Spectroscopic and Molecular Dynamics Evidence for a Sequential Mechanisms for the A-to-B Transition in DNA

Ishita Mukerji

Wesleyan University, imukerji@wesleyan.edu

Follow this and additional works at: http://wesscholar.wesleyan.edu/div3facpubs

Part of the Molecular Biology Commons

Recommended Citation

http://wesscholar.wesleyan.edu/div3facpubs/83

This Article is brought to you for free and open access by the Natural Sciences and Mathematics at WesScholar. It has been accepted for inclusion in Division III Faculty Publications by an authorized administrator of WesScholar. For more information, please contact dschnaidt@wesleyan.edu, ljjohnson@wesleyan.edu.
Spectroscopic and Molecular Dynamics Evidence for a Sequential Mechanism for the A-to-B Transition in DNA

Kelly M. Knee,* Surjit B. Dixit,† Colin Echeverría Aitken,* Sergei Ponomarev, †, D. L. Beveridge, † and Ishita Mukerji*†

*Molecular Biology and Biochemistry Department, and †Chemistry Department and Molecular Biophysics Program, Wesleyan University, Middletown, Connecticut 06459

ABSTRACT The A-to-B form transition has been examined in three DNA duplexes, d(CGCGAATTCGCG)2, d(CGCGAATTGCGC), and d(CGCAAATTTCGC), using circular dichroism spectroscopy, ultraviolet resonance Raman (UVRR) spectroscopy, and molecular dynamics (MD) simulation. Circular dichroism spectra confirm that these molecules adopt the A form under conditions of reduced water activity. UVRR results, obtained under similar conditions, suggest that the transition involves a series of inter-mediate forms between A and B. Cooperative and distinct transitions were observed for the bases and the sugars. Independent MD simulations on d(CGCGAATTCGCG) show a spontaneous change from the A to B form in aqueous solution and describe a kinetic model that agrees well with UVRR results. Based on these observations, we predict that the mechanism of the transition involves a series of A/B hybrid forms and is sequential in nature, similar to previous crystallographic studies of derivatized duplexes. A simulation in which waters were restrained in the major groove of B DNA shows a rapid, spontaneous change from B to A at reduced water activity. These results indicate that a quasiergodic sampling of the solvent distribution may be a problem in going from B to A at reduced water activity in the course of an MD simulation.

INTRODUCTION

The B form of DNA is well known to be predominant in biological systems. However, conformational changes from B in the direction of A DNA are implicated in protein-DNA and drug-DNA interactions, and there is leading evidence for a functional role for A as well as B forms in genome structure and function (1). The A-to-B transition in DNA has also served as a prototype case for testing out and validating empirical energy functions and force fields used in molecular dynamics simulations on nucleic acids (2–4). A number of recent experimental (5–8) and theoretical (4,9–12) studies have reported on this system. Most studies to date have focused on the preferential stability of A and B DNA as a function of water activity. These results indicate that a quasiergodic sampling of the solvent distribution may be a problem in going from B to A at reduced water activity in the course of an MD simulation.

BACKGROUND

The structural difference between B and A forms of DNA resides essentially in the sugar puckers (φ) and the helix-base parameters x-displacement (XDP), inclination (INC), and slide (SLD), according to the definition of these parameters (13). In B DNA the nucleotide basepairs are perpendicular to the helix axis (INC = 0) and centered (XDP, SLD = 0), and

Submitted July 17, 2007, and accepted for publication January 10, 2008.

Kelly M. Knee and Surjit B. Dixit contributed equally to this work.
Address reprint requests to D. L. Beveridge, Chemistry Department and Molecular Biophysics Program, or Ishita Mukerji, Molecular Biology and Biochemistry Department, Wesleyan University, Middletown, CT 06459. E-mail: dbeveridge@wesleyan.edu.
Editor: Jonathan B. Chaires.

© 2008 by the Biophysical Society
0006-3495/08/07/257/16 $2.00
doi: 10.1529/biophysj.107.117606
the sugar pucker is $C_2$-endo. In the A form, the basepairs are 
tilted ($\text{INC} = 20^\circ$) such that they are displaced from the helix 
axis by $-4.0$ Å, and the sugar pucker is $C_3$-endo.

DNA is unique in the extent to which its structure depends 
on solvation. The early fiber diffraction experiments that 
provided experimental data crucial to the discovery of the 
double helix also revealed that DNA structure was sensitive 
to the relative humidity of the sample fibers (14). The B form 
is stable at relative humidities of 95% and the A form is 
preferentially stabilized below 74%. Subsequent fiber dif-
fraction by Arnott and co-workers (15) elucidated the mo-
lecular structures of the canonical B and A forms of DNA.
The preferential stability of B- and A-form DNA in solution 
was first reported by Ivanov and co-workers (16) based on 
CD spectroscopy, and was found to depend on water activity 
in a manner that parallels the results on fibers. Malenkov et al. 
reported the B-to-A transition to be cooperative (17).

There has been considerable research into the forces in-
volved in the preferential stability of right-handed DNA 
helices and the A-to-B transition. Solvent accessibility (18), 
base-stacking interactions (19), the economics of phosphate 
hydration (20), hydrophobic pressure, enthalpic stabilization 
of B form by the minor-groove spine of hydration (21,22), 
and electrostatic effects associated with the explicit organi-
zation of mobile counterions in the vicinity of the polyanionic 
double helix (23–26) have all been invoked as plausible and 
possible explanations. Based on analysis of MD simulation 
trajectories, Jayaram et al. (27) proposed that the molecular 
origins of the conformational preferences of A and B DNA in 
water and 85% EtOH lie primarily in the differential free-
energy contributions from interphosphate repulsion, coun-
terion condensation, and solvation.

