Expanding the Known Structure Space for RNA Binding: A Test of 2,5-Diketopiperazine

Diego M. Arévalo,[1] Viktoriya S. Anokhina,[2] Oliver L. R. Swart,[1] Benjamin L. Miller[2,3,*]

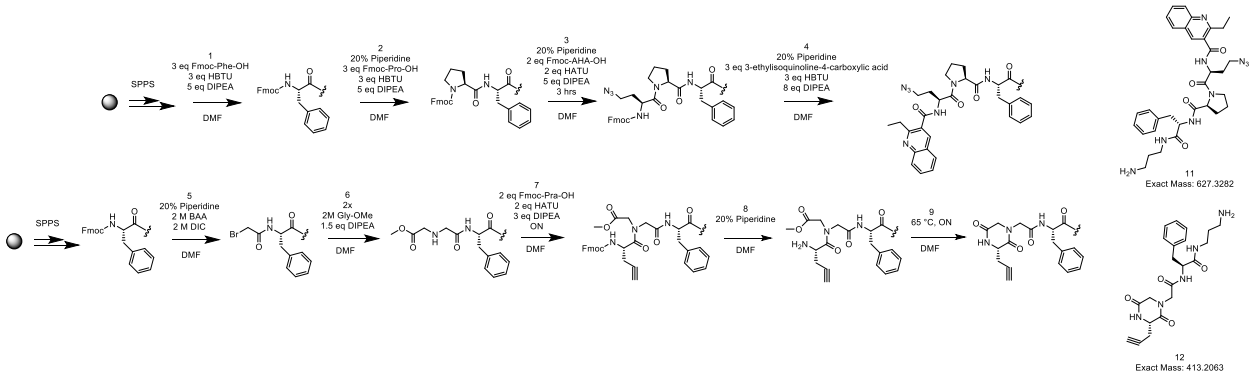
Departments of [1]Chemistry, [2]Biochemistry and Biophysics, and [3]Dermatology, University of Rochester, Rochester, NY 14642, USA

*Corresponding Author

Benjamin_Miller@URMC.Rochester.edu

Supplementary information

1. Synthesis

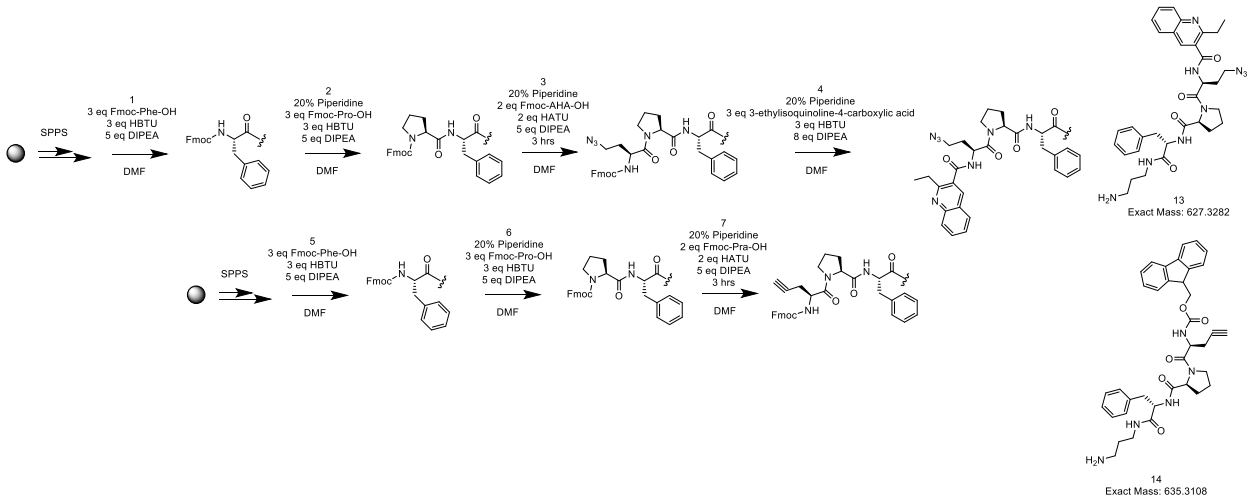


Scheme 1. Synthesis of the azide (compound **11**) and alkyne (compound **12**) precursors for DKP compound **(8)**.

Azide and alkyne-bearing precursors to **7** and **8** were synthesized by solid phase peptide synthesis (SPPS), and followed similar procedures, unless noted otherwise. Scheme 1 and 2 show the synthesis of monomers for the DKP (**8**) and MQ (**7**) compounds, respectively. In the following procedures, "agitated in a reaction vessel" means the synthesis resin was placed in a 10 mL Poly-Prep® Chromatography Column (Bio-Rad Laboratories, Inc) attached to a LabQuake laboratory rotator, and allowed to rotate at room temperature.

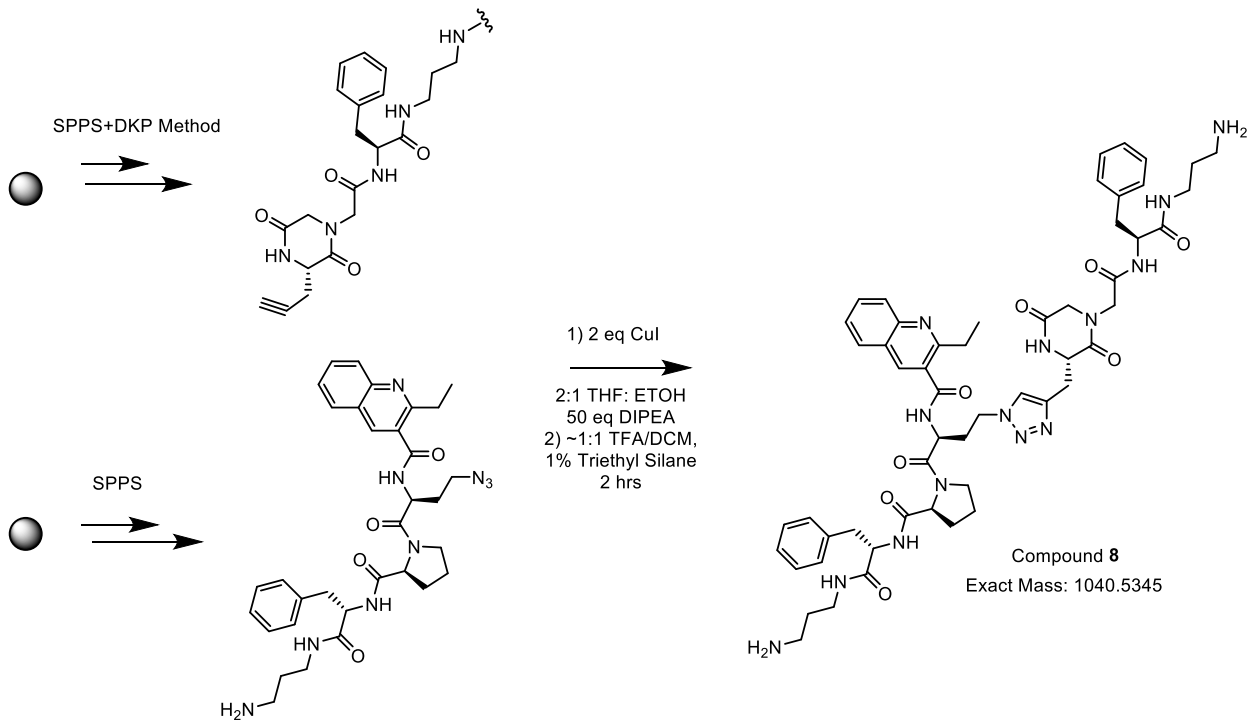
Wang resin (0.96 mmol/g) was activated with 1,1'-carbonyl-di-imidizole (CDI, 10 eq) in dimethylformamide (DMF) for 6 hours, and then 1,3-diaminopropane (10 eq) was added and allowed to react for an additional 6 hours. The conditions used for the remainder of standard SPPS steps were as follows: for each coupling reaction 3 eq Fmoc- Phe or Pro, 3 eq of O-(Benzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate (HBTU), and 5 eq of N,N-Diisopropylethylamine (DIPEA) were agitated in a reaction vessel for 2 hours. For L-azidohomoalanine (AHA) coupling for the azide monomer (**11**), 2 eq of AHA, 2 eq O-(7-Azabenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate (HATU) and 5 eq DIPEA were agitated for 3 hours at room temperature, as shown in step 3. Deprotection of Fmoc protected amino acids on resin beads was performed by 2 additions of 20% piperidine in DMF for 20 minutes each. In order to cap the azide-containing monomer, 3 eq of 3-ethylisoquinoline-4-carboxylic acid, 3 eq HBTU and 8 eq DIPEA were added to the resin and allowed to react for 2 hours of agitation. The conditions for incorporation of the alkyne residue diverge for both the DKP and MQ compounds. For incorporation of the alkyne residue on the DKP containing monomer (**12**), conditions for the synthesis of the DKP begin after coupling and deprotection of the Phe residue on Wang resin by first activating 2 M Bromoacetic acid (BAA) with 2 M N,N'-Diisopropylcarbodiimide (DIC) in DMF. To accomplish this, reagents were mixed in a syringe

fitted with a Kimwipe being used as a filter at the outlet. This step is represented in step 5, Scheme 1. This caused the immediate exothermic formation of urea. After allowing the reaction to cool to room temp, the plunger of the syringe was used to push the activated BAA solution through the Kimwipe (filtering out the urea precipiatate) and into a 10 mL microwave (MW) tube. The Wang resin bearing the deprotected Phe residue was dried and transferred to the MW tube, and the BAA coupling was then facilitated by 2 x 10 min MW reactions (45 ˚C, 20 W, Scheme 1, step 5). Next, 2 M glycine methyl ester (Gly-OMe), 2 M DIPEA in DMF was coupled via 2 x 10 min microwave reactions (65 ˚C, 20 W, Scheme 1, step 6). This produced the glycine peptoid intermediate. Next, 2 eq L-Propargylglycine (Pra) was activated with 2 eq HATU and 5 eq DIPEA, and the resin was agitated for 12 hours. This allowed the secondary amine from the peptoid to act as a nucleophile (Scheme 1, step 7). After washing the resin with 3x 8 mL DCM / 8 mL DMF, the fmoc protecting group on the Pra residue was deprotected via 20% piperidine in DMF for 2x 20 min. After deprotection and washing, the resin was heated in DMF at 50-60 °C overnight to obtain the DKP monomer (Scheme 1, Compound **12**).



Scheme 2. Synthesis of half compounds **13** and **14** for monoquinoline compound **7**. Note that the alkyne compound retains the Fmoc protecting group during the following click reaction.

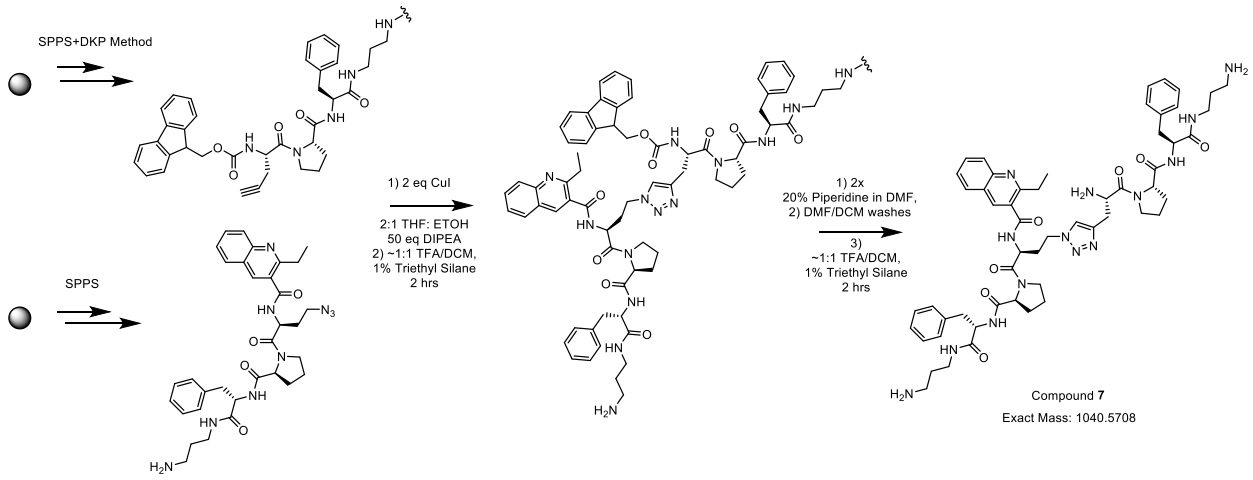
The reaction conditions for coupling the alkyne residue on the Wang resin for the MQ compound followed a standard SPPS procedure: 2 eq Pra, 2 eq HATU and 5 eq DIPEA were agitated for 3 hours. After coupling the Pra residue, the Wang resin was washed with 3x 8 mL DMF/ 8 mL DCM and the Fmoc protecting group was left on the Pra residue for the following click reaction.



Scheme 3. Huisgen dipolar cycloaddition (click) reaction conditions to yield DKP compound **8**.

Once both monomers for the DKP compound were prepared, the resin with the azide containing monomer (**11**) was cleaved by exposure to a combination of 50% trifluoroacetic acid (TFA) with 1% triethyl silane as a scavenger in DCM on a shaker for 2 hours. The azide containing monomer (**11**) was dried via rotary evaporation followed by final drying under laboratory vacuum. In order to form the 1,4 substituted triazole, 2 eq of copper (I) iodide (CuI)

was dissolved in 2:1 THF:EtOH (≤5 mL) and added to a 10 mL round bottom (RB) flask. Dried resin carrying the alkyne monomer was added to the RB followed by a solution of the dissolved azide-containing monomer **11** in 2:1 THF:EtOH to the RB. Finally, 50 eq DIPEA were added to the RB. After allowing the mixture to stir slowly on a magnetic stir plate for 20 min, the reaction was heated to 40 ˚C for overnight. The resin was then washed 3 times with DMF and 3 times with DCM and then the product was cleaved from the resin with 50% trifluoroacetic acid (TFA), 1% triethyl silane in DCM, as shown in Scheme 3. The precipitate was dried under vacuum, yielding the DKP compound (**8**).



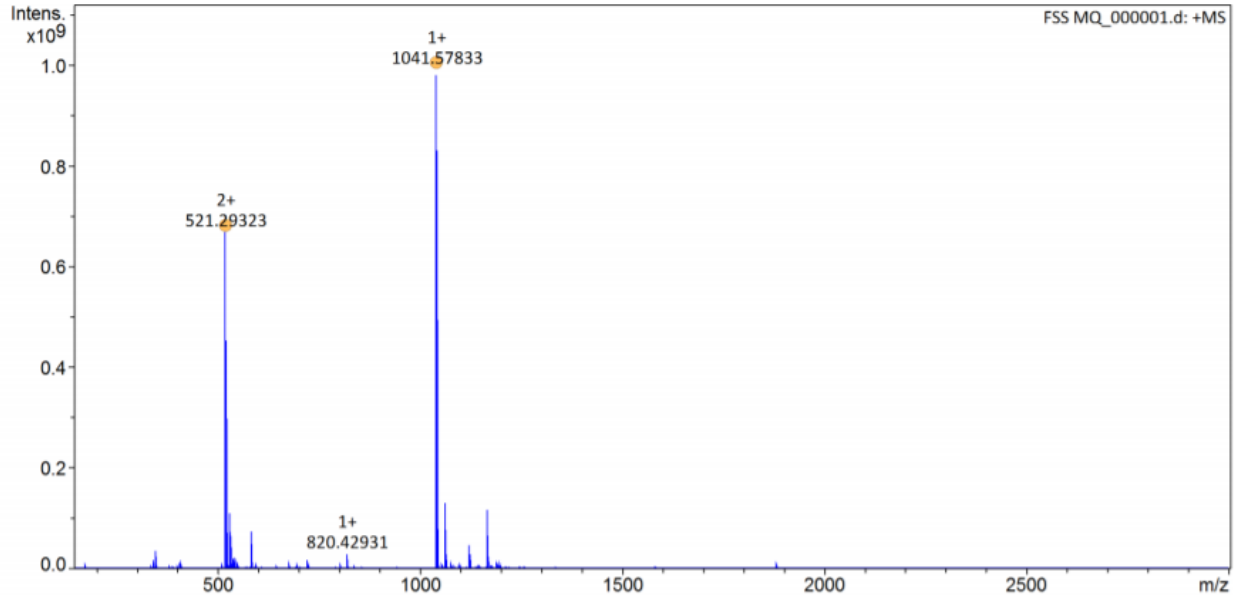
Scheme 4. Huisgen dipolar cycloaddition (click) reaction conditions for monoquinoline (MQ) compound **7**.

The synthesis of MQ compound **7** used similar click and washing step procedures as previously described for DKP compound **8**. After the click reaction and washing step, the remaining Fmoc protecting group was removed with two additions of 20% piperidine for 20 min each, followed by washing via 3x 8 mL DMF/ 8 mL DCM. The resin was then cleaved with 50% trifluoroacetic acid (TFA) and 1% triethyl silane as a scavenger in DCM, as shown in scheme 4. The precipitate was dried under vacuum, yielding the MQ compound (**7**).

