DNA cytosine-5-methyltransferase \(HhaI\) recognizes the GCGC sequence and flips the inner cytosine out of DNA helix and into the catalytic site for methylation. The 5'-phosphate of the flipped out cytosine is in contact with the conserved Thr-250 from the target recognition domain. We have produced 12 mutants of Thr-250 and examined their methylation potential in \(vivo\). Six active mutants were subjected to detailed biochemical and structural studies. Mutants with similar or smaller side chains (Ser, Cys, and Gly) are very similar to wild-type enzyme in terms of steady-state kinetic parameters \(K_{cat}\), \(K_m\) for \(\text{DNA} \cdot \text{AdoMet}\). In contrast, the mutants with bulkier side chains (Asn, Asp, and His) show increased \(K_m\) values for both substrates. Fluorescence titrations and stopped-flow kinetic analysis of interactions with duplex oligonucleotides containing 2-aminopurine at the target base position indicate that the T250G mutation leads to a more polar but less solvent-accessible position of the flipped out target base. The x-ray structure of the ternary \(M.HhaI(T250G)\cdot\text{DNA} \cdot \text{AdoHe}\) complex shows that the target cytosine is locked in the catalytic center of enzyme. The space created by the mutation is filled by water molecules and the adjacent DNA backbone atoms dislocate slightly toward the missing side chain. In aggregate, our results suggest that the side chain of Thr-250 is involved in constraining the conformation the DNA backbone and the target base during its rotation into the catalytic site of enzyme.

Recognition and modification of specific residues in DNA is a key event in many cellular processes. DNA methyltransferases (MTases)\(^1\) which transfer a methyl group from the cofactor S-adenosyl-L-methionine (AdoMet) onto a defined residue in a specific sequence of DNA are attractive model systems for enzymes that integrate molecular recognition and catalytic activity. MTases fall into three classes based on the type of methylation produced: \(N^6\)-methyladenine, \(N^4\)-methylcytosine, and 5-methylcytosine. The DNA cytosine-5-methyltransferases (C5-MTases) are unique in that they are found both in prokaryotes and eukaryotes. In higher organisms, DNA methylation is the focus of renewed interest due to its documented roles in the control of a number of cellular processes including transcription, genomic imprinting, embryonic development, and chromatin organization (for reviews, see Refs. 1–3). MTases are obligatory enzyme components of restriction-modification systems found in most bacterial organisms (4) and have recently been shown to play a role in bacterial virulence (5).

Another distinctive feature of prokaryotic C5-MTases is their common two-domain architecture. A consecutive set of six strongly conserved and four less conserved sequence motifs (I-X) are found in bacterial as well as eukaryotic enzymes (6, 7). Nine of them are located in the larger domain and participate in cofactor binding and catalytic methylation of the target base. A large so-called “variable” region between motifs VIII and IX folds to make the bulk of the smaller domain (8) and serves to define the sequence specificity of C5-MTases (9, 10). This apparent separation of the recognition and catalytic functions into distinct domains offers obvious evolutionary advantages for creating wide diversity of specificities. On the other hand, spatial separation of the two functions demands structural flexibility of the reaction components during the catalytic cycle. Indeed, the x-ray structure of the bacterial C5-MTase \(HhaI\) (11) in complex with DNA revealed that sequence-specific binding is accompanied by rotation of the target nucleotide out of the DNA helix and into the catalytic pocket in the large domain. Subsequent x-ray analyses of other DNA-modification and DNA-repair enzymes confirmed that this novel distortion of DNA, termed base-flipping, is a fundamental mode of DNA-protein interaction in situations when ordinary or damaged bases need to be covalently modified or removed (12, 13).

Many efforts have been devoted to elucidating the mechanism of base-flipping in a variety of systems. \(M.HhaI\), which recognizes the sequence GCGC and methylates the inner cytosine (boldface), serves as a base-flipping paradigm for this class of enzymes (14). Mutational analysis of Gln-237 located in the target recognition domain of \(M.HhaI\) showed that this residue is important for stabilizing the reaction complex upon flipping of the target base out of the DNA helix (15). \(^{13}P\) NMR studies in solution revealed base-flipping intermediates in which the target base is found neither in its original stacked position in the DNA helix, nor in its final destination in the catalytic pocket of the enzyme (16), but is rather exposed on the exterior of the molecule (17). Moreover, it was demonstrated that a variety of bases or even an abasic deoxyribose sugar alone at the target position can be rotated out of the DNA by \(M.HhaI\) (14, 18–20). These findings offered a clear separation of two
mechanistic events, base-flipping and subsequent base-binding (catching) in the catalytic pocket, implying that the initial flipping motion must be actively promoted by the enzyme. However, the details of exactly how this active flipping process is achieved remain obscure.

In addition to the 10 conserved motifs found in C5-MTases, a conserved T(L,V,I) dipeptide can be identified near the C terminus of the variable region (7, 21). X-ray studies of the HhaI (11) and HaeIII (22) MTases in complex with DNA indicate that this conservation is also preserved at the level of tertiary structure. The side chain atoms of the conserved threonine residue superimpose very well in the two structures and pack against the 5′-phosphodiester group of the extrahelical 2′-deoxyctydine (Fig. 1, A and B). Interestingly, although the neighboring amino acid residues are not conserved in the two MTases, the backbone atoms superimpose well in a stretch of 14 residues (amino acids 240–253 in M.HhaI) (22). This segment of the target recognition domain runs parallel with the target DNA backbone atoms superimpose well in a stretch of 14 residues (22). This segment of the target recognition domain runs parallel with the target DNA backbone atoms superimpose well in a stretch of 14 residues (amino acids 240–253 in M.HhaI) (22).

To assess the functional importance of this structural conservation we have replaced Thr-250 in M.HhaI by other amino acid residues. The effects of individual replacements on in vivo methylation activity, catalytic parameters, DNA-binding, and base-flipping were studied using a protection assay, fluorescence spectroscopy, X-ray crystallography, and steady-state kinetics.

