doi:10.1006/jmbi.2000.4167 available online at http://www.idealibrary.com on IDEAL J. Mol. Biol. (2000) 304, 803–820

JMB

DNA Structure: What's in Charge?

Kevin J. McConnell and D. L. Beveridge*

Chemistry Department and Molecular Biophysics Program Wesleyan University, Middletown, CT 06459, USA

DNA structure is well known to be sensitive to hydration and ionic strength. Recent theoretical predictions and experimental observations have raised the idea of the intrusion of monovalent cations into the minor groove spine of hydration in *B*-form DNA. To investigate this further, extensions and further analysis of molecular dynamics (MD) simulations on d(CGCCGAATTCGCG), d(ATAGGCAAAAAATAGG-CAAAAATGG) and d(G_5-($GA_4T_4C)_2$-C_5), including counterions and water, have been performed. To examine the effective of minor groove ions on structure, we analyzed the MD snapshots from a 15 ns trajectory on d(CGCGAATTCGCG) as two subsets: those exhibiting a minor groove water spine and those with groove-bound ions. The results indicate that Na^+ at the ApT step of the minor groove of d(CGCCGAATTCGCG) makes only small local changes in the DNA structure, and these changes are well within the thermal fluctuations calculated from the MD. To examine the effect of ions on the differential stability of a *B*-form helix, further analysis was performed on two longer oligonucleotides, which exhibit A-tract-induced axis bending localized around the CpG step in the major groove. Plots of axis bending and proximity of ions to the bending locus were generated as a function of time and revealed a strong linear correlation, supporting the idea that mobile cations play a key role in local helix deformations of DNA and indicating ion proximity just precedes the bending event. To address the issue of ''what's in charge?'' of DNA structure more generally, the relative free energy of *A* and *B*-form d(CGCGAATTCGCG) structures from MD simulations under various environmental circumstances were estimated using the free energy component method. The results indicate that the dominant effects on conformational stability come from the electrostatic free energy, but not exclusively from groove bound ions *per se*, but from a balance of competing factors in the electrostatic free energy, including phosphate repulsions internal to the DNA, the electrostatic component of hydration (i.e. solvent polarization), and electrostatic effects of the counterion atmosphere. In summary, free energy calculations indicate that the electrostatic component is dominant, MD shows temporal proximity of mobile counterions to be correlated with A-track-induced bending, and thus the mobile ion component of electrostatics is a significant contributor. However, the MD structure of the dodecamer d(CGCGAATTCGCG) is not highly sensitive to whether there is a sodium ion in the minor groove.

Keywords: molecular Dynamics; DNA; Ions; Electrostatics. d(CGCGAATTCGCG); free energy

**Corresponding author*

Introduction

The sensitivity of DNA structure and conformational stability to hydration and salt effects is well established (Saenger, 1984). However, the deeper reasons behind this sensitivity and the observed conformational preferences of DNA at

Abbreviations used: CI, counterion; MD, molecular dynamics; NOESY, nuclear Overhauser enhancement spectroscopy.

E-mail address of the corresponding author: dbeveridge@wesleyan.edu

0022-2836/00/050803–18 $35.00/0

the molecular level are considerably less well understood. This issue has received particular attention in the recent literature, particularly with regard to the possible effects on DNA stability of the fractional occupation of counterions in the grooves (Chiu *et al.*, 1999; Denisov & Halle, 2000; Hud & Feigon, 1997; Hud *et al.*, 1999; McFail-Isom *et al.*, 1999; Young *et al.*, 1997a). We provide here a critical overview of the problem, and describe molecular dynamics (MD) simulations on the sequences d(CGCCGAATTCGCG), d(ATAGGCAAAAAA-TAGGCAAAAATGG) and d[G_5-($GA_4T_4C)_2$-C_5] in solution, which contribute further to understanding recent crystal structure results. Analysis of the results provides some additional ideas about the role of mobile counterions and electrostatic effects in general in determining the thermodynamic stability of the DNA double helix under various environmental circumstances.

Background

The Watson-Crick double helix, proposed in 1953, is now commonly referred to as the right-handed *B*-form DNA. *B*-DNA is the predominant structural form of DNA *in vivo*, and is well known to be preferentially stabilized under conditions of high water activity. The right handed *A*-form is favored under conditions of lower water activity, and the left-handed *Z*-form is favored by G + C-rich sequences at high salt concentrations. Thus DNA structure is observed to be highly sensitive to the environmental effects of solvent water, mobile cations necessary to achieve electroneutrality with anionic DNA, and excess salt. For the extensive evidence supporting these assertions, see the reviews by Saenger (1984), Westhof & Beveridge (1989) and Schneider & Berman (1995).

The DNA double helix from four independent crystal structures reported over the last 20 years for the *B*-form sequence d(CGCGAATTCGCG), the *Eco*RI dodecamer, is shown in Figure 1. The four DNA structures are superimposable within 0.436-0.685 Å, but show some small local differences. With respect to DNA solvation, the salient features from the crystal structures are ''cones'' of localized electron density around the phosphate groups and ordered peaks in the major and minor grooves of the structure. The characteristic cone patterns have been assigned to phosphate hydration (Pullman *et al.*, 1975; Schneider *et al.*, 1998). Further study of ions around the phosphate groups shows discrete patterns of Na^+, K^+, and Mg^{2+} (Schneider & Kabelae, 1998). A notable solvent feature, discovered in the first structure of a *B*-form oligonucleotide by Dickerson and co-workers (Drew & Dickerson, 1981), is the so-called spine of hydration extending throughout the central CAATTG tract of the minor groove. The spine of hydration has been considered a key feature stabilizing *B*-form DNA at high water activity, based in part on the observation that when *A*-form DNA is titrated with the antibiotic ligand netropsin, a minor groove binder, the structure of the DNA in the complex is pulled over into the *B*-form (Fritzsche *et al.*, 1984, 1992; Luck & Zimmer, 1973; Zimmer *et al.*, 1983).

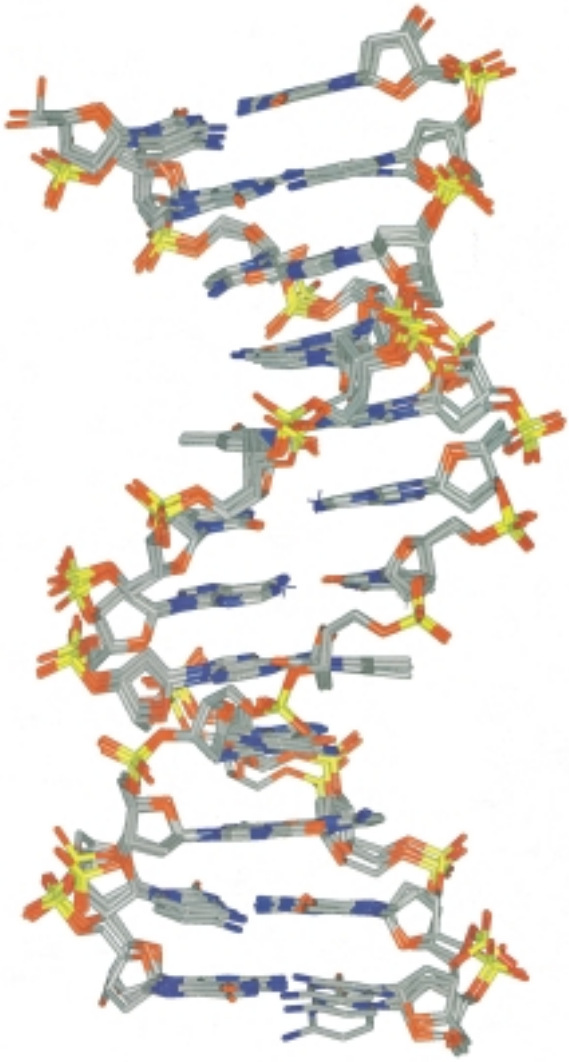

Figure 1. Superposition of four independent crystal structures for the *Eco*RI dodecamer (Chiu *et al.*, 1999; Shui *et al.*, 1998b; Tereshko *et al.*, 1999a; Wing *et al.*, 1980) The structures are all within 0.8 Å rmsd of each other.

Recently, a series of theoretical and experimental studies (Chiu *et al.*, 1999; Denisov & Halle, 2000; Hud & Feigon, 1997; Hud *et al.*, 1999; McFail-Isom *et al.*, 1999; Young *et al.*, 1997a) have generated the idea that the spine of hydration in *B*-DNA is not in fact all water, but should be described as a fractional occupancy problem, with counterions occasionally intruding and spending some fraction of time in particularly electronegative regions (pockets) in the grooves of DNA. This phenomenon was first noted in an early fiber diffraction

study of Cs^+ DNA system (Bartenev *et al*., 1983). Fractional occupation of ions in the DNA grooves has a number of possible implications, both structural and functional. From a structural point of view, an appreciable fractional occupancy of ions in the minor (or major) groove would effect phosphate repulsions different from intervening water molecules, and become a previously unanticipated source of sequence-dependent alterations in base-pair morphology and axis bending. The role of ions as bending loci has been described by Rouzina & Bloomfield (1998) and MD support for this hypothesis has been provided in GROMOS studies by Bonvin on Trp repressor DNA (Bonvin, 2000).

One particularly attractive site for ion occupancy in *B*-form DNA is the ApT step in the minor groove, which is a region of demonstrably low electronegativity (Lavery & Pullman, 1981). An early crystal structure of rApU, an RNA base-pair step analogous to dApT, showed a sodium ion in the minor groove (Seeman *et al*., 1976). The corresponding structure of rCpG showed none (Rosenberg *et al*., 1976). The idea of counterions ''intruding'' into the minor groove spine of hydration and the fractional occupancy problem was re-opened in 1995 by Young *et al*. (Young *et al*., 1997a,b) based on all-atom MD simulation, including water and counterions. Independent experimental evidence for ion occupancy in DNA grooves was presented by Hud and co-workers (Hud & Feigon, 1997; Hud *et al*., 1999) based on NMR spectroscopy. Both the Williams and Egli groups have recently reported ca 1.1-1.5 Å crystal structures on the *Eco*RI dodecamer in the presence of Na^+ (Shui *et al*., 1998a), K^+ (Shui *et al*., 1998b), Cs^+ (Tereshko *et al*., 1999b; Woods *et al*., 2000) and Rb^+ (Tereshko *et al*., 1999b) salts, and provide evidence for the extent of fractional occupation of ions in the minor groove. Egli and co-workers support this idea with a crystal structure of the Rb^+ salt, and noted that the minor groove exhibited a ''hydrat-ion'' spine. While the analysis for the Na^+ salt (Shui *et al*., 1998a), reporting ca 10% fractional occupancy of Na^+ at the ApT step, has been disputed (Chiu *et al*., 1999; Tereshko *et al*., 1999a), the effect is more pronounced in the case of the ions of higher atomic number. Recent NMR studies by Denisov & Halle (2000) have employed a new technique to estimate the residence time of Na^+ in the minor groove, and report an ion occupancy of ca 3% at the ApT step of the *Eco*RI dodecamer. The propensity of other base-pair steps for fractional occupancy of ions in both the minor and major grooves has been noted (Young *et al*., 1997a). With the idea of fractional occupation of ions in the grooves of DNA at electronegative pockets increasingly viable, the extent of ion occupancy in solution, the effect of ions on structure, and the relationship of both to free energy need to be addressed.

