Spontaneous base flipping in DNA and its possible role in methyltransferase binding

Y. Z. Chen,[1,2] V. Mohan,[1] and R. H. Griffey[1]
[1]*ISIS Pharmaceuticals, 2292 Faraday Avenue, Carlsbad, California 92008*
[2]*Department of Computational Science, National University of Singapore, Singapore 119260*
(Received 15 February 2000)

Recent crystallographic studies showed that *HhaI* and other methyltransferases flip their target DNA base completely out of a DNA helix. This base flipping is also a key feature in a number of other enzyme-catalyzed processes involving DNA. The mechanism of base flipping by these enzymes remains elusive. Based on a full atomic level description of bond rotational motions we have studied the energetics of flipping a base in a B-DNA duplex in the absence of the enzyme. We have also investigated the effect of the restraints from enzyme-distorted DNA backbone on the movement of a flipped base in several methytransferase bound DNA crystal structures. Our study on crystal B-DNA helices showed that a base could be flipped at an energy cost close to the enthalpy observed for base pair opening in premelting thermal fluctuations. This suggests that spontaneous base flipping in DNA due to thermal fluctuation may be achieved. Analysis of several crystal *HhaI* and *HaeIII* methyltransferase DNA duplex structures showed that the enzyme induced DNA backbone distortion severely restricts the movement of the flipped base, which indicates that during base flipping the backbone needs to adopt a substantially different conformation than that observed in the x-ray (enzyme-bound) structures. Our results suggest the possible role of thermally induced transient base opening in facilitating recognition and binding of methyltransferases and other enzymes.

PACS number(s): 87.14.Gg, 87.10.+e, 87.15.By

I. INTRODUCTION

Nature elegantly manifests the sequence specific molecular recognition process via the formation of double stranded nucleic acids. The unique power of base pairing enables molecular events such as replication and transcription. Polymerase enzymes open nucleic acid helices as a prerequisite for their catalytic action. Similarly repair enzymes need to recognize and correct mismatched or damaged bases. In principle, the process of nucleobase flipping (rotation by 180° of the base from the initial ''closed'' helix state) could represent the first step in these key fundamental biological processes. The atomic level detail of how this enzyme-nucleic acid recognition is accomplished is not well understood and deserves further investigation. Determination of the precise mechanism of base flipping is important in understanding such processes as enzyme-catalyzed DNA methylation, repair, mismatch recognition, and initiation of transcription and replication [1–3]. In addition, such studies would form the basis for elucidation of general principles that govern the protein-DNA interactions and will also offer insight for drug development approaches targeting nucleic acids.

The structural aspect of base flipping has been probed by x-ray crystallographic studies of cytosine-5 *HhaI* methyltransferase (*HhaI* MTase) complexed with its DNA substrate [1,4–6]. The target cytosine was found to flip completely out of the helix and into the catalytic site of the binding enzyme without seriously disturbing the rest of the DNA helix. Details of how this base is trapped outside the helix and the role of the binding enzyme is unclear. The base flipping may be induced or facilitated by enzyme binding (for instance, through deformation of DNA backbone or introduction of specific enzyme-DNA interactions). Alternatively, the enzyme may recognize and trap a single base transiently flipped open from the helix [7].

The estimated rate constant of methyltransferase reaction is 0.02/s [8], while the measured base pair lifetime of DNA duplex is 10 ms [9]. Therefore an active involvement of an enzyme in accelerating base flipping appears to be unnecessary. This is confirmed by nuclear magnetic resonance (NMR) study of the dynamic modes of base flipping during *HhaI* MTase-DNA interactions in solution [10]. This NMR study also showed that no trapping of a target base by *HhaI* MTase occurs in the absence of cofactor. Given that no significant acceleration in base pair opening rate was detected, the interpretation of the findings from this experiment is difficult. While a more active role of cofactor as well as the enzyme is indicated, it does not rule out the possibility that transient base flipping is involved in the process.

