Free Energy Analysis of the Conformational Preferences of A and B Forms of DNA in Solution

B. Jayaram,† D. Sprous, M. A. Young, and D. L. Beveridge*

Contribution from the Department of Chemistry and Program in Molecular Biophysics, Wesleyan University, Middletown, Connecticut 06459

Received April 17, 1998. Revised Manuscript Received August 19, 1998

Abstract: Right-handed DNA duplexes assume a B form at high water activity and an A form at reduced levels, but the molecular origins of this behavior are in debate. Four large-scale molecular dynamics simulations performed on sodium salts of the A and B forms of DNA [d(CGCGAATTCGCG)] in water and in ethanol/water mixtures form the basis for a molecular level explanation of the origins of environmental sensitivity of DNA conformation. The trends observed in conformational preferences experimentally are accounted for by the calculations. Analysis of the results indicates the free energy associated with the explicit organization of the mobile counterions around the A and B forms of DNA to be the key feature in the resolution of otherwise paradoxical observed trends.

I. Introduction

Conformational changes in DNA are an important aspect of drug−DNA and protein−DNA interactions, and understanding the environmental sensitivity of DNA structure is necessary for a full account of structure−function relationships in nucleic acids. Right-handed DNA duplexes of mixed sequence are generally expected to assume a B form at high water activity and an A form at reduced levels.[1−3] If the intramolecular energetics of the DNA favored the A form and hydration favored the B form, explaining the observed conformational preferences would be straightforward. At high water activity, the contribution of hydration would dominate the total free energy, preferentially stabilizing the B form. At low water activity, with the hydration energy correspondingly reduced, the intramolecular term would dominate, favoring the A form. However, this explanation is not immediately consistent with oligonucleotide crystal structure data,[4,5] which show the A form of DNA to be more compact than B, with interphosphate distances shorter by ca. 0.7 Å. It follows that phosphate−phosphate repulsions are larger in A DNA than in B, and that A form DNA is be electrostatically destabilized with respect to B. Thus, to explain the observed behavior, a deeper molecular level analysis of the problem is required.

Factors historically identified with conformational preferences of right-handed DNA helices include solvent accessibility,[6] base stacking interactions,[2] hydrophobic effects,[3] the "economics" of phosphate hydration,[7] and the minor groove spine of hydration.[4] However, an explanation of the observed behavior in which structures are linked explicitly with free energy and thermodynamic stability has not yet been achieved. In this article, we present a detailed analysis of the molecular origins of the conformational stability of the A and B forms of a DNA oligonucleotide in solution based on theoretical calculations of free energies. The experimentally observed conformational preferences are accounted for by the calculations. Analysis of the results indicate that the explicit organization of the mobile counterions around the A and B forms of DNA is a key feature in the resolution of otherwise paradoxical observed trends.

II. Methods

Conformational preferences of DNA oligonucleotides are currently the focus of several large-scale molecular dynamics (MD) simulations, including solvent and counterions explicitly and extending well into the nanosecond time scale.[8−13] All simulations to date have emphasized structure over thermodynamics, since full free energy determinations, while well-defined in principle,[14] are not computationally feasible for the size of the systems under consideration. To proceed beyond this obstacle, we obtained a select set of DNA and associated counterion *structures* from MD trajectories describing A and B DNA in various environmental circumstances as well as an A-to-B transition[9,10] and used them as a basis for post facto estimation and analysis of the conformational free energies. The calculation of intramolecular enthalpies is based on the empirical energy functions used in the MD simulations,[15] with corresponding entropies calculated from the ensemble of MD structures by the quasiharmonic method.[16] The free

† On leave from the Department of Chemistry, Indian Institute of Technology, Hauz Khas, New Delhi-110016, India.

