VILNIUS UNIVERSITY
INSTITUTE OF BIOTECHNOLOGY

Rūta Gerasimaitė

A DIRECTED EVOLUTION DESIGN OF TARGET SPECIFICITY AND KINETIC ANALYSIS OF CONFORMATIONAL TRANSITIONS IN THE HhaI METHYLTRANSFERASE

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

Vilnius, 2011

This study has been carried out at the Institute of Biotechnology during 2003-2009.
Dissertation is maintained by extern.

Scientific consultant

Prof. habil. dr. **Saulius Klimašauskas** (Institute of Biotechnology, Vilnius University; physical sciences, biochemistry – 04 P)

Evaluation board of dissertation of Biochemistry trend

Chairman:

Prof. habil. dr. **Vida Mildažienė** (Vytautas Magnus University; physical sciences, biochemistry – 04 P)

Members:

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

Dr. **Daumantas Matulis** (Institute of Biotechnology, Vilnius University; physical sciences, biochemistry – 04 P)

Dr. **Rolandas Meškys** (Institute of Biochemistry, Vilnius University; physical sciences, biochemistry – 04 P)

Dr. **Gintautas Tamulaitis** (Institute of Biotechnology, Vilnius University; physical sciences, biochemistry – 04 P)

Official opponents:

Habil. dr. **Narimantas Čėnas** (Institute of Biochemistry, Vilnius University; physical sciences, biochemistry – 04 P)

Dr. **Arūnas Lagunavičius** (Thermo Fisher Scientific; physical sciences, biochemistry – 04 P)

The thesis defence will take place at 11 a.m. on 2nd of May, 2011, at the Institute of Biotechnology.
Address: V. A. Graičiūno 8, Vilnius LT-02241, Lithuania

The abstract of thesis was distributed on 2nd of April, 2011.

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

VILNIAUS UNIVERSITETAS
BIOTECHNOLOGIJOS INSTITUTAS

Rūta Gerasimaitė

DNR METILTRANSFERAZĖS HhaI ATPAŽĮSTAMOS SEKOS INŽINERIJA IR KONFORMACINIŲ VIRSMŲ KINETIKOS TYRIMAS

Daktaro disertacijos santrauka
Fiziniai mokslai, biochemija (04 P)

Vilnius, 2011

Disertacija rengta 2003–2009 metais Biotechnologijos Institute.
Disertacija ginama eksternu.

Mokslinis konsultantas:

Prof. habil. dr. **Saulius Klimašauskas** (Biotechnologijos Institutas; fiziniai mokslai, biochemija – 04 P)

Disertacija ginama Vilniaus universiteto Biochemijos mokslo krypties taryboje.

Pirmininkė:

Prof. habil. dr. **Vida Mildažienė** (Kauno Vytauto Didžiojo universitetas; fiziniai mokslai, biochemija – 04 P)

Nariai:

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

Dr. **Daumantas Matulis** (Vilniaus universitetas, Biotechnologijos institutas; fiziniai mokslai, biochemija – 04 P)

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

Dr. **Gintautas Tamulaitis** (Vilniaus universitetas, Biotechnologijos institutas; fiziniai mokslai, biochemija – 04 P)

Oponentai:

Habil. dr. **Narimantas Čėnas** (Vilniaus universitetas, Biochemijos institutas; fiziniai mokslai, biochemija – 04 P)

Dr. **Arūnas Lagunavičius** (Thermo Fisher Scientific; fiziniai mokslai, biochemija – 04 P)

Disertacija bus ginama viešame Biochemijos mokslo krypties tarybos posėdyje 2011 m. gegužės mėn. 2 d. 11 val. Biotechnologijos instituto aktų salėje.

Adresas: Graičiūno 8, Vilnius, LT–02241, Lietuva.

Disertacijos santrauka išsiuntinėta 2011 m. balandžio mėn. 2 d.

Disertaciją galima peržiūrėti Biotechnologijos instituto ir Vilniaus universiteto bibliotekose.

CONTENTS

INTRODUCTION

Enzymological characterization of catalytic mechanisms of single enzymes is important even in the post-genomic era, when an increasing number of studies rely on high-throughput methods. The detailed knowledge of the catalytic mechanisms facilitates the development of new drugs, discovery of new drug targets and help to engineer enzymes with the improved properties to be used in biotechnological applications. In addition to merely practical goals, the investigation of the catalytic mechanism bears a wider significance. Deciphering mechanism of a new enzyme (or enzyme family) contributes to creating new techniques and improving the existing ones, thereby facilitating investigation of yet uncharacterized proteins.

DNA cytosine-C5 methyltransferases (C5-MTases) have been intensively studied during several decades. DNA C5-MTases catalyze the transfer of the methyl group from the cofactor *S*-adenosyl-L-methionine (AdoMet) to their target nucleotides within 2–8 bp sequences in DNA (Dryden, 1999). In most eukaryotes 5-methylcytosine serves as epigenetic mark, thereby expanding the content of the genome-encoded information. Although the molecular mechanisms are not always clear, the changes in DNA methylation levels are related to cancer, schizophrenia, diabetes, as well as to other complex diseases and genetic syndromes in humans (Cedar and Bergman, 2009; Pogribny and Beland, 2009; Cheng and Blumenthal, 2010). A mechanism-based inhibitor of DNA C5-MTase, 5-azacytosine, has been recently approved for treatment of mielodisplastic syndrome (Kaminskas *et al.*, 2005).

Despite the utmost importance, the investigation of molecular mechanisms of eukaryotic DNA C5-MTases is hampered by their large size and complex regulation of their activity by interactions with other cellular proteins (Goll and Bestor, 2005; Cheng and Blumenthal, 2008). Therefore most mechanistic studies of C5-methylation are performed using smaller and less complex prokaryotic enzymes as model systems. Prokaryotic DNA C5-MTases share high sequence and structure similarities among themselves, as well as with the catalytic domain of eukaryotic DNA MTases. Due to their high structural similarity, eukaryotic and prokaryotic DNA C5-MTases are believed to share many aspects of the catalytic mechanism. Several years ago the analogs of cofactor AdoMet were synthesized, allowing the MTase-catalyzed transfer of the functional groups onto a base in the target sequence. This allows site-specific labeling of DNA molecules and paves a way for prokaryotic DNA MTases to become important tools in modern molecular biology (Klimašauskas and Weinhold, 2007).

The paradigm model for structural and mechanistic studies of cytosine C5-methylation is the HhaI MTase (M.HhaI), a component of a type II restriction-modification system from *Haemophylus haemolyticus* (Sankpal and Rao, 2002). M.HhaI recognizes the tetranucleotide sequence G<u>C</u>GC and methylates the inner cytosine residue (underlined). Crystallographic studies revealed that M.HhaI folds into two domains. The larger (catalytic) domain contains most of the highly conserved residues and forms the cofactor binding pocked and the active site (Cheng *et al.*, 1993). The smaller (target recognition) domain is formed by the highly variable protein region and by making contacts to the DNA bases is responsible for the recognition of the specific target (Klimašauskas *et al.*, 1994). Thus the catalytic and DNA recognition functions are largely segregated in two distinct domains. Furthermore, the target recognition domain (TRD) itself has an in-built modularity. Crystal structures of M.HhaI (Klimašauskas *et*

al., 1994) and M.HaeIII (Reinisch *et al.*, 1995) show that these enzymes recognize their targets via two distinct recognition loops (Loop-1 and Loop-2) which form multiple base-specific contacts with two distinct segments of their target sequences. These features together with a great variety of target specificities make prokaryotic DNA C5-MTases an attractive system for studying mechanisms of specific DNA recognition.

DNA binds in a cleft of the bilobal protein leading to dramatic conformational changes in both the DNA and the enzyme itself (Cheng *et al.*, 1993; Klimašauskas *et al.*, 1994). A MTase-mediated rotation of the target nucleotide out of the DNA helix (base flipping) serves to deliver the base into a concave catalytic site of the enzyme. Base flipping is accompanied by a massive movement of the catalytic loop (residues 81-100) towards the DNA which locks the flipped-out base for methylation. The enzymatic catalysis involves the transient formation of a Michael adduct between the catalytic cysteine (Cys81) and the C6 of the cytosine, permitting a direct S_N2 transfer of the methyl group onto C5. Although the rates of methyl transfer (Vilkaitis *et al.*, 2001), cofactor binding and final product release (Merkienė and Klimašauskas, 2005) are known, the order and kinetics of initial key steps such as target cytosine flipping, loop closure, covalent bond formation have not yet been established. The elucidation of these reaction steps have been impeded by lack of the adequate methods. For example, kinetics of the target base flipping was investigated using DNA substrates with a fluorescent base 2-aminopurine (P) in the target position (Vilkaitis *et al.*, 2000). However P:G is a weak base pair and the mode of 2-aminopurine binding in the active site differs from that of cytosine (Neely *et al.*, 2005). Therefore it is unclear whether mechanistic details derived from these studies are applicable to M.HhaI reaction with the cognate DNA substrate.

During this work two aspects of the M.HhaI reaction have been investigated. First, taking advantage of vast structural and biochemical information available to date, the target specificity of M.HhaI was changed from G<u>C</u>GC to G<u>C</u>G, thereby creating the MTase of a unique specificity. The newly constructed MTase may become a useful tool for studying the mechanisms of eukaryotic maintenance DNA methyltransferases and for site-specific DNA labeling. This study also provided new insights into the mechanism of specific DNA recognition.

Second, the kinetics of yet poorly characterized pre-catalytic conformational transitions in the MTase and DNA has been investigated in great detail. To this end, a new method to follow the target cytosine flipping and its subsequent covalent activation has been introduced, which allows a direct real-time observation of these processes by monitoring associated UV absorbance changes.

<u>Specific aims</u>

1. Experimental validation of a new strategy for the construction of catalytically efficient DNA cytosine-5 methyltransferases with novel target specificities by functional elimination of one of the two target recognition elements:
    a. functional deletion the target recognition Loop-2 by rational protein design and directed evolution approaches, in order to alter the target specificity of M.HhaI (G<u>C</u>GC to G<u>C</u>G);
    b. characterization of the catalytic properties and target specificity of the constructed MTases.

c. determination of the impact of the introduced changes in the protein on its interactions with DNA, cofactor and synthetic cofactor analogs.

2. Development of a method for a direct real-time observation of the methyltransferase-induced flipping of the target cytosine out of DNA helix and its covalent activation in catalytic centre by monitoring hyperchromicity and other associated UV absorbance changes.
3. Kinetic characterization of the target cytosine flipping and its covalent activation in relation to the catalytic loop closure and the chemical methyl transfer during the catalytic cycle of M.HhaI.

Scientific novelty

The present work for the first time provides an experimental demonstration of a design strategy for altering the target specificity of DNA C5-MTases by functional elimination of one target recognition loop in an enzyme. This provides an illustration to the general principle of DNA-enzyme interaction that impaired catalytic efficiency due to the loss of specific contacts to the target nucleobases can be recovered by enhancing non-specific protein contacts to the DNA backbone.

This work for the first time describes a method for direct real-time observation of the cytosine flipping and its covalent activation in a chemically unperturbed DNA by monitoring associated UV absorbance changes. It is the first to show that the flipping of the target cytosine and the closure of the catalytic loop in M.HhaI proceed simultaneously. The rate of these processes (~10-20 s^{-1}) is strongly dependent on the stability of the target base-pair. The rate of the covalent activation of the target cytosine by M.HhaI (0.3-0.6 s^{-1}) has been determined, which is closely followed by, but is not coincident, with the methyl transfer (0.1-0.2 s^{-1}).

Practical value

Due to a non-palindromic nature of the GCG target, the engineered MTase can be used for creating hemimethylated CpG sites in large natural DNA molecules, which are unique substrates for mechanistic studies of the maintenance DNA methylation in higher eukaryotes. The GCG-specific MTase is also compatible with mTAG (methyltransferase-directed transfer of activated groups) technology and can be used for site-specific labeling of DNA with reporter groups.

The present work also describes the first method that allows a direct real-time observation of enzymatically induced cytosine flipping and its covalent activation in a chemically unperturbed system. Since this new approach is based on the general phenomenon of hyperchromicity, it is thus applicable to many other base-flipping systems.

Findings presented for defense

1. A novel strategy has been demonstrated (proof of principle) for the construction of catalytically efficient and sequence-selective DNA C5-MTases via functional elimination of one of the two target recognition elements.
2. Enhancing of non-specific protein contacts to the DNA backbone can compensate for the loss of base-specific DNA interactions incurred upon removal of a recognition loop.

3. The catalytic efficiency and sequence fidelity of the engineered GCG-specific hemimethylase are comparable to those of other wild-type MTases and are thus sufficient for its use as a tool in molecular biology.
4. The GCG-specific MTase catalyzes efficient transfer of extended groups from synthetic AdoMet analogs onto DNA and therefore can be used for site-specific labeling or hemimethylation of natural DNA.
5. Flipping of the target cytosine out of DNA double helix and its covalent activation in the active site of M.HhaI can be directly observed by monitoring small changes in UV absorbance in chemically unperturbed (natural) DNA substrates. Since the hyperchromicity is a general feature of double helical nucleic acids, this method is applicable to many other base-flipping systems.
6. Flipping of the target cytosine and the closure of the catalytic loop in M.HhaI occur simultaneously; the rate of this concerted motion is highly dependent on the strength of the target base-pair.
7. The covalent activation of the target cytosine and the transfer of the methyl group are temporally distinct steps in the catalytic cycle of M.HhaI.

MATERIALS AND METHODS

Materials

All restriction enzymes, DNA polymerases, exonuclease I, proteinase K, shrimp alkaline phosphatase, λ DNA (*dam⁻, dcm⁻*), pBR322 DNA, pUC19 DNA and kits for molecular biology were obtained from *Fermentas* and were used according to manufacturer's instructions.

Cofactor and cofactor analogs

AdoMet: *S*-(5'-adenosyl)-L-methionine chloride (purity ~70 %), and S-(5'-adenosyl)-L-methionine p-sulfonate salt (purity 91 %) were from *Sigma Aldrich*. Commercial AdoMet was further purified to remove contaminating AdoHcy by passing through a short column of C_{18}-Reversed phase Silica gel 100 (*Fluka*), using 25 mM $HCOONH_4$ pH 3.5 as an eluent, or by cation-exchange chromatography on a Dowex 1 column using 0.1 mM HCOOH as an eluent. The resulting AdoMet was at least 95% pure and contained no detectable AdoHcy. It was concentrated to 21-23 mM and stored at -20ºC in small aliquots. **[*methyl*-^{3}H]AdoMet** (specific activity 70-84 Ci/mmol (1 mCi/ml); 15-16,1 Ci/mmol (1 mCi/ml) and 0.5 Ci/mmol (0,5 mCi/ml) was from *Amersham Pharmacia Biotech.*

AdoHcy (*S*-(5'-adenosyl)-L-homocysteine) was from *Sigma Aldrich*. **AdoPentyn** (S isomer, purity >90%) and **AdoHxGABANH$_2$** (racemic mixture, purity >90%) were synthesized by dr. G. Lukinavičius.

DNA and oligonucleotides

poly[dG-dC]·poly[dG-dC] was from *Sigma Aldrich*. Molar concentration of total nucleotides was determined using the published $\varepsilon_{254} = 8.4 \cdot 10^3\ M^{-1} \cdot cm^{-1}$ (Wells *et al.*, 1970). The concentration of double-stranded GCGC sites was calculated as one-fourth of total nucleotides concentration (Wu and Santi, 1987).

DNA oligonucleotides were from *IDT Technologies*, *MWG Biotech* or *Fermentas* (HPSF grade or HPLC purified). For kinetic experiments oligonucleotides of ≥80% purity were used. DNA duplexes for biochemical studies were produced by annealing 20-150 μM of individual strands by slowly cooling from 90ºC to 20°C over 2 h.

Table 1. DNA oligonucleotides for amplification of bisulfite-converted DNA

Fragment	Oligonucleotide name	Oligonucleotide sequence	Strand	Length of analyzed region*, nt	Location on pBR322**
I	NG-VP1	GTTGTGTAGATAATTATGATATGGGAGGG	upper	237	3369-3364
	NG-VA7	AAACCAAAACTAAATAAAACCATACCAAAC			
II	NG-AP2	TTTCATTCATCCATAATTACCTAACTCCCC	lower	294	3339-3695
	NG-AA8	GGGGGATTATGTAATTTGTTTTGATTGTTGGGA			
III	NG-VP5	GTTGTTGTTATTGTTATAGGTATTGTGGTG	upper	312	3594-3967
	NG-VA3	CATTTTCCAATAATAAACACTTTTAAAATTCT			
IV	NG-AP6	CAACATTATTACCATTACTACAAACATC	lower	309	3590-3958
	NG-AA4	ATGATGATGATTTTTAAAGTTTTGTTATGTGG			

V	**MV-VP1** **MV-VA2**	TTAGGGGATAATGTAGGAAAGAATATGTGAG AACAACCTACACCAAACTAAAATACCTACA	upper	257	2451-2768
VI	**MV-AP3** **MV-AA4**	CTCAAAAACAATAATACAATTATCCACAAAATC GAAGGGAGAAAGGTGGATAGGTATTTGGTAAG	lower	223	2420-2707

* excluding primer hybridization sites.

** amplicon position on pBR322, including primer hybridization sites.

Table 2. DNA duplexes for biochemical analysis of M.HhaI

Target site	Target base pair	DNA sequence
GCGC/GMGC*	**C:G**	5′-TACAGTATCAG**GCGC**TGACCCACAA TGTCATAGTC**CGMG**ACTGGGTGTTG-5′
GCGC/GMTC	**C:T**	5′-TACAGTATCAG**GCGC**TGACCCACAA TGTCATAGTC**CTMG**ACTGGGTGTTG-5′
GCGC/GMCC	**C:C**	5′-TACAGTATCAG**GCGC**TGACCCACAA TGTCATAGTC**CCMG**ACTGGGTGTTG-5′
GsGC/GMGC*	**s:G**	5′-TACAGTATCAG**GsGC**TGACCCACAA TGTCATAGTC**CGMG**ACTGGGTGTTG-5′
GPGC/GMTC*	**P:T**	5′-TACAGTATCAG**GPGC**TGACCCACAA TGTCATAGTC**CTMG**ACTGGGTGTTG-5′
GPGC/GMGC*	**P:G**	5′-TACAGTATCAG**GPGC**TGACCCACAA TGTCATAGTC**CGMG**ACTGGGTGTTG-5′
GFGC/GMGC*	**F:G**	5′-TACAGTATCAG**GFGC**TGACCCACAA TGTCATAGTC**CGMG**ACTGGGTGTTG-5′
GFGC/GFGC-13*	**F:G**	5′-TGATA**GFGC**TATC CTAT**CGFG**ATAGT-5′

* M – 5-methylcytosine, s – abasic site (sugar), P – 2-aminopurine, F – 5-fluorocytosine.

Methyltransferases

M.HhaI C81S was prepared by dr. G. Lukinavičius. **M.HhaI W41F** was prepared by dr. E. Merkienė.

Buffers for DNA methyltransferase assays

Methylation buffer I: 50 mM Tris·HCl pH 7.4, 15 mM NaCl, 5 mM EDTA, 0.2 mg/ml BSA, 2 mM 2-mercaptoethanol. **Methylation buffer II:** 50 mM MOPS pH 7.4, 15 mM NaCl, 0.5 mM EDTA, 0.2 mg/ml BSA, 2 mM 2-mercaptoethanol. **Reaction buffer:** 10 mM Tris·HCl pH 7.4, 50 mM NaCl, 0.5 mM EDTA, 2 mM 2-mercaptoethanol.

Plasmids, strains and protein expression

pHH553 – M.HhaI overexpression vector; derivative of pUHE25-2: Ap^r Cm^r, pMB1 replicon; expression of target gene is under control of hybrid T7 early promoter $P_{A1\text{-}04/03}$ (H.Bujard, unpublished observations). *hhaIM* gene contains six unique REase sites introduced by silent mutations (Acc65I, Ecl136II, Cfr9I (SmaI), BspTI, Eco91I, BpiI) (Merkienė and Klimašauskas, 2005). Derivatives of **pHH553: pHHΔ324G** codes for the enhanced-solubility M.HhaI variant, in which C-terminal FKPY tetrapeptide is substituted with a single glycine residue (Δ324G) (Daujotytė *et al.*, 2003); **pL2Bsp** – codes for the hybrid MTase Hha-L2Bsp, in which recognition Loop-2 of M.HhaI is substituted with a predicted analogous fragment from M.Bsp6I (Vilkaitis and Klimašauskas, 1998).

Expression of mutant MTases was carried out in *E.coli* strains **ER2267** (*New England Biolabs*), **ER1727** (Dila *et al.*, 1990), and **ER2566 *lacI*^{q}**, which was obtained by transferring an episome from **ER2267** to endow ER2566 (*New England Biolabs*) cells with a *lacI*^{q} gene (G. Mitkaitė, personal communication).

Unless indicated otherwise, *E.coli* cells were grown in minimal M9 medium supplemented with histidine, methionine, tryptophan (10 μg/ml each), thiamine (6 μg/ml), carbenicillin (100 μg/ml) and kanamycin (30 μg/ml) at 37ºC overnight. For protein overexpression and purification, the ER2267 strain bearing an appropriate plasmid was grown in the same medium at 37ºC until OD600 nm 0.6-0.8, then the culture was cooled down to 16ºC and IPTG added to a 0.4 mM concentration and incubation continued at 16ºC overnight. For the determination of L2Bsp specificity *in vivo*, mutant MTase was overexpressed in ER1727: the cells were grown in LB medium supplemented with ampicillin (100 μg/ml) and tetracycline (10 μg/ml) until OD600 nm 0.6-0.8, then IPTG was added to a final concentration 0.4 mM and the cells were cultivated at 37°C for additional 2 h.

Molecular cloning

If not indicated otherwise, plasmid DNA isolation, protein and DNA electrophoresis, *E.coli* transformation and other molecular cloning procedures were performed according to well established protocols (Ausubel, 1995; Sambrook and Russel, 2001).

