J|A|C|S
COMMUNICATIONS

Published on Web 04/05/2003

Nucleic Acid Base Pair Dynamics: The Impact of Sequence and Structure Using Free-Energy Calculations

Emmanuel Giudice 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 January 9, 2003; E-mail: rlavery@ibpc.fr

Proton exchange is a powerful tool for probing the dynamic properties of biological macromolecules. In the case of double-stranded nucleic acids, NMR studies of imino proton exchange of thymine and guanine have shown that base pairs open spontaneously on a millisecond time scale.[1−4] Such exchange requires that the imino protons become accessible to the surrounding solvent by rotating the base into either of the grooves of the duplex, breaking the Watson−Crick hydrogen bonds and strongly diminishing the π-stacking with neighboring base pairs. It is thus not surprising that the stronger hydrogen bonding of GC pairs is reflected in their longer lifetimes, which are typically 15−25 ms as compared to 5−10 ms for AT pairs.[1,3] Although a number of studies have shown that the surrounding base sequence seems to have only moderate effects on base pair lifetimes,[2−4] there is one notable exception involving so-called A-tracts (oligo-adenine runs containing at least three successive AT pairs) which induce a change in duplex structure and, when juxtaposed with other sequences, induce helix bending. A-tracts are characterized by strong propeller twists and groove narrowing in the 5′ → 3′ direction, leading to a structural variant of B-DNA known as B′.[5,6] Although these structural changes do not seem to be very significant, the opening lifetime of AT pairs within A-tracts increases dramatically and, in some cases, exceeds 100 ms.[3] Why AT base pair opening is affected so strongly by the apparently minor perturbation from B → B′ DNA is not known. This result may appear even more surprising, when it is noted that the passage from B-DNA to A-RNA has a comparatively small impact on opening lifetimes, although the structural perturbation is much greater, in terms of both helical parameters (base pairs inclined and strongly displaced toward the minor groove creating a deep and narrow major groove) and backbone conformation (decreased glycosidic torsion angles coupled to sugar pucker transitions from C2′-endo in the B-form to C3′-endo in the A-form). In RNA, the lifetimes of GC pairs, 40−50 ms, are roughly twice those in DNA, while those of AU pairs are shorter ($\leq$1 ms) or, in cases such as AU pairs sandwiched between GC pairs, virtually identical to those in DNA.[4]

It has recently become feasible to carry out molecular dynamics free-energy calculations to clarify the details of base pair opening.[7−10] We have been able to study AT and GC base pair opening for the central base pairs within a B-DNA 13-mer, d(GAGAGAGAGAGAG), hereafter termed GA_{ref}, explicitly taking into account both solvent and counterions.[7,10] Free-energy variations were obtained using umbrella sampling with an angular restraint to control the rotation of a chosen base into either of the grooves of the duplex.[7] The results show that at least 50° of rotation is necessary to make the imino protons of thymine or guanine sterically accessible. Energetically, major groove opening is easier for purines (A, G), but pyrimidines (T, C) show almost identical energy profiles toward either groove. It was also noted that base opening is coupled to bending, which increases as a base is removed from the helical stack. This bending is dominantly in the direction of the major groove for both opening pathways.

We now extend these calculations in an attempt to understand the effects of sequence and structure discussed above. First, we have studied a modified GA_{ref} oligomer involving two GC → AT substitutions (underlined), and creating a five base pair A-tract, d(GAGAGAAAAAGAG). We have also built an A-RNA, using the original GA alternating sequence, r(GAGAGAGAGAGAG). Thymine opening has been studied for the A_8T_{19} pair at the center of the B-DNA A-tract and for the equivalent A_8U_{19} pair of the A-RNA (numbering both strands in the 5′ → 3′ direction and starting with the homopurine strand). Simulations were carried out using AMBER 6.0 with the parm99 force field[11] in a truncated octahedral unit cell containing the solute molecule solvated by between 5600 and 6000 water molecules and neutralized by 24 Na^+ counterions. After equilibrating each system, we continued the simulations for 2 ns using a constant temperature and pressure ensemble and using particle mesh Ewald summations to treat long-range electrostatic interactions. Opening was then induced in 5° steps, and free energies were obtained by 50 ps of equilibration and 150 ps of sampling in each window (see refs 8 and 10 for other details of the protocol).

