VILNIUS UNIVERSITY

Zita Liutkevičiūtė

DNA CYTOSINE METHYLTRANSFERASE-DIRECTED REACTIONS INVOLVING NON-COFACTOR-LIKE COMPOUNDS

Summary of doctoral dissertation
Physical science, biochemistry (04 P)

Vilnius, 2012

This study has been carried out at the Institute of Biotechnology, Vilnius University, during 2007-2011.

Scientific supervisor Prof. habil. dr. **Saulius Klimašauskas** (Vilnius University, physical sciences, biochemistry – 04 P)

Evaluation board of dissertation of Biochemistry trend

Chairman

Prof. dr. **Virginijus Šikšnys** (Vilnius University, physical sciences, biochemistry – 04 P)

Reviewers

Prof. dr. **Vida Kirvelienė** (Vilnius University, physical sciences, biochemistry – 04 P)

Prof. habil. dr. **Eugenijus Butkus** (Vilnius University, physical sciences, chemistry – 03 P)

Prof. dr. **Vilmantė Borutaitė** (Lithuanian University of Health Sciences, physical sciences, biochemistry – 04 P)

Assoc. prof. dr. **Jolanta Sereikatė** (Vilnius Gediminas Technical University, physical sciences, biochemistry – 04 P)

Opponents

Dr. **Rolandas Meškys** (Vilnius University, physical science, biochemistry – 04 P)

Dr. **Eglė Merkienė** (Thermo Fisher Scientific, physical science, biochemistry – 04 P).

The thesis defence will take place at the Institute of Biotechnology, Vilnius University (Graičiūno 8, LT-02241 Vilnius, Lithuania) on 29th of March, 2011, at 11 a.m.

The summary of doctoral dissertation was sent on of 20th of February, 2012.

The thesis is available at the Library of Institute of Biotechnology and at the Library of Vilnius University.

VILNIAUS UNIVERSITETAS

Zita Liutkevičiūtė

DNR CITOZINO METILTRANSFERAZIŲ REAKCIJOS, NEDALYVAUJANT KOFAKTORIUI

Daktaro disertacijos santrauka
Fiziniai mokslai, biochemija (04 P)

Vilnius, 2012

Disertacija rengta 2007-2011 m. Vilniaus universiteto Biotechnologijos institute.

Mokslinis vadovas Prof. habil. dr. **Saulius Klimašauskas** (Vilniaus universitetas, fiziniai mokslai, biochemija – 04 P).

Disertacija ginama Vilniaus universiteto Biochemijos mokslo krypties taryboje:

Pirmininkas Prof. dr. **Virginijus Šikšnys** (Vilniaus universitetas, biochemija – 04 P)

Nariai Prof. dr. **Vida Kirvelienė** (Vilniaus universitetas, fiziniai mokslai, biochemija – 04 P)

Prof. habil. dr. **Eugenijus Butkus** (Vilniaus universitetas, fiziniai mokslai, chemija – 03 P)

Prof. dr. **Vilmantė Borutaitė** (Lietuvos sveikatos mokslų universitetas, fiziniai mokslai, biochemija – 04 P)

Doc. dr. **Jolanta Sereikaitė** (Vilniaus Gedimino technikos universitetas, fiziniai mokslai, biochemija – 04 P)

Oponentai Dr. **Rolandas Meškys** (Vilniaus universitetas, fiziniai mokslai, biochemija – 04 P)

Dr. **Eglė Merkienė** (Thermo Fisher Scientific, fiziniai mokslai, biochemija – 04 P)

Disertacija bus ginama viešame Biochemijos mokslo krypties posėdyje 2012 m. kovo 29 d. 11 val. Vilniaus Universiteto Biotechnologijos instituto konferencijų salėje.
Adresas: Graičiūno 8, LT-02241 Vilnius, Lietuva.

Disertacijos santrauka išsiuntinėta 2012 m. vasario mėn. 20 d.
Disertaciją galima peržiūrėti Vilniaus universiteto Biotechnologijos instituto ir Vilniaus universiteto bibliotekose.

CONTENT

INTRODUCTION

DNA is composed of four major heterocyclic bases: thymine (T), adenine (A), guanine (G) and cytosine (C). The fifth minor base – 5-methylcytosine (mC) – in eukaryotic DNA was discovered few years before the famous DNA structure model was proposed by Watson and Crick (Hotchkiss, 1948; Watson and Crick, 1953) and remained an intensive research object till nowadays (Razin and Riggs, 1980; Zemach et al., 2010). In eukaryotes, mC is an important epigenetic marker, that is involved in regulation of gene expression, silencing of parasitic genomic elements and X chromosome inactivation in mammals; it also participates in imprinting, mutagenesis, DNA lesions repair and recombination (Goll and Bestor, 2005). In genomic DNA, mC originates by action of methyltransferases (MTases), which transfer the methyl group from cofactor S-adenosyl-L-methionine (SAM) to cytosine in DNA. In contrast to the quite well studied genetics and biology of DNA methylation, the enzymatic mechanism of demethylation remains elusive and controversial (Wu and Zhang, 2010). Three years ago researches on DNA demethylation became very intensive when the sixth base – 5-hydroxymethylcytosine (hmC) – was discovered in eukaryotic DNA (Kriaučionis and Heintz, 2009; Tahiliani et al., 2009). Shortly, other two minor DNA bases 5-formylcytosine (fC) and 5-carboxycytosine (caC) were found in embryonic stem cell though in very small amounts (He et al., 2011; Ito et al., 2011). The later three modifications (hmC, fC and caC) originate from mC in the presence of newly discovered eukaryotic oxygenases Tet1,2,3 (Tahiliani et al., 2009). The biological function of hmC in DNA demethylation and epigenetic regulation is still to be determined, however such studies are hampered by the lack of adequate methods that could unequivocally distinguish mC from hmC (Nabel et al., 2011).

The prokaryotic MTase M.HhaI was first shown to flip the target cytosine out from the DNA helix and into the active site of the enzyme (Klimašauskas et al., 1994), and M.HhaI since then became a paradigm for structural and mechanistic studies of DNA cytosine-5 methyltransferases. Subsequent studies showed that base flipping is a fundamental mechanism in DNA modification and repair, and is also used by proteins responsible for the opening of the DNA or RNA helix during replication, transcription and recombination (Cheng and Blumenthal, 1996). Since many more proteins are expected to employ this mechanistic feature in their interactions with DNA, a fast method for initial screening would be very useful. In this work, we assessed the suitability of 2-chloroacetaldehyde (Caa), which was previously shown to react with unpaired bases (Kohwi-Shigematsu and Kohwi, 1992), for chemical detection of flipped out bases in protein–DNA complexes. Our results confirmed that DNA base flipping by MTases or restriction endonucleases (REases) can indeed be detected by treatment of preformed protein-DNA complexes with Caa. Detailed analysis of the reaction products revealed that different cytosine adducts are formed in single stranded DNA and in C5-MTase-DNA complexes upon treatment with Caa. This finding prompted a more systematic examination of several classes of electrophilic compounds. Unexpectedly, we found that, in the absence of cofactor SAM, the C5-MTases catalyze reversible covalent addition of small aliphatic aldehydes to the C5 position of target cytosine producing 5-(1-hydroxyalkyl)cytosine residues in DNA. Importantly, C5-MTases perform the formaldehyde coupling reaction with high efficiency yielding hmC

– the base that was recently found in vertebrate genomic DNA. Furthermore, we discovered that C5-MTases can direct the condensation of exogenous thiols and selenols with hmC in DNA to yield corresponding 5-chalcogenomethyl derivatives *in vitro*. At this point it is not clear if similar reactions occur *in vivo*, and therefore the roles of MTases in DNA modification (methylation and hydroxymethylation) and its reversal need to be carefully revised.

The specific aims of this study were:

1) To investigate the reactivity of 2-haloacetaldehydes in DNA-protein complexes and to adapt these reactions to study DNA base flipping by DNA cytosine-5 methyltransferases and restriction endunucleases.

2) To examine the conditions, products and mechanism of the newly discovered methyltransferase-directed reversible coupling of small electrophilic compounds to target cytosine residues in DNA.

3) To examine the reaction conditions, products and mechanism of the newly discovered methyltransferase-directed condensation of nucleophilic compounds with 5-hydroxymethylcytosine residues in DNA.

4) To examine if the newly discovered atypical reactions of DNA cytosine-5 methyltransferases can be used as molecular tools for genomic research.

Scientific novelty

It is shown here for the first time that DNA cytosine-5 methyltransferases can perform two atypical reactions in the absence of cofactor *in vitro*: i) reversible covalent addition of short aliphatic aldehydes to the target cytosine producing 5-(1-hydroxyalkyl)cytosine, and ii) coupling thiols and selenols to 5-hydroxymethylated cytosine residues producing corresponding 5-chalcogenomethylcytosines. These are the first demonstrations of reactions involving non-cofactor-like substrates catalyzed by SAM-dependent enzymes, which reveal atypical versatility of DNA cytosine-5 methyltransferases. Mechanistic insights into of these reactions indicate that the key step in both cases is the covalent activation of the flipped out target residue by the enzyme.

It is also demonstrated here that 2-chloroaldehyde can be used for chemical detection of flipped out nucleotides in MTase-DNA and REase-DNA complexes.

Practical value

The new discovered atypical reactions open new ways for developing molecular tools for epigenome analysis. First of all, the aldehyde coupling reaction can be used to introduce hmC residues into DNA substrates in a sequence-specific manner, which may often be a useful alternative to oligonucleotide synthesis or PCR-based approaches. Such hmC-containing DNA substrates serve as model systems in a variety of functional and genomic studies. Secondly, the MTase-directed reversal reaction can erase hmC modifications at CpG sites, and thus bisulfite sequencing of the original and MTase-treated DNA can reveal the position of hmC residues in genomic DNA. Most importantly, it is demonstrated here that the MTase-directed condensation of thiols with hmC in DNA can be exploited for covalent labelling of the hmC residues permitting subsequent enrichment of corresponding fragments from genomic pools. hmC is a recently discovered cytosine modification in mammals whose biological function and distribution are not yet established.

MATERIALS AND METHODS

Reagents

2-mercaptoethanol, L-cysteine, L-selenocystine, cystamine, selenocystamine were purchased from "Sigma-Aldrich"; 1,4-dithiothreitol (DTT) was purchased from "Fermentas". Cysteamine, selenocysteamine and L-selenocysteine were prepared from corresponding commercial disulfides and diselenides by incubation with a 4-fold excess of DTT in water and were used without further purification. [γ-^{33}P]ATP, [α-^{33}P]CTP (specific activity >2500 Ci/mmol) and [γ-^{32}P]ATP (specific activity >5000 Ci/mmol were obtained from "Hartmann analytic".

Enzymes

M.HhaI and its mutants were prepared as described earlier (Vilkaitis et al., 2000). Mouse Dnmt1-dN, a truncated form of Dnmt1 lacking the N-terminal 290 amino acid residue region, was prepared as previously described (Vilkaitis et al., 2005). Restriction endonucleases R.Ec18kI, R.PspGI and R.BcnI were prepared by dr. Gintautas Tamulaitis as described earlier (Tamulaitis et al., 2002; Pingoud and kt, 2003; Sokolowska et al., 2007). M.SssI was prepared by Giedrė Urbanavičiūtė (Kriukienė et al., manuscript in submission). M.AluI, M.SssI, M.HaeIII, M.HpaII and M.MspI were purchased from "New England Biolabs". Restriction endonucleases, CIAP, FastAP, nuclease BAL31, T4 DNA ligase, T4 polynucleotide kinase, Klevow Fragment exo$^-$ and Taq DNA polymerases were obtained from "Fermentas". Nuclease P1 was purchased from "Sigma-Aldrich".

Buffers and other solutions

M.AluI and other MTases buffer for reactions with aldehydes: 25 mM MOPS, 25 mM MES (pH 7.5), 1 mM Na_2EDTA, 15 mM NaCl, 0.2 mg/ml BSA.

Restriction endonucleases buffer: 33 mM Tris-OAc (pH 7.9), 66 mM K-OAc, 10 mM $CaCl_2$, 0.2 mg/ml BSA.

P1 nuclease buffer: 10 mM Na-OAc (pH 5.2), 1 mM $Zn\text{-}OAc_2$.

0.5x TBE – 44.5 mM Tris-HCl (pH 8.0), 44.5 mM H_3BO_3, 2 mM EDTA.

Tris-OAc/$Ca\text{-}OAc_2$ – 40 mM Tris-OAc (pH 8.3), 10 mM $Ca\text{-}OAc_2$.

M.HhaI buffer for reactions with aldehydes: 50 mM MOPS, 50 mM MES (pH 7.0), 1 mM Na_2EDTA, 15 mM NaCl, 0.2 mg/ml BSA.

M.HhaI reaction buffer (TEN): 10 mM Tris-HCl (pH 7.4), 50 mM NaCl, 0.5 mM Na_2EDTA 0.2 mg/ml BSA.

M.HhaI storage buffer: 10 mM $K\text{-}PO_4$ (pH 7.4), 5 mM EDTA, 10 mM 2-mercaptoethanol, 100 mM NaCl, 50% glycerol.

Buffers used for thiols and selenols coupling:

Tris-HCl buffer: 50 mM Tris-HCl (pH 7.5, 8.0 and 8.0), 0.2 mg/mL BSA, 0.5 mM EDTA.

Citrate buffer: 15 mM Na-citratas (pH 5.0, 5.5, 6.0 and 6.5), 0.2 mg/mL BSA.

Phosphate buffer: 10 mM $Na\text{-}PO_4$ (pH 6.5, 7.0, 7.5 and 8.0), 0.5 mM Na_2EDTA, 0.2 mg/ml BSA.

Acetate buffer: 50 mM sodium acetate (pH 5.0, 5.5, 6.0, 6.5 and 7.4), 0.2 mg/ml BSA.

Oligonucleotides

DNA oligonucleotides (HPLC-purified grade) were obtained from IDT (USA), IBA (Germany) or Metabion (Germany). Synthetic hmC oligonucleotides always contained a substantial fraction of the O-cyanoethyl protected precursor and were subjected to an additional round of deprotection with 0.1 N sodium hydroxide at 65 °C for 7 hours.

Table 1. Oligonucleotides duplex substrates. M, 5-methylcytosine; H, 5-hydroxymethylcytosine (synthetically incorporated); MTases target sites are in bold or highlighted in grey, target nucleotides or ^{33}P-labeled cytosine nucleotides in TLC DNA substrates are underlined, polymerase-extended nucleotides are italicized.

Sequence	Number of base	Short name
DNA used in denaturing PAA gels		
5‘-GTAATAGACTGCACGACGCGCCAGGCCGGCGAGCTTTACGAT 3‘-ATTATCTGACGTGCTGCGCGGTCCGGCCGCTCGAAATGCTAT	42	uCG/lCG-42
5'-TACAGTATCAG**GCGC**TGACCCACAA 3′- TGTCATAGTC**CGMG**ACTGGGTGTTG	25	GCGC/GMGC-25
5'-TACAGTATCAG**GCGC**TGACCCACAA 3′- TGTCATAGTC**CCMG**ACTGGGTGTTG	25	GCGC/GMCC-25
5′-TGACCCACGCTCG**CCCGG**CGACACATTACGT 3′-ACTGGGTGCGAGC**GGGCC**GCTGTGTAATGCA	31	CCCGG/GGGCC-31
5′-TGACCCACGCTCG**CCCGG**CGACACATTACGT 3′-ACTGGGTGCGAGC**GGCCC**GCTGTGTAATGCA	31	CCCGG/GGCCC-31
DNA used in HPLC experiments		
5'-TAATAAT**GCGC**TAATAATAATAAT 3'-ATTATTA**CGCG**ATTATTATTATTA	24	GCGC/GCGC-24
DNA used in TLC experiments		
5'-CCAGCTGCATTAATGAATC	19	*CCN-19
5'-TGACCCACGCTCG**CC*CGG***AGATAAATTATGT 3'-ACTGGGTGCGAGC**GGGCC**TCTATTTAATACA	31	CC*CGG/CCGGG-31
5'-TCGGATGTTGTGGGTCA***GCGC***ATGATAGTGTA 3'-AGCCTACAACACCCAGT**CGCG**TACTATCACAT	32	G*CG*C/GCGC-32
5'-TCGGATGTTGTGGGTCA***GCGC***ATGATAGTGTA 3'-AGCCTACAACACCCAGT**CGMG**TACTATCACAT	32	G*CG*C/GMGC-32
5'-CGCGCCATTCCTGCG**AG*CT***TTAGGATA 3'-GCGCGGTAAGGACGC**TCGA**AATCCTAT	27	AG*CT/AGCT-27
5‘-CGGATGTTGTGGGTCA**GHGC**ATCATACTCTA 3‘-GCCTACAACACCCAGT**CGMG**TAGTATGAGAT	31	G*HGC/GMGC-31

Preparation of DNA substrates

If necessary oligonucleotides were 5’-labeled using a T4 polynucleotide kinase and [γ-^{33}P]ATP. DNA duplexes were prepared by annealing appropriate oligonucleotide strands. Internally labeled CC*CGG/CCGGG-31, G*CG*C/GCGC-32, G*CG*C/GMGC-32 and AG*CT/AGCT-27 duplex substrates for TLC experiments were prepared by Klenow Fragment exo$^-$ Polymerase extension of synthetic DNA duplexes in the presence of labeled [α-^{33}P]dCTP. Labeled G*HGC/GMGC duplex for TLC experiments was prepared as follows: synthetic DNA oligonucleotide 5‘-

HGCATCATACTCTA was 5'-labeled using T4 polynucleotide kinase, purified by ethanol precipitation, annealed with DNA oligonucleotides 5'TAGAGTATGATGCGCTGACCCACAACATCCG and 3'-GACTGGGTGTTGTAGGC and ligated using T4 DNA ligase.

618 bp DNA fragment was prepared by standard PCR method using pUC19 plasmid DNA template and primers 5-AACGTTGTTGCCATTGCTAC and 5-GCTCATGAGACAATAACCCTGA. PCR fragments were purified with PCR Purification kit („Qiagen").

To monitor the efficiency and specificity of the hmC labeling reaction, a pair of two DNA fragments (specific (230 bp) and control (200 bp)) was designed. The specific DNA fragment contains two M.SssI target site, and the control DNA fragment contains no M.SssI sites. Both DNA probes were prepared by PCR amplification using any mouse line DNA template for specific probe and C57BL/6J mouse line DNA template for non-specific probe and two sets of primers: (1) 5'-GCCCATGTCGCTGTG and 5'-AAGATGTGTCCHGGCTGGCCTATACTCAGHGC; 2) 5'-GCCCACTTCACTTCTTGTG and 5'-AGGCCAAAAGAAAGAAGAGAT (H – 5-hydroxymethylcytosine) for the specific and non-specific probe, respectively. DNA probes were purified from agarose gel with Gel Extraction kit (Qiagen). These probes were prepared by dr. Edita Kriukienė.

Electrophoretic gel mobility shift assays

Typically, reactions were incubated for 20 min at 20°C or 37°C, and 1–3 μl samples were then loaded onto an 8% polyacrylamide (PAA) gel (37.5:1 crosslink ratio). Electrophoresis was performed in 45 mM Tris-borate (pH 8.3), 1 mM Na_2EDTA for 1–2 h at 10 V/cm (for M.HhaI, M.AluI, M.SssI and M.HpaII) or 40 mM Tris-Acetate (pH 8.3), 10 mM Ca-Acetate for 2–3 h at 5 V/cm (for R.Ecl18kI, R.PspGI and R.BcnI).

Denaturing polyacrylamide (PAA) gel

After the incubation with Caa and enzyme DNA was precipitated with ethanol, redissolved in 1 M piperidine, heated at 90°C for 30 min, and lyophilized. Lyophilized material was resuspended in gel loading solution (STOP solution, "Fermentas") and applied to a 15-20% PAA gel (19:1 crosslink ratio) with 7 M urea. Electrophoresis was performed in 90 mM Tris-borate (pH 8.3), 2 mM Na_2EDTA at 60 W constant power for 1-2 h. Gels were dried on Whatman 3MM paper and radioactive bands were autoradiographed to an imaging plate ("Fujifilm") followed by scanning with a FLA-5100 phosphoimager. DNA bands were quantitated using MultiGauge software ("Fujifilm"). Purine (G+A)-specific tracks were generated by traditional Maxam-Gilbert chemical sequencing reactions (Sambrook and Russell, 2001).

Modification reactions

Reactions with exogenous compound containing 20 nM DNA and 1-20 times excess of MTase (in TLC and denaturing PAA gels experiments), 13 μM DNA and M.HhaI (for HPLC experiments) or 820 nM concentration of target sites (0.12 μg/μl λ phage DNA) and 3.3 μM M.HhaI (in λ phage DNA labeling experiment) and exogenous reagent (typically, 13 mM formaldehyde (FA), 800 mM acetaldehyde (AA), 200 mM propionaldehyde (PA), 200 mM 2-chloroacetaldehyde (Caa), 12-50 mM L-cysteine, cysteamine or DTT, 1.2-12 mM L-selenocysteine or selenocysteamine, 150-500 mM 2-

mercaptoethanol or 0.5 mM adenosyl-5'-thiol) were incubated for 1-2 hours at room temperature. For modification reversal reactions, aldehyde-modified DNA was incubated with MTase for 2 hours at 37°C in TEN buffer.

Thin layer chromatography (TLC) analysis

Precipitated DNA was incubated for 1-3 hours with BAL31 (0.1-0.05 u/µl; 30°C) or P1 nuclease (0.02-0.01 u/µl; 55°C). Aliquots of hydrolyzate were spotted on TLC plates (PEI Cellulose F, 20 x 20 cm, "Merck") and analyzed by eluting with isobutyric acid/ water/ conc. ammonia, (66:17:4,vol/vol/vol) and, if required in second dimension, with saturated ammonium sulfate/ sodium acetate (1 M)/ isopropanol (40:9:1, vol/vol/vol). Plates were dried, autoradiographed to phosphorimager screens and analyzed with a FLA-5100 scanner and MultiGauge software ("Fujifilm).

