JCTC Journal of Chemical Theory and Computation

Electrostatic Polymer Condensation and the A/B Polymorphism in DNA: Sequence Effects

Alexy K. Mazur*

CNRS UPR9080, Institut de Biologie Physico-Chimique, 13, rue Pierre et Marie Curie, Paris 75005, France

Received October 1, 2004

Abstract: Dynamics of the polymorphic A↔B transitions in DNA is compared for two polypurine sequences, poly(dA).poly(dT) and poly(dG).poly(dC), long known to exhibit contrasting properties in experiments. In free molecular dynamics simulations reversible transitions are induced by changing the size of a water drop around DNA neutralized by sodium ions. In poly(dG).poly(dC) the B↔A transitions are easy, smooth and perfectly reversible. In contrast, a B→A transition in a poly(dA).poly(dT) dodecamer fragment could not be obtained even though its A-form is stable under low hydration. An intermediate range of hydration numbers is identified where opposite transitions are observed, namely, A→B in poly(dA).poly(dT) and B→A in poly(dG).poly(dC). The two sequences exhibit qualitatively different counterion distributions, with a characteristic accumulation of sodium in the major groove of poly(dG).poly(dC) and the B→A transition driven by the electrostatic condensation mechanism. The resistance of the poly(dA).poly(dT) sequence to adopting the A-form is traced to the specific steric interactions of thymine methyl groups in the major groove. With these methyls replaced by hydrogens, reversible B↔A transitions become possible and the difference between the two molecules is significantly reduced. The good overall agreement with experimental data corroborates the general role of the electrostatic condensation mechanism in the A/B polymorphism in DNA.

Introduction

The double helical DNA may adopt a number of different structural forms.[1] Forms A and B are the most interesting because they are observed in vivo. The B-form is the dominant biological conformation, whereas the A-form is sometimes found in protein-DNA complexes and is considered as a high energy state adopted temporarily.[2,3] Reversible B↔A transitions probably represent one of the modes for governing protein-DNA interactions. The B-helix is long and narrow, with its core formed by stacked base pairs. In contrast, the A-helix is much shorter, and its core represents a 6 Å solvent accessible hole, with strongly inclined base pairs wrapping around it. Despite these very different shapes, both forms are right-handed helical duplexes with identical topologies and hydrogen bonding; transitions between them do not require base pair opening or destacking and involve relatively small conformational barriers.[4]

The B↔A transitions can be also induced in vitro by changing the DNA environment.[5−8] In condensed preparations, that is, in crystalline and amorphous fibers as well as in films, DNA always adopts the B-form under high relative humidity, but it can be reversibly driven to the A-form by placing the samples under relative humidity below 80%.[5,7,8] In aqueous solutions, single DNA molecules exhibit reversible B↔A transitions when certain organic solvents are added.[2,6] In both cases, the transition occurs at about the same water activity suggesting that the B↔A conformational switch is driven by the hydration state of the double helix.[9] The most known molecular mechanism proposed by Saenger et al.[10] postulates that, in A-DNA, water molecules form bridges between consecutive phosphate groups, which is not possible for the B-form where the interphosphate distances are too long. Therefore, under low water activity, hydration

* Corresponding author e-mail: alexey@ibpc.fr.

10.1021/ct049926d CCC: $30.25
Published on Web 01/29/2005

forces impose a more "economical" A-form.[10] This elegant mechanism contradicts, however, some long-known observations,[11] and it does not explain why, in both experiment and simulations, the B form generally remains stable in high salt where water activity is also reduced.[12−14] Some alternative mechanisms assume that the B→A equilibrium is governed by sequence dependent base pair stacking[15] as well as the hydrophobic effect.[16,17] Interbase interactions should be somehow involved because B↔A transitions are strongly sequence dependent. Notably, poly(dA).poly(dT) never goes to the A-form in condensed state[7] and is strongly A-phobic in solution.[18] In contrast, the poly(dG).poly(dC) duplex reportedly exhibits A-like features even under normal environment conditions.[19−21] Better understanding of the driving forces of the in vitro B↔A transition is essential for interpretation of its biological manifestations.

A number of theoretical approaches were applied in this field during its long history.[12,15,17,22,23] In the recent years, significant progress has been achieved by using all atom molecular dynamics (MD) simulations.[24−31] Spontaneous A→B transitions were shown to occur after less than a nanosecond of conventional simulations with periodical boundaries.[24,30] However, simulations of A-DNA under similar conditions encountered serious difficulties. In a homogeneous media composed of experimental concentrations of water, ions, and EtOH the A-DNA duplex was found to be unstable.[26,30] Its dynamics could be observed during a few nanoseconds under biphasic solvation, with only water and ions in the inner shell. Transitions from B- to A-DNA were obtained only with multivalent macroions initially placed at specific positions around DNA.[27] The energetic balance between the two forms in solution is very delicate and force field dependent.[25,26,32] An important recurrent feature revealed in these simulations consisted of the accumulation of free metal cations in the major groove of A-DNA suggesting that the electrostatic interactions across the major groove are essential for stabilization of the A-form[26,27,30,33] and that they can cause the B→A transition at least in some cases.

Stable dynamics of A-DNA and reversible B↔A transitions can be studied in MD simulations with vacuum boundaries.[31] In this approach, a DNA fragment is placed in a water drop with the desired number of free ions.[34] The drop size is maintained during simulations by periodical artificial return of the evaporated water molecules. The relative stability of A and B forms is modulated by changing the size of the drop, with spontaneous transitions observed in both directions and reversed, if necessary. The estimated critical hydration in the transition point was reasonably close to experiment without any additional fitting of the force field,[31] suggesting that the physics of the transitions is reasonably well reproduced even though water drop conditions do not correspond to any earlier experiments. It was shown that the major driving force of the B→A transition comes from accumulation of free metal cations in the opening of the major groove that results in inversion of electrostatic interactions between the phosphates of opposite backbone strands. This transition, therefore, is similar to the well-known phenomenon of free-ion-mediated electrostatic condensation of polyelectrolytes, but here it occurs inside the DNA duplex. With certain assumptions, this mechanism can be applied to experimental B↔A transitions, notably, it offers a detailed interpretation of the long known cooperative effects.[31] However, the relevance of MD simulation results to experimental B↔A transitions remains an open issue, and some authors consider that in reality this process should be too slow to be modeled in a computer.[35]

To get further insight in the A/B polymorphism in DNA it is necessary to rationalize the mechanisms of its sequence effects. A number of such effects were earlier discovered and studied in detail experimentally.[2,7,18,36−38] The relative A/B propensities of different base pair steps appear such that they tend to compensate one another, and generic DNA sequences commonly are neither A- nor B-philic.[18] There are, however, a few sequence motives characterized by very different A/B propensities. Among them stretches of consecutive guanines and adenines (G-tracts and A-tracts, respectively) represent the most known example. Disclosing the physical origin of these differences is an important challenge because these sequences also play special roles in vivo. They are overrepresented in eukariotic genomes as well as prokaryotes and archaebacteria[39,40] and have a long record of involvement in various biological processes. Contrasting physical properties of A- and G-tract DNA duplexes are well characterized.[7,20,36,41−54] A systematic review of these data exists in the literature.[55] In G-tracts, the B→A transition is very easy. The poly(dG).poly(dC) is generally prone to adopt the A-DNA conformation in conditions where random sequences stay firmly in the B-form.[7,45,48] In contrast, in A-tracts, the B→A transition is difficult,[48,54] and it was never observed for poly(dA).poly(dT) in condensed phase.[7,36]

The mechanism by which the base pair sequence can affect A↔B transitions is not clear. Early calculations[22] demonstrated that stacking in different base pair steps can partially account for their A- or B-philicity. A-tracts may be specifically stabilized in the B-form by specific minor groove hydration patterns[56] as well as cross-strand hydrogen bonding.[57] There are also evidences suggesting that the exceptional resistance of poly(dA).poly(dT) to the B→A transition is related to the thymine methyl groups that form a continuous hydrophobic cluster in the major DNA groove and provide several other stereochemical factors that may specifically stabilize B- versus A-form.[18,58]

Here we study the aforementioned sequence effects in B↔A transitions by using all atom MD simulations with free vacuum boundaries. DNA fragments with A- and G-tract motives are found to be distinctly different as regards their relative preferences toward B- and A-forms in good qualitative agreement with experiment. In poly(dG).poly(dC) dodecamer, the B→A transition is easy, smooth and perfectly reversible. In contrast, only A→B transitions could be observed in a similar poly(dA).poly(dT) fragment, whereas attempts of the B→A transition always lead to denaturation. An intermediate range of hydration numbers is found where the B-form poly(dA).poly(dT) dodecamer is stable and the A→B transition is observed, while the poly(dG).poly(dC) duplex is stable in A-form and makes the opposite transition. The denaturation of poly(dA).poly(dT) in intermediate

