Circular RNA and Vectors thereof

Technical field

The present invention relates generally to the field of molecular biology. In particular, the invention is directed to a vector for generating a circular RNA. Methods for generating circular RNA from a precursor RNA transcribed from the vector are also provided herein.

Background

Gene therapy has witnessed a paradigm shift with the emergence of circular mRNA as a promising tool. Circular mRNA, distinct from its linear counterpart, offers unique advantages in the field of gene therapy. Its circular structure not only enhances stability and resistance to degradation, but also enables prolonged translation in target cells. This opens new avenues for sustained therapeutic protein expression. Furthermore, circular mRNA can be engineered to contain specific regulatory elements that fine-tune expression levels, duration, and tissue specificity. These attributes address some of the challenges posed by traditional linear mRNA delivery, such as transient expression and the need for repeated administrations. The use of circular mRNA holds immense potential for precise and durable gene therapies, revolutionizing the treatment landscape for various genetic and acquired disorders. As research in this field progresses, harnessing the capabilities of circular mRNA could usher in a new era of safer, more effective, and long-lasting gene therapies.

Recent studies have pioneered an effective technique to engineer circular mRNA molecules by utilizing group 1 introns, a class of ribozymes known for their self-splicing capability. This method involves exploiting the intrinsic properties of group 1 introns to catalyze self-catalytic circularization of linear precursors containing permuted group 1 introns. Specifically, they utilized group 1 introns from *Anabaena* pre-tRNA, featuring a homologous domain and a spacer flanking each ribozyme fragments. Their strategic design facilitated effective circularization of gene sizes greater than 5000 nt, achieving an impressive circularization efficiency of up to 90%. Nevertheless, the robust catalytic efficacy of group 1 introns comes at a cost. It demands elevated temperatures, GTP as an energy source, and bivalent salts to sustain ribozyme activity. This circumstance results in

over 20% of undesired nicks, as evidenced by recent study. Regrettably, these nicks can trigger potent immunogenic responses, making them unsuitable for downstream applications like gene therapy or vaccine development. Considering these limitations, alternative circularization strategies have been explored, including the use of RNA ligases and Twister ribozymes. However, it is noteworthy that, as of now, none of these alternatives surpasses the efficiency and yield demonstrated by the group 1 Anabaena pre-tRNA approach.

It would be desirable to overcome or alleviate at least one of the above-described problems, or at least to provide a useful alternative.

Summary

Disclosed herein is a vector for generating a circular RNA, the vector comprising the following elements operably connected to each other and arranged in the following sequence: a) a modified 5' Group II intron fragment comprising a 3' splice site and a first purification fragment sequence positioned at or close to the splicing junction point, b) a nucleic acid sequence comprising i) an internal ribosome entry site (IRES) region and ii) a protein coding or noncoding region, and c) a modified 3' Group II intron fragment comprising a 5' splice site and a second purification fragment sequence positioned at or close to the splicing junction point, wherein the second purification sequence comprises an intron-binding site 1 (IBS1) sequence that is complementary to an exon-binding site 1 (EBS1) sequence in the modified 3' Group II intron, wherein the first and second purification fragment sequences are capable of being retained in the circular RNA upon Group II intron-mediated back-splicing to form an RNA aptamer.

Disclosed herein is a linear RNA precursor encoded by a vector as defined herein.

Disclosed herein is a circular RNA produced by a vector as defined herein.

Disclosed herein is a method of purifying circular RNA from a sample, the method comprising: a) providing a vector as defined herein; b) performing in vitro transcription to obtain a linear RNA precursor from the vector; c) incubating the linear RNA precursor under conditions to form a circular RNA; and d) purifying the circular RNA.

Disclosed herein is a method of expressing protein coding or non-coding region in a cell, said method comprising transfecting the circular RNA as defined herein into the cell.

Disclosed herein is a nanoparticle composition comprising a circular RNA produced by a vector as defined herein.

Brief description of the drawings

Embodiments of the present invention will now be described, by way of non-limiting example, with reference to the drawings in which:

Figure 1. Group 1 intron- mediated RNA circularization yields circular RNAs of 3-5 kb in size, with 10-30% exhibiting nicks or breaks. (A)(B) Electrophoresis analysis of 200 ng in vitro transcribed (IVT) products from a 3 kb group 1 intron RNA at various time points (1 hr, 2 hrs, and 3 hrs), using different concentrations of MgCl2 (1.5 mM, 15 mM, and 30 mM). The IVT products were resolved on both 1% E-Gel EX and 1% Agarose gel. A control (ctrl) group from 15 mM MgCl2 IVT was included to facilitate comparison of the percentages of circular, precursor, and nicked RNA species (A). The proportions of nicked and circular 3 kb RNA were quantified using ImageJ software and graphically represented using GraphPad Prism (B). (C)(D) Electrophoretic analysis of 200 ng in vitro transcribed (IVT) and folded (FB) products of 5 kb group 1 intron RNA, utilizing different MgCl2 concentrations (5 mM, 10 mM, 15 mM, 30 mM), on a 1% E-Gel EX. Successive images at 10-minute and 20-minute intervals during electrophoresis were captured, with circular, precursor, and nicked RNA species indicated by corresponding arrows (C). The proportions of nicked and circular 5 kb RNA were quantified using ImageJ software, and the data were visualized through graphical representation using GraphPad Prism (D).

Figure 2. Engineered group 2 intron circularized at 90% and contains <2% nicked byproducts. (A) Schematic representations of different designs of permuted intron-exon (PIE) circular RNA utilizing various versions (version 1-5) of group 2 intron. Elements, such as 5'G2i and 3'G2i (depicted in orange), spacer (in green), IRES-GOI (in blue), and homologous domains (in grey and black), were illustrated. The corresponding average circular efficiency (Circular%) values were provided on the right side of the schematic. (B) Electrophoretic analysis of 100 ng *in vitro* transcribed (IVT) and folded (FE) products of 2

kb group 2 intron RNA (version 1-5) on a 1% Agarose gel, precursor, circular, and spliced intron RNA species indicated by corresponding arrows on the right. (C)(E) Electrophoresis analysis of 100 ng *in vitro* transcribed (IVT), folded (FE) and RNaseR (RR) products of 2 kb (C) and 5 kb (E) group 2 intron RNA (version 5), precursor, circular, and spliced intron RNA species indicated by corresponding arrows on the right, experiment were repeated 3 times. (D)(F) Comparative analysis of IVT, FE, and RR-treated group 1 (AN_2K, AN_5K) and group 2 (TET_2K, TET_5K) circular RNA, separated on a 1% E-Gel EX. Circular and nicked RNA species were highlighted with arrows.

Figure 3. Generation of scarless (SL) and streptavidin aptamer (S1) tagged circular mRNA using permuted group 2 introns at the junction point. (A) Schematic depictions of scarless (SL) and S1-tagged (S1) circular RNA generated through PIE group 2 intron. (B) Visualization of IVT, FE and one-step purified circular SL/S1 GLUC RNA on a 2% E-gel EX. Circular RNA species are demarcated by arrows, while linear TETH RNA serves as a nonbinding control. (C) One-step purification of S1-tagged 2k (GLUC) circular RNA using the NicOPURE platform. Precursor, circular and spliced intron RNA species were visualized on 1% agarose gel, with respective species indicated by arrows. (D) One-step purification of S1-tagged 2k (GLUC) circular RNA using the NicOPURE platform. Precursor, circular and spliced intron RNA species were visualized on 1% E-gel EX. Arrows indicate the respective RNA species. (E) One-step purification of S1-tagged 5k (SPIKE) circular RNA using the NicOPURE platform. Precursor, circular and spliced intron RNA species were visualized on 1% agarose gel, with respective species indicated by arrows. (F) 1% E-gel EX visualization of one-step purification of S1-tagged 5k (SPIKE) circular RNA using the NicOPURE platform.

Figure 4. Structure probing indicates Group 1 and Group 2 intron- generated circular RNA have similar secondary structures *in vitro*. (A) Comparative SHAPE-Map reactivity profiles of circGLUC generated through group 1 (top, circGLUC-AN) or group 2 (bottom, circGLUC-TET) PIE constructs. The SHAPE reactivity is determined using the formula $[(\text{Modified}_{\text{MutR}} - \text{Untreated}_{\text{MutR}})/\text{Denatured}_{\text{MutR}}]$, and subsequently normalized using a model-free box-plot methodology. Regions with reactivities below 0.4 are depicted in black, those between 0.4 and 0.8 in yellow, and reactivities exceeding 0.80 in red. The standard deviation (SD) of SHAPE reactivities is calculated, serving as an indicator of the methods' efficacy in distinguishing single- or double-stranded regions. (B)(C) Secondary structures of

group 1-generated circGLUC (B) and group 2- (NicOPURE) generated circGLUC RNA (C) were drawn using RNA structure-ProbKnot with SHAPE reactivity annotated in colored bases (-0.15 blue to 1.5 red for better visualization).

Figure 5. Significant boost in translation efficiency from NicOPURE generated circular mRNA due to reduced immunogenic responses. (A)(B) Assessment of group 1-generated circGLUC and group 2- (NicOPURE) generated circGLUC RNA through *in vitro* transcription/translation method (IVTT). Luminescence measurements were conducted either after 2 hours of IVTT (A) or monitored throughout the IVTT process (B). (C)(D) Hela cells were transfected with group 1- (spin columns) or group 2- (NicOPURE) generated circGLUC RNA, culture media were collected at indicated time points (1-5 days) for luminescence measurements (C), or IFN-a ELISA measurement (D). Statistical significance was determined using a two-tailed t test: ***$p < 0.001$.

Figure 6. Schematics of Group 1 and Group 2 intron secondary structure. (A) Schematic for group 1 intron, including the 5' splicing site (5' ss), the 3' splicing site (3' ss), the pair regions (P1 to P9), the GTP binding site at P7 and the cutting site for generation of PIE at P6. (B) Schematic of group 2 intron, including the 5' splicing site (5' ss), the 3' splicing site (3' ss), the exon binding sites (EBS1,2), the intron binding sites (IBS1,2), six typical stem-loop domains (D1 to D6), the adenine bulge is located at D6 and the cutting site for generation of PIE is at D4.

Figure 7. RNA circularization methods and conditions. (A) Schematic for conventional circular RNA production and purification steps. (B) Schematic for the NicOPURE platform technology. (C) Table for RNA circularization buffer recipe, temperature and incubation time.

Figure 8. *Flaviviridae* cyclizing UTRs significantly boost fluorescent RNA aptamer Broccoli folding by over sixfold. (A) *In vitro* fluorescence assay using the Flaviviridae cyclization UTRs fused to Broccoli (Denv_cUTR_Broc) or the Dengue cyclization sequence alone (Denv_cUTR_mock). Broccoli (Broc and repeated Broccoli (2xBroc) was also included for control. (B) *In vitro* fluorescence assay of various RNA concentrations of Denv_cUTR_Broc and Broccoli, Michaelis constant (Km) was calculated to be 2.19 and 1.13 for Denv_cUTR_Broc and Broccoli respectively.

Figure 9. Pulldown efficiency testing using affinity aptamers. (A) Schematic for Sephadex D8 tag and Streptavidin S1 tag. (B) S1 tag mediated RNA pulldown testing using a linear 300nt RNA tagged either at 5'- or 3'- end, the RNA bound to Dynabead MyOne Streptavidin C1 beads was eluted with Monarch RNA binding buffer (NEB, Cat.: T2040) and visualized on 2% agar gel. (C) S1 tag mediated RNA pulldown efficiency tested across 7 types of Streptavidin Magnetic beads from N-Lab, and the Streptavidin C1, M280 beads from Invitrogen.

Figure 10. Generation of group 1 and group 2 intron circular mRNA SHAPE-Map library. (A) 200ng of the 100mg group 1-generated circGLUC (AN 2K) and group 2-generated circGLUC (TET 2K) were visualized on 2% EX Gels. (B) Small scale PCR determining the cycle number for generating optimum cDNA library fragment size (200-300bp) for Illumina deep-sequencing. (C) Tapestation analysis of group 1-generated and group 2-generated circGLUC, before and after the fragmentation step in SHAPE-Map. (D) Sample name, index numbers, and final cDNA yield of library submitted for deep sequencing.

Figure 11. Quality assessment of group 1 and group 2 intron circular mRNA SHAPE-Map sequencing data. (A)(B) Sequencing depth of NAI modified, untreated or denatured group 1 circGLUC (A) or group 2 circGLUC (B). (C)(D) Mutation rate (%) of NAI modified, untreated or denatured group 1 circGLUC (C) or group 2 circGLUC (D). (E)(F) Reactivity distribution and Standard errors (Stderrs) of NAI modified, untreated or denatured group 1 circGLUC (E) or group 2 circGLUC (F).

