Electronic Supplementary Information (ESI)

A library screening approach identifies naturally occurring RNA sequences for a G-quadruplex binding ligand

Gayan Mirihana Arachchilage,[a] Mark J. Morris[b] and Soumitra Basu*[a]

[a] Department of Chemistry and Biochemistry, Kent State University, Kent, Ohio 44242 USA
[b] Base Pair Biotechnologies, Inc., 8058 El Rio St. Houston, TX 77054 USA

1. Materials and Methods

General methods

All the solutions for buffers were prepared with DEPC-treated nanopure water. All DNA oligonucleotides were purchased from Integrated DNA Technologies (IDT, Coralville, IA) and were purified by denaturing 17 % polyacrylamide gel electrophoresis (PAGE) prior to use.

Preparation of RNA Sequences

RNA sequences were synthesized by *in vitro* transcription using T7 polymerase enzyme. All RNAs were purified by 17 % denaturing PAGE. The RNA was extracted via the crush and soak method by incubating the gel slice at 4 °C in a solution of 300 mM NaCl, 10 mM Tris-HCl (pH 7.5) and 0.1 mM EDTA. Samples were concentrated with 2-butanol, precipitated with 100% ethanol and the salts were removed by 2X 70% ethanol wash. The RNA pellet was dissolved in 10 mM Tris-HCl (pH 7.5) and 0.1 mM EDTA. Concentrations of RNAs were determined based on their UV absorbance values at 260 nm by using NanoDrop® ND-1000 spectrophotometer.

Radiolabeling of RNA

RNA was treated with Calf-intestinal alkalinephosphatase (New England Biolabs) and 5′-end was radiolabeled by treating with T4 polynucleotide kinase (New England Biolabs), [γ-^{32}P] ATP (Pelkin Elmer) and incubated for 45 min at 37 °C. The radiolabeled full-length RNA was isolated by 17% denaturing PAGE. The RNA was extracted from the gel via the crush and soak method and precipitated as described above.

Attachment of kanamycin A to N-hydroxyl succinimide-activated sepharose

The immobilization of kanamycin was performed using previously published procedures.[1, 2] Briefly, a suspension (2.4 mL) of commercially available resin in isopropanol (1.2 mL of resin, loading 18 mM) was placed in a Bio Rad chromatography column. The column was washed with water (2x10 mL). 50 mg of Kanamycin A was dissolved in Na_2CO_3 buffer (5 mL, pH 8.5) and added to the washed resin and incubated overnight with shaking in a 4 °C cold room. After

incubation, the resin was washed with Na_2CO_3 buffer (3x10 mL), and the unreacted succinimide esters were quenched with an ethanolamine solution (5 mL, 10 mM) in 5% Na_2CO_3 (1 h at room temperature with continuous shaking). The resin was washed with water and then washed first with ammonium acetate (10 mL, 10 mM, pH 4.5) followed by Tris-HCl (10 mL, 10 mM, pH 7.5). The resin was stored in Tris-HCl (10 mM, pH 7.5) at 4°C for future use. As is true for other aminoglycosides Kanamycin A contains several amines, however, it is known that the 6'-amino is more reactive.[3-5] To preferentially attach Kanamycin A to the resin via the more reactive 6'-amino group a huge excess of Kanamycin A to succinimide esters loaded on the resin was used. The resin-Kanamycin A mixture (100 µL) was placed in the Bio Rad column and washed with the binding buffer (3X, 100 mM KCl, 10 mM Tris-HCl, pH 7.5) and the Kanamycin A derivatized resin was used directly for the selection assay.

Selection of RNA G-quadruplexes that recognize kanamycin A:

Radiolabeled RNA was folded to form RNA G-quadruplexes as described above. Folded RNA (~ 50 pmoles) containing BSA (200 µg) was added to the column in binding buffer 100 mM KCl, 10 mM Tris-HCl (pH 7.5), and 0.1 mM EDTA. The resin was then washed with 2.5 mL portions of binding buffer until it reached background radioactivity. The radioactivity was being eluted from the column was measured using liquid scintillation counting. Bound RNAs were eluted by washing the column with elution buffer containing 1 mM kanamycin A, 100 mM KCl, 10 mM Tris-HCl (pH 7.5) and 0.1 mM EDTA. Radioactivity of the elute was measured which showed that only a small percentage of folded RNA is bound to the immobilized Kanamycin A. Elute was concentrated to 500 µL with 2-butanol and precipitated as described above with glycogen (100 µg) as a carrier. The sample was resuspended in 10 mM Tris-HCl (pH 7.5) and 0.1 mM EDTA and treated with DNase I (RNase-free, Promega; 1 µl) by incubation at 37 °C for 4 hrs. Stop buffer (20 mM EGTA, pH 8.0) was added to the reaction and incubated at 65 °C for 10 min to inactivate the DNase I. Then the sample was used as a template in RT-PCR reactions.

