# REGULATION OF GENE EXPRESSION BY APTAMER-MEDIATED MODULATION OF ALTERNATIVE SPLICING

## CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to U.S. Application No. 62/110,919 filed February 2, 2015, which is incorporated herein by reference in its entirety.

## FIELD OF THE INVENTION

[0002] The invention provides a platform and methods of using the platform for the regulation of the expression of a target gene using exposure to a small molecule. The platform features a polynucleotide cassette that is placed in the target gene and includes a synthetic riboswitch positioned in the context of a 5′ intron–alternative exon–3′ intron. The riboswitch comprises an effector region and an aptamer that binds a ligand (e.g., a small molecule) and provides control of target gene expression by exposure to the ligand.

## BACKGROUND OF THE INVENTION

[0003] Splicing refers to the process by which intronic sequence is removed from the nascent pre-messenger RNA (pre-mRNA) and the exons are joined together to form the mRNA. Splice sites are junctions between exons and introns, and are defined by different consensus sequences at the 5′ and 3′ ends of the intron (i.e., the splice donor and splice acceptor sites, respectively). Alternative pre-mRNA splicing, or alternative splicing, is a widespread process occurring in most human genes containing multiple exons. It is carried out by a large multi-component structure called the spliceosome, which is a collection of small nuclear ribonucleoproteins (snRNPs) and a diverse array of auxiliary proteins. By recognizing various *cis* regulatory sequences, the spliceosome defines exon/intron boundaries, removes intronic sequences, and splices together the exons into a final translatable message (i.e., the mRNA). In the case of alternative splicing, certain exons can be included or excluded to vary the final coding message thereby changing the resulting expressed protein.

[0004] Regulation of the expression of a target gene (e.g., a therapeutic transgene) is necessary in a variety of situations. In the context of the therapeutic expression of genes, techniques that enable regulated expression of transgenes have the potential to enhance safety

by regulating the level of expression and its timing. A regulated system to control protein expression has practical and, in some cases, essential roles for safe and effective therapeutic applications.

## SUMMARY

[0001] In one aspect, there is provided a polynucleotide cassette for the regulation of the expression of a target gene comprising (a) a riboswitch and (b) an alternatively-spliced exon, flanked by a 5′ intron and a 3′ intron, wherein the riboswitch comprises (i) an effector region comprising a stem having 7 to 20 base pairs that includes the 5′ splice site sequence of the 3′ intron, and (ii) an aptamer, wherein: the 5' splice site sequence of the 3' intron comprises the sequence GTAATG or GTRAGT, wherein R is A or G; and the alternatively-spliced exon comprises a stop codon that is in-frame with the target gene when the alternatively-spliced exon is spliced into the target gene mRNA. In one example, the aptamer specifically binds a small molecule ligand.

[0002] In one example, the polynucleotide for the regulation of the expression of a target gene ("gene regulation cassette" "regulatory cassette" or "polynucleotide cassette") contains 5′ and 3′ introns that are derived from an endogenous intron from the target gene. In one example, the 5′ and 3′ introns are exogenous to the target gene. In one example, the 5′ and 3′ introns are derived from intron 2 of the human β-globin gene. In one example, the 5′ intron comprises a stop codon in-frame with the target gene. In one example, the 5′ and 3′ introns are each independently from about 50 to about 300 nucleotides in length. In one example, the 5′ and 3′ introns are each independently from about 125 to about 240 nucleotides in length. In one example, the 5′ and/or 3′ introns have been modified to include, or alter the sequence of, an intron splice enhancer, an intron splice enhancer, a 5′ splice site, a 3′ splice site, or the branch point sequence.

[0003] In one example, the effector region stem of the riboswitch is about 7 to about 20 base pairs in length. In one example, the effector region stem is 8 to 11 base pairs in length.

[0004] In one example, the alternatively-spliced exon is derived from exon 2 of the human dihydrofolate reductase gene (DHFR), mutant human Wilms tumor 1 exon 5, mouse calcium/calmodulin-dependent protein kinase II delta exon 16, or SIRT1 exon 6. In one example, the alternatively-spliced exon is the modified DHFR exon 2 from SEQ ID NO:15. In one example, the alternatively-spliced exon has been modified in one or more of the group consisting of altering the sequence of an exon splice silencer, altering the

sequence of an exon splice enhancer, adding an exon splice enhancer, and adding an exon splice donor. In one example, the alternatively-spliced exon is synthetic (i.e., not derived from a naturally-occurring exon).

[0005] In another aspect, there is provided the polynucleotide cassette of the present invention for use of regulation of the expression of the target gene in therapy. In one example, there is provided use of the polynucleotide cassette as described herein for modulating the expression of a target gene, wherein (a) the polynucleotide cassette of the present invention (as, e.g., described above and herein) is to be inserted into the target gene, (b) the target gene comprising the polynucleotide cassette is to be introduced into a cell, and (c) the cell is to be exposed to a small molecule ligand that specifically binds the aptamer in an amount effective to induce expression of the target gene.

[0006] In one example, expression of the target gene is greater than about 5-fold higher when the small molecule ligand is present than the expression levels when the small molecule ligand is absent. In one example, the expression of the target gene is greater than about 10-fold higher when the small molecule ligand is present than the expression levels when the small molecule ligand is absent.

[0007] In one example, the polynucleotide cassette is inserted into the protein coding region of the target gene. In one example, two or more of the polynucleotide cassettes are inserted into the target gene. In one example, the two or more polynucleotide cassettes comprise different aptamers that specifically bind to different small molecule ligands. In another example, the two or more polynucleotide cassettes comprise the same aptamer of different aptamers that specifically bind the same ligand.

[0008] In one example, the target gene comprising the polynucleotide cassette is incorporated in a vector for the expression of the target gene. In one example, the vector is a viral vector. In further examples, the viral vector is selected from the group consisting of adenoviral vector, adeno-associated virus vector, and lentiviral vector.

[0009] In another aspect, there is provided a vector comprising a target gene that contains a polynucleotide cassette, for use in the regulation of the expression of the target gene in therapy, wherein (a) the vector is to be introduced into the eye; and (b)

a small molecule ligand in an amount effective to induce expression of the target gene is to be provided to the mammal; wherein the polynucleotide cassette comprises (i) a riboswitch and (ii) an alternatively-spliced exon flanked by a 5′ intron and a 3′ intron, wherein the riboswitch comprises an effector region comprising a stem having 7 to 20 base pairs that includes the 5′ splice site sequence of the 3′ intron, and an aptamer that binds a small molecule ligand, wherein: i. the 5' splice site sequence of the 3' intron comprises the sequence GTAATG or GTRAGT, wherein R is A or G; and ii. the alternatively-spliced exon comprises a stop codon that is in-frame with the target gene. In one example, the ligand is a small molecule.

[0010] In one example, the vector is introduced into the eye by intraocular injection. In one example, the vector is a viral vector. In one example, the viral vector is selected from the group consisting of adenoviral vector, adeno-associated virus vector, and lentiviral vector.

[0011] In one example, the polynucleotide for the regulation of the expression of a target gene in the eye contains 5′ and 3′ introns that are derived from an endogenous intron from the target gene. In one example, the 5′ and 3′ introns are exogenous to the target gene. In one example, the 5′ and 3′ introns are derived from intron 2 of the human β-globin gene. In one example, the 5′ intron comprises a stop codon in-frame with the target gene. In one example, the 5′ and 3′ introns are each independently from about 50 to about 300 nucleotides in length. In one example, the 5′ and 3′ introns are each independently from about 125 to about 240 nucleotides in length. In one example, the 5′ and/or 3′ introns have been modified to include, or alter the sequence of, an intron splice enhancer, an intron splice enhancer, a 5′ splice site, a 3′ splice site, or the branch point sequence. In one example, the effector region stem of the riboswitch is about 7 to about 20 base pairs in length. In one example, the effector region stem is 8 to 11 base pairs in length. In one example, the alternatively-spliced exon is derived from exon 2 of the human dihydrofolate reductase gene (DHFR) mutant human Wilms tumor 1 exon 5, mouse calcium/calmodulin-dependent protein kinase II delta exon 16, or SIRT1 exon 6. In one example, the alternatively-spliced exon is the modified DHFR exon 2 from SEQ ID NO:15, a modified exon 2 from human DHFR. In one example, the alternatively-spliced exon has been modified in one or more of the group consisting of altering the sequence of an exon splice silencer, altering the sequence of an exon splice enhancer, adding an exon splice enhancer, and adding an exon splice donor. In one

example, the alternatively-spliced exon is synthetic (i.e., not derived from a naturally-occurring exon).

[0012] In one aspect, there is provided a recombinant polynucleotide comprising a target gene containing the polynucleotide cassette of the present invention. In one example, the polynucleotide cassette is located in the protein coding sequence of the target gene. In one aspect, there is provided an *in vitro* method of modulating the expression of a target gene, the method comprising: a. inserting the polynucleotide cassette of the present invention into a target gene, b. introducing the target gene comprising the polynucleotide cassette into a cell, and c. exposing the cell to a small molecule ligand that specifically binds the aptamer in an amount effective to induce expression of the target gene. In another example, there is provided use of the polynucleotide cassette of the present invention in the manufacture of a medicament for the treatment of a disease selected from the group consisting of: cystic fibrosis, hemophilia, muscular dystrophy, thalassemia, and sickle cell anemia.

[0013] In one aspect, there is provided a vector comprising a target gene that contains a polynucleotide cassette of the present invention . In one example, the vector is a viral vector. In one example, the viral vector is selected from the group consisting of adenoviral vector, adeno-associated virus vector, and lentiviral vector.

## DESCRIPTION OF THE FIGURES

[0014] **Fig. 1a.** Schematic view of the splicing construct "Con 1" with the truncated human beta-globin intron 2 (IVS2Δ) inserted in the coding sequence of the luciferase gene. The designations "Luci Exon 1" and "Luci Exon 2" refer to the division of the luciferase gene into 5′ and 3′ coding sequences. Splicing of the inserted intron sequence IVS2Δ results in full length mRNA which is translated into full-length protein.

[0015] **Fig. 1b.** The effect of intron insertion and splice site sequence on luciferase expression. Con 1 through Con 7 have different intronic splice sites (*see* Table 1). Con 1 has IVS2Δ with its original IVS2 5′ splice site and 3′ splice site ("5'ss" and "3'ss" respectively). Con 2 to Con 7 have IVS2Δ inserted but with 5'ss and 3'ss sequence differences as listed in Table 1. Con 8 has no IVS2Δ intron. Con1 through Con 3 demonstrated no effects with intron insertion on luciferase expression compared to a luciferase control with no intron inserted (Con 8). Con 4 to Con 7, with weaker splice sites, displayed reduced luciferase expression.

[0016] **Fig. 2a.** Schematic diagram of the Intron-Exon-Intron cassette with exon inclusion (I) and exclusion (II) splicing patterns depicted. The star (✦) denotes a stop codon in DHFR exon 2. When the alternative DHFR exon is included in the target gene (luciferase) mRNA by splicing (I), the resulting transcript contains an in-frame stop codon, which blocks luciferase gene expression. Only when the alternative DHFR exon containing the stop codon is excluded from the final mRNA is the target gene expressed (II).

[0017] **Fig. 2b.** The effect of inclusion or exclusion of the DHFR alternative exon on expression of the luciferase gene. A luciferase assay was performed on HEK 293 cells transfected with different luciferase reporter constructs containing an Intron-Exon-Intron cassette as shown in Fig. 2a. The DHFR exon 2 with wild-type splice sites sequences (DHFR_wt) was compared to the DHFR exon containing mutations in the 5′ ss (DHFR_wt5ssC) or with the native 5' ss replaced with the stronger Con 1 5'ss

(DHFR_Con15ss), or with the weak 5'ss from Con 4 (DHFR_Con45ss). The construct used in lane 1 in 2b and 2c is Con 1, which is shown in Fig. 1 and Example 1. Insertion of DHFR exon 2 into the luciferase mRNA caused a decrease in luciferase expression, which did not occur when the 5'ss (i.e., the splice donor) of DHFR exon 2 was mutated (compare DHFR_wt to DHFR_wt5ssC). When the 5′ ss of the DHFR exon was replaced with the stronger 5′ ss from Con 1 (construct DHFR-Con1 5′ss) the inclusion of the DHFR exon was enhanced, leading to a 545-fold lower luciferase expression compared with Con 1. When a weak 5'ss (from Con 4 in Example 1) was used to replace the wild type 5'ss reduced splicing at this site prevented DHFR exon inclusion, thereby allowing increased luciferase expression.

[0018] **Fig. 2c.** Exonic splicing enhancer (ESE) or suppressor (ESS) elements within the DHFR exon sequence influence splicing of the alternative DHRF exon and modulate expression of the target gene. Mutation of the SRp40 binding site within the DHFR exon 2 resulted in dramatic decrease in luciferase expression: 2,982-fold difference between DHFR_wtmtSRp40 expression and that of Con 1 (Fig. 2c DHFR_wtmtSRp40; Table 2). Mutation of a putative binding site for the splicing enhancer SC35, to generate a stronger SC35 binding site (Table 2, StrSC35), further decreased luciferase expression (139-fold) compared to Con 1 (Fig. 2c, DHFR_wtStrSC35), presumably due to increased efficiency of inclusion of the DHFR exon. Replacing the binding site for the splicing enhancer SC35 with that of the splicing inhibitor hnRNP A1 (Table 2, SC35hnRNPA1), led to a 4.3-fold increase in luciferase expression compared to the wild type DHFR exon 2 (Fig. 2c, DHFR_wt SC35hnRNPA1).

[0019] **Fig 3a.** Schematic diagram of the Intron-Exon-Intron cassette containing a hairpin structure at the 5'ss of the alternative DHFR exon 2. When the DHFR 5'ss is embedded within the hairpin structure the DHFR exon will not be included in the transcript, thus allowing luciferase to be expressed (x represents the DHFR exon 5'ss buried in the hairpin).

[0020] **Fig. 3b.** Sequences and structures of four different hairpins tested in the Intron-Exon-Intron cassette illustrated in Fig. 3a.

[0021] **Fig. 3c.** Effect of hairpin structure at the 5'ss of DHFR exon 2 on target gene expression. The construct containing the DHFR exon with the Con 1 5'ss sequence efficiently suppresses luciferase expression due to DHFR exon inclusion in the spliced mRNA (DHFR_Con15ss, Fig. 3c). However, embedding the 5'ss of the DHFR exon in a

hairpin structure, efficiently prevents the inclusion of the DHFR exon and allows luciferase expression (DHFR_Con15ss_HP15 Fig. 3c). A hairpin sequence with a disrupted stem does not restore luciferase expression (Fig. 3c. DHFR_Con15ss_HP15x). The DHFR_wtmtSRp40 construct (Example 2) does not express luciferase unless the 5'ss of the DHFR exon is stably sequestered in a hairpin structure (DHFR_wtmtSRp40_HP15). Destabilization of the hairpin prevents expression of luciferase, even in the context of a mutant SRp40 binding site with strong splicing activity (DHFR_wtmtSRp40_HP15x).

[0022] **Fig. 4a. and Fig. 4b.** Schematic diagram of gene regulation by the Intron-Exon-Intron regulatory cassette containing synthetic riboswitch. In the absence of aptamer/ligand binding, the aptamer sequence disrupts hairpin stem formation, leaving the DHFR exon 5'ss accessible and leading to inclusion of the DHFR exon, thus preventing translation and blocking protein expression (Fig. 4a). When aptamer/ligand binding occurs, ligand-dependent conformational changes in the aptamer stabilize stem formation, sequestering the DHFR exon 5'ss, resulting in DHFR exon exclusion and luciferase gene expression (Fig. 4b).

[0023] **Fig. 4c.** Hairpin stem and theophylline aptamer configurations with different connecting stem lengths. The stem of the theophylline aptamer was directly linked to the stem of the hairpin sequestering the DHFR exon 5'ss, generating a 20 bp synthetic stem. The stem sequence was truncated, generating a series of hairpins with different stem lengths. Shown are the stem structures for DHFR_Theo1, 12, 13 and 14 with stem lengths of 20bp, 9bp, 8bp and 7bp respectively. Theophylline is symbolized as (▲).

[0024] **Fig. 4d.** Effect of different stem lengths using the theophylline aptamer on target gene expression in the presence and absence of theophylline. Graph showing luciferase expression regulated by theophylline aptamer containing regulatory cassettes that were generated as described in Example 4 (Fig. 4c). In constructs Theo 1 through 12, with stem lengths of 20 bp down to 9 bp, the hairpin stem was of sufficient length to form a stable structure in the absence of aptamer/ligand binding. DHFR_Theo13, does not form a stable hairpin stem in the absence of ligand, thus the DHFR exon 5'ss is not hidden resulting in inclusion of DHFR exon, blocking luciferase expression. In the presence of theophylline the hairpin is stabilized and the DHFR exon 5'ss is inaccessible to splicing machinery. This results in exclusion of the DHFR exon allowing luciferase expression. In the presence of 3 mM theophylline, DHFR_Theo_13 demonstrates 43-fold induction over the un-induced baseline level of expression. DHFR_Theo_14 shows no luciferase expression either with or

without theophylline present, suggesting that this 7 bp stem is too short to form a stable hairpin even when the aptamer binds to its ligand. As a result, the DHFR exon is spliced into the transcript and luciferase expression is blocked.

[0025] **Fig. 5a.** Sequences of synthetic stems connecting xpt-guanine aptamer and DHFR exon 5'ss sequence generated by serial truncation of the hairpin stem and the aptamer P1 stem. Guanine is symbolized as (●).

[0026] **Fig. 5b.** Effect of the stem length on the ability of a riboswitch to regulate luciferase expression in response to aptamer ligand binding. Eighteen riboswitches of different stem lengths were inserted into the regulatory cassette and the constructs were transfected into HEK 293 cells which were grown in the presence or absence of 500 μM guanine. In the absence of guanine, constructs G14 through G18 demonstrate reduced luciferase expression compared to the unregulated Con 1 control. In the presence guanine, luciferase expression was restored to varying extents.

[0027] **Fig. 5c and 5d.** Further analysis of effect of the stem length on ability of riboswitch to regulate luciferase expression in response to aptamer ligand binding. Construct G11 through G18 were validated using luciferase assay. Fig. 5d shows the basal and induced level of luciferase relative to Con1. In the absence of guanine, constructs G14 through G17 demonstrate clear regulation of luciferase expression by the aptamer ligand (in this case guanine). In the absence of guanine, luciferase expression levels are low. In the presence of guanine, luciferase expression is significantly activated. Fig. 5d shows the % of Con 1 control expression achieved for these regulated constructs on induction with guanine.

[0028] **Fig. 5e.** The regulatory cassette containing xpt-G17 riboswitch allowed regulation of gene expression in response to ligand in a number of different mammalian cell types. DHFR_G17 was transfected into HepG2, AML12, RD and C2C12 cell lines, and assayed for induced luciferase expression on treatment with guanine. The fold induction of luciferase expression when the cells were grown in the presence of ligand compared to the un-induced baseline level of expression when no guanine was added to the cell culture media is shown.

[0029] **Fig. 5f.** Regulation of luciferase by the xpt-G17 containing regulatory cassette in the context of a viral vector. A construct with the luciferase gene containing the xpt-G17 regulatory cassette was transferred to an AAV vector backbone and used to transfect cells. Cells were grown in the presence and absence of guanine. The fold induction of luciferase in

the presence of the guanine is shown, 1687 fold induction of expression was seen on treatment with guanine.

[0030] **Fig. 5g.** Regulation of antibody by the regulatory cassette in response to the aptamer binding ligand. The Intron-Exon-Intron regulatory cassette with the xpt-G17 riboswitch was inserted into the leader peptide sequence of the anti-KDR antibody sequence, and the resulting construct was transfected into HEK 293 cells. As assayed by ELISA, there was an 80-fold induction of antibody protein expression upon treatment of the transfected cells with ligand, compared to untreated cells. The induced level of expression reached about 40% of control construct containing the Con 1 intron sequences.

