

# Two-dimensional combinatorial screening enables the bottom-up design of a microRNA-10b inhibitor

Sai Pradeep Velagapudi<sup>a,b</sup> and Matthew D. Disney<sup>b,\*</sup>

<sup>‡</sup>*Department of Chemistry, The State University of New York at Buffalo, Buffalo, New York 14260*

<sup>†</sup>*Department of Chemistry, The Scripps Research Institute, Scripps Florida, 130 Scripps Way #3A1, Jupiter, FL 33458*

*\* author to whom correspondence is addressed: Disney@scripps.edu*

<b>Table of Contents:</b>	<b>Pages</b>
1. Synthetic procedures for ligands used in 2DCS	2-6
2. Characterization of newly synthesized compounds ( <sup>1</sup> H and <sup>13</sup> C NMR spectra and analytical HPLC chromatograms)	7-14
3. 2DCS selection of RNA motifs that bind to G Neo B and G Kan A	15
4. Sequences of selected RNAs and RNA-PSP analysis.	16-23
5. Determination of binding affinity using fluorescence-based binding assays	24
6. Representative binding curves	25-30
7. Representative binding curve for G Neo B and 5'AUACC/3'UAAGG, an internal loop found in miR-10b	31
8. Secondary structure of control miR-149	32
9. G Neo B affects miR-10b biogenesis	33
10. References	34

## Section 1: Synthetic procedures for ligands used in 2DCS

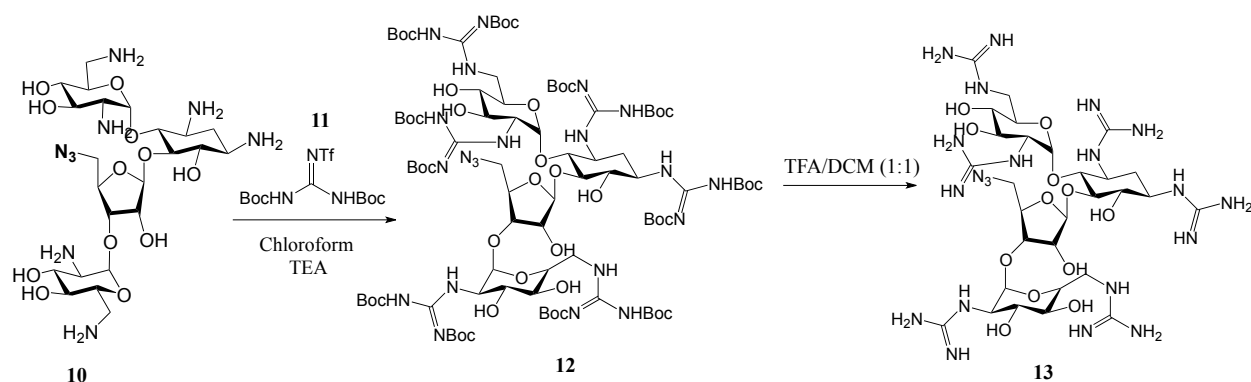
**Abbreviations:** Boc, tert-butyloxycarbonyl; BSA, bovine serum albumin; DCM, dichloromethane; DMSO, dimethyl sulfoxide; HPLC, high performance liquid chromatography; HRMS, high resolution mass spectrometry; MALDI, matrix-assisted laser desorption/ionization; MS-ESI, mass spectrometry-electrospray ionization; TBTA, tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine; TFA, trifluoroacetic acid; TLC, thin layer chromatography

**Instrumentation:** Mass spectra were collected on a Varian 500-MS IT mass spectrometer. NMR spectra were acquired using Varian NMR spectrometer (operating at 100 MHz on <sup>13</sup>C and 400 MHz on <sup>1</sup>H). Preparative HPLC and analytical HPLC separations were completed on a Waters 1525 binary HPLC pump equipped with a Waters 2487 dual absorbance detector system.

**HPLC:** Preparative HPLC purifications were completed at room temperature on a Waters Atlantis Prep C18 column (5 μM, 19x150 mm) using a linear gradient of 0% to 100% B in A and a flow rate of 5 mL/min. The purity of the product was assessed by analytical HPLC using a Waters Symmetry® C18 (5 μm, 4.6x150 mm) column at room temperature, a flow rate of 1 mL/min and a linear gradient of 0% to 100% B in A over 50 min. Absorbance was monitored at 218 and 256 nm. A is 0.1% (v/v) TFA in water and B is 0.1% (v/v) TFA in acetonitrile.

**Synthesis of GuanidinoBoc<sub>12</sub>-5''-azido Neomycin B (12):** The overall reaction scheme for the synthesis of guanidino-5''-azido neomycin B (**13**) is shown in Scheme 1. Briefly, 5''-azido neomycin B (**10**) and *N,N'*-diBoc-*N''*-trifylguanidine (**11**) were synthesized according to previously published methods.<sup>1, 2</sup> To a solution of **10** (50 mg, 78 μmoles) in 25 mL of chloroform was added **11** (1500 mg, 3.8 mmoles, 49 eq) and 5 mL of triethylamine. The reaction mixture was stirred at room temperature, and reaction progress was monitored by TLC (5% methanol in DCM).<sup>3, 4</sup> Once the reaction reached completion, the mixture was concentrated under vacuum and purified by flash chromatography using a gradient of 0-20% methanol in DCM. The product was recovered (4 mg) at 7% methanol. HRMS MS-ESI calculated: 2092.0787; observed: 1047.0477 (M+2H/2); 698.3647 (M+3H/3). <sup>1</sup>H NMR (CDCl<sub>3</sub>,

400 MHz)  $\delta$  11.42(s), 11.33, 11.28, 9.21, 8.83 (d), 8.38, 8.19 (d), 5.75 (s), 5.66 (d), 5.10 (s), 5.03 (s), 4.34-4.22 (m), 4.10-3.94 (m), 3.91-3.80 (m), 3.64-3.56 (m), 3.40-3.24(m), 2.88, 2.81,, 2.28, 1.71, 1.43, 1.19.  $^{13}\text{C}$  NMR ( $\text{D}_2\text{O}$ , 100 MHz)  $\delta$  163.54, 163.45, 163.12, 162.99, 162.93, 157.08, 156.08, 155.87, 152.92, 152.62, 152.02, 128.23, 96.65, 83.84, 83.46, 82.60, 79.66, 79.49, 79.37, 72.80, 54.57, 52.03, 51.48, 50.02, 49.03, 39.98, 29.70, 28.21, 28.10, 28.01.

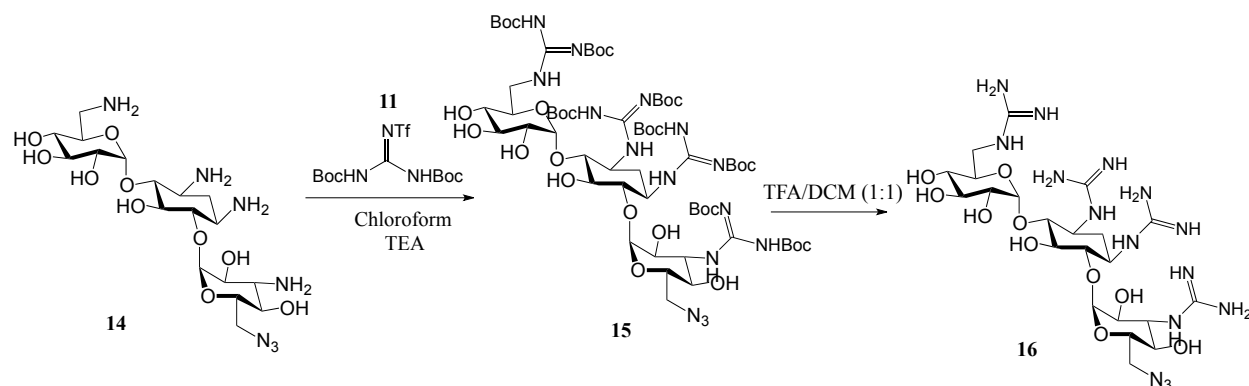


**Scheme S-1:** Synthesis of Guanidino-5''-azido Neomycin B (G Neo B; **13**).

**Synthesis of Guanidino-5''-azido Neomycin B (**13**):** A sample of **12** was de-protected in 1 mL of DCM:TFA (1:1), and the reaction was stirred for 2 h. The product-containing liquid was dried under air, dissolved in 5 mL of ether, and tumbled for 1 h. The ether was decanted, and the solid was dissolved in water and lyophilized, yielding 1.7 mg of product. HRMS MS-ESI calculated: (M+H) 892.4567; observed: 446.7324 (M+2H/2), 298.1575 (M+3H/3).  $^1\text{H}$  NMR ( $d_6$ -DMSO, 400 MHz)  $\delta$  8.15 (d, J=8.1 Hz), 8.09 (d, J= 6 Hz), 7.75(s), 7.52-7.48 (m), 7.36-7.03 (m), 6.88 (d, J=9.06 Hz), 5.95 (s), 5.88 (s), 5.86 (s), 5.77 (s), 5.72 (d, J=3.62 Hz), 5.37 (s), 5.01 (s), 4.83 (s), 4.23 (d, J= 3.68 Hz), 4.16(m), 3.96-3.46 (m), 3.81-3.14 (m), 2.08 (s), 1.98 (d, J=8.99 Hz), 1.52 (d, J=11.8 Hz), 1.23 (s):  $^{13}\text{C}$  NMR ( $\text{D}_2\text{O}$ , 100 MHz)  $\delta$  163.32, 163.11, 162.76, 162.41, 157.89, 157.37, 157.28, 154.11, 151.54, 117.78, 114.88, 111.98, 111.19, 98.30, 95.84, 79.80, 77.19, 72.79, 71.24, 69.09, 66.76, 65.99, 55.46, 53.43, 50.39, 41.86, 38.70, 31.96.

**Synthesis of GuanidinoBoc<sub>3</sub>-6''-azido Kanamycin A (**15**):** The overall reaction scheme for the synthesis of guanidino-6''-azido kanamycin A (**16**) is shown in Scheme 2. Briefly, 6''-azido kanamycin A (**14**) was synthesized according to a previously published method.<sup>1</sup> To a solution

of **14** (20 mg, 39  $\mu$ moles) in 25 mL of chloroform was added **11** (400 mg, 1 mmole, 26 eq) and 5 mL of triethylamine. The reaction mixture was stirred at room temperature, and reaction progress was monitored by TLC (5% methanol in DCM).<sup>3, 4</sup> Once the reaction reached completion, the mixture was concentrated under vacuum and purified by flash chromatography using a gradient of 0-20% methanol in DCM. The product was recovered (10 mg) at 7% methanol. HRMS MS-ESI calculated: (M+H) 1478.7584; observed: 1478.7559, 739.8828 (M+2H/2), 493.5913 (M+3H/3). <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz)  $\delta$  11.48 (s), 11.42 (s), 11.40 (s), 11.29 (s), 8.84 (s), 8.35 (s), 8.24 (d, J = 4 Hz), 7.96 (s), 7.59 (s), 5.64 (s), 5.38 (s), 5.04-2.81 (m), 2.29 (s), 2.01-1.19 (m). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz)  $\delta$  163.61, 163.32, 162.46, 162.05, 158.47, 157.62, 156.57, 155.82, 153.29, 152.74, 152.64, 150.91, 148.95, 101.80, 98.51, 85.98, 83.75, 83.55, 83.33, 83.14, 82.91, 79.96, 79.69, 79.60, 74.76, 72.91, 72.40, 72.10, 71.88, 71.64, 69.58, 69.19, 57.76, 51.33, 49.46, 48.76, 40.88, 34.74, 29.70, 29.56, 29.33, 28.24, 28.14, 27.98.



