Supplemental Information

RE-SELEX: Restriction Enzyme-Based Evolution of Structure-Switching Aptamer Biosensors

Aimee A Sanford¹, Alexandra E Rangel², Trevor A Feagin², Robert G Lowery³, Hector S Argueta-Gonzalez¹, and Jennifer M Heemstra¹*

¹ Department of Chemistry, Emory University, Atlanta, Georgia 30322, USA

² Department of Chemistry and the Center for Cell and Genome Science, University of Utah, Salt Lake City, Utah 84112, USA

³ BellBrook Labs, LLC, Madison, Wisconsin 53711, USA

* To whom correspondence should be addressed. Tel: +1 404 727 7766; Email: jen.heemstra@emory.edu

MATERIALS AND METHODS

Design of restriction enzyme libraries

EcoRI-HF Library or BamHI-HF library was hybridized using slow cooling from 95 °C to 25 °C over 30 minutes in a thermal cycler to varying lengths of capture sequences (8-14 nt) in a 1:1 ratio in 1X Cut Smart Buffer (50 mM potassium acetate, 20 mM tris-acetate, 10 mM magnesium acetate, 100 µg/mL bovine serum albumin, pH 7.9, New England Biolabs) to a final concentration of 1 μ M library and 1 μ M capture sequence. The hybridized complex was then divided into 10 μL portions and incubated with either 1.5 U of EcoRI-HF or BamHI-HF (New England Biolabs). The solutions were incubated for 2 h at 37 °C, followed by 20 min at 65 °C. Samples were analyzed by denaturing 10 % PAGE to monitor digestion of the biosensor. The gels were imaged using a GE Amersham Typhoon RGB scanner and a 488 nm excitation laser and the Cy2 525BP20 emission filter. Digestion efficiency was determined by the percent cleaved (fluorescence band volume for the cleaved product/total lane volume) using ImageJ.

Bead assisted restriction enzyme SELEX

For the bead assisted restriction enzyme SELEX, capture strand provided directly on the oligo-affinity polymeric support (26-4001-01, Glen Research) as PS Bead-spacer18-spacer18-photocleavable linker (10-4913, Glen Research) -3'- AGTCTTAAGTAA-5' from the University of Utah DNA/Peptide Synthesis Core Facility. 500 μL of 1 mg/mL of beads were washed 2X by vortexing and centrifugation (16 xg, 2 minutes) with 500 μL of the selection buffer (50 % 1X Cut Smart Buffer (New England Biolabs) 50 mM potassium acetate, 20 mM tris-acetate, 10 mM magnesium acetate, 100 µg/mL bovine serum albumin (pH 7.9); 1X artificial cerebrospinal fluid buffer, 0.05 % Tween-20). The beads were resuspended in with 1 nmol of library in a total of 500 μL selection buffer. After round 1, 100 picomoles were used with 100 μL beads in a total of 100 μL selection buffer. The library was hybridized to the beads for 1 h at 25 °C. The beads were then washed 3X with 500 μL selection buffer as described above. Biosensor was collected following UV irradiation (365 nm) for 45 minutes and resuspended in 10 mM target in selection buffer and incubated for 1 hr at 25 °C.

Bulk biosensor test

Round 10 and 11 Library and 3'-BHQ-1 (12 nt) and capture strand were diluted in selection buffer (50 % 100 mM sodium chloride, 20 mM tris-hydrochloride, 2 mM magnesium chloride, 5 mM potassium chloride, and 1 mM calcium chloride (pH 7.6); 50 % 1X Cut Smart Buffer (New England Biolabs) 50 mM potassium acetate, 20 mM tris-acetate, 10 mM magnesium acetate, 100 µg/mL bovine serum albumin (pH 7.9)). Biosensors were constructed by combining library (1 μM) with capture strand (2 μM) in selection buffer and heating to 95 °C and slow-cooling to 25 °C over 30 mins in a thermal cycler. Biosensor was then added to a Corning Costar 384-well black plate. Buffer or kanamycin A (1 mM) was then added and re-incubated for 1 h at 25 °C, covered in foil. Samples were normalized to a well with library in selection buffer alone. Percent displacement was calculated using equation 1.

