REDUCING INTRON RETENTION

CROSS-REFERENCE

[0001] This application claims the benefit of UK Patent Application NO: 1410693.4, filed on 16 June 2014, which is incorporated herein by reference in its entirety.

BACKGROUND

[0002] Alternative splicing can be a frequent phenomenon in the human transcriptome. Intron retention is one example of alternative splicing in which a partially processed mRNA retains retention of at least one intron after undergoing partial splicing. In some instances, the presence of a retained intron in a partially processed mRNA can prevent or reduce translation of a functional protein.

SUMMARY

[0003] This invention relates to a method of reducing or preventing intron retention in a transcript (e.g., a partially processed mRNA transcript) and treatment or prevention of diseases related to inadvertent intron retention.

[0004] In some aspects, the invention discloses a method of prevention or treatment of a disease in a subject comprising reducing the incidence of intron retention in gene transcripts, wherein the disease is induced by defective protein expression caused by the intron retention in the gene transcripts.

[0005] In some aspects, the invention discloses a method of modulating intron splicing in a cell, comprising hybridizing a polynucleic acid polymer to a region of pre-mRNA, wherein the region comprises or consists of SEQ ID NO: 46, or a region having at least 95% identity to SEQ ID NO: 46.

[0006] In some aspects, the invention discloses a method of modulating intron splicing in a cell, comprising hybridizing a polynucleic acid polymer to a region of pre-mRNA, wherein the region comprises or consists of SEQ ID NO: 46, or a region having at least 95% identity to SEQ ID NO: 46.

[0007] In some aspects, the invention discloses a method of modulating intron splicing in a cell, comprising hybridizing a polynucleic acid polymer to a region of pre-mRNA, wherein the region comprises or consists of SEQ ID NO: 3, or a region having at least 95% identity to SEQ ID NO: 3.

[0008] In some aspects, the invention discloses a method of modulating intron splicing in a cell, comprising hybridizing a polynucleic acid polymer to a region of pre-mRNA, wherein the region

comprises or consists of a sequence complementary to a sequence having at least 95% identity to any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof.

[0009] In some aspects, the invention discloses a polynucleic acid polymer which is antisense to at least part of a region of polynucleic acid polymer comprising or consisting of SEQ ID NO: 46, or a region of polynucleic acid polymer comprising or consisting of a sequence having at least 95% sequence identity to SEQ ID NO: 46.

[0010] In some aspects, the invention discloses a polynucleic acid polymer which is antisense to at least part of a region of polynucleic acid polymer comprising or consisting of SEQ ID NO: 3, or a region of polynucleic acid polymer comprising or consisting of a sequence having at least 95% sequence identity to SEQ ID NO: 3.

[0011] In some aspects, the invention discloses a polynucleic acid polymer which is antisense to at least part of a region of polynucleic acid polymer, wherein the region comprises or consists of a sequence complementary to any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof; or optionally a region of polynucleic acid polymer comprising or consisting of a sequence having at least 95% sequence identity to SEQ ID NOs: 47 to 434.

[0012] In some aspects, the invention discloses a polynucleic acid polymer comprising or consisting of a nucleic acid sequence having at least 95% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 28; SEQ ID NO: 29; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 34; SEQ ID NO: 35; SEQ ID NO: 37; SEQ ID NO: 38; SEQ ID NO: 40; SEQ ID NO: 41; SEQ ID NO: 43; and SEQ ID NO: 44; or combinations thereof; and optionally wherein the uracil nucleotides are substituted with thymine nucleotides

[0013] In some aspects, the invention discloses a polynucleic acid polymer comprising or consisting of a nucleic acid sequence having at least 95% identity to a sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof; and optionally wherein the uracil nucleotides are substituted with thymine nucleotides.

[0014] In some aspects, the invention discloses a pharmaceutical composition which comprises a polynucleic acid polymer that hybridizes to a target sequence of a partially processed mRNA transcript which encodes a protein and which comprises a retained intron, wherein the target sequence is in between two G quadruplexes, wherein the polynucleic acid polymer is capable of

inducing splicing out of the retained intron from the partially processed mRNA transcript; and a pharmaceutically acceptable excipient and/or a delivery vehicle. In some instances, the polynucleic acid polymer hybridizes to the retained intron of the partially processed mRNA transcript.

[0015] In some aspects, the invention discloses a pharmaceutical composition which comprises a polynucleic acid polymer that hybridizes to a target sequence of a partially processed mRNA transcript which encodes a protein and which comprises a retained intron, wherein the polynucleic acid polymer hybridizes to an intronic splicing regulatory element of the partially processed mRNA transcript, wherein the intronic splicing regulatory element comprises a first CCC motif, and wherein the polynucleic acid polymer is capable of inducing splicing out of the retained intron from the partially processed mRNA transcript; and a pharmaceutically acceptable excipient and/or a delivery vehicle.

[0016] In some instances, the intronic splicing regulatory element further comprises a second CCC motif. In some instances, the first CCC motif is about 3 or more nucleotide bases from the second CCC motif. In some instances, the polynucleic acid polymer hybridizes to an intronic splicing regulatory element comprising a CCCAG or an AGGCC motif.

[0017] In some aspects, the invention discloses a pharmaceutical composition which comprises a polynucleic acid polymer that hybridizes to a target sequence of a partially processed mRNA transcript which encodes a protein and which comprises a retained intron, wherein the polynucleic acid polymer hybridizes to a binding motif of the partially processed mRNA transcript, wherein the binding motif does not form a G quadruplex, and wherein the polynucleic acid polymer is capable of inducing splicing out of the retained intron from the partially processed mRNA transcript; and a pharmaceutically acceptable excipient and/or a delivery vehicle.

[0018] In some instances, the polynucleic acid polymer comprises a pyridine nucleotide at the 3' terminal position and/or at the 5' terminal position. In some instances, the polynucleic acid polymer comprises two consecutive pyridine nucleotides at the 3' terminal position and/or at the 5' terminal position.

[0019] In some aspects, the invention discloses a pharmaceutical composition which comprises a polynucleic acid polymer that hybridizes to a target sequence of a partially processed mRNA transcript which encodes a protein and which comprises a retained intron, wherein the polynucleic acid polymer hybridizes to a binding motif of the partially processed mRNA transcript, and wherein the binding motif forms a hairpin structure, wherein the polynucleic acid polymer is capable of

inducing splicing out of the retained intron from the partially processed mRNA transcript; and a pharmaceutically acceptable excipient and/or a delivery vehicle.

[0020] In some instances, the polynucleic acid polymer is further capable of destabilizing the hairpin structure. In some instances, the delivery vehicle comprises a cell penetrating peptide or a peptide-based nanoparticle. In some instances, the delivery vehicle is complexed with the polynucleic acid polymer by ionic bonding.

[0021] In some instances, the polynucleic acid polymer is between about 10 and about 50, about 10 and about 45, about 10 and about 40, about 10 and about 30, about 10 and about 25, or about 10 and about 20 nucleotides in length. In some instances, the sequence of the polynucleic acid polymer is at least 60%, 70%, 80%, 90%, or 95% or 100% complementary to a target sequence of the partially processed mRNA transcript. In some instances, the sequence of the polynucleic acid polymer has 4 or less, 3 or less, 2 or less, or 1 or less mismatches to a target sequence of the partially processed mRNA transcript.

[0022] In some instances, the polynucleic acid polymer is modified at the nucleoside moiety or at the phosphate moiety. In some instances, the polynucleic acid polymer comprises one or more artificial nucleotide bases. In some instances, the one or more artificial nucleotide bases comprises 2'-O-methyl, 2'-O-methoxyethyl (2'-O-MOE), 2'-O-aminopropyl, 2'-deoxy, T-deoxy-2'-fluoro, 2'-O-aminopropyl (2'-O-AP), 2'-O-dimethylaminoethyl (2'-O-DMAOE), 2'-O-dimethylaminopropyl (2'-O-DMAP), T-O- dimethylaminoethyloxyethyl (2'-O-DMAEOE), or 2'-O-N-methylacetamido (2'-O-NMA) modified, locked nucleic acid (LNA), ethylene nucleic acid (ENA), peptide nucleic acid (PNA), 1', 5'- anhydrohexitol nucleic acids (HNA), morpholino, methylphosphonate nucleotides, thiolphosphonate nucleotides, or 2'-fluoro N3-P5'-phosphoramidites. In some instances, the polynucleic acid polymer is modified at the 2' hydroxyl group of the ribose moiety of the nucleoside moiety of the polynucleic acid polymer. In some instances, the modification at the 2' hydroxyl group is by a 2'-O-methyl, 2'-O-methoxyethyl (2'-O-MOE), 2'-O-aminopropyl, 2'-deoxy, T-deoxy-2'-fluoro, 2'-O-aminopropyl (2'-O-AP), 2'-O-dimethylaminoethyl (2'-O-DMAOE), 2'-O-dimethylaminopropyl (2'-O-DMAP), T-O- dimethylaminoethyloxyethyl (2'-O-DMAEOE), or 2'-O-N-methylacetamido (2'-O-NMA) moiety. In some instances, a methyl group is added to the 2' hydroxyl group of the ribose moiety to generate a 2'-O-methyl ribose moiety. In some instances, a methoxyethyl group is added to the 2' hydroxyl group of the ribose moiety to generate a 2'-O-methoxyethyl ribose moiety. In some instances, the modification at the 2' hydroxyl group is linked to the 4' carbon by a methylene group. In some instances, the ribose ring is substituted with a six

member morpholino ring to generate a morpholino artificial nucleotide analogue. In some instances, the phosphate backbone is substituted with an oligoglycine-like moiety to generate peptide nucleic acid (PNA). In some instances, the phosphate backbone is modified by a thiol group or a methyl group. In some instances, the 5' terminus, 3' terminus, or a combination thereof is modified. In some instances, the modification protects the polynucleic acid polymer from endogenous nucleases in the subject. In some instances, the modified polynucleic acid polymer does not induce or has a reduced ability to induce RNase H cleavage of RNA. In some instances, the polynucleic acid polymer is modified to increase its stability. In some instances, the hybridization is specific hybridization.

[0023] In some instances, the polynucleic acid polymer hybridizes to an mRNA transcript comprising at least 80%, 85%, 90%, or 95% or 100% sequence identity to at least 13 contiguous bases of SEQ ID NO: 46. In some instances, the polynucleic acid polymer hybridizes to a mRNA transcript comprising at least 10 contiguous bases of SEQ ID NO: 46. In some instances, the polynucleic acid polymer hybridizes to a mRNA transcript comprising at least 63%, 70%, 80%, 90%, or 95% sequence identity to a sequence selected from the group consisting of SEQ ID NO: 6, SEQ ID NO: 9, SEQ ID NO: 12, SEQ ID NO: 15, SEQ ID NO: 18, SEQ ID NO: 21, SEQ ID NO: 24, SEQ ID NO: 27, SEQ ID NO: 30, SEQ ID NO: 33, SEQ ID NO: 36, SEQ ID NO: 39, SEQ ID NO: 42, SEQ ID NO: 45, or combinations thereof. In some instances, the polynucleic acid polymer hybridizes to a mRNA transcript comprising at least 55%, 60%, 70%, 80%, 90%, or 95% sequence identity to SEQ ID NO: 3. In some instances, the polynucleic acid polymer comprises at least 63%, 70%, 80%, 90%, or 95% or 100% sequence identity to a sequence selected from the group consisting of SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 43, and SEQ ID NO: 44, or combinations thereof, and optionally wherein uracil nucleotides are substituted with thymine nucleotides. In some instances, the polynucleic acid polymer comprises at least 55%, 60%, 70%, 80%, 90%, or 95%, or comprises 100% sequence identity to a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, or combinations thereof, and optionally wherein uracil nucleotides are substituted with thymine nucleotides. In some instances, the polynucleic acid polymer hybridizes to an mRNA transcript comprising at least 80%, 85%, 90%, or 95% sequence

identity to at least 13 contiguous bases of a sequence selected from the group consisting of SEQ ID NOs: 47-434, or combinations thereof.

[0024] In some instances, the polynucleic acid polymer is a synthesized polynucieic acid polymer.

[0025] In some instances, the pharmaceutical composition comprising the polynucleic acid polymer is administered for intravenous or subcutaneous administration.

[0026] In some aspects, the invention discloses a composition for use in the treatment of a disease or condition in a patient in need thereof, comprising administering to the patient a pharmaceutical composition disclosed herein. In some instances, the disease or condition is associated with an impaired production of a protein or is characterized by a defective splicing. In some instances, the disease or condition is a hereditary disease. In some instances, a subject with the hereditary disease has a genome that comprises a copy of a gene that comprises an exon that when properly transcribed into fully processed mRNA encodes the full-length functional form of the protein. In some instances, a subject with the hereditary disease has a genome that comprises a copy of a gene that comprises a set of exons that when properly transcribed into fully processed mRNA encodes the full-length functional form of the protein. In some instances, a subject with the hereditary disease has a genome that comprises a defective copy of the gene, which is incapable of producing a full-length functional form of the protein. In some instances, the disease or condition is diabetes. In some instances, the disease or condition is cancer. In some instances, the composition for use further comprises selecting a subject for treatment of a disease or condition associated with an impaired production of a protein which comprises (a) determining if the subject has the disease or condition associated with an impaired production of the protein; and (b) administering to the subject a pharmaceutical composition of claims 42-82 if the subject has the disease or condition associated with an impaired production of the protein. In some instances, the composition for use further comprises selecting a subject for treatment of a disease or condition characterized by a defective splicing, which comprises (a) determining if the subject has the disease or condition characterized by the defective splicing; and (b) administering to the subject a pharmaceutical composition of claims 42-82 if the subject has the disease or condition characterized by the defective splicing.

[0027] In some aspects, the invention discloses a method of treating a disease or condition characterized by impaired production of a full-length functional form of a protein in a subject in need thereof, comprising: (a) administering to the subject a pharmaceutical composition comprising: a therapeutic agent that induces an increase in splicing out of an intron in a partially processed mRNA transcript; and a pharmaceutically acceptable excipient and/or a delivery vehicle; wherein the

subject has a pool of partially processed mRNA transcripts, which are capable of encoding copies of the full-length functional form of the protein and each of which comprise at least one retained intron that inhibits translation of the partially processed mRNA transcripts; and (b) contacting a target cell of the subject with the therapeutic agent to induce a portion of the pool of the partially processed mRNA transcripts to undergo splicing to remove the at least one retained intron from each of the partially processed mRNA transcripts in the portion, to produce fully processed mRNA transcripts, wherein the fully processed mRNA transcripts are translated to express copies of the full-length functional form of the protein, which treat the disease or condition.

[0028] In some instances, the therapeutic agent causes activation of one or more splicing protein complexes in the cell to remove the at least one retained intron from each of the partially processed mRNA transcripts in the portion of the pool of the partially processed mRNA transcripts. In some instances, the therapeutic agent inhibits a protein that regulates intron splicing activity. In some instances, the therapeutic agent activates a protein that regulates intron splicing activity. In some instances, the therapeutic agent binds to a protein that regulates intron splicing activity. In some instances, the therapeutic agent binds to target polynucleotide sequence of the partially processed mRNA transcripts. In some instances, the therapeutic agent is a polynucleic acid polymer. In some instances, the therapeutic agent is a small molecule.

[0029] In some instances, the pharmaceutical composition is the pharmaceutical composition described herein.

[0030] In some instances, the impaired production of a full-length functional form of the protein comprises sub-normal production of the full-length functional form of the protein. In some instances, the impaired production of a full-length functional form of the protein is due to an absence of expression of the full-length functional form of the protein or a level of expression of the full-length functional form of the protein that is sufficiently low so as to cause the disease or condition. In some instances, the impaired production of a full-length functional form of the protein comprises absence of production of the protein or production of a defective form of the protein. In some instances, the defective form of the protein is a truncated form of the protein, a mis-folded form of the protein or a form of the protein with aberrant target binding. In some instances, treating the subject results in increased expression of the full-length functional form of the protein.

[0031] In some aspects, the invention discloses a method of inducing processing of a partially processed mRNA transcript to remove a retained intron to produce a fully processed mRNA transcript that encodes a full-length functional form of a protein, comprising: (a) hybridizing an

isolated polynucleic acid polymer to the partially processed mRNA transcript, which is capable of encoding the full-length functional form of the protein and which comprises at least one retained intron; (b) removing the at least one retained intron from the partially processed mRNA transcript to produce a fully processed mRNA transcript that encodes a full-length functional form of the protein; and (c) translating the full-length functional form of the protein from the fully processed mRNA transcript. In some instances, the method further comprises administering a pharmaceutical composition comprising the isolated polynucleic acid polymer to a subject in need thereof.

[0032] In some instances, the impaired production of the full-length functional form of the protein is correlated to a disease or condition. In some instances, the disease or condition is a hereditary disease. In some instances, the subject with the hereditary disease has a genome that comprises a copy of a gene that comprises an exon that when properly transcribed into fully processed mRNA encodes the full-length functional form of the protein. In some instances, the subject with the hereditary disease has a genome that comprises a copy of a gene that comprises a set of exons that when properly transcribed into fully processed mRNA encodes the full-length functional form of the protein. In some instances, the subject with the hereditary disease has a genome that comprises a defective copy of the gene, which is incapable of producing a full-length functional form of the protein. In some instances, the disease or condition is diabetes. In some instances, the disease or condition is cancer.

[0033] In some instances, the polynucleic acid polymer is an anti-sense sequence. In some instances, the anti-sense sequence hybridizes to a retained intron of the partially processed mRNA transcript. In some instances, the anti-sense sequence hybridizes to an intronic splicing regulatory element of the partially processed mRNA transcript. In some instances, the intronic splicing regulatory element comprises a CCC motif. In some instances, the anti-sense sequence hybridizes to a binding motif of the partially processed mRNA transcript wherein the binding motif does not form a G quadruplex. In some instances, the anti-sense sequence hybridizes to a binding motif of the partially processed mRNA transcript wherein the binding motif is between two G quadruplexes. In some instances, the anti-sense sequence hybridizes to a binding motif of the partially processed mRNA transcript wherein the binding motif has a first CCC motif and a second CCC motif. In some instances, the first CCC motif is about 3 or more nucleotide bases from the second CCC motif. In some instances, the sequence of the polynucleic acid polymer is at least 60%, 70%, 80%, 90%, or 95% or 100% complementary to a target sequence of the partially processed mRNA transcript. In some instances, the sequence of the polynucleic acid polymer has 4 or less, 3 or less, 2 or less, or 1 or less

mismatches to a target sequence of the partially processed mRNA transcript. In some instances, the polynucleic acid polymer is between about 10 and about 50, about 10 and about 45, about 10 and about 40, about 10 and about 30, about 10 and about 25, or about 10 and about 20 nucleotides in length.

[0034] In some instances, the subject is a eukaryote. In some instances, the subject is a eukaryote selected from a human, mouse, rat, non-human primate, or non-primate mammal.

[0035] In some aspects, described is a pharmaceutical composition which comprises a polynucleic acid polymer that hybridizes to a target sequence of a partially processed mRNA transcript which encodes a protein and which comprises a retained intron, wherein the polynucleic acid polymer induces splicing out of the retained intron from the partially processed mRNA transcript; and a pharmaceutically acceptable excipient and/or a delivery vehicle.

BRIEF DESCRIPTION OF THE DRAWINGS

[0036] The novel features of the invention are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present invention will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the invention are utilized, and the accompanying drawings of which:

[0037] Figure 1. Location of SSOs in the human proinsulin gene. (**A**) Schematics of the *INS* reporter and its mRNA products. SSOs are shown as black horizontal bars below exons (numbered boxes) and below intron 1 (line); their sequences are in Table 2. Start and stop codons are denoted by arrowheads. Canonical (solid lines) and cryptic (dotted lines) splicing is shown above the primary transcript; designation of cryptic splice sites is in grey. SSOs targeting intron 1 segments del4-del7 are shown in the lower panel. (**B**) mRNA isoforms (numbered 1–6) generated by the *INS* reporter construct. Description of isoforms that do not produce proinsulin is labelled with *.

[0038] Figure 2. SSO-induced inhibition of *INS* intron 1 retention. (**A**) Cotransfection of the *INS* reporter construct (IC D-F) with the indicated SSOs into HEK293 cells. Spliced products described in Figure 1B are shown to the right. Bars represent percentage of intron 1-containing isoforms relative to natural transcripts (upper panel) or percentage of splicing to the cryptic 3' splice site of intron 2 relative to the total (lower panel). Error bars denote SD; sc, scrambled control; SSO-, 'no SSO' control. Final concentration of SSOs was 1, 3, 10 and 30 nM, except for SSO6 and SSO8 (10 and 30 nM). (**B**) SSO21-mediated promotion of intron 1 splicing in clones lacking the cryptic 3ss of intron 2. RNA products are to the right. (**C**) A fold change in SSO21-induced intron 1 retention in

transcripts containing and lacking the cryptic 3'ss of intron 2. The final concentration of SSO21 was 30 nM in duplicate transfection. Designation of the reporter constructs is at the bottom.

[0039] Figure 3. *INS* SSOs targeting cryptic 3' splice sites. (**A**) Activation of cryptic 3'ss of intron 2 (cr3'ss+126; Figure 1A) by SSO6 and promotion of exon 2 skipping by SSO8. Concentration of each SSO was 2, 10, 50 and 250 nM. SSOs are shown at the top, spliced products to the right, reporter at the bottom. (**B**) A predicted stable hairpin between the authentic and cryptic 3'ss of *INS* intron 2. Bases targeted by SSO6 are denoted by asterisks and predicted splicing enhancer hexamers (listed to the right) are denoted by a dotted line. (**C**) SSO4 does not prevent activation of cryptic 3'ss 81 base pairs downstream of its authentic counterpart (cr3'ss+81) in cells depleted of U2AF35 but induces exon skipping. The final concentration of each SSO in COS7 cells was 5, 20 and 80 nM. The final concentration of the siRNA duplex U2AF35ab (77) was 70 nM. The reporter was the same as in panel A.

[0040] Figure 4. Optimization of the intron retention target by antisense microwalk at a single-nucleotide resolution. (**A**) Location of oligoribonucleotides. Microwalk SSOs and oligos used for CD/NMR are represented by horizontal black bars below and above the primary transcript, respectively. Intron 1 sequences predicted to form RNA G-quadruplexes are highlighted in grey. Microwalk direction is shown by grey arrows; winner oligos are highlighted in black. A box denotes a single nucleotide polymorphism reported previously (20). (**B**) Intron retention levels of each microwalk SSO in two cell lines. Error bars denote SDs obtained from two independent cotransfections with reporter IC D-F.

[0041] Figure 5. Biophysical characterization of RNA secondary structure formation. (**A**) Far-UV CD spectrum at 25°C for CD1 (19-mer) and CD2 (20-mer) RNAs, revealing ellipticity maximum at 265 and 270 nm, respectively. (**B**) 1H NMR spectra of CD1 and CD2 recorded at 800 MHz and 298 K showing characteristic groups of resonances from H-bonded G bases. (**C**) Sigmoidal CD melting curves for the two RNAs showing a transition mid-point at 56.8 ± 0.2°C and 69.0 ± 0.45°C, respectively. The two curves have been displaced slightly from each other for clarity. (**D**) The proposed parallel quadruplex structure with two stacked G-tetrads connected by short loop sequences for CD1 (top panel). Predicted hairpin structures for CD2 are shown at the bottom panel. G→C mutations are in red.

[0042] Figure 6. Conformational quadruplex/hairpin transitions involving the antisense target. (**A**) Schematic equilibrium between hairpin (black) and quadruplex (dark blue) structures proposed to form within the G-rich motif encompassing oligoribonucleotide CD3. CD4 contains a CC→UU

mutation (highlighted by *). (**B**) The NMR spectrum in the 9–15 ppm region reveals imino proton signals corresponding to hydrogen bonded bases. The signals between 10 and 12 ppm are characteristic of Hoogsteen hydrogen bonded Gs within a G-tetrad (Q box), while signals > 12 ppm are indicative of Watson–Crick A-U and G-C base pairs within hairpin structures (H box). In CD3, hairpin H1 is significantly populated, but mutations in CD4 destabilize H1 making H2 the major species, with both in equilibrium with the quadruplex structure. (**C**) Mfold predictions of two possible hairpins, consistent with the NMR data. (D) Reduction of intron retention upon destabilization of the hairpin structure by the CC→UU mutation. Error bars denote SD of a duplicate experiment with reporter IC D-C. Del5, the IC D-C reporter lacking segment del5 (Figure 1A); M1, a reporter containing two substitutions (Table 1A) to destabilize both the G-quadruplex and the stem-loop.

[0043] Figure 7. Identification of proteins that interact with pre-mRNAs encompassing the antisense target for intron retention. (**A**) Intron retention levels for wild type and mutated reporter constructs (IC D-C) following transient transfections into HEK293T cells. Mutations are shown in Table 1A. RNA products are to the right. The presence of predicted RNA quadruplexes, hairpins H1/H2 and the upstream and downstream C4 run are indicated below the gel figure. Error bars denote SDs obtained from two replicate experiments. (**B**) Intron retention levels of tested RNAs correlate with their predicted stabilities across the antisense target. (**C**)Western blot analysis of a pull-down assay with antibodies indicated to the right. NE, nuclear extracts; B, beads-only control; AV3, control RNA oligo containing a cytosine run and a 3'ss AG (7). The sequence of CD5 RNA is shown in Figure 4A.

[0044] Figure 8. Splicing pattern of quadruplex-rich and -poor minigenes upon DHX36 depletion. (**A**) Schematics of reporter constructs. Predicted quadruplexes are denoted by black rectangles; their densities are shown in Table 1. Exons (boxes) are numbered; forward slash denotes shortening of *F9* intron 3 (24). The *F9* and *TSC2* minigenes contain branch point substitutions c.253–25C and c.5069–18C, respectively, that impair splicing (24). Cr5'ss-104; cryptic 5'ss 104 upstream of authentic 5'ss of intron 2. (**B**) Immunoblot with antibodies against DHX36. sc, scrambled siRNA; c, untreated cells. Error bars are SDs of two transfection experiments. (**C–E**) Intron retention and exon skipping of the indicated reporters. The final concentration of DHX36 siRNA was 50 nM. RNA products are shown schematically to the right. Error bars are SDs of two transfection experiments.

DETAILED DESCRIPTION

[0045] Genes such as eukaryotic genes contain intervening sequences or introns that must be accurately removed from primary transcripts to create functional mRNAs capable of encoding proteins (1). This process modifies mRNA composition in a highly dynamic manner, employing interdependent interactions of five small nuclear RNAs (snRNAs) (e.g., U1, U2, U4, U5, and U6) and a large number of proteins with conserved but degenerate sequences in the pre-mRNA (2). In particular, introns can be defined by the core splice site elements that comprises the 5' splice site (5'ss) which can comprise a conserved GU motif, the 3' splice site (3'ss) which can terminates with an invariant AG motif, the branchpoint sequence which can comprise a conserved adenine base, and the polypyrimidine (Py) tract. Splicing reaction can be initiated upon binding of the U1 snRNP to the conserved GU motif on the 5'ss, followed by binding of U2 snRNP at the conserved adenine base of the branchpoint, and finally U4, U5, and U6 snRNPs interactions near the 5' and 3' splice site. The complex formed by snRNPs and the respective intron can be referred to as a spliceosome. Additional splicing factors such as U2 small nuclear RNA auxiliary factor 1 (U2AF35), U2AF2 (U2AF65) and splicing factor 1 (SF1) can contribute to the spliceosome assembly and facilitate the splicing event.

[0046] Intron splicing generally promotes mRNA accumulation and protein expression across species (3-5). This process can be altered by intronic mutations or variants that may also impair coupled gene expression pathways, including transcription, mRNA export and translation. This is exemplified by introns in the 5'UTR where natural variants or mutations modifying intron retention alter the relative abundance of transcripts with upstream open reading frames (uORFs) or other regulatory motifs and dramatically influence translation (6,7). Further, impaired protein translation due to a defective splicing such as intron retention has led to development of diseases and/or progression of diseases such as genetic disorders or conditions (e.g., hereditary diseases or cancer). However, successful sequence-specific strategies to normalize gene expression in such situations have not been developed.

[0047] Splice-switching oligonucleotides (SSOs) are antisense reagents that modulate intron splicing by binding splice-site recognition or regulatory sequences and competing with *cis*- and *trans*-acting factors for their targets (8). They have been shown to restore aberrant splicing, modify the relative expression of existing mRNAs or produce novel splice variants that are not normally expressed (8). Improved stability of targeted SSO-RNA duplexes by a number of SSO modifications, such as 2'-O-methyl and 2'-O-methoxyethyl ribose, facilitated studies exploring their therapeutic potential for a growing number of human disease genes, including DMD in muscular dystrophy (9,10), SMN2 in

spinal muscular atrophy (11), ATM in ataxia-telangiectasia (12) and BTK in X-linked agammaglobulinaemia (13). Although such approaches are close to achieving their clinical potential for a restricted number of diseases (8), >300 Mendelian disorders resulting from mutation-induced aberrant splicing (14) and a growing number of complex traits may be amenable to SSO-mediated correction of gene expression.

[0048] Etiology of type 1 diabetes has a strong genetic component conferred by human leukocyte antigens (HLA) and a number of modifying non-HLA loci (15). The strongest modifier was identified in the proinsulin gene (INS) region on chromosome 11 (termed IDDM2) (15). Further mapping of this area suggested that INS is the most likely IDDM2 target (16), consistent with a critical role of this autoantigen in pathogenesis (17). Genetic risk to this disease at IDDM2 has been attributed to differential steady-state RNA levels from predisposing and protective INS haplotypes, potentially involving a minisatellite DNA sequence upstream of this gene (18,19). However, systematic examination of naturally occurring INS polymorphisms revealed haplotype-specific proinsulin expression levels in reporter constructs devoid of the minisatellite sequence, resulting from two variants in intron 1 (7), termed IVS1+5ins4 (also known as rs3842740 or INS-69) and IVS1-6A/T (rs689, INS-27 or HphI+/-) (16,20). The former variant activates a cryptic 5' splice site of intron 1 whereas adenine (A) at the latter variant, which resides 6 nucleotides upstream of the 3' splice site (3'ss), promotes intron retention, expanding the relative abundance of transcripts with extended 5'UTR (21). As compared to thymine (T), the A allele at IVS1-6A/T decreases affinity to pyrimidine-binding proteins in vitro and renders the 3'ss more dependent on the auxiliary factor of U2 small nuclear ribonucleoprotein (U2AF) (7), a heterodimer required for U2 binding, spliceosome assembly and 3'ss selection (22). Intron 1-containing transcripts are overrepresented in IVS1-6A-derived cDNA libraries prepared from insulin producing tissues (21), are exported from the nucleus (23), and contain a short, Homininae-specific uORF that co-evolved with relaxation of the 3'ss of intron 1 in higher primates (7). The lower proinsulin expression conferred by the A allele may lead to suboptimal presentation of proinsulin peptides in the fetal thymus and inadequate negative selection of autoreactive T cells, culminating in autoimmune destruction of insulin-producing β cells in the pancreas (7).

[0049] An aim of the invention is to induce processing of a partially processed mRNA transcript to remove a retained intron to produce a fully processed mRNA transcript that encodes a full-length functional form of a protein. An additional aim is to treat a disease or condition characterized by

impaired production of a protein and/or a disease or condition characterized by a defective splicing in a subject in need thereof.

[0050] Another aim of the invention is to correct the low efficiency of INS intron 1 removal from the IVS1-6A-containing pre-mRNAs and reduce intron retention to the levels observed for the disease-protective T allele. A further aim of the invention is to provide new therapy approaches to genetic diseases including cancer that are characterized by (or associated with) irregular or aberrant intron retention.

[0051] According to a first aspect of the invention, there is provided a method of prevention or treatment of a disease in a subject comprising correction of intron retention in mature gene transcripts, wherein the disease is induced by defective protein expression caused by the intron retention in the gene transcripts.

[0052] Retained intron

[0053] Retained intron is one of five types of alternative splicing that can also include exon skipping, alternative 5' splice site, alternative 3' splice site, and mutually exclusive exons. Exon skipping can occur when an exon is skipped over or is spliced out of the processed mRNA and can be the most common type of alternative splicing. Alternative 5' ss and alternative 3' ss can signify alternate splice sites such as cryptic splice sites or pseudo splice sites. Mutually exclusive exons can occur when only one of two exons is retained in the processed mRNA after splicing. Although intron retention can be less common than exon skipping, intron retention has been shown to occur more frequently than previously realized. Indeed, a study of 21,106 human genes by the De Souza group has shown that about 15% of the genes tested showed intron retention (*see*, Galante et al, "Detection and evaluation of intron retention events in the human transcriptome," *Bioinformatics* **10**:757-765 (2004)). Further, a study by the Moore group has shown that about 35% of human 5'-UTRs and about 16% of 3'-UTRs harbor introns (Bicknell et al., "Introns in UTRs: Why we should stop ignoring them," *Bioessays* **34**:1025-1034 (2012)). As such, intron retention can occur in a coding region, a non-coding region, at the 5' UTR, or at the 3' UTR. In the coding region, the retained intron can encode amino acids in frame, or can be in misalignment which can generate truncated proteins or non-functional proteins due to stop codon or frame shifts. Further, intron can be in between two exons, located at the 5' UTR, or located at the 3' UTR.

[0054] Compared to a non-retained intron, a retained intron can be characterized to have a shorter sequence length. The sequence length of the retained intron can be less than 5 kb, less than 4 kb, less

than 3 kb, less than 2.5 kb, less than 2 kb, less than 1.5 kb, less than 1 kb, less than 0.5 kb, less than 0.4 kb, less than 0.3 kb, less than 0.2 kb, or less than 0.1 kb.

[0055] Further, compared to a non-retained intron, a retained intron can be characterized to have a G/C content of between about 40% to about 60%. The G/C content of the retained intron can be about 40%, 45%, 50%, 55%, or about 60%.

[0056] The retained intron can also be flanked by a weak 5' splice site, a weak 3' splice site, or both. A weak splice site can refer to a splice site that may require a regulatory protein such as an intronic splicing enhancer for function.

[0057] Furthermore, a retained intron may comprise a lower presence of a GGG motif relative to a non-retained intron. In some instances, the GGG motif is an intronic splicing enhancer.

[0058] A partially processed mRNA transcript is a mRNA transcript that has undergone partial splicing and comprises at least one retained intron. The at least one intron can be within the 5' UTR, 3' UTR, or at an internal position in between two exons. The partially processed mRNA transcript can be unable to be translated to produce a functional or full-length protein.

[0059] A partially processed mRNA transcript can comprise an intron that is characterized by a short sequence length. A partially processed mRNA transcript can comprise an intron that is characterized by a sequence of less than 3 kb, less than 2.5 kb, less than 2 kb, less than 1.5 kb, less than 1 kb, less than 0.5 kb, less than 0.4 kb, less than 0.3 kb, less than 0.2 kb, or less than 0.1 kb.

[0060] A partially processed mRNA transcript can comprise an intron that is characterized to have a G/C content of between about 40% to about 60%. A partially processed mRNA transcript can comprise an intron that is characterized to have a G/C content of about 40%, 45%, 50%, 55%, or about 60%.

[0061] A partially processed mRNA transcript can comprise an intron that is flanked by a weak 5' splice site, a weak 3' splice site, or both.

[0062] A partially processed mRNA transcript can comprise an intron that may comprise a lower presence of a GGG motif relative to a non-retained intron.

[0063] In some cases, one or more introns are retained in a partially processed mRNA transcript. The partially processed mRNA can comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more retained introns.

[0064] A fully processed mRNA transcript is one that has undergone splicing to remove introns, such as retained introns and is capable of being translated to produce a protein such as a full length functional protein. A partially processed mRNA transcript with one or more retained introns can be spliced so as to become a fully processed mRNA.

[0065] A full-length functional protein has the same length and function as the wild-type form of the protein. The full-length functional protein can be the wild-type form of the protein. The full-length functional protein can be an isoform of the wild type protein with the same length as the wild type protein. The full-length functional protein can comprise a mutation such as a substitution of an amino acid for an alternative amino acid residue with similar properties so that the phenotype (e.g., the function) of the protein is not altered by the mutation.

[0066] Exemplary amino acids with similar properties can include amino acids with electrically charged side chains: arginine, histidine, and lysine which are positively charged; or aspartic acid and glutamic acid which are negatively charged; polar amino acid residues: serine, threonine, asparagine, glutamine, cysteine, and methionine; nonpolar amino acids: glycine, alanine, valine, leucine, isoleucine, and proline; and aromatic amino acids: phenylalanine, tyrosine, and tryptophan.

[0067] In some instances, the partially processed mRNA transcript leads to impaired production of a protein. The impaired production of the protein can be the impaired production of a full-length functional form of the protein which can comprise sub-normal production of the full-length functional form of the protein. The impaired production of a full-length functional form of the protein can comprise production of a defective form of the protein. The defective form of the protein can be a truncated form of the protein, a mis-folded form of the protein, or a form of the protein that comprises an aberrant target binding site.

[0068] Polynucleic acid polymer

[0069] Polynucleic acid polymers described herein can be used to hybridize to a partially processed mRNA transcript to initiate removal of a retained intron. In some instances, the polynucleic acid polymer is an antisense polynucleic acid polymer. The antisense polynucleic acid polymer may hybridize to a region of a partially processed mRNA transcript with a high G/C content, such as a region that comprises between about 40% to about 60%. The antisense polynucleic acid polymer may comprise a high G/C content, such as between about 40% to about 60%. The antisense polynucleic acid polymer may hybridize to a region of a partially processed mRNA transcript that is at, next to, or near a weak 5' splice site, or a weak 3' splice site. The antisense polynucleic acid polymer may hybridize to a region of a partially processed mRNA transcript that may comprise a lower presence of a GGG motif.

[0070] The antisense sequence of the polynucleic acid polymer may hybridize to a retained intron of the partially processed mRNA transcript. The antisense sequence may hybridize to an intronic splicing regulatory element of the partially processed mRNA transcript. The intronic splicing

regulatory element may include an intronic splicing enhancer or an intronic splicing silencer. The anti-sense sequence may hybridize to an intronic splicing silencer. The intronic splicing regulatory element may modulate splicing by affecting the early and/or intermediate steps of spliceosomal assembly, such as when U1 and U2 snRNPs pair at splice sites across an exon during exon definition or during subsequent transition to intron-spanning complex. Exemplary intronic splicing regulatory elements may include polypyrimidine-tract-binding protein (PTB), U2AF65 and/or U2AF35, hnRNP L and hnRNP A1.

[0071] The polynucleic acid polymer may hybridize at the 5' splice site, 3' splice site, branchpoint, polypyrimidine tract, or an intron enhancer of the intron. The polynucleic acid polymer may also hybridize at a distance of about 30 bases away, 25 bases away, 20 bases away, 15 bases away, 10 bases away, or 5 bases away from a 5' splice site, 3' splice site, branchpoint, polypyrimidine tract or an intron enhancer to promote or enhance splicing. Hybridization of the polynucleic acid polymer at or near the splice sites may promote or enhance splicing by recruiting splicing factors toward the splice sites, thereby initiating splicing.

[0072] A polynucleic acid polymer may hybridize at an intron silencer site (e.g., *de novo* intron silencer site), a cryptic intron splice site, or a pseudo splice site. The intron silencer site can be recognized by a silencer which suppresses the splicing reaction. The cryptic intron splice site can be created by a mutation, such as a substitution, a deletion, or an insertion. The cryptic intron splice site can be a cryptic 5' splice site or a cryptic 3' splice site. The cryptic intron splice site can be a cryptic 5' splice site. The pseudo splice site can be a weak splice site in which its activation can be resulted from a mutation of the canonical splice site. The pseudo splice site can be a pseudo 5' splice site or a pseudo 3' splice site. Hybridization at an intron silencer site may sterically block a silencer from binding and may help from preventing disruption of the assembly and processing of the spliceosome. Hybridization at a cryptic intron splice site may redirect interaction toward a canonical splice site such as a canonical weak splice site.

[0073] A polynucleic acid polymer may hybridize to an internal region of the intron. The internal region of the intron may not encompass a splice site, such as a 5' splice site (e.g., canonical, cryptic, or pseudo 5' splice site), 3' splice site (e.g., canonical, cryptic, or pseudo 3' splice site), an enhancer site or a silencer site. Hybridization of the polynucleic acid polymer at an internal region may promote or enhance splicing by recruiting splicing factors toward the splice sites, such as an intron enhancer site.