**Experimental Procedures**

**Mutagenesis**—Restriction endonucleases, DNA modification enzymes, and kits were obtained from MBI Fermentas. The following degenerate duplex oligodeoxyribonucleotides (obtained from MBI Fermentas) were used for cassette mutagenesis of the 250 codon: 5′-CCGGGCAATGGATCAGATCCG-3′ and 3′-CCCTATGAATGGATCAGATCCG-5′; 5′-CCGGGCAATGGATCAGATCCG-3′ and 3′-CCCTATGAATGGATCAGATCCG-5′; 5′-CCGGGCAATGGATCAGATCCG-3′ and 3′-CCCTATGAATGGATCAGATCCG-5′.

The four double-mutant codons (250 in bold) were ligated into the large BspI fragment of pHF52.2 The transformants of the ER1727 strain were plated on selective agar, and individual clones were analyzed by DNA sequencing of the regions concerned (10). Thirteen out of 15 possible variants of codon 250 were identified: G (Val), GCC (Ala), GAC (Asp), GCC (Gly), ATC (Ile), AAC (Asn), AGC and TCC (Ser), CCC (Pro), CAC (His), CGC (Arg), TAC (Tyr), and TGC (Cys). No mutations were found in other regions of the genes.

**In Vivo Methylation Activities**—Analysis was carried out essentially as described previously (10). Controlled expression of the mutant proteins was achieved by adjusting the concentration of IPTG (0 or 0.4 mM) in the growth medium. Plasmids isolated from induced and uninduced cells were challenged with an excess of B.HhaI (MBI Fermentas), an isochizomer of R.HhaI (24) and analyzed on a 1% agarose gel.

**Protein Expression and Purification**—Mutant proteins and wild-type M.HhaI were expressed in Erichschia coli strain ER1727 and extracted as described previously (16). All proteins appeared as sole bands (>95%) in Coomassie-stained polyacrylamide gels (not shown). Protein concentrations were estimated using a Coomassie G-250 assay (Bio-Rad) with bovine serum albumin as standard. Exact concentrations were determined by active site titration with a 37-mer fluorescent duplex oligonucleotide (see below) or 3′-labelled hemimethylated 37-mer duplex in the presence of AdoHcy (not shown) and the values amounted to ~60% of those obtained in the Coomassie measurements. In subsequent preparations of all variants, Coomassie readings corrected by the factor of 0.6 were used for all mutants with reproducible results.

**Steady-state Kinetics**—Methylation reactions were carried out in Methylation buffer (50 mM Tris-HCl, pH 7.4, 10 mM EDTA, 6 mM 2-mercaptoethanol, 0.2 mg/ml bovine serum albumin (25)) in the presence of 333 nM [methyl-3H]AdoMet (15 or 70 Ci/mmol) and poly(dG-dC) (Amersham Pharmacia Biotech) at double-strand recognition site concentrations ranging from 0.5 to 200 nM. 32P-[γ-32P] Measurements were performed with 120 nM as standard and concentrations of 250, 500, 1000, and 2000 nM of methyl-[3H]AdoMet concentration from 5 nM to 25.6 μM. After 10 min incubation at 37 °C, the reactions were supplemented with at least 100-fold excess of non-radioactive AdoMet and AdoHcy mixture (100:1). Duplicate samples were spotted onto 2.5-mm DE-81 filters (Whatman), washed 4 times with 50 ml sodium phosphate buffer for 10 min, 2 times with H2O, and 2 times with ethanol and dried. The radioactivity on the filters was measured in 4 ml of CytoScint (Fisher Biotech) using a Beckman LS 8015 liquid scintillation spectrometer. Data were analyzed by nonlinear regression fitting to the Michaelis-Menten equation using the computer program Grafit, version 3.01 (26), or by global fitting of the entire data set of initial velocities to a Bi-Bi kinetic mechanism using Dynafit (27).

**Fluorescence Spectroscopy**—The 37-mer deoxyribonucleo- toides were obtained from MWG-Biotech AG (HPSF grade) or from MBI Fermentas (gel-purified): 5′-GACTGGTGACATGACGACGACCGACACAACTATCCG (GPPGCC); 5′-GACTGGTGATGATGACGACGACGGCTGACACAACTATCCG (GAGCC); 5′-GCCCTGATACTGTACCAGT (GMGC).

The M.HhaI recognition site is underlined and nucleotides at the target base position are boldface. DNA duplexes were produced by annealing appropriate 37-mer strands (abbreviations are shown in brackets) as described previously (14).

Fluorescence intensity measurements were performed at 25 °C on an SLM-Amino AB2 spectrophotometer at excitation wavelength (λEx) of 320 nm and emission wavelength (λEm) of 381 nm as described previously (14). Titration of 100 nM 37-mer GPPGC/GMGC duplex was performed by incremental addition of 4 μM M.HhaI in Reaction buffer (50 mM NaCl, 10 mM Tris-HCl, pH 7.4, 0.5 mM EDTA, 2 mM 2-mercaptoethanol) containing 0.2 mg/ml bovine serum albumin. Emission and excitation spectra were recorded at 22 °C on a PerkinElmer Life Sciences LC50-B luminescence spectrometer equipped with a Xe lamp and a 4-mm rectangular quartz cell with a mixer. Emission spectra were recorded at λEm = 370 nm with excitation and emission bandwidths of 10 and 2.5 nm, respectively; excitation spectra were recorded at 5 μM = 370 nm with emission and excitation bandwidths of 10 and 5 nm, respectively. 200 nM GPPGC/GMGC in Reaction buffer was titrated by incremental addition of 25 μM M.HhaI until saturation of fluorescence level was achieved. Three scans were averaged for each spectrum. Further excitation spectra (not shown) were collected under identical conditions except that GAGC/GMGC duplex was used instead of the fluorescent DNA. Corrected excitation and emission spectra of 2AP were obtained by subtracting the control spectra to eliminate Raman scatter peaks (observed at 328 and 360 nm, respectively) and contributions from the protein and other components. The magnitude of spectral corrections was in the range of 10–35%.

Acrylamide fluorescence quenching titrations were performed at 22 °C with the PerkinElmer Life Sciences LC50-B luminescence spectrometer. Emission intensities were recorded at excitation wavelength 320 nm, bandwidth 2.5 nm, and emission wavelength 370 nm, bandwidth 10 nm, sampling time >2 min. Acrylamide solution was added incrementally into 200 nM GPPGC/GMGC and 350 nM M.Tase in Reaction buffer. Relative intensities F/F0 corrected for background fluorescence were fitted to the general Stern-Volmer equation F/F0 = 1 + Ksv [acrylamide] (28).

