nature chemistry

SUPPLEMENTARY INFORMATION

DOI: 10.1038/NCHEM.2058

# Recognition and Sensing of Low-Epitope Targets via Ternary Complexes with Oligonucleotides and Synthetic Receptors

**Kyung-Ae Yang[1], Michaela Barbu[1], Marlin Halim[1], Payal Pallavi[1], Benjamin Kim[2], Dmitry Kolpashchikov[1], Stevan Pecic[1], Steven Taylor[1], Tilla S. Worgall[2]*, & Milan N. Stojanovic[1,3]***

## Table of contents:

## Experimental Section

**General.** All oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA, USA). Standard desalted oligonucleotides were used for the library and primers. Modified oligonucleotides (e.g. biotinylation, fluorophore conjugates) were purified by reverse-phase HPLC. Nuclease-free water was used for all purposes, e.g., oligonucleotide dissolution and buffer preparation. Anthracene-9,10-dicarbaldehyde was purchased from Matrix Scientific (Columbia, SC, USA), and 2-bromomethylphenylboronic acid pinacol ester was purchased from Combi-Blocks (San Diego, CA, USA). All other compounds were purchased from Sigma-Aldrich Co. (St. Louis, MO, USA) unless otherwise noted.

**Boronic acid synthesis.**

**1,1'-(anthracene-9,10-diyl)bis(N-methylmethanamine)** Anthracene-9,10-dicarbaldehyde (0.117 g, 0.50 mmol) was dissolved in methanol (15 mL) and cooled to 0°C, followed by slow addition of 2 M methylamine in methanol (0.75 mL, 1.50 mmol). The reaction mixture was stirred overnight at room temperature under an argon atmosphere. The resulting mixture was then cooled to 0°C, sodium borohydride (0.095 g, 2.50 mmol) was added in one portion, and the reaction mixture stirred for an additional 6 h at room temperature. The resulting mixture was poured onto 5 mL of ice-water and the aqueous layer extracted with chloroform (3 x 15 mL). The combined organic extracts were dried over $Na_2SO_4$, filtered, and concentrated under reduced pressure. The crude product was purified by silica gel chromatography eluting with 10% methanol in dichloromethane to afford the title compound as an orange powder (0.110 g, 83 %), mp 139-140°C. $^1$H NMR (400 MHz, $CDCl_3$): $\delta$ 8.39-8.38 (d, $J$ = 6.8 Hz, 4H), 7.55-7.53 (d, $J$ = 6.4 Hz, 4H), 4.68 (s, 4H), 2.67 (s, 6H), 1.51 (bs, 2H); $^{13}$C NMR (100 MHz, $CDCl_3$): $\delta$ 132.1, 130.3, 125.8, 125.0, 48.2, 37.3.

**((((anthracene-9,10-diylbis(methylene))bis(methylazanediyl))bis(methylene))bis(2,1-phenylene))diboronic acid:** Diamine from the previous step (0.100 g, 0.38 mmol) and 2-bromomethylphenylboronic acid pinacol ester (0.425 g, 1.5 mmol) were dissolved in 2.0 mL of DMF, followed by addition of N,N-diisopropylethylamine (0.20 mL, 1.13 mmol). The reaction mixture was stirred for 18 h at room temperature followed by 2 h at 40°C. The reaction mixture was diluted with 20 mL of water and then the aqueous layer extracted with chloroform (2 x 25 mL). The organic layers were combined, dried over $Na_2SO_4$, filtered, and concentrated under reduced pressure. Yellow oil was obtained and purified by silica gel column chromatography eluting with 5% methanol in dichloromethane. The title compound was obtained as pale yellow powder (0.110 g, 55% yield). $^1$H NMR (400 MHz, Methanol-$D_4$): $\delta$ 8.28-8.26 (m, 4H), 7.64-7.62 (d, $J$ = 7.2 Hz, 4H), 7.58-7.56 (dd, $J$ = 7.0 Hz, 3.2 Hz, 4H), 7.35-7.26 (m, 6H), 4.98 (s, 4H), 4.28 (s, 4H), 2.36 (s, 6H). $^{13}$C NMR (100 MHz, Methanol-$D_4$): $\delta$ 135.4, 132.1, 131.8, 128.9, 127.8, 125.9, 64.3, 50.1, 40.5. NMR spectra were recorded at 25°C, on a Varian 400 NMR spectrometer $^{13}$C (100 MHz) and $^1$H (400 MHz).

**Preparation of buffers and target solutions.** Three buffers were prepared; (1) SELEX buffer (20 mM HEPES, 1 M NaCl, 10 mM $MgCl_2$, 5 mM KCl, pH 7.5); (2) no $MgCl_2$ buffer for strand separation (20 mM HEPES, 300 mM NaCl, pH 7.5,), and (3) 2X SELEX buffer (40 mM HEPES, 2 M NaCl, 20 mM $MgCl_2$, 10 mM KCl, pH 7.5). The target complexes for SELEX are in two groups; (1) monosaccharides with the bisboronic acid and (2) amino acids with pentamethylcyclopentadienyl-rhodium(III) chloride dimer ($[Cp^*RhCl_2]_2$). For the three monosaccharide targets – appropriate amounts of freshly prepared 1 M stock solutions in water of glucose, fructose, or galactose were combined with the bisboronic-acid (2.5 mM stock solution in methanol) to give a final concentration of 50 μM bisboronic-acid and 40 mM glucose (or 100 mM for fructose and galactose) when diluted with SELEX buffer. The resulting mixtures were incubated for at least 40 minutes to establish equilibrium. For the amino acid-Cp*Rh (III) targets – 50 mM stock solutions of phenylalanine and tryptophan were prepared in nuclease-free water; for tyrosine, a 2 mM of stock solution was prepared directly in the SELEX buffer due to its low solubility. A 2.5 mM of $[Cp^*RhCl_2]_2$ stock solution was prepared in nuclease-free water and kept at room temperature. The final concentration of the target complex components in SELEX buffer was 100 μM of Cp*Rh(III) with 1 mM of amino acid.

***In vitro* selection process.** The procedure by Yang *et al.*[11] was largely followed. The oligonucleotides used for SELEX were as follows:

(1) a random ($N_{30}$) library (72-mer): 5'– GGAGGCTCTCGGGACGAC($N_{30}$)GTCGTCCCGATGCT-GCAATCGTAA-3',

(2) forward-primer (18-mer): 5'-GGAGGCTCTCGGGACGAC-3',

(3) reverse-primer (22-mer): 5'-TTACGATTGCAGCATCGGGACG-3',

(4) biotinylated reverse-primer (22-mer): 5'-biotin-TTACGATTGCAGCATCGGGACG-3',

(5) biotinylated column immobilizing capture (18-mer): 5'- GTCGTCCCGAGAGCCATA-BioTEG-3'.

For the 1st round of SELEX, an oligonucleotide mixture containing 0.6 nmoles of capture strand and 0.15 nmoles random library was prepared in 250 μL of SELEX buffer and incubated at 95°C for 5 min. The mixture was cooled down slowly to room temperature (to ~24°C, over at least 10 min), then added to a streptavidin agarose column. The streptavidin agarose column was prepared with 250 μL of the streptavidin agarose resin (1-3 mg biotinylated BSA/ml resin; Thermo scientific, IL, USA) in a micro bio-spin chromatography column (Bio-rad, CA, USA). To equilibrate the column, the streptavidin agarose was washed 5 times with the same volume of SELEX buffer, and then the oligonucleotide mixture passed through the column. The mixture was collected and passed through the column two more times. The column was then washed ten times with SELEX buffer. Eluent from each of these washes was collected into separate tubes. We then eluted the DNA-target complex with target solution three times, collecting each 250 μL fraction separately. The eluted samples were used as a template for PCR after concentrating to 50 μL by Amicon Ultra centrifugal filter (3K MWCO; Millipore, Cork, Ireland).

