Structure, Vol. 12, 1047–1055, June, 2004, ©2004 Elsevier Ltd. All rights reserved. DOI 10.1016/j.str.2004.04.007

HhaI DNA Methyltransferase Uses the Protruding Gln237 for Active Flipping of Its Target Cytosine

Dalia Daujotytė,[1] Saulius Serva,[1,3]
Giedrius Vilkaitis,[1,4] Eglė Merkienė,[1]
Česlovas Venclovas,[1,2] and Saulius Klimašauskas[1,*]
[1]Laboratory of Biological DNA Modification
Institute of Biotechnology
LT-02241 Vilnius
Lithuania
[2]Biology and Biotechnology Research Program
Lawrence Livermore National Laboratory
Livermore, California 94551

Summary

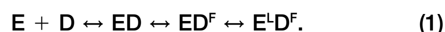
Access to a nucleotide by its rotation out of the DNA helix (base flipping) is used by numerous DNA modification and repair enzymes. Despite extensive studies of the paradigm HhaI methyltransferase, initial events leading to base flipping remained elusive. Here we demonstrate that the replacement of the target C:G pair with the 2-aminopurine:T pair in the DNA or shortening of the side chain of Gln237 in the protein severely perturb base flipping, but retain specific DNA binding. Kinetic analyses and molecular modeling suggest that a steric interaction between the protruding side chain of Gln237 and the target cytosine in B-DNA reduces the energy barrier for flipping by 3 kcal/mol. Subsequent stabilization of an open state by further 4 kcal/mol is achieved through specific hydrogen bonding of the side chain to the orphan guanine. Gln237 thus plays a key role in actively opening the target C:G pair by a "push-and-bind" mechanism.

Introduction

Rotation of a nucleotide out of the DNA helix (base flipping) is a key mechanistic feature employed by DNA modification and repair enzymes (Cheng and Blumenthal, 1996; Cheng and Roberts, 2001), and by proteins responsible for the opening of DNA or RNA during replication, transcription (Lim et al., 2001), and recombination (Gupta et al., 1999). Base flipping was first observed in a reaction complex of the HhaI DNA cytosine-5 methyltransferase (Klimasauskas et al., 1994), and many efforts have been since devoted to elucidating its mechanism in a variety of systems. DNA cytosine-5 methyltransferases (C5-MTases), which catalyze the transfer of methyl groups from the ubiquitous donor *S*-adenosyl-L-methionine (AdoMet) onto cytosine bases, are found in diverse organisms from bacteria to humans (Cheng and Roberts, 2001). In higher organisms, DNA methylation is essential for a number of cellular processes, including transcriptional regulation, genomic imprinting, silencing of retroviruses, and chromatin remodeling. Aberrations in cytosine-5 methylation correlate with human genetic disease, and, therefore, C5-MTases are potent candidate targets for developing new therapies (Robertson and Wolffe, 2000; Tycko, 2000).

A characteristic feature of all C5-MTases is an ordered set of conserved sequence motifs that fold into a bilobal structure. The bacterial MTase HhaI (M.HhaI), which methylates the inner cytosine (underlined) in the sequence G<u>C</u>GC, is the most studied and serves as a structural paradigm for this class of enzymes. A series of X-ray structures for the ternary reaction complexes provided atomic details of the target base flipped out into a catalytic pocket in the enzyme (Klimasauskas et al., 1994; Kumar et al., 1997; O'Gara et al., 1998). The departure of the cytosine from the helix is accompanied by a large conformational motion of the catalytic loop (residues 81–98) in the larger domain of the protein, which locks the flipped-out cytosine and the bound cofactor in the catalytic site. Gln237 from the smaller domain and Ser87 from the catalytic loop approach the DNA helix from the major and minor groove side, respectively, and occupy the original position of the flipped-out cytosine (see Figure 3C). However, these end-point structures provided scarce information concerning initial events on the flipping pathway. It is not clear, whether the infiltration of Gln237 accelerates the opening of the target base pair (active base flipping) or whether it merely occupies the space vacated by the cytosine to stabilize the flipped-out state; nor is it clear whether the base or the sugar is primarily targeted for rotation by the enzyme (Cheng and Roberts, 2001). Both the protruding Gln237 residue (Cheng and Blumenthal, 1996; Mi et al., 1995) and the mobile Ser87 (Huang et al., 2003) have been implicated in the initiation of base flipping; however, experimental evidence of this has not yet been obtained.

The major steps along the reaction pathway (association, base flipping, loop locking) lead to binary intermediates ED (enzyme-DNA), ED^F (flipped out), and E^LD^F (flipped out and locked):



$$E + D \leftrightarrow ED \leftrightarrow ED^F \leftrightarrow E^LD^F. \quad (1)$$

Solution ^{19}F NMR studies found that these species are in rapid exchange in the binary MTase-DNA complex, but the equilibrium is shifted to the right in the presence of cofactor AdoMet or its product AdoHcy (Klimasauskas et al., 1998). Early studies involving mismatched DNA duplexes found that the affinity of binary interaction inversely correlates with the stability of the target base pair, suggesting that binding energy is used to open up the DNA helix (Klimasauskas and Roberts, 1995; Yang et al., 1995). The nature of the target base itself proved much less important for flipping (Klimasauskas and Roberts, 1995; O'Gara et al., 1998), implying that no specific recognition of the target base occurs during initial interaction with the enzyme. To this end, valuable probes for base flipping studies have been DNA duplexes con-

*Correspondence: klimasau@ibt.lt
[3]Present address: Department of Plant Pathology, University of Kentucky, Lexington, Kentucky 40503.
[4]Present address: Institute for Protein Research, Osaka University, Osaka 565-0871, Japan.

Table 1. Structure and Abbreviation of DNA Duplexes

Sequence	Target Base Pair/ Abbreviation
37-mer hemimethylated duplexes	
5′-GACTGGTACAGTATCAG**GCGC**TGACCCACAACATCCG TGACCATGTCATAGTC**CGMG**ACTGGGTGTTGTAGGCT-5′	C:G
5′-GACTGGTACAGTATCAG**GPGC**TGACCCACAACATCCG TGACCATGTCATAGTC**CGMG**ACTGGGTGTTGTAGGCT-5′	P:G
5′-GACTGGTACAGTATCAG**GAGC**TGACCCACAACATCCG TGACCATGTCATAGTC**CGMG**ACTGGGTGTTGTAGGCT-5′	A:G
5′-GACTGGTACAGTATCAG**GPGC**TGACCCACAACATCCG TGACCATGTCATAGTC**CIMG**ACTGGGTGTTGTAGGCT-5′	P:I
5′-GACTGGTACAGTATCAG**GPGC**TGACCCACAACATCCG TGACCATGTCATAGTC**CAMG**ACTGGGTGTTGTAGGCT-5′	P:A
5′-GACTGGTACAGTATCAG**GPGC**TGACCCACAACATCCG TGACCATGTCATAGTC**CCMG**ACTGGGTGTTGTAGGCT-5′	P:C
5′-GACTGGTACAGTATCAG**GPGC**TGACCCACAACATCCG TGACCATGTCATAGTC**CTMG**ACTGGGTGTTGTAGGCT-5′	P:T
24-mer unmethylated duplex	
5′-TAATAAT**GCGC**TAATAATAATAAT ATTATTA**CGCG**ATTATTATTATTA-5′	GCGC

HhaI recognition site is boldface, nucleotide at the target position is underlined; M, 5-methylcytosine; P, 2-aminopurine.

taining a fluorescent base analog, 2-aminopurine (P), at the target position. Interaction of M.HhaI with such DNA leads to a dramatic enhancement of the fluorescence intensity (Holz et al., 1998), permitting direct observation of base flipping in equilibrium titrations or kinetic experiments (Vilkaitis et al., 2000).