Recently, Vargason et al. (8,28) reported a crystallo-
graphic map of the conversion of B to A DNA based on a set 
of 13 crystal structures of d(GGC GCC), with various struc-
tural intermediates trapped by methylating or brominating the 
cytosine bases. This study provided a model for the B-to-A 
conformational transition in terms of a set of distinct con-
formational intermediates, each of which is an equilibrium 
state with respect to the free energy of the particular crystal.
The duplex in some instances was indicated to be partly A 
form and partly B form, depending on local sequence com-
position, and some of the intermediate structures bear little 
resemblance to either the starting or ending states. If the 
mechanism of the transition can be successfully modeled 
based on these intermediate structures, the conversion of 
SLD precedes the change in base-pair inclination in going 
from B to A, and changes in XDP occur continuously. In 
these results, sugar pucker shows a reasonably sharp transi-
tion point from B-like $C_2$-endo to A-like $C_3$-endo structures 
(with some intermediate structures exhibiting a mixture of 
$C_2$- and $C_3$-endo-type sugars). In the intermediate structures, 
an extension and unwinding of the helix occurs along with 
the change in sugar pucker. It remains to be established 
whether these results from crystallography carry over from 
equilibrium states to dynamics and are applicable to the 
solution as well as to the crystalline state. Experimental data on 
the dynamics of the B-to-A transition are relatively sparse. A 
study by Jose and Porshke (29) reports on the dynamics of the 
A-to-B transition based on stopped-flow, electric-field-jump 
experiments. They report time constants for the transition in 
the range of 10 µs, and present evidence for a significant 
activation-energy barrier in going from B- to A-form DNA.

A number of CD and related studies on DNA have been 
reported which characterized the sequence dependence of the 
preferential stability of A and B forms of DNA (20). In gen-
eral, ApA and ApT steps are relatively rigid and resist trans-
ition from B to A, and the B $\rightarrow$ A transition is viable only for

![FIGURE 1 Structural representation of the d(CGCGAATTTCGCG)2 DNA sequence.](image-url)
sequences with a sufficient number of GC basepairs (30,31). Similarly, conversion to the A form using Co$^{3+}$\((\text{NH}_3)_6\) requires GpG steps and has not been demonstrated with sequences containing a majority of AT basepairs (25,32). As noted above, the B form is favored under conditions of high water activity in solution or high relative humidity in fibers. When the water activity is lowered, or the relative humidity of a DNA fiber is $<\sim75\%$, B-form sequences convert spontaneously to A form. Several groups (33,34) have examined the experimental data on the preferential stability as a function of sequence and B/A philicity has been expressed based on dimeric and trimeric base-pair steps. It is important to note that the solution structures of the A- and B-DNA endpoints of the transition are expected to contain some structural differences relative to the ideal canonical A and B forms of DNA inferred from fiber diffraction (35).

In addition to CD, infrared (36–38), and Raman (5,39) spectroscopic methods have been applied to the study of B- and A-form DNA. Raman spectroscopic markers for the B and A forms of DNA are well established (5,39). Certain structural parameters of DNA can be obtained from these methods, particularly the configuration of the sugar rings, the angles of phosphates, and overall base geometries. A disadvantage of these methods, including CD spectroscopy, is that the information on sequence dependence cannot generally be resolved. NMR studies involving 2D nuclear Overhauser effect spectroscopy do not resolve the DNA helical parameters, but the recent addition of reduced dipolar coupling has measurably improved these determinations. However, elucidation of a reliable molecular structure from NMR data remains challenging.

UVRR has introduced some new dimensions to the study of nucleic acid structure (40–42). The resonance enhancement intrinsic to UVRR spectroscopy makes it possible to selectively investigate the behavior of individual types of nucleotide bases and, in the case of propitious sequences, individual bases. As a consequence of this resonance selective enhancement and detection, stacking interactions and H-bonding of relevant functional groups can be attributed to distinct residue types. Thus, UVRR spectroscopy can be a source of more detailed information regarding local structure of DNA as a function of sequence and conformation.

In previous studies, the UVRR method has proven useful in evaluating the structure and stacking of DNA molecules, cation coordination, and H-bonding (43,44). From measurements of the premelting transition in DNA A-tracts, the relative strength of cross-strand 3-centered H-bonds could be determined (45). More recently, UVRR spectroscopy has been applied to the study of protein-DNA interactions, and subtle changes in DNA structure associated with protein binding were readily detected (46–48).

MD computer simulation including explicit consideration of solvent has been applied extensively to model the dynamic structure of DNA oligonucleotides (Fig. 1) (49–54). MD simulations are found to be generally in close accord with NMR-derived solution structures (7,55–59). MD simulations have provided new insights into the solution structures of these molecules, which include essentially straight A-tracts (60), highly flexible purine-pyrimidine steps that can serve as a locus for axis bending (56), and support for the "non-A-tract" model of DNA curvature (49). MD studies of sequence effects on B-form DNA structures have been advanced by recent simulations on all 136 unique tetranucleotide base-pair steps (61,62) and has led also to improvements in nucleic acid force fields (63).

