VIP DNA Labeling

DOI: 10.1002/anie.201007169

Methyltransferase-Directed Derivatization of 5-Hydroxymethylcytosine in DNA**

*Zita Liutkevičiūtė, Edita Kriukienė, Indrė Grigaitytė, Viktoras Masevičius, and Saulius Klimašauskas**

The modification of cytosine by *S*-adenosylmethionine-dependent DNA methyltransferases is part of an intricate epigenetic regulation mechanism in vertebrates. DNA cytosine-5 methyltransferases (C5-MTases) catalyze the transfer of a methyl group from *S*-adenosyl-L-methionine (AdoMet or SAM, **2**)[1] to the cytosine (**1**) residue in CpG dinucleotides. Recent studies of genomic DNA from mouse embryonic stem cells, neurons, and the brain found that a substantial fraction of 5-methylcytosine (**3**) in CG sequences is converted into 5-hydroxymethylcytosine (hmC, **5**) by the action of 2-oxoglutarate- and Fe^{2+}-dependent oxygenases of the TET family.[2–6] As interactions of the 5-methyl- and 5-hydroxymethyl groups with cellular proteins in DNA are distinct,[7,8] hmC residues may play an independent role in yet unknown epigenetic pathways during embryonic development, brain functioning, and cancer progression. However, further studies of these intriguing phenomena are hampered by the lack of efficient analytical techniques for mapping hmC residues in the genome.[8] Herein we show that C5-MTases can direct the condensation of exogenous thiols and selenols with hmC in DNA to yield the corresponding 5-chalcogenomethyl derivatives. These transformations open new possibilities for the sequence-specific derivatization and analysis of this newly discovered epigenetic mark in mammalian DNA.

The production of hmC and its homologues in vitro was recently demonstrated by the methyltransferase-directed addition of exogenous aliphatic aldehydes, such as formaldehyde (**4**), to cytosine;[9] it was also found that DNA C5-MTases are capable of removing the 5-hydroxymethyl group from their target residues in vitro. Natural methylation by C5-MTases proceeds through the direct transfer of the sulfonium-bound methyl group from SAM to the C5 position of the cytosine ring, which is activated by the nucleophilic addition of a conserved cysteine residue in the enzyme to the 6-position of the ring (Scheme 1 a).[10–12]

To better understand the mechanism of these novel enzymatic reactions, we examined the behavior of hmC residues in DNA duplexes (see Table 1) in the presence of C5-MTases under a variety of reaction conditions. Remarkably, we found that the presence of an exogenous thiol compound, such as 2-mercaptoethanol (**6**), cysteamine (**7**), or L-cysteine (**8**), at a millimolar concentration led to the appearance of a new product at the expense of the original hmC (Figure 1). ESI-MS and UV spectral analyses of the new compounds showed that the thiol replaces a hydroxy group in the hmC residue, and the sulfur atom is directly attached to the nucleoside (see Table S1 and Figures S1 and S2 in the Supporting Information). Taking into account the observation that substrates containing C or mC at the target position were completely inert in the reaction (data not shown), and on the basis of the documented reactivity of the 5-hydroxymethyl group towards a variety of nucleophiles,[7,13,14] we presumed that addition of the thiol occurs at the 5-hydroxymethyl group to yield 5-(2-hydroxyethyl)thiomethylcytosine (**13**), 5-(2-aminoethyl)thiomethylcytosine (**14**), and *S*-(5-methylcytosinyl)-L-cysteine (**15**) residues in DNA, respectively (Scheme 1). The identity of the isolated nucleoside **13** was confirmed by direct chromatographic comparison of its nitrous acid deamination product with a chemically synthesized 5-(2-hydroxyethyl)thiomethyl-2′-deoxyuridine (see Figure S3 in the Supporting Information). Similar reactivity but even at lower

Table 1: Structure of short hmC-containing duplex substrates.

