

Supporting Information

Screening and Application of DNA Aptamers for Heparin-Binding Protein

Xi Zhou, Yingying Cao, Xiaocui Huang, Shuqian Qiu, Xinran Xiang, Huimin Niu, Li Chen, Shuiliang Wang, Zhenyu Li and Shenghang Zhang

SECTION A: SUPPORTING TABLES

Table S1. Libraries, primers and probe sequences.

Name	Sequence (5'-3')
Upstream fixed sequence	TTCAGCACTCCACGCATAGC
Downstream fixed sequence	CCTATGCGTGCTACCGTGAA
S1	FAM - TTCAGCACTCCACGCATAGC
A2	PloyA (19) - Spacer 18 - TTCACGGTAGCACGCATAGG
Apt-01(74)	TTCAGCACTCCACGCATAGCTTCACCTACCGCAATCCGTTGCTCA GTGCTACCGCACCTATGCGTGCTACCGTG
C2-13	CTTCACGGTAGCACGC
Apt-01(74)-BHQ1	TTCAGCACTCCACGCATAGCTTCACCTACCGCAATCCGTTGCTCA GTGCTACCGCACCTA/iBHQ1dT/GCGTGCTACCGTG
C2-13-FAM	CTTCACGGTAGCACGC - FAM
Padlock primer	GACTGGGCCTCAGTCCGTTTCTCATGGTTC 5'P-
Padlock	GACTGAGGCCAGTCCTTGAATTTGAC <u>GCGTGCTACCGTGAAGG</u> AACCATGAGAAACG

Note: Sequences of letters with corresponding underscores or italics are complementary.

Table S2. Control the incubation time and number of wash time of the screen to ensure the affinity of the library to the target.

SELEX rounds	Negative SELEX		Positive SELEX	
	Incubation time/min	wash times	Incubation time/min	wash times
1			60	4
2	60	4	60	4
3	60	4	60	4
4	60	4	60	4
5	60	4	60	4
6	60	4	60	4
7	60	4	60	4
8	60	4	60	4
9	60	4	45	6
10	60	4	45	6
11	60	4	45	6

Table S3. Five kinds of cDNA were designed to hybridize to each of the five regions of Apt-01 to explore the regions that bind to aptamer. Sequences underlined are the Apt-01 complementary sites to the cDNA.

cDNA name	Sequence (5'-3')	Aptamer (5'-3')
C1	TTCACGGTAG-FAM	TTCAGCACTCCACGCATAGCTTCACCTACCGCAATCCGTTGCTCAGTGCTACCGCACC/iBHQ1dT/ATGCGTGCT <u>TACCGTGA</u> A
C2	GGTAGCACGC-FAM	TTCAGCACTCCACGCATAGCTTCACCTACCGCAATCCGTTGCTCAGTGCTACCGCACC/iBHQ1dT/ATGCGTGCT <u>TACCGTGA</u> A
C3	CATAGGTGCG-FAM	TTCAGCACTCCACGCATAGCTTCACCTACCGCAATCCGTTGCTCAGTGCTACCGCACC/iBHQ1dT/ATGCGTGCTACCGTGA A
C4	ATTGCGGTAG-FAM	TTCAGCACTCCACGCATAGCT/iBHQ1dT/CACCTACCGCAATCCGTTGCTCAGTGCTACCGCACCTATGCGTGCTACCGTGA A
C5	TGAAGCTATG-FAM	TTCAGCACTCCACGCATAGCT/iBHQ1dT/CACCTACCGCAATCCGTTGCTCAGTGCTACCGCACCTATGCGTGCTACCGTGA A

Note: Sequences of letters with corresponding underscores complement each other.

Table S4 High-throughput sequencing top 100 sequence results.