Analysis of the unperturbed DNA oligomer shows an overall B-like conformation. However, the A-tract exhibits strong propeller twisting of the AT pairs (with average magnitudes increasing 5′ → 3′, with −11° at A_6 and −17° at A_9), and a narrow minor groove (with average widths decreasing 5′ → 3′ with 6.7 Å at A_6 and 4.8 Å at A_9). The overall bending of the oligomer (average value 14.5°, measured using CURVES[12]) is slightly smaller than that of the canonical GA_{ref} (17.4°), and it is more rigid (with a standard deviation of bending of 7.5°, smaller than that of GA_{ref} by 2.4°). This result is in line with earlier MD simulations which also led to A-tracts with B′-like structures, but found that overall bending only exceeded that of canonical B-DNA in the presence of excess salt.[13,14] It is also in line with experimental results which show reduced A-tract gel retardation in the presence of Na^+.[15]

The A-RNA simulations lead to a canonical A-form conformation with average base pair inclinations of 9°, Xshifts of −4.5 Å (toward the minor groove), and dominantly C3′-endo sugars. As expected, the major groove is very narrow, its average width, 6.8 Å, being only 0.5 Å larger than the minor groove of B-DNA.

We now turn to base opening. Figure 1 shows the free energy for opening T_{19} within the B-DNA A-tract as a function of our angular constraint.[7,8] As with our earlier results,[7,8,10] this curve shows a quadratic zone associated with elastic deformation of the AT pair, followed by a quasi-linear energy increase as the thymine moves out of stack. In the zone where the thymine imino proton becomes accessible ($\sim\pm$50°), there is a roughly 1 kcal mol^{-1} advantage for major groove opening.

Comparison with the corresponding free-energy curve from GA_{ref} (dotted line) shows that it is indeed more difficult to open thymine

10.1021/ja034095r CCC: $25.00

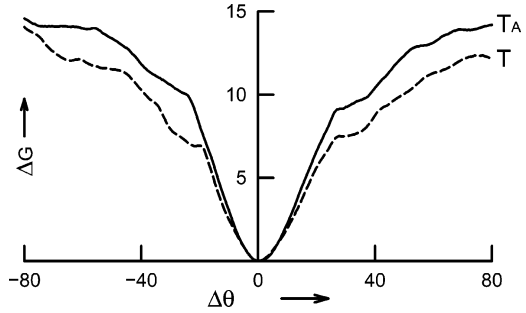

Figure 1. Free-energy variation (kcal mol^{-1}) as a function of the base opening angle (deg). Solid lines refer to opening thymine (T_A) at the center of a B-DNA A-tract, and dotted lines refer to opening thymine (T) at the center of a canonical B-DNA oligomer (GA_{ref}). Positive values refer to major groove opening, and negative values refer to minor groove opening.

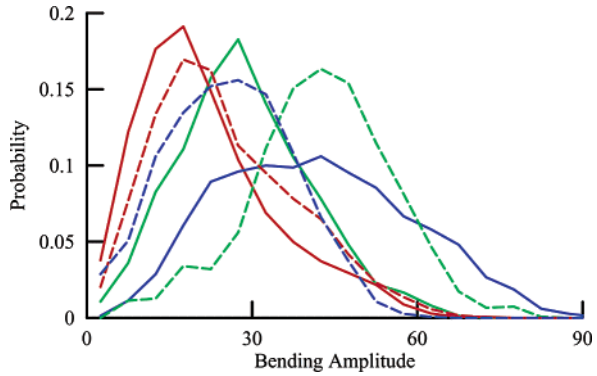

Figure 2. Histograms of bending amplitude. Red, relaxed DNA oligomer; green, thymine opened into the major groove by >50°; blue, thymine opened into the minor groove by <−50°. Solid lines refer to opening thymine at the center of a B-DNA A-tract, and dotted lines refer to opening thymine at the center of a canonical B-DNA oligomer (GA_{ref}).

within the A-tract by roughly 2 kcal mol^{-1} toward either groove. The conformational changes associated with opening thymine are very similar to those seen with GA_{ref},[7] with the exception of the induced bending (Figure 2), which is now weaker for major groove opening and stronger for minor groove opening.