Recently, Williams and co-workers have reviewed diverse ways in which groove-bound ions can cause ''electrostatic collapse'' and narrowing of the minor groove by modulating anionic repulsions of DNA phosphate and might influence DNA structure and bending (Shui *et al*., 1998b). Evidence supporting this idea comes from earlier work by Maher (Strauss & Maher, 1994; Strauss-Soukup *et al*., 1998) showing that DNA sequences bend when the phosphate backbone is partially neutralized with proper phasing akin to phased A-track experiments. Williams *et al*. (McFail-Isom *et al*., 1999), taking the idea a step further, subsequently argued the case for ''cations in charge'' of DNA structure. In response, Dickerson and co-workers reported on the structure of a cross-linked *B*-form DNA oligonucleotide crystal prepared under conditions where no appreciable amount of monovalent cations was present (Chiu *et al*., 1999), only divalent cations. The structure turned out to be essentially identical with that obtained by Drew and Dickerson in 1981 (Drew *et al*., 1981). Since, presumably, no groove-bound monovalent cations are present in the cross-linked structure, they concluded ''cations, contrary to what has been claimed, are not in charge'' (of DNA structure) (Chiu *et al*., 1999). Two independently solved structures of the calcium salt of the *Eco*RI dodecamer actually crystallize in a different space group, *R*3 instead of $P2_12_12_1$ (Liu *et al*., 1998; Liu & Subirana, 1999; Minasov *et al*., 1999) In the *R*3 space group, the end to end interaction of the dodecamers are very different and the terminal residues are disordered as a consequence.

All arguments described above, for or against ''cations in charge'', are based solely on structure determinations. However, the question is not only one of structure but of the relationship between the structure and free energy. Here, we revisit the question of what's in charge of DNA structure, and clarify the proposals made on the basis of MD modeling about fractional occupation of ions in the minor groove spine. The sensitivity of *B*-form DNA fine structure to ion intrusion into the spine of hydration is examined based on an extended MD trajectory on d(CGCGCAATTCGCG) in solution. Analysis of the effect of counterions on the development of local axis deformation in sequences that exhibit A-tract-induced bending provides additional perspectives on the possible role of mobile counterions on helix deformations relative to *B*-form structure. Finally, we comment on the question of ''what's in charge?'' based theoretical estimates of various terms that contribute to conformational free energy of *B*-form and *A*-form DNA.

Results

Results from MD simulations: d(CGCGAATTCGCG)

The dynamical structure of the *Eco*RI dodecamer including solvent and ion density based on 15 ns of MD trajectory is shown in Figure 2. On extend-

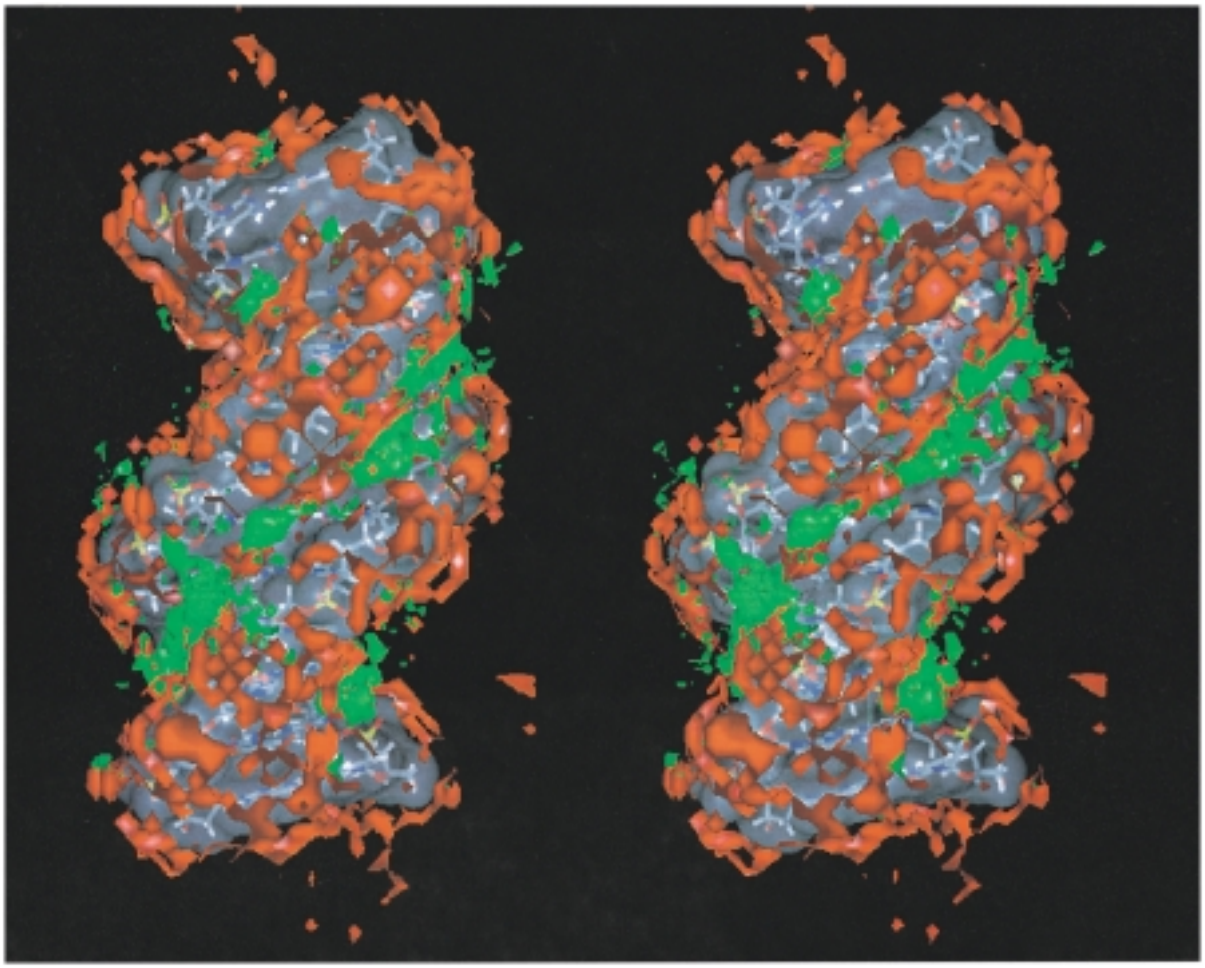

Figure 2. The average structure of the *Eco*RI dodecamer from MD simulation, surrounded by calculated solvent density both normalized to their respective bulk density. Water density, blue; ion density, green.

ing the trajectory from 5 to 15 ns, the dynamical structure of the DNA was observed to remain virtually unchanged, indicating the MD structure to be oscillating in a bound state clearly identifiable as *B*-form DNA. The root-mean-square deviation (rmsd) based on all heavy-atoms of the average MD structure was 2.8 Å from canonical *B*-form DNA and 2.6 Å from the Drew-Dickerson crystal form. The MD structure resides 2.4 Å rms from a recent NMR structure determination for the *Eco*RI sequence in solution based on 2D nuclear Overhauser spectroscopy (2D-NOESY). We have carried out a back-calculation of the 2D-NOESY intensities from the MD structure (McConnell *et al*., 2000a), and found close accord with those observed from NMR, supporting the AMBER MD model as a good representation of DNA structure in solution. The local dielectric behavior of the solvent in the vicinity of the DNA based on the 15 ns trajectory has been calculated using Kirkwood-Grunwald theory (Young *et al*., 1998), and shows a pattern quite similar to that observed experimentally (Jin & Breslauer, 1988); see also Lamm & Pack (1997).

One of the main issues to be addressed here is the effect of fractional occupancy of ions on the DNA structure, particularly in the minor groove, and the effect of counterions in general on thermodynamic stability. The MD solvent density (Figure 2) is composed from individual snapshots from the trajectory. A preponderant number of these snapshots show an intact spine of hydration in the minor groove of DNA (Figure 3, additional water molecules removed for clarity). The MD water spine is in close accord with the spine of hydration described in the original crystal structure by Drew & Dickerson (1981). However, as noted above, some 5-10 % of the MD structures show sodium ions intruding into the AATT region of the minor groove spine of hydration, as shown for the ApT step in Figure 4 (water molecules removed for clarity). The MD counterion density from the first shell was used to calculate the fractional occupancy of sodium ions at each of the base-pair steps in the major and minor grooves with respect to the 15 ns of MD trajectory. The assignment of ions to nucleotide base atoms is based on the proximity method (Mehrotra & Beveridge, 1980; Mezei & Beveridge,

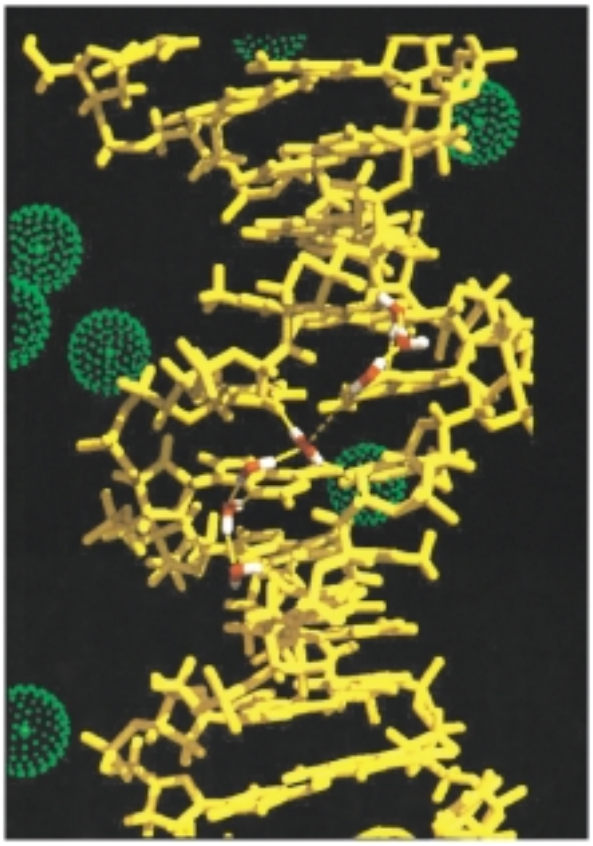

Figure 3. Snapshot from the MD of the *Eco*RI DNA showing an intact spine of hydration.

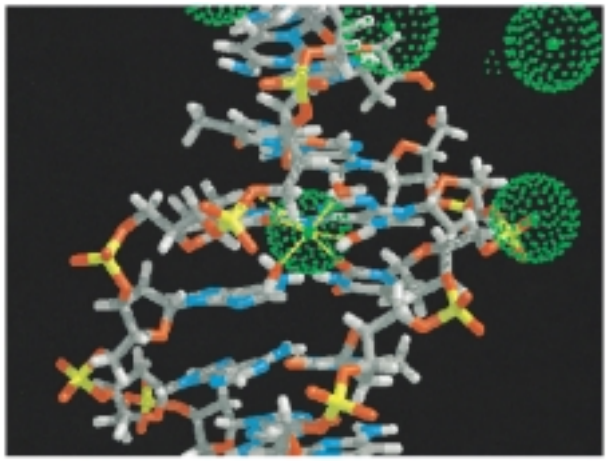

Figure 4. Snapshot from the MD on the *Eco*RI DNA showing a sodium ion intruding into the minor groove spine of the ApT step.

1986). What is actually defined as the region of a base-pair or base-pair step (and thus the results) depends on this definition, and the calculated value of fractional occupancy is best interpreted as indicating ''in the region of'' rather than precise positions. The calculated fractional occupancies are shown Figure 5. Sequence termini are to be ignored due to end effects. For the minor groove, we find 5-10 % fractional occupancy in the vicinity of the AATT region of the sequence. The calculated fractional occupancy is in reasonable accord with the reports based on the crystallography (Shui *et al*., 1998b) and NMR spectroscopy (Denisov & Halle, 2000). In the major groove, we find greater than 10 % fractional Na^+ occupancy in the vicinity of the CpG step. L. Williams (personal communication) informs us that this corresponds well with new results just obtained on the crystal structure of the Th^+ salt of d(CGCGAATTCGCG).