[0074] A polynucleic acid polymer may hybridize to an intronic splicing regulatory element of the partially processed mRNA transcript. The intronic splicing regulatory element may comprise a CCC motif. The anti-sense sequence may hybridize to a binding motif of the partially processed mRNA transcript wherein the binding motif does not form a G quadruplex. The anti-sense sequence may hybridize to a binding motif of the partially processed mRNA transcript wherein the binding motif is between two G quadruplexes. The anti-sense sequence may also hybridize to a binding motif of the partially processed mRNA transcript wherein the binding motif has a first CCC motif and a second CCC motif. The first CCC motif may be about 3 or more nucleotide bases from the second CCC motif.

[0075] A polynucleic acid polymer may hybridize to a binding motif of the partially processed mRNA transcript in which the binding motif forms a hairpin structure. The polynucleic acid polymer may further be capable of destabilizing the hairpin structure.

[0076] A polynucleic acid polymer may comprise a pyridine nucleotide at the 3' terminal position and/or at the 5' terminal position. The polynucleic acid polymer may further comprise two consecutive pyridine nucleotides at the 3' terminal position and/or at the 5' terminal position.

[0077] In some instances, a polynucleic acid polymer may be a splice-switching oligonucleotide (SSO). The splice-switching oligonucleotide may hybridize to a retained intron of the partially processed mRNA transcript. The splice-switching oligonucleotide may hybridize to an intronic splicing regulatory element of the partially processed mRNA transcript.

[0078] Splice-switching oligonucleotides (SSOs) have been widely used to inhibit exon usage but antisense strategies that promote removal of entire introns to increase splicing-mediated gene expression have not been developed. Using a series of splicing reporters containing the human proinsulin gene, it has been shown that *INS* intron 1 retention by SSOs that bind transcripts derived from a human haplotype expressing low levels of proinsulin can be reduced. The SSO-assisted promotion of weak intron removal from the 5'UTR through competing noncanonical and canonical RNA structures facilitates development of sequence-based antisense strategies to enhance gene expression.

[0079] The term "correction of intron retention" is understood to the correction of irregular or aberrant intron retention. The correction may be complete correction or partial correction. Correction may comprise reducing the incidence of intron retention. Correction may comprise reducing aberrant intron retention.

[0080] Reference to "defective" in the context of protein expression caused by the intron retention in the gene transcripts, may comprise inadequate, defective, or aberrant protein expression.

[0081] Correction of intron retention may comprise administering a polynucleic acid polymer arranged to hybridize with the gene transcript. Correction of intron retention may comprise administering a polynucleic acid polymer arranged to hybridize with the gene transcript in order to alter higher-order structures in the gene transcript. Correction of intron retention may comprise administering a polynucleic acid polymer arranged to hybridize with the gene transcript in order to interfere with one or more of conformational transitions of canonical (stem loops); noncanonical (G quadruplex) RNA structures; interactions with trans-acting factors; and the rate of RNA–protein complex formation. Correction of intron retention may comprise administering a polynucleic acid polymer arranged to hybridize with the gene transcript in order to interfere with, such as block, conformational transitions of canonical (stem loops) and/or noncanonical (G quadruplex) RNA structures. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of intronic splicing regulatory elements. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of overlapped intronic splicing regulatory elements conserved in mammals. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of intronic segments containing short penta- to heptamer splicing regulatory motifs. Splicing regulatory motifs may comprise CCCAG or AGGCC. Splicing regulatory motifs may comprise any of the motifs provided by Yeo GW, et al (2007). PLoS Genet 3:e85 and/or Voelker RB, & Berglund JA (2007). Genome Res 17:1023-1033, both documents incorporated herein by reference. The target region may not comprise C runs in order to avoid G-quadruplex formation. The target region may not be proximal to both 5' and 3' splice sites, polypyrimidine tracts, branch sites and/or suprabranch regions.

[0082] The polynucleic acid polymer may provide binding platforms for splicing factors which have been shown to influence *INS* intron 1 and exon 2 splicing, including Tra2, SRSF3 or U2AF35, or other peptides, sense or antisense nucleic acids, small molecules, or other chemicals to facilitate delivery of the polynucleic acid polymer and/or target the nucleic acid to a specific tissue, cell or a developmental stage. Such antisense strategy may help reduce pervasive intron retention in cancer cells, particularly those that contain somatic mutations of splicing factor genes, as first shown for specific substitutions in the zinc finger domain of U2AF35 in myeloproliferative diseases (76). These mutations occur in many tumours and result in splicing defects that may contribute to

malignant growth. The invention may help control cell proliferation and reduce malignant growth in a significant fraction of cancer patients that carry mutations in splicing factors involved in 3' splice site recognition, currently estimated at >15% (76).

[0083] The sequence of the polynucleic acid polymer may be at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99%, or 99.5% complementary to a target sequence of the partially processed mRNA transcript. The sequence of the polynucleic acid polymer may be 100% complementary to a target sequence of the partially processed mRNA transcript.

[0084] The sequence of the polynucleic acid polymer may have 4 or less mismatches to a target sequence of the partially processed mRNA transcript. The sequence of the polynucleic acid polymer may have 3 or less mismatches to a target sequence of the partially processed mRNA transcript. The sequence of the polynucleic acid polymer may have 2 or less mismatches to a target sequence of the partially processed mRNA transcript. The sequence of the polynucleic acid polymer may have 1 or less mismatches to a target sequence of the partially processed mRNA transcript.

[0085] The polynucleic acid polymer may specifically hybridize to a target sequence of the partially processed mRNA transcript. The specificity may be a 95%, 98%, 99%, 99.5% or 100% sequence complementarity of the polynucleic acid polymer to a target sequence of the partially processed mRNA transcript. The hybridization may be under high stringent hybridization conditions.

[0086] The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of SEQ ID NO: 46. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 99% identity with SEQ ID NO: 46. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 98% identity with SEQ ID NO: 46. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 95% identity with SEQ ID NO: 46. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 90% identity with SEQ ID NO: 46. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 85% identity with SEQ ID NO: 46. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 80% identity with SEQ ID NO: 46. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 75% identity with SEQ ID NO: 46. The

polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 70% identity with SEQ ID NO: 46. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 65% identity with SEQ ID NO: 46. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 60% identity with SEQ ID NO: 46. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 55% identity with SEQ ID NO: 46. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 50% identity with SEQ ID NO: 46.

[0087] A polynucleic acid polymer may hybridize to an mRNA transcript comprising at least 80%, 85%, 90%, 95%, or 99% sequence identity to at least 13 contiguous bases of SEQ ID NO: 46. The polynucleic acid polymer may hybridize to a mRNA transcript comprising 100% sequence identity to at least 13 contiguous bases of SEQ ID NO: 46.

[0088] A polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of SEQ ID NO: 3. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 99% identity with SEQ ID NO: 3. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 98% identity with SEQ ID NO: 3. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 95% identity with SEQ ID NO: 3. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 90% identity with SEQ ID NO: 3. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 85% identity with SEQ ID NO: 3. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 80% identity with SEQ ID NO: 3. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 75% identity with SEQ ID NO: 3. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 70% identity with SEQ ID NO: 3. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript

comprising or consisting of a sequence having at least 65% identity with SEQ ID NO: 3. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 60% identity with SEQ ID NO: 3. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 55% identity with SEQ ID NO: 3. The polynucleic acid polymer may be antisense to at least part of a target region of the transcript comprising or consisting of a sequence having at least 50% identity with SEQ ID NO: 3.

[0089] Reference to antisense to at least part of a target region of the transcript may comprise at least 5 consecutive nucleotides, or at least 10 consecutive nucleotides. The polynucleic acid polymer may be antisense to at least 5 consecutive nucleotides of a target region of the transcript comprising or consisting of SEQ ID NO: 46. The polynucleic acid polymer may be antisense to at least 10 consecutive nucleotides of a target region of the transcript comprising or consisting of SEQ ID NO: 46. The polynucleic acid polymer may be antisense to at least 15 consecutive nucleotides of a target region of the transcript comprising or consisting of SEQ ID NO: 46. The polynucleic acid polymer may be antisense to at least 20 consecutive nucleotides of a target region of the transcript comprising or consisting of SEQ ID NO: 46.

[0090] A polynucleic acid polymer may hybridize to a mRNA transcript comprising at least 10 contiguous bases of SEQ ID NO: 46. The polynucleic acid polymer may hybridize to a mRNA transcript consisting of at least 10 contiguous bases of SEQ ID NO: 46.

[0091] The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of a sequence selected from any of the group comprising SEQ ID NO: 3; SEQ ID NO: 6; SEQ ID NO: 9; SEQ ID NO: 12; SEQ ID NO: 15; SEQ ID NO: 18; SEQ ID NO: 21; SEQ ID NO: 24; SEQ ID NO: 27; SEQ ID NO: 30; SEQ ID NO: 33; SEQ ID NO: 36; SEQ ID NO: 39; SEQ ID NO: 42; SEQ ID NO: 45; or combinations thereof.

[0092] A polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of a sequence having at least 99% sequence identity to a sequence selected from any of the group comprising SEQ ID NO: 3; SEQ ID NO: 6; SEQ ID NO: 9; SEQ ID NO: 12; SEQ ID NO: 15; SEQ ID NO: 18; SEQ ID NO: 21; SEQ ID NO: 24; SEQ ID NO: 27; SEQ ID NO: 30; SEQ ID NO: 33; SEQ ID NO: 36; SEQ ID NO: 39; SEQ ID NO: 42; SEQ ID NO: 45; or combinations thereof. The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of a sequence having at least 98% sequence identity to a sequence selected from any of the group comprising SEQ ID NO: 3; SEQ ID NO: 6; SEQ ID NO: 9; SEQ ID NO: 12;

SEQ ID NO: 15; SEQ ID NO: 18; SEQ ID NO: 21; SEQ ID NO: 24; SEQ ID NO: 27; SEQ ID NO: 30; SEQ ID NO: 33; SEQ ID NO: 36; SEQ ID NO: 39; SEQ ID NO: 42; SEQ ID NO: 45; or combinations thereof. The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of a sequence having at least 95% sequence identity to a sequence selected from any of the group comprising SEQ ID NO: 3; SEQ ID NO: 6; SEQ ID NO: 9; SEQ ID NO: 12; SEQ ID NO: 15; SEQ ID NO: 18; SEQ ID NO: 21; SEQ ID NO: 24; SEQ ID NO: 27; SEQ ID NO: 30; SEQ ID NO: 33; SEQ ID NO: 36; SEQ ID NO: 39; SEQ ID NO: 42; SEQ ID NO: 45; or combinations thereof. The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of a sequence having at least 90% sequence identity to a sequence selected from any of the group comprising SEQ ID NO: 3; SEQ ID NO: 6; SEQ ID NO: 9; SEQ ID NO: 12; SEQ ID NO: 15; SEQ ID NO: 18; SEQ ID NO: 21; SEQ ID NO: 24; SEQ ID NO: 27; SEQ ID NO: 30; SEQ ID NO: 33; SEQ ID NO: 36; SEQ ID NO: 39; SEQ ID NO: 42; SEQ ID NO: 45; or combinations thereof. The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of a sequence having at least 85% sequence identity to a sequence selected from any of the group comprising SEQ ID NO: 3; SEQ ID NO: 6; SEQ ID NO: 9; SEQ ID NO: 12; SEQ ID NO: 15; SEQ ID NO: 18; SEQ ID NO: 21; SEQ ID NO: 24; SEQ ID NO: 27; SEQ ID NO: 30; SEQ ID NO: 33; SEQ ID NO: 36; SEQ ID NO: 39; SEQ ID NO: 42; SEQ ID NO: 45; or combinations thereof. The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of a sequence having at least 80% sequence identity to a sequence selected from any of the group comprising SEQ ID NO: 3; SEQ ID NO: 6; SEQ ID NO: 9; SEQ ID NO: 12; SEQ ID NO: 15; SEQ ID NO: 18; SEQ ID NO: 21; SEQ ID NO: 24; SEQ ID NO: 27; SEQ ID NO: 30; SEQ ID NO: 33; SEQ ID NO: 36; SEQ ID NO: 39; SEQ ID NO: 42; SEQ ID NO: 45; or combinations thereof. The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of a sequence having at least 75% sequence identity to a sequence selected from any of the group comprising SEQ ID NO: 3; SEQ ID NO: 6; SEQ ID NO: 9; SEQ ID NO: 12; SEQ ID NO: 15; SEQ ID NO: 18; SEQ ID NO: 21; SEQ ID NO: 24; SEQ ID NO: 27; SEQ ID NO: 30; SEQ ID NO: 33; SEQ ID NO: 36; SEQ ID NO: 39; SEQ ID NO: 42; SEQ ID NO: 45; or combinations thereof. The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of a sequence having at least 70% sequence identity to a sequence selected from any of the group comprising SEQ ID NO: 3; SEQ ID NO: 6; SEQ ID NO: 9; SEQ ID NO: 12; SEQ ID NO: 15; SEQ ID NO: 18; SEQ ID NO: 21; SEQ ID NO: 24; SEQ ID NO: 27; SEQ ID NO: 30; SEQ ID NO: 33; SEQ ID NO: 36; SEQ ID NO: 39; SEQ ID NO: 42; SEQ ID

NO: 45; or combinations thereof. The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of a sequence having at least 65% sequence identity to a sequence selected from any of the group comprising SEQ ID NO: 3; SEQ ID NO: 6; SEQ ID NO: 9; SEQ ID NO: 12; SEQ ID NO: 15; SEQ ID NO: 18; SEQ ID NO: 21; SEQ ID NO: 24; SEQ ID NO: 27; SEQ ID NO: 30; SEQ ID NO: 33; SEQ ID NO: 36; SEQ ID NO: 39; SEQ ID NO: 42; SEQ ID NO: 45; or combinations thereof. The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of a sequence having at least 63% sequence identity to a sequence selected from any of the group comprising SEQ ID NO: 3; SEQ ID NO: 6; SEQ ID NO: 9; SEQ ID NO: 12; SEQ ID NO: 15; SEQ ID NO: 18; SEQ ID NO: 21; SEQ ID NO: 24; SEQ ID NO: 27; SEQ ID NO: 30; SEQ ID NO: 33; SEQ ID NO: 36; SEQ ID NO: 39; SEQ ID NO: 42; SEQ ID NO: 45; or combinations thereof. The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of a sequence having at least 60% sequence identity to a sequence selected from any of the group comprising SEQ ID NO: 3; SEQ ID NO: 6; SEQ ID NO: 9; SEQ ID NO: 12; SEQ ID NO: 15; SEQ ID NO: 18; SEQ ID NO: 21; SEQ ID NO: 24; SEQ ID NO: 27; SEQ ID NO: 30; SEQ ID NO: 33; SEQ ID NO: 36; SEQ ID NO: 39; SEQ ID NO: 42; SEQ ID NO: 45; or combinations thereof. The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of a sequence having at least 50% sequence identity to a sequence selected from any of the group comprising SEQ ID NO: 3; SEQ ID NO: 6; SEQ ID NO: 9; SEQ ID NO: 12; SEQ ID NO: 15; SEQ ID NO: 18; SEQ ID NO: 21; SEQ ID NO: 24; SEQ ID NO: 27; SEQ ID NO: 30; SEQ ID NO: 33; SEQ ID NO: 36; SEQ ID NO: 39; SEQ ID NO: 42; SEQ ID NO: 45; or combinations thereof.

[0093] The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of a sequence selected from any of the group comprising SEQ ID NO: 3; SEQ ID NO: 6; SEQ ID NO: 9; and SEQ ID NO: 36; or combinations thereof. The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of a sequence having at least 99% identity to a sequence selected from any of the group comprising SEQ ID NO: 3; SEQ ID NO: 6; SEQ ID NO: 9; and SEQ ID NO: 36; or combinations thereof. The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of a sequence having at least 98% identity to a sequence selected from any of the group comprising SEQ ID NO: 3; SEQ ID NO: 6; SEQ ID NO: 9; and SEQ ID NO: 36; or combinations thereof. The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of a sequence having at least

95% identity to a sequence selected from any of the group comprising SEQ ID NO: 3; SEQ ID NO: 6; SEQ ID NO: 9; and SEQ ID NO: 36; or combinations thereof.

[0094] The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of SEQ ID NO: 3. The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of SEQ ID NO: 6. The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of SEQ ID NO: 9. The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of SEQ ID NO: 36.

[0095] The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of a sequence having at least 99%, at least 98%, at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 65%, at least 63%, at least 60%, at least 55%, or at least 50% identity to SEQ ID NO: 3. The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of a sequence having at least 99%, at least 98%, at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 65%, at least 63%, at least 60%, at least 55%, or at least 50% identity to SEQ ID NO: 6. The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of a sequence having at least 99%, at least 98%, at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 65%, at least 63%, at least 60%, at least 55%, or at least 50% identity to SEQ ID NO: 9. The polynucleic acid polymer may be antisense to a target region of the transcript comprising or consisting of a sequence having at least 99%, at least 98%, at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 65%, at least 63%, at least 60%, at least 55%, or at least 50% identity to SEQ ID NO: 36.

[0096] The polynucleic acid polymer may be antisense to a region of the transcript comprising or consisting of a sequence complementary to any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof. The polynucleic acid polymer may be antisense to a region of the transcript comprising or consisting of a sequence complementary to a sequence having at least 99% identity with any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof. The polynucleic acid polymer may be antisense to a region of the transcript comprising or consisting of a sequence complementary to a sequence having at least 98% identity with any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof. The polynucleic acid polymer may be antisense to a region of the transcript comprising or consisting of a sequence complementary to a sequence having at least 95% identity with any of the group of

sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof. The polynucleic acid polymer may be antisense to a region of the transcript comprising or consisting of a sequence complementary to a sequence having at least 90% identity with any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof. The polynucleic acid polymer may be antisense to a region of the transcript comprising or consisting of a sequence complementary to a sequence having at least 85% identity with any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof. The polynucleic acid polymer may be antisense to a region of the transcript comprising or consisting of a sequence complementary to a sequence having at least 80% identity with any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof. The polynucleic acid polymer may be antisense to a region of the transcript comprising or consisting of a sequence complementary to a sequence having at least 75% identity with any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof. The polynucleic acid polymer may be antisense to a region of the transcript comprising or consisting of a sequence complementary to a sequence having at least 70% identity with any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof. The polynucleic acid polymer may be antisense to a region of the transcript comprising or consisting of a sequence complementary to a sequence having at least 65% identity with any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof. The polynucleic acid polymer may be antisense to a region of the transcript comprising or consisting of a sequence complementary to a sequence having at least 60% identity with any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof. The polynucleic acid polymer may be antisense to a region of the transcript comprising or consisting of a sequence complementary to a sequence having at least 55% identity with any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof. The polynucleic acid polymer may be antisense to a region of the transcript comprising or consisting of a sequence complementary to a sequence having at least 50% identity with any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof. The skilled person will understand that uracil nucleotides may be substituted with thymine nucleotides (e.g. the DNA form of the RNA of such sequences).

[0097] The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO:

22; SEQ ID NO: 23; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 28; SEQ ID NO: 29; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 34; SEQ ID NO: 35; SEQ ID NO: 37; SEQ ID NO: 38; SEQ ID NO: 40; SEQ ID NO: 41; SEQ ID NO: 43; and SEQ ID NO: 44; or combinations thereof.

[0098] The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 99% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 28; SEQ ID NO: 29; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 34; SEQ ID NO: 35; SEQ ID NO: 37; SEQ ID NO: 38; SEQ ID NO: 40; SEQ ID NO: 41; SEQ ID NO: 43; and SEQ ID NO: 44; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 98% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 28; SEQ ID NO: 29; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 34; SEQ ID NO: 35; SEQ ID NO: 37; SEQ ID NO: 38; SEQ ID NO: 40; SEQ ID NO: 41; SEQ ID NO: 43; and SEQ ID NO: 44; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 95% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 28; SEQ ID NO: 29; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 34; SEQ ID NO: 35; SEQ ID NO: 37; SEQ ID NO: 38; SEQ ID NO: 40; SEQ ID NO: 41; SEQ ID NO: 43; and SEQ ID NO: 44; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 90% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 28; SEQ ID NO: 29; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 34; SEQ ID NO: 35; SEQ ID NO: 37; SEQ ID NO: 38; SEQ ID NO: 40; SEQ ID NO: 41; SEQ ID NO: 43; and SEQ ID NO: 44; or

combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 85% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 28; SEQ ID NO: 29; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 34; SEQ ID NO: 35; SEQ ID NO: 37; SEQ ID NO: 38; SEQ ID NO: 40; SEQ ID NO: 41; SEQ ID NO: 43; and SEQ ID NO: 44; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 80% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 28; SEQ ID NO: 29; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 34; SEQ ID NO: 35; SEQ ID NO: 37; SEQ ID NO: 38; SEQ ID NO: 40; SEQ ID NO: 41; SEQ ID NO: 43; and SEQ ID NO: 44; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 75% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 28; SEQ ID NO: 29; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 34; SEQ ID NO: 35; SEQ ID NO: 37; SEQ ID NO: 38; SEQ ID NO: 40; SEQ ID NO: 41; SEQ ID NO: 43; and SEQ ID NO: 44; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 70% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 28; SEQ ID NO: 29; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 34; SEQ ID NO: 35; SEQ ID NO: 37; SEQ ID NO: 38; SEQ ID NO: 40; SEQ ID NO: 41; SEQ ID NO: 43; and SEQ ID NO: 44; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 65% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 5;

SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 28; SEQ ID NO: 29; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 34; SEQ ID NO: 35; SEQ ID NO: 37; SEQ ID NO: 38; SEQ ID NO: 40; SEQ ID NO: 41; SEQ ID NO: 43; and SEQ ID NO: 44; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 60% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 28; SEQ ID NO: 29; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 34; SEQ ID NO: 35; SEQ ID NO: 37; SEQ ID NO: 38; SEQ ID NO: 40; SEQ ID NO: 41; SEQ ID NO: 43; and SEQ ID NO: 44; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 55% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 28; SEQ ID NO: 29; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 34; SEQ ID NO: 35; SEQ ID NO: 37; SEQ ID NO: 38; SEQ ID NO: 40; SEQ ID NO: 41; SEQ ID NO: 43; and SEQ ID NO: 44; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 50% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 28; SEQ ID NO: 29; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 34; SEQ ID NO: 35; SEQ ID NO: 37; SEQ ID NO: 38; SEQ ID NO: 40; SEQ ID NO: 41; SEQ ID NO: 43; and SEQ ID NO: 44; or combinations thereof.

[0099] The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO 5; SEQ IN NO: 7; SEQ ID NO: 8; SEQ ID NO: 34; and SEQ ID NO: 35; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 99% identity to a sequence selected from any of the group

comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO 5; SEQ IN NO: 7; SEQ ID NO: 8; SEQ ID NO: 34; and SEQ ID NO: 35; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 98% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO 5; SEQ IN NO: 7; SEQ ID NO: 8; SEQ ID NO: 34; and SEQ ID NO: 35; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 95% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO 5; SEQ IN NO: 7; SEQ ID NO: 8; SEQ ID NO: 34; and SEQ ID NO: 35; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 90% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO 5; SEQ IN NO: 7; SEQ ID NO: 8; SEQ ID NO: 34; and SEQ ID NO: 35; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 85% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO 5; SEQ IN NO: 7; SEQ ID NO: 8; SEQ ID NO: 34; and SEQ ID NO: 35; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 80% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO 5; SEQ IN NO: 7; SEQ ID NO: 8; SEQ ID NO: 34; and SEQ ID NO: 35; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 75% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO 5; SEQ IN NO: 7; SEQ ID NO: 8; SEQ ID NO: 34; and SEQ ID NO: 35; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 70% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO 5; SEQ IN NO: 7; SEQ ID NO: 8; SEQ ID NO: 34; and SEQ ID NO: 35; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 65% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO 5; SEQ IN NO: 7; SEQ ID NO: 8; SEQ ID NO: 34; and SEQ ID NO: 35; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 60% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2;

SEQ ID NO: 4; SEQ ID NO 5; SEQ IN NO: 7; SEQ ID NO: 8; SEQ ID NO: 34; and SEQ ID NO: 35; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 55% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO 5; SEQ IN NO: 7; SEQ ID NO: 8; SEQ ID NO: 34; and SEQ ID NO: 35; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 50% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO 5; SEQ IN NO: 7; SEQ ID NO: 8; SEQ ID NO: 34; and SEQ ID NO: 35; or combinations thereof.

[00100] The polynucleic acid polymer may comprise or consist of SEQ ID NO: 1 or SEQ ID NO: 2. The polynucleic acid polymer may comprise or consist of SEQ ID NO: 3 or SEQ ID NO: 4. The polynucleic acid polymer may comprise or consist of SEQ ID NO: 7 or SEQ ID NO: 8. The polynucleic acid polymer may comprise or consist of SEQ ID NO: 34 or SEQ ID NO: 35.

[00101] The polynucleic acid polymer may comprise or consist of a sequence having at least 99% identity to SEQ ID NO: 1 or SEQ ID NO: 2. The polynucleic acid polymer may comprise or consist of a sequence having at least 98% identity to SEQ ID NO: 1 or SEQ ID NO: 2. The polynucleic acid polymer may comprise or consist of a sequence having at least 95% identity to SEQ ID NO: 1 or SEQ ID NO: 2. The polynucleic acid polymer may comprise or consist of a sequence having at least 90% identity to SEQ ID NO: 1 or SEQ ID NO: 2. The polynucleic acid polymer may comprise or consist of a sequence having at least 85% identity to SEQ ID NO: 1 or SEQ ID NO: 2. The polynucleic acid polymer may comprise or consist of a sequence having at least 80% identity to SEQ ID NO: 1 or SEQ ID NO: 2. The polynucleic acid polymer may comprise or consist of a sequence having at least 75% identity to SEQ ID NO: 1 or SEQ ID NO: 2. The polynucleic acid polymer may comprise or consist of a sequence having at least 70% identity to SEQ ID NO: 1 or SEQ ID NO: 2. The polynucleic acid polymer may comprise or consist of a sequence having at least 65% identity to SEQ ID NO: 1 or SEQ ID NO: 2. The polynucleic acid polymer may comprise or consist of a sequence having at least 60% identity to SEQ ID NO: 1 or SEQ ID NO: 2.

[00102] The polynucleic acid polymer may comprise or consist of a sequence having at least 99% identity to SEQ ID NO: 3 or SEQ ID NO: 4. The polynucleic acid polymer may comprise or consist of a sequence having at least 98% identity to SEQ ID NO: 3 or SEQ ID NO: 4. The polynucleic acid polymer may comprise or consist of a sequence having at least 95% identity to SEQ ID NO: 3 or SEQ ID NO: 4. The polynucleic acid polymer may comprise or consist of a sequence

having at least 90% identity to SEQ ID NO: 3 or SEQ ID NO: 4. The polynucleic acid polymer may comprise or consist of a sequence having at least 85% identity to SEQ ID NO: 3 or SEQ ID NO: 4. The polynucleic acid polymer may comprise or consist of a sequence having at least 80% identity to SEQ ID NO: 3 or SEQ ID NO: 4. The polynucleic acid polymer may comprise or consist of a sequence having at least 75% identity to SEQ ID NO: 3 or SEQ ID NO: 4. The polynucleic acid polymer may comprise or consist of a sequence having at least 70% identity to SEQ ID NO: 3 or SEQ ID NO: 4. The polynucleic acid polymer may comprise or consist of a sequence having at least 65% identity to SEQ ID NO: 3 or SEQ ID NO: 4. The polynucleic acid polymer may comprise or consist of a sequence having at least 60% identity to SEQ ID NO: 3 or SEQ ID NO: 4.

[00103] The polynucleic acid polymer may comprise or consist of a sequence having at least 99% identity to SEQ ID NO: 7 or SEQ ID NO: 8. The polynucleic acid polymer may comprise or consist of a sequence having at least 98% identity to SEQ ID NO: 7 or SEQ ID NO: 8. The polynucleic acid polymer may comprise or consist of a sequence having at least 95% identity to SEQ ID NO: 7 or SEQ ID NO: 8. The polynucleic acid polymer may comprise or consist of a sequence having at least 90% identity to SEQ ID NO: 7 or SEQ ID NO: 8. The polynucleic acid polymer may comprise or consist of a sequence having at least 85% identity to SEQ ID NO: 7 or SEQ ID NO: 8. The polynucleic acid polymer may comprise or consist of a sequence having at least 80% identity to SEQ ID NO: 7 or SEQ ID NO: 8. The polynucleic acid polymer may comprise or consist of a sequence having at least 75% identity to SEQ ID NO: 7 or SEQ ID NO: 8. The polynucleic acid polymer may comprise or consist of a sequence having at least 70% identity to SEQ ID NO: 7 or SEQ ID NO: 8. The polynucleic acid polymer may comprise or consist of a sequence having at least 65% identity to SEQ ID NO: 7 or SEQ ID NO: 8. The polynucleic acid polymer may comprise or consist of a sequence having at least 60% identity to SEQ ID NO: 7 or SEQ ID NO: 8.

[00104] The polynucleic acid polymer may comprise or consist of a sequence having at least 99% identity to SEQ ID NO: 34 or SEQ ID NO: 35. The polynucleic acid polymer may comprise or consist of a sequence having at least 98% identity to SEQ ID NO: 34 or SEQ ID NO: 35. The polynucleic acid polymer may comprise or consist of a sequence having at least 95% identity to SEQ ID NO: 34 or SEQ ID NO: 35. The polynucleic acid polymer may comprise or consist of a sequence having at least 90% identity to SEQ ID NO: 34 or SEQ ID NO: 35. The polynucleic acid polymer may comprise or consist of a sequence having at least 85% identity to SEQ ID NO: 34 or SEQ ID NO: 35. The polynucleic acid polymer may comprise or consist of a sequence having at least 80% identity to SEQ ID NO: 34 or SEQ ID NO: 35. The polynucleic acid polymer may comprise or

consist of a sequence having at least 75% identity to SEQ ID NO: 34 or SEQ ID NO: 35. The polynucleic acid polymer may comprise or consist of a sequence having at least 70% identity to SEQ ID NO: 34 or SEQ ID NO: 35. The polynucleic acid polymer may comprise or consist of a sequence having at least 65% identity to SEQ ID NO: 34 or SEQ ID NO: 35. The polynucleic acid polymer may comprise or consist of a sequence having at least 60% identity to SEQ ID NO: 34 or SEQ ID NO: 35.

[00105] The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof, wherein uracil nucleotides are substituted with thymine nucleotides (e.g. the DNA form of the RNA of such sequences).

[00106] The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 99% identity with a sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 99% identity with a sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof, wherein uracil nucleotides are substituted with thymine nucleotides (e.g. the DNA form of the RNA of such sequences). The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 98% identity with a sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 98% identity with a sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof, wherein uracil nucleotides are substituted with thymine nucleotides (e.g. the DNA form of the RNA of such sequences). The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 95% identity with a sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof. The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 95% identity with a sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof, wherein uracil nucleotides are substituted with thymine nucleotides (e.g. the DNA form of the RNA of such sequences). The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 90% identity with a sequence selected from any of the group comprising

SEQ ID NOs: 47 to 434; or combinations thereof, wherein uracil nucleotides are substituted with thymine nucleotides (e.g. the DNA form of the RNA of such sequences). The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 85% identity with a sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof, wherein uracil nucleotides are substituted with thymine nucleotides (e.g. the DNA form of the RNA of such sequences). The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 80% identity with a sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof, wherein uracil nucleotides are substituted with thymine nucleotides (e.g. the DNA form of the RNA of such sequences). The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 75% identity with a sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof, wherein uracil nucleotides are substituted with thymine nucleotides (e.g. the DNA form of the RNA of such sequences). The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 70% identity with a sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof, wherein uracil nucleotides are substituted with thymine nucleotides (e.g. the DNA form of the RNA of such sequences). The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 65% identity with a sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof, wherein uracil nucleotides are substituted with thymine nucleotides (e.g. the DNA form of the RNA of such sequences). The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 60% identity with a sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof, wherein uracil nucleotides are substituted with thymine nucleotides (e.g. the DNA form of the RNA of such sequences). The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 55% identity with a sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof, wherein uracil nucleotides are substituted with thymine nucleotides (e.g. the DNA form of the RNA of such sequences). The polynucleic acid polymer may comprise or consist of a polynucleic acid polymer sequence having at least 50% identity with a sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof, wherein uracil nucleotides are substituted with thymine nucleotides (e.g. the DNA form of the RNA of such sequences).

[00107] Advantageously, the invention identifies SSOs reducing the relative abundance of intron transcripts (e.g., intron 1-retaining transcripts) and delineates the optimized antisense target at a single-nucleotide resolution.

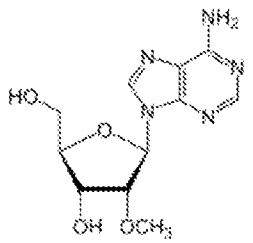
[00108] A polynucleic acid polymer may comprise a splice-switching oligonucleotide (SSO). The polynucleic acid polymer may be about 50 nucleotides in length. The polynucleic acid polymer may be about 45 nucleotides in length. The polynucleic acid polymer may be about 40 nucleotides in length. The polynucleic acid polymer may be about 35 nucleotides in length. The polynucleic acid polymer may be about 30 nucleotides in length. The polynucleic acid polymer may be about 25 nucleotides in length. The polynucleic acid polymer may be about 20 nucleotides in length. The polynucleic acid polymer may be about 19 nucleotides in length. The polynucleic acid polymer may be about 18 nucleotides in length. The polynucleic acid polymer may be about 17 nucleotides in length. The polynucleic acid polymer may be about 16 nucleotides in length. The polynucleic acid polymer may be about 15 nucleotides in length. The polynucleic acid polymer may be about 14 nucleotides in length. The polynucleic acid polymer may be about 13 nucleotides in length. The polynucleic acid polymer may be about 12 nucleotides in length. The polynucleic acid polymer may be about 11 nucleotides in length. The polynucleic acid polymer may be about 10 nucleotides in length. The polynucleic acid polymer may be between about 10 and about 50 nucleotides in length. The polynucleic acid polymer may be between about 10 and about 45 nucleotides in length. The polynucleic acid polymer may be between about 10 and about 40 nucleotides in length. The polynucleic acid polymer may be between about 10 and about 35 nucleotides in length. The polynucleic acid polymer may be between about 10 and about 30 nucleotides in length. The polynucleic acid polymer may be between about 10 and about 25 nucleotides in length. The polynucleic acid polymer may be between about 10 and about 20 nucleotides in length. The polynucleic acid polymer may be between about 15 and about 25 nucleotides in length. The polynucleic acid polymer may be between about 15 and about 30 nucleotides in length. The polynucleic acid polymer may be between about 12 and about 30 nucleotides in length.

[00109] A polynucleic acid polymer, such as the SSOs, may comprise RNA or DNA. The polynucleic acid polymer, such as the SSOs, may comprise RNA. The polynucleic acid polymer, such as the SSOs, may comprise natural or synthetic or artificial nucleotide analogues or bases, having equivalent complementation as DNA or RNA. The polynucleic acid polymer, such as the SSOs, may comprise combinations of DNA, RNA and/or nucleotide analogues. Nucleotide

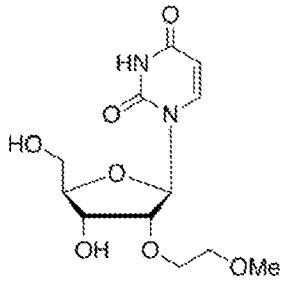
analogues may comprise PNA or LNA. In another embodiment, the nucleic acid, such as the SSOs, may comprise or consist of PMO.

[00110] In some instances, the synthetic or artificial nucleotide analogues or bases can comprise modifications at one or more of ribose moiety, phosphate moiety, nucleoside moiety, or a combination thereof.

[00111] Nucleotide analogues or artificial nucleotide base may comprise a nucleic acid with a modification at a 2' hydroxyl group of the ribose moiety. The modification can be a 2'-O-methyl modification or a 2'-O-methoxyethyl (2'-O-MOE) modification. The 2'-O-methyl modification can add a methyl group to the 2' hydroxyl group of the ribose moiety whereas the 2'O-methoxyethyl modification can add a methoxyethyl group to the 2' hydroxyl group of the ribose moiety. Exemplary chemical structures of a 2'-O-methyl modification of an adenosine molecule and 2'O-methoxyethyl modification of an uridine are illustrated below.

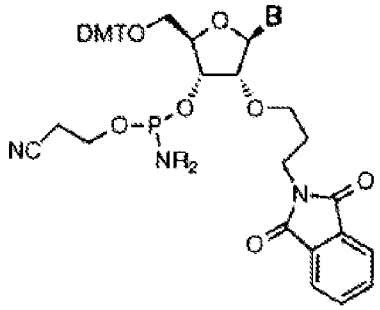


2'-O-methyl-adenosine



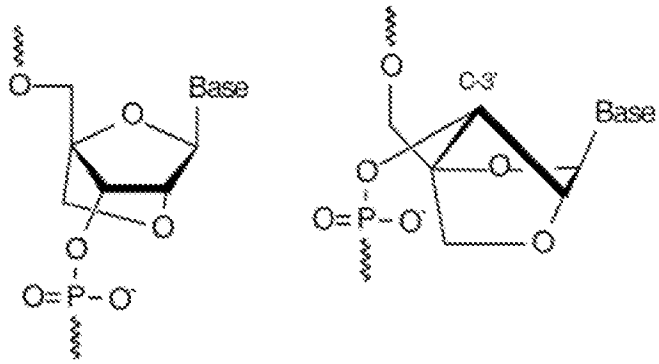
2'-O-methoxyethyl uridine

[00112] An additional modification at the 2' hydroxyl group can include a 2'-O-aminopropyl sugar conformation which can involve an extended amine group comprising a propyl linker that binds the amine group to the 2' oxygen. This modification can neutralize the phosphate derived overall negative charge of the oligonucleotide molecule by introducing one positive charge from the amine group per sugar and can thereby improve cellular uptake properties due to its zwitterionic properties. An exemplary chemical structure of a 2'-O-aminopropyl nucleoside phosphoramidite is illustrated below.



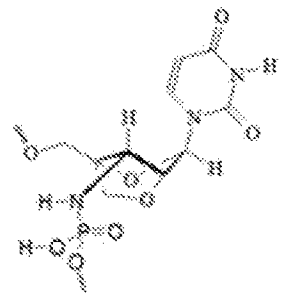
2'-O-aminopropyl nucleoside phosphoramidite

[00113] Another modification at the 2' hydroxyl group can include a locked or bridged ribose conformation (e.g., locked nucleic acid or LNA) where the 4' ribose position can also be involved. In this modification, the oxygen molecule bound at the 2' carbon can be linked to the 4' carbon by a methylene group, thus forming a 2′-*C*,4′-*C*-oxy-methylene-linked bicyclic ribonucleotide monomer. Exemplary representations of the chemical structure of LNA are illustrated below. The representation shown to the left highlights the chemical connectivities of an LNA monomer. The representation shown to the right highlights the locked 3′-endo (^{3}E) conformation of the furanose ring of an LNA monomer.

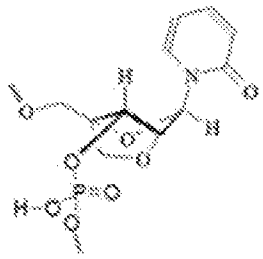


LNA (Locked Nucleic Acids)

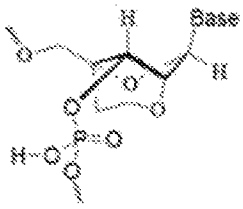
[00114] A further modification at the 2' hydroxyl group may comprise ethylene nucleic acids (ENA) such as for example 2'-4'-ethylene-bridged nucleic acid, which locks the sugar conformation into a C_3'-endo sugar puckering conformation. ENA are part of the bridged nucleic acids class of modified nucleic acids that also comprises LNA. Exemplary chemical structures of the ENA and bridged nucleic acids are illustrated below.



