# Supporting Information

# Mutagenesis-Based Evolution Enables Rapid Retargeting of L-Aptamers for Structured RNA Recognition

Xuan Han[1], Tyler Guilbault[1], and Jonathan T. Sczepanski[1,2]*

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

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

## MATERIAL AND METHODS

### General

Commercial DNA oligonucleotides were purchased from either Integrated DNA Technologies (Coralville, IA) or GENEWIZ (South Plainfield, NJ). All RNA oligonucleotides were prepared in house by solid-phase oligonucleotide synthesis using an Expedite 8909 DNA Synthesizer. D-DNA phosphoramidites, solid supports, and other oligonucleotide synthesis reagents were purchased from Glen Research (Sterling, VA). L-ribonucleoside phosphoramidites and solid supports were purchased from ChemGenes (Wilmington, MA). Superscript II reverse transcriptase, TURBO DNase, and SYBR Gold Nucleic Acid Gel Stain were purchased from ThermoFisher (Waltham, MA). Recombinant T7-RNA polymerase was expressed in house as previously described.[1] FAM N-Hydroxysuccinimide (NHS) ester and Sulfo-Cy5 NHS esters were purchased from Lumiprobe (Cockeysville, MD). All other reagents were purchased from Sigma Aldrich (St. Louis, MO).

### Oligonucleotide synthesis and purification

All oligonucleotides used in this study are shown in Table S1. RNA chemical synthesis, deprotection, and purification was described in our previous work.[2]

### Preparation of mutagenized DNA library D-Lib14

A binomial model was used to predict the abundance of total abundance of sequences with mutation number of $k$, shown. The probability or relative abundance ($P_k$) of a library sequence with a $l$-nt mutagenized domain at a mutation rate of $P_{mut}$, can be expressed with Equation 1. Abundance of each unique sequences with a mutation number of $k$, $P_k^{seq}$, can be expressed using Equation 2. The sequence distribution predicted using both equations was thought to be ideal when $P_{mut} = 14\%$.

$$P_k = \binom{l}{k} \times {P_{Mut}}^k \times (1 - P_{Mut})^{l-k}$$

**Equation 1.** Binomial distribution model used to calculate percent abundance ($P_k$) of sequences with a mutation count of k in a library with a mutagenized region. $l$: Length of mutagenized region. $P_{Mut}$: Mutation rate.

$$P_k^{seq} = \frac{P_k}{\binom{l}{k} \times 3^k}$$

**Equation 2.** Equation for calculating probability $P_k^{seq}$ of each unique sequence with a mutation count of k.

To introduce a 14% mutation rate across the aptamer region, a doping strategy was employed using all four deoxyribonucleotide phosphoramidites (A, C, G, and T). Each phosphoramidite was dissolved in anhydrous acetonitrile at a final concentration of 133 mM. To ensure that the overall probability of mutation at each nucleotide position was precisely 14% (with an 86% probability of retaining the wild-type base), a randomized mixture of the four phosphoramidites was prepared, proportioned according to the molar fractions listed below. These ratios account for differences in relative coupling efficiencies among the phosphoramidites to ensure equimolar incorporation potential. Next, this randomized mixture was doped into each individual phosphoramidite solution. The doping volumes were calculated using Equation 4, which was derived from Equation 3. This adjusts for both the desired mutation rate and the relative coupling efficiencies. This process yielded four separate amidite solutions, each containing a defined proportion of the randomized mixture. As a result, each solution was calibrated to introduce a 14% chance of base substitution (mutation) at each target site, while retaining an 86% chance for the original base (N).

Relative coupling efficiency of D-DNA phosphoramidites and molar fraction used to accommodate for different coupling efficiencies:

| ***Name*** | ***Rel. Eff.*** | ***Mol. Frac.*** |
|---|---|---|
| Pac-dA-CE | 1 | 31% |
| AC-dC-CE | 1.2 | 25% |
| iPr-Pac-dG-CE | 1.3 | 24% |
| dT-CE | 1.5 | 20% |

$$P_{Mut} = \frac{n_{mut}}{n_B + n_{mut}} = \frac{V_D \times E_D \times (1 - 25\%)}{V_N \times E_N + V_D \times E_D}$$

**Equation 3.** Equation describing how rate of mutation within the mutagenized region ($P_{Mut}$) is achieved through mixing of phosphoramidite solutions. A certain volume ($V_D$) of mixture containing all four phosphoramidites was doped into a solution containing a single phosphoramidite ($N$) with a volume of $V_N$. Differences between coupling efficiencies are normalized through using efficiency of $N$ ($E_N$) and efficiency of mixture used to dope in ($E_D$). The mixture containing all four bases has equal efficiency for each base, *i.e.* $E_D = (E_A + E_G + E_C + E_T)/4$.

$$V_D = \frac{4P_{Mut} \times V_N}{3 - 4P_{Mut}} \times \frac{E_N}{E_D}$$

**Equation 4.** Derived equation for calculation of volume of doping mixture needed ($V_D$) for a certain volume of $N$ phosphoramidite ($V_N$). Generate by solving Equation 3.

The mutagenized aptamer domain was synthesized using these doped amidite solutions following the manufacturer's recommended solid-phase synthesis protocols. For the synthesis of primer binding domains, only pure, undoped phosphoramidite solutions were used to ensure sequence fidelity. Mutagenesis SELEX RNA library, D-Lib$_{14}$, preparation was described previously[3] except for Turbo DNase (0.1 U/μL) was introduced after 2 hour incubation.

**Mutagenesis-based SELEX**

A 100 μL reaction mixture containing 1 nmol of either D-Lib14 RNA library, 5 mM $MgCl_2$, 25 mM NaCl, and 50 mM Tris (pH 7.6) was annealed at 70 °C for 1 minute and allowed to cool slowly to room temperature. The reaction mixture was added onto 1 mg of streptavidin coated magnetic beads (washed and pre-blocked following a previously established protocol[4]) and incubated at 23 °C for 30 minutes. The supernatant was removed and the beads discarded to remove any bead-binding RNAs. To the retained supernatant was added 200 pmol of both 3′- biotinylated L-TAR$_A$ and L-TAR$_U$. The solution was then incubated at room temperature for 30 minutes. At this point, 1 mg of pre-blocked streptavidin coated magnetic beads was added. After incubating for 15 minutes at 23 °C, the beads were washed five times with 1 mL of the same buffer in order to remove weakly bound species. The retained RNA was eluted using two 150 μL portions of a stripping solution containing 25 mM NaOH and 1 mM EDTA. Both aliquots were quickly combined into a solution containing 30 μL of 3 M NaOAc (pH 5.4), 30 μL of 1 M Tris (pH 7.6), and 5 μL glycogen (1 mg/mL). The RNA was then ethanol precipitated and used directly in a 50 μL reverse transcription reaction containing 10 U/μL RT, 3 mM $MgCl_2$, 75 mM KCl, 10 mM DTT, 50 mM Tris (pH 8.3), and 0.5 mM each of the four dNTPs, which was incubated at 42 °C for 1 hour. The resulting cDNA was added as a template into a 100 μL scale PCR and amplified using FWD.primer and REV.primer (Table S1). The amplified DNA was ethanol precipitated and approximately half of the material was used to amplify further in a 10 mL PCR using the same procedures. Half of the further amplified material was used to generate a new RNA library for the next round of selection. Subsequent selection rounds were performed in the same manner except for using 200 pmol library and single target in deconvolution rounds.

### Next generation sequencing

Products of 100 µL scale PCR reactions from each selection round along with that of the original library were PCR amplified using EZseq.FWD and EZSeq.REV primers containing Illumina adapters. Resulted PCR product was purified through agarose gel and extracted using QIAquick Gel Extraction Kit (QIAGEN, Hilden, Germany) then sequenced by GENEWIZ (South Plainfield, NJ) through Amplicon-EZ service.

Result files from sequencing were processed using *Pear*[5] and *Cutadapt*[6] tools on usegalaxy.org[7]. Trimmed data merged using *FASTAptameR 2.0.*[8]

### *In vitro* transcription of RNA

DNA templates used to prepare cross-chiral aptamer candidates were generated by a cross-extension of two overlapping synthetic oligonucleotides (Table S2), which were subsequently transcribed into RNA and purified as described previously.[2]

### Electrophoretic mobility assays

The dissociation constant ($K_d$) of aptamer sequences for TAR variants was determined by EMSA as described previously.[3] Briefly, a trace amount (2 nM) of the indicated TAR hairpin ligand was mixed with various concentration of aptamer candidates or library in a 20 µL reaction mixture containing 10 mM $MgCl_2$, 50 mM NaCl, 25 mM Tris (pH 7.6), and 8% glycerol. After incubating for 30 minutes at 23 °C, an aliquot was removed and analyzed by 10% native PAGE (29:1 acrylamide:bisacrylamide) containing 10 mM $Mg(OAc)_2$, 50 mM NaOAc and 25 mM HEPES (pH 7.6). Electrophoresis was carried out at 30 mA for 1 h at 4 °C and the gel visualized by fluorescence using either a Typhoon FLA-9500 Molecular Imager (General Electric Co., Boston, MA) or ChemiDoc MP Imaging System (Bio-Rad Laboratories, Inc., Fort Worth, TX).

### RNA pulldown assay

A 300 µL reaction mixture containing 40 pmol of annealed biotinylated aptamer (L-$5_A$-16t, L-$5_U$-1t, or L-6-4t), 5 mM $MgCl_2$, 25 mM NaCl, and 50 mM Tris (pH 7.6) incubated with 0.2 mg of streptavidin coated magnetic beads (washed following a previously established protocol[4]) and incubated at 23 °C for 30 minutes. After degrading supernatants, each bead preparation was washed 5 times with a buffer containing 5 mM $MgCl_2$, 25 mM NaCl, and 50 mM Tris (pH 7.6). A

250 μL mixture containing 250 pmol of each annealed transcript (D-HP$_{80}$, D-HP$_{60}$, and D-HP$_{40}$), 5 mM MgCl$_2$, 25 mM NaCl, and 50 mM Tris (pH 7.6) was incubated with L-5$_A$-16t, L-5$_U$-1t, and L-6-4t coated beads sequentially for 15 minutes at 4 °C each. The beads were washed one time and eluted as described previously. Gel was stained with 1X SYBR™ Gold Nucleic Acid Gel Stain and visualized by fluorescence using ChemiDoc MP Imaging System (Bio-Rad Laboratories, Inc., Fort Worth, TX).

