BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2011,** *50,* 1465–1473 1465
DOI: 10.1021/bi101446g

Determinants of Precatalytic Conformational Transitions in the DNA Cytosine Methyltransferase M.HhaI[†]

Douglas M. Matje,[‡] Dylan F. Coughlin,[‡] Bernard A. Connolly,[§] Frederick W. Dahlquist,[‡] and Norbert O. Reich*[,‡]

[‡]*Department of Chemistry and Biochemistry, University of California, Santa Barbara, California 93106-9510, United States, and* [§]*Institute for Cell and Molecular Biosciences, University of Newcastle, Newcastle upon Tyne NE2 4HH, U.K.*

Received September 7, 2010; Revised Manuscript Received December 6, 2010

ABSTRACT: The DNA methyltransferase M.HhaI is an excellent model for understanding how recognition of a nucleic acid substrate is translated into site-specific modification. In this study, we utilize direct, real-time monitoring of the catalytic loop position via engineered tryptophan fluorescence reporters to dissect the conformational transitions that occur in both enzyme and DNA substrate prior to methylation of the target cytosine. Using nucleobase analogues in place of the target and orphan bases, the kinetics of the base flipping and catalytic loop closure rates were determined, revealing that base flipping precedes loop closure as the rate-determining step prior to methyl transfer. To determine the mechanism by which individual specific hydrogen bond contacts at the enzyme−DNA interface mediate these conformational transitions, nucleobase analogues lacking hydrogen bonding groups were incorporated into the recognition sequence to disrupt the major groove recognition elements. The consequences of binding, loop closure, and catalysis were determined for four contacts, revealing large differences in the contribution of individual hydrogen bonds to DNA recognition and conformational transitions on the path to catalysis. Our results describe how M.HhaI utilizes direct readout contacts to accelerate extrication of the target base that offer new insights into the evolutionary history of this important class of enzymes.

The first structurally characterized DNA cytosine methyltransferase, M.HhaI,[1] has provided many relevant insights into how DNA-modifying enzymes recognize and modify DNA sequence-specifically (*1*−*3*). M.HhaI transfers a methyl group to C^5 of the underlined cytosine in the GCGC sequence as part of the restriction modification system in *Haemophilus hemolyticus* using the cofactor *S*-adenosyl-L-methionine (AdoMet). In prokaryotes, DNA methylation is an epigenetic process involved in gene regulation, replication timing, and host defense (*4*, *5*). Cytosine methylation in eukaryotes plays critical roles in developmental regulation and tumorigenesis and is conducted by enzymes with a high degree of structural homology to M.HhaI (*6*−*8*). A better mechanistic understanding of the substrate recognition pathway of DNA cytosine methyltransferases is expected to form the basis of improved therapeutics (*9*) and novel nucleic acid manipulation techniques (*10*,*11*).

Despite numerous protein−DNA cocrystal structures for various sequence-specific DNA-modifying enzymes, our understanding of how recognition is achieved and translated into catalysis remains limited. Comparison of the ternary M.HhaI cognate DNA−cofactor−methyltransferase and the binary methyltransferase−cofactor structures in the presence of nonspecific DNA shows both DNA and enzyme undergo large structural rearrangements upon formation of the cognate complex (*12*, *13*). The ternary structure shows the target cytosine flipped 180° from the native helical position and stabilized in the active site (*1*). Residues 80−100 comprise a long flexible catalytic loop that exists in two distinct states both in the crystal structures and, as confirmed by NMR, in solution (*2*, *14*). Upon binding the cognate sequence, the tip of the loop, Ser87, travels 26 Å to occupy the space vacated in the DNA helix following base flipping. Loop closure is critical for tight binding of the enzyme to the cognate sequence as well as for stabilizing the target cytosine in position for methyl addition (Figure 1A) (*15*). Tryptophan residues engineered into the catalytic loop have provided a basis for real-time tracking of this induced-fit mechanism (*13*, *16*), revealing that individual hydrogen bond contacts between the enzyme and its substrate make dramatically different contributions to loop repositioning. Our interests lie in determining the basis for this induced-fit mechanism, one of the largest conformational rearrangements observed for an enzyme.

Here we determined the individual contributions of four previously unexamined hydrogen bond contacts to ternary complex formation, catalytic loop positioning, and methyl transfer. Similar to prior observations that describe the effects of removing two hydrogen bond contacts (*13*), the four recognition elements examined in this work have dramatically different contributions to binding, loop positioning, and stabilization of the catalytically competent intermediate. Furthermore, we show that loop closure is apparently limited by base flipping steps and reveal the kinetics of each step, providing insights into how DNA recognition generates the conformational transitions of base flipping and loop closure on the path to catalysis.

[†]This research was partially supported by University of California Systemwide Biotechnology Research and Education Program GREAT Training Grant 2008-25 awarded to D.M.M. and N.O.R.

*To whom correspondence should be addressed: Department of Chemistry and Biochemistry, University of California, Santa Barbara, CA 93106-9510. E-mail: Reich@chem.ucsb.edu. Phone: (805) 893-8368. Fax: (805) 893-4120.

[1]Abbreviations: M.HhaI, DNA cytosine C^5 methyltransferase from *Haemophilus hemolyticus*; AdoHcy, *S*-adenosylhomocysteine; AdoMet, *S*-adenosylmethionine; FRET, Förster resonance energy transfer; IPTG, isopropyl β-D-1-thiogalactopyranoside; EDTA, ethylenediaminetetraacetic acid; PDB, Protein Data Bank; Z, aebularine; M, 5-methylcytosine; X, abasic site.

