The EMBO Journal Vol.17 No.1 pp.317–324, 1998

Dynamic modes of the flipped-out cytosine during *Hha*I methyltransferase–DNA interactions in solution

Saulius Klimašauskas, Thomas Szyperski[1,2], Saulius Serva and Kurt Wüthrich[1]

Institute of Biotechnology, LT-2028 Vilnius, Lithuania and [1]Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule Hönggerberg, CH-8093 Zürich, Switzerland

[2]Corresponding author
e-mail: thom@mol.biol.ethz.ch

Flipping of a nucleotide out of a B-DNA helix into the active site of an enzyme has been observed for the *Hha*I and *Hae*III cytosine-5 methyltransferases (M.*Hha*I and M.*Hae*III) and for numerous DNA repair enzymes. Here we studied the base flipping motions in the binary M.*Hha*I–DNA and the ternary M.*Hha*I–DNA–cofactor systems in solution. Two 5-fluorocytosines were introduced into the DNA in the places of the target cytosine and, as an internal control, a cytosine positioned two nucleotides upstream of the recognition sequence 5′-GCGC-3′. The ^{19}F NMR spectra combined with gel mobility data show that interaction with the enzyme induces partition of the target base among three states, i.e. stacked in the B-DNA, an ensemble of flipped-out forms and the flipped-out form locked in the enzyme active site. Addition of the cofactor analogue *S*-adenosyl-L-homocysteine greatly enhances the trapping of the target cytosine in the catalytic site. Distinct dynamic modes of the target cytosine have thus been identified along the reaction pathway, which includes novel base-flipping intermediates that were not observed in previous X-ray structures. The new data indicate that flipping of the target base out of the DNA helix is not dependent on binding of the cytosine in the catalytic pocket of M.*Hha*I, and suggest an active role of the enzyme in the opening of the DNA duplex.

Keywords: cytosine-5 methyltransferase/DNA base flipping/DNA methylation/^{19}F NMR/5-fluorocytosine

Introduction

Methylation of cytosine by DNA methyltransferases (MTases) is used by nature to expand the information content of the genome in organisms ranging from bacteria to mammals. Besides the ensuing specific interest in their physiological functions (Jost and Saluz, 1993; Bird, 1996), the cytosine-5 MTases are also attractive models for studies of general structural aspects of protein–DNA recognition (Anderson, 1993; Kumar *et al.*, 1994). For example, the DNA cytosine-5 methyltransferase *Hha*I (M.*Hha*I) recognizes the dyad tetranucleotide site 5′-GCGC-3′ and transfers a methyl group from *S*-adenosyl-L-methionine (AdoMet) to the inner cytosine (underlined) on each strand of the DNA duplex. This enzyme was the first to be shown to operate via an intermediate in which the target residue is flipped out of the DNA helix into a concave active site in the enzyme (Klimasauskas *et al.*, 1994) (Figure 1). Given the high conservation of key primary sequence motifs, all members of the cytosine-5 methyltransferase family, which includes MTases from higher eukaryotes, are expected to share the same catalytic mechanism (Kumar *et al.*, 1994; Reinisch *et al.*, 1995). Studies of numerous other enzymes confirmed that base flipping is also a common and crucial event in DNA repair (for reviews, see Roberts, 1995; Vassylyev and Morikawa, 1997).

Although the crystal structures reveal details of the M.*Hha*I–DNA interface at atomic resolution, they provide only static glimpses of events along the reaction pathway (O'Gara *et al.*, 1996a). Mutational analysis of position 237, which had seemed to be a key residue during the *Hha*I–DNA interaction, indicated only a secondary role for Gln237 in the base-flipping mechanism (Mi *et al.*, 1995). Therefore, the driving force for the expulsion of the target base out of the DNA helix still remains unclear. A widely discussed issue is whether the enzyme has an active role in this process, or just captures a transiently extrahelical base as the DNA breathes (Winkler, 1994; Reinisch *et al.*, 1995; Slupphaug *et al.*, 1996). NMR measurements of imino-proton exchange indicate that lifetimes of individual G–C base pairs in free DNA are of the order of 10 ms (Guéron and Leroy, 1995). In view of the low turnover numbers of MTases (for M.*Hha*I, k_{cat} is 0.02/s in the methyltransferase reaction, and 0.15/s in the C5-proton exchange reaction; Wu and Santi, 1987), an active participation of the enzyme in accelerating the intrinsic rate of base pair opening appears not to be mandatory.

To obtain further insight into dynamic aspects of base flipping, we have undertaken an NMR study of the behaviour of the target cytosine residue during its interaction with the *Hha*I MTase in solution. Since ^{1}H spectral analysis would have been difficult due to the high molecular weight of the complexes (46 kDa), we incorporated ^{19}F nuclei into two specific positions of the DNA substrate. Comparisons of the free DNA duplex, the binary M.*Hha*I–DNA complex and a ternary M.*Hha*I–DNA–cofactor complex, which all form along the reaction pathway, were based on the chemical shifts and spin relaxation parameters of the two fluorine labels. The NMR experiments were complemented with gel mobility studies of MTase complexes involving both canonical and fluorinated DNA substrates.

Results

Design of the ^{19}F-labelled DNA for NMR studies

The ^{19}F-labelled DNA 12mer duplex used for the NMR measurements (Figure 2) contains in one strand 5-fluoro-