[0031] **Fig. 5h.** Regulation of secreted erythropoietin protein (EPO) by the regulatory cassette in response to the aptamer /ligand binding. The Intron-Exon-Intron regulatory cassette with the xpt-G17 riboswitch was inserted in the murine erythropoietin (Epo) gene and the resulting construct was transfected into HEK 293 cells. Low level expression of EPO was observed in the absence of guanine, as assayed by ELISA. In the presence of guanine, 140-fold induction of EPO expression was observed.

[0032] **Fig. 6a** Structures of the different purine riboswitch stems tested in the regulatory cassette. . Purine is symbolized as ●.

[0033] **Fig. 6b–6e.** Dose responses of constructs with regulatory cassettes containing different aptamer based riboswitches (the riboswitch stems illustrated in Fig. 6a).

[0034] **Fig.6b.** Guanine aptamer containing regulatory cassettes response to guanine.

[0035] **Fig. 6c.** Guanine aptamer containing regulatory cassettes response to guanosine.

[0036] **Fig. 6d.** Guanine aptamer containing regulatory cassettes response to 2'dG.

[0037] **Fig. 6e.** Adenine aptamer containing regulatory cassettes response to adenine.

[0038] **Fig. 7a.** Induction of EGFP with the xpt-G17 containing regulatory cassette by guanine. The Intron-Exon-Intron cassette with xpt-G17 riboswitch was cloned into EGFP gene. HEK 293 cells stably transfected with the construct were treated with 500 μM guanine and assayed by flow cytometry analysis for GFP expression 6 hr after treatment. Guanine treatment resulted in increased EGFP expression, Fig 7a, (B).

[0039] **Fig. 7b.** Target gene expression responds to the presence or absence of ligand. HEK 293 cells stably transfected EGFP with Intron-Exon-Intron cassette containing xpt-G17

riboswitch were treated for 3 days with 500 μM guanine, and assayed by flow cytometry analysis every 24 hr for 3 days. Guanine containing media was washed from the cells, and they continued to be grown without guanine treatment for a further 10 days, and the EGFP expression was monitored. EGFP expression increased when guanine was present in the cell culture media. On withdrawal of guanine EGFP, expression was lost.

[0040] **Fig. 8a.** Luciferase expression regulated by two copies of the xpt-G15 containing regulatory cassette. The graph shows the guanine dose response of constructs with a single xpt- G17 or xpt-G15 containing regulatory cassette inserted into the target gene, and a construct with two copies of xpt-G15 containing regulatory cassette (xpt-G15 double).

[0041] **Fig. 8b.** EGFP expression in tissue culture cells regulated by different regulatory cassettes. One copy of the xpt-G17 containing cassette (EGFP-xpt-G17) results in low un-induced baseline expression (A), and reaches a lower induced level of expression compared to the cells containing the EGFP-xpt-G15 construct (D). One copy of the xpt-G15 containing cassette (EGFP-xpt-G15) gives higher un-induced baseline expression (B) as well as higher induced expression (E). With the construct containing two copies of the xpt-G15 containing cassettes (EGFP-xpt-G15 double), the background un-induced expression is reduced (C) without reducing the induced level of expression (F), thus the fold induction is increased. Cells were treated with guanine and imaged 24 hr after treatment.

[0042] **Fig. 8c.** Regulatory cassettes containing the xpt-G15 and xpt-G17 riboswitches respond to both guanine and guanosine. Quantification of EGFP expression (mean fluorescence intensity) was analysed by flow cytometry, and the fold induction was calculated as mean fluorescent intensity (MFI) obtained with guanine or guanosine treatment divided by MFI obtained without treatment. Treatment with guanine and guanosine gave similar levels and fold induction.

[0043] **Fig. 8d.** Luciferase expression from a construct containing one copy of the xpt-G17 containing regulatory cassette in addition to one copy of the Ydhl-A5 containing regulatory cassette. HEK 293 cells transfected with this construct were treated with either guanine or adenine, or both. The highest induction of luciferase was seen with combined use of both ligands at their highest concentrations.

[0044] **Fig. 9a and 9b.** The effects of intron truncations on luciferase expression regulated by the Intron-Exon-Intron regulatory cassette containing the xpt-G17 riboswitch.

Fig. 9a shows the fold induction, and 9b shows the percent of luciferase expression compared to Con 1.

[0045] **Fig. 9c.** Diagram of sequences deleted from the Intron-Exon-Intron regulatory construct DHFR_G17. The deleted sequence is depicted by the open bar, remaining sequence depicted by the solid bar.

[0046] **Figs. 9d and 9e.** The effect of different intron deletions, depicted in Fig 9c, on gene regulation. Sequences within the introns flanking the alternative DHFR exon modified exon splicing and relative gene regulation. For example, DHFR-G17_2IR_3 shows intron deletions that result in a significant increase in the fold of target gene expression. Fig. 9d shows fold induction, whereas Fig. 9e shows absolute level of protein expression relative to Con 1 control.

[0047] **Figs. 10a and 10b.** Different exons can act as the alternative exon in the Intron-Exon-Intron regulatory cassette to regulate gene expression. Fig. 10a shows that constructs with different exons have various un-induced baseline and induced (500 μM guanine) levels of luciferase expression. Fig. 10b shows the induction fold with these constructs, with CaMKIId-e16 generating equivalent fold induction to the DHFR exon with SRp40 activating mutation (mtDHFR).

[0048] **Figs. 11a-c.** Regulation of luciferase expression *in vivo* in mice. The construct containing two copies of the xpt-G15 regulatory cassette (xpt-G15 double, Example 8, Fig. 8a) was delivered to the liver of mice by hydrodynamic injection. Mice were dosed orally with different doses of guanosine at 2 hr and 12 hr after DNA delivery, and then were imaged. Oral dosing of the ligand resulted in dose related activation of expression of the regulated target gene in the liver of the mice (Fig. 11a and Fig. 11b).

[0049] In another experiment, guanosine was administered intra-peritoneally (Fig. 11c). .Images show luciferase expression before and after guanosine treatment with either 100 mg/kg or 300 mg/kg dose. In the graph (Fig. 11d), the luciferase activity was expressed as mean photon/sec/$mm^2$ ± s.d. (n=5).

[0050] **Fig. 12.** EGFP transgene expression mediated by riboswitch-based AAV vectors in the murine retina. Fluorescent fundus photography showing EGFP transgene expression in the retina mediated by AAV2/8-GTX7, 8 days post subretinal injection (exposure time: 30s).

[0051] **Figs. 13a and 13b.** Representative fundus images of a single murine eye subretinally injected with AAV2/8-GTX7 showing variation in EGFP transgene expression in the retina over time. A-E: Images taken under white light illumination with an exposure time of 200ms at 2, 8, 9, 10 and 12 days post vector injection. Circle shows area of retina visible through the pupil that was taken as the ROI for quantification. F-J: Images taken under 475±25nm light illumination with an exposure time of 30s. showing eGFP fluorescence at 2, 8, 9, 10 and 12 days post vector injection. K-O: Images taken under 475±25nm light illumination with an exposure time of 30s at 2, 8, 9, 10 and 12 days post vector injection highlighting pixels above an intensity threshold of 50 within the ROI (circle). Fig. 13b shows high-resolution images pre (A) and post (B) induction.

[0052] **Fig. 13c.** EGFP transgene expression in the murine retina quantified over time after subretinal injection of AAV2/8-GTX7. Fluorescent fundus photographs were taken at the following time points: 2, 8, 9, 10 and 12 days post subretinal injection of AAV2/8-GTX7. Exposure time: 30s, pixel intensity threshold for analysis: 50. Intraperitoneal induction was carried out after imaging at 8, 9 and 10 days post subretinal injection of AAV2/8-GTX7. In addition, intravitreal induction was carried out at 11 days post subretinal injection of AAV2/8-GTX7. Statistical significance shown based upon 1-way ANOVA with Dunnetts correction and 8 days post injection as the control point.

[0053] **Fig. 13d.** EGFP transgene expression in the murine retina quantified over time post subretinal injection of AAV2/8-GTX5 (positive control). Fluorescent fundus photographs were taken at the following time points: 2, 8, 9, 10 and 12 days post subretinal injection of AAV2/8-GTX5. Exposure time: 10s, pixel intensity threshold for analysis: 190. Intraperitoneal administration of guanosine was carried out after imaging at 8, 9 and 10 days post subretinal injection of AAV2/8-GTX5. In addition, intravitreal administration of guanosine was carried out at 11 days post subretinal injection of AAV2/8-GTX5. One-way ANOVA with Bonferroni correction was applied, and no statistically significant differences in expression of EGFP were found on treatment with guanosine.

## DETAILED DESCRIPTION

[0054] The present invention provides a gene regulation cassette that comprises a riboswitch in the context of a 5′ intron–alternative exon–3′ intron. The gene regulation cassette refers to a recombinant DNA construct that when incorporated into the DNA of a

target gene provides the ability to regulate expression of the target gene by aptamer/ligand mediated alternative splicing of the resulting pre-mRNA. The riboswitch in the context of the present invention contains a sensor region (e.g., an aptamer) and an effector region that together are responsible for sensing the presence of a small molecule ligand and altering splicing to an alternative exon. In one embodiment, the target gene's expression is increased when the aptamer ligand is present and decreased when the ligand is absent.

[0055] **Riboswitch**

[0056] The term "riboswitch" as used herein refers to a regulatory segment of a RNA polynucleotide. A riboswitch in the context of the present invention contains a sensor region (e.g., an aptamer) and an effector region that together are responsible for sensing the presence of a ligand (e.g., a small molecule) and altering splicing to an alternative exon. In one embodiment, the riboswitch is recombinant, utilizing polynucleotides from two or more sources. The term "synthetic" as used herein in the context of a riboswitch refers to a riboswitch that is not naturally occurring. In one embodiment, the sensor and effector regions are joined by a polynucleotide linker. In one embodiment, the polynucleotide linker forms a RNA stem (i.e., a region of the RNA polynuceotide that is double-stranded).

[0057] **Effector region**

[0058] In one embodiment, the effector region comprises the 5′ splice site ("5'ss") sequence of the 3′ intron (i.e., the intronic splice site sequence that is immediately 3′ of the alternative exon). The effector region comprises the 5′ ss sequence of the 3′ intron and sequence complimentary to the 5′ ss sequence of the 3′ intron. When the aptamer binds its ligand, the effector region forms a stem and thus prevents splicing to the splice donor site at the 3′ end of the alternative exon (*see, e.g.*, Fig. 4b). Under certain conditions (for example, when the aptamer is not bound to its ligand), the effector region is in a context that provides access to the splice donor site at the 3′ end of the alternative exon leading to inclusion of the alternative exon in the target gene mRNA (*see, e.g.*, Fig. 4a).

[0059] The stem portion of the effector region should be of a sufficient length (and GC content) to substantially prevent alternative splicing of the alternative exon upon ligand binding the aptamer, while also allowing access to the splice site when the ligand is not present in sufficient quantities. In embodiments of the invention, the stem portion of the effector region comprises stem sequence in addition to the 5′ ss sequence of the 3′ intron and its complementary sequence. In embodiments of the invention, this additional stem sequence

comprises sequence from the aptamer stem. The length and sequence of the stem portion can be modified using known techniques in order to identify stems that allow acceptable background expression of the target gene when no ligand is present and acceptable expression levels of the target gene when the ligand is present (*see, e.g.*, Examples 4 and 5 and Figs. 4c and 4d, 5a, 5b, 5c, 5d). If the stem is, for example, too long it may hide access to the 5′ ss sequence of the 3′ intron in the presence or absence of ligand. If the stem is too short, it may not form a stable stem capable of sequestering the 5′ ss sequence of the 3′ intron, in which case the alternative exon will be spliced into the target gene message in the presence or absence of ligand. In one embodiment, the total length of the effector region stem is between about 7 base pairs to about 20 base pairs. In some embodiments, the length of the stem is between about 8 base pairs to about 11 base pairs. In some embodiments, the length of the stem is 8 base pairs to 11 base pairs. In addition to the length of the stem, the GC base pair content of the stem can be altered to modify the stability of the stem.

[0060] **Aptamer/Ligand**

[0061] The term "aptamer" as used herein refers to an RNA polynucleotide that specifically binds to a ligand. The term "ligand" refers to a molecule that is specifically bound by the aptamer. In one embodiment, the ligand is a low molecular weight (less than about 1,000 Daltons) molecule including, for example, lipids, monosaccharides, second messengers, other natural products and metabolites, nucleic acids, as well as most therapeutic drugs In one embodiment the ligand is a polynucleotide with 2 or more nucleotide bases.

[0062] Aptamers have binding regions, which are capable of forming complexes with an intended target molecule (i.e., the ligand). The specificity of the binding can be defined in terms of the comparative dissociation constants (Kd) of the aptamer for its ligand as compared to the dissociation constant of the aptamer for unrelated molecules. Thus, the ligand is a molecule that binds to the aptamer with greater affinity than to unrelated material. Typically, the Kd for the aptamer with respect to its ligand will be at least about 10-fold less than the Kd for the aptamer with unrelated molecules. In other embodiments, the Kd will be at least about 20-fold less, at least about 50-fold less, at least about 100-fold less, and at least about 200-fold less. An aptamer will typically be between about 15 and about 200 nucleotides in length. More commonly, an aptamer will be between about 30 and about 100 nucleotides in length.

[0063] The aptamers that can be incorporated as part of the riboswitch can be a naturally occurring aptamer, or modifications thereof, or aptamers that are designed *de novo* or synthetic screened through systemic evolution of ligands by exponential enrichment (SELEX). Examples of aptamers that bind small molecule ligands include, but are not limited to theophylline, dopamine, sulforhodamine B, and cellobiose kanamycin A, lividomycin, tobramycin, neomycin B, viomycin, chloramphenicol, streptomycin, cytokines, cell surface molecules, and metabolites. For a review of aptamers that recognize small molecules, *see, e.g.,* Famulok, Science 9:324-9 (1999) and McKeague, M. & DeRosa, M.C. J. Nuc. Aci. 2012. In another embodiment, the aptamer is a complementary polynucleotide.

[0064] In one embodiment, the aptamer is designed to bind a particular small molecule ligand. Methods for designing aptamers include for example SELEX. Methods for designing aptamers that selectively bind a small molecule using SELEX are disclosed in, e.g., U.S. Patent Nos. 5,475,096, 5,270,163, and Abdullah Ozer, et al. Nuc. Aci. 2014, which are incorporated herein by reference. Modifications of the SELEX process are described in U.S. Patent Nos. 5,580,737 and 5,567,588, which are incorporated herein by reference.

[0065] Selection techniques for identifying aptamers generally involve preparing a large pool of DNA or RNA molecules of the desired length that contain a region that is randomized or mutagenized. For example, an oligonucleotide pool for aptamer selection might contain a region of 20-100 randomized nucleotides flanked by regions of defined sequence that are about 15-25 nucleotides long and useful for the binding of PCR primers. The oligonucleotide pool is amplified using standard PCR techniques, or other means that allow amplification of selected nucleic acid sequences. The DNA pool may be transcribed *in vitro* to produce a pool of RNA transcripts when an RNA aptamer is desired. The pool of RNA or DNA oligonucleotides is then subjected to a selection based on their ability to bind specifically to the desired ligand. Selection techniques include, for example, affinity chromatography, although any protocol which will allow selection of nucleic acids based on their ability to bind specifically to another molecule may be used. Selection techniques for identifying aptamers that bind small molecules and function within a cell may involve cell based screening methods. In the case of affinity chromatography, the oligonucleotides are contacted with the target ligand that has been immobilized on a substrate in a column or on magnetic beads. The oligonucleotide is preferably selected for ligand binding in the presence of salt concentrations, temperatures, and other conditions which mimic normal physiological conditions. Oligonucleotides in the pool that bind to the ligand are retained on the column or

bead, and nonbinding sequences are washed away. The oligonucleotides that bind the ligand are then amplified (after reverse transcription if RNA transcripts were utilized) by PCR (usually after elution). The selection process is repeated on the selected sequences for a total of about three to ten iterative rounds of the selection procedure. The resulting oligonucleotides are then amplified, cloned, and sequenced using standard procedures to identify the sequences of the oligonucleotides that are capable of binding the target ligand. Once an aptamer sequence has been identified, the aptamer may be further optimized by performing additional rounds of selection starting from a pool of oligonucleotides comprising a mutagenized aptamer sequence.

[0066] *In vivo* aptamer screening may be used following one or more rounds of in vitro selection (e.g., SELEX). For example, Konig, J. et al. (RNA. 2007, 13(4):614–622, incorporated herein by reference) describe combining SELEX and a yeast three-hybrid system for *in vivo* selection of aptamer.

[0067] **The alternative exon**

[0068] The alternative exon that is part of the gene regulation polynucleotide cassette of the present invention can be any polynucleotide sequence capable of being transcribed to a pre-mRNA and alternatively spliced into the mRNA of the target gene. The alternative exon that is part of the gene regulation cassette of the present invention contains at least one sequence that inhibits translation such that when the alternative exon is included in the target gene mRNA, expression of the target gene from that mRNA is prevented or reduced. In a preferred embodiment, the alternative exon contains a stop codon (TGA, TAA, TAG) that is in frame with the target gene when the alternative exon is included in the target gene mRNA by splicing. In embodiments, the alternative exon comprises, in addition to a stop codon, or as an alternative to a stop codon, other sequence that reduces or substantially prevents translation when the alternative exon is incorporated by splicing into the target gene mRNA including, e.g., a microRNA binding site, which leads to degradation of the mRNA. In one embodiment, the alternative exon comprises a miRNA binding sequence that results in degradation of the mRNA. In one embodiment, the alternative exon encodes a polypeptide sequence which reduces the stability of the protein containing this polypeptide sequence. In one embodiment, the alternative exon encodes a polypeptide sequence which directs the protein containing this polypeptide sequence for degradation.

[0069] The basal or background level of splicing of the alternative exon can be optimized by altering exon splice enhancer (ESE) sequences and exon splice suppressor (ESS) sequences and/or by introducing ESE or ESS sequences into the alternative exon. Such changes to the sequence of the alternative exon can be accomplished using methods known in the art, including, but not limited to site directed mutagenesis. Alternatively, oligonucleotides of a desired sequence (e.g., comprising all or part of the alternative exon) can be obtained from commercial sources and cloned into the gene regulation cassette. Identification of ESS and ESE sequences can be accomplished by methods known in the art, including, for example using ESEfinder 3.0 (Cartegni, L. et al. ESEfinder: a web resource to identify exonic splicing enhancers. Nucleic Acid Research, 2003, 31(13): 3568-3571) and/or other available resources.

[0070] In one embodiment, the alternative exon is exogenous to the target gene, although it may be derived from a sequence originating from the organism where the target gene will be expressed. In one embodiment the alternative exon is a synthetic sequence (see Example 10).

[0071] In one embodiment, the alternative exon is a naturally-occurring exon (see Example 10). In another embodiment, the alternative exon is derived from all or part of a known exon (see Example 10). In this context, "derived" refers to the alternative exon containing sequence that is substantially homologous to a naturally occurring exon, or a portion thereof, but may contain various mutations, for example, to introduce a stop codon that will be in frame with the target gene sequence, or to introduce or delete an exon splice enhancer, and/or introduce delete an exon splice suppressor. In one embodiment, the alternative exon is derived from exon 2 of the human dihydrofolate reductase gene (DHFR), mutant human Wilms tumor 1 exon 5, mouse calcium/calmodulin-dependent protein kinase II delta exon 16, or SIRT1 exon 6.

