Single Site Discrimination of Cytosine, 5-Methylcytosine and 5-Hydroxymethylcytosine Nucleobases in DNA using Anthracene-Tagged Fluorescent Probes

Jean-Louis H. A. Duprey, Gemma A. Bullen, Zheng-yun Zhao, Dario M. Bassani, Anna F. A. Peacock, John Wilkie, and James H. R. Tucker

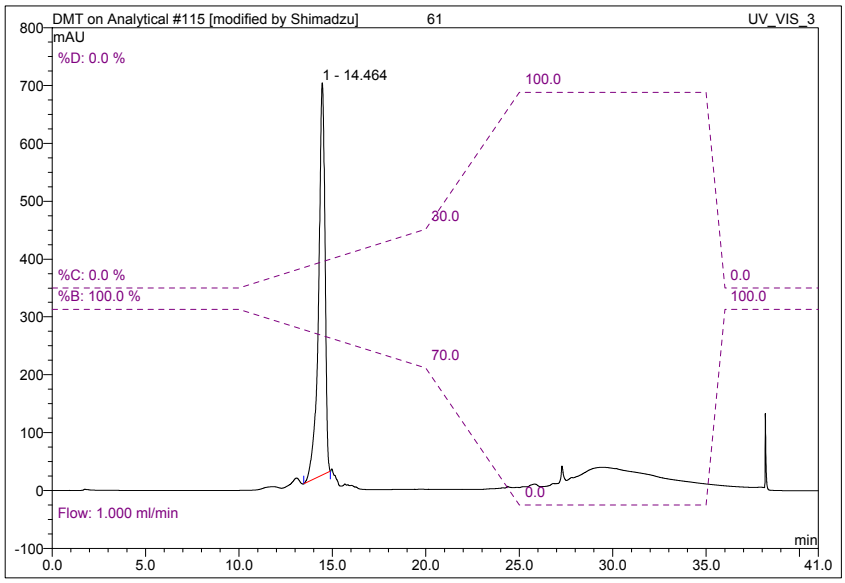
Index:

S1. HPLC of modified strands

Buffer B - 0.1M TEAA (pH 7.0), 15% CH_3CN
Buffer C - CH_3CN

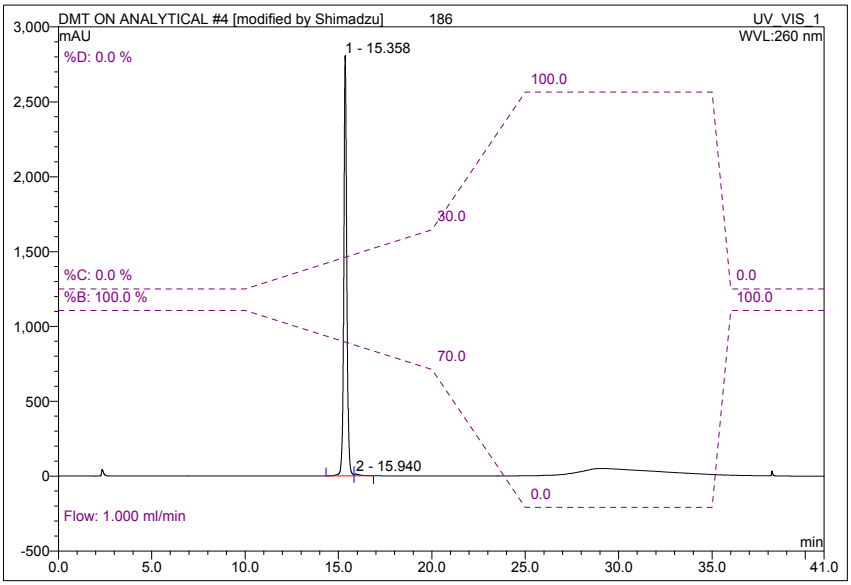
Gradient 0-10 mins, 100% B; 10-20 mins, 0% C – 30% C (remainder B); 20-25 mins, 30% C – 100% C; 25-35 mins, 100% C; 35-36 mins, 100% C – 0% C; 36-41 mins, 100% B