RT-PCR reactions

In order to conform that there was no contamination in the samples, negative controls without template and template without AMV reverse transcriptase (New England Biolabs) were used. Primers used for RT-PCR experiments were: PCR primer, 5'- TAATACGACTCACTATAG

GGAGACAAGAATAAACGCTCAA, containing a T7 promoter; RT primer, 5'- GCCTGT TGTGAGCCTCCTGTCGAA. RT primer was allowed to anneal with DNase treated RNA (22.5 µL) by incubating with RT primer (1 µL, 100 µM) at 95°C for 3 min followed by incubation on ice for 3 min. Then dNTPs (1.5 µL, 10 mM), 10x RT buffer (3 µL, supplied by the manufacturer), and reverse transcriptase or H_2O for no RT controls (2 µL) were added and incubated at 37 °C for 20 min and quenched the reaction by heating at 95 °C for 10 min. PCR amplification was completed by adding PCR primer (2 µM), RT primer (2 µM), dNTPs (1mM each), $MgCl_2$ (7.5 mM), Taq polymerase (25 units), H_2O , and 10x PCR buffer . Thirty cycles of PCR were completed at 95 °C for 30 sec, 42 °C for 30 sec and 72 °C for 30 sec. PCR products were analyzed by running an aliquot on a denaturing 8% polyacrylamide gel stained with ethidium bromide. Only experiments in which the negative controls didn't show products were used for cloning. RT-PCR products were then cloned into a pUC18 vector. The plasmids were transformed into JM109 competent E. coli and plated on Luria Bertani (LB) plates containing ampicillin, X-gal, and IPTG. White colonies were grown overnight at 37 °C. Plasmids were isolated from bacteria using a PureYield™ Plasmid Miniprep System. Sequencing reactions were completed by the Ohio State University Plant microbe Genomics Facility.

NCBI Blast

G-quadruplex forming region of selected RNA was subjected to standard nucleotide BLAST against 'Human genomic + transcript database' in NCBI. Search results are shown below (with the accession numbers);

For S3 RNA

- Homo sapiens acyl-CoA synthetase long-chain family member 6 (ACSL6), transcript variant 1, 2 and 3, mRNAs (Accession: NM_015256.3, NM_001009185.2 and NM_001205247.1)
- Homo sapiens potassium voltage-gated channel, subfamily H (eag-related), member 2 (KCNH2), transcript variant 1 and 2, mRNAs (Accession: NM_000238.3 and NM_172056.2)
- Homo sapiens SMAD specific E3 ubiquitin protein ligase 1 (SMURF1), transcript variant 1, 2 and 3 mRNAs (Accession: NM_020429.2, NM_181349.2 and NM_001199847.1)
- Homo sapiens glucose-6-phosphate isomerase (GPI), transcript variant 2, mRNA (Accession: NM_000175.3)

- Homo sapiens neogenin 1 (NEO1), transcript variant 1, 2 and 3, mRNAs (Accession: NM_002499.3, NM_001172623.1 and NM_001172624.1)

- Homo sapiens elongation factor RNA polymerase II (ELL), mRNA (Accession: NM_006532.3)
- Homo sapiens family with sequence similarity 69, member B (FAM69B), mRNA (Accession: NM_152421.3)

- Homo sapiens corticotropin releasing hormone receptor 1 (CRHR1), transcript variant 1,2,3 and 4, mRNA (Accession: NM_001145146.1, NM_004382.4, NM_001145147.1 and NM_001145148.1)

- Homo sapiens zinc finger protein 385B (ZNF385B), transcript variant 1, mRNA (Accession: NM_152520.4)

For S6 RNA

- Homo sapiens potassium voltage-gated channel, subfamily H (eag-related), member 2 (KCNH2), transcript variant 1and 3, mRNAs (Accession: NM_000238.3 and NM_172057.2)

- Homo sapiens zinc finger protein 669 (ZNF669), transcript variant 1, mRNA (Accession: NM_024804.2)

- Homo sapiens Na+/K+ transporting ATPase interacting 3 (NKAIN3), mRNA (Accession: NM_173688.2)

For S7 RNA

- Homo sapiens aminopeptidase-like 1 (NPEPL1), transcript variant 1, mRNA (Accession: NM_024663.3)

Circular Dichroism (CD) Studies

RNA was folded to form G-quadruplex by heating the samples in 100 mM KCl, 10 mM Tris-HCl (pH 7.5), and 0.1 mM EDTA at 95 °C for 5 min followed by cooling to room temperature over a 90 min period. The circular dichroism (CD) spectra were recorded using a Jasco J-810 spectropolarimeter with a 0.1 cm path-length cell at a scan speed of 50 nm/min and a response

time of 1 s. The spectra were averaged over three scans. A buffer baseline was collected in the same cuvette and subtracted from the average scan for each experiment.