Some insight into the mechanism of base flipping may be obtained by investigating the feasibility of base flipping in a B-DNA helix in the absence of an enzyme. If a base can be shown to flip at an energy cost comparable to observed enthalpy for premelting thermal fluctuational base pair opening, it may indicate that enzyme binding may not be a prerequisite for base flipping. On the other hand, transient flipping that occurs spontaneously due to thermal fluctuation may play a role in enzyme-base recognition and enzyme binding. In the present paper, the pathway and energetics of base flipping in a B-DNA crystal structure has been examined to determine whether it is possible to flip a B-DNA base in the absence of an enzyme, and whether this flipping can be induced by thermal fluctuation. The flipped base in *HhaI* MTase-DNA complex has been found to orient towards the minor groove [4–6]. This has been used as one of the evidences to support the hypothesis of enzyme induced minor groove pathway. The orientation of the flipped bases has been compared to that in the crystal structure of *HhaI* MTase-DNA complex to see whether they are similar to the observed.

Comparison of the enzyme-DNA structure with the free enzyme crystal structure [4,10] indicate that *HhaI* MTase binding is initiated and maintained through sequence specific contacts in the major groove. While the major groove is completely blocked by these contacts, sufficient room is available on the minor groove side to allow base movement before a large conformation change in a loop of the binding enzyme closes the minor groove. Based on this and other observations, a minor groove pathway had been proposed for base flipping [4]. However, modeling studies [11–13] indicate that the base pair opens primarily through the major groove. This fact has been used as an argument to support the hypothesis that an active participation of enzyme is needed to create a pathway for base flipping through the minor groove.

Given that motion along a pathway is reversible (principle of microscopic reversibility), the existence of an enzyme induced flipping pathway may be probed. Such a study may shed light on whether the restraint of an enzyme distorted DNA backbone is moderate enough to allow a flipped base to return to the closed state through the minor groove. Our earlier investigations on low-energy single base opening in DNA indicated that base movement is controlled by backbone restraint as well as energy barrier along the grooves [12,13]. Backbone restraint sets the upper limit of base displacement. In particular, when the orientation of the O_3'-P bond becomes perpendicular to the helix axis, the magnitude of base movement has been found to be diminished and the base is locked in its original position [13]. In the present paper the upper limit of displacement of a flipped base in MTase-DNA crystal structures have been determined and the implication of our results on enzyme induced flipping pathway will be discussed.

II. METHOD

Base flipping in crystal B-DNA structures and the movement of a flipped base in crystal enzyme-DNA structures are studied by using a method developed earlier to probe the pathway and energetics of low-energy single base opening in DNA [12,13]. In this approach, base opening is accomplished by means of simultaneous rotation of a minimum number of rotatable bonds in the local backbone and glycosidic bond of the base under study, while the rest of the helix is held rigid. Bond rotation is assumed to only involve torsion angle variation while the length and angle of all the bonds are kept fixed. It has been established that ζ is the principal productive torsion angle in driving the base to open out of the helix stack. It has also been shown that a single base can be opened to the observed average opening extent by simultaneously rotating only five backbone as well as the glycosidic bond torsion angle [12,13].

The flipping of a base involves angular displacement of 180° [1,4], which is significantly larger than the observed average displacement of 20°–30° for premelting thermal fluctuational base pair opening [14]. Therefore, for base flipping and inverse flipping, more rotatable bonds than those for low-energy single base opening are expected to be involved. A structural comparison between flipped and unflipped crystal B-DNA helix showed that the displacement of the adjacent nucleotides to the flipped base is relatively small [1,4]. Our own analysis indicates that it is possible to flip a base by rotation of eight torsion angles in the local backbone as well as the glycosidic torsion angle without displacement of its neighboring nucleotides. For the nth base, these eight backbone torsion angles are located between C4′ of $(n-1)$th nucleotide and O3′ of $(n+1)$th nucleotide and they are represented by a Greek letter on each bond:

$$\mathrm{C3'}(n-1)\text{-}\mathrm{C4'}\overset{\gamma}{\text{—}}\mathrm{C5'}\overset{\beta}{\text{—}}\mathrm{O5'}\overset{\alpha}{\text{—}}\mathrm{P}(n)\overset{\zeta}{\text{—}}\mathrm{O3'}\overset{\varepsilon}{\text{—}}\mathrm{C3'}\text{-}$$

$$\mathrm{C4'}\overset{\gamma}{\text{—}}\mathrm{C5'}\overset{\beta}{\text{—}}\mathrm{O5'}\overset{\alpha}{\text{—}}\mathrm{P}(n+1)\text{-}\mathrm{O3'}.$$

Here the δ torsion angle (O3′-C3′-C4′-C5′) is kept unchanged so as to maintain the sugar pucker of the flipped nucleotide in the original value. This ensures that the overall conformation is in the B-form for B-DNA crystal structures and that in the enzyme-deformed conformation for the enzyme-DNA crystal structures.

In the present paper, in order to investigate the base flipping and closing process, ζ has been employed as the principal torsion promoting the motion. In addition a geometric constraint (which mimics helix-restoring forces) has been used to determine the change or response of the rest of torsion angles [12]. This geometric constraint ensures that the variation of these torsion angles only changes local backbone conformation. The trajectory of these torsion angles is determined by a change of ζ followed by the variation of the other seven local backbone torsion angles. This is achieved by a grid search in the relevant torsion angle space, such that the $(n-1)$C3′ end is fixed and the displacement at the $(n+1)$O3′ end is kept at a minimum. A maximum allowed displacement of 0.1 Å is imposed, which defines the upper limit of base flipping. This upper limit together with the energy barrier along the grooves determines the maximum extent of base flipping [12]. In addition to the eight backbone angles, the glycosidic bond torsion angle (O4′-C1′-N1-C2 for cytosine) is also varied to keep the flipped base as much in plane as possible.

The energy barrier of base flipping or closing can be estimated by the following empirical energy forms:

$$\begin{aligned} V = {} & \Sigma_{\text{torsions}} 1/2 V_n [1+\cos(n\phi-\gamma')] \\ & + \Sigma_{\text{H bonds}} [V_0 (1-e^{-a(r-r_0)})^2 - V_0] \\ & + \Sigma_{\text{nonbonded}} [A_{ij}/r_{ij}^{12} - B_{ij}/r_{ij}^{6} + q_i q_j/\varepsilon_r r_{ij}], \end{aligned} \quad (1)$$

where ϕ denotes a torsion angle, V_n, n, and γ' are torsion potential parameters, r is the H-bond donor-acceptor distance, V_0, a, and r_0 are H-bond potential parameters, A_{ij} and B_{ij} are nonbonded van der Waals parameters, ε_r is the dielectric constant, q_i and q_{ij} are the partial charges of the ith and jth atoms, and r_{ij} is the distance between them.

Except for the H-bond terms, the potentials and their parameters are taken from the assisted-model-building-with-energy- refinement force field [15]. In order to circumvent the difficulty in solving for the dynamics of hydrogen atoms along the flipping pathway, an implicit hydrogen atom Morse potential [16] is used for H-bond energy terms. This potential has been shown to give reasonable DNA interbase hydrogen bond energy and hydrogen bond breaking probabilities [17,18]. Bond stretch and angle bending terms are