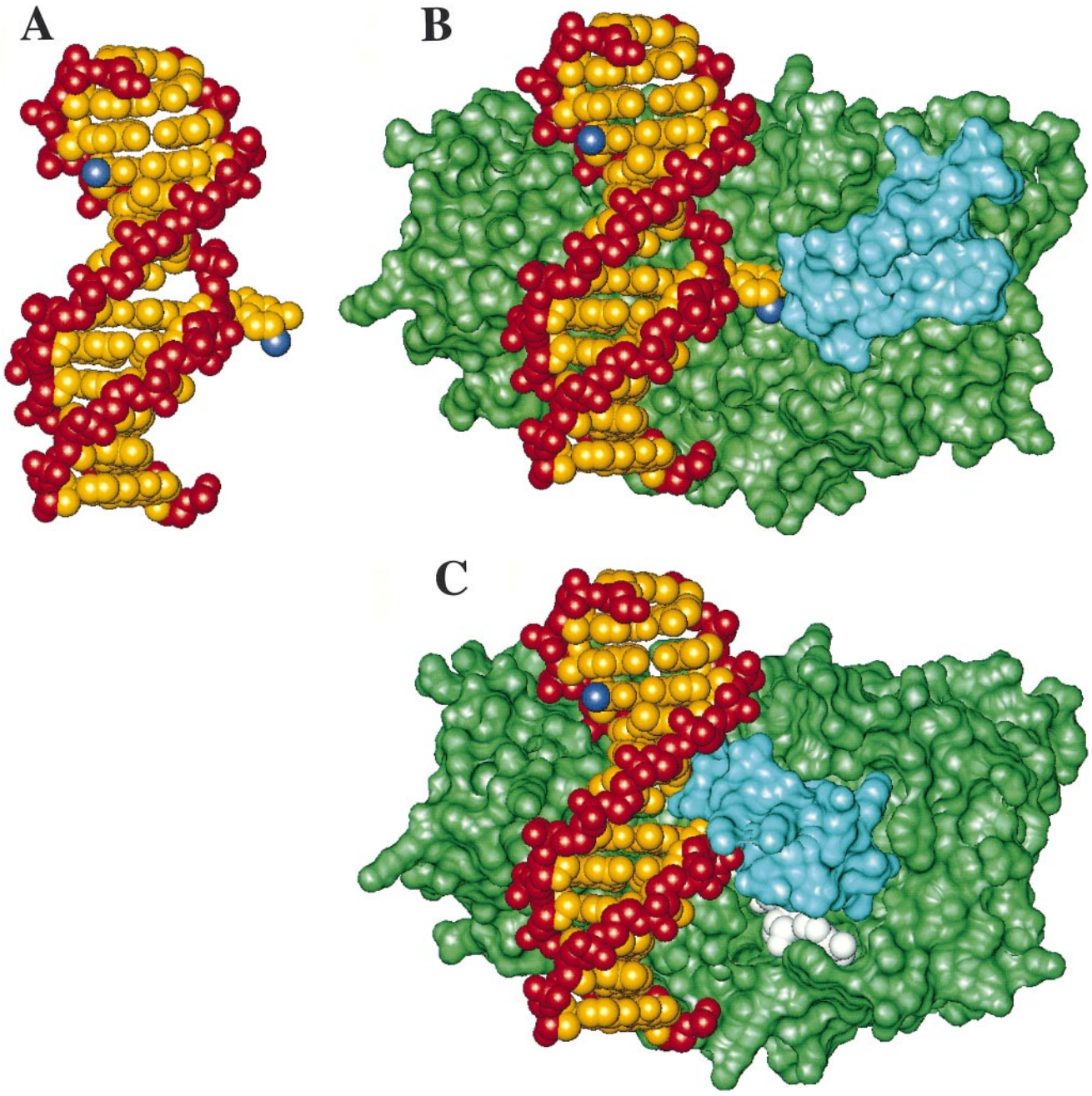

Fig. 1. Molecular models of the *Hha*I MTase complex with the self-complementary 13mer DNA duplex d(TGATAGCGCTATC)$_2$. The locations corresponding to the positions of the ^{19}F labels in the related DNA duplex used for the present NMR studies (F^1 and F^2 in Figure 2) are indicated by blue spheres. (**A**) DNA duplex from the structure shown in (C). (**B**) Putative model for a binary M.*Hha*I–DNA complex obtained by replacing the protein part in the ternary complex with that from the binary M.*Hha*I–AdoMet complex (Cheng *et al.*, 1993). The open form of M.*Hha*I allows flipping of the target cytosine in and out of the B-DNA helix. (**C**) Structure for the ternary M.*Hha*I–DNA–AdoHcy complex (Klimasauskas *et al.*, 1994). The closed conformation of the active site loop precludes flipping of the target base in and out of the active site. The DNA and AdoHcy are represented by van der Waals spheres for individual atoms, and a surface display was generated for the M.*Hha*I. The following colour code is used: DNA backbone, red; DNA bases, orange; protein surface, green, except for the active site loop of residues 80–100, which is cyan; AdoHcy, white.

cytosine in the target position of the dyad recognition site (underlined), which serves as a probe for monitoring the base-flipping motion under the influence of M.*Hha*I. A second 5-fluorocytosine residue is positioned three nucleotides upstream from the target nucleotide and serves as an internal reference that reflects the behaviour of a base that has no direct contacts with the enzyme. The second strand of the substrate DNA is pre-methylated in order to direct the docking of the MTase such that the other strand is targeted. Catalytic competence of this modified substrate has been confirmed by observation of irreversible formation of a ternary complex (Chen *et al.*, 1993; Klimasauskas *et al.*, 1994; Klimasauskas and Roberts, 1995) in a gel shift assay (Figure 3B, lane 4). As a reference, binding experiments with model 37mer duplexes show qualitatively similar behaviour of the