Kanamycin immobilization on beads and binding test

PierceTM NHS-activated magnetic beads (300 μL) were washed with 1 mL of ice-cold 1 mM hydrochloric acid. After discarding the supernatant, 300 μL of 10 mg/mL kanamycin A was added to the beads and incubated with gentle shaking for 18 h at 25 °C. The beads were then collected using a magnetic stand and washed twice with 1 mL 50 mM borate buffer, pH 8.5 followed by ultrapure water once. Quenching buffer (1 mL) was then added and incubated for 2 h at 25 °C. The beads were then collected, washed twice with ultrapure water and stored in 300 μL selection buffer. For bead binding, 20 µL beads were diluted in 100 µL selection buffer containing 100-200 pmol of each aptamer candidate (n=1). The samples were incubated at 25 °C for 1 h, then washed twice using selection buffer before eluting twice for 5 min at 95 °C. Absorbance was measured using a BioTek Take3 plate set to ssDNA mode. Percent bound was calculated from quantifying elution compared to input. We included #3-19 to benchmark positive hits.¹ We also included a random ssDNA as a negative control to set a baseline for non-specific binding.

Next Generation Sequencing of re-ligated vs. library samples

Library and re-ligated samples were amplified in 50 μL PCR reactions containing 0.2 μM template, 0.5 μM primers (forward primer, 5'-CGCATACCAGCTTAGTTCAGAATTCATT-3'; reverse primer, 5'-GCCGAGATTGCACTTACTATCT-3') and 1X Hot Start Master Mix (10 mM Tris-HCl, 50 mM KCl, 1.5 mM MgCl₂, 0.2 mM dNTPs, 5 % Glycerol, 25 units/mL Hot Start Taq DNA polymerase, pH 8.3). (New England Biolabs). The template was amplified with an initial

denaturation at 95 °C for 3 min, 25 cycles of (95 °C for 30 s, 55 °C for 30 s, and 72 °C for 20 s), and a final extension 72 °C for 2 min. The amplified double stranded DNA was purified using a MinElute PCR cleanup column (Qiagen). PCR products were quantified via Nanodrop and normalized to 20 ng/µL in nuclease free water. The samples were sent to Genewiz (South Plainfield, NJ) for Amplicon-EZ (150 – 500 bp) sequencing. Raw sequencing data were uploaded to the Emory Galaxy server. Bioinformatic analysis was carried out as described in Alam et. al.² Briefly, the reads were first trimmed to only contain the 90 nt library sequence and filtered by quality with a minimum score of 30 %. Sequences were then isolated that contained the forward primer and reverse primer binding sites through the barcode function. The FASTAptamer-Count and FAST-QC function were run to gather information about pool uniqueness and nucleotide distribution.

RESULTS

Supplementary Figure 1. DNA sequences used in restriction enzyme library optimization. FAM = fluorescein

Supplementary Figure 2. DNA sequences used for SELEX and biosensor characterization. FAM = fluorescein; Sp9 = spacer 9; BHQ1 = black hole quencher 1

Supplementary Figure 3. Maximizing digestion efficiency of restriction enzyme-based libraries. Restriction enzyme meditated digestion of 5'-fluorescein library (1 μ M) hybridized to capture strands of varying length (2 μ M) for EcoRI-HF and BamHI-HF. Digestion was carried out for 2 h at 37 °C by adding EcoRI-HF (100,000 U/mL) and BamHI-HF (20,000 U/mL) in a 10:1 ratio of complex to restriction enzyme. (a) The cleavage products were analyzed by 10 % denaturing PAGE. (b) Band intensity was used to quantify percent cleaved using ImageJ.

Supplementary Figure 4. EcoRI-HF cleavage in common SELEX buffers. EcoRI-HF meditated digestion of 5' fluorescein library (1 µM) hybridized to 12 nt capture strand (2 µM) with (a) varying % CutSmart in Kanamycin buffer (100 mM NaCl, 20 mM Tris-HCl, 2 mM MgCl₂, 5 mM KCl, 1 mM CaCl₂, pH 7.6) or with (b) common SELEX buffers, PBS (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na₂HPO₄, 1.47 mM KH₂PO₄, pH 7.4) and OTA buffer (10 mM Tris-HCl, 120 mM NaCl, 5 mM KCl, 1 mM CaCl₂, pH 8.5),³ with (+) or without (-) 2 mM C₄H₆MgO₄. The cleavage products were monitored by 10 % denaturing PAGE.

Supplementary Figure 5. Fidelity of PCR amplification for uncleaved products. (a) EcoRI-HF digestion results in three cleavage products of ranging size (90 nt, 70 nt, 20 nt). (b) PCR amplification of 90 nt library at 0 cycles (std) and 25 cycles (90 nt; 20 nt + 70 nt; no template control) The PCR products were monitored by 10 % denaturing PAGE. (c) Band intensity was used to quantify relative fluorescence units using ImageJ.