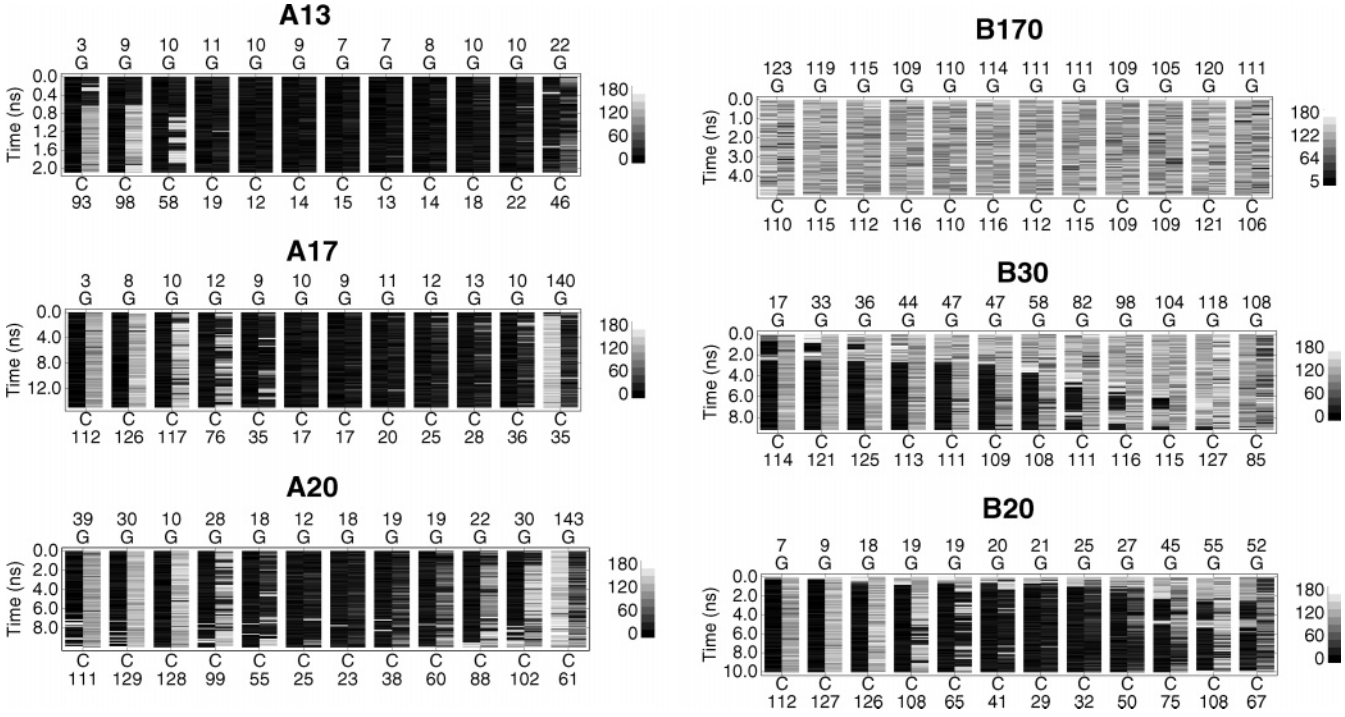

Figure 1. A↔B transitions in C12 shown by the dynamics of sugar pucker pseudorotation.[92] The boundary pucker values (ca. 0° and 180°) are assigned the black and white colors, respectively, with intermediate values mapped linearly to the gray scale levels. Therefore, the C3′-endo conformation of the A-form is seen as a nearly black field, whereas typical B-form C2′-endo and C1′-exo conformations correspond to nearly white and light gray, respectively. Each base pair step is characterized by a column consisting of two subcolumns, with the left subcolumns referring to the sequence written above in 5′-3′ direction from left to right and the time averaged phases given on top. The right subcolumns refer to the complementary sequence shown below together with the corresponding time averaged phases. The starting A- and B-DNA states and the corresponding hydration numbers (Γ) are included in the labels.

conditions can be prevented by flanking A-tracts with GC pairs. In this case the B→A transition becomes possible in long duplexes, but dodecamer fragments stay in B-form down to very low hydration and then collapse without denaturation, which makes this process reversible and allows testing the role of different factors, notably, the thymine methyls. It appears that with methyls replaced by hydrogens, i.e., thymine residues substituted by uracils, the B→A transitions become possible, and the difference in the relative A/B philicity between A-tracts and G-tracts is drastically reduced.

Methods

A series of MD simulations was carried out for double helical dodecamer fragments poly(dA).poly(dT) and poly(dG).poly-(dC) referred to as T12 and C12, respectively, as well as a few derivatives of these sequences. The sizes of the water drop were chosen to provide a hydration number (Γ) of 13, 17, 20, 25, 30, 40, 60, and 80 water molecules per nucleotide. Additional simulations for $\Gamma \approx 170$ were carried out by using periodical boundary conditions. The simulations either started from standard A- and B-DNA states prepared as explained below or continued from a final state of a previous trajectory with water molecules added or removed to reach a desired hydration level. Similar results are obtained with both strategies. In this way a quasi-static pattern of B↔A transitions can be reproduced that mimics in vitro titration experiments.

The simulation protocols were similar to the earlier water drop simulations.[34,31] We use the internal coordinate molecular dynamics (ICMD) method[59,60] adapted for DNA[61,62] with the time step of 0.01 ps. In this approach, the DNA molecule has all bond lengths and almost all bond angles fixed at their standard values. The only variable bond angles are those centered at the sugar C1′,...,C4′, and O4′ atoms, which ensures the flexibility of the furanose rings. In contrast, bases, thymine methyls, and phosphate groups move as articulated rigid bodies, with only rotations around single bonds allowed. The highest frequencies in thus obtained models are additionally balanced by increasing rotational inertia of the lightest rigid bodies as described earlier.[61,63] The possible physical effects of the above modifications have been discussed elsewhere.[60,64] The electrostatic interactions are treated with the SPME method[65] adapted for infinite vacuum boundary conditions.[34] The common values of Ewald parameters were used, that is 9 Å truncation for the real space sum and $\beta \approx 0.35$. Reference simulations with periodical boundaries were carried out as described before,[34] with the standard SPME method in NVT ensemble conditions with a rectangular unit cell of $45 \times 45 \times 65$ Å under normal water density.

The standard initial states were prepared as follows. Either A- or B-DNA canonical structures[66] were first immersed in a large rectangular TIP3P[67] water box and next external solvent molecules were removed by using a spherical distance cutoff from DNA atoms. The cutoff radius was

adjusted to obtain the desired number of water molecules remaining. The drop was neutralized by randomly placing Na^+ ions at water positions selected so that their distances from DNA were 5 Å or larger. The initial counterion distribution was preequilibrated by running 1 ns dynamics in water drops of 500 molecules for A-DNA and 800 molecules for B-DNA, with DNA atoms restrained to their initial positions. The final drop size was adjusted by adding or removing water from the surface. The described procedure was intended to ensure the start of dynamics from closely similar states regardless of the drop size.

Every system was energy minimized first with the solute held rigid and then with all degrees of freedom. Dynamics were initiated with the Maxwell distribution of generalized momenta at low temperature. The system was next slowly heated to 250 K and equilibrated during several picoseconds. Production trajectories were computed with the temperature bound to 293 K by the Berendsen algorithm[68] with a relaxation time of 10 ps. For better comparison with earlier simulations of A↔B transitions, the original Cornell et al. force filed[69] was used. Duration of production runs varied from 2 to 25 ns depending upon the observed character of dynamics. The conformations were saved with a 2.5 ps interval. The results were analyzed with in-house routines and the Curves program.[70]

Results

Comparative Dynamics of A↔B Transitions. The B↔A transition dynamics for C12 are shown in Figure 1 and Figure 2. Simulations starting from A- and B-DNA converged to identical ensembles of structures. This was checked for several Γ values covering the whole range of interest. The A-form loses stability with $\Gamma \gtrsim 20$, which is similar to earlier results for the dodecamer CGCGAATTCGCG.[31] In contrast, the low limit of the B-form is shifted to higher Γ values. With $\Gamma \approx 30$, sugars in the purine strand concertedly switch to C3′-endo, and for $20 < \Gamma < 30$ an intermediate state is observed, with one strand in nearly A-DNA conformation and the other closer to the B-form. As a result, the B↔A transition looks easier and much smoother than for CGCGAATTCGCG.[31] The transition is complete in the center, and the A-DNA zone gradually spreads toward the termini as long as Γ is reduced. From the experiment, the B-form is known to be particularly resistant near the termini.[71] For the CGCGAATTCGCG sequence this B-philicity involved three base pairs, and it did not change with the duplex length increased.[31] Here this effect is much smaller. All this agrees with the known A-philicity of the poly(dG).poly(dC).[72] Figure 2 shows a quasi-static titration-like pattern of these transitions as monitored by different structural parameters. All traces exhibit S- or Z-shaped profiles sometimes with a very distinct transition zone.[73] The A-DNA structures obtained under low hydration are very close to the canonical conformation, with the final RMSD values below 2 Å. In contrast, under the highest hydration, an underwound B-DNA is observed with a strong bias toward A-form. Similar deviations were reported earlier by others,[21,74,75] and they are partially due to a known force field bias.[76]

