Long-Range Structural and Dynamical Changes Induced by Cofactor Binding in DNA Methyltransferase M.HhaI‡

Hongjun Zhou, Whitney Shatz, Matthew M. Purdy, Nick Fera, Frederick W. Dahlquist,* and Norbert O. Reich*

Department of Chemistry and Biochemistry, University of California, Santa Barbara, California 93106-9510

Received December 30, 2006; Revised Manuscript Received March 27, 2007

ABSTRACT: The bacterial DNA cytosine methyltransferase M.HhaI sequence-specifically modifies DNA in an *S*-adenosylmethionine dependent reaction. The enzyme stabilizes the target cytosine (GCGC) into an extrahelical position, with a concomitant large movement of an active site loop involving residues 80−99. We used multidimensional, transverse relaxation-optimized NMR experiments to assign nearly 80% of all residues in the cofactor-bound enzyme form, providing a basis for detailed structural and dynamical characterization. We examined details of the previously unknown effects of the cofactor binding with M.HhaI in solution. Addition of the cofactor results in numerous structural changes throughout the protein, including those decorating the cofactor binding site, and distal residues more than 30 Å away. The active site loop is involved in motions both on a picosecond to nanosecond time scale and on a microsecond to millisecond time scale and is not significantly affected by cofactor binding except for a few N-terminal residues. The cofactor also affects residues near the DNA binding cleft, suggesting a role for the cofactor in regulating DNA interactions. The allosteric properties we observed appear to be closely related to the significant amount of dynamics and dynamical changes in response to ligand binding detected in the protein.

S-Adenosylmethionine (AdoMet)[1] dependent DNA methyltransferases sequence-specifically modify DNA in a wide range of organisms. In bacteria, modification can occur at cytosine N^4, cytosine C^5, and adenine N^6 (*1*). DNA methylation regulates gene transcription, DNA mismatch repair, and DNA replication, and some DNA methyltransferases form part of restriction−modification systems. DNA adenine methylation controls the virulence of several human pathogens, including *Salmonella typhimurium* (*2*−*4*), suggesting that DNA methyltransferase inhibitors may form a novel class of antibiotics (*3*). In eukaryotes, DNA methylation occurs exclusively at cytosine C^5 and is essential for embryogenesis (*5*). Alterations in DNA methylation are an early, nonmutagenic, step in tumorigenesis; up to 65% of all cancers have an epigenetic basis (*6*−*8*).

All DNA methyltransferases require the methyl donor AdoMet and flip the target nucleoside out of the DNA helix (*1*, *9*, *10*). M.HhaI was the first DNA cytosine methyltransferase whose structure was determined; the enzyme uses AdoMet to methylate C^5 of the inner cytosine in the cognate sequence 5′-GCGC-3′ and generates *S*-adenosylhomocysteine (AdoHcy). The first crystal structure of M.HhaI (∼37 kDa) revealed a protein consisting of three parts, a large domain (residues 1−193 and 304−327), a small domain (residues 194−275), and a hinge region (residues 276−303) (*11*). The key catalytic residue, Cys81, is located in an extended active site loop (residues 80−99) in the large domain, above a deep, wide DNA-binding cleft formed by the two domains. Subsequent crystal structures of various DNA−protein−cofactor complexes (*12*−*20*) showed a striking feature of the protein−DNA interaction: cognate DNA binding causes the active site loop to swing a distance of some 30 Å, from an "open" conformation in the DNA-free state to a "closed" conformation in the DNA-bound state. This change also appears to be accompanied by a concerted reorientation of the domains toward each other.

While interactions between M.HhaI and various DNA sequences have been studied extensively in binary complexes without a cofactor or in ternary complexes with a cofactor, the structural effect of the cofactor on the protein is less well understood. Cofactor binding stabilizes the protein−DNA complex and confers some DNA sequence specificity (*21*−*23*), and cofactor (product) release is a rate-limiting step in enzyme turnover (*23*, *24*). The lack of information about the effects of the cofactor on M.HhaI structure is partly due to difficulties in crystallizing the free protein and solubility issues that arise when working with the protein using NMR techniques (*25*). Additionally, the active site loop exhibits high thermal factors in the crystal structure of the protein−cofactor binary complex (*11*), consistent with this region being highly flexible. These motions are significantly reduced to normal levels in the ternary complex bound with cognate DNA (*19*). These interesting ligand-induced changes in loop structure and dynamics observed in X-ray structures are likely to be important for various aspects of the enzyme's function. Here we describe our efforts to overcome solubility and

‡ Resonance assignments will be deposited in BioMagResBank as entry BMRB 15170.

* To whom all correspondence should be addressed. E-mail: reich@chem.ucsb.edu or dahlquist@chem.ucsb.edu. Phone: (805) 893-8368. Fax: (805) 893-4120.

[1] Abbreviations: AdoMet, *S*-adenosylmethionine; AdoHcy, *S*-adenosyl-L-homocysteine; NMR, nuclear magnetic resonance; TROSY, transverse relaxation-optimized spectroscopy; HSQC, heteronuclear single-quantum coherence.

10.1021/bi602662e CCC: $37.00
Published on Web 05/25/2007

stability issues and assign the backbone resonances using high-resolution NMR techniques. We also describe the structural changes that occur when M.HhaI binds to its cofactor, the dynamic properties of the protein, and the correlation of this work with mutational studies addressing cofactor and DNA binding activities.

MATERIALS AND METHODS

Protein Expression and Purification. PCR primers were obtained from Integrated DNA Technologies. DNA sequencing was performed by the University of California facility at Berkeley. Restriction endonucleases and *Escherichia coli* strains ER2566 (for protein expression) and ER1727 (for plasmid production) were obtained from New England Biolabs. Isotopically labeled compounds were purchased from Cambridge Isotope Laboratories. Isopropyl D-thiogalactopyranoside (IPTG) was Biotech grade from Fisher. Ni-NTA agarose and cellulose phosphate P11 resin were from Qiagen and Whatman, respectively. *S*-Adenosyl-L-homocysteine (AdoHcy) was obtained from Sigma-Aldrich.

A His_6-tagged version of the previously described solubility-enhanced Δ324G M.HhaI truncation (*25*) was prepared and used throughout this study. The M.HhaI gene (without codons for the four C-terminal amino acids) was PCR amplified from the pHSHW-5 plasmid (provided by S. Kumar, New England Biolabs) using primers 5′-GAGGAGAAATTAATCATGATTGAAATAAAAG-3′ and 5′-CAATTTAATATGGCTCGAGATTTAATGATGAACCAATG-3′. After digestion with BspHI and XhoI endonucleases, the gene was ligated into NcoI- and XhoI-digested pET28a plasmid using the compatible overhangs from the NcoI and BspHI digestions. This produced the expression plasmid M.HhaIΔ324-His_6/pET28a encoding amino acid residues 1−324 of M.HhaI fused to a C-terminal His_6 tag through a Leu-Glu linker.

