

COMPOSITIONS AND METHODS FOR PRODUCING CIRCULAR POLYRIBONUCLEOTIDES

Background

There is a need for methods of producing, purifying, and using circular polyribonucleotides.

5

Summary

The disclosure provides compositions and methods for producing, purifying, and using circular RNA.

10 In one example, the disclosure features a linear polyribonucleotide having the formula 5'-(A)-(B)-(C)-(D)-(E)-(F)-(G)-3'. The linear polyribonucleotide includes, from 5' to 3', (A) a 3' half of Group I catalytic intron fragment; (B) a 3' splice site; (C) a 3' exon fragment; (D) a polyribonucleotide cargo; (E) a 5' exon fragment; (F) a 5' splice site; and (G) a 5' half of Group I catalytic intron fragment. The polyribonucleotide includes a first annealing region that has from 2 to 50, e.g., 5 to 50, e.g., 6 to 50, e.g., 7 to 50, e.g., 8 to 50 (e.g., from 10 to 30, 10 to 20, or 10 to 15, e.g., at least 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50) ribonucleotides and is present within (A) the 3' half of Group I catalytic intron fragment; (B) the 3' splice site; or (C) the 3' exon fragment. The polyribonucleotide also includes a second annealing region that has from 2 to 50, e.g., 5 to 50, e.g., 6 to 50, e.g., 7 to 50, e.g., 8 to 50 (e.g., from 10 to 30, 10 to 20, or 10 to 15, e.g., at least 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50) ribonucleotides and is present within (E) the 5' exon fragment; (F) the 5' splice site; or (G) the 5' half of Group I catalytic intron fragment. The first annealing region has from 80% to 100% (e.g., 85% to 100%, e.g., 90% to 100%, e.g., 80%, 85%, 90%, 95%, 97%, 99%, or 100%) complementarity with the second annealing region or has from zero to 10 (e.g., 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10), mismatched base pairs.

25

In one aspect, the present disclosure refers to a linear polyribonucleotide having the formula 5'-(A)-(B)-(C)-(D)-(E)-(F)-(G)-3', wherein:

(A) comprises a 3' half of Group I catalytic intron fragment;

(B) comprises a 3' splice site;

(C) comprises a 3' exon fragment;

(D) comprises a polyribonucleotide cargo;

(E) comprises a 5' exon fragment;

(F) comprises a 5' splice site; and

(G) comprises a 5' half of Group I catalytic intron fragment;

and wherein (A), (B), or (C) comprises a first annealing region comprising from 8 to 50 ribonucleotides and (E), (F), or (G) comprises a second annealing region comprising from 8 to 50 ribonucleotides,

wherein the first annealing region and the second annealing region have from 80% to 100% complementarity, or

the first annealing region and the second annealing region comprises from zero to 10 mismatched base pairs.

In one aspect, the present disclosure refers to a linear polyribonucleotide having the formula 5'-(A)-(B)-(C)-(D)-(E)-(F)-(G)-3', wherein:

(A) comprises a 3' portion of Group I catalytic intron fragment;

(B) comprises a 3' splice site;

(C) comprises a 3' exon fragment comprising a first annealing region comprising from 8 to 50 ribonucleotides, wherein the first annealing region is flanked on each side by a portion of the 3' exon fragment;

(D) comprises a polyribonucleotide cargo;

(E) comprises a 5' exon fragment comprising a second annealing region comprising from 8 to 50 ribonucleotides that can hybridize to the first annealing region;

(F) comprises a 5' splice site; and

(G) comprises a 5' portion of Group I catalytic intron fragment.