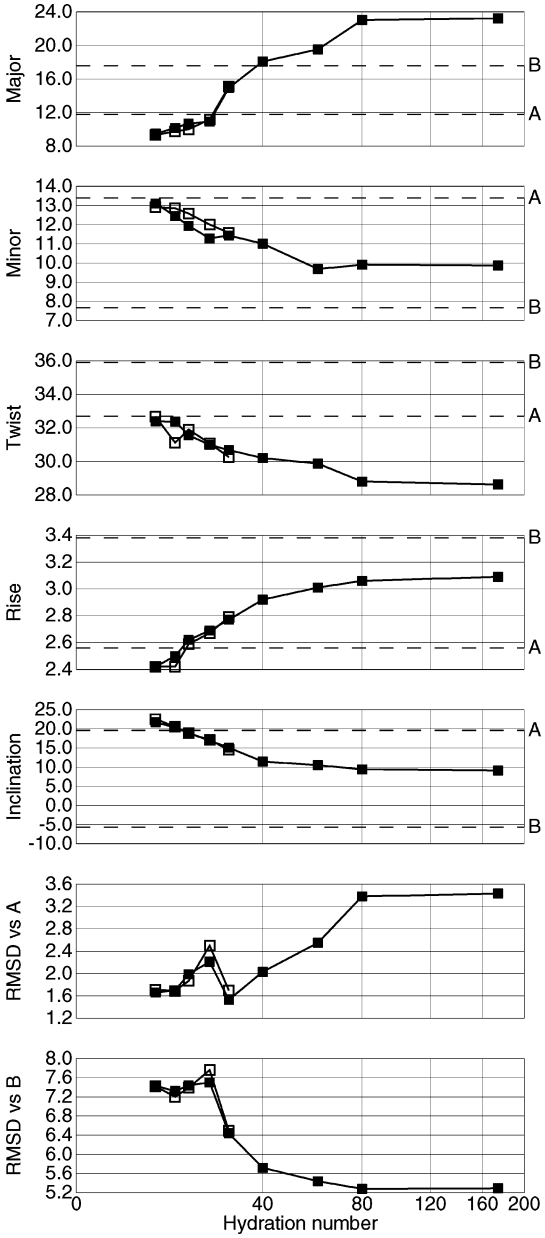

Figure 2. A↔B transitions in C12 monitored by different parameters. The data were obtained by averaging over the last 1 ns of trajectories computed with different hydration numbers. Open and closed squares show the results obtained from trajectories starting from A- and B-DNA, respectively. The two top plates show the average groove widths measured as described elsewhere.[31,93] The helical parameters are computed with program Curves.[70] All distances are in angstroms and angles in degrees. The dotted lines mark canonical A- and B-DNA levels.

Similar results for T12 are shown in the next two figures. In this case we failed to obtain a B→A transition despite many attempts. Trajectories starting from B-DNA with $\Gamma < 20$ lead to deformed structures with some base pairs opened. The corresponding A-form trajectories were stable and showed no structural trends during reasonable simulation time (up to 30 ns). For $\Gamma \geq 20$ normal A→B transitions were observed, and trajectories converged to similar structures starting from A- and B-DNA states. Because of the

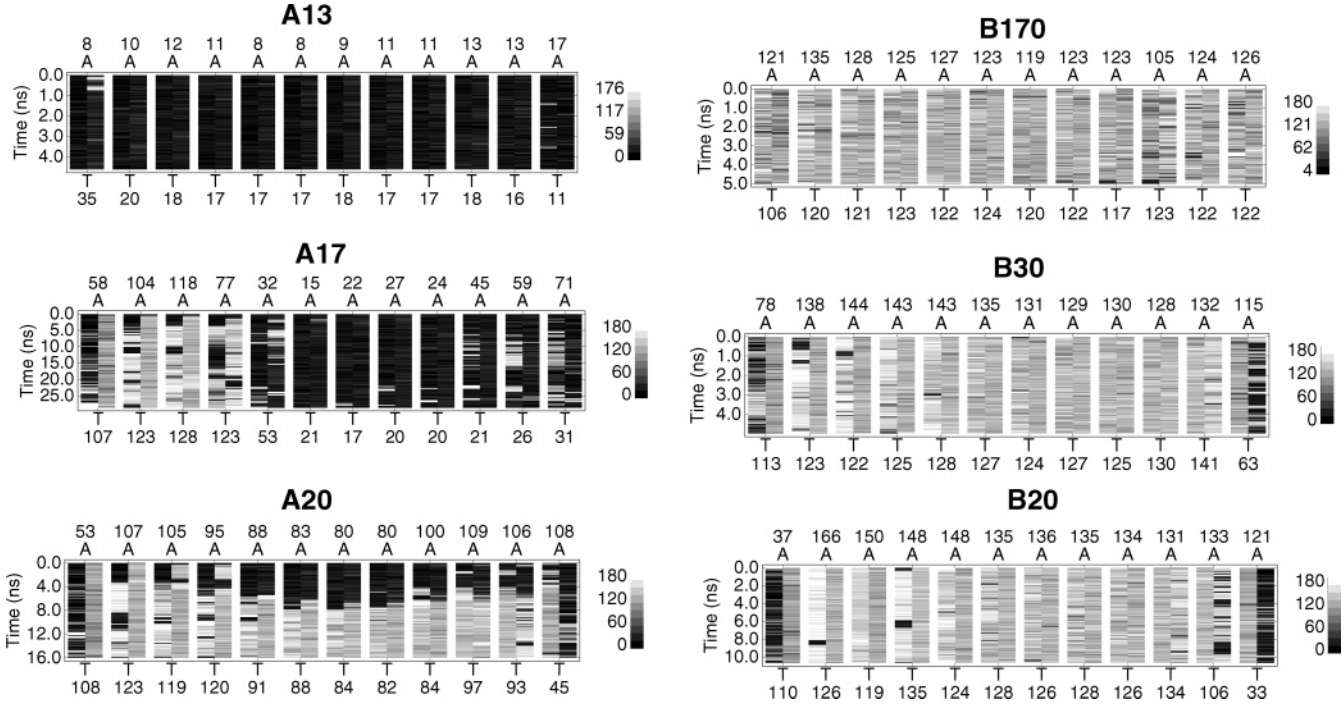

Figure 3. A↔B transitions in T12 shown by the dynamics of sugar pucker pseudorotation. The notation is as in Figure 1.

divergence of trajectories one cannot say which of the final states is more probable below $\Gamma \approx 20$. However, comparison of plates A20 and B20 in Figure 1 and Figure 3 shows that with $\Gamma \gtrsim 20$ opposite transitions are observed in C12 and T12. In this range of hydration, trajectories starting from A- and B-DNA were convergent for both duplexes; therefore, we can conclude that the experimental difference between poly(dG).poly(dC) and poly(dA).poly(dT) is qualitatively reproduced. Figure 5 shows the snapshots for T12 and C12 at the end of dynamics with $\Gamma = 25$. For T12 this is a B-DNA with a narrow minor groove near both ends and a widening in the middle. A significant number of the Na^+ ions are seen in front of the minor groove. In contrast, C12 gives a typical A-DNA conformation with a layer of Na^+ sandwiched between the opposite phosphate groups in the narrow major groove. This "electrostatic sandwich" provides the main driving force of the B A transition according to the free counterion driven condensation mechanism.[26,33,31]

The time averaged pattern of the transitions for T12 is shown in Figure 4. Compared to that in Figure 2 it shows similarities as well as differences. The low hydration A-form structures are very close to those for C12 and the canonical A-DNA. Unlike for C12, however, the computed high hydration structures represent rather good B-DNA. Only the twist remains clearly underestimated in agreement with earlier reports.[76] The traces in Figure 4 are interrupted to indicate that they perhaps do not sample from a single continuous transition pathway. Note that when Γ changes from 60 to 30 the structures show no trend toward the A-form. Instead some parameters are nearly stable and other even exhibit an opposite trend. A shift toward the A-form is observed under lower hydration, but it is relatively small and inconsistent in different parameters. This behavior is very different from C12 as well as the CGCGAATTCGCG dodecamer,[31] suggesting that the T12 structure may be trapped in a pathway that leads away from the A-form and the B→A transition would never occur if the corresponding trajectories were continued.

