Supporting Information

Interrogation of Mirror-Image L-RNA-Protein Interactions Reveals Key Mechanisms of Single-Stranded G-Rich L-RNA Cytotoxicity and a Potential Mitigation Strategy

Chen-Hsu Yu[1], Xiaomei He[2], Rosemarie Elloisa P. Acero[1,†], Xuan Han[1,†], Yinsheng Wang[2], and Jonathan T. Sczepanski[1,3,*]

[1] Department of Chemistry, Texas A&M University, College Station, Texas, 77843, USA

[2] Department of Chemistry, University of California Riverside, Riverside, California, 92521-0403, USA

[3] Department of Biochemistry and Biophysics, Texas A&M University, College Station, Texas, 77843, USA

[†]These authors contributed equally to this work.

S1. Supplementary Text.

MATERIALS AND METHODS

General. Solid-phase oligonucleotide synthesis was carried out on an Expedite 8909 DNA/RNA Synthesizer using manufacturer recommended protocols. D-Nucleoside phosphoramidites, CPG solid supports, and all oligonucleotide synthesis reagents were purchased from Glen Research (Sterling, Va). L-Nucleoside phosphoramidites were purchased from ChemGenes (Wilmington, Ma). EZ-Link NHS-LC-Biotin, Lipofectamine 3000, cell culture media (DMEM, OptiMEM, SILAC kit) and all antibodies were purchased from ThermoFisher, (Waltham, MA). Sulfo-Cy5 N-hydroxysuccinimide (NHS) ester was purchased from Lumiprobe Corp. (Hunt Valley, MD). Camptothecin were purchased from Abcam (Cambridge, United Kingdom). Purelink RNA Mini Kit was purchased from Invitrogen (San Diego, CA). Recombinant nucleolin (NCL) was purchased from Antibodies.com (St. Louis, MO; Cat. # ABIN7455408). Recombinant FUS was purchased from Raybiotech (Peachtree Corners, GA; Cat. # 230-01094-10). Recombinant heterogeneous nuclear ribonucleoprotein R (HNRNPR) was purchased from Sinobiological (Houston, TX; Cat. # 14309-H20B). All commercial proteins were used without further purification. All other reagents were purchased from Sigma Aldrich (St. Louis, MO).

Oligonucleotide purification and labeling. All oligonucleotides used in this study are shown in Table S1. All oligonucleotides were purified by denaturing polyacrylamide gel electrophoresis (PAGE; 20% 19:1 acrylamide:bisacrylamide) prior to use. The band corresponding to the correct oligonucleotide product was cut from the gel and eluted overnight at room temperature in a buffer consisting of 200 mM NaCl, 10 mM EDTA, and 10 mM Tris (pH 7.6). Gel fragments were removed from the solution by filtration, and eluted oligonucleotides were concentrated using a 3 kDa pore size Amicon Ultra Centrifugal Filter device (MilliporeSigma, Burlington, MA). The concentrated samples were desalted by ethanol precipitation, and the concentration of the oligonucleotide was determined by absorbance at 260 nm on a NanoDrop 2000c (ThermoFisher, Waltham, MA). The identity of all oligonucleotides was confirmed using a Thermo Scientific Q Exactive Focus ESI mass spectrometer (Figure S22 – S32).

Oligonucleotides for the SILAC-proteomics experiment (Figure 1) were synthesized from a 3′-Biotin-TEG CPG resin. All other oligonucleotides were synthesized with a 3′-PT-Amino-Modifier C6 CPG resin (Glen Research Co., Sterling, VA). Sulfo-Cy5 N-hydroxysuccinimide (NHS) ester (Lumiprobe Corp., Hunt Valley, MD) or EZ-Link NHS-LC-

Biotin (ThermoFisher, Waltham, MA) was conjugated to the 3′ end of the indicated L-RNA via a 3′ amino modification installed at the time of synthesis. The conjugation reaction has been described in our previous work.[1]

Cell culture and oligonucleotide transfection. HeLa cells were obtained from ATCC (Manassas, VA) and were cultured in Dulbecco's Modified Eagle's Medium (DMEM; Thermo Fisher Scientific) supplemented with 10 mM HEPES, 1 mM GlutaMax (Thermo Fisher), and 10% fetal bovine serum (FBS; Thermo Fisher). Metabolic incorporation of heavy amino acids was carried out in HeLa cells using the SILAC Protein Quantification Kit (ThermoFisher #A33972).[2] [$^{13}C_6{}^{15}N_2$]-L-Lysine•2HCl and [$^{13}C_6{}^{15}N_4$]-L-Arginine•2HCl were incorporated as the heavy amino acids. HeLa cells were cultured in complete heavy SILAC media for 10 cell-doublings to ensure complete labeling. All cells were maintained at 37 °C in a humidified CO_2 (5%) atmosphere.

For oligonucleotide delivery, cells were plated at 5000 cells/well (HeLa) in a 96-well plate one day prior to transfection. Oligonucleotide transfection was carried out using Lipofectamine 3000 (Thermo Fisher Scientific) according to manufacturer's protocol. Transfection solutions were prepared by adding the required amount of a concentrated solution of oligonucleotide (10 μM) to a mixture containing 0.5 μL Lipofectamine 3000 and 4.5 μL Opti-MEM. This mixture was further diluted to 30 μL using Opti-MEM. The 30 μL transfection mixture was then added into each well already containing 70 μL of DMEM. Cells were maintained at 37 °C in a humidified CO_2 (5%) atmosphere for 2 hours following transfection. The transfection mixture was subsequently removed and each well was washed with 2 × 100 μL fresh DMEM and incubated under 100 μL DMEM for the indicated duration of the experiment.

Nuclear protein lysate generation. Cells cultured in SILAC medium in T75 flask were harvested at 80% confluency. Cells were washed with PBS twice and detached using a 0.25% trypsin solution at 37°C for 3 minutes. DMEM supplemented with 10% FBS was then added to quench trypsin and cells were pelleted by centrifugation. Cell pellets were washed with PBS twice and nuclear lysate was extracted using NE-PER nuclear and cytoplasmic extraction reagents (Thermo Fisher Scientific) following the manufacturer's guidelines. The protein concentrations were measured by absorbance at 280 nm on a NanoDrop 2000c (ThermoFisher, Waltham, MA), and the nuclear lysate was stored at -80 °C until use.

Pull-down of L-oligonucleotide binding proteins. For each pull-down experiment, 60 µL Dynabeads MyOne Streptavidin C1 (ThermoFisher Scientific) was used. The beads were washed with Solution A (0.1 M NaOH 0.05 M NaCl) twice and Solution B (0.1 M NaCl) once according to manufacturer's protocol. The beads were then resuspended in 60 µL binding buffer (10 mM Tris-HCl, pH 7.5, 150 mM KCl, 1.5 mM $MgCl_2$, 0.05% (v/v) IGEPAL-CA620, 0.5 mM DTT). 30 µL washed beads were added to 500 µg nuclear lysate in 1 mL binding buffer supplement with 1× Halt Protease Inhibitor Cocktail (ThermoFisher Scientific) to pre-clear the nuclear lysate. The mixture was incubated at 4 °C with rocking for 1 hour. Beads were removed and 50 pmol biotinylated RNA probe (Table S1) was added into each pre-cleared lysate and incubated at 4 °C with rocking for 2 hours. 30 µL washed beads were added into each binding mixture and incubated at 4 °C with rocking for 2 hours. After the incubation, the supernatant was removed and beads were washed three times with 1 mL binding buffer supplemental with increasing concentrations of NaCl (50, 100, and 200 mM). After the final washing, the beads were resuspended in 500 µL PBS. The corresponding light and heavy beads were combined into a new 1.7 mL tube. The PBS was removed and the bound proteins were eluted from beads by boiling in 30 µL of 2× SDS-PAGE loading buffer (100 mM Tris-Cl (pH 6.8), 4% (w/v) SDS, 0.2% (w/v) bromophenol blue, 20% (v/v) glycerol, 200 mM DTT) for 10 minutes (vortexing every 5 minutes). The resulting supernatant was saved into a new tube and stored at -20°C until use. Triplicate forward pull-downs and one reverse pull-down were performed for proteomics studies. In the forward direction, heavy nuclear lysates were pulled down with L-r$(GA)_{20}$. In the reverse direction, light nuclear lysates were pulled down with L-r$(GA)_{20}$.

In-gel digestion. Eluted proteins were resolved by 15% SDS-PAGE for approximately 1 cm and the gel was subsequently stained with Coomassie blue. Band with a molecular weight above 20 kDa were excised for in-gel tryptic digestion, as described previously.[3] Briefly, the gel band was cut into 1 mm^3 cubes and de-stained by shaking in 25% and 50% acetonitrile in 50 mM ammonium bicarbonate repeatedly until gel becomes colorless. Cysteine reduction and alkylation were conducted by incubating the gel pieces in 10 mM DTT at 37 °C for 1 hour and 55 mM iodoacetamide in the dark at room temperature for 20 minutes, respectively. Proteins were digested with 500 ng trypsin (Thermo Scientific) at 37 °C for 16 hours. The tryptic peptides were eluted by shaking the gel pieces at 37 °C in 5% formic acid, 50% acetonitrile/5% formic acid and 70% acetonitrile/5% formic acid, successively, for 15 minutes each. The resulting peptide mixtures were desalted with C_{18} ziptip and reconstituted in 2% acetonitrile/0.1% formic acid for LC-MS/MS analysis.

LC-MS/MS analysis. The LC-MS/MS experiments were conducted on a Q Exactive Plus Quadrupole-Orbitrap Mass Spectrometer (Thermo Fisher) coupled with an Ultimate 3000 UPLC (Thermo Fisher). Desalted peptides were resuspended in 15 μL of 2% acetonitrile/0.1% formic acid. Peptides were loaded onto a 3-cm capillary column (150 μm i.d.) packed in-house with C_{18} resin (5 μm in particle size and 120 Å in pore size, Dr. Maisch GmbH HPLC) with buffer A (0.1% formic acid in water) at a flow rate of 3 μL/min. The peptides were eluted from the trapping column and separated on a 25-cm analytical column (5 μm in particle size and 120 Å in pore size, Dr. Maisch GmbH HPLC) packed in-house with C_{18} resin (3 μm in particle size and 120 Å in pore size, Dr. Maisch GmbH HPLC) at a flow rate of 300 nL/min using a linear gradient of 5–37% buffer B (80% acetonitrile and 0.1% formic acid in water) over 135 minutes.

Peptides eluted from the analytical column were ionized with a Flex nanoelectrospray ion source (Thermo Fisher). The spray voltage and the capillary inlet temperature were set at 2 kV and 325 °C, respectively. Full-scan MS in the *m/z* range of 350-1200 were acquired at a resolution of 35 k. Maximal injection time for full-scan MS was set to 100 ms with an AGC of $1e^6$. Representative mass spectra of tryptic peptides from FUS, NCL, HNRNPR, P54nrb, SFPQ, RBM15 and ALYREF are depicted in Figure S2–S8. For MS/MS acquisition, top 25 precursor ions were isolated at a width of 1.6 m/z unit and subsequently fragmented by higher-energy collisional dissociation (HCD) with a normalized collision energy (NCE) of 28. MS/MS were recorded at a resolution of 17.5 k. Maximal ion injection time for MS/MS was 75 ms with an AGC of $1e^5$. The dwell time was 3 seconds, the minimal AGC to trigger MS/MS acquisition was set at $1e^3$, and the duration for dynamic exclusion was set at 35 seconds.

Proteomic data processing. Raw LC-MS/MS data were processed in MaxQuant (version 2.1.2.0).[4] MaxQuant multiplicity was set to 2, and [$^{13}C_6$,$^{15}N_2$]-L-lysine and [$^{13}C_6$,$^{15}N_4$]-L-arginine were selected as heavy amino acids. Methionine oxidation and N-terminal acetylation were set as variable modifications; cysteine carbamidomethylation was specified as a fixed modification. Mass tolerance for full-scan MS was set as 20 and 4.5 ppm for the first and main search, respectively. The mass tolerance for MS/MS was set at 20 ppm. Two missed cleavages were allowed for trypsin. Peptide spectra were searched against target-decoy Uniprot human proteome database (UP000005640_9606) and the proteins were subsequently filtered at 1% false discovery rate (FDR). The “match between runs” option was enabled with alignment windows and minimum protein ratio counts being 0.7 minutes and 2.0 minutes, respectively. The potential contamination and the decoys were removed from the output files. Triplicate forward and one reverse replicate were

processed by limma (version 3.54.2)[5] on Rstudio (version 1.4.1717) with R version 4.1.0. Proteins with adjusted p value <0.05 and $\log_2$ fold change (heavy/light) >1.5 or <-1.5 were considered as enriched. To identify the pathway that related with the enriched proteins, Gene Ontology (GO) enrichment of enriched proteins was carried out using clusterProfiler (version 4.0.5).[6] Identified GO terms with p-value <0.05 were considered as significant enriched. Data are available via ProteomeXchange with identifier PXD054780.

RNA Binding Domain (RBD) analysis. Amino acid sequences of enriched proteins were acquired from Uniprot. RBD analysis was conducted using NCBI batch-CD search and referenced to the pfam database.[7] Classical and nonclassical RNA binding domains from pfam accession under specific hit type was quantified for the enriched proteins.

Immunofluorescence. Immunofluorescence staining was carried out as described previously.[1] The identity of the antibodies used in this study, as well as their working dilutions, are listed in Table S2. Confocal images were obtained using a Leica SP8 confocal microscope using a HC PL APO 40×/1.10 W motCORR CS2 water immersion objective in conjunction with a 405 nm CW laser and a 470nm–670 nm white pulsed laser. Data analysis was conducted in ImageJ (v1.53q). Data analysis was conducted in ImageJ (v1.53q). The freehand selection and measure feature in ImageJ was used to determine the mean area fluorescence intensity within the nucleus and nucleolus. Using freehand selection, nucleus areas were selected based on Hoechst signal; nucleoli areas were selected based on fibrillarin signal. The selected areas were then applied to the protein of interest channel by the ROI manager. The fluorescence intensity was measured using the measure function in ROI manager. Mean fluorescence ratio of nucleoli/nucleus was calculated by the following equation.

$$\text{Mean fluorescence ratio} = \frac{\frac{\sum A_{nucleolus} \times I_{nucleolus}}{\sum A_{nucleolus}}}{I_{nucleus}}$$

Where $A_{nucleolus}$ represents the area of each nucleolus; $I_{nucleolus}$ represents the measured mean fluorescence intensity of each nucleolus; $I_{nucleus}$ represents the mean fluorescence intensity of each nucleus. All images were acquired under identical microscopy settings for a given experiment.

Western blotting. Pull-down experiments were carried out as described above for the proteomics experiments using the indicated 3′-biotinylated L-RNA (Figure 3c and Table S1). Proteins enriched during the pull-down were resolved on a 10% SDS-PAGE. Proteins

were transferred to a nitrocellulose membrane using a Trans-Blot Turbo Transfer System (Bio-rad). The membrane was blocked with 5% non-fat milk in PBST (blocking solution) for 1 hour on a rocker at room temperature. The membrane was then rinsed with 3× 4 mL PBST (1 minute each) and incubated with the primary antibody (Table S2) in blocking solution overnight at 4 °C on a rocker. The membrane was then further washed with 3× 4 mL PBST (5 minutes each) and incubated with the secondary antibody in blocking solution for 1 hour at room temperature on a rocker. Following an additional wash with 3× 4 mL PBST (5 minutes each), the membrane was imaged by fluorescence emission (Cy5; excitation/emission: 635 nm/670 nm) using Typhoon FLA9500 Multimode Imager (GE Healthcare Lifesciences).