In the present study, we used similar DNA duplexes that contained both the fluorescent flipping marker on the target strand and structural variations at the "orphan" guanine position on the methylated strand. These structural alterations were correlated with the DNA binding affinity and the extent of base flipping by M.HhaI, which highlighted the importance of steric interactions in the major groove of DNA. Mutations in the catalytic loop and the Gln237 residue in the enzyme combined with kinetic analyses and molecular modeling indicate that the protruding side chain of Gln237 plays a key active role in expelling the target cytosine out of the DNA helix and suggest that such straining of the DNA substrate may be a general mechanism used by this class of enzymes.

Results

Base Pair Stability and Base Flipping

Previous studies found that the affinity of DNA binding by M.HhaI inversely correlates with the stability of the target base pair (Klimasauskas and Roberts, 1995). In parallel with such binding studies, we have now measured the extent of base flipping using fluorescence spectroscopy (Holz et al., 1998). For this, we prepared similar 37-mer duplexes (Table 1) in which the target cytosine was replaced with 2-aminopurine (P). On the opposite strand, the complementary orphan base was varied (G, I, A, C, or T) to obtain different base pair stabilities. Based on published data, the base pairing strength of P diminishes in the order T>C>A>G (Guest et al., 1991; Law et al., 1996). The stability of the P:I pair has not been reported, but in many cases the pairing potential of hypoxanthine is intermediate between those of G and A (Martin et al., 1985; Oda et al., 1991). As expected, our titration of duplexes containing purines as the orphan base showed a linear increase in fluorescence followed by a plateau (Figure 1A). Surprisingly, the maximal fluorescence amplitudes at saturation turned out to be substantially different for each DNA substrate (60%–80% for P:I and 8%–15% for P:A, as compared with the P:G duplex), although their excitation and emission spectra appeared virtually identical (Figure 1B). Moreover, no change in fluorescence was observed upon addition of the enzyme to both substrates containing pyrimidines (C or T) as the orphan base (Figure 1B). Our data thus clearly indicate that base flipping in the HhaI MTase can be perturbed or even abolished by alterations in the orphan base and the stability of its pairing with the target base (Figure 2).

A further prediction from this conclusion would be that the binding affinities should vary in accordance with the extent of base flipping since the flipping equilibrium would affect the overall binding (Equation 1) if other parameters remained the same. Gel mobility shift analysis of the binary complexes showed that the binding affinities decrease in the order G>I∼T>A>C (Figure 2). Overall, our analysis thus confirms the expected inverse correlation between the strength of target base pair in the DNA substrates and the affinity of DNA binding and the extent of base flipping by M.HhaI. However, the P:T duplex proved a clear outlier, showing no flipping but a relatively strong binding ($K_D = 0.5$ nM). To obtain clues for this anomalous behavior we have looked at the geometrical features of the base pairs involving P. It turns out that the P:T base pair is the only one with a correct Watson-Crick geometry. The P:C forms a wobble pair (Fagan et al., 1996) and the rest (P:A, P:G, and P:I) are bulkier purine:purine base pairs with nonstandard geometry (Fazakerley et al., 1987). Several other base pairs with Watson-Crick geometry including A:T, T:P, and T:A were also tested; however, all of them exhibited poor interaction with the WT M.HhaI (see below).

Another peculiarity of the P:T pair is that the target

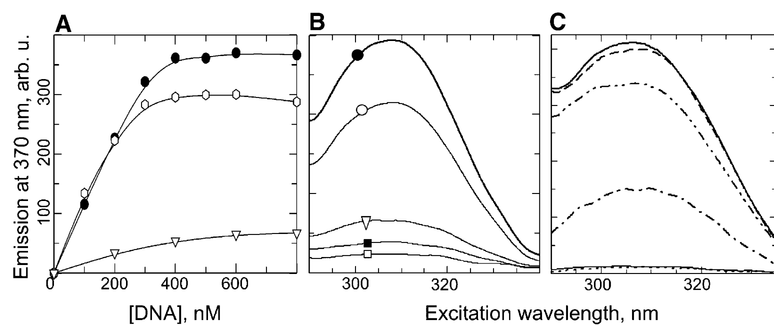


Figure 1. Fluorescence Analysis of Base Flipping by M.HhaI upon Interaction with 37-mer DNA Duplexes Containing 2-Aminopurine at the Target Base Position

(A) Corrected fluorescence intensity curves obtained upon titration of 400 nM WT MTase with increasing amounts of the P:G (closed circles), P:I (open circles), and P:A (open triangles) DNA duplex.

(B) Corrected excitation spectra of binary complexes involving M.HhaI (W41F) and the P:G (closed circles), P:I (open circles), P:A (open triangles), P:C (closed squares), or P:T (open squares) DNA duplex.

(C) Corrected excitation spectra of binary complexes involving the P:G duplex and the WT (solid line), Q237L (broken-dotted), Q237A (solid gray), Q237G (dotted), Δ(82–97) (broken-double dotted), and Δ(81–99)GG (broken) variants of M.HhaI.

base carries no exocyclic group on the major groove side, as opposed to the four canonical DNA base pairs. To explore if this concave geometry of the base may be important for interactions with the protein, we modeled B-DNA helices containing nonflipped target base pairs into the DNA binding site of M.HhaI (see Experimental Procedures). We assumed that, in such "initial" M.HhaI-DNA complex, a DNA helix would be bound in a similar manner (except for interactions with the flipped-out nucleotide and adjacent phosphates) as in the ternary M.HhaI-DNA-AdoHcy complex (c.a. 6MHT [Kumar et al., 1997]). The P:T pair (Figure 3A) or other canonical pairs (data not shown) were modeled by replacing the nonflipped target C:G pair in the modeled complex (Figure 3B), but preserving the same orientation of the N-glycosyl bonds. Notably, the overall conformation (C_α chain) of the target recognition domain (residues 194–275) in all crystal structures of M.HhaI remains essentially the same (the average rms deviation of the C_α atoms is less than 0.5 Å and the B factor values are relatively low). We therefore reasoned that the rigid protein backbone should retain its fold in the modeled "initial" complex. Inspection of the model revealed that

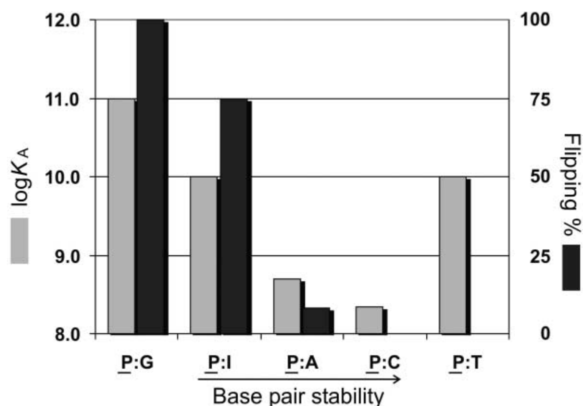


Figure 2. DNA Binding Affinity and 2-Aminopurine Base Flipping by WT M.HhaI

Binding affinities were determined by gel shift titration analysis and are shown as $\log K_A$. The extent of base flipping is shown in percent of the total fluorescence amplitude observed with P:G. Both the DNA binding and base flipping decrease with increasing the target base stability, but the P:T duplex shows an anomalously high binding affinity with no detectable base flipping.

