J|A|C|S
COMMUNICATIONS

Published on Web 05/29/2002

Base Flipping in DNA: Pathways and Energetics Studied with Molecular Dynamic Simulations

Péter Várnai and Richard Lavery*

Laboratoire de Biochimie Théorique, CNRS UPR 9080, Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, Paris 75005, France

Received February 19, 2002

The DNA double helix is a flexible biopolymer that can be deformed in many ways during its interactions with proteins. Certain proteins are notably able to break specific base pairs to perform chemical reactions on normally inaccessible sites. *Hha*I DNA C5-methyltransferase carries out this process on the central cytosine within a GCGC sequence.[1] Crystallographic data on a blocked reaction intermediate of the protein−DNA complex show the cytosine completely flipped out of the helical stack.[2] Although the base flipping rate of C5-methyltransferase[3] and the spontaneous opening of base pairs (as measured in NMR studies of imino proton exchange[4]) both occur on millisecond time scales, it is not known whether the protein induces base pair flipping or traps a spontaneously created open state. It should, however, be noted that imino proton exchange appears to require much smaller opening angles[5] (∼60°) than those seen in the flipped state. Whether the flipped state occurs via opening into the minor or major grooves of the double helix is also unknown, although the latter route is generally assumed to be preferred for steric reasons.[6]

As a first step to understanding this process, we have simulated the opening of the central cytosine of a DNA dodecamer 5′-dGTCAGCGCATGG-3′ containing the target sequence of *Hha*I DNA C5-methyltransferase.[2b] Molecular dynamics simulations were carried out at constant temperature (300 K) and pressure (1 bar), using AMBER with the parm99 force field.[7] DNA was solvated with roughly 5000 water molecules and 22 neutralizing Na^+ counterions in a truncated octahedral box with periodic boundaries. After equilibrating the dodecamer for 1.4 ns, starting from a B-DNA conformation, cytosine opening was induced by modifying the opening angle in 5° steps by using a soft biasing potential (0.05 kcal/mol/deg^2), carrying out 50 ps of equilibration and 150 ps of sampling at each step. The biasing restraint (added to AMBER) positions the glycosidic bond of the opening base with respect to the C1′−C1′ axis of the corresponding base pair.[8] For other details of the protocol see ref 5. Opening angles are given as differences with respect to the relaxed position of the base, positive angles leading into the major groove and negative angles into the minor groove. During 15 ns of dynamics, the base was opened by 180° into the major groove and by −200° into the minor groove. Since the opening angle seen in the methyltransferase−DNA crystal is 170°, our simulation passes through this state by two distinct routes. Nanosecond simulations at selected opening angles confirmed the stability of these results.

A free energy profile (Figure 1) was constructed for the opening pathway by using the weighted histogram analysis method[9] (WHAM). Fifty ps block and cumulative data of opening angle indicated convergence to about 1 kcal/mol. As the base pair is deformed from its relaxed state, the Watson−Crick (WC) hydrogen bonds begin to break and the free energy rises sharply. There are, however, considerable differences in the pathway depending on the direction of opening. Rotation of −25° into the minor groove results in the rupture of the three WC hydrogen bonds at a cost of 12 kcal/mol. Toward the major groove, 25° rotation breaks two WC hydrogen bonds, but creates a new bond between the cytosine carbonyl group and the imino proton of the paired guanine. This leads to a free energy shoulder at 5 kcal/mol, which persists until this bond breaks at 50° opening. The final WC hydrogen bond gives way only after 80° opening. In line with the minor groove route, the base pair is finally broken at a cost of 12 kcal/mol.

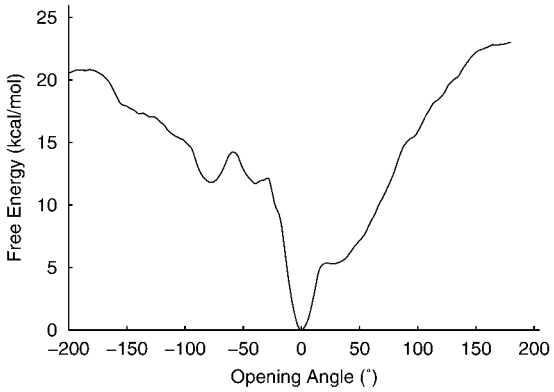

Figure 1. Free energy changes as a function of the base opening angle into the major (positive angles) and minor (negative angles).