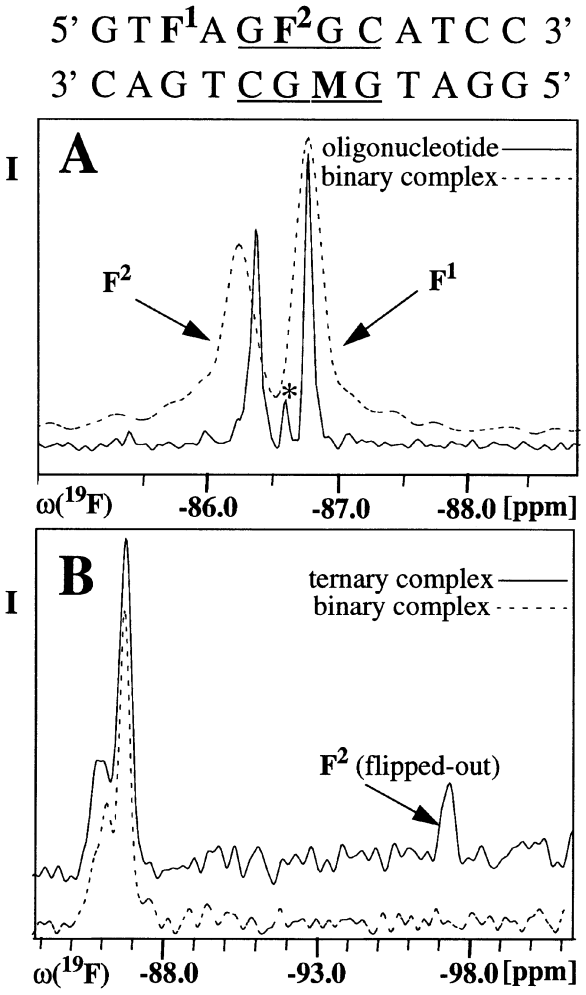

Fig. 2. One-dimensional ^{19}F NMR spectra of the DNA duplex shown at the top, where F and M denote 5-fluorocytosine and 5-methylcytosine, respectively, and the recognition site of the methyltransferase is underlined. The ^{19}F chemical shifts are relative to trifluoroacetic acid (TFA) calculated with Ξ (^{19}F/^{1}H) = 0.940 866 982 (Maurer and Kalbitzer, 1996). (**A**) Spectral region from –85 to –88 p.p.m. of ^{19}F NMR spectra of a free DNA duplex (solid line) and its binary M.*Hha*I complex (dashed line) recorded at a ^{19}F resonance frequency of 470.5 MHz. The concentration was 200 μM for both the free DNA and the complex (solvent 90% H_2O/10% D_2O, pH = 7.5; 100 mM NaCl; 20 mM KPO_4). The signal intensity, I, is plotted against the chemical shift in p.p.m., ω(^{19}F). The small peak marked with an asterisk is due to an impurity. (**B**) Spectral region from –85 to –100 p.p.m. of ^{19}F NMR spectra of the binary M.*Hha*I–DNA complex (dashed line) and the ternary M.*Hha*I–DNA–AdoHcy complex (solid line). The spectra were recorded and processed as described for (A), except that the concentration of the complex was 120 μM and the salt concentration was reduced to 50 mM NaCl, 10 mM KPO_4.

canonical targets and the corresponding fluorinated targets, although the affinity for the latter is slightly reduced (Figure 3A).

Base flipping in the binary MTase–DNA complex

The ^{19}F NMR spectra of the free duplex dodecanucleotide and its binary M.*Hha*I complex both exhibit two well-resolved signals corresponding to the two fluorine labels (Figure 2A). Consistent with the nearly 6-fold increase in the molecular weight, the resonance lines of the complex are significantly broadened. Furthermore, upon enzyme binding, the resonance at lower field is shifted by ~60 Hz (0.13 p.p.m.) downfield, suggesting that it corresponds to the fluorine label in the recognition site. The shift of the resonance at higher field remains unaltered, which is consistent with the assumption, based on the available crystal structures, that the F^1 atom (Figure 1) has no direct contacts with the protein in the specific complex. To confirm these resonance assignments, we recorded 1D 'truncated-driven' ^{19}F{^{1}H}-nuclear Overhauser enhancement (NOE) experiments (Wagner and Wüthrich, 1979) of both the free and the M.*Hha*I-bound dodecanucleotide duplex. From the DNA sequence (Figure 2), one expects a sequential TCH_3–C^{19}F NOE for the ^{19}F spin upstream of the recognition sequence, but no NOE with methyl groups for the ^{19}F spin of the target cytosine. When the spectral range of the TCH_3 groups, which extends from 1 to 2 p.p.m., was pre-saturated with a WALTZ-l6 sequence (Shaka *et al.*, 1983) for 30 ms, an ~5% attenuation of the higher field ^{19}F signal in the M.*HhaI*-bound DNA was observed (data not shown). The assignment inferred from the chemical shift difference between free and bound DNA (Figure 2A) could thus be confirmed, since neither the crystal structure of the ternary M.*Hha*I–DNA–AdoHcy complex (Klimasauskas *et al.*, 1994; O'Gara *et al.*, 1996a) (Figure 1C), nor a model of the binary MTase–DNA complex (Figure 1B) contains methyl groups near the target cytosine that could be the source of a methyl–^{19}F Overhauser effect. [In the free oligonucleotide, where the NOE transfer is less efficient because of the lower molecular weight, the anticipated sequential NOE turned out to be too weak to be detected when using pre-saturation times $<$100 ms; this was not unexpected since the TCH_3–C^{19}F distance in B-DNA is ~5.0 Å (Wüthrich, 1986).]

Since ^{19}F spins are sensitive probes for detecting changes in their environment (Rastinejad *et al.*, 1995), flipping of a ^{19}F-labelled base out of a DNA duplex would be expected to manifest itself by a large shift of the ^{19}F resonance. The magnitude of the shift observed upon formation of the binary 'complex I' (Figure 2A) corresponds to data from other ^{19}F NMR studies of protein–DNA complexes where flipping-out of target nucleotides could be excluded (Metzler and Lu, 1989; Rastinejad *et al.*, 1993), which suggests that the resonance at –86.2 p.p.m. in Figure 2A represents a 5-fluorocytosine residue that remains stacked in the B-DNA duplex. Figure 2A shows further that the ratio of the intensities of the target signal and the reference signal for the binary complex is at most slightly lower than that observed for the free duplex. This ratio is, however, significantly smaller for the binary complex at reduced ionic strength (Figure 2B), where enhanced substrate binding was also observed in a gel assay (data not shown). 1D ^{19}F NMR spectra recorded with the free DNA showed that this ionic strength-dependent variation is specific for the protein–DNA complex and is not observed for the free B-DNA duplex (data not shown). In view of the peak overlap and the poor signal-to-noise ratio obtained at reduced ionic strength, we refrained from extending these qualitative observations with a quantitative determination of the change of the ratio. The enzyme-induced loss of signal intensity from the target cytosine could be explained by the conformational equilibria between complex I and one or several less abundant species, where the signals of the latter would

have escaped detection either because of insufficient sensitivity of the experiment or due to line broadening by the conformational exchange.