Published on Web 01/13/2011

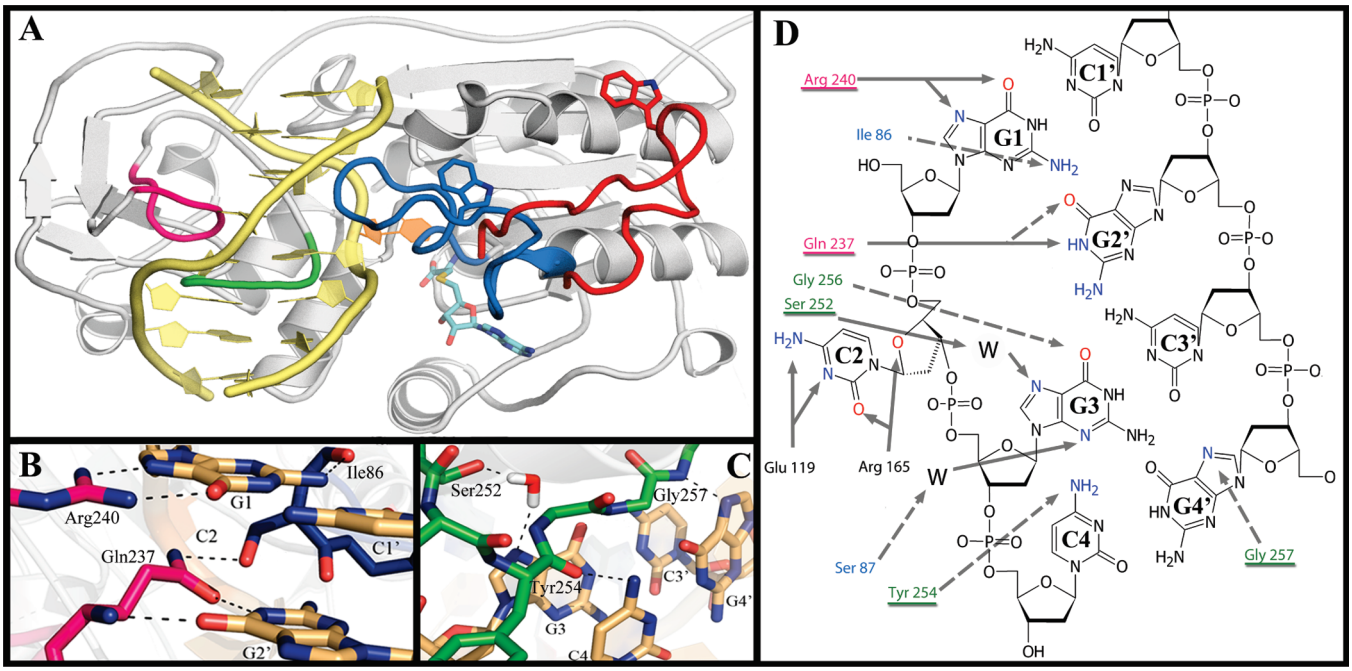

FIGURE 1: (A) Overlay of the binary (PDB entry 1HMY) and ternary (PDB entry 2HR1) crystal structures revealing a relatively stable structure in both forms, except for the catalytic loop consisting of residues 80−100 (root-mean-square deviation of 0.82 Å excluding loop residues). The open form of the loop from the binary structure is colored red, and the closed loop position in complex with cognate DNA is colored blue. Shown in cyan is the cofactor AdoHcy, and the orange extrahelical base is the target cytosine. Residues 236−240 are colored magenta, while residues comprising the 252−257 loop are colored green. Also visible is the strong distortion of the DNA backbone resulting from enzyme binding. (B) Recognition of G1 and stabilization of the orphaned G2′ by the 236−240 loop. The critical contacts from Arg240 to G1 are shown, along with the orphan G2′ and loop-closed stabilizing contacts made by Gln237. (C) Numerous contacts of the 252−257 loop. Only the water-mediated Ser252−G3 contact involves a side chain interaction. Panels A−C were generated with Pymol. (D) Hydrogen bond recognition network between M.HhaI and the cognate site. Solid lines represent side chain interactions, and dashed lines represent backbone contacts. Contacts by underlined residues were probed in this study, and residue colors correspond to loop colors in panel A.

MATERIALS AND METHODS

Unless noted, all reagents were purchased from Fisher Scientific. AdoHcy and AdoMet were purchased from Sigma-Aldrich and used without further purification. Radiolabeled reagents were purchased from Perkin-Elmer.

DNA Preparation. DNA substrates were purchased from Midland DNA and purified by HPLC or PAGE, except for sGCGZ, which was prepared, purified, and characterized as described previously (*17*, *18*, *19*). Concentrations were determined by the absorbance at 260 nm using the molar extinction values calculated using the nearest neighbor method with the algorithm available at http://biophysics.idtdna.com using the Cavaluzzi−Borer correction. Table 1 lists the synthetic oligodeoxynucleotides used in this study. Oligodeoxynucleotides were annealed by being heated to 65 °C followed by being slowly cooled to room temperature. Table 2 lists the duplex forms of the substrates used in the assays along with the contacts to the protein that are disrupted with the duplexes. Annealed substrates were examined by PAGE and observed to be >95% duplex via staining with SybrGold (Invitrogen). All duplex substrates were used in a hemimethylated complex to direct single-orientation binding.

Protein Preparation. The previously described K91W/W41F and E94W/W41F M.HhaI mutants (*13*, *16*) were used for this study. The plasmid containing the gene was transformed into and expressed in NEB T7 Express I^q *Escherichia coli* cells by IPTG induction after reaching an OD_{600} of 0.8. Induction was allowed to proceed for 4 h at 22 °C, when the cells were pelleted by centrifugation. Cell pellets were lysed by sonication in 20 mM sodium phosphate (pH 7.0), 400 mM NaCl, 10 mM imidazole, 0.1% Triton X-100, 200 μM PMSF, and 0.01 mg/mL DNase I. The lysate was clarified by centifugation and the supernatant applied to 1 mL of Ni-NTA (Qiagen) pre-equilibrated in lysis buffer. After the lysate had been loaded, the column was washed with 50 mL of lysis buffer with no PMSF or DNase. This was followed by washing with 15 mL of a buffer containing 20 mM sodium phosphate, 20 mM imidazole, and 800 mM NaCl. The enzyme was then eluted with 15 mL of a buffer containing 20 mM sodium phosphate, 250 mM imidazole, and 200 mM NaCl.

The eluted fraction was directly loaded onto a column of 10 mL of phosphocellulose P-11 resin (Whatman) pre-equilibrated with 20 mM sodium phosphate and 200 mM NaCl, washed with 50 mL of equilibration buffer, and eluted with 20 mM sodium phosphate and 800 mM NaCl. The eluted material was dialyzed 100-fold several times to move residual DNA and cofactor into a storage

Table 1: Single-Stranded Oligonucleotides Used in This Paper

name	sequence (5′−3′)	substitution
s-GCGC	CATGGCGCAGTG	none
s-G^{7C}CGC	CATGG^{7C}CGCAGTG	G1 → 7-dezaguanosine
s-GCG^{7C}C	CATGGCG^{7C}CAGTG	G3 → 7-deazaguanosine
s-GCGZ	CATGGCGZAGTG	C4 → zebularine
s-GXGC	CATGGXGCAGTG	C2 → abasic
s-CGMG	CACTGMGCCATG	C3′ → 5-methylcytosine
s-CGMG^{7C}	CACTG^{7C}MGCCATG	G4′ → 7-deazaguanosine
s-CIMG	CACTGMICCATG	G3′ → inosine
s-ACGC	CATGACGCAGTG	G1 → A
s-TGMG	CACTCGMGTCATG	C1′ → T
s-NS top	CATGAGCTAGTG	G1 → A, C4 → T
s-NS bot	CACTAGCTCATG	C1′ → T, G4′ → A

Table 2: Duplex Substrates Used To Probe Hydrogen Bond Contributions[a]

duplex substrate (5′ → 3′/3′ → 5′)	constituents	contact removed
GCGC/CGMG	s-GCGC + s-GMGC	cognate substrate
G^{7C}CGC/CGMG	s-G^{7C}CGC + s-GMGC	Arg240 to G1 N^7
GCG^{7C}C/CGMG	s-GCG^{7C}C + s-GMGC	Ser252 to G3 N^7
GCGZ/CGMG	s-GCGZ + s-GMGC	Tyr254 to C4 N^4
GCGC/CGMG^{7C}	s-GCGC + s-G^{7C}MGC	Gly257 to G4′ N^7
GCGZ/CGMG^{7C}	s-GCGZ + s-G^{7c}MGC	Tyr254 to C4 N^4 and Gly257 to G4′ N^7
GCG^{7C}C/CGMG^{7C}	s-GCG^{7C}C + s-G^{7C}MGC	Ser252 to G3 N^7 and Gly257 to G4′ N^7
GXGC/CGMG	s-GXGC + s-GMGC	target cytosine removed
GCGC/CIMG	s-GCGC + s-GMIC	C2 N^4 to G2′ O^6
ACGC/TGMG	s-ACGC + s-GMGT	noncogate substrate
AGCT/TCGA	s-NS top + s-NS bot	nonspecific substrate

[a]M = 5-methylcytosine. G = 7-deazaguanosine. Z = zebularine = 2-pyrimidinone 1-β-2′-deoxyriboside.

buffer containing 25 mM sodium phosphate (pH 7.5), 200 mM NaCl, 1 mM DTT, 1 mM EDTA, and 10% glycerol. The enzyme was concentrated using a YM-30 Centriprep (Millipore). Aliquots were frozen at −80 °C. Concentrations were determined by measuring the molar extinction coefficient of 25500 M^{-1} cm^{-1} at 280 nm.

