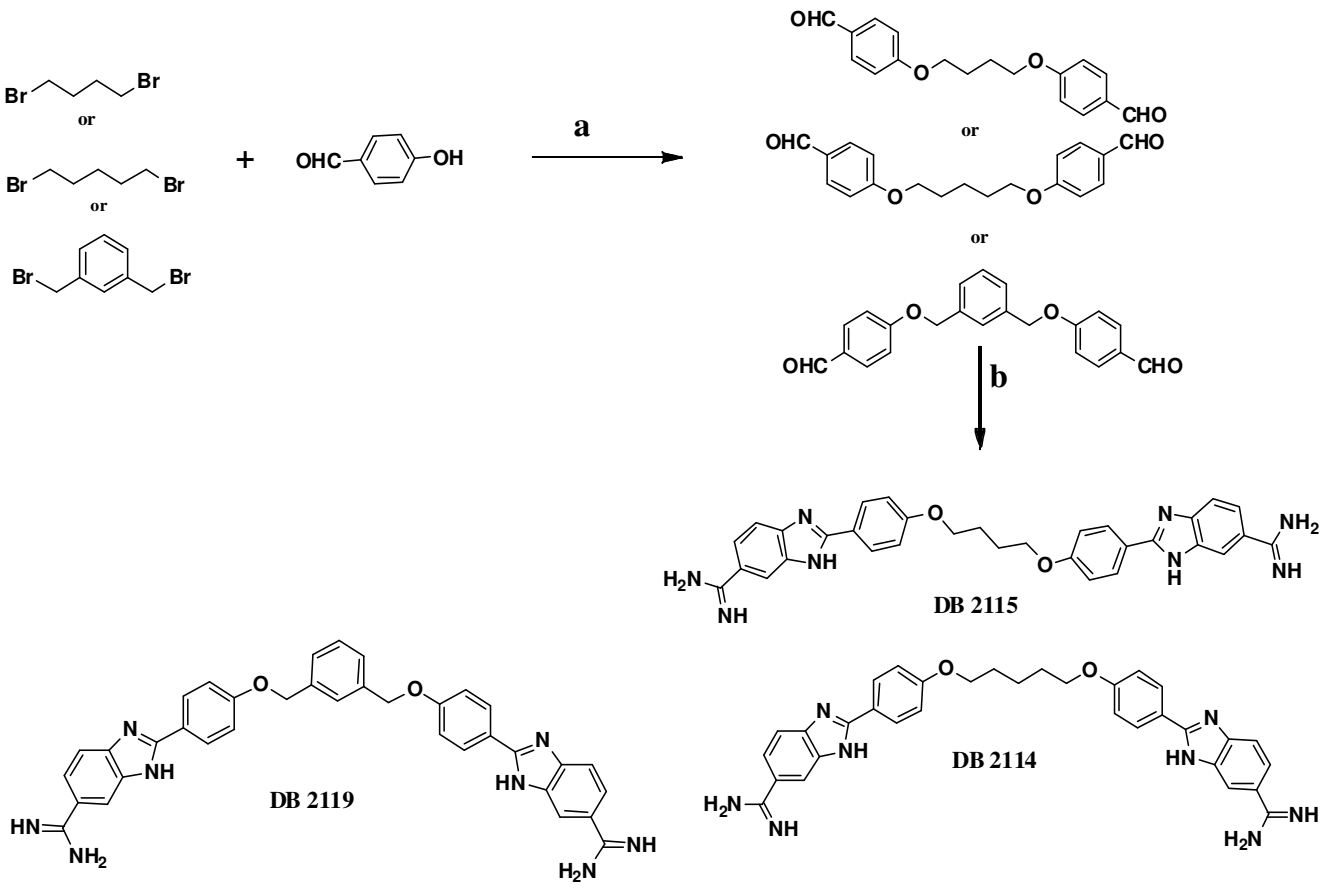
Designed compounds for recognition of ten base pairs of DNA with two AT binding sites

Yang Liu[†], Yun Chai[†], Arvind Kumar[†], Richard R. Tidwell[£], David W. Boykin[†], W. David Wilson[†]*

[*]To whom correspondence should be addressed. E-mail: wdw@gsu.edu

Contents

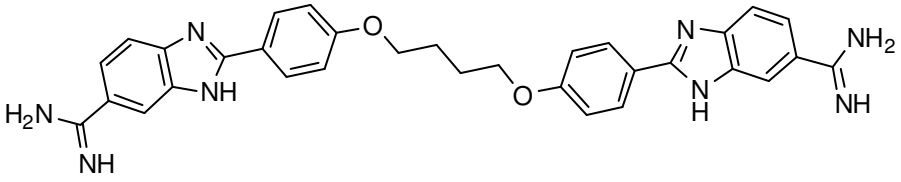
Synthesis of Linked Benzimidazoles



a) K2CO3 anhyd., DMF anhyd., 80^{0}C, 12h; b) 4-amidino-1,2-phenylenediamine.HCl, 1,4-benzoquinone, anhd. ethanol, 80^{0}C

Scheme S1: Synthesis of symmetric compounds.

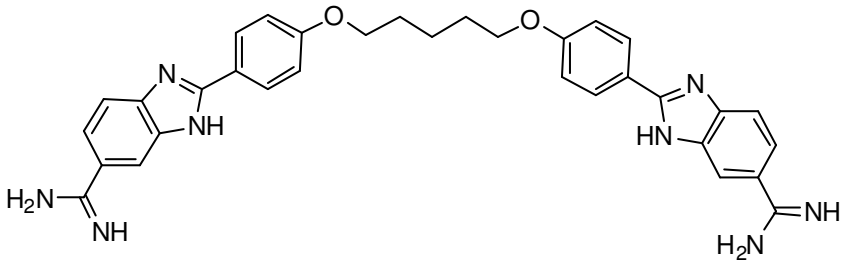
Synthesis procedure and compound characterization:



DB2115:
1,4-Bis{4[4(5)amidino-benzimidazolyl]phenoxy]}-butane tetrahydrochloride

A solution of 4,4"-[butane-1,4-diylbis(oxy)] dibenzaldhyde[1]