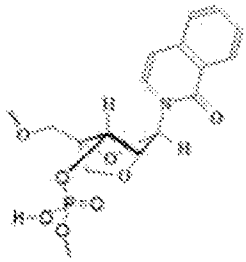
3'-amino-2',4'-BNA



2',4'-BNA-2-pyridone



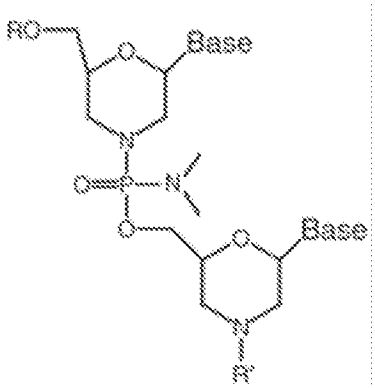
2',4'-ENA



2',4'-BNA-1-isoquinolone

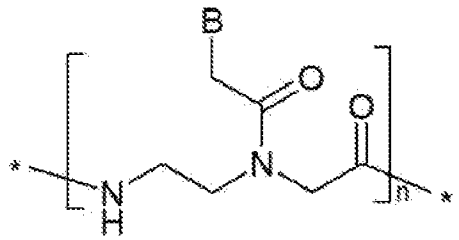
[00115] Still other modifications at the 2' hydroxyl group can include 2'-deoxy, T-deoxy-2'-fluoro, 2'-O-aminopropyl (2'-O-AP), 2'-O-dimethylaminoethyl (2'-O-DMAOE), 2'-O-dimethylaminopropyl (2'-O-DMAP), T-O- dimethylaminoethyloxyethyl (2'-O-DMAEOE), or 2'-O-N-methylacetamido (2'-O-NMA).

[00116] Nucleotide analogues may further comprise Morpholinos, peptide nucleic acids (PNAs), methylphosphonate nucleotides, thiolphosphonate nucleotides, 2'-fluoro N3-P5'-phosphoramidites, 1', 5'- anhydrohexitol nucleic acids (HNAs), or a combination thereof. Morpholino or phosphorodiamidate morpholino oligo (PMO) comprises synthetic molecules whose structure mimics natural nucleic acid structure by deviates from the normal sugar and phosphate structures. Instead, the five member ribose ring can be substituted with a six member morpholino ring containing four carbons, one nitrogen and one oxygen. The ribose monomers can be linked by a phosphordiamidate group instead of a phosphate group. These backbone alterations can remove all positive and negative charges making morpholinos neutral molecules that can cross cellular membranes without the aid of cellular delivery agents such as those used by charged oligonucleotides.



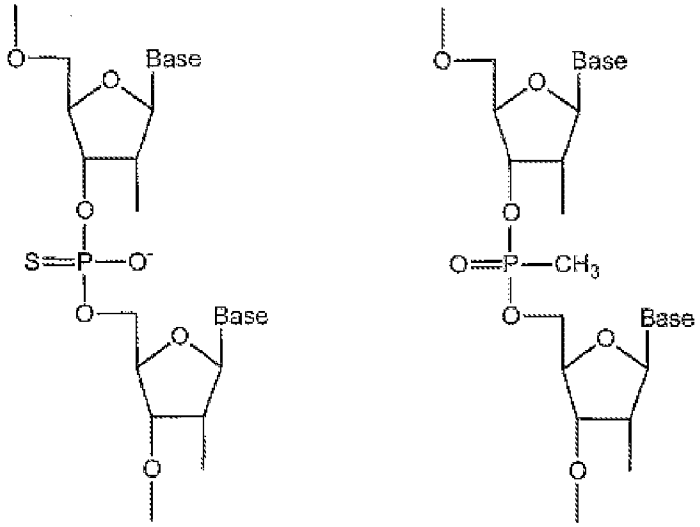
Morpholino

[00117] Peptide nucleic acid (PNA) does not contain sugar ring or phosphate linkage. Instead, the bases can be attached and appropriately spaced by oligoglycine-like molecules, therefore, eliminating a backbone charge.

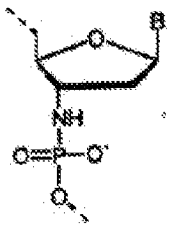


PNA

[00118] Modification of the phosphate backbone may also comprise methyl or thiol modifications such as methylphosphonate nucleotide and. Exemplary thiolphosphonate nucleotide (left) and methylphosphonate nucleotide (right) are illustrated below.

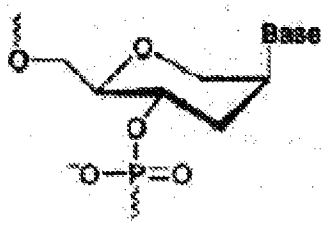


[00119] Furthermore, exemplary 2'-fluoro N3-P5'-phosphoramidites is illustrated as:



N3'-P5' Phosphoroamidate

[00120] And exemplary hexitol nucleic acid (or 1', 5'- anhydrohexitol nucleic acids (HNA)) is illustrated as:



Hexitol Nucleic Acid

[00121] In addition to modification of the ribose moiety, phosphate backbone and the nucleoside, the nucleotide analogues can also be modified by for example at the 3' or the 5' terminus. For example, the 3' terminus can include a 3' cationic group, or by inverting the nucleoside at the 3'-terminus with a 3'-3' linkage. In another alternative, the 3'-terminus can be blocked with an aminoalkyl group, e.g., a 3' C5-aminoalkyl dT. The 5' -terminus can be blocked with an aminoalkyl group, e.g., a 5'-O-alkylamino substituent. Other 5' conjugates can inhibit 5'-3' exonucleolytic cleavage. Other 3' conjugates can inhibit 3'-5' exonucleolytic cleavage.

[00122] In some cases, one or more of the artificial nucleotide analogues described herein are resistant toward nucleases such as for example ribonuclease such as RNase H, deoxyribunuclease such as DNase, or exonuclease such as 5'-3' exonuclease and 3'-5' exonuclease when compared to natural polynucleic acid polymers. In some instances, artificial nucleotide analogues comprising 2'-O-methyl, 2'-O-methoxyethyl (2'-O-MOE), 2'-O-aminopropyl, 2'-deoxy, T-deoxy-2'-fluoro, 2'-O-aminopropyl (2'-O-AP), 2'-O-dimethylaminoethyl (2'-O-DMAOE), 2'-O-dimethylaminopropyl (2'-O-DMAP), T-O- dimethylaminoethyloxyethyl (2'-O-DMAEOE), or 2'-O-N-methylacetamido (2'-O-NMA) modified, LNA, ENA, PNA, HNA, morpholino, methylphosphonate nucleotides, thiolphosphonate nucleotides, 2'-fluoro N3-P5'-phosphoramidites, or combinations thereof are resistant toward nucleases such as for example ribonuclease such as RNase H, deoxyribunuclease such as DNase, or exonuclease such as 5'-3' exonuclease and 3'-5' exonuclease. 2'-O-methyl modified polynucleic acid polymer may be nuclease resistance (e.g., RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). 2'O-methoxyethyl (2'-O-MOE) modified polynucleic acid polymer may be nuclease resistance (e.g., RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). 2'-O-aminopropyl modified polynucleic acid polymer may be nuclease resistance (e.g., RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). 2'-deoxy

modified polynucleic acid polymer may be nuclease resistance (e.g., RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). T-deoxy-2'-fluoro modified polynucleic acid polymer may be nuclease resistance (e.g., RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). 2'-O-aminopropyl (2'-O-AP) modified polynucleic acid polymer may be nuclease resistance (e.g., RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). 2'-O-dimethylaminoethyl (2'-O-DMAOE) modified polynucleic acid polymer may be nuclease resistance (e.g., RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). 2'-O-dimethylaminopropyl (2'-O-DMAP) modified polynucleic acid polymer may be nuclease resistance (e.g., RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). T-O-dimethylaminoethyloxyethyl (2'-O-DMAEOE) modified polynucleic acid polymer may be nuclease resistance (e.g., RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). 2'-O-N-methylacetamido (2'-O-NMA) modified polynucleic acid polymer may be nuclease resistance (e.g., RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). LNA modified polynucleic acid polymer may be nuclease resistance (e.g., RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). ENA modified polynucleic acid polymer may be nuclease resistance (e.g., RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). HNA modified polynucleic acid polymer may be nuclease resistance (e.g., RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). Morpholinos may be nuclease resistance (e.g., RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). PNA can be resistant to nucleases (e.g., RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). Methylphosphonate nucleotides modified polynucleic acid polymer may be nuclease resistance (e.g., RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). Thiolphosphonate nucleotides modified polynucleic acid polymer may be nuclease resistance (e.g., RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance). Polynucleic acid polymer comprising 2'-fluoro N3-P5'-phosphoramidites may be nuclease resistance (e.g., RNase H, DNase, 5'-3' exonuclease or 3'-5' exonuclease resistance).

[00123] In some instances, one or more of the artificial nucleotide analogues described herein have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. The one or more of the artificial nucleotide analogues comprising 2'-O-methyl, 2'-O-methoxyethyl (2'-O-MOE), 2'-O-aminopropyl, 2'-deoxy, T-deoxy-2'-fluoro, 2'-O-aminopropyl (2'-0-AP), 2'-O-dimethylaminoethyl (2'-O-DMAOE), 2'-O-dimethylaminopropyl (2'-O-DMAP), T-O- dimethylaminoethyloxyethyl (2'-O-DMAEOE), or 2'-O-N-methylacetamido (2'-O-NMA) modified, LNA, ENA, PNA, HNA, morpholino, methylphosphonate nucleotides,

thiolphosphonate nucleotides, or 2'-fluoro N3-P5'-phosphoramidites can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. 2'-O-methyl modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. 2'-O-methoxyethyl (2'-O-MOE) modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. 2'-O-aminopropyl modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. 2'-deoxy modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. T-deoxy-2'-fluoro modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. 2'-O-aminopropyl (2'-O-AP) modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. 2'-O-dimethylaminoethyl (2'-O-DMAOE) modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. 2'-O-dimethylaminopropyl (2'-O-DMAP) modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. T-O-dimethylaminoethyloxyethyl (2'-O-DMAEOE) modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. 2'-O-N-methylacetamido (2'-O-NMA) modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. LNA modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. ENA modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. PNA modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. HNA modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. Morpholino modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. Methylphosphonate nucleotides modified polynucleic acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. Thiolphosphonate nucleotides modified polynucleic

acid polymer can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. Polynucleic acid polymer comprising 2'-fluoro N3-P5'-phosphoramidites can have increased binding affinity toward their mRNA target relative to an equivalent natural polynucleic acid polymer. The increased affinity can be illustrated with a lower Kd, a higher melt temperature (Tm), or a combination thereof.

[00124] In additional instances, a polynucleic acid polymer described herein may be modified to increase its stability. In an embodiment where the polynucleic acid polymer is RNA, the polynucleic acid polymer may be modified to increase its stability. The polynucleic acid polymer may be modified by one or more of the modifications described above to increase its stability. The polynucleic acid polymer may be modified at the 2' hydroxyl position, such as by 2'-O-methyl, 2'-O-methoxyethyl (2'-O-MOE), 2'-O-aminopropyl, 2'-deoxy, T-deoxy-2'-fluoro, 2'-O-aminopropyl (2'-0-AP), 2'-O-dimethylaminoethyl (2'-O-DMAOE), 2'-O-dimethylaminopropyl (2'-O-DMAP), T-O- dimethylaminoethyloxyethyl (2'-O-DMAEOE), or 2'-O-N-methylacetamido (2'-O-NMA) modification or by a locked or bridged ribose conformation (e.g., LNA or ENA). The polynucleic acid polymer may be modified by 2'-O-methyl and/or 2'-O-methoxyethyl ribose. The polynucleic acid polymer may also include morpholinos, PNAs, HNA, methylphosphonate nucleotides, thiolphosphonate nucleotides, or 2'-fluoro N3-P5'-phosphoramidites to increase its stability. Suitable modifications to the RNA to increase stability for delivery will be apparent to the skilled person.

[00125] A polynucleic acid polymer described herein can be constructed using chemical synthesis and/or enzymatic ligation reactions using procedures known in the art. For example, a polynucleic acid polymer can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the polynucleic acid polymer and target nucleic acids. Exemplary methods can include those described in: US5,142,047; US5,185,444; WO2009099942; or EP1579015. Additional exemplary methods can include those described in: Griffey et al., "2'-O-aminopropyl ribonucleotides: a zwitterionic modification that enhances the exonuclease resistance and biological activity of antisense oligonucleotides," J. *Med. Chem.* **39**(26):5100-5109 (1997)); Obika, et al. "Synthesis of 2′-O,4′-C-methyleneuridine and -cytidine. Novel bicyclic nucleosides having a fixed C3, -endo sugar puckering". *Tetrahedron Letters* **38** (50): 8735 (1997); Koizumi, M. "ENA oligonucleotides as therapeutics". *Current opinion in molecular therapeutics* **8** (2): 144–149 (2006); and Abramova et al., "Novel oligonucleotide analogues based

on morpholino nucleoside subunits-antisense technologies: new chemical possibilities," Indian Journal of Chemistry 48B:1721-1726 (2009). Alternatively, the polynucleic acid polymer can be produced biologically using an expression vector into which a polynucleic acid polymer has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted polynucleic acid polymer will be of an antisense orientation to a target polynucleic acid polymer of interest).

[00126] A polynucleic acid polymer may be bound to any nucleic acid molecule, such as another antisense molecule, a peptide, or other chemicals to facilitate delivery of the polynucleic acid polymer and/or target the nucleic acid to a specific tissue, cell type, or cell developmental stage. The polynucleic acid polymer may be bound to a protein or RNA. The protein tethered to the polynucleic acid polymer may comprise a splicing factor to enhance, inhibit or modulate splicing and intron removal. RNA tethered to the polynucleic acid polymer may comprise an aptamer or any structure that enhance, inhibit or modulate splicing and intron removal. The polynucleic acid polymer may be isolated nucleic acid.

[00127] A polynucleic acid polymer may be conjugated to, or bound by, a delivery vehicle suitable for delivering the polynucleic acid polymer to cells. The cells may be a specific cell type, or specific developmental stage. The delivery vehicle may be capable of site specific, tissue specific, cell specific or developmental stage-specific delivery. For example, the delivery vehicle may be a cell specific viral particle, or component thereof, alternatively, the delivery vehicle may be a cell specific antibody particle, or component thereof. The polynucleic acid polymer may be targeted for delivery to beta cells in the pancreas. The polynucleic acid polymer may be targeted for delivery to thymic cells. The polynucleic acid polymer may be targeted for delivery to malignant cells. The polynucleic acid polymer may be targeted for delivery to pre-malignant cells (that are known to develop into overt malignant phenotypes within a foreseeable future, such as pre-leukaemias and myelodysplastic syndromes or histopathologically defined precancerous lesions or conditions.

[00128] In one embodiment the polynucleic acid polymer may be bound to a chemical molecule (e.g. non-peptide or nucleic acid based molecule), such as a drug. The drug may be a small molecule (e.g. having a MW of less than 900Da).

[00129] In one embodiment of the invention, the delivery vehicle may comprise a cell penetrating peptide (CPP). For example, the polynucleic acid polymer may be bound or complexed with a CPP. The skilled person will understand that any suitable CPP may be conjugated with the polynucleic acid polymer to aid delivery of the polynucleic acid polymer to and/or into cells. Such CPPs may be any suitable CPP technology described by Boisguérin *et al.* (2015, Advanced Drug

Delivery Reviews doi: 10.1016/j.addr.2015.02.008), which is herein incorporated by reference. Suitable delivery vehicles for conjugation to the polynucleic acid polymer are also described in Lochmann *et al* ((*European Journal of Pharmaceutics and Biopharmaceutics* 58 (2004) 237–251), which is herein incorporated by reference).

[00130] The CPP may be an arginine and/or lysine rich peptide, for example, wherein the majority of residues in the peptide is either lysine or arginine. The CPP may comprise a poly-L-lysine (PLL). Alternatively, the CPP may comprise a poly-arginine. Suitable CPPs may be selected from the group comprising Penetratin; R6-Penetratin; Transportan; oligo-arginines; F-3; B-peptide; B-MSP; Pip peptides, such as Pip1, Pip2a, Pip2b, Pip5e, Pip5f, Pip5h, Pip5j; Pip5k, Pip5l, Pip5m, Pip5n, Pip5o, Pip6a, Pip6b, Pip6c, Pip6d, Pip6e, Pip6f, Pip6g, or Pip6h; peptide of sequence PKKKRKV; Penatratin; Lys_4; SPACE; Tat; Tat-DRBD (dsRNA-binding domain); $(RXR)_4$; $(RFF)_3RXB$; $(KFF)_3K$; R_gF_2; T-cell derived CPP; Pep-3; PEGpep-3; MPG-8; MPG-8-Chol; PepFect6; P5RHH; R_{15}; and Chol-R_9; or functional variants thereof (e.g. see Boisguérin *et al.* (2015, Advanced Drug Delivery Reviews doi: 10.1016/j.addr.2015.02.008).

[00131] In one embodiment, the CPP comprises or consists of a Pip peptide. The Pip peptide may be selected from the group comprising Pip1, Pip2a, Pip2b, Pip5e, Pip5f, Pip5h, Pip5j; Pip5k, Pip5l, Pip5m, Pip5n, Pip5o, Pip6a, Pip6b, Pip6c, Pip6d, Pip6e, Pip6f, Pip6g, and Pip6h.

[00132] In one embodiment of the invention, the delivery vehicle may comprise a peptide-based nanoparticle (PBN), wherein a plurality of CPPs (for example one or more suitable CPPs discussed herein) form a complex with the polynucleic acid polymer through charge interactions. Such nanoparticles may be between about 50nm and 250nm in size. In one embodiment the nanoparticles may be about 70-200nm in size. In another embodiment the nanoparticles may be about 70-100nm in size or 125-200nm in size.

[00133] In one embodiment, the polynucleic acid polymer may be complexed with a delivery vehicle, for example by ionic bonding. Alternatively, the polynucleic acid polymer may be covalently bound to the delivery vehicle. Conjugation/binding methods are described in Lochmann *et al* ((*European Journal of Pharmaceutics and Biopharmaceutics* 58 (2004) 237–251), which is herein incorporated by reference). For example, a conjugation method may comprise introducing a suitable tether containing a reactive group (e.g. –NH2 or –SH2) to the polynucleic acid polymer and to add the delivery vehicle, such as a peptide, post-synthetically as an active intermediate, followed by carrying out the coupling reaction in aqueous medium. An alternative method may comprise carrying out the conjugation in a linear mode on a single solid-phase support.

[00134] The delivery vehicle and polynucleic acid polymer may be thiol and/or maleimide linked, such as thiol-maleimide linked. The conjugation of the polynucleic acid polymer and the delivery vehicle may be by click-chemistry, such as reaction of azido or 2'-*O*-propyargyl functional groups and alkyne groups on the respective molecules to be conjugated. In one embodiment, the delivery vehicle and polynucleic acid polymer may be linked by a thioether bridge. In another embodiment, the delivery vehicle and polynucleic acid polymer may be linked by a disulphide bridge. The skilled person will readily identify suitable linking groups or reactions for conjugation of polynucleic acid polymer and the delivery vehicle, such as a peptide.

[00135] The gene transcript may encode pro-insulin. The gene transcript may be transcribed from *INS* gene. The gene transcripts may be derived from a human haplotype expressing low levels of proinsulin. The intron may comprise *INS* intron 1.

[00136] The gene transcript may be transcribed from a gene or ORF selected from any of the genes or ORFs comprising ABCD4; ABCF3; ACADVL; ALKBH6; AP1G2; APEX1; ARFRP1; ATHL1; ATP1A3; ATP5D; ATP13A1; BAX; BDH2; BRD2; C1orf63; C1orf630; C1orf631; C1orf124; C2orf49; C8orf82; C16orf59; CAPRIN2; CDCA7; CEP164; CEP170; CLCN7; CPNE1; CPSF3L; DCXR; DENND4B; DFFA; DIS3L2; DNAJB12; DPF1; DRG2; DSN1; EML3; EWSR1; EWSR10; FGFR4; FTSJ1; GBAP1; GMPPA; GMPR2; GNPTG; GORASP1; GPATCH4; HGS; HMG20B; IFFO1; ISYNA1; KRI1; LOC148413; LZTR1; MAN2C1; MAP4K2; MCOLN1; MDP1; MIB2; MITD1; MOK; MOV10; MRPL35; MTMR11; MUS81; NAPEPLD; NBEAL2; NDRG4; NDUFB10; NFATC4; NFKBIB; NIT1; NKTR; NPRL2; NSUN5P1; NUDT22; PAN2; PDDC1; PDLIM4; PHF1; PIK3CD; PITPNM1; PPIL2; PPP1R35; PPP4C; PQLC2; PRPF39; PSME2; PTPMT1; QARS; RAD52; RHOT2; RMND5B; RNF123; RPL10A; RPP21; RPS6KB2; RUSC1; SCRN2; SCYL1; SFR1; SGSM3; SIRT7; SLC25A3; SLC25A3; SLC30A7; SLC37A4; STK19; STX10; TCF25; TOMM40; TP53I3; TRIM41; TRPT1; TSTA3; TTC14; TTC140; TUBGCP6; U2AF1L4; UCK1; UNC45A; VAMP1; VAMP10; VARS; VPS28; WDR24; WDR90; WRAP53; YDJC; YIPF3; YIPF3; ZCCHC8; ZCCHC18; ZFAND1; ZNF131; ZNF300; ZNF317; ZNF692; ZNF711; ZNRD1; ZWINT; or combinations thereof.

[00137] The terms "polynucleic acid polymer" and "nucleic acid" can be used interchangeable and can refer to a polynucleic acid polymer that is between about 10 to about 50 nucleotides in length.

[00138] **Diseases**

[00139] The methods and compositions described herein can be used to treat a disease or condition characterized by an impaired production of a protein. The methods and compositions described herein can also be used to treat a disease or condition characterized by a defective splicing. The disease or condition can be a genetic disorder or condition. The genetic disorder or condition can be characterized by an impaired production of a protein. The genetic disorder or condition can also be characterized by a defective splicing. The genetic disorder or condition can be a hereditary disorder, or a nonhereditary defect within one or more locations within the genome. The genetic disorder can be a hereditary disease. The hereditary disease can be characterized by an impaired production of a protein. The hereditary disease can be characterized by a defective splicing. A subject with a hereditary disease can have a genome that can comprise a copy of a gene that comprises an exon that when properly transcribed into fully processed mRNA can encode the full-length functional for of the protein. A subject with a hereditary disease can have a genome that comprises a copy of a gene that can comprise a copy of a gene that comprises a set of exons that when properly transcribed into fully processed mRNA can encode the full-length functional form of the protein. A subject with a hereditary disease can have a genome that can comprise a defective copy of the gene, which can be incapable of producing a full-length functional form of the protein.

[00140] The genetic disorder or condition can be a nonhereditary defect within one or more locations within the genome. The nonhereditary defect can be a point mutation, a deletion, an insertion, or a frame shift. The genetic disorder or condition associated with the nonhereditary defect can be characterized by an impaired production of a protein. The genetic disorder or condition associated with the nonhereditary defect can be characterized by a defective splicing. A subject with a nonhereditary defect can have a genome that can comprise a copy of a gene that comprises an exon that when properly transcribed into fully processed mRNA can encode the full-length functional for of the protein. A subject with a nonhereditary defect can have a genome that comprises a copy of a gene that can comprise a copy of a gene that comprises a set of exons that when properly transcribed into fully processed mRNA can encode the full-length functional form of the protein. A subject with a nonhereditary defect can have a genome that can comprise a defective copy of the gene, which can be incapable of producing a full-length functional form of the protein.

[00141] The genetic disorder or condition can be an autosomal dominant disorder, an autosomal recessive disorder, X-linked dominant disorder, X-linked recessive disorder, Y-linked disorder, mitochondrial disease, or multifactorial or polygenic disorder. Sometimes, a hereditary disease can also be characterized as an autosomal dominant, autosomal recessive, X-linked

dominant, X-linked recessive, Y-linked, mitochondrial, or multifactorial or polygenic hereditary disease. Autosomal dominant disorder, an autosomal recessive disorder, X-linked dominant disorder, X-linked recessive disorder, Y-linked disorder, mitochondrial disease, or multifactorial or polygenic disorder can be characterized by an impaired production of a protein. Autosomal dominant disorder, an autosomal recessive disorder, X-linked dominant disorder, X-linked recessive disorder, Y-linked disorder, mitochondrial disease, or multifactorial or polygenic disorder can be characterized by a defective splicing. A subject with an autosomal dominant disorder, an autosomal recessive disorder, X-linked dominant disorder, X-linked recessive disorder, Y-linked disorder, mitochondrial disease, or multifactorial or polygenic disorder can have a genome that can comprise a copy of a gene that comprises an exon that when properly transcribed into fully processed mRNA can encode the full-length functional for of the protein. A subject with an autosomal dominant disorder, an autosomal recessive disorder, X-linked dominant disorder, X-linked recessive disorder, Y-linked disorder, mitochondrial disease, or multifactorial or polygenic disorder can have a genome that comprises a copy of a gene that can comprise a copy of a gene that comprises a set of exons that when properly transcribed into fully processed mRNA can encode the full-length functional form of the protein. A subject with an autosomal dominant disorder, an autosomal recessive disorder, X-linked dominant disorder, X-linked recessive disorder, Y-linked disorder, mitochondrial disease, or multifactorial or polygenic disorder can have a genome that can comprise a defective copy of the gene, which can be incapable of producing a full-length functional form of the protein.

[00142] Exemplary hereditary disease can include achondroplasia, hereditary hemochromatosis, Down Syndrome, hereditary spherocytosis, Tay-Sachs Disease, Usher syndrome, hereditary fructose intolerance, hemophilia, muscular dystrophy (e.g., Duchenne muscular dystrophy or DMD), polygenic disorders, breast cancer, ovarian cancer, Parkinson's disease, Bardet-Biedl syndrome, Prader-Willi syndrome, diabetes, heart disease, arthritis, motor neuron disease, albinism, Cri-du-Chat syndrome, cystic fibrosis, fragile X syndrome, galactosemia, Huntington's disease, Jackson-Weiss syndrome, Klinefelter syndrome, Krabbe disease, Langer-Giedion syndrome, Lesch-Nyhan syndrome, Marfan syndrome, myotonic dystrophy, Nail-Patella syndrome, neurofibromatosis, Noonan syndrome, triple X syndrome, osteogenesis imperfecta, Patau syndrome, phenylketonuria, porphyria, retinoblastoma, Rett syndrome, sickle cell disease, Turner syndrome, Usher syndrome, Von Hippel-Lindau syndrome, Waardenburg syndrome, Wilson's disease, xeroderma pigmentosum, XXXX syndrome, or YY syndrome.

[00143] A hereditary disease such as for example: achondroplasia, hereditary hemochromatosis, Down Syndrome, hereditary spherocytosis, Tay-Sachs Disease, Usher syndrome, hereditary fructose intolerance, hemophilia, muscular dystrophy (e.g., Duchenne muscular dystrophy or DMD), polygenic disorders, breast cancer, ovarian cancer, Parkinson's disease, Bardet-Biedl syndrome, Prader-Willi syndrome, diabetes, heart disease, arthritis, motor neuron disease, albinism, Cri-du-Chat syndrome, cystic fibrosis, fragile X syndrome, galactosemia, Huntington's disease, Jackson-Weiss syndrome, Klinefelter syndrome, Krabbe disease, Langer-Giedion syndrome, Lesch-Nyhan syndrome, Marfan syndrome, myotonic dystrophy, Nail-Patella syndrome, neurofibromatosis, Noonan syndrome, triple X syndrome, osteogenesis imperfecta, Patau syndrome, phenylketonuria, porphyria, retinoblastoma, Rett syndrome, sickle cell disease, Turner syndrome, Usher syndrome, Von Hippel-Lindau syndrome, Waardenburg syndrome, Wilson's disease, xeroderma pigmentosum, XXXX syndrome, or YY syndrome can be characterized by an impaired production of a protein, or by a defective splicing. A hereditary disease such as for example: achondroplasia, hereditary hemochromatosis, Down Syndrome, hereditary spherocytosis, Tay-Sachs Disease, Usher syndrome, hereditary fructose intolerance, hemophilia, muscular dystrophy (e.g., Duchenne muscular dystrohy or DMD), polygenic disorders, breast cancer, ovarian cancer, Parkinson's disease, Bardet-Biedl syndrome, Prader-Willi syndrome, diabetes, heart disease, arthritis, motor neuron disease, albinism, Cri-du-Chat syndrome, cystic fibrosis, fragile X syndrome, galactosemia, Huntington's disease, Jackson-Weiss syndrome, Klinefelter syndrome, Krabbe disease, Langer-Giedion syndrome, Lesch-Nyhan syndrome, Marfan syndrome, myotonic dystrophy, Nail-Patella syndrome, neurofibromatosis, Noonan syndrome, triple X syndrome, osteogenesis imperfecta, Patau syndrome, phenylketonuria, porphyria, retinoblastoma, Rett syndrome, sickle cell disease, Turner syndrome, Usher syndrome, Von Hippel-Lindau syndrome, Waardenburg syndrome, Wilson's disease, xeroderma pigmentosum, XXXX syndrome, or YY syndrome can comprise a copy of a gene that comprises an exon that when properly transcribed into fully processed mRNA can encode the full-length functional for of the protein, can comprises a copy of a gene that can comprise a copy of a gene that comprises a set of exons that when properly transcribed into fully processed mRNA can encode the full-length functional form of the protein, or can comprise a defective copy of the gene, which can be incapable of producing a full-length functional form of the protein.

[00144] As described above, the genetic disorder or condition can be an autosomal dominant disorder, an autosomal recessive disorder, X-linked dominant disorder, X-linked recessive disorder, Y-linked disorder, mitochondrial disease, or multifactorial or polygenic disorder. The genetic

disorder or condition can be an autosomal dominant disorder, an autosomal recessive disorder, X-linked dominant disorder, X-linked recessive disorder, Y-linked disorder, mitochondrial disease, or multifactorial or polygenic disorder can be a disorder characterized by an impaired production of a protein or by a defective splicing.

[00145] Exemplary autosomal dominant disorder can include Huntington's disease, neurofibromatosis type 1, neurofibromatosis type 2, Marfan syndrome, hereditary nonpolyposis colorectal cancer, hereditary multiple exostoses, Tuberous sclerosis, Von Willebrand disease, or acute intermittent porphyria.

[00146] An autosomal dominant disorder such as Huntington's disease, neurofibromatosis type 1, neurofibromatosis type 2, Marfan syndrome, hereditary nonpolyposis colorectal cancer, hereditary multiple exostoses, Tuberous sclerosis, Von Willebrand disease, or acute intermittent porphyria can be characterized by an impaired production of a protein or by a defective splicing. An autosomal dominant disorder such as Huntington's disease, neurofibromatosis type 1, neurofibromatosis type 2, Marfan syndrome, hereditary nonpolyposis colorectal cancer, hereditary multiple exostoses, Tuberous sclerosis, Von Willebrand disease, or acute intermittent porphyria can comprise a copy of a gene that comprises an exon that when properly transcribed into fully processed mRNA can encode the full-length functional for of the protein, can comprises a copy of a gene that can comprise a copy of a gene that comprises a set of exons that when properly transcribed into fully processed mRNA can encode the full-length functional form of the protein, or can comprise a defective copy of the gene, which can be incapable of producing a full-length functional form of the protein.

[00147] Exemplary autosomal recessive disorder can include albinism, Medium-chain acyl-CoA dehydrogenase deficiency, cystic fibrosis, sickle-cell disease, Tay-Sachs disease, Niemann-Pick disease, spinal muscular atrophy, or Roberts syndrome.

[00148] An autosomal recessive disorder such as: albinism, Medium-chain acyl-CoA dehydrogenase deficiency, cystic fibrosis, sickle-cell disease, Tay-Sachs disease, Niemann-Pick disease, spinal muscular atrophy, or Roberts syndrome can be characterized by an impaired production of a protein or by a defective splicing. An autosomal recessive disorder such as: albinism, Medium-chain acyl-CoA dehydrogenase deficiency, cystic fibrosis, sickle-cell disease, Tay-Sachs disease, Niemann-Pick disease, spinal muscular atrophy, or Roberts syndrome can comprise a copy of a gene that comprises an exon that when properly transcribed into fully processed mRNA can encode the full-length functional for of the protein, can comprises a copy of a gene that can comprise

a copy of a gene that comprises a set of exons that when properly transcribed into fully processed mRNA can encode the full-length functional form of the protein, or can comprise a defective copy of the gene, which can be incapable of producing a full-length functional form of the protein.

[00149] Exemplary X-linked dominant disorder can include X-linked hypophosphatemic rickets, Rett syndrome, incontinentia pigmenti type 2, Aicardi syndrome, or Klinefelter syndrome.

[00150] An X-linked dominant disorder such as: X-linked hypophosphatemic rickets, Rett syndrome, incontinentia pigmenti type 2, Aicardi syndrome, or Klinefelter syndrome can be characterized by an impaired production of a protein or by a defective splicing. An X-linked dominant disorder such as: X-linked hypophosphatemic rickets, Rett syndrome, incontinentia pigmenti type 2, Aicardi syndrome, or Klinefelter syndrome can comprise a copy of a gene that comprises an exon that when properly transcribed into fully processed mRNA can encode the full-length functional for of the protein, can comprises a copy of a gene that can comprise a copy of a gene that comprises a set of exons that when properly transcribed into fully processed mRNA can encode the full-length functional form of the protein, or can comprise a defective copy of the gene, which can be incapable of producing a full-length functional form of the protein.

[00151] Exemplary X-linked recessive disorder can include hemophilia A, Duchenne muscular dystrophy, Lesch-Nyhan syndrome, or Turner syndrome.

[00152] An X-linked recessive disorder such as: hemophilia A, Duchenne muscular dystrophy, Lesch-Nyhan syndrome, or Turner syndrome can be characterized by an impaired production of a protein or by a defective splicing. An X-linked recessive disorder such as: hemophilia A, Duchenne muscular dystrophy, Lesch-Nyhan syndrome, or Turner syndrome can comprise a copy of a gene that comprises an exon that when properly transcribed into fully processed mRNA can encode the full-length functional for of the protein, can comprises a copy of a gene that can comprise a copy of a gene that comprises a set of exons that when properly transcribed into fully processed mRNA can encode the full-length functional form of the protein, or can comprise a defective copy of the gene, which can be incapable of producing a full-length functional form of the protein.

[00153] Exemplary Y-linked disorder can include Swyer syndrome or a form of retinitis pigmentosa.

[00154] A Y-linked disorder such as Swyer syndrome or a form of retinitis pigmentosa can be characterized by an impaired production of a protein or by a defective splicing. A Y-linked disorder such as Swyer syndrome or a form of retinitis pigmentosa can comprise a copy of a gene that

comprises an exon that when properly transcribed into fully processed mRNA can encode the full-length functional for of the protein, can comprises a copy of a gene that can comprise a copy of a gene that comprises a set of exons that when properly transcribed into fully processed mRNA can encode the full-length functional form of the protein, or can comprise a defective copy of the gene, which can be incapable of producing a full-length functional form of the protein.

[00155] Exemplary mitochondrial disease can include Leber's hereditary optic neuropathy.

[00156] Mitochondrial disease such as Leber's hereditary optic neuropathy can be characterized by an impaired production of a protein or by a defective splicing. Mitochondrial disease such as Leber's hereditary optic neuropathy can comprise a copy of a gene that comprises an exon that when properly transcribed into fully processed mRNA can encode the full-length functional for of the protein, can comprises a copy of a gene that can comprise a copy of a gene that comprises a set of exons that when properly transcribed into fully processed mRNA can encode the full-length functional form of the protein, or can comprise a defective copy of the gene, which can be incapable of producing a full-length functional form of the protein.

[00157] Exemplary multifactorial or polygenic disorder can include heart disease, diabetes, autoimmune diseases such as multiple sclerosis, inflammatory bowel disease, or cancer.

[00158] Multifactorial or polygenic disorder such as heart disease, diabetes, autoimmune diseases such as multiple sclerosis, inflammatory bowel disease, or cancer can be characterized by an impaired production of a protein or by a defective splicing. Multifactorial or polygenic disorder such as heart disease, diabetes, autoimmune diseases such as multiple sclerosis, inflammatory bowel disease, or cancer can comprise a copy of a gene that comprises an exon that when properly transcribed into fully processed mRNA can encode the full-length functional for of the protein, can comprises a copy of a gene that can comprise a copy of a gene that comprises a set of exons that when properly transcribed into fully processed mRNA can encode the full-length functional form of the protein, or can comprise a defective copy of the gene, which can be incapable of producing a full-length functional form of the protein.

[00159] In some instances, compositions and methods described herein is used to treat a genetic disorder or condition such as a hereditary disease. Compositions and methods described herein can be used to treat a genetic disorder or condition such as a hereditary disease that is characterized by an impaired production of a protein. Compositions and methods described herein can be used to treat a genetic disorder or condition such as a hereditary disease that is characterized by a defective splicing.

[00160] Compositions and methods described herein can also be used to treat a genetic disorder or condition such as an autosomal dominant disorder, an autosomal recessive disorder, X-linked dominant disorder, X-linked recessive disorder, Y-linked disorder, mitochondrial disease, or multifactorial or polygenic disorder. Compositions and methods described herein can be used to treat an autosomal dominant disorder, an autosomal recessive disorder, X-linked dominant disorder, X-linked recessive disorder, Y-linked disorder, mitochondrial disease, or multifactorial or polygenic disorder, in which the disorder or condition is characterized by an impaired production of a protein. Compositions and methods described herein can also be used to treat an autosomal dominant disorder, an autosomal recessive disorder, X-linked dominant disorder, X-linked recessive disorder, Y-linked disorder, mitochondrial disease, or multifactorial or polygenic disorder, in which the disorder or condition is characterized by a defective splicing.

[00161] In some instances, a disease or condition includes muscular dystrophy, spinal muscular atrophy (SMA), ataxia-telangiectasia, X-linked agammaglobulinaemia, diabetes, or cancer.

[00162] Muscular dystrophy is a group of muscle diseases that can weaken the musculoskeletal system and can hamper locomotion. It can be characterized by the progressive skeletal muscle weakness, defects in muscle proteins, and the death of muscle cells and tissues. One common form of muscular dystrophy can be Duchenne muscular dystrophy (DMD). Additional forms of muscular dystrophy can include Becker, limb-girdle, congenital, facioscapulohumeral, myotonic, oculopharyngeal, distal, and Emery-Dreifuss muscular dystrophy.

[00163] Spinal muscular atrophy (SMA) is an autosomal recessive disorder that is one of the most common genetic causes of childhood mortality. The main characteristic of the disease can be a progressive loss of spinal cord motor neurons, resulting in skeletal muscle denervation with subsequent weakness, atrophy, and paralysis of voluntary muscles. The SMA locus can be mapped to a complex inverted repeat of ~500 kb on Chromosome 5q13 that contains several genes. The main cause of SMA can be homozygous loss of the telomeric copy of the survivor of motor neuron gene (*SMN1*) located within the inverted repeat. A duplicated gene within the centromeric copy of the inverted repeat (*SMN2*) can also be transcribed, but the *SMN2* gene does not completely compensate for loss of *SMN1* function.

[00164] Ataxia-telangiectasia (A-T) or Louis-Bar syndrome is a rare neurodegenerative inherited disease. The term "ataxia" refers to poor coordination and the term "telangiectasia" refers to small dilated blood vessels. Both terms characterize the hallmarks of this disease. AT can impair the cerebellum and additional areas of the brain which can cause impaired movement and

coordination. AT can also weaken the immune system thereby increasing infection and can impair DNA repair, thereby increase the risk of cancer. A-T can be associated with a defect in the gene *ATM*, which is responsible for managing cellular response to multiple form of stress.

[00165] X-linked agammaglobulinaemia, also known as X-linked hypogammaglobulinemia, XLA, Bruton type agammaglobulinemia, Bruton syndrome, or Sex-linked agammaglobulinemia, is an X-linked genetic disorder that can affect the body's ability to fight infection. XLA patients lack mature B cells and as a result, lack the necessary antibodies to combat infection. Bruton's tyrosine kinase (BTK) can be associated with mediating B cell development and maturation and the *BTK* gene can be associated with XLA.