(1) Franklin, R. E.; Gosling, R. G. *Acta Crystallogr.* **1953**, *6*, 673.
(2) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.
(3) Ivanov, V. I.; Krylov, D. Y. *Methods Enzymol.* **1992**, *211*, 111.
(4) Dickerson, R. E. *Methods Enzymol.* **1992**, *211*, 67.
(5) Berman, H. M.; Olson, W. K.; Beveridge, D. L.; Westbrook, J.; Gelbin, A.; Demeny, T.; Hsieh, S.-H.; Srinivasan, A. R.; Schneider, B. *Biophys. J.* **1992**, *63*, 751.
(6) Alden, C. J.; Kim, S. H. *J. Mol. Biol.* **1979**, *132*, 411.
(7) Saenger, W.; Hunter, W. N.; Kennard, O. *Nature* **1986**, *324*, 385.
(8) Beveridge, D. L.; Ravishanker, G. *Curr. Opin. Struct. Biol.* **1994**, *4*, 246.
(9) Young, M. A.; Ravishanker, G.; Beveridge, D. L. *Biophys. J.* **1997**, *73*, 2313.
(10) Sprous, D.; Young, M. A.; Beveridge, D. L. *J. Phys. Chem.* **1998**, *102*, 4658.
(11) Feig, M.; Pettitt, B. M. *J. Phys. Chem.* **1997**, *101*, 7361.
(12) Cheatham, T. E., III; Crowley, M. F.; Fox, T.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 9626.
(13) MacKerell, A. D. J. *J. Phys. Chem.* **1997**, *101*, 646.
(14) Beveridge, D. L.; DiCapua, F. M. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431.
(15) Cornell, W. D.; et al. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
(16) Karplus, M.; Kushick, J. N. *Macromolecules* **1981**, *14*, 325.