Single-Turnover Analysis. All measurements were taken at 22 °C in 25 mM potassium phosphate (pH 7.5), 100 mM NaCl, and 1 mM EDTA [fluorescence buffer (FB)] in triplicate. Substrates for which turnover is slower than 0.05 s^{-1} were denoted by hand, while substrates with faster k_{chem} values were observed using a Kin-Tek RQF-3 apparatus. All assays contained 2.0 μM enzyme, 0.5 μM substrate, and 5.0 μM AdoMet with a 1:10 ratio of ^{3}H-labeled AdoMet to unlabeled AdoMet. Aliquots were quenched with an excess of 1% SDS in FB. Upon completion of the assay, quenched aliquots were spotted onto DE-81 filter paper and processed as previously described (*20*). Counts per minute (Y) were plotted as a function of time and the data fitted to the equation $Y = Y_0 + (Y_{max} - Y_0)(1 - e^{-kt})$, where Y_0 was determined by spotting an aliquot containing no enzyme and the plateau value (Y_{max}) calculated from the best fit to the data with Prism 5 (Graphpad).

Equilibrium Binding. K_D values were determined by the gel-shift method with varying enzyme concentrations. Double-stranded substrates were end labeled using T4 PNK (NEB) and [γ-^{32}P]ATP. After completion of the reaction, the enzyme was inactivated by being heated at 65 °C for 20 min, followed by slow cooling to anneal the DNA. Samples were diluted 20-fold in FB and run through a pre-equilibrated $P\overline{6}$ size exclusion resin (Bio-Rad) to purify the DNA from the labeling reagents.

The enzyme at varying concentrations encompassing the preliminarily determined K_D values was incubated for 15 min with the labeled substrate (~5 pM) and 500 μM AdoHcy in a buffer containing 25 mM sodium phosphate (pH 7.5), 50 mM NaCl, 0.5 mM EDTA, 0.2 mg/mL bovine serum albumin, and 5% glycerol at room temperature; 12% PAGE gels were cast with 45 mM Tris-Borate (pH 7.6) and 0.1 mM EDTA and run at 10 V/cm for 90 min after a 15 min prerun. Gels were dried and exposed to a Molecular Dynamics phosphor screen for imaging on a Typhoon Trio (GE). The ratio of bound to free DNA was plotted as a function of enzyme concentration and fitted to the equation $Y = X/(K_D + X)$, where Y is the fraction bound and X is the DNA concentration, using Prism. All values were determined by performing the gel-shift experiments in triplicate.

Time-Resolved Fluorescence. Real-time loop closing was observed using an Applied Photophysics SX.18MV stopped-flow reaction analyzer equipped with a single-channel emission photomultiplier tube positioned 90° from the excitation beam, powered to 400 V with a 320 nM cutoff filter. Samples were excited at 290 nm with a 5 mm slit width with 1000 time points taken for each experiment. Preincubated AdoHcy and enzyme were mixed with DNA substrates in FB at a final concentration of 50, 0.5, or 1.0 μM. At least five traces for each sample were averaged and analyzed using Prism and the equation $Y = Y_0 + (Y_{max} - Y_0)(1 - e^{-kt})$.

Equilibrium Fluorescence. All spectra were recorded at 22 °C using a Perkin-Elmer LB55 fluorimeter. Spectra were recorded with 1.0 μM enzyme and 50 μM AdoHcy. Samples were excited at 290 nm, and the fluorescence signal was recorded from 310 to 420 nm with an excitation slit width of 2.5 mm and an emission slit width of 4.5 mm. Spectra were recorded at 800 nm/min, and five spectra were recorded and averaged for each sample. Data were processed and averaged using Spekwin32 (*41*). Inner-filter effects were corrected by measuring the absorbance at 290 nm for each sample on a Shimadzu UV1700 spectrophotometer and applying the equation described in ref *21* in Microsoft Excel. Final graphs were generated using Prism.

RESULTS

To investigate the contributions of individual hydrogen bond recognition contacts in the precatalytic events of M.HhaI, we prepared synthetic DNA substrates with base analogues lacking the hydrogen bonding functional groups. Cognate, noncognate (one base pair different from cognate), nonspecific (substrates containing no sequence resembling cognate), and cognatelike (containing base analogues) sequences were annealed and used as substrates for all studies. Figure 1D shows a stylized summary of the contacts between the enzyme and cognate sequence observed in the ternary crystal structure (*1*). 7-Deazaguanosine (G^{7C}) was used at positions G1, G3, and G4′, while zebularine (Z), which lacks the exocyclic amine at position 4 of cytosine, was substituted at position C4. To investigate the base flipping and loop closure steps, the guanosine analogue inosine, lacking the 2-amino group, was substituted at position G2′. The abasic substrate retains the phosphate backbone and ribose ring at the target C2 but lacks the cytidine nucleobase (abbreviated X). All substrates contained 5-methylcytosine (M) at the target C on the bottom strand of the palindromic site (C3′) to ensure a single functional binding orientation.

Hydrogen Bonding Interactions at the Orphan Guanosine: The Relationship between Base Flipping and Loop Kinetics. To probe the relationship between the base flipping

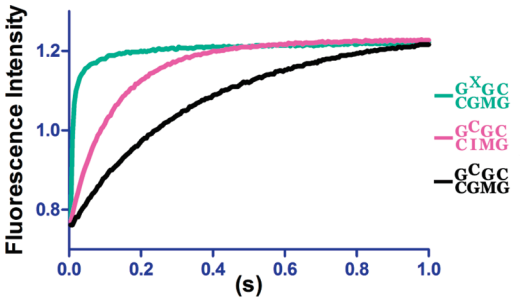

FIGURE 2: Comparison of the rates of loop closure for substrates differing at the target cytosine. GXGC (cyan) contains only the deoxyribose and phosphate at the target site, eliminating the base flipping step from the kinetic mechanism and allowing a more direct observation of the rate of loop closure. The magenta substrate contains an inosine base in place of G2′, removing a single hydrogen bond from the Watson−Crick pair while not perturbing any enzyme contacts. Kinetic parameters are listed in Table 3.

event and loop closure, we altered the barriers to base flipping at the target cytosine·guanosine base pair through selective disruption of the Watson−Crick base pairing interactions, which has been shown to increase substrate affinity in base flipping enzymes (*34*, *39*, *40*). Weakening the cytosine·guanosine base pair will enhance the rate of loop closure observed if this step is limited by base flipping. Rates of loop closure as reported by the fluorescence change were compared among the cognate substrate, a substrate lacking one hydrogen bond from the orphaned guanosine to the target cytosine, and a substrate lacking the target base altogether (Figure 2 and Table 3). Removing a single hydrogen bond between G2′ and C2 with the base analogue inosine results in a 2.8-fold increase in k_{loop}^{obs}, suggestive of base flipping limiting loop closure and a reduction in the barrier to flipping of 0.6 kcal/mol, less than the nominal value of 3 kcal/mol for the O^6-N^2 bond (*35*).