Construction of mutant hhaIM genes

WT M.HhaI, HhaI-L2Bsp and HhaI-ΔL2 mutants were expressed as enhanced solubility M.HhaI variant Δ324G (Daujotytė *et al.*, 2003). The C-terminal His_6 tag was added to Δ324G gene by PCR mutagenesis using oligonucleotides 5'-TTTTCGCAATGATCTCAATATT C-3' (direct) and 5'-TTAGTGGTGGTGGTGGTGGCCATTTAATGATGAAC-3' (reverse; 21 nts coding for a His_6 tag and a stop codon are underlined) and pHHΔ324G as a template, resulting in plasmid **pΔ324GH$_6$**. R.BpiI and R.HindIII (filled-in) sites were used for cloning. **pL2BspΔ324GH$_6$** (coding for His_6-tagged version of HhaI-L2Bsp) was constructed by transferring 600 bp fragment from pL2Bsp into pΔ324GH$_6$, using REases Eco81I and Eco91I. **pHHΔBE** was constructed by digesting pΔ324GH$_6$ with R.BspTI and R.Eco91I, blunt-ending with Klenow fragment and recirculization to give a 36 nt in-frame deletion in the *hhaIM* gene, involving recognition Loop-2 of M.HhaI.

Protein purification and characterization

Affinity purification of his$_6$-tagged MTases

Cells were disrupted by sonication in a Buffer A (20 mM Na-PO4 pH 7.4, 1 M NaCl and 1 mM PMSF). The supernatant was loaded onto 5 ml HiTrap Chelating™ column (*Amersham Pharmacia Biotech*) and eluted with a 3–300 mM linear gradient of imidazole in Buffer A. After extensive dialysis to remove AdoMet, the protein was concentrated to 100-200 μM and stored at -20°C in a buffer containing 20 mM Na-PO_4 pH 7.4, 0.5 mM EDTA, 100 mM NaCl, 2 mM 2-mercaptoethanol and 50% glycerol.

M.HhaI purification by ion-exchange chromatography

WT M.HhaI and its point mutants with no affinity tag were purified as described previously (Kumar *et al.*, 1992). Briefly, cells were disrupted by sonication in the low ionic strength buffer (10 mM Hepes pH 7.2, 5 mM EDTA, 10%glycerol, 2 mM 2-mercaptoethanol), where M.HhaI precipitated together with the cell debris. Then the MTase was selectively enriched by back-extraction with high ionic strength buffer (10 mM K-PO_4 pH 7.4, 400 mM NaCl, 5 mM EDTA, 10% glycerol, 2 mM 2-mercaptoethanol) and extensively dialyzed to remove AdoMet. Then solution was loaded onto Q-Sepharose column and the flow-through containing M.HhaI was fractionated on the SP-Sepharose column with 0.1-0.5 M linear gradient of NaCl. The purified proteins were concentrated to 200-300 μM and stored at -20°C in a buffer containing 10 mM K-PO_4 pH 7.4, 0.5 mM EDTA, 100 mM NaCl, 2 mM 2-mercaptoethanol and 50% glycerol.

Characterization of protein preparations

Protein concentrations were determined by a Coomassie G-250 assay with BSA as a standard. The concentrations of WT M.HhaI and variants with point mutations were further refined by active site titration using a fluorescent DNA duplex GPGC/GMGC, as previously described (Vilkaitis *et al.*, 2000).

The molecular size of all proteins was verified by ESI mass spectrometry. Briefly, 30 μl of concentrated protein solution (100-300 μM) was dialyzed against 0.1% acetic acid at 8°C overnight and centrifuged in the bench-top centrifuge for 10 min. at the maximum speed. 10 μl of supernatant was loaded into Hewlett-Packard 1100 series single quadruple electrospray ionization mass spectrometer and eluted into the ionization chamber with 50% isopropanol and 1% acetic acid solution at a flow rate 30 μl/min. Mass spectra were acquired in 700-2000 m/z range in a positive ion mode. Ionization capillary voltage was 5000 V, fragmentor voltage was 120 V, quadruple temperature was 350°C, drying gas temperature was 300°C and flow rate was 10 l/min. The spectra were analyzed with HP Chemstation software (version A.05.01).

Using DNA protection assay (Dalhoff *et al.*, 2006) W41F/I86W and C81S protein preparations were shown to contain <<1% AdoMet; WT M.HhaI contained less than 2% AdoMet. The altered-specificity MTase preparations were shown to contain no residual AdoMet by HPLC. Briefly, 20 μl of protein (total 1.4-2.8 pmol) was denatured by adding 4 μl of 5.5 N HCl. The mixture was centrifuged and the supernatant was mixed with 140 μl H_2O. The protein was remove by ultracentrifugation and the sample was split into to aliquots of 55 μl. 10 μl H_2O was added to one of them, the other was supplemented with 10 μl 9 mM AdoMet (internal standard). 60 μl of each sample was analyzed by ion-exchange HPLC on the Nucleosil 100-5SA column (4×150 mm; *Macherey-Nagel*) using 100 mM TEA-HCOOH pH 4.0 as elution buffer (flow rate 0.9 ml/min, 25°C). Retention time of AdoMet under these conditions was ~6.5 min. The amount of AdoMet in the sample was determined by integrating peak area.

Characterization of Methyltransferases *in vitro*

Analysis of methylation specificity by DNA protection assay

1.2 μg λ DNA was methylated in 20 μl of Methylation buffer II containing 300 μM AdoMet. The amount of MTase was varied from 3.1 μM to 12 nM in twofold

dilutions which corresponded to MTase : GCG sites ratio from 1:1 to 1:512 (equivalent to MTase : GCGC ratios from 8:1 to 1:64). Methylation reaction was allowed to proceed for 1 h then MTase was inactivated by heating at 80°C for 15 min. Reactions with AdoMet analogs were performed under the same conditions with AdoMet substituted by 300 μM cofactor analog. The restriction reactions were performed as follows: 5 μl of mixture containing 1×REase buffer (*Fermentas*), 20 mM $MgCl_2$ and 1-3 u of 5mC-sensitive REase was added to 10 μl of methylation mixture. Incubation was continued at 37°C for additional 2 hours and the REase was inactivated by incubation with SDS (0.25%) and proteinase K (0.4 mg/ml) for ≥30 min. at 55°C. The mixture then was fractionated in the 1% agarose gel, DNA was visualized by ethidium bromide staining and the least amount of MTase required for full protection from digestion was determined.

Site specific labeling of GCG sites in plasmid DNA

5 μg pUC19 DNA was methylated by 1.55 μM ΔL2-14 and 300 μM AdoHeksinGABANH_2 in the Methylation buffer II at 37°C for 1 h. The reaction was stopped by phenol, phenol-chloroform and chloroform extractions followed by precipitation with ethanol. 3.6 μg of modified DNA in 50 μl 150 mM $NaHCO_3$ pH 9.0 was incubated with 105 mM Cy5-NHS for 1 h at room temperature then precipitated with isopropanol. The pellet was washed twice with 75% ethanol, dried, dissolved in water and further purified with QIAquick DNA purification kit (*Qiagen*).

Analysis of methylation specificity by bisulphite sequencing

For the determination of HhaI-L2Bsp specificity *in vivo*, pL2Bsp plasmid isolated from cells overexpressing HhaI-L2Bsp were analyzed. For the analysis of methylation specificity *in vitro*, 4.9 μg of pBR322 in 120 μl of reaction mixture (3.1 μM GCG sites) was methylated with 775 nM and 48 nM HhaI-L2Bsp or HhaI-ΔL2 mutant which corresponded to MTase : GCG ratio of 1:4 and 1:64, respectively. After the reaction the mixture was treated with 0.2 mg/ml proteinase K in the presence of 0.25% SDS for 1 hour at 55°C. Then the sample was extracted with equal volumes of phenol-chloroform-isoamyl alcohol mixture, then with chloroform and DNA was precipitated with ethanol. In order to facilitate denaturation methylated DNA was fragmented either with R.Alw44I (pL2Bsp) or with R.PagI (pBR322).

Bisulphite solution was prepared immediately before use by dissolving 7.74 g of sodium metabisulphite (67% SO_2, ACS reagent, *Sigma*) in 7.5 ml of H_2O, 1.5 ml of 2 N NaOH and 1.5 ml of 1 M hydroquinone. The mixture was incubated at 50-60°C while vortexing until solid sodium bisulphite completely dissolved. Approximately 0.5–1 μg DNA in 20 μl was denatured by adding 3 μl of 2 N NaOH to a final concentration of 0.3 M and incubating for 30 min at 37°C. Then 500 μl of freshly prepared sodium bisulphite solution was added. Samples were incubated for 5 hours at 55°C. Each hour the temperature was raised to 95°C for 3 min. in order to ensure complete denaturation of the DNA. Samples were desalted using DNA Extraction Kit (*Fermentas*) and desulphonation of DNA was carried out by incubating for 10 min at room temperature in the presence of 0.2 N NaOH. The reaction mixture was neutralized by addition of ammonium acetate (pH 7.0) to 2.5 M, DNA was precipitated with ethanol and dissolved in 50 μl H_2O.

PCR was performed using modified-DNA specific primers (**Table 1)**. 1-2 µl of bisulphite modified DNA was used as a template for a 50 µl PCR reaction. Cycling conditions: denaturation at 94°C for 3 min, followed by 26 cycles of denaturation at 94°C for 1 min, annealing at 48°C for 1 min, and extension at 72°C for 1 min and a final extension was performed for another 5 min at 72°C. When methylation specificity *in vitro* was investigated, primers NG-VP1 and NG-VA3 were used to amplify a 599 bp upper strand fragment corresponding to the 3369–3967 positions of pBR322. A single strand-specific PCR product was obtained. It was purified by incubation with SAP and exonuclease I and directly sequenced using NG-VP1 primer (read length of 538 nt). The positions of methylated cytosines were identified by examination of sequencing traces and characterized as methylated, partially methylated or non-methylated. Alternatively PCR product was purified by agarose gel electrophoresis, blunt-ended by incubation with T4 DNA polymerase and sub cloned into SmaI-digested pUC19. DNA from individual clones was sequenced using primers M13/pUC dir (-46), 22-mer and/or M13/pUC rev (-46), 24-mer (*Fermentas*). The sequencing results were analyzed with the BiQ Analyzer software v0.91 beta (Bock *et al.*, 2005).

Enzymology methods

Analysis of MTase-cofactor interaction

Affinity of mutant MTases towards cofactor AdoMet was determined by measuring tryptophan fluorescence quenching upon cofactor binding as described previously (Vilkaitis *et al.*, 2000) Fluorescence measurements were performed at 25°C on Perkin-Elmer LS-50B luminescence spectrometer. The fluorescence intensity was measured at an excitation wavelength of 295 nm (bandwidth 2.5 nm) and emission wavelength of 350 nm (bandwidth 5 nm). WT M.HhaI or mutant MTases (1 µM) were titrated by addition of AdoMet (0.2-550 µM) in the Reaction buffer. Titration data were fitted into the equation for single site binding using Grafit 5.0.6.

Steady-state kinetic analysis

Methylation reactions were carried out in the Methylation buffer I at 37°C. K_M^{AdoMet} measurements of ΔL2-MTases were performed with constant MTase (4 nM) and poly[dG-dC]·poly[dG-dC] DNA (1.5 µM double-stranded GCGC sites) and varying [*methyl*-^{3}H]AdoMet (0.5 Ci/mmol) concentrations. Reactions were allowed to proceed for 10 min (ΔL2-14) or 25 min (ΔL2-6, ΔL2-9 and L2Bsp). K_M^{DNA} measurements were performed with constant MTase (25 or 50 pM ΔL2-6 and ΔL2-9 or 1 nM L2Bsp) and [*methyl*-^{3}H]AdoMet (500 µM or 750 nM (16.1 Ci/mmol) for ΔL2-6 and ΔL2-9; 10 µM for L2Bsp (0.5 Ci/mmol)) concentrations and varying poly[dG-dC]·poly[dG-dC] concentration. Reactions were allowed to proceed for 30-60 min and were stopped by adding HCl to 0.5 N concentration. Samples were spotted onto Whatman DE-81 filters and processed as described previously (Vilkaitis *et al.*, 2000; Vilkaitis *et al.*, 2001). Kinetic constants of WT HhaI were determined as described previously using 16.1 or 76 Ci/mmol [*methyl*-^{3}H]AdoMet (Vilkaitis *et al.*, 2000). Data were analyzed by non-linear regression fitting to a Michaelis–Menten equation using Grafit 5.0.6 (*Erithacus Software*) or Dynafit software (Kuzmic, 1996).

Single turnover assay

Single turnover reactions were performed in the Methylation buffer I at 25ºC in a Rapid-Quench-Flow instrument RQF-3 (*KinTek*). MTase preincubated with DNA was mixed rapidly with [*methyl*-^{3}H]AdoMet, and after a specified period (0.1-300 s) the reactions were quenched with 0.7 N HCl (final concentration). The amount of radioactivity incorporated into DNA was measured by filter binding assay as described (Vilkaitis *et al.*, 2000). In case of WT M.HhaI 200 nM MTase, 100 nM GCGC/GMGC and 2.5 μM [*mehyil*-^{3}H]AdoMet (16.1 Ci/mmol) were used. In case of M.HhaI W41F/I86W mutant the reactions contained 1 μM MTase, 0.5 μM GCGC/GMGC and 30 μM [*methyl*-^{3}H]AdoMet (1.62 Ci/mmol). 1.62 Ci/mmol [*methyl*-^{3}H]AdoMet was obtained by mixing 16.1 Ci/mmol [*methyl*-^{3}H]AdoMet with 5 mM nonradioactive AdoMet at the v/v ratio of 9:1.

Stopped-flow kinetics

Measurements were performed at 25°C in the Reaction buffer (without BSA) using an Applied Photophysics SX.18MV spectrometer equipped with a xenon lamp, 20 μl cell and a dual detection accessory allowing simultaneous acquisition of absorbance and fluorescence traces. Excitation wavelength was set by monochromator to 280 nm (slit-width of 4.65 mm). Experiments were performed using 10 mm or 2 mm path lengths when higher sensitivity in absorption mode or lower inner-filter effect in the fluorescence regime, respectively, was required. Typically, 3–7 traces were acquired and averaged from reactions containing 3 μM MTase, 2.5 μM DNA and 0–200 μM AdoMet or AdoHcy. Where necessary, the fluorescence traces were corrected for inner filter effect, as described (Palmier and Van Doren, 2007). The obtained traces were analyzed using Kinet Asyst (*Hi-Tech*) or Dynafit (Version 3.28.059) (Kuzmic, 1996) software. Multicurve analysis was performed using Origin 8.1 SS2 software (*OriginLab*). Normalized absorbance and fluorescence curves were fitted into three-exponential functions with two shared rate constants.

RESULTS AND DISCUSSION

1. Engineering MTases with a new target specificity

During recent years prokaryotic DNA methyltransferases have become important tools in molecular biology and nanotechnology (Buryanov and Shevchuk, 2005; Klimašauskas and Weinhold, 2007). For example, DNA MTases recognizing short sequences have been used as non-destructive chromatin foot-printing agents *in vivo* or *in vitro* (Kilgore *et al.*, 2007). Targeted heritable gene silencing can be achieved by MTases fused with sequence-specific Zn-finger domains (Carvin *et al.*, 2003; Smith and Ford, 2007). Irreversible covalent inhibitory complexes between C5-MTases and nucleotide analogs such as 5-fluorocytosine in DNA can be used for the design of covalently functionalized DNA-based nanostructures (Singer and Smith, 2006). Recently, novel approaches for sequence-specific functionalization and labeling of DNA using synthetic analogs of the cofactor AdoMet have been proposed (Klimašauskas and Weinhold, 2007). More than 900 different MTases are known recognizing over 200 different DNA targets. Since the repertoire of naturally occurring enzymes still lacks many useful sequences and sequence types, engineering of enzymes with novel predetermined specificities is increasingly desirable.

Crystal structures of the C5-MTases HhaI (Klimašauskas *et al.*, 1994) and HaeIII (Reinisch *et al.*, 1995) show that these enzymes recognize their targets by means of two distinct recognition loops (Loop-1 and Loop-2) from the target recognition domain (TRD), which form multiple base-specific contacts with two distinct segments of the target sequence (**Figure 1**). Modular organization of TRDs suggested that novel specificities could potentially be created by swapping loop regions between different MTases, however, such experiments typically yielded enzymes with diminished catalytic activity and relaxed target specificity (Lange *et al.*, 1996; Vilkaitis and Klimašauskas, 1998). During earlier work on swapping segments of TRD it was found that hybrid MTases, in which the C-terminal recognition loop (Loop-2) of M.HhaI was exchanged, often retained the ability to methylated DNA although with lower efficiency (Vilkaitis and Klimašauskas, 1998). One such hybrid, M.HhaI-L2Bsp (**Figure 1**), obtained by replacing recognition Loop-2 of M.HhaI with a short fragment from M.Bsp6I (target site G<u>C</u>NGC) showed a marked preference to methylated G<u>C</u>G targets. The asymmetric nature of the G<u>C</u>G sequence allows the generation of hemimethylated <u>C</u>G sites in DNA, which can be uniquely used to study the action of eukaryotic maintenance DNA methyltransferases (Vilkaitis *et al.*, 2005). These features of the hybrid MTase prompted further studies.

In this work a detailed analysis of the sequence specificity of the hybrid M.HhaI-L2Bsp MTase (further referred to as L2Bsp) is reported. Next, using M.HhaI as a model system the new strategy for the construction of catalytically efficient and sequence-selective DNA C5-MTases via functional elimination of one of the two target recognition elements is demonstrated.

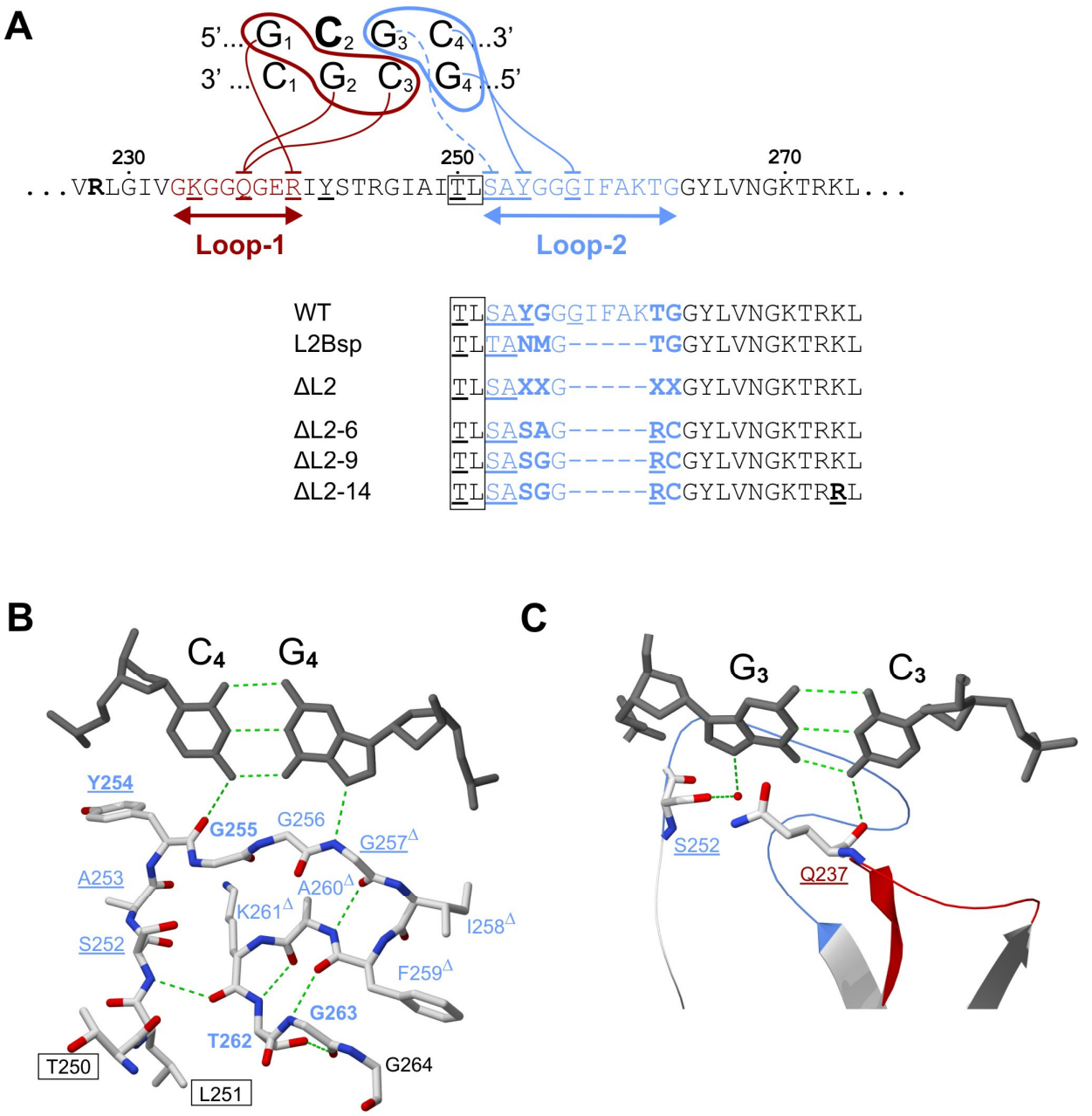


Figure 1. Recognition of the GCGC target sequence by the HhaI methyltransferase. **A**, *top*– schematic representation of contacts between the target recognition loops (Loop-1 and Loop-2) of M.HhaI and the DNA bases in the GCGC target site; DNA contacting residues are underlined; lines represent direct H-bonds to the DNA bases, dotted lines indicate nucleobase contacts through a water molecule; residues of the conserved TL dipeptide are boxed; target cytosine C_2 is shown in bold; residues in Loop-1 and Loop-2 are colored red and blue, respectively; **A**, *bottom* – aligned Loop-2 sequences of from WT M.HhaI and its engineered variants with altered target specificity; ΔL2 represents the randomized library; randomized positions are marked as X (bold); an additional mutation outside Loop-2 is bold underlined. **B** and **C** – stick models depicting interactions of M.HhaI with the 4th and 3rd target G:C base pair, respectively, based on a crystal structure of the M.HhaI-DNA-AdoHcy complex (PDB code 3mht (Klimašauskas *et al.*, 1994)). Deleted residues are marked with Δ, H-bonds are shown as green dotted lines, other coding as in **A**.