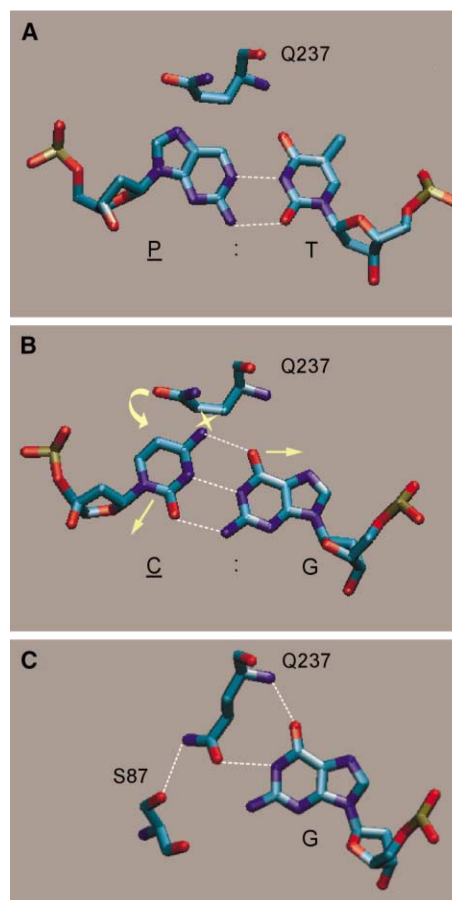


Figure 3. Molecular Models Illustrating Interactions between Gln237 and the Target Base Pair in DNA

(A) The bent conformation of the side chain of Gln237 avoids clashing with the concave P:T base pair, but not with the canonical C:G base pair.

(B) The proposed steric clash (×) expels the target base into a groove leading to the disruption of the base pairing and stacking interactions. Arrows point toward the final positions of the side chain and the bases.

(C) Experimental structures (1MHT–9MHT) show the embrace of the orphan guanine by the side chain of Gln to form two stabilizing H bonds. The catalytic loop closes to form a H bond between the side chains of Ser87 and Gln237.

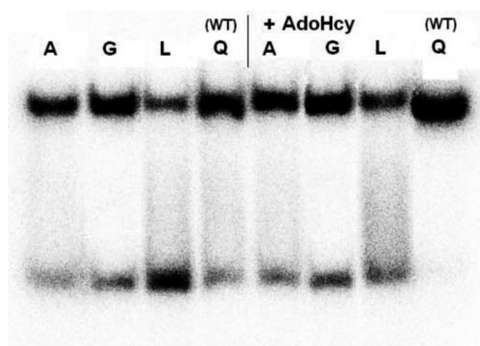


Figure 4. Gel Shift Analysis of DNA Binding by the Gln237 Mutants of HhaI MTase

Samples containing 7 nM 37-mer C:G DNA duplex and 10 nM M.HhaI, supplemented with 100 nM AdoHcy, if any, were analyzed as described in Experimental Procedures.

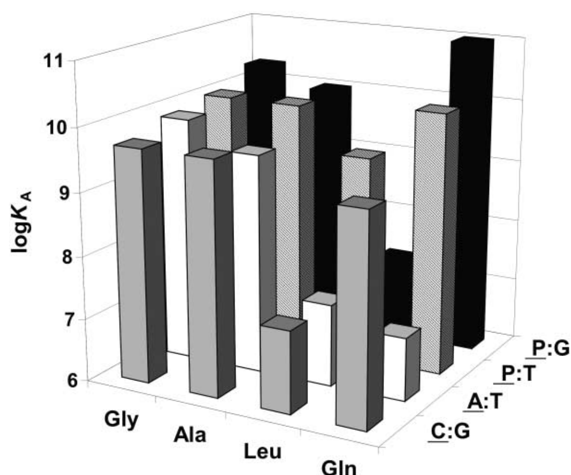


Figure 5. Affinity of Binary Interaction of the Gln237 Variants with DNA

Binding affinities were determined by gel shift titration analysis and are shown as logK_A. All binary complexes involving the Q237G, Q237A mutants and/or the P:T duplex show high binding affinity. The WT M.HhaI binds duplexes containing G at the orphan position much better than the Q237L mutant.

most protein residues interacting with the target GCGC site were in acceptable positions. The only clear problem was Gln237, whose side chain clashed with the base to be flipped out. Remarkably, it turned out that in the case of the P:T target pair, Gln237 can be easily accommodated in the major groove by simply selecting another rotamer in which the side chain steers away from the DNA (Figure 3A). A simple analysis showed that a nonconflicting accommodation of the side chain is possible only for this particular base pair due to its unique concave geometry. In contrast, no conformation of the Gln237 side chain can be found to avoid steric conflicts with the exocyclic groups of the regular base pairs (Figure 3B). Similarly, the wobble pairs formed between P and other bases (A or C) cannot accommodate the side chain without steric clashes. This crude model thus helps to illustrate a possible reason for the observed higher stability of the binary M.HhaI-(P:T) complex in the absence of base flipping.

Role of Gln237 in DNA Binding and Base Flipping

To verify if indeed the target base pair interacts with the side chain of Gln237, we resorted to mutants of M.HhaI with alterations at the implicated residue. Gln237 mutants in which the amide functionality is removed (Leu) and the side chain is shortened (Ala and Gly) were produced by site-directed mutagenesis and purified to homogeneity by established procedures (Vilkaitis et al., 2000). The mutant proteins were examined for their ability to bind DNA and flip the target base employing the same experimental assays as described above for the WT enzyme. The behavior of the mutants was determined using DNA substrates with the following target base pairs: P:T, A:T, C:G, and P:G (summarized in Figure 5). The P:T and A:T are strong Watson-Crick-type base pairs that share the same overall geometry but differ in the location of the exocyclic amino group (minor versus major groove). P:G and C:G are a weak and strong pair, respectively, both containing the orphan guanine base.