Further analysis was pursued by dividing the MD snapshots into two classes: water spine, and ion bound (defined as having an ion <2.9 Å from a minor groove base atom). The helicoidal and groove width values of the two classes are then compared. The results are presented in Figures 6-8. The calculated distribution of DNA minor groove widths at the ApT base-pair step for water spine and ion-bound structures is shown in Figure 6(a). The results indicate the distribution of minor groove widths to be slightly displaced in the ion-bound forms compared with water spine structures. To examine if this is statistically significant, a plot of the MD average values of groove widths as a function of sequence for ion-bound and water spine snapshots was calculated, with the corresponding thermal fluctuations about the mean in the dynamic structure displayed as vertical bars of one standard deviation. The results (Figure 7) show the MD minor groove width (Figure 7(a)) to be greater in the CGCG region and less in the AATT region for both water spine and ion-bound structures, with an incipient narrowing 5′ to 3′ in the ApA region in the sequence. Note that this is a well-documented property of A-tracts in general (Shafer *et al*., 1989) and is discussed from an MD perspective elsewhere (Sprous *et al*., 1999; Young & Beveridge, 1998). However, the results in

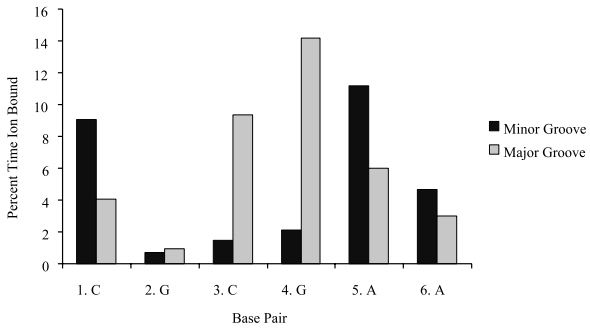

Figure 5. MD calculated fractional occupancy of the ions in the major and minor groove base-pairs of the *Eco*RI dodecamer; palindromic symmetry was invoked in calculating the fraction.

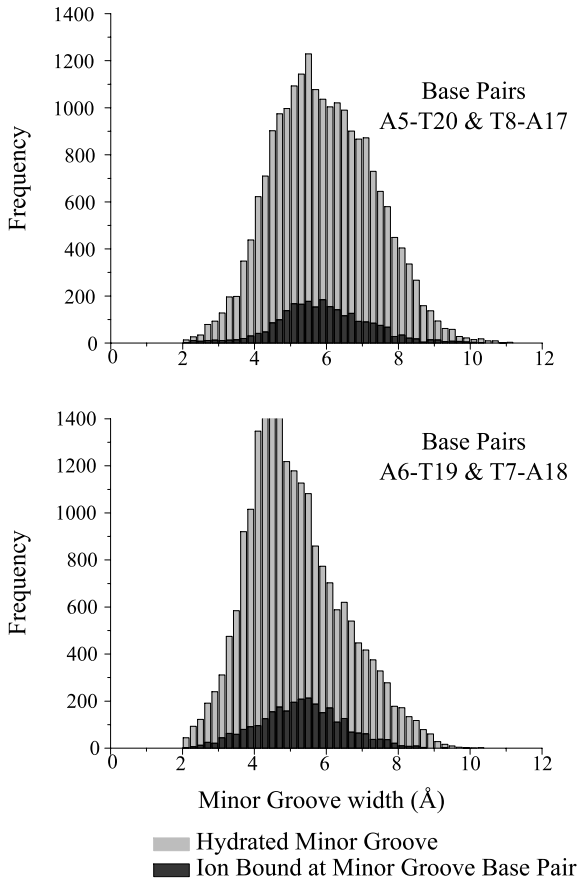

Figure 6. MD calculation of the distribution of groove width for MD structures with intact water spine and ion bound for (a) A5-T20 base-pair plus T8-A17 base-pair and (b) A6-T19 base-pair plus T7-A18 base-pair.

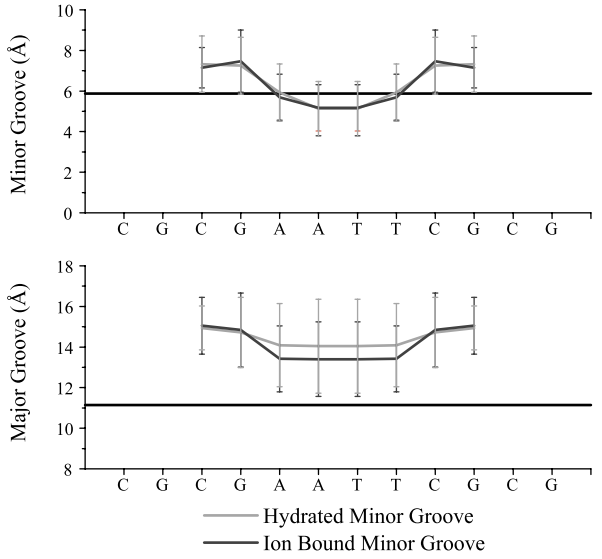

Figure 7. MD average values of groove width as a function of sequence for water spine and ion bound in the minor groove for a given base-pair of the *Eco*RI DNA.

Figure 7(a) indicate that any incipient difference between the minor groove width for ion-bound and water spine structures in the MD model of *Eco*RI DNA is well within the thermal fluctuations in groove width at room temperature. The calculated major groove width (Figure 7(b)) is correspondingly greater in the AATT region compared with the CGCG, and narrows in going from water spine to ion-bound structures. The differences are also well within the thermal fluctuations of the MD but the trend indicated in this graph may be significant. The effect of ions on the selected helicoidal properties of the DNA structure is shown in Figure 8. Here, calculated values of roll, tilt and twist for minor groove water spine and ion-bound forms are presented as a function of sequence. The results, as in the case of the major and minor groove widths, show that there is a slight effect of ions on the calculated mean values, but the difference is well within one standard deviation of thermal fluctuations. In particular, we note a slight reduction in helix twist at the ApT step of ion-bound structures.

The MD calculated fractional occupations in Figure 2 indicate that favorable positions in the major groove are also worthy of consideration. The prime example is the CpA step as shown in the MD snapshot (Figure 9). The effect of ions residing in the vicinity of the CpG step of the major groove is shown in Figure 10. The ion-bound forms display a slight narrowing effect on the major groove, but never beyond one standard deviation on thermal fluctuation. There is a significant difference in major groove width when an ion is bound in the major groove ApA and ApT steps, the closest thing we have seen to an example of electrostatic collapse. Again, a reduction in helix twist is observed at the CpG step in ion-bound forms.

Results from MD simulations: d(ATAGGCAAAAAATAGCCAAAAATGG) and d[GGGGG(GA_4T_4C)$_2$CCCCC]

Snapshots representative of the MD simulations on sequences exhibiting A-tract-induced axis bending are shown in Figure 11. The MD exhibit spontaneous axis bending in a concerted direction, with an average magnitude of -16.5° per A-tract (ca 33° overall). This result compares qualitatively with the axis bending anticipated for A-tracts phased by a full helix turn, and bending per turn of approximately 17°-21° inferred from cyclization experiments. The MD models exhibits a distinct, progressive 5′ to 3′ narrowing of the minor groove, a feature inferred from extensive results from DNA footprinting studies by Tullius and co-workers (Burkhoff & Tullius, 1988) of phased A-tract

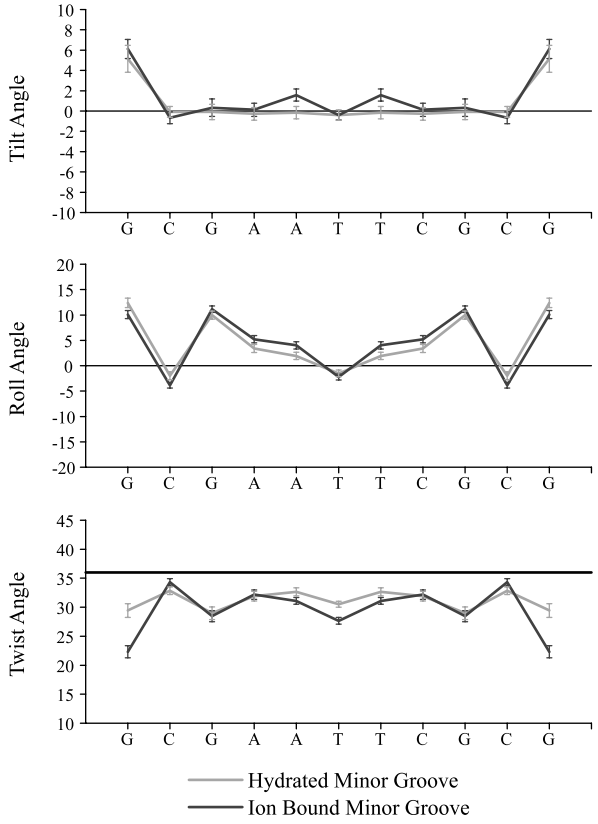

Figure 8. MD average values of helicoidal tilt, roll and twist as a function of sequence for minor groove high trait and minor groove ion-bound base-pair steps of *Eco*RI DNA.

sequences, whereas MD structures of control sequences lacking phased A-tracks did not show this effect. Analysis of the MD results supports a bending model with essentially straight, relatively rigid A-tracts and axis deformations in the non-A-tract regions. Bending is not localized at junctions of A-tracts but extends somewhat into the flanking sequences as well. The direction of helix curvature is toward the major groove with respect to non A-tract regions and towards the minor groove from the point of view of A-tracts, consistent with the results of gel retardation experiments (Koo *et al.*, 1986; Zinkel & Crothers, 1987). Both sequences exhibit a single bending locus in MD, and are thus propitious cases to examine the effect of ions on differential changes in helix structure relative to *B*-form DNA.

A plot of axis bending and the proximity of mobile counterions to the bending locus calculated as a function of time from the MD on d[G_5-(GA_4T_4C)$_2$-C_5] is shown in Figure 12(a). The data, as expected, are quite noisy, and correlations difficult to discern. We proceeded to smooth the data by Fourier transformation, and obtained the results shown in Figure 12(b). Here, one can discern a relationship between the occurrence of extrema in the time-series for axis bending and ion proximity. A plot of the time of occurrence of extrema in axis bending *versus* the time of occurrence of extrema in the ion proximity time-series is shown in Figure 13. Here, we find a strong correlation between occurrence of axis bending and the proximity of mobile counterions, supporting the hypothesis that cations can effect DNA structure. Note that the linear relationship in Figure 13 does not pass through the origin but is displaced slightly, indicating that the extrema in ion proximity just precede that of axis bending. The MD model is thus consistent with the idea that ion proximity is the cause and axis bending the effect in these systems.

Results from free energy component analysis

We next consider the influence of hydration and counterions on the conformational stability of the DNA. In previous work (Jayaram *et al.*, 1998b), we calculated the relative stability of *A* and *B*-forms of DNA under conditions of relatively high and low water activity, represented as a dilute aqueous solution of DNA and the DNA in a mixed solvent system of 85% (v/v) ethanol/water, respectively. Here, free energy component analysis was applied to MD snapshots from various simulations, treating the DNA and its complement of sodium counterions explicitly using the theory and methodology described in the Appendix. From these calculations, we estimated the relative conformational free energies of *B* and *A*-form DNA in water and 85% ethanol, and decomposed the results into contributions from sources intrinsic to the Na DNA complex and solvation. The calculations were shown to reproduce the conformational preferences of *B*-form in water, and of the *A*-form in 85% ethanol. Since the issue of ''what's in charge?'' can be addressed in terms of the conformational preference of *B* with respect to *A*-form in water, and *A* with respect to *B*-form in 85% ethanol, we performed further analysis of these calculations with special attention to the effects of ions and electrostatics.