High-performance liquid chromatography – mass spectrometry (HPLC-MS) analysis

DNA was digested to nucleosides by incubation with P1 nuclease (0.01 u/µl) for 2 hours at 55°C and calf intestine alkaline phosphatase CIAP or FastAP (0.01 u/µl) overnight at 37°C. Hydrolyzate was analyzed on an integrated HPLC/ESI-MS Agilent 1100 series system equipped with a Discovery C18 column (75 x 2.1 mm, "Supelco") by elution with a linear gradient of solvents A (20 mM ammonium formate pH 3.5 or ammonium acetate pH 5.5) and B (80% aqueous methanol) at a flow of 0.3 ml/min at 30°C as follows: 0–20 min, 0–20% B; 20–22 min, 20–100% B, 22-27 min, 100% B; 27-29 min, 0-100% A. UV spectra (190–400 nm) were acquired by an in-line diode array UV absorbance detector at peak maxima. High-resolution mass spectra of modification products were acquired on an LTQ Orbitrap mass spectrometer ("Thermo Electron") equipped with a Proxeon NanoSpray ESI source or Agilent Q-TOF 6520 mass spectrometer ("Agilent Technologies") equipped with Dual-ESI.

Modification of plasmid DNA

Modification reactions containing 0.2 mg/ml pUC19 DNA (2 µM of M.HhaI target sites) and 8 µM M.HhaI were incubated with 13 mM formaldehyde for 1 hour at room temperature in 30 µl of TEN buffer. Then, freshly diluted L-cysteine was added to the final concentration of 50 mM and reactions were left for 1 more hour. Afterwards samples were diluted 6-fold with water and extracted once with Roti®-Phenol, two times with Roti®-Phenol/C/I and three times with chloroform. Aqueous phase was separated and DNA was precipitated with isopropanol (0.9 volume). The amino-modified pUC19 DNA was treated with 250 µM fluorescein-NHS (6-[Fluorescein-5(6)-carboxamido]hexanoic acid N-hydroxysuccinimide (NHS) ester, "Fluka") in 0.15 M sodium bicarbonate (pH 9.0) at room temperature for 1.5 hour and purified by using Nucleotide Removal kit ("Qiagen"). Labeled DNA was fragmented with R.FspBI and analyzed by 2% agarose gel electrophoresis. Gels were scanned with a Fuji FLA-5100 imaging system using a 473 nm laser and then stained with ethidium bromide.

Biotin-labeling and enrichment of hmC-containing DNA fragments

The experiment was performed by dr. Edita Kriukienė. A mixture of the hmC-containing (specific) and non-specific fragment (50 ng each) was incubated with 12 mM cysteamine in the presence of 4 µM M.SssI in acetate buffer (pH 6.0) at room

temperature for 1 hour. The resulting amino-derivatized DNA was treated with a 1,500 molar excess of biotin disulfide *N*-hydroxysuccinimide (NHS) ester ("Sigma") in 0.15 M sodium bicarbonate (pH 9.0) buffer at room temperature for 2.5 h. DNA samples were purified with Nucleotide Removal kit ("Qiagen") and then incubated with streptavidin-coated magnetic beads Dynabeads M-280 ("Invitrogen") in 10 mM Tris-HCl (pH 8.5), 1 M NaCl buffer on a tumbling wheel at room temperature for 3 hours. Beads were collected with a magnetic rack, washed three times with the incubation buffer (10 mM Tris-HCl (pH 8.5), 1 M NaCl) and finally re-suspended in 10 mM Tris-HCl (pH 8.5). On-beads DNA samples were immediately used for quantitation by multiplex real-time PCR on a Rotor-Gene™ 6000 real-time PCR instrument ("Corbett Research") using Maxima™ Probe qPCR Master Mix ("Fermentas"). 0.9 μM primers for the specific DNA fragment, 0.3 μM primers for the non-specific probe and 0.25 μM of the respective dual-labeled probe ("Metabion") were used in each reaction in a final volume of 25 μL (see table below for primer sequence details). The amplification program was set as: 95 °C for 10 min, 40 cycles 95°C for 15 s, 60°C for 1 min. Data were analyzed by Rotor-Gene™ software and reported as percentage of starting material.

Table 2. Primers used for RT-PCR.

Fragment	Primer	Primer sequence (5'→3')
Specific probe	Specific-dir	ATGTGTTGGAGTGTGCCTGA
	Specific-rev	GTGGCTCTGATTGATGGCTC
	TaqMan probe	FAM-TCCCTGTGTGATCACCCCTATGCTTG-BHQ1
Control probe	Control-dir	CAGGCCTCTTCAAGGGTCA
	Control-rev	AAGAGATGAGGGCCTGGG
	TaqMan probe	JOE-TGGCCCATACCTCTTCAAGGGCA-BHQ1

Synthesis of modified nucleosides

Standard of 5-hydroxymethyl-2'-deoxycytosine (hmC), 5-hydroxymethyl-2'-deoxyuridine (hmU) and 5-(2-hydroxyethylthio)methyl-2'-deoxyuridine (hetmU) was chemically synthesized according to (Tardy-Planechaud et al., 1997; LaFrancois et al., 1998) with modifications. The synthesis was performed by dr. Viktoras Masevičius.

Adenosyl-5'-thiol was synthesized according to (Pignot et al., 2000) by dr. Milda Malvina Burbulienė and dr. Viktoras Masevičius.

5-hydroxymethyl-2'-deoxycytidine (hmC). 4 ml of triethylamine was added to a suspension of 0.5 g (1.90 mmol) of 2'-deoxycytidine hydrochloride in 3 ml of 20% aqueous formaldehyde. Reaction mixture was stirred at r.t. for 30 min, then placed in a steel bomb and heated for 7 days at 65 ºC. The conversion of dC reached ~50%, yielding 5-hydroxymethyl-2'-deoxyuridine as major product. Reaction mixture was cooled; aqueous layer was separated and passed through 4 g of cation exchange resin (Dowex 50Wx8, or similar, H^+-form). Product was eluted with water and a combined aqueous fraction was concentrated under reduced pressure. Column chromatography (silica gel 40–60 μm, chloroform: methanol 1:1) was used to isolate ~26 mg of hmC (9%). ^{1}H NMR spectrum (300 MHz, DMSO-D_6) of the obtained material was fully consistent with reported spectra of 2'-deoxycytidine and 5-hydroxymethyl-2'-deoxycytidine.

5-hydroxymethyl-2'-deoxyuridine (hmU). 4 ml of triethylamine was added to a suspension of 0.5 g (2.19 mmol) 2'-deoxyuridine in 3 ml 20% aqueous formaldehyde. Reaction mixture was stirred at r.t. for 30 min, then placed in a steel bomb and heated for 4 days at 65 ºC. Reaction mixture was cooled to r.t., an aqueous layer was separated and

passed through 4 g of cation exchange resin (Dowex 50Wx8, or similar, H^+-form). Product was eluted with water and a combined aqueous fraction was concentrated under reduced pressure. Residue was purified by column chromatography (silica gel 40–60 μm) by elution with 90% aqueous acetonitrile to yield 0.19 g (34%) of compound hmU. ^{1}H NMR spectrum (300 MHz, DMSO-D_6) of the synthesized compound matches published data (Shiau et al., 1980); ^{13}C NMR (75 MHz, DMSO-D_6): δ = 40.47, 56.70, 62.14, 71.25, 84.59, 87.99, 114.96, 137.46, 150.07, 163.35.

5-(2-hydroxyethylthio)methyl-2'-deoxyuridine (hetmU). A suspension of 0.042 g (0.16 mmol) 5-hydroxymethyl-2'-deoxyuridine (hmU) in 400 μL 2-mercaptoethanol and 0.015 g (0.13 mmol, 10 μL) of trifluoroacetic acid (TFA) was heated in a sealed polypropylene vial at 130ºC for 1 h, then cooled to r.t. and poured into 5 ml of CH_2Cl_2. Suspension was loaded onto a 2 cm layer of silica gel (40–60 μm), washed with 3x15 ml of CH_2Cl_2 to remove most of remaining 2-mercaptoethanol and TFA, and then eluted with 15 ml of methanol. Methanol was removed under reduced pressure and residue was repurified on a silica gel column with 95% aqueous acetonitrile to yield 28 mg (54%) of compound hetmU. ^{1}H NMR (300 MHz, DMSO-D_6): δ = 2.11 (dd, $J1_{C2'\text{-}C1'}$ = 6.8Hz, $J2_{C2'\text{-}C3'}$ = 4.6Hz, 2H, C2'), 2.54 (t, $J_{SCH2\ -\ CH2O}$ = 6,8Hz, 2H, S**CH_2**CH_2), 3.38 (s, 2H, C5CH_2S), 3.54 (t, $J_{OCH2\ -\ CH2S}$ = 6,8Hz 2H, CH_2**CH_2**O), 3.58 (t, $J_{C5'\text{-}C4'}$ = 3.9Hz 2H, C5'), 3.80 (td, $J1_{C4'\text{-}C5'}$ = 3.9Hz, $J2_{C4'\text{-}C3'}$ = 2.8Hz, 1H, C4'), 4.26 (td, $J1_{C3'\text{-}C2'}$ = 4.6Hz, $J2_{C3'\text{-}C4'}$ = 2.8Hz, 1H, C3'), 4,83 (brs, 1H, OH), 5.01 (brs, 1H, OH), 5.30 (brs, 1H, OH), 6.18 (t, JC1'-C2' = 6.8Hz, 1H, C1'), 7.85 (s, 2H, C5), 10.65 (brs, 1H, NH). ^{13}C NMR (75 MHz, DMSO-D_6): δ = 27.95, 34.49, 40.38, 61.33, 62.08, 71.17, 84.75, 88.08, 111.76, 137.95, 150.96, 163.25. HR-MS: MS^1 ($[M+Na]^+$, $C_{12}H_{18}N_2O_6SNa$), calculated: 341,0778, found: 341,0779; MS^2 ([MS^1 minus deoxyribose], $C_7H_{10}N_2O_3SNa$) calculated: 225.0304, found: 225.0298; MS^3 ([MS^2 minus SCH_2CH_2OH], $C_5H_4N_2O_2Na$), calculated: 147.0165, found: 147.0158

Deamination of enzymatically produced 5-(2-hydroxyethylthio)methyl-2'-deocytidine (hetmC) to 5-(2-hydroxyethylthio)methyl-2'-deoxyuridine (hetmU). hmtC obtained in a 1.8 ml enzymatic reaction as described above was isolated by HPLC using an analytical RP column as described. The isolated fraction was evaporated to dryness and treated with 3 M sodium nitrite in 0.3 M sodium acetate pH 4.5 for 1 hour at 4 ºC. Aqueous ammonia was added to pH 8.0 and the reaction was incubated for 4 hours at 37 ºC.

RESULTS AND DISCUSSION

1 Caa reacts with unpaired cytosine

Although X-ray crystallography of protein-DNA complexes can provide the ultimate proof of base flipping, co-crystallization of proteins with their DNA substrates is often tedious or even impossible. Alternative methods to detect base flipping in aqueous solution are essential for extending studies of this phenomenon. One such method is fluorescence spectroscopy; however it requires modified bases to be introduced in the DNA. 2-aminopurine, a close structural analogue of adenine, is one of the most widely used fluorescent probes (Holz et al., 1998; Neely et al., 2005). However fluorescent probes gained limited popularity for studies of enzyme-induced base-flipping due to substantial alterations of the base-pairing potential and/or steric bulk as compared with the natural cytosine base. On the other hand, fluorescence spectroscopy requires specialized equipment and may not always be sufficiently sensitive to serve as a routine laboratory tool. Previously it was found that $KMnO_4$ is an efficient probe for detection of flipped out thymines in DNA-methyltransferase (Serva et al., 1998) and DNA-transposase complexes (Bischerour and Chalmers, 2007). However none of the above described approaches can directly detect flipping of cytosine, which is a natural target base for numerous DNA methyltransferases. Therefore, a cytosine-specific probe would clearly be a useful method for initial analysis of novel base-flipping systems.

Chloro- and bromoacetaldehyde are known to react with unpaired adenine and cytosine bases in DNA (Kusmierek and Singer, 1982; Kohwi-Shigematsu and Kohwi, 1992) yielding N1,N^6-ethenoadenine or N^4,N3-ethenocytosine derivatives, respectively (Fig. 1). Such haloacetaldehyde-modified residues can be detected by piperidine-induced strand cleavage (Kohwi-Shigematsu and Kohwi, 1992).

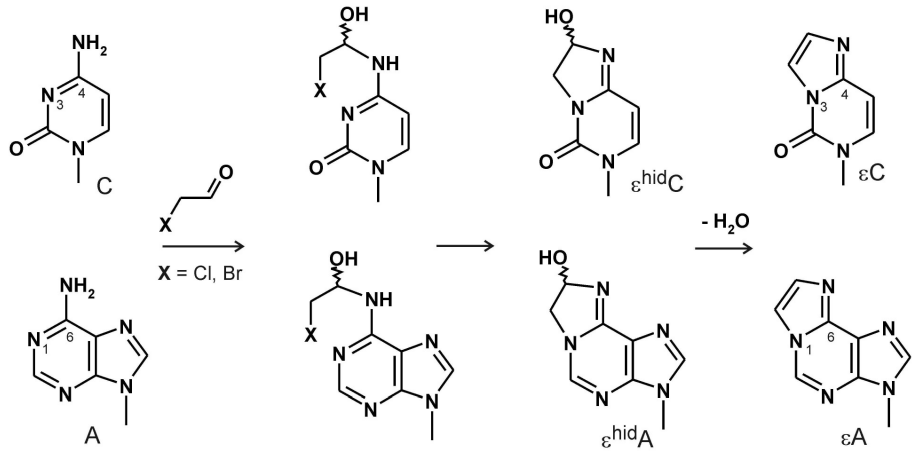

Figure 1. 2-haloacetaldehydes (2-chloroacetaldehyde and 2-bromoacetaldehyde) react with unpaired cytosine and adenine nucleobases in DNA. In the first step the aldehyde group of haloacetaldehyde reacts with exo amino group and in the second step the halomethyl group reacts with endo nitrogen atom. Reaction is proceeding via long-lasting intermediate stage (ε^{hid}C and ε^{hid}A) producing N^4,N3-ethenocytosine (εC) and N1,N^6-ethenoadenine (εA).

1.1 Caa reactivity of cytosines in complexes with DNA cytosine-5 methyltransferases

A 25-mer duplex DNA GCGC/GMGC-25 containing a hemimethylated target sequence for M.HhaI and 5'-labeled on the upper strand was used for screening. M.HhaI-DNA complexes were formed by adding a slight access of M.HhaI to ensure complete binding, and were then treated with Caa for 15–105 min at r.t. or 37°C. DNA strand cleavage at Caa-modified sites was achieved by heating with 1 M piperidine and followed by denaturing gel electrophoresis. In parallel, the integrity of the M.HhaI-DNA complexes was verified by analyzing reaction aliquots using gel mobility shift assay under native conditions.

Addition of M.HhaI leads to a selective and significant enhancement of the reactivity of the target base, whereas a control reaction containing no enzyme shows background levels of reactivity (Fig. 2a). Importantly, addition of the Q237G mutant, which was indirectly shown to have abolished base flipping activity but retained binding affinity (Daujotytė et al., 2004), shows no enhanced reactivity with CAA, serving as a non-flipped control reaction (Fig. 2a).

Notably, there was no detectable Caa modification of the flipped out target cytosines in the ternary M.HhaI complex with DNA and cofactor product SAH (SAM was not used as it causes enzymatic turnovers) (Fig. 2a).

In addition, we analyzed the Caa reactivity of a mismatched target cytosine paired with C, A or T on the complementary strand of the DNA duplex. The weaker non-canonical base pair may exist in unpaired, partially flipped out conformations and is thus likely to become accessible to chemical modification. However, no Caa modification of the mismatched target cytosine was detected under the assay conditions in the absence of enzyme. This observation indicates that a cytosine residue is modified with Caa only when stabilized in an extrahelical conformation by the enzyme (Fig. 2a). On the other hand, cytosines in single-stranded DNA (ssDNA; absence of complementary strand) are readily detectable under these conditions (Fig. 2a).

Having successfully demonstrated chemical mapping of extrahelical cytosines in a well proven system, we went on to examine the generality of this approach by analyzing a series of DNA cytosine-5 MTases: M.SssI, M.AluI, M.HaeIII, M.HpaII, M.MspI, M.Dmt1, M.Eco31IC and two 4mC MTases: M.MvaI and M.BfiC2. Reaction conditions were again optimized to achieve maximal cytosine modification with Caa in a control ssDNA, while maintaining the integrity of the MTase-DNA complexes (examined in a gel mobility shift assay, not shown).

Analysis with 50-300 mM Caa showed an enhanced reactivity of their respective target cytosine residues in the presence of saturating amounts of M.AluI and M.SssI. (Fig. 2b-d). Best results were achieved with 50 mM Caa at 37°C for M.AluI and 100 mM Caa at 20°C for M.SssI. As observed for M.HhaI, addition of SAH abolished the Caa reaction in both cases. Interestingly, M.SssI showed a marked preference for a centrally located target site among a total of five CG sites present in the DNA duplex. The nature of this preferential binding has not been investigated further.

Similar experiments with other MTases using 50-500 mM Caa gave no detectable modification at the target cytosine (not shown). All the enzyme preparations contained no detectable amount of bound endogenous SAM.

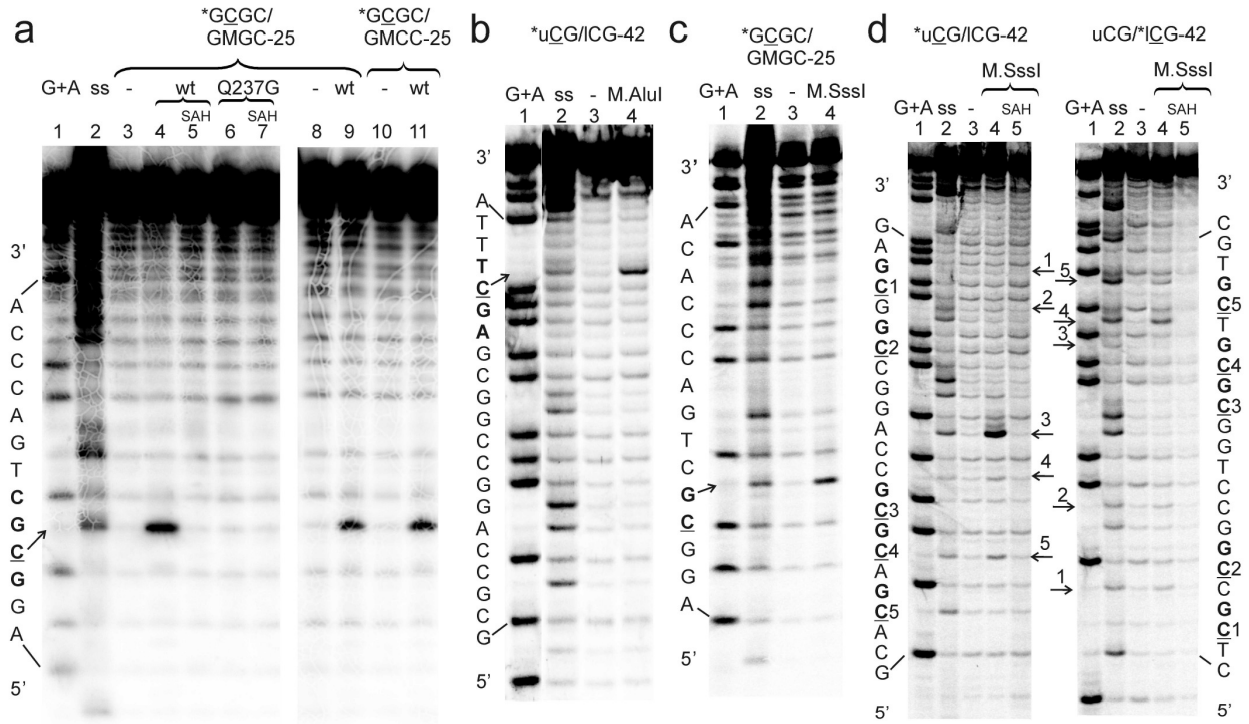

Figure 2. Caa reactivity of cytosines in DNA duplexes upon interaction with methyltransferases. Autoradiograph of degradation products of DNA duplexes 5'-labeled on the target strand after separation by electrophoresis on a 15% sequencing gel. Lane 1, G+A nucleotide cleavage marker; lanes 2–11, reactions with Caa; line 2, single stranded DNA (ss); "-" denote double stranded DNA without MTase. **a** Caa reactivity of cytosines in cognate GCGC/GMGC-25 (lines 3-9) or mismatch GCGC/GMCC-25 (lines 10-11) DNA duplexes upon interaction with M.HhaI wt (lines 4-5, 9, 11) or Q237G (lines 6-7). Lines 2-11, reactions with 200 mM Caa for 45 min at 20°C (line 2 – 37°C) containing 300 nM M.HhaI and 280 nM DNA (lines 3-7) or 30 nM M.HhaI and 10 nM DNA (lines 8-11); lines 5 and 7 also contained 0.1 mM SAH. **b** CAA reactivity of cytosines in the 42-mer oligonucleotide uCG/lCG-42 upon interaction with M.AluI. Lanes 2-4, reactions with 50 mM Caa for 50 min at 37°C, containing 10 u/µl M.AluI and 105 nM DNA. **c-d** Caa reactivity of cytosines in the duplexes GCGC/GMGC-25 containing one hemimethylated target site (c) or CG/lCG-42 containing 5 unmethylated target sites (d) upon interaction with M.SssI. Lines 2-5, reactions with 100 mM Caa for 55 min at 20° (line 2 – 37°C), containing 4 u/µl (c) or 8 u/µl (d) M.SssI, 20 nM DNA and 0.1 mM SAH (line 5). In **d** upper (*uCG-42; left panel) or lower (*lCG-42; right panel) DNA strands were 5'-labeled.