Electrophoretic mobility assays (Figure 4) revealed similar mobilities of the complexes, indicating that no gross alterations were introduced by the mutations. Control experiments confirmed that such specific complexes are observed only in the context of the GXGC recognition site. The binary binding affinities determined in titration experiments (data not shown) are largely in the nanomolar range (Figure 5). In a few cases, limitations of the gel shift assay (band smearing at [MTase] $>$50 nM), precluded accurate determination of the K_D values that are larger than 50–100 nM (corresponding K_A values are thus shown as their upper limits in Figure 5). The observed nanomolar binding affinity of the Ala and Gly mutants to the hemimethylated substrates is in discord with a previous report (Mi et al., 1995), in which no specific complexes could be identified using similar unmethylated DNA duplexes. In our hands, the unmethylated duplexes formed MTase-DNA complexes in the presence of AdoHcy that are readily detectable by gel shift analysis (data not shown). A likely reason for this discrepancy is the preparation of the proteins by immunoprecipitation in the previous study, which may interfere with DNA binding.

We find that the Q237G and Q237A mutants bind all types of the DNA duplexes examined with high affinity (Figure 5), but no flipping of P is observed (Figure 1C). In contrast, the Leu mutant shows poor DNA binding, except for the concave P:T substrate, however, shows partial flipping of P upon binary interaction with the P:G DNA (Figure 1C). The extent of base flipping and the affinity of binding in the Q237L-P:G complex is similar to that observed with the WT protein and P:A duplex (Figure 1B). In both cases, no stabilizing interaction from the side chain of residue 237 to the orphan base is possible. A clear illustration of the effect of the presence of the exocyclic amino group on the target base pair comes from direct comparison of the behavior of the MTase variants with respect to the P:T and A:T duplexes (Figure 5). None of the MTase variants can open the rigid and concave P:T base pair, but all of them do bind the substrate quite efficiently. Similar molecular

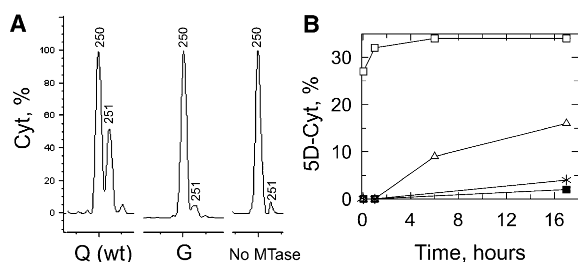


Figure 6. Analysis of Hydrogen-5 Exchange in the Binary M.HhaI-DNA Complexes

(A) Mass spectra of the dCyd-Na^+ ions obtained after treatment of the unlabeled GCGC duplex in 80% D_2O for 1 hr with WT M.HhaI (left), the Q237G mutant (center), and for 16 hr in the absence of M.HhaI (right). Peaks scaled to 100% of the 250 m.u. ion. The untreated control (right) shows the presence of 5% of the 251 ion due to natural occurrence of ^{13}C.

(B) Time dependence of deuterium incorporation into dCyd at 20°C by the WT (open squares), Q237L (open triangles), Q237A (asterisk), and Q237G (closed squares) M.HhaI.

modeling of the Q237L-(P:T) complex (data not shown) indicates that the side chain of Leu can adopt a nonclashing conformation by bending off the DNA. Moving the amino group from the minor groove (P) into the major groove (A) changes the picture in a dramatic manner, as neither the Gln (WT) nor the Leu variant can bind the DNA. As mentioned above, the "short" mutants, lacking the C_γ (Ala) or both C_γ and C_β (Gly) moieties in the side chain, do bind with high affinity. These findings are completely consistent with our hypothesis that the side chain of residue 237 is involved in a steric interaction with the DNA.

Cytosine Flipping during Catalysis

The experiments described above used noncatalytic substrates containing P, which permitted monitoring of base flipping using fluorescence. Although there is no comparable direct assay to detect flipping of cytosine in binary MTase-DNA complexes, we took advantage of a mechanism-based property of M.HhaI to catalyze the exchange of the hydrogen-5 atom on the target cytosine for protons of water in the absence of cofactor (Wu and Santi, 1987). The H5-exchange reaction requires the formation of a transient covalent bond at C6 in the cytosine ring to form a 5,6-dihydrocytosine derivative, which is only possible when the target cytosine is flipped out and bound in the catalytic site of the enzyme (E^LD^F in Equation 1). A newly developed assay was used to monitor the exchange reaction between the unlabeled 24-mer duplex GCGC (see Table 1) and deuterated solvent (80% D_2O). The extent of the H→D conversion was determined by mass-spectrometric analysis of the 2′-deoxycytidine fraction obtained after enzymatic hydrolysis of the modified DNA. As expected, treatment with WT M.HhaI leads to a fast build up of the N+1 peak (Figure 6A, left), reaching the level of 34% of the total dCyd, or 68% for the target cytosine. Our estimated apparent rate of cytosine-C5 deuteration (~1 min^{-1}) in the 24-mer duplex in D_2O at 20°C thus compares reasonably with the previously observed exchange of 5-tritium (9 min^{-1}) in copolymer poly(dG-dC)·poly(dG-dC) for protons of H_2O at 37°C (Wu and Santi, 1987). Remarkably, the time course analysis shows a dramatic reduction in the exchange rate for the mutants (Figure 6B). A ~1000-fold slower exchange is observed for the Leu mutant. The Q237A and Q237G mutants barely incorporate any D above background even after 16 hr incubation (Figure 6), precluding a meaningful kinetic assessment. Since the mutants contain all other catalytic residues and are active in the methylation reaction in the presence of AdoMet ([Mi et al., 1995] and see below), this observation strongly suggests that deletion of the side chain of Gln237 dramatically impairs the flipping of the target cytosine during binary interaction with DNA.

Interactions of Gln237 with the target C:G base pair in the ternary complex were analyzed by examining the behavior of the Gln237 mutants under catalytic conditions. First, we determined their steady-state kinetic parameters using the copolymer poly(dG-dC)·poly(dG-dC) as a substrate (summarized in Table 2). The lack of the terminal amide in the mutants results in considerably higher values of K_M^{AdoMet} (70-fold for Leu, and 300- to 700-fold for the shorter side chains). This effect may derive from a weaker locking capacity of the catalytic loop due to the disrupted H bond interaction between the amide of Gln237 and the Ser87 residue (for Leu) and/or a strongly reduced population of the flipped-out state. The K_M^{DNA} values go up by 1–2 orders of magnitude, but the effects are lower for the shorter side chains, in line with the trend established in the binding experiments (Figure 5). Interestingly, k_{cat} is slightly higher (2- to 3-fold) for the Leu and Ala mutants, but the Gly shows a 40-fold drop in the turnover rate.