1.1. Sequence specificity of L2Bsp *in vivo*

In preliminary studies, it was found that the previously constructed L2Bsp hybrid MTase preferentially methylates GCG targets (Vilkaitis and Klimašauskas, 1998). To determine the *in vivo* specificity of the L2Bsp variant at single nucleotide resolution, we performed bisulfite sequencing of plasmid DNA isolated from *E. coli* ER1727 cells overexpressing L2Bsp. Plasmid DNA was treated with bisulfite and PCR-amplified fragments were cloned into R.SmaI-digested pUC19. Both strands of three fragments from the pMB1 replicon and the β-lactamase gene were analyzed in individual clones (**Table 1**). The total length of analyzed regions spanned 1577 nucleotides which included 32 different GCG sites. Control sequencing of a plasmid which does not encode an active methyltransferase found no non-converted cytosines besides those in three CCWGG sites methylated by the endogenous EcoDcm MTase. Consistent with earlier results (Vilkaitis and Klimašauskas, 1998), the GCG sites proved the major targets of L2Bsp (**Figure 2**), although the extent of methylation at individual GCG sites varied from 40 to 100%, no clear correlation between the methylation efficiency and the nature of flanking sequences could be established. However, significant methylation was observed at non-GCG sites: certain GCA sites were methylated as efficiently as 40%, and GCC and GCT were occasionally found to be methylated up to 25%. We concluded that the specificity of L2Bsp is more degenerate than was thought previously and should be defined as GC[G/a]. Therefore the utility of L2Bsp as a molecular tool was limited due to its promiscuous specificity, which manifested as substantial methylation of sequences outside the desired CG consensus.

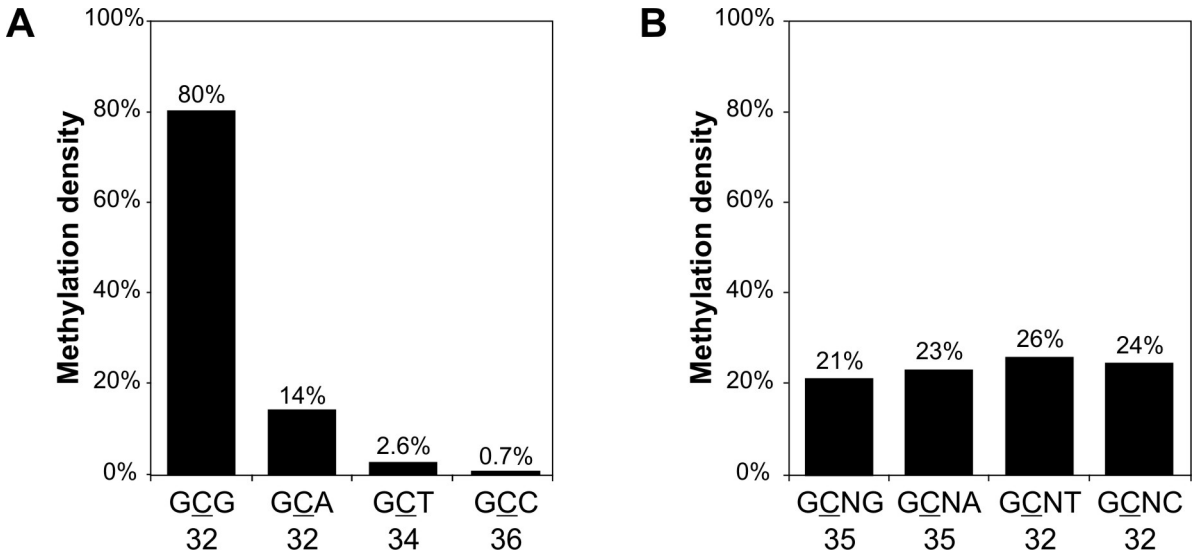


Figure 2. Bisulfite sequencing analysis of the *in vivo* sequence specificity of L2Bsp. Plasmid DNA was isolated from *E. coli* cells overexpressing L2Bsp and analyzed by bisulfite sequencing. The analyzed region spanned 1577 bases and contained 134 GCN sites. Methylation of each site was assayed by sequencing 16-32 independent clones and the methylation density at each position was determined as a ratio of methylated cytosines over the total number of sequence reads. An average methylation density at GCX (left) and GCNX (right) sequences is shown (the target cytosine residues are underlined) and the number of individual target sites is indicated underneath each sequence.

1.2. Strategy for designing catalytically efficient GCG-specific methyltransferase

In this work, our ultimate target sequence specificity was GCG. Starting from M.HhaI, our task appeared as a "functional deletion" of Loop-2, which is responsible for recognition of the 4th base pair (**Figure 1**). This recognition is mediated by two hydrogen bond contacts from the main chain atoms of the Loop-2. Directed evolution of DNA-contacting residues in Loop-2 did not yield any active mutants with altered target specificities (Lee *et al.*, 2002). However this was in part achieved in the L2Bsp hybrid, where the recognition Loop-2 of M.HhaI had become shorter by 5 amino acids. Based on the crystal structure of the M.HhaI-DNA complex we concluded that the truncation of the Loop-2 and the associated loss of recognition contacts to the fourth G:C pair is most likely responsible for the change in sequence specificity of the L2Bsp hybrid towards the GCG sites. We also assumed that the new sequence elements that come from M.Bsp6I are too short and out of context to form any structural elements for DNA target recognition. Thus, in order to obtain a more efficient GCG-specific MTase, we decided to retain a five-residue deletion and to optimize the sequence of the truncated Loop-2 by directed molecular evolution. We went on to create a protein library in which four residues in the truncated Loop-2 were randomized; a Gly residue was retained in the center of the random region (XXGXX) (**Figure 1**) to ensure a certain degree of folding flexibility in the redesigned loop.

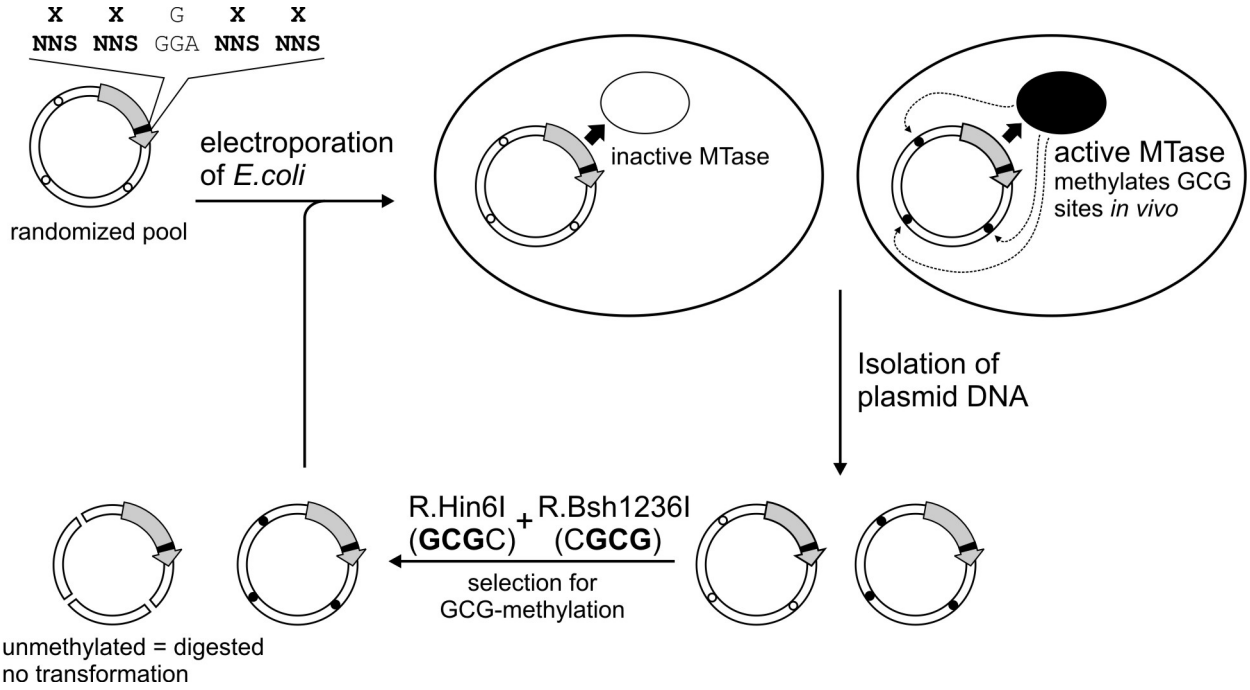


Figure 3. Directed evolution of GCG-specific C5-MTases. A schematic representation of random library selection for active GCG-specific MTases. The MTase encoding gene is shown as a grey arrow, the recognition Loop-2 is shown in black; open and filled circles indicate nonmethylated and methylated GCG sites, respectively.

Mutagenesis was carried out by PCR with a degenerate mutagenic primer carrying four NNS codons in positions to be randomized (**Figure 3**). Linearized pHHΔBE, which contained a 36 nt deletion in the Loop-2 and encoded a catalytically inactive protein, was used as the PCR template. The lack of homology to the degenerate part of the primer precluded potential sequence bias from the template. The randomized

PCR fragment was inserted in-frame into pHHΔBE, and transformation of ER2566 *lacI*^{q} *E.coli* cells with the resulting plasmid pool yielded a total of ~$6 \cdot 10^6$ clones in liquid culture (determined by aliquot plating). The cells were grown at 37°C in M9 minimal medium with background level of transcription (no IPTG added). These growth conditions permitted complete methylation of plasmid DNA (protection against R.Hin6I cleavage *in vitro*) in cells expressing WT M.HhaI, but little methylation was rendered by L2Bsp (not shown). The total plasmid DNA was isolated, and sequencing of the randomized region showed no significant nucleotide bias in the four randomized codons. Restriction endonuclease mapping confirmed the correct overall structures of the plasmid DNA (not shown).

Selection was carried out by digesting the total plasmid DNA with two 5-methylcytosine sensitive restriction endonucleases, R.Hin6I (GCGC; cytosines whose methylation blocks DNA cleavage are underlined, methylation sensitivity profiles provided in ref. (Roberts *et al.*, 2010)) and R.Bsh1236I (CGCG) (14 and 8 target sites on the pHhaI-ΔL2 plasmid, respectively). The digested plasmid DNA was again transformed into *E. coli* cells and ~$2 \cdot 10^4$ clones were obtained. Plasmid DNA from 9 individual colonies was analyzed with R.Bsh1236I, which showed only 4 transformants expressing active methyltransferases. The remaining clones were combined, a total plasmid DNA was isolated and the selection procedure was repeated. After the second round of selection, ~1500 transformants were obtained. 35 individual clones were tested with R.Bsh1236I and most showed a higher degree of protection than the original L2Bsp expressing plasmid. Sequencing of 28 clones revealed only two sequence variants (SGGRC – 22 clones and SAGRC – 1 clone) in the randomized region (**Figure 1A**). Notably, nearly all possible codons for Ser, Gly, and Arg were found in the mutant genes indicating that the selected variants did not arise from a single precursor or due to nucleotide bias in the library. A third protein variant (5 clones) contained an additional inadvertent mutation, K273R, 10 codons outside of the randomized region.

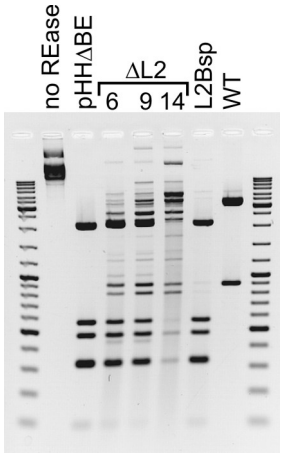


Figure 4. *In vivo* methylation activity of mutant MTases. *E. coli* ER2267 cells carrying plasmids that encode WT or mutant MTases were grown in the minimal medium at 37ºC overnight, plasmid DNA was isolated and digested with R.Bsh1236I, which recognizes CGCG target and is sensitive to methylation of either cytosine; pHHΔBE – control plasmid, coding for inactive MTase.

Representatives of all three variants were selected (referred to as ΔL2-6, ΔL2-9 and ΔL2-14) and corresponding plasmids were analyzed with R.Bsh1236I to compare

their methylation efficiency *in vivo* under background level of transcription with that of the original L2Bsp (**Figure 4**). All three selected variants showed a much stronger protection against the cleavage as compared to L2Bsp; notably, the highest degree of protection came from the ΔL2-14 variant, which contained the K273R mutation. All three ΔL2-mutant MTases (clones ΔL2-6, ΔL2-9 and ΔL2-14), WT M.HhaI and the L2Bsp controls were expressed in *E. coli* and purified to near-homogeneity by Ni^{2+} chelating affinity column chromatography for further characterization *in vitro*.

1.3. Catalytic efficiency and sequence-specificity of selected MTases

For initial characterization of the catalytic activity and sequence-specificity of the M.HhaI mutants, a DNA protection assay was used (Dalhoff *et al.*, 2006). Serial two-fold dilutions starting with equimolar amounts of MTase and target sites were used to methylate bacteriophage lambda DNA, which was then challenged with a set of individual 5-methylcytosine-sensitive restriction endonucleases and analyzed by agarose gel electrophoresis. Enzymatic turnover rates were estimated based on the minimal molar ratio of MTase to its target sites that is required for complete protection of DNA in 1 hour (end-point assay).

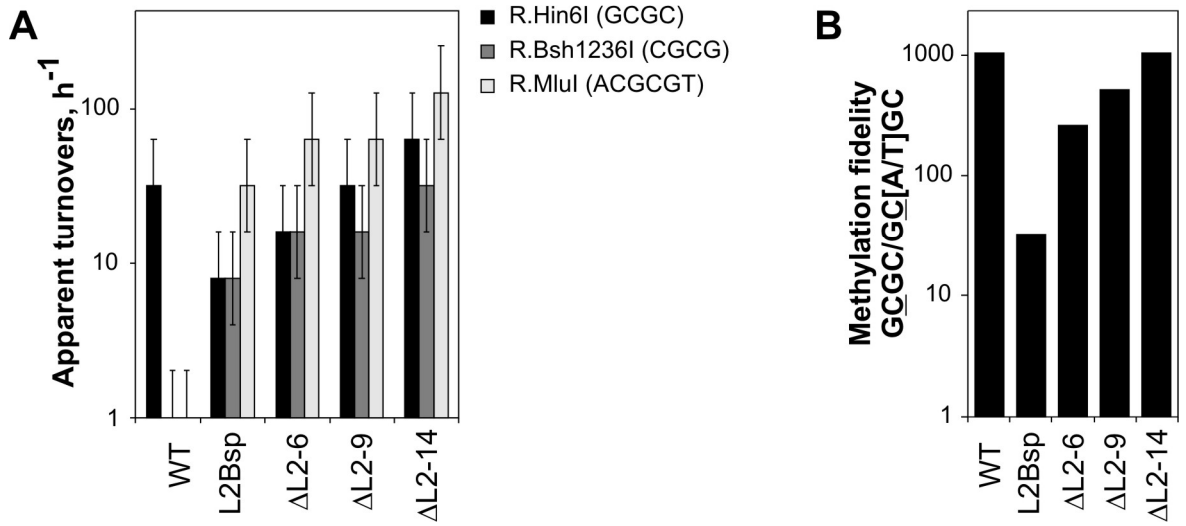


Figure 5. DNA protection analysis of the *in vitro* specificity of the ΔL2-MTases. A – apparent number of enzymatic turnovers executed at different target sites. Serial two-fold dilutions starting with equimolar amounts of MTase and target sites were used to methylate bacteriophage λ DNA, which was then challenged with a set of 5-methylcytosine-sensitive restriction endonucleases. Enzymatic turnover rates (turnovers per hour) were estimated based on a minimal molar ratio of MTase to its target sites that is required for a complete protection of DNA in 1 hour. **B** – the sequence fidelity of the ΔL2-MTases, expressed as the ratio of methylation turnover rates at GCGC to GC[A/T]GC sites.

All mutant MTases methylated DNA *in vitro* and complete protection from R.Hin6I could be readily achieved although at different enzyme dilutions (**Figure 5A**). In support of our *in vivo* results, L2Bsp proved the least active mutant, whereas the ΔL2-14 was the most efficient enzyme. The mutant MTases methylated a broader range of targets than the WT HhaI. A complete protection from REases that do not have

GCGC in their sites (R.Bsh1236I (CGCG) and R.MluI (ACGCGT)) was achieved, consistent with the specificity switch from GCGC (WT) to GCGN. Again, ΔL2-14 was found to be the most efficient MTase, whereas ΔL2-6 and ΔL2-9 showed very similar methylation efficiency.

Similarly, off-target methylation at GC[A/T] sites was assessed by digestion with R.BseXI (GC[A/T]GC). We found that methylation of these sites was evident only at high MTase concentrations with the selected mutants. Since no complete protection could be achieved at the non-cognate sites, direct comparison of the methylation efficiency at specific and non-specific sites was not possible by theoretical calculation of end-point dilutions assuming that, on average, five two-fold dilutions separate a starting and a full protection point (not shown). The derived ratios of the target/off-target methylation, which can only provide rough estimates are shown in **Figure 1B**, indicating that the ΔL2-14 mutant is similarly faithful as the WT M.HhaI, whereas the L2Bsp hybrid shows a ten-fold lower sequence fidelity.

When DNA MTases are used as molecular biology tools, it is convenient to express their quantity in activity units (u). 1 u of DNA MTase is defined as the amount of enzyme required to protect 1 μg of λ DNA in a total reaction volume of 50 μl in 1 hour at 37°C against cleavage by cognate restriction endonuclease. In order to assay specific activity of L2Bsp and ΔL2-14 R.Bsh1236I was used, because this REase probed the largest fraction of GCG sites (157 out of 1730 present on λ DNA). Specific activity (u / pmol) of ΔL2-14 was 16 times lower than that of WT M.HhaI (**Figure 6**). However, as there are 8 times more GCG than GCGC sites on λ DNA, one can conclude that both enzymes operate with comparable efficiency under conditions of this experiment. On the other hand the specific activity of L2Bsp is 6 lower than that of ΔL2-14.

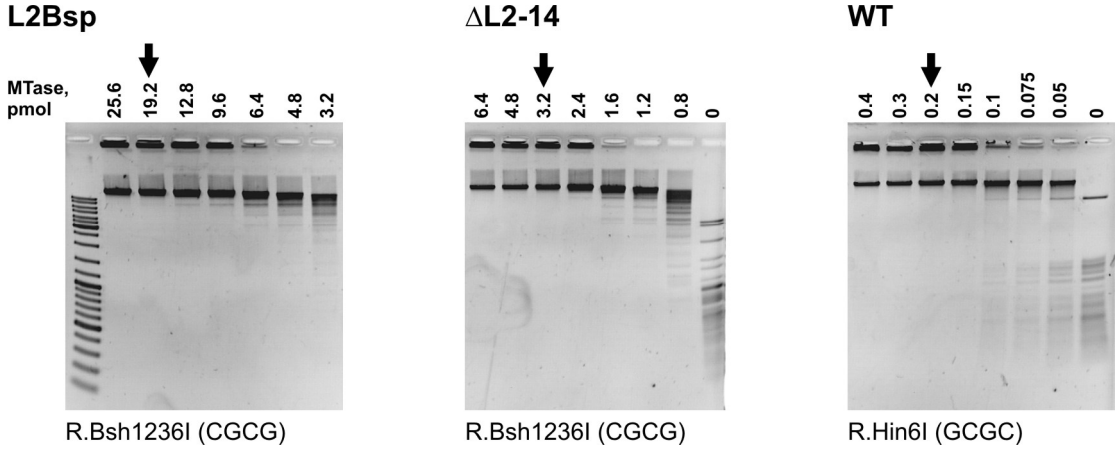


Figure 6. Determination of the specific activity of WT and truncated M.HhaI variants. 1 μg λ DNA was incubated with the indicated amounts of methyltransferases in 50 μl Methylation buffer II containing 100 μM AdoMet for 1 h at 37ºC and the minimum amount of MTase, required for complete protection from R.Bsh1236I or R.Hin6I was determined. Lanes containing 1 u of MTase are marked with arrows.

Since the use of restriction endonucleases for specificity assessment is undermined by the limited repertoire of available sequence specificities and by significant errors inherent in the determination of end-point fragmentation patterns, bisulfite sequencing of the modified DNA was employed to define more precisely the recognition targets of the ΔL2-MTases. pBR322 DNA was methylated *in vitro*, treated

with bisulfite and the region of interest was amplified with converted-DNA specific primers. We analyzed a 538 bp fragment in the β-lactamase gene (upper strand positions 3398–3935 on pBR322) which contains 9 GCG sites and 37 GC[A/T/C] sites.

For initial analysis, the PCR-amplified fragment was directly sequenced which allowed us qualitatively describe sites as methylated, unmethylated or partially methylated but gave no information about the methylation density of a particular site. When small amount of MTases were used for reaction (MTase : GCG molar ratio 1:64), 8, 7 and 5 out of 9 GCG sites were completely methylated by ΔL2-14, ΔL2-9, ΔL2--6, respectively (other GCG sites were nearly fully or partially methylated) (**Figure 7A**). No methylation in other positions was observed, and no non-converted cytosines were present in a control sample derived from pBR322 incubated without MTase. In contrast, L2Bsp methylated only 3 out of nine GCG motifs completely, while weak partial methylation at several GCA positions was clearly evident. Such a wide spread of modification densities observed with the different enzymes appeared suitable for quantitative analysis of their methylation specificity.