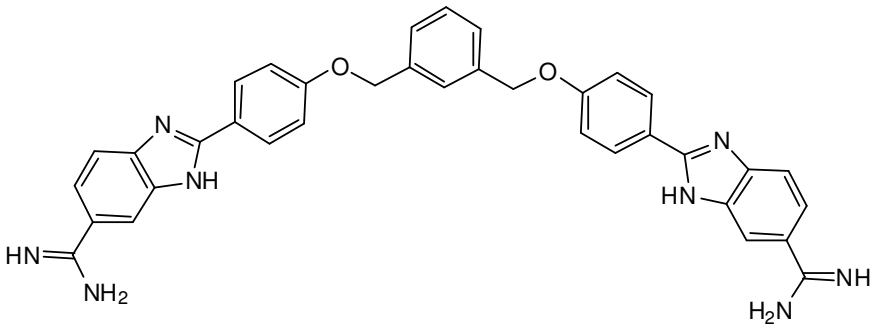
(0.2 g, 0.67 mmol), 4-amidino-1,2-phenylene diamine hydrochloride (0.25 g,1.34 mmol), and 1,4-benzoquinone (0.145 g, 1.34 mmol) in ethanol (30 ml) was heated at reflux for 8 h (under nitrogen). The reaction mixture was cooled to room temperature and the dark solid was collected by filtration, washed with dry acetone, cold ethanol and anhydrous ether, and dried to yield 0.3 g (71%) of the dihydrochloride. This solid was dissolved slowly in hot ethanol (300 ml) and filtered. The filtrate volume was reduced to 70 ml and was acidified with HCl-saturated ethanol. After stirring overnight and diluting with dry ether the purple solid was collected by filtration, washed with anhydrous ether, and dried under vacuum to yield 0.34 g (72% yield) of a purple tetrahydrochloride salt; mp >300°C dec. ^{1}H NMR (DMSO-d6): 9.40 (s, 4H); 9.11 (s, 4H), 8.34 (d, 4H, *J* = 1.5 Hz), 8.2(s, 2H), 7.84 (d, 2H, J = 5Hz), 7.78 (dd, 2H, J = 5Hz, J = 1.5), 7.22 (d, 4H, J=8.5Hz), 4.23 (t, 4H, J=6.5Hz), 1.97 (t, 4H, J=6.5Hz). MS (FAB): *m/z* 559 (M^++1); HRMS: calculated mass (free base), 559.6333 (M^++1); observed mass, 559.6336. Anal. calculated for $C_{32}H_{30}N_8O_2$. 4HCl: C, 54.55; H, 4.86; N, 15.90; Found: C, 54.62; H, 4.79; N, 15.76.



DB2114:
1,5-Bis{4[4(5)amidino-benzimidazolyl]phenoxy]}-pentane tetrahydrochloride

Reaction of 4,4"-[pentane-1,5-diylbis(oxy)]dibenzaldehyde[2]

(0.21g ,.67mmole) , 4-amidino-1,2-phenylenediamine hydrochloride (0.25g, 1.34 mmol), and 1,4-benzoquinone (0.145g, 1.34 mmol) in ethanol (30 ml) was heated at reflux for 8h (under nitrogen) following standard workup and treating with HCl-saturated ethanol yielded 0.37g (75%) violet tetrahydrochloride salt; mp >320°C dec. ^{1}H NMR (DMSO-d6): 9.47 (s, 4H); 9.21 (s, 4H), 8.40 (d, 4H, *J* =8.5 Hz), 8.22 (d, 2H, J=1.5Hz), 7.87 (d, 2H, J=8.5Hz), 7.82 (dd, 2H, J=8.5Hz, J=1.5Hz), 7.22 (d, 4H, J=8.5Hz), 4.18 (t, 4H, J=6.5Hz), 1.87 (quint., 4H, J=6.5Hz), 1.64 (t, 2H, J=6.5Hz). MS (FAB): *m/z* 573 (M^++1); HRMS: calculated mass (free base), 573.6678 (M^++1); observed mass, 573.686. Anal. calculated for $C_{33}H_{32}N_8O_2$. 4HCl: C, 55.16; H, 5.05; N, 15.59; Found: C, 55.19; H, 4.89; N, 15.48.



DB2119:
1,3-Bis{4[4(5)amidino-benzimidazolyl]phenoxy]}-xylene tetrahydrochloride

Reaction of 4,4'-[1,3-phenylenebis(methylene)]bis(oxy)dibenzaldehyde[3] (0.23 g, 0.67 mmole), 4-amidino-1,2-phenylenediamine hydrochloride (0.25g, 1.34 mmol), and 1,4-benzoquinone (0.145g, 1.34 mmol) in ethanol (30 ml) was heated at reflux for 8h (under nitrogen),following standard workup and treating with HCl-saturated ethanol yielded 0.40g (78%) purplish grey tetrahydrochloride salt; mp >340°C dec. ^{1}H NMR (DMSO-d6): 9.34 (s, 4H); 9.07 (s, 4H), 8.31 (d, 4H, *J* =8.5 Hz), 8.17 (s, 2H), 7.81 (d, 2H, J=8.5Hz), 7.74 (d, 2H, J=8.5Hz), 7.62 (s, 1H), 7.48 (br, 3H), 7.27 (d, 4H), 5.29 (s, 4H); MS (FAB): *m/z* 607 (M^++1); HRMS: calculated mass (free base), 607.6840 (M^++1); observed mass, 607.6969. Anal. calculated for $C_{33}H_{30}N_8O_2$. 4HCl: C, 56.77; H, 4.63; N, 14.71; Found: C, 56.62; H, 4.79; N, 14.62.