Alternative splicing analysis. RNA-seq trimmed reads were obtained from our previous work (GEO accession number GSE205338).[1] The clean reads of all 15 samples were aligned to human reference genome (hg38) using STAR (version 2.7.7a) with 2 pass mode for higher accuracy at splicing junctions.[8] Triplicate bam files were then analyzed by rMATs (version 4.2.0)[9-10] for differential alternative splicing counting. The reference human genome file and gene annotation file were both obtained from the ENSEMBL database. Both genome alignment and alternative splicing analysis were conducted using the high-performance research computing resources provided by Texas A&M University (http://hprc.tamu.edu) in the Linux operating system. Significant alternative splicing events were determined by filtering events with FDR <0.05 and percent spliced in (PSI) changes >0.2 or <-0.2. Alternative splicing events were visualized by rmats2sashimiplot (version 3.0.0)[11] to generate the sashimi plot for specific exons (Figure S15).

***K*-mer analysis.** Sequences 150 bp upstream and 150 bp downstream of all differential skipped exons were extracted from human reference genome (hg38). *K*-mer analysis was carried out on these 150 bp sequences by Python package fastas2kmers with n set as 5. Exon skipping control was obtained from The Cancer Genome Atlas Program (TCGA) Exon Skipping list in the Exon Skipping annotation Database.[12]

RT-PCR analysis of exon skipping. HeLa cells were transfected with 200 nM of the indicated L-RNA or Lipofectamine 3000 only using the transfection protocol describe above. Where indicated, cells were also incubated with 10 μM camptothecin or 5 μM As^{3+}. All experiments were conducted in biological triplicate. After a 12-hour-treatment, total RNA was extracted using PureLink RNA Mini Kit (Invitrogen) and quantified using NanoDrop 2000c. Total RNA (1 μg) was reverse transcribed into cDNA using High-

Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Waltham, MA). The cDNA was amplified in a 20 µL PCR containing 0.02 U/µL Phusion High-Fidelity DNA Polymerase (NEB, Ipswich, MA), 1X Phusion GC buffer, 0.5 µM forward and 0.5 µM reverse primers, 3% (v/v) DMSO, and 50 µM each of the four dNTPs (Table S3). All PCRs were conducted with the following program: i) 98°C for 30s, (ii) 98°C for 10 s, (iii) 50°C for 30 s, (iv) 72°C for 1 min, and (v) 72°C for 10 min, with repeating steps (ii) to (iv) for N cycles as needed. The amplified transcripts were resolved by 2% agarose gel and imaged by fluorescence emission (EtBr; excitation/emission: 532 nm/605 nm) using Typhoon FLA9500 Multimode Imager and quantified using ImageQuant TL software (Cytiva).

Electrophoretic mobility shift assays (EMSA). The binding affinity of oligonucleotides for proteins was determined by EMSA (e.g., Figure S9b). The indicated concentration of protein or lysate was incubated with dye-labeled (Cy3/Cy5) D-RNA or L-RNA probe (Table S1) in a 10 µL binding mixture containing 50 mM Tris pH 7.4, 100 mM KCl, 2.5 mM $MgCl_2$, 0.1 mM $ZnCl_2$, 2 mM BME, 0.1 mg/mL BSA, and 5% glycerol. 1 nM oligonucleotide was used in EMSAs for individual proteins and 200 nM oligonucleotide was used in EMSAs for cell lysate. The binding reactions were incubated at 30 °C for 30 minutes then resolved by 0.5% agarose gel (0.2 × TBE supplemented with 10 mM KOAc), which were run at 50 mV for 1 hour at 4 °C. The gel was imaged by fluorescence emission (Cy5; excitation/emission: 635 nm/670 nm or Cy3; excitation/emission: 532 nm/570 nm) using Typhoon FLA9500 Multimode Imager and quantified using ImageQuant TL software (Cytiva).

Toxicity assays. The indicated concentrations of L-RNA were transfected into cells as described above using the same amount of Lipofectamine 3000 for a given experiment. After 2 hours, the cells were washed with 2 × 100 µL fresh DMEM and incubated under 100 µL DMEM for 46 hours at 37°C. The Cell Counting Kit-8 reagent (CCK-8; Sigma Aldrich) was then added to the cells. After a 1-hour incubation at 37°C, the absorbance at 450 nm was measured on a GloMax Discover multi-well plate reader (Promega Corp., Madison, WI). A control well containing only the CCK-8 reagent in DMEM was used as the background signal, which was subtracted from each sample. Viability data was normalized to HeLa cells treated with Lipofectamine 3000 only. The half maximal effective concentration (EC_{50}) was determined using the equation:

$$Viability\ (\%) = \frac{100}{(1 + \frac{Concentration\ (nM)}{EC_{50}})}$$

Cellular uptake analysis. Cy5-labeled L-RNAs (Table S1) were transfected into HeLa cells following the protocol described above. The media was removed 2-hours later, and the cells were washed with 2 × 50 µL PBS and detached from the well using a 0.25% trypsin solution (50 µL) at 37°C for 3 minutes. At this point, DMEM (50 µL) supplemented with 10% FBS was added to quench the trypsin and the fluorescence signal corresponding to L-RNA uptake was measured on an Accuri C6 Flow Plus Cytometer (BD Biosciences, San Jose, CA) using the FL4–APC filter (Ex: 640 nm; Em:BP 675/25 nm). A flow rate of 66 µL per minute was used for data collection and a minimum of 10,000 events were collected for each experiment.

Statistical Analysis. Statistical analysis was carried out on GraphPad Prism (v.9.3.1). An unpaired one-way analysis of variance (ANOVA) was used to compare datasets within a given experiment and a Tukey's multiple comparisons test ($\alpha_{threshold} = 0.05$) was used to compare significant differences between each condition.

Data Availability. Raw proteomics data was submitted to ProteomeXchange under the accession number PXD054780:
Username: reviewer_pxd054780@ebi.ac.uk; Password: sRRItAU3e00o
The complete list of enriched proteins can be accessed in Supplementary File 1.

S2. Supplementary Figures.

Figure S1

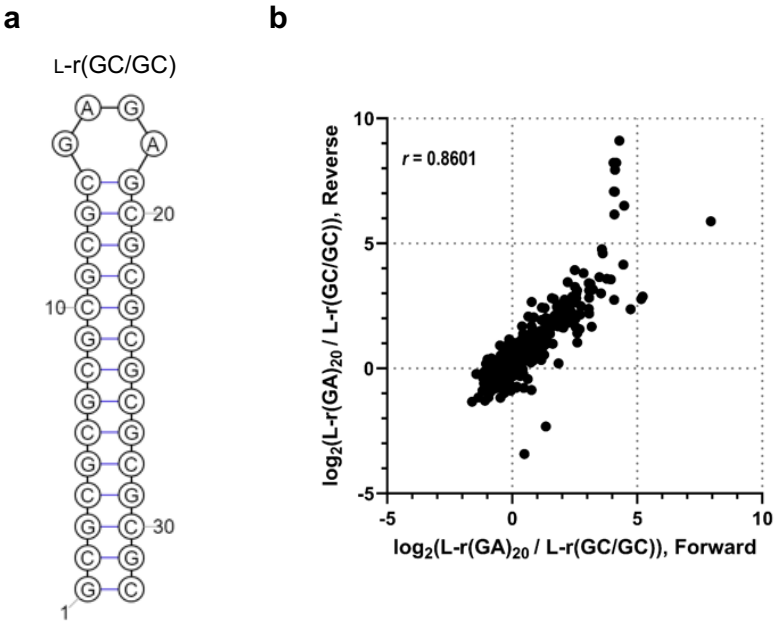

Figure S1. (a) The sequence and secondary structure of hairpin L-r(GC/GC). (b) A scatter plot showing the correlation of protein quantification results obtained from forward and reverse SILAC pull-down assays.

Figure S2

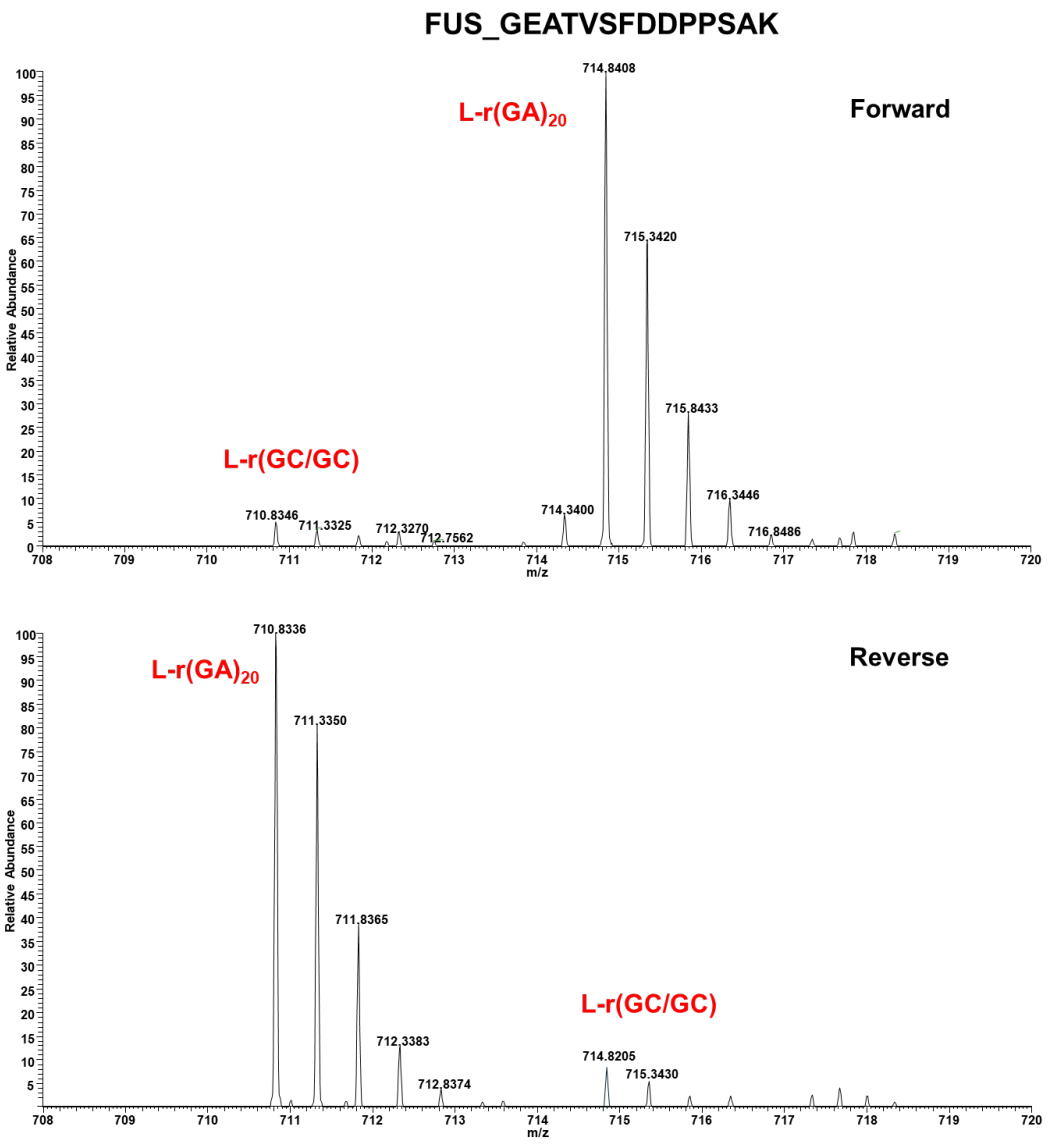

Figure S2. ESI-MS showing the $[M + 2H]^{2+}$ ions of light and heavy lysine-containing peptides, GEATVSFDDPPSAK, derived from FUS in forward and reverse SILAC experiments.

Figure S3

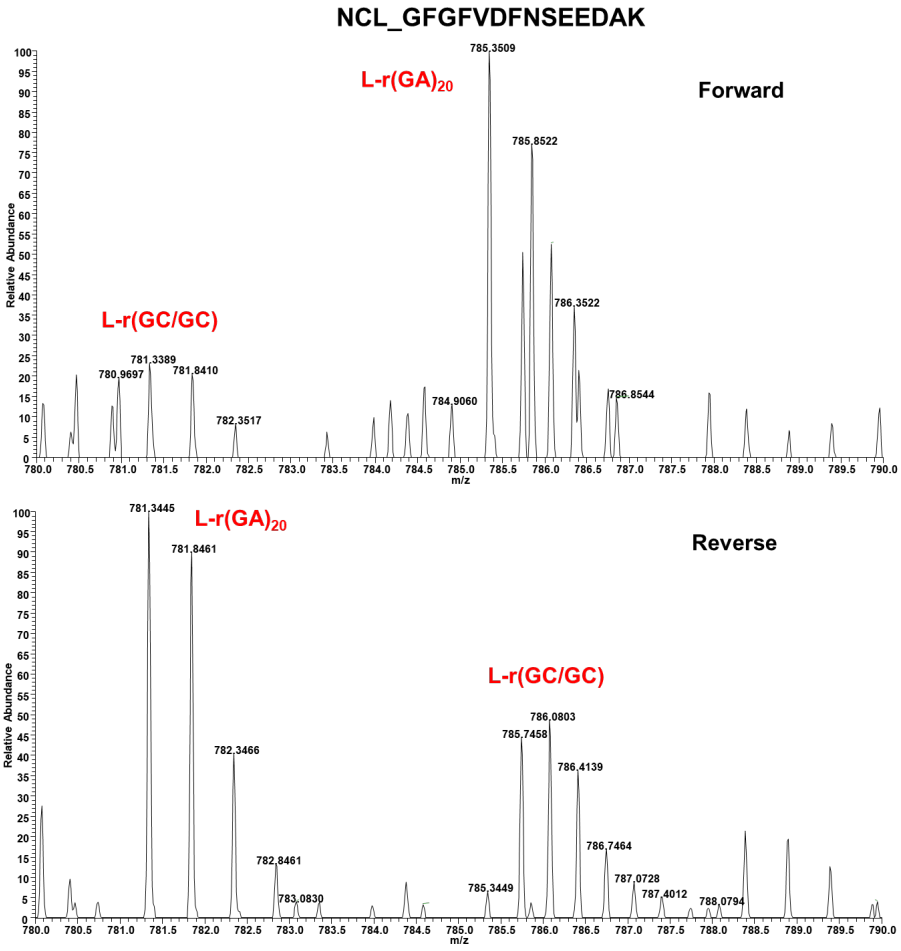

Figure S3. ESI-MS showing the $[M + 2H]^{2+}$ ions of light and heavy lysine-containing peptides, GFGFVDFNSEEDAK, derived from NCL in forward and reverse SILAC experiments.