Stopped-flow experiments were performed on a Hi-Tech Scientific SF 61MX apparatus (single mixing mode) equipped with a Xe-Hg UV lamp. The excitation wavelength was 312 nm (2 mm slit); emission light was passed through a 389-nm cut-on filter. The 37-mer DNA duplex (75 nM) was rapidly mixed with 30, 60, 120, 240, 450, and 960 nM M.HhaI or its mutant in Reaction buffer (see above). 500 logarithmically spaced data points were collected for each progress curve in the time window from 1.3 to 250 ms at signal averaging constant of ≤100 μs. Multiple time courses (4–6 runs) were averaged and analyzed by fitting to double exponential equations using KinetAsyst2, version 2.0 (Hi-Tech) (29), or Grafit, version 3.01 (28) software packages. Multicurve fitting was performed with Dynafit (27). Combined data set from WT, T250D, and T250G mutants (100 data points × 5 progress curves × 3 variants) was fitted to the 2-step binding mechanism (Scheme 1). The fluorescence intensity ratio FMTase-DNA/M.Tase-DNA was arbitrarily selected (1.5, 1.7, 2, or 3) and used as a fixed global parameter. The four rate constants and progress curve offsets (background fluorescence intensity) were allowed to refine.
Cryotransfer—The structure of the T250G mutant of M.HhaI in a ternary complex with AdoHcy and a short DNA duplex was determined by x-ray crystallography (Protein Data Bank entry code 1FJX). A 13-mer palindromic oligonucleotide (5'-TGATAGGGCCTATC-3') was annealed to form a 12-base pair symmetric duplex with 5' single thymine overhangs. Crystals were grown by mixing the T250G-DNA:AdoHcy complex with an equal volume of 2.7 M ammonium sulfate, 20% (v/v) glucose, 50 mM sodium citrate at pH 5.6, and equilibrating the mixture against 1 ml of the latter solution at 16 °C. The 1FJX crystals formed in the same space group R32 as the previously reported ternary complex structures with cell dimensions of a = b = 98.46 Å, c = 332.22 Å. One complex was present per asymmetric unit. X-ray diffraction intensities were measured from a single frozen crystal at 95 K and a wavelength of 1.1 Å on a ADSC Q4 CCD at the beamline X26C of the National Synchrotron Light Source, Brookhaven National Laboratory. The crystal was transferred to a cryobuffer (15% (w/v) glycerol, 2.7 M ammonium sulfate, 50 mM sodium citrate, pH 6.5) prior to freezing. Data acquisition yielded 24,152 unique reflections out of 67,597 total observations (Rmerge = [S|I|/|S|] = 0.056, <I/I>} = 13.2, where <I/} is the measured intensity averaged from multiple observations of symmetry related reflections and <I/} is the root mean square deviation from the mean).

The structure was solved by the difference Fourier method using the previously solved 3MHT ternary structure (WT M.HhaI with AdoHcy and DNA identical in sequence to the one used here) as the initial model (30). The model was refined against diffraction data measured from 2.26 Å using X-PLOR (31). After several rounds of manual building and placement of well ordered solvent molecules (interpreted as water) by examination of difference electron density, the cryotransfer R-free and R-factor (\( R = \sum\) \( |F_o| - |F_c|/\sum |F_o| \)) were reduced to 28.3% for 1,896 (approximately 8% of total unique reflections) and 20.7% for 22,169 unique reflections, respectively.

RESULTS

Conservation of Thr in the Variable Region of C5-MTases—To assess the conservation level of the TL dipeptide in mono- and DNA identical in sequence to the one used here) as the initial model (30). The thymidine residue interferes with the conserved Thr position (the plane for aligning sequences that do not allow an ambiguous selection based on dipeptide analysis alone. In certain cases, pairwise similarities between several sequences were used in the construction of the alignment. The results of this analysis indicate that, depending on a similarity level, the conserved dipeptide TL can be identified in 67–71 out of 79 bacterial sequences analyzed (85–90%). The other sequences presumably contain Cys, Ser, or Ala at the Thr positions.

Eukaryotic C5-MTases seems to have a more even distribution of the 4 amino acids at the first position (Fig. 1C), although the number of sequences currently available is too small for meaningful quantitative evaluation. An interesting regularity is observed for both pairs of so-called hemi-methylases (M.Bpu10I and M.HgaI), enzymes that occur in the same cell and methylate complementary strands of the same non-palindromic target site (24). All four enzymes match the broad consensus. However, one MTase in each pair contains a normal doublet (TI), whereas a scrambled version (AI or AR) appears at the equivalent position of the accompanying enzyme (Fig. 1C).

Construction and in Vivo Characterization of Thr-250 Mutants—Twelve Thr-250 of M.HhaI mutants were constructed by cassette mutagenesis of the hhaM gene through built-in unique restriction endonuclease sites. Their in vivo methylation efficiencies of the GCGC sites was assessed by challenging plasmid DNA isolated from the corresponding clones with a methylation sensitive restriction endonuclease, R.HinI. Table I summarizes the methylation activities of the mutant MTases in vivo. Most substitutions at residue 250 lead to protection levels comparable to that of WT enzyme. Some versions, however, exhibited a significantly reduced activity (T250P, T250H), whereas two mutants (T250R and T250Y) showed a very low modification level even in cells overexpressing the proteins. Since the spectrum of mutations examined in this activity study represented all structural types of amino acids (small, large, aliphatic, aromatic, polar, charged etc.), we can generalize that large structural perturbations in this position interfere with the catalytic function of the enzyme.

Based on the methylation efficiencies in vivo, several mutants were selected for further analysis. We purified six of them (T250C, T250D, T250G, T250H, T250N, and T250S) to homogeneity using a procedure that involves a selective salt back-extraction of M.HhaI from cell debris (16, 32). Attempts to isolate some hydrophobic variants (T250A, T250P, T250Y, and T250V) by this method proved unsuccessful due to low solubility of these proteins at a variety of salt conditions examined.