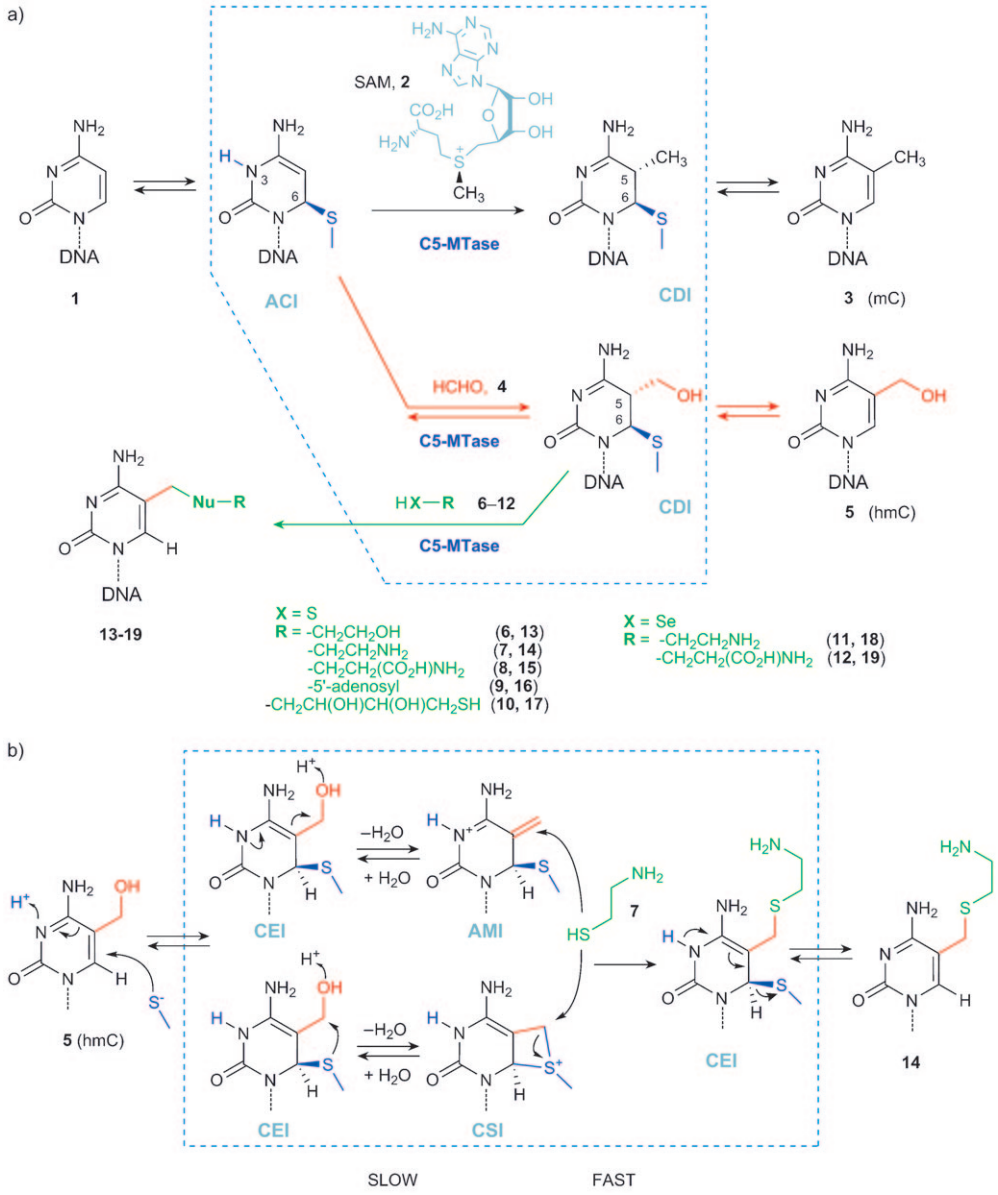
Sequence[a]	Short name
5′–TAATAATGhGCTAATAATAATAAT	GhGC/GhGC
3′–ATTATTACGhGATTATTATTATTA	
5′–TAGAGTATGATGMGCTGACCCACAACATCCG	G*HGC/GMGC
3′–ATCTCATACTACGHGACTGGGTGTTGTAGGC	
5′–TCGGATGTTGTGGGTCAG*hG**C**ATGATAGTGTA*	G*hG*C/GMGC
3′–AGCCTACAACACCCAGTCGMGTACTATCACAT	

[a] M, 5-methylcytosine; H, 5-hydroxymethylcytosine (**5**; synthetically incorporated); h, 5-hydroxymethylcytosine (introduced by the enzymatic hydroxymethylation with formaldehyde and M·HhaI). M·HhaI/M·SssI target sites are highlighted in gray, target nucleotides are underlined, polymerase-extended nucleotides are italicized, and ^{33}P-labeled cytosine nucleotides are shown in boldface.