Supplementary Figure 6. Bead assisted EcoRI SELEX to generate structure-switching aptamers. (a) Progression of bead assisted EcoRI-HF SELEX to generate structure-switching aptamers. (b) Following digestion, cleavage products were monitored by 10 % denaturing PAGE. Band intensity was used to quantify percent uncleaved. (c) Potential structure-switching aptamer candidates and their corresponding enrichment values from round 4 to round 7. The EcoRI-HF restriction site is underlined. The N40 restriction site is underlined with a single mismatch (bold, red).

Supplementary Figure 7. N40 mismatch affords selection advantage. (a) Highly enriched sequences did not correctly hybridize the capture strand through the built in EcoRI-HF recognition site (underlined blue). A highly enriched sequence contained a partial recognition site in the N40 region (underlined black) with one mismatch (red, bold) as denoted by PBS + N40*. This sequence and a sequence without the built in EcoRI-HF recognition with the partial recognition site in the N40 region as denoted by N40* (1 μ M) was hybridized to 12 nt capture strand (2 μ M) and subsequently digested. The cleavage products were monitored by 10 % denaturing PAGE. Band intensity was used to quantify percent cleaved using ImageJ.

Supplementary Figure 8. Optimization of EcoRI-based SELEX steps. (a) Scheme overview of EcoRI-based SELEX. Steps b, c, and d were optimized to reduce background enrichment. (b) The initial N40 5'-fluorescein labelled ssDNA library was hybridized to capture strand in a 1:2 ratio in selection buffer. This complex was digested with EcoRI-HF. The 70 nt cleaved product was gel excised and purified following 10 % denaturing PAGE, 1X SYBR Gold staining, and visualization on a UV-transilluminator. (c) Full-length product was recovered through ligation by T4 DNA ligase and visualized using 10 % denaturing PAGE. Percent ligated was calculated by the band intensity using ImageJ and was used to calculate final library concentration in the biosensor. (d) 10 % denaturing PAGE analysis of EcoRI digestion of purified ligation from (b) after hybridization to free capture strand in a 1:2 ratio. Percent cleaved was calculated from band intensity using ImageJ.

 $\mathbf c$

Supplementary Figure 9. Investigation of potential biasfollowing ligation. (a) Distribution of nucleotides in the N40 calculated from the FASTAptamer-Count function.² Percent random was calculated from the abundance of unique sequences divided by the total counted. (b) Sequence content distribution at each base pair position in the N40 region for library (left) and re-ligated (right) samples calculated using the FAST-QC function.²

Supplementary Figure 10. 10 % denaturing PAGE analysis of SELEX rounds. Digestion of the 5'-fluorescein labelled libraries was carried out for 2 h at 37 °C. Cleavage products for buffer (-) and 10 mM kanamycin (+) were analyzed by 10 % denaturing PAGE. Percent cleaved was quantified using band intensity in ImageJ using library as a positive control lane. Rounds increase from left to right.

a

 $\mathsf b$

Supplementary Figure 11. Bulk biosensor activity of SELEX rounds 10 and 11. PCR amplified library (1 µM) from rounds 10 (n=1) and 11 (n=2) were hybridized to the 12 nt BHQ1 capture strand (2 µM). 1 mM kanamycin A or selection buffer was added to the biosensors and incubated for 1 h at 25 °C. Displacement was quantified by measuring the fluorescence intensity on a Cytation 5 multi-mode plate reader (BioTek) using excitation at 490 nm and emission at 520 nm (bandwidth 9, read height 6.5 mm). All samples were normalized to wells with aptamer alone. Percent displacement was calculated using equation 1 and plotted using GraphPad Prism.

Supplementary Figure 12. Bioinformatics analysis of SELEX rounds 9, 10, and 11. (a) Percent unique sequences over SELEX rounds calculated from the total sequences and number of unique sequences populated using the FASTAptamer-Count function.² The count files were then further analyzed using the FASTAptamer-Compare function to show scatter plots of read distribution from (b) replicate reads for round 11 (c) rounds 9 and 10 and d) rounds 10 and 11.

Free energy of secondary structure: -13.60 kcal/mol

Free energy of secondary structure: -14.20 kcal/mol

Free energy of secondary structure: -12.54 kcal/mol

 $K4-1$

Free energy of secondary structure: -16.48 kcal/mol

Free energy of secondary structure: -15.31 kcal/mol

 $K4-2$

Free energy of secondary structure: -23.05 kcal/mol

Free energy of secondary structure: -16.40 kcal/mol

Free energy of secondary structure: -13.66 kcal/mol

Free energy of secondary structure: -12.90 kcal/mol

Free energy of secondary structure: -15.31 kcal/mol

Supplementary Figure 13. **Secondary structure analysis of aptamer sequences using NUPACK.**⁴ Structures were predicted at 25 °C using standard conditions.