Number	Random sequence (5'-3')	R11
1	TTCACCTACCGCAATCCGTTGCTCAGTGCTACCGCA	11404
2	TTCACCTACCGCGCAGCTCAGTCGTTGCCCACACGA	11608
3	TTCACCTACGAACGCCAGTGACAGCACCGATCAGTG	9823
4	TTCATCTACCCCGCCAAGGTCAGCCACGTCAGTGTG	8127
5	TTCACCTACACCGCGACCGCGTTTCAGCCTCTGCAGG	8966
6	TTCATCTACCGCGTTAGCTCAGTCGTCTCGTAACGA	8945
7	CTTGCCTACCGGACGAAGTACTCAGTCACTCCGTGA	5902
8	TCATCGCGAGCGACACGTTACCGTGCTTTACTACCG	8462
9	CTTGCCTACCGATCCTGGATCTCAGCTGTACCACAG	7003
10	CTTGCCTACCCACAGCGTGTAGTGTACTGCGTACA	2379
11	TTTACTACCGCGCACTTAGTCGCTCAGCCGTTCCGG	5812
12	TTCACCTACTGGCATATGCCTCAGTCCACTCCTGGA	5850
13	TCATCGCCTGCGACATGCTAGCATGCTTTACTACTG	5695
14	TTCACCTACCGCCAATCAGCTCAGTGTGTCTCCACA	3640
15	TTCACCTACTGCCCTAGCTCAGCCGTCTCCATCCGG	5248
16	TTTACCTACCGCGCAGCTCAGTCGTTGCCCACACGA	5810
17	CTTGCCTACCGACAACCGTCTCAGTTCGTCATCGAA	4828
18	TTTACCTACCGCATCGCTCAGTCAACCATTTGGTGGA	4631
19	TCATGCCCTGTGCACGGCCCAACCGCTTTACCTACG	3666
20	TTCACCTACCGCCGAAGGCTCAGTCCGTGCTTCGGA	6199
21	TTTACCTACTGCCACAACCTGGGGCTCAGCCGCGAGG	5609

22	TCATCCGGATCGGACACATCTTGTGCCTTGCTACTG	2286
23	TTCACCTACCGCCCTGCTCAGCCGACTCGGATCGG	5004
24	TTCATCTACCGCGAGCAGCTCAGCATGCCCAGCATG	3454
25	TTCACCTACTGCCACAACCTGGGGCTCAGCCGCGAGG	2844
26	TTCATCTACCGTGCCTCTCGCACTCAGCCCCTGTGG	4138
27	TCATCCCATGGGACCGTTATACGGCTTACCTACCG	2718
28	TTCATCTACTGCGACTAGCTCAGCCTATCCACTAGG	3825
29	TTCATCTACCGCACCAGCTCAGCCGACCTCCGGAGG	3514
30	CTTGCTACCGGGAACCTCAGTCAGTGTGTCAACTGA	2449
31	TTCACCTACTGCCTCGGCTCAGATTGATTCCGAAT	2987
32	TTCACCTACCGCCTTGCTCAGCCCATTGGAATGGG	2217
33	TTCACTACCGCGCTAAGCTCAGTTTGGGTGACCAAA	1733
34	CTTGTCTACCCGGGAACCGTCAGTCCGTCACACGGA	1824
35	TTTACCTACCGCAATCCGTTGCTCAGTGCTACCGCA	2915
36	TTCATCTACCGCCCCAGCTCAGCCGACCTCCGGAGG	3092
37	CTTGCTACCGTCCTTGACTCAGTTCAGCGCCGTGAA	1275
38	CTTGCCTACCGGGAGAATCCCTCAGACACAGCGTGT	1278
39	TTTACTACGACCACCGGGACAGATTGTGCAGACGAT	1667
40	TTCATCTACCGGTCGCGAGACCTCAGTCACACGTGA	1914
41	CTTGCTACGAGCCTGCGACAGTCGCAAACCTTTCGA	1741
42	TTTACTACCGCCCTGGCTCAGATGGTTCCCGGCCAT	1319
43	CTTGCCTACCCACACCGTGTCTCAGTCTGTACGTCAGA	2596
44	TTCACCTACGAGCGCAGTGACAGTGGTCGCAGTCCA	1449
45	TTTACCTACCGCCCTGCTCAGCCGACTCGGATCGG	2896
46	TTTACCTACCGCACCTCGGGCTCAGTTGCATGACAA	2289
47	TTTACCTACCGCACCCGCTCAGTCGTCTCCGCGA	2777
48	TTCACCTACCGCACCTCGGGCTCAGTTGCATGACAA	1007
49	TTCACCTACCGCTACCCTGCTCAGCTGCCAAAGCAG	2220
50	CTTGCCTACCGACATTGTTGTCTCAGTTCCTACGAA	1579
51	CTTGCCTACTGCCTATCGGGCTCAGTCGTTCTCGA	1850
52	TTCACCTACCGCTGTGATTAGCAGCTCAGCCATCGG	2664
53	TTCATCTACCGCGCAGCTCAGTCGTTGCCACACGA	2252
54	TTCACCTACCGCGCAGATCGCTCAGCCACTCATTGG	1930
55	TTTACTACCGCCACATGCTGGCTCAGTCCCTGAGGA	1624
56	CTCGCCTACCGACAACCGTCTCAGTTCGTCATCGAA	811
57	TTCACTACCGCGCCAGCTCAGTCGGACCAGCTGCGA	944
58	TTTACCTACCGCCGTGCTCAGTTGGACTTGTCCCAA	1772
59	CTTGCCTACCGCGCCAGCTCAGGGAACGTACCGTTGC	1536
60	CTTGCTACACCGACAACGTGTTTCAGTGGTCCACCA	955
61	CTTGTCTACCCGCCTACTGGACGTCAGTCGTTGAG	837
62	TCATTGCCCCGCAACAACCTCGTTTGCTTCATCTACG	1240
63	TTTACCTACTGCGGAACGCTCAGACGATGCCACCGT	1270
64	TTTACCTACCTCGCAGTCCGCCCCGCTTGGCATCGGA	861
65	CTTGCTACCGTCGCCGACTCAGTTGGATTGCCCCAA	894