Table 3. Reaction conditions for 2-chloroacetaldehyde probing of extrahelical cytosine residues in MTase-DNA complexes

MTase	**Target site** (modified cytosine is underline)	**MTase-DNA complex** (verified by gel mobility shift assay)	**Caa conc., M**	**T,°C**	**Time, min**
M.HhaI	G<u>C</u>GC	without Caa – all DNA bounded, with Caa – >1/2	0.01-0.2	20, 37	15-105
M.AluI	AG<u>C</u>T	without Caa >1/2 of DNA bounded, with Caa – small amount of DNA bounded	0.05-0.2	20, 37	40-60
M.SssI	<u>C</u>G	with or without Caa – small amount of DNA bounded	0.1-0.3	20	60

1.2 Caa reactivity of cytosines in complexes with restriction endonucleases

To further assess the generality of the new assay, we extended our studies to a recently discovered class of base-flipping enzymes – restriction endonucleases (REases). Three REases were tested.

Homologous restriction endonucleases R.Ecl18kI and R.PspGI recognize DNA sequences CCNGG and CCWGG (N – any base; W – A or T), respectively, and cleaves them before the outer C. The enzymes has recently been proven using X-ray crystallography and other methods (Bochtler et al., 2006; Tamulaitis et al., 2007; Szczepanowski et al., 2008) to flip out both nucleotides (N or W) of the central base-pair upon binding the substrate. Therefore we examined the susceptibility of the central cytosine in the sequence CCCGG to the Caa reagent. We used a non-cognate DNA substrate for R.PspGI, since it has recently been shown that R.PspGI can interact with this DNA substrate but does not cleave it (Tanulaitis et al., 2008). In contrast, the R.BcnI, which recognizes and cuts the CCSGG (S – C or G) target site, exerts no base flipping upon binding its substrate DNA (Sokolowska et al., 2007). This system was therefore expected to provide a negative control.

Caa reaction conditions were optimized to achieve maximal cytosine modification, at the same time maintaining the R.Ecl18kI-DNA, R.PspGI-DNA or R.BcnI-DNA complexes during the modification reaction. Fragmentation at the "target" cytosine sites in R.Ecl18kI-DNA and R.PspGI-DNA complexes were observed, indicating an extrahelical conformation of those cytosines in solution (Fig. 3). Notably, the modification was observed only in a relatively high Caa concentration range of 0.2–1.2 M. However gel binding experiments confirmed that the substrate DNA was bound with high affinity under these conditions (not shown). In both cases, the strand cleavage was Caa-dependent and occurred at positions different from those observed during nucleolytic cleavage in the presence of Mg^{2+} ions (not shown).

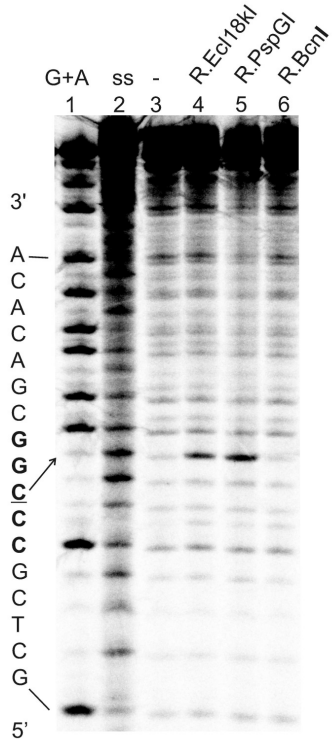

Figure 3. Caa reactivity of cytosines in the 31-mer CCCGG/CCGGG-31 duplex (100 nM) upon interaction with restriction endonucleases R.Ecl18kI, R.PspGI and R.BcnI (1000 nM). Complexes were treated with 400 mM Caa for 40 min at 37ºC (ssDNA was treated with 200 mM Caa for 10 min at 37ºC, lane 2), followed by piperidine treatment and strand separation by electrophoresis on a 15% sequencing gel. Lane 1, purine nucleotide cleavage marker; lane 2, upper CCCGG-31 strand alone; lane 3, CCCGG/CCGGG-31; lane 4, CCCGG/CCGGG-31 + R.Ecl18kI; lane 5, CCCGG/CCGGG-31 + R.PspGI; lane 6, CCCGG/CCGGG-31 + R.BcnI.

The control reactions involving R.BcnI showed no Caa-dependent modification and strand cleavage at the expected cytosine residue under similar conditions (Fig. 3, line 6). Since R.BcnI interacts with all its target base pairs without much disturbing the double helical DNA structure (Sokolowska et al., 2007), our observation confirms that the Caa modification reaction is specific for unpaired extrahelical cytosines.

2 Atypical reactions of DNA cytosine-5 methyltransferases

Surprisingly, our analyses involving M.HhaI mutants containing individual replacements of essential residues in the catalytic centre (C81S, E119A, R165A) (Mi and Roberts, 1993; Shieh et al., 2006; Shieh and Reich, 2007) or a complete deletion of the catalytic loop (Daujotytė et al., 2004) all showed no reaction with Caa (Fig. 4a). It was shown that the integrity of the MTase-DNA complexes was maintained during the Caa treatment (Fig. 4b). These observations clearly indicate that the covalent activation of the target cytosine in the catalytic centre of the DNA cytosine-5 MTase is somehow important for its reactivity with Caa. We decided to analyze the reaction products of cytosine and Caa reaction in ssDNA, DNA-MTase and DNR-REase complexes using 2D-TLC and HPLC-MS.

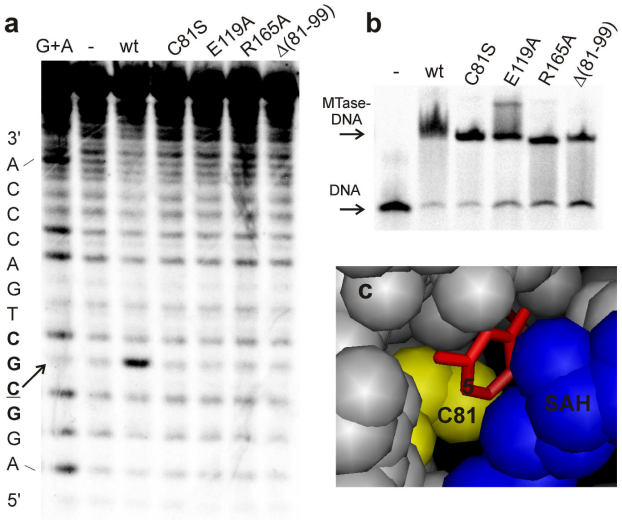

Figure 4. a Autoradiograph of degradation products of a 25-mer duplex 5'-[^{33}P]-labeled on the target (upper) strand after separation by electrophoresis on a 15% sequencing gel. Lane 1, G+A nucleotide cleavage marker; lanes 2-7, reactions with 200 mM Caa for 45 min at 20°C, containing 500 nM protein and 370 nM DNA *GCGC/GMGC-25. **b** DNA binding efficiency of M.HhaI wt and mutants upon incubation with Caa for 10 min (conditions like in a). **c** A space-fill model of the ternary M.HhaI-DNA-SAH complex (PDB code (1mht). Activated cytosine intermediate (ACI) is shown in red, the cofactor SAH – in blue, the covalently bound catalytic Cys81 – in yellow and other residues of M.HhaI – in grey. Cytosine-5 position is blocked for exogenous aldehydes in the presence of SAH.

2.1 Analysis of Caa and cytosine reaction products by 2D-TLC

First, we compared the products formed upon reaction of Caa with unpaired cytosines in ssDNA with those produced in the MTase-DNA complexes. Selective analysis of modification at the target site was achieved with oligonucleotide duplexes that were internally ^{33}P-labeled at one or two cytosine nucleotides in the target site (see Table 1). Following a Caa modification reaction, the DNA was digested to 5'-mononucleotides and analyzed by two-dimensional TLC (2D-TLC) (Kuchino et al., 1987). The analyses showed completely different cytosine modification products in free single strand DNA vs duplex DNA bound to M.HhaI, M.SssI or M.AluI (Fig. 5). In the former case, Caa treatment leads to three chromatographically distinct nucleotides (Fig. 5b), which converge to a single modified product upon incubation at 85°C for 2 hours (Fig. 5c) which is further degraded upon incubation with piperidine (Fig. 5d). This pattern is consistent with the known modification mechanism whereby cytosine is first

converted into N^4,N3-(1-hydroxyethan)cytosine ($\varepsilon^{hid}C$) which then slowly dehydrate to give N^4,N3-ethenocytosine (εC) (Fig. 1) (after the first dimention the TLC plate is dried overnight, we propose that during this period of time some amount of $\varepsilon^{hid}C$ is dehydrated to εC and in the second dimension is moving like εC; accordingly two spots of $\varepsilon^{hid}C$ are seen in Fig. 5b) (Krzyzosiak et al., 1981; Kusmierek and Singer, 1982). In contrast, Caa treatment of the MTase-DNA complex gave rise to a new major product, which was partially separated from the original cytosine nucleotide in the TLC system (Fig. 5f). The new nucleotide undergoes further transformation upon heating and piperidine treatment (Fig. 5g,h). Clearly, the modification of the target cytosine in the active site of the DNA cytosine-5 MTases occurs via a different pathway yielding new products. It should be pointed that treatment with Caa of R.Ecl18kI-DNA and R.PspGI-DNA complexes showed the same Caa modified cytosines as in ssDNA (Fig. 5b,j,k). The intensity of εC spot in DNA-R.Ecl18kI and DNA-R.PspGI complexes increased about 3 times compared to controls (free DNA and DNA-R.BcnI complex).

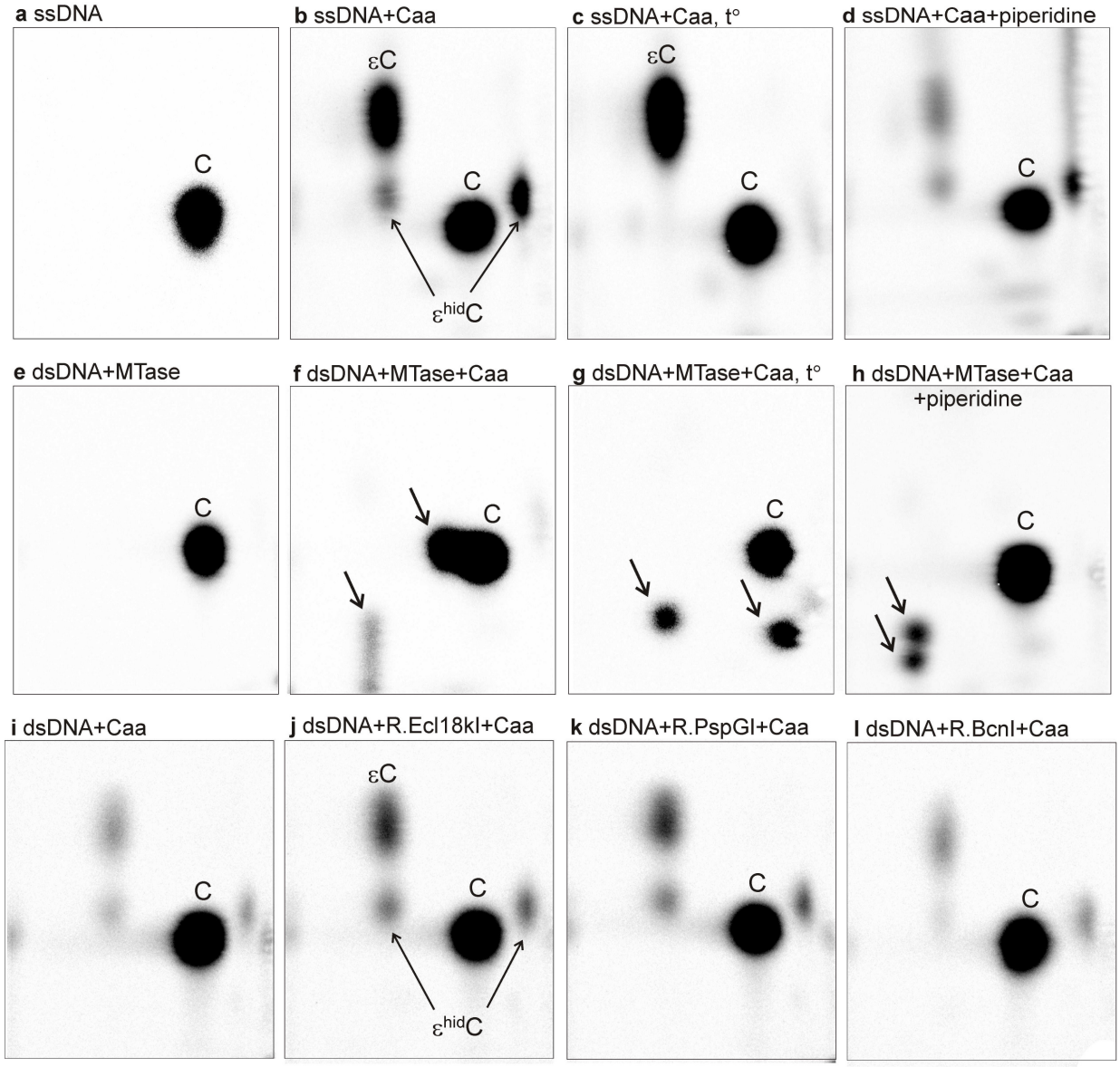

Figure 5. TLC analysis of products formed upon Caa modification of cytosine residues in free ssDNA (**a-d**) and in duplex DNA complexes with M.SssI (**e-h**) or REases (**i-l**). Reactions containing 20 nM *CCN-19 (a-d) or 20 nM CC*CGG/CCGGG-31 DNA and 0.4 u/µl M.SssI (e-h) or 100 nM

CC*CGG/GGGCC-31 DNA and 1 μM REases dimmer were incubated with 400 mM Caa for 1 h at 37°C (b-d, i-l) or with 100 mM Caa for 50 min at 20°C (f-h). Control sample **a** and **d** lacked Caa; samples **c** and **f** were additionally incubated for 2 hour at 85°C and samples **d** and **h** were incubated with 1 M piperidine for 30 min at 90°C. Modified DNA was then enzymatically digested to mononucleotides, analyzed using 2D-TLC on PEI-cellulose and autoradiographed to reveal [^{33}P]-labeled nucleotides and its modification products. Spots of εC and ε^{hid}C were ascribed according (Krzyzosiak et al., 1981).

2.2 Methyltransferase-directed aldehyde coupling to target cytosines producing 5-(1-hydroxyalkyl)cytosines

To confirm the assumption that aldehyde rather than chloromethyl is the reactive group in Caa, a formaldehyde (FA) and chloroacetamide were compared for their ability to modify the target cytosine in the binary M.HhaI-DNA complex (Fig. 6). The modification reactions were monitored using both systems: the internally labeled DNA substrate followed by one-dimensional TLC analysis of ^{33}P-labeled nucleotides and HPLC-MS analysis of unlabeled deoxynucleosides. Indeed, the TLC analysis of the DNA-M.HhaI complex treated with 13 mM FA showed that nearly 50% of radioactivity from the two ^{33}P-labeled C nucleotides on the upper strand was shifted indicating a nearly complete modification of the target cytosine (Fig. 7a). No new reaction product was observed when chloroacetamide was used (not shown). Either no new product were noted when FA, catalytically active M.HhaI, or the cognate DNA substrate for M.HhaI was missing, or when the target cytosine resides was enzymatically premethylated at the 5-position. HPLC-MS analysis of the FA modified DNA also revealed the formation of a novel nucleoside that eluted right following dC (Fig. 7f). The UV spectrum and the molecular mass of the new compound suggested an addition of Fa to the cytosine ring. Additional evidence for cytosine-5 coupling was obtained from analyses of a DNA duplex containing 5-dueterated cytosine at the target position, which was prepared by incubating the DNA substrate in $^{2}H_2O$ in the presence of M.HhaI (Wu and Santi, 1987; Daujotytė et al., 2004). Indeed, the 5-[^{2}H] label is completely lost upon M.HhaI-directed coupling of FA (and acetaldehyde, see below) in $^{2}H_2O$, which indirectly implicates carbon-5 as the target (Fig. 8c). Finally, direct chromatographic comparison of the modified nucleoside with authentic 5-hydroxymethyl-2'-deoxycytidine in a spike-experiment confirmed their identity (Fig. 8a). Therefore, we can conclude with confidence that the target cytosine in the specific M.HhaI-DNA complex is converted to 5-hydroxymethyl-cytosine (hmC) in the presence of FA.

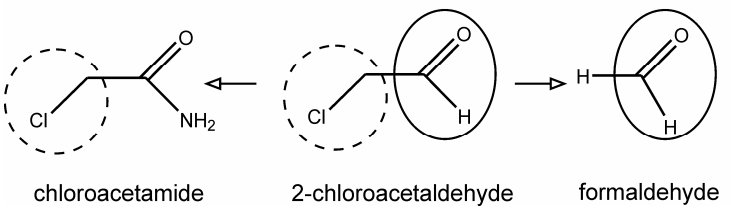

Fig 6. 2-chloroacetaldehyde (center) and its mono-functional mimics chloroacetamide (left) and formaldehyde (right). Chloromethyl and aldehyde groups are shown by dotted and full circles, respectively.

Similar experiments with other aldehydes such as acetaldehyde (ethanal, AA) and propionaldehyde (propanal, PA) also showed MTase-dependent formation of cytosine-5 adducts (Fig. 7). Typically, two coupling products of the same molecular mass were observed in HPLC traces consistent with the formation of an asymmetric centre at the α-carbon in the attached side chains. Our results thus indicate that the addition reaction is not limited to a unique example but is general for aliphatic aldehydes as a class, although compounds with bulkier chains showed lower coupling yields (Table 4). TLC analyses

also showed minor amounts of modification products upon treatment with benzyloxyacetaldehyde and glycine betaine aldehyde (Fig. 7a,b), however, corresponding compounds were not detectable with our HPLC-MS system. Other electrophiles such as aliphatic ketones, conjugated aldehydes and alkenes were inactive in our hands.

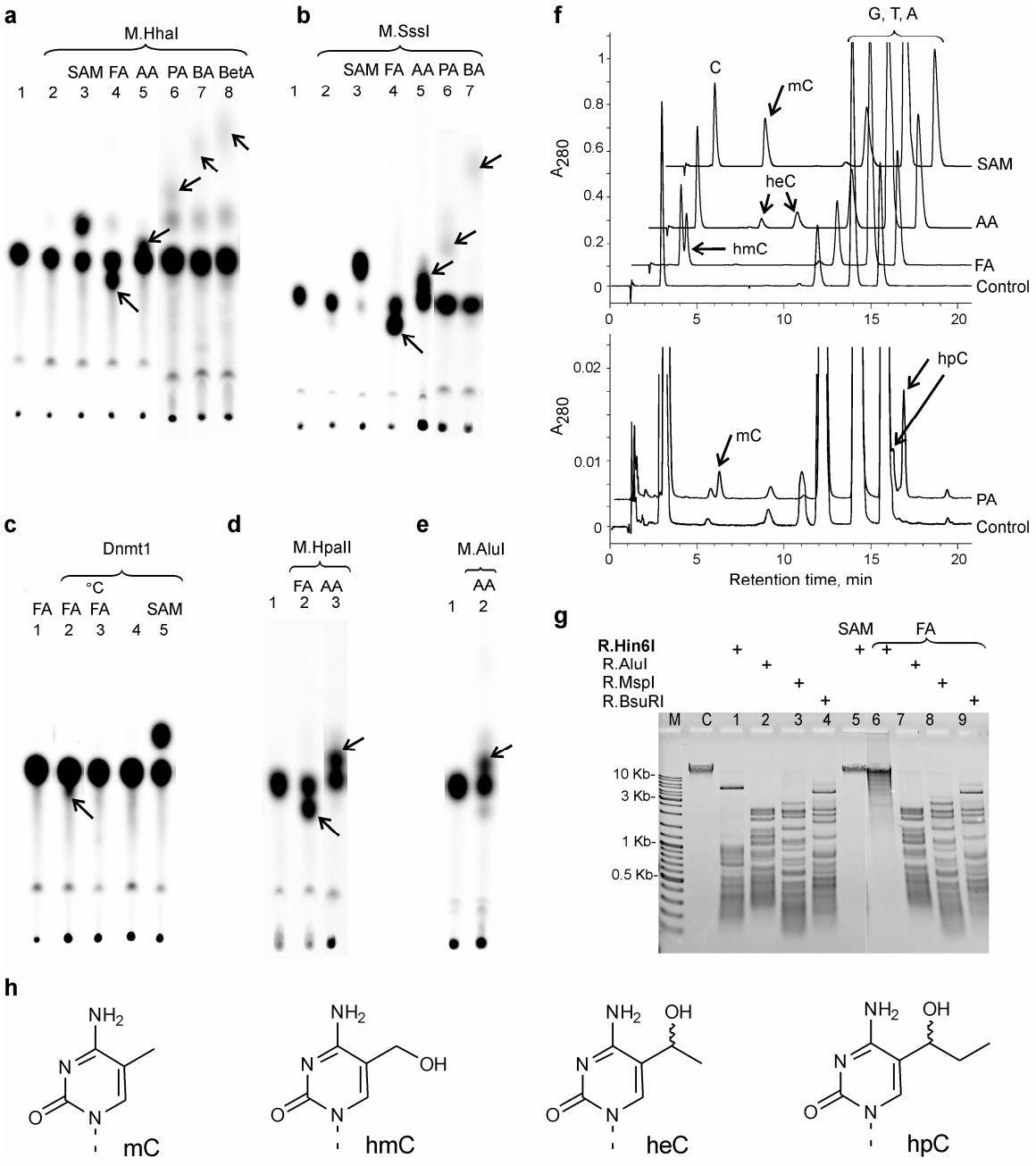

Figure 7. Covalent modifications of cytosines with aliphatic aldehydes in the presence of DNA cytosine-5 MTase. Reactions, containing internally [^{33}P]-labeled 20 nM DNA duplexes G*CG*C/GCGC-32 and 125 nM M.HhaI (**a**), 20 nM CC*CGG/CCGGG-31 and 0.8 u/µl M.SssI (**b**), 20 nM G*CG*C/GMGC-32 and 6.6 µM Dnmt1-ΔN (**c**; in lane 3, the sample was incubated at 75°C for 15 min (heat inactivated MTase) before FA was added), 40 nM CC*CGG/CCGGG-31 and 600 nM M.HpaII (**d**), or 105 nM AG*CT/AGCT-27 and 10 u/µl M.AluI (**e**) were treated with 200 µM SAM, 13 mM formaldehyde (FA), 800 mM acetaldehyde (AA), 200 mM propanal (PA), 100 mM benzyloxyacetaldehyde (BA) or 100 mM glycine betaine aldehyde (BetA) for 1 hour at r.t. Modified DNA samples were digested to 5'-dNMPs and analyzed by TLC/autoradiography. **f** GCGC/GCGC-24 duplex (13 µM) was incubated, in the presence of 13 µM M.HhaI, with aldehydes at conditions described above. Control reactions contained no exogenous

reagent. Modified DNA was enzymatically fragmented to nucleosides and analyzed by reversed-phase HPLC. Arrows point at peaks corresponding to new modification products. **g** Modification reactions containing bacteriophage λ DNA (0.82 μM HhaI target sites), M.HhaI (3.3 μM) and cofactor as indicated (200 μM SAM or 13 mM FA) were incubated for 1 hour at room temperature. The modified DNA was then fragmented with an indicated REases and analyzed by agarose gel electrophoresis. Lane M, DNA length marker (GeneRuler^{TM} DNA Ladder); lane C, bacteriophage λ DNA control. **h** Structure of obtained modified cytosines.