Figure 12. Group 2 intron folding condition comparison. (A) Circular mRNA precursor self-splicing efficiency comparison under Fold E, with either increase concentration of monovalent salt MgSo4 (10mM, 100mM, 200mM, 225mM), or different cofactors (ATP, CTP, GTP, UTP) at 100mM MgSo4.

Detailed description

The present specification teaches a vector for generating a circular RNA, the vector comprising the following elements operably connected to each other and arranged in the

following sequence: a) modified 5' Group II intron fragment comprising a 3' splice site, b) a nucleic acid sequence comprising i) an internal ribosome entry site (IRES) region and ii) a protein coding or noncoding region, and d) a modified 3' Group II intron fragment comprising a 5' splice site.

Disclosed herein is a vector for generating a circular RNA, the vector comprising the following elements operably connected to each other and arranged in the following sequence: a) a modified 5' Group II intron fragment comprising a 3' splice site and a first purification fragment sequence positioned at or close to the splicing junction point, b) a nucleic acid sequence comprising i) an internal ribosome entry site (IRES) region and ii) a protein coding or noncoding region, and c) a modified 3' Group II intron fragment comprising a 5' splice site and a second purification fragment sequence positioned at or close to the splicing junction point, wherein the second purification sequence comprises an intron-binding site 1 (IBS1) sequence that is complementary to an exon-binding site 1 (EBS1) sequence in the modified 3' Group II intron, wherein the first and second purification fragment sequences are capable of being retained in the circular RNA upon Group 2 intron-mediated back-splicing of the modified 5' and 3' Group II intron fragments to form an RNA aptamer.

Without being bound by theory, the inventors introduced modified group 2 introns from *Clostridium tetani*, incorporating the upstream translation regions (UTRs) identified in *Flaviviridae*. These modifications elevated the circularization efficiency of Clostridium tetani's group 2 introns from 50% to an impressive 90%, comparable to the efficiency achieved by Anabaena's group 1 intron. Notably, the heightened efficiency of circularization with group 2 introns also carries distinct advantages. The enhanced efficiency within the context of group 2 introns does not entail the production of undesirable RNA byproducts with nicks. Consequently, this improvement contributes to an increased overall yield and purity of circular RNA. Moreover, the inherent interchangeable sequences of intron-binding sites (IBS) and exon-binding sites (EBS) at the splicing junctions of group 2 introns offer a distinctive advantage. This enables the creation not only of seamless circular RNA but also of circular RNA embedded with a distinct positive selection tag, formed precisely at the junction point following group 2 intron-mediated back-splicing. This innovation leads to an efficient and streamlined downstream purification process, utilizing a single-step affinity-based positive selection. This approach has the potential to seamlessly integrate into the

existing linear mRNA purification workflows within the industrial setting, thereby enhancing process efficiency.

In one embodiment, there is provided a vector for generating a circular RNA, the vector comprising the following elements operably connected to each other and arranged in the following sequence: a) a modified 5' Group II intron fragment comprising a 3' splice site and a first purification fragment sequence positioned at or close to the splicing junction point, b) a nucleic acid sequence comprising i) an internal ribosome entry site (IRES) region and ii) a protein coding or noncoding region, and c) a modified 3' Group II intron fragment comprising a 5' splice site and a second purification fragment sequence positioned at or close to the splicing junction point, wherein the first and second purification fragment sequences are capable of being retained in the circular RNA upon Group 2 intron-mediated back-splicing of the modified 5' and 3' Group II intron fragments to form an RNA aptamer. The second purification sequence may comprise an intron-binding site 1 (IBS1) sequence that is complementary to an exon-binding site 1 (EBS1) sequence in the modified 3' Group II intron.

The terms "complementary" and "complementarity" refers to the relationship between two nucleic acid sequences or nucleic acid monomers having the capacity to form hydrogen bond(s) with one another by either traditional Watson-Crick base-paring or other non-traditional types of pairing. The degree of complementarity between two nucleic acid sequences can be indicated by the percentage of nucleotides in a nucleic acid sequence which can form hydrogen bonds (e.g., Watson-Crick base pairing) with a second nucleic acid sequence (e.g., about 50%, about 60%, about 70%, about 80%, about 90%, and 100% complementary). Two nucleic acid sequences are "perfectly complementary" if all the contiguous nucleotides of a nucleic acid sequence will hydrogen bond with the same number of contiguous nucleotides in a second nucleic acid sequence. Two nucleic acid sequences are "substantially complementary" if the degree of complementarity between the two nucleic acid sequences is at least 60% (e.g., at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, at least 99%, or 100%) over a region of at least 8 nucleotides (e.g., at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 30, at least 35, at least 40, at least 45, at least 50, or more nucleotides), or if the two nucleic acid sequences hybridize under at

least moderate, or, in some embodiments high, stringency conditions.

As used herein, the term "hybridization" or "hybridizes" refers to the process in which two single-stranded polynucleotides bind non-covalently to form a stable double-stranded polynucleotide, i.e., a duplex. The term "hybridization" may also refer to triple-stranded hybridization. The resulting (usually) double-stranded polynucleotide is a "hybrid". The proportion of the population of polynucleotides that forms stable hybrids is referred to herein as the "degree of hybridization".

Hybridization conditions will typically include salt concentrations of less than about 1M, more usually less than about 500 mM and less than about 200 mM. Hybridization temperatures are typically greater than 22°C, more typically greater than about 30°C, and preferably in excess of about 37°C. In one embodiment, hybridization takes place at about 65°C. Hybridizations are usually performed under stringent conditions, i.e., conditions under which a probe will hybridize to its target. Stringent conditions are sequence-dependent and are different under different circumstances. Longer fragments may require higher hybridization temperatures for specific hybridization. As other factors may affect the stringency of hybridization, including base composition and length of the complementary strands, presence of organic solvents and extent of base mismatching, the combination of parameters is more important than the absolute measure of any one alone. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength, pH and nucleic acid composition) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium.

The 5' and 3' Group II intron fragments may, for example, be from *Clostridium tetani, Pylaiella littoralis, Bacillus subtilis, Escherichia coli, Lactococcus lactis, Pseudomonas aeruginosa, Salmonella enterica, Staphylococcus aureus, Saccharomyces cerevisiae, Streptococcus pneumoniae, Synechocystis sp. PCC 6803, Thermus thermophiles, Yersinia pestis, Saccharomyces cerevisiae* or *Arthrospira platensis*. In one embodiment, the modified 5' and 3' Group II intron fragments are from *Clostridium tetani*.

As used herein, a "vector" means a piece of DNA, that is synthesized (e.g., using PCR), or that is taken from a virus, plasmid, or cell of a higher organism into which a foreign DNA

fragment can be or has been inserted for cloning and/or expression purposes. In some embodiments, a vector can be stably maintained in an organism. A vector can comprise, for example, an origin of replication, a selectable marker or reporter gene, such as antibiotic resistance or GFP, and/or a multiple cloning site (MCS). The term includes linear DNA fragments (e.g., PCR products, linearized plasmid fragments), plasmid vectors, viral vectors, cosmids, bacterial artificial chromosomes (BACs), yeast artificial chromosomes (YACs), and the like.

As used herein, the elements of a vector are "operably connected" if they are positioned on the vector such that they can be transcribed to form a precursor RNA that can then be circularized into a circular RNA using the methods provided herein.

As used herein, the elements of a vector are "operably connected" if they are positioned on the vector such that they can be transcribed to form a precursor RNA that can then be circularized into a circular RNA using the methods provided herein.

As used herein, "precursor RNA" refers to a linear RNA molecule created by *in vitro* transcription (e.g., from a vector provided herein). This precursor RNA molecule contains the entirety of the circRNA sequence, plus splicing sequences (e.g. intron fragments or homology sequences) necessary to circularize the RNA. These splicing sequences (intron fragments or homology sequences) are substantially removed from the precursor RNA during circularization, yielding circRNA plus two intron/homology sequence linear RNA fragments. Precursor RNA can be unmodified, partially modified or completely modified. In one embodiment, the precursor RNA contains only naturally occurring nucleotides.

The terms "polynucleotide," "polynucleotide sequence," "nucleotide sequence," "nucleic acid" or "nucleic acid sequence as used herein designate mRNA, RNA, cRNA, cDNA or DNA. The term typically refers to polymeric form of nucleotides of at least 10 bases in length, either ribonucleotides or deoxynucleotides or a modified form of either type of nucleotide. The term includes single and double stranded forms of RNA or DNA.

As used herein, an "RNA" refers to a ribonucleic acid that may be naturally or non-naturally occurring. For example, an RNA may include modified and/or non-naturally occurring components such as one or more nucleobases, nucleosides, nucleotides, or linkers. An RNA

may include a cap structure, a chain terminating nucleoside, a stem loop, a polyA sequence, and/or a polyadenylation signal. An RNA may have a nucleotide sequence encoding a polypeptide of interest. For example, an RNA may be a messenger RNA (mRNA). Translation of an mRNA encoding a particular polypeptide, for example, *in vivo* translation of an mRNA inside a mammalian cell, may produce the encoded polypeptide. RNAs may be selected from the non-liming group consisting of small interfering RNA (siRNA), asymmetrical interfering RNA (aiRNA), microRNA (miRNA), Dicer-substrate RNA (dsRNA), small hairpin RNA (shRNA), mRNA, and mixtures thereof.

As used herein, "modified" means non-natural. For example, an RNA may be a modified RNA. That is, an RNA may include one or more nucleobases, nucleosides, nucleotides, or linkers that are non-naturally occurring. A "modified" species may also be referred to herein as an "altered" species. Species may be modified or altered chemically, structurally, or functionally. For example, a modified nucleobase species may include one or more substitutions that are not naturally occurring.

"Polypeptide," "peptide," "protein" and "proteinaceous molecule" are used interchangeably herein to refer to molecules comprising or consisting of a polymer of amino acid residues and to variants and synthetic analogues of the same.

As used herein, the term "encode" refers broadly to any process whereby the information in a polymeric macromolecule is used to direct the production of a second molecule that is different from the first. The second molecule may have a chemical structure that is different from the chemical nature of the first molecule.

In one embodiment, the junction sequence of the modified 5' Group II intron fragment has been replaced with the first purification fragment sequence and the junction sequence of the modified 3' Group II intron fragment has been replaced with the second purification fragment sequence.

The 3' Group II intron fragment may be modified to have an EBS1 sequence that is complementary to the IBS1 sequence in the second purification fragment sequence.

In one embodiment, the modified 5' Group II intron fragment comprises or consists of a nucleic acid sequence having at least 70% sequence identity to SEQ ID NO: 3 or 5 and

wherein the modified 3' Group II intron fragment comprises or consists of a nucleic acid sequence having at least 70% sequence identity to SEQ ID NO: 4 or 6.

In one embodiment, the modified 5' Group II intron fragment comprises or consists of a nucleic acid sequence having at least 70% sequence identity to SEQ ID NO: 3 and wherein the modified 3' Group II intron fragment comprises or consists of a nucleic acid sequence having at least 70% sequence identity to SEQ ID NO: 4.

In one embodiment, the modified 5' Group II intron fragment comprises or consists of a nucleic acid sequence having at least 70% sequence identity to SEQ ID NO: 5 and wherein the modified 3' Group II intron fragment comprises or consists of a nucleic acid sequence having at least 70% sequence identity to SEQ ID NO: 6.

The vector may further comprise at least one (or at least two) homology sequences.

In one embodiment, the vector further comprises at least one homology sequence positioned 5' (i.e. upstream) of the modified 5' Group II intron fragment and at least one homology sequence positioned 3' (i.e. downstream) of the modified 3' Group II intron fragment.

In one embodiment, the vector further comprises at least one homology sequence positioned 3' (i.e. downstream) of the modified 5' Group II intron fragment and at least one homology sequence positioned 5' (i.e. upstream) of the modified 3' Group II intron fragment.

The at least one homology sequence may be a Flaviviridae UTR sequence. The at least one homology sequence may be a Dengue UTR sequence.

In one embodiment, the vector further comprises a nucleic acid sequence having at least 70% sequence identity to SEQ ID NO: 9 positioned 5' (i.e. upstream) of the modified 5' Group II intron fragment and a nucleic acid sequence having at least 70% sequence identity to SEQ ID NO: 10 3' (i.e. downstream) of the modified 3' Group II intron fragment.

The term "RNA aptamer" may refer to a single-stranded or double-stranded RNA molecule that has been engineered to bind to a specific target molecule with high affinity and selectivity. For example, the RNA aptamer may have been engineered to bind to a

streptavidin or sephadex molecules.

The RNA aptamer may be an S1 streptavidin aptamer or a D8 sephadex aptamer.

The term "at least 70% sequence identity" may refer to at least 70%, 80%, 90%, 95%, 99% or 100% sequence identity.