Complete removal of the base but retention of the phosphate backbone and deoxyribose ring with the abasic substrate results in much faster loop closing with an apparent rate constant of 150 s^{-1}, ~50-fold faster than the rate for cognate DNA. The rate constant observed for the abasic DNA complex is very close to the dead time of this stopped-flow instrument; thus, the actual rate of loop closure for this DNA substrate may be even faster. These observations are highly suggestive of a mechanism in which base flipping is slower than loop closure and dominates k_{loop}^{obs} in the cognate recognition pathway. Further investigation of the base flipping rate was conducted with native cofactor AdoMet that showed very similar results (Figure 4 of the Supporting Information), with the native cofactor resulting in a slightly faster k_{loop}^{obs}.

Effects of the Removal of Interactions on the Rate of Loop Closure. The previously characterized tryptophan loop reporter enzymes (K91W/W41F and E94W/W41F) allow for real-time, direct monitoring of loop repositioning with minimal perturbation of the kinetic and thermodynamic properties of the enzyme (*13*). As proposed in Scheme 1 and supported by the results described above, loop closure takes place after substrate−cofactor binding and base flipping and is the last known step before covalent attack by Cys81. Therefore, measurement of the observed rate of loop closure (k_{loop}^{obs}) reports on the rate of base flipping for the substrates studied.

Observation of loop closure was conducted in a stopped-flow apparatus by mixing an enzyme/AdoHcy solution with substrate and measuring the rate of change in the tryptophan fluorescence

Scheme 1: Precatalytic Steps of M.HhaI[a]

$$E^oS_M + D^H \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} E^oS_MD^H \underset{k_{stack}}{\overset{k_{flip}}{\rightleftharpoons}} E^oS_MD^F \underset{k_{open}}{\overset{k_{loop}}{\rightleftharpoons}} E^cS_MD^F \xrightarrow{\text{methyl transfer}} E^cS_HD^{F*}$$

[a]Enzyme in the loop open form, AdoMet, and fully helical DNA assemble at a rate close to the diffusion limit in random order (1×10^8 M^{-1} s^{-1}). Induced fit between the enzyme and target DNA results in capture of the extrahelical base (k_{flip} = 2 s^{-1}) in the active site, which is quickly followed by closing of the catalytic loop (k_{loop} = 120 s^{-1}). This prevents the flipped target cytosine from restacking into the helix and positions the catalytic Cys81 for nucleophilic attack and subsequent methyl transfer from AdoMet (k_{chem} = 0.2 s^{-1}). Because $k_{on} \gg k_{flip}$ or k_{loop}, the fluorescent signal observed upon cognate substrate binding is dominated by k_{flip}. Removal of this step with the abasic substrate analogue allows direct observation of k_{loop} as well revealing the rate of base flipping due to the difference in the rates of loop closure between the substrates containing and lacking a target base.

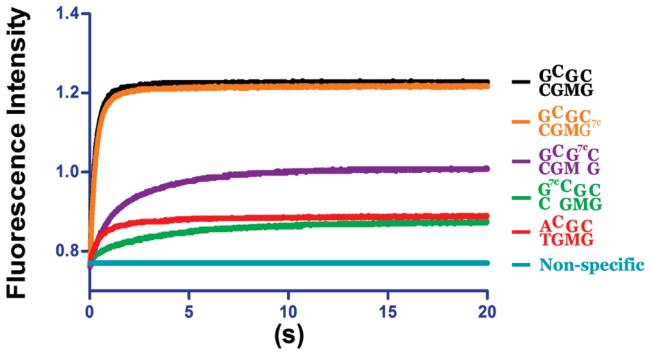

FIGURE 3: Real-time fluorescence monitoring of loop position in response to various substrates. The loop tryptophan of the K91W/W41F mutant undergoes a 60% increase in quantum yield after moving into the loop closed position in response to binding of cognate DNA in the presence of AdoHcy. Nonspecific DNA shows no change in fluorescence.

signal as a function of time. Changing the order of addition had no effect on the observed rates because of the near-diffusion limit formation of the ternary complex and mixed-order association (Figure 2 of the Supporting Information). Figure 3 shows the change in fluorescence signal that is observed when K91W/W41F M.HhaI and AdoHcy are mixed with individual substrates in a stopped-flow apparatus. Use of the cofactor product AdoHcy results in k_{loop}^{obs} rates that are 2−10-fold slower for substrates containing a target base and identical equilibrium positioning of the loop but without the complications of substrate turnover and polyphasic traces.

No change in fluorescence signal is observed with nonspecific DNA (Figure 3). Cognate DNA binding results in a 45% increase in the magnitude of the fluorescence signal that does not change in magnitude or rate above a 1:1 enzyme:substrate ratio at these concentrations (Figure 4 of the Supporting Information), showing that the bimolecular binding step is too fast to measure. Results with the alternative E94W/W41F mutant that shows a decrease in fluorescence upon binding to cognate DNA are shown in Figure 1 of the Supporting Information, demonstrating fluorescence equilibrium and rate changes complementary but reciprocal to those of the K91W/W41F mutant.

The apparent rate of loop closure with the GCGC/$CGMG^{7C}$ sequence is approximately equal to that of the cognate sequence, demonstrating the minimal role the Gly254−G4′ interaction plays in regulating loop positioning. Interruption of a contact

Table 3: k_{loop}^{obs} Values for Substrates with Lower Barriers to Base Flipping

substrate	orphan base	target base	K91W k_{loop} (s^{-1})	E94W k_{loop} (s^{-1})
GCGC/CGMG	guanine	cytosine	2.3 ± 0.01	5.0 ± 0.4
GCGC/CIMG	inosine	cytosine	6.5 ± 0.03	14 ± 0.6
GXGC/CGMG	guanine	none	120 ± 0.63	260 ± 11

adjacent to the target base with substrate GCG^{7C}C/CGMG results in a 5-fold reduction in k_{loop}^{obs}. Again, the substrate G^{7C}CGC/CGMG shows the largest change in the measured rate, a 10-fold decrease in the rate of loop closure illustrating the importance of this hydrogen bond to base flipping and loop closure.

The stopped-flow traces shown in Figure 3 reveal information about the equilibrium position of the loop in the ternary complex with cognatelike substrates. All traces shown in Figure 3 involve saturating substrate concentrations. The cognate trace shows the maximal change in fluorescence by the K91W/W41F mutant in response to binding, where it is expected that all or nearly all of the enzyme will be substrate-bound with the loop closed ($E^CS_MD^F$). These observations suggest that traces that do not reach this plateau represent ternary complexes that have altered equilibrium between the $E^OS_MD^H$ and $E^CS_MD^F$ states, likely affecting the ability of the enzyme to maintain a stable extrahelical position for the target cytosine.