[0072] **5′ and 3′ intronic sequences**

[0073] The alternative exon is flanked by 5′ and 3′ intronic sequences. The 5′ and 3′ intronic sequences that can be used in the gene regulation cassette of the present invention can be any sequence that can be spliced out of the target gene creating either the target gene mRNA or the target gene comprising the alternative exon in the mRNA, depending upon the presence or absence of a ligand that binds the aptamer. The 5′ and 3′ introns each has the sequences necessary for splicing to occur, i.e., splice donor, splice acceptor and branch point

sequences. In one embodiment, the 5′ and 3′ intronic sequences of the gene regulation cassette are derived from one or more naturally occurring introns or a portion thereof. In one embodiment, the 5′ and 3′ intronic sequences are derived from a truncated human beta-globin intron 2 (IVS2Δ). In other embodiements the 5′ and 3′ intronic sequences are derived from the SV40 mRNA intron (used in pCMV-LacZ vector from Clonetech), intron 6 of human triose phosphate isomerase (TPI) gene (Nott Ajit, et al. RNA. 2003, 9:6070617), or an intron from human factor IX (Sumiko Kurachi et al. J. Bio. Chem. 1995, 270(10), 5276), the target gene's own endogenous intron, or any genomic fragment or synthetic introns (Yi Lai, et al. Hum Gene Ther. 2006:17(10):1036) that contain elements that are sufficient for regulated splicing (Thomas A. Cooper, Methods 2005 (37):331).

[0074] In one embodiment, the alternative exon and riboswitch of the present invention are engineered to be in an endogenous intron of a target gene. That is, the intron (or substantially similar intronic sequence) naturally occurs at that position of the target gene. In this case, the intronic sequence immediately upstream of the alternative exon is referred to as the 5′ intron or 5′ intronic sequence, and the intronic sequence immediately downstream of the alternative exon is referred to as the 3′ intron or 3′ intronic sequence. In this case, the endogenous intron is modified to contain a splice acceptor sequence and splice donor sequence flanking the 5′ and 3′ ends of the alternative exon.

[0075] The splice donor and splice acceptor sites in the gene regulation cassette of the present invention can be modified to be strengthened or weakened. That is, the splice sites can be modified to be closer to the consensus for a splice donor or acceptor by standard cloning methods, site directed mutagenesis, and the like. Splice sites that are more similar to the splice consensus tend to promote splicing and are thus strengthened. Splice sites that are less similar to the splice consensus tend to hinder splicing and are thus weakened. The consensus for the splice donor of the most common class of introns (U2) is A/C A G || G T A/G A G T (where || denotes the exon/intron boundary). The consensus for the splice acceptor is C A G || G (where || denotes the exon/intron boundary). The frequency of particular nucleotides at the splice donor and acceptor sites are described in the art (*see, e.g.*, Zhang, M.Q., Hum Mol Genet. 1988. 7(5):919-932). The strength of 5′ss and 3′ splice sites can be adjusted to modulate splicing of the alternative exon.

[0076] Additional modifications to 5′ and 3′ introns in the gene regulation cassette can be made to modulate splicing including modifying, deleting, and/or adding intronic

splicing enhancer elements and/or intronic splicing suppressor elements, and/or modifying the branch site sequence.

[0077] In one embodiment, the 5′ intron has been modified to contain a stop codon that will be in frame with the target gene. The 5′ and 3′ intronic sequences can also be modified to remove cryptic slice sites, which can be identified with publicly available software (*see, e.g.*, Kapustin, Y. et al. Nucl. Acids Res. 2011. 1-8). The lengths of the 5′ and 3′ intronic sequences can be adjusted in order to, for example, meet the size requirements for viral expression constructs. In one embodiment, the 5′ and 3′ intronic sequences are independently from about 50 to about 300 nucleotides in length. In one embodiment, the 5′ and 3′ intronic sequences are independently from about 125 to about 240 nucleotides in length.

[0078] **Target genes**

[0079] The gene regulation cassette of the present invention is a platform that can be used to regulate the expression of any target gene that can be expressed in a target cell, tissue or organism. The term "target gene" refers to a polynucleotide that is introduced into a cell and is capable of being transcribed into RNA and translated and/or expressed under appropriate conditions. Alternatively, the target gene is endogenous to the target cell and the gene regulation cassette of the present invention is positioned into the target gene (for example into an existing intron of the endogenous target gene). An example of a target gene is a polynucleotide encoding a therapeutic polypeptide. In one embodiment, when the target gene is expressed using the gene regulation cassette of the present invention, the target gene is not expressed as a fusion protein comprising the alternative exon. Inclusion of the alternative exon minimizes translation of the mRNA by, e.g., causing degradation of the message containing the alternative exon, or otherwise prevents expression of a functional target gene due, e.g., to its premature truncation. In one embodiment, the target gene is exogenous to the cell in which the recombinant DNA construct is to be transcribed. In another embodiment, the target gene is endogenous to the cell in which the recombinant DNA construct is to be transcribed. The alternative exon, in one embodiment, may contain a stop codon in frame with the coding sequence of the target gene. In other embodiments, the alternative exon may contain other sequences that drive transcript degradation and/or block translation of the target gene.

[0080] The target gene according to the present invention may be a gene encoding a protein, or a sequence encoding a non-protein coding RNA. The target gene may be, for example, a gene encoding a structural protein, an enzyme, a cell signaling protein, a mitochondrial protein, a zinc finger protein, a hormone, a transport protein, a growth factor, a cytokine, an intracellular protein, an extracellular protein, a transmembrane protein, a cytoplasmic protein, a nuclear protein, a receptor molecule, an RNA binding protein, a DNA binding protein, a transcription factor, translational machinery, a channel protein, a motor protein, a cell adhesion molecule, a mitochondrial protein, a metabolic enzyme, a kinase, a phosphatase, exchange factors, a chaperone protein, and modulators of any of these. In embodiments, the target gene encodes erythropoietin (Epo), human growth hormone (hGH), transcription activator-like effector nucleases (TALEN), human insulin, CRISPR associated protein 9 (cas9), or an immunoglobulin (or portion thereof), including, e.g., a therapeutic antibody.

[0081] **Expression Constructs**

[0082] The present invention contemplates the use of a recombinant vector for introduction into target cells a polynucleotide encoding a target gene and containing the gene regulation cassette of the present invention. In many embodiments, the recombinant DNA construct of this invention includes additional DNA elements including DNA segments that provide for the replication of the DNA in a host cell and expression of the target gene in that cell at appropriate levels. The ordinarily skilled artisan appreciates that expression control sequences (promoters, enhancers, and the like) are selected based on their ability to promote expression of the target gene in the target cell. "Vector" means a recombinant plasmid, yeast artificial chromosome (YAC), mini chromosome, DNA mini-circle or virus (including virus derived sequences) that comprises a polynucleotide to be delivered into a host cell, either *in vitro* or *in vivo*. In one embodiment, the recombinant vector is a viral vector or a combination of multiple viral vectors.

[0083] Viral vectors for the expression of a target gene in a target cell, tissue, or organism are known in the art and include adenoviral (AV) vectors, adeno-associated virus (AAV) vectors, retroviral and lentiviral vectors, and Herpes simplex type 1 (HSV1) vectors.

[0084] Adenoviral vectors include, for example, those based on human adenovirus type 2 and human adenovirus type 5 that have been made replication defective through deletions in the E1 and E3 regions. The transcriptional cassette can be inserted into the E1

region, yielding a recombinant E1/E3-deleted AV vector. Adenoviral vectors also include helper-dependent high-capacity adenoviral vectors (also known as high-capacity, "gutless" or "gutted" vectors), which do not contain viral coding sequences. These vectors, contain the cis-acting elements needed for viral DNA replication and packaging, mainly the inverted terminal repeat sequences (ITR) and the packaging signal (Ψ). These helper-dependent AV vector genomes have the potential to carry from a few hundred base pairs up to approximately 36 kb of foreign DNA.

[0085] Recombinant adeno-associated virus "rAAV" vectors include any vector derived from any adeno-associated virus serotype, including, without limitation, AAV-1, AAV-2, AAV-3, AAV-4, AAV-5, AAV-7 and AAV-8, AAV-9, AAV-10, and the like. rAAV vectors can have one or more of the AAV wild-type genes deleted in whole or in part, preferably the Rep and/or Cap genes, but retain functional flanking ITR sequences. Functional ITR sequences are retained for the rescue, replication, packaging and potential chromosomal integration of the AAV genome. The ITRs need not be the wild-type nucleotide sequences, and may be altered (e.g., by the insertion, deletion or substitution of nucleotides) so long as the sequences provide for functional rescue, replication and packaging.

[0086] Alternatively, other systems such as lentiviral vectors can be used in embodiments of the invention. Lentiviral-based systems can transduce nondividing as well as dividing cells making them useful for applications targeting, for examples, the non-dividing cells of the CNS. Lentiviral vectors are derived from the human immunodeficiency virus and, like that virus, integrate into the host genome providing the potential for very long-term gene expression.

[0087] Polynucleotides, including plasmids, YACs, minichromosomes and minicircles, carrying the target gene containing the gene regulation cassette can also be introduced into a cell or organism by nonviral vector systems using, for example, cationic lipids, polymers, or both as carriers. Conjugated poly-L-lysine (PLL) polymer and polyethylenimine (PEI) polymer systems can also be used to deliver the vector to cells. Other methods for delivering the vector to cells includes hydrodynamic injection and electroporation and use of ultrasound, both for cell culture and for organisms. For a review of viral and non-viral delivery systems for gene delivery see Nayerossadat, N. et al. (Adv Biomed Res. 2012; 1:27) incorporated herein by reference.

[0088] **Methods of Modulating Expression of a Target Gene**

[0089] In one aspect, this invention provides a method of modulating expression of a target gene (e.g., a therapeutic gene), by (a) inserting the gene regulation cassette of the present invention into a target gene; (b) introducing the target gene comprising the gene regulation cassette into a cell; and (c) exposing the cell to a ligand that binds the aptamer. In one embodiment, the ligand is a small molecule. In aspects, expression of the target gene in target cells confers a desired property to a cell into which it was introduced, or otherwise leads to a desired therapeutic outcome.

[0090] In a preferred embodiment, the gene regulation cassette is inserted into the protein coding sequence of the target gene (rather than in the 5′ or 3′ untranslated regions). In one embodiment, a single gene regulation cassette is inserted into the target gene. In other embodiments 2, 3, 4, or more gene regulation cassettes are inserted in the target gene. In one embodiment, two gene regulation cassettes are inserted into the target gene. When multiple gene regulation cassettes are inserted into a target gene, they each can contain the same aptamer such that a single ligand can be used to modulate alternative splicing at the multiple cassettes and thereby modulate target gene expression. In other embodiments, multiple gene regulation cassettes are inserted into a target gene, each can contain a different aptamer so that exposure to multiple different small molecule ligands modulates target gene expression. In other embodiments, multiple gene regulation cassettes are inserted into a target gene, each containing different 5' intron, alternative exon, and 3' intron sequences. This may be useful in reducing recombination and improving ease of incorporation into viral vectors.

[0091] **Methods of Treatment and Pharmaceutical Compositions**

[0092] One aspect of the invention provides a method of regulating the level of a therapeutic protein delivered by gene therapy. In this embodiment, the "target gene" may encode the therapeutic protein. The "target gene" may encode a protein that is endogenous or exogenous to the cell.

[0093] The therapeutic gene sequence containing the regulatory cassette with aptamer-driven riboswitch is delivered to the target cells in the body, e.g., by a vector. The cell specificity of the "target gene" may be controlled by promoter or other elements within the vector. Delivery of the vector construct containing the target gene, and the transfection of the target tissues resulting in stable transfection of the regulated target gene, is the first step in producing the therapeutic protein.

[0094] However, due to the presence of the regulatory cassette within the target gene sequence, the target gene is not expressed at significant levels, i.e., it is in the "off state" in the absence of the specific ligand that binds to the aptamer contained within in the regulatory cassette riboswitch. Only when the aptamer specific ligand is administered is the target gene expression activated.

[0095] The delivery of the vector construct containing the target gene and the delivery of the activating ligand generally are separated in time. The delivery of the activating ligand will control when the target gene is expressed, as well as the level of protein expression. The ligand may be delivered by a number of routes including, but not limited to, oral, intramuscular (IM), intravenous (IV), intraocular, or topically.

[0096] The timing of delivery of the ligand will depend on the requirement for activation of the target gene. For example, if the therapeutic protein encoded by the target gene is required constantly, an oral small molecule ligand may be delivered daily, or multiple times a day, to ensure continual activation of the target gene, and thus continual expression of the therapeutic protein. If the target gene has a long acting effect, the inducing ligand may be dosed less frequently.

[0097] This invention allows the expression of the therapeutic transgene to be controlled temporally, in a manner determined by the temporal dosing of the ligand specific to the aptamer within the regulatory cassette. The expression of the therapeutic transgene only on ligand administration, increases the safety of a gene therapy treatment by allowing the target gene to be off in the absence of the ligand.

[0098] Different aptamers can be used to allow different ligands to activate target genes. In certain embodiments of the invention, each therapeutic gene containing a regulatory cassette will have a specific aptamer within the cassette that will be activated by a specific small molecule. This means that each therapeutic gene can be activated only by the ligand specific to the aptamer housed within it. In these embodiments, each ligand will only activate one therapeutic gene. This allows for the possibility that several different "target genes" may be delivered to one individual and each will be activated on delivery of the specific ligand for the aptamer contained within the regulatory cassette housed in each target gene.

[0099] This invention allows any therapeutic protein whose gene can be delivered to the body (such as erythropoietin (EPO) or a therapeutic antibody) to be produced by the body

when the activating ligand is delivered. This method of therapeutic protein delivery may replace the manufacture of such therapeutic proteins outside of the body which are then injected or infused, e.g., antibodies used in cancer or to block inflammatory or autoimmune disease. The body containing the regulated target gene becomes the biologics manufacturing factory, which is switched on when the gene-specific ligand is administered.

[00100] Dosing levels and timing of dosing of a therapeutic protein may be critical to therapeutic effect. For example in the delivery of AVASTIN® (anti-VEGF antibody) for cancer. The present invention increases the ease of dosing in response to monitoring for therapeutic protein levels and effects.

[00101] In one embodiment, the target protein may be a nuclease that can target and edit a particular DNA sequence. Such nucleases include Cas9, zinc finger containing nucleases, or TALENs. In the case of these nucleases, the nuclease protein may be required for only a short period of time that is sufficient to edit the target endogenous genes. However, if an unregulated nuclease gene is delivered to the body, this protein may be present for the rest of the life of the cell. In the case of nucleases, there is an increasing risk of off-target editing the longer the nuclease is present. Regulation of expression of such proteins has a significant safety advantage. In this case, vector containing the nuclease target gene containing a regulatory cassette could be delivered to the appropriate cells in the body. The target gene is in the "off" state in the absence of the cassette-specific ligand, so no nuclease is produced. Only when the activating ligand is administered, is the nuclease produced. When sufficient time has elapsed allowing sufficient editing to occur, the ligand will be withdrawn and not administered again. Thus the nuclease gene is thereafter in the "off" state and no further nuclease is produced and editing stops. This approach may be used to correct genetic conditions, including a number of inherited retinopathies such as LCA10 caused by mutations in CEP290 and Stargardts disease caused by mutations in ABCA4.

[00102] Administration of a regulated target gene encoding a therapeutic protein which is activated only on specific ligand administration may be used to regulate therapeutic genes to treat many different types of diseases, e.g., cancer with therapeutic antibodies, immune disorders with immune modulatory proteins or antibodies, metabolic diseases, rare diseases such as PNH with anti-C5 antibodies or antibody fragments as the regulated gene, or ocular angiogenesis with therapeutic antibodies, and dry AMD with immune modulatory proteins.

[00103] A wide variety of specific target genes, allowing for the treatment of a wide variety of specific diseases and conditions, are suitable for use in the present invention. For example, insulin or an insulin analog (preferably human insulin or an analog of human insulin) may be used as the target gene to treat type I diabetes, type II diabetes, or metabolic syndrome; human growth hormone may be used as the target gene to treat children with growth disorders or growth hormone-deficient adults; erythropoietin (preferably human erythropoietin) may be used as the target gene to treat anemia due to chronic kidney disease, anemia due to myelodysplasia, or anemia due to cancer chemotherapy.

[00104] The present invention may be especially suitable for treating diseases caused by single gene defects such as cystic fibrosis, hemophilia, muscular dystrophy, thalassemia, or sickle cell anemia. Thus, human β-, γ-, δ-, or ζ-globin may be used as the target gene to treat β-thalassemia or sickle cell anemia; human Factor VIII or Factor IX may be used as the target gene to treat hemophilia A or hemophilia B.

[00105] The ligands used in the present invention are generally combined with one or more pharmaceutically acceptable carriers to form pharmaceutical compositions suitable for administration to a patient. Pharmaceutically acceptable carriers include solvents, binders, diluents, disintegrants, lubricants, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, generally used in the pharmaceutical arts. Pharmaceutical compositions may be in the form of tablets, pills, capsules, troches, and the like, and are formulated to be compatible with their intended route of administration. Examples of routes of administration include parenteral, e.g., intravenous, intradermal, intranasal, subcutaneous, oral, inhalation, transdermal (topical), transmucosal, and rectal.

[00106] The pharmaceutical compositions comprising ligands are administered to a patient in a dosing schedule such that an amount of ligand sufficient to desirably regulate the target gene is delivered to the patient. When the ligand is a small molecule and the dosage form is a tablet, pill, or the like, preferably the pharmaceutical composition comprises from 0.1 mg to 10 g of ligand; from 0.5 mg to 5 g of ligand; from 1 mg to 1 g of ligand; from 2 mg to 750 mg of ligand; from 5 mg to 500 mg of ligand; or from 10 mg to 250 mg of ligand.

[00107] The pharmaceutical compositions may be dosed once per day or multiple times per day (e.g., 2, 3, 4, 5, or more times per day). Alternatively, pharmaceutical compositions may be dosed less often than once per day, e.g., once every 2, 3, 4, 5, 6, 7, 8, 9,

10, 11, 12, 13, or 14 days, or once a month or once every few months. In some embodiments of the invention, the pharmaceutical compositions may be administered to a patient only a small number of times, e.g., once, twice, three times, etc.

[00108] The present invention provides a method of treating a patient in need of increased expression of a therapeutic protein encoded by a target gene, the method comprising administering to the patient a pharmaceutical composition comprising a ligand for an aptamer, where the patient previously had been administered a recombinant DNA comprising the target gene, where the target gene contains a gene regulation cassette of the present invention that provides the ability to regulate expression of the target gene by the ligand of the aptamer by alternative splicing of pre-mRNA of the target gene, thereby increasing expression of the therapeutic protein.

[00109] **Articles of Manufacture and Kits**

[00110] Also provided are kits or articles of manufacture for use in the methods described herein. In aspects, the kits comprise the compositions described herein (e.g., for compositions for delivery of a vector comprising the target gene containing the gene regulation cassette) in suitable packaging. Suitable packaging for compositions (such as ocular compositions for injection) described herein are known in the art, and include, for example, vials (such as sealed vials), vessels, ampules, bottles, jars, flexible packaging (e.g., sealed Mylar or plastic bags), and the like. These articles of manufacture may further be sterilized and/or sealed.

[00111] The present invention also provides kits comprising compositions described herein and may further comprise instruction(s) on methods of using the composition, such as uses described herein. The kits described herein may further include other materials desirable from a commercial and user standpoint, including other buffers, diluents, filters, needles, syringes, and package inserts with instructions for performing the administration, including e.g., any methods described herein. For example, in some embodiments, the kit comprises rAAV for expression of the target gene comprising the gene regulation cassette of the present invention, a pharmaceutically acceptable carrier suitable for injection, and one or more of: a buffer, a diluent, a filter, a needle, a syringe, and a package insert with instructions for performing the injections. In some embodiments the kit is suitable for intraocular injection, intramuscular injection, intravenous injection and the like.