P-3 5’ TGGACTC**3**CTCAATG 3’



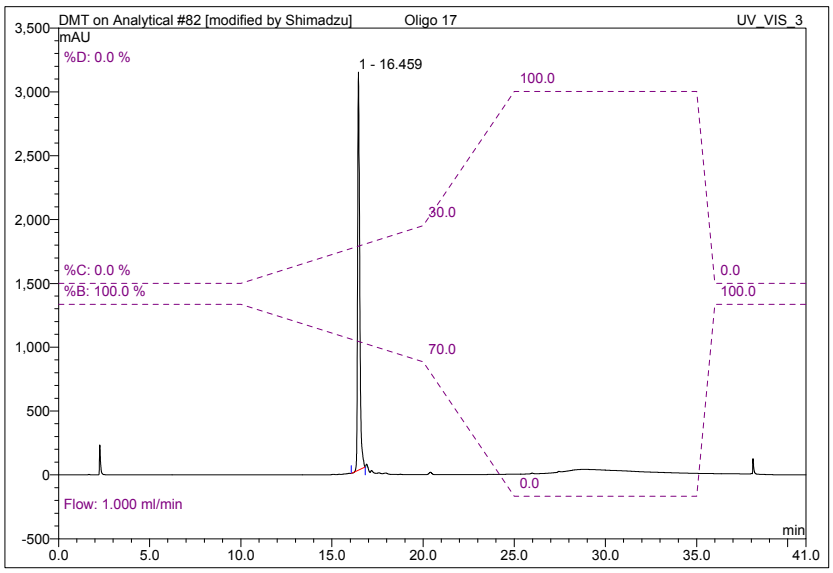
No.	Ret.Time min	Peak Name	Height mAU	Area mAU*min	Rel.Area %	Amount	Type
1	14.46	n.a.	677.849	271.246	100.00	n.a.	BMB
Total:			677.849	271.246	100.00	0.000	

P-4 5’ TGGACTC**4**CTCAATG 3’



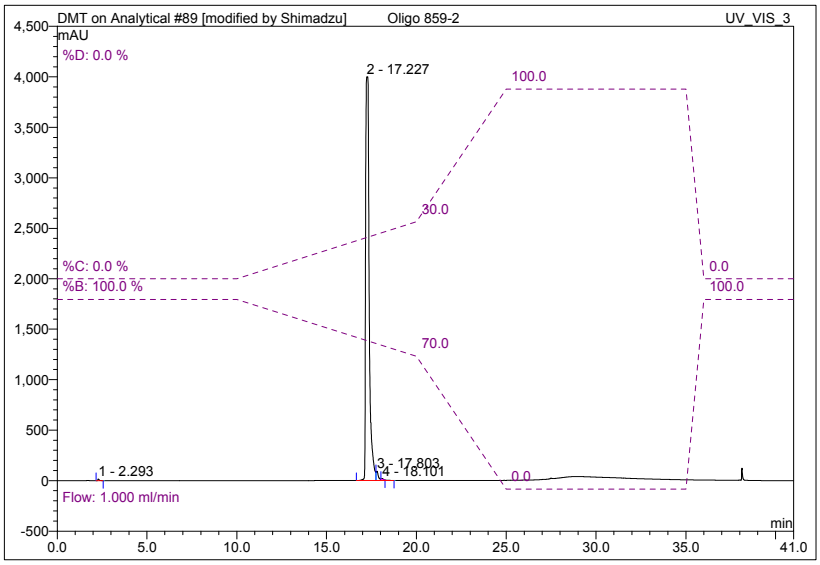
No.	Ret.Time min	Peak Name	Height mAU	Area mAU*min	Rel.Area %	Amount	Type
1	15.36	n.a.	2806.039	532.031	99.24	n.a.	BM
2	15.94	n.a.	12.168	4.096	0.76	n.a.	MB
Total:			2818.207	536.127	100.00	0.000	