M.HhaI protein was expressed in *E. coli* ER2566 with 30 μg/mL kanamycin present in all growth media. Uniform labeling was accomplished using a modification of the protocol of Marley et al. (*26*). Cells were grown in shaker flasks at 37 °C in LB medium to an optical density of 0.7 at 600 nm and harvested by centrifugation. The cells were washed by resuspension in unlabeled M9 medium followed by centrifugation. The washed cells were resuspended in either singly labeled ([^{15}N]ammonium chloride) or triply labeled ([^{15}N]ammonium chloride, [U-^{13}C;1,2,3,4,5,6,6-^{2}H]-D-glucose, and ∼99% deuterium oxide) M9 medium and shaken at 37 °C for 1 h. Protein expression was induced by addition of 1 mM IPTG followed by shaking for 12−18 h at 22−26 °C. The cell paste was frozen at −20 °C. Isotopic labeling at specific amino acid types was accomplished by substituting a defined medium (*27*) (containing either 115 mg of [^{15}N]-L-phenylalanine or 65 mg of [^{15}N]-L-leucine per liter) for M9 medium in the protocol described above. Thiamin, niacin, L-tryptophan, $CaCl_2$, $ZnSO_4$, and $MnSO_4$ were used as described previously with minor differences (*27*).

Cells were thawed and resuspended in a buffer containing 20 mM potassium phosphate (pH 7.0) with 400 mM NaCl, 10 mM imidazole, 0.1% Triton X-100, 200 μM phenylmethanesulfonyl fluoride, and 10 mM β-mercaptoethanol (βMe) added and lysed by sonication. The lysate was clarified by centrifugation at 28000*g* for 60 min, and the supernatant was applied to a Ni-NTA agarose column. The column was washed with a buffer containing 20 mM potassium phosphate, 1.0 M NaCl, and 20 mM imidazole (pH 7.0) with 10 mM βMe. The high NaCl concentration was necessary for the removal of DNA bound to the protein. The protein was eluted with a buffer containing 20 mM potassium phosphate, 200 mM NaCl, and 250 mM imidazole (pH 7.0) with 10 mM βMe. The elutant was applied to a cellulose phosphate P-11 column. This column was washed with 10 mM potassium phosphate, 200 mM NaCl, and 1 mM EDTA (pH 6.8) with 10 mM βMe followed by elution with the same buffer containing 800 mM NaCl.

The purified protein was dialyzed into NMR buffer containing 50 mM sodium phosphate, 100 mM NaCl, 1 mM EDTA, 50 mM L-arginine, 50 mM L-glutamic acid, and 0.75% glycerol (pH 6.5) with 5 mM dithiothreitol. This combination of arginine and glutamic acid enhances the solubility and diminishes aggregation with several proteins (*28*). The protein concentration was estimated using a NanoDrop ND-1000 UV−vis spectrophotometer using an extinction coefficient of 25 000 M^{-1} cm^{-1} at 280 nm calculated by the method of Gill and von Hipple (*29*) as implemented by the website http://www.scripps.edu/∼cdputnam/protcalc.html. The protein−cofactor binary complexes were formed by adding 1.5−2 equiv of AdoHcy. The M.HhaI−AdoHcy complex or free M.HhaI samples were then concentrated to 0.3−0.6 mM by ultrafiltration with a centrifuge using Millipore's Centriprep (YM-10, 10 kDa molecular mass cutoff) or Ultrafree (with a 5 kDa molecular mass cutoff) concentration devices. The final protein concentration was determined by a Bradford assay using uncomplexed M.HhaI as a standard.

NMR Experiments and Data Analysis. All NMR data were collected at 25 °C on a Varian 600 MHz spectrometer equipped with a four-channel (^{1}H, ^{13}C, ^{15}N, and ^{2}H) cryoprobe and *Z*-axis pulsed field gradients. Transverse relaxation-optimized spectroscopy (TROSY) (*30*) with improved sensitivity (*31*, *32*) was used in all experiments, unless otherwise stated. For backbone and sequential assignments, we have performed the following pairs of three-dimensional (3D) TROSY-based experiments: HNCA and HN(CO)CA, HNCACB and HN(CO)CACB (*31*, *33*, *34*), and HNCO and HN(CA)CO (*31*, *33*, *35*). The resulting pairs of spectra allowed sequential connections through CA, CB, or CO resonances and amino acid type identification, leading to resonance assignments for the primary sequence.

Longitudinal and transverse relaxation times T_1 and T_2 (or inverse rates R_1 and R_2) and $\{^{1}H\}^{15}N$ NOE factors of backbone amide ^{15}N nuclei were measured using inverse-detected two-dimensional (2D) experiments (*36*−*39*). The following relaxation delays were used: 55.5, 277.5, 610.0, 943.5, and 1110.0 ms for T_1 measurements and 16.5, 32.9, 49.3, and 65.8 ms for T_2 measurements. Relaxation rates R_1 and R_2 were extracted by fitting the peak intensities with a single-exponential decay function. The $\{^{1}H\}^{15}N$ NOE factor was taken as the ratio of the peak intensities with and without ^{1}H saturation during 3 s of the total 8 s recycle delay period.

Data processing was carried out with the nmrPipe program package (*40*). Data analysis and resonance assignments were made with a modified version of ANSIG (*41*, *42*) running under the Linux operating system.