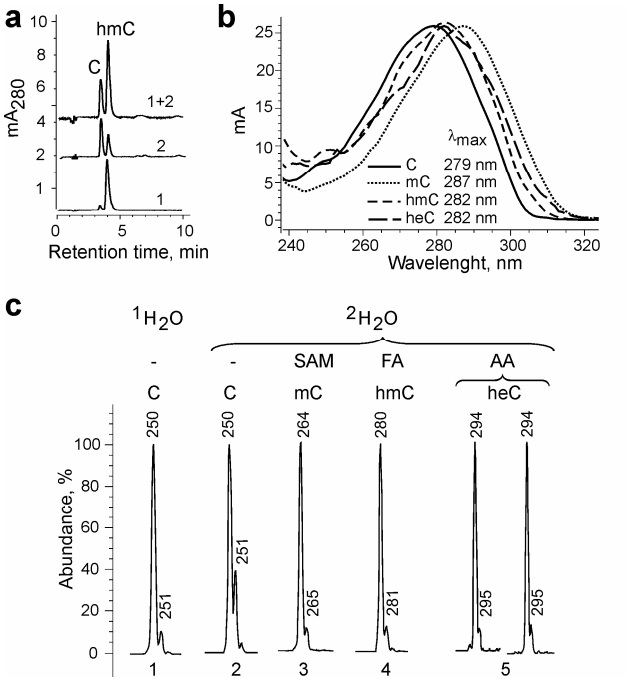

Figure 8. MTase-directed formation of 5-(1-hydroxyalkyl)cytosines in DNA. **a** RP-HPLC chromatogram of chemically synthesized 5-hydroxymethyl-2'-deoxycytidine (1), DNA nucleosides obtained upon incubation of 13 μM GCGC/GCGC-24 duplex with 13 μM M.HhaI and 13 mM FA (2) and the two samples combined together (1+2). **b** UV absorption spectra of MTase-directed coupling products at pH 3.5. **c** MS spectra of modified nucleosides, obtained in reactions performed in 60% ^2H_2O. Lane1, unmodified control in 0% ^2H_2O; line 2, unmodified control in 60% ^2H_2O; lines 3-5, modified nucleosides obtained in the presence of SAM or aldehydes in 60% ^2H_2O

Table 4. Methyltransferase-directed modification of target cytosine residues in DNA with exogenous reagents.

Exogenous reagent/ reaction product	Estimated extent of target cytosine modification observed by TLC (HPLC), %				HRMS analysis of M.HhaI-directed reaction product	
	HhaI	SssI	AluI	HpaII	Fomular (M^+)	Found/Calculated
Formaldehyde (FA)/ 5-hydroxymehyl-dC (hmC)	60 (90)	65	-	50	$C_{10}H_{16}N_3O_5$	258.1085/ 258.1084
Acetaldehyde (AA)/ 5-(1-hydroksyethyl)-dC (heC)	40 (60)	60	30	40	$C_{11}H_{18}N_3O_5$	272.1243/ 272.1241
Propionaldehyde (PA)/ 5-(1-hydroxypropyl)-dC (hpC)	6 (8)	4	3	0	$C_{12}H_{20}N_3O_5$	286.1398/ 286.1398
Chloroacetaldehyde (Caa)/ 5-(1-hydroxy-2-chloroethyl)-dC (hcC)	8 (10)	30 (8)	10	0	$C_{11}H_{17}N_3O_5Cl$	306.0852(32)*/ 306.0851(33)
Benzyloxyacetaldehyde	2 (0)	5	0	0	$C_{18}H_{24}N_3O_6$	378.166/ not found
Glycine betaine aldehyde	2 (0)	-	-	-	$C_{14}H_{25}N_4O_5$	329.1819/ not found

‚-' – experiment was not performed;
‚*' – intensity (%) relative to the most abundant peak is shown in parenthesis.

Having demonstrated the reactivity of the target cytosine in a model MTase, we tested the generality of this reaction with other DNA cytosine-5 MTases: M.AluI,

M.HpaII and M.SssI. Remarkably, all of the interrogated MTases showed substantial catalytic activity towards the aldehydes as determined by TLC of labeled target nucleotides (Fig. 7b,d,e). Examination of the mouse Dnmt1 methyltransferase also showed detectable formation of the modified product in the presence of FA (Fig. 7c).

Figure 7g demonstrates that the FA coupling by M.HhaI is directed exclusively to GCGC sites in lambda DNA, since the modified DNA becomes resistant to cleavage with a GCGC-specific restriction endonuclease (R.Hin6I), but is readily fragmented by restriction nucleases targeting other sites.

2.3 HPLC-MS analysis of products formed upon Caa treatment of MTase-DNA complexes

After an incubation of MTase-DNA complex with Caa in 2D-TLC experiment we saw at least two Caa-cytosine products which mutate to other adducts after heating and piperidine treatment (Fig. 5e-h). To identify these products we decided to analyse them by HPLC-MS. 24 bp DNA oligonucleotide duplex containing one unmethylated target site was incubated with M.HhaI and Caa. After the reaction, DNA was enzymatically degradated to nucleosides and analysed on RF column equipted with mass spectrometer. Similarly like with other aldehydes (Fig. 7f), in HPLC chromatograms we could identify Caa-cytosine product: 5-(1-hydroxy-2-chloroethyl)-2‘-deoxycytoine (hcC). After heating the reaction for 2 hour at 85°C, the peak of hcC disappeared and we could identify one additional cytosine adduct. MS, fluorescence and UV data suggest it is C5,N^4-etheno-2‘-deoxycytosine (Z) (Fig. 9a-d; Table 5) (Woo et all, 1996). To identify other compounds which were seen in 2D-TLC experiments, we purified hcC on RF column and then heated it for 2 hours at 85°C. After heating, next to Z, two potential cytosine adducts 5-(1,2-dihydroxyethyl)-2‘-deoxycytosine (X) and C5,N^4-(1-hydroxyethyl)-2‘-deoxycytosine (Y_1) or 5-(1-oxoethyl)-2‘-deoxycytosine (Y_2) (Y_1 and Y_2 have the same molecular mass) were detected by UV, HRMS and MS/MS (Fig. 9e, Table 5). We expect one or some of Caa-cytosine adducts react with piperidine to produce DNA strand cleavage. Z could be the most plausible variant because of the big double aromatic ring which makes N-glycosylic bond less stable (like in purines compared to pyrimidines). At higher temperature Z base could be released from DNA producing abasic site which originate to DNA strand cleavage upon piperidine treatment (Fig. 10).

Table 5. Cytosine adducts obtained after M.HhaI-DNR complex treatment with Caa.

Product	HRMS		
	Fomular	Calculated	Found
5-(1-hydroxy-2-chloroethyl)-dC (hcC)	$C_{11}H_{17}ClN_3O_5+H^+$	306.0851/ 308.0826 (33)*	306.0852/ 308.0820 (32)*
X 5-(1,2-dihydroxyethyl)-dC	$C_{11}H_{17}N_3O_6+H^+$	288.1190	288.1188
Y1 C5,N^4-(1-hydroxyethyl)-dC or Y2 5-(1-oxoethyl)-dC	$C_{11}H_{15}N_3O_5+H^+$	270.1084	270.1082
Z C5,N^4-etheno-dC	$C_{11}H_{13}N_3O_4+Na^+$	274.0798	274.0796

,*‘ – intensity (%) relative to the most abundant peak is shown in parenthesis.

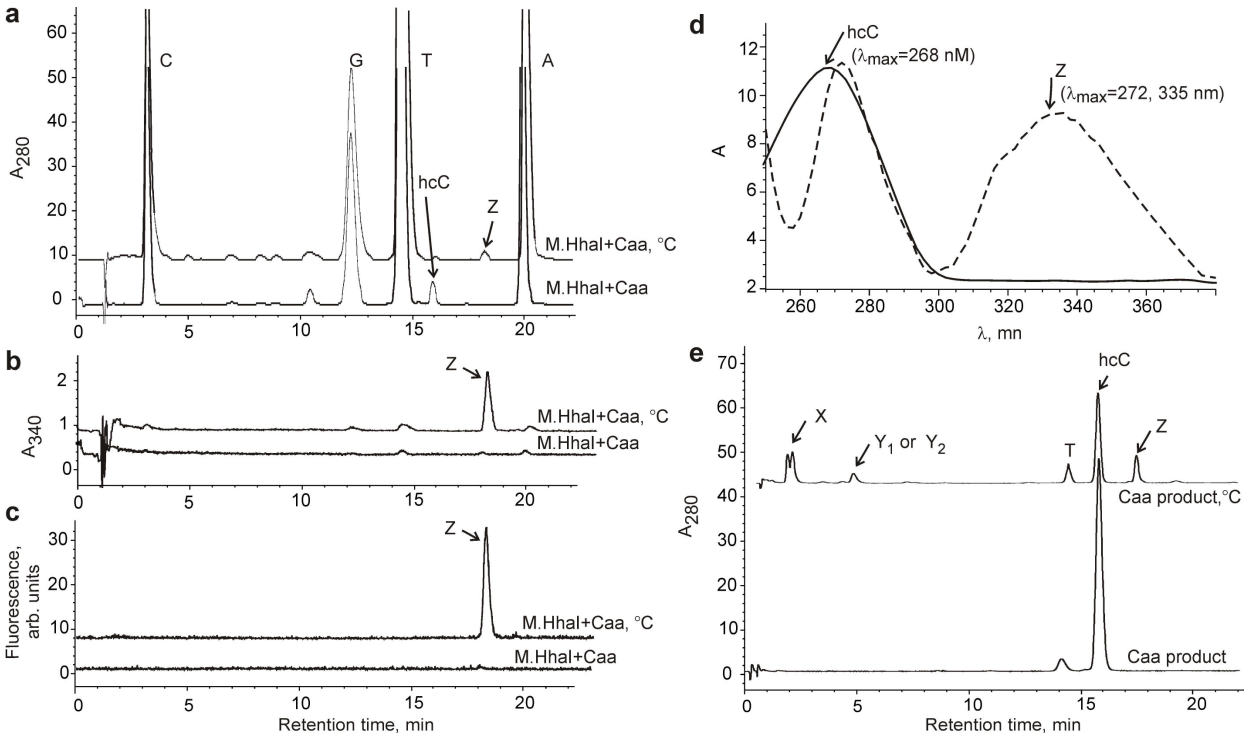

Figure 9. Cytosine adducts obtained after M.HhaI-DNR complex treatment with Caa. **a-c** RP-HPLC chromatograms at 280 nm (**a**), 340 nm (**b**) and fluorecence chromatogram (**c**; λ_{ex}= 345 nm, λ_{em}= 450 nm) of DNA modification products obtained upon treating 13 μM GCGC/GCGC-24 duplex with 13 μM M.HhaI and 200 mM Caa for 1 hour at room temperature (bottom traces) and after additional incubation at 85°C for 2 hours (top traces). Modified DNA was enzymatically digested to nucleosides and analyzed. Due to limited resolution/ sensitivity, only one of the two expected compounds hcC (either a dominant epimer or a mixture of both epimers) is observed). **d** UV absorption spectrum of cytosine adducts (pH 5.5). **e** RP-HPLC UV chromatograms at 280 nm of purified hcC (bottom trace) and after additional incubation for 2 hours at 85°C (top trace). Heating of the purified hcC for 2 hours at 85°C leads to the appearance of three new compounds X (two epimers are separated), Y (at least two different formulas with the same mass are possible) and Z. Also see Table 5; Fig. 10.

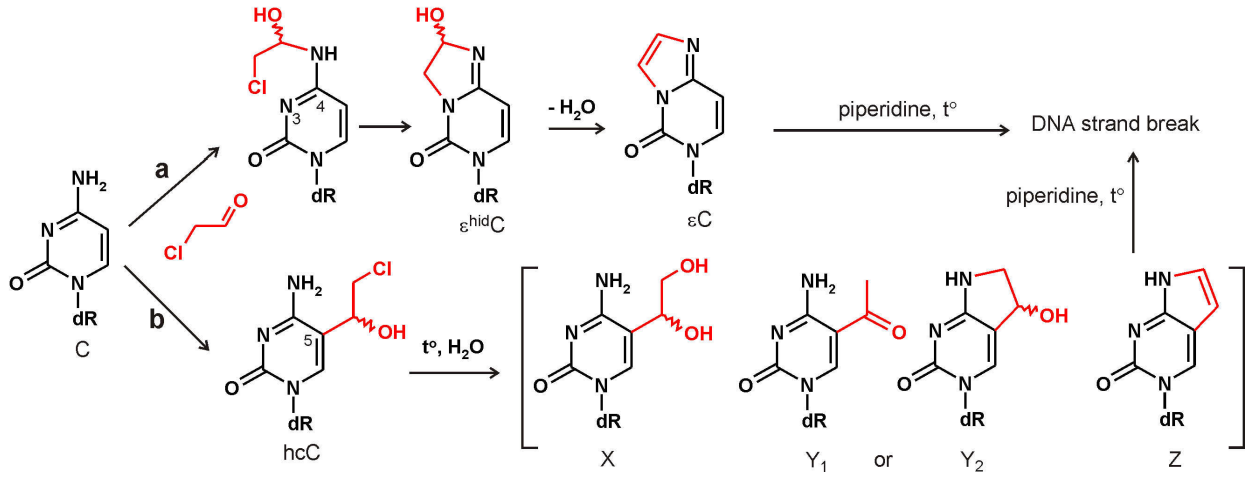

Figure 10. Reactios of 2-chloroacetaldehyde with unpaired cytosines**. a** Caa reacts with unpaired cytosine in single-stranded DNA producing N^4,N3-(1-hydroxyethyl)cytosine which dehydrate to N^4,N3-ethenocytosine (εC). εC reacts with piperidine producing DNA strand cleavage. **b** Proposed scheme of MTase reaction with Caa: primarily 5-(1-hydroxy-2-chlorethyl)-dC (hcC) is produced, which after heating converts to three different product X (5-(1,2-dihydroxyethyl)-dC), Z (C5,N^4-etheno-dC) and Y1 (5-(1-oxoethyl)-dC) or Y2 (C5,N^4-(1-hydroxyethyl)-dC); some of them or all generate DNA strand cleavage after piperidine treatment. Structures derived based on UV, HRMS and MS/MS spectra.

2.4 MTase-dependent conversion of 5-(1-hydroxyalkyl)cytosines to unmodified cytosines

We next examined whether DNA cytosine-5 MTases are capable of promoting the reverse reaction: the removal of formaldehyde from hmC. For this, the DNA duplex that contained enzymatically produced hmC residues at the target position was purified and then treated with a fresh portion of the same MTase. Fig. 11a,c shows that the amount of hmC was substantially reduced after such treatment whereas the amount of cytosine nucleotide was increased. Fig. 11b demonstrates that upon such treatment the DNA became partially unprotected against the action of a cognate restriction endonuclease. Only catalytically active MTase can convert 5-(1-hydroxyalkyl)cytosine to unmodified cytosine (Fig. 11a,b, line 4). Other tested MTases – M.SssI and M.HpaII – were able to catalyze similar reaction. Similar results were obtained with 5-hydroxyethyl-dC (Fig. 11b bottom). Our results thus demonstrate that DNA cytosine-5 MTases promote removal of a coupled aldehyde restoring the original target cytosine residue in DNA.

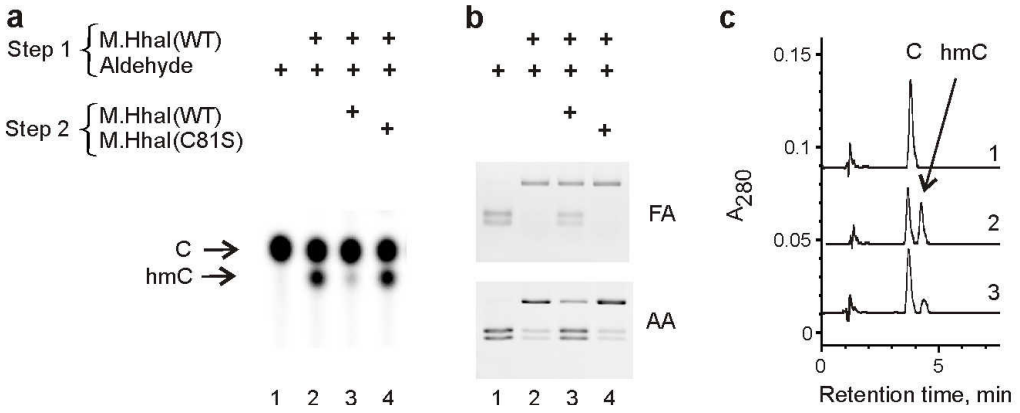

Figure 11. Catalytic reversal of sequence-specific 5-hydroxymethylation of cytosines in DNA by a DNA cytosine-5 methyltransferase. **a** TLC analysis of modified [^{33}P]-labeled nucleotides. Step 1: 20 nM G*CG*C/GCGC-32 duplex was incubated with 125 nM M.HhaI and 13 mM FA for 1 hour at r.t. Step 2 (lanes 3 and 4): FA modified DNA was purified by ethanol precipitation and treated with 750 nM M.HhaI (WT or its catalytic mutant C81S) for 2 hours at 37°C. Modified DNA duplex was enzymatically digested to 5'-dNMPs and analyzed by TLC/autoradiography. **b** Protection of DNA cleavage by R.Hin6I endonuclease. Step 1: 100 nM 618 bp PCR fragment was incubated with 50 nM M.HhaI and 13 mM FA (top) or 0.8 M AA (bottom) for 1 hour at r.t. Step 2 (lanes 3 and 4): aldehyde-modified DNA was purified by ethanol precipitation and treated with 800 nM M.HhaI (WT or its catalytic mutant C81S) for 2 hours at 37°C. Modified DNA was subjected to cleavage with R.Hin6I and analyzed by agarose gel electrophoresis. **c** HPLC analysis of modified 2'-deoxynucleosides. Reactions containing 13 μM GCGC/GCGC-24 duplex and 15 μM M.HhaI were treated with 13 mM FA for one hour at r.t. Modified DNA was purified by ethanol precipitation, redissolved and incubated with 15 μM M.HhaI for 1 hour at r.t. DNA samples were enzymatically fragmented to nucleosides and analyzed by HPLC. Traces 1-3 are denoted like in a.

2.5 Methyltransferase-directed coupling of thiols and selenols with 5-hydroxymethylcytosine

We found that the presence of an exogenous thiol compound such as 2-mercaptoethanol, cysteamine or L-cysteine (Cys) at millimolar concentrations led to the appearance of a new product at the expense of the original hmC (Fig. 12). HRMS (Table 6) and UV spectral analyses (not shown) of the new compounds show that the thiol replaces a hydroxyl group in the hmC residue. MS/MS analyses (Fig. 13) clearly

indicated that the sulphur atom is directly attached to the nucleoside. Taking into account that substrates containing C or mC at the target position were completely inert in the reaction (data not shown), and based on the documented reactivity of the 5-hydromethyl group towards a variety of nucleophiles (Tardy-Planechaud et all., 1997; Cline et all., 1959) we presumed that the thiol addition occurs at the 5-hydroxymethyl group yielding 5-(2-hydroxyethylthio)methyl-cytosine (hetmC), 5-(2-aminoethylthio)methyl-cytosine and 5-(2-amino-2-carboxyethylthio)methyl-cytosine (aketmC) residues in DNA, respectively. The identity of the isolated nucleoside hetmC was confirmed by direct chromatographic comparison of its nitrous acid deamination product with a chemically synthesized 5-(2-hydroxyethylthio)methyl-2'-deoxyuridine (hetmU; Fig. 14). Similar reactivity but even at lower concentrations was observed with corresponding selenols such as selenocysteamine and L-selenocysteine (Fig. 12). Adenosyl-5'-thiol, which can be regarded as a simple cofactor-mimic due to a potential anchoring interaction of the adenosine moiety with the enzyme, was also reactive at submilllimolar concentrations. Other types of nucleophiles such as hydrazine, p-nitrophenol, phenol, methanol, sodium sulfide, sodium azide, potassium bromide or sodium iodide showed no similar reactions. in our hands.

Besides M.HhaI, another bacterial cytosine-5 MTase, M.SssI, was examined and also showed clearly detectable catalytic activity at its target sites (not shown). The generality of this phenomenon indicates that the role of the enzyme is obviously to direct the reaction by flipping out and exposing in the catalytic centre a residue that occurs at its target position. Moreover, we find that the reaction requires the presence of the catalytic cysteine residue in the enzyme (Fig. 12a, lines 9-10).