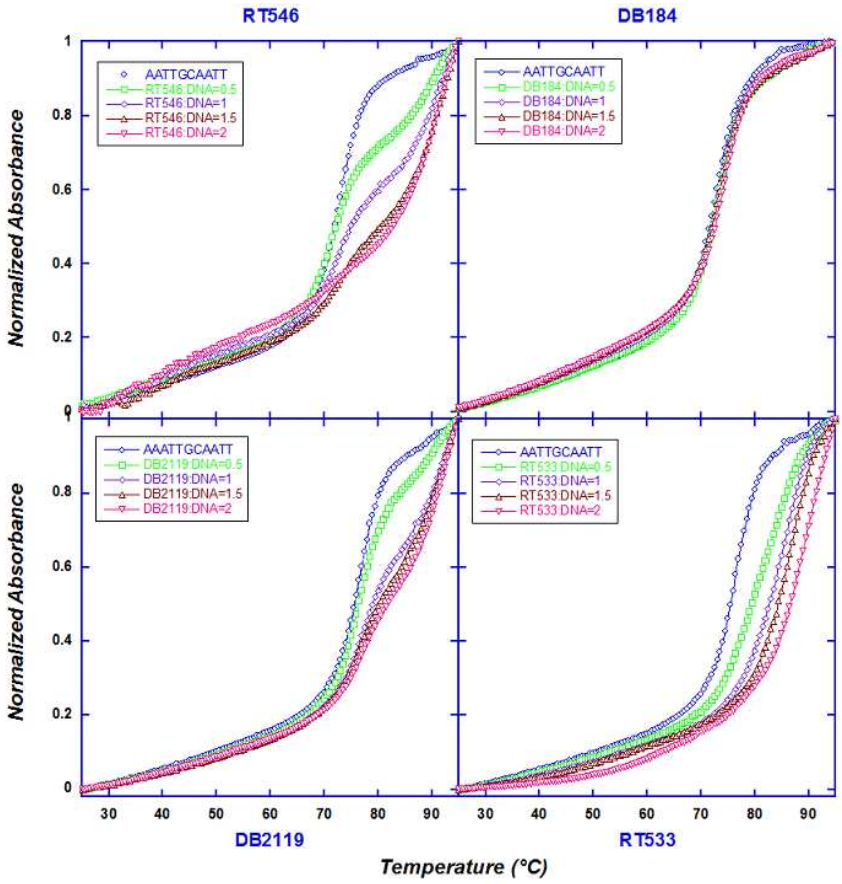
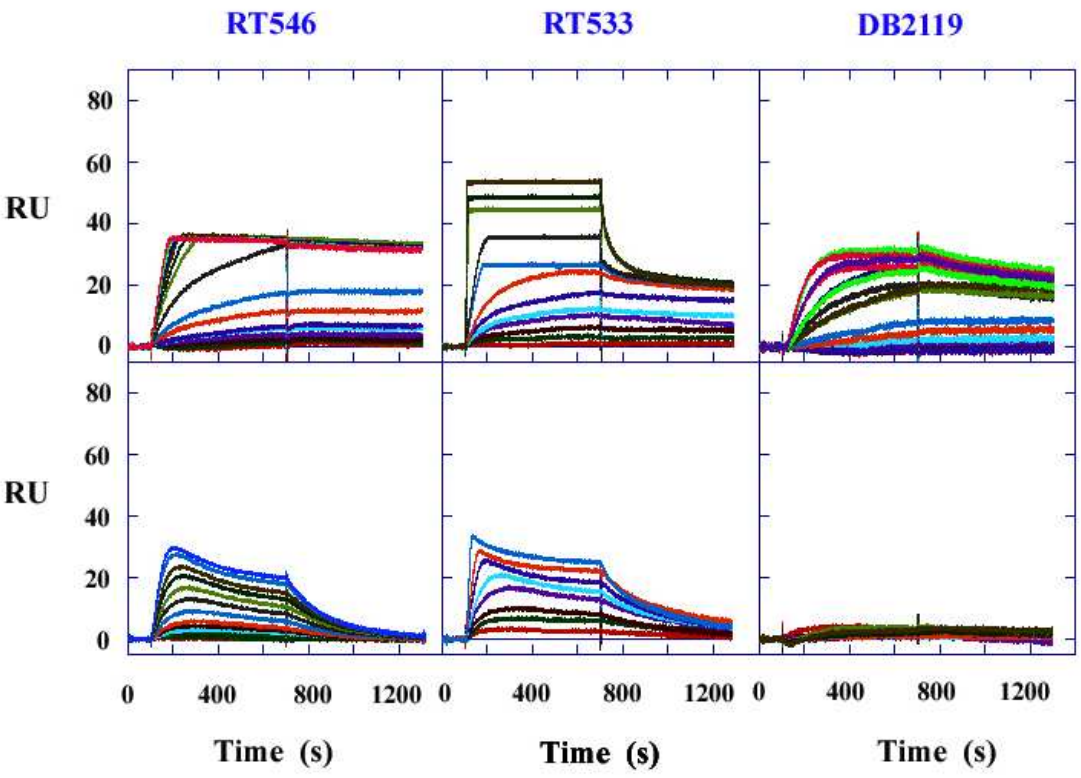


Figure S1. UV melting profiles at 260nm of the AATTGCAATT hairpins in the absence and presence of compounds. (A) Thermal melting curves of RT546 at the indicated ratio of compound per DNA hairpin duplex. (B) Thermal melting curves of DB184 at the indicated ratio of compound per DNA hairpin duplex. (C) Thermal melting curves of DB2119 at the indicated ratio of compound per DNA hairpin duplex. (D) Thermal melting curves of RT533 at the indicated ratio of compound per DNA hairpin duplex.

Thermal Melting: Relative affinity ranking. DNA thermal melting studies provide a rapid, qualitative method for ranking compounds according to their binding affinity[4] and provide a test for SPR conclusions. Large increases in melting temperature are observed for RT546 binding with the AATTGCAATT sequence. In the unbound state the AATTGCAATT has monophasic melting curves but exhibits biphasic curves below the

saturation ratio and monophasic curves above the 2:1 ratio of RT546. At the 1:1 ratio, the two phases have approximately equal intensity as would be expected for a 2:1 