**PCR protocol:** 1 cycle of 95°C, 2 min, N cycle of [95°C, 15 sec; 60°C, 30 sec; 72°C, 45 sec], and 1 cycle

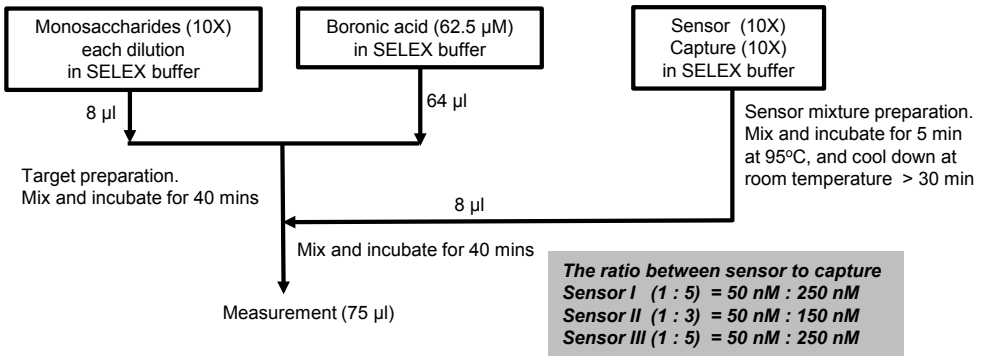
of 72°C, 2 min. The number of PCR cycles were between 9 and 13; as SELEX rounds increased less cycles were applied. The PCR amplicons were concentrated by centrifugal filtering (10K MWCO). Following concentrating, the double strand PCR amplicons were captured on a streptavidin agarose column via the biotinylated antisense strand. 250 µL of streptavidin agarose resin was used and washed three times with strand separation buffer (20 mM HEPES, 300 mM NaCl, pH 7.5). After loading the sample into the column, the resin was washed five times with the strand separation buffer, and incubated with 250 µL of 0.2 M NaOH. The flow-out drops were collected, neutralized by adding 0.2 M HCl and concentrated using a centrifugal filter (3K MWCO) up to 100 µL. The buffer pH and ionic condition was optimized by adding 100 µL of 2X reaction buffer to the concentrated flow-out to be used for the next round of SELEX.

**Cloning and sequencing.** During the selection step, the elution profile of the SELEX was monitored by PCR amplification of the flow-out collected during the washes and the target addition. After obtaining a clear elution profile (marked by a large increase in the PCR band intensity when the target was added) cloning was performed. The cloning and sequencing procedure by Yang *et al*[11] was followed.

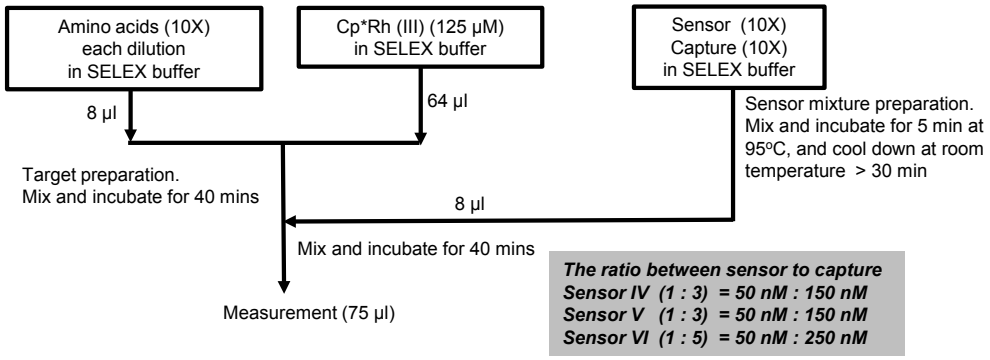
**Aptameric sensor binding assay.** About two sensor candidates from each SELEX were selected based on the redundancy of clone copy; the better responsive sensor between the candidates was selected for further analysis, one sensor from each SELEX. The composition of the sensor sequence and the optimized capture sequence is given in the 2D structure of sensors in **Fig. 2** and **Fig. 3**.

All measurements were performed in SELEX buffer with the concentrations indicated on the plots below. For complexation - the target and cofactor were mixed then incubated for 40 mins in the reaction buffer, and separately a mixture of sensor and capture strand was incubated at 95°C for 5 min then cooled down slowly to room temperature (at ~24°C, > 40 min). For monosaccharides, 8 µL of ten time concentrated monosaccharide stock for each testing concentration and 64 µL of 62.5 µM of bisboronic acid solution were mixed in SELEX buffer (50 µM bisboronic acid in final). After inducing the target complex, 8 µL of the ten time concentrated sensor-capture mixture (e.g. 500 nM sensor & 2500 nM capture mixture was added and incubated for 40 min at room temperature. The final concentrations of all sensors is 50 nM and that of capture is 250 nM for glucose-bisboronic acid, galactose-bisboronic acid and tryptophan-Cp*Rh(III). 150 nM of capture was applied for fructose-bisboronic acid, phenylalanine-Cp*Rh(III) sensor and tyrosine-Cp*Rh(III) sensor. The mixtures were transferred into a 384-well non-binding surface, flat bottom; black polystyrene assay plates (Corning, NY, USA). The fluorescence in each well was measured with a 485-nm excitation filter and a 535-nm emission filter on a Perkin-Elmer Victor II microplate reader (Shelton, CT, USA), with RFUs (relative fluorescence units) directly used as read out from the instrument. All measurements were done in at least triplicate. 2D plots were drawn on Prism 5.0 (GraphPad Software, CA, USA), respectively. The schematic drawing of the process is shown below.

**For monosaccharides**




**For amino acids**




**Serum mix-and-measure assays and self-experiments.** The phenylalanine oral administration method was from Rampini *et al* [22]. 20 μL of fresh capillary blood samples or serum was diluted in 780 μL of 0.9% NaCl solution (1/40X) for further use (1/200 or 1/400 dilution in the testing samples). The assay procedure was identical with the sensor binding assay as described above.

**Sensor affinity measurement.** The $K_d$ is determined following Hu and Easley's method[s1]. There were two separated binding affinity measurements for each sensor, $K_{d,\,eff1}$ and $K_{d,\,eff2}$. $K_{d,\,eff1}$ is the dissociation constants between sensor and capture oligonucleotide; it is calculated based on the decreasing fluorescent intensity (quenching) of sensor (50 nM fixed) in the presence of 0 ~ 500 nM range of the capture strand concentrations (capture strand contains DABCYL). The binding curve is plotted using 'one phase decay'. $K_{d,\,eff2}$ is a unitless constant between the sensor-capture complex and its target ($K_{d,\,eff2}$ = [capture strand][sensor-target] /[sensor-capture][target]). For the $K_{d,eff1}$ each binding curve is plotted using non-linear least-squares fitting using prism 5.0 (GraphPad Software). $K_d$ is derived from $K_{d,\,eff1}/K_{d,\,eff2}$, each value is shown in **Supplementary Figures 3** and **8**.

**Reference**

S1. Hu, J., Easley, C. J. A simple and rapid approach for measurement of dissociation constants of DNA aptamers against proteins and small molecules via automated microchip electrophoresis. *Analyst,* **136,** 3461-3468 (2011).

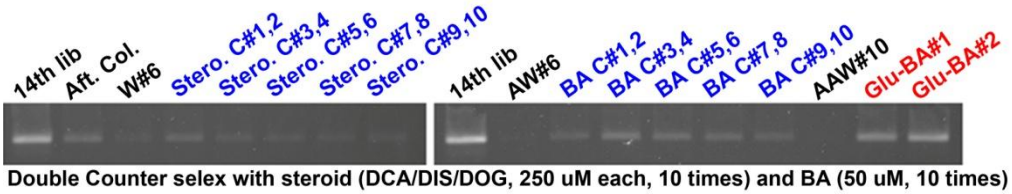
**Table S1. List of cloned sequences from six SELEX experiments.** From each SELEX experiment, about 20 clones were randomly chosen for sequencing and the redundancy of copy number was analyzed. The small letters represent the conserved sequence region of the library, and capital letters indicate sequences from the random region.