In another example, the disclosure features a linear polyribonucleotide having the formula 5'-(A)-(B)-(C)-(D)-(E)-(F)-(G)-3'. The linear polyribonucleotide includes, from 5' to 3', (A) a 3' half of Group I catalytic intron fragment; (B) a 3' splice site; (C) a 3' exon fragment; (D) a polyribonucleotide cargo; (E) a 5' exon fragment; (F) a 5' splice site; and (G) a 5' half of Group I catalytic intron fragment, wherein the 3' half of Group I catalytic intron fragment of (A) and the 5' half of Group I catalytic intron fragment of (G) are from a Cyanobacterium Anabaena pre-tRNA-Leu gene. The polyribonucleotide includes a first annealing region that has from 5 to 50, e.g., 6 to 50 (e.g., from 10 to 30, 10 to 20, or 10 to 15, e.g., at least 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50) ribonucleotides and is present within (A) the 3' half of Group I catalytic intron fragment; (B) the 3' splice site; or (C) the 3' exon fragment. The polyribonucleotide also includes a second annealing region that has from 5 to 50, e.g., 6 to 50 (e.g., from 10 to 30, 10 to 20, or 10 to 15, e.g., at least 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50) ribonucleotides and is present within (E) the 5' exon fragment; (F) the 5' splice site; or (G) the 5' half of Group I catalytic intron fragment. The first annealing region has from 80% to 100% (e.g., 85% to 100%, e.g., 90% to 100%, e.g., 80%, 85%, 90%, 95%, 97%, 99%, or 100%) complementarity with the

second annealing region or has from zero to 10 e.g., (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10) mismatched base pairs.

In another example, the disclosure features a linear polyribonucleotide having the formula 5'-(A)-(B)-(C)-(D)-(E)-(F)-(G)-3'. The linear polyribonucleotide includes, from 5' to 3', (A) a 3' half of Group I catalytic intron fragment; (B) a 3' splice site; (C) a 3' exon fragment; (D) a polyribonucleotide cargo; (E) a 5' exon fragment; (F) a 5' splice site; and (G) a 5' half of Group I catalytic intron fragment, wherein the 3' half of Group I catalytic intron fragment of (A) and the 5' half of Group I catalytic intron fragment of (G) are from a Tetrahymena pre-rRNA. The polyribonucleotide includes a first annealing region that has from 6 to 50, e.g., 7 to 50 (e.g., from 10 to 30, 10 to 20, or 10 to 15, e.g., at least 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50) ribonucleotides and is present within (A) the 3' half of Group I catalytic intron fragment; (B) the 3' splice site; or (C) the 3' exon fragment. The polyribonucleotide also includes a second annealing region that has from 6 to 50, e.g., 7 to 50 (e.g., from 10 to 30, 10 to 20, or 10 to 15, e.g., at least 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50) ribonucleotides and is present within (E) the 5' exon fragment; (F) the 5' splice site; or (G) the 5' half of Group I catalytic intron fragment. The first annealing region has from 80% to 100% (e.g., 85% to 100%, e.g., 90% to 100%, e.g., 80%, 85%, 90%, 95%, 97%, 99%, or 100%) complementarity with the second annealing region or has from zero to 10 e.g., (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10) mismatched base pairs.

In some examples, (A) or (C) includes the first annealing region and (E) or (G) includes the second annealing region.

In some examples, the 3' exon fragment of (C) includes the first annealing region and the 5' exon fragment of (E) includes the second annealing region.

In some examples, the 3' exon fragment of (C) includes the first annealing region and the 5' half of Group I catalytic intron fragment of (G) includes the second annealing region.

In some examples, the 3' half of Group I catalytic intron fragment of (A) includes the first annealing region and the 5' exon fragment of (E) includes the second annealing region.

In some examples, first annealing region and the second annealing region include zero or one mismatched base pair.

In some examples, the first annealing region and the second annealing region are 100% complementary.

In some examples, the 3' half of Group I catalytic intron fragment of (A) and the 5' half of Group I catalytic intron fragment of (G) are from a cyanobacterium Anabaena pre-tRNA-Leu gene, a Tetrahymena pre-rRNA, or a T4 phage td gene.