stoichiometry. The biphasic melting is characteristic of compounds that bind very strongly to DNA and indicate that RT546 binds to the AATTGCAATT sequence as a strong dimer. DB2119 also has biphasic curves in agreement with its positive binding cooperativity. DB184, the compound without a linker, only has monophasic melting curves up to a ratio of 2:1 and much smaller changes in the ΔT_m values. RT533 has a monophonic curve in agreement with its lack of binding cooperativity. These observations are in excellent agreement with the SPR results.

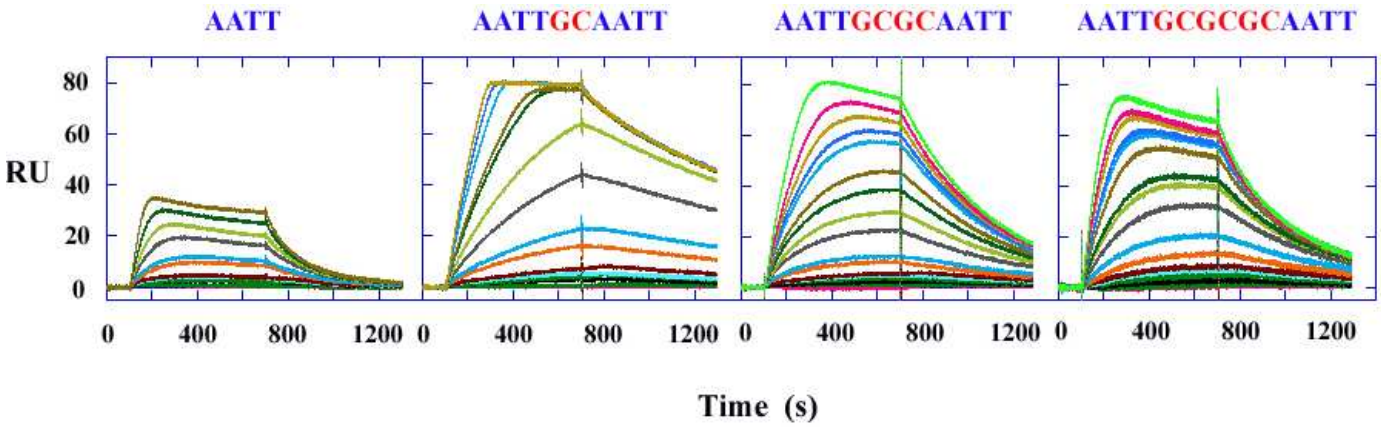


DNA	RT546 K^a (mol^{-1})	RT533 K^a (mol^{-1})	DB2119 K^a (mol^{-1})
AAAAATTTTT	$K = 9.3 \times 10^7$	$K_1 = 1.0 \times 10^8$ $K_2 = 4.2 \times 10^6$ $K = 2.1 \times 10^7$	$K = 5.9 \times 10^7$
AATTGCAGTC	$K = 7.2 \times 10^6$	$K = 1.5 \times 10^8$	—

a $K = (K_1 * K_2)^{0.5}$

Figure S2. SPR binding affinity: A) SPR sensorgrams for the interaction of selected compounds with A_5T_5 (upper panel) and AATTGCAGTC (lower panel) DNA sequences. B) Comparison of the SPR binding affinity for A_5T_5 and AATTGCAGTC DNA sequences with different compounds using steady-state analysis.