2. Compound characterization

^{1}H NMR spectra were recorded at 25 °C on a Bruker Avance 500 (500 MHz) instrument and processed using MestReNova NMR processing software. Chemical shifts (δ) are reported in parts per million (ppm) and referenced to the residual protium signal in the NMR solvents. Data are reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, m = multiplet, and q = quartet), coupling constant (J) in Hertz (Hz), and integration. ^{13}C spectra were recorded at 25 °C on a Bruker Avance 500 instrument operating at 126 MHz. Chemical shifts (δ) are reported in ppm and referenced to the primary carbon resonance in the NMR solvent. High-resolution mass spectra (HRMS) were acquired by the University of Buffalo Chemistry Department Mass Spectrometry Facility, Buffalo, NY.

Compound 7:



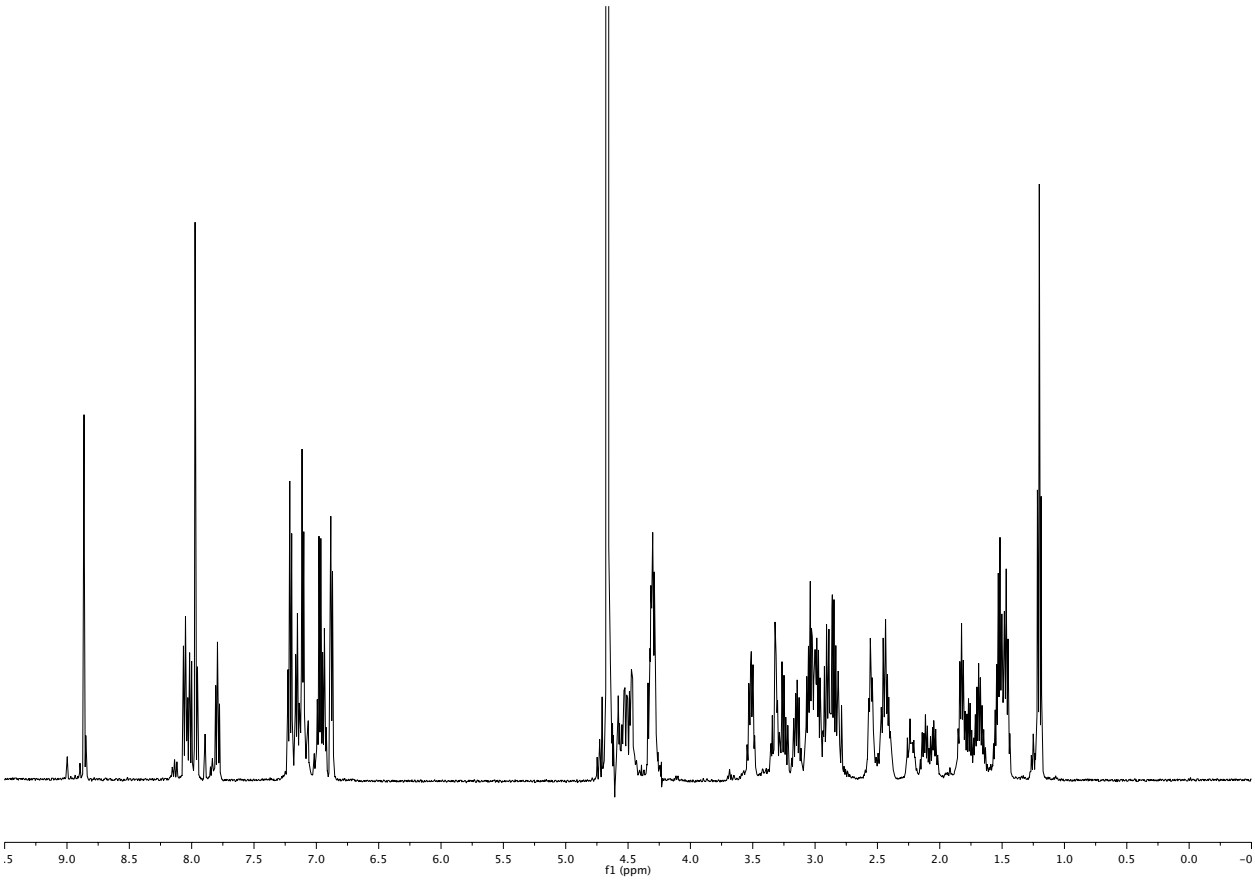
HRMS m/z calculated for $C_{55}H_{72}N_{14}O_7$ $[M+H]^+$: 1041.5708, found: 1041.5783

^{1}H NMR (500 MHz, D_2O) δ 8.86 (s, 1H), 8.09 – 7.99 (m, 3H), 7.97 (s, 2H), 7.79 (t, J = 7.6 Hz, 1H), 7.21 (t, J = 7.5 Hz, 3H), 7.16 (d, J = 7.2 Hz, 1H), 7.11 (d, J = 7.4 Hz, 3H), 6.96 (dq, J = 14.3,

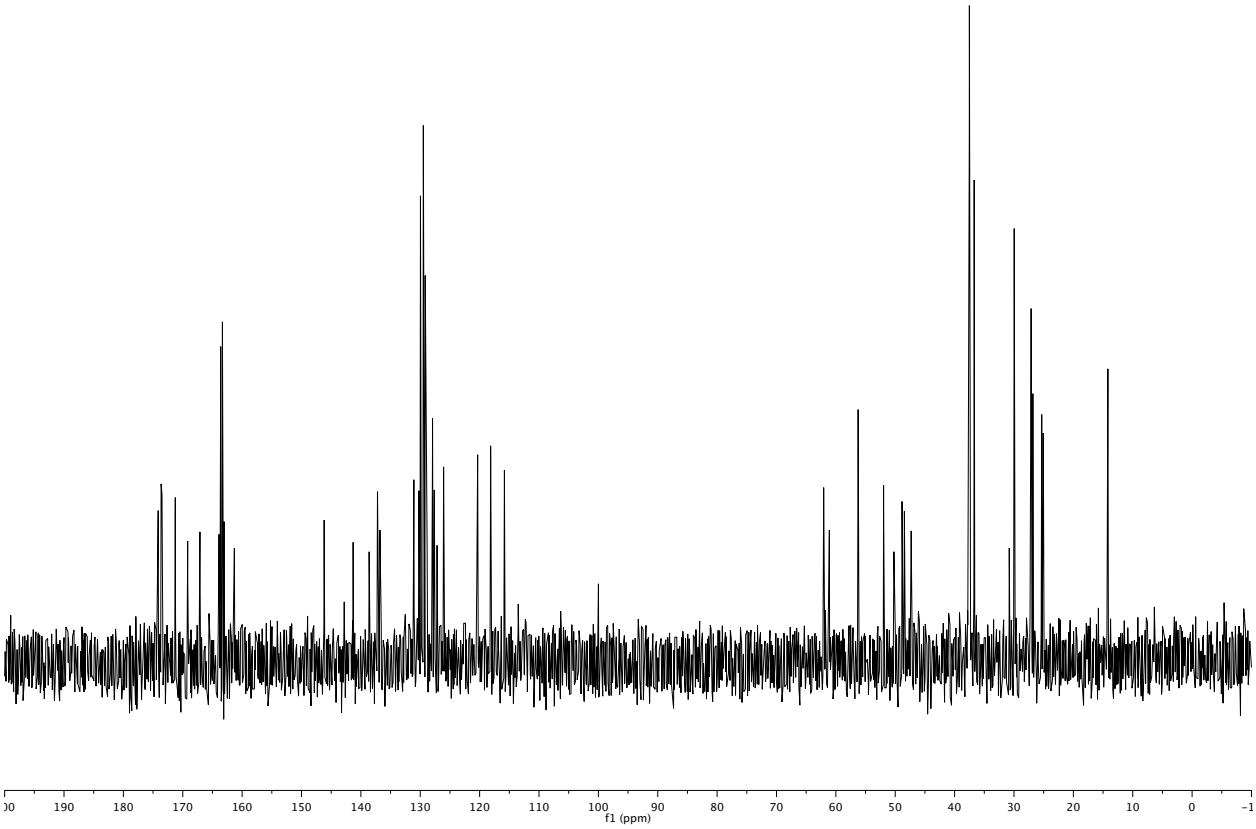
7.4 Hz, 4H), 6.88 (d, *J* = 7.3 Hz, 2H), 4.60 – 4.42 (m, 3H), 4.35 – 4.25 (m, 5H), 3.51 (dt, *J* = 13.8, 7.0 Hz, 2H), 3.37 – 3.28 (m, 2H), 3.24 (dd, *J* = 15.3, 8.1 Hz, 1H), 3.15 (dq, *J* = 15.2, 7.7 Hz, 1H), 3.10 – 2.94 (m, 7H), 2.94 – 2.77 (m, 6H), 2.61 – 2.50 (m, 3H), 2.44 (tt, *J* = 11.6, 5.9 Hz, 4H), 2.23 (td, *J* = 10.7, 5.4 Hz, 1H), 2.16 – 2.00 (m, 2H), 1.82 (p, *J* = 7.3 Hz, 2H), 1.76 (q, *J* = 6.4 Hz, 1H), 1.68 (tq, *J* = 13.5, 6.8 Hz, 2H), 1.58 – 1.43 (m, 7H), 1.20 (t, *J* = 7.6 Hz, 4H).

^{13}C NMR (126 MHz, D_2O) δ 174.18, 173.52, 171.28, 169.19, 167.12, 163.90, 163.62, 163.34, 163.05, 161.32, 146.20, 141.32, 138.62, 137.22, 136.80, 131.09, 130.23, 129.98, 129.68, 129.49, 129.17, 128.85, 127.94, 127.66, 127.18, 126.08, 120.34, 118.16, 115.84, 62.06, 61.11, 56.25, 51.97, 50.23, 48.86, 48.46, 47.32, 37.74, 37.52, 36.71, 30.79, 29.97, 27.09, 26.82, 25.33, 25.07, 14.19.

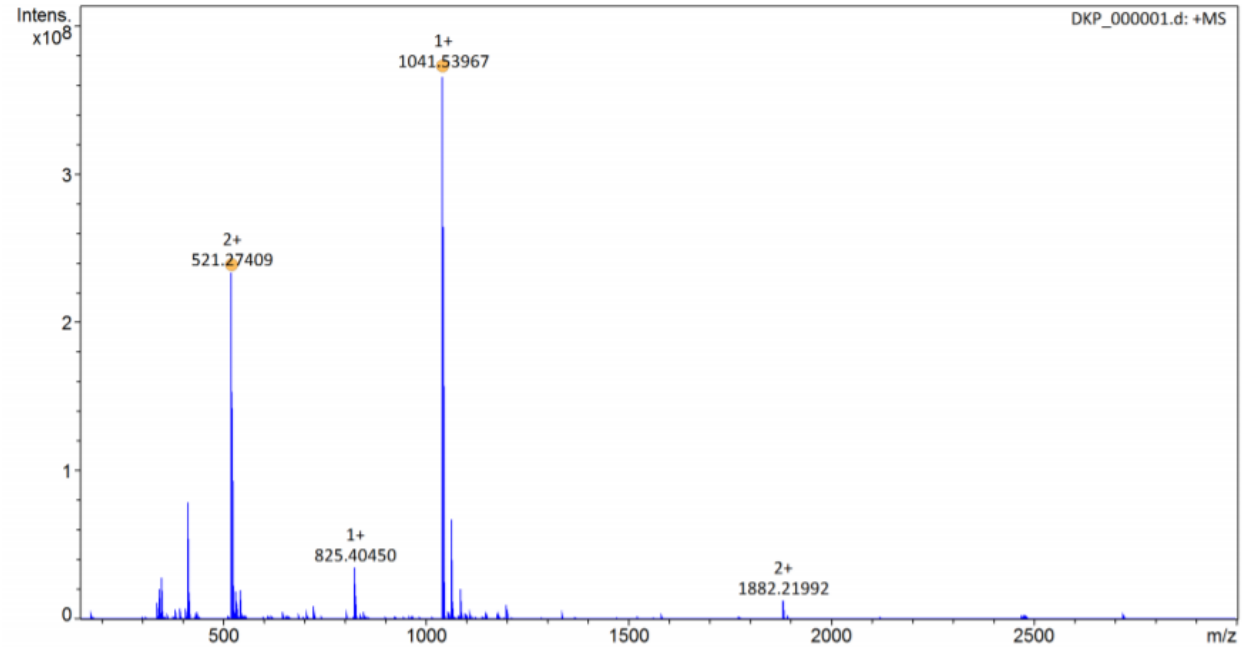
^{1}H NMR spectrum:



^{13}C NMR spectrum:



Compound 8:



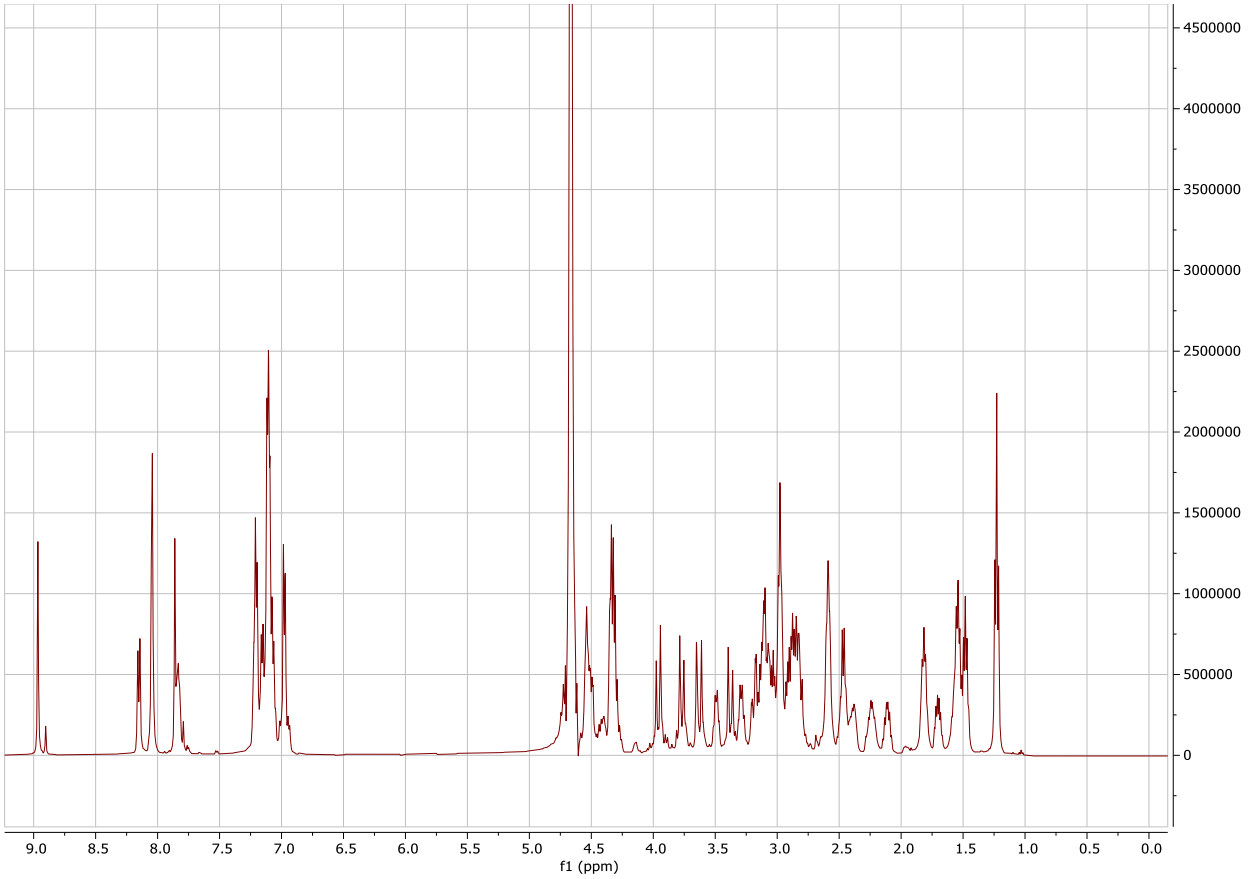
HRMS m/z calculated for $C_{54}H_{68}N_{14}O_8$ $[M+H]^+$: 1041.5345, found: 1041.5397

^{1}H NMR (500 MHz, D_2O) δ 8.97 (s, 1H), 8.15 (d, *J* = 8.3 Hz, 1H), 8.04 (s, 2H), 7.86 (s, 1H), 7.85 – 7.78 (m, 2H), 7.21 (t, *J* = 7.1 Hz, 3H), 7.16 (d, *J* = 7.1 Hz, 1H), 7.11 (t, *J* = 6.4 Hz, 6H), 7.07 (d, *J* = 7.4 Hz, 1H), 6.98 (d, *J* = 7.4 Hz, 2H), 4.57 – 4.47 (m, 3H), 4.37 – 4.27 (m, 5H), 3.96 (d, *J* = 16.8 Hz, 1H), 3.77 (d, *J* = 16.6 Hz, 1H), 3.64 (d, 2H), 3.49 (q, *J* = 7.7 Hz, 1H), 3.38 (d, *J* = 18.0 Hz, 1H), 3.29 (q, *J* = 7.7 Hz, 1H), 3.23 – 3.01 (m, 5H), 2.98 (t, *J* = 6.7 Hz, 3H), 2.93 – 2.77 (m, 4H), 2.59 (t, *J* = 7.7 Hz, 4H), 2.47 (q, *J* = 7.4 Hz, 2H), 2.11 (dq, *J* = 14.6, 7.4 Hz, 1H), 1.88 – 1.76 (m, 3H), 1.71 (dt, *J* = 12.9, 6.2 Hz, 1H), 1.53 (dd, *J* = 14.7, 7.4 Hz, 4H), 1.47 (q, *J* = 7.2 Hz, 2H), 1.23 (t, *J* = 7.6 Hz, 4H).