Figure S4

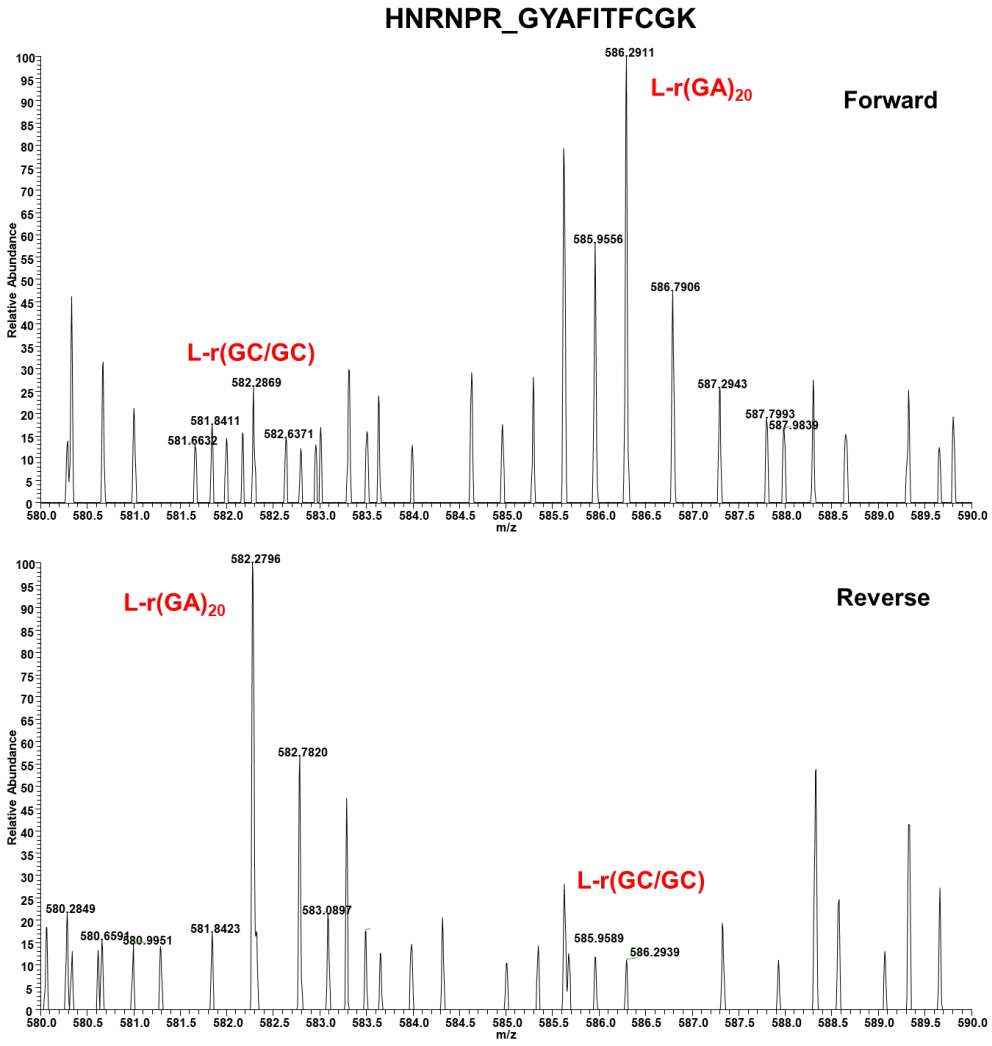

Figure S4. ESI-MS showing the $[M + 2H]^{2+}$ ions of light and heavy lysine-containing peptides, GYAFITFCGK, derived from HNRNPR in forward and reverse SILAC experiments.

Figure S5

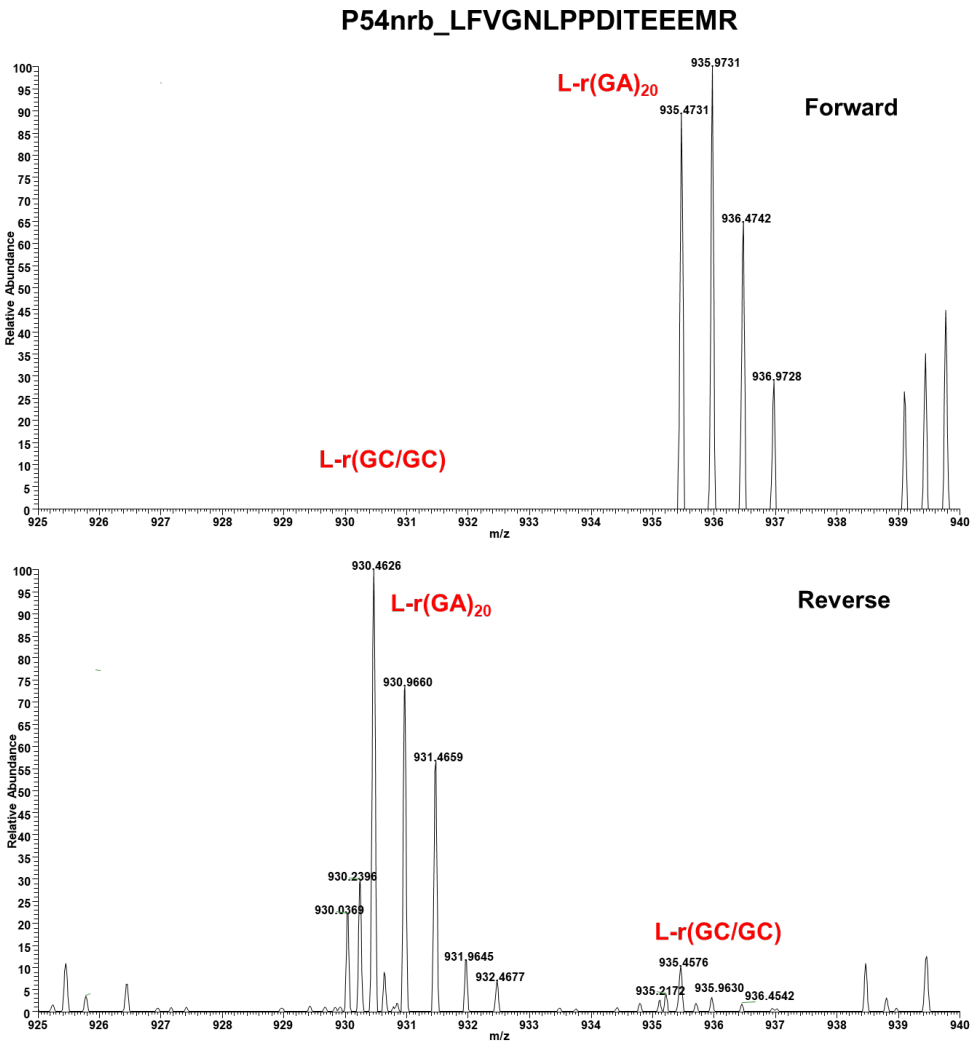

Figure S5. ESI-MS showing the $[M + 2H]^{2+}$ ions of light and heavy arginine-containing peptides, LFVGNLPPDITEEEMR, derived from P54nrb in forward and reverse SILAC experiments.

Figure S6

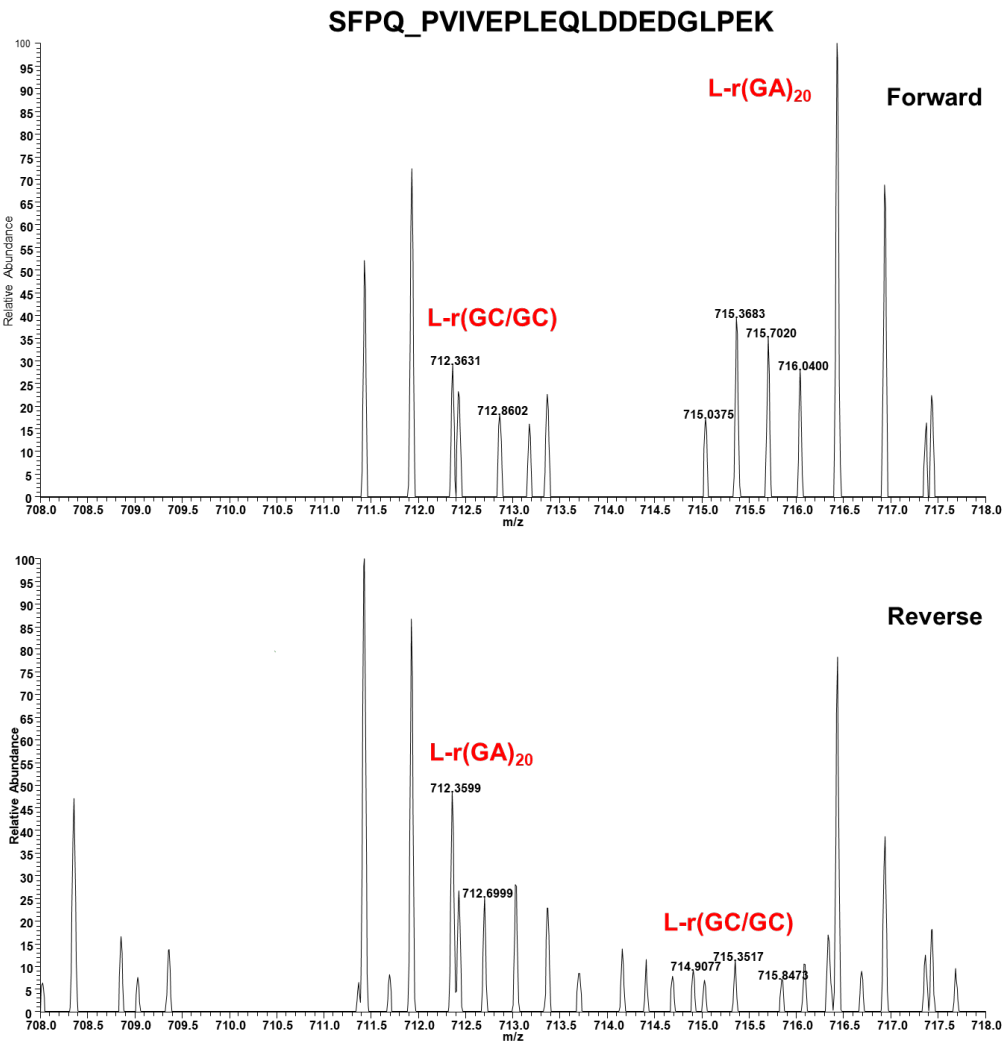

Figure S6. ESI-MS showing the $[M + 2H]^{2+}$ ions of light and heavy lysine-containing peptides, PVIVEPLEQLDDEDGLPEK, derived from SFPQ in forward and reverse SILAC experiments.

Figure S7

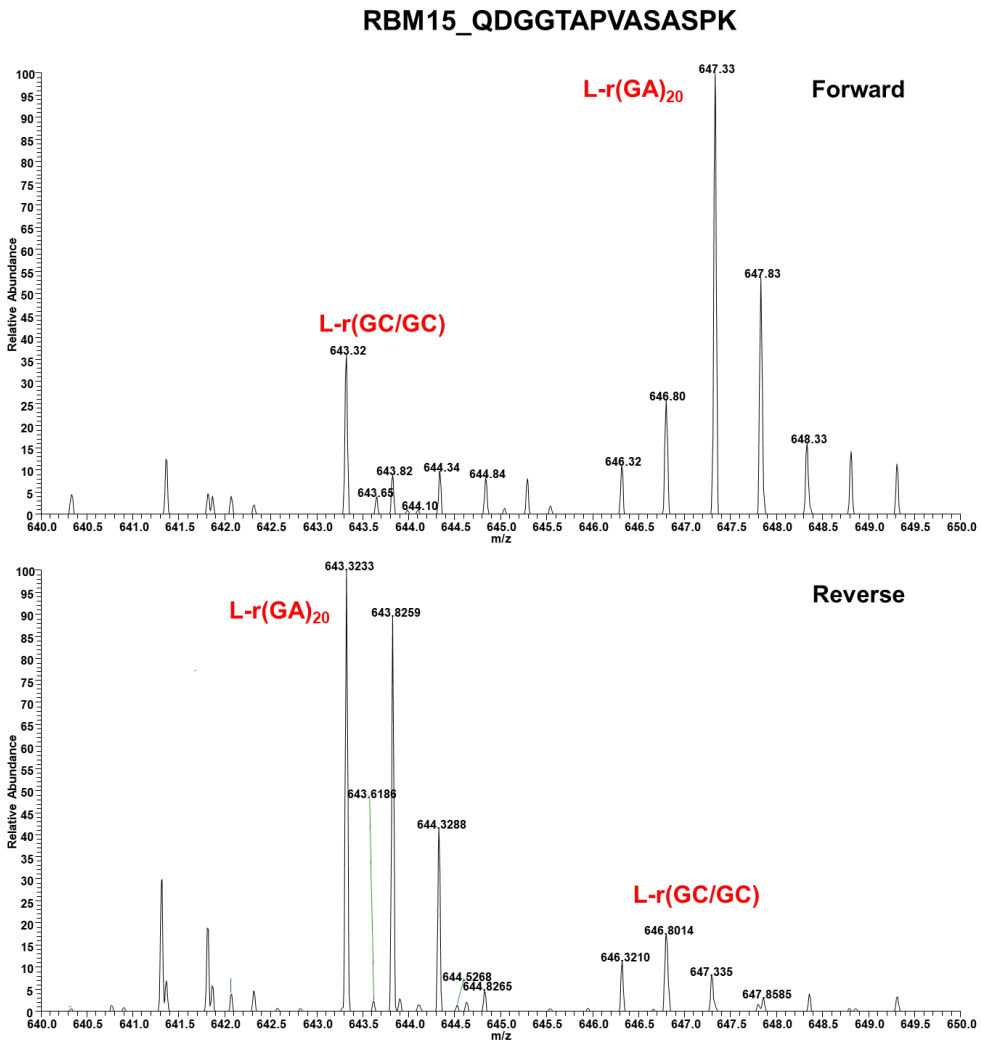

Figure S7. ESI-MS showing the $[M + 2H]^{2+}$ ions of light and heavy lysine-containing peptides, QDGGTAPVASASPK, derived from RBM15 in forward and reverse SILAC experiments.

Figure S8

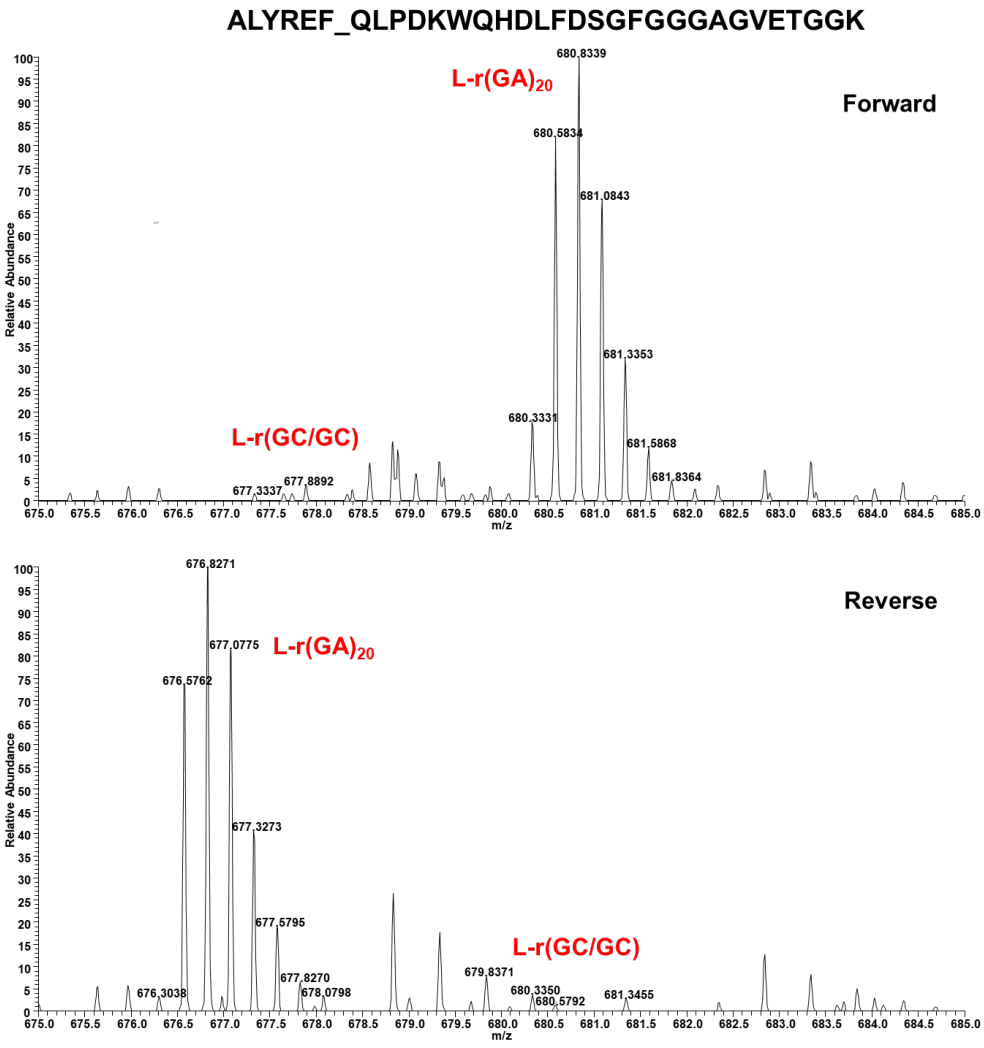

Figure S8. ESI-MS showing the $[M + 4H]^{4+}$ ions of light and heavy lysine-containing peptides, QLPDKWQHDLFDSGFGGGAGVETGGK, derived from ALYREF in forward and reverse SILAC experiments.