In structural terms, minor groove opening should be more difficult due to steric clashes of the exocyclic groups and the proximity of the sugar−phosphate backbones on the minor groove side of the base pair. In fact, this difficulty is avoided by a local increase in helical rise within the cytosine strand that allows this base to slide under its paired guanine until it has overcome possible hindrance at −40° opening. This maneuver, however, explains the rapid loss of the three WC hydrogen bonds for opening toward the minor groove. Beyond −40° opening, cytosine loses its stacking interactions with its neighbors, as reflected by the increasing free energy curve. This causes the cytosine to turn out-of-plane to establish stabilizing hydrogen bonds with the base sites lining the minor groove. One such interaction is responsible for the local minimum occurring in the free energy at −80° opening (Figure 1) and consists of a base triplet with hydrogen bonds between the opening cytosine and the amino group of the guanine and the carbonyl group of the cytosine belonging to the pair on its 5′ side. Other such interactions occur as cytosine opens further, involving both base and backbone sites up to three pairs away from the opening base. This form of stabilization, whose details are certainly dependent on the neighboring base sequence, continues beyond −140° opening when the cytosine has left the minor groove and

* Corresponding author. E-mail: rlavery@ibpc.fr.

10.1021/ja025980x CCC: $22.00

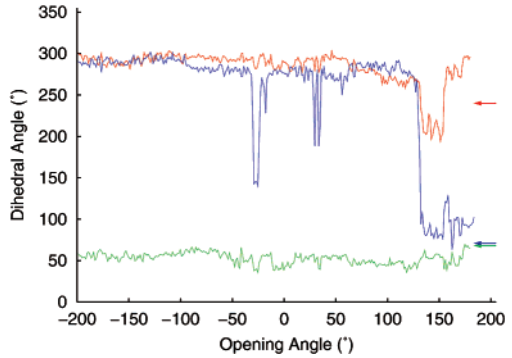

Figure 2. Dihedral angles α (red), γ (green), and ζ (blue) at the 3′ side of the opening base along the pathways. Arrows indicate experimental values.[2b]

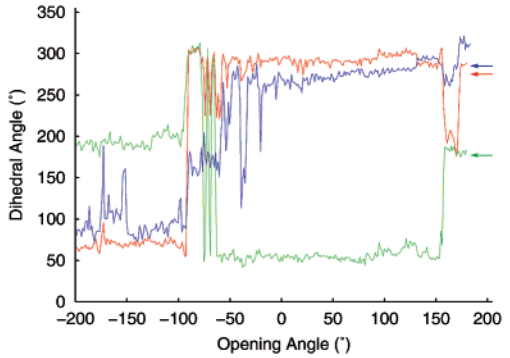

Figure 3. Dihedral angles α (red), γ (green), and ζ (blue) at the 5′ side of the opening base along the pathways. Arrows indicate experimental data.

only ceases when the base rotates upward and points freely into the surrounding solution at −175° opening (ΔG = 21 kcal/mol).

No such stabilization occurs for opening toward the major groove and, as reflected by the free energy curve, the base rotates smoothly out of stack. Another important contrast with the minor groove route is that a significant backbone conformational change occurs after 125° opening, which leads the cytidine sugar to point outward. This change allows the amino group of the opening cytosine to bond to an anionic oxygen of the neighboring 5′ phosphate. To compensate for the water-filled gap created in the double helix, the unpaired guanine moves across to stack principally with the guanines which flank the opening base and, by 160° opening, also forms a hydrogen bond between its amino group and an anionic phosphate oxygen of the opening cytidine. This change weakens the GC pair on the 5′ side of the opening base which buckles and loses two of its hydrogen bonds; the cytosine, however, forms a new hydrogen bond with a phosphate group two levels before. Following these backbone and stacking transitions, the opening cytosine points into solution at 175° opening (23 kcal/mol).