66	CTTGCTACCTCGACAGATCAGTCACACGCCCTGGGA	641
67	CTTGCCTACCGCATCACATGCTCAGTGTCCGCGACA	2066
68	TTTACCTACCGCGATCGCTCAGTCCGCCTTGCTGA	2315
69	CTTGCCTACTGGCCCCGCTGCCTCAGTGTTCTAACA	1038
70	CTTGCCTACCGTCCTGAACTCAGTCTCGTTCCGAGA	851
71	TTTACCTACCGCCGTGGCTCAGTTGTCTAGGAACAA	1333
72	CTTGCTACTCGGAACGTCAGCTCGCACTCCCGCGAG	560
73	TCATCGTCCCGGACGACGTCCCGCCTTGTCTACCG	932
74	TTTACCTACCGCGATCAGCTCAGTCTCACCGTGAGA	2181
75	CTTGCTACGAGTCTGAACGACAGTCCGGTTGACGGA	630
76	TTTACCTACCGCCAATCAGCTCAGTGTGTCTCCACA	1545
77	TCATCCGCTCGGACGGTGACCACCGCTTTACCTACG	2025
78	TTTACCTACTGCCCCGCTCAGTCACGCCACCGTGA	2018
79	CTTGCTACCCGCGAATAGCGTCAGTCGGCCTTGCGA	341
80	CTTCCTGGTGTCTGGGCAGTTATAACGGTGCGTCTGA	423
81	CTTGCCTACCCCGGTCTCGGTCTCAGTCGTTGTGTCGA	330
82	CTTGCCTACCGATCTAGGATCTCAGTGTTGCCGACA	783
83	TTTACCTACCGCGAGGTACCGACGCTCAGCCTTTGG	1743
84	TTTACCTACCGCGGATTCTGCTCAGTCTCCAGAGA	1632
85	TTCACCTACTGCCTTAGCTCAGGTCTGCATACTGAC	889
86	TTTACCTACCGCACCGACGTCGCTCAGCCCTTGGG	1785
87	TTCATCTACCGCATCCCGCTCAGTCGGCCGGACCGA	1483
88	TCATCGGCACGACCGCTATTGCTGCTTTACCTACCG	1479
89	CTTGCTACCGTTCCCTACCCTGAACTCAGTCCGTGGA	455
90	CTTGCCTACCGCGACTCGATCTCGCTCAGTCTCTGA	911
91	CACTTCACCTACGACCACCGGGACATCGCGCTCGATG	482
92	TTCACCTACTGCCTAGCTCAGCCGTCTCCATCCGG	1590
93	TTCATCTACCGGTCCCACAGACCTCAGTGCCTCGCA	1278
94	TTTACCTACCGCACACTGTTGCTCAGTCGCTTTCGA	1680
95	TTCACCTACCGCGCCAGCTCAGGGAACGTACCGTTGC	1355
96	TTTACTACCGTGGCCACACTCAGTTGCAGTTCGCAA	942
97	TTCATCTACGACCTCCCGGGACAGTCACGCCTGTGA	1049
98	TTCACCTACCGCGTTAGCTCAGTCGTCTCGTAACGA	1315
99	TTTACCTACCTCGCAGTCTGTCAGCACGTTGCCAGA	737
100	CTTGCTACCGAGCCGTGCTCTCAGTTCCTTCCAGAA	687