[00112] "Homology" and "homologous" as used herein refer to the percent of identity between two polynucleotide sequences or between two polypeptide sequences. The correspondence between one sequence to another can be determined by techniques known in the art. For example, homology can be determined by a direct comparison of two polypeptide molecules by aligning the sequence information and using readily available computer programs. Alternatively, homology can be determined by hybridization of polynucleotides under conditions which form stable duplexes between homologous regions, followed by digestion with single-stranded-specific nuclease(s), and size determination of the digested fragments. Two polynucleotide or two polypeptide sequences are "substantially homologous" to each other when, after optimally aligned with appropriate insertions or deletions, at least about 80%, at least about 85%, at least about 90%, and at least about 95% of the nucleotides or amino acids, respectively, match over a defined length of the molecules, as determined using the methods above.

[00113] "Percent sequence identity" with respect to a reference polypeptide or nucleic acid sequence is defined as the percentage of amino acid residues or nucleotides in a candidate sequence that are identical with the amino acid residues or nucleotides in the reference polypeptide or nucleic acid sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent amino acid or nucleic acid sequence identity can be achieved in ways known to the ordinarily-skilled artisan, for example, using publicly available computer software programs including BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software.

[00114] The term "polynucleotide" or "nucleic acid" as used herein refers to a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides. Thus, this term includes, but is not limited to, single-, double- or multi- stranded DNA or RNA, genomic DNA, cDNA, DNA-RNA hybrids, or a polymer comprising purine and pyrimidine bases, or other natural, chemically or biochemically modified, non-natural, or derivatized nucleotide bases.

[00115] "Heterologous" or "exogenous" means derived from a genotypically distinct entity from that of the rest of the entity to which it is compared or into which it is introduced or incorporated. For example, a polynucleotide introduced by genetic engineering techniques into a different cell type is a heterologous polynucleotide (and, when expressed, can encode a heterologous polypeptide). Similarly, a cellular sequence (e.g., a gene or portion thereof) that

is incorporated into a viral vector is a heterologous nucleotide sequence with respect to the vector.

[00116] It is to be understood and expected that variations in the principles of invention herein disclosed can be made by one skilled in the art and it is intended that such modifications are to be included within the scope of the present invention. The following Examples further illustrate the invention, but should not be construed to limit the scope of the invention in any way. All references cited herein are hereby incorporated by reference in their entirety.

## EXAMPLES

[00117] **EXAMPLE 1. Effects of splice site strength on the expression of regulated genes**

[00118] **Experimental procedures**

[00119] *Plasmid constructs:* Luci-BsaI-acceptor: A DNA fragment containing a CMV promoter was released from pHAGE-CMV-eGFP-W (Harvard University) vector by restriction enzymes SpeI and NotI, and this fragment was cloned into pHDM-G (Harvard University) vector digested with SpeI and NotI. A fragment containing SV40 Ori sequence in the resulting vector was deleted by digesting with BsmI and BstXI, removing the 3' overhang and ligating. The subsequent vector was subjected to site directed mutagenesis (Agilent) to delete the BsaI site in the AmpR gene. The resulting vector was then digested with NotI and BamHI, and ligated with a fragment containing NotI-BsaI-BamHI sites to generate the final Luci-BsaI-acceptor vector. pHDM-G was used as template for the human beta-globin intron 2 ("IVS2Δ") containing a deletion of the middle portion considered non-crucial to splicing (*see* Table 5, SEQ ID NO:1). pGL3-promoter (Promega) was used as template for the firefly luciferase gene. Construct 8: Luciferase gene was amplified by PCR using primers Luc-For-BsaI and Luci-Rev-BsaI. The PCR products were digested with BsaI and cloned into BsaI-digested Luci-BsaI-acceptor vector. Splicing Constructs 1-7 (Con 1-7, SEQ ID NOS. 1-7) expressing the luciferase gene inserted with intron IVS2Δ that has different 5′ss and 3′ss at each end of the intron sequence were made by ligating 3 BsaI-digested PCR products into BsaI-digested Luci-BsaI-acceptor. pGL3-promoter vector was used as Luciferase template, and pHDM-G was used as template for IVS2Δ. Primer pairs used to amplify PCR fragments for Con 1-7 are as follows: Con 1: Luci-For-BsaI/Luci-

Splice-Rev_1, IVS2-BsaI-For/IVS2-BsaI-Rev_1 and Luci-Splice-For_1/Luci-Rev-BsaI; Con 2: Luci-For-BsaI/Luci-Splice-Rev_2, IVS2-BsaI-For/IVS2-BsaI-Rev_2, and Luci-Splice-For_2/Luci-Rev-BsaI; Con 3: Luci-For-BsaI/Luci-Splice-Rev_3, IVS2-BsaI-For/IVS2-BsaI-Rev_3, and Luci-Splice-For_3/Luci-Rev-BsaI; Con 4: Luci-For-BsaI/Luci-Splice-Rev_4, IVS2-BsaI-For/IVS2-BsaI-Rev_1, and Luci-Splice-For_4/Luci-Rev-BsaI; Con 5: Luci-For-BsaI/Luci-Splice-Rev_1, IVS2-BsaI-For/IVS2-BsaI-Rev_1 and Luci-Splice-For_5/Luci-Rev-BsaI; Con 6: Luci-For-BsaI/Luci-Splice-Rev_1, IVS2-BsaI-For/IVS2-BsaI-Rev_1 and Luci-Splice-For61/Luci-Rev-BsaI; Con 7: Luci-For-BsaI/Luci-Splice-Rev_1, IVS2-BsaI-For/IVS2-BsaI-Rev_1 and Luci-Splice-For71/Luci-Rev-BsaI. All constructs were verified by DNA sequencing.

Table 1. Splice sites of the splicing constructs (Con 1-7). The intron/exon boundaries are marked by ||.

| **Construct** | **5′ splice site** | **3′ splice site** |
|---|---|---|
| Con 1 | AGG\|\|**GT**GAGT | TCTTATCTTCCTCCCAC**AG**\|\|C |
| Con 2 | AAA\|\|**GT**AAGC | TCTTATCTTCCTCCCAC**AG**\|\|C |
| Con 3 | GCA\|\|**GT**AAGT | TCTTATCTTCCTCCCAC**AG**\|\|C |
| Con 4 | GAG\|\|**GT**GTGG | TCTTATCTTCCTCCCAC**AG**\|\|C |
| Con 5 | AGG\|\|**GT**GAGT | CTTTACTTCTATGACTGT**AG**\|\|C |
| Con 6 | AGG\|\|**GT**GAGT | GTGACTGTGTGTATGCAC**AG**\|\|C |
| Con 7 | AGG\|\|**GT**GAGT | ATTGTGATCGCAGCCAAT**AG**\|\|C |

[00120] *Transfection:* 3.5 x10^4 HEK 293 cells were plated in 96-well flat bottom plate the day before transfection. Plasmid DNA (500 ng) was added to a tube or a 96-well U-bottom plate. Separately, TransIT-293 reagent (Mirus; 1.4 uL) was added to 50 µL Opti-mem I media (Life Technologies), and allowed to sit for 5 minutes at room temperature (RT). Then, 50 uL of this diluted transfection reagent was added to the DNA, mixed, and incubated at room temperature ("RT") for 20 min. Finally, 7 µL of this solution was added to a well of cells in the 96-well plate.

[00121] *Firefly luciferase assay of cultured cells:* 24 hours after media change, plates were removed from the incubator, and equilibrated to RT for several minutes on a lab bench, then aspirated. Glo-lysis buffer (Promega, 100 µL, RT) was added, and the plates allowed to remain at RT for at least 5 minutes. Then, the well contents were mixed by 50 µL trituration, and 20 µL of each sample was mixed with 20 µL of bright-glo reagent (Promega) that had been diluted to 10% in glo-lysis buffer. 96 wells were spaced on an opaque white 384-well plate. Following a 5 min incubation at RT, luminescence was measured using Tecan machine

with 500 mSec read time. The luciferase activity was expressed as mean relative light unit (RLU)±S.D.

[00122] Results

[00123] In order to build a splicing-based gene regulation platform, we first tested (i) the effect of inserting an intron into the coding sequence (CDS) of a gene of interest, in this case the firefly luciferase gene (Fig. 1a), and (ii) the effects of different 5' ss and 3' ss on gene expression. The truncated human beta-globin intron 2 (IVS2Δ) with different 5' ss and 3' ss at each end was inserted in the coding sequence of the firefly luciferase gene to test efficiency of splicing. Construct Con 8 has no IVS2Δ intron, and Con 1 (SEQ ID No.: 1) has IVS2Δ with its original IVS2 5' and 3' ss sequences. Constructs Con 2 to Con 7 (SEQ ID No.: 2-7) have IVS2Δ with different 5' and 3' ss sequences as listed in Table 1. As shown in Fig. 1b, insertion of IVS2Δ with native IVS2 splice sites into the luciferase gene did not affect gene expression (compare Con 1 vs Con 8). However, replacement of IVS2 splice sites in IVS2Δ with splice site sequences having different strength significantly impaired luciferase expression. As shown in Fig. 1b, Con 2 and Con 3 with altered 5' ss have expression levels similar to Con 1 and Con 8, however the 5'ss changes in Con 4, and 3' ss changes in Con 5 through Con 7, did significantly reduce luciferase expression (compare Con 4 to Con 7 with Con 8). Therefore, differences in splice sites affect target gene expression. Con 1 was used for further development.

[00124] **EXAMPLE 2. An Intron-Exon-Intron cassette and the effect of cis-elements on splicing in modulating target gene expression.**

[00125] Experimental Procedures

[00126] Putative exon splice enhancer (ESE) sequences were predicted using ESEfinder 3.0. The wild type human dihydropholate reductase (DHFR) exon 2 with intronic flanking sequences, either with native 5'ss (DHFR-wt; (Table 2); SEQ ID NO.: 47), native 5'ss with four nucleotides mutated to C (DHFR-wt5ssC; (Table 2); SEQ ID NO.:48), 5'ss sequences from Con 1 (DHFR-Con 15ss; Table 2 SEQ ID NO.:49) or Con4 (DHFR-Con45ss, SEQ ID NO.:50) were synthesized (IDT). To test the effect of ESE and exon splice suppressor (ESS) sequences within the DHFR exon 2, different DHFR exon 2 mutants were synthesized (listed in Table 2).

[00127] All of these different DHFR exon 2 sequences were cloned into the approximate center of the IVS2Δ intron in the Con 1 construct using the Golden Gate cloning strategy (NEB).

[00128] The constructs were verified by DNA sequencing (Genewiz). DNA was transfected in HEK 293 cells and assayed for luciferase activity as described in Example 1.

Table 2. DHFR exon 2 containing modified splice regulatory sequences. The underlined sequence indicates the modified splicing regulatory sequences within the DHFR exon 2.

| | |
|---|---|
| DHFR-wt | GAATGAATTCAGATATTTCCAGAGAATGACCACAACCTCTTCAGT AGAAG |
| mtSRp40 | GAATGAATTCAGATATTTCCAGAGAATGAAAAAAAAAATCTTCAGT AGAAG |
| StrSC35 | GAATGGCCCCTGATATTTCCAGAGAATGACCACAACCTCTTCAGT AGAAG |
| SC35hnRNPA1 | GAATGTAGGGAGATATTTCCAGAGAATGACCACAACCTCTTCAG TAGAAG |

[00129] Results

[00130] The wild type human DHFR exon 2 and adjacent intronic sequences (SEQ ID NO.: 8) was inserted into the approximate center of the IVS2Δ intron in the Con 1 construct. This configuration generates a platform in which an exogenous exon in the intron sequence of a target gene can serve as an alternative exon allowing the expression of the target gene to be regulated through modulating alternative exon splicing. In this configuration (Fig. 2a), splicing events presumably occur between the 5′ portion of the target gene and the DHFR exon, as well as between the DHFR exon and the 3′ portion of target gene, resulting in inclusion of the DHFR exon into the target gene mRNA. As the alternative DHFR exon contains an in-frame premature stop codon when the DHFR exon is included in the mRNA, thereby reducing luciferase gene expression. However, when the 5′ ss of the alternative DHFR exon (i.e., the splice donor site at the 5′ end of the 3′ intron) is mutated or inaccessible preventing splicing at this site, the DHFR exon is excluded from the mRNA, and the mRNA is efficiently translated and the target gene protein is expressed (Fig. 2a).

[00131] We first tested the splicing of DHFR exon 2 with its native cis-elements unchanged, as well as various other versions in which the 5′ ss sequences were either

strengthened or weakened. Insertion of the DHFR exon with native 5'ss and 3'ss (SEQ ID NO.: 8) into the intronic sequence in Con 1 to create DHFR_wt, significantly decreased luciferase expression compared to Con1 which contains no alternative DHFR exon. Expression from the DHFR_wt construct is 116-fold lower than Con 1(Fig 2b).

[00132] When the 5′ ss of the DHFR exon is mutated to a non-functional sequence (DHFR_wt5ssC; SEQ ID NO.: 48), DHFR exon inclusion is blocked and luciferase expression is restored to the level of Con 1 (Fig 2aII, 2b, 2c)

[00133] When the 5′ ss of the DHFR exon is replaced with the stronger 5′ ss, in this case the 5'ss from Con 1 (DHFR_Con1 5ss; SEQ ID NO.: 49), the inclusion of the DHFR exon is increased, leading to a 545-fold decrease in luciferase gene expression compared with Con 1 (Fig 2b). However, when the weak 5'ss from Con 4 was used (DHFR_Con4 5ss; SEQ ID NO.: 50), the DHFR exon is not included and luciferase expression is increased (Fig. 2b).

[00134] Exonic splicing enhancer (ESE) or suppressor (ESS) elements play crucial roles in splicing, and their functions often are context dependent. The effect of putative splicing regulatory sequences located within the DHFR exon were tested. When a putative splicing enhancer, the SRp40 binding site located within the DHFR exon was mutated (Table 2, DHFR_mtSRp40; SEQ ID NO.: 51), DHFR exon inclusion was dramatically enhanced, resulting in a 2982-fold decrease in luciferase expression compared to Con 1 (Fig. 2c, DHFR_wt and DHFR mtSRp40).

[00135] When another splicing enhancer, the SC35 binding site within the DHFR exon (predicted by ESE finder) was mutated to a stronger SC35 binding sequence (Table 2, DHFR_StrSC35; SEQ ID NO.: 52), DHFR exon inclusion was enhanced, decreasing luciferase expression by 139-fold compared to Con1 (Fig 2c, DHFR_wtStrSC35). This is a slightly greater decrease than was seen with the construct containing the native DHFR exon (DHFR_wt Fig. 2b)

[00136] When the splicing enhancer, SC35 binding site was mutated to a splicing inhibitor (hnRNP A1 binding site) (Table 2, DHFR_wtSC35hnRNPA1; SEQ ID NO.: 53), the inclusion of the DHFR exon was less efficient leading to increased luciferase expression (Fig 2c, DHFR_wt and DHFR_wtSC35hnRNPA1).

[00137] An Intron-Exon-Intron cassette has been created in which the expression of the target gene, in this case luciferase, can be switched on or off depending on the inclusion or exclusion of an alternative exon, in this case the alternative DHFR exon containing an in-

frame stop codon. Splicing that results in inclusion of the alternative exon reduces gene expression, while gene expression increases when splicing excludes the alternative exon. The strength or weakness of the alternative exon 5' ss as well as sequences within the exon that modulate splicing alter the level of target expression via their impact on the inclusion or exclusion of the exogenous exon.

[00138] **EXAMPLE 3. The effects of hairpin formation at the alternative exon 5' splice site on regulating target gene expression.**

[00139] Experimental Procedures

[00140] Sequences containing the DHFR exon 2 with native 3' and 5' ss sequences, in which the 5'ss was embedded in a hairpin structure were synthesized (IDT), and cloned into the indicated vector using Golden Gate cloning strategy (NEB). Constructs were transfected into HEK 293 cells and assayed for luciferase activity as described in Example 1.

[00141] Results

[00142] We tested whether embedding the 5' ss of the DHFR exon into a hairpin stem structure could affect splicing and inclusion of the alternative DHFR exon and thus alter target gene expression (illustrated in Fig. 3a).

[00143] Inclusion of alternative DHFR exon with Con1 5' splice site (DHFR_Con15ss; SEQ ID NO.: 49) sequences abolishes luciferase expression compared to Con 1 (Fig. 3c, DHFR_Con15ss). A 15 base pair (bp) hairpin structure embedding the entire sequence of the 5′ ss was engineered into DHFR_Con 15ss to create DHFR_Con 15ss_HP15 (SEQ ID NO.: 54) (Fig. 3a). The presence of the 15 bp hairpin structure completely restores luciferase expression to the level of Con 1, indicating that the hairpin has abolished accessibility of the 5' ss and thereby abolished inclusion of the alternative DHFR exon. (Fig 3c, Con 1, DHFR_Con15ss and DHFR_Con15ss_HP15)

[00144] In contrast, a 15 bp hairpin with a "broken stem" (Fig. 3b, Con15ss_15HPx; SEQ ID NO.: 55) was not able to restore luciferase expression (Fig. 3c, DHFR_Con15ss_HP15x), indicating the intact stem is an important component of the RNA secondary structure in regulating the accessibility of 5′ splice site and thereby determining the inclusion or exclusion of the alternative exon.

[00145] The same experiments were carried out using the construct containing the DHFR exon with mutant SRp40 binding site that has increased splicing efficiency

(DHFR_wtmtSRp40, see Example 2). Embedding the 5' ss in a hairpin restored luciferase expression, whereas breaking the hairpin blocked luciferase expression (Fig 3c, DHFR_wtmtSRp40, DHFR_wtmtSRp40 _HP15 and DHFR_wtmtSRp40_HP15x)

[00146] Thus, embedding the 5'ss of an alternative exon in a hairpin structure can restore target gene expression by blocking accessibility of that 5′ ss and, thereby preventing inclusion of the alternative exon into the mRNA, and allowing target gene protein expression. A gene expression platform was created in which target gene protein expression can be modulated by altering availability of the 5′ ss of an exogenous alternative exon through secondary RNA structure.

[00147] The construct DHFR_wtmtSRp40 (SEQ ID NO.: 58) (referred to as "mtDHFR" below) was used for further riboswitch development.

[00148] **EXAMPLE 4. Use of a theophylline aptamer to regulate target gene expression via alternative splicing.**

[00149] Experimental Procedures

[00150] A DHFR-acceptor vector was constructed to facilitate the cloning of an aptamer sequence attached to the hairpin stems with different length. The theophylline aptamer sequence used was: ggcgatacCAGCCGAAAGGCCCTTGgcagcgtc (SEQ ID NO:9). Theophylline aptamer oligonucleotides ("oligos") with 4 nucleotide overhang at 5′ end were synthesized (IDT), annealed and ligated to BsaI-digested DHFR-acceptor vector. HEK 293 cells were transfected with the luciferase reporter constructs with the regulation cassette containing the theophylline aptamer, as described in Example 1. Four hours after transfection, the media was aspirated, and new media with or without 3 mM theophylline was added, and luciferase was assayed 20 to 24 hours after theophylline treatment. The fold induction was expressed as the quotient of luciferase activity obtained in the presence of aptamer ligand divided by the value obtained in the absence of the aptamer ligand. The level of the luciferase activity was expressed as percent of level of luciferase activity (referred as maximal expression) produced by Con1 construct that does not IVS2Δ intron in the CDS of luciferase gene.

[00151] Results

[00152] In order to regulate the accessibility of the 5′ splice site of the stop codon-containing alternative exon, and thereby regulate target gene protein expression, aptamer

sequences were attached to the stem of a hairpin structure that embeds the intronic portion of the DHFR 5′ ss and its complementary sequence In this configuration, insertion of aptamer sequences disrupts the formation of the hairpin stem, leaving DHFR 5' ss accessible, thus resulting in the inclusion of alternative DHFR exon and preventing target gene protein expression (Fig. 4a). When aptamer/ligand binding occurs, as depicted in Fig. 4b, conformational change in aptamer triggered by ligand binding brings together the DHFR 5′ ss and its complementary sequence for stable stem formation, thus hiding the DHFR 5′ ss and resulting in DHFR exon exclusion and target gene protein expression.