10.1021/ja981307p CCC: $15.00
Published on Web 10/01/1998

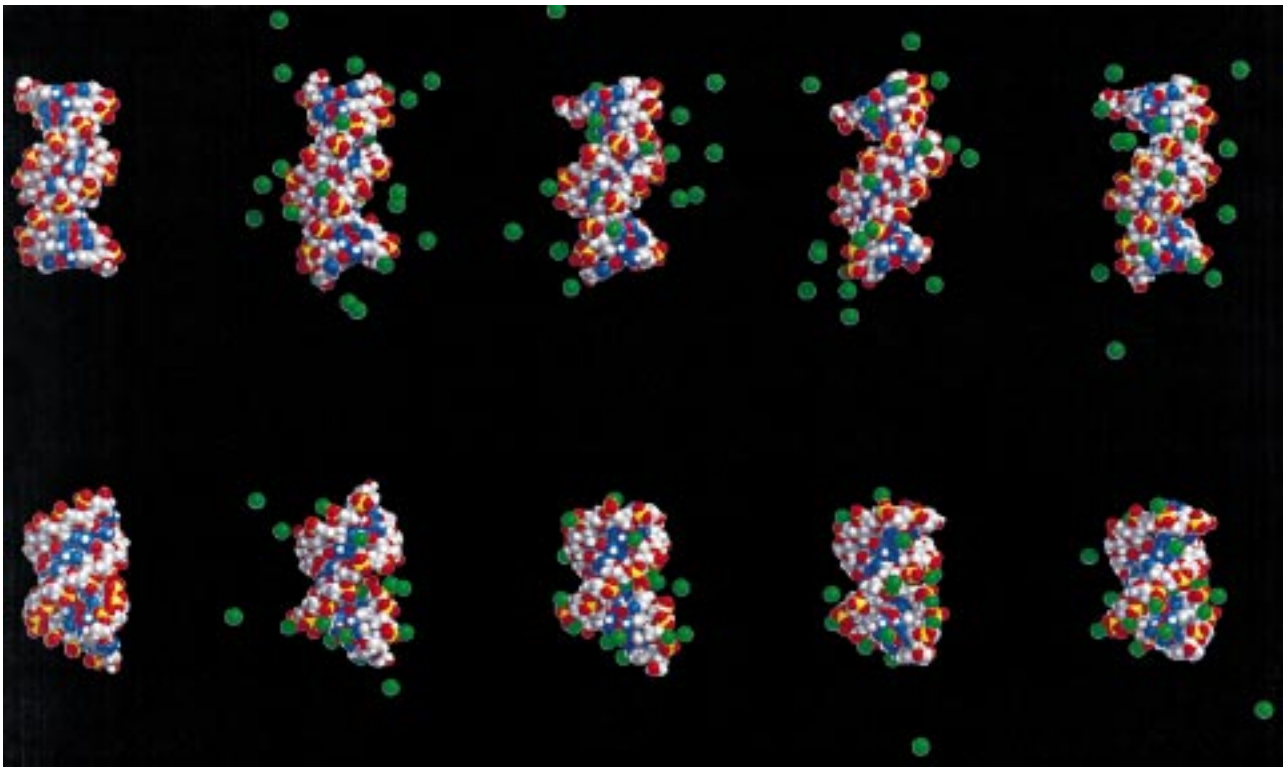


Figure 1. (i) Canonical form of B DNA (top row, first structure) and MD view of B DNA in a medium of explicit water and counterions (top row, last four structures). (ii) Canonical form of A DNA (bottom row, first structure) and MD view of A DNA in a mixed solvent system containing 85% ethanol, water, and counterions (bottom row, last four structures).

energies of solvation were evaluated for these structures using a modified version of the generalized Born solvent accessibility (GBSA) method,[17−19] demonstrated previously to predict the solvation free energy of a large number of small organic molecules and ions within ∼5% of the observed values.

The sodium salt of the d(CGCGAATTCGCG) duplex (hereafter referred to as "NaDNA") is taken as the focus for investigation. The oligonucleotide d(CGCGAATTCGCG) presents a full turn of DNA double helix and has been characterized extensively by crystallography[4] and NMR[20] and via a series of MD simulations.[8,9,10,21] Our free energy analysis is based on four MD computer simulations performed using the AMBER 4.1 program,[22] utilizing the empirical force field recently proposed for nucleic acids in solution by Cornell et al.[15] The simulations are as follow: (i) free dynamics of B form NaDNA in water, (ii) constrained dynamics of A form NaDNA in water, and free dynamics of both (iii) B form NaDNA and (iv) A form NaDNA in 85% (v/v) ethanol/water mixture (hereafter abbreviated as 85% EtOH). Simulation protocols follow those described in detail by Young et al.[9] and are maintained as similar as possible for each case. Some 4000 solvent molecules and 22 sodium counterions, sufficient to achieve electrical neutrality, were explicitly included in the aqueous MDs, while the mixed solvent cases involved 877 EtOH and 501 water molecules. The MD trajectories are developed in a (T, P, N) ensemble, using particle mesh Ewald[23] to treat long-range electrostatics, and each was extended into the nanosecond regime based on a 2-fs (10^{-15} s) time step. A constraint function was necessary in case ii, since the A form of the oligonucleotide duplex readily converts to B form in water during the time scale of simulations.[10,12] Some 100 structures of DNA together with the counterions were culled from each of the above four simulations. Representative snapshots from the A and B form trajectories are shown in Figure 1; the leftmost panel in each row shows the canonical DNA starting structure.

For analysis, the ensemble averages of the intramolecular energy differences were pooled separately into bonded (bond, angle, dihedral) and nonbonded van der Waals and electrostatic terms (which include the 1−4 contributions). Counterion (CI) electrostatics are grouped separately. These energy terms collectively provide an estimate of the intramolecular conformational enthalpy. The corresponding intramolecular entropy contribution to the free energy ($-T\Delta S$) is estimated for the ensemble of structures using the quasiharmonic method.[16] The absolute entropy of Na^+ ions in water of 9.1 cal mol^{-1} K^{-1} (excluding the electrostatic contribution to the entropy of solvation from its gas-phase value)[24] translates to a TS ($T = 298$ K) of 2.7 kcal/mol for each counterion free in solution. After examining the entropies of Na^+ ions in water, ethanol, and crystals,[25] we adopted a TS value of 2 and 3 kcal/mol for each ion free in water and 85% EtOH solution, respectively. Counterions beyond the second shell of DNA, as noted from the DNA−Na^+ radial distribution functions in MD simulations, are treated as free. The contribution of free counterions to $T\Delta S$ terms is small, and that of condensed counterions is still smaller, and thus $T\Delta S$ terms are found to be dominated by DNA intramolecular quasi-harmonic entropies.