Second, single-turnover and presteady state burst kinetic experiments (Vilkaitis et al., 2001) were performed using the C:G duplex as a substrate to determine the rate of the chemical methyl transfer step, k_{chem}, and the rate of enzyme release, k_{off}. Our results indicate that the apparent rate of methyl transfer decreases gradually from 11–2 min^{-1} in the order Gln>Leu>Ala, and further drops to 0.2 min^{-1} (60-fold as compared with WT) in the case of Gly (Table 2). The length of the side chain thus correlates with the apparent rate of methyl transfer, but not to the extent observed in the binary complex. In addition to the reduction of k_{chem} in the Gly mutant, the derived k_{off} value is also significantly lower than in the other variants. The lack of a steric clash with the target base pair in Q237G confers an enhanced stability of the binary MTase-DNA complexes regardless of the nature of the target base (Figure 5) or the methylation status of the target cytosine. The catalytic efficiency of the enzyme, which is proportional to k_{cat}/K_M for a single-substrate enzyme (Fersht, 1999), or $k_{cat}/(K_M^{AdoMe}\cdot K_M^{DNA})$ in this case, diminishes continuously by 3–5 orders of magnitude as the side chain is shortened (Table 2). In aggregate, our kinetic analyses indicate that the Q237 mutations do not dramatically affect the rate of the catalytic step; rather, perturbed base flipping is partially compensated for by other interactions, as manifested by higher K_M values for both substrates.

Role of the Catalytic Loop

To assess a possible role of the catalytic loop in the base flipping process, we constructed two mutants,

Table 2. Kinetic Parameters of the Gln237 Mutants of HhaI Methyltransferase

Residue 237	K_M^{AdoMet}, μM	K_M^{DNA}, nM	k_{cat}, min^{-1}	k_{chem}, min^{-1}	k_{off}, min^{-1}	$k_{cat}/(K_M^{AdoMet} \cdot K_M^{DNA})$ WT/Mutant
Gly	7.5 ± 1.0	3.1 ± 0.3	0.03 ± 0.001	0.2 ± 0.02	0.072 ± 0.012	68,000
Ala	17 ± 2.6	7.1 ± 1.6	2.1 ± 0.13	2.4 ± 0.3		5,000
Leu	1.8 ± 0.28	43 ± 7.2	3.0 ± 0.1	10 ± 2.3	4.2 ± 1.1	2,300
Gln (WT)	0.025 ± 0.002	0.5 ± 0.07	1.1 ± 0.02	11 ± 2.1	3 ± 0.8	1

Δ(81–99)GG and Δ(82–97), bearing large deletions in the catalytic loop. The Δ(82–97) mutant retains the catalytic Pro-Cys motif (residues 80–81), but both mutants lack many other conserved residues including the implicated Ser87 (Huang et al., 2003). The proteins were purified to homogeneity and appeared comparably soluble as the WT protein. As expected, they showed no catalytic activity in vivo or in vitro (data not shown). Remarkably, the proteins retained the ability to form specific protein-DNA complexes with high affinity (data not shown). Fluorescence equilibrium titrations with DNA duplexes containing P at the target base position showed a strong increase in fluorescence, as in the case of the WT enzyme (Figure 1C), indicating that the mutants were proficient in flipping the P base. The similarity in fluorescence intensities as well as the excitation and emission spectra suggests that the target base (P) flips out into a similar environment (solvent) in both mutants and the WT protein. These experiments confirm a key role for the loop in the catalytic methylation of cytosine, and clearly demonstrate no such role with respect to target base flipping, at least in the case of the mismatched DNA substrate.

Discussion

The present study for the first time reveals a direct link between the stability of the target base pair on one hand and the extent of base flipping and the affinity of binary interaction with DNA on the other. Combined with mutational analysis, this illuminates a molecular mechanism by which binding and specific recognition of the target site by the enzyme is intimately connected with the opening of the target base pair. Upon formation of the specific contacts between the three recognition base pairs and the protein, the protruding side chain of Gln237 comes in contact with the target base in the major groove of the DNA (Figure 3B). The rigid fold of the target recognition domain observed in numerous X-ray structures (Klimasauskas et al., 1994; O'Gara et al., 1999) implies that the deformation is absorbed by the bound DNA, and the target base is dislodged from its original position to accommodate the space requirements of Gln237. The significance of this steric interaction is supported by the facts that no specific recognition of the target base appears necessary for flipping (Klimasauskas and Roberts, 1995; O'Gara et al., 1998), whereas the removal of either the side chain of Gln237 or the exocyclic group of the target base (in P:T) leads to the formation of stable nonflipped binary complexes (Figures 2 and 5). Consistent with our model, previous ^{19}F NMR studies (Klimasauskas et al., 1998) revealed a dynamic equilibrium of multiple flipped-out states in the binary M.HhaI-DNA complex, while the existence of a species with an intrahelical target base (equivalent of ED) is only detectable under conditions of a higher ionic strength where the overall equilibrium is greatly shifted to the left (Equation 1). On the other hand, our observation that the concave geometry of 2-aminopurine may perturb or even abolish an important interaction with a protein is of special importance for the utility of this analog as a reporter in base-flipping studies (Allan et al., 1998; Gupta et al., 1999; Holz et al., 1998; Stivers and Jiang, 2003; Vilkaitis et al., 2000). An obvious example of the failure to open the P:T pair has been recently observed during promoter-RNA polymerase interaction (Lim et al., 2001). Yet another candidate system may be the EcoRV adenine-N6 MTase, in which no fluorescence enhancement was observed in complexes when the target adenine was replaced with P (Gowher and Jeltsch, 2000).

In energetic terms, the bulk from the protruding residue 237 serves to destabilize the initial specific binary MTase-DNA complex (ED in Figure 7), which in turn lowers the energy barrier for the subsequent base-flipping step. This mechanism thus exemplifies the concept of substrate "strain" whereby the binding energy is utilized for maximizing the rate of enzymatic catalysis (Fersht, 1999). The ratio of the binary binding affinities of the Gly/Leu mutants (>100-fold) gives an estimated free energy difference ($\Delta\Delta G = -RT \ln(K_A^{Leu}/K_A^{Gly})$) of at

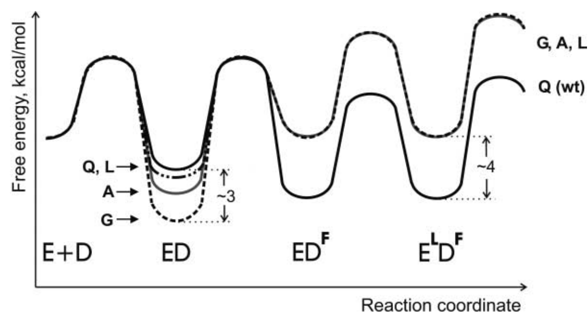


Figure 7. Energetic Effects of Gln237 Mutations in HhaI MTase
Free energy profiles of the WT (solid line), Q237L (broken-dotted line), Q237A (solid gray), and Q237G (dotted) variants are shown along the major binary intermediates ED (enzyme-DNA), ED^F (flipped out) and E^LD^F (flipped out and locked) (see Equation 1). Cofactor interactions are not shown for clarity. The bulk from residue 237 (in the Gln [WT] and Leu variants) destabilizes ED due a steric clash with the target base pair (Figure 3B). Unlike the Gln237 mutants, the WT enzyme can stabilize the flipped-out state ED^F by providing two H bonds from the side chain to the orphan guanine base; E^LD^F may be further stabilized by a H bond to Ser87 (see Figure 3C). Following methyl transfer, the complex is disassembled in reverse order.