The interconversion from B to A DNA has been investigated via molecular dynamics simulations as described in several recent articles (2,4,11,12,64). Cheatham and Kollman (2) were the first to report MD studies of A- and B-form conformational stability in explicit solvent. Subsequently, Sprous et al. (4) reported MD simulations on d(CGCGAA-TTCGCG)$_2$ in A and B forms. The motivation for these studies was directed more toward assessing the performance of force fields, with the A-to-B transition as a prototype case. Both Cheatham et al. (2) and Sprous et al. (4) independently observed that in MD based on the AMBER parm94 force field, B-form DNA was stable in water and the A form was stable in a mixed ethanol/water solvent with reduced water activity. The MD model A-form DNA in aqueous solution was observed to convert spontaneously into the B form. However, both studies reported that B-form DNA did not spontaneously convert to the A form in the low-water-activity mixed solvent at 300 K. This was initially addressed as a force-field limitation and a subsequent local alteration of the force field was designed to make this transition more favorable (65). However, the performance of the resulting MD model DNA deteriorated with respect to other structural features. In these studies, the expectation of a B-to-A transition at low water activity was based on generic patterns of environmental effects on DNA conformational stability rather than specific behavior of a DNA sequence and the question of whether in fact this sequence was observed experimentally to form A DNA at low water activity was raised by Pichler et al. based on infrared studies of hydrated films (38). The exact solvent conditions that favored A- or B-form DNA for this sequence were not well delineated. Elucidation of these issues, as well as the mechanism underlying them, motivated the study described here.

Recently, Pastor studied the B-to-A transition in a TATA box sequence using MD based on the CHARMM nucleic acids force field (12) and found evidence in support of a sequential "slide first, roll later" mechanism that is opposed by DNA electrostatics and favored by increasing condensation of sodium ions. The latter phenomenon was noted earlier by Sprous et al. (4) and figured significantly in the free energy calculations of Jayaram et al. (27), with cross correlations between various theoretical studies providing some enhanced validation of the results. Banavali and Roux (64) recently reported a free energy profile for the B-to-A transition of d(CTCGAG) in water with minimal salt based on the

Biophysical Journal 95(1) 257–272
CHARMM force field. The free energy difference between canonical A and B forms was found to be at least 2.8 kcal/mol, and the simulation with root mean-square deviation constraints sampled a continuum of hybrid structures with no well defined local minima between the two forms. The details of this calculation raise a number of additional issues, since some base-flipping events and other effects were noted that are presumably artifacts. Also the MD structures of A and B DNA deviate significantly from the idealized canonical forms, which are not appropriate endpoints for a determination of the free energy difference in solution.

In this study, the mechanism of the A-to-B transition is revisited to address the relative importance of base step, solvation, and electrostatics in the transition. The combined methodologies of UVRR spectroscopy and MD simulation provide new insights into the transition, since the UVRR results identify regions of change and stability, and the MD simulations provide a structural basis for interpretation of Raman data. These results particularly point to a sequential mechanism for the transition, in which reorganization of water plays a central role.

Methods

Oligonucleotides
Experiments were carried out on three dodecamer sequences, d(5'-CCGGAATTCGCG-3'), d(5'-CCGGAATTCGCTGG-3'), and d(5'-CCGAAATTCGCG-3'), as a function of water activity. The 5' to 3' sequences are shown; however, all experiments were performed on duplexes with the appropriate complementary strands. Oligonucleotides were synthesized in 1-mmol quantities (Integrated DNA Technologies, Coralville, IA). The deblocked and desalted oligomers were purified by polyacrylamide gel electrophoresis. Purity was checked by analytical gel electrophoresis. For all spectroscopic experiments, samples were dialyzed against a buffer containing 0.1 M NaCl and 0.67 mM NaPO₄, pH 7.0. For the trifluoroethanol (TFE) experiments, only buffer was used, as minimal salt concentrations were chosen to promote oligonucleotide solubility. Duplex samples were prepared by heating to 90°C in a water bath for 5 min, followed by slow cooling to room temperature. Annealed samples were subsequently diluted 10-fold with either TFE or in a water bath for 5 min, followed by slow cooling to room temperature. Samples were examined at room temperature and were normalized to the TFE band occurring at 1459 cm⁻¹. Curve fitting of the spectra was accomplished with a mixed Lorentzian and Gaussian function, using constant peak frequencies. All data manipulation and analyses were performed using GramsAI (ThermoGalactic, Salem, NH).

Percent relative molar hydration (MH) was calculated as the mole fraction of water at each TFE concentration using the equation

\[ X_{MH} = \frac{(n_{w} / (n_{w} + n_{TFE})) \times 100}{100} \]  

where \( n_{w} \) refers to the number of moles of water or TFE. Concentrations of DNA and salt were neglected, as they remained constant. Transition midpoints were treated independently and analyzed using a Boltzmann function with Origin v. 6.0 (MicroCal, Northampton, MA): \( y = \frac{(A_1 - A_2)}{1 + e^{(x-x_0)/\theta}} + A_2 \).