Figure S9

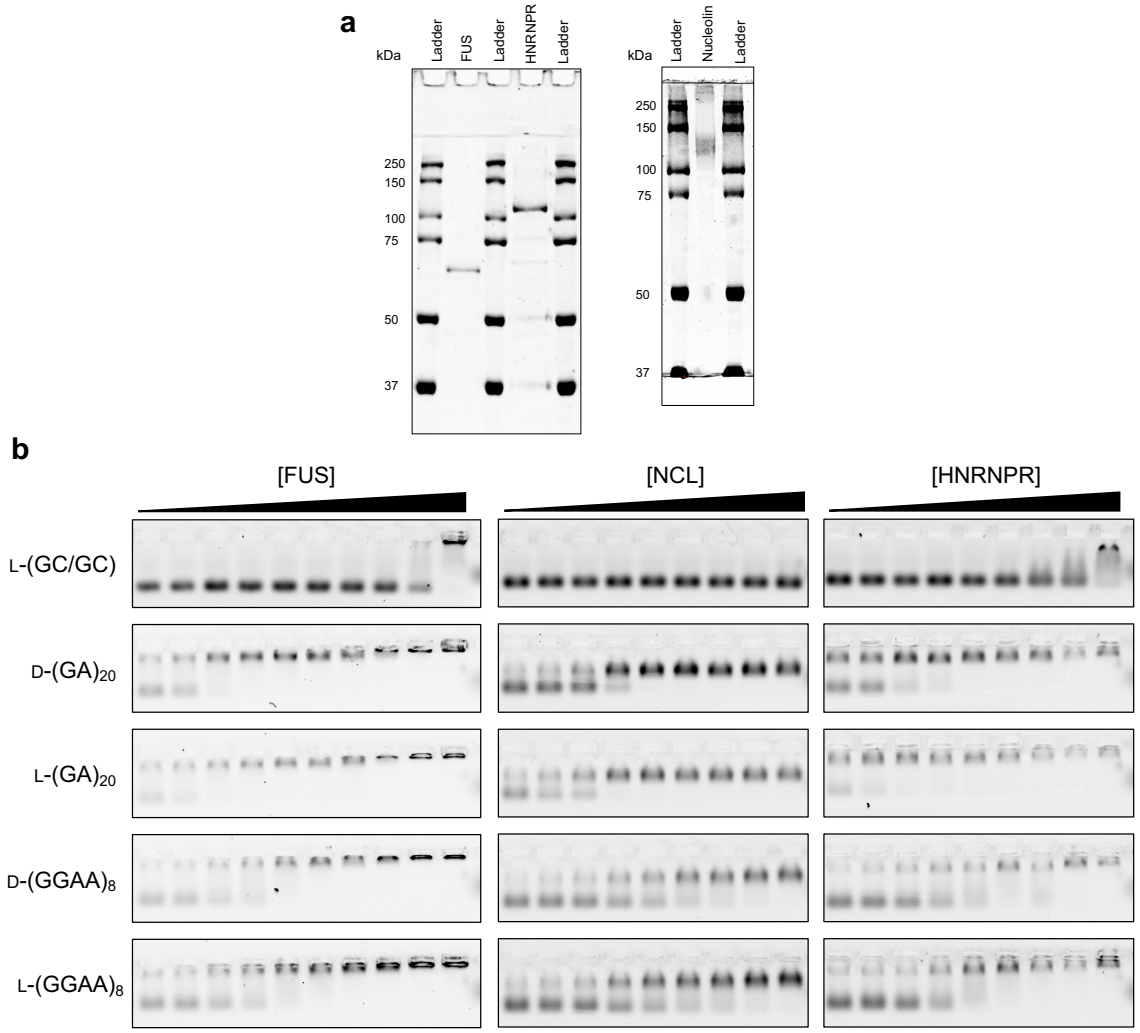

Figure S9. (a) SDS-PAGE (10%) analysis of recombinant proteins used in this work. FUS and HNRNPR have molecular masses of approximately 75 and 98 kDa, respectively. Recombinant HNRNPR migrates as an approximately 114 kDa band in SDS-PAGE under reducing conditions. Nucleolin has a predicted molecular mass of 77.3 kDa. Recombinant nucleolin migrates as an approximate 100-130 kDa band due to glycosylation. All recombinant proteins were obtained from commercial sources and used without further purification. (b) Representative EMSA gels for the different D- and L-RNA sequences binding to FUS, NCL and HNRNPR. A 2-fold serial dilution was carried out for each protein with the maximum concentrations being 500, 200, and 250 nM for FUS, NCL and HNRNPR, respectively. Uncropped gel images are presented in Figure S33.

Figure S10

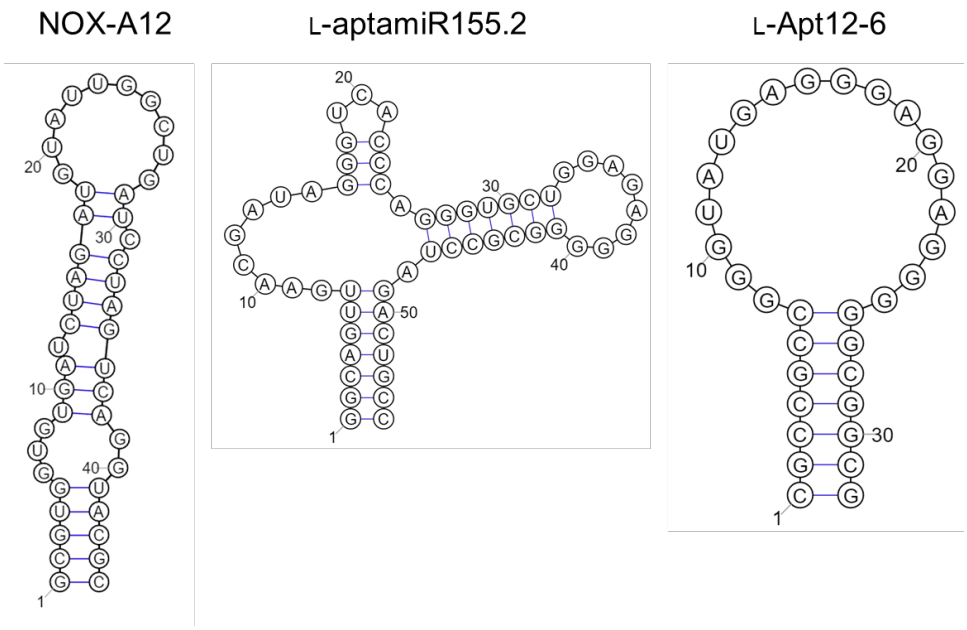

Figure S10. The sequence and secondary structure of the Spiegelmers tested in this work. Secondary structures were predicted by RNAfold[13] and plotted by RiboSketch.[14]

Figure S11

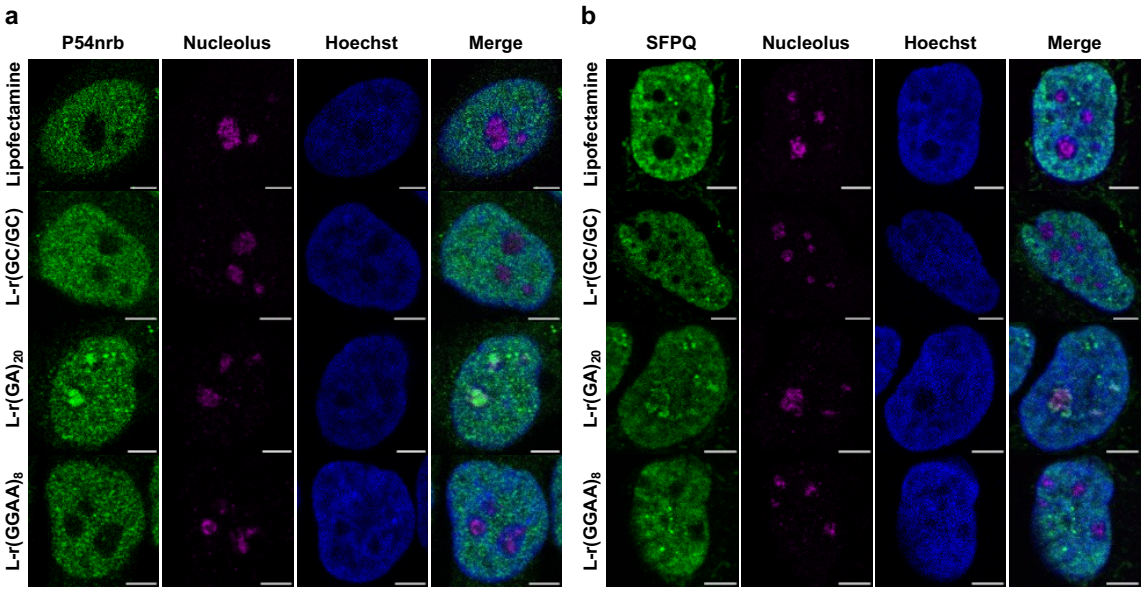

Figure S11. Cytotoxic G-rich L-RNAs elicit delocalization of paraspeckle proteins to nucleoli. p54nrb (a) and SFPQ (b) were stained with the corresponding primary antibody and Cy3-labeled secondary antibody. The nucleolus was stained with an anti-fibrillarin antibody and Alexa488-labeled secondary antibody. The nucleus was stained with Hoechst. Scale bar: 5 μm.

Figure S12

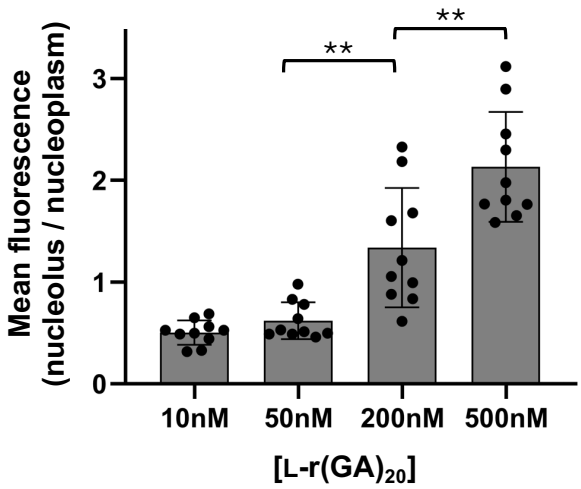

Figure S12. Nucleolar enrichment of SFPQ is dependent on L-r$(GA)_{20}$ concentration. Data is the mean fluorescent intensity within the nucleolus divided by the mean fluorescent intensity within the nucleoplasm (n = 10 cells). $^{**}P < 0.01$.

Figure S13

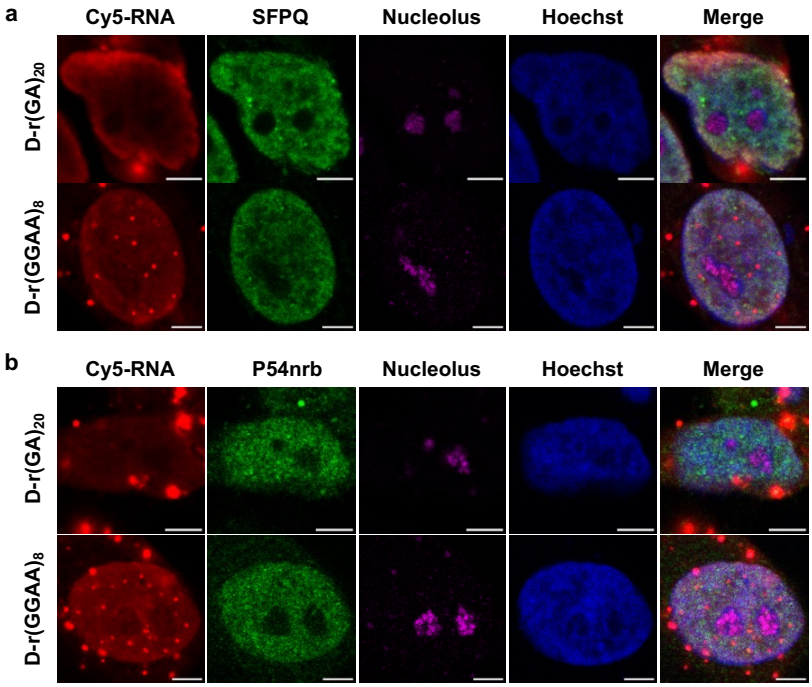

Figure S13. Representative confocal fluorescent microscopy images of HeLa cells transfected with 200 nM of the indicated Cy5-labeled D-RNAs for 2 hours. SFPQ (a) and p54nrb (b) were stained with the corresponding primary antibody and Cy3-labeled secondary antibody. The nucleolus was stained with an anti-fibrillarin antibody and Alexa488-labeled secondary antibody. The nucleus was stained with Hoechst. Scale bar: 5 μm.

Figure S14

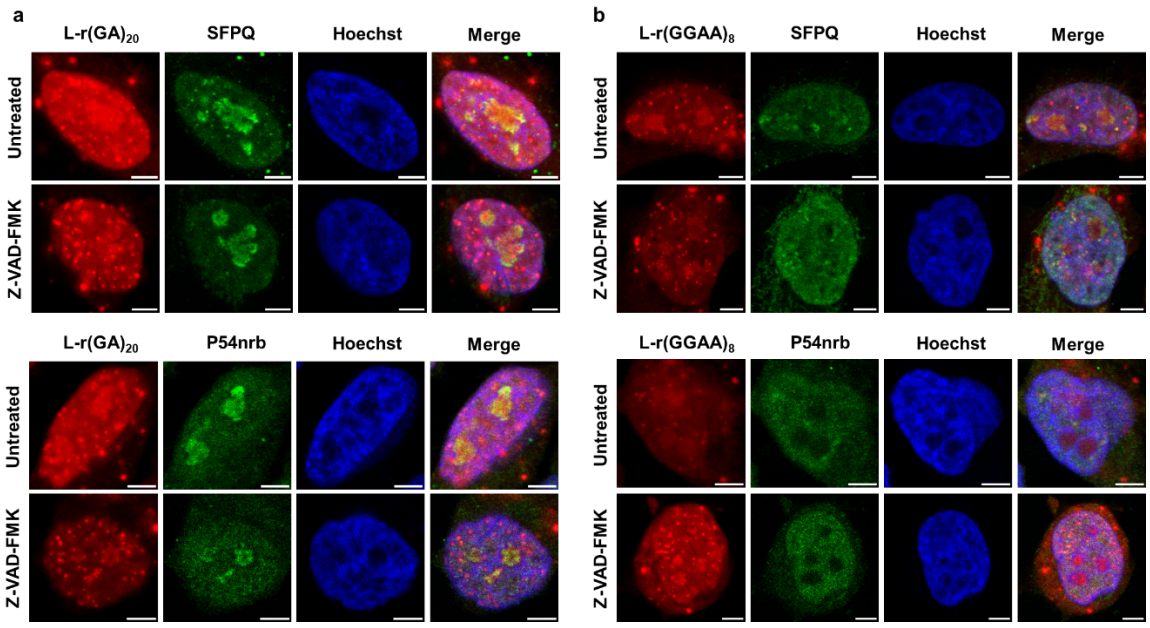

Figure S14. Delocalization of p54nrb and SFPQ to nucleoli was not the result of apoptosis. HeLa cells were treated with 20 μM Z-VAD-FMK for 1 hour followed by transfection of 200 nM of either Cy5-labeled L-r$(GA)_{20}$ (a) or Cy5-labeled L-r$(GGAA)_8$ (b) for 2 hours. SFPQ and p54nrb were stained with the corresponding primary antibody and Cy3-labeled secondary antibody. The nucleus was stained with Hoechst. Scale bar: 5 μm.