Terms used to describe sequence relationships between two or more polynucleotides or polypeptides include “reference sequence,” “comparison window”, “sequence identity,” “percentage of sequence identity” and “substantial identity”. A “reference sequence” is at least 12 but frequently 15 to 18 and often at least 25 monomer units, inclusive of nucleotides and amino acid residues, in length. Because two polynucleotides may each comprise (1) a sequence (i.e., only a portion of the complete polynucleotide sequence) that is similar between the two polynucleotides, and (2) a sequence that is divergent between the two polynucleotides, sequence comparisons between two (or more) polynucleotides are typically performed by comparing sequences of the two polynucleotides over a “comparison window” to identify and compare local regions of sequence similarity. A “comparison window” refers to a conceptual segment of at least 6 contiguous positions, usually about 50 to about 100, more usually about 100 to about 150 in which a sequence is compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. The comparison window may comprise additions or deletions (i.e., gaps) of about 20% or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Optimal alignment of sequences for aligning a comparison window may be conducted by computerized implementations of algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package Release 7.0, Genetics Computer Group, 575 Science Drive Madison, WI, USA) or by inspection and the best alignment (i.e., resulting in the highest percentage homology over the comparison window) generated by any of the various methods selected. Reference also may be made to the BLAST family of programs as for example disclosed by Altschul et al., 1997, Nucl. Acids Res.25:3389. A detailed discussion of sequence analysis can be found in Unit 19.3 of Ausubel et al., “Current Protocols in Molecular Biology”, John Wiley & Sons Inc, 1994-1998, Chapter 15.

The vector as defined herein may comprises at least one spacer domain.

In one embodiment, the vector comprises i) a first spacer domain positioned between the modified 5' Group II intron fragment and the IRES sequence, and ii) a second spacer domain positioned between the protein coding or noncoding region and the modified 3' Group II intron fragment.

In one embodiment, the vector comprises i) a first spacer domain positioned between the modified 5' Group II intron fragment and the protein coding or noncoding region, and ii) a second spacer domain positioned between the IRES sequence and the modified 3' Group II intron fragment.

As used herein, a "spacer" refers to a region of a polynucleotide sequence ranging from 1 nucleotide to hundreds or thousands of nucleotides separating two other elements along a polynucleotide sequence. The sequences can be defined or can be random. A spacer is typically non-coding. In some embodiments, spacers include duplex forming regions. The "spacer" may refer to any contiguous nucleotide sequence that is 1) predicted to avoid interfering with proximal structures, for example, from the IRES, coding or noncoding region, or intron 2) at least 7 nucleotides long (and optionally no longer than 100 nucleotides) 3) located downstream of and adjacent to the 3' intron fragment and/or upstream of and adjacent to the 5' intron fragment and/or 4) contains one or more of the following: a) an unstructured region at least 5nt long b) a region predicted base pairing at least 5nt long to a distal (i.e., non-adjacent) sequence, including another spacer, and/or c) a structured region at least 7nt long limited in scope to the sequence of the spacer.

As used herein, "interfering" with regard to sequences refers to sequence(s) predicted or empirically determined to alter the folding of other structures in the RNA, such as the IRES or group II intron-derived sequences.

As used herein, "unstructured" with regard to RNA refers to an RNA sequence that is not predicted by the RNAFold software or similar predictive tools to form a structure (e.g., a hairpin loop) with itself or other sequences in the same RNA molecule.

As used herein, "structured" with regard to RNA refers to an RNA sequence that is predicted by the RNAFold software or similar predictive tools to form a structure (e.g., a hairpin loop) with itself or other sequences in the same RNA molecule.

In some embodiments, the vector comprises an IRES sequence. The IRES sequence can be selected from, but not limited to, an IRES sequence of a Taura syndrome virus, Triatoma virus, Theiler's encephalomyelitis virus, simian Virus 40, Solenopsis invicta virus 1, Rhopalosiphum padi virus, Reticuloendotheliosis virus, fuman poliovirus 1, Plautia stali intestine virus, Kashmir bee virus, Human rhinovirus 2, Homalodisca coagulata virus- 1, Human Immunodeficiency Virus type 1, Homalodisca coagulata virus- 1, Himetobi P virus, Hepatitis C virus, Hepatitis A virus, Hepatitis GB virus, foot and mouth disease virus, Human enterovirus 71, Equine rhinitis virus, Ectropis obliqua picorna-like virus, Encephalomyocarditis virus (EMCV), Drosophila C Virus, Crucifer tobamo virus, Cricket paralysis virus, Bovine viral diarrhea virus 1, Black Queen Cell Virus, Aphid lethal paralysis virus, Avian encephalomyelitis virus, Acute bee paralysis virus, Hibiscus chlorotic ringspot virus, Classical swine fever virus, Human FGF2, Human SFTPA1, Human AML1/RUNX1, Drosophila antennapedia, Human AQP4, Human AT1R, Human BAG-1, Human BCL2, Human BiP, Human c-IAP1 , Human c-myc, Human eIF4G, Mouse NDST4L, Human LEF1, Mouse HIF1 alpha, Human n.myc, Mouse Gtx, Human p27kipl, Human PDGF2/c-sis, Human p53, Human Pim-1, Mouse Rbm3, Drosophila reaper, Canine Scamper, Drosophila Ubx, Human UNR, Mouse UtrA, Human VEGF-A, Human XIAP, Drosophila hairless, S. cerevisiae TFIID, S. cerevisiae YAP1, Human c-src, Human FGF-1, Simian picomavirus, Turnip crinkle virus, an aptamer to eIF4G, Coxsackievirus B3 (CVB3) or Coxsackievirus A (CVB1/2). Wild-type IRES sequences can also be modified and be effective in the invention. In some embodiments, the IRES sequence is about 50 nucleotides in length.

In one embodiment, the IRES is an IRES sequence from viral CVB3, human SAT1, human HK1, viral RhPV, human eIF4GI, viral HCV, viral HallV or circIRES9128.

In one embodiment, the IRES is an IRES sequence from viral CVB3.

The vector as defined herein may comprise a protein coding or noncoding region.

In one embodiment, the vector comprises a protein coding region. The protein coding region can encode a protein for therapeutic use or diagnostic use. In some embodiments, the protein can be any protein for therapeutic use or diagnostic use. For example, the protein coding region can encode human protein or antibodies. In some embodiments, the protein can be selected from, but not limited to, hFIX, SP-B, VEGF-A, human methylmalonyl-CoA mutase (hMUT), CFTR, cancer self-antigens, and additional gene editing enzymes like Cpf1, zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs). The protein can be a protein from bacteria, a virus, a fungus or a parasite. In some embodiments, the protein can be a viral protein. The viral protein can be SARS-CoV-2 spike protein, dengue EV71 protein, influenza haemagglutinin, neuraminidase nucleoprotein or ion channel protein, Zika prM-E protein, HIV gp120 or gp41, RSV Matrix protein or glycoprotein, Ebola glycoprotein, Rabies glycoprotein, *plasmodium gametocyte*, PMIF or PfGARP protein. The protein can be an antigen or agent that can stimulate a body's immune system to recognize the antigen or agent as a foreign invader and/or to generate antibodies that bind to the antigen or agent.

The antibodies as referred to herein may include monoclonal antibodies, polyclonal antibodies, recombinantly produced antibodies, human antibodies, humanized antibodies, chimeric antibodies, synthetic antibodies, tetrameric antibodies comprising two heavy chain and two light chain molecules, antibody light chain monomers, antibody heavy chain monomers, antibody light chain dimers, antibody heavy chain, antibody heavy chain dimers, antibody light chain-heavy chain pairs, intrabodies, heteroconjugate antibodies, monovalent antibodies, antigen-binding fragments of full-length antibodies, and fusion proteins of the above. Such antigen-binding fragments include, but are not limited to, single-domain antibodies (variable domain of heavy chain antibodies (VHHs) or nanobodies), Fabs, F (ab') 2S, and scFvs (single-chain variable fragments).

In one embodiment, the protein coding region is more than 1kb in size. The protein coding region may encode a viral protein (such as a SARS-CoV-2 spike protein or dengue EV71 protein).

In one embodiment, the nucleic acid sequence in b) may comprise i) an IRES region that is positioned 5' of ii) the protein coding or noncoding region. In another embodiment, the

nucleic acid sequence in b) may comprise i) an IRES region that is positioned 3' of ii) the protein coding or noncoding region.

In some embodiments, the protein is a chimeric antigen receptor (CAR) or T cell receptor (TCR) complex protein. The CAR or TCR complex protein may comprise an antigen binding domain specific for a tumor antigen.

In some embodiments, the protein coding region encodes an mRNA sequence. The mRNA may encode any polypeptide of interest, including any naturally or non-naturally occurring or otherwise modified polypeptide. A polypeptide encoded by an mRNA may be of any size and may have any secondary structure or activity. In some embodiments, a polypeptide encoded by an mRNA may have a therapeutic effect when expressed in a cell.

In another embodiment, the vector comprises a noncoding region. In some embodiments, the noncoding regions can encode sequences that alter cellular behaviour, such as e.g., lymphocyte behaviour. In some embodiments, the noncoding sequences are antisense to cellular RNA sequences.

In some embodiments, the noncoding region encodes an siRNA. An siRNA may be capable of selectively knocking down or down regulating expression of a gene of interest. For example, an siRNA could be selected to silence a gene associated with a particular disease, disorder, or condition upon administration to a subject in need thereof of a nanoparticle composition including the siRNA. An siRNA may comprise a sequence that is complementary to an mRNA sequence that encodes a gene or protein of interest. In some embodiments, the siRNA is an immunomodulatory siRNA.

In some embodiments, the noncoding region encodes an shRNA. An shRNA may be produced inside a target cell upon delivery of an appropriate construct to the nucleus. Constructs and mechanisms relating to shRNA are well known in the relevant arts.

In one embodiment, the vector further comprises an RNA polymerase promoter. In one embodiment, the vector further comprises an RNA polymerase positioned upstream of the 5' Group II intron fragment.

The term "promoter" is used herein to refer to a nucleotide region comprising a DNA regulatory sequence, wherein the regulatory sequence is derived from a gene that is capable of binding to an RNA polymerase and allowing for the initiation of transcription of a downstream (3′ direction) coding sequence. It may contain genetic elements at which regulatory proteins and molecules may bind, such as RNA polymerase and other transcription factors, to initiate the specific transcription of a nucleic acid sequence. The phrases "operatively positioned," "operatively linked," "under control," and "under transcriptional control" mean that a promoter is in a correct functional location and/or orientation in relation to a nucleic acid sequence to control transcriptional initiation and/or expression of that sequence.

In another embodiment, the RNA polymerase promoter is a T7 virus RNA polymerase promoter, T6 virus RNA polymerase promoter, SP6 virus RNA polymerase promoter, T3 virus RNA polymerase promoter, or T4 virus RNA polymerase promoter.

In one embodiment, there is provided a precursor RNA comprising:

a) a 5' Group II intron fragment,
b) a nucleic acid sequence comprising i) an internal ribosome entry site (IRES) and ii) a protein coding or noncoding region, and
c) a 3' Group II intron fragment.

Provided herein is a linear RNA precursor encoded by a vector as defined herein.

Also provided herein is a circular RNA produced by a vector as defined herein.

The terms "scar" refers to the length of the region in a circular product excluding the target sequence. A scarless cirRNA contains 0 nucleotide scar sequence. A near-scarless cirRNA contains a scar sequence that is equal to or less than 20 nucleotides in length.

Also provided here is a circular RNA as defined herein for use as a medicament. In one embodiment, there is provided a circular RNA as defined herein for use in treating a disease or a condition in a subject.

As used herein, the term "treating" refers to partially or completely alleviating, ameliorating, improving, relieving, delaying onset of, inhibiting progression of, reducing severity of, and/or reducing incidence of one or more symptoms or features of a particular infection, disease, disorder, and/or condition. For example, "treating" cancer may refer to inhibiting survival, growth, and/or spread of a tumor. Treatment may be administered to a subject who does not exhibit signs of a disease, disorder, and/or condition and/or to a subject who exhibits only early signs of a disease, disorder, and/or condition for the purpose of decreasing the risk of developing pathology associated with the disease, disorder, and/or condition.

The terms "subject", "patient", "host" or "individual", used interchangeably herein, refer to any subject, particularly a vertebrate subject, and even more particularly a mammalian subject, for whom therapy or prophylaxis is desired. Suitable vertebrate animals that fall within the scope of the invention include, but are not restricted to, any member of the subphylum Chordata including primates (e.g., humans, monkeys and apes, and includes species of monkeys such as from the genus Macaca (e.g., cynomolgus monkeys such as Macaca fascicularis, and/or rhesus monkeys (Macaca mulatta)) and baboon (Papio ursinus), as well as marmosets (species from the genus Callithrix), squirrel monkeys (species from the genus Saimiri) and tamarins (species from the genus Saguinus), as well as species of apes such as chimpanzees (Pan troglodytes)), rodents (e.g., mice rats, guinea pigs), lagomorphs (e.g., rabbits, hares), bovines (e.g., cattle), ovines (e.g., sheep), caprines (e.g., goats), porcines (e.g., pigs), equines (e.g., horses), canines (e.g., dogs), felines (e.g., cats), avians (e.g., chickens, turkeys, ducks, geese, companion birds such as canaries, budgerigars etc.), marine mammals (e.g., dolphins, whales), reptiles (snakes, frogs, lizards etc.), and fish. In one embodiment, the subject is a human subject.