[00166] Diabetes mellitus (DM) (commonly known as diabetes) is a group of metabolic diseases characterized by a high blood sugar level over a prolonged period. Symptoms of diabetes can include weight loss, polyuria or increased urination, polydipsia or increased thirst, and polyphagia or increased hunger. Diabetes can be classified into four categories: type 1, type 2, gestational diabetes, and other specific types of diabetes. Type 1 diabetes can be characterized by a loss of insulin-producing beta cells of the islets of Langerhans in the pancreas, leading to insulin deficiency. Type 2 diabetes can be characterized by insulin resistance, which can also be combined with a reduced insulin secretion. Gestational diabetes can resemble type 2 diabetes, and can involve combination of inadequate insulin secretion and responsiveness. Other specific types of diabetes can include prediabetes, latent autoimmune diabetes of adults (LADA) and congenital diabetes.

[00167] Cancer can be a solid tumor or a hematologic malignancy. A solid tumor can be a sarcoma or a carcinoma. Sarcoma can be a cancer of bone, cartilage, fat muscle, vascular or hematopoietic tissues. Exemplary sarcoma can include alveolar rhabdomyosarcoma, alveolar soft part sarcoma, ameloblastoma, angiosarcoma, chondrosarcoma, chordoma, clear cell sarcoma of soft tissue, dedifferentiated liposarcoma, desmoid, desmoplastic small round cell tumor, embryonal rhabdomyosarcoma, epithelioid fibrosarcoma, epithelioid hemangioendothelioma, epithelioid sarcoma, esthesioneuroblastoma, Ewing sarcoma, extrarenal rhabdoid tumor, extraskeletal myxoid chondrosarcoma, extraskeletal osteosarcoma, fibrosarcoma, giant cell tumor, hemangiopericytoma, infantile fibrosarcoma, inflammatory myofibroblastic tumor, Kaposi sarcoma, leiomyosarcoma of bone, liposarcoma, liposarcoma of bone, malignant fibrous histiocytoma (MFH), malignant fibrous histiocytoma (MFH) of bone, malignant mesenchymoma, malignant peripheral nerve sheath tumor, mesenchymal chondrosarcoma, myxofibrosarcoma, myxoid liposarcoma, myxoinflammatory fibroblastic sarcoma, neoplasms with perivascular epitheioid cell differentiation, osteosarcoma,

parosteal osteosarcoma, neoplasm with perivascular epitheioid cell differentiation, periosteal osteosarcoma, pleomorphic liposarcoma, pleomorphic rhabdomyosarcoma, PNET/extraskeletal Ewing tumor, rhabdomyosarcoma, round cell liposarcoma, small cell osteosarcoma, solitary fibrous tumor, synovial sarcoma, telangiectatic osteosarcoma.

[00168] Carcinoma can be a cancer developed from epithelial cells. Exemplary carcinoma can include adenocarcinoma, squamous cell carcinoma, adenosquamous carcinoma, anaplastic carcinoma, large cell carcinoma, small cell carcinoma, anal cancer, appendix cancer, bile duct cancer (i.e., cholangiocarcinoma), bladder cancer, brain tumor, breast cancer, cervical cancer, colon cancer, cancer of Unknown Primary (CUP), esophageal cancer, eye cancer, fallopian tube cancer, gastroenterological cancer, kidney cancer, liver cancer, lung cancer, medulloblastoma, melanoma, oral cancer, ovarian cancer, pancreatic cancer, parathyroid disease, penile cancer, pituitary tumor, prostate cancer, rectal cancer, skin cancer, stomach cancer, testicular cancer, throat cancer, thyroid cancer, uterine cancer, vaginal cancer, or vulvar cancer.

[00169] Hematologic malignancy is a malignancy of the blood system and can include T-cell based and B-cell based malignancies. Exemplary hematologic malignancy can include myeloid leukaemia, myeloproliferative neoplasias, peripheral T-cell lymphoma not otherwise specified (PTCL-NOS), anaplastic large cell lymphoma, angioimmunoblastic lymphoma, cutaneous T-cell lymphoma, adult T-cell leukemia/lymphoma (ATLL), blastic NK-cell lymphoma, enteropathy-type T-cell lymphoma, hematosplenic gamma-delta T-cell lymphoma, lymphoblastic lymphoma, nasal NK/T-cell lymphomas, treatment-related T-cell lymphomas, chronic lymphocytic leukemia (CLL), small lymphocytic lymphoma (SLL), high risk CLL, non-CLL/SLL lymphoma, prolymphocytic leukemia (PLL), follicular lymphoma (FL), diffuse large B-cell lymphoma (DLBCL), mantle cell lymphoma (MCL), Waldenstrom's macroglobulinemia, multiple myeloma, extranodal marginal zone B cell lymphoma, nodal marginal zone B cell lymphoma, Burkitt's lymphoma, non-Burkitt high grade B cell lymphoma, primary mediastinal B-cell lymphoma (PMBL), immunoblastic large cell lymphoma, precursor B-lymphoblastic lymphoma, B cell prolymphocytic leukemia, lymphoplasmacytic lymphoma, splenic marginal zone lymphoma, plasma cell myeloma, plasmacytoma, mediastinal (thymic) large B cell lymphoma, intravascular large B cell lymphoma, primary effusion lymphoma, or lymphomatoid granulomatosis.

[00170] In some instances, compositions and methods described herein is used to treat muscular dystrophy, spinal muscular atrophy (SMA), ataxia-telangiectasia, X-linked agammaglobulinaemia, diabetes, or cancer. Compositions and methods described herein can be used

to treat muscular dystrophy, spinal muscular atrophy (SMA), ataxia-telangiectasia, X-linked agammaglobulinaemia, diabetes, or cancer in which the disease or disorder is associated with an impaired production of a protein. Compositions and methods described herein can be used to treat muscular dystrophy, spinal muscular atrophy (SMA), ataxia-telangiectasia, X-linked agammaglobulinaemia, diabetes, or cancer in which the disease or disorder is characterized by a defective splicing.

[00171] The disease may be any genetic condition caused by mutations leading to retention of entire introns in mature transcripts. The disease may be diabetes. The disease may be diabetes type I. The disease may be diabetes type II. In another embodiment, the disease may be cancer. The cancer may be myeloid leukaemia or myeloproliferative neoplasias. The cancer may sustain mutations in any of the spliceosomal components that facilitate recognition of 3' splice sites. The polynucleic acid polymer may be used to increase endogenous expression in a subject with residual β-cell activity. The polynucleic acid polymer may be used to increase expression of proinsulin in other diabetes patients, including those who received transplanted β-cells. The polynucleic acid polymer may be used as antisense therapy of malignant tumours containing mutations in genes encoding U2 components (e.g., >20% of myeloid leukaemias).

[00172] In an embodiment where the disease is diabetes, reducing the incidence of intron retention may be during fetal development of the subject. For example a pregnant mother may be administered with the polynucleic acid polymer in order to reduce the intron retention in the fetus.

[00173] The subject may be eukaryote. The subject may be mammalian. The subject may be human. The subject may be a non-human primate. The subject may be a non-primate mammal such as a rat, mouse, ferret, dog, cat, or pig. The subject may be a fetus, such as a human fetus.

[00174] The method may comprise a step of determining if a disease pathology is caused by an intron retention in a gene transcript prior to treatment. The determination may use any suitable assay or genetic analysis available to the skilled person.

[00175] In some instances, detection is done at a nucleic acid level with nucleic acid-based techniques such as in situ hybridization and RT-PCR. Sequencing technologies can include next-generation sequencing technologies such as Helicos True Single Molecule Sequencing (tSMS) (Harris T.D. et al. (2008) Science 320:106-109); 454 sequencing (Roche) (Margulies, M. et al. 2005, Nature, 437, 376-380); SOLiD technology (Applied Biosystems); SOLEXA sequencing (Illumina); single molecule, real-time (SMRT™) technology of Pacific Biosciences; nanopore sequencing (Soni GV and Meller A. (2007) Clin Chem 53: 1996-2001); semiconductor sequencing (Ion Torrent;

Personal Genome Machine); DNA nanoball sequencing; sequencing using technology from Dover Systems (Polonator), and technologies that do not require amplification or otherwise transform native DNA prior to sequencing (e.g., Pacific Biosciences and Helicos), such as nanopore-based strategies (e.g. Oxford Nanopore, Genia Technologies, and Nabsys). Sequencing technologies can also include Sanger sequencing, Maxam-Gilbert sequencing, Shotgun sequencing, bridge PCR, mass spectrometry based sequencing, microfluidic based Sanger sequencing, microscopy-based sequencing, RNAP sequencing, or hybridization based sequencing.

[00176] Sequencing of a gene transcript of interest may also include an amplification step. Exemplary amplification methodologies include, but are not limited to, polymerase chain reaction (PCR), nucleic acid sequence based amplification (NASBA), self-sustained sequence replication (3SR), loop mediated isothermal amplification (LAMP), strand displacement amplification (SDA), whole genome amplification, multiple displacement amplification, strand displacement amplification, helicase dependent amplification, nicking enzyme amplification reaction, recombinant polymerase amplification, reverse transcription PCR, ligation mediated PCR, or methylation specific PCR.

[00177] Additional methods that can be used to obtain a nucleic acid sequence include, e.g., whole-genome RNA expression array, enzyme-linked immunosorbent assay (ELISA), genome sequencing, de novo sequencing, Pacific Biosciences SMRT sequencing, immunohistochemistry (IHC), immunoctyochemistry (ICC), mass spectrometry, tandem mass spectrometry, matrix-assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF MS), in-situ hybridization, fluorescent in-situ hybridization (FISH), chromogenic in-situ hybridization (CISH), silver in situ hybridization (SISH), digital PCR (dPCR), reverse transcription PCR, quantitative PCR (Q-PCR), single marker qPCR, real-time PCR, nCounter Analysis (Nanostring technology), Western blotting, Southern blotting, SDS-PAGE, gel electrophoresis, and Northern blotting.

[00178] In some cases, detection can be done at a protein level, using, for example, immunoprecipitation based assays such as Western blot, or ELISA. Additionally, methods such as electrophoresis and mass spectrometry analysis can also be utilized for detection of a protein of interest.

[00179] According to another aspect of the invention, there is provided a method of modulating intron splicing in a cell, comprising hybridizing a polynucleic acid polymer to a region of pre-mRNA, wherein the region comprises or consists of SEQ ID NO: 46, or optionally, a region

having at least 95% identity to SEQ ID NO: 46. The region may have at least 98% or 99% identity to SEQ ID NO: 46.

[00180] According to another aspect of the invention, there is provided a method of modulating intron splicing in a cell, comprising hybridizing a polynucleic acid polymer to a region of pre-mRNA, wherein the region comprises or consists of SEQ ID NO: 3, or optionally, a region having at least 95% identity to SEQ ID NO: 3. The region may have at least 98% or 99% identity to SEQ ID NO: 3.

[00181] According to another aspect of the invention, there is provided a method of modulating intron splicing in a cell, comprising hybridizing a polynucleic acid polymer to a region of pre-mRNA, wherein the region comprises or consists of a sequence complementary to any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof.

[00182] According to another aspect of the invention, there is provided a method of modulating intron splicing in a cell, comprising hybridizing a polynucleic acid polymer to a region of pre-mRNA, wherein the region comprises or consists of a sequence complementary to a sequence having at least 95% identity to any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof. The region may comprise or consist of a sequence complementary to a sequence having at least 98% identity to any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof. The region may comprise or consist of a sequence complementary to a sequence having at least 99% identity to any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof.

[00183] The cell may be *in vitro*. The cell may be *ex vivo*. The cell may be a eukaryotic cell or a prokaryotic cell. The cell may be a eukaryotic cell. The cell may be a mammalian cell from human; non-human primate; or non-primate mammals such as cat, rat, mouse, dog, ferret, or pigs. The cell may be a human cell such as from an epithelial cell, connective tissue cell, hormone secreting cell, a nerve cell, a skeletal muscle cell, a blood cell, or an immune system cell. The cell may be a tumor cell such as a solid tumor cell or a hematologic malignant cell.

[00184] According to another aspect of the invention, there is provided polynucleic acid polymer which is antisense to at least part of a region of polynucleic acid polymer comprising or consisting of SEQ ID NO: 46, or optionally a region of polynucleic acid polymer comprising or consisting of a sequence having at least 95% sequence identity to SEQ ID NO: 46. The region may have at least 98% or 99% identity to SEQ ID NO: 46.

[00185] According to another aspect of the invention, there is provided polynucleic acid polymer which is antisense to at least part of a region of polynucleic acid polymer comprising or consisting of SEQ ID NO: 3, or optionally a region of polynucleic acid polymer comprising or consisting of a sequence having at least 95% sequence identity to SEQ ID NO: 3. The region may have at least 98% or 99% identity to SEQ ID NO: 3.

[00186] According to another aspect of the invention, there is provided polynucleic acid polymer which is antisense to at least part of a region of polynucleic acid polymer, wherein the region comprises or consists of a sequence complementary to any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof.

[00187] According to another aspect of the invention, there is provided polynucleic acid polymer which is antisense to at least part of a region of polynucleic acid polymer, wherein the region comprises or consists of a sequence complementary to any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof; or optionally a region of polynucleic acid polymer comprising or consisting of a sequence having at least 95% sequence identity to SEQ ID NOs: 47 to 434. The region may have at least 98% or 99% identity to SEQ ID SEQ ID NOs: 47 to 434.

[00188] Reference to being antisense to at least part of a region of polynucleic acid polymer may be understood by the skilled person to mean a region of at least 5 consecutive nucleotides. Reference to being antisense to at least part of a region of polynucleic acid polymer may be understood by the skilled person to mean a region of at least 10 consecutive nucleotides.

[00189] According to another aspect of the invention, there is provided polynucleic acid polymer comprising or consisting of a nucleic acid sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 28; SEQ ID NO: 29; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 34; SEQ ID NO: 35; SEQ ID NO: 37; SEQ ID NO: 38; SEQ ID NO: 40; SEQ ID NO: 41; SEQ ID NO: 43; and SEQ ID NO: 44; or combinations thereof.

[00190] According to another aspect of the invention, there is provided polynucleic acid polymer comprising or consisting of a nucleic acid sequence having at least 99% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 13;

SEQ ID NO: 14; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 28; SEQ ID NO: 29; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 34; SEQ ID NO: 35; SEQ ID NO: 37; SEQ ID NO: 38; SEQ ID NO: 40; SEQ ID NO: 41; SEQ ID NO: 43; and SEQ ID NO: 44; or combinations thereof. According to another aspect of the invention, there is provided polynucleic acid polymer comprising or consisting of a nucleic acid sequence having at least 98% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 28; SEQ ID NO: 29; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 34; SEQ ID NO: 35; SEQ ID NO: 37; SEQ ID NO: 38; SEQ ID NO: 40; SEQ ID NO: 41; SEQ ID NO: 43; and SEQ ID NO: 44; or combinations thereof. According to another aspect of the invention, there is provided polynucleic acid polymer comprising or consisting of a nucleic acid sequence having at least 95% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 28; SEQ ID NO: 29; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 34; SEQ ID NO: 35; SEQ ID NO: 37; SEQ ID NO: 38; SEQ ID NO: 40; SEQ ID NO: 41; SEQ ID NO: 43; and SEQ ID NO: 44; or combinations thereof.

[00191] According to another aspect of the invention, there is provided polynucleic acid polymer comprising or consisting of a nucleic acid sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof. It is understood that uracil nucleotides may be substituted with thymine nucleotides (e.g. the DNA form of the RNA of such sequences).

[00192] According to another aspect of the invention, there is provided polynucleic acid polymer comprising or consisting of a nucleic acid sequence having at least 99% identity to a sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof. According to another aspect of the invention, there is provided polynucleic acid polymer comprising or consisting of a nucleic acid sequence having at least 98% identity to a sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof. According to another aspect of the invention, there is provided polynucleic acid polymer comprising or consisting of a nucleic acid sequence having at least 95% identity to a sequence selected from any

of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof. It is understood that uracil nucleotides may be substituted with thymine nucleotides (e.g. the DNA form of the RNA of such sequences).

[00193] The polynucleic acid polymer may be isolated polynucleic acid polymer. The polynucleic acid polymer may be conjugated to, or bound by, a delivery vehicle suitable for delivering the polynucleic acid polymer to cells. The delivery vehicle may be capable of site specific, tissue specific or cell specific delivery. For example, the delivery vehicle may be a cell specific viral particle, or component thereof, alternatively, the delivery vehicle may be a cell specific antibody particle, or component thereof. The polynucleic acid polymer may be targeted for delivery to beta cells in the pancreas. The polynucleic acid polymer may be targeted for delivery to thymic cells. The polynucleic acid polymer may be targeted for delivery to malignant cells. The polynucleic acid polymer may be modified to increase its stability. In an embodiment where the polynucleic acid polymer is RNA the polynucleic acid polymer may be modified to increase its stability. The polynucleic acid polymer may be modified by 2'-O-methyl and 2'-O-methoxyethyl ribose. Suitable modifications to the RNA to increase stability for delivery will be apparent to the skilled person.

[00194] The polynucleic acid polymer may be part of a plasmid vector. The polynucleic acid polymer may be part of a viral vector. The polynucleic acid polymer may be encoded on a viral vector.

[00195] According to another aspect of the invention, there is provided a vector comprising the polynucleic acid polymer of the invention. The vector may comprise a viral vector. The viral vector may comprise adeno-associated viral vector. The vector may comprise any virus that targets the polynucleic acid polymer to malignant cells or specific cell type.

[00196] According to another aspect of the invention, there is provided a delivery vehicle comprising, or bound to, the polynucleic acid polymer of the invention. The delivery vehicle may comprise a lipid-based nanoparticle; a cationic cell penetrating peptide (CPP); or a linear or branched cationic polymer; or a bioconjugate, such as cholesterol, bile acid, lipid, peptide, polymer, protein, or an aptamer, which is conjugated to the polynucleic acid polymer for intracellular delivery and/or improved stability. The delivery vehicle may be cell or tissue specific or developmental stage specific. The delivery vehicle may comprise an antibody, or part thereof. The antibody may be specific for a cell surface marker on the cell of interest for delivery of the polynucleic acid polymer to the specific cell. For example, the antibody may comprise an anti-GAD antibody for targeted

non-viral polynucleic acid polymer delivery to islet beta cells according to Ji Hoon Jeong et al (Journal of Controlled Release 107 (2005) 562– 570), incorporated herein by reference. For example, the anti-GAD antibody conjugating anti-GAD Fab' fragment to PEI via a PEG linker (PEI-PEG-Fab'). Other specific antibodies may be used in such a conjugation for targeted tissue delivery of the polynucleic acid polymer.

[00197] Pharmaceutical Composition/Formulations, Dosing, and Treatment Regimens

[00198] According to one aspect of the invention, there is provided a therapeutic agent for the treatment of a disease or condition characterized by impaired production of a functional form of a protein which comprises administering to the subject a pharmaceutical composition comprising a therapeutic agent that induces an increase in splicing out of an intron in a partially processed mRNA transcript, wherein the subject has a pool of partially processed mRNA transcripts, which are capable of encoding copies of the full-length functional form of the protein and each of which comprise at least one retained intron that inhibits translation of the partially processed mRNA transcripts; and contacting a target cell of the subject with the therapeutic agent to induce a portion of the pool of the partially processed mRNA transcripts to undergo splicing to remove the at least one retained intron from each of the partially processed mRNA transcripts in the portion, to produce fully processed mRNA transcripts, wherein the fully processed mRNA transcripts are translated to express copies of the full-length functional form of the protein, which treat the disease or condition.

[00199] The therapeutic agent can causes activation of one or more splicing protein complexes in the cell to remove the at least one retained intron from each of the partially processed mRNA transcripts in the portion of the pool of the partially processed mRNA transcripts. The therapeutic agent can inhibit a protein that regulates intron splicing activity. The therapeutic agent can activate a protein that regulates intron splicing activity. The therapeutic agent may interact or bind to a protein that regulates intron splicing activity. The therapeutic agent may interact or bind to target polynucleotide sequence of the partially processed mRNA transcripts. In some embodiments the therapeutic agent can be a polynucleic acid polymer, such as the polynucleic acid polymers described herein. In some embodiments, the therapeutic agent can be a small molecule.

[00200] The small molecule can be a molecule of less than 900 Daltons, and can initiate one or more splicing protein complexes in the cell to remove the at least one retained intron from each of the partially processed mRNA transcripts in the portion of the pool of the partially processed mRNA transcripts. The small molecule can inhibit a protein that regulates intron splicing activity. The small molecule can activate a protein that regulates intron splicing activity. The small molecule may

interact or bind to a protein that regulates intron splicing activity, or may interact or bind to target polynucleotide sequence of the partially processed mRNA transcripts.

[00201] According to another aspect of the invention, there is provided a composition comprising the polynucleic acid polymer of the invention. The composition may be a pharmaceutically acceptable composition. The composition may comprise a pharmaceutically acceptable carrier. The composition may comprise an additional active agent, such as a drug or pro-drug. The composition may comprise combinations of different polynucleic acid polymers, such as SSOs, for therapy.

[00202] The composition may comprise at least one other biologically active molecule in addition to the polynucleic acid polymer. The biologically active molecule may be drug or a pro-drug. The biologically active molecule may comprise nucleic acid or amino acid. The biologically active molecule may comprise a small molecule (e.g. a molecule of <900 Daltons).

[00203] A pharmaceutical composition described herein may comprise a polynucleic acid polymer that hybridizes to a target sequence of a partially processed mRNA transcript which encodes a protein and which comprises a retained intron, wherein the target sequence is in between two G quadruplexes, wherein the polynucleic acid polymer is capable of inducing splicing out of the retained intron from the partially processed mRNA transcript; and a pharmaceutically acceptable excipient and/or a delivery vehicle. The pharmaceutical composition described herein may also comprise a polynucleic acid polymer that hybridizes to a target sequence of a partially processed mRNA transcript which encodes a protein and which comprises a retained intron, wherein the polynucleic acid polymer hybridizes to an intronic splicing regulatory element of the partially processed mRNA transcript, wherein the intronic splicing regulatory element comprises a first CCC motif, and wherein the polynucleic acid polymer is capable of inducing splicing out of the retained intron from the partially processed mRNA transcript; and a pharmaceutically acceptable excipient and/or a delivery vehicle. In addition, the pharmaceutical composition described herein may comprise a polynucleic acid polymer that hybridizes to a target sequence of a partially processed mRNA transcript which encodes a protein and which comprises a retained intron, wherein the polynucleic acid polymer hybridizes to a binding motif of the partially processed mRNA transcript, wherein the binding motif does not form a G quadruplex, and wherein the polynucleic acid polymer is capable of inducing splicing out of the retained intron from the partially processed mRNA transcript; and a pharmaceutically acceptable excipient and/or a delivery vehicle.

[00204] A pharmaceutical composition described herein may further comprise a polynucleic acid polymer that hybridizes to a target sequence of a partially processed mRNA transcript which encodes a protein and which comprises a retained intron, wherein the polynucleic acid polymer hybridizes to a binding motif of the partially processed mRNA transcript, and wherein the binding motif forms a hairpin structure, wherein the polynucleic acid polymer is capable of inducing splicing out of the retained intron from the partially processed mRNA transcript; and a pharmaceutically acceptable excipient and/or a delivery vehicle.

[00205] A pharmaceutical formulations described herein can be administered to a subject by multiple administration routes, including but not limited to, parenteral (e.g., intravenous, subcutaneous, intramuscular), oral, intranasal, buccal, topical, rectal, or transdermal administration routes. In some instances, the pharmaceutical composition describe herein is formulated for parenteral (e.g., intravenous, subcutaneous, intramuscular) administration. In other instances, the pharmaceutical composition describe herein is formulated for oral administration. In still other instances, the pharmaceutical composition describe herein is formulated for intranasal administration.

[00206] A pharmaceutical formulations described herein may include, but are not limited to, aqueous liquid dispersions, self-emulsifying dispersions, solid solutions, liposomal dispersions, aerosols, solid dosage forms, powders, immediate release formulations, controlled release formulations, fast melt formulations, tablets, capsules, pills, delayed release formulations, extended release formulations, pulsatile release formulations, multiparticulate formulations, and mixed immediate and controlled release formulations.

[00207] A pharmaceutical formulations may include a carrier or carrier materials which may include any commonly used excipients in pharmaceutics and should be selected on the basis of compatibility with the composition disclosed herein, and the release profile properties of the desired dosage form. Exemplary carrier materials include, e.g., binders, suspending agents, disintegration agents, filling agents, surfactants, solubilizers, stabilizers, lubricants, wetting agents, diluents, and the like. Pharmaceutically compatible carrier materials may include, but are not limited to, acacia, gelatin, colloidal silicon dioxide, calcium glycerophosphate, calcium lactate, maltodextrin, glycerine, magnesium silicate, polyvinylpyrrollidone (PVP), cholesterol, cholesterol esters, sodium caseinate, soy lecithin, taurocholic acid, phosphotidylcholine, sodium chloride, tricalcium phosphate, dipotassium phosphate, cellulose and cellulose conjugates, sugars sodium stearoyl lactylate, carrageenan, monoglyceride, diglyceride, pregelatinized starch, and the like. See, e.g., *Remington:*

The Science and Practice of Pharmacy, Nineteenth Ed (Easton, Pa.: Mack Publishing Company, 1995); Hoover, John E., *Remington's Pharmaceutical Sciences*, Mack Publishing Co., Easton, Pennsylvania 1975; Liberman, H.A. and Lachman, L., Eds., *Pharmaceutical Dosage Forms*, Marcel Decker, New York, N.Y., 1980; and *Pharmaceutical Dosage Fo*rms *and* Drug Delivery Systems, Seventh Ed. (Lippincott Williams & Wilkins1999).

[00208] A pharmaceutical formulations may include dispersing agents, and/or viscosity modulating agents which may include materials that control the diffusion and homogeneity of a drug through liquid media or a granulation method or blend method. In some embodiments, these agents also facilitate the effectiveness of a coating or eroding matrix. Exemplary diffusion facilitators/dispersing agents include, e.g., hydrophilic polymers, electrolytes, Tween ® 60 or 80, PEG, polyvinylpyrrolidone (PVP; commercially known as Plasdone®), and the carbohydrate-based dispersing agents such as, for example, hydroxypropyl celluloses (e.g., HPC, HPC-SL, and HPC-L), hydroxypropyl methylcelluloses (e.g., HPMC K100, HPMC K4M, HPMC K15M, and HPMC K100M), carboxymethylcellulose sodium, methylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropylmethylcellulose phthalate, hydroxypropylmethylcellulose acetate stearate (HPMCAS), noncrystalline cellulose, magnesium aluminum silicate, triethanolamine, polyvinyl alcohol (PVA), vinyl pyrrolidone/vinyl acetate copolymer (S630), 4-(1,1,3,3-tetramethylbutyl)-phenol polymer with ethylene oxide and formaldehyde (also known as tyloxapol), poloxamers (e.g., Pluronics F68®, F88®, and F108®, which are block copolymers of ethylene oxide and propylene oxide); and poloxamines (e.g., Tetronic 908®, also known as Poloxamine 908®, which is a tetrafunctional block copolymer derived from sequential addition of propylene oxide and ethylene oxide to ethylenediamine (BASF Corporation, Parsippany, N.J.)), polyvinylpyrrolidone K12, polyvinylpyrrolidone K17, polyvinylpyrrolidone K25, or polyvinylpyrrolidone K30, polyvinylpyrrolidone/vinyl acetate copolymer (S-630), polyethylene glycol, e.g., the polyethylene glycol can have a molecular weight of about 300 to about 6000, or about 3350 to about 4000, or about 7000 to about 5400, sodium carboxymethylcellulose, methylcellulose, polysorbate-80, sodium alginate, gums, such as, e.g., gum tragacanth and gum acacia, guar gum, xanthans, including xanthan gum, sugars, cellulosics, such as, e.g., sodium carboxymethylcellulose, methylcellulose, sodium carboxymethylcellulose, polysorbate-80, sodium alginate, polyethoxylated sorbitan monolaurate, polyethoxylated sorbitan monolaurate, povidone, carbomers, polyvinyl alcohol (PVA), alginates, chitosans and combinations thereof. Plasticizers such as cellulose or triethyl cellulose can also be used as dispersing agents. Dispersing agents particularly

useful in liposomal dispersions and self-emulsifying dispersions are dimyristoyl phosphatidyl choline, natural phosphatidyl choline from eggs, natural phosphatidyl glycerol from eggs, cholesterol and isopropyl myristate.

[00209] A pharmaceutical formulations may include pH adjusting agents or buffering agents which may include acids such as acetic, boric, citric, lactic, phosphoric and hydrochloric acids; bases such as sodium hydroxide, sodium phosphate, sodium borate, sodium citrate, sodium acetate, sodium lactate and tris-hydroxymethylaminomethane; and buffers such as citrate/dextrose, sodium bicarbonate and ammonium chloride. Such acids, bases and buffers are included in an amount required to maintain pH of the composition in an acceptable range.

[00210] A pharmaceutical formulation may also include one or more salts in an amount required to bring osmolality of the composition into an acceptable range. Such salts may include those having sodium, potassium or ammonium cations and chloride, citrate, ascorbate, borate, phosphate, bicarbonate, sulfate, thiosulfate or bisulfite anions; suitable salts include sodium chloride, potassium chloride, sodium thiosulfate, sodium bisulfite and ammonium sulfate.

[00211] A pharmaceutical formulations may further include diluent which may also be used to stabilize compounds because they can provide a more stable environment. Salts dissolved in buffered solutions (which also can provide pH control or maintenance) can be utilized as diluents in the art, including, but not limited to a phosphate buffered saline solution. In certain instances, diluents increase bulk of the composition to facilitate compression or create sufficient bulk for homogenous blend for capsule filling. Such compounds can include e.g., lactose, starch, mannitol, sorbitol, dextrose, microcrystalline cellulose such as Avicel®; dibasic calcium phosphate, dicalcium phosphate dihydrate; tricalcium phosphate, calcium phosphate; anhydrous lactose, spray-dried lactose; pregelatinized starch, compressible sugar, such as Di-Pac® (Amstar); mannitol, hydroxypropylmethylcellulose, hydroxypropylmethylcellulose acetate stearate, sucrose-based diluents, confectioner's sugar; monobasic calcium sulfate monohydrate, calcium sulfate dihydrate; calcium lactate trihydrate, dextrates; hydrolyzed cereal solids, amylose; powdered cellulose, calcium carbonate; glycine, kaolin; mannitol, sodium chloride; inositol, bentonite, and the like.

[00212] A pharmaceutical formulations may include disintegration agents or disintegrants to facilitate the breakup or disintegration of a substance. The term "disintegrate" can include both the dissolution and dispersion of the dosage form when contacted with gastrointestinal fluid. Examples of disintegration agents can include a starch, e.g., a natural starch such as corn starch or potato starch, a pregelatinized starch such as National 1551 or Amijel®, or sodium starch glycolate such as

Promogel® or Explotab®, a cellulose such as a wood product, methylcrystalline cellulose, e.g., Avicel®, Avicel® PH101, Avicel® PH102, Avicel® PH105, Elcema® P100, Emcocel®, Vivacel®, Ming Tia®, and Solka-Floc®, methylcellulose, croscarmellose, or a cross-linked cellulose, such as cross-linked sodium carboxymethylcellulose (Ac-Di-Sol®), cross-linked carboxymethylcellulose, or cross-linked croscarmellose, a cross-linked starch such as sodium starch glycolate, a cross-linked polymer such as crospovidone, a cross-linked polyvinylpyrrolidone, alginate such as alginic acid or a salt of alginic acid such as sodium alginate, a clay such as Veegum® HV (magnesium aluminum silicate), a gum such as agar, guar, locust bean, Karaya, pectin, or tragacanth, sodium starch glycolate, bentonite, a natural sponge, a surfactant, a resin such as a cation-exchange resin, citrus pulp, sodium lauryl sulfate, sodium lauryl sulfate in combination starch, and the like.

[00213] A pharmaceutical formulations may include filling agents such as lactose, calcium carbonate, calcium phosphate, dibasic calcium phosphate, calcium sulfate, microcrystalline cellulose, cellulose powder, dextrose, dextrates, dextran, starches, pregelatinized starch, sucrose, xylitol, lactitol, mannitol, sorbitol, sodium chloride, polyethylene glycol, and the like.

[00214] A pharmaceutical formulations may include flavoring agents and/or sweeteners" such as for example, acacia syrup, acesulfame K, alitame, anise, apple, aspartame, banana, Bavarian cream, berry, black currant, butterscotch, calcium citrate, camphor, caramel, cherry, cherry cream, chocolate, cinnamon, bubble gum, citrus, citrus punch, citrus cream, cotton candy, cocoa, cola, cool cherry, cool citrus, cyclamate, cylamate, dextrose, eucalyptus, eugenol, fructose, fruit punch, ginger, glycyrrhetinate, glycyrrhiza (licorice) syrup, grape, grapefruit, honey, isomalt, lemon, lime, lemon cream, monoammonium glyrrhizinate (MagnaSweet®), maltol, mannitol, maple, marshmallow, menthol, mint cream, mixed berry, neohesperidine DC, neotame, orange, pear, peach, peppermint, peppermint cream, Prosweet® Powder, raspberry, root beer, rum, saccharin, safrole, sorbitol, spearmint, spearmint cream, strawberry, strawberry cream, stevia, sucralose, sucrose, sodium saccharin, saccharin, aspartame, acesulfame potassium, mannitol, talin, sylitol, sucralose, sorbitol, Swiss cream, tagatose, tangerine, thaumatin, tutti fruitti, vanilla, walnut, watermelon, wild cherry, wintergreen, xylitol, or any combination of these flavoring ingredients, e.g., anise-menthol, cherry-anise, cinnamon-orange, cherry-cinnamon, chocolate-mint, honey-lemon, lemon-lime, lemon-mint, menthol-eucalyptus, orange-cream, vanilla-mint, and mixtures thereof.

[00215] Lubricants and glidants may also be included in the pharmaceutical formulations described herein which can prevent, reduce or inhibit adhesion or friction of materials. Exemplary lubricants can include, e.g., stearic acid, calcium hydroxide, talc, sodium stearyl fumerate, a

hydrocarbon such as mineral oil, or hydrogenated vegetable oil such as hydrogenated soybean oil (Sterotex®), higher fatty acids and their alkali-metal and alkaline earth metal salts, such as aluminum, calcium, magnesium, zinc, stearic acid, sodium stearates, glycerol, talc, waxes, Stearowet®, boric acid, sodium benzoate, sodium acetate, sodium chloride, leucine, a polyethylene glycol (e.g., PEG-4000) or a methoxypolyethylene glycol such as Carbowax™, sodium oleate, sodium benzoate, glyceryl behenate, polyethylene glycol, magnesium or sodium lauryl sulfate, colloidal silica such as Syloid™, Cab-O-Sil®, a starch such as corn starch, silicone oil, a surfactant, and the like.

[00216] Plasticizers can be compounds used to soften the microencapsulation material or film coatings to make them less brittle. Suitable plasticizers include, e.g., polyethylene glycols such as PEG 300, PEG 400, PEG 600, PEG 1450, PEG 3350, and PEG 800, stearic acid, propylene glycol, oleic acid, triethyl cellulose and triacetin. Plasticizers can also function as dispersing agents or wetting agents.

[00217] Solubilizers can include compounds such as triacetin, triethylcitrate, ethyl oleate, ethyl caprylate, sodium lauryl sulfate, sodium doccusate, vitamin E TPGS, dimethylacetamide, N-methylpyrrolidone, N-hydroxyethylpyrrolidone, polyvinylpyrrolidone, hydroxypropylmethyl cellulose, hydroxypropyl cyclodextrins, ethanol, n-butanol, isopropyl alcohol, cholesterol, bile salts, polyethylene glycol 200-600, glycofurol, transcutol, propylene glycol, and dimethyl isosorbide and the like.

[00218] Stabilizers can include compounds such as any antioxidation agents, buffers, acids, preservatives and the like.

[00219] Suspending agents can include compounds such as polyvinylpyrrolidone, e.g., polyvinylpyrrolidone K12, polyvinylpyrrolidone K17, polyvinylpyrrolidone K25, or polyvinylpyrrolidone K30, vinyl pyrrolidone/vinyl acetate copolymer (S630), polyethylene glycol, e.g., the polyethylene glycol can have a molecular weight of about 300 to about 6000, or about 3350 to about 4000, or about 7000 to about 5400, sodium carboxymethylcellulose, methylcellulose, hydroxypropylmethylcellulose, hydroxymethylcellulose acetate stearate, polysorbate-80, hydroxyethylcellulose, sodium alginate, gums, such as, e.g., gum tragacanth and gum acacia, guar gum, xanthans, including xanthan gum, sugars, cellulosics, such as, e.g., sodium carboxymethylcellulose, methylcellulose, sodium carboxymethylcellulose, hydroxypropylmethylcellulose, hydroxyethylcellulose, polysorbate-80, sodium alginate, polyethoxylated sorbitan monolaurate, polyethoxylated sorbitan monolaurate, povidone and the like.

[00220] Surfactants can include compounds such as sodium lauryl sulfate, sodium docusate, Tween 60 or 80, triacetin, vitamin E TPGS, sorbitan monooleate, polyoxyethylene sorbitan monooleate, polysorbates, polaxomers, bile salts, glyceryl monostearate, copolymers of ethylene oxide and propylene oxide, e.g., Pluronic® (BASF), and the like. Additional surfactants can include polyoxyethylene fatty acid glycerides and vegetable oils, *e.g.*, polyoxyethylene (60) hydrogenated castor oil; and polyoxyethylene alkylethers and alkylphenyl ethers, *e.g.*, octoxynol 10, octoxynol 40. Sometimes, surfactants may be included to enhance physical stability or for other purposes.

[00221] Viscosity enhancing agents can include, e.g., methyl cellulose, xanthan gum, carboxymethyl cellulose, hydroxypropyl cellulose, hydroxypropylmethyl cellulose, hydroxypropylmethyl cellulose acetate stearate, hydroxypropylmethyl cellulose phthalate, carbomer, polyvinyl alcohol, alginates, acacia, chitosans and combinations thereof.

[00222] Wetting agents can include compounds such as oleic acid, glyceryl monostearate, sorbitan monooleate, sorbitan monolaurate, triethanolamine oleate, polyoxyethylene sorbitan monooleate, polyoxyethylene sorbitan monolaurate, sodium docusate, sodium oleate, sodium lauryl sulfate, sodium doccusate, triacetin, Tween 80, vitamin E TPGS, ammonium salts and the like.

[00223] Injectable Formulations

[00224] Formulations suitable for intramuscular, subcutaneous, or intravenous injection may include physiologically acceptable sterile aqueous or non-aqueous solutions, dispersions, suspensions or emulsions, and sterile powders for reconstitution into sterile injectable solutions or dispersions. Examples of suitable aqueous and non-aqueous carriers, diluents, solvents, or vehicles include water, ethanol, polyols (propyleneglycol, polyethylene-glycol, glycerol, cremophor and the like), suitable mixtures thereof, vegetable oils (such as olive oil) and injectable organic esters such as ethyl oleate. Proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersions, and by the use of surfactants. Formulations suitable for subcutaneous injection may also contain additives such as preserving, wetting, emulsifying, and dispensing agents. Prevention of the growth of microorganisms can be ensured by various antibacterial and antifungal agents, such as parabens, chlorobutanol, phenol, sorbic acid, and the like. It may also be desirable to include isotonic agents, such as sugars, sodium chloride, and the like. Prolonged absorption of the injectable pharmaceutical form can be brought about by the use of agents delaying absorption, such as aluminum monostearate and gelatin.

[00225] For intravenous injections, compounds described herein may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hank's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art. For other parenteral injections, appropriate formulations may include aqueous or nonaqueous solutions, preferably with physiologically compatible buffers or excipients. Such excipients are generally known in the art.

[00226] Parenteral injections may involve bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, *e.g.*, in ampoules or in multi-dose containers, with an added preservative. The pharmaceutical composition described herein may be in a form suitable for parenteral injection as a sterile suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, *e.g.*, sterile pyrogen-free water, before use.

