ctccgatcgt tgtcagaagt aagttggccg cagtgttatc 7920
actcatgggt atggcagcac tgcataattc tcttactgtc atgccatccg taagatgctt 7980
ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag 8040
ttgctcttgc cgggcgtcaa tacgggataa taccgcgcca catagcagaa ctttaaaagt 8100
gctcatcatt ggaaaacgtt cttcgggcg aaaactctca aggatcttac cgctgttgag 8160
atccagttcg atgtaacca ctcgtgcacc caactgatct tcagcatctt ttactttcac 8220
cagcgtttct gggtagacaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc 8280
gacacggaaa tgttgaatac tcatactctt cttttttcaa tattattgaa gcatttatca 8340
gggttattgt ctcattgagc gatacatatt tgaatgtatt tagaaaaata aacaaatagg 8400
ggttcgcgc acatttcccc gaaaagtgcc acctgacgcg ccctgtagcg gcgcattaag 8460
cgcggcgggt gtggtggtta cgcgacgcgt gaccgctaca cttgccagcg ccctagcgcc 8520
cgctcctttc gctttcttcc cttcctttct cgccacgttc gccggctttc cccgtcaagc 8580
tctaaatcgg gggctccctt taggggtccg atttagtgct ttacggcacc tcgaccccaa 8640
aaaacttgat taggggtgat gttcacgtag tgggccatcg ccctgataga cggtttttcg 8700
ccctttgacg ttggagtcca cgttctttaa tagtggaactc ttgttccaaa ctggaacaac 8760
actcaaccct atctcggtct attcttttga tttataaggg attttgccga tttcggccta 8820
ttggttaaaa aatgagctga ttttaaaaa atttaacgcg aatttt 8866

```

<210> 9
 <211> 9665
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

```

<400> 9
agatctgcgc agcaccatgg cctgaaataa cctctgaaag aggaacttgg ttaggtacct 60
accggaagga acccgcgcta tgacggcaat aaaaagacag aataaaacgc acggtgttct 120
tataatggtt acaataaaag caatagcatc acaaatttca caaataaagc atttttttca 180
ctgcattcta gttgtggtta taaaatatct ttattttcat tacatctgtg tgttggtttt 240
ttgtgtgtgg cttcccaaag tgctgggatt acaggcatga gccatcgagc ccaaccaat 300
tttttttttt ttttaatttta ctttctgcaa tcattcatcc attcagccag tgcggtattt 360
ctgaggtgtg ttcgatcgcg gatccatgcc tgccgcagta cagttgtgag ccaaatagaga 420
ctgagactag ttcccgccct ccaagagctt gcaagacccg cagtggcgta aaaactacta 480
catcttttag tgatcgattc tgcactccag gggttttcaa tctactacaa gagtgaataa 540
gagttcgcct ttgtctgata tctgtgtgta ttctctctcg cttctttaac tgattttttc 600

```

NaDoc 002-PCT-1_seq listing_ST25.txt

tcagctaata aaacatccac ccacaacccc ccgaacgccc gcaaacacca ggccactcta	660
gcaaaacctc tctcactccg cctgcgcaat ccagctgact tccggttaca gataaccacg	720
tgattgggaa cccttgctgc gcatgtctag taggaagtcg gactatacca ctttccttac	780
ggaaggggta cttttttatg tttttaagtt taaaaccgat ttctgatatt tgacttttat	840
catttcaggc ctatatggag gctatgagtg agtttagtgt ggcagaagat gaaagaaccg	900
gacaggaata cggacgaaat tggagcaggg tttgggctct ccccttcgca gataatcgga	960
ggagccgggc ccgagcgagc tctttccttt cgctgctgcg gccgcagccg tgaggtgagg	1020
gcgagctggt ctccatcagg cgctgacgcg tgctgacaag ggactgtcgg tcttgggacc	1080
gcagctgggg ttgggggaga tgaaatggag gccgcctaa agcggccggt cccgggggtt	1140
ggggtaggcc ggagcacttt cgtcccgggc ctccggagtg agggggggcg gggagcgctg	1200
cagcaactga gaccaggaaa agtctgcccc ggctgggtgcc gcaccgcaca cgtgtccggt	1260
cgaccacgc gagcagagca aacggagcga acaagaccaa gccgtgggcc ctttcttgct	1320
tggcacaccc ggagcggagc cgatctctgc ttccacgtga tgtagggcaa gcctagtgt	1380
ggccccaggc ctccgactgc cgagagaggt gatctctaac tcttgactcc attcactcct	1440
ttggcctctc ataaaggaaa tctctgcgaa tagccgaacg aggcttggtta ctgtgataaa	1500
acagggaaat aagcccagaa aacagagtaa cttgcctgca ttcctagact agaaatcagg	1560
tctactcacc tcgaatattc tttaaagcgt gagtaccaga aatggcataa cccccctatt	1620
caatccaata agtccttggc ttgactttcc agaggagaaa tgcaacatg aggtccgag	1680
aggtgaaggc atagcgtggg ttttgaagtc ttaaacccaa gggggccagc tgcatagccc	1740
agagccttaa agatgattta gggaagagtc ttatttcgcg gctgtggtgt gggtcacaaa	1800
gggcaggctt tgatggggac gttcattctt gcccaggatt ggctttcaga gtctaatacat	1860
gttttctgtg tgtctagtat cctcaggctt cagaagaggc tcgcctctag tgcctccgc	1920
tgtggcaaga agaaggtctg gaccggtcaa gatggcggcc aacaacaaca acaacaaca	1980
caacaacaac aacaacaaca acaacaagaa gatggcggca aacaacaaca caacaacaac	2040
aacaacaaca acaacaaca ccaacaacaa gatggcggcc aacaacaaca acaacaaca	2100
caacaagaag atggcggcaa caacaacaac aacaacaaca acaaccaaga tggcggccaa	2160
caacaacaag aagatggcgg caacaacaac aaccaagatg gcggccaaca acaacaagaa	2220
gatggcggca acaacaaca ccaagatggc ggcacgcgtc ggtccggcta gccgtacgct	2280
ccttagcgac gaaatctact gccccctga gagccaccat ggcttgggggt cctacgctgt	2340
gcaggccaag tttggagatt acaacaaaga aggccgcat ggtgggcacc tcagctctga	2400
gcggctcatc cgccaccatg gggtggacca gcacaaactt accagggacc gccgcatgg	2460
ccggacccag gcgtgccacc atggacaccg tgggttgcg cgccatggtg ctctgttga	2520
gtgccaccat ggtgctcagg acctgggccg ccatggaata cctgataact gataagccac	2580
catgggaaca gaccttggc ttggagttga cgcccttga ctcaacattt acgaggccgc	2640

NaDoc 002-PCT-1_seq listing_ST25.txt

catggagttc accccaaaga ttggctttcc ttggagtga atcaggaaca tctctgccac	2700
catggaaaag tttgtcatca agcccatcga caaggccgcc atggactttg tgttttacgc	2760
cccacgtctc acagccacca tggggaccct gcagctcgcc gccatggacc acgagttgta	2820
cgccaccatg gggaagcctg acaccgccgc catggagcag acgaaggccg ccaccatgga	2880
ggctgataag ctgataagcc gccatgggct ggaaacagag aagaaaagga gagaaaccgt	2940
ggagagagag aaagagcgcc accatggcga gaaggaggag ttgttgctgc ggctgcagga	3000
ctacgaggag aagacaagcc gccatgggag agacctctcg gagcagattc agagggggcca	3060
ccatggggag gaggagagga agcgggcaca ggagggccgc catggcccag aggctgaccg	3120
ccaccatgga ctgctgggcta agggccgcc tgggagacag gcggtgggcc accatgggag	3180
ccaggagcag cgccgccatg ggctacctga taactgataa gccaccatgg tggaagaggc	3240
gcggaggcgc aaggaggacg aagttgaaga gtggcagcaa gccgccatgg aagcccagga	3300
cgacctgtc aagaccaagg aggagctgca cctggtgccg gccaccatgg cgccaccacc	3360
accacccgtg tacgagccgg ccgccatgga cgtccaggag agcttgcaag acgaggggtgc	3420
caccatggcg ggctacagcg cagccgccat ggctgacggc atccgggcca ccatggacga	3480
ggagaagcgt gccgccatgg cagagaagaa cgaggccacc atggggcctg ataagctgat	3540
aagccgccat gggggccgag acgagaacaa gaggaccac aacgacatca tccacaacga	3600
gagccaccat ggaggccggg acaagtacaa gacgctgcgg cagatccggc agggcaacac	3660
cagccgccat ggcgacgagt tcgaggccct gcaacagcca ggccaccatg gagggcagag	3720
gggtgtcat agcgggcgct gccgccatgg ccacgcttgt gtctgccacc atggaagtct	3780
cggaactcgc cgccatggca gttcctttcg aagccaccat ggcaacagaa acattcgccg	3840
ccatggacca cctgataact gataagccac catggttgca atcgtgccaa gcaggcctga	3900
ttctcgcgat tactcgcgaa tcaccgccgc catggtgctg ggagcaggac tcattgaatt	3960
acggaaaacg cctgtcaagt ctacggccac catgggggaaac tggcctgtgt catacaagag	4020
tcaggccgcc atgggggaaac gtggcaggac ttccatctgt gccgccacca tgggtgtattc	4080
gaaacgagcc gccatggatt ttctcatctc tgccaccatg gcatctttgt acattgccgc	4140
catgggaggg gtcaaaattg ccaccatggt ggctgataag ttgatagtaa ccgccatggt	4200
gtttcatcca gtcgccacca tgggctggca gagagcagcc gccatggcag cgtcagtgg	4260
ggccaccatg gcttggattt ttttttttgt tttttttttt ttgtctaac aattttacaa	4320
cacattgtgt cgacgagctc aagcttcccg gcgcgccccg gtccgtccgg tcccacgcgt	4380
caattggaaa acttacgctg agtacttcga tctccctacg gcaagctgac cctgaagttc	4440
aacagatctc gccgccatgg gagctgatga tgtggttgat tcttcgaaat cttttgtcat	4500
ggaaaacttt tcttcgtacc acgggacgaa acctggttat gtggattcca ttcaaaaagg	4560
catacaaaag ccaaaatctg gtacacaagg aaactatgac gatgattgga aagggtttta	4620
tagtaccgac aacaaatatg acgctgcggg atactctgtg gataatgaaa acccgctctc	4680

NaDoc 002-PCT-1_seq listing_ST25.txt

tggaagagct ggaggcgtgg tcaaagtgac gtatccagga ctgacgaagg ttctcgact	4740
aaaggtggat aatgccgaaa ctattaagaa agagttaggt ttaagtctca ctgaaccgct	4800
catggagcaa gtcggaacgg aagagtttat caaaagattc ggtgatggtg cttcgcgtgt	4860
agtgtcagc cttcccttcg ctgaggggag ttctagcgtt gagtacatca acaactggga	4920
acaggcgaaa gcgttaagcg tagaacttga gattaacttt gaaacccgtg gaaaacgtgg	4980
ccaagatgcg atgtatgagt atatggctca agcctgtgca ggaatcgtg tcaggcgata	5040
gtgaactagt atccggaatc tagagcggcc gctggccgca ataaaatatc tttattttca	5100
ttacatctgt gtgttggtt tttgtgtgag gatctaaatg agtcttcgga cctcgcgggg	5160
gccgcttaag cgggtggttag ggtttgtctg acgcgggggg agggggaagg aacgaaacac	5220
tctcattcgg aggcggctcg gggtttggtc ttggtggcca cgggcacgca gaagagcgcc	5280
gcgatcctct taagcacccc cccgccctcc gtggaggcgg gggtttggtc ggcgggtggt	5340
aactggcggg ccgctgactc gggcgggtcg cgcgccccag agtgtgacct tttcggctcg	5400
ctcgcagacc cccgggcggc gccgccgcgg cggcgacggg ctgctgggt cctaggctcc	5460
atggggaccg tatacgtgga caggctctgg agcatccgca cgactgcggt gatattaccg	5520
gagaccttct gcgggacgag ccgggtcacg cggtgacgc ggagcgtccg ttgggcgaca	5580
aacaccagga cggggcacag gtacactatc ttgtcacccg gaggcgcgag ggactgcagg	5640
agcttcaggg agtggcgcag ctgcttcac cccgtggccc gttgctcgcg tttgctggcg	5700
gtgtccccgg aagaaatata tttgcatgtc tttagttcta tgatgacaca aaccccgcgc	5760
agcgtcttgt cattggcgaa ttcgaacacg cagatgcagt cggggcggcg cgggtcccagg	5820
tccacttcgc atattaaggt gacgcgtgtg gcctcgaaca ccgagcgacc ctgcagcgac	5880
ccgcttaaaa gcttggcatt ccggtactgt tggtaaagcc accatggccg atgctaagaa	5940
cattaagaag ggccctgctc ccttctaccc tctggaggat ggcaccgctg gcgagcagct	6000
gcacaaggcc atgaagaggt atgccctggt gcctggcacc attgccttca ccgatgccca	6060
cattgaggtg gacatcacct atgccgagta cttcgagatg tctgtgcgcc tggccgaggc	6120
catgaagagg tacggcctga acaccaacca ccgcatcgtg gtgtgctctg agaactctct	6180
gcagttcttc atgccagtgc tgggcgccct gttcatcggg gtggccgtgg cccctgctaa	6240
cgacatttac aacgagcgcg agctgctgaa cagcatgggc atttctcagc ctaccgtggt	6300
gttcgtgtct aagaagggcc tgcagaagat cctgaacgtg cagaagaagc tgcctatcat	6360
ccagaagatc atcatcatgg actctaagac cgactaccag ggcttcaga gcatgtacac	6420
attcgtgaca tctcatctgc ctctggctt caacgagtac gacttcgtgc cagagtcttt	6480
cgacagggac aaaaccattg ccctgatcat gaacagctct ggggtctaccg gcctgcctaa	6540
gggcgtggcc ctgcctcatc gcaccgcctg tgtgcgcttc tctcacgccc gcgaccctat	6600
tttcggcaac cagatcatcc ccgacaccgc tattctgagc gtggtgccat tccaccacgg	6660
cttcggcatg ttaccacccc tgggctacct gatttgcggc ttccgggtgg tgctgatgta	6720

NaDoc 002-PCT-1_seq listing_ST25.txt

ccgcttcgag gaggagctgt tcctgcgag cctgcaagac tacaaaattc agtctgccct	6780
gctggtgcca accctgttca gcttcttcgc taagagcacc ctgatcgaca agtacgacct	6840
gtctaacctg cacgagattg cctctggcgg cgccccactg tctaaggagg tgggcgaagc	6900
cgtggccaag cgctttcatc tgccaggcat ccgccagggc tacggcctga ccgagacaac	6960
cagcgccatt ctgattacc cagagggcga cgacaagcct ggcgccgtgg gcaagggtgg	7020
gccattcttc gaggccaagg tgggtggacct ggacaccggc aagaccctgg gagtgaacca	7080
gcgcggcgag ctgtgtgtgc gcggccctat gattatgtcc ggctacgtga ataaccctga	7140
ggccacaaac gccctgatcg acaaggacgg ctggctgcac tctggcgaca ttgcctactg	7200
ggacgaggac gagcacttct tcatcgtaga ccgcctgaag tctctgatca agtacaagg	7260
ctaccagggtg gccccagccg agctggagtc tatcctgctg cagcaccta acattttcga	7320
cgccggagtg gccggcctgc ccgacgacga tgccggcgag ctgcctgccg ccgtcgtcgt	7380
gctggaacac ggcaagacca tgaccgagaa ggagatcgtg gactatgtgg ccagccaggt	7440
gacaaccgcc aagaagctgc gcggcgaggt ggtgttcgtg gacgaggtgc ccaagggcct	7500
gaccggcaag ctggacgccc gcaagatccg cgagatcctg atcaaggcta agaaaggcg	7560
caagatcgcc gtgtaataat tctagagtcg gggcgccgg ccgcttcgag cagacatgat	7620
aagatacatt gatgagtttg gacaaaccac aactagaatg cagtgaataa aatgctttat	7680
ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca ataaacaagt	7740
taacaacaac aattgcattc attttatgtt tcaggttcag ggggaggtgt gggaggtttt	7800
ttaaagcaag taaaacctct acaaatgtgg taaaatcgat aaggatccag gtggcacttt	7860
tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta	7920
tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa ggaagagtat	7980
gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt gccttcctgt	8040
ttttgctcac ccagaaacgc tgggtgaaagt aaaagatgct gaagatcagt tgggtgcacg	8100
agtgggttac atcgaactgg atctcaacag cggtgaagatc cttgagagtt ttcgccccga	8160
agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcgg tattatccc	8220
tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga atgacttggt	8280
tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa gagaattatg	8340
cagtgtgcc ataaccatga gtgataaac tgccggccaac ttacttctga caacgatcgg	8400
aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa ctcgccttga	8460
tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca ccacgatgcc	8520
tgtagcaatg gcaacaacgt tgccgaaact attaactggc gaactactta ctctagcttc	8580
ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcgctc	8640
ggcccttccg gctggctgggt ttattgctga taaatctgga gccgggtgagc gtgggtctcg	8700
cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac	8760

NaDoc 002-PCT-1_seq listing_ST25.txt

```

gacggggagt caggcaacta tggatgaacg aaatagacag atcgctgaga taggtgcctc      8820
actgattaag cattggtaac tgtcagacca agtttactca tatatacttt agattgattt      8880
aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac      8940
caaaatccct taacgtgagt tttcgttcca ctgagcgta gaccccgtag aaaagatcaa      9000
aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc      9060
accgctacca gcggtgggtt gtttgccgga tcaagagcta ccaactcttt ttccgaagggt      9120
aactggcttc agcagagcgc agataccaaa tactgttctt ctagtgtagc cgtagttagg      9180
ccaccacttc aagaactctg tagcaccgcc tacatacctc gctctgctaa tcctgttacc      9240
agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt      9300
accggataag gcgcagcggg cgggctgaac ggggggttcg tgcacacagc ccagcttgga      9360
gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct      9420
tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa caggagagcg      9480
cacgaggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca      9540
cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa      9600
cgccagcaac gcggcctttt tacgggttcct ggccttttgc tggccttttg ctcacatggc      9660
tcgac                                     9665

```

<210> 10
 <211> 8290
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