Provided herein is a method of purifying circular RNA from a sample, the method comprising: a) providing a vector as defined herein; b) performing *in vitro* transcription to obtain a linear RNA precursor from the vector; c) incubating the linear RNA precursor under conditions to form a circular RNA; and d) purifying the circular RNA.

The term "in vitro transcription" or "IVT" is used herein to refer to the method of producing RNA *in vitro* that uses an RNA polymerase, ribonucleotides, and appropriate buffer conditions to synthesis RNA from a DNA template.

The circular RNA produced by vector or method of the present invention may be purified. For example, the purification means is selected from one or more of a group of: enzymatic treatment; chromatography, including but not limited to affinity column chromatography, reversed-phase silica gel column liquid chromatography, and gel exclusion liquid chromatography; and electrophoresis, including but not limited to gel electrophoresis such as agarose gel electrophoresis, and capillary electrophoresis.

In one embodiment, purifying the circular RNA comprises capturing the circular RNA onto a surface with an immobilized ligand. The immobilized ligand may be streptavidin or sephadex.

Provided herein is a method of expressing protein coding or non-coding region in a cell, said method comprising transfecting the circular RNA as defined herein into the cell.

Provided herein is a nanoparticle composition comprising a circular RNA produced by a vector as defined herein.

As used herein, a "nanoparticle composition" is a composition comprising one or more lipids. Nanoparticle compositions are typically sized on the order of micrometers or smaller and may include a lipid bilayer. Nanoparticle compositions encompass lipid nanoparticles (LNPs), liposomes (e.g., lipid vesicles), and lipoplexes. For example, a nanoparticle composition may be a liposome having a lipid bilayer with a diameter of 500 nm or less.

As used herein, the phrase "lipid nanoparticle" refers to a transfer vehicle comprising one or more lipids (e.g., cationic lipids, non-cationic lipids, and PEG-modified lipids). Preferably, the lipid nanoparticles are formulated to deliver one or more mRNA to one or more target cells. Examples of suitable lipids include, for example, the phosphatidyl compounds (e.g., phosphatidylglycerol, phosphatidylcholine, phosphatidylserine, phosphatidylethanolamine, sphingolipids, cerebrosides, and gangliosides). Also contemplated is the use of polymers as transfer vehicles, whether alone or in combination with other transfer vehicles. Suitable polymers may include, for example, polyacrylates, polyalkycyanoacrylates, polylactide, polylactide -polyglycolide copolymers, polycaprolactones, dextran, albumin, gelatin, alginate, collagen, chitosan, cyclodextrins, dendrimers and polyethylenimine.

As used herein, "transfer vehicle" includes any of the standard pharmaceutical carriers, diluents, excipients, and the like, which are generally intended for use in connection with the administration of biologically active agents, including nucleic acids.

In some embodiments, the invention also relates to compositions, e.g., compositions comprising a circular RNA and a pharmaceutically acceptable carrier. Pharmaceutical compositions of the present disclosure may comprise a circular RNA as described herein, in combination with one or more pharmaceutically or physiologically acceptable carriers, excipients or diluents. In some embodiments, pharmaceutical compositions of the present disclosure may comprise a circular RNA expressing cell as described herein, in combination with one or more pharmaceutically or physiologically acceptable carriers, excipients or diluents.

In some embodiments, a pharmaceutically acceptable carrier can be an ingredient in a pharmaceutical composition, other than an active ingredient, which is nontoxic to the subject.

A pharmaceutically acceptable carrier can include, but is not limited to, a buffer, excipient, stabilizer, or preservative. Examples of pharmaceutically acceptable carriers are solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like that are physiologically compatible, such as salts, buffers, saccharides, antioxidants, aqueous or non-aqueous carriers, preservatives, wetting agents, surfactants or emulsifying agents, or combinations thereof. The amounts of pharmaceutically acceptable carrier(s) in the pharmaceutical compositions may be determined experimentally based on the activities of the carrier(s) and the desired characteristics of the formulation, such as stability and/or minimal oxidation.

In some embodiments, such compositions may comprise buffers such as acetic acid, citric acid, histidine, boric acid, formic acid, succinic acid, phosphoric acid, carbonic acid, malic acid, aspartic acid, Tris buffers, HEPPSO, HEPES, neutral buffered saline, phosphate buffered saline and the like; carbohydrates such as glucose, sucrose, mannose, or dextrans, mannitol; proteins; polypeptides or amino acids such as glycine; antioxidants; chelating agents such as EDTA or glutathione; adjuvants (e.g., aluminum hydroxide); antibacterial and antifungal agents; and preservatives.

In certain embodiments, compositions of the present disclosure can be formulated for a variety of means of parenteral or non-parenteral administration. In one embodiment, the compositions can be formulated for infusion or intravenous administration. Compositions disclosed herein can be provided, for example, as sterile liquid preparations, e.g., isotonic aqueous solutions, emulsions, suspensions, dispersions, or viscous compositions, which may be buffered to a desirable pH. Formulations suitable for oral administration can include liquid solutions, capsules, sachets, tablets, lozenges, and troches, powders liquid suspensions in an appropriate liquid and emulsions.

In one embodiment, there is provided a method of preventing or treating a disease or condition in a subject, the method comprising administering an effective amount of a vector, a circular RNA, nanoparticle composition or pharmaceutical composition as described herein to the subject.

In one embodiment, the disease or condition is a viral infection such as a viral infection or dengue infection. In another embodiment, the disease or condition is a cancer.

The term "administering" refers to contacting, applying, injecting, transfusing or providing a drug as referred to herein to a subject.

By "effective amount", in the context of treating or preventing a condition is meant the administration of an amount of an agent or composition to an individual in need of such treatment or prophylaxis, either in a single dose or as part of a series, that is effective for the prevention of incurring a symptom, holding in check such symptoms, and/or treating existing symptoms, of that condition. The effective amount will vary depending upon the health and physical condition of the individual to be treated, the taxonomic group of individual to be treated, the formulation of the composition, the assessment of the medical situation, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials.

As used herein, "and/or" refers to and encompasses any and all possible combinations of one or more of the associated listed items, as well as the lack of combinations when interpreted in the alternative (or).

As used in this application, the singular form “a,” “an,” and “the” include plural references unless the context clearly dictates otherwise. For example, the term “an agent” includes a plurality of agents, including mixtures thereof.

By “about” is meant a quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length that varies by as much 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 % to a reference quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length.

Throughout this specification and the claims which follow, unless the context requires otherwise, the word “comprise”, and variations such as “comprises” and “comprising”, will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

Throughout this specification and the claims which follow, unless the context requires otherwise, the phrase "consisting essentially of", and variations such as "consists essentially of" will be understood to indicate that the recited element(s) is/are essential i.e. necessary elements of the invention. The phrase allows for the presence of other non-recited elements which do not materially affect the characteristics of the invention but excludes additional unspecified elements which would affect the basic and novel characteristics of the method defined.

The reference in this specification to any prior publication (or information derived from it), or to any matter which is known, is not, and should not be taken as an acknowledgment or admission or any form of suggestion that that prior publication (or information derived from it) or known matter forms part of the common general knowledge in the field of endeavor to which this specification relates.

Those skilled in the art will appreciate that the invention described herein is susceptible to variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications, which fall within the spirit and scope. The invention also includes all of the steps, features, compositions and

compounds referred to or indicated in this specification, individually or collectively, and any and all combinations of any two or more of said steps or features.

Unless otherwise defined, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs.

Certain embodiments of the invention will now be described with reference to the following examples which are intended for the purpose of illustration only and are not intended to limit the scope of the generality hereinbefore described.

EXAMPLES

EXAMPLE 1

Group 1 intron- mediated RNA circularization yields circular RNAs of 3-5 kb in size, with 10-30% exhibiting nicks or breaks.

Earlier studies have showed group 1 intron (Fig. 6A) produces nicked byproducts during in vitro circularization. To investigate the percentage of nicked products generated during in vitro transcription (IVT), we constructed the *Anabaena* PIE IVT templates encompassing various length of gene inserts, ranging from 1.0 kb, 3.0 kb to a maximum size of 5.0 kb. Since the average eukaryotes protein size is ~400 aa, which is ~3.0 kb in nucleotides (21). We examined the IVT products of a 3.0 kb PIE construct on both 1% EX-gel (capable of differentiating nicks) and 1% agarose gel (Fig. 1A). Results suggested that bivalent salt (Mg2+) is critical for both the IVT product yield and the generation of nicks. Low (1.5mM) Mg2+ reduces RNA yield significantly but high (15mM) Mg2+ and long incubation time (3 hours) would lead to accumulation of nicks and reduction of overall circularity (Fig. 1B).

The inventors further tested the effects of different folding conditions (Fig. 7C) on nicks and circularization percentage using a 5.0 kb construct, which is highly nick-prone (Fig. 1C & D). Ex-gel suggested that high Mg2+ (30mM) induces co-transcriptional RNA fragmentation (nicks, Fig. 1C, 20min Ex gel, lane 2 vs. lane 4), and high temperature (55°C) during the folding step (FB) is the culprit of RNA fragmentation because effects of Mg2+

on nick percentage during folding step is marginal (Fig. 1C, lane 6-8). Lower temperature folding (37°C, 40°C) which are ineffective for group 1 based circularization (data not shown) were also tested. Above results indicate that there are limited choices available for reducing nick production using the group 1 based circularization approach. Hence, alternative approaches for RNA circularization that might provide improved performance in terms of both purity and quantity were chosen.

EXAMPLE 2

Engineered Group 2 intron circularized at 90% and contains low nicked byproducts than Group 1 intron circRNA

Recognizing the evident limitations of group 1 introns in circularizing RNA, the inventors subsequently delved into the prospect of employing alternative options. Group 2 introns are intricate RNA molecules commonly found in prokaryotes and lower eukaryotes. Unlike their splicing counterparts, group 1 introns, these molecules self-splice without the need for external cofactor (GTP), showcasing a distinct mechanism of RNA processing. To enhance the efficiency of circularization mediated by group 2 introns, a range of constructs incorporating diverse RNA elements, such as spacers, synthetic homologous sequences, and naturally occurring cyclizing UTRs from Flaviviridae (Fig. 2A, v1-v5) was formulated. The Group 2 PIE Constructs (v1-v5) RNAs were synthesized using the NEB (Cat.: E2040S) HiScribe T7 High Yield RNA Synthesis Kit (Fig. 2B, lanes 1-5) and subsequently circularized under folding condition E (Fold E). Through direct observation on a 1% agarose gel, it is evident that the cyclizing UTRs sourced from Flaviviridae effectively elevate the circularization of a 2.0 kb RNA transcript from 40% to a commendable 90% (Fig. 2B, comparing lane 6 to lane 10).

In order to substantiate the reliability of the 90% circularization efficiency attained with group 2 PIE construct v5, further validation procedures involving both the 2.0 kb and a larger 5.0 kb construct were carried out. This validation encompassed subjecting the samples to RNase R (RR) treatment to verify their circular structure, as illustrated in Fig. 2C-2F. For the 2.0 kb construct, clear separation of the IVT precursor, circular RNA, spliced 3'- intron and spliced 5'- intron were resolved on 1% agarose gel (Fig. 2C, top to lower). To contrast the distinctions in splicing products (precursors, circular forms, introns, nicks) between

group 1 and group 2 intron PIE, the RNA products from AN_2K (group 1 intron) and TET_2K (group 2 intron) were simultaneously analyzed through IVT, Folding, and RR treatments, employing a 1% EX-gel for resolution (depicted in Fig. 2D). Notably, it is evident from the results that group 2 intron PIE v5 (TET_2K) achieves a comparable circularization percentage to that of group 1 intron PIE (AN_2K), while maintaining purity and yield without compromise (indicated by the top arrows in Fig. 2D).

The susceptibility of the 5K constructs to nicks was also examined using similar methodologies. Specifically, a 5 kb group 2 intron PIE (TET_5K) construct was synthesized, folded under FE conditions, and treated with RNaseR (RR) before being visualized on a 1% agarose gel. The results from biological triplicates indicated that TET_5K v5 (incorporating cyclizing UTRs) could achieve efficient circularization of over 80% without requiring any cofactors (depicted in Fig. 2E). Notably, a direct comparison between group 1 and group 2 PIE constructs on a 1% EX-gel demonstrated a significant reduction in nick percentages from the 5 kb constructs, dropping to less than 2% from the 25% observed in the group 1 PIE construct (indicated by the bottom arrows in Fig. 2F). In summary, the engineered group 2 intron PIE demonstrates the capacity to achieve a high circularization percentage while simultaneously mitigating the generation of byproducts (nicks).