The results are given in Figure 14. The calculated free energy can be decomposed into free energy contributions intrinsic to the Na DNA (valence, van der Waals and electrostatics of counter ion (CI) and DNA) and the free energy of solvation (electrostatic and non-electrostatic contributions). The essential trend in the net results, *B*-form preferred at high water activity and *A*-form at lower water activity, is recovered correctly. Examination of the component analysis shows that DNA electrostatics favors the *B*-form over the *A*-form. This comes about because phosphate separations are ca 0.7 Å greater in the *B*-form and the anion-anion repulsions are lower. The electrostatic free energy originating in DNA cation interactions favors the *A*-form, since the *A*-form is a more compact structure and the attractions are greater, more than offsetting the unfavorable cation-cation repulsions. Internal entropy effects of the *A*-form relative to

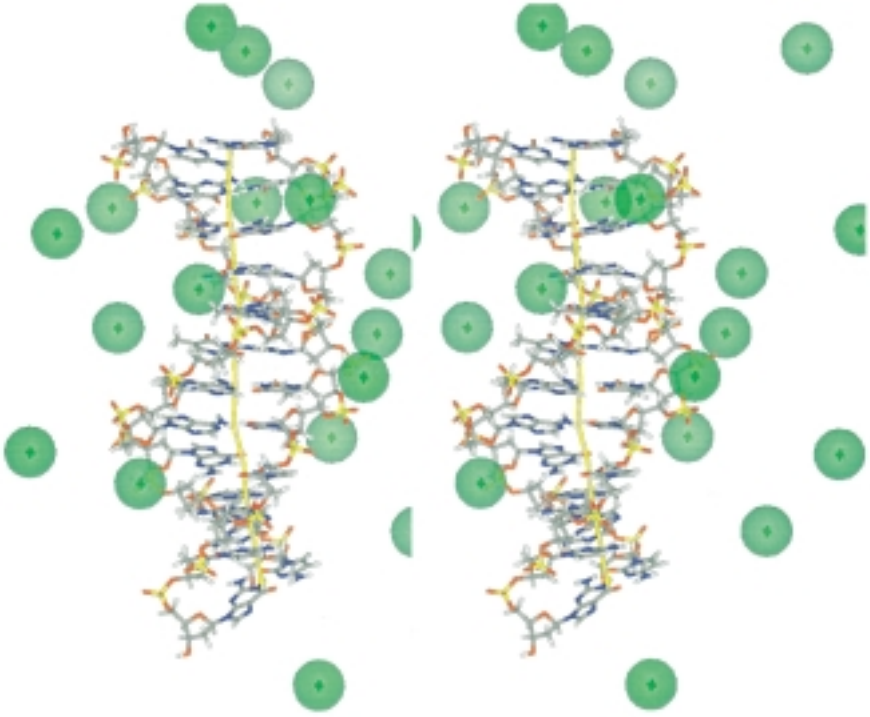

Figure 9. Picture of snapshot of the ion bound at the CG step of the major groove of *Eco*RI DNA.

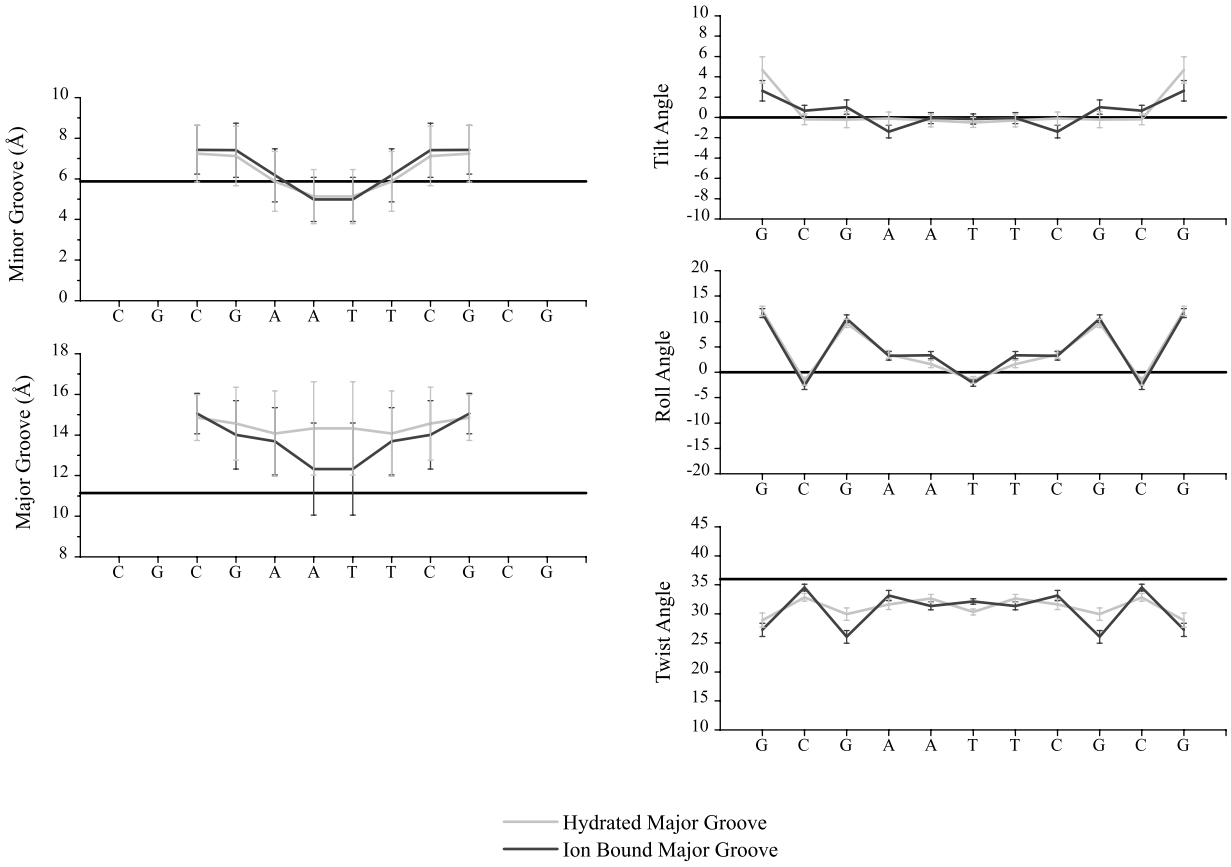

Figure 10. MD average values of groove width and helical tilt, roll and twist as a function of sequence for major groove high trait and major groove ion bound base-pairs of *Eco*RI DNA.

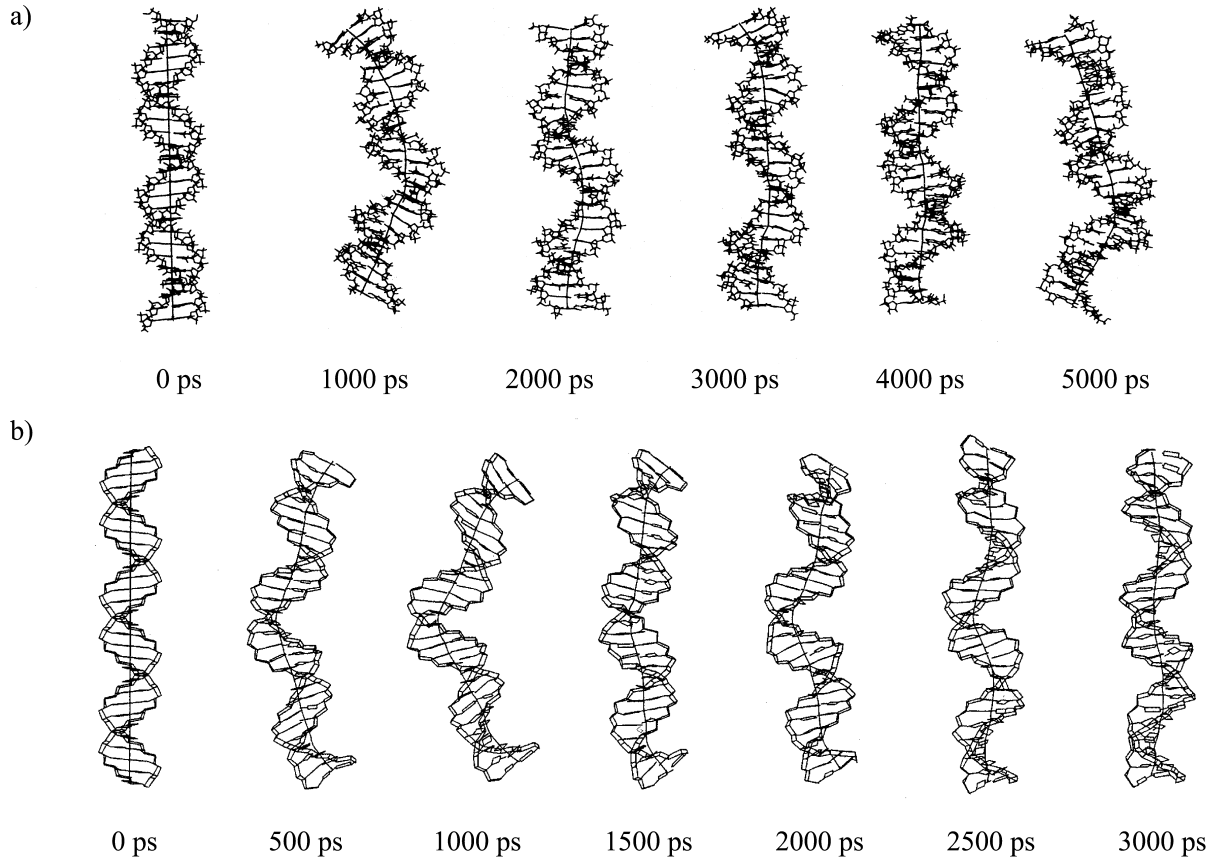

Figure 11. Snapshots of DNA structures from MD simulations of (a) d(ATAGGCAAAAAATAGGCAAAAATGG) (Young & Beveridge, 1998) and (b) d[G_5-($GA_4T_4C)_2$-C_5] (Sprous *et al.*, 1999) including counterions and water.

the *B*-form, treated with quasiharmonic theory (Karplus & Kushick, 1981), make only minor contributions to the net free energy difference. Examining the components associated with the thermodynamics of hydration of Na DNA shows that the electrostatic term (solvent polarization) favors the *B*-form, since the charges are more exposed. The preference of DNA for the *B*-form at high water activity arises as the internal electrostatics and hydration more than offset the tendency of the sum total of counterion interactions to favor the *A*-form.

The crux of the stability issue comes from comparing the MD results on d(CGCGAATTCGCG) in water and in 85% ethanol (Figure 14(a) and (b)), respectively. Intrinsic electrostatic effects are reduced in going from water, the high dielectric medium, to 85% ethanol, the lower dielectric. The reasons behind this can be appreciated by examining several representative snapshots of Na DNA in water (Figure 15(a)), and in ethanol (Figure 15(b)). These structures indicate that the more compact *A*-form, with anionic charge from phosphate groups more concentrated, draws corresponding complex of counterions in much more tightly. The net electrostatic free energy arising from interactions intrinsic to the DNA is reduced in magnitude in the *A*-form as compared with the *B*-form, since the anionic sites are closer. The counterion interactions (attraction and repulsion) are lower for the *A*-form in ethanol compared with the *B*-form in water. While the effects for the *B*-form in water and the *A*-form in ethanol are parallel, the magnitude of the various contributions are altered in the two cases, at least as described by this model. As a consequence, the model predicts that *B*- form DNA is preferentially stabilized in water, and the *A*-form is preferred under conditions of lower water activity.