In principle, the reduction of the $^{19}F^2$ signal intensity might alternatively be due to line broadening induced by the dynamic equilibrium between free DNA and complex I. However, this can be ruled out based on the following considerations: since the equilibrium concentration of the free DNA is much smaller than the concentration of the complexed DNA, the magnitude of the maximal possible line broadening of the major resonance (corresponding to complex I) can be estimated for the limiting case with two highly unequally populated states (Sandström, 1982). This yields values <10 Hz within the range of possible K_B values (10^5–10^9/M) and the chemical shift difference of 60 Hz observed for $^{19}F^2$ in the free DNA and complex I (Figure 2A), which compares with a line broadening of ~150 Hz that would be required to account for the observed reduction of the intensity by about a factor of 2 (Table I, Figure 2).

In apparent contrast with the quantitative formation of the binary complex inferred from the ^{19}F NMR spectrum (Figure 2A), gel mobility shift experiments show only weak complex formation in the absence of a cofactor (Figure 3B, lane 2). This apparent discrepancy is resolved if one assumes that complexes with the unaltered B-DNA elude electrophoretic detection due to fast exchange with the individual components, while complexes with the flipped-out base are longer lived and thus appear as retarded bands. Support for this notion comes from earlier investigations of the reaction kinetics by Wu and Santi (1987), which imply that the decomposition of complex I is faster than the maximal forward reaction velocity. Furthermore, observation of an inverse correlation between the stability of the target base pair in a DNA substrate and the amount of complex II detected in a gel shift assay suggests that formation of these binary complexes involves disruption of the target base pair in the DNA duplex (Klimasauskas and Roberts, 1995).

Prolonged separation resolves the band of the M.*Hha*I–DNA complex into two bands (Figure 3B, lane 5), indicating the presence of two types of species with different

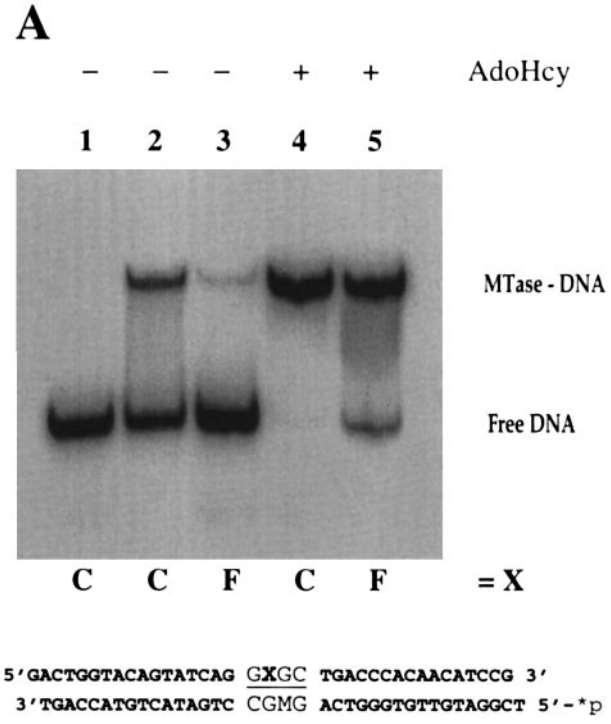

Table I. ^{19}F spin relaxation times in the free and M.*Hha*I-bound fluorinated DNA duplex[a]

	12mer DNA duplex (8 kDa)[b]		DNA–M.*Hha*I complex (46 kDa)[b]	
	F^1	F^2	F^1	F^2
T_1	528 ± 12	562 ± 11	838 ± 35	815 ± 55
T_2	10.7 ± 0.8	12.3 ± 1.5	1.71 ± 0.27	2.11 ± 0.40
$T_{1\rho}$	25.8 ± 2.2	23.5 ± 3.0	3.58 ± 0.25	4.11 ± 0.31

[a]In milliseconds with standard deviations. Data were collected at a ^{19}F resonance frequency of 470.5 MHz; rotating frame relaxation times, $T_{1\rho}$, were measured at the spin-lock field strength of 8181 rad/s.
[b]F^1 and F^2 indicate the sequence positions of the two ^{19}F labels as defined in Figure 2, with F^2 being the target nucleotide. The molecular weight is indicated in brackets.

Fig. 3. Gel mobility shift analysis of interactions between M.*Hha*I and DNA duplexes. 5′-^{32}P-labelled duplexes (5 nM) were incubated with 60 nM M.*Hha*I and, where applicable, 50 μM cofactor. (**A**) Binding of 37mer hemimethylated (M = 5-methylcytosine) duplexes with either cytosine (X = C) or 5-fluorocytosine (X = F) at the target site. Electrophoresis was performed in 8% (19:1) gel for 2 h. Lane 1, DNA control; lanes 2–5 with M.*Hha*I and, where indicated at the top, with AdoHcy. (**B**) Binding of the fluorinated 12mer duplex. Electrophoresis was performed in 7% (37.5:1) gel for 1 h. Lane 1, DNA control; lane 2, DNA + M.*Hha*I; lane 3, DNA + M.*Hha*I + AdoHcy; lane 4, DNA + M.*Hha*I + AdoMet (after overnight incubation). Lanes 5 and 6 are expanded presentations of the experiments in lanes 2 and 3.

electrophoretic mobilities (denoted 'complex II' and 'complex III'). Similar heterogeneity in electrophoretic behaviour was also observed for the M.*Hha*I complexes with canonical 37mer duplexes (Klimasauskas and Roberts, 1995), where it was attributed to conformational variations in the protein. The most prominent rearrangement involves the so-called active site loop (residues 80–99), which undergoes a movement of nearly 25 Å at its tip. The faster moving band of complex III would then represent the more compact, 'closed' form of the enzyme, in which the loop locks the flipped-out target cytosine into the active site pocket (Figure 1C). Altogether, the present observations imply that at least three different MTase–DNA complexes are formed during the methylation reaction:

$$\text{M.}Hha\text{I} + \text{DNA} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \text{complex I} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \text{complex II} \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} \text{complex III} \qquad [1]$$