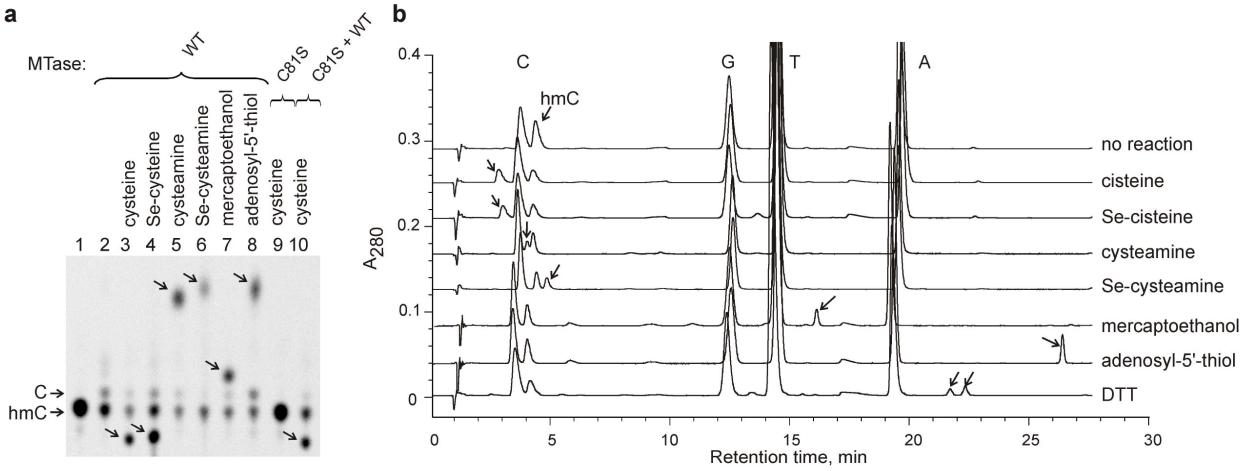

Figure 12. Methyltransferase-directed coupling of thiols and selenols with 5-hydroxymethylcytosine residues in DNA. **a** TLC analysis of a modified ^{33}P-labeled hmC nucleotide. A G*HGC/GMGC-31 duplex was treated with 12 mM L-cysteine, L-selenocysteine, cysteamine, selenocysteamine, 150 mM 2-mercaptoethanol, or 0.5 mM adenosyl-5'-thiol in the presence of 200 nM M.HhaI (wild type (WT), a catalytic C81S mutant, or both) for 1.5 h at r.t. Modified DNA was digested to 5'-mononucleotides, analyzed by TLC, and autoradiographed. **b** Reversed-phase HPLC analysis of modified 2'-deoxynucleosides. A GCGC/GCGC-24 duplex (13 μM) was treated with 13 mM FA and M.HhaI (13 μM) for 1 hour at r.t. Modified DNA was purified by ethanol precipitation, redissolved and incubated with an exogenous reagent as above (50 mM Cys, Se-Cys, dithiothreitol (DTT); 12 mM cysteamine, selenocysteamine; 0.5 mM adenosyl-5'-thiol, 500 mM 2-mercaptoethanol) in the presence of 13 mM M.HhaI for 1 h at r.t. Modified DNA was digested to nucleosides and analyzed by HPLC-MS method. Arrows indicate peaks corresponding to modified nucleosides.

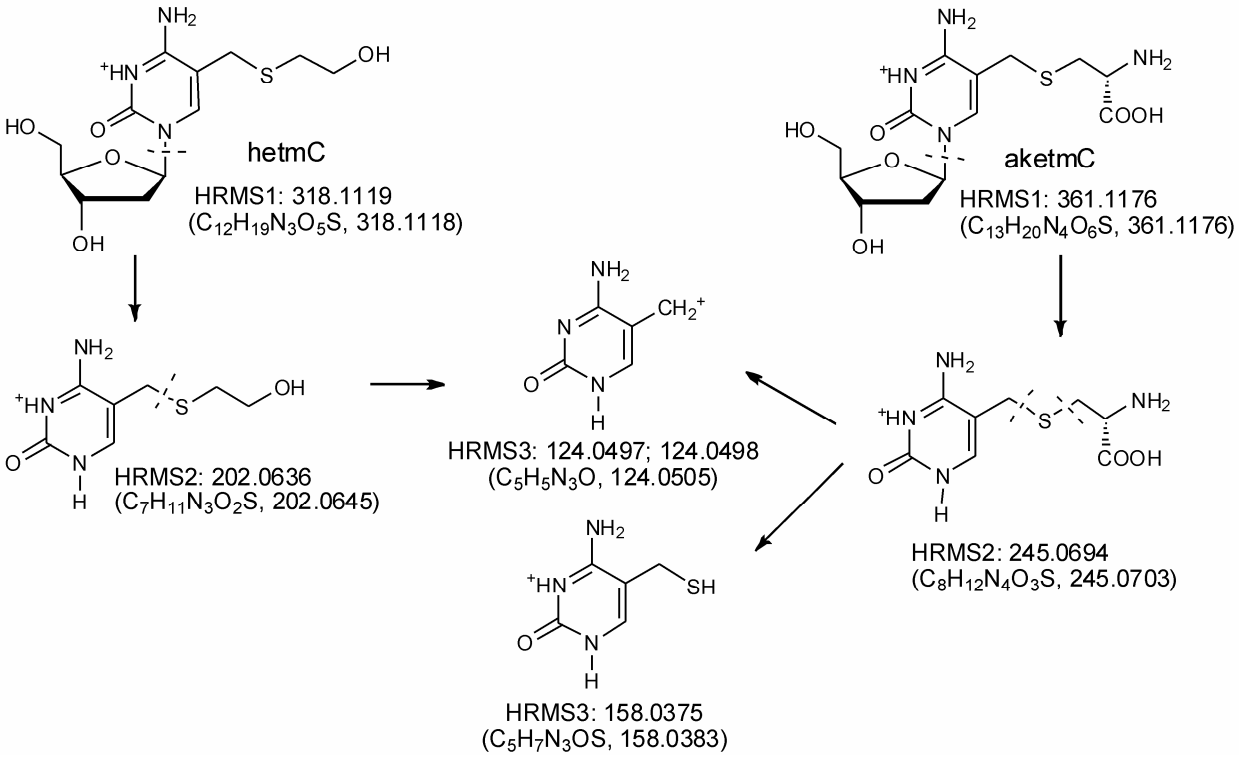

Figure 13. Identification of reaction products and their fragment ions using high resolution MS/MS analysis. Molecular formulas and counted masses are shown in parenthesis.

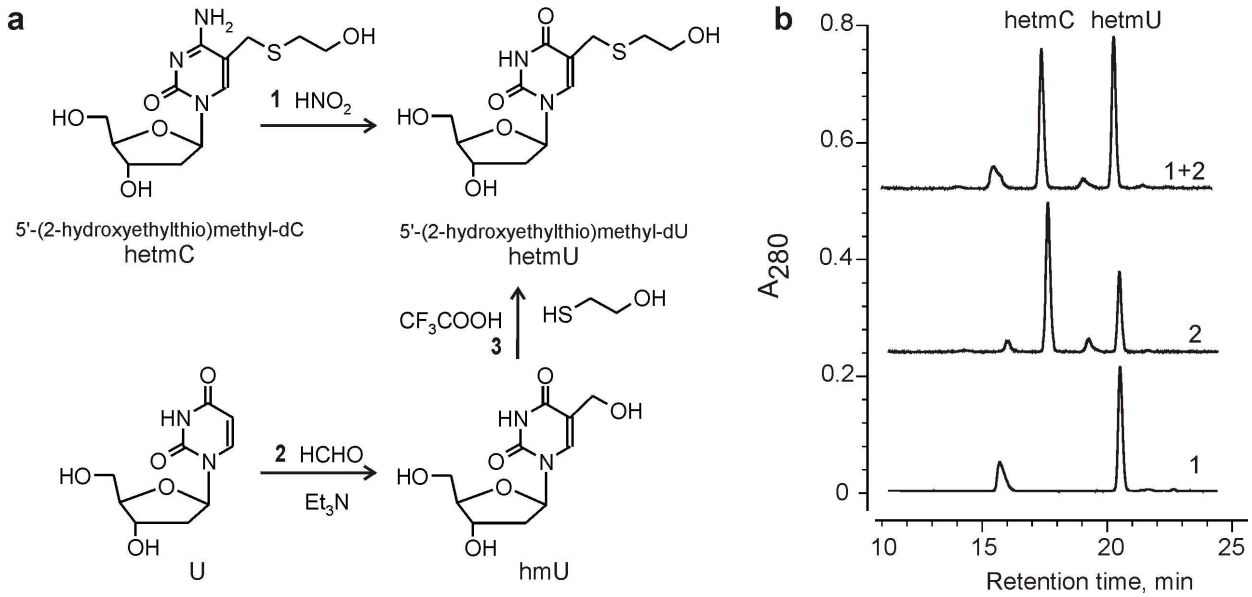

Figure 14. The proof of methyltransferase-directed modification of hmC in DNA with exogenous reagents. **a** The enzymatic reaction product 5'-(2-hydroxyethylthio)methyl-dC (hetmC) was deaminated with nitrous acid to a corresponding uracil derivative 5-(2-hydroxyethylthio)methyl-dC (hetmU) (1). This derivative was compared with synthetically produced hetmU (2, 3). **b** HPLC analysis of synthetically produced hetmU (1), products of partial nitrous acid deamination of the purified hetmC (2) and the two samples combined together (1+2).

Table 6. Methyltransferase-directed modification of hmC in DNA with exogenous reagents.

Exogenous reagent / Reaction product	Extent of target hmC modification observed by TLC (HPLC), %		HRMS of reaction product, m/z	
	HhaI	SssI	Formula (M^+)	Found / Calculated
2-mercaptoethanol / 5-(2-hydroxyethylthio)methyl-dC (hetmC)	55 (22)	10	$C_{12}H_{19}N_3O_5S$	318.1119/ 318.1118
L-cysteine (Cys)/ 5-(2-amino-2-carboxyethylthio)methyl-dC (aketmC)	60 (36)	10	$C_{13}H_{20}N_4O_6S$	361.1176/ 361.1176
L-Selenocysteine (Se-Cys)/ 5-(2-amino-2-carboxyethylseleno)methyl -dC	60 (27)	-	$C_{13}H_{20}N_4O_6Se$	409.0621/ 409.0622
cysteamine/ 5-(2-aminoethylthio)methyl -dC	75 (25)	60	$C_{12}H_{20}N_4O_4S$	317.1277/ 317.1278
Se-cysteamine/ 5-(2-aminoethylseleno)methyl-dC	75 (41)	14	$C_{12}H_{20}N_4O_4Se$	365.0725/ 365.0723
adenosyl-5'-thiol/ 5-(5‘-adenosylthio)methyl-dC	50 (24)	-	$C_{20}H_{26}N_8O_7S$	523.1718/ 523.1718
1,4-dithiothreitol (DTT)/ 5-(2,3-dihydroxy-4-mercaptobutylthio)methyl-dC	n. (30)	-	$C_{14}H_{23}N_3O_6S_2$	394.1105/ 394.1101

"-" – not performed;
"n." – not determed because additional spots were not seen in TLC.

2.6 Mechanism of atypical reactions of cytosine-5 methyltransferases

The MTase-directed aldehyde adding to target cytosines, thiol and selenol coupling to hmC and 5-(1-hydroxyalkyl)cytosine conversation to unmodified cytosine occur under mild conditions and with high sequence and base specificity. The triple role of an MTase is thus to 1) recognize a specific DNA sequence, 2) present a target nucleobase in the catalytic center and 3) covalently activate the C5 position in the base.

Naturally DNA cytosine-5 MTases catalyze site-specific transfers of a methyl group from the ubiquitous cofactor SAM onto the 5-position of their target cytosine residues in DNA (Fig. 15). The C5 position of the cytosine ring is activated by nucleophilic addition of a conserved cysteine residue of the enzyme to the 6-position of the ring, forming activated cytosine intermediate (ACI) (Wu and Santi, 1987; Klimašauskas et all., 1994). ACI can be formed without any cofactor and is responsible for cytosine-5-H exchange reaction (Wu and Santi, 1987) and also aldehyde coupling. The C5 atom of cytosine in ACI is nucleophilic enough to attack electrophilic center of SAM methyl group or aldehyde. When methyl group from SAM or hydroxyalkyl group from aldehyde is added on the C5 atom, the covalent 5,6-dihydrocytosine intermediate (CDI) appears which undergoes β-elimination producing modified cytosine, free enzyme and in the former case SAH. The reactive aldehydes are not bona fide cofactors of SAM-dependent MTases because they lack an anchor moiety (such as adenosyl) that would assist in the formation of a discrete, specific complex with the enzyme. The chemical reaction (nucleophilic addition) itself is different from the S_N2 transfer (nucleophilic substitution) naturally catalyzed by MTases (Fig. 15).

Normally, at much higher concentrations aldehydes attack exocyclic amino groups of nucleobases in DNA, which leads to N-hydroxymethyl derivatives and further disubstituted products (Beland et all, 1984). These reactions are largely responsible for

the cytotoxicity of formaldehyde *in vivo*, and are exploited for crosslinking of interacting proteins (Solomon and Varshavsky, 1985; Wang et al., 2007).

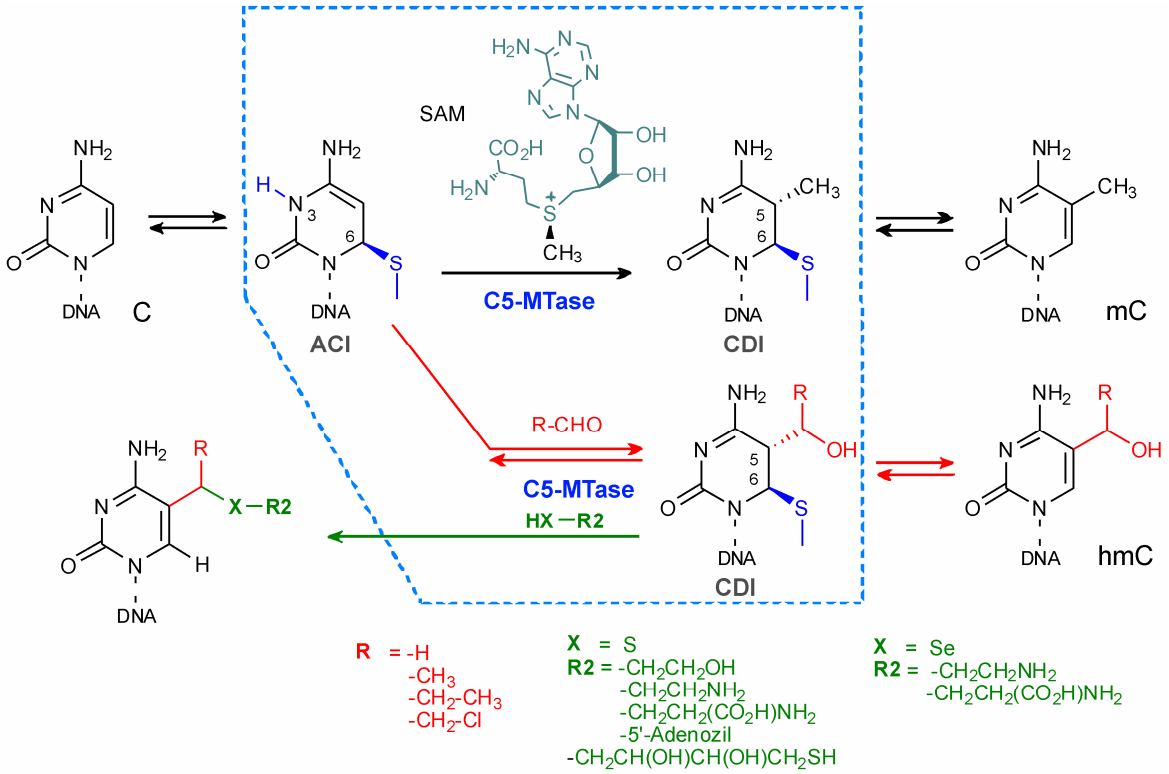

Figure 15. Transformations of a target residue by DNA cytosine-5 methyltransferases. Biological methylation by C5-MTases occurs through an S_N2 reaction between an activated cytosine intermediate (ACI) and the cofactor SAM to give a covalent 5,6-dihydrocytosine intermediate (CDI) and then 5-methylcytosine (mC). The activated cytosine residue can also participate in reversible addition reactions (red-colored route) with aldehydes to yield 5-(1-hydroxyalkyl)cytosine. The latter compound can undergo further methyltransferase-directed condensation (green-colored route) with thiol or selenol reagents to give stable 5-chalcogenomethyl derivatives. Modifying reagents are shown in red (aldehydes) or green (thiol and selenols), C5-MTase and its catalytic moieties are shown in blue, boxed areas denote species and reactions within the catalytic centre of the enzyme.

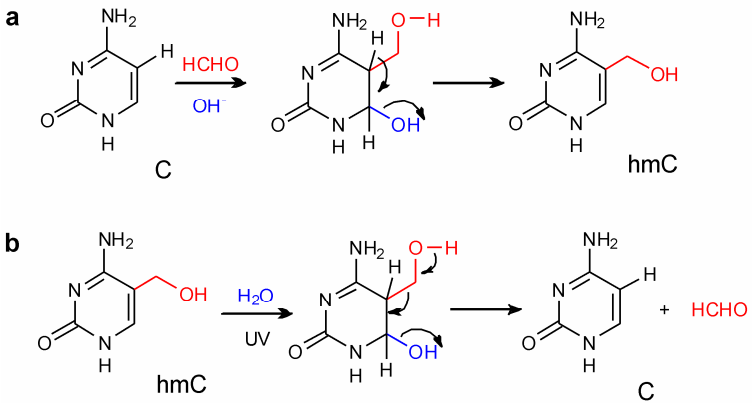

Figure 16. Non-enzyme directed cytosine reaction with formaldehyde and 5-hydroxymethylcytosine dehydroxymethylation. **a** Chemical synthesis of hmC from cytosine and FA in the presence of high pH. **b** Exposing to UV, hmC dehydroxymethylates to cytosine. Intermediate of both reaction is 6-hydroxy-6,5-dihydrocytosine (Cline et al., 1959; Alegria, 1967; Privat and Sowers, 1996).

Addition of formaldehyde to cytosine can also be achieved synthetically in the absence of enzyme. In basic solution, pyrimidines react with formaldehyde to form the corresponding 5-hydroxymethyl derivatives (Cline et al., 1959; Alegria, 1967). The reaction conditions are severe and would lead to DNA degradation therefore only bases,

nucleosides or nucleotides can be used. The mechanism of this reaction is similar to MTases and involves the formation of a covalent 5,6-dihydro-6-hydroxy intermediate. The same intermediate was proposed for the direct formation of cytosine from hmC by releasing the formaldehyde in the present of UV or high pH (Fig. 16) (Cline et al., 1959; Alegria, 1967; Privat and Sowers, 1996).

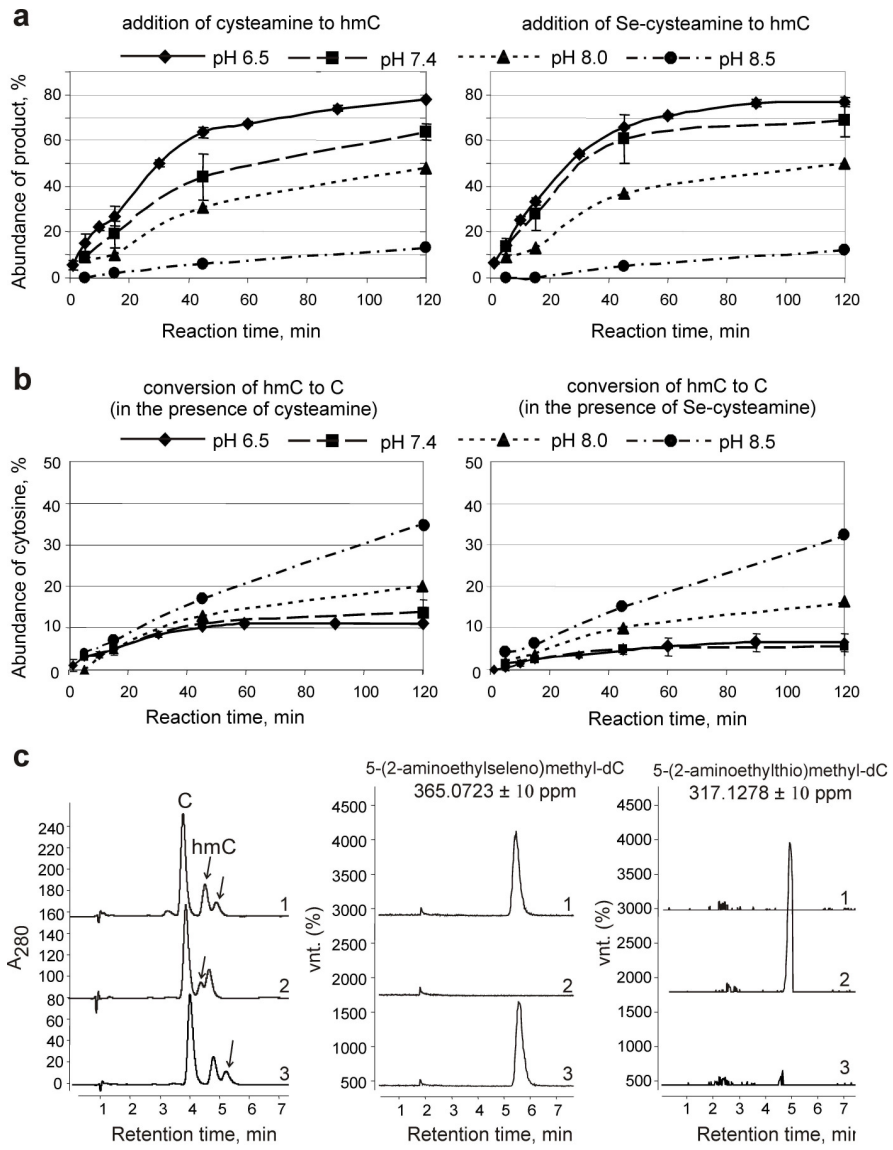

Figure 17. Kinetics of M.HhaI-directed addition of thiol and selenol to hmC (**a**) and dehydroxymethylation of hmC in the presence of tiols and selenols (**b**). 20 nM internally ^{33}P-labeled DNA duplex G*HGC/GMGC-31 was incubated with 12 mM cysteamine (left panel) or selenocysteamine (right panel) in the presence of 200 nM M.HhaI Q82A/Y254S/N304A mutant in different buffers: sodium acetate (pH 6.5 and 7.4), sodium phosphate (pH 8.0) and Tris-HCl (pH 8.5). Modified DNA was digested to 5'-mononucleotides, analyzed by TLC, autoradiographed and analyzed with MultiGauge software. **c** HPLC-MS analysis of modified 2'-deoxynucleosides. Reactions containing 13 μM GCGC/GCGC-24 duplex and 13 μM M.HhaI were incubated for 1 hour at room temperature with 13 mM FA. Modified DNA was purified by ethanol precipitation, redissolved and incubated with 1.2 mM cysteamine, 1.2 mM selenocysteamine or 0.6 mM of each cysteamine and selenocysteamine. Modified DNA was digested to nucleosides and analyzed by HPLC-MS. Left – HPLC chromatogram at 280 nm; middle and right – chromatograms of mass extraction of product 5-(2-aminoethylthio)methyl-dC and 5-(2-aminoethylseleno)methyl-dC, respectively.