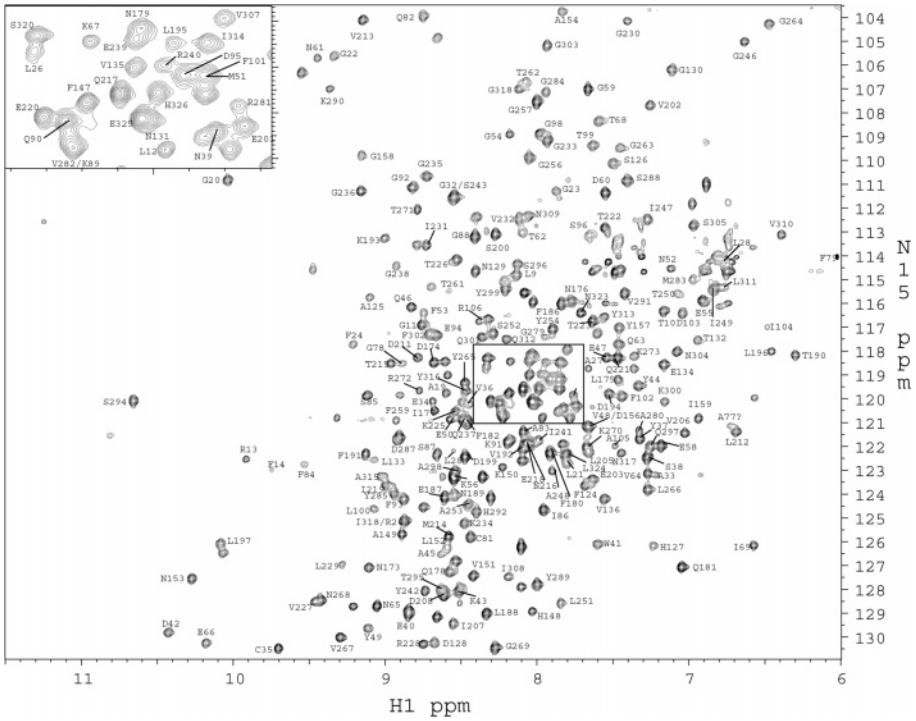

FIGURE 1: $^{1}H-^{15}N$ TROSY-HSQC spectrum of 0.3 mM ^{15}N-, ^{13}C-, and ^{2}H-labeled M.HhaI denoted with backbone amide assignments. The area enclosed by a solid square in the middle of the spectrum is expanded and labeled in the top left corner. Aliased peaks are denoted with asterisks next to the labels. The buffer condition is as follows: 50 mM $NaPO_4$, 100 mM NaCl, 50 mM Arg, 50 mM Glu, 5 mM DTT, 1 mM EDTA, 0.75% glycerol, and 0.5 mM AdoHcy (pH 6.5) in a 92% H_2O/8% D_2O mixture.

RESULTS

Sequential Assignments and Secondary Structure. The deuterated and ^{15}N- and ^{13}C-labeled M.HhaI spectrum exhibits mostly sharp resonances with excellent dispersion (Figure 1). Using a set of 3D TROSY experiments correlating the backbone amide NH with previous and/or intraresidue CA−CB resonances, we reliably assigned 242 (76%) of 313 NH resonances which exclude 12 prolines and 6 histidines at the C-terminus, 260 (79%) CA resonances, and 220 (73%) CB resonances excluding 28 glycines which do not have CB atoms. Several critical regions of the protein have been assigned, including the entire catalytic loop residues, Cys81−Thr99, and two glycine-rich surface loops that make sequence-specific DNA contacts in the major groove, Gly233−Arg240 and Thr250−Gly257, except for Gly255. Assignments were made primarily with uniformly labeled protein saturated with the cofactor AdoHcy and were aided and confirmed by data from [^{15}N]Leu and [^{15}N]Phe specifically labeled protein samples. Approximately 40, many weak, resonances in the HSQC spectrum remain unassigned due to the lack of extended sequential connectivity through CA−CB resonances.

The longest stretches of unassigned residues include Ile15−Leu18, Asp71−Cys76, Ile107−Asn123, Lys137−Ser146, and Glu161−Arg172. Most of the unassigned residues reside in the central β-sheet of the cofactor binding or large domain. The weakening or absence of the resonances from this region may suggest that some of the amide deuterons are not fully exchanged with water protons. Several attempts to exchange any buried amide deuterons for protons through a refolding process were unsuccessful due to aggregation and misfolding problems at micromolar protein concentrations. However, several pieces of evidence suggest that exchange broadening dynamics likely contributes the most to the missing resonances and difficulty in making their assignments. Comparing the $^{1}H-^{15}N$ TROSY-HSQC spectra of the uniformly ^{15}N-, ^{13}C-, and ^{2}H-labeled protein and the [^{15}N]Phe or [^{15}N]Leu specifically labeled protein, we observed 17 of a total of 19 Phe resonances and 24 of a total of 26 Leu resonances; furthermore, the peak intensities vary similarly for most peaks between uniformly (perdeuterated) and amino acid specifically labeled (fully protonated) samples (see Figure 1 of the Supporting Information). Comparing the spectra from uniformly ^{15}N-, ^{13}C-, and ^{2}H-labeled and ^{15}N-labeled protein samples, we have not found a large number of missing resonances from deuterated protein. Thus, although the lack of complete deuteron−proton exchange may have caused a few missing amide resonances, this is most likely due to dynamics and other chemical exchange processes.

In general, the secondary structure of the protein identified by NMR matches well that detected in the crystal structure (Figure 2) (*11*). This conclusion is based on the common method of secondary structure prediction using chemical shift deviations of typical C_α and CO nuclei from values for random coil peptides (*43*−*45*). Consecutive positive C_α shift deviations indicate helical structure, while consecutive negative deviations indicate extended, typically β-strand structure. The NMR data reveal more details of the secondary structure of the active site loop. Notably, this extended loop region (residues 80−99) exhibits negative C_α chemical shift deviations of more than −1.0 ppm from residue 79 to 85, except for Cys81, and small deviations (±0.5 ppm) from residue 86 to 92 indicative of random coil structure, followed by positive deviations from residue 93 to 97 exceeding 1.0 ppm. These data indicate residues 79−85 are formed in an

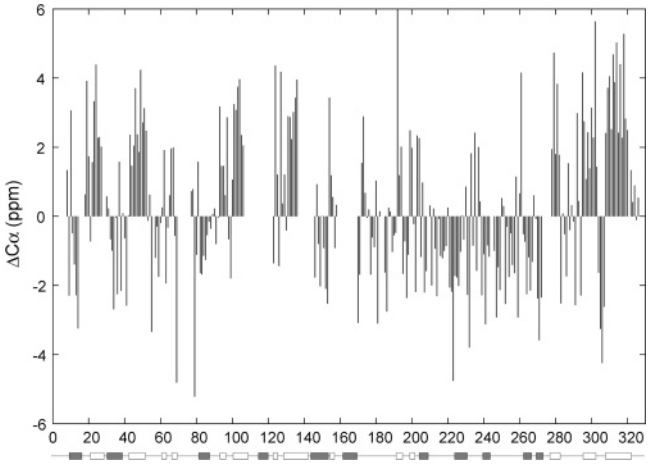

FIGURE 2: C_α chemical shift deviations for M.HhaI from values for random coil peptides. The secondary structure derived from a crystal structure bound with the cofcator AdoMet (PDB entry 1HMY) is shown below the residue numbers with white rectangles representing helices and gray rectangles strands.

extended, possibly β-strand-like structure. The helical character of residues 93−97 was detected in the X-ray structure of the protein−cofactor binary complex (*11*). The moderate magnitude of the C_α shift deviations and the short length of this surface helix suggest that the helical structure may not be well-formed and stable.