The major driving force of the B→A transitions is arguably due to accumulation of free solvent ions in the major DNA groove.[31] Therefore, the difference between T12 and C12 should be also visible in the behavior of counterions. The corresponding cylindric radial distribution functions are compared in Figure 6. These patterns provide useful information about the structure-specific interactions between DNA and free solvent cations. Because the duplexes are entirely covered by water the radial distributions of water oxygens effectively show the free space around DNA. All this space is, in principle, accessible for Na^+ ions since their size is not very different from that of water. In the absence of specific DNA-ion interactions, the solid and dotted traces in Figure 6 should have similar peak positions. In contrast, sites of strong Na^+-DNA interactions produce separate peaks. Note also that the height of each peak in a cylindrical distribution should be multiplied by the corresponding distance when their relative weights are estimated.

The characteristic distributions for B-DNA are best seen in the top three plates of T12. The three broad water peaks at approximately 4, 8 and 12.5 Å correspond, respectively, to the first water layer in the major groove, the next few layers in both major and minor grooves and the bulk water outside DNA radius which is about 10 Å for both A- and B-DNA. The characteristic A-DNA distributions can be seen in the bottom two plates of T12 as well as in several plates of C12. In this case the water radial distributions exhibit two broad maxima, one in the interior and another in the exterior of DNA. Each of the many peaks in the Na^+ traces commonly involve contributions from more than one type of ion-DNA contacts. For instance, the prominent peak at 5 Å from the center of B-DNA results from direct as well as

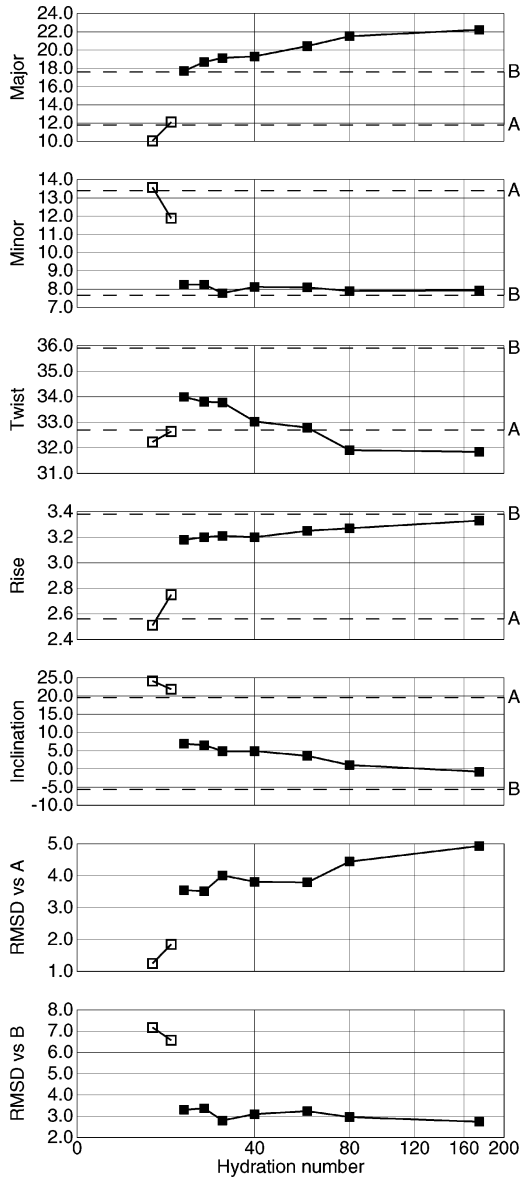

Figure 4. A↔B transitions in T12 monitored by different parameters. Solid squares show results for $\Gamma \geq 20$ that were similar for trajectories starting from A- as well as B-DNA. Open squares show results obtained form trajectories starting from A-DNA only. Other notation is as in Figure 2.

water-mediated contacts in depth of both major an minor DNA grooves. Most interesting for our purposes are the outer peaks that contain the major part of ions. For B-DNA this is the broad peak at about 12.5 Å which forms the helical axis. It corresponds to highly mobile ions involved in nonspecific phosphate screening. In A-DNA, the major Na^+ peak is found at 10 Å, and it corresponds to the ions sandwiched in the opening of the major groove as seen in Figure 5.

For C12 even under high hydration the Na^+ distribution is rather different from typical B-DNA. These structures have a strong negative Xdisp; therefore, the helical axis is shifted to the major groove and is accessible to water and ions. Nevertheless, the large majority of counterions are found in the single broad outer peak characteristic of B-form. This peak does not change when the amount of water is further increased. In fact, ion distributions obtained in these conditions are identical to those observed in infinite media with periodic boundaries.[34] In contrast, with reduced Γ, these counterions are progressively pushed closer to DNA, and, eventually, they accumulate in the characteristic A-DNA peak at 10 Å. This peak is predominant with $\Gamma \leq 30$. The foregoing scenario corresponds to the relatively smooth B→A transition shown in Figure 1 and Figure 2. The T12 patterns in the right column are radically different. In this case both water and Na^+ distributions retain the characteristic B-DNA shapes even with $\Gamma \approx 20$. The ions are not pushed inside DNA when the outer water shell is reduced, and the Na^+ peak at 12.5 Å remains dominant. As shown in Figure 5 many of these ions are aligned along the minor groove between the two phosphate strands. Such ions are essentially immobilized because there is not enough water for them to diffuse elsewhere.

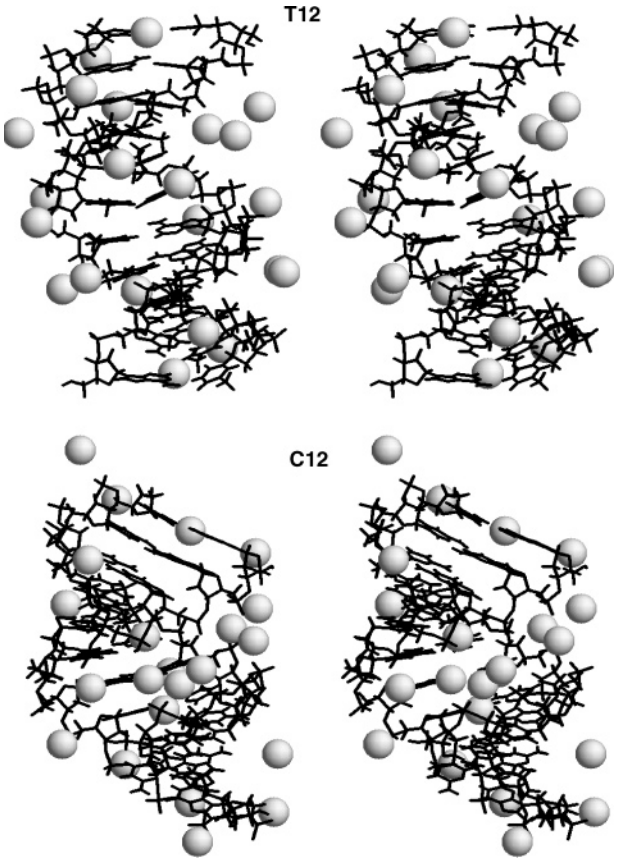

Figure 5. Stereo snapshots of the convergent final states of T12 and C12 in a water drop corresponding to $\Gamma = 25$. The Na^+ ions are shown by spheres.