Note: R11 denotes the number of corresponding sequence entries in the eleventh round of screening libraries.

Table S5 Stability and affinity characterization of 14 candidate aptamers.

Aptamer	Full-length sequence (5'-3')	Family	ΔG (kcal/mol)	ΔH (kcal/mol)	ΔS (kcal/mol)	T_m (°C)	RU
Apt-01	TTCAGCACTCCACGCAT AGCTTCACCTACCGCAA TCCGTTGCTCAGTGCTA CCGCACCTATGCGTGCT ACCGTGAA	5	-12.48	-145.20	-445.15	53.0	424.1
Apt-02	TTCAGCACTCCACGCAT AGCTTCACCTACCGCGC AGCTCAGTCGTTGCCCA CACGACCTATGCGTGCT ACCGTGAA	5	-11.43	-134.50	-412.78	52.7	161.7
Apt-03	TTCAGCACTCCACGCAT AGCTTCACCTACGAACG CCAGTGACAGCACCGAT CAGTGCCTATGCGTGCT ACCGTGAA	3	-11.67	-131.10	-400.57	54.1	118.6
Apt-04	TTCAGCACTCCACGCAT AGCTTCATCTACCCCGC CAAGGTCAGCCACGTC AGTGTGCCTATGCGTGC TACCGTGAA	3	-9.16	-139.80	-438.17	45.9	121.1
Apt-05	TTCAGCACTCCACGCAT AGCTTCACCTACACCGC GACCGCGTTCAGCCTCT GCAGGCCTATGCGTGCT ACCGTGAA	2	-12.07	-138.80	-425.05	53.4	132.2
Apt-07	TTCAGCACTCCACGCAT AGCCTTGCCTACCGGAC GAAGTACTCAGTCACTC CGTGACCTATGCGTGCT ACCGTGAA	6	-10.24	-143.60	-447.29	47.9	76.2
Apt-08	TTCAGCACTCCACGCAT AGCTCATCGCGAGCGAC ACGTTACCGTGCTTTAC TACCGCCTATGCGTGCT ACCGTGAA	1	-12.14	-129.80	-394.63	55.8	79
Apt-09	TTCAGCACTCCACGCAT AGCCTTGCCTACCGATC CTGGATCTCAGCTGTAC CACAGCCTATGCGTGCT ACCGTGAA	6	-11.14	-141.60	-437.57	50.5	87.5