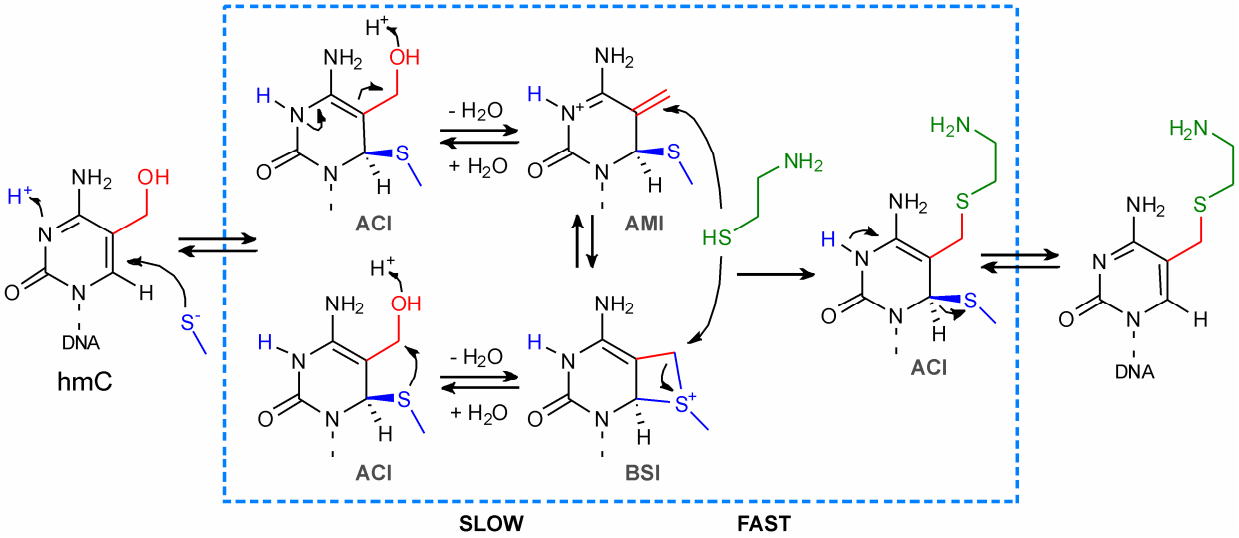

Figure 18. Proposed mechanism of nucleophile condensation at the target hmC residue catalyzed by a cytosine-5 MTase. The reaction is thought to proceed via a activated cytosine intermediate (ACI) followed by either a highly active 5-methylene intermediate (AMI) or a bicyclic sulfonium intermediate (BSI), which would then undergo the fast addition of a nucleophilic reagent, such as cysteamine. The C5-MTase and its catalytic moieties are shown in blue; boxed areas denote species and reactions within the catalytic centre of the enzyme; formaldehyde and cysteamine are shown in red and green, respectively.

The most straightforward mechanism of thiol/selenol coupling to hmC 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 (Fig. 17a). On the other hand, a seleno product always prevailed over a thio product even in the presence of an excess amount of thiol (Fig. 17c). 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 (Barrett and Maltby, 1998) and deoxycytidylate hydroxymethylase (Graves et al., 1992), 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; Fig. 18). 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 (BSI) containing a four-membered thietane ring (Fig. 18). 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 (Law and Bruice, 1999) but whether the proposed bicyclic structure is well-compatible with the active sites of other MTases 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. At higher pH, the rate of the nucleophile condensation goes down whereas a competing reaction, dehydroxymethylation via deprotonation at C5, becomes faster (Fig. 17a,b). This reverse reaction also requires

catalysis, which suggests that it proceeds via a covalent 5,6-dihydrocytosine intermediate (CDI, Fig. 19). A mechanism for this reaction can be derived from analogy with the light- and alkali-induced conversions of hmC to cytosine, which occur via corresponding 5,6-dihydro-6-hydroxy derivatives (Fig. 16; discussed above).

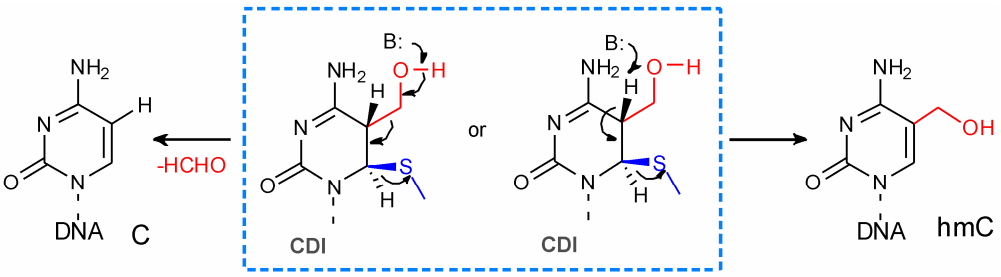

Figure 19. Proposed mechanism of decay of CDI (center). Depending on whether the exocyclic hydroxyl group or the ring 5 position is deprotonated, the release of enzyme yields either the unmodified cytosine and free formaldehyde (left) or the coupling product hmC (right), respectively. Formaldehyde is shown in red, MTase and its catalytic moieties are shown in blue, boxed areas denote species and reactions within the catalytic center of the enzyme.

2.7 Practical applications of the atypical reactions

The described novel atypical reaction of DNA cytosine-5 MTases to catalyze the condensation of exogenous nucleophiles to the 5-hydroxymethyl group of hmC residues in DNA offers new ways for sequence-specific derivatization and labeling of DNA. Since the L-cysteine and cysteamine condensation product contained an aliphatic primary amino group, we examined if this or similar modifications can be exploited as chemical anchors for covalent attachment of reporter groups to DNA (Fig. 20a). Moreover, since the MTase-directed condensation of nucleophiles is not possible at 5-methylated and unmethylated cytosine, the derivatization reaction is well suited to query the hydroxymethylation status of CG sites in mammalian genomic DNA. This can be easily achieved by combining cysteamine or selenocysteamine addition with amino-selective biotinylation. Such labeled DNA is selectively retained on streptavidine beads, and the enriched hmC fraction can then be analyzed by RT-PCR (Fig. 20b) or other techniques (DNA microarrays or sequencing). The above examples thus demonstrate a unique potential of the new chemo-enzymatic approach for genome-wide mapping of epigenetic cytosine modifications in DNA.

Similarly, the MTase-directed reactions open new ways for sequence-specific labeling of DNA which is applied in scientific and medical laboratories (Ghosh et al, 2006; Juskowiak, 2010). DNA labeling approach using MTases and double activated synthetic SAM analogues was developed before (Dalhoff et al., 2006; Lukinavičius et al. 2007). These reactions are highly efficient and target-specific however the cofactor analogues are not yet commercially available and their synthesis can only be accomplished in specialized laboratories. The coupling reactions involving small aldehydes are simple and fast, and are thus suitable for routine laboratory applications. Modification reactions with longer aldehydes carrying useful functional groups proved inefficient in our hands, but this may improve by steric engineering of the enzymes. For example, protein engineering (side chain shortening) of the cofactor pocket in M.HhaI mutant Q82A/N304A restores the FA reaction in the presence of SAH and enhances acetaldehyde coupling in the absence of SAH (Fig. 20c). DNA labeling can also be

achieved in three steps using simple reagents – formaldehyde and L-cysteine. Plasmid pUC19 DNA was first 5-hydromethylated at the target cytosine residue in the presence of M.HhaI and FA. Then it was treated with L-cysteine and M.HhaI and labeled with a fluorescein by incubation with an NHS-ester. The fluorescence intensity distribution in four pUC19-FspBI fragments was fully consistent with the positions and numbers of the M.HhaI sites in the original plasmid (Fig. 20d).

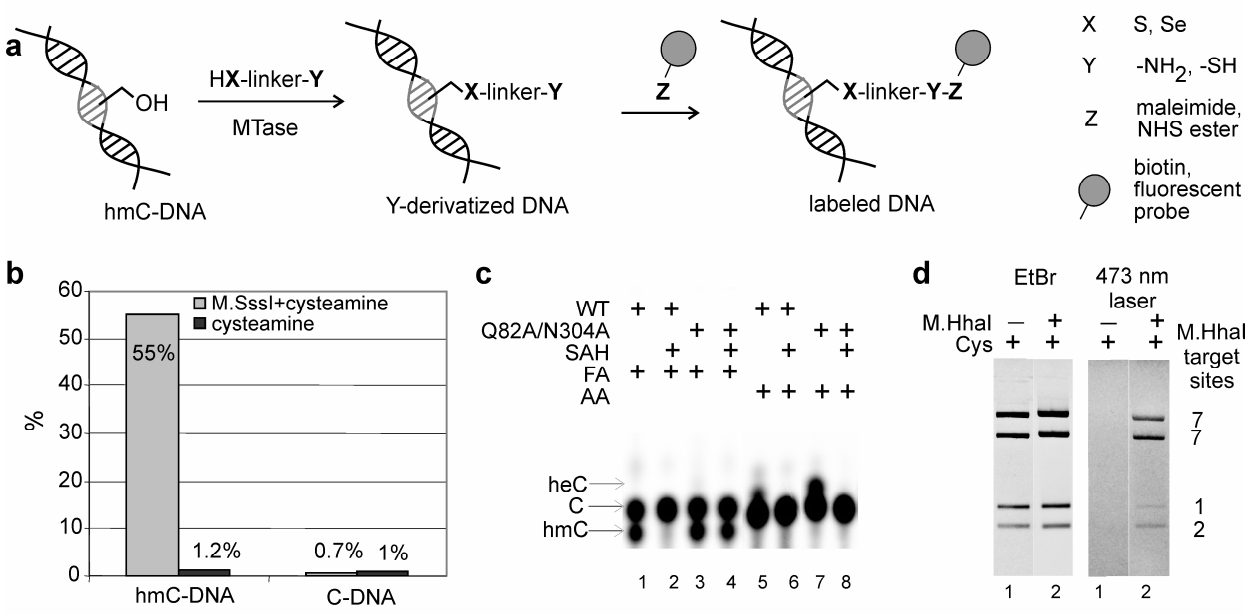

Figure 20. a General approach to the selective derivatization and labeling of hmC-containing DNA with a reporter group (shown as a ball). DNA bases in M.HhaI target site are shown as grey sticks. **b** A mixture of a hmC-containing (hmC-DNR) and non-hmC DNA (C-DNA) fragments was treated with cysteamine in the presence of M.SssI, followed by labeling with biotin NHS 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. **c** Reactions, containing a [^{33}P]-labeled 20 nM DNA duplex G*CG*C/GMGC-32, 560 μM SAH (lanes 2, 4, 6, 8), 125 nM M.HhaI WT (lanes 1, 2, 5, 6) or Q82A/N304A mutant (lanes 3,4, 7, 8), were treated with 13 formaldehyde (FA; lanes 1–4) or 0.8 M acetaldehyde (AA; lanes 5–8) for 1 h at r.t. Modified DNA was digested to 5'-dNMPs and analyzed by TLC/autoradiography to reveal 5'-[^{33}P]-labeled products. **c** Fluorescence labeling of a pUC19 plasmid containing hmC residues at the HhaI target sites. The DNA was amino-derivatized by treatment with Lcysteine in the presence of M.HhaI, and the derivatized DNA was labeled with fluorescein NHS 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). Lane 1 is the control with M.HhaI omitted.

3 Discussion

3.1 Chemical mapping of cytosines enzymatically flipped out of the DNA helix

Our results demonstrate the applicability of the Caa modification reaction for probing single extrahelical cytosines in MTase-DNA and REase-DNA complexes. This assay is cheap, simple and sensitive and can be easily performed in a standard laboratory. While studying MTase base flipping by Caa we found that both target base flipping and covalent activation of the target cytosine by the MTase are necessary for

getting positive results. The Caa specificity to cytosine is different in MTase-DNA complexes compared to single stranded DNA and REase-DNA complexes thus different reaction products are formed, respectively. One should bear in mind that Caa is a highly reactive electrophile with respect to numerous nucleophilic centres present in proteins. Thereby we observed dose-dependent inactivation of the DNA methyltransferases under the assay conditions.

3.2 Atypical DNA cytosine-5 methyltransferase reactions

Our findings reveal a previously unknown ability of cytosine-5 MTases to catalyze the reversible covalent addition of short aliphatic aldehydes to the 5th position of target cytosine residues in DNA yielding 5-(1-hydroxyalkyl)cytosines. The latter cytosine adducts proceeds the condensation of exogenous alkylthiols and alkylselenols forming 5-chalcogenomethylcytosine *in vitro* in the presence of active methyltransferase. To our knowledge, this is the first demonstration of wild-type cofactor-dependent enzymes catalyzing an atypical chemical reaction using non-cofactor-like exogenous substrates. The reactive aldehydes, thiol and selenols are not bona fide cofactors of SAM-dependent MTases because they lack an anchor moiety (such as adenosyl) that would assist in the formation of a discrete, specific complex with the enzyme. Moreover, methyltransferases are able to recognise and modified changed target base: 5-hydromethylcytosine instead of cytosine. The triple role of an MTase is thus to (i) recognize a specific DNA sequence, (ii) present a target nucleobase in the catalytic centre and (iii) covalently activate the C5 position in the base.

Atypical methyltransferase reactions were discovered at the same time when hmC was found in eukaryotic DNA. This conjunction enables us to offer an original technique to determine hmC positions in genomic DNA and also let us to propose new participants in DNA demethylation and hydroxymethylation pathways in mammals.

3.2.1 Link to DNA demethylation and hydroxymethylation

It was shown recently that 2-oxoglutarate- and Fe^{2+}-dependent oxygenases Tet proteins convert mC to hmC *in vivo* and *in vitro*. Supposedly, this is the main if not the only way of hmC origin in eukaryotic DNA (Tahiliani et al., 2009). Methyltransferases could also contribute to higher amount of hmC in mammals DNA us formaldehyde occurs in millimolar concentrations in certain tissues of rats and humans (Heck and Casanova, 2004). Our results show that mouse methyltransferase Dnmt1 can add formaldehyde to cytosine *in vitro* producing hmC although at a much lower efficiency compared to prokaryotic MTases (Fig. 7c). MTases catalyze the coupling of aldehydes provided if it is not in complex with cofactor SAM. The concentration of SAM in the cell is high because SAM is vitally essential to many processes of cell biochemistry. However, eukaryotic methyltransferases are large multiple-domain proteins and thus could possess specific mechanisms (conformational change or binding of other proteins) to block the entrance of SAM to the active site of MTases. Additionally, in some cellular compartments the concentration of SAM may be low.

Interestingly, mC and hmC was also found in eukaryotic mitochondrial DNA. Dnmt1 gene has an alternative translation start site coding the special peptide which is responsible for the import to the mitochondria. Tet proteins do not have similar peptides thus it is not known how hmC appear in the mitochondrial DNA (Shock et al., 2011). Thereby Dnmt1 could be a potential player.

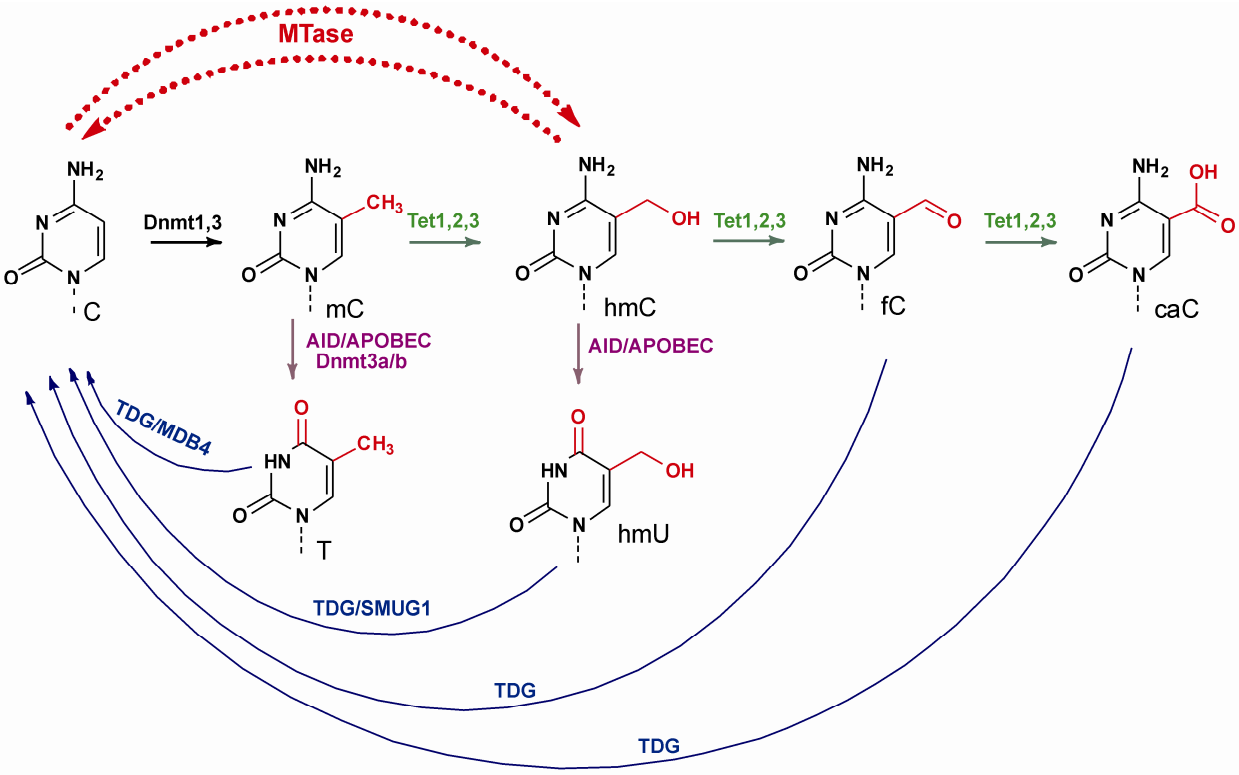

Figure 21. Formation and removal of epigenetic marks in mammalian DNA. Cytosine (C) is converted to 5-methylcytosine (mC) by action of endogenous DNA MTases of Dnmt1 and Dnmt3 families. Several mechanisms for DNA demethylation, in which 5-methylcytosine (mC) is converted back to C, have been proposed. Green arrows represent oxidation-based pathways performed by Tet proteins: methyl group of mC is consecutive oxidized to hydroxymethyl, formyl and carboxy groups forming 5-hydroxymethylcytosine (hmC), 5-formylcytosine (fC) and 5-carboxycytosine (caC), respectively. Lilac arrows show deamination-based pathways where mC and hmC are deaminated to thymine (T) and 5-hydroxymethyluracil (hmU), respectively, in the presence of AID, APOBEC and maybe DNA methyltransferases Dnmt3a and Dnmt3b. Blue arrows denote base excision repair (BER) pathways involving TDG, MBD4 and SMUG1 glycosylases. Red dotted arrows denote the newly discovered hydroxymethylation and dehydroxymethylation reactions performed by cytosine-5 methyltransferases *in vitro*.

In contrast to the well studied genetics and biology of DNA methylation, the enzymatic mechanism of demethylation remains elusive and controversial (Wu and Zhang, 2010). The direct removal of methyl group of mC requires an enzyme with enormous catalytic power because of the strength of the carbon–carbon bond that needs to be broken. If the methyl group is oxidized to hydroxymethyl, aldehyde or carboxy group the energy of removing those groups would be lower. Certain DNA repair enzymes can reverse N-alkylation damage to DNA via a two-step mechanism, which involves an enzymatic oxidation of N-alkylated nucleobases (N3-alkylcytosine, N1-alkyladenine) to corresponding N-(1-hydroxyalkyl)derivatives. These intermediates then undergo spontaneous hydrolytic release of an aldehyde from the ring nitrogen to directly generate the original unmodified base (Sedgwick et al., 2007). Histone demethylation is proceeding through the similar route whereas in thymidine salvage pathway C-C bong is enzymatically broken starting with thymine-7-hydroxylase, which carries out three consecutive oxidation reactions of thymine methyl group to hydroxymethyl, formyl and carboxy groups and finishing with decarboxylase, which releases the carbon dioxide to

produce uridine. Tet proteins could also operate in the similar way as they need the same cofactor as thymine-7-hydroxylase: 2-oxoglutarate, Fe^{2+} and oxygen (Wu and Zhang, 2010). Tet proteins efficiently oxidise the methyl group of mC producing hydroxymethyl group and hmC but further oxidation steps to formyl and carboxy groups forming 5-formylcytosine (fC) and 5-carboxycytosine (caC), respectively, are inefficient and were not observed initially (He et al., 2011; Ito et al., 2011) either fC and caC were not detected in eukaryotic DNA (Globisch et al., 2010). Afterwards more sensitive methods were invoked, which enabled to detected fC and caC thought at very low amounts (more than 100 time less then hmC) (Ito et al., 2011; Pfaffeneder et al., 2011) and few mechanisms were proposed and demonstrated concerning further hmC, fC and caC conversion to cytosine *in vivo* (Fig. 21) (He et al., 2011; Maiti and Drohat, 2011; Cortellino et al., 2011; Guo et al., 2011). Our finding that DNA cytosine-5 MTases can promote conversion of hmC to C (Fig. 11) provides a plausible precedent that hmC could be directly transformed to C. This reaction could be carried by special protein, which proceeds through the similar catalytic mechanism like methyltransferases. Although DNA methyltransferases could also perform this reaction in the cell as it was shown that methyltransferases are involved in DNA demethylation *in vivo* (Metivier et al., 2008).