```

<400> 10
aacaaaatat taacgcttac aatttccatt cgccattcag gctgcgcaac tgttgggaag      60
ggcgatcggg gcgggcctct tcgctattac gccagctggc gaaaggggga tgtgctgcaa      120
ggcgattaag ttgggtaacg ccagggtttt ccagtcacg acgttgtaaa acgacggcca      180
gtgccaagct gatctataca ttgaatcaat attggcaatt agccatatta gtcattggtt      240
atatagcata aatcaatatt ggctattggc cattgcatac gttgtatcta tatcataata      300
tgtacattta tattggctca tgtccaatat gaccgccatg ttgacattga ttattgacta      360
gttattaata gtaatcaatt acgggggtcat tagttcatag cccatatatg gagttccgcg      420
ttacataact tacggtaaatt ggccgcctg gctgaccgcc caacgacccc cgccattga      480
cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat tgacgtcaat      540
gggtggagta ttacggtaa actgcccact tggcagtaca tcaagtgtat catatgccaa      600
gtccgcccc tattgacgtc aatgacggta aatggcccg cttggcattat gccagtaca      660
tgaccttacg ggactttcct acttggcagt acatctacgt attagtcac gctattacca      720
tggtgatgcg gttttggcag tacaccaatg ggcgtggata gcggtttgac tcacggggat      780

```

NaDoc 002-PCT-1_seq listing_ST25.txt

ttccaagtct ccacccatt gacgtcaatg ggagtttgtt ttggcaccaa aatcaacggg	840
actttccaaa atgtcgtaat aaccccgccc cgttgacgca aatgggcggg aggcgtgtac	900
ggtgggaggt ctatataagc agagctcggt tagtgaaccg tcagaatttt gtaatacgac	960
tcactatagg gcgggccgga attcgctgac tggatccggg acctagctag gtagcaattg	1020
accggtcaag atggcggcca acaacaacaa caacaacaac aacaacaaca acaacaacaa	1080
caacaagaag atggcggcaa caacaacaac aacaacaaca acaacaacaa caacaacaac	1140
caacaacaag atggcggcca acaacaacaa caacaacaac aacaagaaga tggcggcaac	1200
aacaacaaca acaacaacaa caaccaagat ggcggccaac aacaacaaga agatggcggc	1260
aacaacaaca accaagatgg cggccaacaa caacaagaag atggcggcaa caacaacaac	1320
caagatggcg gcacgcgtcg gtccggctag ccgtacgctc cttagcgacg aaatctactg	1380
ccccctgag agccaccatg gcttggggtc ctacgctgtg caggccaagt ttggagatta	1440
caacaaagaa ggccgccatg gtgggcacct cagctctgag cggctcatcc gccaccatgg	1500
gttgaccag cacaaactta ccagggaccg ccgccatggc cggaccagg cgtgccacca	1560
tggacaccgt gggttgcgcc gccatggtgc tctgttgag tgccaccatg gtgctcagga	1620
cctgggccgc catggaatac ctgataactg ataagccacc atgggaacag acctttggct	1680
tggagttgac gcccttgag tcaacattta cgaggccgcc atggagttca ccccaaagat	1740
tggctttcct tggagtgaat tcaggaacat ctctgccacc atggaaaagt ttgtcatcaa	1800
gcccatcgac aaggccgcca tggactttgt gttttacgcc ccacgtctca cagccaccat	1860
ggggaccctg cagctcgccg ccatggacca cgagttgtac gccaccatgg ggaagcctga	1920
caccgccgcc atggagcaga cgaaggccgc caccatggag gctgataagc tgataagccg	1980
ccatgggctg gaaacagaga agaaaaggag agaaaccgtg gagagagaga aagagcgcca	2040
ccatggcgag aaggaggagt tgttgctgcg gctgcaggac tacgaggaga agacaagccg	2100
ccatgggaga gacctctcgg agcagattca gaggggccac catggggagg aggagaggaa	2160
gcgggcacag gagggccgcc atggcccaga ggctgaccgc caccatggac tgcgggctaa	2220
gggccgccat ggagacagg cgggtgggcca ccatgggagc caggagcagc gccgccatgg	2280
gctacctgat aactgataag ccaccatggt ggaagaggcg cggaggcgca aggaggacga	2340
agttgaagag tggcagcaag ccgccatgga agcccaggac gacctggtca agaccaagga	2400
ggagctgcac ctggtgccgg ccaccatggc gccaccacca ccaccgtgt acgagccggc	2460
cgccatggac gtccaggaga gcttgcaaga cgagggtgcc accatggcgg gctacagcgc	2520
agccgccatg gctgacggca tccgggccac catggacgag gagaagcgtg ccgccatggc	2580
agagaagaac gagggcacca tggggcctga taagctgata agccgccatg gggcccagga	2640
cgagaacaag aggaccaca acgacatcat ccacaacgag agccaccatg gagggcggga	2700
caagtacaag acgctgcggc agatccggca gggcaacacc agccgccatg gcgacgagtt	2760
cgaggccctg caacagccag gccaccatgg agggcagagg ggtgctcata gcgggcgctg	2820

NaDoc 002-PCT-1_seq listing_ST25.txt

ccgccatggc cacgcttgtg tctgccacca tggaagtctc ggaactcgcc gccatggcag	2880
ttcctttcga agccaccatg gcaacagaaa cattcgccgc catggaccac ctgataactg	2940
ataagccacc atggttgcaa tcgtgccaag caggcctgat tctcgcgatt actcgcgaat	3000
caccgccgcc atggtgctgg gagcaggact cattgaatta cggaaaacgc ctgtcaagtc	3060
tcaggccacc atggggaact ggcctgtgtc atacaagagt caggccgccca tggggaaacg	3120
tggcaggact tccatctgtg ccgccaccat ggtgtattcg aaacgagccg ccatggattt	3180
tctcatctct gccaccatgg catctttgta cattgccgcc atgggagggg tcaaaattgc	3240
caccatggtg gctgataagt tgatagtaac cgccatggtg tttcatccag tcgccaccat	3300
gggctggcag agagcagccg ccatggcagc gtcagtgggt gccaccatgg cttggatttt	3360
tttttttgtt tttttttttt ttgctcaaca attttacaac acattgtgtc gacgagctca	3420
agcttcccgg cgcgccccgg tccgtccgga ctacggcaag ctgaccctga agttcatccc	3480
aaaacttacg ctgagtactt cgatctggtc accccggatc cgtgatagta acctgatagt	3540
aacctgataa tagcagatct gcggccgcac tcgaggttta aacggccggc cgcggtcata	3600
gctgtttcct gaacagatcc cgggtggcat ccctgtgacc cctccccagt gcctctcctg	3660
gccctggaag ttgccactcc agtgcccacc agccttgtcc taataaaatt aagttgcatc	3720
attttgtctg actagggtgtc cttctataat attatggggg ggaggggggt ggtatggagc	3780
aaggggcaag ttgggaagac aacctgtagg gcctgcgggg tctattggga accaagctgg	3840
agtgcagtgg cacaatcttg gctcactgca atctccgcct cctgggttca agcgattctc	3900
ctgcctcagc ctcccagatt gttgggattc caggcatgca tgaccaggct cagctaattt	3960
ttgttttttt ggtagagacg gggtttcacc atattggcca cgctgggtct caactcctaa	4020
tctcaggtga tctaccacc ttggcctccc aaattgctgg gattacaggc gtgaaccact	4080
gctcccttcc ctgtccttct gattttaaaa taactatacc agcaggagga cgtccagaca	4140
cagcataggc tacctggcca tgcccaaccg gtgggacatt tgagttgctt gcttggcact	4200
gtcctctcat gcgttgggtc cactcagtag atgcctgttg aattgggtac gcggccagct	4260
tggctgtgga atgtgtgtca gttaggggtg ggaaagtccc caggctcccc agcaggcaga	4320
agtatgcaaa gcatgcatct caattagtca gcaaccagggt gtggaaagtc cccaggctcc	4380
ccagcaggca gaagtatgca aagcatgcat ctcaattagt cagcaaccat agtcccggcc	4440
ctaactccgc ccatcccgcc cctaactccg ccagttccg cccattctcc gccccatggc	4500
tgactaattt tttttattta tgcagaggcc gaggcgcct cggcctctga gctattccag	4560
aagtagtgag gaggtttttt tggaggccta ggcttttgca aaaagctccc gggagcttgt	4620
atatccattt tcggatctga tcaagagaca cgtacgacca tggagagcga cgagagcggc	4680
ctgcccgccca tggagatcga gtgccgcac accggcacc tgaacggcgt ggagttcgag	4740
ctggtgggcg gcggagaggg ccccccgag cagggccgca tgaccaacaa gatgaagagc	4800
accaaaggcg ccctgacctt cagcccctac ctgctgagcc acgtgatggg ctacggcttc	4860

NaDoc 002-PCT-1_seq listing_ST25.txt

taccacttcg gcacctaccc cagcggctac gagaaccctt tcctgcacgc catcaacaac	4920
ggcggctaca ccaacacccg catcgagaag tacgaggacg gcggcgtgct gcacgtgagc	4980
ttcagctacc gctacgaggc cggccgcgtg atcggcgact tcaaggatgat gggcaccggc	5040
ttccccgagg acagcgtgat cttcaccgac aagatcatcc gcagcaacgc caccgtggag	5100
cacctgcacc ccatgggcca taacgatctg gatggcagct tcacccgcac cttcagcctg	5160
cgcgacggcg gctactacag ctccgtgggtg gacagccaca tgcacttcaa gagcgccatc	5220
cacccagca tcctacagaa cgggggcccc atgttcgcct tccgccgcgt ggaggaggat	5280
cacagcaaca ccgagctggg catcgtggag taccagcacg cttcaagac cccggatgca	5340
gatgccggtg aagaataact gcagcgggac tctgggggttc gaaatgaccg accaagcgac	5400
gccccacctg ccatcacgag atttcgattc caccgccgcc ttctatgaaa ggttgggctt	5460
cggaatcgtt ttccgggacg ccggctggat gatcctccag cgcggggatc tcatgctgga	5520
gttcttcgcc caccccaact tgtttattgc agcttataat ggttacaaat aaagcaatag	5580
catcacaat ttacaaata aagcattttt ttcactgcat tctagttgtg gtttgtccaa	5640
actcatcaat gtatcttatt atgtctgtat accgtcgacc tctagctaga gcttggcgta	5700
atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg ctcaaatc cacacaacat	5760
acgagccgga agcataaagt gtaaagcctg ggggtgcctaa tgagttagct aactcacatt	5820
aattgcgttg cgctcactgc ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta	5880
atgaatcggc caacgcgcgg ggagaggcgg ttgtcgtatt gggcgctctt ccgcttcctc	5940
gctcactgac tcgctgcgct cggctgttcg gctgcggcga gcggtatcag ctactcaaa	6000
ggcggttaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa	6060
aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct	6120
ccgccccctt gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac	6180
aggactataa agataccagg cgtttcccc tggaagctcc ctctgctgct ctctgttcc	6240
gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttc	6300
tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttcgtcca agctgggctg	6360
tgtgcacgaa cccccgttc agcccagccg ctgcgcctta tccggttaact atcgtcttga	6420
gtccaacccg gtaagacag acttatcgcc actggcagca gccactggta acaggattag	6480
cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tgggtggccta actacggcta	6540
cactagaaga acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag	6600
agttggtagc tcttgatccg gcaaaaaaac caccgctggt agcgggtggt tttttgttg	6660
caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac	6720
ggggtctgac gctcagtgga acgaaaactc acgttaaggg attttgggtca tgagattatc	6780
aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga agttttaaat caatctaaag	6840
tatatatgag taaacttggg ctgacagtta ccaatgctta atcagtgagg cacctatctc	6900

NaDoc 002-PCT-1_seq listing_ST25.txt

agcgatctgt ctatttcggt catccatagt tgcctgactc cccgtcgtgt agataactac	6960
gatacgggag ggcttaccat ctggccccag tgctgcaatg ataccgagag acccacgctc	7020
accggctcca gatttatcag caataaacca gccagccgga agggccgagc gcagaagtgg	7080
tcctgcaact ttatccgcct ccatccagtc tattaattgt tgccgggaag ctagagtaag	7140
tagttcgcca gttaatagtt tgcgcaacgt tggtgccatt gctacaggca tcgtggtgtc	7200
acgctcgtcg tttggtatgg cttcattcag ctccggttcc caacgatcaa ggcgagttac	7260
atgatcccc atgtttgtgca aaaaagcggg tagctccttc ggtcctccga tcgttgctag	7320
aagtaagttg gccgcagtgt tatcactcat gggttatggca gactgcata attctcttac	7380
tgtcatgcca tccgtaagat gcttttctgt gactggtgag tactcaacca agtcattctg	7440
agaatagtgt atgcggcgac cgagttgctc ttgcccggcg tcaatacggg ataataccgc	7500
gccacatagc agaactttta aagtgtcat cattggaaaa cgttcttcgg ggcgaaaact	7560
ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa cccactcgtg caccacaactg	7620
atcttcagca tcttttactt tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa	7680
tgccgcaaaa aagggaataa gggcgacacg gaaatgttga atactcatac tcttcctttt	7740
tcaatattat tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg	7800
tatttagaaa aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga	7860
cgcgccctgt agcggcgcat taagcgcggc ggggtgtggtg gttacgcgca gcgtgaccgc	7920
tacacttgcc agcgccctag cgcccgtcc ttctgctttc ttcccttcct ttctcgccac	7980
gttcgccggc tttccccgtc aagctctaaa tcgggggctc cctttagggt tccgatttag	8040
tgctttacgg cacctcgacc caaaaaact tgattagggt gatggttcac gtagtgggcc	8100
atcgccctga tagacggttt ttcgcccttt gacgttgag tccacgttct ttaatagtgg	8160
actcttgctc caaactggaa caacactcaa ccctatctcg gtctattctt ttgatttata	8220
agggattttg ccgatttcgg cctattgggt aaaaaatgag ctgatttaac aaaaatttaa	8280
cgcgattttt	8290

<210> 11
 <211> 6441
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 11	
aacaaaatat taacgcttac aatttccatt cgccattcag gctgcgcaac tgttggaag	60
ggcgatcggg ggggctct tcgctattac gccagctggc gaaaggggga tgtgctgaa	120
ggcgattaag ttgggtaacg ccagggtttt cccagtcacg acgttgtaaa acgacggcca	180
gtgccaagct gatctataca ttgaatcaat attggcaatt agccatatta gtcattggtt	240
atatagcata aatcaatatt ggctattggc cattgcatac gttgtatcta tatcataata	300

NaDoc 002-PCT-1_seq listing_ST25.txt

tgtacattta tattggctca tgtccaatat gaccgccatg ttgacattga ttattgacta	360
gttattaata gtaatcaatt acgggggcat tagttcatag cccatatatg gagttccgcg	420
ttacataact tacggtaaatt ggcccgctg gctgaccgcc caacgacccc cgccattga	480
cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat tgacgtcaat	540
gggtggagta ttacggtaa actgcccact tggcagtaca tcaagtgtat catatgccaa	600
gtccgcccc tattgacgtc aatgacggta aatggccgc ctggcattat gcccagtaca	660
tgaccttacg ggactttcct acttggcagt acatctacgt attagtcatc gctattacca	720
tggtgatgcg gttttggcag tacaccaatg ggcgtggata gcggtttgac tcacggggat	780
ttccaagtct ccacccatt gacgtcaatg ggagtttgtt ttggcaccaa aatcaacggg	840
actttccaaa atgtcgtaat aacccgccc cgttgacgca aatgggcggt aggcgtgtac	900
ggtgggaggt ctatataagc agagctcgtt tagtgaaccg tcagaatttt gtaatacgac	960
tcactatagg gcggccggga attcgtcgac tggatccggt accccaagc ttaccggtcc	1020
cacgcgtcaa ttggaaaact tacgtgagt acttcgatct ccctacggca agctgacct	1080
gaagttcaac agatctcgcc gccatgggag ctgatgatgt ggttgattct tcgaaatctt	1140
ttgtcatgga aaacttttct tcgtaccacg ggacgaaacc tggttatgtg gattccattc	1200
aaaaaggcat acaaaagcca aaatctggta cacaaggaaa ctatgacgat gattggaaag	1260
ggttttatag taccgacaac aaatatgacg ctgcgggata ctctgtggat aatgaaaacc	1320
cgctctctgg aaaagctgga ggcgtggta aagtgacgta tccaggactg acgaagggtc	1380
tcgcactaaa ggtggataat gccgaaacta ttaagaaaga gttaggttta agtctcactg	1440
aaccgctcat ggagcaagtc ggaacggaag agtttatcaa aagattcggt gatggtgctt	1500
cgcggtgagt gctcagcctt cccttcgctg aggggagttc tagcgttgag tacatcaaca	1560
actgggaaca ggcgaaagcg ttaagcgtag aacttgagat taactttgaa acccgaggaa	1620
aacgtggcca agatgcgatg tatgagtata tggctcaagc ctgtgcagga aatcgtgtca	1680
ggcgatagtg aactagtatc cggaatctag agcggccgca ctcgaggttt aaacggccgg	1740
ccgcggtcat agctgtttcc tgaacagatc ccgggtggca tccctgtgac ccctccccag	1800
tgctctcct ggccctggaa gttgccactc cagtgccac cagccttgtc ctaataaaat	1860
taagttgcat cattttgtct gactaggtgt cttctataa tattatgggg tggagggggg	1920
tggtatggag caaggggcaa gttgggaaga caacctgtag ggcctgcggg gtctattggg	1980
aaccaagctg gagtgcagtg gcacaatctt ggctcactgc aatctccgcc tcctgggttc	2040
aagcgattct cctgcctcag cctcccaggt tgttgggatt ccaggcatgc atgaccaggc	2100
tcagctaatt tttgtttttt tggtagagac ggggtttcac catattggcc acgctggtct	2160
ccaactccta atctcaggtg atctaccac cttggcctcc caaattgctg ggattacagg	2220
cgtgaaccac tgctcccttc cctgtccttc tgattttaaa ataactatac cagcaggagg	2280
acgtccagac acagcatagg ctacctggcc atgcccacc ggtgggacat ttgagttgct	2340

NaDoc 002-PCT-1_seq listing_ST25.txt

tgcttggcac	tgtcctctca	tgcgttgggt	ccactcagta	gatgcctggt	gaattgggta	2400
cgcgccagc	ttggctgtgg	aatgtgtgtc	agttagggtg	tggaaagtcc	ccaggctccc	2460
cagcaggcag	aagtatgcaa	agcatgcata	tcaattagtc	agcaaccagg	tgtggaaagt	2520
ccccaggctc	cccagcaggc	agaagtatgc	aaagcatgca	tctcaattag	tcagcaacca	2580
tagtccccgc	cctaactccg	cccatccccg	ccctaactcc	gccagttcc	gcccattctc	2640
cgccccatgg	ctgactaatt	ttttttat	atgcagaggc	cgaggccgcc	tcggcctctg	2700
agctattcca	gaagtagtga	ggaggctttt	ttggaggcct	aggcttttgc	aaaaagctcc	2760
cgggagcttg	tatatccatt	ttcgatctg	atcaagagac	acgtacgacc	atggagagcg	2820
acgagagcgg	cctgccccgc	atggagatcg	agtgccgcac	caccggcacc	ctgaacggcg	2880
tggagttcga	gctggtgggc	ggcggagagg	gcacccccga	gcagggccgc	atgaccaaca	2940
agatgaagag	caccaaaggc	gccctgacct	tcagccccta	cctgctgagc	cacgtgatgg	3000
gctacggctt	ctaccacttc	ggcacctacc	ccagcggcta	cgagaacccc	ttcctgcacg	3060
ccatcaacaa	cggcggctac	accaacaccc	gcatacgaaa	gtacgaggac	ggcggcgtgc	3120
tgcacgtgag	cttcagctac	cgctacgagg	ccggccgcgt	gatcggcgac	ttcaagggtga	3180
tgggcaccgg	cttccccgag	gacagcgtga	tcttcaccga	caagatcatc	cgcagcaacg	3240
ccaccgtgga	gcacctgcac	ccatggggcg	ataacgatct	ggatggcagc	ttcaccgcga	3300
ccttcagcct	gcgcgacggc	ggctactaca	gctccgtggt	ggacagccac	atgcacttca	3360
agagcgccat	ccacccagc	atcctacaga	acggggggccc	catgttcgcc	ttccgccgcg	3420
tggaggagga	tcacagcaac	accgagctgg	gcatacgtga	gtaccagcac	gccttcaaga	3480
ccccgatgc	agatgccggt	gaagaataac	tgcagcggga	ctctgggggt	cgaaatgacc	3540
gaccaagcga	cgcccaacct	gccatcacga	gatttcgatt	ccaccgccgc	cttctatgaa	3600
aggttgggct	tcggaatcgt	tttccgggac	gccggctgga	tgatcctcca	gcgcggggat	3660
ctcatgctgg	agttcttcgc	ccaccccaac	ttgtttattg	cagcttataa	tggttacaaa	3720
taaagcaata	gcatacaaaa	tttcacaaat	aaagcatttt	tttactgca	ttctagttgt	3780
ggtttgtcca	aactcatcaa	tgtatcttat	catgtctgta	taccgtcgac	ctctagctag	3840
agcttggcgt	aatcatggtc	atagctgttt	cctgtgtgaa	attgttatcc	gctcacaatt	3900
ccacacaaca	tacgagccgg	aagcataaag	tgtaaagcct	ggggtgccta	atgagtgagc	3960
taactcacat	taattgcgtt	gcgctcactg	cccgccttcc	agtcgggaaa	cctgtcgtgc	4020
cagctgcatt	aatgaatcgg	ccaacgcgcg	gggagaggcg	gtttgcgtat	tgggcgctct	4080
tccgcttcct	cgctcactga	ctcgtgcgc	tcggtcgttc	ggctgcggcg	agcggtatca	4140
gctcactcaa	aggcggtaat	acggttatcc	acagaatcag	gggataacgc	aggaaagaac	4200
atgtgagcaa	aaggccagca	aaaggccagg	aaccgtaaaa	aggccgcgtt	gctggcgttt	4260
ttccataggc	tccgcccccc	tgacgagcat	cacaaaaatc	gacgctcaag	tcagaggtgg	4320
cgaaacccga	caggactata	aagataccag	gcgtttcccc	ctggaagctc	cctcgtgcgc	4380

NaDoc 002-PCT-1_seq listing_ST25.txt

tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc	4440
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cgggtgtaggt cgttcgctcc	4500
aagctgggct gtgtgcacga accccccgtt cagcccgacc gctgcgcctt atccggtaac	4560
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt	4620
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct	4680
aactacggct aactagaag aacagtatct ggatatctgcg ctctgctgaa gccagttacc	4740
ttcggaaaaa gaggttgtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt	4800
ttttttgttt gcaagcagca gattacgcgc agaaaaaag gatctcaaga agatcctttg	4860
atcttttcta cggggtctga cgctcagtg aacgaaaact cacgttaagg gattttgttc	4920
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa	4980
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtga	5040
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgctcg	5100
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga	5160
gaccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag	5220
cgcagaagtg gtcctgaac tttatccgcc tccatccagt ctattaattg ttgccgggaa	5280
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc	5340
atcgtggtgt cagcctcgtc gtttggtatg gcttcattca gctccgggtc ccaacgatca	5400
aggcgagtta catgatcccc catgttgtag aaaaaagcgg ttagctcctt cggtcctccg	5460
atcgttgta gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat	5520
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc	5580
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg	5640
gataataccg cgccacatag cagaacttta aaagtgtcga tcattggaaa acgttcttcg	5700
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt	5760
gcaccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca	5820
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aataactcata	5880
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac	5940
atatttgaat gtatttagaa aaataaacia ataggggttc cgcgcacatt tccccgaaaa	6000
gtgccacctg acgcgccctg tagcggcgca ttaagcgcgg cgggtgtggt ggttacgcgc	6060
agcgtgaccg ctacacttg cagcgcccta gcgcccgtc ctttcgcttt cttcccttcc	6120
tttctcgcca cgttcgccgg ctttccccgt caagctctaa atcgggggct ccctttaggg	6180
ttccgattta gtgctttacg gcacctcgac cccaaaaaac ttgattaggg tgatggttca	6240
cgtagtgggc catcgccctg atagacggtt tttcgccctt tgacgttga gtccacgttc	6300
tttaatagtg gactcttggt ccaaactgga acaacactca accctatctc ggtctattct	6360
tttgatttat aagggttttt gccgatttcg gcctattggg taaaaaatga gctgatttaa	6420

NaDoc 002-PCT-1_seq listing_ST25.txt

caaaaattta acgcaattt t

6441

<210> 12
 <211> 6368
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 12
 aacaaaatat taacgcttac aatttccatt cgccattcag gctgcgcaac tgttggaag 60
 ggcgatcggg ggggctct tcgctattac gccagctggc gaaagggga tgtgctgcaa 120
 ggcgattaag ttgggtaacg ccagggtttt ccagtcacg acgttgtaa acgacggcca 180
 gtgccaagct gatctataca ttgaatcaat attggcaatt agccatatta gtcattgggt 240
 atatagcata aatcaatatt ggctattggc cattgcatac gttgtatcta taccataata 300
 tgtacattta tattgggtca tgtccaatat gaccgccatg ttgacattga ttattgacta 360
 gttattaata gtaatcaatt acggggctcat tagttcatag cccatatatg gagttccgcg 420
 ttacataact tacggtaaat ggccgcctg gctgaccgcc caacgacccc cgccattga 480
 cgtcaataat gacgtatggt cccatagtaa cgccaatagg gactttccat tgacgtcaat 540
 ggggtggagta ttacggtaa actgcccact tggcagtaca tcaagtgtat catatgcca 600
 gtccgcccc tattgacgtc aatgacggta aatggccgc ctggcattat gccagtaga 660
 tgaccttacg ggactttcct acttggcagt acatctacgt attagtcac gctattacca 720
 tgggtgatgcg gttttggcag tacaccaatg ggcgtggata gcggttgac tcacggggat 780
 ttccaagtct ccacccatt gacgtcaatg ggagtttgtt ttggcacaa aatcaacggg 840
 actttccaaa atgtcgtaat aacccgccc cgttgacgca aatgggagggt aggcgtgtac 900
 ggtgggagggt ctatataagc agagctcgtt tagtgaaccg tcagaatttt gtaatacgac 960
 tcactatagg gcggccggga attcgtcgac tggatctgct agcggcgcg ccccggtacc 1020
 tgataagcct agcagcgaat gcctggggca gacgatgtcg tcgacagtag caagagcttt 1080
 gtgatggaga attttagtag ctatcatggt actaagccgg gatacgtaga tagtatccag 1140
 aagggaatcc agaaaccaa gagcgggaact cagggaact acgatgacga ctggaagggt 1200
 ttctactcga ccgataaaa atatgatgca gccgggtaca gcgtggacaa cgagaatcct 1260
 ttgagcggca aggcaggcgg ggtcgtcaag gtcacctacc ccggtttaac caaagtgtta 1320
 gctttgaagg tggacaacgc ggagacaatc aaaaaggaa tcggactctc gtcacggag 1380
 cctcttatgg aacaggctcg caccgaggaa ttcataaagc gttttggaga tggagcaagt 1440
 agggttgtct tatcattgcc atttgcggaa ggctcgagct cagtggagta cataaacaat 1500
 tgggagcaag ccaaggcact ctcaattgag ctggagatca acttcgagac aagaggcaag 1560
 agagggcagg acgcatgta cgagtacatg gcacaggcgt gcgctggcaa cagagtccgt 1620
 aggtgaacat aagcataggc ggccgactc gaggtttaaa cgccggccg cggtcatagc 1680

NaDoc 002-PCT-1_seq listing_ST25.txt

tgtttcctga acagatcccc ggtggcatcc ctgtgacccc tccccagtg cttctctggc	1740
cctggaagtt gccactccag tgcccaccag cttgttccta ataaaattaa gttgcatcat	1800
tttgtctgac taggtgtcct tctataatat tatgggggtgg aggggggtgg tatggagcaa	1860
ggggcaagtt ggggaagacaa cctgtagggc ctgcggggtc tattgggaac caagctggag	1920
tgcagtggca caatcttggc tctactgcaat ctccgcctcc tgggttcaag cgattctcct	1980
gcctcagcct cccgagttgt tgggattcca ggcattgcac accaggctca gctaattttt	2040
gttttttttg tagagacggg gtttcacat attggccacg ctggtctcca actcctaata	2100
tcaggtgatc taccacattt ggcctcccaa attgtggga ttacaggcgt gaaccactgc	2160
tcccttccct gtccttctga ttttaaaata actataccag caggaggacg tccagacaca	2220
gcataggcta cctggccatg cccaaccggt gggacatttg agttgcttgc ttggcactgt	2280
cctctcatgc gttgggtcca ctacagtagat gcctgttgaa ttgggtacgc ggccagcttg	2340
gctgtggaat gtgtgtcagt tagggtgtgg aaagtcccca ggctccccag caggcagaag	2400
tatgcaaagc atgcatctca attagtacgc aaccagggtgt ggaaagtccc caggctcccc	2460
agcaggcaga agtatgcaaa gcatgcatct caattagtca gcaaccatag tcccgccctt	2520
aactccgcc atcccgcccc taactccgcc cagttccgcc cattctccgc cccatggctg	2580
actaattttt tttatttatg cagaggccga ggccgcctcg gcctctgagc tattccagaa	2640
gtagtgagga ggcttttttg gaggcctagg cttttgcaaa aagctcccgg gagcttgat	2700
atccattttt ggatctgatc aagagacacg tacgaccatg gagagcgacg agagcggcct	2760
gcccgccatg gagatcgagt gccgcatcac cggcaccctg aacggcgtgg agttcgagct	2820
ggtgggcggc ggagagggca cccccagca gggccgcatg accaacaaga tgaagagcac	2880
caaaggcgcc ctgaccttca gcccctacct gctgagccac gtgatgggct acggcttcta	2940
ccacttcggc acctacccca gcggctacga gaacccctt ctgcacgcca tcaacaacgg	3000
cggctacacc aacaccgca tcgagaagta cgaggacggc ggcgtgctgc acgtgagctt	3060
cagctaccgc tacgaggccg gccgcgtgat cggcgacttc aaggatgatg gcaccggctt	3120
ccccgaggac agcgtgatct tcaccgacaa gatcatccgc agcaacgcca ccgtggagca	3180
cctgcacccc atgggcgata acgatctgga tggcagcttc acccgcacct tcagcctgag	3240
cgacggcggc tactacagct ccgtgggtgga cagccacatg cacttcaaga gcgccatcca	3300
ccccagcatc ctacagaacg ggggccccat gttcgccttc cgccgcgtgg aggaggatca	3360
cagcaacacc gagctgggca tcgtggagta ccagcacgcc ttcaagaccc cggatgcaga	3420
tgccggtgaa gaataactgc agcgggactc tgggggttca aatgaccgac caagcgacgc	3480
ccaacctgcc atcacgagat ttcgattcca ccgcccctt ctatgaaagg ttgggcttcg	3540
gaatcgtttt ccgggacgcc ggctggatga tcctccagcg cggggatctc atgctggagt	3600
tcttcgcccc cccaacttg tttattgcag cttataatgg ttacaaataa agcaatagca	3660
tcacaaattt cacaataaaa gcattttttt cactgcattc tagttgtggt ttgtccaaac	3720

NaDoc 002-PCT-1_seq listing_ST25.txt

tcatcaatgt atcttatcat gtctgtatac cgtcgacctc tagctagagc ttggcgtaat	3780
catggtcata gctgtttcct gtgtgaaatt gttatccgct cacaattcca cacaacatac	3840
gagccggaag cataaagtgt aaagcctggg gtgcctaata agtgagctaa ctcacattaa	3900
ttgcgttgcg ctactgccc gctttccagt cgggaaacct gtcgtgccag ctgcattaat	3960
gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc	4020
tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg	4080
cggtaatacgc gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag	4140
gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc	4200
gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag	4260
gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga	4320
ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaaagcgtg gcgctttctc	4380
atagctcacg ctgtaggtat ctcaattcgg tgtaggtcgt tcgctccaag ctgggctgtg	4440
tgcacgaacc ccccgctcag cccgaccgct gcgccttate cggtaaactat cgtcttgagt	4500
ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca	4560
gagcgaggta tgtaggcggg gctacagagt tcttgaagtg gtggcctaac tacggctaca	4620
ctagaagaac agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag	4680
ttggtagctc ttgatccggc aaacaaacca ccgctggtag cgggtggtttt tttgtttgca	4740
agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg	4800
ggctctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa	4860
aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta	4920
tatatgagta aacttggctc gacagttacc aatgcttaat cagtgaggca cctatctcag	4980
cgatctgtct atttcgttca tccatagttg cctgactccc cgctcgttag ataactacga	5040
tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac ccacgctcac	5100
cggctccaga tttatcagca ataaaccagc cagccggaag ggccgagcgc agaagtggc	5160
ctgcaacttt atccgcctcc atccagtcta ttaattgttg ccgggaagct agagtaagta	5220
gttcgccagt taatagtttg cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac	5280
gctcgtcgtt tggatggct tcattcagct ccggttccca acgatcaagg cgagttacat	5340
gatcccccat gttgtgcaa aaagcgggta gctccttcgg tcctccgatc gttgtcagaa	5400
gtaagttagc cgagtggtta tcactcatgg ttatggcagc actgcataat tctcttactg	5460
tcatgccatc cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag tcattctgag	5520
aatagtgtat gcggcgaccg agttgctctt gccggcgctc aatacgggat aataccgcgc	5580
cacatagcag aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct	5640
caaggatctt accgctgttg agatccagtt cgatgtaacc cactcgtgca cccaactgat	5700
cttcagcatc ttttactttc accagcgttt ctgggtgagc aaaaacagga aggcaaaatg	5760

NaDoc 002-PCT-1_seq Listing_ST25.txt