[00227] Oral Formulations

[00228] Pharmaceutical preparations for oral use can be obtained by mixing one or more solid excipient with one or more of the compounds described herein, optionally grinding the resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients can include, for example, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methylcellulose, microcrystalline cellulose, hydroxypropylmethylcellulose, sodium carboxymethylcellulose; or others such as: polyvinylpyrrolidone (PVP or povidone) or calcium phosphate. If desired, disintegrating agents may

be added, such as the cross-linked croscarmellose sodium, polyvinylpyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate.

[00229] Dragee cores can be provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

[00230] Solid dosage forms may be in the form of a tablet, (including a suspension tablet, a fast-melt tablet, a bite-disintegration tablet, a rapid-disintegration tablet, an effervescent tablet, or a caplet), a pill, a powder (including a sterile packaged powder, a dispensable powder, or an effervescent powder) a capsule (including both soft or hard capsules, e.g., capsules made from animal-derived gelatin or plant-derived HPMC, or "sprinkle capsules"), solid dispersion, solid solution, bioerodible dosage form, controlled release formulations, pulsatile release dosage forms, multiparticulate dosage forms, pellets, granules, or an aerosol. In other instances, the pharmaceutical formulation is in the form of a powder. In still other instances, the pharmaceutical formulation is in the form of a tablet, including but not limited to, a fast-melt tablet. Additionally, pharmaceutical formulations described herein may be administered as a single capsule or in multiple capsule dosage form. In some cases, the pharmaceutical formulation is administered in two, or three, or four, capsules or tablets.

[00231] The pharmaceutical solid dosage forms can include a composition described herein and one or more pharmaceutically acceptable additives such as a compatible carrier, binder, filling agent, suspending agent, flavoring agent, sweetening agent, disintegrating agent, dispersing agent, surfactant, lubricant, colorant, diluent, solubilizer, moistening agent, plasticizer, stabilizer, penetration enhancer, wetting agent, anti-foaming agent, antioxidant, preservative, or one or more combination thereof. In still other aspects, using standard coating procedures, such as those described in *Remington's Pharmaceutical Sciences*, 20th Edition (2000).

[00232] Suitable carriers for use in the solid dosage forms can include, but are not limited to, acacia, gelatin, colloidal silicon dioxide, calcium glycerophosphate, calcium lactate, maltodextrin, glycerine, magnesium silicate, sodium caseinate, soy lecithin, sodium chloride, tricalcium phosphate, dipotassium phosphate, sodium stearoyl lactylate, carrageenan, monoglyceride, diglyceride, pregelatinized starch, hydroxypropylmethylcellulose, hydroxypropylmethylcellulose acetate stearate, sucrose, microcrystalline cellulose, lactose, mannitol and the like.

[00233] Suitable filling agents for use in the solid dosage forms can include, but are not limited to, lactose, calcium carbonate, calcium phosphate, dibasic calcium phosphate, calcium sulfate, microcrystalline cellulose, cellulose powder, dextrose, dextrates, dextran, starches, pregelatinized starch, hydroxypropylmethycellulose (HPMC), hydroxypropylmethycellulose phthalate, hydroxypropylmethylcellulose acetate stearate (HPMCAS), sucrose, xylitol, lactitol, mannitol, sorbitol, sodium chloride, polyethylene glycol, and the like.

[00234] Binders impart cohesiveness to solid oral dosage form formulations: for powder filled capsule formulation, they aid in plug formation that can be filled into soft or hard shell capsules and for tablet formulation, they ensure the tablet remaining intact after compression and help assure blend uniformity prior to a compression or fill step. Materials suitable for use as binders in the solid dosage forms described herein include, but are not limited to, carboxymethylcellulose, methylcellulose (e.g., Methocel®), hydroxypropylmethylcellulose (e.g. Hypromellose USP Pharmacoat-603, hydroxypropylmethylcellulose acetate stearate (Aqoate HS-LF and HS), hydroxyethylcellulose, hydroxypropylcellulose (e.g., Klucel®), ethylcellulose (e.g., Ethocel®), and microcrystalline cellulose (e.g., Avicel®), microcrystalline dextrose, amylose, magnesium aluminum silicate, polysaccharide acids, bentonites, gelatin, polyvinylpyrrolidone/vinyl acetate copolymer, crospovidone, povidone, starch, pregelatinized starch, tragacanth, dextrin, a sugar, such as sucrose (e.g., Dipac®), glucose, dextrose, molasses, mannitol, sorbitol, xylitol (e.g., Xylitab®), lactose, a natural or synthetic gum such as acacia, tragacanth, ghatti gum, mucilage of isapol husks, starch, polyvinylpyrrolidone (e.g., Povidone® CL, Kollidon® CL, Polyplasdone® XL-10, and Povidone® K-12), larch arabogalactan, Veegum®, polyethylene glycol, waxes, sodium alginate, and the like.

[00235] Suitable lubricants or glidants for use in the solid dosage forms can include, but are not limited to, stearic acid, calcium hydroxide, talc, corn starch, sodium stearyl fumerate, alkali-metal and alkaline earth metal salts, such as aluminum, calcium, magnesium, zinc, stearic acid, sodium stearates, magnesium stearate, zinc stearate, waxes, Stearowet®, boric acid, sodium benzoate, sodium acetate, sodium chloride, leucine, a polyethylene glycol or a methoxypolyethylene glycol such as Carbowax™, PEG 4000, PEG 5000, PEG 6000, propylene glycol, sodium oleate, glyceryl behenate, glyceryl palmitostearate, glyceryl benzoate, magnesium or sodium lauryl sulfate, and the like.

[00236] Suitable diluents for use in the solid dosage forms can include, but are not limited to, sugars (including lactose, sucrose, and dextrose), polysaccharides (including dextrates and maltodextrin), polyols (including mannitol, xylitol, and sorbitol), cyclodextrins and the like.

[00237] Suitable wetting agents for use in the solid dosage forms can include, for example, oleic acid, glyceryl monostearate, sorbitan monooleate, sorbitan monolaurate, triethanolamine oleate, polyoxyethylene sorbitan monooleate, polyoxyethylene sorbitan monolaurate, quaternary ammonium compounds (e.g., Polyquat 10®), sodium oleate, sodium lauryl sulfate, magnesium stearate, sodium docusate, triacetin, vitamin E TPGS and the like.

[00238] Suitable surfactants for use in the solid dosage forms can include, for example, sodium lauryl sulfate, sorbitan monooleate, polyoxyethylene sorbitan monooleate, polysorbates, polaxomers, bile salts, glyceryl monostearate, copolymers of ethylene oxide and propylene oxide, e.g., Pluronic® (BASF), and the like.

[00239] Suitable suspending agents for use in the solid dosage forms can include, but are not limited to, polyvinylpyrrolidone, e.g., polyvinylpyrrolidone K12, polyvinylpyrrolidone K17, polyvinylpyrrolidone K25, or polyvinylpyrrolidone K30, polyethylene glycol, e.g., the polyethylene glycol can have a molecular weight of about 300 to about 6000, or about 3350 to about 4000, or about 7000 to about 5400, vinyl pyrrolidone/vinyl acetate copolymer (S630), sodium carboxymethylcellulose, methylcellulose, hydroxy-propylmethylcellulose, polysorbate-80, hydroxyethylcellulose, sodium alginate, gums, such as, e.g., gum tragacanth and gum acacia, guar gum, xanthans, including xanthan gum, sugars, cellulosics, such as, e.g., sodium carboxymethylcellulose, methylcellulose, sodium carboxymethylcellulose, hydroxypropylmethylcellulose, hydroxyethylcellulose, polysorbate-80, sodium alginate, polyethoxylated sorbitan monolaurate, polyethoxylated sorbitan monolaurate, povidone and the like.

[00240] Suitable antioxidants for use in the solid dosage forms include, for example, e.g., butylated hydroxytoluene (BHT), sodium ascorbate, and tocopherol.

[00241] Liquid formulation dosage forms for oral administration can be aqueous suspensions selected from the group including, but not limited to, pharmaceutically acceptable aqueous oral dispersions, emulsions, solutions, elixirs, gels, and syrups. See, e.g., Singh et al., *Encyclopedia of Pharmaceutical Technology*, 2nd Ed., pp. 754-757 (2002). In addition the liquid dosage forms may include additives, such as: (a) disintegrating agents; (b) dispersing agents; (c) wetting agents; (d) at least one preservative, (e) viscosity enhancing agents, (f) at least one sweetening agent, and (g) at least one flavoring agent. In some embodiments, the aqueous dispersions can further include a crystalline inhibitor.

[00242] The aqueous suspensions and dispersions described herein can remain in a homogenous state, as defined in The USP Pharmacists' Pharmacopeia (2005 edition, chapter 905),

for at least 4 hours. The homogeneity should be determined by a sampling method consistent with regard to determining homogeneity of the entire composition. In one embodiment, an aqueous suspension can be re-suspended into a homogenous suspension by physical agitation lasting less than 1 minute. In another aspect, an aqueous suspension can be re-suspended into a homogenous suspension by physical agitation lasting less than 45 seconds. In yet another aspect, an aqueous suspension can be re-suspended into a homogenous suspension by physical agitation lasting less than 30 seconds. In still another embodiment, no agitation is necessary to maintain a homogeneous aqueous dispersion.

[00243] In another aspect, dosage forms may include microencapsulated formulations. In some embodiments, one or more other compatible materials are present in the microencapsulation material. Exemplary materials include, but are not limited to, pH modifiers, erosion facilitators, anti-foaming agents, antioxidants, flavoring agents, and carrier materials such as binders, suspending agents, disintegration agents, filling agents, surfactants, solubilizers, stabilizers, lubricants, wetting agents, and diluents.

[00244] Exemplary microencapsulation materials useful for delaying the release of the formulations including compounds described herein, include, but are not limited to, hydroxypropyl cellulose ethers (HPC) such as Klucel® or Nisso HPC, low-substituted hydroxypropyl cellulose ethers (L-HPC), hydroxypropyl methyl cellulose ethers (HPMC) such as Seppifilm-LC, Pharmacoat®, Metolose SR, Methocel®-E, Opadry YS, PrimaFlo, Benecel MP824, and Benecel MP843, methylcellulose polymers such as Methocel®-A, hydroxypropylmethylcellulose acetate stearate Aqoat (HF-LS, HF-LG,HF-MS) and Metolose®, Ethylcelluloses (EC) and mixtures thereof such as E461, Ethocel®, Aqualon®-EC, Surelease®, Polyvinyl alcohol (PVA) such as Opadry AMB, hydroxyethylcelluloses such as Natrosol®, carboxymethylcelluloses and salts of carboxymethylcelluloses (CMC) such as Aqualon®-CMC, polyvinyl alcohol and polyethylene glycol co-polymers such as Kollicoat IR®, monoglycerides (Myverol), triglycerides (KLX), polyethylene glycols, modified food starch, acrylic polymers and mixtures of acrylic polymers with cellulose ethers such as Eudragit® EPO, Eudragit® L30D-55, Eudragit® FS 30D Eudragit® L100-55, Eudragit® L100, Eudragit® S100, Eudragit® RD100, Eudragit® E100, Eudragit® L12.5, Eudragit® S12.5, Eudragit® NE30D, and Eudragit® NE 40D, cellulose acetate phthalate, sepifilms such as mixtures of HPMC and stearic acid, cyclodextrins, and mixtures of these materials.

[00245] Plasticizers may include polyethylene glycols, e.g., PEG 300, PEG 400, PEG 600, PEG 1450, PEG 3350, and PEG 800, stearic acid, propylene glycol, oleic acid, and triacetin are

incorporated into the microencapsulation material. In other embodiments, the microencapsulating material useful for delaying the release of the pharmaceutical compositions is from the USP or the National Formulary (NF). In yet other embodiments, the microencapsulation material is Klucel. In still other embodiments, the microencapsulation material is methocel.

[00246] Microencapsulated compositions may be formulated by methods known by one of ordinary skill in the art. Such known methods include, e.g., spray drying processes, spinning disk-solvent processes, hot melt processes, spray chilling methods, fluidized bed, electrostatic deposition, centrifugal extrusion, rotational suspension separation, polymerization at liquid-gas or solid-gas interface, pressure extrusion, or spraying solvent extraction bath. In addition to these, several chemical techniques, e.g., complex coacervation, solvent evaporation, polymer-polymer incompatibility, interfacial polymerization in liquid media, in situ polymerization, in-liquid drying, and desolvation in liquid media could also be used. Furthermore, other methods such as roller compaction, extrusion/spheronization, coacervation, or nanoparticle coating may also be used.

[00247] Intranasal Formulations

[00248] Intranasal formulations are known in the art and are described in, for example, U.S. Pat. Nos. 4,476,116 and 6,391,452. Formulations that include the compositions described herein, which are prepared according to the above described and other techniques well-known in the art are prepared as solutions in saline, employing benzyl alcohol or other suitable preservatives, fluorocarbons, and/or other solubilizing or dispersing agents known in the art. See, for example, Ansel, H. C. et al., Pharmaceutical Dosage Forms and Drug Delivery Systems, Sixth Ed. (1995). Preferably these compositions and formulations are prepared with suitable nontoxic pharmaceutically acceptable ingredients. These ingredients are known to those skilled in the preparation of nasal dosage forms and some of these can be found in Remington: The Science and Practice of Pharmacy, 21st edition, 2005, a standard reference in the field. The choice of suitable carriers is highly dependent upon the exact nature of the nasal dosage form desired, e.g., solutions, suspensions, ointments, or gels. Nasal dosage forms generally contain large amounts of water in addition to the active ingredient. Minor amounts of other ingredients such as pH adjusters, emulsifiers or dispersing agents, preservatives, surfactants, gelling agents, or buffering and other stabilizing and solubilizing agents may also be present. The nasal dosage form should be isotonic with nasal secretions.

[00249] For administration by inhalation described herein may be in a form as an aerosol, a mist or a powder. Pharmaceutical compositions described herein are conveniently delivered in the

form of an aerosol spray presentation from pressurized packs or a nebulizer, with the use of a suitable propellant, *e.g.*, dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol, the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, such as, by way of example only, gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound described herein and a suitable powder base such as lactose or starch.

[00250] Therapeutic Regimens

[00251] The compositions may be administered for therapeutic applications or as a maintenance therapy, for example for a patient in remission. The composition may be administered once per day, twice per day, three times per day or more. The composition may be administered daily, every day, every alternate day, five days a week, once a week, every other week, two weeks per month, three weeks per month, once a month, twice a month, three times per month, or more. The composition may be administered for at least 1 month, 2 months, 3 months, 4 months, 5 months, 6 months, 7 months, 8 months, 9 months, 10 months, 11 months, 12 months, 18 months, 2 years, 3 years, or more.

[00252] In the case wherein the patient's status does improve, upon the doctor's discretion the administration of the compounds may be given continuously; alternatively, the dose of drug being administered may be temporarily reduced or temporarily suspended for a certain length of time (i.e., a "drug holiday"). The length of the drug holiday can vary between 2 days and 1 year, including by way of example only, 2 days, 3 days, 4 days, 5 days, 6 days, 7 days, 10 days, 12 days, 15 days, 20 days, 28 days, 35 days, 50 days, 70 days, 100 days, 120 days, 150 days, 180 days, 200 days, 250 days, 280 days, 300 days, 320 days, 350 days, or 365 days. The dose reduction during a drug holiday may be from 10%-100%, including, by way of example only, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100%.

[00253] Once improvement of the patient's conditions has occurred, a maintenance dose is administered if necessary. Subsequently, the dosage or the frequency of administration, or both, can be reduced, as a function of the symptoms, to a level at which the improved disease, disorder or condition is retained. Patients can, however, require intermittent treatment on a long-term basis upon any recurrence of symptoms.

[00254] The amount of a given agent that will correspond to such an amount will vary depending upon factors such as the particular compound, the severity of the disease, the identity (e.g., weight) of the subject or host in need of treatment, but can nevertheless be routinely

determined in a manner known in the art according to the particular circumstances surrounding the case, including, e.g., the specific agent being administered, the route of administration, and the subject or host being treated. The desired dose may conveniently be presented in a single dose or as divided doses administered simultaneously (or over a short period of time) or at appropriate intervals, for example as two, three, four or more sub-doses per day.

[00255] The pharmaceutical composition described herein may be in unit dosage forms suitable for single administration of precise dosages. In unit dosage form, the formulation is divided into unit doses containing appropriate quantities of one or more compound. The unit dosage may be in the form of a package containing discrete quantities of the formulation. Non-limiting examples are packaged tablets or capsules, and powders in vials or ampoules. Aqueous suspension compositions can be packaged in single-dose non-reclosable containers. Alternatively, multiple-dose reclosable containers can be used, in which case it is typical to include a preservative in the composition. By way of example only, formulations for parenteral injection may be presented in unit dosage form, which include, but are not limited to ampoules, or in multi-dose containers, with an added preservative.

[00256] The foregoing ranges are merely suggestive, as the number of variables in regard to an individual treatment regime is large, and considerable excursions from these recommended values are not uncommon. Such dosages may be altered depending on a number of variables, not limited to the activity of the compound used, the disease or condition to be treated, the mode of administration, the requirements of the individual subject, the severity of the disease or condition being treated, and the judgment of the practitioner.

[00257] Toxicity and therapeutic efficacy of such therapeutic regimens can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, including, but not limited to, the determination of the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between the toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD50 and ED50. Compounds exhibiting high therapeutic indices are preferred. The data obtained from cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with minimal toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized.

[00258] According to another aspect of the invention, there is provided a method of selecting a subject for treatment, comprising determining if the subject has a disease induced by defective protein expression caused by the intron retention in gene transcripts, wherein the subject is selected for treatment upon positive confirmation; and optionally treating the subject.

[00259] The treatment may comprise correction of intron retention in the gene transcripts. The treatment may comprise hybridizing a polynucleic acid polymer according to the invention to the gene transcript in order to induce intron removal from the gene transcript.

[00260] According to another aspect of the invention, there is provided the use of antisense polynucleic acid polymer to normalize gene expression by correction of retention of introns in cancer cells.

[00261] According to another aspect of the invention, there is provided the polynucleic acid polymer according to the invention, the composition according to the invention, the vector according to the invention, or the delivery vehicle according to the invention, for use in the treatment or prevention of a disease.

[00262] According to another aspect of the invention, there is provided the polynucleic acid polymer according to the invention, the composition according to the invention, the vector according to the invention, or the delivery vehicle according to the invention, for use in the manufacture of a medicament for the treatment or prevention of a disease.

[00263] The disease may be diabetes or cancer.

[00264] Where reference is made to a polynucleic acid polymer sequence, the skilled person will understand that one or more substitutions may be tolerated, optionally two substitutions may be tolerated in the sequence, such that it maintains the ability to hybridize to the target sequence, or where the substitution is in a target sequence, the ability to be recognized as the target sequence. References to sequence identity may be determined by BLAST sequence alignment (www.ncbi.nlm.nih.gov/BLAST/) using standard/default parameters. For example, the sequence may have 99% identity and still function according to the invention. In other embodiments, the sequence may have 98% identity and still function according to the invention. In another embodiment, the sequence may have 95% identity and still function according to the invention.

[00265] Where reference is made to reducing or correcting intron retention, the reduction may be complete, e.g. 100%, or may be partial. The reduction may be clinically significant. The reduction/correction may be relative to the level of intron retention in the subject without treatment, or relative to the amount of intron retention in a population of similar subjects. The

reduction/correction may be at least 10% less intron retentions relative to the average subject, or the subject prior to treatment. The reduction may be at least 20% less intron retentions relative to an average subject, or the subject prior to treatment. The reduction may be at least 40% less intron retentions relative to an average subject, or the subject prior to treatment. The reduction may be at least 50% less intron retentions relative to an average subject, or the subject prior to treatment. The reduction may be at least 60% less intron retentions relative to an average subject, or the subject prior to treatment. The reduction may be at least 80% less intron retentions relative to an average subject, or the subject prior to treatment. The reduction may be at least 90% less intron retentions relative to an average subject, or the subject prior to treatment.

[00266] **Kits/Articles of Manufacture**

[00267] Kits and articles of manufacture are provided herein for use with one or more methods described herein. The kits can contain one or more of the polynucleic acid polymers described herein, such as the polynucleic acid polymers identified as SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 43, or SEQ ID NO: 44. The kits can also contain one or more of the polynucleic acid polymers that are antisense to polynucleic acid polymers described herein, such as for example SEQ ID NO: 3, SEQ ID NO: 6, SEQ ID NO: 9, SEQ ID NO: 12, SEQ ID NO: 15, SEQ ID NO: 18, SEQ ID NO: 21, SEQ ID NO: 24, SEQ ID NO: 27, SEQ ID NO: 30, SEQ ID NO: 33, SEQ ID NO: 36, SEQ ID NO: 39, SEQ ID NO: 42, SEQ ID NO: 45, SEQ ID NO: 46, or SEQ ID NOs: 47-434. The kits can further contain reagents, and buffers necessary for the makeup and delivery of the polynucleic acid polymers.

[00268] The kits can also include a carrier, package, or container that is compartmentalized to receive one or more containers such as vials, tubes, and the like, each of the container(s) comprising one of the separate elements, such as the polynucleic acid polymers and reagents, to be used in a method described herein. Suitable containers include, for example, bottles, vials, syringes, and test tubes. The containers can be formed from a variety of materials such as glass or plastic.

[00269] The articles of manufacture provided herein contain packaging materials. Examples of pharmaceutical packaging materials include, but are not limited to, bottles, tubes, bags,

containers, bottles, and any packaging material suitable for a selected formulation and intended mode of administration and treatment.

[00270] A kit typically includes labels listing contents and/or instructions for use, and package inserts with instructions for use. A set of instructions will also typically be included.

[00271] The skilled person will understand that optional features of one embodiment or aspect of the invention may be applicable, where appropriate, to other embodiments or aspects of the invention.

[00272] Embodiments of the invention will now be described in more detail, by way of example only, with reference to the accompanying drawings.

EXAMPLES

[00273] These examples are provided for illustrative purposes only and not to limit the scope of the claims provided herein.

[00274] **Example 1**

[00275] Most eukaryotic genes contain intervening sequences or introns that must be accurately removed from primary transcripts to create functional mRNAs capable of encoding proteins (1). This process modifies mRNP composition in a highly dynamic manner, employing interdependent interactions of five small nuclear RNAs and a large number of proteins with conserved but degenerate sequences in the pre-mRNA (2). Intron splicing generally promotes mRNA accumulation and protein expression across species (3–5). This process can be altered by intronic mutations or variants that may also impair coupled gene expression pathways, including transcription, mRNA export and translation. This is best exemplified by introns in the 5'untranslated region (5'UTR) where natural variants or mutations modifying intron retention alter the relative abundance of transcripts with upstream open reading frames (uORFs) or other regulatory motifs and dramatically influence translation (6,7). However, successful sequence-specific strategies to normalize gene expression in such situations have not been developed.

[00276] Splice-switching oligonucleotides (SSOs) are antisense reagents that modulate intron splicing by binding splice-site recognition or regulatory sequences and competing with *cis*- and *trans*-acting factors for their targets (8). They have been shown to restore aberrant splicing, modify the relative expression of existing mRNAs or produce novel splice variants that are not normally expressed (8). Improved stability of targeted SSO-RNA duplexes by a number of SSO modifications, such as 2'-*O*-methyl and 2'-*O*-methoxyethyl ribose, facilitated studies exploring their

therapeutic potential for a growing number of human disease genes, including DMD in muscular dystrophy (9,10), SMN2 in spinal muscular atrophy (11), ATM in ataxia-telangiectasia (12) and BTK in X-linked agammaglobulinemia (13). Although such approaches are close to achieving their clinical potential for a restricted number of diseases (8), >300 Mendelian disorders resulting from mutation-induced aberrant splicing (14) and a growing number of complex traits may be amenable to SSO-mediated correction of gene expression.

[00277] Etiology of type 1 diabetes has a strong genetic component conferred by human leukocyte antigens (HLA) and a number of modifying non-HLA loci (15). The strongest modifier was identified in the proinsulin gene (*INS*) region on chromosome 11 (termed IDDM2) (15). Further mapping of this area suggested that *INS* is the most likely IDDM2 target (16), consistent with a critical role of this autoantigen in pathogenesis (17). Genetic risk to this disease at IDDM2 has been attributed to differential steady-state RNA levels from predisposing and protective *INS* haplotypes, potentially involving a minisatellite DNA sequence upstream of this gene (18,19). However, systematic examination of naturally occurring *INS* polymorphisms revealed haplotype-specific proinsulin expression levels in reporter constructs devoid of the minisatellite sequence, resulting from two variants in intron 1 (7), termed IVS1+5ins4 (also known as *rs3842740* or INS-69) and IVS1–6A/T (*rs689*, INS-27 or *Hph*I+/–) (16,20). The former variant activates a cryptic 5' splice site of intron 1 whereas adenine (A) at the latter variant, which resides 6 nucleotides upstream of the 3' splice site (3'ss), promotes intron retention, expanding the relative abundance of transcripts with extended 5'UTR (21). As compared to thymine (T), the A allele at IVS1– 6A/T decreases affinity to pyrimidine-binding proteins *in vitro* and renders the 3'ss more dependent on the auxiliary factor of U2 small nuclear ribonucleoprotein (U2AF) (7), a heterodimer required forU2 binding, spliceosome assembly and 3'ss selection (22). Intron 1-containing transcripts are overrepresented in IVS1-6A-derived cDNA libraries prepared from insulin producing tissues (21), are exported from the nucleus (23) and contain a short, *Homininae*-specific uORF that co-evolved with relaxation of the 3'ss of intron 1 in higher primates (7). The lower proinsulin expression conferred by the A allele may lead to suboptimal presentation of proinsulin peptides in the fetal thymus and inadequate negative selection of autoreactive T cells, culminating in autoimmune destruction of insulin-producing β cells in the pancreas (7). However, no attempts have been made to correct the low efficiency of *INS* intron 1 removal from the IVS1-6A-containing pre-mRNAs and reduce intron retention to the levels observed for the disease-protective T allele.

[00278] This study set out to search for SSOs that increase the efficiency of *INS* intron 1 splicing and repress splicing silencers or decoy splice sites in the pre-mRNA to enhance proinsulin expression. SSOs reducing the relative abundance of intron 1-retaining transcripts were identified, delineation of the optimized antisense target at a single-nucleotide resolution is shown, and evidence is shown for formation of a parallel G-quadruplex adjacent to the antisense target sequence and identification of proteins that bind to this region.

MATERIALS AND METHODS

Antisense oligonucleotides

[00279] SSOs were purchased from the MWG Biotech (Germany). All SSOs and scrambled controls had a full-length phos-phorothioate backbone with 2'-*O*-methyl ribonucleotides at the second ribose position. Apart from *INS* SSOs and their scrambled versions, we employed SSOs that target other human genes as additional controls, as described (13). Location of each SSO is shown in Figure 1A and their sequences in Table 2.

Splicing reporter constructs

[00280] The wild-type splicing reporter carrying the type 1 diabetes associated haplotype termed IC was reported previously (7,21). Each construct contains all *INS* exons and unabridged introns but differ in the length of the last exon. The IC reporters were cloned using primers D-C, D-F and D-B; IC D-B lacks the cryptic 3'ss of intron 2. The relative abundance of isoforms spliced to this site is lower for IC DF than for IC D-C (7,21). To test SSOs targeting the cryptic 5' splice site of intron 1, the IC construct was modified by a 4-nt insertion at *rs3842740* to create a reporter termed ICIVS1+ 5ins4. *TSC2* and *F9* constructs were reported previously (24). Plasmids were propagated in the *E. coli* strain DH5α and plasmid DNA was extracted using the Wizard Plus SV Miniprep kit (Promega, USA). Their inserts were completely sequenced to confirm the identity of each of the 14 intragenic natural variants and to exclude undesired mutations.

Cell cultures and transfections

[00281] Human embryonic kidney 293 (HEK293), human hepatocellular liver carcinoma HepG2 and African green monkey COS7 cells were cultured in Dulbecco's modified Eagle medium, 10% fetal calf serum and penicillin/streptomycin (Life technologies, USA). Transient transfections were carried out as described (13), using jetPRIME (Polyplus, USA) according to manufacturer's recommendations. Downregulation of U2AF35 by RNA interference (RNAi) to induce cryptic 3'ss of intron 1 was performed with two hits of small interfering RNA (siRNA) U2AF35ab, as reported earlier (7,25); siRNA duplex targeting DHX36 was as described (26). The second hit was applied 24

h before the addition of SSOs and/or reporter. Cell cultures were harvested 24 h after addition of reporter constructs.

Analysis of spliced products

[00282] Total RNA was extracted with TRI Reagent and treated with DNase (Life technologies, USA). The first-strand cDNA was reverse transcribed using oligo-(dT)15 primers and Moloney murine virus reverse transcriptase (Promega, USA). Polymerase chain reaction (PCR) was carried out with a combination of a vector-specific primer PL3 and primer E targeting the 3'UTR, as reported previously (7). PCR products were separated on polyacrylamide gels and their signal intensity was measured as described (27). The identity of each mRNA isoform was confirmed by Sanger nucleotide sequencing.

Circular dichroism and nuclear magnetic resonance spectroscopy

[00283] Oligoribonucleotides for circular dichroism (CD) and nuclear magnetic resonance (NMR) were purchased from Thermo Scientific, deprotected according to manufacturer's instructions, lyophilized and stored at −20◦C. Stock solutions were prepared from the desalted, lyophilized samples by resuspending in milliQ water or KCl buffer (100 mM KCl, 10mMK_2HPO4/KH2PO_4, pH 7.0, milliQ water) to a final concentration of 2–4 μM. CDspectra were acquired using a PiStar-180 spectrophotometer (Applied Photophysics Ltd, Surrey, UK), equipped with a LTD6G circulating water bath (Grant Instruments, UK) and thermoelectric temperature controller (Melcor, USA). Samples were heated in the cell to 95°C for a total period of 15 min, samples were then annealed by allowing to cool to room temperature for a minimum period of 4 h. CD spectra were recorded over a wavelength range of 215– 340 nm using a 1 cm path length strain-free quartz cuvette and at the temperatures indicated. Data points recorded at 1 nm intervals. A bandwidth of 3 nm was used and 5000 counts acquired at each point with adaptive sampling enabled. Each trace is shown as the mean of three scans (±SD). CD temperature ramps were acquired at 265 nm corresponding to the band maxima of the folded quadruplex species. Ranges between 5 and 99°C were used, with points acquired at 0.5°C intervals with a 120–180 s time step between 0.5°C increments. Points were acquired with 10,000 counts and adaptive sampling enabled. Heating and cooling studies were compared to check for hysteresis and overall reversibility. NMR spectra (1H) were collected at 800 MHz using a Bruker Avance III spectrometer with a triple resonance cryoprobe. Standard Bruker acquisition parameters were used. Data were collected using Topspin (v. 3.0) and processed in CCPN Analysis (v. 2.1).

Pull-down assays and western blotting

[00284] *In vitro* transcription was carried out using MEGAshortscriptTM T7 (LifeTechnologies, USA) and T7-tagged PCR products amplified with primers 5'-ATTAATACGACTCACTATAGGGCTCAGGGTTCCAGG and 5'- TGCAGCAGGGAGGACG, and DNA of the indicated plasmids as a template. Indicated synthetic RNAs were purchased from Eurofins UK. Five hundred pmols of each RNA was treated with 5 mM sodium *m*-periodate and bound to adipic acid dehydrazide agarose beads (Sigma, USA). Beads with bound RNAs were washed three times in 2 ml of 2 M NaCl and three times in buffer D (20 mM HEPES–KOH, pH 7, 6.5% v/v glycerol, 100 m M KCl, 0.2 mM EDTA, 0.5 mM dithiothreitol), incubated with HeLa nuclear extracts and buffer D with heparin at a final concentration of 0.5 mg/ml. Unbound proteins were washed five times with buffer D. Bound proteins were separated on 10% sodium dodecyl sulphate polyacrylamide gel electrophoresis, stained by Coomassie blue and/or blotted nitrocellulose membranes.

[00285] Western blotting was carried out as described (7). Antibodies were purchased from Sigma (hnRNP E1/E2, product number R4155, U2AF65, product number U4758 and SFRS2, product number S2320), Abcam (DHX36, product number ab70269) and Millipore (SC35, clone 1SC-4F11). Antiserum against hnRNP F and hnRNP H provided by Prof. Douglas Black, UCLA.

Mass spectrometry analysis

[00286] Following trypsin digestion, samples were freeze dried and resuspended with 25 µl of 5% ACN/0.1% formic acid for mass spectrometry (MS). Peptides were analysed by LC/MS/MS using a Surveyor LC system and LCQ Deca XP Plus (ThermoScientific). The raw data files were converted into mascot generic files using the MassMatrix File Conversion Tool (Version 2.0; http://www.massmatrix.net) for input into the Mascot searching algorithm (Matrix Science).

Enzymatic structural probing

[00287] RNA secondary structure determination with the use of limited V1 RNAse (Ambion), T1 RNAse (Ambion) and S1 nuclease (Fermentas) digestion has been described in detail elsewhere (28). Briefly, 1 µg aliquots of RNAs from the insertion (ins) and deletion (del) pre-mRNAs were digested with 0.002 U of RNAse V1, 0.05 U of RNAse T1 and 19 U of S1 nuclease in a 100 µl at 30◦C for 10 min. An enzyme free aliquot was used as a control (C). The cleaved RNAs were retro-transcribed according to standard protocols using antisense primers labeled with [^{32}P]-ATP at the 5' end.

RESULTS

Antisense oligonucleotides that promote pre-mRNA splicing of a weak intron in 5'UTR

[00288] To identify SSOs capable of reducing retention of *INS* intron 1 and increase splicing-mediated translational enhancement, we designed a series of 2'-*O*-methyl-modifed phosphorothioate SSOs, individually co-expressed each SSO with a splicing reporter construct carrying haplotype IC in HEK293 cells and examined the relative abundance of exogenous mRNA products (Figure 1A and B). The IC haplotype in the reporter was devoid of the minisatellite sequence and contained a total of 14 polymorphic sites (7,20), including the A allele at *rs689*. This allele inhibits intron 1 splicing and yields lower proinsulin levels as compared to the more common T allele (21). SSOs targeting intron 1 and exon 2 were chosen in regions that showed the most prominent alterations of exon inclusion or intron retention in previous systematic deletion analyses of these sequences (7). SSOs in exon 3 were located between authentic 3'ss of intron 2 and a strong competing cryptic 3'ss 126 nt downstream to identify pre-mRNA motifs that modify their usage (Figure 1A). Of the initial set of 15 *INS* SSOs tested in HEK293 cells, 11 showed reproducible alterations in the relative abundance of mRNA isoforms (Table 2). Intron 1 retention was significantly reduced by a single oligoribonucleotide SSO21 ($P < 0.01$, Mann-Whitney rank sum test; Figure 2A). SSO21 targeted intron 1 positions 59–74, encompassing a motif (termed del5) previously found to confer the largest reduction of intron retention upon deletion (7). The decrease in intron retention levels induced by SSO21 was dose-dependent (Figure 2A) and was also observed in HepG2 cells and *Chlorocebus aethiops* COS7 cells, consistent with ubiquitous expression and a high degree of evolutionary conservation of spliceosome components that employ auxiliary splicing sequences (1,2). In addition to reducing intron 1 retention, SSO21 promoted cryptic 3'ss of intron 2 (Figure 2A). However, this effect was also seen for other *INS* SSOs and for scrambled controls (Figure 3 and Table 2), suggesting non-specific interactions. To confirm that the SSO21-induced enhancement of intron 1 splicing is not facilitated by the cryptic 3'ss of intron 2, we co-transfected this SSO with a shorter reporter lacking this site and retaining only the first 89 nucleotides of exon 3. Figure 2B shows that SSO21 was capable of promoting intron 1 splicing to the same extent as the reporter with longer exon 3. In contrast, the SSO21-induced decrease of intron retention was not observed for the reporter lacking the del5 segment. Apart from intron retention, an increase of exon 2 skipping was observed for five SSOs, including SSO8 that bound downstream of the cryptic 3'ss of intron 1 (cr3'ss+81; Figures 1 and 3C, Table 2). This cryptic 3'ss was induced by RNAi-mediated depletion of the small subunit of U2AF (U2AF35) and was not reversed by a bridging oligoribonucleotide (SSO4) in cells lacking U2AF35; instead exon 2 skipping was observed (Figure 3C). Depletion of

U2AF35 also repressed the cryptic 3'ss of intron 2. Taken together, a single SSO was identified that reduced *INS* intron 1 retention in several primate cell lines.

Optimization of the intron retention target at the single nucleotide level

[00289] Interestingly, other SSOs designed to target the del5 segment did not reduce intron 1 retention, except for a small effect of SSO20 (Figures 1A and 2A). To test the importance of nucleotides flanking SSO21 and to map the optimal target at a single-base resolution, a detailed antisense microwalk was carried out in this region. The *INS* reporter was individually co-transfected with additional eighteen 16-mers bound 1–9 nucleotides 5' and 3' of SSO21 into HEK293 cells and their RNA products examined. Intron 1 retention was most repressed by SSO21 and by SSOs that were shifted by 1–2 nucleotides in each direction (Figure 4). In agreement with the initial screen, SSOs targeting more than one cytosine in the upstream run of four Cs (C4, see SSO1 and SSO2, Figure 1A) were not effective (SSO21–3r through SSO21–10r, Figure 4). In the opposite direction, SSOs targeting consecutive Gs, which are often found in intronic splicing enhancers (29–31), increased intron retention. Thus, the optimal antisense target for reducing retention of *INS* intron 1 was mapped at a single nucleotide resolution to a region previously identified as the most repressive by a systematic deletion analysis of the entire intron (7).

Antisense target for intron retention is adjacent to a parallel RNA quadruplex

[00290] It was noticed that the target was sandwiched between two intronic segments predicted to form stable RNA guanine (G) quadruplexes (intron 1 nucleotides 36–61 and 78–93; highlighted in Figure 4A). These structures are produced by stacking G-quartets that consist of four Gs organized in a cyclic Hoogsteen hydrogen bonding arrangement (32) and have been implicated in important cellular processes, including replication, recombination, transcription, translation (33,34) and RNA processing (35–39). To test if they are formed *in vitro*, synthetic ribonucleotides derived from this region were employed in CD spectroscopy that has been used widely to characterize DNA and RNA quadruplex structures in vitro (40–43). The CD spectrum of a downstream 19-mer (termed CD1) recorded between 215 and 330 nm at 25°C revealed strong positive ellipticity at 265 nm with negative intensity at around 240 nm, indicative of a parallel quadruplex (Figure 5A). To confirm the presence of a quadruplex, rather than other stable secondary structure motifs, UV absorbance spectra was recorded at 5°C and 95°C. The UV absorbance difference spectrum at the two temperatures (below and above the melting transition point) showed the characteristic hyperchromic shift at ~295 nm and a double maximum at 240 nm and 280 nm, providing evidence for formation of a stable parallel-stranded RNA quadruplex *in vitro*. This was confirmed by 1H

NMR studies of CD1 (Figure 5B) which showed a characteristic envelope of signals between 10 and 12 ppm corresponding to Hoogsteen H-bonded Gs within G-tetrad structures. Thermal stability measurements by CD produced a highly reversible sigmoidal co-operative unfolding transition with a $T_m = 56.8 \pm 0.2°C$ (Figure 5C). Figure 5D (upper panel) shows a possible arrangement of the 19-mer into two stacked G-tetrads connected by relatively short loop sequences of 1–4 nucleotides.

Conformational transition model for splicing inhibitory sequences in *INS* intron 1

[00291] CD of a synthetic 20-mer derived from a region upstream of the antisense target (termed CD2) also showed evidence of stable structure formation, giving a broader absorption envelope centered around 270 nm and a sigmoidal thermal unfolding transition ($T_m = 69.0 \pm 0.45°C$; Figure 5A). Unlike the downstream oligo CD1, no hyperchromic shift in the UV was found in the thermal difference spectrum. However, a well-defined set of sharp signals in the 1H NMR spectrum between 12 and 14 ppm that differed from those for CD1 showed the formation of Watson–Crick H-bonded base pairs characteristic of double-stranded RNA (Figure 5B). Secondary structure predictions of overlapping intronic segments using Mfold suggested that the pre-mRNA forms stable local stem-loops; one of them was further stabilized by a G→C mutation (termed G2; Figure 5D, lower panel) that increased intron 1 retention (7). Another G→C substitution (termed G3) located further downstream and destabilizing the quadruplex structure (Figure 5D, upper panel) also repressed intron splicing (7). Finally, CD2 oligonucleotides containing either A or G at a single-nucleotide polymorphism (Figure 4A and (20) exhibited very similar CD spectra with well-defined melting transitions and T_m values, suggesting that the G and A alleles form the same structure.