The results for opening U_{19} within the A-RNA oligomer also show the characteristic quadratic and quasi-linear domains, with, once again, a small preference (∼1 kcal mol^{-1}) for major groove opening (Figure 3). Comparison with the free-energy curve for T_{21} opening within GA_{ref}, however, indicates almost identical results for opening angles up to and beyond the domain where the pyrimidine imino protons become accessible (∼±50°). The most striking conformational change coupled to opening is not bending, as in the case of DNA, but variations in the major groove width at the level of the opening pair, which increases by up to 8 Å for base opening into the major groove (although only 3−4 Å is necessary to reach 70° of opening) and diminishes by roughly 4 Å for base opening into the minor groove.

The free energies calculated for T opening within the DNA A-tract and for U opening within RNA therefore correlate with the experimental observations of base pair lifetimes. Do these simulations provide any insight into the underlying causes? The answer appears to be connected with changes in groove widths and flexibilities.

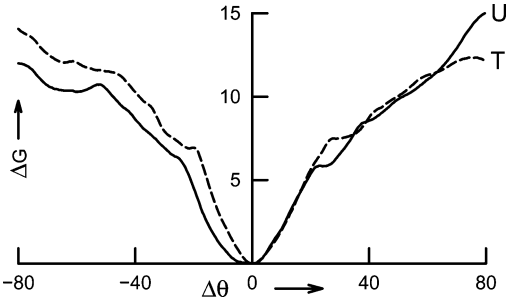

Figure 3. Free-energy variation (kcal mol^{-1}) as a function of the base opening angle (deg). Solid lines refer to opening uracil (U) at the center of an A-RNA oligomer, and dotted lines refer to opening thymine (T) at the center of a canonical B-DNA oligomer (GA_{ref}). Positive values refer to major groove opening, and negative values refer to minor groove opening.

In canonical B-DNA, both grooves are large enough to allow base opening, and the molecule is flexible enough to respond by bending. In contrast, the narrow and rather rigid minor groove of A-tracts (the standard deviation in width being 1.2 Å at the center of the A-tract as compared to an average of 1.5 Å in GA_{ref}) leads to a modified response to thymine opening. Toward the major groove, the narrow minor groove remains intact and hinders bending. Toward the minor groove, bending occurs once base opening has perturbed the minor groove. Overall, the major groove pathway is preferred, but opening is more costly than in canonical DNA. In A-RNA, major groove opening requires significant major groove stretching, but this occurs easily (the average standard deviation in width for the unperturbed RNA being 2.2 Å), a property reflected in the significant major groove contraction occurring during minor groove opening. This flexibility explains why base opening toward the major groove is preferred despite its narrowness. We conclude that the base pair lifetimes for canonical B-DNA and A-RNA directly reflect the difficulty of breaking Watson−Crick hydrogen bonds and local stacking interactions, while the unusually long lifetimes for AT pairs within A-tracts reflect the additional constraints imposed by a narrow and rigid minor groove.

References

(1) Guéron, M.; Leroy, J.-L. *Methods Enzymol.* **1995**, *261*, 383−413.
(2) Folta-Stogniew, E.; Russu, I. M. *Biochemistry* **1994**, *33*, 11016−11024.
(3) Leroy, J. L.; Charretier, E.; Kochoyan, M.; Guéron, M. *Biochemistry* **1988**, *27*, 8894−8898.
(4) Snoussi, K.; Leroy, J. L. *Biochemistry* **2001**, *40*, 8898−8904.
(5) Chuprina, V. P. *Nucleic Acids Res.* **1987**, *15*, 293−311.
(6) Dickerson, R. E.; Goodsell, D. S.; Neidle, S. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3579−3583.
(7) Giudice, E.; Várnai, P.; Lavery, R. *ChemPhysChem* **2001**, *2*, 673−677.
(8) Varnai, P.; Lavery, R. *J. Am. Chem. Soc.* **2002**, *124*, 7272−7273.
(9) Banavali, N. K.; MacKerell, A. D., Jr. *J. Mol. Biol.* **2002**, *319*, 141−160.
(10) Giudice, E.; Várnai, P.; Lavery, R. *Nucleic Acid Res.* **2003**, *31*, 1434−1443.
(11) (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, 1999. (b) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049−1074.
(12) Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1989**, *6*, 655−667.
(13) Young, M. A.; Beveridge, D. L. *J. Mol. Biol.* **1998**, *281*, 675−687.
(14) McConnell, K. J.; Beveridge, D. L. *J. Mol. Biol.* **2001**, *314*, 23−40.
(15) Diekmann, S.; Wang, J. C. *J. Mol. Biol.* **1985**, *186*, 1−11.

JA034095R