The target base would be stacked in the B-DNA duplex in complex I, flipped out in complex II (Figure 1B) and locked in the active site in complex III (Figure 1C). $K_B = k_1/k_{-1}$, $K_F = k_2/k_{-2}$ and $K_L = k_3/k_{-3}$ (B, F and L stand for 'binding', 'flipping' and 'locking', respectively) denote the three equilibrium constants which determine the relative abundances of free enzyme, free DNA duplex and the complexes I, II and III in solution. In the case of the fluorinated dodecamer duplex, complex I constitutes a major species, as evidenced by NMR (Figure 2A), and that this complex represents specific recognition of the DNA is indicated by the fact that only the chemical shift of $^{19}F^2$ but not of $^{19}F^1$ is affected by complex formation (Figure 2A), as one would predict from the X-ray structure of the specific M.*Hha*I–DNA complex (Figure 1). If the intensities of the two upper bands in the gel mobility assay are used to estimate the relative abundances of the complexes II and III, one arrives at a value for K_L of ~0.25. For the 37mer DNA duplex that contains a canonical hemimethylated site, only one band is observed that corresponds to complex II. This band does not split up even after prolonged electrophoresis (Klimasauskas and Roberts, 1995), suggesting that there is either negligible formation of complex III (small value of K_L), or fast exchange between the complexes II and III during electrophoresis. The latter possibility appears more likely because transient formation of complex III is implied by the capacity of M.*Hha*I to catalyse two side reactions, i.e. hydrolytic deamination of the target cytosine to uracil (Yang *et al.*, 1995) and the exchange of the C5-proton of the target cytosine with solvent protons (Wu and Santi, 1987). Both of these reactions occur only in the absence of cofactor, and both require covalent activation of the cytosine ring.

Base flipping in the ternary MTase–DNA–AdoHcy complex

A key observation resulted from ^{19}F NMR experiments recorded after addition of *S*-adenosyl-L-homocysteine (AdoHcy) to the M.*Hha*I–DNA complex. Binding of this competitive inhibitor (and metabolic product of AdoMet) leads to the appearance of a new ^{19}F resonance at –97.3 p.p.m. that is shifted from the lines of the stacked 5-fluorocytosines by ~11 p.p.m. (the shift is to the right in Figure 2B). This large chemical shift change reflects a major alteration of the environment of the target nucleotide, such as one expects between the stacked and flipped-out states of fluorinated cytosine. Since the new resonance line is not significantly broadened when compared with the other signals (Figure 2B), the lifetime of the target base in the state manifested by this line must be in the range of milliseconds or longer. The observation that amplitude and chemical shift of the resonance line near –86.2 p.p.m., which corresponds to complex I, undergo no significant alterations upon cofactor binding suggests that the new signal at high field largely derives from the ensemble of flipped-out conformers (complex II). We thus arrive at the conclusion that binding of AdoHcy locks the different flipped-out conformers into a single, stable species (complex III). Concurrently, gel mobility assays show enhanced formation of complex III after addition of AdoHcy (Figure 3B, lane 6), indicating a significant increase in K_L. This effect is even more pronounced in the case of the canonical hemimethylated substrate, for which complex III is the sole species observed after addition of AdoHcy (Figure 3A, lane 4; see also Klimasauskas and Roberts, 1995). These observations are fully consistent with earlier kinetic and gel binding studies, which had shown that AdoHcy strongly inhibits the decay of the M.*Hha*I–DNA complex (Wu and Santi, 1987; Yang *et al.*, 1995).

Dynamics of base pair opening

To complement the observations of multiple distinct environments of the target 5-fluorocytosine with data on the exchange dynamics between stacked and flipped-out states, we determined the T_1, T_2 and $T_{1\rho}$ spin relaxation times of ^{19}F for the free duplex and the binary complex. [Unfortunately, no measurements of corresponding ^{19}F spin relaxation times were obtained for the ternary complex, because of the lower solubility and concomittantly reduced long-term stability of the system (see also Materials and methods).] Nearly identical relaxation times were observed for the two ^{19}F spins in both the free DNA and the complex (Table I). The $T_{1\rho}$ measurements performed at a spin-lock field strength of $\omega_1 = 8181$ rad/s yielded significantly longer transverse relaxation times than the T_2 measurements, indicating the presence of conformational exchange processes in the frequency range from ~1 Hz to 10 kHz. To estimate the effective correlation time, we determined $T_{1\rho}$ for the free dodecanucleotide duplex also as a function of the spin-lock power, with ω_1 varying from 13 500 to 850 rad/s. Since $T_{1\rho}$ did not vary noticeably over this range of the spin-lock power, the effective correlation time for the motional modes influencing T_2 must be longer than ~1 ms (Deverell *et al.*, 1970; Szyperski *et al.*, 1993). The allowed range thus includes the lifetime of ~10 ms measured for individual G–C base pairs in DNA by NMR studies of imino-proton exchange (Guéron and Leroy, 1995), and most likely reflects motions of the 5-fluorocytosines that are associated with spontaneous DNA breathing. Overall, the close similarity of the relaxation parameters for the two ^{19}F labels indicates that base pair lifetimes of the target 5-fluorocytosine (F^2) and the reference 5-fluorocytosine (F^1) are comparable in

free DNA as well as in the M.*Hha*I–DNA complex, and hence no significant acceleration of the internal motional processes in the DNA duplex could be detected that would be due to the binding of M.*Hha*I. This does not rule out that there might be conformational exchange with scarcely populated states, which would exhibit an effective correlation time >1 ms. Such motional modes could be dramatically accelerated upon complex formation without being manifested in the ^{19}F spin relaxation data.