[00153] A theophylline aptamer was tested by linking the lower stem of theophylline aptamer directly to the hairpin stem (Fig. 4c, DHFR_Theo1). If the stem is too long, it may form a stable structure in the absence of aptamer/ligand binding, while if it is too short it may never form a stable stem, even when the ligand is present. Therefore the length of the stem needs to be optimized such that a stable secondary structure is only formed on aptamer/ligand binding. To determine the optimal stem length that allows stem formation in the presence but not absence of the ligand, a number of constructs were made in which the theophyline aptamer was cloned into mtDHFR (described in Example 2, Table 2) and serial truncations of the stem were carried out. Fig. 4c shows four constructs from this serial truncation.

[00154] Fig. 4d shows the expression of luciferase from this deletion series of constructs in the presence and absence of theophylline. In constructs Theo 1 through Theo 12 with stem lengths from 20bp down to 9bp, the stem length was sufficient to form a stable secondary structure even in the absence of aptamer /ligand binding. Thus luciferase expression is seen at levels similar to Con 1 in both the absence and the presence of thyophylline..

[00155] With construct DHFR_Theo13, luciferase expression is suppressed in the absence of theophylline. This indicates the availability of the mtDHFR exon 5'ss, leading to inclusion of the alternative DHFR exon and suppression of gene expression. However, in the presence of theophylline, luciferase expression was switched on, resulting in a 43-fold induction over the expression level without theophylline, and about 56% of the luciferase level expressed by the Con1 control vector. Therefore, a mammalian on-riboswitch was generated, which turns on target gene protein expression in the presence of the aptamer ligand, theophylline.

[00156] **EXAMPLE 5. Use of xpt guanine aptamer to regulate target gene expression via alternative splicing.**

[00157] Experimental Procedures

[00158] Xpt-guanine aptamer with the following sequence: cactcatataatCGCGTGGATATGGCACGCAAGTTTCTACCGGGCACCGTAAATGTCcgact atgggtg (SEQ ID NO.: 10), was used to construct a riboswitch. Oligos containing the sequence of guanine aptamer and hairpin stem with 4 nucleotide 5′ overhang were synthesized (IDT), annealed and then ligated to BsaI-digested DHFR-acceptor vector. HEK 293 cells were transfected as described in Example 1. Four hours after transfection, the media was aspirated, and new media with or without 500 μM guanine was added. Luciferase expression was assayed 20 to 24 hrs after guanine treatment as described in Example 1 and Example 4. HepG2, AML12, RD and C2C12 (ATCC) were cultured using the protocol recommended by ATCC. The intron-exon-intron cassette containing the xpt-G17 riboswitch (SEQ ID NO.: 15) was inserted in the leader peptide sequences of the anti-KDR antibody gene and into the StuI site in mouse erythropoietin gene using Gibson cloning strategy (NEB). Constructs containing mouse erythropoietin (Epo) or anti-KDR antibody were transfected into HEK 293 cells. Four hour after transfection, the media was aspirated, and new media with or without 500 μM guanine was added. Supernatants were subjected to ELISA assay for the production of either anti-KDR antibody or the production of mouse Epo (R&D Systems).

[00159] Results

[00160] The use of additional aptamer/ligands to control target gene expression by aptamer-mediated modulation of alternative splicing was studied by attaching an xpt-guanine aptamer, derived from *Bacillus subtilis*, through stem P1 to the hairpin stem (Figure 5a, DHFR_G1). Similar to Example 4, 18 constructs were made by serial truncation of the connecting stem (Figure 5a and 5b; DHFR-G1 through G18, also referred to as xpt-G1 through G18 containing regulatory cassettes) to obtain the optimal length of hairpin stem in connection with the guanine aptamer, thus allowing the communication of aptamer/ligand binding to 5'ss accessibility and exon splicing. As shown in Fig. 5b, with constructs DHFR-G1 through G13, with stem lengths from 24 bp down to 12 bp, luciferase expression is not affected by the insertion of the alternative DHFR exon and xpt—guanine aptamer in the presence or absence of the aptamer ligand guanine. This suggests that the length of the stem

is sufficient to form a stable structure in both the absence and the presence of the ligand, preventing inclusion of the alternative exon in the mRNA. However, in constructs DHFR_G14 through DHFR_G18, luciferase expression was suppressed in the absence of added guanine. When μM guanine was added, luciferase expression from these constructs was induced (Fig 5b).

[00161] A further stringent validation of constructs G11 through G18 again showed clear regulation of luciferase expression upon guanine treatment (Fig. 5c). Construct DHFR_G17 containing xpt-G17 (SEQ ID NO.: 15) (Fig. 5a) gave 2000-fold induction of expression, resulting in about 65% of the level of luciferase expressed by Con1 (referred as maximal expression). This high dynamic range of induction resulted from activation of expression from a very low un-induced baseline level in the absence of the ligand. Construct DHFR_G16 (Fig. 5a) gave about 800-fold induction over un-induced baseline expression to a level that was 83% of maximal expression (Fig. 5c and 5d). In addition, constructs DHFR_G14 and G15 showed nearly 100% of maximal expression with a lower fold induction due to higher un-induced baseline expression of luciferase.

[00162] To test the general functionality and applicability of the synthetic riboswitch in the Intron-Exon-Intron cassette, we transfected the xpt-G17 containing construct (DHFR_G17) into multiple human and mouse cell lines. In these different cell lines, guanine treatment generated significant induction of gene expression, more than 500-fold induction in HepG2 cells, with lower fold induction in other cell lines (Fig 5e). The different fold induction in different cell lines may reflect differences in transfection efficiency as well as in the cell-type specific splicing regulator expression profile. Further, the luciferase gene with the regulatory cassette containing the xpt-G17 riboswitch (DHFR_G17) yielded similar level of induction when transferred to an AAV backbone (Figure 5f), indicating that the gene regulating effect is not vector backbone dependent.

[00163] In addition to regulating the luciferase gene, the xpt-G17 containing regulatory cassette was also tested in regulating secreted proteins, anti-KDR antibody and erythropoietin (Epo). The xpt-G17 containing regulatory cassette was inserted into the coding sequence of anti-KDR antibody and of erythropoietin. As shown in Figure 5g and 5h, guanine treatment yielded 80-fold induction in anti-KDR antibody production and140-fold induction in Epo production, when compared to the production of each molecule from cells in the absence of ligand.

[00164] These results demonstrate the general functionality and applicability in regulating protein expression of a potential therapeutic target gene, as well as the application of this gene regulation cassette in AAV-mediated gene delivery. Thus, we have created a synthetic mammalian "on"-riboswitch which is capable of switching on target gene protein expression in response to the presence of an aptamer specific ligand in mammalian cells.

[00165] **EXAMPLE 6. Different purine aptamers may be used to regulate target gene expression via alternative splicing.**

[00166] Experimental Procedures

[00167] The following aptamer sequences, listed in Table 3, were used to build riboswitches:

Table 3.

| | |
|---|---|
| Ydhl-G<br>SEQ ID NO.: 11 | ttgtataacctcaataatatggtttgagggtgtctaccaggaaccgtaaaatcctgactacaa |
| Ydhl-A<br>SEQ ID NO.: 12 | ttgtataacctcaataatatggtttgagggtgtctaccaggaaccgtaaaatcctgattacaa |
| addA-G<br>SEQ ID NO.: 13 | tcatataatcctaatgatatggtttgggagtttctaccaagagccttaaactcttgactatga |
| addA-A<br>SEQ ID NO.: 14 | tcatataatcctaatgatatggtttgggagtttctaccaagagccttaaactcttgattatga |

[00168] Results

[00169] To test additional aptamers in our gene regulation system, we used the same strategy and method as described in previous Examples to generate multiple guanine and adenine responsive riboswitches by linking different guanine and adenine aptamers to the intron-mtDHFR-intron cassette (Fig. 6a). The guanine riboswitches that were tested efficiently regulated the expression of the luciferase gene in response to guanine (Fig. 6b). Additionally, we discovered that these guanine riboswitches regulated the expression of the target gene in response not only to guanine (Figure 6b), but also to guanosine (Fig. 6c), and 2'deoxyguanosine (2'dG) (Fig. 6d).

[00170] A number of adenine riboswitches (Fig. 6a) were generated, and also demonstrated gene regulation functionality (Fig. 6e).

[00171] The differences in the regulation of gene expression for the different aptamer containing constructs tested could reflect differences in aptamer/ligand binding affinity and aptamer secondary structure that may affect the accessibility of 5′ ss, alternative exon inclusion, and therefore target gene expression. The Intron-Exon-Intron gene regulation cassette can be optimized by changing aptamer sequences to achieve desired level of gene regulation.

[00172] **EXAMPLE 7. The on/off state of target gene expression regulated by mammalian guanine riboswitch.**

[00173] Experimental Procedures

[00174] The intron-mtDHFR-aptamer-intron cassette was PCR amplified and cloned using Golden Gate cloning strategy (NEB) into pEGFP-C1 vector. To obtain a cell line stably expressing EGFP with riboswitch, HEK-293 cells were electroporated with 20 ng of plasmid DNA. Forty eight hours after electroporation, cell culture was selected with 800 μg/ml G418 for 2 weeks for cells that stably express the cassette. Cells were trypsinized and cell suspension was subjected to flow cytometric analysis of intensity of GFP fluorescence using a Guava EasyCyte 8HT machine. The resulting data was analyzed using GuavaSoft2.2.2.

[00175] Results

[00176] To further demonstrate that expression of a target gene containing our Intron-Exon-Intron regulatory cassette can be regulated by exposure to the ligand specific to the aptamer contained within the riboswitch, the intron-mtDHFR-intron cassette containing the xpt-G17 riboswitch (SEQ ID NO.: 15) was inserted into the EGFP gene, and stably transfected HEK 293 cells. In the presence of guanine, EGFP expression was switched on (Fig 7a). The fluorescence was detected as early as 6 hours after guanine treatment and increased over 3 days of guanine treatment, reaching close to 300-fold induction compared to untreated cells (Fig. 7b), indicating the “on” status of target gene expression in the presence of aptamer ligand. When guanine was withdrawn from the cell culture medium, EGFP expression diminished, indicating the “off” status of the target gene expression in the absence of the aptamer specific ligand (Fig. 7b). Thus, we have created a gene regulation platform, comprised of an Intron-Exon-Intron cassette containing a synthetic riboswitch, through which

the expression of a target gene is regulated, in mammalian cells, by the presence or absence of a specific aptamer ligand.

[00177] **EXAMPLE 8. Effects of multiple regulation cassettes on regulating target gene expression.**

[00178] Experimental Procedures

[00179] Constructs were made using Golden Gate cloning strategy (NEB). HEK 293 cells were transfected with the indicated constructs, treated with the 500 μM of guanine or 1 mM guanosine (Sigma) 4 hr after transfection. Luciferase activity was assayed as described in Example 5.

[00180] Results

[00181] The construct with the xpt-G15 (SEQ ID NO.: 46) containing regulatory cassette (DHFR_G15; Example 5), showed 60-fold induction of luciferase expression in response to guanine treatment, compared to the un-induced basal expression level and reached nearly 100% of level of luciferase expressed by Con 1 (Fig. 8a). This is a useful feature when regulating a therapeutic protein that is required at high levels.

[00182] In contrast, the construct with the xpt-G17 containing regulatory cassette (DHFR_G17) had significantly higher fold induction of 2181-fold, due to the lower un-induced baseline expression , but a considerably lower maximal level of expression upon induction compared to Con1 (Fig 8a).

[00183] To test whether two copies of the xpt-G15 containing regulatory cassette (xpt-G15 double; SEQ ID NO.: 64) could reduce basal levels of expression, without compromising the maximal expression level of luciferase upon induction, two copies of the xpt-G15 containing regulatory cassette were embedded into the luciferase gene, each copy at a different location in the gene sequence. When two copies of xpt-G15 containing regulatory cassette were present, the un-induced baseline expression was decreased resulting in a significantly higher induction fold (from 60-fold to 1008-fold), without compromising the maximum expression level (Figure 8a).

[00184] The EC50 of guanine for the xpt-G15 double cassette (xpt-G15 double; SEQ ID NO.: 64) was 5 times lower than the EC50 of guanine for the construct containing a single copy of the the more stringent xpt-G17 containing cassette (43 μM v.s. 206 μM), thus increasing the sensitivity of ligand response (Fig. 8a).

[00185] The strategy of using two copies of a less stringent regulatory cassette to enhance fold induction and maximal induced gene expression, was also applied to the EGFP gene. As shown in Fig. 8b and 8c, consistent with the results for luciferase regulation (Fig 8a), a single copy of the xpt-G15 regulatory cassette in the EGFP gene (EGFP-xpt-G15) generated higher un-induced baseline level of EGFP expression when compared with the xpt-G17 regulatory cassette containing construct (EGFP-xpt-G17). However, when two copies of xpt-G15 regulation cassette were inserted into the EGFP gene at different locations (EGFP-xpt-G15 double), the un-induced baseline expression level of was decreased to that of EGFP-xpt-G17with the induced level of EGFP even higher than Con1-EGFP control (Fig. 8c).

[00186] Further, one copy of the xpt-G17 containing regulatory cassette and one copy of a Ydhl-A5 adenine riboswitch containing regulatory cassette were embedded into the luciferase gene. Luciferase expression was induced by either addition of adenine (25-fold) or addition of guanine (120-fold) alone, however, a significantly higher level of induction was achieved (up to 2966-fold) with the combined use of adenine and guanine at each of their highest concentration used (Fig. 8d). These results demonstrate the modularity of the alternative splicing based riboswitches in regulating target gene expression.

[00187] In order to reduce recombination and increase the ease of production of viral vectors containing two or more regulatory cassettes, regulatory cassettes with different intron and exon sequences may be used in a single target gene, and these may contain either the same or different ligand responsive aptamers.

[00188] **EXAMPLE 9. Effects of intron size and sequence on regulating target gene expression via aptamer-mediated alternative splicing.**

[00189] Experimental Procedures

[00190] The Con1 construct was used as a template for PCR amplification of intron fragments that have either upstream or downstream intron deletions. To generate constructs that have single intron deletions, PCR products were cloned into the constructs containing the xpt-G17 riboswitch using Golden Gate cloning strategy (NEB). To generate constructs with both upstream and downstream intron deletions, fragments released by EcoRI and BamHI from constructs with single deletions in the downstream intron sequence were cloned into EcoRI and BamHI-digested constructs with single deletions in the upstream intron sequence.

[00191] Results

[00192] Introns contain elements that mayeither promote (intronic splicing enhancer, ISE) or suppress (intronic splicing suppressor, ISS) exon splicing. Among all the riboswitches we have generated, xpt-G17 demonstrated the best regulating ability in terms of both the induction fold and the level of induced gene expression. Using the xpt-G17 riboswitch in the Intron-Exon-Intron cassette, we made a series of modification in the intron sequences and intron length and in the splice sites to further optimize the system.

[00193] First, the effect of intron modification was tested by introducing single deletions in intron sequences either upstream or downstream of the mtDHFR exon (Figure 9a, 9b) 16 constructs with the xpt-G17 containing riboswitch were generated xpt-G17-IR-1 through xpt-G17-IR-16 (sequences of 13 of these constructs are given in Table 5, SEQ ID NOS: 16-28). Then, upstream and downstream intron deletions were combined to generate larger intron deletions, as depicted in Figure 9c. As shown in Fig. 9d and 9e, of the 16 constructs made with two intron deletions (2IR), constructs 2IR-1 through 2IR-10 (SEQ ID NOS.: 29-38) showed significantly higher induction folds, without compromising the induced expression level of luciferase, with 2IR-3 having the greatest improvement in fold induction (4744-fold). In addition, we also made constructs with a mutated 3′ ss upstream of mtDHFR exon and also reduced the size of the downstream intron. As shown in Figure 9d and 9e (constructs DHFR_3ssC_1 to 5) these modifications further improved the relative fold induction, however in this case a reduction in the level of induced expression was observed (from 64% to 32% for 3ssC_3).

[00194] These results indicate that the gene regulating ability of the Intron-Exon-aptamer-Intron cassette can be optimized through modifying the intron sequences flanking the alternative exon in order to achieve the desired level of gene regulation.

[00195] **EXAMPLE 10. The use of multiple natural exons as well as synthetic exons in the gene regulation cassette.**

[00196] Experimental Procedures

[00197] Sequences of mutant human Wilms tumor 1 exon 5 (mutWT1-e5, SEQ ID NO.: 61), SIRT1 exon 6 (SIRT1-e6, SEQ ID NO.: 62), mouse calcium/calmodulin-dependent protein kinase II delta exon 16, or 17 (Camk2d-e16 or e17, SEQ ID NOs.: 59, 60), and synthetic exon ENEEE (SEQ ID NO.: 63) were synthesized (IDT) and cloned into the DHFR-G17 vector in place of the DHFR exon using Gibson Cloning kit (NEB). Plasmid

DNA was transfected into HEK 293 cells, treated with 500 μM guanine, and luciferase assay performed as described in Example 1. The sequences of each exon with 5′ and 3′ splice sites are shown with exonic sequences in uppercase letters (Table 5, SEQ ID NOS.: 59 to 63).

[00198] Results

[00199] In order to determine that the regulating function of our Intron-Exon-aptamer-Intron cassette is not limited to a specific exon sequence, we replaced the mtDHFR exon in the construct containing the guanine xpt-G17 riboswitch (DHFR-xpt-G17), with multiple different natural and mutant exons as well as synthetic exons that contain known exonic splice enhancer sequences (ESE). As shown in Figure 10, the regulation cassette with the CamkIId-e16 exon generates nearly equivalent fold induction compared to DHFR-xpt-G17 but with a lower level of both basal and induced luciferase expression. Cassettes containing other exons also showed variable levels of both basal and induced luciferase expression. Thus, the aptamer-mediated alternative splicing gene regulation cassette is not exon-specific, and not limited to the mDHFR-e2 exon.

[00200] Exons that can generate efficient alternative splicing events are suitable for the aptamer-mediated gene regulation cassette. These results further indicate that the gene regulating capability of this Intron-Exon-aptamer-Intron cassette can be optimized by modifying the sequences in the alternative exon as well as surrounding intron sequences, for example the splicing strength of the 5'ss and 3' ss sequences of the alternative exon as well as ESE and ESS sequences in the alternative exon as described herein.

[00201] **EXAMPLE 11. Regulation of target gene expression by aptamer-mediated alternative splicing *in vivo* in mice.**

[00202] Experimental Procedures

[00203] *Hydrodynamic DNA delivery and drug treatment:* 5μg or 10μg of endotoxin-free plasmid DNA containing the luciferase gene with two copies of the xpt-G15 containing regulatory cassette (xpt-G15 double, SEQ ID NO.: 64; Example 8, Fig. 8a ), diluted in saline (Qiagen Endofree kit) was injected through tail vein in a volume of 10% body weight over 5 to 10 seconds to 6-7 weeks old CD-1 female mice. Guanosine (Sigma) was suspended in 0.5% methylcellulose/0.25% Tween 80 (Sigma) in water freshly and administrated orally at 2 hr and 12hr post DNA injection, or, delivered through intraperitoneal injection (IP) at 5 hr, 12 hr, 16 hr and 24 hr post DNA delivery.

[00204] *Noninvasive live animal bioluminescence imaging:* before imaging, mice were anesthesized with 2% isoflurane, and injected with 150 gm/kg body weight of luciferin, and images were taken within 2 to 5 minutes after luciferin injection using Bruker Xtreme system at the indicated time point post DNA injection. Luciferase activity was expressed as mean photon/sec ± s.d. The induction fold was calculated as the quotient of photon/sec obtained in mice treated with guanosine divided by value obtained in mice without guanosine treatment.