Figure S15

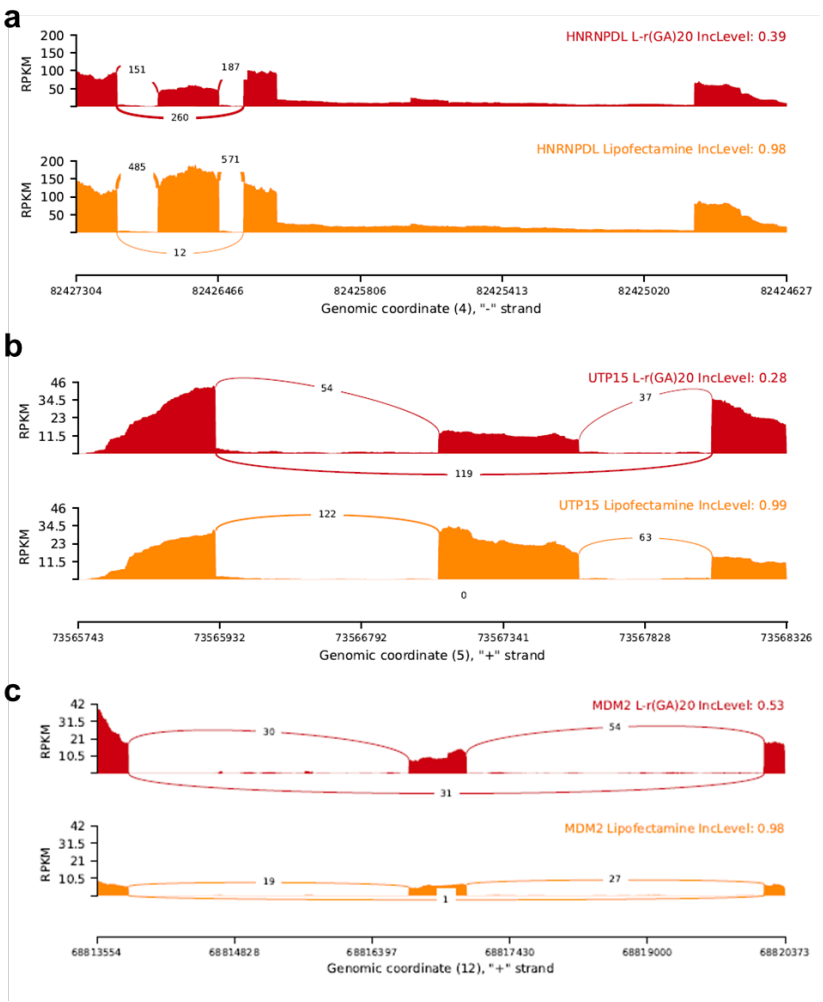

Figure S15. Sashimi plot for SE events at HNRNPDL exon 6 (a), UTP15 exon 2 (b) and MDM2 exon 4 (c) upon treatment with L-r(GA)$_{20}$. Reads obtained following L-r(GA)$_{20}$ treatment are colored in red and reads obtained from the lipofectamine only control treated cells are colored in orange. The numbers between each exon indicate the number of transcripts spanning the splice junction. Inclusion level represents the percent spliced in (PSI); the lower the PSI, the more the skipped exon isoform transcript is observed. Sashimi plots were generated by rmats2sashimiplot (version 3.0.0).[11]

Figure S16

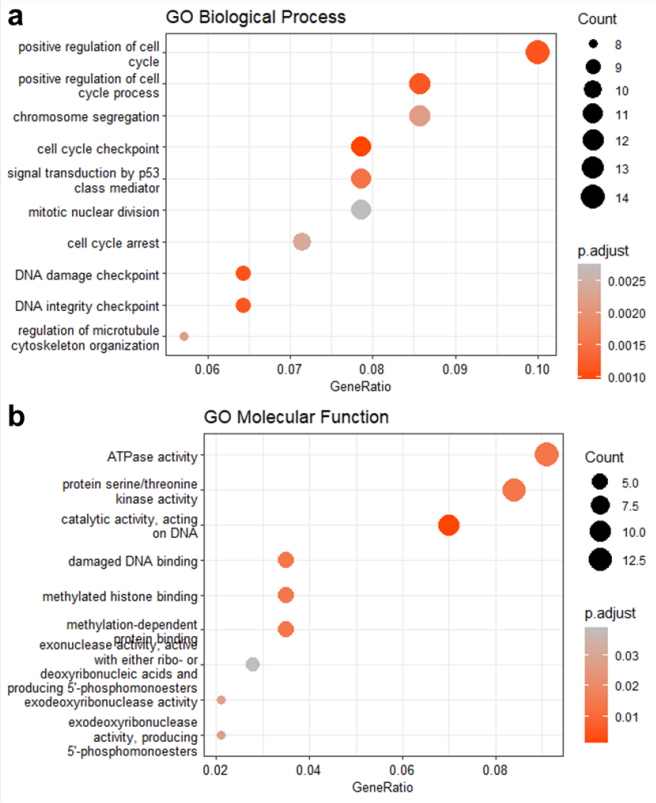

Figure S16. Gene ontology analysis of 145 common skipped exons between L-r$(GA)_{20}$ and L-r$(GGAA)_8$. The top ten GO terms in biological process (a) and molecular function (b) are listed.

Figure S17

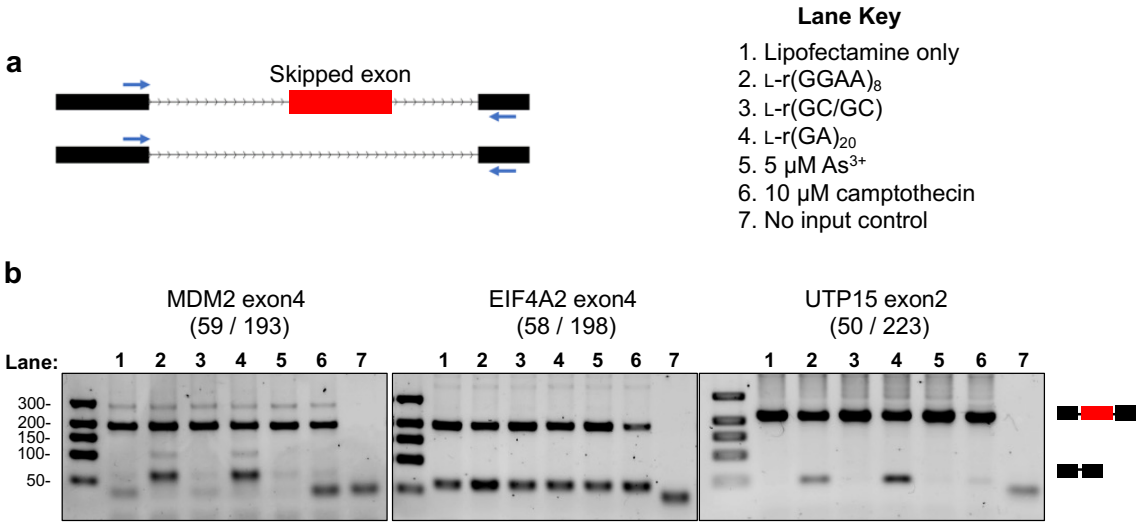

Figure S17. (a) Schematic of the primer design for assessing alternative splicing by RT-PCR. The forward primer covered the upstream exon and the reverse primer covered the downstream exon. (b) Representative agarose gels of exon skipping at the indicated site upon various treatments (see Lane Key). Cells were treated as indicated for 12 hours and the total RNA was isolated and analyzed by RT-PCR. The expected length of the skipped exon and full-length PCR products are indicated in parentheses above the gel (skipped bp / full-length bp).

Figure S18

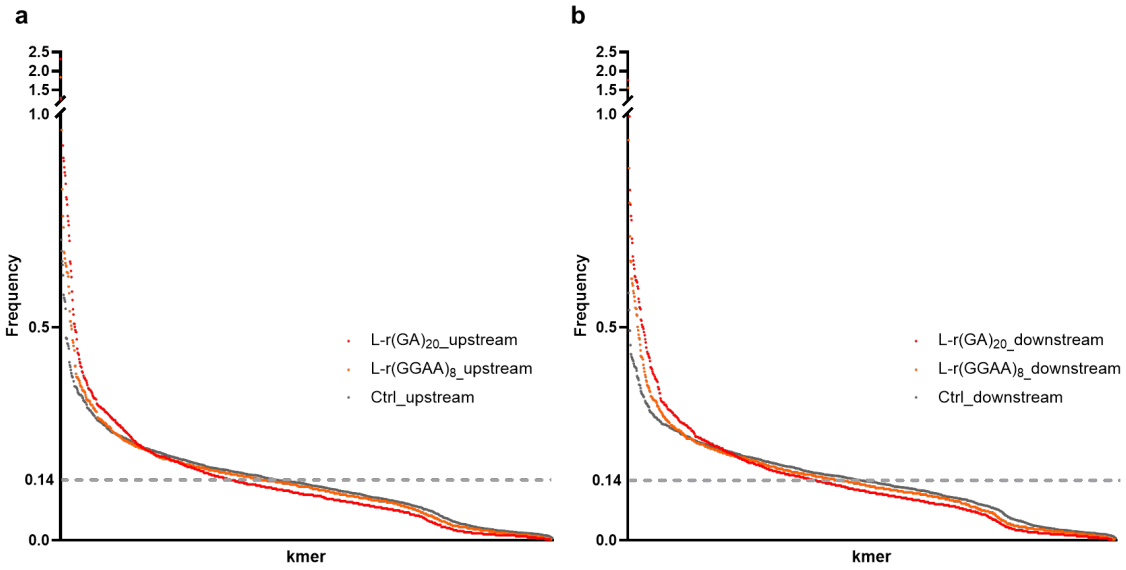

Figure S18. *K*-mer (*k* = 5) distribution for the 150 bp upstream (a) and 150 bp downstream (b) of skipped exons affected by the indicated L-RNA. *K*-mers from each treatment were ranked based on frequency. The horizontal dotted line indicates the average frequency (0.14) in a 150 bp region for each possible *k*-mer. For the control (Ctrl), the same analysis was carried out on all known exon skipping events from the exon skipping database ExonSkipDB.

Figure S19

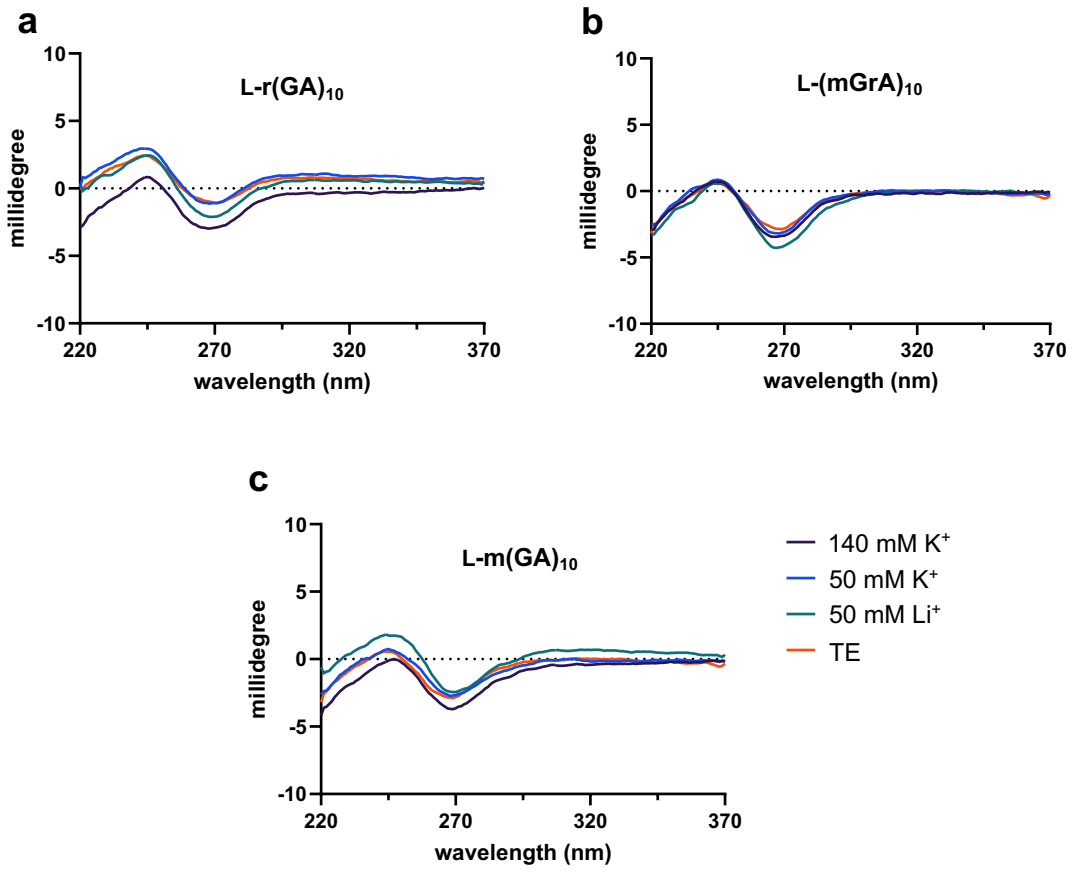

Figure S19. Circular dichroism (CD) spectra of L-r(GA)$_{10}$ (a), L-(mGrA)$_{10}$ (b), and L-m(GA)$_{10}$ (c) in buffer containing 25 mM Tris (pH 7.4), 1 mM EDTA, and either 50 mM KCl (50 mM K^+), 140 mM KCl (140 mM K^+), or 50 mM LiCl (50 mM Li^+). TE indicates a buffer consisting of only 25 mM Tris (pH 7.4) and 1 mM EDTA**.** The CD spectra for these sequences lack features of a G-quadruplex and are mostly unaffected by K^+ and Li^+, which promote and inhibit G-quadruplex formation, respectively.