P-5 5’ TGGACTC**5**CTCAATG 3’



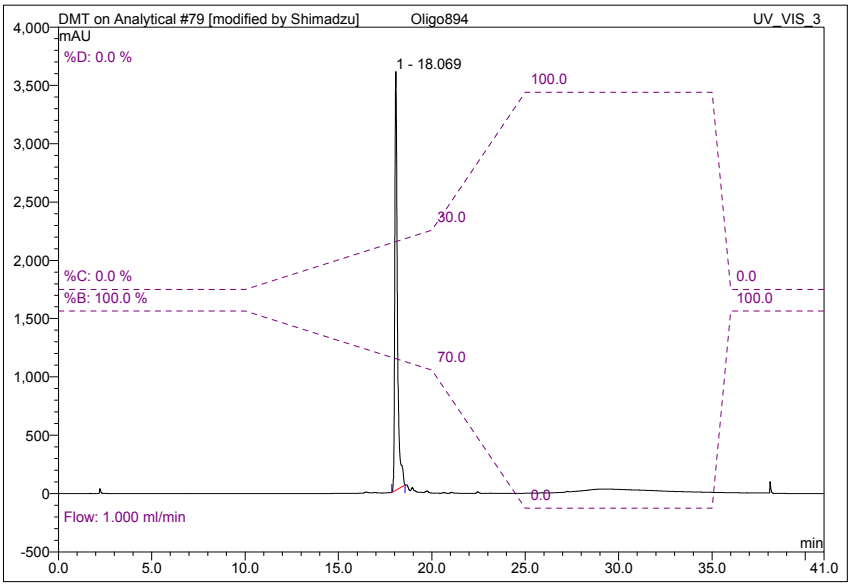
No.	Ret.Time min	Peak Name	Height mAU	Area mAU*min	Rel.Area %	Amount	Type
1	16.46	n.a.	3116.743	408.915	100.00	n.a.	BMB*
Total:			3116.743	408.915	100.00	0.000	

P-6 5’ TGGACTC**6**CTCAATG 3’



No.	Ret.Time min	Peak Name	Height mAU	Area mAU*min	Rel.Area %	Amount	Type
1	2.29	n.a.	16.477	1.193	0.13	n.a.	BMB
2	17.23	n.a.	3998.416	913.838	98.09	n.a.	BM
3	17.80	n.a.	90.749	15.510	1.66	n.a.	MB
4	18.10	n.a.	10.340	1.046	0.11	n.a.	Rd
Total:			4115.982	931.587	100.00	0.000	

P-7 5’ TGGACTC**7**CTCAATG 3’



No.	Ret.Time min	Peak Name	Height mAU	Area mAU*min	Rel.Area %	Amount	Type
1	18.07	n.a.	3589.356	605.355	100.00	n.a.	BMB*
Total:			3589.356	605.355	100.00	0.000	

S2. Mass Spectrometry

Probe **P-n** 5'-TGGACTC**n**CTCAATG -3'
P-3 *m/z* Calculated for $C_{158}H_{197}N_{51}O_{89}P_{14}$ 4668 (M), (ES-) Found 4668
P-4 *m/z* Calculated for $C_{159}H_{199}N_{51}O_{89}P_{14}$ 4682 (M), (ES-) Found 4682
P-5 *m/z* Calculated for $C_{160}H_{201}N_{51}O_{89}P_{14}$ 4696 (M), (ES-) Found 4696
P-6 *m/z* Calculated for $C_{161}H_{203}N_{51}O_{89}P_{14}$ 4710 (M), (ES-) Found 4710
P-7 *m/z* Calculated for $C_{162}H_{205}N_{51}O_{89}P_{14}$ 4724 (M), (ES-) Found 4724