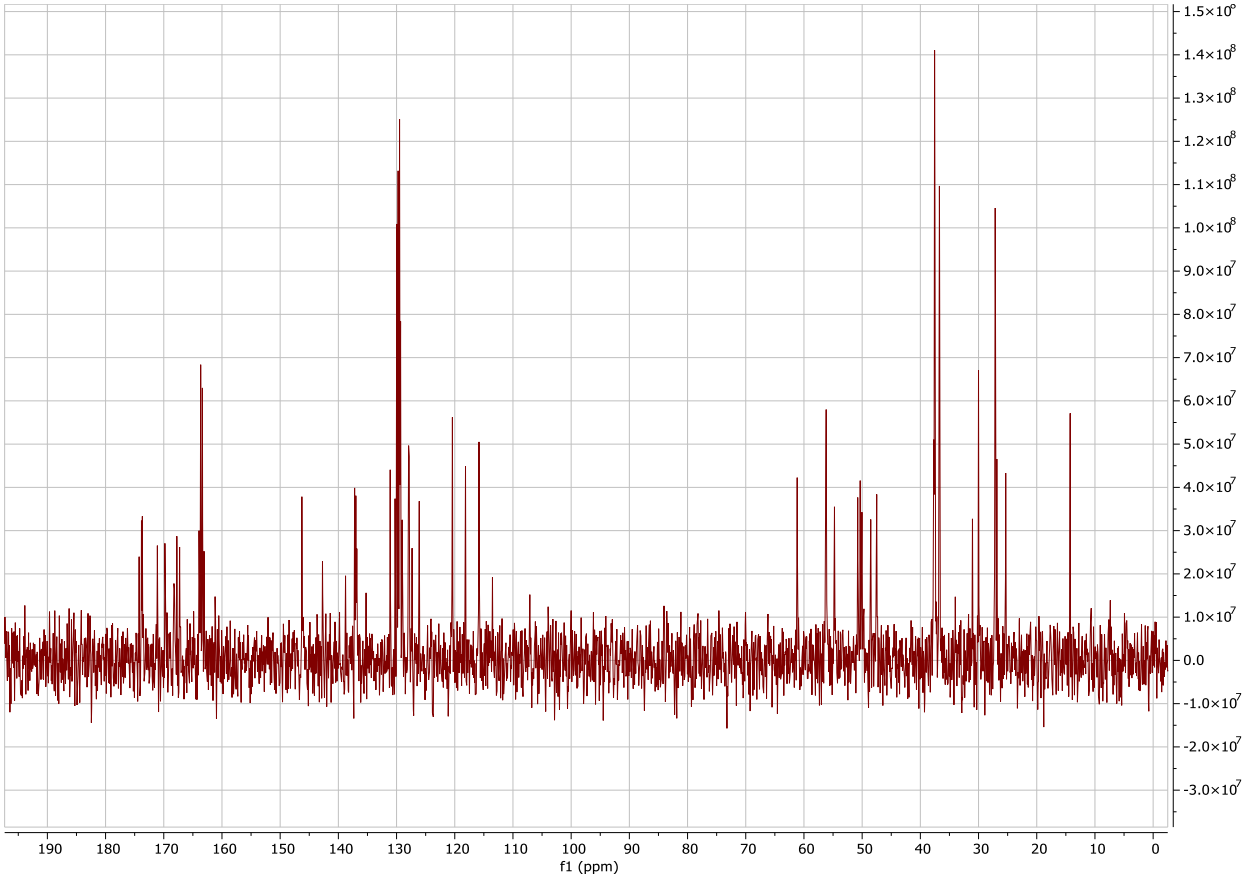
^{13}C NMR (126 MHz, D_2O) δ 173.68, 171.09, 169.80, 167.75, 167.28, 163.66, 163.37, 163.08, 146.27, 137.21, 137.00, 131.09, 130.26, 129.97, 129.73, 129.49, 129.34, 129.02, 127.93, 127.34,

126.10, 120.40, 118.16, 115.84, 61.16, 56.19, 54.76, 50.76, 50.34, 50.04, 48.53, 47.50, 37.54, 36.72, 31.05, 30.01, 27.14, 26.85, 25.33, 14.25.

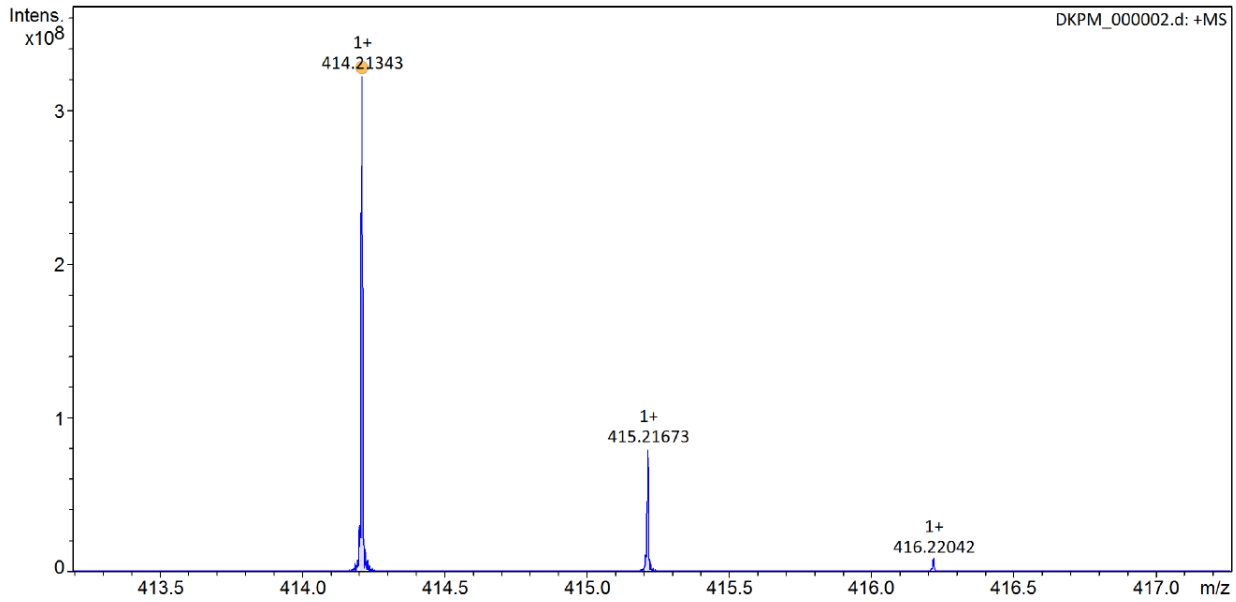
^{1}H NMR spectrum:



^{13}C NMR spectrum:



Compound 12 (DKP Monomer)

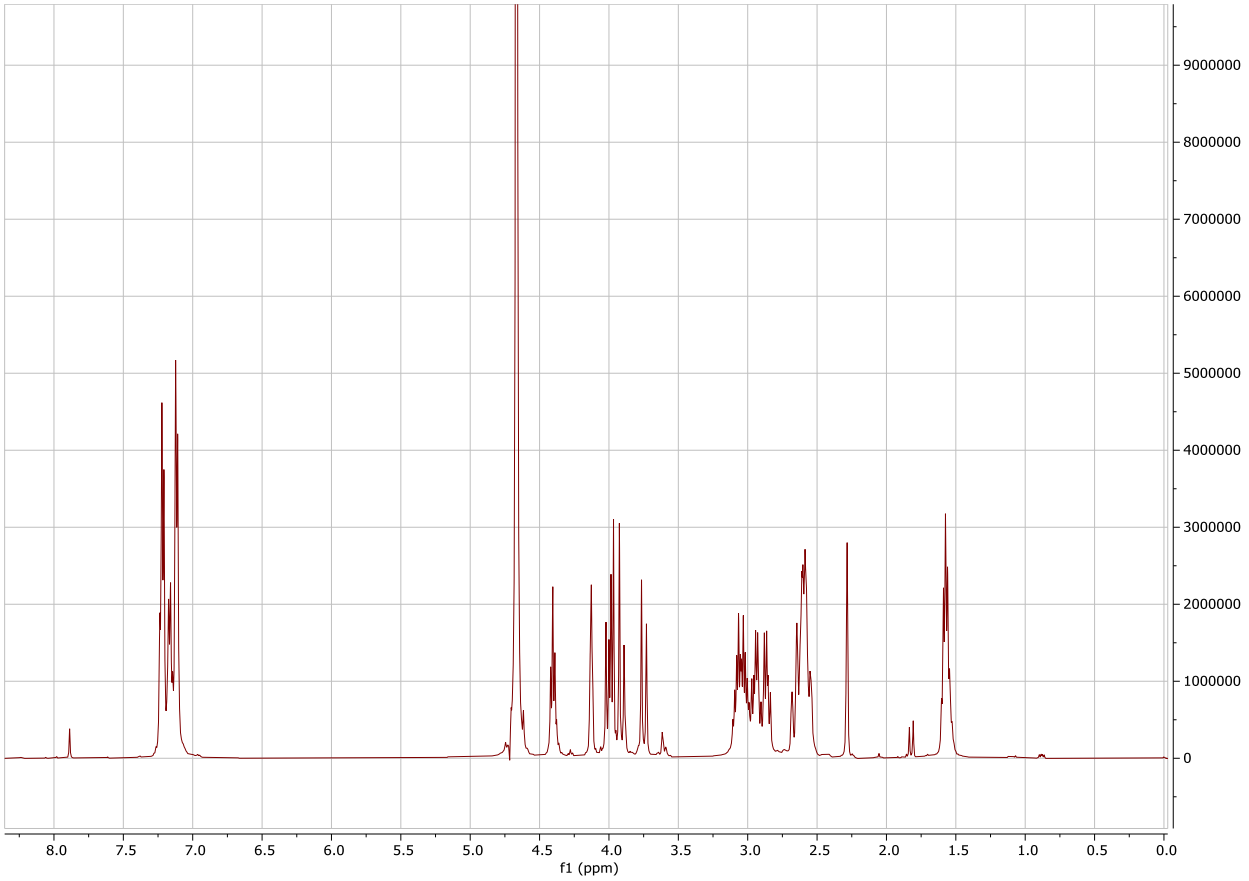


HRMS m/z calculated for $C_{21}H_{27}N_5O_4$ $[M+H]^+$: 414.2063, found: 414.2134

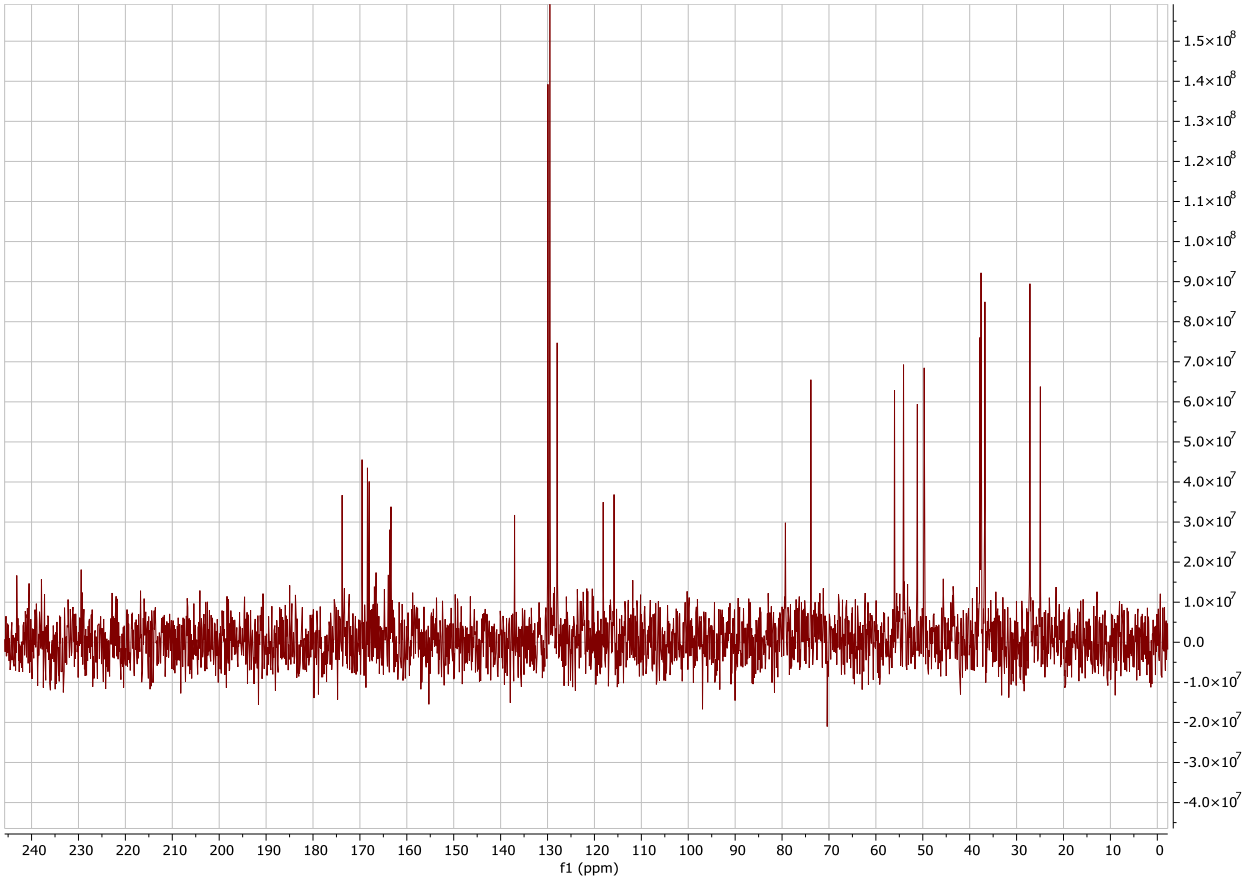
^{1}H NMR (500 MHz, D2O) δ 7.22 (t, J = 7.5 Hz, 3H), 7.17 (d, J = 7.1 Hz, 1H), 7.12 (d, J = 7.4 Hz, 3H), 4.40 (t, J = 7.7 Hz, 1H), 4.13 (d, J = 4.5 Hz, 1H), 4.03 – 3.86 (m, 3H), 3.75 (d, J = 17.9 Hz, 1H), 3.15 – 2.79 (m, 6H), 2.71 – 2.49 (m, 5H), 2.29 (d, J = 2.6 Hz, 1H), 1.57 (h, J = 8.2 Hz, 3H).

^{13}C NMR (126 MHz, D_2O) δ 173.79, 169.54, 168.39, 168.02, 163.65, 163.37, 137.03, 129.92, 129.50, 127.95, 118.14, 115.82, 79.29, 73.85, 56.02, 54.11, 51.16, 49.70, 37.82, 37.58, 36.75, 27.16, 24.94.