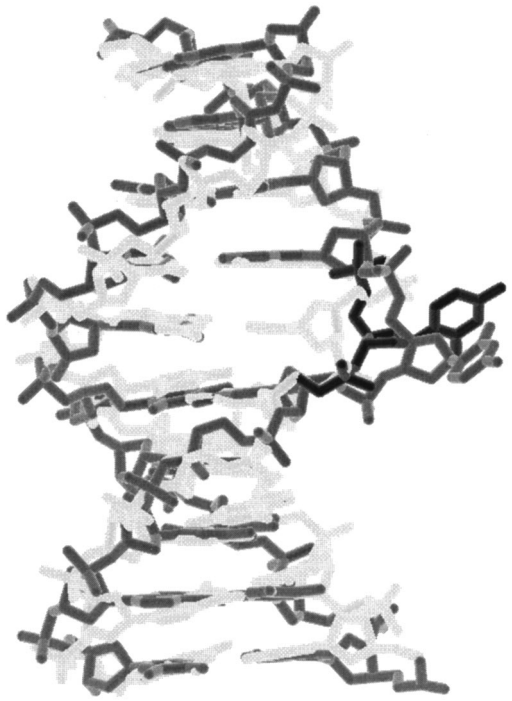

FIG. 1. Comparison of the crystal structure of B-DNA $d(CCGGCGCCGG)_2$ (with the bases paired and intact in the closed form shown in light gray), with the C5 base flipped using the present method (black), and the DNA portion of $d(GATAGCGCTATC)_2$ complexed with *HhaI* methyltransferase in which the C18 base is flipped out (gray).

excluded in the potential-energy function because they do not contribute to the lower-energy base flipping process. The base stacking interactions are implicitly included in the nonbonded van der Waals and electrostatic energy terms. A distance-dependent dielectric constant [19] is used. As the contribution of electrostatic terms to the total energy of base flipping is relatively small, our results are relatively insensitive to the choice of dielectric constant.

III. RESULTS AND DISCUSSIONS

A. Base flipping in B-DNA duplex

Figure 1 shows a crystal B-DNA $d(CCGGCGCCGG)_2$ [20] structure with its C5 base fully flipped out by means of the method described above. Similar flipping behavior has been found for other bases in this and other crystal B-DNA structures [21] selected from the nucleic acids database [22]. For direct comparison of the structural features of the flipped base (obtained using the present approach) with the x-ray structure of *HhaI* MTase bound DNA, Fig. 1 includes the DNA portion of the crystal structure of $d(GATAGCGCTATC)_2$ complexed with *HhaI* MTase [4] in which the C18 base is flipped. In addition, the crystal B-DNA $d(CCGGCGCCGG)_2$ with unflipped C5 base [20] is also included. The *GCGC* central portions of the sequences of the three structures are superimposed. The configuration or orientation of the flipped base in B-DNA is similar to that found in the enzyme bound system. However, the local backbone of the former is not as fully stretched out as later. It is noted that the local backbone of the strand

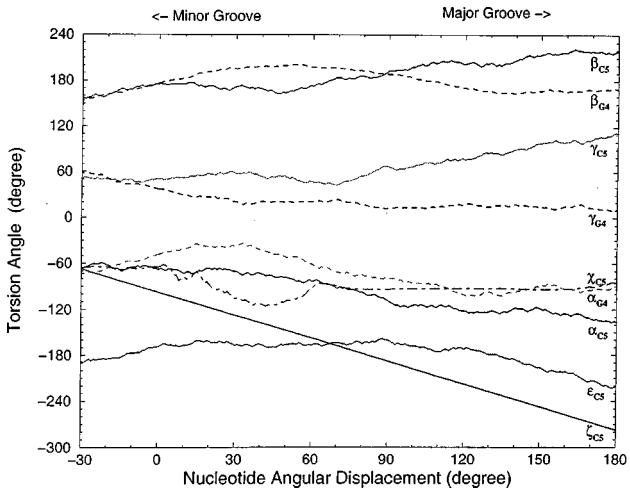

FIG. 2. Variation of rotatable torsion angles for flipping the C5 base in B-DNA $d(CCGGCGCCGG)_2$.

opposite to the flipped base in the enzyme-bound DNA is also stretched out as compared to that in the crystal B-DNA structure. Moreover the sugar pucker of the flipped base in the enzyme-bound structure is changed from the standard B-form C2′ endo to C4′ exo. Therefore, further stretching in the enzyme-bound structure likely results from enzyme-DNA interactions.