3.2.2 Practical employment of atypical methyltransferase reactions

Discovery of a new base – hmC – in eukaryotic DNA three years ago stimulated the intensive research on hmC biological function and distribution in genomic DNA. Unfortunately, the gold standard in mC research – DNA sequencing after bisulfite treatment – can not distinguish mC from hmC (unmodified cytosines deaminate to uridines which are seen as „A“ in sequencing, but C5 modified cytosines do not deaminate thus are seen as normal „G“) (Fig. 22). Restriction endonuclease which are sensitive to cytosine-5 modification also do not discriminate mC from hmC. Only methods which are based on antibodies or methyl-DNA binding proteins can differentiate the methyl and hydroxymethyl group at C5 position of cytosine (Huang et al., 2010; Jin et al., 2010; Nestor et al., 2010). Thus, the data of the earlier mC distribution in genomic DNA should be interpreted carefully and new methods are needed for hmC, mC and C differentiation. During the 3 years we and others proposed several different techniques:

1) R.MspI, R.GlaI and R.MspJI cleave the target DNA if cytosine, mC or hmC is present, but glucosylated hmC is not cleaved. Effective glucosylation of hmC is achieved by specific enzymes of T4 phage: α and β glucosyltransferases (AGT and BGT), which use uridinyl-(1-glucosyl)diphosphate (UDP-glc) as a cofactor (Fig. 23a) (Zheng et al., 2010; Davis and Vaisvila, 2011). Other REases hydrolysed DNA if it is hydroxymethylated (Szwagierczak et al., 2011; Wang et al., 2011).

2) Single molecular DNA sequencing which can discriminate hmC from mC without any additional steps (Flusberg et al., 2010; Wallace et al., 2010; Wanunu et al., 2010).

3) Special protein from Trypanosoma which can interact with glucosylated hmC (Robertson et al., 2011).

4) Chemical oxidation of hmC which leads to DNA strand cleavage in the presence of piperidine (Okamoto et al., 2011).

5) Oxidation of glucosylated hmC by $NaIO_4$ gives two aldehyde groups which can be used to attach reporter molecule (Fig. 23a) (Pastor et al., 2011).

6) Using hmC or 5-sulfonylmethylcytosine (which is formed by treating hmC with bisulfite (Fig. 22c)) antibodies (Ficz et al., 2011; Jin et al., 2011; Pastor et al., 2011).

7) Special modified cofactor analogues of T4 phage glucosyltransferase having azide (Song et al., 2010) or aldehyde group (Song et al., 2011, b) instead of UDP-Glc gliucose motif 6-OH group (Fig. 23b). Functional azide and aldehyde groups are used to couple biotine which is further used to separate hmC containing DNA.

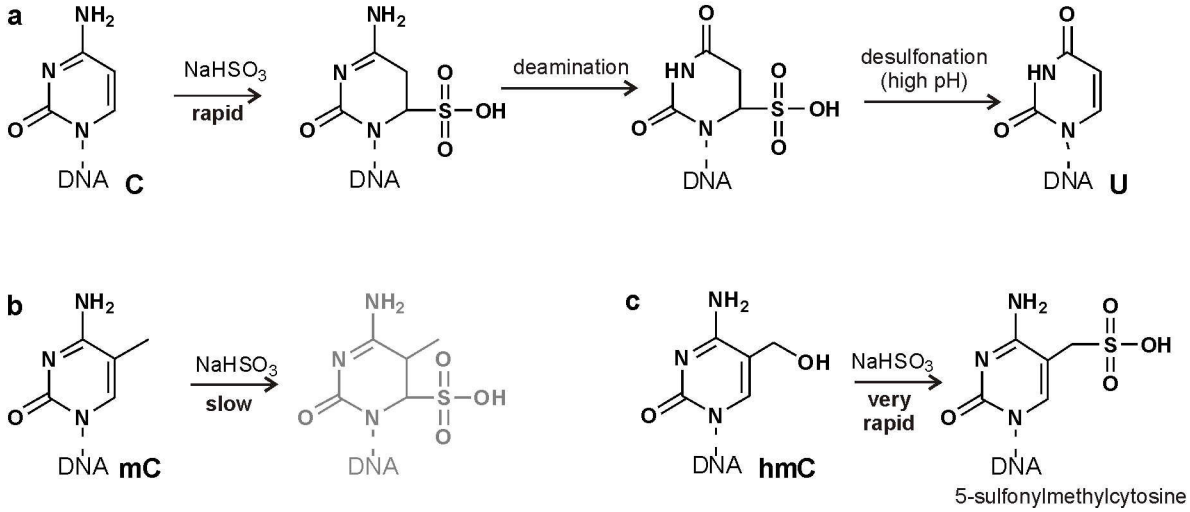

Figure 22. Reactions of cytosine (**a**), 5-methylcytosine (**b**) and 5-hydroxymethylcytosine (**c**) with sodium bisulfite ($NaHSO_3$) (Huang et al., 2010).

The majority of these approaches were demonstrated to work with model DNA substrates and do need further optimization and validation on genomic DNA. Only the last three methods were tested with eukaryotic DNA. The main desirable features of assays are one nucleotide resolution and potential to analyse all DNA of organism, for instance using deep DNA sequencing or microplates. These two characteristics now are covered only by the last approach in combination with single-molecule real-time (SMRT) DNA sequencing where DNA polymerase stops beside modified hmC (Song et al., 2011, a). Likely, in our method we could get similar result if the functional groups would be further modified with a big chemical group. The other advantage of our method is a potential to analyse all CG dinucleotides using M.SssI. In mammalian DNA, mC and hmC are mainly found in these dinucleotides. Together using the older mTAG (methyltransferase-directed transfer of activated groups) technology (see below), both mC and hmC positions could be determined applying similar experimental procedure.

Another field were atypical reactions of cytosine-5 methyltransferase can be adapted – DNA labeling at specific target sites which is applied in scientific and medical diagnostic laboratories (Ghosh et al., 2006). A few years ago MTases and double-activated synthetic analogues of cofactors were started to use for DNA labeling (Dalhoff et al., 2006; Lukinavičius et al., 2007). The mTAG approach comprises of two steps: 1) a functional (amino, mercapto, alkyne or other) group from the cofactor analogue is transfered on DNA by a MTase, 2) a reporter (biotine, fluorescent dye) is attached to functional group applying well-known chemistry. mTAG method is selective and efficient but requires expensive cofactor analogues. Using two readily available reagents – formaldehyde and L-cysteine – the functional amino group can be transferred by a

MTase onto DNA in two steps (Fig. 23c, 20d). Thus atypical reactions of cytosine-5 methyltransferases could be a cheaper but more elaborate alternative for sequence-specific labeling of DNA.

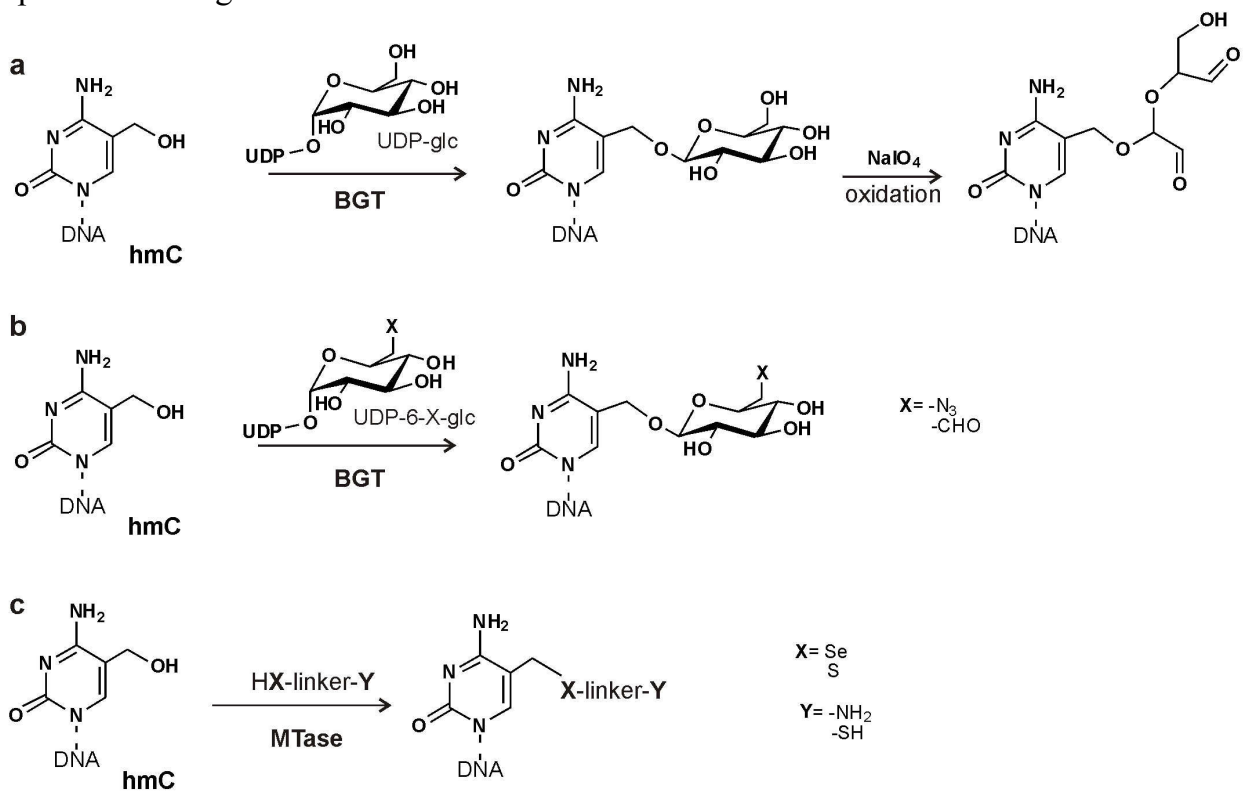

Figure 23. Enzymatic labeling of 5-hydroxymethylcytosine in DNA. **a, b** T4 phage β-glucosyltranferase (BGT) transfer the glucose motif from natural cofactor (**a**, UDP-glc) or synthetic cofactor analogue having functional azide or aldehyde group (**b**, UDP-6-X-glc) to hmC. The attached glucose can be oxidized with $NaIO_4$ originating two aldehyde groups (**a**). **c** DNA cytosine-5 methyltransferase modifies hmC being at the target site by coupling thiols or selenols having functional mercapto or amino group. The functional groups (aldehyde, azide, amino or mercapto group) are exploited for biotin labeling which enables to separate hmC containing DNA (Song et al., 2010; Munzel et al., 2011; Pastor et al., 2011; Song et al., 2011, b).

CONCLUSIONS

1. 2-chloroacetaldehyde can be used for chemical detection of extahelical cytosines induced in DNA by DNA cytosine-5 methyltransferases or restriction endonucleases, leading to the formation of N^4,5- or N^4,N3-ethenocytosine adducts, respectively.

2. DNA cytosine-5 methyltransferases catalyze the reversible covalent addition of short aliphatic aldehydes to the 5^{th} position of target cytosine residues in DNA yielding 5-(1-hydroxyalkyl)cytosines and the conversion of 5-(1-hydroxyalkyl)cytosines to unmodified cytosines *in vitro*.

3. DNA cytosine-5 methyltransferases catalyze the condensation of exogenous alkylthiols and alkylselenols to 5-hydroxymethylcytosine residues in DNA yielding 5-chalcogenomethylcytosine *in vitro*. The reaction proceeds by an acid-assisted dehydration at the 5-hydroxymethyl group to give a highly electrophilic 5-methylene (or a bicyclic cytosine-5,6-thietane) intermediate, followed by a fast addition of an exogenous nucleophile whereby the relative strength of the nucleophile determines the nature of the final product.

4. The methyltransferase-directed coupling of alkylthiols and alkylselenols to 5-hydroxymethylcytosine in DNA can be used for selective covalent incorporation of functional and reporter groups permitting facile analysis of the 5-hydroxymethyl modification in genomic DNA.

LIST OF PUBLICATIONS

Journal publications

1. **Liutkevičiūtė, Z.**, Kriukienė, E., Grigaitytė, I., Masevičius, V. and Klimašauskas, S. (2011) Methyltransferase-directed derivatization of 5-hydroxymethylcytosine in DNA. Angew Chem Int Ed Engl. 50(9):2090-2093; Angew Chem 123: 2138-2141.
2. **Liutkevičiūtė, Z.**, Lukinavičius, G., Masevičius, V., Daujotytė, D. and Klimašauskas, S. (2009) Cytosine-5 methyltransferases add aldehydes to DNA. Nature Chem. Biol., 5(6): 400-402.
3. Daujotytė, D., **Liutkevičiūtė, Z.**, Tamulaitis, G. and Klimašauskas, S. (2008) Chemical mapping of cytosines enzymatically flipped out of the DNA helix. Nucleic Acids Res. 36(10): e57.

Patent applications

4. Klimašauskas, S., **Liutkevičiūtė, Z.** and Kriukienė, E. Derivatization of biomolecules by covalent coupling of non-cofactor compounds using methyltransferases. International patent application PCT/EP2010/054436.
5. Klimašauskas, S., **Liutkevičiūtė, Z.** and Kriukienė, E. Conversion of alpha-hydroxyalkylated residues in biomolecules using methyltransferases. International patent application PCT/EP2010/054437.

Book chapters

6. Klimašauskas, S. and **Liutkevičiūtė, Z.** (2009) Experimental approaches to study DNA base flipping. „DNA and RNA Modification Enzymes: Structure, Mechanism, Function and Evolution“. ed. H. Grosjean, Landes Bioscience: 19-32.
7. Klimašauskas, S., **Liutkevičiūtė, Z**. and Daujotytė, D. (2009) Biophysical approaches to study DNA base flipping. „Biophysics and the challenges of emerging threats“. ed. J. D. Puglisi, Springer Netherlands: 51-64.

CONFERENCE PRESENTATIONS

1. **Liutkevičiūtė, Z.**, Kriukienė, E., Grigaitytė, I., Vainorius, G., Masevičius, V. and Klimašauskas, S. Atypical reactions of DNA cytosine-5 methyltransferase. COST action TD09/05 meeting „Epigenetics – Bench to Bedside“, Croatia, Split. 2011.04.28-05.30.
2. **Liutkevičiūtė, Z.**, Kriukienė, E., Masevičius, V., Urbanavičiūtė, G. and Klimašauskas, S. Methyltransferase-directed sequence-specific coupling of exogenous electrophiles to DNA. 6th New England Biolabs Meeting on Restriction/Modification, Jacobs University, Bremen, Germany, 2010.09.01-06.
3. **Liutkevičiūtė, Z.**, Lukinavičius, G., Masevičius, V., Daujotytė, D. and Klimašauskas S. Methyltransferase-directed site-specific coupling of exogenous electrophiles to DNA. XI international meeting of Lithuania biochemical society "LBS 50", Molėtai, Lithuania, 2010.06.15-17.
4. **Liutkevičiūtė, Z.**, Daujotytė, D., Tamulaitis, G., Klimašauskas, S. Chemical identification of flipped out cytosines in enzyme-DNA complexes. Meeting/Conference of Lithuanian Biochemical society "Biochemistry and system biology”, Molėtai, Lithuania, 2008.06.20-22.

FINANCIAL SUPPORT

This work was in part supported by grants from the National Institutes of Health (USA), the Lithuanian Research Council and the Fp7-RegPot-2009 program (project MoBiLi).

CURRICULUM VITAE

Name Zita Liutkevičiūtė

Date of birth 1982 04 03

Work address Department of Biological DNA Modification Vilnius University Institute of Biotechnology
V. A. Graičiūno 8
Vilnius LT-02241
Lithuania

Phone +370 5 2602118
+370 650 49419

Fax +370 5 2602116

E-mails zita@ibt.lt
zitulis@gmail.com

Education and professional background

2004 B.Sc. Biology
Vilniaus University (Vilnius, Lietuva)

2006 M.Sc. Biochemistry
Vilniaus University (Vilnius, Lietuva)

2007-2011 PhD student of Biochemistry at
Vilniaus University Institute of Biotechnology (Vilnius, Lietuva)

Since 2011 Junior Scientist
at Vilniaus University Institute of Biotechnology, Department of Biological DNA Modification

ACKNOWLEDGEMENTS

I am especially grateful to my supervisor Saulius Klimašauskas for his huge input to this work: for helping to plan experiments, to interpret the results and to write thesis. For his patience while explaining reaction mechanisms of MTases and for encouragement to stay in scientific work five years ago.

I am grateful to all my colleagues who worked with me at Institute of Biotechnology for useful advices, discussions and goodwill. I thank Viktoras Masevičius for his valuable assistance in organic chemistry; Edita Kriukienė for the experiment of DNA labelling and enrichment; Gražvydas Lukinavičius for showing HPLC-MS peculiarities; Rūta Gerasimaitė for helping with the first TLC experiment; Giedrė Urbanavičiūtė for DNA labeling consultations and huge number of TLC plates which were removed from the tank early in the morning; Gintautas Tamulaitis for his help in working with restriction endonucleases and the samples of the enzymes; Giedrius Vilkaitis, Eglė Merkienė, Miglė Tomkuvienė, Lina Leinartaitė and Zdislav Staševskij for the samples of the proteins; my first supervisor in laboratory Dalia Daujotytė who trained me to work on my own, to analyse the results critically and never to give up; my students Indrė Grigaitytė and Gintautas Vainorius for their valuable experimental contribution and sunshine; Ervinas Gaidamauskas and Romanas Chaleckis for providing hundreds of articles.

I am grateful to my first biology teacher Birutė Diliautienė who directed me to the natural sciences. I thank my chemistry teacher Liuda Pivoriūnienė who in two months managed to prepare me for the chemistry exam.

I am grateful to all my friends of mountains, travels and studies for remarkable leisure, interest in this work, scientific and not scientific discussions.

Finally my warmest appreciation goes to my parents for the opportunity to spend my wonderful childhood very close to the nature and to be able to study the living world from the first days of my life. For the big and beautiful family and other things that are not easily expressed in words. I am grateful to my sisters, brother and their families who were always by my side, supported me and were interested in this work. Special thanks goes to my twin sister Edita for her patience while reading the thesis and searching for grammatical mistakes.

REZIUMĖ

Be keturių pagrindinių heterociklinių bazių eukariotinėje DNR yra randama penkta minorinė bazė – 5-metilcitozinas (mC) – epigenetinis žymuo, lemiantis chromatino aktyvumą (Goll ir Bestor, 2005). mC atsiranda dėl fermentų DNR citozino-5 metiltransferazių (MTazių), kurios perneša metilgrupę nuo kofaktoriaus S-adenozil-L-metionino ant citozino jau susintetintoje DNR. DNR metilinimas yra intensyviai tyrinėjamas jau daugelį metų, tačiau vienos jo pusės – demetilinimo – mechanizmas ir veikėjai nebuvo tiksliai nustayti (Wu ir Zhang, 2010). Demetilinimo tyrimai ypač suintensyvėjo prieš kelis metus, kai eukariotinėje DNR buvo surasta dar viena modifikuota bazė – 5-hidroksimetilcitozinas (hmC) (Kriaučionis ir Heintz, 2009; Tahiliani ir kt., 2009). Netrukus kamieninėse embrioninėse ląstelėse buvo identifikuoti 5-formilcitozinas (fC) ir 5-karboksicitozinas (caC) (He ir kt., 2011; Ito ir kt., 2011), tiesa, labai mažais kiekiais, bet galutinai sugriaunantys mitą apie DNR nekintamumą. Pastarosios trys modifikacijos (hmC, fC ir caC) genominėje DNR atsiranda iš mC, veikiant fermentams oksogenazėms Tet1,2,3 (Tahiliani ir kt., 2009). Parodyta, kad šios modifikacijos yra tarpiniai mC demetilinimo produktai. Iš šios trijulės reikia išskirti hmC, kuris yra stabili modifikacija, tam tikrose vystymosi stadijose ir kai kuriuose audiniuose randama dideliais kiekiais (Li ir Liu, 2011; Munzel ir kt., 2011). Tiksli hmC reikšmė epigenetikoje ir DNR demetilinime bus atskleista ateityje, nes šiuo metu trūksta metodų, galinčių nustatyti hidroksimetilintas vietas DNR (Nabel ir kt., 2011).