| Sequence ID | Copies | Sequence (5' → 3') |
|---|---|---|
| | | ctctcgggacgac NNNNNNNNNNNNNNNNNNNNNNNNNNNNNN gtcgtccc |
| Glucose-boronic acid_01 | 5 | ctctcgggacga**c** **AGCCGAG**CACTACATTAGTTGGG**CAGCCGA** **g**tcgtccc |
| Glucose-boronic acid_02 | 2 | ctctcgggacgac **CAGCCGAG**ATAGGTCGTTCTAT **CAGCCGA** **G**gtcgtccc |
| Glucose-boronic acid_03 | 1 | ctctcgggacga**c** **AGCCGAG**TTCAGGGATTCCCTAA**CAGCCGA** **g**tcgtccc |
| Glucose-boronic acid_04 | 1 | ctctcgggacga**c** **AGCCGAG**GAAACAAACTTTTTTC**CAGCCGA** **g**tcgtccc |
| Glucose-boronic acid_05 | 1 | ctctcgggacga**c** **AGCCGAG**TCAAAGTTTAACTTGA**CAGCCGA** **g**tcgtccc |
| Glucose-boronic acid_06 | 1 | ctctcgggacga**c** **AGCCGAG**TTATGACATTCAATAA**CAGCCGA** **g**tcgtccc |
| Glucose-boronic acid_07 | 1 | ctctcgggacgac **CAGCCGAG**ATTTTGCATAAAAA **CAGCCGAG** gtcgtccc |
| Glucose-boronic acid_08 | 2 | ctctcgggacgac GACCGTAGGGGTAGCTGTATATGCGGATGA gtcgtccc |
| Glucose-boronic acid_09 | 1 | ctctcgggacgac AACCGAGTAGGATACTAAGCATCCAGCGGA gtcgtccc |
| Glucose-boronic acid_10 | 1 | ctctcgggacgac AGGGGGTAGGGGGCCCGGACTGTTAAGGGT gtcgtccc |
| Glucose-boronic acid_11 | 1 | ctctcgggacgac CGCGGGAGCAATGCGATGACGAAGGACGGG gtcgtccc |
| Glucose-boronic acid_12 | 1 | ctctcgggacgac CGGAGCTCCTGCGATTGACTAAAAGGAAG gtcgtccc |
| Glucose-boronic acid_13 | 1 | ctctcgggacgac GGGACCAACCGGGATGAGCATAAGTGCGAC gtcgtccc |
| Glucose-boronic acid_14 | 1 | ctctcgggacgac GGGGAACGTTTTCGTGGATGAGCTGGCGAC gtcgtccc |
| Glucose-boronic acid_15 | 1 | ctctcgggacgac GGGGGGAGAGTATTGGATGACCGCAGGGACC gtcgtccc |
| Glucose-boronic acid_16 | 1 | ctctcgggacgac GGGGGGAACATTGTGATCTCGTTAGAGCAC gtcgtccc |
| Galactose-boronic acid_01 | 7 | ctctcgggacgac CCAGGTGTCCTGCTTCTCAGTAGTAGGTTA gtcgtccc |
| Galactose-boronic acid_02 | 6 | ctctcgggacgac CGAGTAGGTGTCCTGGATGCAGGTTTGGAG gtcgtccc |
| Galactose-boronic acid_03 | 4 | ctctcgggacgac CACTACGCATAGTTTCTATCGCCAGGAAGG gtcgtccc |
| Galactose-boronic acid_04 | 2 | ctctcgggacgac CGGAGGGATCTCGCAATGTTGTGGGTTTAG gtcgtccc |
| Galactose-boronic acid_05 | 1 | ctctcgggacgac CGAGTAGGTGTCCTGGATATAGGTTTGGAG gtcgtccc |
| Galactose-boronic acid_06 | 1 | ctctcgggacgac GGGAGGGTCGGTTAGCACTAACTGACGGCT gtcgtccc |
| Galactose-boronic acid_07 | 1 | ctctcgggacgac GGGTTGGGATTAACATGACTTAGGGTTCCT gtcgtccc |
| Fructose-boronic acid_01 | 12 | ctctcgggacgac CGGAGTGGGTTCGAGCGTGCGCTACATCAG gtcgtccc |
| Fructose-boronic acid_02 | 3 | ctctcgggacgac GGCTGGCACGTTTGGTTCAAGAATGTGGGT gtcgtccc |
| Fructose-boronic acid_03 | 3 | ctctcgggacgac GGACAGAGGTTCGAGCGTGCGCTCTAGGAA gtcgtccc |
| Fructose-boronic acid_04 | 1 | ctctcgggacgac GGGAAGGGGTTCGAGCGTGCGCTACAGGAA gtcgtccc |
| Fructose-boronic acid_05 | 1 | ctctcgggacgac CGGTTCAGGTTCGAGCGTGCGCTACATCAG gtcgtccc |
| Fructose-boronic acid_06 | 1 | ctctcgggacgac GGGCTGGCATGTTACTCTTGTAATGTGGGTT gtcgtccc |
| Phenylalanine-Cp*Rh(III)_01 | 9 | ctctcgggacgac GGACGCTAATCTTACAAGGGCGTAGTGTAT gtcgtccc |
| Phenylalanine-Cp*Rh(III)_02 | 10 | ctctcgggacgac CGCCGATAATCTCACAAGGGCGTATCAAAG gtcgtccc |
| Phenylalanine-Cp*Rh(III)_03 | 2 | ctctcgggacgac GGGTAGGGATGTCTAATCCCGGCGGGAGCT gtcgtccc |
| Phenylalanine-Cp*Rh(III)_04 | 2 | ctctcgggacgac CAGCGAAGTTTCAAGGGAATGTTGAAGTAG gtcgtccc |
| Tyrosine-Cp*Rh(III)_01 | 3 | ctctcgggacgac **GGCCCG**TAGATATT----------**AGTA** gtcgtccc |
| Tyrosine-Cp*Rh(III)_02 | 3 | ctctcgggacgac **GGCCCG**AATGTGTA----------**AGTA** gtcgtccc |
| Tyrosine-Cp*Rh(III)_03 | 3 | ctctcgggacgac **GGCCCG**ATCTCAG----------**AGTA** gtcgtccc |
| Tyrosine-Cp*Rh(III)_04 | 3 | ctctcgggacgac **GGCCCG**CATTAATT----------**AGTA** gtcgtccc |
| Tyrosine-Cp*Rh(III)_05 | 1 | ctctcgggacgac **GGCCCG**AACATATGTA----------**AGTA** gtcgtccc |
| Tyrosine-Cp*Rh(III)_06 | 1 | ctctcgggacgac **GGCCCG**ATAGTAG----------**AGTA** gtcgtccc |
| Tyrosine-Cp*Rh(III)_07 | 1 | ctctcgggacgac **GGCCCG**ATATGTAATT----------**AGTA** gtcgtccc |
| Tyrosine-Cp*Rh(III)_08 | 1 | ctctcgggacgac **GGCCCG**ACATCATCATCAGTATATAG**AGTA** gtcgtccc |
| Tyrosine-Cp*Rh(III)_09 | 1 | ctctcgggacgac **GGCCCG**ATGTTCCAG----------**AGTA** gtcgtccc |
| Tyrosine-Cp*Rh(III)_10 | 1 | ctctcgggacgac **GGCCCG**AGATAATCA----------**AGTA** gtcgtccc |
| Tyrosine-Cp*Rh(III)_11 | 6 | ctctcgggacgac **GGCCCG**ATCTCAG----------**AGTA** gtcgtccc |
| Tyrosine-Cp*Rh(III)_12 | 5 | ctctcgggacgac **GGCCCG**AATGTGTA----------**AGTA** gtcgtccc |
| Tyrosine-Cp*Rh(III)_13 | 3 | ctctcgggacgac **GGCCCG**ATGTTCCAG----------**AGTA** gtcgtccc |
| Tyrosine-Cp*Rh(III)_14 | 2 | ctctcgggacgac **GGCCCG**ATGATGTATTCG-- -----**AGTA** gtcgtccc |
| Tyrosine-Cp*Rh(III)_15 | 1 | ctctcgggacgac **GGCCCG**AAACTG----------**AGTA** gtcgtccc |
| Tyrosine-Cp*Rh(III)_16 | 1 | ctctcgggacgac **GGCCCG**TAGATATT----------**AGTA** gtcgtccc |
| Tyrosine-Cp*Rh(III)_17 | 1 | ctctcgggacgac **GGCCCG**CAATTAATT----------**AGTA** gtcgtccc |
| Tyrosine-Cp*Rh(III)_18 | 1 | ctctcgggacgac **GGCCCG**AGCACTAGG----------**AGTA** gtcgtccc |
| Tryptophan-Cp*Rh(III)_01 | 4 | ctctcgggacgac CGCGGTAGTCTTAACCTAAAGCGGTGTCAG gtcgtccc |
| Tryptophan-Cp*Rh(III)_02 | 3 | ctctcgggacgac GGGTAGCAGGGTTTATCCTAAGCAGTGTAC gtcgtccc |
| Tryptophan-Cp*Rh(III)_03 | 1 | ctctcgggacgac CACGTAGAGCCAGTATGTCTCGGTGATTGG gtcgtccc |
| Tryptophan-Cp*Rh(III)_04 | 1 | ctctcgggacgac CACGTAGCATAGGTACTTAGCGGTGATTGG gtcgtccc |
| Tryptophan-Cp*Rh(III)_05 | 1 | ctctcgggacgac CGCGGTAGCTAGATATCAAGCGGTGTTCAG gtcgtccc |
| Tryptophan-Cp*Rh(III)_06 | 1 | ctctcgggacgac GGCAGGTAGCTCGAAGAAGTGGTGGGTGCA gtcgtccc |
| Tryptophan-Cp*Rh(III)_07 | 1 | ctctcgggacgac GGGGTCACAGGGGTCCGGGTGTGGGTGGTT gtcgtccc |
| Tryptophan-Cp*Rh(III)_08 | 1 | ctctcgggacgac GGGTAGTGAGTGATAGACCGAACAGTGTAC gtcgtccc |

*The consensus region between sequences are marked (red, bold in glucose-boronic acid complex, and tyrosine-Cp*Rh(III), these motifs form the common pocket shapes in the predicted secondary structures.