In some examples, the 3' half of Group I catalytic intron fragment of (A) and the 5' half of Group I catalytic intron fragment of (G) are from a Cyanobacterium Anabaena pre-tRNA-Leu gene, and the 3' exon fragment of (C) includes the first annealing region and the 5' exon fragment of (E) includes the second annealing region. The first annealing region may include, e.g., from 5 to 50, e.g., from 10 to 15 (e.g., 10, 11, 12, 13, 14, or 15) ribonucleotides and the second annealing region may include, e.g., from 5 to 50, e.g., from 10 to 15 (e.g., 10, 11, 12, 13, 14, or 15) ribonucleotides.

In some examples, the 3' half of Group I catalytic intron fragment of (A) and the 5' half of Group I catalytic intron fragment of (G) are from a Tetrahymena pre-rRNA, and the 3' half of Group I catalytic intron fragment of (A) includes the first annealing region and the 5' exon fragment of (E) includes the second annealing region. In some examples, the 3' exon fragment of (C) includes the first annealing region and the 5' half of Group I catalytic intron fragment of (G) includes the second annealing region. The first annealing region may include, e.g., from 6 to 50, e.g., from 10 to 16 (e.g., 10, 11, 12, 13, 14, 15, or 16) ribonucleotides, and the second annealing region may include, e.g., from 6 to 50, e.g., from 10 to 16 (e.g., 10, 11, 12, 13, 14, 15, or 16) ribonucleotides.

In some examples, the 3' half of Group I catalytic intron fragment of (A) and the 5' Group I catalytic intron fragment of (G) are from a T4 phage td gene. The 3' exon fragment of (C) may include the first annealing region and the 5' half of Group I catalytic intron fragment of (G) may include the second annealing region. The first annealing region may include, e.g., from 2 to 16, e.g., 10 to 16 (e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or 16) ribonucleotides, and the second annealing region may include, e.g., from 2 to 16, e.g., 10 to 16 (e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or 16) ribonucleotides.

In some examples, the 3' half of Group I catalytic intron fragment of (A) is the 5' terminus of the linear polynucleotide.

In some examples, the 5' half of Group I catalytic intron fragment of (G) is the 3' terminus of the linear polyribonucleotide.

In some examples, the linear polyribonucleotide does not include a further annealing region.

In some examples, the linear polyribonucleotide does not include an annealing region 3' to (A) that includes partial or complete nucleic acid complementarity with an annealing region 5' to (G).

In some examples, the polyribonucleotide cargo of (D) includes an expression sequence, a non-coding sequence, or an expression sequence and a non-coding sequence.

In some examples, the polyribonucleotide cargo of (D) includes an expression sequence encoding a polypeptide.

In some examples, the polyribonucleotide cargo of (D) includes an IRES operably linked to an expression sequence encoding a polypeptide.

In some examples, the IRES is located upstream of the expression sequence. In some examples, the IRES is located downstream of the expression sequence.

In some examples, the polyribonucleotide cargo of (D) includes an expression sequence that encodes a polypeptide that has a biological effect on a subject.

In some examples, the linear polyribonucleotide further includes a first spacer region between the 3' exon fragment of (C) and the polyribonucleotide cargo of (D). The first spacer region may be, e.g., at least 5 (e.g., at least 10, at least 15, at least 20) ribonucleotides in length. In some examples, the linear polyribonucleotide further includes a second spacer region between the polyribonucleotide cargo of (D) and the 5' exon fragment of (E). The second spacer region may be, e.g., at least 5 (e.g., at least 10, at least 15, at least 20) ribonucleotides in length. In some examples, each spacer region is at least 5 (e.g., at least 10, at least 15, at least 20) ribonucleotides in length. Each spacer region may be, e.g., from 5 to 500 (e.g., 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, or 500) ribonucleotides in length. The first spacer region, the second spacer region, or the first spacer region and the second spacer region may include a polyA sequence. The first spacer region, the second spacer

region, or the first spacer region and the second spacer region may include a polyA-C sequence. The first spacer region, the second spacer region, or the first spacer region and the second spacer region may include a polyA-G sequence. The first spacer region, the second spacer region, or the first spacer region and the second spacer region may include a polyA-T sequence. The first spacer region, the second
5 spacer region, or the first spacer region and the second spacer region may include a random sequence.