A)



B)

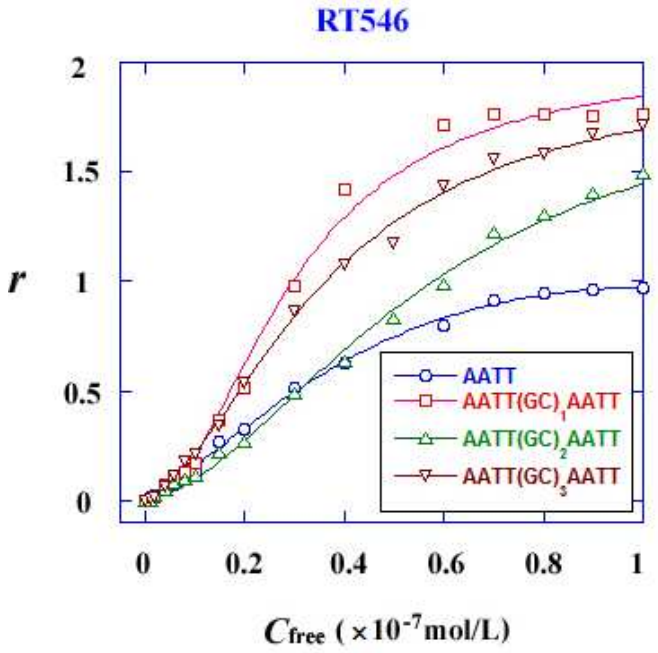


Figure S3. The effect of GC length change on RT546 binding: A) SPR sensorgrams for the interaction of RT546 with AATT (0 GC), AATTGCAATT (1 GC), AATTGCGCAATT (2 GC) and AATTGCGCGCAATT (3 GC). B) Comparison of the SPR binding affinity for RT546 binding with the two sites AATT DNA sequences with different GC length (0 to 3 GC). The lines are the best fit values to a single site or two sites interaction models.

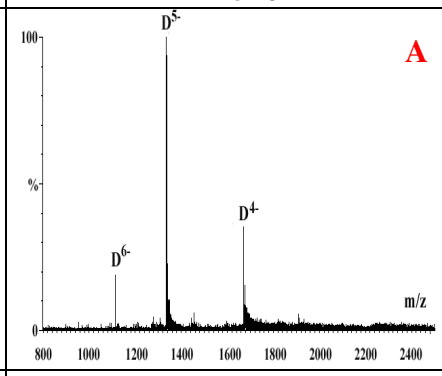
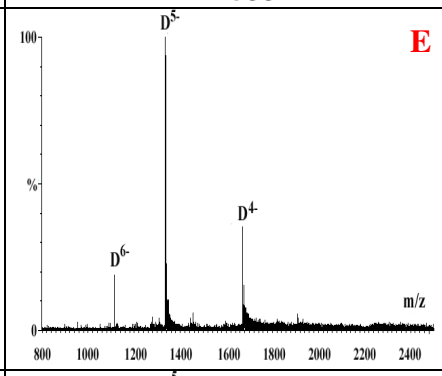
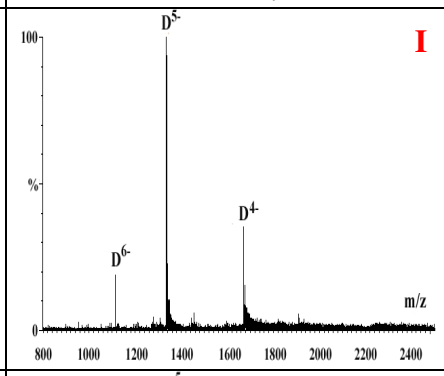
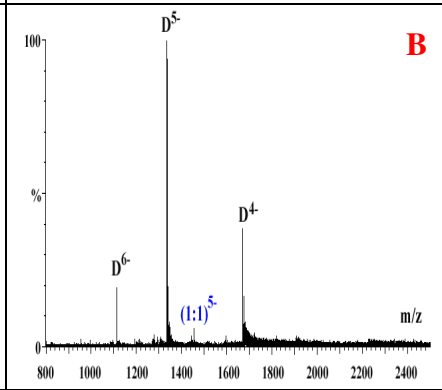
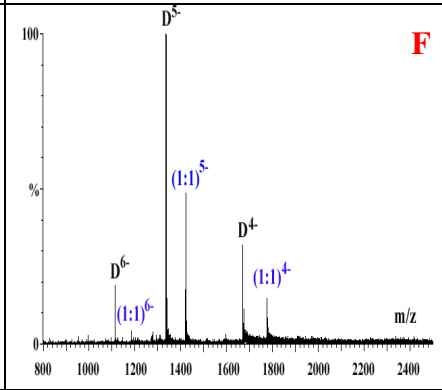
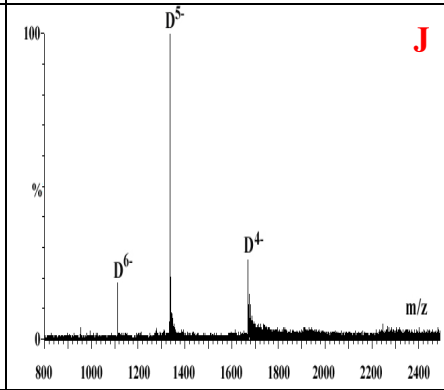
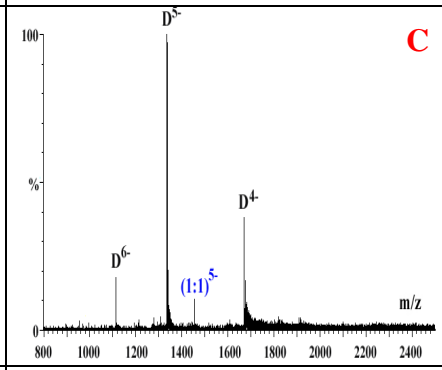
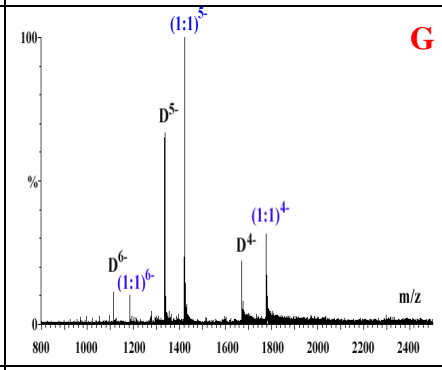
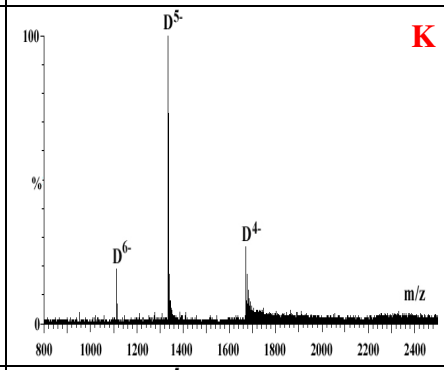
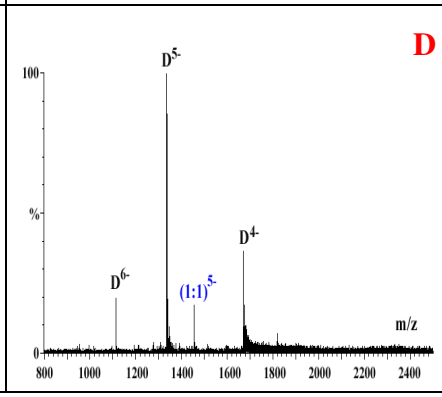
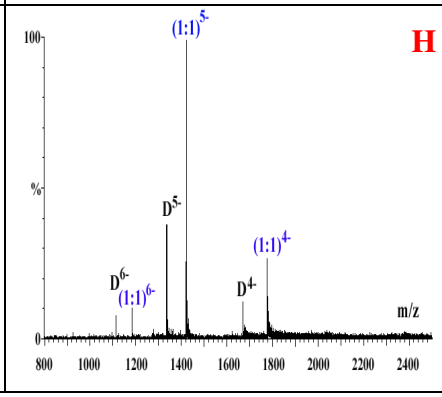
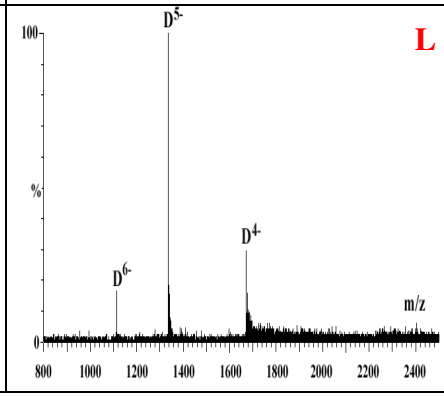
Ratio (Comp. / DNA)	AATT (D)		
	RT546	RT533	DB2119
0:1			
1:1			
2:1			
3:1			