Figure S20

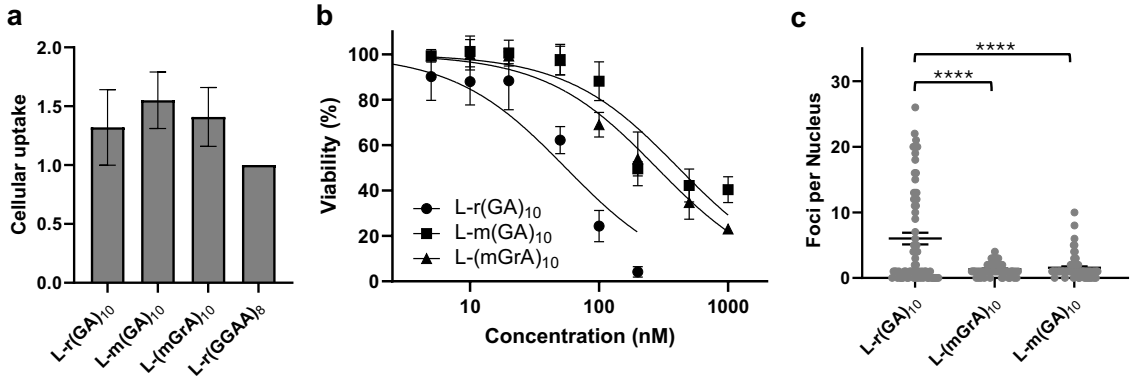

Figure S20. (a) Cellular uptake of the indicated L-RNAs. HeLa cells were transfected with 200 nM L-RNA and uptake was measured 2 hours later by flow cytometry. Data are mean ± S.D. (n = 3 biological replicates). (b) Dose response curves for the 2´-OMe modified L-r(GA)$_{10}$ based on the CCK-8 assay. HeLa cells were transfected with indicated concentration of L-RNA for 2 hours, washed, and the incubated in fresh DMEM for 46 hours as described in the methods section. Data are mean ± S.D. (n = 3 biological replicates). (c) Quantification of nuclear foci formation following the treatment of HeLa cells with 200 nM of the indicated L-RNA. (n = 60 cells). ****P < 0.0001

Figure S21

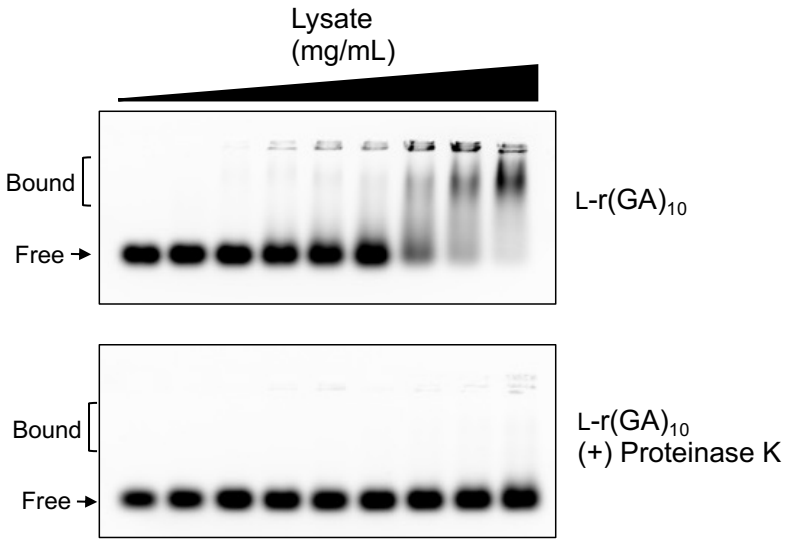

Figure S21. Proteinase K treatment of HeLa cell lysates abolishes the shift of L-r(GA)$_{10}$ in the EMSA. Representative EMSA data for L-r(GA)$_{10}$ binding to HeLa cell nuclear lysate (0 – 10 mg/mL) before (top) and after (bottom) treatment with proteinase K (2 mg/mL). Uncropped gel images are presented in Figure S34.

Figure S22

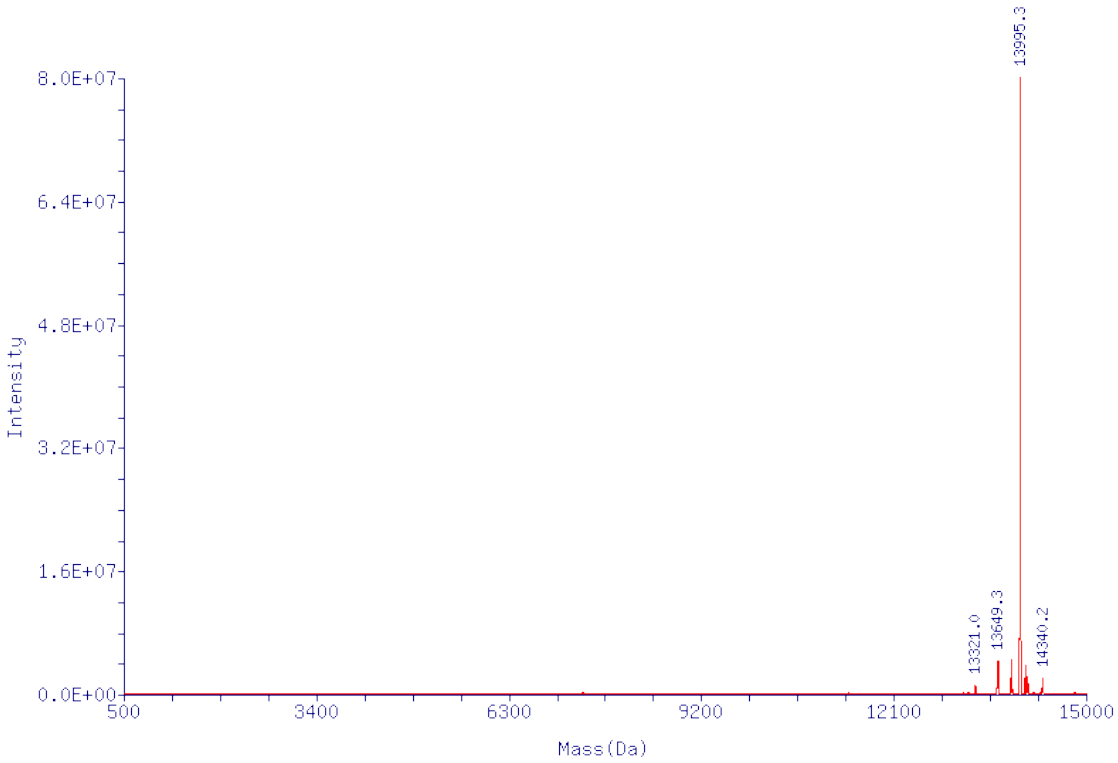

Figure S22. ESI-MS of L-r(GA)$_{20}$_Bio prepared by solid-phase synthesis. Mass calculated: 13995.9 Da; Mass found: 13995.3 Da.

Figure S23

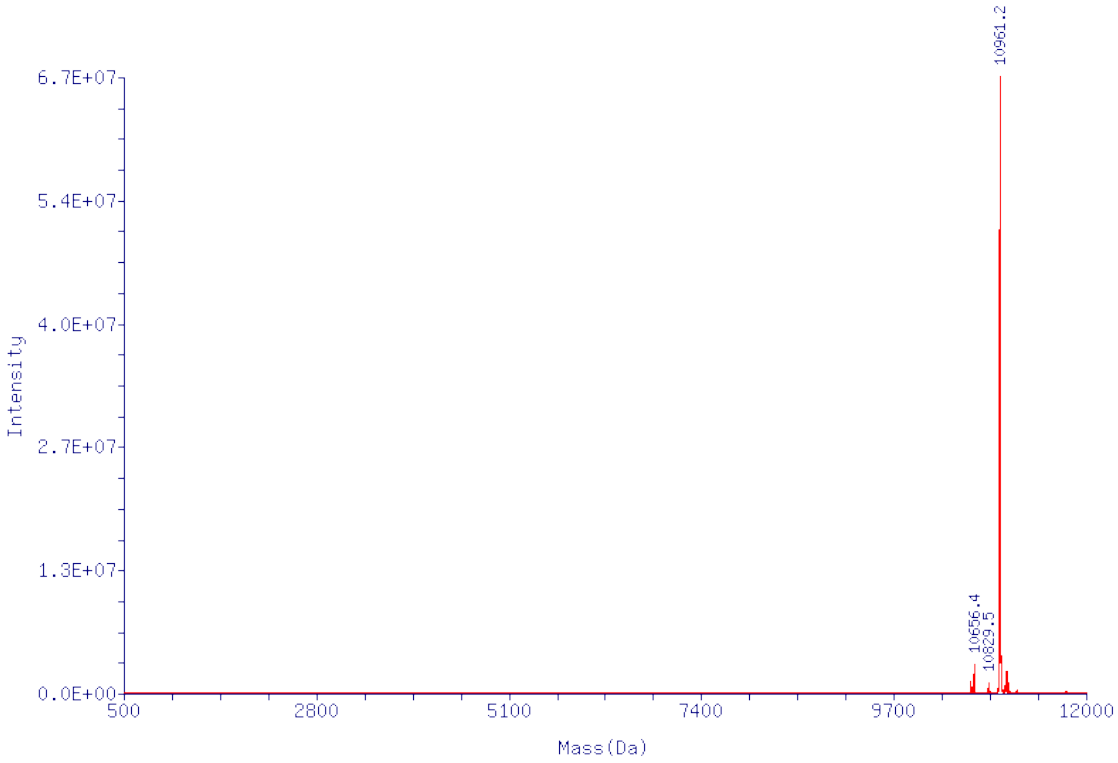

Figure S23. ESI-MS of L-r(GC/GC)_Bio prepared by solid-phase synthesis. Mass calculated: 10961.9 Da; Mass found: 10961.2 Da.

Figure S24

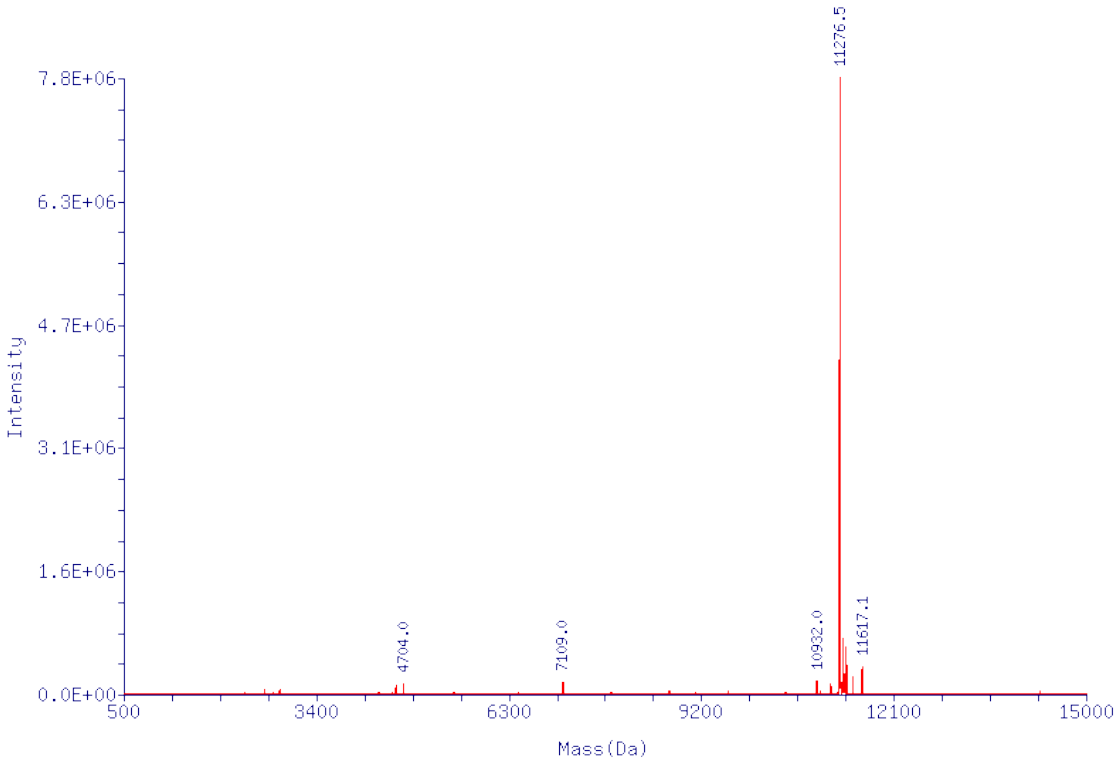

Figure S24. ESI-MS of L-r(GGAA)$_8$_Bio prepared by solid-phase synthesis. Mass calculated: 11277.4 Da; Mass found: 11276.5 Da.

Figure S25

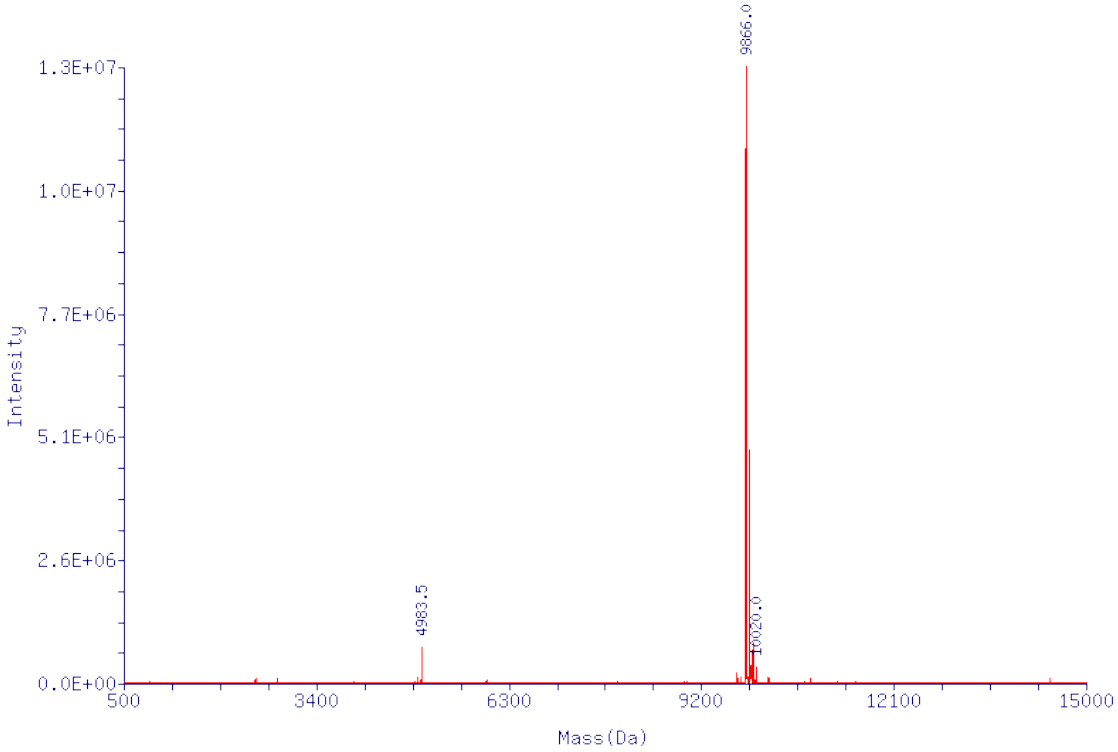

Figure S25. ESI-MS of L-r(G_3A_4)$_4$_Bio prepared by solid-phase synthesis. Mass calculated: 9866.5 Da; Mass found: 9666.0 Da.

Figure S26

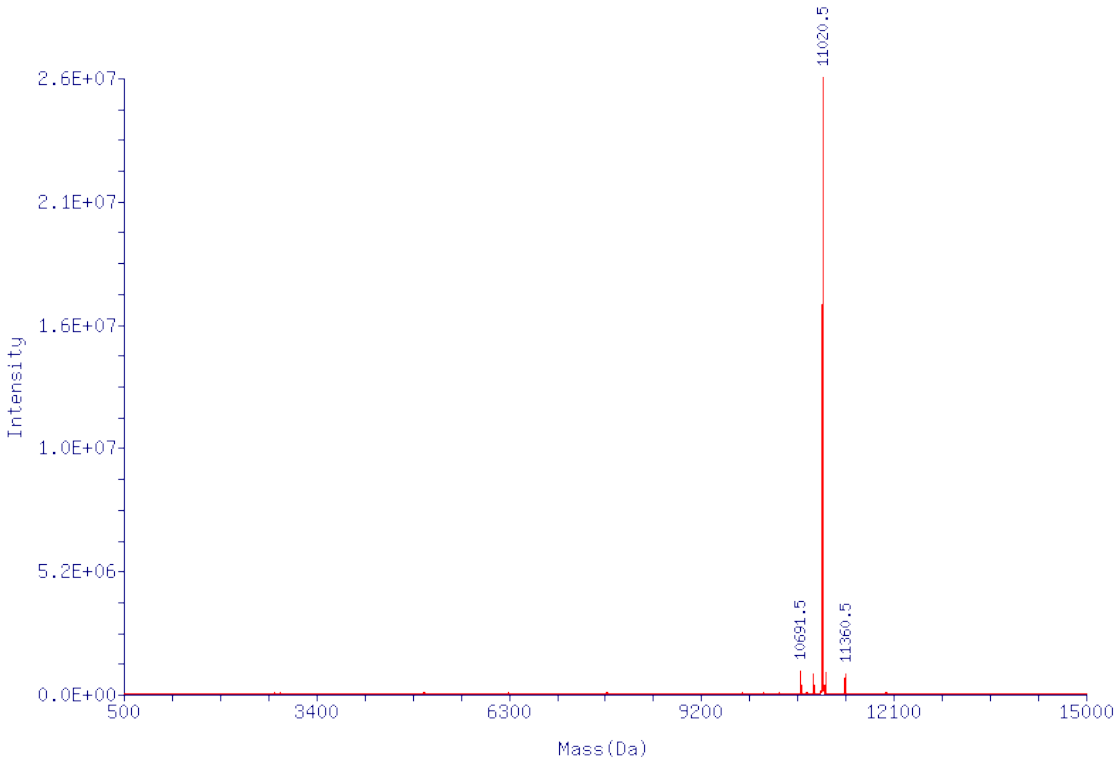

Figure S26. ESI-MS of L-rA$_{32}$_Bio prepared by solid-phase synthesis. Mass calculated: 11021.4 Da; Mass found: 11020.5 Da.