## SUPPLEMENTARY FIGURES

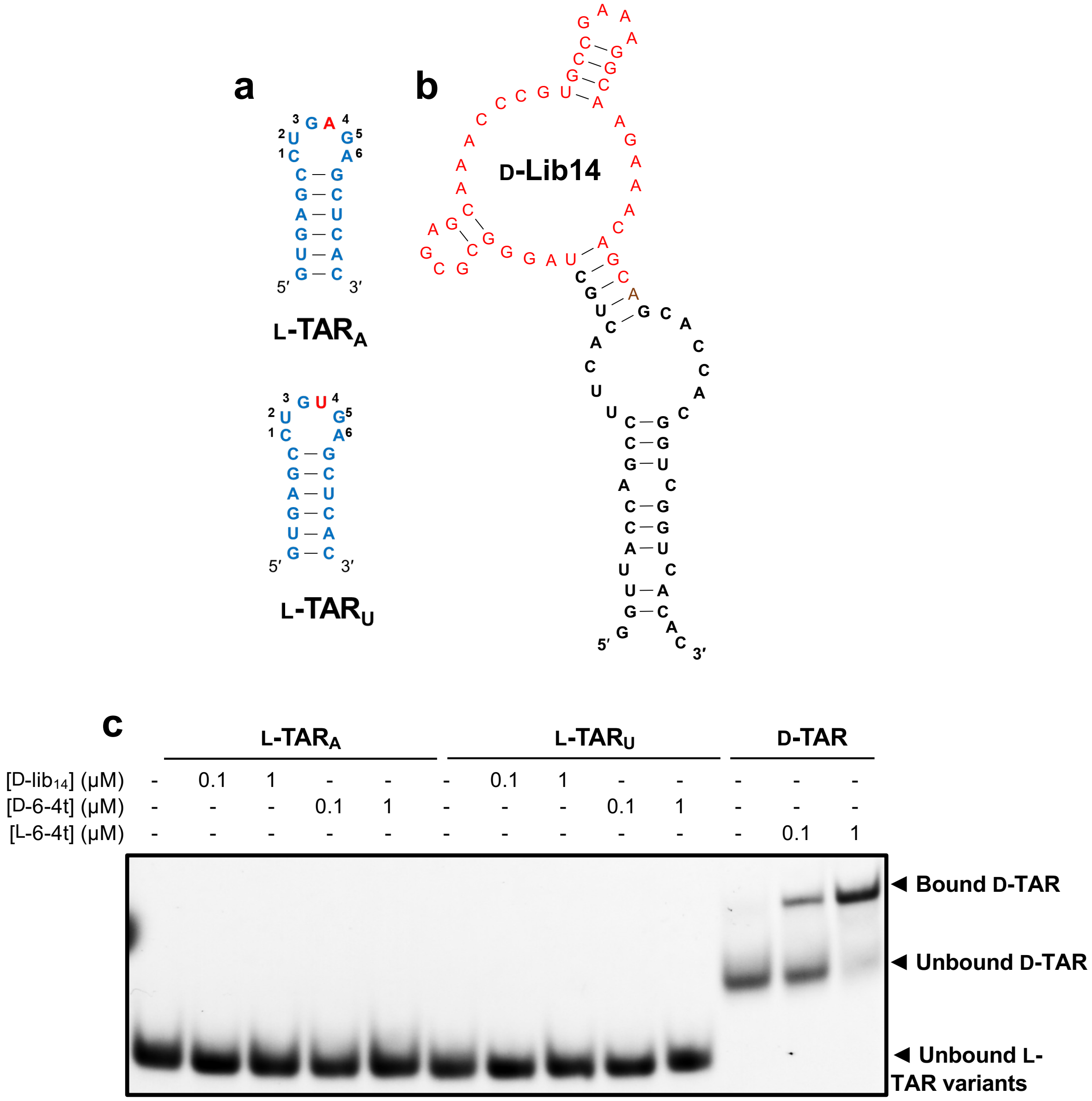


**Figure S1.** (a) Sequence and predicted secondary structure of L-RNA TAR variants L-TAR$_A$ and L-TAR$_U$. 3′-biotinylated L-TAR$_A$ or L-TAR$_U$ are referred to as L-TAR$_A$-BIO and L-TAR$_U$-BIO, respectively, and were used during the SELEX process. See Table S1. (b) Sequence and predicted secondary structure of D-Lib14. Red nucleotides were mutagenized. (c) EMSA analysis (10% native PAGE; 19:1 acrylamide:bis-acrylamide) of L-TAR variants binding to D-6-4t. Reactions employed 1 μM of either D-Lib14 or 6-4t aptamer, 5 nM either L-TAR$_A$ or L-TAR$_U$, 50 mM NaCl, 10 mM MgCl2, 8% v/v glycerol, and 25 mM Tris (pH 7.6) and were incubated at 23 °C for 30 minutes before analyzed via native PAGE. Uncropped gel image is shown in Figure S16.

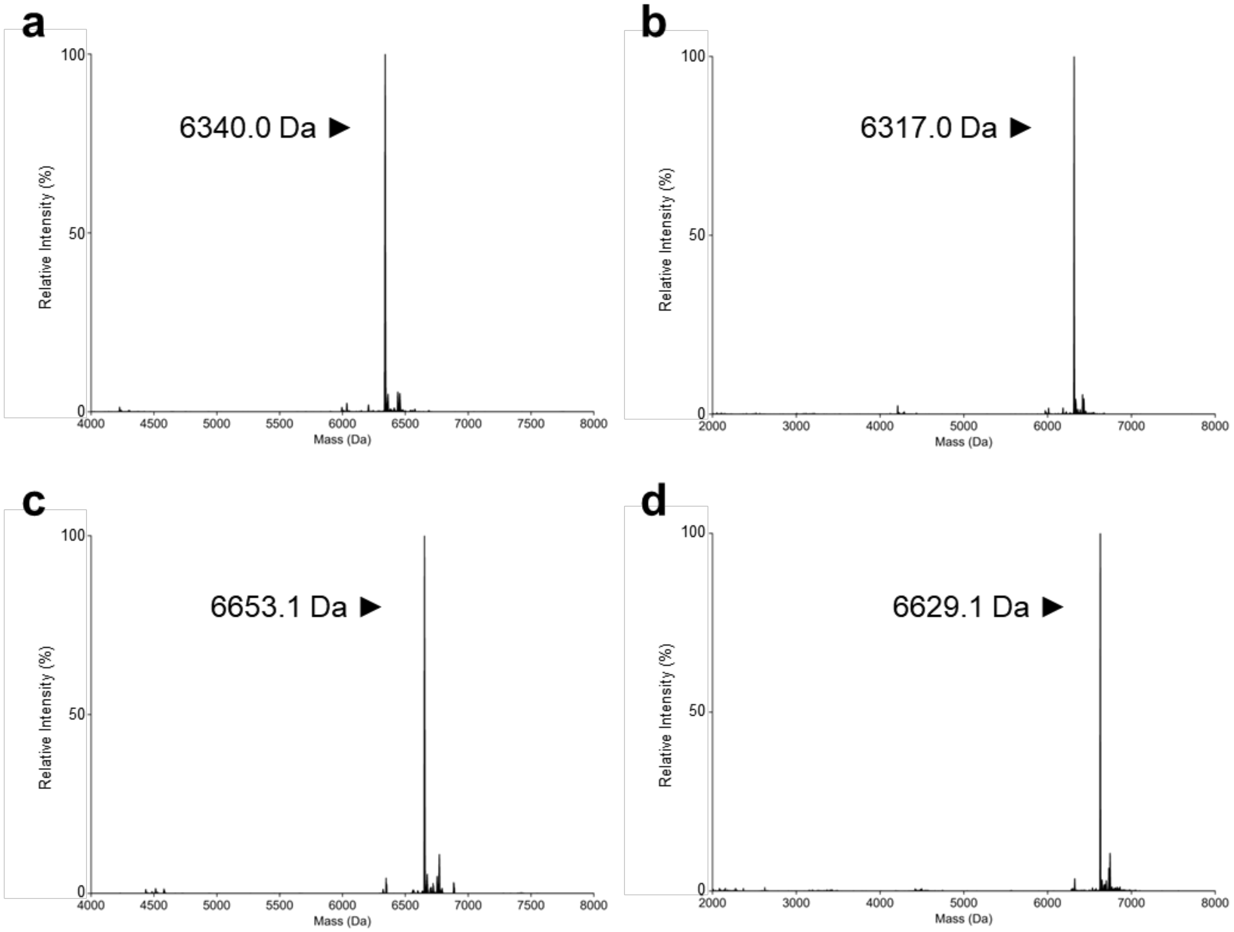


**Figure S2.** ESI-MS characterization of L-TAR variants used during SELEX and EMSAs. (a) ESI-MS data for L-TAR$_{A}$-BIO. Calculated: 6340.1 Da, observed: 6340.0 Da. (b) ESI-MS data for L-TAR$_{U}$-BIO. Calculated: 6317.1 Da, observed: 6317.0 Da. (c) ESI-MS data for L-TAR$_{A}$. Calculated: 6652.3 Da, observed: 6653.1 Da. (d) ESI-MS data for L-TAR$_{U}$. Calculated:6629.3 Da, observed: 6629.1 Da.

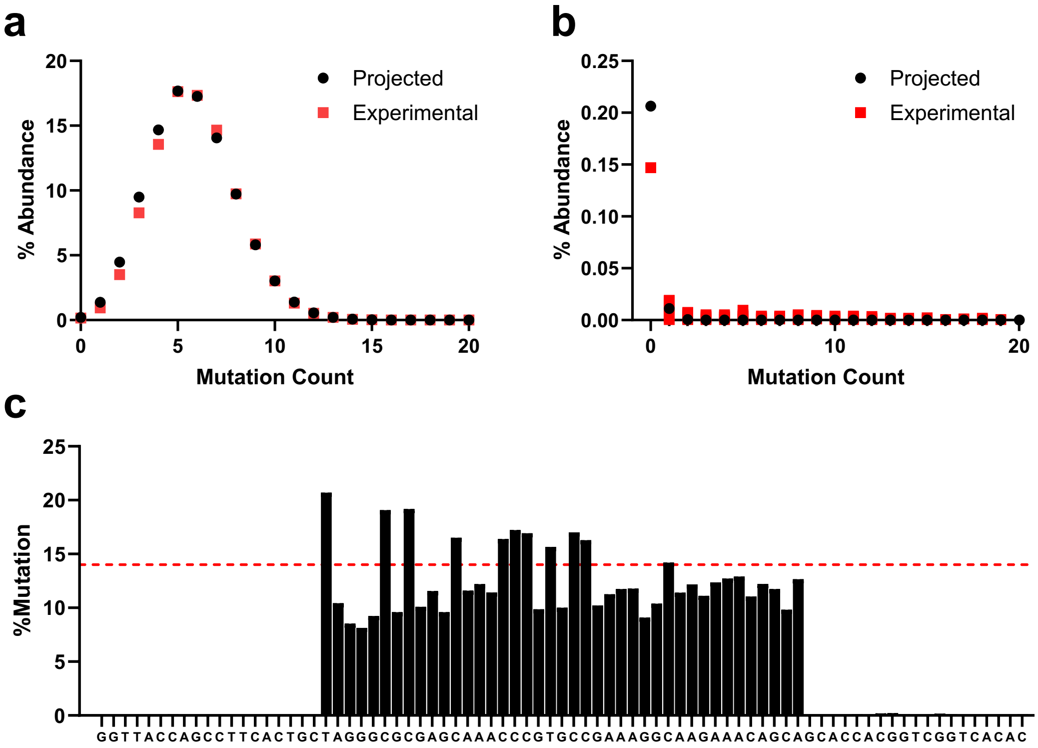
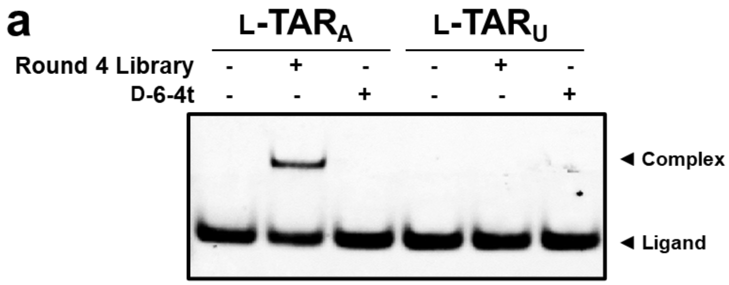


**Figure S3.** Comparison of projected and experimentally observed properties of the protected mutagenized library. (a) Projected and experimentally measured total abundance of sequences in the D-Lib14 library as a function of the number of mutations per sequence. (b) Projected and experimentally measured abundance of unique sequences in the round 0 library as a function of the number of mutations per sequence. (c) Position-specific mutation frequencies observed in the round 0 library. The red dashed line indicates the designed theoretical mutation rate of 14%.



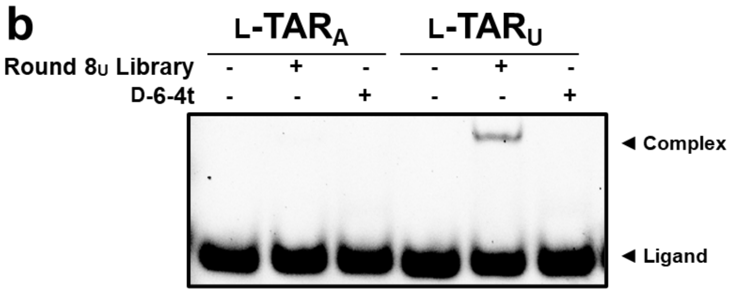


**Figure S4.** EMSA analysis (10% native PAGE; 19:1 acrylamide:bis-acrylamide) of enriched D-Lib14 binding to L-TAR variants following SELEX round 4 (a) or 8 (b). Reactions employed 1 μM of either D-Lib14 or D-6-4t RNA, 5 nM of the indicated L-TAR substrate, 50 mM NaCl, 10 mM MgCl2, 8% v/v glycerol, and 25 mM Tris (pH 7.6) and were incubated at 23 °C for 30 minutes prior to gel loading. Uncropped gel images are shown in Figure S17.