Supplementary Figure 14. Initial bead binding screen for top aptamer candidates. Candidates were incubated with 10 mg/mL kanamycin A magnetic beads for 1 h at 25 °C, followed by two washes. Bound sequences were eluted at 95 °C for 5 min, repeated once. The eluted ssDNA was quantified using absorbance at 260 nm using a Biotek Take3 plate on the plate reader. Percent bound was calculated from (eluted ssDNA/input) *100 %. Sequences that had higher % bound than the negative control (random 84 bp sequence) were carried forward for biosensor testing. (n=1)

Supplementary Figure 15. Increased ratio of 9 nt capture strand does not increase K16-1 stability. Displacement as a function of kanamycin A concentration using increasing K16-1:capture strand ratios using (a) 9 nt and (b)10 nt capture strands. Samples resulting in negative displacement were denoted 0 % displacement. Error bars denote standard error (n=3).

Supplementary Figure 16. Biosensor activity of #3-19. (a) Structure of #3-19, a top binding candidate from Capture SELEX hybridized to the 12 nt capture strand ^{1,5}. The docking site is in highlighted in red. Structure estimated using NUPACK at 25°C using standard conditions.⁴ (b) Biosensor concentration impacts signal quenching from hybridization of #3-19 to 12 nt capture strand. (c) Normalized RFU to show there was no displacement as a function of kanamycin A concentration using capture stand lengths. Error bars denote standard error (n=3).

Supplementary Figure 17. Binding isotherm for K16-1 and #3-19 determined by MST. Graphs are displayed as fraction bound versus kanamycin A concentration. Errors represent the standard error of three independent trials. The data was graphed using GraphPad Prism and fit according to the law of mass action to determine K_{D} .

Supplementary Figure 18. Binding isotherm for K16-1a, K16b, and K16c determined by MST. Graphs are displayed as fraction bound versus kanamycin A concentration. Errors represent the standard error of three independent trials. The data was graphed using GraphPad Prism and fit according to the law of mass action to determine $K_{D.}$

Supplementary Figure 19. Kanamycin A aptamersequences and their corresponding *K***Dvalues determined by MST.** Errors represent the standard error of three independent trials. The data was graphed using GraphPad Prism and fit according to the law of mass action to determine K_{D} .

Supplementary Figure 20. Optimization of K16-1c as a biosensor. (a) Biosensor concentration was optimized by monitoring concentration-dependent hybridization for K16-1c with 9 nt capture strand. (b) Capture strand displacement as a function of kanamycin A concentration using 100 nM K16-1c at a 1:1 ratio with 9 nt capture strand. (c) Displacement as a function of kanamycin A concentration using decreasing K16-1c:capture strand (9 nt) ratios at 500 nM K16-1c. Samples resulting in negative displacement were denoted 0 % displacement. Error bars denote standard error (n=3).

Supplementary Figure 21. Orthogonal method of determining K16-1 and K16-1c binding to kanamycin A. Normalized hybridization of increasing concentration of 9 nt capture strand to (a) 100 nM K16-1 and (b) 500 nM K16- 1c. Dose-dependent displacement of capture strand as result of kanamycin A concentration for (c) K16-1 and (d) K16-1c. These data was graphed using GraphPad Prism and fit according to the law of mass action to determine K_D and K_{sens}. Equation 3 was used to calculated target binding K_D for (e) K16-1 and (f) K16-1c.

Supplementary Figure 22. Normalized RFU of K16-1 with aminoglycosides. Normalized RFU to show there was a concentration dependent drop in fluorescence as a function of aminoglycoside concentration. Error bars denote standard error (n=3).

Supplementary Figure 23. Tabular data for K16-1 hybridization with 12 nt capture strand in percent quenched and raw fluorescence values. Errors represent the standard error of three independent trials.

Supplementary Figure 24. Tabular data for K16-1 dose-displacement with 9, 10, and 12 nt capture strand in percent displacement. Errors represent the standard error of three independent trials.

Supplementary Figure 25. Tabular data for K16-1 dose-displacement with 9, 10, and 12 nt capture strand in raw fluorescence values. Errors represent the standard error of three independent trials.

Supplementary Figure 26. Tabular data for K1-1 dose-displacement with 9, 10, and 12 nt capture strand in percent displacement. Errors represent the standard error of three independent trials.

Supplementary Figure 27. Tabular data for K1-1 dose-displacement with 9, 10, and 12 nt capture strand in raw fluorescence values. Errors represent the standard error of three independent trials.