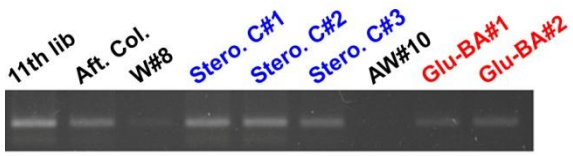
**Table S2. The SELEX/counter-SELEX process by rounds**. The number of rounds of SELEX is determined by the elution profile which has a significantly brighter band post-PCR on an agarose gel for the target in comparison to the last wash and the "counter-target" addition. Often, when counter-SELEX would be started, the brighter band would disappear. In this case, we would increase the number of PCR cycles to get sufficient amount of DNA for the next round. In most cases, the brighter band would reappear after 2-3 rounds of continued SELEX.

| Target | Co-factor | SELEX round process |
|---|---|---|
| **Glucose** | Boronic acid | 1. Glucose-boronic acid complex only SELEX (1st ~ 5th round)<br>2. Boronic acid counter SELEX (6th round to 10th round)<br>3. Steroid mixture counter SELEX (11th round to 14th round) |
| **Fructose** | Boronic acid | 1. Fructose-boronic acid complex only SELEX (1st ~ 3rd round)<br>2. Boronic acid counter SELEX (4th ~12th round)<br>3. Steroid mixture counter SELEX (7th round to 14th round)<br>4. Glucose-boronic acid complex counter SELEX (15th round to 32nd round)<br>5. Galactose-boronic acid complex counter SELEX (15th round to 32nd round) |
| **Galactose** | Boronic acid | 1. Galactose-boronic acid complex only SELEX (1st ~ 3rd round)<br>2. Boronic acid counter SELEX (4th ~12th round)<br>3. Steroid mixture counter SELEX (7th round to 14th round)<br>4. Glucose-boronic acid complex SELEX (15th round to 32nd round)<br>5. Fructose-boronic acid complex SELEX (15th round to 32nd round) |
| **Phenylalanine** | Cp*Rh(III) | 1. Phenylalanine-Cp*Rh (III) complex SELEX (1st ~ 3rd round)<br>2. Cp*Rh(III) counter SELEX (4th~11th round)<br>3. Tyrosine-Cp*Rh(III) complex counter SELEX (12th~18th round) |
| **Tyrosine** | Cp*Rh(III) | 1. Tyrosine-Cp*Rh (III) complex SELEX (1st ~ 6th round)<br>2. Two separated SELEX from 7th round. (A) Cp*Rh(III) counter SELEX (7th ~ 12th round), and (B tyrosine only counter SELEX (7th ~12th round) |
| **Tryptophan** | Cp*Rh(III) | 1. Tryptophan-Cp*Rh(III) complex SELEX (1st ~ 4th round)<br>2. Counter SELEX with combining of tyrosine & phenylalanine-Cp*Rh(III) complex (5th ~ 7th rou<br>3. Separated counter SELEX with Cp*Rh (III), tyrosine-Cp*Rh(III), and phenylalanine-Cp*Rh(III (8th ~ 15th round) |



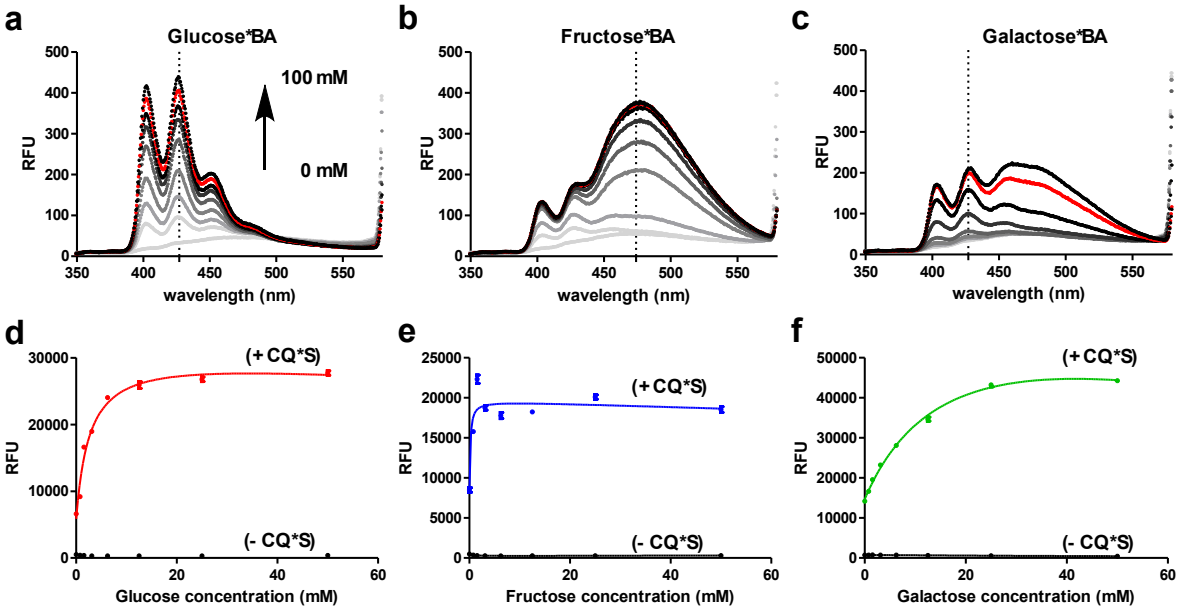

**Double Counter selex with steroid (DCA/DIS/DOG, 250 uM each, 10 times) and BA (50 uM, 10 times)**

An example ('crude') of a PCR-amplifed DNA-elution profile from column in glucose SELEX (14th cycle); "Stero" is counterselection with steroids; "BA" bisboronic acid counterselection, "W", "AW", and "AAW" are representative washes with only buffer, "#" are column volumes of elutions, while "Glu-BA" are elutions that were used for cloning. Earlier round (11th cycle) showing stronger elutions with steroids:

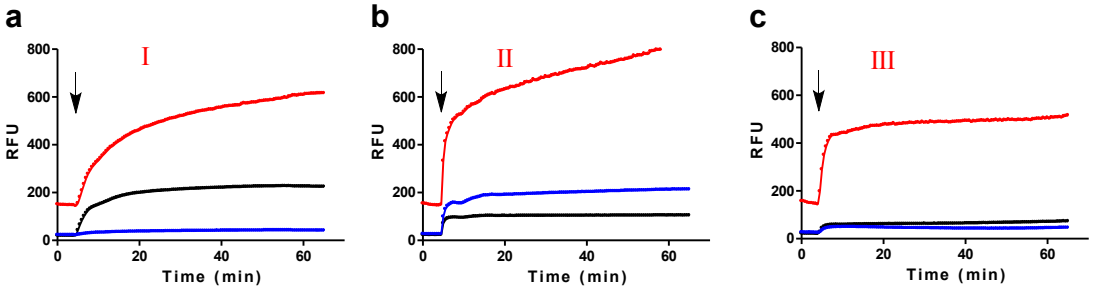



**Counter selex with steroid (DCA/DIS/DOG, 250 uM each, 5 times)**

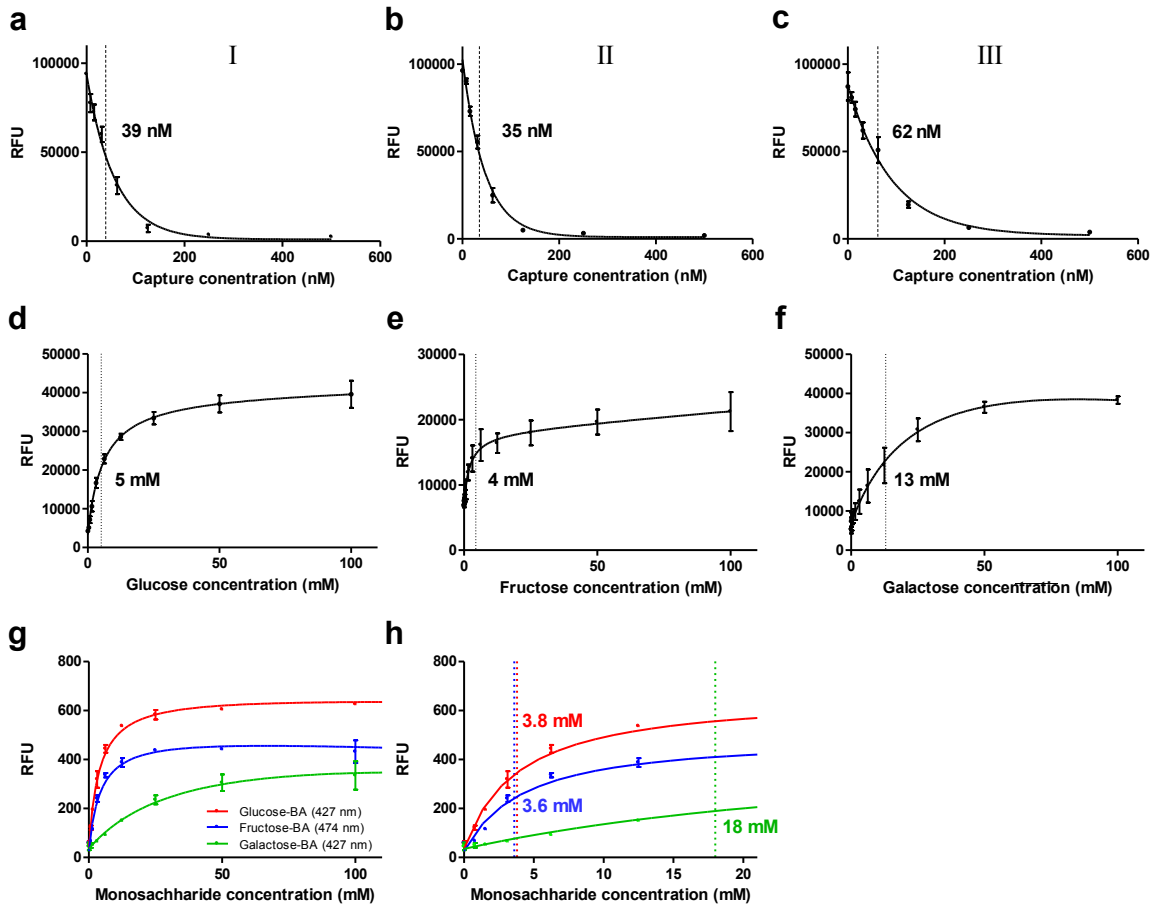
Of note, relative amounts of amplifiable DNA in steroid-eluted and target eluted fractions changed significantly.




**Supplementary Figure 1. Fluorescence emission spectra from the formation of the monosaccharide-bisboronic acid (BA) complexes (obtained on Perkin Elmer LS55 instrument). (a)** Glucose-bisboronic acid complex formation, **(b)** fructose-bisboronic acid complex formation, and **(c)** galactose-bisboronic acid complex formation, in SELEX buffer. The monosaccharide concentrations used were from 0 mM-to-100 mM (0 mM, 1.56 mM, 3.13 mM, 6.25 mM, 12,5 mM, 25 mM, 50 mM (red dashed line), and 100 mM) under a fixed bisboronic acid concentration (50 μM). Each sample was incubated at least 40 minutes at room temperature then excited at 300 nm and the emission spectra obtained from 350-to-580 nm. The vertical dashed lines indicate the emission maxima of each complex and the wavelength used to create **Supplementary Figure. 2.** These differences in spectra support the hypothesis of different conformations with different sugars, and this may have led to successful SELEXes. **(d-f)** Control experiments in which we compared the signal with and without fluorescent aptameric sensor; in this way we demonstrate that monosaccharide-bisboronic acid complex alone is not responsible for the signal observed and reported in our Figure 2 (measured at ex/em = 485-nm/535-nm on a Perkin-Elmer Victor II microplate reader).



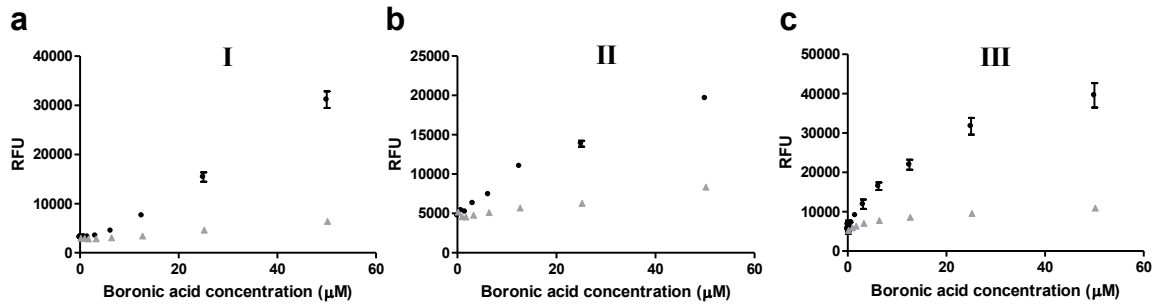

**Supplementary Figure 2. Time-course monitoring of the monosaccharide-bisboronic acid aptamer-sensors demonstrating that complexation with bisboronic acid (Shinkai's) receptor (blue and black lines) coincides with increase in aptameric sensor response (red line).** Pre-incubated sensor-strand (i.e. aptamer) and capture-strand mixture was added to the bisboronic acid (50 μM) solution, then monitored for 5 min before adding glucose to sensor **I (a)**, fructose to sensor **II (b)**, or galactose to sensor **III (c)**. The arrows indicate the point of each monosaccharide addition. After adding monosaccharide solution, three different wavelengths were monitored in parallel: wavelength 1 (■: ex/em=300/427) is for monitoring glucose-boronic acid complex formation and galactose-bisboronic acid complex formation (maxima for this complex is at this wavelength), wavelength 2 (■: ex/em=300/474) for fructose-bisboronic acid complex formation (maxima for this complex), and wavelength 3 (■: ex/em=495/520) for monitoring the binding of fluorescein-conjugated aptameric sensors **I-III** to their targets. The fluorescein signals from the sensors (red) increase upon the formation of monosaccharide-bisboronic acid complex while no sensor response was observed with bisboronic acid only.



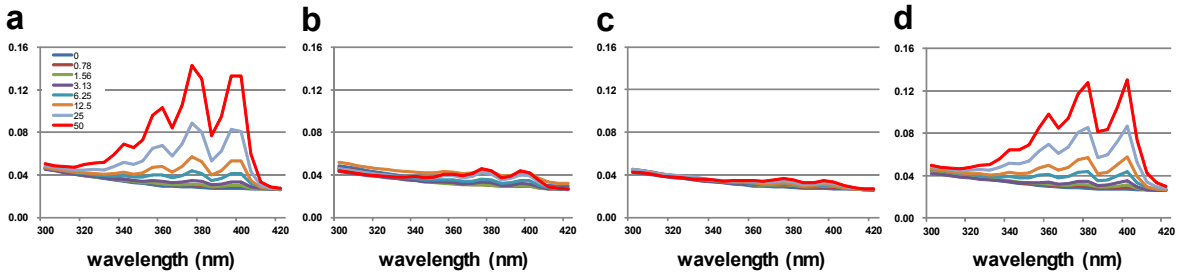

**Supplementary Figure 3. Determination of $K_d$ through a competitive binding assay for three aptameric sensors (I-III). (a, b, c)** are quenching plots for individual sensors in the presence of increasing amounts of competitor oligonucleotide-dabcyl quencher (capture), used to determine $K_d$ values of competitor oligonucleotides to aptamers; These plots are obtained by adding an increasing concentration of a capture strand, and observing the increase in quenching, with $K_d$ approximated by half-maximum value; **(d, e, f)** are plots showing response in competitive assays of sensor mixtures (fluorescein-labeled aptamer + competitor dabcyl-labeled oligonucleotide) to the addition of Shinkai's receptor in the presence of increasing amounts of sugar, with the actual concentration of a receptor*complex at any particular glucose concentrations calculated from the $K_d$ of Shinkai's receptor. The vertical dotted lines in each graph indicate the concentration at the half-maximum value of signal, and the numerical values are shown in the plots. These $K_d$'s are approximate and in all our determinations we neglect solubility issues and additional interactions with oligonucleotides.