Figure 2 gives the variation of all the nine rotatable torsion angles along the base flipping pathway for C5 base in $d(CCGGCGCCGG)_2$. Our analysis indicates that ζ is the only torsion angle that drives the base flipping process. The variation of ζ has been found to be approximately proportional to the translational displacement of the flipped base. Hence the change in ζ is directly correlated to the extent of rotation (angular motion) of the base, out of the helical arrangement. On the other hand the other seven torsion angles show no linear relationship with the translational displacement of the base. The magnitude of variation of these torsion angles along the base flipping pathway is substantially smaller than that of ζ.

Our analysis indicates that base flipping most likely proceeds through opening into the major groove. As shown in

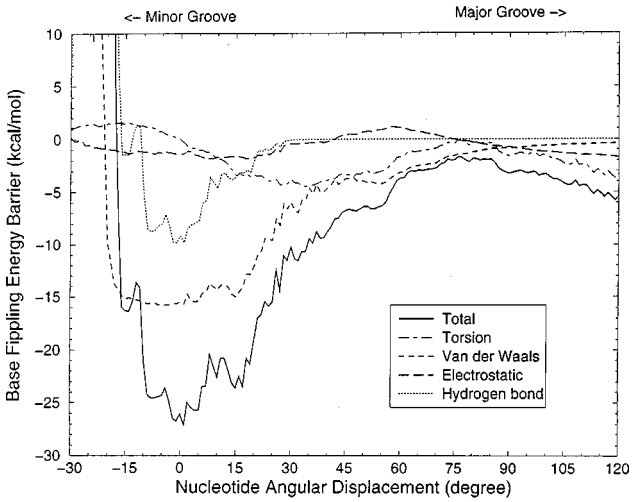

FIG. 3. Energy barrier for the flipping of the C5 base in B-DNA $d(CCGGCGCCGG)_2$.

TABLE I. Upper-limit of displacement ΔR_m for a flipped base (shown as bold letters in column 1) from flipped position towards the minor groove in several crystal *HhaI* and *HaeIII* MT ase-DNA structures. ΔR_m is defined as the displacement of the N1 or N3 atom for purine or pyrimidine base, respectively. The energy barrier of the binding enzyme along the minor groove is neglected in computing ΔR_m. For comparison ΔR_M of a base moving through the major groove in several B-DNA crystal structures are included. The literature references for the crystal structures are included in parentheses in the NDB ID column.

System	NDB ID	Flipped base	ΔR_M (Å)	ΔR_m (Å)
HhaI MTase+$d(GATAG\mathbf{C}GCTATC)_2$	PDEB08 (4)	C18		4.87
HhaI MTase+$d(GATAGCGCTATC\text{-}).$ $.d(CTATCG\mathbf{C}GATAGT)$	PDE0121 (5)	C19		6.70
HhaI MTase+$d(GATAGCGCTATC\text{-}).$ $.d(CTATCG\mathbf{C}GATAGT)$	PDED122 (5)	C19		6.30
*HaeIII*MTase+$d(ACCAGCAGG\mathbf{C}CACCAGTG\text{-}).$ $.d(\text{-}GGTCGTCCGGTGGTCACT)$	PDEB19 (25)	C10		1.23
HhaI MTase+$d(CCATG\mathbf{C}GCTGAC)_2$	PDEB123 (26)	C18		3.23
*HhaI*MTase+$d(GTCAGC\mathbf{G}CATGG).$ $.d(CAGTCG\mathbf{A}GTACC)$	PD0017 (27)	A18		4.42
B-form $d(CCGG\mathbf{C}GCCGG)_2$	BDJ039 (20)	C5	12.61	
B-form $d(CG\mathbf{C}G\mathbf{AA}TTCGCG)_2$	BDL001 (21)	C3	12.13	
		A5	14.17	
		A6	13.88	

Fig. 3, rotation of a base towards its minor groove encounters a high-energy barrier due to steric clash first with its complementary base and then with one of its neighboring bases. In contrast, no steric clash is found along the major groove.