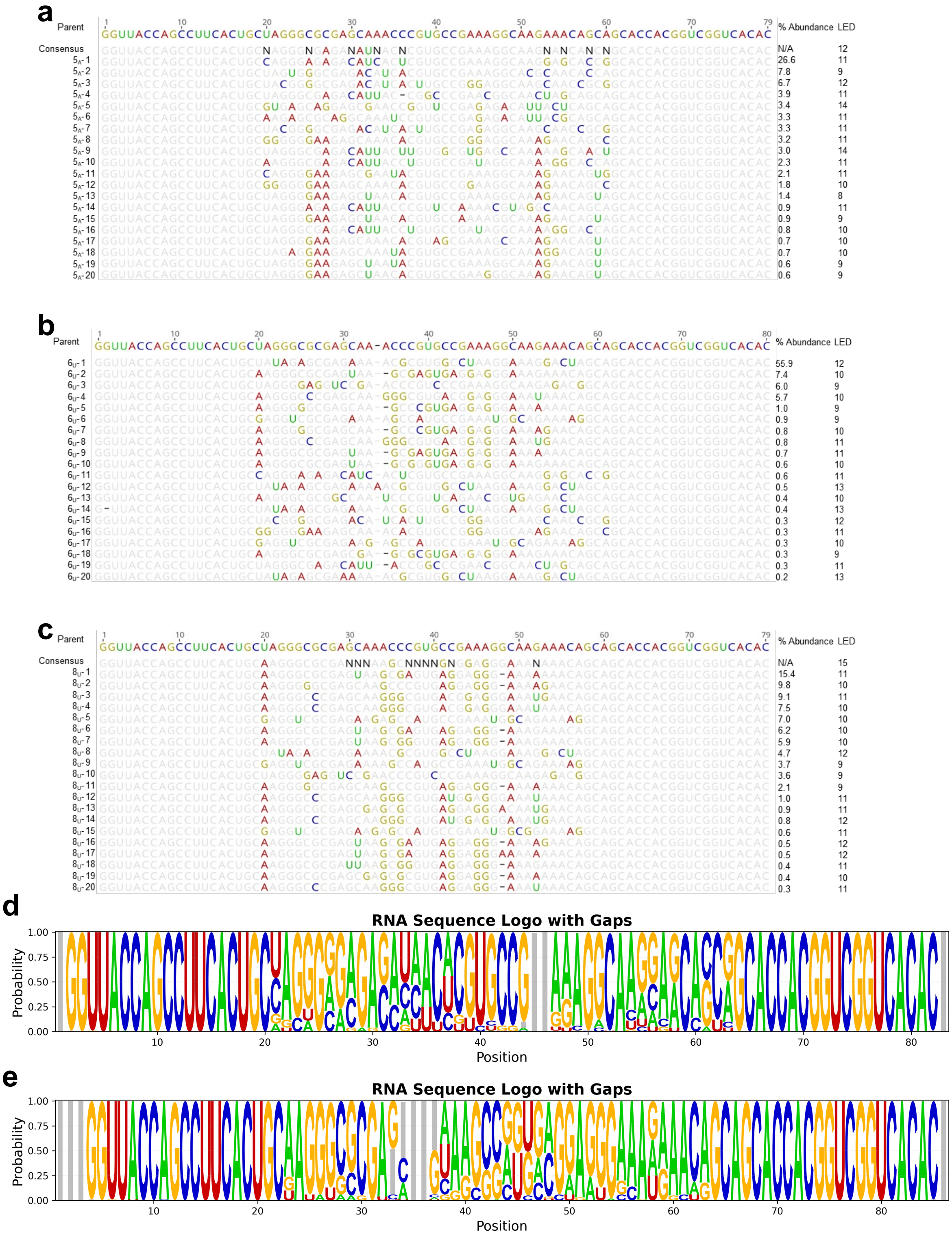
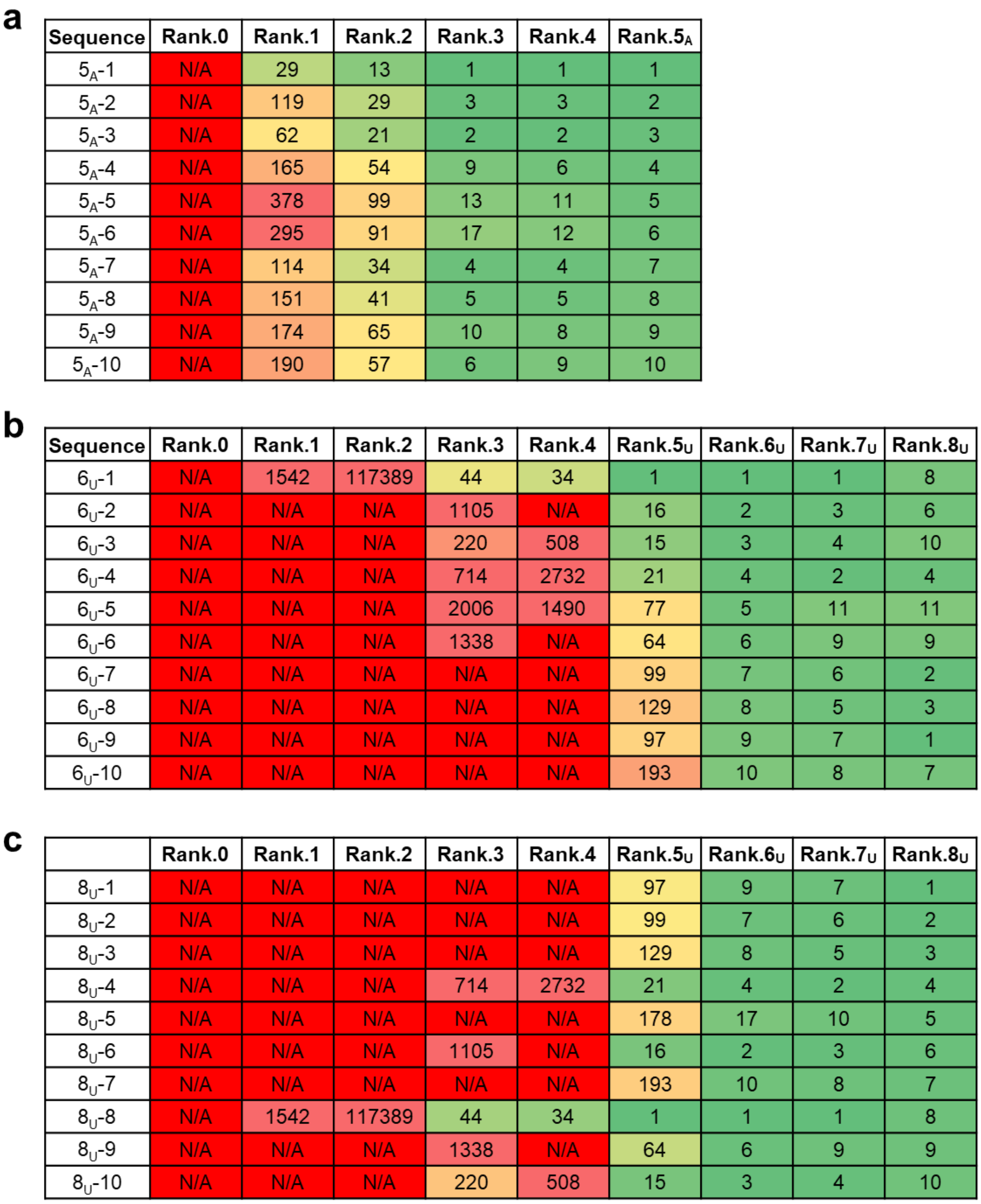
a
Parent
GGUUACCAGCCUUCACUGCUAGGGCGCGAGCAAACCCGUGCCGAAAGGCAAGAAACAGCAGCACCACGGUCGGUCACAC
% Abundance
LED
Consensus
GGUUACCAGCCUUCACUGCNAGGGNGAGANAUNACNCGUGCCGAAAGGCAAGNANCANCNGCACCACGGUCGGUCACAC
N/A 12
$5_A$-1 26.6 11
$5_A$-2 7.8 9
$5_A$-3 6.7 12
$5_A$-4 3.9 11
$5_A$-5 3.4 14
$5_A$-6 3.3 11
$5_A$-7 3.3 11
$5_A$-8 3.2 11
$5_A$-9 3.0 14
$5_A$-10 2.3 11
$5_A$-11 2.1 11
$5_A$-12 1.8 10
$5_A$-13 1.4 8
$5_A$-14 0.9 11
$5_A$-15 0.9 9
$5_A$-16 0.8 10
$5_A$-17 0.7 10
$5_A$-18 0.7 10
$5_A$-19 0.6 9
$5_A$-20 0.6 9
b
Parent
GGUUACCAGCCUUCACUGCUAGGGCGCGAGCAA-ACCCGUGCCGAAAGGCAAGAAACAGCAGCACCACGGUCGGUCACAC
% Abundance
LED
$6_U$-1 55.9 12
$6_U$-2 7.4 10
$6_U$-3 6.0 9
$6_U$-4 5.7 10
$6_U$-5 1.0 9
$6_U$-6 0.9 9
$6_U$-7 0.8 10
$6_U$-8 0.8 11
$6_U$-9 0.7 11
$6_U$-10 0.6 10
$6_U$-11 0.6 11
$6_U$-12 0.5 13
$6_U$-13 0.4 10
$6_U$-14 0.4 13
$6_U$-15 0.3 12
$6_U$-16 0.3 11
$6_U$-17 0.3 10
$6_U$-18 0.3 9
$6_U$-19 0.3 11
$6_U$-20 0.2 13
c
Parent
GGUUACCAGCCUUCACUGCUAGGGCGCGAGCAAACCCGUGCCGAAAGGCAAGAAACAGCAGCACCACGGUCGGUCACAC
% Abundance
LED
Consensus
GGUUACCAGCCUUCACUGCAAGGGCGCGANNNAAGCNNNNGNGGAGGGAAANAAACAGCAGCACCACGGUCGGUCACAC
N/A 15
$8_U$-1 15.4 11
$8_U$-2 9.8 10
$8_U$-3 9.1 11
$8_U$-4 7.5 10
$8_U$-5 7.0 10
$8_U$-6 6.2 10
$8_U$-7 5.9 10
$8_U$-8 4.7 12
$8_U$-9 3.7 9
$8_U$-10 3.6 9
$8_U$-11 2.1 9
$8_U$-12 1.0 11
$8_U$-13 0.9 11
$8_U$-14 0.8 12
$8_U$-15 0.6 11
$8_U$-16 0.5 12
$8_U$-17 0.5 12
$8_U$-18 0.4 11
$8_U$-19 0.4 10
$8_U$-20 0.3 11
d
RNA Sequence Logo with Gaps
Probability
0.00
0.25
0.50
0.75
1.00
Position
10
20
30
40
50
60
70
80
e
RNA Sequence Logo with Gaps
Probability
0.00
0.25
0.50
0.75
1.00
Position
10
20
30
40
50
60
70
80

**Figure S5.** (Previous page) Sequence identities of the top-20 most abundant sequences in round $5_A$ (a), $6_U$ (b), and $8_U$ (c). Grey nucleotides are unchanged compared to parent sequence on the top of the table. Abundance (%) and LED are shown to the right of each sequence. (d) Sequence LOGO for the 1,000 most abundant sequences in round $5_A$. (e) Sequence LOGO for the 1,000 most abundant sequences in round $8_U$.