In some examples, the linear polyribonucleotide is from 50 to 20,000, e.g., 100 to 20,000, e.g., 200 to 20,000, e.g., 300 to 20,000 (e.g., 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1,000, 1,100, 1,200, 1,300, 1,400, 1,500, 1,600, 1,700, 1,800, 1,900, 2,000, 2,500, 3,000, 3,500, 4,000, 5,000, 6,000, 7,000, 8,000, 9,000, 10,000, 11,000, 12,000, 13,000, 14,000, 15,000, 16,000, 17,000, 18,000, 19,000, or
10 20,000) ribonucleotides in length. In examples, the linear polyribonucleotide is, e.g., at least 50, at least 100, at least 200, at least 300, at least 400, at least 500, at least 1,000, at least 2,000, at least 3,000, at least 4,000, or at least 5,000 ribonucleotides in length.

In another example, the disclosure features a DNA vector including an RNA polymerase promoter operably linked to a DNA sequence that encodes the linear polyribonucleotide of any of the examples
15 described herein. In one aspect, the present disclosure refers to a DNA vector comprising an RNA polymerase promoter operably linked to a DNA sequence that encodes any one of the linear polyribonucleotide described herein.

In another example, the disclosure features a circular polyribonucleotide (e.g., a covalently closed circular polyribonucleotide) produced from the linear polyribonucleotide or the DNA vector of any of the
20 examples described herein. In one aspect, the present disclosure refers to a circular polyribonucleotide produced from any one of the linear polyribonucleotide described herein or from the DNA vector described herein.

In another example, the disclosure features a circular polyribonucleotide (e.g., a covalently closed circular polyribonucleotide) having a splice junction joining a 5' exon fragment and a 3' exon fragment.
25 The 3' exon fragment includes a first annealing region including 2 to 50, e.g., 5 to 50, e.g., 6 to 50, e.g., 7 to 50, e.g., 8 to 50 (e.g., from 10 to 30, 10 to 20, or 10 to 15, e.g., 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50) ribonucleotides, and the 5' exon fragment includes a second annealing region including 2 to 50, e.g., 5 to 50, e.g., 6 to 50, e.g., 7 to 50, e.g., 8 to 50 (e.g., from 10 to 30, 10 to 20, or 10 to 15,
30 e.g., 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50) ribonucleotides. In example, the first annealing region and the second annealing region include from 80% to 100% (e.g., 80%, 85%, 90%, 95%, 97%, 99%, or 100%) complementarity. In examples the first annealing region and the second annealing region include from zero to 10 (e.g., 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10) mismatched base pairs
35 (bp). In examples, the circular polynucleotide further include a polyribonucleotide cargo. In examples, the polyribonucleotide cargo includes an expression (or coding) sequence, a non-coding sequence, or a combination of an expression (or coding) sequence and a non-coding sequence. In examples, the polyribonucleotide cargo includes an expression (coding) sequence encoding a polypeptide. In examples, the polyribonucleotide includes an IRES operably linked to an expression sequence encoding
40 a polypeptide. In some examples, the circular polyribonucleotide further includes a spacer region between the IRES and the 3' exon fragment or the 5' exon fragment. The spacer region may be, e.g., at least 5 (e.g., at least 10, at least 15, at least 20) ribonucleotides in length ribonucleotides in length. The spacer region may be, e.g., from 5 to 500 (e.g., 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, or 500) ribonucleotides. In some examples, the spacer region includes a polyA
45 sequence. In some examples, the spacer region includes a polyA-C sequence. In

some examples, the spacer region includes a polyA-G sequence. In some examples, the spacer region includes a polyA-T sequence. In some examples, the spacer region includes a random sequence.