The Origin of the Resistance of poly(dA).poly(dT) to Adopting the A-Form. The resistance of T12 to the B→A transition turned out to be only relative because in longer poly(dA).poly(dT) fragments such transitions were reproduced without major difficulties (the results not shown). The tendency of terminal base pairs to break was overcome by adding GC pairs at both duplex termini. Nevertheless, the special case of T12 deserves attention because its distinction from C12 is in clear correspondence with long-known experimental data, and it is interesting to clarify the origin of such behavior at least in simulations. To this end, we carried out a series of additional calculations with varied

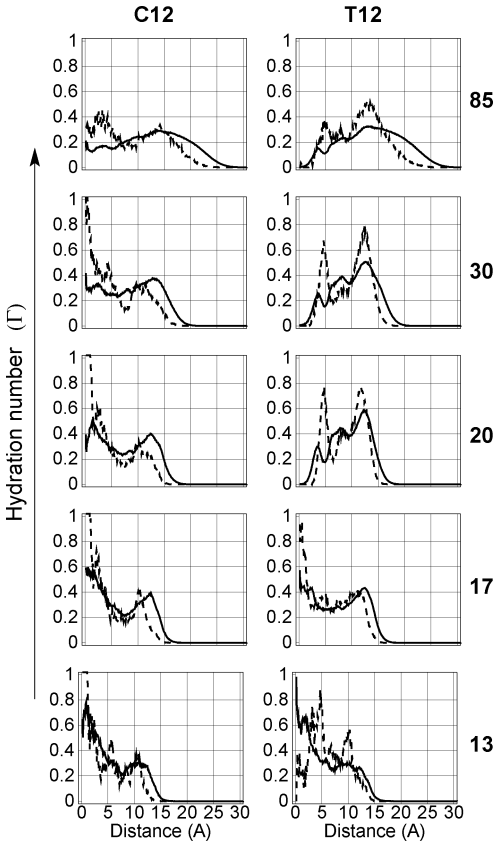

Figure 6. Cylindrical radial distribution functions for water oxygens (solid lines) and Na^+ ions (dashed lines) around C12 and T12. DNA structures saved in the last nanosecond of dynamics together with surrounding water and counterions were superimposed with the canonical canonical B-DNA with the global coordinate OZ direction as the common helical axis. The Na^+ ions and water oxygens were counted in coaxial 0.1 Å thick cylinders. The distributions are volume normalized, that is scaled with a factor of 1/*r*, and the final plots were area normalized to fit the same scale. The corresponding hydration numbers are shown on the right.

conditions. (i) It was checked that the A-form of poly(dA).poly(dT) is sufficiently stable with respect to the base pair opening. The forced breaking of one base pair did not perturb the overall structure even in very long simulations under intermediate hydration. (ii) The collapse of T12 under low hydration is reproduced when the size of the water drop is reduced gradually in several steps by removing each time only outer molecules not in contact with ions. This proves that the observed onset of denaturation is not caused by the initial conditions when the B-form is immediately placed under low hydration. (iii) The B→A transition still does not occur when the base pair opening is prevented by placing GC pairs at both termini (CT10C dodecamer). Starting from the B-form this duplex remained stable even under $\Gamma = 17$. With $\Gamma = 13$ water started to leave the minor groove and the structure exhibited irregular deformations that drove it to a shrunk collapsed conformation but not to the A-form. At the same time, the CT10C duplex exhibited higher stability in the A-form, with AB transitions observed only beyond $\Gamma = 25$.

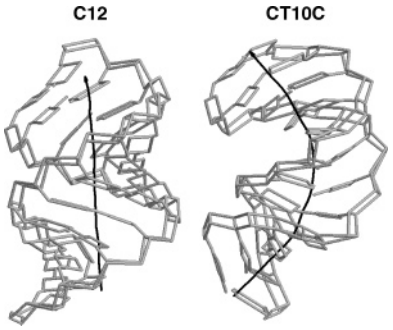

Figure 7. Stable conformations of C12 and CT10C under $\Gamma = 13$.

Figure 7 compares the conformations of C12 and CT10C observed in simulations starting from the canonical B-DNA under $\Gamma = 13$. For C12 this is an A-form structure with a very straight helical axis. In contrast, the CT10C structure represents a strongly curved B-like duplex with the minor groove closed near one end. The strong bend was evidently due to several Na^+ ions sandwiched between the opposite phosphates in the locally narrowed major groove also seen in Figure 7. In this region, the major groove width is similar to that in A-DNA; therefore, the main driving factor of the B→A transition is present and it is sufficiently strong to bring the opposite phosphates together, but the overall duplex structure prefers to deform rather than go to the A-form. The CT10C transition from straight B-DNA to the shrunk structure in Figure 7 is easily reversible. With a few outer water shells added, the Na^+ ions rapidly leave the major groove, the helical axis straightens up, and water invades the collapsed region of the minor groove.

Earlier it was noticed that the thymine methyl groups affect a number of different interactions involved in B↔A transitions.[18,58] The array of thymine methyls in the major groove of poly(dA).poly(dT) strongly stabilize the B-DNA conformation,[58,77] both mechanically and via the hydrophobic effect, which would hinder the B→A transition. The distinct behavior of the CT10C duplex gives an opportunity to check this hypothesis. To this end we carried out a series of additional simulations for an analogous dodecamer duplex with thymine methyls replaced by hydrogens (CU10C). To avoid the possible additional effect of the charge redistribution inside pyrimidine rings all atoms in the newly obtained uracil residues had parameters identical to those in thymine, whereas the H6 hydrogens carried a residual charge identical to that of the thymine methyl group. The results obtained appear in Figures 8−10. It is readily seen that they agree with the above hypothesis. With C6 pyrimidine methyls replaced by hydrogens the B→A transition becomes possible, and reversible B↔A transition are observed in the same range of hydration numbers as in C12 and earlier in CGCGAATTCGCG.[31] The transition patterns in Figure 8 resemble those in Figure 1 in that the sugars in the purine strand switch from C2′-endo to C3′-endo early in the B→A transition. The B→A transition occurs with somewhat lower Γ values suggesting that the CU10C duplex is less A-philic, which is also seen in Figure 9. Unlike for C12, the ion and water radial distributions obtained in large drops as well as with periodical boundaries exhibit clear B-DNA patterns with the

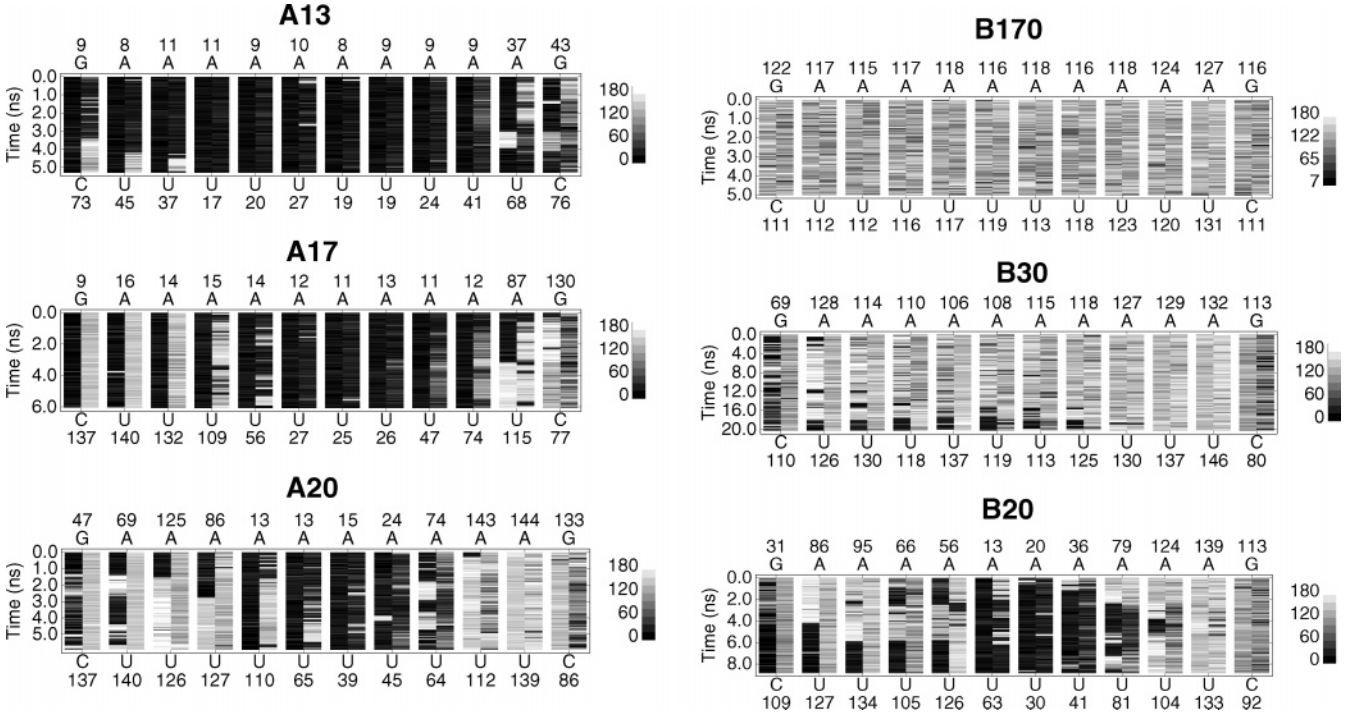

Figure 8. A↔B transitions in CU10C shown by the dynamics of sugar pucker pseudorotation. The notation is as in Figure 1.

center of the duplex inaccessible for the solvent (see Figure 10). Overall this behavior resembles that of C12 as well as CGCGAATTCGCG,[31] but it is radically different from T12 and CT10C in agreement with an earlier suggested key role of the thymine methyls in the poly(dA).poly(dT) resistance to the B→A transformation. These results show also that the thymine methyls affect the B-DNA structure under high hydration and not just the B↔A transition state.