**a**

| Sequence | Rank.0 | Rank.1 | Rank.2 | Rank.3 | Rank.4 | Rank.$5_A$ |
|---|---|---|---|---|---|---|
| $5_A$-1 | N/A | 29 | 13 | 1 | 1 | 1 |
| $5_A$-2 | N/A | 119 | 29 | 3 | 3 | 2 |
| $5_A$-3 | N/A | 62 | 21 | 2 | 2 | 3 |
| $5_A$-4 | N/A | 165 | 54 | 9 | 6 | 4 |
| $5_A$-5 | N/A | 378 | 99 | 13 | 11 | 5 |
| $5_A$-6 | N/A | 295 | 91 | 17 | 12 | 6 |
| $5_A$-7 | N/A | 114 | 34 | 4 | 4 | 7 |
| $5_A$-8 | N/A | 151 | 41 | 5 | 5 | 8 |
| $5_A$-9 | N/A | 174 | 65 | 10 | 8 | 9 |
| $5_A$-10 | N/A | 190 | 57 | 6 | 9 | 10 |

**b**

| Sequence | Rank.0 | Rank.1 | Rank.2 | Rank.3 | Rank.4 | Rank.$5_U$ | Rank.$6_U$ | Rank.$7_U$ | Rank.$8_U$ |
|---|---|---|---|---|---|---|---|---|---|
| $6_U$-1 | N/A | 1542 | 117389 | 44 | 34 | 1 | 1 | 1 | 8 |
| $6_U$-2 | N/A | N/A | N/A | 1105 | N/A | 16 | 2 | 3 | 6 |
| $6_U$-3 | N/A | N/A | N/A | 220 | 508 | 15 | 3 | 4 | 10 |
| $6_U$-4 | N/A | N/A | N/A | 714 | 2732 | 21 | 4 | 2 | 4 |
| $6_U$-5 | N/A | N/A | N/A | 2006 | 1490 | 77 | 5 | 11 | 11 |
| $6_U$-6 | N/A | N/A | N/A | 1338 | N/A | 64 | 6 | 9 | 9 |
| $6_U$-7 | N/A | N/A | N/A | N/A | N/A | 99 | 7 | 6 | 2 |
| $6_U$-8 | N/A | N/A | N/A | N/A | N/A | 129 | 8 | 5 | 3 |
| $6_U$-9 | N/A | N/A | N/A | N/A | N/A | 97 | 9 | 7 | 1 |
| $6_U$-10 | N/A | N/A | N/A | N/A | N/A | 193 | 10 | 8 | 7 |

**c**

|  | Rank.0 | Rank.1 | Rank.2 | Rank.3 | Rank.4 | Rank.$5_U$ | Rank.$6_U$ | Rank.$7_U$ | Rank.$8_U$ |
|---|---|---|---|---|---|---|---|---|---|
| $8_U$-1 | N/A | N/A | N/A | N/A | N/A | 97 | 9 | 7 | 1 |
| $8_U$-2 | N/A | N/A | N/A | N/A | N/A | 99 | 7 | 6 | 2 |
| $8_U$-3 | N/A | N/A | N/A | N/A | N/A | 129 | 8 | 5 | 3 |
| $8_U$-4 | N/A | N/A | N/A | 714 | 2732 | 21 | 4 | 2 | 4 |
| $8_U$-5 | N/A | N/A | N/A | N/A | N/A | 178 | 17 | 10 | 5 |
| $8_U$-6 | N/A | N/A | N/A | 1105 | N/A | 16 | 2 | 3 | 6 |
| $8_U$-7 | N/A | N/A | N/A | N/A | N/A | 193 | 10 | 8 | 7 |
| $8_U$-8 | N/A | 1542 | 117389 | 44 | 34 | 1 | 1 | 1 | 8 |
| $8_U$-9 | N/A | N/A | N/A | 1338 | N/A | 64 | 6 | 9 | 9 |
| $8_U$-10 | N/A | N/A | N/A | 220 | 508 | 15 | 3 | 4 | 10 |

**Figure S6.** Per-round ranking (RPM) of the top-10 most abundant sequence in round $5_A$ (a), round $6_U$ (b), and round $8_U$ (c). N/A indicates that the sequence was not present in the data.

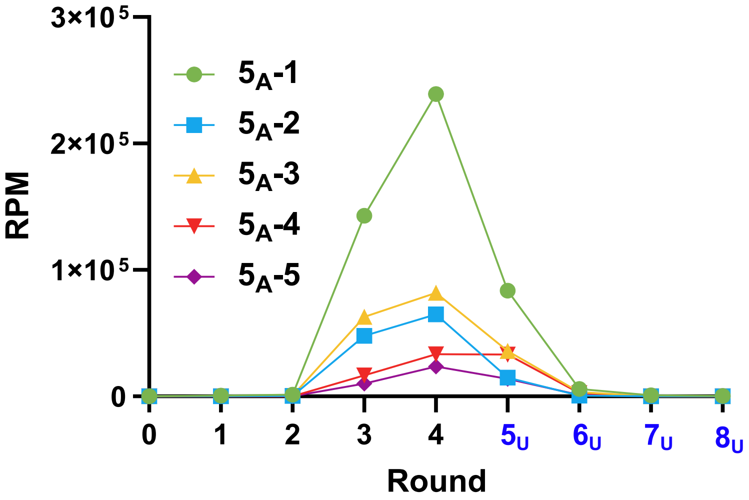


**Figure S7.** Enrichment behavior of the top-ranked L-TAR$_A$-binding sequences during deconvolution for L-TAR$_U$. Changes in reads per million (RPM) across deconvolution rounds are shown for the highest-ranking sequences identified in round 5$_A$.

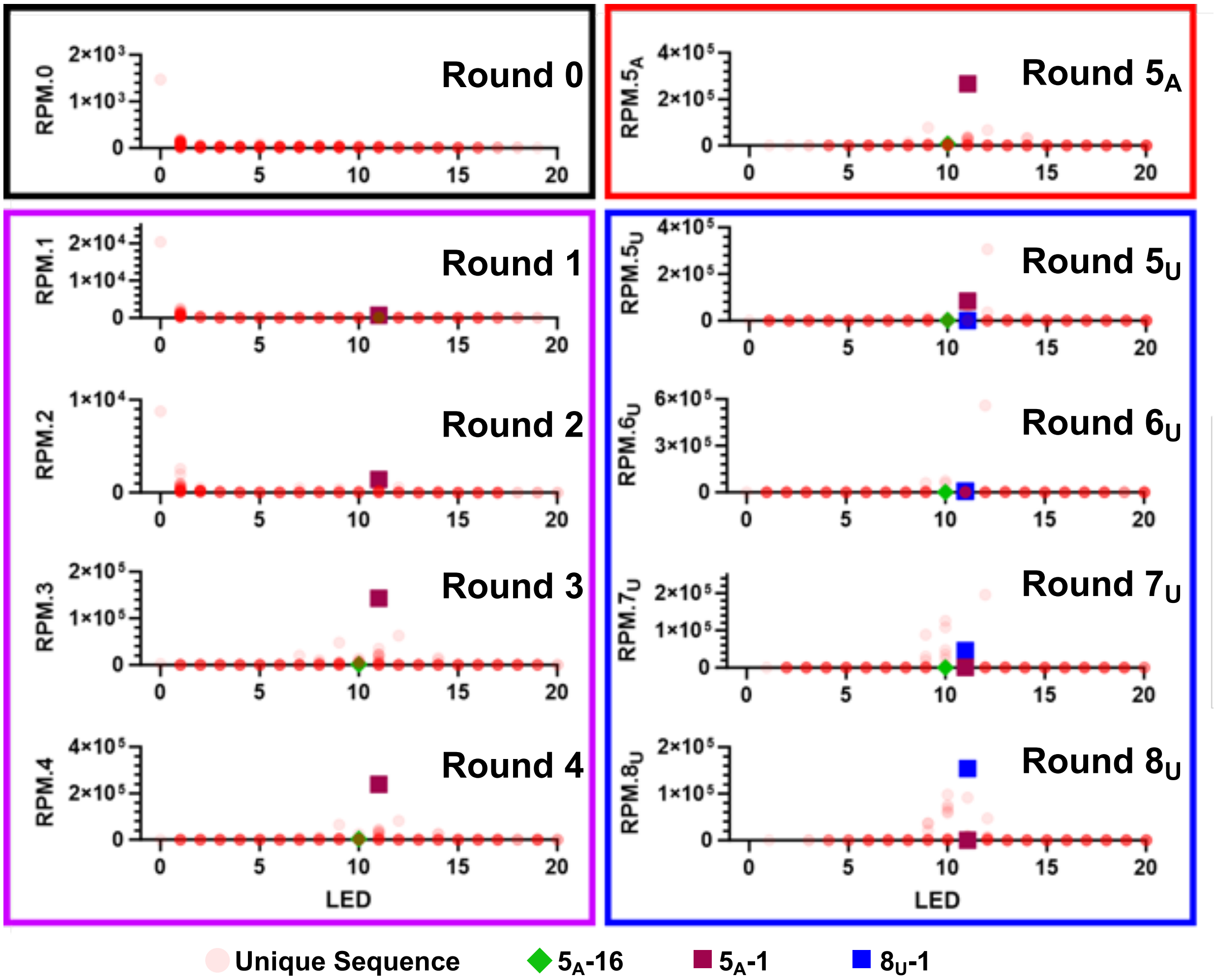


**Figure S8.** Distribution of sequence abundance as a function of Levenshtein edit distance (LED) across selection rounds. Scatter plots show reads per million (RPM) plotted against LED relative to the parent aptamer for individual sequences at each selection round. LED is defined as the minimum number of single-nucleotide substitutions, insertions, or deletions required to convert a given sequence into the parental aptamer. Each semi-transparent red circle represents a unique sequence, with increased color intensity indicating accumulation of distinct sequences with similar RPM and LED values. Panel outlines denote selection phase: black, initial library (round 0); purple, combined early rounds; red, deconvoluted rounds selected against L-TAR$_A$; blue, deconvoluted rounds selected against L-TAR$_U$. Highlighted sequences include 5A-1 (magenta square), 5A-16 (green diamond), and 8U-1 (blue square).

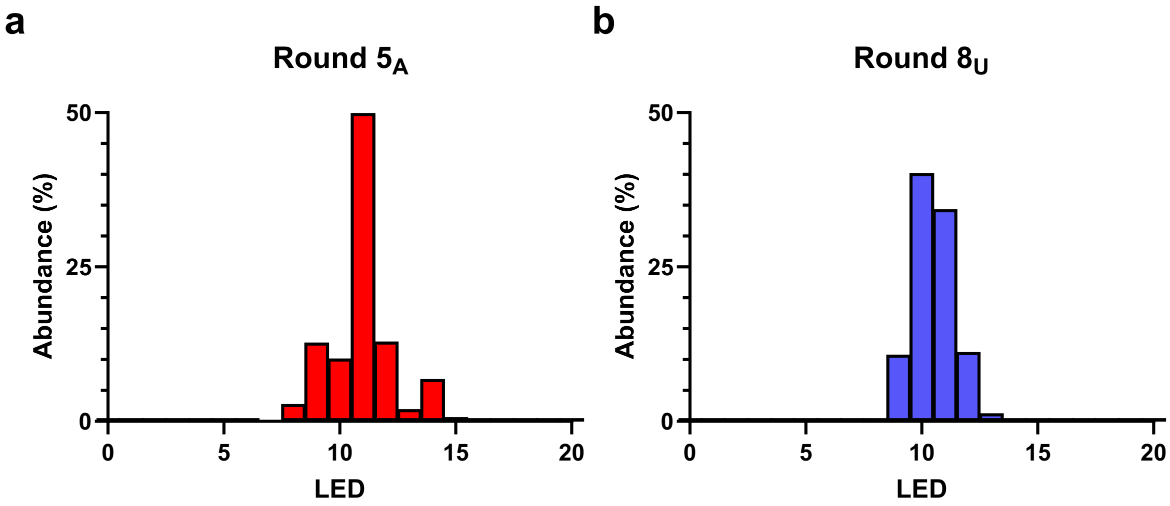


**Figure S9.** Levenshtein edit distance (LED) histograms for the 1,000 most abundant sequences from round $5_A$ (a) and $8_U$ (b).