Discussion

The M.*Hha*I cytosine-5 methyltransferase operates by a sequential kinetic scheme in which the DNA first binds to the enzyme, and this binary complex then binds the cofactor AdoMet (Wu and Santi, 1987). The binary MTase–DNA complex is thus a discrete step in the reaction pathway. The present NMR data now demonstrate that the interactions with the enzyme in the binary complex lead to a dynamic equilibrium between an ensemble of flipped-out states of the target nucleotide (complexes II and III) and the stacked state (complex I), as described by K_F (Equation [1]). Taking account of the implicated conformational diversity of the flipped-out base in complex II, at least on the ^{19}F chemical shift time scale, provides compelling evidence that there is essentially no specific recognition or trapping of the target cytosine by M.*Hha*I in the absence of cofactor. This is supported further by the fact that formation of complex II is insensitive to the nature of the target base *per se*, but inversely correlates with the strength of the base pair (Klimasauskas and Roberts, 1995; Yang *et al.*, 1995). A major mechanistic implication from this finding is that it is not crucial for the flipping-out of the target base that it is trapped in the catalytic pocket of M.*Hha*I. It also suggests that certain MTase–DNA contacts observed in the ternary complexes (Figure 1C), such as those from the residues in the active site loop and the catalytic site, may play a less important role in the flipping mechanism than was earlier anticipated based on the X-ray structures alone (Klimasauskas *et al.*, 1994; O'Gara *et al.*, 1996a). On the other hand, the lack of a clearly preferred conformation for the flipped-out base is remarkable in light of the large stabilization of the flipped-out state achieved upon interaction with M.*Hha*I, where the flipping constant K_F increases the value of 10^{-5}–10^{-6} observed for the free DNA (Gueron and Leroy, 1995) to ~1 as estimated for the binary complex (see above). Clearly, the numerous protein contacts to both the recognition bases and the phosphodiester backbone are needed to keep the DNA in its strained conformation. As no structures are yet available for binary M.*Hha*I–DNA complexes, we propose the putative model of Figure 1B as an illustration of the nature of the MTase–DNA interactions. Nonetheless, it seems likely that the 10^6-fold stabilization of the open state versus the stacked state of the target base is achieved not simply by virtue of the favourable protein–DNA interactions exerted in complex II, but to a certain extent through the destabilization of the B-helical conformation in complex I.

Formation of the ternary M.*Hha*I–DNA–AdoMet complex is a subsequent discrete step in the physiological methylation reaction. To avoid catalytic turnover during the NMR measurements, we used AdoHcy in place of AdoMet. AdoHcy is a potent competitive inhibitor with respect to AdoMet (Wu and Santi, 1987), and is expected to mimic closely the binding of AdoMet in the cofactor pocket (O'Gara *et al.*, 1996b). Upon addition of AdoHcy, the target cytosine is trapped in a unique conformation outside of the DNA helix (Figure 2B), and the protein itself undergoes a conformational rearrangement to a more compact structure (Figure 3, lane 6). These observations are in full agreement with the X-ray structures of the ternary complexes containing AdoHcy, where the target cytosine is firmly locked in the active site and the return path for the target base is blocked in the closed form of the active site loop (Figure 1B and C). The locking of the extrahelical base in the enzyme active site initiates transient formation of a covalent link between the Cys81 thiolate of M.*Hha*I and the C6 atom of the target cytosine, thus providing the nucleophilic activation of the cytosine ring that is required for the subsequent methyl transfer from the cofactor AdoMet (reviewed in Kumar *et al.*, 1994).

The ^{19}F spin relaxation data suggest that M.*Hha*I does not accelerate base pair opening beyond the intrinsic rate of DNA breathing, which corresponds to a lifetime for G–C base pairs of ~10 ms (Guéron and Leroy, 1995). The lower limit for the rate constant k_2 of 0.15/s, which corresponds to a lifetime of ~4 s at 37°C, can be derived from the turnover rate of the M.*Hha*I in the C5-hydrogen exchange reaction (Wu and Santi, 1987). These numbers would be consistent with the hypothesis that cytosine-5 methyltransferases might function by trapping spontaneous base pair openings as initial flipped-out intermediates (Winkler, 1994). However, we showed that catching of the target base is not essential for its flipping out of the helix, but rather serves for proper spatial positioning of the extrahelical base, AdoMet and the catalytic groups on the enzyme during the methyl transfer reaction. Another caveat to this assumption is that based on structural models for M.*Hha*I and M.*Hae*III, these enzymes flip the target base through the minor groove of the DNA. Since sequence-specific contacts occur on the major groove side in the complex III (Figure 1C), it appears straightforward to expect that these interactions will be largely maintained during the base-flipping process in the specific complexes I and II (Figure 1B) (Kumar *et al.*, 1994; Reinisch *et al.*, 1995). In contrast, model calculations indicate that movements towards the major groove should account for most of the spontaneous base openings, since base rotation in the opposite direction would be much less favourable energetically and lead to steric crowding of the atoms involved in the base pair hydrogen bonding (Ramstein and Lavery, 1988). Therefore, an active participation of the enzyme seems necessary to create an exit path for the target base on the minor groove side. Flipping through the major groove cannot be excluded by our data, but this would invoke the existence of an intermediate enzyme–DNA complex in which the recognition contacts would be substantially different from those observed in complex III. Recently, an active role in the flipping mechanism has also been suggested for uracil DNA glycosylase, an enzyme which exerts no sequence specificity during catalysis (Slupphaug *et al.*, 1996). Surprisingly, although this enzyme flips uracil residues through the major groove, the conformational distortions of the DNA helix in the reaction complex appear to be quite similar to those

observed for M.*Hha*I (Cheng and Blumenthal, 1996; Vassylyev and Morikawa, 1997). Both enzymes induce substantial interstrand separation of the phosphodiester backbone in the 5′ direction from the target nucleotide, which may serve to lower the activation energy barrier required for base flipping. This similarity most likely relates to the intrinsic deformability of the DNA double helix, and there seem to be at least two alternative strategies for this conformation to be induced and stabilized by enzymes with two-domain topologies. Consistent with the opposite directions of base flipping by the two enzymes, the active site pocket that binds the target base and the group that fills in the void in the DNA helix would then be located on the opposite sides relative to the target DNA strand.

Materials and methods

Purified oligodeoxyribonucleotides were obtained from New England BioLabs except for the methylated dodecanucleotide strand that was synthesized at MBI Fermentas (Vilnius, Lithuania). Oligonucleotide concentrations were determined spectrophotometrically based on the A_{260}/A_{320} values, with molar extinction coefficients calculated as the sum of the contributions from the individual deoxynucleotides. Oligonucleotides were 5′-labelled using T4 polynucleotide kinase (Fermentas) and [γ-^{32}P]ATP (Amersham). DNA duplexes were prepared by annealing of the individual strands from 80 to 20°C over several hours.

M.*Hha*I was isolated as described by Kumar *et al.* (1992); exhaustive dialysis prior to chromatography yielded the AdoMet-free enzyme. The enzyme was purified to homogeneity by passing it through a pre-column of Q-Sepharose followed by column chromatography on S-Sepharose.