Molecular dynamics simulations
Four simulations on the duplex d(CCGGAATTCGCG) are considered here. Simulations on both the B and A forms of this sequence reported earlier were all redone and extended to much longer trajectories. Two of these involve the simulation of the canonical A-form structure in an ~85% (v/v) ethanol/water mixture, i.e., ~30% relative MH, and in water, and the third and fourth are of the canonical B-form structure in water and at 30% MH. Considerable new analysis of solvation was carried out for this project. Details of the calculations are as follows: The AMBER utility NUCGEN was employed to create the canonical A- and B-DNA starting structures based on fiber diffraction data (35). For simulations in the ethanol/water mixture, the united-atom OPLS-ethanol model (66), which is computationally effective and closely approximates the correct density, heat capacity, and heat of vaporization observed for ethanol (67), together with the TIP3P water molecules, was used at a concentration of 30% MH relative to ethanol. The simulation protocol is similar to those reported earlier (4,9). A “biphasic” initial configuration, as described by Cheatham et al. (9), was constructed by adding TIP3P water molecules in the first 6 Å from the DNA surface followed by the requisite number of ethanol molecules to make up the 30% MH ethanol/water solvent mixture. For the simulations in water, the TIP3P (68) water model was employed for the solvent such that the solvent extends up to 13 Å from the surface of the solute in all the directions. Electroneutrality was established by adding 22 sodium cations (69) to the system. The positions of the ions were randomized such that they were at least 5 Å from the DNA and 5 Å away from each other. All MD simulations were performed in TPN ensemble using AMBER version 7 or 8 and the parm94 version of the Cornell et al. force field (70), and employing the particle-mesh Ewald procedure (71,72) for the treatment of long-range interactions. The MD protocol of minimization, heating, equilibration, and production dynamics was carried out in the following steps. Minimization involved 100 steps of steepest descent followed by 250 steps of conjugate gradient method. Heating from 0 to 300 K was done over 10 ps, followed by equilibration at 300 K for an additional 40 ps with SHAKE constraints (73). During the equilibration phase, a flat well restraint on the C1'-C2'-C3'-C4' torsion was introduced to maintain the angle between 30° and 40° in the A-form structure, set at 25 kcal/mol/deg² for the first 30 ps and then reduced to 5 kcal/mol/deg². The production simulations were pursued without any restraints for simulation times of 10 ns and longer using 2-fs time steps.

The analysis of the composition of a molecular fluid requires an interpretation of the statistical distribution functions in structural and energetic terms. A theoretical approach for this problem was mapped out several years ago for pure liquids by Ben-Naim (74) based on the generalized molecular distribution functions and the closely related quasiconstituent distribution function involves developing the distribution of particles with certain well-defined values of a compositional characteristic on the statistical state of the system (75). The basis for a general compositional analysis of the statistical state of molecular fluids must be a unique definition of the local solution.
environment of each identifiable substructure-atom, function group, or subunit of the solute. The proximity criterion accomplishes this by uniquely identifying each solvent molecule with a well-defined solute entity in each configuration. The proximity analysis partitions the solvent according to the proximal solute atoms and calculates distribution properties for each partition. The solute atoms can be partitioned independently or into functional groups defined by their chemical identity. All the proximity calculations were performed using the MMC program (76). The solvent-accessible surface area calculations on the snapshots in the MD trajectory have been performed using the program SurfRacer 3.0 (77). The atomic radius set reported by Alden and Kim (18) was employed. A spherical probe radius of 1.4 Å corresponding to the regular radius of a water molecule was employed in all the solvent-accessible surface area calculations.

RESULTS

CD spectra

CD spectra for all three duplexes considered in this study were examined to make sure they each exhibited the expected preferential stability of the B and A forms at high and low water activities and were suitable for our study. The CD spectrum of the d(CGCGAATTCCGC)2 duplex is shown as a function of increasing relative percentages of TFE in Fig. 2 a. At relatively low concentrations of TFE (>77% MH), the spectrum exhibits a maximum at 283 nm and a minimum at 252 nm of comparable magnitude. These conservative features above 220 nm are consistent with those previously observed for B-form DNA (78). As the TFE concentration is increased, the maximum at 283 nm decreases in intensity and shifts to 270 nm (Fig. 2, inset). In previous work, spectroscopic features of A-form DNA at low ionic strength included a peak at 270 nm, the shape of which was strongly sequence-dependent (16,31). To ensure that the transition was not associated with the palindromic nature of the d(CGCGAATTCCGC) sequence, similar CD studies were performed with other DNA dodecamers, which are of similar base composition but different sequence. For all duplexes examined, a decrease in ellipticity at 250 nm coupled with a shift in maximum to 270 nm is observed with increasing concentrations of TFE (Fig. S2 in Supplementary Material, Data S1). The magnitude of the ellipticity at 270 nm measured at 41% MH lies in the 10–20 Δε (M⁻¹ cm⁻¹) range. This is in good accord with previous measurements of the B-to-A conversion in TFE, in which peaks of similar ellipticity were observed at 270 nm and were identified as being diagnostic of the A form (16,31). For all duplexes examined, the spectrum obtained at 41% MH exhibits an increase in ellipticity at 270 nm, a decrease at 250 nm and an increase at 210 nm relative to the B-form spectrum. These spectral features have been previously assigned to A-form DNA (33,78) and indicate that the d(CGCGAATTCCGC)2 duplex has adopted the A form with increasing concentrations of TFE.

The CD intensity at 252 nm for d(CGCGAATTCCGC)2 is plotted as a function of molar hydration in Fig. 2 b. The change in intensity occurs over a relatively narrow range of hydration, consistent with previous studies. The results indicate that the global change in conformation from B- to A-form DNA occurs at 65% molar hydration. The d(CGCGAATTCCGC)2 duplex, which is exactly the same base composition as the d(5′-CGCGAATTCCGC-3′) duplex, has a transition midpoint of 68.3%, whereas the d(5′-CGCGAATTCCGC-3′) duplex exhibits a transition midpoint at 63.2% molar hydration (Fig. S2 in Data S1). As expected, the increase in number of AT basepairs in going to d(5′-CGCGAATTCCGC-3′) results in a transition midpoint at a significantly lower level of hydration, that is, the more AT-rich sequence converts to the B form at lower water activity, consistent with conventional wisdom on the relative propensities of the basepairs to form A- and B-form DNA. In summary, the CD studies indicate that A/B interconversion does occur in each of the sequences studied. The average transition midpoint is consistent with the general idea of the preferential stability of B form at high water activity and A form at low water activity and of A/B philicity as a function of sequence.