[*] Z. Liutkevičiūtė, Dr. E. Kriukienė, I. Grigaitytė, Dr. V. Masevičius, Prof. S. Klimašauskas
Department of Biological DNA Modification
Institute of Biotechnology, Vilnius University
V. A. Graičiūno 8, 02241 Vilnius (Lithuania)
Fax: (+370) 5-260-2116
E-mail: klimasau@ibt.lt
Homepage: http://www.ibt.lt/dmtl_en.html
Dr. V. Masevičius
Faculty of Chemistry, Vilnius University
Naugarduko 24, 03225 Vilnius (Lithuania)

[**] We thank G. Urbanavičiūtė and G. Vainorius for enzyme preparations, and M. Krenevičienė for technical assistance. We also thank S. Tumkevičius, G. Lukinavičius, D. Daujotytė, and A. Petronis for fruitful discussions. This research was supported by the Lithuanian State Science and Study Foundation (P-07003), the Research Council of Lithuania (student research fellowship to I.G.), the FP7-REGPOT-2009-1 program (project 245721 MoBiLi), and the US National Institutes of Health (1R21HG005758).

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201007169.



Scheme 1. Transformations of a target residue by DNA cytosine-5 methyltransferases. a) Biological methylation by C5-MTases occurs through an S_N2 reaction between an activated cytosine intermediate (ACI) and the cofactor SAM (**2**) to give a covalent 5,6-dihydrocytosine intermediate (CDI) and then 5-methylcytosine (mC, **3**). The activated cytosine residue can also participate in reversible addition reactions (red-colored route) with formaldehyde (**4**) to yield 5-hydroxymethylcytosine (**5**, hmC). The latter compound, including hmC residues in natural DNA, can undergo further methyltransferase-directed condensation (green-colored route) with thiol or selenol reagents (**6**–**12**) to give stable 5-alkyl chalcogenomethyl derivatives (**13**–**19**). b) Proposed mechanism of nucleophile condensation at the target hmC (**5**) residue catalyzed by a C5-MTase. The reaction is thought to proceed via a covalent enamine intermediate (CEI) followed by either a highly active 5-methylene intermediate (AMI) or a bicyclic sulfonium intermediate (CSI), which would then undergo the fast addition of a nucleophilic reagent, such as **7**. The C5-MTase and its catalytic moieties are shown in blue; boxed areas denote species and reactions within the catalytic center of the enzyme.

concentrations was observed (Figure 1) with selenols, such as selenocysteamine (**11**) and L-selenocysteine (**12**). 5′-Deoxy-5′-thioadenosine (**10**), which can be regarded as a simple cofactor mimic owing to a potential anchoring interaction of the adenosine moiety with the enzyme, was also reactive at submillimolar concentrations. Other types of nucleophiles, such as hydrazine, *p*-nitrophenol, phenol, methanol, sodium sulfide, sodium azide, potassium bromide, or sodium iodide, showed no detectable reaction.

Besides M·HhaI, another bacterial C5-MTase, M·SssI, was examined and also showed clearly detectable catalytic activity at its target sites (see Figure S4 in the Supporting Information). The generality of this phenomenon indicates that one role of the enzyme is to direct the reaction by flipping out and exposing in the catalytic center a residue that occurs at its target position.[11] Moreover, we found that the reaction requires the presence of the catalytic cysteine residue in the enzyme (Figure 1 a), a result that suggests a covalent reaction intermediate (Scheme 1).