^{1}H NMR Spectrum:



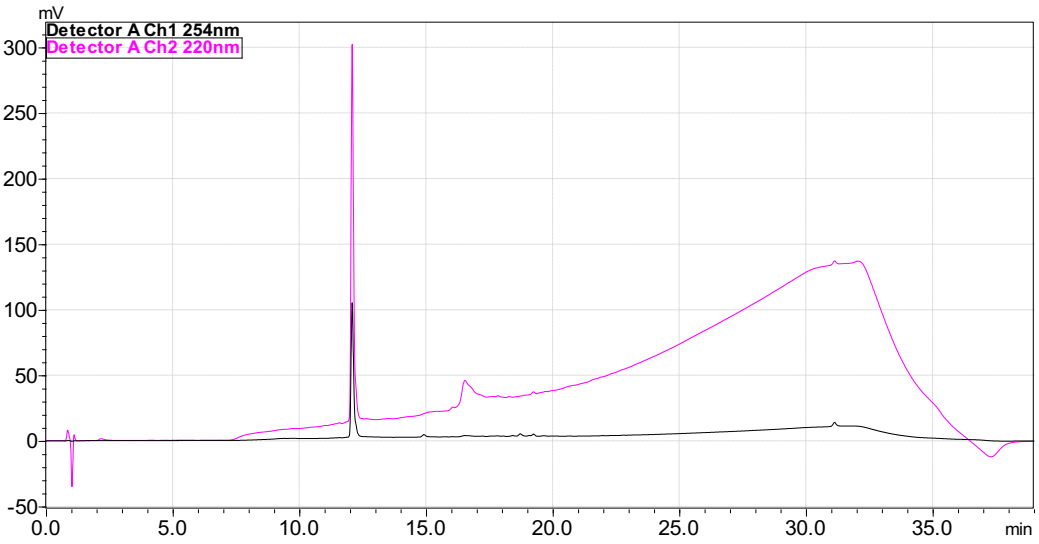
^{13}C NMR spectrum:



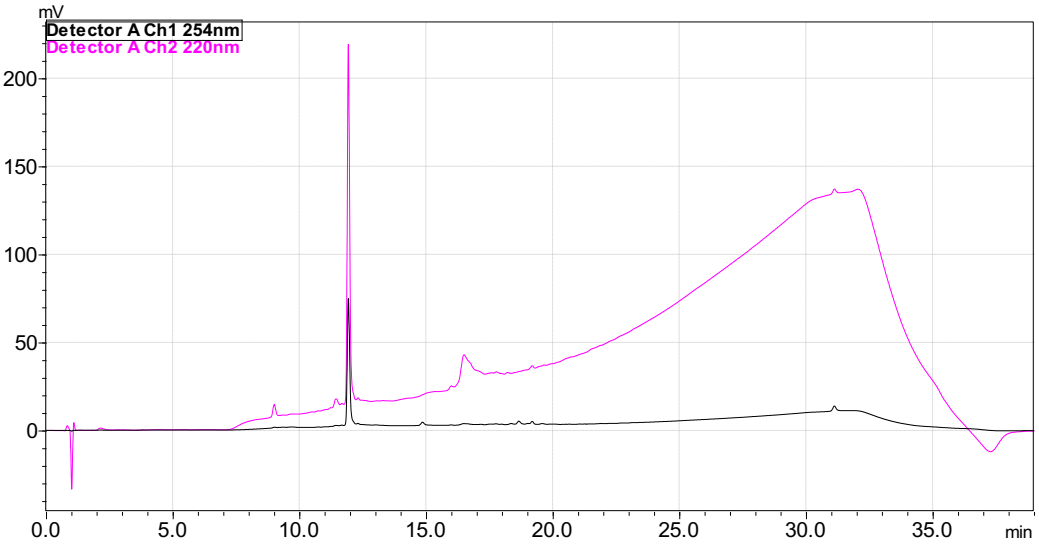
Confirmation of purity – Analytical HPLC:

Note: Peak visible in the 220 nm trace at a retention time of 17 minutes is also observed in blank runs.

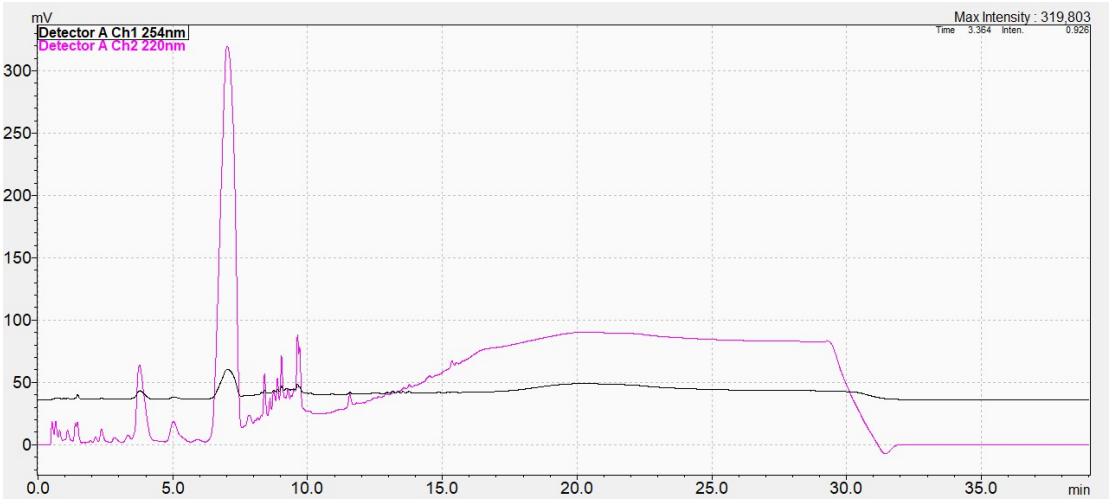
Compound **7**: Purity 98% (254 nm)



Compound **8**: Purity 92% (254 nm)



Compound **12** (DKP Monomer): Purity 85% (254 nm)

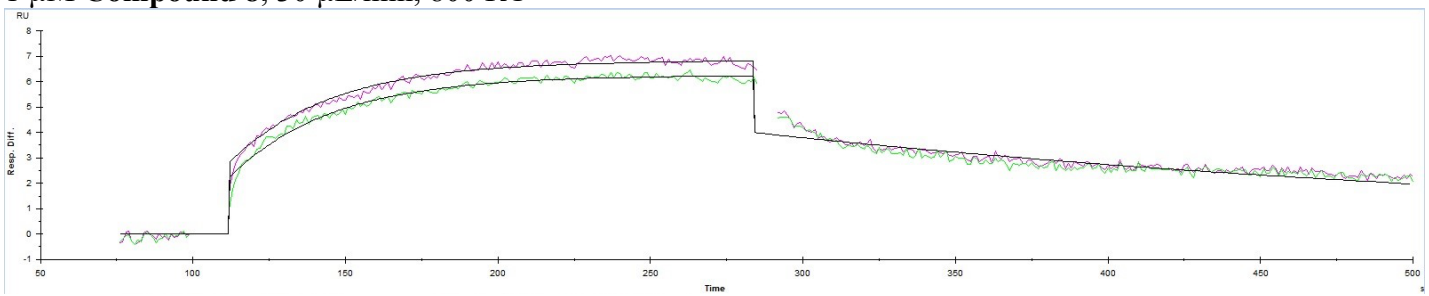


3. Detailed SPR data

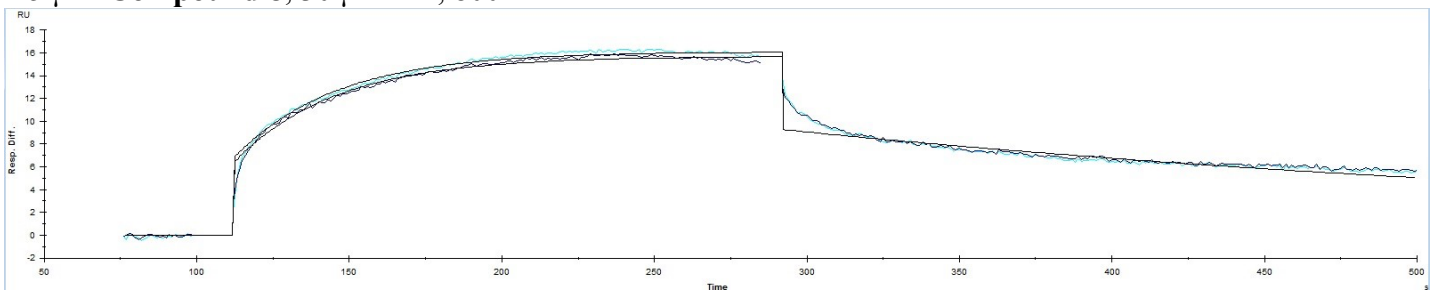
SPR experiments were performed on a Biacore X100 instrument (GE Healthcare Bio-Sciences AB) using a CM5 chip (GE). Before preparing the chip, 5'-amine terminated HIV-1 FSS RNA was refolded in HEPES buffer (10 mM HEPES, 150 mM NaCl, 3 mM EDTA and 0.005% Tween 20) at pH 7.4 by heating the RNA in a heat block at 95 ºC for 5 min and allowed to reach room temperature in the heat block over the course of approximately 1 hour. Flow cell (FC) one (FC-1) was used as the reference cell, while a 1:1 mixture of 0.4 M EDC/ 0.1 M NHS in RNAse-free water was injected to FC-2. Next, the refolded 5'-amine termianted HIV-1 FSS RNA was injected into FC-2 with concentrations of 20 µM and 1 µM on separate CM5 chips, resulting in approximately 1500 and 800 response units (RU) of immobilized 5'-amine terminated HIV-1 FSS RNA, respectively. Compounds were injected at flow rates of 30 µL/min and 70 µL/min in running buffer (10 mM HEPES, 150 mM NaCl, 3 mM EDTA and 0.005% Tween-20) with a contact/injection time of 180 and 60 seconds, respectively. 500 mM and 1 M injections of NaCl was used to regenerate the chip for consecutive runs.

800 RU Loading Capacity Chip
DKP Dimer – **Compound 8**, 30 µL/min, 800 RU

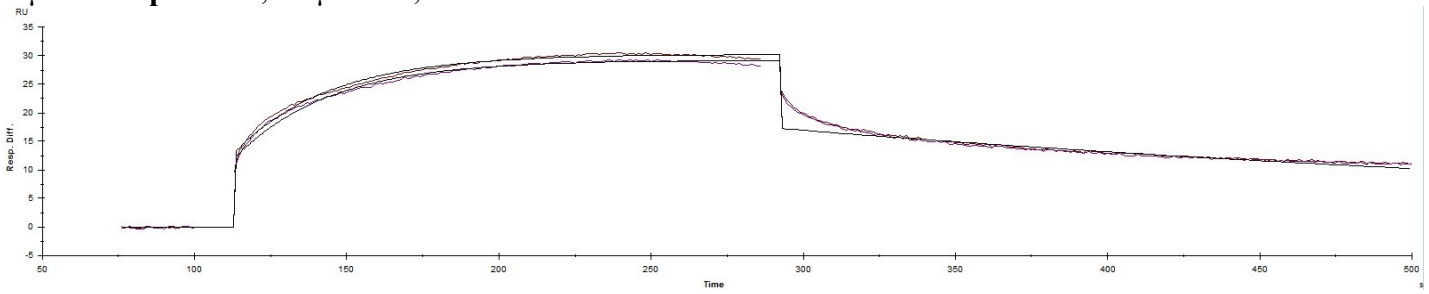
1 µM **Compound 8**, 30 µL/min, 800 RU



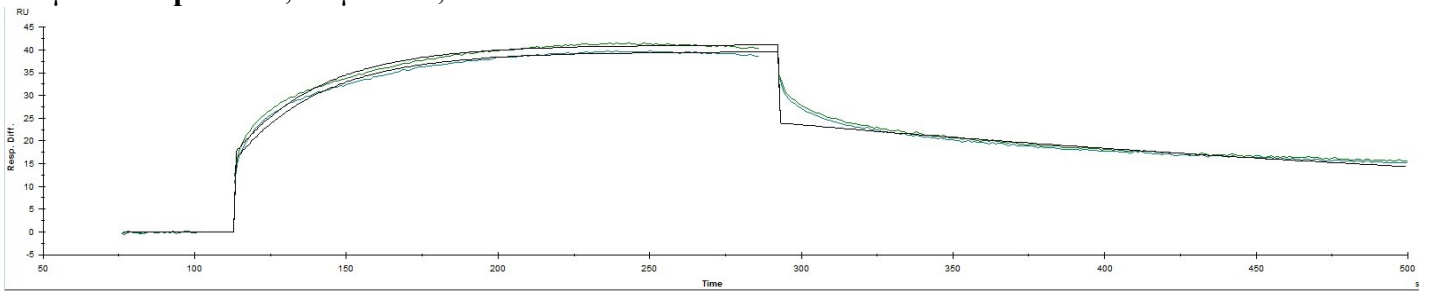
2.5 µM **Compound 8**, 30 µL/min, 800 RU