Target **S-B** 5'-CATTGAG**B**GAGTCCA-3'
S-C *m/z* Calculated for $C_{146}H_{184}N_{58}O_{87}P_{14}$ (M) 4577 Found (ES-) 4577
S-mC *m/z* Calculated for $C_{147}H_{186}N_{58}O_{87}P_{14}$ (M) 4591 Found (ES-) 4591
S-hmC *m/z* Calculated for $C_{147}H_{186}N_{58}O_{88}P_{14}$ (M) 4607 Found (ES-) 4607

S3. CD Spectroscopy

CD spectra were acquired on a Jasco J810 spectropolarimeter at 21 °C with a cuvette path length of 1 cm. A total of 10 scans were measured between 200 and 450 nm at a scan rate of 200 nm/min, data pitch 0.5 nm and normal sensitivity.

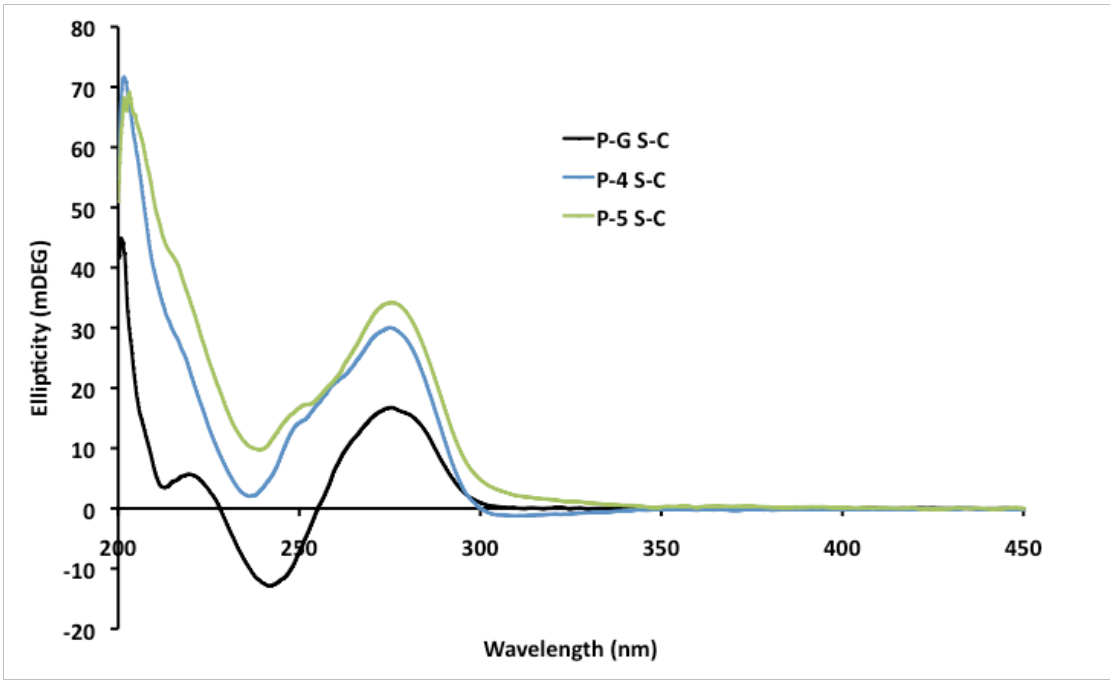


Figure S1. CD spectra of **P-4** (blue) and **P-5** (green) duplexed with **S-C**. Unmodified control (**P-G•S-C)** is also shown (black). [DNA] = 5 μM in 10 mM sodium phosphate buffer pH 7.0, 100 mM NaCl.