Figure S27

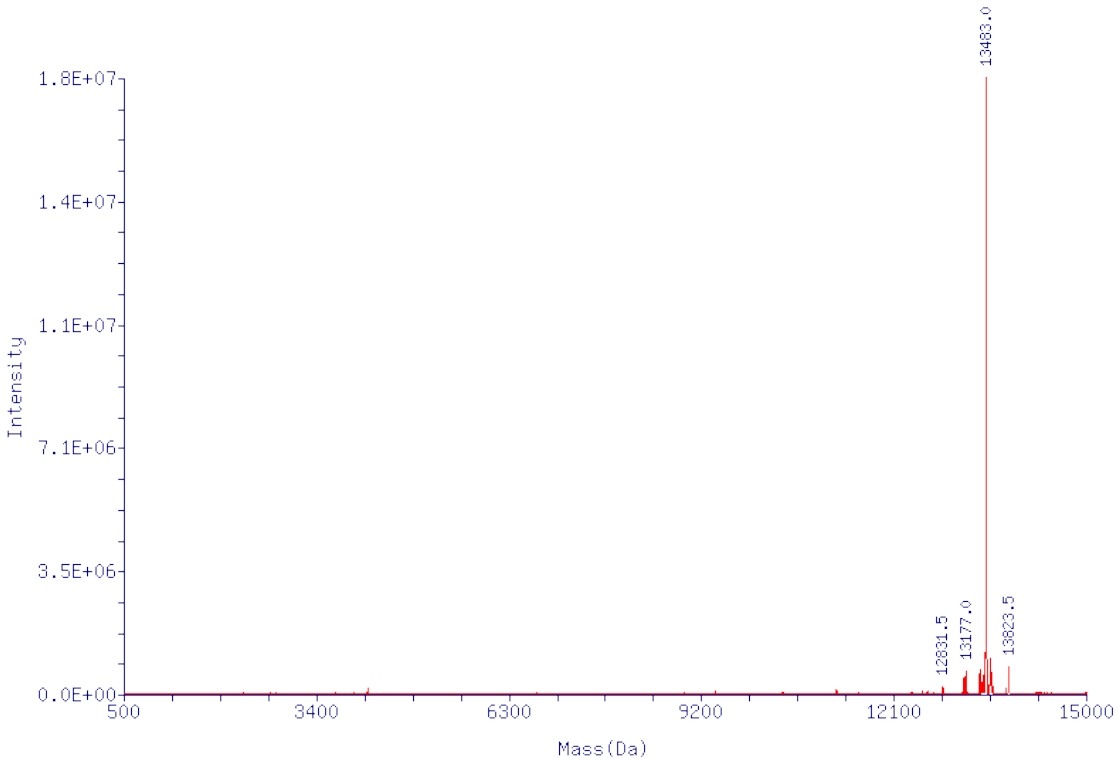

Figure S27. ESI-MS of L-r(GU)$_{20}$_Bio prepared by solid-phase synthesis. Mass calculated: 13484.2 Da; Mass found: 13483.0 Da.

Figure S28

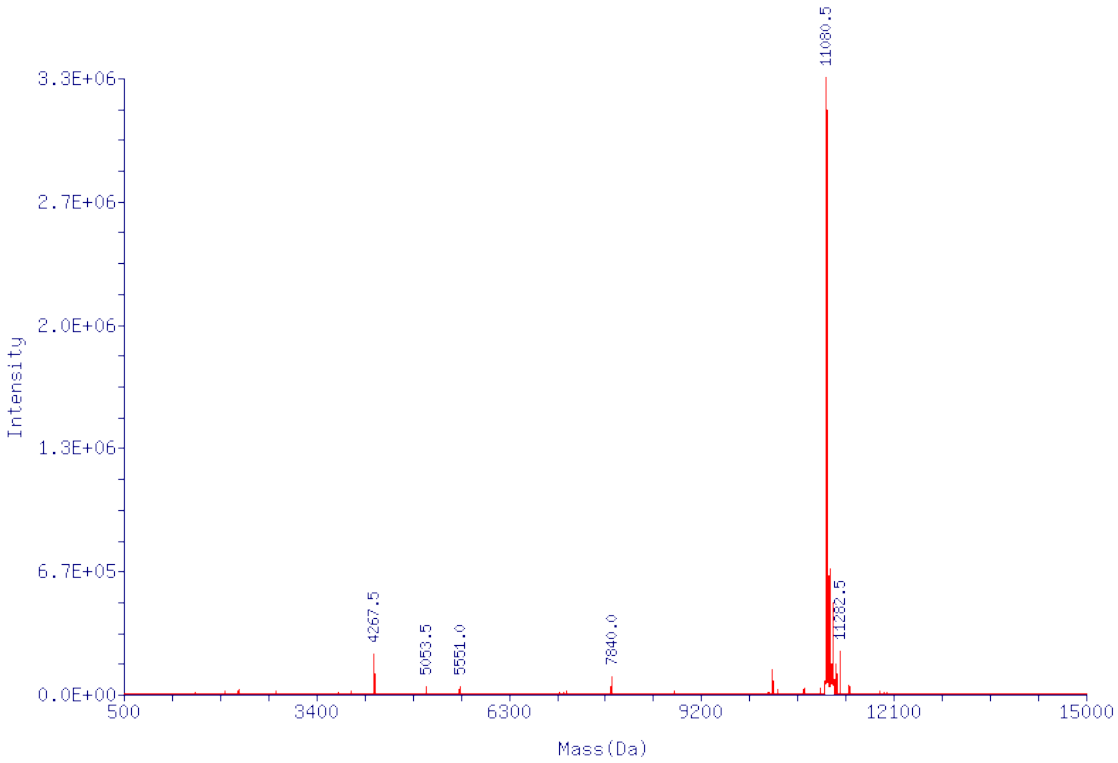

Figure S28. ESI-MS of L-Apt12-6_Bio prepared by solid-phase synthesis. Mass calculated: 11081.1 Da; Mass found: 11080.5 Da.

Figure S29

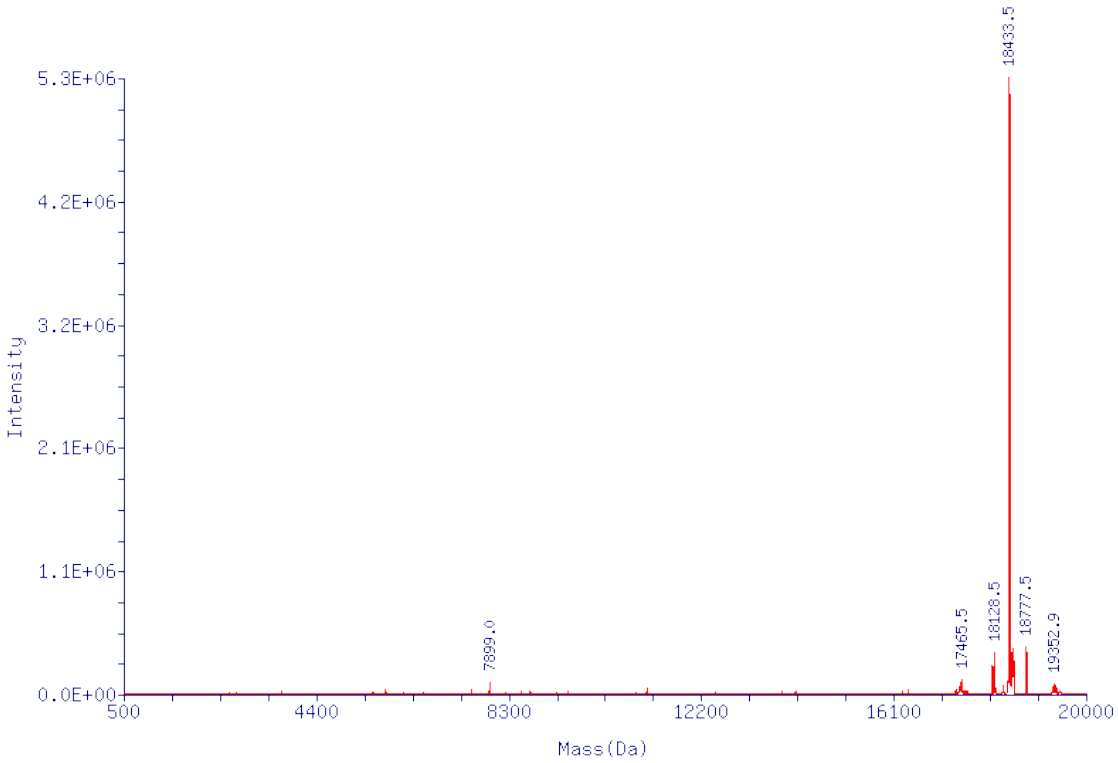

Figure S29. ESI-MS of L- aptamiR155.2_Bio prepared by solid-phase synthesis. Mass calculated: 18434.5 Da; Mass found: 18433.5 Da.

Figure S30

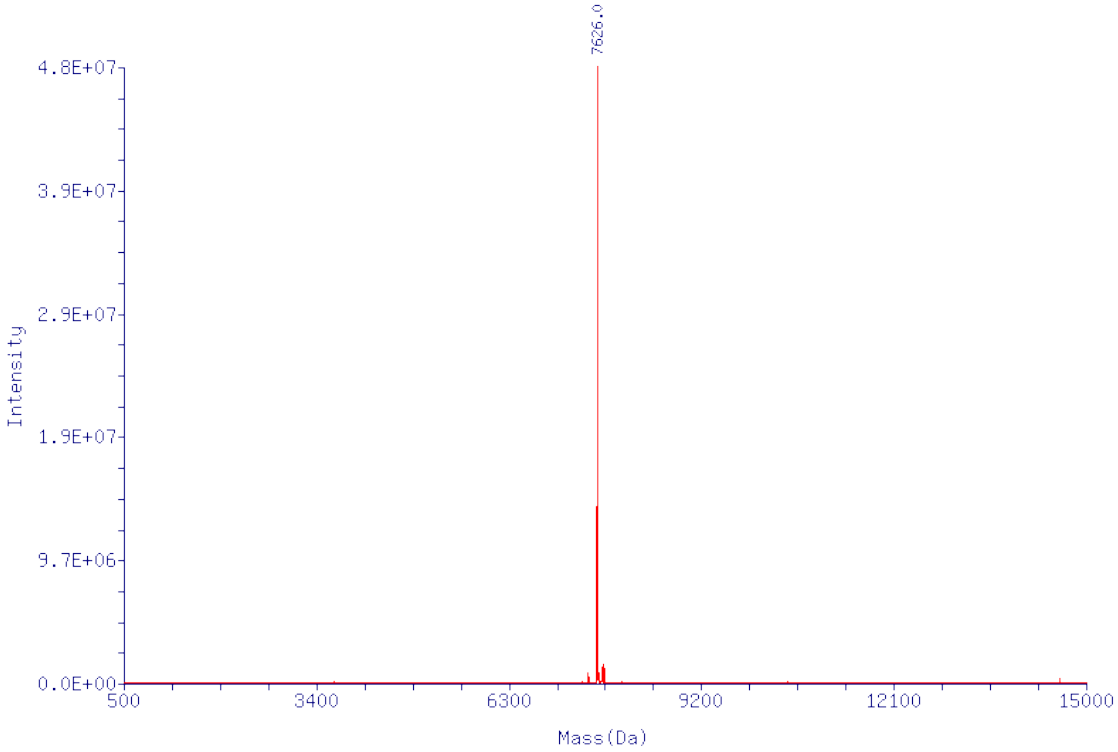

Figure S30. ESI-MS of L-$(mGrA)_{10}$ prepared by solid-phase synthesis. Mass calculated: 7626.9 Da; Mass found: 7626.0 Da.

Figure S31

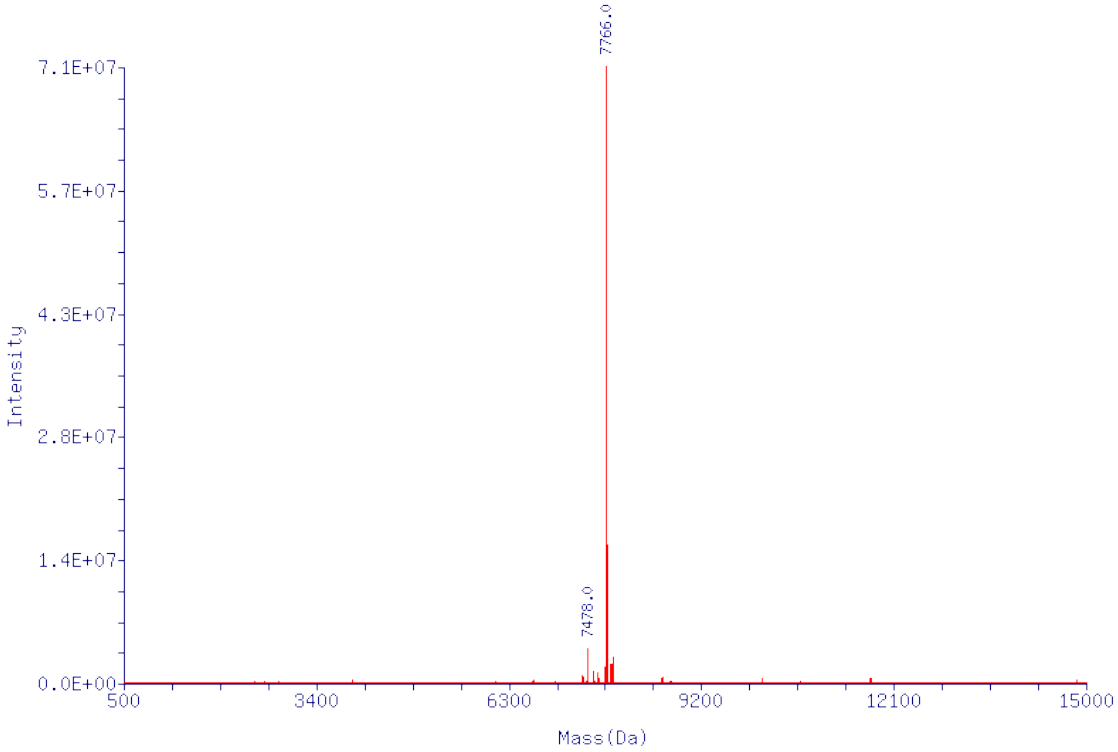

Figure S31. ESI-MS of L-m$(GA)_{10}$ prepared by solid-phase synthesis. Mass calculated: 7767.1 Da; Mass found: 7766.0 Da.