Example 3

Generation of scarless and streptavidin aptamer (S1) tagged circular mRNA using permuted group 2 introns at the junction point.

From a bioprocessing standpoint, linear mRNA can be readily purified using affinity selection resin designed to bind to the Poly A tail. On the contrary, the endless circular mRNA lacks both a Cap structure and a Poly A tail, leading to extended stability and decreased immunogenicity as inherent features. However, this characteristic also presents challenges in downstream purification because there is no distinctive affinity tag available to separate circularized RNA from its precursors and associated by-products, such as spliced introns and nicked RNA. To address these purification challenges, the inventors capitalized on the adaptable intron-binding sites (IBS) and exon-binding sites (EBS) situated at the splicing junction point of group 2 introns (Fig. 6B). These sites were substituted with two segments of S1 streptavidin RNA aptamers (Fig. 9A). This arrangement ensures that upon self-circularization, a complete S1 tag is established at the junction site of the circular RNA.

Furthermore, a comparable design approach can be employed to create a seamless junction (scarless) or other varieties of aptamer tags at the same point (Fig. 3A).

Prior to assessing circular RNAs, a benchmark of the purification efficiency of the S1 aptamer on linear RNAs was conducted, utilizing various magnetic streptavidin beads, as depicted in Figure 9B and 9C. Next, the efficacy of purifying S1-tagged circular mRNA using Dynabead MyOne Streptavidin C1 beads (Invitrogen, Cat.:65001) was evaluated. As demonstrated on a 2% EX-gel, the S1-tagged 2kb (S1TET2K) circular mRNA exhibited strong affinity for the C1 beads, as indicated by the top arrows in Figure 3B. Conversely, neither the scarless circular mRNA nor the untagged linear RNA (TETH) displayed binding to the beads. It is worth noting that nonspecific bands were also observed in the S1TET2K lane, likely attributable to the mild wash conditions commonly employed in commercial kits, as indicated by the bottom arrows in Figure 3B. Subsequently, a comprehensive validation of the nickless RNA circularization was undertaken and a one-step purification (NicOPURE, Nickless One-step Purification of Engineered circular mRNA) system (Fig. 7) was developed by using tailored wash buffers (Fig. 3C-F). Outcomes gleaned from a 2 kb construct (S1TET2K) underscore the efficacy of the S1 tag, achieving a purity exceeding 95% in the production of circular RNA. This successful implementation effectively eliminates precursor RNA and spliced introns, as confirmed by the absence of observable nicks on the 1% EX-gel (as shown in Fig. 3C, 3D).

Notably, the NicOPURE technology also demonstrated impressive competence in cyclizing and purifying larger 5 kb RNA molecules. This adeptness enabled the efficient removal of detrimental byproducts while preserving a satisfactory yield. Quality assessment using both 1% agarose and 1% EX gel systems demonstrated clear results, with no nicks or precursors observed following one-step purification, as shown in Figure 3E and 3F (lane 'Bound'). This outcome aligns exceptionally well with the prevailing industry standards for affinity-based mRNA purification strategies, thus strongly indicating a high level of compatibility with established methodologies and reinforcing the efficacy of the approach.

EXAMPLE 4

Structure probing indicates Group 1 and Group 2 intron- generated circular RNA have similar secondary structures *in vitro*.

The inventors then aimed to understand the structural and functional distinctions existing within Group 1 and Group 2 intron circular mRNAs. 100 µg of circular mRNA (circGLUC) for the same target gene were generated using two distinct approaches: one utilizing a Group 1 intron (*Anabaena*) and the other employing the NicOPURE system with Group 2 intron (*Clostridium tetani*) (library prep, fragmentation, library yield; Fig. 10A). The group 1 and group 2 circular mRNA secondary structure were studied using RNA structure probing method: Selective 2'-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) (sequencing quality check; Fig. 11). SHAPE reactivities were computed for the GLUC open-reading frame (ORF) in both group 1 and group 2 circular mRNA using Shapemapper2 (Fig. 4A). The analysis revealed only minor variations in reactivity within the GLUC gene region spanning from nucleotide positions 30 to 60 and 360 to 390 (as indicated by the blue rectangles in Fig. 4A, comparing the top and bottom panels). Importantly, these observed differences were relatively insignificant, suggesting that, overall, the structural distinctions between group 1 and group 2 circular mRNA are not substantial.

Consequently, secondary structures, derived using RNA structure-ProbKnot and incorporating SHAPE reactivity data with annotations exceeding 1.5 for enhanced visualization, are depicted in Figure 4B and 4C. The color-coded bases within these structures indicate pair probabilities. Structural modeling indicates that circular mRNAs generated by both Group 1 and Group 2 introns exhibit similar secondary structures within the GLUC ORF region. The distinct stem-loop formation observed in the 139nt-150nt region of group 2 circGLUC (Figure 4C, highlighted in yellow), representing the sole structural divergence between the two groups. Above finding suggests that the method employed for RNA circularization exerts minimal influence on functional outcomes.

EXAMPLE 5

Significant boost in translation efficiency from NicOPURE generated circular mRNA due to reduced immunogenic responses.

To further assess circular mRNA functionality, two distinct approaches were employed: a fully *in vitro* transcription/translation method (IVTT, Thermo Scientific Cat. 8881) and a

cell-based protein expression assay using HeLa cells, which harbor nucleic acid sensors from the innate immune system. In IVTT assays with spin-column purified Group 1 (*Anabaena*) and Group 2 intron (*Clostridium tetani*, NicOPURE) circular mRNAs encoding GLUC, the quantification of protein production yield based on luminescence revealed comparable outcomes for both samples (as depicted in Fig. 5A, Fig. 5B kinetic assay). These results suggest that, within an *in vitro* system, the protein output remains consistent irrespective of the method used to prepare the circular mRNAs, indicating that equivalent amounts of input circular mRNA yield similar levels of protein production.

Nonetheless, during a cell-based protein expression assay utilizing HeLa cells, a system equipped with innate immune sensors including retinoic acid-inducible gene I (RIG-I) and melanoma differentiation-associated gene 5 (MDA5), the translation efficiency of circular mRNA prepared using group 1 introns (nick-prone) was significantly hindered (as illustrated in Fig. 5C, group 1) upon delivery into the cells using a lipid-based carrier (MessengerMAX, Thermofisher, Cat.: LMRNA001). In contrast, the NicOPURE circular mRNA consistently outperformed the group 1 intron counterpart right from day 1 of transfection ($P<0.001$, ***, Fig. 5D, group 2). By day 5, the NicOPURE circular mRNA from group 2, benefiting from its enhanced purity and resistance to degradation, continued to surpass the expression levels of linear mRNA. This observation underscores its superior stability compared to linear RNA counterparts. The production of interferon alpha (IFN-α), a pivotal component of the anti-RNA viral response, upon delivery of group 1 and group 2 circular mRNA using MessengerMAX was further evaluated. The results revealed that group 1 circular mRNA elicited a robust IFN-α response in contrast to group 2 NicOPURE circular mRNA (as shown in Fig. 5D, $P<0.001$, ***). This factor has been proven to be harmful to protein production, as illustrated in Fig. 5C.

Name of construct	**Sequence**
Clostridium tetani 5' Group II intron fragment	CTATTATCGAGCGAACGCCTTATGCGATGAAAGT CGCACGTAGGGTGTAGACCAAGCGAAATCCTATG CATTTAGGATAGTGAGGTATAGCAAAGGAGAA (SEQ ID NO: 1)
Clostridium tetani 3' Group II intron fragment	GAATATACTATAGCCATACAATAAAAGTGCGAAA CGTTATCCTATAAGTAAGAAAGTTTAAAATTTTCT TACGAAAAGGATAGAACTTAAAAGTTCTAACTGT TCTACTAAAGTAATAAGTGAAAATGTTATTTAAA GCAAACAACCAAGTAGCTTTAAGTCTAAGTCCCC TACACAAGTTTTATACTACTATGCAAAACTTGTGA

	AGCTAGGTAAGGTCGTAATCCGTGAAAGTCGGAT GCGGGGCTCCTTAAAAGATTACTATGGTAAACAT AAGCTAATCCATTAAGATGCGATTTATATGTATTT TATACTGTAAATATTTTTGTGCTTGTGGCTTGGTA TAAAACAGTTAAGATGAAGTACTTAACTGGTTTT GGAATAATTGGTTGTTAAACTAAAACATTATAAA TCGTTAGTGGATACCTAAGGTAATCAAAAATAGG GATAGGTAGAATGGAACGTTTGATGCTGTATATG AAGAGGTTTAGTAGAACCTAGGACACATATACGG GCTCAGCAGGTTCATAGTAGCTATGATACTCAGC CGGAAGTCAAATTAATTTTGAAATACTTCTATGGT AACATAGGAGAAGGATAAAACTGAGTGAGCCAA GGAACCTAGTCGGTAATAG (SEQ ID NO: 2)
Clostridium tetani S1 modified 5' Group II intron fragment	ctattatcgagcgaacgccttatgcgatgaaagtcgcacgtagggtgtagaccaag cgaaatcctatgcatttaggatagtgaggtatGCGTAAGATAGTCGC GGGCCGGGG (SEQ ID NO: 3)
Clostridium tetani S1 modified 3' Group II intron fragment	ACCGACCAGAATCATGCAAGTcaataaaagtgcgaaacgttat cctataagtaagaaagtttaaaattttcttacgaaaaggatagaacttaaaagttctaa ctgttctactaaagtaataagtgaaaatgttatttaaagcaaacaaccaagtagcttta agtctaagtcccctacacaagttttatactactatgcaaaacttgtgaagctaggtaa ggtcgtaatccgtgaaagtcggatgcggggctccttaaaagattactatggtaaac ataagctaatccattaagatgcgatttatatgtattttatactgtaaatatttttgtgctA GTTGCtggtataaaacagttaagatgaagtacttaactggttttggaataattggt tgttaaactaaaacattataaatcgttagtggatacctaaggtaatcaaaaatagggga taggtagaatggaacgtttgatgctgtatatgaagaggtttagtagaacctaggaca catatacgggctcagcaggttcatagtagctatgatactcagccggaagtcaaatta attttgaaatacttctatggtaacataggagaaggataaaactgagtgagccaagga acctagtcggtaatag (SEQ ID NO: 4)
Clostridium tetani D8 modified 5' Group II intron fragment	ctattatcgagcgaacgccttatgcgatgaaagtcgcacgtagggtgtagaccaag cgaaatcctatgcatttaggatagtgaggtat TGATACGGTTGCGGA (SEQ ID NO: 5)
Clostridium tetani D8 modified 3' Group II intron fragment	TCCGAGTAATTTACGTTTcaataaaagtgcgaaacgttatcctata agtaagaaagtttaaaattttcttacgaaaaggatagaacttaaaagttctaactgttc tactaaagtaataagtgaaaatgttatttaaagcaaacaaccaagtagctttaagtct aagtcccctacacaagttttatactactatgcaaaacttgtgaagctaggtaaggtcg taatccgtgaaagtcggatgcggggctccttaaaagattactatggtaaacataagc taatccattaagatgcgatttatatgtattttatactgtaaatatttttgtgctAAACG Tttggtataaaacagttaagatgaagtacttaactggttttggaataattggttgttaa actaaaacattataaatcgttagtggatacctaaggtaatcaaaaatagggataggt agaatggaacgtttgatgctgtatatgaagaggtttagtagaacctaggacacatat acgggctcagcaggttcatagtagctatgatactcagccggaagtcaaattaatttt gaaatacttctatggtaacataggagaaggataaaactgagtgagccaaggaacc tagtcggtaatag (SEQ ID NO: 6)

Spacer 1	GCCGGAAACGCAATAGCCGAAAAACAAAAAACA AAAAAAACAAAAAAAAAACCAAAAAAACAAAAC ACA (SEQ ID NO: 7)
Spacer 2	ACCGGTAAAAAACAAAAAACAAAACGGCTATTAT GCGTTACCGGC (SEQ ID NO: 8)
Denv 5'UTR	AGTTGTTAGTCTACGTGGACCGACAAAGACAGAT TCTTTGAGGGAGCTAAGCTCAACGTAGTTCTAAC AGTTTTTTAATTAGAGAGCAGATCTCTGATGAATA ACCAACGGAAAAAGGCGAAAAACACGCCTTTCAA TATGCTGAA (SEQ ID NO: 9)
Denv 3'UTR	AACAGCATATTGACGCTGGGAAAGACCAGAGATC CTGCTGTCTCCTCAGCATCATTCCAGGCACAGAAC GCCAGAAAATGGAATGGTGCTGTTGAATCAACAG GTTCT (SEQ ID NO: 10)