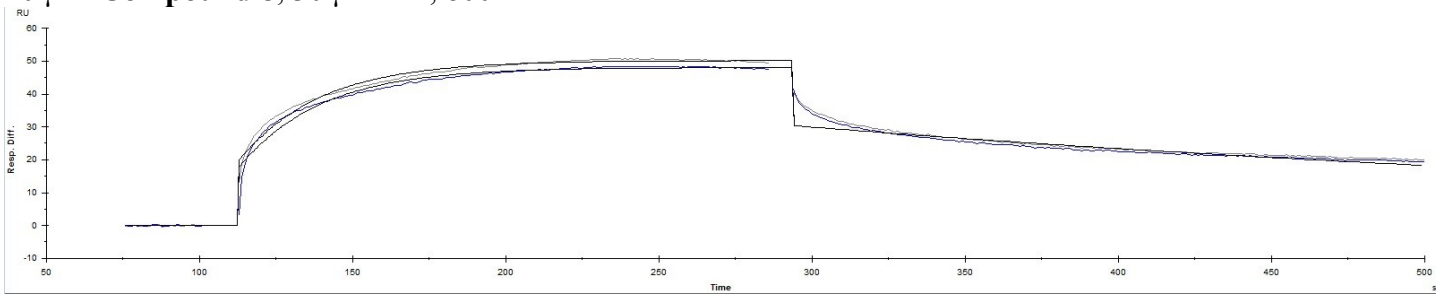
5 µM **Compound 8**, 30 µL/min, 800 RU



7.5 µM **Compound 8**, 30 µL/min, 800 RU

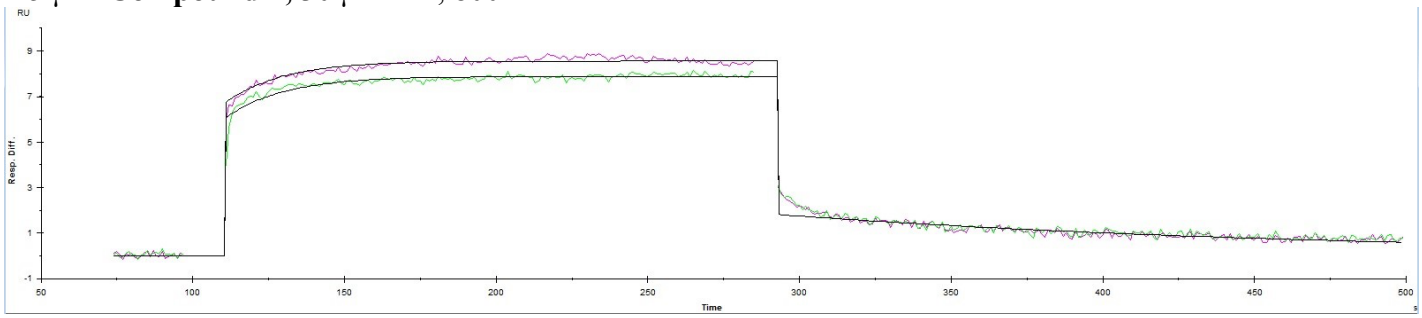


10 µM **Compound 8**, 30 µL/min, 800 RU

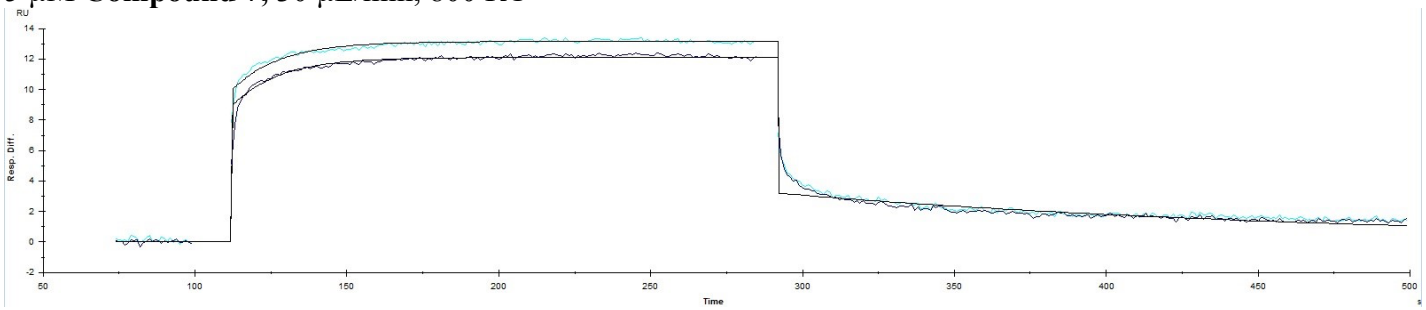


MQ- **Compound 7,** 30 µL/min, 800 RU

2.5 µM **Compound 7,** 30 µL/min, 800 RU

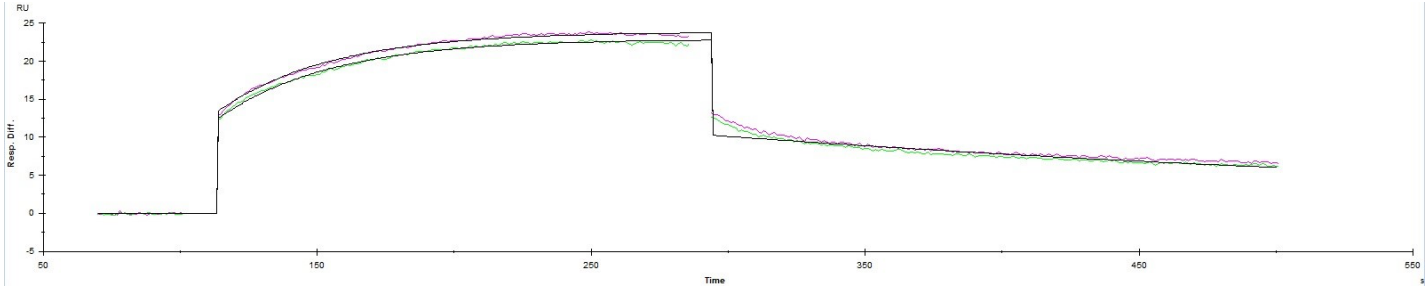


5 µM **Compound 7,** 30 µL/min, 800 RU

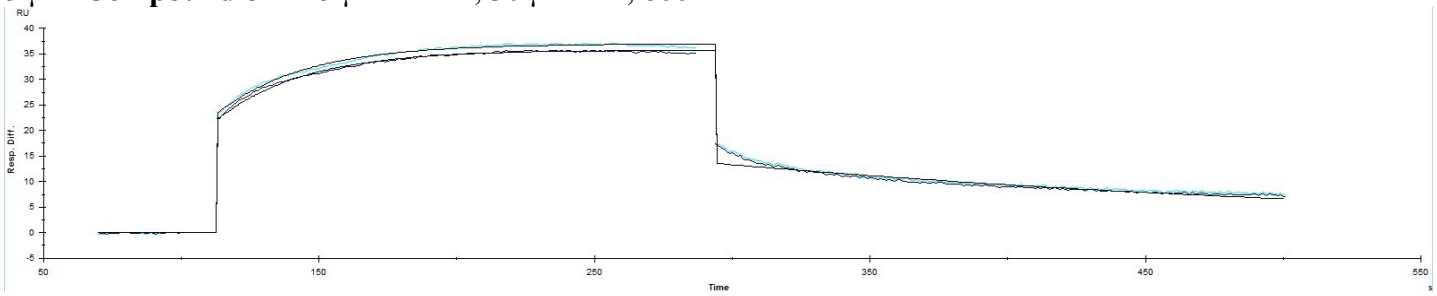


DKP (**Compound 8**) + tRNA, 30 µL/min, 800 RU

2.5 µM **Compound 8** + 12.5 µM tRNA, 30 µL/min, 800 RU

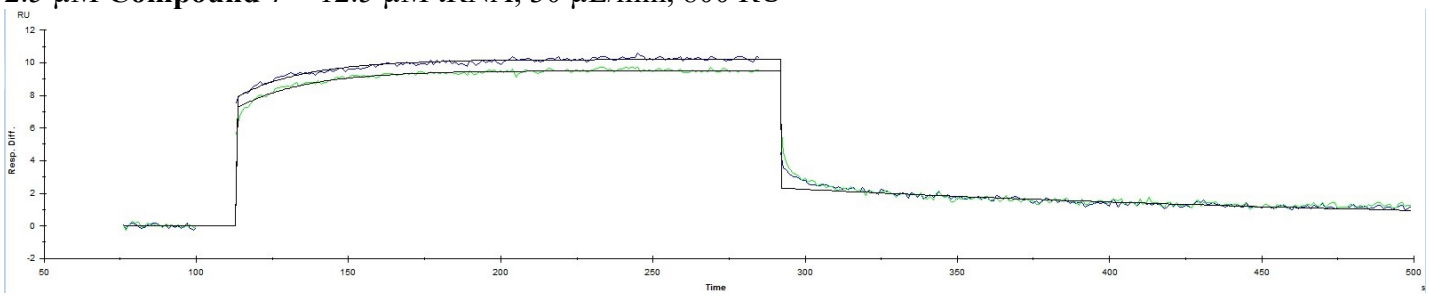


5 µM **Compound 8** + 25 µM tRNA, 30 µL/min, 800 RU

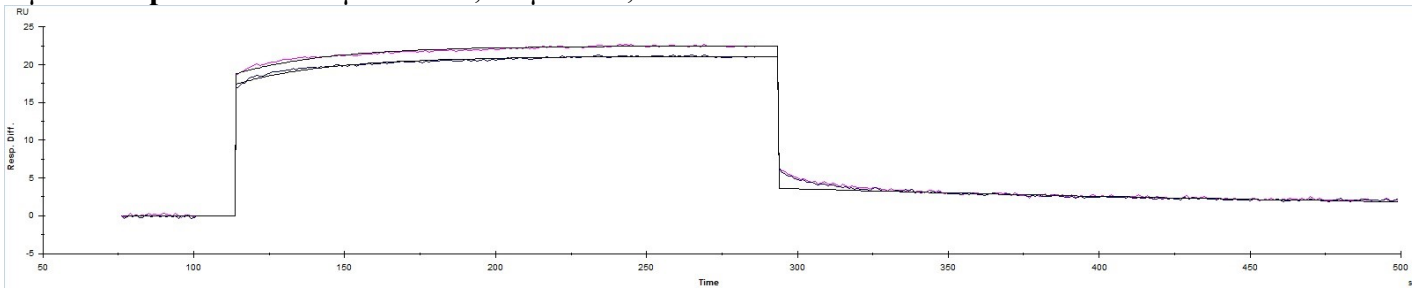


MQ (**Compound 7**) + tRNA, 30 µL/min, 800 RU

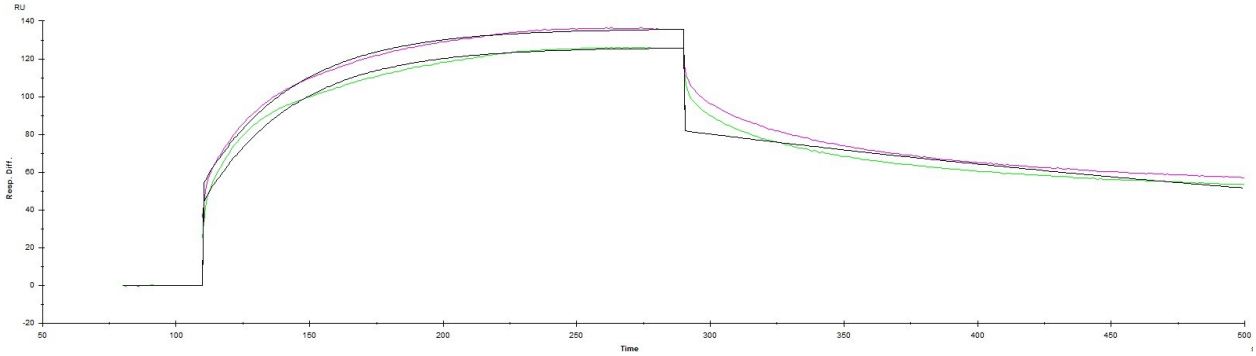
2.5 µM **Compound 7** + 12.5 µM tRNA, 30 µL/min, 800 RU



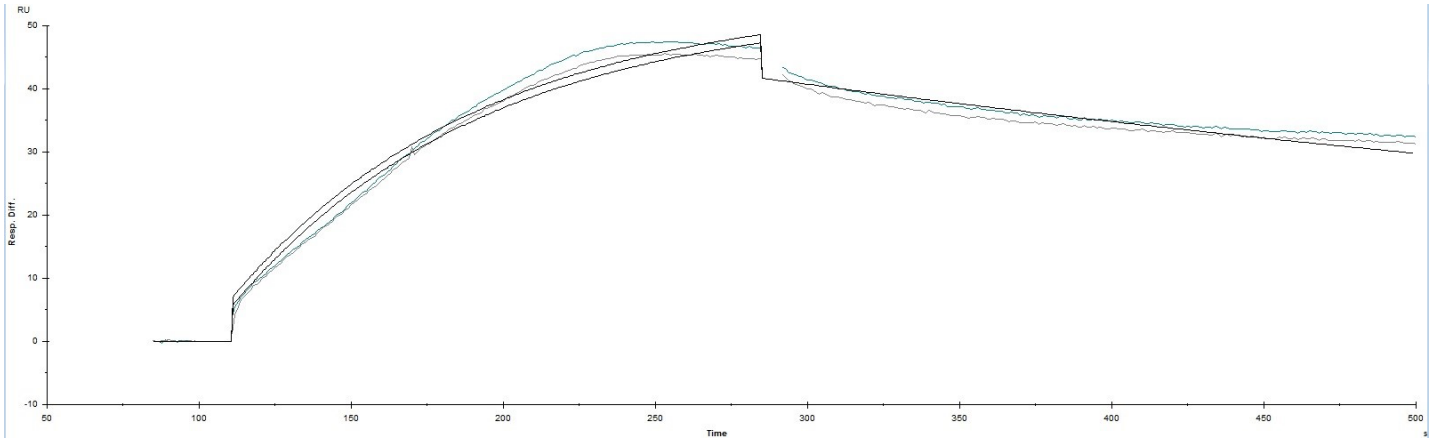
5 µM **Compound 7** + 25 µM tRNA, 30 µL/min, 800 RU



0.5 µM Tz0 (**Compound 2**) 30 µL/min, 800 RU

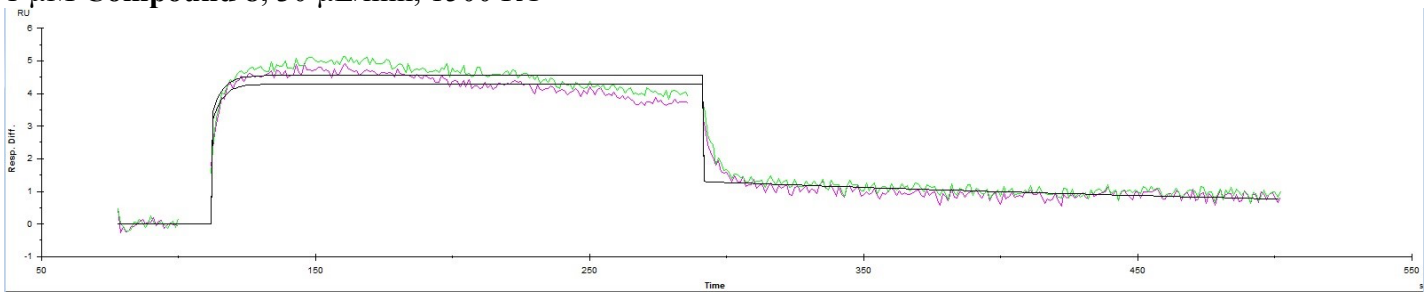


0.5 µM Tz02 (**Compound 3**) 30 µL/min, 800 RU

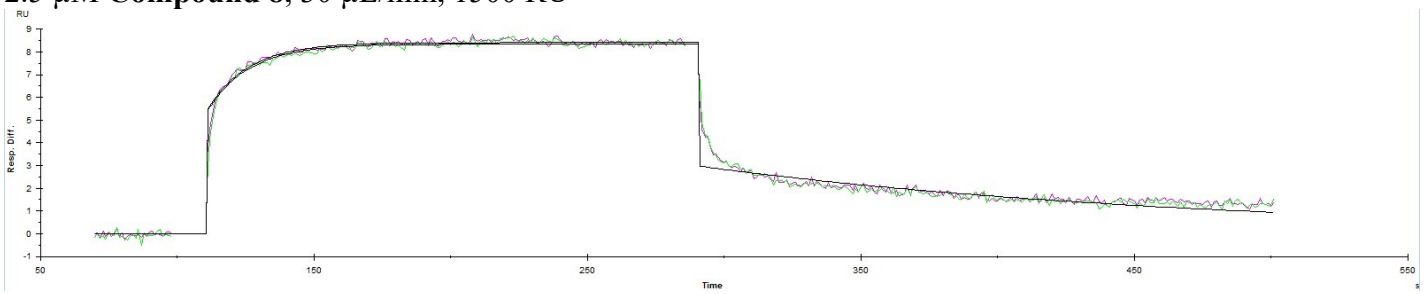


RNA Loading Capacity - 1500 RU
DKP Dimer – **Compound 8**, 30 µL/min, 1500 RU

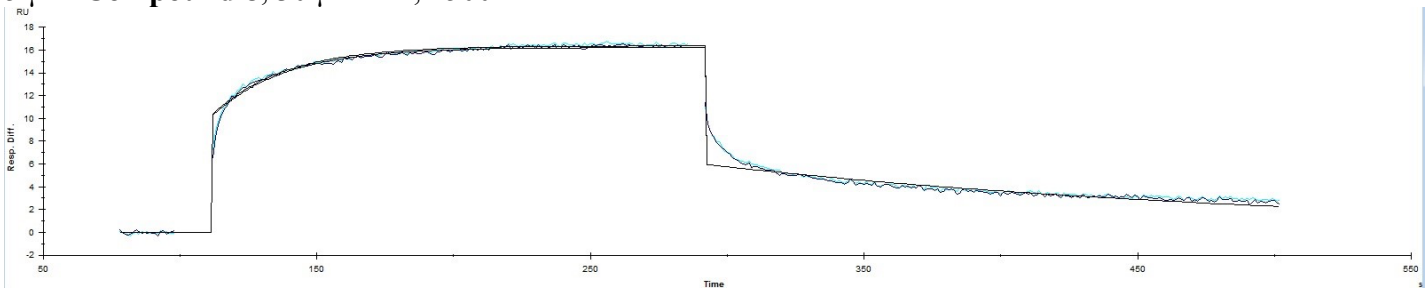
1 µM **Compound 8**, 30 µL/min, 1500 RU



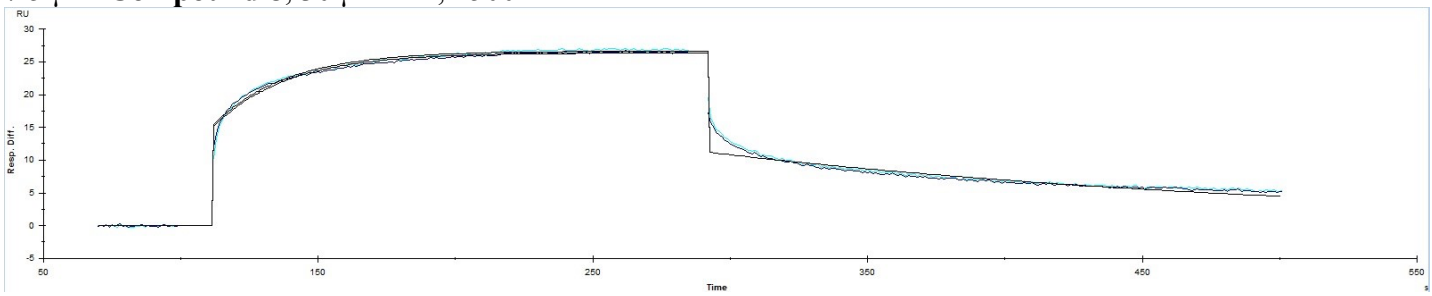
2.5 µM **Compound 8**, 30 µL/min, 1500 RU