Gel mobility shift analysis

5′-^{32}P-labelled duplexes (5 nM) were incubated with 60 nM M.*Hha*I and, where applicable, 50 μM cofactor. Aliquots of 20 μl of the solutions were incubated at 20°C for 30 min in 20 mM KPO_4 (pH 7.5), 5 mM Na_2EDTA and 10% glycerol. Five μl samples were loaded onto polyacrylamide:bisacrylamide gels and resolved by electrophoresis in 45 mM Tris–borate/1 mM Na_2EDTA at 10 V/cm for 1–2 h. Dried gels were autoradiographed to an X-ray film and the films were scanned with a ScanMaker E_6 densitometer (Microtek).

NMR measurements

All NMR data were collected at T = 25°C on a Bruker DRX500 spectrometer equipped with a ^{19}F{^{1}H} probehead and operating at a ^{19}F resonance frequency of 470.5 MHz. The one-dimensional (1D) ^{19}F NMR spectra of the free DNA and the binary M.*Hha*I–DNA complex shown in Figure 2A were recorded at 200 μM concentration (solvent 90% H_2O/10% D_2O; pH = 7.5; 100 mM NaCl; 20 mM KPO_4). The M.*Hha*I–DNA complex was prepared with a 1.5-fold excess of the protein to ensure quantitative binding of the DNA. The acquisition parameters for the free DNA (for the complex) were: 27 (13.6) ms acquisition time, 4096 (32 384) scans, 1.4 (11.2) h of measurement time. ^{1}H decoupling during acquisition was achieved with DIPSI-2 (Shaka *et al.*, 1988). Prior to Fourier transformation, the spectra were multiplied with a sine-squared (sine) function shifted by 45° (22.5°) (DeMarco and Wüthrich, 1976). The 1D ^{19}F NMR spectra of the binary M.*Hha*I–DNA complex and the ternary M.*Hha*I–DNA–AdoHcy complex (shown in Figure 2B) were recorded and processed as above, except that the concentration of the complex was 120 μM and the ionic strength was reduced to one half by lowering the salt concentration to 50 mM NaCl and 10 mM KPO_4. A total of 154 000 scans were acquired for the binary complex, yielding a measurement time of 48 h. Following addition of 300 μM AdoHcy, 385 000 additional scans were acquired in 120 h to observe the ternary complex. During the second part of the experiment, a precipitate from part of the dissolved ternary complex was formed, as observed by inspection of the sample at the end of the NMR experiment.

Relaxation times were obtained by non-linear fits of a single exponential function to the experimental data. For the free DNA duplex, longitudinal relaxation times, T_1, were measured with ten 1D inversion recovery experiments, [180°(^{19}F)–T_{rel}–90°(^{19}F)–Acq.], with relaxation delays T_{rel} = 10, 50, 100, 150, 200, 300, 400, 800, 1200 and 1600 ms. The acquisition parameters were the same as for the 1D ^{19}F spectra in Figure 2A, except that 320 scans were acquired per spectrum and the relaxation delay between scans was set to 5 s, resulting in a total measurement time of 5.1 h. For the M.*Hha*I–DNA complex, eight inversion recovery experiments were recorded, with relaxation delays T_{rel} = 10, 100, 200, 300, 400, 800, 1200 and 1600 ms, acquiring 2048 scans with a total recording time of 28 h. Transverse relaxation times, T_2, for the free duplex were measured with six 1D spin-echo spectra [90°(^{19}F)–T_{rel}/2–180°(^{19}F)–T_{rel}/2–Acq.], with relaxation delays T_{rel} = 1, 10, 15, 20, 30 and 40 ms. The same acquisition parameters were used as for T_1, except that 256 scans were acquired, and the total measurement time was 2.3 h. ^{1}H decoupling during the relaxation delay was achieved with DIPSI-2 (Shaka *et al.*, 1988). For the complex, five 1D spin-echo spectra with relaxation delays T_{rel} = 0.1, 1, 3, 5 and 7 ms were measured, accumulating 2048 scans with a total recording time of 15.3 h. Rotating frame relaxation times, $T_{1\rho}$, for the free oligonucleotide were measured with ten 1D spin-lock spectra [90°$_x$(^{19}F)–spin-lock$_y$(T_{rel})–Acq.]. The spin-lock field strength was 8181 rad/s, T_{rel} = 1, 5, 10, 15, 20, 30, 40, 50, 60 and 80 ms, the acquisition parameters were the same as for T_1, and the total measurement time was 3.0 h. For the complex, eight 1D spin-lock spectra with T_{rel} = 0.1, 1, 3, 5, 7, 9, 11 and 13 ms were recorded, accumulating 2048 scans with a total recording time of 24.5 h. In addition, the lock field dependence of $T_{1\rho}$ for the free duplex was measured with a series of experiments with ω_1 = 13 500, 6750, 3375, 1700, 1200 and 850 rad/s; the total measuring time was 70 h.

Molecular models

The molecular models of Figure 1 were generated using VMD (Humphrey *et al.*, 1996). The protein surface was computed with a probe radius of 1.4 Å. The putative model for the binary M.*Hha*I–DNA complex in Figure 1B was obtained by replacing the protein in the ternary M.*Hha*I–DNA–AdoHcy complex (Klimasauskas *et al.*, 1994; PDB code 1MHT) with that from the binary M.*Hha*I–AdoMet complex (Cheng *et al.*, 1993; PDB code 1HMY). For this, the two proteins were superimposed manually and then subjected to two cycles of refinement using the LSQ routine of O (Jones *et al.*, 1991). The calculation resulted in an r.m.s. deviation of 0.5 Å between the core residues of the large domain, excluding the active site loop (residues 1–79, 100–195 and 285–327).

Acknowledgements

The authors thank Dr R.J.Roberts for a gift of the fluorinated oligonucleotide strands, Dr D.Moskau of Spectrospin A.G., Fällanden, Switzerland for help in setting up the experiments on a Bruker DRX500 spectrometer equipped for ^{19}F NMR measurements, Dr Z.Maneliene of MBI Fermentas, Vilnius, Lithuania, for the synthesis of the methylated dodecanucleotide strand, and Mrs E.Ulrich and Mrs E.Zollinger for the careful processing of the text. This work was supported by an EMBO fellowship (ASTF 8424), a Howard Hughes Medical Institute International Research Scholarship grant to S.K. and the Schweizerischer Nationalfonds (project 31.32033.91).