```

ccgcaaaaaa ggaataagg ggcacacgga aatgttgaat actcatactc ttcctttttc 5820
aatattattg aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta 5880
tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg 5940
cgccctgtag cggcgcatga agcgcggcgg gtgtggtggt tacgcgcagc gtgaccgcta 6000
cacttgccag cgccctagcg cccgctcctt tcgctttcct cccttccttt ctcgccacgt 6060
tcgccggctt tccccgtcaa gctctaaatc gggggctccc tttagggttc cgatttagtg 6120
ctttacggca cctcgacccc aaaaaacttg attagggtag tggttcacgt agtgggcat 6180
cgccctgata gacggttttt cgcccttga cgttgagtc cacgttcttt aatagtggac 6240
tcttgttcca aactggaaca aactcaacc ctatctcggg ctattctttt gatttataag 6300
ggattttgcc gatttcggcc tattggttaa aaaatgagct gatttaaca aaatttaacg 6360
cgaatttt 6368

```

<210> 13
 <211> 6015
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

```

<400> 13
agatctgcgc agcaccatgg cctgaaataa cctctgaaag aggaacttgg ttaggtacct 60
tctgaggcgg aaagaaccag ctgtggaatg tgtgtcagtt aggggtgtgga aagtccccag 120
gctccccagc aggcagaagt atgcaaagca tgcattctca ttagtcagca accagggtgtg 180
gaaagtcccc aggtccccca gcaggcagaa gtatgcaaag catgcatctc aattagtcag 240
caaccatagt cccgccccta actccgcccc tccgcccctt aactccgccc agtccgccc 300
attctccgcc ccatggctga ctaatttttt ttatttatgc agaggccgag gccgcctcgg 360
cctctgagct attccagaag tagtgaggag gcttttttgg aggcctaggc ttttgcaaaa 420
agcttgattc ttctgacaca acagtctcga acttaagctg cagaagttgg tcgtgaggca 480
ctgggcaggt aagtatcaag gttacaagac aggtttaagg agaccaatag aaactgggct 540
tgtcgagaca gagaagactc ttgcgtttct gataggcacc tattggtctt actgacatcc 600
actttgcctt tctctccaca ggtgtccact cccagttcaa ttacagctct taaggctaga 660
gtacttaata cgactcacta taggctagac gcggtaccta gctaggtagc aattgaccgg 720
tcccacgcgt caattgaaa acttacgctg agtacttca tctccctacg gcaagctgac 780
cctgaagtgc aacagatctc gccgccatgg gagctgatga tgtggttgat tcttcgaaat 840
ctttgtcat ggaaaacttt tcttcgtacc acgggacgaa acctgggtat gtggattcca 900
ttcaaaaagg catacaaaag ccaaaatctg gtacacaagg aaactatgac gatgattgga 960
aagggtttta tagtaccgac acaaatatg acgctgcggg atactctgtg gataatgaaa 1020
acccgctctc tggaaaagct ggaggcgtgg tcaaagtac gtatccagga ctgacgaagg 1080

```

NaDoc 002-PCT-1_seq listing_ST25.txt

ttctcgcact aaaggtggat aatgccgaaa ctattaagaa agagttaggt ttaagttctca	1140
ctgaaccgct catggagcaa gtcggaacgg aagagtttat caaaagattc ggtgatggtg	1200
cttcgcgtgt agtgctcagc cttcccttcg ctgaggggag ttctagcggt gagtacatca	1260
acaactggga acaggcgaaa gcgttaagcg tagaacttga gattaacttt gaaacccgtg	1320
gaaaacgtgg ccaagatgcg atgtatgagt atatggctca agcctgtgca ggaaatcgtg	1380
tcaggcgata gtgaactagt atccggaatc tagagcggcc gctggccgca ataaaatatc	1440
tttattttca ttacatctgt gtgttggttt tttgtgtgag gatctaaatg agtcttcgga	1500
cctcgcgggg gccgcttaag cgggtggtag ggtttgtctg acgcgggggg agggggaagg	1560
aacgaaacac tctcattcgg aggcggctcg gggtttggtc ttggtggcca cgggcacgca	1620
gaagagcgcc gcgactctct taagcacccc cccgcctcc gtggaggcgg gggtttggtc	1680
ggcgggtggt aactggcggg ccgctgactc gggcgggtcg cgcgccccag agtgtgacct	1740
tttcggtctg ctcgcagacc cccgggcggc gccgcgcgg cggcgacggg ctcgctgggt	1800
cctaggctcc atggggaccg tatacgtgga caggctctgg agcatccgca cgactgcggt	1860
gatattaccg gagaccttct gcgggacgag ccgggtcacg cggtgacgc ggagcgtccg	1920
ttgggcgaca aacaccagga cggggcacag gtacactatc ttgtcaccg gaggcgcgag	1980
ggactgcagg agcttcaggg agtggcgcag ctgcttcac cccgtggccc gttgctcgcg	2040
tttgctggcg gtgtccccg aagaaatata tttgcatgtc tttagttcta tgatgacaca	2100
aaccccgccc agcgtcttgt cattggcgaa ttcgaacacg cagatgcagt cggggcggcg	2160
cgggtcccagg tccacttcgc atattaaggt gacgcgtgtg gcctcgaaca ccgagcgacc	2220
ctgcagcgac ccgcttaaaa gcttggcatt ccggtactgt tggtaaagcc accatggccg	2280
atgctaagaa cattaagaag ggccctgctc ctttctaccc tctggaggat ggcaccgctg	2340
gcgagcagct gcacaaggcc atgaagaggt atgccctggt gcctggcacc attgccttca	2400
ccgatgccc cattgagggt gacatcacct atgccgagta cttcgagatg tctgtgcgcc	2460
tggccgaggg catgaagagg tacggcctga acaccaacca ccgcatcgtg gtgtgctctg	2520
agaactctct gcagttcttc atgccagtgc tgggcgccct gttcatcgga gtggccgtgg	2580
cccctgctaa cgacatttac aacgagcgcg agctgctgaa cagcatgggc atttctcagc	2640
ctaccgtggt gttcgtgtct aagaagggcc tgcagaagat cctgaacgtg cagaagaagc	2700
tgccatcat ccagaagatc atcatcatgg actctaagac cgactaccag ggcttccaga	2760
gcatgtacac attcgtgaca tctcatctgc ctctggctt caacgagtac gacttcgtgc	2820
cagagtcttt cgacagggac aaaaccattg ccctgatcat gaacagctct gggcttaccg	2880
gcctgcctaa gggcgtggcc ctgcctcatc gcaccgcctg tgtgcgcttc tctcagccc	2940
gcgaccctat tttcggcaac cagatcatcc ccgacaccgc tattctgagc gtggtgccat	3000
tccaccacgg cttcggcatg ttcaccaccc tgggctacct gatttgcggc tttcgggtgg	3060
tgctgatgta ccgcttcgag gaggagctgt tcctgcgcag cctgcaagac taaaaattc	3120

NaDoc 002-PCT-1_seq listing_ST25.txt

agtctgccct gctggtgcca accctgttca gcttcttcgc taagagcacc ctgatcgaca	3180
agtacgacct gtctaacctg cacgagattg cctctggcgg cgccccactg tctaaggagg	3240
tgggcgaagc cgtggccaag cgctttcatc tgccaggcat ccgccagggc tacggcctga	3300
ccgagacaac cagcgccatt ctgattaccc cagagggcga cgacaagcct ggcgccgtgg	3360
gcaaggtggt gccattcttc gaggccaagg tgggtggacct ggacaccggc aagaccctgg	3420
gagtgaacca gcgcggcgag ctgtgtgtgc gcggccctat gattatgtcc ggctacgtga	3480
ataaccctga ggccacaaac gccctgatcg acaaggacgg ctggctgcac tctggcgaca	3540
ttgcctactg ggacgaggac gagcacttct tcatcgtgga ccgcctgaag tctctgatca	3600
agtacaaggg ctaccagggtg gccccagccg agctggagtc tatcctgctg cagcacccta	3660
acattttcga cgccggagtg gccggcctgc ccgacgacga tgccggcgag ctgcctgccg	3720
ccgtcgtcgt gctggaacac ggcaagacca tgaccgagaa ggagatcgtg gactatgtgg	3780
ccagccagggt gacaaccgcc aagaagctgc gcggcggagt ggtgttcgtg gacgaggtgc	3840
ccaagggcct gaccggcaag ctggacgccc gcaagatccg cgagatcctg atcaaggcta	3900
agaaaggcgg caagatcgcc gtgtaataat tctagagtcg gggcgccgg ccgcttcgag	3960
cagacatgat aagatacatt gatgagtttg gacaaaccac aactagaatg cagtgaaaaa	4020
aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca	4080
ataaacaagt taacaacaac aattgcattc attttatgtt tcaggttcag ggggaggtgt	4140
gggaggtttt ttaaagcaag taaaacctct acaaagtggg taaaatcgat aaggatccag	4200
gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt	4260
caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa	4320
ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt	4380
gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt	4440
tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaaagatc cttgagagtt	4500
ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcg	4560
tattatcccc tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga	4620
atgacttggt tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa	4680
gagaattatg cagtgtctgc ataaccatga gtgataacac tgccggccaac ttacttctga	4740
caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa	4800
ctgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca	4860
ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaactggc gaactactta	4920
ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac	4980
ttctgcgctc ggcccttccg gctggctggt ttattgctga taaatctgga gccggtgagc	5040
gtgggtctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag	5100
ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga	5160

NaDoc 002-PCT-1_seq listing_ST25.txt

taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca tatatacttt	5220
agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata	5280
atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag	5340
aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa	5400
caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt	5460
ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgttctt ctagtgtagc	5520
cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatacctc gctctgctaa	5580
tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa	5640
gacgatagtt accggataag gcgcagcggc cgggctgaac ggggggttcg tgcacacagc	5700
ccagcttggg gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa	5760
gcgccacgct tcccgaaggg agaaaggcgg acagggtatcc ggtaagcggc agggctcgaa	5820
caggagagcg cacgagggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg	5880
ggtttcgcca cctctgactt gagcgtcgat ttttgatg ctcgtcaggg gggcggagcc	5940
tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccctttgc tggccttttg	6000
ctcacatggc tcgac	6015

<210> 14
 <211> 6352
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 14	
aacaaaatat taacgcttac aatttccatt cgccattcag gctgcgcaac tgttggaag	60
ggcgatcggg gcgggcctct tcgctattac gccagctggc gaaaggggga tgtgctgcaa	120
ggcgattaag ttgggtaacg ccagggtttt ccagtcacg acgttgtaaa acgacggcca	180
gtgccaagct gatctataca ttgaatcaat attggcaatt agccatatta gtcattggtt	240
atatagcata aatcaatatt ggctattggc cattgcatac gttgtatcta ttcataata	300
tgtacattta tattggctca tgtccaatat gaccgccatg ttgacattga ttattgacta	360
gttattaata gtaatcaatt acgggggtcat tagttcatag cccatatatg gagttccgcg	420
ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc cgccattga	480
cgtaataat gacgtatggt cccatagtaa cgccaatagg gactttccat tgacgtcaat	540
gggtggagta ttacggtaa actgcccact tggcagtaca tcaagtgtat catatgcaa	600
gtccgcccc tattgacgtc aatgacggta aatggcccgc ctggcattat gccagtaca	660
tgaccttacg ggactttcct acttggcagt acatctacgt attagtcac gctattacca	720
tggtgatgcg gttttggcag tacaccaatg ggcgtggata gcggtttgac tcacggggat	780
ttccaagtct ccaccccat gacgtcaatg ggagtttgtt ttggcaccaa aatcaacggg	840

NaDoc 002-PCT-1_seq listing_ST25.txt

actttccaaa atgtcgtaat aacccccgcc cgttgacgca aatgggagggt aggcgtgtac	900
ggtagggagggt ctatataagc agagctcggt tagtgaaccg tcagaatttt gtaatacgac	960
tcactatagg gcggccggga attcgtcgac tggatcttgt acattcgaac gccgccatgg	1020
gcgctgatga tgttggtgat tcttctaaat cttttgtcat ggaaaacttt tcttcgtacc	1080
acgggactaa acctggttat gtggattcca ttcaaaaagg tatacaaaag ccaaaatctg	1140
gtacacaagg aaattatgac gatgattgga aagggtttta tagtaccgac aataaatacg	1200
acgctgcggg atactctgtg gataatgaaa acccgctctc tggaaaagct ggaggcgtgg	1260
tcaaagtgac gtatccagga ctgacgaagg ttctcgact aaaagtggat aatgccgaaa	1320
ctattaagaa agagttaggt ttaagtctca ctgaaccgtt gatggagcaa gtcggaacgg	1380
aagagtttat caaaagggtt ggtgatgggt cttcgcgtgt agtgctcagc cttcccttcg	1440
ctgaggggag ttctagcggt gaatatatta ataactggga acaggcgaaa gcgttaagcg	1500
tagaacttga gattaatttt gaaacccgtg gaaaacgtgg ccaagatgag atgtatgagt	1560
atatggctca agcctgtgca ggaaatcgtg tcaggcgata gtgaactagt tccggatcta	1620
gagcgccgc actcgagggt taaacggccg gccgcggtca tagctgtttc ctgaacagat	1680
cccgggtggc atccctgtga cccctcccca gtgcctctcc tggccctgga agttgccact	1740
ccagtgccca ccagccttgt cctaataaaa ttaagttgca tcattttgtc tgactagggtg	1800
tccttctata atattatggg gtggaggggg gtggtatgga gcaaggggca agttgggaag	1860
acaacctgta gggcctgcgg ggtctattgg gaaccaagct ggagtgcagt ggcacaatct	1920
tggctcactg caatctccgc ctctgggtt caagcgattc tcctgcctca gcctcccgag	1980
ttgttgggat tccaggcatg catgaccagg ctacagctaat ttttgttttt ttggtagaga	2040
cggggtttca ccatattggc cagcgtggtc tccaactcct aatctcaggt gatctacca	2100
ccttggcctc ccaaattgct gggattacag gcgtgaacca ctgctccctt ccctgtcctt	2160
ctgattttta aataactata ccagcaggag gacgtccaga cacagcatag gctacctggc	2220
catgccaac cggtagggaca tttgagttgc ttgcttggca ctgtcctctc atgcgttggg	2280
tccactcagt agatgcctgt tgaattgggt acgcggccag cttggctgtg gaatgtgtgt	2340
cagttagggt gtggaaagtc cccaggctcc ccagcaggca gaagtatgca aagcatgcat	2400
ctcaattagt cagcaaccag gtgtggaaag tccccaggct cccagcagg cagaagtatg	2460
caaagcatgc atctcaatta gtcagcaacc atagtccgc ccctaactcc gcccatcccg	2520
cccctaactc cgcccagttc cgccattctt ccgccccatg gctgactaat tttttttatt	2580
tatgcagagg ccgaggccgc ctggcctct gagctattcc agaagtagtg aggaggcttt	2640
tttggaggcc taggcttttg caaaaagctc ccgggagctt gtatatccat tttcgatct	2700
gatcaagaga cacgtacgac catggagagc gacgagagcg gcctgcccgc catggagatc	2760
gagtgccgca tcaccggcac cctgaacggc gtggagttcg agctggtggg cggcggagag	2820
ggcacccccg agcagggccg catgaccaac aagatgaaga gcaccaaagg cgccctgacc	2880

NaDoc 002-PCT-1_seq listing_ST25.txt

ttcagccctt acctgctgag ccacgtgatg ggctacggct tctaccactt cggcacctac	2940
cccagcggct acgagaaccc cttcctgcac gccatcaaca acggcggcta caccaacacc	3000
cgcacgcaga agtacgagga cggcggcgctg ctgcacgtga gcttcagcta ccgctacgag	3060
gccggcccg tgcacggcga cttcaagggtg atgggcaccg gcttccccga ggacagcgtg	3120
atcttcaccg acaagatcat ccgcagcaac gccaccgtgg agcacctgca ccccatgggc	3180
gataacgata tggatggcag cttcaccgc accctcagcc tgcgcgacgg cggctactac	3240
agctccgtgg tggacagcca catgcacttc aagagcgcca tccacccag catcctacag	3300
aacgggggccc ccatgttcgc cttccgccgc gtggaggagg atcacagcaa caccgagctg	3360
ggcatcgtgg agtaccagca cgccttcaag acccggatg cagatgccgg tgaagaataa	3420
ctgcagcggg actctgggggt tcgaaatgac cgaccaagcg acgccaacc tgccatcacg	3480
agatttcgat tccaccgccg cttctatga aagggtgggc ttcggaatcg ttttccggga	3540
cgccggctgg atgatcctcc agcgcgggga tctcatgctg gagttcttcg cccaccccaa	3600
cttgtttatt gcagcttata atggttataa ataaagcaat agcatcaca atttcacaaa	3660
taaagcattt ttttactgc attctagttg tggttgtcc aaactcatca atgtatctta	3720
tcatgtctgt ataccgtcga cctctagcta gagcttgccg taatcatggt catagctggt	3780
tcctgtgtga aattgttatc cgctcacaat tccacacaac atacgagccg gaagcataaa	3840
gtgtaaagcc tggggtgcct aatgagtgag ctaactcaca ttaattgcgt tgcgctcact	3900
gcccgtttc cagtcgggaa acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc	3960
ggggagaggc ggtttgcgta ttggcgctc ttccgcttc tcgctcactg actcgtcgcg	4020
ctcggctcgtt cggctgcggc gagcggatc agctcactca aaggcggtaa tacggttatc	4080
cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag	4140
gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg ctccgcccc ctgacgagca	4200
tcacaaaaat cgacgtcaa gtcagagggtg gcgaaacccg acaggactat aaagatacca	4260
ggcgtttccc cctggaagct cctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg	4320
atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcatagct cacgctgtag	4380
gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aacccccgt	4440
tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtaagaca	4500
cgacttatcg cactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg	4560
cggtgctaca gagttcttga agtggtggcc taactacggc tacactagaa gaacagtatt	4620
tggtatctgc gctctgctga agccagttac cttcgaaaa agagttggta gctcttgatc	4680
cggcaaaaca accaccgctg gtagcgggtg tttttttgtt tgcaagcagc agattacgcg	4740
cagaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg acgctcagtg	4800
gaacgaaaac tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta	4860
gatcctttta aattaaaaat gaagttttta atcaatctaa agtatatatg agtaaacttg	4920

NaDoc 002-PCT-1_seq listing_ST25.txt

```

gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg      4980
ttcatccata gttgcctgac tccccgtcgt gtagataact acgatacggg agggccttacc      5040
atctggcccc agtgctgcaa tgataccgcg agaccacgc tcaccggctc cagattttatc      5100
agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc      5160
ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag      5220
tttgcgcaac gttgttgcca ttgctacagg catcgtggtg tcacgctcgt cgtttggtat      5280
ggcttcattc agtccgggtt cccaacgata aaggcgagtt acatgatccc ccatgttggtg      5340
caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc agaagtaagt tggccgcagt      5400
gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag      5460
atgcttttct gtgactgggt agtactcaac caagtcattc tgagaatagt gtatgcggcg      5520
accgagttgc tcttgcccgg cgtcaatacg ggataatacc gcgccacata gcagaacttt      5580
aaaagtgctc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct      5640
gttgagatcc agttcgatgt aaccactcgc tgcacccaac tgatcttcag catcttttac      5700
tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaaat      5760
aaggcgaca cggaatgtt gaatactcat actcttcctt tttcaatatt attgaagcat      5820
ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga aaaataaaca      5880
aataggggtt ccgcgcacat ttccccgaaa agtgccacct gacgcgccct gtagcggcgc      5940
attaagcgcg gcgggtgtgg tggttacgcg cagcgtgacc gctacacttg ccagcgccct      6000
agcgcgccgt cctttcgctt tcttccttc ctttctcgcc acgttcgccg gctttccccg      6060
tcaagctcta aatcgggggc tccctttagg gttccgattt agtgctttac ggcacctcga      6120
ccccaaaaaa cttgattagg gtgatggttc acgtagtggg ccatcgccct gatagacggt      6180
ttttcgccct ttgacgttgg agtccacgtt cttaatagt ggactcttgt tccaaactgg      6240
aacaacactc aaccctatct cggctctattc ttttgattta taagggattt tgccgatttc      6300
ggcctattgg ttaaaaaatg agctgattta acaaaaattt aacgcgaatt tt      6352

```

<210> 15
 <211> 6658
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