RNA Aptamers

D8 Sephadex	UCCGAGUAAUUUACGUUUUGAUACGGUUGCGGA (SEQ ID NO: 11)
S1m enhanced streptavidin	AUGCGGCCGCCGACCAGAAUCAUGCAAGUGCGU AAGAUAGUCGCGGGUCGGCGGCCGCAU (SEQ ID NO: 12)
4,4'-methylenedianiline (MDA)	CUGCGAUCAGGGGUAAAUUUCCGCGCAGGCUCC ACGCCGC (SEQ ID NO: 13)
Acetylcholine receptor (AChR)	GCUAGUAGCCUCAGCAGCAUAGUUUCGCCGCUA UGCAGUA (SEQ ID NO: 14)
African trypanosome	AUCGCUACUGCGCCGGUUGCGGCUUGCGGUUGC AACGCCA (SEQ ID NO: 15)
CD4	CUCAGACAGAGCAGAAACGACAGUUCAAGCCGA A (SEQ ID NO: 16)
EGFR	UGCCGCUAUAAUGCACGGAUUUAAUCGCCGUAG AAAAGCAUGUCAAAGCCG (SEQ ID NO: 17)
L-selectin	UAACAACAAUCAAGGCGGGUUCACCGCCCCAGU AUGAGUA (SEQ ID NO: 18)
Neurotensin-1	ACAGAUACGGAACUACAGAGGUCAAUUACGGUG GCCACGC (SEQ ID NO: 19)
NF-κB	CAUACUUGAAACUGUAAGGUUGGCGUAUG (SEQ ID NO: 20)

Minimal S1 motif	ACCGACCAGAAUCAUGCAAGUGCGUAAGAUAGU CGCGGGCCGGG (SEQ ID NO: 30)

Gene sequences

Name of construct	**Sequence**
C. tetani Group II intron Clostridium tetani	GAATATACTATAGCCATACAATAAAAGTGCGAAA CGTTATCCTATAAGTAAGAAAGTTTAAAATTTTCT TACGAAAAGGATAGAACTTAAAAGTTCTAACTGT TCTACTAAAGTAATAAGTGAAAATGTTATTTAAA GCAAACAACCAAGTAGCTTTAAGTCTAAGTCCCC TACACAAGTTTTATACTACTATGCAAAACTTGTGA AGCTAGGTAAGGTCGTAATCCGTGAAAGTCGGAT GCGGGGCTCCTTAAAAGATTACTATGGTAAACAT AAGCTAATCCATTAAGATGCGATTTATATGTATTT TATACTGTAAATATTTTTGTGCTTGTGGCTTGGTA TAAAACAGTTAAGATGAAGTACTTAACTGGTTTT GGAATAATTGGTTGTTAAACTAAAACATTATAAA TCGTTAGTGGATACCTAAGGTAATCAAAAATAGG GATAGGTAGAATGGAACGTTTGATGCTGTATATG AAGAGGTTTAGTAGAACCTAGGACACATATACGG GCTCAGCAGGTTCATAGTAGCTATGATACTCAGC CGGAAGTCAAATTAATTTTGAAATACTTCTATGGT AACATAGGAGAAGGATAAAACTGAGTGAGCCAA GGAACCTAGTCGGTAATAG(311NT)CTATTATCGA GCGAACGCCTTATGCGATGAAAGTCGCACGTAGG GTGTAGACCAAGCGAAATCCTATGCATTTAGGAT AGTGAGGTATAGCAAAGGAGAA (SEQ ID NO: 21)
T7 PIE C. tetani CVB3 GLUC XhoI	TAATACGACTCACTATAGGGCTATTATCGAGCGA ACGCCTTATGCGATGAAAGTCGCACGTAGGGTGT AGACCAAGCGAAATCCTATGCATTTAGGATAGTG AGGTATAGCAAATTAAAACAGCCTGTGGGTTGAT CCCACCCACAGGCCCATTGGGCGCTAGCACTCTG GTATCACGGTACCTTTGTGCGCCTGTTTTATACCC CCTCCCCCAACTGTAACTTAGAAGTAACACACAC CGATCAACAGTCAGCGTGGCACACCAGCCACGTT TTGATCAAGCACTTCTGTTACCCCGGACTGAGTAT CAATAGACTGCTCACGCGGTTGAAGGAGAAAGCG TTCGTTATCCGGCCAACTACTTCGAAAAACCTAGT AACACCGTGGAAGTTGCAGAGTGTTTCGCTCAGC ACTACCCCAGTGTAGATCAGGTCGATGAGTCACC GCATTCCCCACGGGCGACCGTGGCGGTGGCTGCG TTGGCGGCCTGCCCATGGGGAAACCCATGGGACG CTCTAATACAGACATGGTGCGAAGAGTCTATTGA GCTAGTTGGTAGTCCTCCGGCCCCTGAATGCGGCT AATCCTAACTGCGGAGCACACACCCTCAAGCCAG AGGGCAGTGTGTCGTAACGGGCAACTCTGCAGCG

	GAACCGACTACTTTGGGTGTCCGTGTTTCATTTTA TTCCTATACTGGCTGCTTATGGTGACAATTGAGAG ATCGTTACCATATAGCTATTGGATTGGCCATCCGG TGACTAATAGAGCTATTATATATCCCTTTGTTGGG TTTATACCACTTAGCTTGAAAGAGGTTAAAACATT ACAATTCATTGTTAAGTTGAATACAGCAAAGCCA CCATGGGAGTCAAAGTTCTGTTTGCCCTGATCTGC ATCGCTGTGGCCGAGGCCAAGCCCACCGAGAACA ACGAAGACTTCAACATCGTGGCCGTGGCCAGCAA CTTCGCGACCACGGATCTCGATGCTGACCGCGGG AAGTTGCCCGGCAAGAAGCTGCCGCTGGAGGTGC TCAAAGAGATGGAAGCCAATGCCCGGAAAGCTGG CTGCACCAGGGGCTGTCTGATCTGCCTGTCCCACA TCAAGTGCACGCCCAAGATGAAGAAGTTCATCCC AGGACGCTGCCACACCTACGAAGGCGACAAAGA GTCCGCACAGGGCGGCATAGGCGAGGCGATCGTC GACATTCCTGAGATTCCTGGGTTCAAGGACTTGG AGCCCATGGAGCAGTTCATCGCACAGGTCGATCT GTGTGTGGACTGCACAACTGGCTGCCTCAAAGGG CTTGCCAACGTGCAGTGTTCTGACCTGCTCAAGAA GTGGCTGCCGCAACGCTGTGCGACCTTTGCCAGC AAGATCCAGGGCCAGGTGGACAAGATCAAGGGG GCCGGTGGTGACTAAGCCATACAATAAAAGTGCG AAACGTTATCCTATAAGTAAGAAAGTTTAAAATT TTCTTACGAAAAGGATAGAACTTAAAAGTTCTAA CTGTTCTACTAAAGTAATAAGTGAAAATGTTATTT AAAGCAAACAACCAAGTAGCTTTAAGTCTAAGTC CCCTACACAAGTTTTATACTACTATGCAAAACTTG TGAAGCTAGGTAAGGTCGTAATCCGTGAAAGTCG GATGCGGGGCTCCTTAAAAGATTACTATGGTAAA CATAAGCTAATCCATTAAGATGCGATTTATATGTA TTTTATACTGTAAATATTTTTGTGCTTGTGGCTTGG TATAAAACAGTTAAGATGAAGTACTTAACTGGTT TTGGAATAATTGGTTGTTAAACTAAAACATTATAA ATCGTTAGTGGATACCTAAGGTAATCAAAAATAG GGATAGGTAGAATGGAACGTTTGATGCTGTATAT GAAGAGGTTTAGTAGAACCTAGGACACATATACG GGCTCAGCAGGTTCATAGTAGCTATGATACTCAG CCGGAAGTCAAATTAATTTTGAAATACTTCTATGG TAACATAGGAGAAGGATAAAACTGAGTGAGCCA AGGAACCTAGTCGGTAATAGAAACTCGAGAAA (SEQ ID NO: 22)
T7 PIE C. tetani STREP S1 CVB3 GLUC XhoI	TAATACGACTCACTATAGGGCTATTATCGAGCGA ACGCCTTATGCGATGAAAGTCGCACGTAGGGTGT AGACCAAGCGAAATCCTATGCATTTAGGATAGTG AGGTATGCGTAAGATAGTCGCGGGCCGGGGTTAA AACAGCCTGTGGGTTGATCCCACCCACAGGCCCA TTGGGCGCTAGCACTCTGGTATCACGGTACCTTTG TGCGCCTGTTTTATACCCCCTCCCCCAACTGTAAC

	TTAGAAGTAACACACACCGATCAACAGTCAGCGT GGCACACCAGCCACGTTTTGATCAAGCACTTCTGT TACCCCGGACTGAGTATCAATAGACTGCTCACGC GGTTGAAGGAGAAAGCGTTCGTTATCCGGCCAAC TACTTCGAAAAACCTAGTAACACCGTGGAAGTTG CAGAGTGTTTCGCTCAGCACTACCCCAGTGTAGAT CAGGTCGATGAGTCACCGCATTCCCCACGGGCGA CCGTGGCGGTGGCTGCGTTGGCGGCCTGCCCATG GGGAAACCCATGGGACGCTCTAATACAGACATGG TGCGAAGAGTCTATTGAGCTAGTTGGTAGTCCTCC GGCCCCTGAATGCGGCTAATCCTAACTGCGGAGC ACACACCCTCAAGCCAGAGGGCAGTGTGTCGTAA CGGGCAACTCTGCAGCGGAACCGACTACTTTGGG TGTCCGTGTTTCATTTTATTCCTATACTGGCTGCTT ATGGTGACAATTGAGAGATCGTTACCATATAGCT ATTGGATTGGCCATCCGGTGACTAATAGAGCTATT ATATATCCCTTTGTTGGGTTTATACCACTTAGCTT GAAAGAGGTTAAAACATTACAATTCATTGTTAAG TTGAATACAGCAAAGCCACCATGGGAGTCAAAGT TCTGTTTGCCCTGATCTGCATCGCTGTGGCCGAGG CCAAGCCCACCGAGAACAACGAAGACTTCAACAT CGTGGCCGTGGCCAGCAACTTCGCGACCACGGAT CTCGATGCTGACCGCGGGAAGTTGCCCGGCAAGA AGCTGCCGCTGGAGGTGCTCAAAGAGATGGAAGC CAATGCCCGGAAAGCTGGCTGCACCAGGGGCTGT CTGATCTGCCTGTCCCACATCAAGTGCACGCCCAA GATGAAGAAGTTCATCCCAGGACGCTGCCACACC TACGAAGGCGACAAAGAGTCCGCACAGGGCGGC ATAGGCGAGGCGATCGTCGACATTCCTGAGATTC CTGGGTTCAAGGACTTGGAGCCCATGGAGCAGTT CATCGCACAGGTCGATCTGTGTGTGGACTGCACA ACTGGCTGCCTCAAAGGGCTTGCCAACGTGCAGT GTTCTGACCTGCTCAAGAAGTGGCTGCCGCAACG CTGTGCGACCTTTGCCAGCAAGATCCAGGGCCAG GTGGACAAGATCAAGGGGGCCGGTGGTGACTAAA CCGACCAGAATCATGCAAGTCAATAAAAGTGCGA AACGTTATCCTATAAGTAAGAAAGTTTAAAATTTT CTTACGAAAAGGATAGAACTTAAAAGTTCTAACT GTTCTACTAAAGTAATAAGTGAAAATGTTATTTAA AGCAAACAACCAAGTAGCTTTAAGTCTAAGTCCC CTACACAAGTTTTATACTACTATGCAAAACTTGTG AAGCTAGGTAAGGTCGTAATCCGTGAAAGTCGGA TGCGGGGCTCCTTAAAAGATTACTATGGTAAACA TAAGCTAATCCATTAAGATGCGATTTATATGTATT TTATACTGTAAATATTTTTGTGCTAGTTGCTGGTA TAAAACAGTTAAGATGAAGTACTTAACTGGTTTT GGAATAATTGGTTGTTAAACTAAAACATTATAAA TCGTTAGTGGATACCTAAGGTAATCAAAAATAGG GATAGGTAGAATGGAACGTTTGATGCTGTATATG AAGAGGTTTAGTAGAACCTAGGACACATATACGG