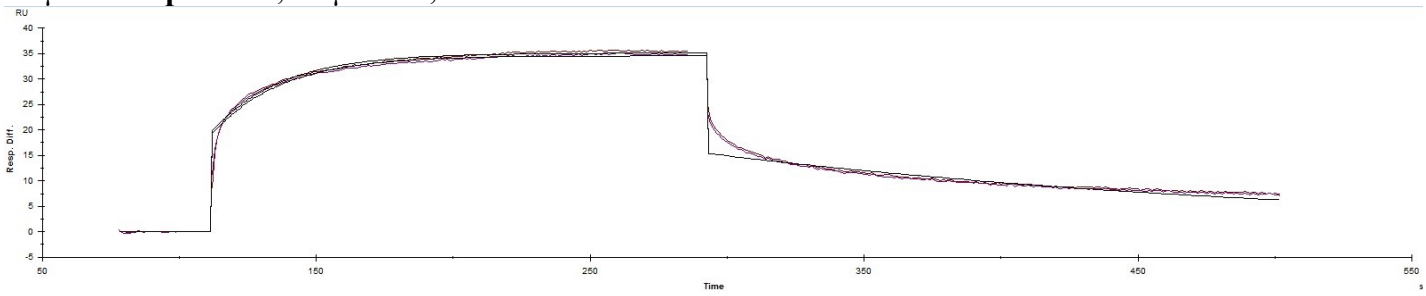
5 µM **Compound 8**, 30 µL/min, 1500 RU



7.5 µM **Compound 8**, 30 µL/min, 1500 RU

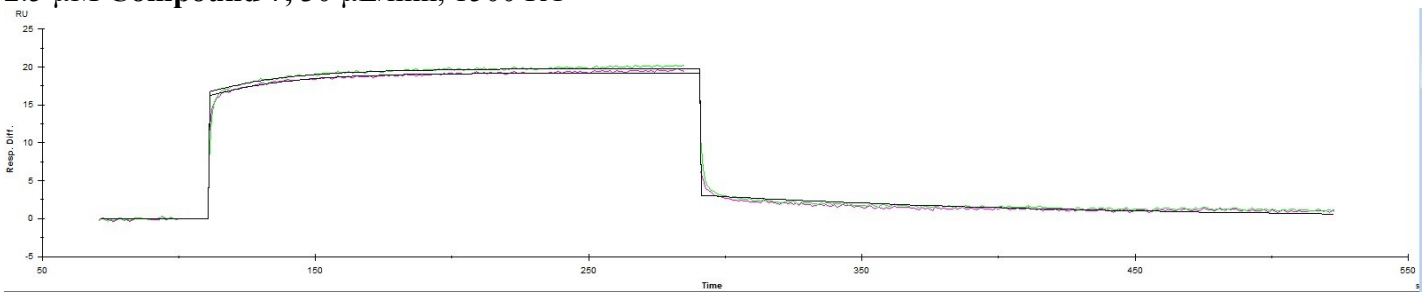


10 µM **Compound 8**, 30 µL/min, 1500 RU



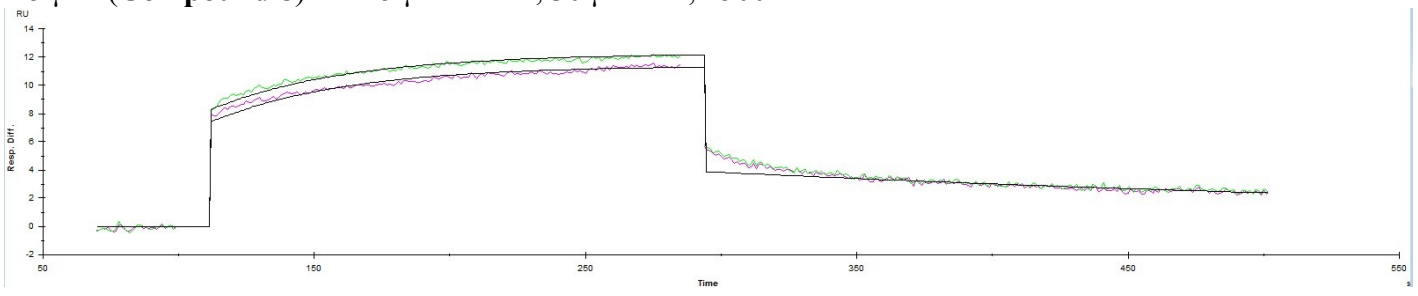
MQ- **Compound 7,** 30 µL/min, 1500 RU

2.5 µM **Compound 7,** 30 µL/min, 1500 RU

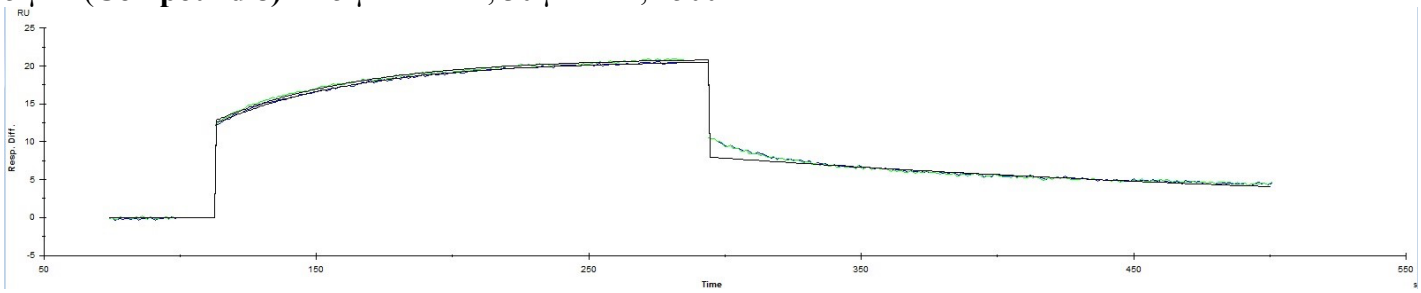


DKP (**Compound 8**)+ tRNA, 30 µL/min, 1500 RU

2.5 µM (**Compound 8**) + 12.5 µM tRNA, 30 µL/min, 1500 RU

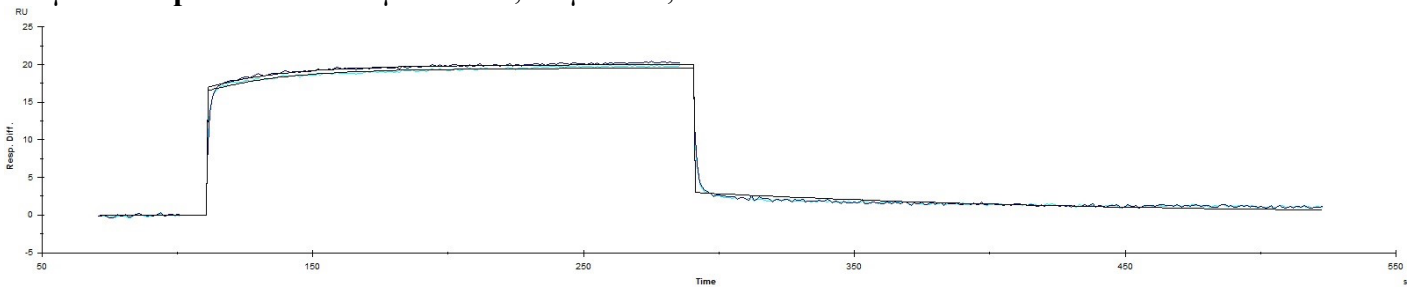


5 µM (**Compound 8)** + 25 µM tRNA, 30 µL/min, 1500 RU



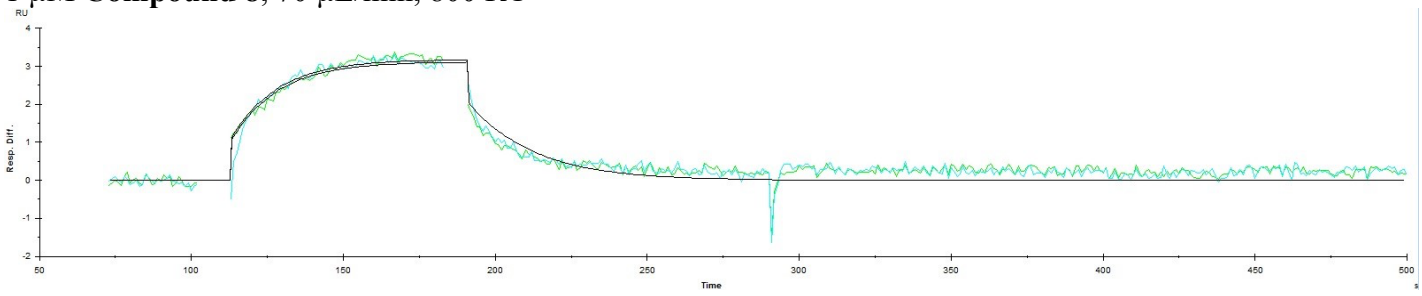
MQ (**Compound 7**) +tRNA, 30 µL/min, 1500 RU

2.5 µM **Compound 7** + 12.5 µM tRNA, 30 µL/min, 1500 RU

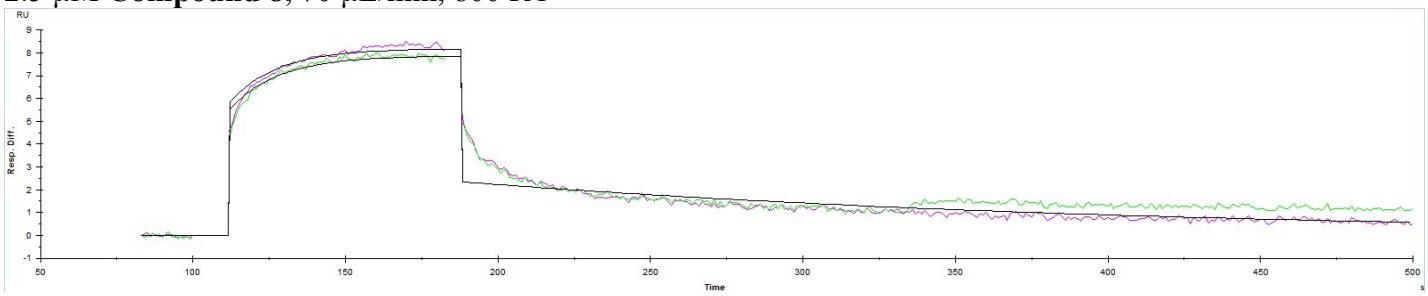


DKP (**Compound 8**), 70 µL/min, 800 RU

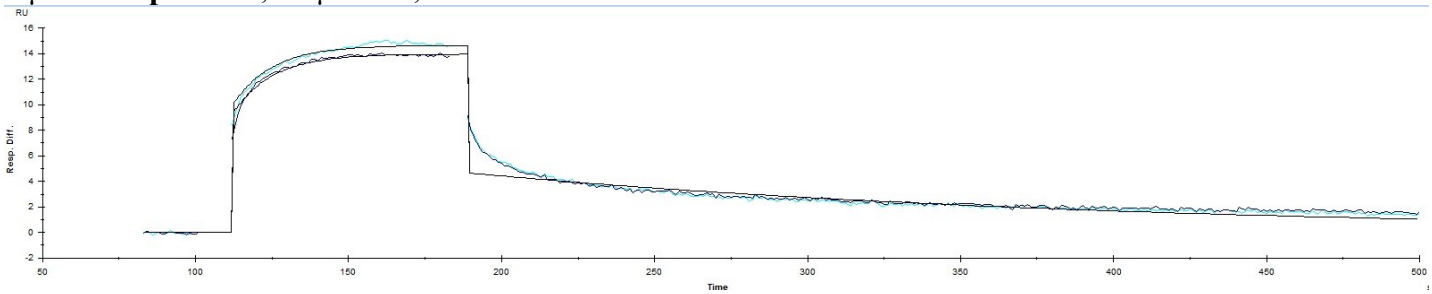
1 µM **Compound 8**, 70 µL/min, 800 RU



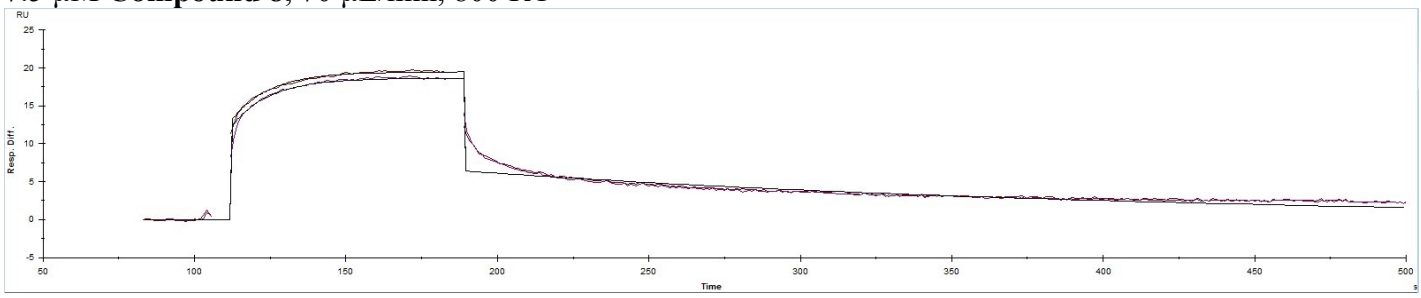
2.5 µM **Compound 8**, 70 µL/min, 800 RU