```

<400> 15
aacaaaatat taacgcttac aatttcatt cgccattcag gctgcgcaac tgttggaag      60
ggcgcgctgt gcgggcctct tcgctattac gccagctggc gaaaggggga tgtgctgcaa      120
ggcgattaag ttgggtaacg ccagggtttt ccagtcacg acgttgtaaa acgacggcca      180
gtgccaagct gatctataca ttgaatcaat attggcaatt agccatatta gtcattgggt      240
atatagcata aatcaatatt ggctattggc cattgcatac gttgtatcta tatcataata      300

```

NaDoc 002-PCT-1_seq listing_ST25.txt

tgtacattta tattggctca tgtccaatat gaccgccatg ttgacattga ttattgacta	360
gttattaata gtaatcaatt acgggggtcat tagttcatag cccatatatg gagttccgcg	420
ttacataact tacggtaaatt ggcccgccctg gctgaccgcc caacgacccc cgccattga	480
cgtaataat gacgtatggt cccatagtaa cgccaatagg gactttccat tgacgtcaat	540
gggtggagta ttacggtaa actgcccact tggcagtaca tcaagtgtat catatgccaa	600
gtccgcccc tattgacgtc aatgacggt aatggccgc ctggcattat gccagtaca	660
tgacctacg ggactttcct acttggcagt acatctacgt attagtcac gctattacca	720
tggatgacg gttttggcag tacaccaatg ggcgtggata gcggtttgac tcacggggat	780
ttccaagtct ccacccatt gacgtcaatg ggagtttgtt ttggcacc aaatcaacggg	840
actttccaaa atgtcgtaat aaccccgccc cgttgacgca aatggcggt aggcgtgtac	900
ggtgggaggt ctatataagc agagctcgt tagtgaaccg tcagaatttt gtaatacgac	960
tcactatagg gcggccggga attcgtcgac tggatccggt accatgggag ccgacgatgt	1020
ggtcgattct tcgaaatctt ttgtcatgga aaacttttct tcgtaccacg ggacgaaacc	1080
tggttatgtg gattccattc aaaaaggcat acaaaagcca aaatctggta cacaaggaaa	1140
ctacgacgac gattggaaag ggttttacag taccgacaac aaatacgacg ctgagggata	1200
ctctgtggac aacgaaaacc cgctctctgg aaaagctgga ggcgtggta aagtcacgta	1260
tccaggtgag tctctagccc tgcctttgcc tgcctctca gcacttccat tagccagcta	1320
cctacttcca tccactcca aacttcagg ctctgcctgc cccagaggc acaggactta	1380
gttctgggac caggatcag gccgcagccc tggcctgctg ttgcttctgt cagggacttg	1440
cctttgaccc cagcctctct gaccctcagg gtctccttgg ggagctcttc tgaatttggg	1500
ctggcagata cccacccag accaggtctg ccggtgcggc agggccagtg gggcaggtg	1560
gctgtggctg ctgtgcccta gtctgccctt tctgacttgc agggctcacg aaggttctcg	1620
cactcaaggt ggacaatgcc gaaactatca agaaagagt ggggtctcagc ctcaccgaac	1680
cgctcatgga gcaagtcgga acggaagagt ttatcaaaag attcggatg ggtgcttcgc	1740
gtgtagtgct cagccttccc ttcgctgagg ggagttctag cgttgagtac atcaacaact	1800
gggaacaggc gaaagcgta agcgtagaac ttgagattaa ctttgaaacc cgtggaaaac	1860
gtggccaaga tgcgatgtat gagtatatgg ctcaagcctg tgcaggaaat cgtgtcaggc	1920
gatagtgagc ggccgcactc gaggtttaaa cgcccgccg cggtcatagc tgtttctga	1980
acagatccc ggtggcatcc ctgtgacccc tccccagtgc ctctcctggc cctggaagtt	2040
gccactccag tgcccaccag ccttgtccta ataaaattaa gttgcatcat tttgtctgac	2100
taggtgtcct tctataatat tatgggtgg aggggggtgg tatggagcaa ggggcaagtt	2160
gggaagacaa cctgtagggc ctgcggggtc tattgggaac caagctggag tgcagtggca	2220
caatcttggc tcaactgaat ctccgctcc tgggttcaag cgattctcct gcctcagcct	2280
cccagttgt tgggattcca ggcattcatg accaggctca gctaattttt gtttttttgg	2340

NaDoc 002-PCT-1_seq listing_ST25.txt

tagagacggg gtttcacat attggccacg ctggtctcca actcctaate tcaggatgac	2400
taccacattt ggcctcccaa attgctggga ttacaggcgt gaaccactgc tcccttcctt	2460
gtccttctga ttttaaaata actataccag caggaggacg tccagacaca gcataggcta	2520
cctggccatg cccaaccggt gggacatttg agttgcttgc ttggcactgt cctctcatgc	2580
gttgggtcca ctcatagat gcctgttgaa ttgggtacgc ggccagcttg gctgtggaat	2640
gtgtgtcagt taggggtgtg aaagtcccca ggctcccgag caggcagaag tatgcaaagc	2700
atgcatctca attagtcagc aaccagggtg ggaaagtccc caggctcccc agcaggcaga	2760
agtatgcaaa gcatgcatct caattagtca gcaaccatag tcccgcccc aactccgccc	2820
atcccccccc taactccgcc cagttccgcc cattctccgc cccatggctg actaattttt	2880
tttatttatg cagaggccga ggccgcctcg gcctctgagc tattccagaa gtagtgagga	2940
ggcttttttg gaggcctagg cttttgcaaa aagctcccg gagcttgat atccattttc	3000
ggatctgac aagagacacg tacgaccatg gagagcgacg agagcggcct gcccggcatg	3060
gagatcgagt gccgcatcac cggcaccctg aacggcggtg agttcgagct ggtgggcggc	3120
ggagagggca ccccgagca gggccgcatg accaacaaga tgaagagcac caaaggcgcc	3180
ctgaccttca gcccctacct gctgagccac gtgatgggt acggcttcta ccacttcggc	3240
acctaccca gcggctacga gaacccctt ctgcacgcca tcaacaacgg cggtacacc	3300
aacacccgca tcgagaagta cgaggacggc ggcgtgctgc acgtgagctt cagctaccgc	3360
tacgaggccg gccgcgtgat cggcgacttc aagggtgatg gcaccggctt ccccgaggac	3420
agcgtgatct tcaccgacaa gatcatccgc agcaacgcca ccgtggagca cctgcacccc	3480
atgggcgata acgatctgga tggcagcttc acccgcacct tcagcctgcg cgacggcggc	3540
tactacagct ccgtgggtga cagccacatg cacttcaaga gcgccatcca cccagcatc	3600
ctacagaacg ggggccccat gttcgcttc cgccgcgtg aggaggatca cagcaacacc	3660
gagctgggca tcgtggagta ccagcacgcc ttcaagaccc cggtatgaga tgccggtgaa	3720
gaataactgc agcgggactc tgggggttcga aatgaccgac caagcgacgc ccaacctgcc	3780
atcacgagat ttcgattcca ccgccgctt ctatgaaagg ttgggcttcg gaatcgtttt	3840
ccgggacgcc ggctggatga tcctccagcg cggggatctc atgctggagt tcttcgccc	3900
ccccacttg tttattgcag cttataatgg ttacaaataa agcaatagca tcacaaattt	3960
cacaaataaa gcattttttt cactgcattc tagttgtggt ttgtccaaac tcatcaatgt	4020
atcttatcat gtctgtatac cgtcgacctc tagctagagc ttggcgtaat catggtcata	4080
gctgtttcct gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag	4140
cataaagtgt aaagcctggg gtgcctaata agtgagctaa ctacattaa ttgcgttgcg	4200
ctcactgccc gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca	4260
acgcgcgggg agaggcgggt tgcgtattgg gcgctcttcc gcttctcgc tcaactgactc	4320
gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatagc	4380

NaDoc 002-PCT-1_seq listing_ST25.txt

gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa	4440
ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccctga	4500
cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag	4560
ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct	4620
taccggatac ctgtccgcct ttctcccttc gggagcgtg gcgctttctc atagctcacg	4680
ctgtaggtat ctgagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc	4740
ccccgttcag cccgaccgct gcgccttatt cggttaactat cgtcttgagt ccaacccggt	4800
aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta	4860
tgtaggcggg gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac	4920
agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc	4980
ttgatccggc aaacaaacca ccgctggtag cgggtggtttt tttgtttgca agcagcagat	5040
tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc	5100
tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt	5160
cacctagatc cttttaaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta	5220
aacttgggtc gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct	5280
atttcgttca tccatagttg cctgactccc cgtcgtgtag ataactacga tacgggaggg	5340
cttaccatct ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga	5400
tttatcagca ataaaccagc cagccggaag ggccgagcgc agaagtgggc ctgcaacttt	5460
atccgcctcc atccagtcta ttaattgttg ccgggaagct agagtaagta gttcgccagt	5520
taatagtttg cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt	5580
tggtatggct tcattcagct ccggttccca acgatcaagg cgagttacat gatccccat	5640
gttgtgcaaa aaagcgggta gtccttcggt tcctccgatc gttgtcagaa gtaagttggc	5700
cgcagtgtta tcaactcatg ttatggcagc actgcataat tctcttactg tcatgccatc	5760
cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat	5820
gcggcgaccg agttgctctt gcccggcgtc aatacgggat aataccgcgc cacatagcag	5880
aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt	5940
accgctgttg agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc	6000
ttttactttc accagcgttt ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa	6060
gggaataagg gcgacacgga aatgttgaat actcatactc ttcctttttc aatattattg	6120
aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa	6180
taaacaaata ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg cgccctgtag	6240
cggcgcatta agcgcggcgg gtgtggtggt tacgcgcagc gtgaccgcta cacttgccag	6300
cgccctagcg cccgctcctt tcgctttctt cccttccttt ctcgccacgt tcgcccgtt	6360
tccccgtcaa gctctaaatc gggggctccc tttagggttc cgatttagtg ctttacggca	6420

NaDoc 002-PCT-1_seq listing_ST25.txt

cctcgacccc aaaaaacttg attaggggtga tggttcacgt agtgggcat cgccctgata	6480
gacggttttt cgccctttga cggttgagtc cacgttcttt aatagtggaac tcttggtcca	6540
aactggaaca acactcaacc ctatctcggt ctattctttt gatttataag ggattttgcc	6600
gatttcggcc tattgggttaa aaaatgagct gatttaacaa aaatttaacg cgaatttt	6658

<210> 16
 <211> 6964
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 16	
aacaaaatat taacgcttac aatttccatt cgccattcag gctgcgcaac tgttggaag	60
ggcgatcggt gcgggcctct tcgctattac gccagctggc gaaaggggga tgtgctgcaa	120
ggcgattaag ttgggtaacg ccagggtttt ccagtcacg acgttgtaaa acgacggcca	180
gtgccaagct gatctataca ttgaatcaat attggcaatt agccatatta gtcattgggt	240
atatagcata aatcaatatt ggctattggc cattgcatac gttgtatcta taccataata	300
tgtacattta tattggctca tgtccaatat gaccgccatg ttgacattga ttattgacta	360
gttattaata gtaatcaatt acggggctcat tagttcatag cccatatatg gagttccgcg	420
ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc cgccattga	480
cgtcaataat gacgtatggt cccatagtaa cgccaatagg gactttccat tgacgtcaat	540
gggtggagta ttacggtaa actgcccact tggcagtaca tcaagtgtat catatgccaa	600
gtccgcccc tattgacgtc aatgacggta aatggccgc ctggcattat gccagtaca	660
tgaccttacg ggactttcct acttggcagt acatctacgt attagtcac gctattacca	720
tgggtgatgcg gttttggcag tacaccaatg ggcgtggata gcggtttgac tcacggggat	780
ttccaagtct ccacccatt gacgtcaatg ggagtttggt ttggcaccaa aatcaacggg	840
actttccaaa atgtcgtaat aaccccgccc cggtgacgca aatgggcggg aggcgtgtac	900
ggtgggaggt ctatataagc agagctcggt tagtgaaccg tcagaatttt gtaatacgac	960
tcactatagg gcggccggga attcgtcgac tggatccggg accgaggaga tctgccgccg	1020
cgatcgccgg aagcgaatgg gagccgacga tgtggtcgat tcttcgaaat cttttgtcat	1080
ggaaaacttt tcttcgtacc acgggacgaa acctggttat gtggattcca ttcaaaaagg	1140
taggtttaat gttcgtaga tatagttgca gtttctaaca aacatcaaaa ctgattatgc	1200
ttagggtttt tctttttatt tttaacagg catacaaaag ccaaaatctg gtacacaagg	1260
aaactacgac gacgattgga aaggtgaggc actcaggggtg caggacttgg actataaacc	1320
caatggagaa gatagccctt caacctctgt gacttttcta aagctacttt cccccctttt	1380
tgccctaggg ttttacagta ccgacaacaa atacgacgct gcgggatact ctgtggacaa	1440
cgaaaacccg ctctctggaa aagctggagg cgtgggtcaaa gtcacgtatc cagggtcaag	1500

NaDoc 002-PCT-1_seq listing_ST25.txt

gaaataaatt tttagaatcc atttatttgt actgaagtaa aagttcacat atgcaacttc	1560
tatttaatat gtttaacttca caaacctatt ctgtaccata gggctcacga aagttctcgc	1620
actcaaagtg gacaatgccg aaactatcaa gaaagagttg ggtctctctc tcaccgaacc	1680
gctcatggag caagtcggaa cggaagagtt tatcaaaaga ttcggcgatg gtgcttcgcg	1740
tgctgtgctc agccttcctc tcgccgaggg gagttccagc gtcgagtaca tcaacaactg	1800
ggaacaggta tgaatgcaat tggtggcatc tttttttaa gttatgttta agatatgaag	1860
ttaaaattat tttcaaatct gtagttaggc tagtcattaa aactttttcc aggtcagaac	1920
ttacgacctg cttttatttc caaataggcg aaagcgctca gcgtcgaact cgagatcaac	1980
ttcgaaacct gtggaaaacg tggccaagat gcgatgtacg agtatatggc tcaagcctgt	2040
gcaggtgggc agctcatgag cccaggagat tctgtcttgt ttctgtgcct agtggagttt	2100
gttagtttgc tgtgattagc tggcaacgga aactggattc atgttgaga gggtttttct	2160
catctgggta ttcttggttt tccacttaca ctttccccgt cttttctgta ggaaatcgtg	2220
tcaggcgata gtgagcggcc gcactcgagg tttaaacggc cggccgcggc catagctgtt	2280
tcctgaacag atcccgggtg gcattccctgt gacctctccc cagtgcctct cctggccctg	2340
gaagttgcca ctccagtgcc caccagcctt gtcctaataa aattaagttg catcattttg	2400
tctgactagg tgtccttcta taatattatg ggggtggagg ggggtggtatg gagcaagggg	2460
caagttggga agacaacctg tagggcctgc ggggtctatt ggggaaccaag ctggagtga	2520
gtggcacaat cttggctcac tgcaatctcc gcctcctggg ttcaagcgat tctcctgcct	2580
cagcctcccc agttgttggg attccaggca tgcattacca ggctcagcta atttttgttt	2640
ttttggtaga gacgggggtt caccatattg gccacgctgg tctccaactc ctaatctcag	2700
gtgatctacc caccttgccc tcccaaattg ctgggattac aggcgtgaac cactgtctcc	2760
ttccctgtcc ttctgatttt aaaataacta taccagcagg aggacgtcca gacacagcat	2820
aggctacctg gccatgccc accggtggga catttgagtt gcttgcttgg cactgtcctc	2880
tcatgcgttg ggtccactca gtagatgcct gttgaattgg gtacgcggcc agcttggtg	2940
tggaatgtgt gtcagttagg gtgtggaaag tccccaggct cccagcagg cagaagtatg	3000
caaagcatgc atctcaatta gtcagcaacc aggtgtggaa agtccccagg ctccccagca	3060
ggcagaagta tgcaaagcat gcattctaat tagtcagcaa ccatagtccc gcccctaact	3120
ccgccccatc cggcccctaac tccgcccagt tccgcccatt ctccgcccc tggctgacta	3180
atttttttta ttatgcaga ggccgaggcc gcctcggcct ctgagctatt ccagaagtag	3240
tgaggaggct tttttggagg cctaggcttt tgcaaaaagc tcccgggagc ttgtatatcc	3300
attttcggat ctgatcaaga gacacgtacg accatggaga gcgacgagag cggcctgccc	3360
gccatggaga tcgagtgccg catcaccggc accctgaacg gcgtggagtt cgagctgggtg	3420
ggcggcgag agggcacccc cgagcagggc cgcatgacca acaagatgaa gagcaccaaa	3480
ggcgccctga ccttcagccc ctacctgtg agccacgtga tgggctacgg cttctaccac	3540

NaDoc 002-PCT-1_seq listing_ST25.txt

ttcggcacct accccagcgg ctacgagaac cccttcctgc acgccatcaa caacggcggc	3600
tacaccaaca cccgcatcga gaagtacgag gacggcggcg tgctgcacgt gagcttcagc	3660
taccgctacg aggccggccg cgtgatcggc gacttcaagg tgatgggcac cggcttcccc	3720
gaggacagcg tgatcttcac cgacaagatc atccgcagca acgccaccgt ggagcacctg	3780
caccccatgg gcgataacga tctggatggc agcttcaccc gcaccttcag cctgcgcgac	3840
ggcggctact acagctccgt ggtggacagc cacatgcact tcaagagcgc catccacccc	3900
agcatcctac agaacggggg ccccatgttc gccttcgcc gcgtggagga ggatcacagc	3960
aacaccgagc tgggcatcgt ggagtaccag cacgccttca agaccccgga tgcagatgcc	4020
ggtgaagaat aactgcagcg ggactctggg gttcgaaatg accgaccaag cgacgccccaa	4080
cctgccatca cgagatttcg attccaccgc cgccttctat gaaagggttg gcttcggaat	4140
cgttttccgg gacgccggct ggatgatcct ccagcgcggg gatctcatgc tggagttctt	4200
cgcccccccc aacttgttta ttgcagctta taatggttac aaataaagca atagcatcac	4260
aaatttcaca aataaagcat ttttttact gcattctagt tgtggtttgt ccaaactcat	4320
caatgtatct tatcatgtct gtataccgtc gaccttagc tagagcttgg cgtaatcatg	4380
gtcatagctg tttcctgtgt gaaattgtta tccgctcaca attccacaca acatacgagc	4440
cggaagcata aagtgtaaag cctgggggtgc ctaatgagtg agctaactca cattaattgc	4500
gttgcgctca ctgcccgtt tccagtcggg aaacctgtcg tgccagctgc attaatgaat	4560
cggccaacgc gcggggagag gcggtttgcg tattgggcgc tcttccgtt cctcgtcac	4620
tgactcgctg cgctcggtcg ttcggctgcg gcgagcggtg tcagctcact caaaggcggt	4680
aatacgggta tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca	4740
gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc	4800
ccctgacgag catcacaaaa atcgacgtc aagtcagagg tggcgaaacc cgacaggact	4860
ataaagatac caggcgtttc cccctggaag ctccctcgtg cgctctcctg ttccgacctt	4920
gccgcttacc ggatacctgt cgcctttct ccttcggga agcgtggcg tttctcatag	4980
ctcacgctgt aggtatctca gttcgggtga ggtcgttcgc tccaagctgg gctgtgtgca	5040
cgaaccccc gttcagcccc accgctgcgc cttatccggt aactatcgtc ttgagtccaa	5100
cccggtaaga cagcacttat cgccactggc agcagccact ggtaacagga ttagcagagc	5160
gaggtatgta ggcggtgcta cagagttctt gaagtgggtg cctaactacg gctacactag	5220
aagaacagta tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg	5280
tagctcttga tccggcaaac aaaccaccgc tggtagcggg ggtttttttg tttgcaagca	5340
gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc	5400
tgacgctcag tggaacgaaa actcacgtta agggattttg gtcatgagat tatcaaaaag	5460
gatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata	5520
tgagtaaact tgggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat	5580

NaDoc 002-PCT-1_seq listing_ST25.txt

ctgtctatTT cgttcatcca tagttgcctg actccccgtc gtgtagataa ctacgatacg 5640
 ggagggcTTa ccatctggcc ccagtgtgtc aatgataccg cgagaccac gctcaccggc 5700
 tccagattTa tcagcaataa accagccagc cggaagggcc gagcgcagaa gtggtcctgc 5760
 aactttatcc gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc 5820
 gccagttaat agtttgcgca acgttggtgc cattgctaca ggcatcgtgg tgtcacgctc 5880
 gtcgtttggT atggcttcat tcagctccgg ttcccaacga tcaaggcgag ttacatgatc 5940
 ccccatgttg tgcaaaaaag cggttagctc cttcggctct ccgatcgttg tcagaagtaa 6000
 gttggccgca gtgttatcac tcatggttat ggcagcactg cataattctc ttactgtcat 6060
 gccatccgTa agatgctttt ctgtgactgg tgagtactca accaagtcac tctgagaata 6120
 gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaata cgggataata ccgcgccaca 6180
 tagcagaact ttaaaagtgc tcatcattgg aaaacgttct tcggggcgaa aactctcaag 6240
 gatcttaccg ctgttgagat ccagttcgat gtaaccact cgtgcacca actgatcttc 6300
 agcatctttt actttcacca gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc 6360
 aaaaaaggga ataagggcga cacggaaatg ttgaatactc atactcttcc ttttcaata 6420
 ttattgaagc atttatcagg gttattgtct catgagcgga tacatatttg aatgtattta 6480
 gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgccac ctgacgcgcc 6540
 ctgtagcggc gcattaagcg cggcgggtgt ggtggttacg cgcagcgtga ccgctacact 6600
 tgccagcgcc ctagcgccc ctcctttcgc tttcttcct tcctttctcg ccacgttcgc 6660
 cggctttccc cgtcaagctc taaatcgggg gctcccttta gggttccgat ttagtgcttt 6720
 acggcacctc gaccccaaaa aacttgatta ggggtgatgg tccagtagtg ggccatcgcc 6780
 ctgatagacg gtttttcgcc ctttgacgtt ggagtccacg ttctttaata gtggactctt 6840
 gttccaaact ggaacaacac tcaacctat ctcggtctat tcttttgatt tataagggat 6900
 tttgccgatt tcggcctatt ggttaaaaaa tgagctgatt taacaaaaat ttaacgcgaa 6960
 tttt 6964

<210> 17
 <211> 7733
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 17
 agatctgcgc agcaccatgg cctgaaataa cctctgaaag aggaacttgg ttaggtacct 60
 tctgaggcgg aaagaaccag ctgtggaatg tgtgtcagtt aggggtgtgga aagtccccag 120
 gctccccagc aggcagaagt atgcaaagca tgcatctcaa ttagtcagca accaggtgtg 180
 gaaagtcccc aggtcccca gcaggcagaa gtatgcaaag catgcatctc aattagtcag 240
 caaccatagt cccgccccta actccgcca tcccgcccct aactccgcc agttccgccc 300

NaDoc 002-PCT-1_seq listing_ST25.txt

attctccgcc ccatggctga ctaatttttt ttatttatgc agaggccgag gccgcctcgg	360
cctctgagct attccagaag tagtgaggag gcttttttgg aggcctaggc ttttgcaaaa	420
agcttgattc ttctgacaca acagtctcga acttaagctg cagaagttgg tcgtgaggca	480
ctgggcaggt aagtatcaag gttacaagac aggtttaagg agaccaatag aaactgggct	540
tgtcgagaca gagaagactc ttgcgtttct gataggcacc tattggtctt actgacatcc	600
actttgcctt tctctccaca ggtgtccact cccagttcaa ttacagctct taaggctaga	660
gtacttaata cgactcacta taggctagac gcggtaccta gctaggtagc aattgaccgg	720
tcaagatggc ggccaacaac aacaacaaca acaacaaca caacaacaac aacaacaaca	780
agaagatggc ggcaacaaca acaacaaca caacaacaac aacaacaaca acaaccaaca	840
acaagatggc ggccaacaac aacaacaaca acaacaaca gaagatggcg gcaacaaca	900
caacaacaac aacaacaacc aagatggcgg ccaacaaca caagaagatg gcggcaaca	960
caacaacca gatggcggcc aacaacaaca agaagatggc ggcaacaaca acaaccaaga	1020
tggcggcacg cgtcggtcgg gctagccgta cgctccttag cgacgaaatc tactgcccc	1080
ctgagagcca ccatggcttg gggtcctacg ctgtgcaggc caagtttggg gattacaaca	1140
aagaaggccg ccatgggtgg cacctcagct ctgagcggct catccgccac catggggttg	1200
accagcaca acttaccagg gaccgccgcc atggccggac ccaggcgtgc caccatggac	1260
accgtgggtt gcgccgccat ggtgctctgt tggagtggca ccatggtgct caggacctgg	1320
gccgccatgg aatacctgat aactgataag ccaccatggg aacagacctt tggcttggag	1380
ttgacgccct tggactcaac atttacgagg ccgccatgga gttcaccca aagattggct	1440
ttccttggag tgaaatcagg aacatctctg ccaccatgga aaagtgtgtc atcaagccca	1500
tcgacaaggc cgccatggac tttgtgtttt acgccccacg tctcacagcc accatgggga	1560
ccctgcagct cgccgccatg gaccacgagt tgtacgccac catggggaag cctgacaccg	1620
ccgccatgga gcagacgaag gccgccacca tggaggctga taagctgata agccgccatg	1680
ggctggaaac agagaagaaa aggagagaaa ccgtggagag agagaaagag cgccaccatg	1740
gcgagaagga ggagttgttg ctgaggctgc aggactacga ggagaagaca agccgccatg	1800
ggagagacct ctcgagcag attcagaggg gccaccatgg ggaggaggag aggaagcggg	1860
cacaggaggg ccgccatggc ccagaggctg accgccacca tggactgcgg gctaagggcc	1920
gccatgggag acaggcgggtg ggccaccatg ggagccagga gcagcgccgc catgggctac	1980
ctgataactg ataagccacc atggtggaag aggcgcggag gcgcaaggag gacgaagttg	2040
aagagtggca gcaagccgcc atggaagccc aggacgacct ggtcaagacc aaggaggagc	2100
tgcacctggt gccggccacc atggcgccac caccaccacc cgtgtacgag ccggccgcca	2160
tggacgtcca ggagagcttg caagacgagg gtgccaccat ggcgggctac agcgagccg	2220
ccatggctga cggcatccgg gccaccatgg acgaggagaa gcgtgccgcc atggcagaga	2280
agaacgaggg caccatgggg cctgataagc tgataagccg ccatggggcc cgagacgaga	2340

NaDoc 002-PCT-1_seq listing_ST25.txt

acaagaggac ccacaacgac atcatccaca acgagagcca ccatggaggc cgggacaagt	2400
acaagacgct gcggcagatc cggcagggca acaccagccg ccatggcgac gagttcgagg	2460
ccctgcaaca gccaggccac catggagggc agaggggtgc tcatagcggg cgctgccgcc	2520
atggccacgc ttgtgtctgc caccatggaa gtctcggaac tcgccgccat ggcagttcct	2580
ttcgaagcca ccatggcaac agaaacattc gccgccatgg accacctgat aactgataag	2640
ccaccatggt tgcaatcgtg ccaagcaggc ctgattctcg cgattactcg cgaatcacccg	2700
ccgccatggt gctgggagca ggactcattg aattacggaa aacgcctgtc aagtctcagg	2760
ccaccatggg gaactggcct gtgtcataca agagtcaggc cgccatgggg aaacgtggca	2820
ggacttccat ctgtgccgcc accatggtgt attcgaaacg agccgccatg gattttctca	2880
tctctgccac catggcatct ttgtacattg ccgccatggg aggggtcaaa attgccacca	2940
tgggtggctga taagttgata gtaaccgcca tgggttttca tccagtcgcc accatgggct	3000
ggcagagagc agccgccatg gcagcgtcag tgggtggccac catggccttg atttttttt	3060
ttgtttttt ttttttgct caacaatttt acaacacatt gtgtcgagcc cgggaattcg	3120