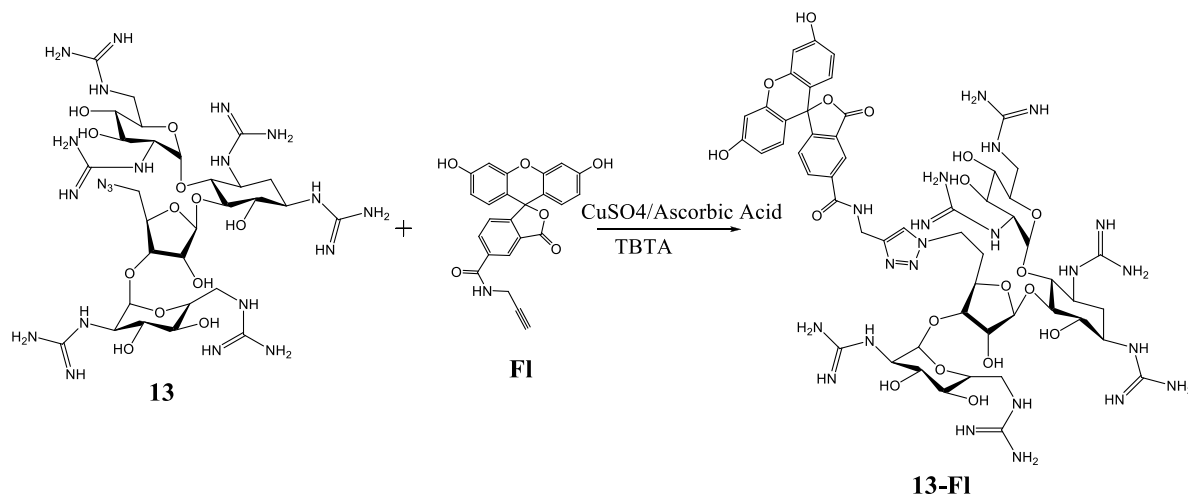
**Scheme S-2:** Synthesis of Guanidino-6''-azido Kanamycin A (G Kan A; **16**).

**Synthesis of Guanidino-6''-azido Kanamycin A (**16**):** A sample of **15** was de-protected in 1 mL of DCM:TFA (1:1), and the reaction was stirred for 2 h. The product-containing liquid was dried under air, dissolved in 5 mL of ether, and tumbled for 1 h. The ether was decanted, and the solid was dissolved in water and lyophilized, yielding 4.3 mg of product. HRMS MS-ESI calculated: (M+H) 678.3390; observed: 678.3398, 339.6734 (M+2H/2), 226.7846 (M+3H/3). <sup>1</sup>H NMR (d<sub>6</sub>-DMSO, 400 MHz) 8.17 (d, J=6.7 Hz), 7.97 (d, J= 4.4 Hz), 7.81 (d, J= 8.1 Hz), 7.34-7.15 (m), 5.89 (s), 5.72 (d, J= 8.0 Hz), 5.65 (s), 5.42 (s), 5.15-4.95 (m), 4.32 (d, J= 8.0 Hz), 3.64-3.08 (m), 2.95 (s), 2.55 (s), 2.51 (s), 2.09 (s), 1.49 (m), 1.28-1.18 (m). <sup>13</sup>C NMR (D<sub>2</sub>O, 100 MHz)

163.49, 163.13, 162.77, 162.42, 158.31, 157.92, 156.61, 156.47, 156.40, 140, 120.68, 117.78, 114.89, 111.99, 99.89, 98.49, 82.01, 80.81, 74.68, 71.26, 71.01, 69.44, 68.65, 56.94, 51.40, 50.73, 50.22, 41.86, 38.70, 32.80.

### Synthesis of fluorescein labeled conjugates of **13** and **16**: **13-FI** & **16-FI**

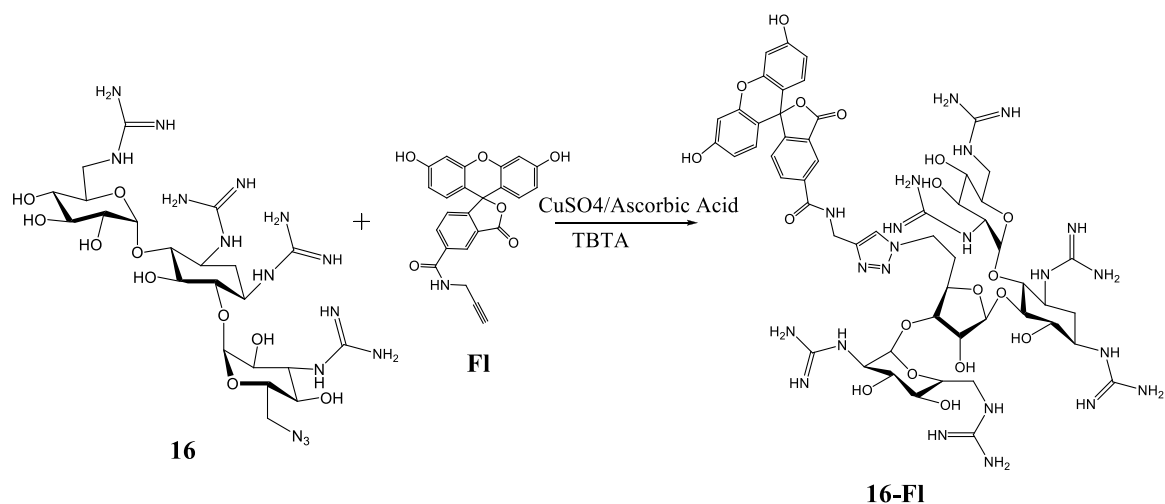
**Synthesis of 13-FI.** 5-(*N*-(2-propyne)-formamide)-fluorescein (**FI**) was synthesized according to previous published method.<sup>5</sup> A 200 nmol sample of **FI** dissolved in methanol was added to a solution containing 600 nmol of **13**, 180 nmol of TBTA (dissolved in 4:1 2-butanol: DMSO), 1  $\mu$ mol CuSO<sub>4</sub>, and 5  $\mu$ mol freshly prepared ascorbic acid. The final volume was brought to 1 mL with methanol. The reaction mixture was microwaved for 4 h at 110°C with stirring in an Emrys<sup>TM</sup> microwave system (Biotage). The product was purified by preparative HPLC as described above and characterized by ESI-MS. HRMS: calculated: 1305.5473 (M+H); observed: 435.8543 (M+3H/3). The purity of the product was determined by analytical HPLC as described above;  $t_R$  = 33 min; 35% yield as determined by absorbance at 496 nm in 1X PBS, pH 7.4 using an extinction coefficient of 45,000 M<sup>-1</sup>cm<sup>-1</sup>.<sup>6</sup>



**Scheme S-3:** Synthesis of **13-FI**.

**Synthesis of 16-FI.** A 500 nmol sample of **FI** in methanol was added to a solution containing 1.5  $\mu$ mol of **16**, 450 nmol of TBTA (dissolved in 4:1 2-butanol: DMSO), 2.5  $\mu$ mol CuSO<sub>4</sub>, and 5  $\mu$ mol freshly prepared ascorbic acid. The final volume was brought to 1 mL with methanol. The reaction mixture was microwaved for 4 h at 110°C with stirring in an Emrys<sup>TM</sup> microwave

system (Biotage). The product was purified by preparative HPLC as described above and characterized by MALDI-MS. HRMS calculated: 1091.4289 (M+H); observed: 1091.4234 (M+H). The purity of the product was determined by analytical HPLC as described above;  $t_R = 35$  min; 30% yield as determined by absorbance at 496 nm in 1X PBS, pH 7.4 using an extinction coefficient of  $45,000 \text{ M}^{-1}\text{cm}^{-1}$ .<sup>6</sup>



**Scheme S-4: Synthesis of 16-FI.**

Section 2: NMR spectral characterization of **12**, **13**, **15** and **16** and analytical HPLC chromatograms of **13-FI** and **16-FI**

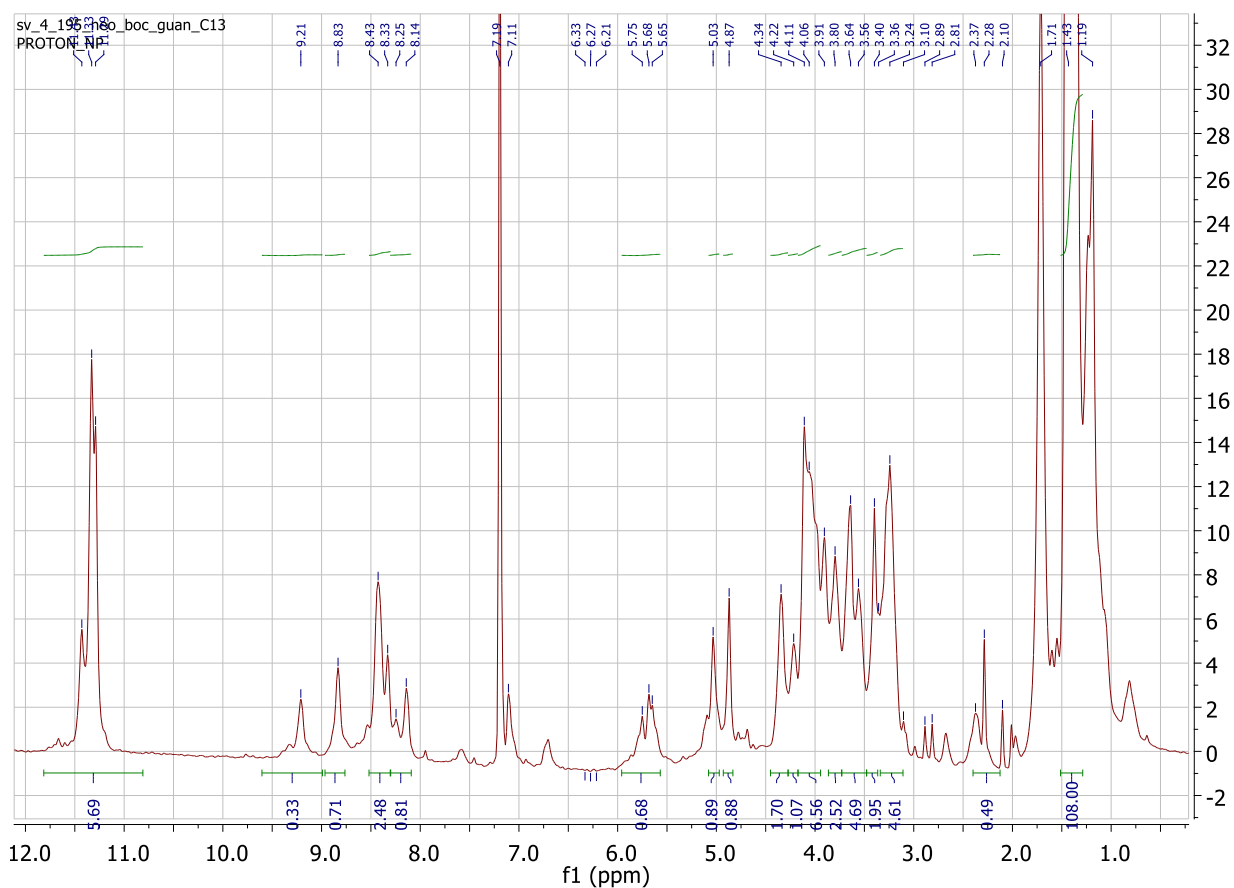


Figure S-1:  $^1\text{H}$  NMR of **12** in  $\text{CDCl}_3$ .