	GCTCAGCAGGTTCATAGTAGCTATGATACTCAGC CGGAAGTCAAATTAATTTTGAAATACTTCTATGGT AACATAGGAGAAGGATAAAACTGAGTGAGCCAA GGAACCTAGTCGGTAATAGAAACTCGAGAAA (SEQ ID NO: 23)
T7 PIE C. tetani V5 Spacer CVB3 GLUC XhoI	TAATACGACTCACTATAGGGAGTTGTTAGTCTACG TGGACCGACAAAGACAGATTCTTTGAGGGAGCTA AGCTCAACGTAGTTCTAACAGTTTTTTAATTAGAG AGCAGATCTCTGATGAATAACCAACGGAAAAAGG CGAAAAACACGCCTTTCAATATGCTGAACTATTAT CGAGCGAACGCCTTATGCGATGAAAGTCGCACGT AGGGTGTAGACCAAGCGAAATCCTATGCATTTAG GATAGTGAGGTATAGCAAAGCCGGAAACGCAATA GCCGAAAAACAAAAAACAAAAAAAACAAAAAAA AAACCAAAAAAACAAAACACATTAAAACAGCCT GTGGGTTGATCCCACCCACAGGCCCATTGGGCGC TAGCACTCTGGTATCACGGTACCTTTGTGCGCCTG TTTTATACCCCCTCCCCCAACTGTAACTTAGAAGT AACACACACCGATCAACAGTCAGCGTGGCACACC AGCCACGTTTTGATCAAGCACTTCTGTTACCCCGG ACTGAGTATCAATAGACTGCTCACGCGGTTGAAG GAGAAAGCGTTCGTTATCCGGCCAACTACTTCGA AAAACCTAGTAACACCGTGGAAGTTGCAGAGTGT TTCGCTCAGCACTACCCCAGTGTAGATCAGGTCG ATGAGTCACCGCATTCCCCACGGGCGACCGTGGC GGTGGCTGCGTTGGCGGCCTGCCCATGGGGAAAC CCATGGGACGCTCTAATACAGACATGGTGCGAAG AGTCTATTGAGCTAGTTGGTAGTCCTCCGGCCCCT GAATGCGGCTAATCCTAACTGCGGAGCACACACC CTCAAGCCAGAGGGCAGTGTGTCGTAACGGGCAA CTCTGCAGCGGAACCGACTACTTTGGGTGTCCGTG TTTCATTTTATTCCTATACTGGCTGCTTATGGTGAC AATTGAGAGATCGTTACCATATAGCTATTGGATTG GCCATCCGGTGACTAATAGAGCTATTATATATCCC TTTGTTGGGTTTATACCACTTAGCTTGAAAGAGGT TAAAACATTACAATTCATTGTTAAGTTGAATACAG CAAAGCCACCATGGGAGTCAAAGTTCTGTTTGCC CTGATCTGCATCGCTGTGGCCGAGGCCAAGCCCA CCGAGAACAACGAAGACTTCAACATCGTGGCCGT GGCCAGCAACTTCGCGACCACGGATCTCGATGCT GACCGCGGGAAGTTGCCCGGCAAGAAGCTGCCGC TGGAGGTGCTCAAAGAGATGGAAGCCAATGCCCG GAAAGCTGGCTGCACCAGGGGCTGTCTGATCTGC CTGTCCCACATCAAGTGCACGCCCAAGATGAAGA AGTTCATCCCAGGACGCTGCCACACCTACGAAGG CGACAAAGAGTCCGCACAGGGCGGCATAGGCGA GGCGATCGTCGACATTCCTGAGATTCCTGGGTTCA AGGACTTGGAGCCCATGGAGCAGTTCATCGCACA GGTCGATCTGTGTGTGGACTGCACAACTGGCTGC

	CTCAAAGGGCTTGCCAACGTGCAGTGTTCTGACCT GCTCAAGAAGTGGCTGCCGCAACGCTGTGCGACC TTTGCCAGCAAGATCCAGGGCCAGGTGGACAAGA TCAAGGGGGCCGGTGGTGACTAAACCGGTAAAAA ACAAAAAACAAAACGGCTATTATGCGTTACCGGC GCCATACAATAAAAGTGCGAAACGTTATCCTATA AGTAAGAAAGTTTAAAATTTTCTTACGAAAAGGA TAGAACTTAAAAGTTCTAACTGTTCTACTAAAGTA ATAAGTGAAAATGTTATTTAAAGCAAACAACCAA GTAGCTTTAAGTCTAAGTCCCCTACACAAGTTTTA TACTACTATGCAAAACTTGTGAAGCTAGGTAAGG TCGTAATCCGTGAAAGTCGGATGCGGGGCTCCTT AAAAGATTACTATGGTAAACATAAGCTAATCCAT TAAGATGCGATTTATATGTATTTTATACTGTAAAT ATTTTTGTGCTTGTGGCTTGGTATAAAACAGTTAA GATGAAGTACTTAACTGGTTTTGGAATAATTGGTT GTTAAACTAAAACATTATAAATCGTTAGTGGATA CCTAAGGTAATCAAAAATAGGGATAGGTAGAATG GAACGTTTGATGCTGTATATGAAGAGGTTTAGTA GAACCTAGGACACATATACGGGCTCAGCAGGTTC ATAGTAGCTATGATACTCAGCCGGAAGTCAAATT AATTTTGAAATACTTCTATGGTAACATAGGAGAA GGATAAAACTGAGTGAGCCAAGGAACCTAGTCGG TAATAGAACAGCATATTGACGCTGGGAAAGACCA GAGATCCTGCTGTCTCCTCAGCATCATTCCAGGCA CAGAACGCCAGAAAATGGAATGGTGCTGTTGAAT CAACAGGTTCTAAACTCGAGAAA (SEQ ID NO: 24)
T7 S1 TET_DC_Spacer_CVB3 dSPIKE XhoI (V5)	TAATACGACTCACTATAGGGAGTTGTTAGTCTACG TGGACCGACAAAGACAGATTCTTTGAGGGAGCTA AGCTCAACGTAGTTCTAACAGTTTTTTAATTAGAG AGCAGATCTCTGATGAATAACCAACGGAAAAAGG CGAAAAACACGCCTTTCAATATGCTGAACTATTAT CGAGCGAACGCCTTATGCGATGAAAGTCGCACGT AGGGTGTAGACCAAGCGAAATCCTATGCATTTAG GATAGTGAGGTATGCGTAAGATAGTCGCGGGCCG GGGGCCGGAAACGCAATAGCCGAAAAACAAAAA ACAAAAAAAACAAAAAAAAAACCAAAAAAACAA AACACATTAAAACAGCCTGTGGGTTGATCCCACC CACAGGCCCATTGGGCGCTAGCACTCTGGTATCA CGGTACCTTTGTGCGCCTGTTTTATACCCCCTCCC CCAACTGTAACTTAGAAGTAACACACACCGATCA ACAGTCAGCGTGGCACACCAGCCACGTTTTGATC AAGCACTTCTGTTACCCCGGACTGAGTATCAATA GACTGCTCACGCGGTTGAAGGAGAAAGCGTTCGT TATCCGGCCAACTACTTCGAAAAACCTAGTAACA CCGTGGAAGTTGCAGAGTGTTTCGCTCAGCACTA CCCCAGTGTAGATCAGGTCGATGAGTCACCGCAT TCCCCACGGGCGACCGTGGCGGTGGCTGCGTTGG CGGCCTGCCCATGGGGAAACCCATGGGACGCTCT

	AATACAGACATGGTGCGAAGAGTCTATTGAGCTA GTTGGTAGTCCTCCGGCCCCTGAATGCGGCTAATC CTAACTGCGGAGCACACACCCTCAAGCCAGAGGG CAGTGTGTCGTAACGGGCAACTCTGCAGCGGAAC CGACTACTTTGGGTGTCCGTGTTTCATTTTATTCCT ATACTGGCTGCTTATGGTGACAATTGAGAGATCG TTACCATATAGCTATTGGATTGGCCATCCGGTGAC TAATAGAGCTATTATATATCCCTTTGTTGGGTTTA TACCACTTAGCTTGAAAGAGGTTAAAACATTACA ATTCATTGTTAAGTTGAATACAGCAAAGCCACCA TGTTCGTGTTCCTGGTGCTGCTGCCTCTGGTGTCC AGCCAGTGTGTGAACCTGCGCACCAGAACACAGC TGCCTCCAGCCTACACCAACAGCTTTACCAGAGG CGTGTACTACCCCGACAAGGTGTTCAGATCCAGC GTGCTGCACTCTACCCAGGACCTGTTCCTGCCTTT CTTCAGCAACGTGACCTGGTTCCACGCCATCCACG TGTCCGGCACCAATGGCACCAAGAGATTCGACAA CCCCGTGCTGCCCTTCAACGACGGGGTGTACTTTG CCAGCACCGAGAAGTCCAACATCATCAGAGGCTG GATCTTCGGCACCACACTGGACAGCAAGACCCAG AGCCTGCTGATCGTGAACAACGCCACCAACGTGG TCATCAAAGTGTGCGAGTTCCAGTTCTGCAACGA CCCCTTCCTGGACGTCTACTACCACAAGAACAAC AAGAGCTGGATGGAAAGCGGGGTGTACAGCAGC GCCAACAACTGCACCTTCGAGTACGTGTCCCAGC CTTTCCTGATGGACCTGGAAGGCAAGCAGGGCAA CTTCAAGAACCTGCGCGAGTTCGTGTTTAAGAAC ATCGACGGCTACTTCAAGATCTACAGCAAGCACA CCCCTATCAACCTCGTGCGGGATCTGCCTCAGGGC TTCTCTGCTCTGGAACCCCTGGTGGATCTGCCCAT CGGCATCAACATCACCCGGTTTCAGACACTGCTG GCCCTGCACAGAAGCTACCTGACACCTGGCGATA GCAGCAGCGGATGGACAGCTGGTGCCGCCGCTTA CTATGTGGGCTACCTGCAGCCTAGAACCTTCCTGC TGAAGTACAACGAGAACGGCACCATCACCGACGC CGTGGATTGTGCTCTGGATCCTCTGAGCGAGACA AAGTGCACCCTGAAGTCCTTCACCGTGGAAAAGG GCATCTACCAGACCAGCAACTTCCGGGTGCAGCC CACCGAATCCATCGTGCGGTTCCCCAATATCACCA ATCTGTGCCCCTTCGGCGAGGTGTTCAATGCCACC AGATTCGCCTCTGTGTACGCCTGGAACCGGAAGC GGATCAGCAATTGCGTGGCCGACTACTCCGTGCT GTACAACTCCGCCAGCTTCAGCACCTTCAAGTGCT ACGGCGTGTCCCCTACCAAGCTGAACGACCTGTG CTTCACAAACGTGTACGCCGACAGCTTCGTGATCC GGGGAGATGAAGTGCGGCAGATTGCCCCTGGACA GACAGGCAAGATCGCCGACTACAACTACAAGCTG CCCGACGACTTCACCGGCTGTGTGATTGCCTGGA ACAGCAACAACCTGGACTCCAAAGTCGGCGGCAA CTACAATTACCGGTACCGGCTGTTCCGGAAGTCC

	AATCTGAAGCCCTTCGAGCGGGACATCTCCACCG AGATCTATCAGGCCGGCAGCAAGCCTTGTAACGG CGTGGAAGGCTTCAACTGCTACTTCCCACTGCAGT CCTACGGCTTTCAGCCCACAAATGGCGTGGGCTA TCAGCCCTACAGAGTGGTGGTGCTGAGCTTCGAA CTGCTGCATGCCCCTGCCACAGTGTGCGGCCCTAA GAAAAGCACCAATCTCGTGAAGAACAAATGCGTG AACTTCAACTTCAACGGCCTGACCGGCACCGGCG TGCTGACAGAGAGCAACAAGAAGTTCCTGCCATT CCAGCAGTTTGGCCGGGATATCGCCGATACCACA GACGCCGTTAGAGATCCCCAGACACTGGAAATCC TGGACATCACCCCTTGCAGCTTCGGCGGAGTGTCT GTGATCACCCCTGGCACCAACACCAGCAATCAGG TGGCAGTGCTGTACCAGGGCGTGAACTGTACCGA AGTGCCCGTGGCCATTCACGCCGATCAGCTGACA CCTACATGGCGGGTGTACTCCACCGGCAGCAATG TGTTTCAGACCAGAGCCGGCTGTCTGATCGGAGC CGAGCACGTGAACAATAGCTACGAGTGCGACATC CCCATCGGCGCTGGAATCTGCGCCAGCTACCAGA CACAGACAAACAGCCGTCGGAGAGCCAGAAGCG TGGCCAGCCAGAGCATCATTGCCTACACAATGTC TCTGGGCGCCGAGAACAGCGTGGCCTACTCCAAC AACTCTATCGCTATCCCCACCAACTTCACCATCAG CGTGACCACAGAGATCCTGCCTGTGTCCATGACC AAGACCAGCGTGGACTGCACCATGTACATCTGCG GCGATTCCACCGAGTGCTCCAACCTGCTGCTGCA GTACGGCAGCTTCTGCACCCAGCTGAATAGAGCC CTGACAGGGATCGCCGTGGAACAGGACAAGAAC ACCCAAGAGGTGTTCGCCCAAGTGAAGCAGATCT ACAAGACCCCTCCTATCAAGGACTTCGGCGGCTT CAATTTCAGCCAGATTCTGCCCGATCCTAGCAAGC CCAGCAAGCGGAGCTTCATCGAGGACCTGCTGTT CAACAAAGTGACACTGGCCGACGCCGGCTTCATC AAGCAGTATGGCGATTGTCTGGGCGACATTGCCG CCAGGGATCTGATTTGCGCCCAGAAGTTTAACGG ACTGACAGTGCTGCCTCCTCTGCTGACCGATGAG ATGATCGCCCAGTACACATCTGCCCTGCTGGCCG GCACAATCACAAGCGGCTGGACATTTGGAGCAGG CGCCGCTCTGCAGATCCCCTTTGCTATGCAGATGG CCTACCGGTTCAACGGCATCGGAGTGACCCAGAA TGTGCTGTACGAGAACCAGAAGCTGATCGCCAAC CAGTTCAACAGCGCCATCGGCAAGATCCAGGACA GCCTGAGCAGCACAGCAAGCGCCCTGGGAAAGCT GCAGAACGTGGTCAACCAGAATGCCCAGGCACTG AACACCCTGGTCAAGCAGCTGTCCTCCAACTTCG GCGCCATCAGCTCTGTGCTGAACGATATCCTGAG CAGACTGGACCCTCCTGAGGCCGAGGTGCAGATC GACAGACTGATCACAGGCAGACTGCAGAGCCTCC AGACATACGTGACCCAGCAGCTGATCAGAGCCGC CGAGATTAGAGCCTCTGCCAATCTGGCCGCCACC