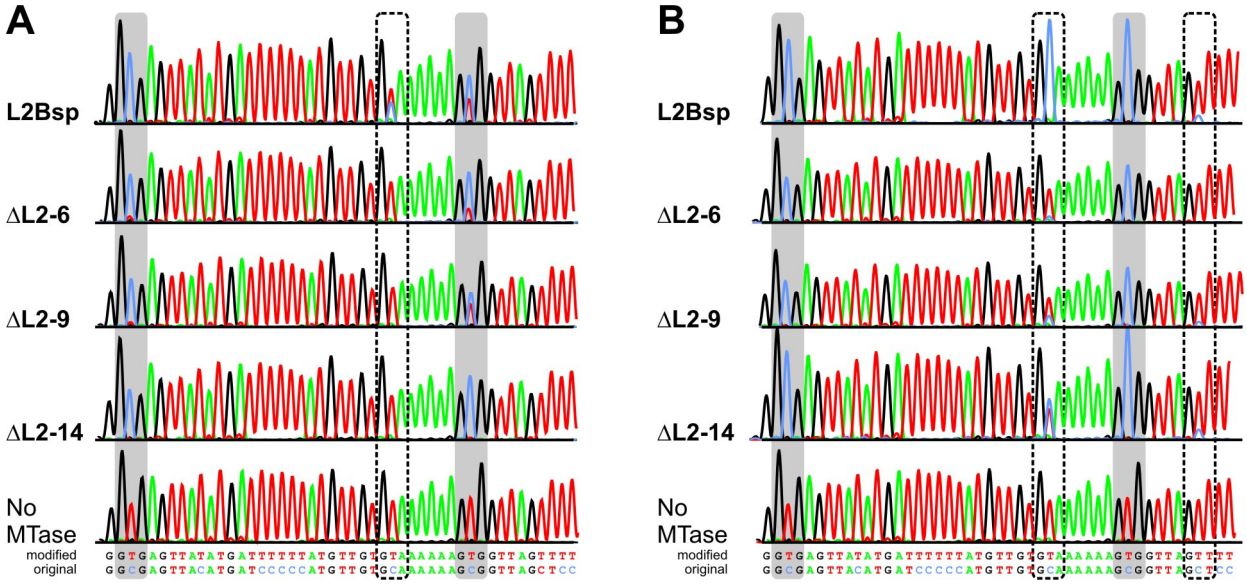


Figure 7. Analysis of the *in vitro* specificity of the ΔL2-MTases by direct sequencing of bisulfite-modified DNA. pBR322 DNA was methylated with ΔL2-MTases at a MTase : GCG target sites ratio of 1:64 (**A**) or 1:4 (**B**) and subjected to bisulfite modification. The bisulfite-modified DNA was amplified by PCR and the obtained PCR product was directly sequenced. After the reaction with bisulfite and PCR unmethylated cytosines are converted to thymines; cytosines are only found in the positions that contain 5mC in the original sequence. Fragments of representative electrophoretograms are shown. GCG sites are shaded; GCA and GCT sites are framed by dotted line.

The specificity of L2Bsp and ΔL2-14 was investigated more thoroughly by cloning the PCR amplified fragments and sequencing individual clones. The methylation density at a particular site was determined as the ratio of number of methylated cytosines found to a total number of reads. The bisulfite sequencing data are summarized in **Figure 8**. In addition to GCG, L2Bsp methylated GCA sites quite efficiently and displayed several cytosines in GCC and GCT sites. Consistently with the direct PCR product sequencing data, only two out of nine GCG sites were methylated completely under these conditions.

5-methylcytosines were detected almost exclusively in GCN sites indicating that the third position of the target was the most degenerate. Analysis of DNA methylation by ΔL2-14 revealed only a tiny fraction of methylated cytosines outside the GCG targets (4 occurrences in 703 reads through GC[A/T/C] sites) (**Figure 8**). Comparison of relative methylation densities of GCG versus non-GCG sites, R_{spec}, again shows that the engineered MTase is approximately ten-fold more specific than the original L2Bsp hybrid (R_{spec} =170 and 17, respectively).

As the amount of MTases in the methylation reaction was increased (MTase : GCG ratio 1:4) some methylated cytosines were found outside the GCG sequences in DNA methylated by all variants (**Figure 7B**).

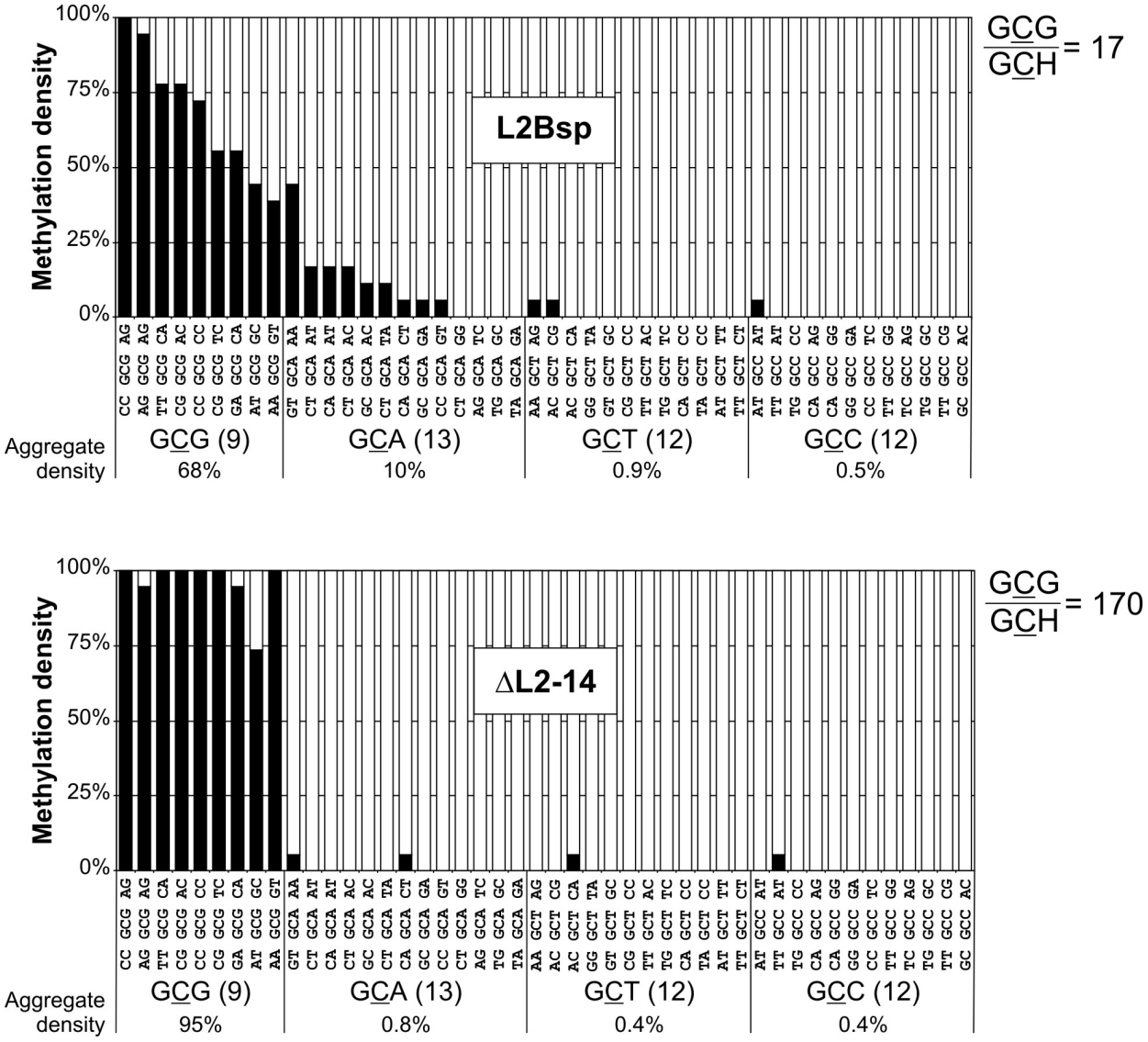


Figure 8. Bisulfite sequencing analysis of *in vitro* specificity of the L2Bsp and ΔL2-14 MTases. pBR322 DNA was methylated with L2Bsp (top) or ΔL2 14 (bottom) at a MTase to GCG target sites ratio of 1:64 and subjected to bisulfite modification. Methylation densities at individual 46 GCN sites were determined by sequencing of a 538 nt pBR322 fragment in 18-19 individual clones obtained after cloning the bisulfite-converted DNA. The methylation density is expressed as a ratio of methylated cytosines observed to a total number of sequence reads. H – adenine, cytosine or thymine.

1.4. Kinetic characterization of selected MTases

In order to better understand the observed changes in catalytic efficiency and specificity, kinetic parameters were determined for L2Bsp, the three ΔL2-mutants and WT M.HhaI. Steady-state kinetic analyses were carried out with poly[dG-dC]·poly[dG-dC] DNA as previously described (Wu and Santi, 1987; Vilkaitis *et al.*, 2000). Our data presented in **Table 3** show that all variants, including the WT HhaI, have similar k_{cat} values (smaller than 4-fold difference), whereas the most prominent differences between the mutants are manifested in K_M^{DNA}. With its K_M^{DNA} of nearly four orders of magnitude higher than that of WT M.HhaI, L2Bsp clearly indicated its DNA binding capacity to be severely impaired. This notion is in a good agreement with our DNA binding studies, which showed no detectable MTase-DNA complexes in gel shift assays even in the presence of AdoHcy (not shown). Upon optimization of the Loop-2, the interaction between MTases and DNA has improved significantly, as K_M^{DNA} decreases in the order L2Bsp >> ΔL2-6 ≥ ΔL2-9 > ΔL2-14 (**Table 3**). The same trend was observed when the ability of MTases to form covalent complex with 5-fluorocytosine containing DNA was investigated: no denaturation-resistant complex was formed by L2Bsp even after overnight incubation. ΔL2-MTases were able to form such complex, albeit at lower efficiency than WT M.HhaI (**Figure 9**). Altogether, our experiments show that the ΔL2-14 variant is the most catalytically efficient GCG-specific MTase.

Table 3. Kinetic and thermodynamic parameters of WT and truncated variants of M.HhaI

MTase	K_D^{AdoMet}, μM	K_M^{AdoMet}, μM	K_M^{DNA}, nM	k_{cat}, min^{-1}	k_{cat} / K_M^{DNA}, $M^{-1}\cdot s^{-1}$
WT	4.2 ± 0.2	0.03 ± 0.01	0.17 ± 0.02	0.89 ± 0.05	$9\cdot10^7$
L2Bsp	6.1 ± 0.3	4.0 ± 0.5	1300 ± 200	0.27 ± 0.01	$3\cdot10^3$
ΔL2-6	8.4 ± 0.9	1.2 ± 0.1	13.5 ± 4.2	0.45 ± 0.01	$6\cdot10^5$
ΔL2-9	7.5 ± 0.8	1.3 ± 0.2	4.8 ± 1.4	0.23 ± 0.01	$8\cdot10^5$
ΔL2-14	6.9 ± 0.5	0.9 ± 0.1	1.7 ± 0.3	0.98 ± 0.03	$1\cdot10^7$

In addition to increased K_M^{DNA}, all mutants showed substantially higher K_M^{AdoMet} values as compared to WT M.HhaI (**Table 3**). This may appear somewhat surprising, as changes in the TRD are not expected to disrupt any protein contacts with the cofactor. To rule out this possibility, we determined $K_D^{AdoMet(binary)}$ by monitoring Trp41 fluorescence changes upon cofactor binding (Vilkaitis *et al.*, 2000). Indeed we found that $K_D^{AdoMet(binary)}$ is virtually unaffected by the deletion of the Loop-2 (**Table 3**) and is similar for all mutants and the WT M.HhaI. This suggested that the observed increase in K_M^{AdoMet} may be related to a faster release of cofactor from the closed ternary complex. It was previously shown that the rate limiting step in the catalytic cycle of M.HhaI is the dissociation of the ternary product complex (MTase-methylated DNA-AdoHcy), which leads to faster product formation in the first turnover (burst) as compared to the steady state rate (Vilkaitis *et al.*, 2001). We therefore, performed a similar pre-steady-state kinetic experiment (**Figure 10**), but observed no pre-steady-state burst with the ΔL2-14 mutant. Altogether, our findings indicate that the rate limiting step in this mutant is different from WT M.HhaI, consistently with an enhanced cofactor exchange in the ternary complex (Merkienė and Klimašauskas, 2005).

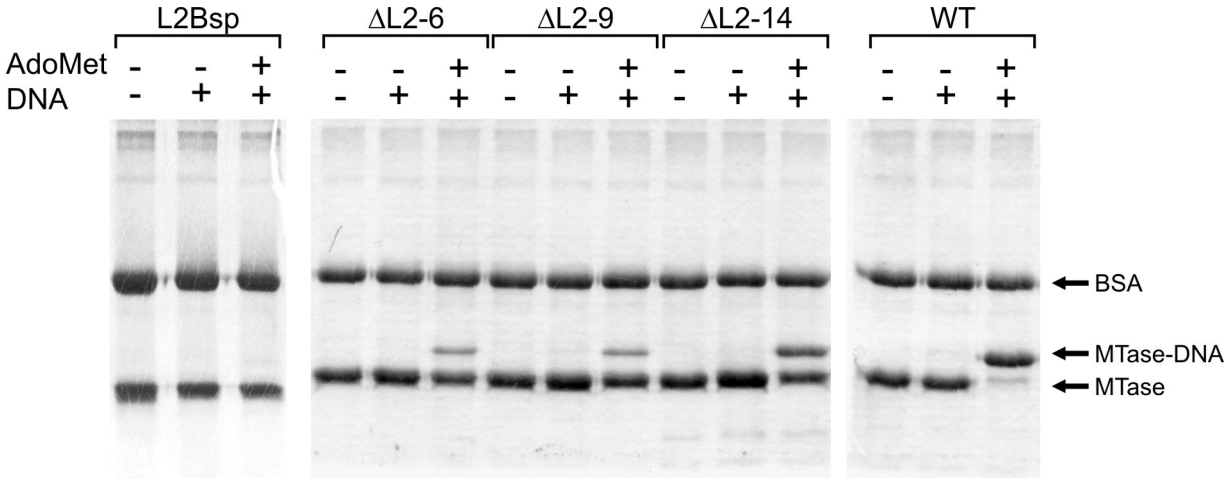


Figure 9. Interaction of WT and truncated M.HhaI variants with 5-fluorocytosine containing DNA. 5 μM of MTase was incubated with 5 μM GFGC/GFGC-13 DNA in the presence of 200 μM AdoMet (if any) in the Methylation buffer I at room temperature overnight. The MTase was inactivated by heating 5 min at 95ºC in the presence of SDS and the samples were resolved in the 10% polyacrylamide gel with SDS. Protein containing bands were visualized by Comassie staining.

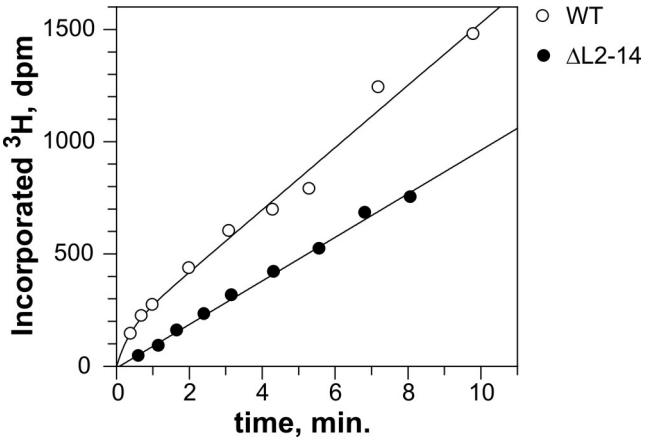


Figure 10. Pre-steady state burst analysis of M.HhaI. Reactions were performed in the methylation buffer at 37°C, containing 4 nM ΔL2-14 or WT M.HhaI, poly[dG-dC]·poly[dG-dC] DNA (500 nM double-stranded GCGC sites) and 10 μM [methyl-^{3}H]AdoMet (0.5 Ci/mmol). Aliquots were withdrawn at indicated time points, and incorporated radioactivity was measured by filter-binding assay as described (Vilkaitis *et al.*, 2000). Open circles, WT M.HhaI; closed circles, ΔL2-14.

1.5. DNA transalkylation using synthetic AdoMet analogs

Synthetic AdoMet analogs with extended sulfonium-bound propargyllic side chains have been used for methyltransferase-directed sequence-specific derivatization and labeling of DNA (Dalhoff *et al.*, 2006; Lukinavičius *et al.*, 2007). The novel approach, named mTAG, envisions many useful applications. However, as the side chain is extended in length, wild-type DNA MTases, such as M.HhaI, become increasingly inefficient. Simple steric engineering of the cofactor binding pocket (replacing bulky amino acids with Ala or Ser residues) enabled M.HhaI to use such cofactor analogs as alkyl donors (Lukinavičius *et al.*, 2007). Tyr254, which is part of the Loop-2, was one of such replacements around the cofactor pocket (Lukinavičius *et al.*, in preparation).

Meanwhile, the ΔL2-MTases inherently contain a Ser at this position, suggesting that they may be more active than the WT M.HhaI in the transalkylation reactions. Therefore, the ΔL2-14 mutant was tested for its ability to transfer a pentynyl chain from a synthetic cofactor analog, AdoPentyn (**Figure 11A**). DNA protection assays showed that the apparent alkylation rate of the engineered MTase is ~8 turnovers per hour, which is at least 100-fold improvement as compared to WT M.HhaI (compare **Figure 11B** and **C**), and is only ~8 fold lower than the rate observed with AdoMet (see **Figure 5**). Digestion of the modified DNA with R.Bsh1236I indicated that the sequence specificity of ΔL2-14 remained unaltered with both cofactors (**Figure 11C**). Similar results were obtained with the cofactor analog AdoHxGABANH$_2$, which carries an aliphatic aminogroup and is suitable for site-specific labeling of DNA (**Figure 11D**). Notably, WT M.HhaI does not catalyze the reaction with this cofactor analog (Lukinavičius *et al.*, in preparation). These experiments demonstrate that the designed MTase can catalyze an efficient transfer of extended linear chains to the GCG sites.

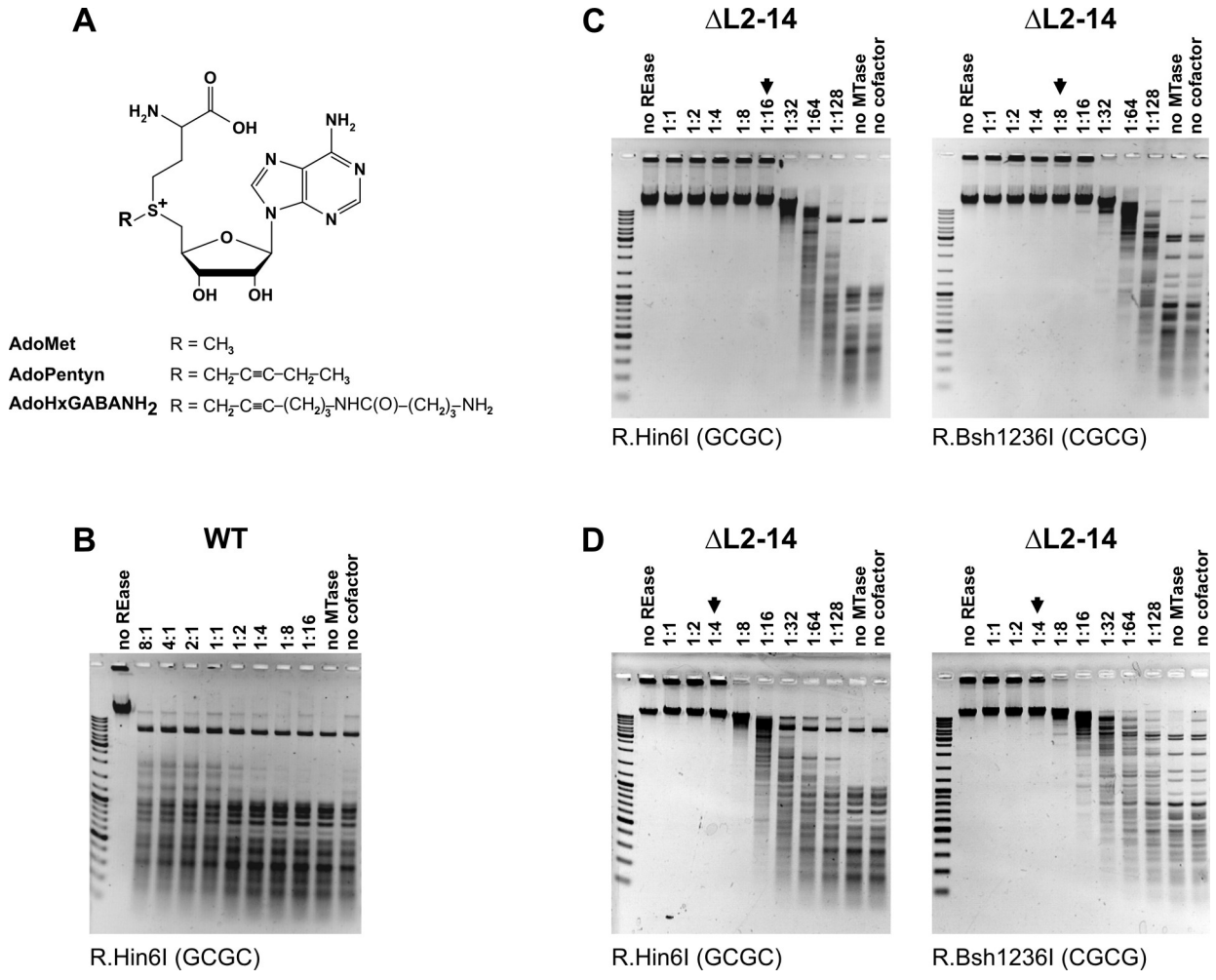


Figure 11. Enzymatic transalkylation of DNA using synthetic cofactor analogs. A – chemical structures of the AdoMet cofactor and its synthetic analogs. **B** – WT M.HhaI reaction with AdoPentyn. **C** and **D** – ΔL2-14 reactions with AdoPentyn and AdoHxGABANH$_2$, respectively. λ DNA was incubated with decreasing amounts (two-fold serial dilutions) of MTase in the presence of 300 μM cofactor analog for 1 hour at 37°C, then digested with R.Hin6I or R.Bsh1236I and analyzed by agarose gel electrophoresis. Numbers above lanes indicate molar ratios of MTases to their target

sites (GCGC or GCG). Arrows mark the minimal ΔL2-14 : GCG ratio sufficient to render a complete protection (end-point). This ratio corresponds to a number of apparent enzyme turnovers (h^{-1}).