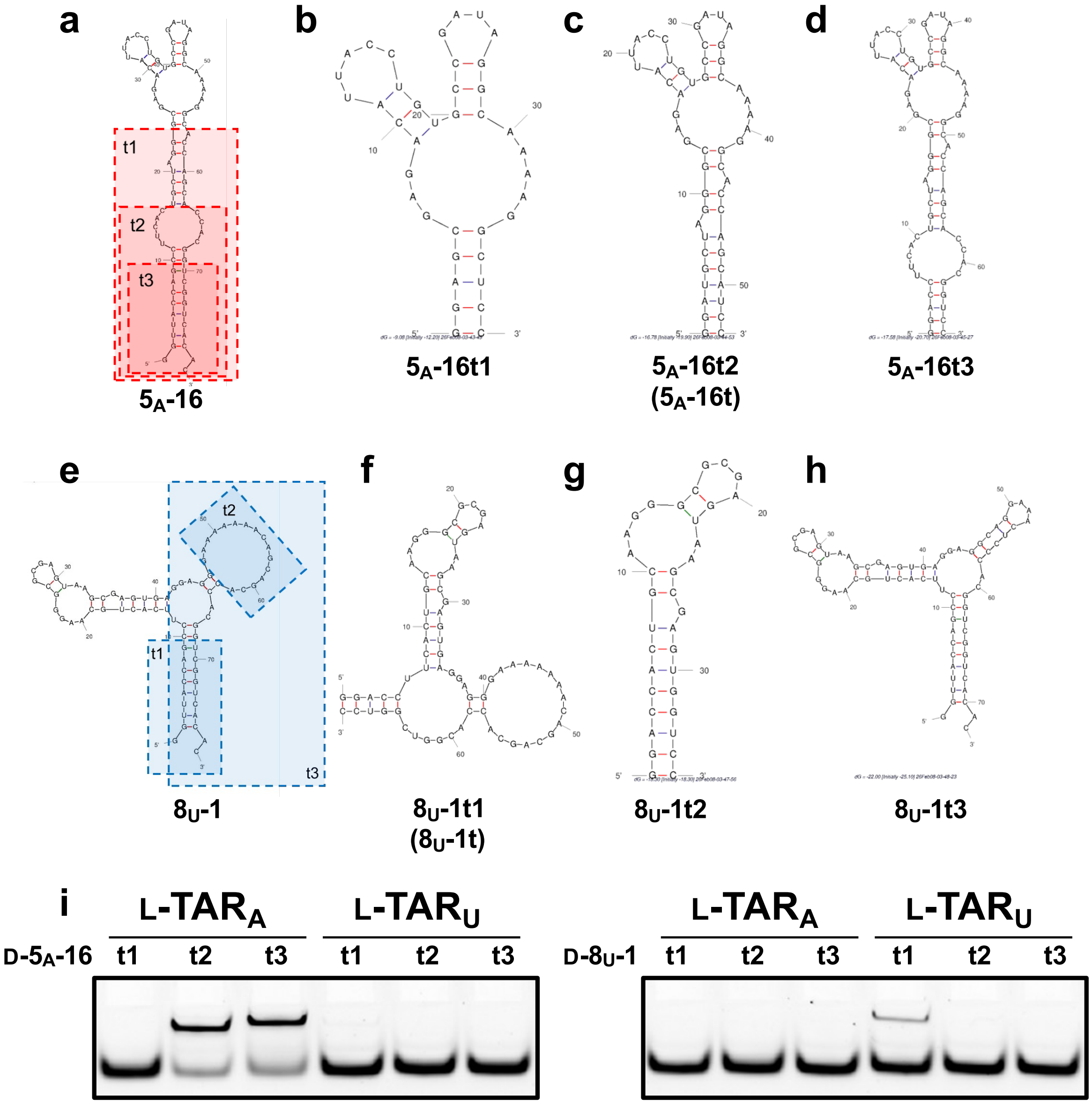


**Figure S10.** Secondary structure–guided truncation of evolved aptamer candidates. (a) Predicted secondary structure of aptamer 5$_A$-16 highlighting regions selected for truncation. Dashed boxes indicate segments removed to generate truncated variants 5$_A$-16t1 (b), 5$_A$-16t2 (c), and 5$_A$-16t3 (d). (e) Predicted secondary structure of aptamer 8$_U$-1 highlighting regions selected for truncation. Dashed boxes indicate segments removed to generate truncated variants 8$_U$-1t1 (f), 8$_U$-1t2 (g), 8$_U$-1t3 (i). Preliminary screening of truncated aptamer variants by EMSA (10% native PAGE; 19:1 acrylamide:bis-acrylamide). Reactions contained 1 μM D-aptamer, 5 nM FAM-labeled cognate target (L-TAR$_A$ or L-TAR$_U$), 50 mM NaCl, 10 mM $MgCl_2$, 8% (v/v) glycerol, and 25 mM Tris (pH 7.6), and were incubated at 23 °C for 30 min prior to separation. Uncropped gel images are shown in Figure S18.

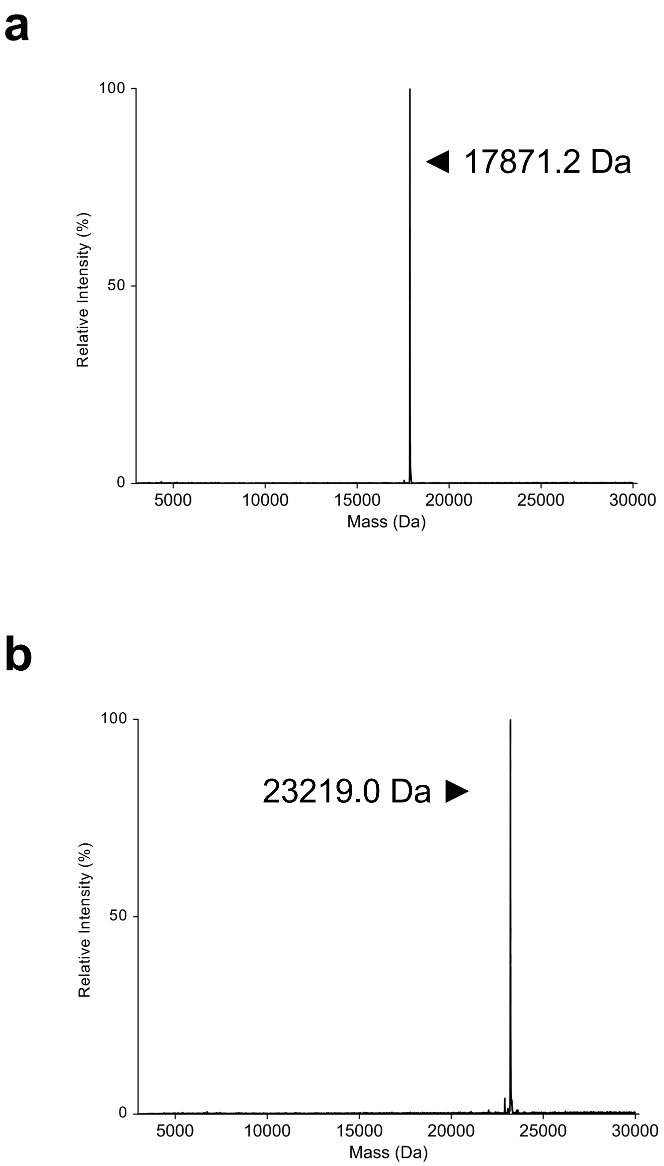


**Figure S11.** (a) ESI-MS data for L-$5_{\mathrm{A}}$-16t. Calculated: 17871.2 Da, observed: 17871.2 Da. (b) ESI-MS data for L-$8_{\mathrm{U}}$-1t. Calculated: 23217.5 Da, observed: 23219.0 Da.

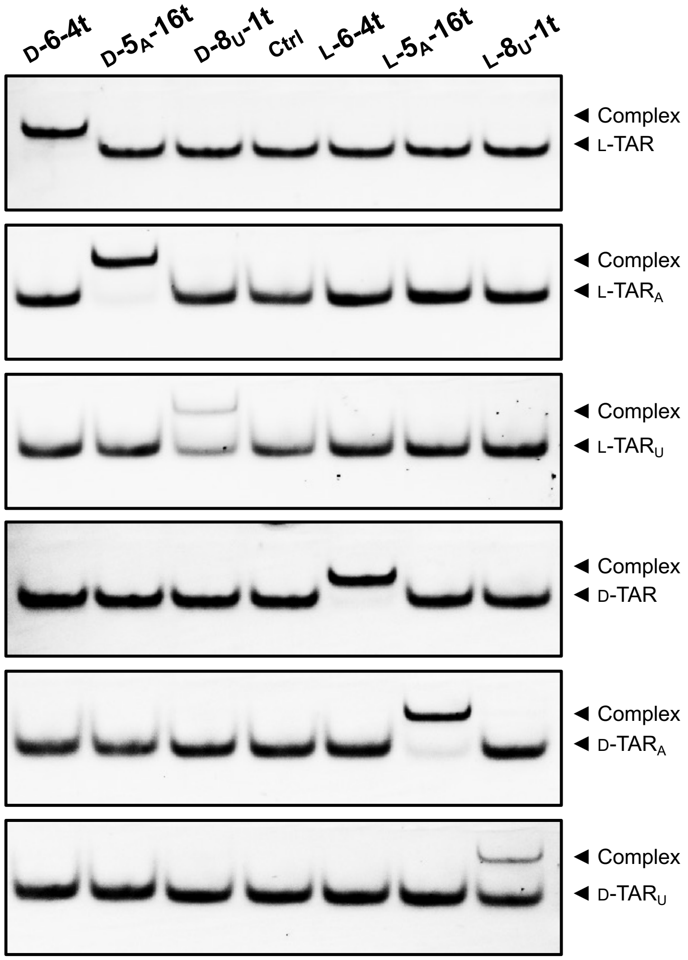


**Figure S12.** Comprehensive cross-reactivity evaluation of selected aptamers by EMSA (10% native PAGE; 19:1 acrylamide:bis-acrylamide). Reactions employed 1 μM of the indicated RNA aptamer candidate, 5 nM either L-TAR$_{\text{A}}$ or L-TAR$_{\text{A}}$, 50 mM NaCl, 10 mM MgCl$_2$, 8% v/v glycerol, and 25 mM Tris (pH 7.6) and were incubated at 23 °C for 30 minutes prior to separation. Uncropped gel images are shown in Figure S19.