Steady-state Kinetic Analysis—Previous steady-state kinetic study of the HhaI MTase (25) found that the enzyme operates by an ordered Bi Bi mechanism in which DNA binds before AdoMet. Our multiple-turnover kinetic analysis of the M.HhaI mutants was performed under very similar reaction conditions as used by Wu and Santi (25). Methylation reactions were performed with poly(dG-dC) and [3H]AdoMet as substrates at 37 °C. Initial velocity measurements were performed in the linear phase of product formation (substrate consumption was maintained below 10 and 3% for DNA and AdoMet, respectively). Concentration of DNA and cofactor were varied to obtain Km values for each substrate (Fig. 2). The kinetic parameters observed for the WT enzyme were in excellent agreement with the results reported previously (25). As can be seen in Table II, the multiple turnover km values of the mutants are very similar and cluster within a factor of 2 around that of the WT enzyme. The K

\[ \text{K}^{\text{DNA}} \] values obtained are quite
uniform too and follow in general the $K_D$ values (1–5 nM) for binary MTase-DNA interaction measured in gel mobility shift experiments; the largest (40-fold) increase in $K_{m}^{DNA}$ is observed for the T250D mutant, consistent with the lack of the corresponding binary complex in the gel analysis (not shown). Overall, the catalytic efficiencies $k_{cat}/K_{m}^{DNA}$ of T250H and T250D are

---

**FIG. 1.** Sequence and structure conservation of dipeptide TL in bacterial and eukaryotic C5-MTases. A, schematic of structural elements observed in the target recognition domain of bacterial MTases HhaI (11) and HaeIII (22): target recognition Loop1 and Loop2, and the two common structural elements termed “TL” and “HP” (boxed in the sequence alignment). B, structure of the M.HhaI(DNA)AdoHcy complex (3MHT) illustrates the position of Thr-250 (space-fill, green and red) relative to DNA (yellow sticks) and AdoHcy (red sticks) and the rest of protein (shown as ribbons). C, sequence alignment of bacterial and eukaryotic C5-MTases anchored at the conserved dipeptide TL (shown as white in dark background). MTases are compiled from Ref. 24 along with their recognition sequences (target cytosine underlined, if known). Numbers on the right show separation of the TL dipeptide from the subsequent conserved motif IX (not shown). Light shading shows residues conforming to the sequence consensus proposed in this work. Boldface residues are non-matching candidate TLs. Footnote a, previously published sequence (44) (accession number X55139) has been amended by the authors leading to the AGG(Arg) → ACG(Thr) change at the conserved TL (M. Kroger, personal communication). Footnote b, alignment based on the recently determined protein structure (A. Dong, J. Yoder, X. Zhang, L. Zhou, T. Bestor, and X. Cheng, submitted for publication).
Threonine-250 Mutants of HhaI DNA Methyltransferase

FIG. 2. Steady-state kinetic analysis of T250 mutants of M.HhaI. Reactions containing poly(dG-dC) and [methyl-3H]AdoMet were incubated at 37 °C for 10 min in the presence of WT (●), T250N (○), T250D (●), T250C (□), T250G (○), T250H (▲), or T250S (△) M.HhaI. The normalized initial velocities were fit by non-linear regression to the Michaelis-Menten equation and are presented as plots against AdoMet concentration.

TABLE II

Steady-state kinetic parameters of Thr-250 mutants of M.HhaI with poly(dG-dC) and [methyl-3H]AdoMet as substrates

| MTase | k_{cat} | k_{DNA} | k_{AdoMet} | k_{cat}/k_{DNA} | k_{cat}/k_{AdoMet} |
|-------|---------|---------|------------|----------------|-------------------|
|       | s^{-1}  | nm      | μM^{-1} s^{-1} | WT/Mut^a       | WT/Mut            |
| WT    | 0.020 ± 0.001 | 0.9 ± 0.2 | 13 ± 1 | 22 | 1 | 1.5 | 1 |
| T250G | 0.022 ± 0.001 | 1.4 ± 0.3 | 12 ± 1 | 16 | 1 | 1.4 | 0.8 |
| T250S | 0.042 ± 0.001 | 1.8 ± 0.6 | 30 ± 1 | 23 | 1 | 1.4 | 1.1 |
| T250C | 0.024 ± 0.001 | 0.7 ± 0.1 | 25 ± 2 | 34 | 0.65 | 0.96 | 1.6 |
| T250N | 0.036 ± 0.001 | 2.9 ± 0.5 | 300 ± 10 | 12 | 1.8 | 0.12 | 13 |
| T250D | 0.018 ± 0.001 | 42 ± 9 | 480 ± 23 | 0.43 | 52 | 0.04 | 38 |
| T250H | 0.009 ± 0.001 | 17 ± 3 | 6500 ± 280 | 0.53 | 42 | 0.0014 | 1100 |

^a^ Previously reported values are k_{cat} = 0.02 s^{-1}, k_{DNA} = 2.3 nm, k_{AdoMet} = 15 nm (25).

^b^ Values relative to those of WT enzyme.

down by a factor of 40 and 50, respectively, whereas the other mutants are almost as efficient as the native enzyme.

Unexpectedly, the K_{AdoMet} values showed an even wider variation. Replacements of Thr-250 with its closest structural homologues (Ser, Cys, and Gly) gave MTases that were very similar to the WT enzyme, however, the T250N, T250D, and T250H mutants exhibited a 23-, 35-, and 500-fold larger K_{DNA} values, respectively. The observed trend suggests that an increased size of the side chain may have important implications for cofactor interactions during catalysis.

Base-flipping Activity—Although methylation by M.HhaI is restricted to cytosine and, under certain conditions, uracil, the enzyme exerts its base-flipping activity toward a wide variety of residues present at the target position (18, 19). This property of the enzyme offers a broad selection of base chemistries for the DNA helix due to stacking interactions with neighboring bases (33–35). Interaction of M.HhaI with a hemimethylated 37-mer DNA duplex containing 2AP at the target position (GPGC/GMGC) leads to a 50-fold enhancement of fluorescence intensity due to enzyme-mediated flipping of the target base out of the DNA helix (14). In this study, similar titrations were performed with six Thr-250 mutants to determine their capability to unstack the target base upon interaction with DNA (Fig. 3A). As expected, addition of saturating amounts of the MTases to GPGC/GMGC results in a strong increase of fluorescence intensity. Fluorescence amplitudes ΔF = F_{max} – F_0 of four mutants were similar to those of the WT enzyme, suggesting that these proteins are fully competent in base-flipping. Surprisingly, a further 2-fold gain in fluorescence intensity was observed with the T250G mutant. The attained amplitude approaches the fluorescence level of free 2AP-deoxyribose in solution (14). In contrast, the T250D mutant showed a clear reduction (60% of the WT) in fluorescence intensity under identical conditions.