References

Anderson,J.E. (1993) Restriction endonucleases and modification methylases. *Curr. Opin. Struct. Biol.*, **3**, 24–30.

Bird,A.P. (1996) The relationship of DNA methylation to cancer. *Cancer Surv.*, **28**, 87–101.

Chen,L., MacMillan,A.M., Chang,W., Ezaz-Nikpay,K., Lane,W.S. and Verdine,G.L. (1991) Direct identification of the active site nucleophile in a DNA (cytosine-5)-methyltransferase. *Biochemistry*, **30**, 11018–11025.

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

Cheng,X., Kumar,S., Pósfai,J., Pflugrath,J.W. and Roberts,R.J. (1993) Crystal structure of the *Hha*I DNA methyltransferase complexed with *S*-adenosyl-L-methionine. *Cell*, **74**, 299–307.

DeMarco,A. and Wüthrich,K. (1976) Digital filtering with a sinusoidal window function: an alternative technique for resolution enhancement in FT NMR. *J. Magn. Resonance*, **24**, 201–204.

Deverell,C., Morgan,R.E. and Strange,J.H. (1970) Studies of chemical exchange by nuclear magnetic relaxation in the rotating frame. *Mol. Phys.*, **18**, 553–559.

Guéron,M. and Leroy,J.L. (1995) Studies of base pair kinetics by NMR measurement of proton exchange. *Methods Enzymol.*, **261**, 383–413.

Humphrey,W., Dalke,A. and Schulten,K. (1996) VMD—visual molecular dynamics. *J. Mol. Graphics*, **14**, 33–38.

Jones,T.A., Zou,J.Y., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for building protein models in electron density maps and the location of errors in the models. *Acta Crystallogr.*, **A47**, 110–119.

Jost,J.P. and Saluz,H.P. (1993) *DNA Methylation: Molecular Biology and Biological Significance*. Birkhäuser, Basel.

Klimasauskas,S. and Roberts,R.J. (1995) M.*Hha*I 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) *Hha*I methyltransferase flips its target base out of the DNA helix. *Cell*, **76**, 357–369.

Kumar,S., Cheng,X., Pflugrath,J.A. and Roberts,R.J. (1992) Purification, crystallization and preliminary X-ray diffraction analysis of an M.*Hha*I–AdoMet complex. *Biochemistry*, **31**, 8648–8653.

Kumar,S., Cheng,X., Klimasauskas,S., Mi,S., Pósfai,J., Roberts,R.J. and Wilson,G.G. (1994) The DNA (cytosine-5) methyltransferases. *Nucleic Acids Res.*, **22**, 1–10.

Maurer,T. and Kalbitzer,H.R. (1996) Indirect referencing of ^{31}P and ^{19}F NMR spectra. *J. Magn. Resonance.*, **B113**, 177–178.

Metzler,W.J. and Lu,P. (1989) λ *cro* repressor complex with O_R3 operator DNA: ^{19}F NMR observations. *J. Mol. Biol.*, **205**, 149–164.

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

O'Gara,M., Klimasauskas,S., Roberts,R.J. and Cheng,X. (1996a) Enzymatic C5-cytosine methylation of DNA: mechanistic implications of new crystal structures for *Hha*I methyltransferase–DNA–AdoHcy complexes. *J. Mol. Biol*., **261**, 634–645.

O'Gara,M., Roberts,R.J. and Cheng,X. (1996b) A structural basis for the preferential binding of hemimethylated DNA by *Hha*I methyltransferase. *J. Mol. Biol*., **263**, 597–606.

Ramstein,J. and Lavery,R. (1988) Energetic coupling between DNA bending and base pair opening. *Proc. Natl Acad. Sci. USA*, **85**, 7231–7235.

Rastinejad,F., Artz,P. and Lu,P. (1993) Origin of the asymmetrical contact between *lac* repressor and *lac* operator DNA. *J. Mol. Biol.*, **233**, 389–399.

Rastinejad,F., Evilla,C. and Lu,P. (1995) Studies of nucleic acids and their protein interactions by ^{19}F NMR. *Methods Enzymol.*, **261**, 560–574.

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

Roberts,R.J. (1995) On base flipping. *Cell*, **82**, 9–12.

Sandström,J. (1982) *Dynamic NMR Spectroscopy.* Academic Press, New York.

Shaka,A.J., Lee,C.J. and Pines,A.J. (1983) An improved sequence for broadband decoupling: WALTZ-16. *J. Magn. Resonance*, **52**, 335–338.

Shaka, A.J., Lee,C.J. and Pines,A.J. (1988) Iterative schemes for bilinear operators: application to spin decoupling. *J. Magn. Resonance*, **77**, 274–293.

Slupphaug,G., Mol,C.D., Kavli,B., Arvai,A.S., Krokan,H.E. and Tainer,J.A. (1996) A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA. *Nature*, **384**, 87–92.

Szyperski,T., Luginbühl,P., Otting,G. and Wüthrich,K. (1993) Protein dynamics studied by rotating frame ^{15}N spin relaxation times. *J. Biomol. NMR*, **3**, 151–164.

Vassylyev,D. and Morikawa,K. (1997) DNA-repair enzymes. *Curr. Opin. Struct. Biol*., **7**, 103–109.

Wagner,G. and Wüthrich,K. (1979) Truncated driven nuclear Overhauser effect (TOE): a new technique for studies of selective ^1H–^1H Overhauser effects in the presence of spin diffusion. *J. Magn. Resonance*, **33**, 675–680.

Winkler,F.K. (1994) DNA totally flipped out by methylase. *Structure*, **2**, 79–83.

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

Wüthrich,K. (1986) *NMR of Proteins and Nucleic Acids.* Wiley, New York.

Yang,A.S., Shen,J.-C., Zingg,J.-M., Mi,S. and Jones,P.A. (1995) *Hha*I and *Hpa*II methyltransferases bind DNA mismatches, methylate uracil and block DNA repair. *Nucleic Acids Res.*, **23**, 1381–1388.

Received May 26, 1997; revised July 29, 1997; accepted October 17, 1997