The most straightforward mechanism is a direct S_N2 attack of the nucleophile at the protonated 5-hydroxymethyl group. However, this mechanism does not agree with our finding that the rate of adduct formation is fairly independent of whether a thiol or the corresponding selenol was used in the reaction. On the other hand, a seleno product always prevailed over a thio product even in the presence of an excess amount of thiol (see Figures S5 and S6 in the Supporting Information). These observations clearly point to a mechanism in which the formation of an active intermediate is the slow rate-limiting step, whereas the relative strength of the nucleophile determines the nature of the final product. On the basis of previously proposed mechanisms for thymidylate synthase[15] and deoxycytidylate hydroxymethylase,[16] one could assume that the reaction proceeds by an acid-assisted dehydration at the 5-hydroxymethyl group to give a highly electrophilic 5-methylene intermediate (activated methylene intermediate, AMI;

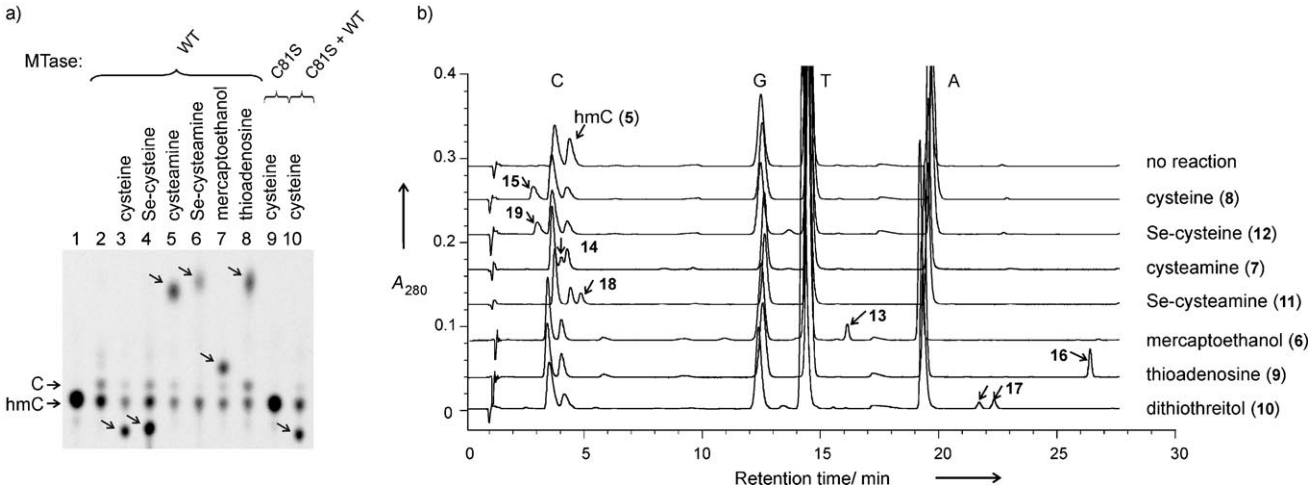


Figure 1. Methyltransferase-directed coupling of thiols and selenols with 5-hydroxymethylcytosine residues in DNA. a) TLC analysis of a modified ^{33}P-labeled hmC nucleotide. DNA containing hmC at the target site of M·HhaI was treated with 12 mM L-cysteine, 12 mM L-selenocysteine, 12 mM cysteamine, 12 mM selenocysteamine, 150 mM 2-mercaptoethanol, or 0.5 mM 5′-thio-5′-deoxyadenosine in the presence of 200 nM M·HhaI (wild type (WT), a catalytic C81S mutant, or both) for 1.5 h at room temperature. Modified DNAs were digested to 5′-mononucleotides, analyzed by TLC, and autoradiographed. b) Reversed-phase HPLC analysis of modified 2′-deoxynucleosides. A GhGC/GhGC duplex (13 μM) was treated with an exogenous reagent as above (50 mM L-cysteine, 50 mM L-selenocysteine, 12 mM cysteamine, 12 mM selenocysteamine, 0.5 mM 5′-thio-5′-deoxyadenosine, 500 mM 2-mercaptoethanol, or 50 mM dithiothreitol (DTT)) in the presence of 13 μM M·HhaI for 1 h at room temperature. Modified DNAs were digested to nucleosides and analyzed by HPLC–MS methods. Small arrows indicate peaks corresponding to modified nucleosides.