UVRR spectra: ribosyl conformation

Three different excitation wavelengths were used to gain a molecular picture of the DNA oligonucleotides changing from A to B form in solution (Fig. 3). To selectively monitor the deoxyribose-phosphate backbone, an excitation wavelength of 210 nm was used (43,45). At this excitation wavelength, the peak intensities of ribose modes for dG residues are preferentially enhanced (41,42). The vibrational frequencies of ribose modes are sensitive to conformation, with the C2′-endo configuration associated with a peak at 685 cm⁻¹ and the C3′-endo configuration associated with a peak at 665 cm⁻¹ (5,39). This feature is often used as an indicator of DNA conformation, but the results presented below indicate that ribose conformation does not necessarily reflect an overall A- or B-form structure. In this work, we find that this mode shifts to lower frequency as the relative hydration is decreased (Fig. 3). A transition midpoint of 68.3 ± 2.3% is obtained from a plot of the frequency as a function of relative hydration (Fig. 4); this value is within the range of that determined by CD spectroscopy (see above) (Table 1).

To estimate the relative populations of the C2′-endo and C3′-endo configurations, the Raman bands were fit assuming that only these two components were contributing to the observed band. This fitting indicates that even at 41.5% MH, the population of dG ribose modes in the C3′-endo configuration is 94% and not 100% (Fig. 5 a). This reflects a situation wherein the dynamic structure of the duplex in solution involves both B-form and A-form ribose structures in the Boltzmann ensemble. Because UVRR spectroscopy superimposes all dG ribose modes, it is not possible to distinguish which individual ribose moieties have adopted the C3′-endo configuration at 41.5% MH and which ones have remained in the C2′-endo configuration. The sugar conformation of other
bases cannot be examined because of spectral interference from TFE bands.

**UVRR spectra: base interactions**

An advantage of the UVRR technique is the ability to selectively enhance contributions from the different bases through judicious choice of excitation wavelength. Thus, dG residues are enhanced through an excitation wavelength of 240 nm and dA residues contribute strongly to the spectrum when excited at a wavelength of 260 nm. Contributions from dT bases are also observed using an excitation wavelength of 260 nm. Under these excitation conditions the modes that are observed primarily arise from ring-stretching vibrations.
Because of the effect of Raman hypochromism, which is directly related to absorption hypochromism, the intensity of the base modes can be related to base-stacking interactions. For all of the modes examined, the intensity of the modes increases upon adopting the B form (Fig. 3). The increase in intensity is attributed to a reduction in base-stacking interactions because of the greater distance or rise in base steps (3.4 Å for B vs. 2.8 Å for A) and the more pronounced helical twist between base steps (36° for B vs. 33° for A) in B-form DNA. The increased intensity upon adoption of the B form has also been observed in other Raman studies of DNA conformation (5,39,47).

The intensity changes of the 1362 and 1485 cm⁻¹ modes obtained with an excitation wavelength of 240 nm are shown as a function of molar hydration in Fig. 4. Based on a comparison of the relative Raman cross sections (41,42,79), these modes arise primarily from dG stretching motions rather than dA residues because of the 240-nm excitation wavelength. The mode occurring at 1362 cm⁻¹ arises from dG C2-N3-C4-N9 stretching motions, with contributions from N7 ring-stretching vibrations and the 1485-cm⁻¹ mode arises from C2-H and C8-H bending modes, coupled with C8-C9 stretching motions. Since these motions lie primarily in the plane of the ring, the enhancement mechanism is associated with the absorption of the base itself and changes in intensity can be related to changes in hypochromicity associated with stacking interactions. The transition midpoints obtained from the intensity changes as a function of molar hydration yield a value of 85 ± 1.0% (Table 2). Interestingly, this transition midpoint differs significantly from that obtained for the change in ribose conformation and suggests that the base-stacking geometry of the dG residues changes at a higher molar hydration compared to the ribose.

Using an excitation wavelength of 260 nm, the intensities of modes at 1332 and 1575 cm⁻¹ can also be monitored as a function of hydration (Fig. 4). These modes, which arise primarily from dA residues at this excitation wavelength, result from imidazole and pyrimidine ring-stretching vibrations, respectively. The intensity change as a function of molar hydration yields transition midpoints of 69% and 75%, respectively. These midpoints are suggestive of a mean transition midpoint for the AT basepair of 72% MH, which differs significantly from that obtained for the ring-stretching modes arising from dG residues (Fig. 4). Thus, these results indicate that dA residues do not change their base-stacking geometries under the same conditions as dG residues and that the dA residues require significantly lower levels of molar hydration for transition.
hydration to convert to the A form. For both base types, the transition midpoints observed are at higher water activity relative to that observed for the ribose conformation. These findings are in good qualitative agreement with the structure of the d(CCCCGGGG)2 duplex sequence determined by CD and NMR spectroscopic techniques (6), in which the base-stacking properties were found to be A-like and the sugar pucker was observed to be B-like. The NMR-determined structure was interpreted to be an intermediate in the A-to-B transition (6).

UVRR spectra of the A and B forms of DNA also provide some insight into the relative hydration of the grooves through measurement of H-bonding strength. The relative strength of H-bonding interactions is inferred from the carboxyl stretching frequency of relevant functional groups. A decrease in frequency results from a reduction in the force constant of the carboxyl bond and is correlated with increased strength of H-bonding interactions (45,80). The thymine C4=O group points into the major groove and is readily monitored using an excitation wavelength of 210 nm (41,45). Under conditions of increasing concentrations of TFE, the frequency of the C4=O shifts from 1662 to 1653 cm⁻¹ (Fig. 4). The downshift in frequency is indicative of stronger H-bonding in A-form than in B-form DNA. Crystallographic structures have revealed a more ordered water structure in the major groove of A-form DNA, and the decrease in frequency of the C4=O is attributed to stronger H-bonding with locally ordered water molecules. Examination of the MD simulations and canonical structures also indicates that H-bonding geometries of the A and B forms are not substantively different.