The essence of the results from free energy component analysis is that the electrostatic effects of DNA, the hydration and the distribution of mobile counterions are all key players in determining net conformational stability. The answer to question of ''what's in charge?'' is thus (for this model) ''electrostatic free energy.'' The hypothesis of ''cations in charge'' is certainly a component of this quantity. However, we do not find the role of specific site-bound or fractionally resident cations to be dominant in the preferential stabilization of *A versus B* DNA. Our studies indicate that thermodynamic stability is dominated by the electrostatic component of free energy, arising as a consequence of a balance of contributions originating in anion-anion repulsions in the DNA, the electrostatics of hydration, and the cation-anion attractions and cat-

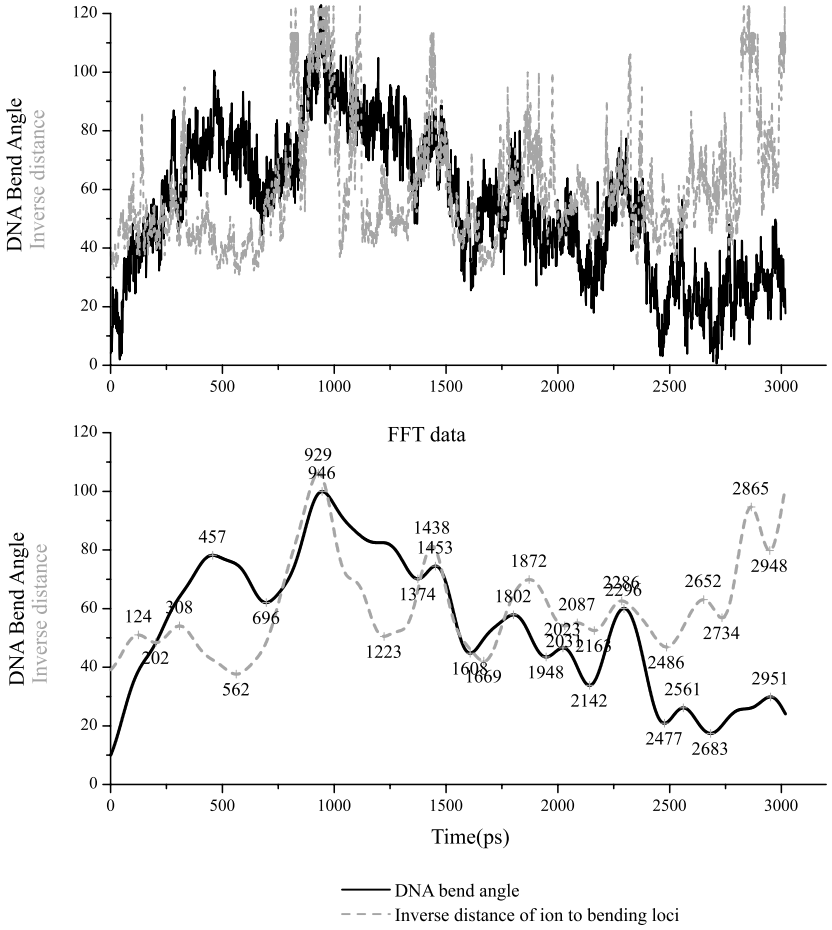

Figure 12. Axis bending and ion proximity as a function of time from MD on duplex d[G_5-(GA_4T_4C)$_2$-C_5], including counterions and water.

ion-cation repulsions in the counterion atmosphere of the molecule.

Discussion

Analysis of MD on d(CGCGAATTCGCG) shows that the idea of fractional occupation of ions in the grooves of DNA is viable but the extent of fractional occupation is predicted to be only around 10-15 % for the case of sodium in the major or minor groove. However, our estimates of fractional occupancy are based on only 15 ns of MD trajectory, and should thus be considered only a preliminary result. A corresponding MD study of the *Eco*RI dodecamer in the $P2_12_12_1$ crystalline unit cell shows that sodium ions remain bound at the ApT step throughout the trajectory; this study will be described further elsewhere (McConnell *et al*., 2000b). The difference in DNA structure for MD snapshots of exclusively water spine structure *versus* exclusively ion-bound forms shows in incipient structural electrostriction around the ions, but not to the extent that we would term it electrostatic collapse; minor perturbation is a more accurate description. Moreover, the extent of change the ions make on groove widths and helicoidal parameters of DNA is indicated to be well within the thermal fluctuations in the structure. Thus we find that the fact that the crystal structure of the cross-linked dodecamer determined by Chiu *et al*. (1999), under conditions where no appreciable amount of monovalent cations is present, remains essentially the same as that reported earlier for the *Eco*RI dodecamer by Drew and Dickerson (Drew *et al*., 1981), and is consistent with our result that even groove-bound Na^+ makes only a minor perturbation. Thus, their result cannot be interpreted as unequivocal evidence that ''cations.... are not in charge. '' Furthermore, in the absence of monovalent ions, the structure requires a cross-link to be stable, and ions many indeed be important to stability of an unmodified form.

With respect to the deformation of longer DNA sequences from *B*-form structures, analysis of the MD results on d(ATAGGCAAAAAATAGG-CAAAAATGG) and d[G_5-(GA_4T_4C)$_2$-C_5] shows a strong correlation between ion proximity and axis bending, and leading evidence that the arrival of a mobile counterion at a point in the sequence amenable to bending is a causative event. This supports the idea that cations are substantively involved in local stability of the DNA double helix.

From the free energy component analysis of the conformational stability of *A*-form and *B*-form

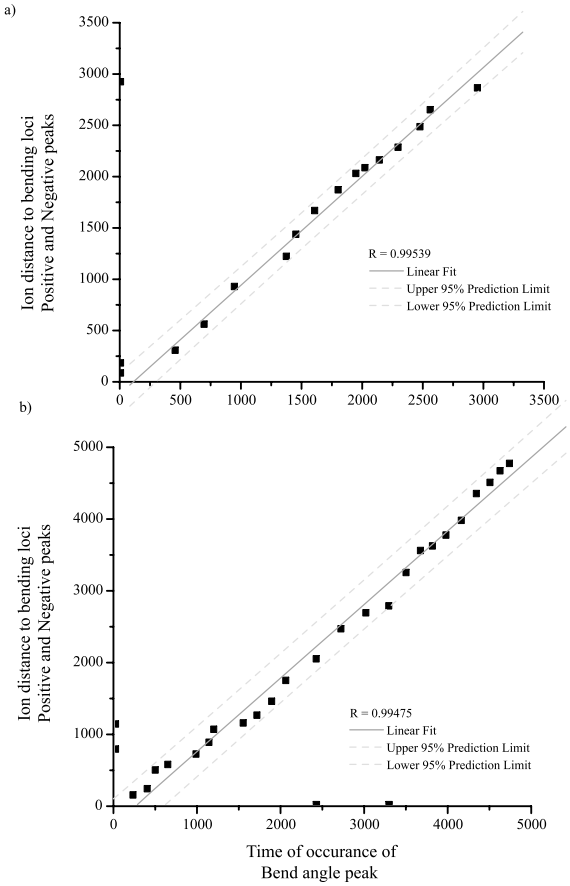

Figure 13. Scatter plots of occurrence of extrema in the time-series for axis bending and the time series for ion proximity to the bending locus: (a) from MD on d(ATAGGCAAAAAATAGCICAAAAATGG) and (b) from MD on $d[G_5\text{-}(GA_4T_4C)_2\text{-}C_5]$.

DNA, we extract ultimately a fairly simple picture. The electrostatic free energy is the dominant factor. The DNA electrostatics favor the *B*-form over the *A*-form because the anionic charges are further apart, and solvation free energy of the Na DNA complex is lower, since the charges are more exposed and polarize the solvent better than in the *A*-form. Counterion contributions favor the more compact *A*-form in which the ion atmosphere is correspondingly more compacted, effectively modulating the stronger anionic repulsions of the phosphate groups. In going from high water activity to lower water activity, essentially the same terms contribute to stability but their relative magnitudes are altered, and the *A*-form wins out in the balance of terms. In all cases, the electrostatic free energy is the responsible party, but no one term associated with either the DNA or the counterions dominates and can be considered ''in charge.'' The anionic charges on the DNA, the electrostatic polarization of solvent, and effects originating in the organization of the counterion atmosphere all are players in the relative stability of *A versus B*, not just fractional occupancy of ions in the grooves.

The reliability of the MD results from the AMBER/Parm 95 simulation protocol has been documented elsewhere (Beveridge & McConnell, 2000). One effect of note is a tendency towards 3-4° of underwinding in the dynamic structure of *B*-form DNA. We do not feel that this will materially affect the outcome of our analysis. The approximation of additivity of free energy components, critically considered in a recent article by Dill (1997), is possibly significant. Furthermore, to ''explain'' conformational changes, one naturally turns to examining physicochemical contributions such as valence, electrostatic, hydrophobic, van der Waals and entropic forces. This inevitably leads to a sum of rather a large number of terms, some large and some small, some favorable and some unfavorable with respect to a given process as written. The net free energy change arises as the resultant of these forces, i.e. as a small difference between often large opposing terms. The magnitudes of these terms in a computation are inevitably sensitive to free parameters. Thus claims of good agreement with experiment on a particular system are always suspect; reporting a large number of cases treated by a consistent and well-defined protocol carries much more theoretical credibility. Over and above this, the uncertainties in each of the terms, which may be large and in

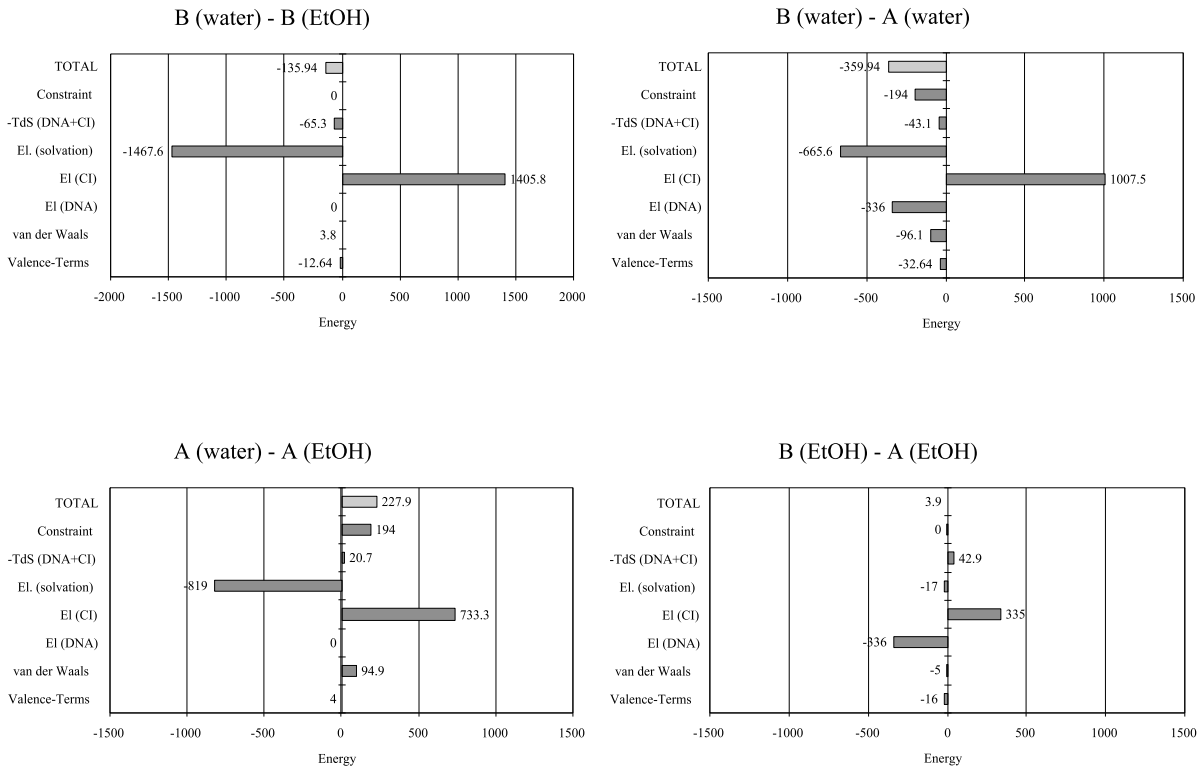

Figure 14. Calculated free energy component analysis for the *A* and *B*-forms of *Eco*RI DNA in water and 85 % ethanol. See the Appendix for definition of terms.

some cases (such as valence forces) difficult to estimate reliably at all, may propagate in such way as to make the magnitude and even sign of calculated net free energies correspondingly uncertain. Mitigating this effect in free energy component analysis is an offset of terms associated with the free energy of a macromolecular solute and the corresponding free energy of solvation. For relatively compact structures, the intrinsic free energy is more stabilizing than for more open forms. However, for the solvation free energies, the opposite is found, i.e. open structures are favored as a consequence of the increase in exposed surface areas. This effect is ubiquitous; it is found in electrostatic, van der Waals and hydrophobic contributions. If a particular computational scheme treats these effects in an internally consistent manner, some of the systematic uncertainties may cancel. However, one never knows how much, and *caveat emptor*. In summary, with regard to free energy component analysis, the assumptions and approximations in this computational schema render the calculations reliable only as a basis for examining trends. Given this, a clear implication of our results is that processes as complex as those we need to consider in molecular biophysics arise as a consequence of competing factors that are often easy to enumerate but difficult to assess reliably in terms of relative magnitudes.