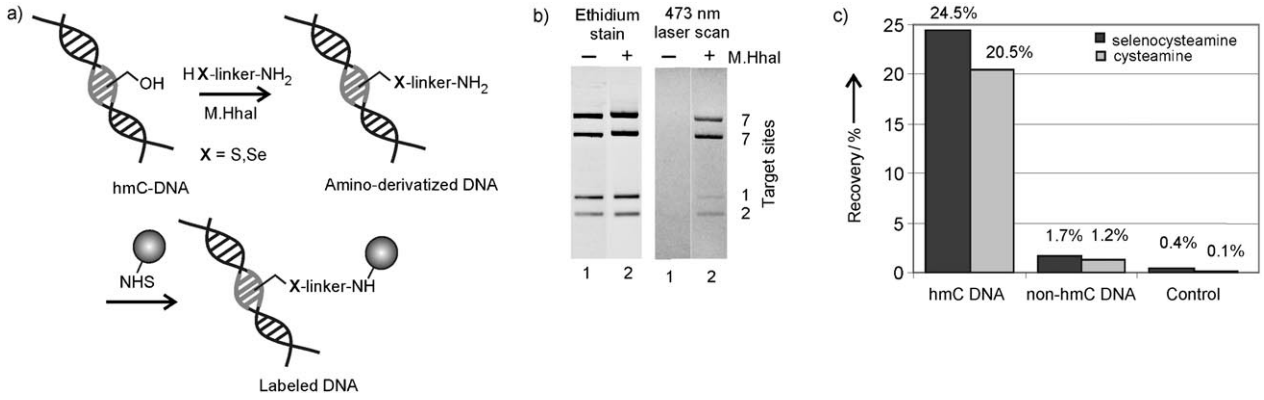


Figure 2. Sequence-specific covalent derivatization and labeling of hmC-containing DNA. a) General approach to the selective derivatization and labeling of hmC-containing DNA with a reporter group (shown as a ball). DNA bases in the HhaI target site are shown as gray sticks. b) Fluorescence labeling of a pUC19 plasmid containing hmC residues at the HhaI target sites. The DNA was amino-derivatized by treatment with L-cysteine in the presence of M·HhaI, and the derivatized DNA was labeled with fluorescein *N*-hydroxysuccinimidyl ester, fragmented with R·FspBI, and analyzed by agarose gel electrophoresis. Fluorescein imaging of the labeled fragments was performed with a 473 nm laser scanner (right panel); bulk DNA fragments were visualized after staining with ethidium bromide (left panel). In each case, lane 1 is the control with M·HhaI omitted. c) A mixture of a hmC-containing and non-hmC DNA fragments was treated with selenocysteamine or cysteamine in the presence of M·HhaI (no M·HhaI in the control), followed by labeling with biotin *N*-hydroxysuccinimidyl ester. Streptavidin-coated magnetic beads were added and then washed to remove unbound DNA. Fragment recovery was determined by on-bead real-time qPCR analysis.

Scheme 1b). However, this high-energy exomethylene compound could in principle be avoided by an intramolecular attack of the enzyme-borne C6-bound sulfur atom onto the protonated 5-hydroxymethyl group. This reaction would give a bicyclic sulfonium intermediate (CSI) containing a four-membered thietane ring. This route requires a substantial degree of conformational plasticity in the active site for the catalytic sulfur atom to approach the exocyclic target. Such conformational flexibility can be predicted from molecular-dynamics studies of M·HhaI,[17] but whether the proposed bicyclic structure is well-compatible with the active sites of other C5-MTases[11] and related enzymes (e.g. thymidylate synthase, deoxycytidylate hydroxymethylase) remains to be determined. In both cases, subsequent addition/attack of a nucleophile would readily give the observed product, whereas the addition of water would promote the reverse reaction to give the original hydroxymethylated cytosine residue.