[00205] Results

[00206] We assessed the gene regulating function of the Intron-Exon- Intron regulatory cassette in mice, *in vivo*. Endotoxin-free plasmid DNA of a construct containing two copies of xpt-G15 riboswitch in the luciferase gene (xpt-G15 double) was delivered to the liver in mice through hydrodynamic injection, and guanosine was administrated intraperitoneally. We tested two routes of guanosine delivery. In one experiment (Fig 11a and 11b) mice were administered with different doses of guanosine orally 2 hr and 12hr after DNA delivery, then were imaged. As shown in Fig. 11a, mice treated with guanosine showed higher luciferase expression at 9 hours post DNA. The luciferase expression in guanosine treated mice increased over time reaching the highest level at 48 hours post DNA injection, after which expression declined.

[00207] In a separate experiment (Fig. 11c and Fig. 11d), guanosine was administered intraperitoneally. At 4 hours post DNA injection (P.I.) and before guanosine treatment, mice in each group showed similar level of basal luciferase activity (Figure 11). Then, mice were treated with either vehicle as control, or guanosine. At 11 hours P.I., luciferase activity increased in all the mice, consistent with the report that luciferase gene expression peaks 12 hours post hydrodynamic DNA injection in the liver. However, in mice treated with guanosine, there is a significantly higher level of luciferase expression compared to that in untreated mice, 4.7-fold and 16.2-fold induction compared to the uninduced baseline expression was seen with 100 mg/kg and 300mg/kg of guanosine, respectively.

[00208] Thus, the splicing-based gene regulation cassette was shown to regulate gene expression *in vivo* in animals, in a dose dependent manner, in response to the administration of the ligand specific for the aptamer contained within the regulatory cassette.

[00209] **EXAMPLE 12. Delivery of riboswitch constructs to the murine retina via Adeno Associated Viral (AAV) Vectors.**

[00210] **Experimental Procedures**

[00211] *AAV Plasmid constructs:* Two riboswitch expression constructs (described in the table below) were adapted via molecular cloning into a format able to be packaged as an AAV genome.

Table 4.

| Name | Riboswitch inducible element | Promotor | Transgene reporter |
|---|---|---|---|
| GTX5 | None (control) | CMV | Enhanced Green Fluorescent Protein (eGFP) |
| GTX7 | G15 | CMV | Enhanced Green Fluorescent Protein (eGFP) |

[00212] Expression constructs based around the EGFP transgene (GTX 5 – 7) were digested with restriction enzymes MfeI and NheI releasing an ~1400bp DNA fragment containing the Riboswitch inducible element and EGFP transgene. A pD10 AAV genome plasmid was also digested with MfeI and NheI, releasing a 4475bp fragment containing the AAV ITRs, CMV promotor and SV40 polyadenylation signal. The two fragments were ligated using T4 DNA ligase resulting in plasmids containing sequence with the following structure, able be packaged as an AAV2 genome:

[ITR]–[CMV]–[5' EGFP]–[Riboswitch Element]–[3' EGFP]–[SV40]–[ITR]

[00213] All resulting plasmid constructs were verified by DNA sequencing and named according to the following convention: pD10-GTX#.

[00214] *AAV Vector Production and titration:* Adeno-associated virus (AAV) was produced *in vitro* by transient transfection of HEK-293T cells with three plasmids.

(i) Viral Genomic plasmid based upon pD10 backbone

(ii) AAV Packaging plasmid containing the AAV2 Rep78 gene and a viral capsid gene. Many different serotypes of AAV can be produced by varying the capsid gene sequence but in this case an AAV8 capsid was used.

(iii) Helper plasmid (pHGTI-Adeno1). This plasmid provides a near minimal set of the Adenovirus genes that AAV requires to package and assemble.

[00215] These plasmids were transfected into HEK-293T cells in the ratio 1:1:3, with a total of 50 μg of plasmid DNA was transfected per 80-90% confluent 150 $cm^2$ plate. A typical production run consisted of 20 such plates. The transfection reagent used was polyethylenimine (PEI) at a PEI to DNA ratio of 2.25 : 1 (w/w). Seventy-two hours after transfection, cells were physically detached from the plates and pelleted by centrifugation; the resulting cell pellet was resuspended in 20mL of TRIS density buffer. The pellet was then lysed by repeated Freeze / Thaw / Vortex cycles and any non-packaged DNA remaining in the lysate was destroyed by Benzonase digestion. The lysate was then clarified by dead-end filtration and centrifugation before being diluted up to a total volume of 50 mL.

[00216] Clarified lysate was then purified via an affinity based FPLC procedure using an AVB column on an AKTA Pure instrument (both GE Healthcare) run according to pre-programmed protocols. The final AAV containing eluate from the FPLC column was concentrated down to a volume of ~200 μL by centrifugation at 5000 x g in a 10,000 MW cut off Vivaspin 4 centrifugal concentrator (GE Healthcare), 2 mL of PBS-MK was added (to dilute out high salt elution buffer), and the eluate re-concentrated back to ~200μL using the same concentrator. This material constituted the purified AAV virus, and was aliquoted as appropriate and stored at -80°C.

[00217] Vector titer was established using qPCR (targeted against the SV40 polyadenylation signal) directly upon a sample of purified vector. The resulting cycle threshold value was compared against a known standard curve and the number of vector genomes per mL was calculated.

[00218] Riboswitch AAV vectors were named according to the following convention: AAV2/[capsid serotype #]-GTX#

[00219] *Murine subretinal injections:* Injections of vector into the subretinal space were performed upon mice under general anesthesia using a manually guided 10mm, 34-gauge needle mounted on a 5μL Hamilton syringe. Needle tip was guided into injection position by observation of the retina via an operating microscope. In all eyes receiving vector, 2 x 2μL injections were performed, with one injection placed in the superior hemisphere of the eye and another in the inferior hemisphere. After injection, the quality of the resulting retinal detachment and any reflux of injected material was recorded.

[00220] *Fluorescent fundus photography:* Following subretinal injection, EGFP transgene expression was periodically assessed by Fundus photography using a slit lamp (SC-

16, Keeler) with an attached Leica DC500 digital camera. Animals were placed under general anesthesia and their pupils dilated with 1 % topical tropicamide. Corneal refractive power was neutralized by placing a coverslip on the cornea covered with a coupling medium solution (Viscotears). Under bright white light, the instrument was adjusted and the animal positioned so that the retina was in sharp focus and the optic disk was centered in the field of view, a bright field image was then taken using a 200ms exposure time. Transgene (EGFP) fluorescence was assessed by filtering the light source (475±25nm) and taking two further images with 10 and 30s exposures.

[00221] Results

[00222] The riboswitch constructs (Table 4) were successfully cloned into a format able to be packaged as an AAV genome as shown by DNA sequencing of the ligation products. Of the resulting constructs, pd10-GTX7 and pd10-GTX5 were further produced as AAV2/8 viral vectors. The vectors produced were shown to have the following titers by qPCR:

AAV2/8-GTX7: $1.17 \times 10^{13}$ Vector Genomes / mL

AAV2/8-GTX5: $1.73 \times 10^{13}$ Vector Genomes / mL

[00223] These two vectors were then injected subretinally and left for 8 days for EGFP transgene expression to develop before assessing expression by fluorescent fundus photography. Fig. 12 shows that EGFP is expressed in a retina injected with AAV2/8-GTX7. Transgene expression is low, but substantial expression from AAV2/8-GTX7 would only be expected after induction via aptamer-mediated alternative splicing in response to ligand (which was not added).

[00224] **EXAMPLE 13. Regulation of target gene expression by regulation cassette-mediated alternative splicing *in vivo* in the murine retina following AAV delivery.**

[00225] Procedures

[00226] *Quantification of Fluorescent fundus photography (EGFP signal):* All manipulation and analysis of images was performed using GNU Image Manipulation Program (GIMP, open source). As described above, three images of each retina were taken at each point of imaging: White light (200ms), 475±25nm (10s) and 475±25nm (30s). First these three images were superimposed as layers, and using the white light image as a guide, a

region of interest (ROI) was defined to encompass the entire retina visible through the pupil. Upon the two 475±25nm (EGFP fluorescence) images the threshold tool was used to highlight only those pixels with an intensity value above a defined threshold. The threshold value was selected upon the basis of defining clean separation of EGFP signal from background, and to provide an appropriate dynamic range for analysis. The number of pixels above threshold within the ROI was recorded for each image. To correct for variable dilation of the pupil leading to variation in the area of retina visible between eyes the number of pixels above threshold was divided by the total number of pixels within the ROI.

[00227] *Induction:* Riboswitch-mediated induction of target gene expression was carried out via two routes of administration as described below:

[00228] *Intraperitoneal injection (I.P.):* A volume of 100 µL of [75mg / ml guanosine + 0.5% w/v Methyl Cellulose + 0.25% v/v Tween 80 in Water] was injected into the intraperitoneal cavity using a 13mm, 30-gauge needle. This equates to a dose of guanosine of 300mg/kg in an adult mouse weighing 25g.

[00229] *Intravitreal injection (I-Vit.):* A volume of 2 µL of [1mM guanosine + 2.5% DMSO in PBS-MK] was injected intravitreally using a manually guided 10mm, 34-gauge needle. The needle tip position upon injection was below the lens directly above the optic disk, having been guided into this position by observation of the retina via an operating microscope.

[00230] Results

[00231] A total of 9 eyes were injected subretinally as described in Example 12 as follows on day 00:

- 6 eyes with AAV2/8-GTX7 (EGFP transgene expression from the CMV promotor regulated by the G15 riboswitch element)
- 3 eyes with AAV2/8-GTX5 (Positive control construct, unregulated EGFP transgene expression from the CMV promotor)

[00232] Fluorescent fundus photography as described in Example 12 was performed on days 02, 08, 09, 10 and 12.

[00233] All eyes received induction via Intraperitoneal injection after fluorescent fundus photography on days 08, 09 and 10. All eyes received induction via Intravitreal

injection on day 11. Fluorescent signal was quantified as described above and example images are shown in Fig. 13a.

[00234] No induction was carried out during the first 8 days post vector injection as gene expression from AAV2/8 is known to take up to 7 days to become maximal. The expression level on day 8 was therefore taken as the pre-induction base line. On day 10 post vector injection after 2 rounds of I.P. induction, transgene expression had increased by ~3.5x compared to this baseline ($P \leq 0.05$, 1-way ANOVA, Dunnetts) as shown in Fig. 13c and Fig. 13a (L vs N).

[00235] On day 12 post vector injection, 24hrs following intravitreal induction and 48 hrs after the last I.P. induction, transgene expression had increased by ~9 x compared to baseline ($P \leq 0.001$, 1-way ANOVA, Dunnetts) as shown in Fig. 13c and Fig. 13a (L vs O). This much larger induction following intravitreal induction implies (but does not definitively show) that this route of induction might be more effective than intraperitoneal injection.

[00236] Higher resolution images showing the difference in EGFP transgene expression pre and post induction are shown in Fig. 13b.

[00237] Over the same time period and under the same induction regime the EGFP expression levels mediated by the unregulated control vector AAV2/8-GTX5 did not vary significantly (1-way ANOVA, Bonferroni), remaining roughly constant as shown in Fig. 13d. Due to the large difference in expression level mediated by GTX7 vs GTX5, each set of images required a different exposure time (30s and 10s respectively) and threshold (50 and 190 respectively).

[00238] This data clearly shows that transgene expression from the G15 based GTX7 construct was being regulated via aptamer-mediated alternative splicing in the murine retina. The maximum level of transgene expression induced from GTX7 was lower than that mediated by the uninducable positive control construct GTX5

Table 5. Description and associated sequences. Exon sequence is in uppercase letters and intron sequence in in lowercase letters unless otherwise stated.

| SEQ ID NO. | Description | Sequence |
|---|---|---|

| SEQ ID NO. | Description | Sequence |
|---|---|---|
| SEQ ID NO.: 1 | Luci-IVSΔ-Luci in construct Con1<br>Human beta-globin intron 2 containing a deletion ("IVS2Δ") is in lowercase and flanking luciferase sequence is in uppercase. | GAGGTTCCATCTGCCAGGTATCAGGgtgagtctat gggacccttgatgttttctttccccttctttttctatggttaagttcatgtcatag gaaggggagaagtaacagggtacacatattgaccaaatcagggtaatttt gcatttgtaattttaaaaaatgctttcttctttttaatatacttttttgtttatcttattt ctaatactttccctaatctctttctttcagggcaataatgatacaatgtatcat gcctctttgcaccattctaaagaataacagtgataatttctgggttaaggca atagcaatatttctgcatataaatatttctgcatataaattgtaactgatgtaa gaggtttcatattgctaatagcagctacaatccagctaccattctgcttttatt ttatggttgggataaggctggattattctgagtccaagctaggccctttgct aatcatgttcatacctcttatcttcctcccacagCAAGGATATGG GCTCACTGAGACTACATCAGCTATTCT |
| SEQ ID NO.: 2 | Luci-IVSΔ-Luci in construct Con2 | GATTACACCCGAGGGGGATGATAAAGtaagcct atgggacccttgatgttttctttccccttctttttctatggttaagttcatgtcata ggaaggggagaagtaacagggtacacatattgaccaaatcagggtaatt ttgcatttgtaattttaaaaaatgctttcttctttttaatatacttttttgtttatcttat ttctaatactttccctaatctctttctttcagggcaataatgatacaatgtatca tgcctctttgcaccattctaaagaataacagtgataatttctgggttaaggc aatagcaatatttctgcatataaatatttctgcatataaattgtaactgatgta agaggtttcatattgctaatagcagctacaatccagctaccattctgctttta ttttatggttgggataaggctggattattctgagtccaagctaggccctttg ctaatcatgttcatacctcttatcttcctcccacagCCGGGCGCGG TCGGTAAAGT |
| SEQ ID NO.: 3 | Luci-IVSΔ-Luci in construct Con3 | TTCTTCGCCAAAAGCAgtaagtctatgggacccttgatgttt tctttccccttctttctatggttaagttcatgtcataggaaggggagaagta acagggtacacatattgaccaaatcagggtaattttgcatttgtaattttaaa aaatgctttcttctttttaatatacttttttgtttatcttatttctaatactttccctaat ctctttctttcagggcaataatgatacaatgtatcatgcctctttgcaccattc taaagaataacagtgataatttctgggttaaggcaatagcaatatttctgca tataaatatttctgcatataaattgtaactgatgtaagaggtttcatattgctaa tagcagctacaatccagctaccattctgcttttatttttatggttgggataagg ctggattattctgagtccaagctaggccctttgctaatcatgttcatacctct tatcttcctcccacagCTCTGATTGACAAATACG |
| SEQ ID NO.: 4 | Luci-IVSΔ-Luci in construct Con4 | AAGAGCTGTTTCTGAGGAGgtgtggctatgggaccctt gatgttttctttccccttctttttctatggttaagttcatgtcataggaaggggga gaagtaacagggtacacatattgaccaaatcagggtaattttgcatttgtaa ttttaaaaaatgctttcttctttttaatatacttttttgtttatcttatttctaatactttc cctaatctctttctttcagggcaataatgatacaatgtatcatgcctctttgca ccattctaaagaataacagtgataatttctgggttaaggcaatagcaatattt ctgcatataaatatttctgcatataaattgtaactgatgtaagaggtttcatatt gctaatagcagctacaatccagctaccattctgcttttatttttatggttgggat aaggctggattattctgagtccaagctaggccctttgctaatcatgttcata cctcttatcttcctcccacagCCTTCAGGATTACAAGATT CAA |

| SEQ ID NO. | Description | Sequence |
|---|---|---|
| SEQ ID NO.: 5 | Luci-IVSΔ-Luci in construct Con5 | CATCTGCCAGGTATCAGGgtgagtctatgggacccttga tgttttctttccccttctttttctatggttaagttcatgtcataggaaggggaga agtaacagggtacacatattgaccaaatcagggtaattttgcatttgtaattt taaaaaatgctttcttcttttaatatactttttgtttatcttatttctaatactttcc ctaatctctttctttcagggcaataatgatacaatgtatcatgcctctttgcac cattctaaagaataacagtgataatttctgggttaaggcaatagcaatatttc tgcatataaatatttctgcatataaattgtaactgatgtaagaggtttcatatt gctaatagcagctacaatccagctaccattctgcttttatttttatggttgggat aaggctggattattctgagtccaagctaggccctttttgctaatcatgttcata ccctttacttctatgactgtagCAAGGATATGGGCTCACT GAGACT |
| SEQ ID NO.: 6 | Luci-IVSΔ-Luci in construct Con6 | TCCATCTGCCAGGTATCAGGgtgagtctatgggaccct tgatgttttctttccccttctttttctatggttaagttcatgtcataggaaggggga gaagtaacagggtacacatattgaccaaatcagggtaattttgcatttgtaa ttttaaaaaatgctttcttcttttaatatactttttgtttatcttatttctaatactttc cctaatctctttctttcagggcaataatgatacaatgtatcatgcctctttgca ccattctaaagaataacagtgataatttctgggttaaggcaatagcaatattt ctgcatataaatatttctgcatataaattgtaactgatgtaagaggtttcatatt gctaatagcagctacaatccagctaccattctgcttttatttttatggttgggat aaggctggattattctgagtccaagctaggccctttttgctaatcatgttcata ccgtgactgtgtgtatgcacagCAAGGATATGGGCTCAC TGAGACT |
| SEQ ID NO.: 7 | Luci-IVSΔ-Luci in construct Con7 | ATCTGCCAGGTATCAGGgtgagtctatgggacccttgatg ttttctttccccttctttttctatggttaagttcatgtcataggaaggggagaag taacagggtacacatattgaccaaatcagggtaattttgcatttgtaatttta aaaaatgctttcttcttttaatatactttttgtttatcttatttctaatactttccct aatctctttctttcagggcaataatgatacaatgtatcatgcctctttgcacca ttctaaagaataacagtgataatttctgggttaaggcaatagcaatatttctg catataaatatttctgcatataaattgtaactgatgtaagaggtttcatattgct aatagcagctacaatccagctaccattctgcttttatttttatggttgggataa ggctggattattctgagtccaagctaggccctttttgctaatcatgttcatacc attgtgatcgcagccaatagCAAGGATATGGGCTCACT GAGACT |
| SEQ ID NO.: 8 | DHFR exon 2 with flanking intronic sequence | gagtaacgctgtttctctaacttgtagGAATGAATTCAGATA TTTCCAGAGAATGACCACAACCTCTTCAGTA GAAGgtaatgtg |
| SEQ ID NO.: 9 | Theophylline aptamer | ggcgataccagccgaaaggcccttggcagcgtc |
| SEQ ID NO.: 10 | Xpt-guanine aptamer | cactcatataatcgcgtggatatggcacgcaagtttctaccgggcaccgt aaatgtccgactatgggtg |
| SEQ ID NO.: 11 | Ydhl-Guanine aptamer | ttgtataacctcaataatatggtttgagggtgtctaccaggaaccgtaaaat cctgactacaa |
| SEQ ID NO.: 12 | Ydhl-Adenine aptamer | ttgtataacctcaataatatggtttgagggtgtctaccaggaaccgtaaaat cctgattacaa |