All mutants exhibited similar emission spectra at the primary excitation wavelength of 2AP λ_{Ex} = 320 nm (Fig. 3B, Table III). Excitation spectra at λ_{Em} = 370 nm show maxima for the protein tryptophan band near 284 nm (not shown) and the primary 2AP maximum at around 310 nm. The 2AP maximum is similar in the presence of WT and most mutants. However, the T250S and T250G mutants show peaks that are red-shifted by 3 and 8 nm, respectively. This observation is consistent with the notion that the excitation maximum is a sensitive indicator of the environment of 2AP (35). However, scarce information is available to correlate fluorescence excitation spectra with structural peculiarities of flipped out 2AP residues in protein-DNA complexes (36). Duplex-to-coil transition in DNA leads to a 2–5-fold increased fluorescence intensity (33, 37) and a blue shift of the excitation maximum from ~315 to ~305 nm (35, 38). Since melting of duplex DNA disrupts inter-strand base pair interactions, but leaves a substantial amount of intra-strand base stacking interactions, it was suggested that the excitation blue shift is due to the lost hydrogen bonding to N-1 of the 2AP base (38). For comparison, base-flipping by WT M.HhaI leads to a 50-fold higher intensity (14) and a similar excitation blue shift (Fig. 3B), consistent with both targeted base pair disruption and unstacking of the base. The T250G mutant confers an additional 2-fold gain in fluorescence intensity, but no blue shift in the excitation spectrum is observed as compared with the base-paired 2AP base. Therefore, the strongly enhanced intensity and the absence of the excitation blue shift in the M.HhaI(T250G)-GPGC/GMGC complex would be consistent with a model whereby the target base is located in a more polar environment and is hydrogen bonded at N-1 (to a nearby residue in the protein, the DNA, or a bound water molecule). However, it should be noted that the magnitude of the blue shift observed upon loss of a pairing partner can vary in the range from 0 to 8 nm, depending on a sequence context (35).

To assess the solvent accessibility of the flipped out 2AP base in the binary MTase-DNA complexes, fluorescence quenching experiments with acrylamide were performed. In all cases, the quenching curves followed the Stern-Volmer function fairly well. Control gel mobility shift binding experiments (not shown) indicated that the stability of corresponding binary complexes is not affected by the presence of acrylamide at
concentrations as high as 400 mM confirming that the observed fluorescence changes were largely due to quenching of the fluorophore. Remarkably, the $K_s$ values (Table III) obtained from linear slopes of $F_0/F$ versus acrylamide concentration (Fig. 4) showed a clear correlation with the side chain size of residue 250: the 2AP base is best exposed to bulk solvent in the T250H mutant, whereas the T250G and T250S mutants confer a 5–6-fold lower accessibility of the flipped-out base.

**Kinetics of Base-flipping**—The unexpected difference in the equilibrium fluorescence signal from the target base in some of the mutants led us to compare their base-flipping kinetics. Stopped-flow experiments were performed in which the fluorescence intensity was measured following rapid mixing of enzyme with the GPGC/GMGC duplex in a stopped-flow device. A series of fluorescence time courses at different protein and fixed fluorophore concentrations were measured. The obtained progress curves (see Fig. 5) were fitted to exponential equations using a least-squares approximation. In most cases, biphasic transitions were observed. The first phase was dominant (75–80% of the total signal) and its rate varied linearly with protein concentration up to the high limit of the stopped-flow apparatus, ~1000 s$^{-1}$. The second smaller phase ($k_{obs} = 20–30$ s$^{-1}$) was essentially concentration independent. Detailed kinetic analysis of interactions between M.HhaI and the fluorescent DNA duplex containing 2AP is presented elsewhere.$^4$ The binding kinetics is best described by a two-step sequential mechanism. The first step is a combination of diffusion-controlled binding of MTase to its target site ($k_{on} = 10^9$ M$^{-1}$ s$^{-1}$) followed by almost instantaneous ($k_{flip} > 2000$ s$^{-1}$) flipping of the target 2AP base to form complex [MTase-DNA$^1$]; these two events cannot be kinetically separated by the present data. The second step is described by the second transient, which we attribute to a conformational rearrangement of the initial flipped-out complex [MTase-DNA$^1$] into complex [MTase-DNA$^2$] with a higher fluorescence intensity of the flipped-out 2AP base,

\[
\text{MTase} + \text{DNA} \leftrightarrow [\text{MTase-DNA}^1] \leftrightarrow [\text{MTase-DNA}^2]
\]

$^4$ S. Serva, G. Vilkaitis, E. Weinhold, and S. Klimašauskas, manuscript in preparation.
A similar two-phase kinetics of 2AP flipping is also characteristic of the adenine MTase EcoRI (39). Analysis of the mutants showed that four of them (T250C, T250H, T250N, and T250S) were fairly similar to the WT enzyme in this respect (not shown), and the total signal amplitudes paralleled those observed in fluorescence titrations. Remarkably, the T250G mutant behaved in a different manner. Two transients with similar rates and concentration dependence were again observed, however, their amplitudes were dramatically different. The first phase showed an intensity of 50–60% of that of the WT; the second amplitude was 6–10-fold bigger than that of the WT MTase comprising 60–70% of the total signal. In agreement with the titration data and the observed differences in excitation spectra (Fig. 3B), the total fluorescence amplitude at λex = 312 nm was 16-fold higher than that of the WT MTase. In the case of the T250D mutant, the second transient was barely detectable, although kinetics of the first phase was similar to those described above.