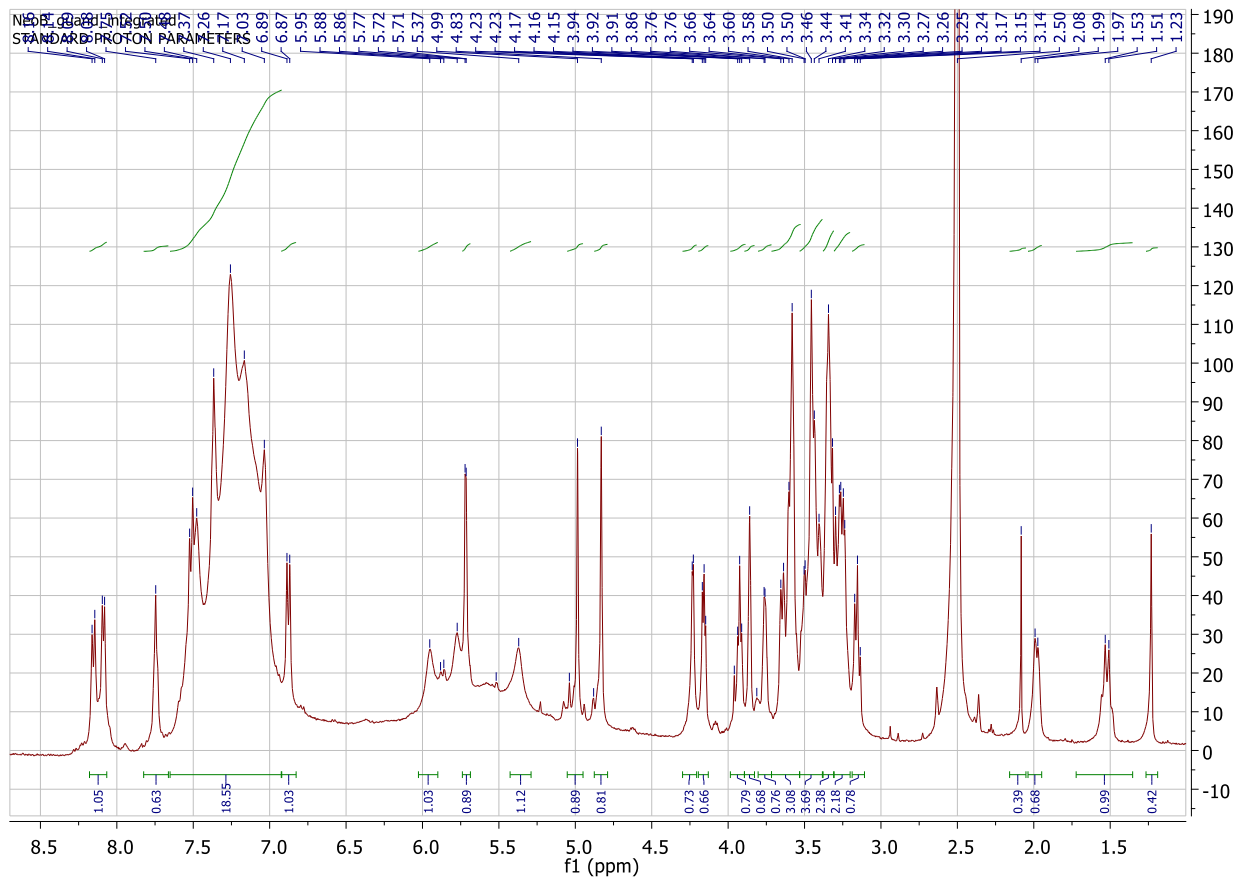


Figure S-2:  $^1\text{H}$  NMR of **13** in  $\text{d}_6\text{-DMSO}$ .



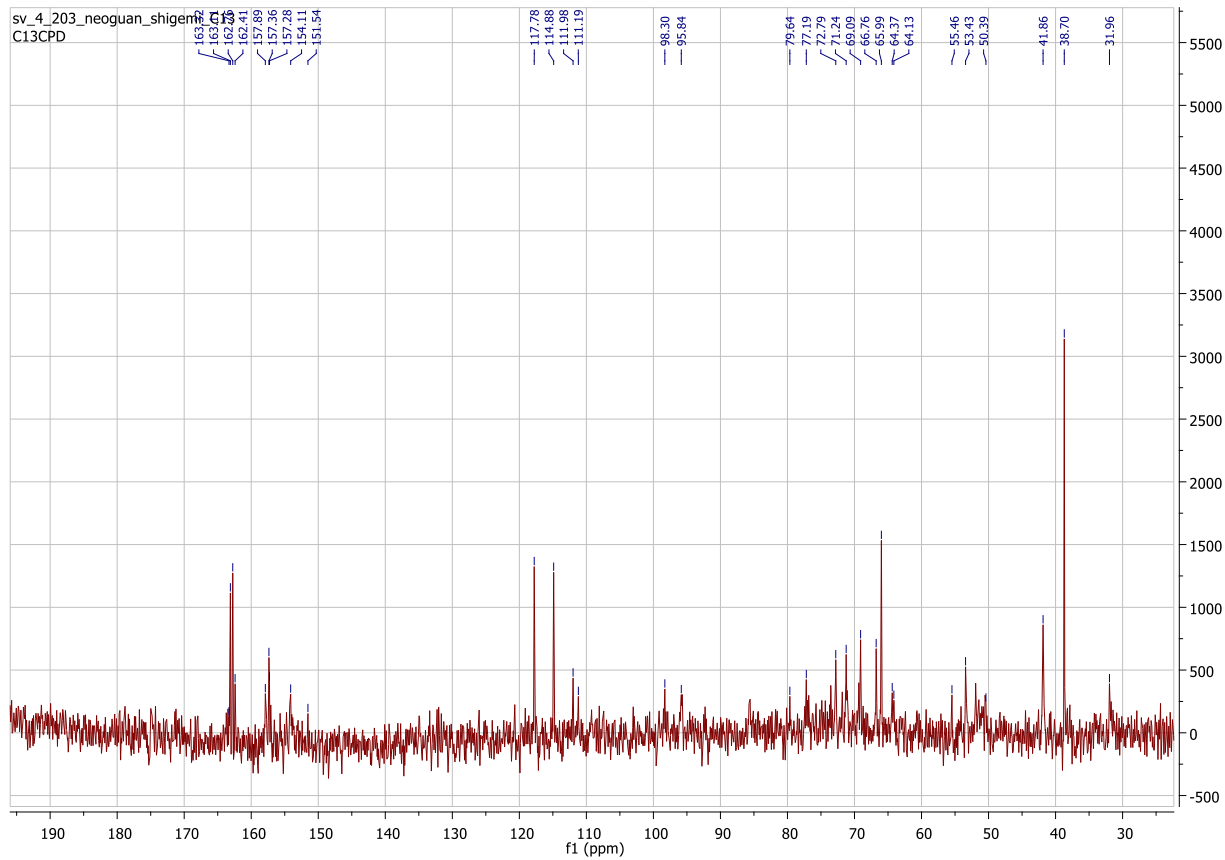


Figure S-3:  $^{13}\text{C}$  NMR of **13** in  $\text{D}_2\text{O}$ .

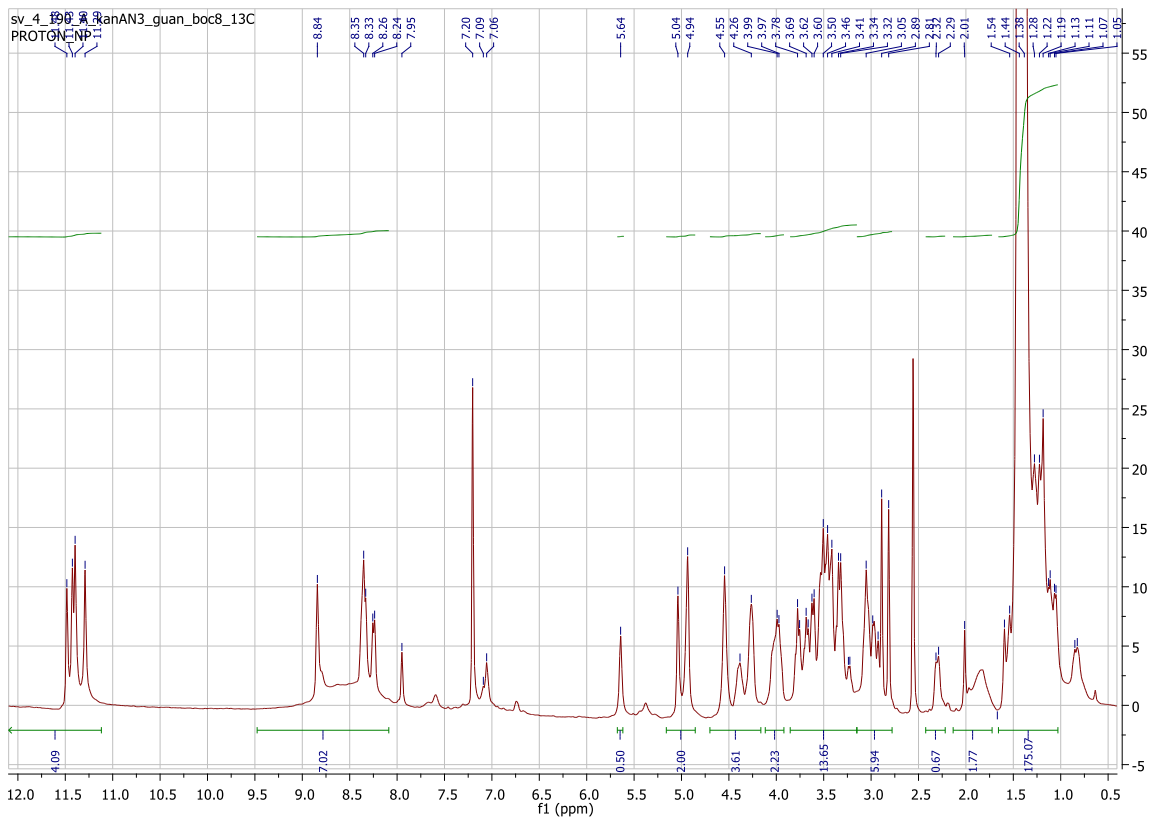


Figure S-4:  $^1\text{H}$  NMR of **15** in  $\text{CDCl}_3$ .

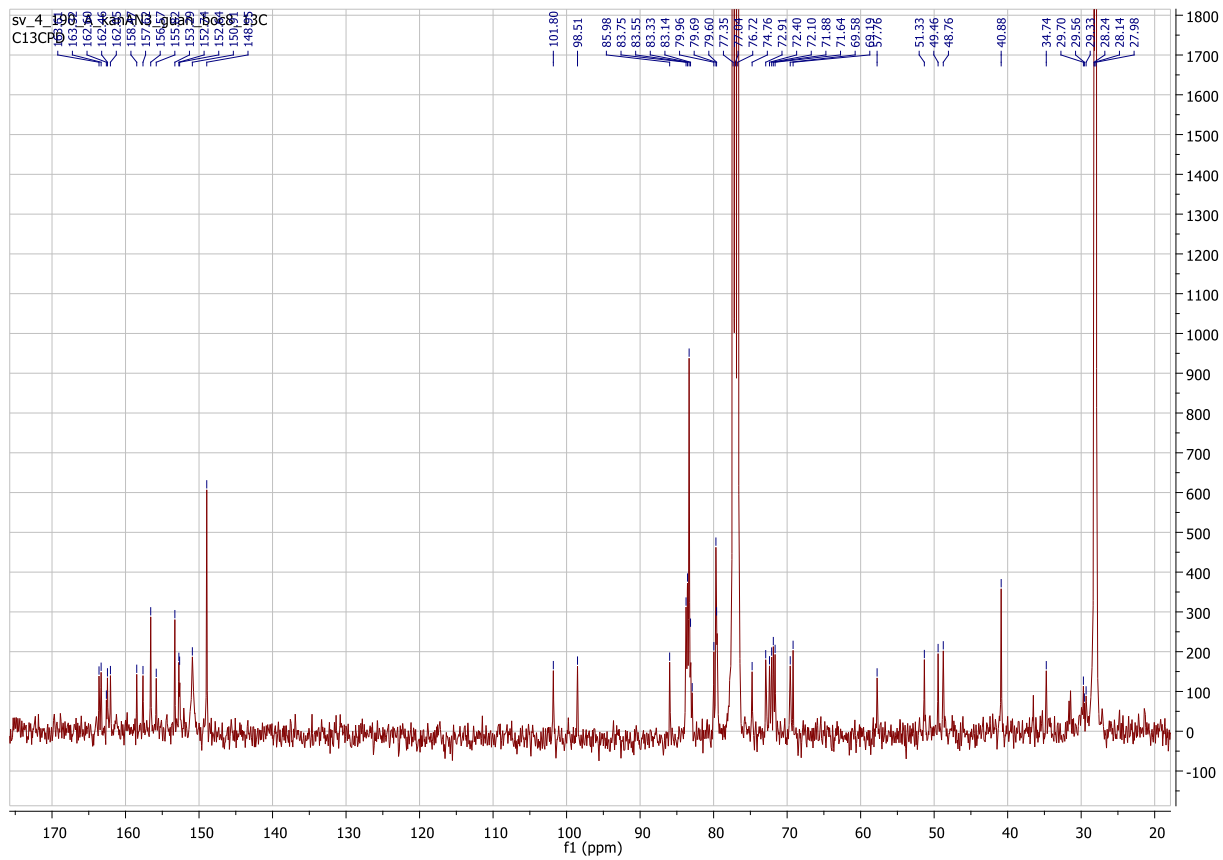


Figure S-5:  $^{13}\text{C}$  NMR of **15** in  $\text{CDCl}_3$ .