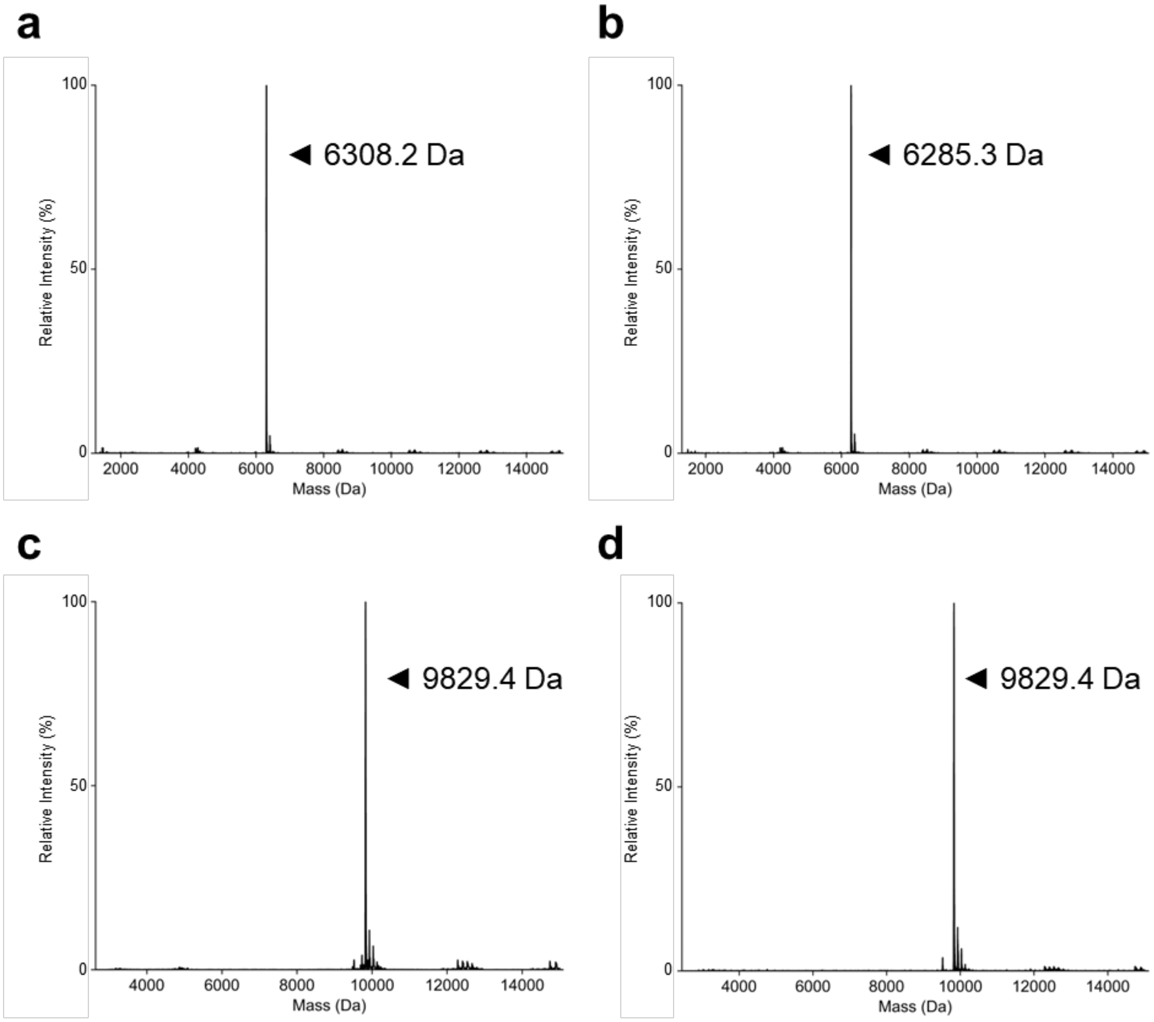


**Figure S13.** ESI-MS characterization of synthetic RNA hairpin used in this work. (a) ESI-MS data for D-TAR$_A$. Calculated: 6308.7 Da, observed: 6308.2 Da. (b) ESI-MS data for D-TAR$_U$. Calculated: 6285.7 Da, observed: 6285.3 Da. (c) ESI-MS data for D-TAR-FAM. Calculated: 9829.8 Da, observed: 9829.4 Da. (d) ESI-MS data for L-TAR-FAM. Calculated: 9829.8 Da, observed: 9829.4 Da.

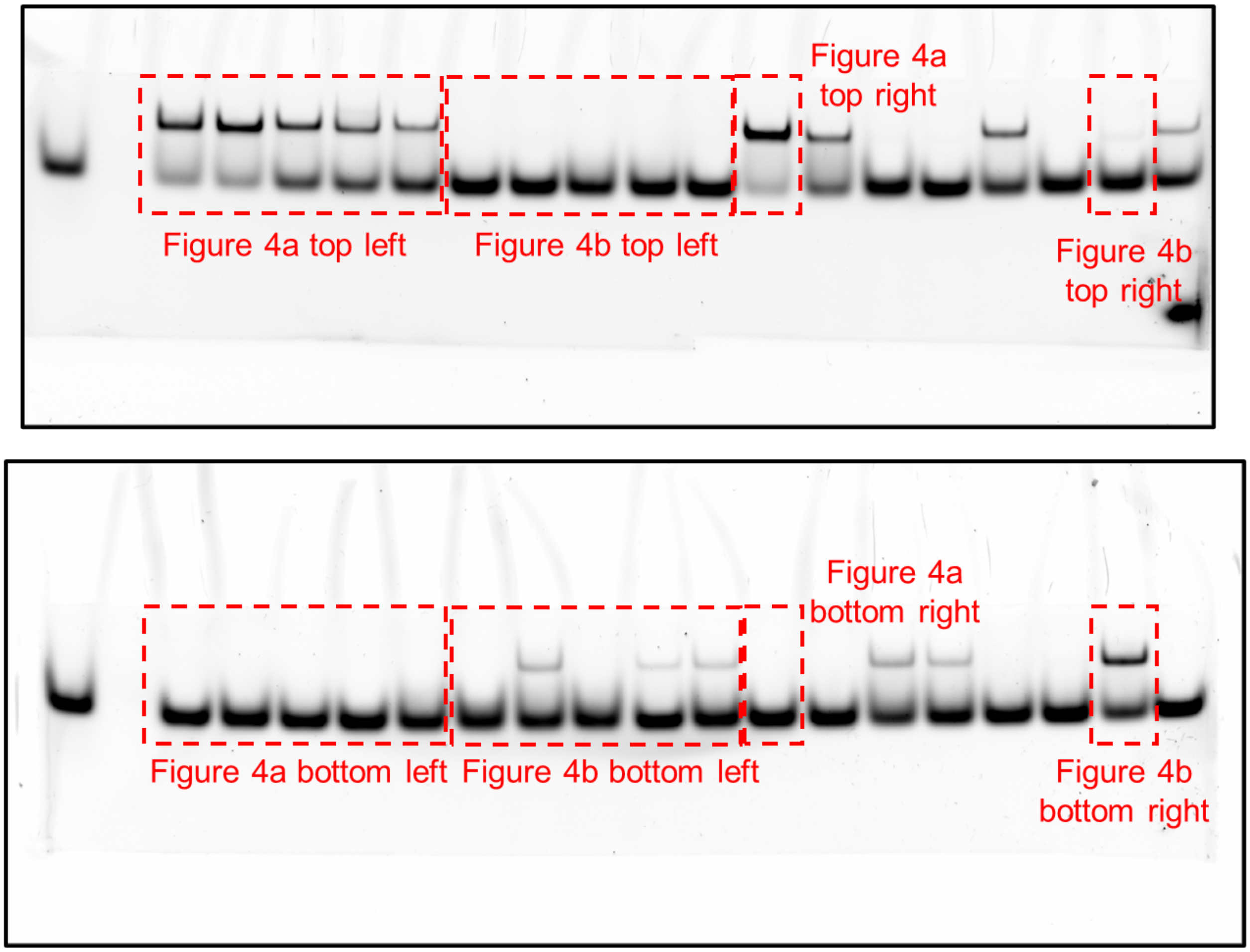


**Figure S14.** Uncropped gel image for Figure 4. Frame indicates the cropped region shown in Figure 4. Unboxed region contains unrelated experiments.

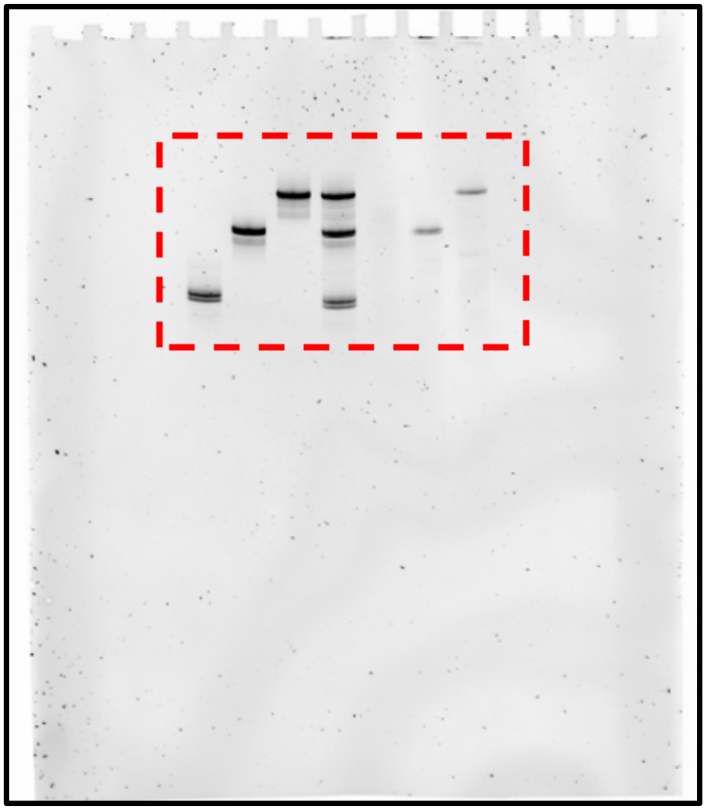


**Figure S15.** Uncropped gel image for Figure 5b. Frame indicates the cropped region shown in Figure 5b.

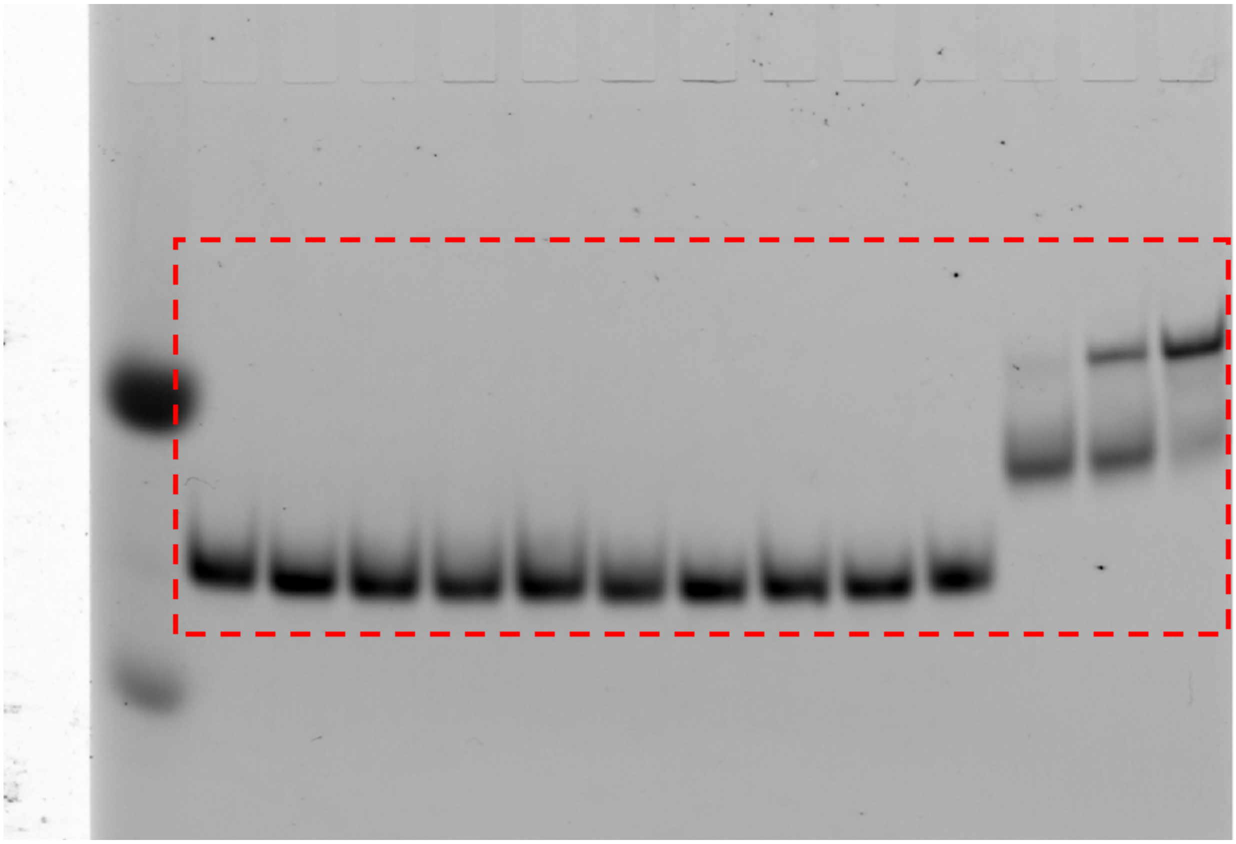


**Figure S16.** Uncropped gel image for Figure S1c. Frame indicates the cropped region shown in Figure S1c. Unboxed region contains loading dye.