	TTCAGCACTCCACGCAT						
	AGCTTCACCTACTGGCA						
Apt-12	TATGCCTCAGTCCACTC	7	-13.07	-144.00	-439.14	54.8	107.2
	CTGGACCTATGCGTGCT						
	ACCGTGAA						
	TTCAGCACTCCACGCAT						
	AGCTCATCGCCTGCGAC						
Apt-13	ATGCTAGCATGCTTTACT	1	-11.76	-129.10	-393.56	54.9	169.4
	ACTGCCTATGCGTGCTA						
	CCGTGAA						
	TTCAGCACTCCACGCAT						
	AGCTTCACCTACCGCCA						
Apt-14	ATCAGCTCAGTGTGTCT	4	-11.34	-129.50	-396.31	53.6	50.3
	CCACACCTATGCGTGCT						
	ACCGTGAA						
	TTCAGCACTCCACGCAT						
	AGCTTCACCTACTGCCC						
Apt-15	TAGCTCAGCCGTCTCCA	2	-11.21	-125.00	-381.65	54.4	72.5
	TCCGGCCTATGCGTGCT						
	ACCGTGAA						
	TTCAGCACTCCACGCAT						
	AGCTTTACCTACCGCAT						
Apt-18	CGCTCAGTCAACCATTG	7	-9.57	-103.30	-314.37	55.4	41.2
	GTGGACCTATGCGTGCT						
	ACCGTGAA						
	TTCAGCACTCCACGCAT						
	AGCTTCACCTACCGCCG						
Apt-20	AAGGCTCAGTCCGTGCT	4	-15.44	-146.80	-440.58	60.0	99
	TCGGACCTATGCGTGCT						
	ACCGTGAA						

Note: The Gibbs free energy showed that the 14 selected candidate aptamers were all stable with little variation between sequences, but seven candidate aptamers showed response values (RU) of more than 100 RU to HBP, of which Apt-01, Apt-02, and Apt-13 showed response values of more than 150 RU.

SECTION B: SUPPORTING FIGURES

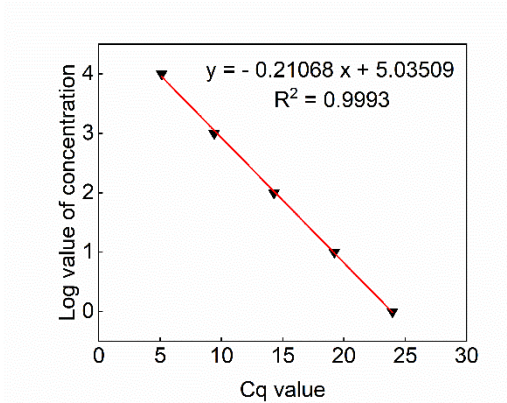


Figure S1. Characterization of the library. Standard curves for initial libraries alongside log values of their concentrations.

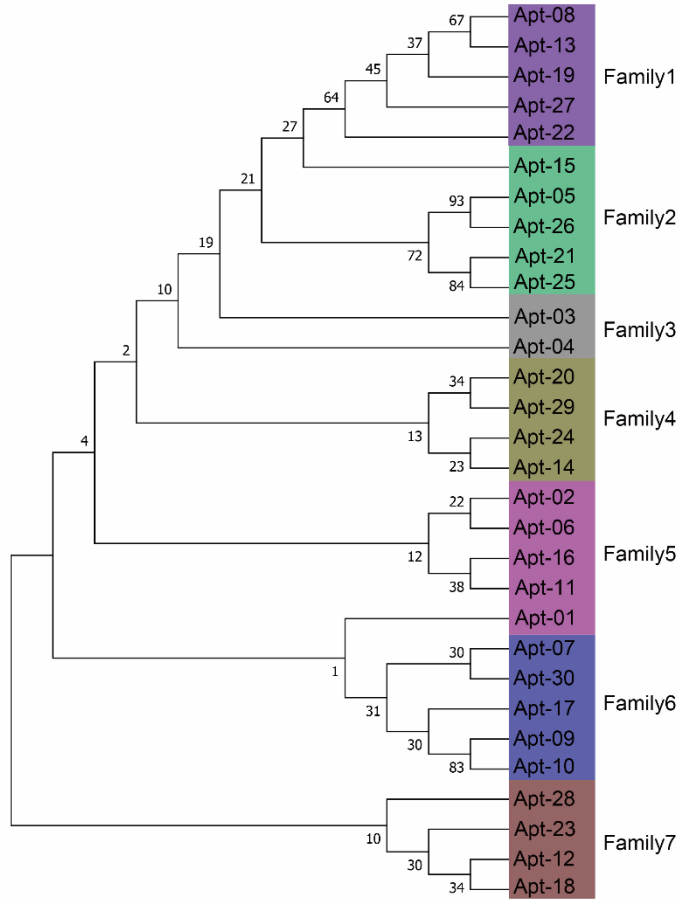


Figure S2. Following sequencing, the initial 30 sequences were examined for homology and separated into seven families according to their homolog.

