Hydroxylation of Methylated CpG Dinucleotides Reverses Stabilisation of DNA Duplexes by Cytosine 5-Methylation

Armin Thalhammer[a], Anders S. Hansen[a‡], Afaf H. El-Sagheer[b,c], Tom Brown[b] and Christopher J. Schofield[a]*

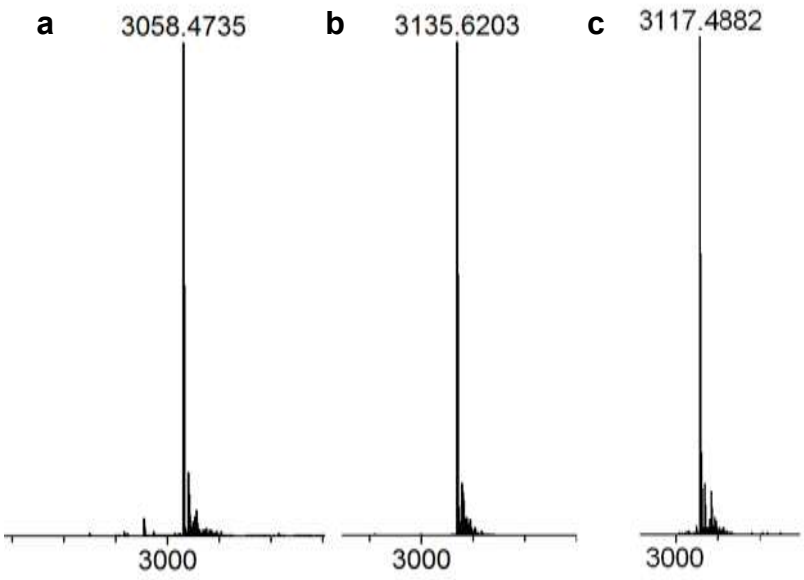
[a] *Department of Chemistry and the Oxford Centre for Integrative Systems Biology, Chemistry Research Laboratory, Oxford, U.K. Fax: +44 (0)1865 285002; Tel: +44 (0)1865 275625;*

[b] *School of Chemistry, University of Southampton, Southampton, UK.*

[c] *Chemistry Branch , Department of Science and Mathematics, Faculty of Petroleum and Mining Engineering, Suez Canal University, Suez, Egypt.*

**E-mail: Christopher.schofield@chem.ox.ac.uk*

[‡] *Current Address: Dept. of Chemistry and Chemical Biology, Harvard University, Cambridge, MA, USA.*



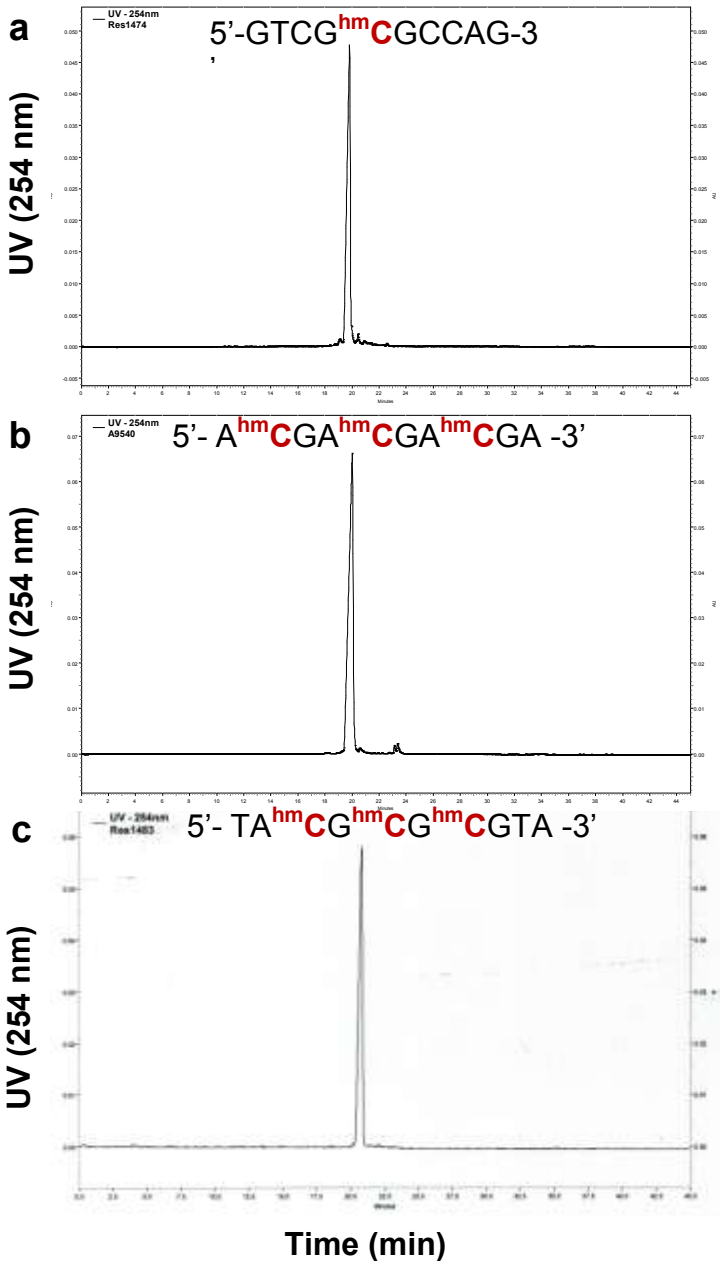
Supplementary Figure S1 Mass spectra of 5hmC-containing ODNs.

Spectra were recorded using a Bruker microTOF^{TM} II Focus ESI-TOF MS instrument in negative ion electrospray ionisation mode. Data were processed using the MaxEnt algorithm as implemented in the MassLynx software suite.

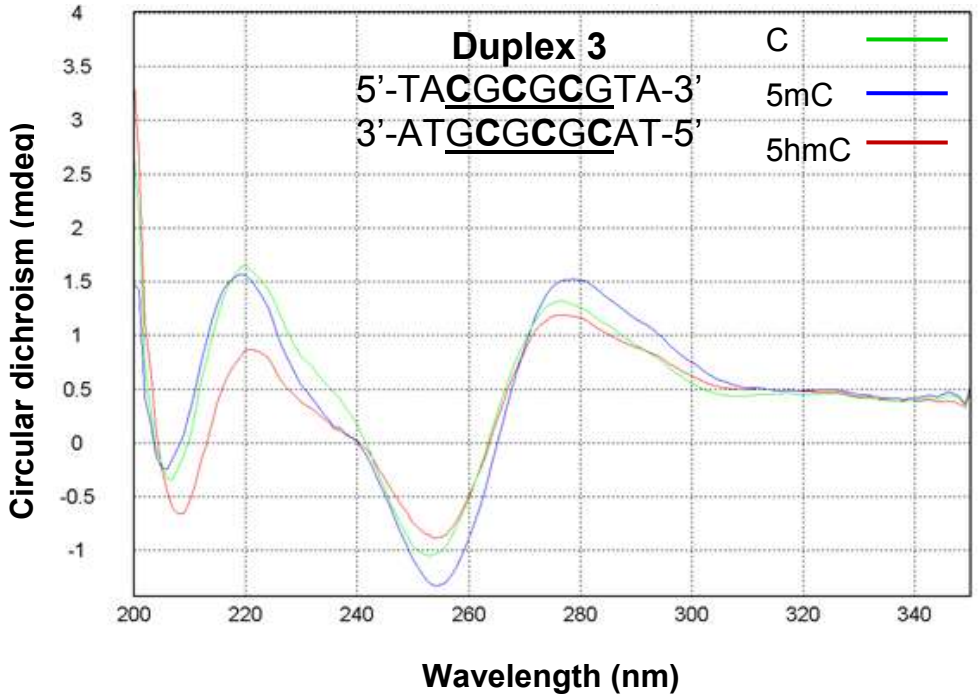
(**a**) Representative ODN containing a single 5hmC residue, 5'-GTCG^{hm}**C**GCCAG-3', calcd: 3059, found: 3058.

(**b**) ODN containing multiple 5hmC residues, 5'-A^{hm}**C**GA^{hm}**C**GA^{hm}**C**GA-3', calcd: 3136, found: 3135.

(**c**) ODN containing multiple 5hmC residues, 5'-TA^{hm}**C**G^{hm}**C**G^{hm}**C**GTA-3' calcd: 3118, found: 3117.



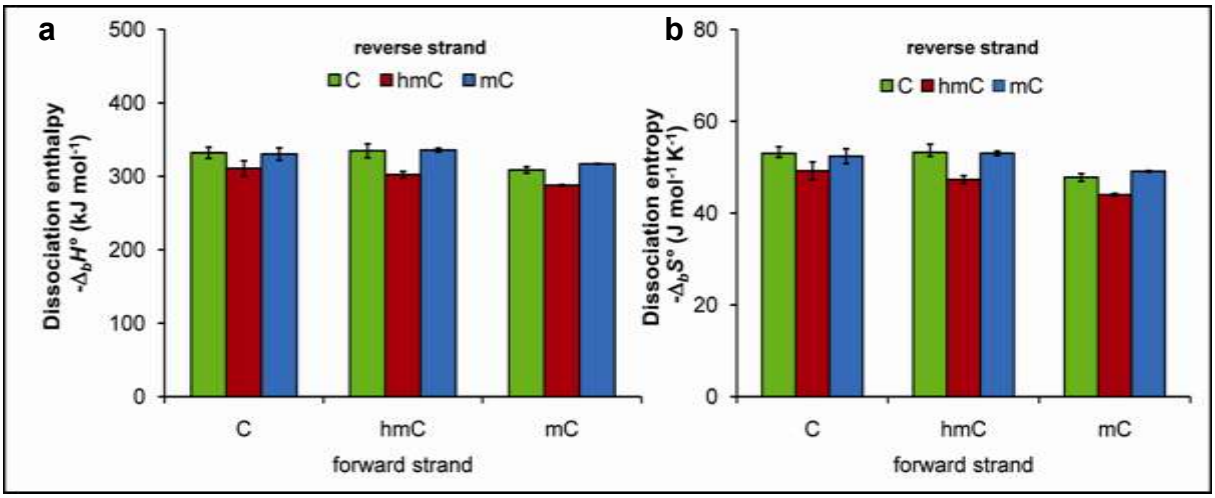
Supplementary Figure S2 Capillary electrophoresis (CE) analysis of 5hmC-containing ODNs.
(**a**) ODN containing a single 5hmC residue.
(**b**) and (**c**) ODNs containing three 5hmC residues.
The purity of 5hmC-containing ODNs was assessed by injection of 0.4 OD_{260} units of ODN in 100 µL water into a Beckman Coulter P/ACE™ MDQ capillary electrophoresis system, using an ssDNA 100-R Gel, Tris-Borate-7 M Urea (Kit No 477480). An injection voltage of 10 kV and a separation voltage of 9 kV were used (run duration: 45 min). Detection was performed by monitoring the UV absorbance at 254 nm. Data were analysed using 32 Karat software.



Supplementary Figure S3 CD spectrophotometry of C/hmC/mC-containing variants of duplex 3 (Supplementary Table S1).

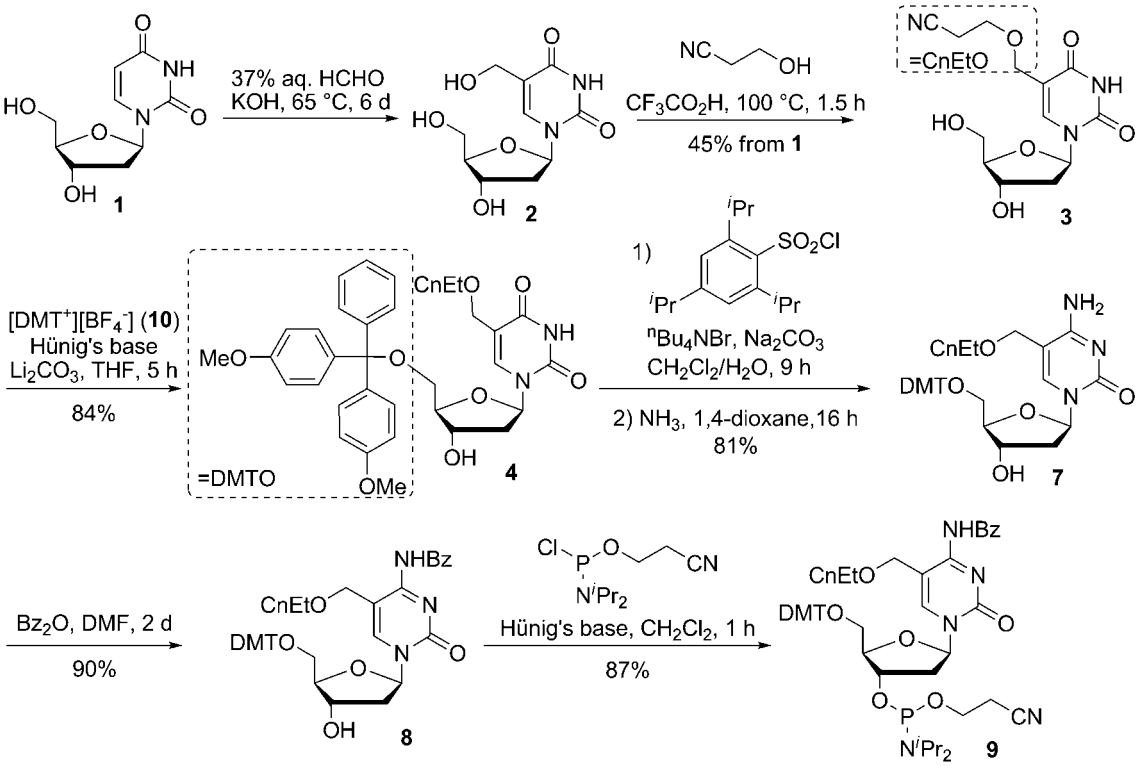
CD spectra were recorded in 10 mM sodium phosphate pH 7.0, 200 mM NaCl, 1 mM EDTA at 8 μM single-strand concentration, after annealing of self-complementary strands. The shape of the titration curves suggests that the tested duplexes exist predominantly in B-DNA form.[1]

Duplex 4
5'-A**C**GA**C**GA**C**GA-3'
3'-TCGT**C**GT**C**GT-5'



Supplementary Figure S4 ITC-derived enthalpic ($\Delta_bH°$) and entropic ($\Delta_bS°$) contributions to the binding free energy in C/hmC/mC-containing variants of duplex 4 bearing six modifications (Supplementary Table S1).

Raw data sets are given in Supplementary Table S3 and Supplementary Table S4, and data bars represent mean ± s.d. of at least two independent experiments. See main text (Fig. 4) for the corresponding binding constants (K_b), representative titration curves and discussion.



Supplementary Figure S5 Preparation of β-cyanoethyl-protected 5-hydroxymethyl-2'-deoxycytidine phosphoramidite.[2]

Supplementary Tables

Supplementary Table S1 Sequences of ODN duplexes investigated in this study.

(Bold: modification site; underlined: CpG dinucleotides)

Entry	Sequence	Modification pattern		CpG sites	Self-complementary ?	Method of Analysis
		Forward strand	Reverse strand			
1	5′-CTGA**C**GTGAC-3′ 3′-GACTG**C**ACTG-5′	C	C	2	No	UV melting
			hmC			
			mC			
		hmC	C			
			hmC			
			mC			
		mC	C			
			hmC			
			mC			
2	5′-CTGA**C**GAGAC-3′ 3′-GACTG**C**TCTG-5′	C	C	2	No	
			hmC			
			mC			
		hmC	C			
			hmC			
			mC			
		mC	C			
			hmC			
			mC			
3	5′-TA**C**G**C**G**C**GTA-3′ 3′-ATG**C**G**C**G**C**AT-5′	C		6	Yes	Concentration-dependent UV melting; Circular dichroism (CD)
		hmC				
		mC				
4	5′-A**C**GA**C**GA**C**GA-3′ 3′-TG**C**TG**C**TG**C**T-5′	C	C	6	No	Isothermal titration calorimetry (ITC)
			hmC			
			mC			
		hmC	C			
			hmC			
			mC			
		mC	C			
			hmC			
			mC			

Supplementary Table S2 Thermodynamic parameters for C/hmC/mC-containing variants of duplex 3 (Supplementary Table S1) determined by concentration-dependent UVM.

nucleotide	concentration range	number of concentrations tested	total number of data points	Best-fit values[1]				Thermodynamic parameters[1]			
				Slope	*Slope 95% CI*	*Y intercept*	*Y intercept 95% CI*	$-\Delta_b H°$	$-\Delta_b S°$	$-\Delta_b G°^{,calcd,2}$	$K_b^{calcd,2}$
	(μM)			(K)	(K)	(K^{-1})	(K^{-1})	(kJ mol^{-1})	(J mol^{-1} K^{-1})	(kJ mol^{-1})	(10^7 M^{-1})
C	1 - 80	7	57	$3.07 \cdot 10^{-5}$	$4.3 \cdot 10^{-7}$	$2.61 \cdot 10^{-3}$	$4.9 \cdot 10^{-6}$	271 ± 4	707 ± 10	52 ± 3	5
hmC	1 - 80	8	74	$3.01 \cdot 10^{-5}$	$5.4 \cdot 10^{-7}$	$2.62 \cdot 10^{-3}$	$6.2 \cdot 10^{-6}$	276 ± 5	723 ± 13	52 ± 4	6
mC	1 - 80	7	96	$3.22 \cdot 10^{-5}$	$7.0 \cdot 10^{-7}$	$2.54 \cdot 10^{-3}$	$8.1 \cdot 10^{-6}$	258 ± 6	658 ± 14	54 ± 4	14

[1]Data are given as mean ± 95% confidence interval, as calculated in GraphPad Prism v5.0.
[2]Calculated at 37 °C (310.15 K).

Supplementary Table S3 Individual ITC titration experiments and derived thermodynamic quantities for C/hmC/mC-containing variants of duplex 4 (Supplementary Table S1). Entries refer to independent biological repeats. Modification patterns refer to the type of modification present in the forward and complementary strands of the DNA duplex.

Entry	Modification pattern	*n*	K_b [$\cdot 10^6$ M^{-1}]	$-\Delta_b H°$ [kJ mol^{-1}]	$-\Delta_b S°$ [J mol^{-1} K^{-1}]
1	**C/C**	0.98	20 ± 2	336	54
2		0.96	29 ± 10	321	51
3		0.98	27 ± 4	335	54
4		0.99	20 ± 2	337	54
5	**C/hmC**	0.99	19 ± 2	307	49
6		0.97	19 ± 1	303	48
7		0.93	17 ± 2	323	51
8	**C/mC**	0.94	46 ± 8	321	51
9		0.91	44 ± 5	333	53
10		0.93	43 ± 3	337	54
11	**hmC/C**	1.05	61 ± 5	327	52
12		1.05	52 ± 4	333	53
13		0.91	51 ± 3	323	51
14		0.92	48 ± 7	342	55
15		0.99	31 ± 3	337	54
16		0.96	42 ± 5	348	56
17	**hmC/hmC**	1.01	40 ± 3	299	47
18		1.02	36 ± 2	305	48
19	**hmC/mC**	0.97	102 ± 9	333	53
20		0.96	99 ± 7	338	54
21		1.00	85 ± 10	336	53
22	**mC/C**	0.99	185 ± 54	310	48
23		0.97	162 ± 29	305	47
24		0.96	133 ± 27	315	49
25		1.06	181 ± 36	305	47
26		1.01	207 ± 68	309	48
27	**mC/hmC**	0.99	137 ± 31	289	44
28		1.00	174 ± 45	287	44
29	**mC/mC**	0.96	272 ± 62	317	44
30		0.97	274 ± 72	317	44

Supplementary Table S4 Averages of thermodynamic parameters calculated from independent ITC titrations (Supplementary Table S3) based on duplex 4 (Supplementary Table S1). Values are mean ± s.d. of independent experiments.

Entry	Modification pattern	Replicates	*n*	K_b [$\cdot 10^6$ M^{-1}]	$-\Delta_b H°$ [kJ mol^{-1}]	$-\Delta_b S°$ [J mol^{-1} K^{-1}]
1	**C/C**	4	0.98 ± 0.01	24 ± 5	332 ± 8	53.0 ± 1.5
2	**C/hmC**	3	0.96 ± 0.03	18 ± 1	311 ± 10	49.2 ± 1.9
3	**C/mC**	3	0.93 ± 0.01	44 ± 2	330 ± 8	52.4 ± 1.6
4	**hmC/C**	6	0.98 ± 0.06	48 ± 10	335 ± 9	53.2 ± 1.8
5	**hmC/hmC**	2	1.02 ± 0.01	38 ± 3	302 ± 4	47.3 ± 0.8
6	**hmC/mC**	3	0.98 ± 0.02	95 ± 9	336 ± 3	53.0 ± 0.5
7	**mC/C**	5	1.00 ± 0.04	174 ± 28	309 ± 4	47.8 ± 0.8
8	**mC/hmC**	2	0.99 ± 0.01	156 ± 26	288 ± 1	44.0 ± 0.3
9	**mC/mC**	2	0.96 ± 0.00	273 ± 1	317 ± 0	49.1 ± 0.1

Supplementary Methods

ODN synthesis, purification and analysis

Standard DNA phosphoramidites, solid supports and additional reagents were from Link Technologies and Applied Biosystems Ltd. ODNs were synthesised using an Applied Biosystems 394 automated DNA/RNA synthesiser using a standard 1.0 μmole phosphoramidite cycle of acid-catalysed detritylation, coupling, capping, and iodine oxidation. Stepwise coupling efficiencies and overall yields were determined by an automated trityl cation conductivity monitoring facility and in all cases were >98.0%. β-Cyanoethyl phosphoramidite monomers were dissolved in anhydrous acetonitrile to a concentration of 0.1 M immediately prior to use. The coupling time for normal A, G, C, and T monomers was 3 s, and the coupling time for the 5-hydroxmethyl-2'-deoxycytidine monomer (5hmC) was extended to 360 s. Cleavage of the ODNs from the solid support and deprotection was achieved by exposure to concentrated aqueous ammonia solution (60 min. room temp.) followed by heating in a sealed tube (60 h, 65 °C). Crude products were purified by HPLC using a C8 reversed-phase column (Brownlee Aquapore) with a gradient of 15% to 40% MeCN in 0.1 M ammonium acetate over 20 min. The purified ODNs were characterised by MS and capillary gel electrophoresis (Supplementary Figures S1 and S2).

Determination of UV melting temperatures

Data were collected on a Cary 4000 Scan UV-Visible Spectrophotometer (Varian) at an ODN concentration of 5.0 μM (unless stated otherwise) in PBS. Spectra were recorded at 260 nm. The samples were denatured by heating to 85 °C at 10 °C/min then cooled to 20 °C at 1 °C/min and heated to 85 °C at 1 °C/min. Up to six melting curves were measured and the average T_m values were calculated using Cary Win UV Thermal application Software.

Analysis of thermodynamic data based on UV melting experiments

Self-complementary ODNs with modified bases were dissolved at 1–80 μM final concentration in 10 mM phosphate buffer pH 7.0, 200 mM NaCl, 1 mM EDTA, and subjected to multiple melting-annealing cycles while monitoring UV absorbance. Melting transitions were assumed to proceed in a two-state manner, and to obey van't Hoff's equation,

$$\frac{1}{T_m} = \frac{\mathrm{R}}{\Delta_b H^\circ} \ln c_T + \frac{\Delta_b S^\circ}{\Delta_b H^\circ}$$

where $\Delta_b H°$ and $\Delta_b S°$ denote the enthalpy and entropy of binding, respectively.

Plotting $1/T_m$ *versus* ln c_T gives a straight line with slope $R/\Delta_b H°$ and y-axis intercept $\Delta_b S°/\Delta_b H°$. Data were fitted using linear least-squares minimisation as implemented in GraphPad Prism 5.0. The free energy of binding, $\Delta_b G°$, and the binding constant, K_b (Supplementary Table S2) were calculated at 37 °C using

$$\Delta_b G° = \Delta_b H° - T\Delta_b S°$$

and

$$K_b = e^{-\frac{\Delta_b G°}{RT}}$$

Values for derived thermodynamic quantities are stated as best fit value ± 95% confidence interval.

CD spectroscopy

For CD analysis, self-complementary ODNs were diluted to 8 μM in PBS buffer, heated to 80 °C and re-annealed by placement on ice to afford duplex DNA (4 μM). CD spectra were recorded in quartz cuvettes (path length 1 mm, vol. 400 μL) from 200 nm to 350 nm at 20 °C using an Applied Photophysics Chirascan CD spectrophotometer. Data were processed by baseline subtraction (PBS buffer) and smoothed using the Savitsky-Golay algorithm (polynomial order 10).

Determination and accurate matching of ODN concentrations for ITC experiments

Phosphate-buffered saline (PBS) was used throughout and contained 10 mM sodium phosphate, 200 mM NaCl and 1 mM EDTA sodium salt, pH 7.0 (all Sigma-Aldrich). Due to the sensitivity of the derived thermodynamics parameters of ODN binding to salt concentration and pH,[3] all experiments were conducted using the same batch of degassed, sterile-filtered buffer.

Lyophilised ODNs (c=100-500 μM) were reconstituted in PBS and concentrations were initially determined by UV absorption at 260 nm. However, ODN extinction coefficients are susceptible to bias from sequence context-specific base stacking interactions, which can result in deviations from the expected 1:1 binding stoichiometry in ITC experiments.[4] To minimise these effects, ODNs (10–50 μM) were hydrolysed[4] with phosphodiesterase I (1 U/mL, from *Crotalus adamanteus*) in 200 mM Tris-Cl, 3 mM $MgCl_2$, pH 9.3 for 30 min at room temperature prior to absorption measurement (A_{260}) using a NanoDrop ND-1000 UV-Vis spectrophotometer (NanoDrop products, Wilmington, USA). Extinction coefficients for the hydrolysed mixtures

were calculated by addition of the literature values[5] for nucleoside 5'-monophosphates (A: 15.4, C: 7.4, G, 11.5 and T, 8.7 × 10^3 M^{-1} cm^{-1}),

To further increase the accuracy of ITC-derived thermodynamic quantities, and to rule out interference from potentially inaccurate extinction coefficients for 5hmC, an initial test set of ITC titrations was performed (in duplicate, data not shown) for each of the nine variants of duplex 4, using concentrations derived from phosphodiesterase digests.[4] Then, the apparent concentrations of ODN solutions were calculated, based on the average binding stoichiometry obtained across all titrations that contained a particular ODN. Apparent concentrations were within 15% of the expected values.

For ITC titrations used to calculate thermodynamic parameters (Supplementary Table S3), ODN stock solutions were diluted according to apparent concentrations derived from the above test titrations, to minimise the possible impact of inaccurate ODN concentrations on the derived thermodynamic quantities.

Isothermal titration calorimetry (ITC)

The heat released upon duplex formation from single-stranded ODNs was measured by isothermal titration calorimetry (ITC) at 37 °C, using an iTC200 microcalorimeter (MicroCal, Northampton, MA, USA), according to established procedures.[6, 7]

ODN concentrations were chosen such that the shape of the resulting titration curve allows derivation of the thermodynamic parameters of interest, *i.e.* $\Delta_b H°$, $\Delta_b S°$ and K_b, with comparable accuracy, at 37° C. High ODN concentrations result in larger heat amounts released per injection and increase the accuracy of calculated $\Delta_b H°$ values, yet at the same time give rise to steep transitions at the equivalence point, thus decreasing the accuracy of the calculated values of K_b, $\Delta_b G°$ and, consequently, $\Delta_b S°$.

ITC was performed by titration of a CpG dinucleotide-containing triple-modified ODN (10 μM) with a solution (100 μM) of its reverse complement. A 40 μL syringe was used to inject the titrant in 1.3 μL aliquots, until a >2-fold molar excess of titrant was reached, with stirring at 1000 rpm. A smaller volume (0.6 μL) was chosen for the first injection to compensate for the backslash of the syringe injector mechanism upon initial use. The first data point was excluded from all analyses. Stock solutions of PBS buffer were equilibrated at 37 °C for at least 15 min,

degassed under stirring and used to dilute ODN stock solutions, immediately before the start of titrations.

Thermodynamic data analysis from ITC

Automatic baseline correction was applied to the raw power/time data as implemented in the ORIGIN v7.0 software provided with the microcalorimeter. To correct for heats of dilution of the tested ODNs, the constant part of the titration curve in the power/time diagram, after the equivalence point, was fitted with a straight line parallel to the y-axis and subtracted from the raw data.

These corrected power/time curves were integrated to yield relative heat values (kJ per mol of injectant), and plotted against the molar ratio of reverse and forward strand.[6, 7] Values for the binding stoichiometry (n), binding enthalpy ($\Delta_b H°$), entropy ($\Delta_b S°$) and the binding constant (K_b) were determined by non-linear least-squares fitting to a two-state single-site binding model, as implemented in ORIGIN v7.0.

Thermodynamic parameters derived from fitting of independent replicates of titration curves and are stated in Supplementary Table S3. Binding constants are given as (best-fit value ± standard error). K_b, $\Delta_b H°$ and $\Delta_b S°$ for combinations of C/hmC/mC-containing ODNs are reported as mean ± s.d from at least two independent replicates (Supplementary Table S4). The average binding stoichiometry obtained across all ODN combinations was n=0.98 ± 0.04 (mean ± s.d.).

Supplementary References

1. D. M. Gray, R. L. Ratliff and M. R. Vaughan, in *Methods in Enzymology*, eds. D. M. J. Lilley and J. E. Dahlberg, Academic Press, 1992, pp. 389-406.
2. A. S. Hansen, A. Thalhammer, A. H. El-Sagheer, T. Brown and C. J. Schofield, *Bioorg. Med. Chem. Lett.*, 2011, **21**, 1181-1184.
3. B. E. Lang and F. P. Schwarz, *Biophys. Chem.*, 2007, **131**, 96-104.
4. G. Kallansrud and B. Ward, *Anal. Biochem.*, 1996, **236**, 134-138.
5. P. N. Borer, in *Handbook of Biochemistry and Molecular Biology, Nucleic Acids*, ed. G. D. Fasman, CRC Press, Boca Raton, 1975, p. 589.
6. T. Wiseman, S. Williston, J. F. Brandts and L. N. Lin, *Anal. Biochem.*, 1989, **179**, 131-137.
7. F. P. Schwarz, K. D. Puri, R. G. Bhat and A. Surolia, *J. Biol. Chem.*, 1993, **268**, 7668-7677.