Figure S32

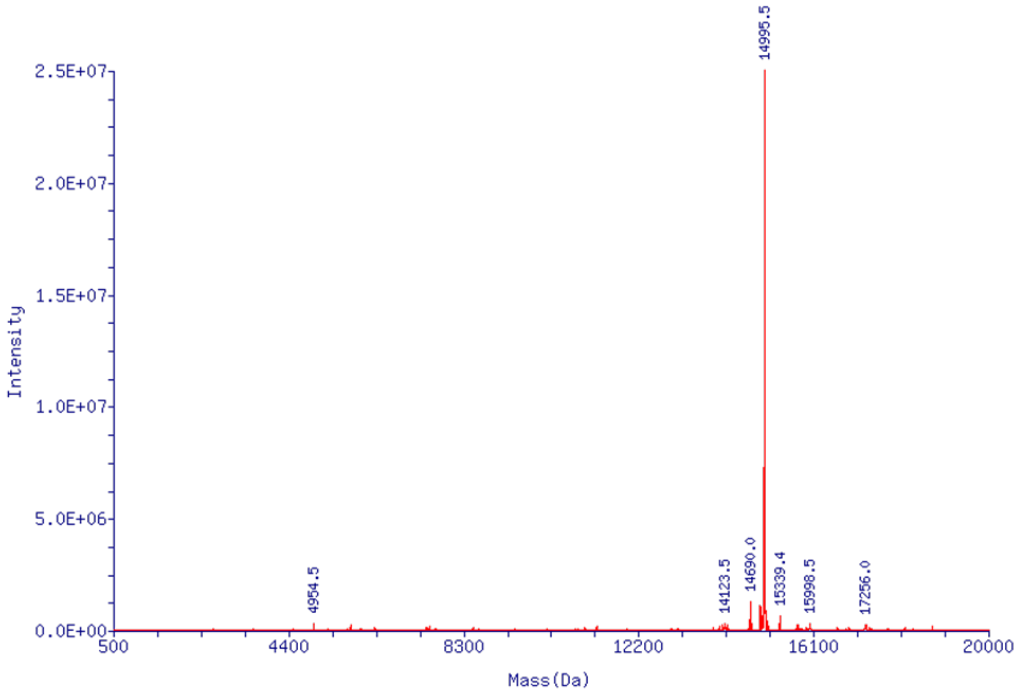

Figure S32. ESI-MS of NOX-A12_Bio prepared by solid-phase synthesis. Mass calculated: 14995.9 Da; Mass found: 14995.5 Da.

Figure S33

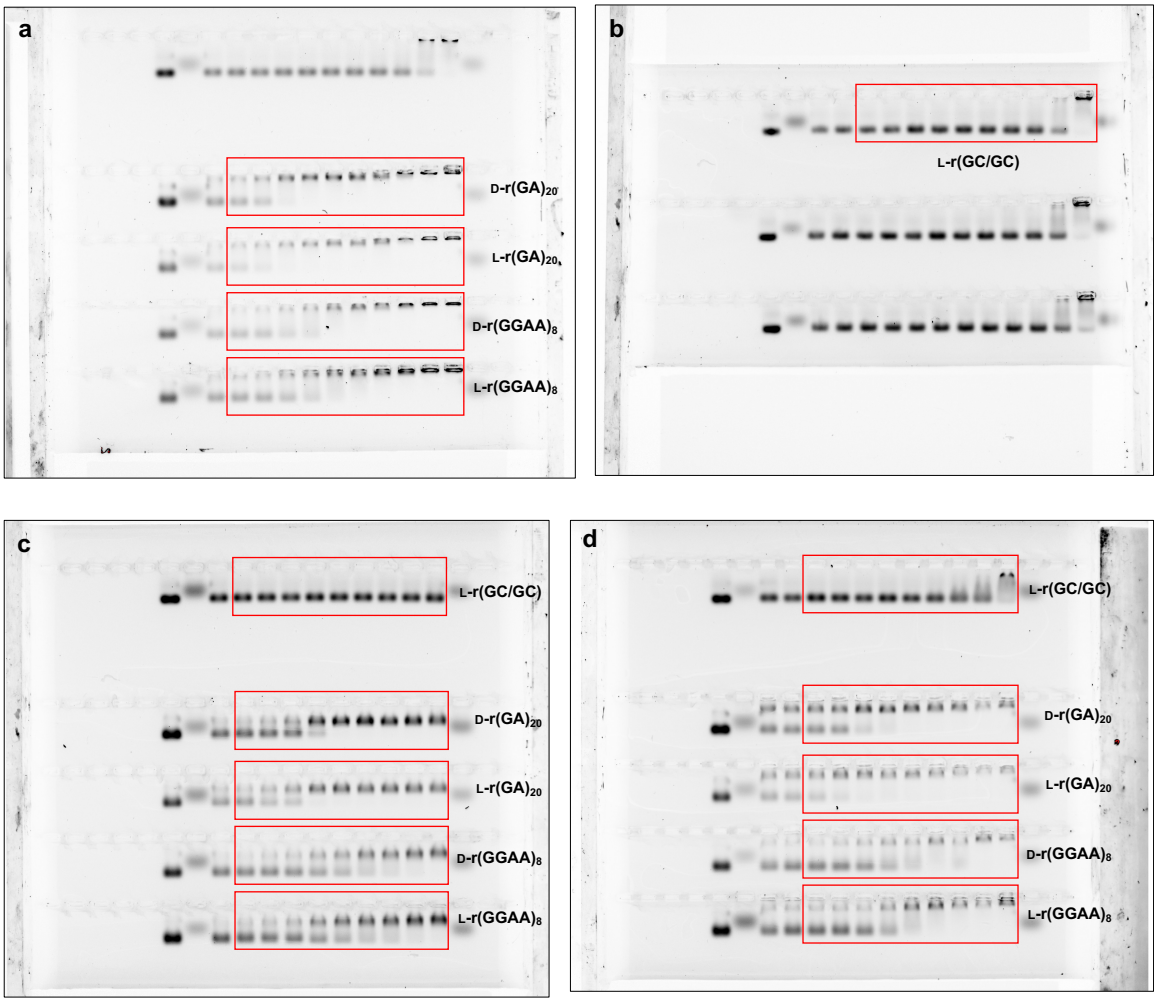

Figure S33. Uncropped gel image for Figure S9b. (a,b) Frames indicate the cropped regions used to assemble the column of images for FUS in Figure S9b. (c) Frames indicate the cropped regions used to assemble the column of images for NCL in Figure S9b. (d) Frames indicate the cropped regions used to assemble the column of images for HNRNPR in Figure S9b. The first lane in each set (unboxed) is the no-protein control.

Figure S34

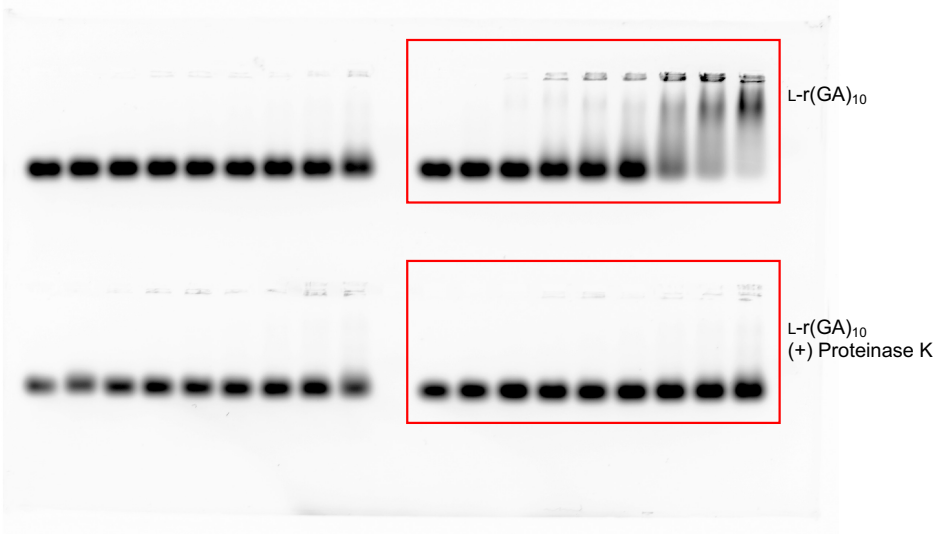

Figure S34. Uncropped gel image for Figure S21. Frames indicate the cropped regions shown in Figure S21.

S3. Supplementary Tables.

Table S1. Names and sequences of oligonucleotides used in this work. L-RNA (blue) and D-RNA (black) are indicated by color. 2´-OMe modified nucleotides are underlined. /Biotin/ = long-chain Biotin; /Cy3/ = sulfo-cyanine 3; /Cy5/ = sulfo-cyanine 5.

Sequence Name	Sequence Identity 5′→3′
L-r$(GA)_{20}$_Bio	GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA/Biotin/
*L-r$(GA)_{20}$_Cy5	GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA/Cy5/
*D-r$(GA)_{20}$_Cy3	GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA/Cy3/
L-r(GC/GC)_Bio	GCGCGCGCGCGCGCGAGAGCGCGCGCGCGCGC/Biotin/
*L-r(GC/GC)_Cy5	GCGCGCGCGCGCGCGAGAGCGCGCGCGCGCGC/Cy5/
L-r$(GGAA)_8$_Bio	GGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAA/Biotin/
*L-r$(GGAA)_8$_Cy5	GGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAA/Cy5/
*D-r$(GGAA)_8$_Cy5	GGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAA/Cy5/
L-r$(G_3A_4)_4$_Bio	GGGAAAAGGGAAAAGGGAAAAGGGAAAA/Biotin/
L-rA_{32}_Bio	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA/Biotin/
L-r$(GU)_{20}$_Bio	GUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGU/Biotin/
L-Apt12-6_Bio	CGCCGCCGGGUAUGAGGGAGGAGGGGGCGGCG/Biotin/
L-aptamiR155.2_Bio	GGCAGUUGAACGAUAGGGUCACCCAGGGUGCUGGAGAGGGGGCGCCUAGACUGCC/Biotin/
NOX-A12_Bio	GCGUGGUGUGAUCUAGAUGUAUUGGCUGAUCCUAGUCAGGUACGC/Biotin/
L-r$(GA)_{10}$	GAGAGAGAGAGAGAGAGAGA-/Cy5/
L-$(mGrA)_{10}$	GAGAGAGAGAGAGAGAGAGA-/Cy5/
L-m$(GA)_{10}$	GAGAGAGAGAGAGAGAGAGA-/Cy5/

*These oligonucleotides were prepared and characterized previously.[1]

Table S2. Antibodies used in this work.

Antibody	Host	Dilution (WB)	Dilution (IF)	Source
anti-p54nrb	Rabbit	1:500	1:50	Proteintech; cat.# 11058-1-AP
anti-SFPQ	Rabbit	1:2000	1:50	Proteintech; cat.# 15585-1-AP
anti-HNRNPR	Rabbit	1:500	1:50	Invitrogen; cat.# PA5-109826
anti-RBM15	Mouse	1:500	1:200	Proteintech; cat.# 66059-1-IG
anti-FUS	Mouse	1:5000	1:20	Proteintech; cat.# 60160-1-IG
anti-NCL	Rabbit	1:5000	1:500	Proteintech; cat.# 10556-1-AP
anti-FBL	Mouse		1:100	antibodies.com; cat.# A85370
anti-ALYREF	Rabbit	1:500		Proteintech; cat.# 16690-1-AP
anti-HIST3	Rabbit	1:1000		Invitrogen; cat.# PA5-16183
anti-G3BP1	Mouse	1:5000		Proteintech; cat.# 66486-1-IG
anti-RBMX	Rabbit	1:1000		Invitrogen; cat.# PA5-49468
Cy3 anti-mouse IgG	Goat		1:500	Invitrogen; cat.# A10521
Cy3 anti-rabbit IgG	Goat		1:500	Invitrogen; cat.# A10520
Alexa Fluor™488 anti-mouse IgG	Goat		1:2000	Invitrogen; cat.# A-11001
Alexa Fluor™ Plus 647 anti-rabbit IgG	Goat	1:10000		Invitrogen; cat.# A32733
Alexa Fluor™ Plus 647 anti-mouse IgG	Goat	1:10000		Invitrogen; cat.# A32728

Table S3. Primers used to validate exon skipping by RT-PCR.

Name	Sequence Identity 5′→3′
MDM2E4_FWD	CAAAAAGACACTTATACTATGAAAGAG
MDM2E4_REV	CCAAGTTCCTGTAGATCATGGT
UTP15E2_FWD	CCTTTGGGGCTCAGTGGAG
UTP15E2_REV	CTGCACCAAATTCCTTAATCTGAAC
EIF4AE4_FWD	GAGAGCTATTATTCCCTGTATTAAAG
EIF4AE4_REV	TAGTCTCCAAGTGCCAGAATT

S4. References

1. Yu, C.-H.; Sczepanski, J. T., The influence of chirality on the behavior of oligonucleotides inside cells: revealing the potent cytotoxicity of G-rich L-RNA. *Chem. Sci.* **2023,** *14* (5), 1145-1154.
2. Harsha, H. C.; Molina, H.; Pandey, A., Quantitative proteomics using stable isotope labeling with amino acids in cell culture. *Nature Protocols* **2008,** *3* (3), 505-516.
3. Shevchenko, A.; Tomas, H.; Havlis, J.; Olsen, J. V.; Mann, M., In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols* **2006,** *1* (6), 2856-2860.
4. Tyanova, S.; Temu, T.; Cox, J., The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols* **2016,** *11* (12), 2301-2319.
5. Ritchie, M. E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C. W.; Shi, W.; Smyth, G. K., limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **2015,** *43* (7), e47.
6. Wu, T.; Hu, E.; Xu, S.; Chen, M.; Guo, P.; Dai, Z.; Feng, T.; Zhou, L.; Tang, W.; Zhan, L.; Fu, X.; Liu, S.; Bo, X.; Yu, G., clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation* **2021,** *2* (3), 100141.
7. Mistry, J.; Chuguransky, S.; Williams, L.; Qureshi, M.; Salazar, Gustavo A.; Sonnhammer, E. L. L.; Tosatto, S. C. E.; Paladin, L.; Raj, S.; Richardson, L. J.; Finn, R. D.; Bateman, A., Pfam: The protein families database in 2021. *Nucleic Acids Res.* **2021,** *49* (D1), D412-D419.
8. Dobin, A.; Davis, C. A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T. R., STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **2013,** *29* (1), 15-21.
9. Shen, S.; Park, J. W.; Lu, Z. X.; Lin, L.; Henry, M. D.; Wu, Y. N.; Zhou, Q.; Xing, Y., rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc. Natl. Acad. Sci. USA* **2014,** *111* (51), E5593-5601.
10. Park, J. W.; Tokheim, C.; Shen, S.; Xing, Y., Identifying differential alternative splicing events from RNA sequencing data using RNASeq-MATS. *Methods Mol. Biol.* **2013,** *1038*, 171-179.

11. Wang, Y.; Xie, Z.; Kutschera, E.; Adams, J. I.; Kadash-Edmondson, K. E.; Xing, Y., rMATS-turbo: an efficient and flexible computational tool for alternative splicing analysis of large-scale RNA-seq data. *Nature Protocols* **2024,** *19* (4), 1083-1104.
12. Kim, P.; Yang, M.; Yiya, K.; Zhao, W.; Zhou, X., ExonSkipDB: functional annotation of exon skipping event in human. *Nucleic Acids Res.* **2020,** *48* (D1), D896-D907.
13. Gruber, A. R.; Lorenz, R.; Bernhart, S. H.; Neubock, R.; Hofacker, I. L., The Vienna RNA websuite. *Nucleic Acids Res.* **2008,** *36* (Web Server issue), W70-74.
14. Lu, J. S.; Bindewald, E.; Kasprzak, W. K.; Shapiro, B. A., RiboSketch: versatile visualization of multi-stranded RNA and DNA secondary structure. *Bioinformatics* **2018,** *34* (24), 4297-4299.