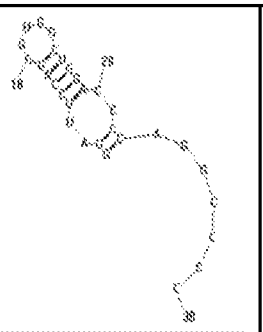
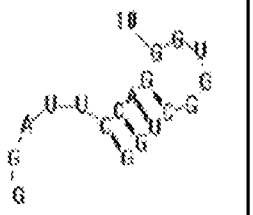
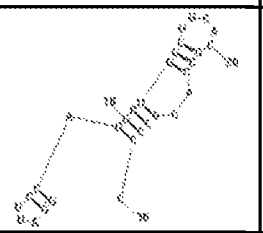
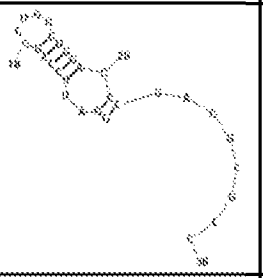
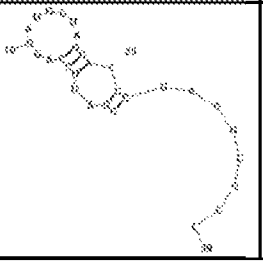
[00292] To test further the importance of a tentative equilibrium between canonical and noncanonical structures in intron splicing, a combination of CD, NMR and mutagenesis experiments was used (Figure 6). An oligoribonucleotide CD3 encompassing the 5' end of the intron retention target was synthesized and stem-loops/quadruplex were predicted (Figures 4A and 6A). A mutated version CD4 was also synthesized, which carried two C→U transitions destabilizing the hairpin but maintaining stability of the quadruplex. The same mutation was also introduced in the IC reporter construct transfected into HEK293 cells. The NMR spectrum of CD3 revealed the co-existence of signals for both G-tetrad and canonical base-paired hairpin structures (termed H1 and H2) in equilibrium (Figure 6B and C). The effects of Mg2+ on the conformational equilibrium between quadruplex and hairpin were investigated by adding 2 mM and then 6 mM $MgCl_2$ to the buffered solution containing 100 mM KCl. As reported by Bugaut *et al.* (57), the conformational equilibrium was not significantly perturbed by the addition of Mg2+ in the presence of KCl. Thus, we observed

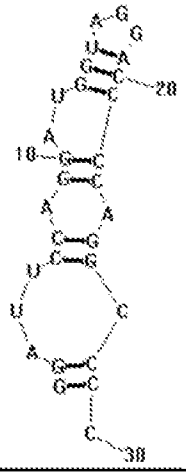
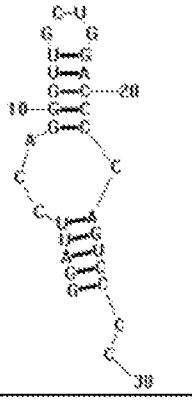
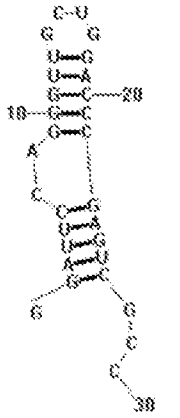
formation of the RNA hairpin and quadruplex structures in an environment that mimics the cellular context where both K^+ and Mg^{2+} ions were present at high concentrations. The CD melting curve showed a broad transition (T_m = 79.9◦C), consistent with multiple conformational states with different stabilities. The CC→UU mutation in CD4 resulted in the loss of NMR signals for H1 (Figure 6B) and a reduction in the *Tm* by 13°C, consistent with the selective destabilization of the more stable hairpin H1, leading to an increase in the population of H2 in equilibrium with the quadruplex. Transient transfections showed that the CC→UU mutation improved intron 1 splicing while a mutation termed M1 predicted to destabilize both the quadruplex and the hairpin had only a small effect (Figure 6D, Table 1A).

[00293] To explore how the equilibrium of these structures affects intron splicing more systematically, a series of mutated constructs were prepared in order to destabilize/maintain predicted quadruplex, H1/H2 structures and two cytosine runs (Table 1A). Their transcripts showed significant differences in intron retention levels (Figure 7; P = 0.0001, Kruskal–Wallis one-way ANOVA on ranks). First, elimination of the G-quadruplex increased intron 1 retention, which was further enhanced by removing each cytosine run (cf. mutations 4–6 with the wild-type, P = 0.0004). These mutations appeared to have additive effects on intron retention (cf. wild-type versus mutations 1 or 9; 3 versus 2 and 4 versus 5). Second, the increased intron retention in the absence of the G-quadruplex was not altered by removing H1 and H2, but their elimination enhanced exon skipping (cf. isoform 2 for mutations 4 versus 6). Third, when only one of the two C4 runs was present, removal of H1 somewhat improved intron 1 splicing (cf. 8 versus 9), consistent with a statistically significant correlation between intron retention and predicted stability of tested RNAs (Figure 7B). The efficiency of intron splicing was thus controlled by conformational transitions between canonical and noncanonical structures in equilibrium.

TABLE 1A *INS* intron 1 mutations altering predicted RNA G quadruplexes, stem loops and two cytosines runs in plasmid constructs

Mutation	Input sequence for computation predictions1	Predicted RNA quadruplex	H2	H1	C runs	The most stable RNA structure	Free energy (kcal/mmol)

Wildtype sequence	GGAUUCCAGGG UGGCUGGACCC CAGGCCCC	+	+	+	+		-9.8
Del5	GGAUUCCAGGG UGGCUGG--------	+	+	-	-		-2.7
M1	GGAUUCCAGGG UGGCUGCACGC CAGGCCCC	+	-	-	-		-9.3
1	GGAUUCCAGGG UGGCUGGACCC GAGGCGCC	+	+	+	-		-8.4
2	GGAUUCCAGGA UGGUAGGACCC GAGGCGCC	+	-	-	-		-2.8

3	GGAUUCCAGGA UGGUAGGACCC CAGGCCCC	+	-	-	+		-4.3
4	GGAUUCCAGGG UUGCUGGACCC CAGUCCCC	-	+	+	+		-11.1
5	GGAUUCCAGGG UUGCUGGACCC GAGUCGCC	-	+	+	-		-11.3

Protein–RNA interactions in the region targeted by winner SSOs

[00294] To identify proteins that interact with RNAs encompassing the antisense target and/or associated canonical and noncanonical structures, pull-down assays were carried out using wild type and del5 RNAs transcribed from T7-tagged PCR products, a synthetic RNA(CD5) representing the target sequence, and a control oligo containing a 3'ss CAG, termed AV3.Western blotting showed that both wild type and del5 transcripts bound hnRNPs F/H but this binding was absent for CD5

(Figure 7C). These proteins were also detected by MS/MS analysis of differentially stained fragments from pull down gels with wild type and del5 RNAs as compared to beads-only controls. Two antibodies against SRSF2, which showed the highest score for putative binding activity among several SR proteins, failed to detect any specific interaction (Figure 7C). Although the signal from hnRNP E1/E2, which constitute a major poly(C) binding activity in mammalian cells (44), was above background for del5 (Figure 7C), no change in intron retention was observed in cells lacking hnRNP E1/E2.

Splicing pattern of G-rich and G-poor reporters upon DHX36 depletion

[00295] RNA G-quadruplexes bind helicase DHX36, which is capable of converting quadruplex RNA to a stable duplex and is a major source of quadruplex-resolving activity in HeLa cells (26,45). DHX36 was crosslinked to an intronic splicing enhancer in the *ATM* pre-mRNA (46) and could unwind the quadruplex structure within the 5' region of *TERC* (26). To test if DHX36 depletion can influence *INS* splicing, G-quadruplex-poor and -rich reporters were transiently transfected (Figure 8A, Table 1) into depleted cells. Control constructs were chosen to give approximately equal representation of spliced products, which was achieved by weakening the branch site (24), thus providing a sensitive *ex vivo* splicing assay. However, despite efficient DHX36 depletion (Figure 8B), statistically significant alterations of *INS* intron 1 retention were not seen in either short or long constructs, nor did we observe major changes in G-poor and G-rich controls (Figure 8C–E). These results are in agreement with a previous lack of significant enrichment of quadruplex sequences among transcripts down regulated in DHX36-depleted cells (47) and with the absence of *ATM* response to the knockdown (46).

SSO-induced repression of a population-specific cryptic 5' splice site of *INS* intron 1

[00296] In addition to *rs689, INS* intron 1 splicing is influenced by a polymorphic TTGC insertion at *rs3842740* located in the vicinity of the natural 5'ss (21). This insertion is present in a quarter of all African chromosomes but is absent on Caucasian IC haplotypes (20). The insertion activates a downstream cryptic 5'ss (Figure 1A), extending the 5'UTR of the resulting mRNAs by further 26 nucleotides and repressing proinsulin expression (7,21). To test if the new 5'ss can be efficiently inhibited by SSOs, the same insertion was introduced in the IC construct and the wild type and mutated reporters co-expressed with a bridging oligoribonucleotide termed SSO10. Although the cryptic splicing was inhibited, canonical splicing of intron 1 was not completely restored even at high SSO10 concentrations, most likely as a result of suboptimal recognition of the authentic 5'ss weakened by the insertion.

[00297] To gain initial insights into folding of 5'UTR sequences in the presence and absence of the insertion, enzymatic structural probing was carried out using partial RNA digestion with single- and double-strand specific RNAses. The overall cleavage positions and intensities detected for the wild-type RNA were broadly consistent with mfold predictions, in which two major stem loop regions (SL1 and SL2) were interrupted by several internal bulges. Both the structural probing and mfold predictions suggested that the insertion at *rs3842740* extended the central bulge in SL1 as the number of T1 and S1 cleavages in this region increased in contrast to the remaining portions of SL1 and in SL2. Finally, transcripts were not digested by RNase V1 in regions showing quadruplex formation *in vitro*.

DISCUSSION

Antisense intron retention target in a splicing silencer of *INS* intron 1

[00298] Here it is demonstrated the first use of antisense technology to reduce retention of the entire intron in mature transcripts and to modify the haplotype-dependent *INS* expression using SSOs. Identification of winner SSOs that compensate the adverse impact of the A allele at *rs689* on efficient RNA processing was facilitated by systematic mutagenesis of intron 1 (7), and by macro- (Figure 1) and micro-walk (Figure 4) strategies. Interestingly, the target sequence contains a tandem CAG(G/C) motif, which resembles a 3'ss consensus (Figure 4). Such 'pseudo-acceptors' were previously implicated in splice-site repression experimentally (27) and are overrepresented in splicing silencers. For example, the two tetramers are more common among high-confidence 102 intronic splicing silencers (49) and are depleted in 109 enhancers (50) identified by fluorescence activated screen of random 10-mers. The YAG motifs were also more frequent than expected among QUEPASA splicing silencers (51), suggesting that they are important functional components of the retention target. The intervening cytosine tract may also play an important role as the frequency of C4 runs among QUEPASA silencers is ~2 times higher than expected. These motifs were also found in 4% of intronic splicing regulatory elements identified by a systematic screening of sequences inserted at positions –62/–51 relative to a tested 3'ss (52). This study identified an element termed ISS22 (AAATAGAGGCCCCAG) that shared a 3' nonamer (underlined) with the optimal intron retention target. However, unlike an optimal 3'ss recognition sequence of AV3, our pull-down assay coupled with western blotting revealed only a very weak binding if any to U2AF65 (Figure 7C).

Conformational transition between quadruplex and hairpins in RNA processing control

[00299] The antisense target was identified just upstream of a potential G-quadruplex forming RNA whose structure was subsequently confirmed by CD and NMR analysis (Figures 1A and 5). RNA quadruplexes are more stable than their DNA counterparts, have been increasingly implicated in regulation of RNA metabolism (33,34,41,42) and offer unique avenues for drug development (53). The 2-quartet quadruplexes are thermodynamically less stable than their 3- or 4-quartet counterparts and are probably kinetically more labile, yet they still display pronounced stability and may serve as more compliant and dynamic switches between quadruplex and non-quadruplex structures in response to cellular environment (54–56). The winner SSOs may block interactions with *trans*-acting factors, alter higher-order structures, the rate of RNA–protein complex formation or impair conformational transition between the 2-quartet quadruplex and H1/H2 (Figure 5). A similar transition has been recently described for a quadruplex not predicted *ab initio* (57), raising a possibility that additional sequences in the G-rich intron 1 may participate in the equilibria near the antisense target, possibly involving multiple quadruplex motifs and competing stem-loops.

[00300] The binding (Figure 7C) and functional experiments showing the increased intron 1 retention upon hnRNP F/H depletion and the opposite effect upon hnRNP F/H overexpression (7) indicate that these proteins interact with key splicing auxiliary sequences in this intron. In contrast to a previous report concluding that hnRNP F binds directly to the RNA quadruplex (58), hnRNP F has been shown to prevent formation of RNA quadruplexes by binding exclusively single-stranded G-tracts (59). Predictions based on primate genomes suggest that the majority of putative quadruplexes are likely to fold into canonical structures (60). Decreased pre-mRNA occupancy by these proteins, presumably promoting quadruplex formation (59), and potentially reducing splicing efficiency.

RNA quadruplexes in coupled splicing and translational gene expression control

[00301] RNA quadruplexes were predicted in ~8.0% of 5'UTR and were proposed to act as general inhibitors of translation (60,61). *INS* intron 1 is weakly spliced and U2AF35-dependent (7) and a significant fraction of intron 1-containing transcripts is exported from the nucleus (23). This suggests that the RNA G-quadruplex formed by CD1 could influence translation of these mRNAs, which contain a three-amino acid uORF specific for *Homininae* (7). This uORF markedly inhibits proinsulin expression and is located just a few base-pairs downstream, prompting a concept that the G-quadruplexes can promote translation by sequestering uORFs. Functional 2-quartet quadruplexes are required for activity of internal ribosomal entry sites (54).

Table 1. Density of predicted RNA G-quadruplexes in reporter constructs

Reporter	TSC2	F9	INS
G-quadruplexes per nucleotide[a]	0.25	0.05	0.27
G score per nucleotide[a]	0.20	0.04	0.22

[a]The length of non-overlapping quadruplex sequences and their G scores were computed as described (78).

Antisense strategies for dependencies in splice-site selection

[00302] Apart from canonical mRNA isoform 4, isoforms 2, 3 and 6 (Figure 1B) have been found in expressed sequence tag databases derived from cDNA libraries from insulin producing tissues (21). This suggests that cryptic splice sites produced by the reporter construct are recognized *in vivo* and that our haplotype-dependent reporter system recapitulates these events accurately in cultured cells no matter whether the cells express or not endogenous insulin. Apart from repressing intron 1-retaining transcripts, optimal SSOs increased utilization of cryptic 3'ss of exon 3 (Figure 2). This undesired effect could be explained by coordination of splicing of adjacent exons and introns, which was observed previously for individual genes and globally (63–67). Also, G-richness downstream transcription start sites have been associated with RNA polymerase II pausing sites (68). Although the two robustly competing 3'ss of intron 2 are likely to respond to non-specific signals that influence RNA folding (Figure 3, Table 2), it might be possible to alleviate the observed dependencies and reduce cryptic 3'ss activation using SSO combinations at linked splice sites and examine their synergisms or antagonisms, benefiting from the use of full-gene constructs as opposed to minigenes.

Multifunctional antisense oligonucleotides to reduce *INS* intron 1 retention

[00303] Since the first use of 2'-*O*-methyl-phosphorothioate SSOs (69), this type of chemical modification has been successfully exploited for many *in vitro* and *in vivo* applications (9,10,70). To further fine-tune expression of mRNA isoforms, optimized SSOs can be designed to tether suitable *trans*-acting splicing factors to their target sequences (11,71). An obvious candidate for this system is U2AF35 because intron 1 is weak as a result of relaxation of the 3'ss in higher primates and is further undermined by the A allele at *rs689*, which renders this intron highly U2AF35-dependent (Figure 3) (7). Apart from U2AF35, future bi- or multifunctional antisense strategies can employ binding platforms for splicing factors previously shown to influence *INS* intron 1 and exon 2

splicing, such as Tra2β or SRSF3 (7). Tra2β is likely to bind the SSO6 target which forms a predicted stable hairpin structure with a potent GAA splicing enhancer in a terminal loop (Figure 3B). SRSF3 is required for repression of the cryptic 3'ss of intron 2 (7) and binds pyrimidine-rich sequence with a consensus (A/U)C(A/U)(A/U)C (72). The CAUC motif, which interacts with the RNA-recognition motif of SRSF3 (73), is present just upstream of the cryptic 3'ss.

Normalizing intron retention levels in human genetic disease

[00304] These results provide an opportunity to use non-genetic means to compensate less efficient splicing and lower *INS* expression from haplotypes predisposing to type 1 diabetes.

[00305] Common variants such as *rs689* contribute to a great extent to the heritability of complex traits, including autoimmune diseases (74), but their functional and structural consequences are largely unknown. If optimized *INS* SSOs can be safely and efficiently introduced into the developing thymus, this approach may offer a novel preventive approach to promote tolerance to the principal self-antigen in type 1 diabetes. The most obvious candidates for such intervention are mothers who had an affected child homozygous for disease-predisposing alleles at both HLA and *INS* loci. Such genotypes were associated with an extremely high disease risk for siblings (75). Apart from primary prevention of type 1 diabetes, future SSO-based therapeutics might be applicable to patients with significant residual β-cell activity at diagnosis and to those who are eligible to receive β-cell transplants and may benefit from increased intron-mediated enhancement of proinsulin expression from transplanted cells. It is also possible to envisage use of this therapeutic modality for other patients with diabetes through a more dramatic enhancement of intron splicing and proinsulin expression by targeting multiple splicing regulatory motifs with multifunctional SSOs. The SSOs may have utility in thymic epithelial cells and 13-cells that may provide a more natural system for testing their impact on both exo- and endogenous proinsulin expression. Finally, similar antisense strategies may help reduce pervasive intron retention in cancer cells resulting from somatic mutations of splicing factor genes, as illustrated by specific substitutions in the zinc finger domain of U2AF35 in myeloproliferative diseases (76).

Reducing intron retention in malignant cells

[00306] A set of 146 intronic sequences that are preferentially retained in U2AF-deficient HEK293 cells was selected using RNAseq data from replicated, polyA-selected samples, followed by inspection of each intron retention event in genome browsers. These sequences were repeat-masked using a sensitive version of RepeatMasker, available at http://www.repeatmasker.org/cgi-

bin/WEBRepeatMasker. Because the optimal antisense target for reducing *INS* intron 1 retention [Kralovicova J et al (2014). *Nucleic Acids Res* doi: 10.1093/nar/gku507, published on 17 June 2014] overlapped intronic splicing regulatory elements conserved in mammals [Yeo GW, et al (2007). *PLoS Genet* 3:e85], including CCCAG, AGGCC (Fig. 4 in Kralovicova et la. 2014), antisense targets in intronic segments containing these short penta- to heptamer motifs and an independently derived set of intronic splicing regulatory motifs [Voelker RB, & Berglund JA (2007). *Genome Res* 17:1023-1033], were selected, thus increasing the chance of oligonucleotides interacting with these targets to influence RNA processing. The target sequences were subjected to routine antisense oligonucleotides design, including removal of sequences containing C runs to avoid G-quadruplex formation. The proximity of both 5' and 3' splice sites, polypyrimidine tracts, branch sites and suprabranch regions were also avoided. This selection yielded a set of 388 compounds (Table 3), covering ~15% of the total lengths of U2AF-sensitive introns and representing ~0.001% of all human intronic sequences. Thus, this set of oligonucleotides target regions enriched for splicing inhibitory sequences of U2-dependent introns, which do not have sufficient help from auxiliary factors in malignant cells that sustain mutations or deletions in the U2 pathway.

INSULIN RELATED SEQUENCES

Candidate SSO sequences

Note: candidates marked with "*" are the winner oligos discussed in Figure 4.

CD5

[00307] RNA form - CUGCAGAGCUGGGGCCUG (SEQ ID NO: 1)

[00308] DNA form - CTGCAGAGCTGGGGCCTG (SEQ ID NO: 2)

[00309] Binds to caggccccagcucugcag (SEQ ID NO: 3)

SSO21*

[00310] RNA form - UGCAGAGCUGGGGCCU (SEQ ID NO: 4)

[00311] DNA form - TGCAGAGCTGGGGCCT (SEQ ID NO: 5)

[00312] Binds to aggccccagcucugca (SEQ ID NO: 6)

SSO21-2r*

[00313] RNA form - GCAGAGCUGGGGCCUG (SEQ ID NO: 7)

[00314] DNA form - GCAGAGCTGGGGCCTG (SEQ ID NO: 8)

[00315] Binds to caggccccagcucugc (SEQ ID NO: 9)

SSO21-3r

[00316] RNA form - CAGAGCUGGGGCCUGG (SEQ ID NO: 10)

[00317] DNA form - CAGAGCTGGGGCCTGG (SEQ ID NO: 11)

[00318] Binds to ccaggccccagcucug (SEQ ID NO: 12)

SSO21-4r

[00319] RNA form - AGAGCUGGGGCCUGGG (SEQ ID NO: 13)

[00320] DNA form - AGAGCTGGGGCCTGGG (SEQ ID NO: 14)

[00321] Binds to cccaggccccagcucu (SEQ ID NO: 15)

SSO21-5r

[00322] RNA form - GAGCUGGGGCCUGGGG (SEQ ID NO: 16)

[00323] DNA form - GAGCTGGGGCCTGGGG (SEQ ID NO: 17)

[00324] Binds to ccccaggccccagcuc (SEQ ID NO: 18)

SSO21-6r

[00325] RNA form - AGCUGGGGCCUGGGGU (SEQ ID NO: 19)

[00326] DNA form - AGCTGGGGCCTGGGGT (SEQ ID NO: 20)

[00327] Binds to accccaggccccagcu (SEQ ID NO: 21)

SSO21-7r

[00328] RNA form - GCUGGGGCCUGGGGUC (SEQ ID NO: 22)

[00329] DNA form - GCTGGGGCCTGGGGTC (SEQ ID NO: 23)

[00330] Binds to gaccccaggccccagc (SEQ ID NO: 24)

SSO21-8r

[00331] RNA form - CUGGGGCCUGGGGUCC (SEQ ID NO: 25)

[00332] DNA form - CTGGGGCCTGGGGTCC (SEQ ID NO: 26)

[00333] Binds to ggaccccaggccccag (SEQ ID NO: 27)

SSO21-9r

[00334] RNA form - UGGGGCCUGGGGUCCA (SEQ ID NO: 28)

[00335] DNA form - TGGGGCCTGGGGTCCA (SEQ ID NO: 29)

[00336] Binds to uggaccccaggcccca (SEQ ID NO: 30)

SSO21-10r

[00337] RNA form - GGGGCCUGGGGUCCAG (SEQ ID NO: 31)

[00338] DNA form - GGGGCCTGGGGTCCAG (SEQ ID NO: 32)

[00339] Binds to cuggaccccaggcccc (SEQ ID NO: 33)

SSO21-14f*

[00340] RNA form - CUGCAGAGCUGGGGCC (SEQ ID NO: 34)

[00341] DNA form - CTGCAGAGCTGGGGCC (SEQ ID NO: 35)

[00342] Binds to ggccccagcucugcag (SEQ ID NO: 36)

SSO21-15f

[00343] RNA form - GCUGCAGAGCUGGGGC (SEQ ID NO: 37)

[00344] DNA form - GCTGCAGAGCTGGGGC (SEQ ID NO: 38)

[00345] Binds to gccccagcucugcagc (SEQ ID NO: 39)

SSO21-16f

[00346] RNA form - UGCUGCAGAGCUGGGG (SEQ ID NO: 40)

[00347] DNA form - TGCTGCAGAGCTGGGG (SEQ ID NO: 41)

[00348] Binds to ccccagcucugcagca (SEQ ID NO: 42)

SSO21-17f

[00349] RNA form - CUGCUGCAGAGCUGGG (SEQ ID NO: 43)

[00350] DNA form - CTGCTGCAGAGCTGGG (SEQ ID NO: 44)

[00351] Binds to cccagcucugcagcag (SEQ ID NO: 45)

Sequence of target region of pre-mRNA transcript (e.g. for binding SSOs)

[00352] cuggaccccaggccccagcucugcagcag (SEQ ID NO: 46)

Table 2

SSO	**Location1**	**Sequence (5'-3')**	**Effects on the relative abundance of *INS* mRNA isoforms**
1	Intron 1 (del5, del6)	AGCUGGGGCCUGGGGU	Activation of the cryptic 3'ss of intron 2
2	Intron 1 (del5, del6)	UGCAGAGCUGGGGCCUGGGGU	Activation of the cryptic 3'ss of intron 2
3	Intron 1 (del8, del9)	CAUGCUUCACGAGCCCAGCC	Increased exon 2 skipping
4	Exon 2 (cryptic 3'ss +81, del8, del9)	AAGGCUGCGGCUGGGUC	Increased exon 2 skipping
5	Exon 3	UGGUAGAGGGAGCAGAUGCUG	Decreased efficiency of intron 2 splicing; Activation of the cryptic 3'ss of intron 2
6	Exon 3	UGGUACAGCAUUGUUCCACA	Activation of the cryptic 3'ss of intron 2 at high concentration
8	Exon 2 (del13-15)	CGCACACUAGGUAGAGAGC	Increased exon 2 skipping
9	Exon 1	GAUGCAGCCUGUCCUGGAG	None
10	Intron 1 (del1, del2, cryptic 5' splice site	GAGCCCACCUGACGCAAAGGC	Partial restoration of authentic 5' splice site

	+30)		
16	Exon 1	UGGAGGGCUGAGGGCUGCU	None
17	Exon 1	AUGGCCUCUUCUGAUGCA	None
18	Intron 1 (del9, del10)	UCACCCCCACAUGCUUC	Increased exon 2 skipping
19	Intron 1 (del9)	ACAUGCUUCACGAG	Increased exon 2 skipping
20	Intron 1 (del5)	CUGGGGCCUGGGGU	Minor reduction of intron 1 retention; activation of the cryptic 3'ss of intron 2
21	Intron 1 (del5, del6)	UGCAGAGCUGGGGCCU	Reduction of intron 1 retention; activation of the cryptic 3'ss of intron 2
1sc	Scrambled control	AGGUGCUCGCGGGUGG	None
2sc	Scrambled control	GGGUGGAAGCGUCCGGUCGUG	Stimulation of the cryptic 3'ss of intron 2
3sc	Scrambled control	ACACACUGUGCCUCGCCAGC	None
6sc	Scrambled control	GACUCACUUGCCGUAGUUAA	Stimulation of the cryptic 3'ss of intron 2
8sc	Scrambled control	CACGCUCAGUAGAGAAGGC	None

[1], sequence of deleted segments (del) is shown in Fig. 2

CANCER RELATED SEQUENCES

[00353] The sequences of the following table (Table 3) may also be provided with thymine residues substituting the uracil residues (e.g. in DNA form). Each sequence of the following table may be an embodiment of the polynucleic acid polymer of the invention. Each gene or ORF referred to in the table below (Table 3) under "name of compound", may comprise the gene target for correction of intron retention.

Table 3

Name of compound	Corresponding nucleotide sequence of antisense oligoribonucleotide	
ABCD4-1	SEQ ID NO: 47.	UAGAGAGGUGUGGGAAGGGAAGCAGA
ABCD4-2	SEQ ID NO: 48.	AAUUCCUUCAUCAUGGCACAUUUAUCCUUGCAGACAGG
ABCD4-3	SEQ ID NO: 49.	CCUGAGGAUACUCACAGAAAGGCAACAG
ABCF3-1	SEQ ID NO: 50.	UUUCCCCAACACACUCCAGCA
ACADVL-1	SEQ ID NO: 51.	GGGCCGCUGCCCACCGUC
ALKBH6-1	SEQ ID NO: 52.	CAGCACAGCUCAGAAGUCUGAG
ALKBH6-2	SEQ ID NO: 53.	CCAAGCCAGGGACAGGGAGGUGAAUGCC
AP1G2-1	SEQ ID NO: 54.	CUUCUGCCCAGCUCUCUGACUG

APEX1-1	SEQ ID NO: 55.	UCUUCACAAACCCCUGCAAAAAUGAG
ARFRP1-1	SEQ ID NO: 56.	CCCAAAGCCCCCGCAGGUGCAGCC
ATHL1-1	SEQ ID NO: 57.	CCCCUCCCCACGCUCUGGAAA
ATHL1-2	SEQ ID NO: 58.	GCAGCACCGGGAGGCUCAGACAAC
ATHL1-3	SEQ ID NO: 59.	GAGCCUCAUCAAAGAAACGG
ATP13A1-1	SEQ ID NO: 60.	GCUCCCACUGGGACUGAGCG
ATP1A3-1	SEQ ID NO: 61.	AGAUGGGAAGAGAGAGAAGAG
ATP1A3-2	SEQ ID NO: 62.	AGAGACAAGGAAACCACACAGACAGAGACC
ATP1A3-3	SEQ ID NO: 63.	GCCGCCCAGCAGAGAGAGG
ATP5D-1	SEQ ID NO: 64.	AGCUGGCUGGGCCCACCUGGCAU
ATP5D-2	SEQ ID NO: 65.	GGGCCCAGGCAGAAGCCU
ATP5D-3	SEQ ID NO: 66.	UCCCCAGAGCUUUCAACACAG
ATP5D-4	SEQ ID NO: 67.	GCAGCCACAGCUCAAAGCUGAGGA
BAX-1	SEQ ID NO: 68.	GAUCAGACACGUAAGGAAAACGCAUUAUA
BAX-2	SEQ ID NO: 69.	GCAGAAGGCACUAAUCAAGUCAAGGUCACA
BAX-3	SEQ ID NO: 70.	CGGGCAUUAAAGAGCUGGACUCAG
BDH2-1	SEQ ID NO: 71.	ACCAAUUUUGAAAAAAGCAG
BDH2-2	SEQ ID NO: 72.	CCACAUUUUAAUUUAAUUUUAC
BDH2-3	SEQ ID NO: 73.	CCAUUAGAAAGAAUAAAAG
BDH2-4	SEQ ID NO: 74.	UAUUUUAAAUUAAUUAAAUGUUAAAUGG
BDH2-5	SEQ ID NO: 75.	AUUUCAUUUUAAACUCACAGAU
BDH2-6	SEQ ID NO: 76.	AUCCUUGCAAAGAGAAGAAAUG
BDH2-7	SEQ ID NO: 77.	UCCUUCAACUUGACUUCUUGCUGAUGGCUCAGAUCAACU
BRD2-1	SEQ ID NO: 78.	UAUUUUAUAAAAGUAAAAUGCCAAGAACCAAAGACU
BRD2-2	SEQ ID NO: 79.	UUCAAACUCCAAGAAAUACAAAUUCUCAAAACAC
BRD2-3	SEQ ID NO: 80.	UUUUCUCAAGACAAAGAAACCC
C16orf59-1	SEQ ID NO: 81.	GGGUGGAGCAGUCAAGCC
C16orf59-2	SEQ ID NO: 82.	ACUUCCCAACCCACACACACAGAC
C1orf124-1	SEQ ID NO: 83.	GUCACAUAAAAAUCAGAAGAAU
C1orf124-2	SEQ ID NO: 84.	CAAAUAUUAUCAGAGAUUGAA
C1orf124-3	SEQ ID NO: 85.	CUUGAAUUAUUGUUUUUAUUUUGACAAUC
C1orf124-4	SEQ ID NO: 86.	ACUCAAUAAUUAAAGAUUUGGGAAAUAU
C1orf124-5	SEQ ID NO: 87.	AGGCAACAUUUACCUUGAAAAU
C1orf124-6	SEQ ID NO: 88.	GAGGGCAAUCUUCAGAAUUCAG

C1orf63-1	SEQ ID NO: 89.	UAAUCAGAUUUGACAGUUGGCUUUCUGAAAGUUUU
C1orf63-10	SEQ ID NO: 90.	ACAUUUCUGGAGAAUUAUAAUAAACUUAU
C1orf63-11	SEQ ID NO: 91.	CAAUUACACAGAUUCAUUUAGAUA
C1orf63-2	SEQ ID NO: 92.	UCAGAUUUGCUACUUUGAAUUUAGCACAUUAU
C1orf63-3	SEQ ID NO: 93.	AAAUAAAGCUCAUUAAUCUCCCAUUUUCAUG
C1orf63-4	SEQ ID NO: 94.	UGAAAAUGAAAAAAAUAAAUGU
C1orf63-5	SEQ ID NO: 95.	GCUACAAACACUCUGUAAAUAGCUUAGAAAAACU
C1orf63-6	SEQ ID NO: 96.	CAUGAUUUCUAUAAGACAGAAAUAGAGCAGAUAA
C1orf63-7	SEQ ID NO: 97.	CAAUUACCAACAGAUUUUCUUCAUCAAUG
C1orf63-8	SEQ ID NO: 98.	ACAUAAACUUCAAAUUAAACCU
C1orf63-9	SEQ ID NO: 99.	GUACCUUUGCUUAGUUUAAAAAUUG
C2orf49-1	SEQ ID NO: 100.	GGAAUUUGAUAAUUUUCUAAAGG
C8orf82-1	SEQ ID NO: 101.	CGGAAGGGAGAAAGAAGGG
C8orf82-2	SEQ ID NO: 102.	CCUGGCCUCACUCAGCG
CAPRIN2-1	SEQ ID NO: 103.	UAAAGAAAUAAUGCUUACUGGU
CAPRIN2-3	SEQ ID NO: 104.	UGUGGUAAUCAAAGCAAAUAGA
CAPRIN2-4	SEQ ID NO: 105.	AUGAUUUAGAACAGCAUGAAAAAUCAAAAUA
CAPRIN2-5	SEQ ID NO: 106.	CUUAAAUUUAAAUUAAGAAAUGAG
CAPRIN2-6	SEQ ID NO: 107.	UAAAAGAAAAUGGAUUCUAAUUAAUAU
CASKIN2-1	SEQ ID NO: 108.	GCAAAGCCACAGCUGAGGGUGACAGCACG
CASKIN2-2	SEQ ID NO: 109.	CCAGCCAGAGGAGAAAAGGCA
CDCA7-1	SEQ ID NO: 110.	CACACAAAUAAAGAAAUUAGAUUU
CDCA7-3	SEQ ID NO: 111.	UUUUCUUCUUUUAUUUUCAUUCUCCAAUUUUAAA
CDCA7-4	SEQ ID NO: 112.	AAGCCAGGAAAAAGAAAUCUUUUCUAUCA
CDCA7-5	SEQ ID NO: 113.	AGAAACACAUUCAGUUUCUAC
CDCA7-6	SEQ ID NO: 114.	UCUAAAAAAAAAAAUUUUCUC
CDCA7-7	SEQ ID NO: 115.	UGCAUAAUGCAUGGCAAAAUGAGC
CEP164-1	SEQ ID NO: 116.	GCUAGAGAAGCUAUGACUCUGAGGUCAAGGAC
CEP170-1	SEQ ID NO: 117.	CUUCAUCAAAGAAUGCAAUCA
CEP170-2	SEQ ID NO: 118.	ACUUUGAGUAAAAGAAU
CEP170-3	SEQ ID NO: 119.	CUUUGCUUUCUCAAGUUUUGUAUGU
CLCN7-1	SEQ ID NO: 120.	CCAGCAGAGGCAGGCAGAGAAGGAAG
CLCN7-2	SEQ ID NO: 121.	CUGAAAUGAGAAACAGAAGACACAUAAGAGAACCC
CLCN7-3	SEQ ID NO: 122.	GCCGCGUACAUACACACAGAACAACC

CPNE1-1	SEQ ID NO: 123.	UGAGCAUCCCUUGGGCCUCAACCCUACUCACAUCAGGGAAA GGUGAAAGGGUAAACU
CPNE1-2	SEQ ID NO: 124.	AGGCUUAGAGGAAAAGGUGAGCAU
CPNE1-3	SEQ ID NO: 125.	UAUUUCAUGCUCAAGAACCCAACCA
CPNE1-IVSB-1	SEQ ID NO: 126.	CACAUCAGGGAAAGGUGAAAGGGUAAA
CPSF3L-1	SEQ ID NO: 127.	CCCACGCCGCCCGCCCG
CPSF3L-1	SEQ ID NO: 128.	UCUGAGGCCCAGGGUCCAGCUGUGGAUG
CPSF3L-2	SEQ ID NO: 129.	CAGCCAUCCAAGCACAACCACUGCU
CPSF3L-4	SEQ ID NO: 130.	CUCACUGACAGAUGUGAGCUGGAAGCUGA
CPSF3L-5	SEQ ID NO: 131.	GGGUUCUAUGUGCAGACUCAG
DCXR-1	SEQ ID NO: 132.	CAUCACUCACGAGAAUUCC
DENND4B-1	SEQ ID NO: 133.	ACAGACCAGGGAUCACCCAGA
DFFA-1	SEQ ID NO: 134.	UUUAGAUUAAUGAGAUUUUUGC
DFFA-2	SEQ ID NO: 135.	UGCAUUUUUCUUUAAAGCUAUUUUG
DFFA-3	SEQ ID NO: 136.	AAGACCCAGAAGCCAUCUCAGAAGAUUG
DFFA-4	SEQ ID NO: 137.	AUGACAGGGACAAGGACAAUGAAUCAGAAGUAG
DFFA-5	SEQ ID NO: 138.	UUUUCUUACAACACCAACAGGAAGAAGU
DFFA-6	SEQ ID NO: 139.	GUUUAUGUUACCUCUUUACACUGAAAUG
DIS3L2-1	SEQ ID NO: 140.	GGGACACAGAUGAAGGAAUGAG
DIS3L2-2	SEQ ID NO: 141.	CAAGGAAGGGAAGGUGGUGCCAGAAAGCAGG
DIS3L2-3	SEQ ID NO: 142.	AGGCUUAUGAAACACAACC
DNAJB12-1	SEQ ID NO: 143.	AGGGCCAAAGCUGCCAGGAGU
DNAJB12-2	SEQ ID NO: 144.	CUCCCUUUCUCCCCCUCCCUCCUCUGCUCA
DNAJB12-3	SEQ ID NO: 145.	CUGGAGCCAGGGAGCAGAGCG
DNAJB12-4	SEQ ID NO: 146.	CUCAGCAACAGUUUCAAGUUCCCAC
DNAJB12-5	SEQ ID NO: 147.	CCGCCACCAAGACUGCCAGCUCCCACCCACCUC
DNAJB12-6	SEQ ID NO: 148.	AGUGCCUCAGAUCCCACCAGAGG
DNAJB12-7	SEQ ID NO: 149.	GCCUGCUACCAGCAACUCUCAUUUCC
DNAJB12-8	SEQ ID NO: 150.	CACAGAGAAGAACCUUCACUGCUUCUGC
DNAJB12-9	SEQ ID NO: 151.	GAGGACACAGGCAAAGGAGGG
DNAJB12-IVSB-1	SEQ ID NO: 152.	CCAAAGCUGCCAGGAGUUGCA
DNAJB12-IVSB-2	SEQ ID NO: 153.	GCUGGAGGUCAGGCUGGG