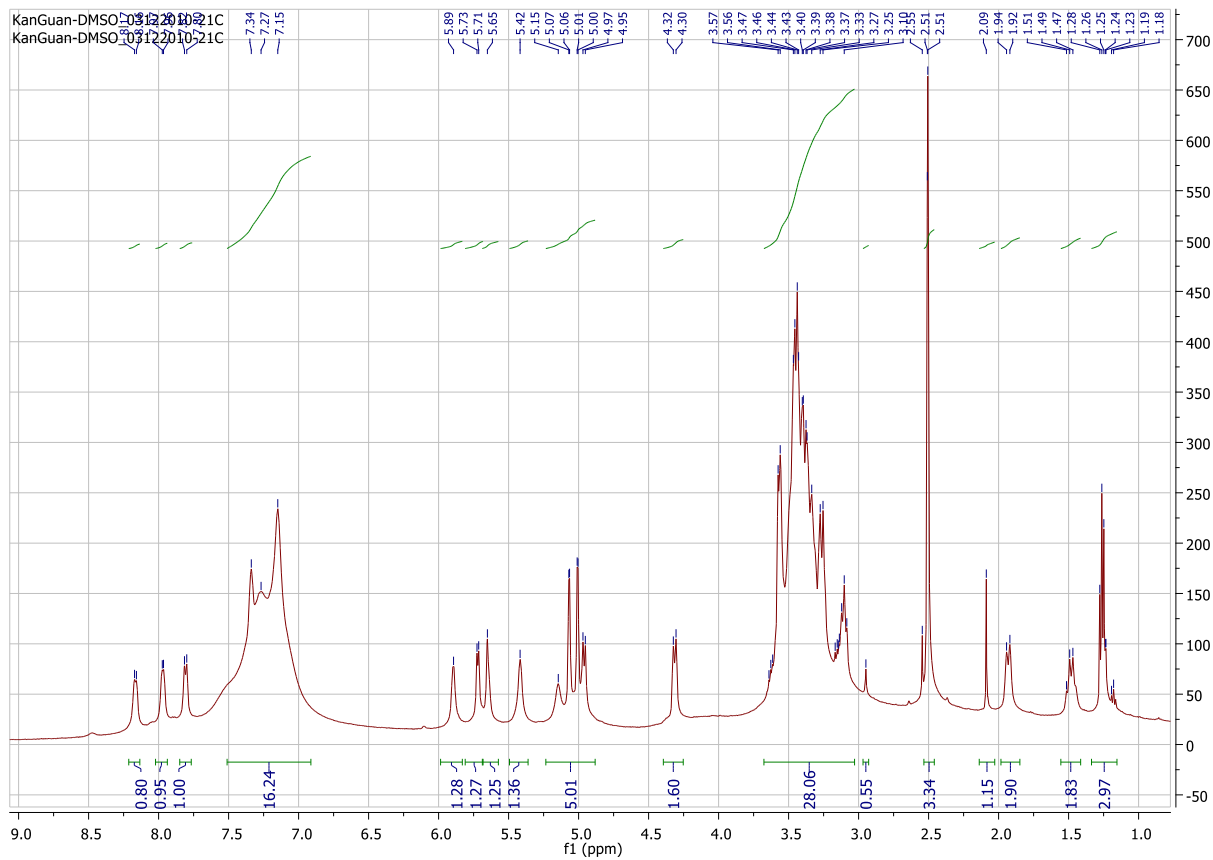


Figure S-6:  $^1\text{H}$  NMR of **16** in  $\text{d}_6\text{-DMSO}$ .

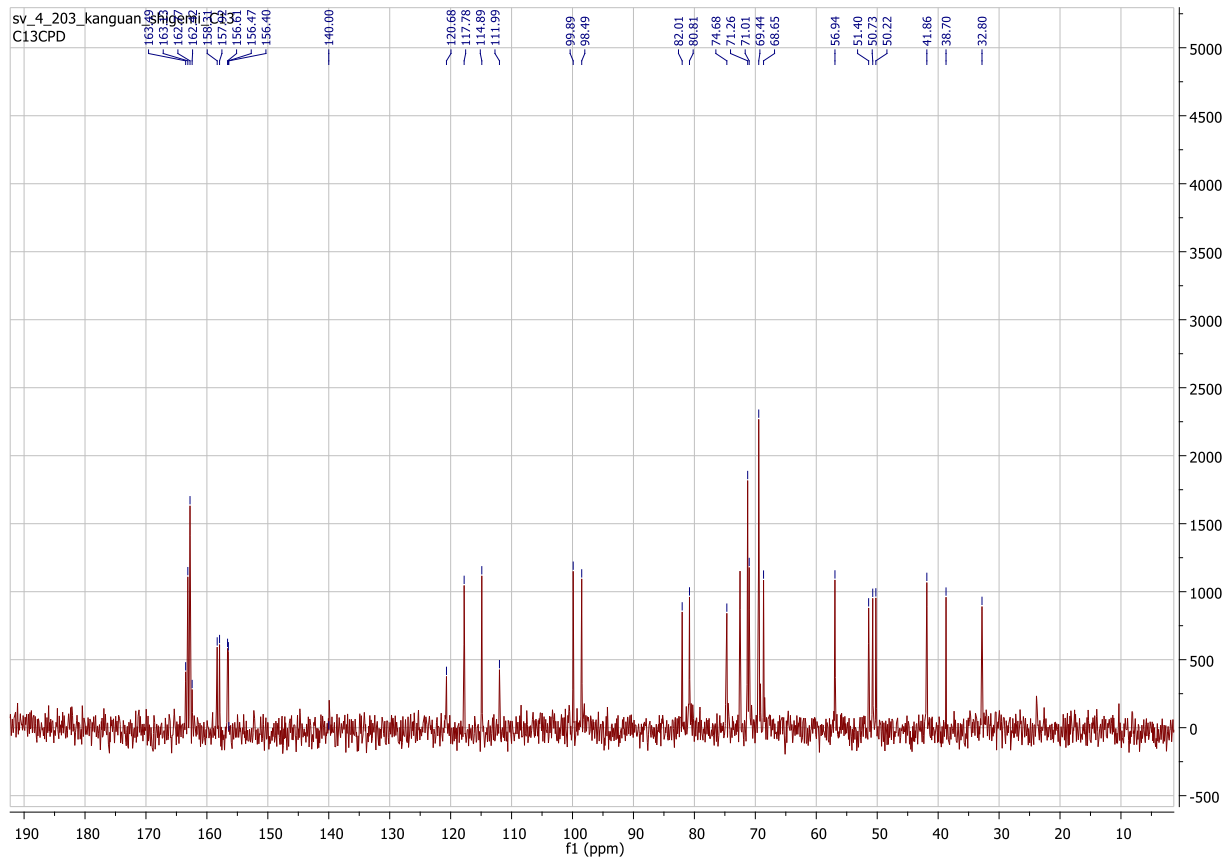
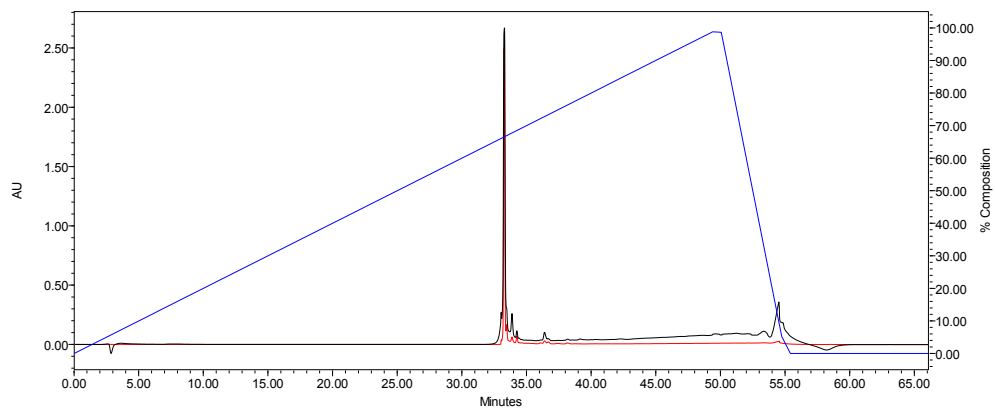
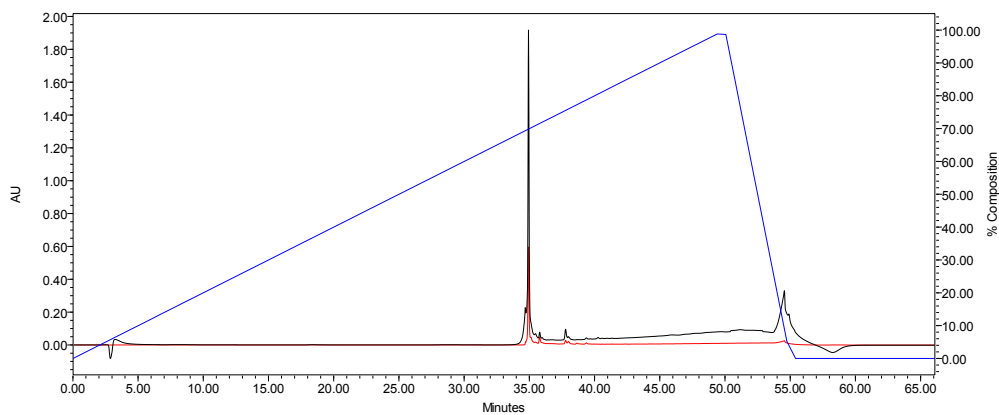


Figure S-7:  $^{13}\text{C}$  NMR of **16** in  $\text{D}_2\text{O}$ .

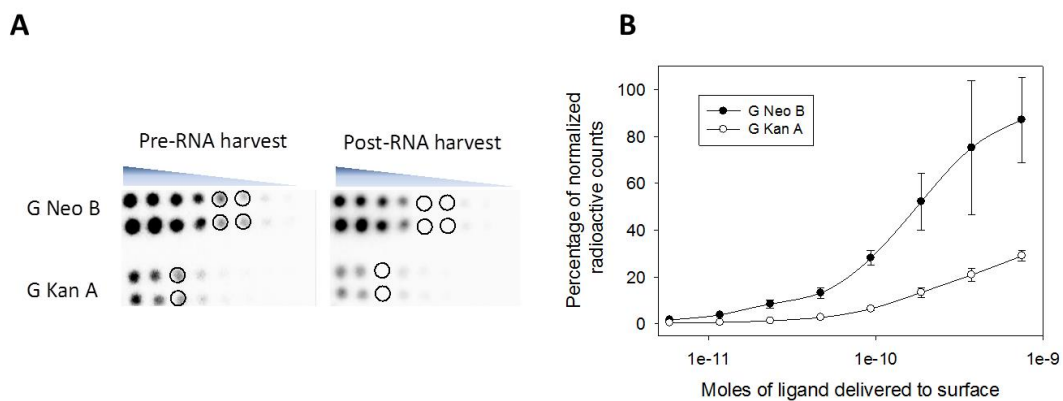


**Figure S-8:** Analytical HPLC chromatogram of **13-FI**



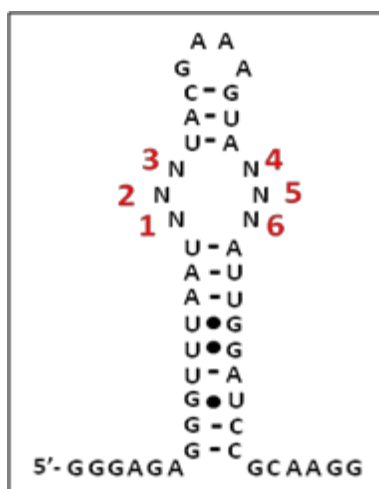
**Figure S-9:** Analytical HPLC chromatogram of **16-FI**.