least 3 kcal/mol for the ED intermediate that comes from the aliphatic atoms of residue 237, assuming that the energies of the other species at equilibrium (ED^F and E^LD^F) are approximately the same. This value corresponds to the activation energy reduction at the base-flipping step (Figure 7), implying a direct active role of the enzyme in accelerating the flipping process. As proposed previously (Klimasauskas et al., 1994; Mi et al., 1995), another important role of Gln237 is the stabilization of the flipped-out intermediates ED^F and E^LD^F via the H bond between the side chain carbonyl and the N1 of the orphan guanine (Figure 3C). In addition, the side chain amino group seems to form a H bond to Ser87 (distance of ~3 Å) in E^LD^F. All of the Gnl237 mutants lack these interactions and therefore are not able to make stable flipped-out complexes (Figure 5), nor are such complexes possible between the WT enzyme and duplexes containing other bases at the orphan base position (Figure 2). The ratio of the binding affinities of the bulky variants (Leu/WT >100-fold, Figure 5) permits an assessment of the (lower limit of) free energy that is gained by these specific H bond interactions (≥3 kcal/mol). Taking into consideration that the 5-hydrogen reaction is ~1000-fold slower in the Leu mutant as compared to the WT enzyme, one can conclude that the free energy difference is around 4 kcal/mol (Figure 7). In total, the side chain of Gln237 appears to provide 7 kcal/mol on the flipping pathway: ≥3 kcal/mol by destabilizing the closed base pair ("push") and another 4 kcal/mol by stabilizing the open base pair ("bind"). This value is internally consistent with the ~10^5-fold drop (corresponds to 6.8 kcal/mol in free energy) in the catalytic efficiency of the Gly mutant (Table 2) as derived from kinetic analysis of the ternary complex. Therefore, the side chain of Gln237 plays a key role by providing nearly a half of the estimated 15–20 kcal/mol required to completely flip out the target cytosine (Banavali and MacKerell, 2002; Varnai and Lavery, 2002).

Additional contributions favoring the flipped-out state may come from yet unidentified phosphate backbone interactions (Cheng and Blumenthal, 1996; Klimasauskas et al., 1994). This might explain the flipped-out conformation of an abasic target nucleotide observed in the crystal structure of the ternary complex involving modified DNA (O'Gara et al., 1998). Since the energetic cost is much reduced for such a nucleotide (no stacking and base pairing interactions), its rotation should readily occur without engaging the base-pushing mechanism (see also below). In the ternary complex with AdoHcy or AdoMet, binding of the flipped-out cytosine in the catalytic site and locking of the catalytic loop (catching) provides a further ~1000-fold stabilization of the MTase-DNA complex (Lindstrom et al., 2000; Vilkaitis et al., 2001), thereby extending the lifetime of the flipped-out conformer into the minute/hour time scale (Kumar et al., 1997; Vilkaitis et al., 2001). It is not surprising that the Q237 mutants still show a considerable catalytic activity despite a highly impaired base flipping, since the flipping deficiency is partially compensated for by other interactions, such as binding of substrates (increase in K_M). However, the catalytic efficiency of the enzyme, which is proportional to $k_{cat}/(K_M^{AdoMe} \cdot K_M^{DNA})$, goes down by 3–5 orders of magnitude as the side chain is shortened (Table 2). Interestingly, the removal of the terminal amido group (in Leu) has no effect on the apparent rate of the methyl transfer step k_{chem} (Table 2), whereas further shortening of the side chain slows down the methylation by a factor of 5 (Ala) and 60 (Gly). In other words, the stabilization of the flipped-out state thus appears less important than the reduction of the activation barrier for flipping. This clearly highlights a key role of the protruding Q237 side chain in maintaining proper populations of the flipped-out intermediates during catalysis.

Our model implies that any of the four normal base pairs within the context of the GNGC tetranucleotide would be tried by pushing upon binding to the enzyme. Specific stabilization and further opening of the pre-flipped state will take place only in the case of the correctly oriented C:G pair, while the other cases will most often result in reversal to the normal state and dissociation of enzyme. It should be noted that our molecular models of the initial complexes just serve to illustrate the existence of a steric clash rather than describe atomic details of this interaction. Therefore, the actual conformation and interactions of the side chain that occur upon docking of the DNA may be different from those depicted in Figure 3B. Based on the general topological argument, one may expect that the target base exits via the minor groove, although other pathways cannot be excluded by the present data. Recently, a "pincer pinch" model has been suggested based on molecular dynamics calculations, in which specific H bond interactions from the side chains of Ser87 and Gln237 open up the target C:G base pair via the major groove inside the closed protein (Huang et al., 2003). Without specifically addressing the role of Ser87, the presented findings (Figure 1C) indicate that flipping of the P base occurs even if the whole catalytic loop is removed. Moreover, phylogenetic data indicate a weak conservation of Ser at this position (usual occurrences Ala, Gly, and Ile) in the otherwise well conserved catalytic loop (Klimasauskas et al., 1994).

Inspection of other available cocrystal structures suggests that "pushing" in the major groove of DNA may be not unique to M.HhaI. The HaeIII cytosine-C5 MTase, which flips and methylates the inner cytosine in the target site GGCC (Reinisch et al., 1995), contains a protruding residue, Ile221, whose side chain juts into the major groove of DNA. Similar modeling of the initial M.HaeIII-DNA complex (data not shown) suggests a potential steric conflict with the partner guanine base in the target C:G pair. In this case, the stabilization of a flipped intermediate is achieved by shifting the H-bonding partners among the remaining bases. Due to the high overall similarity of C5-MTases (Cheng and Roberts, 2001), it seems likely that the base pushing mechanism may be common for many enzymes of this class, including the physiologically important eukaryotic homologs. To this end, a number of DNA repair enzymes (UDG, MUG, and AlkA) (Barrett et al., 1998; Lau et al., 1998; Stivers and Jiang, 2003) have also been suggested to employ "pushing" as an initial step for the base pair disruption.

Conclusion

Upon binding to its cognate recognition site, the HhaI MTase actively opens the target C:G base pair em-

ploying the protruding Gln237 residue from the target recognition domain. The proposed "push-and-bind" mechanism involves (1) a steric clash between the side chain of Gln237 and the exocyclic group of the target cytosine in the major groove of DNA (Figure 3B) pushes the cytosine toward a groove leading to (partial) disruption of the base pairing and stacking interactions; and (2) the side chain of Gln237 occupies the space left by the target base and obstructs its return by binding to the orphan guanine base via specific stabilizing H bonds. Subsequently, the target nucleotide is further rotated out of the DNA helix and into the catalytic site. The catalytic loop closes to lock the cytosine and extends a trans-helical H bond from the side chain hydroxyl of Ser87 to the side chain amino group of Gln237 (Figure 3C).

Experimental Procedures

DNA duplexes for fluorescence and binding experiments were produced by annealing oligonucleotide strands (HSPF grade, MWG, Germany), as listed in Table 1.