DNAJB12-IVSB-3	SEQ ID NO: 154.	CCCUCAGCAACAGUUUCAAGUUC
DNAJB12-IVSB-4	SEQ ID NO: 155.	AAUAGUCUGCUGUGCUGGAGAAAGGG
DNAJB12-IVSB-5	SEQ ID NO: 156.	UUCCUCCUAGCUGGAGGGAUGGAGAAAG
DNAJB12-IVSB-6	SEQ ID NO: 157.	AGAGAGUGCCUCAGAUCCCACCAGA
DNAJB12-IVSB-7	SEQ ID NO: 158.	AGAAGGAGGGAGCCUGCUACCAGCAACUCUCAUUUC
DNAJB12-IVSB-8	SEQ ID NO: 159.	CACACAGAGAAGAACCUUCAC
DNAJB12-IVSB-9	SEQ ID NO: 160.	CAGCACAGAGGCAGGCACAAAUG
DPF1-1	SEQ ID NO: 161.	UCUGGAACGGGAGGGAGAGGG
DPF1-2	SEQ ID NO: 162.	CUCAGCCAGAGACCUGAGCAGC
DRG2-1	SEQ ID NO: 163.	CAAUUUCAACGAUCAGUAACAGAGC
DRG2-2	SEQ ID NO: 164.	UUCUGGAAAGCGGGAUAAUGGAC
DRG2-3	SEQ ID NO: 165.	CAUCAUAAAAGGAGUAACAGGAUAAUA
DRG2-4	SEQ ID NO: 166.	CUUAUUUCAGAAGAAAAUCCGA
DRG2-5	SEQ ID NO: 167.	CAAGCUUGGCAUUUUUCUUUAAUCCA
DSN1-1	SEQ ID NO: 168.	GUGGAAACAUAAAGAAAGCAUC
DSN1-2	SEQ ID NO: 169.	UGCAAAAAAGUGGAAAAAGUAAAUGUA
EML3-1	SEQ ID NO: 170.	AUCUUCAGGUUUCUGGACUCUCACCCA
EWSR1-1	SEQ ID NO: 171.	AAACAAAAUUAGGUAAAAGGAG
EWSR1-10	SEQ ID NO: 172.	CUUUAAACACAAAAGUUUACA
EWSR1-2	SEQ ID NO: 173.	GGAAAUGCAGAAAUUAAUUUCUUAUG
EWSR1-3	SEQ ID NO: 174.	AUUUCAAGACAACCAUUCAAAGGCAGUUAGUUAACAA
EWSR1-4	SEQ ID NO: 175.	CUAAACAAAGUUUUCUAAACCAGAUU
EWSR1-5	SEQ ID NO: 176.	GGACAGAACACACACAGAAC
EWSR1-6	SEQ ID NO: 177.	AGUUAAAAAUCAACUUUAAUUUUGAAG
EWSR1-7	SEQ ID NO: 178.	UUUUCCAAAUCAGAAGAUUG
EWSR1-8	SEQ ID NO: 179.	UAUUUUAAAACAUCCAAAAAGAAGU
EWSR1-9	SEQ ID NO: 180.	GACAAAGCAUGUUAAAAAGUUUCCA
FGFR4-1	SEQ ID NO: 181.	AUCAGAUGAGCAGCAGCGG

FTSJ1-1	SEQ ID NO: 182.	GGGUCAAGGCAGGCUGAGAG
FTSJ1-2	SEQ ID NO: 183.	CCAGAAACCAUGAGAUUUGGGUCAGAAAAAGGCA
FTSJ1-IVSB-1	SEQ ID NO: 184.	CAGUCGGCGUCCCAGAGAUCC
GBAP1-1	SEQ ID NO: 185.	CAUUUAAGUAGCAAAUUCUGGGC
GBAP1-2	SEQ ID NO: 186.	CUCAUCUUCUUCAGAGAAGU
GBAP1-3	SEQ ID NO: 187.	CCAAAGAAUUGGCAAAGAAAAG
GBAP1-4	SEQ ID NO: 188.	AUUUCACUGGCAUUAAGACAG
GBAP1-5	SEQ ID NO: 189.	GUCCGUAGCAGUUAGCAGAUGA
GBAP1-6	SEQ ID NO: 190.	GUCUGAGUCAGGGCCAAAAGGAA
GMPPA-1	SEQ ID NO: 191.	GGGAAACAGCAUGAAGAUAAGCAGG
GMPPA-2	SEQ ID NO: 192.	AUGAGAAACUAGAUUAGGG
GMPPA-3	SEQ ID NO: 193.	GAAAAGCAAUAAAGAAAUGAGCAACA
GMPPA-4	SEQ ID NO: 194.	AAGUCCAGAAACCAGUUUCAGUC
GMPR2-1	SEQ ID NO: 195.	GAGCUGGGAAAGGGUUGUGAGAGAC
GMPR2-2	SEQ ID NO: 196.	GGUCCCUGAAGCCUGUCACC
GMPR2-3	SEQ ID NO: 197.	CGCUUAAGUUGUGGAAGGUCG
GNPTG-1	SEQ ID NO: 198.	AGCACUACAGGGCCUCCAGCAGGGC
GORASP1-2	SEQ ID NO: 199.	ACAAAACCAGACACUUCUCAUGGACAGCA
GPATCH4-1	SEQ ID NO: 200.	AUCUGAAGACAUCUCUUCCCACAUU
GPATCH4-2	SEQ ID NO: 201.	CCAGUCAAGCAUUAGAUUUAGC
GPATCH4-3	SEQ ID NO: 202.	UCCUCCUUCUAAAACAUU
HGS-1	SEQ ID NO: 203.	AGGAUGCACCCCAUGCU
HMG20B-1	SEQ ID NO: 204.	CGGAGCCACAAGCAAUUCAAAUCCAGC
HMG20B-2	SEQ ID NO: 205.	GUCAGCAGUCGGGACACGGUGGGUUAGA
IFFO1-1	SEQ ID NO: 206. GAA	UGGUUAAAGAAACUGGAGAAAGAAAAGCAAAAGGAUAAAG
IFFO1-2	SEQ ID NO: 207.	CAAGUCAGGGAGAGAGAGAGAGAGG
ISYNA1-1	SEQ ID NO: 208.	AGCCGCCCCGCUCUCCCCAGC
KRI1-2	SEQ ID NO: 209.	GAGGAUGAAAGAGGAAAGG
KRI1-3	SEQ ID NO: 210.	CUGAGGGCACAAGAGAGACAG
KRI1-4	SEQ ID NO: 211.	GGGAAGACAAAGACUUGACAAGG
KRI1-5	SEQ ID NO: 212.	AGGUCAAACAGGUGGUCAAACAGCAGGA
LOC148413-1	SEQ ID NO: 213.	UAAGGACUGAAGACACGACG
LOC148413-	SEQ ID NO: 214.	GAGUGUUGAAGGCAAGACUUUGCAG

2		
LZTR1-1	SEQ ID NO: 215.	CCCACUCAGUGGGAGCUGCAGCCAU
MAN2C1-1	SEQ ID NO: 216.	GGAAGACCCAUUUCUCCAUGCC
MAP4K2-1	SEQ ID NO: 217.	CCCAGAGCUCUGAGGGUGCCCUGGGC
MCOLN1-2	SEQ ID NO: 218.	GUGCUCACCCAGCAGGCA
MCOLN1-3	SEQ ID NO: 219.	GCCACGUGCUGACUCUGCAGCUGGCAGG
MDP1-1	SEQ ID NO: 220.	UCGCCCCCAGUCUUCCCU
MIB2-1	SEQ ID NO: 221.	GGCAGCACAGCAAGAGG
MITD1-1	SEQ ID NO: 222.	CAAAAACAGUGCUACACAUUUACUCA
MOK-1	SEQ ID NO: 223.	UCAGAAAGCCUGUGACAAAUCUU
MOK-2	SEQ ID NO: 224.	AAGAAGAGUCCAAAAUGGUU
MOK-3	SEQ ID NO: 225.	UGAGAAGAAUGAAAUAAAAUUUAACAAAA
MOK-4	SEQ ID NO: 226.	UGUUAUGCUAAAAAUGUAAGAAAAC
MOV10-4	SEQ ID NO: 227.	AUCAGAAUUUCCAAGAGAGAGGCC
MOV10-5	SEQ ID NO: 228.	UAAGGAAAGAAAACAGCAUUGCAAAGAACACG
MRPL35-1	SEQ ID NO: 229.	AGUUUUAAAACUUUUCUAAGUUUAAUGU
MRPL35-5	SEQ ID NO: 230.	AAUGAAAACAUGAAAUCUGA
MRPL35-6	SEQ ID NO: 231.	GAAAAUUUGUGGGAAAAGUUUAUCCUUAC
MRPL35-7	SEQ ID NO: 232.	UCUGAAACAGUAAUUCAUGCAUAAUUCU
MRPL35-8	SEQ ID NO: 233.	UGCAGAACUUCAAUUUCAUAAUUUU
MTMR11-1	SEQ ID NO: 234.	AAACAAAUCAAGACCAAACUUCAGAGAGU
MTMR11-2	SEQ ID NO: 235.	CCUGAAAAUGAGAAUAAAUCUCC
MTMR11-3	SEQ ID NO: 236.	GACAAAUCAUGAGAUUCUCACC
MUS81-1	SEQ ID NO: 237.	UCCCUGCCACUCCCUCCA
MUS81-2	SEQ ID NO: 238.	CUGCAGGAAGAGAGGCAGCGA
NAPEPLD-1	SEQ ID NO: 239.	GCCUUUUUCAUUAAAAG
NAPEPLD-2	SEQ ID NO: 240.	UUUCAUUUGUUUUUAAACUUAGAU
NAPEPLD-3	SEQ ID NO: 241.	UAUUCAUGAAUUUCUAA
NAPEPLD-4	SEQ ID NO: 242.	UUUCCAAAUGUAAAAUAAUCACA
NAPEPLD-5	SEQ ID NO: 243.	CACAAAACAUAAAACAUAAAC
NAPEPLD-6	SEQ ID NO: 244.	UACUAGGAAGCAAGUUAUUA
NAPEPLD-7	SEQ ID NO: 245.	AAUUCAUUAUUUAAAUGAC
NAPEPLD-8	SEQ ID NO: 246.	AUGAAAUUUAAAAUCCACAUUAGC
NBEAL2-1	SEQ ID NO: 247.	ACAUUCUGAUUAGGGAGG

NDRG4-1	SEQ ID NO: 248.	GAAGGCAACAGAGGUGAGUGUGA
NDRG4-2	SEQ ID NO: 249.	CCAGAGGGCAGGCAAGGCAGAAGUG
NDUFB10-1	SEQ ID NO: 250.	GGAAGAUUUGCAAUGGUUCUG
NFATC4-1	SEQ ID NO: 251.	ACACACAGACAAAAGAGUUGCAAGAGACAGAGAC
NFATC4-2	SEQ ID NO: 252.	GGCAAACUAGAAUAGAAAGA
NFATC4-3	SEQ ID NO: 253.	CCAGAGCAGAGAGAGGGUUAAACAGG
NFATC4-4	SEQ ID NO: 254.	UCAGCAGUAGACACACAAAUAAACCAG
NFKBIB-1	SEQ ID NO: 255.	GUCGGUGCCUAAUUAUCUUCUUGGG
NFKBIB-2	SEQ ID NO: 256.	AGUUUUUCAGCCACUUCU
NFKBIB-3	SEQ ID NO: 257.	UCUUGCUGCCUAAAAUCAC
NFKBIB-4	SEQ ID NO: 258.	UGCCUUUACCCAAAUUCCUC
NFKBIB-5	SEQ ID NO: 259.	UUCAAGGUCAUUUCUACAGACCAAUUUCU
NIT1-1	SEQ ID NO: 260.	GGACACUGUCCAACAAAGAUUCUAC
NIT1-2	SEQ ID NO: 261.	CUGGCAACCCAGGGACAC
NKTR-1	SEQ ID NO: 262.	AAUAAAAUUGAGUUUAUAGAAUUA
NKTR-2	SEQ ID NO: 263.	AUUUGCCAGAUUUCAAUUUAAAGUUUAAAAAG
NKTR-3	SEQ ID NO: 264.	AAACUGAAAACACACAAAUCUUUGAAAUGAAAUGC
NKTR-6	SEQ ID NO: 265.	CUUUUUUAUUUUAAGAGUUCCA
NKTR-7	SEQ ID NO: 266.	AUGAUUUUCACAAAGAGAACAAUA
NKTR-8	SEQ ID NO: 267.	AUUUCAUAAUAAAAGCACAUAAAAUUAGU
NPRL2-1	SEQ ID NO: 268.	CCUGCCACCCACCGCUCACCC
NPRL2-2	SEQ ID NO: 269.	CCUUCCUCCUCCUGGGACAA
NSUN5P1-1	SEQ ID NO: 270.	AUUAAAGUGUCAGAACUAAGACCAAAACAGAUG
NSUN5P1-2	SEQ ID NO: 271.	CCUGAAAUCCUUGCCUCACAGAGGAGAACU
NSUN5P1-3	SEQ ID NO: 272.	GCCUCAGUCCUGAAAUCCU
NUDT22-1	SEQ ID NO: 273.	GGCAGUAAAACGUGCCAUCUUC
NUDT22-2	SEQ ID NO: 274.	UGUCGCAGACCUCCUGAGGG
PAN2-1	SEQ ID NO: 275.	UCUUCCUUUUCCCUCUGCUAAGUUU
PAN2-2	SEQ ID NO: 276.	GUGACUAUGGAAAAUCCCCUAACAG
PDDC1-1	SEQ ID NO: 277.	GUGCAGCUCUGAUGUGGCAGG
PDLIM4-1	SEQ ID NO: 278.	UGCAGGGAGUGGGAAGGCAGAU
PDLIM4-2	SEQ ID NO: 279.	GGGCCGCAGAGACCGAAGAGGGCAGGUG
PDLIM4-3	SEQ ID NO: 280.	GAAGCCAGGGCGUAGCAAGGUUGUAGCAA
PDLIM4-4	SEQ ID NO: 281.	GGGCAACCUGGGCACUGCA

PHF1-1	SEQ ID NO: 282.	UUUUUCCUUCAUUUCCUGGGAU
PHF1-2	SEQ ID NO: 283.	GUCCCAAACCCUAAACUUACCUC
PIK3CD-1	SEQ ID NO: 284.	CUGGGAUUCCCACAGAACGG
PIK3CD-2	SEQ ID NO: 285.	GCUGGAAACGUCCCCAGUGGCCUUCC
PITPNM1-1	SEQ ID NO: 286.	GGCGGAGCCCUCCCGCAGAGGC
PPIL2-1	SEQ ID NO: 287.	GCAGCAGGCAAGCAAUUUAUUG
PPIL2-2	SEQ ID NO: 288.	GCCCUUGGCAACAGGUUAAGGGA
PPIL2-3	SEQ ID NO: 289.	CAGGUCCUGGAAACAGGGCCAA
PPIL2-5	SEQ ID NO: 290.	GGGUAAGAAAACCAGACAUA
PPIL2-7	SEQ ID NO: 291.	GCACAUUUAACAGAAAAAUG
PPIL2-8	SEQ ID NO: 292.	UGAAGACGAAGAAAAAGCCAGCCAGG
PPP1R35-1	SEQ ID NO: 293.	CGCACGCGGCCGGCCGCCCGC
PPP4C-1	SEQ ID NO: 294.	CCACCCCCAAAAGCAGAAU
PPP4C-2	SEQ ID NO: 295.	CUGCCCCUCCCAGAAUGCUG
PPP4C-3	SEQ ID NO: 296.	UCUUUCACCUACCAGACACAGAC
PPP4C-4	SEQ ID NO: 297.	CCUCCAGAGAAUGUAAAGCUGA
PQLC2-1	SEQ ID NO: 298.	GGAGAGGGCUGGAAGGAUGUGGCA
PQLC2-2	SEQ ID NO: 299.	AAAAACGAAGCCAUCAGAUGCCAAG
PRPF39-1	SEQ ID NO: 300.	GUGACAAAUGCAAAUAAAUAC
PRPF39-2	SEQ ID NO: 301.	CUGCCAACAAAGAGAGAAAAUAUUAGCU
PRPF39-3	SEQ ID NO: 302.	UGUUUGGAAAAUGAGAAAUAAAUGU
PRPF39-4	SEQ ID NO: 303.	UAGCAAAUGUGACUAGCAAACCAAC
PRPF39-5	SEQ ID NO: 304.	CUAAUUACUGGAAUUUUGUUUAAAUAAUC
PSME2-1	SEQ ID NO: 305.	UUGUUAGCUAGAGAGGGUGGGCAAAGGG
PSME2-2	SEQ ID NO: 306.	CCUAAUCCACUAUUUGAAAC
PSME2-3	SEQ ID NO: 307.	CAUGCCUCACGCCAUCCUAAUG
PTPMT1-1	SEQ ID NO: 308.	GACAGGGACGGAGCGGCGG
QARS-1	SEQ ID NO: 309.	ACCUCCCUCACCCCAAACC
RAD52-1	SEQ ID NO: 310.	GGCCGCAGAGGAAAGGAGG
RAD52-2	SEQ ID NO: 311.	GCAGCCCCGUGACACAGGAG
RHOT2-1	SEQ ID NO: 312.	CACAGGCCGCGCCGCCCC
RHOT2-2	SEQ ID NO: 313.	CCAUGCUGGGCCAGAUCUGCCAGG
RMND5B-1	SEQ ID NO: 314.	CUGAGAGGUCGAAGCAGAAUGC
RMND5B-2	SEQ ID NO: 315.	GUGAAAUGAAGACCACAGUCAAGCCC

RMND5B-3	SEQ ID NO: 316.	GAGACAGGGCUGCAGGCAAGUCAAGUA
RNF123-1	SEQ ID NO: 317.	ACACACACAACCAAACACGCACAACAC
RNF123-2	SEQ ID NO: 318.	GGCAGCAGGAGCAGAAACCAG
RNF123-3	SEQ ID NO: 319.	CACAACAGUCAGCAGGUCAGACUG
RPL10A-1	SEQ ID NO: 320.	GAAGGGUCUGGGACCGCAGCA
RPP21-1	SEQ ID NO: 321.	UACAGUGAGAAAGGCGCU
RPP21-2	SEQ ID NO: 322.	AGGAACUUAAUCCAAACCCGAAGAAGGAAGAC
RPP21-3	SEQ ID NO: 323.	CCUCUUAAAAGUUAUUAUUUAUU
RPP21-5	SEQ ID NO: 324.	AAUUUCAAUGAGAAUAAUGAAU
RPP21-7	SEQ ID NO: 325.	UCUUUAAGAUAAAGUUCAAAC
RPP21-8	SEQ ID NO: 326.	CAAUUUGAAUGCACAUUUGAU
RPS6KB2-1	SEQ ID NO: 327.	CGACAGACGUGGCCAAGGCA
RPS6KB2-2	SEQ ID NO: 328.	AGACACAGCAACCGAAGCCAACACU
RPS6KB2-3	SEQ ID NO: 329.	ACACAGGCCGCGGGCUCCACAAAC
RUSC1-1	SEQ ID NO: 330.	GAGCUCCAUUACUCUCCUCAU
RUSC1-2	SEQ ID NO: 331.	CACCUCCCGCCAACCAUUCC
SCRN2-1	SEQ ID NO: 332.	UUCCUUCAUAUUUCCAGAGUC
SCRN2-2	SEQ ID NO: 333.	UCCCCAGCUCUGAAAUCUCU
SCRN2-3	SEQ ID NO: 334.	CUCACACAAGCAGGAGAAAGGAGAU
SCYL1-1	SEQ ID NO: 335.	CUAGUCUUCAGCCCACCCAG
SFR1-1	SEQ ID NO: 336.	ACAAUACUUAGAAACACAUAAUGG
SFR1-2	SEQ ID NO: 337.	CGUAGAAUUUAAACCACC
SFR1-3	SEQ ID NO: 338.	CACAUUAUGUUAAUUAACAAC
SFR1-4	SEQ ID NO: 339.	AGAAGAAAAACAAAAUUAUUUAAUAAAAU
SFR1-5	SEQ ID NO: 340.	UAACUGAAAUGAAUUCAUUCAAGAGGAAAAUAUGGAA
SFR1-6	SEQ ID NO: 341.	UCAGAAUUACAGAGUAAGGAAAAGACCU
SFR1-7	SEQ ID NO: 342.	GGCAUCACAAAAUGACUUUAAUUUCUGGA
SGSM3-1	SEQ ID NO: 343.	CUAACCCCAGAGAGGUCUCUA
SIRT7-1	SEQ ID NO: 344.	UGGAGACCCUGGGUCCCUGCAG
SLC25A3-1	SEQ ID NO: 345.	CCACAGGAGCUCUGGGCU
SLC25A3-2	SEQ ID NO: 346.	UGGGCCCACCGCCAAAGCAGCG
SLC25A3-3	SEQ ID NO: 347.	UCCACGCCCUUGAAGAGGUCACGGCGG
SLC30A7-1	SEQ ID NO: 348.	AUUUCUCUCUUUUAAAAAGCUG
SLC30A7-2	SEQ ID NO: 349.	GCACAAAAGAAAAGACCAAAAGU

SLC30A7-3	SEQ ID NO: 350.	CAGAAGUCAAAAAGAUUUGGAGGAAAG
SLC30A7-4	SEQ ID NO: 351.	AAACCUCAGAAGUCAAAAAGAU
SLC37A4-1	SEQ ID NO: 352.	UAUGACAAUCCAAACAGGC
SLC37A4-2	SEQ ID NO: 353.	UAAGAAAGGGCGCUCCCACAUGCUCUUUAGG
SLC37A4-3	SEQ ID NO: 354.	UCCUAAAAUAUCUUGACAAGCAAU
SLC37A4-4	SEQ ID NO: 355.	AAGCUCACAUUACAGGGAAGAGGGA
STK19-1	SEQ ID NO: 356.	UCAUUUUAUUAACAAGAAGAGUC
STK19-2	SEQ ID NO: 357.	ACCAAGAACUGAAUUCUAUUUCAGG
STK19-3	SEQ ID NO: 358.	GAAACACGGGCAACCAUGCAAGAGAGACU
STX10-1	SEQ ID NO: 359.	CCCACCAGGACUGACCCCUCCC
TCF25-1	SEQ ID NO: 360.	CCCUCCUGCUGCUGGAAGCAGGUCC
TCF25-2	SEQ ID NO: 361.	GUCACAGAAAUGUGAAAAUGCACC
TCF25-3	SEQ ID NO: 362.	UUCUUUAGGAAGCAGGACUGA
TOMM40-1	SEQ ID NO: 363.	GACUCAGCCCCAGCAAAUCCGC
TOMM40-2	SEQ ID NO: 364.	GCACCCGGCUCCGGCCCC
TP53I3-1	SEQ ID NO: 365.	GCAAAUCACACUCCCUCUGAGUUGGAAGC
TP53I3-2	SEQ ID NO: 366.	CCGCCUCCAGACCGAUCCCACCCGGAACACAGAUGGG
TRIM41-1	SEQ ID NO: 367.	AUACCGAAGAGAAGCAGGGAC
TRIM41-3	SEQ ID NO: 368.	CCCAGAGGGAAAAGCAAAAGCUGAGG
TRPT1-1	SEQ ID NO: 369.	GCAGACAGGCUCACGUUUCUCU
TRPT1-2	SEQ ID NO: 370.	CCCAGACAAGAACUCUCCUCAG
TSTA3-1	SEQ ID NO: 371.	GCUGGGCCUCAGCAGGA
TSTA3-3	SEQ ID NO: 372.	CUUACUGAGGCUGGCACGAAGACC
TTC14-1	SEQ ID NO: 373.	CCUUAAGUUUAAAAAUACUGA
TTC14-10	SEQ ID NO: 374.	AAAUGUUUCUAAAUUAUUCAUAAAGAUG
TTC14-2	SEQ ID NO: 375.	AAUACUUUCAUAUUUUUAUUUACUUUACCUCC
TTC14-3	SEQ ID NO: 376.	UCUUUAAUAAGAAAAUACAUGGAACACA
TTC14-4	SEQ ID NO: 377.	UAUUCUAUAUUUUAAUUCUAAGAUACUCU
TTC14-7	SEQ ID NO: 378. GAC	UGAAAGACAGACUUUUUUCAACACUACCUUAAAAACUUAA
TTC14-8	SEQ ID NO: 379.	AAGAUCUAAUUUUACUAUUAAGCAC
TTC14-9	SEQ ID NO: 380.	UAUUUGUUUCCUUUAAAGAUUUUAUAAAAGCU
TUBGCP6-1	SEQ ID NO: 381.	CCUGCCAACAGCAACUGC
TUBGCP6-2	SEQ ID NO: 382.	ACGUGCUGGGAACCAGCCAGC
TUBGCP6-3	SEQ ID NO: 383.	UCCGCCCCCAUCCACAGGAGAUG

U2AF1L4-1	SEQ ID NO: 384.	GGCUUAGGGUUAGGCUCAUCUGAGGAU
U2AF1L4-2	SEQ ID NO: 385.	CUGAAAUAACUAGAGUUCUAAGACACGA
UCK1-1	SEQ ID NO: 386.	CAGGACCUGCCGCCAGCCUCGGCCAGGCAGGCACGG
UNC45A-1	SEQ ID NO: 387.	CCACAGAAGCCCUACAGCUCC
UNC45A-2	SEQ ID NO: 388.	CGGUGCAGCGGUCCCAGAGUCC
VAMP1-1	SEQ ID NO: 389.	AGGCUUGUCCAUCAAAGAAAUC
VAMP1-10	SEQ ID NO: 390.	AGGGCGAAAGGAAAGGAAGGAUG
VAMP1-2	SEQ ID NO: 391.	AGCCCCACUUCCUCAGAACAGG
VAMP1-3	SEQ ID NO: 392.	GGAAAAGAGAAAGAGACAGGAGAAAACAAGAGGGU
VAMP1-4	SEQ ID NO: 393.	AACUUGAGAGUACAGAAAAAGCAGG
VAMP1-5	SEQ ID NO: 394.	CCAGUGGCCAGGUUUUCUAGA
VAMP1-6	SEQ ID NO: 395.	ACGAACAGAUUAGAAAUAACU
VAMP1-7	SEQ ID NO: 396.	CUGUAGAAAAUGUAAAGAAGAGAAAGC
VAMP1-8	SEQ ID NO: 397.	UAGAAUUCAGACAGGAAAGGG
VAMP1-9	SEQ ID NO: 398.	CAAACCAUGCAAAGAGGAGGAAGAGAAA
VARS-1	SEQ ID NO: 399.	CCUCCAGACCCUCAAAGC
VPS28-1	SEQ ID NO: 400.	CCGCCUGGCUGGGAGGG
WDR24-1	SEQ ID NO: 401.	AGCAGCCCCAGCCCCUGG
WDR90-1	SEQ ID NO: 402.	CCCCACCCACAGUGCCAG
WRAP53-1	SEQ ID NO: 403.	CUCAGGGAUCCGACGCAGAG
YDJC-1	SEQ ID NO: 404.	UGUUUGAAUGCGGAAGUCAUCC
YIPF3-1	SEQ ID NO: 405.	AUCCUCAGGCAGCUUUCAACC
YIPF3-2	SEQ ID NO: 406.	UGAUCUCAGCCUCACCUAG
ZCCHC18-1	SEQ ID NO: 407.	CACAGAUUUAUGAUAAUAAGAAACCAUUA
ZCCHC18-2	SEQ ID NO: 408.	CUUCUAAUUCUAGAUGACAUAG
ZCCHC18-3	SEQ ID NO: 409.	GCCGCUUCCGUUUAAUAAAAGCAUC
ZCCHC18-4	SEQ ID NO: 410.	CUGGUAGAAAGAGACUGAGC
ZCCHC8-1	SEQ ID NO: 411.	CUUAGUGGCAAGAUGCAUAAAAG
ZCCHC8-2	SEQ ID NO: 412.	UGCAAAAUUUGGAAAUUGUUUUAA
ZFAND1-1	SEQ ID NO: 413.	CACUUAAACAGAUAUACAAAGUGUGAA
ZNF131-1	SEQ ID NO: 414.	UGACAGCUGAAGUUCCACAA
ZNF131-2	SEQ ID NO: 415.	AUGGAACAAGUCCUUCACAU
ZNF300-1	SEQ ID NO: 416.	UUCAGGAAAGACAACAAAUAUAAACA
ZNF300-2	SEQ ID NO: 417.	UAUUUGACAUUUAAUUUAAUACA

ZNF300-3	SEQ ID NO: 418.	UAAUUUUCUCUGAACUUCUAAAACAGU
ZNF300-5	SEQ ID NO: 419.	CAACUAACAAAUAAUAGAAAAAUCCAA
ZNF300-6	SEQ ID NO: 420.	UUAAUUUCAUUUAUAUUAUAAAUCAUGU
ZNF300-7	SEQ ID NO: 421.	GACAGACAAGAAUGUUAAACAGAAAUA
ZNF317-1	SEQ ID NO: 422.	GAAGCUCUGCAAGAAUUCCAGCAUGCAC
ZNF317-2	SEQ ID NO: 423.	GGAAACAGAUGCUACAUAAAUC
ZNF317-3	SEQ ID NO: 424.	GAGCAAGGGCCUGAGAUUUUGCAAGCAUG
ZNF317-4	SEQ ID NO: 425.	CUUCAGAUGCAACCCUGACAAGGGACUAAU
ZNF692-1	SEQ ID NO: 426.	GCCCCUGCCCUUUCUGUCUCA
ZNF711-1	SEQ ID NO: 427.	GUUAAAACAUAGGUUAUAAAAGAAGAAC
ZNF711-2	SEQ ID NO: 428.	UAGAAGAAAGCAAAACAACAAAACU
ZNF711-3	SEQ ID NO: 429.	AGUAAACCAAAAAUAAUGG
ZNF711-4	SEQ ID NO: 430.	UUUGAGAAAAAAAUGCAAUUGACAA
ZNRD1-1	SEQ ID NO: 431.	AUUCUGUCCCAGGACCUAGGAGU
ZWINT-1	SEQ ID NO: 432.	UGCAGAGCAGCUUGUCUUUCUUCUGAGAG
ZWINT-2	SEQ ID NO: 433.	UACUCACGGCUCGUGUCUUCAGAAGCCAAGG
ZWINT-3	SEQ ID NO: 434.	CCUUCCCCACUCAGGUCAGCUGCUA

[00354] While preferred embodiments of the present invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the invention. It should be understood that various alternatives to the embodiments of the invention described herein may be employed in practicing the invention. It is intended that the following claims define the scope of the invention and that methods and structures within the scope of these claims and their equivalents be covered thereby.

[00355] The particular methods, compositions, and kits described can vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting, since the scope of the present invention will be limited only by the appended claims.

[00356] When values are provided, it can be understood that each value can be expressed as "about" a particular value or range. "About" can also include an exact amount. For example, "about 5 µL" can mean "about 5 µL" or "5 µL." Generally, the term "about" can include an amount that would be expected to be within 10% of a recited value.

[00357] The section headings used herein are for organizational purposes only and are not to be construed as limiting the subject matter described.

[00358] Unless defined otherwise, all technical and scientific terms used herein can have the same meaning as is commonly understood by one of skill in the art to which the claimed subject matter belongs. The descriptions herein are exemplary and explanatory only and are not restrictive of any subject matter claimed. In this application, the use of the singular can include the plural unless specifically stated otherwise. As used in the specification and the appended claims, the singular forms "a," "an" and "the" can include plural referents unless the context clearly dictates otherwise. In this application, the use of "or" can mean "and/or" unless stated otherwise. Furthermore, use of the term "including" as well as other forms, such as "include", "includes," and "included," is not limiting.

INCORPORATION BY REFERENCE

[00359] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference in their entirety to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

1. Smith, C.W. and Valcarcel, J. (2000) Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends Biochem. Sci.*, **25**, 381–388.
2. Wahl, M.C., Will, C.L. and Luhrmann, R. (2009) The spliceosome: design principles of a dynamic RNP machine. *Cell*, **136**, 701–718.
3. Callis, J., Fromm, M. and Walbot, V. (1987) Introns increase gene expression in cultured maize cells. *Genes Dev.*, **1**, 1183–1200.
4. Buchman, A.R. and Berg, P. (1988) Comparison of intron-dependent and intron-independent gene expression. *Mol. Cell. Biol.*, **8**, 4395–4405.
5. Le Hir, H., Nott, A. and Moore, M.J. (2003) How introns influence and enhance eukaryotic gene expression. *Trends Biochem. Sci.*, **28**, 215–220.
6. Cazzola, M. and Skoda, R.C. (2000) Translational pathophysiology: a novel molecular mechanism of human disease. *Blood*, **95**, 3280–3288.
7. Kralovicova, J. and Vorechovsky, I. (2010) Allele-dependent recognition of the 3' splice site of *INS* intron 1. *Hum. Genet.*, **128**, 383–400.

8. Kole, R., Krainer, A.R. and Altman, S. (2012) RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nat. Rev. Drug Discov.*, **11**, 125–140.

9. Aartsma-Rus, A. and van Ommen, G.J. (2007) Antisense-mediated exon skipping: a versatile tool with therapeutic and research applications. *RNA*, **13**, 1609–1624.

10. Goyenvalle, A., Seto, J.T., Davies, K.E. and Chamberlain, J. (2012) Therapeutic approaches to muscular dystrophy. *Hum. Mol. Genet.*, **20**, R69–78.

11. Hua, Y., Sahashi, K., Hung, G., Rigo, F., Passini, M.A., Bennett, C.F. and Krainer, A.R. (2010) Antisense correction of SMN2 splicing in the CNS rescues necrosis in a type III SMA mouse model. *Genes Dev.*, **24**, 1634–1644.

12. Du, L., Pollard, J.M. and Gatti, R.A. (2007) Correction of prototypic ATM splicing mutations and aberrant ATM function with antisense morpholino oligonucleotides. *Proc. Natl Acad. Sci. U.S.A.*, **104**, 6007–6012.

13. Kralovicova, J., Hwang, G., Asplund, A.C., Churbanov, A., Smith, C.I. and Vorechovsky, I. (2011) Compensatory signals associated with the activation of human GC 5' splice sites. *Nucleic Acids Res.*, **39**, 7077–7091.

14. Buratti, E., Chivers, M.C., Hwang, G. and Vorechovsky, I. (2011) DBASS3 and DBASS5: databases of aberrant 3' and 5' splice sites in human disease genes. *Nucleic Acids Res.*, **39**, D86–D91.

15. Davies, J.L., Kawaguchi, Y., Bennett, S.T., Copeman, J.B., Cordell, H.J., Pritchard, L.E., Reed, P.W., Gough, S.C., Jenkins, S.C., Palmer, S.M. *et al.* (1994) A genome-wide search for human type 1 diabetes susceptibility genes. *Nature*, **371**, 130–136.

16. Barratt, B.J., Payne, F., Lowe, C.E., Hermann, R., Healy, B.C., Harold, D., Concannon, P., Gharani, N., McCarthy, M.I., Olavesen, M.G. *et al.* (2004) Remapping the insulin gene/IDDM2 locus in type 1 diabetes. *Diabetes*, **53**, 1884–1889.

17. Zhang, L., Nakayama, M. and Eisenbarth, G.S. (2008) Insulin as an autoantigen in NOD/human diabetes. *Curr. Opin. Immunol.*, **20**, 111–118.

18. Vafiadis, P., Bennett, S.T., Todd, J.A., Nadeau, J., Grabs, R., Goodyer, C.G., Wickramasinghe, S., Colle, E. and Polychronakos, C. (1997) Insulin expression in human thymus is modulated by *INS* VNTR alleles at the IDDM2 locus. *Nat. Genet.*, **15**, 289–292.

19. Pugliese, A., Zeller, M., Fernandez, A. Jr, Zalcberg, L.J., Bartlett, R.J., Ricordi, C., Pietropaolo, M., Eisenbarth, G.S., Bennett, S.T. and Patel, D.D. (1997) The insulin gene is transcribed in the human thymus and transcription levels correlated with allelic variation at the INS VNTR-IDDM2 susceptibility locus for type 1 diabetes. *Nat. Genet.*, **15**, 293–297.

20. Stead, J.D., Hurles, M.E. and Jeffreys, A.J. (2003) Global haplotype diversity in the human insulin gene region. *Genome Res.*, **13**, 2101–2111.

21. Kralovicova, J., Gaunt, T.R., Rodriguez, S., Wood, P.J., Day, I.N.M. and Vorechovsky, I. (2006) Variants in the human insulin gene that affect pre-mRNA splicing: is -23*Hph*I a functional single nucleotide polymorphism at *IDDM2*? *Diabetes*, **55,** 260–264.

22. Ruskin, B., Zamore, P.D. and Green, M.R. (1988) A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. *Cell*, **52**, 207–219.

23. Wang, J., Shen, L., Najafi, H., Kolberg, J., Matschinsky, F.M., Urdea, M. and German,M. (1997) Regulation of insulin preRNA splicing by glucose. *Proc. Natl Acad. Sci. U.S.A.*, **94**, 4360–4365.

24. Kralovicova, J., Haixin, L. and Vorechovsky, I. (2006) Phenotypic consequences of branchpoint substitutions. *Hum. Mutat.*, **27**, 803–813.

25. Pacheco, T.R., Gomes, A.Q., Barbosa-Morais, N.L., Benes, V., Ansorge, W., Wollerton, M., Smith, C.W., Valcarcel, J. and Carmo-Fonseca, M. (2004) Diversity of vertebrate splicing factor U2AF35: identification of alternatively spliced U2AF1 mRNAS. *J. Biol. Chem.*, **279**, 27 039–27 049.

26. Booy, E.P., Meier, M., Okun, N., Novakowski, S.K., Xiong, S., Stetefeld, J. and McKenna, S.A. (2012) The RNA helicase RHAU (DHX36) unwinds a G4-quadruplex in human telomerase RNA and promotes the formation of the P1 helix template boundary. *Nucleic Acids Res.*, **40**, 4110–4124.

27. Lei, H. and Vorechovsky, I. (2005) Identification of splicing silencers and enhancers in sense *Alu*s: a role for pseudo-acceptors in splice site repression. *Mol. Cell. Biol.*, **25**, 6912–6920.

28. Buratti, E., Muro, A.F., Giombi, M., Gherbassi, D., Iaconcig, A. and Baralle, F.E. (2004) RNA folding affects the recruitment of SR proteins by mouse and human

polypurinic enhancer elements in the fibronectin EDA exon. *Mol. Cell. Biol.*, **24**, 1387–1400.

29. Nussinov, R. (1988) Conserved quartets near 5' intron junctions in primate nuclear pre-mRNA. *J. Theor. Biol.*, **133**, 73–84.
30. Sirand-Pugnet, P., Durosay, P., Brody, E. and Marie, J. (1995) An intronic (A/U)GGG repeat enhances the splicing of an alternative intron of the chicken beta-tropomyosin pre-mRNA. *Nucleic Acids Res.*, **23**, 3501–3507.
31. Kralovicova, J. and Vorechovsky, I. (2006) Position-dependent repression and promotion of DQB1 intron 3 splicing by GGGG motifs. *J. Immunol.*, **176**, 2381–2388.
32. Neidle, S. and Balasubramanian, S. (2006) *Quadruplex Nucleic Acids*. RSC Biomolecular Sciences, Cambridge, UK.
33. Bugaut, A. and Balasubramanian, S. (2012) 5'-UTR RNA G-quadruplexes: translation regulation and targeting. *Nucleic Acids Res.*, **40**, 4727–4741.
34. Millevoi, S., Moine, H. and Vagner, S. (2012) G-quadruplexes in RNA biology. *Wiley Interdiscip. Rev. RNA*, **3**, 495–507.
35. Gomez, D., Lemarteleur, T., Lacroix, L.,Mailliet, P., Mergny, J.L. and Riou, J.F. (2004) Telomerase downregulation induced by the G-quadruplex ligand 12459 in A549 cells is mediated by hTERT RNA alternative splicing. *Nucleic Acids Res.*, **32**, 371–379.
36. Didiot, M.C., Tian, Z., Schaeffer, C., Subramanian, M., Mandel, J.L. and Moine, H. (2008) The G-quartet containing FMRP binding site in FMR1 mRNA is a potent exonic splicing enhancer. *Nucleic Acids Res.*, **36**, 4902–4912.
37. Hai, Y., Cao, W., Liu, G., Hong, S.P., Elela, S.A., Klinck, R., Chu, J. and Xie, J. (2008) A G-tract element in apoptotic agents-induced alternative splicing. *Nucleic Acids Res.*, **36**, 3320–3331.
38. Marcel, V., Tran, P.L., Sagne, C., Martel-Planche, G., Vaslin, L., Teulade-Fichou, M.P., Hall,J., Mergny, J.L., Hainaut, P. and Van Dyck, E. (2011) G-quadruplex structures in TP53 intron 3: role in alternative splicing and in production of p53 mRNA isoforms. *Carcinogenesis*, **32**, 271–278.
39. Melko, M., Douguet, D., Bensaid, M., Zongaro, S., Verheggen, C., Gecz, J. and Bardoni, B. (2011) Functional characterization of the AFF (AF4/FMR2) family of

RNA-binding proteins: insights into the molecular pathology of FRAXE intellectual disability. *Hum. Mol. Genet.*, **20**, 1873–1885.