### Section 3: 2DCS selection of RNA motifs that bind to G Neo B and G Kan A



**Figure S-10:** 2DCS selection of RNA motifs that bind to G Neo B and G Kan A. **A:** left, representative image of microarrays after hybridization with nucleic acids ( $^{32}\text{P}$ -labeled **1** and unlabeled **2-8**); right, image of microarray after harvesting RNA. Circles indicate positions that were excised. **B,** Plot of radioactive counts of  $^{32}\text{P}$ -labeled **1** bound as a function of ligand loading.

#### Section 4: Sequences of selected RNAs and their corresponding Sum Z-scores



**Table S-1:** Sequences of RNAs selected for G Neo B and the corresponding Sum Z-scores.

Sequence ID	1	2	3	4	5	6	$\Sigma$ Zscore
G Neo B IL1	A	G	A	A	U	C	57.68
G Neo B IL2	A	C	A	A	G	C	52.85
G Neo B IL3	C	U	C	C	C	U	51.20
G Neo B IL4	G	U	G	A	C	G	47.66
G Neo B IL5	A	C	A	C	C	U	45.38
G Neo B IL6	G	G	C	C	G	C	40.24
G Neo B IL7	G	G	C	U	A	G	38.16
G Neo B IL8	G	U	G	C	C	U	34.54
G Neo B IL9	G	U	C	U	G	G	33.47
G Neo B IL10	A	U	A	A	C	C	31.38
G Neo B IL11	G	G	C	U	C	G	29.82
G Neo B IL12	U	G	U	C	G	C	23.21
G Neo B IL13	A	G	G	A	G	A	17.86
G Neo B IL14	A	U	C	C	A	C	11.71
G Neo B IL15	C	A	A	C	G	G	5.85
G Neo B IL16	U	C	C	G	A	A	0.00



G Neo B IL17	A	G	A	A	U	C	57.68
G Neo B IL18	A	C	A	C	C	U	45.38
G Neo B IL19	C	G	C	C	C	U	44.61
G Neo B IL20	A	C	A	A	G	A	43.02
G Neo B IL21	A	C	A	A	G	U	42.13
G Neo B IL22	G	U	G	A	C	C	39.41
G Neo B IL23	G	U	G	C	C	G	38.14
G Neo B IL24	A	G	G	G	G	A	34.10
G Neo B IL25	G	G	C	C	U	C	30.30
G Neo B IL26	C	G	A	A	C	C	26.99
G Neo B IL27	G	U	C	U	G	U	26.98
G Neo B IL28	A	C	C	G	C	G	26.78
G Neo B IL29	A	C	C	G	C	G	26.78
G Neo B IL30	A	G	G	G	G	U	25.75
G Neo B IL31	C	G	C	G	C	U	24.15
G Neo B IL32	C	C	U	C	C	U	21.62
G Neo B IL33	G	G	U	C	A	G	21.59
G Neo B IL34	G	G	U	C	G	G	21.59
G Neo B IL35	U	G	G	G	U	C	21.31
G Neo B IL36	C	G	A	C	G	C	21.17
G Neo B IL37	C	U	G	A	C	U	20.45
G Neo B IL38	U	G	U	G	C	C	19.55
G Neo B IL39	U	G	G	A	A	C	17.95
G Neo B IL40	U	G	U	U	A	C	16.62
G Neo B IL41	G	G	C	G	A	C	16.52
G Neo B IL42	C	U	C	C	U	C	16.24
G Neo B IL43	C	U	C	C	A	G	15.80
G Neo B IL44	G	C	C	U	G	A	15.27

G Neo B IL45	U	G	G	G	A	G	14.20
G Neo B IL46	G	C	C	A	C	U	12.81
G Neo B IL47	A	A	G	G	G	C	12.01
G Neo B IL48	G	U	G	G	G	U	11.71
G Neo B IL49	C	G	G	G	A	G	11.28
G Neo B IL50	U	A	A	A	C	C	11.24
G Neo B IL51	C	U	C	A	G	U	10.84
G Neo B IL52	A	C	U	G	G	A	10.41
G Neo B IL53	G	C	G	U	U	G	10.31
G Neo B IL54	G	C	G	A	A	G	10.31
G Neo B IL55	A	A	C	C	A	C	9.22
G Neo B IL56	G	A	C	A	C	A	7.82
G Neo B IL57	C	C	C	C	G	G	6.29
G Neo B IL58	A	U	G	U	U	G	5.85
G Neo B IL59	G	A	U	U	G	C	5.42
G Neo B IL60	C	A	A	C	A	U	5.42
G Neo B IL61	U	A	U	G	A	C	5.42
G Neo B IL62	A	U	A	C	U	G	5.32
G Neo B IL63	G	A	A	G	G	G	5.32
G Neo B IL64	U	U	G	G	C	A	4.22
G Neo B IL65	A	C	G	U	C	A	3.36
G Neo B IL66	C	A	G	G	U	G	2.93
G Neo B IL67	G	C	U	U	C	C	2.93
G Neo B IL68	A	U	C	U	U	A	2.93
G Neo B IL69	G	A	A	G	U	C	2.93
G Neo B IL70	C	A	C	U	C	A	2.49
G Neo B IL71	C	C	G	U	A	G	2.49
G Neo B IL72	A	A	A	U	A	G	2.40

G Neo B IL73	A	A	C	A	U	G	0.00
G Neo B IL74	A	U	U	G	A	U	0.00
G Neo B IL75	G	U	A	A	A	A	0.00
G Neo B IL76	U	C	U	G	U	G	0.00
G Neo B IL77	U	A	U	C	U	U	0.00

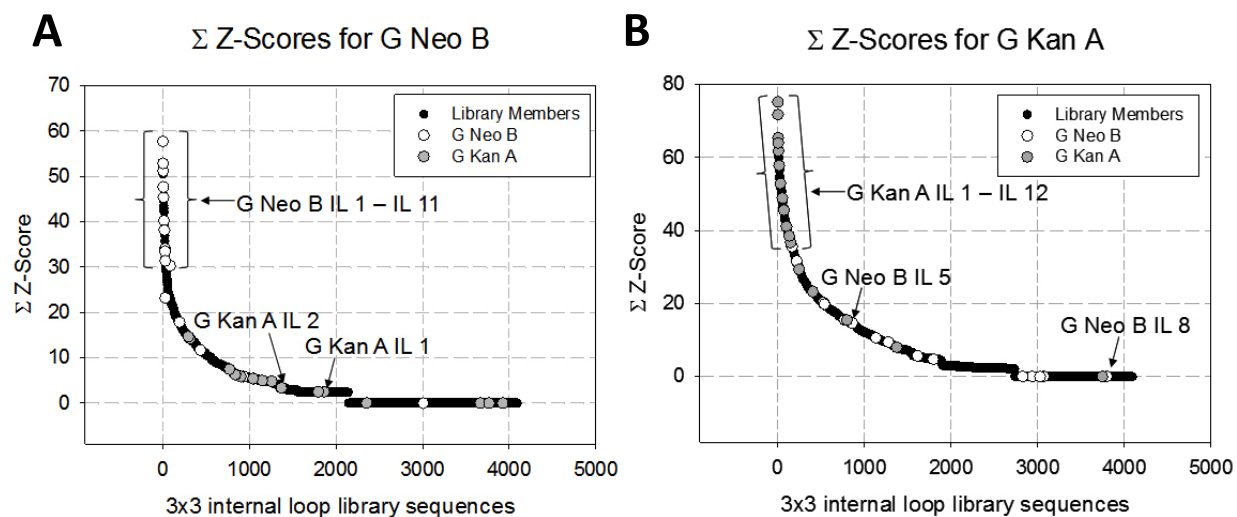
**Table S-2:** Sequences of RNAs selected for G Kan A and the corresponding Sum Z-scores.

Sequence ID	1	2	3	4	5	6	$\Sigma$ Zscore
G Kan A IL 1	C	C	C	U	A	U	75.20
G Kan A IL 2	C	A	C	C	G	A	71.73
G Kan A IL 3	G	A	C	G	A	G	65.36
G Kan A IL 4	C	U	C	A	C	A	63.97
G Kan A IL 5	C	G	C	C	A	U	61.78
G Kan A IL 6	C	A	C	A	C	C	57.82
G Kan A IL 7	C	G	C	A	G	U	52.84
G Kan A IL 8	C	A	C	G	C	G	48.86
G Kan A IL 9	C	G	C	C	C	G	45.54
G Kan A IL 10	C	G	C	A	U	U	41.10
G Kan A IL 11	A	G	C	U	G	U	38.50
G Kan A IL 12	U	U	C	C	C	C	36.69
G Kan A IL 13	A	C	U	A	U	G	29.36
G Kan A IL 14	A	C	C	U	C	A	23.25
G Kan A IL 15	A	C	U	G	C	A	15.47
G Kan A IL 16	C	C	G	U	G	C	7.99
G Kan A IL 17	G	U	U	A	G	C	0.00
G Kan A IL 18	C	A	C	A	G	C	58.30
G Kan A IL 19	U	C	C	C	C	C	53.50

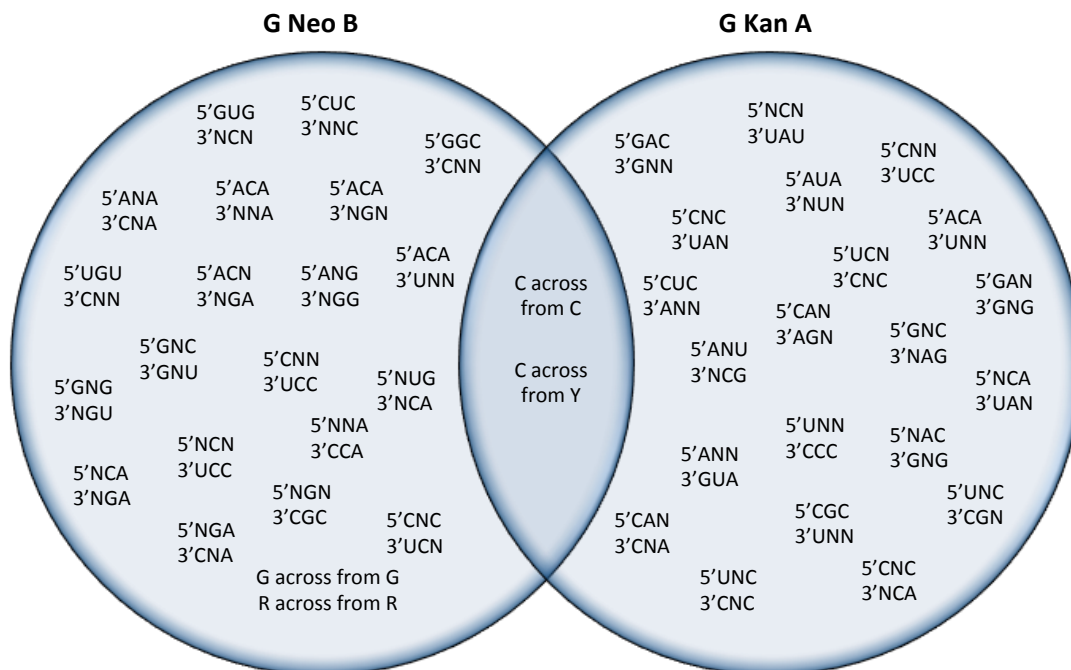
G Kan A IL 20	C	U	C	C	A	U	51.01
G Kan A IL 21	U	C	C	C	G	C	50.53
G Kan A IL 22	G	A	C	C	U	G	46.10
G Kan A IL 23	U	G	C	G	G	C	44.52
G Kan A IL 24	U	G	C	G	G	C	44.52
G Kan A IL 25	G	A	C	U	C	G	39.61
G Kan A IL 26	C	G	C	U	G	G	38.51
G Kan A IL 27	A	C	A	A	A	U	34.98
G Kan A IL 28	U	A	C	A	A	U	34.26
G Kan A IL 29	A	C	C	A	C	G	34.21
G Kan A IL 30	A	C	A	G	A	U	29.33
G Kan A IL 31	G	C	C	G	A	A	28.40
G Kan A IL 32	A	A	U	G	C	C	28.12
G Kan A IL 33	A	A	U	G	C	C	28.12
G Kan A IL 34	A	U	A	U	U	G	28.09
G Kan A IL 35	C	U	C	U	U	A	27.97
G Kan A IL 36	A	U	A	A	U	G	26.28
G Kan A IL 37	G	G	C	C	A	C	25.49
G Kan A IL 38	G	C	C	A	U	A	24.61
G Kan A IL 39	G	C	A	U	A	U	23.77
G Kan A IL 40	G	A	G	C	U	G	21.25
G Kan A IL 41	U	G	C	C	U	G	20.76
G Kan A IL 42	U	C	U	C	C	C	20.66
G Kan A IL 43	G	A	G	G	C	G	20.37
G Kan A IL 44	A	C	A	U	C	U	20.30
G Kan A IL 45	A	U	A	U	U	U	19.02
G Kan A IL 46	C	A	A	G	G	A	18.76
G Kan A IL 47	A	U	C	A	G	C	18.60