Discussion

The striking contrast in the A/B polymorphism of poly(dA).poly(dT) and poly(dG).poly(dC) is a very long-known unexplained phenomenon.[78] Owing to the recent progress in the force field development[69,79] it is now possible to reproduce this effect in atomistic detail in a computer. The DNA structures computed in the course of this study are reasonably close to experiment, and several important earlier findings are qualitatively reproduced. (i) There is a range of hydration numbers where poly(dA).poly(dT) and poly(dG).poly(dC) are stable in B- and A-forms, respectively, with reproducible opposite transitions observed for the two sequences. Even with no additional force field fitting applied, the characteristic hydration values are close to experiment. (ii) The C12 dodecamer exhibits a very easy and smooth B↔A transition under relatively high humidity, with its B-form featuring a strong A-DNA bias. (iii) The T12 dodecamer refuses to take the B-to-A transition pathway and is prone to denaturation. At the same time, longer poly(dA).poly(dT) fragments with GC termini can reversibly transform. These features resemble the experimental behavior of poly(dA).poly(dT) which is B-philic and tends to collapse instead of a B→A transition but still can go to the A-form in long A-tracts in solution.[80]

Under low humidity, all trajectories converge to similar A-form conformations that seem to be independent of the sequence and are always very close to the fiber canonical A-DNA. In contrast, a much more significant and sequence dependent divergence from experimental structures is observed for B-DNA. This result is in a surprising correspondence with experimental data. The high regularity and the absence of sequence effects for A-form was noticed long ago for X-ray fiber diffraction patterns[7] and later confirmed in the ensemble of single-crystal A-DNA structures.[81] Thus, in both experiment and simulations, the A-form of DNA is virtually insensitive to the base pair sequence. One should note, in addition, that it is much less sensitive to the accuracy of the force field than the B-form. Really, despite specific fitting of empirical potentials, at present, B-DNA structures obtained in free simulations only rarely approach experimental conformations closer than 2.5−3.0 Å RMSD for dodecamer duplexes. This is significantly higher than the RMSD numbers observed for A-DNA without any additional parameter fitting. All these observations are consistently explained if we assume that the A-form is dominated by the ion/phosphate "electrostatic sandwich" in the major groove. The strong interactions of phosphate groups with metal cations effectively impose geometric constraints upon the interphosphate distances and suppress all other factors that might affect the overall structure. The same interactions can maintain A-DNA in protein complexes. Interestingly, in all such structures refined until now the major DNA groove is exposed to solvent, with protein-DNA contacts limited to the minor groove only.

The observed systematic bias in the computed B-DNA conformations is similar to earlier reports.[21,74,75] The Cornell et al. parameters[69] tend to underestimate the helical twist in B-DNA,[76] and for the sequences studied here this bias is

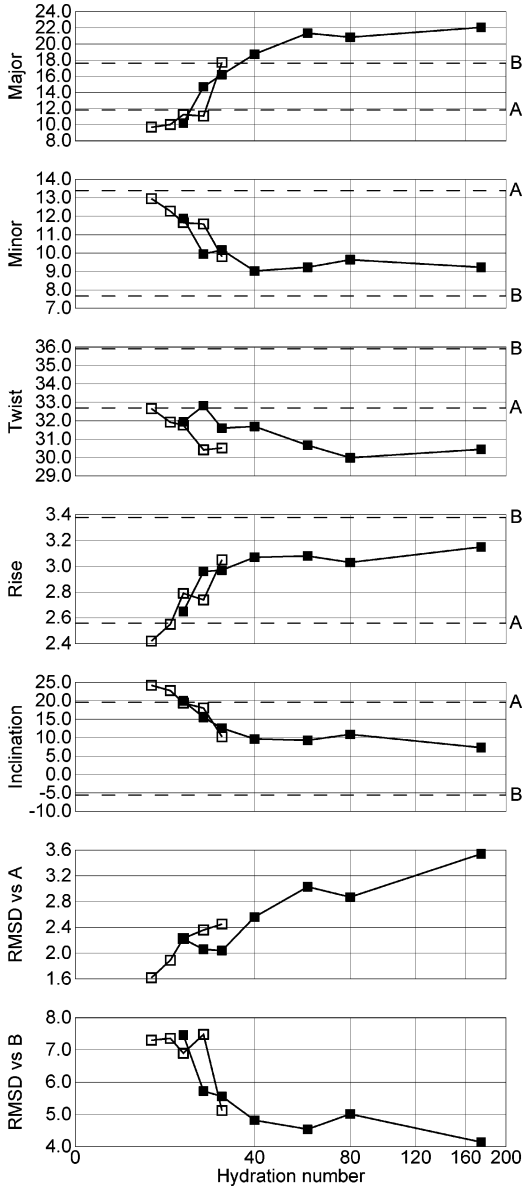

Figure 9. A↔B transitions in CU10C monitored by different parameters. Open and closed squares show the results obtained from trajectories starting from A- and B-DNA, respectively. Other notation is as in Figure 2.

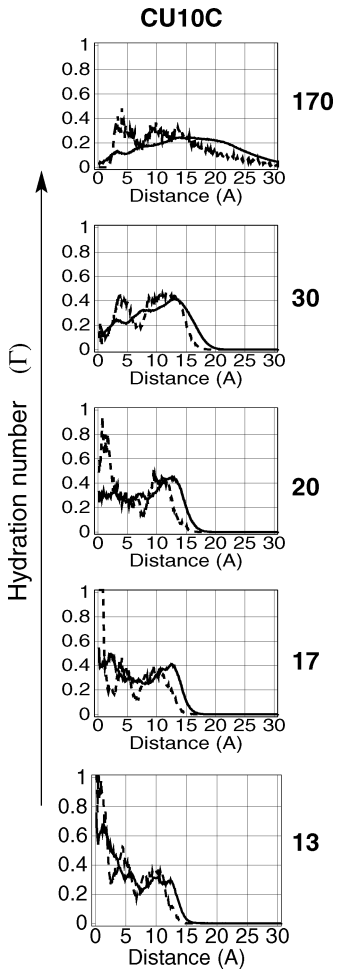

Figure 10. Cylindrical radial distribution functions for water oxygens (solid lines) and Na^+ ions (dashed lines) around CU10C. See details in the legend to Figure 6.

perhaps the largest. The computed poly(dA).poly(dT) conformations are relatively close to the canonical B-DNA structure although, as earlier,[82] it is underwound by ca. 4° with respect to experiment in solution.[41−43] For poly(dG).poly(dC) a similar bias results in structures with very A-like helical parameters although the sugar pseudorotation dynamics features mainly south/east phases characteristic of B-DNA (see Figure 1). Earlier such structures were considered as strongly underwound B-DNA or as A-DNA with B-like backbone.[21,74,75] It is difficult to tell exactly how strong these deviations are with respect to experimental data. Compared to free G-tracts in crystals the computed structures deviate toward the B-form. However, these X-ray data are hardly appropriate for comparison because they all are obtained for a unique A-DNA specific packing characterized by very special DNA-DNA interactions in the minor groove.[83] In contrast, compared to G-tracts in protein-DNA complexes, the computed structures deviate toward the A-form. However, it is also not clear if these data are relevant for comparison. For instance, there is a G5-tract in the complex of the refined structure of DNA polymerase Y.[84] This G-tract is protruding from the enzyme active center and makes no contacts with the protein. Its structure features a normal B-DNA conformation with south sugar puckers and the average twist of 34.2°, i.e., rather close to that of generic B-DNA. This value, however, diverges from experiments in solution (ca. 32.7° for long G-tracts[41]) suggesting that the structure may still be affected by the crystal environment.

The foregoing arguments show how challenging is the problem of modeling of sequence effects in DNA. On the other hand, NMR and circular dichroism (CD) studies in solution qualitatively agree with the character of G-tract structures obtained here and earlier by others with the same force field. The corresponding CD data suggest that the high hydration conformation of poly(dG).poly(dC) already has an A-like base pair stacking.[21,74] Nevertheless, these structures cannot be assigned to the A-form because the sugar phases determined by NMR are all in the south and because a cooperative transition to a genuine A-form is still distinguishable in solution as well as in crystalline fibers.[19] Since the corresponding spectral changes are relatively small, they can well result from a transition pattern similar to that in

Figure 2, with sugar phases switching to the north and grooves changing their width, but relatively minor shifts in helical parameters.