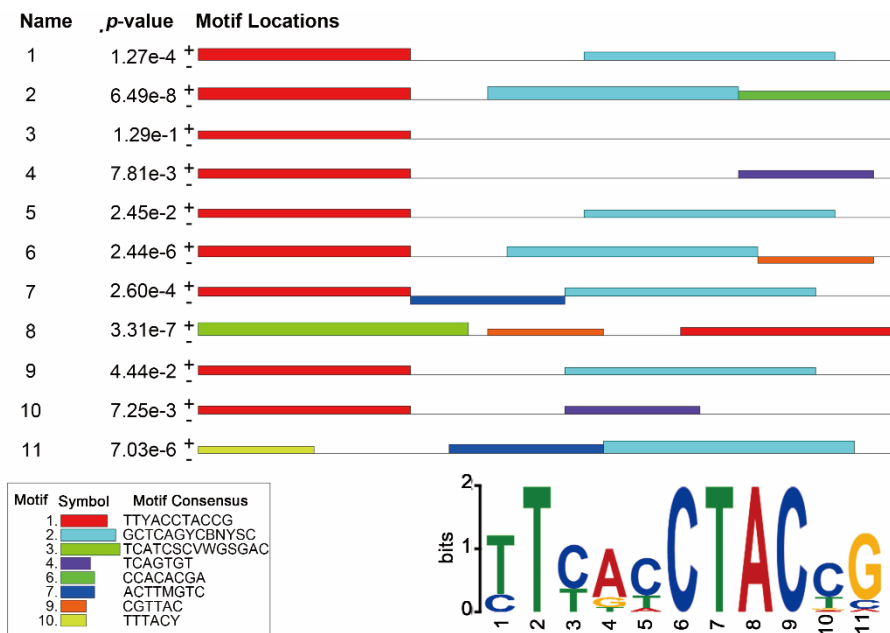


Figure S3. Conserved and variant site assessment of the top 30 sequences with comparatively high abundance.

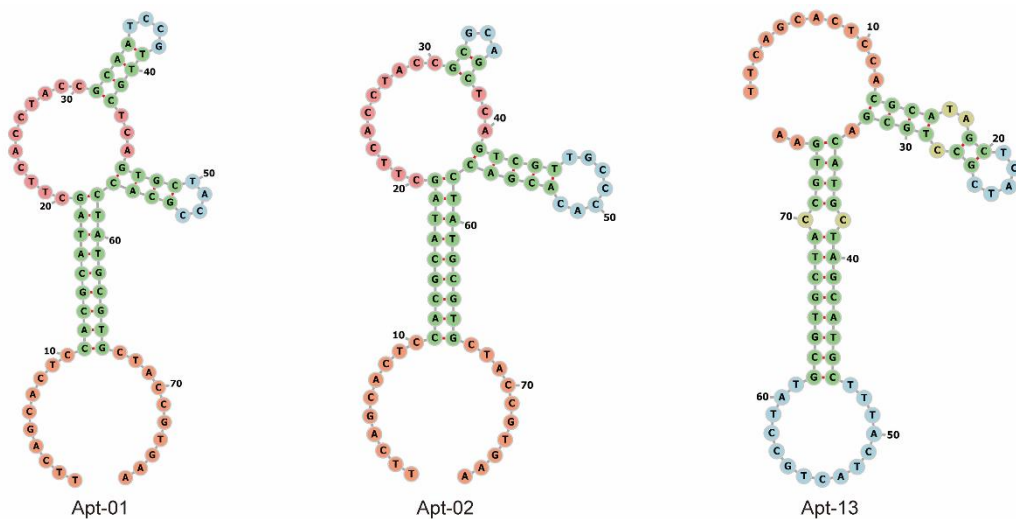


Figure S4. Modeling of the secondary structure of three aptamers using Mfold.