G Kan A IL 48	U	U	C	A	A	G	18.59
G Kan A IL 49	A	G	G	A	U	G	18.15
G Kan A IL 50	C	A	G	G	A	U	17.13
G Kan A IL 51	C	A	A	U	U	A	17.12
G Kan A IL 52	C	A	U	C	U	A	16.14
G Kan A IL 53	G	U	C	C	G	U	15.56
G Kan A IL 54	U	A	U	A	A	U	14.77
G Kan A IL 55	A	C	U	A	A	A	13.83
G Kan A IL 56	A	C	G	U	U	G	13.38
G Kan A IL 57	C	A	G	A	A	C	12.63
G Kan A IL 58	G	G	C	U	U	A	11.53
G Kan A IL 59	C	C	A	U	U	C	10.61
G Kan A IL 60	C	U	G	A	U	G	9.69
G Kan A IL 61	G	C	U	C	U	A	9.58
G Kan A IL 62	G	C	G	G	G	A	9.58
G Kan A IL 63	G	C	G	C	C	A	9.58
G Kan A IL 64	G	A	U	U	A	U	8.29
G Kan A IL 65	C	U	A	C	A	C	5.37
G Kan A IL 66	C	U	U	U	C	C	4.96
G Kan A IL 67	G	C	G	A	U	C	3.03
G Kan A IL 68	C	C	U	G	U	U	2.34
G Kan A IL 69	A	G	U	C	U	C	0.00
G Kan A IL 70	A	A	A	C	G	C	0.00

## Sum Z-scores of selected RNAs and all members of 1



**Figure S-11:** Sum Z-scores for all members of the RNA library **1** for binding G Neo B and G Kan A (Figure 1). **A**, RNAs selected to bind G Neo B are shown in open circles (high Sum Z-scores for G Neo B). Loops selected to bind G Kan A are shown in grey (low Sum Z-scores for G Neo B). The remaining library members are depicted in black. The affinities of G Kan A IL 2 and G Kan A IL 1 for G Neo B were measured in order to study selectivity. Both RNAs bind tightly to G Kan A and weakly to G Neo B, as expected based on their Sum Z-scores. **B**, Sum Z-scores for all members of the RNA library **1** for binding G Kan A (Figure 1). RNAs selected to bind G Kan A are shown in grey (high Sum Z-scores for G Kan A). Loops selected to bind G Neo B are shown in open circles (low Sum Z-scores for G Kan A). The remaining library members are depicted in black. The affinities of G Neo B IL5 and G Neo B IL8 for G Kan A were measured in order to study selectivity. Both RNAs bind tightly to G Neo B and weakly to G Kan A, as expected based on their Sum Z-scores.



**Figure S-12:** Venn diagrams that illustrate the RNA motif space that imparts binding affinity for each aminoglycoside. Only trends with >99% confidence levels are shown. Each aminoglycoside has unique RNA motif binding space. The nucleotide adjacent to the “5” label is nucleotide 1 and the nucleotide adjacent to the “3” label is nucleotide 6 as indicated in the schematic on page S-16.

## Section 5: Determination of binding affinity using fluorescence-based binding assays

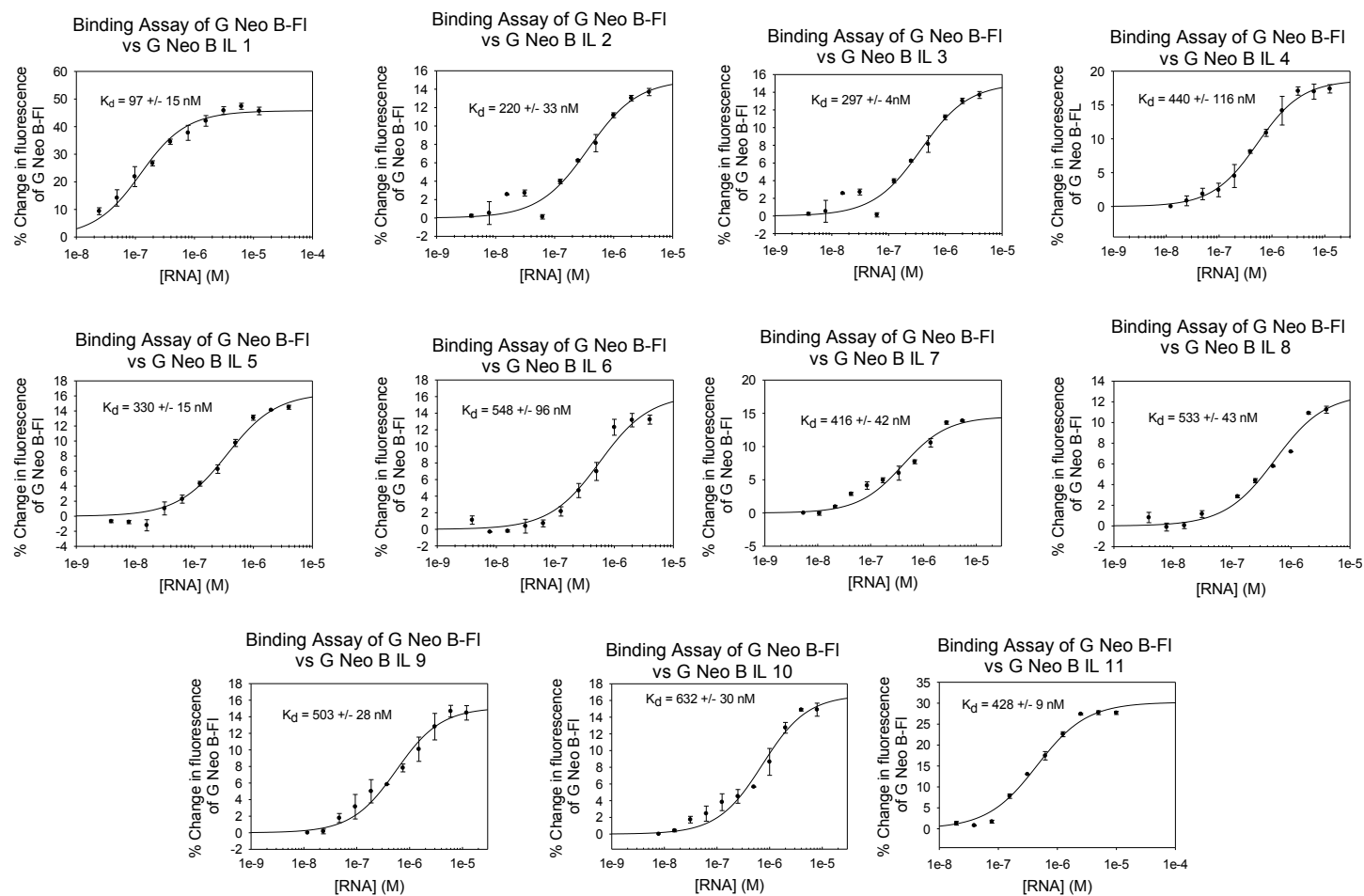
The affinities of RNAs for G Neo B-FI (**13-FI**) or G Kan A-FI (**16-FI**) were determined using an in solution, fluorescence-based assay.<sup>1</sup> A selected RNA was folded in 1× Assay Buffer (20 mM Hepes, 150 mM NaCl, 5 mM KCl, 1 mM MgCl<sub>2</sub>, and 40 µg/mL BSA) by heating at 60°C for 5 min followed by slowly cooling to room temperature. G Neo B-FI or G Kan A-FI was then added to a final concentration of 50 nM. Serial dilutions (1:2) were completed in 1× Assay Buffer supplemented with 50 nM of G Neo B-FI or G Kan A-FI. The solutions were incubated for 30 min at room temperature and then transferred to a well of 96-well black plate. Fluorescence intensity was measured on a Bio-Tek FLx-800 plate reader. The change in fluorescence intensity as a function of RNA concentration was fit to the following equation<sup>7</sup>:

$$I = I_0 + 0.5\Delta\varepsilon \left( ([FL]_0 + [RNA]_0 + K_t) - \left( ([FL]_0 + [RNA]_0 + K_t)^2 - 4[FL]_0[RNA]_0 \right)^{0.5} \right)$$

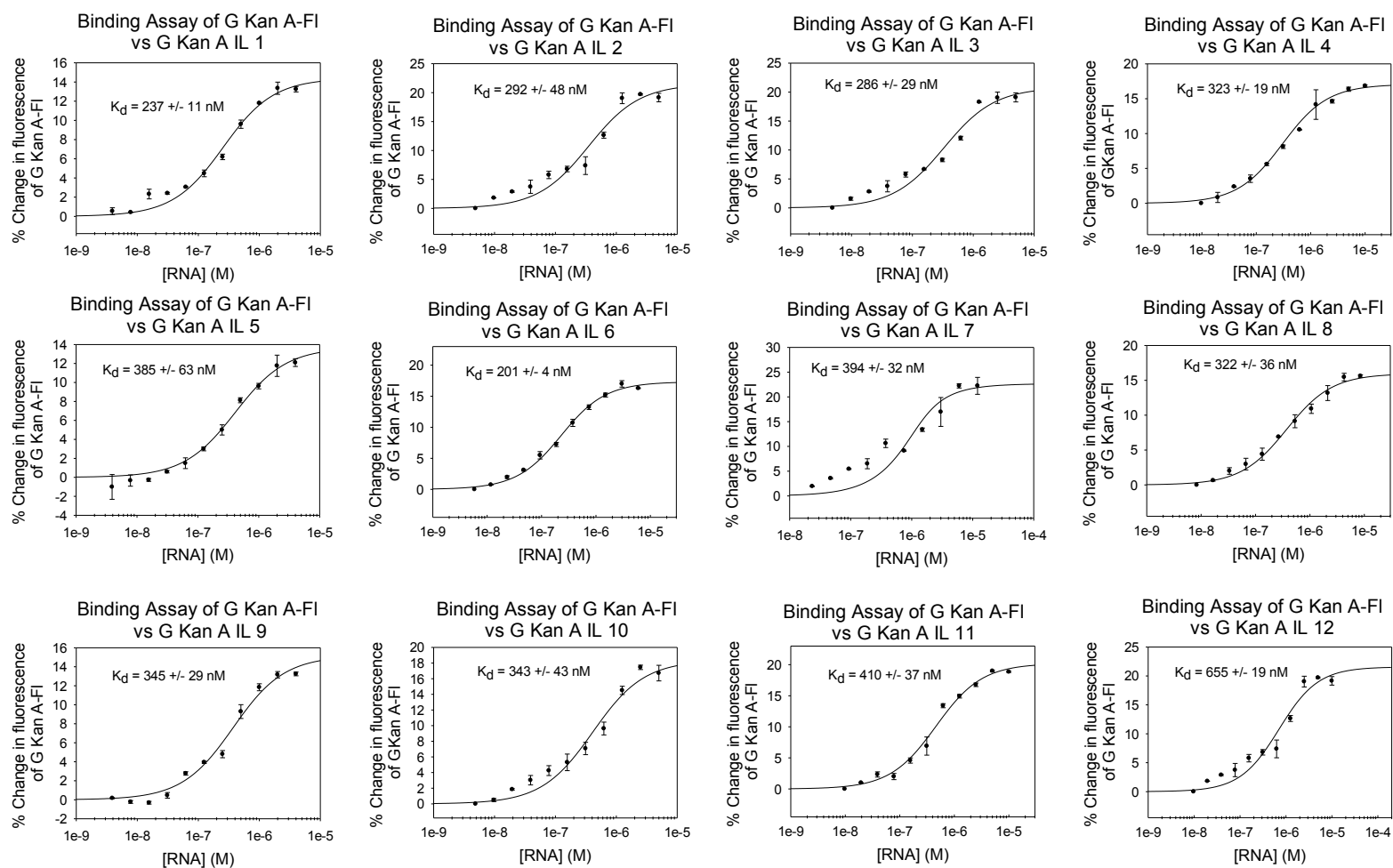
where  $I$  and  $I_0$  are the observed fluorescence intensity in the presence and absence of RNA respectively,  $\Delta\varepsilon$  is the difference between the fluorescence intensity in the absence of RNA and in the presence of infinite RNA concentration,  $[FL]_0$  is the concentration of G Neo B-FI or G Kan A-FI,  $[RNA]_0$  is the concentration of RNA, and  $K_t$  is the dissociation constant.