tttaaaccta gagcggccgc tggccgcaat aaaatatctt tattttcatt acatctgtgt	3180
gttggttttt tgtgtgagga tctaaatgag tcttcggacc tcgcgggggc cgcttaagcg	3240
gtggttaggg tttgtctgac gcggggggag ggggaaggaa cgaaacactc tcattcggag	3300
gcggctcggg gtttggctct ggtggccacg ggcacgcaga agagcggcgc gatcctctta	3360
agcaccctccg cgccctccgt ggaggcgggg gtttggctcg cgggtggtaa ctggcgggcc	3420
gctgactcgg gcgggtcgcg cgcccagag tgtgacctt tcggtctgct cgcagacccc	3480
cgggcggcgc cgccgcggcg gcgacgggct cgctgggtcc taggctccat ggggaccgta	3540
tacgtggaca ggctctggag catccgcacg actgcggtga tattaccgga gaccttctgc	3600
gggacgagcc gggtcacgcg gctgacgcg agcgtccgtt gggcgacaaa caccaggacg	3660
gggcacaggt acactatctt gtcacccgga ggcgcgagg actgcaggag cttcagggag	3720
tggcgcagct gcttcatccc cgtggcccgt tgctcgcgtt tgctggcgtt gtccccggaa	3780
gaaatatatt tgcattgtct tagttctatg atgacacaaa ccccgcccag cgtcttgtca	3840
ttggcgaatt cgaacacgca gatgcagtcg gggcggcgcg gtcccaggtc cacttcgcat	3900
attaaggtga cgcgtgtggc ctcgaacacc gagcgaccct gcagcgaccc gcttaaaagc	3960
ttggcattcc ggtactgttg gtaaagccac catggccgat gctaagaaca ttaagaagg	4020
ccctgctccc ttctaccctc tggaggatgg caccgctggc gagcagctgc acaaggccat	4080
gaagaggtat gccctggtgc ctggcaccat tgccttcacc gatgcccaca ttgaggtgga	4140
catcacctat gccgagtact tcgagatgtc tgtgcgctg gccgaggcca tgaagaggt	4200
cggcctgaac accaaccacc gcatcgtggt gtgctctgag aactctctgc agttcttcat	4260
gccagtgtcgg ggcgcctgt tcatcggagt ggccgtggcc cctgctaacg acatttaca	4320
cgagcgcgag ctgctgaaca gcatgggcat ttctcagcct accgtggtgt tcgtgtctaa	4380

NaDoc 002-PCT-1_seq listing_ST25.txt

gaagggcctg cagaagatcc tgaacgtgca gaagaagctg cctatcatcc agaagatcat	4440
catcatggac tctaagaccg actaccaggg cttccagagc atgtacacat tcgtgacatc	4500
tcattctgcct cctggcttca acgagtacga cttcgtgccg gagtcttttcg acagggacaa	4560
aaccattgcc ctgatcatga acagctctgg gtctaccggc ctgcctaagg gcgtggccct	4620
gcctcatcgc accgcctgtg tgcgcttctc tcacgcccgc gaccctatctt tcggcaacca	4680
gatcatcccc gacaccgcta ttctgagcgt ggtgccattc caccacggct tcggcatgtt	4740
caccaccctg ggctacctga tttgcggctt tcgggtggtg ctgatgtacc gcttcgagga	4800
ggagctgttc ctgctcagcc tgcaagacta caaaattcag tctgccctgc tgggtgccaac	4860
cctgttcagc ttcttcgcta agagcaccct gatcgacaag tacgacctgt ctaacctgca	4920
cgagattgcc tctggcggcg cccactgtc taaggaggtg ggcgaagccg tggccaagcg	4980
ctttcatctg ccaggcatcc gccagggcta cggcctgacc gagacaacca gcgccattct	5040
gattacccca gagggcgacg acaagcctgg cgccgtgggc aagggtggtgc cattcttcga	5100
ggccaaggtg gtggacctg acaccggcaa gaccctggga gtgaaccagc gcggcgagct	5160
gtgtgtgctg gccctatga ttatgtccgg ctacgtgaat aaccctgagg ccacaaacgc	5220
cctgatcgac aaggacggct ggctgcactc tggcgacatt gcctactggg acgaggacga	5280
gcacttcttc atcgtggacc gcctgaagtc tctgatcaag tacaagggt accaggtggc	5340
cccagccgag ctggagtcta tctgtctgca gcaccctaac attttcgacg ccggagtggc	5400
cggcctgccc gacgacgatg ccggcgagct gcctgccgcc gtcgtcgtgc tggaacacgg	5460
caagaccatg accgagaagg agatcgtgga ctatgtggcc agccaggtga caaccgcaa	5520
gaagctgctg ggcggagtgg tgctcgtgga cgaggtgccc aagggcctga ccggcaagct	5580
ggacgcccgc aagatccgcg agatcctgat caaggctaag aaaggcggca agatcgccgt	5640
gtaataattc tagagtcggg gcggccggcc gcttcgagca gacatgataa gatacattga	5700
tgagtttgga caaaccacaa ctagaatgca gtgaaaaaaa tgctttatctt gtgaaatttg	5760
tgatgctatt gctttatctt taaccattat aagctgcaat aaacaagtta acaacaacaa	5820
ttgcattcat tttatgtttc aggttcaggg ggaggtgtgg gaggtttttt aaagcaagta	5880
aaacctctac aaatgtggta aaatcgataa ggatccaggt ggcacttttc ggggaaatgt	5940
gcgcggaacc cctatttggt tttttttcta aatacattca aatatgtatc cgctcatgag	6000
acaataaccc tgataaatgc ttcaataata ttgaaaaagg aagagtatga gtattcaaca	6060
tttcctgtgc gcccttattc ctttttttgc ggcattttgc cttcctgttt ttgctcacc	6120
agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcacgag tgggttacat	6180
cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgccccgaag aacgttttcc	6240
aatgatgagc acttttaaag ttctgctatg tggcgcggtg ttatcccgtg ttgacgccgg	6300
gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttggttg agtactcacc	6360
agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca gtgctgccat	6420

NaDoc 002-PCT-1_seq listing_ST25.txt

aaccatgagt gataaactg cggccaactt acttctgaca acgatcggag gaccgaagga	6480
gctaaccgct tttttgcaca acatggggga tcatgtaact cgccttgatc gttgggaacc	6540
ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg tagcaatggc	6600
aacaacgttg cgaaactat taactggcga actacttact ctagcttccc ggcaacaatt	6660
aatagactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg cccttccggc	6720
tggctggttt attgctgata aatctggagc cggtgagcgt gggctctcgc gtatcattgc	6780
agcactgggg ccagatggta agccctcccg tatcgtagtt atctacacga cggggagtca	6840
ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac tgattaagca	6900
ttggtaactg tcagaccaag tttactcata tatacttttag attgatttaa aacttcattt	6960
ttaatttaaa aggatctagg tgaagatcct ttttgataat ctcatgacca aaatccctta	7020
acgtgagttt tcgttccact gagcgtcaga ccccgtagaa aagatcaaag gatcttcttg	7080
agatcctttt tttctgcgcg taatctgctg cttgcaaaca aaaaaaccac cgctaccagc	7140
ggtggtttgt ttgccggatc aagagctacc aactcttttt ccgaaggtaa ctggcttcag	7200
cagagcgcag ataccaaata ctgttcttct agtgtagccg tagttaggcc accacttcaa	7260
gaactctgta gcaccgccta catacctcgc tctgctaata ctgttaccag tggctgctgc	7320
cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga cgatagttac cggataaggc	7380
gcagcggctg ggctgaacgg ggggttcgtg cacacagccc agcttgagc gaacgaccta	7440
caccgaactg agatacctac agcgtgagct atgagaaagc gccacgcttc ccgaaggag	7500
aaaggcggac aggtatccgg taagcggcag ggtcggaaac ggagagcgca cgaggagct	7560
tccaggggga aacgcctggt atctttatag tcctgtcggg tttcgccacc tctgacttga	7620
gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta tggaaaaacg ccagcaacgc	7680
ggccttttta cggttcctgg ccttttgctg gccttttgct cacatggctc gac	7733

<210> 18
 <211> 6679
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 18	
aacaaaatat taacgcttac aatttccatt cgccattcag gctgcgcaac tgttgggaag	60
ggcgatcggg gcgggcctct tcgctattac gccagctggc gaaaggggga tgtgctgcaa	120
ggcgattaag ttgggtaacg ccagggtttt cccagtcacg acgttgtaaa acgacggcca	180
gtgccaaagt gatctataca ttgaatcaat attggcaatt agccatatta gtcattgggt	240
atatagcata aatcaatatt ggctattggc cattgcatac gttgtatcta taccataata	300
tgtacattta tattggctca tgtccaatat gaccgccatg ttgacattga ttattgacta	360
gttattaata gtaatcaatt acgggggtcat tagttcatag cccatatatg gagttccgcg	420

NaDoc 002-PCT-1_seq listing_ST25.txt

ttacataact tacggtaa	ggccgcctg gctgaccgcc	caacgacccc cgcccattga	480
cgtcaataat gacgtatgtt	cccatagtaa cgccaatagg	gactttccat tgacgtcaat	540
gggtggagta ttacggtaa	actgcccact tggcagtaca	tcaagtgtat catatgccaa	600
gtccgcccc tattgacgtc	aatgacggta aatggccgc	ctggcattat gcccagtaca	660
tgaccttacg ggactttcct	acttggcagt acatctacgt	attagtcac gctattacca	720
tggatgatgcg gttttggcag	tacaccaatg ggcgtggata	gcggtttgac tcacggggat	780
ttccaagtct ccacccatt	gacgtcaatg ggagtttggt	ttggcaccaa aatcaacggg	840
actttccaaa atgtcgtaat	aacccgcgcc cgttgacgca	aatgggcggg aggcgtgtac	900
gggtgggaggt ctatataagc	agagctcggt tagtgaaccg	tcagaatttt gtaatacgac	960
tcactatagg gcggccggga	attcgtcgac tggatcttgt	acattcgaac gccgccatgg	1020
gcgctgatga tgttggtgat	tcttctaaat cttttgtcat	ggaaaacttt tcttcgtacc	1080
acgggactaa acctggttat	gtggattcca ttcaaaaagg	tatacaaaag ccaaactctg	1140
gtacacaagg aaattatgac	gatgattgga aagggtttta	tagtaccgac aataaatacg	1200
acgctgcggg atactctgtg	gataatgaaa acccgctctc	tggaaaagct ggaggcgtgg	1260
tcaaagtgac gtatccagga	ctgacgaagg ttctcgact	aaaagtggat aatgccgaaa	1320
ctattaagaa agagttaggt	ttaagtctca ctgaaccgtt	gatggagcaa gtcggaacgg	1380
aagagtttat caaaagggtc	ggtgatgggt cttcgctgtg	agtgtctcagc cttcccttcg	1440
ctgaggggag ttctagcggt	gaatatatta ataactggga	acaggcgaaa gcgttaagcg	1500
tagaacttga gattaatttt	gaaacccgtg gaaaacgtgg	ccaagatgcg atgtatgagt	1560
atatggctca agcctgtgca	ggaaatcgtg tcaggcgata	gtgaactagt tccggatcta	1620
gagcggccgc actcgagggt	taaacggccg gccgcgttca	tagctgtttc ctgaacagat	1680
cccggtggc atccctgtga	cccctcccca gtgcctctcc	tggccctgga agttgccact	1740
ccagtgtcca ccagccttgt	cctaataaaa ttaagttgca	tcattttgtc tgactagggt	1800
tccttctata atattatggg	gtggaggggg gtggtatgga	gcaaggggca agttgggaag	1860
acaacctgta gggcctgcgg	ggtctattgg gaaccaagct	ggagtgcagt ggcacaatct	1920
tggctcactg caatctccgc	ctcctgggtt caagcgattc	tcctgcctca gcctcccag	1980
ttgttgggat tccaggcatg	catgaccagg ctcagcta	at tttgtttt ttggtagaga	2040
cggggtttca ccatattggc	caggctggtc tccaactcct	aatctcaggt gatctacca	2100
ccttggcctc ccaaattgct	gggattacag gcgtgaacca	ctgctccctt ccctgtcctt	2160
ctgattttaa aataactata	ccagcaggag gacgtccaga	cacagcatag gctacctggc	2220
catgcccaac cgggtgggaca	tttgagttgc ttgcttggca	ctgtcctctc atgcgttggg	2280
tccactcagt agatgcctgt	tgaattgggt acgcggccag	cttggctgtg gaatgtgtgt	2340
cagttagggt gtggaaagtc	cccaggctcc ccagcaggca	gaagtatgca aagcatgcat	2400
ctcaattagt cagcaaccag	gtgtggaaag tccccaggct	ccccagcagg cagaagtatg	2460

NaDoc 002-PCT-1_seq listing_ST25.txt

caaagcatgc atctcaatta gtcagcaacc atagtcccg ccctaactcc gcccatcccg	2520
cccctaactc cgcccagttc cgccattct cgccccatg gctgactaat tttttttatt	2580
tatgcagagg ccgaggccgc ctggcctct gagctattcc agaagtagtg aggaggcttt	2640
tttgagggcc taggcttttg caaaaagctc ccgggagctt gtatatccat tttcggatct	2700
gatcaagaga cacgtacgac catgaaaaag cctgaactca ccgcgacgtc tgttgagaag	2760
tttctgatcg aaaagttcga cagcgtctcc gacctgatgc agctctcgga gggcgaagaa	2820
tctcgtgctt tcagcttcga tgtaggagg cggtgatatg tcctgcgggt aaatagctgc	2880
gccgatgggt tctacaaaga tcgttatgtt tatcggcact ttgcatcggc cgcgctcccg	2940
attccggaag tgcttgacat tggggaattt agcgagagcc tgacctattg catctccgc	3000
cgtgcacagg gtgtcacgtt gcaagacctg cctgaaaccg aactgcccgc tgttctgcaa	3060
ccggtcgcgg aggccatgga tgcaatcgct gcggccgatc ttagccagac gagcgggttc	3120
ggccattcg gaccgcaagg aatcgggtaa tacactacat ggcgtgattt catatgcgcg	3180
attgctgatc cccatgtgta tcaactggca actgtgatgg acgacaccgt cagtgcgtcc	3240
gtcgcgcagg ctctcgatga gctgatgctt tgggccgagg actgccccga agtccggcac	3300
ctcgtgcacg cggatttcgg ctccaacaat gtcctgacgg acaatggccg cataacagcg	3360
gtcattgact ggagcgaggc gatgttcggg gattcccaat acgaggtcgc caacatcttc	3420
ttctggaggc cgtggttggc ttgtatggag cagcagacgc gctacttcga gcggaggcat	3480
ccggagcttg caggatcgcc gcggctccgg gcgtatatgc tccgcattgg tcttgaccaa	3540
ctctatcaga gcttggttga cggaatttc gatgatgcag cttgggcgca gggtcgatgc	3600
gacgcaatcg tccgatccgg agccgggact gtcgggcgta cacaatcgc ccgcagaagc	3660
gcggccgtct ggaccgatgg ctgtgtagaa gtactcgccg atagtggaaa ccgacgcccc	3720
agcactcgtc cgagggcaaa ggaatagctg cagcgggact ctggggttcg aaatgaccga	3780
ccaagcgacg cccaacctgc catcacgaga ttctgattcc accgccgctt tctatgaaag	3840
gttgggcttc ggaatcgttt tccgggacgc cggctggatg atcctccagc gcggggatct	3900
catgctggag ttcttcgccc accccaactt gtttattgca gcttataatg gttacaaata	3960
aagcaatagc atcacaatt tcacaaataa agcatttttt tcaactgcatt ctagtgtg	4020
tttgtccaaa ctcatcaatg tatcttatca tgtctgtata ccgtcgacct ctagctagag	4080
cttggcgtaa tcatggtcat agctgtttcc tgtgtgaaat tggtatccgc tcacaattcc	4140
acacaacata cgagccggaa gcataaagtg taaagcctgg ggtgcctaata gaggtagcta	4200
actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc tgtcgtgcca	4260
gctgcattaa tgaatcggcc aacgcgcggg gagaggcggg ttgcgtattg ggcgctcttc	4320
cgcttcctcg ctcaactgact cgctgcgctc ggtcgttcgg ctgcggcgag cggatcagc	4380
tcaactcaag gcggtaatac gggtatccac agaatcaggg gataacgcag gaaagaacat	4440
gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcggtttt	4500

NaDoc 002-PCT-1_seq listing_ST25.txt

ccataggctc cgccccctg acgagcatca caaaaatcga cgctcaagtc agaggtggcg	4560
aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc tcgtgcgctc	4620
tcctgttccg accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt	4680
ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa	4740
gctgggctgt gtgcacgaac cccccgttca gcccgaccgc tgcgccttat ccggttaacta	4800
tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag ccaactggtaa	4860
caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa	4920
ctacggctac actagaagaa cagtatttgg tatctgcgct ctgctgaagc cagttacctt	4980
cggaaaaaga gttggtagct cttgatccgg caaacaacc accgctggta gcggtggttt	5040
ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat	5100
cttttctacg ggtctgacg ctcaagtggaa cgaaaactca cgttaaggga ttttggctcat	5160
gagattatca aaaaggatct tcacctagat ctttttaa ataaaaatgaa gttttaaatc	5220
aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgaaggc	5280
acctatctca gcgatctgtc tttttcgttc atccatagtt gcctgactcc ccgtcgtgta	5340
gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga taccgcgaga	5400
cccacgctca ccggtccag atttatcagc aataaaccag ccagccggaa gggccgagcg	5460
cagaagtggc cctgcaactt tatccgcctc catccagtct attaatgtt gccgggaagc	5520
tagagtaagt agttcgccag ttaatagttt gcgcaacggt gttgccattg ctacaggcat	5580
cgtggtgtca cgctcgtcgt ttggtatggc ttcatcagc tccggttccc aacgatcaag	5640
gcgagttaca tgatccccc tgttgtgcaa aaaagcgggt agctccttcg gtcctccgat	5700
cgttgtcaga agtaagtgg ccgcagtgt atcactcatg gttatggcag cactgcataa	5760
ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt actcaaccaa	5820
gtcattctga gaatagtgtg tgccggcacc gagttgctct tgcccggcgt caatacggga	5880
taataccgcg ccacatagca gaactttaaa agtgctcatc attggaaaac gttcttcggg	5940
gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac ccaactcgtc	6000
acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag caaaaacagg	6060
aaggcaaaat gccgcaaaaa aggaataag ggcgacacgg aaatgttgaa tactcatact	6120
cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga gcggatacat	6180
atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt	6240
gccacctgac gcgccctgta gcggcgcatc aagcgcggcg ggtgtggtgg ttacgcgcag	6300
cgtgaccgct acacttgcca gcgccctagc gcccgctcct ttcgctttct tcccttcctt	6360
tctcgccacg ttcgccggct ttccccgtca agctctaaat cgggggctcc ctttaggggt	6420
ccgatttagt gctttacggc acctcgaccc caaaaaactt gattaggggt atggttcacg	6480
tagtgggcca tcgccctgat agacggtttt tcgccctttg acgttgaggt ccacgttctt	6540

NaDoc 002-PCT-1_seq listing_ST25.txt

taatagtgga ctcttgttcc aaactggaac aacactcaac cctatctcgg tctattcttt 6600
 tgatttataa gggatttttgc cgatttcggc ctattgggta aaaaatgagc tgatttaaca 6660
 aaaatttaac gcgaatttt 6679

<210> 19
 <211> 870
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 19
 cgcggatcca ccggtcaatt gtatcaactc tgagatgcag gtacatccag ctgatgagtc 60
 ccaaatagga cgaaacgcgc ttcggtgcgt cctggattcc actgctatcc actattcatc 120
 tacttgact gcacccgata ccctgtcacc ggatgtgctt tccggtctga tgagtccgtg 180
 aggacgaaac aggactggaa cgtactacga caggaacttg tcctgagatg caggtacatc 240
 ccactgatga gtcccaaata ggacgaaacg cgcttcggtg cgtctgggat tccactgcta 300
 tccacacgcg tcggtccgaa gcttgtcgac cgccggtgca aagatctgaa ttcacctgat 360
 agctgatagc tgatagcccg gggctctctgt ggatagacca gagcggagcc tgggagctct 420
 ctggctatct acggaacca ctgcttaagc ctcaataccg cttgccttga gtgcttcaag 480
 tagtgtgtgc cgaacacaag ctcacgaccc actacacaag ctcacgaccc actacacaag 540
 ctcacgaccc actacacgag cttggggcgc gtgggtggcg ctgcagccgc caccacgcgc 600
 cccggatcgg agattgtcag gagctaagga agctaaacaa cgactacagc aggcttttgc 660
 aaaaagctcc accacggccc aacgttgggc cgtgggtggag cttggattgt acttgactg 720
 catacacaac gagatcgga cgtactacga caggaactac gaccctgcgg tccaccacgg 780
 ccgatatcac ggccgtggtg gaccgcaggg aagaacaacg tctccgatct ttggtaccaa 840
 cacatctaga caaagtactg gcgcgcccaa 870

<210> 20
 <211> 4950
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic polynucleotide

<400> 20
 gcggccgcaa taaaatatct ttattttcat tacatctgtg tgttggtttt ttgtgtgaat 60
 cgtaactaac atacgctctc catcaaaaca aaacgaaaca aaacaaacta gcaaaatagg 120
 ctgtccccag tgcaagtgca ggtgccagaa catttctcta tcgaaggatc tgcgatcgct 180
 ccggtgcccg tcagtgggca gagcgcacat cggccacagt ccccgagaag ttggggggag 240
 gggtcggcaa ttgaacgggt gcctagagaa ggtggcgcg ggtaaaactgg gaaagtgatg 300
 tcgtgtactg gctccgcctt tttcccaggg gtgggggaga accgtatata agtgcagtag 360

NaDoc 002-PCT-1_seq listing_ST25.txt

tcgccgtgaa cgttcttttt cgcaacgggt ttgccgccag aacacagctg aagcttcgag	420
gggctcgcac ctctccttca cgcgccgcc gccctacctg aggccgccat ccacgccggt	480
tgagtcgcgt tctgccgctt cccgcctgtg gtgcctcctg aactgcgtcc gccgtctagg	540
taagtttaaa gctcaggtcg agaccgggcc tttgtccggc gctcccttgg agcctaccta	600
gactcagccg gctctccacg ctttgccctga ccttgcttgc tcaactctac gtctttgttt	660
cgttttctgt tctgcgccgt tacagatcca agctgtgacc ggccgctacc tgagatcacc	720
ggattcgaaa gatctgccac catacgttgc cgcgcagcgg actgcccgcc aggatatgga	780
tcctgatgat gttgttgatt cttctaaatc tttgttatg gaaaactttt cttcgtacca	840
cgggactaaa cctggttatg tggattccat tcaaaaagggt atacaaaagc caaaatctgg	900
taccaagga aattatgacg atgattggaa agggttttat agtaccgaca ataaatacga	960
cgctgcggga tactctgtag ataataaaaa cccgctctct ggaaaagctg gaggcgtggt	1020
caaagtgcg tatccaggac tgacgaaggt tctgcacta aaagtggata atgccgaaac	1080
tattaagaaa gagttagggt taagtctcac tgaaccgttg atggagcaag tcggaacgga	1140
agagtttatc aaaagggtcg gtgatgggtg ttcgcgtgta gtgctcagcc ttcccttcgc	1200
cgaggggagt tctagcgttg aatatattaa taactgggaa caggcgaaag cgtraagcgt	1260
agaacttgag attaatcttg aaaccctgtg aaacgtggc caagatgcga tgtatgagta	1320
tatggctcaa gcctgtgcag gaaatcgtgt caggcgatct ctttgtgaag gaaccttact	1380
tctgtggtgt gacataattg gacaaactac ctacagagat ttaaagctct aatgactcga	1440
gccatgggac ccacactttt ctactgtggc agacatgata agatacattg atgagtttgg	1500
acaaaccaca actagaatgc agtgaaaaaa atgctttatt tgtgaaattt gtgatgctat	1560
tgctttattt gtaaccatta taagctgcaa taaacaagtt aacaacaaca attgcattca	1620
ttttatgttt caggttcagg gggagggtgt ggagggtttt taaagcaagt aaaacctcta	1680
caaatgtggt atggaattct aaaatacagc atagcaaaac tttaacctcc aaatcaagcc	1740
tctacttgaa tccttttctg agggatgaat aaggcatagg catcaggggc tgttgccaat	1800
gtgcattagc tgtttgcagc ctcaccttct ttcattggagt ttaagatata gtgtattttc	1860
ccaaggtttg aactagctct tcatttcttt atgttttaaa tgcactgacc tcccacattc	1920
cctttttagt aaaatattca gaaataattt aaatacatca ttgcaatgaa aataaatgtt	1980
ttttattagg cagaatccag atgctcaagg ccttcataa tatccccag tttagtagtt	2040
ggacttaggg aacaaaggaa cttttaatag aaattggaca gcaagaaagc gagcttctag	2100
ctttagtcct gttcctcagc taaaaatgg acacaatttc cagcagggtc tctgagggca	2160
aattcccttc cccaagggtt ttcaccaatt tctgtcatgg ctgggccaga ggcacccctg	2220
aaatttgtgc tgactacttc tgaccattct gcataaagct catctaggcc tctgacccag	2280
acccaagcaa ggggtgtgtc agggacaact tggtcctgaa ctgctgagat gaagaggggtg	2340
acatcatctc tgacaacacc agcaaaatca tcttcaacaa agtctctgga gaatccta	2400

NaDoc 002-PCT-1_seq listing_ST25.txt

ctgtcagtcc agaactctac agccccctgca acatccccttg ctgtgaggac tgggactgca	2460
gaagtgagtt tggccatgat ggctcctcct gtcaggagag gaaagagaag aaggtagta	2520
caattgctat agtgagttgt attatactat gcttatgatt aattgtcaaa ctagtgggtt	2580
catagtgccca ctttctctgc actgccccat ctctgcccc ccctttccca ggcatagaca	2640
gtcagtgact tacccttgta cagctcatcc attcccagag taattcctgc tgctgtcaca	2700
aactccagga ggaccatgtg gtctcttttc tcattagggc ctttgacag agcagattga	2760
gtgctgagat agtgattatc tgggaggaga actgggccat caccaatagg ggtgttctgc	2820
tggaatggc ctgccagttg gacagatcca tctcaatgt tgtgtctaatt cttgaaatta	2880
gccttaattc cattcctctg cttatctgcc ataattgaaa cattgtgaga attatagttg	2940
tactccagct tgtgaccag aatgtttcca tcttccttaa aatcaatgcc tttcagctca	3000
attctgttaa ccagtgtatc acctcaaac ttcacttctg cccttgtctt ataatttcca	3060
tcattcctaa agaagattgt cctctcctga acataacctt ctggcattgc agatttaaag	3120
aagtcagtct gcttcatgtg gtcagggtat ctgctgaaac attgaacacc ataagtcagg	3180
gtggtcacca gagttggcca aggcactggc agctttcctg ttgtacaaat gaacttcaga	3240
gtcagctttc cataagttgc atctccttca ccttcaccag acacagagaa tttgtggcca	3300
ttcacatcac catccagctc aaccagaatt gggacaacac cagtaaagag ttcttctccc	3360
ttgctcatgg tggcttgat ctgtaacggc gcagaacaga aaacgaaaca aagacgtaga	3420
gttgagcaag cagggtcagg caaagcgtgg agagccggct gagtctaggt aggcctccaag	3480
ggagcgccgg acaaaaggccc ggtctcgacc tgagctttaa acttacctag acggcggacg	3540
cagttcagga ggcaccacag gcgggaggcg gcagaacgag actcaaccgg cgtggatggc	3600
ggcctcaggt agggcgccgg gcgctggaag gagagatgag agcccctga agctgatctg	3660
acggttcact aaacgagctc tgcttatata gacctccac cgtacacgcc taccgcccac	3720
ttgctgaat gggcgaggat tggtacgaca ttttgaaag tcccgttgat ttactagtca	3780
aaacaaactc ccattgacgt caatggggtg gagacttggc aatccccgtg agtcaaaccg	3840
ctatccacgc ccattgatgt actgcaaaa ccgcatcatc atggtaatag cgatgactaa	3900
tacgtagatg tactgccaag taggaaagtc ccataaggtc atgtactggg cataatgcc	3960
ggcgggcat ttaccgtcat tgacgtcaat agggggcgta cttggcatat gatacacttg	4020
atgtactgcc aagtgggcag ttaccgtaa atactccacc cattgacgtc aatggaaagt	4080
ccctattggc gttactatgg gaacatacgt cattattgac gtcaatgggc gggggtcgtt	4140
ggcggtcag ccaggcgggc cattaccgt aagttatgta acgcctgcag gttaatgaag	4200
aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg	4260
ttttccata ggctccgcc ccctgacgag catcacaata atcgacgctc aagtcagagg	4320
tggcgaaacc cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgtg	4380
cgctctcctg ttccgacct gccgcttacc ggatacctgt ccgcctttct cccttcggga	4440

NaDoc 002-PCT-1_seq listing_ST25.txt