Mutagenesis and Protein Purification

Mutagenesis of Gln237 was performed as previously described (Daujotyte et al., 2003) using the following degenerate 26-mer primers: ATTCGTTCACCGSMCCCACCTTTTCC and AATTCGTTCACC TRKCCCACCTTTTCC (the mutated codons underlined). Deletion mutations in the catalytic loop were introduced as described (Vilkaitis et al., 2000) using the following oligonucleotide primers: CAAAAAAGAGGGTACCGCACGGAAACCCTGC for the Δ(82–97) variant (the Cys81 and Gly98 codons underlined) and GCAATAT CAAAAAAGAG*GCCTCC*AAACCCTGCACATAAAATG for the Δ(81–99)GG variant (the Phe79 and Leu100 codons underlined and the Gly-Gly linker italicized). M.HhaI variants were expressed in *E. coli* and purified as described previously (Klimasauskas et al., 1998). Protein concentrations were determined photometrically (ϵ_{280} = 25,610 M^{-1} cm^{-1}). The molecular size of the mutant proteins was verified by ESI mass spectrometry.

Fluorescence Spectroscopy

Steady-state fluorescence data were collected on an SLM Aminco AB2 or Perkin Elmer LS-50B Luminescence Spectrometer at 25°C at excitation and emission wavelengths of 320 and 370 nm, and bandwidths of 2.5 and 10 nm, respectively, as described previously (Vilkaitis et al., 2000). Titrations were performed by addition of 4 μM DNA and 400 nM M.HhaI into 500 μl of 400 nM M.HhaI solution. Titration curves were corrected by subtracting fluorescence intensity values from control titrations, in which thermally inactivated M.HhaI was used. Excitation spectra of the binary complexes were recorded with 300 nM MTase and 250 nM DNA duplex at 370 nm with excitation and emission bandwidths of 2.5 and 5 nm, respectively, and corrected as described (Vilkaitis et al., 2000). W41F mutant, which is very similar to the WT enzyme with respect to base flipping (E.M. and S.K., unpublished data), was used in certain experiments to improve the spectral resolution at excitation wavelengths below 300 nm.

Electrophoretic Gel Mobility Shift Assay

Analyses were performed at room temperature as described (Vilkaitis et al., 2001). Typical reactions contained 7 nM 5′-^{32}P-labeled 37-mer duplex, 10 nM protein, and 100 nM AdoHcy, as required. For K_D analysis, 5′-^{32}P-labeled 37-mer duplex oligonucleotides (0.01 or 0.1 nM) were titrated with increasing protein concentrations (0.01–100 nM) and processed as described (Vilkaitis et al., 2001).

Kinetic Analysis

Steady-state methylation reactions were performed in the presence of 0.1–10 μM [*methyl*-^{3}H]AdoMet (15 or 85 Ci/mmol) and poly(dG-dC)·poly(dG-dC) at double strand recognition site concentrations from 0.1 nM to 10 μM as described previously (Vilkaitis et al., 2000). K_M^{AdoMet} measurements were carried out with 1 μM DNA for the mutant proteins and 50 nM DNA for the WT MTase and varying the [*methyl*-^{3}H]AdoMet concentrations from 2.5 nM to 100 μM (0.5 or 85 Ci/mmol). After incubation, the reactions were quenched with 0.5 N HCl and analyzed as described previously (Vilkaitis et al., 2000). Transient kinetic analyses were carried out in a Rapid-Quench-Flow instrument RQF-3 (KinTek) as described previously (Vilkaitis et al., 2001). Single-turnover reactions contained 100–200 nM hemimethylated 37-mer duplex C:G preincubated with 200–400 nM MTase and 1.3–100 μM [*methyl*-^{3}H]AdoMet (0.5 or 15 Ci/mmol).

Deuterium Exchange Assay

Reactions containing 12 μM of both MTase and unlabeled 24-mer duplex GCGC (Table 1) were incubated in the reaction buffer containing 80% D_2O for specified periods of time at 22°C. Samples were then treated with Nuclease P1 and alkaline phosphatase (Fermentas) and analyzed on an integrated HPLC/ESI-MS ion trap Hewlett-Packard 1100 series system. A sample of 200–300 pmol of total nucleosides was injected into a Discovery C18 column (150 × 2.1 mm; Supelco) and was eluted with 20 mM ammonium formate (pH 3.5) for 5 min followed by a linear gradient to 55% methanol over 15 min at a flow rate of 0.3 ml/min at 45°C. The dCyd peak (elution time 3.4 ± 0.1 min) was analyzed in a mass detector operating at a capillary voltage of 5000 V. Spectra were collected in the positive ion mode and ion trap mass range of 108–600 M/Z to monitor the presence of Cyt-Na^+ (134 m.u.) and dCyd-Na^+ (250 m.u.) ions.

Computational Modeling

Modeling and analysis of M.HhaI interaction with DNA was done using InsightII (Accelrys). The best resolved structure for a ternary M.HhaI-DNA-AdoHcy complex (PDB code 6MHT [Kumar et al., 1997]) was used as a framework for representing M.HhaI-DNA interaction. The initial (nonflipped) position of the target base pair in M.HhaI-DNA complex was modeled using the structure of B-DNA helix containing the GCGC recognition site (PDB code 1CGC [Heinemann et al., 1992]). The B-DNA was fitted onto DNA bound within the M.HhaI ternary complex by the least-squares procedure using base pairs of the M.HhaI recognition sequence except the one with the flipped-out cytosine. Interaction of M.HhaI with target bases other than the "native" cytosine was explored by replacing the C:G pair with the base pairs of interest.

Acknowledgments

The authors are indebted to Laura Manelytė for her help with kinetic analyses and to Gražvydas Lukinavičius for his expert assistance with mass spectrometry. We are also grateful to Richard J. Roberts and Peter Varnai for comments on the manuscript, Elmar Weinhold for numerous discussions, and MBI Fermentas for continued support. This work was supported by grants form the Howard Hughes Medical Institute, Volkswagen-Stiftung, and the Lithuanian State Science and Studies Foundation.

Received: December 19, 2003
Revised: March 25, 2004
Accepted: April 13, 2004
Published: June 8, 2004

References

Allan, B.W., Beechem, J.M., Lindstrom, W.M., and Reich, N.O. (1998). Direct real time observation of base flipping by the *Eco*RI DNA methyltransferase. J. Biol. Chem. *273*, 2368–2373.

Banavali, N.K., and MacKerell, A.D., Jr. (2002). Free energy and structural pathways of base flipping in a DNA GCGC containing sequence. J. Mol. Biol. *319*, 141–160.

Barrett, T.E., Savva, R., Panayotou, G., Barlow, T., Brown, T., Jiricny, J., and Pearl, L.H. (1998). Crystal structure of a G:T/U mismatch-specific DNA glycosylase: mismatch recognition by complementary-strand interactions. Cell *92*, 117–129.

Cheng, X., and Blumenthal, R.M. (1996). Finding a basis for flipping bases. Structure *4*, 639–645.

Cheng, X., and Roberts, R.J. (2001). AdoMet-dependent methylation, DNA methyltransferases and base flipping. Nucleic Acids Res. *29*, 3784–3795.