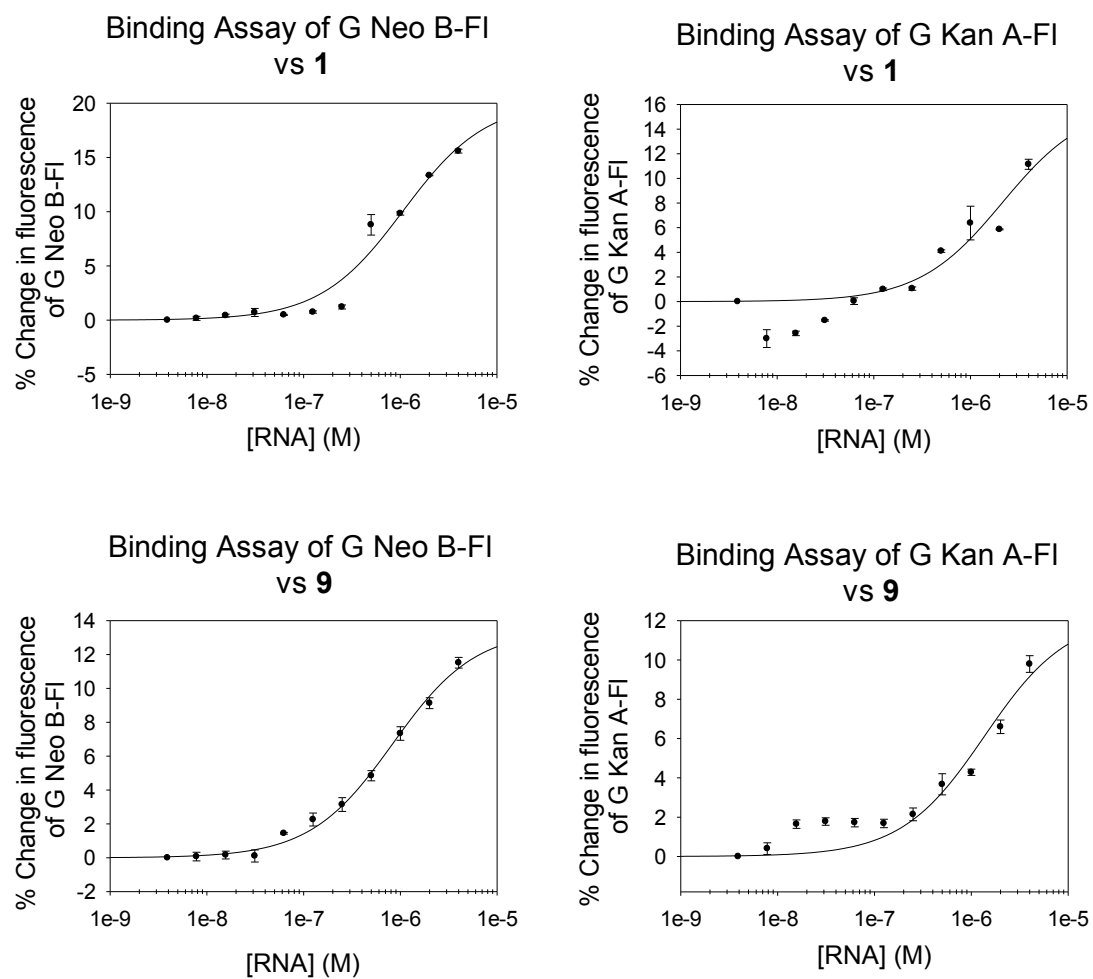
## Section 6: Representative binding curves



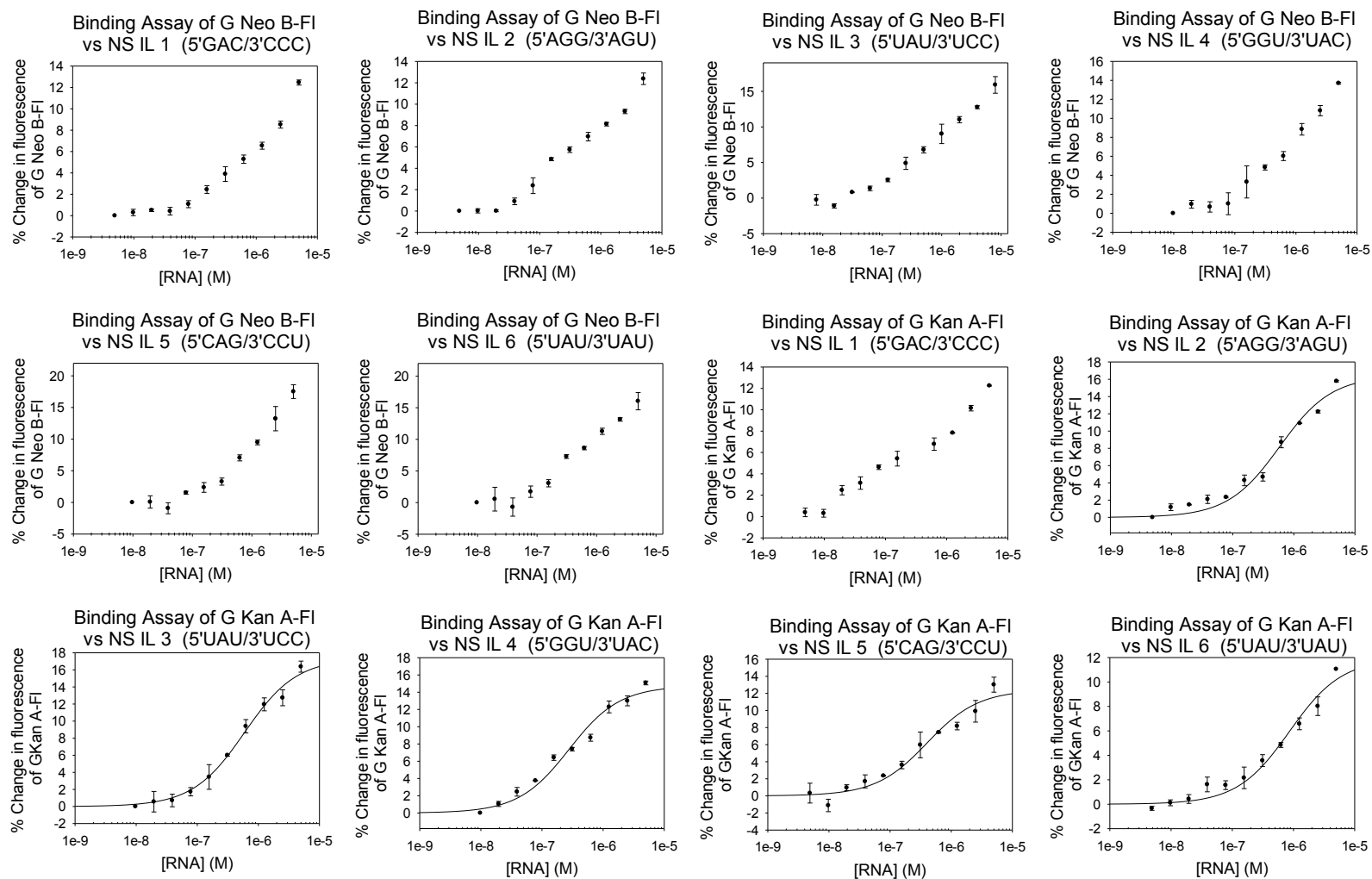
**Figure S-13:** Representative binding curves for a subset of RNAs selected to bind G Neo B.



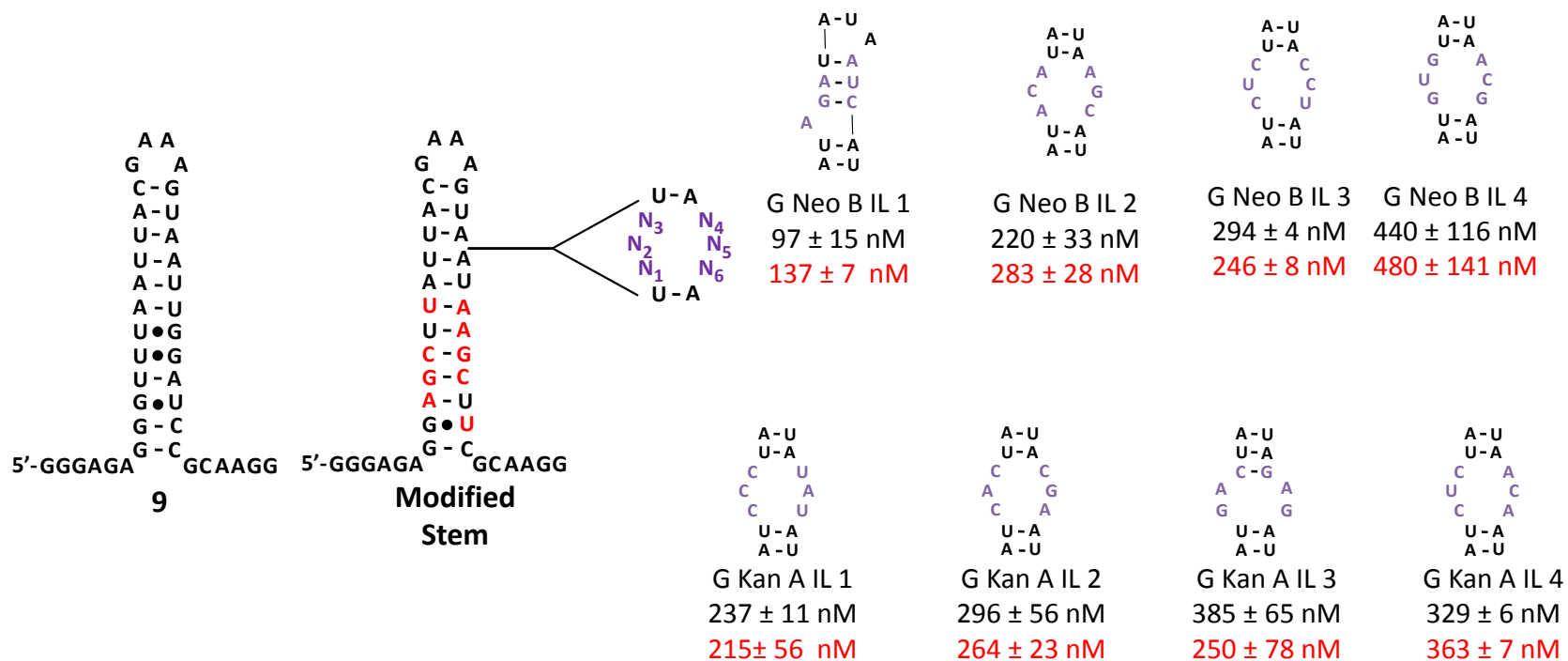
**Figure S-14:** Representative binding curves for a subset of RNAs selected to bind G Kan A.