If the cognate substrate is taken to represent 100% of the enzyme in the loop closed state, the GCGC/CGMG^{7C} substrate shows 98% of the enzyme to exist in the loop closed form, in agreement with the minimal disruption in binding, k_{loop}^{obs}, and the single-turnover rate. Substrate GCG^{7C}C/CGMG results in 44% loop closed species, also in agreement with the other measured parameters. G^{7C}CGC/CGMG and ACGC/TGMG show 27 and 20% of the bound enzyme, respectively, to exist in loop closed form. The decreases in the $E^CS_MD^F$ form of the ternary complex are likely reflective of the non-loop closed enzyme species partitioning between the $E^OS_MD^H$ and $E^OS_MD^F$ forms because of an inability to maintain the stable extrahelical base resulting from binding and DNA distortion energy lost because of the removal of contacts.

Unfortunately, zebularine is itself weakly fluorescent with an absorption at 300−320 nm and emission at 370 nm (*17*) and may generate a FRET pair with the engineered tryptophan functioning as the donor. This resulted in quenching of tryptophan fluorescence that prevents characterization with the loop position reporting mutants using the DNA sequence containing this modified base.

Changes in Protein−DNA Binding Affinity via Removal of a Single Hydrogen Bonding Interaction. To determine the binding energy contributions made by individual hydrogen bond interactions and confirm that the base analogue substitutions did not result in large-scale disruptions of the DNA structure, equilibrium dissociation constants (K_D) were determined using electrophoretic mobility shift assays in the presence of AdoHcy. Table 4 shows changes in K_D for each of the hydrogen bonds probed in this study and corresponding changes in the free energy of dissociation. The cognate sequence, GCGC/CGMG, was found to have a high affinity for the enzyme with a K_D of 10 pM, in agreement with several previous publications (*15*, *28*, *34*). In contrast, noncognate (ACGC/TGMG) and nonspecific (AGCT/TCGA) DNA binds with affinities between 100 and 1000 nM (*22*). Substrate GCGC/CGMG^{7C} results in a minor 2.7-fold loss of binding. Similarly, substrate GCGZ/CGMG also has a weak, ~6-fold loss of binding affinity (K_D), while GCG^{7C}C/CGMG has a 17-fold increase in the dissociation constant that demonstrates the moderate contribution to substrate recognition this water-mediated interaction plays.

The most significant disruption in binding comes from removing N^7 of G1, resulting in an 80000-fold decrease in binding affinity. In the ternary crystal structure, the side chain of Arg240 is precisely positioned to recognize O^6 and N^7 of G1 (Figure 1B). The decrease in binding affinity is equal in magnitude to that realized via removal of two hydrogen bonding pairs (e.g., GCGZ/CGMG^{7C}) or switching the first base pair from G1-C1′ to A1-T1′, a noncognate sequence. The two substrates that lack two hydrogen bonding partners (GCGZ/CGMG^{7C} and GCG^{7C}C/CGMG^{7C}) result in a very large decreases in binding affinity to approximately the same value exhibited by noncognate DNA. Both show an effect on the binding constant much greater than a simple additive effect predicted for the removal of the two bonds individually.

Effects of the Removal of Interactions on the Rate of Chemistry. Second-order formation of the ternary complex is very fast, approaching the diffusion limit for the cognate sequence (*27*) (data not shown). As the steady-state catalytic rate for M.HhaI is dominated by product release (*28*), measurement of the steps up to and including methyl transfer via single-turnover experiments is a more appropriate measure for how hydrogen bond perturbations translate to formation of a catalytically competent complex.

Single-hydrogen bond removal results in modest decreases in single-turnover kinetics for all but the G^{7C}CGC/CGMG substrate, again highlighting the importance of this contact toward binding and catalysis (Table 4). Removal of this interaction results in a 40-fold slower rate of methylation. GCGZ/CGMG, GCGC/CGMG^{7C}, and GCG^{7C}C/CGMG all show a <4-fold effect on k_{chem}. All three of these substrates have decreases in their chemistry kinetics that correspond approximately with their decrease in binding affinity. Thus, disruption of some recognition interactions independently alters base flipping and/or loop closure as well as the correct assembly of the active site prior to methyl transfer.

Surprisingly, GCGZ/CGMG^{7C}, which lacks both the G4′ and C4 hydrogen bonding interactions, exhibits an only 3.1-fold decrease in k_{chem}, suggesting that these recognition elements are nonessential for catalysis. In marked contrast, GCG^{7C}C/CGMG^{7C}, which lacks the N^7 of both G4′ and G3, shows a substantial 600-fold decrease in k_{chem}.

The single-turnover rate measured in these experiments is also reflective of the ability of the enzyme to maintain the extrahelical base because the transfer of a methyl group from AdoMet is the rate-limiting step in the single-turnover experiment. Thus, the percent reduction in the equilibrium population of the loop closed enzyme shown in Figure 3 is expected to alter the observed rates of methylation during single-turnover measurements (k_{chem}), if this alteration in the equilibrium population reflects the relative stabilization of the extrahelical base in the catalytically competent complex. This is indeed observed in our results, where substrate GCGC/CGMG^{7C} shows 98% loop closed but a k_{chem} that is only 68% of that of the cognate form. Substrate GCG^{7C}C/CGMG results in 44% loop closed enzyme but a 25% rate of single turnover; G^{7C}CGC/CGMG has 27% closed and a 2.2% k_{chem}, and ACGC/TGMG gives 20% loop closed enzyme with a single-turnover value of 5.7% as fast as that of the cognate substrate.

DISCUSSION

How DNA methyltransferases and endonucleases translate sequence-specific recognition into catalysis remains a subject of great interest for these important enzymes (*25*, *31*, *32*). M.HhaI−AdoHcy−cognate DNA crystal structures reveal approximately 10 direct readout contacts from amino acids involved in recognition of the DNA sequence that are complemented by another 12 phosphate contacts. Comparison of the ternary complexes with the M.HhaI−AdoMet binary complex reveals rigid catalytic and recognition domains except for a large translocation of residues 80−100 present in the catalytic loop (Figure 1A). Previous studies of M.HhaI have shown that some amino acid−base contacts are critical to substrate recognition, loop repositioning, and catalysis, while others can be removed with little consequence (*16*, *25*). This is at odds with studies of the restriction endonuclease EcoRI, where it was observed that most amino acid−base contacts have equal contributions toward binding and catalysis (*23*), but in agreement with data of EcoRV endonuclease, where large differences between contact contributions were observed (*38*). The tryptophan loop reporter enzymes K91W/W41F and E94W/W41F allow observation of the effects of hydrogen bond perturbation in the induced-fit mechanism that are not revealed by other means. This study reveals how single recognition elements in the cognate DNA sequence contribute to the conformational rearrangements in the substrate and enzyme preceding and including catalysis.