In the simplest model, the observed differences in fluorescence intensities might be accounted for different occupancy of the F2 state (Scheme 1) that would be enhanced in the binary complex of T250G and reduced in the binary complex of T250D. A quantitative assessment of internal F1 ↔ F2 equilibrium requires knowledge of the relative fluorescence intensities of both complexes which cannot be determined directly with our measurements. However, if one presumes that [MTase-DNAF1] and [MTase-DNAF2] are structurally similar for the MTases, then these complexes will have similar fluorescence intensities in all MTases under consideration. Taking also into account that the total fluorescence of the WT binary complex is ~1/2 of that of free 2AP in solution (14), the ratio of fluorescence intensities of the two species F2/F1 is ≤2. Based on this assumption we performed simultaneous fitting of entire data sets of the WT, T250G, and T250D MTases to the two-step kinetic model (Scheme 1) using DynaFit kinetic analysis program (27). Typically, the residuals of fits were in the order of 10−4 for F2/F1 values in the range from 1.3 to 3. For F2/F1 = 2 (Fig. 5), the occupancy of the F2 state defined as YF2 = [F2]/([F1]+[F2]) = k2/(k+ + k−2) refined at 18% for WT, 90% for T250G, and 10% for the T250D mutant (Table IV). Although the absolute YF2 and k5 values were somewhat dependent on the selected F2/F1 ratio, the relative numbers essentially remained constant. Therefore, we conclude that the observed behavior of the mutants can be described by a two-step kinetic mechanism with two types of binary MTase-DNA complexes (two states of the flipped-out 2AP base) in equilibrium with each other. The first species (F1) is dominant in the case of WT enzyme and the T250C, T250N, T250H, and T250S mutants. The removal of the side chain of Thr-250 (in T250G) leads to a severalfold higher conversion of the flipped-out 2AP base to the second state (F2). The T250D shows the opposite effect, i.e. a suppressed transition to the second conformational state F2.

Crystallographic Analysis of Ternary M.HhaI(T250G)DNAAdoHcy Complex—To obtain structural insights into the effects of the Gly for Thr substitution, the structure of the T250G mutant in a ternary complex with a 13-mer palindromic DNA duplex and AdoHcy was determined by x-ray crystallography (Protein Data Bank entry 1FJX). An all-atom superposition of the protein components of 1FJX and 3MHT (containing native M.HhaI, AdoHcy, and a DNA identical in sequence to the DNA in 1FJX) yields an root mean square deviation of 0.4 Å. The nucleic acid components of 1FJX and 3MHT are also similar, with an root mean square deviation of 0.5 Å over all atoms. The target cytosine is rotated out of the helix and into the catalytic pocket of the enzyme as seen in each of the previously solved DNA-M.HhaIcofactor co-crystal structures. The introduction of Gly at position 250 did not grossly distort the structure of the complex. The space left by the side chain of Thr-250 is occupied by two water molecules (Fig. 6). These water molecules mediate hydrogen bonding interactions between protein (side chain of Arg-165 and main chain carbonyls of Lys-162 and Gly-250) and DNA backbone (O-5′-oxygen of the target cytosine and the phosphate of the upstream guanine). The waters, however, make no direct contacts with the flipped-out base and are very unlikely to account for the large increase in fluorescence quantum yield of 2AP observed in the binary complex. The largest distortions occur at the interface of the complex. The space left by the side chain of Thr-250 is occupied by two water molecules (Fig. 6). These water molecules mediate hydrogen bonding interactions between protein (side chain of Arg-165 and main chain carbonyls of Lys-162 and Gly-250) and DNA backbone (O-5′-oxygen of the target cytosine and the phosphate of the upstream guanine). The waters, however, make no direct contacts with the flipped-out base and are very unlikely to account for the large increase in fluorescence quantum yield of 2AP observed in the binary complex. The largest distortions occur at the interface of the ApG step on the target strand of DNA and the contacting protein residue: 1.7 Å by the NH3+ group of Lys-162, 1.2 Å by the 5′-Gua phosphate oxygen, 1.5 Å by the main chain carbonyl oxygen of Lys-162. Although these changes are relatively small, both the site and direction of the observed dislocations are consistent with the absence of the side chain in the mutant protein.

**DISCUSSION**

Conservation of an isolated amino acid residue or a short stretch of residues surrounded by variable sequences in a protein often indicates that the conserved element plays an im-

**TABLE IV**

| MTase   | Microscopic rate constants | Occupancy of state F2 |
|---------|---------------------------|----------------------|
|         | k1 | k2 | k−2 | YF2 | k2/(k1+k−2) |
| WT      | 1.3 ± 0.04 | 3 ± 0.4 | 14 ± 1.7 | 0.22 | 0.18 |
| T250G   | 1.0 ± 0.03 | 28 ± 1.1 | 3 ± 0.34 | 9.5 | 0.90 |
| T250D   | 0.9 ± 0.04 | 1.4 ± 0.4 | 13 ± 4 | 0.11 | 0.10 |

* Rate constants k1, k2, and k−2 correspond to mechanism shown in Scheme 1; k−1 is not reliably defined by the data.
important role for its function or structural organization. Sequence alignment of C5-MTases at the C-terminal sections of their variable regions indicate that in over 80% of sequences a consensus T(L,V,I) dipeptide (Fig. 1C) can be identified within 12–40 residues upstream of the sequence motif IX. The x-ray structures available for two enzymes from this class (11, 22) confirm that this conservation is preserved at the structural level (Fig. 1A), very likely, in majority of C5-MTases. The present study examines the functional importance of the conserved Thr-250 in the HhaI MTase by a detailed examination of six mutant proteins.