| SEQ ID NO. | Description | Sequence |
|---|---|---|
| SEQ ID NO.: 13 | addA-Guanine aptamer | tcatataatcctaatgatatggtttgggagtttctaccaagagccttaaactc ttgactatga |
| SEQ ID NO.: 14 | addA-Adenine aptamer | tcatataatcctaatgatatggtttgggagtttctaccaagagccttaaactc ttgattatga |
| SEQ ID NO.: 15 | xpt-G17 riboswitch<br>The aptamer is underlined and the stem is double underlined.<br>A modified DHFR exon 2 is in capital letters. | gtgagtctatgggacccttgatgttttctttccccttctttttctatggttaagttc atgtcataggaaggggagaagtaacagggtacacatattgaccaaatca gggtaattttgcatttgtaatttttaaaaaatgctttcttctttttaatatacttttttg tttatcttatttctaatactttccctaatctctttctttcagggcaataatgataca atgtatcatgccgagtaacgctgtttctctaacttgtagGAATGAAT TCAGATATTTCCAGAGAATGAAAAAAAAAT CTTCAGTAGAAGgtaatgtataatcgcgtggatatggcacgca agtttctaccgggcaccgtaaatgtccgactacattacgcaccattctaaa gaataacagtgataatttctgggttaaggcaatagcaatatttctgcatata aatatttctgcatataaattgtaactgatgtaagaggtttcatattgctaatag cagctacaatccagctaccattctgcttttatttatggttgggataaggctg gattattctgagtccaagctaggccctttgctaatcatgttcatacctcttat cttcctcccacag |
| SEQ ID NO.: 16 | xpt-G17-IR-1 | gtgagtctatgggacccttgatgttttctttccctgctcaaatcagggtaattt tgcatttgtaattttaaaaaatgctttcttctttttaatatacttttttgtttatcttatt tctaatactttccctaatctctttctttcagggcaataatgatacaatgtatcat gccgagtaacgctgtttctctaacttgtagGAATGAATTCAGA TATTTCCAGAGAATGAAAAAAAAATCTTCA GTAGAAGgtaatgtataatcgcgtggatatggcacgcaagtttcta ccgggcaccgtaaatgtccgactacattacgcaccattctaaagaataac agtgataatttctgggttaaggcaatagcaatatttctgcatataaatatttct gcatataaattgtaactgatgtaagaggtttcatattgctaatagcagctac aatccagctaccattctgcttttatttatggttgggataaggctggattattct gagtccaagctaggccctttgctaatcatgttcatacctcttatcttcctccc acag |
| SEQ ID NO.: 17 | xpt-G17-IR-2 | gtgagtctatgggacccttgatgttttctttccctgctctttcagggcaataat gatacaatgtatcatgccgagtaacgctgtttctctaacttgtagGAAT GAATTCAGATATTTCCAGAGAATGAAAAAA AAATCTTCAGTAGAAGgtaatgtataatcgcgtggatatgg cacgcaagtttctaccgggcaccgtaaatgtccgactacattacgcacca ttctaaagaataacagtgataatttctgggttaaggcaatagcaatatttctg catataaatatttctgcatataaattgtaactgatgtaagaggtttcatattgct aatagcagctacaatccagctaccattctgcttttatttatggttgggataa ggctggattattctgagtccaagctaggccctttgctaatcatgttcatacc tcttatcttcctcccacag |

| SEQ ID NO. | Description | Sequence |
|---|---|---|
| SEQ ID NO.: 18 | xpt-G17-IR-3 | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtgctcaaatcagggtaattttgcatttgtaatttttaaaaaatgctttcttctt ttaatatacttttttgtttatcttatttctaatactttccctaatctctttctttcagg gcaataatgatacaatgtatcatgccgagtaacgctgtttctctaacttgta gGAATGAATTCAGATATTTCCAGAGAATGAA AAAAAAATCTTCAGTAGAAGgtaatgtataatcgcgtg gatatggcacgcaagtttctaccgggcaccgtaaatgtccgactacatta cgcaccattctaaagaataacagtgataatttctgggttaaggcaatagca atatttctgcatataaatatttctgcatataaattgtaactgatgtaagaggttt catattgctaatagcagctacaatccagctaccattctgcttttattttatggtt gggataaggctggattattctgagtccaagctaggccctttgctaatcatg ttcatacctcttatcttcctcccacag |
| SEQ ID NO.: 19 | xpt-G17-IR-4 | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtgctctttcagggcaataatgatacaatgtatcatgccgagtaacgct gtttctctaacttgtagGAATGAATTCAGATATTTCCA GAGAATGAAAAAAAAATCTTCAGTAGAAGgt aatgtataatcgcgtggatatggcacgcaagtttctaccgggcaccgtaa atgtccgactacattacgcaccattctaaagaataacagtgataatttctgg gttaaggcaatagcaatatttctgcatataaatatttctgcatataaattgtaa ctgatgtaagaggtttcatattgctaatagcagctacaatccagctaccatt ctgcttttattttatggttgggataaggctggattattctgagtccaagctag gccctttgctaatcatgttcatacctcttatcttcctcccacag |
| SEQ ID NO.: 20 | xpt-G17-IR-5 | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtcataggaaggggagaagtaacagggtactgctcaaatcagggtaa ttttgcatttgtaattttaaaaaatgctttcttcttttaatatacttttttgtttatctt atttctaatactttccctaatctctttctttcagggcaataatgatacaatgtat catgccgagtaacgctgtttctctaacttgtagGAATGAATTCA GATATTTCCAGAGAATGAAAAAAAAATCTT CAGTAGAAGgtaatgtataatcgcgtggatatggcacgcaagttt ctaccgggcaccgtaaatgtccgactacattacgcaccattctaaagaat aacagtgataatttctgggttaaggcaatagcaatatttctgcatataaatat ttctgcatataaattgtaactgatgtaagaggtttcatattgctaatagcagct acaatccagctaccattctgcttttattttatggttgggataaggctggattat tctgagtccaagctaggccctttgctaatcatgttcatacctcttatcttcct cccacag |
| SEQ ID NO.: 21 | xpt-G17-IR-6 | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtcataggaaggggagaagtaacagggtactgctctttcagggcaat aatgatacaatgtatcatgccgagtaacgctgtttctctaacttgtagGA ATGAATTCAGATATTTCCAGAGAATGAAAA AAAAATCTTCAGTAGAAGgtaatgtataatcgcgtggat atggcacgcaagtttctaccgggcaccgtaaatgtccgactacattacgc accattctaaagaataacagtgataatttctgggttaaggcaatagcaatat ttctgcatataaatatttctgcatataaattgtaactgatgtaagaggtttcata ttgctaatagcagctacaatccagctaccattctgcttttattttatggttggg ataaggctggattattctgagtccaagctaggccctttgctaatcatgttca tacctcttatcttcctcccacag |

| SEQ ID NO. | Description | Sequence |
|---|---|---|
| SEQ ID NO.: 22 | xpt-G17-IR-7 | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtcataggaagtgctcaaatcagggtaattttgcatttgtaattttaaaaa atgctttcttcttttaatatacttttttgtttatcttatttctaatactttccctaatct ctttctttcagggcaataatgatacaatgtatcatgccgagtaacgctgtttc tctaacttgtagGAATGAATTCAGATATTTCCAGAG AATGAAAAAAAAATCTTCAGTAGAAGgtaatgt ataatcgcgtggatatggcacgcaagtttctaccgggcaccgtaaatgtc cgactacattacgcaccattctaaagaataacagtgataatttctgggttaa ggcaatagcaatatttctgcatataaatatttctgcatataaattgtaactgat gtaagaggtttcatattgctaatagcagctacaatccagctaccattctgctt ttattttatggttgggataaggctggattattctgagtccaagctaggccctt ttgctaatcatgttcatacctcttatcttcctcccacag |
| SEQ ID NO.: 23 | xpt-G17-IR-8 | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtcataggaagtgctctttcagggcaataatgatacaatgtatcatgcc gagtaacgctgtttctctaacttgtagGAATGAATTCAGATA TTTCCAGAGAATGAAAAAAAAATCTTCAGT AGAAGgtaatgtataatcgcgtggatatggcacgcaagtttctaccg ggcaccgtaaatgtccgactacattacgcaccattctaaagaataacagt gataatttctgggttaaggcaatagcaatatttctgcatataaatatttctgca tataaattgtaactgatgtaagaggtttcatattgctaatagcagctacaatc cagctaccattctgcttttattttatggttgggataaggctggattattctgag tccaagctaggccctttgctaatcatgttcatacctcttatcttcctcccaca g |
| SEQ ID NO.: 24 | xpt-G17-IR-9 | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtcataggaaggggagaagtaacagggtacacatattgaccaaatca gggtaattttgcatttgtaattttaaaaaatgctttcttcttttaatatacttttttg tttatcttatttctaatactttccctaatctctttctttcagggcaataatgataca atgtatcatgccgagtaacgctgtttctctaacttgtagGAATGAAT TCAGATATTTCCAGAGAATGAAAAAAAAAT CTTCAGTAGAAGgtaatgtataatcgcgtggatatggcacgca agtttctaccgggcaccgtaaatgtccgactacattacgcaccattctaaa gaataacagtgataatttctgggttaaggcaatagctgctgctggattattc tgagtccaagctaggccctttgctaatcatgttcatacctcttatcttcctcc cacag |
| SEQ ID NO.: 25 | xpt-G17-IR-10 | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtcataggaaggggagaagtaacagggtacacatattgaccaaatca gggtaattttgcatttgtaattttaaaaaatgctttcttcttttaatatacttttttg tttatcttatttctaatactttccctaatctctttctttcagggcaataatgataca atgtatcatgccgagtaacgctgtttctctaacttgtagGAATGAAT TCAGATATTTCCAGAGAATGAAAAAAAAAT CTTCAGTAGAAGgtaatgtataatcgcgtggatatggcacgca agtttctaccgggcaccgtaaatgtccgactacattacgcaccattctaaa gaataacagtgataatttctgggttaaggcaatagcaatatttctgcatata aatatttctgcatataaattgtaactgatgtaagaggtttcatattgctaatag cagctacaatccagctgctgctggattattctgagtccaagctaggccctttt gctaatcatgttcatacctcttatcttcctcccacag |

| SEQ ID NO. | Description | Sequence |
|---|---|---|
| SEQ ID NO.: 26 | xpt-G17-IR-11 | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtcataggaaggggagaagtaacagggtacacatattgaccaaatca gggtaattttgcatttgtaatttttaaaaaatgctttcttcttttaatatacttttttg tttatcttatttctaatactttccctaatctctttctttcagggcaataatgataca atgtatcatgccgagtaacgctgtttctctaacttgtagGAATGAAT TCAGATATTTCCAGAGAATGAAAAAAAAAT CTTCAGTAGAAGgtaatgtataatcgcgtggatatggcacgca agtttctaccgggcaccgtaaatgtccgactacattacgcaccattctaaa gaataacagtgataatttctgggttaaggcaatagctgctgaggtttcatat tgctaatagcagctacaatccagctaccattctgcttttattttatggttggga taaggctggattattctgagtccaagctaggccctttgctaatcatgttcat acctcttatcttcctcccacag |
| SEQ ID NO.: 27 | xpt-G17-IR-13 | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtcataggaaggggagaagtaacagggtacacatattgaccaaatca gggtaattttgcatttgtaatttttaaaaaatgctttcttcttttaatatacttttttg tttatcttatttctaatactttccctaatctctttctttcagggcaataatgataca atgtatcatgccgagtaacgctgtttctctaacttgtagGAATGAAT TCAGATATTTCCAGAGAATGAAAAAAAAAT CTTCAGTAGAAGgtaatgtataatcgcgtggatatggcacgca agtttctaccgggcaccgtaaatgtccgactacattacgcaccattctaaa gaataacagtgataatttctgggttaaggcaatagctgctctaccattctgc ttttattttatggttgggataaggctggattattctgagtccaagctaggccc ttttgctaatcatgttcatacctcttatcttcctcccacag |
| SEQ ID NO.: 28 | xpt-G17-IR-15 | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtcataggaaggggagaagtaacagggtacacatattgaccaaatca gggtaattttgcatttgtaatttttaaaaaatgctttcttcttttaatatacttttttg tttatcttatttctaatactttccctaatctctttctttcagggcaataatgataca atgtatcatgccgagtaacgctgtttctctaacttgtagGAATGAAT TCAGATATTTCCAGAGAATGAAAAAAAAAT CTTCAGTAGAAGgtaatgtataatcgcgtggatatggcacgca agtttctaccgggcaccgtaaatgtccgactacattacgcaccattctaaa gaataacagtgataatttctgggttaaggcaatagctgctgcagctacaat ccagctaccattctgcttttattttatggttgggataaggctggattattctga gtccaagctaggccctttgctaatcatgttcatacctcttatcttcctcccac ag |
| SEQ ID NO.: 29 | xpt-G17-2IR-1 | gtgagtctatgggacccttgatgttttctttccctgctcaaatcagggtaattt tgcatttgtaatttttaaaaaatgctttcttcttttaatatacttttttgtttatcttatt tctaatactttccctaatctctttctttcagggcaataatgatacaatgtatcat gccgagtaacgctgtttctctaacttgtagGAATGAATTCAGA TATTTCCAGAGAATGAAAAAAAAATCTTCA GTAGAAGgtaatgtataatcgcgtggatatggcacgcaagtttcta ccgggcaccgtaaatgtccgactacattacgcaccattctaaagaataac agtgataatttctgggttaaggcaatagctgctgctggattattctgagtcc aagctaggccctttgctaatcatgttcatacctcttatcttcctcccacag |

| SEQ ID NO. | Description | Sequence |
|---|---|---|
| SEQ ID NO.: 30 | xpt-G17-2IR-2 | gtgagtctatgggacccttgatgttttctttccctgctcaaatcagggtaatttgcatttgtaattttaaaaaatgctttcttcttttaatatacttttttgtttatcttatttctaatactttccctaatctctttctttcagggcaataatgatacaatgtatcatgccgagtaacgctgtttctctaacttgtagGAATGAATTCAGATATTTCCAGAGAATGAAAAAAAAATCTTCAGTAGAAGgtaatgtataatcgcgtggatatggcacgcaagtttctaccgggcaccgtaaatgtccgactacattacgcaccattctaaagaataacagtgataatttctgggttaaggcaatagcaatatttctgcatataaatatttctgcatataaattgtaactgatgtaagaggtttcatattgctaatagcagctacaatccagctgctgctggattattctgagtccaagctaggccctttgctaatcatgttcatacctcttatcttcctcccacag |
| SEQ ID NO.: 31 | xpt-G17-2IR-3 | gtgagtctatgggacccttgatgttttctttccctgctcaaatcagggtaatttgcatttgtaattttaaaaaatgctttcttcttttaatatacttttttgtttatcttatttctaatactttccctaatctctttctttcagggcaataatgatacaatgtatcatgccgagtaacgctgtttctctaacttgtagGAATGAATTCAGATATTTCCAGAGAATGAAAAAAAAATCTTCAGTAGAAGgtaatgtataatcgcgtggatatggcacgcaagtttctaccgggcaccgtaaatgtccgactacattacgcaccattctaaagaataacagtgataatttctgggttaaggcaatagctgctgaggtttcatattgctaatagcagctacaatccagctaccattctgcttttatttttatggttgggataaggctggattattctgagtccaagctaggccctttgctaatcatgttcatacctcttatcttcctcccacag |
| SEQ ID NO.: 32 | xpt-G17-2IR-4 | gtgagtctatgggacccttgatgttttctttccctgctcaaatcagggtaatttgcatttgtaattttaaaaaatgctttcttcttttaatatacttttttgtttatcttatttctaatactttccctaatctctttctttcagggcaataatgatacaatgtatcatgccgagtaacgctgtttctctaacttgtagGAATGAATTCAGATATTTCCAGAGAATGAAAAAAAAATCTTCAGTAGAAGgtaatgtataatcgcgtggatatggcacgcaagtttctaccgggcaccgtaaatgtccgactacattacgcaccattctaaagaataacagtgataatttctgggttaaggcaatagctgctctaccattctgcttttatttttatggttgggataaggctggattattctgagtccaagctaggccctttgctaatcatgttcatacctcttatcttcctcccacag |
| SEQ ID NO.: 33 | xpt-G17-2IR-5 | gtgagtctatgggacccttgatgttttctttccctgctcaaatcagggtaatttgcatttgtaattttaaaaaatgctttcttcttttaatatacttttttgtttatcttatttctaatactttccctaatctctttctttcagggcaataatgatacaatgtatcatgccgagtaacgctgtttctctaacttgtagGAATGAATTCAGATATTTCCAGAGAATGAAAAAAAAATCTTCAGTAGAAGgtaatgtataatcgcgtggatatggcacgcaagtttctaccgggcaccgtaaatgtccgactacattacgcaccattctaaagaataacagtgataatttctgggttaaggcaatagctgctgcagctacaatccagctaccattctgcttttatttttatggttgggataaggctggattattctgagtccaagctaggccctttgctaatcatgttcatacctcttatcttcctcccacag |

| SEQ ID NO. | Description | Sequence |
|---|---|---|
| SEQ ID NO.: 34 | xpt-G17-2IR-6 | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtgctcaaatcagggtaattttgcatttgtaattttaaaaaatgctttcttctt ttaatatacttttttgtttatcttatttctaatactttccctaatctctttctttcagg gcaataatgatacaatgtatcatgccgagtaacgctgtttctctaacttgta gGAATGAATTCAGATATTTCCAGAGAATGAA AAAAAAATCTTCAGTAGAAGgtaatgtataatcgcgtg gatatggcacgcaagtttctaccgggcaccgtaaatgtccgactacatta cgcaccattctaaagaataacagtgataatttctgggttaaggcaatagct gctgctggattattctgagtccaagctaggccctttgctaatcatgttcata cctcttatcttcctcccacag |
| SEQ ID NO.: 35 | xpt-G17-2IR-7 | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtgctcaaatcagggtaattttgcatttgtaattttaaaaaatgctttcttctt ttaatatacttttttgtttatcttatttctaatactttccctaatctctttctttcagg gcaataatgatacaatgtatcatgccgagtaacgctgtttctctaacttgta gGAATGAATTCAGATATTTCCAGAGAATGAA AAAAAAATCTTCAGTAGAAGgtaatgtataatcgcgtg gatatggcacgcaagtttctaccgggcaccgtaaatgtccgactacatta cgcaccattctaaagaataacagtgataatttctgggttaaggcaatagca atatttctgcatataaatatttctgcatataaattgtaactgatgtaagaggttt catattgctaatagcagctacaatccagctgctgctggattattctgagtcc aagctaggccctttgctaatcatgttcatacctcttatcttcctcccacag |
| SEQ ID NO.: 36 | xpt-G17-2IR-8 | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtgctcaaatcagggtaattttgcatttgtaattttaaaaaatgctttcttctt ttaatatacttttttgtttatcttatttctaatactttccctaatctctttctttcagg gcaataatgatacaatgtatcatgccgagtaacgctgtttctctaacttgta gGAATGAATTCAGATATTTCCAGAGAATGAA AAAAAAATCTTCAGTAGAAGgtaatgtataatcgcgtg gatatggcacgcaagtttctaccgggcaccgtaaatgtccgactacatta cgcaccattctaaagaataacagtgataatttctgggttaaggcaatagct gctgaggtttcatattgctaatagcagctacaatccagctaccattctgcttt tattttatggttgggataaggctggattattctgagtccaagctaggcccttt tgctaatcatgttcatacctcttatcttcctcccacag |
| SEQ ID NO.: 37 | xpt-G17-2IR-9 | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtgctcaaatcagggtaattttgcatttgtaattttaaaaaatgctttcttctt ttaatatacttttttgtttatcttatttctaatactttccctaatctctttctttcagg gcaataatgatacaatgtatcatgccgagtaacgctgtttctctaacttgta gGAATGAATTCAGATATTTCCAGAGAATGAA AAAAAAATCTTCAGTAGAAGgtaatgtataatcgcgtg gatatggcacgcaagtttctaccgggcaccgtaaatgtccgactacatta cgcaccattctaaagaataacagtgataatttctgggttaaggcaatagct gctctaccattctgcttttattttatggttgggataaggctggattattctgagt ccaagctaggccctttgctaatcatgttcatacctcttatcttcctcccaca g |