Next, the possibility to use ΔL2-14 for the site-specific DNA labeling was investigated. In the presence of AdoHxGABANH$_2$ ΔL2-14 transfers the extended linear chain containing an aliphatic aminogroup from the cofactor analog to the DNA therefore allowing specific labeling of GCG sites. Various reporter groups may be coupled to the introduced aminogroups in a chemical reaction with commercially available NHS-esters. pUC19 plasmid DNA was labeled by incubation with AdoHxGABANH$_2$ and ΔL2-14, followed by reaction with Cy5-NHS ester. Then the DNA was digested with R.BamHI and R.PagI and analyzed by agarose gel electrophoresis. Fragments containing the fluorescent dye Cy5 were visualized by scanning with 635 nm laser (**Figure 12A**). Fluorescence intensity distribution in the obtained bands correlated well with the numbers of GCG sites in the corresponding restriction fragments (**Figure 12B**). Therefore ΔL2-14 MTase is suitable for the site-specific DNA labeling in the reactions utilizing cofactor analogs.

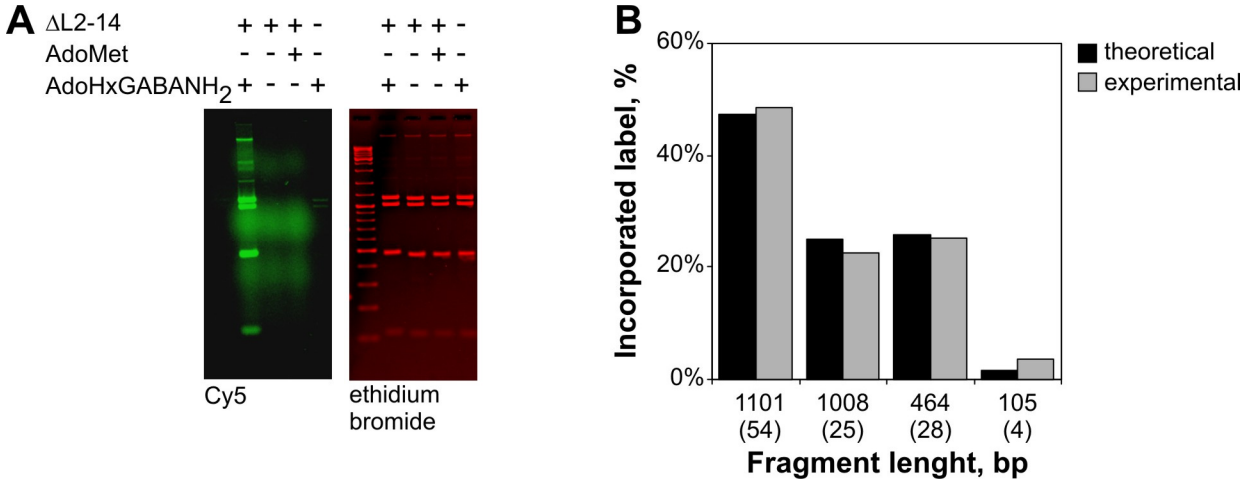


Figure 12. Two-step specific labeling of plasmid DNA at GCG sites. pUC19 plasmid DNA was amino-modified by incubation with ΔL2-14 in the presence of AdoHxGABANH$_2$ and then treated with Cy5-NHS ester. Labeled DNA was digested with R.BamHI and R.PagI and the resulting fragments were resolved by agarose gel electrophoresis. **A** (*left*) – visualization of Cy5-labeled DNA fragments with 635 laser; **A** (*right*) visualization of all DNA in ethidium bromine stained gel with 488 nm laser. **B** – theoretical (black bars) and experimentally-determined (grey bars) distribution of Cy5 fluorescence intensity in the DNA bands; number in parentheses indicate numbers of GCG sites in the corresponding fragment.

1.6. Discussion

It came as no surprise that elimination of two hydrogen bonds from the protein-DNA interface upon truncation of the Loop-2 resulted in a decreased affinity of the L2Bsp hybrid towards DNA. This is manifested by its poor catalytic efficiency, a lack of a detectable MTase-DNA complex band in gel-shift assays (not shown), and by increased K_M^{DNA}. Although such changes seem inevitable when an enzyme with a lower specificity is designed, a possible solution to this problem is to compensate for the lost

specific contacts by introducing new nonspecific contacts (or enhancing existing ones) between the DNA and the protein. This is well illustrated by the occurrence of an arginine residue in all selected ΔL2-MTases (see **Figure 1**). Although we cannot predict the exact conformation of the truncated Loop-2, the proximity of the added Arg to the phosphodiester backbone is likely to account for ~100-fold lower K_M^{DNA} values in ΔL2-6 or ΔL2-9 *versus* L2Bsp. Moreover, the most active variant, ΔL2-14, contained another substitution outside the randomized region, which resulted in a several-fold improvement in both k_{cat} and K_M^{DNA}. The latter mutation maps to the IX conserved motif, where either Arg or Lys is typically present in other DNA MTases. M.HhaI-DNA crystal structures show that Lys273 points towards the DNA although its positively charged nitrogen atom is ~5 Å apart from the phosphodiester backbone (not shown). The longer side chain of arginine may bring the positively charged moiety closer to the phosphate leading to an enhanced interaction with the DNA.

In addition to increased K_M^{DNA}, all the mutants showed 100-fold higher K_M^{AdoMet} values as compared to WT M.HhaI (**Table 3**). Higher K_M^{AdoMet} may result from destabilizing the closed conformation of the catalytic loop (residues 81–100 in M.HhaI) or otherwise enhancing accessibility of the cofactor pocket both leading to a faster cofactor exchange in the ternary reaction/product complex (Vilkaitis *et al.*, 2000; Merkienė and Klimašauskas, 2005). Inspection of the crystal structures (Klimašauskas *et al.*, 1994) indicates that such Loop-2 truncations and replacement of Tyr254 with a smaller residue (Thr or Ser) in particular, would surely create a wider solvent channel in the cofactor pocket. Moreover, these mutations remove a stabilizing contact between Gln82 in the catalytic loop and Tyr254 in TRD, which is likely to shift the equilibrium of the catalytic loop towards an open conformer. Since essentially no improvement of K_M^{AdoMet} occurred upon evolution of the truncated Loop-2 (only 4-fold lower values in ΔL2-MTases as compared to L2Bsp), one can conclude that either the lost structural feature cannot be recovered in the current structural framework, or further rounds of directed evolution under conditions of low AdoMet concentrations are required. From the point of practical utility, this parameter is quite satisfactory for both *in vivo* and *in vitro* applications. In addition, this endows the ΔL2-MTases with a valuable feature to accommodate AdoMet analogs in the active site for targeted transfer of extended groups to DNA (**Figure 11**) (Klimašauskas and Weinhold, 2007; Lukinavičius *et al.*, 2007).

The GCG specificity is unique as no natural C5-MTase is known to recognize this target. Methylation of this asymmetric target leads to the formation of hemimethylated CG sites, which are preferred substrates for eukaryotic maintenance DNA methyltransferases. Previously, *in vivo* methylated plasmid DNA obtained from cells overexpressing L2Bsp was used as a substrate to study the processivity of the mouse Dnmt1 MTase (Vilkaitis *et al.*, 2005). Thus in the context of recently reported specificity changes of M.SinI (Timar *et al.*, 2004) and M.HaeIII (Cohen *et al.*, 2004), the obtained GCG-MTase has a high potential to become a valuable molecular tool for studies of various aspects of eukaryotic DNA methylation. Although the catalytic efficiency (k_{cat}/K_m^{DNA}, see **Table 3**) of the designed MTase is ten-fold lower than that of the WT M.HhaI (Vilkaitis *et al.*, 2000; Youngblood *et al.*, 2006), it is comparably efficient or even supersedes certain C5-MTases such as M.HaeIII ($3 \cdot 10^4$) (Cohen *et al.*, 2004), M.SinI ($3 \cdot 10^5$) (Timar *et al.*, 2004) or M.SssI (10^4–10^5) (Subach *et al.*, 2006; Rathert *et al.*, 2007), some of which are widely used to produce methylated DNA molecules for epigenetic and biochemical studies. Its sequence fidelity *in vitro* is also comparable with

that of currently characterized WT C5-MTases (Cohen *et al.*, 2004; Timar *et al.*, 2004; Youngblood *et al.*, 2006). The methylation fidelity defined as the ratio of methylation of target/off-target sites is typically around two orders of magnitude when nano- to micromolar concentrations of enzyme is used (**Figure 5** and **8**). At these practically useful concentrations and typical K_M^{DNA} values in the nM range ([MTase] > K_M^{DNA}), the definition of fidelity as a ratio of k_{cat}/K_m^{DNA} values for specific over non-specific sites is not particularly informative, since the specificity is largely controlled by k_{cat} with little contribution from the DNA binding affinity.

In summary, this work represents the first example of enzyme engineering effort leading to a dual specificity change in a DNA methyltransferase. The newly designed MTase is an efficient sequence-specific enzyme that is able to use synthetic AdoMet analogs for the transfer of extended groups onto DNA. The endowed unique features envision useful practical applications of the designer MTase, such as i) *in vitro* or *in vivo* generation of hemimethylated CpG-sites for studies of maintenance methylation in mammals and ii) attaching larger chemical entities to DNA for synthesis of DNA-based nanoparticles.

2. Kinetic analysis of conformational transitions during the M.HhaI-catalyzed reaction

2.1. Direct observation of cytosine flipping and covalent bond formation

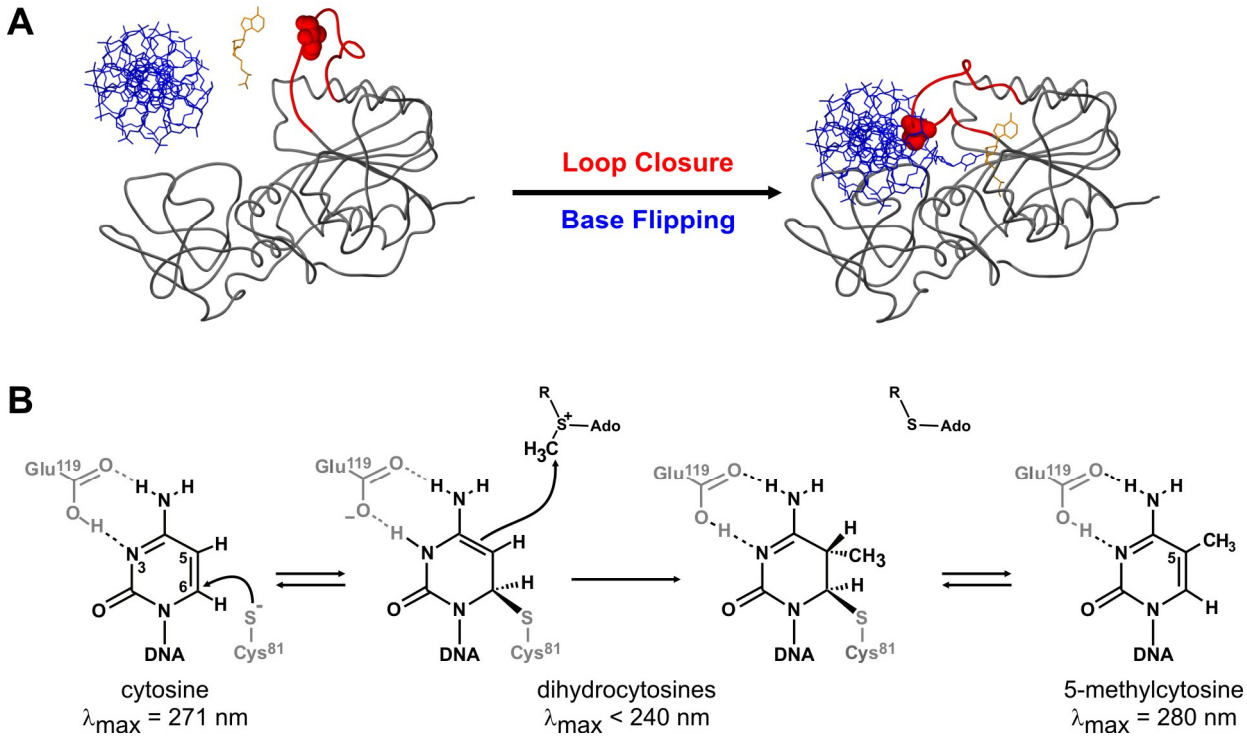


Figure 13. Conformational transitions and covalent catalysis by M.HhaI. A – upon binding of DNA and cofactor, M.HhaI flips its target cytosine out of the DNA helix into the active site; the catalytic loop in the protein makes a large motion to lock the target base and the bound cofactor. M.HhaI is shown as backbone trace (grey), the catalytic loop (residues 81–100) is red, the engineered Ile86 residue is shown as spacefill, DNA and cofactor are represented as sticks models in blue and yellow, respectively. **B** – the mechanism of catalytic target base activation and methyl group transfer by M.HhaI along with associated spectral changes of the target base.

Binding of M.HhaI to its GCGC target in duplex DNA induces flipping of the target cytosine base (underlined) out of the DNA helix with little disturbance to the rest of the duplex. Since the target cytosine is unstacked completely, and the two adjacent guanines are unstacked on one face (Klimašauskas *et al.*, 1994), we presumed that these changes might give a detectable hyperchromic change for direct observation of the flipping motion. Subsequent covalent catalysis by M.HhaI involves the formation of a covalent bond from the Cys81 residue to the C6 atom of the flipped out target cytosine (**Figure 13**). The attack at C6 disrupts the aromatic system in the cytosine ring and dramatically alters the spectral properties of the chromophore (λ_{max} shifts from 270 nm in cytosine to 225–240 nm in 5,6-dihydrocytosines) (Skaric *et al.*, 1974; Sverdlov *et al.*, 1974; Ulanov *et al.*, 1976). We therefore examined if the UV absorbance changes deriving from the target cytosine can be detected in a 25-mer duplex substrate using stopped flow techniques. Remarkably, a clearly defined A_{280} signal comprising a 25-

30 mAU increase followed by a 18-20 mAU decrease (in the background of total absorbance of ~1 AU) was obtained upon fast mixing of a 2.5 μM cognate hemimethylated 25-mer DNA duplex with an excess of M.HhaI (3 μM) and AdoHcy (100 μM) (**Figure 14A**) in a 10 mm cuvette. Control experiments with DNA duplexes containing an abasic residue replacing the target cytosine, a non-flippable 2-aminopurine : thymine base pair (Daujotytė *et al.*, 2004), a non-cognate target site all showed none or minor changes in absorbance (amplitudes of 3-6 mAU) (**Figure 14A**). A wavelength scan of the absorbance signal in the cognate system revealed a maximal change at about 280-290 nm (not shown). Our observation of a similar absorbance trace (**Figure 16C**) in the W41F mutant excludes the possibility that the observed signal originates from the unique tryptophan residue (W41) in the protein. The observed upward amplitude (**Figure 14A, Table 4**) represents a 150% hyperchromic change based on the absorbance of one 2'-deoxycytidine residue (derived from a molar extinction coefficient in neutral milieu of $\varepsilon_{280} = 7{,}200\ M^{-1}\cdot cm^{-1}$). Taking into account that a much smaller absorbance change is observed upon flipping of an abasic target residue one can conclude that the ascending transients in the absorbance traces largely represent the outgoing conformational transitions of the target nucleotide.

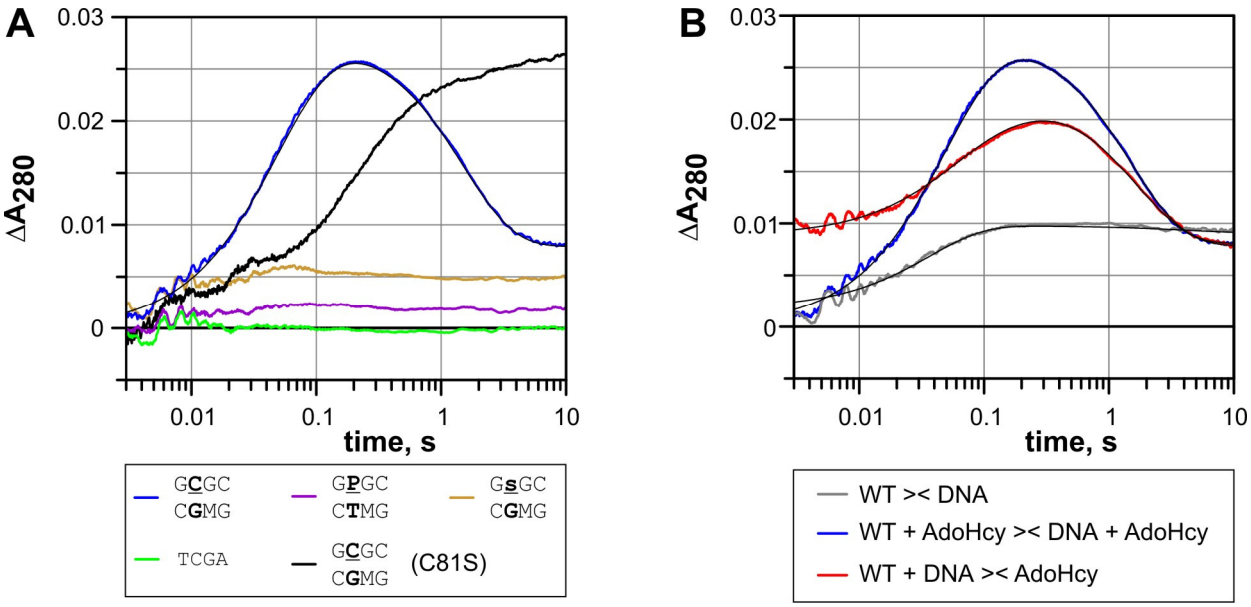


Figure 14. Stopped-flow absorption traces during M.HhaI-DNA interaction. Reactions contained 3 μM MTase, 2.5 μM DNA and 100 μM AdoHcy (final concentrations). **A** – WT M.HhaI was rapidly mixed with different DNA substrates in the presence of AdoHcy: blue – GCGC/GMGC (cognate target site), yellow – GsGC/GMGC (abasic target site), green – TCGA (non-cognate target site), purple – GPGC/GMTC (non-flippable target site); black – the catalytic C81S mutant rapidly mixed with GCGC/GMGC in the presence of AdoHcy. **B** – reaction profiles at different mixing orders: blue – MTase was rapidly mixed with GCGC/GMGC in the presence of AdoHcy, red – MTase was first premixed with GCGC/GMGC and then mixed with AdoHcy, grey – MTase rapidly mixed with G**C**GC/GMGC in the absence of cofactor. Parameters derived from fitting experimental traces to multiexponential equations are shown in **Table 4**.

Table 4. Kinetics of target base flipping and covalent bonding upon M.HhaI interaction with cognate DNA

Reaction	HhaI >< DNA		HhaI + AdoHcy >< DNA + AdoHcy		HhaI + DNA >< AdoHcy	
Parameters	**Rate, s^{-1}**	**Amplitude, mAU**	**Rate, s^{-1}**	**Amplitude, mAU**	**Rate, s^{-1}**	**Amplitude, mAU**
Cytosine flipping	27.4 ± 0.3 -	8.3 ± 0.04 -	19.1 ± 0.1 2.5 ± 0.2	28 ± 0.1 2.5 ± 0.1	19.7 ± 0.9 5.5 ± 0.4	7.6 ± 0.5 7.1 ± 0.4
Covalent bonding	0.23 ± 0.06	-0.7 ± 0.07	0.71 ± 0.003	-23 ± 0.1	0.60 ± 0.01	-16 ± 0.1

Apparent kinetic parameters derived from fitting absorption traces to multi-exponential functions for reactions involving 3 μM M.HhaI, 2.5 μM DNA and 100 μM cofactor as shown in **Figure 14B**. Premixed components in two chambers (separated by ><) were rapidly mixed in a stopped-flow cell to observe reaction-associated changes in absorbance. Values are reported with standard errors of the fit.

To confirm the nature of the descending absorbance change, we performed a control experiment with a catalytic mutant of M.HhaI, C81S, which lacks the catalytic thiol group required for the covalent bonding to the target cytosine (Mi and Roberts, 1993). As expected, the downward signal was absent in the mutant system (**Figure 14A**). The observed downward amplitude of 16-20 mAU agrees well with the loss of one cytosine chromophore; an equally high although significantly slower (0.0003 s^{-1}) absorbance change was also observed upon formation of an irreversible covalent M.HhaI-DNA complex involving 5-fluorocytosine (**Figure 15**) (Osterman *et al.*, 1988). This result agrees well with the published observation that M.HhaI reaction with 5-fluorocytosine is ~400 times slower than with cytosine. The descending part of stopped flow absorbance trace fitted reasonably well into a single-exponential decay function; the determined amplitude (13 mAU) corresponded to ~65% of cytosine being covalently bound (estimation based on the molar extinction coefficient ε_{280} = 8,000 $M^{-1} \cdot cm^{-1}$ (Tanaka *et al.*, 1981)). When this reaction was performed with ^{33}P-labeled DNA and the resulting products were resolved on the denaturing PAA gel, ~70% of DNA was found in the denaturation-resistant complex with protein (**Figure 15**), which is fully consistent with the stopped-flow results.

In the absence of cofactor, interaction with DNA gives similar transients but several-fold reduced amplitudes as compared to the ternary complex (**Figure 14B**). The remaining signal is regained upon addition of AdoHcy to the preformed M.HhaI-DNA complex. This suggests that in the binary complex, a fast equilibrium is achieved in which $\sim^1/_3$ of the target cytosine occurs in the flipped-out state. This observation is consistent with the previously described dynamic nature of the complex (Klimašauskas *et al.*, 1998) and catalysis of fast C5-hydrogen exchange in the absence of cofactor (Wu and Santi, 1987).

These findings clearly indicate that the descending transient is a reporter of the covalent bond formation in the M.HhaI-DNA complex. Altogether, this set-up thus permits a direct real-time observation of the target cytosine flipping and subsequent covalent bond formation by the HhaI methyltransferase in a chemically unperturbed system.