S4. Fluorescence Spectra and Quantum Yields

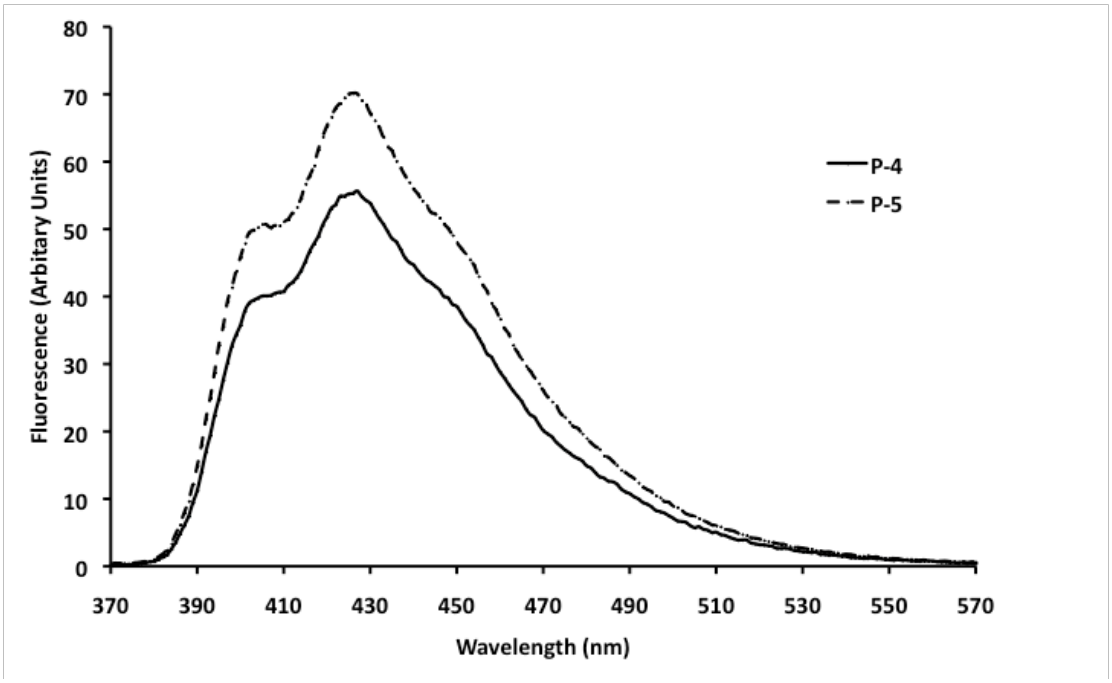


Figure S2. Fluorescence spectra of **P-4** (solid) and **P-5** (dashed) λ_{ex} = 350 nm. [DNA] = 1 μM in 10 mM sodium phosphate buffer pH 7.0, 100 mM NaCl.

Quantum yields (QY) of the probes and duplexes were determined using quinine sulfate in 1 N sulfuric acid as a standard for fluorescence emission intensity upon excitation at 350 nm (Φ_f = 0.546).[1]

Table S1. Comparison of the quantum yields of **P-n** single strands and duplexes with target strands **S-C**, **S-mC** and **S-hmC**. λ_{ex} = 350 nm. [DNA] = 1 μM in 10 mM sodium phosphate buffer pH 7.0, 100 mM NaCl.

	Number of carbons				
	n = 3	n = 4	n = 5	n = 6	n = 7
P-n	0.040	0.029	0.039	0.041	0.052
S-C	0.031	0.035	0.047	0.054	0.049
S-mC	0.028	0.024	0.042	0.042	0.045
S-hmC	0.026	0.020	0.031	0.036	0.039

S5. Fluorescence Lifetimes

Table S2. Fluorescence lifetimes (ns) of single- and double-stranded systems involving P-n, S-C, S-mC and S-hmC. Recorded at room temperature (293 K), λ_{ex} 371 nm, λ_{em} 426 nm, [DNA] = 5 μM, 10 mM pH 7 phosphate buffer, 100 mM NaCl.