The described novel reactions of C5-MTases to catalyze the condensation of non-cofactor-like compounds with the exocyclic hydroxymethyl group of a modified cytosine residue provide another spectacular example of the catalytic versatility of these cofactor-dependent enzymes.[9] Moreover, the reactions occur under mild conditions and retain the high sequence and base specificity characteristic of bacterial DNA MTases. They thus offer new approaches to the sequence-specific derivatization and labeling of DNA. For example, the cysteine condensation product **15** contains an aliphatic primary amino group (Scheme 1). We thus examined whether this or similar modifications can be exploited as chemical anchors for the covalent attachment of reporter groups to DNA (Figure 2a). Plasmid pUC19 DNA that was previously 5-hydroxymethylated at the target cytosine residue was treated with L-cysteine in the presence of M·HhaI and then labeled with fluorescein by treatment with an *N*-hydroxysuccinimide (NHS) ester. The fluorescence-intensity distribution in four pUC19-FspBI fragments was consistent with the positions and numbers of the HhaI sites in the original plasmid (Figure 2b). Moreover, since the MTase-directed condensation of nucleophiles is not possible at 5-methylated and unmodified cytosine residues, the derivatization reaction is well-suited to query the hydroxymethylation status of CG sites in mammalian genomic DNA. Such investigations can be readily carried out by combining the addition of cysteamine or selenocysteamine with amino-selective biotinylation. The labeled DNA is selectively retained on streptavidin beads, and the enriched hmC fractions can then be analyzed by real-time quantitative PCR (qPCR; Figure 2c), the use of DNA microarrays, or sequencing (manuscript in preparation). The above examples demonstrate the unique potential of this chemoenzymatic approach for the genome-wide mapping of epigenetic cytosine modifications in DNA.

Received: November 15, 2010
Published online: January 26, 2011

Keywords: DNA labeling · enzyme mechanisms · epigenetics · methyltransferases · nucleophilic addition

[1] M. G. Goll, T. H. Bestor, *Annu. Rev. Biochem.* **2005**, *74*, 481–514.
[2] S. Kriaucionis, N. Heintz, *Science* **2009**, *324*, 929–930.
[3] M. Tahiliani, K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind, A. Rao, *Science* **2009**, *324*, 930–935.
[4] S. Ito, A. C. D'Alessio, O. V. Taranova, K. Hong, L. C. Sowers, Y. Zhang, *Nature* **2010**, *466*, 1129–1133.
[5] M. Münzel, D. Globisch, T. Brückl, M. Wagner, V. Welzmiller, S. Michalakis, M. Müller, M. Biel, T. Carell, *Angew. Chem.* **2010**, *122*, 5503–5505; *Angew. Chem. Int. Ed.* **2010**, *49*, 5375–5377.
[6] A. Szwagierczak, S. Bultmann, C. S. Schmidt, F. Spada, H. Leonhardt, *Nucleic Acids Res.* **2010**, *38*, e181.
[7] S. Tardy-Planechaud, J. Fujimoto, S. S. Lin, L. C. Sowers, *Nucleic Acids Res.* **1997**, *25*, 553–558.
[8] S.-G. Jin, S. Kadam, G. P. Pfeifer, *Nucleic Acids Res.* **2010**, *38*, e125.
[9] Z. Liutkevičiūtė, G. Lukinavičius, V. Masevičius, D. Daujotytė, S. Klimašauskas, *Nat. Chem. Biol.* **2009**, *5*, 400–402.
[10] J. C. Wu, D. V. Santi, *J. Biol. Chem.* **1987**, *262*, 4778–4786.
[11] S. Klimasauskas, S. Kumar, R. J. Roberts, X. Cheng, *Cell* **1994**, *76*, 357–369.
[12] M. Peräkylä, *J. Am. Chem. Soc.* **1998**, *120*, 12895–12902.
[13] R. E. Cline, R. M. Fink, K. Fink, *J. Am. Chem. Soc.* **1959**, *81*, 2521–2527.
[14] H. Hayatsu, M. Shiragami, *Biochemistry* **1979**, *18*, 632–637.
[15] J. E. Barrett, D. A. Maltby, D. V. Santi, P. G. Schultz, *J. Am. Chem. Soc.* **1998**, *120*, 449–450.
[16] K. L. Graves, M. M. Butler, L. W. Hardy, *Biochemistry* **1992**, *31*, 10315–10321.
[17] E. Y. Lau, T. C. Bruice, *J. Mol. Biol.* **1999**, *293*, 9–18.