A more detailed clarification, based on the reference '*A simple and rapid approach for measurement of dissociation constants of DNA aptamers against proteins and small molecules via automated microchip electrophoresis'* by Hu and Easley and using their terminology: There are two separate dose-response measurements for each sensor, leading to $K_{d,\,eff1}$ and $K_{d,\,eff2}$. $K_{d,\,eff1}$ is the dissociation constant between sensor and capture and is calculated based on the decreasing fluorescent intensity of the sensor under the

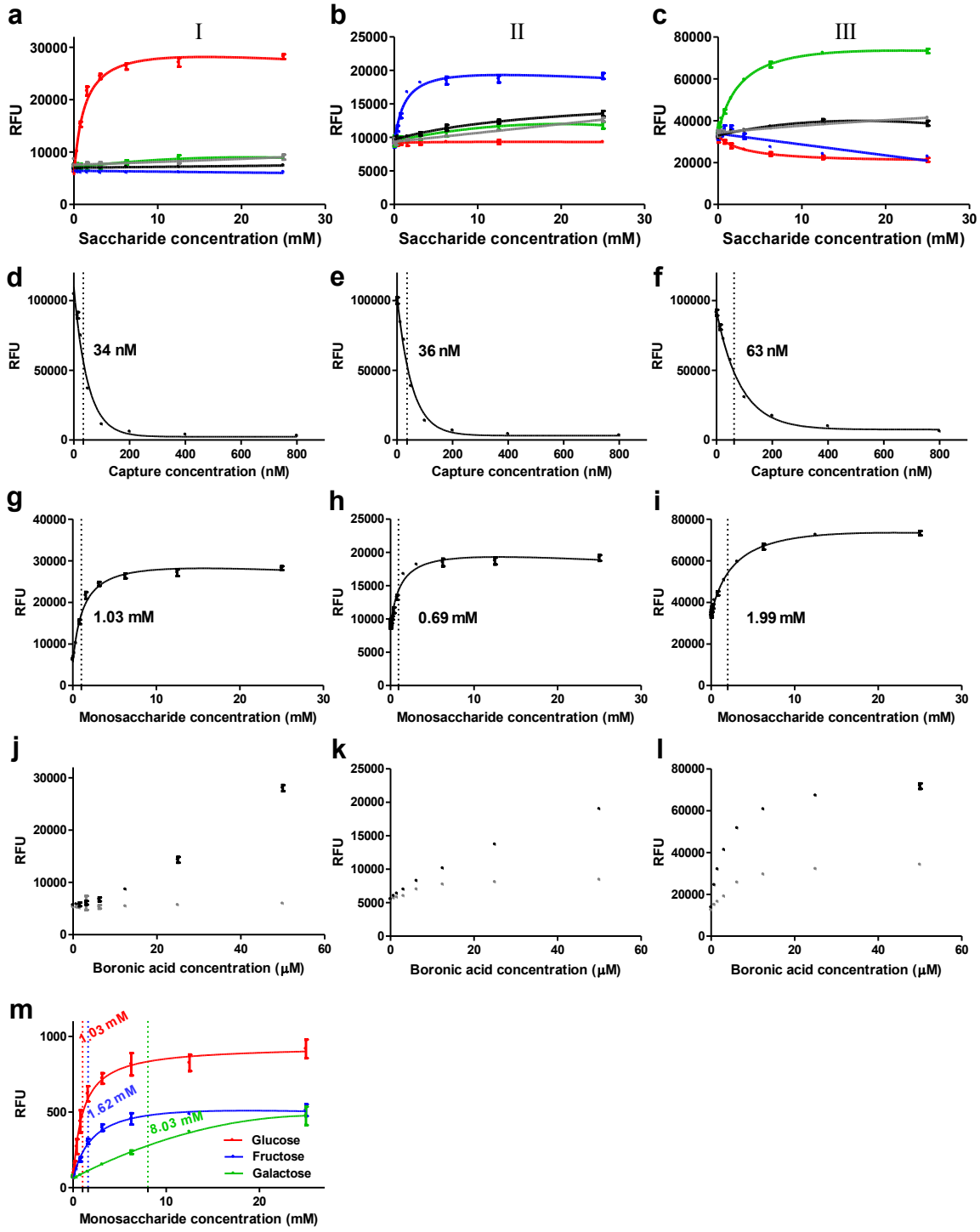
fixed concentration of sensor (50 nM) and 0~ 500 nM range of the capture concentration (**a-c**), the dashed lines indicate 50% of the sensor fluorescence level. We define $K_{d,\ eff1}$ as: $K_{d,\ eff1}$ =[free sensor][free capture strand]/[sensor-capture strand]. $K_{d,\ eff2}$ is a unitless constant between the sensor-capture complex and its target ($K_{d,\ eff2}$ = [capture strand][aptameric sensor-target]/[aptameric sensor-capture][target]). To calculate the target (monosaccharide-bisboronic acid complex) concentration, [target] = [bisboronic acid $_{total}$]·[free monosaccharide]/ ($K_d$ of Shinkai's sensor + [free monosaccharide]) is applied, based on the assumption that if the concentration of total monosaccharide >> total bisboronic acid, then we could make the approximation that free monosaccharide is equal to total monosaccharide. $K_d$ is derived from $K_{d,\ eff1}/K_{d,\ eff2}$. (**a, d**) represent the plots for glucose, (**b, e**) for fructose, and (**c, f**) for galactose. **(g, h)** $K_d$ value of Shinkai's sensor are ≈ 3.8 mM (glucose), ≈ 3.6 mM (fructose), and ≈ 18 mM (galactose) under our experiment conditions; these fluorescence measurements with Perkin Elmer LS55 spectrofluorimeter. Estimated $K_d$ values for our sensors (i.e., their ability to recognize complexes of Shinkai's receptor with individual sugars) are then ≈ 1.7 μM (glucose (●)), ≈ 2.0 μM (fructose (●)), and ≈ 3.3 μM (galactose (●)). (**h**) is the enlarged plot showimg the changed X-axis scale setting from the same experiment from (**g**). All measurements were performed in triplicate, according to the methods in the experimental section above.




**Supplementary Figure 4. Interactions of the three monosaccharide aptameric sensors with bis-boronic acid and sugar-bisboronic acid complex at constant sugar concentrations in SELEX buffer with 2% MeOH. (a)** glucose-bisboronic acid sensor (**I**) with a fixed concentration of glucose (50 mM, black dot) or without glucose (grey dot) under increasing bisboronic acid concentrations. (**b**) fructose-bisboronic acid sensor (**II**), and (**c**) galactose sensor (**III**). Results show that these aptameric sensors are responding to the monosaccharide-bisboronic acid complex and poorly to the bisboronic acid alone; we cannot exclude the possibility that this is due in part to lack of solubility without sugar. The shape of RFU-concentration dependence cannot be satisfactorily fit to a 1:1 binding model (thus curve fitting was not performed), and this observation is supporting the hypothesis that more than one Shinkai's receptor binds to each aptamer. Attempts to perform Job plot analysis lead to the difficult-to-interpret results.

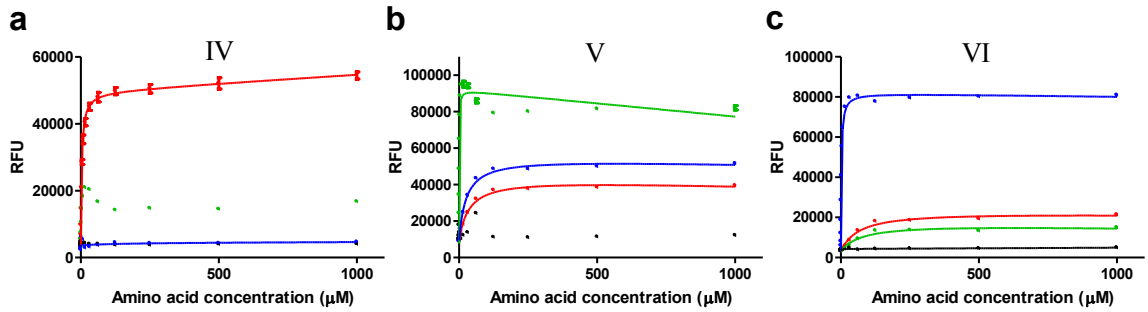



**Supplementary Figure 5. UV/Vis spectra used in estimating solubility behavior of the bisboronic acid receptor under different conditions used in this work.** To check the extent of solubility of the bisboronic acid under our SELEX conditions we measured UV/VIS spectra of solutions after dissolution, and centrifugation to remove any insoluble material: **(a)** bisboronic acid dissolves readily in 100% methanol (MeOH) used here as a positive control for solubility, **(b)** SELEX buffer with 20% MeOH, **(c)** SELEX buffer with 2% MeOH before centrifugation, and **(d)** SELEX buffer with 2% MeOH and glucose (50 mM); buffer with 2% MeOH without centrifugation was used for the SELEX process. The bisboronic acid was prepared in the different solvent mixtures, incubated for 40 min at room temperature, centrifuged at 13,200 rpm for 5 min to remove any insoluble matter, then the supernatant measured by SPECTRAMAX (Molecular Devices, CA, USA) (sixty four-fold serial dilutions from 50 μM – red line is this starting dilution- each concentration is indicated in the graph legend of **a**). The bandwidth of the light source was 5 nm, scanning from 300-to-420 nm. In conclusion, and in concordance with previous reports of Shinkai's group, the receptor is sparingly soluble in water mixtures without larger methanol content, while the complex with glucose (and other sugars) is fully soluble.