Figure S4.A. ESI-MS spectra of AATT (5 µM) with the three linked compounds RT546, RT533 and DB2119 at different compound to DNA ratios (0:1, 1:1, 2:1 and 3:1).

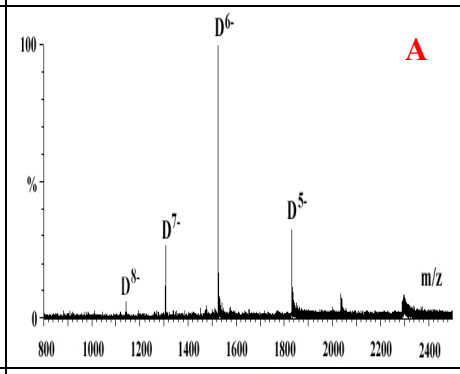
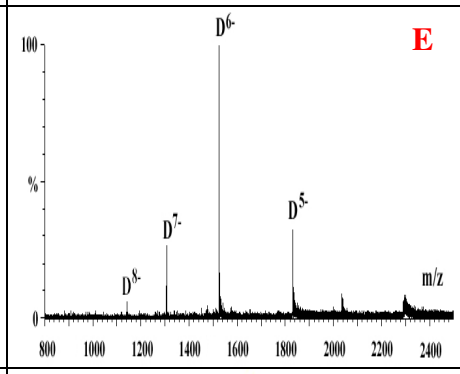
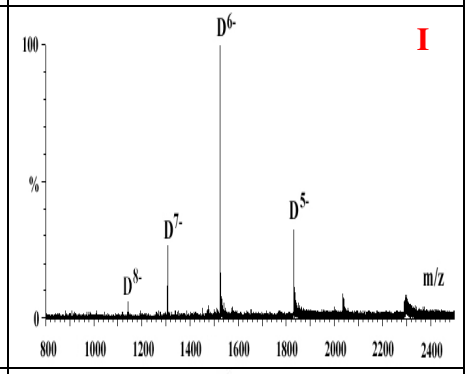
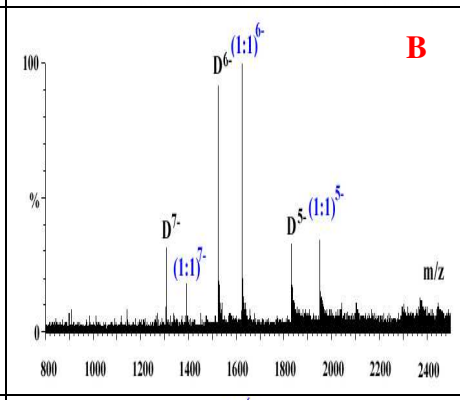
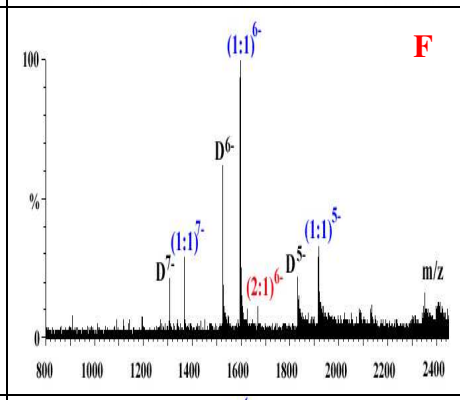
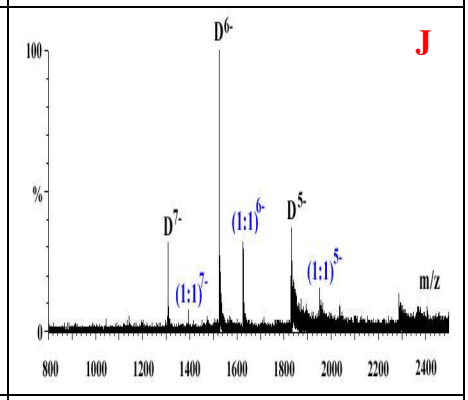
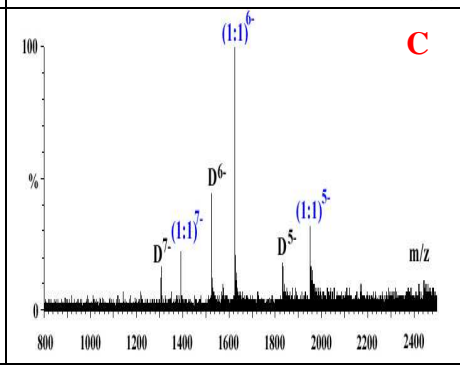
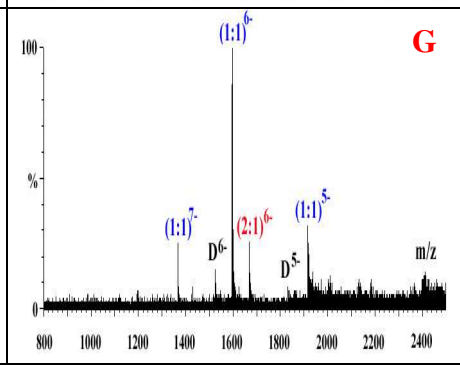
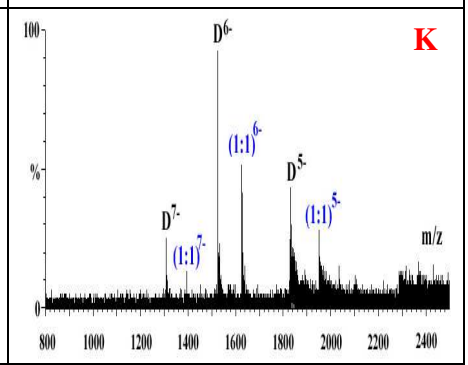
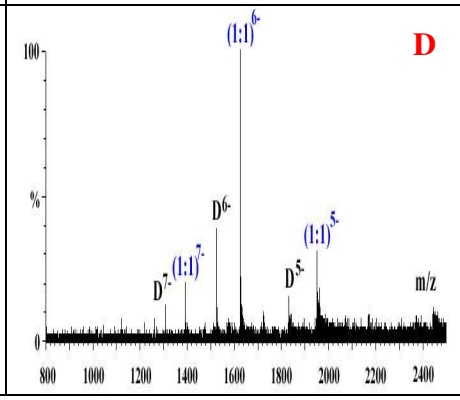
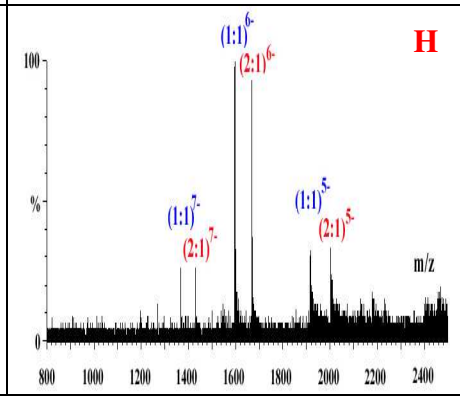
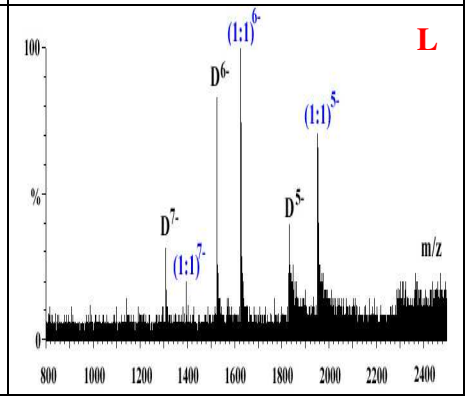
Ratio (Comp. / DNA)	AATTGAATT (D)		
	RT546	RT533	DB2119
0:1			
1:1			
2:1			
3:1			

Figure S4.B. ESI-MS spectra of AATTGAATT (5 µM) with the three linked compounds RT546, RT533 and DB2119 at different compound to DNA ratios (0:1, 1:1, 2:1 and 3:1).