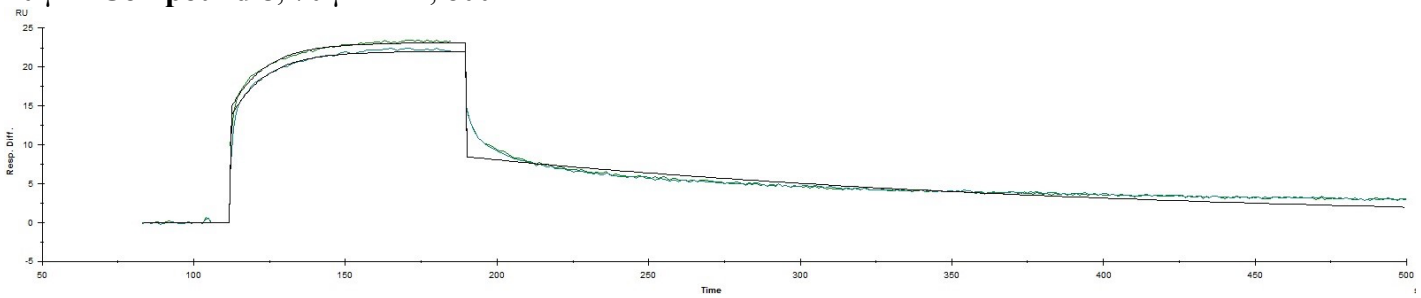
5 μM **Compound 8**, 70 μL/min, 800 RU



7.5 μM **Compound 8**, 70 μL/min, 800 RU

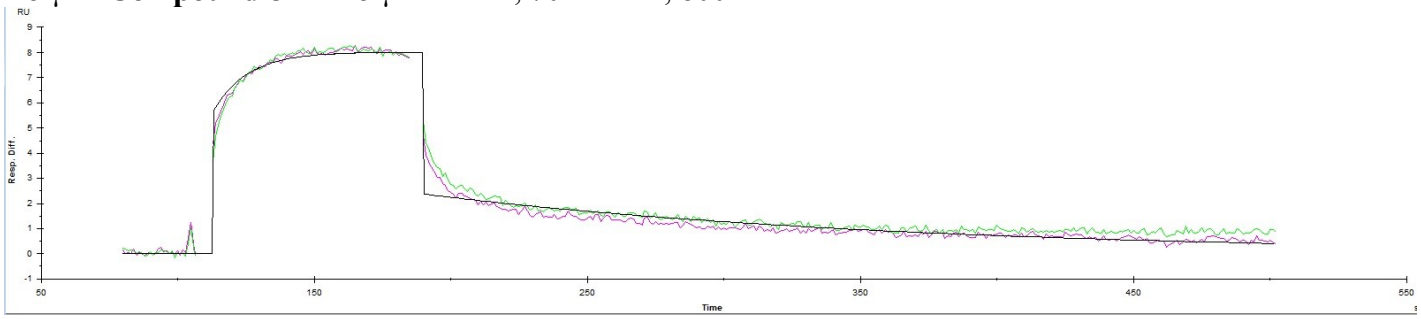


10 μM **Compound 8**, 70 μL/min, 800 RU

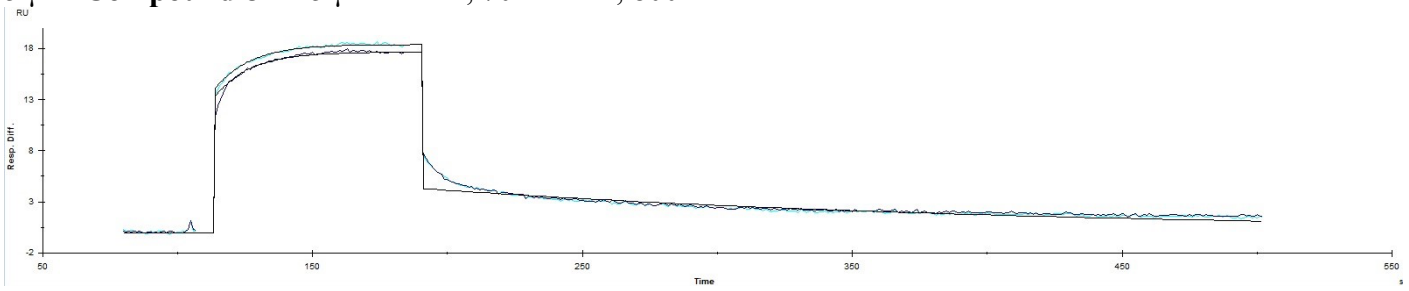


DKP (**Compound 8**) + tRNA, 70 uL/min, 800 RU

2.5 μM **Compound 8** + 12.5 μM tRNA, 70 uL/min, 800 RU

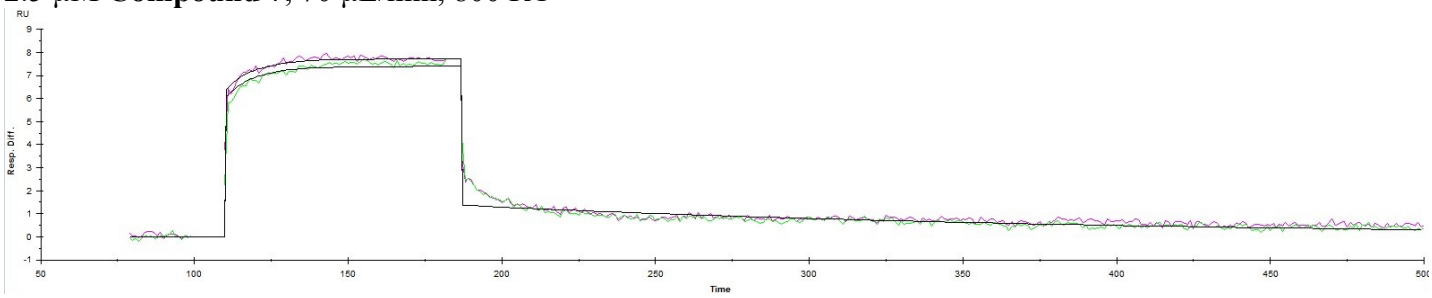


5 µM **Compound 8** + 25 µM tRNA, 70 uL/min, 800 RU

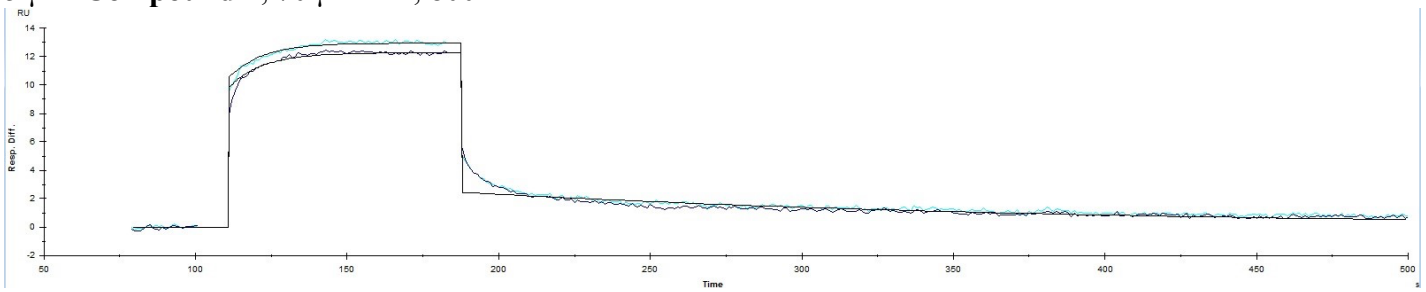


MQ (**Compound 7**), 70 µL/min, 800 RU

2.5 µM **Compound 7**, 70 µL/min, 800 RU

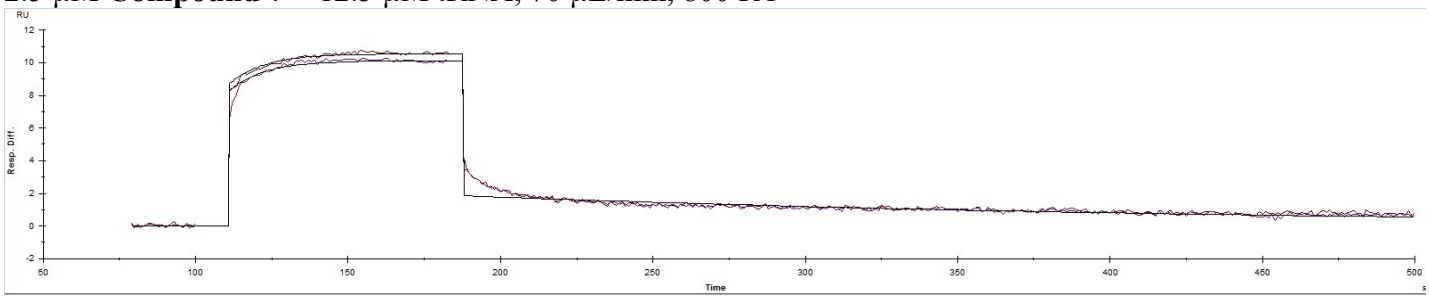


5 µM **Compound 7**, 70 µL/min, 800 RU

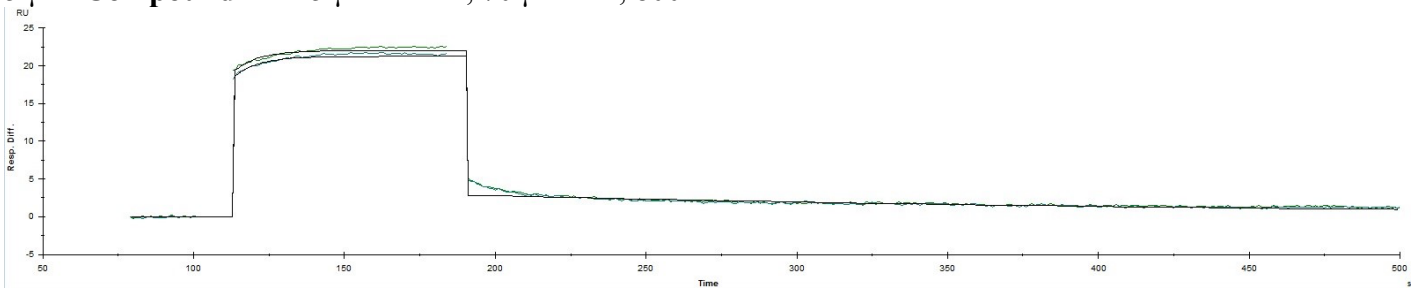


MQ (**Compound 7)** + tRNA, 70 µL/min, 800 RU

2.5 µM **Compound 7** + 12.5 µM tRNA, 70 µL/min, 800 RU

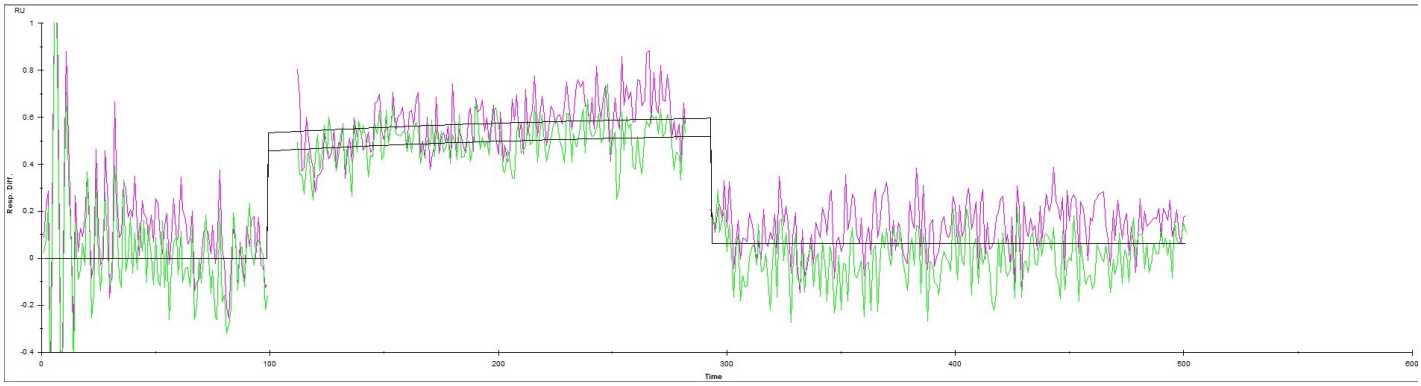


5 µM **Compound 7** + 25 µM tRNA, 70 µL/min, 800 RU

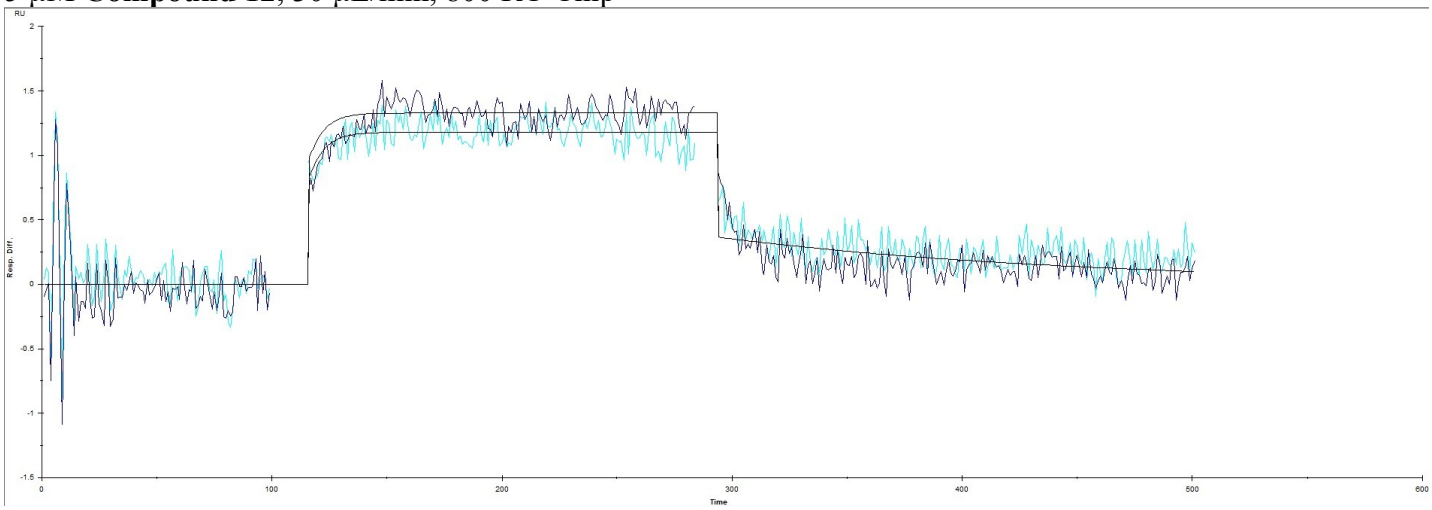


DKP Monomer (**Compound 12**), 30 µL/min, 800 RU Chip

2.5 µM **Compound 12**, 30 µL/min, 800 RU Chip

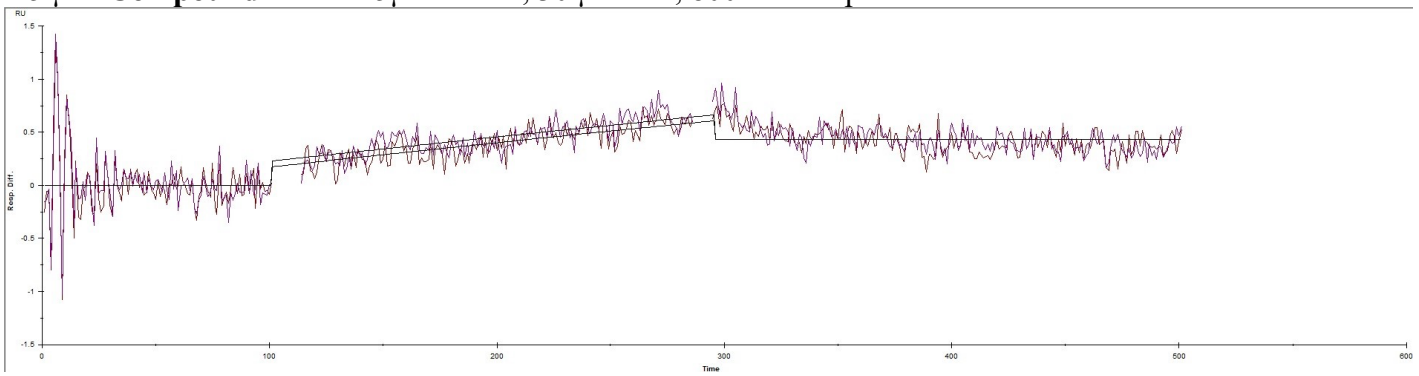


5 µM **Compound 12**, 30 µL/min, 800 RU Chip

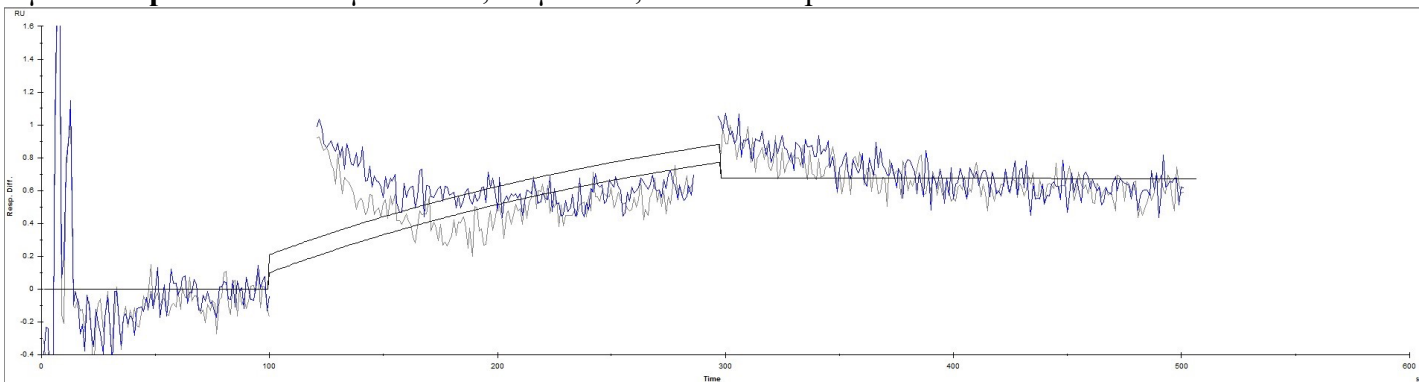


DKP Monomer (**Compound 12**) + tRNA, 30 µL/min, 800 RU Chip

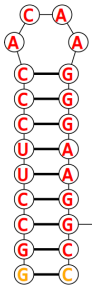
2.5 µM **Compound 12** + 12.5µM tRNA, 30 µL/min, 800 RU Chip



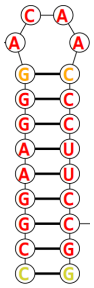
5 µM **Compound 12** + 25 µM tRNA, 30 µL/min, 800 RU Chip



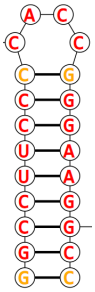
4. Competition SPR with closely related sequences. An SPR chip bearing 800 RU HIV-1 FSS RNA throughout these experiments. Compound **8** was flowed using parameters analogous to those shown in the previous section, but in the presence of a 5-fold excess of one of three RNA sequences: the HIV-1 FSS RNA, a "stem-swapped" sequence, or a "loop mutant" sequence. All sequences are shown below. The stem-swapped RNA was able to compete off 18% of bound compound (vs. 40% for the identical sequence to that immobilized on the chip); the loop mutant only competed off 4%. This suggests that the most important interactions with **8** are in the loop region of the HIV-1 FSS.



HIV-1 FSS



"stem-swapped" RNA

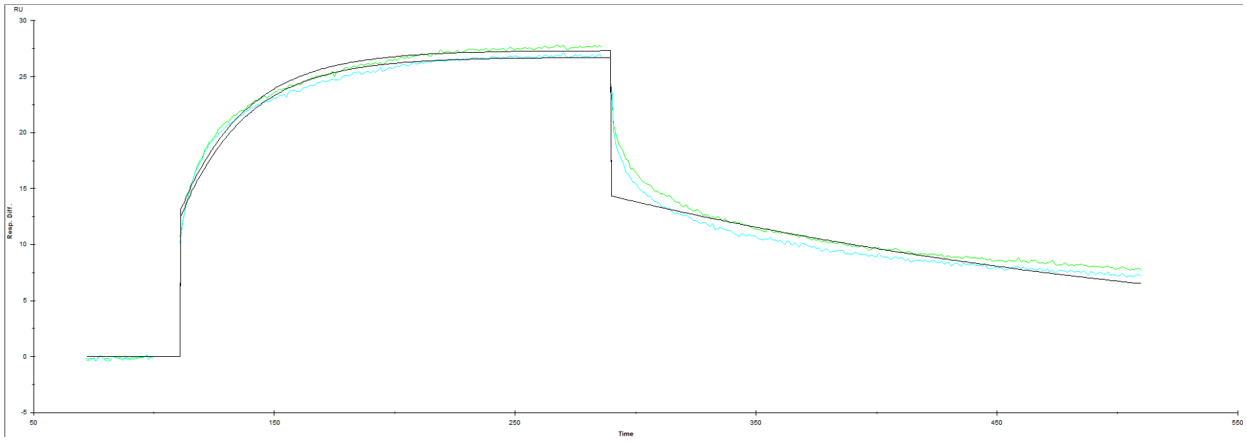


"loop mutant" RNA

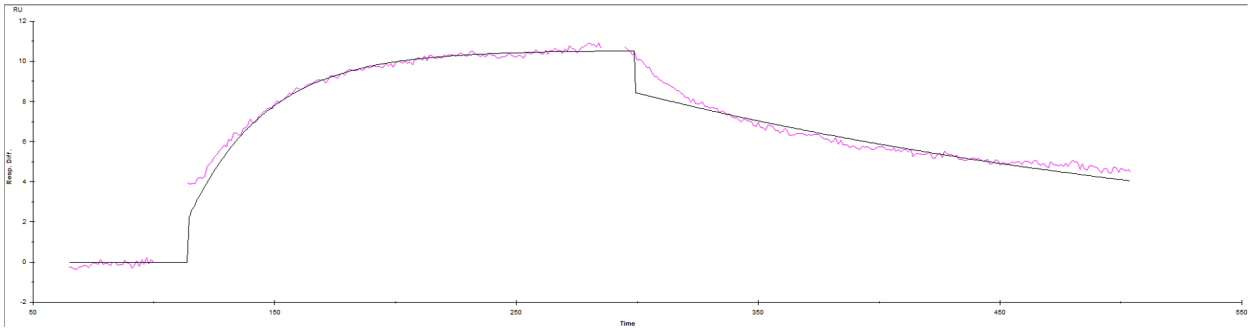
Results:

Conditions	K_D	R(max)	Δ[R(max)], %
5 μM **8**	5.04e-7	15.9	0
5 μM **8** + 25 μM HIV-1 FSS	6.5e-7	9.6	40
5 μM **8** + 25 μM "stem-swapped" RNA	6.4e-7	13	18
5 μM **8** + 25 μM "loop mutant"	6.3e-7	15.2	4