CD Melting Experiments

CD-melt spectra were recorded using a 0.1 cm path-length cell. Samples were folded as described above in the presence 10 mM Tris-HCl, 0.1 mM EDTA (pH7.5) and 10 mM or 100 mM KCl. Folded RNA (5 μM) was melted by gradually heating to 97 °C at a gradient of 15 °C/hr. Mineral oil was placed on top of the sample to prevent the evaporation. The melting curves were obtained by monitoring the change of CD intensity at 263 nm. Thermodynamic parameters and T_m values were calculated according to van't Hoff method by using Origin 8.[6, 7]

G4-Flurescence intercalator displacement assay

RNA was prefolded as described above in the presence of 100 mM KCl, 10 mM Tris-HCl (pH 7.5), and 0.1 mM EDTA. Folded RNA (0.25 μM) was incubated with 0.50 μM Thiazole Orange for 30 min prior to each titration. All fluorescence titrations were done in 100 mM KCl, 10 mM Tris-HCl (pH 7.5), and 0.1 mM EDTA. Kanamycin A concentration was increased from 0-2 mM by adding stock solutions to RNA-TO mixture and incubated for 4 min before taking the spectrum. Fluorescence spectra were obtained using Varian Cary Eclipse fluorescence spectrophotometer upon excitation at 501 nm and data was collected from 510 nm to 650 nm. The $^{G4}DC_{50}$ values were obtained by fitting the experimental data using Origin 8. The $^{G4}DC_{50}$ represents the amount of kanamycin A required to displace 50% of bound TO from G-quadruplex.

Footprinting by RNase T1

The 5'-end radiolabeled RNA was folded as described above in the presence or absence of 100 mM KCl. Folded RNA was digested with 0.02 U of RNase T1 (Ambion) for 2 min at 37 °C. The reactions were terminated by using an equal volume of 2X urea loading buffer containing 7 M urea followed by heating to 95 °C for 2 min. Treated RNA was electrophoresed on 10% denaturing gel, dried on Whatman paper, and exposed to a phosphorimager screen overnight.

The gel images were visualized by scanning the screen on a Phosphorimager Typhoon FLA 9500 (GE).

Complete sequence of DNA template used to synthesis the library

5'GCCTGTTGTGAGCCTCCTGTCGAACCCNNCCCNNCCCNNCCCTTGAGCGTTTATTCTTGT CTCCCTATAGTGAGTCGTATTA 3'

Complete sequence of mutant RNA

5'GGGAGACAAGAAUAAACGCUCAAGCACCACUCUGAGACAUGUUCGACAGGAGGCUCAC AACAGGC 3'

2. Supplementary Figures and Tables

Table S1. Complete sequences of selected RNAs. Flanking regions are written in blue color.

Name	Sequence (5' to 3')
S 1	GGGAGACAAGAAUAAACGCUCAAGGGACGGGCCGGGCCGGGUUCGA CAGGAGGCUCACAACAGGC
S 2	GGGAGACAAGAAUAAACGCUCAAGGGCCGGGCAGGGCAGGGUUCGA CAGGAGGCUCACAACAGGC
S 3	GGGAGACAAGAAUAAACGCUCAAGGGCCGGGCCGGGCCGGGUUCGA CAGGAGGCUCACAACAGGC
S 4	GGGAGACAAGAAUAAACGCUCAAGGGCAGGGCCGGGACGGGUUCGA CAGGAGGCUCACAACAGGC
S 5	GGGAGACAAGAAUAAACGCUCAAGGGCCGGGUAGGGACGGGUUCGA CAGGAGGCUCACAACAGGC
S 6	GGGAGACAAGAAUAAACGCUCAAGGGCCGGGCCGGGCGGGGUUCGA CAGGAGGCUCACAACAGGC
S 7	GGGAGACAAGAAUAAACGCUCAAGGGCCGGGCAGGGCCGGGUUCGA CAGGAGGCUCACAACAGGC

Table S2: T_m Values and thermodynamic parameters for the folding of S7 RNA. Experiments were performed in the presence of 10 mM KCl.