Interaction with the Cofactor AdoHcy. Detailed structural changes upon cofactor binding were previously unknown because the available X-ray structures contain either AdoMet or AdoHcy. While the sequential assignments were made with protein samples in the presence of the cofactor to gain more long-term stability and sensitivity, we also collected and compared $^1H-^{15}N$ correlation spectra with and without the cofactor. AdoHcy appears to saturate the protein at a 1:1 equivalent, consistent with a K_d of a few micromolar (*23*, *46*−*48*). The NMR assay of this binding revealed extensive changes in the $^1H-^{15}N$ correlation spectrum (Figure 3). Four closely spaced regions in the sequence, residues 20−31, 41−47, 59−65, and 79−84, were strongly perturbed, followed by a long stretch of sporadic disruptions between residues 270 and 312 (Figure 4). Several weaker patterns of disruption were also observed in Figure 4, including residues 99−102 and 247−253. These interacting residues match those identified in the protein−cofactor complex (*11*). The majority of the residues affected by AdoHcy are located in an area surrounding the cofactor binding site in the large domain, including the interdomain surface (Figure 5). Some of the amide resonances moved significantly so that their bound state peaks cannot be identified in the spectrum by simply comparing the two spectra, particularly in crowded areas.

The active site loop region (residues 80−99) is perturbed by cofactor binding because of its proximity to the binding site, with the largest effects seen at the beginning of the loop where the catalytic Cys81 residue resides. The combined amide peak displacement for Phe84 is approximately 100 Hz, while residues 79−81 are displaced more than 60 Hz, beyond reliable identification in the spectrum. The N-terminus of the loop resides adjacent to the adenosyl moiety of the cofactor, with Phe79 experiencing the biggest impact of the ring current effect upon cofactor binding (*49*).

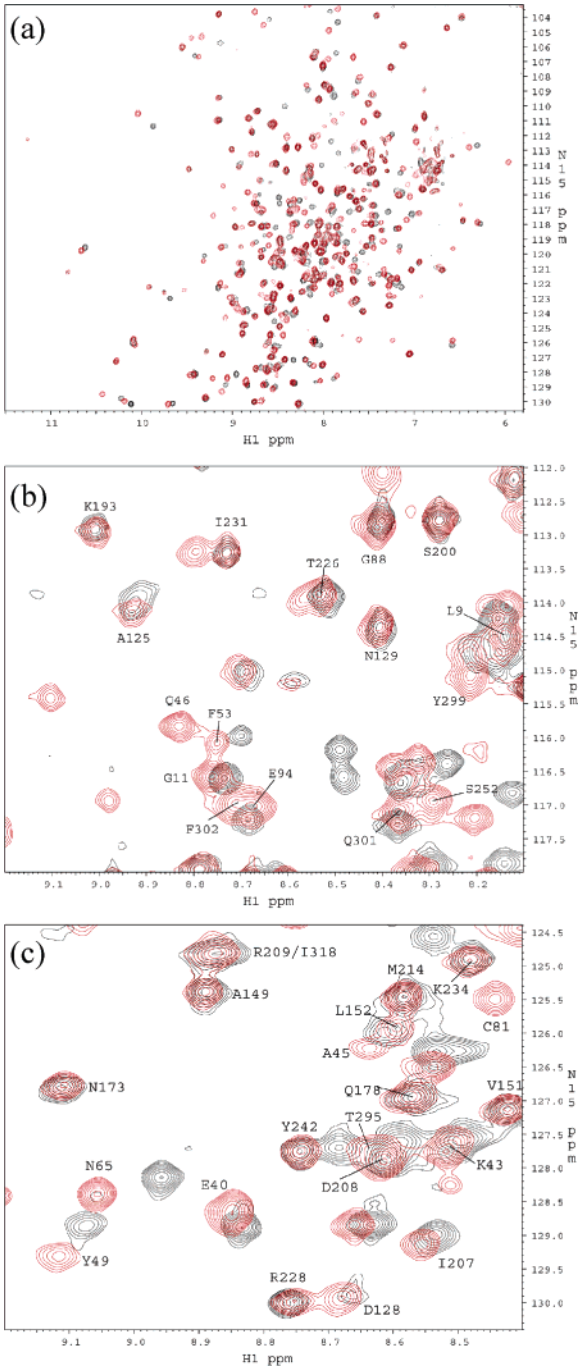

FIGURE 3: Superposition of $^1H-^{15}N$ TROSY-HSQC spectra of HhaI in the presence (red) and absence (black) of the cofactor AdoHcy. Assignment labels are for the spectrum with AdoHcy. A number of new peaks appear or are strengthened in the presence of AdoHcy. Assignment labels are for the spectrum in the presence of the cofactor.

The central region (residues 85−96) of the loop points away from the cofactor and interacts with a surface helix (residues 132−143). The central loop region is also affected by the cofactor, but to a much smaller extent, with peak displacements of <20 Hz. Several residues in the interacting helical region also have peak displacements on the same order of magnitude, with the maximum at 20 Hz. The C-terminal end of the loop folds back toward the cofactor with stronger perturbations compared to the loop center, reaching a maximum of 37 Hz for Leu100, still significantly weaker than the perturbations seen for the N-terminal end of the

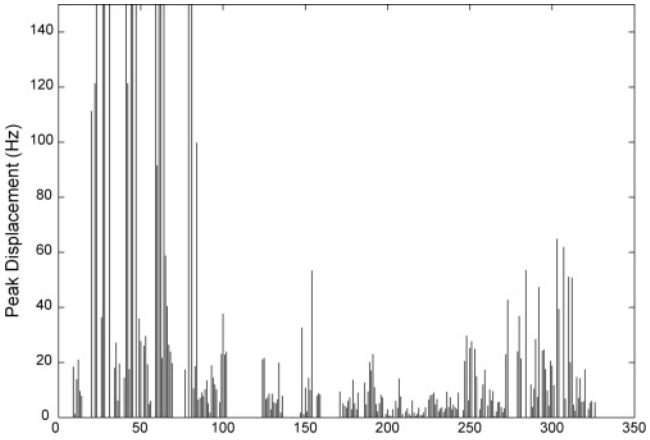

FIGURE 4: Chemical shift perturbation (or peak displacement) upon addition of the cofactor AdoHcy. The total peak displacement is calculated as the combined movement in hertz in the $^{1}H-^{15}N$ correlation spectrum. Residues 22, 23, 27, 28, 31, 41, 44, 45, 47, 59−62, 64, 79, and 81 have shifted significantly but cannot be reliably identified in the spectrum. These residues are assigned an arbitrary 150 Hz displacement in the figure for more complete data presentation.

loop. The pattern of chemical shift perturbations in the loop region by the cofactor indicates the central region of the loop is positioned in a similar environment in the free and cofactor-bound protein.