Enzymatic cytosine-C5-methylation involves many molecular events including massive conformational rearrangements in both the DNA and the protein (8, 11). The majority of these changes occur upon association of M.HhaI with DNA to form a binary complex. Solution studies employing 31P NMR, gel mobility shift analysis (16), and H5-proton exchange kinetics (25) show that the binary MTase-DNA complex is a dynamic molecule with multiple internal equilibria: reversible flipping of the target base out of the DNA helix and motions of the 20-residue active site loop toward the DNA (11). Consequently, the target base populates multiple conformational states spanning from its original position in the DNA helix to its final position in the catalytic pocket. Overall, the flipped-out base appears in contact with bulk solvent, especially if the catalytic loop is displaced from the target base. This notion is indirectly supported by the x-ray analysis of the ternary T250G-DNA:AdoHcy complex (see below). The other five variants (Ser, Cys, Thr, Asn, and His) are neutral in this respect, suggesting that the size of the side chain can be varied widely, and a group as small as hydroxymethyl is sufficient for maintaining a proper conformation of the nucleotide. In contrast, the T250D mutant shows suppressed transition of the target base into the second conformational state observed for T250G and several 10-fold decreased affinity toward the DNA (higher $K_{m}^{DNA}$ and $K_{m}^{DNA}$ values). Taking into account that the T250N mutant shows none of these effects, the observed differences can be attributed to the negatively charged carboxylate group. The negative charge should cause repulsive interactions with the nearby backbone phosphates on the target DNA strand, and/or attractive interactions with nearby Lys-162 and Arg-165 residues. This might in turn lead to a displacement of the DNA backbone and the target base in a different direction as compared with that observed in the M.HhaI(T250G):GPGC complex. From this analysis we conclude that the side chain of Thr-250 plays a role in shaping the conformation of the DNA backbone in the vicinity of the target base which in turn affects the selection of a preferred conformational state of the flipped out base.

However, what can we tell about the location of the flipped-out base in the binary M.HhaI(T250G):GPGC complex? The enhanced quantum yield of the 2AP residue in the T250G mutant is a clear indication that the target base is located in a more polar environment (33). The red-shifted fluorescence excitation maximum might be indicative of hydrogen bonding at N-1 of the 2AP base (38). And finally, the base turns out to be much less accessible to exogenous quenchers, which points to its more hidden position in the molecule. From a structural standpoint, the above three requirements may seem incompatible. One possible solution to this enigma appears to be that the target base is bound in the catalytic site of enzyme. The previous crystal structure for WT M.HhaI in complex with a mismatched DNA substrate that contains adenine at the target position (20) indicates that the adenine can be accommodated in the catalytic pocket of the enzyme. An analogous stable complex with guanine could not be obtained, most likely due to a presumed steric clash between the 2-amino group of the base and the side chain of Arg-165 (20). In this context, one would assume that 2AP would not fit in the catalytic site of the WT enzyme, because the base also contains an amino group at the 2-position of the ring. In the T250G mutant, the side chain of Arg-165 should be able to move slightly toward the newly available space, thereby creating more room in the catalytic site and permitting the 2-amino group of the 2AP base to avoid the aforementioned unfavorable interactions. If bound in the active site in the manner observed for adenine, the N-1 atom of the base would face the side chain of the nearby Glu-119 which would potentially lead to a hydrogen bond similar to that observed at N-3 of cytosine in the native structures (40). Moreover, this model also suggests that the base would be largely protected from bulk solvent, especially if the catalytic loop adopts a closed conformation as detected in the adenine structure (20). Consistent with this speculation, our acrylamide fluorescence quenching studies show an inverse correlation between the size of the side chain of residue 250 and the accessibility of the target base in the binary complex.6 The high

6 Corresponding binary complexes containing thymine as the target base show permanganate reactivity of the target base (17) diminishing in the order T250H > WT > T250G (S. Serva and S. K., unpublished observations).
fluorescence intensity of the bound 2AP base can be explained by the fact that the base is held in the catalytic site by several hydrogen bonds directed to the edges of the ring. There are no direct contacts (shorter than 4 Å) from non-polar residues to the faces of the ring (see Fig. 3B in Ref. 20), and, therefore, one might expect that relaxation of the excited fluorophore via static energy transfer or collisional encounters with solvent will be relatively weak (28). Altogether, these considerations suggest that besides the effects on the conformation of DNA backbone, the removal of the side chain of residue 250 might have substantial effects on structural rigidity of the active site, and its accessibility for binding by non-canonical bases.

Upon binding of the cofactor AdoMet or its analogues such as the product AdoHcy, the reaction complex becomes more compact (18) with clearly defined structural features (16) that have been visualized crystallographically (11, 20, 30, 41). In the ternary T250G-DNA-AdoHcy complex, the target cytosine is firmly locked in the catalytic complex and superimposes with that in the native molecule (Fig. 6). Although the small dislocation of the DNA backbone in the vicinity of the missing side chain of residue 250 may be reminiscent of the backbone distortion envisioned for the binary complex, there are no structural indications that subsequent steps during catalysis would be substantially affected by the T250G mutation. Indeed, biochemical and kinetic studies show that replacement of Thr-250 with structurally similar or smaller residues (Cys, Ser, and Gly) leads to MTases that are very similar to the WT enzyme.

Based on the overall catalytic efficiency $k_{cat} / K_m$ (also called specificity constant) and other criteria such as $K_m$, $K_{DNA}$ the four mutants and the WT enzyme are virtually indistinguishable (see Table II). The catalytic efficiency of an enzyme is thought to be one of the most important parameters at which evolutionary pressure is exerted (42). In line with this reasoning, our sequence analysis indicates that besides Thr, its closest homologues Ser and Cys can also be found at this position. However, it is not quite clear why Thr so much prevails over Ser or Cys in bacterial C5-MTases, in contrast to a quite even distribution of the 3 amino acids in eukaryotic enzymes (Fig. 1C). We cannot exclude that M.HhaI is more tolerant to structural variations at the conserved threonine residue than most other members of the C5-MTase class. Interestingly, high mutability of conserved residues in the catalytic site of DNA polymerases has recently been reported (43).

On the other hand, substitutions that introduce bulkier side chains at position 250 (Asn, Asp, and His) lead to more pronounced and diverse effects. From analyzing x-ray structures of the ternary M.HhaI complexes, one would expect that these latter substitutions introduce moderate perturbations in the catalytic center of the enzyme. Indeed, we find that these mutations manifest themselves at the level of ternary complex formation and affect the catalytic efficiency largely via increased $K_m$ for AdoMet and DNA. From this perspective, it is clear why the C5-MTases avoid residues with bulky side chains at position 250. Other factors that may be important for the selective advantage are stability and solubility of a protein. Although the present study did not address this question directly, our failure to solubilize the T250A, T250V, and T250P mutants may indirectly indicate that variants with hydrophobic side chains are prone to misfolding and/or aggregation and are thus selected against during adaptive evolution.

Acknowledgments—We are grateful to B. Holz for discussions, V. Siksna for comments on the manuscript, and MBI Fermentors for continuous support.