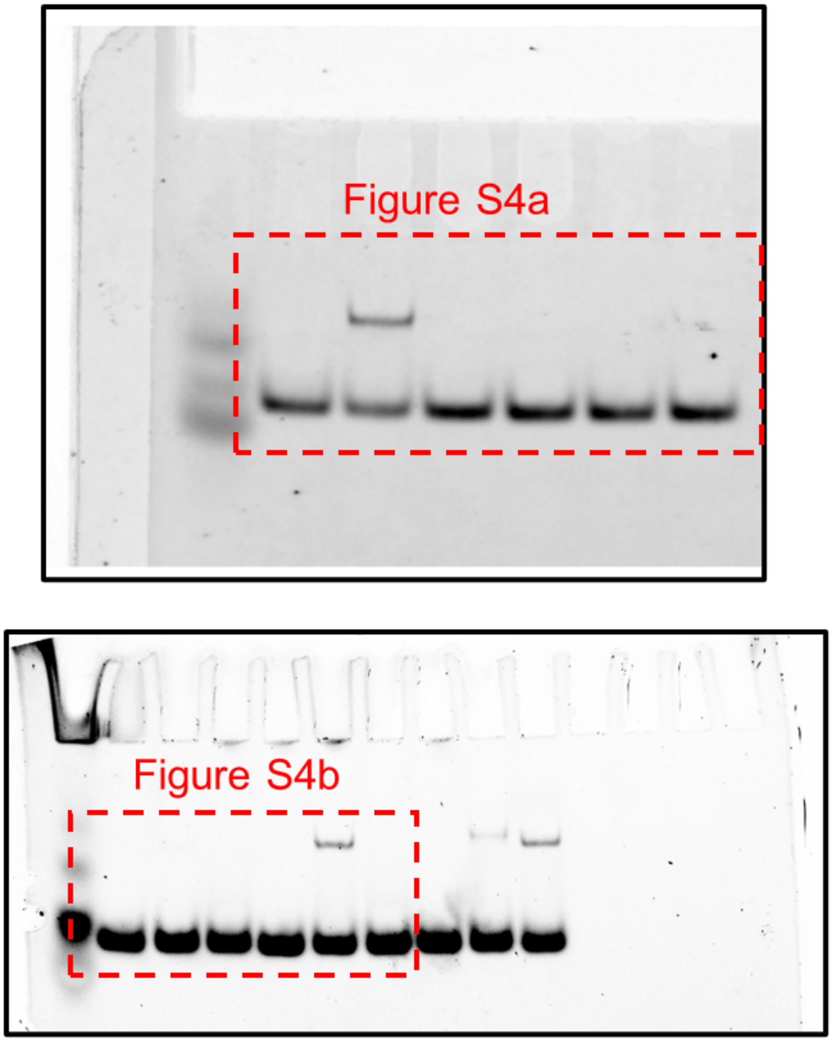


**Figure S17.** Uncropped gel image for Figure S4. Frame indicates the cropped region shown in Figure S4. Unboxed region contains unrelated experiments.

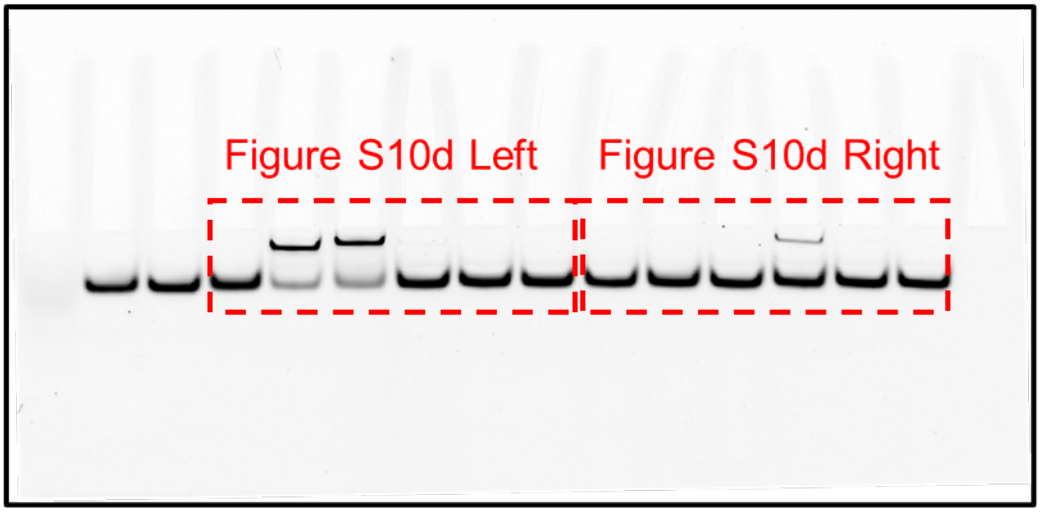


**Figure S18.** Uncropped gel image for Figure S10i. Frame indicates the cropped region shown in Figure S10i. Unboxed region contains control experiments.

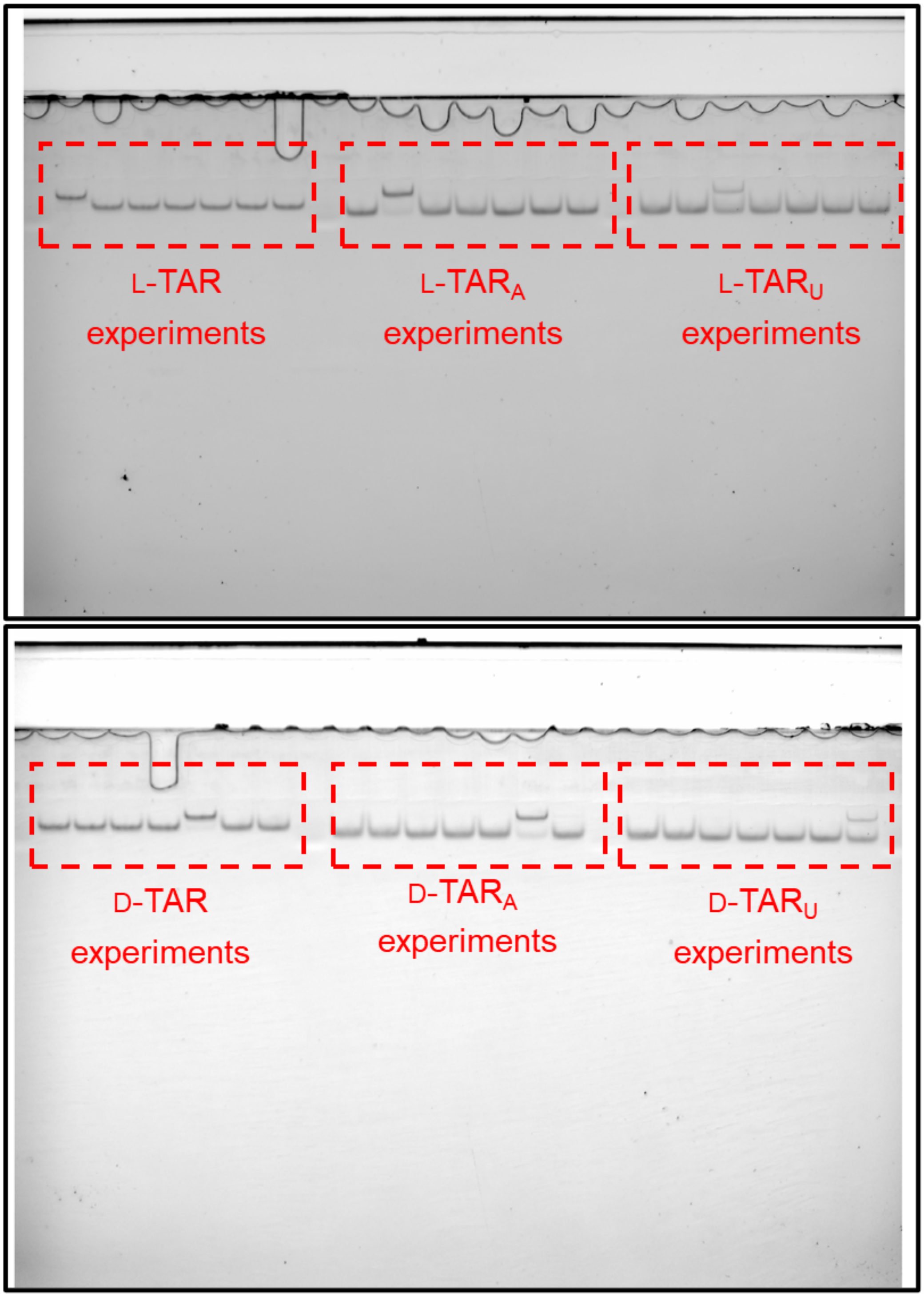


**Figure S19.** Uncropped gel image for Figure S12. Frame indicates the cropped region shown in Figure S12. Unboxed region contains unrelated experiments.

## SUPPLEMENTARY TABLES

**Table S1.** Names and sequences of oligonucleotides used in this work. L-RNA (blue) and D-RNA/DNA (black) are indicated by blue color. Mutagenized regions are indicated by brown text. /BioTEG/ = triethylene glycol biotin; /Bio/ = 5′ biotin; /sp18/ = hexaethylene glycol; /AmMC6/ = C6 amino modifier; /56-FAM/ = 5′ 6-fluorescein (phosphoramidite); /FAM/ = fluorescein 6-isomer (NHS ester); /Cy5/ = Sulfo-Cy5.

| Name | Sequence Identity (5'→3') | Type |
|---|---|---|
| D-dLib14 | GGTTACCAGCCTTCACTGC**TAGGGCGCGAGCAAACCCGTGCCGAAAGGCAAGAAACAGCA**GCACCACGGTCGGTCACAC | D-DNA |
| D-Lib14 | GGUUACCAGCCUUCACUGC**UAGGGCGCGAGCAAACCCGUGCCGAAAGGCAAGAAACAGCA**GCACCACGGUCGGUCACAC | D-RNA |
| FWD.primer | TTCTAATACGACTCACTATAGGTTACCAGCCTTCACTGC | D-DNA |
| REV.primer | GTGTGACCGACCGTGGTGC | D-DNA |
| EZSeq.FWD | ACACTCTTTCCCTACACGACGCTCTTCCGATCTAACGCCTGTTGTGAGCCTAACTGCGCTCGGGCAAGACTCAAATAGGTTACCAGCCTTCACTG | D-DNA |
| EZSeq.REV | GACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCAGCGTCACTAACGTCACGTCGCTCACGTACAATAGACAAGTCGTGTGACCGACCGTGGTGC | D-DNA |
| L-TAR$_A$-BIO | GUGAGCCUGAGAGCUCAC-/BioTEG/ | L-RNA |
| L-TAR$_U$-BIO | GUGAGCCUGUGAGCUCAC-/BioTEG/ | L-RNA |
| L-TAR$_A$ | /56-FAM/-/sp18/-GUGAGCCUGAGAGCUCAC | L-RNA |
| L-TAR$_U$ | /56-FAM/-/sp18/-GUGAGCCUGUGAGCUCAC | L-RNA |
| D-TAR-Cy5 | GCCAGAUUUGAGCCUGGGAGCUCUCUGGC-/AmMC6/-/Cy5/ | D-DNA |
| D-TAR$_A$-FAM | GUGAGCCUGAGAGCUCAC-/AmMC6/-/FAM/ | D-RNA |
| D-TAR$_U$-FAM | GUGAGCCUGUGAGCUCAC-/AmMC6/-/FAM/ | D-RNA |
| L-TAR-FAM | GCCAGAUUUGAGCCUGGGAGCUCUCUGGC-/AmMC6/-/FAM/ | L-RNA |
| D-TAR-FAM | GCCAGAUUUGAGCCUGGGAGCUCUCUGGC-/AmMC6/-/FAM/ | D-RNA |
| D-6-4t | GGACUAGGGCGCGAGCAAACCCGUGCCGAAAGGCAAGAAACAGUCC | D-RNA |
| L-6-4t | /Bio/−/sp18/−GGACUAGGGCGCGAGCAAACCCGUGCCGAAAGGCAAGAAACAGUCC | L-RNA |
| L-5$_A$-16t | /Bio/−/sp18/−GGAUGCUAGGGCGAGACAUUACCUGUGCCGAUAGGCAAAAGGCACCAGCAUCC | L-RNA |
| L-8$_U$-1t | /Bio/−/sp18/−GGACCUUCACUGCAAGGGCGCGAGUAAGCGAGUGAGGAGGGAAAAAAACAGCAGCACCACGGUCGGUCC | L-RNA |