Supplementary Figure 28. Tabular data for K1-2 dose-displacement with 9, 10, and 12 nt capture strand in percent displacement. Errors represent the standard error of three independent trials.

Supplementary Figure 29. Tabular data for K1-2 dose-displacement with 9, 10, and 12 nt capture strand in raw fluorescence values. Errors represent the standard error of three independent trials.

Supplementary Figure 30. Tabular data for K2-1 dose-displacement with 9, 10, and 12 nt capture strand in percent displacement. Errors represent the standard error of three independent trials.

Supplementary Figure 31. Tabular data for K2-1 dose-displacement with 9, 10, and 12 nt capture strand in raw fluorescence values. Errors represent the standard error of three independent trials.

Supplementary Figure 32. Tabular data for K16-1 dose-displacement with 9 nt capture strand in percent displacement and raw fluorescence values. Errors represent the standard error of nine independent trials.

Supplementary Figure 33. Tabular data for #3-19 hybridization with 12 nt capture strand in percent quenched and raw fluorescence values. Errors represent the standard error of three independent trials.

Supplementary Figure 34. Tabular data for #3-19 dose-displacement with 9, 10, and 12 nt capture strand in raw fluorescence values. Errors represent the standard error of three independent trials.

Supplementary Figure 35. Tabular data for K16-1 dose-displacement at 1:1, 1:2, and 1:3 ratio aptamer: 9 nt capture strand in percent displacement. Errors represent the standard error of three independent trials.

Supplementary Figure 36. Tabular data for K16-1 dose-displacement at 1:1, 1:2, and 1:3 ratio aptamer: 9 nt capture strand in raw fluorescence values. Errors represent the standard error of three independent trials.

Supplementary Figure 37. Tabular data for K16-1 dose-displacement at 1:1, 1:2, and 1:3 ratio aptamer: 10 nt capture strand in percent displacement. Errors represent the standard error of three independent trials.

Supplementary Figure 38. Tabular data for K16-1 dose-displacement at 1:1, 1:2, and 1:3 ratio aptamer: 10 nt capture strand in raw fluorescence values. Errors represent the standard error of three independent trials.

Supplementary Figure 38. Tabular data for K16-1c hybridization with 9 nt capture strand in raw fluorescence values and percent quenched. Errors represent the standard error of three independent trials.

Supplementary Figure 39. Tabular data for K16-1c (100 nM) dose-dependent displacement with 9 nt capture strand in percent displacement and raw fluorescence values. Errors represent the standard error of three independent trials.

Supplementary Figure 40. Tabular data for K16-1c (500 nM) dose-dependent displacement with 9 nt capture strand at 125 nM and 250 nM in percent displacement and raw fluorescence values. Errors represent the standard error of three independent trials.

Supplementary Figure 41. Tabular data for K16-1c dose-displacement with 9 nt capture strand in percent displacement and raw fluorescence values. Errors represent the standard error of nine independent trials.

Supplementary Figure 42. Tabular data for K16-1 hybridization (100 nM) with 9 nt capture strand in percent quenched and raw fluorescence values. Errors represent the standard error of three independent trials.

Supplementary Figure 43. Tabular data for K16-1c hybridization (500 nM) with 9 nt capture strand in percent quenched and raw fluorescence values. Errors represent the standard error of three independent trials.

Supplementary Figure 44. Tabular data for K16-1 raw fluorescence in the presence of aminoglycosides. Errors represent the standard error of three independent trials.

Supplementary Figure 45. Tabular data for K16-1 and K16-1c displacement with 9 nt capture strand in percent displacement for aminoglycosides at 1 mM in percent displacement and normalized fluorescence. Errors represent the standard error of three independent trials.

REFERENCES

- 1 R. Stoltenburg, N. Nikolaus and B. Strehlitz, *J. Anal. Methods Chem.*, 2012, **1**, 14.
- 2 K. K. Alam, J. L. Chang and D. H. Burke, *Mol. Ther. - Nucleic Acids*, 2015, **4**, e230.
- 3 T. A. Feagin, D. P. V. Olsen, Z. C. Headman and J. M. Heemstra, *J. Am. Chem. Soc.*, 2015, **137**, 4198–4206.
- 4 J. N. Zadeh, C. D. Steenberg, J. S. Bois, B. R. Wolfe, M. B. Pierce, A. R. Khan, R. M. Dirks and N. A. Pierce, *J. Comput. Chem.*, 2011, **32**, 170–173.
- 5 N. Nikolaus and B. Strehlitz, *Sensors (Switzerland)*, 2014, **14**, 3737–3755.