The solvation free energy of each structure is estimated using the GBSA method with modifications introduced by Jayaram et al.[19] To be consistent with the MD force field, the GB calculations were reparametrized to accurately reproduce solvation free energies of 32 small molecules comprising the constituents of proteins and nucleic acids.[26] A dielectric constant of 80 is employed for analyzing the solutes in water. For the mixed solvent system of 85% EtOH, the

(17) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127.
(18) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1996**, *100*, 19824.
(19) Jayaram, B.; Liu, Y.; Beveridge, D. L. *J. Chem. Phys.* **1998**, *109*, 1465.
(20) Lane, A.; Jenkins, T. C.; Brown, T.; Neidle, S. *Biochemistry* **1991**, *30*, 1372.
(21) Beveridge, D. L. In *Encyclopedia Of Computational Chemistry*; Schleyer, P. v. R., Ed.; John Wiley and Sons: New York, 1997.
(22) Pearlman, D. A.; et al. *AMBER 4.1*; UCSF: San Francisco, CA, 1995.
(23) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1995**, *98*, 10089.
(24) Friedman, H. L.; Krishnan, C. V. In *Water. A Comprehensive Treatise*; Franks, F., Ed.; Plenum Press: New York, 1973.
(25) Krestov, G. A. *Thermodynamics of Solvation*; Ellis-Horwood: New York, 1991.
(26) Jayaram, B.; Sprous, D.; Beveridge, D. L. Manuscript in preparation.

Table 1. Energetics Seen for theVarious Legs of the Conformational Change Scheme

(a) Process I

	B(water)	B(85% eth)	ΔE, B(85% eth) → B(water)
ΔH(bonded)	1022.36	983.43	38.93
ΔH(vdw)	−204.2	−201.71	−2.49
ΔH(electrostatics-DNA)	954	1144.86	−190.86
ΔH(electrostatics-CI)	−4882.2	−6260	1377.8
ΔH(constraint)	0	0	0
ΔG(electrostatics solvation)	−3388.6	−2157.30225	−1231.29775
ΔG(hydrophobic solvation)	35	34.6	0.4
$-T\Delta S$(quasiharmonic)	0	65.3	−65.3
$-T\Delta S$(CI-release)	−24	−24	0
ΔG (total)	−6487	−6414.6	−72.4

(b) Process II

	A(water)	A(85% eth)	ΔE, A(85% eth) → A(water)
ΔH(bonded)	995.22	1051	−55.78
ΔH(vdw)	−115	−203	88
ΔH(electrostatics-DNA)	2047	1507	540
ΔH(electrostatics-CI)	−6266	−6732.03	466.03
ΔH(constraint)	194	0	194
ΔG(electrostatics solvation)	−2986	−2121.8	−864.2
ΔG(hydrophobic solvation)	33.6	34.9	−1.3
$-T\Delta S$(quasiharmonic)	43.1	22.4	20.7
$-T\Delta S$(CI-release)	−20	−18	−2
ΔG (total)	−6074	−6459.5	−385.5

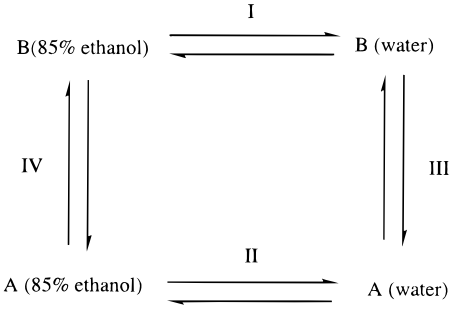
(c) Process III

	B(water)	A(water)	ΔE, A(water) → B(water)
ΔH(bonded)	1022.36	995.22	27.14
ΔH(vdw)	−204.2	−115	−89.2
ΔH(electrostatics-DNA)	954	2047	−1093
ΔH(electrostatics-CI)	−4882.2	−6266	1383.8
ΔH(constraint)	0	194	−194
ΔG(electrostatics solvation)	−3388.6	−2986	−402.6
ΔG(hydrophobic solvation)	35	33.6	1.4
$-T\Delta S$(quasiharmonic)	0	43.1	−43.1
ΔG (total)	−6463.64	−6054.08	−409.56