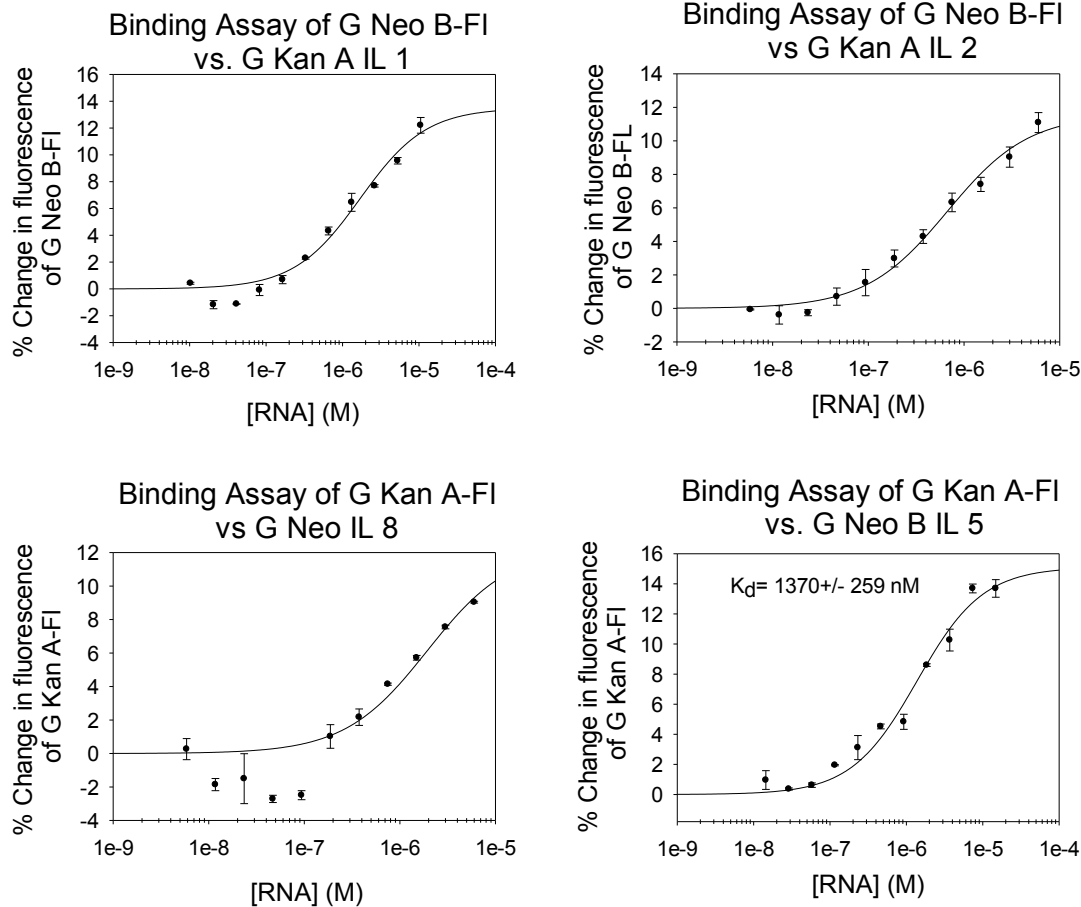
**Figure S-15:** Representative curves for the binding of **1** and **9** to G Neo B and G Kan A. Both compounds bind weakly to the entire RNA library (**1**) and the cassette into which the randomized region was embedded (**9**).



**Figure S-16:** Representative binding curves for a subset of RNAs not selected to bind G Neo B and G Kan A and that have low Sum Z-scores for both compounds.

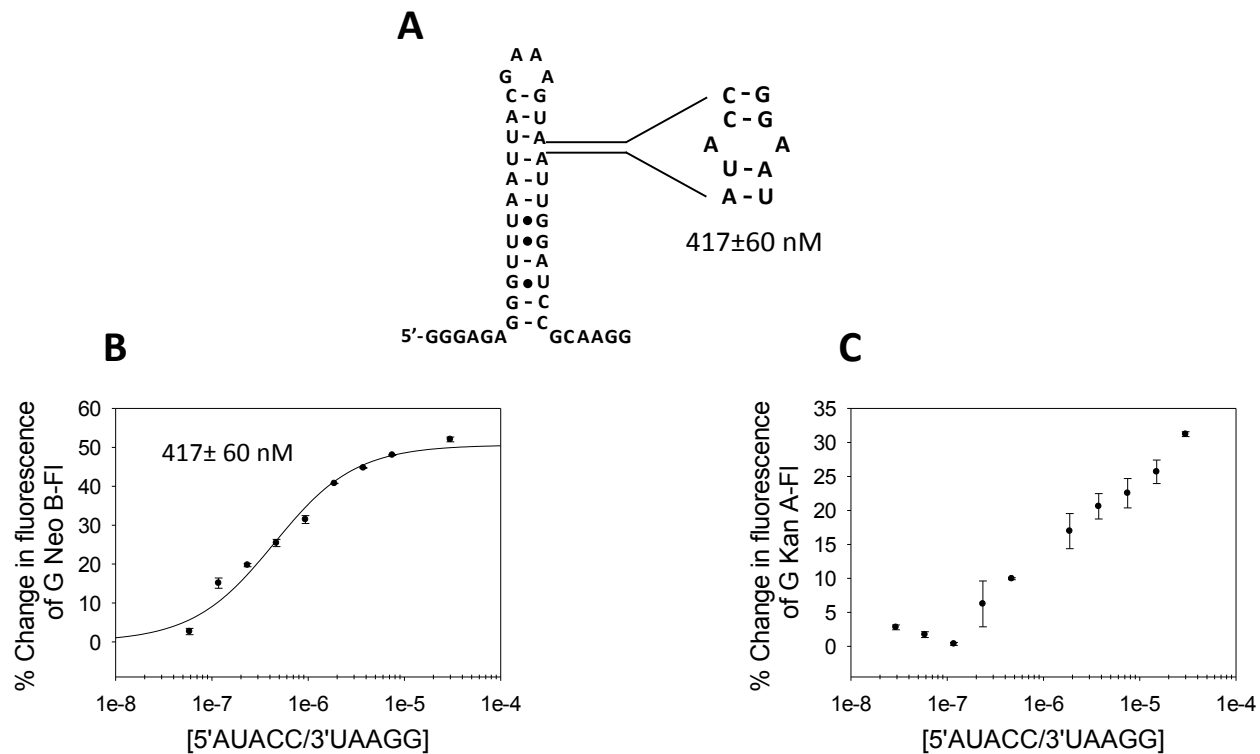


**Figure S-17:** Secondary structures of cassettes used to display internal loop nucleotides. The loop identifier is indicated below the loop secondary structure. The affinity of the loop displayed in **9** is reported in black while the affinity of the loop displayed in the modified stem is reported in red.



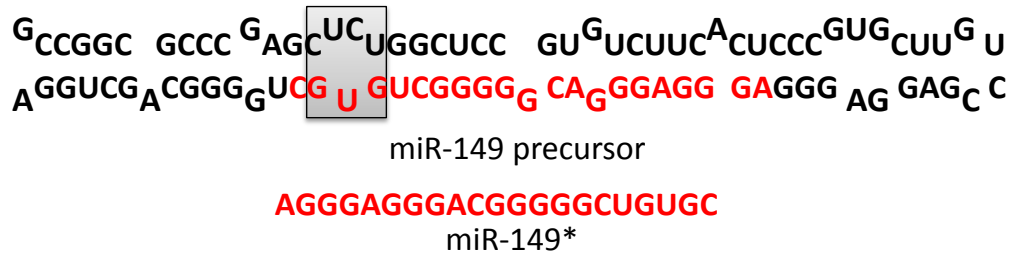
**Figure S-18:** Representative binding curves to determine the affinities of two RNAs selected by G Neo B (G Neo IL 5 and G Neo IL 8) for G Kan A and two RNAs selected by G Kan A (G Kan IL1 and G Kan IL2) for G Neo B. As expected, RNAs bind weakly to their non-cognate ligand; that is, they are selective for the small molecule for which they were selected.

**Section 7: Representative Binding Curve for G Neo B and 5'AUACC/3'UAAGG, an internal loop found in miR-10b**



**Figure S-19:** G Neo B, but not G Kan A, binds the internal loop in miR-10b as predicted by 2DCS and StARTS analysis. **A**, secondary structure of the RNA that contains 5'AUACC/3'UAAGG from miR-10b embedded in **9** (Figure 1). **B**, Representative binding curve of G Neo B-FI and 5'AUACC/3'UAAGG ( $K_d = 417 \pm 60$  nM). **C**, Representative binding curve of G Kan A-FI and 5'AUACC/3'UAAGG.

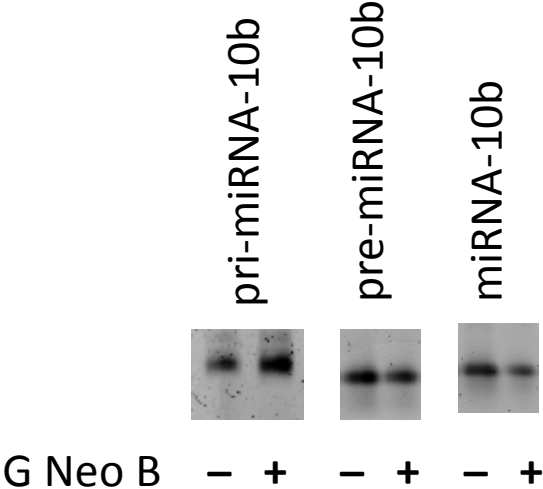
**Section 8: Secondary structure of the control precursor miRNA, miR-149**



**Figure S-20:** Secondary structure of miR-149 precursor and sequence of miR-149\*. miR-149 was used as a control to test specificity of G Neo B towards miR-10b. G Neo B is not predicted to bind the Drosha or Dicer sites in the miR-149 precursor. MiR-149\* is not predicted to bind the 3' UTR of miR-10b's downstream target (*HOXD10*).



Section 9: G Neo B affects microRNA-10b biogenesis



**Figure S-21:** Effect of G Neo B (100  $\mu$ M dosage) on pri-, pre- and mature miR-10b levels in cells in which the miRNA is overexpressed. qRT-PCR products were verified by agarose gel electrophoresis, stained with ethidium bromide.

## Section 10: References

1. J. L. Childs-Disney, M. Wu, A. Pushechnikov, O. Aminova and M. D. Disney, *ACS Chem. Biol.*, 2007, **2**, 745-754.
2. K. Feichtinger, C. Zapf, H. L. Sings and M. Goodman, *J. Org. Chem.*, 1998, **63**, 3804-3805.
3. T. J. Baker, N. W. Luedtke, Y. Tor and M. Goodman, *J. Org. Chem.*, 2000, **65**, 9054-9058.
4. D. W. Staple, V. Venditti, N. Niccolai, L. Elson-Schwab, Y. Tor and S. E. Butcher, *ChemBioChem*, 2008, **9**, 93-102.
5. S. P. Velagapudi, S. J. Seedhouse, J. French and M. D. Disney, *J. Am. Chem. Soc.*, 2011, **133**, 10111-10118.
6. T. Durroux, M. Peter, G. Turcatti, A. Chollet, M.-N. Balestre, C. Barberis and R. Seyer, *J. Med. Chem.*, 1999, **42**, 1312-1319.
7. M. D. Disney, L. P. Labuda, D. J. Paul, S. G. Poplawski, A. Pushechnikov, T. Tran, S. P. Velagapudi, M. Wu and J. L. Childs-Disney, *J. Am. Chem. Soc.*, 2008, **130**, 11185-11194.