In summary, the UVRR results make it possible to monitor the ribose conformations, as well as the base stacking of dG, dA, and dT. An examination of the series of equilibrium states of the system as the relative hydration is increased reveals that the sugar moieties change to C2-endo followed by A- to B-type changes in base-stacking interactions. There is evidence of stronger H-bonding to solvent in the major groove of A-form DNA than in the B-form. The UVRR results thus support a cooperative but sequential mechanism for the B-to-A transition, as opposed to an all-or-none, simple two-state mechanism.

### MD simulations

Since earlier reported MD involved relatively short trajectories, MD simulations on B and A forms of d(CGCGAATTCCGCG) were extended to 60 ns and 20 ns, respectively. The overall behavior of the extended simulations was essentially similar to that reported (Fig. 1) (4), and provides a more rigorous basis for detailed analysis of the solvation. The MD simulation most relevant to the mechanism of the B-to-A transition is a simulation on the d(CGCGAATTCCGCG) duplex in aqueous solution, beginning with the sequence in the canonical A form. A comparison of the canonical A- and B-form structures with the ensemble of structures generated by the MD simulations is given in Fig. 1. During the initial phase of the MD, there is a rapid interconversion of the sequence from A- to B-form, and a detailed structural analysis provides the MD-predicted mechanism. In going from A to B, the sugar pucker and the helical parameter INC transition to B-form values within 500–750 ps. The basepair displacement XDP converts from A to B values within the 750–1250 ps time range. The A-to-B interconversion is complete by 1500 ps of MD, and the resulting B-form structure is quite similar to that found in the simulation of the sequence in aqueous solution that began with the canonical B form. This finding agrees with experiment, i.e., the B form is the preferred conformation in aqueous solution.

To be more quantitative, we have monitored the relaxation times of these properties in the course of the A-to-B transition observed during the simulation starting with the A-form.
structure in water. The relaxation times are calculated on the basis of exponential fit of the decay in the autocorrelation function of these parameters (Table 2). Neglecting the outlier observed in sugar pucker relaxation time of the G10 nucleotide, we observe that, on average, the sugar transition from C$_{3'}$-endo to C$_{2'}$-endo conformations relax faster than XDP and INC. With respect to sequence effects, the AATT tract converts to B form first, followed at a later point by the GC tracts. These observations are consistent with those obtained from UVRR spectroscopy, described above, which indicate that the mechanism of the A-to-B transition is sequential, with sugars converting from A form to B form, followed later by the changes in base stacking. The structure passes through an A/B hybrid form, with the AT tract converting at lower relative hydration than the GC tracts. The MD results point to populations of mixed ribose conformation for A DNA in ethanol and B DNA in water but also indicate that small fractions of other ribose conformations, not resolved in the UVRR spectra, may be present. Interestingly, MD simulations predict that even A DNA in ethanol/water contains $\sim$20% of the C$_{2'}$-endo configuration and that the ribose conformation is more heterogeneous in this conformation than in B DNA (Fig. 5 b). In summary, the MD and UVRR results agree well on the issue of a sequential mechanism, i.e., one in which changes happen in sequence as opposed to “all or none.” Since water activity is the agent of change, we now look to the simulation results to obtain an understanding of the manner by which solvation effects the transition.

Our analysis of solvation examines the general nature of the A and B forms of d(CGCGAATTCGCG) based on calculation of solvent accessibility, and is followed by a detailed analysis of the MD results using the proximity method. Calculations of the solvent-accessible surface area of the four nucleotides in the A- and B-form structures, computed from the MD trajectories, are shown in Fig. 6. The solvent accessibility analysis is broken down in terms of the constituent atoms in the sugar, phosphate, and major and minor groove regions. Although the grooves together contribute $\sim$20% of the net solvent-accessible surface area, the related change on going from the A- to the B-form structures is large, and comparable to that found for the phosphate and sugar groups (Fig. 6 c). There is a net reduction in the solvent-exposed surface area for both the AT and GC basepairs in the A-form structures, compared to the B-form, consistent with the idea of reduced hydration in A DNA. Moreover, the change in solvent-accessible surface area for the two forms is distinctly different for the AT and GC basepairs. Specifically, the solvent accessibility of AT basepairs in the major groove in the

FIGURE 6 (a and b) Percentage composition of the solvent-accessible surface areas of the G, A, C, and T nucleotides in the MD simulation of the central 10 basepairs of the CGCGAATTCGCG sequence in A-form DNA (a) and B-form DNA (b). (c) Difference in the average solvent-accessible surface area (in Å$^2$) between the A and B conformations for the four nucleotides and the AT and GC basepairs.
A-form structure is greatly reduced. This finding is indicative of a higher level of water rearrangement in the immediate proximity of the AT basepairs, compared to the GC basepairs, during the B-to-A transition. This finding is in good qualitative agreement with experimental results, which indicates that AT-rich regions convert from the A to B form at a lower level of relative hydration. Thus, the differences in accessible surface area for GC and AT basepairs can be correlated with their relative differences in energy in the transition, and these differences are reflected in the experimental data.