```

agcgtggcgc tttctcatag ctcacgctgt aggtatctca gttcgggtgta ggtcgttcgc      4500
tccaagctgg gctgtgtgca cgaaccccc gttcagcccc accgctgcgc cttatccggt      4560
aactatcgtc ttgagtccaa cccggtaaga cacgacttat cgccactggc agcagccact      4620
ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtgggtg      4680
cctaactacg gctacactag aagaacagta tttggtatct gcgctctgct gaagccagtt      4740
accttcggaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcgg      4800
ggtttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct      4860
ttgatctttt ctacggggtc tgacgctcag tggaacgaaa actcacgtta agggattttg      4920
gtcatggcta gttaattaac atttaaatca      4950

```

<210> 21
 <211> 7241
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

```

<400> 21
agatctgcgc agcaccatgg cctgaaataa cctctgaaag aggaacttgg ttaggtacct      60
accggaagga acccgcgcta tgacggcaat aaaaagacag aataaaacgc acggtgttct      120
tataatgggtt acaaataaag caatagcatc acaaatttca caaataaagc atttttttca      180
ctgcattcta gttgtggtaa taaaatatct ttattttcat tacatctgtg tgttgggttt      240
ttgtgtgtgg cctcccaaag tgctgggatt acaggcatga gccatcgagc ccaacccaat      300
tttttttttt tttaatttta ctttctgcaa tcattcatcc attcagccag tgcggtattt      360
ctgaggtgtg ttcatcgcg gatccatgcc tgccgcagta cagttgtgag ccaaatgaga      420
ctgagactag ttcccgccct ccaagagctt gcaagacccg cagtggcgta aaaacactaa      480
catcttttag tgatcgattc tgcactccag gggttttcaa tctactacaa gagtgaataa      540
gagttcgctt ttgtctgata tctgttgtca ttctctctcg cttctttaac tgattttttc      600
tcagctaata aaacatccac ccacaacccc ccgaacgccc gcaaacacca ggccactcta      660
gcaaaacctc tctcactccg cctgcgcaat ccagctgact tccggttaca gataaccacg      720
tgattgggaa cccttgctgc gcatgtctag taggaagtgc gactatacca ctttccctac      780
ggaaggggta cttttttatg tttttaagtt taaaaccgat ttctgatatt tgacttttat      840
catttcaggc ctatatggag gctatgagtg agtttagtgt ggcagaagat gaaagaaccg      900
gacaggaata cggacgaaat tggagcaggg tttgggctct ccccttcgca gataatcgga      960
ggagccgggc ccgagcgagc tctttccttt cgctgctgcg gccgcagccg tgaggtgagg      1020
gcgagctggg ctccatcagg cgctgacgcg tgcgacaag ggactgtcgg tcttgggacc      1080
gcagctgggg ttgggggaga tgaaatggag gccgccctaa agcggccggg cccgggggtt      1140
ggggtaggcc ggagcacttt cgtccccggc ctccggagtg agggggggcg gggagcgctg      1200

```

NaDoc 002-PCT-1_seq listing_ST25.txt

cagcaactga gaccaggaaa agtctgcccc ggctggtgcc gcaccgcaca cgtgtccggt	1260
cgacccacgc gagcagagca aacggagcga acaagaccaa gccgtgggcc ctttcttgct	1320
tggcacaccc ggagcggagc cgatctctgc ttccacgtga tgtagggcaa gcctagtgtg	1380
ggccccaggc ctccgactgc cgagagaggt gatctctaac tcttgactcc attcactcct	1440
ttggcctctc ataaaggaaa tctctgcgaa tagccgaacg aggccttgta ctgtgataaa	1500
acagggaaat aagcccagaa aacagagtaa cttgcctgca ttcctagact agaaatcagg	1560
tctactcacc tcgaatattc tttaaagcgt gaggaccaga aatggcataa cccccctatt	1620
caatccaata agtccttggc ttgactttcc agaggagaaa tgcgaacatg aggcctccgag	1680
aggtaaggc atagcgtggg ttttgaagtc ttaaacccaa gggggccagc tgcataagccc	1740
agagccttaa agatgattta ggaagagtc ttatttcgcg gctgtggtgt gggtcacaaa	1800
gggcaggctt tgatggggac gttcattctt gccaggatt ggctttcaga gtctaatacat	1860
gttttctgtg tgtctagtat cctcaggctt cagaagaggc tcgcctctag tgcctccgc	1920
tgtggcaaga agaaggtctg gaccggtccc acgcgtcaat tggaaaactt acgctgagta	1980
cttcgatctc cctacggcaa gctgaccctg aagttcaaca gatctcgccg ccatgggagc	2040
tgatgatgtg gttgattctt cgaaatcttt tgtcatggaa aacttttctt cgtaccacgg	2100
gacgaaacct ggttatgtgg attccattca aaaaggcata caaaagccaa aatctggtac	2160
acaaggaaac tatgacgatg atttgaaagg gttttatagt accgacaaca aatatgacgc	2220
tgcgggatac tctgtggata atgaaaaccc gctctctgga aaagctggag gcgtgggtcaa	2280
agtgacgtat ccaggactga cgaaggttct cgcactaaag gtggataatg ccgaaactat	2340
taagaaagag ttaggtttta gtctcactga accgctcatg gagcaagtcg gaacggaaga	2400
gtttatcaaa agattcgggtg atggtgcttc gcgtgtagtg ctcagccttc ctttcgctga	2460
ggggagttct agcgttgagt acatcaacaa ctgggaacag gcgaaagcgt taagcgtaga	2520
acttgagatt aactttgaaa cccgtggaaa acgtggccaa gatgcgatgt atgagtatat	2580
ggctcaagcc tgtgcaggaa atcgtgtcag gcgatagtga actagtatcc ggaatctaga	2640
gcggccgctg gccgcaataa aatatcttta ttttcattac atctgtgtgt tggttttttg	2700
tgtgaggatc taaatgagtc ttcggacctc gcgggggccg cttaaagcgt ggttaggggt	2760
tgctgacgc ggggggaggg ggaaggaaacg aaacactctc attcggaggc ggctcggggt	2820
ttggtcttgg tggccacggg cacgcagaag agcgccgca tcctcttaag ccccccccg	2880
ccctccgtgg aggcgggggt ttggtcggcg ggtggttaact ggcggggccg tgactcgggc	2940
gggtcgcgcg cccagagtg tgaccttttc ggtctgctcg cagacccccg ggcggcgccg	3000
ccgcggcggc gacgggctcg ctgggtccta ggctccatgg ggaccgtata cgtggacagg	3060
ctctggagca tccgcacgac tgcggtgata ttaccggaga cttctgctcg gacgagccgg	3120
gtcacgcggc tgacgcggag cgtccgttgg gcgacaaaca ccaggacggg gcacaggtac	3180
actatcttgt caccggagg cgcgaggagc tgcaggagct tcaggagtg gcgcagctgc	3240

NaDoc 002-PCT-1_seq listing_ST25.txt

ttcatccccg tggcccgttg ctgcgctttg ctggcggtgt ccccggaaga aatatatttg	3300
catgtcttta gttctatgat gacacaaacc ccgcccagcg tcttgtcatt ggcgaattcg	3360
aacacgcaga tgcagtcggg gcggcgcggt cccaggtcca cttcgcatat taaggtgacg	3420
cgtgtggcct cgaacaccga gcgacctgc agcgaccgc ttaaaagctt ggcattccgg	3480
tactgttggg aaagccacca tggccgatgc taagaacatt aagaagggcc ctgctccctt	3540
ctaccctctg gaggatggca ccgctggcga gcagctgcac aaggccatga agaggtatgc	3600
cctggtgcct ggcaccattg ccttcaccga tgccacatt gaggtggaca tcacctatgc	3660
cgagtacttc gagatgtctg tgcgcctggc cgaggccatg aagaggtacg gcctgaacac	3720
caaccaccgc atcgtggtgt gctctgagaa ctctctgcag ttcttcatgc cagtgtggg	3780
cgccctgttc atcggagtgg ccgtggcccc tgctaacgac atttacaacg agcgcgagct	3840
gctgaacagc atgggcattt ctacgcctac cgtggtgttc gtgtctaaga agggcctgca	3900
gaagatcctg aacgtgcaga agaagctgcc tatcatccag aagatcatca tcatggactc	3960
taagaccgac taccagggtt tccagagcat gtacacattc gtgacatctc atctgcctcc	4020
tggcttcaac gagtacgact tcgtgccaga gtctttcgac agggacaaaa ccattgccct	4080
gatcatgaac agctctgggt ctaccggcct gcctaagggc gtggccctgc ctcatcgac	4140
cgcctgtgtg cgcttctctc acgcccgcga ccctattttc ggcaaccaga tcatccccga	4200
caccgctatt ctgagcgtgg tgccattcca ccacggcttc ggcattgtca ccaccctggg	4260
ctacctgatt tgcggctttc ggggtggtgct gatgtaccgc ttcgaggagg agctgttctt	4320
gcgcagcctg caagactaca aaattcagtc tgccctgctg gtgccaaccc tgttcagctt	4380
cttcgctaag agcaccctga tcgacaagta cgacctgtct aacctgcacg agattgcctc	4440
tggcggcgcc ccactgtcta aggaggtggg cgaagccgtg gccaaagcgt ttcacttgcc	4500
aggcatccgc cagggctacg gcctgaccga gacaaccagc gccatttctga ttaccccaga	4560
gggcgacgac aagcctggcg ccgtgggcaa ggtggtgcc a ttcttcgagg ccaaggtggt	4620
ggacctggac accggcaaga ccctgggagt gaaccagcgc ggcgagctgt gtgtgcgcgg	4680
ccctatgatt atgtccggct acgtgaataa ccctgaggcc acaaacgccc tgatcgacaa	4740
ggacggctgg ctgcactctg gcgacattgc ctactgggac gaggacgagc acttcttcat	4800
cgtggaccgc ctgaagtctc tgatcaagta caagggtac caggtggccc cagccgagct	4860
ggagtctatc ctgctgcagc accctaacat ttctgacgcc ggagtggccg gcctgcccga	4920
cgacgatgcc ggcgagctgc ctgccgccgt cgtcgtgctg gaacacggca agaccatgac	4980
cgagaaggag atcgtggact atgtggccag ccaggtgaca accgccaaga agctgcgcgg	5040
cggagtgggtg ttcgtggacg aggtgcccaa gggcctgacc ggcaagctgg acgcccgcaa	5100
gatccgcgag atcctgatca aggctaagaa aggcggcaag atcgccgtgt aataattcta	5160
gagtcggggc ggccggccgc ttcgagcaga catgataaga tacattgatg agtttgaca	5220
aaccacaact agaatgcagt gaaaaaatg ctttatttgt gaaatttgtg atgctattgc	5280

NaDoc 002-PCT-1_seq listing_ST25.txt

tttatttgta accattataa gctgcaataa acaagttaac aacaacaatt gcattcattt	5340
tatgtttcag gttcaggggg aggtgtggga ggttttttta agcaagtaaa acctctacaa	5400
atgtggtaaa atcgataag atccaggtgg cacttttcgg ggaaatgtgc gcggaacccc	5460
tatttgttta tttttctaaa tacattcaaa tatgtatccg ctcatgagac aataaccctg	5520
ataaatgctt caataatatt gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc	5580
ccttattccc ttttttgcgg ctttttgcct tcctgttttt gtcacccag aaacgctgg	5640
gaaagtaaaa gatgctgaag atcagttggg tgcacgagtg gggtacatcg aactggatct	5700
caacagcggg aagatccttg agagttttcg ccccgaagaa cgttttccaa tgatgagcac	5760
ttttaagtt ctgctatgtg gcgcggtatt atcccgtatt gacgccgggc aagagcaact	5820
cggtcgccgc atacactatt ctcaaatga cttggttgag tactcaccag tcacagaaaa	5880
gcattcttac gatggcatga cagtaagaga attatgcagt gctgccataa ccatgagtga	5940
taacactgcg gccaaacttac ttctgacaac gatcggagga ccgaaggagc taaccgcttt	6000
tttgacaaac atgggggatc atgtaactcg ctttgatcgt tgggaaccgg agctgaatga	6060
agccatacca aacgacgagc gtgacaccac gatgcctgta gcaatggcaa caacgttg	6120
caaactatta actggcgaac tacttactct agcttcccgg caacaattaa tagactggat	6180
ggaggcggat aaagtgtcag gaccacttct gcgctcggcc cttccggctg gctggtttat	6240
tgctgataaa tctggagccg gtgagcgtgg gtctcgcgg atcattgcag cactggggcc	6300
agatggtaag ccttcccgtg tcgtagttat ctacacgagc gggagtcagg caactatgga	6360
tgaacgaaat agacagatcg ctgagatagg tgcctcactg attaagcatt ggtaactgtc	6420
agaccaagtt tactcatata tacttttagat tgatttaaaa cttcattttt aatttaaaag	6480
gatctaggtg aagatccttt ttgataatct catgaccaa atcccttaac gtgagttttc	6540
gttccactga gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag atcctttttt	6600
tctgcgcgta atctgtgct tgcaaaaaaaa aaaaccaccg ctaccagcgg tggtttgttt	6660
gccggatcaa gagctaccaa ctctttttcc gaaggtaact ggcttcagca gagcgcagat	6720
accaaatact gttcttctag tgtagccgta gttaggccac cacttcaaga actctgtagc	6780
accgcctaca tacctcgctc tgctaatacct gttaccagtg gctgctgcca gtggcgataa	6840
gtcgtgtctt accgggttgg actcaagacg atagttaccg gataaggcgc agcggtcggg	6900
ctgaacgggg gggttcgtgca cacagcccag cttggagcga acgacctaca ccgaactgag	6960
atacctacag cgtgagctat gagaaagcgc cagcgttccc gaaggagaaa aggcggacag	7020
gtatccggta agcggcaggg tcggaacagg agagcgcacg agggagcttc cagggggaaa	7080
cgcctggat ctttatagtc ctgtcgggtt tcgccacctc tgacttgagc gtcgattttt	7140
gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg	7200
gttcctggcc ttttgctggc cttttgctca catggctcga c	7241

NaDoc 002-PCT-1_seq listing_ST25.txt

<210> 22
 <211> 6326
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 22
 agatctgcgc agcaccatgg cctgaaataa cctctgaaag aggaacttgg ttaggtacct 60
 tctgaggcgg aaagaaccag ctgtggaatg tgtgtcagtt aggggtgtgga aagtccccag 120
 gctccccagc aggcagaagt atgcaaagca tgcattctca ttagtcagca accaggtgtg 180
 gaaagtcccc aggtctccca gcaggcagaa gtatgcaaag catgcatctc aattagtcag 240
 caaccatagt cccgccccta actccgcca tcccgccct aactccgcc agttccgccc 300
 attctccgcc ccatggctga ctaatttttt ttatttatgc agaggccgag gccgcctcgg 360
 cctctgagct attccagaag tagtgaggag gcttttttgg aggcctaggc ttttgcaaaa 420
 agcttgatct ttctgacaca acagtctcga acttaagctg cagaagttgg tcgtgaggca 480
 ctgggcaggt aagtatcaag gttacaagac aggtttaagg agaccaatag aaactgggct 540
 tgtcgagaca gagaagactc ttgcgtttct gataggcacc tattggtctt actgacatcc 600
 actttgcctt tctctccaca ggtgtccact cccagttcaa ttacagctct taaggctaga 660
 gtacttaata cgactcacta taggctagcc cgaccgggtca acaacgtctc agatcttccc 720
 ccaaggcgcg ccgaactcta gccaccatgg cttccaaggt gtacgacccc gagcaacgca 780
 aacgcatgat cactgggcct cagtgggtgg ctcgctgcaa gcaaatgaac gtgctggact 840
 ctttcatcaa ctactatgat tccgagaagc acgccgagaa cgccgtgatt tttctgcatg 900
 gtaacgctgc ctccagctac ctgtggaggc acgtcgtgcc tcacatcgag cccgtggcta 960
 gatgcatcat ccctgatctg atcggaatgg gtaagtccgg caagagcggg aatggctcat 1020
 atcgctctct ggatcactac aagtacctca ccgcttggtt cgagctgctg aaccttccaa 1080
 agaaaatcat ctttgtgggc cagactggg gggcttgtct ggcctttcac tactcctacg 1140
 agcaccaaga caagatcaag gccatcgtcc atgctgagag tgtcgtggac gtgatcgagt 1200
 cctgggacga gtggcctgac atcgaggagg atatcgccct gatcaagagc gaagagggcg 1260
 agaaaatggg gcttgagaat aacttcttcg tcgagaccat gctccaagc aagatcatgc 1320
 ggaaactgga gcctgaggag ttcgctgcct acctggagcc attcaaggag aagggcgagg 1380
 ttagacggcc taccctctcc tggcctcgcg agatccctct cgtaaggga ggcaagcccg 1440
 acgtcgtcca gattgtccgc aactacaacg cctaccttcg ggccagcgac gatctgccta 1500
 agatgttcat cgagtccgac cctgggttct tttccaacgc tattgtcgag ggagctaaga 1560
 agttccctaa caccgagttc gtgaagggtga agggcctcca cttcagccag gaggacgctc 1620
 cagatgaaat gggtaagtac atcaagagct tcgtggagcg cgtgctgaag aacgagcagt 1680
 aattctaggc gatcgtcga gcccggaat tcgtttaaac ctagagcggc cgctggccgc 1740
 aataaaatat ctttattttc attacatctg tgtgttggtt ttttgtgtga ggatctaaat 1800

NaDoc 002-PCT-1_seq listing_ST25.txt

gagtccttcgg acctcgcggg ggccgcttaa gcggtggtta gggtttgtct gacgcggggg	1860
gagggggaag gaacgaaaca ctctcattcg gaggcggctc ggggttttgtt cttggtggcc	1920
acgggcacgc agaagagcgc cgcgatactc ttaagcaccc ccccgccctc cgtggaggcg	1980
ggggtttgtt cggcgggtgg taactggcgg gccgctgact cgggcgggtc gcgcgcccc	2040
gagtgtgacc ttttcggtct gctcgcagac ccccgggcgg cgcgcgcgcg gcggcgacgg	2100
gctcgcgtgg tcctaggctc catggggacc gtatacgtgg acaggctctg gagcatccgc	2160
acgactgcgg tgatattacc ggagaccttc tgcgggacga gccgggtcac gcggctgacg	2220
cggagcgtcc gttgggacgc aaacaccagg acggggcaca ggtacactat cttgtcacc	2280
ggaggcgcga gggactgcag gagcttcagg gagtggcgca gctgcttcac ccccggtggc	2340
cgttgctcgc gtttgctggc ggtgtccccg gaagaaatat atttgcatgt ctttagttct	2400
atgatgacac aaaccccgcc cagcgtcttg tcattggcga attcgaacac gcagatgcag	2460
tcggggcggc gcgggtccag gtccacttcg catattaagg tgacgcgtgt ggccctgaac	2520
accgagcgac cctgcagcga cccgcttaaa agcttggtcat tccggtactg ttggtaaagc	2580
caccatggcc gatgctaaga acattaagaa gggccctgct cccttctacc ctctggagga	2640
tggcaccgct ggcgagcagc tgcacaaggc catgaagagg tatgccctgg tgcctggcac	2700
cattgccttc accgatgccc acattgaggt ggacatcacc tatgccgagt acttcgagat	2760
gtctgtgcgc ctggccgagg ccatgaagag gtacggcctg aacaccaacc accgcatcgt	2820
ggtgtgtctt gagaactctc tgcagttctt catgccagtg ctgggcgccc tgttcacg	2880
agtggccgtg gccctgcta acgacattta caacgagcgc gagctgctga acagcatggg	2940
catttctcag cctaccgtgg tgttcgtgtc taagaaggcg ctgcagaaga tcctgaacgt	3000
gcagaagaag ctgcctatca tccagaagat catcatcatg gactctaaga ccgactacca	3060
gggcttccag agcatgtaca cattcgtgac atctcatctg cctcctggct tcaacgagta	3120
cgacttcgtg ccagagtctt tcgacaggga caaaaccatt gccctgatca tgaacagctc	3180
tgggtctacc ggccctgcta agggcgtggc cctgcctcat cgcaccgcct gtgtgcgctt	3240
ctctcacgcc cgcgacccta ttttcggcaa ccagatcatc cccgacaccg ctattctgag	3300
cgtggtgcca ttccaccacg gcttcggcat gttcaccacc ctgggctacc tgatttgcg	3360
ctttcgggtg gtgctgatgt accgcttcga ggaggagctg ttcctgcgca gcctgcaaga	3420
ctacaaaatt cagtctgccc tgctggtgcc aaccctgttc agcttcttcg ctaagagcac	3480
cctgatcgac aagtacgacc tgtctaacct gcacgagatt gcctctggcg gcgccccact	3540
gtctaaggag gtgggcgaag ccgtggccaa gcgctttcat ctgccaggca tccgccaggg	3600
ctacggcctg accgagacaa ccagcgccat tctgattacc ccagaggcg acgacaagcc	3660
tggcgccgtg ggcaagggtg tgccattctt cgaggccaag gtggtggacc tggacaccgg	3720
caagaccctg ggagtgaacc agcgcggcga gctgtgtgtg cgcggcccta tgattatgtc	3780
cggctacgtg aataaccctg aggccacaaa cgccctgatc gacaaggacg gctggctgca	3840

NaDoc 002-PCT-1_seq listing_ST25.txt

ctctggcgac attgcctact gggacgagga cgagcacttc ttcacgtgg accgcctgaa	3900
gtctctgata aagtacaagg gctaccaggt ggccccagcc gagctggagt ctatcctgct	3960
gcagcaccct aacatttttcg acgccggagt ggccggcctg cccgacgacg atgccggcga	4020
gctgcctgcc gccgtcgtcg tgctggaaca cggcaagacc atgaccgaga aggagatcgt	4080
ggactatgtg gccagccagg tgacaaccgc caagaagctg cgcggcggag tgggtgttcgt	4140
ggacgagggtg cccaagggcc tgaccggcaa gctggacgcc cgcaagatcc gcgagatcct	4200
gatcaaggct aagaaaggcg gcaagatcgc cgtgtaataa ttctagagtc ggggcggccg	4260
gccgcttcga gcagacatga taagatacat tgatgagttt ggacaaacca caactagaat	4320
gcagtgaaaa aaatgcttta tttgtgaaat ttgtgatgct attgctttat ttgtaaccat	4380
tataagctgc aataaacaag ttaacaacaa caattgcatt cttttatgt ttcaggttca	4440
gggggagggtg tgggagggtt tttaaagcaa gtaaaacctc taaaatgtg gtaaaatcga	4500
taaggatcca ggtggcactt ttcggggaaa tgtgcgcgga acccctatit gtttattttt	4560
ctaaatacat tcaaatatgt atccgctcat gagacaataa ccctgataaa tgcttcaata	4620
atattgaaaa aggaagagta tgagtattca acatttccgt gtcgccctta ttcccttttt	4680
tgccggcattt tgccttcctg tttttgctca cccagaaacg ctggtgaaag taaaagatgc	4740
tgaagatcag ttgggtgcac gagtgggtta catcgaactg gatctcaaca gcggtaatag	4800
ccttgagagt tttcgccccg aagaacgtt tccaatgatg agcactttta aagtctctgt	4860
atgtggcgcg gtattatccc gtattgacgc cgggcaagag caactcggtc gccgcataca	4920
ctattctcag aatgacttgg ttgagtactc accagtcaca gaaaagcatc ttacggatgg	4980
catgacagta agagaattat gcagtgtcgc cataaccatg agtgataaca ctgcggccaa	5040
cttacttctg acaacgatcg gaggaccgaa ggagctaacc gcttttttgc acaacatggg	5100
ggatcatgta actcgccttg atcgttggga accggagctg aatgaagcca taccaaacga	5160
cgagcgtgac accacgatgc ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg	5220
cgaactactt actctagctt cccggcaaca attaatagac tggatggagg cggataaagt	5280
tgcaggacca cttctgcgct cggcccttcc ggctggctgg tttattgctg ataaatctgg	5340
agccggtgag cgtgggtctc gcggtatcat tgcagcactg gggccagatg gtaagccctc	5400
ccgtatcgta gttatctaca cgacggggag tcaggcaact atggatgaac gaaatagaca	5460
gatcgtgag ataggtgcct cactgattaa gcattggtaa ctgtcagacc aagtttactc	5520
atatatactt tagattgatt taaaacttca tttttaattt aaaaggatct aggtgaagat	5580
cctttttgat aatctcatga ccaaataccc ttaacgtgag ttttcgttcc actgagcgtc	5640
agaccccgta gaaaagatca aaggatcttc ttgagatcct ttttttctgc gcgtaatctg	5700
ctgcttgcaa acaaaaaaac caccgctacc agcgggtggt tgtttgccgg atcaagagct	5760
accaactctt tttccgaagg taactggctt cagcagagcg cagataccaa atactgttct	5820
tctagtgtag ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct	5880

NaDoc 002-PCT-1_seq listing_ST25.txt

cgctctgcta atcctgttac cagtggctgc tgccagtggc gataagtcgt gtcttaccgg	5940
gttggactca agacgatagt taccggataa ggcgagcgg tcgggctgaa cggggggttc	6000
gtgcacacag cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga	6060
gctatgagaa agcgccacgc ttcccgaagg gagaaaggcg gacaggatc cggtaacgcg	6120
cagggtcgga acaggagagc gcacgagga gcttccagg ggaaacgcct ggtatcttta	6180
tagtcctgtc gggtttcgcc acctctgact tgagcgtcga tttttgtgat gctcgtcagg	6240
ggggcggagc ctatggaaaa acgccagcaa cgcggccttt ttacggttcc tggccttttg	6300
ctggcctttt gctcacatgg ctcgac	6326

<210> 23
 <211> 9195