(d) Process IV

	B(85% eth)	A(85% eth)	ΔE, A(85% eth) → B(85% eth)
ΔH(bonded)	983.43	1051	−67.57
ΔH(vdw)	−201.71	−203	1.29
ΔH(electrostatics-DNA)	1144.86	1507	−362.14
ΔH(electrostatics-CI)	−6260	−6732.03	472.03
ΔH(constraint)	0	0	0
ΔG(electrostatics solvation)	−2157.30225	−2121.8	−35.50225
ΔG(hydrophobic solvation)	34.6	34.9	−0.3
$-T\Delta S$(quasiharmonic)	65.3	22.4	42.9
$-T\Delta S$(CI-release)	−24	−18	−6
ΔG (total)	−6414.6	−6459.5	44.9

dielectric constant is set at 30, consistent with experiment.[27] The nonelectrostatic contribution to the solvation free energy is computed via SA contributions weighted with a coefficient of 7.2 cal/Å^2,[17] which accounts for both van der Waals interactions of the solute with solvent and cavity formation expense in the solvent. A probe radius of 1.4 Å was used for SA calculations. A larger probe radius and a different value for the coefficient in the case of mixed solvent system could be contemplated, but the MD simulations show that the first solvation shell of DNA is dominated by waters.[10] The computed nonelectrostatic energies are 35.0 kcal for $[\text{B DNA}]_{H_2O}$, 34.9 kcal for $[\text{B DNA}]_{85\%\ \text{EtOH}}$, and 34.6 kcal for $[\text{A DNA}]_{85\%\ \text{EtOH}}$. The difference between A and B forms in these latter terms is thus negligibly small in the overall analysis.



III. Results

The complete analysis involves four steps, as indicated in the following reaction scheme:

Results on steps I and II are provided to establish that the observed preference of B and A forms of DNA for water and 85% EtOH, respectively, are successfully reproduced in the calculations. A consideration of the relative stability of A and B forms in water and 85% EtOH follows from an analysis of

(27) Hasted, J. B. In *Water*: *A Comprehensive Treatise*; Franks, F., Ed.; Plenum: New York, 1973; p 421.

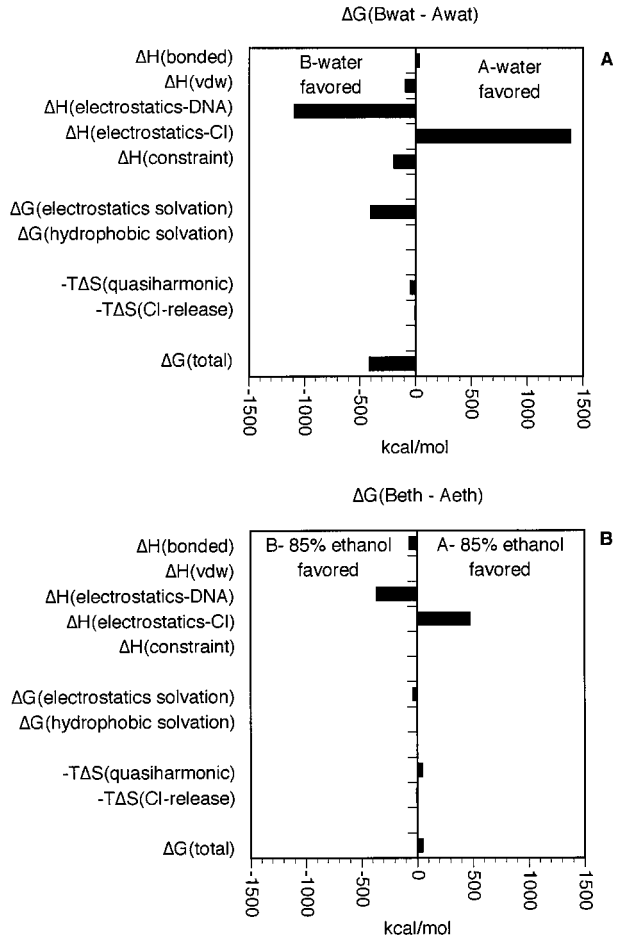


Figure 2. Intramolecular contributions to the enthalpy and entropy, and electrostatic component of the solvation free energy shown as differences between the respective components for A and B forms of NaDNA in water and in ethanol/water mixture: (A) step I of reaction scheme, $[B\ NaDNA]_{H_2O} - [B\ NaDNA]_{85\%\ EtOH}$; (B) step II of reaction scheme, $[A\ NaDNA]_{H_2O} - [A\ NaDNA]_{85\%\ EtOH}$.