REFERENCES

1. Vertino, P. M. (1999) in S. Adenosylmethionine-dependent Methytransferases: Structures and Functions (Cheng, X., and Blumenthal, R. M., eds) pp. 341–372, World Scientific, Singapore.
2. Colot, V., and Rossignol, J. L. (1998) Bioessays 19, 702–708.
3. Robertson, R. D., and Jones, P. A. (2000) Carcinogenesis 21, 461–467.
4. Dryden, D. F. F. (1999) in S. Adenosylmethionine-dependent Methytransferases: Structures and Functions (Cheng, X., and Blumenthal, R. M., eds) pp. 283–340, World Scientific, Singapore.
5. Hittinger, J. M., Silverstein, R. L., Low, D. A., and Mahan, J. M. (1999) Science 284, 967–970.
6. Posfai, J., Bhagwat, A. S., Posfai, G., and Roberts, R. J. (1989) Nucleic Acids Res. 17, 2421–2435.
7. Kumar, S., Cheng, X., Klismasauskas, S., Mi, S., Posfai, J., Roberts, R. J., and Wilson, G. G. (1994) Nucleic Acids Res. 22, 1–10.
8. Cheng, X., Kumar, S., Klismasauskas, S., and Roberts, R. J. (1993) Cold Spring Harbor Symp. Quant. Biol. 58, 331–338.
9. Trautner, T. A., Balganes, T. S., and Pawlek, B. (1988) Nucleic Acids Res. 16, 6649–6658.
10. Klismasauskas, S., Nelson, J. L., and Roberts, R. J. (1991) Nucleic Acids Res. 19, 6183–6190.
11. Klismasauskas, S., Kumar, S., Roberts, R. J., and Cheng, X. (1994) Cell 76, 357–369.
12. Vassylyev, D. G., and Morikawa, K. (1997) Curr. Opin. Struct. Biol. 7, 103–109.
13. Roberts, R. J., and Cheng, X. (1998) Annu. Rev. Biochem. 67, 181–198.
14. Holz, B., Klismasauskas, S., Serva, S., and Weinhold, E. (1998) EMBO J. 17, 317–324.
15. Mi, S., Alonso, D., and Roberts, R. J. (1995) Nucleic Acids Res. 23, 620–627.
16. Klismasauskas, S., Sayers, T. K., Serva, S., and Wuithrich, K. (1998) EMBO J. 17, 317–324.
17. Serva, S., Weinhold, E., Roberts, R. J., and Klismasauskas, S. (1998) Nucleic Acids Res. 26, 3473–3479.
18. Klismasauskas, S., and Roberts, R. J. (1995) Nucleic Acids Res. 23, 1388–1395.
19. Yang, A. S., Shen, J.-C., Ziegler, J.-M., Mi, S., and Jones, P. A. (1995) Nucleic Acids Res. 23, 1380–1387.
20. O’Gara, M., Horton, J. R., Roberts, R. J., and Cheng, X. (1998) Nat. Struct. Biol. 5, 872–877.
21. Lauster, R., Trautner, T. A., and Noyer-Weidner, M. (1989) J. Mol. Biol. 206, 305–312.
22. Reinisch, K. M., Chen, L., Verdine, G. L., and Lipscomb, W. N. (1995) Cell 82, 143–153.
23. Cheng, X., and Blumenthal, R. M. (1996) Structure 4, 639–645.
24. Roberts, R. J., and Marelis, D. (2000) Nucleic Acids Res. 28, 306–307.
25. Wu, J. C., and Santi, D. V. (1987) J. Biol. Chem. 262, 4778–4786.
26. Leatherbarrow, R. J. (1992) Graphit, Version 3.0, Erithacus Software Ltd., Staines, United Kingdom.
27. Kuzmic, P. (1996) Anz Biochem. 237, 260–271.
28. Lakowicz, J. R. (1999) Principles of Fluorescence Spectroscopy, Kluwer Academic/Plenum Publishers, New York.
29. Cousina, J. (1997) In KinetAssyst, Version 2.0. Hi-Tech Scientific Ltd., Salisbury, United Kingdom.
30. O’Gara, M., Roberts, R. J., and Cheng, X. (1996) J. Mol. Biol. 261, 597–606.
31. Brünger, A. T. (1992) X-FLOR, A System for X-ray Crystallography and NMR, Version 3.1, Yale University Press, New Haven, CT.
32. Kuzmic, P., Cheng, X., Blumenthal, R. M., and Roberts, R. J. (1992) Biochemistry 31, 8648–8653.
33. Ward, D. C., Reich, E., and Stryer, L. (1989) J. Biol. Chem. 264, 2368–2373.
34. Allan, B. W., Beechem, J. M., Lindstrom, W. M., and Reich, N. O. (1998) Proc. Natl. Acad. Sci. U.S.A. 95, 6164–6169.
35. Raney, K. D., Sowers, L. C., Miller, D. P., and Benkovic, S. J. (1994) Proc. Natl. Acad. Sci. U.S.A. 91, 6637–6642.
36. Xu, D., Evans, K. O., and Nordlund, T. M. (1994) Biochemistry 33, 9792–9799.
37. Allan, B. W., and Reich, N. O. (1998) J. Biol. Chem. 273, 2691–2697.
38. McCullough, A. K., Dodson, M. L., Schaefer, O. D., and Lloyd, R. S. (1997) J. Biol. Chem. 272, 2710–2712.
39. Kuzmic, P. (1996) Anz Biochem. 237, 260–271.
40. Erich, T. F. X., Flager, T. A., and Roberts, R. J. (1992) Science 250, 1076–1083.
41. Erdmann, D., Horst, G., Dasterhoft, A., and Krüger, M. (1992) Gene (Amst.) 117, 15–22.
Functional Roles of the Conserved Threonine 250 in the Target Recognition Domain of HhaI DNA Methyltransferase
Giedrius Vilkaitis, Aiping Dong, Elmar Weinhold, Xiaodong Cheng and Saulius Klimasauskas

J. Biol. Chem. 2000, 275:38722-38730.
doi: 10.1074/jbc.M005278200 originally published online August 23, 2000

Access the most updated version of this article at doi: 10.1074/jbc.M005278200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 37 references, 9 of which can be accessed free at http://www.jbc.org/content/275/49/38722.full.html#ref-list-1