Interestingly, we observed small but noticeable and systematic chemical shift perturbations on the order of 10−20 Hz to residues remote from the cofactor binding site. These residues include distal sites such as Ile207 and Arg272, approximately 30 Å from the cofactor across the small domain, and Leu9 and Gly11 approximately 25 Å from the cofactor in the cofactor binding domain (Figure 5). Such small, long-range perturbations are beyond experimental uncertainties and can be explained by structural changes at the distal sites induced by the cofactor. The region encompassing the DNA binding cleft is also notably perturbed. The perturbations include areas at the bottom and side of the cleft not in direct contact with the cofactor.

Backbone Dynamics. In the free protein, the resonances in the $^{1}H-^{15}N$ correlation spectrum have an uneven distribution of intensity and line width, indicating the presence of microsecond to millisecond dynamics. Addition of the cofactor to the protein not only causes widespread chemical shift perturbations but also results in sharpening of some resonances and the appearance of ∼10 new peaks (Figure 3). For this reason, our protein sample used in assignment experiments was saturated with AdoHcy so that more resonances could be observed. Notably, the amide resonances from residues 100−106 immediately after the active site loop are either extremely weak or absent in the free protein but are strengthened with good intensity and line width upon addition of the cofactor, indicating the cofactor significantly alters microsecond to millisecond dynamics, reducing exchange broadening (R_{ex}) present in the free protein. Residues 107−112 are not assigned due to the absence of resonances, indicating that the reduction of resonance broadening dynamics is not complete even in the presence of the cofactor.

To learn more about the dynamic nature of the protein, we measured the relaxation parameters of the backbone ^{15}N nuclei in M.HhaI in the presence of the cofactor (Figure 6).

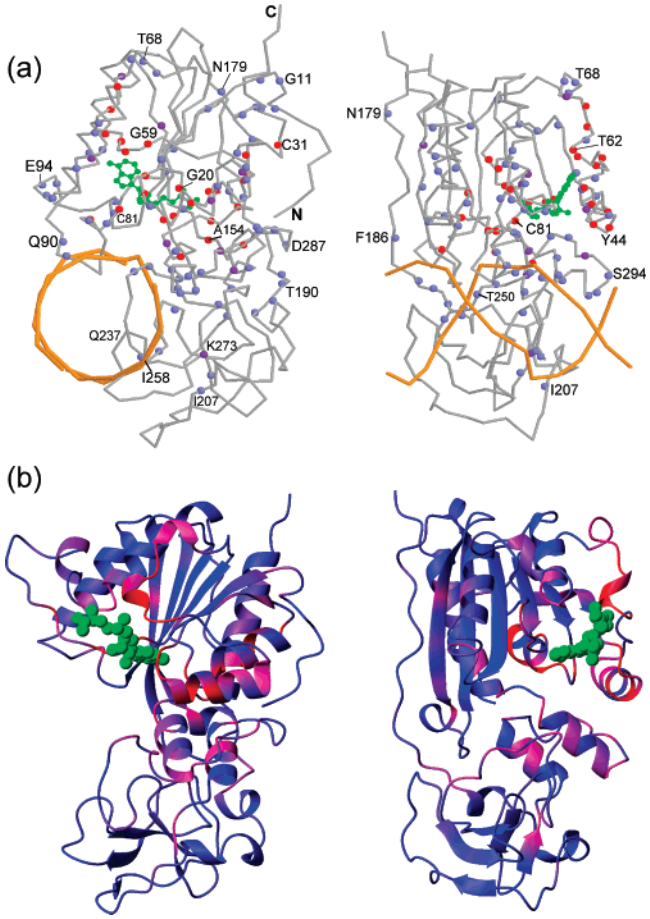

FIGURE 5: Chemical shift perturbations to HhaI with the addition of the cofactor AdoHcy. (a) Chemical shift changes upon binding to AdoHcy are depicted as red, purple, and blue dots at the backbone amide nitrogen positions representing changes between 10 and 30 Hz (blue), 30 and 50 Hz (purple), and >50 Hz (red), respectively. The structure is a ternary complex of the protein with the cofactor (green) and cognate DNA (orange) (PDB entry 2HR1). (b) The peak displacements are schematically shown in a smooth color scheme ranging from the biggest changes (∼120 Hz) in red, the intermediate changes in purple, to the smallest changes in blue. This structure is a binary complex between the protein and the cofactor AdoMet (green) (PDB entry 1HMY). In both panels a and b, the right structures resulted from a 90° right-handed rotation around the vertical axis from the left structures.

The average longitudinal relaxation rate R_1 and transverse relaxation rate R_2 are 0.68 ± 0.12 and 27.8 ± 4.9 s^{-1}, respectively. The average R_2/R_1 value is 42.7 ± 11.9. For macromolecules of this size, fast internal motions on the picosecond to nanosecond time scale usually lead to reduced R_2 and concurrent elevated R_1 values (*36*, *37*, *50*). Such motions are clearly shown by the peaks in the R_1, R_2, and R_2/R_1 data around residue 90. The high flexibility of the central loop region is consistent with their chemical shifts resembling random coil peptides and elevated *B*-factors observed for residues 85−95 in the X-ray structure (PDB entry 1HMY) (*11*). Motions on the microsecond to millisecond time scale, often resulting from conformational exchange leading to line broadening, are reflected in significantly increased R_2 but average R_1 values. These motions are shown by the data near residue 100. Together, the dynamics data indicate that the active site loop region possesses flexibility on the picosecond to nanosecond time scale with increased mobility on the microsecond to mil-

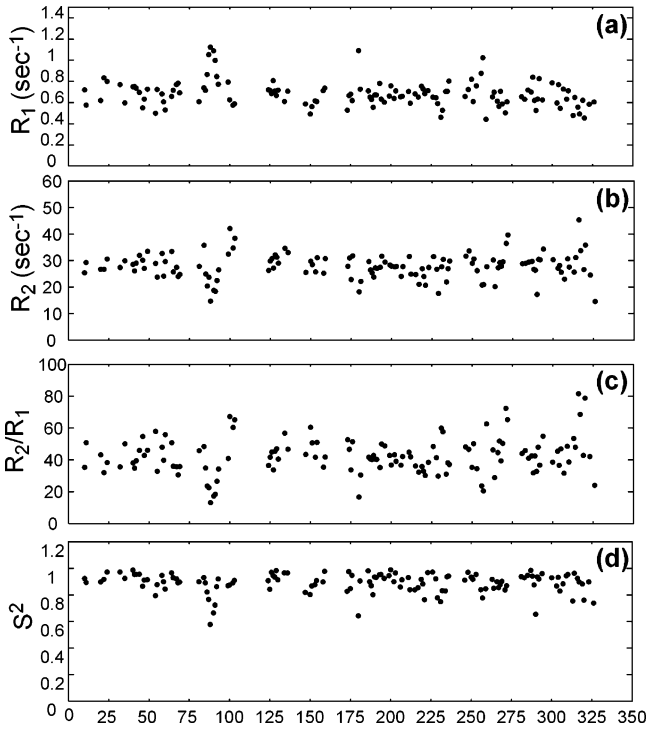

FIGURE 6: Backbone amide ^{15}N relaxation data from M.HhaI as a function of residue number. The longitudinal relaxation rate R_1 (a), the transverse relaxation rate R_2 (b), and {^{1}H}^{15}N NOEs (see the Supporting Information) were measured using a 0.3 mM ^{15}N-, ^{13}C-, and ^{2}H-labeled protein sample in a 92% H_2O/8% D_2O mixture in the presence of an excess of the cofactor AdoHcy. The calculated R_2/R_1 ratio (c) and extracted order parameter S^2 (d) are also shown.

lisecond time scale toward the region of residues 100−125. The microsecond to millisecond mobility observed suggests that severe resonance broadening is likely responsible for the lack of assignments for the region around residue 100.