In the B-form structures, phosphate groups are the most solvent-accessible components of the nucleotides, contributing 49.5% of the total, followed by sugars at 30.5%. The corresponding contribution of the phosphates and sugars in the A-form structure is 47.25 and 34.25%, respectively. For all the nucleotides, the solvent-accessible surface area of the sugar and minor groove regions is larger in the A form, but is consistently lower for the phosphate group and the major groove. The hydrophobic sugar moieties are more solvent-exposed in A-form structures, whereas the charged phosphate groups experience decreased solvent exposure since the interphosphate distance is only sufficient to accommodate one water molecule. The increased accessibility of the surface to the solvent is a result of the shallow nature of the minor groove region in the A-form structure, whereas the concomitant decrease in major groove accessibility is a result of the deeper, narrower groove dimension (81).

The details of the MD-calculated water structure around the A and B forms of DNA have been examined based on proximity analysis of water molecules in MD trajectories of A-form DNA in an ethanol/water mixture and of B-form DNA in water (82). The analysis is presented in Table 3 in terms of the four structural moieties, i.e., sugar, phosphate, and major and minor grooves, for each of the central eight basepairs in the dodecamer DNA. In particular, we report the volume of the first solvation shell, the corresponding coordination numbers, and binding energies for the stable A- and B-form structures. Since the ethanol/water mixture contained 528 water molecules, the number of water molecules closest to that in the B-DNA simulation in aqueous solution was achieved through addition of TFE, whereas in the computations, the cosolvent used was ethanol. Since ethanol can precipitate DNA, TFE was used in solution experiments. Although hydrogen-bonding properties of the two solvents are different, the changes observed are mainly attributed to a reduction in water activity. This conclusion is supported by our analyses, which indicates that the MD and UVRR results agree well on the issue of a sequential mechanism for the A-to-B transition in DNA and the detailed nature of it at the molecular level. The UVRR and MD results thus collectively support the idea of a cooperative but sequential mechanism for the B-to-A transition, as opposed to an all-or-none, simple two-state mechanism.

In comparing experimentally observed and theoretically calculated results, it is important to note that the experimental results report the A/B ratios for a series of equilibrium conditions.
mixtures, whereas the MD simulation models a kinetic pathway between the calculated solution structures of the A and B forms. The features of the A-to-B transition indicated by the UVRR spectroscopy and the MD simulations are consistent with the diverse experimental results reviewed in Background, which pertain to the general aspects of conformational stability. At a more detailed level, we have compared the UVRR/MD sequential model of the transition to the mechanism proposed on the basis of x-ray diffraction of various derivatized intermediates (see Fig 9 for our MD-calculated results). A comparison of the equilibrium model from experiment and the kinetic model from MD in terms of the order of events agrees well with the idea of a sequential mechanism. Comparison with the crystal structure model indicates that the results obtained on the crystalline solid carry over to the solution state.

A new perspective on the A-to-B transition was indicated by the observations that H-bonding to the dT C₄=O group is stronger in the A form and the transition midpoint occurs at a high molar hydration (87.5% MH) (Fig. 4; Table 2). The increased H-bonding strength was attributed to greater solvent interaction and MD simulations point to increased localization of water and ions in the major groove in the A form. These results and the accessible surface area analyses support the idea that reorganization of the solvent is a major determinant in the transition.

We turn now to a methodological issue relevant to MD simulations on DNA. Many groups have independently

### TABLE 3 Proximity analysis of A- and B-form DNA trajectories

<table>
<thead>
<tr>
<th>Basepair position</th>
<th>Basepair</th>
<th>Volume of first solvation shell*</th>
<th>Coordination number ((k))¹</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>B DNA</td>
<td>A DNA</td>
</tr>
<tr>
<td>3</td>
<td>CG</td>
<td>Sugar</td>
<td>315.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phosphate</td>
<td>321.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Major groove</td>
<td>282.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Minor groove</td>
<td>234.3</td>
</tr>
<tr>
<td>4</td>
<td>GC</td>
<td>Sugar</td>
<td>327.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phosphate</td>
<td>319.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Major groove</td>
<td>275.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Minor groove</td>
<td>229.6</td>
</tr>
<tr>
<td>5</td>
<td>AT</td>
<td>Sugar</td>
<td>307.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phosphate</td>
<td>308.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Major groove</td>
<td>325.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Minor groove</td>
<td>229.9</td>
</tr>
<tr>
<td>6</td>
<td>AT</td>
<td>Sugar</td>
<td>302.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phosphate</td>
<td>322.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Major groove</td>
<td>321.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Minor groove</td>
<td>225.8</td>
</tr>
<tr>
<td>7</td>
<td>TA</td>
<td>Sugar</td>
<td>299.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phosphate</td>
<td>326.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Major groove</td>
<td>327.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Minor groove</td>
<td>231.0</td>
</tr>
<tr>
<td>8</td>
<td>TA</td>
<td>Sugar</td>
<td>311.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phosphate</td>
<td>320.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Major groove</td>
<td>335.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Minor groove</td>
<td>236.6</td>
</tr>
<tr>
<td>9</td>
<td>CG</td>
<td>Sugar</td>
<td>321.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phosphate</td>
<td>328.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Major groove</td>
<td>266.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Minor groove</td>
<td>242.6</td>
</tr>
<tr>
<td>10</td>
<td>GC</td>
<td>Sugar</td>
<td>308.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phosphate</td>
<td>331.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Major groove</td>
<td>292.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Minor groove</td>
<td>235.4</td>
</tr>
</tbody>
</table>

*Volume of primary solvent shell (Å³) derived by the proximity method.