Removal of Base Flipping Barriers Accelerates k_{loop}^{obs} To Reveal Base Flipping and Loop Closure Rates. While it is perhaps the defining feature of DNA methyltransferases, kinetic characterization of the base flipping step has been described with only the nonisosteric base analogue 2-aminopurine in place of the target cytosine (*27*, *36*). We envisioned that reduction of the kinetic barriers to base flipping would result in changes in k_{loop}^{obs} that could be assigned to individual microscopic precatalytic rate constants. For example, decreases in the base flipping barrier should have no impact on loop kinetics if base flipping follows or is faster than loop rearrangement. This was accomplished via modification of the target C2·G2′ base pair, first removing one hydrogen bond with inosine in place of G2′ and then removing the cytidine of C2 to eliminate all but the ribose puckering barrier to base flipping (*34*). Substrates containing mismatches or the abasic site at the target base pair are bound tightly by M.HhaI, and the corresponding ternary structure with the abasic site (PDB entry 7MHT) is identical to the structure with native DNA (*34*). As shown in Figure 2 and Table 3, complete removal of the target base causes k_{loop}^{obs} to be increased 50-fold in comparison to the native cognate sequence. Similarly, moderate reduction of the energetic barrier of base flipping by removal of a single Watson−Crick hydrogen bond involved in base pairing results in an approximately 3-fold increase in the loop closure rate. The acceleration of k_{loop}^{obs} resulting from the reduction of the base flipping barrier is supportive of the precatalytic scheme shown in Scheme 1 in which base flipping precedes loop closure and is the rate-limiting step prior to nucleophilic attack by Cys81 at the target cytosine, in agreement with previous experimental studies (*6*, *16*, *27*). k_{loop}^{obs} for the cognate substrate is dominated by the rate of base flipping, which occurs at approximately 2−5 s^{-1} with AdoHcy at 25 °C. Removal of the base flipping barrier reveals the catalytic loop closure rate to be on the order of 120−260 s^{-1}, slower than loop closure rates observed with other enzymes, but not surprising for such a large conformational shift (*7*), particularly where this transition is not rate-limiting. Assignment of the rates of the base flipping and loop closure rates reveals the order of M.HhaI's precatalytic conformational transitions as well as informing our interpretation of the impact of single-hydrogen bond contact removal, revealing that the energy necessary for reduction of the base flipping barriers arises from the specific hydrogen bond contacts (see below).

Previous work with 2-aminopurine as the target base revealed a flipping/stabilization rate of 3 s^{-1}, in excellent agreement with our findings (*27*). The observed energy difference for the base flipping step, 2.3 kcal/mol between the abasic and cognate site, is an approximately 1 order of magnitude reduction of the 25 kcal/mol barrier presented by the native Watson−Crick pair, suggesting that distortion of the DNA structure facilitated by M. HhaI promotes flipping of the base into the enzyme pocket (*29*, *30*). Our results support a mechanism in which the enzyme greatly reduces the barrier to base flipping via distortion of the native DNA structure without directly utilizing the catalytic loop to push the base from the helix. Direct readout contacts from the enzyme are optimized to stabilize the DNA helix in the distorted form that promotes base flipping, and we show below that individual contacts can make various contributions to the achievement and stabilization of the base flipped, loop closed state.

Individual Hydrogen Bond Contacts Contribute Discrete Energy to Formation of the Ternary Complex. Substrates containing base analogues lacking hydrogen bonding recognition elements were used to investigate loop closure during the precatalytic steps (Figure 3 and Table 3). The engineered Trp reporter mutants allow direct observation of the transition from the $E^{O}S_{M}D^{H}$ form to the $E^{C}S_{M}D^{F}$ complex that encompasses the base flipping and loop closure steps (Scheme 1). Assembly of the initial $E^{O}S_{M}D^{H}$ complex (loop open, helical base) is faster than 500 s^{-1} with the saturated conditions used for these experiments (Figure 4 of the Supporting Information), in agreement with previous studies showing the cognate complex forms near the diffusion limit (*27*). Because base flipping dominates k_{loop}^{obs} (Table 3 and the discussion above), changes to the equilibrium between loop open and closed observed in the stopped-flow traces are representative of alterations to the base flipping equilibrium and rate. The congruence in the equilibrium positioning of the K91W/W41F and E94W/W41F traces validates our interpretation of the fluorescence signal being a direct reporter of loop position as opposed to microscopic environmental differences of the tryptophan in response to different substrates.

Figure 3 shows the fluorescence traces of the K91W/W41F mutant upon binding DNA. Recognition of cognate DNA results in a large increase in the magnitude of the signal, while noncognate DNA has no effect on the fluorescent signal. GCGC/CGMG^{7C}, which is bound and methylated at near-cognate rates, shows little perturbation in the stopped-flow trace. Substrates with poor affinity and catalytic parameters show corresponding reductions in the loop position equilibrium and k_{loop}^{obs} values. In all cases, the rate of loop closure observed for an individual substrate is 10−20 times faster than the rate of a single turnover (Tables 4 and 5). All traces show at least minor increases in the equilibrium population of the loop closed species resulting from the addition of DNA, in agreement with results obtained via NMR studies showing a distribution of loop open and closed species in the presence of the ACGC/TGMG sequence (*33*). Of interest is removal of the G3 N^{7}−Ser252 interaction (Figure 3, purple trace), where the trace shows a decreased base flipping rate

Table 4: Kinetic and Thermodynamic Values for Substrates Used in This Study

substrate	K_D (pM)	$\Delta\Delta G°_{diss}$ (kcal/mol)	k_{chem} (s^{-1})	$\Delta\Delta G°_{chem}$ (kcal/mol)	k_{chem}/K_D (mol^{-1} s^{-1})
GCGC/CGMG	10 ± 2.3	0.0	$(1.9 \pm 0.2) \times 10^{-1}$	0.0	1.9×10^7
GCG**Z**/CGMG	60 ± 17	1.1	$(8.9 \pm 0.8) \times 10^{-2}$	0.4	1.5×10^6
GCGC/CGM**G**^{7C}	28 ± 9.0	0.6	$(1.3 \pm 0.1) \times 10^{-1}$	0.2	4.6×10^6
GC**G**^{7C}C/CGMG	170 ± 80	1.7	$(4.8 \pm 0.5) \times 10^{-2}$	0.8	2.8×10^5
G^{7C}CGC/CGMG	$(8.3 \pm 1.0) \times 10^5$	6.7	$(4.3 \pm 0.7) \times 10^{-3}$	2.2	5.2
GCG**Z**/CGM**G**^{7C}	$(5.0 \pm 1.1) \times 10^5$	6.4	$(6.1 \pm 0.1) \times 10^{-2}$	0.7	1.2×10^2
GC**G**^{7C}C/CGM**G**^{7C}	$(5.0 \pm 0.9) \times 10^5$	6.4	$(3.2 \pm 0.9) \times 10^{-4}$	3.8	0.6
ACGC/**T**GMG	$(4.8 \pm 1.0) \times 10^5$	6.4	$(1.1 \pm 0.9) \times 10^{-2}$	1.5	2.3 × 10
AGCT/TCGA	$(7.8 \pm 1.0) \times 10^5$	6.7	not determined	–	–