Mūsų tyrimai prasidėjo nuo modelinės prokariotinės MTazės M.HhaI, kuri yra intensyviai tyrinėjama jau 20 metų. Išsprendus trinario komplekso M.HhaI-DNR-SAH struktūrą pirmą kartą buvo parodytas elegantiškas DNR deformacijos būdas: taikinio citozinas yra išsukamas 180° kampu iš dvigrandės DNR į aktyvų fermento centrą (Klimašauskas ir kt., 1994). Vėliau buvo nustatyta, kad daugelis su DNR sąveikaujančių baltymų, panašiai kaip ir M.HhaI, išsuka bazes iš DNR spiralės (Cheng ir Blumenthal, 1996). Šie atradimai paskatino kurti naujus metodus, įgalinančius tirti bazių išsukimą. Nusprendėme, kad aprašytos 2-chloracetaldehido (Caa) reakcijos su nesuporuotomis bazėmis galėtų pasitarnauti ir baltymų-DNR tyrimuose, paprastai ir greitai identifikuojant fermentų išsuktus citozinus (Kohwi-Shigematsu ir Kohwi, 1992). Eksperimentai parodė, kad šiuo metodu gali būti tyrinėjamos tiek MTazės, tiek restrikcijos endonukleazės (Daujotytė ir kt., 2008). Detalesni susidariusių produktų tyrimai po Caa reakcijos atskleidė, kad MTazių išsukamas citozinas yra modifikuojamas netradiciškai. Išbandžius keletą elektrofilų paaiškėjo, kad MTazės geba prijungti trumpus alifatinius aldehidus prie taikinio citozino-5 atomo, susidarant 5-(1-hidroksialkil)citozinams (Liutkevičiūtė ir kt., 2009). Pažymėtina, kad MTazės dideliu efektyvumu prijungia formaldehidą, susidarant hmC – šeštai DNR bazei, neseniai surastai žinduolių DNR. MTazės taip pat geba katalizuoti grįžtamą reakciją: hmC virtimą į citoziną. Dar viena atrasta netipinė citozino-5 MTazių reakcija – taikinyje esančio hmC modifikacija alkiltioliais ir alkilselenoliais, susidarant 5-chalkogenometilcitozinams. Pastaroji reakcija gali būti panaudota praktiškai genominės DNR tyrimuose, nustatant hmC pozicijas genome (Liutkevičiūtė ir kt., 2011). Atrastos netipinės DNR citozino-C5 metiltransferazių reakcijos rodo, kad tam tikromis sąlygomis fermentai gali katalizuoti jiems nebūdingas reakcijas. Šiame etape sunku įvertinti, ar panašios reakcijos galėtų vykti ir *in vivo*, tačiau MTazių svarba, modifikuojant DNR gyvuosiuose organizmuose (ne tik metilinant, bet ir hidroksimetilinant bei demetilinant), turėtų būti atidžiai peržiūrėta.

REFERENCES

1. Alegria, A. H. (1967). "Hydroxymethylation of pyrimidine mononucleotides with formaldehyde." Biochim. Biophys. Acta **149**: 317-324.
2. Barrett, J. E. and Maltby (1998). "Trapping of C5 methylene intermediate in thymidylate synthase." J. Am. Chem. Soc. **120**: 449-450.
3. Beland, F. A., N. F. Fullerton and R. H. Heflich (1984). "Rapid isolation, hydrolysis and chromatography of formaldehyde-modified DNA." J Chromatogr **308**: 121-31.
4. Bischerour J. and R. Chalmers (2007). "Base-flipping dynamics in a DNA hairpin processing reaction". Nucleic Acids Res. **35**: 2584-2595.
5. Bochtler, M., R. H. Szczepanowski, G. Tamulaitis, S. Gražulis, H. Czapinska, E. Manakova and V. Šikšnys (2006). "Nukleotide flips determine the specificity of the Ecl18kl restriction endonuklease." EMBO J. **25**(10): 2219-2229.
6. Cheng, X. and R. M. Blumenthal (1996). "Finding a basis for flipping bases." Structure **4**: 639-645.
7. Cline, R. E., R. M. Fink and K. Fink (1959). "Synthesis of 5-substituted pyrimidines via formaldehyde addition." J. Am. Chem. Soc. **81**: 2521-2527.
8. Cortellino, S., J. F. Xu, M. Sannai, R. Moore, E. Caretti, A. Cigliano, M. Le Coz, K. Devarajan, A. Wessels, D. Soprano, L. K. Abramowitz, M. S. Bartolomei, F. Rambow, M. R. Bassi, T. Bruno, M. Fanciulli, C. Renner, A. J. Klein-Szanto, Y. Matsumoto, D. Kobi, I. Davidson, C. Alberti, L. Larue and A. Bellacosa (2011). "Thymine DNA Glycosylase Is Essential for Active DNA Demethylation by Linked Deamination-Base Excision Repair." Cell **146**(1): 67-79.
9. Dalhoff, C., G. Lukinavičius, S. Klimašauskas and E. Weinhold (2006). "Direct transfer of extended groups from synthetic cofactors by DNA methyltransferases." Nat Chem Biol **2**(1): 31-2.
10. Daujotytė, D., Z. Liutkevičiūtė, G. Tamulaitis and S. Klimašauskas (2008). "Chemical mapping of cytosines enzymatically flipped out of the DNA helix." Nucleic Acids Res. 36(10): e57.
11. Daujotytė, D., S. Serva, G. Vilkaitis, E. Merkienė, C. Venclovas and S. Klimašauskas (2004). "HhaI DNA methyltransferase uses the protruding Gln237 for active flipping of its target cytosine." Structure **12**: 1047-1055.
12. Davis, T. and R. Vaisvila (2011). "High sensitivity 5-hydroxymethylcytosine detection in Balb/C brain tissue." J Vis Exp (48).
13. Ficz, G., M. R. Branco, S. Seisenberger, F. Santos, F. Krueger, T. A. Hore, C. J. Marques, S. Andrews and W. Reik (2011). "Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation." Nature **473**(7347): 398-402.
14. Flusberg, B. A., D. R. Webster, J. H. Lee, K. J. Travers, E. C. Olivares, T. A. Clark, J. Korlach and S. W. Turner (2010). "Direct detection of DNA methylation during single-molecule, real-time sequencing." Nat Methods **7**(6): 461-5..
15. Ghosh, I., C. I. Stains, A. T. Ooi and D. J. Segal (2006). "Direct detection of double-stranded DNA: Molecular methods and applications for DNA diagnostics." Mol Biosyst **2**(11): 551-60.
16. Globisch, D., M. Munzel, M. Muller, S. Michalakis, M. Wagner, S. Koch, T. Bruckl, M. Biel and T. Carell (2010). "Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates." PLoS One **5**(12): e15367.
17. Goll, M. G. and T. H. Bestor (2005). "Eukaryotic cytosine methyltransferases." Annu. Rev. Biochem. **74**: 481-514.
18. Graves, K. L., M. M. Butler and L. W. Hardy (1992). "Roles of Cys148 and Asp179 in catalysis by deoxycytidylate hydroxymethylase from bacteriophage T4 examined by site-directed mutagenesis." Biochemistry **31**(42): 10315-21.
19. Guo, J. U., Y. J. Su, C. Zhong, G. L. Ming and H. J. Song (2011). "Hydroxylation of 5-Methylcytosine by TET1 Promotes Active DNA Demethylation in the Adult Brain." Cell **145**(3): 423-434.
20. He, Y. F., B. Z. Li, Z. Li, P. Liu, Y. Wang, Q. Y. Tang, J. P. Ding, Y. Y. Jia, Z. C. Chen, L. Li, Y. Sun, X. X. Li, Q. Dai, C. X. Song, K. L. Zhang, C. He and G. L. Xu (2011). "Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA." Science **333**(6047): 1303-1307.
21. Heck, H. and M. Casanova (2004). "The implausibility of leukemia induction by formaldehyde: a critical review of the biological evidence on distant-site toxicity." Regul Toxicol Pharmacol **40**(2): 92-106.
22. Holz, B., S. Klimašauskas, S. Serva and E. Weinhold (1998). "2-Amino purine as a fluorescence probe for DNA base flipping by methyltransferases." Nucleic Acids Res. **26**(4): 1076-1083.
23. Hotchkiss, R. (1948). "The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography." J Biol Chem **175**(1): 315-32.

24. Huang, Y., W. A. Pastor, Y. Shen, M. Tahiliani, D. R. Liu and A. Rao (2010). "The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing." PLoS One **5**(1): e8888.
25. Ito, S., L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He and Y. Zhang (2011). "Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine." Science **333**(6047): 1300-1303.
26. Jin, S. G., S. Kadam and G. P. Pfeifer (2010). "Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine." Nucleic Acids Res **38**(11): e125.
27. Jin, S. G., X. Wu, A. X. Li and G. P. Pfeifer (2011). "Genomic mapping of 5-hydroxymethylcytosine in the human brain." Nucleic Acids Res.
28. Juskowiak, B. (2010). "Nucleic acid-based fluorescent probes and their analytical potential." Anal Bioanal Chem. **399**(9):3157-76.
29. Klimašauskas, S., S. Kumar, R. J. Roberts and X. Cheng (1994). "Hhal methyltransferase flips its target base out of the DNA helix." Cell **76**: 357-369.
30. Kohwi-Shigematsu, K. and Y. Kohwi (1992). Detection of non-B-DNA structures at specific sites in supercoiled plasmid DNA and chromatin with haloacetaldehyde and diethyl pyrocarbonate. Methods Enzymology. D. M. J. Lilley and J. E. Dahlberg. San Diego, Academic Press, INC. **212:** 155-180.
31. Kriaučionis, S. and N. Heintz (2009). "The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain." Science **324**(5929): 929-30.
32. Krzyzosiak, W. J., J. Biernat, J. Ciesiolka, K. Gulewicz and M. Wiewiorowski (1981). "The reactions of adenine and cytosine residues in tRNA with chloroacetaldehyde." Nucleic Acids Res. **9**(12): 2841-2851.
33. Kuchino Y., N. Hanyu and S. Nishimura (1987). "Analysis of modified nucleosides and nucleotide sequence of tRNA." Methods Enzymol. **155**:379-96.
34. Kusmierek, J. T. and B. Singer (1982). "Chloroacetaldehyde-treated ribo- and deoxyribopolynucleotides. 1. Reaction products." Biochemistry **21**(22): 5717-5722.
35. LaFrancois, C. J., J. Fujimoto and L. C. Sowers (1998). "Synthesis and characterization of isotopically enriched pyrimidine deoxynucleoside oxidation damage products." Chem Res Toxicol **11**(1): 75-83.
36. Lau, E. Y. and T. C. Bruice (1999). "Active site dynamics of the HhaI methyltransferase: insights from computer simulation." J. Mol. Biol. **293**: 9-18.
37. Liutkevičiūtė, Z., E. Kriukiene, I. Grigaityte, V. Masevicius and S. Klimašauskas (2011). "Methyltransferase-Directed Derivatization of 5-Hydroxymethylcytosine in DNA." Angewandte Chemie-International Edition **50**(9): 2090-2093.
38. Liutkevičiūtė, Z., G. Lukinavičius, V. Masevicius, D. Daujotytė and S. Klimašauskas (2009). "Cytosine-5-methyltransferases add aldehydes to DNA." Nat Chem Biol **5**(6): 400-2.
39. Lukinavičius, G., V. Lapiene, Z. Stasevskij, C. Dalhoff, E. Weinhold and S. Klimašauskas (2007). "Targeted labeling of DNA by methyltransferase-directed transfer of activated groups (mTAG)." J Am Chem Soc **129**(10): 2758-9.
40. Maiti, A. and A. Drohat (2011). "Thymine DNA Glycosylase Can Rapidly Excise 5-Formylcytosine and 5-Carboxylcytosine: POTENTIAL IMPLICATIONS FOR ACTIVE DEMETHYLATION OF CpG SITES." J. Biol. Chem. **286**(41): 35334-8.
41. Metivier, R., R. Gallais, C. Tiffoche, C. Le Peron, R. Z. Jurkowska, R. P. Carmouche, D. Ibberson, P. Barath, F. Demay, G. Reid, V. Benes, A. Jeltsch, F. Gannon and G. Salbert (2008). "Cyclical DNA methylation of a transcriptionally active promoter." Nature **452**(7183): 45-U2.
42. Mi, S. and R. J. Roberts (1993). "The DNA binding affinity of HhaI methylase is increased by a single amino acid substitution in the catalytic center." Nucleic Acids Res. **21**: 2459-2464.
43. Nabel, C., S. Manning and R. Kohli (2011). "The Curious Chemical Biology of Cytosine: Deamination, Methylation and Oxidation as Modulators of Genomic Potential." ACS Chem. Biol.
44. Neely, R. K., D. Daujotytė, S. Gražulis, S. W. Magennis, D. T. Dryden, S. Klimašauskas and A. C. Jones (2005). "Time-resolved fluorescence of 2-aminopurine as a probe of base flipping in M.HhaI-DNA complexes." Nucleic Acids Res. **33**(22): 6953-60.
45. Nestor, C., A. Ruzov, R. Meehan and D. Dunican (2010). "Enzymatic approaches and bisulfite sequencing cannot distinguish between 5-methylcytosine and 5-hydroxymethylcytosine in DNA." Biotechniques **48**(4): 317-9.
46. Okamoto, A., K. Sugizaki, A. Nakamura, H. Yanagisawa and S. Ikeda (2011). "5-Hydroxymethylcytosine-selective oxidation with peroxotungstate." Chem. Commun. (Camb). **47**(40): 11231-3.
47. Pastor, W. A., U. J. Pape, Y. Huang, H. R. Henderson, R. Lister, M. Ko, E. M. McLoughlin, Y. Brudno, S. Mahapatra, P. Kapranov, M. Tahiliani, G. Q. Daley, X. S. Liu, J. R. Ecker, P. M. Milos, S. Agarwal and A. Rao (2011). "Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells." Nature **473**(7347): 394-7.

48. Pfaffeneder, T., B. Hackner, M. Truss, M. Munzel, M. Muller, C. A. Deiml, C. Hagemeier and T. Carell (2011). "The Discovery of 5-Formylcytosine in Embryonic Stem Cell DNA." Angewandte Chemie-International Edition **50**(31): 7008-7012.
49. Pignot, M., G. Pljevaljcic and E. Weinhold (2000). "Efficient synthesis of S-adenosyl-L-homocysteine natural product analogues and their use to elucidate the structural determinant for cofactor binding of the DNA methyltransferase M.HhaI." Eur. J. Org. Chem: 549-555.
50. Pingoud V., C. Conzelmann, S. Kinzebach, A. Sudina, V. Metelev, E. Kubareva, J. M. Bujnicki, R. Lurz, G. Luder, S. Y. Xu and A. Pingoud. (2003). "PspGI, a type II restriction endonuclease from the extreme thermophile Pyrococcus sp.: structural and functional studies to investigate an evolutionary relationship with several mesophilic restriction enzymes." J. Mol. Biol. **329**:913-29.
51. Privat, E. and L. C. Sowers (1996). "Photochemical Deamination and Demethylation of 5-Methylcytosine." Chem. Res. Toxicol. **9**(4): 745-750.
52. Razin, A. and A. Riggs (1980). "DNA Methylation and Gene Function." Science **210**(4470): 604-10.
53. Robertson, A. B., J. A. Dahl, C. B. Vagbo, P. Tripathi, H. E. Krokan and A. Klungland (2011). "A novel method for the efficient and selective identification of 5-hydroxymethylcytosine in genomic DNA." Nucleic Acids Res.
54. Sambrook, J. and D. W. Russell (2001). Molecular Cloning. Cold Spring Harbor, New York, Cold Spring Harbor Laboratory Press.
55. Sedgwick, B., P. A. Bates, J. Paik, S. C. Jacobs and T. Lindahl (2007). "Repair of alkylated DNA: recent advances." DNA Repair (Amst) **6**(4): 429-42.
56. Serva, S., E. Weinhold, R. J. Roberts and S. Klimašauskas (1998). "Chemical display of thymine residues flipped out by DNA methyltransferases." Nucleic Acids Res. **26**(15): 3473-3479.
57. Shiau, G. T., R. F. Schinazi, M. S. Chen and W. H. Prusoff (1980). "Synthesis and biological activities of 5-(hydroxymethyl, azidomethyl, or aminomethyl)-2'-deoxyuridine and related 5'-substituted analogues." J Med Chem **23**(2): 127-33.
58. Shieh, F. K., B. Youngblood and N. O. Reich (2006). "The role of Arg165 towards base flipping, base stabilization and catalysis in M.HhaI." J Mol Biol **362**(3): 516-27.
59. Shieh, F. K. and N. O. Reich (2007). "AdoMet-dependent methyl-transfer: Glu119 is essential for DNA C5-cytosine methyltransferase M.HhaI." J Mol Biol **373**(5): 1157-68.
60. Shock, L. S., P. V. Thakkar, E. J. Peterson, R. G. Moran and S. M. Taylor (2011). "DNA methyltransferase 1, cytosine methylation, and cytosine hydroxymethylation in mammalian mitochondria." Proc Natl Acad Sci U S A.
61. Sokolowska, M., M. Kaus-Drobek, H. Czapinska, G. Tamulaitis, R. H. Szczepanowski, C. Urbanke, V. Šikšnys and M. Bochtler (2007). "Monomeric restriction endonuclease BcnI in the apo form and in an asymmetric complex with target DNA." J. Mol Biol. **369**(3): 722-34.
62. Solomon, M. J. and A. Varshavsky (1985). "Formaldehyde-mediated DNA-protein crosslinking: a probe for in vivo chromatin structures." Proc Natl Acad Sci U S A **82**(19): 6470-4.
63. a) Song, C. X., T. A. Clark, X. Y. Lu, A. Kislyuk, Q. Dai, S. W. Turner, C. He and J. Korlach (2011). "Sensitive and specific single-molecule sequencing of 5-hydroxymethylcytosine." Nat Methods.
64. b) Song, C. X., Y. Sun, Q. Dai, X. Y. Lu, M. Yu, C. G. Yang and C. He (2011). "Detection of 5-Hydroxymethylcytosine in DNA by Transferring a Keto-Glucose by Using T4 Phage beta-Glucosyltransferase." Chembiochem **12**(11): 1682-1685.
65. Song, C. X., K. E. Szulwach, Y. Fu, Q. Dai, C. Yi, X. Li, Y. Li, C. H. Chen, W. Zhang, X. Jian, J. Wang, L. Zhang, T. J. Looney, B. Zhang, L. A. Godley, L. M. Hicks, B. T. Lahn, P. Jin and C. He (2010). "Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine." Nat Biotechnol **29**(1): 68-72.
66. Szczepanowski R. H., M. A. Carpenter, H. Czapinska, M. Zaremba, G. Tamulaitis, V. Šikšnys, A. S. Bhagwat and M. Bochtler (2008). "Central base pair flipping and discrimination by PspGI". Nucleic Acids Res. **36**:6109-17
67. Szwagierczak, A., A. Brachmann, C. S. Schmidt, S. Bultmann, H. Leonhardt and F. Spada (2011). "Characterization of PvuRts1I endonuclease as a tool to investigate genomic 5-hydroxymethylcytosine." Nucleic Acids Res.
68. Tahiliani, M., 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 (2009). "Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1." Science **324**(5929): 930-5.
69. Tamulaitis G., A. S. Solonin and V. Šikšnys (2002). "Alternative arrangements of catalytic residues at the active sites of restriction enzymes". FEBS Lett. **518**: 17-22.
70. Tamulaitis, G., M. Zaremba, R. H. Szczepanowski, M. Bochtler and V. Šikšnys (2008). "How PspGI, catalytic domain of EcoRII and Ecl18kI acquire specificities for different DNA targets." Nucleic Acids Res. **36**:6101-8

71. Tamulaitis, G., M. Zaremba, R. H. Szczepanowski, M. Bochtler and V. Šikšnys (2007). "Nucleotide flipping by restriction enzymes analyzed by 2-aminopurine steady-state fluorescence." Nucleic Acids Res. **35**(14): 4792-9.
72. Tardy-Planechaud, S., J. Fujimoto, S. S. Lin and L. C. Sowers (1997). "Solid phase synthesis and restriction endonuclease cleavage of oligodeoxynucleotides containing 5-(hydroxymethyl)-cytosine." Nucleic Acids Res **25**(3): 553-9.
73. Vilkaitis, G., A. Dong, E. Weinhold, X. Cheng and S. Klimašauskas (2000). "Functional roles of the conserved threonine 250 in the target recognition domain of HhaI DNA methyltransferase." J. Biol. Chem. **275**(49): 38722-38730.
74. Vilkaitis, G., I. Suetake, S. Klimašauskas and S. Tajima (2005). "Processive methylation of hemimethylated CpG sites by mouse Dnmt1 DNA methyltransferase." J. Biol. Chem. **280**: 64-72.
75. Wallace, E. V., D. Stoddart, A. J. Heron, E. Mikhailova, G. Maglia, T. J. Donohoe and H. Bayley (2010). "Identification of epigenetic DNA modifications with a protein nanopore." Chem Commun (Camb) **46**(43): 8195-7.
76. Wang, H., S. Guan, A. Quimby, D. Cohen-Karni, S. Pradhan, G. Wilson, R. Roberts, Z. Zhu and Y. Zheng (2011). "Comparative characterization of the PvuRts1I family of restriction enzymes and their application in mapping genomic 5-hydroxymethylcytosine." Nucl. Acids Res.: 1-12.
77. Wang, M., G. Cheng, P. W. Villalta and S. S. Hecht (2007). "Development of liquid chromatography electrospray ionization tandem mass spectrometry methods for analysis of DNA adducts of formaldehyde and their application to rats treated with N-nitrosodimethylamine or 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone." Chem Res Toxicol **20**(8): 1141-8.
78. Wanunu, M., D. Cohen-Karni, R. R. Johnson, L. Fields, J. Benner, N. Peterman, Y. Zheng, M. L. Klein and M. Drndic (2010). "Discrimination of Methylcytosine from Hydroxymethylcytosine in DNA Molecules." J Am Chem Soc.
79. Watson, J. and F. Crick (1953). "Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid." Nature **171**(4356): 737-8.
80. Woo, J., R. B. Meyer, Jr. and H. B. Gamper (1996). "G/C-modified oligodeoxynucleotides with selective complementarity: synthesis and hybridization properties." Nucleic Acids Res **24**(13): 2470-5.
81. Wu, J. C. and D. V. Santi (1987). "Kinetic and catalytic mechanism of HhaI methyltransferase." J. Biol.Chem. **262**(10): 4778-4786.
82. Wu, S. C. and Y. Zhang (2010). "Active DNA demethylation: many roads lead to Rome." Nat Rev Mol Cell Biol **11**(9): 607-20.
83. Zemach, A., I. E. McDaniel, P. Silva and D. Zilberman (2010). "Genome-wide evolutionary analysis of eukaryotic DNA methylation." Science **328**(5980): 916-9.
84. Zheng, Y., D. Cohen-Karni, D. Xu, H. G. Chin, G. Wilson, S. Pradhan and R. J. Roberts (2010). "A unique family of Mrr-like modification-dependent restriction endonucleases." Nucleic Acids Res **38**(16): 5527-34.