40. Balagurumoorthy, P., Brahmachari, S.K., Mohanty, D., Bansal, M. and Sasisekharan, V. (1992) Hairpin and parallel quartet structures for telomeric sequences. *Nucleic Acids Res.*, **20**, 4061–4067.
41. Derecka, K., Balkwill, G.D., Garner, T.P., Hodgman, C., Flint, A.P. and Searle, M.S. (2010) Occurrence of a quadruplex motif in a unique insert within exon C of the bovine estrogen receptor alpha gene (ESR1). *Biochemistry (Mosc.).* **49**, 7625–7633.
42. Balkwill, G.D., Derecka, K., Garner, T.P., Hodgman, C., Flint, A.P. and Searle, M.S. (2009) Repression of translation of human estrogen receptor alpha by G-quadruplex formation. *Biochemistry (Mosc.).* **48**, 11487–11495.
43. Garner, T.P., Williams, H.E., Gluszyk, K.I., Roe, S., Oldham, N.J., Stevens, M.F., Moses, J.E. and Searle, M.S. (2009) Selectivity of small molecule ligands for parallel and anti-parallel DNA G-quadruplex structures. *Org. Biomol. Chem.*, **7**, 4194–4200.
44. Thisted, T., Lyakhov, D.L. and Liebhaber, S.A. (2001) Optimized RNA targets of two closely related triple KH domain proteins, heterogeneous nuclear ribonucleoprotein K and alphaCP-2KL, suggest Distinct modes of RNA recognition. *J. Biol. Chem.*, **276**, 17484–17496.
45. Creacy, S.D., Routh, E.D., Iwamoto, F., Nagamine, Y., Akman, S.A. and Vaughn, J.P. (2008) G4 resolvase 1 binds both DNA and RNA tetramolecular quadruplex with high affinity and is the major source of tetramolecular quadruplex G4-DNA and G4-RNA resolving activity in HeLa cell lysates. *J. Biol. Chem.*, **283**, 34626–34634.
46. Pastor, T. and Pagani F. (2011) Interaction of hnRNPA1/A2 and DAZAP1 with an Alu-derived intronic splicing enhancer regulates ATM aberrant splicing. *PLoS One*, **6**, e23349.
47. Iwamoto, F., Stadler, M., Chalupnikova, K., Oakeley, E. and Nagamine, Y. (2008) Transcription-dependent nucleolar cap localization and possible nuclear function of DExH RNA helicase RHAU. *Exp. Cell Res.*, **314**, 1378–1391.
48. Singh, N.N., Hollinger, K., Bhattacharya, D. and Singh, R.N. (2010) An antisense microwalk reveals critical role of an intronic position linked to a unique long-distance interaction in pre-mRNA splicing. *RNA*, **16**, 1167–1181.

49. Wang, Y., Xiao, X., Zhang, J., Choudhury, R., Robertson, A., Li, K., Ma, M., Burge, C.B. and Wang, Z. (2012) A complex network of factors with overlapping affinities represses splicing through intronic elements. *Nat. Struct. Mol. Biol.*, **20**, 36–45.

50. Wang, Y., Ma, M., Xiao, X. and Wang, Z. (2012) Intronic splicing enhancers, cognate splicing factors and context-dependent regulation rules. *Nat. Struct. Mol. Biol.*, **19**, 1044–1052.

51. Ke, S., Shang, S., Kalachikov, S.M.,Morozova, I., Yu, L., Russo, J.J., Ju, J. and Chasin, L.A. (2011) Quantitative evaluation of all hexamers as exonic splicing elements. *Genome Res.*, **21**, doi10.1101/gr.119628.119110.

52. Culler, S.J., Hoff, K.G., Voelker, R.B., Berglund, J.A. and Smolke, C.D. (2010) Functional selection and systematic analysis of intronic splicing elements identify active sequence motifs and associated splicing factors. *Nucleic Acids Res.*, **38**, 5152–5165.

53. Collie, G.W. and Parkinson, G.N. (2011) The application of DNA and RNA G-quadruplexes to therapeutic medicines. *Chem. Soc. Rev.*, doi 10.1039/c1cs15067g.

54. Morris, M.J., Negishi, Y., Pazsint, C., Schonhoft, J.D. and Basu, S. (2010) An RNA G-quadruplex is essential for cap-independent translation initiation in human VEGF IRES. *J. Am. Chem. Soc.*, **132**, 17831–17839.

55. Wieland, M. and Hartig, J.S. (2007) RNA quadruplex-based modulation of gene expression. *Chem. Biol.*, **14**, 757–763.

56. Zhang, A.Y. and Balasubramanian, S. (2012) The kinetics and folding pathways of intramolecular g-quadruplex nucleic acids. *J. Am. Chem. Soc.*, **134**, 19297– 19308.

57. Bugaut, A., Murat, P. and Balasubramanian, S. (2012) An RNA hairpin to g-quadruplex conformational transition. *J. Am. Chem. Soc.*, **134**, 19953–19956.

58. Decorsiere, A., Cayrel, A., Vagner, S. and Millevoi, S. (2011) Essential role for the interaction between hnRNP H/F and a G quadruplex in maintaining p53 pre-mRNA 3'-end processing and function during DNA damage. *Genes Dev.*, **25**, 220– 225.

59. Samatanga, B., Dominguez, C., Jelesarov, I. and Allain , F.H. (2013) The high kinetic stability of a G-quadruplex limits hnRNP F qRRM3 binding to G-tract RNA. *Nucleic Acids Res.*, **41**, 2505–2516.

60. Lorenz, R., Bernhart, S.H., Qin, J., Honer, Zu, Siederdissen, C., Tanzer, A., Amman, F., Hofacker, I.L. and Stadler, P.F. (2013) 2D meets 4G: G-quadruplexes in RNA secondary structure prediction. *IEEE/ACM Trans. Comput. Biol. Bioinform.*

61. Beaudoin, J.D. and Perreault, J.P. (2010) 5'-UTR G-quadruplex structures acting as translational repressors. *Nucleic Acids Res.*, **38**, 7022–7036.

62. Fred, R.G., Sandberg, M., Pelletier, J. and Welsh, N. (2011) The human insulin mRNA is partly translated via a cap- and eIF4A-independent mechanism. *Biochem. Biophys. Res. Commun.*, **412**, 693–698.

63. Schwarze, U., Starman, B.J. and Byers, P.H. (1999) Redefinition of exon 7 in the *COL1A1* gene of type I collagen by an intron 8 splice-donor-site mutation in a form of osteogenesis imperfecta: influence of intron splice order on outcome of splice-site mutation. *Am. J. Hum. Genet.*, **65**, 336–344.

64. Xing, Y., Resch, A. and Lee, C. (2004) The multiassembly problem: reconstructing multiple transcript isoforms from EST fragment mixtures. *Genome Res.*, **14**, 426–441.

65. Fededa, J.P., Petrillo, E., Gelfand, M.S., Neverov, A.D., Kadener, S., Nogues, G., Pelisch, F., Baralle, F.E., Muro, A.F. and Kornblihtt, A.R. (2005) A polar mechanism coordinates different regions of alternative splicing within a single gene. *Mol. Cell*, **19**, 393–404.

66. Emerick, M.C., Parmigiani, G. and Agnew, W.S. (2007) Multivariate analysis and visualization of splicing correlations in single-gene transcriptomes. *BMC Bioinformatics*, **8**, 16.

67. Peng, T., Xue, C., Bi, J., Li, T., Wang, X., Zhang, X. and Li, Y. (2008) Functional importance of different patterns of correlation between adjacent cassette exons in human and mouse. *BMC Genomics*, **9**, 191.

68. Eddy, J., Vallur, A.C., Varma, S., Liu, H., Reinhold, W.C., Pommier, Y. and Maizels, N. (2011) G4 motifs correlate with promoter-proximal transcriptional pausing in human genes. *Nucleic Acids Res.*, **39**, 4975–4983.

69. Mayeda, A., Hayase, Y., Inoue, H., Ohtsuka, E. and Ohshima, Y. (1990) Surveying *cis*-acting sequences of pre-mRNA by adding antisense 2_-O-methyl oligoribonucleotides to a splicing reaction. *J. Biochem.*, **108**, 399–405.

70. Dominski, Z. and Kole, R. (1993) Restoration of correct splicing in thalassemic pre-mRNA by antisense oligonucleotides. *Proc. Natl Acad. Sci. U.S.A.*, **90**, 8673– 8677.

71. Baughan, T.D., Dickson, A., Osman, E.Y. and Lorson, C.L. (2009) Delivery of bifunctional RNAs that target an intronic repressor and increase SMN levels in an animal model of spinal muscular atrophy. *Hum. Mol. Genet.*, **18**, 1600–1611.

72. Cavaloc, Y., Bourgeois, C.F., Kister, L. and Stevenin, J. (1999) The splicing factors 9G8 and SRp20 transactivate splicing through different and specific enhancers. *RNA*, **5**, 468–483.

73. Hargous, Y., Hautbergue, G.M., Tintaru, A.M., Skrisovska, L., Golovanov, A.P., Stevenin, J., Lian, L.Y., Wilson, S.A. and Allain, F.H. (2006) Molecular basis of RNA recognition and TAP binding by the SR proteins SRp20 and 9G8. *EMBO J.*, **25**, 5126–5137.

74. Hunt, K.A., Mistry, V., Bockett, N.A., Ahmad, T., Ban, M., Barker, J.N., Barrett, J.C., Blackburn, H., Brand, O., Burren, O. *et al.* (2013) Negligible impact of rare autoimmune-locus coding-region variants on missing heritability. *Nature*, **498**, 232–235.

75. Aly, T.A., Ide, A., Jahromi, M.M., Barker, J.M., Fernando, M.S., Babu, S.R., Yu, L., Miao, D., Erlich, H.A., Fain, P.R. *et al.* (2006) Extreme genetic risk for type 1A diabetes. *Proc. Natl Acad. Sci. U.S.A.*, **103**, 14074–14079.

76. Yoshida, K., Sanada, M., Shiraishi, Y., Nowak, D., Nagata, Y., Yamamoto, R., Sato, Y., Sato-Otsubo, A., Kon, A., Nagasaki, M. *et al.* (2011) Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature*, **478**, 64–69.

77. Pacheco, T.R., Moita, L.F., Gomes, A.Q., Hacohen, N. and Carmo-Fonseca, M. (2006) RNA interference knockdown of hU2AF35 impairs cell cycle progression and modulates alternative splicing of Cdc25 transcripts. *Mol. Biol. Cell*, **17**, 4187– 4199.

78. Kikin, O., D_Antonio, L. and Bagga, P.S. (2006) QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res.*, **34**, W676–W682.

CLAIMS

WHAT IS CLAIMED IS:

1. A method of prevention or treatment of a disease in a subject comprising correction of intron retention in mature gene transcripts, wherein the disease is induced by defective protein expression caused by the intron retention in the gene transcripts.
2. The method according to claim 1, wherein correction of intron retention comprises administering a polynucleic acid polymer arranged to hybridize with the gene transcript.
3. The method according to claim 1 or claim 2, wherein correction of intron retention comprises administering a polynucleic acid polymer arranged to hybridize with the gene transcript and arranged to interfere with one or more of conformational transitions of canonical (stem loops); noncanonical (G quadruplex) RNA structures; interactions with trans-acting factors; and the rate of RNA–protein complex formation.
4. The method according to any preceding claim, wherein the polynucleic acid polymer is antisense to at least part of a target region of the transcript comprising or consisting of intronic splicing regulatory elements.
5. The method according to any preceding claim, wherein the polynucleic acid polymer is antisense to at least part of a target region of the transcript comprising or consisting of SEQ ID NO: 46; or a sequence having at least 95% identity with SEQ ID NO: 46.
6. The method according to any preceding claim, wherein the polynucleic acid polymer is antisense to at least part of a target region of the transcript comprising or consisting of SEQ ID NO: 3; or a sequence having at least 95% identity with SEQ ID NO: 3.
7. The method according to any preceding claim, wherein the polynucleic acid polymer is antisense to a target region of the transcript comprising or consisting of a sequence having at least 95% sequence identity to a sequence selected from any of the group comprising SEQ ID NO: 3; SEQ ID NO: 6; SEQ ID NO: 9; SEQ ID NO: 12; SEQ ID NO: 15; SEQ ID NO: 18; SEQ ID NO: 21; SEQ ID NO: 24; SEQ ID NO: 27; SEQ ID NO: 30; SEQ ID NO: 33; SEQ ID NO: 36; SEQ ID NO: 39; SEQ ID NO: 42; SEQ ID NO: 45; or combinations thereof.
8. The method according to any of claims 1 to 4, wherein the polynucleic acid polymer is antisense to a region of the transcript comprising or consisting of a sequence complementary to a sequence having at least 95% identity with any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof.

9. The method according to any of claims 1 to 7, wherein the polynucleic acid polymer comprises or consists of a polynucleic acid polymer sequence having at least 95% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 28; SEQ ID NO: 29; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 34; SEQ ID NO: 35; SEQ ID NO: 37; SEQ ID NO: 38; SEQ ID NO: 40; SEQ ID NO: 41; SEQ ID NO: 43; and SEQ ID NO: 44; or combinations thereof; and optionally wherein uracil nucleotides are substituted with thymine nucleotides.

10. The method according to any of claims 1 to 4, or 8, wherein the polynucleic acid polymer comprises or consists of a polynucleic acid polymer sequence having at least 95% identity with a sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof; and optionally wherein uracil nucleotides are substituted with thymine nucleotides.

11. The method according to any preceding claim, wherein the polynucleic acid polymer is between about 10 and about 30 nucleotides in length.

12. The method according to any preceding claim, wherein the polynucleic acid polymer is modified to increase its stability.

13. The method according to any preceding claim, wherein the polynucleic acid polymer is tethered to a protein or aptamer, wherein the protein is a splicing factor to enhance, inhibit or modulate splicing and intron removal.

14. The method according to any preceding claim, wherein the polynucleic acid polymer is conjugated to, or bound by, a delivery vehicle suitable for delivering the polynucleic acid polymer to and/or into cells or tissue.

15. The method according to any of claims 1 to 7, 9, or 11 to 14, wherein the gene transcript encodes pro-insulin.

16. The method according to any of claims 1 to 4, 8, or 10 to 14, wherein the gene transcript is transcribed from a gene or ORF selected from any of the genes or ORFs comprising ABCD4 ABCF3; ACADVL; ALKBH6; AP1G2; APEX1; ARFRP1; ATHL1; ATP1A3; ATP5D; ATP13A1; BAX; BDH2; BRD2; C1orf63; C1orf630; C1orf631; C1orf124; C2orf49; C8orf82; C16orf59; CAPRIN2; CASKIN2; CDCA7; CEP164; CEP170; CLCN7; CPNE1;

CPSF3L; DCXR; DENND4B; DFFA; DIS3L2; DNAJB12; DPF1; DRG2; DSN1; EML3; EWSR1; EWSR10; FGFR4; FTSJ1; GBAP1GMPPA; GMPR2; GNPTG; GORASP1; GPATCH4; HGS; HMG20B; IFFO1ISYNA1; KRI1; LOC148413; LZTR1; MAN2C1; MAP4K2; MCOLN1; MDP1; MIB2; MITD1; MOK; MOV10; MRPL35; MTMR11; MUS81; NAPEPLD; NBEAL2; NDRG4NDUFB10; NFATC4; NFKBIB; NIT1; NKTR; NPRL2; NSUN5P1; NUDT22; PAN2; PDDC1; PDLIM4; PHF1; PIK3CD; PITPNM1; PPIL2; PPP1R35; PPP4C; PQLC2; PRPF39; PSME2; PTPMT1; QARS; RAD52; RHOT2; RMND5B; RNF123; RPL10A; RPP21; RPS6KB2RUSC1; SCRN2; SCYL1; SFR1; SGSM3; SIRT7; SLC25A3; SLC30A7; SLC37A4; STK19; STX10; TCF25; TOMM40; TP53I3; TRIM41; TRPT1; TSTA3; TTC14; TTC140; TUBGCP6; U2AF1L4UCK1; UNC45A; VAMP1; VAMP10; VARS; VPS28; WDR24; WDR90; WRAP53; YDJC; YIPF3; ZCCHC8; ZCCHC18; ZFAND1; ZNF131; ZNF300ZNF317; ZNF692; ZNF711; ZNRD1; ZWINT; or combinations thereof.

17. The method according to any of claims 1 to 7, 9, or 10 to 15, wherein the disease is diabetes.

18. The method according to any of claims 1 to 4, 8, or 10 to 14, or 16, wherein the disease is cancer.

19. The method according to any preceding claims, wherein the method further comprises a step of determining if a disease pathology is caused by an intron retention in a gene transcript prior to treatment.

20. A method of modulating intron splicing in a cell, comprising hybridizing a polynucleic acid polymer to a region of pre-mRNA, wherein the region comprises or consists of SEQ ID NO: 46, or a region having at least 95% identity to SEQ ID NO: 46.

21. A method of modulating intron splicing in a cell, comprising hybridizing a polynucleic acid polymer to a region of pre-mRNA, wherein the region comprises or consists of SEQ ID NO: 3, or a region having at least 95% identity to SEQ ID NO: 3.

22. A method of modulating intron splicing in a cell, comprising hybridizing a polynucleic acid polymer to a region of pre-mRNA, wherein the region comprises or consists of a sequence complementary to a sequence having at least 95% identity to any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof.

23. A polynucleic acid polymer which is antisense to at least part of a region of polynucleic acid polymer comprising or consisting of SEQ ID NO: 46, or a region of polynucleic acid

polymer comprising or consisting of a sequence having at least 95% sequence identity to SEQ ID NO: 46.

24. A polynucleic acid polymer which is antisense to at least part of a region of polynucleic acid polymer comprising or consisting of SEQ ID NO: 3, or a region of polynucleic acid polymer comprising or consisting of a sequence having at least 95% sequence identity to SEQ ID NO: 3.

25. A polynucleic acid polymer which is antisense to at least part of a region of polynucleic acid polymer, wherein the region comprises or consists of a sequence complementary to any of the group of sequences comprising SEQ ID NOs: 47 to 434; or combinations thereof; or optionally a region of polynucleic acid polymer comprising or consisting of a sequence having at least 95% sequence identity to SEQ ID NOs: 47 to 434.

26. A polynucleic acid polymer comprising or consisting of a polynucleic acid polymer sequence having at least 95% identity to a sequence selected from any of the group comprising SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 28; SEQ ID NO: 29; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 34; SEQ ID NO: 35; SEQ ID NO: 37; SEQ ID NO: 38; SEQ ID NO: 40; SEQ ID NO: 41; SEQ ID NO: 43; and SEQ ID NO: 44; or combinations thereof; and optionally wherein the uracil nucleotides are substituted with thymine nucleotides.

27. A polynucleic acid polymer comprising or consisting of a polynucleic acid polymer sequence having at least 95% identity to a sequence selected from any of the group comprising SEQ ID NOs: 47 to 434; or combinations thereof; and optionally wherein the uracil nucleotides are substituted with thymine nucleotides.

28. The polynucleic acid polymer according to any of claims 23 to 27, wherein the polynucleic acid polymer is conjugated to, or bound by, a delivery vehicle suitable for delivering the polynucleic acid polymer to cells.

29. A vector comprising the polynucleic acid polymer of any of claims 23 to 27.

30. A delivery vehicle comprising, or bound to, the polynucleic acid polymer of any of claims 23 to 27.

31. A composition comprising the polynucleic acid polymer of any of claims 23 to 28, or the vector according to claim 29, or the delivery vehicle according to claim 30.

32. The composition according to claim 31, further comprising an additional biologically active molecule.

33. A method of selecting a subject for treatment, comprising determining if the subject has a disease induced by defective protein expression caused by the intron retention in gene transcripts, wherein upon positive confirmation the subject is selected for treatment with the polynucleic acid polymer according to any of claims 23 to 28, or the vector according to claim 29, or the delivery vehicle according to claim 30, or the composition according to claims 31 or 32; and optionally treating the subject.

34. Use of antisense polynucleic acid polymer to normalize gene expression by correction of retention of introns in cancer cells.

35. The method according to any preceding claims, wherein the gene transcript is an intronic splicing regulatory element.

36. The method according to any preceding claims, wherein the intronic splicing regulatory element is in between two G quadruplexes.

37. The method according to any preceding claims, wherein the intronic splicing regulatory element comprises a first CCC motif.

38. The method of claim 37, wherein the intronic splicing regulatory element further comprises a second CCC motif.

39. The method according to any preceding claims, wherein the intronic splicing regulatory element further comprises a CCCAG or AGGCC motif.

40. The method according to any preceding claims, wherein the intronic splicing regulatory element does not form a G quadruplex.

41. The method according to any preceding claims, wherein the polynucleic acid polymer specifically hybridizes with the gene transcript.

42. A pharmaceutical composition comprising:

a polynucleic acid polymer that hybridizes to a target sequence of a partially processed mRNA transcript which encodes a protein and which comprises a retained intron, wherein the target sequence is in between two G quadruplexes, wherein the polynucleic acid polymer is capable of inducing splicing out of the retained intron from the partially processed mRNA transcript; and

a pharmaceutically acceptable excipient and/or a delivery vehicle.

43. The pharmaceutical composition of claim [0014], wherein the polynucleic acid polymer hybridizes to the retained intron of the partially processed mRNA transcript.

44. A pharmaceutical composition comprising:

a polynucleic acid polymer that hybridizes to a target sequence of a partially processed mRNA transcript which encodes a protein and which comprises a retained intron, wherein the polynucleic acid polymer hybridizes to an intronic splicing regulatory element of the partially processed mRNA transcript, wherein the intronic splicing regulatory element comprises a first CCC motif, and wherein the polynucleic acid polymer is capable of inducing splicing out of the retained intron from the partially processed mRNA transcript; and

a pharmaceutically acceptable excipient and/or a delivery vehicle.

45. The pharmaceutical composition of claim [0014], wherein the intronic splicing regulatory element further comprises a second CCC motif.

46. The pharmaceutical composition of claim [0015], wherein the first CCC motif is about 3 or more nucleotide bases from the second CCC motif.

47. The pharmaceutical composition of any one of the claims [0014]-[0016], wherein the polynucleic acid polymer hybridizes to an intronic splicing regulatory element comprising a CCCAG or an AGGCC motif.

48. A pharmaceutical composition comprising:

a polynucleic acid polymer that hybridizes to a target sequence of a partially processed mRNA transcript which encodes a protein and which comprises a retained intron, wherein the polynucleic acid polymer hybridizes to a binding motif of the partially processed mRNA transcript, wherein the binding motif does not form a G quadruplex, and wherein the polynucleic acid polymer is capable of inducing splicing out of the retained intron from the partially processed mRNA transcript; and

a pharmaceutically acceptable excipient and/or a delivery vehicle.

49. The pharmaceutical composition of any one of the claims [0014]-[0016], wherein the polynucleic acid polymer comprises a pyridine nucleotide at the 3' terminal position and/or at the 5' terminal position.

50. The pharmaceutical composition of any one of the claims [0014]-[0017], wherein the polynucleic acid polymer comprises two consecutive pyridine residues at the 3' terminal position and/or at the 5' terminal position.

51. A pharmaceutical composition comprising:

a polynucleic acid polymer that hybridizes to a target sequence of a partially processed mRNA transcript which encodes a protein and which comprises a retained intron, wherein the polynucleic acid polymer hybridizes to a binding motif of the partially processed mRNA transcript, and wherein the binding motif forms a hairpin structure, wherein the polynucleic acid polymer is capable of inducing splicing out of the retained intron from the partially processed mRNA transcript; and

a pharmaceutically acceptable excipient and/or a delivery vehicle.

52. A pharmaceutical composition comprising:

a polynucleic acid polymer that hybridizes to a target sequence of a partially processed mRNA transcript which encodes a protein and which comprises a retained intron, wherein the polynucleic acid polymer induces splicing out of the retained intron from the partially processed mRNA transcript; and

a pharmaceutically acceptable excipient and/or a delivery vehicle.

53. The pharmaceutical composition of claim [0018], wherein the polynucleic acid polymer is further capable of destabilizing the hairpin structure.

54. The pharmaceutical composition of any one of the claims [0014]-53, wherein the delivery vehicle comprises a cell penetrating peptide or a peptide-based nanoparticle.

55. The pharmaceutical composition of any one of the claims [0014]-[0020], wherein the delivery vehicle is complexed with the polynucleic acid polymer by ionic bonding.

56. The pharmaceutical composition of any one of the claims [0014]-[0020], wherein the polynucleic acid polymer is between about 10 and about 50, about 10 and about 45, about 10 and about 40, about 10 and about 30, about 10 and about 25, or about 10 and about 20 nucleotides in length.

57. The pharmaceutical composition of any one of the claims [0014]-[0020], wherein the sequence of the polynucleic acid polymer is at least 60%, 70%, 80%, 90%, or 95% or is 100% complementary to a target sequence of the partially processed mRNA transcript.

58. The pharmaceutical composition of any one of the claims [0014]-[0021], wherein the sequence of the polynucleic acid polymer has 4 or less, 3 or less, 2 or less, or 1 or less mismatches to a target sequence of the partially processed mRNA transcript.

59. The pharmaceutical composition of any one of the claims [0014]-[0021], wherein the polynucleic acid polymer is modified at the nucleoside moiety or at the phosphate moiety.

60. The pharmaceutical composition of any one of the claims [0014]-[0021], wherein the polynucleic acid polymer comprises one or more artificial nucleotide bases.

61. The pharmaceutical composition of claim [0021] or [0022], wherein the one or more artificial nucleotide bases comprises 2'-O-methyl, 2'-O-methoxyethyl (2'-O-MOE), 2'-O-aminopropyl, 2'-deoxy, T-deoxy-2'-fluoro, 2'-O-aminopropyl (2'-O-AP), 2'-O-dimethylaminoethyl (2'-O-DMAOE), 2'-O-dimethylaminopropyl (2'-O-DMAP), T-O-dimethylaminoethyloxyethyl (2'-O-DMAEOE), or 2'-O-N-methylacetamido (2'-O-NMA) modified, locked nucleic acid (LNA), ethylene nucleic acid (ENA), peptide nucleic acid (PNA), 1', 5'- anhydrohexitol nucleic acids (HNA), morpholino, methylphosphonate nucleotides, thiolphosphonate nucleotides, or 2'-fluoro N3-P5'-phosphoramidites.

62. The pharmaceutical composition of claim [0022], wherein the polynucleic acid polymer is modified at the 2' hydroxyl group of the ribose moiety of the nucleoside moiety of the polynucleic acid polymer.

63. The pharmaceutical composition of claim [0022], wherein the modification at the 2' hydroxyl group is by a 2'-O-methyl, 2'-O-methoxyethyl (2'-O-MOE), 2'-O-aminopropyl, 2'-deoxy, T-deoxy-2'-fluoro, 2'-O-aminopropyl (2'-O-AP), 2'-O-dimethylaminoethyl (2'-O-DMAOE), 2'-O-dimethylaminopropyl (2'-O-DMAP), T-O- dimethylaminoethyloxyethyl (2'-O-DMAEOE), or 2'-O-N-methylacetamido (2'-O-NMA) moiety.

64. The pharmaceutical composition of claim [0022], wherein a methyl group is added to the 2' hydroxyl group of the ribose moiety to generate a 2'-O-methyl ribose moiety.

65. The pharmaceutical composition of claim [0022], wherein a methoxyethyl group is added to the 2' hydroxyl group of the ribose moiety to generate a 2'-O-methoxyethyl ribose moiety.

66. The pharmaceutical composition of claim [0022], wherein the modification at the 2' hydroxyl group is linked to a 4' carbon by a methylene group.

67. The pharmaceutical composition of any one of the claims [0021]-[0022], wherein the ribose ring is substituted with a six member morpholino ring to generate a morpholino artificial nucleotide analogue.

68. The pharmaceutical composition of any one of the claims [0021]-[0022], wherein the phosphate backbone is substituted with an oligoglycine-like moiety to generate peptide nucleic acid (PNA).

69. The pharmaceutical composition of any one of the claims [0021]-[0022], wherein the phosphate backbone is modified by a thiol group or a methyl group.

70. The pharmaceutical composition of any one of the claims [0021]-[0022], wherein the 5' terminus, 3' terminus, or a combination thereof is modified.

71. The pharmaceutical composition of any one of the claims [0021]-[0022], wherein the modification protects the polynucleic acid polymer from endogenous nucleases in the subject.

72. The pharmaceutical composition of any one of the claims [0021]-[0022], wherein the modified polynucleic acid polymer does not induce or has a reduced ability to induce RNase H cleavage of RNA.

73. The composition of any one of the claims [0021]-[0022], wherein the polynucleic acid polymer is modified to increase its stability.

74. The pharmaceutical composition of any one of the claims [0014]-[0022], wherein the polynucleic acid polymer hybridizes to a mRNA transcript comprising at least 80%, 85%, 90%, or 95% or comprising 100% sequence identity to at least 13 contiguous bases of SEQ ID NO: 46.

75. The pharmaceutical composition of any one of the claims [0014]-[0022], wherein the polynucleic acid polymer hybridizes to a mRNA transcript comprising at least 10 contiguous bases of SEQ ID NO: 46.

76. The pharmaceutical composition of any one of the claims [0014]-[0022], wherein the polynucleic acid polymer hybridizes to a mRNA transcript comprising at least 63%, 70%, 80%, 90%, or 95% sequence identity to a sequence selected from the group consisting of SEQ ID NO: 6, SEQ ID NO: 9, SEQ ID NO: 12, SEQ ID NO: 15, SEQ ID NO: 18, SEQ ID NO: 21, SEQ ID NO: 24, SEQ ID NO: 27, SEQ ID NO: 30, SEQ ID NO: 33, SEQ ID NO: 36, SEQ ID NO: 39, SEQ ID NO: 42, SEQ ID NO: 45, or combinations thereof.

77. The pharmaceutical composition of any one of the claims [0014]-[0022], wherein the polynucleic acid polymer hybridizes to a mRNA transcript comprising at least 55%, 60%, 70%, 80%, 90%, or 95% sequence identity to SEQ ID NO: 3.

78. The pharmaceutical composition of any one of the claims [0014]-[0022], wherein the polynucleic acid polymer comprises at least 63%, 70%, 80%, 90%, or 95% or 100% sequence identity to a sequence selected from the group consisting of SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 28, SEQ

ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 43, and SEQ ID NO: 44, or combinations thereof, and optionally wherein uracil nucleotides are substituted with thymine nucleotides.

79. The pharmaceutical composition of any one of the claims [0014]-[0022], wherein the polynucleic acid polymer comprises at least 55%, 60%, 70%, 80%, 90%, or 95%, or comprises 100% sequence identity to a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, or combinations thereof, and optionally wherein uracil nucleotides are substituted with thymine nucleotides.

80. The pharmaceutical composition of any one of the claims [0014]-[0022], wherein the polynucleic acid polymer hybridizes to an mRNA transcript comprising at least 80%, 85%, 90%, or 95% sequence identity to at least 13 contiguous bases of a sequence selected from the group consisting of SEQ ID NOs: 47-434, or combinations thereof.

81. The pharmaceutical composition of any one of the claims [0014]-[0023], wherein the hybridization is specific hybridization.

82. The pharmaceutical composition of any one of the claims [0014]-[0022], wherein the polynucleic acid polymer is a synthesized polynucleic acid polymer.

83. The pharmaceutical composition of any one of the claims [0014]-[0023], wherein the pharmaceutical composition comprising the polynucleic acid polymer is administered for intravenous or subcutaneous administration.

84. A composition for use in the treatment of a disease or condition in a patient in need thereof, comprising administering to the patient a pharmaceutical composition of claims [0014]-[0024].

85. The composition for use of claim [0025], wherein the disease or condition is associated with an impaired production of a protein or is characterized by a defective splicing.

86. The composition for use of claim [0025] or [0026], wherein the disease or condition is a hereditary disease.

87. The composition for use of claim [0026], wherein a subject with the hereditary disease has a genome that comprises a copy of a gene that comprises an exon that when properly transcribed into fully processed mRNA encodes the full-length functional form of the protein.

88. The composition for use of claim [0026], wherein a subject with the hereditary disease has a genome that comprises a copy of a gene that comprises a set of exons that when properly transcribed into fully processed mRNA encodes the full-length functional form of the protein.

89. The composition for use of claim [0026], wherein a subject with the hereditary disease has a genome that comprises a defective copy of the gene, which is incapable of producing a full-length functional form of the protein.

90. The composition for use of any one of the claims [0025]-[0026], wherein the disease or condition is diabetes.

91. The composition for use of any one of the claims [0025]-[0026], wherein the disease or condition is cancer.

92. The composition for use of any one of the claims [0025]-[0026], further comprising selecting a subject for treatment of a disease or condition associated with an impaired production of a protein, comprising:

a) determining if the subject has the disease or condition associated with the impaired production of the protein; and

b) administering to the subject a pharmaceutical composition of claims [0014]-[0024] if the subject has the disease or condition associated with the impaired production of the protein.

93. The composition for use of any one of the claims [0025]-[0026], further comprising selecting a subject for treatment of a disease or condition characterized by a defective splicing, comprising:

a) determining if the subject has the disease or condition characterized by the defective splicing; and

b) administering to the subject a pharmaceutical composition of claims [0014]-[0024] if the subject has the disease or condition characterized by the defective splicing.

94. A method of treating a disease or condition characterized by impaired production of a full-length functional form of a protein in a subject in need thereof, comprising:

a) administering to the subject a pharmaceutical composition comprising:

a therapeutic agent that induces an increase in splicing out of an intron in a partially processed mRNA transcript; and

a pharmaceutically acceptable excipient and/or a delivery vehicle;

wherein the subject has a pool of partially processed mRNA transcripts, which are capable of encoding copies of the full-length functional form of the protein and each of which comprise at least one retained intron that inhibits translation of the partially processed mRNA transcripts; and

b) contacting a target cell of the subject with the therapeutic agent to induce a portion of the pool of the partially processed mRNA transcripts to undergo splicing to remove the at least one retained intron from each of the partially processed mRNA transcripts in the portion, to produce fully processed mRNA transcripts, wherein the fully processed mRNA transcripts are translated to express copies of the full-length functional form of the protein, which treat the disease or condition.

95. The method of claim [0026], wherein the therapeutic agent causes activation of one or more splicing protein complexes in the cell to remove the at least one retained intron from each of the partially processed mRNA transcripts in the portion of the pool of the partially processed mRNA transcripts.

96. The method of claim [0026] or [0027], wherein the therapeutic agent inhibits a protein that regulates intron splicing activity.

97. The method of any one of the claims [0026]-[0028], wherein the therapeutic agent activates a protein that regulates intron splicing activity.

98. The method of any one of the claims [0026]-[0028], wherein the therapeutic agent binds to a protein that regulates intron splicing activity.

99. The method of any one of the claims [0026]-[0028], wherein the therapeutic agent binds to target polynucleotide sequence of the partially processed mRNA transcripts.

100. The method of any one of the claims [0026]-[0028], wherein the therapeutic agent is a polynucleic acid polymer.

101. The method of any one of the claims [0026]-[0028], wherein the therapeutic agent is a small molecule.

102. The method of any one of the claims [0026]-[0028], wherein the pharmaceutical composition is the pharmaceutical composition of claims [0014]-[0024].

103. The method of any one of the claims [0026]-[0028], wherein the impaired production of a full-length functional form of the protein comprises sub-normal production of the full-length functional form of the protein.

104. The method of any one of the claims [0026]-[0029], wherein the impaired production of a full-length functional form of the protein is due to an absence of expression of the full-length functional form of the protein or a level of expression of the full-length functional form of the protein that is sufficiently low so as to cause the disease or condition.

105. The method of any one of the claims [0026]-[0030], wherein the impaired production of a full-length functional form of the protein comprises absence of production of the protein or production of a defective form of the protein.

106. The method of claim [0030], wherein the defective form of the protein is a truncated form of the protein, a mis-folded form of the protein or a form of the protein with aberrant target binding.

107. The method of any one of the claims [0026]-[0030], wherein treating the subject results in increased expression of the full-length functional form of the protein.

108. A method of inducing processing of a partially processed mRNA transcript to remove a retained intron to produce a fully processed mRNA transcript that encodes a full-length functional form of a protein, comprising:

a) hybridizing an isolated polynucleic acid polymer to the partially processed mRNA transcript, which is capable of encoding the full-length functional form of the protein and which comprises at least one retained intron;

b) removing the at least one retained intron from the partially processed mRNA transcript to produce a fully processed mRNA transcript that encodes a full-length functional form of the protein; and

c) translating the full-length functional form of the protein from the fully processed mRNA transcript.

109. The method of claim [0030], further comprising administering a pharmaceutical composition comprising the isolated polynucleic acid polymer to a subject in need thereof.

110. The method of claim [0030] or [0031], wherein the impaired production of the full-length functional form of the protein is correlated to a disease or condition.

111. The method of any one of the claims [0030]-[0031], wherein the disease or condition is a hereditary disease.

112. The method of any one of the claims [0030]-[0032], wherein the subject with the hereditary disease has a genome that comprises a copy of a gene that comprises an exon that when

properly transcribed into fully processed mRNA encodes the full-length functional form of the protein.

113. The method of any one of the claims [0030]-[0032], wherein the subject with the hereditary disease has a genome that comprises a copy of a gene that comprises a set of exons that when properly transcribed into fully processed mRNA encodes the full-length functional form of the protein.

114. The method of any one of the claims [0030]-[0032], the subject with the hereditary disease has a genome that comprises a defective copy of the gene, which is incapable of producing a full-length functional form of the protein.

115. The method of any one of the claims [0030]-[0032], wherein the disease or condition is diabetes.

116. The method of any one of the claims [0030]-[0032], wherein the disease or condition is cancer.

117. The method of any one of the claims [0030]-[0032], wherein the polynucleic acid polymer is an anti-sense sequence.

118. The method of any one of the claims [0030]-[0032], wherein the anti-sense sequence hybridizes to a retained intron of the partially processed mRNA transcript.

119. The method of any one of the claims [0030]-[0070], wherein the anti-sense sequence hybridizes to an intronic splicing regulatory element of the partially processed mRNA transcript.

120. The method of claim [0070], wherein the intronic splicing regulatory element comprises a CCC motif.

121. The method of any one of the claims [0030]-[0070], wherein the anti-sense sequence hybridizes to a binding motif of the partially processed mRNA transcript wherein the binding motif does not form a G quadruplex.

122. The method of any one of the claims [0030]-[0074], wherein the anti-sense sequence hybridizes to a binding motif of the partially processed mRNA transcript wherein the binding motif is in between two G quadruplexes.

123. The method of any one of the claims [0030]-[0074], wherein the anti-sense sequence hybridizes to a binding motif of the partially processed mRNA transcript wherein the binding motif has a first CCC motif and a second CCC motif.

124. The method of claim [0074], wherein the first CCC motif is about 3 or more nucleotide bases from the second CCC motif.

125. The method of any one of the claims 101-[0074], wherein the sequence of the polynucleic acid polymer is at least 60%, 70%, 80%, 90%, or 95% or is 100% complementary to a target sequence of the partially processed mRNA transcript.

126. The method of any one of the claims [0030]-[0033], wherein the sequence of the polynucleic acid polymer has 4 or less, 3 or less, 2 or less, or 1 or less mismatches to a target sequence of the partially processed mRNA transcript.

127. The method of any one of the claims [0030]-[0033], wherein the polynucleic acid polymer is between about 10 and about 50, about 10 and about 45, about 10 and about 40, about 10 and about 30, about 10 and about 25, or about 10 and about 20 nucleotides in length.

128. The method of any one of the claims [0030]-[0033], wherein the subject is a eukaryote or prokaryote.

129. The method of any one of the claims [0026]-[0033], wherein the subject is a eukaryote selected from a human, mouse, rat, non-human primate, or non-primate mammal.