Although the base flipping occurs through the major groove in a rather facile way, the base appears to orient towards the minor groove for ζ displacements greater than 100° (which corresponds to a base angular displacement of more than 160°). As shown in Fig. 1, the orientation of the flipped base by application of the current method is towards the minor groove, which is similar to that of the base flipped by the *HhaI* MTase. This shows that the orientation of a flipped base may not be used as a sole indicator of the pathway of flipping. Our result suggests that it may not be nec essary to assume the existence of a minor groove flipping pathway by invoking an active participation of the enzyme.

The computed energy barrier for base flipping along the major groove is 25.3 kcal/mol, which is comparable to the observed enthalpy of 17–26 kcal/mol for premelting thermal fluctuational base pair opening [23,24]. This suggests the possibility of spontaneous base flipping by thermal fluctuation in the absence of an enzyme. Such a thermally induced transient base flipping may play an important role in facilitating enzyme-base recognition and enzyme binding. For instance, a flipped base can provide more recognition sites than that in a closed helix. The steric clash between the base and binding enzyme along the major groove flipping pathway may also be avoided if the flipping occurs prior to the binding.

B. Base closing in MTase bound DNA duplexes

To investigate the effect of MTase-induced backbone distortion on the movement of a flipped base, several *HhaI* and *HaeIII* MTase-DNA crystal structures [4,5,25–27] are used as the starting structure. Earlier studies on some of these structures have shown that sequence specific contacts of these enzymes with DNA occur in the major groove [10]. Relatively insignificant conformation changes in the enzyme domain on the major groove side have been found from structural comparison between the free and DNA bound *HhaI* MTase. Thus, assuming that base flipping can be induced by enzyme binding, major groove contacts are expected to be largely maintained during the flipping process [6,10,25].

In the present paper the movement of the flipped base along the minor groove is attempted to examine the feasibility of its return to the stacked helix. In particular, we estimate the upper limit of this return movement. The movement is accomplished by rotating the ζ torsion backward starting from the flipped position, followed by simultaneous variation of other backbone torsions as outlined in the methods section. Our analysis indicates that such an operation has an insignificant effect on the major groove contacts. Thus, the enzyme-induced backbone distortion can be assumed to be largely maintained. For the purpose of estimating the upper limit of base movement, the enzyme-DNA interaction along the minor groove can be neglected. This interaction affects only the energetics of base movements.

Table I gives the upper limit of the movement of a flipped base along the minor groove in several crystal *HhaI* and *HaeIII* MTase bound-DNA structures (together with their respective nucleic acid database ID and the relevant reference). For comparison, the relevant data for some selected bases flipped by application of our method in several B-DNA crystal structures, are also included. The movement of the B-DNA bases is along the major groove, as the minor groove pathway has been shown to be blocked. The movement of the flipped base in all the MTase bound systems studied is restricted to a range between 1–7 Å, which is insufficient for the base to return to the helix stack.

Our analysis indicates that the limited freedom of the flipped bases results from the specific orientation of the O3′-P bond of the flipped nucleotide. The backbone distortion induced by each of the binding MTases appears to render the O3′-P bond of the flipped base towards an orientation perpendicular to the helix axis. As a result, the magnitude of displacement is largely diminished. We have shown in an earlier study that, if the orientation of O3′-P bond approaches that perpendicular to the helix axis, the motion in the ζ torsion becomes ineffective in promoting base angular movement in and out of the base stack [12]. In the MTase bound systems, the orientation of the O3′-P bond becomes more perpendicular to the helix as the flipped base moves towards the minor groove. Therefore, a minor groove pathway for base closing and thus flipping seems to be unlikely.