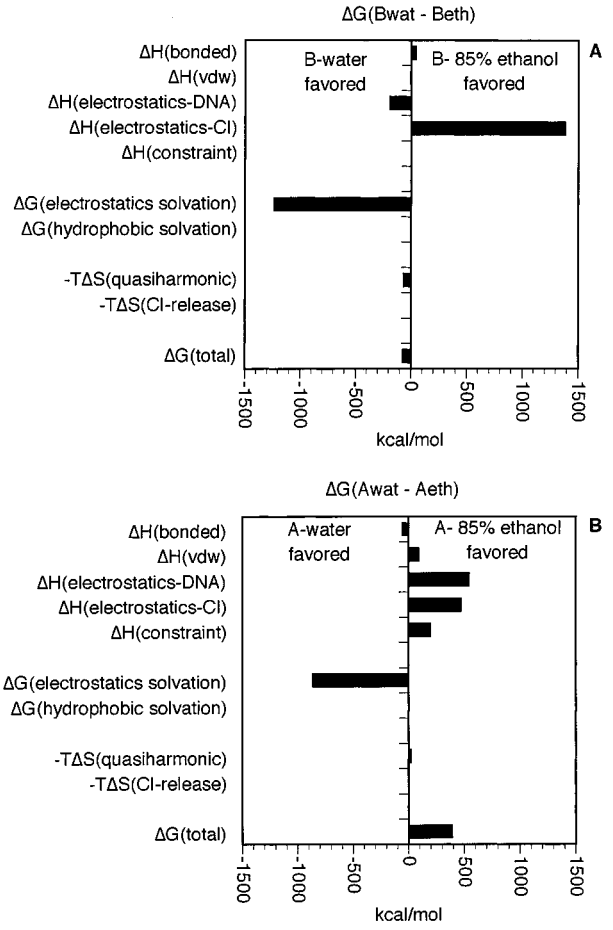


Figure 3. Relative energetics of A and B forms of DNA (as in Figure 2): (A) step III of reaction scheme, $[B\ NaDNA]_{H_2O} - [A\ NaDNA]_{H_2O}$; (B) step IV of reaction scheme, $[B\ NaDNA]_{85\%\ EtOH} - [A\ NaDNA]_{85\%\ EtOH}$.

steps III and IV in the reaction scheme. Calculated results for each leg of the reaction scheme are given in Table 1.

The relative energetics of step I, B form DNA in water vis-à-vis 85% EtOH, are shown in Figure 2A. The bonded interactions (bond stretching, angle bending, dihedral angle variations) and the van der Waals interactions for the DNA are much the same in the two solvents. The intramolecular electrostatics of DNA, dominated by the phosphate repulsions, is also essentially independent of the solvent. In contrast, the DNA counterion interactions are decidedly more favorable in 85% EtOH solution than in water, a result of an increased extent of counterion condensation in the reduced dielectric environment. Thus, with counterions included (system B NaDNA), the *intra*molecular preference of the B form is for 85% EtOH solvent over water. However, the solvation free energy of the B form NaDNA complex in water is larger than that in 85% EtOH and overcomes the intramolecular preference for the lower dielectric alternative. Thus, the calculations support the result that water is the preferred solvent condition for the B form structure.

Results for A DNA in the two solvents, step II of the reaction scheme, are shown in Figure 2B. The differences in bonded and van der Waals interactions are small for the A form as for the B. The counterion electrostatics and solvation components are both reduced in magnitude, but the tendency of the former to be more favorable in 85% EtOH and the latter in water remains intact. Thus, for the A form of NaDNA, the intramolecular preference for the lower dielectric competes *successfully* with the solvation energy, making 85% EtOH the preferred solvent condition. The collective results presented in Figure 2 thus establish that our calculations are in accord with the expected conformational preferences: B form more stable in water and A form more stable in 85% EtOH.