Further analysis of the dynamics data was performed but was treated qualitatively due to the limited availability of data, significant conformational exchange, and large degree of motional anisotropy. Analysis of the data using the "model-free" approach (*51*, *52*) assuming isotropic overall motions yielded an overall correlation time of 20.5 ns, with a pronounced dip in order parameters in the active site loop region (Figure 6d), further evidence that the central active site loop region is flexible on the picosecond to nanosecond time scale. Analysis was also carried out with a contribution from the conformational exchange term (R_{ex}) to the transverse relaxation rate R_2 (*39*) as well as using an axially symmetric tumbling model (*50*, *53*). Although these approaches yielded general qualitative conclusions that are the same as those discussed above, more detailed, quantitative analysis of the dynamics data incorporating fast and intermediate time scale dynamics and the distinct nonspherical shape for this elongated protein would require data collected at multiple fields and/or independent measurement of the microsecond to millisecond motional terms. It should be noted that analysis using an anisotrpic diffusion model based on the crystal structure does not change our qualitative conclusions concerning the dynamic properties of the active site loop.

DISCUSSION

The sequence-specific DNA cytosine C5 methyltransferase M.HhaI has been under intensive study in an attempt to understand the basis of its catalytic and specificity mechanisms. However, to date studies using high-resolution NMR techniques have been limited (*21*, *25*), partly hampered by its relatively large size (~37 kDa) and limited wild-type solubility (*25*). Recent developments in TROSY techniques (*30*) and cryogenic probes have significantly lessened the difficulty in studying such systems. Additionally, solubility at NMR concentrations (several hundred micromolar to millimolar) was improved via removal of the C-terminal hydrophobic tail (*25*). Following these developments, we have successfully engineered a wild-type variant of M.HhaI with increased solubility and unaltered activity (*25*, *54*), allowing us to assign ~80% of the residues. Crucial residues involved in enzyme activity, including the active site loop, were almost completely assigned, and structural perturbations with the cofactor AdoHcy, previously unknown, were evaluated.

M.HhaI Is an Allosteric Protein. The cofactor induced widespread structural changes centered at the binding site, but reaching areas 30 Å distant. The active site loop region was affected, but the relatively small changes in the central region of the loop are consistent with the loop conformation and dynamics being largely intact. Although the overall magnitudes of these changes are smaller than those caused by DNA binding as determined by NMR (data not shown), the allosteric effects are similar. Examination of the DNA-bound and DNA-free protein structures shows that active site loop closing is accompanied by a concerted reorientation of the two domains around the narrow domain junction toward each other. The ability of the cofactor to induce changes in distal elements of both the small and large domains (the bottom surface of the small domain and the top surface of the large domain in Figure 6) indicates that M.HhaI acts through an allosteric mechanism involving a delicate set of internally coupled interactions adapted to respond to external perturbations. The extensive perturbations seen at the domain boundary are good indications that the cofactor is capable of transducing signals between the domains and may affect domain structure and orientation, likely toward a state that is more capable of DNA binding, base flipping, and selectivity. Indeed, an area making part of the DNA binding cleft, including Ser251−Lys261, is perturbed by the cofactor with perturbations of 20−30 Hz around Thr250 and 10−20 Hz around I258. This area inserts into the DNA duplex and is not in direct contact with the cofactor. These changes appear to propagate through a helix (Ser296−Asn304) proximal to the cofactor and connecting the small and large domains.

Changes around an extensively studied residue, Thr250, hint at the effects of the cofactor on methylation activity and DNA selectivity. Thr250 is a highly conserved residue which interacts with the 5′-phosphate of the flipped-out cytosine and appears to stabilize the flipped-out form (*54*). Residues around Thr250 were clearly perturbed by cofactor binding, although the perturbations are smaller relative to the major areas that are in direct contact with the cofactor (Figure 5). The backbone amide of Thr250 is approximately 15 Å from the nearest backbone position of the cofactor, and there is no direct contact through space between this area and the cofactor. The subtle changes caused by cofactor binding are consistent with the requirement for the cofactor in stabilizing the DNA-bound, base flipped-out state (*21*)

and in the dramatic enhancement in selectivity with the addition of the cofactor (*22*, *54*).

A less conserved area around Gln237 is not perturbed by the cofactor. Like Thr250, this residue has no direct contact with the cofactor, and the distance between its backbone nitrogen and the cofactor is approximately 21 Å. Gln237, along with Ser87, occupies the space in the DNA duplex left by the flipped-out cytosine and forms a hydrogen bond to Ser87 and the orphan guanosine (Figure 6). While mutations of Gln237 to 19 other amino acids were found to affect the stability of the protein−DNA−cofactor complex and severely perturb base flipping, they did not affect target sequence specificity (*55*, *56*). Our data indicate that this region is not responsive to the cofactor binding either, arguing that it is not used by the cofactor to regulate interactions with DNA.

M.HhaI Is a Dynamic Protein. Ligand-induced changes in areas remote from the ligand-binding site are characteristics of allosteric proteins, and they are often closely linked to the dynamic nature of the protein (*57*−*60*). In M.HhaI, ligand binding has profound effects on both the structure and dynamics of the protein. Cofactor binding significantly altered dynamics, indicated by the reduction in the peak line width and the appearance of new resonances in the $^{1}H-^{15}N$ correlation spectrum. The ^{15}N relaxation data also indicate that the active site loop possesses a significant amount of dynamics both on the picosecond to nanosecond time scale and on the microsecond to millisecond time scale.