Although the minor and major groove 170° base-flipped states only differ by 2 kcal/mol, they are associated with different backbone and stacking conformations. Nanosecond simulations of these states show no spontaneous interconversion and suggest that they are separated by a significant energy barrier. The major groove route leads to a state characterized by a turned-out sugar and a coupled ζ dihedral (O3′−P) transition from g^- to g^+ of this nucleotide (Figure 2), not normally seen in B-DNA. The ζ on the 5′ side of the opening cytidine also makes a transition from g^- to g^+ (through a *t* intermediate state) but only for the minor groove route (Figure 3). ζ has already been linked to base opening,[10] but we find that it only intervenes for flipped states with significant backbone rearrangement. Other differences involve the α dihedral (P−O5′) of the opening cytidine, which makes a g^- to g^+ transition along the minor groove route, but not toward the major groove. Finally, these differences are also coupled to changes in sugar puckers, with the pucker of the opening base moving to higher phase angles for minor groove opening, while major groove opening produces a transition to low phase angles. It is remarked that recent simulation studies (communicated to us by the authors) do not appear to show similar differences between the minor and major groove flipped states and also show virtually flat free energy curves once the base pair hydrogen bonds have been broken.[11]

If we compare our flipped states with the trapped intermediate seen in the C5-methyltransferase complex,[2b] it is clear that the major groove route leads to a structure in surprisingly good agreement with the protein-bound form. This is true for the overall geometry, but also for all the backbone dihedrals shown in Figures 2 and 3. Minor groove opening does not induce this backbone rearrangement spontaneously and also leads to intermediate states with the open base bound in the minor groove, which may or may not hinder complexation with the protein, but almost certainly influence opening kinetics. Given our free energy profile, minor groove flipping is feasible, but clearly requires a further conformational step to reach the experimentally observed flipped state. Irrespective of the route, a spontaneous occurrence of a fully flipped state prior to protein binding is improbable on energy grounds. Spontaneous base opening (∼60°) combined with enzyme participation is possible and would favor the major groove direction. It is finally noted that neither opening route leads to a local energy minimum for the fully flipped state, which points to an active stabilizing role for the enzyme.

Acknowledgment. P.V. thanks The Wellcome Trust for an International Prize Travelling Research Postdoctoral Fellowship (grant reference 060078). We acknowledge a generous allocation of computer time from CINES (France) and thank Dr. Alex MacKerell for providing results prior to publication.

References

(1) Cheng, X.; Roberts, R. J. *Nucleic Acids Res.* **2001**, *29*, 3784−3795.
(2) (a) Klimašauskas, S.; Kumar, S.; Roberts, R. J.; Cheng, X. *Cell* **1994**, *76*, 357−369. (b) Kumar, S.; Horton, J. R.; Jones, G. D.; Walker, R. T.; Roberts, R. J.; Cheng, X. *Nucleic Acids Res.* **1997**, *25*, 2773−2783.
(3) (a) Wu, J. C.; Santi, D. V. *J. Biol. Chem.* **1987**, *262*, 4778−4786. (b) Klimašauskas, S.; Szyperski, T.; Serva, S.; Wüthrich, K. *EMBO J.* **1998**, *17*, 317−324. (c) Vilkaitis, G.; Merkiene, E.; Serva, S.; Weinhold, E.; Klimašauskas, S. *J. Biol. Chem.* **2001**, *276*, 20924−20934.
(4) Guéron, M.; Leroy, J.-L. *Methods Enzymol.* **1995**, *261*, 383−413.
(5) Giudice, E.; Várnai, P.; Lavery, R. *Chem. Phys. Chem.* **2001**, *11*, 673−677.
(6) Ramstein, J.; Lavery, R. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 7231−7235.
(7) (a) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weimer, P. K.; Kollman, P. A. *AMBER 6*; University of California: San Francisco, CA, 1999. (b) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049−1074.
(8) Bernet, J.; Zakrzewska, K.; Lavery, R. *J. Mol. Struct. (THEOCHEM)* **1997**, *398*, 473−482.
(9) (a) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011−1021. (b) Boczko, E. M.; Brooks, C. L., III *J. Phys. Chem.* **1993**, *97*, 4509−4513.
(10) Chen, Y. Z.; Mohan, V.; Griffey, R. H. *Phys. Rev. E* **2000**, *62*, 1133−1137
(11) Banavali, N. K.; MacKerell, A. D., Jr. *J. Mol. Biol.* In press.

JA025980X