Daujotyte, D., Vilkaitis, G., Manelyte, L., Skalicky, J., Szyperski, T., and Klimasauskas, S. (2003). Solubility engineering of the *Hha*I methyltransferase. Protein Eng. *16*, 295–301.

Fagan, P.A., Fabrega, C., Eritja, R., Goodman, M.F., and Wemmer, D.E. (1996). NMR study of the conformation of the 2-aminopurine: cytosine mismatch in DNA. Biochemistry *35*, 4026–4033.

Fazakerley, G.V., Sowers, L.C., Eritja, R., Kaplan, B.E., and Goodman, M.F. (1987). NMR studies of an oligonucleotide containing 2-aminopurine opposite adenine. Biochemistry *26*, 5641–5646.

Fersht, A. (1999). Enzyme-substrate complementarity and the use of binding energy in catalysis. In Structure and Mechanism in Protein Science (New York: W.H. Freeman and Co.), pp. 349–376.

Gowher, H., and Jeltsch, A. (2000). Molecular enzymology of the EcoRV DNA-(Adenine-N6)-methyltransferase: kinetics of DNA binding and bending, kinetic mechanism and linear diffusion of the enzyme on DNA. J. Mol. Biol. *303*, 93–110.

Guest, C.R., Hochstrasser, R.A., Sowers, L.C., and Millar, D.P. (1991). Dynamics of mismatched base pairs in DNA. Biochemistry *30*, 3271–3279.

Gupta, R.C., Folta-Stogniew, E., O'Malley, S., Takahashi, M., and Radding, C.M. (1999). Rapid exchange of A:T base pairs is essential for recognition of DNA homology by human Rad51 recombination protein. Mol. Cell *4*, 705–714.

Heinemann, U., Alings, C., and Bansal, M. (1992). Double helix conformation, groove dimensions and ligand binding potential of a G/C stretch in B-DNA. EMBO J. *11*, 1931–1939.

Holz, B., Klimasauskas, S., Serva, S., and Weinhold, E. (1998). 2-aminopurine as a fluorescence probe for DNA base flipping by methyltransferases. Nucleic Acids Res. *26*, 1076–1083.

Huang, N., Banavali, N.K., and MacKerell, A.D. (2003). Protein-facilitated base flipping in DNA by cytosine-5-methyltransferase. Proc. Natl. Acad. Sci. USA *100*, 68–73.

Klimasauskas, S., and Roberts, R.J. (1995). M.HhaI binds tightly to substrates containing mismatches at the target base. Nucleic Acids Res. *23*, 1388–1395.

Klimasauskas, S., Kumar, S., Roberts, R.J., and Cheng, X. (1994). HhaI methyltransferase flips its target base out of the DNA helix. Cell *76*, 357–369.

Klimasauskas, S., Szyperski, T., Serva, S., and Wüthrich, K. (1998). Dynamic modes of the flipped-out cytosine during *Hha*I methyltransferase-DNA interactions in solution. EMBO J. *17*, 317–324.

Kumar, S., Horton, J.R., Jones, G.D., Walker, R.T., Roberts, R.J., and Cheng, X. (1997). DNA containing 4'-thio-2′-deoxycytidine inhibits methylation by HhaI methyltransferase. Nucleic Acids Res. *25*, 2773–2783.

Lau, A.Y., Schaerer, O.D., Samson, L., Verdine, G.L., and Ellenberger, T. (1998). Crystal structure of a human alkylbase-DNA repair enzymes complexed to DNA: mechanisms for nucleotide flipping and base excision. Cell *95*, 249–258.

Law, S.M., Eritja, R., Goodman, M.F., and Breslauer, K.J. (1996). Spectroscopic and calorimetric characterizations of DNA duplexes containing 2-aminopurine. Biochemistry *35*, 12329–12337.

Lim, H.M., Lee, H.J., Roy, S., and Adhya, S. (2001). A "master" in base unpairing during isomerisation of a promoter upon RNA polymerase binding. Proc. Natl. Acad. Sci. USA *98*, 14849–14852.

Lindstrom, W.M., Jr., Flynn, J., and Reich, N.O. (2000). Reconciling structure and function in HhaI DNA cytosine-C-5 methyltransferase. J. Biol. Chem. *275*, 4912–4919.

Martin, F.H., Castro, M.M., Aboul-Ela, F., and Tinoco, I. (1985). Base pairing involving deoxyinosine: implications for probe design. Nucleic Acids Res. *13*, 8927–8938.

Mi, S., Alonso, D., and Roberts, R.J. (1995). Functional analysis of Gln-237 mutants of HhaI methyltransferase. Nucleic Acids Res. *23*, 620–627.

Oda, Y., Uesugi, S., Ikehara, M., Kawase, Y., and Ohtsuka, E. (1991). NMR studies for identification of dI:dG mismatch base-pairing structure in DNA. Nucleic Acids Res. *19*, 5263–5267.

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.

O'Gara, M., Zhang, X., Roberts, R.J., and Cheng, X. (1999). Structure of a binary complex of HhaI methyltransferase with S-adenosyl-L-methionine formed in the presence of a short non-specific DNA oligonucleotide. J. Mol. Biol. *287*, 201–209.

Reinisch, K.M., Chen, L., Verdine, G.L., and Lipscomb, W.N. (1995). The crystal structure of HaeIII methyltransferase covalently complexed to DNA: an extrahelical cytosine and rearranged base pairing. Cell *82*, 143–153.

Robertson, K.D., and Wolffe, A.P. (2000). DNA methylation in health and disease. Nat. Rev. Genet. *1*, 11–19.

Stivers, J.T., and Jiang, Y.L. (2003). A mechanistic perspective on the chemistry of DNA repair glycosylases. Chem. Rev. *103*, 2729–2759.

Tycko, B. (2000). Epigenetic gene silencing in cancer. J. Clin. Invest. *105*, 401–407.

Varnai, P., and Lavery, R. (2002). Base flipping in DNA: pathways and energetics studied with molecular dynamic simulations. J. Am. Chem. Soc. *124*, 7272–7273.

Vilkaitis, G., Dong, A., Weinhold, E., Cheng, X., and Klimasauskas, S. (2000). Functional roles of the conserved threonine 250 in the target recognition domain of *Hha*I DNA methyltransferase. J. Biol. Chem. *275*, 38722–38730.

Vilkaitis, G., Merkiene, E., Serva, S., Weinhold, E., and Klimasauskas, S. (2001). The mechanism of DNA cytosine-5 methylation: kinetic and mutational dissection of *Hha*I methyltransferase. J. Biol. Chem. *276*, 20924–20934.

Wu, J.C., and Santi, D.V. (1987). Kinetic and catalytic mechanism of HhaI methyltransferase. J. Biol. Chem. *262*, 4778–4786.

Yang, A.S., Shen, J.-C., Zingg, J.-M., Mi, S., and Jones, P.A. (1995). HhaI and HpaII DNA methyltransferases bind DNA mismatches, methylate uracil and block DNA repair. Nucleic Acids Res. *23*, 1380–1387.