T_m (°C)	ΔH (kJ/mol)	ΔS (kJ mol^{-1} K^{-1})	$\Delta G°_{37°C}$ (kJ/mol)
62.6 ± 0.4	-81.10 ± 5.91	-0.24± 0.02	-6.20 ± 0.51

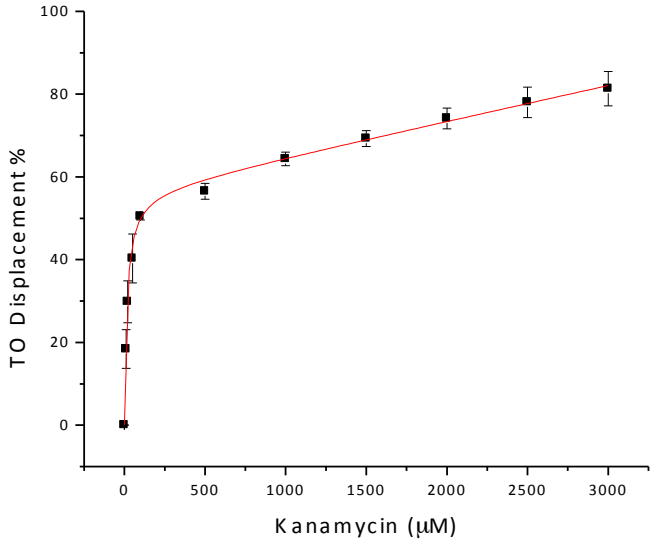

Fig S1: (A) FID titration plot of MT3-MMP mRNA with kanamycin A in 0.1 mM EDTA and 10 mM Tris-HCl (pH 7.5) in the presence of 100 mM KCl. Prefolded RNA (0.25 μM) was mixed with thiazole orange (0.50 μM) in a 1:2 ratio. The mixture was incubated for 30 min and titrated with kanamycin A until the fluorescence level went down to near to the baseline. Thiazole orange excitation used was at 501 nm and the emission spectrum was recorded from 510 nm to 650 nm. $^{G4}DC_{50}$ was calculated as the concentration of kanamycin A needed to displace 50% of bound TO. $^{G4}DC_{50}$ was calculated to be 139.32 ± 25.59 μM

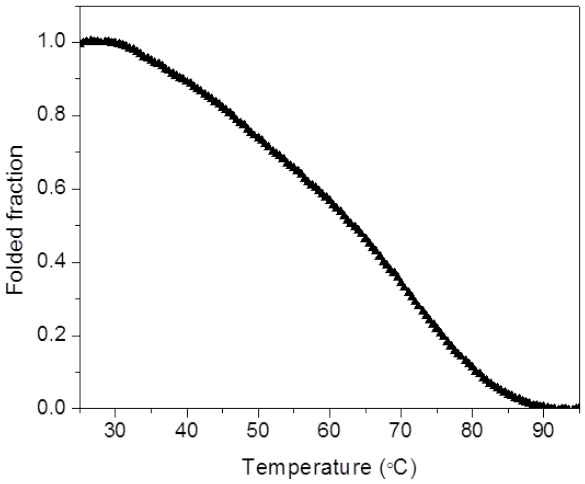

Fig S2: Circular dichroism melting curve of S7 RNA at a strand concentration of 5 μM in 0.1mM EDTA and 10mM Tris-HCl (pH 7.5) in the presence of 10 mM KCl.

3. References

1. M. D. Disney and J. L. Childs-Disney, *ChemBioChem*, 2007, **8**, 649-656.
2. A. Krishna Mallia , P. K. Smith and G. T. Hermanson, Immobilized Affinity Ligand Techniques, *Academic Press*, 1992.
3. M. D. Disney, S. Magnet, J. S. Blanchard and P. H. Seeberger, *Angew. Chem.-Int. Edit.*, 2004, **43**, 1591-1594.
4. Y. Wang and R. R. Rando, *Chemistry & Biology*, 1995, **2**, 281-290.
5. Y. Wang, K. Hamasaki and R. R. Rando, *Biochemistry*, 1997, **36**, 768-779.
6. L. A. Marky and K. J. Breslauer, *Biopolymers*, 1987, **26**, 1601-1620.
7. J. L. Mergny and L. Lacroix, *Oligonucleotides*, 2003, **13**, 515-537.