Similar ligand-induced changes have been observed in other allosteric proteins. In the phosphorylation-activated signaling protein NtrC, a strong correlation between microsecond to millisecond time scale motions and activation was detected, and activation through phosphorylation shifts a pre-existing equilibrium between inactive and active conformations toward the active state (*61*). In the case of ATP-binding cassette (ABC) transporters, ADP−Mg binding not only perturbs the structure but also alters conformational dynamics on the microsecond to millisecond time scale to areas 30 Å from the binding site, while picosecond to nanosecond dynamics are largely unaffected (*62*). Motions on the microsecond to millisecond time scale were detected in the apoprotein and their magnitudes reduced by ligand binding, and multiple conformations were observed in X-ray crystal structures for the ligand-binding loop in the apoprotein. Activation or ligand binding in both the NtrC and ABC proteins fits well with a "selected-fit" model. In this model, it is assumed that there is a pre-existing equilibrium of conformations in the apo form and an activation or binding event drives the equilibrium toward the active or bound state. While structural data on pre-existing multiple conformations of M.HhaI in the absence of DNA are sparse, it is tempting to speculate that the elevated dynamics in and near the active site loop and distal structural changes induced by cofactor binding, including areas affecting DNA binding and domain interaction, may be correlated and relevant to the cofactor and/or DNA binding mechanism.

SUPPORTING INFORMATION AVAILABLE

One figure comparing the $^{1}H-^{15}N$ correlation spectra from uniformly ^{15}N- and ^{13}C-labeled and perdeuterated, and [^{15}N]-Phe or [^{15}N]Leu specifically labeled, fully protonated M.HhaI samples; one table listing measured longitudinal and transverse relaxation rates R_1 and R_2, respectively, $\{^{1}H\}^{15}N$ NOE factors, calculated order parameters (S^2), and correlation times for internal motions (τ_e), and the corresponding estimated uncertainties. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Jeltsch, A. (2002) Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases, *ChemBioChem 3*, 274−293.
2. Heithoff, D. M., Sinsheimer, R. L., Low, D. A., and Mahan, M. J. (1999) An essential role for DNA adenine methylation in bacterial virulence, *Science 284*, 967−970.
3. Julio, S. M., Heithoff, D. M., Sinsheimer, R. L., Low, D. A., and Mahan, M. J. (2002) DNA adenine methylase overproduction in *Yersinia pseudotuberculosis* alters YopE expression and secretion and host immune responses to infection, *Infect. Immun. 70*, 1006−1009.
4. Low, D. A., Weyand, N. J., and Mahan, M. J. (2001) Roles of DNA adenine methylation in regulating bacterial gene expression and virulence, *Infect. Immun. 69*, 7197−7204.
5. Egger, G., Liang, G., Aparicio, A., and Jones, P. A. (2004) Epigenetics in human disease and prospects for epigenetic therapy, *Nature 429*, 457−463.
6. Jones, P. A. (2003) Epigenetics in carcinogenesis and cancer prevention, *Ann. N.Y. Acad. Sci. 983*, 213−219.
7. Jones, P. A. (2002) DNA methylation and cancer, *Oncogene 21*, 5358−5360.
8. Jones, P. A., and Takai, D. (2001) The role of DNA methylation in mammalian epigenetics, *Science 293*, 1068−1070.
9. Cheng, X., and Roberts, R. J. (2001) AdoMet-dependent methylation, DNA methyltransferases and base flipping, *Nucleic Acids Res. 29*, 3784−3795.
10. Roberts, R. J., and Cheng, X. (1998) Base flipping, *Annu. Rev. Biochem. 67*, 181−198.
11. Cheng, X., Kumar, S., Posfai, J., Pflugrath, J. W., and Roberts, R. J. (1993) Crystal structure of the HhaI DNA methyltransferase complexed with S-adenosyl-L-methionine, *Cell 74*, 299−307.
12. 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 HhaI DNA methyltransferase, *J. Biol. Chem. 275*, 38722−38730.
13. Sheikhnejad, G., Brank, A., Christman, J. K., Goddard, A., Alvarez, E., Ford, H., Jr., Marquez, V. E., Marasco, C. J., Sufrin, J. R., O'Gara, M., and Cheng, X. (1999) Mechanism of inhibition of DNA (cytosine C5)-methyltransferases by oligodeoxyribonucleotides containing 5,6-dihydro-5-azacytosine, *J. Mol. Biol. 285*, 2021−2034.
14. 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.
15. 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.
16. O'Gara, M., Roberts, R. J., and Cheng, X. (1996) A structural basis for the preferential binding of hemimethylated DNA by HhaI DNA methyltransferase, *J. Mol. Biol. 263*, 597−606.
17. O'Gara, M., Klimasauskas, S., Roberts, R. J., and Cheng, X. (1996) Enzymatic C5-cytosine methylation of DNA: Mechanistic implications of new crystal structures for HhaL methyltransferase-DNA-AdoHcy complexes, *J. Mol. Biol. 261*, 634−645.
18. 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.
19. Shieh, F. K., Youngblood, B., and Reich, N. O. (2006) The role of Arg165 towards base flipping, base stabilization and catalysis in M.HhaI, *J. Mol. Biol. 362*, 516−527.
20. Youngblood, B., Shieh, F. K., De Los, R. S., Perona, J. J., and Reich, N. O. (2006) Engineered extrahelical base destabilization enhances sequence discrimination of DNA methyltransferase M.HhaI, *J. Mol. Biol. 362*, 334−346.