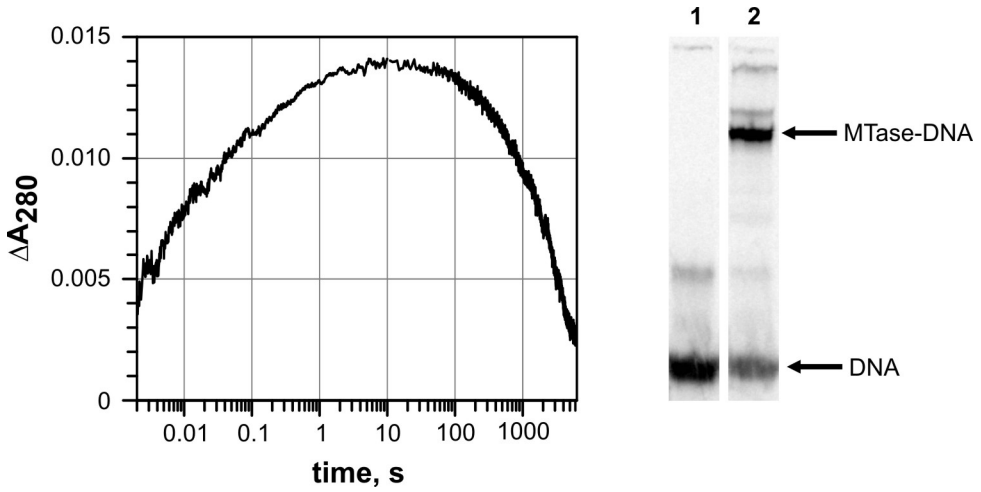


Figure 15. Stopped flow analysis of M.HhaI interaction with DNA containing 5-fluorocytosine. *Left*, 3 μM WT M.HhaI was rapidly mixed with 2.5 μM GFGC/GMGC duplex in the presence of 100 μM AdoMet at 25ºC and absorbance data collected for 10 s or 6000 s. *Right*, 3 μM WT M.HhaI was mixed with 2.5 μM ^{33}P-labeled GFGC/GMGC duplex in the presence of 100 μM AdoMet, reactions were incubated 27 h at 25ºC and the products were resolved by denaturing polyacrylamide gel electrophoresis. The DNA-containing bands were visualized with PhosphorImager (*Fuji*). **1, 2** – reactions without and with MTase, respectively.

2.2. Simultaneous sensing of the catalytic loop motion with built-in fluorophores

In addition to base flipping, M.HhaI itself undergoes major conformational changes during the catalytic cycle (Klimašauskas *et al.*, 1994). A large motion of the catalytic loop (residues 81–100), engulfs the bound DNA and the catalytic nucleophile Cys81 is brought in contact with the flipped out cytosine. To follow the motions of the catalytic loop during catalysis, we have redesigned the fluorophore locations in the HhaI methyltransferase by site directed mutagenesis of two positions. The unique intrinsic Trp residue (Trp41) in the cofactor binding pocket has been replaced by a Phe, whereas the bulky Ile86 residue in the catalytic loop was replaced with a Trp. The double W41F/I86W mutant showed similar kinetic parameters with the previously described W41F variant (**Table 5**) (Merkienė and Klimašauskas, 2005); it differed from the WT M.HhaI mainly in the lower affinity towards the cofactor AdoMet or its product AdoHcy, which can be compensated by increased cofactor concentration in the reaction. A similar behavior has recently been noted for another two engineered single-tryptophan variants of M.HhaI, W41F/K91W and W41F/E94W (Estabrook and Reich, 2006; Estabrook *et al.*, 2009).

Table 5. Kinetic parameters of WT M.HhaI and W41F I86W mutant

MTase	$K_M^{AdoMet\ (37°C)}$, μM	$k_{cat}^{(37°C)}$, s^{-1}	$k_{chem}^{(25°C)}$, s^{-1}
WT	0.035[a] ± 0.003	0.018[a] ± 0.001	0.12 ± 0.004
W41F/I86W	2.7[a] ± 0.4	0.056[a] ± 0.002	0.09 ± 0.004

[a] determined by dr. E. Merkienė

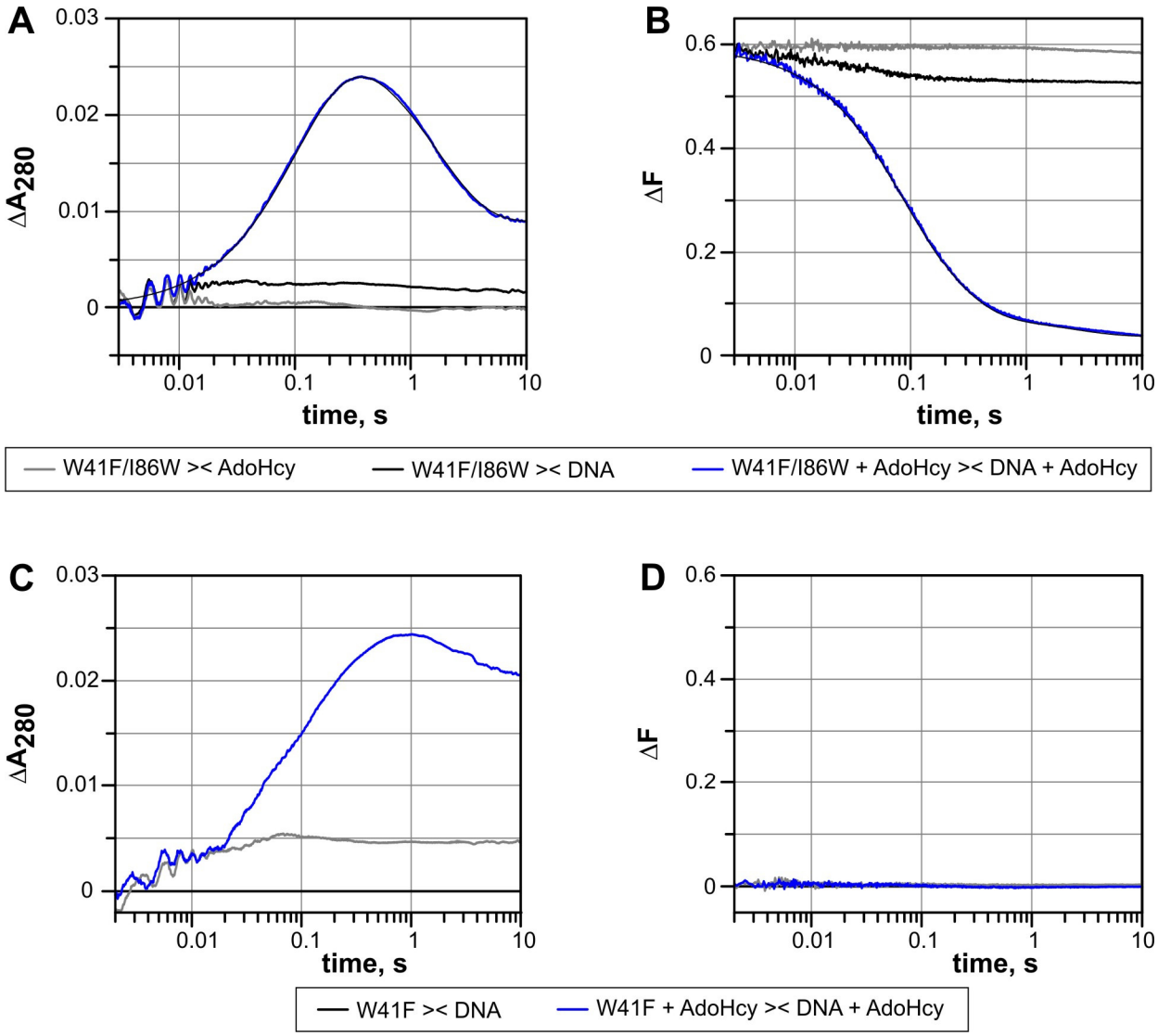


Figure 16. Flipping of the target cytosine and the closure of the catalytic loop in W41F/I86W mutant. A, B – 3 μM W41F/I86W was mixed with 100 μM AdoHcy (grey) or 2.5 μM GCGC/GMGC (black); 3 μM W41F/I86W was mixed with 2.5 μM GCGC in the presence of 100 μM AdoHcy (blue); **A** – absorption traces, **B** – fluorescence traces. **C, D** – 3 μM W41F was mixed 2.5 μM GCGC/GMGC in the absence of cofactor (grey) or in the presence 100 μM AdoHcy (blue). **C** – absorption traces, **D** – fluorescence traces. Parameters derived from fitting experimental traces to multiexponential equations are shown in **Table 6**. The fits are shown as thin black lines in **A** and **B**.

Stopped flow experiments with the W41F/I86W mutant M.HhaI showed a strong decrease in the Trp fluorescence (excitation at 280 nm, emission at > 335 nm) intensity upon formation of the ternary complex with the cognate DNA duplex and AdoHcy (**Figure 16B**). Null or small fluorescence changes were detected when either the DNA or cofactor was omitted from the reaction, indicating that the signal is dependent on the formation of the closed ternary complex (Klimašauskas and Roberts, 1995; Klimašauskas *et al.*, 1998; Estabrook *et al.*, 2009). The fluorescence signal was also absent in the catalytically active W41F mutant, which contains no tryptophan residues (**Figure 16D**). In the presence of the AdoMet cofactor, the fluorescence decrease was

followed by an ascending fluorescence signal (see below and **Figure 22B**), consistent with the reversal of the conformational change after the methylation reaction. Thus we conclude that the observed quenching of the tryptophan fluorescence derives from the closure of the catalytic loop whereby the Trp86 residue approaches the major groove of the bound DNA duplex; similarly with the previously reported K91W and E94W variants (Estabrook and Reich, 2006; Estabrook *et al.*, 2009), the engineered Trp86 residue is thus capable of reporting on conformational transitions of the catalytic loop through changes in fluorescence intensity. Although this engineered system differs slightly form the WT enzyme, it permits direct comparison of the base flipping and the loop motion. Incidentally, both the target cytosine and the Trp residue in the catalytic loop can be analyzed at the same excitation wavelength (typically 280 nm), and therefore both conformational changes can be simultaneously followed in real time as absorbance and fluorescence signals (emission at a 180º and 90º angle from the incident beam, respectively) in each individual stopped-flow measurement.

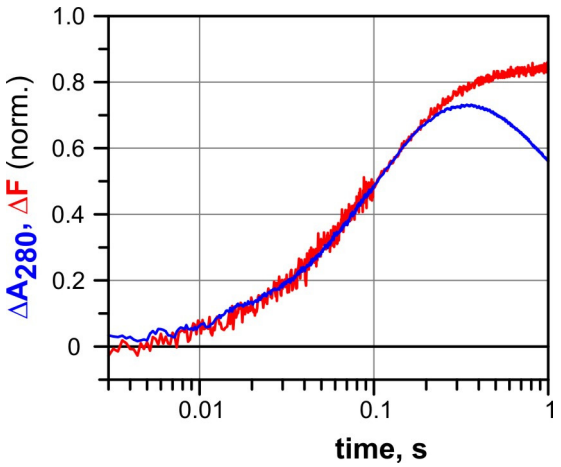


Figure 17. Flipping of the target cytosine is tightly coupled to the catalytic loop motion in W41F/I86W. 3 μM W41F/I86W was rapidly mixed with 2.5 μM GCGC/GMGC in the presence of 100 μM AdoHcy; absorption (blue) and fluorescence (red) traces were recorded simultaneously. The traces were normalized; the polarity of fluorescence was inverted for the demonstration purpose.

In reaction with cognate DNA and 100 μM AdoHcy the absorbance time course of the W41F/I86W mutant (**Figure 16A**) was similar to that of WT M.HhaI. The data fitted well into a three-exponential function yielding a slightly lower rates of base flipping (9 s^{-1} *vs* 19 s^{-1} for WT M.HhaI) whereas the rate of covalent bond formation was nearly identical with that of WT M.HhaI (0.64 s^{-1}). The fluorescence intensity also followed a three-exponential decay with similar apparent rates. Notably, absorption and fluorescence traces could be globally fitted into a three-exponential function with two identical rate constants, implying that cytosine flipping and closure of the catalytic loop proceed simultaneously (**Table 6**). This is well illustrated by the **Figure 17**, where simultaneously acquired absorbance and fluorescence traces are shown. The third transient in the fluorescence profiles showed substantially smaller amplitude as compared with the preceding two steps, and it was not fully synchronous with the covalent bond formation in the absorbance profile. Since this minor signal varied with DNA and M.HhaI preparations, its nature has not been investigated further.

Table 6. Kinetics of cytosine flipping and catalytic loop motion in Trp-engineered M.HhaI

Reaction	Target base pair = C:G Figure 16			Target base pair = C:T Figure 20		
$\chi^2 \times 10^5$	12.5			8.65		
Parameters	**Rate, s^{-1}**	**Amplitude, relative units**		**Rate, s^{-1}**	**Amplitude, relative units**	
		Abs.	**Fluor.**		**Abs.**	**Fluor.**
Cytosine flipping/ Loop closure	19.0 ± 1.1 6.9 ± 0.2	0.26 ± 0.03 0.81 ± 0.03	-0.32 ± 0.03 -0.57 ± 0.03	63 ± 2 19.6 ± 0.6	0.53 ± 0.02 0.44 ±0.02	-0.56 ± 0.02 -0.38 ± 0.02
Covalent bonding	0.64 ± 0.01/ 0.76 ± 0.05*	-0.74 ± 0.005	-0.10 ± 0.004*	0.60 ± 0.004/ 0.68 ± 0.06*	0.70 ± 0.001	-0.05 ± 0.001*

Apparent kinetic parameters derived from global fitting normalized absorption and fluorescence traces to three-exponential functions for reactions involving 3 μM M.HhaI, 2.5 μM DNA and 100 μM AdoHcy as shown in **Figure 16** and **Figure 20**. Values are reported with standard errors of the fit.

* unassigned change in fluorescence intensity.

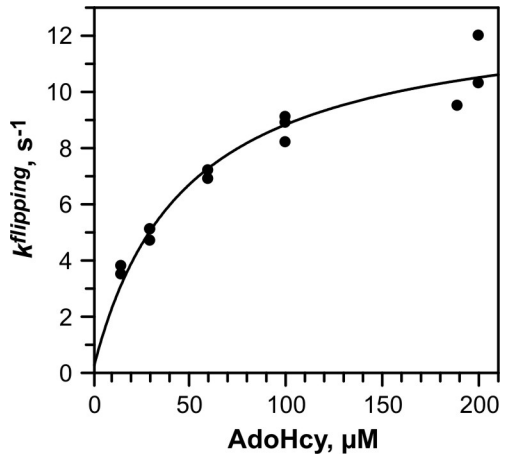
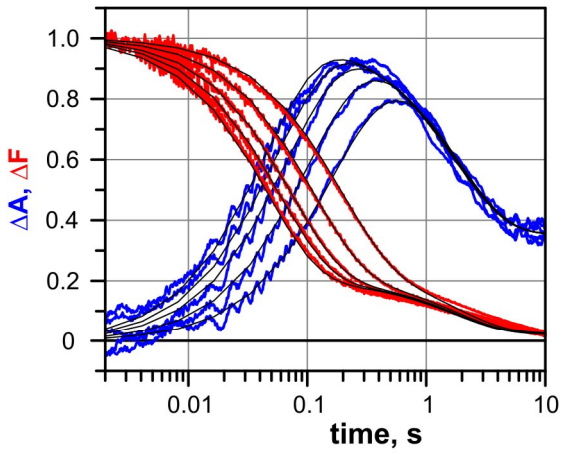


Figure 18. Rate of cytosine flipping as a function of AdoHcy concentration. 3 μM W41F/I86W was mixed with 2.5 μM GCGC/GMGC in the presence of 15-200 μM AdoHcy and the A_{280} was monitored. The rate of cytosine flipping was derived by fitting the obtained traces into two-exponential function.

Since the W41F/I86W mutant is somewhat impaired in cofactor binding, we investigated the possibility that the slower base flipping rate is caused by incomplete saturation of the W41F/I86W-DNA complex with AdoHcy. In contrast to the WT enzyme, the reaction profile showed a marked dependence on AdoHcy concentration up to the highest attainable values (15–200 μM). The cytosine flipping rate dependence on AdoHcy concentration followed a hyperbola, with the extrapolated maximum rate ~13 s^{-1} (**Figure 18**), which is only slightly slower than that of WT M.HhaI. The apparent K_D for the cofactor binding was estimated to be ~50 μM, which is higher than the K_M (~1 μM) and expected $K_D^{AdoHcy(binary)}$ (~20 μM, assuming identical to W41F). This may be rationalized by assuming that upon binding of AdoHcy to a binary HhaI-DNA complex a weak initial HhaI-DNA-AdoHcy complex is formed, in which the cofactor is only loosely bound, the catalytic loop is in an open conformation and the target cytosine is not yet inserted into the active site. Structurally this complex should resemble a weak complex with unspecific DNA. This initial complex isomerizes to form a highly stable complex with the closed catalytic loop and the target cytosine inserted into the active site.

The minimal mechanism which well describes the series of absorption and fluorescence traces is shown in **Figure 19**. According to this mechanism cytosine flipping and closure of the catalytic loop proceed simultaneously and two stable conformations with the flipped-out cytosine and the closed catalytic loop exist along the reaction pathway. The attempts to fit these data into more complex mechanism did not yield satisfactory fits. However it is still possible that this simplest scheme does not fully describe the reaction catalyzed by WT M.HhaI, as only one of the several possible routes may dominate the reaction catalyzed by the W41F/I86W mutant.



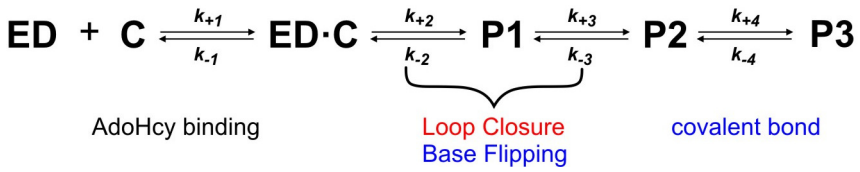


Figure 19. Detailed analysis of W41F/I86W and DNA conformational transitions in the reaction with AdoHcy. A preformed W41F/I86W (3 μM) complex with GCGC/GMGC (2.5 μM) was mixed with 15-200 μM AdoHcy. *Top* – normalized absorption (blue) and fluorescence (red) traces with the global fit (black lines) into the mechanism, shown *below*. Data fitting was performed using Dynafit 3.28.059 software (Kuzmic, 1996). ED – binary complex MTase-DNA, ED·C – a weak initial ternary complex MTase-DNA-AdoHcy, in which the target cytosine is stacked in the DNA helix and the catalytic loop is open. P1-P3 – transient conformations of the complex where the catalytic loop is closed and the target cytosine is flipped-out (P1-P2) or flipped-out and covalently bound (P3) to the Cys81. The derived rates are shown in the **Table 7**.

Table 7. Rate constants derived from fitting W41F/I86W reaction into the minimal reaction mechanism

	AdoHcy binding (i=1)	Cytosine flipping/ Loop closure (i=2, i=3)		Covalent bonding (i=4)
k_{+i}	29 $\mu M^{-1} \cdot s^{-1}$	29 s^{-1}	14 s^{-1}	3.2 s^{-1}
k_{-i}	1820 s^{-1}	0.2 s^{-1}	70 s^{-1}	0.07 s^{-1}

2.3. Interaction with mismatched DNA substrates

To better understand the relationship between the base flipping and the catalytic loop motion, we investigated interactions of the W41F/I86W mutant with mispaired substrates in the presence of AdoHcy. It has been reported that M.HhaI and other DNA MTases bind strongly DNA substrates containing mismatched bases at the target position and the increased affinity of M.HhaI towards mismatched DNA was correlated to the decreased stability of the base pair to be disrupted (Klimašauskas and Roberts, 1995; Daujotytė *et al.*, 2004). Using different partner bases we examined how the motions of the catalytic loop and the target cytosine are influenced by the nature of the target base pair. We found that the rate of cytosine flipping as well as the rate of the closure of the

catalytic loop increased as the stability of the target base pair decreased (**Figure 20A, B**). As with the cognate substrate, the base flipping and loop closure could be fitted with identical rate constants (**Table 6**), suggesting a synchronized nature of both processes. Notably, no closure of the catalytic loop was observed with the GPGC/GMTC DNA, containing a non-flippable target base pair.

An increased rate in the reaction with mismatched substrates was also observed in the case of WT M.HhaI, when base flipping and covalent complex formation was monitored (**Figure 21**, **Table 8**). In addition to the dominant flipping rate observed with correctly paired substrate (19 s^{-1}), a faster step was detectable with mismatched target base pairs (130 s^{-1} for C:T and 154 s^{-1} for C:C base pairs, respectively). These results agree well with a model in which a mismatched base pair promotes the accumulation of the target cytosine in a flipped-out state which facilitates the transition of the complex into the fully closed conformation.

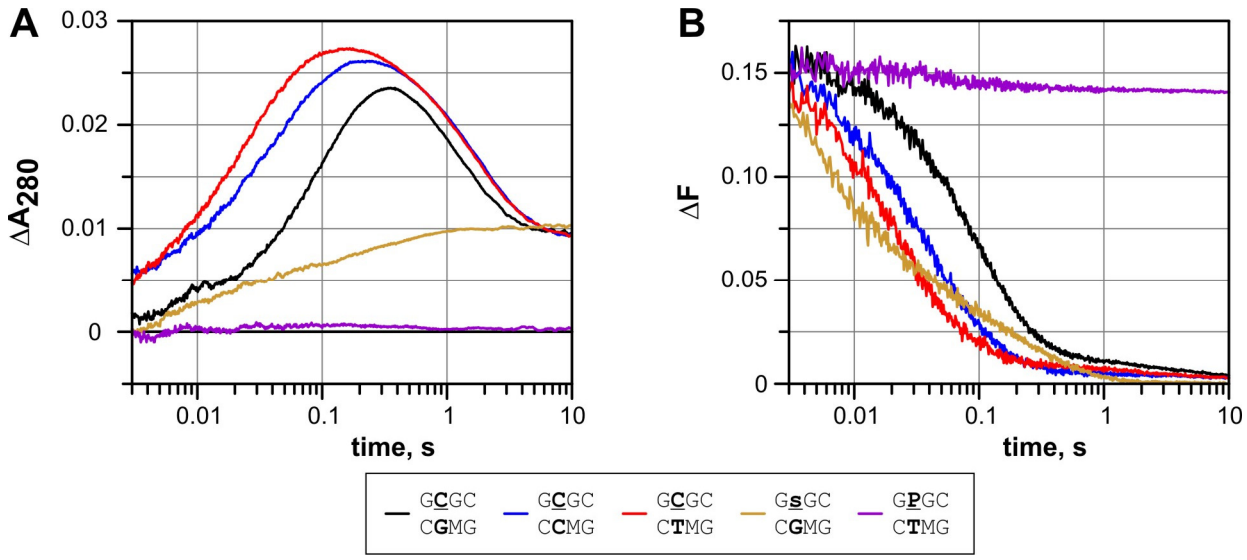


Figure 20. Interaction of W41F/I86W with mismatched DNA substrates. 3 μM W41F/I86W was mixed with 2.5 μM DNA in the presence of 100 μM AdoHcy. **A** – stopped-flow absorption traces, **B** – stopped-flow fluorescence. Black – GCGC/GMGC (target base pair C:G), blue – GCGC/GMCC (target base pair C:C), red – GCGC/GMTC (target base pair C:T), yellow – GsGC/GMGC (abasic target site), purple – GPGC/GMTC (non-flippable target site).