Strand[a]	τ1 Wt (ns) (%)	τ2 Wt (ns) (%)	τ3 Wt (ns) (%)	χ2
P3	0.8 (34)	3.0 (40)	9.3 (26)	1.02
P3:SC	0.8 (47)	2.6 (40)	7.7 (13)	0.93
P3:SmC	0.8 (62)	2.5 (32)	9.7 (6)	1.02
P3:ShmC	0.8 (56)	3.3 (44)		1.18
P4	0.6 (34)	2.6 (42)	8.3 (24)	0.97
P4:SC	1.1 (50)	2.6 (50)		0.98
P4:SmC	0.8 (33)	2.6 (60)	19 (7)	1.02
P4:ShmC	0.9 (37)	2.2 (63)		0.91
P5	0.8 (42)	3.2 (37)	10.9 (21)	0.99
P5:SC	0.7 (25)	2.3 (55)	4.6 (20)	1.24
P5:SmC	1.0 (25)	2.5 (53)	5.2 (22)	1.06
P5:ShmC	0.9 (34)	2.0 (59)	5.8 (7)	1.19
P6	0.6 (28)	2.6 (41)	8.3 (31)	0.95
P6:SC	1.4 (29)	2.7 (64)	8.5 (7)	1.06
P6:SmC	0.9 (29)	3.0 (71)		1.01
P6:ShmC	0.9 (22)	2.7 (78)		0.95
P7	0.8 (26)	3.1 (43)	11.2 (31)	0.97
P7:SC	1.0 (30)	3.2 (50)	9.1 (20)	1.08
P7:SmC	0.8 (22)	2.7 (58)	7.3 (20)	1.04
P7:ShmC	1.3 (52)	3.4 (48)		1.03

S6. Molecular Modelling

Molecule design was carried out using the insightII modelling software whereby an unmodified duplex model of the appropriate sequence (**P-G•S-C**) was manually modified by replacing the G3 nucleotide unit of **P-G** with the anthracene monomer (n = 4 or 5). This monomer unit was parameterised using the antechamber module of AMBER to generate prep and frcmod files. For the cytosine variant duplexes (**P-n•S-mC/hmC**) prep and frcmod files for nucleosides mC[2] and hmC[3] were obtained from the literature. Water molecules were added around the DNA structure with a 15 Å cut-off using the TIP3PBOX[4] format and sodium ions were added to neutralise the backbone charges to create a neutral molecule. Inpcrd and prmtop files were generated using the LEaP module of AMBER. Energy minimization was performed on the structures followed by a 10 ns molecular dynamics simulation, with 500 ps equilibration, using the sander module of AMBER with the ff03[5] and GAFF[6] force fields. A 12 Å non-bonding cut-off was used and the simulation was performed at a constant temperature (300K) and volume. All calculations were carried out on the University of Birmingham Bluebear computer cluster. The output files were visualized using VMD.[7]

P-4 and **P-5** with **S-C**, **S-mC** and **S-hmC**

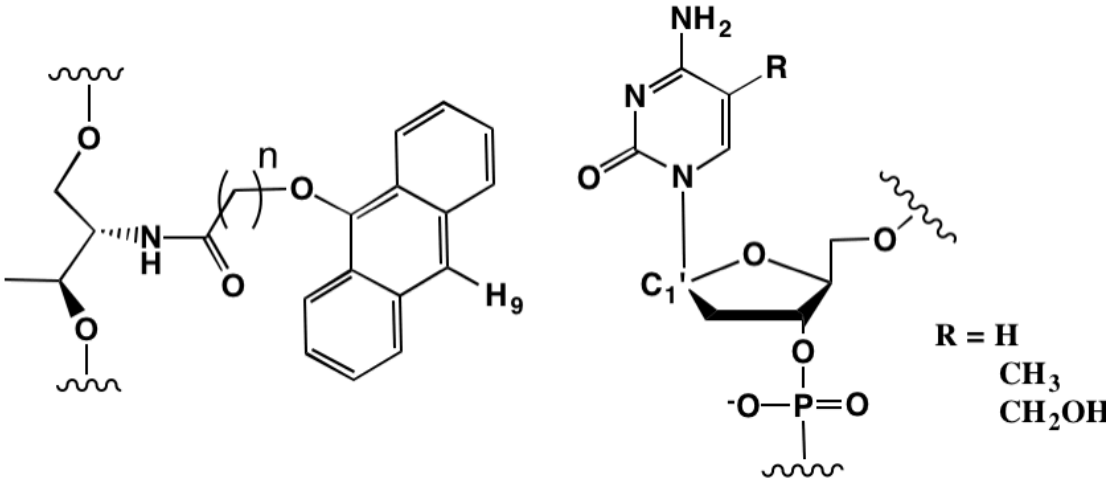


Figure S3. H9 of anthracene to 1'C of base opposite distance (distance X)