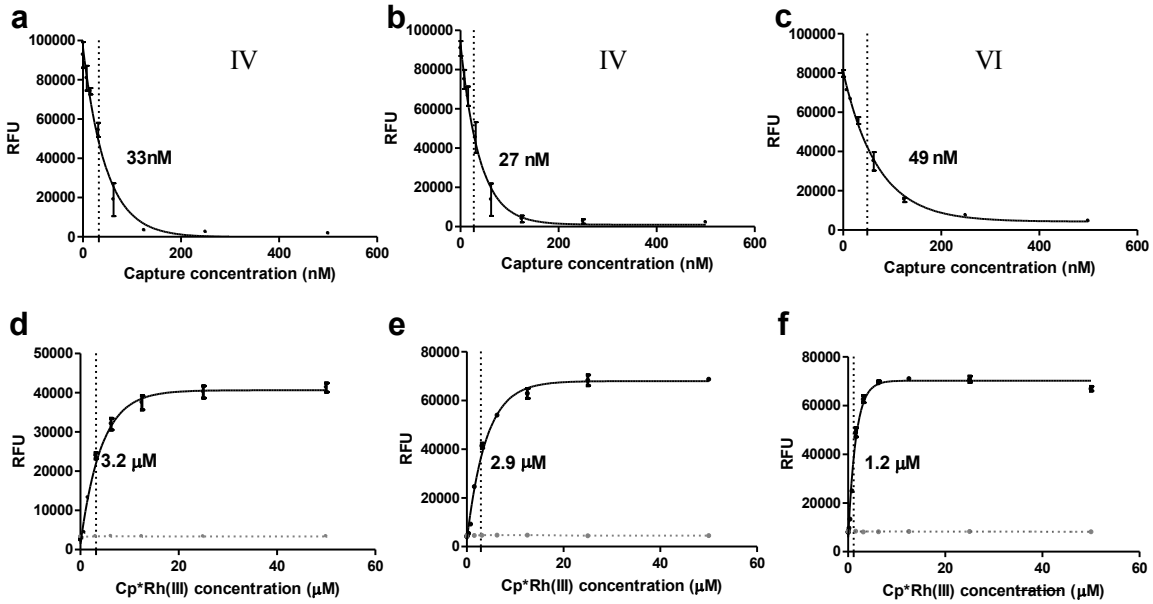



**Supplementary Figure 6. The response of monosaccharide aptameric sensors I-III under buffer conditions containing 20% of MeOH.** We carried out identical experiments to **Fig 2** in the main text, and **Supplementary Figure 3** and **Supplementary Figure 4,** under buffer conditions containing a higher concentration of MeOH (20%) (compared to 2% that was used elsewhere). **(a, b, c)** are the responses of sensor **I-III** and the tested saccharides glucose(●), fructose (●), galactose (●), mannose (●), and disaccharide mixture (●). **(d, e, f)** are quenching plots for individual sensors, **(g, h, i)** are plots showing response in

competitive assays of sensor mixtures (fluorescein-labeled aptamer + competitor dabcyl-labeled oligonucleotide) to the addition of Shinkai's receptor in the presence of increasing amounts of sugar. **(j, k, i)** are the interactions of the three monosaccharide aptameric sensors for sugar-bisboronic acid complex versus bisboronic acid alone. **(j)** glucose-boronic acid sensor (**I**) with fixed concentration of glucose (50 mM, black dot) or without glucose (grey) under the differing boronic acid concentrations. (**k**) is from the fructose-boronic acid sensor (**II**), and **(i)** is from galactose sensor (**III**). **(m)** The $K_d$ value of Shinkai's sensor under buffer conditions containing 20% MeOH is ≈ 1.03 mM (glucose), ≈ 1.62 mM (fructose), and ≈ 8.03 mM (galactose). Again, these $K_d$ values are approximate, and we neglect any solubility or multiple binding site issues.

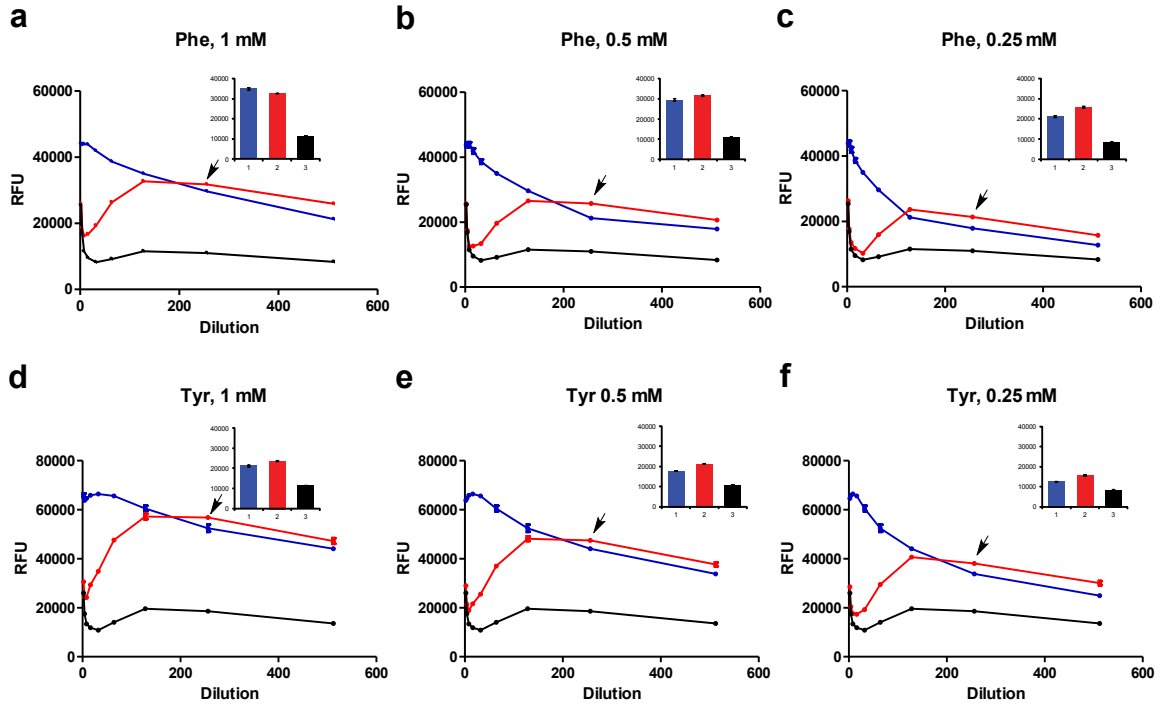



**Supplementary Figure 7. Response of sensors to amino acids (full concentration range up to 1 mM).** **(a-c)** The response of sensors **IV-VI**, i.e., for phenylalanine (red), tyrosine (blue), tryptophan (green), and a mixture of eight amino acids (black): While the sensors for Tyr and Trp sense their own targets with high selectivity at low concentrations **(Fig. 3.** in main text**,** 1-5 μM**),** poorer selectivity was observed at higher concentrations. Response to Trp for both Phe (**IV**) and Tyr (**V**) sensors was unusual, showing a spike, and could not be fitted properly in 1:1 binding. While we speculate that this effect is due to quenching, we are uncertain at this stage of its cause.




**Supplementary Figure 8. Determination of $K_d$ through a competitive assay for the three aptameric amino acid sensors.** We used the same method as described for bisboronic receptors. (**a-c**) are quenching plots for individual sensors, used to determine $K_d$ values of competitor oligonucleotides to aptamers; **(d, e, f)** are plots showing response in competitive assays of sensor mixtures (fluorescein-labeled aptamer + competitor dabcyl-labeled oligonucleotide) to the addition of Cp*Rh(III) in the presence of increasing amounts of amino acids. The grey dotted lines in (**d-f**) show the Cp*Rh(III) concentration-dependent response of the sensor without the indicated amino acids. (**a, d**) shown for the phenylalanine-Cp*Rh(III) sensor, (**b, e**) tyrosine-Cp*Rh(III) sensor, and (**c, f**) for the tryptophan-Cp*Rh(III) sensor. All measurements were performed in triplicate. The dashed lines indicate half-

maximum value of the sensor fluorescence level. The derived $K_d$ of each is ≈ 180 nM (phenyalanine), ≈ 60 nM (tyrosine), and ≈ 120 nM (tryptophan). This determination was different from the one with bisboronic acid (Shinkai's) receptor, because Cp*Rh(III) was fully soluble, and showed no clear indications of multiple binding interactions with aptamers; however, at higher concentrations Cp*Rh(III) quenched fluorescein. In curve fitting (1:1 model), we neglected points that showed quenching (e.g., in graph f with VI 50 μM CpRh(III) is the most obvious example of quenching); we also assumed that no additional interactions with oligonucleotides were occurring (at least these were invisible to our measurements). These assumptions may lead to errors in measuring affinity.