The fastest closure of the catalytic loop was observed upon interaction with a DNA duplex containing abasic target site (initial rate > 200 s^{-1}, **Figure 20A, B**). The rate of the loop closure with this substrate was essentially cofactor independent, which is consistent with the previously noted ability of M.HhaI to form highly compact binary complexes with abasic DNA substrates (Wang *et al.*, 2000). This suggests that the motion of the catalytic loop from an open to a closed conformation is potentially faster than is observed with natural DNA substrates, and is largely limited by the target nucleotide flipping.

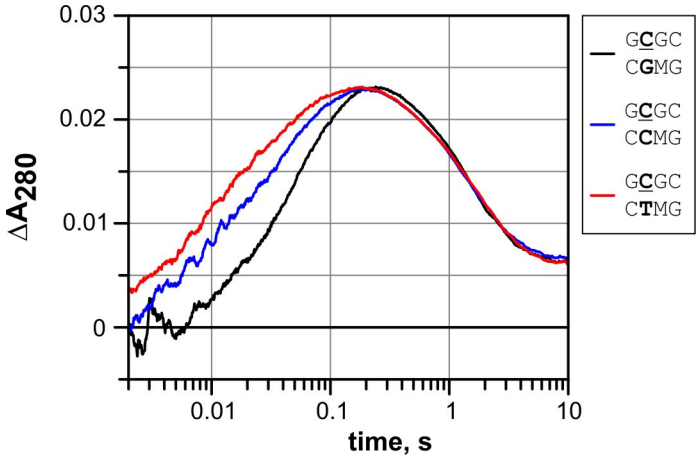


Figure 21. Interaction of WT M.HhaI with mismatched DNA substrates. 3 μM WT was mixed with 2.5 μM DNA in the presence of 30 μM AdoHcy and changes in A_{280} were monitored. Black – GCGC/GMGC (target base pair C:G), blue – GCGC/GMCC (target base pair C:C), red – GCGC/GMTC (target base pair C:T). Parameters derived from fitting experimental traces to three-exponential equation are shown in **Table 8**.

Table 8. Kinetics of WT M.HhaI interaction with the mismatched DNA substrates

Target base pair	C:G		C:C		C:T	
Parameters	**Rate, s^{-1}**	**Amplitude, mAU**	**Rate, s^{-1}**	**Amplitude, mAU**	**Rate, s^{-1}**	**Amplitude, mAU**
Cytosine flipping	19.5 ± 0.7 5.5 ± 1.1	24 ± 1 6 ± 1	154 ± 5 19.1 ± 0.2	8 ± 0.1 19 ± 0.1	130 ± 2 18.6 ± 0.2	13 ± 0.1 13 ± 0.1
Covalent bonding	0.70 ± 0.02	-22 ± 0.4	0.64 ± 0.004	-19 ± 0.04	0.60 ± 0.002	-19 ± 0.02

Apparent kinetic parameters derived from fitting absorption traces shown in **Figure 21** to three-exponential functions are reported with standard errors of the fit.

2.4. Ternary complexes with AdoMet: the complete reaction cycle

In the presence of AdoMet, the reaction can proceed through the full catalytic cycle yielding the ultimate reaction products: methylated DNA and. The initial steps, the target base flipping and covalent cytosine activation, can be readily discerned in the absorbance and fluorescence profiles of the reaction, since they show similar signals and apparent rates as observed in the corresponding AdoHcy profiles (**Figure 22B, Table 9**). As observed in the reaction with AdoHcy, cytosine flipping and loop closure proceeded simultaneously. Consequently the absorption and fluorescence traces could be fitted into the four-exponential function with common rate constants.

In addition to the processes observed with AdoHcy, an ascending transient of ~0.1 s^{-1} is observed in both the fluorescence and absorbance traces. Obviously, the reverse fluorescence change of matching amplitude can be safely assigned to the opening of the catalytic loop after the catalytic reaction. However the interpretation of a much smaller increase in absorbance is not as straightforward, mostly because the methyl transfer step itself does not give a detectable optical change. Although the C5-methylation of cytosine results in a substantial bathochromic shift (see **Figure 13B**), it remains obscure until the subsequent β-elimination step occurs in which the transient covalent bond is broken and the aromatic system of the cytosine ring is restored. However, subsequent back-flipping of the methylated target base in the presence of high

concentrations of AdoMet (Estabrook *et al.*, 2009) leads to a downward signal due to the hypochromic effect, which nearly cancels out the upward absorbance changes. Altogether, a single step with a small, AdoMet concentration-dependent increase in absorbance is observed at 280 nm (**Figure 23**). Since the observed transient closely follows the chemically determined methyl transfer step and is coincident with the major upward transient in the fluorescence trace (reopening of the catalytic loop) in the Trp-engineered M.HhaI and (see **Figure 22B** and **Table 9** and **Table 5.** Kinetic parameters of WT M.HhaI and W41F I86W mutant), it appears that the above three steps occur nearly simultaneously and cannot be resolved by the present data. Therefore, combining absorption, fluorescence and chemical methylation data, the four transients can be reliably characterized in the engineered M.HhaI variant.

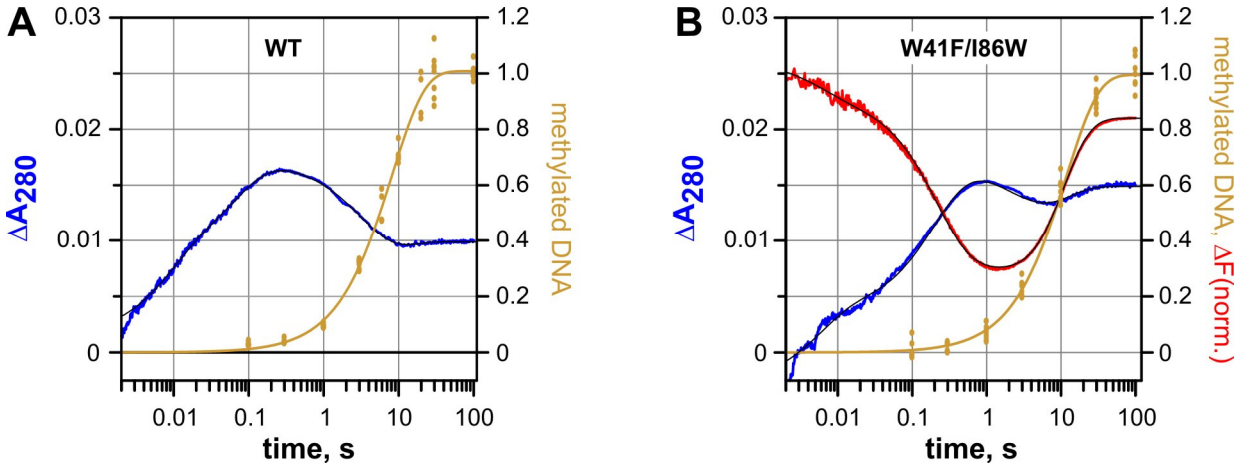


Figure 22 Optical and covalent changes observed during catalytic turnover of M.HhaI. WT (**A**) or W41F/I86W (**B**) variant of M.HhaI (3 μM) was mixed with 2.5 μM cognate DNA duplex in the presence of 100 μM AdoMet. Stopped-flow absorbance traces are shown in blue, fluorescence traces are shown in red; the catalytic transfer of methyl groups (yellow) was measured under single turnover conditions using [*methyl*-^{3}H]-AdoMet. Absorption traces are shown in actual scale, fluorescence and single turnover traces are normalized to unity. Parameters derived from fitting experimental traces to exponential equations are shown in **Table 9**.

Table 9. Kinetics of WT M.HhaI and W41F/I86W reaction in the presence of AdoMet

	WT + AdoMet >< DNA AdoMet		W41F/I86W + AdoMet >< DNA + AdoMet		
	Rate, s^{-1}	Amplitude, mAU	Rate, s^{-1}	Amplitude, rel. u.	
				Abs.	Fluor.
Cytosine flipping/ loop closure	120 ± 3 14.4 ± 0.2	7.1 ± 0.1 8.5 ± 0.1	180 ± 5 4.68 ± 0.03	0.23 ±0.004 0.46 ± 0.002	-0.13 ± 0.004 -0.59 ± 0.002
Covalent binding	0.26 ± 0.003	-12.8 ± 0.4	0.36 ± 0.01	-0.23 ± 0.01	-0.44 ± 0.02*
Methyl transfer/ flip-back	0.16 ± 0.004	5.6 ± 0.4	0.12 ± 0.002	0.16 ± 0.01	0.95 ± 0.02

Apparent kinetic parameters derived from fitting traces depicted in **Figure 22** to multi-exponential functions are reported with standard errors of the fit.

* unassigned change.

The WT enzyme shows a very similar absorption profile with four-exponential kinetics (**Figure 22A, Table 9**). The slowest transient again closely resembles the rate of methyl transfer determined under single-turnover conditions in a rapid-quench experiment (**Table 5**). Although the amplitude of the slowest transient is quite low at 280 nm, traces collected at 290–300 nm give somewhat higher upward signals (due to a higher spectral difference between C and 5mC), permitting a fairly reliable kinetic assignment of this final step. Therefore methyl transfer closely follows, but is not coincident with the covalent activation of the target cytosine; on the other hand, β-elimination is kinetically indistinguishable from the opening of the catalytic loop and restacking of the 5-methylcytosine after the reaction.

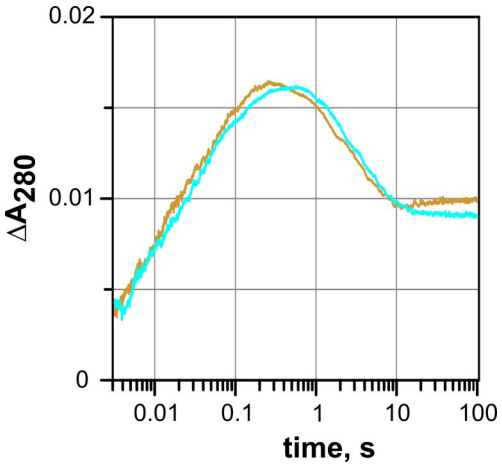


Figure 23. Cofactor concentration dependence of M.HhaI interactions with cognate DNA. 3 μM WT M.HhaI was mixed with 2.5 μM GCGC/GMGC the presence 30 μM (cyan) or 100 μM (yellow) AdoMet.

2.5. Discussion

Here we provide the first direct real-time observation of DNA base flipping in an unperturbed enzymatic system at physiological conditions. Since the approach is based on the general phenomenon of hyperchromicity, it is thus applicable to many other base-flipping systems. Although such subtle absorbance changes may not always permit a unequivocal discrimination between base flipping and other distortions of DNA or RNA duplexes that involve unstacking of nucleobases (such as DNA kinks or RNA strand looping (Moller *et al.*, 1979)), there is no inherent limitation with respect to the nature of the target base analyzed.

To our knowledge, the hyperchromic effect of a single nucleotide flipping in DNA has not been experimentally characterized before. It's magnitude (150% at 280 nm) turned out to be 2–3-fold higher than the average contribution of a disrupted base pair (60–80% at 260 nm), suggesting that a single base gap in the double helical stack may impact its electronic structure to a larger extent than can be anticipated from the nearest-neighbor effects alone (Kelley and Barton, 1999; Rajski *et al.*, 1999). As the signal largely originates from the loss of stacking interactions between the target cytosine and the adjacent guanines, it is important to note that in general these effects are largest when A or G are involved as neighbors (Cantor *et al.*, 1970; Cavaluzzi and Borer, 2004) and that C:G base pairs have a maximum hypochromicity at 280 nm (Puglisi and Tinoco, 1989). In addition, the absorbance of the flipped-out cytosine can be influenced

by the protein environment since it is known that the N3 protonation of the cytosine leads to a 9 nm red shift and a 1.5-fold increase in molar absorbance (ε_{max} = 9,000 $M^{-1} \cdot cm^{-1}$ at λ_{max} = 271 nm under physiological conditions, ε_{max} = 13,200 $M^{-1} \cdot cm^{-1}$ at λ_{max} = 280 nm at pH < 3.0 (Fox and Shugar, 1952)). Crystal structures, biochemical and computational studies of M.HhaI-DNA complexes indeed suggest extended protonation and/or hydrogen bonding of the N3 and at O2 positions of the cytosine ring in the catalytic site (Klimašauskas *et al.*, 1994).

The observed base flipping in the ternary complex involving M.HhaI, cognate DNA and cofactor generally occurs at a rate of 10-20 s^{-1}. Previous studies of *N*6-adenine MTases EcoRI and Ecodam (Allan *et al.*, 1999; Liebert *et al.*, 2004) as well as the HhaI (Vilkaitis *et al.*, 2000) MTases using 2-aminopurine fluorescence have not provided reliable estimates for the outgoing motion of the natural target nucleotide. The base flipping is synchronous with the motion of the catalytic loop in the M.HhaI mutant (**Figure 17** and **Figure 22**), and is very likely to be the same in the WT enzyme. Indirectly, similar estimates for the rate of loop closure were derived from the 2-aminopurine flipping experiments (Vilkaitis *et al.*, 2000) as well as observed using other M.HhaI catalytic loop mutants (Estabrook *et al.*, 2009). The concerted flip-lock motion is consistent with the notion that efficient extrahelical trapping of the flipped out base occurs only within a fully assembled catalytic site (Klimašauskas and Roberts, 1995; Klimašauskas *et al.*, 1998). Although the intrinsic rate of the loop closure can be higher with abasic DNA substrates, the assembly of the closed conformer is dependent on the strength of the target base pair. It is thus possible that the accumulation of a (partially) unpaired base-flipping intermediate would give rise to a faster base flipping phase, thereby explaining biphasic kinetics of the flip-lock motion. On the other hand, the biphasic behavior may also derive from an intermediate stable conformer in the trajectory of the catalytic loop or the target cytosine. Notably, crystal structures of the open and closed protein show extensive internal rearrangements in the twenty-residue loop during the locking motion (Klimašauskas *et al.*, 1994).

The established tight coupling between the target base flipping and the catalytic loop closure is consistent with two alternative models for the M.HhaI-induced base flipping. The first model suggests that the target base flipping occurs via the minor groove of the DNA, which then induces the catalytic loop closure (Klimašauskas and Roberts, 1995). This hypothesis is largely based on steric considerations (Klimašauskas *et al.*, 1994), and solvent accessibility of the target nucleotide replaced with mismatched bases such 2-aminopurine or thymine (Serva *et al.*, 1998; Vilkaitis *et al.*, 2000). Alternatively, a molecular dynamics study suggested that the closure of the catalytic loop upon the bound DNA would induce target base flipping via the major groove inside the protein (Huang *et al.*, 2003); this point of view gained some indirect support from crystallographic studies (Horton *et al.*, 2004). Although the Ser87 residue of M.HhaI implicated in the base-pair opening turned out to be dispensable for both base flipping and catalytic activity (Daujotytė *et al.*, 2004), this counter-intuitive (due to steric constrains to base flipping inside the closed ternary complexes) model cannot be decisively excluded or confirmed by the present data either.

Our first direct observation of the covalent bond formation to the target cytosine in solution provides evidence on several important aspects of the reaction mechanism. First, the observed rate of the covalent bonding in the presence of AdoHcy or AdoMet (0.3-0.7 s^{-1}) is considerably slower than the cytosine flipping and locking on the ternary

complex. This suggests that the formation of the covalent activated intermediate in the cognate reaction complex is largely dependent on the intrinsic rate of nucleophilic addition of Cys81, rather than the rate of cytosine locking in the catalytic site. However, the cytosine locking may be an important mechanistic factor that contributes to the catalytic rate of the MTase upon interaction with near-cognate or non-specific DNA substrates (Svedruzic and Reich, 2004; Estabrook *et al.*, 2009).

In the presence of product AdoHcy, the reaction reaches a dead-end equilibrium state. The extent of the covalent bond that is formed under these conditions has not been known. Numerous X-ray structures of the ternary M.HhaI complex with AdoHcy show that the bond length between the catalytic sulfur and the C6 atom of the cytosine is around 2.6–2.8 Å (Kumar *et al.*, 1997). This represents a weighed average between the length of a covalent C-S bond (1.8 Å) observed in a irreversible covalent complex involving 5-fluorocytosine (Klimašauskas *et al.*, 1994)) and a van der Waals distance (3.6 Å) (Kumar *et al.*, 1997), suggesting a partial (approximately 50%) formation of the covalent bond in the crystals. We find that the downhill amplitude of the absorbance signal in the presence of AdoHcy, which reflects the covalent bond formation (**Figure 14A, Table 4**), approximately matches the absorbance value of one cytosine chromophore at physiological conditions. Taking into account that the spectral properties of the bound cytosine may be affected (prior to the covalent bond formation) by the interactions with residues in the active site (see discussion above), the extent of covalent bonding to the target base in the ternary complex in solution can be estimated to be in the range of 70–100%.

In the presence of AdoMet, the covalent complex undergoes subsequent transfer of the methyl groups from the cofactor. We observe that the formation of the covalent bond temporally precedes the methyl transfer step, i.e. the two events appear as distinct kinetic steps in the reaction cycle (**Figure 22**, **Table 9**). These findings agree with a recent quantum-mechanical (QM) study (Zangi *et al.*, 2010) and disprove a QM calculations-derived concerted mechanism with simultaneous catalytic activation and methyl transfer steps (Zhang and Bruice, 2006).

CONCLUSIONS

1. The target specificity of M.HhaI G<u>C</u>GC has been changed to G<u>C</u>G by functional elimination of the recognition Loop-2 using rational protein design and directed evolution approaches.
2. Impaired catalytic efficiency of a methyltransferase due to eliminated sequence-specific contacts to the target nucleobases, can be largely compensated by enhancing non-specific protein contacts to the DNA backbone.
3. The catalytic efficiency and sequence fidelity of the engineered GCG-specific hemimethylase are comparable to those of other wild-type MTases and are thus sufficient for its application in molecular biology experiments.
4. The GCG-specific MTase catalyze the transfer of extended groups from synthetic AdoMet analogs onto DNA and therefore can be used for site-specific labeling of or hemimethylation of CG sites in natural DNA molecules.
5. Monitoring small changes in UV absorbance of native DNA allows direct observation of cytosine flipping out of DNA double helix by hyperchromicity and its covalent activation in the active site of M.HhaI by the loss of the cytosine chromophore.
6. In the reaction catalyzed by M.HhaI the flipping of the target cytosine and the closure of the catalytic loop proceed simultaneously, with the average rate of 10-20 s^{-1}. The rate of these processes highly depends on the strength of the target base-pair.
7. The covalent activation of the target cytosine (apparent rate 0.3-0.6 s^{-1}) and the transfer of the methyl group (apparent rate 0.1-0.2 s^{-1}) are distinct kinetic steps in the catalytic cycle of M.HhaI.

LIST OF PUBLICATIONS

The doctoral dissertation is based on the following original publications:

1. **Gerasimaitė, R.**, Merkienė, E. and Klimašauskas, S. (2011). Direct observation of cytosine flipping and covalent catalysis in a DNA methyltransferase. *Nucleic Acids Res* DOI: gkq1329 [pii] 10.1093/nar/ gkq1329.
2. **Gerasimaitė, R.**, Vilkaitis, G. and Klimašauskas, S. (2009). A directed evolution design of a GCG-specific DNA hemimethylase. *Nucleic Acids Res* **37**(21): 7332-7341.

Conference abstracts

1. **Gerasimaitė, R.**, Merkienė, E. and Klimašauskas, S. (2010). Direct real-time observation of cytosine flipping and covalent catalysis in a DNA methyltransferase. Abstracts of Chemical Biology 2010, Heidelberg, Germany; 22-25 September 2010; p. 82
2. **Gerasimaitė, R.** and Klimašauskas, S. (2007). Novel approaches to engineering sequence-specificity of DNA cytosine-5 methyltransferases. Abstracts of 32nd FEBS Congress, Vienna, Austria; 7-13 June, 2007; *FEBS J* **274** (Suppl. 1) p. 259.
3. **Gerasimaitė, R.** and Klimašauskas, S. (2006). DNA Methyltransferases with novel target specificity. Abstracts of the 9th Meeting/Conference of Lithuanian Biochemical Society, Molėtai, Lithuania; 16-18 June 16-18 2006; p 34.

Other publications

Klimašauskas, S., **Gerasimaitė, R.**, Vilkaitis, G. and Kulakauskas, S. (2002). N4,5-dimethylcytosine, a novel hypermodified base in DNA. *Nucleic Acids Res Suppl*(2): 73-74.

ACKNOWLEDGEMENTS

I thank prof. Saulius Klimašauskas for the opportunity to work in the Laboratory of Biological DNA Modification, his guidance and scientific discussions. I thank the coauthors of the publications dr. Eglė Merkienė and dr. Giedrius Vilkaitis for sharing their knowledge and the fruitful collaboration.