The distance X was chosen to assess whether the anthracene tag could be intercalated inside the duplex or located outside; a short distance is indicative of the former (less than 4 angstroms reflects the minimum possible van der Waals radius). The data for each probe (Figs S4 and S5 below) show no dependence on the methylation status of the cytosine opposite the tag, with intercalation possible in all cases, which is in general agreement with the melting temperature data. The models indicate that the base opposite is also highly dynamic, with each epigenetic variant able to "flip-out" of the duplex. However, the variation in distance for the different targets across various runs is greater for the **P-5** probe sequence than the **P-4**, particularly for **P-5•ShmC**. These results support the idea that a tag system with a finely tuneable size and flexibility is key to differentiating very closely matching epigenetic variations.

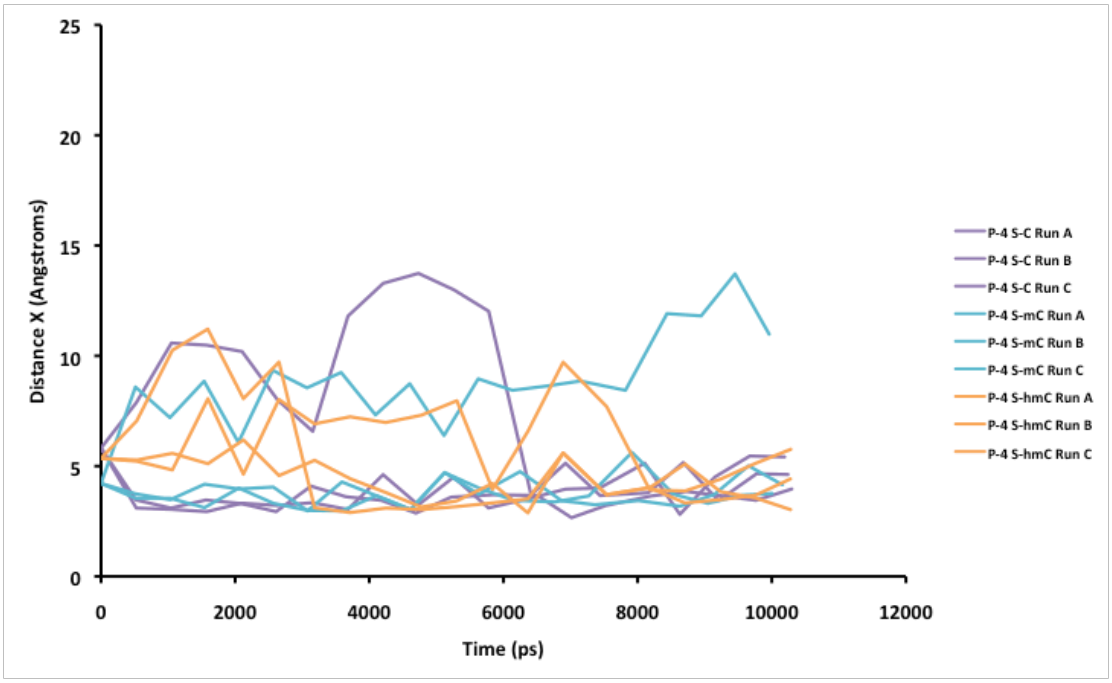


Figure S4. Graph of the variation in distance (X) over time for the **P-4** anthracene probe hybridized with different targets **S-C (**purple), **S-mC** (light blue) and **S-hmC** (orange) over three runs each.