Summary and Conclusions

The issue of ''what's in charge?'' of DNA structure has been investigated based on MD simulations and free energy component analysis. For d(CGCGAATTCGCG), MD predictions of fractional occupancies of sodium ions at the ApT step is approximately 5 %, neighboring ApA step shows ca 10 % Na^+ occupancy. Sodium ions in the minor groove make only small, local changes in minor groove width, and these changes are well within the thermal fluctuations in structure calculated from the MD. Another MD-predicted site of fractional Na^+ occupancy is the CpG step of the major groove, correlated with a local under-twisting of the helix. Ion occupancy in the AATT region of either groove results in little change in the width of the minor groove, but a narrowing of the major groove in this region. Analysis of the MD results on two 25-mer oligonucleotides that exhibit a single locus of A-tract-induced axis bending at CGC step in the major groove show that occurrence time of extrema in the axis bending *versus* ions at bending loci exhibit a strong linear correlation, supporting the idea that mobile cations play a key role in local helix deformations of DNA, as proposed by Williams and co-workers (Shui *et al.*, 1998b). The relative free energy of *A* and *B*-form d(CGCGAATTCGCG) structures from MD

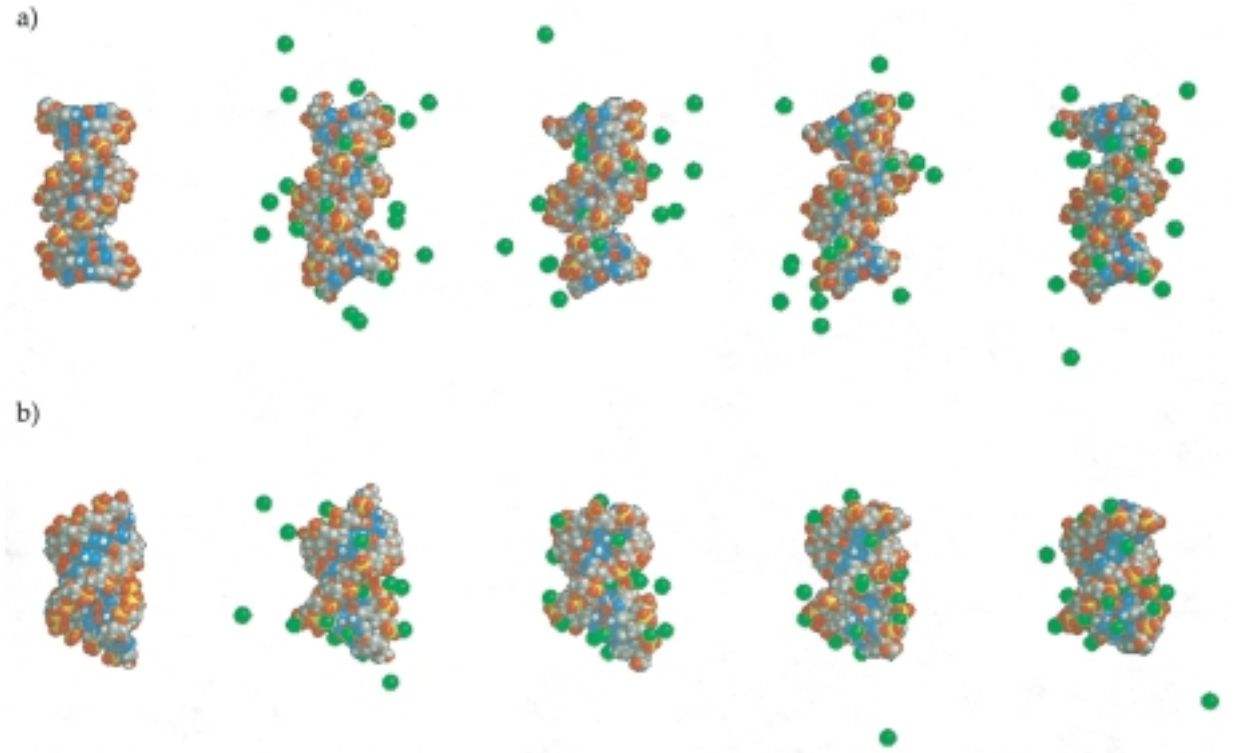

Figure 15. Snapshots from the in the simulation of (a) the *B*-form of *Eco*RI DNA in water, (b) the *A*-form of *Eco*RI DNA in ethanol.

simulations under various environmental circumstances estimated using the free energy component method (Jayaram *et al*., 1999) was subjected to detailed analysis. The results indicate that the dominant effects on conformational stability come from the electrostatic free energy. Further analysis reveals that this originates not exclusively from groove-bound ions *per se*, but from a balance of competing factors in the electrostatic free energy including phosphate repulsions internal to the DNA, the electrostatic component of hydration (i.e. solvent polarization), and electrostatic effects of the counterion atmosphere. In conclusion, we note that comprehensive theories of sequence effects on DNA structure so far (Calladine & Drew, 1997; El Hassan & Calladine, 1996; Suzuki *et al*., 1997) are generally derived from base-pair clash models, whereas we identify the operational quantity in stability to be electrostatic free energy with a significant solvent component originating, in part, from counterions.

Methods

The calculations involved in this study are: (a) 15 ns MD simulations on the *Eco*RI DNA dodecamer as described (Young *et al*., 1997b), which provide an ensemble of ''snapshots'' of DNA structures in solution under various environmental conditions; (b) MD trajectories carried out under a similar simulation protocol for two 25-mer oligonucleotide duplexes, d(ATAGGCAAAAAATAGGCAAAAATGG) (Young & Beveridge, 1998) and d[G_5-(GA_4T_4C)$_2$-C_5] (Sprous *et al*., 1999), both of which exhibit A-tract-induced axis bending; and (c) free energy component analysis (Jayaram *et al*., 1998b), a procedure applied *post hoc* to a representative subset of the MD snapshots. All MD simulations were carried out using the AMBER 5.0 suite of programs (Case *et al*., 1997) the parm95 AMBER force-field developed by Cornell *et al*. (1995), and the particle mesh Ewald (PME) treatment of long-range forces (Darden *et al*., 1993). The simulation cell in each case was comprised of canonical *B*-form of DNA (Arnott & Hukins, 1972), counterions, and TIP3P water molecules. In view of the possibility that the motions of solvent water and mobile ions are slower to stabilize in MD than DNA structure, our trajectory on the *B*-form d(CGCGAATTCGCG) duplex was extended from 5 ns to 15 ns. The dynamic structure of the DNA did not change appreciably between 5 and 15 ns of trajectory, but the stabilization of the dynamic properties of the solvent molecules improved considerably (Young *et al*., 1998). For simulations on d(CGCGAATTCGCG) at lower water activity, a RESP potential was used for the cosolvent, ethanol (Bayly *et al*., 1993). Details of the MD protocols have been described (Sprous *et al*., 1998). MD on *B*-form DNA in water and both *B*-form and *A*-form DNA in water and 85 % (v/v) ethanol solution are from Young *et al*. (1997b) and Sprous *et al*. (1998). The MD simulations on d(ATAGGCAAAAAATAGCICAAAAATGG) and d[G_5-(GA_4T_4C)$_2$-C_5] were configured in water and 10 mM Mg^{2+}, 50 mM K^+ and 70 mM Cl^- above simple neutralization of the DNA

backbone, solution conditions targeted to match a ligase buffer system of a particular series of phased A-tract experiments (Gartenberg & Crothers, 1991; Koo *et al*., 1986).

MD simulation based on all-atom models of DNA and solvent is computationally intensive, with each trajectory requiring several hundred hours of supercomputer time. Full free energy simulations, using thermodynamic integration or the perturbation method, require multiple trajectories and are thus expensive and not always readily interpretable (Beveridge & DiCapua, 1989; Kollman, 1993). Recently, several studies have explored a more tractable approach to estimating conformational free energy, based on calculations carried out a set of MD snapshots. In this approach, free energy is treated by component analysis (Jayaram *et al*., 1999), i.e. written as a sum of terms identified with the various chemical and thermodynamic forces on the DNA and its corresponding solvation. Internal energies are computed using the MD force-field (Cornell *et al*., 1995). The solvation free energy is treated by the method of generalized born-solvent accessibility (GBSA) (Still *et al*., 1990), which has been demonstrated to give reasonably good agreement with observed solvation energies for a large number of small molecules and ions, including prototypes of the sugars, phosphate ions and nucleotide bases of DNA (Jayaram *et al*., 1998a). Free energy component analysis applied to MD structures of DNA has been described recently by Srinivasan *et al*. (1998, 1999), and Jayaram *et al*. (1998b). In related work, Tsui & Case (2000) describe MD on *A* and *B*-forms of d(CCAACGTTGG) and r(CCAACGUUGG) duplexes in a generalized Born solvent model. Computational details and approximations inherent in free energy component analysis are relevant to a critical perspective on the arguments we advance here, a so further specifics and methodological details are provided in the Appendix.

Acknowledgments

The work of and discussions with Matthew A. Young, Dennis Sprous and Professor B. Jayaram are gratefully acknowledged. The authors thank Professors John Kuriyan and Barry Honig for an invitation to present material at the fall 1999 meeting of the New York Structural Biology Group at Rockefeller University, from which this article is derived. We acknowledge helpful conversations with Professors Loren Williams, Steve Harvey, Peter Kollman, Helen Berman and Haribabu Arthanari. This research is supported by NIH grant # GM37909 (to D.L.B.) and an NIH Traineeship in Molecular Biophysics, grant # GM 08271, (to K.J.M.).

References

Arnott, S. & Hukins, D. W. L. (1972). Optimized parameters for *A* and *B* DNA. *Biochem. Biophys. Res. Commun.* **47**, 1504-1510.

Bartenev, V. N., Golovamov Eu, I., Kapitonova, K. A., Mokulskii, M. A., Volkova, L. I. & Skuratovskii, I. (1983). Structure of the *B* DNA cationic shell as revealed by an X-ray diffraction study of CsDNA. Sequence-specific cationic stabilization of *B*-form DNA. *J. Mol. Biol.* **169**, 217-234.

Bayly, C. I., Cieplak, P., Cornell, W. D. & Kollman, P. A. (1993). A well behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **97**, 10269-10280.

Beveridge, D. L. & DiCapua, F. M. (1989). Free energy *via* molecular simulation: applications to chemical and biomolecular systems. *Annu. Rev. Biophys. Biophys. Chem.* **18**, 431-492.

Beveridge, D. L. & McConnell, K. J. (2000). Nucleic acids: theory and computer simulation, Y2 K. *Curr. Opin. Struct. Biol.* **10**, 182-196.

Bonvin, A. M. (2000). Localisation and dynamics of sodium counterions around DNA in solution from molecular dynamics simulation. *Eur. Biophys. J.* **29**, 57-60.