In some examples, the circular polyribonucleotide is from 50 to 20,000, e.g., 100 to 20,000, e.g., 200 to 20,000, e.g., 300 to 20,000 (e.g., 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1,000, 1,100, 1,200, 1,300, 1,400, 1,500, 1,600, 1,700, 1,800, 1,900, 2,000, 2,500, 3,000, 3,500, 4,000, 5,000, 6,000, 7,000, 8,000, 9,000, 10,000, 11,000, 12,000, 13,000, 14,000, 15,000, 16,000, 17,000, 18,000, 19,000, or 20,000) ribonucleotides in length. In examples, the circular polyribonucleotide is, e.g., at least 500, at least 1,000, at least 2,000, at least 3,000, at least 4,000, or at least 5,000 ribonucleotides in length.

In one aspect, the present disclosure refers to a circular polyribonucleotide comprising a splice junction joining a 5' exon fragment and a 3' exon fragment,

wherein the 3' exon fragment comprises a first annealing region comprising 8 to 50 ribonucleotides, and the 5' exon fragment comprises a second annealing region comprising 8 to 50 ribonucleotides,

and wherein the first annealing region and the second annealing region comprises from 80% to 100% complementarity, or the first annealing region and the second annealing region comprises from zero to 10 mismatched base pairs.

In some examples, the circular polyribonucleotide is produced from a linear polyribonucleotide or vector as described herein.

In another example, the disclosure features a method of expressing a polypeptide in a cell by providing a linear polyribonucleotide, a DNA vector, or a circular polyribonucleotide as described herein to the cell. In one aspect, the present disclosure refers to a method of expressing a polypeptide in a cell, the method comprising providing any one of the linear polyribonucleotide described herein, the DNA vector described herein, or any one of the circular polyribonucleotide described herein to the cell. The method further includes allowing the cellular machinery to express the polypeptide from the polyribonucleotide.

In another example, the disclosure features a method of producing a circular polyribonucleotide as described herein by providing a linear polyribonucleotide as described herein under conditions suitable for self-splicing of the linear polyribonucleotide to produce the circular polyribonucleotide. In one aspect, the present disclosure refers to a method of producing a circular polyribonucleotide from any one of the linear polyribonucleotide described herein, the method comprising providing the linear polyribonucleotide under conditions suitable for self-splicing of the linear polyribonucleotide to produce the circular polyribonucleotide.

Definitions

To facilitate the understanding of this disclosure, a number of terms are defined below. Terms defined herein have meanings as commonly understood by a person of ordinary skill in the areas relevant to the disclosure. Terms such as "a", "an," and "the" are not intended to refer to only a singular entity but include the general class of which a specific example may be used for illustration. The term "or" is used to mean "and/or" unless explicitly indicated to refer to alternatives only or the alternative are mutually exclusive, although the disclosure supports a definition that refers to only alternatives and "and/or". The terminology herein is used to describe specific embodiments, but their usage is not to be taken as limiting, except as outlined in the claims.

As used herein, any values provided in a range of values include both the upper and lower bounds, and any values contained within the upper and lower bounds.

As used herein, the term "about" refers to a value that is within $\pm 10\%$ of a recited value.

As used herein, the term "carrier" is a compound, composition, reagent, or molecule that facilitates the transport or delivery of a composition (e.g., a circular polyribonucleotide) into a cell by a covalent modification of the circular polyribonucleotide, via a partially or completely encapsulating agent, or a combination thereof. Non-limiting examples of carriers include carbohydrate carriers (e.g., an anhydride-modified phytoglycogen or glycogen-type material), nanoparticles (e.g., a nanoparticle that encapsulates or is covalently linked binds to the circular polyribonucleotide), liposomes, fusosomes, ex vivo differentiated reticulocytes, exosomes, protein carriers (e.g., a protein covalently linked to the circular polyribonucleotide), or cationic carriers (e.g., a cationic lipopolymer or transfection reagent).