Table 5: Observed Rates of Loop Closure for K91W/W41F and E94W/W41F Mutants

substrate	K91W k_{loop} (s^{-1})	$\Delta\Delta G°k_{loop}$ (kcal/mol)	E94W k_{loop} (s^{-1})	$\Delta\Delta G°k_{loop}$ (kcal/mol)
GCGC/CGMG	3.0 ± 0.02	0.0	4.9 ± 0.01	0.0
GCG**Z**/CGMG	not determined	–	–	–
GCGC/CGMG^{7C}	2.8 ± 0.02	0.1	4.6 ± 0.03	0.0
GC**G**^{7C}C/CGMG	0.46 ± 0.01	1.1	0.42 ± 0.02	1.5
G^{7C}CGC/CGMG	0.26 ± 0.01	1.4	0.23 ± 0.05	1.8
GCG**Z**/CGM**G**^{7C}	not determined	–	–	–
GC**G**^{7C}C/CGM**G**^{7C}	not determined	–	–	–
ACGC/**T**GMG	0.58 ± 0.02	1.5	0.24 ± 0.02	1.8
AGCT/TCGA	< 0.02	> 3	< 0.01	> 3.65

but an equilibrium in favor of stabilization of the base flipped state compared to the ACGC/TGMG trace (Figure 3, red trace), which shows that the rate of base flipping is moderately impacted but the degree of stabilization of the extrahelical base is significantly reduced. These findings are consistent with the implication of this residue in base flipping and its high degree of conservation across cytosine methyltransferases (*37*).

The Arginine 240−G1 Interaction Dominates Formation of the Ternary Complex and Methyl Transfer. Cognatelike substrates lacking single hydrogen bonding groups were used to probe the contributions of the direct readout interactions to ternary complex formation and catalysis without perturbing the structure of the substrate or enzyme. Noncognate sites show a narrow binding regime, with K_D values of ~10−300 nM compared to 10−100 pM for the cognate site (*22*). Nonspecific DNA containing no GpC sites binds with a K_D of ~400−1000 nM (Table 4) (*2*). Our results show a large range in contributions of individual recognition elements to binding and catalytic turnover. Of particular interest are the minimal contribution provided by the individual interactions between the C4·G4′ base pair to Try254 and Gly257 and the profound importance of the Arg240−G1 interaction (Table 4).

Table 4 shows the dissociation and single-turnover constants that were obtained for the substrates tested in this study. The data show the small contribution to binding by the C4 and G4′ interactions, as probed by substrates GCGZ/CGMG and GCGC/CGMG^{7C}. Both of the interactions are derived from the peptide backbone along the 252−257 loop in the major groove (Figure 1C). Simultaneous removal of both of these contacts results in a reduction in binding affinity to the level of noncognate DNA, but a minimal reduction in k_{chem}. These results are in agreement with the generation of the M.HhaI fusion protein lacking these contacts having a near-wild-type k_{cat} on the GCG sequence but a greatly reduced K_{MDNA} and relaxed specificity (*24*, *25*). In all other cases, the decrease in the association energy seen with cognatelike substrates parallels energetic reductions in the single-turnover rate, but 2−3-fold lower in magnitude.

Simultaneous removal of the interactions of G3 and G4′ with the GCG^{7C}C/CGMG^{7C} substrate results in dissociation constants comparable to those of noncognate DNA or removal of both the C4−G4′ contacts with GCGZ/CGMG^{7C}. However, the GCG^{7C}C/CGMG^{7C} substrate is methylated very poorly, 2 orders of magnitude slower than the noncognate ACGC/TGMG substrate. It is likely that the multiplicative reduction in k_{chem} observed with this substrate relative to the reduction in the single-mutant rate results from disruption of both contacts at the ends of the 252−257 region that greatly destabilizes the ability of this recognition loop to orient into its preferred position along the major groove of the target sequence (Figure 1A, green). Misplacement of this region would also break the Tyr254−C4 contact, thus perturbing at least three sequence-specific hydrogen bond contacts.

Most significantly, removal of N^7 from G1 results in pronounced changes in K_D and k_{chem}. Disruption of the Arg240−O^6 interaction (Figure 1B) at this base was previously shown to have consequences similar to those of binding (*13*). k_{chem} is reduced 7 orders of magnitude for this substrate. Removal of each single interaction between Arg240 and G1 has an impact on binding as significant as mutation of residue 240 to alanine (*13*). The dramatic reductions in the levels of binding and catalysis with removal of the Arg240−G1 N^7 interaction probed in this study and the previously characterized Arg240−G1 O^6 interaction are clear evidence that the correct positioning of the arginyl guanidinium group to donate a pair of hydrogen bonds to the O^6 and N^7 groups of G1 is an important recognition element for substrate binding, loop closure, and catalysis.

Our results are in good agreement with earlier efforts to define M.HhaI recognition elements (*13*, *22*) but describe the effect of removing four contacts that have not been probed by any previous investigations. Random mutagenesis showed that Ser252 and Tyr254 mutations were well tolerated (*24*), and

directed evolution efforts allowed both C4−G4′ contacts to be removed, generating a GCG/CGC methylating mutant with catalytic efficiency only 10-fold lower than that of the wild-type enzyme (*25*). Molecular dynamics-based examination of the energetics of substrate recognition and base flipping also found the Arg240−G1 interaction to provide the largest contribution to binding energy (*3*). Similarly, base substitution at G1 or C4 of the GGCC sequence recognized by the highly homologous M.HaeIII enzyme perturbs catalysis only 10−100-fold, while base substitution at G2 results in no methylation (*26*). The strong energetic and catalytic bias displayed by contacts at G1 in M.HhaI and G2 in M.HaeIII are suggestive of an ancestral C^5 DNA methyltransferase in which an induced-fit mechanism involved the recognition of the simple palindromic GpC sequence. Base flipping is driven by recognition of this simple two-base site to distort the DNA and stabilize the extrahelical position, while divergence of the target recognition domain toward larger recognition sequences has allowed prokaryotic speciation via coevolution of restriction enzymes without necessitating large changes in the catalytic mechanism to gain new specificities.

The results presented here demonstrate the order and rate of the precatalytic steps and offer new insights into how the hydrogen bonding network at the recognition interface is translated into conformational rearrangements on a path to DNA modification. Understanding how recognition drives sequence-specific catalysis can guide future efforts aimed at engineering new specificity into DNA methyltransferases while improving our understanding of this important enzyme family.

ACKNOWLEDGMENT

We thank Dr. August Estabrook for assistance in interpreting results and Celeste Holz-Schietinger for insightful discussions about and review of the manuscript.