Analysis of the relative stability of A and B forms of NaDNA in water and 85% EtOH (steps III and IV of the reaction scheme, respectively) is shown in Figure 3. In water (step III of the reaction scheme and Figure 3A), the DNA electrostatics favors B since phosphate repulsions are lower. In contrast, the counterion−DNA electrostatics strongly favor the A form. The reason for this can readily be seen from the MD snapshots in Figure 1. The equilibrium distribution of the counterions in A DNA, responding to the more compact A form structure, is also more compact. The resultant attractive CI-DNA attractions counterbalance the repulsive interphosphate and sodium−sodium repulsions. By contrast, solvation of NaDNA favors the B form, since the less compact counterion structure leaves the ions and DNA more exposed to hydration. Overall, the preference for B form over A form in water wins out due to a combination of reduced phosphate repulsions and improved hydration.

In 85% EtOH, step IV (Figure 3b), the tendencies of all the various terms remain intact, but their magnitudes are greatly diminished. The difference in the solvation term for A and B structures in 85% EtOH favors B, but only by a negligible

amount. DNA electrostatics favors the B form. The energetic effects of more compact counterion condensation favor A. The net result in 85% EtOH is a preference, albeit slight, for the A form. The magnitudes of the net free energy changes are small and subject to some uncertainties, but from the trends in the contributing factors we can successfully construct a conceptual basis for the observed behavior.

Examination of the results in Figures 2 and 3 shows that, in the component analysis, the net free energy is of the same order of magnitude as a number of the smaller contributions, and this raises a question about the interpretation of results in terms of one factor as opposed to another. However, Figures 2 and 3 also show that the large components—electrostatics and solvation—are also the most sensitive and variable to the structures and environmental effects as studied. Thus, we feel it is appropriate to place the origins of the explanation at the molecular level in these terms. This study also demonstrates that the concept of a "high-energy" form of DNA, as used in the literature,[28] depends on which part of the system one is examining, suggesting by implication how it may be altered by changes in parts of the system. For the polyanion, B DNA is the low-energy form by virtue of the lesser interphosphate repulsions. With counterions included in the system, A DNA is the lower energy form. In water, the free energy analysis shows B DNA to be the preferred form, with A DNA becoming preferred when the water activity is reduced. In the final analysis, the net result is a consequence of a fine balance of a number of competing terms—the electrostatics, the van der Waals attractions, the torsional flexibility, the entropies of counterions and DNA conformation, and solvation.

The ability of computations based on atomic models and the laws of physics to reproduce the trend in A/B conformational preferences successfully with no adjusted parameters is gratifying but is sensitive to assumptions. This study succeeds in the identification of the main contributing factors to the preferential stability of A and B forms of d(CGCGAATTCGCG) in solution and the development of a coherent, plausible explanation of the conformational preferences at the molecular level. Sequence effects, size dependence, salt, and other factors will introduce additional variables that must be considered in a more comprehensive treatment of the problem.

(28) Ivanov, V. I.; Krylov, D. *Methods Enzymol*. **1992**, *211*, 111.

IV. Conclusions

In summary, we find the molecular origins of the conformational preferences of A and B DNA in water and 85% EtOH to lie primarily in the *differential* free energy contributions from interphosphate repulsion, counterion condensation, and solvation. Lower interphosphate repulsions favor the B form independent of solvent conditions. Counterion−DNA interactions favor the A form as a consequence of its more compact structure and higher charge density, but in approaching closer to the A form than B, counterions pay a heavier desolvation penalty in the free energy. In water, solvation favors the B form NaDNA complex, with a magnitude that wins out in the balance of terms. In 85% EtOH, the magnitudes of these terms are reduced and the balance is shifted. At lower water activity, the solvation free energy is greatly reduced, and the free energy contribution originating in the more compact and energetically more favorable organization of counterions wins out and stabilizes the A form structure.

Acknowledgment. This research is supported by Grant No. GM-37909 (to D.L.B.) from the National Institutes of Health. B.J. acknowledges the financial support and encouragement received from the Council of Scientific and Industrial Research (CSIR) and the Indo-French Centre for the Promotion of Advanced Research. M.A.Y. is the recipient of a Graduate Traineeship in molecular biophysics under NIH Grant No. GM-08271, and D.S. is supported by NIH Postdoctoral Fellowship No. GM-18117.

JA981307P