| SEQ ID NO. | Description | Sequence |
|---|---|---|
| SEQ ID NO.: 38 | xpt-G17-2IR-10 | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtgctcaaatcagggtaattttgcatttgtaattttaaaaaatgctttcttctt ttaatatacttttttgtttatcttatttctaatactttccctaatctctttctttcagg gcaataatgatacaatgtatcatgccgagtaacgctgtttctctaacttgta gGAATGAATTCAGATATTTCCAGAGAATGAA AAAAAAATCTTCAGTAGAAGgtaatgtataatcgcgtg gatatggcacgcaagtttctaccgggcaccgtaaatgtccgactacatta cgcaccattctaaagaataacagtgataatttctgggttaaggcaatagct gctgcagctacaatccagctaccattctgcttttattttatggttgggataag gctggattattctgagtccaagctaggccctttgctaatcatgttcatacct cttatcttcctcccacag |
| SEQ ID NO.: 39 | xpt-G17-2IR-11 | gtgagtctatgggacccttgatgttttctttccctgctctttcagggcaataat gatacaatgtatcatgccgagtaacgctgtttctctaacttgtagGAAT GAATTCAGATATTTCCAGAGAATGAAAAAA AAATCTTCAGTAGAAGgtaatgtataatcgcgtggatatgg cacgcaagtttctaccgggcaccgtaaatgtccgactacattacgcacca ttctaaagaataacagtgataatttctgggttaaggcaatagctgctgctgg attattctgagtccaagctaggccctttgctaatcatgttcatacctcttatct tcctcccacag |
| SEQ ID NO.: 40 | xpt-G17-2IR-12 | gtgagtctatgggacccttgatgttttctttccctgctctttcagggcaataat gatacaatgtatcatgccgagtaacgctgtttctctaacttgtagGAAT GAATTCAGATATTTCCAGAGAATGAAAAAA AAATCTTCAGTAGAAGgtaatgtataatcgcgtggatatgg cacgcaagtttctaccgggcaccgtaaatgtccgactacattacgcacca ttctaaagaataacagtgataatttctgggttaaggcaatagctgctctacc attctgcttttattttatggttgggataaggctggattattctgagtccaagct aggccctttgctaatcatgttcatacctcttatcttcctcccacag |
| SEQ ID NO.: 41 | xpt-G17-2IR-13 | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtgctctttcagggcaataatgatacaatgtatcatgccgagtaacgct gtttctctaacttgtagGAATGAATTCAGATATTTCCA GAGAATGAAAAAAAAATCTTCAGTAGAAGgt aatgtataatcgcgtggatatggcacgcaagtttctaccgggcaccgtaa atgtccgactacattacgcaccattctaaagaataacagtgataatttctgg gttaaggcaatagctgctgctggattattctgagtccaagctaggccctttt gctaatcatgttcatacctcttatcttcctcccacag |
| SEQ ID NO.: 42 | xpt-G17-2IR-14 | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtgctctttcagggcaataatgatacaatgtatcatgccgagtaacgct gtttctctaacttgtagGAATGAATTCAGATATTTCCA GAGAATGAAAAAAAAATCTTCAGTAGAAGgt aatgtataatcgcgtggatatggcacgcaagtttctaccgggcaccgtaa atgtccgactacattacgcaccattctaaagaataacagtgataatttctgg gttaaggcaatagctgctctaccattctgcttttattttatggttgggataagg ctggattattctgagtccaagctaggccctttgctaatcatgttcatacctct tatcttcctcccacag |

| SEQ ID NO. | Description | Sequence |
|---|---|---|
| SEQ ID NO.: 43 | xpt-G17-2IR-15 | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtcataggaaggggagaagtaacagggtactgctctttcagggcaat aatgatacaatgtatcatgccgagtaacgctgtttctctaacttgtagGA ATGAATTCAGATATTTCCAGAGAATGAAAA AAAAATCTTCAGTAGAAGgtaatgtataatcgcgtggat atggcacgcaagtttctaccgggcaccgtaaatgtccgactacattacgc accattctaaagaataacagtgataatttctgggttaaggcaatagctgctg ctggattattctgagtccaagctaggccctttgctaatcatgttcatacctct tatcttcctcccacag |
| SEQ ID NO.: 44 | xpt-G17-2IR-16 | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtcataggaaggggagaagtaacagggtactgctctttcagggcaat aatgatacaatgtatcatgccgagtaacgctgtttctctaacttgtagGA ATGAATTCAGATATTTCCAGAGAATGAAAA AAAAATCTTCAGTAGAAGgtaatgtataatcgcgtggat atggcacgcaagtttctaccgggcaccgtaaatgtccgactacattacgc accattctaaagaataacagtgataatttctgggttaaggcaatagctgctc taccattctgcttttattttatggttgggataaggctggattattctgagtcca agctaggccctttgctaatcatgttcatacctcttatcttcctcccacag |
| SEQ ID NO.: 45 | xpt-G17-3ssC-1 | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtcataggaaggggagaagtaacagggtacacatattgaccaaatca gggtaattttgcatttgtaattttaaaaaatgctttcttcttttaatatacttttttg tttatcttatttctaatactttccctaatctctttctttcagggcaataatgataca atgtatcatgccgagtaacgctgtttctctaacttccccGAATGAAT TCAGATATTTCCAGAGAATGAAAAAAAAAT CTTCAGTAGAAGgtaatgtataatcgcgtggatatggcacgca agtttctaccgggcaccgtaaatgtccgactacattacgcaccattctaaa gaataacagtgataatttctgggttaaggcaatagcaatatttctgcatata aatatttctgcatataaattgtaactgatgtaagaggtttcatattgctaatag cagctacaatccagctaccattctgcttttattttatggttgggataaggctg gattattctgagtccaagctaggccctttgctaatcatgttcatacctcttat cttcctcccacag |
| SEQ ID NO.: 46 | xpt-G15 riboswitch | gtgagtctatgggacccttgatgttttctttccccttcttttctatggttaagttc atgtcataggaaggggagaagtaacagggtacacatattgaccaaatca gggtaattttgcatttgtaattttaaaaaatgctttcttcttttaatatacttttttg tttatcttatttctaatactttccctaatctctttctttcagggcaataatgataca atgtatcatgccgagtaacgctgtttctctaacttgtagGAATGAAT TCAGATATTTCCAGAGAATGAAAAAAAAAT CTTCAGTAGAAGgtaatgtgtataatcgcgtggatatggcacg caagtttctaccgggcaccgtaaatgtccgactacacattacgcaccattc taaagaataacagtgataatttctgggttaaggcaatagcaatatttctgca tataaatatttctgcatataaattgtaactgatgtaagaggtttcatattgctaa tagcagctacaatccagctaccattctgcttttattttatggttgggataagg ctggattattctgagtccaagctaggccctttgctaatcatgttcatacctct tatcttcctcccacag |
| SEQ ID NO.: 47 | DHFR WildType 5' single-strand | gagtaacgctgtttctctaacttgtagGAATGAATTCAGATA TTTCCAGAGAATGACCACAACCTCTTCAGTA GAAGgtaatgtg |

| SEQ ID NO. | Description | Sequence |
|---|---|---|
| SEQ ID NO.: 48 | DHFR 5' single-strandC | gagtaacgctgtttctctaacttgtagGAATGAATTCAGATA TTTCCAGAGAATGACCACAACCTCTTCAGTA GAAGcccctgtg |
| SEQ ID NO.: 49 | DHFR-Con1 5' single-strand | gagtaacgctgtttctctaacttgtagGAATGAATTCAGATA TTTCCAGAGAATGACCACAACCTCTTCAGTA GAGGgtgagttg |
| SEQ ID NO.: 50 | DHFR-Con4 5' single-strand | gagtaacgctgtttctctaacttgtagGAATGAATTCAGATA TTTCCAGAGAATGACCACAACCTCTTCAGTA GGAGgtgtggtg |
| SEQ ID NO.: 51 | DHFR WildType mtSRp40 | gagtaacgctgtttctctaacttgtagGAATGAATTCAGATA TTTCCAGAGAATGAAAAAAAAATCTTCAGT AGAAGgtaatgtg |
| SEQ ID NO.: 52 | DHFR WildType strongSC35 | gagtaacgctgtttctctaacttgtagGAATGGCCCCTGATA TTTCCAGAGAATGACCACAACCTCTTCAGTA GAAGgtaatgtg |
| SEQ ID NO.: 53 | DHFR WildType SC35hnRNPA1 | gagtaacgctgtttctctaacttgtagGAATGTAGGGAGATA TTTCCAGAGAATGACCACAACCTCTTCAGTA GAAGgtaatgtg |
| SEQ ID NO.: 54 | DHFR-Con1 5' single-strand HP15 | gagtaacgctgtttctctaacttgtagGAATGAATTCAGATA TTTCCAGAGAATGACCACAACCTCTTCAGTA GAGGgtgagttggcgaaagccaactcaccctct |
| SEQ ID NO.: 55 | DHFR-Con1 5' single-strand HP15x | gagtaacgctgtttctctaacttgtagGAATGAATTCAGATA TTTCCAGAGAATGACCACAACCTCTTCAGTA GAGGgtgagttggcgaaaaacagcataaagtat |
| SEQ ID NO.: 56 | DHFR WildType mtSRp40 HP15 | gagtaacgctgtttctctaacttgtagGAATGAATTCAGATA TTTCCAGAGAATGAAAAAAAAATCTTCAGT AGAAGgtaatgtggcgaaagccacattaccttct |
| SEQ ID NO.: 57 | DHFR WildType mtSRp40 HP15X | gagtaacgctgtttctctaacttgtagGAATGAATTCAGATA TTTCCAGAGAATGAAAAAAAAATCTTCAGT AGAAGgtaatgtggcgaaaaacagaactgagtat |
| SEQ ID NO.: 58 | Mutant DHFR-e2 (mtDHFR) | gagtaacgctgtttctctaacttgtagGAATGAATTCAGATA TTTCCAGAGAATGAAAAAAAAATCTTCAGT AGAAGgtaatgt |
| SEQ ID NO.: 59 | Camk2d-e16 | gagtaacgctgtttctctaacttgtagTGAGCCCCAAACTAC TGTAATCCACAACCCTGACGGAAACAAGgtaa tgt |
| SEQ ID NO.: 60 | Camk2d-e17 | gagtaacgctgtttctctaacttgtagGAGTCAACTGAGAGC TCAAACACCACCATTGAGGATGAAGACGTG AAAGgtaatgt |
| SEQ ID NO.: 61 | mutWT1-e5 | gagtaacgctgtttctctaacttgtagAGTTGCTGCTGAGAG CTCCAGCTCAGTGAAATGGACAGAAGGGCA GAGCAAgtaatgt |
| SEQ ID NO.: 62 | SIRT1-e6 | tgtggtgtgttcaagaaacagaaatacttctttaataaagcatatatatgttgt ttgtttttagGTTCCTTTGCAACAGCATCTTGCCTG ATTTGTAAATACAAAGTTGACTGTGAAGCTG TACGAGGAGATATTTTTAATCAGgtaatgt |

| SEQ ID NO. | Description | Sequence |
|---|---|---|
| SEQ ID NO.: 63 | ENEEE synthetic exon | gagtaacgctgtttctctaacttgtagACAATCCTCGAACCA AACAACCAAACAACCAAACAATCCTCGAAC CAAACAATCCTCGAACCAAACAATCCTCGA ACCAAgtaatgt |
| SEQ ID NO.: 64 | xpt-G15-double | GCCAAGAGGTTCCATCTGCCAGGTATCAGGg tgagtctatgggacccttgatgttttctttccccttctttctatggttaagttca tgtcataggaaggggagaagtaacagggtacacatattgaccaaatcag gtaattttgcatttgtaattttaaaaaatgctttcttctttttaatatacttttttgtt tatcttatttctaatactttccctaatctctttctttcagggcaataatgatacaa tgtatcatgccgagtaacgctgtttctctaacttgtagGAATGAATT CAGATATTTCCAGAGAATGAAAAAAAAATC TTCAGTAGAAGgtaatgtgtataatcgcgtggatatggcacgc aagtttctaccgggcaccgtaaatgtccgactacacattacgcaccattct aaagaataacagtgataatttctgggttaaggcaatagcaatatttctgcat ataaatatttctgcatataaattgtaactgatgtaagaggtttcatattgctaat agcagctacaatccagctaccattctgcttttatttatggttgggataaggc tggattattctgagtccaagctaggccctttgctaatcatgttcatacctctt atcttcctcccacagCAAGGATATGGGCTCACTGAG ACTACATCAGCTATTCTGATTACACCCGAGG GGGATGATAAACCGGGCGCGGTCGGTAAAG TTGTTCCATTTTTTGAAGCGAAGGTTGTGGA TCTGGATACCGGGAAAACGCTGGGCGTTAA TCAAAGAGGCGAACTGTGTGTGAGAGGTCC TATGATTATGTCCGGTTATGTAAACAATCCG GAAGCGACCAACGCCTTGATTGACAAGGAT GGATGGCTACATTCTGGAGACATAGCTTACT GGGACGAAGACGAACACTTCTTCATCGTTGA CCGCCTGAAGTCTCTGATTAAGTACAAAGGg tgagtctatgggacccttgatgttttctttccccttctttctatggttaagttca tgtcataggaaggggagaagtaacagggtacacatattgaccaaatcag gtaattttgcatttgtaattttaaaaaatgctttcttctttttaatatacttttttgtt tatcttatttctaatactttccctaatctctttctttcagggcaataatgatacaa tgtatcatgccgagtaacgctgtttctctaacttgtagGAATGAATT CAGATATTTCCAGAGAATGAAAAAAAAATC TTCAGTAGAAGgtaatgtgtataatcgcgtggatatggcacgc aagtttctaccgggcaccgtaaatgtccgactacacattacgcaccattct aaagaataacagtgataatttctgggttaaggcaatagcaatatttctgcat ataaatatttctgcatataaattgtaactgatgtaagaggtttcatattgctaat agcagctacaatccagctaccattctgcttttatttatggttgggataaggc tggattattctgagtccaagctaggccctttgctaatcatgttcatacctctt atcttcctcccacagCTATCAGGTGGCTCCCGCTGAA TTGGAATCCATCTTGCTCC |

Claims:

1. A polynucleotide cassette for the regulation of the expression of a target gene comprising

a. a riboswitch

b. an alternatively-spliced exon, flanked by a 5′ intron and a 3′ intron,

wherein the riboswitch comprises (i) an effector region comprising a stem having 7 to 20 base pairs that includes the 5′ splice site sequence of the 3′ intron, and (ii) an aptamer, wherein:

the 5' splice site sequence of the 3' intron comprises the sequence GTAATG or GTRAGT, wherein R is A or G; and

the alternatively-spliced exon comprises a stop codon that is in-frame with the target gene when the alternatively-spliced exon is spliced into the target gene mRNA.

2. The polynucleotide cassette of claim 1, wherein the aptamer binds a small molecule ligand.

3. The polynucleotide cassette of claim 1 or 2 for use in the regulation of the expression of the target gene in therapy.

4. The polynucleotide cassette for use of claim 3, wherein

a. the polynucleotide cassette of claim 1 or 2 is to be inserted into a target gene,

b. the target gene comprising the polynucleotide cassette of claim 1 or 2 is to be introduced into a cell, and

c. the cell is to be exposed to a small molecule ligand that specifically binds the aptamer in an amount effective to induce expression of the target gene.

5. The polynucleotide cassette for use of claim 4, wherein the polynucleotide cassette is to be inserted into the protein coding region of the target gene.

6. The polynucleotide cassette for use of claim 4 or 5, wherein the expression of the target gene is greater than 5-fold higher, or is greater than 10-fold higher when the small molecule ligand is present than the expression levels when the small molecule ligand is absent.

7. The polynucleotide cassette for use of any one of claims 4-6, wherein two or more of the polynucleotide cassettes are to be inserted into the target gene.

8. The polynucleotide cassette for use of claim 7, wherein the two or more polynucleotide cassettes comprise different aptamers that specifically bind to different small molecule ligands.

9. The polynucleotide cassette for use of claim 7, wherein the two or more polynucleotide cassettes comprise the same aptamer.

10. The polynucleotide cassette for use of any one of claims 4 to 9, wherein the target gene comprising the polynucleotide cassette is to be incorporated in a vector for the expression of the target gene.

11. A vector comprising a target gene that contains a polynucleotide cassette, for use in the regulation of the expression of the target gene in therapy, wherein:

a. the vector is to be introduced into the eye; and

b. a small molecule ligand in an amount effective to induce expression of the target gene is to be provided to the mammal;

wherein the polynucleotide cassette comprises (i) a riboswitch and (ii) an alternatively-spliced exon flanked by a 5′ intron and a 3′ intron, wherein the riboswitch comprises an effector region comprising a stem having 7 to 20 base pairs that includes the 5′ splice site sequence of the 3′ intron, and an aptamer that binds a small molecule ligand, wherein:

i. the 5' splice site sequence of the 3' intron comprises the sequence GTAATG or GTRAGT, wherein R is A or G; and

ii. the alternatively-spliced exon comprises a stop codon that is in-frame with the target gene.

12. The vector for use of claim 11, wherein the vector is to be introduced into the eye by intraocular injection.

13. The vector for use of claim 11 or 12, wherein the vector is a viral vector.

14. The vector for use of claim 13, wherein the viral vector is selected from the group consisting of adenoviral vector, adeno-associated virus vector, and lentiviral vector.

15. The polynucleotide cassette for use of any one of claims 3 to 10, or the vector for use of any one of claims 11 to 14, wherein the polynucleotide cassette is located in the protein coding sequence of the target gene.

16. The polynucleotide cassette of claim 1 or 2, or the polynucleotide cassette for use of any one of claims 3 to 10 and 15, or the vector for use of any one of claims 11 to 15, wherein the 5′ and 3′ introns

(i) are derived from an endogenous intron from the target gene;

(ii) are exogenous to the target gene; or

(iii) are derived from intron 2 of the human β-globin gene.

17. The polynucleotide cassette of claim 1, 2 or 16, or the polynucleotide cassette for use of any one of claims 3 to 10, 15 and 16, or the vector for use of any one of claims 11 to 16, wherein the 5′ intron comprises a stop codon in-frame with the target gene.

18. The polynucleotide cassette of claim 1, 2, 16 or 17, or the polynucleotide cassette for use of any one of claims 3 to 10 and 15 to 17, or the vector for use of any one of claims 11 to 17, wherein the 5′ and 3′ introns are each independently from 50 to 300 nucleotides in length.

19. The polynucleotide cassette, or the polynucleotide cassette for use, or the vector for use of claim 18, wherein the 5′ and 3′ introns are each independently from 125 to 240 nucleotides in length.

20. The polynucleotide cassette of any one of claims 1, 2, and 16-19, or the polynucleotide cassette for use of any one of claims 3 to 10 and 15 to 19, or the vector for use of any one of claims 11 to 19, wherein the effector region stem is 8 to 11 base pairs in length.

21. The polynucleotide cassette of any one of claims 1, 2, and 16-20, or the polynucleotide cassette for use of any one of claims 3 to 10 and 15 to 20, or the vector for use of any one of claims 11 to 20, wherein the alternatively-spliced exon

(i) is derived from exon 2 of the human dihydrofolate reductase gene, mutant human Wilms tumor 1 exon 5, mouse calcium/calmodulin-dependent protein kinase II delta exon 16, or SIRT1 exon 6;

(ii) has been modified by one or more of the group consisting of altering the sequence of an exon splice enhancer, altering the sequence of exon splice silencer, adding an exon splice enhancer, and adding an exon splice silencer;

(iii) is the modified DHFR exon 2 from SEQ ID NO:15; or

(iv) is synthetic.

22. A recombinant polynucleotide comprising a target gene containing the polynucleotide cassette of any one of claims 1 to 2 and 16 to 21.

23. The recombinant polynucleotide of claim 22, wherein the polynucleotide cassette is located in the protein coding sequence of the target gene.

24. A vector comprising a target gene that contains a polynucleotide cassette of any one of claims 1 to 2 and 16 to 21.

25. The vector of claim 24, wherein the vector is a viral vector.

26. The vector of claim 25, wherein the viral vector is selected from the group consisting of adenoviral vector, adeno-associated virus vector, and lentiviral vector.

27. An *in vitro* method of modulating the expression of a target gene, the method comprising:

a. inserting the polynucleotide cassette of claim 1 or 2 into a target gene,

b. introducing the target gene comprising the polynucleotide cassette into a cell, and

c. exposing the cell to a small molecule ligand that specifically binds the aptamer in an amount effective to induce expression of the target gene.