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 23	
agatctgcgc agcaccatgg cctgaaataa cctctgaaag aggaacttgg ttaggtacct	60
tctgaggcgg aaagaaccag ctgtggaatg tgtgtcagtt aggggtgtga aagtccccag	120
gctccccagc aggcagaagt atgcaaagca tgcattctca ttagtcagca accagggtgtg	180
gaaagtcccc aggtctcccca gcaggcagaa gtatgcaaag catgcatctc aattagtcag	240
caaccatagt cccgccccta actccgcca tcccgcctt aactccgccc agttccgccc	300
attctccgcc ccatggctga ctaatttttt ttatttatgc agaggccgag gccgcctcgg	360
cctctgagct attccagaag tagtgaggag gcttttttgg aggcctaggc ttttgcaaaa	420
agcttgattc ttctgacaca acagtctcga acttaagctg cagaagttgg tcgtgaggca	480
ctgggcaggt aagtatcaag gttacaagac aggtttaagg agaccaatag aaactgggct	540
tgctgagaca gagaagactc ttgcgtttct gataggcacc tattggctctt actgacatcc	600
actttgcctt tctctccaca ggtgtccact cccagttcaa ttacagctct taaggctaga	660
gtacttaata cgactcacta taggctagcc cgaccggtca acaacgtctc agatcccggt	720
ccgactagtc gtacgcaccg gcggtcgacc acgtgcctgc aggaaagtgc gcacggtcac	780
cgagctcaaa acgccaagaa cctcatcatt ttccctacgg caagctgacc ctgaagttca	840
tcaactacac aaatcagcga tttccagctc agcgggaccc gcatccggag gcgcgccgaa	900
ctctagccac catggcttcc aagggtgtacg accccgagca acgcaaacgc atgatcactg	960
ggcctcagtg gtgggctcgc tgcaagcaaa tgaacgtgct ggactccttc atcaactact	1020
atgattccga gaagcacgcc gagaacgccg tgatttttct gcatggtaac gctgcctcca	1080
gctacctgtg gaggcacgtc gtgcctcaca tcgagcccgt ggctagatgc atcatccctg	1140
atctgatcgg aatgggtaag tccggcaaga gcgggaatgg ctcatatcgc ctcctggatc	1200
actacaagta cctcaccgct tggttcgagc tgctgaacct tccaaagaaa atcatctttg	1260

NaDoc 002-PCT-1_seq listing_ST25.txt

tgggccacga ctggggggct tgtctggcct ttcactactc ctacgagcac caagacaaga	1320
tcaaggccat cgtccatgct gagagtgtcg tggacgtgat cgagtcctgg gacgagtggc	1380
ctgacatcga ggaggatata gccctgatca agagcgaaga gggcgagaaa atggtgcttg	1440
agaataactt cttcgtcgag accatgctcc caagcaagat catgcggaaa ctggagcctg	1500
aggagttcgc tgcctacctg gagccattca aggagaaggg cgaggttaga cggcctaccc	1560
tctcctggcc tcgcgagatc cctctcgta agggaggcaa gcccgcgtc gtccagattg	1620
tccgcaacta caacgcctac cttcgggcca gcgacgatct gcctaagatg ttcacgagt	1680
ccgaccctgg gttcttttcc aacgctattg tcgagggagc taagaagttc cctaaccacg	1740
agttcgtgaa ggtgaagggc ctccacttca gccaggagga cgctccagat gaaatgggta	1800
agtacatcaa gagcttcgtg gagcgcgtgc tgaagaacga gcagtaattc taggcgatcg	1860
tctagagcct gacctagctg acctagccgc caccatgcag aggctgcagg tagtgctggg	1920
ccacctgagg ggtccggccg attccggctg gatgccgcag gcagcgcctt gcctgagatc	1980
tactagtggg gacctagctg acctagccgt caccatggac cctgttgtgc tgcaaaggag	2040
agactgggag aacctgggag tcacccagct cactgaactt aacctagcc ctgccaccat	2100
ggcttgagg aactccgagg aagccaggac tgacagtaac agtaggcagc gccgccatgg	2160
caacggagag tggaggtttg cctggtttga cgctaaggat agtgtggcca ccatggggct	2220
ggagtgcgac ctcccagagg cggtgacgt taaagttagc agcagccgcc atgggcacgg	2280
ctacgacgcg cccatctaca gtgacgttag cactaagatc gccaccatgg ccccttttgt	2340
gcccaccgag aacctgacta actgtagcag tgacacctgc cgccatggcg agagctggct	2400
gcaagaagga cagactaaga ttagctgtga cggagccacc atggccttcc acctctgggt	2460
caacggcagg tgtaacggtg gcggtgaaga cagccgcat ggctccgagt tcgaccttc	2520
tgcttccgt aacgctgaag ataacagggc caccatgggt gtgctcaggt ggtccgacgg	2580
cagctatagg gctgaccgta acatgtgccg ccatgggtggc atcttcaggg acgtcagcct	2640
gcttagcact gagactaacc aggccaccat ggtccacgtt gccacgaggt tcaacgacga	2700
tagcagtga gctaagctgg gccgccatgg gcagatgtgt ggagaactca gagagtctga	2760
cagtagcact aagagcgcca ccatgggcga gaccaggtg gcctctggca cagctgactt	2820
tagagctaag atcagccgcc atggaggagg ctacgccgac agagccaccc ttgagcttag	2880
cgttaagaac gccaccatgg ggtctgccga gaccccaac ctctacagta acgttagggc	2940
tgagcacagc cgccatggca cgctcatcga agccgaagcc tgcgataacg gtgacagtag	3000
agtcgccacc atggacggcc tgctgctgct caacggcaag cctaagcttg acagtagagt	3060
cagccgccat gggcaccatc ctctgcacgg acaagtcatt aaggctgaga ctagggtggc	3120
caccatgggt ctacgaagc agaacaactt caacgctaag agtgactcta gtaccgccc	3180
ccatgggtctc tggtaacccc tgtgcgacag gaataacctt agggttgacg aggccaccat	3240
ggtcgagaca cacggcatgg tgcccacgaa taagcttaga gctgaccca gccgccatgg	3300

NaDoc 002-PCT-1_seq listing_ST25.txt

tgccatgtcc gagagagtca ccaggattaa gcatagagct gagaacgcca ccatgggtcat	3360
catctgggtct ctgggcaacg agtctaagca tagagctgac cacggccgcc atggcagggtg	3420
gatcaagtct gccgacccca gtagacctga gcataacgaa gccaccatgg cagacaccac	3480
agccacagac atcatctgta gcattgaggc taaggctggc cgccatgggc ccttcctg	3540
tgtgccaag tggagtagca ctgagtgtaa ctctgccacc atggaaacga gacctctcat	3600
cctgtgagag tatagacctg aaagtaaacc cggccgccat ggctttgcca agtactggca	3660
agccttcact gagtatagca ctaagcaagc caccatggcc cgcgactggg tggaccagtc	3720
actcattgag tatagcgata acggcagccg ccatgggtgcc tacggaggag actttggcga	3780
cactgagaat agcactaagt tcgccaccat gggcctggtc tttgccgacc ggactccgcc	3840
tgaccctagc actaaggcca gccgccatgg acagttcttc ccgttcacgc tgtctggtga	3900
aactaacgat agcacagcca ccatggtctt cagacactcc gacaacgagc tgcttgactg	3960
taaggctagg ctgggcccgc atggtctggc ttctggcgag gtgcctctgg ctgaggctaa	4020
tcatagaaag gccaccatgg aactgcccga gctgcctcag ccagagtctg acgataacct	4080
taggctcagc cgccatgggg tttagcccaa cgcaacagct tggctctgag ctagccataa	4140
ctctgccacc atggagtgga ggctggccga gaacctctcg gttgacctta gggctaactc	4200
tcgccgccat ggtcacctca caacatccga aatggagttt gacattaggc ttaacaacgc	4260
caccatggag ttcaacaggc agtctggctt cctgtctgag attaggacta aagacagccg	4320
ccatggcctc tctcctctcc gagacctgtt cactagggct gagcttaaca ctgccaccat	4380
gggtgtcagag gccaccagga tcgacccaaa tagttgtgag gataagtga gccgccatgg	4440
acactaccag gccgaggctg ccctgcttag ctgtgaagct aaccagcgaa tgcttggggc	4500
tctcatcacc acagcccacg cttgtagccc tgaagctaa acagcgaatg cctggggcaa	4560
gacctacaga atcgacggcc ataggaaacc tagagcggcc gctggccgca ataaaatatc	4620
tttattttca ttacatctgt gtgttggttt tttgtgtgag gatctaaatg agtcttcgga	4680
cctcgcgggg gccgcttaag cgggtggttag ggtttgtctg acgcgggggg agggggaagg	4740
aacgaaacac tctcattcgg aggcggctcg gggtttggtc ttggtggcca cgggcacgca	4800
gaagagcgcc gcgacacctt taagcaccac cccgccctcc gtggaggcgg gggtttggtc	4860
ggcgggtggt aactggcggg ccgctgactc gggcgggtcg cgcgccccag agtgtgacct	4920
tttcggtctg ctgcgagacc cccgggcggc gccgcgcggc cggcgacggg ctgcgtgggt	4980
cctaggctcc atggggaccg tatacgtgga caggctctgg agcatccgca cgactgcggt	5040
gatattaccg gagaccttct gcgggacgag ccgggtcacg cggctgacgc ggagcgtccg	5100
ttgggcgaca aacaccagga cggggcacag gtacactatc ttgtcaccg gaggcgcgag	5160
ggactgcagg agcttcaggg agtggcgag ctgcttcac cccgtggccc gttgctcgcg	5220
tttgctggcg gtgtccccgg aagaaatata tttgcatgtc tttagttcta tgatgacaca	5280
aaccccgccc agcgtcttgt cattggcgaa ttcgaacacg cagatgcagt cggggcggcg	5340

NaDoc 002-PCT-1_seq Listing_ST25.txt

cggtcccagg tccacttcgc atattaaggt gacgcgtgtg gcctcgaaca ccgagcgacc	5400
ctgcagcgac ccgcttaaaa gcttggcatt ccggtactgt tggtaaagcc accatggccg	5460
atgctaagaa cattaagaag ggccctgctc ctttctaccc tctggaggat ggcaccgctg	5520
gcgagcagct gcacaaggcc atgaagaggt atgccctggt gcctggcacc attgccttca	5580
ccgatgcca cattgaggtg gacatcacct atgccgagta cttcgagatg tctgtgcgcc	5640
tggccgaggg catgaagagg tacggcctga acaccaacca ccgcatcgctg gtgtgctctg	5700
agaactctct gcagttcttc atgccagtgc tgggcgcctt gttcatcgga gtggccgtgg	5760
cccctgctaa cgacatttac aacgagcgcg agctgctgaa cagcatgggc atttctcagc	5820
ctaccgtggt gttcgtgtct aagaaggggc tgcagaagat cctgaacgtg cagaagaagc	5880
tgcctatcat ccagaagatc atcatcatgg actctaagac cgactaccag ggcttccaga	5940
gcatgtacac attcgtgaca tctcatctgc ctcttggtt caacgagtac gacttcgtgc	6000
cagagtcttt cgacagggac aaaaccattg ccctgatcat gaacagctct ggggtctaccg	6060
gcctgcctaa gggcgtggcc ctgcctcatc gcaccgcctg tgtgcgcttc tctcacgccc	6120
gcgaccctat tticggcaac cagatcatcc ccgacaccgc tattctgagc gtggtgccat	6180
tccaccacgg cttcggcatg ttcaccaccc tgggctacct gatttgcggc tticgggtgg	6240
tgctgatgta ccgcttcgag gaggagctgt tcctgcgcag cctgcaagac taaaaattc	6300
agtctgccct gctggtgcca accctgttca gcttcttcgc taagagcacc ctgatcgaca	6360
agtagcacct gtctaacctg cacgagattg cctctggcgg cgcctcactg tctaaggagg	6420
tgggcgaagc cgtggccaag cgctttcatc tgccaggcat ccgccagggc tacggcctga	6480
ccgagacaac cagcgccatt ctgattaccc cagagggcga cgacaagcct ggcgccgtgg	6540
gcaaggtggt gccattcttc gaggccaagg tgggtggacct ggacaccggc aagaccctgg	6600
gagtgaacca gcgcggcgag ctgtgtgtgc gcggccctat gattatgtcc ggctacgtga	6660
ataaccctga ggccacaaac gccctgatcg acaaggacgg ctggctgcac tctggcgaca	6720
ttgcctactg ggacgaggac gagcacttct tcatcgtgga ccgcctgaag tctctgatca	6780
agtacaaggg ctaccaggtg gccccagccg agctggagtc tatcctgctg cagcacccta	6840
acattttcga cgccggagtg gccggcctgc ccgacgacga tgccggcgag ctgcctgccg	6900
ccgtcgtcgt gctggaacac ggcaagacca tgaccgagaa ggagatcgctg gactatgtgg	6960
ccagccaggt gacaaccgcc aagaagctgc gcggcggagt ggtgttcgtg gacgaggtgc	7020
ccaagggcct gaccggcaag ctggacgccc gcaagatccg cgagatcctg atcaaggcta	7080
agaaaggcgg caagatcgcc gtgtaataat tctagagtcg gggcggccgg ccgcttcgag	7140
cagacatgat aagatacatt gatgagtttg gacaaaccac aactagaatg cagtgaaaaa	7200
aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca	7260
ataaacaagt taacaacaac aattgcattc attttatgtt tcaggttcag ggggaggtgt	7320
gggaggtttt ttaaagcaag taaaacctct acaaatgtgg taaaatcgat aaggatccag	7380

NaDoc 002-PCT-1_seq listing_ST25.txt

gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt	7440
caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa	7500
ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt	7560
gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt	7620
tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaagatc cttgagagtt	7680
ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcg	7740
tattatcccg tattgacgcc gggcaagagc aactcggctg ccgcatacac tattctcaga	7800
atgacttggt tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa	7860
gagaattatg cagtgtgccc ataaccatga gtgataacac tgcggccaac ttacttctga	7920
caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa	7980
ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca	8040
ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaaact attaaactggc gaactactta	8100
ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac	8160
ttctgcgctc ggcccttccg gctggctggt ttattgctga taaatctgga gccggtgagc	8220
gtgggtctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag	8280
ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga	8340
taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca tatatacttt	8400
agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata	8460
atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag	8520
aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa	8580
caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt	8640
ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgttctt ctagttagc	8700
cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatacctc gctctgctaa	8760
tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa	8820
gacgatagtt accggataag gcgcagcggc cgggctgaac ggggggttcg tgcacacagc	8880
ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa	8940
gcgccacgct tcccgaaggg agaaaggcgg acaggatatcc ggtaagcggc agggtcggaa	9000
caggagagcg cacgaggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg	9060
ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc	9120
tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc tggccttttg	9180
ctcacatggc tcgac	9195

<210> 24
 <211> 13
 <212> RNA
 <213> Artificial Sequence

NaDoc 002-PCT-1_seq listing_ST25.txt

<220>
 <223> Synthetic

<220>
 <221> misc_feature
 <222> (7)..(7)
 <223> R is A or G

<400> 24
 gccgccrcca ugg

13

<210> 25
 <211> 7
 <212> RNA
 <213> Homo sapiens

<400> 25
 accaugg

7

<210> 26
 <211> 9
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic

<220>
 <221> MISC_FEATURE
 <222> (1)..(1)
 <223> Xaa is Arginine or Lysine

<220>
 <221> MISC_FEATURE
 <222> (2)..(2)
 <223> Xaa is Leucine or Valine or Isoleucine

<220>
 <221> MISC_FEATURE
 <222> (3)..(7)
 <223> Xaa is any amino acid

<220>
 <221> MISC_FEATURE
 <222> (8)..(8)
 <223> Xaa is Glutamine or Histidine

<220>
 <221> MISC_FEATURE
 <222> (9)..(9)
 <223> Xaa is Leucine or Alanine

<400> 26

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1 5

<210> 27
 <211> 9
 <212> PRT
 <213> Homo sapiens

NaDoc 002-PCT-1_seq listing_ST25.txt

<400> 27

Arg Leu Arg Val Leu Ser Gly His Leu
 1 5

<210> 28

<211> 31

<212> RNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 28

uggagcagag gcucuggcag cuuuugcagc g

31

<210> 29

<211> 21

<212> RNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 29

gcccaaggagc cagagagcau g

21

<210> 30

<211> 11

<212> RNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 30

gcccaaggagc c

11

<210> 31

<211> 5

<212> RNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 31

auuuu

5

<210> 32

<211> 9

<212> RNA

<213> Artificial Sequence

<220>

<223> Synthetic

<220>

<221> misc_feature

<222> (8)..(9)

<223> N is U or A

NaDoc 002-PCT-1_seq listing_ST25.txt

<400> 32
uuuuuuann

9

<210> 33
<211> 4
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 33
auuu

4

<210> 34
<211> 16
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 34
ggcucuuuuc agagcc

16

<210> 35
<211> 6
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 35
uuuuuau

6

<210> 36
<211> 7
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 36
uuuuuau

7

<210> 37
<211> 7
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 37
uuuuuau

7

<210> 38
<211> 9
<212> RNA
<213> Artificial Sequence

NaDoc 002-PCT-1_seq listing_ST25.txt

<220>
 <223> Synthetic

 <400> 38 9
 uuuuuuuuuu

 <210> 39
 <211> 7
 <212> RNA
 <213> Artificial Sequence

 <220>
 <223> Synthetic

 <400> 39 7
 uuuuuuuu

 <210> 40
 <211> 11
 <212> RNA
 <213> Artificial Sequence

 <220>
 <223> Synthetic

 <400> 40 11
 uuuuuuauaaa g

 <210> 41
 <211> 21
 <212> RNA
 <213> Artificial Sequence

 <220>
 <223> Synthetic

 <400> 41 21
 uauuuuuauuc uuggccgccc g

 <210> 42
 <211> 36
 <212> RNA
 <213> Artificial Sequence

 <220>
 <223> Synthetic

 <400> 42 36
 gcggcccgcc guauuuuuau cuuggccgcc cgccac

 <210> 43
 <211> 49
 <212> RNA
 <213> Artificial Sequence

 <220>
 <223> Synthetic

 <400> 43 49
 gauuuuuuuu cgggggggccg ccaagagcgg cccgccguau uauuuuuuu

NaDoc 002-PCT-1_seq listing_ST25.txt

<210> 44
 <211> 21
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 44
 gauaauaaua cggcgggccg c 21

<210> 45
 <211> 27
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 45
 guggcgggcg gccagauaa uauauuu 27

<210> 46
 <211> 25
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 46
 uauuauuaua uggccgccc gccac 25

<210> 47
 <211> 21
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 47
 ggcggccaag auaauauau u 21

<210> 48
 <211> 21
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 48
 cgggcggcca agauauaua a 21

<210> 49
 <211> 725
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

NaDoc 002-PCT-1_seq listing_ST25.txt

<400> 49
 ccgcgccguga gaugcaggua caucccacug augaguccca aauaggacga aacgcgcuc 60
 ggugcgucug ggauuccacu gcuauccacg cggccgcgac cauggaugga accggcgggc 120
 ggccaagaua auaauacacc augggcgcug augauguugu ugaucucucu aaauuuuug 180
 ucauggaaaa cuuuucucg uaccacggga cuaaaccugg uuauguggau uccaaucaaa 240
 aagguauaca aaagccaaaa ucugguacac aaggaaauua ugacgaugau uggaaagggu 300
 uuuauaguac cgacaauaaa uacgacgcug cgggauacuc uguggauaau gaaaaccgc 360
 ucucuggaaa agcuggagggc guggucaaag ugacguaucc aggacugacg aagguucucg 420
 cacuaaaagu ggauaaugcc gaaacuauua agaaagagu agguuuuagu cucacugaac 480
 cguugaugga gcaugcggga acggaagagu uuaucaaaag guucggugau ggugcuucgc 540
 gugugugcu cagccuuccc uucgcugagg ggaguucug cguugaauau auuaauaacu 600
 gggaaacaggc gaaagcguua agcguagaac uugagauuaa uuugaaacc cguggaaaac 660
 guggccaaga ugcgauguau gaguauaugg cucaagccug ugcaggaaau cgugucaggc 720
 gauag 725

<210> 50
 <211> 599
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 50
 gauaaauau caccaugggc gcugaugaug uuguugauuc uucuaaaucu uuugucaugg 60
 aaaacuuuuc uucguaccac gggacuaaac cugguuauu ggauuccauu caaaaaggua 120
 uacaaaagcc aaaaucuggu acacaaggaa auuaugacga ugaauaggaa ggguuuuaua 180
 guaccgacaa uaaauacgac gcugcgggau acucugugga uaaugaaaac ccgcucucug 240
 gaaagcugg aggcgugguc aaagugacgu auccaggacu gacgaagguu cucgcacuaa 300
 aaguggauaa ugccgaaacu auuaagaaag aguuaagguu aagucucacu gaaccguuga 360
 uggagcaagu cggaacggaa gaguuuauca aaagguucgg ugauggugcu ucgcguguag 420
 ugcucagccu ucccuucgcu gaggggaguu cuagcguuga auauauuaa aacugggaac 480
 aggcgaaagc guuaagcguu gaacuugaga uuaauuuuga aacccgugga aaacguggcc 540
 aagaugcgau guaugaguau auggcucaag ccugugcagg aaauuguguc aggcgauag 599

<210> 51
 <211> 21
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 51
 cucuguccac uuggagcccu u

NaDoc 002-PCT-1_seq listing_ST25.txt

<210> 52
 <211> 50
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> synthetic

<400> 52
 ucagaagaga ccuucucugu ccacuuggag cccuuuguau acuccuacug

50

<210> 53
 <211> 50
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> synthetic

<400> 53
 ggaguauaca aagggcucca agggcuccaa guggacagag aaggucucuu

50

<210> 54
 <211> 21
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> synthetic

<400> 54
 ggagcccuuu guauacuccu u

21

<210> 55
 <211> 21
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> synthetic

<400> 55
 ggaguauaca aagggcucca a

21

<210> 56
 <211> 29
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> synthetic

<400> 56
 gggcuccaag uggacagaga aggucucuu

29

<210> 57
 <211> 21
 <212> RNA
 <213> Artificial Sequence

<220>

NaDoc 002-PCT-1_seq listing_ST25.txt

<223> Synthetic

<400> 57

gggcuccaag uggacagaga a

21

<210> 58

<211> 21

<212> RNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 58

aagggcucca aguggacaga g

21

<210> 59

<211> 722

<212> RNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 59