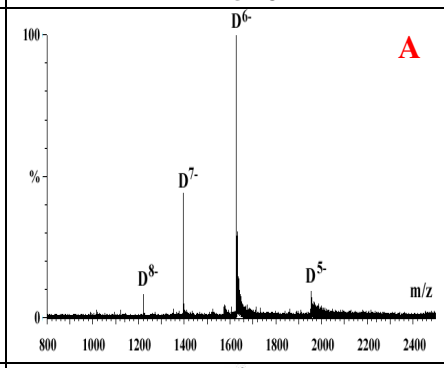
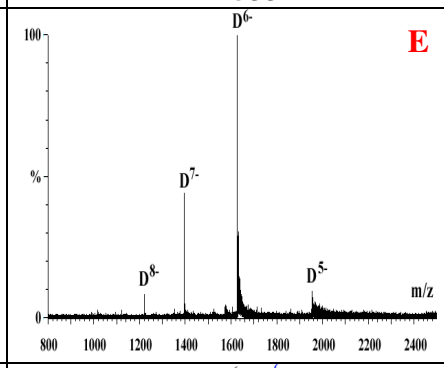
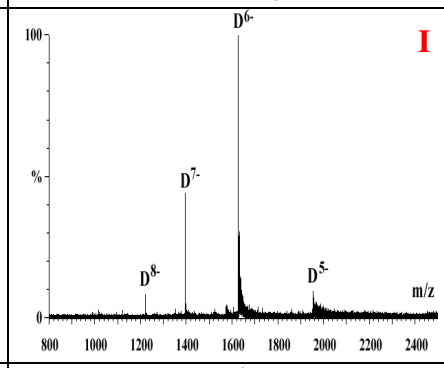
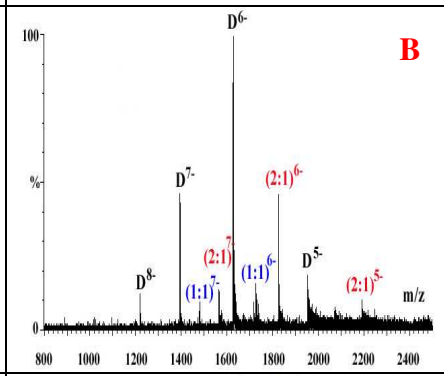
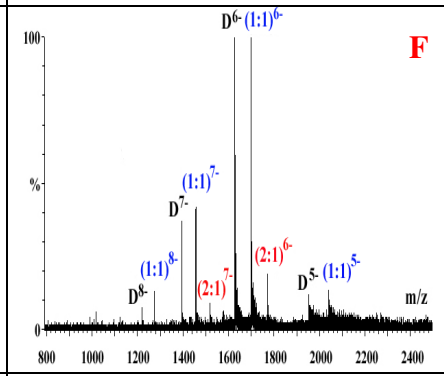
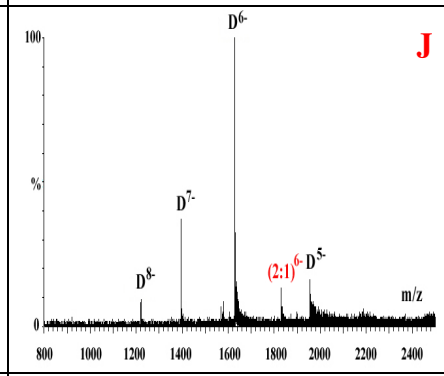
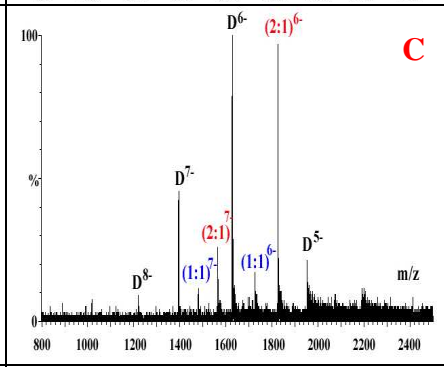
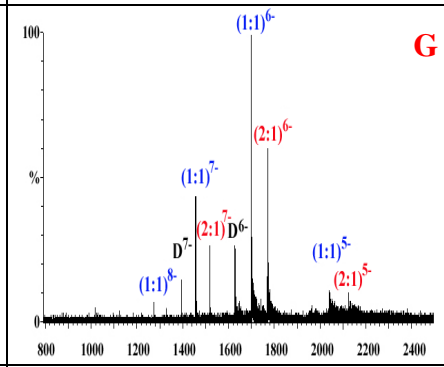
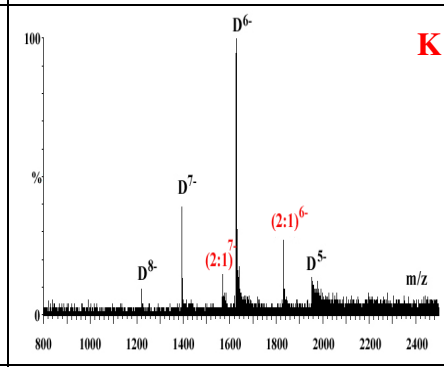
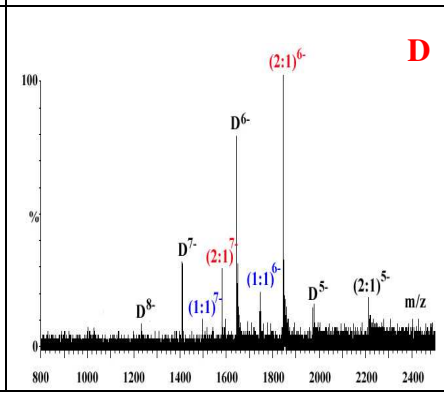
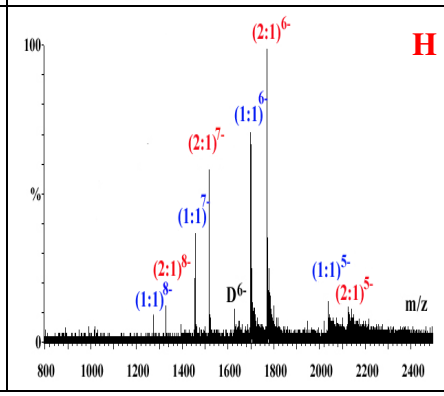
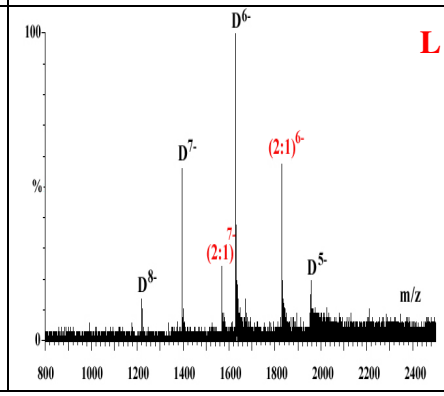
Ratio (Comp. / DNA)	AATTGCAATT (D)		
	RT546	RT533	DB2119
0:1			
1:1			
2:1			
3:1			

Figure S4C ESI-MS spectra of AATTGCAATT (5 μM) with the three linked compounds RT546, RT533 and DB2119 at different compound to DNA ratios (0:1, 1:1, 2:1 and 3:1).

A)

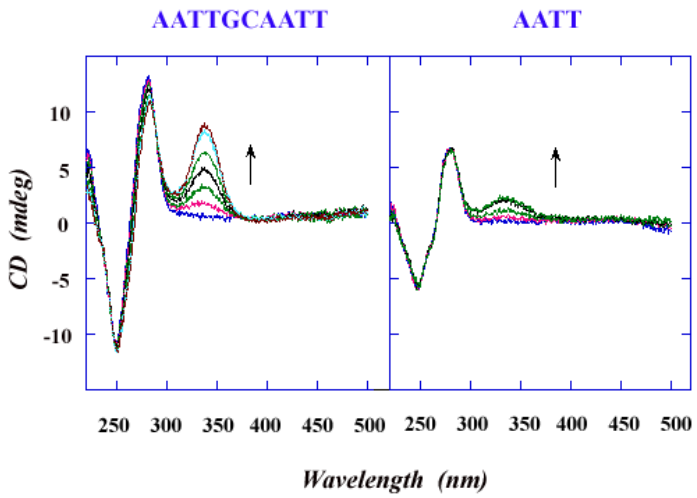


Figure S5A. Comparison of CD spectra of AATTGCAATT and AATT sequences in complex with RT546 at various mixing ratios. The experiments were conducted in CCL10 buffer at 25 °C with 10 μM DNA hairpin duplex. The ratios of compound to DNA are 0, 0.25, 0.5, 1, 1.5, 2.0 and 2.5 (from bottom to top) for AATTGCAATT and 0, 0.25, 0.5, 0.75,1 (from bottom to top) for AATT sequence.