A number of earlier observations indicate that, during the B↔A transition, the double helical DNA temporarily loses its stability. With reduced water content in solution the melting temperature of B-DNA first decreases and then goes up, with a minimum passed before the B→A transition occurs.[85] In some sequences even a temporary denaturation takes place followed by renaturation in the A-form when the intermediate range of hydration is passed.[86] The base pair opening in the AT-alternating sequence may also explain the reportedly very slow relaxation kinetics of B↔A transitions in this duplex.[35] Therefore, the onset of denaturation in T12 in our simulations is not completely surprising. This DNA fragment as well as CT10C may represent examples of duplexes that are stable in A- as well as B-forms but cannot sustain the intermediate conditions. Similar experimental examples exist in the literature.[86] It is understood that, if the transition pathway passes through a denaturated state, it could not be reproduced in calculations.

A number of possible mechanisms were earlier discussed in relation to the resistance of poly(dA).poly(dT) to the B→A transition. This sequence effect can originate from specific base stacking as well as DNA-solvent interactions, but all such factors are mixed and difficult to separate.[15,22,17] For instance, the distinctive narrow minor groove of the B′-form of poly(dA).poly(dT)[87] is probably due to internal interactions, but it is also sealed by a very stable water spine hydration structure in the minor groove.[88] Moreover, the narrow groove should produce a 2-fold effect upon the counterion distribution around DNA.[33,55] The free positive ions should tend to accumulate in front of the narrow minor groove, and, simultaneously, their concentration in the major groove is reduced. All these features should stabilize the B-form and hinder the B→A transition. Different authors earlier pointed to thymine methyls as the probable principal cause of this effect.[17,18,31] We are not aware of published experiments on B↔A transitions in poly(dA).poly(dU), but a generally strong effect of C6 substitutions in the uracil residue is well-known.[77,86] Here this hypothesis has been confirmed in realistic atom-level simulations. The thymine methyls generally do not hinder B↔A transitions in water drop simulations with sequences other than poly(dA).poly(dT), for instance, poly(dAT) (unpublished results of the author). In contrast, for short fragments of poly(dA).poly(dT) their effect is very strong, which allowed us to set up a direct "yes or no" test.

Three different mechanisms of the thymine methyl effect were proposed earlier. (i) These groups may hinder a negative slide movement involved in the B→A transition.[17] (ii) They form a continuous nonpolar cluster in the major groove of B-DNA that should provide additional hydrophobic stabilization.[18] (iii) They reduce the accessible volume of the major groove, which may prevent accumulation of free solvent cations.[31] Our results agree with the first two explanations better than with the last one. Really, it turns out that, with removed methyl groups, the properties of B-DNA under high hydration are strongly affected and not just the B→A transition state. It should be noted that all electrostatic properties of uracil bases here were nearly identical to those of thymines. Therefore the effect cannot be due to different ring stacking angles and should be due to either major groove hydration or some steric factors related to the arrays of methyl groups in tracts of stacked thymines.

Many years ago, the contrasting propensities of poly(dA).poly(dT) and poly(dG).poly(dC) to adopt A- and B-forms presented the first experimental demonstration of sequence dependent properties of the double helical DNA structure.[78] Since then the repertory of reported sequence effects has many times increased, and yet the exact physical origin of this particular difference remains controversial. It is shown here that A↔B transitions observed in water drop simulations exhibit clear trends qualitatively similar to the long known experimental observations. These results corroborate the putative general role of the intraduplex electrostatic condensation mechanism for A↔B transitions in DNA in vitro and suggest that future studies in the same direction can give more definite answers to the issues discussed here.

References

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

(2) Ivanov, V. I.; Minchenkova, L. E. *Mol. Biol.* **1995**, *28*, 780−788.

(3) Lu, X. J.; Shakked, Z.; Olson, W. K. *J. Mol. Biol.* **2000**, *300*, 819−840.

(4) Foloppe, N.; Nilsson, L.; MacKerell, A. E., Jr. *Biopolymers* **1999**, *61*, 61−76.

(5) Franklin, R. E.; Gosling, R. G. *Nature* **1953**, *171*, 740−741.

(6) Tunis-Schneider, M. J.; Maestre, M. F. *J. Mol. Biol.* **1970**, *52*, 521−541.

(7) Leslie, A. G. W.; Arnott, S.; Chandrasekaran, R.; Ratliff, R. L. *J. Mol. Biol.* **1980**, *143*, 49−72.

(8) Piskur, J.; Rupprecht, A. *FEBS Lett.* **1995**, *375*, 174−178.

(9) Malenkov, G.; Minchenkova, L.; Minyat, E.; Schyolkina, A.; Ivanov, V. *FEBS Lett.* **1975**, *51*, 38−42.

(10) Saenger, W.; Hunter, W. N.; Kennard, O. *Nature* **1986**, *324*, 385−388.

(11) Fuller, W.; Mahensdasingam, A.; Forsyth, V. T. *Nature* **1988**, *335*, 596.

(12) Ivanov, V. I.; Minchenkova, L. E.; Minyat, E. E.; Frank-Kametetskii, M. D.; Schyolkina, A. K. *J. Mol. Biol.* **1974**, *87*, 817−833.

(13) Zimmerman, S. B.; Pheiffer, B. H. *J. Mol. Biol.* **1980**, *142*, 315−330.

(14) Cheatham, T. E., III.; Kollman, P. A. in *Interactions and Expression of Biological Macromolecules. Proceedings of the 10th Conversation, State University of New York, Albany, N. Y. 1998*; Sarma, R. H., Sarma, M. H., Eds.; Adenine Press: New York, 1998; pp 99−116.

(15) Calladine, C. R.; Drew, H. R. *J. Mol. Biol.* **1984**, *178*, 773−782.

(16) Ivanov, V. I.; Minchenkova, L. E.; Schyolkina, A. K.; Poletaev, A. I. *Biopolymers* **1973**, *12*, 89−110.

(17) Hunter, C. A. *J. Mol. Biol.* **1993**, *230*, 1025−1054.

(18) Tolstorukov, M. Y.; Ivanov, V. I.; Malenkov, G. G.; Jernigan, R. L.; Zhurkin, V. B. *Biophys. J.* **2001**, *81*, 3409−3421.

(19) Arnott, S.; Selsing, E. *J. Mol. Biol.* **1974**, *88*, 551−2.

(20) Sarma, M. H.; Gupta, G.; Sarma, R. H. *Biochemistry* **1986**, *25*, 3659−3665.

(21) Trantirek, L.; Stefl, R.; Vorlickova, M.; Koca, J.; Sklenar, V.; Kypr, J. *J. Mol. Biol.* **2000**, *297*, 907−22.

(22) Mazur, J.; Sarai, A.; Jernigan, R. L. *Biopolymers* **1989**, *28*, 1223−1233.

(23) Ivanov, V. I.; Zhurkin, V. B.; Zavriev, S. K.; Lysov, Y. P.; Minchenkova, L. E.; Minyat, E. E.; Frank-Kamenetskii, M. D.; Schyolkina, A. K. *Int. J. Quantum Chem.* **1979**, *16*, 189−201.

(24) Cheatham, T. E., III.; Kollman, P. A. *J. Mol. Biol.* **1996**, *259*, 434−444.

(25) Yang, L.; Pettitt, B. M. *J. Phys. Chem. B* **1996**, *100*, 2564−2566.

(26) Cheatham, T. E., III.; Crowley, M. F.; Fox, T.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 9626−9630.

(27) Cheatham, T. E., III.; Kollman, P. A. *Structure* **1997**, *5*, 1297−1311.

(28) Cieplak, P.; Cheatham, T. E., III.; Kollman, P. A. *J. Am. Chem. Soc.* **1997**, *119*, 6722−6730.

(29) Jayaram, B.; Sprous, D.; Young, M. A.; Beveridge, D. L. *J. Am. Chem. Soc.* **1998**, *120*, 10629−10633.

(30) Sprous, D.; Young, M. A.; Beveridge, D. L. *J. Phys. Chem. B* **1998**, *102*, 4658−4667.

(31) Mazur, A. K. *J. Am. Chem. Soc.* **2003**, *125*, 7849−7859.

(32) Feig, M.; Pettitt, B. M. *Biophys. J.* **1998**, *75*, 134−149.

(33) Rouzina, I.; Bloomfield, V. A. *Biophys. J.* **1998**, *74*, 3152−3164.

(34) Mazur, A. K. *J. Am. Chem. Soc.* **2002**, *124*, 14707−14715.

(35) Jose, D.; Porschke, D. *Nucleic Acids. Res.* **2004**, *32*, 2251−2258.

(36) Zimmerman, S. B. *Annu. Rev. Biochem.* **1983**, *51*, 395−427.

(37) Becker, M. M.; Wang, Z. *J. Biol. Chem.* **1989**, *265*, 4163−4167.

(38) Basham, B.; Schroth, G. P.; Ho, P. S. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 6464−6468.