| | | |
|---|---|---|
| D-$5_A$-1 | GGUUACCAGCCUUCACUGCCAGGGAGAGACAUCACUCGUGCCGAAAGGCAAGGAGCACCGGCACCACGGUCGGUCACAC | D-RNA |
| D-$5_A$-2 | GGUUACCAGCCUUCACUGCUAGUGGGCGAGACAUCACGUGCCGAAAGGCACGCAACACCAGCACCACGGUCGGUCACAC | D-RNA |
| D-$5_A$-3 | GGUUACCAGCCUUCACUGCUACGGGGCGAGACAUCACUUGCCGGGAGGCAAGCAACCGCGGCACCACGGUCGGUCACAC | D-RNA |
| D-$5_A$-4 | GGUUACCAGCCUUCACUGCUAGGGCGAGACAUUACCGGCCCGAACGGCAACUAGCAGCAGCACCACGGUCGGUCACAC | D-RNA |
| D-$5_A$-5 | GGUUACCAGCCUUCACUGCGUGAGCAGGAGCGAACCGGUUCCGAGAGACAUUACUCAGCAGCACCACGGUCGGUCACAC | D-RNA |
| D-$6_U$-1 | GGUUACCAGCCUUCACUGCUAUAGAGCGAGAAAACGCGUGGCCUAAGGAAAGGACUAGCAGCACCACGGUCGGUCACAC | D-RNA |
| D-$6_U$-2 | GGUUACCAGCCUUCACUGCAAGGGCGCGAGUAAGCGAGUGAGGAGGGAAAGAAACAGCAGCACCACGGUCGGUCACAC | D-RNA |
| D-$6_U$-3 | GGUUACCAGCCUUCACUGCUAGGGGAGGUCCGAACCCGUCCCGAAAGGCAAGAGACGGCAGCACCACGGUCGGUCACAC | D-RNA |
| D-$6_U$-4 | GGUUACCAGCCUUCACUGCAAGGGCCCGAGCAAGGGCGUGACGGAGGGAAAUAAACAGCAGCACCACGGUCGGUCACAC | D-RNA |
| D-$6_U$-5 | GGUUACCAGCCUUCACUGCAAGGGGGCGAGCAAGCCCGUGAGGAGGGAAAAAAACAGCAGCACCACGGUCGGUCACAC | D-RNA |
| D-$5_A$-16 | GGUUACCAGCCUUCACUGCUAGGGCGAGACAUUACCUGUGCCGAUAGGCAAAAGGCACCAGCACCACGGUCGGUCACAC | D-RNA |
| D-$8_U$-1 | GGUUACCAGCCUUCACUGCAAGGGCGCGAGUAAGCGAGUGAGGAGGGAAAAAAACAGCAGCACCACGGUCGGUCACAC | D-RNA |
| D-$5_A$-16t1 | GGAGCGAGACAUUACCUGUGCCGAUAGGCAAAAGGCUCC | D-RNA |
| D-$5_A$-16t(2) | GGAUGCUAGGGCGAGACAUUACCUGUGCCGAUAGGCAAAAGGCACCAGCAUCC | D-RNA |
| D-$5_A$-16t3 | GGACCUUCACUGCUAGGGCGAGACAUUACCUGUGCCGAUAGGCAAAAGGCACCAGCACCACGGUCC | D-RNA |
| D-$8_U$-1t(1) | GGACCUUCACUGCAAGGGCGCGAGUAAGCGAGUGAGGAGGGAAAAAAACAGCAGCACCACGGUCGGUCC | D-RNA |
| D-$8_U$-1t2 | GGACCACUGCAAGGGCGCGAGUAAGCGAGUGGUCC | D-RNA |
| D-$8_U$-1t3 | GGUUACCAGCCUUCACUGCAAGGGCGCGAGUAAGCGAGUGAGGAGGCAGGAAACUCCCACGGUCGGUCACAC | D-RNA |
| D-hp$_{40}$ | GGCUACUUUUCUGCACUAUCGUGACUGCCUGUGAGCAGUC | D-RNA |
| D-hp$_{60}$ | GGCUACUUUUCUGCACUAUCGUCCCUGUUACGACAUAUCUCACUAGGCCUGAGAGCCUAG | D-RNA |
| D-hp$_{80}$ | GGCUACUUUUCUGCACAUUGAACUGAUCUCUGUUAAAUUUUUUGCUAAUCUACUAUGCAUCUGUGAGCCUGGGAGCUCAC | D-RNA |

**Table S2.** DNA oligonucleotides used in cross extension reactions to generate the dsDNA templates for *in vitro* transcription of the indicated D-RNA sequences in this work. The T7 promoter sequence is underlined. FWD.primer in Table S1 was used as forward sequence unless indicated otherwise

| RNA seq | Strand | Sequence Identity (5'→3') |
|---|---|---|
| D-5$_{A}$-1 | Reverse | GTGTGACCGACCGTGGTGCCGGTGCTCCTTGCCTTTCGGCACGAGTGATGTCTCTCCCTGGCAGTGAAGGCTGGTAACC |
| D-5$_{A}$-2 | Reverse | GTGTGACCGACCGTGGTGCTGGTGTTGCGTGCCTTTCGGCACGTGATGTCTCGCCCACTAGCAGTGAAGGCTGGTAACC |
| D-5$_{A}$-3 | Reverse | GTGTGACCGACCGTGGTGCCGCGGTTGCTTGCCTCCCGGCAAGTGATGTCTCGCCCCGTAGCAGTGAAGGCTGGTAACC |
| D-5$_{A}$-4 | Reverse | GTGTGACCGACCGTGGTGCTGCTGCTAGTTGCCGTTCGGGCCGGTAATGTCTCGCCCTAGCAGTGAAGGCTGGTAACC |
| D-5$_{A}$-5 | Reverse | GTGTGACCGACCGTGGTGCTGCTGAGTAATGTCTCTCGGAACCGGTTCGCTCCTGCTCACGCAGTGAAGGCTGGTAACC |
| D-6$_{U}$-1 | Reverse | GTGTGACCGACCGTGGTGCTGCTAGTCCTTTCCTTAGGCCACGCGTTTTCTCGCTCTATAGCAGTGAAGGCTGGTAACC |
| D-6$_{U}$-2 | Reverse | GTGTGACCGACCGTGGTGCTGCTGTTTCTTTCCCTCCTCACTCGCTTACTCGCGCCCTTGCAGTGAAGGCTGGTAACC |
| D-6$_{U}$-3 | Reverse | GTGTGACCGACCGTGGTGCTGCCGTCTCTTGCCTTTCGGGACGGGTTCGGACCTCCCCTAGCAGTGAAGGCTGGTAACC |
| D-6$_{U}$-4 | Reverse | GTGTGACCGACCGTGGTGCTGCTGTTTATTTCCCTCCGTCACGCCCTTGCTCGGGCCCTTGCAGTGAAGGCTGGTAACC |
| D-6$_{U}$-5 | Reverse | GTGTGACCGACCGTGGTGCTGCTGTTTTTTTCCCTCCTCACGGGCTTGCTCGCCCCCTTGCAGTGAAGGCTGGTAACC |
| D-5$_{A}$-16 | Reverse | GTGTGACCGACCGTGGTGCTGGTGCCTTTTGCCTATCGGCACAGGTAATGTCTCGCCCTAGCAGTGAAGGCTGGTAACC |
| D-8$_{U}$-1 | Reverse | GTGTGACCGACCGTGGTGCTGCTGTTTTTTTCCCTCCTCACTCGCTTACTCGCGCCCTTGCAGTGAAGGCTGGTAACC |
| D-5$_{A}$-16t1 | Forward | TTCTAATACGACTCACTATAGGAGCGAGACATTACC |
|  | Reverse | GGAGCCTTTTGCCTATCGGCACAGGTAATGTCTCGCTCC |
| D-5$_{A}$-16t2 | Forward | TTCTAATACGACTCACTATAGGATGCTAGGGCGAGACATTACC |
|  | Reverse | GGATGCTGGTGCCTTTTGCCTATCGGCACAGGTAATGTCTCGCCCTAGC |
| D-5$_{A}$-16t3 | Forward | TTCTAATACGACTCACTATAGGACCTTCACTGCTAGGGCGAGACATTACCTGTG |

| | | |
|---|---|---|
| | Reverse | GGACCGTGGTGCTGGTGCCTTTTGCCTATCGGCACAGGTAATGTCTCGCCCTAGC |
| D-8$_U$-1t1 | Forward | TTCTAATACGACTCACTATAGGACCTTCACTGCAAGGGCGCGAGTAAGCGAGTG |
| | Reverse | GGACCGACCGTGGTGCTGCTGTTTTTTTCCCTCCTCACTCGCTTACTCGCGCCCTTGC |
| D-8$_U$-1t2 | Forward | TTCTAATACGACTCACTATAGGACCACTGCAAGGGCG |
| | Reverse | GGACCACTCGCTTACTCGCGCCCTTGCAGTGGTCC |
| D-8$_U$-1t3 | Forward | TTCTAATACGACTCACTATAGGACCACTGCAAGGGCG |
| | Reverse | GGACCACTCGCTTACTCGCGCCCTTGCAGTGGTCC |
| D-hp$_{40}$ | Forward | TTCTAATACGACTCACTATAGGCTACTTTTCTGCAC |
| | Reverse | GACTGCTCACAGGCAGTCACGATAGTGCAGAAAAGTAGCC |
| D-hp$_{60}$ | Forward | TTCTAATACGACTCACTATAGGCTACTTTTCTGCAC |
| | Reverse | CTAGGCTCTCAGGCCTAGTGAGATATGTCGTAACAGGGACGATAGTGCAGAAAAGTAGCC |
| D-hp$_{80}$ | Forward | TTCTAATACGACTCACTATAGGCTACTTTTCTGCAC |
| | Reverse | GTGAGCTCCCAGGCTCACAGATGCATAGTAGATTAGCAAAAAATTTAACAGAGATCAGTTCAATGTGCAGAAAAGTAGCC |

## S4. References

1. A. M. Kabza and J. T. Sczepanski, An l-RNA Aptamer with Expanded Chemical Functionality that Inhibits MicroRNA Biogenesis, *ChemBioChem*, 2017, **18**, 1824-1827.
2. X. Han and J. T. Sczepanski, An expanded substrate scope for cross-chiral ligation enables efficient synthesis of long l-RNAs, *RSC Chemical Biology*, 2025, **6**, 209-217.
3. J. T. Sczepanski and G. F. Joyce, Binding of a Structured d-RNA Molecule by an l-RNA Aptamer, *J. Am. Chem. Soc.*, 2013, **135**, 13290-13293.
4. S. Dey and J. T. Sczepanski, In vitro selection of l-DNA aptamers that bind a structured d-RNA molecule, *Nucleic Acids Res.*, 2020, **48**, 1669-1680.
5. J. Zhang, K. Kobert, T. Flouri and A. Stamatakis, PEAR: a fast and accurate Illumina Paired-End reAd mergeR, *Bioinformatics*, 2013, **30**, 614-620.
6. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads, *2011*, 2011, **17**, 3.
7. T. G. Community, The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update, *Nucleic Acids Res.*, 2024, **52**, W83-W94.
8. S. T. Kramer, P. R. Gruenke, K. K. Alam, D. Xu and D. H. Burke, FASTAptameR 2.0: A web tool for combinatorial sequence selections, *Molecular Therapy - Nucleic Acids*, 2022, **29**, 862-870.