Burkhoff, A. M. & Tullius, T. D. (1988). Structural details of an adenine tract that does not cause DNA to bend. *Nature,* **331**, 455-457.

Calladine, C. R. & Drew, H. R. (1997). *Understanding DNA: The Molecule and How it Works*, Academic Press, San Diego, CA.

Case, D. A., Pearlman, D. A., Caldwell, J. W., Cheatham, T. E., III, Ross, W. S., Simmerling, C., Darden, T., Merz, K. M., Stanton, R. V., Cheng, A., Vincent, J. J., Crowley, M., Ferguson, D. M., Radmer, R. & Seibel, G. L., *et al*. (1997). *AMBER: Version 5*, 5.0 edit, University of California, San Francisco.

Chiu, T. K., Kaczor-Grzeskowiak, M. & Dickerson, R. E. (1999). Absence of minor groove monovalent cations in the crosslinked dodecamer C-G-C-G-A-A-T-T-C-G-C-G. *J. Mol. Biol.* **292**, 589-608.

Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Jr, Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W. & Kollman, P. A. (1995). A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **117**, 5179-5197.

Darden, T. A., York, D. M. & Pedersen, L. G. (1993). Particle mesh Ewald: an *N* log (*N*) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089-10092.

Denisov, V. P. & Halle, B. (2000). Sequence-specific binding of counterions to *B*-DNA. *Proc. Natl Acad. Sci. USA,* **97**, 629-633.

Dill, K. A. (1997). Additivity principles in biochemistry. *J. Biol. Chem.* **272**, 701-704.

Drew, H. R. & Dickerson, R. E. (1981). Structure of a *B*-DNA dodecamer. III. Geometry of hydration. *J. Mol. Biol.* **151**, 535-556.

Drew, H. R., Wing, R. M., Takano, T., Broka, C., Tanaka, S., Itikura, K. & Dickerson, R. E. (1981). Structure of a *B* DNA dodecamer. I: conformation and dynamics. *Proc. Natl Acad. Sci. USA,* **78**, 2179-2183.

El Hassan, M. A. & Calladine, C. R. (1996). Propeller-twisting of base-pairs and the conformational mobility of dinucleotide steps in DNA. *J. Mol. Biol.* **259**, 95-103.

Fritzsche, H., Rupprecht, A. & Richter, M. (1984). Infrared linear dichroism of oriented DNA-ligand complexes prepared with the wet-spinning method. *Nucl. Acids Res.* **12**, 9165-9177.

Fritzsche, H., Brandes, R., Rupprecht, A., Song, Z., Weidlich, T. & Kearns, D. R. (1992). The formation of *A*-DNA in NaDNA films is suppressed by netropsin. *Nucl. Acids Res.* **20**, 1223-1228.

Gartenberg, M. R. & Crothers, D. M. (1991). Synthetic DNA bending sequences increase the rate of *in vitro* transcription initiation at the *Escherichia coli* lac promoter. *J. Mol. Biol.* **219**, 217-230.

Hud, N. V. & Feigon, J. (1997). Localization of divalent metal ions in the minor groove of DNA *A*-tracts. *J. Am. Chem. Soc.* **119**, 5756-5757.

Hud, N. V., Sklenar, V. & Feigon, J. (1999). Localization of ammonium lons in the minor groove of DNA duplexes in solution and the origin of DNA *A*-tract bending. *J. Mol. Biol.* **286**, 651-660.

Jayaram, B., Sprous, D. & Beveridge, D. L. (1998a). Solvation free energy of biomacromolecules: parameters for a modified generalized born model consistent with the AMBER force field. *J. Phys. Chem.* **102**, 9571-9576.

Jayaram, B., Sprous, D., Young, M. A. & Beveridge, D. L. (1998b). Free energy analysis of the conformational preferences of *A* and *B* forms of DNA in solution. *J. Am. Chem. Soc.* **120**, 10629-10633.

Jayaram, B., McConnell, K., Dixit, S. B. & Beveridge, D. L. (1999). Free energy analysis of protein-DNA binding: the *Eco*RI endonuclease-DNA complex. *J. Comput. Phys.* **151**, 333-357.

Jin, R. & Breslauer, K. J. (1988). Characterization of the minor groove environment in a drug-DNA complex: bisbenzimide bound to the poly[d(AT)] poly[d(AT)]duplex. *Proc. Natl Acad. Sci. USA,* **85**, 8939-8942.

Karplus, M. & Kushick, J. N. (1981). Method for estimating the configurational entropy of macromolecules. *Macromolecules,* **14**, 325-332.

Kollman, P. A. (1993). Free energy calculations: applications to chemical and biochemical phenomena. *Chem. Rev.* **93**, 2395-2417.

Koo, H.-S., Wu, H.-M. & Crothers, D. M. (1986). DNA bending at adeninethymine tracts. *Nature,* **320**, 501-506.

Lamm, G. & Pack, G. R. (1997). Calculation of dielectric constants near polyelectrolytes in solution. *J. Phys. Chem. ser. B,* **101**, 959-965.

Lavery, R. & Pullman, B. (1981). The molecular electrostatic potential, steric accessibility and hydration of Dickerson's *B*-DNA dodecamer d(CpGpCpGpApApTpTpCpGpCpG). *Nucl. Acids Res.* **9**, 3765-3777.

Liu, J. & Subirana, J. A. (1999). Structure of d(CGCGAATTCGCG) in the presence of Ca^{2+} ions. *J. Biol. Chem.* **274**, 24749-24752.

Liu, J., Malinina, L., Huynh-Dinh, T. & Subirana, J. A. (1998). The structure of the most studied DNA fragment changes under the influence of ions: a new packing of d(CGCGAATTCGCG). *FEBS Letters,* **438**, 211-214.

Luck, G. & Zimmer, C. (1973). Interaction of netropsin with DNA in the course of the *B-A* transition. *Stud. Biophys.* **40**, 9-12.

McConnell, K. J., Arthanari, H., Young, M. A., Beger, R., Bolton, P. H. & Beveridge, D. L. (2000a). NMR assessment of molecular dynamics of DNA in solution: calculated and observed NOESY intensities and scalar coupling constants, vol. Submitted.

McConnell, K. J., Young, M. A. & Beveridge, D. L. (2000b). Molecular dynamics simulation on crystallographic unit cell of $d(CGCGAATTCGCG)_2$ comparison with predicted solution structures and studies of packing effects, vol. Submitted.

McFail-Isom, L., Sines, C. C. & Williams, L. D. (1999). DNA structure: cations in charge? *Curr. Opin. Struct. Biol.* **9**, 298-304.

Mehrotra, P. K. & Beveridge, D. L. (1980). Structural analysis of molecular solutions based on quasi-component distribution functions. Application to $[H_2CO]_{aq}$ at 25 °C. *J. Am. Chem. Soc.* **102**, 4287-4294.

Mezei, M. & Beveridge, D. L. (1986). Structural chemistry of biomolecular hydration: the proximity criterion. *Methods Enzymol.* **127**, 21-47.

Minasov, G., Tereshko, V. & Egli, M. (1999). Atomic-resolution crystal structures of *B*-DNA reveal specific influences of divalent metal ions on conformation and packing. *J. Mol. Biol.* **291**, 83-99.

Pullman, A., Berthod, H. & Gresh, N. (1975). Quantum mechanical studies of environmental effects on biomolecules. *Ab initio* study of the hydration of dimethylphosphate. *Chem. Phys. Letters,* **33**, 11-14.

Rosenberg, J. M., Seeman, N. C., Day, R. O. & Rich, A. (1976). RNA double-helical fragments at atomic resolution. II. The crystal structure of sodium guanylyl -3′,5′-cytidine nonahydrate. *J. Mol. Biol.* **104**, 145-167.

Rouzina, I. & Bloomfield, V. A. (1998). DNA bending by small mobile, multivalent cations. *Biophys. J.* **74**, 3152-3164.

Saenger, W. (1984). *Principles of Nucleic Acid Structure*, Springer Verlag, New York.

Schneider, B. & Berman, H. M. (1995). Hydration of the DNA bases is local. *Biophys. J.* **69**, 2661-2669.

Schneider, B. & Kabelae, M. (1998). Stereochemistry of binding of metal cations and water to a phosphate group. *J. Am. Chem. Soc.* **120**, 161-165.

Schneider, B., Patel, K. & Berman, H. M. (1998). Hydration of the phosphate group in double-helical DNA. *Biophys. J.* **75**, 2422-2434.

Seeman, N. C., Rosenberg, J. M., Suddath, F. L., Kim, J. J. P. & Rich, A. (1976). RNA double-helical fragments at atomic resolution. I. The crystal and molecular structure of sodium adenylyl-3′,5′-uridine hexahydrate. *J. Mol. Biol.* **104**, 109-144.

Shafer, G. E., Price, M. A. & Tullius, T. D. (1989). Use of the hydroxyl radical and gel electrophoresis to study DNA structure. *Electrophoresis,* **10**, 397-404.

Shui, X., McFail-Isom, L., Hu, G. G. & Williams, L. D. (1998a). The *B*-DNA dodecamer at high resolution reveals a spine of water on sodium. *Biochemistry,* **37**, 8341-8355.

Shui, X., Sines, C. C., McFail-Isom, L., VanDerveer, D. & Williams, L. D. (1998b). Structure of the potassium form of CGCGAATTCGCG: DNA deformation by electrostatic collapse around inorganic cations. *Biochemistry,* **37**, 16877-16887.

Sprous, D., Young, M. A. & Beveridge, D. L. (1998). Molecular dynamics studies of the conformational preferences of a DNA double helix in water and in an ethanol/water mixture: theoretical considerations of the *A/B* transition. *J. Phys. Chem.* **102**, 4658-4667.

Sprous, D., Young, M. A. & Beveridge, D. L. (1999). Molecular dynamics studies of axis bending in $d(G_5\text{-}(GA_4T_4C)_2\text{-}C_5)$ and $d(G_5\text{-}(GT_4A_4C)_2\text{-}C_5)$: effects of sequence polarity on DNA curvature. *J. Mol. Biol.* **285**, 1623-1632.

Srinivasan, J., Cheatham, T. E., III, Cieplak, P., Kollman, P. A. & Case, D. A. (1998). Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate - DNA helices. *J. Am. Chem. Soc.* **120**, 9401-9409.