21. Klimasauskas, S., Szyperski, T., Serva, S., and Wuthrich, K. (1998) Dynamic modes of the flipped-out cytosine during HhaI methyltransferase-DNA interactions in solution, *EMBO J. 17*, 317−324.
22. 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.
23. 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.
24. Merkiene, E., and Klimasauskas, S. (2005) Probing a rate-limiting step by mutational perturbation of AdoMet binding in the HhaI methyltransferase, *Nucleic Acids Res. 33*, 307−315.
25. Daujotyte, D., Vilkaitis, G., Manelyte, L., Skalicky, J., Szyperski, T., and Klimasauskas, S. (2003) Solubility engineering of the HhaI methyltransferases, *Protein Eng. 16*, 295−301.
26. Marley, J., Lu, M., and Bracken, C. (2001) A method for efficient isotopic labeling of recombinant proteins, *J. Biomol. NMR 20*, 71−75.
27. Muchmore, D. C., McIntosh, L. P., Russell, C. B., Anderson, D. E., and Dahlquist, F. W. (1989) Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance, *Methods Enzymol. 177*, 44−73.
28. Golovanov, A. P., Hautbergue, G. M., Wilson, S. A., and Lian, L. Y. (2004) A simple method for improving protein solubility and long-term stability, *J. Am. Chem. Soc. 126*, 8933−8939.
29. Gill, S. C., and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data, *Anal. Biochem. 182*, 319−326.
30. Pervushin, K., Riek, R., Wider, G., and Wuthrich, K. (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution, *Proc. Natl. Acad. Sci. U.S.A. 94*, 12366−12371.
31. Yang, D. W., and Kay, L. E. (1999) Improved lineshape and sensitivity in the HNCO-family of triple resonance experiments, *J. Biomol. NMR 14*, 273−276.
32. Rance, M., Loria, J. P., and Palmer, A. G., III (1999) Sensitivity improvement of transverse relaxation-optimized spectroscopy, *J. Magn. Reson. 136*, 92−101.
33. Salzmann, M., Pervushin, K., Wider, G., Senn, H., and Wuthrich, K. (1998) TROSY in triple-resonance experiments: New perspectives for sequential NMR assignment of large proteins, *Proc. Natl. Acad. Sci. U.S.A. 95*, 13585−13590.
34. Yamazaki, T., Lee, W., Arrowsmith, C. H., Muhandiram, D. R., and Kay, L. E. (1994) A suite of triple-resonance NMR experiments for the backbone assignment of ^{15}N, ^{13}C, ^{2}H labeled proteins with high-sensitivity, *J. Am. Chem. Soc. 116*, 11655−11666.
35. Kay, L. E., Ikura, M., Tschudin, R., and Bax, A. (1990) Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins, *J. Magn. Reson. 89*, 496−514.
36. Kay, L. E., Torchia, D. A., and Bax, A. (1989) Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: Application to staphylococcal nuclease, *Biochemistry 28*, 8972−8979.
37. Clore, G. M., Driscoll, P. C., Wingfield, P. T., and Gronenborn, A. M. (1990) Analysis of the backbone dynamics of interleukin-1β using two-dimensional inverse detected heteronuclear ^{15}N-^{1}H NMR spectroscopy, *Biochemistry 29*, 7387−7401.
38. Palmer, A. G., III, Rance, M., and Wright, P. E. (1991) Intramolecular motions of a zinc finger DNA-binding domain from Xfin characterized by proton-detected natural abundance carbon-13 heteronuclear NMR spectroscopy, *J. Am. Chem. Soc. 113*, 4371−4380.
39. Mandel, A. M., Akke, M., and Palmer, A. G., III (1995) Backbone dynamics of *Escherichia coli* ribonuclease HI: Correlations with structure and function in an active enzyme, *J. Mol. Biol. 246*, 144−163.
40. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR 6*, 277−293.
41. Kraulis, P. J. (1989) ANSIG: A program for the assignment of protein ^{1}H 2D NMR spectra by interactive graphics, *J. Magn. Reson. 24*, 627−633.
42. Kraulis, P. J., Domaille, P. J., Campbell-Burk, S. L., van Aken, T., and Laue, E. D. (1994) Solution structure and dynamics of Ras p21.GDP determined by heteronuclear three- and four-dimensional NMR spectroscopy, *Biochemistry 33*, 3515−3531.
43. Wishart, D. S., and Sykes, B. D. (1994) The ^{13}C chemical-shift index: A simple method for the identification of protein secondary structure using ^{13}C chemical-shift data, *J. Biomol. NMR 4*, 171−180.
44. Wishart, D. S., Richards, F. M., and Sykes, B. D. (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure, *J. Mol. Biol. 222*, 311−333.
45. Spera, S., and Bax, A. (1991) Emperical correlation between protein backbone conformation and C_α and C_β ^{13}C nuclear magnetic resonance chemical shifts, *J. Am. Chem. Soc. 113*, 5490−5492.
46. Swaminathan, C. P., Sankpal, U. T., Rao, D. N., and Surolia, A. (2002) Water-assisted dual mode cofactor recognition by HhaI DNA methyltransferase, *J. Biol. Chem. 277*, 4042−4049.
47. Vilkaitis, G., Merkiene, E., Serva, S., Weinhold, E., and Klimasauskas, S. (2001) The mechanism of DNA cytosine-5 methylation. Kinetic and mutational dissection of Hhai methyltransferase, *J. Biol. Chem. 276*, 20924−20934.
48. Sankpal, U. T., and Rao, D. N. (2002) Mutational analysis of conserved residues in HhaI DNA methyltransferase, *Nucleic Acids Res. 30*, 2628−2638.
49. Wuthrich, K. (1986) *NMR of proteins and nucleic acids*, Wiley-Interscience, New York.
50. Barbato, G., Ikura, M., Kay, L. E., Pastor, R. W., and Bax, A. (1992) Backbone dynamics of calmodulin studied by ^{15}N relaxation using inverse detected two-dimensional NMR spectroscopy: The central helix is flexible, *Biochemistry 31*, 5269−5278.
51. Lipari, G., and Szabo, A. (1982) Model-Free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules. 1. Theory and range of validity, *J. Am. Chem. Soc. 104*, 4546−4559.
52. Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules. 2. Analysis of experimental results, *J. Am. Chem. Soc. 104*, 4559−4570.
53. Tjandra, N., Wingfield, P., Stahl, S., and Bax, A. (1996) Anisotrpic rotational diffusion of perdeuterated HIV protease from ^{15}N NMR relaxation measurements at two magnetic fields, *J. Biomol. NMR 8*, 272−284.
54. 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 HhaI DNA methyltransferase, *J. Biol. Chem. 276*, 38722−38730.
55. Mi, S., Alonso, D., and Roberts, R. J. (1995) Functional analysis of Gln-237 mutants of HhaI methyltransferase, *Nucleic Acids Res. 23*, 620−627.
56. Daujotyte, D., Serva, S., Vilkaitis, G., Merkiene, E., Venclovas, C., and Klimasauskas, S. (2004) HhaI DNA methyltransferase uses the protruding Gln237 for active flipping of its target cytosine, *Structure 12*, 1047−1055.
57. Jardetzky, O. (1996) Protein dynamics and conformational transitions in allosteric proteins, *Prog. Biophys. Mol. Biol. 65*, 171−219.
58. Kern, D., and Zuiderweg, E. R. P. (2003) The role of dynamics in allosteric regulation, *Curr. Opin. Struct. Biol. 13*, 748−757.
59. Gunasekaran, K., Ma, B., and Nussinov, R. (2004) Is allostery an intrinsic property of all dynamic proteins? *Proteins 57*, 433−443.
60. Popovych, N., Sun, S., Ebright, R. H., and Kalodimos, C. G. (2006) Dynamically driven protein allostery, *Nat. Struct. Mol. Biol. 13*, 831−838.
61. Volkman, B. F., Lipson, D., Wemmer, D. E., and Kern, D. (2001) Two-state allosteric behavior in a single-domain signaling protein, *Science 291*, 2429−2433.
62. Wang, C., Karpowich, N., Hunt, J. F., Rance, M., and Palmer, A. G. (2004) Dynamics of ATP-binding cassette contribute to allosteric control, nucleotide binding and energy transduction in ABC transporters, *J. Mol. Biol. 342*, 525−537.

BI602662E