(39) Dechering, K. J.; Cuelenaere, K.; Konings, R. N.; Leunissen, J. A. *Nucleic Acids Res.* **1998**, *26*, 4056−4062.

(40) Vashakidze, R. P.; Prangishvili, D. A. *FEBS Lett.* **1987**, *216*, 217−220.

(41) Peck, L. J.; Wang, J. C. *Nature* **1981**, *292*, 375−378.

(42) Rhodes, D.; Klug, A. *Nature* **1981**, *292*, 378−380.

(43) Strauss, F.; Gaillard, C.; Prunell, A. *Eur. J. Biochem.* **1981**, *118*, 215−222.

(44) Hogan, M.; LeGrange, J.; Austin, B. *Nature* **1983**, *304*, 752−754.

(45) Nishimura, Y.; Torigoe, C.; Tsuboi, M. *Nucleic Acids Res.* **1986**, *14*, 2737−2749.

(46) Behling, R. W.; Kearns, D. R. *Biochemistry* **1986**, *25*, 3335−3346.

(47) Benevides, J. M.; Wang, A. H.; Rich, A.; Kyogoku, Y.; van der Marel, G. A.; van Boom, J. H.; Thomas, G. J., Jr. *Biochemistry* **1986**, *25*, 41−50.

(48) Peticolas, W. L.; Wang, Y.; Thomas, G. A. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 2579−2583.

(49) Buckin, V. A.; Kankiya, B. I.; Bulichov, N. V.; Lebedev, A. V.; Gukovsky, I. Y.; Chuprina, V. P.; Sarvazyan, A. P.; Williams, A. R. *Nature* **1989**, *340*, 321−322.

(50) Buckin, V. A.; Kankiya, B. I.; Sarvazyan, A. P.; Uedaira, H. *Nucleic Acids Res.* **1989**, *17*, 4189−4203.

(51) Brahms, S.; Fritsch, V.; Brahms, J. G.; Westhof, E. *J. Mol. Biol.* **1992**, *223*, 455−476.

(52) Chalikian, T. V.; Plum, G. E.; Sarvazyan, A. P.; Breslauer, K. J. *Biochemistry* **1994**, *26*, 8629−8640.

(53) Vorlickova, M.; Subriana, J. A.; Chladkova, J.; Tejralova, I.; Huynh-Dinh, T.; Arnold, L.; Kypr, J. *Biophys J.* **1986**, *71*, 1530−1538.

(54) Ivanov, V. I.; Minchenkova, L. E.; Burckhardt, G.; Birch-Hirschfeld, E.; Fritzsche, H.; Zimmer, C. *Biophys. J.* **1996**, *71*, 3344−3349.

(55) Hud, N. V.; Plavec, J. *Biopolymers* **2003**, *69*, 144−159.

(56) Drew, H. R.; Dickerson, R. E. *J. Mol. Biol.* **1981**, *151*, 535−556.

(57) Nelson, H. C. M.; Finch, J. T.; Luisi, B. F.; Klug, A. *Nature* **1987**, *330*, 221−226.

(58) Umezawa, Y.; Nishio, M. *Nucleic Acids. Res.* **2002**, *30*, 2183−2192.

(59) Mazur, A. K. *J. Comput. Chem.* **1997**, *18*, 1354−1364.

(60) Mazur, A. K. In *Computational Biochemistry and Biophysics*; Becker, O. M., MacKerell, A. D., Jr., Roux, B., Watanabe, M., Eds.; Marcel Dekker: New York, 2001; pp 115−131.

(61) Mazur, A. K. *J. Am. Chem. Soc.* **1998**, *120*, 10928−10937.

(62) Mazur, A. K. *J. Chem. Phys.* **1999**, *111*, 1407−1414.

(63) Mazur, A. K. *J. Phys. Chem. B* **1998**, *102*, 473−479.

(64) Mazur, A. K.; Sumpter, B. G.; Noid, D. W. *Comput. Theor. Polym. Sci.* **2001**, *11*, 35−47.

(65) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

(66) Arnott, S.; Hukins, D. W. L. *Biochem. Biophys. Res. Commun.* **1972**, *47*, 1504−1509.

(67) Jorgensen, W. L.; Chandreskhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys* **1983**, *79*, 926−935.

(68) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(69) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.

(70) Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1988**, *6*, 63−91.

(71) Minchenkova, L. E.; Schyolkina, A. K.; Chernov, B. K.; Ivanov, V. I. *J. Biomol. Struct. Dyn.* **1986**, *4*, 463−476.

(72) The current AMBER parameters reportedly overestimate the stability of cytosine C2′-endo sugar pucker with respect to the C3′-endo.[76,89] This force field defect is perhaps responsible for a delayed transition of cytosine sugars to C3′-endo, and it may somewhat reduce the A-philicity of poly(dG).poly(dC) DNA.

(73) The quasi-static profiles in our calculations should be distinguished from familiar S-shaped experimental plots of cooperative A↔B transitions.[12] In the latter case they result from two different contributions: (i) the shift in relative populations of A- and B-forms in the ensemble of DNA conformations and (ii) small deformations of the A- and B-conformations under varied hydration. Our simulations probe only the second contribution. Its relative weight in experiments is not known well. The Ising model of cooperative A↔B transitions considers only the first contribution.[12] The second one is effectively neglected by assuming that in small DNA fragments the transition is described by a step function. It is known, however, that, beyond the transition zone of hydration numbers, conformations of both B- and A-DNA change in a smooth and noncooperative manner and that these changes are rather significant.[90,91] Evaluation of equilibrium populations of A- and B-forms from these types of simulations rests beyond the current possibilities; nevertheless, reversion of A↔B transitions in a single MD trajectory can be observed in very long simulations under intermediate hydration (unpublished results).

(74) Stefl, R.; Trantirek, L.; Vorlickova, M.; Koca, J.; Sklenar, V.; Kypr, J. *J. Mol. Biol.* **2001**, *307*, 513−24.

(75) Lankas, F.; Cheatham, T. E., III.; Spaskova, N.; Hobza, P.; Langowskii, J.; Sponer, J. *Biophys. J.* **2002**, *82*, 2592−2609.

(76) Cheatham, T. E., III.; Cieplak, P.; Kollman, P. A. *J. Biomol. Struct. Dyn.* **1999**, *16*, 845−862.

(77) Wang, S.; Kool, E. T. *Biochemistry* **1995**, *34*, 4125−4132.

(78) Pilet, J.; Blicharski, J.; Brahms, J. *Biochemistry* **1975**, *14*, 1869−1876.

(79) MacKerell, A. D., Jr.; Wiórkiewicz-Kuczera, J.; Karplus, M. *J. Am. Chem. Soc.* **1995**, *117*, 11946−11975.

(80) Ivanov, V. I.; Krylov, D. Y. *Methods. Enzymol.* **1992**, *211*, 111−27.

(81) Suzuki, M.; Amano, N.; Kakinuma, J.; Tateno, M. *J. Mol. Biol.* **1997**, *274*, 421−35.

(82) McConnell, K. J.; Beveridge, D. L. *J. Mol. Biol.* **2001**, *314*, 23−40.

(83) Wahl, M. C.; Sundaralingam, M. In *Oxford Handbook of Nucleic Acid Structure*; Neidle, S., Ed.; Oxford University Press: New York, 1999; pp 117−144.

(84) Ling, H.; Boudsocq, F.; Woodgate, R.; Yang, W. *Cell* **2001**, *107*, 91−102.

(85) Ivanov, V. I.; Krylov, D. Y.; Minyat, E. E. *J. Biomol. Struct. Dyn.* **1985**, *3*, 43−55.

(86) Vorlickova, M.; Sagi, J.; Hejtmankova, I.; Kypr, J. *J. Biomol. Struct. Dyn.* **1991**, *9*, 571−578.

(87) Alexeev, D. G.; Lipanov, A. A.; Skuratovskii, I. Y. *Nature* **1987**, *325*, 821−823.

(88) Dickerson, R. E.; Drew, H. R.; Conner, B. N.; Wing, R. M.; Fratini, A. V.; Kopka, M. L. *Science* **1982**, *216*, 475−485.

(89) Olson, W. K.; Zhurkin, V. B. *Curr. Opin. Struct. Biol.* **2000**, *10*, 286−297.

(90) Lee, C.-H.; Mizusawa, H.; Kakefuda, T. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 2838−2842.

(91) Harmouchi, M.; Albiser, G.; Premilat, S. *Eur. Biophys. J.* **1990**, *19*, 87−92.

(92) Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1972**, *94*, 8205−8212.

(93) Mazur, A. K. *J. Mol. Biol.* **1999**, *290*, 373−377.

CT049926D