```

ccugucaccg gauguguuuu ccggucugau gaguccguga ggacgaaaca ggaccauggg      60
aaccauggaa agggcuccaa guggacagag aggggugggg guugggggaa ggggaugggc      120
gcugaugaug uuguugauuc uucuaaaucu uuugucaugg aaaacuuuuc uucguaccac      180
gggacuaaac cugguuauuu ggauuccauu caaaaaggua uacaaaagcc aaaaucuggu      240
acacaaggaa auuaugacga ugauuggaaa ggguuuuuaua guaccgacaa uaaauacgac      300
gcugcgggau acucugugga uaaugaaaac ccgcucucug gaaaagcugg aggcgugguc      360
aaagugacgu auccaggacu gacgaagguu cucgcacuaa aaguggauaa ugccgaaacu      420
auuaagaaag aguuagguuu aagucucacu gaaccguuga uggagcaagu cggaacggaa      480
gaguuuauca aaagguucgg ugauggugcu ucgcguguag ugcucagccu ucccuucgcu      540
gaggggaguu cuagcguuga auauuuuauu aacugggaac aggcgaaagc guuaagcgua      600
gaacuugaga uuaauuuuga aacccgugga aaacguggcc aagaugcgau guaugaguau      660
auggcucaag ccugugcagg aaaucguguc aggcgauagc ccuuccccc aacccccacc      720
cc                                                                 722

```

<210> 60

<211> 631

<212> RNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 60

```

guggacagag agggguyaag gggauuggcg cugaugaugu uguugauucu ucuaaaucuu      60
uugucaugga aaacuuuucu ucguaccacg ggacuaaacc ugguuaugug gauuccauuc      120
aaaaagguau acaaaagcca aaaucuggua cacaaggaaa uuaugacgau gauuggaaag      180

```

NaDoc 002-PCT-1_seq listing_ST25.txt

gguuuuuauag uaccgacaau aaauacgacg cugcgggaua cucuguggau aaugaaaacc 240
 cgcucucugg aaaagcugga ggcgugguca aagugacgua uccaggacug acgaagguuc 300
 ucgcacuaaa aguggauaa gccgaaacua uuaagaaaga guuagguua agucucacug 360
 aaccguugau ggagcaaguc ggaacggaag aguuuaucaa aagguucggu gauggugcuu 420
 cgcguguagu gcucagccuu cccuucgcug aggggaguuc uagcguugaa uauauuaaua 480
 acugggaaca ggcgaaagcg uuaagcguag aacuugagau uaauuuugaa acccguggaa 540
 aacguggcca agaugcgaug uaugaguaua uggcucaagc cugugcagga aaucguguca 600
 ggcgauagcc ccuuccccca acccccaccc c 631

<210> 61
 <211> 21
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 61
 uaccaaugcu gcuugugccu g 21

<210> 62
 <211> 51
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 62
 uagcaauaca gcagcuacca augcugcuug ugccuggcua gaagcacaag a 51

<210> 63
 <211> 21
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 63
 caggcacaag cagcauuggu a 21

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

<220>
 <223> Synthetic

<400> 64
 agcauuggua gcugcuguau u 21

<210> 65
 <211> 21

NaDoc 002-PCT-1_seq listing_ST25.txt

<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 65
ugcuucuagc caggcacaag c 21

<210> 66
<211> 681
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 66
augaccaugg gaaccaugcg ccuccgcguu cucucaggcc accucggcca ggcacaagca 60
gcauugguag caugggcgcu gaugauguug uugauucuuc uaaaucuuuu gucauggaaa 120
acuuuuucuuc guaccacggg acuaaaccug guuaugugga uuccauuca aaagguauac 180
aaaagccaaa aucugguaca caaggaaauu augacgauga uuggaaaggg uuuuauagua 240
ccgacaauaa auacgacgcu gcgggauacu cuguggauaa ugaaaacccg cucucuggaa 300
aagcuggagg cguggucaa gugacguauc caggacugac gaagguucuc gcacuaaaag 360
uggauaaugc cgaaacuauu aagaaagagu uagguuuuag ucucacugaa ccguugaugg 420
agcaagucgg aacggaagag uuuaucaaaa gguucgguga uggugcuucg cguguagugc 480
ucagccuucc cuucgcugag gggaguucua gcguugaaua uauuaauaac ugggaacagg 540
cgaaagcguu aagcguagaa cuugagauua auuuugaaac ccguggaaaa cguggccaag 600
augcgaugua ugaguauaug gcucaagccu gugcaggaaa ucgugucagg cgauaggcca 660
ggcacaagca gcauugguag c 681

<210> 67
<211> 21
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 67
ggcacaagca gcauugguau u 21

<210> 68
<211> 610
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 68
agcauuggua gcaugggcgc ugaugauguu guugauucuu cuaaaucuuu ugcauggaa 60
aacuuuuucu cguaccacgg gacuaaaccu gguuaugugg auuccauuca aaaagguaua 120

NaDoc 002-PCT-1_seq listing_ST25.txt

caaaagccaa aaucugguac acaaggaaau uaugacgaug auuggaaagg guuuuuauagu 180
accgacaaua aaucagacgc ugcgggauac ucuguggaua augaaaaccc gcucucugga 240
aaagcuggag ggcuggucaa agugacguau ccaggacuga cgaagguucu cgcacuaaaa 300
guggauaaug ccgaaacuau uaagaaagag uuagguuuuaa gucucacuga accguugaug 360
gagcaagucg gaacggaaga guuuaucaaa agguucggug auggugcuuc gcguguagug 420
cucagccuuc ccuucgcuga ggggaguucu agcguugaau auauuaauaa cugggaacag 480
gcgaaagcgu uaagcguaga acugagauu aauuuugaaa cccguggaaa acguggccaa 540
gaugcgaugu augaguauau ggcucaagcc ugugcaggaa aucgugucag gcgauaggcc 600
aggcacaagc 610

<210> 69
<211> 21
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 69 21
aagcgccggc cggccgcugg u

<210> 70
<211> 40
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 70 40
uccccuaag cgccggccgg ccgcuggucu guuuuuucgu

<210> 71
<211> 63
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 71 60
gaaaaaacag accagcggcc gcagcggccg gccggcgcuu aggccgcugg ucuguuuuuu
cuu 63

<210> 72
<211> 21
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 72 21
cagcggccgg ccggcgcuua g

NaDoc 002-PCT-1_seq listing_ST25.txt

<210> 73
 <211> 21
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 73
 gaaaaaacag accagcggcc g 21

<210> 74
 <211> 21
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 74
 gccgcugguc uguuuuuuucu u 21

<210> 75
 <211> 28
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 75
 uccccuaag cgccggccgg ccgcuggu 28

<210> 76
 <211> 21
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 76
 accagcggcc ggccggcgcu u 21

<210> 77
 <211> 813
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<220>
 <221> misc_feature
 <222> (635)..(716)
 <223> n is a, c, g, or u

<400> 77
 augaccaugg aaccauggac cagcggccgg ccggcgcuuc augggcgcuug augauguugu 60

ugaauucuucu aaauuuuug ucauggaaaa cuuuucuucg uaccacggga cuaaaccugg 120

NaDoc 002-PCT-1_seq listing_ST25.txt

```

uuauugggau uccauucaaa aagguauaca aaagccaaaa ucugguacac aaggaaauua      180
ugacgaugau uggaaagggg uuuauaguac cgacaauaaa uacgacgcug cgggauacuc      240
uguggauaa uaaaacccgc ucucuggaaa agcuggaggc guggucaaag ugacguaucc      300
aggacugacg aagguucucg cacuaaaagu ggauaaugcc gaaacuauua agaaagagu      360
agguuuuagu cucacugaac cguugaugga gcaagucgga acggaagagu uuaucaaaa      420
guucggugau ggugcuucgc gugugugcu cagccuuccc uucgcugagg ggaguucuag      480
cguugaauau auuaauaacu gggaacaggc gaaagcguaa agcguagaac uugagauua      540
uuuugaaacc cguggaaaac guggccaaga ugcgauguau gaguauaugg cucaagccug      600
ugcaggaaa uugucaggc gauaguuuuu uauunnnnnn nnnnnnnnnn nnnnnnnnnn      660
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn      720
auugaccagc ggccggccgg cguuaagaa ccuggagcag aggcucuggc aagcuuuugc      780
agcaggaaa cgcgaagga gccagagagc aug                                          813

```

```

<210> 78
<211> 706
<212> RNA
<213> Artificial Sequence

```

```

<220>
<223> Synthetic

```

```

<220>
<221> misc_feature
<222> (606)..(687)
<223> n is a, c, g, or u

```

```

<400> 78
ccggcgacca gaugggcgcu gaugauguug uugauucuuc uaaauuuuu gucauggaaa      60
acuuuuucuuc guaccacggg acuaaaccug guuaugugga uuccauucaa aaagguauac      120
aaaagccaaa aucugguaca caaggaaaau augacgauga uuggaaaggg uuuuauagua      180
ccgacaauaa auacgacgcu gcggggaucu cuguggauaa ugaaaacccg cucucuggaa      240
aagcuggagg cguggucaa gugacguauc caggacugac gaagguucuc gcacuaaaag      300
uggauaaugc cgaaacuauu aagaaagagu uagguuuuag ucucacugaa ccguugaugg      360
agcaagucgg aacggaagag uuuaucaaaa gguucgguga uggugcuucg cguguagugc      420
ucagccuucc cuucgcugag gggaguucua gcguugaaua uauuaauaac ugggaacagg      480
cgaaagcgua aagcguagaa cuugagauua auuuugaaac ccguggaaaa cguggccaag      540
augcgaugua ugaguauaug gcucaagccu gucaggaaa ucgugucagg cgauaguuuu      600
uuauunnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn      660
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn uauuauucgc ggccgg                      706

```

```

<210> 79
<211> 21

```

NaDoc 002-PCT-1_seq listing_ST25.txt

<212> RNA
 <213> Artificial Sequence

 <220>
 <223> Synthetic

 <400> 79
 aaauaaguu augaacgggu c 21

 <210> 80
 <211> 42
 <212> RNA
 <213> Artificial Sequence

 <220>
 <223> Synthetic

 <400> 80
 cgggaccgggu aaauaaguu uaugaacggg ucgauagcca gc 42

 <210> 81
 <211> 36
 <212> RNA
 <213> Artificial Sequence

 <220>
 <223> Synthetic

 <400> 81
 ggcuauccug augaggccga aaggccgaaa cccguu 36

 <210> 82
 <211> 32
 <212> RNA
 <213> Artificial Sequence

 <220>
 <223> Synthetic

 <400> 82
 cgggaccgggu aaauaaguu uaugaacggg uc 32

 <210> 83
 <211> 27
 <212> RNA
 <213> Artificial Sequence

 <220>
 <223> Synthetic

 <400> 83
 cccguucaua aacuuauua accgguc 27

 <210> 84
 <211> 36
 <212> RNA
 <213> Artificial Sequence

 <220>
 <223> Synthetic

 <400> 84
 gaacgggcug augaggccga aaggccgaaa cccguu 36

NaDoc 002-PCT-1_seq listing_ST25.txt

<210> 85
 <211> 21
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 85
 cccguucaua aacuuauua a 21

<210> 86
 <211> 21
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 86
 gacccguuca uaaacuuaau u 21

<210> 87
 <211> 657
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 87
 augaccaugg accauggucg acccguucau aaacuuaauu ggcaaugggc gcugaugaug 60
 uuguugauuc uucuaaauc uugucaugg aaaacuuauc uucguaccac gggacuaaac 120
 cugguuauug ggaauccauu caaaaaggua uacaaaagcc aaaaucuggu acacaaggaa 180
 auuaugacga ugauuggaaa ggguuuuaua guaccgacaa uaaauacgac gcugcgggau 240
 acucugugga uaaugaaaac ccgcucucug gaaaagcugg aggcgugguc aaagugacgu 300
 auccaggacu gacgaagguu cugcacuaa aaguggauaa ugccgaaacu auuaagaaag 360
 aguuagguuu aagucucacu gaaccguuga uggagcaagu cggaacggaa gaguuauca 420
 aaagguucgg ugauggugcu ucgcguguag ugcucagccu ucccuucgu gaggggaguu 480
 cuagcguuga auauauuaa aacugggaac aggcgaaagc guuaagcgua gaacuugaga 540
 uuaauuuuga aacccgugga aaacguggcc aagaugcgau guaugaguau auggcucaag 600
 ccugucagg aaaucguguc aggcgauagc ccaaaggcuc uuucagagc ccccuca 657

<210> 88
 <211> 627
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 88
 aaacuuaauu ggcaaugggc gcugaugaug uuguugauuc uucuaaauc uugucaugg 60

NaDoc 002-PCT-1_seq listing_ST25.txt

```

aaaacuuuuc uucguaccac gggacuaaac cugguuauugu ggauuccauu caaaaaggua      120
uacaaaagcc aaaaucuggu acacaaggaa auuaugacga ugauuggaaa ggguuuuaua      180
guaccgacaa uaaauacgac gcugcgggau acucugugga uaaugaaaac ccgcucucug      240
gaaaagcugg aggcgugguc aaagugacgu auccaggacu gacgaagguu cucgcacuaa      300
aaguggauaa ugccgaaacu auuaagaaag aguuagguuu aagucucacu gaaccguuga      360
uggagcaagu cggaacggaa gaguuuauca aaagguucgg ugauggugcu ucgcguguag      420
ugcucagccu ucccuucgcu gaggggaguu cuagcguuga auauauuaau aacugggaac      480
aggcgaaagc guuaagcgua gaacuugaga uuaauuuuga aacccgugga aaacguggcc      540
aagaugcgau guaugaguau auggcucaag ccugugcagg aaaucguguc aggcgauagc      600
ccaaaggcuc uuucagagc ccccccua      627

```

```

<210> 89
<211> 23
<212> RNA
<213> Artificial Sequence

```

```

<220>
<223> Synthetic

```

```

<400> 89
cugaugaggc cgaaaggccg aaa      23

```

```

<210> 90
<211> 33
<212> RNA
<213> Artificial Sequence

```

```

<220>
<223> Synthetic

```

```

<400> 90
accagagaaa cacacguugu gguauuuuac ugg      33

```

```

<210> 91
<211> 52
<212> RNA
<213> Artificial Sequence

```

```

<220>
<223> Synthetic

```

```

<400> 91
ccugucaccg gauguguuuu ccggucugau gaguccguga ggacgaaaca gg      52

```

```

<210> 92
<211> 97
<212> RNA
<213> Artificial Sequence

```

```

<220>
<223> Synthetic

```

```

<400> 92
ccgcggcuga gaugcaggua caucccacug augaguccca aaaggacga aacgcgcuuc      60

```

NaDoc 002-PCT-1_seq listing_ST25.txt

ggugcgucug ggauuccacu gcuauccacg cggccgc 97

<210> 93
 <211> 29
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 93
 ugggauucca cugcuaucca cgcggccgc 29

<210> 94
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic RNA/DNA

<400> 94
 aaacaugcag aaaaugcugt t 21

<210> 95
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic RNA/DNA

<400> 95
 cagcauuuuc ugcauguut t 21

<210> 96
 <211> 23
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic RNA/DNA

<400> 96
 aagcuaccug uucauggcc att 23

<210> 97
 <211> 23
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic RNA/DNA

<400> 97
 tggccatgga acaggtagct ttt 23

<210> 98
 <211> 21
 <212> RNA
 <213> Artificial Sequence

NaDoc 002-PCT-1_seq listing_ST25.txt

<220>
<223> Synthetic

<400> 98
gaguucccgga cgguccuag c 21

<210> 99
<211> 21
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 99
uaggacgcu cggaacucg c 21

<210> 100
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 100
aactacacaa atcagcgatt t 21

<210> 101
<211> 21
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 101
cuacacaaau cagcgauuuu u 21

<210> 102
<211> 21
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 102
aaaucgcuga uuuguguagu u 21

<210> 103
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 103
aaaacttacg ctgagtactt cgatct 26

NaDoc 002-PCT-1_seq listing_ST25.txt

<210> 104
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic RNA/DNA

<400> 104
 cuuacgcuga guacuucgat t 21

<210> 105
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic RNA/DNA

<400> 105
 ucgaaguacu cagcguaagt t 21

<210> 106
 <211> 28
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 106
 ctacggcaag ctgaccctga agttcatc 28

<210> 107
 <211> 22
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 107
 gcaagcugac ccugaaguuc au 22

<210> 108
 <211> 22
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

<400> 108
 gaacuucagg gucagcuugc cg 22

<210> 109
 <211> 24
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic

NaDoc 002-PCT-1_seq listing_ST25.txt

<400> 109
aatcgcttac cgattcagaa tcgc 24

<210> 110
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic RNA/DNA

<400> 110
ucgcuuaccg auucagaaut t 21

<210> 111
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic RNA/DNA

<400> 111
auucugaauc gguaagcgat t 21

<210> 112
<211> 7
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<222> (1)..(1)
<223> R is A or G

<220>
<221> misc_feature
<222> (2)..(3)
<223> n is a, c, g, or u

<400> 112
rnnauagg 7

<210> 113
<211> 47
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 113
auuuuuuuua uuauuuuuua uuauuuuuuu uuauuuuuuu auuuuuuu 47

<210> 114
<211> 4
<212> RNA
<213> Artificial Sequence

NaDoc 002-PCT-1_seq listing_ST25.txt

<220>

<223> Synthetic

<400> 114

aaga

4