SUPPORTING INFORMATION AVAILABLE

Traces of the E94W/W41F mutant with various substrates lacking hydrogen bonds, K91W/W41F fluorescence traces demonstrating that the k_{loop}^{obs} values measured are not dependent on the order of addition, primary data for the K91W/W41F mutant fluorescence changes when combined with SAM and DNA substrates, and a representative series of traces describing the K91W/W41F mutant fluorescence response when various concentrations of cognate DNA are titrated with the enzyme and cofactor. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Klimasauskas, 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.
2. Zhou, H., Purdy, M. M., Dahlquist, F. W., and Reich, N. O. (2009) The Recognition Pathway for the DNA Cytosine Methyltransferase M.HhaI. *Biochemistry 48*, 7807–7816.
3. Huang, N., and MacKerell, A. D., Jr. (2005) Specificity in Protein-DNA Interactions: Energetic Recognition by the (Cytosine-C5)-methyltransferase from HhaI. *J. Mol. Biol. 345*, 265–274.
4. Casadesús, J., and Low, D. (2006) Epigenetic Gene Regulation in the Bacterial World. *Microbiol. Mol. Biol. Rev. 70*, 830–856.
5. Cheng, X. (1995) Structure and Function of DNA Methyltransferases. *Annu. Rev. Biophys. Biomol. Struct. 24*, 293–318.
6. Jeltsch, A. (2002) Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. *ChemBioChem 3*, 274–293.
7. Jia, D., Jurkowska, R. Z., Zhang, X., Jeltsch, A., and Cheng, X. (2007) Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature 449*, 248–251.
8. Cheng, X., and Blumenthal, R. M. (2008) Mammalian DNA Methyltransferases: A Structural Perspective. *Structure 16*, 341–50.
9. Mashhoon, N., Pruss, C., Carroll, M., Johnson, P. H., and Reich, N. O. (2006) Selective Inhibitors of Bacterial DNA Adenine Methyltransferases. *J. Biomol. Screening 11*, 497–510.
10. Buryanov, Y., and Shevchuk, T. (2005) The use of prokaryotic DNA methyltransferases as experimental and analytical tools in modern biology. *Anal. Biochem. 338*, 1–11.
11. Pljevaljcic, G., Schmidt, F., and Weinhold, E. (2004) Sequence-specific Methyltransferase-Induced Labeliing of DNA (SMILing DNA). *ChemBioChem 5*, 265–269.
12. 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.
13. Estabrook, R. A., Nguyen, T. T., Fera, N., and Reich, N. O. (2009) Coupling Sequence-specific Recognition to DNA Modification *J. Biol. Chem. 284*, 22690–22696.
14. Zhou, H., Shatz, W., Purdy, M. M., Fera, N., Dahlquist, F. W., and Reich, N. O. (2007) Long-range structural and dynamical changes induced by cofactor binding in DNA methyltransferase M.HhaI. *Biochemistry 46*, 7261–7268.
15. Estabrook, R. A., Lipson, R., Hopkins, B., and Reich, N. (2004) The Coupling of Tight DNA Binding and Base Flipping. *J. Biol. Chem. 279*, 31419–31428.
16. Estabrook, R. A., and Reich, N. (2006) Observing an Induced-fit Mechanism during Sequence-specific DNA Methylation. *J. Biol. Chem. 281*, 37205–37214.
17. Connolly, B. A. (1992) Synthetic oligodeoxynucleotides containing modified bases. *Methods Enzymol. 211*, 36–53.
18. Zhou, L., Cheng, X., Connolly, B., Dickman, M., Hurd, P., and Hornby, D. (2002) Zebularine: A Novel DNA Methylation Inhibitor that Forms a Covalent Complex with DNA Methyltransferases. *J. Mol. Biol. 321*, 591–599.
19. Connolly, B. A., and Newman, P. C. (1989) Synthesis and properties of oligonucleotides containing 4-thiothymidine, 5-methyl-2-pyrimidinone-1-β-D-(2′-deoxyriboside) and 2-thiothymidine. *Nucleic Acids Res. 17*, 4957–4974.
20. Coffin, S. R., and Reich, N. O. (2008) Modulation of *Escherichia coli* DNA Methyltransferase Activity by Biologically Derived GATC-flanking Sequences. *J. Biol. Chem. 283*, 20106–20116.
21. Lakowicz, J. R. (1999) Principles of Fluorescence Spectropscopy, 2nd ed., Kluwer Academic/Plenum Publishers, New York.
22. 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*, 15563–15572.
23. Lesser, D. R., Kurpiewski, M. R., and Jen-Jacobson, L. (1990) The energetic basis of specificity in the Eco RI endonuclease−DNA interaction. *Science 250*, 776–786.
24. Lee, Y., 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*, 4937–4944.
25. Gerasimaite, R., Vilkaitis, G., and Klimasauskas, S. (2009) A directed evolution design of a GCG-specific DNA hemimethylase. *Nucleic Acids Res. 37*, 7332–7341.
26. Cohen, H. M., Tawfik, D. S., and Griffiths, A. D. (2002) Promiscuous methylation of non-canonical DNA sites by HaeIII methyltransferase. *Nucleic Acids Res. 30*, 3880–3885.
27. 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. 275*, 38722–38730.
28. Wu, J. C., and Santi, D. V. (1987) Kinetic and catalytic mechanism of HhaI methyltransferase. *J. Biol. Chem. 262*, 4778–4786.
29. Mo, Y. (2006) Probing the nature of hydrogen bonds in DNA base pairs. *J. Mol. Model. 12*, 665–672.
30. Huang, N., Banavali, N. K., and MacKerell, A. D. (2003) Protein-facilitated base flipping in DNA by cytosine-5-methyltransferase. *Proc. Natl. Acad. Sci. U.S.A. 100*, 68–73.
31. Ashworth, J., Havranek, J. J., Duarte, C. M., Sussman, D., Monnat, R. J., Stoddard, B. L., and Baker, D. (2006) Computational redesign of endonuclease DNA binding and cleavage specificity. *Nature 441*, 656–659.
32. Chahar, S., Elsawy, H., Ragozin, S., and Jeltsch, A. (2010) Changing the DNA recognition specificity of the EcoDam DNA-(adenine-N6)-methyltransferase by directed evolution. *J. Mol. Biol. 395*, 79–88.
33. Zhou, H., Matje, D. M., Reich, N. O., and Dahlquist, F. W. (2011) manuscript in preparation

34. O'Gara, M., Horton, J. R., Roberts, R. J., and Cheng, X. (1998) Structures of HhaI methyltransferase complexed with substrates containing mismatches at the target base. *Nat. Struct. Biol. 5*, 872–877.
35. Bloomfield, V. A., Crothers, D. M., and Tinoco, I. (2000) Nucleic acids: Structures, properties, and functions, University Science Books, Sausalito, CA.
36. Daujotyte, D., Grazulis, S., Magennis, S. W., Dryden, D. T. F., Klimasauskas, 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*, 6953–6960.
37. Horton, J. R., Ratner, G., Banavali, N. K., Huang, N., Choi, Y., Maier, M. A., Marquez, V. E., MacKerell, A. D., 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*, 3877–3886.
38. Parry, D., Moon, S. A., Liu, H., Heslop, P., and Connolly, B. A. (2003) DNA recognition by the EcoRV restriction endonuclease probed using base analogues. *J. Mol. Biol. 331*, 1005–1016.
39. Cal, S., and Connolly, B. A. (1997) DNA Distortion and Base Flipping by the EcoRV DNA Methyltransferase. *J. Biol. Chem. 272*, 490–496.
40. Klimasauskas, S., and Roberts, R. J. (1995) MHhal binds tightly to substrates containing mismatches at the target base. *Nucleic Acids Res. 23*, 1388–1395.
41. Menges, F. (2010) Spekwin32: Free optical spectroscopy software, version 1.71.0, http://www.effemm2.de/spekwin/.