B)

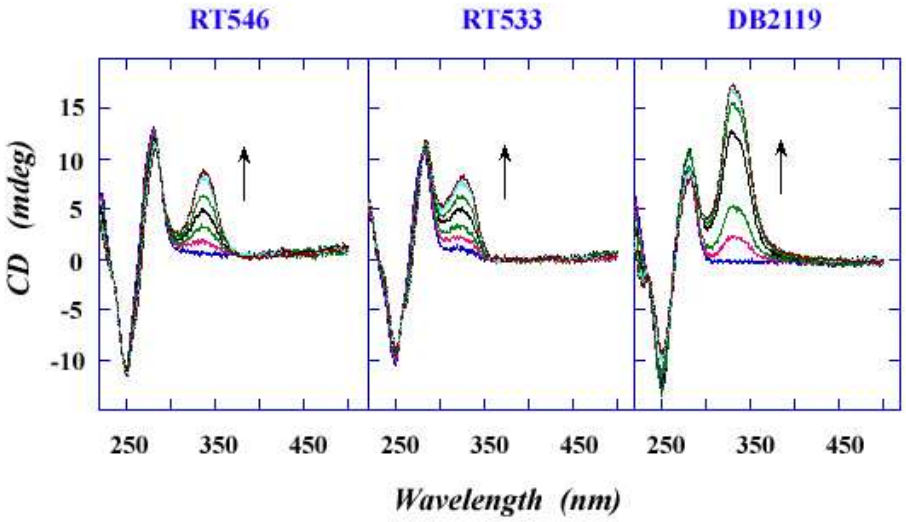


Figure S5B. Comparison of CD spectra of AATTGCAATT DNA sequence in complex with selected compounds (RT546, RT533 and DB2119) at various mixing ratios. The experiments were conducted in CCL10 buffer at 25 °C with 10 μM DNA hairpin duplex. The ratios of compound to DNA are 0, 0.25, 0.5, 1, 1.5, 2.0 and 2.5 (from bottom to top).

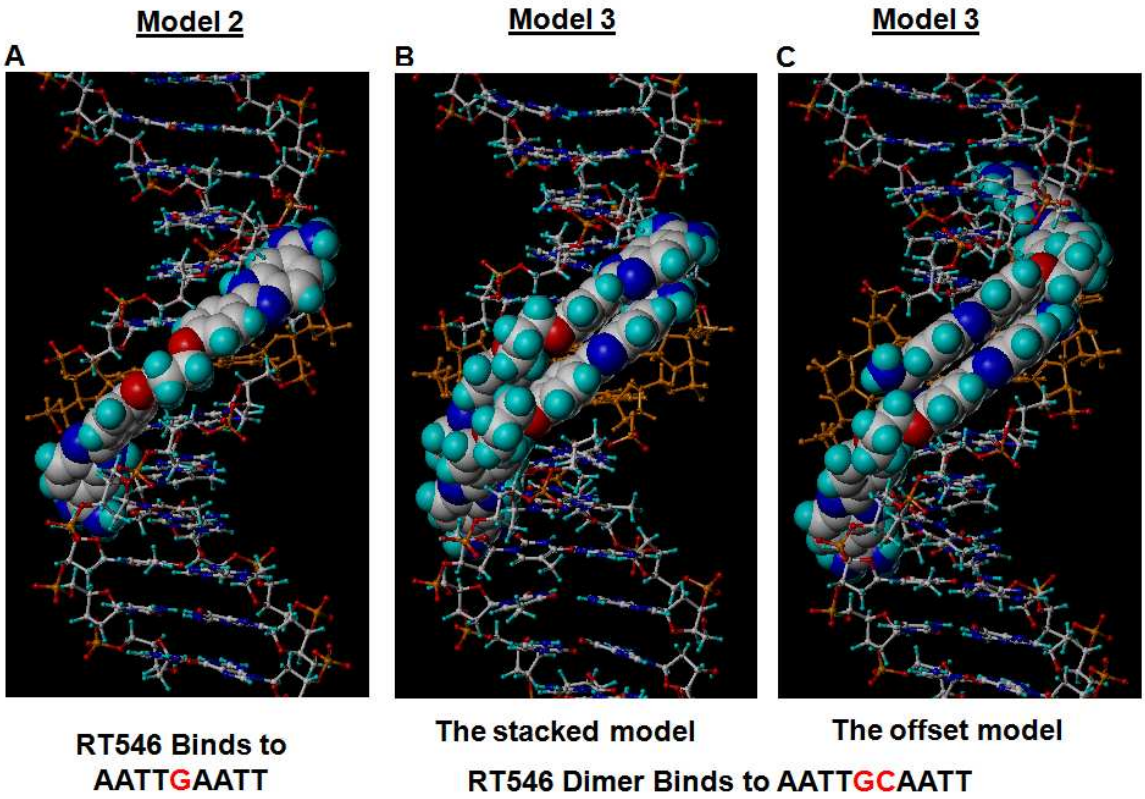


Figure. S6. Models for RT546 docked into the (A) AATTGAATT site, Model 2; and the AATTGCAATT site in (B) a highly overlapped and an offset complex (C), Model 3. The DNA sequences are displayed by ball and stick type in CPK colors and the GC base pairs are displayed in orange. The ligand is shown as spacefill type in CPK colors.

Table S1. Summary of SPR binding affinity and cooperativity for the interaction of DNA sequences with compounds that have different GC linker lengths.

Compound	AATT		AATTGAATT		AATTGCAATT	
	K^a (mol^{-1})	CF[b]	K^a (mol^{-1})	CF[b]	K^a (mol^{-1})	CF[b]
RT546 **[-O-$(CH_2)_3$-O-]**	$K = 6.8 \times 10^6$	—	$K = 5.1 \times 10^7$	—	$K_1 = 1.1 \times 10^6$ $K_2 = 9.4 \times 10^8$ $K = 3.2 \times 10^7$	3.4×10^3
DB2115 **[-O-$(CH_2)_4$-O-]**	—	—	$K = 9.5 \times 10^7$	—	$K_1 = 3.5 \times 10^6$ $K_2 = 1.9 \times 10^8$ $K = 2.6 \times 10^7$	2.2×10^3
DB2114 **[-O-$(CH_2)_5$-O-]**	—	—	$K = 3.5 \times 10^7$	—	$K_1 = 4.2 \times 10^6$ $K_2 = 2.5 \times 10^7$ $K = 1.0 \times 10^7$	23.8

a $K = (K_1 * K_2)^{0.5}$

b **CF** (Cooperativity Factor) = $(K_2 / K_1) \times 4$

Reference:

1. Borges, M.N., Messeder, J.C. & Figueroa-Villar, J.D. *European Journal of Medicinal Chemistry* **39***,* 925–929 (2004).
2. Wright, M.E. & Mullick, S. *Macromolecules* **25,** 6045-6049 (1992).
3. Rajakumar, P. & Murali, V. *Tetrahedron* **60,** 2351–2360 (2004).
4. Wojdacz, T.K. & Dobrovic, A. DNA Methylation: Methods and Protocols 2nd Edition, page 229, (Humana Press, 2009).