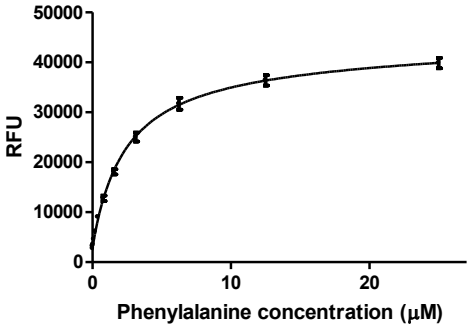



**Supplementary Figure 9. Aptameric sensor response in serially diluted human serum.** In this figure we show that we can detect amino acids at high dilutions in serum, under the conditions in which Cp*Rh(III) complex is present at sufficient quantities, so other ligands (e.g., other amino acids in serum) are not strongly competing with the target which is then recognized by sensor.

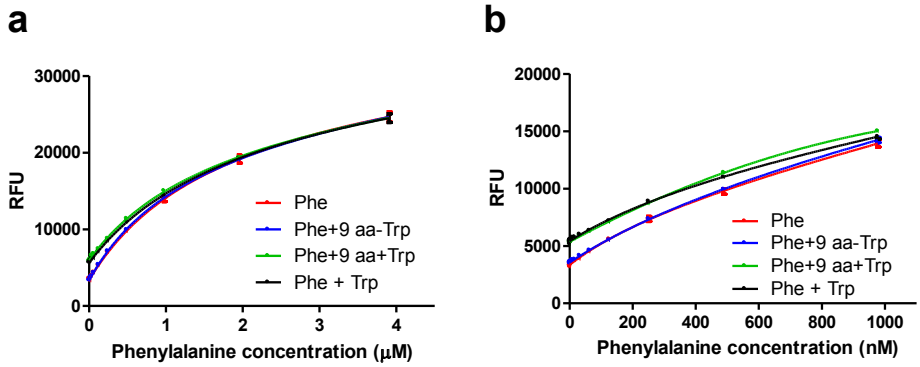
**Background:** Serum was spiked with known concentrations of Phe or Tyr (e.g., 1 mM, 0.5 mM, and 0.25 mM; these are reasonable concentrations to assume for sera of patients with inborn errors of metabolism). The three starting solutions, serum, spiked serum, and buffer (with the same spiked concentrations), were then serially diluted into 100 μM of Cp*Rh(III). Buffer dilutions show expected effects (drop in signal), while serum and spiked serum, after initial drop, show an increase, and then drop again. This is due to the presence of other ligands in solution, which compete for Cp*Rh(III) with targets, but do not compete for sensor (i.e., after certain dilution there is enough derivatization agent so we start seeing the signal). In non-spiked sera, we see the background signal from naturally present Phe in healthy controls.

**Legend:** (**a-c**) is the aptameric sensor responding to phenylalanine-Cp*Rh(III) complex at different concentrations in serum, and (**d-f**) is for tyrosine-Cp*Rh(III) complex.

A line and scatter plot of the sensor in serially diluted 1 mM (**a, d**), 500 μM (**c, e**) and 250 μM (**c, f**) of phenylalanine or tyrosine solution in buffer (blue), diluted serum spiked with Phe or Tyr to the final concentrations as shown (red), and the same serum (black). A 100 μM concentration of monomeric Cp*Rh(III) was added to all samples to derivatize all bi- and tri-dentate ligands. At certain dilutions, the difference between spiked and non-spiked serums was observed, despite the presence of all interferences (including other amino acids). The inserted bar graphs show data for the 1:256 dilution samples. The data indicates that phenylalanine and tyrosine can be detected in a dilute-and-measure assay at clinically relevant concentrations.

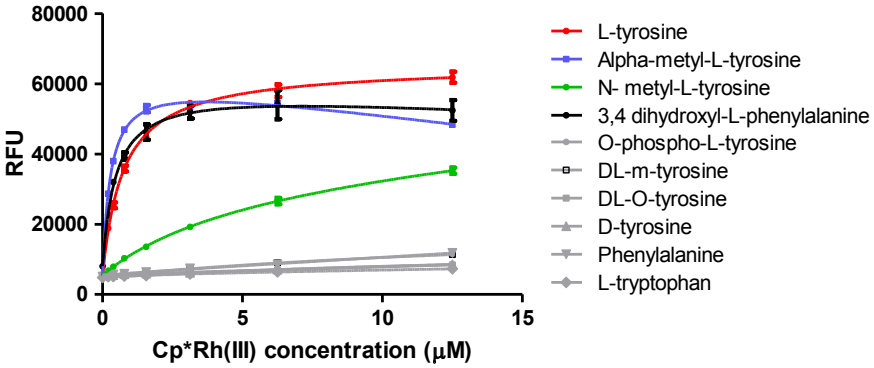



**Supplementary Figure 10. Standard curve for phenyalanine quantification.** The resulting fluorescent intensity of phenylalanine-Cp*Rh(III) with sensor (**VI**) was plotted by serial dilution of phenylalanine from 0 μM to 25 μM in 100 μM concentration of monomeric Cp*Rh(III). Based on this curve, the concentration of phenylalanine in the diluted sera was determined (**Figure 4b** in main text**)**. Standard deviations of three measurements are shown. No further corrections for other amino acids or serum albumin were performed; therefore the concentrations in **4b** are approximate (see next Figure).

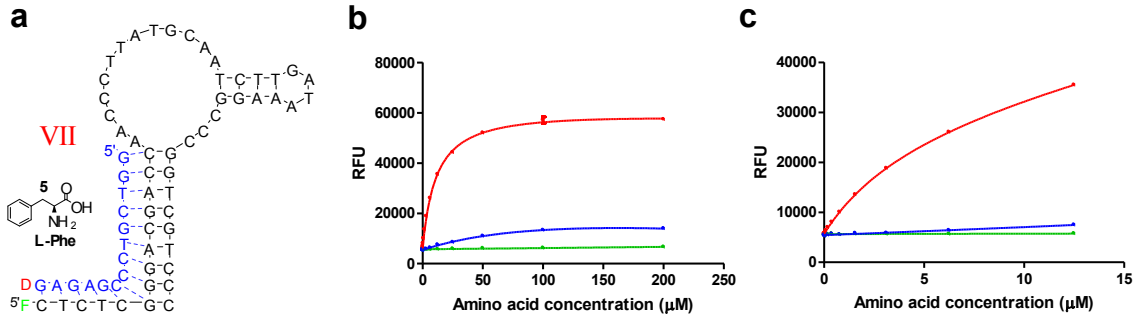



**Supplementary Figure 11. Phenylalanine-Cp*Rh(III) sensor cross-reactivity with tryptophan and other amino acids .** The range of tryptophan in serum of the human adult is ≈ 30 ~ 90 μM, when the serum is diluted 200 times, that is between 150 nM ~ 450 nM. To check the cross-reactivity of the phenylalanine-Cp*Rh(III) sensor with tryptophan, the indicated phenyalanine concentration on the plot

was measured under the fixed concentration of spiked tryptophan (375 nM, equivalent to 75 μM of tryptophan concentration in the 1/200 dilution serum). **(a, b)** show the changed X-axis scale setting from the same experiment, four different conditions were tested: phenylalanine (●), phenylalanine including spiked with 9 different amino acids except tryptophan (●), phenylalanine including spiked with tryptophan plus 9 different amino acids (●), and phenylalanine including spiked with tryptophan (●). The spiked 9 amino acids are alanine, arginine, glutamine, glycine, histidine, leucine, lysine, serine and valine, the spiked concentration of each amino acid is 375 nM. Plot b shows an expansion of submicromolar region; above this concentration, there is no difference in calibration curves.




**Supplementary Figure 12. Tyrosine-Cp*Rh(III) sensor response to tyrosine analogs.** This is a plot of tyrosine-Cp*Rh(III) sensor response to various tyrosine analogs. In the presence of 1 mM concentration of tyrosine analogs, Cp*Rh(III)-dependent behaviors were analyzed. All measurements were performed in triplicate with standard deviations shown.




**Supplementary Figure 13. A phenylalanine-Cp*Rh(III) sensor variant.** This is the sensor variant (**VII**) from another SELEX in which we applied strong tryptophan-Cp*Rh(III) counterselection. This shows the decreased tryptophan responses in comparison to sensor **IV** in the main manuscript. In the presence of 100 μM concentration of Cp*Rh(III), Phe-Cp*Rh(III) (●), Tyr-Cp*Rh(III) (●), and Trp-Cp*Rh(III) (●)-dependent behaviors were analyzed, showing that cross-reactivity with Trp has been completely eliminated. **(b, c)** show the expanded x-axis scale from the same experiment. All measurements were performed in triplicate with standard deviations shown. This result was obtained after the initial submission of the manuscript, and was not available to the initial reviewers.