IV. CONCLUSION

Base flipping in B-DNA crystal structures and the movement of a flipped base in MTase-DNA crystal structures has been studied by means of a computational method with full description of all chemical bonds. Our results seem to indicate the feasibility of spontaneous base flipping in the absence of an enzyme. The movement of the flipped bases towards the minor groove in all the MTase bound systems have been found to be severely restricted by the enzyme-induced backbone distortion such that during base flipping the backbone needs to adopt a substantially different conformation than the observed x-ray structures. Studies on motions involving multiple bases and a model involving full enzyme-DNA interactions in complete detail are needed to further our understanding of this fascinating fundamental biological process.

[1] R. J. Roberts, Cell **82**, 9 (1995).
[2] G. Slupphaug, C. D. Mol, B. Kavli, A. S. Arvai, H. E. Krokan, and J. A. Tainer, Nature (London) **384**, 87 (1996).
[3] D. G. Vassylyev and K. Morikawa, Curr. Opin. Struct. Biol. **7**, 103 (1997).
[4] S. Klimasauskas, S. Kumar, R. J. Roberts, and X. Cheng, Cell **76**, 357 (1994).
[5] M. O'Gara, S. Klimasauskas, R. J. Roberts, and X. Cheng, X. J. Mol. Biol. **261**, 634 (1996).
[6] S. Kumar, X. Cheng, S. Klimasauskas, S. Mi, J. Posfai, R. J. Roberts, and G. G. Wilson, Nucl. Acids Res. **22**, 1 (1994).
[7] F. K. Winkler, Structure **2**, 79 (1994).
[8] J. C. Wu and D. V. Santi, J. Biol. Chem. **262**, 4778 (1987).
[9] M. Gueron and J. L. Leroy, Nature (London) **328**, 89 (1987).
[10] S. Klimasauskas, T. Szyperski, S. Serva, and K. Wuthrich, EMBO J. **17**, 317 (1998).
[11] J. Ramstein and R. Lavery, Proc. Natl. Acad. Sci. USA **85**, 7231 (1988).
[12] Y. Z. Chen, V. Mohan, and R. H. Griffey, J. Biomol. Struct. Dyn. **15**, 765 (1998).
[13] Y. Z. Chen, V. Mohan, and R. H. Griffey, Chem. Phys. Lett. **287**, 570 (1998).
[14] W. Eimer, J. R. Williamson, S. G. Boxer, and R. Pecora, Biochemistry **29**, 799 (1990).
[15] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. M. Caldwell, and P. A. Kollman, J. Am. Chem. Soc. **117**, 5179 (1995).
[16] N. C. Baird, Int. J. Quantum Chem., Symp. **1**, 49 (1974).
[17] Y. Z. Chen, W. Zhuang, and E. W. Prohofsky, Biopolymers **31**, 1273 (1991).
[18] Y. Z. Chen and E. W. Prohofsky, Nucleic Acids. Res. **20**, 415 (1992).
[19] J. A. McCammon and M. Karplus, Nature (London) **268**, 765 (1977).
[20] U. Heinemann, C. Alings, and E. Bansal, EMBO J. **11**, 1931 (1992).
[21] H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura, and R. E. Dickerson, Proc. Natl. Acad. Sci. USA **78**, 2179 (1981).
[22] H. M. Berman, W. K. Olson, D. L. Beveridge, J. W. A. Gelbin, T. Demeny, S. H. Hsieh, A. R. Srinivasan, and B. Schneider, Biophys. J. **63**, 751 (1992).
[23] D. Xu, K. O. Evans, and T. M. Nordlund, Biochemistry **33**, 9592 (1994).
[24] E. Folta-Stogniew and I. M. Russu, Biochemistry **33**, 11016 (1994).
[25] K. M. Reinisch, L. Chen, G. L. Verdine, and W. N. Lipscomb, Cell **82**, 143 (1995).
[26] M. O'Gara, R. J. Roberts, and X. Cheng, J. Mol. Biol. **263**, 597 (1996).
[27] M. O'Gara, J. R. Horton, R. Roberts, and X. Cheng, Nat. Struct. Biol. **5**, 872 (1998).