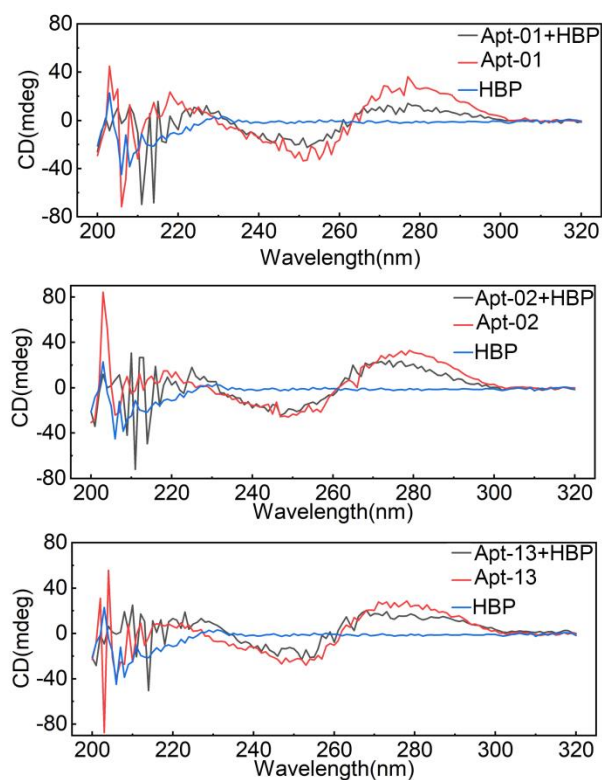


Figure S5. Aptamer and HBP of CD spectra in PBS buffer..

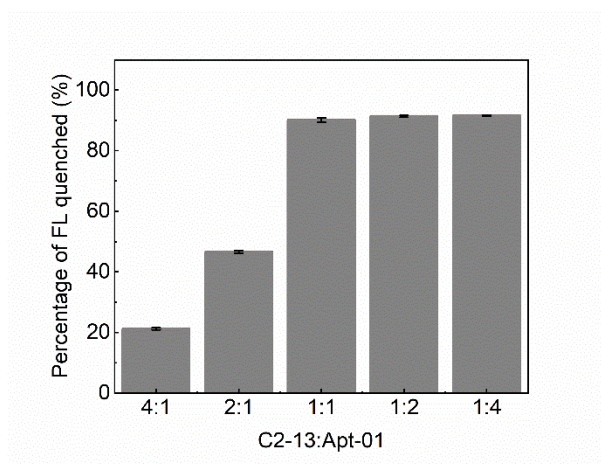


Figure S6. The FRET effect occurred following the hybridization of FAM modified C2-13 with BHQ1 modified Apt-01.

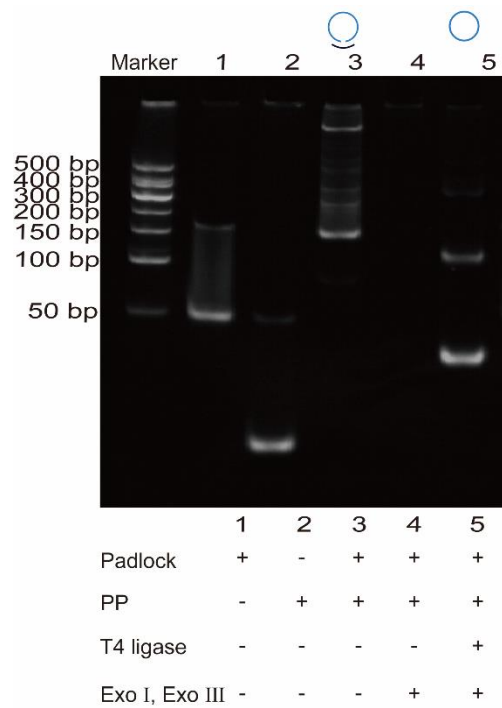


Figure S7. Characterization of CT formation by polypropylene gel electrophoresis.