	AAGATGTCTGAGTGTGTGCTGGGCCAGAGCAAGA GAGTGGACTTTTGCGGCAAGGGCTACCACCTGAT GAGCTTCCCTCAGTCTGCCCCTCACGGCGTGGTGT TTCTGCACGTGACATATGTGCCCGCTCAAGAGAA GAATTTCACCACCGCTCCAGCCATCTGCCACGAC GGCAAAGCCCACTTTCCTAGAGAAGGCGTGTTCG TGTCCAACGGCACCCATTGGTTCGTGACACAGCG GAACTTCTACGAGCCCCAGATCATCACCACCGAC AACACCTTCGTGTCTGGCAACTGCGACGTCGTGAT CGGCATTGTGAACAATACCGTGTACGACCCTCTG CAGCCCGAGCTGGACAGCTTCAAAGAGGAACTGG ACAAGTACTTTAAGAACCACACAAGCCCCGACGT GGACCTGGGCGATATCAGCGGAATCAATGCCAGC GTCGTGAACATCCAGAAAGAGATCGACCGGCTGA ACGAGGTGGCCAAGAATCTGAACGAGAGCCTGAT CGACCTGCAAGAACTGGGGAAGTACGAGCAGTAC ATCAAGTGGCCCTGGTACATCTGGCTGGGCTTTAT CGCCGGACTGATTGCCATCGTGATGGTCACAATC ATGCTGTGTTGCATGACCAGCTGCTGTAGCTGCCT GAAGGGCTGTTGTAGCTGTGGCAGCTGCTGCAAG TTCGACGAGGACGATTCTGAGCCCGTGCTGAAGG GCGTGAAACTGCACTACACATGAACCGGTAAAAA ACAAAAAACAAAACGGCTATTATGCGTTACCGGC ACCGACCAGAATCATGCAAGTCAATAAAAGTGCG AAACGTTATCCTATAAGTAAGAAAGTTTAAAATT TTCTTACGAAAAGGATAGAACTTAAAAGTTCTAA CTGTTCTACTAAAGTAATAAGTGAAAATGTTATTT AAAGCAAACAACCAAGTAGCTTTAAGTCTAAGTC CCCTACACAAGTTTTATACTACTATGCAAAACTTG TGAAGCTAGGTAAGGTCGTAATCCGTGAAAGTCG GATGCGGGGCTCCTTAAAAGATTACTATGGTAAA CATAAGCTAATCCATTAAGATGCGATTTATATGTA TTTTATACTGTAAATATTTTTGTGCTACTTGCTGGT ATAAAACAGTTAAGATGAAGTACTTAACTGGTTT TGGAATAATTGGTTGTTAAACTAAAACATTATAA ATCGTTAGTGGATACCTAAGGTAATCAAAAATAG GGATAGGTAGAATGGAACGTTTGATGCTGTATAT GAAGAGGTTTAGTAGAACCTAGGACACATATACG GGCTCAGCAGGTTCATAGTAGCTATGATACTCAG CCGGAAGTCAAATTAATTTTGAAATACTTCTATGG TAACATAGGAGAAGGATAAAACTGAGTGAGCCA AGGAACCTAGTCGGTAATAGAACAGCATATTGAC GCTGGGAAAGACCAGAGATCCTGCTGTCTCCTCA GCATCATTCCAGGCACAGAACGCCAGAAAATGGA ATGGTGCTGTTGAATCAACAGGTTCTAAACTCGA GAAA (SEQ ID NO: 25)
DENV1CyclBroc (301nt)	AGTTGTTAGTCTACGTGGACCGACAAGAACAGTT TCGAATCGGAAGCTTGCTTAACGTAGTTCTAACA GTTTTTTATTAGAGAGCAGATCTCTGATGAACAAC

	CAACGAAAAAAGACGGCTCGACCGTCTTTCAATA TGCTGAAGAGACGGTCGGGTCCAGATATTCGTAT CTGTCGAGTAGAGTGTGGGCTCAACAGCATATTG ACGCTGGGAGAGACCAGAGATCCTGCTGTCTCTA CAGCATCATTCCAGGCACAGAACGCCAGAAAATG GAATGGTGCTGTTGAATCAACAGGTTCT (SEQ ID NO: 26)
DENV2CyclBroc (303nt)	AGTTGTTAGTCTACGTGGACCGACAAAGACAGAT TCTTTGAGGAAGCTAAGCTTAACGTAGTTCTAACA GTTTTTTAATTAGAGAGCAGATCTCTGATGAATAA CCAACGGAAAAAGGCGAGAAATACGCCTTTCAAT ATGCTGAAGAGACGGTCGGGTCCAGATATTCGTA TCTGTCGAGTAGAGTGTGGGCTCAACAGCATATT GACGCTGGGAAAGACCAGAGATCCTGCTGTCTCC TCAGCATCATTCCAGGCACAGAACGCCAGAAAAT GGAATGGTGCTGTTGAATCAACAGGTTCT (SEQ ID NO: 27)
DENV3CyclBroc(300nt)	AGTTGTTAGTCTACGTGGACCGACAAGAACAGTT TCGATTCGGAAGCTTGCTTAACGTAGTGCTGACA GTTTTTTATTAGAGAGCGGATCTCTGATGAACAAC CAACGGAAGAAGACGGGAAAACCGTCTATCAATA TGCTGAAGAGACGGTCGGGTCCAGATATTCGTAT CTGTCGAGTAGAGTGTGGGCTCAACAGCATATTG ACGCTGGGAGAGACCAGAGATCCTGCTGTCTCCT CAGCATCATTCCAGGCACAGAACGCCAGAAAATG GAATGGTGCTGTTGAATCAACAGGTTC (SEQ ID NO: 28)
DENV4CyclBroc (306nt)	AGTTGTTAGTCTGTGTGGACCGACAAGGACAGTT CCAAATCGGAAGCTTGCTTAACACAGTTCTAACA GTTTGTTTTAAATAGAGAGCAGATCTCTGGAAAA ATGAACCAACGAAAAAAGGTGGTCAGACCACCTT TCAATATGCTGAAGAGACGGTCGGGTCCAGATAT TCGTATCTGTCGAGTAGAGTGTGGGCTCAACAGC ATATTGACGCTGGGAAAGACCAGAGATCCTGCTG TCTCTGCAACATCAATCCAGGCACAGAGCGCCGC GAGATGGATTGGTGTTGTTGATCCAACAGGTTCT (SEQ ID NO: 29)

CLAIMS

1. A vector for generating a circular RNA, the vector comprising the following elements operably connected to each other and arranged in the following sequence:
   a) a modified 5' Group II intron fragment comprising a 3' splice site and a first purification fragment sequence positioned at or close to the splicing junction point,
   b) a nucleic acid sequence comprising i) an internal ribosome entry site (IRES) region and ii) a protein coding or noncoding region, and
   c) a modified 3' Group II intron fragment comprising a 5' splice site, a second purification fragment sequence positioned at or close to the splicing junction point, wherein the second purification sequence comprises an intron-binding site 1 (IBS1) sequence that is complementary to an exon-binding site 1 (EBS1) sequence in the modified 3' Group II intron;
   wherein the first and second purification fragment sequences are capable of being retained in the circular RNA upon Group II intron-mediated back-splicing of the modified 5' and 3' Group II intron fragments to form an RNA aptamer.

2. The vector according to claim 1, wherein the vector further comprises at least one spacer domain.

3. The vector according to claim 1 or 2, wherein the vector comprises i) a first spacer domain positioned between the modified 5' Group II intron fragment and the IRES sequence and ii) a second spacer domain positioned between the protein coding or noncoding region and the modified 3' Group II intron fragment.

4. The vector according to any one of claims 1 to 3, wherein the modified 5' and 3' Group II intron fragments are from *Clostridium tetani.*

5. The vector according to any one of claims 1 to 4, wherein the junction sequence of the modified 5' Group II intron fragment has been replaced with the first purification fragment sequence and the junction sequence of the modified 3' Group II intron fragment has been replaced with the second purification fragment sequence.

6. The vector according to any one of claims 1 to 5, wherein the 3' Group II intron fragment has been modified to have an EBS1 sequence that is complementary to the IBS1 sequence in the second purification fragment sequence.

7. The vector according to any one of claims 1 to 6, wherein the modified 5' Group II intron fragment comprises or consists of a nucleic acid sequence having at least 70% sequence identity to SEQ ID NO: 3 or 5 and wherein the modified 3' Group II intron fragment comprises or consists of a nucleic acid sequence having at least 70% sequence identity to SEQ ID NO: 4 or 6.

8. The vector according to any one of claims 1 to 7, wherein the vector further comprises at least one homology sequence positioned 5' of the modified 5' Group II intron fragment and at least one homology sequence positioned 3' of the modified 3' Group II intron fragment.

9. The vector of claim 8, wherein the at least one homology sequence is a *Flaviviridae* UTR sequence.

10. The vector of any one of claims 1 to 9, wherein the RNA aptamer is an S1 streptavidin aptamer or a D8 sephadex aptamer.

11. The vector of any one of claims 1 to 10, wherein the IRES is an IRES sequence from viral CVB3, human SAT1, human HK1, viral RhPV, human eIF4GI, viral HCV, viral HallV or circIRES9128.

12. The vector of any one of claims 1 to 11, wherein the vector further comprises an RNA polymerase positioned upstream of the 5' Group II intron fragment.

13. The vector of claim 12, wherein the RNA polymerase is an RNA polymerase promotor from T7 virus, T6 virus, SP6 virus, T3 virus, or T4 virus.

14. A vector for generating a circular RNA, the vector comprising the following elements operably connected to each other and arranged in the following sequence: a) a modified 5' Group II intron fragment comprising a 3' splice site and a first purification fragment sequence positioned at or close to the splicing junction point, b) a nucleic acid sequence comprising i) an internal ribosome entry site (IRES)

region and ii) a protein coding or noncoding region, and

c) a modified 3' Group II intron fragment comprising a 5' splice site and a second purification fragment sequence positioned at or close to the splicing junction point, wherein the first and second purification fragment sequences are capable of being retained in the circular RNA upon Group 2 intron-mediated back-splicing of the modified 5' and 3' Group II intron fragments to form an RNA aptamer.

15. A linear RNA precursor encoded by a vector according to any one of claims 1 to 14.

16. A circular RNA produced by a vector according to any one of claims 1 to 14.

17. A method of purifying circular RNA from a sample, the method comprising:
    a) providing a vector according to any one of claims 1 to 14;
    b) performing *in vitro* transcription to obtain a linear RNA precursor from the vector;
    c) incubating the linear RNA precursor under conditions to form a circular RNA; and
    d) purifying the circular RNA.

18. The method of claim 17, wherein purifying the circular RNA comprises capturing the circular RNA onto a surface with an immobilized ligand.

19. The method of claim 18, wherein the immobilized ligand is streptavidin.

20. A method of expressing protein coding or non-coding region in a cell, said method comprising transfecting the circular RNA of claim 16 into the cell.

21. A nanoparticle composition comprising a circular RNA produced by a vector of any one of claims 1 to 14.