Srinivasan, J., Trevathan, M. W., Beroza, P. & Case, D. A. (1999). Application of a pairwise generalized

Born model to proteins and nucleic acids: inclusion of salt effects. *Theor. Chem. Accts.* **101**, 426-434.
Still, W. C., Tempczyk, A., Hawley, R. C. & Hendrickson, T. (1990). Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **112**, 6127-6129.
Strauss, J. K. & Maher, L. J., III (1994). DNA bending by asymmetric phosphate neutralization. *Science,* **266**, 1829-1834.
Strauss-Soukup, J. K., Rodrigues, P. D. & Maher, L. J., III (1998). Effect of base composition on DNA bending by phosphate neutralization. *Biophys. Chem.* **72**, 297-306.
Suzuki, M., Amano, N., Kakinuma, J. & Tateno, M. (1997). Use of a 3D structure data base for understanding sequence-dependent conformational aspects of DNA. *J. Mol. Biol.* **274**, 421-435.
Tereshko, V., Minasov, G. & Egli, M. (1999a). The Dickerson-Drew *B*-DNA dodecamer revisited at atomic resolution. *J. Am. Chem. Soc.* **121**, 470-471.
Tereshko, V., Minasov, G. & Egli, M. (1999b). A ''hydrat-ion'' spine in a *B*-DNA minor groove. *J. Am. Chem. Soc.* **121**, 3590-3595.
Tsui, V. & Case, D. A. (2000). Molecular dynamics simulations of nucleic acids with a generalized Born solvation model. *J. Am. Chem. Soc.* **122**, 2489-2498.
Westhof, E. & Beveridge, D. L. (1989). Hydration of nucleic acids. In *Water Science Reviews: The Molecules of Life* (Franks, F., ed.), vol. 5, pp. 24-136, Cambridge University Press, Cambridge.
Wing, R. M., Drew, H. R., Takano, T., Broka, C., Tanaka, S., Itakura, I. & Dickerson, R. E. (1980). Crystal structure analysis of a complete turn of *B*-DNA. *Nature,* **287**, 755-758.
Woods, K. K., McFail-Isom, L., Sines, C. C., Howerton, S. B., Stephens, R. K. & Williams, L. D. (2000). Monovalent cations sequester within the *A*-tract minor groove of $[d(CGCGAATTCGCG)]_2$. *J. Am. Chem. Soc.* **122**, 1546-1547.
Young, M. A. & Beveridge, D. L. (1998). Molecular dynamics simulations of an oligonucleotide duplex with adenine tracts phased by a full helix turn. *J. Mol. Biol.* **281**, 675-687.
Young, M. A., Jayaram, B. & Beveridge, D. L. (1997a). Intrusion of counterions into the spine of hydration in the minor groove of *B*-DNA: fractional occupancy of electronegative pockets. *J. Am. Chem. Soc.* **119**, 59-69.
Young, M. A., Ravishanker, G. & Beveridge, D. L. (1997b). A 5-nanosecond molecular dynamics trajectory for *B*-DNA: analysis of structure, motions and solvation. *Biophys. J.* **73**, 2313-2336.
Young, M. A., Jayaram, B. & Beveridge, D. L. (1998). Local dielectric environment of *B*-DNA in solution: results from a 14 nanosecond molecular dynamics trajectory. *J. Phys. Chem. ser. B,* **102**, 7666-7669.
Zimmer, C., Marck, C. & Guschlbauer, W. (1983). *Z*-DNA and other non-*B*-DNA structures are reversed to *B*-DNA by interaction with netropsin. *FEBS Letters,* **154**, 156-160.
Zinkel, S. S. & Crothers, D. M. (1987). DNA bend direction by phase sensitive detection. *Nature,* **328**, 178-181.

Appendix

Conformational Free Energy Calculations *via* Component Analysis

The relationship of free energy component analysis to statistical mechanics and thermodynamics is described by Gilson *et al*. (1997). The ''dynamic structure'' we refer to is the ensemble of structures associated with a given familial form, i.e. *A*-form or *B*-form, of the DNA. Conformational preferences at a given temperature, pressure and solvent composition are reflected in the thermodynamics of the process:

$$A = B \tag{1}$$

Postulating conformational free energies, G_A and G_B:

$$\Delta\Delta G^0_{A\to B} = \Delta G^0_B - \Delta G^0_A \tag{2}$$

Estimates of the free energy of a DNA molecule as it assumes a particular dynamic structure of *A* or *B* DNA in solution were obtained using free energy component analysis. Component analysis assumes additivity, and that the free energy of a DNA conformation in solution can be effectively approximated as:

$$G^0 = G^0_{\text{int}} + g^0_{\text{solv}} \tag{3}$$

Here, we distinguish the free energy intrinsic/internal to the DNA and any other components treated explicitly in the model, such as counterions, as upper case (G) and solvation free energies with lower case (g). For the intrinsic free energy:

$$\Delta G^0_{\text{int}} = \Delta H^0_{\text{int}} - T\Delta S^0_{\text{int}} \tag{4}$$

where T is the absolute temperature. For the enthalpy term:

$$\Delta H^\circ_{\text{int}} \approx \Delta E_{\text{int}} \tag{5}$$

where E_{int} denotes energies from the Born Oppenheimer energy surface (BOES) of the molecule. Estimation of these energies can be made using the conventional expressions for an empirical forcefield (Leach, 1996):

$$E_{\text{int}} = E_{\text{val}} + E_{\text{es}} + E_{\text{vdW}} \tag{6}$$

Here, E_{val} is the valence term, and E_{es} and E_{vdW} are the electrostatic and van der Waals components, respectively. The valence term is taken as:

$$E_{\text{val}} = E_{\text{bonds}} + E_{\text{angles}} + E_{\text{dihedrals}} \tag{7}$$

with the various terms on the right-hand side of the equation describing the energy costs of bond stretching, angle bending and torsional displacements, respectively. The electrostatic and van der Waals components are:

$$E_{es} = \sum_{i<j} \frac{q_i q_j}{r_{ij}} \quad (8)$$

with q_i and q_j representing net atomic charges on atoms considered explicitly and r_{ij} is the interatomic separation, and:

$$E_{vdW} = \sum_{i<j} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \quad (9)$$

with ε and σ representing the van der Waals attractive energy and collision diameter, respectively. All calculations of the above quantities are carried out in this study using the AMBER parm.95 force-field as specified by Cornell *et al*. (1995). The intrinsic entropy terms *S* are calculated directly from the MD trajectories on DNA using the quasiharmonic method:

$$S_0 = \frac{1}{2}k_B \ln Q + \frac{1}{2}k_B(3N-6)(1+\ln 2\pi) \quad (10)$$

For the conformational free energy of salvation, we write:

$$\Delta g^0_{solv} = \Delta g^0_{es} + \Delta g^0_{nes} \quad (11)$$

where the terms on the right-hand side represent the electrostatic (el) and non-electrostatic (nel) contributions, respectively. The electrostatic free energy of solvation of a DNA conformation is estimated using the modified generalized Born (GB) method (Still, 1990) or Jayaram (1998):

$$\Delta g^0_{el} = \Delta g^0_{pol} + \Delta g^0_{scr} \quad (12)$$

where the f_{m2GB} are effective separations contain semi-empirical fitting parameters. In generalized Born theory, the salvation free energy can be decomposed further into contributions from solvent polarization:

$$g^0_{pol} = -166\left(1 - \frac{1}{\varepsilon}\right)\sum_{i=1}^{n}\frac{q_i^2}{\alpha_i} \quad (13)$$

and from the modulation of intramolecular electrostatics due to solvent screening:

$$g^0_{scr} = -166\left(1 - \frac{1}{\varepsilon}\right)\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{q_i q_j}{f_{m2GB}} \quad (14)$$

where the α_i and f_{m2GB} are related to radii of each of the atoms and include parameters fit with respect to experimental free energies on prototype cases (Jayaram, 1998). The development of the GBSA parameter set used in this study is described by Jayaram *et al*. (1998).

The non-electrostatic free energy of solvation is written as:

$$g^0_{nel} = \gamma_{nel}\Delta A \quad (15)$$

Where *SA* is solvent-accessible surface area and $\gamma = 7.2$ (Still, 1990). For the convenience of further analysis, we write:

$$\gamma_{nel} = \gamma_{vdW} + \gamma_{cav} \quad (16)$$

which permits separate estimates of the free energy of DNA cavitation and the free energy associated with DNA-solvent van der Waals interactions. In our calculations, following Jayaram *et al*. (1999), we set:

$$\gamma_{vdW} = +47 \text{ cal/A}^2 \quad (17)$$

$$\gamma_{cav} = -39 \text{ cal/A}^2 \quad (18)$$

Note that the net conformational free energy sees only γ; the further decomposition is introduced only for the purposes of analysis.

The treatment of free energy *via* component analysis, albeit approximate, has the material advantage of being readily decomposable into contributions of terms readily identifiable with physicochemical forces (valence, van der Waals, electrostatic and hydrophobic) for purposes of analysis and interpretation. For further perspectives and other current studies from this vantage point see Gilson (1997), Honig (Proloff, 1997), Jayaram *et al*. (1998, 1999) and the review of applications to ligand binding by Ajay & Murko (Ajay, 1995) and Case (Srinivasan, 1999).

The terms reported in the histograms in Figure 14 of the main text are defined and obtained as follows. The intramolecular energies in equation (6) are calculated from the ensemble averages and separately into bonded (bond, angle, dihedral), ΔH (bonded) and non-bonded van der Waals, ΔH(vdw) and electrostatic terms (which include the 1-4 contributions) ΔH (electrostatics-DNA). The electrostatics of equation (8) for counterions (CI) are grouped separately, ΔH(electrostatics-CI). These energy terms collectively provide an estimate of the intramolecular conformational enthalpy as described by equation (5). The corresponding intramolecular entropy contribution to the free energy, $-T\Delta S$ (quasiharmonic) is estimated for the ensemble of structures using the quasiharmonic method using equation (10). The experimentally determined absolute entropy of Na^+ in water of 9.1 cal mol^{-1} K^{-1} (excluding the electrostatic contribution to the entropy of solvation from its gas-phase value) (Friedman, 1973) translates to a *TS* ($T = 298$ K) of 2.7 kcal/mol for each counterion free in solution. After examining the entropies of Na^+ in water, ethanol, and crystals, (Krestov, 1991) we adopted a *TS* value of 2 and 3 kcal/mol for each ion free in water and 85% EtOH solution, respectively and the results are described in the $-T\Delta S$ (CI-release) term. 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 quasiharmonic entropies. The ensemble average is reported for the ΔG (electrostatic solvation) using the GBSA method as described in equation (14) and the ΔG(hydrophobic solvation) is determined based on equation (16). The ΔH(constraint) term is required for the *A*(water) simulation to account for the energy constraining the DNA in the *A*-form conformation when simulated in water. Without this constraint a simulation of *A*-form DNA in water is observed to undergo an *A* to *B* conversion as previously described.

References

Ajay, A. & Murcko, M. A. (1995). Computational methods to predict binding free energies in ligand-receptor complexes. *J. Med. Chem.* **38**, 4953-4967.

Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Jr, Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W. & Kollman, P. A. (1995). a second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **117**(19), 5179-5197.

Friedman, H. L. & Krishnan, C. V. (1973). Water - a comprehensive treatise. in water. *A Comprehensive Treatise* (Franks, F., ed.), vol. 3. Plenum Press, New York

Gilson, M. K., Given, J. A., Bush, B. L. & McCammon, J. A. (1997). The statistical thermodynamic basis for computation of binding affinities: a critical review. *Biophys. J.* **72**, 1047-1069.

Jayaram, B., Liu, Y. & Beveridge, D. L. (1998a). A modification of the generalized born theory for improved estimates of solvation energies and pKa shifts. *J. Chem. Phys.* **109**, 1465-1471.

Jayaram, B., Sprous, D. & Beveridge, D. L. (1998b). Solvation free energy of biomacromolecules: parameters for a modified generalized born model consistent with the AMBER force field. *J. Phys. Chem.* **102**(47), 9571-9576

Jayaram, B., Sprous, D., Young, M. A. & Beveridge, D. L. (1998c). Free energy analysis of the conformational preferences of A and B forms of DNA in solution. *J. Am. Chem. Soc.* **120**(41), 10629-10633.

Jayaram, B. McConnell, K., Dixit, S. B. & Beveridge, D. L. (1999). Free energy analysis of protein-DNA binding: the *Eco*RI endonuclease - DNA Complex. *J. Comput. Phys.* **151**(1), 333-357.

Krestov, G. A. (1991). *Thermodynamics of Solvation*, Ellis-Horwood, New York.

Leach, A. R. (1996). *Molecular Modeling: Principles and Applications*, Addison Wesley Longman Ltd, Essex, UK.

Proloff, N., Windemuth, A. & Honig, B. (1997). On the calculation of binding free energies using continuum methods: application to MHC Class I protein - peptide interactions. *Protein. Sci.* **6**, 1293-1301.

Srinivasan, J., Trevathan, M. W., Beroza, P. & Case, D. A. (1999). Application of a pairwise generalized Born model to proteins and nucleic acids: inclusion of salt effects. *Theoret. Chem. Accounts*, **101**(6), 426-434.

Still, W. C., Tempczyk, A., Hawley, R. C. & Hendrickson, T. (1990). Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **112**, 6127-6129.

Edited by B. Honig

(Received 5 May 2000; received in revised form 1 September 2000; accepted 5 September 2000)