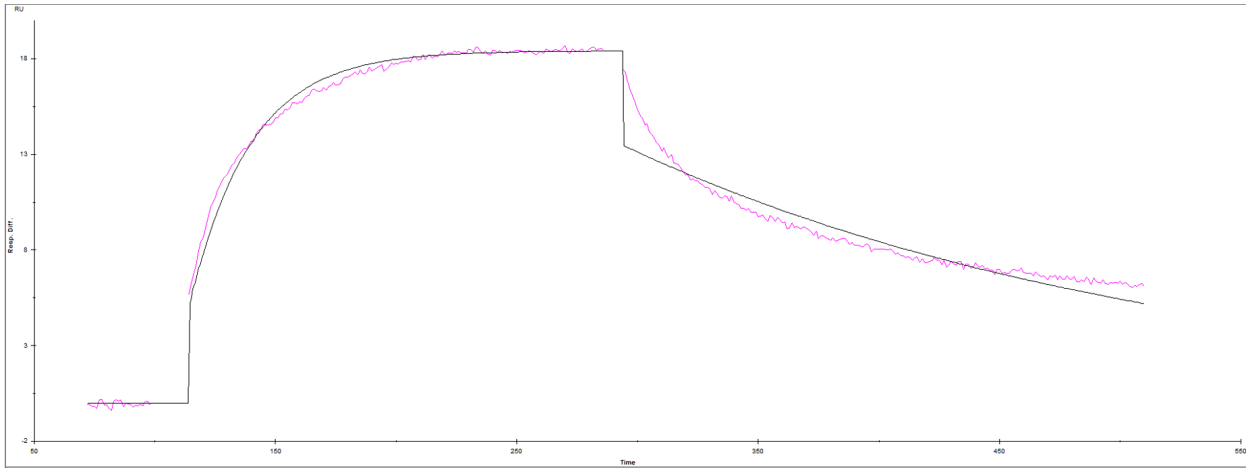
5 μM **Compound 8,** 30 μL/min, 800 RU Chip



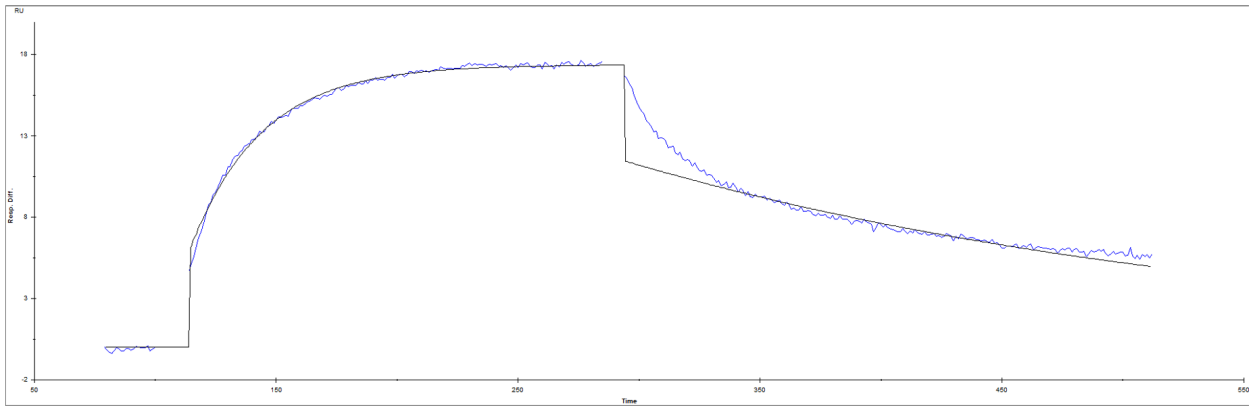
5 µM **Compound 8** + 25 µM **HIV-1 FSS RNA**, 30 µL/min, 800 RU Chip



5 µM **Compound 8** + 25 µM **Loop Mutant RNA**, 30 µL/min, 800 RU Chip

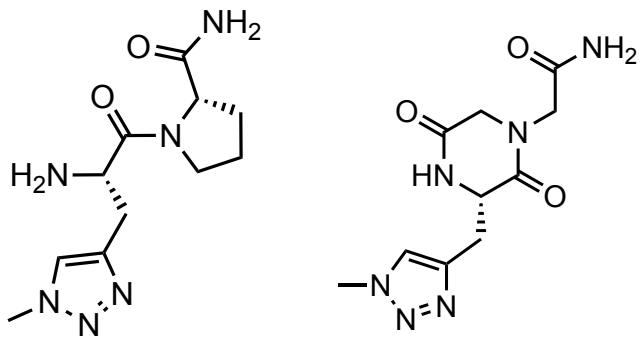


5 µM **Compound 8** + 25 µM **Stem-Swapped Mutant RNA**, 30 µL/min, 800 RU Chip



5. Conformational analysis of 9 and 10

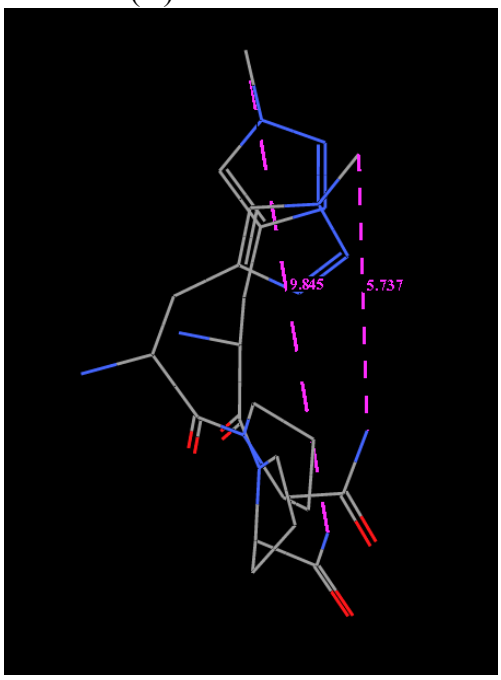
Molecular mechanics conformational analysis was used to evaluate the flexibility and preferred conformational orientations of simplified structures **9** and **10**.



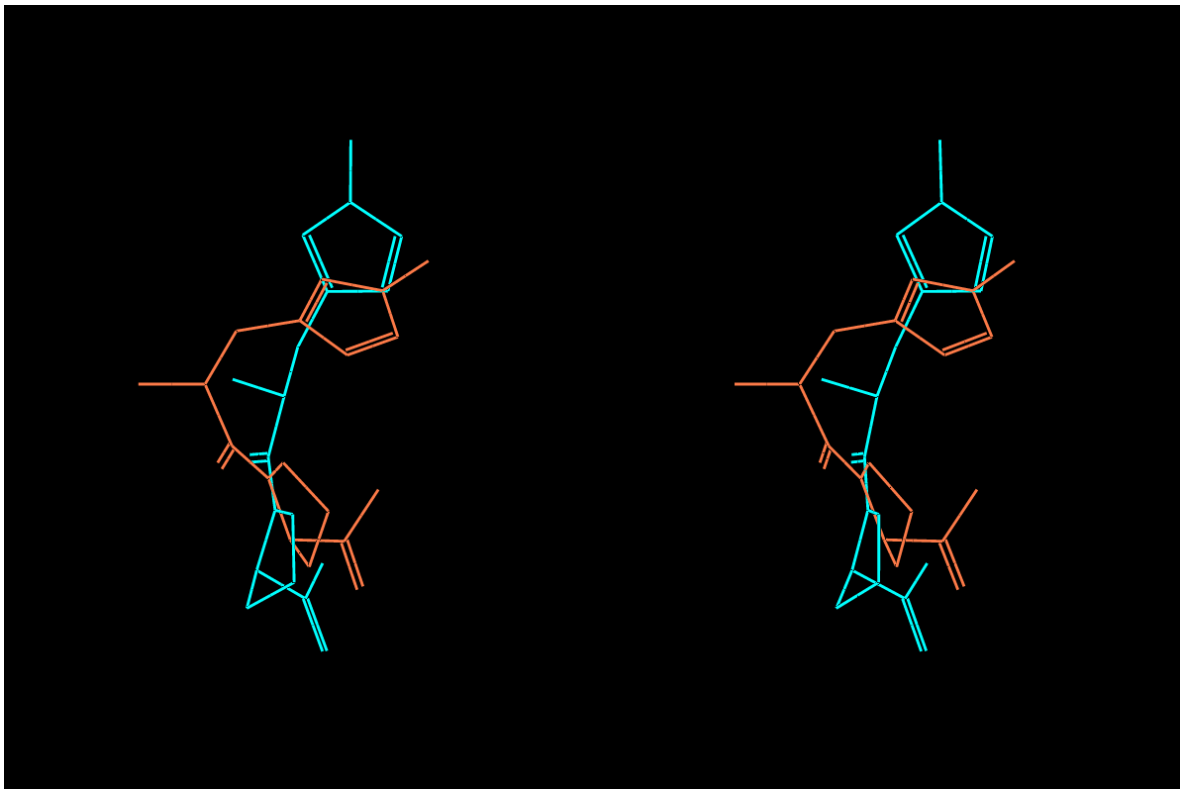
9 **10**

Local-minimum energy structures optimized by DFT methods (see main text) were used as the input in each case for the search. All calculations were conducted using version 9.1.107 of Maestro (Schrödinger, LLC) with the MMFFs* force field, and GB/SA water. 2000 conformers were generated for each compound, wither structures saved within a 3 kcal/mol window. These conditions allowed global minimum structures to be found a minimum of 10 times, indicating sufficient coverage of conformer space.

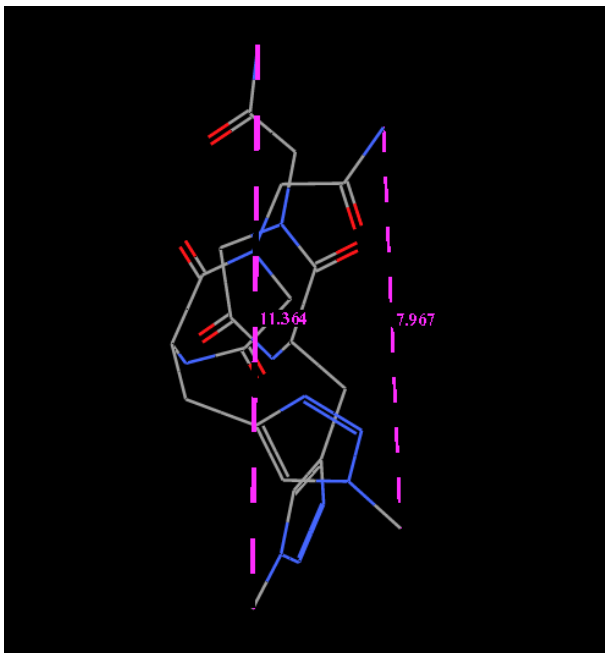
1: Dipeptide **9**, maximum and minimum energy conformers, superimposed and showing distances (Å):



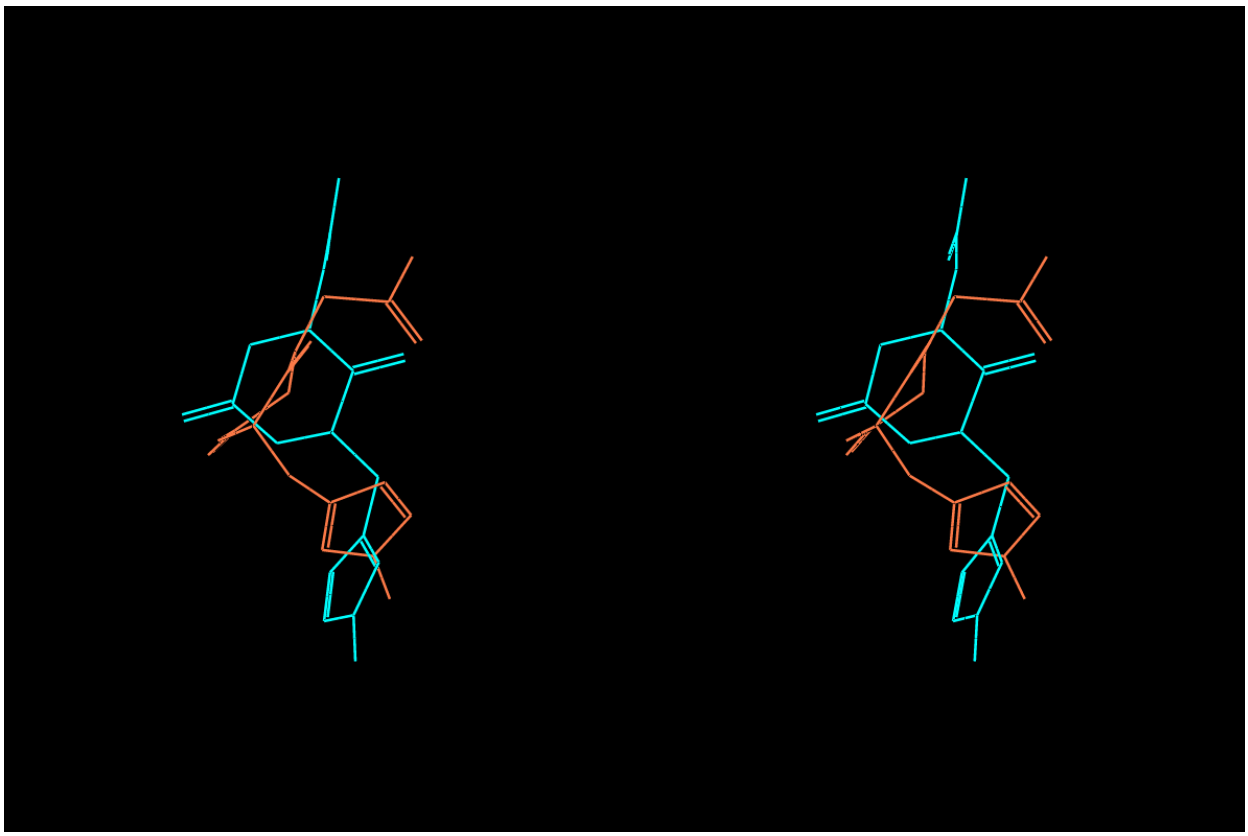
2: Dipeptide **9**, maximum (teal) and minimum (orange) energy conformers within the 3 kcal/mol window, relaxed-eye stereoview:



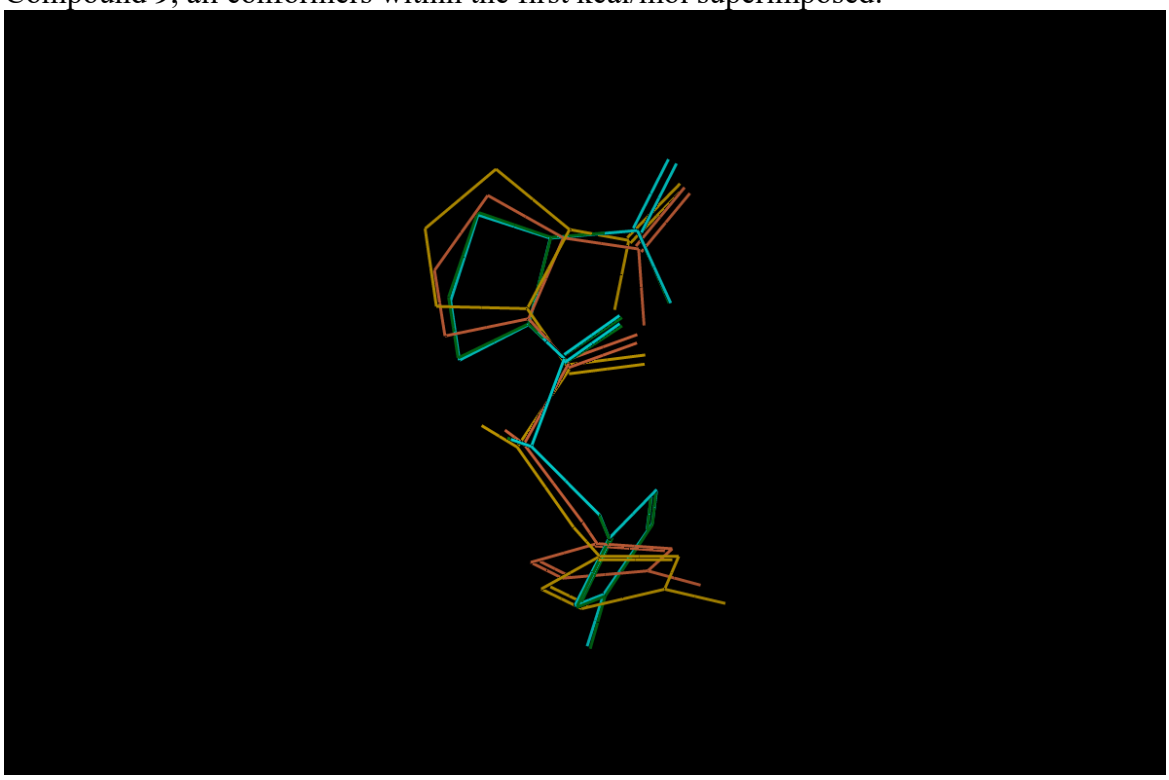
3. Compound **10**, maximum and minimum energy conformers within the 3 kcal/mol window, superimposed and showing distances (Å):



Compound **10**, maximum (teal) and minimum (orange) energy conformers within the 3 kcal/mol window, relaxed-eye stereoview:



Compound **9**, all conformers within the first kcal/mol superimposed:



Compound **10**, all conformers within the first kcal/mol superimposed: