ChemComm

View Online
Dynamic Article Links

Cite this: *Chem. Commun.*, 2011, **47**, 5325–5327

www.rsc.org/chemcomm

COMMUNICATION

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,[bc] Tom Brown[b] and Christopher J. Schofield*[a]

Received 20th December 2010, Accepted 16th March 2011
DOI: 10.1039/c0cc05671e

Cytosine-5-methylation stabilises DNA duplexes and is associated with transcriptional repression; 5-methylcytosine undergoes hydroxylation to 5-hydroxymethylcytosine, a modification of unknown biological function. Spectroscopic and calorimetric analyses show that 5-hydroxymethylcytosine introduction reverses the stabilising effect of 5-methylcytosine, suggesting that in some contexts, 5-methylcytosine hydroxylation may, along with other factors, contribute to the alleviation of transcriptional repression.

Cytosine C-5 methylation of CpG dinucleotides (Fig. 1a) is a DNA modification associated with inhibition of transcription (Fig. 1b) and epigenetic gene silencing.[1] The presence of 5-methylcytosine (5mC) can directly inhibit binding of the transcriptional machinery (*e.g.* by blocking transcription factor binding) or recruit chromatin-remodelling complexes for histone-mediated gene silencing. Aberrant DNA methylation is associated with cancer.[2,3] Whereas direct enzymatic reversal is well established for some protein and DNA modifications including histone (de)acetylation and (de)methylation and DNA *N*-demethylation, there have been no reports of a single-step reversion of 5mC to cytosine.

Recently, it was reported that 5-hydroxymethylcytosine (5hmC, Fig. 1a) occurs in human Purkinje neuron DNA.[4] 5hmC comprises up to ~20% of total 5mC levels in some regions, and its formation is catalysed by TET1 in human cell lines.[5] TET1 and its homologues, TET2[6] (implicated in myeloid malignancies[7,8]) and TET3,[6] are 2-oxoglutarate (2OG) and Fe^{II}-dependent oxygenases, other examples of which have important roles in transcriptional regulation by transcription factor hydroxylation and histone demethylation.[9] 5hmC is a chemically stable DNA modification but its biological function is unknown.

We considered the possibility that 5mC-hydroxylation might alter the chemical stability of DNA. In addition to the stabilising thermodynamic effects of 5mC on DNA,[10,11] precedent for this possibility comes from observations on protein-hydroxylation, *i.e.* that the collagen triple helix[12] and ankyrin repeat domain folds[13] are stabilised by 2OG oxygenase-catalysed prolyl-4- and asparaginyl-3-hydroxylations, respectively.

To investigate the effect of 5hmC, we prepared β-cyanoethyl-protected 5hmC phosphoramidite using a procedure based on a conversion of 2′-deoxyuridine to 2′-deoxycytidine in the absence of temporary protecting groups (Fig. S5†).[14] Oligodeoxyribonucleotides (ODNs) were prepared by phosphoramidite synthesis, purified and analysed by LC-MS (Fig. S1, ESI†) and electrophoresis (Fig. S2†).

We initially performed UV melting (UVM) analyses on 2 DNA duplex forming sequences (1 and 2, Fig. 2) containing all combinations of C/5hmC/5mC on forward and complementary strands (Table S1†). Selection criteria included: (i) the presence of a CpG dinucleotide, (ii) an appropriate length both to form duplexes and investigate effects of modifications, (iii) an absence of predicted secondary structures (*e.g.* hairpin loops), and (iv) an even distribution of AT- and GC-rich regions to enable single-stage duplex melting transitions.

In accord with prior work,[10,11] our results (Fig. 2) show that duplexes with CpG dinucleotides methylated on both strands (mC/mC) have higher duplex melting temperatures (T_m) than

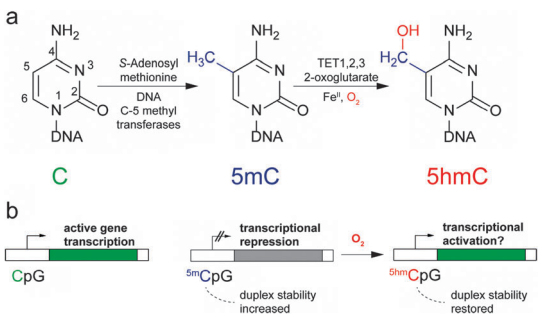

Fig. 1 (a) Post-replicatory cytosine C-5 modifications in human DNA. 5-Methylcytosine at CpG dinucleotides in promoter regions is associated with transcriptional silencing. 5mC and 5hmC formation are catalysed by DNA methyltransferases and 2OG oxygenases, respectively. (b) Potential roles of 5mC and 5hmC in transcriptional repression and activation.

[a] *Department of Chemistry and the Oxford Centre for Integrative Systems Biology, Chemistry Research Laboratory, Oxford, UK. E-mail: Christopher.schofield@chem.ox.ac.uk; 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*

† Electronic supplementary information (ESI) available: Sequences, synthesis, purification and analysis (LC-HRMS, CE, CD) of ODNs, biophysical data (UVM/ITC), and detailed procedures. See DOI: 10.1039/c0cc05671e

‡ Current Address: Dept. of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts, USA.

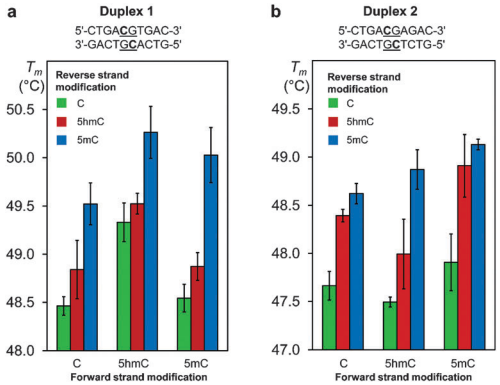

Fig. 2 Analyses of melting temperatures of duplexes containing single C, 5hmC or 5mC nucleotides (Table S1). *x*-axis, forward strand modification; colour: complementary strand modification.

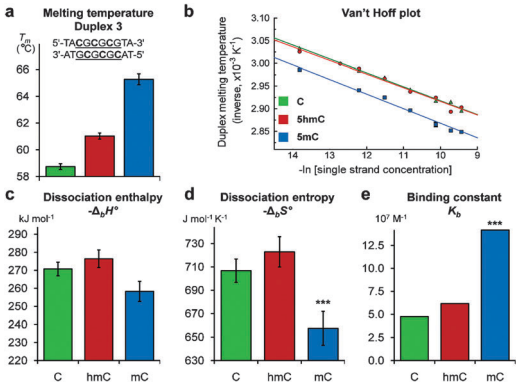

Fig. 3 Concentration-dependent UV-studies on 10mer duplexes with six modifications (bold). (a) Effect of C/5hmC/5mC at single concentrations. (b) Van't Hoff analysis. Thermodynamic parameters were derived using best-fit values of linear regression (mean ± 95% CI, ESI†), (c) Dissociation enthalpy ($-\Delta_bH^\circ$), (d) Dissociation entropy ($-\Delta_bS^\circ$). (e) Binding constants (calcd. from Δ_bG°, 37 °C). (*** denotes $P < 0.01$).

the cytosine equivalents (C/C). With the exception of duplexes based on sequence 1 with an hmC on the forward strand, these initial results suggested that, duplexes containing 5hmC have a T_m resembling that of unmodified DNA, with T_m increasing in the order C ≤ 5hmC < 5mC. They also suggest that, at least with singly modified duplexes, the relative degree of stabilisation depends on sequence context, and, when studying heterogeneously modified duplexes, on which strand the modification is placed. However, T_m differences for single modifications are small. Hence, we investigated the effect of increasing the number of 5mC and 5hmC modifications to reduce the significance of sequence context.

Our model sequence comprised a 10mer self-complementary CpG dinucleotide-based sequence with modified bases on both strands (Fig. 3). CD analyses for these duplexes were characteristic of B-DNA rather than the A- or Z-conformations (Fig. S3), indicating that 5hmC is a stable DNA component that does not induce overall major conformational changes (in short sequences). As with the results for singly modified duplexes, 5mC caused duplex stabilisation (ΔT_m = +1.1 °C per 5mC). In contrast to 5mC, 5hmC elicited a much smaller effect (ΔT_m = +0.4 °C per 5hmC) relative to unmodified C. These results suggest that the relatively large stabilising effect of cytosine methylation may, at least partly, be reversed by hydroxylation.

To investigate the basis of the effects of 5hmC, the enthalpic (Δ_bH°) and entropic (Δ_bS°) components of the free binding energy (Δ_bG°) were initially determined using concentration-dependent UV melting (UVM). Analysis of the concentration-dependent UVM data utilises the van't Hoff equation to derive thermodynamic parameters from the experimentally determined relationship between total ODN concentration and T_m (Fig. 3b). We used self-complementary 10mer ODNs with three modified nucleotides per strand to maximise potential T_m differences and ensure equal quantities of forward/complementary strands.

The derived $\Delta_bH^\circ/\Delta_bS^\circ/K_b$ values (Fig. 3c,d,e and Table S2†) imply that the stabilising effect of 5mC relative to 5hmC is primarily due to a less unfavourable change in binding entropy (Fig. 3d), which counteracts the less exothermic binding enthalpy (Fig. 3c); 5mC causes a significant reduction in Δ_bS° ($P < 10^{-4}$) relative to unmodified duplexes. These data concur with observations of a methylation-induced increase[10] in the T_m of the hairpin 5′-CGCGCGTTTCGCGCG-3′ without affecting Δ_bH°, comparisons of the stability of poly(G·C) to that of poly(G·5mC),[11] and work on dye-assisted melting-temperature-based mapping of CpG methylation.[15] Analyses to date, including our work, imply that duplex stabilisation by 5mC is predominantly due to entropic rather than enthalpic effects. The overall results for 5mC also demonstrate that data derived from model sequences, including those used in our work, may be of general significance. It is reasonable to propose that the same applies to 5hmC.

In contrast to 5mC-containing DNA, the enthalpic and entropic terms as well as the binding constants (K_b) for 5hmC-containing DNA were remarkably similar to those of unmodified DNA, as apparent from the van't Hoff plot (Fig. 3b), *i.e.* the stabilising effect of 5mC is substantially reduced when substituted for 5hmC (Fig. 3e).

To obtain additional insight into the determining factors of the effects of 5hmC on duplex stability, we then performed analyses (Fig. 4) in a different sequence context using isothermal titration calorimetry (ITC) under more biologically relevant conditions (37 °C, 200 mM NaCl).§ The concentrations of single-stranded DNA were chosen to enable reliable curve fitting (non-linear regression, ESI†) and therefore a simultaneous determination of K_b, Δ_bH° and Δ_bS° (Fig. S4a and S4b†) with comparable accuracy, and the sequence was selected to minimise secondary structure formation.

The titration curves for unmodified (C/C, Fig. 4a), 5hmC-containing (5hmC/5hmC, Fig. 4b) and 5mC-containing homoduplexes (5mC/5mC, Fig. 4c) were consistent with a single-site binding model.¶ The derived binding constants (Fig. 4d) show that methylation on both DNA strands causes stabilisation of ~10-fold (5mC/5mC *vs.* C/C, $P < 10^{-4}$) relative to unmodified duplex, whereas for 5hmC the binding constants were similar to those of unmethylated duplexes (~1.6-fold stabilisation of 5hmC/5hmC *vs.* C/C).

As for the stabilising effect of 5mC *vs.* C, the destabilising effect of 5hmC relative to 5mC was evident in multiple sequence contexts, being apparent in the binding constants derived from UVM (Fig. 3e) and ITC data (Fig. 4d). This trend was not limited to fully modified duplexes, but also

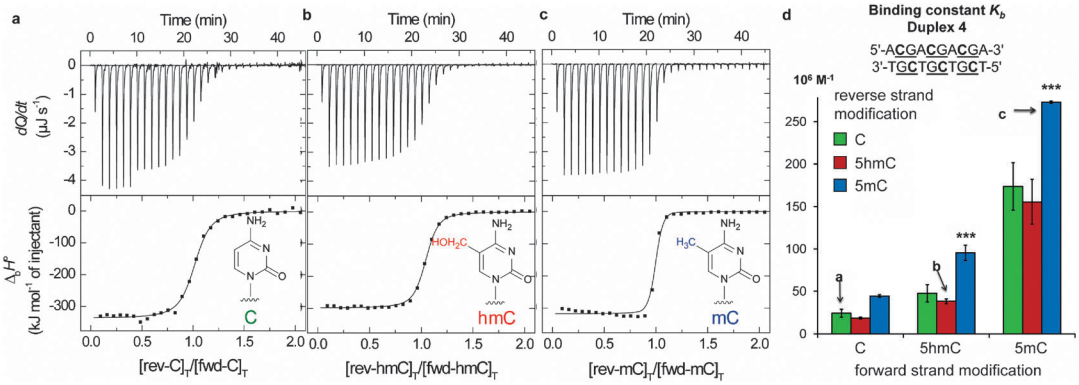

Fig. 4 ITC of duplexes with (a) unmodified C, (b) 5hmC and (c) 5mC. (a, b, c) Upper panels, data after baseline correction; Lower panels, binding isotherms from integration. (d) K_b values (mean ± s.d. of at least 2 independent expts., ESI†). Titration curves correspond to duplexes with same modification on both strands and are representative of expts. giving the indicated K_b values. (***denotes $P < 0.01$).

observed within ODN sets bearing a particular modification on the forward (groups of data bars in Fig. 4d) or reverse strands (data bars of same colour in Fig. 4d), with the binding constants of 5hmC-modified duplexes resembling those of unmodified rather than methylated duplexes. In the case of the duplex-stabilising effect of 5mC, high-level quantum calculations suggest enhanced stacking interactions[16,17] conferred by the hydrophobic methyl group may be responsible, thus raising the possibility of the reversal of these effects by hydroxylation.

Overall, the combined analyses reveal the potential of 5mC hydroxylation to, at least in part, reverse the thermodynamic duplex stabilisation conferred by 5mC in multiple sequence contexts, *without a requirement for demethylation or further modification*. Although the roles of proteins in the regulation of transcription have been extensively studied, to our knowledge the effect that methylation has 'directly' on DNA duplex stability has not been widely considered as a contributing factor in transcriptional regulation. With all other factors being equal, our results lead to the proposal that 5mC hydroxylation is more likely associated with transcriptional activation than repression (Fig. 1b). Consistent with our proposal, two recent studies have associated TET activity with transcriptional activation.[18,19] Of course, as for 5mC, the effects of 5hmC may be differentially manifested in the presence of additional factors, including, for example, DNA-binding[20] proteins such the CXXC domains which are part of the TET enzymes.[18] Given the established regulatory role of chromatin structure on transcription, TET-catalysed 5mC-hydroxylation of CpG islands in promoter regions of transcriptionally inactive genes could act to reduce duplex stability, facilitate strand separation or alleviate 5mC-controlled transcriptional repression by enabling recruitment of transcriptional machinery. Finally, given the role of the hypoxia inducible factor 2OG-dependent hydroxylases in the oxygen dependent regulation of transcription, it is possible that the 5hmC formation represents a mechanism for oxygen dependent regulation of expression, including in an epigenetic manner.

We thank the Biotechnology and Biological Sciences Research Council, Cancer Research UK (to A.T.), the European Community's 7th Framework Programme [FP7/2007-2013] grant [HEALTH-F4-2008-201418] READNA (A.H. E-S and T.B.), Pfizer and the Wellcome Trust for funding. We thank Kevin Dack for discussions.

Notes and references

§ The absolute values from UVM and ITC differ, because the techniques measure thermodynamic parameters at T_m, and at the exptl. temp., respectively. While parameters obtained at one temp. can be extrapolated to another, additional conformational 'melting' transitions of single strands occurring between these temperature need to be taken into account, precluding direct quantitative comparison. ¶ Table S3 gives data for titrations/accuracy of parameter fitting; Table S4 gives data for group averages of all unique combinations of C/5hmC/5mC duplexes.

1 J. T. Attwood, R. L. Yung and B. C. Richardson, *Cell. Mol. Life Sci.*, 2002, **59**, 241–257.
2 A. Portela and M. Esteller, *Nat. Biotechnol.*, 2010, **28**, 1057–1068.
3 M. Esteller, *Hum. Mol. Genet.*, 2007, **16**, R50–59.
4 S. Kriaucionis and N. Heintz, *Science*, 2009, **324**, 929–930.
5 M. Tahiliani, K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind and A. Rao, *Science*, 2009, **324**, 930–935.
6 S. Ito, A. C. D'Alessio, O. V. Taranova, K. Hong, L. C. Sowers and Y. Zhang, *Nature*, 2010, **466**, 1129–1133.
7 U. Bacher, C. Haferlach, S. Schnittger, A. Kohlmann, W. Kern and T. Haferlach, *Ann. Hematol.*, 2010, **89**, 643–652.
8 U. Bacher, S. Schnittger and T. Haferlach, *Curr. Opin. Oncol.*, 2010, **22**, 646–655.
9 M. A. McDonough, C. Loenarz, R. Chowdhury, I. J. Clifton and C. J. Schofield, *Curr. Opin. Struct. Biol.*, 2010, **20**, 659–672.
10 L. E. Xodo, G. Manzini, F. Quadrifoglio, G. A. van der Marel and J. H. van Boom, *Nucleic Acids Res.*, 1991, **19**, 5625–5631.
11 H. H. Klump and R. Loffler, *Biol. Chem. Hoppe-Seyler*, 1985, **366**, 345–353.
12 D. J. Prockop and K. I. Kivirikko, *Annu. Rev. Biochem.*, 1995, **64**, 403–434.
13 A. P. Hardy, I. Prokes, L. Kelly, I. D. Campbell and C. J. Schofield, *J. Mol. Biol.*, 2009, **392**, 994–1006.
14 A. S. Hansen, A. Thalhammer, A. H. El-Sagheer, T. Brown and C. J. Schofield, *Bioorg. Med. Chem. Lett.*, 2011, **21**, 1181–1184.
15 C. M. R. Lopez, B. G. Asenjo, A. J. Lloyd and M. J. Wilkinson, *Anal. Chem.*, 2010, **82**, 9100–9108.
16 C. Acosta-Silva, V. Branchadell, J. Bertran and A. Oliva, *J. Phys. Chem. B*, 2010, **114**, 10217–10227.
17 N. Tretyakova, R. Guza and B. Matter, *Nucleic Acids Symp. Ser.*, 2008, **52**, 49–50.
18 S. Ito, A. C. D'Alessio, O. V. Taranova, K. Hong, L. C. Sowers and Y. Zhang, *Nature*, 2010, **466**, 1129–1133.
19 H. Zhang, X. Zhang, E. Clark, M. Mulcahey, S. Huang and Y. G. Shi, *Cell Res.*, 2010, **20**, 1390–1393.
20 V. Valinluck, H.-H. Tsai, D. K. Rogstad, A. Burdzy, A. Bird and L. C. Sowers, *Nucleic Acids Res.*, 2004, **32**, 4100–4108.