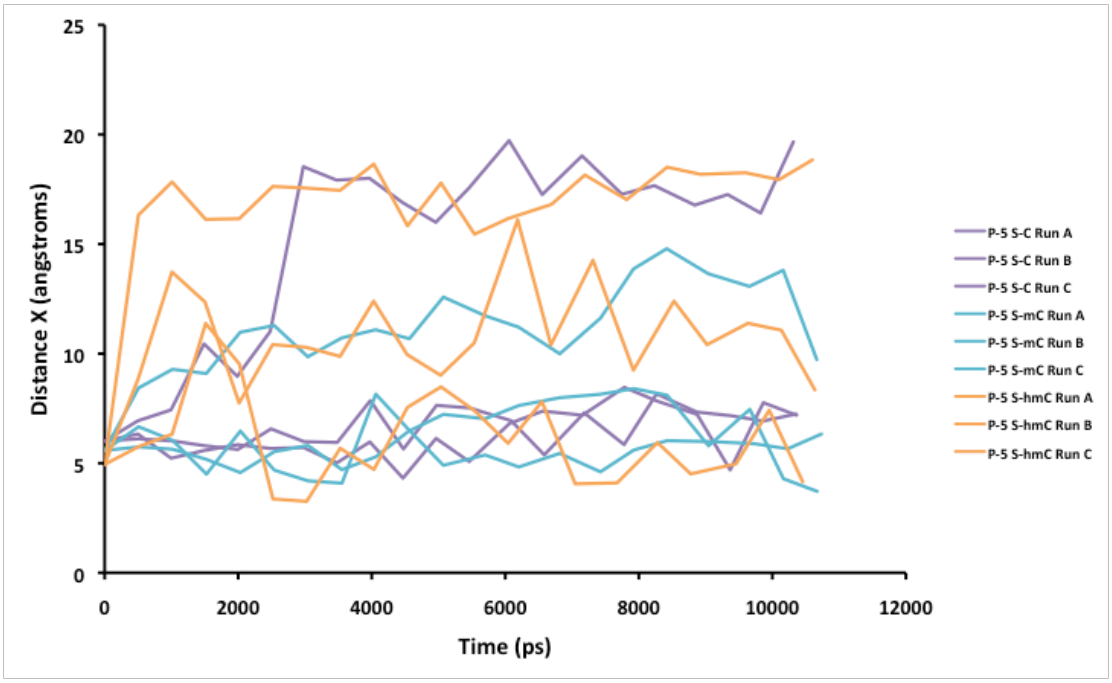


Figure S5. Graph of the variation in distance (X) over time for the **P-5** anthracene probe hybridised with different targets **S-C (**purple), **S-mC** (light blue) and **S-hmC** (orange) over three runs each.

References

[1] Brouwer, A. M. (2011) Standards for photoluminescence quantum yield measurements in solution. *Pure Appl. Chem.*, *83*, 2213–2228.

[2] Lankas, F., T., Cheatham III, T. E., Spackova, N., Hobza, P., Langowski, J. Sponer, J. (2002) Critical Effect of the N2 Amino Group on Structure, Dynamics, and Elasticity of DNA Polypurine Tracts. *Biophys. J. 82*, 2592-2609

[3] See Aduri, R., Psciuk, B. T., Saro, P., Taniga, H., Schlegel, H. B., SantaLucia Jr, J. (2007) AMBER Force Field Parameters for the Naturally Occurring Modified Nucleosides in RNA. *J. Chem. Theory Comput*, *3,* 1464–1465 and references therein.

[4] Jorgensen, W. L., Chandrasekhar J., Madura J., Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys. 79*, 926–935.

[5] Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T. (2003) A point- charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem. 24*, 1999–2012.

[6] Wang, J., Wolf, R., Caldwell, J., Kollman, P., Case, D. (2004) Development and testing of a general amber force field. *J. Comput. Chem. 25*, 1157–1174.

[7] Humphrey, W., Dalke, A., Schulten, K. (1996) *VMD: Visual molecular dynamics*. *J. Mol. Graphics*, *14*, 33-38.