¹Number of water molecules in the primary proximity region.
shown that MD simulations in water using the popular AMBER parm94 force field equilibrate in a B-form-like structure regardless of whether the simulation starts in an A- or B-form structure (4,65). In an ethanol/water solvent mixture of 30% relative MH, a simulation started from the canonical A-form structure remains A-like (4,65). According to conventional wisdom, a B-form structure in such a low-water-activity state should transition to an A-like structure, but in this case it was originally reported, and confirmed in this study, that the MD structure remains B-like in this case. This discrepancy has raised questions about the quality of the force field, the length of the simulation, and the sequence dependence of the transition (4,65). To pursue this issue, we have extended our MD solvent analysis using a numerical procedure to determine the average solvent positions, taking into consideration the diffusional interchanges that would occur as a result of solvent motion. These so-called generic sites (84) represent the locations of the peaks in the 3D density distribution of the solvent. Solvation sites computed from the trajectories of A- and B-DNA structures in the 30% MH mixture are shown in Fig. 7. The stable A-form structure in the ethanol/water mixture shows a strong preference to localize water and ions in the major groove, in accord with our Raman spectroscopic results and the solute-solvent interaction energy calculated using the proximity analysis. However, the solvent analysis for the corresponding B-form structure in the ethanol/water mixture indicates that the solvent distribution in the groove region has not achieved localization of water and ions in the major groove equivalent to that observed for the A-form structure.

FIGURE 7 Representation of generic solvent site analysis (84) of the MD trajectories of A- and B-form d(CGCGAATTCGCG)$_2$ DNA in the ethanol/water mixture with the water locations color-coded by fractional occupancy. A gradient coloring scheme is employed for the generic water positions, with the sites of high occupancy (>0.9) shaded blue, those of low occupancy (~0.3) shaded red, and those of intermediate occupancy shaded green. The C, G, T, and A nucleotides are colored yellow, green, blue, and red, respectively. Sodium ions are colored magenta. The figures have been generated using the program PyMOL (DeLano Scientific, Palo Alto, CA).
To assess the importance of water structure in the major groove of the DNA in facilitating the transition from B to A structures, we constructed a simulation “experiment” in the ethanol/water mixture that began with a small cluster of 30 water molecules restrained to the vicinity of major-groove surface atoms of the B-form starting structure. The results of this “computer experiment” are shown in Fig. 8. The XDP of the 10 central nucleotide pairs in the dodecamer sequence moves rapidly toward more negative values, which is indicative of a transition toward A-like structures. Reversible transitions of the sugar pseudorotation phase angle from the C2'-endo toward the C3'-endo conformations is also observed, indicating that the selective hydration of the major groove does lead to an A-like sugar puckering. (Fig. S3 in Data S1).

We also observe that in this simulation of the CGCGA-ATTGC sequence, the sugars attached to C and G bases are more likely to be in the C3'-endo state than are those attached to the central A and T bases. This result suggests that the problem with converting B to A at reduced water activity, observed in MD simulations, may be a matter of sampling rather than, as previously suspected, a force field problem. The idea is that the B-to-A transition evidently proceeds if the solvent is structured properly in the major groove, but this is a highly improbable event on the nanosecond timescale, and thus the transition is not observed in the MD simulations. On the timescale of the
simulation, this is a quasiergodic problem. In time frames from microseconds to seconds, even improbable events at the nanosecond level can occur and nucleate the conformational change, analogous to what happens in the observation of particles undergoing Brownian motion. Although this is only a leading idea, and not unequivocally demonstrated by the analysis, our results do indicate that quasiergodic problems deserve serious consideration in assessing future MD studies of the B-to-A transition at reduced water activity.

The time actually required for the reorganization of the solvent remains an open question. The measurements of Jose and Porschke (29) suggest that the timescale of the A-to-B transition is 10 ms and that a significant transition barrier exists. The results presented here indicate that the barrier for the transition lies in the organization of the solvent. Mazur had concluded that occupation of ions in the major groove was required for the transition (11); however, the results of this study suggest that reorganization of water is equally important in this transition. The ultrafast dynamic measurements of Zewail and co-workers (85) demonstrate that solvent reorganization can occur on a timescale of 20 ps. However, these measurements were confined to a relatively localized area, and a complete reorganization of solvent in the major groove may require a longer timescale. Further studies are needed to appropriately address this issue.

**SUMMARY AND CONCLUSIONS**

A combination of methodologies has been utilized to examine the A-to-B transition. The comparison of UVRR results with information gleaned from MD simulations has proven illuminating with respect to the nature and mechanism of the DNA A-to-B transition. Of significance are the correlations between UVRR, MD, and crystallography results, which all point to a sequential mechanism for the A-to-B transition (Fig. 9) (8). In addition, UVRR and MD results also reveal that sugar conformation may not be the most reliable measure of A-to-B DNA conformation, as suggested in earlier NMR studies (6). Finally, these measurements highlight the significance of water and ion positions in facilitating the A-to-B transition.

**SUPPLEMENTARY MATERIAL**

To view all of the supplemental files associated with this article, visit www.biophysj.org.

This work was supported by National Science Foundation grant MCB 0316625 (I.M.) and an investigator grant from the Patrick and Catherine Weldon Donaghue Medical Research Foundation (I.M.). D.L.B. acknowledges support from National Institute of General Medical Sciences 7909, National Institutes of Health Molecular Biophysics training grant GM 008272 and the Howard Hughes Medical Institute supported K.M.K. and C.E.A., respectively.

**REFERENCES**


Biophysical Journal 95(1) 257–272

---

**FIGURE 9** Molecular dynamics snapshots of the A-to-B transition. As indicated, the sugar puckers and the helical parameter INC transition to B-form values within 500–750 ps. The basepair XDP converts from A to B values in the 750–1250 ps time range. The A-to-B interconversion is complete by 1500 ps of MD. NP and SP correspond to C3-endo and C2-endo ribosyl conformations, respectively.
Mechanism for DNA A-to-B Transition