I am especially grateful to my husband and colleague dr. Gražvydas Lukinavičius for his support and understanding; for a kind gift of cofactor analogs, M.HhaI C81S protein and always insightful scientific discussions. I thank my former students Miglė Tomkuvienė and Anastasija Chomič for their inspiring optimism and stimulating questions. I would like to thank all the colleagues from the Laboratory of Biological DNA Modification for their valuable advices, suggestions and moral support at different stages of this project. Special thanks for Zdislav Staševskij for the scientific advices and extremely effective help in solving all the possible technical problems.

I thank Goda Mitkaitė for constructing the *E. coli* ER2566 *lacI*^{q} strain and Eglė Rudokienė for DNA sequencing.

I am grateful to prof. Andreas Mayer (University of Lausanne, Switzerland) and all the members of his team for a very friendly working environment, which greatly facilitated the final stages of preparing my doctoral thesis.

Many thanks to my family, especially my mother and grandmother, for their concern, understanding and continuous moral support.

This work was in part supported by grants from the Howard Huges Medical Institute, the Ministry of Education and Science of Lithuania and the Lithuanian State Science and Study foundation.

CURRICULUM VITAE

Name:	Rūta Gerasimaitė
Date & place of birth:	March 5, 1977, Vilnius
Address:	Avenue du 1er Mai 10, 1020 Renens, Switzerland
E-mail:	ruta.gerasimaite@unil.ch

Education:

2001	*Master of Sciences* (Biochemistry), Vilnius University, Lithuania
1999	*Bachelor of Sciences* (Biochemistry), Vilnius University, Lithuania

Employment:

Form 2009	*Researcher*, University of Lausanne, Lausanne, Switzerland
2007 – 2009	*Junior research fellow*, Institute of Biotechnology, Vilnius, Lithuania
2007 – 2008	*Lecturer*, Vilnius Gediminas Technical University, Vilnius, Lithuania
2003 – 2007	*Ph.D. student* in Biochemistry, Institute of Biotechnology and Vilnius University, Vilnius, Lithuania
2001 – 2003	*Junior research fellow*, Institute of Biotechnology, Vilnius, Lithuania

REZIUMĖ

DNR citozino C5-metiltransferazės (C5-MTazės) atpažįsta 2-8 bp ilgio DNR sekas (taikinius) ir perneša metilgrupę nuo kofaktoriaus *S*-adenozil-L-metionino (AdoMet) ant taikinyje esančio citozino. Kaip modelinė sistema DNR citozino metilinimo tyrimuose dažniausiai naudojama ir išsamiausiai charakterizuota prokariotinė MTazė HhaI (M.HhaI), atpažįstanti seką GCGC ir metilinanti joje pirmąjį citoziną. Kristalografiniai M.HhaI tyrimai parodė, kad už katalizę ir specifinę sąveiką su DNR atsakingi baltymo elementai yra atskirti erdvėje. Didžiajame (katalitiniame) baltymo domene sutelkta dauguma konservatyvių, aktyvų centrą ir kofaktoriaus surišimo kišenę sudarančių aminorūgščių. Dvi paviršinės kilpos iš mažojo (taikinio atpažinimo) domeno sudaro beveik visus specifinius kontaktus su DNR bazėmis. Su konkrečia baze kontaktuoja tik vienai kilpai priklausančios aminorūgštys, taigi taikinio atpažinimo domenui būdinga modulinė struktūra.

Šio darbo metu M.HhaI taikinys GCGC pakeistas į GCG, baltymų inžinerijos ir kryptingos evoliucijos metodais pašalinant vienos iš dviejų taikinio atpažinimo kilpų funkciją. Ši asimetrinį taikinį atpažįstanti MTazė (hemimetilazė) naudojama eukariotų DNR metilinimo tyrimuose. Darbo metu nustatyta, kad specifinių kontaktų su DNR bazėmis praradimas pablogina fermento katalitinį efektyvumą, tačiau tai galima kompensuoti sustiprinant nespecifinius kontaktus tarp baltymo ir DNR fosfodiesterinio karkaso. Naujo specifiškumo MTazė katalizuoja didesnių negu metilas grupių pernešimą nuo AdoMet analogų, todėl gali būti naudojama mTAG technologijoje (*methyltransferase-directed transfer of activated groups*) gamtinių DNR žymėjimui. Taigi pirmą kartą vienu metu pakeistas DNR C5-MTazės specifiškumas ir DNR taikiniui, ir kofaktoriui.

Susidarant M.HhaI kompleksui su DNR stipriai keičiasi tiek baltymo, tiek DNR konformacija: metilinamas citozinas išsukamas iš DNR dvigrandės spiralės o paviršinė didžiojo domeno kilpa (katalitinė kilpa) užsidaro užrakindama surištą DNR ir baigdama formuoti aktyvų centrą. Reakcijos metu citozino žiedas aktyvuojamas susidarant kovalentiniam tarpiniam kompleksui tarp katalitinio cisteino ir citozino žiedo C6-atomo. Nors daugelis metilinimo reakcijos aspektų išsamiai charakterizuoti, iki šiol mažai ištyrinėti greiti procesai, vykstantys iki metilgrupės pernašos – citozino išsukimas, katalitinės kilpos užsidarymas ir citozino kovalentinė aktyvacija. Šių stadijų tyrimus didele dalimi ribojo adekvačių metodų nebuvimas. Antrojoje darbo dalyje buvo tyrinėjami greiti baltymo ir DNR konformaciniai virsmai, vykstantys prieš metilgrupės pernašą. Buvo sukurtas naujas metodas, leidžiantis tiesiogiai stebėti citozino išsukimą iš DNR spiralės, matuojant šio proceso sukeliamą hiperchrominį efektą, ir kovalentinės jungties su baltymu susidarymą, matuojant chromoforo išnykimą. Kadangi hiperchrominis efektas – fundamentali nukleorūgščių molekulių savybė, šį metodą galima pritaikyti daugelio bazę iš RNR arba DNR išsukančių baltymų tyrimui, nepriklausomai nuo bazės prigimties. Pirmą kartą parodyta, kad M.HhaI katalizuojamoje reakcijoje citozinas išsukamas ir katalitinė kilpa užsidaro sinchroniškai ir šio virsmo greitį ženkliai sąlygoja taikinio bazių poros stiprumas. Taip pat nustatytas citozino kovalentinės aktyvacijos greitis (0,3-0,6 s^{-1}) ir parodyta, kad metilgrupės pernaša yra vėlesnė katalitinio ciklo stadija.

REFERENCES

1. Allan, B. W., Reich, N. O. and Beechem, J. M. (1999). Measurement of the absolute temporal coupling between DNA binding and base flipping. *Biochemistry* **38**(17): 5308-5314.
2. Ausubel, F. M. (1995). *Short protocols in molecular biology*, Harvard Medical School, USA.
3. Bock, C., Reither, S., Mikeska, T., Paulsen, M., Walter, J. and Lengauer, T. (2005). BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing. *Bioinformatics* **21**(21): 4067-4068.
4. Buryanov, Y. and Shevchuk, T. (2005). The use of prokaryotic DNA methyltransferases as experimental and analytical tools in modern biology. *Anal Biochem* **338**(1): 1-11.
5. Cantor, C. R., Warshaw, M. M. and Shapiro, H. (1970). Oligonucleotide interactions. 3. Circular dichroism studies of the conformation of deoxyoligonucleotides. *Biopolymers* **9**(9): 1059-1077.
6. Carvin, C. D., Parr, R. D. and Kladde, M. P. (2003). Site-selective in vivo targeting of cytosine-5 DNA methylation by zinc-finger proteins. *Nucleic Acids Res* **31**(22): 6493-6501.
7. Cavaluzzi, M. J. and Borer, P. N. (2004). Revised UV extinction coefficients for nucleoside-5'-monophosphates and unpaired DNA and RNA. *Nucleic Acids Res* **32**(1): e13.
8. Cedar, H. and Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. *Nat Rev Genet* **10**(5): 295-304.
9. Cheng, X. and Blumenthal, R. M. (2008). Mammalian DNA methyltransferases: a structural perspective. *Structure* **16**(3): 341-350.
10. Cheng, X. and Blumenthal, R. M. (2010). Coordinated chromatin control: structural and functional linkage of DNA and histone methylation. *Biochemistry* **49**(14): 2999-3008.
11. Cheng, X., Kumar, S., Posfai, J., Pflugrath, J. W. and Roberts, R. J. (1993). Crystal structure of the HhaI DNA methyltransferase complexed with S-adenosyl-L-methionine. *Cell* **74**: 299-307.
12. Cohen, H. M., Tawfik, D. S. and Griffiths, A. D. (2004). Altering the sequence specificity of HaeIII methyltransferase by directed evolution using in vitro compartmentalization. *Protein Eng Des Sel* **17**(1): 3-11.
13. Dalhoff, C., Lukinavičius, G., Klimašauskas, S. and Weinhold, E. (2006). Direct transfer of extended groups from synthetic cofactors by DNA methyltransferases. *Nat Chem Biol* **2**(1): 31-32.
14. Daujotytė, D., Serva, S., Vilkaitis, G., Merkienė, E., Venclovas, C. and Klimašauskas, S. (2004). HhaI DNA methyltransferase uses the protruding Gln237 for active flipping of its target cytosine. *Structure* **12**(6): 1047-1055.
15. Daujotytė, D., Vilkaitis, G., Manelytė, L., Skalicky, J., Szyperski, T. and Klimašauskas, S. (2003). Solubility engineering of the HhaI methyltransferase. *Protein Eng* **16**(4): 295-301.
16. Dila, D., Sutherland, E., Moran, L., Slatko, B. and Raleigh, E. A. (1990). Genetic and sequence organization of the mcrBC locus of Escherichia coli K-12. *J Bacteriol* **172**(9): 4888-4900.

17. Dryden, D. T. (1999). *Bacterial DNA Methyltransferases*. S-Adenosylmethionine-Dependent Methyltransferases: Structures and Functions. Cheng, X. and R. M. Blumenthal. Singapore, World Scientific Publishing Company.
18. Estabrook, R. A., Nguyen, T. T., Fera, N. and Reich, N. O. (2009). Coupling sequence-specific recognition to DNA modification. *J Biol Chem* **284**(34): 22690-22696.
19. Estabrook, R. A. and Reich, N. (2006). Observing an induced-fit mechanism during sequence-specific DNA methylation. *J Biol Chem* **281**(48): 37205-37214.
20. Fox, J. J. and Shugar, D. (1952). Spectrophotometric studies of nucleic acid derivatives and related compounds as a function of pH. II. Natural and synthetic pyrimidine nucleosides. *Biochim Biophys Acta* **9**(4): 369-384.
21. Goll, M. G. and Bestor, T. H. (2005). Eukaryotic cytosine methyltransferases. *Annu Rev Biochem* **74**: 481-514.
22. Horton, J. R., Ratner, G., Banavali, N. K., Huang, N., Choi, Y., Maier, M. A., Marquez, V. E., MacKerell, A. D., Jr. and Cheng, X. (2004). Caught in the act: visualization of an intermediate in the DNA base-flipping pathway induced by HhaI methyltransferase. *Nucleic Acids Res* **32**(13): 3877-3886.
23. Huang, N., Banavali, N. K. and MacKerell, A. D., Jr. (2003). Protein-facilitated base flipping in DNA by cytosine-5-methyltransferase. *Proc Natl Acad Sci U S A* **100**(1): 68-73.
24. Youngblood, B., Buller, F. and Reich, N. O. (2006). Determinants of sequence-specific DNA methylation: target recognition and catalysis are coupled in M.HhaI. *Biochemistry* **45**(51): 15563-15572.
25. Kaminskas, E., Farrell, A. T., Wang, Y. C., Sridhara, R. and Pazdur, R. (2005). FDA drug approval summary: azacitidine (5-azacytidine, Vidaza) for injectable suspension. *Oncologist* **10**(3): 176-182.
26. Kelley, S. O. and Barton, J. K. (1999). Electron transfer between bases in double helical DNA. *Science* **283**(5400): 375-381.
27. Kilgore, J. A., Hoose, S. A., Gustafson, T. L., Porter, W. and Kladde, M. P. (2007). Single-molecule and population probing of chromatin structure using DNA methyltransferases. *Methods* **41**(3): 320-332.
28. Klimašauskas, S., Kumar, S., Roberts, R. J. and Cheng, X. (1994). Hhal methyltransferase flips its target base out of the DNA helix. *Cell* **76**: 357-369.
29. Klimašauskas, S. and Roberts, R. J. (1995). M.HhaI binds tightly to substrates containing mismatches at the target base. *Nucleic Acids Res.* **23**: 1388-1395.
30. Klimašauskas, S., Szyperski, T., Serva, S. and Wuthrich, K. (1998). Dynamic models of the flipped-out cytosine during HhaI methyltransferase-DNA interactions in solution. *EMBO J.* **17**: 317-324.
31. Klimašauskas, S. and Weinhold, E. (2007). A new tool for biotechnology: AdoMet-dependent methyltransferases. *Trends Biotechnol* **25**(3): 99-104.
32. Kumar, S., Cheng, X., Pflugrath, J. W. and Roberts, R. J. (1992). Purification, crystallization, and preliminary X-ray diffraction analysis of an M.HhaI-AdoMet complex. *Biochemistry* **31**(36): 8648-8653.
33. Kumar, S., Horton, J. R., Jones, G. D., Walker, R. T., Roberts, R. J. and Cheng, X. (1997). DNA containing 4'-thio-2'-deoxycytidine inhibits methylation by HhaI methyltransferase. *Nucleic Acids Res* **25**(14): 2773-2783.

34. Kuzmic, P. (1996). Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. *Anal Biochem* **237**(2): 260-273.
35. Lange, C., Wild, C. and Trautner, T. A. (1996). Identification of a subdomain within DNA-(cytosine-C5)-methyltransferases responsible for the recognition of the 5' part of their DNA target. *EMBO J.* **15**: 1443-1450.
36. Lee, Y. F., Tawfik, D. S. and Griffiths, A. D. (2002). Investigating the target recognition of DNA cytosine-5 methyltransferase HhaI by library selection using in vitro compartmentalisation. *Nucleic Acids Res* **30**(22): 4937-4944.
37. Liebert, K., Hermann, A., Schlickenrieder, M. and Jeltsch, A. (2004). Stopped-flow and mutational analysis of base flipping by the Escherichia coli Dam DNA-(adenine-N6)-methyltransferase. *J Mol Biol* **341**(2): 443-454.
38. Lukinavičius, G., Lapienė, V., Staševskij, Z., Dalhoff, C., Weinhold, E. and Klimašauskas, S. (2007). Targeted labeling of DNA by methyltransferase-directed transfer of activated groups (mTAG). *J Am Chem Soc* **129**(10): 2758-2759.
39. Merkienė, E. and Klimašauskas, S. (2005). Probing a rate-limiting step by mutational perturbation of AdoMet binding in the HhaI methyltransferase. *Nucleic Acids Res* **33**(1): 307-315.
40. Mi, S. and Roberts, R. J. (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.
41. Moller, A., Wild, U., Riesner, D. and Gassen, H. G. (1979). Evidence from ultraviolet absorbance measurements for a codon-induced conformational change in lysine tRNA from Escherichia coli. *Proc Natl Acad Sci U S A* **76**(7): 3266-3270.
42. Neely, R. K., Daujotytė, D., Gražulis, S., Magennis, S. W., Dryden, D. T., Klimašauskas, S. and Jones, A. C. (2005). Time-resolved fluorescence of 2-aminopurine as a probe of base flipping in M.HhaI-DNA complexes. *Nucleic Acids Res* **33**(22): 6953-6960.
43. Osterman, D. G., DePillis, G. D., Wu, J. C., Matsuda, A. and Santi, D. V. (1988). 5-Fluorocytosine in DNA is a mechanism-based inhibitor of HhaI methylase. *Biochemistry* **27**(14): 5204-5210.
44. Palmier, M. O. and Van Doren, S. R. (2007). Rapid determination of enzyme kinetics from fluorescence: overcoming the inner filter effect. *Anal Biochem* **371**(1): 43-51.
45. Pogribny, I. P. and Beland, F. A. (2009). DNA hypomethylation in the origin and pathogenesis of human diseases. *Cell Mol Life Sci* **66**(14): 2249-2261.
46. Puglisi, J. D. and Tinoco, I., Jr. (1989). Absorbance melting curves of RNA. *Methods Enzymol* **180**: 304-325.
47. Rajski, S. R., Kumar, S., Roberts, R. J. and Barton, J. K. (1999). Protein-Modulated DNA Electron Transfer. *Journal of the American Chemical Society* **121**(23): 5615-5616.
48. Rathert, P., Rasko, T., Roth, M., Slaska-Kiss, K., Pingoud, A., Kiss, A. and Jeltsch, A. (2007). Reversible inactivation of the CG specific SssI DNA (cytosine-C5)-methyltransferase with a photocleavable protecting group. *Chembiochem* **8**(2): 202-207.
49. Reinisch, K. M., Chen, L., Verdine, G. L. and Lipscomb, W. N. (1995). The crystal structure of HaeIII methyltransferase covalently complexed to DNA: an extrahelical cytosine and rearranged base pairing. *Cell* **82**: 143-153.

50. Roberts, R. J., Vincze, T., Posfai, J. and Macelis, D. (2010). REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Research* **38**(suppl 1): D234-D236.
51. Sambrook, J. and Russel, D. W. (2001). *Molecular cloning. A laboratory manual*, CSH laboratory press, USA.
52. Sankpal, U. T. and Rao, D. N. (2002). Structure, function, and mechanism of HhaI DNA methyltransferases. *Crit Rev Biochem Mol Biol* **37**(3): 167-197.
53. Serva, S., Weinhold, E., Roberts, R. J. and Klimašauskas, S. (1998). Chemical display of thymine residues flipped out by DNA methyltransferases. *Nucleic Acids Res* **26**(15): 3473-3479.
54. Singer, E. M. and Smith, S. S. (2006). Nucleoprotein assemblies for cellular biomarker detection. *Nano Lett* **6**(6): 1184-1189.
55. Skaric, V., Gaspert, B., Hohnjec, M. and Lacan, G. (1974). Some dihydro-cytidines and -isocytidines. *Journal of the Chemical Society, Perkin Transactions 1*: 267-271.
56. Smith, A. E. and Ford, K. G. (2007). Specific targeting of cytosine methylation to DNA sequences in vivo. *Nucleic Acids Res* **35**(3): 740-754.
57. Subach, O. M., Baskunov, V. B., Darii, M. V., Maltseva, D. V., Alexandrov, D. A., Kirsanova, O. V., Kolbanovskiy, A., Kolbanovskiy, M., Johnson, F., Bonala, R., Geacintov, N. E. and Gromova, E. S. (2006). Impact of benzo[a]pyrene-2'-deoxyguanosine lesions on methylation of DNA by SssI and HhaI DNA methyltransferases. *Biochemistry* **45**(19): 6142-6159.
58. Svedruzic, Z. M. and Reich, N. O. (2004). The mechanism of target base attack in DNA cytosine carbon 5 methylation. *Biochemistry* **43**(36): 11460-11473.
59. Sverdlov, E. D., Monastyrskaya, G. S., Guskova, L. I., Levitan, T. L., Sheichenko, V. I. and Budowsky, E. I. (1974). Modification of cytidine residues with a bisulfite--O-methylhydroxylamine mixture. *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis* **340**(2): 153-165.
60. Tanaka, M., Yoshida, S., Saneyoshi, M. and Yamaguchi, T. (1981). Utilization of 5-Fluoro-2′-deoxyuridine Triphosphate and 5-Fluoro-2′-deoxycytidine Triphosphate in DNA Synthesis by DNA Polymerases α and β from Calf Thymus. *Cancer Research* **41**(10): 4132-4135.
61. Timar, E., Groma, G., Kiss, A. and Venetianer, P. (2004). Changing the recognition specificity of a DNA-methyltransferase by in vitro evolution. *Nucleic Acids Res* **32**(13): 3898-3903.
62. Ulanov, B. P., Matorina, T. I. and Zmanuzl, N. M. (1976). [Double modification of cytidine residues in DNA]. *Mol Biol (Mosk)* **10**(6): 1211-1220.
63. Vilkaitis, G., Dong, A., Weinhold, E., Cheng, X. and Klimašauskas, S. (2000). Functional roles of the conserved threonine 250 in the target recognition domain of HhaI DNA methyltransferase. *J. Biol. Chem.*: in press.
64. Vilkaitis, G. and Klimašauskas, S. (1998). Construction and analysis of monospecific DNA cytosine-C5 methyltransferases with chimeric target recognition domains. *Biologija*(1): 51-54.
65. Vilkaitis, G., Merkienė, E., Serva, S., Weinhold, E. and Klimašauskas, S. (2001). The mechanism of DNA cytosine-5 methylation. Kinetic and mutational dissection of Hhai methyltransferase. *J Biol Chem* **276**(24): 20924-20934.

66. Vilkaitis, G., Suetake, I., Klimašauskas, S. and Tajima, S. (2005). Processive methylation of hemimethylated CpG sites by mouse Dnmt1 DNA methyltransferase. *J Biol Chem* **280**(1): 64-72.
67. Wang, P., Brank, A. S., Banavali, N. K., Nicklaus, M. C., Marquez, V. E., Christman, J. K. and MacKerell, A. D. (2000). Use of Oligodeoxyribonucleotides with Conformationally Constrained Abasic Sugar Targets To Probe the Mechanism of Base Flipping by HhaI DNA (Cytosine C5)-methyltransferase. *Journal of the American Chemical Society* **122**(50): 12422-12434.
68. Wells, R. D., Larson, J. E., Grant, R. C., Shortle, B. E. and Cantor, C. R. (1970). Physicochemical studies on polydeoxyribonucleotides containing defined repeating nucleotide sequences. *J Mol Biol* **54**(3): 465-497.
69. Wu, J. C. and Santi, D. V. (1987). Kinetic and catalytic mechanism of HhaI methyltransferase. *J. Biol. Chem.* **262**: 4778-4786.
70. Zangi, R., Arrieta, A. and Cossio, F. P. (2010). Mechanism of DNA methylation: the double role of DNA as a substrate and as a cofactor. *J Mol Biol* **400**(3): 632-644.
71. Zhang, X. and Bruice, T. C. (2006). The mechanism of M.HhaI DNA C5 cytosine methyltransferase enzyme: a quantum mechanics/molecular mechanics approach. *Proc Natl Acad Sci U S A* **103**(16): 6148-6153.